Model Adaptive Image Restoration

Wai Ho Pun and Brian D. Jeffs
Department of Electrical and Computer Engineering
Brigham Young University
459 Clyde Building, Provo Utah 84602
phone: (801) 378-3062, email: bjeffs@ee.byu.edu

Abstract

A new model adaptive method is proposed for restoration of blurred and noise corrupted images. This approach exploits information available from observed data to choose the appropriate optimization criterion and produce an approximate maximum likelihood solution. The generalized p-Gaussian family of probability distributions is used to model a wide range of observed noise classes. Distribution shape parameters are estimated from the image, and the resulting maximum likelihood optimization problem is solved. A fast iterative algorithm for this method is presented and analyzed. Experimental results indicate that this method outperforms the least squares method by taking advantage of the non-Gaussian characteristics of the noise data.

1 Introduction

Image restoration is the process of attempting to reconstruct a degraded image using some prior knowledge of the degradation phenomenon. Over the years, algorithms such as constrained least squares [1], maximum entropy [2], minimum norm [3, 4], etc. have been developed to solve this problem. These methods usually involve formulating a criterion as a measure of goodness that will yield some optimal estimate of the original image. Oftentimes, the criterion of goodness is chosen because of its computational simplicity instead of its optimal performance.

The choice of a particular deterministic algorithm is often equivalent to applying a specific prior model on the noise process or image itself, leading to expression of the restoration in terms of an optimization problem with an optimality criterion that dictates the form of the algorithm. For example, the least squares solution is optimal in the maximum likelihood sense when the image is deterministic, and the additive noise is independent and identically distributed (i.i.d.) Gaussian at each pixel.

In this paper, we propose a model adaptive method that uses information available from the observed image data to select the best restoration method based on the maximum likelihood principle. Rather than making an explicit or implicit assumption of Gaussian noise, this method exploits information in the image itself to achieve better results in the presence of additive non-Gaussian noise. When the additive noise is Gaussian, the performance of the model adaptive method is comparable with that of a least squares algorithm.

The basic idea behind the model adaptive method is that when precise knowledge of the noise process is absent, a distribution family is used to model the noise process. This distribution family must be chosen such that it includes (at least approximately) any probability distribution likely to be encountered in the image noise field. Selection of a particular noise model from the family is accomplished by adjusting a relatively small set of shape parameters, which are estimated from the observed data, thus adapting the noise model to the actual image. From this statistical model, a maximum likelihood estimater is applied to recover the original image.

2 Theoretical Development

2.1 Generalized p-Gaussian Distribution

Success of the model adaptive method depends on the accuracy of the noise model chosen. The distribution family used in modeling the noise process should be general enough to include a wide range of important distributions. In addition, the model fitting process or estimation of model parameters should also be

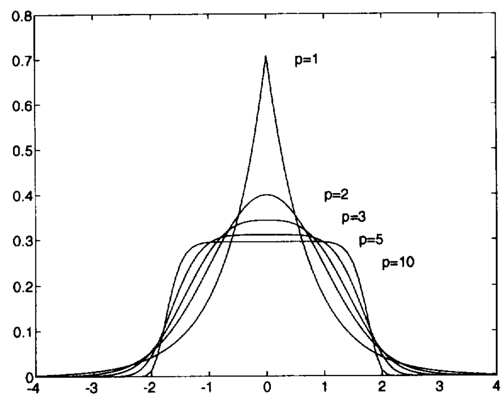


Figure 1: Generalized Gaussian density function with zero mean , unit variance and various values of p.

relatively easy and use only available information such as the observed image.

One such distribution family is the generalized p-Gaussian (gpG) distribution, also known as Box-Tiao [5] or power exponential distribution, which includes a wide range of shapes and can be used to represent distributions such as Gaussian, double exponential, uniform, Laplacian, Cauchy, and others [6].

The probability density function of the generalized Gaussian distribution family is defined as

$$f(x;p,\beta) = \frac{p}{2\beta\Gamma(\frac{1}{p})} \exp\left\{-\left(\frac{|x|}{\beta}\right)^p\right\}, \qquad (1)$$

where $\Gamma(\cdot)$ is the standard gamma function. This is a two-sided symmetric density with two distributional parameters, p and β, which control the shape and standard deviation of the density respectively. As shown in Figure 1, this family is very flexible. For example, with $p = 2, \beta = \sqrt{2}$, it becomes a standard normal distribution. For $p = 1$, we have a double exponential, and for $0 < p < 1$, we have heavy tailed distributions, while as $p \to \infty$, the uniform distribution is approximated. Since this distribution family has only two parameters, their estimation can be carried out without difficulties, as shown in Section 3.1.

