A Super-Exponential Method for Blur Identification and Image Restoration

Thomas J. Kostas, Laurent Mugnier, Aggelos K. Katsaggelos, and Alan V. Sahakian
Department of Electrical Engineering and Computer Science,
McCormick School of Engineering and Applied Science
Northwestern University, Evanston, IL, 60208-3118

Abstract

This paper examines a super-exponential method for blind deconvolution. Possibly non-minimal phase point spread functions (PSFs) are identified. The PSF is assumed to be low pass in nature. No other prior knowledge of the PSF or the original image is necessary to assure convergence of the algorithm. Results are shown using synthetically degraded satellite images in order to demonstrate the accuracy of the PSF estimates. In addition, radiographic images are restored with no knowledge of the PSF of the x-ray imaging system. These experiments suggest a promising application of this algorithm to a variety of blur identification problems.

Keywords: blur identification, cumulants, higher order statistics, higher order spectra, blind deconvolution

1. INTRODUCTION

Blur identification is an important and necessary step in the restoration of images having linear, space invariant degradations. Since the 1960s, research in blind deconvolution, where there is no assumed knowledge of the degradation system's point spread function (PSF), has centered around communication signals. Initial research dealt with systems whose PSF could be modeled parametrically using zero crossing information preserved in the spectrum of the blurred signal. Only certain classes of blurs, such as cylindrical or 1-D motion blurs, could be successfully extracted using these methods. Stochastic methods were soon introduced. Autoregressive Moving-Average (ARMA) models along with maximum likelihood based approaches became successful due to their ability to deconvolve signals with low to moderate noise [1, 2].

Recently, the use of higher order statistics/spectra (HOS) has become popular in PSF identification in communication channels [3, 4, 5, 6, 7]. HOS preserve important phase information which is lost when using second order spectra. Another attractive feature is that given enough realizations of a signal containing additive Gaussian noise, the use of HOS allows us to easily filter out the noise. These properties permit us to identify non-minimum phase systems, such as non-causal blurs.

In this work we extend and apply a blur identification algorithm based on HOS, which was previously used for the 1-D communication signal problem [8], to 2-D radiographic images. Our primary motivation in this work is to provide fast and accurate restorations of blurred radiographs of fatigued aircraft structures. To do this, we use a blind deconvolution algorithm that converges at a super-exponential rate independent of initialization, as proven in [8]. The algorithm can be applied to the blind deconvolution of any type of image, such as those obtained by the Hubble Space Telescope, satellites, etc.

The general blind deconvolution problem is illustrated in Figure 1. In this model, $a_{m,n}$ is the input, H represents the transfer function of the degradation system, $y_{m,n}$ is the output of the system, C is the inverse filter transfer function, $z_{m,n}$ is the output of the inverse filter, and S represents the overall transfer function of the system. Thus,

$$s_{m,n} = h_{m,n} * c_{m,n} = \sum_l \sum_i h_{l,i} c_{m-l,n-i}, \tag{1}$$

where $h_{m,n}$ and $c_{m,n}$ represent respectively the PSFs of the blurring system and its inverse, and $*$ denotes convolution. Since recognition of the blur, and therefore the original image, is desired, $z_{m,n}$ should equal $a_{m,n}$ with the exception of a possible phase shift and delay. Therefore:

$$s_{m,n} = e^{j\phi} \delta_{m-u,n-v}, \tag{2}$$

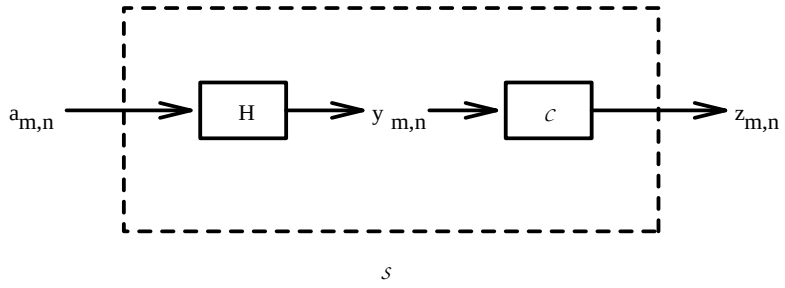


Figure 1: Blind deconvolution model

where u and v represent the two dimensional delay and ϕ the phase shift is the system response being sought. H, C, and S can have either finite or infinite length.

