A Nonparametric Procedure For Blind Image Deblurring

Peihua Qiu
School of Statistics
University of Minnesota
313 Ford Hall
224 Church St. SE
Minneapolis, MN 55455

Abstract

Observed images are usually blurred versions of the true images, due to imperfections of the imaging devices, atmospheric turbulence, out of focus lens, motion blurs, and so forth. The major purpose of image deblurring is to restore the original image from its blurred version. A blurred image can be described by the convolution of the original image with a point spread function (psf) that characterizes the blurring mechanism. Thus, one essential problem for image deblurring is to estimate the psf from the observed but blurred image, which turns out to be a challenging task, due to the "ill-posed" nature of the problem. In the literature, most existing image deblurring procedures assume that either the psf is completely known or it has a parametric form. Motivated by some image applications, including handwritten text recognition and calibration of imaging devices, we suggest a method for estimating the psf nonparametrically, in cases when the true image has one or more line edges, which is usually satisfied in the applications mentioned above and which is not a big restriction in some other image applications, because it is often convenient to take pictures of objects with line edges, using the imaging device under study. Both theoretical justifications and numerical studies show that the proposed method works well in applications.

Key Words: Camera calibration; Circularly symmetric function; Deconvolution; Handwritten text recognition; Image deblurring; Image restoration; Least squares estimation; Line edges; Spatial degradation; Test patterns.

1 Introduction

Observed images generated by image acquisition devices are usually not exactly the same as the true images, but are instead degraded versions of their true images (cf., Qiu 2005, Chapter 7).

Degradations can occur in the entire process of image acquisition, and there are many different sources of degradation. For instance, in aerial reconnaissance, astronomy, and remote sensing, images are often degraded by atmospheric turbulence, aberrations of the optical system, or relative motion between the camera and the object. Image degradations can be classified into several categories, among which *point degradations* (or, noise) and *spatial degradations* (or, blurring) are most common in applications. Other types of degradations involve chromatic or temporal effects. For a detailed discussion about formation and description of various degradations, please read books such as Andrews and Hunt (1977) and Bates and McDonnell (1986).

Image restoration is a process to restore an original image f from its observed but degraded version Z. In the literature, a commonly used model for describing the relationship between f and Z is as follows:

$$Z(x,y) = h \otimes f(x,y) + \varepsilon(x,y), \quad \text{for } (x,y) \in \Omega, \tag{1.1}$$

where h is a 2-D function representing the spatial blurring mechanism, $\varepsilon(x,y)$ is a pointwise noise at (x,y), Ω is the design space, and $h \otimes f$ denotes the convolution between h and f, defined by

$$h \otimes f(x,y) = \int\int_{R^2} h(u,v) f(x-u, y-v)\, dudv. \tag{1.2}$$

In model (1.1), it is assumed that the true image f is degraded spatially by h and pointwise by ε, the spatial degradation is linear and location invariant, and the point degradation is additive. These assumptions are valid (at least approximately) for most applications. But for some applications, one or more such assumptions may not hold. For instance, photocopies are often nonlinear transformations of their original documents. For more discussion about these assumptions, see Rosenfeld and Kak (1982, Chapters 6 and 7).

When the true image intensity function f is a delta (or, impulse) function having an area of unity in any infinitesimal neighborhood of a given point (x^*, y^*), then, from expression (1.2), its blurred version is just $h(x-x^*, y-y^*)$. That is, the function h actually describes the degraded version of a point source image. For this reason, h is called the *point spread function* (psf) in the literature.

If the psf h is not zero at the origin and zero everywhere else, then there is no blurring in the observed image. In such cases, the true image can be restored from the observed image by removing noise, or *denoising*, and the major issue becomes how to preserve edges when removing noise from the observed image. Several edge-preserving image restoration procedures have been

proposed in the literature. For instance, see Besag (1986), Geman and Geman (1984), Gijbels *et al.* (2006), Hillebrand and Müller (2007), Perona and Malik (1990), Polzehl and Spokoiny (2000, 2003), Saint-Marc *et al.* (1991), Tomasi and Manduchi (1998), Qiu (1998, 2004), and Wang (1998). Overviews about these methods can be found in Li (1995) and Qiu (2005, 2007).

When there is blurring in the observed image, in order to restore the true image from the observed one, the psf h should be specified properly, besides denoising. The research problem to restore the true image from its blurred-and-noisy version is often called image *deblurring*. Generally speaking, the image deblurring problem is ill-posed, in the sense that there might be two or more different sets of h and f corresponding to a same blurred image and that the inverse problem to estimate f from Z often involves some singularities (cf., equation (1.3) below). Therefore, *it is impossible to estimate both h and f properly from the observed image Z without using any extra information about either f, or h, or both.*

In the literature, many image deblurring procedures assume that the psf h is known. This assumption is reasonable in some cases, because h can be specified (at least approximately) based on our knowledge about the image acquisition device (e.g., camera). For instance, the *relative motion blur model* is appropriate when blurring is mainly caused by relative motion between the image acquisition device and the object. The *Gaussian blur model* is often used for describing deblurring caused by atmospheric turbulence in remote sensing and aerial imaging. For detailed introduction about these blurring models, see Goodman (1968) and Bates and McDonnell (1986).

After h is specified, f can be estimated based on the relationship that

$$\mathcal{F}\{Z\}(u,v) = \mathcal{F}\{h\}(u,v)\mathcal{F}\{f\}(u,v) + \mathcal{F}\{\varepsilon\}(u,v), \quad \text{for } (u,v) \in R^2, \tag{1.3}$$

where $\mathcal{F}\{f\}$ denotes the Fourier transformation of f. By equation (1.3), many methods have been proposed in the literature for estimating f, including some non-iterative methods, e.g., inverse filtering, Wiener filtering, and constrained least squares filtering procedures (cf., e.g., Gonzalez and Woods 1992, Chapter 5), and several iterative methods, e.g., Lucy-Richardson procedure, Landweber procedure, Tikhonov-Miller procedure, maximum *a posteriori* (MAP) procedure, maximum entropy procedure, procedures based on EM algorithm, and so forth (cf., e.g., Skilling 1989, Figueiredo and Nowak 2003).

In many applications, however, it is difficult to specify the psf h completely, based on our prior knowledge about the image acquisition device. Image restoration when h is unknown is called the

blind image restoration problem. In the literature, a number of procedures have been proposed for blind image restoration. One type of such procedures assumes that h can be described by a parametric model with one or more unknown parameters, and the parameters together with the true image are estimated by some algorithms, most of which are iterative (e.g., Cannon 1976, Katsaggelos and Lay 1990, Carasso 2001, Joshi and Chaudhuri 2005, Hall and Qiu 2007a). Another type of procedures assumes that the true image comprises of an object with known finite support, the background is uniformly black, gray, or white, and the psf h satisfies various conditions (e.g., Yang *et al.* 1994, Kundur and Hatzinakos 1998, Hall and Qiu 2007b). For instance, Kundur and Hatzinakos (1998) assumes that h has its inverse, both h and its inverse have finite integrations over the real plane, and both the true image f and the psf h are *irreducible* in the sense that each of them can not be expressed as a convolution of two or more component images with finite support.

Based on the above introduction, we can see that it is important to estimate the psf h properly, from observed images, for two major reasons. One is that it is essential for restoring the true image from its blurred version. The second one is that it is helpful for understanding the blurring mechanism of an imaging device, which could lead to possible quality improvement of the imaging device.

In this paper, we propose a nonparametric estimator of the psf h. In our procedure, only some mild conditions are imposed on h. More specifically, h is assumed to be circularly symmetric, or, confined to one specific direction, and be a member of $L^2(\mathcal{D})$ in the sense that $\int\int_{\mathcal{D}} h^2(x,y)\,dxdy < \infty$, where $\mathcal{D}$ is a connected subset of R^2 including the origin and with finite diameter. These conditions on h are commonly used in the image restoration literature (cf., e.g., Rosenfeld and Kak 1982, Chapters 6 and 7), and psf functions satisfying these conditions should include most ones commonly used in applications, including the relative motion blur and Gaussian blur models mentioned above and the circular-exponential blur model discussed by Carasso (2001).

In order to estimate h properly, we further assume that the true image has one or more regions in which line edges are surrounded by uniform backgrounds. This condition is usually satisfied in the image applications of handwritten text recognition (e.g., Plamondon and Srihari 2000, Arica and Yarman-Vural 2001) and calibration of imaging devices (e.g., Koren 2004), because the handwritten texts can be regarded as line edges, and imaging devices are often calibrated using test patterns consisting of lines. In certain other image applications, this condition can also be roughly satisfied, because it is often possible to take images of the scenes with line edges, using the image acquisition

device under study, in order to study its blurring mechanism (cf., related discussion in Chapters 6 and 7 of Rosenfeld and Kak 1982).

The proposed procedure is substantially easier to implement than the iterative algorithms used in most earlier nonparametric techniques. It is described in detail in next section. Some of its statistical properties are discussed in Section 3. A simulation study is presented in Section 4. Two image restoration examples are discussed in Section 5. Several remarks conclude the article in Section 6. The proof of a theorem is provided in Appendix A.

2 Nonparametric estimation of the point spread function

This section is organized in three parts. In Section 2.1, we first formulate the problem in the case that the true image f has only one line edge surrounded by a uniform background in a region. Then, our procedure for estimating the point spread function (psf) h in such a case is described in Section 2.2. In Section 2.3, some generalizations are discussed.