2.2 Maximum Likelihood Solution

The image degradation process being considered in this paper is given by the equation $\mathbf{y} = \mathbf{Hx} - \mathbf{n}$. Assuming the image data, $\mathbf{x}$, are deterministic and the noise $\mathbf{n}$ comes from an i.i.d. generalized p-Gaussian process, the maximum likelihood solution is given by

$$\hat{\mathbf{x}}_{ML} = Arg\{\max_{\mathbf{x}} f_{\mathbf{y}}(\mathbf{y}|\mathbf{x})\}, \qquad (2)$$

where

$$f_{\mathbf{y}}(\mathbf{y}|\mathbf{x};p) = \left[\frac{p}{2\beta\Gamma(\frac{1}{p})}\right]^N \exp\left\{-\sum_{i=1}^{N}\left[\frac{|n_i|}{\beta}\right]^p\right\}. \qquad (3)$$

The noise vector element n_i can be replaced by $y_i - h_i^T\mathbf{x}$ where $\mathbf{H} = [h_1, h_2, \cdots, h_N]$. The log-likelihood function in terms of $\mathbf{x}$, β, and p is then

$$\mathcal{L}(\mathbf{x}, p, \beta) = N \ln\left[\frac{p}{2\beta\Gamma(\frac{1}{p})}\right] - \frac{1}{\beta^p}\sum_{i=1}^{N}|y_i - h_i^T\mathbf{x}|^p. \qquad (4)$$

If both p and β are known, then the maximum likelihood estimate of the original image can be found as

$$\hat{\mathbf{x}}_{ML} = Arg\{\max_{\mathbf{x}} \mathcal{L}(p, \beta, \mathbf{x})\}. \qquad (5)$$

Since the only unknown in equation (4) is $\mathbf{x}$, the estimate can be found as [7]

$$\hat{\mathbf{x}}_{ML} = Arg\{\min_{\mathbf{x}} \sum_{i=1}^{N}|y_i - h_i^T\mathbf{x}|^p\}. \qquad (6)$$

Notice that the term $\sum_{i=1}^{N}|y_i - h_i^T\mathbf{x}|^p$ is monotonically related to the l_p vector norm $\|\mathbf{y} - \mathbf{Hx}\|_p$. Thus l_p norm minimization yields the maximum likelihood solution to the image restoration problem under the assumption that the observation error can be modeled as a generalized p-Gaussian distribution for some value of p [7]. Note that if $p = 2$, which corresponds to a Gaussian noise case, the maximum likelihood solution becomes the least squares solution.

3 Algorithm Implementation

3.1 Parameter Estimation

Based on the assumption that the noise data can be modeled by the generalized p-Gaussian distribution, we have shown that the maximum likelihood estimate of the image depends on the shape parameter p, and the standard deviation parameter β. In practice, β can be absorbed into the restoration process as a scale factor and its estimation is not required. On the other hand, knowing the value of p is crucial to the model adaptive approach and the result of the restoration relies on using the correct value of p.

The maximum likelihood solution derived in the previous section assumed p was a deterministic known parameter. For a fully model adaptive restoration, the ideal approach would be a joint maximum likelihood estimation of $\mathbf{x}$, p, and β

$$\hat{\mathbf{x}}_{ML} = Arg\{\max_{\mathbf{x},p,\beta} \mathcal{L}(\mathbf{x},p,\beta)\}. \tag{7}$$

Although this has been done successfully with overdetermined problems with only a few parameters [6], equation (7) poses a most difficult nonlinear optimization problem in the high dimensionality environment of image restoration. A more practical approach, used in our work, is to form a prior estimate of p from image data, then compute the maximum likelihood restoration of $\mathbf{x}$ given that p.

