# Blind Image Separation Through Kurtosis Maximization

Ning Chen and Phillip De Leon
New Mexico State University
Klipsch School of Electrical and Computer Engineering
Las Cruces, NM  88003
{nchen, pdeleon}@nmsu.edu

## Abstract

*Blind Source Separation has been an extremely active area of research for the last few years. Most of the research has been focused on separation of sources from one-dimensional mixture signals such as speech. More recently, separation of two-dimensional sources (images) has been also examined to a limited extent using second-order statistics, information theoretic models, and neural networks. In this paper, we extend a simple kurtosis maximization algorithm, successfully used in separation of instantaneous speech signals, to images. The higher-order statistics-based algorithm is simple and performs relatively well.*

## 1  Introduction

Blind Source Separation (BSS) and Independent Component Analysis (ICA) have been extremely active areas of research for the last few years [1]. Both of these methods seek to separate out one or more individual source signals from one or more mixture signals using either Higher Order Statistics (HOS), Information-Theoretic (IT) Models, or Neural Networks (NN). Mixture signals are modeled as either an instantaneous mix of scaled sources or as a more general convolutional mix of filtered sources.

In instantaneous mixing, source signals $s_1(\eta)$ and $s_2(\eta)$ are projected onto a mixing matrix, $\mathbf{A}$ composed of scalar elements to produce the mixture signals, $x_1(\eta)$ and $x_2(\eta)$. In matrix form we have

$$\left[ \begin{array}{c} x_1(\eta) \\ x_2(\eta) \end{array} \right] \;=\; \left[ \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right] \left[ \begin{array}{c} s_1(\eta) \\ s_2(\eta) \end{array} \right] \qquad (1)$$

or

$$\mathbf{x}(\eta) \;=\; \mathbf{A}\mathbf{s}(\eta) \qquad (2)$$

where $\mathbf{s} = [s_1(\eta)\ s_2(\eta)]^T$ and $\mathbf{x}(\eta) = [x_1(\eta)\ x_2(\eta)]^T$. The independent variable set $\eta$ can be either a sample index $n$ for a one dimension signal, like speech, or a spatial coordinate $(m, n)$ for a two dimension signal, like an image.

In the BSS problem, we seek to build a de-mixing matrix, $\mathbf{W}$ which is an estimate of $\mathbf{A}^{-1}$ such that the output signals, $y_1(\eta)$ and $y_2(\eta)$ approximate the original, unknown sources. In matrix form we have,

$$\left[ \begin{array}{c} y_1(\eta) \\ y_2(\eta) \end{array} \right] \;=\; \left[ \begin{array}{cc} w_{11} & w_{12} \\ w_{21} & w_{22} \end{array} \right] \left[ \begin{array}{c} x_1(\eta) \\ x_2(\eta) \end{array} \right] \qquad (3)$$

or

$$\mathbf{y}(\eta) \;=\; \mathbf{W}\mathbf{x}(\eta) \qquad (4)$$

where $\mathbf{y}(\eta) = [y_1(\eta)\ y_2(\eta)]^T$. In convolutional mixing, the sources are convolved with a mixing matrix whose elements are now digital filters, $\mathbf{a}_{ij}$ and we seek a de-mixing matrix composed of elements which are also digital filters, $\mathbf{w}_{ij}$ in order to minimize co-channel interference from the other (filtered) source.

Most of the BSS/ICA research has been focused on separation of sources from one-dimensional mixture signals such as speech. In the speech application, we assume a convolutional mixing (as would be expected in a reverberant environment) with individual mixing filters having on the order of thousands of coefficients. Although this problem appears (and is) extremely difficult, some success has been reported [2]. To a more limited extent, signal separation has been extended to cases with two-dimensional sources(images). Applications include the separation of original thermal wave images from the noise-deteriorated images for nondestructive evaluation (NDE) of materials and the extraction of the foetal electrocardiogram (FECG) from skin electrode signals in biomedical signal processing. Noteworthy is the work by Sahlin and Broman which separate simple convolutional mixtures ($2 \times 2$ spatial filter and filters with only one non-zero coefficient) using second order statistics [3], Cichochi and Amari which use NNs [4], and Miskin and MacKay which apply IT methods (Ensemble Learning) to separate both instantaneous and convolutional mixtures of images [5].