The blind deconvolution algorithm discussed in this paper can first be described in terms of the system response S. Through the formulation of this algorithm in terms of the system response, we gain insight into how S converges to the desired response described in Eq. (2). Consider the following iteration:

$$(s^k_{m,n})' = (s^{k-1}_{m,n})^p((s^{k-1}_{m,n})^*)^q \quad (3)$$

$$s^k_{m,n} = \frac{(s^k_{m,n})'}{\|(s^k)'\|}, \quad (4)$$

where $*$ stands for complex conjugate, $(s^k_{m,n})'$ represent the values of the coefficients during iteration k, and $s^k_{m,n}$ are the normalized coefficients. The norm $\|\cdot\|$ represents the l_2 norm. The coefficients, p and q, should be non-negative such that $p+q \geq 2$.

At the end of each iteration, the values of the system response are normalized such that each is less than or equal to one. During the next iteration, these coefficients are raised to the p+q power. Thus, the smaller coefficients become exponentially smaller compared to the larger values. If there is one maximum coefficient value, then the algorithm converges to a delta function of weight one at that location.

If there are M coefficients that have the same maximum value, then this iterative process converges to a function containing M delta functions of weight $\frac{1}{M}$ at those specific locations. In practice, due to errors caused by finite arithmetic and noise present in the signal, it is highly unlikely for the algorithm to converge to such functions.

2. A BLIND DECONVOLUTION ALGORITHM

The iteration described in Eqs. (3) and (4) is implemented by using empirical cumulants to estimate the blur, H. Since these cumulants are determined from observation and there is a finite amount of data, errors will be introduced which can affect the convergence of the algorithm. Errors will also occur due to the finite filter lengths that will be assumed.

2.1 Cumulants

Let $x_1, x_2, ..., x_n$ be a set of random variables and $f(x)$ be the probability density function. Then, the characteristic function of a random variable is defined by the integral (in the continuous case) [9]:

$$\Phi(\omega) = \int_{-\infty}^{\infty} f(x)e^{j\omega x}dx. \quad (5)$$

The second characteristic function of x is defined as:

$$\Psi(\omega) = ln\Phi(\omega). \quad (6)$$

The cumulant of $x_{i_1}, x_{i_2}, ..., x_{i_m}$ where $i\epsilon\{1, 2, ...n\}$ is therefore defined as:

$$cum(x_{i_1}; x_{i_2}; ...; x_{i_m}) = (-j)^m \frac{\partial^m \Psi(\omega)}{\partial\omega_{i_1}\partial\omega_{i_2}\cdots\partial\omega_{i_m}}|_{\omega=0}. \tag{7}$$

Realizing that:

$$\Phi(\omega)|_{\omega=0} = \int_{-\infty}^{\infty} f(x)dx = 1, \tag{8}$$

the cumulants can be expanded as follows:

$$\begin{aligned}
cum(x_{i_1}; x_{i_2}; ...; x_{i_m}) &= (-j)^m \frac{\partial^{m-1}}{\partial\omega_{i_1}\partial\omega_{i_2}\cdots\partial\omega_{i_{m-1}}}(\frac{1}{\Phi(\omega)}\frac{\partial\Phi(\omega)}{\partial\omega_{i_m}})|_{\omega=0} &(9)\\
&= (-j)^m \frac{\partial^{m-2}}{\partial\omega_{i_1}\partial\omega_{i_2}\cdots\partial\omega_{i_{m-2}}}(\frac{1}{\Phi(\omega)}\frac{\partial^2\Phi(\omega)}{\partial\omega_{i_m}\partial\omega_{i_{m-1}}} \\
&\quad -\frac{1}{\Phi^2(\omega)}\frac{\partial\Phi(\omega)}{\partial\omega_{i_m}}\frac{\partial\Phi(\omega)}{\partial\omega_{i_{m-1}}})|_{\omega=0} &(10)\\
&= (-j)^m \frac{\partial^{m-3}}{\partial\omega_{i_1}\partial\omega_{i_2}\cdots\partial\omega_{i_{m-3}}}(\frac{1}{\Phi(\omega)}\frac{\partial^3\Phi(\omega)}{\partial\omega_{i_m}\partial\omega_{i_{m-1}}\partial\omega_{i_{m-2}}} \\
&\quad -\frac{1}{\Phi^2(\omega)}\frac{\partial^2\Phi(\omega)}{\partial\omega_{i_m}\partial\omega_{i_{m-1}}} - \frac{1}{\Phi^2(\omega)}\frac{\partial^2\Phi(\omega)}{\partial\omega_{i_m}\partial\omega_{i_{m-2}}} - \frac{1}{\Phi^2(\omega)}\frac{\partial^2\Phi(\omega)}{\partial\omega_{i_{m-1}}\partial\omega_{i_{m-2}}} \\
&\quad +\frac{2}{\Phi^3(\omega)}\frac{\partial\Phi(\omega)}{\partial\omega_{i_m}}\frac{\partial\Phi(\omega)}{\partial\omega_{i_{m-1}}}\frac{\partial\Phi(\omega)}{\partial\omega_{i_{m-2}}})|_{\omega=0}. &(11)
\end{aligned}$$

