Blind Restoration/Superresolution with Generalized Cross-Validation Using Gauss-Type Quadrature Rules

Nhat Nguyen and Gene Golub
SCCM Program
Stanford University
Stanford, CA 94305-9025
{nguyen,golub}@sccm.stanford.edu

Peyman Milanfar
Department of Electrical Engineering
University of California
Santa Cruz, CA 95064
milanfar@cse.ucsc.edu

Abstract

In many restoration/superresolution applications, the blurring process, i.e., point spread function (PSF) of the imaging system, is not known. We estimate the PSF and regularization parameters for this ill-posed inverse problem from raw data using the generalized cross-validation method (GCV). To reduce the computational complexity of GCV, we propose efficient approximation techniques based on the Lanczos algorithm and Gauss quadrature theory. Data-driven blind restoration/superresolution experiments are presented to demonstrate the effectiveness and robustness of our method.

1 Introduction

Given a sequence of aliased low resolution frames $\mathbf{b}_i, i = 1, ..., k$, superresolution reconstructs an enhanced, high resolution image $\mathbf{x}$ by extracting subpixel information from the given frames. We model the low resolution frames as blurred and down-sampled, shifted version of the high resolution image we wish to estimate. To simplify the problem, we model the point spread function (PSF) as a parametric blur with parameter set σ. Then with given σ, the forward model for superresolution is

$$\mathbf{b} = A(\sigma)\mathbf{x} + \epsilon, \tag{1}$$

where $A(\sigma)$ is a very large, sparse, ill-conditioned and typically underdetermined matrix generated from the blur parameter set σ, and ϵ is an additive noise vector [10]. In [9], we showed that restoration is a special case of superresolution. Thus, (1) is applicable as a restoration model equation as well. The regularized minimum norm least squares solution to (1) can be expressed as follows:

$$\mathbf{x}(\sigma, \lambda) = A(\sigma)^T (A(\sigma)A(\sigma)^T + \lambda I)^{-1}\mathbf{b}, \tag{2}$$

where λ is the regularization parameter. In many applications, the camera's characteristics may be unknown. Our approach to the blind restoration/superresolution problem first estimates the unknown PSF parameter set σ and regularization parameter λ from raw data. Once estimates $\hat{\sigma}$, $\hat{\lambda}$ are available, a computationally inexpensive preconditioned conjugate gradient algorithm is used to solve the non-blind problem (2), see [10]. In Section 2, we describe our parametrized blur estimation technique using generalized cross-validation. In Section 3, we propose a method based on quadrature rules and the Lanczos algorithm, which bounds the GCV criterion value accurately and efficiently. Blur estimation and blind restoration/superresolution results are shown in Section 4.

2 Cross-Validation

Generalized cross-validation is a popular method for computing regularization parameters [4]. More recently, Reeves and Mersereau have used GCV for blur identification under an autoregressive moving average (ARMA) model [11]. In a recent study by Chardon, Vozel, and Chehdi [2], GCV has been shown to be an effective tool in parametric blur estimation. Motivated by these successes, we apply GCV to estimate both the PSF and regularization parameters for blind superresolution:

$$\left\{\hat{\sigma}, \hat{\lambda}\right\} = \text{argmin}_{\{\sigma,\lambda\}} \frac{\|(A(\sigma)A(\sigma)^T + \lambda I)^{-1}\mathbf{b}\|_2}{\text{tr}((A(\sigma)A(\sigma)^T + \lambda I)^{-1})}. \tag{3}$$

Reeves and Mersereau simplified the objective function in (3) by assuming the system matrix $A(\sigma)$ to be circulant. In this paper, we neither assume ARMA model or circulant structure for $A(\sigma)$. Reeves and Mersereau also simultaneously estimated the optimal blur and regularization parameters while keeping the image model parameters fixed. We found, however, that by setting the regularization parameter to some small number, the PSF parameters can be better

