# ITERATIVE ADAPTIVE $l_p$ RESTORATION OF BLURRED IMAGES

*Wai Ho Pun and Brian D. Jeffs*
*Department of Electrical and Computer Engineering*
*Brigham Young University*
*459 Clyde Building, Provo Utah 84602*

## ABSTRACT

A new model adaptive method is presented for restoration of blurred and noise corrupted images by exploiting information available from observed data to choose the appropriate optimization criterion. The derived maximum likelihood solution is based on the $l_p$ minimization criterion that naturally arises from the adoption of the generalized p-Gaussian family of probability distributions as an additive noise model. A fast and efficient iterative algorithm for this adaptive method is developed and analyzed. Experimental results indicate that this method adapts to the non-Gaussian nature of the noise process and outperforms the least squares method, which lacks the flexibility of the former method.

## 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. The choice of a particular deterministic algorithm is often equivalent to applying a specific prior model on the noise process or image itself, expressing the restoration as an optimization problem with an optimality criterion that dictates the form of the algorithm. For example, the least squares solution is optimum 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 the noise process, this method adapts itself to the observed data to achieve optimum results.

The basic idea behind the model adaptive method is that when precise knowledge of the noise process is absent, a parameter distribution family is used. This 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. From this statistical model, a maximum likelihood estimater is applied to recover the original image.

## 2. ADAPTIVE $l_p$ RESTORATION

### 2.1. Generalized p-Gaussian Distribution

The generalized p-Gaussian (gpG) distribution family, also known as Box-Tiao or power exponential distribution [5], is used to model the unknown noise process because it has a wide range of shapes. The probability density function of the generalized Gaussian distribution family is defined as

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

where $\Gamma(\cdot)$ is the standard gamma function. This is a two-sided symmetric density with two distributional parameters, $p$ and $\beta$, which control the shape and standard deviation of the density respectively. As shown in Figure 1, this family is very flexible. For example, with $p = 2$, and $\beta = \sqrt{2}$, 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.

### 2.2. Maximum Likelihood Solution

Consider the image degradation model $\mathbf{y} = \mathbf{Hx} - \mathbf{n}$. Assuming the image data, $\mathbf{x}$, are deterministic and the noise $\mathbf{n}$ comes from an i.i.d. generalized p-Gaussian process, the fully adaptive maximum likelihood solution is given by

$$\hat{\mathbf{x}}_{ML} = Arg\{\max_{\mathbf{X},p,\beta} f_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})\}, \tag{2}$$

where

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

The noise vector element $n_i$ can be replaced by $y_i - h_i^T\mathbf{x}$ where $\mathbf{H} = [h_1, h_2, \cdots, h_N]$. Equivalently, the solution can be obtained by solving

$$\hat{\mathbf{x}}_{ML} = Arg\{\max_{\mathbf{X},p,\beta} \mathcal{L}(\mathbf{x},p,\beta)\}, \tag{4}$$

where $\mathcal{L}(\mathbf{x},p,\beta)$ is the log-likelihood function. However, equation (4) still poses a difficult nonlinear optimization problem in the high dimensionality environment of image restoration. A more practical approach, used in our work,



0-7803-1775-0/94 $3.00 

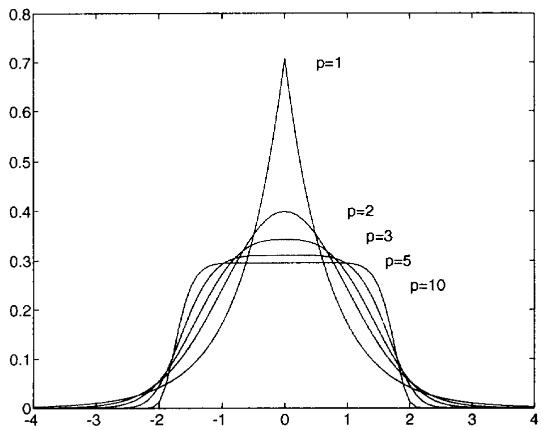



Figure 1. Generalized p-Gaussian density function with zero mean , unit variance and various values of $p$.

is to form a prior estimate of $p$ from image data, then compute the maximum likelihood restoration of $\mathbf{x}$ given that $p$. A heuristic approach for estimation of $p$ can be found in [6]. Since the parameter $\beta$ can be absorbed into the restoration process as a scaling factor, once the estimate of $p$ is obtained, equation (4) reduces to

$$\hat{\mathbf{x}}_{ML} = Arg\{\min_{\mathbf{x}} \sum_{i=1}^{N} |y_i - h_i^T \mathbf{x}|^p\}. \tag{5}$$

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 $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_p$. 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]. If $p = 2$, which corresponds to a Gaussian noise case, the maximum likelihood solution becomes the least squares solution.

## 3. ALGORITHM IMPLEMENTATION

### 3.1. Iterative Solution

Although the solution given by equation (5) is straight forward, closed form solutions are not known, and solving the problem using general purpose nonlinear optimization algorithms involves a tremendous computational load and impractical in many cases. 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 amplification can be managed [8].

