AUTONOMOUS DIRECTIVITY MICROPHONE SYSTEM BASED ON THE SPATIO-TEMPORAL BLIND SOURCE SEPARATION ALGORITHM

Kenbu TERAMOTO *Department of Mechanical Engineering, SAGA University, SAGA university, 1-Honjo Saga-shi, JAPAN*
E-mail: tera@me.saga-u.ac.jp
Kohsuke TSURUTA *Department of Mechanical Engineering, SAGA University, SAGA university, 1-Honjo Saga-shi, JAPAN*
E-mail: 02td73@edu.cc.saga-u.ac.jp
Hans-Elias de BREE *Microflown Technologies BV, Netherland*

Keywords: Spatio-temporal gradient method, Array signal processing, Particle velocity sensing, Blind signal separation

ABSTRACT

This paper presents an autonomous directivity microphone system based on the newly proposed spatio-temporal blind source separation. The blind source separation principally uses no a priori knowledge about parameters of convolution, filtering and mixing. In the simplest case of the blind source separation problems, observed mixed signals are linear combinations of unknown mutually statistically independent, zero-mean source signals. The spatio-temporal blind signal separation algorithm utilizes the linearity among the four signals: (1) the sound pressure, (2)x, (3)y, and (4)z-directional particle velocities, all of which are governed by the equation of motion. The proposed method, therefore, has an ability to simplify the convolution blind source separation problems into the instantaneous blind source separation over the spatio-temporal gradient space. Several acoustical experiments are performed with the particle velocity microphone (Microflown) successfully instead of sound pressure microphones. Because, x and y-directional spatial gradients of the sound pressure are equivalent the corresponding directional particle acceleration theoretically.

INTRODUCTION

Recently, the methods of blind source separation and multi-deconvolution of source signals have been proposed in various fields [1, 2, 3]. Especially, acoustic applications include the cocktail - party problem, sonar problem, and biomedical problem which extract the fetus heart beat signal from several acoustic transducers outside the mother's body. Equally the spatio-temporal gradient method has been proposed as one of the acoustical signal processing algorithm and developed for the non-destructive evaluation [4, 5]. Combining these methods, the spatio-temporal blind source separation algorithm make the convolution blind source separation problems into the simplest instantaneous mixture problems.

This paper is organized as follows: the second section presents the spatio-temporal gradient blind source separation algorithm, the third section evaluates the proposed velocimetry via acoustical experiment with Microflown system.

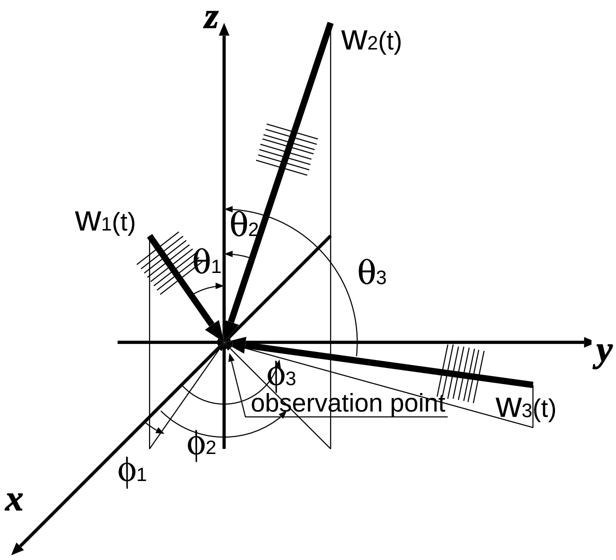


Fig. 1: Geometry of the three independent incident plane waves and the observation point

PROBLEM FORMULATION

An Overlap of three plane waves

Assuming that three band-limited plane waves propagates in the three lineally independent directions respectively as shown in Fig.1, it can be found that the acoustic wave fields are denoted at (x, y, z) as follows:

$$\begin{aligned} f_i(x,y,z,t) &= w_i\big(t - \frac{1}{v}(x\sin\theta_i\cos\phi_i \\ &+ y\sin\theta_i\sin\phi_i + z\cos\theta_i)\big), \\ &\text{where } i = 1,2,3. \end{aligned} \tag{1}$$

Here, $w_i(t)$ denotes the band-limited source signal:

$$w_i(t) = \int_{-\omega_i}^{\omega_i} A_i(\omega) e^{j\omega t} \mathrm{d}\omega, \tag{2}$$

where ω_i denotes the bandwidth of the each source signal. Angles, ϕ_i and θ_i denote the each propagating direction.

