

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
9 September 2005 (09.09.2005)

PCT

(10) International Publication Number  
WO 2005/083706 A1

(51) International Patent Classification<sup>7</sup>: G11B 20/10

(21) International Application Number:  
PCT/KR2005/000526

(22) International Filing Date: 26 February 2005 (26.02.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10-2004-0013016  
26 February 2004 (26.02.2004) KR

GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

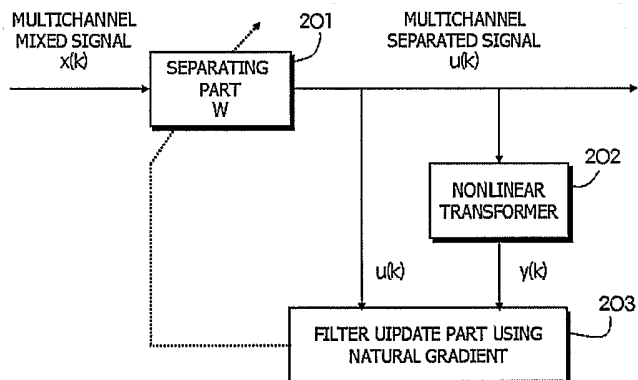
- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(71) Applicant and  
(72) Inventor: NAM, Seung Hyon [KR/KR]; #7-1503 Hatnim Apartment, Dunsan-dong, Seo-gu, Daejeon 302-737 (KR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: THE METHODS AND APPARATUS FOR BLIND SEPARATION OF MULTICHANNEL CONVOLUTIVE MIXTURES IN THE FREQUENCY DOMAIN



(57) Abstract: A method and apparatus is disclosed for performing blind source separation using frequency-domain normalized multichannel blind deconvolution. In an embodiment of this invention, the multichannel mixed signals are formed as frames of  $N$  samples, which consist of  $r$  consecutive blocks of  $M$  samples. The frames of mixed signals are separated using separating filters in the frequency domain in an overlap-save manner using a discrete Fourier transform (DFT). The separated signals are then converted back into the time domain using the inverse DFT to be applied to a nonlinear function. The cross-power spectra between separated signals and nonlinear-transformed signals are computed and are normalized by the power spectra of separated signals and the power spectra of nonlinear-transformed signals to have

flat spectra. The invention then applies the time domain constraint to preserve the first  $L$  cross-correlations. These alias-free normalized cross-power spectra are further constrained by nonholonomic constraints. The invention then computes natural gradient by convolving alias-free normalized cross-power spectra with separating filters. After the length of separating filters is constrained to  $L$ , separating filters are updated using the natural gradient and normalized to have unit norm. The terminating conditions are checked to see if separating filters converged.

WO 2005/083706 A1

**【DESCRIPTION】****【Invention Title】**

The Methods and Apparatus for Blind Separation of Multichannel Convolutional Mixtures in the Frequency-domain

5

**【Technical Field】**

This invention relates to signal processing, more particularly to a method, apparatus, and storage medium that contains a program for performing blind signal separation of multichannel convolutional mixtures in the frequency domain.

10

**【Background Art】**

In the art of speech processing, it is necessary to separate mixtures of multiple signals (including speech signals) from multiple sensors in a multipath environment. Such a separation of the mixtures without a priori knowledge of signals is known as blind source separation (BSS). BSS is very useful to separate signals that are from independent sources such as multiple speakers and sonar arrays. BSS techniques may be applied to speaker location tracking, speech recognition, speech coding, 3-D object-based audio signal processing, acoustic echo cancellers, channel equalization, estimation of direction of arrival, and detection of various biological signals such as EEG and MEG.

20

Most BSS techniques try to recover the original signals by nullifying the effect of multipath effects. Although filters of infinite length are required for this purpose in general, filters of finite length also provide sufficient separation in most real world environments.

25

There are two popular approaches to this BSS problem: (i) multiple decorrelation (MD) methods that exploit the second order statistics of signals as independence measure and (ii) multichannel blind deconvolution (MBD) methods that exploit the higher order statistics.

30

The MD methods decorrelate mixed signals by diagonalizing second order statistics. [See, e.g. E. Weinstein, M. Feder, and A. V. Oppenheim, "Multi-channel signal separation by decorrelation," *IEEE Trans. Speech Audio Processing*, vol. 1, no. 4, pp. 405-413, Apr. 1993; Lucas Parra and Clay Spence, "Convolutional blind source separation of nonstationary sources", *IEEE Trans. Speech Audio Processing*, pp.320-327, May, 2000; D.W.E. Schobben and P.C.W. Sommen, "A frequency-domain blind signal separation method based on decorrelation," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp. 1855-1865, Aug. 2002; N. Murata and S. Ikeda, and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signal," *Neurocomputing*, vol. 41, no. 4, pp.1-24, 2001] Diagonalization should be performed at multiple time instants for successful separation of signals. For this reason, these methods are only applied to

35

nonstationary signals. These methods are quite fast and stable. The MBD methods, on the other hand, separate signals by minimizing mutual information of nonlinear-transformed separated signals which are transformed by a nonlinear function matched to statistical distributions of signals.

