REGULARIZATION PARAMETER TRIMMING FOR ITERATIVE IMAGE RECONSTRUCTION

Haoyi Liang and Daniel S. Weller

University of Virginia, Department of ECE, Charlottesville, VA, 22904, USA

ABSTRACT

Conventional automatic parameter choosing involves testing many parameter values, increasing computing time for iterative image reconstructions. The proposed approach first measures the image quality after each iteration and then predicts the convergence trend corresponding to each value of the parameter. Values unlikely to achieve the best quality upon convergence are trimmed from successive iterations to save time. Experimental results show that our parameter trimming method could reduce the running time of total variation parameter selection solved by Split Bregman iteration by more than 50% when the numbers of iterations and parameter candidates are large.

***Index Terms*—**Parameter trimming, Iterative algorithm, Image reconstruction, Perceptual image quality assessment

1. INTRODUCTION

In many regularized image processing algorithms [1]-[9], parameter tuning is pivotal to attaining satisfactory image quality or speed of the algorithm. In this paper, we focus on image reconstruction because image reconstruction is an ill-posed problem and a regularized iterative algorithm is the common approach. Compared with denoising [3], [10], the regularization parameter, as an estimation of the sharpness in the original image, is even more important to image reconstruction because not only noise is added, but also a portion of the image information may be lost. Conventional approaches to parameter tuning based on data discrepancy, e.g., generalized cross validation (GCV) [11], L-curve [12], are widely used in regularization parameter choosing. But unlike perceptual image metrics, GCV and the L-curve method do not consistently yield realistic images. The common approach to automatic parameter selection samples the parameters over a possible range and runs the algorithm with each of these values to convergence before the parameter value yielding the best quality is chosen. This idea of incorporating quality-based parameter selection with an iterative algorithm has been mentioned before [4]-[7], [10]. However, in these works, time efficiency is not the key point, since either the algorithm converges quickly [7], [10] or the quality assessment itself is time-consuming [6]. For instance, the denoising parameter selection in [10] involves experiments with 30 parameter candidates and 20 iterations/candidate. In situations where algorithms converge slowly or the number of parameter candidates is large, assessing the image quality and identifying up the best result after all algorithm instances converge would be too time-consuming to be practical.

In this paper, we explore the convergence process of an iterative reconstruction algorithm with different parameters and trim the potential parameter candidates during the convergence process. Our work includes three key elements: image quality assessment, image reconstruction and parameter trimming. As for image quality assessment, numerous methods [10], [13]-[18] have been put forward. For image reconstruction, only no-reference methods [10], [16]-[18] are of interest. We finally choose Metric Q [10] as our no-reference image quality index because it is consistent with full-reference image quality index like structural similarity index measure (SSIM) [13] and does not require intensive computation. With the help of quality assessment after each iteration of the reconstruction process, simple and effective criteria are proposed to determine if a regularization parameter candidate can be trimmed before its convergence. By doing so, considerable time is saved finding the best regularization parameter.

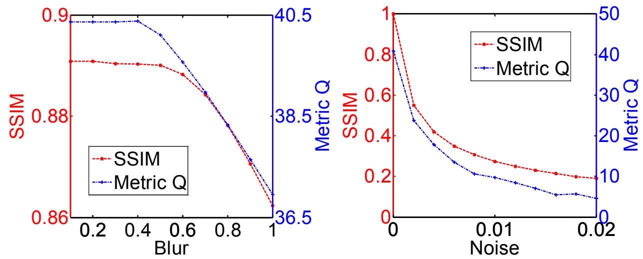
The rest of this paper is organized as follows. In Section 2, we briefly introduce the no-reference Metric Q and the image reconstruction algorithm. Then, we develop the model of parameter trimming during the reconstruction process. Next, Section 3 gives experiments to verify the effectiveness of the proposed method. On six total variation reconstruction examples using Split Bregman iteration, we present the time difference of parameter selection with and without parameter trimming. Finally, we conclude and discuss our contribution and further work in Section 4.

2. METHODOLOGY

We now illustrate the three main parts of our work: image quality assessment, iterative reconstruction algorithm and regularization parameter trimming criteria.

