



Proposed Integration Algorithm to Optimize the Separation of Audio Signals Using the ICA and Wavelet Transform

Enrique San Juan¹, Ali Dehghan Firoozabadi^{2(✉)}, Ismael Soto¹,
Pablo Adasme¹, and Lucio Cañete¹

¹ Electrical Engineering Department, Universidad de Santiago de Chile,
Santiago, Chile

{enrique.sanjuan, ismael.soto, pablo.adasme,
lucio.canete}@usach.cl

² Department of Electricity, Universidad Tecnológica Metropolitana,
Av. Jose Pedro Alessandri 1242, 7800002 Santiago, Chile
adehghanfiroozabadi@utem.cl

Abstract. In the present work, an integration of two combined methodologies is developed for the blind separation of mixed audio signals. The mathematical methodologies are the independent component analysis (ICA) and the discrete Wavelet transform (DWT). The DWT optimizes processing time by decreasing the amount of data, before that signals are processed by ICA. A traditional methodology for signal processing such as Wavelet is combined with a statistical process as ICA, which assumes that the source signals are mixed and they are statistically independent of each other. The problem refers to very common situations where the human being listens to several sound sources at the same time. The human brain being able to pay attention to the message of a particular signal. The results are very satisfactory, effectively achieving signal separation, where only a small background noise and a attenuation in the amplitude of the recovered signal are noticed, but that nevertheless the signal message is identified in such a way.

Keywords: Voice recognition · Wavelet transform · Voice processing · Independent component analysis

1 Introduction

In the last two decades the wavelet transform has positioned itself as a powerful tool for digital signal processing complementing the Fourier transform and in some cases more powerful, since the wavelet transform performs an analysis both in time and in the frequency at the same time. On the other hand, the independent component analysis (ICA) provides an analysis perspective of the signals from the statistical point of view. In relation to the combined use of the Wavelet transform and the ICA, some related publications are shown below.

In [1], an adaptive hybrid algorithm based on DWT and ICA is proposed to remove noise from images obtained by means of magnetic resonance imaging, where this

combined technique is compared to conventional techniques, such as DWT, undecimated discrete wavelet transforms (UDWT) and ICA. In [2], the idea of energy efficiency state identification is proposed and the monitoring strategy of energy efficiency state is established for a metal cutting process. A combined application method of continuous wavelet transform (CWT) and fast independent component analysis (FICA) is proposed for feature extraction of low or high energy efficiency state. On the other hand in [3] an electrocardiogram (ECG) signal noise elimination method is proposed based on wavelet transformation and ICA. First, two-channel ECG signals are acquired. We decompose these two ECG signals by wavelet and adding the useful wavelet coefficients separately, obtaining two-channel ECG signals with fewer interference components. Second, these two-channel ECG signals are processed and a channel signal constructed to perform an additional process with ICA, obtaining the separate ECG signal. In [4], a possible solution to the problem is proposed. Forensic speaker verification systems show severe performance degradation in the presence of noise when the signal to noise ratio (SNR) is low. Also use a combination of feature warped Mel frequency Cepstral coefficients (MFCCs) and feature warped MFCC extracted from the DWT of the enhanced speech signals as the feature extraction. In [5], a method for separation of mixed signals using ICA and Wavelet transform is proposed. This problem is solved using DWT based parallel architecture, which is a combined system consisting of two sub-over complete ICA. One process takes the high-frequency wavelet part of observations as its inputs and the other process takes the low-frequency part. Then, the final results are generated by merged these results. In [6], authors proposed a new wavelet based ICA method using Kurtosis for blind audio source separation. In this method, the observations are transformed into an adequate representation using wavelet packets decomposition and Kurtosis criterion.

2 Independent Component Analysis (ICA)

There is an essentially statistical method for signals which simulates the way that human brain is able to differentiate a particular signal when it is listening to several signals at the same time (Fig. 1). We use the premise of the statistical independence from the origin signals. It means the signals are mixed with a noise which has a probability function and subsequently through ICA to obtain the original independent signals.

