Comparative analysis of kurtosis and negentropy principles for music elements separation using Independent Component Analysis

Rithesh Kumar R[1] and Mohanaprasad k[1],

[1] VIT University, Vellore
rithesh.kumar7493@gmail.com

Abstract. Independent component analysis (ICA), also known as Blind Source Separation (BSS) is one of the widely used methods in separating a signal into its components. A music source with varying number of elements has been considered. The emphasis is on separating one of these tracks from the mixture of elements, i.e., the piece of music. This can be done using ICA by maximizing the non-gaussianity of the signal through the gradient algorithms of Negentropy and Kurtosis. Further, Singular Value Decomposition (SVD) technique has been used for whitening the signal. Their comparative results, with different number of sources have been analyzed using MATLAB simulation. It shows that, from the signal to interference ratio (SIR), Negentropy based ICA gives better separation with more number of elements, compared to Kurtosis principle.

Keywords: ICA, Kurtosis, Negentropy, SIR, BSS, SVD

1 Introduction

Independent Component Analysis (ICA) is a statistical method for transforming an observed multidimensional random vector into components that are as statistically independent from each other as possible. It has been used to perform signal separation in multiple domains such as Audio [7], Electrocardiogram (ECG) [6] and [8], Electroencepaholography (EEG) [9], signal and image analysis like feature extraction and clustering, Compression and redundancy reduction [11]. In this paper an attempt is made to separate the music elements which combine to form the music, that we hear in a composition.

Any source of music can be modeled as a composition of various elements, or in other words, instruments, in general. Even though these elements may have different principal frequencies, they are actually interspersed with other elements, across the frequency spectrum. This makes them an inseparable mixture, in terms of filtering. This makes us look for different mathematical, iterative algorithms that are used to separate the components. Out of the various methods that have been developed for this purpose, which include principal component analysis, projection pursuit, independent component analysis, factor analysis, etc., Independent component analysis is chosen due to its simplicity and robustness.

2. Independent Component Analysis

In the real world, many data often do not follow a Gaussian distribution, unlike the usual assumptions. They have super Gaussian distributions [1] and [10]. Hence, the distribution of the wave form and thereby its components can be found out, by finding out the gaussianity. Independent component analysis assumes that a given signal has a non-Gaussian distribution and is a mixture of statistically mutually independent components that are mixed in the following form:

$$x_i = a_{i1}s_1 + a_{i2}s_2 \dots\dots.. + a_{in}s_n \quad \text{for all } i = 1 \dots..n \qquad (1)$$

where a_{ij}, i,j=1…n are some real coefficients. The mixture is thus defined by:

$$X = A \times s \qquad (2)$$

where X is the random vector whose elements are the mixtures $x_1, x_2 \dots..x_n$; s is the random vector where $s_1, s_2 \dots. s_n$; represent the individual elements of the music signal and A is the vector of the real coefficients that represent the random mixing of the elements. As is evident, only the Vector X is known and both the vectors s and A are unknown. Now, our interest is to find out the vector s, where there exists a weight matrix W with coefficients w_{ij} such that,

$$S_i = w_{i1}x_1 + w_{i2}x_2 \dots\dots. + w_{in}x_n \quad \text{for all } i = 1 \dots..n \qquad (3)$$

It must be noted that, this model is only applicable if the signals are non-Gaussian and statistically independent. Now, the task is to find out the weight matrix that makes the sources as statistically independent as possible, which implies that this method is a special case of redundancy reduction [2].

ICA assumes that the independent components of the signal are maximally non-Gaussian. Hence, the weight matrix is the one that represents the maximum non-gaussianity point in the signal space distribution. Two of such methods available to find the maximized points of non-gaussianity are negentropy and Kurtosis.

The generalized algorithm for ICA is as follows:

1. Centralize the data: $X = X - E[X]$

   This step is carried out to make the mean of the signal zero, as this assumption simplifies the algorithms quite a lot.
2. Whitening: The next step is to make the signals uncorrelated, because independent signals are uncorrelated.
3. Finding the maximized non-Gaussian points.