[See, e.g. S. Amari, S.C. Douglas, A. Cichocki, H.H. Yang, "Novel on-line adaptive learning

5 algorithm for blind deconvolution using the natural gradient approach", *Proc. IEEE 11th IFAC Symposium on System Identification*, Japan, 1997, pp.1057-1062; A. J. Bell and T. J. Sejnowski,

"An information maximization approach to blind separation and blind deconvolution," *Neural Computation*, 7, no. 6, pp. 1129-1159, Nov. 1995; L. Zhang, A. Cichocki, and S. Amari,

10 "Geometrical structures of FIR manifolds and their application to multichannel blind deconvolution," *Proc of Int. IEEE Workshop on Neural Networks and Signal Processing*, pp. 303-312, Madison, Wisconsin, USA, Aug. 23-25, 1999]

**【Disclosure】****【Technical Problem】**

In the prior art, the separation performances are significantly limited due to their shortcomings such as frequency permutation, whitening, and filter types employed.

5           The MD methods suffer from the frequency permutation problem – the separated sources are differently ordered in each frequency bin so that the resulting separated signals are still mixed. Although there are some solutions to this permutation problem, separation performance is degraded as the length of separating filters increase. On the other hand, the MBD methods suffer from whitening effect – the spectra of separating signals are whitened (or flattened). The linear  
10       predictive method for speech signals has been proposed as a solution for this shortcoming of the MBD methods. [See, e.g., S.C. Douglas, "Blind separation of acoustic signals", in *Microphone Arrays: Signal processing techniques and applications*, M. Brandstein and D. Ward Eds, Springer, pp. 355-380, 2001.] This method employs bidirectional filters that may be inappropriate normal mixing environments in practice. In addition, parts of room impulse response may be treated as  
15       vocal track response of human speech signals.

          Therefore, there is a need for a BSS technique that separates speech signals fast and accurately with high speech quality.

**【Technical Solution】**

20           This invention provides a method and apparatus of multichannel blind deconvolution, that estimates unidirectional separating filters for blind signal separation, with normalized natural gradient in the block frequency domain.

          Figure 1 depicts a system 100 for executing signal separation of the invention. The system 100 comprises an input device 126 that supplies the mixed signals that are to be separated and a  
25       computer system 108 that executes the frequency-domain normalized multichannel blind deconvolution routine 124 of the present invention. The input device 126 may contain any types of devices, but is illustratively shown to contain a sensor array 102, a signal processor 104 and a recorded signal source 106. The sensor array 102 contains one or more transducers 102A, 102B, 102C such as microphones. The signal processor 108 digitizes a (convolutive) mixed signal.

30           The computer system 108 comprises a central processing unit (CPU) 114, a memory 122, input/output (I/O) interface 120, and support circuits 116. The computer system is generally connected to the input device 110 and various input/output devices such as a monitor, a mouse, and a keyboard through the I/O interface 120. The support circuit 116 comprises well-known circuits such as power supplies, cache, timing circuits, a communication circuit, bus and the like. The  
35       memory 122 may include random access memory (RAM), read only memory (ROM), disk drive, tape drive, flash memory, compact disk (CD), and the like, or some combination of memory

devices. The invention is implemented as the frequency-domain normalized multichannel blind  
deconvolution routine 124 that is stored in memory 122 and executed by the CPU 114 to process  
the signals from the input devices 126. As such, the computer system 108 is a general purpose  
computer system that becomes a specific purpose computer system when executing the routine 124  
5 of the present invention. The invention can also be implemented in software, hardware or a  
combination of software and hardware such as application specific integrated circuits (ASIC),  
digital signal processor, and other hardware devices.

The illustrative computer system 108 further contains speech recognition processor 118,  
such as a speech recognition circuit card or a speech recognition software, that is used to process  
10 the separated signals that are extracted from the mixed signal by the invention. As such, mixed  
signals in a room having more than two persons speaking simultaneously with background noise or  
music can be captured by the microphone array 102. The speech signals captured by the  
microphones 102 are mixed signals that should be separated into individual components for speech  
recognition. The mixed signal is sent to the computer system 108 after filtered, amplified, and  
15 digitized by the signal processor 104. The CPU 114, executing the frequency-domain normalized  
multichannel blind deconvolution routine 124, separates the mixed signals into its component  
signals. From these component signals, background noise can be removed easily. The component  
signals without noise are then applied to the speech recognition processor 118 to process the  
component signals into computer text or computer commands. In this manner, the computer system  
20 108, executing the frequency-domain normalized multichannel blind deconvolution routine 124, is  
performing signal preprocessing or conditioning for speech recognition processor 118.