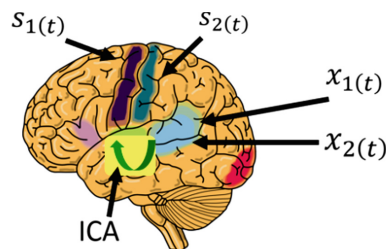


Fig. 1. The human brain according to ICA

2.1 Operation of the ICA Method

Usually when there is a social meeting and two persons are talking, the brain of the person who is listening only concentrates on the voice of the person with whom he is talking and keeping his main characteristics of the signal [7]. The human brain is able to isolate and separate the different sounds that it is listening in order to concentrate only on the voice that is interested (Fig. 2). This ability to separate and recognize a sound source from a noisy environment is known as the Cocktail Party effect technique. This acoustic phenomenon of human psychology was initially proposed by Colin Cherry [8].

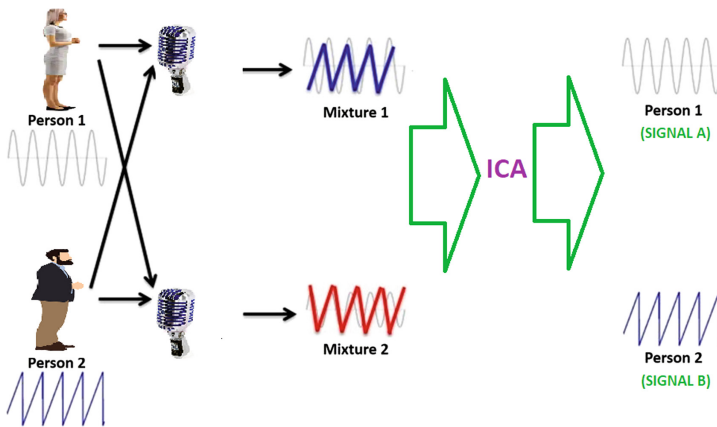


Fig. 2. Cocktail Party effect applied with ICA

In the following example there are two speakers **A** and **B** that each of them is speaking independently. The signals are separated by a mathematical process where two acoustic signals s_a and s_b are generated. Also, The signals are mixed and next to them there is a microphone that record these signals independently [9]. The ICA method takes these signals (generated independently) which are mixed with random noise and then using a statistical process (ICA) to recover the two original signals separately (Fig. 2).

This process can be observed in greater detail when the signals **A** and **B** (Fig. 3) are mixed linearly through a linear function $s = 0.7 * A - 0.11 * B$ to generate two new mixed signals M_A and M_B [10].

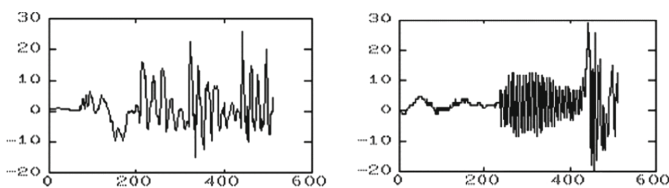


Fig. 3. Independent signals A and B

Once they are separated, the ICA method is applied to the processed signals M_A and M_B . When the process is finished, it is recovered with the two independent signals except for a very low noise factor called scalar which can be positive or negative (Fig. 4).

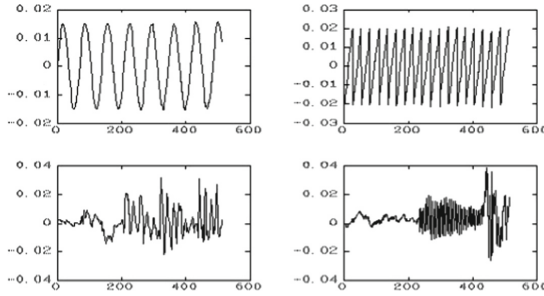


Fig. 4. Recovered signal

2.2 Formulation of the ICA Method for Discrete Signals

Let's consider $x = [x_1, x_2, \dots, x_n]$ as a vector of n components that are the samples of a signal. Also, we assume this vector is produced as a linear combination of n independent denoted signals $s = [s_1, s_2, \dots, s_n]$.

This can be expressed in Eq. 1:

$$\begin{aligned} x_1 &= a_{1,1}s_1 + a_{1,2}s_2 + \dots + a_{1,m}s_m \\ x_2 &= a_{2,1}s_1 + a_{2,2}s_2 + \dots + a_{2,m}s_m \\ x_m &= a_{m,1}s_1 + a_{m,2}s_2 + \dots + a_{m,m}s_m \end{aligned} \quad (1)$$

where the coefficients a_n determine a matrix A which is known as the mixing matrix, vector x is called the vector of mixtures and vector s is the vector of the independent components to be processed. Then, it can be described using matrix notation as:

$$x = As \quad (2)$$

In practice, only the vector of mixtures x that is generated when sampling a dialogue event for example is known. Therefore, the ICA method consists in applying the algorithm that allows to find a mixed desiccation matrix W in such a way that $y = Wx$ is a good approximation of the vector s ($y \cong s$).