2.1 Whitening using Singular Value Decomposition

To find the independent components, we must first uncorrelate the signal. This step will greatly simplify the ICA, if the observed mixture vectors are whitened [1]. A signal is said to be uncorrelated if it has unit variance and if the covariance matrix is identity. We must now find a linear transformation V that transforms the original signal X into a whitened vector z

$$z = U \times X \qquad (4)$$

where U is the unitary matrix obtained by SVD. We go for SVD because of its simplicity and versatility [3].

The algorithm [3] and [4] is given by
1. Find the non-zero Eigen values, λ_i of the matrix $X^T X$ and arrange them in descending order.
2. Find the orthogonal eigenvectors of the matrix $X^T X$ corresponding to the obtained Eigen values, and arrange them in the same order to form the column-vectors of the matrix V
3. Form a diagonal matrix S placing the square roots $s_i = \sqrt{\lambda_i}$ the p=*min*{M,N} first eigen values of the matrix $X^T X$ found in step (1) in descending order on theleading diagonal.
4. Find the first column-vectors of the matrix U: $u_i=s_i^{-1}Xv_i$ $(i=1: p)$.
5. Add to the matrix U the rest of M-p vectors using the Gram-Schmidt orthogonalization process

2.2 Kurtosis

It is the fourth order cumulant of the whitened signal and it is the measure of its non-gaussianity. The reason that a cumulant, rather than a moment, is measured stems from the fact that:

$$C_n(x_1+x_2)=C_n(x_1)+C_n(x_2) \tag{5}$$

$$M_n(x_1+x_2)\neq M_n(x_1)+M_n(x_2) \tag{6}$$

Where $C_n(.)$ and $M_n(.)$ represent the cumulant and moment of the random variable(music) respectively. This means, that the cumulant of the component mixture is equal to the sum of cumulant of each component (music sources/elements) [LMK]. Thus finding that point in the signal space where one of these cumulant is maximized, gives us one of the music elements, which is our goal. The kurtosis, which is the fourth order cumulant, is expressed as

$$kurt(y)=E\{y^4\}-3(E\{y^2\})^2 \tag{7}$$

with the assumption that the random variable is zero mean. Further, since we have whitened the signal, its variance is equal to 1. Which implies that $E\{y^2\}=1$ (zero mean). Thereby, we may also ignore the second term on the equation and it simplifies to

$$kurt(y)=E\{y^4\} \tag{8}$$

The algorithm for kurtosis is given by[10]
1. Initially, randomly assume a normalized value for the unit norm for w_p
2. Let $w_p \leftarrow E\{z[w_p{}^T z]^3\} - 3w_p\|w_p\|^2$, where z is the whitened signal.
3. Update the value of $w_p = w_p + 2\mathrm{e}^{-2}w_p$
4. Normalize $w_p \leftarrow w_p/\|w_p\|$
5. Check if the value has converged, i.e $\ll w_p w_p{}^T \gg = 1$. This gives the intended weight matrix w. Else, go to step 2

2.3 Negentropy

Again, this is also a measure of non-gaussianity for a random variable and it also includes higher order statistical information [1&5]. Negentropy is zero for a Gaussian variable. In this work, the gradient algorithm using negentropy has been applied. This algorithm is used to find out the maximized negentropy. This value corresponds to that value in the distribution space of the signal, where one of the independent components has the highest weight while the others have the least or close to zero values.

The negentropy algorithm [1] followed is:

1. Initially, randomly assume a normalized value for the unit norm for w_p
2. Let $w_p = E\{zg({w_p}^T z)\} - E\{g'({w_p}^T z)\}w_p$
   where z is the whitened signal, $g(y) = \tanh(y)$ and $g^{'}(y) = 1 - tanh^2(y)$
3. Normalize $w_p \leftarrow w_p/\|w_p\|$
4. Check if the value has converged, i.e. $\ll w_p {w_p}^T \gg = 1$. This gives the intended weight matrix w. Otherwise, repeat the algorithm

3 Simulation Procedures

The recording from multiple sources are taken. They represent a random mixture of the individual independent components. Now to separate these mixtures, we first center the signal. By the process of centering, it means that the Mean of the signal is made zero by subtracting every term with the mean value of the original signal. We do centering in order to whiten the signal. Whitening prepares the signal for the process of ICA, by making all the components uncorrelated with each other. This process is validated by the fact that, independent signals are uncorrelated signals. So, by un-correlating the obtained signal, we are distinguishing the individual components before separating them. This situation is validated by the fact that the co-variance matrix of a whitened signal is and Identity matrix, because the independent components are only correlated with themselves.