2.1 Image restoration when the true image has a single line edge

We start with a simple case when the true image f has only one line edge surrounded by a uniform background in a region. Without loss of generality, we assume that this region is $[0,1] \times [0,1] \in \Omega$, and f has the following expression in this region:

$$f(x,y) = A\,\delta_{0.5}(y), \text{ for } (x,y) \in [0,1] \times [0,1], \tag{2.1}$$

where $A > 0$ is a constant, and $\delta_{0.5}(y)$ is an *impulse* (or *delta*) function of y, which has an area of unity in an infinitesimal neighborhood of $y = 0.5$ and which takes the value of 0 everywhere else. In such a case, f has one line edge that is parallel to the x-axis at $y = 0.5$.

In the region $[0,1] \times [0,1]$, it is assumed that there are n^2 observed image intensities generated from the model

$$Z_{ij} = h \otimes f(x_i, y_j) + \varepsilon_{ij}, \quad \text{for } i,j = 1,2,\ldots,n, \tag{2.2}$$

where $\{(x_i, y_j) = (i/n, j/n),\ i,j = 1,2,\ldots,n\}$ are equally spaced pixels in the region and $\{\varepsilon_{ij},\ i,j = 1,2,\ldots,n\}$ are independent and identically distributed (i.i.d.) random errors with mean 0 and fi-

nite variance σ^2. Obviously, model (2.2) is a discrete version of model (1.1), constrained in the region $[0,1] \times [0,1]$.

The psf h involved in the observed image is assumed to be a circularly symmetric, or, confined to one specific direction, real function in $L^2(\mathcal{D})$, where $\mathcal{D}$ is a connected subset of R^2 that includes the origin and has a finite diameter. In most references, h is also assumed to be a 2-D density function on $\mathcal{D}$ because it is believed that any blurring process does not change the image mass, which is also adopted here. That is, we assume that $h(x,y) \geq 0$, for $(x,y) \in \mathcal{D}$, and

$$\int\int_{\mathcal{D}} h(x,y)\, dxdy = 1. \tag{2.3}$$

Our goal is to estimate the psf h from observations $\{Z_{ij},\ i,j = 1,2,\ldots,n\}$. Once h is estimated, the true image f can be estimated accordingly in the entire design space Ω, using some existing procedures, such as those mentioned in Section 1.

2.2 Estimation of the psf h

In many applications, h is defined in a circle $\mathcal{D} = \{(x,y) : \sqrt{x^2+y^2} \leq r\}$ with radius r (cf., discussion about the Gaussian blur model in Section 1); in such cases, $2r$ denotes the extent of blurring and this number is usually small. Next, we first discuss the case when h is a circularly symmetric function in $L^2(\mathcal{D})$. In such cases, h can be expressed as

$$h(x,y) = \xi\left[(x^2+y^2)/r^2\right],$$

where $\xi \in L^2([0,1])$. The function ξ can be written as a linear combination of a basis of $L^2([0,1])$. Suppose that $\{b_k(x),\ k = 0,1,\ldots,\infty\}$ is a basis of $L^2([0,1])$ and that $h(x,y)$ has the following expression:

$$h(x,y) = \sum_{k=0}^{\infty} c_k b_k\left[(x^2+y^2)/r^2\right],\ \text{for } (x,y) \in \mathcal{D}. \tag{2.4}$$

Then, when the true image f has expression (2.1) and the psf h has property (2.3), it can be checked that

$$h \otimes f(x,y) = A \sum_{k=0}^{\infty} c_k b_k^*(y-0.5), \tag{2.5}$$

where

$$b_k^*(y-0.5) = 2\int_0^{\sqrt{r^2-(y-0.5)^2}} b_k\left[(u^2+(y-0.5)^2)/r^2\right]\, du,$$

which does not depend on x. Our goal is to estimate the unknown coefficients A and $\{c_k,\ k = 0, 1, \ldots, \infty\}$ based on the observed image intensities $\{Z_{ij},\ i, j = 0, 1, \ldots, n\}$ in the region $[0, 1]\times[0, 1]$ (cf., model (2.2)).

It should be pointed out that, in deriving the above formula for $b_k^*(y - 0.5)$, we have used the assumption that the design region $[0, 1] \times [0, 1]$ used in (2.1), which includes a line edge surrounded by a uniform background, is a central part of a little larger region with the same property. More specifically, we assume that there exists a region, which includes the same line edge surrounded by a uniform background, with size at least $[-r, 1 + r] \times [-r, 1 + r]$. In such cases, there is no "boundary" problem in defining $b_k^*(y - 0.5)$. Otherwise, the quantity $b_k^*(y - 0.5)$ when x is in the boundary regions $[0, r]$ and $[1 - r, 1]$ would be different from its value when x is in the interior region $(r, 1 - r)$. Since r is usually small, the above mentioned assumption can be achieved in most cases. Otherwise, the estimation procedure introduced below should be modified in the same way as that when the constant A in (2.1) depends on x, which is discussed at the end of Section 2.3.

With a given sample size n^2, it is unrealistic to estimate all the details of h from the data. That is, some terms on the right-hand-side of (2.4) should be dropped from the estimation process. Usually, the index k of the basis functions $\{b_k(x),\ k = 0, 1, \ldots, \infty\}$ is related to the support of $b_k(x)$, or the frequency that $b_k(x)$ switches between positive and negative values in a given interval. When k is larger, the support of $b_k(x)$ would be smaller, or its frequency to switch between positive and negative values in a given interval would be larger (cf., several specific sets of basis functions discussed below). In such cases, it is natural to only include the first J terms in the estimation process, where $0 \le J \le n$ is an integer, because regular estimators of the remaining terms often correspond to the high-frequency part of the data and are mainly contributed by noise. That is, we consider using

$$h_J(x, y) = \sum_{k=0}^{J-1} c_k b_k \left[(x^2 + y^2)/r^2\right], \quad \text{for } (x, y) \in \mathcal{D}. \tag{2.6}$$

to approximate h. In (2.6), integer J works as a smoothing parameter, with smaller J implying more smoothing and vice versa. In the literature, there are some other ways to drop terms for function approximation. See, for instance, Qiu (2005, Section 2.6) for introduction about the hard-thresholding and soft-thresholding schemes in the wavelet context.

Because functions $b_k^*(y - 0.5)$ in equation (2.5) do not depend on x, parameter estimation from the original data $\{Z_{ij},\ i, j = 1, 2, \ldots, n\}$ is equivalent to parameter estimation from the averaged

data $\{\overline{Z}_{.j},\ j = 1, 2, \ldots, n\}$, where $\overline{Z}_{.j} = (1/n)\sum_{i=1}^{n} Z_{ij}$, for $j = 1, 2, \ldots, n$. Let

$$\begin{aligned} \overline{\mathbf{Z}} &= \left(\overline{Z}_{.1}, \overline{Z}_{.2}, \ldots, \overline{Z}_{.n}\right)', \\ \mathbf{c} &= (c_0, c_1, \ldots, c_{J-1})', \\ \mathbf{b}^*(y - 0.5) &= \left(b_0^*(y - 0.5), b_1^*(y - 0.5), \ldots, b_{J-1}^*(y - 0.5)\right)' \end{aligned}$$

and

$$\mathbf{B} = (\mathbf{b}^*(y_1 - 0.5), \mathbf{b}^*(y_2 - 0.5), \ldots, \mathbf{b}^*(y_n - 0.5))'.$$

Then, the relationship between the averaged data $\overline{\mathbf{Z}}$ and the parameter vector $\mathbf{c}$ can be approximated by the following linear model:

$$\overline{\mathbf{Z}} = (A\mathbf{B})\mathbf{c} + \overline{\varepsilon}, \tag{2.7}$$

where $\overline{\varepsilon} = (\overline{\varepsilon}_{.1}, \overline{\varepsilon}_{.2}, \ldots, \overline{\varepsilon}_{.n})'$, and $\overline{\varepsilon}_{.j} = (1/n)\sum_{i=1}^{n} \varepsilon_{ij}$. In the case when the inverse $(\mathbf{B}'\mathbf{B})^{-1}$ exists, the parameter vector $\mathbf{c}$ can be estimated by its least squares (LS) estimator

$$\widehat{\mathbf{c}}_A = \frac{1}{A}\left(\mathbf{B}'\mathbf{B}\right)^{-1}\mathbf{B}'\overline{\mathbf{Z}}. \tag{2.8}$$

In equation (2.8), the parameter A is often unknown and should be estimated from data. Toward this end, the regularity condition (2.3) can be used. By this condition, we define

$$\widehat{A} = \mathbf{b}'\left(\mathbf{B}'\mathbf{B}\right)^{-1}\mathbf{B}'\overline{\mathbf{Z}}, \tag{2.9}$$

where

$$\mathbf{b} = \left(\int\int_{\mathcal{D}} b_0[(x^2 + y^2)/r^2]\, dxdy, \int\int_{\mathcal{D}} b_1[(x^2 + y^2)/r^2]\, dxdy, \ldots, \int\int_{\mathcal{D}} b_{J-1}[(x^2 + y^2)/r^2]\, dxdy\right)'.$$

By (2.8) and (2.9), the parameter vector $\mathbf{c}$ can be estimated by

$$\widehat{\mathbf{c}} = \frac{(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\overline{\mathbf{Z}}}{\mathbf{b}'\,(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\overline{\mathbf{Z}}}. \tag{2.10}$$

Finally, the psf h can be estimated by

$$\widehat{h}(x, y) = \sum_{k=0}^{J-1} \widehat{c}_k b_k\left[(x^2 + y^2)/r^2\right], \text{ for } (x, y) \in \mathcal{D}. \tag{2.11}$$

The estimator $\widehat{h}(x, y)$ defined in equation (2.11) has unit integration in $\mathcal{D}$; but it may not be nonnegative in $\mathcal{D}$. If the nonnegativity of the estimator is desired, then it can be replaced by the following modified version:

$$\widehat{h}_M(x, y) = \frac{\widehat{h}(x, y) - \min_{\mathcal{D}} \widehat{h}(s, t)}{\int\int_{\mathcal{D}} \left(\widehat{h}(u, v) - \min_{\mathcal{D}} \widehat{h}(s, t)\right)\, dudv} \tag{2.12}$$

