OPTIMAL CHOICE OF LOCAL REGULARIZATION WEIGHTS IN ITERATIVE IMAGE RESTORATION

Steven S. O. Choy, Yuk-Hee Chan and Wan-Chi Siu

Department of Electronic Engineering
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

ABSTRACT

In the study of space-variant regularization for image restoration, little effort has been devoted to the search of optimal local regularization weights. In this paper, we address how to derive the optimal local regularization weights in the context of iterative image restoration. The optimal relationship between the two weight matrices for local regularization is derived, and, based on that relationship, a proper choice of the weight matrices is then presented. The results we derived provide a mathematical backup of the viability of some heuristic solutions suggested in the literature.

1. INTRODUCTION

Image restoration refers to the problem of estimating the original image from its degraded observation. In many practical situations, image degradation process can be generally formulated by the following matrix-vector equation

$$\tilde{y} = y + n = Dx + n \qquad (1)$$

where x and $\tilde{y}$ are the lexicographically ordered original and degraded images, n is white Gaussian noise uncorrelated with x, and matrix D represents the degradation operator. It is well-known that image restoration is an ill-posed inverse problem owing to the presence of noise in the observed image. Hence, the original image cannot be restored satisfactorily by applying direct inverse operator to the observed image. Regularization is a well-established approach to compensate the ill-posedness of the restoration problem [1]. Regularized image restoration is usually formulated as finding an solution $\hat{x}$ which minimizes the functional [2]

$$\|\tilde{y} - D\hat{x}\|_R^2 + \alpha\|C\hat{x}\|_S^2, \qquad (2)$$

where the norms are defined as $\|C\hat{x}\|_S^2 = (C\hat{x})^T S(C\hat{x})$ and $\|\tilde{y} - D\hat{x}\|_R^2 = (\tilde{y}-D\hat{x})^T R(\tilde{y}-D\hat{x})$. In above expressions, matrix C is the regularization operator and α, R and S are used to control over the amount of regularization. The scalar α is applied to regulate the solution globally and is usually termed as regularization parameter. On the other hand, diagonal matrices R and S are used to weight the relative amount of regularization at each point of the solution. In such a case, regularization is space-variant and adaptivity can be introduced by adapting regularization to the local image properties [3]. It has been shown that space-variant (adaptive) regularization can achieve a better performance in restoration than space-invariant regularization does since it can reduce the restoration artifacts more effectively [3-4].

The regularized solution that minimizes (2) with respect to $\hat{x}$ can be written explicitly as

$$\hat{x} = (D^TRD + \alpha C^TSC)^{-1} D^TR\tilde{y} \qquad (3)$$

However, direct computation of the above solution is impractical as it involves an inversion of a huge matrix. Therefore, the solution is usually obtained by successive approximation using iteration [2]

$$\hat{x}_{k+1} = \hat{x}_k + \beta[D^TR(\tilde{y}-D\hat{x}_k) - \alpha C^TSC\hat{x}_k] \qquad (4)$$

where β is the relaxation parameter of the iteration. The above equation is the general formulation of iterative adaptive regularized image restoration algorithm. The realization of this iterative algorithm necessitates the choices of the regularization operator C, the regularization parameter α, and the weight matrices R and S. The determination of the global regularization weights, C and α, have been studied by a number of researchers [5-8]. Specifically, it has been found that the Discrete Laplacian operator is an adequate approximation of the optimal regularization operator [5-6]. Various methods have been proposed to choose the regularization parameter and different techniques were applied to estimate

Acknowledgment: This work was supported by the PolyU Research Grant: Project account 340.813.A3.420.

the optimal regularization parameter [6-8]. While the proper choice of the global regularization weights has been addressed adequately in the literature, few studies have been undertaken for searching the optimal local regularization weights in R and S.

In this paper, we address how to derive the optimal weight matrices R and S. Their optimal relationship is derived directly from the optimal solution in iterative image restoration. Based on their optimal relationship, we then present a proper choice of weight matrices. The derived solution provides a mathematical backup of the viability of the heuristic solution suggested in the literature [3-4]. Experimental results are provided for comparative studies and they justify our proposed solution.