0-7803-5700-0/99/$10.00©1999 IEEE

estimated even in the presence of noise. We then use the computed PSF to determine the appropriate regularization parameter. Our intuition is that with under-regularization (small λ), the noise effect is exacerbated, moving the GCV criterion away from possible local minima. Furthermore, the estimated PSF is less biased away from the actual PSF even though the variance of the estimates is larger. We used $\lambda_0 = 10^{-3}$ in our experiments so that the estimated PSF parameters can be found by

$$\hat{\sigma} = \text{argmin}_\sigma \frac{\|(A(\sigma)A(\sigma)^T + \lambda_0 I)^{-1}\mathbf{b}\|_2}{\text{tr}((A(\sigma)A(\sigma)^T + \lambda_0 I)^{-1})}. \quad (4)$$

In the simplest case, the parameter set σ consists of one parameter describing the smoothness of the blur, e.g. the standard deviation of a Gaussian hump or the radius of a pillbox (out-of-focus) blur. Once a blur estimate $\hat{\sigma}$ is available, we compute the regularization parameter from

$$\hat{\lambda} = \text{argmin}_\lambda \frac{\|(A(\hat{\sigma})A(\hat{\sigma})^T + \lambda I)^{-1}\mathbf{b}\|_2}{\text{tr}((A(\hat{\sigma})A(\hat{\sigma})^T + \lambda I)^{-1})}. \quad (5)$$

3 Quadrature Rules with Lanczos Algorithm

For large systems, the numerators and denominators of (4) and (5) are very expensive to evaluate directly. We first approximate the denominators using an unbiased trace estimator by Hutchinson [8]: Let U be a discrete random variable which takes the values -1 and +1 each with probability $\frac{1}{2}$, and let $\mathbf{u}$ be a vector whose entries are independent samples from U. Then the term $\mathbf{u}^T(AA^T + \lambda I)^{-1}\mathbf{u}$ is an unbiased estimator of $tr((AA^T + \lambda I)^{-1})$.

Now, in order to estimate numerators and denominators in (4) and (5), we need to estimate quadratic forms $\mathbf{v}^T f(M)\mathbf{v}$, where M is some symmetric positive definite matrix and $f(\xi) = \xi^{-p}$, $p = 1, 2$. There is extensive literature on the application of Gauss quadrature rules to bound bilinear forms; see papers by Golub and collaborators [1, 3, 5, 6, 7]. We will briefly summarize the details in the following.

Let the eigendecomposition of M be given by $M = Q^T \Xi Q$, where Q is an orthogonal matrix and Ξ is a diagonal matrix of eigenvalues in increasing order. Then

$$\begin{aligned} \mathbf{v}^T f(M)\mathbf{v} &= \mathbf{v}^T Q^T f(\Xi) Q\mathbf{v} \\ &= \tilde{\mathbf{v}}^T f(\Xi)\tilde{\mathbf{v}} \\ &= \textstyle\sum_{i=1}^n f(\xi_i)\tilde{v}_i^2, \end{aligned}$$

where $\tilde{\mathbf{v}} = (\tilde{v}_i) \equiv Q\mathbf{v}$. Suppose that we have bounds on the spectrum of M, e.g. by Gershgorin circle theorem, $a \le \xi_1 \le ... \le \xi_n \le b$. The last sum can be considered as a Riemann-Stieltjes integral with piecewise constant measure

$$\sum_{i=1}^n f(\xi_i)\tilde{v}_i^2 = \int_a^b f(\xi)d\mu(\xi), \quad (6)$$

where $\mu(\xi)$ is defined as

$$\mu(\xi) = \begin{cases} 0, & \text{if } \xi < \xi_1 \\ \sum_{j=1}^i \tilde{v}_j^2, & \text{if } \xi_i \le \xi < \xi_{i+1} \\ \sum_{j=1}^n \tilde{v}_j^2, & \text{if } \xi_n \le \xi. \end{cases}$$

We can approximate the Riemann-Stieltjes integral (6) with Gauss-type quadrature rules. The general form for quadrature rules is

$$\begin{aligned} I[f] &= \sum_{i=1}^k \omega_i f(\theta_i) + \sum_{j=1}^l \nu_j f(\tau_j), \\ \int_a^b f(\xi)d\mu(\xi) &= I[f] + R[f], \end{aligned} \quad (7)$$

where the weights ω_i, ν_j and the nodes θ_i are unknown, the nodes τ_j are predetermined, and $R[f]$ is the remainder term. The Gauss-type quadrature rules differ from one another by the number of prescribed nodes. If there are no prescribed nodes, then we obtain the standard Gauss quadrature.

