Separation of Gravity Anomaly Data Considering Statistical Independence of Source Signals

Prem Prakash KHATRI[1] · Riki HONDA[2] · Hitoshi MORIKAWA[3]

[1]Master course, Dept. of Civil Engineering, University of Tokyo,
E-mail: khatri-p@ip.civil.t.u-tokyo.ac.jp
[2]Member of JSCE, University of Tokyo,
E-mail: rhonda@k.u-tokyo.ac.jp
[3]Member of JSCE, Tokyo Institute of Technology,
E-mail: morika@enveng.titech.ac.jp

Gravity anomaly is one of efficient methods to evaluate underground structure, which is essential for estimation of ground motion due to earthquakes. Data observation is, however, costly since it requires expensive devices. In order to overcome this problem, Morikawa et al. have been working to develop a mobile gravimeter that uses force-balance (FB) accelerometer. In comparison to the conventional spring type gravimeters, it is less costly, compact and can be carried by relatively small carriers. However, it raised problems that the observed data are severely contaminated by various kinds of disturbances such as engine vibration and carrier motion. Therefore an appropriate data processing method for extracting gravity anomaly signal from such observed data is required.

For that purpose, we propose to use the statistical independence property of gravity anomaly and other noisy data. The gravity anomaly and other noises are generated from different sources and it can be safely assumed that they are independent.

As a scheme of considering independence of signals, blind source separation techniques are used. Second Order Blind Identification method (SOBI) separates the target sources by assuming that source and noises are un-correlated at various time-lags. Similarly, Independent Component Analysis (ICA) separates the sources by maximizing the independence of linearly transformed observed signals. An ICA algorithm namely ThinICA is proposed that implements the maximization of independence among source signals at various time-lags and thus incorporates the advantages of both SOBI and ICA.

The proposed method is applied to the data observed at Toyama Bay, Japan. It is observed that the motion of carrier (ship) influences the performance of de-noising algorithm. Under certain favorable data acquisition environment, the proposed method was able to salvage the gravity anomaly data from the noise-contaminated data with the accuracy sufficient for the purpose of identification of gravity anomaly distribution.

Key Words : *Gravity Anomaly, Force-Balanced Accelerometer, Statistical Independence, Blind Source Separation, Independent Component Analysis.*

1. INTRODUCTION

The information of local subsurface structure is essential for the evaluation of seismic ground motion [6]. For the survey of subsurface rock structure, gravity method has been one of the useful methods [8]. For the purpose of improvement of usability and applicability of gravity method, Morikawa et al. have been working to develop a gravity observation system using force-balance (FB) accelerometer. It makes the system compact and implements high mobility, because it can be carried by relatively small carriers.

However, it raised problems that the observed data are severely contaminated by various kinds of disturbances in a small size carrier like engine vibration, carrier acceleration and carrier tilting in addition to conventional noises such as sensor drifts, electrical noise etc. This happens as a result of the high sensitivity of FB sensor and its vulnerability to the high frequency noises. The frequency range of these noises is wide and their amplitudes can be 100,000 times larger than the gravity anomaly signal. Therefore an appropriate data processing method for extracting the gravity anomaly signal from such observed data is required.

For that purpose, we propose to use statistical independence property of gravity anomaly data and other noisy data components. Since the gravity anomaly and the other noises are generated by different physical processes it is

not unsafe to assume that they are statistically independent.

The Blind source separation (BSS) techniques are known for data separation considering statistical independence of signals. An advanced ThinICA algorithm is proposed that combines the merits of powerful SOBI method and ICA principles. ThinICA separates the sources by maximizing independence of linearly transformed observed data at several time-lags. It uses a peculiar contrast function (refer section 6) to measure the independence of signals.

The proposed method is applied to the real data observed at Toyama bay, Japan. Low pass filtering is employed as a pre-processing to ThinICA in order to filter out the high frequency components in the observed data. The reference data is generated for the same place with the help of gravity map provided by National Institute of Advanced Industrial Science and Technology (AIST), Japan. This reference data is used for checking the performance of proposed scheme by comparing it with the signal separated by proposed method.

This paper is organized as follows. Section 1 gives the brief introduction to the paper. The gravity anomaly, development of prototype mobile gravimeter and incurring problems of severe noise contamination followed by conventional data processing method are explained in section 2. The Statistical Independence and the Blind Source Separation (BSS) technique are defined in sections 3 and 4. SOBI and ICA are described in section 5 and section 6, respectively. Section 7 presents the description of proposed ThinICA algorithm. Section 8 presents the observed data obtained by using the prototype gravimeter. The application of proposed ThinICA method for gravity anomaly data separation and the results are presented in section 9. Section 10 concludes the paper.