The following relationships can be established between cumulants and moments:

$$\begin{aligned}
m &= 1 : cum(x_{i_1}) = E(x_{i_1}), &(12)\\
m &= 2 : cum(x_{i_1}; x_{i_2}) = E((x_{i_1} - E(x_{i_1}))(x_{i_2} - E(x_{i_2}))), &(13)\\
m &= 3 : cum(x_{i_1}; x_{i_2}; x_{i_3}) = E((x_{i_1} - E(x_{i_1}))(x_{i_2} - E(x_{i_2}))(x_{i_3} - E(x_{i_3}))). &(14)
\end{aligned}$$

For $m = 2$ and $m = 3$, these are the central moments if $x_{i_1} = x_{i_2} = x_{i_3}$.

2.2 Blur Estimation

Some assumptions about the input must be made to assure convergence of this super-exponential blind deconvolution method to the desired solution. First, the input must be non-Gaussian since Gaussian cumulants of order greater than two are equal to zero. This property of cumulants can be advantageous if the signal is plagued by additive Gaussian white noise. Second, the input, $a_{m,n}, (m, n)\epsilon\{1, 2, 3, ...\}$, must be a sequence of independent identically distributed (i.i.d.) random variables. Since neighboring pixels tend to be strongly correlated, this assumption poses some problems which are evident in the results later reported. These problems are discussed in the experimental section. Third, the inverse system should be stable. In the case of blurs where the inverse is not stable (such as motion blurs), a good estimate (in the least squares sense) of the desired inverse should result. Last, the blur and its inverse are assumed to have finite length so that the combined system response S also has finite length. In a case where the inverse blur or blur have infinite length, most of the energy is assumed to be conserved by any truncation that is performed.

The steps in the blind deconvolution algorithm are described as follows:

a) Estimate the PSF $h_{m,n}$ and denote this estimate by $d^k_{m,n}$, using HOS of the restored image, $z^{k-1}_{m,n}$, and the blurred image, $y_{m,n}$.
(In the first iteration, let $z^0_{m,n} = y_{m,n}$.)

b) Compute the inverse blur, $(c^k_{m,n})'$.

c) Compute the normalized the inverse blur, $c^k_{m,n}$.

d) Restore $y_{m,n}$, the blurred image to generate $z^k_{m,n}$.

These four steps are now described in more detail. For step (a), we use:

$$d^k_{i,j} = \frac{cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; y^*_{m-i,n-j})}{cum(a_{m,n}; a^*_{m,n}; a^*_{m,n})}. \tag{15}$$

Equation (15) is derived by using:

$$y_{m,n} = a_{m,n} * h_{m,n} = \sum_i \sum_j a_{i,j} h_{m-i,n-j}. \tag{16}$$

We assume that the PSF is shift invariant. Since cumulants are linear operators, the numerator in Eq. (15) is defined by:

$$cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; y^*_{m-a,n-b}) = \sum_i \sum_j h_{i-a,j-b} cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; a^*_{m-i,n-j}) \tag{17}$$

where

$$\begin{aligned} cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; a^*_{m-i,n-j}) &= cum(\sum_e \sum_f s^{k-1}_{e,f} a_{m-e,n-f}; \sum_k \sum_l (s^{k-1}_{k,l})^* a^*_{m-k,n-l}; a^*_{m-i,n-j}) \\ &= \sum_e \sum_f \sum_k \sum_l s^{k-1}_{e,f} (s^{k-1}_{k,l})^* cum(a_{m-e,n-f}; a^*_{m-k,n-l}; a^*_{m-i,n-j}). \end{aligned} \tag{18}$$

