J. Vis. Commun. Image R. 33 (2015) 115–122

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Mallows' statistics C_L: A novel criterion for parametric PSF estimation ☆

Feng Xue *, Jiaqi Liu, Zhifeng Li, Shengdong Liu, Gang Meng, Li Zhang

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

ARTICLE INFO

Article history:
Received 1 October 2014
Accepted 7 September 2015
Available online 25 September 2015

Keywords:
Blind deconvolution
Parametric PSF estimation
Mallows' statistics
Smoother filtering
Frequency-dependent regularizer
Expected prediction error
Gaussian kernel
jinc function

ABSTRACT

Considering blind image deconvolution as a statistical estimation problem, we propose an unbiased estimator of the prediction error – Mallows' statistics C_L – 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 smoother filterings (with frequency-dependent regularization term). We then perform non-blind deconvolution using the popular BM3D algorithm. The C_L-based framework is exemplified with a number of parametric PSF's, involving a scaling factor that controls the blur size. A typical example of such parametrization is the Gaussian kernel.

The experimental results show that the C_L-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 demonstrate the great potential of developing more powerful blind deconvolution algorithms based on the C_L-estimator.

1. Introduction

As a standard linear inverse problem, blind image deconvolution has been an important image processing topic for several decades. The task amounts to estimating both original image and point spread function (PSF) from the observed image only. Notable examples of real applications can be found in medical imaging [1], microscopy [2], astronomy [3,4], remote sensing [5] and infrared imaging [6]. To address the ill-posedness of the problem, it is typical to introduce a certain consumption or priors of the original image and the PSF by means of regularization techniques [7–9] or Bayesian approaches [10,11], which utilize optimization methods to simultaneously estimate both original image and the PSF.

Recently, it is preferable to estimate the original image and the PSF separately: firstly estimate the PSF, and then, perform non-blind deconvolution. This strategy is more appealing, since it allows to use any developed high-quality non-blind deconvolution algorithms. In this work, we choose this procedure, and since there are already several excellent non-blind deconvolution algorithms available (see BM3D [12] for example), we are going to focus on PSF estimation.

Non-parametric PSF estimation – The PSF is represented by discrete pixel values. In this regime, it is crucial to incorporate a certain assumption of the PSF, within regularization or Bayesian framework [10,11,9,8].

Parametric PSF estimation – In specific applications, the parametric forms of the PSF can be either theoretically available or practically assumed, from the physical description of the image acquisition [4,5,2,6,13,14]. Typical examples of the parametric approach can be found in the applications of optical imaging [14], fluorescence microscopy [2,15,13], atmospheric turbulence [16,17] and astronomy [4], interferometry [18]. However, due to the limitations, the PSF parameters are unknown or imperfectly known, and thus, need to be estimated, in addition to the original image.

In parametric representation, the PSF is completely characterized by a small number of parameters, which dramatically reduces the degrees of freedom of PSF estimation [15,19]. Moreover, the parametric form confines the solution of PSF to a predefined function space, and avoids delta function as the trivial solution [19]. To this end, the present paper is devoted to parametric estimation – estimating the PSF parameters, from the observed image only.

A number of methods have been proposed for (particular) PSF types. The PSF parameters can be estimated by kurtosis minimization of the restored image [17]. Chen et al. estimated PSF parameter by selecting to be at the maximum point of the differential coefficients of restored image Laplacian ℓ^1-norm curve [20]. However, all the approaches mentioned above need to manually adjust regularization parameter for restoration. In addition, kurtosis [17]

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author.
E-mail addresses: fxue2012@gmail.com (F. Xue), ljq006@vip.sina.com (J. Liu), zfli1976@163.com (Z. Li), liushdbit1019@163.com (S. Liu), mgang2012@yeah.net (G. Meng), zhangli60@163.com (L. Zhang).

http://dx.doi.org/10.1016/j.jvcir.2015.09.001
1047-3203/

and DL1C [20] methods merely provide empirical solutions, being lack of theoretical justification of the accuracy of PSF estimation. APEX method estimated the PSF parameters by matching the PSF to the blurred image in frequency domain [21]. This method is only applied to a special type of PSF, which restricts a wider range of applications.