2. GRAVITY ANOMALY

(1) Introduction

The acceleration due to gravity 'g' varies by a subtle amount with the lateral variation in density of rocks. This is known as Gravity anomaly. This concept is represented in the schematic diagram (Fig. 1) [8] If the layers 1,2,3,4 with increasing density values lay flat uniform laterally, there will be no gravity anomaly and the 'g' value is constant. However, if the density contrast occurs laterally as a result of structural uplift as seen in the central portion of the figure, gravity anomaly is observed. Thus, the density of the various components of the geologic column and

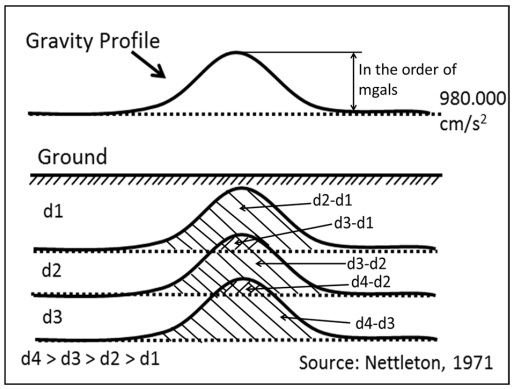

Figure 1 Density layers, Density contrasts and Gravity anomaly [8]

the resulting density contrasts among the rocks produced by structures developed in these rocks are related to gravity anomalies. Utilizing this relation, the structure of subsurface rocks can be estimated and sub-surface modeling can be done [1]. In fact, there are several factors that are responsible for Gravity anomaly such as Eötvös effect, latitude, altitude, Bouguer's effect, terrain effect etc. Eötvös effect is most significant and that occurs due to carrier motion and the rotation of earth. In order to map the gravity anomaly due to variation in rock density in a spatial plan, the data should be corrected for the factors that are significant. Resultant gravity anomaly data after necessary corrections can be correlated to the variation in densities of subsurface rock strata [8].

(2) Mobility of Gravity method

Conventionally, spring-type gravimeters have been used for gravity anomaly data observation. These gravimeters can provide accurate data with resolution of about 1 microGals. They are not sensitive to high-frequency disturbances and observed data is not distorted severely. However, they are expensive, difficult to handle, require large carriers and consume long time for data observation. Since they require large carriers, the data observation at shallow sea area may not be possible. In order to solve these problems and improve the mobility of gravity method, Morikawa et al. have been working to develop a compact, less expensive gravimeter that can be carried in a small carrier. The data observations are carried out by using the prototype mobile gravimeter and its performance is presented in this paper.

(3) Severe Contamination of Data

The prototype mobile gravimeter consists of FB sensor. Unlike spring-type gravimeter, FB sensor is highly sensitive to high frequency noises. As a result, the observed gravity anomaly data is added with the unwanted acceleration components that incur due to engine vibration and ship tilting. The sensor drift and electrical noise also add up to the data. The frequency ranges of noise are wide and their amplitudes are upto 100,000 times larger than gravity anomaly. An appropriate data processing methodology is needed in order to obtain gravity anomaly data from such observed data.

(4) Conventional Low pass filtering (LPF)

The noise contamination in the observed data obtained by conventional spring-type gravimeter was not severe. Low pass filtering (LPF) was often enough to remove the noises. Normally, the gravity anomaly data is expected to have low frequency and LPF filters out high frequency components. But it may impose the risk of loss of some useful data in unusual conditions. Moreover, LPF is unable to distinguish the unwanted data at lower frequencies. Since the data observed by prototype gravimeter consists of noises with wide frequency range, LPF alone is not enough and further improvement in the methodology is expected.

The simple Finite Impulse Response Low Pass Gaussian Filter (FIR LPGF or LPF) algorithm is presented as follows. Let us consider a time series $y(t)$ in which LPF is applied and y_f is the signal after filtering. Assuming a cut-off period t_c, time interval of data Δt and gaussian function f used for distribution of smoothing weights with standard deviation σ,

$$\sigma = t_c/(2\Delta t)$$
$$f = \frac{1}{K} e^{\frac{-x^2}{2\sigma^2}} \quad (1)$$

where K is a constant that normalizes sum of weights into 1, range of x can be chosen according to the need (such as -3σ to $+3\sigma$), n is the number of time intervals of $y(t)$.

3. STATISTICAL INDEPENDENCE

Two random variables y_1 and y_2 are said to be uncorrelated if their covariance is zero:

$$E\{y_1 y_2\} - E\{y_1\}E\{y_2\} = 0 \quad (2)$$

The concept of independence is stronger than uncorrelatedness. If y_1 and y_2 are two scalar-valued random variables, they are said to be independent if information on the value of y_1 does not give any information on the value of y_2, and vice versa. The joint probability density function (pdf) of y_1 and y_2, $p(y_1, y_2)$, is given by the product of marginal pdf $p_1(y_1)$ and $p_2(y_2)$ as

$$p(y_1, y_2) = p_1(y_1)p_2(y_2) \quad (3)$$

If $f_1(y_1)$ and $f_2(y_2)$ are two functions of two independent random variables y_1 and y_2 respectively, we have

$$E\{f_1(y_1)f_2(y_2)\} = E\{f_1(y_1)\}E\{f_2(y_2)\} \quad (4)$$

where E{...} denotes the expected value.