Since $a_{m,n}$ is assumed to be an i.i.d. sequence, then:

$$cum(a_{m-e,n-f}; a^*_{m-k,n-l}; a^*_{m-i,n-j}) = cum(a_{m,n}; a^*_{m,n}; a^*_{m,n}). \tag{19}$$

Substituting (19) into (18), the following is obtained:

$$cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; a^*_{m-i,n-j}) = s^{k-1}_{i,j} (s^{k-1}_{i,j})^* cum(a_{m,n}; a^*_{m,n}; a^*_{m,n}). \tag{20}$$

Substituting (20) into (17), the following results:

$$cum(z^{k-1}_{m,n}; (z^{k-1}_{m,n})^*; y^*_{m-a,n-b}) = cum(a_{m,n}; a^*_{m,n}; a^*_{m,n}) \sum_i \sum_j h_{i-a,j-b} s^{k-1}_{i,j} (s^{k-1}_{i,j})^*, \tag{21}$$

thus giving us the means to generate the estimate of $h_{m,n}$ in Eq. (15).

Cumulants are also used to compute the inverse blur in step (b). Consider the following:

$$(c^k_{m,n})' = R^{-1}_{m,n} * d^k_{m,n}, \tag{22}$$

where each component of the two dimensional function $R_{m,n}$ at position (i,j) is given by:

$$R_{i,j} = \frac{cum(y_{m-i,n-j}; y^*_{m,n})}{cum(a_{m,n}; a^*_{m,n})} = \frac{cov(y)_{i,j}}{var(a_{m,n})} \tag{23}$$

We derive Eq. (22) by considering Eq. (16) in the following:

$$\begin{aligned} cum(y_{m-a,n-b}; y^*_{m-c,n-d}) &= cum(\sum_i \sum_j h_{i-a,j-b} a_{m-i,n-j}; \sum_u \sum_v h^*_{u-c,v-d} a^*_{m-u,n-v}) \\ &= \sum_i \sum_j \sum_u \sum_v h_{i-a,j-b} h^*_{u-c,v-d} cum(a_{m-i,n-j}; a^*_{m-u,n-v}) \\ &= \sum_i \sum_j \sum_u \sum_v h_{i-a,j-b} h^*_{u-c,v-d} cum(a_{m,n}; a^*_{m,n}). \end{aligned} \tag{24}$$

Next we consider the normalization described in step (c). If we lexicographically order $s^k_{m,n}$ and $c^k_{m,n}$, Eq. (1) can be written as:

$$s^k = Hc^k, \tag{25}$$

where s^k and c^k are vectors and H is the appropriately sized matrix. The goal is to choose the correct c^k such that $s^k = \delta^{(r)}$ where $\delta^{(r)}$ is a vector with a one in the r^{th} position and zeros elsewhere. Since c^k is forced to be of finite length, the ideal solution may not be reached as the number of iterations is increased. Therefore, c^k must be chosen to meet some error criterion. If we solve the problem by using a linear least squares approach, the solution is:

$$\begin{aligned} \min_{c^k} \|Hc^k - \delta^{(r)}\|^2 \quad &\rightarrow \quad c^k = (H^+H)^{-1}H^+\delta^{(r)} \\ &\rightarrow \quad s^k = H(H^+H)^{-1}H^+\delta^{(r)}, \end{aligned} \tag{26}$$

where + denotes the conjugate transpose of a matrix or a vector. According to the iteration described in Eq. (3), we would like to find a vector c^k such that:

$$\min_{c^k} \|Hc^k - s^k\|^2 \rightarrow c^k = (H^+H)^{-1}H^+s^k. \tag{27}$$

Given the above, the normalization in Eq. (4) takes the form:

$$c^k = \frac{(c^k)'}{\sqrt{(c^k)'^+H^+H(c^k)'}}, \tag{28}$$

or standard notation:

$$c^k_{m,n} = \frac{(c^k_{m,n})'}{\sqrt{\sum_e \sum_f \sum_i \sum_j (c^k_{i,j})'^+ R_{e-i,f-j} (c^k_{e,f})'}}. \tag{29}$$

An inverse filter is used for step (d). If c^k is long enough, $H(H^+H)^{-1}H^+$ will approach the identity matrix and s^k will converge to the desired solution. A proof of the global convergence of the algorithm subject to finite-length restrictions is in [8].

