Brain Activity Measurement during Program Comprehension with NIRS

Conference Paper *in* International Journal of Networked and Distributed Computing · June 2014

DOI: 10.1109/SNPD.2014.6888727

2 authors:

Yoshiharu Ikutani
Nara Institute of Science and Technology

Hidetake Uwano
Nara Institute of Science and Technology

International Journal of Networked and Distributed Computing, Vol. 2, No. 4 (October 2014), 259-268

Brain Activity Measurement during Program Comprehension with NIRS

Yoshiharu Ikutani [1] , Hidetake Uwano [2]

[1] *Dept. of Information Engineering,*
National Institute of Technology, Nara college,
22 Yata,Yamatokoriyama, Nara, Japan
E-mail: ikutani@info.nara-k.ac.jp

[2] *Dept. of Information Engineering,*
National Institute of Technology, Nara college,
22 Yata,Yamatokoriyama, Nara, Japan
E-mail: uwano@info.nara-k.ac.jp

Abstract

Near infrared spectroscopy (NIRS) has been used as a low cost, noninvasive method to measure brain activity. In this paper, we measure the effects of variables and controls in a source code on brain activity during program comprehension. The measurement results are evaluated after noise reduction and normalization to statistical analysis. As the result of the experiment, significant differences in brain activity were observed at a task that requires memorizing variables to understand a code snippet. On the other hand, no significant difference was observed between different levels of mental arithmetic tasks. We conclude that the frontal pole reflects workload to short-term memory caused by variables without affected from calculation.

Keywords: Program comprehension, Brain activity measurement, NIRS, Frontal pole

1. Introduction

Near infrared spectroscopy (NIRS) has been used in various research fields as a low cost, noninvasive method to measure brain activity. Using the differential of light absorptivity between oxy-hemoglobin (Oxy-Hb) and deoxy-hemoglobin (deOxy-Hb), NIRS measures cerebral blood flow to estimate a regional brain activity[1].

In the domain of program comprehension, Siegmunt et al. insisted on the need to identify the brain areas that are activated during program comprehension[2,3]. To quantify an action on program comprehension, Nakagawa et al. measured cerebral blood flow of participants who simulate source code psychologically using NIRS[4,5]. NIRS is an suitable method especially for a program comprehension research in various devices to measure brain activity, because of high temporal resolution and low restriction on participant. Program comprehension consists of many factors such as number calculation, variable memorizing, and understanding of conditional branches. However, effects of each factors in program comprehension to brain activity is still unclear. To understand how programmers comprehend program source code, identifying the effect of each factor is required. We think brain activity measurement in program comprehension proceeds a quantitative analysis of program comprehension processes based on Neuroscience and Neuropsychology.

In this paper, we investigate the effects of each factors in source code from programmer's brain activity. We focus on two factors in the experiment; variables and conditional branches. Variables and conditional branches are primary elements in source code, and essential factors in program comprehension. We measure participant's brain activity during tasks that have them read code snippets, and compare measurements statistically.

2. Related Work

2.1. Brain Measurement with NIRS

NIRS is a low cost and noninvasive method to measure brain activity. In various research fields, the device has been used to measure brain activities related to language, auditory, and motor functions[6]. NIRS is light weight and tolerant towards electrical noises, hence, the device is applied to BMI (Brain Machine Interface) domain as a method to measure brain activity[7].

NIRS has higher temporal resolution and lower restriction on participant in comparison with other methods such as PET, fMRI, EEG, and MEG[8]. On the other hand, spatial resolution of NIRS is coarse and the measurement depth is limited to a surface of the brain. Also the most of NIRS devices measure brain activity as a relative value based from an initial value. Therefore comparison of measurements under different conditions is difficult. Additionally, the method is sensitive to physiological noises including body motion, heartbeat, and respiration[9]. Therefore, proper noise reduction is crucial to analyze experiment results.

To cope with such problems, several methods have been proposed. Mitsuya et al. focused on time series of measurements obtained from NIRS, and proposed a method to eliminate biological noises using trend analysis and moving average method[10]. Tsunashima et al. proposed a noise reduction/normalization method using discrete wavelet transform and Z-score transformation to statistical processing[11]. In this study, we use the Tsunashima's method for noise reduction and statistical analysis.