If the variables are independent, $E\{y_1 y_2\} = E\{y_1\}E\{y_2\}$ and they are always uncorrelated. However, the opposite is not always true as uncorrelatedness does not always imply independence.

4. BLIND SOURCE SEPARATION

Blind Source Separation (BSS) is a powerful tool that separates the sets of source signals blindly from the available mixed sets of signals with the help of no or very little information about the source signals and the mixing parameters. Let us imagine the sets of observed data $x_i(t)$ with the help of n sensors where t is the time index. Let us denote these source signals by $s_j(t)$. Assuming the source signals are linearly mixed, the observed data can be expressed as:

$$x_i(t) = \sum a_{ij} s_j(t) \quad (5)$$

where a_{ij} with $i, j = 1, ..., n$ are some unknown parameters that depend on the medium between the source of data and the sensor, i being the number of sensors n and j being the number of source data n ($i = j$ is not true always). Here, only x_i are known and both the matrix a_{ij} and source s_j are unknowns. This is the BSS problem and it does not have unique solution.

Assuming that mixing matrix a_{ij} is invertible, there exists a de-mixing matrix w_{ij} such that the sources s_i are separated as

$$s_j(t) = \sum w_{ij} x_j(t) \quad (6)$$

For the separation, the mixing parameters w_{ij} are arbitrarily chosen. The mixed set of signals are linearly combined with these arbitrary mixing parameters. There are certain measures of statistical independence that are used to measure these transformed mixed signals $w_{ij}x_j$. Then the measures of independence are maximized to make them mutually independent. Once their mutual independence is maximized, they are expected to be close to the source signals s_i.

5. SECOND ORDER BLIND IDENTIFICATION (SOBI)

SOBI is an advanced blind source separation technique that exploits the time coherence of the source signals [2]. It assumes that the random variables (or true source and noises in the context of this paper) are not correlated at all the time-lags with each other. This approach relies on stationary second-order statistics that are based on joint diagonalization of a set of covariance matrices that is based upon eigen-decompositon [2].

The data model is $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{n}$, which is similar to ICA model explained in section 5 and that fits to many situations of practical interest. Here, $\boldsymbol{x}$ is the observed data vector, $\boldsymbol{A}$ being the mixing matrix, $\boldsymbol{s}$ being the source vector and $\boldsymbol{n}$ being the noise vector.

Firstly, the whitening of observed data vector $\boldsymbol{x}$ is done into whitened data vector $\boldsymbol{z}$. This means the observed data are converted into signals that are uncorrelated with each other and have unit variances each. Thus the co-variance matrix of whitened data vector $\boldsymbol{z}$ is an identity matrix. Then the further sample co-variance matrices are computed for a certain set of time lags between the whitened data i.e., $z(t+\tau)$ and $z(t)$. for this whitened data vector $\boldsymbol{z}$. Then, these co-variance matrices are jointly diagonalized. This means all the sets of matrices are converted into identity matrices jointly. The diagonalizer matrix is combined with whitening matrix to determine the mixing matrix $\boldsymbol{A}$ and finally the source data vector is estimated by inversion.

6. INDEPENDENT COMPONENT ANALYSIS (ICA)

Independent Component Analysis (ICA) is a statistical and computational BSS technique for separating latent source data from its mixture with other signals. It was first introduced in 1980s in the context of neural network modeling [7]. Some highly successful new algorithms were introduced in mid-1990s. It has wide applications in the fields like biomedical signal processing, audio signal separation, telecommunication, financial time series analysis, etc. The features of ICA and its principle for the separation of sources are explained in the following based on the references[7].

(1) ICA Model

The random variables $x_i(t)$ with $i = 1, ..., n$ are observed which are modeled as linear combination of n random variables $s_i(t)$:

$$x_i = \sum a_{ij} s_j \tag{7}$$

It is also represented in a matrix form as

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \tag{8}$$

(2) Assumptions in ICA

In ICA, it is assumed that

1. The independent components are assumed statistically independent.
2. The independent components must have non-gaussian distributions. Since the higher order cumulants are zero for Gaussian distributions, ICA is impossible if the observed variables have gaussian distributions.
3. The unknown mixing matrix is assumed to be square. This simplifies the ICA estimation. By estimating the mixing matrix A, its inverse matrix W can be estimated and the independent components are easily computed as

$$\boldsymbol{s} = \boldsymbol{A}^{-1}\boldsymbol{x} = \boldsymbol{W}\boldsymbol{x} \tag{9}$$

It is also assumed that mixing matrix is invertible.

(3) Ambiguities of ICA

ICA has some ambiguity property.

The amplitude of the independent components cannot be determined. Since both $\boldsymbol{s}$ and $\boldsymbol{A}$ of ICA model in equation (10) are unknown, the infinite number of solutions for $\boldsymbol{s}$ and $\boldsymbol{A}$ are possible that give the same product $\boldsymbol{x}$. Any scalar multiplier in one of the sources s_i could be canceled by dividing the corresponding column a_i of $\boldsymbol{A}$ by the same scalar α_i as shown in

$$x = \sum_i (\frac{1}{\alpha_i} a_i)(s_i \alpha_i) \tag{10}$$

This problem is fixed by normalizing the independent components into unit variance, i.e., $E\{s_i^2\} = 1$. This still leaves the ambiguity of the sign, so it might often be necessary to multiply the independent components with -1, which does not affect the model.