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 (5) can be minimized by the following iterative procedure:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \hat{\mathbf{x}}_k - \frac{\alpha_k}{p} \nabla_{\mathbf{x}} \Phi(\mathbf{x})|_{\hat{\mathbf{x}}_k} \\ &= \hat{\mathbf{x}}_k + \alpha_k \sum_{i=1}^{N} (y_i - h_i^T \hat{\mathbf{x}}_k) |y_i - h_i^T \hat{\mathbf{x}}_k|^{p-2} h_i \\ &= \hat{\mathbf{x}}_k + \alpha_k U_k \mathbf{H}^T (\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_k), \end{aligned} \tag{6}$$

where $\alpha_k$ 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 [9].

### 3.2. Convergence Analysis

Two major concerns of using an iterative algorithm, such as equation (6), are 1) The convergence criteria, and 2) The limiting solution if the algorithm converges. To investigate these problems, let $v_i$ and $\lambda_i$ 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 (6), the estimate $\hat{\mathbf{x}}_{k+1}$ can be represented as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \sum_i \langle \hat{\mathbf{x}}_{k+1}, v_i \rangle v_i \\ &= \sum_i \langle \hat{\mathbf{x}}_k + \alpha_k U_k \mathbf{H}^T (\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_k), v_i \rangle v_i \\ &= \sum_i \left\{ (I - \alpha_k U_k \mathbf{H}^T \mathbf{H}) \langle \hat{\mathbf{x}}_k, v_i \rangle \right. + \\ &\quad \left. \alpha_k U_k \mathbf{H}^T \langle \mathbf{y}, v_i \rangle \right\} v_i, \end{aligned} \tag{7}$$

where $\langle a, b \rangle$ denotes the inner product between $a$ and $b$. Using the eigenvalue and eigenvector relation and by careful examination of the recursion in equation (7), this iteration can be written as:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \sum_i \left\{ (I - \alpha_k U_k |\lambda_i|^2) \langle \hat{\mathbf{x}}_k, v_i \rangle \right. + \\ &\quad \left. \alpha_k U_k \lambda_i \langle \mathbf{y}, v_i \rangle \right\} v_i, \\ &= \sum_i \left\{ \frac{1}{\lambda_i} \left[ I - \sum_{j=0}^{k+1} (-1)^j \alpha_k^j \begin{bmatrix} S_k \\ j \end{bmatrix} |\lambda_i|^{2j} \right] \right\} \\ &\quad \cdot \langle \mathbf{y}, v_i \rangle v_i, \end{aligned} \tag{8}$$

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 $\sum_{j=0}^{k+1} (-1)^j \alpha_k^j \left[\begin{smallmatrix} S_k \\ j \end{smallmatrix}\right] |\lambda_i|^{2j}$ in equation (8) converges to $\mathbf{0} \in \mathbb{R}^{n \times n}$ as $k \to \infty$ if the following condition is satisfied:

$$0 < \alpha_k < \frac{2}{u_{\max} |\lambda_{\max}|^2}, \tag{9}$$

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 (8), the limiting solution is the inverse filtered solution. Note that if $p = 2$ (corresponding to a Gaussian noise case) in equation (6), then the set $S_k = \{I, I, \cdots, I\}$, and $u_{max} = 1$. The convergence condition reduces to $0 < \alpha_k < 2/|\lambda_{max}|^2$ as given by Biemond et al for the least squares method [8]. In this case, the adaptive method turns into a least squares method.

### 3.3. Computational Efficiency

For images of typical size, the iteration of equation (6) becomes computationally impractical. Assuming $\mathbf{x}$ and $\mathbf{y}$ correspond to square images of size $M \times M$, each iteration of (6) involves $O\{2(M^4 + M^2)\}$ operations. Significant improvement in computational efficiency is possible if we recognize the 2-D convolution and correlation operations implicit in equation (6), which may then be rewritten as

$$\begin{aligned} \hat{\mathcal{X}}_{k+1} &= \hat{\mathcal{X}}_k + \alpha \mathcal{H} \otimes |\mathcal{Y} - \mathcal{H} * \hat{\mathcal{X}}_k|^{p-1} \odot \\ & \quad signum\{\mathcal{Y} - \mathcal{H} * \hat{\mathcal{X}}_k\} \end{aligned} \tag{10}$$

where $\hat{\mathcal{X}}, \mathcal{Y}$ and $\mathcal{H}$ are the 2-D image matrices corresponding to $\mathbf{x}, \mathbf{y}$ and $\mathbf{H}$, $\otimes$ and $*$ indicate 2-D deterministic correlation, and convolution respectively, and where $\odot$, $|\cdot|^{p-1}$, and $signum\{\ \}$ are element by element matrix operations of multiplication, absolute exponentiation, and sign retrieval respectively. We assume that $\mathcal{H}$ is $N \times N$, and that convolution and correlation operations are truncated to $M \times M$. Under these conditions, (10) involves $O\{2(M^2N^2 + M^2)\}$ operations. Since the point spread function usually has a small finite region of support, we note that $N \ll M$, so the computational load for equation (10) can be many orders of magnitude less than for equation (6). Results presented in Section 4 were computed using equation (10), and we have found this algorithm to be sufficiently efficient for restoration of 1024 by 1024 pixel images on a modest desktop workstation computer in a few minutes.

