SELECTION OF VARYING SPATIALLY ADAPTIVE REGULARIZATION PARAMETER FOR IMAGE DECONVOLUTION

Dmitriy Paliy[a], *Vladimir Katkovnik*[a], *Sakari Alenius*[b], *Karen Egiazarian*[a]

[a]Instutite of Signal Processing, Tampere University of Technology,
P.O. Box 553, 33101 Tampere, Finland, e-mail: firstname.lastname@tut.fi,
[b]Nokia Research Center, Tampere, Finland, e-mail: firstname.lastname@tut.fi.

ABSTRACT

The deconvolution in image processing is an inverse ill-posed problem which necessitates a trade-off between fidelity to data and smoothness of a solution adjusted by a regularization parameter. In this paper we propose two techniques for selection of a varying regularization parameter minimizing the mean squared error for every pixel of the image. The first algorithm uses the estimate of the squared point-wise bias of the regularized inverse. The second algorithm is based on direct multiple statistical hypothesis testing for the estimates calculated with different regularization parameters. The simulation results on images illustrate the efficiency of the proposed technique.

Introduction

A variety of image capturing principles can be modeled by the integral $z(x) = \int_X v(x,t)y(t)dt$, $x,t \in X \subset \mathbb{R}^2$, where v is a point-spread function (PSF) of the system, y is an image intensity function, and z is an observed image. A natural simplification is that the PSF v is shift-invariant which leads to the convolution operation in the observation model. We consider a discrete noisy approximation of the problem so that the observation z is given in the following form:

$$z(x) = (v \circledast y)(x) + \varepsilon(x), \tag{1}$$

where " $\circledast$ " denotes the discrete convolution, the argument x is defined on the regular $n_1 \times n_2$ lattice, $x \in X = \{(x_1,x_2) : x_i = 0,1,...,n_i-1,\ i=1,2\}$, and ε is a random noise. It is assumed that the noise is white Gaussian with zero-mean and variance σ^2, $\varepsilon(x) \sim \mathbb{N}(0, \sigma^2)$.

In the 2D frequency domain for the circular convolution the model (1) takes a form:

$$Z(\omega) = V(\omega)Y(\omega) + \varepsilon(\omega), \tag{2}$$

where the notation $\mathcal{F}\{\cdot\}$ is used for the discrete Fourier transform (DFT) $Z(\omega) = \mathcal{F}\{z(x)\}$, $V(\omega) = \mathcal{F}\{v(x)\}$, $Y(\omega) = \mathcal{F}\{y(x)\}$, $\varepsilon(\omega) = \mathcal{F}\{\varepsilon(x)\}$, and $\omega \in W$, $W = \{(\omega_1,\omega_2) : \omega_i = k_i/n_i,\ k_i = 0,1,...,n_i-1,\ i=1,2\}$, is the normalized 2D discrete frequency.

The estimation of y from the observation z is a removal of the degradation caused by a PSF. It is an inverse problem. Usually this problem is ill-posed which results in instability of the solution which, in particular, is very sensitive with respect to the additive noise.

Stabilizing effects can be introduced by constraints imposed on the solution. A general approach to this kind of constrained estimation refers to the methods of Lagrange multipliers and the Tikhonov regularization [1]. The regularized (constrained) inverse (RI) filter can be obtained as a solution of the least square problem with a penalty term:

$$J = \|Z - VY\|_2^2 + r\,\|Y\|_2^2, \tag{3}$$

where $r \geq 0$ is a regularization parameter and $\|\cdot\|_2$ denotes Euclidean norm. Here, the first term $\|Z - VY\|_2^2$ evaluates the fidelity of the model VY to the available data Z and the second term $\|Y\|_2^2$ bounds the power of this estimate. The regularization parameter r balances these two terms in the criterion J. In (3), and further, we omit the argument ω in the Fourier transforms. We obtain the estimate of the image by minimizing (3):

$$\hat{y}_r(x) = \mathcal{F}^{-1}\{\hat{Y}\},\ \hat{Y} = \frac{V^*}{|V|^2 + r}Z, \tag{4}$$

where the star (*) means the complex-conjugate variable.