3 Wavelets Process

The wavelet transform has as main characteristics that is low computational and algorithmic complexity and the high probability of recovering small details in the inverse transformation [11]. These are very interesting characteristics for human voice signals. In addition, the process for the compression of the sampled data is the most efficient to optimize the computation time of the ICA method [12]. Another reason why wavelet was selected for use in research is its conduction in the areas of medicine, especially for the separation of brain signals where have good results [13] in combination with ICA. The mother wavelet that is selected for this work is the HAAR because it is appropriate to look for the points of inflection and discontinuities in the signals [14].

3.1 HAAR Wavelet Transform

The HAAR wavelet functions provides an optimal balance between the resolutions in time and frequency spaces. It is more efficient tool to compute the wavelet transform. This wavelet is the most used since it is reduced to calculate averages (sums) and changes (differences) between the data. Every mother wavelet has an associated scale function φ . In the case of the HAAR wavelet, the scale function is represented as:

$$\varphi_{u,v}(t) = \begin{cases} 1 & \text{if } u \leq x < v \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The mother scale function is defined as:

$$\varphi_{0,1} = \varphi_{0,\frac{1}{2}} + \varphi_{\frac{1}{2},1} \quad (4)$$

Using the same scale function, the basic HAAR wavelet function is considered as:

$$\varphi_{0,1} = \varphi_{0,\frac{1}{2}} - \varphi_{\frac{1}{2},1} \quad (5)$$

These two previous definitions multiplied by the coefficients 1 and 2 allow to generate the wavelet coefficients with the data of our sample in a way using only averages and differences between the data. If we have a vector s of m data with $m = 2n$, we calculate the first level of wavelet coefficients as:

$$\begin{aligned} c^1 &= \left[\frac{s_1 + s_2}{2}, \frac{s_3 + s_4}{2}, \dots, \frac{s_{m-1} + s_m}{2}, \frac{s_1 - s_2}{2}, \frac{s_3 - s_4}{2}, \dots, \frac{s_{m-1} - s_m}{2} \right] \\ &= \frac{1}{2} [s_1 + s_2, s_3 + s_4 + \dots + s_{m-1} + s_m, s_1 - s_2, s_3 - s_4, \dots, s_{m-1} - s_m] \end{aligned} \quad (6)$$

It is shown that the vector c^1 is divided into two parts. The first part corresponds only to the averages of two elements of s and the second section corresponds to the differences two of the elements of s .

For the next wavelet level, the first section of the vector c^1 is taken and the previous procedure is repeated. The second part of the vector is only multiplied by scale factor. We notice the elements of c^1 as the next level of wavelet coefficients will be the vector c^2 contains in the first quarter, the averages of the elements of s in the second quarter, and the differences of the averages of s and the second half only the differences of the elements of s .

$$c^1 = [c_1^+, c_2^+, \dots, c_m^+, c_1^-, c_2^-, \dots, c_m^-] \quad (7)$$

The procedure can be repeated up to n times, which is the number of subdivisions in half that can be made from the vector s . Therefore, in each subsequent level sums and differences of each sub-vector of averages are calculated. In the following, the algorithm for calculating the coefficients of the wavelet HAAR is presented.

3.2 Inverse Wavelet HAAR

It is possible to return by performing simple addition and subtraction operations again in terms of the mother wavelet function and the mother scale function, when calculating the wavelet coefficients up to some level t . To obtain the original data again, it can be expressed through:

$$\frac{1}{2}(\varphi_{0,1} + \psi_{0,1}) = \varphi_{0,\frac{1}{2}}, \quad \frac{1}{2}(\varphi_{0,1} - \psi_{0,1}) = \varphi_{\frac{1}{2},0} \quad (8)$$

In terms of the data generated by performing a wavelet level (c^1), we obtained the data as:

$$c + 1 = s_1 + s_{12}, \quad c - 1 = s_1 - s_2 \quad (9)$$

The original data is also recovered as:

$$s_1 = C_1^+ + C_1^-, \quad s_1 = C_1^+ - C_1^- \quad (10)$$

This process can be repeated with all the data of a c_t vector until all the data of the previous wavelet level is recovered.

