A cascade form blind source separation connecting source separation and linearization for nonlinear mixtures

著者	Nakayama Kenji, Hirano Akihiro, Nishiwaki Takayuki
journal or publication title	Proceedings of the International Joint Conference on Neural Networks
volume	3
page range	1856-1861
year	2003-07-01
URL	http://hdl.handle.net/2297/6855

A CASCADE FORM BLIND SOURCE SEPARATION CONNECTING SOURCE SEPARATION AND LINEARIZATION FOR NONLINEAR MIXTURES

Kenji Nakayama *Akihiro Hirano* *Takayuki Nishiwaki*

Dept. of Information Systems Eng., Faculty of Eng., Kanazawa Univ.
2-40-20, Kodatsuno, Kanazawa, 920-8667, Japan
nakayama@t.kanazawa-u.ac.jp

ABSTRACT

A network structure and its learning algorithm have been proposed for blind source separation applied to nonlinear mixtures. The network has a cascade form consists of a source separation block and a linearization block in this order. The conventional learning algorithm is employed for the separation block. A new learning algorithm is proposed for the linearization block assuming 2nd-order nonlinearity. After, source separation, the outputs include the nonlinear components for the same signal source. This nonlinearity is suppressed through the linearization block. Parameters in this block are iteratively adjusted based on a process of solving a 2nd-order equation of a single variable. Simulation results, using 2-channel speech signals and an instantaneous nonlinear mixing process, show good separation performance.

1. INTRODUCTION

Recently, many kinds of information are transmitted and processed through world wide communications. Communication terminals are used under a variety of environments. At the same time, high quality is required. For this reason, signal processing including noise cancelation, echo cancelation, equalization of transmission lines, restoration of signals have been becoming very important technology. In some cases, we do not have enough information about signal sources and interference. Furthermore, their mixing process and transmission process are not well known in advance. Under these situations, blind source separation using statistical property of the signal sources has become important.

Jutten et all proposed a blind source separation algorithm based on statistical independence and symmetrical distribution of the signal sources [1]-[8]. Two kinds of stabilization methods have been proposed for Jutten's method [9],[10]. Convolutive mixture models have been discussed [11],[12]. Convergence and separation performances are highly dependent on relation between a probability density function of the output signals and nonlinear functions, which are used in updating coefficients in a separation block. Optimum nonlinearity has been discussed [13], [14], [15], and adaptive nonlinear functions have been proposed [16], [17], [18].

In actual applications, mixing processes include nonlinearity, such as loud speakers. In this case, signal sources are mixed in a complicated manner, and are difficult to be separated. In these problems, both source separation and linearization are simultaneously required. One way to model a nonlinear mixture is a combination of a linear mixing process and a nonlinear transform in a cascade form. In a separation block, a linearization process and a linear separation process are arranged in this order. Spline nonlinear functions or spline neural networks have been applied to the linearization process [19], [20]. Furthermore, a maximum likelihood estimator has been applied [21]. However, separation performance is not enough.

In this paper, an approach is proposed, in which a linear separation process and a linearization process are arranged in this order. First, the signal sources, which include nonlinearity, are separated based on their statistical independency. In the linearization process, the nonlinear components in each signal source are suppressed through an iterative learning algorithm. Simulation using 2-channel speech signals and 2nd-order nonlinearity will be shown to confirm usefulness of the proposed method.

2. CASCADE FORM BLIND SOURCE SEPARATION

2.1. Network Structure

A proposed cascade form blind source separation (BSS) is shown in Fig.1. The nonlinear mixture model is the same as in [19], [20], [21]. First, the signal sources s_i are mixed through linear combination resulting in u_j. After that, they are transmitted through nonlinear functions F_j resulting in x_j. In the BSS block, a linear source separation process and a linearization process are arranged in this order. In the conventional methods, they are arranged in the reverse order.