In this paper, we extend a normalized kurtosis maximization algorithm (KMA), successfully used in separation of instantaneous mixtures of speech signals, to

images. We provide motivation for the use of kurtosis and development of the algorithm, necessary extensions for image separation, as well as the examples using standard test images.

## 2  Kurtosis-Based Separation Algorithm

Most approaches to the problem of source separation assume statistical independence of the source signals. To capture statistical independence, Higher Order Statistics (HOS) are required [6]. Many approaches to BSS involve HOS such as cumulants and kurtosis [7]:

$$\kappa_x \equiv \frac{E\left[x^4\right]}{\{E\left[x^2\right]\}^2}. \qquad (5)$$

Kurtosis is the classical measure of nongaussianity, i.e.

$$
\begin{aligned}
\kappa_x &< 3 \text{ subgaussian (or platykurtic)} \\
\kappa_x &= 3 \text{ gaussian (or mesokurtic)} \qquad (6) \\
\kappa_x &> 3 \text{ supergaussian (or leptokurtic).}
\end{aligned}
$$

Motivation for this approach stems from the long-term observation that many natural source signals of interest (such as speech) have kurtosis lower than their instantaneous mixtures [7]. This observation can be used for separation as follows. Assuming $x$ is of zero mean and unit variance, (5) simplifies to $E\left[x^4\right]$. For the instantaneous mixture $x_i = a_{i1}s_1 + a_{i2}s_2$, with $-1 < a_{i1}, a_{i2} < 1$, we have $\kappa_{x_i} = a_{i1}^4\kappa_{s_1} + a_{i2}^4\kappa_{s_2}$. Without loss of generality, we can assume that $\sigma_{x_i}^2 = 1$, this implies a constraint on $a_{ij}$, i.e. $E\left[x_i^2\right] = a_{i1}^2 + a_{i2}^2 = 1$. It is not hard to show that the maxima of $\kappa_{x_i}$ are at the points when one of the mixing coefficients is zero, and the other one is either $+1$ or $-1$, i.e. $x_i = \pm s_1$ or $\pm s_2$ [8].

We may therefore iteratively construct a de-mixing matrix based on maximizing the output kurtosis. Mathematically,

$$
\begin{aligned}
\mathbf{W}(n+1) &= \mathbf{W}(n) + \mu \bigtriangledown \kappa_{\mathbf{y}} \\
&= \mathbf{W}(n) + \mu \begin{bmatrix} \frac{\partial \kappa_{y_1}}{\partial W_{11}} & \frac{\partial \kappa_{y_1}}{\partial W_{12}} \\ \frac{\partial \kappa_{y_2}}{\partial W_{21}} & \frac{\partial \kappa_{y_2}}{\partial W_{22}} \end{bmatrix} \\
&= \mathbf{W}(n) + \mu \mathbf{C}(n) \qquad (7)
\end{aligned}
$$

where $\mu$ is the step size, $\bigtriangledown \kappa_{\mathbf{y}}$ is the gradient of the kurtosis of the output signals with respect to the elements of the separation matrix, $\kappa_{y_i}$ is the output kurtosis, and $\mathbf{C}(n)$ is the correction matrix used in the update rule. The complete algorithm is given in Fig. 1 where $\hat{\sigma}_i^2$ is an autoregressive (AR) estimate of the variance (assuming zero mean) of $x_i$, $\hat{r}_{12}$ is an AR estimate of

