Czech Pattern Recognition Workshop 2000, Tomáš Svoboda (Ed.)
Peršlák, Czech Republic, February 2–4, 2000
Czech Pattern Recognition Society

Edge Preserving Model-Based Image Restoration

Michal Haindl

Institute of Information Theory and Automation
Academy of Sciences
and Center for Machine Perception, Czech Technical University
Prague, Czech Republic, CZ182 08
haindl@utia.cas.cz

Abstract *This paper presents a derivation of a fast recursive filter for image restoration if degradation obeys a linear degradation model with the unknown possibly non-homogeneous point-spread function. Pixels in the vicinity of steep discontinuities are left unrestored to minimize restoration blurring effect. The degraded image is assumed to follow a causal simultaneous regressive model and the point-spread function is estimated using the local least-square estimate.*

1 Introduction

Physical imaging systems, the recording medium, the atmosphere are imperfect and thus a recorded image represents a degraded version of the original scene. Similarly an image is usually further corrupted during its processing, transmission or storage. Possible examples are lens defocusing or aberration, noisy transmission channels, motion between camera and scene, etc. The image restoration task is to recover an unobservable image given the observed corrupted image with respect to some statistical criterion. Image restoration is the busy research area for already several decades and many restoration algorithms have been proposed. The simplest restoration method is to smooth the data with an isotropic linear or non-linear shift-invariant low-pass filter. Usual filtering techniques (e.g. median filter, Gaussian low pass filter, band pass filters, etc.) tend to blur the location of boundaries. Several methods [16] try to avoid this problem by using a large number of low-pass filters and combining their outputs. Similarly anisotropic diffusion [17],[5] addresses this problem but it is computationally extremely demanding. Image intensity in this method is allowed to diffuse over time, with the amount of diffusion at a point being inversely proportional to the magnitude of local intensity gradient. A nonlinear filtering method developed by Nitzberg and Shiota [15] uses an offset term to displace kernel centers away from presumed edges and thus to preserve them, however it is not easy to propose all filter parameters to perform satisfactory on variety of different images and the algorithm is very slow.

In the exceptional case when the degradation point-spread function is known the Wiener filter [1] or deconvolution methods [11] can be used.

Model-based methods use most often Markov random field type of models either in the form of wide sense Markov (regressive models) or strong Markov models. The noncausal regressive model used in [3],[4] has the main problem in time consuming iterative solution based on the conjugate gradient method. Similarly Markov random field based restoration methods [7], [6], [12] require time consuming application of Markov chain Monte Carlo methods. Besides this both approaches have solve the problem when to stop these iterative processes. A similar combination of causal and noncausal regressive models as in this paper was used in [13]. However they assume the homogeneous point-spread function and they identify all parameters simultaneously using extremely time consuming iterations of the EM algorithm which is not guaranteed to reach the global optimum. It is seldom possible to obtain a degradation model analytically from the physics of the problem. More often a limited prior knowledge supports only some elementary assumptions about this process. Usual assumption, accepted also in this work, is that the corruption process can be modeled using a linear degradation model.

2 Image Model

Suppose Y represents a true but unobservable image defined on finite rectangular $N \times M$ underlying lattice I. The observable data are X, a version of Y distorted by noise independent of the signal. We assume knowledge of all pixels elements from the reconstructed scene. For the treatment of the more difficult problem when some date are missing see [9], [10]. The image degradation is supposed to be approximated by the linear discrete spatial domain degradation model

$$X_r = \sum_{s \in I_r} h_s Y_{r-s} + \epsilon_r \qquad (1)$$

where h is a discrete representation of the unknown point-spread function. The point-spread function is assumed to be either homogeneous or it can be non-homogeneous but in this case we assume its slow changes relative to the size of an image. I_r is some contextual support set, and noise ϵ is uncorrelated with the true image, i.e.,

$$E\{Y\,\epsilon\} = 0 \;. \tag{2}$$

The point-spread function is unknown but such that we can assume the unobservable image Y to be reasonably well approximated by the expectation of the corrupted image

$$\hat{Y} = E\{X\} \tag{3}$$

in regions with gradual pixel value changes. The above method (3) changes all pixels in the restored image and thus blurs discontinuities present in the scene although to much less extent than the classical restoration methods due to adaptive restoration model (11). This excessive blurring can be avoided if pixels with steep step discontinuities are left unrestored, i.e.,

$$\hat{Y}_r = \begin{cases} E\{X_r\} & \text{if (5) holds} \\ X_r & \text{otherwise} \end{cases} \;, \tag{4}$$

where the adaptive condition (5) is

$$|\, E\{X_r\} - X_r \,| < \frac{1}{n_s} \sum_s |\, E\{X_{r-s}\} - X_{r-s} \,| \;. \tag{5}$$

