IEEE SP Workshop on Higher-Order Statistics, Banff, Canada, July 21-23, 1997, pp. 108-111

Deblurring Two-Tone Images by A Joint Estimation Approach Using Higher-Order Statistics

Ta-Hsin Li* and Ke-Shin Lii†
*Department of Statistics, Texas A&M University, College Station, TX 77843
†Department of Statistics, University of California, Riverside, CA 92521

Abstract

A method is proposed for the restoration of linearly blurred two-tone images without requiring the knowledge of the blur parameters. The method jointly estimates the original image and the blur parameters based on some statistical parameters at the output of an inverse filter. Unlike some other blind image restoration procedures, the proposed method does not require the estimation or modeling of the statistical properties of the original image, yet can be justified even for non-i.i.d. images.

1. Introduction

In typical automatic recognition and identification systems, the input images are usually assumed to be blur-free. Sometimes, blurred images can be encountered due to the impairment of imaging systems and environment. If the blur is not properly removed, the recognizer's performance can severely deteriorate. For example, if the blurred text image in Fig. 1 is segmented, without deblurring, for automatic character recognition, the result could become unrecognizable even to trained human eyes (see Fig. 2).

In this article we are concerned with the restoration of linearly blurred two-tone images such as texts and bar codes [1]. We propose a restoration method that requires no prior knowledge of the blur parameters or the statistical properties of the original images. The method is shown to be able to handle not only images with *i.i.d.* (independent and identically distributed) pixels but also images with non-*i.i.d.* pixels that are often encountered in reality. Due to the use of higher-order statistics [2], the method is also capable of handling minimum as well as non-minimum phase blur filters. These features are very desirable for image restoration problems.

This work is an extension of some previous works on blind deconvolution (or equalization) of 1-D signals in the digital communications context [4], [5]. One of the major differences is that unlike the digital communications problem where the alphabets of the transmitted

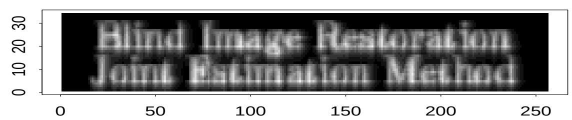

Figure 1. A blurred text image.

digital signal are known *a priori* at the receiver, the two tones of the original image in the image deblurring problem usually are not available. The proposed method in this article is able to take advantage of the two-tone information without requiring knowledge of the specific values of the tones – they are estimated from the blurred image jointly with the blur filter and the original image.

2. Deblurring Method

For linear degradations, the blurred image, denoted by $\mathcal{Y} := \{Y(m,n)\}$, can be regarded as the output from a blur filter

$$\mathcal{Y} = \mathcal{B} \otimes \mathcal{X},$$

where $\mathcal{B} := \{b(j,k)\}$ is the point-spread function (PSF) of the blur filter, $\mathcal{X} := \{X(m,n)\}$ is the original image, and $\otimes$ stands for two-dimensional convolution.

To deblur the image $\mathcal{Y}$ when the blur filter $\mathcal{B}$ is not available, we filter $\mathcal{Y}$ with an inverse filter $\mathcal{F} := \{f(j,k)\}$ to obtain $\mathcal{Z} := \{Z(m,n)\}$, where

$$\mathcal{Z} := \mathcal{F} \otimes \mathcal{Y}.$$

In this article, we propose the following *cost function* for the selection of $\mathcal{F}$:

$$J(\mathcal{F}) := \min_{a_2 = 1} \boldsymbol{a}^T H(\mathcal{Z})\, \boldsymbol{a},$$

where $\boldsymbol{a} := [a_0, a_1, a_2]^T$ is a constant vector and $H(\mathcal{Z})$ is a three-by-three matrix

$$H(\mathcal{Z}) := \begin{bmatrix} 1 & \mu_1(\mathcal{Z}) & \mu_2(\mathcal{Z}) \\ \mu_1(\mathcal{Z}) & \mu_2(\mathcal{Z}) & \mu_3(\mathcal{Z}) \\ \mu_2(\mathcal{Z}) & \mu_3(\mathcal{Z}) & \mu_4(\mathcal{Z}) \end{bmatrix},$$

with $\mu_i(\mathcal{Z})$, defined by

$$\mu_i(\mathcal{Z}) := E\{Z^i(m,n)\},$$

being the ith statistical moment of $\mathcal{Z}$.