In the proposed method, the number of the observations

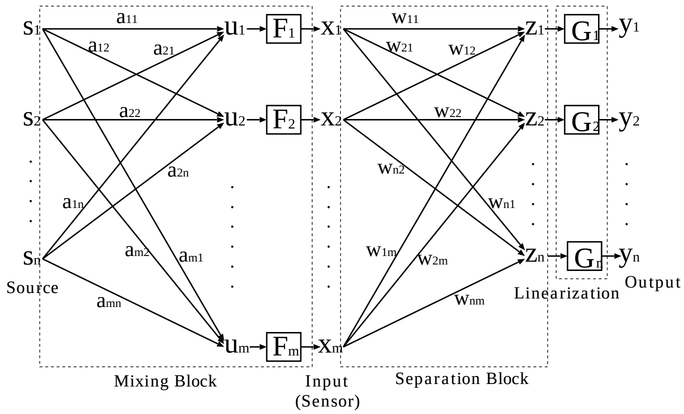

Fig. 1. Network structure of proposed cascade form BSS.

x_j are increased from that of the signa sources s_i in order to increase conditions used in cancelling the nonlinear terms in the linear separation block. For example, s_2, s_2^2 and $s_1 s_2$ are cancelled in z_1, and s_1, s_1^2 and $s_1 s_2$ are cancelled in z_2, respectively.

If linear separation is complete, its outputs z_i include only a single signal source and its nonlinear components. This nonlinearity is suppressed through the linearization block.

2.2. Example of Linear Separation and Linearization

One example is shown here. Two signal sources and 2nd-order nonlinearity are used. In this case, four observations $x_1 \sim x_4$ are required. They are expressed by

$$x_1 = a_{11}s_1 + a_{12}s_2 + a_{13}s_1^2 + a_{14}s_1 s_2 + a_{15}s_2^2 \quad (1)$$
$$x_2 = a_{21}s_1 + a_{22}s_2 + a_{23}s_1^2 + a_{24}s_1 s_2 + a_{25}s_2^2 \quad (2)$$
$$x_3 = a_{31}s_1 + a_{32}s_2 + a_{33}s_1^2 + a_{34}s_1 s_2 + a_{35}s_2^2 \quad (3)$$
$$x_4 = a_{41}s_1 + a_{42}s_2 + a_{43}s_1^2 + a_{44}s_1 s_2 + a_{45}s_2^2 \quad (4)$$

The mixing process is assumed to be linearly independent. Thus, from the above equations, the cross term $s_1 s_2$ can be cancelled, and two independent outputs z_1 and z_2 can be obtained. They still include the 2nd-order term s_1^2 and s_2^2, respectively.

$$z_1 = b_{11}s_1 + b_{12}s_1^2 \quad (5)$$
$$z_2 = b_{21}s_2 + b_{22}s_{2^2} \quad (6)$$

Since z_1 and z_2 include only s_1 and s_2, respectively, they can be linearized through the following nonlinear functions.

$$y_1 = G_1(z_1) = \frac{-b_{11} \pm \sqrt{b_{11}^2 + 4b_{12}z_1}}{2b_{12}} \quad (7)$$
$$y_2 = G_2(z_2) = \frac{-b_{21} \pm \sqrt{b_{21}^2 + 4b_{22}z_2}}{2b_{22}} \quad (8)$$

Finaly, the separated and linearized signal sources are obtained.

$$y_1 = c_{11}s_1 \quad (9)$$
$$y_2 = c_{12}s_2 \quad (10)$$

c_{11} and c_{12} are some constant.

3. LEARNING ALGORITHMS

3.1. Linear Separation Block

If s_1 and s_2 are statistically independent, then $a_{11}s_1 + a_{12}s_1^2$ and $a_{21}s_2 + a_{22}s_2^2$ are also independent. They can be separated through the conventional learning algorithms for linear mixtures. So, the learning algorithm based on likelihood estimation [22], [23], [24] is employed in this paper.

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \eta[\boldsymbol{\Lambda}(t) - \varphi(\boldsymbol{z}(n))\boldsymbol{z}^T(n)]\boldsymbol{W}(n) \quad (11)$$

η is a learning rate, $\boldsymbol{\Lambda}(t)$ is a diagonal matrix, and $\varphi()$ is a nonlinear function [18].

3.2. Linearization Block

Transformations in the linearization block are given by Eqs.(7) and (8). However, in real applications, the coefficients b_{ij} are not known. So, they should be adjusted through an iterative method. Equations (7) and (8) can be expressed by using two parameters as follows:

$$y_i(n) = -\frac{\alpha_i}{2} \pm \sqrt{\frac{\alpha_i^2}{4} + \frac{z_i(n)}{\beta_i}} \quad (12)$$
$$\alpha_i = \frac{b_{i1}}{b_{i2}} \quad (13)$$
$$\beta_i = \frac{1}{b_{i2}} \quad (14)$$

α_i and β_i are adjusted through an iterative method.

