A NOVEL GCV-BASED CRITERION FOR PARAMETRIC PSF ESTIMATION

Feng Xue, Jiaqi Liu, Zhifeng Li, Gang Meng

National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics
Beijing, 100076, China

ABSTRACT

We propose a *generalized cross validation* (GCV) as a novel criterion for estimating a point spread function (PSF) from the degraded image only. The PSF is obtained by minimizing this new objective functional over a family of Wiener processings. Based on this estimated PSF, we then perform deconvolution using our recently developed SURE-LET algorithm. The GCV-based criterion is exemplified with a number of parametric PSF, involving a scaling factor that controls the blur size. A typical example of such parametrization is the Gaussian kernel.

The experimental results demonstrate that the GCV minimization yields highly accurate estimates of the PSF parameters, which also result in a negligible loss of visual quality, compared to that obtained with the exact PSF. The highly competitive results outline the great potential of developing more powerful blind deconvolution algorithms based on this criterion.

Index Terms— Blind deconvoluiton, parametric PSF estimation, GCV, Wiener filtering

1. INTRODUCTION

As a standard linear inverse problem, blind image deconvolution has been an important image processing topic for several decades. It is a highly ill-posed problem due to the ill-conditioned or singular convolution operator and its underdetermined nature [1–3]. Blind deconvolution is often tackled by regularization or Bayesian framework, which enforces a certain regularity conditions (e.g. smoothness, sparsity or total variation) on both original image and PSF, and formulates the problem as minimization of an objective functional, see [1–3] for example.

In some specific applications, the image acquisition of instrument can be modeled by physical description, and thus, the parametric forms of the PSF can be either theoretically available or practically assumed [4, 5]. Typical examples of the parametric approach can be found in the applications of fluorescence microscopy [4] and infrared imaging [5]. However, due to the limitations, the PSF parameters are unknown or imperfectly known. To this end, the present paper is devoted to parametric blind deconvolution, which attempts to estimate original image and PSF parameters, from the observed image only.

Regarding parametric PSF estimation, a number of methods have been proposed for (particular) PSF types. In [6], the authors estimate the PSF parameters by kurtosis minimization of the restored image. The DL1C method estimates PSF parameter by selecting to be at the maximum point of the differential coefficients of restored image Laplacian ℓ^1-norm curve [7]. However, the two methods mentioned above merely provide empirical solutions with no guarantee on the accuracy of PSF estimation. In addition, APEX [8] simply fits the blurred image with a PSF with a low parametric order with fast Fourier decrease, typically a Gaussian function.

In the present paper, we propose a novel criterion for parametric PSF estimation — *generalized cross validation* (GCV). It was first proposed by G. Wahba for optimizing smoothing parameter of spline smoothing, then, applications were extended to the problems of selecting the ridge/regularization parameter [9]. In the work of [10], Reeves *et al.* first applied the GCV as a criterion for blur identification. However, [10] did not specify the approach in details. In this work, we will verify and exemplify the GCV-based criterion for PSF estimation, and then, apply SURE-LET algorithm to perform deconvolution with the estimated PSF [11].

2. GCV: A NOVEL CRITERION FOR BLIND PSF ESTIMATION

2.1. Problem Statement

Consider the linear model

$$\mathbf{y} = \mathbf{H}_0\mathbf{x} + \mathbf{b}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^N$ is the degraded image of the original (unknown) image $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{H}_0$ denotes the latent true (unknown) convolution matrix constructed by the PSF, the vector $\mathbf{b} \in \mathbb{R}^N$ is a zero-mean additive Gaussian white noise with variancex σ^2. Our purpose is to estimate the matrix $\mathbf{H}_0$, such that the estimated $\mathbf{H}$ is as close to the true $\mathbf{H}_0$ as possible, from the observed data $\mathbf{y}$ only.