2.2. Brain Measurement in Program Comprehension

Siegmunt et al. stated the need to identify the brain areas that are activated during program comprehension, and proposed an experimental design to measure program comprehension based on fMRI[2,3]. An introduction of brain activity measurement into program comprehension research allows us to observe directly what is happening inside the brain during program comprehension. The result of the measurement are essential evidence to understand the difference between good programmer and bad programmer, and to get valuable suggestion to develop programmer support systems.

Only a few studies report the measurement results of brain activity during program comprehension[3,5]. Nakagawa et al. measured brain activity during program comprehension using NIRS[4,5]. To quantify program comprehension process, they measured participant's cerebral blood flow during simulating a source code. The result of the experiment showed that brain activity differed according to the task difficulty, and the largest positive blood flow was observed at from initial to middle phase of the task.

Program comprehension consists of many factors such as number calculation, variables memorizing, and understanding of conditional branches. Their code simulation task consist of several factors that affect differently to the brain activity. To understand the relationship between brain activity and program comprehension, measuring the effect of each factors on brain activity is inevitable. In this study, we measure the effects of variable memorizing and conditional branches through an experiment.

3. Measurement with NIRS

3.1. NIRS

Increasing of neural activity is accompanied by blood flow increasing. In order to estimate brain activity, NIRS measures blood flow change in the brain that follows neural activity. Compare with the other methods, NIRS has some advantages such as high temporal resolution and low restriction of participants. The high temporal resolution allows us to analyze brain activity in detail, and the low restriction enables measurement on more practical conditions of programming.

An increase in blood flow causes hemoglobin density changes in the same region. Specifically, Oxy-Hb increase and deOxy-Hb decrease are observed in a region where neural activity increases[12]. Oxy-Hb and deOxy-Hb have different absorptivity to near infra-red light. NIRS measures blood flow changes related to neuronal activity by observe the difference of near infra-red light[6]. In general, Oxy-Hb is considered as the best index of brain measurement experiment with NIRS.

However most of NIRS devices that are now broadly used cannot measure optical path length that is essential for identifying blood flow changes as absolute value. Consequently, measurements become relative value derived from the value that is measured at start-up. It means that comparison between individuals or regions in the brain is inappropriate. Also NIRS is sensitive to physiological noises such as body motion, heartbeat, and respiration[9]. Therefore, proper noise reduction and normalization are crucial to analyze an experiment result.

3.2. *Wavelet-based Multi Resolution Analysis*

Fig. 1 shows an example of the measurements in our experiment. The horizontal axis represents task time, and the vertical axis represents Oxy-Hb value; higher value means higher regional brain activity. The figure shows that the wave contains many noises derived from heartbeats and/or respirations. In this study, we use the wavelet-based noise reduction proposed by Tsunashima et al[11].

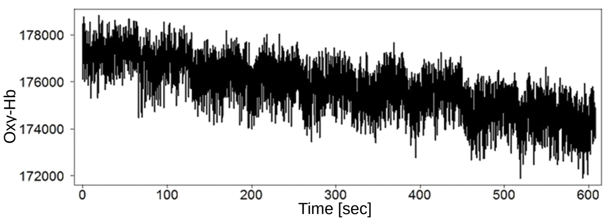

Fig. 1. Original signal from NIRS.

Discrete wavelet transform (DWT) is a method that decomposes signals into an approximated component and detailed components. Fig. 2 shows the decomposing result of the wave shown in Fig. 1. In the figure, $d1$-$d15$ mean the decomposed components of the original wave and each has different frequency. The components that higher than 1Hz (from $d1$ to $d7$) are considered as measurement noises caused by body motion.