Error Function:
In this paper, 2nd-order nonlinearity is assumed. Thus, after the linear source separation, the outputs include 1st-order and 2nd-order terms of the signal sources. Furthermore, if we take speech and music signals into account, their average is almost zero. Therefore, the output average can be used as a cost function.

$$E_i(n) = \frac{1}{M}\sum_{l=0}^{M-1} y_i(n-l) \quad (15)$$

The gradient descent algorithm is used for adjusting the parameters.

$$\alpha_i(n) = \alpha_i(n-1) - \eta \frac{\partial E_i(n)}{\partial \alpha_i(n)} \quad (16)$$

$$\beta_i(n) = \beta_i(n-1) - \eta \frac{\partial E_i(n)}{\partial \beta_i(n)} \quad (17)$$

$$\begin{aligned} \frac{\partial E_i(n)}{\partial \alpha_i(n)} &= \frac{1}{M} \sum_{l=0}^{M-1} \frac{\partial y_i(n-l)}{\partial \alpha_i(n)} \\ &= \frac{1}{M} \sum_{l=0}^{M-1} (-\frac{1}{2} \pm \frac{\alpha_i(n)}{4} (\frac{\alpha_i^2(n)}{4} \\ &\quad + \frac{1}{\beta_i(n)} z(n-l))^{-\frac{1}{2}}) \end{aligned} \quad (18)$$

$$\begin{aligned} \frac{\partial E_i(n)}{\partial \beta_i(n)} &= \frac{1}{M} \sum_{i=0}^{M-1} \frac{\partial y_i(n-l)}{\partial \beta_i(n)} \\ &= \frac{1}{M} \sum_{l=0}^{M-1} (\mp \frac{z(n-l)}{2\beta_i^2} (\frac{\alpha_i(n)^2}{4} \\ &\quad + \frac{1}{\beta_i(n)} z(n-l))^{-\frac{1}{2}}) \end{aligned} \quad (19)$$

Porality Control

In the above update equations, there is a freedom of polarity. It should be judged which polarity should be used. For this purpose, the following conditions are introduced. These conditions do not lose generality in real applications.

1. A linear component is greater than a nonlinear component.
2. The signal source level is limited. Say, for instance $|s_i(n)| < 1$.

Under these conditions, in the linear separation output,

$$z_i(n) = b_{i1} s_i(n) + b_{i2} s_i^2 \quad (20)$$

the following inequality is always held.

$$|b_{i1} s_i(n)| > |b_{i2} s_i^2(n)| \quad (21)$$

This means the porality of $z_i(n)$ is equal to that of $b_{i1} s_i(n)$. So, except for the polarity of b_{i1}, that of the output $y_i(n)$ can be controlled so as to be the same as that of $z_i(n)$. The polarity of b_{i1} does not affect separation performance. Because in blind source separation, constant scaling inherently remains.

3.3. Combination of Both Learning Algorithms

In the proposed method, first the linear separation block is trained. After convergence, the linearization block is adjusted. This separate training is stable.

3.4. Comparison with Conventional Methods

Nonlinear BSS methods have been proposed in [19],[20],[21]. In these methods, the linearization block is used before the linear separation. However, linearization of the mixed signals including the cross terms, like $s_1 s_2$ is difficult. Furthermore, a simultaneous learning for both signal separation and linearization is unstable. In the proposed method, the linearization block is used after the linear separation. Let the mixed signals be expressed with high-order polynomial, the signal sources, including high-order components of themselves like s_i and s_i^2, can be separated through the linear separation process. In other words, s_i^2 is still independent on the $s_j, i \neq j$ components. Since the cross terms are not independent, they can be suppressed through the linear separation process. After that, the linearization in each signal source is easier than that for the mixing signals with nonlinearity.

Although the proposed method is limited to 2nd-order nonlinearity, it can be extend to 3rd-order nonlinearity. In this case, network becomes complicated. Since in a wide range of application fields, order of nonlinearity is mainly up to 3rd-order, this limitation of order does not lose generality.