978-1-4799-2186-7/14/$31.00 ©2014 IEEE

2.2. Prediction error: an *oracle* criterion for the PSF estimation

Denoting a linear function (or processing) by matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, applied to the observed data $\mathbf{y}$, the expected estimation error (referring to $\mathbf{x}$) is defined as $\frac{1}{N}\mathbb{E}\left\{\|\mathbf{W}\mathbf{y}-\mathbf{x}\|^2\right\}$, where $\mathbf{W}\mathbf{y}$ is an estimate of $\mathbf{x}$ by linear processing $\mathbf{W}$, $\mathbb{E}\{\cdot\}$ denotes the mathematical expectation [11]. It is well-known that for the linear model (1) with the known matrix $\mathbf{H}$, the ideal linear processing $\mathbf{W_H}$ that minimizes the expected estimation error is Wiener filtering, expressed as [12]:

$$W_H(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + C(\omega)/S(\omega)} \tag{2}$$

in Fourier domain, where $H(\omega)$ is the Fourier representation of $\mathbf{H}$, $S(\omega)$ and $C(\omega)$ are the mean power spectral densities of signal $\mathbf{x}$ and noise $\mathbf{b}$, respectively. The subscript H of Wiener filtering W emphasizes the dependency of W upon H.

Instead of the expected estimation error, we consider the following expected prediction error (EPE: referring to $\mathbf{H}_0\mathbf{x}$):

$$\mathrm{EPE} = \mathbb{E}\Big\{\underbrace{\frac{1}{N}\left\|\mathbf{H}\mathbf{W_H}\mathbf{y}-\mathbf{H}_0\mathbf{x}\right\|^2}_{\text{prediction error } T(\mathbf{H})}\Big\} \tag{3}$$

as an *oracle*[1] criterion for estimating $\mathbf{H}$. Note that the terms *estimation error* and *prediction error* stem from model selection problem, see [13] for example.

Now, if we base our processing $\mathbf{W}$ on Wiener filtering (2), then the following theorem shows that the solution $H(\omega)$ that minimizes the EPE (3) has the same magnitude as the true $H_0(\omega)$.

Theorem 2.1 *Consider only linear processings* $\mathbf{W_H}$ *in the form of Wiener filtering defined as* (2)*. Minimizing the expected prediction error over* $\mathbf{H}$*:*

$$\min_{\mathbf{H}} \frac{1}{N}\mathbb{E}\left\{\left\|\mathbf{H}\mathbf{W_H}\mathbf{y}-\mathbf{H}_0\mathbf{x}\right\|^2\right\}, \tag{4}$$

yields that $|H(\omega)| = |H_0(\omega)|$*.*

The theorem can be easily proved by Wiener theory for denoising/deconvoluiton [12]. We omit it here. This theorem shows that the minimization of EPE is essentially equivalent to matching $|H(\omega)|$ to the true $|H_0(\omega)|$ in Fourier domain. Hence, if the frequency response of the PSF has positive value with zero phase shift at any frequency ω, the EPE minimization succeeds in estimating the accurate PSF, by seeking its frequency response.

[1] *Oracle* means that this criterion is not accessible in practice, due to the unknown $\mathbf{H}_0\mathbf{x}$ in (3).

2.3. Approximation of the exact Wiener filtering

Notice that the exact Wiener filtering $W_H(\omega)$ in (2) cannot be used in practice, since $C(\omega)/S(\omega)$ is unknown. Consider the basic observation that for natural images with strong low frequencies and weak high frequencies, $C(\omega)/S(\omega)$ increases as the frequency ω goes up. We replace $C(\omega)/S(\omega)$ by $\lambda\omega^2$. Then, we obtain the approximated Wiener filtering $W_{H,\lambda}(\omega)$ as:

$$W_{H,\lambda}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \lambda\omega^2} \Longleftrightarrow \mathbf{W}_{\mathbf{H},\lambda} = (\mathbf{H}^{\mathrm{T}}\mathbf{H} + \lambda\mathbf{R})^{-1}\mathbf{H}^{\mathrm{T}} \tag{5}$$

where $\mathbf{W}_{\mathbf{H},\lambda}$ is the matrix notation, the eigenvalue of the circulant matrix $\mathbf{R}$ is ω^2. Due to the regularization parameter λ introduced, we formulate the PSF estimation as minimizing the prediction error T over both $\mathbf{H}$ and λ, i.e.,

$$\min_{\mathbf{H},\lambda} \underbrace{\frac{1}{N}\left\|\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda}\mathbf{y}-\mathbf{H}_0\mathbf{x}\right\|^2}_{T(\mathbf{H},\lambda)} \tag{6}$$

By Theorem 2.1, the solution $H(\omega)$ to (6) satisfies $|H(\omega)| \approx |H_0(\omega)|$.