2.3 Empirical Cumulants

Since we do not know the exact cumulant values in practice, they must be estimated from the available data. As stated previously in Eqs. (12),(13), and (14), cumulants can be defined in terms of high order moments. If the input is assumed to have zero mean (and be shift invariant), then for an NxN image:

$$cum'(y_{m-u,n-v}; y^*_{m,n}) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} y_{m-u,n-v} y^*_{m,n}, \tag{30}$$

$$cum'(z_{m,n}; z_{m,n}; y^*_{m-u,n-v}) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} |z_{m,n}|^2 y^*_{m-u,n-v}. \tag{31}$$

It is easily seen that the stochastic processes must be ergodic for these definitions to hold true. Asymptotic small error analysis is carried out in Appendix A of [8].

3. RESULTS

Radiographic images are neither stationary nor uncorrelated. But, these images are wide sense stationary (WSS) if they are segmented into smaller sections. These subsections can be viewed as multiple realizations of WSS images blurred with the same degradation operator. We seek to extract the characteristics of this operator by using third order cumulants.

(a) (b)

Figure 2: a) Original satellite image. b) Satellite image blurred with a zero mean Gaussian, variance 7.

Seeing that these subimages are considered WSS, the FFT is used to efficiently implement the blur identification algorithm described in Eqs. (22) and (29).

The strong local correlations of the original image pose another problem to the aforementioned algorithm. These correlations affect the estimated cumulants such that the characteristics of the blur as well as some of the characteristics of the image are extracted at each iteration. Since we are segmenting the image into multiple realizations, the local correlations inherent in each subsection of the original image are minimized by an averaging process. The averaging process is defined by taking the third order cumulant of each subsection and averaging the results.

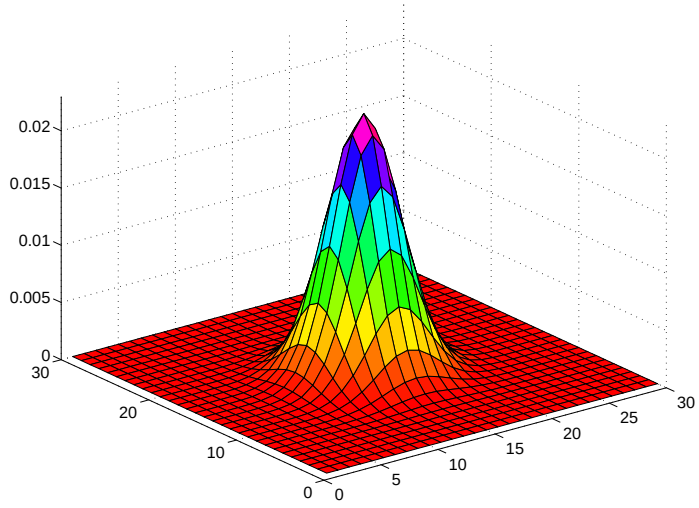
We assume that the original image is convolved with a lowpass filter. Additionally, we assume that the blur function has a DC gain of 1. Given these assumptions, the correlation matrix, R, should also have a DC gain of 1. We can use this fact in the normalization step, Eq. (29), of the algorithm. Therefore, we do not need to know the variance of the original image as used in Eq. (23). Since the blur, d, is normalized in Eq. (29), prior knowledge of the third order cumulant of the original signal, used in Eq. (15), is also unnecessary. In addition, the blur function that is estimated is forced to be non-negative at each iteration.

The satellite image in Fig. (2a) looks fairly decorrelated, especially when it is segmented into smaller portions. This yields a fairly good estimate of the blur. The actual blur we use to filter these images is a zero mean Gaussian with a variance of 7, as seen in Figs. (3a) and (4a). The blurred image is shown in Fig. (2b). We can see in the estimates of the blur, Figs. (3b) and (4b), that the variance of the estimate is slightly larger than that of the actual Gaussian filter. This may be due to the correlations present in the original image.