4. SIMULATIONS AND DISCUSSIONS

4.1. Simulation Conditions

Two signal sources and four observations are used. The signal sources are male speech signals. The mixing matrix is

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 2 \\ 2 & -1 \\ 1 & 2 \end{bmatrix}$$

The learning rate is $\eta = 0.001$. The nonlinear functions in the mixing block are

$$\begin{aligned} F_1(u) &= u + 0.4u^2 \\ F_2(u) &= u + 0.2u^2 \\ F_3(u) &= u - 0.6u^2 \\ F_4(u) &= u + 0.3u^2 \end{aligned}$$

4.2. Linear Separation

The parameters w_{ji} are trained by the learning algorithm described in sec.3.1. The learning curve is shown in Fig.2. The vertical axis indicates SNR in dB, the horizontal axis is the number of update iterations. SNR is defined as follows:

$$\sigma_s^2 = \sum_{i=1}^{2} \text{power of } s_i \text{ in } z_j, \quad \text{where } s_i \text{ is dominant} \tag{22}$$

$$\sigma_n^2 = \sum_{j=1}^{2} \text{power of } z_j \text{ exept for } s_i, \quad \text{where } s_i \text{ is dominant} \tag{23}$$

$$SNR_1 = 10\log_{10}\frac{\sigma_s^2}{\sigma_n^2} \tag{24}$$

$$\sigma_s^2 = \sum_{i=1}^{2} \text{power of } s_i \text{ and } s_i^2 \text{ in } z_j, \quad \text{where } s_i \text{ is dominant} \tag{25}$$

$$\sigma_n^2 = \sum_{j=1}^{2} \text{power of } z_j \text{ exept for } s_i \text{ and } s_i^2, \quad \text{where } s_i \text{ is dominant} \tag{26}$$

$$SNR_2 = 10\log_{10}\frac{\sigma_s^2}{\sigma_n^2} \tag{27}$$

In this figure, both SNR_1 and SNR_2 are shown with a solid line and a dashed line, respectively. In the linear separation process, the 2nd-order components s_i^2 are not suppressed, because they are also independent components against s_j and s_j^2, $i \neq j$. So, SNR_2 has some meaning. The final interferences, that is, s_2, s_2^2 and $s_1 s_2$ in z_1, and s_1, s_1^2 and $s_1 s_2$ in z_2, are reduced.

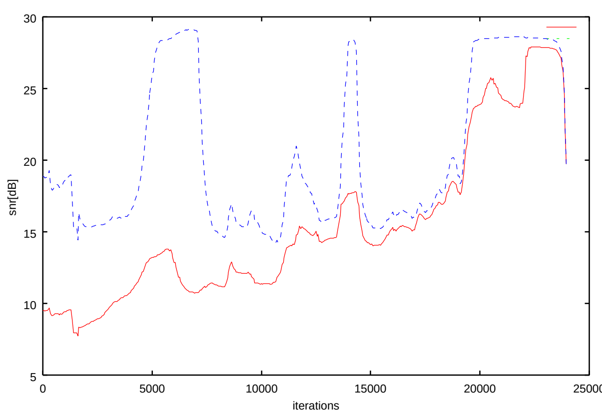

Fig. 2. Learning curve for output of linear separation block. SNR_1 and SNR_2 are shown with solid line and dashed line.

$z_1(n)$ and $z_2(n)$ are shown below.

$$z_1 = 5.28s_1 - 1.19s_2 + 6.05s_1^2 - 1.66s_1s_2 + 3.88s_2^2$$
$$z_2 = -0.21s_1 + 7.0s_2 - 0.03s_1^2 + 1.94s_1s_2 + 3.92s_2^2$$

In $z_1(n)$ and $z_2(n)$, $s_1(n)$ and $s_2(n)$ are extracted, respectively. In $z_1(n)$, $s_1(n)$ and $s_1^2(n)$ remain. On the other hand, in $z_2(n)$, $s_2(n)$ and $s_2^2(n)$ are dominant. In both outputs, the interferences, that is, s_2, s_2^2 and s_1s_2 in z_1, and s_1, s_1^2 and s_1s_2 in z_2, are reduced.