3. Does the Method Work?

To see why minimizing $J(\mathcal{F})$ leads to a reasonable solution, it is instrumental to note that

$$\boldsymbol{a}^T H(\mathcal{Z})\,\boldsymbol{a} = E\{|\psi(Z^i(m,n))|^2\},$$

where $\psi(Z)$ is a two-degree polynomial defined by

$$\psi(Z) := a_0 + a_1 Z + a_2 Z^2.$$

Let β_1 and β_2 represent the zeros of $\psi(Z)$ so that

$$\psi(Z) = a_2\,(Z-\beta_1)\,(Z-\beta_2).$$

Then, it easy to see that if $\mathcal{F}$ equals the inverse of $\mathcal{B}$ or any of its shifted and/or scaled versions, the output image $\mathcal{Z}$ would coincide with $\mathcal{X}$ up to a shift and/or a scale and thus become a two-tone image. With β_1 and β_2 representing the two tones of $\mathcal{Z}$, one would obtain $\psi(Z(m,n)) = 0$, and therefore $J(\mathcal{F})$ would attain its minimum value zero.

On the other hand, if $J(\mathcal{F}) = 0$ for some $\mathcal{F} \neq 0$, then $\psi(Z(m,n)) = 0$ for some $\beta_1 \neq \beta_2$, so that $\mathcal{Z}$ must be a two-tone image. Therefore, for the justification of the method, it suffices to ensure that the two-tone image $\mathcal{Z}$ so obtained can only be a shifted and/or scaled version of the original two-tone image $\mathcal{X}$. Although one should not expect this for every two-tone image $\mathcal{X}$, it can be shown that it is true for a large class of two-tone images [3].

For example, if the pixels of $\mathcal{X}$ are *i.i.d.*, then $\mathcal{Z}$ must be a shifted and/or scaled version of $\mathcal{X}$ as soon as it becomes a two-tone image as a result of the filtering. To see this, we first note that

$$\mathcal{Z} = \mathcal{T} \otimes \mathcal{X},$$

with $\mathcal{T} := \{t(j,k)\} := \mathcal{F} \otimes \mathcal{B}$ being the combination of the deblurring filter with the blur filter. We proceed with the justification by the method of contradiction. Suppose that $\mathcal{Z}$ is a two-tone image but $\mathcal{T}$ is not a shifted and/or scaled delta function; for simplicity, let us assume that $t_0 := t(0,0) \neq 0$ and $t_1 := t(1,1) \neq 0$. Furthermore, without loss of generality, let us assume that $\mathcal{X}$ takes values in $\{0,1\}$. Then, at any location (m,n), we can write

$$\begin{aligned} Z(m,n) \;=\; & t_0\,X(m,n) \\ & + t_1\,X(m-1,n-1) + R(m,n), \end{aligned}$$

where $R(m,n) := \sum' t(j,k)\,X(m-j,n-k)$; the sum $\sum'$ is computed for $(j,k) \neq (0,0),(1,1)$. (For simplicity, let's assume that $\mathcal{T}$ has a finite support.) Because the pixels of $\mathcal{X}$ are independent, it is always possible to find the locations (m_i,n_i), $(i = 1,2,3,4)$, such that $R(m_i,n_i) = \text{const.}$ and $\{X(m_i,n_i), X(m_i-1,n_i-1)\} = \{(0,0),(0,1),(1,0),(1,1)\}$. This implies that the image $\mathcal{Z}$ would at least take all the values in $\{r, t_0+r, t_1+r, t_0+t_1+r\}$ and thus contradicts with the assumption that $\mathcal{Z}$ is a two-tone image.