It is also ICA's ambiguity that the order of the independent components cannot be determined.

(4) Non-gaussian is independent

The central limit theorem states that under certain conditions, the distribution of a sum of independent random variables tends towards a Gaussian distribution. Thus the sum of two independent random variables usually has a distribution that is closer to gaussian than any of the two original random variables. Conversely, if such data close to

gaussian are demixed by maximizing the non-gaussianity we tend to obtain the independent data.

Let us assume that the data vector $\boldsymbol{x}$ is a mixture of independent components as shown in equation (8). According to equation (9) the source independent components are again the linear mixture of $\{x_i\}$. Let us denote the estimated signal vector as $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$ where $\boldsymbol{W}$ is a de-mixing matrix to be determined. If $\boldsymbol{W}$ is equal to inverse of $\boldsymbol{A}$, $\boldsymbol{y}$ would be equal to $\boldsymbol{s}$. Following the principle of central limit theorem, we consider $\boldsymbol{W}$ a matrix of vectors that maximizes the non-gaussianity of $\boldsymbol{W}\boldsymbol{x}$ in order to determine $\boldsymbol{y}$ as close to $\boldsymbol{s}$.

(5) Measures of non-gaussianity

In order to maximize the non-gaussianity, we need a quantitative measure of non-gaussianity of a random variable. For simplification, it is often assumed that the signals have zero mean and unit variance. We do not lose generality and we follow the same assumptions in our approach.

Among various measure of non-gaussianity, one of classical measure is kurtosis or the fourth-order cumulant. Kurtosis of a signal y is defined as

$$\text{Kurt}(y) = \text{E}\{y^4\} - 3(\text{E}\{y^2\})^2 \tag{11}$$

Since y has unit variance, the relation simplifies to

$$Kurt(y) = E\{y^4\} - 3 \tag{12}$$

The kurtosis is zero for a gaussian random variable. For most nongaussian random variables, kurtosis is nonzero. Random variables with negative kurtosis are called subgaussian, and those with positive kurtosis are called supergaussian. However, nongaussianity is typically measured by the absolute value of kurtosis.

(6) Preprocessing for ICA

It is useful to conduct preprocessing, before applying an ICA algorithm on the observed data. Centering, whitening and filtering are the basic preprocessing techniques.

a) Centering

Centering refers to centering the observed vector $\boldsymbol{x}$ by subtracting it with its mean vector $\boldsymbol{m} = E\{\boldsymbol{x}\}$, in order to make $\boldsymbol{x}$ a vector with zero mean variables.

b) Whitening

Whitening refers to the transformation of the observed vector $\boldsymbol{x}$ linearly to make it white vector $\tilde{\boldsymbol{x}}$. A white vector has its components uncorrelated and their variances equal to unity. Naturally the covariance matrix of $\tilde{\boldsymbol{x}}$ is an identity matrix:

$$E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\} = \boldsymbol{I} \tag{13}$$

Whitening can be done by using the eigen-value decomposition (EVD) of the covariance matrix

$$E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\} = \boldsymbol{O}\boldsymbol{D}\boldsymbol{O}^T$$

where $\boldsymbol{O}$ is the orthogonal matrix of eigenvectors of $E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\}$ and $\boldsymbol{D}$ is the diagonal matrix of its eigenvalues, $\boldsymbol{D} = diag(d_1, ..., d_n)$. It is done as

$$\tilde{\boldsymbol{x}} = \boldsymbol{O}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{O}^T\boldsymbol{x} \tag{14}$$

Where the matrix $\boldsymbol{D}^{-\frac{1}{2}} = diag(d_1^{-\frac{1}{2}}, ..., d_n^{-\frac{1}{2}})$ and now $E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\} = \boldsymbol{I}$. Whitening transforms the mixing matrix into a new one, $\tilde{\boldsymbol{A}}$. From equations (8) and (14)

$$\tilde{\boldsymbol{x}} = \boldsymbol{O}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{O}^T\boldsymbol{A}\boldsymbol{s} = \tilde{\boldsymbol{A}}\boldsymbol{s} \tag{15}$$

Whitening makes the new mixing matrix $\tilde{\boldsymbol{A}}$ an orthogonal:

$$E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\} = \tilde{\boldsymbol{A}}E\{\tilde{\boldsymbol{s}}\tilde{\boldsymbol{s}}^T\}\tilde{\boldsymbol{A}}^T = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{A}}^T = \boldsymbol{I} \tag{16}$$

Whitening reduces the mixing matrix into orthogonal and thus eliminates the number of parameters to be identified from n^2 in case of original mixing matrix $\boldsymbol{A}$ to $n(n-1)/2$ in case of orthogonal matrix $\tilde{\boldsymbol{A}}$.

c) Filtering

When we are certain that the observed data consists of unwanted data at certain range of frequencies it might be useful to filter out them before processing by ICA. The time filtering of observed data vector $\boldsymbol{x}$ is done by multiplying it by a filtering matrix $\boldsymbol{F}$ as

$$\boldsymbol{x}* = \boldsymbol{x}\boldsymbol{F} = \boldsymbol{A}\boldsymbol{s}\boldsymbol{F} = \boldsymbol{A}\boldsymbol{s}* \tag{17}$$

which shows that ICA model still remains valid, with the same mixing matrix after applying a filter.

