Hemodynamic Response Latency Analysis Using Wavelet Transform in Event-related Functional MRI

Syoji Kobashi*, Yuri T. Kitamura†, Toshio Yanagida‡,§, and Yutaka Hata¶,∥

1 Graduate School of Medicine, Osaka University, Osaka, Japan
2 Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan
3 WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan
4 Graduate School of Engineering, University of Hyogo, Hyogo, Japan

Received Date (3 March 2008); accepted Date (15 April 2008)

Abstract
Many studies showed that the hemodynamic response (HR) to brief neural activity caused by the blood oxygen level-dependent effect is delayed some seconds, and that the HR latency (= time to peak from neural activation) varies among activation sites. This paper proposes a novel method for estimating HR latency by analyzing event-related functional magnetic resonance images based on a continuous wavelet transform. The proposed method can simultaneously detect activation areas and estimate the HR latency. We also studied variability of the HR latency across subjects within the same activation sites, across acquisition days within a subject, and among activation sites within a subject. The experiments were done on 10 healthy subjects with hand-gripping tasks. We found no significant difference of HR latencies at a given activation site within a subject across acquisition days. And, we found significant differences of HR latency between the supplementary motor area (SMA) and the primary motor cortex (M1) in individual subjects, and among subjects.

Keywords
Event-related fMRI, Hemodynamic Response Latency, Wavelet Transform.

1. INTRODUCTION
Most functional magnetic resonance imaging (fMRI) studies are based on the blood oxygen level-dependent (BOLD) effect [1], which is related to temporal changes of the oxygen content in venules and so is only indirectly linked to neuronal activity via metabolic processes. Because the BOLD effect is rather small (signal change is usually less than 5 %) and data acquired from fMRI are noisy, fMRI data analysis tends to focus on the problem of detecting activation areas, which are sites of brain activation in relation to a given stimulus. In addition, there is a growing awareness that

* Corresponding author information: Syoji Kobashi; Graduate School of Engineering; University of Hyogo; Hyogo; 671-2280; Japan, and WPI Immunology Frontier Research Center; Osaka University; Osaka, Japan; +81 79-267-4989, and kobashi@eng.u-hyogo.ac.jp
determining the kind of activity that occurs is needed to understand the hemodynamics of the brain. The kind of activity can be characterized by the hemodynamic response (HR) of the BOLD effect [2]. The HR latency, which is the time to the peak of the HR from neural activation, does not directly correspond to the timing of the underlying neural activation [3][4]. Rather, it reflects the temporal properties of both the neural events and the relatively sluggish hemodynamic components of the BOLD effect. The HR latency depends on among activation sites, subjects, experimental conditions [4]. Thus, the study of HR latency is important for determining the limitations of fMRI in resolving the chronology of brain activation [5].

A number of methods for detecting activation areas using fMRI data have been proposed. They can be divided into hypothesis-driven analyses and data-driven analyses. Hypothesis-driven analyses such as the t-test [6], correlation analysis [7], z-score [8], etc. require assumptions about the paradigm or the HR function (HRF). One of the hypothesis-driven methods is statistical parametric mapping (SPM) [9], which is based on a statistical test using a general linear model. It assumes that all fMRI voxels have the same HRF, that is, their shapes and latencies are temporally and spatially invariable. However, examinations of fMRI data have revealed variations in HR latency of several seconds [10]. Thus, for complex activation paradigms and brain responses, which are not directly locked to the paradigm, hypothesis-driven analyses are not adequate. Data-driven analyses such as fuzzy cluster analysis [11], principal component analysis [12], independent component analysis [10][13]. Although a few studies have examined the HR latency, their methods have several difficulties, such as the cluster validity problem, which is a problem how to determine the number of clusters.

Wavelet analyses has been applied to fMRI data [14][15]. They apply discrete wavelet transform (DWT) to the fMRI data to decompose the time-course signal into a number of frequency domains. The effect of the wavelet transform is to remove high and low frequencies from the time-course signal to reduce the amount of noise.