$$
\begin{aligned}
\mathbf{y}(n) &= \mathbf{W}(n)\mathbf{x}(n) \\
\hat{\sigma}_i^2(n) &= \lambda_2\hat{\sigma}_i^2(n-1) + (1-\lambda_2)x_i^2(n) \\
\hat{r}_{12}(n) &= \lambda_2\hat{r}_{12}(n-1) + (1-\lambda_2)x_1(n)x_2(n) \\
\alpha_i &= 4y_i^3(n) \\
\beta_i &= -W_{i1}(n)\hat{r}_{12}(n)x_1(n) - W_{i2}(n)\hat{\sigma}_2^2(n)x_1(n) + \\
&\quad W_{i1}(n)\hat{\sigma}_1^2(n)x_2(n) + W_{i2}(n)\hat{r}_{12}(n)x_2(n) \\
\gamma_i &= \left[W_{i1}^2(n)\hat{\sigma}_1^2(n) + 2W_{i1}(n)W_{i2}(n)\hat{r}_{12}(n) + \right. \\
&\quad \left. W_{i2}^2(n)\hat{\sigma}_2^2(n)\right]^{-3} \\
\mathbf{C}(n) &= \begin{bmatrix} -\alpha_1\beta_1\gamma_1 W_{12}(n) & \alpha_1\beta_1\gamma_1 W_{11}(n) \\ -\alpha_2\beta_2\gamma_2 W_{22}(n) & \alpha_2\beta_2\gamma_2 W_{21}(n) \end{bmatrix} \\
\mathbf{W}(n+1) &= \mathbf{W}(n) + \mu\mathbf{C}(n)
\end{aligned}
$$

Figure 1: Kurtosis maximization algorithm (KMA) for speech separation.

the cross-correlation between $x_1$ and $x_2$, and $\lambda$ is the forgetting factor [9].

In order to examine the kurtosis of image data and their mixtures, we assume two $M \times N$ source images are mixed according to (2) to produce two $M \times N$ mixture images. Starting with the two mixture images (no information regarding sources or mixing matrix is available), we reshape $x_1$ and $x_2$ into two, $MN \times 1$ vectors and examine (5). Experimental results have indicated that in many instances, image signals (in vector form) are sub-Gaussian and their instantaneous mixtures are lower than the individual source signals. Thus, we can apply (7) to the reshaped vectors $x_1$ and $x_2$ as one-dimensional signals. The resulting two $MN \times 1$ output signal (vectors) are reshaped back into two $M \times N$ output images.

## 3  Normalizing the HOS-Based Image Separation Algorithm

In practice with images, KMA often produces large, unnatural contrast variations in the separated output images as was observed in previous work with KMA and speech signals[9]. Analysis of these variations led to the observation that $\mu$ is a critical parameter that affects the performance of KMA, since the correction matrix, $\mathbf{C}(n)$ often changes by "large" amounts. Note that in this application, the correction matrix is affected by the amplitude of the image data, which varies considerably across pixels. We therefore utilize four normalization methods:

1. Standardize the sample data of the mixtures

2. Replace the fixed step size, $\mu$ with the normalized step size

$$\mu(n) = \frac{\tilde{\mu}}{||\mathbf{C}(n)||_F}. \qquad (8)$$

where $\tilde{\mu}$ is the normalized step size and

$$||\mathbf{C}(n)||_F = \sqrt{\sum_{i=1}^{L}\sum_{j=1}^{L}|\mathbf{C}_{ij}(n)|^2} \qquad (9)$$

3. Set the diagonal parameters, $W_{ii} = 1$ in the correction matrix to fix the arbitrary scale

4. Linearly stretch the output image data to fill the entire dynamic range using a piece-wise linear gradation function such as

$$h(p) = \begin{cases} 0, & \text{if } f(p) < u_1 \\ G-1, & \text{if } f(p) > u_2 \\ \frac{f(p)-u_1}{u_2-u_1}(G-1), & \text{if } u_1 \le f(p) \le u_2 \end{cases} \qquad (10)$$

where $[u_1, u_2]$ is the gray value interval of one separated output image, and $[0, G-1]$ is the full gray value interval [10]. The functions $f(p)$ and $h(p)$ are the value functions for the image point $p = (m, n)$ in the separated and stretched images, respectively.

The following three observations were made: 1) large, unnatural contrast variations were reduced (informal observation) in the normalized cases, 2) better separation performance was achieved (informal observation and calculated separation ratios) in the normalized cases, and 3) two distinct and fully separated images were produced in the normalized cases more often than in the unnormalized cases where often only one image was separated.

## 4   Simulation

In this section, the performance of the Normalized Kurtosis Maximization Algorithm (NKMA) is demonstrated using two separation examples (case I and case II). Here the mixing matrices are randomly chosen, and the parameter set used in the experiments is listed in Table 1.