At the origin, $x = y = z = 0$, the temporal gradient can be obtained as:

$$f_{it}(0,0,0,t) = \frac{\partial}{\partial t} f_i(x,y,z,t)\Big|_{x=y=z=0} = \dot{w}_i(t). \tag{3}$$

Here, $\dot{w}_i(t)$ is the time-differential of $w_i(t)$ denoted as:

$$\dot{w}_i(t) = \int_{-\omega_i}^{\omega_i} j\omega A_i(\omega) e^{j\omega t} \mathrm{d}\omega \tag{4}$$

Orthogonal triplet of spatial gradients are derived as follows:

$$\begin{aligned} f_{ix}(0,0,0,t) &= \frac{\partial}{\partial x} f_i(x,y,z,t)\Big|_{x=y=z=0} \\ &= -\frac{\sin\theta_i \cos\phi_i}{v} \dot{w}_i(t) \end{aligned} \tag{5}$$

$$\begin{aligned} f_{iy}(0,0,0,t) &= \frac{\partial}{\partial y} f_i(x,y,z,t)\Big|_{x=y=z=0} \\ &= -\frac{\sin\theta_i \sin\phi_i}{v} \dot{w}_i(t) \end{aligned} \tag{6}$$

$$\begin{aligned} f_{iz}(0,0,0,t) &= \frac{\partial}{\partial z} f_i(x,y,z,t)\Big|_{x=y=z=0} \\ &= -\frac{\cos\theta_i}{v} \dot{w}_i(t) \end{aligned} \tag{7}$$

Instantaneous mixture

Based on the above equations (3),...,(7), the orthogonal spatial gradients of the observed signal, $f(x,y,z,t)$, at the origin can be denoted respectively as:

$$\begin{aligned} &\frac{\partial}{\partial x} f(x,y,z,t)\Big|_{x=y=z=0} \\ &= \sum_{i=1}^{3} f_{ix}(0,0,0,t) = \sum_{i=1}^{3} -\frac{\sin\theta_i \cos\phi_i}{v} \dot{w}_i(t), \end{aligned} \tag{8}$$

$$\begin{aligned} &\frac{\partial}{\partial y} f(x,y,z,t)\Big|_{x=y=z=0} \\ &= \sum_{i=1}^{3} f_{iy}(0,0,0,t) = \sum_{i=1}^{3} -\frac{\sin\theta_i \sin\phi_i}{v} \dot{w}_i(t), \end{aligned} \tag{9}$$

$$\begin{aligned} &\frac{\partial}{\partial z} f(x,y,z,t)\Big|_{x=y=z=0} \\ &= \sum_{i=1}^{3} f_{iz}(0,0,0,t) = \sum_{i=1}^{3} -\frac{\cos\theta_i}{v} \dot{w}_i(t). \end{aligned} \tag{10}$$

Therefore, the orthogonal spatial gradients of the observed signal can be denoted as the instantaneous mixture of the source signals as:

$$\nabla f(x,y,z,t)\Big|_{x=y=z=0} = -A \begin{pmatrix} \dot{w}_1(t) \\ \dot{w}_2(t) \\ \dot{w}_3(t) \end{pmatrix}, \tag{11}$$

where, the mixing matrix A can be defined as:

$$A = \begin{pmatrix} \frac{\sin\theta_1 \cos\phi_1}{v} & \frac{\sin\theta_2 \cos\phi_2}{v} & \frac{\sin\theta_3 \cos\phi_3}{v} \\ \frac{\sin\theta_1 \sin\phi_1}{v} & \frac{\sin\theta_2 \sin\phi_2}{v} & \frac{\sin\theta_3 \sin\phi_3}{v} \\ \frac{\cos\theta_1}{v} & \frac{\cos\theta_2}{v} & \frac{\cos\theta_3}{v} \end{pmatrix}. \tag{12}$$