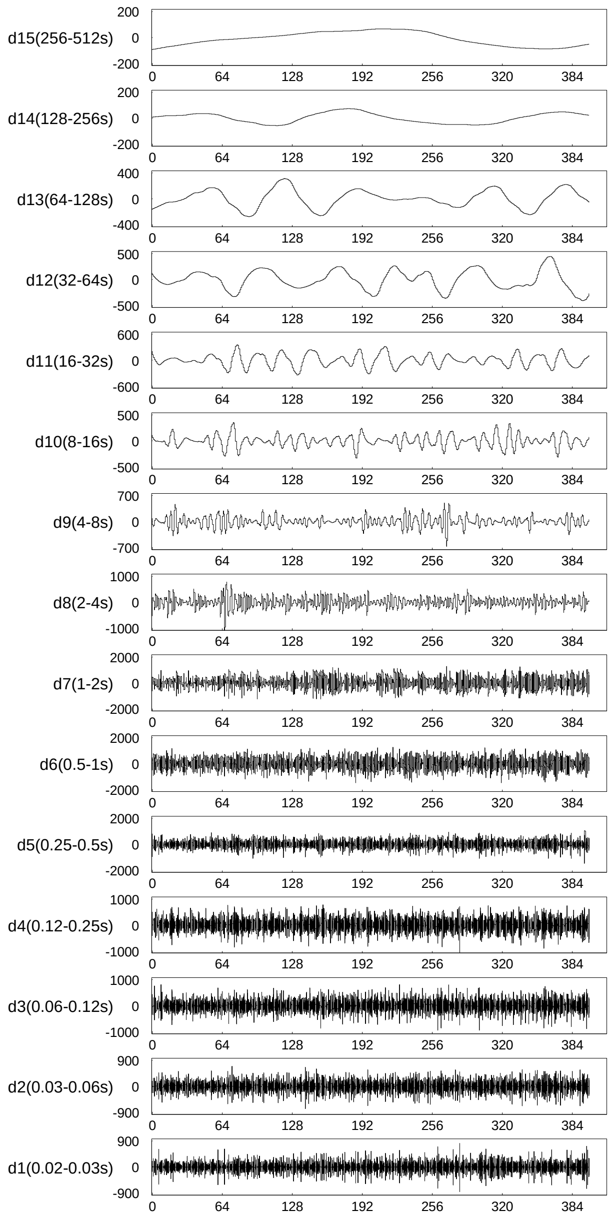

Fig. 2. Decomposed waves.

Also the 0.015-0.50Hz components ($d8$ and $d9$) and the 0.005-0.15Hz components ($d10$ and $d11$) are noises caused by respiration or blood pressure change. In contrast, low frequency components ($d12$, $d13$, and $d14$) are considered as signals derived from brain activity that we focused on. A noise-reduced wave is reconstructed by removing the above-mentioned noise components and combine the residual components. Procedure of the noise reduction method that we use in this paper is described as follows[11]. First, decompose an wave into multiple components that each has different frequency. Then, the components decided as a noise are removed, and finally noise-reduced wave is reconstructed from the residual components.

3.3. Z-score

Signals corrected from NIRS are quantity of relative changes using the start-up value as reference: hence comparison of the signals between subjects is inadequate. In this paper, we normalize signals into Z-scores to compare the signals between subjects and to adopt statistical analysis[11]. The method converts the NIRS signals reconstructed by the procedure described in section 3.2. The method converts signals into Z-scores using the following expression. Here, X is recorded signal during one task, μ and σ are a mean value and a standard deviation respectively.

$$Z = \frac{X-\mu}{\sigma} \tag{1}$$

4. Experiment

We prepared two type of tasks, and measured participant's brain activity during the tasks using NIRS. Eleven male undergraduate students participated in our experiment. All were right-handed and have finished their first programming lecture before the experiment.

4.1. Task

Two types of tasks were prepared in this experiment. In *Program* task, participants read a code snippet to calculate the value of variables. In *Arithmetic* task, participants answer a mental arithmetic question. Both tasks consist from three problem sessions and four rest sessions. At before each problem session, a rest session that is used as a baseline of the subsequent problem session is performed. Both problem and rest session continue 32 seconds. During the rest session, subjects are asked to gaze a cross marker displayed on a center of a computer screen.