Figure 2a is a block diagram of the invention, a frequency-domain normalized  
multichannel blind deconvolution 124. The frequency-domain normalized multichannel blind  
deconvolution of the invention comprises a separation part 201, a nonlinear transformer 202, and a  
25 filter updating part 203 that updates separating filter coefficients using the normalized natural  
gradient. The separation part 201 separates a mixed multichannel signal  $\mathbf{x}(k)$ . The mixed signal  
 $\mathbf{x}(k)$  is observed in a multipath environment as the output of the  $n$  sensors to the  $m$  component  
signals and is defined by the following equation:

$$\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T \quad (1)$$

30 where  $x_j(k)$  is the mixed signal from the  $j$ -th sensor. The separating filter to separate  $\mathbf{x}(k)$  into  
its component signals is an  $m \times n$  matrix  $\mathbf{W}(z, k)$  whose  $(i, j)$  component is represented by the  
following equation:

$$w_{ij}(z, k) = \sum_{p=0}^{L-1} w_{ij,p}(k) z^{-p} \quad (2)$$

where  $L$  is the length of the separating filters. The separated component signal  $\mathbf{u}(k)$  is defined by

the following equation:

$$\mathbf{u}(k) = [u_1(k), u_2(k), \dots, u_m(k)]^T \quad (3)$$

Where  $u_i(k)$  is the  $i$ -th separated signal defined by the following equation:

$$u_i(k) = \sum_{j=1}^n w_{ij,p}(k) x_j(k-p), \quad i = 1, \dots, m \quad (4)$$

5 Figure 2b depicts a separating process for the case of  $m=n=2$ . The separated signal  $\mathbf{u}(k)$  from the separation part 201 is applied to the nonlinear transformer 202.

The nonlinear transformer 202 performs transformation of the separated signal through a memoryless nonlinear function so that the nonlinear-transformed signal has a uniform probability density. The nonlinear transformation is defined by the following equation:

$$10 \quad y_i(k) = f(u_i(k)), \quad i = 1, \dots, m \quad (5)$$

Figure 2c is an illustration of the nonlinear transformation that a signal with Laplacian probability density is mapped into a signal with uniform probability density. A function to be used in the nonlinear transformation is closely related to the probability density. For audio and speech signals,  $\alpha \operatorname{sgn}(u)$  or  $\tanh(u)$  is used in general.

15 The filter updating part 203 updates the separating filter coefficients using the steepest ascent rule with natural gradient by the following equation:

$$w_{ij,p}(k+1) = w_{ij,p}(k) + \mu \Delta w_{ij,p}(k) \quad (6)$$

for  $1 \leq i \leq m, 1 \leq j \leq n, 0 \leq p \leq L-1$ , where  $\mu$  is the step size and  $\Delta w_{ij,p}(k)$  is the natural gradient defined by the following equation:

$$20 \quad \Delta w_{ij,p}(k) = \Delta w_{ij,p}(k) - \sum_{i=1}^m \sum_{q=0}^p \bar{y}_i(k) \bar{u}_i(k-p+q) w_{ij,q}(k) \quad (7)$$

Where  $\bar{y}_i(k)$  and  $\bar{u}_i(k)$  are the frequency-domain normalized versions, having flat spectrum, of  $y_i(k)$  and  $u_i(k)$ , respectively. Note also that the filter lag  $q$  in equation (7) is limited up to  $p$  not up to  $L-1$ . In this invention the separating filter is unidirectional of length  $L$ . Thus no sample delay is required.

25 In this invention, the above mentioned process is performed in the frequency domain in an overlap-save manner to take the advantage of the FFT (Fast Fourier Transform). The filter length, the block length, the frame length are denoted as  $L, M, N$ , respectively. The amount of overlapping between frames is determined by the ratio  $r=N/M$ . In the sequel, 50% overlap is assumed ( $r=2$ ) and the FFT size is assumed to be equal to the frame length for simplicity.

30 Figure 3 depicts a flow chart of an embodiment of this invention, the frequency-domain normalized multichannel blind deconvolution. With reference to the flow chart, the mixed signal  $\mathbf{x}(k)$  is input at step 301. At step 302, the mixed signal forms a current frame of two ( $r=2$ ) consecutive blocks of  $M$  samples as follows:

$$\mathbf{x}_j(b) = [x_j(bM - 2M + 1), \dots, x_j(bM)]^T, j = 1, \dots, n \quad (8)$$

where  $b$  denotes the block index. At step 303, the mixed signal is separated using the separating filters

$$\mathbf{w}_{ij}(b) = [w_{ij,0}, w_{ij,1}, \dots, w_{ij,L-1}]^T \quad (9)$$

5 The separating filters generally initialized as

$$\mathbf{w}_{ij}(0) = [1, 0, \dots, 0]^T, i = j \quad (10a)$$

$$\mathbf{w}_{ij}(0) = [0, \dots, 0]^T, i \neq j \quad (10b)$$

If there is any useful information on the separating filters, however, the information can be utilized into initialization of the separating filters. The separated signal is computed in the frequency