Among all the parametric forms of the PSF [2,15,21], of great significance and applicability is the Gaussian function, since it can be used for modelling many degradation scenarios in real applications. For example, Gaussian function can be used as the approximation of PSF of fluorescence microscopy [22,13], electro-optical systems [5,23,6] and atmospheric turbulence [4,16,24,17]. Thus, the estimation of Gaussian blur variance (the only unknown PSF parameter) becomes particularly interesting. As mentioned above, kurtosis-based [17], DL1C [20] and APEX [21] methods can be applied to this problem.

Another example deserving close attention is *jinc* function,[1] which is often used to describe the PSF under the diffraction-limited condition in optics. The scaling factor of *jinc* function (related to aperture diameter and wavelength [26]) needs to be estimated, if the optical instrument is not exactly clarified.

In the present paper, we propose a novel criterion for parametric PSF estimation – Mallows' statistics C_L. The statistics C_L, which was first proposed by Mallows [27], is an unbiased estimator of prediction error.[2] In statistics, it has been extensively studied and applied for model selection problems, typically, subset-regression and ridge estimator [28,29]. Now, considering deconvolution as estimation problem of linear model, we adopt the statistics C_L as a new criterion for parametric PSF estimation. We show that the C_L-minimization leads to highly accurate estimation of PSF parameters both theoretically and experimentally.

The paper is organized as follows. Section 2 is devoted to the theoretical analysis, where we propose and validate a novel criterion for PSF estimation – Mallows' statistics C_L, incorporating a simple smoother filtering. In Section 3, we exemplify the proposed framework with several particular types of PSF, and report the experimental results for discussion. Some concluding remarks are finally given in Section 4.

Throughout this paper, we use boldface lowercase letters, e.g. $\mathbf{x} \in \mathbb{R}^N$, to denote N-dimensional real vectors, where N is typically the number of pixels in an image. The nth element of $\mathbf{x}$ is written as x_n. The linear transformations (matrices) $\mathbb{R}^N \to \mathbb{R}^M$ are denoted by boldface uppercase letters, e.g. $\mathbf{H} \in \mathbb{R}^{M\times N}$. $\mathbf{H}^{\mathrm{T}} \in \mathbb{R}^{N\times M}$ denotes the transpose of matrix $\mathbf{H}$. Also note that we use the subscript $(\cdot)_0$ to denote the true quantity of $(\cdot)$, for example, matrix $\mathbf{H}_0$ is the true quantity of $\mathbf{H}$.

2. Theoretical background

2.1. Problem statement

Consider the linear model

$$\mathbf{y} = \mathbf{H}_0\mathbf{x} + \mathbf{b}, \quad \mu_0 = \mathbf{H}_0\mathbf{x} \tag{1}$$

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 $\mathbf{h}_0$, the vector $\mathbf{b} \in \mathbb{R}^N$ is a zero-mean additive Gaussian white noise with variances σ^2. For the convenience of the following discussions, we denote the noise-free blurred data $\mathbf{H}_0\mathbf{x}$ by μ_0. For the PSF estimation, our purpose is to accurately estimate the matrix $\mathbf{H}_0$, from the observed data $\mathbf{y}$ only.