4.1.1. Program Task

Subjects are asked to think a consecutive code snippet displayed on a screen silently, and not to answer by voice or keyboard input. Three types of question are used. All types of questions require to calculate three variables (a, b, and c) in their mind. Fig. 3 shows an example of *Program* task. Details of each questions are as follows:

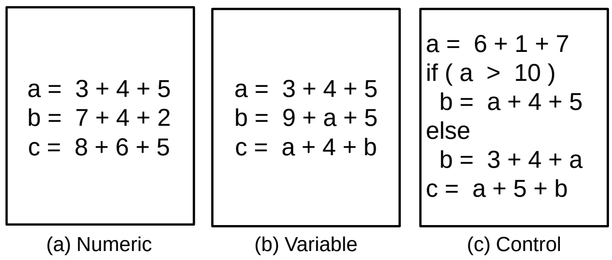

Fig. 3. Example of *Program* task.

- Numeric
  Each question consists of three lines of code. Each line calculates a value of variable from three integers.
- Variable
  Each question consists of three lines of code. Each line calculates a value of variable from integers and other variables. This question requires to memorize the value of the variables.
- Control
  Each question consists of six lines of code include the if – else statement. Subjects read four lines from the code snippet according to the condition. This question requires memorizing the value of the variables and judge the if – else branch.

A number of questions in each session is adjusted according to workload of each question. In the experiment, we assign four questions at *Numeric* session (8 seconds per question), three questions at *Variable* session (10.67 seconds per question), and two questions at *Control* session (16 seconds per question) respectively.

While *Numeric* consists of only numbers, *Variable* consists of numbers and variables. We expect that an effect of a variable on the brain activity is observed between *Numeric* and *variable*. Similarly, *Control* consists of same factors of *Variable* and a conditional branch (if – else branch). We therefore expect that an effect of a conditional branch is observed between *Variable* and *Control*.

4.1.2. Arithmetic Task

Subjects are asked to think a consecutive of mental arithmetic question displayed on a screen silently, not to answer by voice or keyboard input. Three types of question are used.. This task is the same task of Tsunashima's experiment[11]. We employed the task to validate the our experimental setting.

Three difficulty levels of question are used in the task. All types of questions require to calculate an answer of an equation. Fig. 4 shows an example of *Arithmetic* task. Details of each questions are as follows:

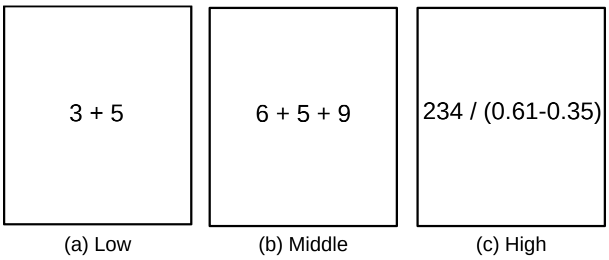

Fig. 4. Example of *Arithmetic* task.

- Low
  Addition of two one-digit numbers.
- Middle
  Addition of three one-digit numbers.
- High
  Subtraction and division of two decimals and one three-digit number.

A number of questions in each session is adjusted according to workload of each question. In the experiment, we assign 16 questions at *Low* session (2 seconds per question), 10 questions at *Middle* (3.2 seconds per question), and two questions at *High* (16 seconds per question) respectively.

4.2. Settings

One-ch NIRS device NeXus10 (TMS international BV) is employed in the experiment to record a brain activity. Fig. 5 and Fig. 6 show the appearance of the device and measurement state. The experiment is performed in a silent room which one subject and two observers are in. In order to restrain artifacts caused by subject's body motion, he sit on a chair which has armrests and a headrest, and are asked to be on steady condition.

We set the device on the forehead of subject, then measure his brain activity at 128Hz sampling frequency. The measured region is the front of the frontal lobe, i.e. "frontal pole". Frontal pole is considered as the region related to short-term memory and higher-order function such as action planning. It is expected that the region activates in *Program* task by memorize the value of variables and by judging the if – else conditions.

The experiment follows the procedure below.

1. Explain the experiment
2. Set a device to the participant
3. Perform *Program* task and recode brain activity
4. Three minutes break
5. Perform *Arithmetic* task and recode brain activity
6. Remove the device

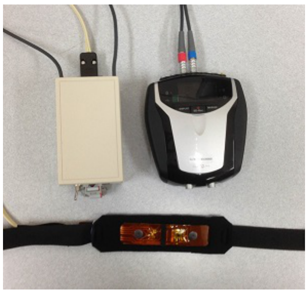

Fig. 5. Measurement device NeXus10.