This procedure is followed by finding either kurtosis value or the negentropy value. Both these process give us that point in the signal density function, where only one component of the signal is maximized, while the others tend to zero. So, filtering the signal through this point must ideally give us any one of the individual components.

However, due to the randomness associated with this process, we cannot obtain the ideal output, and the separation also would not be perfect. The results that are discussed in the following are the desirable results that are achieved through multiple iterations and we go on to compare the different processes of ICA through them.

The same method was applied for a mixture of two, three and four components. The relative quality of separation of both the mixtures was measured using the Signal to Interference Ratio (SIR). SIR is defined as the ratio between the power of the desired signal to the actual output signal. In an ideal system, the SIR is 0 dB. The closer the SIR is towards zero, it obviously implies a better quality of output. The obtained results and hence their inferences are discussed.

4 Results

A piece of music with piano, pad, drum and flute elements was chosen. They were samples at the standard frequency of 44100 *Hz* with a period of 7 seconds. The original signals of those components are shown in figs. 1, 2, 3 and 4.

These signals were originally mixed randomly and they were separated by the algorithm. For a two component input, the input signal that was whitened is shown in fig. 5.

As is observed from the fig. 6, the output obtained is that of the drum loop and it is very close correlation to the original signal, both for negentropy and kurtosis. The SIR in the case of the former is -9.32 dB while that of the latter is -4.24 dB. It is observed that in many a case, kurtosis obtains a better result.

To extend the process to three signals, the signals, piano, drum and pad were mixed randomly. Fig. 8 shows the whitened input for the three component mixture. In this case, both the methods show similar efficiency in results. A close examination suggests that although kurtosis still gives a higher SIR, the perceptional quality is much better in the case of negentropy. The average SIR for the method of Kurtosis is around -5.00 dB, while for negentropy it is around -3.50 dB.

The average SIR for the method of Kurtosis is around -5.00 dB, while for negentropy it is around -3.50 dB. Figs. 9 and 10 show the piano component output for kurtosis and negentropy methods respectively. To further validate the methods for a complex four component mixture, all the elements that have been considered were mixed randomly and subsequently the algorithm was applied.

Even for a mix of four signals, both the principles gave a satisfactory output. Yet the certainty that the separated signal is the source was considerably low. Negentropy gave better results with SIRs as high as -2.66 dB while kurtosis algorithm gave results with SIRs around -5.00 dB as shown in table 1.