The expectation (3) can be expressed as follows:

$$\begin{aligned} E\{X\} &= \int x\, p(x)\, dx \\ &= \int \begin{pmatrix} x_1 & x_2 & \dots & x_M \\ x_{M+1} & x_{M+2} & \dots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{NM-M+1} & x_{NM-M+2} & \dots & x_{NM} \end{pmatrix} \\ &\quad \prod_{r=1}^{NM} p(x_r \,|\, X^{(r-1)})\, dx_1 \dots dx_{NM} \end{aligned} \tag{6}$$

where

$$X^{(r-1)} = \{X_{r-1}, \dots, X_1\} \tag{7}$$

is a set of noisy pixels in some chosen but fixed ordering. For single matrix elements in (6) it holds

$$\begin{aligned} E\{X_j\} &= \int x_j \prod_{r=1}^{NM} p(x_r \,|\, x^{(r-1)})\, dx_1 \dots dx_{NM} \\ &= \int X_j \prod_{r=1}^{j} p(X_r \,|\, X^{(r-1)})\, dX_1 \dots dX_j \\ &= \int E\{X_j \,|\, X^{(j-1)}\} \prod_{r=1}^{j-1} p(X_r \,|\, X^{(r-1)}) \\ &\quad dX_1 \dots dX_{j-1} \\ &= E_{X^{(j-1)}}\{\, E_{X_j}\{X_j \,|\, X^{(j-1)}\}\,\} \end{aligned} \tag{8}$$

Let us approximate after having observed $x^{(j-1)}$ the

$$\hat{Y}_j = E\{X_j\}$$

by the

$$E\{X_j \,|\, X^{(j-1)} = x^{(j-1)})$$

where $x^{(j-1)}$ are known past realization for j. Thus we suppose that all other possible realization $x^{(j-1)}$ than the true past pixel values have negligible probabilities. This assumption implies conditional expectations approximately equal to unconditional ones, i.e.,

Then the expectation (8) is

$$E\{X_j\} \approx E\{X_j \,|\, X^{(j-1)}\} \;, \tag{9}$$

and

$$\hat{Y} = E\{X\} \;\approx \tag{10}$$

$$\begin{pmatrix} E\{X_1 \,|\, x^{(0)}\} & \dots & E\{X_M \,|\, x^{(M-1)}\} \\ E\{X_{M+1} \,|\, x^{(M)}\} & \dots & E\{X_{2M} \,|\, x^{(2M-1)}\} \\ \vdots & \ddots & \vdots \\ E\{X_{NM-M+1} \,|\, x^{(NM-M)}\} & \dots & E\{X_{NM} \,|\, x^{(NM-1)}\} \end{pmatrix}$$

Suppose further that the noisy image can be represented by an adaptive causal simultaneous autoregressive model

$$X_r = \sum_{s \in I_r^c} a_s X_{r-s} + \epsilon_r \;, \tag{11}$$

where ϵ_r is a white Gaussian noise with zero mean and a constant but unknown variance σ^2. The noise is uncorrelated with data from a causal neighbourhood I_r^c. The model adaptivity is introduced using the standard exponential forgetting factor technique in parameter learning part of the algorithm. The model can be written in the matrix form

$$X_r = \gamma Z_r + \epsilon_r \;, \tag{12}$$

where

$$\begin{aligned} \gamma &= [a_1, \dots, a_\eta] \;, \qquad (13) \\ \eta &= card(I_r^c) \qquad (14) \end{aligned}$$

and Z_r is a corresponding vector of X_{r-s}. To evaluate conditional mean values in (10) the one-step-ahead prediction posterior density $p(X_r \,|\, X^{(r-1)})$ is needed. If we assume the normal-gamma parameter prior for parameters in (11) (alternatively we can assume the Jeffreys parameter prior) this posterior density has the form of Student's probability density

$$p(X_r|X^{(r-1)}) = \frac{\Gamma(\frac{\beta(r)-\eta+3}{2})}{\Gamma(\frac{\beta(r)-\eta+2}{2})\; \pi^{\frac{1}{2}}\; (1 + Z_r^T V_{z(r-1)}^{-1} Z_r)^{\frac{1}{2}}\; \lambda_{(r-1)}^{\frac{1}{2}}}$$