Several methods to estimate HRF parameters including HR latencies have been developed. Aguirre et al. developed a method for estimating HRF [16], and investigated its variability. Because the method cannot simultaneously detect activation areas that have different HRFs, it can only derive the HRF within the detected activation areas. Lange and Zeger showed a method for estimating the latency and dispersion parameters by non-linear regression [17]. This is based on time-consuming iterative methods, such as the Gauss-Newton method, which can be unstable because the iteration does not always converge. Purdon et al. also used non-linear regression for a model of HRF based on physiological considerations [18]. Rajapakse et al. developed an intriguing non-iterative method that fits a Gaussian function into the HRF [19]. This method relies on a special property of the Fourier transform of Gaussian functions: the phase is linearly related to the latency, and the log modulus is linearly related to the squared dispersion. Saad et al. proposed a method using the Hilbert transform [20]. They obtained a high-resolution latency estimate by finding the zero-crossing of the cross correlation function. Another method is based on the phase of the response to a periodic stimulus [21], but this does not work for non-periodic stimuli.

This paper introduces a novel analysis, called hemodynamic response analysis using a wavelet transform (HAW, for short), for simultaneously detecting activation areas and estimating HR latencies. Using the continuous wavelet transform (CWT), our method defines the time point of the local maximum wavelet coefficient as the HR latency, and the wavelet coefficient as the degree of activation. The detection accuracy of activation areas was evaluated by comparing with SPM99. Using HAW, we also examine the variability of HR latencies across two activation sites: the supplementary motor area (SMA; Brodmann area (BA) 6) and the primary motor cortex (M1; BA4). The variability was evaluated within a
subject, across subjects within the same activation site, and across data acquisition days within a subject.

2. SUBJECTS AND MATERIALS
Ten healthy male subjects (ages 22.9 ± 2.5 years, mean ± SD) were recruited. Nine were right-handed and one was left-handed. The handedness determined by Edinburgh inventory [22] was 82.98 ± 14.50 for the nine right-handed subjects and –84.6 for the one left-handed subject. Informed consent was obtained in every case.

MRI studies were performed on a 1.5 Tesla scanner (GE Medical Systems, Milwaukee, WI). Echo planar imaging was used to acquire data sensitive to the BOLD signal at a repetition time (TR) of 2000 ms and an echo time of 40 ms. The spatial resolution was set by a 64 by 64 voxel matrix covering 260 × 260 mm² with 5 mm slice thickness. Twenty axial slices were acquired to cover the whole brain. During the data acquisition, 258 images (phases) per slice were obtained in 516 s.

3. METHODS
Experimental Tasks for ER-fMRI
Generally, the minimum TR of EPI is longer than 1.0 s. If we used a TR of 1.0 s, the HR would be sampled with a resolution of 1.0 s. This is not enough to properly sample a time-course signal because the bandwidth of the HR is rather narrow in comparison with the sampling time. Therefore, certain components of the HR, such as the peak, would be missed. This may lead to errors in the estimation of HR latencies because we aim to find the peak of the HR. Conventionally, to improve the resolution of the time-course signal, the acquired signals are interpolated computationally by post-processing.

The temporal behaviour of HRF for repeated trials within the same session is almost stable [16]. In contrast, some studies have reported inter-trial variations with brief stimulation [23]. However, these studies have also shown that the variability of the HR latency is rather small in comparison with other parameters of HRF, such as onset-time and over-shooting time. Consequently, we can assume there are few inter-trial variations of HR latencies.

Under the above assumption, we can improve the apparent time resolution by iterating trials in which there is a short period of lag following the starting stimulus. The period of lag can vary from 0.0 s up to TR depending on the desired time resolution. For example, if TR is 2.0 s, assume that, at the first trial, the MR signal is sampled 0.5 s before the peak as shown in Figure 1(a), i.e., the highest point of the peak is not sampled at the first trial. At the ith trial, by giving the stimulus with a lag of 0.5 s, MR signal is sampled at the top of the peak as shown in Figure 1(b). The trial is iterated, with the lag period changing on each of iteration, so that the HR is properly sampled.

In this study, the subjects performed a given task for duration of 2.0 s. The inter-stimulus interval (ISI) was randomized (i.e., jittered) between 20 s and 30 s by a period of TR (= 2.0 s). Twenty seconds was selected as the minimum ISI because the signal change of the HR returns to the baseline after about 16 s [24]. If the ISI is constant, the subjects will be able to anticipate the occurrence of the next trial, which might produce increases in neural activity prior to the onset of the stimulus [25].