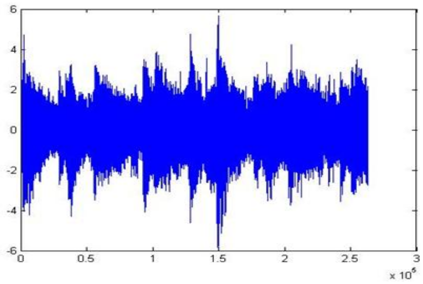

Fig. 1 Source1: Piano

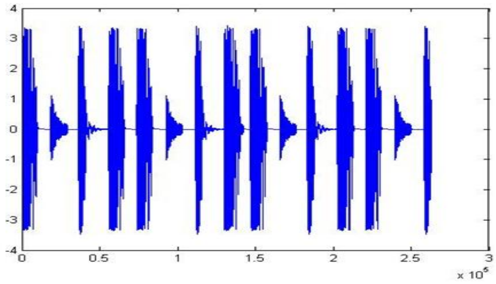

Fig. 2 Source2: Drum Loop

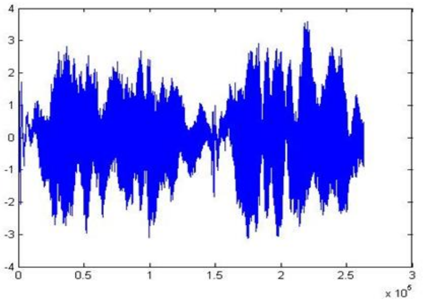

Fig. 3 Source3: Pad

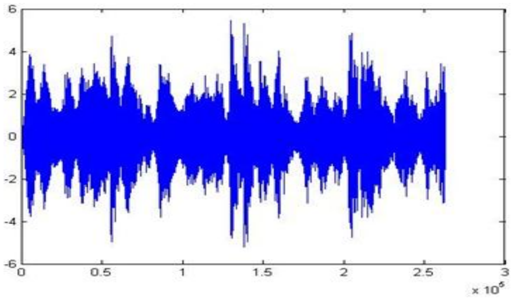

Fig. 4 Source4: Flute

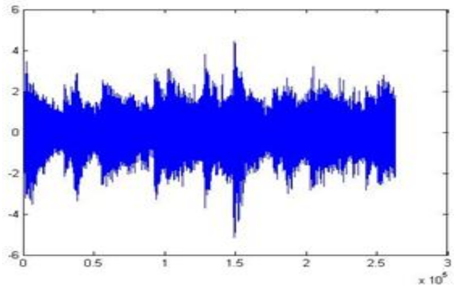

Fig. 5 Whitened Input for two source mixture (drums and piano)

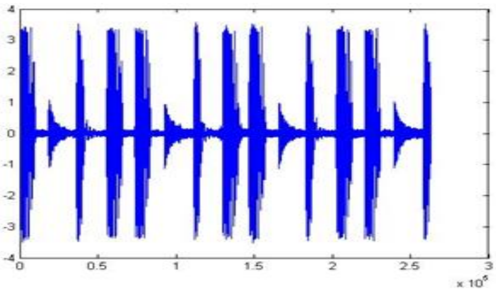

Fig. 6 Output for two source mixture using kurtosis

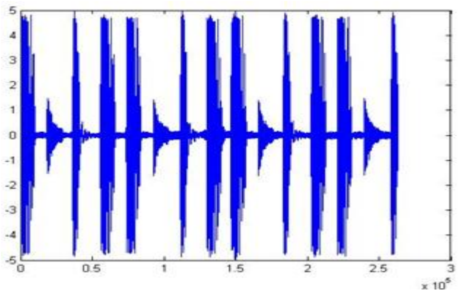

Fig. 7 Output for a mixture of two sources using negentropy

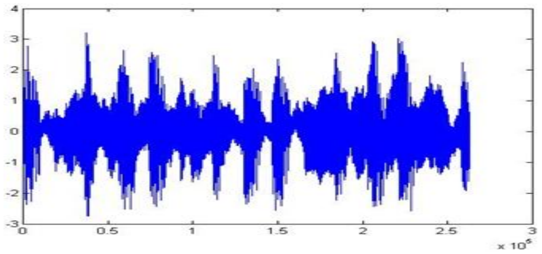

Fig. 8 Whitened Input for three source mixture (piano, drums and pad)

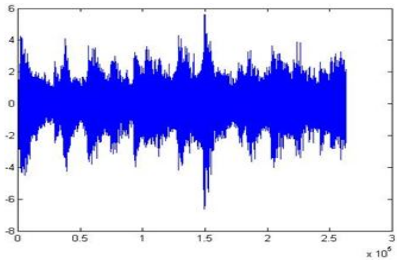

Fig. 9 Output for three source mixture using kurtosis

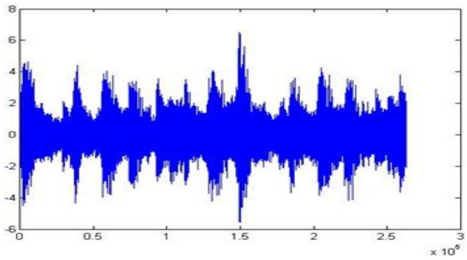

Fig. 10 Output for a mixture of three sources using negentropy

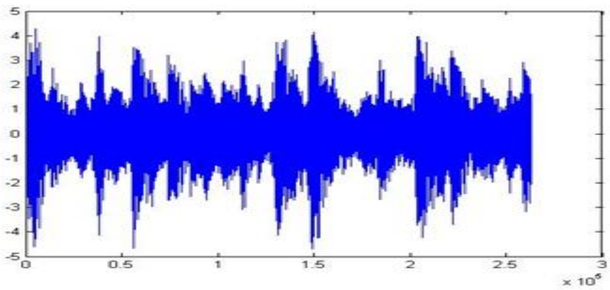

Fig. 11 Whitened Input for four source mixture (piano, drums, pad and flute)