2.1. Image quality index

As mentioned in the last section, parameter trimming relies



(a) **(b)**

Figure 1. (a): SSIM and Metric Q value under blurring; (b): SSIM and Metric Q value under noise.

on the quality assessment of intermediate results after each iteration. In this paper, we adopt Metric Q [10] as the image quality index because it is a robust metric reflecting the degree of noise and blur in an image (as shown in Figure 1). Metric Q is also a practical measurement for its real-time speed. The general idea of Metric Q is splitting the whole image into small patches, and for each patch, a gradient-based quality index is calculated. For the whole image, the global quality is the average of the qualities of the patches which demonstrate meaningful content [10].

The House image [19], as shown in Figure 5(a), is used to show the validity of Metric Q by comparing it with SSIM [13] after the original image suffers different kinds of distortion. Figure 1(a) is the test result of blurring. The variance of the blurring kernel we used varies from 0.1 to 1. The Spearman correlation between SSIM and Metric Q is 0.9909. Figure 1(b) demonstrates that Metric performs well under Gaussian noise too. The variance of Gaussian noise ranges from 0 to 0.02 with mean 0, and the Spearman correlation between SSIM and Metric Q is 0.8902.

2.2. Total Variation reconstruction algorithm

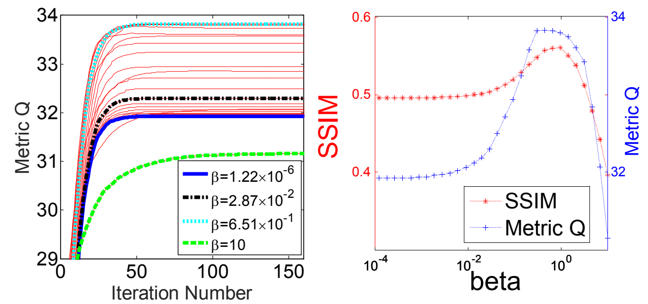
Total variation (TV) reconstruction [22] is aimed at minimizing the cost function (1)

$$E_\beta(x) = \beta\|Dx\|_1 + \frac{1}{2}\|R\mathcal{F}x - y\|_2^2 \qquad (1)$$

where x is the reconstructed image, y is the observed incomplete data set, R represents the subsampling matrix, $\mathcal{F}$ represents the Fourier transform matrix, D represents the finite differences matrix, and the TV regularizer $\|Dx\|_1$ combines directions isotropically. The regularization parameter β controls the sharpness of the reconstructed result. Large β will oversmooth the reconstructed image, while small β will leave residual noise. How to choose a proper β is crucial to the performance of TV reconstruction. Split Bregman iteration [20] is used to solve (1). By making replacements $d \leftarrow Dx$ and introducing the dual variable b, the split formulation of the problem becomes:

$$\min_{x,d} \beta\|d\|_1 + \frac{1}{2}\|R\mathcal{F}x - y\|_2^2 + \frac{\mu\beta}{2}\|d - Dx - b\|_2^2 \qquad (2)$$
$$s.t.\, d = Dx$$

The Split Bregman iteration solving (2) is:



(a) **(b)**

Figure 2. (a): convergence process; (b): result at 1000 iterations.

Initialize: $x^0 = 0, d^0 = b^0 = 0$
While stop criterion is not satisfied

$$x^{k+1} = \mathcal{F}^{-1}K^{-1}\mathcal{F}L$$
$$d^{k+1} = max(s^k - \frac{1}{\mu}, 0)\frac{Dx^k + b^k}{s^k}$$
$$b^{k+1} = b^k + (Dx^k - d^{k+1})$$

End

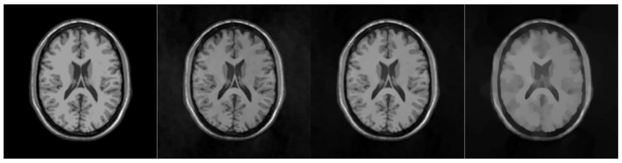
where we use the notation $K = (R^T R - \mu\beta\mathcal{F}D^T D\mathcal{F}^{-1})$, $L = (\mathcal{F}^T R^T y + \mu\beta D^T(d^k - b^k))$, $s^k = \sqrt{|Dx^k + b^k|^2}$ and μ is a Split Bregman penalty parameter, which is set to 0.01 in our implementation. It is worth to point out that since D is circulant, $\mathcal{F}D^T D\mathcal{F}^{-1}$ is diagonal.