Whereas, based on the wave equation of the airborne sound, particle velocity vector $\boldsymbol{v}(x,y,z,t)$ and the spatial gradients of the sound pressure $f(x,y,z,t)$ satisfy the following equation:

$$\rho \frac{\partial \boldsymbol{v}(x,y,z,t)}{\partial t} = -\nabla f(x,y,z,t), \tag{13}$$

where ρ is the density of the air. Therefore, (11) can be denoted as:

$$\rho \frac{\partial \boldsymbol{v}(x,y,z,t)}{\partial t}\Big|_{x=y=z=0} = A \begin{pmatrix} \dot{w}_1(t) \\ \dot{w}_2(t) \\ \dot{w}_3(t) \end{pmatrix}. \tag{14}$$

After the integration of the above equation for time, the following instantaneous mixture can be obtained:

$$\boldsymbol{v}(0,0,0,t) = A \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix}. \tag{15}$$

Blind source separation

The Blind Source Separation estimates the mixing matrix A and independent source signals $w_{i=1,2,3}(t)$ with following steps, from the observed particle velocity vector is broken down as:

$$\begin{pmatrix} v_x(t) \\ v_y(t) \\ v_z(t) \end{pmatrix} = \boldsymbol{v}(0,0,0,t). \tag{16}$$

Pre-whitening

The observed orthogonal triplet of particle velocity, $v_x(t)$, $v_y(t)$, and $v_z(t)$ can be whitened and normalized by the following matrix B [6, 7, 8]:

$$B = \begin{pmatrix} \frac{B_1}{\sqrt{\psi_{xx}} B_2} & \frac{\psi_{yz}\psi_{zx} - \psi_{xy}}{\sqrt{\psi_{xx}} B_1 B_2} & \frac{\psi_{xy}\psi_{yz} - \psi_{zx}}{\sqrt{\psi_{xx}} B_1 B_2} \\ 0 & \frac{1}{\sqrt{\psi_{yy}} B_1} & \frac{-\psi_{yz}}{\sqrt{\psi_{yy}} B_1} \\ 0 & 0 & \frac{1}{\sqrt{\psi_{zz}}} \end{pmatrix}, \tag{17}$$

where B_1 and B_2 are defined as follows:

$$B_1 = \sqrt{1 - \psi_{yz}^2}, \tag{18}$$

$$B_2 = \sqrt{1 - \psi_{xy}^2 - \psi_{yz}^2 - \psi_{zx}^2 + 2\psi_{xy}\psi_{yz}\psi_{zx}}. \tag{19}$$

The auto or mutual correlation functions are denoted as:

$$\begin{aligned} \psi_{\xi\eta} &= \frac{1}{T}\int_0^T v_\xi(t) v_\eta(t) \mathrm{d}t \\ &\quad \text{where} \quad \xi = x,y,z \quad \eta = x,y,z. \end{aligned} \tag{20}$$

The integral duration, T, should be sufficiently long for the statistical stability.

The normalized orthogonal triplet of the particle velocities, $u_1(t)$, $u_2(t)$, and $u_3(t)$ can be obtained as:

$$\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} = B \begin{pmatrix} v_x(t) \\ v_y(t) \\ v_z(t) \end{pmatrix}. \tag{21}$$

The normalized independent triplet of the particle velocities are obtained by homothetic transformations, $C_1(\alpha_1)$, $C_2(\alpha_2)$, and $C_3(\alpha_3)$, as:

$$\begin{pmatrix} u_1^{[\boldsymbol{\alpha}]}(t) \\ u_2^{[\boldsymbol{\alpha}]}(t) \\ u_3^{[\boldsymbol{\alpha}]}(t) \end{pmatrix} = C_1(\alpha_1)C_2(\alpha_2)C_3(\alpha_3) \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}, \quad (22)$$

where, the homothetic transformations are defined as following rotation matrices:

$$C_1(\alpha_1) = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 & 0 \\ \sin\alpha_1 & \cos\alpha_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (23)$$

$$C_2(\alpha_2) = \begin{pmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 \\ 0 & 1 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 \end{pmatrix} \quad (24)$$

$$C_3(\alpha_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_3 & -\sin\alpha_3 \\ 0 & \sin\alpha_3 & \cos\alpha_2 \end{pmatrix} \quad (25)$$