$$\left(1 + \frac{(X_r - \hat{\gamma}_{r-1} Z_r)^T \lambda_{(r-1)}^{-1} (X_r - \hat{\gamma}_{r-1} Z_r)}{1 + Z_r^T V_{z(r-1)}^{-1} Z_r}\right)^{-\frac{\beta(r)-\eta+3}{2}} \;, \tag{15}$$

with $\beta(r) - \eta + 2$ degrees of freedom, where the following notation is used:

$$\begin{aligned}
\beta(r) &= \beta(0) + r - 1 = \beta(r-1) + 1 \ , && (16)\\
\beta(0) &> 1 \ , \\
\hat{\gamma}_{r-1}^T &= V_{z(r-1)}^{-1} V_{zx(r-1)} && (17)\\
V_{r-1} &= \tilde{V}_{r-1} + I \ , \\
\tilde{V}_{r-1} &= \begin{pmatrix} \tilde{V}_{x(r-1)} & \tilde{V}_{zx(r-1)}^T \\ \tilde{V}_{zx(r-1)} & \tilde{V}_{z(r-1)} \end{pmatrix} \ , && (18)\\
\tilde{V}_{x(r-1)} &= \sum_{k=1}^{r-1} X_k^2 \ , && (19)\\
\tilde{V}_{zx(r-1)} &= \sum_{k=1}^{r-1} Z_k X_k \ , && (20)\\
\tilde{V}_{z(r-1)} &= \sum_{k=1}^{r-1} Z_k Z_k^T \ , && (21)\\
\lambda_{(r)} &= V_{x(r)} - V_{zx(r)}^T V_{z(r)}^{-1} V_{zx(r)} \ . && (22)
\end{aligned}$$

If $\beta(r-1) > \eta$ then the conditional mean value is

$$E\{X_r | X^{(r-1)}\} = \hat{\gamma}_{r-1} Z_r \qquad (23)$$

and it can be efficiently computed using the following recursion

$$\begin{aligned}
\hat{\gamma}_r^T &= \hat{\gamma}_{r-1}^T + (1 + Z_r^T V_{z(r-1)}^{-1} Z_r)^{-1} V_{z(r-1)}^{-1} Z_r \\
&\quad (X_r - \hat{\gamma}_{r-1} Z_r)^T \ . \qquad (24)
\end{aligned}$$

3 Optimal Contextual Support

The selection of an appropriate model support (I_r^c) is important to obtain good restoration results. If the contextual neighbourhood is too small it can not capture all details of the random field. Inclusion of the unnecessary neighbours on the other hand add to the computational burden and can potentially degrade the performance of the model as an additional source of noise. The optimal Bayesian decision rule for minimizing the average probability of decision error chooses the maximum posterior probability model, i.e., a model M_i corresponding to

$$\max_j \{p(M_j | X^{(r-1)})\} \ .$$

If we assume uniform prior for all tested support sets (models) the solution can be found analytically. The most probable model given past data is the model M_i $(I_{r,i}^c)$ for which

$$i = \arg\max_j \{D_j\}$$

$$\begin{aligned}
D_j &= \ln\Gamma\left(\frac{\beta(r) - \eta + 2}{2}\right) - \ln\Gamma\left(\frac{\beta(0) - \eta + 2}{2}\right) \\
&- \frac{1}{2}\ln|V_{z(r-1)}| - \frac{\beta(r) - \eta + 2}{2}\ln|\lambda_{(r-1)}| \qquad (25)
\end{aligned}$$

4 Global Estimation of the Point-Spread Function

Similarly with (12) the degradation model (1) can be expressed in the matrix form

$$X_r = \psi W_r + \epsilon_r \ , \qquad (26)$$

where

$$\begin{aligned}
\psi &= [h_1, \dots, h_\nu] \ , && (27)\\
\nu &= card(I_r)
\end{aligned}$$

and W_r is a corresponding vector of Y_{r-s}. The unobservable $\nu \times 1$ image data vector W_r is approximated using (3), (9),(23), i.e.,

$$\hat{W}_r = [\hat{\gamma}_{r-s-1} Z_r]_{s \in I_r}^T \ . \qquad (28)$$

In contrast to the model (11) the degradation model (1) is non-causal and hence it has no simple analytical Bayesian parameter estimate. Instead we use the least square estimate

$$\hat{\psi} = \min_\psi \left\{ \sum_{\forall r \in I} (X_r - \psi_r \hat{W}_r)^2 \right\} \ . \qquad (29)$$

The optimal estimate is

$$\hat{\psi}^T = V_{\hat{W}}^{-1} V_{\hat{W}X} \qquad (30)$$

where the data gathering matrices $V_{\hat{W}}, V_{\hat{W}X}$ are corresponding analogies with the matrices (19),(20).