2.3. Parameter trimming criteria

After specifying the image quality index and reconstruction algorithm, two problems on how to effectively incorporate the quality index into the iterative process arise: how to determine if the algorithm has converged and how to foretell which β is better than the others before convergence. An effective solution to both problems is crucial to accelerate the automatic parameter selection process.

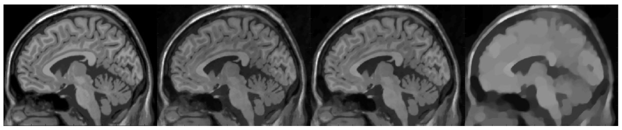
To illustrate the effects of parameter selection on image quality, Brain1 image [21] is reconstructed with 30 values of β in Figure 2, with selected images illustrated in Figure 3. These parameter values are uniformly sampled from 1.22×10^{-6} to 10 in log scale. The image quality index as a function of the number of iterations with different β is plotted in Figure 2(a), among which 4 lines which correspond to the smallest, the largest, the middle and the best of parameter value, are highlighted. Reconstructed image quality in both SSIM and Metric Q is plotted after 160 iterations for each value of β in Figure 2(b).

Inspired by the fact that iterative algorithms mainly introduce gradual global change rather than dramatic local changes, we determine the convergence point by the mean squared difference (MSD) without considering the high level features in an image, like structure.



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

Figure 3. (a): original Brain1 image [21]; (b): reconstructed result with $\boldsymbol{\beta} = 1.22 \times 10^{-6}$; (c): reconstruction with $\boldsymbol{\beta} = 0.446$; (d): reconstructed image with $\boldsymbol{\beta} = 10$.



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

Figure 4. (a): original Brain2 image [21]; (b): reconstructed result with $\boldsymbol{\beta} = 1.22 \times 10^{-6}$; (c): reconstruction with $\boldsymbol{\beta} = 3 \times 10^{-3}$; (d): reconstructed image with $\boldsymbol{\beta} = 10$.

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

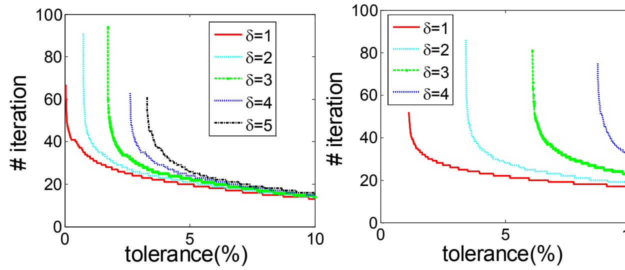
Figure 5. (a): original House [19]; (b): reconstructed result with $\boldsymbol{\beta} = 1.22 \times 10^{-6}$; (c): reconstruction with $\boldsymbol{\beta} = 1$; (d): reconstructed image with $\boldsymbol{\beta} = 10$.

Let $I(m)$ be the reconstructed image after the m^{th} iteration, $S_m = MSD(I(m), I(m-1))$, l_{cfm} is the length of the window to confirm the convergence and T_{cvg} is a user-defined threshold. Convergence of the reconstruction process is determined by the following criterion:

$$\begin{aligned} &if \quad max(S_m, S_{m-1}, \ldots, S_{m-l_{cfm}}) < T_{cvg} \\ &\qquad converge \\ &else \\ &\qquad continue\ iterating \end{aligned}$$

Depending on the accuracy requirement of different applications, the strictness of the convergence criterion can be adjusted by changing l_{cfm} and T_{cvg}. A larger l_{cfm} and a smaller T_{cvg} ensure tighter convergence.