2.2. Expected prediction error: an oracle criterion for PSF estimation

In statistics, it is conventional to refer to the computation of $\mathbf{x}$ as *estimation*, and to the computation of μ_0 as *prediction* [30]. To predict μ_0, denoting a linear function (or processing) by matrix $\mathbf{U} \in \mathbb{R}^{N\times N}$, applied to the observed data $\mathbf{y}$, the expected prediction error (EPE: referring to the estimation of μ_0) is defined as [30]:

$$\mathrm{EPE} = \frac{1}{N}\mathbb{E}\left\{\|\mathbf{U}\mathbf{y} - \mu_0\|^2\right\} \tag{2}$$

where $\mathbf{U}\mathbf{y}$ is an estimate of μ_0 by linear processing $\mathbf{U}$, $\mathbb{E}\{\cdot\}$ denotes the mathematical expectation.

The following theorem shows that given the linear processing $\mathbf{U}$ as an exact smoother filtering (described below), there exists a simple relation between the solution $\mathbf{H}$ to the EPE minimization and the true convolution matrix $\mathbf{H}_0$.

Theorem 2.1. *Consider only linear processings* $\mathbf{U}$ *in the form of exact smoother filtering defined as:*

$$U(\omega) = \frac{|H(\omega)|^2}{|H(\omega)|^2 + C(\omega)/S(\omega)} \tag{3}$$

in frequency domain, where $H(\omega)$ *is the Fourier representation of the PSF,* $C(\omega)$ *and* $S(\omega)$ *are the power spectral densities of signal* $\mathbf{x}$ *and noise* $\mathbf{b}$*, respectively. Minimizing the EPE (2) over* $\mathbf{H}$[3]:

$$\min_{\mathbf{H}} \frac{1}{N}\mathbb{E}\left\{\|\mathbf{U}\mathbf{y} - \mu_0\|^2\right\} \tag{4}$$

yields that $|H(\omega)| = |H_0(\omega)|$*, where* $H_0(\omega)$ *is the frequency representation of the true PSF* $\mathbf{h}_0$.

See Appendix A for the proof. This theorem states that: (1) the EPE minimization is essentially equivalent to matching $|H(\omega)|$ to the true $|H_0(\omega)|$ in Fourier domain; (2) the EPE minimization can only leads to the equality of magnitude frequency response of the PSF: $|H(\omega)| = |H_0(\omega)|$, whereas the phase response is not reflected. Hence, we consider only zero-phase blur models in this work. Since many real-life blurs – linear motion, out-of-focus and atmospheric turbulence blurs – have zero phase, this assumption is rather unrestrictive.

2.3. Mallows' statistics C_L: an unbiased estimator of the prediction error

Notice that we cannot directly minimize the prediction error, since $\mu_0 = \mathbf{H}_0\mathbf{x}$ is unknown in practice. However, based on the linear model (1), the quantity of the prediction error can be replaced by a statistical estimate – Mallows' statistics C_L [27], involving only the measurements $\mathbf{y}$, as summarized in the following theorem.

Theorem 2.2. *Given the linear model (1), Mallows' statistics* C_L *[27]:*

$$C_L = \frac{1}{N}\|\mathbf{U}\mathbf{y} - \mathbf{y}\|^2 + \frac{2\sigma^2}{N}\mathrm{Tr}(\mathbf{U}) - \sigma^2 \tag{5}$$

is an unbiased estimator of the prediction error (2), i.e., $\mathbb{E}\{C_L\} = \mathrm{EPE}$*, where* Tr *denotes matrix trace.*

See Appendix B for the proof, which is similar to the original derivation of Mallows [27]. Note that the first term $\|\mathbf{U}\mathbf{y} - \mathbf{y}\|^2$ of (5) is the residual sum of squares (RSS) in model selection problem [27]. We can see that the statistics C_L depends on the observed data $\mathbf{y}$ and the matrix $\mathbf{U}$ (involving the unknown $\mathbf{H}$ we want to estimate). Thus, it can be a practical substitute of the prediction error.

[1] The terminology *jinc* stems from optics, due to the structural similarity to *sinc* function [25].

[2] In statistics, the *prediction error* is also called *prediction loss* [28].

[3] By Parseval's theorem, EPE minimization in spatial domain is equivalent to that in frequency domain (see Appendix A *for details).*