5 Local Estimation of the Point-Spread Function

If we assume a non-homogeneous slowly changing point-spread function, we can estimate its local value using the local least square estimate

$$\hat{\psi}_r = \min_{\psi_r} \left\{ \sum_{\forall r \in J_r} (X_r - \psi_r \hat{W}_r)^2 \right\} \ . \qquad (31)$$

The locally optimal estimate is

$$\hat{\psi}_r^T = \tilde{V}_{\hat{W}}^{-1} \tilde{V}_{\hat{W}X} \ . \qquad (32)$$

The matrices $\tilde{V}_{\hat{W}}, \tilde{V}_{\hat{W}X}$ are computed from subwindows $J_r \subset I$. This estimator can be efficiently evaluated using the fast recursive square-root filter introduced in [8].

6 Results

The test image of the Cymbidium flower (Fig.1) was corrupted by the white Gaussian noise with $\sigma^2 = 225$ Fig.2. The signal-to-noise ratio for this corrupted image is

$$SNR = 10 \log \left(\frac{var(X)}{\sigma^2} \right) = 13.2 \ dB \ . \qquad (33)$$

The resulting reconstructed image using our method is on the Fig.4 while the image Fig.3 shows reconstruction using identical model but without differentiating discontinuity pixels. Visual comparison of both reconstructructed images demonstrates deblurring effect of the presented algorithm. The second experiment attempted was to reconstruct a uniform motion blur degradation in the vertical direction as can be seen on the Fig.5. Fig.6 presents its reconstrusted version. Finally, the top left image on the Fig.7 shows a range image from the Perceptron laser sensor, the top right image is its reconstructed version while the corresponding images in the lower row of this figure demonstrate an improvement of an edge detector performance due to noise suppression.

Figure 1: Original Cymbidium image.

Figure 2: Corrupted Cymbidium image.

The performance of the both methods is compared on artificially degraded images (so that the unobservable data are known) using the criterion of mean absolute difference between undegraded and restored pixel values

$$MAD = \frac{1}{MN} \sum_{r=1}^{MN} |Y_r - \hat{Y}_r| \tag{34}$$

and the criterion ζ which denotes the improvement in signal-to-noise ratio

$$\zeta = 10 \log \left(\frac{\mu(X)}{\mu(\hat{Y})} \right) \qquad dB \tag{35}$$

Figure 3: The reconstructed Cymbidium image (all pixels reconstructed).

Figure 4: The reconstructed Cymbidium image (method (4)).

where $\mu(X)$ is the mean-square error of X.

Figure 5: A uniform motion blur degraded Cymbidium image.

The Tab. 1 contains mono-spectral restoration results for the Cymbidium image Fig.2 and Fig.5. Both proposed methods are superior over the classical methods using both criteria (33),(34). The edge preserving version of the restoration method demonstrates visible deblurring effect Fig.4 without significantly affecting numerical complexity of the method.

Figure 6: A uniform motion blur restored Cymbidium image.

method	Cymbidium image			
	Gaussian noise		motion blur	
	MAD	ζ	MAD	ζ
AR model (4)	7.6	2.87	8.15	1.8
AR model (3)	7.6	2.35	8.11	1.6
median filter	16.2	-1.79	9.6	0.3
averaging	16.46	-1.86	9.7	0.35
Gaussian averaging	17.24	-2.49	9.8	0.28

Table 1: Restoration results.

7 Conclusions

The proposed recursive blur minimizing reconstruction method is very fast robust and its reconstruction results surpasses some standard reconstruction methods. Causal models such as (11) have obvious advantage to have the analytical solution for parameter estimation, prediction, or model identification tasks. However, this type of models may introduce some artifacts in restored images. These undesirable effects are diminished by introducing adaptivity into the model. This novel formulation allow us to obtain extremely fast adaptive restoration and / or local or global point-spread function estimation which can be easily parallelized. The method can be also easily and naturally generalized for multispectral (e.g. colour, multispectral satellite images) or registered images which is seldom the case for alternative methods. Finally, this method enables to estimate homogeneous or slowly changing non-homogeneous degradation point-spread function.