3.3 Data Compression Process Through Wavelet

The simple way is known as the wavelet HAAR allows to calculate wavelet coefficients. It is possible to perform a compression of the data without losing relevant information of them [15]. As was explained, in the HAAR wavelet, additions and subtractions are made. As several c_t levels of wavelet coefficients are calculated, the energy of the signal s is concentrated in the first section of the vector c_t and to the left of the vector. The result of the vector c_8 shows (8 wavelet levels) the coefficients decay approach to zero (Fig. 5).

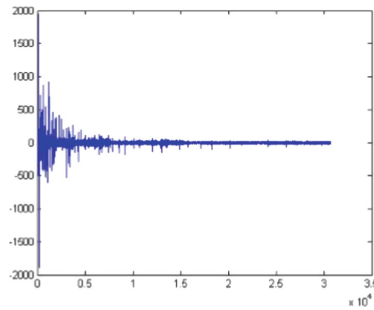


Fig. 5. Density functions of the mixed signals.

4 Implementation Between the ICA and Wavelet Methods

Once these two techniques are described (wavelet transform and the ICA technique), they must be combined in an optimal way for a better compression of the sampled data and trying to avoid the loss of information in the communication between these processes. To obtain the two virtual data streams, an arithmetic mean between data is applied by means of such a way to generate two equal data streams in their creation and an independent requirement to be able to use the ICA process. At the time of processing the information through ICA and in order to reduce the amount of sampled data that have to be calculated by the ICA algorithm, the HAAR wavelet algorithm was applied to further reduce the data that ICA needs.

4.1 Process of Generation of Data Trains

This process must be done using the sample taken from a data source and must create two equal samples, but linearly independent between them (Fig. 6).

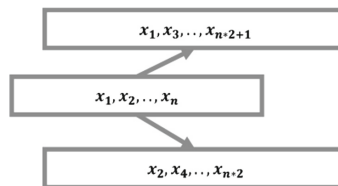


Fig. 6. Mixing process of the implemented algorithm

For this process we should look for the sampled elements that are missing in each of the data streams. These coefficients are searched using a simple arithmetic mean between the previous and above to generate the data stream. When generating the two sampled data streams, they are independent and then processed by the wavelet transform and this generates coefficients that will be taken by the ICA method as shown in Fig. 7.

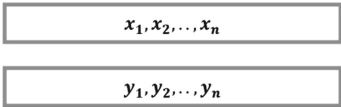


Fig. 7. Arrangements that are created independently

5 Results and Discussions

5.1 Computational Analysis Using ICA and Wavelet Techniques

The algorithm complexity evaluation technique is used to calculate the number of required operations. It focuses on classifying the complexity of the processes and performing the algorithm of the ICA method. It determines that the order is $120\mathbf{m}6\mathbf{n}6$ where \mathbf{n} is the number of signals received and \mathbf{m} the number of iterations of the algorithm. The number of operations necessary for the data reduction process using wavelet HAAR is of the order \mathbf{mn} ($6\mathbf{t}-12*2\mathbf{t}$), where \mathbf{t} is the number of wavelet levels processed [16]. The calculation of the total number of operations for the wavelet analysis was made, but it is negligible compared to the operations necessary for ICA. These results give a linear polynomial as explained above. They show the potential of wavelet analysis to generate good performance, reducing the computational cost.

5.2 Graphic Interface

For this work two algorithms were implemented, one for signal separation using ICA and another wavelet supported by the platform of a graphical interface that allows to visualize and compare part of the compression process of audio signals. The related process is illustrated through a graph to determine the similarity of the signals. However, to compare the transformations and their application in each reading it is more practical to develop each of the steps and display them visually through a graphic.

5.3 Process Time Between Algorithms

The result of the algorithms (the running time in micro second) are presented in Table 1. As seen in this Table, the used time by wavelet is better in larger files and in smaller files the ICA is better. It can be deduced that for a large data the ICA has more calculations to perform, but if the compression of the data with the wavelet technique is used, it has fewer calculations.