Fig. 6. Measurement state.

4.3. Evaluation

The NIRS device measures the changes in Oxy-Hb, deOxy-Hb, and Total-Hb that means the sum of the changes in Oxy-Hb and deOxy-Hb. We use the Oxy-Hb as the evaluation metric for brain activity because the value has a better reflectivity to blood flow changes than the other metrics[12].

First, the original Oxy-Hb signal is decomposed into several components that each has different frequency. A cerebral blood flow reflects a neural activity slowly at a second-order. Therefore the components that have a frequency higher than 1Hz are considered as measurement noises caused by subject's body motion and/or others. Also the 0.15-0.50Hz components are noises caused by respiration, and the 0.05-0.15Hz components are noises caused by blood pressure change[9]. The components regarded as noises are eliminated, and then the residual components ($d12$, $d13$, and $d14$ in Fig. 2) are reconstructed as a task-related signal. Then, converts the noise-reduced signal into Z-score to enable comparison between subjects and for statistical analysis. Finally, calculates brain activity that changes by each session of the task. The brain activity that originated from each problem session is calculated from the difference between the problem session and the former rest session. Let f_0 is a Oxy-Hb value of a session start and f_n is a value of a session end, the brain activity during the problem session F is denoted as following equation.

$$F = \{f_0, f_1, f_2, \cdots, f_{n-1}, f_n\}$$

Also, the brain activity during the rest session which before the F is denoted as following equation.

$$R = \{r_0, r_1, r_2, \cdots, r_{n-1}, r_n\}$$

The brain activity that originated from each problem session *activity* is formulated as follow:

$$\begin{aligned} rest &= \frac{1}{n}\sum_{i=0}^{n} r_i \\ activity &= \sum_{i=0}^{n} (f_i - rest) \cdot \tau \end{aligned}$$

Here, *rest* means the average value of the former rest session, and τ denotes the inverse of sampling frequency.

5. Result

5.1. Program Task

Fig. 7 shows the average brain activity of all subjects in *Program* task. The horizontal axis represents a task time, the vertical axis represents Oxy-Hb value converted into Z-score. Each character on the top shows the session type at each period. The figure describes that the brain activity increased at the problem sessions and decreased at the rest sessions. Also the figure shows that the brain activity was much increased at *Variable* than the other problem sessions.

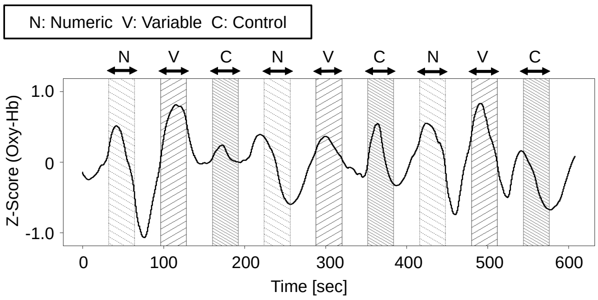

Fig. 7. Brain activity in *Program* task.

Fig. 8 shows the brain activity of all subjects in each problem session at *Program* task. The vertical axis shows the average *activity*; higher the value means higher brain activity. The figure shows the brain activity at *Variable* session is higher than the other sessions. The result of the Ryan's method describes significant differences between *Numeric* – *Variable* (p=0.01) and *Variable* – *Control* (p=0.003). In the *Variable* session, subjects were required to memorize the value of variables. The result suggests a workload to short-term memory that is caused from to memorize the variables increases the brain activity. The result indicates that frontal lobe relates to understand variables in a source code.