A proper selection of the regularization parameter r in (4) is a key point of the regularization technique overall. There are numerous publications concerning this problem.

Roughly speaking there are two types of methods: with a prior knowledge and without a prior knowledge about the noise variance σ^2. The L-curve method, sometimes also called the Tikhonov curve method, belongs to the group of methods with no information on the value of σ^2 (e.g. Miller [2], and Tikhonov and Arsenin [1]). The technique uses a $\log - \log$ plot with the $\log \left\|Z - V\hat{Y}\right\|_2^2$ as an abscise and $\log \left\|\hat{Y}\right\|_2^2$ as an ordinate, with r as a parameter along this curve. The transition between under- and over-regularization corresponds to the "corner" of the L-curve and the corresponding value of r is proposed as an optimal value of the regularization parameter. Further, Hansen has developed this idea in [3] where he has stipulated conditions when the corner exists. The corner is defined as the maximal curvature point of the $\log - \log$ plot. Methods for detection of this point can be seen in [4], [5].

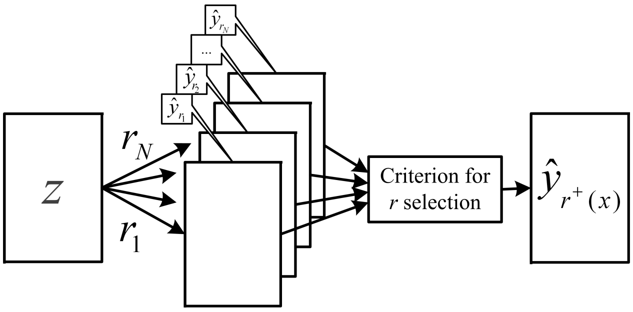


Figure 1. Brief scheme of the regilarization parameter r selection using the proposed techniques.

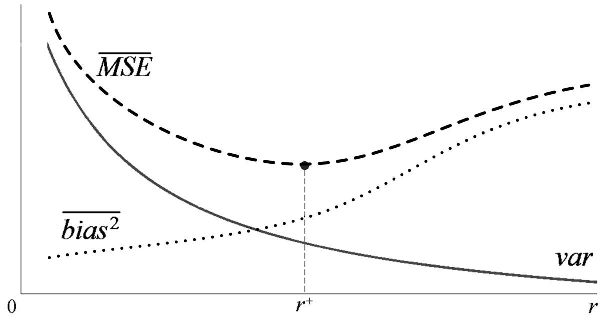


Figure 2. MSE, bias, and variance as a function of the regularization parameter.

Galatsanos and Katsaggelos [6] propose a technique for selection of the asymptotically optimal regularization parameter provided that the variance of noise in (1) is known. This approach is based on calculation of the derivative of the mean squared error (MSE) functional. Similar idea is exploited by Nielamani et al. in [7] where the optimal invariant regularization parameter is found by minimizing an upper bound of MSE calculated in the Fourier-Wavelet domain. Wufan Chen et al. use in [8] the iterative constrained total least-square adaptive procedure, and show its stability and convergence. The review of the methods for invariant regularization parameter selection can be found in [6] and [9].

Berger et al. in [10] propose to use a spatially varying regularization parameter and describe a method based on the local weighted standard deviation analyzing the difference signal of the estimate. Wu et al. in [11] choose both the spatially adaptive regularization parameter and regularization operator by estimation of the local noise variance and detecting edges in the image.

Also, review on the regularization parameter selection methods in inverse problems can be found in [12].

In this paper we develop two methods based on minimization of the point-wise mean squared error. The first algorithm is based on estimation of the squared point-wise bias. The second algorithm is based on direct multiple statistical hypothesis testing for the estimates calculated with different regularization parameters. The intersection of the confidence interval (ICI) rule is used in this testing [13], [14].

The structure of the paper is as follows. In Section 1 we propose two techniques for regularization parameter selection. Results of simulations are given in Section 2, and finally conclusions are done.