2. OPTIMAL LOCAL REGULARIZATION WEIGHTS

By substituting (1) into (4), the iterative restoration algorithm becomes

$$\hat{x}_{k+1} = \hat{x}_k + \beta[D^T R(y+n-D\hat{x}_k) - \alpha C^T S C \hat{x}_k] \quad (5)$$

Rewriting (5) yields

$$\hat{x}_{k+1} = \{\hat{x}_k + \beta D^T(y - D\hat{x}_k)\} + \beta\phi_k \quad (6)$$

where

$$\phi_k = D^T(R-I)D(x-\hat{x}_k) + D^T R n - \alpha C^T S C \hat{x}_k \quad (7)$$

By direct enumeration of (6), it is found that

$$\begin{aligned}\hat{x}_k = (I-\beta D^T D)^k x_0 + [I - (I-\beta D^T D)^k]x \\ +\beta\sum_{i=0}^{k-1}(I-\beta D^T D)^{k-1-i}\phi_i\end{aligned} \quad (8)$$

provided that the matrix $D^T D$ is invertible. The error in $\hat{x}_k$ can then be expressed by

$$\begin{aligned}e_k = \hat{x}_k - x = (I-\beta D^T D)^k(x_0 - x) \\ +\beta\sum_{i=0}^{k-1}(I-\beta D^T D)^{k-1-i}\phi_i\end{aligned} \quad (9)$$

In other words, the restoration error is bounded by

$$\begin{aligned}\|e_k\| &\leq \left\|(I-\beta D^T D)^k(x_0 - x)\right\| \\ &\quad +\beta\left\|\sum_{i=0}^{k-1}(I-\beta D^T D)^{k-1-i}\phi_i\right\| \\ &= E_1(x_0) + E_2(\alpha, C, S, R)\end{aligned} \quad (10)$$

It is observed that the error bound consists of two contributions. First, $E_1(x_0)$ is attributed to the initial estimate x_0. Second, $E_2(\alpha, C, S, R)$ denotes the error due to ϕ_i for $i = 0, \cdots, k-1$. Due to the property that

$$E_1(x_0) \quad \leq \quad \|I - \beta D^T D\|^k \cdot \|x_0 - x\| \quad (11)$$

we can see that, to minimize $E_1(x_0)$, an initial estimate which is as close to x as possible is required. However, if β is chosen in a way that $\|I - \beta D^T D\| < 1$, then

$$\lim_{k\to\infty} E_1(x_0) = 0 \quad (12)$$

Therefore, when the number of iteration is sufficiently large, the choice of initial estimate is insignificant in reducing the restoration error.

It is somewhat complicated to minimize E_2 as it involves a sequence of ϕ_i. To disentangle the complication, the relation between e_k and e_{k-1} is derived as

$$e_k = (I - \beta D^T D)e_{k-1} + \beta\phi_{k-1} \quad (13)$$

It is seen that the error in $\hat{x}_k$ is due to the error in $\hat{x}_{k-1}$ as well as ϕ_{k-1}. Therefore, a solution that minimizes the restoration error can be obtained by minimizing $\|\phi_k\|$ at each iteration.

In ϕ_k, there involves α, C, R and S. The regularization operator C should be chosen to describe some known properties of the original image, and it has been found that the Discrete Laplacian operator is an adequate approximation of the optimal regularization operator [5]. The regularization parameter α is used to trade the fidelity to the original image with the smoothness of the restored image and has to be predetermined. Therefore, both α and C can be treated as fixed factors in ϕ_k. Then the problem of interest is to look for an optimal pair of R and S that minimize $\|\phi_k\|$. In formulation, we have to find a pair of matrices R and S that minimize the functional

$$\begin{aligned}M(R,S) = \\ \left\|D^T R[D(x-\hat{x}_k)+n] - \alpha C^T S C\hat{x}_k - D^T D(x-\hat{x}_k)\right\|^2\end{aligned} \quad (14)$$

The minimization of $M(R,S)$ with respect to R and S lead to, respectively,

$$r_j = \frac{[D(x-\hat{x}_k)]_j + \left[\alpha(D^T)^{-1}C^T S C\hat{x}_k\right]_j}{[D(x-\hat{x}_k)]_j + [n]_j} \quad (15)$$

and

$$s_j = \frac{\left[(C^{-1})^T D^T (R-I) D(x-\hat{x}_k)\right]_j + \left[(C^{-1})^T D^T R n\right]_j}{[\alpha C\hat{x}_k]_j} \quad (16)$$

where r_j and s_j are the j-th diagonal elements of R and S respectively, and $[v]_j$ denotes the j-th element of vector v.

3. DETERMINATION OF THE OPTIMAL WEIGHT MATRICES

In previous section, we have derived the optimal relationship between the two weight matrices. In practical situation of interests, it is impossible to determine an optimal pair of R and S since both x and n is unknown. However, Eqns. (15) and (16) suggest a practical approach to estimate one matrix when another is given so that such pair of matrices can achieve a restoration performance close to the optimal. In this section, we will present the determination of a particular pair of R and S. By setting R to be the identity matrix, we estimate the optimal S based on their relationship. This case is of particular interest since $R = I$ is the typical solution appeared in the literature [3-4,6].