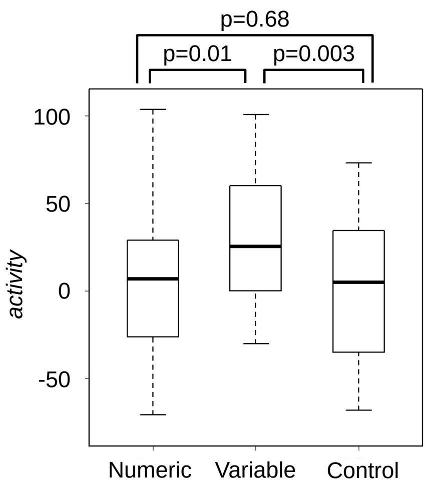

Fig. 8. Brain activity of each question in *Program* task.

In contrast with *Variable*, *Control* shows no differences between *Numeric* session. The possible causes of the result are follows:

- processing the condition of if – else statement causes few brain activity of frontal lobe.
- the fewer number of problems on *Control* session (two problems per session) compared with *Variable* session (three problems per session) reduces the workload.

We assigned the least number of the questions in *Control* session, because each question of *Control* session has the largest code snippet. However the number of lines which subjects actually read during the *Control* question (four lines) is similar to other questions (three lines in *Arithmetic* and *Variable*), hence no difference were observed in the experiment. Clarifying the effect of if – else and other conditional branch to the brain activity is our future work.

5.2. Arithmetic Task

Fig. 9 shows the average brain activity of all subjects in *Arithmetic* task. The figure shows a large increase in brain activity at the first *Middle* session. In contrast, the second session of *Low* and *High* increased largely, and the brain activity in the second *Middle* session was decreased.

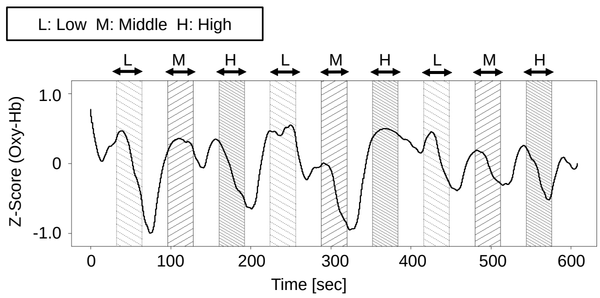

Fig. 9. Brain activity in *Arithmetic* task.

Fig. 10 shows the brain activity of all subjects in each problem session at *Arithmetic* task. There were no significant differences between sessions.

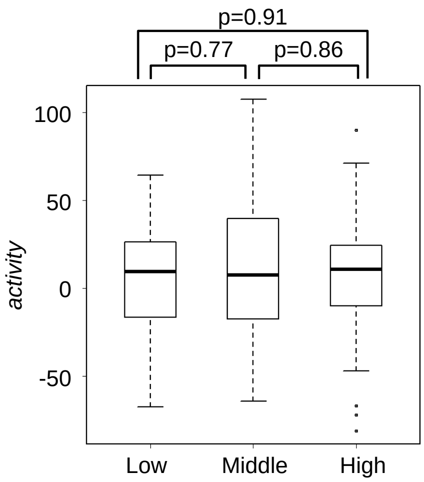

Fig. 10. Brain activity of each question in *Arithmetic* task.

Tsunashima et al. showed that there were significant differences between *Low* – *High*, and *Middle* – *High* of the *Arithmetic* task [11]. However in our experiment there was no significant difference between the difficulties. In our experiment, we select the front of the frontal lobe (frontal pole) as a measurement region, and Tsunashima et al. measured the dorsolateral left prefrontal cortex. Frontal pole is considered as the region related to short-term memory and high level cognitive activity such as action planning. The measurements on the frontal pole showed no significant difference between the mental calculation difficulties[11]. These results indicate that the frontal pole is not activated by calculation in contrast to *Program* task, and that activated brain regions differ by type of tasks; therefore, selection of target region is important at measurement of program comprehension study.

6. Discussion

6.1. Brain Activity in Program Comprehension

The result of the *Program* task describes brain activity on *Variable* was higher than *Numeric* and *Control*. In the *Program* task, both variable memorizing and calculation are the possible cause of the brain activity. However, the result of the *Arithmetic* task shows that workload from calculation does not affect to frontal pole. Hence, the whole result shows that frontal pole reflects workload to short-term memory caused by variables without affected from calculation. Most of program include simple calculation such as increment of index value. Measurement of programmer's frontal pole may allows us to analyze the effect of memorization or judgment in program comprehension.

