

Research Article

An Objective Evaluation Method for Image Restoration

A. Lakshmi and Subrata Rakshit

Center for Artificial Intelligence and Robotics, India

Correspondence should be addressed to A. Lakshmi, lachuv@yahoo.co.in

Received 25 September 2009; Revised 21 January 2010; Accepted 31 January 2010

Academic Editor: Tamal Bose

Copyright © 2010 A. Lakshmi and S. Rakshit. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate image restoration is of paramount importance for low-level vision, computer vision, and various other fields. Numerous complex restoration algorithms have been stated in the literature. Performance of these restoration algorithms differs with the nature of the image and distortion. These algorithms are evaluated either qualitatively or quantitatively by comparing the restored image with the original image. The practical drawback of this quantitative comparison is the requirement for the original image. In this paper, we propose a measure, with theoretical grounding, to objectively evaluate the restoration algorithms with no knowledge about the original image. This measure analyses the deblurring as well as the denoising nature of the restoration method. The main utility of this measure will be in designing automatic restoration systems. Effectiveness of this measure is substantiated with experimental results.

1. Introduction

Images become blurred when there is a smearing of the point spread function caused by certain imaging conditions such as camera motion, object motion, atmospheric turbulence, and so forth. In addition, sensor noise leads to additional degradation of the image. Image restoration [1, 2] is concerned with estimation of the undistorted image from the given blurred and noisy image. It has utmost importance in medical imaging, astronomy, outdoor surveillance photography, and it is a frequent preprocessing step for high-level vision. Image restoration for linear spatially invariant degradation is generally modelled as

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \quad (1)$$

where, \mathbf{g} , \mathbf{f} , and \mathbf{n} are lexicographically ordered vectors of blurred, original, and noisy images, respectively. \mathbf{H} is the convolution matrix of the linear, spatially invariant kernel \mathbf{h} .

Numerous image restoration algorithms [3–8] have emerged utilizing a wide variety of regularizations, derived based on some a priori knowledge about the original image. Image restoration is a complex problem without an exact solution. The performance of a restoration algorithm varies for different sets of images with dynamically changing environment and values chosen for its own parameters [9]. These

algorithms have a regularization parameter that controls the trade off between data fidelity and noise reducing capability of the algorithm. Maximizing the value of the regularization parameter distorts the signal, whereas minimizing its value results in singularity problem followed by noise magnification [1]. Thus, ranking of the algorithms is of paramount importance. Visual evaluation of the restoration algorithm is subjective (depends on the evaluator) and it puts extra load on users. An ability to numerically evaluate multiple image restoration algorithms for the given blurred image is necessary to make image restoration adaptive.

The improvement in the quality of the estimated image $\{\hat{\mathbf{f}}\}$ over the blurred image is measured in terms of improvement in signal-to-noise ratio (ISNR) as [10]

$$\text{ISNR} = \frac{\|\mathbf{g} - \mathbf{f}\|^2}{\|\mathbf{f} - \hat{\mathbf{f}}\|^2}. \quad (2)$$

The ISNR can be computed only if there is access to the original image, thus making this objective evaluation impossible in real-time scenario. Despite significant advances in the field of image restoration, evaluation of these techniques, by and large, has been subjective in the real field conditions.

In this paper, we introduce a new information theoretic approach (based on entropy) to evaluate restoration. Our new method provides a score to evaluate the parameterizations of a single restoration algorithm as well as fundamentally different restoration algorithms. This measure does not demand any knowledge about the original image thus making it feasible to be used to automate the restoration process in the real field. The proposed measure quantifies the noise reducing as well as the deblurring capability of restoration algorithms.

The advantages of the proposed objective evaluation are the following. (i) It automates the restoration process in real time as it does not demand any knowledge of the original image, (ii) it ranks the algorithms based on the data fidelity as well the denoising capability, and (iii) The proposed measure does not assume the statistics of the noise or the image.

Section 2 introduces the proposed empirical measure and justifies the use of entropy as an evaluation function. Sections 3 and 4 furnish the theoretical and empirical results, respectively, to substantiate the use of the given measure in judging restoration algorithms.

2. New Evaluation Method

This section brings out our information theoretic approach that could be used for restoration evaluation and justifies the use of each and every term of the proposed measure in the evaluation process.