Acknowledgements

This research was supported by the GAČR grant no. 102/00/0030 Texture Modelling, the Hewlett Packard project no. ISE-86L8-13 Virtual Reality and partially supported by the INCO Copernicus project no. 960174 VIRTUOUS and the GAČR grant no. 106/00/1715 Computer Aided Quantitative Fractography of Fatigue Failures.

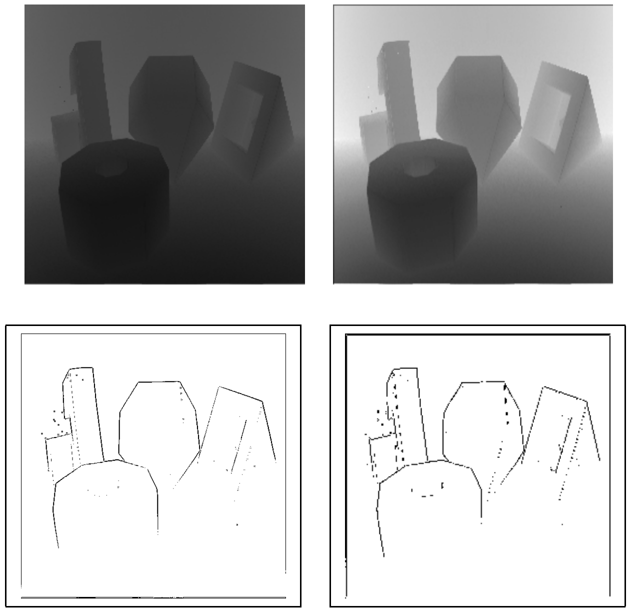

Figure 7: Laser range image and its enhanced version with corresponding edge maps.

References

[1] H. C. Andrews and B. Hunt. *Digital Image Restoration*. Prentice-Hall, Englewood Cliffs, 1977.

[2] B. Chalmond. Image restoration using an estimated markov model. *Signal Processing*, 15(2):115–129, February 1988.

[3] R. Chellappa and R. Kashyap. Digital image restoration using spatial interaction models. *IEEE Trans. Acoustics, Speeech and Sig. Proc.*, 30(3):284–295, June 1982.

[4] K. Deguchi and I. Morishita. Two-dimensional autoregressive model for the representation of random image fields. In *Proc. ICPR Conf.*, pages 90–93, Munich, 1982. IEEE.

[5] B. Fischl and E. Schwartz. Learning an integral equation approximation to nonlinear anisotropic diffusion in image processing. *IEEE Trans. Pattern Anal. Mach. Int.*, 19(4):342–352, April 1997.

[6] D. Geman. *Random fields and inverse problems in imaging*. Springer, Berlin, 1990.

[7] S. Geman and D. Geman. Stochastic relaxation , gibbs distributions and bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Int.*, 6(11):721–741, November 1984.

[8] M. Haindl. Recursive square-root filters. In *submitted to Proc. ICPR Conf.*, Barcelona, 2000. IEEE.

[9] M. Haindl and S. Šimberová. A high - resolution radiospectrograph image reconstruction method. *Astronomy and Astrophysics, Suppl.Ser.*, 115(1):189–193, January 1996.

[10] M. Haindl and S. Šimberová. A scratch removal method. *Kybernetika*, 34(4):423–428, April 1998.

[11] B. Hunt. The application of constraint least square estimation to image restoration by digital computer. *IEEE Trans. Computers*, 22(5):805–812, May 1973.

[12] B. Jeffs and W. Pun. Simple shape parameter estimation from blurred observations for a generalized gaussian mrf image prior used in map restoration. In *Proc. IEEE CVPR Conf.*, pages 465–468, San Francisco, 1996. IEEE.

[13] R. Lagendijk, J. Biemond, and D. Boekee. Identification and restoration of noisy blurred images using the expectation-maximization algorithm. *IEEE Trans. on Acoust., Speech, Signal Processing*, 38(y):1180–1191, xxx 1990.

[14] J. Marroquin and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *J. Am. Stat. Assoc.*, 82(397)(y):76–89, xxx 1987.

[15] M. Nitzberg and T. Shiota. Nonlinear image filtering with edge and corner enhancement. *IEEE Trans. Pattern Anal. Mach. Int.*, 16(8):826–833, August 1992.

[16] P. Perona. Deformable kernels for early vision. *IEEE Trans. Pattern Anal. Mach. Int.*, 17(5):488–489, May 1995.

[17] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Int.*, 12(7):629–639, July 1990.

[18] S. Reeves and R. Mersereau. Identification by the method of generalized cross-validation. *IEEE Trans. Im. Processing*, 1(y):301–311, xx 1992.