We model the mathematics of predicting the converged quality based on two factors: its current quality index and the derivative of its quality curve during iteration at current position. After the i^{th} iteration, for the j^{th} regularization parameter (β) we do the quality assessment V_j^i of the corresponding reconstructed result. V_j represents the convergence process of the j^{th} beta and V^i represents the value of every β after the i^{th} iteration. A map P initialized with zero is maintained to make the trimming decision. $P(m,n) = w$ means in the last w iterations, the m^{th} beta is always better than the n^{th} beta. The predicted quality index of the m^{th} beta is $pre(m) = cur(m) + dev(m) * l_{pre}$. Here, $cur(m)$ is the current quality index of m^{th} beta, $dev(m)$ is the numerical derivative of V_m at i^{th} iteration and l_{pre} is the predicting length in iterations. T_{trim} is introduced to make the trimming method robust to the nonmonotonic part of the reconstruction process. The following criterion is used for parameter trimming:



(a) (b)

Figure 6. (a): relation between convergence and iteration on Brain1 Image; (b): relation between convergence and iteration on House Image.

$$\begin{aligned} &if \left(1 - abs(dev(n))\right) \cdot pre(m) > pre(n) \\ &\qquad P(m,n) = P(m,n) + 1 \\ &else \\ &\qquad P(m,n) = 0 \\ &if\ P(m,n) = T_{trim},\ then^{th}\ beta\ is\ terminated \end{aligned}$$

Since a candidate value will not come back once trimmed and part of the convergence process may not be monotonic, we adopt two criteria to prevent mistrimming. First, parameters that show significant quality change will not be trimmed out by combining $cur(m)$ and $dev(m)$ with l_{pre}, and adding the $\left(1 - abs(dev(n))\right)$ term. Second, T_{trim} ensures that one parameter candidate is trimmed only if it is exceeded by another candidate in the last consecutive T_{trim} iterations. It is the shown in the experiment part that even under the conservative trimming criteria, the proposed approach effectively speeds up the parameter selection process.

3. EXPERIMENTS

We first portray some reconstruction results and then present a real situation in which our parameter trimming algorithm demonstrates obvious advantage over the parameter selection algorithm without it.

In Figures 3-5, we show the reconstructions of two brain images and the House image. Only 50% Fourier transform data are used to reconstruct the image and in order to be more realistic, Fourier transform data are distorted by Gaussian noise. The SNR is kept at 40 dB in all reconstruction experiments.

In order to study the relation between convergence and the number of iterations, we develop a function

$$G_\delta(t) = \min(i_1, i_2, i_3, \ldots, i_m)$$

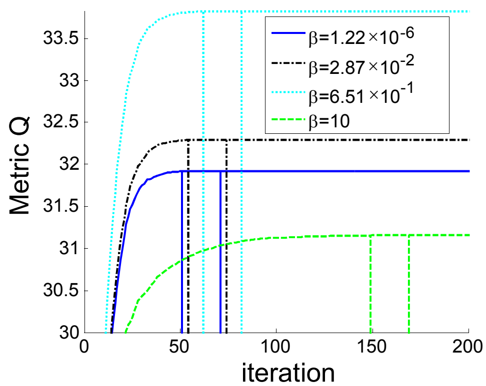


Figure 7. Convergence in Metric Q

in which $i_j, j = 1,2,3, \dots, m$, satisfies

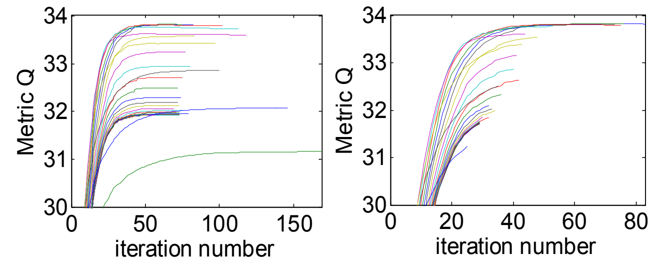
$$\min\left(V_{\beta^*-\delta}^{i_j}, \dots, V_{\beta^*+\delta}^{i_j}\right) > V_{\beta^*}^{N} * (1-t)$$

β^* is best parameter candidate (in order number here), N is the iteration number from where β^* convergences, and t is the relative tolerance of deviation from $V_{\beta^*}^{N}$.