A good restoration method should concentrate on estimating images as close as possible to the original image. The deviation of the estimate from the original image is a quantitative measure for the data fidelity of the restoration algorithm. Thus, the data fidelity of restoration algorithms can be quantified with a function ϕ of the residual image (\mathbf{r}) as given in (3), where $\hat{\mathbf{f}}$ and $\hat{\mathbf{h}}$ are the estimated image and blur, respectively.

$$\mathbf{r} = \mathbf{g} - \hat{\mathbf{H}}\hat{\mathbf{f}}. \quad (3)$$

Maximum likelihood (ML) estimation or least square (LS) estimation could be used as ϕ of the residue. There are disadvantages associated with the use of ML or LS. ML could be applied only when the exact distribution of the residue is known a priori, which is not possible in real-time scenario. LS works best for data that does not contain a large number of random residual errors with extreme values. The main disadvantage of LS is its sensitivity to outliers. Outliers have large influence on the results as squaring the residue magnifies the effect of these extreme data points.

Due to the inherent disadvantages of ML and LS, we use entropy (H) as ϕ of the residue. Entropy is robust to the outliers and it is a nonparametric approach. This is of so much interest in scenarios where the statistics of the underlying residual error is not available a priori. In addition to the entropy of the residue, we consider the entropy of the estimated image and the blur to quantify the

denoising capability of restoration algorithms (explanation in Section 3.2). Thus, the proposed measure turns out to be

$$H(\mathbf{r}) + H(\hat{\mathbf{f}}) + H(\hat{\mathbf{h}}). \quad (4)$$

The restoration algorithm which gives minimum value for the above equation can be concluded as the best restoration algorithm. The role played by each entropy term is given in detail in the forthcoming sections.

3. Justification of the Derived Measure

Equation (4) is comprised of three terms. The first term is a measure of the data faithfulness; the second and third terms are measures of the noise content of the restored image and blur. So, the first term is called Data Fidelity term (DF); the second and third terms are called Noise Assessing terms (NA_1 and NA_2), respectively.

3.1. DF Term. This section proves the significance of the DF term in quantifying the deblurring effect of the given algorithm. The residue \mathbf{r} can be written as

$$\mathbf{r} = \mathbf{x} + \mathbf{n}, \quad (5)$$

where

$$\mathbf{x} = \mathbf{H}\mathbf{f} - \hat{\mathbf{H}}\hat{\mathbf{f}}. \quad (6)$$

If the restoration algorithm completely removes the noise \mathbf{n} from the estimates $\hat{\mathbf{h}}$ and $\hat{\mathbf{f}}$, then \mathbf{x} and \mathbf{n} are independent of each other. In such scenarios, the minimum value of $H(\mathbf{r})$ is $H(\mathbf{n})$, which is proved in the following paragraph.

Let us consider the mutual information between \mathbf{x} and \mathbf{r}

$$\begin{aligned} MI(\mathbf{r}, \mathbf{x}) &= H(\mathbf{r}) - H\left(\frac{\mathbf{r}}{\mathbf{x}}\right) \\ &= H(\mathbf{r}) - H\left(\frac{\mathbf{n}}{\mathbf{x}}\right). \end{aligned} \quad (7)$$

As \mathbf{n} and \mathbf{x} are assumed to be independent of each other,

$$H\left(\frac{\mathbf{n}}{\mathbf{x}}\right) = H(\mathbf{n}). \quad (8)$$

Substituting (8) in (7),

$$MI(\mathbf{r}, \mathbf{x}) = H(\mathbf{r}) - H(\mathbf{n}). \quad (9)$$

The positive property of MI implies

$$H(\mathbf{r}) \geq H(\mathbf{n}). \quad (10)$$

Thus, $H(\mathbf{n})$ is the minimum value that $H(\mathbf{r})$ can attain when the estimates are independent of the input noise \mathbf{n} .

When the image and the blur are estimated perfectly, $H(\mathbf{r})$ becomes $H(\mathbf{n})$ (minimum value of $H(\mathbf{n})$). This proves the legitimacy of $H(\mathbf{r})$ in estimating the data faithfulness. $H(\mathbf{r})$ alone will suffice for assessing the restoration algorithms that remove the input noise completely from the estimated image and the blur. If a restoration algorithm fails to remove the noise totally from the estimates, the assumption of \mathbf{x} and \mathbf{n} being independent will be violated and (10) may not be true and there may be situations (as given in Section 3.2) where a noisy estimate may yield a lower $H(\mathbf{r})$ than that of the perfect estimates.