In either the estimator $\widehat{h}(x,y)$ or its modified version $\widehat{h}_M(x,y)$, the parameter r, which equals half of the blurring extent, and the smoothing parameter J should be chosen properly. In applications, r can be roughly determined from the extent of blurring of some sharp components in the image, such as step edges, corners, line edges, and so on. Both r and J can also be determined by some existing parameter selection procedures, such as the cross-validation (CV), generalized cross-validation (GCV), Mallow's C_p, plug-in procedures, and so forth (cf., e.g., Loader 1999). For instance, they can be determined by the CV procedure as follows. Let $\widehat{A}_{-j}$ and $\widehat{\mathbf{c}}_{-j}$ denote the estimators of A and $\mathbf{c}$, respectively, using the above estimation procedure (2.7)–(2.10), from the observed data with the jth column $\{(x_i, y_j),\ i = 1, 2, \ldots, n\}$ excluded. Then, the CV score is defined by

$$CV(r,J) = \sum_{j=1}^{n} \left\{ \overline{Z}_{\cdot j} - \widehat{A}_{-j}[\mathbf{b}^*(y_j - 0.5)]'\widehat{\mathbf{c}}_{-j} \right\}^2. \tag{2.13}$$

The optimal r and J can be approximated by the minimizers of $CV(r,J)$.

To use the proposed estimation procedure, we still need to choose the basis $\{b_k(x),\ k = 0, 1, \ldots, \infty\}$ of $L^2([0,1])$, used in (2.4). One such basis commonly used in image analysis is the cosine-series basis, by which (2.4) becomes

$$h(x,y) = \sum_{k=0}^{\infty} c_k \cos\left[k\pi(x^2+y^2)/r^2\right], \text{ for } (x,y) \in \mathcal{D}. \tag{2.14}$$

Obviously, expression (2.14) is the well-known inverse Fourier cosine transformation of $h(x,y)$ (e.g., Hankerson *el al.* 1998). Other possible bases include the wavelet basis and various polynomial bases. The polynomial bases can be used only when h is a continuous function.

Now, we briefly discuss the case when the psf h is confined to one specific direction θ and in $L^2(\mathcal{D})$ as well. Obviously, the linear relative motion blur model discussed in Section 1 belongs to this case. Without loss of generality, we assume that h is confined to the y-axis (i.e., $\theta = \pi/2$). Then, it can be expressed as

$$h(x,y) = \delta_0(x)\eta(y/r), \text{ for } y \in [0,r],$$

where $\eta \in L^2([0,1])$. Similar to ξ, η can be expressed as a linear combination of a basis of $L^2([0,1])$. It can be checked that the related parameters can be estimated by procedure (2.7)-(2.12), after some minor modifications, as long as the direction in which h is constant is not perpendicular to the line edge.

2.3 Some generalizations

In the previous subsection, it is assumed that the true image f includes only one line edge which is surrounded by a uniform background in a region and parallel to the x-axis. In this subsection, we discuss several generalizations.

First, if the true image f includes only one line edge which forms an angle θ with the positive x-axis, then the estimation procedure (2.7)–(2.12) can be performed in a same way except that the averaged data $\overline{\mathbf{Z}}$ should be computed within equally spaced bands parallel to the line edge. Or, equivalently, the Cartesian coordinate system is rotated first such that the line edge is parallel to the x-axis after the rotation, and then the procedure (2.7)–(2.12) is applied to the rotated data.

Second, if the true image includes several line edges, each of which is surrounded by a uniform background in a region, then the psf h can be estimated using observations in all these regions as follows. For simplicity, let us assume that there are two such line edges, and the linear models corresponding to model (2.7) for describing averaged data in the two regions that contain the two line edges are

$$\overline{\mathbf{Z}}^{(j)} = (A^{(j)}\mathbf{B}^{(j)})\mathbf{c} + \overline{\varepsilon}^{(j)}, \text{ for } j = 1, 2.$$

For each j, we can obtain an estimator $\widehat{A^{(j)}}$ of $A^{(j)}$, as before. Then the LS estimator of $\mathbf{c}$ can be obtained from the combined linear model:

$$\begin{bmatrix} \overline{\mathbf{Z}}^{(1)} \\ \overline{\mathbf{Z}}^{(2)} \end{bmatrix} = \begin{bmatrix} \widehat{A^{(1)}}\mathbf{B}^{(1)} \\ \widehat{A^{(2)}}\mathbf{B}^{(2)} \end{bmatrix} \mathbf{c} + \begin{bmatrix} \overline{\varepsilon}^{(1)} \\ \overline{\varepsilon}^{(2)} \end{bmatrix}.$$

It can be checked that the LS estimator of $\mathbf{c}$ is

$$\widehat{\mathbf{c}} = \left(\widehat{A^{(1)}}^2 (\mathbf{B}^{(1)})'\mathbf{B}^{(1)} + \widehat{A^{(2)}}^2 (\mathbf{B}^{(2)})'\mathbf{B}^{(2)} \right)^{-1} \begin{bmatrix} \widehat{A^{(1)}}\overline{\mathbf{Z}}^{(1)} \\ \widehat{A^{(2)}}\overline{\mathbf{Z}}^{(2)} \end{bmatrix}. \tag{2.15}$$

Using expressions (2.15) and (2.6), the estimator of the psf h can be defined similarly to the one in equation (2.11), or, to its modified version in equation (2.12).

Third, if the true image f has a single "band" edge parallel to the x-axis and surrounded by a uniform background in a region, i.e., f is defined by

$$f(x, y) = \begin{cases} A, & \text{if } c_1 \le y \le c_2 \\ 0, & \text{otherwise,} \end{cases}$$

where $A > 0$ and $0 < c_1 < c_2 < 1$ are constants, then the estimation procedure (2.7)–(2.12) can be performed in a same way after functions $b_k^*(y - 0.5)$ in equation (2.5) are replaced by

$$b_k^{**}(y) = \int_{-1}^{1} \left(\int_{c_1}^{c_2} b_k(u^2 + (y - v)^2)\, dv \right) du.$$

Other generalizations are also possible. For instance, if the line edges are actually curves, then the above estimation procedure can be executed in a same way except that the data need to be averaged along the curved line edges. Another case that may be of interest is when the constant A in (2.1) depends on x. In such a case, height of the line edge changes with x; the procedure for estimating the psf h should be applied to the original data, instead of the averaged data.

3 Statistical Properties

We give some theoretical properties of the estimators discussed in the previous section, in the case when the true image f has expression (2.1) and the psf h is circularly symmetric. It can be checked that they have similar properties when h is confined to one specific direction that is not perpendicular to the line edge, and when f is in several more general cases discussed in Section 2.3.

Theorem 3.1 *Suppose that the true psf h is circularly symmetric density function and defined in a connected subset $\mathcal{D}$ of R^2 which includes the origin and has finite diameter, that h has first-order derivative at each point in $\mathcal{D}$, and that $\int\int_{\mathcal{D}} (h'(u,v))^2\, dudv < \infty$. The error terms in (2.2) are assumed to be i.i.d. with mean 0 and finite variance σ^2. Then, for estimator $\widehat{h}$ defined in equation (2.11) using the cosine-series basis as in (2.14), we have*

$$\int\int_{\mathcal{D}} \left[\widehat{h}(u,v) - h(u,v)\right]^2 dudv = O\left(J^{-2}\right) + o\left(Jn^{-2}[\log(J^{1/2}n^{-1})]^2\right),\ a.s. \tag{3.1}$$

and, for any $(x,y) \in \mathcal{D}$,

$$\lim_{n\to\infty} \frac{1}{\sigma_{n,\widehat{h}(x,y)}} \left[\widehat{h}(x,y) - h(x,y)\right] \stackrel{D}{=} N(0,1), \tag{3.2}$$

where

$$\begin{aligned} \sigma^2_{n,\widehat{h}(x,y)} &= \frac{\sigma^2}{nA^2} \left[\mathbf{b}'(x,y)(\mathbf{B}'\mathbf{B})^{-1}\mathbf{b}(x,y)\right], \\ \mathbf{b}(x,y) &= (b_0(x,y), b_1(x,y), \ldots, b_{J-1}(x,y))', \end{aligned}$$

and $\stackrel{D}{=}$ *denotes equality in distribution. For estimator* $\widehat{A}$*, we have*

$$\widehat{A} = A + O(J^{-1}) + o\left(J^{1/2}n^{-1}\log(J^{1/2}n^{-1})\right), \ a.s. \tag{3.3}$$

and

$$\lim_{n\to\infty} \frac{1}{\sigma_{n,\widehat{A}}}\left(\widehat{A} - A\right) \stackrel{D}{=} N(0,1), \tag{3.4}$$

where

$$\sigma^2_{n,\widehat{A}} = \frac{\sigma^2}{n}\left[\mathbf{b}'(\mathbf{B}'\mathbf{B})^{-1}\mathbf{b}\right].$$

Theorem 3.1 establishes the strong consistency and asymptotic normality of estimators $\widehat{h}$ and $\widehat{A}$. It can be checked that, when J is chosen to be $O(n^{2/3})$, the integrated squared error (ISE) of $\widehat{h}$ converges to zero almost surely with the rate $O(n^{-4/3}[\log(n)]^2)$, and $\widehat{A}$ converges to A almost surely with the rate $O(n^{-2/3}\log(n))$. These rates are faster than the corresponding optimal convergence rates in 1-D nonparametric regression (cf., Stone 1982), because our estimators are constructed from row averages $\overline{\mathbf{Z}}$ (cf., equations (2.8)–(2.10)), each component of which has variance σ^2/n. It should be pointed out that, in $\sigma^2_{n,\widehat{h}(x,y)}$ and $\sigma^2_{n,\widehat{A}}$ used in equations (3.2) and (3.4), there are unknown parameters A and σ involved. In applications, they can be replaced by their strong consistent estimators, and the corresponding results of asymptotic normality are still true, based on Slusky's theorem. An outline of the proof of Theorem 3.1 is given in Appendix A.