2.4. GCV: an asymptotically unbiased estimator of the prediction error

Notice that we cannot directly minimize the prediction error in practice, since $\mathbf{H}_0\mathbf{x}$ is unknown. However, based on the linear model (1), the quantity of the prediction error T can be replaced by a statistical estimate — GCV, involving only the measurements $\mathbf{y}$, as summarized in the following theorem.

Theorem 2.2 *Given the linear model (1), GCV, expressed by:*

$$V = \frac{\frac{1}{N}\left\|(\mathbf{I}-\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda})\mathbf{y}\right\|^2}{\left[\frac{1}{N}\mathrm{Tr}(\mathbf{I}-\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda})\right]^2} \tag{7}$$

is asymptotically close to $\mathbb{E}\{T\}-\sigma^2$ *as* $N\to\infty$*, if* $\frac{1}{N}\mathrm{Tr}(\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda}) \to 0$ *and* $\frac{[\mathrm{Tr}(\mathbf{H}\mathbf{W_H})/N]^2}{\mathrm{Tr}(\mathbf{H}\mathbf{W_H})^2/N} \to 0$ *as* $N\to\infty$*, where* Tr *denotes matrix trace.*

Refer to [9] for the complete proof. Note that the consistently unbiased estimate of the prediction error T has been derived by C. L. Mallows in [14], namely, Mallows' statistics C_L. Here, we would like to stress that C_L requires the knowledge of noise variance σ^2, but GCV not. This is a key advantage of using GCV. In addition, for image processing, the pixel number N is sufficiently large (typically, at least $256^2 = 65536$). And also fortunately, the conditions $\frac{1}{N}\mathrm{Tr}(\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda}) \to 0$ and $\frac{[\mathrm{Tr}(\mathbf{H}\mathbf{W_H})/N]^2}{\mathrm{Tr}(\mathbf{H}\mathbf{W_H})^2/N} \to 0$ are often satisfied for the blur kernel $\mathbf{H}$. Therefore, GCV is a reliable and convenient estimate of the prediction error, i.e., $\mathbb{E}\{V\} \approx \mathbb{E}\{T\}-\sigma^2$.

2.5. Summary

Finally, we formulate the PSF estimation as minimization of GCV, given as:

$$\min_{\mathbf{H},\lambda} \frac{\frac{1}{N}\left\|(\mathbf{I}-\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda})\mathbf{y}\right\|^2}{\left[\frac{1}{N}\mathrm{Tr}(\mathbf{I}-\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda})\right]^2} \tag{8}$$

The proposed method is summarized in Fig.1. Also note that the minimization of prediction error (6) serves as an *oracle* counterpart of the GCV minimization (8).

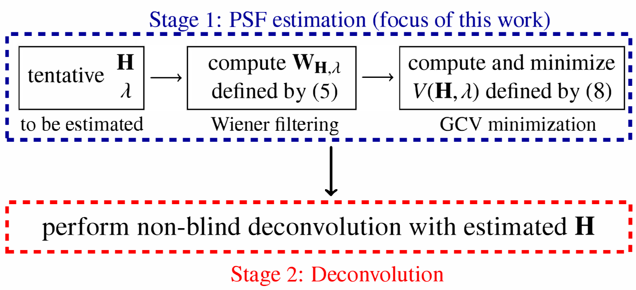

Fig. 1. The flowchart of PSF estimation: joint minimization of the GCV over $\mathbf{H}$ and λ, as shown in (8).

The GCV criterion is especially applicable for parametric PSF estimation. Suppose the PSF is of known parametric form, such that the PSF is completely represented by a small number of unknown parameters $\mathbf{s} = [s_1, s_2, ..., s_P]^{\mathrm{T}}$. Therefore, following (8), we formulate the parameter estimation as:

$$\min_{\mathbf{s},\lambda} \underbrace{\frac{\frac{1}{N}\left\|(\mathbf{I}-\mathbf{H}_{\mathbf{s}}\mathbf{W}_{\mathbf{s},\lambda})\mathbf{y}\right\|^2}{\left[\frac{1}{N}\mathrm{Tr}(\mathbf{I}-\mathbf{H}_{\mathbf{s}}\mathbf{W}_{\mathbf{s},\lambda})\right]^2}}_{\text{GCV: } V(\mathbf{s},\lambda)} \tag{9}$$