TABLE 1: Details of the experiments conducted.

Experiment	Images	Blur	Noise	De-convolution methods
Exp 1a	single chip image	Box blur (7×7 pixels)	No noise	Wiener [16, 17], CLS, Inverse
Exp 1b	single Lena image	Box blur (7×7 pixels)	Poisson noise	Wiener, CLS, Inverse
Exp 2a	Two Rice images	Gaussian blur	Gaussian of variance (in dB) $n_1 = 10.7$ (image 1), $n_2 = 8.83$ (image 2)	Wiener, CLS, inverse filter, Total variation [18], Lucy [19]
Exp 2b	20 Mountain images	Each blurred by Motion blurs of different trajectory.	Gaussian noise	Wiener
Exp 2c	Single Cameraman image	Box blur (7×7 pixels)	Gaussian noise	CLS [9] with lambda 0.1 to 80 in increments of 10

3.2. *NA Terms.* There are restoration algorithms, which are good at deblurring but not at denoising (such as inverse filter). Consider two different algorithms (algo 1 and algo 2) producing estimates of the image and blur (11) and thus leading to $H(\mathbf{r})$ as given in (12). Algo 1 gives perfect estimate of the image and blur, whereas algo 2 has amplified the noise in the process of deblurring

$$\text{Algo1} - \hat{\mathbf{f}}_1 = \mathbf{f}, \quad \hat{\mathbf{H}} = \mathbf{H}, \quad (11)$$

$$\text{Algo2} - \hat{\mathbf{f}}_2 = \mathbf{f} + \mathbf{H}^{-1}\mathbf{n}, \quad \hat{\mathbf{H}} = \mathbf{H},$$

$$H(\mathbf{r}_2) = H(0) = 0, \quad H(\mathbf{r}_1) = H(\mathbf{n}), \quad (12)$$

$$H(\mathbf{r}_2) < H(\mathbf{r}_1). \quad (13)$$

$H(\mathbf{r})$ (13) of algo 2 will be lesser than that of algo 1, indicating algo 2 to be the best algorithm, which is not true. This shows the ineffectiveness of DF term in quantifying the restoration algorithm in the case of noisy estimates. In order to prevent noisy estimates to be judged as good estimates, we need additional NA terms. The NA terms quantify the noise reducing capability of the given algorithm.

The entropy is a measure of disorder in the image. The randomness of noise leads to higher entropy, whereas the natural images which are not as random as that of noise, possess smaller entropy [11, 12], that is, the entropy of image corrupted with independent noise \mathbf{n} will be greater than that of the original image (14).

$$H(\mathbf{f} + \mathbf{n}) > H(\mathbf{f}). \quad (14)$$

To prove (14), let us consider the *MI* between $\mathbf{y} = \mathbf{f} + \mathbf{n}$ and \mathbf{n}

$$\begin{aligned} MI(\mathbf{y}, \mathbf{n}) &= H(\mathbf{y}) - H\left(\frac{\mathbf{y}}{\mathbf{n}}\right) \\ &= H(\mathbf{y}) - H\left(\frac{\mathbf{f}}{\mathbf{n}}\right). \end{aligned} \quad (15)$$

As \mathbf{f} and \mathbf{n} are independent of each other, the above equation becomes

$$MI(\mathbf{y}, \mathbf{n}) = H(\mathbf{y}) - H(\mathbf{f}). \quad (16)$$

The positive nature of *MI* proves that $H(\mathbf{f} + \mathbf{n}) > H(\mathbf{f})$, that is, the entropy of the image with additive noise is greater than

that of the original image. This proves the genuineness of NA terms in the proposed measure to assess the denoising capability of restoration algorithms.

3.3. *Minimum Description Length (MDL) Principle.* An alternate insight of our evaluation method can be obtained by applying MDL [13] principle to balance the trade off between data fidelity and noise removal. Our evaluation formula typically has two description lengths (i) $H(\mathbf{r})$ and (ii) $H(\hat{\mathbf{f}}) + H(\hat{\mathbf{h}})$. Restoration algorithms could be broadly classified as follows. (i) Algorithms that find a very good data fit {smaller $H(\mathbf{r})$ }, but noisy estimates {larger $H(\hat{\mathbf{f}}) + H(\hat{\mathbf{h}})$ }; (ii) algorithms that find rather bad data fit {larger $H(\mathbf{r})$ }, but denoised estimates {smaller $H(\hat{\mathbf{f}}) + H(\hat{\mathbf{h}})$ }. The sum of the two description lengths will be minimized for the restoration algorithm which results in quite (but not too) denoised, with a good (not too perfect) data fit. This has been validated with empirical studies in the forthcoming section.