![Figure 1. Acquisition technique for ER-fMRI. (a) Time-course signal at 1st trial. (b) Time-course signal at ith trial. (c) Sorted time-course signal.](image-url)
This could make it difficult to analyze fMRI data. Because we chose a time resolution of 0.1 s for sampling the proper HR, the trial was repeated 20 times (= TR / desired time resolution) with a random ISI and a lag for each experiment.

The motor cortex activation experiment was conducted using a hand gripping motor task. In the MRI scanner, subjects viewed a backlight projection screen from within the magnet bore through a mirror mounted on the head coil. Normally the image on the screen was a black cross in the center of a white background. The stimulus consisted of changing the cross to a black circle for duration of 2.0 s. The subjects had been instructed to make rapidly opening and closing of fist while the black circle was visible. Figure 2(a) shows the time-course of a typical signal at an activation site in the brain.

Sorting the Time-Course Signals
To increase the efficiency of the analysis, only signals in the brain are analyzed. Voxels with signal intensity lower than a threshold are regarded as background. Motion correction was performed with the automated image registration algorithm [26]. Because SPM requires spatial filtering to obtain adequate statistical parametric maps, we applied a smoothing process using a Gaussian spatial filter in order to compare the results from our method to those obtained by SPM.

Signals for the separate trials are superimposed to make one high-resolution signal. For example, consider the time-course signals at the 1st and \( i \)th trials shown in Figure 1(a) and (b), respectively. The signals at the 1st trial are superimposed on the time-course signal of a single trial with no delay (black dots in Figure 1(c)). Then, the signals at the \( i \)th trial are superimposed with a delay of 0.5 s (gray dots). By superimposing the obtained time-course signals of the separate trials into one time-course signal of one trial, a time-course signal at an interval (= 0.1 s) that is shorter than TR (= 2.0 s) is obtained as shown in Figure 1(c). Therefore, the present method needs no interpolation process during post-processing to estimate the HR latency at high time resolution. Also, we can improve the apparent time resolution by fine-tuning the time lag.

In this study, because the trials were shifted by 0.1 s, the apparent time resolution was 0.1 s by sorting the time-course signals of the separated trials. For example, Figure 2(b) shows the sorted time-course signal. The reconstruction technique will introduce high frequency components when the baseline of the raw time-course signal drifts, although, such high frequency components will be eliminated by applying CWT because CWT acts as a band-pass filter. Finally, we calculate the signal change from the baseline by dividing the time-course signal by its average.

HAW -HR Analysis Using the Wavelet Transform
CWT approximates a time-course signal by convoluting the scaled and shifted mother wavelet into the time-course signal (e.g., [27]). CWT for a given input signal \( s_{xyz}(t) \) at a voxel \((x,y,z)\) is given by:

\[
W(a,b) = \sqrt{a} \int_{-\infty}^{\infty} \Psi\left(\frac{t-b}{a}\right) s_{xyz}(t) dt ,
\]

where \( a \) and \( b \) represent the scaling factor and the time factor, respectively, and \( \Psi(\cdot) \) is the mother wavelet function. The mother wavelet is controlled by the scale factor and the time factor.
We apply this scheme to find HRF from the sorted time-course signal by regarding HRF as the mother wavelet. By changing the scale and time factors of the mother wavelet, we can search for the HRF from the time-course signal for each voxel. Because the wavelet coefficient is the degree of fitness to the signal, we can assume that the wavelet coefficients correspond to the degree of activation, and that the shift and the scale parameters correspond to the HRF latency and the form parameter of the HRF, respectively. Consequently, we define the local maxima of wavelet coefficients in the scale and shift map to be the degrees of activation, and define the shift parameters on the local maxima to be the HR latencies. The mother wavelet used in this study is the Mexican hat (Figure 3(a)):

$$
\Psi(t) = \left(1 - t^2\right) \exp\left(-\frac{t^2}{2}\right),
$$

(2)

because this function is similar in form to the HRF [4]. Of course, we can use another mother wavelet function such as the Gabor or French hat. These functions yielded similar results in our experiments. For example, Figure 3(b) shows wavelet coefficients obtained by applying CWT to the time-course signal given by Figure 2(a). In this figure, \(W(a_{\text{max}}, b_{\text{max}})\) shows the local maximum wavelet coefficient where \(a_{\text{max}}\) and \(b_{\text{max}}\) are the scale and the shift parameters, respectively. According to this definition, in this case, the degree of activation at the voxel of interest is the wavelet coefficient \(W(a_{\text{max}}, b_{\text{max}})\), and the HR latency is the shift parameter \(b_{\text{max}}\).