$$I_G[f] = \sum_{i=1}^k \omega_i f(\theta_i)$$

If one node is prescribed, we get the Gauss-Radau quadrature rule; with two nodes prescribed, the Gauss-Lobatto rule,

$$\begin{aligned} I_R[f] &= \sum_{i=1}^k \omega_i f(\theta_i) + \nu f(\tau) \\ I_L[f] &= \sum_{i=1}^k \omega_i f(\theta_i) + \nu_1 f(\tau_1) + \nu_2 f(\tau_2) \end{aligned}$$

Gauss-Radau rule is often applied with either $\tau = a$ or $\tau = b$. Gauss-Lobatto has both endpoints prescribed, $\tau_1 = a, \tau_2 = b$.

3.1 Quadrature Error

The quadrature error $R[f]$ from (7) can be expressed as

$$R[f] = \frac{f^{(2k+l)}(\eta)}{(2k+l)!} \int_a^b \prod_{j=1}^l (\xi - \tau_j) \prod_{i=1}^k (\xi - \theta_i)^2 d\mu(\xi),$$

for some $\eta \in (a, b)$. We have the following theorem for the Gauss quadrature rule [5].

Theorem 1 *Suppose that $f^{(2k)}(\xi) > 0$, $\forall k$, $\forall \xi$, $a < \xi < b$, then*

$$I_G[f] \le \int_a^b f(\xi)d\mu(\xi).$$

Analogous theorems for Gauss-Radau and Gauss-Lobatto rules are presented in the following.

Theorem 2 *Suppose that* $f^{(2k+1)}(\xi) < 0,\ \forall k,\ \forall \xi,\ a < \xi < b$, *then*

$$I_{Rb}[f] \leq \int_a^b f(\xi) d\mu(\xi) \leq I_{Ra}[f],$$

where I_{Ra} (I_{Rb}) *corresponds Gauss-Radau rule with the prescribed node at* $\tau = a$ (b).

Theorem 3 *Suppose that* $f^{(2k)}(\xi) > 0,\ \forall k,\ \forall \xi,\ a < \xi < b$, *then*

$$\int_a^b f(\xi) d\mu(\xi) \leq I_L[f],$$

with prescribed nodes $\tau_1 = a,\ \tau_2 = b$.

Recall that the bilinear terms which we are interested in approximating have form $\mathbf{v}^T M^{-p} \mathbf{v}$, $p = 1, 2$. The function $f(\xi) = \xi^{-p}$ satisfies the hypotheses of Theorems 1,2, and 3, for M positive definite $(a > 0)$. Hence, we can bound $\int_a^b f(\xi) d\mu(\xi)$ with Gauss-type quadrature rules. Define

$$\begin{aligned} L[f] &:= \max(I_G[f], I_{Rb}[f]), \\ U[f] &:= \min(I_{Ra}[f], I_L[f]), \end{aligned}$$

Then we have the following bounds [5]

$$L[f] \leq \int_a^b f(\xi) d\mu(\xi) \leq U[f].$$

To find quadrature bounds $L[f]$ and $U[f]$ above, we need the unknown weights ω_i, ν_j and nodes θ_i. These quantities can be computed from sequences of orthogonal polynomials associated with the weight measure $d\mu(\xi)$. We can define a sequence of orthonormal polynomials $\{p_i\}_{i=0}^{n-1}$ such that

$$\int_a^b p_i(\xi) p_j(\xi) d\mu(\xi) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

These polynomials satisfy a three-term recurrence relation

$$\xi \mathbf{p}_k(\xi) = T_k \mathbf{p}_k(\xi) + \beta_k p_k(\xi) \mathbf{e}_k, \quad k = 1, ..., n,$$

$$p_{-1}(\xi) = 0, \quad p_0(\xi) = 1.$$

with

$$\mathbf{e}_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_k = \begin{bmatrix} p_0(\xi) \\ \vdots \\ p_{k-1}(\xi) \end{bmatrix},$$

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \alpha_{k-1} & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \end{bmatrix}.$$