1. SPATIALLY ADAPTIVE REGULARIZATION TECHNIQUES

The MSE of the estimator $\hat{y}_r(x)$, $\forall x \in X$, can be decomposed as a sum of the variance of the random component and the squared bias (error of the deterministic component):

$$\mathrm{MSE}(r,x) = E\left\{e^2(x)\right\} = bias^2(r,x) + var(r), \quad (5)$$

where $e(x) = y(x) - \hat{y}_r(x)$.

According to the Parseval theorem the variance term is computed as

$$var(r) = \frac{\sigma^2}{n_1 n_2} \|H_V\|_2^2, \quad (6)$$

where $H_V = \frac{V^*}{|V|^2+r}$. It is a monotonically decreasing function of r with $var(r) \to 0$ as $r \to \infty$.

The bias term $bias^2(r,x)$ depends on the unknown true signal $y(x)$ and calculated as $bias^2(r,x) = ((y \circledast h_r)(x))^2$, where $h_r = \mathcal{F}^{-1}\{H_r\}$, $H_r = \frac{r}{|V|^2+r}$. Under quite mild assumptions on v and y for small r, $bias^2(r,x) \to 0$ as $r \to 0$, and for large r, $bias^2(r,x) \to y^2(x)$ as $r \to \infty$. Assuming a monotonic increasing of the bias as a function of r we arrive to the curves for the variance and the bias shown in Figure 2. In this case always there is a unique minimum value of the MSE achieved at $r^+(x) = \arg\min_r(\mathrm{MSE}(r,x))$.

Our goal is selection of the values of the regularization parameter close to the optimal $r^+(x)$.

In general this optimal value is different for different x as the bias depends on the estimated signal $y(x)$. Thus, the varying point-wise adaptive selection of the regularization parameter has the point-wise optimization of the accuracy of the image reconstruction as a main motivation. In Section

1.1. Bias estimation approach

In this approach we evaluate the criterion (5) and use its minimization for estimation of the optimal regularization parameter $r^+(x)$. Provided a given value of σ^2 the variance $var(r)$ in the criterion (5) can be easily calculated for any r. The situation is much more difficult for the bias as it is signal-dependent.

It can be verified that $E\left\{\left(\mathcal{F}^{-1}\left\{H_r \frac{Z}{V}\right\}\right)^2\right\} - \frac{\sigma^2}{n_1 n_2}\left\|\frac{H_r}{V}\right\|_2^2 = ((y \circledast h_r)(x))^2$. Inserting this expression for the squared bias in (5) we represent the criterion

Test images	Oracle invariant	Bias estimation	ICI rule	Oracle varying
Cheese 128×128	5.90	6.44	6.80	9.19
Cameraman 256×256	5.61	5.88	6.14	8.72
Lena 256×256	6.16	6.31	6.47	9.17
Barbara 512×512	3.76	3.74	4.14	6.50
Boat 512×512	6.39	6.58	6.66	9.35

Table 1. ISNR values in dB for RI technique.

in the form

$$\mathrm{MSE}(r,x) = E\left\{\left(\mathcal{F}^{-1}\left\{H_r\frac{Z}{V}\right\}\right)^2\right\}+ +\frac{\sigma^2}{n_1n_2}\|H_V\|_2^2-\frac{\sigma^2}{n_1n_2}\left\|\frac{H_r}{V}\right\|_2^2.$$

Then a natural random and unbiased estimate of the criterion is

$$\mathrm{MSE}(r,x) \simeq \left(\mathcal{F}^{-1}\left\{H_r\frac{Z}{V}\right\}\right)^2+ +\frac{\sigma^2}{n_1n_2}\|H_V\|_2^2-\frac{\sigma^2}{n_1n_2}\left\|\frac{H_r}{V}\right\|_2^2, \tag{7}$$

where the right hand side can be calculated as it depends only on the observed signal z.

However, it is not so simple as the inverse of V used in (7) is quite similar to the original inverse problem in (1)-(2) and we have here the same calculating difficulties.