The above argument can be generalized to non-*i.i.d.* images, with the relaxed requirement that any finite collection of the two tone values be visited by the pixels of $\mathcal{X}$ with a positive probability [3]. This requirement is analogous to the ergodicity requirements in the classic statistical theory for consistent parameter estimation. For two-tone images that satisfy this requirement, it can be shown that $\mathcal{Z}$ must be a shifted and/or scaled version of $\mathcal{X}$ once $J(\mathcal{F})$ is minimized with some filter $\mathcal{F} \neq 0$.

4. How to Compute the Minimizer?

The minimization of $J(\mathcal{F})$ is a nonlinear problem and therefore has to rely upon some iterative procedures. In the following, we outline a three-step iterative algorithm (TSIA) for the minimization of $J(\mathcal{F})$. More sophisticated variations of TSIA are possible, especially for Step 3, but we do not address them in this article.

1. With $\mathcal{Z}_k := \mathcal{F}_k \otimes \mathcal{Y}$, compute $\boldsymbol{a}_k$ such that

   $$\boldsymbol{a}_k := \arg\min\{\boldsymbol{a}^T H(\mathcal{Z}_k)\,\boldsymbol{a} :\; a_2 = 1\}.$$

2. Compute $\beta_{1,k}$ and $\beta_{2,k}$ as the zeros of the two-degree polynomial $\psi_k(Z) := a_{0,k} + a_{1,k} Z + Z^2$, where $\boldsymbol{a}_k := [a_{0,k}, a_{1,k}, 1]^T$.

3. Update the filter by a steepest descent method, with step size $\mu > 0$, such that

   $$\mathcal{F}_{k+1} := \mathcal{F}_k - \mu\,\nabla J_k(\mathcal{F}_k),$$

   where $J_k(\mathcal{F}) := \boldsymbol{a}_k^T H(\mathcal{Z})\,\boldsymbol{a}_k$ and $\mathcal{Z} = \mathcal{F} \otimes \mathcal{Y}$.

It is clear that TSIA jointly estimates the inverse blur filter, the original image, and the tones by $\mathcal{F}_k$, $\mathcal{Z}_k$, and $(\beta_{1,k}, \beta_{2,k})$, respectively. Note that in order to avoid the trivial solution of $\mathcal{F} = 0$ it may be necessary to impose a constraint, such that $\sum f(j,k) = 1$, on the filter. This constraint does not affect the viability of the method because the tones are automatically adjusted in TSIA according to the scale of $\mathcal{Z}$.

The computation of $\boldsymbol{a}_k$ in Step 1 is quite straightforward. In fact, since $\boldsymbol{a}^T H(\mathcal{Z}_k)\,\boldsymbol{a}$ is a quadratic function of a_0 and a_1 under the constraint $a_2 = 1$, it is easy to show that $\boldsymbol{a}_k$ can be uniquely determined by

$$a_{0,k} = (c_{1,k}c_{3,k} - c_{2,k}^2)/(c_{2,k} - c_{1,k}^2)$$

and

$$a_{1,k} = (c_{1,k}c_{2,k} - c_{3,k})/(c_{2,k} - c_{1,k}^2),$$

where $c_{i,k} := E\{Z_k^i(m,n)\}$ is the ith central moment of $\mathcal{Z}_k$, $(i = 1, 2, 3)$.

As a simple variation of TSIA, one may update the tones after the filter is iterated $M > 1$ times. In other words, one may carry out Steps 1 and 2 once every M iterations of Step 3 with fixed tones. Experience shows that this variation reduces the computational complexity but still achieves reasonable results as long as M is not too large.

5. How Does It Work On Data?

To test the proposed method, let us consider the blurred text image shown in Fig. 1. The blur is caused by an autoregressive filter of the form

$$\begin{aligned} Y(m,n) \;=\; & \rho_1 Y(m-1,n) + \rho_2 Y(m,n-1) \\ & - \rho_1\rho_2 Y(m-1,n-1) + X(m,n), \end{aligned}$$

where ρ_1 and ρ_2 are the filter parameters. In Fig. 1 we use $\rho_1 = \rho_2 = 0.7$.