10 domain using circular convolution as in the following equation:

$$\mathbf{u}_i(f, b) = \sum_{j=1}^n \mathbf{w}_{ij}(f, b) \odot \mathbf{x}_j(f, b) \quad (11)$$

where  $\odot$  denotes the component-wise multiplication, and  $f$  denotes the frequency domain quantity such that

$$\mathbf{w}_{ij}(f, b) = \mathbf{F} \mathbf{w}_{ij}(b) \quad (12a)$$

$$15 \quad \mathbf{x}_j(f, b) = \mathbf{F} \mathbf{x}_j(b) \quad (12b)$$

where  $\mathbf{F}$  is the  $N \times N$  DFT matrix. The separated signal is then transformed back into the time domain in order to discard the first  $L$  aliased samples as in the following equation:

$$\mathbf{u}_i(f, b) = \mathbf{P}_{0, N-L} \mathbf{F} \mathbf{u}_i(f, b) = [0, \dots, 0, u_i(bM - 2M + L + 1), \dots, u_i(bM)]^T \quad (13)$$

where  $\mathbf{P}_{0, N-L}$  is the projection matrix (or window matrix) to make first  $L$  samples to zeros and is

20 defined as follows:

$$\mathbf{P}_{0, N-L} = \begin{pmatrix} \mathbf{0}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N-L} \end{pmatrix} \quad (14)$$

where  $\mathbf{0}_L$  is the  $L \times L$  zero matrix and  $\mathbf{I}_{N-L}$  is the  $(N-L) \times (N-L)$  identity matrix.

At step 304, the separated signal is transformed via a nonlinear function in the time domain. One of two following equations can be used:

$$25 \quad \mathbf{y}_i(b) = f(\mathbf{u}_i(b)) = [0, \dots, 0, f(u_i(bM - 2M + L + 1)), \dots, f(u_i(bM))]^T \quad (15a)$$

$$\mathbf{y}_i(b) = f(\mathbf{u}_i(b)) = [0, \dots, 0, f(u_i(bM - 2M + 2L + 1)), \dots, f(u_i(bM))]^T \quad (15b)$$

The output of this nonlinear function is used to compute the cross-correlations

$f(u_i(k))u_j(k-p)$ ,  $p = 0, 1, \dots, L-1$  at step 306. If equation (15a) is used, the cross-correlations will

be biased. If equation (15b) is used, the cross-correlations will be unbiased.

30 At step 305, the alias-free normalized cross-power spectra are computed. Step 305 is very critical in this invention. The normalized cross-power spectrum is defined by the following

equation:

$$\bar{\mathbf{P}}(b) = \begin{pmatrix} \bar{\mathbf{P}}_{y_1 u_1}(b) & \cdots & \bar{\mathbf{P}}_{y_1 u_m}(b) \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{P}}_{y_m u_1}(b) & \cdots & \bar{\mathbf{P}}_{y_m u_m}(b) \end{pmatrix} \quad (16)$$

Where  $\bar{\mathbf{P}}_{y_i u_j}(f, b)$  is the normalized cross-power spectrum between  $y_i(f, b)$  and  $u_j(f, b)$  to be described below. If  $i = j$ , the expected value is normalized to 1 by Busgang property. At step 306,

5 the cross-power spectra are computed in the frequency domain by the following equation:

$$\mathbf{P}_{y_i u_j}(f, b) = \mathbf{y}_i(f, b) \odot \mathbf{u}_j^*(f, b) \quad (17)$$

where \* denotes the complex conjugation and

$$\mathbf{y}_i(f, b) = \mathbf{F} \mathbf{y}_i(b) \quad (18a)$$

$$\mathbf{u}_j(f, b) = \mathbf{F} \mathbf{u}_j(b) \quad (18b)$$

10 Note that the cross-power spectra in equation (17) are computed using only the samples from the current frame as in equation (18a) and (18b). At step 307, the power spectra of the separated signals and the nonlinear-transformed signals are computed to normalize the cross-power spectra. In order to accommodate time varying nature of the signal, the power spectra are updated at each block as follows:

$$15 \quad \mathbf{P}_{y_i}(f, b) = (1 - \gamma) \mathbf{P}_{y_i}(f, b-1) + \gamma |\mathbf{y}_i(f, b)|^2, \quad i = 1, \dots, m \quad (19a)$$

$$\mathbf{P}_{u_j}(f, b) = (1 - \gamma) \mathbf{P}_{u_j}(f, b-1) + \gamma |\mathbf{u}_j(f, b)|^2, \quad j = 1, \dots, m \quad (19b)$$

Here,  $\gamma$  is a constant between 0 and 1. The power spectra are initialized as

$\mathbf{P}_{y_i}(f, 0) = \mathbf{P}_{u_i}(f, 0) = c[1, \dots, 1]^T$ ,  $i = 1, \dots, m$ , where  $c$  is a small positive constant  $0 < c \ll 1$ . At step