4. Experimental Results

This section brings out the results of the experiments conducted on the restoration results, to test and validate the efficacy of the proposed measure. We have studied the effectiveness of the proposed measure by verifying its consistency with the conventional Distortion Measure (DM - (17))

$$\text{Distortion} = \|\mathbf{f} - \hat{\mathbf{f}}\|^2. \quad (17)$$

Table 1 provides the details of the input images (size 256×256 pixels), blur, type of additive noise, and the iterative deconvolution methods [14, 15] used in different experiments.

The entropies of the residue and blur kernel are estimated as given in (18).

$$H(r) = - \sum_{i=0}^{255} p_r(i) \log\{p_r(i)\}, \quad (18)$$

$$H(\hat{h}) = - \sum_{i=0}^{255} p_h(i) \log\{p_h(i)\},$$

where $p_r(i)$ and $p_h(i)$ are the probabilities that a pixel in the residue and estimated blur respectively has a luminance

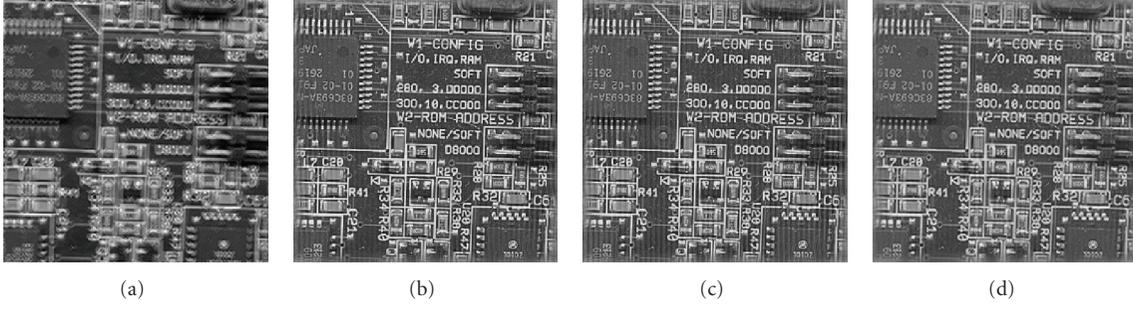


FIGURE 1: Restoration results in no noise scenario (a) Blurred Image, (b) Wiener, (c) Inverse, (d) CLS restored images.

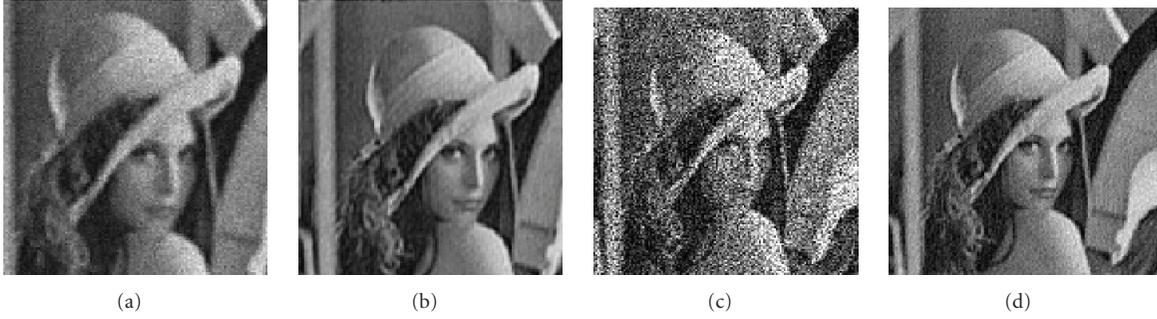


FIGURE 2: Restoration results in Poisson noise scenario (a) Blurred Image, (b) Wiener, (c) Inverse, (d) CLS restored images.

value i . To estimate the NA_1 term, \hat{f} is divided into j (1 to R) subimages of size 16×16 pixels. NA_1 (19) is calculated as the average of the entropy of the subimages (20).

$$H(\hat{f}) = \frac{1}{R} \sum_{j=1}^R H_j(\hat{f}), \quad (19)$$

$$H_j(\hat{f}) = - \sum_{i=0}^{255} p_{f_j}(i) \log\{p_{f_j}(i)\}, \quad (20)$$

where $H_j(\hat{f})$ is the entropy of the sub-image j and $p_{f_j}(i)$ is the probability that a pixel of the sub-image j has luminance value i . Sub-image of any real image will be smoother except for the presence of noise. Thus the entropy of the sub-image gives more appropriate measure of the noise content in the image than the entropy of the image as a whole.

