Non-negative Matrix Factorization Approach to Blind Image Deconvolution

Ivica Kopriva[1] and Danielle Nuzillard[2]

[1] Department of Electrical and Computer Engineering, The George Washington University, 801 22nd St. NW Room 615, Washington DC 20052, USA
ikopriva@gmail.com

[2] CReSTIC, Université de Reims - Chamapgne Ardenne (URCA), Moulin de la Housse, B.P. 1039, 51687 REIMS Cedex 2, France
danielle.nuzillard@univ-reims.fr

Abstract. A novel approach to single frame multichannel blind image deconvolution is formulated recently as non-negative matrix factorization (NMF) problem with sparseness constraint imposed on the unknown mixing vector. Unlike most of the blind image deconvolution algorithms, the NMF approach requires no *a priori* knowledge about the blurring kernel and original image. The experimental performance evaluation of the NMF algorithm is presented with the degraded image by the out-of-focus blur. The NMF algorithm is compared to the state-of-the-art single frame blind image deconvolution algorithm: blind Richardson-Lucy algorithm and single frame multichannel independent component analysis based algorithm. It has been demonstrated that NMF approach outperforms mentioned blind image deconvolution methods.

1 Introduction

The goal of image deconvolution is to reconstruct the original image from an observation degraded by spatially invariant blurring process and noise. Neglecting the noise term the process is modeled as a convolution of a blurring kernel $h(s,t)$ with an original source image $f(x,y)$ as:

$$g(x,y) = \sum_{s=-K}^{K} \sum_{t=-K}^{K} h(s,t) f(x+s, y+t) \tag{1}$$

where K denotes the size of the blurring kernel. If the blurring kernel is known, many so-called non-blind algorithms are available to reconstruct original image $f(x,y)$ [1]. However it is not always possible to measure or obtain information about blurring kernel, which is why blind deconvolution (BD) algorithms are important. Comprehensive comparison of BD algorithms is given in [1]. They can be divided into those that estimate the blurring kernel $h(s,t)$ first and then restore original image by some of the non-blind methods [1], and those that estimate the original image $f(x,y)$ and blurring kernel simultaneously. In order to estimate the blurring kernel, a support size has to be given or estimated. Also, quite often

J. Rosca et al. (Eds.): ICA 2006, LNCS 3889, pp. 966–973, 2006.

a priori knowledge about the nature of the blurring process is assumed to be available in order to use appropriate parametric model of the blurring process [2]. It is not always possible to know the characteristic of the blurring process. Methods that estimate blurring kernel and original image simultaneously use either statistical or deterministic *prior* on the original image, the blurring kernel and the noise [2]. This leads to a computationally expensive maximum likelihood estimation usually implemented by expectation maximization algorithm. In addition to that, exact distributions of the original image required by maximum likelihood algorithm are usually unknown. One of the most representative algorithms from this class is the blind Richardson-Lucy (R-L) algorithm. It has been originally derived for non-blind deconvolution of astronomical images in [3] and [4]. Later on, it was formulated in [5] for BD and then modified by iterative restoration algorithm in [6]. This version of blind R-L algorithm is implemented in MATLAB command '*deconvblind*'. It will be used in the section 3 for the comparison purpose during experimental performance evaluation of the NMF based blind image deconvolution method [7].

In order to overcome difficulties associated with 'standard' BD algorithms an approach was proposed in [8] based on quasi maximum likelihood with an approximate of the probability density function. It however assumed that original image has sparse or super-Gaussian distribution. This is generally not true because image distributions are mostly sub-Gaussian. To overcome that difficulty it was proposed in [8] to apply sparsifying transform to blurred image. However, design of such a transform requires knowledge of at least the typical class of images to which original image belongs. In such a case, training data can be used to design sparsifying transform. Multivariate data analysis methods such as independent component analysis (ICA) [9] might be used to solve BD problem as a blind source separation (BSS) problem. The unknown blurring process is absorbed into what is known as a mixing matrix. The advantage of the ICA approach would be that no *a priori* knowledge about the origin and size of the support of the blurring kernel is required. However, multi-channel image required by ICA is not always available. Even if it is, it would require the blurring kernel to be non-stationary, which is true for blur caused by atmospheric turbulence, but it is not true for out-of-focus blur for example. Therefore, an approach to single frame multi-channel blind deconvolution that requires minimum of *a priori* information about blurring process and original image would be of great interest.