We illustrate the function G with δ from 1 to 5 of Brain1 and House image in Figure 6. Point (t, s) on the line corresponding to δ means after s iterations, all the image quality index in the set of $\{\beta^* - \delta, \dots, \beta^* + \delta\}$ enter our accepted quality range determined by t. Larger δ means we have to make sure more β have entered the accepted quality range. The starting point of each line indicates after how many iterations the quality stop increasing and the best quality this set of β could achieve. In the two examples, we can see that all sets of β stop increasing after about 100 iterations.

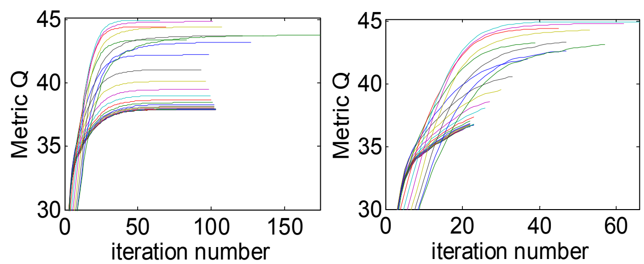
The previous experiments illustrate convergence analysis on all the data we have after 500 iterations. However, we don't have such information in real situations. So how to determine the convergence point and foretell the trend of each β based on existing data is of significance. Figure 7 demonstrates how we determine convergence for different β of the Brain1 image reconstruction with our MSD-based convergence criteria. In Figure 7, the second vertical line indicates the convergence point for each iteration line. The iterations between the first and the second lines are the process of confirming the convergence. In our experiment, we used $T_{cvg} = 10^{-4}$ and $l_{cfm} = 20$ to determine convergence. Please notice that though we do the convergence judgment using MSD values, we illustrate the convergence point in Metric Q values for we use Metric Q to trim parameter candidates.

Figures 8 and 9 illustrate the difference between selecting β with and without parameter trimming. We set $l_{pre} = 5$ and $T_{trim} = 5$ for our trimming predictor here. In Figures 8 (a) and 9 (a), each β continues iterating until convergence criteria are satisfied; in Figures 8(b) and 9(b), β values with poor performance are terminated before convergence. It is



(a) **(b)**

Figure 8. (a): Brain1 image reconstruction process without parameter trimming; (b) Brain1 image reconstruction process with parameter trimming



(a) **(b)**

Figure 9. (a): House image reconstruction process without parameter trimming; (b) House image reconstruction process with parameter trimming

clear that considerably many iterations are saved with parameter trimming.

Table 1 quantitatively compares these two versions of parameter selection on more images (Peppers, Boat and Cameraman are standard images which can be found from [19]). The # Total iteration1 and # Total iteration2 values are the summations of iterations before the best β is selected, without and with parameter trimming respectively. The Best $\beta_1$and best β_2 values are the best β selected without and with parameter trimming, in order number (30 β values are uniformly sampled from 1.22×10^{-6} to 10 in log scale).

Table 1 Effectiveness of parameter trimming

Image	# Total itera-tion1	# Total itera-tion2	**Saved compu-tation**	Best β_1	Best β_2
Brain1	2568	1184	**53.89%**	22th	22th
Brain2	2399	1166	**51.40%**	21th	21th
Pepper	2612	821	**68.57%**	25th	25th
Boat	2946	722	**75.49%**	26th	26th
Cameraman	3837	1517	**60.46%**	28th	28th
House	3073	949	**69.12%**	25th	25th

Clearly, a lot of computation is saved with parameter trimming. In each of our experiments, more than 50% computation is saved and the selected β are the same.

Through these experiments our proposed parameter trimming method meets the two aims: (1) automatically selecting the best parameters among a set of candidates; (2) accelerating the process of parameter selection by foretelling the trend of each parameter.

4. CONCLUSION

We examine the idea of adopting parameter trimming during the convergence process of an iterative reconstruction algorithm and verify this idea by demonstrating significant reduction in total number of iterations while still choosing the best regularization parameter in multiple images. The potential improvement by adopting the idea of parameter trimming is considerable, especially when the number of iterations is large or the accuracy requirement is high. We expect that the idea of parameter trimming would result in noticeably improved computational efficiency in related image processing problems like deblurring or denoising as well.