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

3.1. Experimental setting

Now, we exemplify the proposed approach with two typical parametrized PSF:

- **Gaussian** kernel, with an unknown parameter — blur variance s^2:

$$h(i,j;s) = C \cdot \exp\left(-\frac{i^2+j^2}{2s^2}\right) \tag{10}$$

- *jinc* function, with an unknown scaling factor t:

$$h(i,j;t) = C \cdot \left[\frac{2J_1(r/t)}{r/t}\right]^2 \tag{11}$$

where $J_1(\cdot)$ is first-order Bessel function of first kind, the radius $r = \sqrt{i^2+j^2}$. The constant C in (10) and (11) is a normalization factor, s.t. $\sum_{i,j} h(i,j) = 1$.

The Gaussian kernel has been widely used to model the PSF in many scenarios [5]. The terminology *jinc* is due to the structural similarity to *sinc* function. It is often used to describe optical diffraction-limited condition, where the parameter t is related to wave number and aperture diameter of optical imaging system [4].

We perform the following synthetic experiments: the test image is *Cameraman* of size 256×256, which is blurred by Gaussian kernel (10) with the true value $s_0 = 2.0$ and *jinc* function (11) with the true value $t_0 = 2.0$, respectively, and corrupted by Gaussian white noise with various variances $\sigma^2 = 1.0, 5.0, 10.0, 50.0$.

3.2. Results of parametric PSF estimation

Taking *Cameraman* and noise variance $\sigma^2 = 1.0$ for example, Fig.2 shows the estimated s and t by the GCV minimization (navy blur curves) and by the *oracle* prediction-error minimization $T + \sigma^2$ (dashed red curves), corresponding to Gaussian kernel and *jinc* function, respectively. We can see that the estimated s and t are very close to true value s_0 and t_0.

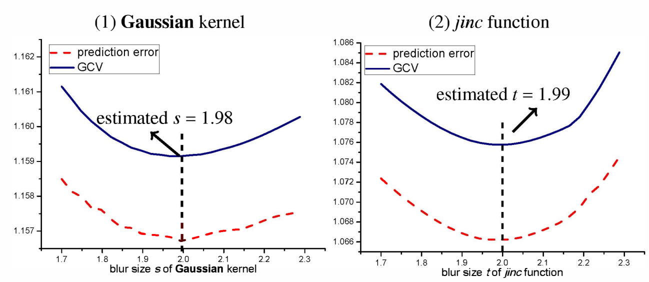

Fig. 2. The GCV minimization under two typical convolution kernels: *Cameraman*, noise variance is $\sigma^2 = 1.0$.

Under various noise levels for *Cameraman*, Table 1 reports the estimation results and present the comparisons with other state-of-the-art methods, including kurtosis [6] and DL1C [7]. The proposed approach is denoted by GCV. We can see that for all cases: (1) the estimated parameters by the GCV minimization are very close to the true values s_0 and t_0; (2) the estimation accuracy of the proposed method outperforms other competitors in average.

Table 1. The estimated blur sizes for *Cameraman*

blur type	Gaussian				*jinc*			
σ^2	1	5	10	50	1	5	10	50
kurtosis	2.04	2.10	2.12	2.17	1.94	1.96	**1.99**	**2.04**
DL1C	2.12	2.21	2.27	—	2.04	2.04	2.04	—
GCV	**1.98**	**1.96**	**1.95**	**1.93**	**1.99**	**2.03**	2.04	**2.04**

3.3. Performance of blind deconvolution

It is important to evaluate the deconvolution performance with the estimated PSF by the GCV minimization. In this work, we perform deconvolution using our recently developed SURE-LET algorithm [11], and the deconvolution performance is measured by PSNR[2]. From Figure 3 we can see that deconvolution using estimated PSF has very similar performance with using exact one: the PSNR loss is within 0.05dB.

Fig. 3. Left: **Gaussian**-blurred image *Cameraman*; Middle: Deconvolution by **Gaussian** function with estimated $s = 1.98$; Right: Deconvolution by exact **Gaussian** function with $s_0 = 2.0$.