Single frame multi-channel representation was proposed in [10]. It was based on a bank of $2-D$ Gabor filters [11] used due to their ability to realize multi-channel filtering. ICA algorithms have been applied in [10] to multichannel image in order to extract the source image and two spatial derivatives along x and y directions. There is however critical condition that source image and their spatial derivatives must be statistically independent. In general this is not true as already observed in [11]. Consequently, quality of the image restoration by proposed single frame multi-channel approach depends on how well each particular image satisfies statistical independence assumption. Therefore, an extension of the ICA approach formulated in [10] is given in [7] where it has been shown that single frame multichannel BD can be formulated as NMF problem with sparseness constraint imposed on the unknown mixing vector. Consequently, no *a priori* knowledge about either the origin or the size of the blurring process is required. Because NMF is deterministic approach no *a priori* information about the statistical nature of the source image is required as well. The rest of the paper is organized as follows. We briefly introduce in section 2 blind R-L algorithm [5][6], ICA approach to single frame multichannel BD [10] and NMF approach to single frame multichannel BD with sparseness constraint [7]. Comparative experimental performance evaluation is given in section 3 for images degraded by out-of-focus blurring process. Conclusion is presented in section 4.

2 Basic Overview of the Compared Blind Image Deconvolution Algorithms

Before proceeding to description of non-bind and blind image deconvolution algorithms, we shall rewrite image observation model given in Eq.1 in the lexicographical notation:

$$g = Hf \tag{2}$$

where $g, f \epsilon \mathbb{Z}_{0+}^{MN}$, $H \epsilon \mathbb{R}_{0+}^{MN \text{x} MN}$ assuming image dimensionality of $M\text{x}N$ pixels. Observed image vector g and original image vector f are obtained by the row stacking procedure. The matrix H is block-circulant matrix, [13], and it absorbs into itself the blurring kernel $h(s,t)$ assuming at least size of it, K, to be known.

2.1 Blind Richardson-Lucy Algorithm

The blind R-L method [5] [6] follows from the non-blind version of the R-L method [3] [4] which itself follows from Bayesian paradigm approach to statistical inference which dictates that inference about true image f should be based on conditional probability $P(f/g)$ given by the Bayes rule. The *prior* knowledge about image degradation process is incorporated in conditional probability $P(g/f)$ and *prior* probability $P(f)$. In low light level imaging such as in astronomy, microscopy and the night vision imaging, the appropriate choice for $P(g/f)$ is Poisson distribution [14]. In the high-brightness conditions the Poisson *prior* should be replaced by the Gaussian one. The R-L algorithm follows when non-informative *prior* is chosen for $P(f)$ *i.e.* $P(f) \prec$ const. The algorithm is obtained through the maximization of the log-likelihood function:

$$\widehat{f} = argmax(logP(g/f)) \tag{3}$$

The EM algorithm is employed to solve problem in Eq.3 yielding numerically efficient multiplicative iterative algorithm known as blind R-L algorithm [5]:

$$\widehat{H}_{i+1}^{(k)} = [(f^{(k-1)})^T (g \varnothing (\widehat{H}_i^{(k)} f^{(k-1)}))] \widehat{H}_i^{(k)} \tag{4}$$

$$\widehat{f}_{i+1}^{(k+1)} = [(f^{(k)})^T \bigotimes (H^T(g \O (H\widehat{f}_i^{(k)} f^{(k-1)}))] \O (\widehat{H}^T 1) \tag{5}$$

where index i is used to denote internal iteration of the blind R-L algorithm and k denotes main iteration index and 1 denotes a column vector with all entries equal to 1. In Eq.4, symbol $'\O'$ denotes component-wise division, and in Eq.5 symbol $'\bigotimes'$ denotes component-wise multiplication.

Multiplicative update rules in Eq.4 and Eq.5, ensure positivity of both blurring kernel and reconstructed image automatically. Problem with blind R-L algorithm is that support size K of the blurring kernel must be known or estimated by some method. This knowledge is not always available *a priori*. This is especially true for non-stationary degradation process such as atmospheric turbulence where the strength of the turbulence, measured by the parameter called scintillation index, will strongly influence the size of the blur.