For the sake of the optimum angles, the steepest decent method minimizes the Kullback-Leibler divergence: $I(u_1^{[\boldsymbol{\alpha}]}, u_2^{[\boldsymbol{\alpha}]}, u_3^{[\boldsymbol{\alpha}]})$:

$$\begin{aligned} &I(u_1^{[\boldsymbol{\alpha}]}, u_2^{[\boldsymbol{\alpha}]}, u_3^{[\boldsymbol{\alpha}]}) \\ &\quad = H(u_1^{[\boldsymbol{\alpha}]}) + H(u_2^{[\boldsymbol{\alpha}]}) + H(u_3^{[\boldsymbol{\alpha}]}) \\ &\quad -H(u_1^{[\boldsymbol{\alpha}]}, u_2^{[\boldsymbol{\alpha}]}, u_3^{[\boldsymbol{\alpha}])}) \end{aligned} \quad (26)$$

where, $H(u_i^{[\boldsymbol{\alpha}]})$ denotes the entropy of $u_i^{[\boldsymbol{\alpha}]}$, $(i = 1, 2, 3)$. The joint entropy $H(u_1^{[\boldsymbol{\alpha}]}, u_2^{[\boldsymbol{\alpha}]}, u_3^{[\boldsymbol{\alpha}])})$ is not affected by rotation matrices. The simplified cost-function, therefore, is introduced as $J(\boldsymbol{\alpha})$:

$$J(\boldsymbol{\alpha}) = H(u_1^{[\boldsymbol{\alpha}]}) + H(u_2^{[\boldsymbol{\alpha}]}) + H(u_3^{[\boldsymbol{\alpha}]}). \quad (27)$$

Consequently, the iterative optimization process is denoted as following steepest decent method:

[step0]initialization

$$\boldsymbol{\alpha} := \mathbf{0} \quad (28)$$

[step1] modification

$$\Delta\boldsymbol{\alpha} = \begin{pmatrix} \Delta\alpha_1 \\ \Delta\alpha_2 \\ \Delta\alpha_3 \end{pmatrix} := -\mu\nabla J(\boldsymbol{\alpha}^{[n]}) \quad (29)$$

$$\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} := C_1(\Delta\alpha_1)C_2(\Delta\alpha_2)C_3(\Delta\alpha_3) \cdot \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}, \quad (30)$$

$$\boldsymbol{\alpha} := \boldsymbol{\alpha} + \Delta\boldsymbol{\alpha} \quad (31)$$

[step2] decision
while $\|\Delta\boldsymbol{\alpha}\| \geq \epsilon$, then [step1] else

$$\begin{pmatrix} u_1^{[\boldsymbol{\alpha}]} \\ u_2^{[\boldsymbol{\alpha}]} \\ u_3^{[\boldsymbol{\alpha}]} \end{pmatrix} = \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}. \quad (32)$$

Directions of arrival signals

From the above equations, (21) and (22),

$$\begin{pmatrix} u_1^{[\boldsymbol{\alpha}]}(t) \\ u_2^{[\boldsymbol{\alpha}]}(t) \\ u_3^{[\boldsymbol{\alpha}]}(t) \end{pmatrix} = C_1(\alpha_1)C_2(\alpha_2)C_3(\alpha_3)B \begin{pmatrix} v_x(t) \\ v_y(t) \\ v_z(t) \end{pmatrix}. \quad (33)$$

is derived. The above mentioned blind source separation yields the normalized independent signals. The separated signals and the original sound pressure, therefore, have following relation:

$$\begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} u_1^{[\boldsymbol{\alpha}]}(t) \\ u_2^{[\boldsymbol{\alpha}]}(t) \\ u_3^{[\boldsymbol{\alpha}]}(t) \end{pmatrix}, \quad (34)$$

where $\sigma_{i(=1,2,3)}$ are the standard deviation of the corresponding original sound pressure $w_{i(=1,2,3)}(t)$. The following identical equation is derived in consequence of equations (11), (33) and (34):

$$\begin{aligned} &\begin{pmatrix} 1/\sigma_1 & 0 & 0 \\ 0 & 1/\sigma_2 & 0 \\ 0 & 0 & 1/\sigma_3 \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix} \\ &\qquad = C_1(\alpha_1)C_2(\alpha_2)C_3(\alpha_3)BA \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix}. \end{aligned} \quad (35)$$

Solving the matrix equations as:

$$I = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} C_1(\alpha_1)C_2(\alpha_2)C_3(\alpha_3)BA, \quad (36)$$

we can estimate the standard deviations, $\sigma_{i(=1,2,3)}$, and the directions of arrivals, $\phi_{i(=1,2,3)}$ and $\theta_{i(=1,2,3)}$.

The orthogonal triplet of the particle velocities at single point can realize the autonomous directivity with the proposed spatio-temporal gradient blind source separation algorithm.

ACOUSTICAL EXPERIMENTS

We have implemented and tested the proposed spatio-temporal blind source separation algorithm with the Proof-Of-Concept (POC) model.

experimental setup

Figure 3(1) shows the external view of the 3-ch probe which has three orthogonally placed particle velocity sensors (Microflown sensor elements). The Microflown which is an acoustic sensor based on a differential hot-wire anemometer used in this experiment consists of two very closely spaced bridges (spacing 350μm) of silicon nitride with a platinum electrode pattern on top of them. Dimensions of the two wires (bridges) are $10000 \times 10 \times 0.5\mu\text{m}(d \times w \times h)$.

Fig.3(2) shows the block diagram of the POC-model. Under our experimental condition, the number of source signals are limited two, which are two women's voices (cf.Table 1).

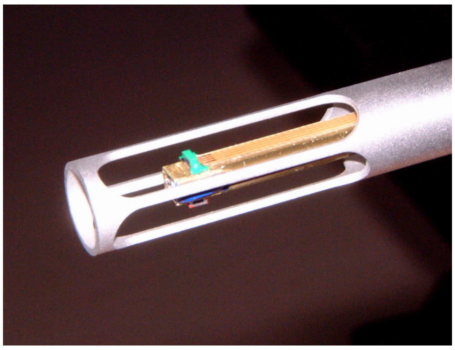


Fig. 2: A view of the 3-channel Microflown probe

Table 1: Specifications of source signals

	No.1	No.2
source	female voice, "NHK Stereo Broadcasting-Opening narration" recoded in 1954	female voice, "National Stereo Hall-Opening narration" recorded in 1961
length	10s	10s
incident directions	$\theta_1 = 0$(rad) $\phi_1 = 2\pi/3$(rad)	$\theta_2 = 0$(rad) $\phi_2 = \pi/3$(rad)
standard deviation normalized by No.1	$\sigma_1 = 1$	$\sigma_2 = 1.03$

Blind source separation

The original source sounds are recoded female voices shown in the Fig.4(1)and (2) respectively. The observed x and y directional particle velocities are shown in Fig.5(1) and (2) respectively. The estimated sound pressures of original sound sources are shown in Fig.6(1) and (2) respectively. Fig.7 shows the convergence plots of the estimated directions of the original sources No.1 and No.2.

Discussion

Table 2 shows that the proposed spatio-temporal gradient blind source separation method has an ability to separate the original sound sources from the orthogonal pair of observed particle velocities.