For automatic character recognition, the raw images are typically required to be segmented into binary ones before they can be fed into a trained recognizer, such as an artificial neural network. A simple segmentation method is to classify each pixel as being black or white by comparing the pixel with a predetermined threshold. If the blurred image in Fig. 1 is segmented without deblurring, one would obtain a binary image that could be difficult for the recognizer to handle.

Fig. 2(a) shows a segmentation result for the blurred image in Fig. 1. As we can see, the direct segmentation of the blurred image produces a result that is unrecognizable perhaps even by trained human eyes. The difficulty is also reflected objectively by the proportion of misclassified pixels, which equals 16.19% for the 33-by-256 image in Fig. 2(a).

To improve the segmentation results, we apply the proposed deblurring method to the blurred image in Fig. 1. The filter $\mathcal{F}$ is restricted to be of size three-by-three, so that $f(j,k) = 0$ if $|j| \geq 2$ or $|k| \geq 2$. The nonzero coefficients of the initial filter $\mathcal{F}_0 := \{f_0(j,k)\}$ are $f_0(0,0) = 1.0$, $f_0(1,0) = f(0,1) = -0.2$, and $f_0(1,1) = 0.04$. Fig. 2(b) shows the segmentation result based on the output from the initial filter – the

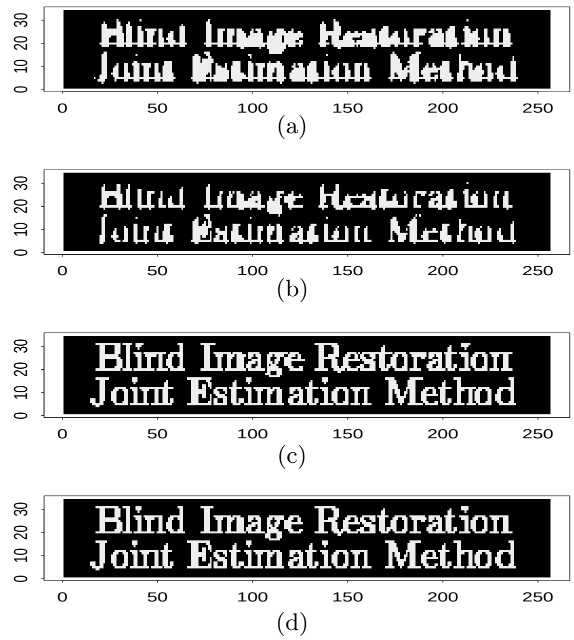

Figure 2. Segmentation results from: (a) the blurred image in Fig. 1; (b) the output of the initial filter; and (c) the deblurred image after 4,000 iterations of TSIA. The original text image is shown in (d).

image is still difficult to recognize, with the proportion of misclassified pixels equal to 10.58%.

Initiated with $\mathcal{F}_0$, we update the deblurring filter using the TSIA algorithm – the step size parameter is taken to be $\mu = 2 \times 10^{-10}$ and the tones are estimated once every $M = 30$ iterations of Step 3. Fig. 2(c) shows the segmentation result based on a deblurred image after 4,000 iterations. Compared with Figs. 2(a)-(b), and with the original two-tone (0-255) image in Fig. 1(d), the improvement on recognizability is quite significant, not only visually but also in terms of the proportion of misclassified pixels, which is now reduced to 0.57%.

The convergence behavior of TSIA is shown in Fig. 3 in terms of the cost function $J(\mathcal{F})$. As we can see, the cost function decreases monotonically as the iteration proceeds. The convergence rate is high at the beginning, but the algorithm slows down after a certain number of iterations (roughly 2,000 in this example). This behavior is typical for the steepest-descent-type algorithms [6].