6.2. Time Resolution of Brain Activity

The noises in the original signals were eliminated by wavelet-based multi resolution analysis in our experiment. Because of the wavelet transform characteristic, low-frequency components have coarse time resolution. The time resolution of reconstructed wave therefore becomes coarse. Hence, the noise reduction method reduces one of the NIRS advantages: high sampling frequency. More fine grained time resolution analysis with noise-reduction will enable to analyze brain activity changes on program comprehension in detail.

In this experiment, we eliminated the components regarded as noises (from $d1$ to $d11$ and $d15$ in Fig. 2) and reconstructed task-related wave from the residual signals ($d12$, $d13$, and $d14$). Cerebral blood flow reflects neural activity slowly at seconds-order. Hence, the components that has frequencies higher than 1Hz (from $d1$ to $d6$) are considered as measurement noises. Also the components of seconds-order contains physiological noises from respiration and/or blood pressure change. On the other hand, these components also include the task related effects since cerebral blood flow reflects a neural activity at seconds-order. In this experiment, $d10$ and $d11$ components (cycle ranged from 8 seconds to 32 seconds) which correspond to the length of each question will contain the effect of the task. Fig. 11 shows the reconstructed wave from $d10$ to $d14$ component. In this figure, brain activity in *Numeric* session (four questions per session) moves up and down more frequently compared with *Control* session (two questions per session). The figure shows that the wave reflects detailed changes of the brain activity than the wave which excludes the $d10$ and

$d11$ shown in Fig. 7.

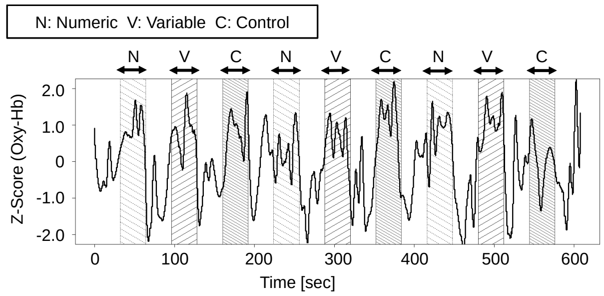

Fig. 11. Reconstruct wave include $d10$ and $d11$ component.

Fig. 12 describes the *activity* value from the detailed waves of all subjects in *Program* task. The figure shows that significant differences ($p < 0.05$) appear between *Numeric* – *Variable* and *Variable* – *Control* in the same as the Fig. 8. The result suggests that the reconstructed wave from components which contain both effects of a task and noises (i.e. $d10$ and $d11$ in this paper) is a useful for fine-grained analysis of the developer's brain activity. To identify task-related components among all decomposed components and establish more efficient analysis method is our future work.

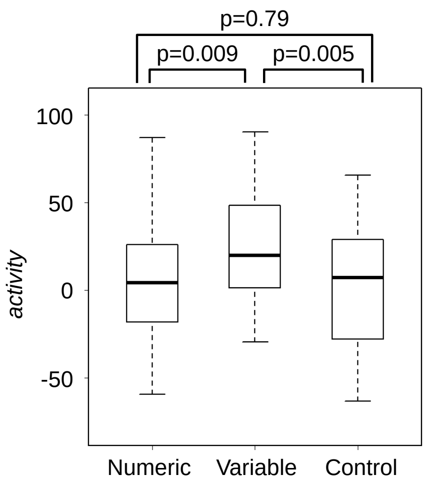

Fig. 12. Evaluation result of the detailed waves.

7. Conclusion

In this paper, we measured the brain activity on two task types to investigate the effects of variables and controls in a source code during program comprehension. We measured blood flow of frontal pole during the tasks that have subjects read code snippet or do mental arithmetic. Noise reduction using wavelet-based multi resolution analysis and normalization by Z-score conversion were used for statistical comparison between the problem sessions in each task.

As a result, significant differences in brain activity were observed between *Variable* and other problem sessions in *Program* task. In contrast, no significant difference was observed in *Arithmetic* task; that means workload that derives from calculation has no effect on the frontal pole. The result suggests that the frontal pole reflects workload to short-term memory caused by variables without affected from calculation. Therefore, measuring the frontal pole activity is an useful method to quantify the workload on short-term memory during the program comprehension task.