By substituting $R = I$ into the optimal relationship shown in (16), we have

$$s_j \left[\alpha C \hat{x}_k\right]_j = \left[(C^{-1})^T D^T n\right]_j \tag{17}$$

In the first place, we need to work out the correlation between $[\alpha C\hat{x}_k]_j$ and $[(C^{-1})^T D^T n]_j$ in order to obtain a good estimate of s_j. It is well-known that the restoration error consists of two independent contributions, namely regularization error and noise magnification error [3]. The first error is due to the use of regularization in restoration and its effect is related strongly to the local structures of the image x [3]. On the other hand, the noise magnification error results from the additive noise n. In view of this, the restored image $\hat{x}_k$ can be decomposed into two uncorrelated components, $(x_k + e_r)$ and e_n, where e_r and e_n denotes respectively the regularization error and the noise magnification error, and $x_k = x + (I - \beta D^T D)^k (x_0 - x)$, which is the ideal restored image after k iterations. Furthermore, in $[\alpha C\hat{x}_k]_j$, only the component $[\alpha C e_n]_j$ correlates with $[(C^{-1})^T D^T n]_j$.

The optimal estimate of s_j is obtained by minimizing the function

$$J_S = E\left\{\left(s_j[\alpha C\hat{x}_k]_j - [(C^{-1})^T D^T n]_j\right)^2\right\} \tag{18}$$

The minimization of J_S with respect to s_j is straightforward and leads to the equation

$$\hat{s}_j = \frac{E\left\{[\alpha C(x_k+e_r+e_n)]_j[(C^{-1})^T D^T n]_j\right\}}{E\left\{[\alpha C(x_k+e_r+e_n)]_j^2\right\}} \tag{19}$$

By utilizing the knowledge that $[\alpha C(x_k + e_r)]_j$ is uncorrelated with both $[\alpha C e_n]_j$ and $[(C^{-1})^T D^T n]_j$, and the property that $E[(C^{-1})^T D^T n]_j = E[\alpha C e_n]_j = 0$, we then have

$$\hat{s}_j = \frac{E\left\{[\alpha C e_n]_j[(C^{-1})^T D^T n]_j\right\}}{E\left\{[\alpha C(x_k+e_r)]_j^2\right\} + E\left\{[\alpha C e_n]_j^2\right\}} \tag{20}$$

Since C is in general a high-pass filter and e_n is the magnified noise originated from n, the term $\alpha C e_n$ can be regarded as Hn, where H is a high-pass operator. In that case, we can make the approximation that $[Hn]_j$ is a scaled version of n_j, i.e. $[\alpha C e_n]_j = [Hn]_j \doteq K_1 n_j$, where K_1 is a constant. By utilizing the above approximation, we have

$$\begin{aligned} E\left\{[\alpha C e_n]_j[(C^{-1})^T D^T n]_j\right\} &\doteq E\{K_1 n_j \sum_i f_{ji} n_i\} \\ &= K_1 \sum_i f_{ji} E\{n_j n_i\} = K_1 f_{jj} E\{n_j^2\} \doteq K_1 f_{jj} \sigma_n^2 \end{aligned} \tag{21}$$

and

$$E\left\{[\alpha C e_n]_j^2\right\} \doteq E\left\{(K_1 n_j)^2\right\} \doteq K_1^2 \sigma_n^2 \tag{22}$$

where σ_n^2 denotes the noise variance and f_{ji} denotes the i-th entry of matrix $(C^{-1})^T D^T$ at the j-th row. By using the approximation that $E\{[\alpha C(x_k+e_r)]_j^2\} \doteq K_2[\sigma_x^2]_j$, where $[\sigma_x^2]_j$ denotes the local spatial variance of x at the j-th location and K_2 is a constant, we have

$$\hat{s}_j \doteq \frac{1}{a + b\sigma_n^{-2}[\sigma_x^2]_j} \tag{23}$$

where $a = K_1/f_{jj}$ and $b = K_2/(K_1 f_{jj})$. Note that f_{jj} is the diagonal entries of matrix $(C^{-1})^T D^T$ and is constant over j as both C and D are Toeplitz.

4. DISCUSSION AND EXPERIMENTAL RESULTS

Lagendijk *et al.* [3] observed qualitatively that, in order to reduce restoration artifacts, the weights in R and S should be based on the local frequency content of the original image. In their experimental studies, R was chosen to be the identity matrix and S were computed as $s_j = 1/\{1 + \mu \cdot \max(0, [\sigma_y^2]_j - \sigma_n^2)\}$, where σ_y^2 is the local variance of the observed image and μ is a tuning parameter. On the other hand, Katsaggelos *et al.* [4] proposed a different formulation of adaptive regularized restoration algorithm, which is based on a set theoretic approach and incorporate the properties of the human visual system. Their solution suggests implicitly that R is the identity matrix and S is defined as $s_j = 1/(1+\theta[\sigma_x^2]_j)^2$, where θ is a tuning parameter. Recently, Reeves [6] adopted a Wiener filter interpretation of the space-variant regularization problem to elucidate the proper choice of the weight matrices. According to that interpretation, R is chosen to be the identity matrix (when the noise is white and identically distributed) and s_j is proportional to the reciprocal of the local variance of the original image.