4.1. Results of Experiment 1. This section gives the results of Exp 1a and 1b (Table 1). Figures 1(a), 1(b), 1(c), and 1(d) show the blurred image, Wiener, Inverse, and CLS restored images of Exp 1a, respectively. Figures 2(a), 2(b), 2(c), and 2(d) show the Blurred image, Wiener, Inverse, and CLS restored images of Exp 1b, respectively.

Table 2 is the display of the proposed measure and the distortion measure calculated on the restored images of Exp 1a and 1b. The proposed and DM conclude Wiener to be the best restoration method in Exp 1a and 1b, thus proving the credibility of the proposed measure.

4.2. Results of Experiment 2. This section furnishes the results of the three experiments Exp 2a, 2b, and 2c (refer Table 1).

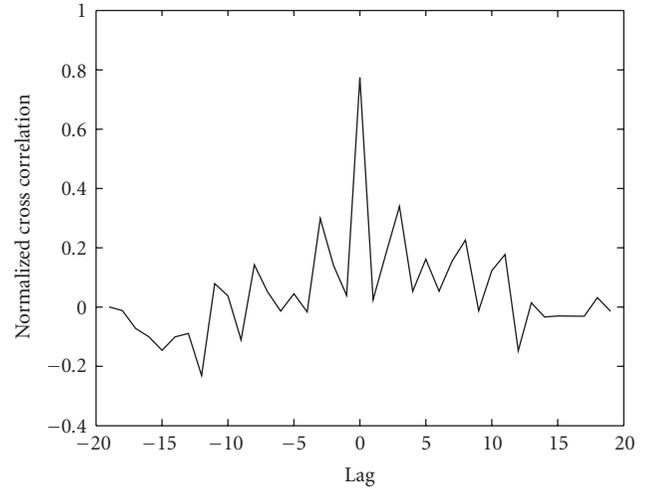


FIGURE 3: Cross correlation between Proposed and DM.

Table 3 gives the values of the proposed and distortion measures for Exp 2a. The proposed measure coincides with the distortion measure, thus proving the efficacy of the proposed measure.

Figure 3 is the plot of normalized cross correlation (21) between the proposed and DM of Exp. 2b.

$$\rho_{pd}(l) = \frac{r_{pd}(l)}{\sqrt{r_{pp}(0)r_{dd}(0)}}, \quad (21)$$

where $r_{pd}(l)$ is the cross correlation between the Proposed measure and distortion measure at lag l . Lag refers to the lag

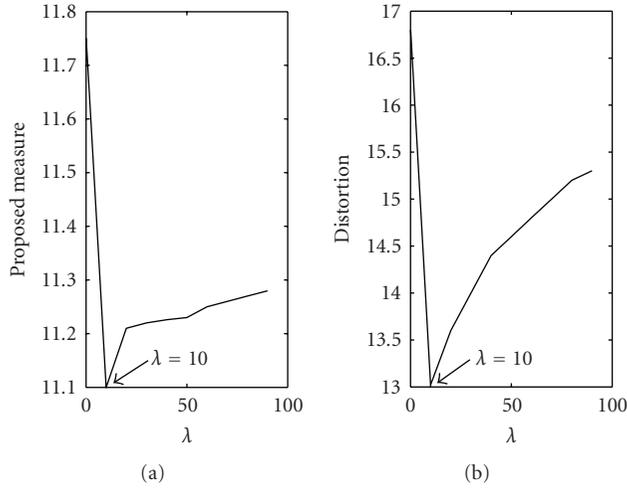


FIGURE 4: Results of Exp 2c. (a) Proposed versus lambda, (b) DM versus lambda.

TABLE 2: Comparison of Proposed and Distortion measure of Experiments 1a and 1b.