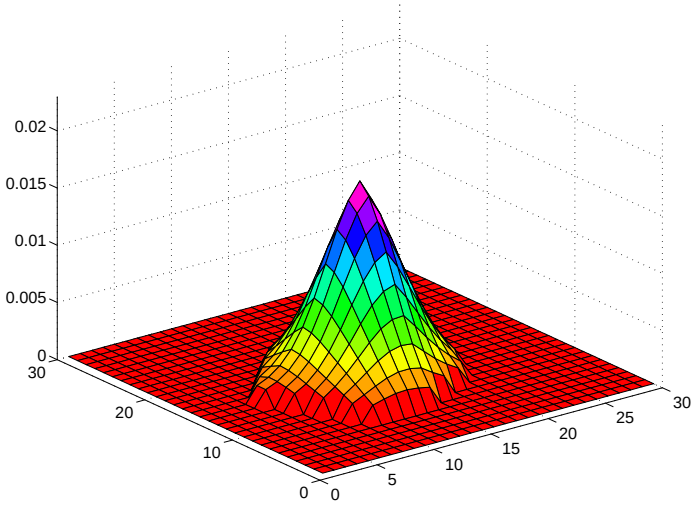
The x-ray phantom image of an aircraft lap joint in Fig. (5a) is blurred due to geometric unsharpness, which is a result of the x-ray imaging system setup. Geometric unsharpness varies with the depth of the object, therefore, a particular depth plane is normally restored. This depth varying PSF is difficult to estimate. The restored image and the estimates of the blur are seen in Figs. (5) and (6). The variance of the estimate is large, which explains the appearance of the restoration. Effects due to the non-symmetric nature of the blur are also seen in this restored image.

4. CONCLUSIONS

The super-exponential, blur identification algorithm which is described in this paper produces good estimates of the blur function. If additional information is known about the PSF, better results are expected by using constraints

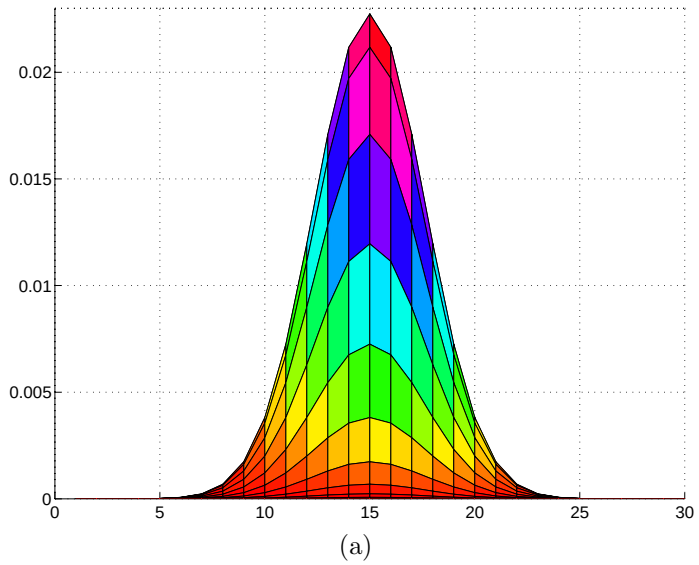


(a)

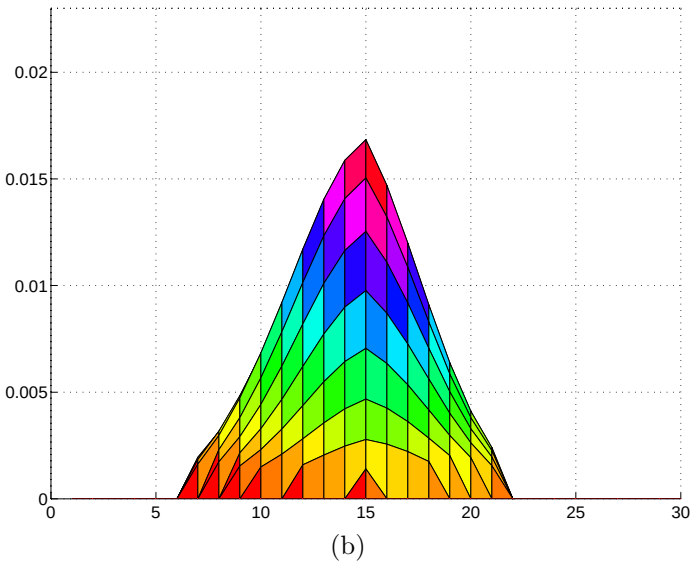


(b)

Figure 3: a) Original blur. b) Estimate of blur after eight iterations.