4 A Simulation Study

In this section, we present some simulation results regarding the numerical performance of the proposed method. The test image used is the one defined by equation (2.1) with $A = 1.0$. Its discrete version can be written as

$$f(x,y) = \begin{cases} n, & \text{if } y \in [0.5 - 1/(2n), 0.5 + 1/(2n)] \\ 0, & \text{otherwise,} \end{cases}$$

where n^2 is the size of the equally spaced observations $\{Z_{ij},\ i,j = 1,2,\ldots,n\}$ generated from model (2.2) at the design points $\{(x_i,y_j) = (i/n, j/n),\ i,j = 1,2,\ldots,n\}$. The error terms in model (2.2) are independent and identically distributed with distribution $N(0,\sigma^2)$. The following two psf

functions are considered:

$$h_1(x,y) = \frac{2}{\pi r^2}\left[1-(x^2+y^2)/r^2\right], \;\; h_2(x,y) = \frac{3}{\pi r^2}\left[1-\sqrt{x^2+y^2}/r\right], \;\; \text{for } x^2+y^2 \le r^2,$$

both of which are circularly symmetric functions. The true test image f, its blurred-and-noisy version by psf h_1 with $r = 0.1$ and $\sigma = 0.5$, and the two psf functions h_1 and h_2 are shown in Figure 4.1. From the figure, it can be seen that h_1 and h_2 are quite different in their shapes: h_1 is quite smooth in the middle while h_2 has a sharp angle in the middle.

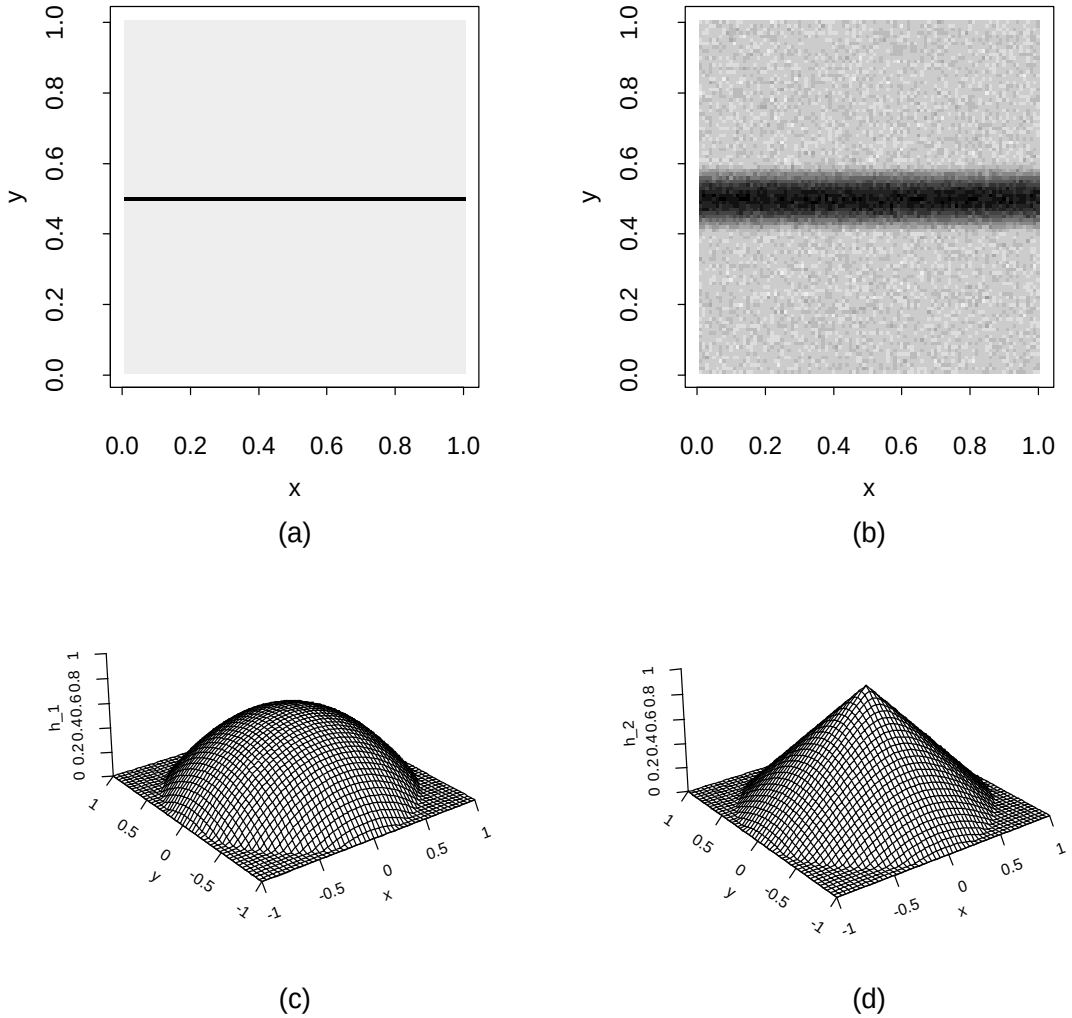

Figure 4.1: (a): Original test image. (b): A blurred-and-noisy version of the original image by psf h_1 with $r = 0.1$ and $\sigma = 0.5$. (c): True psf h_1. (d): True psf h_2.

Then, we let σ change among 0.1, 0.5, and 1.0; r change among 0.05, 0.1, and 0.2; and n change between 100 and 200. For each combination of σ, r, and n, we search for the optimal values of r and J used in procedure (2.7)–(2.12) by minimizing the Mean Integrated Squared Error (MISE)

of $\widehat{h}_M$, based on 100 replications. These optimal values of r and J along with the corresponding MISE values of $\widehat{h}_M$ and Mean Squared Error (MSE) values of $\widehat{A}$ are presented in Table 4.1. The optimal values of r and J are also approximated by the CV procedure (2.13) in each case. The selected r and J values by CV are presented in Table 4.1 in the last line of each entry.

From the table, it can be seen that (i) the optimal MISE value of $\widehat{h}_M$ decreases with the value of r, increases with the value of σ, and decreases with the value of n; (ii) MSE values of $\widehat{A}$ are generally very small; (iii) the optimal values of r selected by MISE equal the true r values in most cases; (iv) the optimal values of J selected by MISE increase with the values of r, decrease with the values of σ, and increase with the values of n; and (v) the selected r and J values by CV are generally close to the optimal values of r and J by MISE. The first result demonstrates that the psf h can be estimated more accurately when the value of r, which represents the blurring extent, is larger. This is reasonable because we have more information about blurring in such cases. The fourth result tells us that the parameter J should be chosen larger when the blurring extent is larger or when the noise level is lower or when the sample size is larger. That is because we can estimate the psf more accurately in these cases by using more terms in function approximation (cf. equations (2.6) and (2.11)). Other results mentioned above are intuitively reasonable.

Next, we demonstrate the proposed estimators of h and A (cf. equations (2.9) and (2.12)) in the case when $n = 100, \sigma = 0.5$, and $r = 0.1$. In Figure 4.2(a), the dotted and dashed curves denote the "typical" estimators of h_1 of the proposed estimation procedure (2.12) using the optimal values of r and J (i.e., $r = 0.1$ and $J = 6$) and their approximations by CV (i.e., $r = 0.11$ and $J = 5$), respectively. Here, the "typical" estimator is defined by the one with median Integrated Squared Error (ISE) among 100 replicated estimators of h_1. Because psf h_1 is circularly symmetric, all results about h_1 are shown only in the cross section of $y = 0$ and $-1 \leq x \leq 0$. In the plot, the solid curve is the true psf h_1 itself. It can be seen from the plot that the typical estimators perform reasonably well, although the estimator with parameters selected by CV is a little worse. The density curves of the estimators of A based on 100 replications are shown in Figure 4.2(b) by the dotted and dashed curves, respectively, in the two cases when r and J take their optimal values and when they are chosen by CV. Obviously, in both cases, the distribution of $\widehat{A}$ looks bell-shaped, and the distribution variability is quite small. The corresponding results when $h = h_2$ are shown in Figures 4.2(c) and 4.2(d). Similar conclusions can be made, except that the upper-right part of the cross section of h_2, corresponding to the angular peak of h_2 in 2-D setup, seems more challenging

Table 4.1: For each combination of σ, r, and n, this table presents the optimal values of r and J (in the third line of each entry), which minimize the MISE value of $\widehat{h}_M$, the corresponding MISE values of $\widehat{h}_M$ (in the first line) and MSE values of $\widehat{A}$ (in the second line), and the selected r and J values by CV (in the fourth line).