It turns out that the nodes θ_i of Gauss quadrature rule are the eigenvalues of T_k, which are also the zeros of the polynomial p_n. The weights ω_i are the square of the first component of the normalized eigenvectors of T_k. For Gauss-Radau rule, we need to adjust the last entry α_k of T_k so that the adjusted tridiagonal matrix $\hat{T}_k$ has an eigenvalue at the prescribed node. For the Gauss-Lobatto rule the last three nonzero entries $\beta_{k-1}, \alpha_k, \beta_{k-1}$ will be adjusted to prescribe eigenvalues at a and b. The orthogonal polynomials $\{p_j(\xi)\}$ and coefficients of their three-term recurrence can be estimated via the Lanczos bidiagonalization algorithm. The follow stopping criterion terminates the Lanczos algorithm

$$\frac{U[f] - L[f]}{U[f]} \leq 0.01. \qquad (8)$$

We use $U[f]$ as the approximate value for $\int_a^b f(\xi) d\mu(\xi) := \mathbf{v}^T f(M) \mathbf{v}$.

4 Experiments and Conclusions

We estimate blur parameters using the GCV criterion with quadrature rules bounds as described above. We use Matlab's CONSTR routine [12] to solve the GCV minimization problem (4). For each function evaluation, we iterate with the Lanczos algorithm until the stopping criterion (8) is satisfied, usually within 70 Lanczos iterations, equivalent to 140 matrix vector multiplies, for our test image sequence. The iteration count is quite low compared to the dimensions of the system matrix (usually in the tens of thousands). In the first set of experiments, 16 low resolution frames are generated by blurring a 172 × 172 pixels high resolution image with a 4 × 4 Gaussian blur and downsampling by a factor of 4 in each dimension. We experiment with blurs of standard deviations 0.75, 1.0, and 2.0. In addition, we consider blind superresolution with 60 dB, 30 dB and without additive Gaussian noise added to the low resolution frames. We simulate blind superresolution for two cases, with all frames given and with 10 randomly choosen of the 16 available frames. When all frames are available, superresolution is equivalent to a restoration problem. Tables 1 and 2 display the mean square error (MSE) (see [2] for precise definition) (in percent) in the PSF estimates. Figure 1 shows the result of blind superresolution using computed PSF and regularization parameters from (4) and (5). The actual blur standard deviation is 0.75. We added white noise to the low resolution frames to realize an SNR of 30 dB. The resulting high resolution estimate is computed from 10 randomly chosen low resolution frames. In the second set of experiments, we ran tests for a pillbox (defocussed) blur. The parameter to be estimated is the radius of the blur. Our experiments tested for out-of-focus blurs with radii 2 and 5. We plotted GCV values for radius taking values from 1 to 10 at 0.2 intervals.

Our experimental results demonstrate the effectiveness of our techniques for identifying unknown blurs in restoration/superresolution problems, even in the presence of noise. Gauss-type quadrature rules bound GCV function values for large systems efficiently. The proposed techniques are a foundation for completely data-driven efficient blind restoration/superresolution algorithms.

References

[1] Z. Bai, M. Fahey, and G. Golub. Some Large-Scale Matrix Computation Problems. *J. Comput. Appl. Math.*, Vol. 74, pp. 71-89, 1996.

[2] S. Chardon, B. Vozel, and K. Chehdi. A Comparative Study between Parametric Blur Estimation Methods. ICASSP '99, Vol. VI, pp. 3233-3236, Phoenix, AZ.

[3] G. Golub. Some Modified Matrix Eigenvalue Problem. *SIAM Review*, Vol. 15, pp. 318-334, 1973.

[4] G. Golub, M. Heath, and G. Wahba. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, Vol. 21, pp. 215-223, 1979.

[5] G. Golub and G. Meurant. Matrices, Moments, and Quadrature. In *Numerical Analysis 1993*, eds. D. F. Griffiths and G. A. Watson, Longman, Essex, England, pp. 105-156, 1994.

[6] G. Golub and Z. Strakos. Estimates in Quadratic Formulas. *Numer. Algor.*, Vol. 8, pp. 241-268, 1994.

[7] G. Golub and J. Welsch. Calculation of Gauss Quadrature Rules. *Math. Comp.*, Vol. 23, pp. 221-230, 1969.

[8] M. F. Hutchinson. A Stochastic Estimator of the Trace of the Influence Matrix for Laplacian Smoothing Splines. *Communications in Statistics, Simulation, and Computation*, 19, pp. 433-450, 1990.