7. THIN INDEPENDENT COMPONENT ANALYSIS (ThinICA)

The contents in this section are referred to [5]. ThinICA is one of the several algorithms used in ICA. It uses a multivariate contrast function for the blind signal extraction of a selected number of independent components from their linear mixture. All the independent components can also be separated if needed. It combines the robustness of the joint approximate diagonalization techniques (such as SOBI) with the flexibility of the methods for blind signal extraction [5]. By maximizing the contrast function it gives two options: a) Hierarchical extraction based upon thinQR factorizations or b) Simultaneous extraction based upon thin Singular Value Decomposition (SVD) factorizations [5].

The signal model is identical to the general ICA model. The basic principle of source estimation is similar to the general principle of ICA described in previous Section 6..

The algorithm can be summarized as follows. The observed data $\boldsymbol{x}$ are first whitened by pre-whitening system $\boldsymbol{M}$. An arbitrary unitary matrix is chosen for initialization. The square matrices are formed using the contrast function. These matrices are diagonalized either by hierarchical or simultaneous approaches to determine new unitary matrix. Unitary matrix is updated that multiplies to whitened data to give the estimates sources $\boldsymbol{y}$ as shown below,

$$\boldsymbol{y}(t) = \boldsymbol{U}z(t) = \boldsymbol{UW}x(t) = \boldsymbol{UWA}s(t) \tag{18}$$

The difference in ThinICA is the measure of independence. It estimates the independent components by jointly maximizing a weighted square sum of cumulants of fixed order $q \geq 2$, determined by positive constants $t_1, ..., t_q$. In other words, the independence of signals measured by the contrast function are maximized at several time-lags defined. The contrast function is given as [5)]

$$\psi_{\Omega}(y) = \sum_{\tau \epsilon \Omega} \omega_{\tau} \mid Cum(y(t_1), ..., y(t_q)) \mid^2$$
$$\text{subject to } \|U\|_2 = 1 \tag{19}$$

where ω_{τ} are positive weighting terms and $\boldsymbol{U}$ is an unitary matrix with vectors of unit 2-norm. The sources are estimated by either hierarchically maximizing or simultaneously maximizing the contrast function. For simplifying the optimization of contrast function in equation (18), alternative similar contrast functions are utilized. Refer to [5)].

When the number of independent components to be extracted are equal to the number of sources and $q = 2$, it is equivalent to SOBI, based on the joint approximate diagonalization (JADE) of a certain set of cumulant slices. However, implementation is different. In other cases, it is superior to other algorithms since it offers both the extraction of selected number of components or separation of all the components. The ThinICA algorithms maximize the contrast function by combining simultaneously the several advantages of powerful techniques like Fast-ICA, JADE and SOBI.

8. DATA OBSERVATION

The gravity data survey was conducted at Toyama Bay, Japan on Oct 31, 2011. The prototype gravimeter set-up was mounted on a mid-size ship. The time history of the gravity data was recorded along a certain length of survey. The gravimeters were synchronized with precise GPS to have the spatial control of the recorded data and know the average speed of motion. The data acquisition environment was not stable throughout the survey. The ship speed and ship tilting frequency were changing from time to time. The latter incurred because of some inevitable circumstances like sea waves or varying wind speeds. The direction of ship was changing frequently near the bay.

Table 1 Features of Prototype EZ-GRAV

Description	EZ-GRAV
Sensor for gravimeter	VSE
Accelerometer (Titan)	2 Horizontal and 1 vertical components
Gradiometer	2 Horizontal components
Recording interval	0.01 sec
Input Voltage Range	± 10 V

Table 2 Basic specification of accelerometer VSE

Description	VSE-156SG
Observable dynamic range	± 50 Gal
Maximum Output Level	±10 V
Sensor Resolution (Accuracy)	$2 \sim 10 \times 10^{-6}$ Gal

This section presents the description of prototype gravimeter setup and its basic specifications, observed data by the sensors and the reference data generated by using gravity map provided by AIST (National Institute of Advanced Industrial Science and Technology).

(1) Prototype Gravimeter (EZ-GRAV)

EZ-GRAV uses VSE-156SG (hereafter VSE) by Tokyo Sokushin Company Limited as its sensor. It includes a gradiometer and an accelerometer. The observed data is recorded in digital format with 24 bit and 0.01 second interval. All the sensors are fixed on an aluminium thick plate and set inside a constant temperature reservoir, because the devices are sensitive to temperature and its fluctuation is much larger than the variation of gravity. The temperature is controlled within ±1°C. The features of prototype EZ-GRAV are listed in Table (1) and the picture of set-up can be seen in Figure (2).