		h_1			h_2		
n	r	$\sigma = 0.1$	$\sigma = 0.5$	$\sigma = 1.0$	$\sigma = 0.1$	$\sigma = 0.5$	$\sigma = 1.0$
100	0.05	$8.2 * 10^{-4}$	$1.1 * 10^{-3}$	$2.0 * 10^{-3}$	$1.5 * 10^{-3}$	$1.5 * 10^{-3}$	$1.5 * 10^{-3}$
		$1.1 * 10^{-5}$	$1.9 * 10^{-5}$	$4.1 * 10^{-5}$	$3.5 * 10^{-4}$	$3.5 * 10^{-4}$	$3.6 * 10^{-4}$
		0.05,4	0.05,4	0.05,4	0.06,4	0.06,4	0.06,4
		0.07,4	0.07,4	0.07,4	0.06,3	0.06,3	0.06,3
	0.10	$3.2 * 10^{-4}$	$5.9 * 10^{-4}$	$7.7 * 10^{-4}$	$7.0 * 10^{-4}$	$8.3 * 10^{-4}$	$9.9 * 10^{-4}$
		$7.3 * 10^{-5}$	$6.8 * 10^{-6}$	$2.3 * 10^{-5}$	$6.1 * 10^{-6}$	$1.1 * 10^{-5}$	$2.2 * 10^{-5}$
		0.09,6	0.10,6	0.10,4	0.11,6	0.11,6	0.10,4
		0.11,7	0.11,5	0.11,3	0.11,5	0.11,5	0.11,5
	0.20	$1.5 * 10^{-4}$	$3.4 * 10^{-4}$	$6.9 * 10^{-4}$	$1.6 * 10^{-4}$	$4.9 * 10^{-4}$	$9.1 * 10^{-4}$
		$5.0 * 10^{-7}$	$1.0 * 10^{-5}$	$4.0 * 10^{-5}$	$3.9 * 10^{-7}$	$9.7 * 10^{-6}$	$3.9 * 10^{-5}$
		0.20,8	0.20,6	0.20,6	0.20,10	0.20,8	0.20,6
		0.22,7	0.21,5	0.21,4	0.21,11	0.21,9	0.21,7
200	0.05	$3.8 * 10^{-4}$	$4.3 * 10^{-4}$	$5.2 * 10^{-4}$	$7.3 * 10^{-4}$	$7.4 * 10^{-4}$	$7.4 * 10^{-4}$
		$1.1 * 10^{-6}$	$2.2 * 10^{-6}$	$5.4 * 10^{-6}$	$1.6 * 10^{-6}$	$2.3 * 10^{-6}$	$4.9 * 10^{-6}$
		0.05,6	0.05,6	0.05,6	0.05,6	0.05,6	0.05,6
		0.06,6	0.06,6	0.06,5	0.04,6	0.05,4	0.05,4
	0.10	$5.3 * 10^{-5}$	$1.9 * 10^{-4}$	$2.9 * 10^{-4}$	$1.5 * 10^{-4}$	$2.2 * 10^{-4}$	$3.6 * 10^{-4}$
		$1.0 * 10^{-7}$	$1.7 * 10^{-6}$	$6.4 * 10^{-6}$	$6.7 * 10^{-8}$	$1.6 * 10^{-6}$	$6.3 * 10^{-6}$
		0.10,12	0.10,8	0.10,8	0.10,10	0.10,10	0.10,8
		0.11,10	0.11,7	0.11,8	0.10,10	0.10,8	0.10,8
	0.20	$5.4 * 10^{-5}$	$2.0 * 10^{-4}$	$4.0 * 10^{-4}$	$6.6 * 10^{-5}$	$2.6 * 10^{-4}$	$5.5 * 10^{-4}$
		$1.3 * 10^{-7}$	$3.2 * 10^{-6}$	$1.3 * 10^{-5}$	$1.3 * 10^{-7}$	$3.2 * 10^{-6}$	$1.3 * 10^{-5}$
		0.20,18	0.20,10	0.20,6	0.20,18	0.20,10	0.20,8
		0.21,15	0.21,11	0.21,5	0.20,18	0.21,11	0.21,7

to estimate, which is true in general for most smoothing techniques not designed for preserving angles and other singularities (see e.g., Härdle 1991, Chapters 5 and 6).

5 Two Image Restoration Examples

In the previous sections, we discussed how to estimate the psf h nonparametrically. After h is estimated, the true image f can be restored from the observed-and-degraded image using various image restoration procedures designed for cases when h is assumed known, which are introduced in Section 1. In this section, we demonstrate the entire image restoration process by two examples. One is related to handwritten text recognition, and the other one is about image restoration when a test pattern is available.

5.1 An example related to handwritten text recognition

Machine recognition of handwriting has practical significance, as in reading postal addresses on envelopes, amounts in bank checks, and so forth. Image restoration techniques, such as the one proposed in this paper, can be used for preprocessing the scanned documents prior to recognizing text (cf., Plamondon and Srihari 2000, Section 4.1). Assume that the observed image is the one presented in Figure 5.1(a), which is generated by the convolution of the text image of the words "line edge" and the psf h_1, defined in Section 4, and additive, independent and identically distributed (i.i.d.) noise with common distribution $N(0, 0.25^2)$. The true text image is of size 400×400, its design space is normalized to be $[0,1] \times [0,1]$, its text pixels have gray levels of 60, and its background pixels have gray levels of 0. In the psf h_1, the true radius r is fixed at 0.05.

We first apply the proposed procedure (2.7)–(2.12) for estimating the true psf h_1. For this purpose, a region consisting of a line edge and a uniform background should be chosen. Among several choices, the one consisting of 60×60 pixels with the letter "I" in the middle is selected. By the CV procedure (2.13), the parameters (r, J) are chosen to be $(0.053, 5)$ in the normalized psf estimator $\widehat{h}_M$ (cf., equation (2.12)). The resulting $\widehat{h}_M$ in the cross section of $y = 0$ and $-1 \le x \le 0$ is shown in Figure 5.2 by the dotted curve. After h is estimated, the following Wiener filter is used for restoring the true image:

$$\widehat{f}(x,y) = \frac{1}{(2\pi)^2}\mathcal{R}\left\{\int\int \frac{\overline{\mathcal{F}\{\widehat{h}_M\}(s,t)}}{|\mathcal{F}\{\widehat{h}_M\}(s,t)|^2 + \alpha(s^2+t^2)^{\beta/2}}\mathcal{F}\{Z\}(s,t)\exp\{i(sx+ty)\}\,dsdt\right\}, \quad (5.1)$$

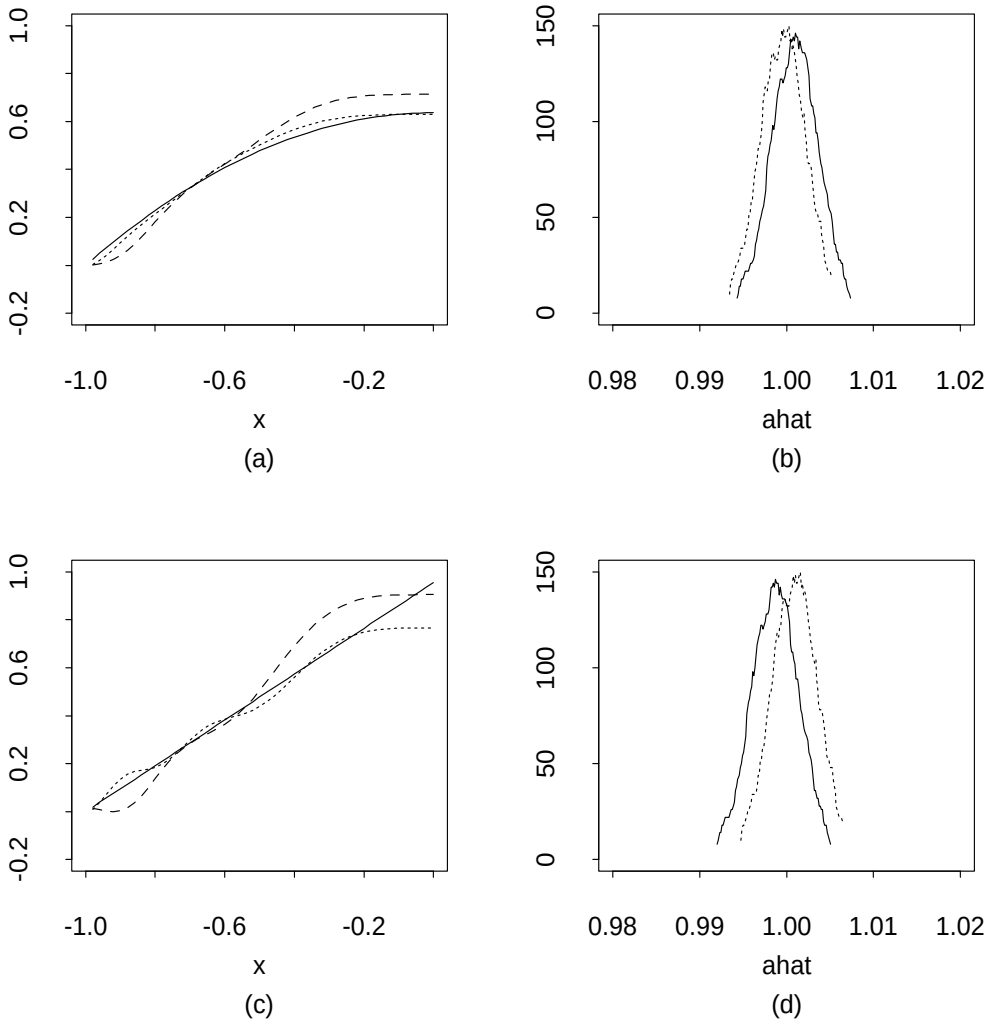

Figure 4.2: (a): Solid, dotted, and dashed curves represent h_1, its estimator with optimal values of r and J, and the estimator with r and J chosen by CV, respectively, in the cross section of $y = 0$ and $-1 \leq x \leq 0$. (b): Solid and dotted curves denote the density curves of $\widehat{A}$ based on 100 replications, when r and J take their optimal values and when they are chosen by CV, respectively. (c) and (d): Corresponding results when $h = h_2$. In plots (a) and (c), the presented estimator is the one with median Integrated Squared Error (ISE) among 100 replications.