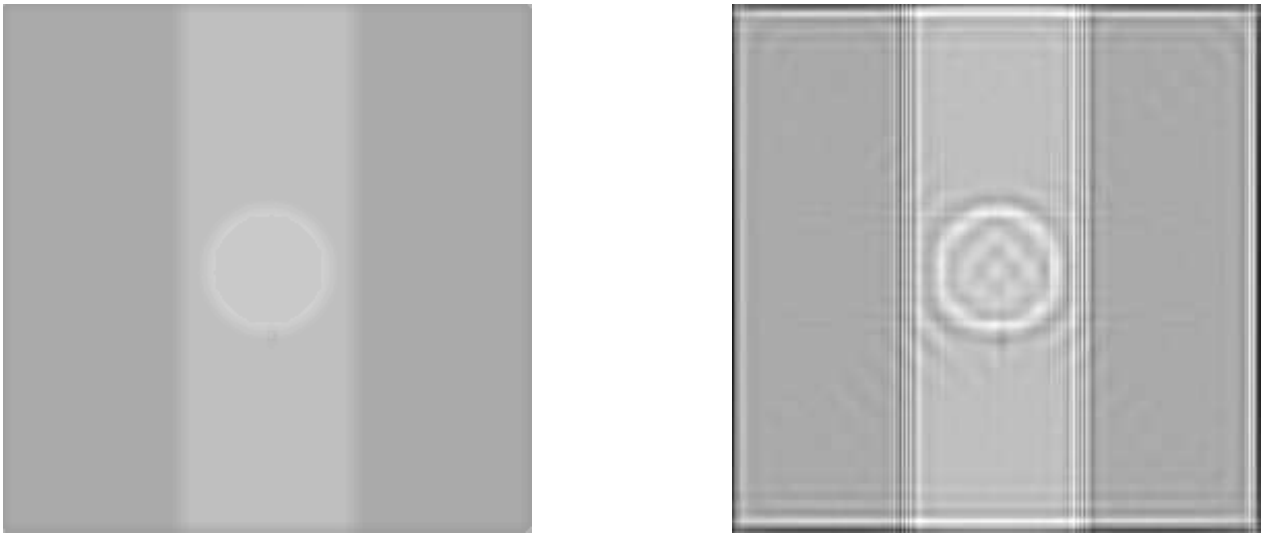


(a)



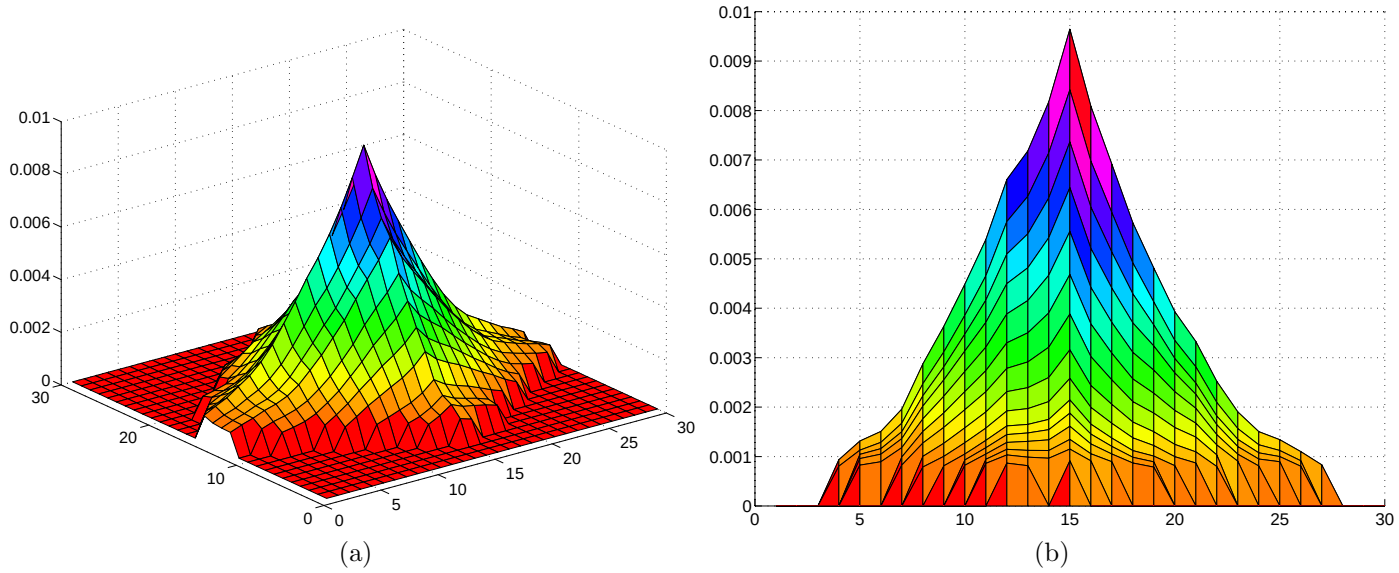
(b)