(2) VSE (Analog Servo)

The basic specification of VSE is listed in Table (1). Owing to its high resolution, the VSE data is supposed to be the major data set amongst all the other data. The time series of recorded data can be observed in Figure (3) below. The data recording was started on 12:27:00. The sampling time was 0.01 second.

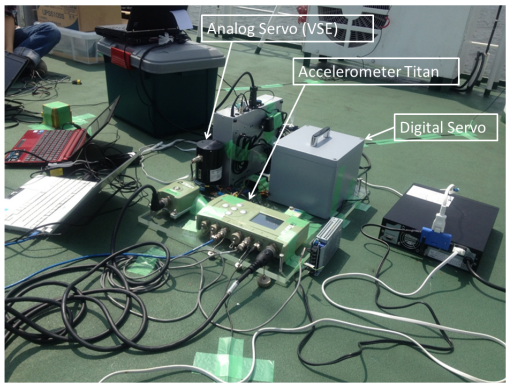

Figure 2 Prototype gravimeter setup (EZ-GRAV: Analog Servo, Accelerometer Titan).

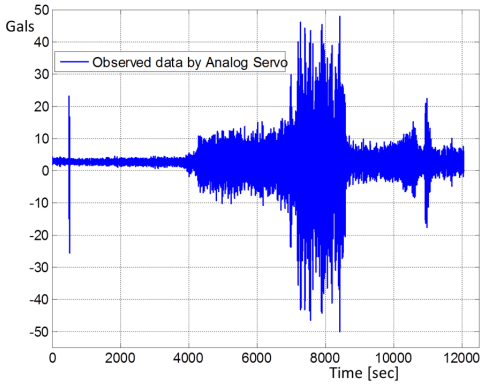

Figure 3 Observed time series gravity data (Gals) by Analog Servo (VSE) starting at time 12:27:00.

(3) Accelerometer Titan

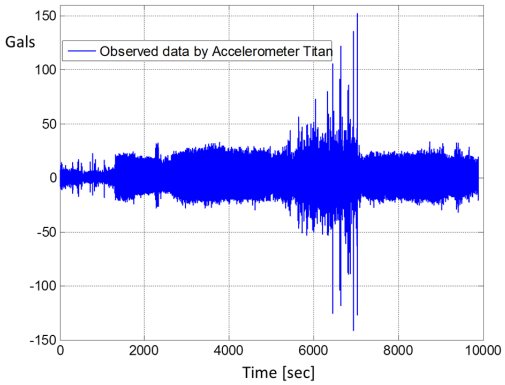

Figure 4 Observed time series gravity data (gals) by Accelerometer Titan (Taurus: vertical component), starting at time 12:53:06.

The accelerometer Titan namely Taurus consists of three sensors that are oriented at three different directions, i.e., North-South (NS component), East-West (EW component) and Up-down (vertical or UD component). These components constitute the component of gravity anomaly data combined with components of carrier accelerations.

The accelerometer Taurus has observable range of $\pm 4 \times g = 4 \times 980 gals$. It has very high observable dynamic range but they do not offer as high resolution as given by VSE. However, we require multiple sets of data for data processing by Blind Signal Separation techniques, and the data observed by this sensor is used as supplementary data to VSE. The vertical (Z) component of Taurus is shown in figure (4).

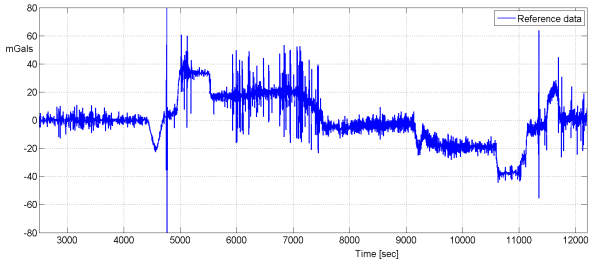

Figure 5 Reference data (Eötvös effect + Free air anomaly) obtained by using Gravity map produced by AIST, Japan, starting time: 12:17:00.

(4) Reference Data from Gravity map

The reference data is obtained by considering the Eötvös effect due to the ship movement and the free air anomaly. The free air anomaly is obtained from the gravity map prepared by AIST, Japan by using a shipboard gravity survey. They used accurate shipborne gravimeter in a much more stable environment using a large ship. This gravimeter was not sensitive to high frequency noise and reads only low frequency gravity data. Besides the accuracy of gravimeter, the survey was performed along the lines making a grid, and so the multiple data at points of intersection were averaged. Thus, the reference data is supposed to be accurate and reliable. And, the Eötvös effect is calculated from the position of the ship obtained by synchronized GPS. The reference data can be observed in the figure (5).