It is natural to replace this inverse by the regularized inverse (4) and rewrite (7) as follows

$$\mathrm{MSE}\,(r,x)\simeq\left(\mathcal{F}^{-1}\left\{H_r\frac{V^*Z}{|V|^2+\alpha}\right\}\right)^2+ \frac{\sigma^2}{n_1n_2}\|H_V\|_2^2-\frac{\sigma^2}{n_1n_2}\left\|H_r\frac{V^*}{|V|^2+\alpha}\right\|_2^2. \tag{8}$$

Here α is the regularization parameter influencing the accuracy of the approximation of $\mathrm{MSE}\,(r,x)$ by the right-hand side of (8).

The new regularization actually means that the bias error calculated originally as $bias\,(r,x)=\mathcal{F}^{-1}\{H_rY\}$ is replaced by $\mathcal{F}^{-1}\{H_r\frac{|V|^2}{|V|^2+\alpha}Y\}$. For small α this change is not significant for the bias and for the following calculations as it has been verified by simulation.

Replacing the criterion $\mathrm{MSE}\,(r,x)$ by the estimate we calculate the estimate of $r^+(x)$ as follows

$$\hat{r}^+(x)=\arg\min_r\left[\left(\mathcal{F}^{-1}\left\{H_r\frac{V^*Z}{|V|^2+\alpha}\right\}\right)^2+ \frac{\sigma^2}{n_1n_2}\|H_V\|_2^2-\frac{\sigma^2}{n_1n_2}\left\|\frac{H_rV^*}{|V|^2+\alpha}\right\|_2^2\right]. \tag{9}$$

In practice the estimates $\widehat{y}_r(x)$ are calculated for the set of regularization parameters $r\in R=\{r_1,r_2,...,r_N\}$, where $r_1<r_2<...<r_N$. Using these N estimates the minimization of criterion function (9) yields the value $\hat{r}^+$ for every pixel x.

In practice, it is implemented as follows. The estimates $\widehat{y}_r(x)$ are calculated for a set of regularization parameters $r\in R=\{r_1,r_2,...,r_N\}$ according to (4) (Fig.1). Using these N estimates, the minimization of criterion function (9) yields the value r^+ for every point x (block "Criterion for r selection" in Fig.1). The final estimate $\widehat{y}_{r^+(x)}(x)$ is built using corresponding estimates.

1.2. ICI rule approach

In this approach selection of the regularization parameter is based on comparison of the signal estimates calculated with different values of the regularization parameter. The ICI concept developed originally for the window size selection in signal/image denoising is described in a number of publications [13], [14], [15]. It gives the window size close to its ideal value minimizing the corresponding MSE.

First, we wish to note that the ICI rule is applicable for selection of the regularization parameter. Indeed, the output $\widehat{y}_r(x)$ of the regularized inverse filter is the estimate of the signal y. As it is illustrated in Fig.2 the bias and the variance of this estimate are increasing and decreasing functions of the regularization parameter, respectively. This behavior of the bias and the variance are the basic assumptions used in derivation and justification of the ICI rule for the window size selection [15], [16]. It follows that application of the ICI rule to the estimate $\widehat{y}_r(x)$ with varying threshold parameter $r(x)$ allows to find the value of the regularization parameter close to the one minimizing the criterion (5).

The idea of the ICI rule is as follows. The estimates $\widehat{y}_r(x)$ are calculated for $r\in R$. The adaptive parameter $\hat{r}^+$ is defined as the largest of those r in R which estimate does not differ significantly from the estimates corresponding to the smaller values of r. This idea is implemented as follows. We consider a sequence of confidence intervals

$$D_s=\left[\widehat{y}_{r_s}(x)-\Gamma_s\sigma_{\hat{y}_{r_s}(x)},\widehat{y}_{r_s}(x)+\Gamma_s\sigma_{\hat{y}_{r_s}(x)}\right],$$

where $s=1,..,N$, $\Gamma_s>0$ is a parameter and $\sigma_{\hat{y}_{r_s}(x)}$ is the standard deviation of the estimate $\widehat{y}_{r_s}$, calculated according to (6) as $\sigma_{\hat{y}_{r_s}(x)}=\sqrt{var(r)}$. The ICI rule is stated as follows: consider the intersection of the confidence intervals $I_s=\bigcap_{i=1}^{s}D_i$ and let s^+ be the largest of the indices s for which I_s is non-empty. Then the adaptive regularization parameter r^+ is defined as $\hat{r}^+ = r_{s^+}$ and the corresponding adaptive estimate is $\widehat{y}_{r^+}(x)$. This procedure is repeated for each x and, in this way, the adaptive $r^+(x)$ is spatially varying.