[9] N. Nguyen. Numerical Techniques for Image Superresolution. PhD thesis, SCCM Program, Stanford University, expected Dec. 1999.

[10] N. Nguyen, P. Milanfar, and G. Golub. Preconditioners for Regularized Image Superresolution. ICASSP '99, Vol. VI, pp. 3249-3252, Phoenix, AZ.

[11] S. Reeves and R. Mersereau. Blur Identification by the Method of Generalized Cross-Validation. *IEEE Trans. Image Processing*, Vol. 1, No. 3, pp. 301-311, July 1997.

[12] The Mathworks, Inc. *MATLAB, High-Performance Numeric Computation and Visualization Software*. Natick, Massachusetts, 1992.

	SNR		
σ	∞ dB	60 dB	30 dB
0.75	0.00	0.00	2.93
1.0	0.04	0.04	0.17
2.0	0.00	0.00	0.09

Table 1. % MSE in PSF estimates for Gaussian blur with all 16 frames available

	SNR		
σ	∞ dB	60 dB	30 dB
0.75	3.43	3.27	4.79
1.0	0.02	0.02	0.53
2.0	0.01	0.01	0.11

Table 2. % MSE in PSF estimates for Gaussian blur with 10 randomly chosen frames

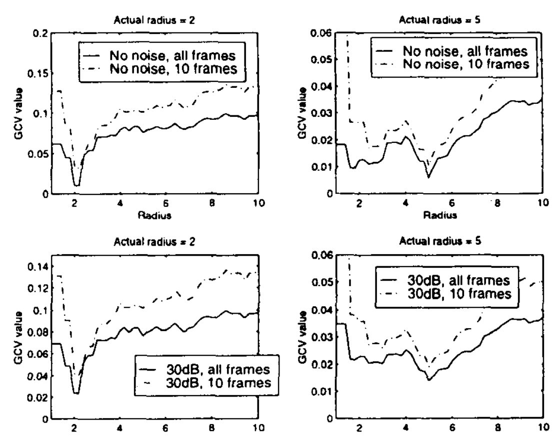

Figure 2. GCV plot for pillbox blur

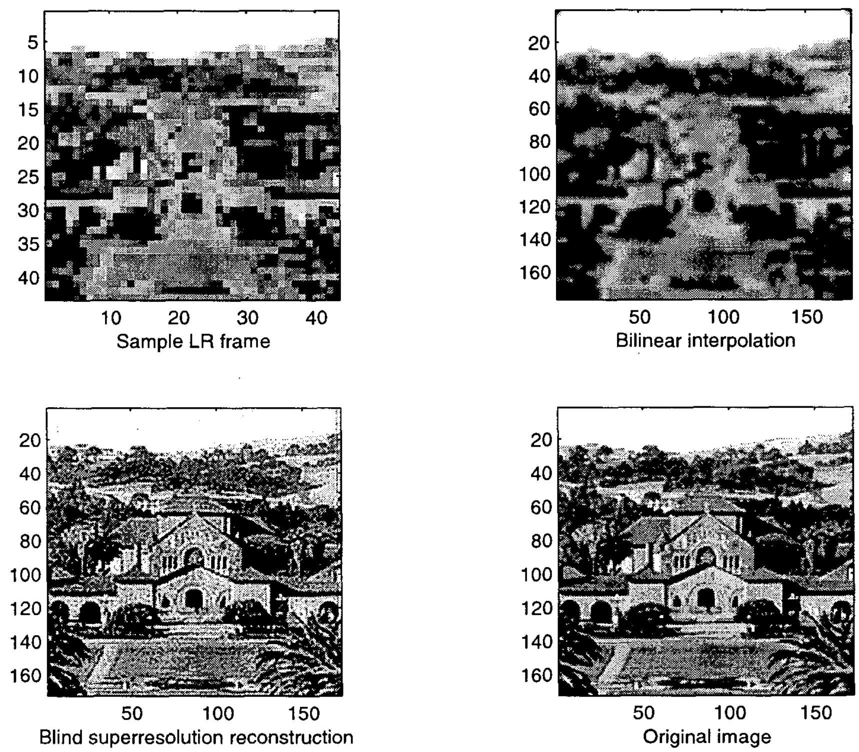

Figure 1. Blind superresolution with noisy LR frames (30 dB)