To obtain the segmented images in Fig. 2, the segmentation threshold is determined so that the proportions of black and white pixels after the segmentation equal those in the original image. Fig. 4 and Table 1

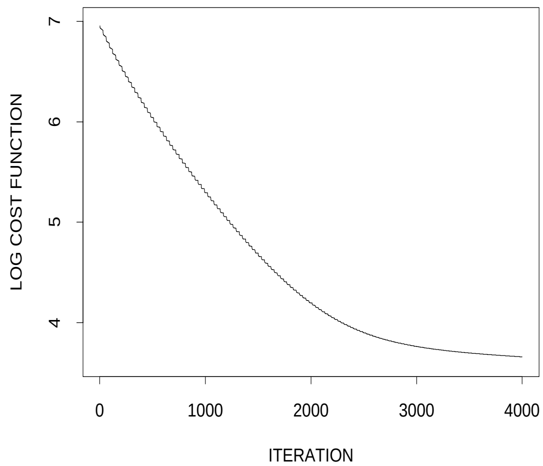

Figure 3. Plot of the cost function $J(\mathcal{F})$ (in logarithmic scale) against the number of iterations. The TSIA parameters are: $\mu = 2 \times 10^{-10}$ and $M = 30$.

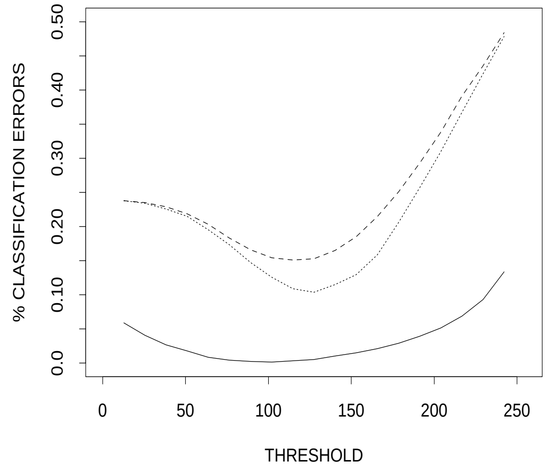

Figure 4. Proportion of mis-classified pixels as a function of the segmentation threshold – dashed line for the blurred image, dotted line for the image from the initial filter, and solid line for the deblurred image after 4,000 iterations.

Table 1. Percentage of mis-classified pixels

Threshold	Blurred	Initial	Deblurred
63	20.33	19.53	0.86
76	18.34	17.34	0.47
89	16.60	14.73	0.24
102	15.42	12.56	0.16
114	15.11	10.90	0.33
127	15.29	10.36	0.54
140	16.50	11.50	1.06

demonstrate the impact of the threshold on the segmentation results in terms of the classification errors. When the images are blurred, the segmentation error depends crucially on the threshold, as shown by the dashed and dotted lines in Fig. 4. The threshold is much less critical when the blur is properly removed, as shown by the solid line in Fig. 4. In this example, the deblurred image produces dramatically improved segmentation results for all the thresholds.

6. Conclusions

In this article we have proposed a method for the restoration of blurred two-tone images when the blur filter is unknown. The method jointly estimates the blur filter along with the original image and the tone values based solely on the blurred image.

The method shows that it is possible to blindly deblur certain two-tone images of correlated pixels without explicitly estimating or modeling their statistical properties. The text image example demonstrates the potential usefulness of the method for front-end processing in automatic character recognition systems.

For the future research, one can certainly extend the results in this article to multi-tone images and explore other possible cost functions and algorithms for the implementation. The bottom line is that if one judiciously takes advantage of the prior two-tone (or multi-tone) information, simple yet powerful deblurring algorithms can be obtained.

References

[1] T. Pavlidis, J. Swartz, and Y. P. Wang, " Information encoding with two-dimensional bar codes," *IEEE Computer*, vol. 25, no. 6, pp. 18–28, 1992.

[2] K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, vol. 10, pp. 1195–1208, 1982.

[3] T. H. Li and K. S. Lii, "A Joint Estimation Approach for Two-Tone Image Deblurring," submitted for publication, 1997.

[4] T. H. Li, "Blind identification and deconvolution of linear systems driven by binary random sequences," *IEEE Trans. Inform. Theory*, vol. 38, no. 1, pp. 26–38, 1992.

[5] T. H. Li, "Estimation and blind deconvolution of autoregressive systems with nonstationary binary inputs," *J. Time Series Anal.*, vol. 14, no. 6, pp. 214–236, 1993.

[6] S. Haykin, *Adaptive Filter Theory*, 3rd Edn., New York: Prentice Hall, 1996.