To obtain an estimate of p, we first extract from the observed image a set of samples, $\hat{\mathbf{n}}$, representing noise data only. A heuristic approach for noise extraction can be found in [8]. Using the extracted noise vector, a density estimate of $\hat{\mathbf{n}}$, $f_{\hat{\mathbf{n}}}$, is constructed by the histogram method [9]. Next, the sum of the squared distance between $f_{\hat{\mathbf{n}}}$ and the density of the generalized p-Gaussian distribution is computed for different values of p. The value of p corresponding to the minimum distance is then taken as the estimate of p. The minimum distance estimater yields a robust estimate of p if $f_{\mathbf{n}}$ is a member of the generalized p-Gaussian distribution family [9].

3.2 Iterative Solution

Although the solution given by equation (6) is straight forward, closed form solutions are not known, and solving the problem using general purpose nonlinear optimization algorithms requires a tremendous computational load and in many cases impractical. The alternative solution presented here is a new iterative algorithm based on steepest descent, which offers the advantage of fast processing speed. In addition, by terminating the iterations prior to convergence, the restoration process can be regularized so that the trade-off between deblurring the image and noise magnification can be managed [10].

Using the steepest descent approach, the objective function, $\Phi(\mathbf{x}) = \sum_{i=1}^{N} |y_i - h_i^T\hat{\mathbf{x}}_k|^p$, as given in equation (6) can be minimized by the following iterative procedure:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \hat{\mathbf{x}}_k - \frac{\alpha}{p}\nabla_{\mathbf{x}}\Phi(\mathbf{x})|_{\hat{\mathbf{x}}_k} \\ &= \hat{\mathbf{x}}_k + \alpha\sum_{i=1}^{N}(y_i - h_i^T\hat{\mathbf{x}}_k)|y_i - h_i^T\hat{\mathbf{x}}_k|^{p-2}h_i \\ &= \hat{\mathbf{x}}_k + \alpha U_k\mathbf{H}^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_k),\end{aligned} \tag{8}$$

where α is a constant that regulates the step size and U_k is a diagonal matrix of the same size as $\mathbf{H}$ and its diagonal elements are $u_{ii} = |y_i - h_i^T\hat{\mathbf{x}}_k|^{p-2}$. This iteration assumes the form of the generalized Landweber iteration with the noise covariance matrix equal I and U_k as the preconditioning matrix [11].

3.3 Convergence Analysis

Two major concerns of using an iterative algorithm, such as equation (8), are 1) The convergence criteria, and 2) The limiting solution if the algorithm converges. To investigate these problems, let v_i and λ_i denote the normalized eigenvectors and eigenvalues associated with the system matrix $\mathbf{H}$ and assume that both $\mathbf{H}$ and $\mathbf{H}^T$ have the same set of eigenvectors. Using the recursive relation of equation (8), the estimate $\hat{\mathbf{x}}_{k+1}$ can be represented as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \sum_i(\hat{\mathbf{x}}_{k+1}, v_i)v_i \\ &= \sum_i(\hat{\mathbf{x}}_k + \alpha U_k\mathbf{H}^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_k), v_i)v_i \\ &= \sum_i\{(I - \alpha U_k\mathbf{H}^T\mathbf{H})(\hat{\mathbf{x}}_k, v_i) \quad + \\ &\quad \alpha U_k\mathbf{H}^T(\mathbf{y}, v_i)\}\, v_i,\end{aligned} \tag{9}$$

where (a, b) denotes the inner product between a and b. Using the eigenvalue and eigenvector relation and by careful examination of the recursion in equation (9), this iteration can be written as:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \sum_i\left\{(I - \alpha U_k|\lambda_i|^2)(\hat{\mathbf{x}}_k, v_i) \quad + \right. \\ &\quad \left.\alpha U_k\lambda_i(\mathbf{y}, v_i)\right\} v_i, \\ &= \sum_i\left\{\frac{1}{\lambda_i}\left[I - \sum_{j=0}^{k+1}(-1)^j\alpha^j\begin{bmatrix}S_k \\ j\end{bmatrix}|\lambda_i|^{2j}\right]\right\} \\ &\quad \cdot(\mathbf{y}, v_i)v_i,\end{aligned} \tag{10}$$