2.2 ICA Approach to Single Frame Multichannel BD (SFMICA)

Single frame multi-channel representation was proposed in [10]. It was based on a bank of 2-D Gabor filters [11] used due to their ability to realize multi-channel filtering and decomposing an input image into sparse images containing intensity variation over narrow range of frequency and orientation. Multichannel version of degraded image is shown to be [10]:

$$G = \begin{pmatrix} g^T \\ g_1^T \\ \dots \\ g_L^T \end{pmatrix} \cong \begin{pmatrix} a_{01} & a_{02} & a_{03} \\ a_{11} & a_{12} & a_{13} \\ \dots & \dots & \dots \\ a_{L1} & a_{L2} & a_{L3} \end{pmatrix} \begin{pmatrix} f^T \\ f_x^T \\ f_y^T \end{pmatrix} = AF \tag{6}$$

where images $g_l, l = 1, .., L$, are produced by Gabor filters, f represents source image and f_x and f_y represent spatial derivatives along x and y directions respectively.

The used Gabor filters had the following real and imaginary respectively:

$$R(x, y) = G(x, y) * cos(\frac{\pi}{\sigma}\varphi(x, y))$$
$$I(x, y) = G(x, y) * sin(\frac{\pi}{\sigma}\varphi(x, y))$$

where

$$G(x, y) = exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{7}$$

$$\varphi(x, y) = x.cos(\frac{\pi}{4}k) + y.sin(\frac{\pi}{4}k) \quad \text{with } k = 0, 1, 2, 3. \tag{8}$$

The parameter k regulates one of the four spatial orientations. The parameter $\sigma = \sqrt{2^n}$ with $n = 1, 2$ regulates one of the two spatial frequencies. Consequently, in SFMICA and later on SFMNMF BD algorithms 16 Gabor filters (8 for real and 8 for imaginary part with 4 spatial orientations and 2 spatial frequencies) were used to obtain multichannel version of the observed image.

The unknown elements a_{lm} of the mixing matrix absorb the blurring kernel assuming no *a priori* information about it including its size. The ICA algorithm has been applied in [10] to image model Eq.6 in order to extract the source image f. There is however critical condition for the source image that must hold in order to ICA algorithm to work. Image f and its spatial derivatives f_x and f_y must be statistically independent. This is in general not true as already observed in [12]. Consequently, quality of the restored image by proposed single frame multichannel approach depends on how well each particular image satisfies statistical independence assumption.

2.3 NMF Approach to Single Frame Multichannel BD (SFMNMF)

It was further shown in [7] that single frame multichannel blind deconvolution can be represented as:

$$G = \begin{pmatrix} g^T \\ g_1^T \\ \dots \\ g_L^T \end{pmatrix} \cong \begin{pmatrix} \widetilde{a_{01}} \\ \widetilde{a_{11}} \\ \dots \\ \widetilde{a_{L1}} \end{pmatrix} \left(f^T \right) = \widetilde{a} f^T \tag{9}$$

where images $g_l, l = 1, ..., L$, were again produced by Gabor filters. Coefficients of the unknown blurring kernel were absorbed into coefficients $\widetilde{a_{lm}}$ of the unknown mixing vector $\widetilde{a}$. Image model Eq.9 suggests the existence of only one source image f in the linear image observation model. Spatially oriented Gabor filters produce images with sparse (super-Gaussian) distributions. If the source image f is sub-Gaussian, which is the case for natural images, an unknown mixing vector must be sparse. Because $\tilde{a}$ and f are non-negative, this enabled in [7] to formulate blind deconvolution problem as an NMF problem with sparseness constraint imposed on the mixing vector [15]:

$$\begin{gathered} (\widehat{\widetilde{a}}, \widehat{f}) = argmin \parallel G - \widehat{\widetilde{a}} \widehat{f}^T \parallel^2 \\ \text{subject to } sparseness(\tilde{a}) = S_a \end{gathered} \tag{10}$$

where 'hat' denotes estimate and the measure of sparseness S_a is number between 0 and 1, with 1 meaning that all components of vector $\tilde{a}$ are small and 0 meaning the opposite [15].