We can see that the solution we derived, as shown in (23), maintains the heuristic appeals made in Refs. [3]

Table 1: SNR Improvement for the Image Blurred by a Horizontal Motion, and with White Gaussian Noise of Various SNRs.

SNR[a] (dB)	α	SNR Improvement[b] (dB)		
		Matrices proposed in [3][c]	Matrices proposed in [4][d]	Matrices we proposed[e]
15	0.6	3.09	3.05	3.16
20	0.5	3.63	3.60	3.69
25	0.3	4.32	4.29	4.46
30	0.0025	4.89	4.89	4.90

[a] SNR is defined as $10\log$(blurred image variance/noise variance).
[b] SNR Improvement is defined as $10\log(\|\tilde{y}-x\|^2/\|\hat{x}_k-x\|^2)$.
[c] Accordingly, $R=I$ and $s_j=1/\{1+\theta\max(0,[\sigma_{\tilde{y}}^2]_j-\sigma_n^2)\}$, with $\theta=0.01, 0.05, 0.1$ and 0.5 for the four various SNRs.
[d] Accordingly, $R=I$ and $s_j=1/(1+\theta[\sigma_x^2]_j)^2$, with $\theta=0.001, 0.001, 0.05$ and 0.1 for the four various SNRs.
[e] $R=I$ and the entries in S are given in (23) with a and b set to be 1.

and [4]. While the tuning parameters in their solutions are experimentally determined, ours is clearly related to the noise variance. Our solution is also consistent with that provided in Ref. [6]. Above all, our proposed weight matrices are derived from a quantitative analysis of the optimal solution in iterative image restoration.

Experiments had been carried out to evaluate the performance of the proposed weight matrices in the adaptive iterative restoration. The well-known image 'Lenna' of size 256×256 was applied. The degraded images were obtained by artificially blurring the original with a horizontal motion blur over 9 pixels and adding white Gaussian noise at 15, 20, 25 and 30 dB SNR. The noisy blurred images were then restored by the iterative algorithm with the proposed weight matrices. For comparative studies, the weight matrices suggested in Refs [3] and [4] were also evaluated. In all experiments, the discrete Laplacian is used as C and β was equal to 1. The local variance of the original image was estimated at each iteration based on the available restored image. The criterion $\|\hat{x}_{k+1}-\hat{x}_k\|^2/\|\hat{x}_k\|^2 \leq 10^{-6}$ was used in our experiments to terminate the iteration. The restoration results obtained are reported in Table 1. From the experimental results, it is shown that better restoration results, in terms of the SNR improvement, were obtained with the proposed weight matrices. That shows the superiority of our proposed solution.

5. CONCLUSIONS

In the study of adaptive regularized image restoration, little effort has been made to analyze and search the optimal local regularization weightings. In this paper, we have facilitated a search of optimal local regularization through a quantitative analysis of the optimal solution in the context of iterative image restoration. An optimal relationship between local regularization weights has been derived. It has been shown that the weight matrices we derived from their optimal relationship can achieve better restoration results than those suggested in the literature.

6. REFERENCES

[1] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-posed Problems.* New York, Winston, 1977.

[2] J. Biemond, R. L. Lagendijk and R. M. Mersereau, "Iterative methods for image deblurring," *Proc. IEEE,* Vol. 78, pp. 856-883, 1990.

[3] R. L. Lagendijk, J. Biemond and D. E. Boekee, "Regularized iterative image restoration with ringing reduction," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. 36, pp. 1874-1887, 1988.

[4] A. K. Katsaggelos, J. Biemond, R. W. Schafer and R. M. Mersereau, "A regularized iterative image restoration algorithm," *IEEE Trans. Signal Processing,* Vol. 39, pp. 914-929, 1991.

[5] N. B. Karayiannis and A. N. Venetsanopoulos, "Regularization theory in image restoration — the stabilizing functional approach," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. 38, pp. 1155-1179, 1990.

[6] S. J. Reeves, "Optimal space-varying regularization in iterative image restoration," *IEEE Trans. Image Processing,* Vol. 3, pp. 319-324, 1994.

[7] 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.

[8] A. M. Thompson, J. C. Brown, J. W. Kay and D. M. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Pattern Anal. and Machine Intell.,* Vol. 13, pp. 326-339, 1991.