## 4. RESULTS

The performance of the model adaptive restoration method is illustrated by an example on an artificially blurred image in this section. This result is compared with 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 over 8 pixels. The discrete point spread function has a $4 \times 8$ 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 p-Gaussian process with shape parameter $p = 8$ and then added to the blurred image, resulting in a $30dB$ signal-to-noise ratio. Before the restoration was carried out, the shape parameter $p$ was estimated to be 3. The difference between this estimate and the actual value is due to the relatively high SNR value in this experiment; however, this value appropriately represents the observable character of the noise data. In experiments (not shown) with higher SNR ($15dB$) $p$ was correctly estimated at $p = 8$. With this estimate, the iterative procedure outlined in section 3 was applied to restore the image. Figure 2 shows how the performances of the adaptive method and least squares method vary as the number of iterations varies. The model adaptive method achieves the best solution with an average pixel error power of 170 after 27 iterations whereas the least squares method achieves its best solution with error power of 293 after 28 iterations. The adaptive method attains an SNR improvement of $5.44dB$ over that of the least squares method in this experiment. Figure 3 shows the restored images using the model adaptive method and the least squares method respectively. Notice that the restored image using least squares has objectionable ringing artifacts near edges where the intensity transitions are steep. On the other hand, the restored image using the adaptive method shows better quality, indicating that the model adaptive method is a better alternative algorithm than the least squares method.

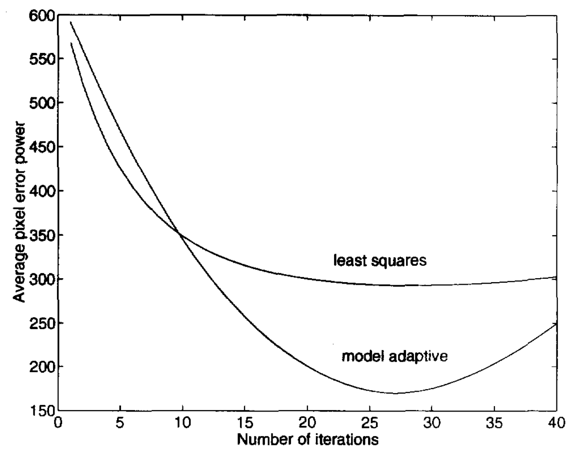
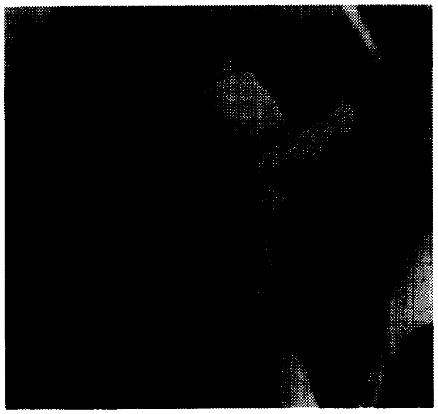


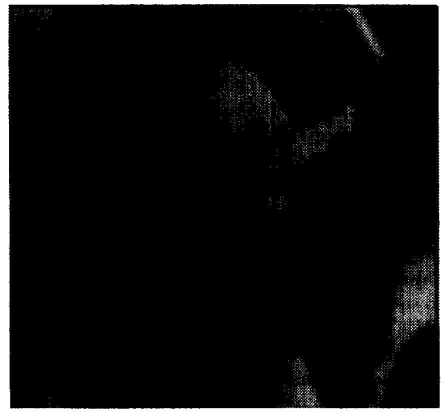

Figure 2. Average pixel power of the error image for least squares method and model adaptive method as a function of the iteration number.

## 5. CONCLUSIONS

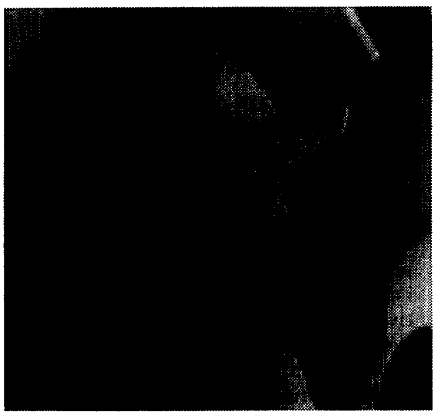
We have shown the development and implementation of the model adaptive image restoration algorithm based on the $l_p$ norm minimization criterion. An efficient iterative procedure is derived to simplify the computation process. In addition, the convergence criterion is also established. Results show 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 to yield an optimum solution.



(a) Original 256 by 240 image



(b) Blurred and noisy image



(c) Model adaptive restoration

(d) Least squares restoration

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

## 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.

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

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

[9] 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.