In general, the ICI allows the maximum degree of smoothing the estimate stopping before oversmoothing begins [17].

The parameter Γ_s is an important element of the algorithm as it says when a difference between estimate deviations is large or small. Too large value of this parameter leads to signal oversmoothing and too small value leads to undersmoothing. Usually the ICI rule is used with a constant Γ_s for all s intervals. Optimization of Γ_s can be produced from some heuristic and theoretical considerations (e.g. [13], [14], [15], [18]). In this paper we use a varying parameters Γ_s with different values for the intervals D_s, $s = 1,..,N$. It is worth to note that this set of the parameters Γ_s is fixed for all our experimental tests.

2. SIMULATIONS

The results are obtained by the Monte-Carlo simulation (100 runs) for the following test images: *Cameraman* of the size 256×256 (Fig.3a), *Lena* 256×256, *Barbara* 512×512, *Boat* 512×512, and *Cheese* 128×128. *Cheese* is a black-and-white image with two regions of constant values of the image intensity.

The blur is defined as the 9×9 boxcar point-spread function (mean filter). The level of the Gaussian noise is calculated as the blurred signal-to-noise ratio (BSNR)

$$\mathrm{BSNR} = 10 \log_{10} \left(\frac{1}{\sigma^2} \left\| v \circledast y - \frac{1}{n_1 n_2} \sum_{x \in X} (v \circledast y)(x) \right\|_2^2 \right)$$

equal to 40 dB in our experiments.

A selection of the set $R = \{r_1, r_2, ..., r_N\}$ of the regularization parameter values is a special problem as it should cover an interval including the optimal value. In general, these optimal value depends on the image, the point-spread function and the noise level. The following construction defines the set R in the manner universally applicable for variety of scenarios:

$$r_i = \tau_i \cdot \frac{n_1 n_2 \sigma^2 \left\| v(x) \right\|_1^2}{\sum_f |Z(f)|^2 - \sigma^2 n_1 n_2}, \; i = 1, ..., 5, \qquad (10)$$

where $\tau \in \{1.3,\ 2.3,\ 6.3,\ 9.6,\ 30\}$ and $\|v(x)\|_1^2 = = (\sum_x |v(x)|)^2$. It is employed for all test images used in our simulation experiments and similar to implementation proposed in [7].

The criterion used to evaluate the performance of the proposed techniques is improved signal-to-noise ratio (ISNR): $\mathrm{ISNR} = 10 \log_{10} \left(\frac{\|y - z\|_2^2}{\|y - \widehat{y}\|_2^2} \right)$ dB.

The results of statistical Monte Carlo experiments are presented in Table 1. The first column gives the names of the test images. The third and the forth columns gives ISNR values obtained by the bias estimation (9) and by the ICI rule, respectively. The parameter α in (9) is set to be 0.001.

We compare these results with the reference ideal (oracle) results which correspond to minimization of MSE by selection of the optimal invariant (single r for all x) and the optimal varying (optimal r for each x) regularization parameter assuming that the true image y is known. The column "Oracle invariant" of Table 1 contains the values of ISNR obtained using the optimal invariant regularization parameter $r_{inv} = \arg\min_{r \in R} \sum_x (y(x) - \widehat{y}_r(x))^2$. The column "Oracle varying" contains the values of ISNR obtained using the varying regularization parameter $r_{var}(x) = \arg\min_{r \in R} (y(x) - \widehat{y}_r(x))^2$.