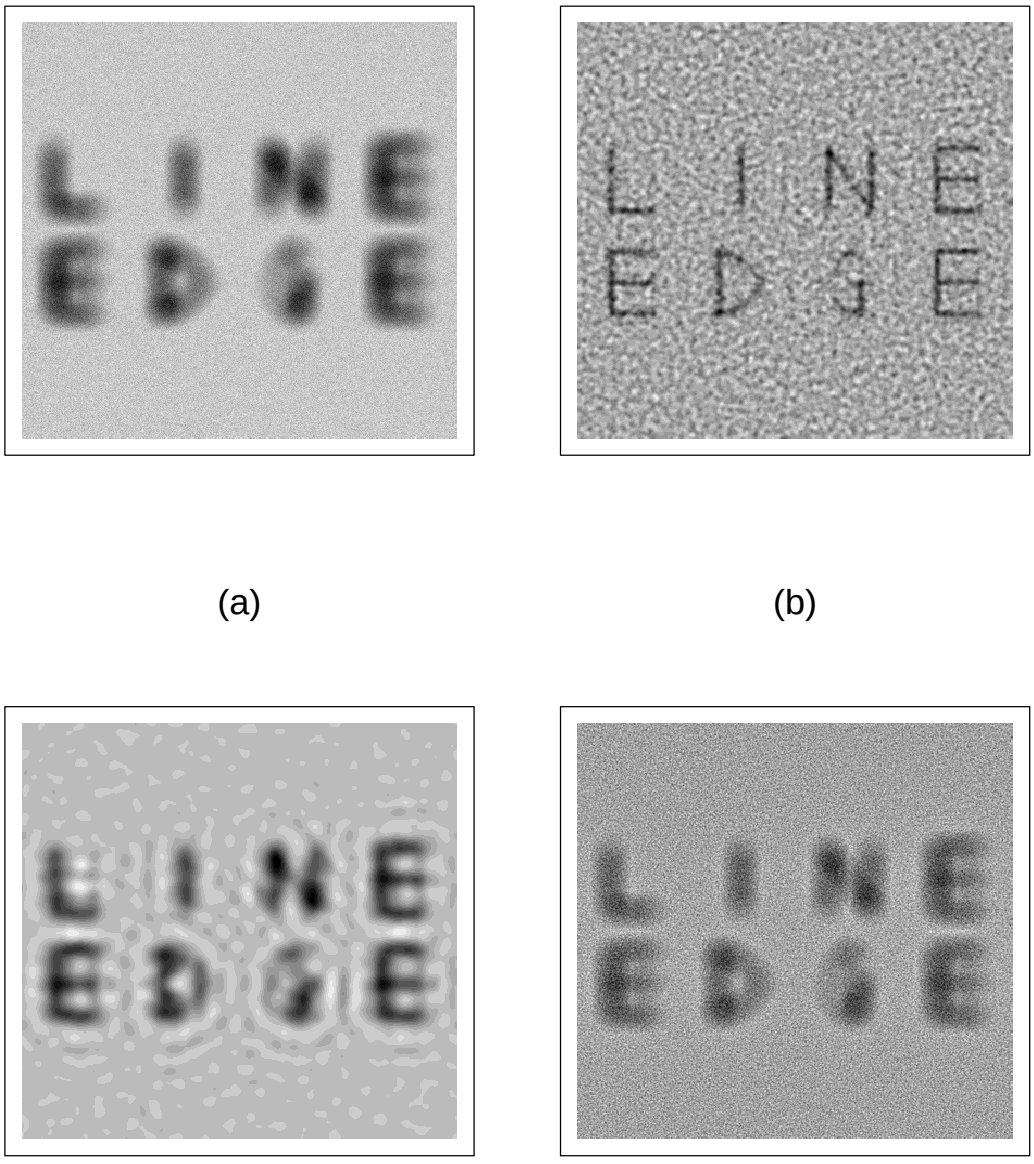

Figure 5.1: (a): Observed blurred-and-noisy image. (b): Restored image by the Wiener filter using the proposed estimator of the psf. (c): Restored image by the Wiener filter using the Gaussian psf. (d): Restored image by Carasso's (2001) blind image restoration procedure.

where $\overline{\mathcal{F}\{\widehat{h}_M\}(s,t)}$ denotes the complex conjugate of $\mathcal{F}\{\widehat{h}_M\}(s,t)$, $\mathcal{R}\{C\}$ denotes the real part of the complex number C, and $\alpha, \beta > 0$ are two parameters. The Wiener filter was proved to be optimal in minimizing the MISE of the restored image when the true psf was assumed known and when the noise was assumed Gaussian, which corresponds to $\beta = 1$ in equation (5.1) (cf., Gonzalez and Woods 1992, Chapter 5). The restored image when $\alpha = 2 \times 10^{-11}$ and $\beta = 1$ is shown in Figure 5.1(b). The above α value is chosen for good visual impression. When α is chosen smaller, the restored text would look sharper, but the background would be noisier, and vice versa.

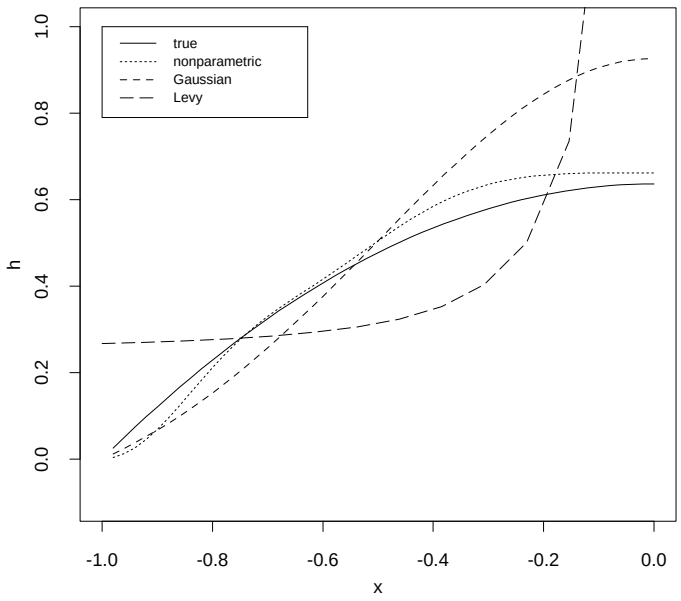

Figure 5.2: Solid curve denotes the true psf. Dotted, dashed and long-dashed curves denote estimated psfs by the nonparametric procedure (2.7)–(2.12), Gaussian approximation, and Carasso's (2001) APEX method using the symmetric Lévy "stable" density family.

In practice, some people would assume that the true psf is Gaussian, which is reasonable in some applications, including some remote sensing and aerial imaging applications. The restored image using the Wiener filter (5.1) and the Gaussian psf with scale parameter $r/2 = 0.025$ is shown in Figure 5.1(c), when $\alpha = 10^{-9}$ and $\beta = 1$. The Gaussian psf is presented in Figure 5.2 by the dashed curve. Its scale parameter is chosen such that its blurring extent matches that of the true psf. From Figure 5.1(c), it can be seen that the restored text is not sharp enough, compared to that in Figure 5.1(b), and the restored background already varies a lot in such a case, especially around the text, which implies that the restored image can not be improved much by changing the value of α in (5.1).

As mentioned in Section 1, there are a few blind image deblurring procedures in the literature, which impose various assumptions on the true psf. We take a different approach in the proposed procedure by avoiding restrictive assumptions on the psf so that it can be used in more applications. However, because of the ill-posed nature of the image restoration problem, we require that the observed image of the imaging device under study includes at least one line edge surrounded by a uniform background in a region, which is not difficult to satisfy in some applications, as discussed in Section 1. Due to their different natures, it may not be appropriate to compare the proposed procedure with these existing blind image deblurring procedures. To further demonstrate this point, next, we apply the blind image deblurring procedure suggested by Carasso (2001) to the degraded image shown in Figure 5.1(a), to investigate the performance of this procedure when its assumptions on the psf are not satisfied. By this approach, the true psf is assumed to be in the symmetric Lévy "stable" density family whose Fourier transformation is defined by

$$\mathcal{F}\{h\}(u,v) = \exp\left\{-\xi(u^2+v^2)^\eta\right\}, \text{ for } (u,v) \in R^2,$$

where $\xi > 0$ and $0 < \eta \leq 1$ are two parameters. To determine the two parameters, Carasso suggested using the so-called APEX method, which chooses ξ and η such that

$$\sum_{u=-n/2}^{n/2} \left[\log\left(\mathcal{F}\{Z\}(u,0)\right) - \left(-\xi|u|^{2\eta} - A\right)\right]^2$$

is minimized, where $2 \leq A \leq 6$ is a constant. By this method, we choose $\xi = 0.4$, $\eta = 0.2$, and $A = 5$. The corresponding psf is shown in Figure 5.2 by the long-dashed curve. Obviously, the approximated psf from the symmetric Lévy "stable" density family is far away from the true psf. After the psf h is specified, by Carasso's procedure, the Fourier transformation of the true image is estimated by

$$\widehat{\mathcal{F}\{f\}}(u,v) = \frac{\overline{\mathcal{F}\{h\}(u,v)}\mathcal{F}\{Z\}(u,v)}{|\mathcal{F}\{h\}(u,v)|^2 + (\epsilon/M)^2 + K^{-2}|1-\mathcal{F}\{h\}^s(u,v)|^2},$$

where ϵ, M, K, and s are constants. By Carasso's suggestions, we choose $\epsilon = 0$, $M = 1$, and $s = 0.005$. The restored image when $K = 0.2$ is shown in Figure 5.1(d), which gives the best visual impression when K changes. We can see that it does not improve much, compared to the observed image, which implies that Carasso's procedure may not be appropriate to use in this example.