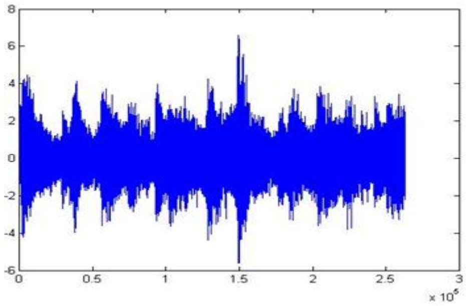

Fig. 12 Output for four source mixture using kurtosis

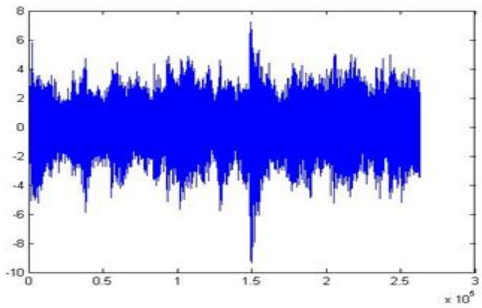

Fig. 13 Output for a mixture of four sources using negentropy

Table 1. SIR analysis

No. of sources	Signal to Interference Ratio (dB)	
	Kurtosis	Negentropy
2	-4.2380	-9.3205
3	-5.8434	-3.4968
4	-5.4540	-2.6590

5 Conclusion

Through numerous iterations and analysis, it is understood that both Kurtosis and Negentropy gradients are efficient in separating the sources from the music mixture. When there are more number if sources (elements), Negentropy gives a better result, while the converse is true for Kurtosis. The average SIR values obtained for all the cases considered is -3.95 dB for Kurtosis and -6.16 dB for negentropy. Kurtosis method is a lot simpler and much more efficient. However, for a complex multi element mixture, the negentropy method proves to be more capable.

References

[1] Aapo Hyvtirinen, JuhaKarhunen, Erkkioja: IndependentComponent Analysis. John Wiley & Sons, Inc., New York, (2001).
[2] Aapo Hyvtirinen: Fast and robust fixed-point algorithms for independent component analysis. IEEE trans. Neural Networks, 10(3), 626-634 (1999)
[3] Dan Kalman: A Singular Valuable Decomposition: The SVD of a Matrix. Lecture document, The American University, Washington(2002)
[4] Clifford G D: Singular Value Decomposition & Independent Component Analysis. J. Biomed. Signal and Image Proce. Spring, 1-42 (2005)
[5] P. Comon: Independent component analysis—a new concept. Signal Processing, 36:287–314 (1994)
[6] Bell AJ, Sejnowski T J: An information-maximization approach to blind separation and blind deconvolution. Neural Computation, 7(6):1129-1159(1995).
[7] Toch B, Lowe D, Saad D: Watermarking of audio signals using ICA. In Third InternationalConference on Web Delivering of Music, 8, 71-74(2003).
[8] Barros A, Mansour A, Ohnishi N: Adaptive blind elimination of artifacts in ECG signals. In Proceedings of I and ANN (1998).
[9] Makeig S, Bell A J, Jung T P, Sejnowski T J: Independent component analysis of electroencephalographic data. In Touretzky DS, Mozer MC, Hasselmo ME (eds.), Advances in Neural Information Processing Systems, volume 8. The MIT Press, 145-151(1996).
[10] Mohanaprasad K, Arulmozhivarman P: Comparison of Independent Component Analysis techniques for Acoustic Echo Cancellation during double talk scenario. Australian Journal of Basics and applied Sciences, 7(4), 108-113(2013).
[11] Hyv¨arinen A. Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation. Neural Computation 11(7):1739-1768(1999).