308, the cross-power spectra are normalized as follows:

$$20 \quad \bar{\mathbf{P}}_{y_i u_j}(f, b) = \frac{\mathbf{P}_{y_i u_j}(f, b)}{\sqrt{\mathbf{P}_{y_i}(f, b) \odot \mathbf{P}_{u_j}(f, b)}} \quad (20)$$

where the division is performed in the component-wise. If the cross-power spectra in equation (20) are transformed back into the time domain, however, the resulting cross-correlations contain aliased parts. Furthermore, only the first  $L$  cross-correlations are required to compute the natural gradient in equation (7). Therefore, only the first  $L$  cross-correlations must be extracted. This is performed

25 at step 309 by applying the proper time domain constraint in the time domain as follows:

$$\tilde{\mathbf{P}}_{y_i u_j}(f, b) = \mathbf{F} \mathbf{P}_{L,0} \mathbf{F}^{-1} \bar{\mathbf{P}}_{y_i u_j}(f, b) \quad (21)$$

where  $\mathbf{F}^{-1}$  is the  $N \times N$  inverse DFT matrix and  $\mathbf{P}_{L,0}$  is the  $N \times N$  projection matrix, which preserves the first  $L$  samples and set the rest  $(N-L)$  samples to zeros, defined as



$$\mathbf{P}_{L,0} = \begin{pmatrix} \mathbf{I}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{N-L} \end{pmatrix} \quad (22)$$

At step 310, the natural gradient is computed using the nonholonomic constraints as follows:

$$\hat{\mathbf{P}}_{y,\mu_j}(f,b) = \begin{cases} \bar{\mathbf{1}} - \tilde{\mathbf{P}}_{y,\mu_j}(f,b), & \text{for } i = j \\ -\tilde{\mathbf{P}}_{y,\mu_j}(f,b), & \text{for } i \neq j \end{cases} \quad (23a)$$

$$\Delta \mathbf{w}_{ij}(f,b) = \sum_{i=1}^m \hat{\mathbf{P}}_{y,\mu_i}(f,b) \odot \mathbf{w}_{ij}(f,b) \quad (23b)$$

where  $\bar{\mathbf{1}} = [1, \dots, 1]^T$ . The nonholonomicity implies that separation is not responding to signal powers but only to statistical dependence between signals.

Note that  $\hat{\mathbf{P}}_{y,\mu_j}(f,b)$  in equation (23a) is approximately nonholonomic since the diagonal components  $\tilde{\mathbf{P}}_{y,\mu_j}(f,b)$  are  $\bar{\mathbf{1}}$  on the average. However, exact nonholonomicity can be attained by forcing the diagonal components to zeros as:

$$\hat{\mathbf{P}}_{y,\mu_j}(f,b) = 0 \quad (24)$$

Although all the components of the separating filters are learned in general, all diagonal components can be omitted in learning so that the diagonal components are absorbed into the off-diagonal components. This is easily achieved in this invention by setting the diagonal components of the gradient to zero as follows:

$$\Delta \mathbf{w}_{ii}(f,b) = 0 \quad (25)$$

If equation (24) and (25) are combined together, the computation can be reduced. Note that, for the special case of  $m=n=2$ , the time-domain constraints in equation (21) are not necessary and the computational burden is significantly reduced. Such flexibility for modifications is one advantage of the present invention.

At step 311, the separating filters are updated as:

$$\mathbf{w}_{ij}(f,b+1) = \mathbf{w}_{ij}(f,b) + \mu \Delta \mathbf{w}_{ij}(f,b) \quad (26)$$

At step 312, the separating filters are normalized in the frequency domain to have unit norm. The separating filters with unit norm preserve signal power during iteration.

At step 313, termination conditions are investigated whether the separating procedure should be terminated or not.

At step 314, the converged separating filters are used to filter the mixed signals to get the separated signals. Equation (11) in step 302 can also be used in this step.

Although various embodiments which incorporate the teaching of the present invention have been shown and described in detail herein, those skilled in the art can readily devise many

other varied embodiments that still incorporate these teachings. Accordingly, it is intended that all such alternatives, modifications, permutations, and variations to the exemplary embodiments can be made without departing from the scope and spirit of the present invention.

#### 5      **【Advantageous Effects】**

Figure 4a shows an example of separating mixed signals recorded in a real-world environment. Speech and music signals are recorded in a room using two microphones and the mixed signals are then separated using the inventive method. Figure 4a shows two mixed signals  $\mathbf{x} = (x_1, x_2)$  and two separating signals  $\mathbf{u} = (u_1, u_2)$  from top to bottom. Parameters used are  $L=128$ ,  
10  $M=2L$ ,  $N=2M$ ,  $\mu = 0.0025$ . Figure 4b shows the final separating filters in this example.

This invention can separate a desired signal from the mixtures with high speech quality so that the separated signal can be directed to a speech recognizer or a speech coder. Figure 5 shows the original signal  $\mathbf{s}$ , the mixed signal  $\mathbf{x}$ , and the separated signal  $\mathbf{u}$  from top to bottom for each channel. Figure 5 demonstrates high quality of the separated speech signals.