Table 1. The result of the algorithms (ICA and ICA+wavelet)

ICA	ICA+WAVELET	Files sizes
1995 μs	2334 μs	10 seg.
4123 μs	4256 μs	20 seg.
6234 μs	5232 μs	30 seg

6 Conclusions

The ICA method is a very effective tool for the blind separation of signals, especially if it is combined with other techniques such as the wavelet transform. In the present study, this statistical mathematical model was developed to separate audio signals using the ICA method, which are generated independently. So that these signals do not have a statistical dependency. This process is supported by the wavelet transform to decrease the processing data. The integration of these methods is performed for the separation of audio signals in which the processing time is optimized using previously the wavelet technique before the use of the ICA algorithm. Because the input signals to the ICA algorithm must be statistically independent, these signals are artificially generated. In this way, the problem is faced where the human being listens to several sound sources at the same time, having the ability to pay attention to the signal coming from one in particular. These signals, which are transported through the ear canal, can be separated and identified by the human being. However, the developed application gives a clear example of the versatility of the wavelet technique properly combined with ICA.

Acknowledgment. The authors acknowledge the financial support of Projects: FONDECYT No. 11180107, FONDECYT Postdoctorado No. 3190147 and Dicyt Project 061713AS.

References

1. Rai, H.M., Chatterjee, K.: Hybrid adaptive algorithm based on wavelet transform and independent component analysis for denoising of MRI images. *Measurement* **144**, 72–82 (2019)
2. Cai, Y., Shi, X., Shao, H., Yuan, J.: Energy efficiency state identification based on continuous wavelet transform-fast independent component analysis. *J. Manuf. Sci. Eng.* **141** (2), 021012 (2019)
3. Liu, M., Liu, M., Dong, L., Zhao, Z., Zhao, Y.: ECG signals denoising using wavelet transform and independent component analysis. In: *International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology*, p. 962213 (2015)
4. Al-Ali, A.K.H., Dean, D., Senadji, B., Baktashmotlagh, M., Chandran, V.: Speaker verification with multi-run ICA based speech enhancement. In: *11th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, pp. 1–7 (2017)
5. Chao, M., Xiaohong, Z., Hongming, X., Dawei, Z.: Separation of mixed signals using DWT based overcomplete ICA estimation. In: *Fourth International Conference on Computational Intelligence and Communication Networks*, Mathura, pp. 416–419 (2012)
6. Mirarab, M.R., Dehghani, H., Pourmohammad, A.: A novel wavelet based ICA technique using Kurtosis. In: *2nd International Conference on Signal Processing Systems*, Dalian, pp. V1-36–V1-39 (2010)
7. Zhao, Z., Xie, T., Zhao, Z., Kong, X.: A blind source separation method based on the time delayed correlations and the wavelet transform. In: *8th International Conference on Signal Processing*, Beijing, China, vol. 1 (2006)

8. Mirarab, M.R., Sobhani, M.A., Nasiri, A.A.: A new wavelet based blind audio source separation using Kurtosis. In: 3rd International Conference on Advanced Computer Theory and Engineering, Chengdu, China, pp. V1-550–V1-553 (2010)
9. Hamada, T., Nakano, K., Ichijo, A.: Wavelet-based underdetermined blind source separation of speech mixtures. In: International Conference on Control, Automation and Systems, Seoul, South Korea, pp. 2790–2794 (2007)
10. Yilmaz, O., Rickard, S.: Blind separation of speech mixtures via time-frequency masking. *IEEE Trans. Sig. Process.* **52**(7), 1830–1847 (2004)
11. Debnath, L.: *Wavelet Transforms and Their Applications*. Birkhäuser, Boston (2002)
12. Osowski, S., Linh, T.H.: ECG beat recognition using fuzzy hybrid neural network. *IEEE Trans. Biomed. Eng.* **48**, 1265–1271 (2001)
13. Prasad, G.K., Sahambi, J.S.: Classification of ECG arrhythmias using multi-resolution analysis and neural networks. In: IEEE Conference on Convergent Technologies (Tecon 2003), Bangalore, India, pp. 227–231 (2003)
14. Abou-Elseoud, A., Starck, T., Remes, J., Nikkinen, J., Tervonen, O., Kiviniemi, V.J.: The effect of model order selection in group PICA. *Hum. Brain Mapp.* **31**(8), 1207–1216 (2010)
15. Addison, P.S.: *The Illustrated Wavelet Transform Handbook*. IOP Publishing, Bristol (2002)
16. Carmona, R., Huang, W.L.: *Practical Time-Frequency Analysis Gabor and Wavelet Transforms with an Implementation in S*. Academic Press, USA (1998)