In this paper, the effect of the if – else statement to brain activity was not observed. To clarify the effect of conditional branch, additional experiments that adjusts the number of the question at *Variable* and *Control* session is required. As a future work, we plan to brain activity analysis of other conditional branch such as "for" and/or "while". Also the measurement employing large size source code such as a function or whole program is an interesting setting.

Acknowledgments

Part of this work was supported by JSPS KAKENHI Grant-in-Aid for Young Scientists (B) Grant Number 24700038.

References

1. E. Watanabe, Y. Yamashita, A. Maki, Y. Ito, H. Koizumi, "Non-invasive Functional Mapping with Multi-channel Near Infra-red Spectroscopic Topography in Humans," Neuroscience Letters, Vol.205, pp.41–44, 1996.
2. J. Siegmund, A. Brechmann, S. Apel, C. Kastner, J. Liebig, T. Leich, G. Saake, "Toward Measuring Program Comprehension with Functional Magnetic Resonance Imaging," In Proceedings of the ACM SIG-

SOFT 20th International Symposium on the Foundations of Software Engineering (FSE '12), No.24, 2012.
3. J. Siegmund, C. Kastner, S. Apel, C. Parnin, A. Bethmann, T. Leich, G. Saake, A. Brechmann, "Understanding Understanding Source Code with Functional Magnetic Resonance Imaging," In Proceedings of the 36th International Conference on Software Engineering (ICSE2014), pp.378–389, 2014.
4. T. Nakagawa, Y. Kamei, H. Uwano, A. Monden, K. Matsumoto, "Quantifying Program Understanding Activities Based on Cerebral Blood Flow Measurement," In proceedings of 20th Workshop on Foundation of Software Engineering (FOSE2013), pp.191–196, 2013 (in Japanese.)
5. T. Nakagawa, Y. Kamei, H. Uwano, A. Monden, K. Matsumoto, D. M. German, "Quantifying Programmers' Mental Workload during Program Comprehension Based on Cerebral Blood Flow Measurement: A Controlled Experiment," In Proceedings of the 36th International Conference on Software Engineering (ICSE2014), pp.448–451, 2014.
6. M. Ferrari, V. Quaresima, "A Brief Review on the History of Human Functional Near-Infrared Spectroscopy (fNIRS) Development and Fields of Application," Neuroimage, Vol.63, No.2, pp.921–35, 2012.
7. S. Coyle, T. Ward, C. Markham, G. McDarby, "On the Suitability of Near-Infrared (NIR) Systems for Next-Generation Brain-Computer Interfaces", Physiological Measurement, Vol.25, pp.815–822, 2004.
8. G. Strangman, J. P. Culver, J. H. Thompson, D. A. Boas, "A Quantitative Comparison of Simultaneous BOLD fMRI and NIRS Recordings during Functional Brain Activation," Neuroimage, Vol.17, No.2, pp.719–731, 2002.
9. S. Coyle, T. Ward, C. Markham, "Physiological Noise in Near-infrared Spectroscopy: Implications for Optical Brain Computer Interfacing," In Proceedings of the 26th Annual International Conference on Engineering in Medicine and Biology, pp.4540–4543, 2004.
10. R. Mitsuya, H. Morikawa, S. Kurihara, T. Kawai, Y. Tamura, "Removal of Organism Noise in Near-Infrared Spectroscopy Signal by Time Series Analysis," The transaction of Human Interface Society, Vol.11, No.2, pp163–172, 2009 (in Japanese.)
11. H. Tsunashima, K. Yanagisawa, M. Iwadate, "Measurement of Brain Function Using Near-Infrared Spectroscopy (NIRS)," Neuroimaging-Methods, 2012.
12. Y. Hoshi, N. Kobayashi, M. Tamura, "Interpretation of Near-Infrared Spectroscopy Signals: A Study with A Newly Developed Perfused Rat Brain Model," Journal of Applied Physiolosy, Vol.90, No.5, pp.1657–1662, 2001.