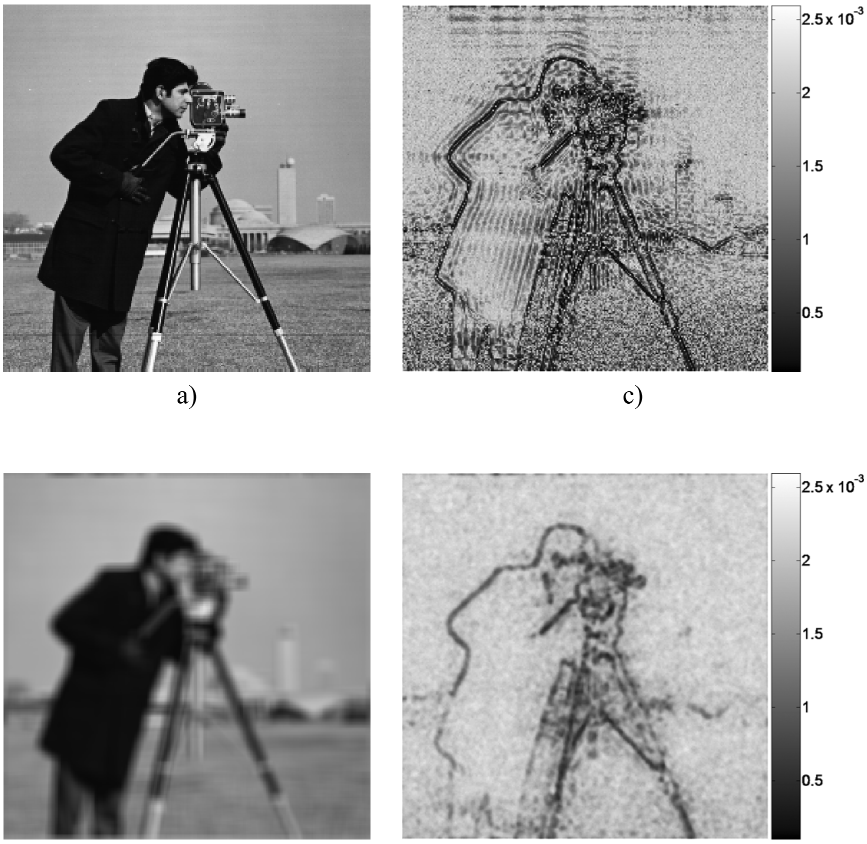
The set of threshold parameters Γ_s used in the ICI rule is $\{0.86,\ 0.5,\ 0.28,\ 0.29,\ 1.1\}$. The results given in the table allow the following conclusions. First of all, the both adaptive estimates show the results which are better than those obtained by the oracle invariant estimate. i.e. the estimates with the varying adaptive regularization parameter work better that the ideal estimate with the best possible selection of the invariant regularization parameter. Comparison of the criterion minimization method versus the ICI rule is in favor of the latter for all test images.

However, we can see that the varying oracle estimates demonstrate much better performance in comparison with all other estimates. A difference in performance of this approach versus the estimate with the invariant oracle selection confirms a general motivation of this paper to improve the estimation by using the varying regularization parameter. In the same time the results obtained by the adaptive regularization parameter estimates are quite far from the varying oracle ones. It says that there are resources for further improvement of the adaptive varying regularization parameter algorithms.

Visually the effects of point-wise regularization are demonstrated in Fig.3c,d, where the results of Monte-Carlo modeling are shown for the blurred Cameraman image (Fig.3b). These images show the mean values of the regularization parameters for different pixels of the Cameraman image. The varying oracle values of the regularization parameter r_{var} are small near edges while larger values correspond to the smoother areas of the image (Fig.3c). The mean values of the varying adaptive regularization parameter selected by the ICI rule are shown in Fig.3d. Comparing the images Fig.3c and Fig.3d we note their similarity. In particular, the ICI rule gives smaller values of the regularization parameter near the edges and larger for smoother areas as the oracle estimator does. The adaptive values of the regularization parameter accurately delineate the edges of the image similarly to Fig.3c.

3. CONCLUSIONS

Two novel approaches for spatially adaptive selection of the regularization parameter for deconvolution problems were proposed in this paper. The first algorithm is based on the bias estimation and MSE minimization for every point of an image. The second one uses the statistical ICI rule. The algorithm based on the ICI rule demonstrated



a) c)

b) d)

Figure 3. Results obtained by Monte-Carlo (100 runs) silmulations: a) true $Cameraman$ test image; b) $Cameraman$ test image blurred with 9x9 boxcar PSF and noisy (white Gaussian noise) with BSNR=40dB. c) values of the varying $oracle$ regularization parameters; d) varying regularization parameters obtained by the ICI rule.

better performance than the algorithm using the bias estimate. Comparison of the proposed adaptive techniques' versus the optimal (oracle) invariant selection technique for various images showed that the proposed techniques performs better than the estimate with the best possible invariant regularization parameter.