A single parameter algorithm is used to evaluate parameter trimming in this paper. However, in many situations where multiple parameters are used, more comprehensive parameter trimming methods should be studied to more accurately predict the convergence trend.

5. REFERENCES

[1] S. Ravishankar and Y. Bresler, “MR image reconstruction from highly undersampled k-space data by dictionary learning,” IEEE Trans. Med. Imaging, vol. 30, no. 5, pp. 1028–1041, May 2011.

[2] V. M. Patel, R. Maleh, A. C. Gilbert, and R. Chellappa, “Gradient-based image recovery methods from incomplete Fourier measurements,” IEEE Trans. Image Process, vol. 21, no. 1, pp. 94–105, Jan. 2012.

[3] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Phys. D, vol. 60, pp. 259–268, 1992.

[4] B. Peng and O. Veksler, “Parameter Selection for Graph Cut Based Image Segmentation,” Procedings Br. Mach. Vis. Conf, pp. 16.1–16.10, 2008.

[5] Y. Yitzhaky and E. Peli, “A Method for Objective Edge Detection Evaluation and Detector Parameter Selection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no.8, pp. 1027-1033, Aug. 2003.

[6] M. S. C. Almeida and M. a T. Figueiredo, “Parameter estimation for blind and non-blind deblurring using residual whiteness measures,” IEEE Trans. Image Process., vol. 22, no. 7, pp. 2751–2763, Jul. 2013.

[7] J. L. Herraiz, M. Ieee, S. Gabarda, G. Cristobal, and S. Member, “Automatic Parameter Selection in PET Image Reconstruction Based on No-reference Image Quality Assessment,” IEEE Nuclear Science Symposiwn and Medical Imaging Conference Record, pp. 3371–3374, 2012.

[8] K. Miller, “Least-squares methods for ill-posed problems with a prescribed bound,” SIAM J. Math.Anal, vol. 1, no. 1, pp. 52–70, Feb. 1970.

[9] G. Demoment, “Image reconstruction and restoration: overview of common estimation structures and problems,” IEEE Trans. Acoust. Sp. Sig. Proc., vol. 37, no. 12, pp. 2024–2036, Dec. 1989.

[10] X. Zhu, S. Member, and P. Milanfar, “Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content,” IEEE Trans. Image Process, vol. 19, no. 12, pp. 3116–3132, Dec. 2010.

[11] P. C. Hansen and D. P. O’Leary, “The use of the L-curve in the regularization of discrete ill-posed problems,” SIAM J. Sci. Comput, vol.14, no. 6, pp. 1487–1503, 1993.

[12] G. H. Golub, M. Heath, and G. Wahba, “Generalized cross-validation as a method for choosing a good ridge parameter,” Technometrics, vol.21, no. 2, pp. 215–223, May 1979.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, “Image Quality Assessment : From Error Visibility to Structural Similarity,” IEEE Trans. Image Process, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[14] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” IEEE Trans. Image Process., vol. 15, no. 2, pp. 430–444, Feb. 2006.

[15] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, “A statistical evaluation of recent full reference image quality assessment algorithms,” IEEE Trans. Image Process, vol. 15, no. 11, pp. 3441–3452, Nov. 2006.

[16] A. K. Moorthy and A. C. Bovik, “Blind image quality assessment: from natural scene statistics to perceptual quality,” IEEE Trans. Image Process, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.

[17] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” IEEE Trans. Image Process, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[18] S. Gabarda and G. Cristóbal, “Blind image quality assessment through anisotropy,” J. Opt. Soc. Am. A, vol. 24, no. 12, pp. B42-B51, Dec. 2007.

[19] http://sipi.usc.edu/database/database.php?volume=misc

[20] T. Goldstein and S. Osher, “The Split Bregman Method for L1-Regularized Problems,” SIAM J. Imaging Sci., vol. 2, no. 2, pp. 323–343, Jan. 2009.

[21] “MRI simulation-based evaluation of image-processing and classification methods” IEEE Trans. Medical Imaging, vol. 18, no. 11, pp. 1085-1097, Nov. 1999.

[22] L. Rudin, and S. Osher. "Total variation based image restoration with free local constraints." in Proc. Int. Conf. Image Process., vol. 1, pp. 31-35, Nov. 1994