Figure 4: a) Original blur, side view. b) Estimate of blur after eight iterations, side view.



(a) (b)

Figure 5: a) Simulated x-ray image of bolt in lap-joint. b) Image restored with estimated blur.



(a) (b)

Figure 6: a) Estimate of blur from x-ray image. b) Estimate of blur from x-ray image, side view.

resulting from this prior information. For example, since many blurs are symmetric, the PSF can be segmented into quadrants and these quadrants can be averaged to produce a symmetric PSF estimate.

The local correlations within the image increase the variance of these blur estimates. Also, this algorithm is not assured to converge to the optimal PSF due to these correlations. Currently, we are experimenting with the use of autoregressive (AR) and other predictive models in order to decorrelate the image data. Preliminary results using causal and semi-causal AR models, and non-causal predictive methods are encouraging.

Future work will also investigate the noise suppression properties of this blur identification algorithm. Specifically, it will address blur estimation in the presence of additive white Gaussian as well as Poisson noise which is present in radiographic images.

5. ACKNOWLEDGEMENTS

The authors would like to thank M. R. Banham and M. G. Kang for their useful discussions. The work of T. J. Kostas and A. V. Sahakian was sponsored in part by the FAA Center for Aviation Systems Reliability, operated by the Ames Laboratory, USDOE, for the Federal Aviation Administration under Contract No. W-7405-ENG-82 for work by Iowa State University and Northwestern University. The work of L. M. Mugnier was supported by a post-doctoral fellowship of the French Research Ministry. The work of A. K. Katsaggelos was supported in part by the Space Telescope Science Institute.

6. REFERENCES

[1] K. T. Lay and A. K. Katsaggelos, "Image Identification and Restoration Based on the Expectation-Maximization Algorithm," *Optical Engineering,* vol. 29, pp. 436-445, May 1990.

[2] A. K. Katsaggelos, editor, *Digital Image Restoration*, Springer Series in Information Sciences, vol. 23, New York: Springer-Verlag, 1991.

[3] D. Hatzinakos, C.L. Nikias, "Blind Equalization Using a Tricepstrum-Based Algorithm," *IEEE Trans. on Comm.,* VOL 39., NO. 5, May 1991, pp. 669-681.

[4] S. Haykin, editor, *Blind Deconvolution*, Prentice Hall Information and System Sciences Series, New Jersey, 1994.

[5] G.B. Giannakis, M.K. Tsatsanis, "A Unifying Maximum-Likelihood View of Cumulant and Polyspectral Measures for Non-Gaussian Signal Classification and Estimation," *IEEE Trans. on Info. Theory,* VOL. 38., NO. 2, March 1992, pp. 386-406.

[6] C. L. Nikias, J. M. Mendel, "Signal Processing with Higer-Order Spectra," *IEEE Signal Processing Magazine* VOL. 10, NO. 3, July 1993, pp. 10-37.

[7] M. R. Raghuveer, C. L. Nikias, "Bispectrum Estimation via AR Modeling," *Signal Processing,* VOL. 10, pp. 35-48, 1986.

[8] O. Shalvi, E. Weinstein, "Super-Exponential Methods for Blind Deconvolution," *IEEE Trans. on Info Theory,* VOL. 39, NO. 2, March 1993, pp. 504-519.

[9] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, 3rd ed.,* New York, NY: McGraw Hill, 1991.

[10] M. G. Kang, *Adaptive Iterative Image Restoration Algorithms,* Ph.D. Thesis, Northwestern University, January 1994.

[11] L. Mugnier, "Bispectrum and Image Restoration," Image Processing Lab Technical Report 393, Northwestern University, December 1993.