5.2 Restoration of the bird image

The bird image shown in Figure 5.3(a) is a standard test image in the literature. The one used here has 256×256 pixels, and its image gray levels are in the range $[0, 255]$. Assume that there are two observed bird images. The one shown in Figure 5.3(b) is a blurred image without pointwise noise, which is the convolution of the original bird image and the psf h_1 used in the previous subsection with $r = 0.05$. The one shown in Figure 5.3(c) is a blurred-and-noisy image, obtained by adding i.i.d. pointwise noise with distribution $N(0, 5^2)$ to the blurred image. For each observed bird image, we further assume that there is a corresponding observed test image having the same setup, except that the true test image is the one shown in Figure 4.1(a).

To restore the true bird image from its blurred version shown in Figure 5.3(b), we can first estimate the true psf from the corresponding observed test image, using the proposed procedure (2.7)–(2.12), and then restore the true image using the Wiener filter (5.1), as discussed in the previous subsection. In procedure (2.9)–(2.14), the two parameters (r, J) should be determined in advance, which can be accomplished in two ways. In image processing practice, such parameters are often determined by repeated experiments using different parameter values, and the ones giving best visual impression can be selected. In this paper, we suggest choosing (r, J) by the CV procedure (2.13). Of course, different people may obtain different results using the first approach. To make it more objective, here, we replace it by the one which searches for the optimal parameter values by minimizing MISE of the estimated psf, as we did in Table 4.1. Although the MISE approach requires information about the true psf and thus can not be used in real applications, its results are the best possible ones by the repeated experiments approach, and it is used here just for demonstrating the repeated experiments approach.

By the MISE and CV procedures, the parameters (r, J) are chosen to be $(0.05, 6)$ and $(0.053, 6)$, respectively, and their estimated psf's are shown in Figure 5.3(f), along with the true psf. Using these results, the restored images by the Wiener filter are shown in Figure 5.3(d) using the parameter values chosen by MISE, and in Figure 5.3(e) using the parameter values chosen by CV. In the Wiener filter, β is fixed at 1 and α is chosen to be 10^{-10} in both cases, which gives the best visual impression, as we did in Figure 5.1. The corresponding results from the observed blurred-and-noisy bird image are shown in Figures 5.3(g)–(i).

From Figure 5.3, it can be seen that (i) the MISE and CV approaches give similar results, as

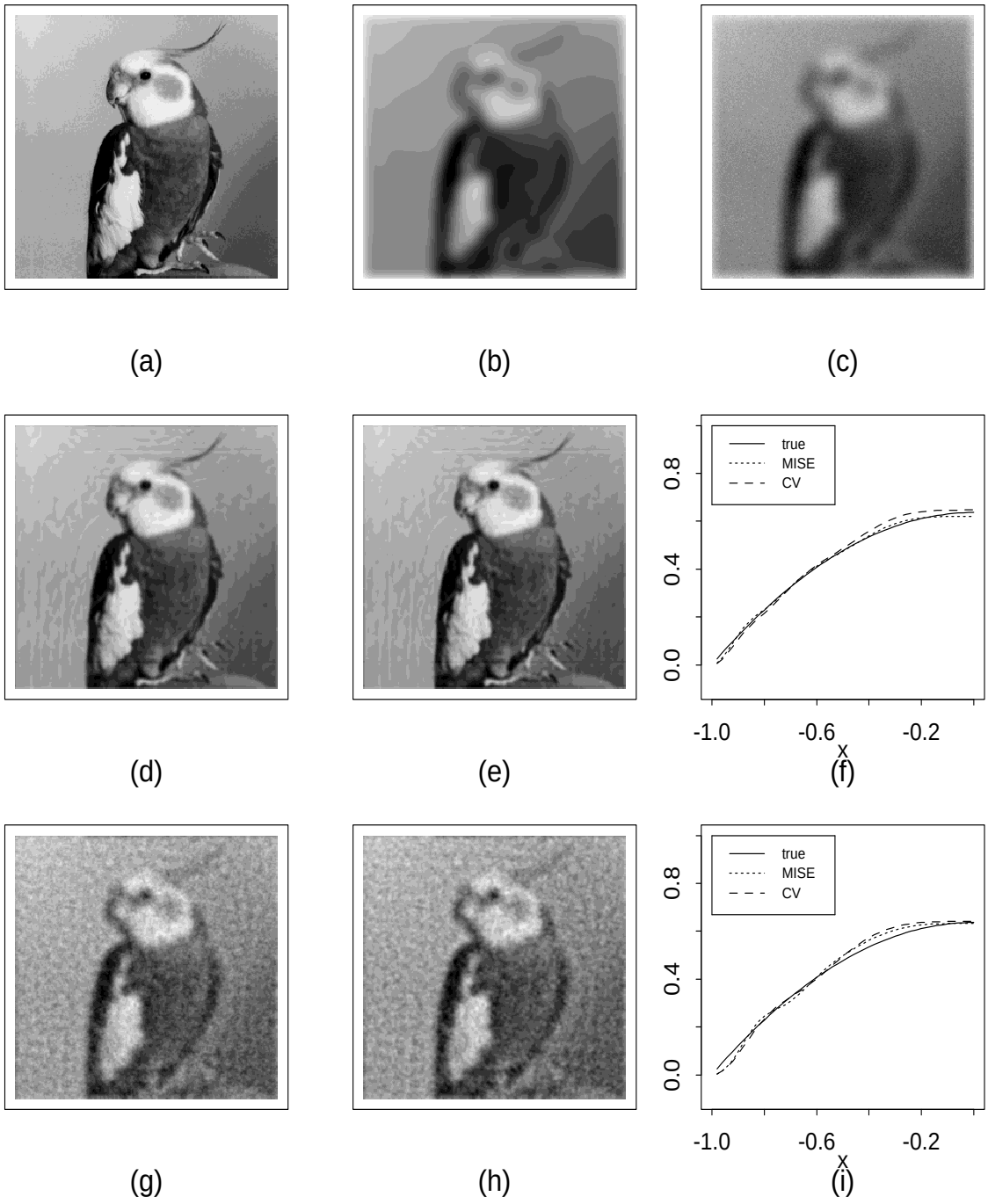

Figure 5.3: (a): True bird image. (b): Observed blurred image. (c): Observed blurred-and-noisy image. (d): Restored image from plot (b), using parameter values chosen by MISE. (e): Restored image from plot (b), using parameter values chosen by CV. (f): True psf, estimated psf using parameters chosen by MISE, and estimated psf using parameters chosen by CV, from the corresponding test image of plot (b). (g)–(i): Corresponding results to those in plots (d)–(f), from the observed blurred-and-noisy image.

demonstrated by Table 4.1 for estimating the psf, (ii) the estimated psf by procedure (2.7)–(2.12) is not sensitive to noise level (cf., plots (f) and (i)), and (iii) the restored image by the Wiener procedure is quite sensitive to noise after the psf is determined. Based on our experience, the third conclusion is also true for other image restoration procedures assuming known psf. Note that the noise level in plot (c) is quite high. When the observed bird image has the same amount of blurring as in plots (b)–(c) and its noise level is between that in plot (b) and that in plot (c), then the restored images should perform between those in plots (d)–(e) and those in plots (g)–(h).

6 Concluding Remarks

We discuss a nonparametric procedure for estimating the psf for blind image restoration, when the true image includes a portion with a line edge surrounded by a uniform background, or, when there is an observed test image having one or more line edges. This procedure can be applied to many applications, including handwritten text recognition and calibration of imaging devices. Numerical results show that it is robust to noise. Together with some existing image restoration procedures assuming known psf, it can restore a true image well from its observed but degraded version.

However, to use the current version of the proposed deblurring procedure, we need to identify a region that includes a line edge surrounded by a uniform background manually in the handwritten text example discussed in Section 5.1. In applications, it might be possible to make this identification process automatic, by first detecting edges using certain edge detectors (cf., e.g., Qiu and Yandell 1997, Qiu 2002, Sun and Qiu 2007) and then automatically figuring out a region with a single detected edge segment. However, detected edge pixels by most existing edge detectors are scattered in the whole design space and they may not form closed edge curves. Therefore, after edge detection, we might need to link detected edge pixels using certain edge linking techniques (cf., Section 6.6.1, Qiu 2005). We may also need to estimate true background intensities and then correct the background so that the corrected background is roughly uniform. It requires much future research to put all these pieces together and make the entire procedure convenient to use.

As discussed in Section 1, the deblurring problem is notoriously challenging. The current deblurring procedure depends on the availability of a test image with one or more line edges surrounded by a uniform background or a portion of the true image with these properties. In certain applications, it might be difficult to obtain a test image or an observed image with such

features. Also, in this paper, we assume that the blurring mechanism is homogeneous in the entire image (i.e., the psf h in model (1.1) does not depend on spatial location). In certain applications, this assumption may not be valid. Image deblurring in these more general cases requires much future research as well.

Acknowledgments: We thank the associate editor and two referees for some helpful comments and suggestions which greatly improve the quality of the paper. This research is supported in part by an NSF grant.

Appendix

A Outline of the Proof of Theorem 3.1

Because estimator $\widehat{A}$ is involved in the definition of estimator $\widehat{h}$, we prove equations (3.3) and (3.4) about estimator $\widehat{A}$ first. By the approximation theory of trigonometric polynomials (cf., e.g., Girosi and Anzellotti 1992), we have

$$\int_0^1 [\xi_J(u) - \xi(u)]^2 \ du = O(J^{-2}),$$

where $\xi_J(u) = \sum_{k=0}^{J-1} c_k \cos(k\pi u)$, and $\xi(u)$ is assumed to have square integrable first-order derivatives in $[0,1]$. By this result, we can draw the following conclusion:

$$\int\int_{\mathcal{D}} [h_J(u,v) - h(u,v)]^2 \ dudv = O(J^{-2}), \tag{A.1}$$

as long as the pdf h has first-order derivatives in $\mathcal{D}$ and $\int\int_{\mathcal{D}}[h'(u,v)]^2 \ dudv < \infty$.