4.3. Linearization

Error valuation

In this process, separation performance is also evaluated by SNR_1 defined by Eq.(24). In the simulation, the s_i component and the other components are discriminated as follows: In the linearization block, $z_i(n)$ is linearized through

$$y_i(n) = -\frac{\alpha_i}{2} + \sqrt{\frac{\alpha_i^2}{4} + \frac{z_i(n)}{\beta_i}} \tag{28}$$

Let

$$\sqrt{\frac{\alpha_i^2}{4} + \frac{z_i(n)}{\beta_i}} = \sqrt{a_i s_i^2(n) + b_i(n)s_i(n) + c_i(n)} \tag{29}$$

Furthermore,

$$\sqrt{a_i s_i^2(n) + b_i s_i(n) + c_i(n)} = d_i s_i(n) + e_i(n) \tag{30}$$

$$a_i s_i^2(n) + b_i s_i(n) + c_i(n) = d_i^2 s_i^2(n) + 2d_i s_i(n) e_i(n) + e_i(n)^2 \tag{31}$$

Comparing the coefficients, the following relations are obtained.

$$d_i^2 = a_i \tag{32}$$
$$2d_i e_i(n) = b_i(n) \tag{33}$$
$$e_i^2(n) = c_i(n) \tag{34}$$

a_i and $c_i(n)$ are calculated using α_i, β_i and $z_i(n)$ at each sample. SNR is evaluated by

$$SNR_1 = 10\log\frac{p(n)}{q(n)} \tag{35}$$

$$p(n) = \frac{1}{M}\sum_{i=0}^{M-1}\left(y_i(n) + \frac{\alpha_i}{2} - e_i(n)\right)^2 \tag{36}$$

$$q(n) = \frac{1}{M}\sum_{i=0}^{M-1}\left(-\frac{\alpha_i}{2} + e_i(n)\right)^2 \tag{37}$$

The learning curves for both the linear separation and the linearization processes are shown in Fig.3, with a solid line and a dashed line, respectively. SNR means SNR_1. SNR after the linearization improved by $15dB$ compared with that after linear separation. Approximately, SNR_1=30dB is guaranteed, which is good separation performance.

4.4. Waveforms

Figures 4, 5 and 6 show waveforms of the speech signal sources, after the linear separation and the linearization, respectively. In Fig.4, the upper and the lower are s_1 and

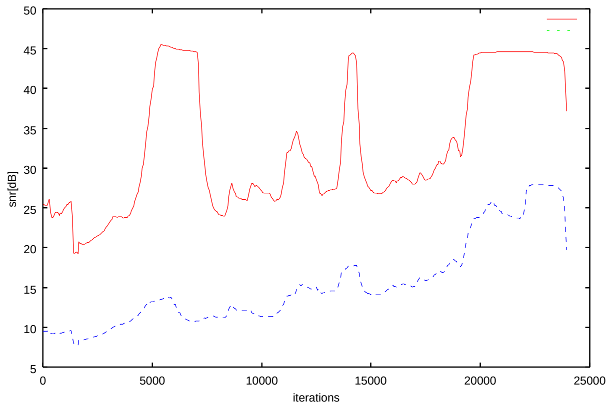

Fig. 3. Learning curve for output of both linear separation and linearization blocks with dashed line and solid line, respectively. SNR means SNR_1.

s_2, respectively. In Figs.5, 6, the upper and the lower are $z_1(n)$ and $y_1(n)$, and $z_2(n)$ and $y_2(n)$, respectively. $s_1(n)$ and $s_2(n)$ are extracted in $z_1(n)$ and $z_2(n)$, respectively. The polarity of $s_2(n)$ is reversed. The waveforms after the linear separation and the linearization are almost the same. However, SNR is slightly improved after the linearization.