15

#### **【Description of Drawings】**

The teaching of the present invention can be readily understood by considering the following description in conjunction with accompanying drawings, in which:

Fig. 1 depicts a system for executing a software implementation of the present invention;

20      Fig. 2a depicts a block diagram of a multichannel blind deconvolution using normalized natural gradient;

Fig. 2b depicts a diagram of separating filters to separate mixed multichannel signals;

Fig. 2c depicts a schematic graph of transforming a separated signal into a signal with uniform probability density using a nonlinear function;

25      Fig. 3 depicts a flow chart of an embodiment of the present invention;

Fig. 4a depicts separated signals, speech and music, from mixtures recorded in a real room by the present inventive method;

Fig. 4b depicts the final converged separating filters  $w_{ij}$  to separate the mixtures recorded in a real room by the present inventive method; and

30      Fig. 5 depicts an original speech signal  $\mathbf{s}$ , a mixed speech signal  $\mathbf{x}$ , and a separated signal  $\mathbf{u}$  for each channel.

#### **【Industrial Applicability】**

The present invention finds application in a speech recognition system as a signal  
35 preprocessor system for deconvolving and separating signals from different sources such that a

speech recognition processor can response to various speech signals without interfering noise sources.

**【CLAIMS】****【Claim 1】**

A method for separating mixed signals into component signals comprising the steps of:

- (a) producing current frames from said mixed signals (302);
- 5 (b) separating said current frames into component signals of the current frame (303);
- (c) transforming said component signals using nonlinear functions into nonlinear-transformed signal (304);
- (d) computing aliasing-free normalized cross-power spectra of said component signals and said nonlinear-transformed signals in (c) (305);
- 10 (e) computing natural gradient using said cross-power spectra (310);
- (f) updating separating filter coefficients using said natural gradient (311);
- (g) normalizing said separating filter coefficients (312);
- (h) determining convergence conditions and iterating from (a) to (g) until convergence (313); and
- 15 (i) separating said mixed signals into component signals using said separating filter coefficients after convergence (314).

**【Claim 2】**

The method of claim 1 wherein step (b) further comprises the substeps of:

- 20 (b1) transforming said mixed frames and said separating filter coefficients into the frequency domain;
- (b2) computing component signals in the frequency domain; and thereby transforming said component signals back into the time domain; and
- (b3) zeroing the first  $L$  samples of said component signals; thereby producing component
- 25 signals.

**【Claim 3】**

The method of claim 1 wherein step (d) further comprises the substeps of:

- 30 (d1) transforming said component signals and said nonlinear-transformed signals into the frequency domain;
- (d2) computing cross-power spectra using said component signals and said nonlinear-transformed signals in the frequency domain (306);
- (d3) computing power spectra of said component signals and the power spectra of nonlinear-transformed signals (307);
- 35 (d4) computing normalized cross-power spectra (308); and

(d5) transforming said normalized cross-power spectra back into the time domain; and applying the time domain constraint for preserving only the first L samples (309).

**【Claim 4】**

5 The method of claim 1 wherein step (e) further comprises the substeps of:

(e1) applying the nonholonomic constraints to said aliasing-free normalized cross-power spectra; and

(e2) computing said natural gradient using said separating filter coefficients and said nonholonomic-constrained cross-power spectra.

10

**【Claim 5】**

An apparatus of claim 1 to claim 4 for separating a plurality of the mixed signals into a plurality of component signals using the frequency-domain normalized multichannel blind deconvolution method.