By equations (2.7) and (2.9), it can be checked that

$$\widehat{A} = A(\mathbf{b}'\mathbf{c}) + \mathbf{b}' \left(\mathbf{B}'\mathbf{B}\right)^{-1} \mathbf{B}'\overline{\boldsymbol{\varepsilon}}. \tag{A.2}$$

Since the psf h is a 2-D density function on $\mathcal{D}$ and h_J has the property of (A.1) (cf., equations (2.3) and (2.4)), we have

$$\mathbf{b}'\mathbf{c} = 1 + O(J^{-1}). \tag{A.3}$$

Therefore, the first term on the right-hand-side of equation (A.2) is $A + O(J^{-1})$. The second term is a linear combination of i.i.d. random variables with mean 0 and variance σ^2/n. By results in Theorem 1 of Lai and Robbins (1977), we know that it is of order $o(J^{1/2}n^{-1}\log(J^{1/2}n^{-1}))$, *a.s.*.

So, equation (3.3) is obtained by this result and equations (A.2) and (A.3). Since $\widehat{A}$ is a linear combination of the observed image intensities, we also have equation (3.4).

To prove equation (3.1), we first notice that

$$\int\int_{\mathcal{D}} \left[\widehat{h}(u,v) - h(u,v)\right]^2 \, dudv \leq 2\int\int_{\mathcal{D}} \left[\widehat{h}(u,v) - h_J(u,v)\right]^2 \, dudv + 2\int\int_{\mathcal{D}} \left[h_J(u,v) - h(u,v)\right]^2 \, dudv. \tag{A.4}$$

By equation (A.1), the second term on the right-hand-side of (A.4) equals $O(J^{-2})$. Regarding the first term, it can be checked that, for any $(x,y) \in \mathcal{D}$,

$$\left[\widehat{h}(x,y) - h_J(x,y)\right]^2 = \mathbf{b}'(x,y)(\widehat{\mathbf{c}} - \mathbf{c})(\widehat{\mathbf{c}} - \mathbf{c})'\mathbf{b}(x,y). \tag{A.5}$$

By equations (2.8)–(2.10), we have

$$(\widehat{\mathbf{c}} - \mathbf{c}) = (\widehat{\mathbf{c}}_A - \mathbf{c})A/\widehat{A} + \mathbf{c}(A/\widehat{A} - 1). \tag{A.6}$$

From equations (2.7) and (2.8), we have $\widehat{\mathbf{c}}_A = c + \frac{1}{A}\left(\mathbf{B}'\mathbf{B}\right)^{-1}\mathbf{B}'\overline{\varepsilon}$. So, again, based on the results in Theorem 1 of Lai and Robbins (1977), we have $(\widehat{\mathbf{c}}_A - \mathbf{c}) = o\left(n^{-1}\log(n)\right)$, *a.s.*. By combining this result and equation (3.3), the first term on the right-hand-side of (A.6) is of order $o\left(n^{-1}\log(n)\right)$ almost surely, and the second term is of order $O(J^{-1}) + o(J^{1/2}n^{-1}\log(J^{1/2}n^{-1}))$ almost surely. Therefore, by this result and equations (A.4)-(A.6), we have equation (3.1). Equation (3.2) can be obtained by the Slusky's theorem based on the fact that (i) $\widehat{h}(x,y)\widehat{A}/A$ is a linear combination of the observations (cf., equations (2.9)–(2.11)), and (ii) $\widehat{A}/A = 1 + o(1)$, *a.s.*.

References

Andrews, H.C., and Hunt, B.R. (1977), *Digital Image Restoration,* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Arica, N., and Yarman-Vural, F.T. (2001), "An overview of character recognition focused on off-line handwriting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 216–233.

Bates, R.H.T., and McDonnell, M.J. (1986), *Image Restoration and Reconstruction,* Oxford: Clarendon Press.

Besag, J. (1986), "On the statistical analysis of dirty pictures (with discussion)," *Journal of the Royal Statistical Society (Series B)*, **48**, 259–302.

Cannon, M. (1976), "Blind deconvolution of spatially invariant image blurs with phase," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **24**, 58–63.

Carasso, A.S. (2001), "Direct blind deconvolution," *SIAM Journal on Applied Mathematics*, **61**, 1980–2007.

Figueiredo, M.A.T., and Nowak, R.D. (2003), "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, **12**, 906–916.

Geman, S., and Geman, D. (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.

Gijbels, I., Lambert, A., and Qiu, P. (2006), "Edge-preserving image denoising and estimation of discontinuous surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**, 1075–1087.

Girosi, F., and Anzellotti, G. (1992), "Rates of convergence of approximation by translates," *A.I. Memo 1288*, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

Gonzalez, R.C., and Woods, R.E. (1992), *Digital Image Processing*, New York: Addison-Wesley Publishing Company.

Goodman, J.W. (1968), *Introduction to Fourier Optics*, New York: McGraw-Hill Book Co.

Hall, P., and Qiu, P. (2007a), "Blind deconvolution and deblurring in image analysis," *Statistica Sinica*, **17**, 1483–1509.

Hall, P., and Qiu, P. (2007b), "Nonparametric estimation of a point spread function in multivariate problems," *The Annals of Statistics*, **35**, 1512–1534.

Hankerson, D., Harris, G.A., and Johnson, P.D., Jr. (1998), *Introduction to Information Theory and Data Compression,* New York: CRC Press.

Härdle, W. (1991), *Smoothing Techniques: With Implementation in S*, New York: Springer Verlag.

Hillebrand, M., and Müller, Ch.H. (2007), "Outlier robust corner-preserving methods for reconstructing noisy images," *The Annals of Statistics*, **35**, 132–165.

Joshi, M.V., and Chaudhuri, S. (2005), "Joint blind restoration and surface recovery in photometric stereo," *Journal of the Optical Society of America, Series A*, **22**, 1066–1076.

Katsaggelos, A.K., and Lay, K.-T. (1990), "Image identification and image restoration based on the expectation-maximization algorithm," *Optical Engineering*, **29**, 436–445.

Koren, N. (2004), Tutorials about the software package *Imatest*, http://www.normankoren.com/Tutorials/MTF5.html.

Kundur, D., and Hatzinakos, D. (1998), "A novel blind deconvolution scheme for image restoration using recursive filtering," *IEEE Transactions on Signal Processing*, **46**, 375–390.

Lai, T.L., and Robbins, H. (1977), "Strong consistency of least squares estimators in regression models," *Proceedings of the National Academy of Sciences of the USA*, **74**, 2667–2669.

Li, S.Z. (1995), *Markov random field modeling in computer vision,* New York: Springer-Verlag.

Loader, C.R. (1999), "Bandwidth selection: classical or plug-in?," *The Annals of Statistics*, **27**, 415–438.

Perona, P., and Malik, J. (1990), "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 629–639.

Plamondon, R., and Srihari, S.N. (2000), "On-line and off-line handwriting recognition: a comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, 63–83.

Polzehl, J., and Spokoiny, V.G. (2000), "Adaptive weights smoothing with applications to image restoration," *Journal of Royal Statistical Society (Series B)*, **62**, 335–354.

Polzehl, J., and Spokoiny, V.G. (2003), "Image denoising: Pointwise adaptive approach," *The Annals of Statistics*, **31**, 30–57.

Qiu, P. (1998), "Discontinuous regression surfaces fitting," *The Annals of Statistics*, **26**, 2218–2245.

Qiu, P. (2002), "A nonparametric procedure to detect jumps in regression surfaces," *Journal of Computational and Graphical Statistics*, **11**, 799–822.

Qiu, P. (2004), "The local piecewisely linear kernel smoothing procedure for fitting jump regression surfaces," *Technometrics*, **46**, 87–98.

Qiu, P. (2005), *Image Processing and Jump Regression Analysis,* New York: John Wiley & Sons.

Qiu, P. (2007), "Jump surface estimation, edge detection, and image restoration," *Journal of the American Statistical Association*, **102**, 745–756.

Qiu, P., and Yandell, B. (1997), "Jump detection in regression surfaces," *Journal of Computational and Graphical Statistics*, **6**, 332–354.

Rosenfeld, A., and Kak, A.C. (1982), *Digital Picture Processing (2nd edition),* New York: Academic Press.

Saint-Marc, P., Chen, J., and Medioni, G. (1991), "Adaptive smoothing: a general tool for early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, 514–529.

Skilling, J. (1989, eds), *Maximum Entropy and Bayesian Methods,* Norwell, MA: Kluwer Academic.

Stone, C.J. (1982), "Optimal global rates of convergence for nonparametric regression," *The Annals of Statistics*, **10**, 1040–1053.

Sun, J., and Qiu, P. (2007), "Jump detection in regression surfaces using both first-order and second-order derivatives," *Journal of Computational and Graphical Statistics*, **16**, 289–311.

Tomasi, C., and Manduchi, R. (1998), "Bilateral Filtering for Gray and Color Images," *Proceedings of the 1998 IEEE International Conference on Computer Vision*, 839–846, Bombay, India.

Wang, Y. (1998), "Change curve estimation via wavelets," *Journal of the American Statistical Association*, **93**, 163–172.

Yang, Y., Galatsanos, N.P., and Stark, H. (1994), "Projection-based blind deconvolution," *Journal of the Optical Society of America A*, **11**, 2401–2409.