### 4.1   Case I

Two images, "Lena" and "Lynda" (original images are at the top in Fig. 2) with $\kappa_{s_{1,2}} = 1.6305, 2.2561$ (respectively), are mixed with

$$\mathbf{A} = \begin{bmatrix} 0.5553 & 0.4447 \\ 0.3864 & 0.6154 \end{bmatrix}, \qquad (11)$$

The images mixtures (middle images in Fig. 2) have kurtosis values of $\kappa_{x_{1,2}} = 1.5092, 1.6892$ (respectively), verifying the assumption on mixed image signals having lower kurtosis.

After reshaping the mixture images to vectors, running them through the NKMA, and reshaping back to

Table 1: Simulation Parameters.

| Parameter | Value |
|---|---|
| Normalized step-size, $\tilde{\mu}$ | 0.0001 |
| Forgetting factor, $\lambda$ | 0.9995 |
| Separation Initialization Matrix | $\mathbf{W}(0) = \begin{bmatrix} 1.0 & 0.1 \\ 0.1 & 1.0 \end{bmatrix}$ |
| Initial moment estimates | $\hat{r}_{12}(0) = 0.01$, $\hat{\sigma}_1(0) = 0.01$, $\hat{\sigma}_2(0) = 0.01$ |

the original image dimensions, we can see very good separation. The Signal-to-Interference Improvement (SIRI) as defined in [9] is measured to be 42dB and 24dB respectively for two separated images (bottom images in Fig. 2)).

### 4.2   Case II

In this example, "APC" and "Peppers" images, with $\kappa_{s_{1,2}} = 1.4058, 1.8024$ (respectively), are mixed with (11). The kurtosis values of the two mixtures are $\kappa_{x_{1,2}} = 1.4259, 1.55$ (respectively). Note here that one mixture violates the critical assumption and as a result, only one source (Image 2) is fully separated with a measured SIRI of 40.55dB as can be seen in Fig. 3. The other output (Image 1) shows no separation and has a measured SIRI of 6.28dB.

During both experiments, we have noted that the performance of NKMA as applied to image mixtures, is very sensitive to the normalized step size, $\tilde{\mu}$. For different pairs of image mixtures, the optimal normalized step size, $\tilde{\mu}$ ranges in value between 0.0001 and 0.001. Even for the same image pair, the best SIRIs for each image may be achieved with different step sizes. As in case II, if the mixtures are processed through two parallel NKMA with two different properly chosen $\tilde{\mu}$, the SIRI can exceed 60dB in both separations. This thought to be due to the crude instantaneous estimate of the stochastic gradient of image data.

## 5   Conclusions

In this paper, we have extended a kurtosis maximization algorithm, successfully used in blind separation of instantaneous mixtures of speech signals, to images. The key features of the algorithm are its simplicity and performance as demonstrated by the examples.
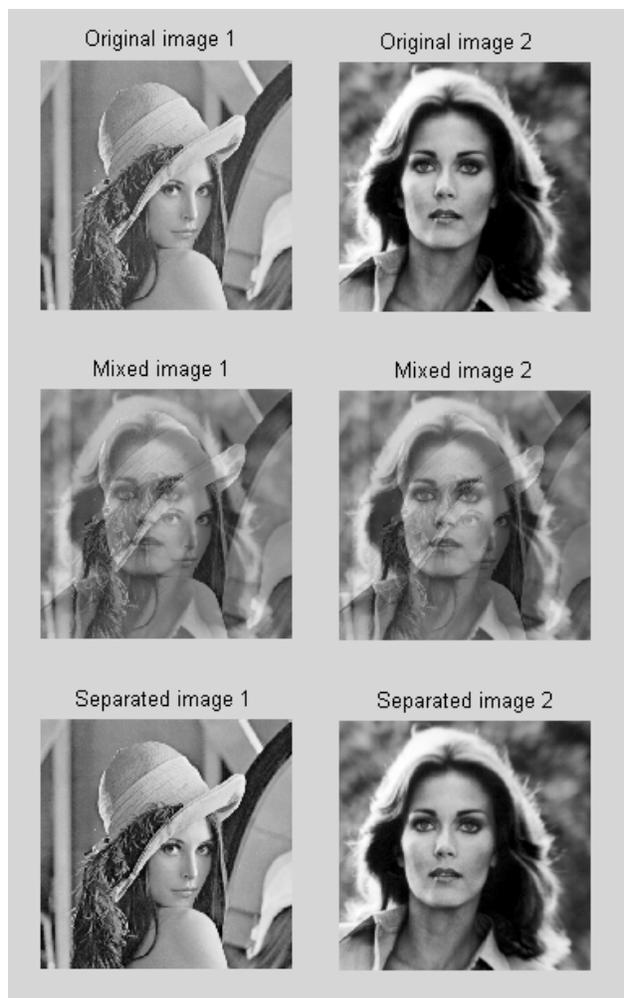
Figure 2: *"Lena" and "Lynda" image separation results.*