15

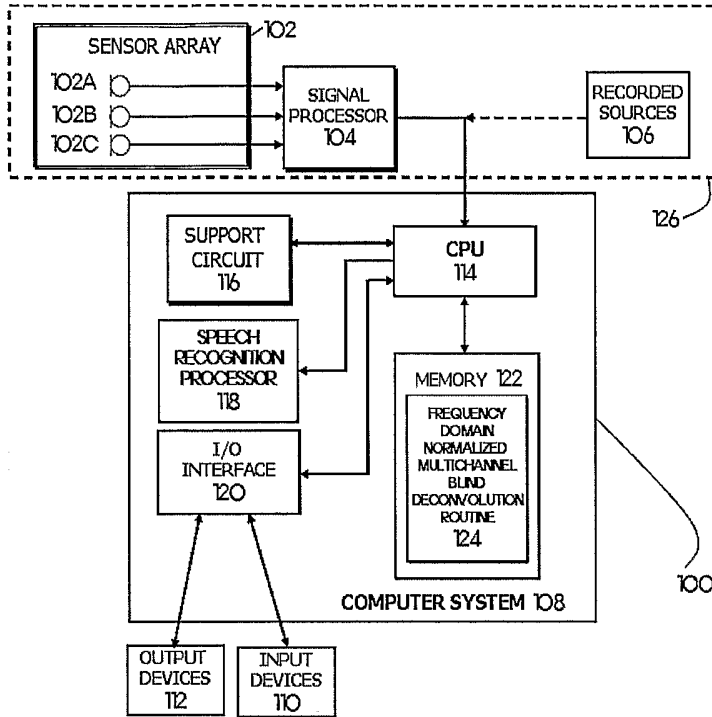
**【Claim 6】**

**【Claim 7】**

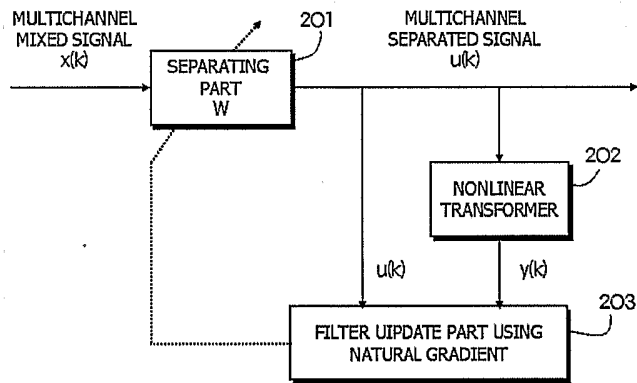
20 A computer readable storage medium of claim 1 to claim 4 containing a program that, when executed upon a general purpose computer system, causes said general purpose computer system to become a specific purpose computer system that separates a plurality of the mixed signals into a plurality of component signals using the frequency-domain normalized multichannel blind deconvolution method.