where S_k is the set $\{U_0, U_1, \cdots, U_k\}$ with $U_i = diag\,[u_{i1}, u_{i2}, \cdots, u_{in}]$, $u_{in} = |y_n - h_n^T\hat{\mathbf{x}}_i|^{p-2}$, and $\left[\begin{smallmatrix}S_k \\ j\end{smallmatrix}\right]$ is the general combination operator defined as the sum of the product of j unique elements out of the set S_k. One can easily show that the term $\sum_{j=0}^{k+1}(-1)^j\alpha^j\left[\begin{smallmatrix}S_k \\ j\end{smallmatrix}\right]|\lambda_i|^{2j}$ in equation (10) converges to $\mathbf{0} \in \mathbb{R}^{n\times n}$ as $k \to \infty$ if the following condition is satisfied:

$$0 < \alpha < \frac{2}{u_{\max}|\lambda_{\max}|^2}, \tag{11}$$

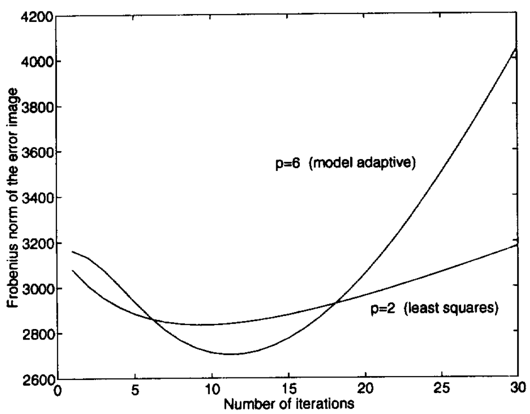


Figure 2: Frobenius or l_2 norm of the error image for least squares method and model adaptive method as a function of the iteration number.

where $u_{\max}$ is the maximum value among all the diagonal elements of the matrix $U_k, \forall k$, and $\lambda_{\max}$ is the maximum eigenvalue among all the eigenvalues of $\mathbf{H}$. Thus, according to equation (10), the limiting solution is the inverse filtered solution. Note that if $p = 2$ (corresponding to a Gaussian noise case) in equation (8), then the set $S_k = \{I, I, \cdots, I\}$, and $u_{\max} = 1$. The convergence condition reduces to $0 < \alpha < 2/|\lambda_{\max}|^2$ as given by Biemond et al for the least squares method [10].

4 Results

The performance of the model adaptive restoration method is illustrated by an example on an artificially blurred image in this section. The result is compared with the one obtained using the least squares restoration method.

This example considers the restoration of a blurred image with additive generalized p-Gaussian noise. The original image was artificially blurred by diagonal linear motion blur over 8 pixels. The discrete point spread function has a 4×8 region of support and the respective coefficients are:

$$d[m,n] = \frac{1}{14}\begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \end{bmatrix}.$$

The noise was generated by a generalized Gaussian process with shape parameter $p = 5$ and then added to the blurred image, resulting with a $15dB$ signal-to-noise ratio. Before the restoration was carried out, the shape parameter was estimated by the procedure outlined in section 3.1. The estimated value for p was 6. With this estimate, the iterative procedure outline in section 3.2 was applied to restore the image. Figure 2 shows how the performances of the model adaptive method and least squares method vary as the number of iterations varies. The model adaptive method achieves the best solution with the Frobenius or l_2 squared norm of the error image at 2703 after 11 iterations whereas the least squares method achieves its best solution with the Frobenius norm at 2834 after 9 iterations. Figure 3 shows the restored images and their corresponding error images using the model adaptive method and the least squares method respectively. Note that the error image from the adaptive method has a smoother appearance without edge degradation when compared to that from the least squares method, showing that the model adaptive method is a better alternative algorithm than the least squares method.