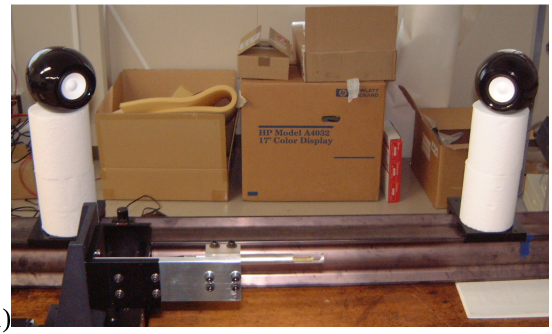
For the statistical stability, the blind source separation process is started 1sec after the signal arrival. From the results, the separation process converged in almost 1sec (cf. Fig.7). Focusing the duration from 3.5s to 4.0s, however, the separation result becomes worse. In this duration, only No.1 female voice exists and another female keeps silent (cf. Fig. 5). The demixing process cannot be derived correctly.

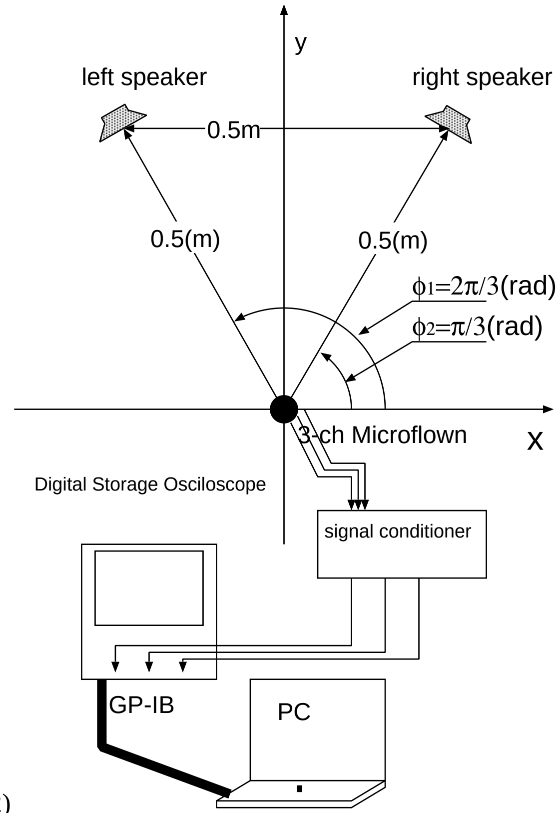
The convergence error is considered to be caused by the steady airflow in the experimental environment.

Table 2: Kullback-Leibler divergence between two signals

	Kullback-Leibler divergence
original voice No.1 and No.2	0.00197
observed particle velocities $v_x(t)$ and $v_y(t)$	0.113
separated signals $\hat{w}_1(t)$ and $\hat{w}_2(t)$	0.00279

(1)





(2)

Fig. 3: (1)A bird's-eye view of the POC model, (2) its block diagram

CONCLUDING REMARKS

This paper proposes the spatio-temporal gradient blind source separation method to be independent of source-locations, and presents simple digital signal processing algorithm characterizing the wave field where incident three plane waves overlap each other.

For detecting the directions of arrivals, spatio-temporal gradient analysis based on the linear dependency among the sound pressure, the orthogonal triplet of the particle velocity, is used. The acoustical experiments realized by Microflown