4. ACKNOWLEDGMENTS

This work was supported by the Finnish Funding Agency for Technology and Innovation (Tekes) and by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence program (2006 - 2011).

5. REFERENCES

[1] Tikhonov, A. N., Arsenin, V. A.: Solutions of Ill-posed Problems. Winston & Sons, Washington (1977).

[2] Miller. K.: Least squares methods for ill-posed problems with a prediscribed bound. SIAM J. Math.Anal., vol. 1, (1970), 52-74.

[3] Hansen, P.C.: Analysis of discrete ill-posed problems by menas of the L-curve. SIAM Rev., vol. 34, (1992), 561-580.

[4] Hansen, P.C., O'Leary, D.P.: The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J. Sci.Comput., vol. 14, no. 6, (1993) 1487-1503.

[5] Oraintara. S., Karl, W.C., Castanon, D.A., Nguyen, T.Q.: A method for choosing the regularization parameter in generalized Tikhonov regularized linear inverse problems. International Conference on Image Processing 2000, Proceedings, vol. 1, 10-13, (2000) 93-96.

[6] Galatsanos, N.P., Katsaggelos, A.K.: Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. IEEE Trans. on Image Processing, vol. 1, issue 3, (1992) 322-336.

[7] Neelamani, R., Choi, H., Baraniuk, R.G.: Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. IEEE Trans. on Signal Processing, vol. 52, issue 2, (2003) 418-433.

[8] Chen, W., Chen, M., Zhou, J.: Adaptively regularized constrained total least-squares image restoration. IEEE Trans on Image Processing, vol. 9, issue 4, (2000) 588-596.

[9] Thompson, A.M., Brown, J.C., Kay, J.W., Titterington, D.M.: A study of methods of choosing the smoothing parameter in image restoration by regularization. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 13, issue 4, (1991) 326-339.

[10] Berger T., Stromberg, J.O., Eltoft, T.: Adaptive regularized constrained least squares image restoration. IEEE Trans. on Image Processing, vol. 8, issue 9, (1999) 1191-1203.

[11] Wu, X., Wang, R., Wang, C.: Regularized image restoration based on adaptively selecting parameter and operator. Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 3, 23-26, (2004) 662-665.

[12] Vogel C.R., Computational Methods for Inverse Problems, Soc. for Industrial and Applied Mathematics, 2002.

[13] Katkovnik, V.: A new method for varying adaptive bandwidth selection. IEEE Trans. on Signal Proc., vol. 47, no. 9, (1999) 2567-2571.

[14] Katkovnik V., Egiazarian, K., Astola, J.: A spatially adaptive nonparametric regression image deblurring. IEEE Trans. on Image Processing, Vol. 14, Issue 10, (2005) 1469 - 1478.

[15] Katkovnik, V., Egiazarian, K., Astola, J.: Local Approximation Techniques in Signal and Image Processing. SPIE Press, Monograph Vol. PM157, Hardcover, 576 pages, ISBN 0-8194-6092-3, (2006).

[16] Goldenshluger, A., Nemirovski, A.: On spatial adaptive estimation of nonparametric regression. Math. Meth. Statistics, vol. 6, (1997) 135 – 170.

[17] Foi, A., Katkovnik, V., Egiazarian, K., Astola, J., Inverse halftoning based on the anisotropic LPA-ICI deconvolution. In: Astola, J. et al. (eds). Proceedings of The 2004 International TICSP Workshop on Spectral Methods and Multirate Signal Processing, SMMSP 2004, Vienna, Austria, (2004) 49 – 56.

[18] Stanković, L.: Performance analysis of the adaptive algorithm for bias-to-variance tradeoff. IEEE Trans. on Signal Proc., vol. 52, No. 5, (2004) 1228 – 1234.