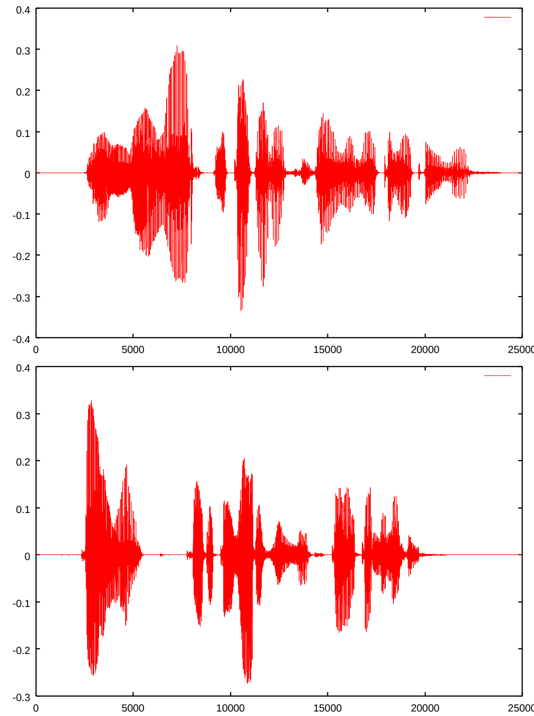

Fig. 4. Waveform of signal sources, which are male voice. Upper and lower are s_1 and s_2, respectively.

5. CONCLUSIONS

In this paper, a blind source separation method has been proposed for instantaneous nonlinear mixtures. It consists of the linear separation and the linearization in a cascade form. Both blocks are separately trained. The conventional learning algorithm of linear mixtures can be used for the former block. The new learning algorithm has been proposed for the latter block. Nonlinearity in the mixture is assumed to be 2nd-order. Simulation, using two speech signals and 2nd-order nonlinearity, shows usefulness of the proposed method.

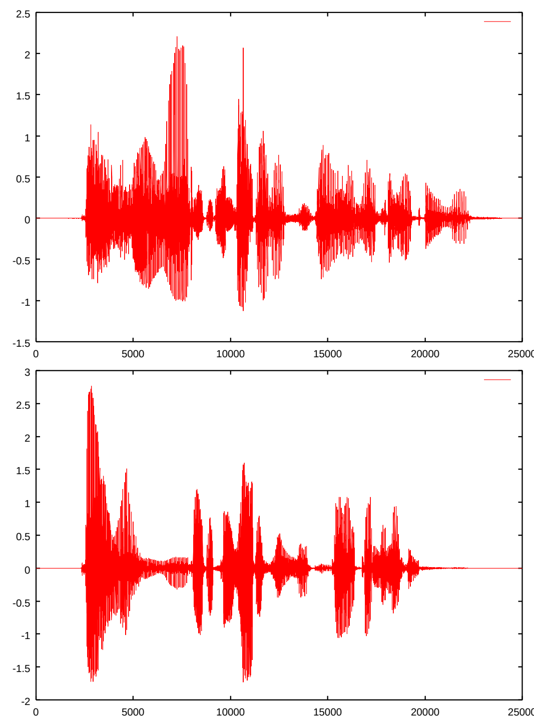

Fig. 5. Signal waveform after linear separation. Upper and lower are $z_1(n)$ and $z_2(n)$, respectively. $s_1(n)$ and $s_2(n)$ are extracted in $z_1(n)$ and $z_2(n)$, respectively.

6. REFERENCES

[1] C.Jutten, J.Herault and A.Guerin, "IIN.C.A: An independent components analyzer based on an adaptive neuromimetic network", in: J.Degmongeot, T.Herve, V.Raille and C.Roche, eds., Artificial Intelligence and Cognitive Science, Manhester Univ. Press, 1988.

[2] C.Jutten and J.Herault, "Analog implementation of permanent unsupervised learning algorithm", Proc. NATO Advanced Research Workshop on Neurocomputing, Les Arcs, France, pp.145-152, Feb.27-March 3, 1989.

[3] P.Comon, "Separation of stochastic process whose linear mixtures observed", Proc. ONR- NSF-IEEE Workshop on