Algo	No noise		With noise	
	Proposed	DM	Proposed	DM
Wiener	12.71	27.7	10.5	12.7
Inverse	12.96	30.9	15.4	18.8
CLS	12.85	28.3	11.1	13.1

TABLE 3: Comparison of Proposed and Distortion measure of Experiment 2a.

Algo	Noise Variance = n_1		Noise Variance = n_2	
	Proposed	DM	Proposed	DM
Wiener	11.8	10.5	10.8	9.7
CLS	12.1	11.3	11.6	10.8
Inverse	15.5	14.2	14.3	13.3
Lucy	13.2	12.1	12.5	11.5
TV	11.5	10.1	11.1	10

in the ordering of the DM values in the experimental table. $r_{pp}(0)$ and $r_{dd}(0)$ are auto correlations of the proposed and distortion measure, respectively at lag 0. Figure 3 reaches its maximum value for lag = 0, thus proving that the proposed and distortion measures follow the same trend in their value fluctuation.

Figure 4 is the result of Exp 2c. As it can be seen, minimum values of the proposed measure and distortion are obtained at the same lambda (=10), certifying the parameter lambda = 10 to be the optimum parameter for the given blurred image.

5. Conclusion

We have proposed a measure to analyse the performance of restoration algorithms for the given blurred image, which (i) does not demand any knowledge about the original image

and (ii) derived without any assumption of the noise or image statistics. This helps in automating the restoration process without the intervention of the user. The given measure has a data fidelity term and noise assessing terms, thus analysing the deblurring as well as the denoising nature of the restoration method. The authenticity of the proposed measure was substantiated with different sets of empirical results.

References

- [1] H. Andrews and B. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood cliffs, NJ, USA, 1977.
- [2] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Magazine*, vol. 14, no. 2, pp. 24–41, 1997.
- [3] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
- [4] A. Tikhonov and V. Arsenin, *Solutions of Ill Posed Problems*, Winston and Sons, New York, NY, USA, 1977.
- [5] A. Tikhonov, "Regularization of incorrectly posed problems," *Soviet Mathematics. Doklady*, vol. 4, pp. 1624–1627, 1963.
- [6] A. K. Katsaggelos, "Iterative image restoration algorithms," *Optical Engineering*, vol. 28, no. 7, pp. 735–748, 1989.
- [7] K. Miller, "Least squares method for ill posed problems with a prescribed bound," *SIAM Journal on Mathematical Analysis*, vol. 1, pp. 52–74, 1970.
- [8] A. K. Katsaggelos, "Iterative image restoration algorithms," *Optical Engineering*, vol. 28, no. 7, pp. 735–748, 1989.
- [9] 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 Transactions on Image Processing*, vol. 1, no. 3, pp. 322–336, 1992.
- [10] A. C. Bovik, *Handbook of Image and Video Processing*, Elsevier Academic Press, San Diego, Calif, USA, 2nd edition, 2005.
- [11] M. Mitzenmacher and E. Upfal, *Probability and Computing, Randomized Algorithms and Probabilistic Analysis*, Press Syndicate of University of Cambridge, Cambridge, Mass, USA, 2005.
- [12] S. Rakshit and A. Mishra, "Estimation of structural information contents in images," in *Proceedings of the 7th Asian Conference on Computer Vision*, Hyderabad, India, January 2006.
- [13] P. D. Grunwald, J. Myung, and M. A. Pitt, *Advances in Minimum Description Length: Theory and Applications*, MIT Press, Cambridge, Mass, USA, 2005.
- [14] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 43–64, 1996.
- [15] D. Kundur and D. Hatzinakos, "Blind image deconvolution revisited," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 61–63, 1996.
- [16] A. D. Hillery and R. T. Chin, "Iterative Wiener filters for image restoration," *IEEE Transactions on Signal Processing*, vol. 39, no. 8, pp. 1892–1899, 1991.
- [17] A. Hillery and R. T. Chin, "Iterative Wiener filters for image restoration," Tech. Rep. TR-89-23, Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1989.

- [18] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, "Total variation-based image deconvolution: a majorization-minimization approach" in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '06)*, vol. 2, pp. 861–864, Toulouse, France, May 2006.
- [19] W. H. Richardson, "Bayesian based iterative method of image restoration," *Journal of Optical society of America*, vol. 62, no. 1, pp. 55–59, 1972.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