3.4. Application to real image

In our last set of experiments, the GCV-based method is applied to real image *Jupiter* captured by telescope, shown in Figure 4. There is no exact expression of the PSF for this image; however, the PSF can be well approximated by Gaussian function [5]. Figure 4 shows the restoration by our proposed method. We can easily see that our proposed approach achieves significant improvement of visual quality.

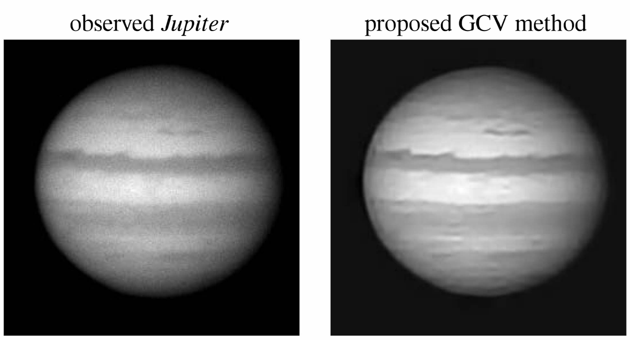

Fig. 4. Restoration of *Jupiter*: the estimated blur size of Gaussian kernel is $s = 2.41$ by the proposed GCV method.

4. CONCLUSIONS

In this paper, we proposed a new PSF estimation method based on a new criterion — GCV: an asymptotically unbiased estimator of the prediction error. We have shown that based on Wiener filtering, the GCV minimization yields highly accurate parametric PSF estimation. The proposed GCV framework is exemplified by two typical parametric forms of PSF.

Results obtained show that the proposed method has significant improvement of quality both numerically and visually. The examples of the blur kernel listed in the paper are but the exemplifications of the GCV framework for PSF estimation. It is worth noting that the GCV minimization itself does not specify any particular parametric form of PSF, as shown in (7). There is huge potential to develop specific algorithms for various applications, e.g. fluorescence microscopy [4], based on the GCV criterion.

[2]PSNR: for 8-bit images, peak signal-to-noise ratio defined as PSNR = $10\log_{10}\frac{255^2}{\|\hat{\mathbf{x}}-\mathbf{x}\|^2/N}$ in dB [11].

5. REFERENCES

[1] T.F. Chan and C.K. Wong, "Total variation blind deconvolution," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998.

[2] R. Molina, J. Mateos, and A.K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3715–3727, 2006.

[3] Haiyong. Liao and Michael K. Ng, "Blind deconvolution using generalized cross-validation approach to regularization parameter estimation," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 3005–3019, 2011.

[4] T. Kenig, Z. Kam, and A. Feuer, "Blind image deconvolution using machine learning for three-dimensional microscopy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2191–2204, 2010.

[5] Feng Xue and Anatoly G Yagola, "Analysis of point-target detection performance based on ATF and TSF," *Infrared Physics & Technology*, vol. 52, no. 5, pp. 166–173, 2009.

[6] Dalong Li and Steve Simske, "Atmospheric turbulence degraded-image restoration by kurtosis minimization," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 2, pp. 244–247, 2009.

[7] Fen Chen and Jianglin Ma, "An empirical identification method of Gaussian blur parameter for image deblurring," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2467–2478, 2009.

[8] A.S. Carasso, "The APEX method in image sharpening and the use of low exponent Lévy stable laws," *SIAM Journal on Applied Mathematics*, pp. 593–618, 2002.

[9] Gene H Golub, Michael Heath, and Grace Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.

[10] Stanley J Reeves and Russell M Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Transactions on Image Processing*, vol. 1, no. 3, pp. 301–311, 1992.

[11] F. Xue, F. Luisier, and T. Blu, "Multi-Wiener SURE-LET deconvolution," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1954–1968, 2013.

[12] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series*, New York: Wiley, 1964.

[13] Trevor Hastie, Robert Tibshirani, Jerome Friedman, T Hastie, J Friedman, and R Tibshirani, *The Elements of Statistical Learning*, vol. 2, Springer, 2009.

[14] Colin L Mallows, "Some comments on C_p," *Technometrics*, vol. 15, no. 4, pp. 661–675, 1973.