A sparseness constraint S_a must be defined for NMF algorithm. In order to obtain truly unsupervised image restoration algorithm, S_a is estimated from the multichannel image G as a ratio between number of sparse images L_s and overall number of images $L + 1$. To estimate L_s, kurtosis of each image in G is estimated. Image g_l is considered to be sparse if $\kappa(g_l) > \delta$. In our experiments we set $\delta = 0.2$.

The SFMNMF algorithm is defined without using any *a priori* information about the blurring process or original image. Because this is a deterministic approach, no assumption about statistical nature of either blur or source image is required. Only sparseness constraint must be imposed on the unknown mixing vector $\widetilde{a}$. First coefficient in $\widetilde{a}$ can initially be approximated by 1, because it represents original blurring process. The rest of the coefficients can initially be set to 0 because they correspond to sparse images. Therefore initial value of the unknown mixing vector is set to $\tilde{a}^{(0)} = [1 \ \ 0 \ \ 0 \ \ ... \ \ 0]^T$.

The SFNMF approach to BD does not have to perform source separation due to the fact that multichannel version G of the observed image g can be approximated by the product of the unknown mixing vector and source image f as shown by Eq.9. Because there is only source image present in the observed image model, there is no need for source separation. This is the main difference with respect to the approach proposed in [10] and by Eq.6. However, it is still not clear at the moment how to apply NMF approach to BD when source image f is sparse. Because the multichannel image G is sparse and original image f is also sparse, it is not obvious in this case whether sparseness constraint must be imposed on the source image f only or it should be imposed on both source image f and unknown mixing vector.

3 Experimental Results

Fig.1 left shows blurred image obtained by digital camera in manually de-focused mode. Fig.1 right shows image reconstructed by SFNMF algorithm. Image reconstructed by SFMICA algorithm is shown in Fig.2 left, where FastICA algorithm with tanh nonlinearity was used. Fig.2 right shows image restored by the blind R-L algorithm after 5 iterations with the circular blurring kernel and radius of $R = 3$ pixels. Because the blurred image, Fig.1, was not highly de-focused blind R-L algorithm with kernel size of $R = 3$ pixels produced good result but still inferior to this produced by SFMNMF algorithm shown in Fig. 1 right. Because the size of the blurring kernel must be known *a priori* for R-L algorithm, the algorithm had to be run several times with the various values for the radius R and then the value

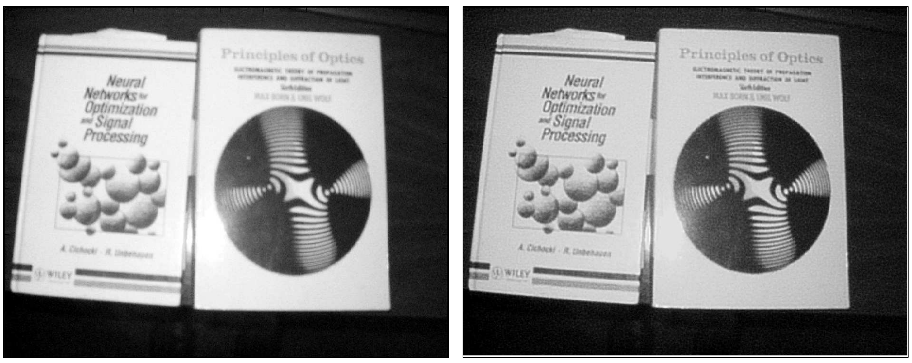

Fig. 1. (left) Image degraded by out-of-focus blur obtained by digital camera in manually de-focused mode; (right) Image reconstructed by SFMNMF algorithm

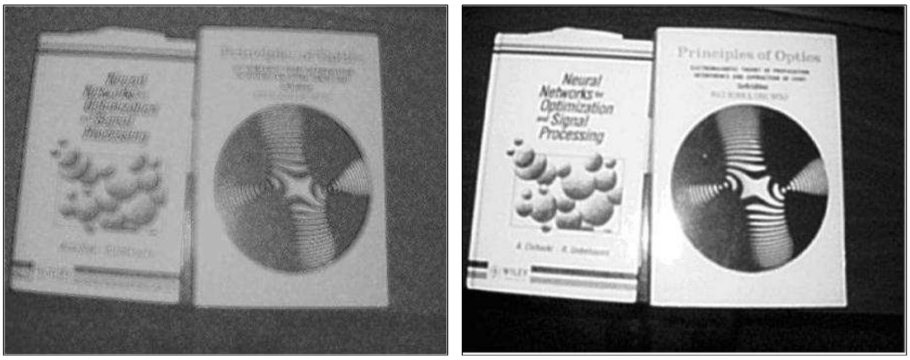

Fig. 2. (left) Image reconstructed by SFMICA algorithm; (right) Image reconstructed by blind Richardson-Lucy algorithm after 5 iterations with the circular blurring kernel with radius of $R = 3$