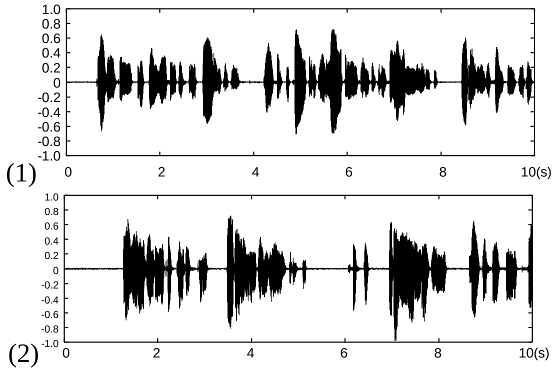


Fig. 4: (1)original sound source No.1, (2) original sound source No.2

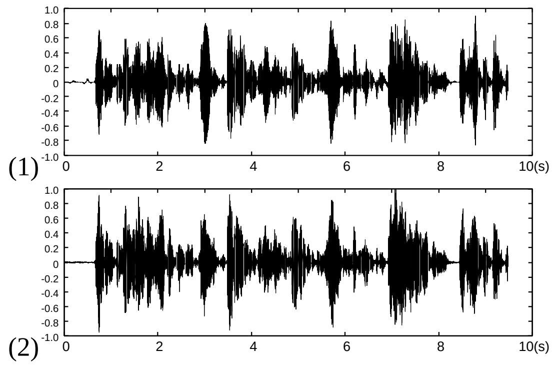


Fig. 5: (1)x-directional particle velocity, (2) y-directional particle velocity

system, were conducted with following conclusions and remarks:

1. the spatio-temporal gradient analysis has an ability to simplify the convolution blind source separation problems into the instantaneous blind source separation over the spatio-temporal gradient space,

2. the directions of wave propagations can be determined with the proposed blind source separation.

3. three sources can be discriminated autonomously even when they are slightly moving.

ACKNOWLEDGMENT

This paper was supported by Grant-in-Aid for Scientific Research(KAKENHI 15560360) from JSPS and MEXT of the Japanese Government.

REFERENCES

1 J.F.Cardoso, "Source separation using higher order moments", Proc. IEEE ICASSP, 4, (1989), 2109-2112.

2 D. T. Pham and J. F. Cardoso, "Blind separation of instantaneous mixtures of non-stationary sources," Proc. IEEE Trans. Signal Processing, **49**, (2001), 1837-1848.

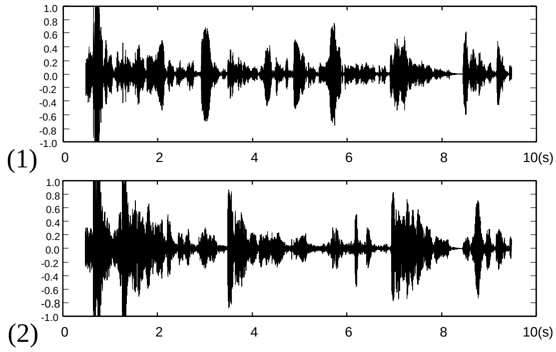


Fig. 6: estimated sound pressures of original sound sources (1)No.1, and (2)No.2

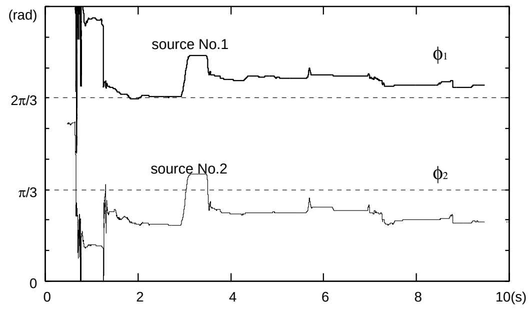


Fig. 7: Convergence plots of the estimated directions of the original sources No.1 and No.2

3 P.Comon, "Independent component analysis, A new concept?", Signal Processing Vol.36, (1994), 287-314.

4 K.Teramoto, "Crack detection by spatio-temporal gradient analysis of the acoustic surface wave field", Acta Acustic, 88-Suppl.1, (2002), 107.

5 K.Teramoto, "Index of Homogeneity for a Criterion Judging Existence of Cracks," Proc. of SICE2002, on CDROM, (2002)

6 H. Ogura and N. Nakasako : "Direct method of blind separation for two mixed signals based on the statistical indepedency from information theoretic approach," Proc. ISNIC-98, (1998), 215-220.

7 H.Ogura, N.Nakasako and N.Seo, "Blind separation of complex signals in terms of complex Hermite moments", 2000 IEEE int., Symp. on Intelligent Signal Processing and Communication Systems, Proceedings Vol.1, (2000), 429-433.

8 K.Tsuruta, N.Nakasako, H.Ogura, A Successive Method of Blind Signal Separation Based on Orthogonalization and Rotational Transformation, Proc. The 8th SICE Chugoku Branch Annual Conference, (1999), 62-63.