Figure 3: *"Apc" and "Peppers" image separation results.*

## References

[1] S. Haykin, ed., *Unsupervised Adaptive Filtering: Volume 1 Blind Source Separation*, John Wiley and Sons, Inc., New York, NY, 2000.

[2] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 3, pp. 320–327, May 2000.
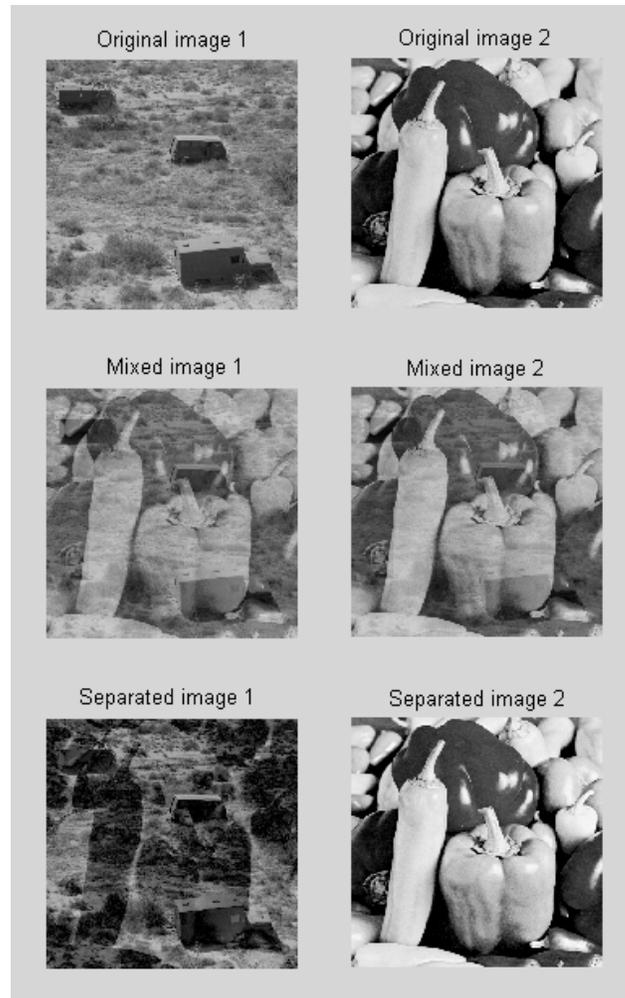
[3] H. Sahlin and H. Broman, "Blind separation of images," *Asilomar Conf. Sigs. , Sys., and Comps.*, 1996.

[4] A. Cichochi, W. Kasprzak, and S. Amari, "Neural network approach to blind separation and enhancement of images," *Proc. EUSIPCO*, 1996.

[5] J. Miskin and D. MacKay, "Ensemble Learning for Blind Image Separation and Deconvolution," in , *Advances in Independent Component Analysis*, M. Girolami, ed., Springer Verlag, New York, NY, 2000.

[6] R. Everson and S. Roberts, ed., *Independent Components Analysis: Principles and Practice*, Cambridge University Press, New York, NY, 2001.

[7] J. LeBlanc and P. De Leon, "Speech separation by kurtosis maximization," *Proc. ICASSP*, vol. 2, pp. 1029–1032, 1998.

[8] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: a deflation approach," *Sigal Processing*, vol. 45, pp. 59–83, 1995.

[9] P. De Leon and Y. Ma, "Normalized, hos-based, blind speech separation algorithms," *Asilomar Conf. Sigs. , Sys., and Comps.*, 2000.

[10] R. Klette and P. Zamperoni, *Handbook of Image Processing Operators*, John Wiley and Sons, Inc., New York, NY, 1996.