that corresponded to the best quality of the restored image was chosen. This was very time consuming process. In addition to that, it is known that either underestimate or overestimate of the size of the blurring kernel leads to severe distortions of the images reconstructed by blind R-L algorithm and other blind algorithms of the similar type [1]. There are no such problem with the SFMNMF algorithm. Image restored by the SFMICA algorithm has poor quality due to the fact that assumption about statistical independence between source image f and its spatial derivatives f_x and f_y does not hold. The SFMNMF algorithm eliminates all these problems due to the fact that no *a priori* knowledge about the size of the blurring kernel or statistical nature of the source image is required.

4 Conclusion

An experimental comparative performance evaluation between novel single frame multichannel blind deconvolution algorithm based on non-negative matrix factorization with sparseness constraint (SFMNMF) and other representative blind image deconvolution algorithms was presented. Image deconvolution methods were compared on image degraded by out-of-focus blur. It has been demonstrated that novel blind image deconvolution algorithm outperforms other methods. We suggest that this result is due to the characteristic of the SFMNMF algorithm which does not require any *a priori* information about the blurring kernel and original image.

References

1. M.R. Banham and A.K. Katsaggelos, "Digital Image Restoration," IEEE Signal Processing Magazine, vol. 14, no. 2, pp. 24-41, March 1997.
2. D. Kundur and D. Hatzinakos, "Blind Image Deconvolution," IEEE Signal Processing Magazine, vol. 13, no. 3, pp. 43-64, May 1996.

3. W.H. Richardson, "Bayesian-based iterative method of image restoration," J.Opt.Soc. Amer., vol. 62, pp. 55-59, 1972.
4. L.B. Lucy, "An iterative technique for rectification of observed distribution," Astron. J., vol. 79, pp.745-754, 1974.
5. D.A. Fish, A.M. Brinicombe, E.R. Pike and J.G. Walker, "Blind deconvolution by means of the Richardson-Lucy algorithm," J. Opt. Soc. Am. A, Vol. 12, No. 1, pp.58-65, January, 1995.
6. D.S.C. Biggs and M. Andrews, "Acceleration of iterative image restoration algorithms," Applied Optics, Vol. 36, No. 8, 1997.
7. I. Kopriva, "Single frame multichannel blind deconvolution by non-negative matrix factorization with sparseness constraints," Optics Letters Vol. 30, No. 23, pp. 3135-3137, December 1, 2005.
8. M.M. Bronstein, A. Bronstein, M. Zibulevsky, and Y.Y. Zeevi, "Blind Deconvolution of Images Using Optimal Sparse Representations"IEEE Tr. on Image Processing, vol. 14, No. 6, pp. 726-736, 2005.
9. A. Hyvärinen, J. Karhunen, and E. Oja, Independent Component Analysis, Wiley Interscience, 2001.
10. S. Umeyama, "Blind Deconvolution of Images by Using ICA," Electronics and Communications in Japan, Part 3, vol. 84, No. 12, 2001.
11. J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression" IEEE Tr. on Acoust., Speech and Sig. Proc., Vol. 36, pp. 1169-1179, 1988.
12. M. Numata, and N. Hamada,"Image Restoration of Multichannel Blurred Images by Independent Component Analysis", Proceedings of 2004 RISP International Workshop on Nonlinear Circuit and Signal Processing (NCSP'04),Hawaii, USA, March 5-7, 2004, pp.197-200.
13. R. L. Lagendijk and J. Biemond, Iterative Identification and Restoration of Images, KAP, 1991.
14. R. Molina, J. Nunez, F.J. Cortijo and J. Mateos, "Image Restoration in Astronomy - A Bayesian Perspective," IEEE Signal Processing Magazine, vol. 18, no. 2, pp. 11-29, March 2001.
15. P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," Journal of Machine Learning Research 5, pp. 1457-1469, 2004.