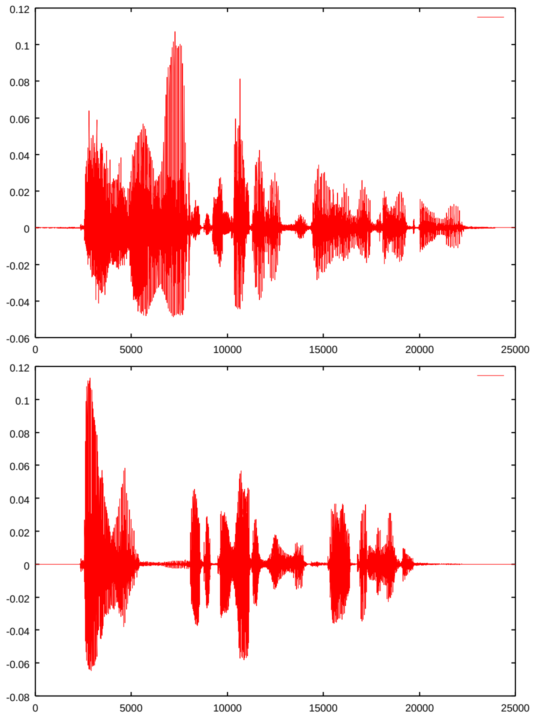

Fig. 6. Signal waveform after linearization. Upper and lower are $y_1(n)$ and $y_2(n)$, respectively. $s_1(n)$ and $s_2(n)$ are extracted in $y_1(n)$ and $y_2(n)$, respectively.

Higher Spectral Analysis Vail, pp.174-179, June 28-30, 1989.

[4] P.Comon, "Separation of sources using higher-order cumulants", SPIE Conference, Vol.1152, Advanced Algorithms and Architectures for Signal Processing IV, San Diego, pp. 170-181, August 6-11, 1989.

[5] J.F.Cardoso, "Eigen structure of the 4th order cumulant tensor with application to the blind source separation problem", ICASSP Proc. pp. 2655-1658.

[6] C.Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture", Signal Proc. 24, pp.1-10, 1991.

[7] P.Comon, C.Jutten and J.Herault, "Blind separation of sources, Part II: Problems statement", Signal Proc. 24, pp.11-20, 1991.

[8] E.Sorouchyari, "Blind separation of sources, Part III: Stability analysis", Signal Proc., 24, pp.21-29, 1991.

[9] K.Nakayama, A.Hirano and M.Nitta, "A constraint learning algorithm for blind source separation", Proc. IJCNN'2000, pp.24-27, July, 2000.

[10] K.Nakayama, A.Hirano and T.Sakai, "A pair-channel learning algorithm with constraints for multi-channel blind separation", Proc. IJCNN'01, July 2001.

[11] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures", Signal Processing, vol.45, no.2, pp.209–229, March 1995.

[12] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm for convolutive blind source separation with transmission delay constraint", Proc. IJCNN'2002, pp.1287-1292, May 2002.

[13] J.F.Cardoso, "Informax and maximum likelihood for source separation", IEEE Signal Processing Letter, vol.4, no.4, pp.112-114, Apr. 1997.

[14] A.Mansour and C.Jutten, "What should we say about the kurtosis", IEEE Signal Processing Letters, vol.6, no.12, pp.321-322, Dec. 1999.

[15] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation", IEEE Proc. ICASSP'01, MULT-P3.3, May 2001.

[16] L.Xu, C.C.Cheung and S.Amari, "Learned parametric mixture based ICA algorithm", Neurocomputing 22, pp.69-80, 1998.

[17] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation", Proc. IJCNN'2002, pp.1234-1239, May. 2002.

[18] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output PDF for blind source separation", Proc. 4th International Symposium on Independent Copmponent Analysis and Blind Signal Separation, ICA'2003, Nara, Japan, April 2003. to be presented.

[19] M.Solazzi, F.Piazza and A.Uncini, "Nonlinear blind source separation by spline neural networks", IEEE Proc. ICASSP'2001, Salt Lake City, MULT-P3.4, May 2001.

[20] F.Milani, M.Solazzi and A.Uncini, "Blind source separation of convolutive nonlinear mixtures by flexible spline nonlinear functions", IEEE Proc. ICASSP'2002, Orlando, Florida, pp.1641-1644, May 2002.

[21] A.Koutras,"Blind separation of non-linear convolved speech mixtures", IEEE Proc. ICASSP'2002, Orlando, Florida, pp.913-916, May 2002.

[22] S.Amari, A.Cichocki and H.H.Yang, "Recurrent neural networks for blind separation of sourdes", Proc. NOLTA-95, Las Vegas, pp.37-42, 1995.

[23] A.Cichocki, S.Amari, M.Adachi, W.Kasprzak, "Self-adaptive neural networks for blind separation of sources", Proc. ISCAS'96, Atlanta, pp.157-161, 1996.

[24] S.Amari, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, vol.10, no.8, pp.1345-1351, 1997.