9. NOISE REDUCTION AND SEPARATION: APPLICATION OF ThinICA AND RESULTS

It is required to have at least two sets of observed data for processing by ICA. Multiple sets of observed data are input to BSS algorithms including proposed ThinICA and the results are compared. ICALAB toolbox [4)] in MATLAB is used for implementing these algorithms [3)].

(1) LPF as pre-processing

The presence of high frequency noise is seen to be unfavorable for the separation of data by using either SOBI or ThinICA. The application of LPF to the observed data is done as a pre-processing to either SOBI or ThinICA. The choice of appropriate cut-off period is important. The results are best for cut-off period around 100 sec.

(2) Combination of sensors

There are several sets of data such as VSE, Titan (Taurus-NS, EW and Z) etc. Out of all possible combinations, the association of VSE and Taurus-Z gave the best results.

(3) Pre-conditions for effective data extraction

The de-noising and data separation is found to be effective at certain pre-conditioning of data acquisition environment. The results are optimal for the portion of data where ship speed is low and the ship direction is uniform and tilting frequency is also low. While the ship is stopped, (ship stoppage time) frequency of tilting motion of the ship is high and the results are deviated from the reference data. Similarly, when the ship changes direction frequently near the bay, the results showed difference from the reference data.

(4) Comparison of Results

The filtered VSE and Taurus-Z data as an input data are shown in Figure (6). The separated output data by ThinICA and SOBI are shown in Figures (7) and (8), respectively. The comparison of LPF VSE data, output by ThinICA, SOBI and the reference data are compared in Figure (9). The results by SOBI and ICA are far improved than filtered VSE and followed the trend of reference data during the most of the period.

(5) Influence of Ship Motion

In the durations when stability of the motion of the ship is lost, results obtained by ThinICA were eccentric to the trend of reference data. For example, when the ship is stopped on the sea, fluctuation of the ship motion is relatively high and the separated singal was very different from the reference data.

While going away from the bay, the average ship speed was around 11 km/hr and the separation result obtained by ThinICA form the data obtained during that duration showed good agreement with the reference data. While returning back to bay, the ship velocity was roughly around

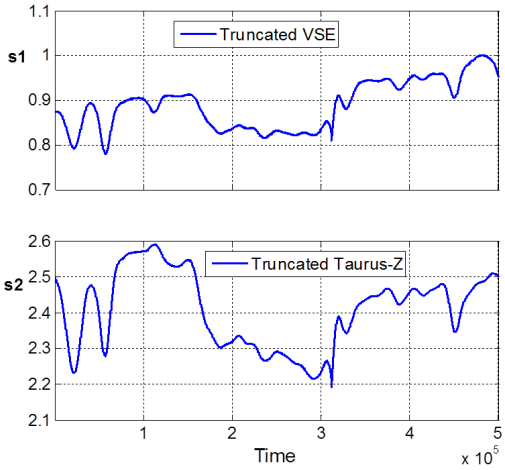

Figure 6 Truncated Input LPF VSE and LPF Taurus-Z data: Ship stoppage time (14:22:00 to 14:50:30) removed.

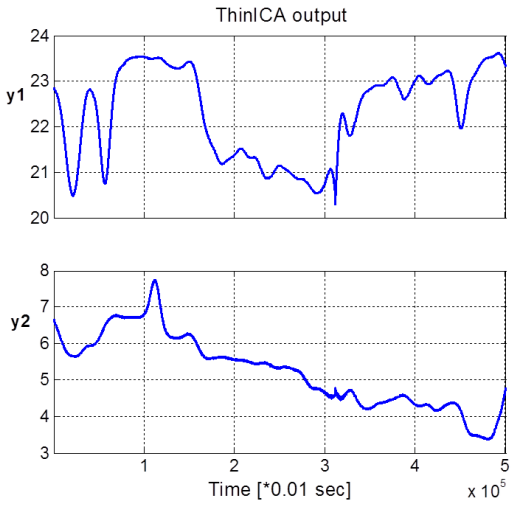

Figure 7 Output by ThinICA for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.

18 km/hr and the ship changed the direction often. Separation results from the data obtained during this period was not as harmonious as the former section when ship speed was lower.

These results indicate that stability of the ship motion, such as stability of velocity and direction, play an important role to describe the effectiveness of ThinICA.

10. CONCLUSIONS

The prototype mobile gravimeter developed by Morikawa et al. uses a force-balace (FB) type accelerometer sensor, which is sensitive to high frequency noises compared to the conventional spring-type gravimeter. The noise can be much larger than the gravity anomaly data

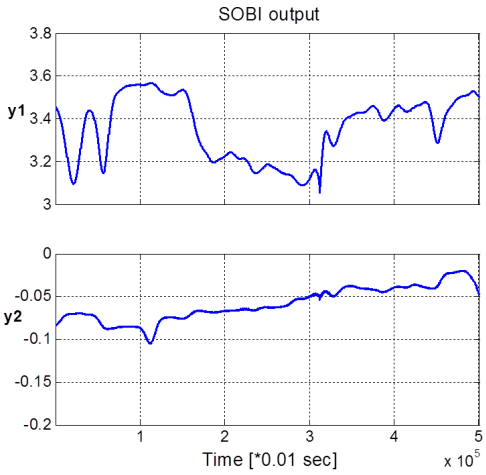

Figure 8 Output by SOBI for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.