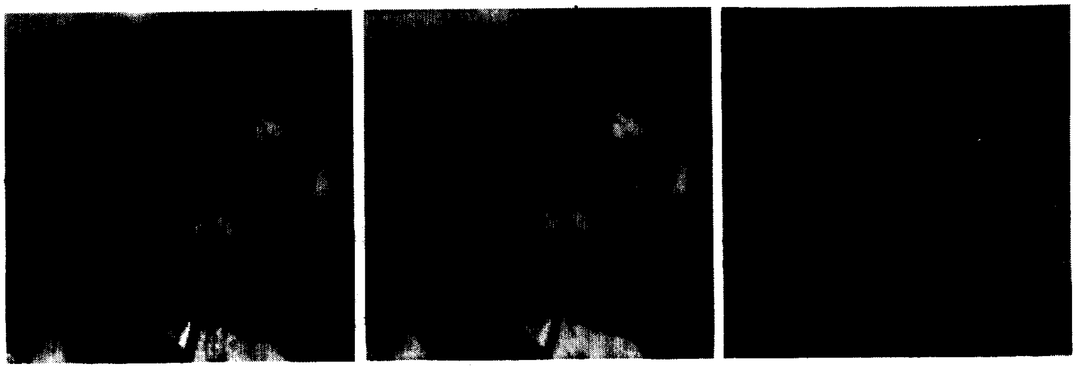
5 Conclusions

We have shown the development and implementation of the model adaptive image restoration algorithm. An iterative procedure is also derived to simplify the computation process. In addition, the convergence criterion is also established. Results show that this method outperforms the least squares method when the implicit assumptions taken by the later fail to match the true nature of the observed data. The major advantage of the model adaptive approach lies in its ability to adapt itself to the observed data and make use of the information available from the data.

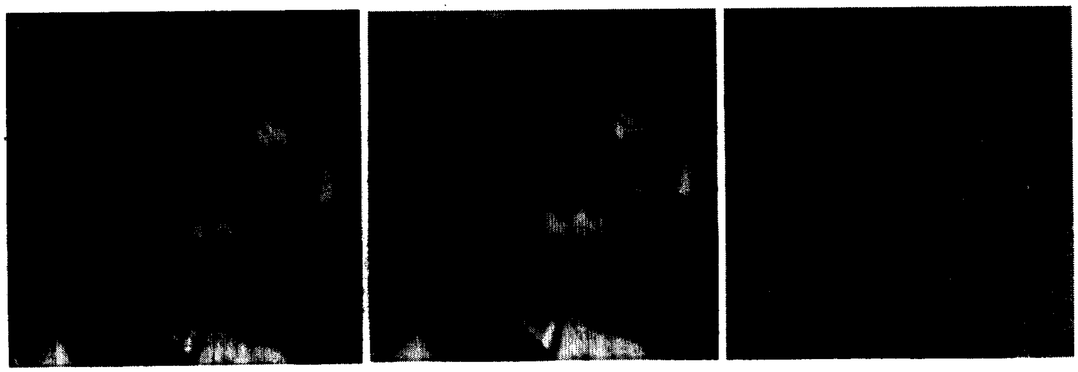
References

[1] B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," *IEEE Transactions on Computers*, vol. C-22, Sept. 1973.

[2] S. F. Gull and J. Skilling, "Maximum entropy method in image processing," *IEEE Proceedings*, vol. 131, pt. F, no. 6, pp. 646–659, 1984.



(a) Original 256 by 256 image (c) Model adaptive restoration (e) Error image of c



(b) Blurred and noisy image (d) Least squares restoration (f) Error image of d

Figure 3: Results of model adaptive and least squares restoration.

[3] Y. Censor, "Finite series-expansion reconstruction methods," *Proceeding of the IEEE*, vol. 71, pp. 409–418, 1983.

[4] A. K. Jain, *Fundamentals of Digital Image Processing.* Englewood Cliffs: Prentice Hall, 1989.

[5] G. E. P. Box and G. C. Tiao, "A further look at robustness via Baye's theorem," *Biometrica*, vol. 49, pp. 419–432, 1962.

[6] J. B. McDonald, "Partially adaptive estimation of ARMA time series models," *International Journal of Forecasting*, vol. 5, pp. 217-230, 1989.

[7] T. T. Pham and R. J. P. deFigueiredo, "Maximum likelihood estimation of a class of non-Gaussian densities with application to l_p deconvolution," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 73–82, Jan. 1989.

[8] B. D. Jeffs and W. H. Pun, "Model adaptive optimal image restoration," in *Proc. SPIE 1771* (A. G. Tescher, ed.), pp. 307–321, 1992.

[9] L. Devroye, *A Course in Density Estimation.* Boston: Birkhäuser, 1987. Edited by Murray Rosenblatt.

[10] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proceeding of the IEEE*, vol. 78, pp. 856–883, 1990.

[11] O. N. Strand, "Theory and methods related to the singular-function expansion and landweber's iteration for integral equations of the first kind," *SIAM J. Numer. Anal.*, vol. 11, no. 4, pp. 798–825, 1974.