【DRAWINGS】

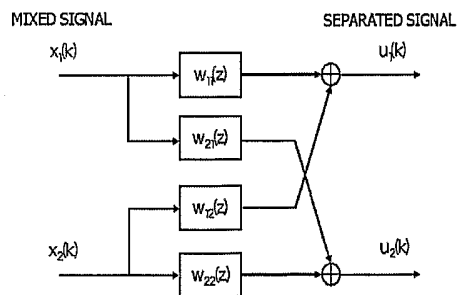
【Figure 1】



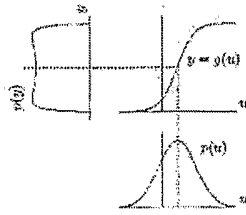
【Figure 2a】



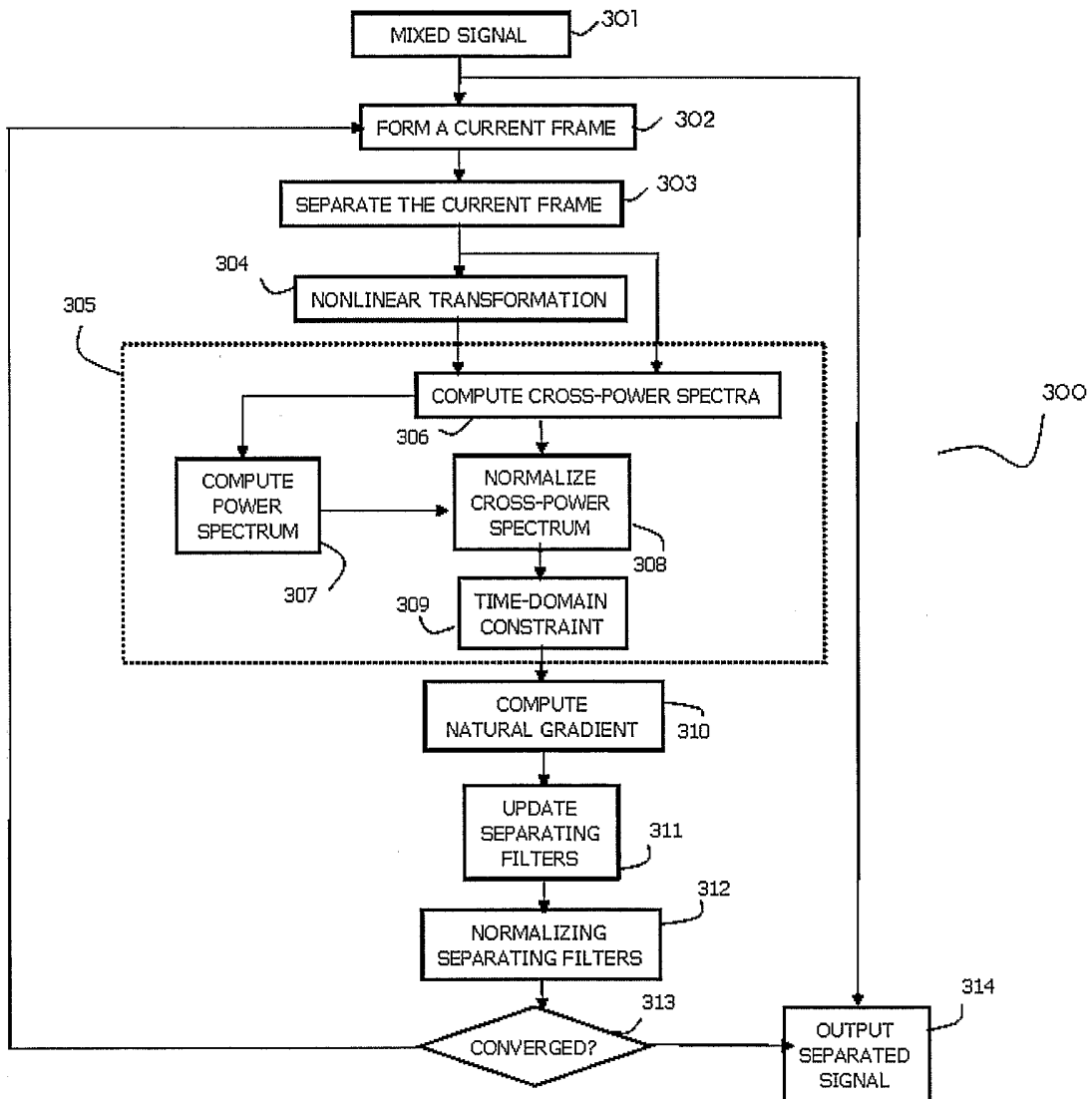
【Figure 2b】



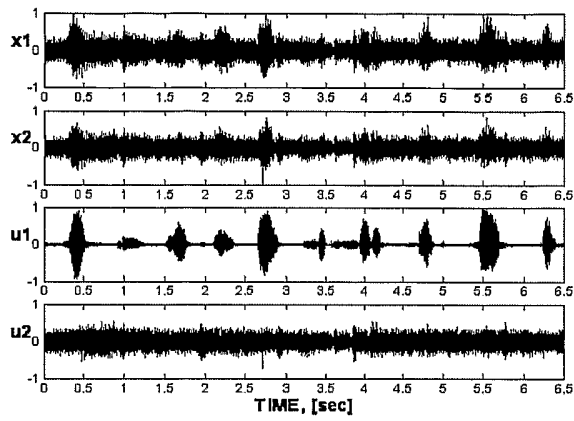
【Figure 2c】



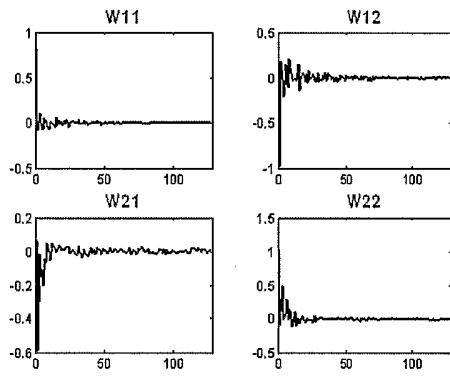
【Figure 3】



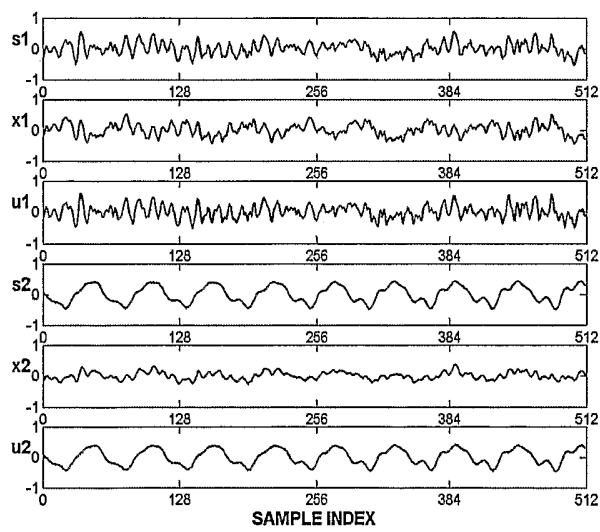
【Figure 4a】



【Figure 4b】



【Figure 5】





**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/KR2005/000526

**A. CLASSIFICATION OF SUBJECT MATTER**

**IPC7 G11B 20/10**

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

G11B 20/10 G06F 15/00 G06F 17/00 H03H 21/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, PAJ "BSS(Blind Source Separation), blind, mixed, component, signal, separate, normalized, cross-power, filter, gradient, convergence, frequency, multi, path, channel"

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6,654,719 B1 (LUCENT TECHNOLOGY INC.) 25 November 2003 See the whole document	1-5, 7
A	US 6,167,417 A (SARNOFF CORPORATION) 26 December 2000 See the whole document	1-5, 7
A	EP 0,779,709 A2 (MOTOROLA INC.) 18 June 1997 See the whole document	1-5, 7
A	JP 2003-78423 A (CLARION CO., LTD.) 14 March 2003 See the whole document	1-5, 7
A	US 5,208,786 A (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 4 May 1993 See the whole document	1-5, 7

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier application or patent but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
 "&" document member of the same patent family


Date of the actual completion of the international search

12 JULY 2005 (12.07.2005)

Date of mailing of the international search report

**12 JULY 2005 (12.07.2005)**

Name and mailing address of the ISA/KR

 Korean Intellectual Property Office  
920 Dunsan-dong, Seo-gu, Daejeon 302-701,  
Republic of Korea  
Facsimile No. 82-42-472-7140

Authorized officer

KIM, Yong Woong  
Telephone No. 82-42-481-5698