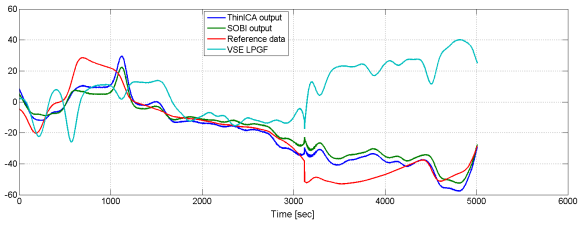

Figure 9 Comparison of output by ThinICA and SOBI with VSE LPF and Reference data (mGals), for truncated input with ship stoppage time removed: Both SOBI and ThinICA show improvements than LPF data with similar performance.

itself. In order to extract the gravity anomaly signal from the noise-contaminated observation data, an appropriate data processing methodology is needed.

Considering such background, we propose to use the advanced blind source separation (BSS) techniques that consider the statistical independence of source signals, because the gravity anomaly data and various noises are expected to be independent. Among various BSS methods, ThinICA algorithm was selected that maximizes the independence among signals at several time-lags. The method are supposed to have the merits of the method called the Second Order Blind Identification (SOBI) and conventional ICA.

In order to verify the performance of proposed method, the presented method was applied to the gravity data obtained by prototype gravimeter at the Toyama bay, Japan. The obtained results were compared with the high quality data generated by AIST.

The prototype by Morikawa et al. consists of multiple sensors. The Analog servo (VSE) is the main sensor and it must be accompanied by other sensor, because BSS requires at least two sets of data. Based on the comparison of computation results, it was suggested that the ThinICA results are good for the combination of VSE data with vertical component of Taurus.

Assuming that gravity anomaly data is dominant at lower frequencies, high frequency components are filtered out using LPF. Such frequency based filtering with an appropriate choice of cut-off period is realized to be important. The presence of high frequency noises is found to be unfavorable to data separation, because BSS methods work only after low pass filtering was conducted.

Also discussed is the influence of data acquisition environment. It was found that in the period when stability of the motion of the ship is lost, results obtained by the presented method deteriorates.

The agreement of ICA separated data with reference data verifies the applicability of the proposed method under certain conditions of data observation environment. The further improvement in data processing methodology is considered to be the part of future works.

The results are at acceptable limit for the purpose of subsurface modeling. Based on the current performance, it can be concluded that the results are encouraging. The consistency in the results in future and further improvements, if possible, will lead to the improvement in mobility of gravity method. The mobility of gravity method will not only facilitate the economic combination of multiple subsurface survey methods but also will facilitate the continuous sets of data leading to abundance of information on subsurface strata. The improved accuracy in subsurface modeling will contribute to improve the quality of GM simulation and seismic design.

ACKNOWLEDGMENT

We would like to thank AIST, Japan, for providing the data of gravity map for the Toyama bay. We acknowledge all the team members Professor Shigekazu Kusumoto (University of Toyama, Japan) for his support on field survey, Professor Hajime Chiba (Toyama National College of Technology, Japan) for allowing to conduct gravity survey in his ship, Mr. Satoshi Tokue and Ms. Yumiko Ogura at Tokyo Institute of Technology for retrieving field data and preparing reference data, and Professor Masayuki Saeki, Tokyo University of Science for his kind lecture about the devices.

We would also like to acknowledge Asian Development Bank Japan Scholarship Program (ADB-JSP) for provid-

ing scholarship to the first author for his study at master course at University of Tokyo. This research was partially supported by JSPS KAKENHI (21671003).

参考文献

1) M. Adachi, T.Noguchi, R.Nishida, I.Ohata, T.Yamashita, and K.Omura. Determination of subsurface structure of izumo plain, southwest japan using microtremors and gravity anomalies. In *The 14th World Conference on Earthquake Engineering*, 2008.
2) A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, and E.Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45:434–444, 1997.
3) A. Cichocki and S. Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Wiley, USA, first edition, 2003.
4) A. Cichocki, S. Amari, K. Siwek, T. Tanaka, and A.H. Phan. Icalab toolboxes version 3, 2007.
5) S. Cruces and A. Cichocki. Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ica. 2004.
6) H. Goto, S. Sawada, H. Morikawa, H. Kiku, and H. Ozalaybey. Modeling of 3d subsurface structure and numerical simulation of strong ground motion in the adazapari basin during 1999 kocaeli earthquake, turkey. *The Bulletin of Seismological Society of America*, 95:6, doi: 10.1785/?0120050002:2197–2215, 2005.
7) A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley and Sons Inc, USA, first edition, 2001.
8) L.L.Nettleton. *Elementary Gravity and Magnetics for Geologists and Seismologists*. The Society of exploration geophysicists, Oklahama, first edition, 1971.

(2012. 09. 22 受付)