Next, we discuss on the meaning of the degree of activation given by using CWT. To do this, we give an example of computing the signal change from a wavelet coefficient. Here, we assume HRF to be a Gaussian function, which is formulated by \(\exp\left(-\frac{x^2}{2}\right)\). CWT is applied to a Gaussian function with a height of 1.0 and shift parameter \(b\) of 0:

$$
W(a, b = 0) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \Psi(t) \exp\left(-\frac{x^2}{2}\right) dt.
$$

(3)

Then the maximum wavelet coefficient is:

$$
\max_a W(a) = \frac{5 \sqrt{2}}{6} \sqrt{\pi} \approx 1.27519.
$$

(4)

Therefore, when we assume the HRF to be a Gaussian function, the signal change of HR, \(\alpha\), can be calculated by:

$$
\alpha = \frac{W(a_{\text{max}}, b_{\text{max}})}{1.27519}.
$$

(5)

Needless to say, we can assume another function to emulate the HRF. In that case, the denominator in Eq. 5, which is the maximum wavelet coefficient, would need to be recalculated in a similar way.

Consequently, by applying the CWT to the time-course signal for each voxel, we can obtain the change in the HR signal from the baseline and the latency time after starting the stimulus. The activation area, which is the site of brain activation in relation to a given stimulus, is derived by thresholding the signal change. Then, a clump of the connecting activation areas is defined as a cluster.

### 4. EXPERIMENTAL RESULTS AND DISCUSSION

#### Detecting ability of activation area

ER-fMRI data acquired from ten subjects that were given right-hand and left-hand gripping tasks were analyzed by HAW and SPM99 for comparison. The average hand-gripping fre-
Figure 4. Comparison of the detected activation areas between HAW and SPM. (a) and (b) Detection results for right-hand and left-hand gripping tasks, respectively, at the same section for subject 1. Upper; signal change map by HAW (red-white shows 0.0% to 3.0% whose wavelet coefficient of 0.00 to 3.83). Middle; the thresholded signal change map. Lower; t-value map by SPM99. (c) Activation areas imposed on 3D rendering of the brain (right-hand gripping task) by HAW (left) and SPM99 (right).

Frequency in the MRI scanner was $4.00 \pm 1.08$ grips. For each ER-fMRI data, Gaussian smoothing with FWHM (full width at half maximum) of 8.12 mm. was applied. The computation time for HAW (not including preprocessing by SPM99) was shorter than 5 minutes running on a personal computer (1GHz Pentium III CPU, 256Mb of memory).

Estimates of the $t$-value map obtained by SPM99, and signal change map obtained by HAW for subject 1 (a right-handed subject) are shown in Figure 4. For HAW, the threshold of a signal change was set at 1.0% (= wavelet coefficient of 1.28) and for SPM99, the threshold of the maximal $p$-value was set at 0.005 (uncorrected). The signal change and $p$-value thresholds were determined experimentally. Both methods successfully detected activation in M1, which is in the contralateral hemisphere in response to the gripping hand, and in SMA. These results show that HAW can detect activation areas that are equivalent to those estimated by SPM99. For other subjects, HAW detected activation in M1, and in SMA that was almost the same as the activation detected by SPM99.

Figure 5. HR delay map of activation areas for Subject 1 with right-hand (upper) and left-hand gripping (lower) tasks. Left-right; $z = 16, 15, 14$. $z$ shows the slice number of coronal images in which the smaller value means the inferior coronal section. Note that $z$ does not correspond to the Talairach coordinate.

**HR latencies and its variability**

We define an HR latency map in which the value at each voxel in the activation areas is the HR latency for the voxel, and the value of the other voxels is zero. For subject 1, HR latency maps were obtained as shown in Figure 5. In the case of subject 1, the HR latency at SMA was shorter than the HR latencies at M1. Figure 6 shows the change in activation areas in
the hemisphere over an approximately 8 s periods. These results led to the hypothesis that HR latencies vary among activation sites. To test this hypothesis, we investigated the variability of HR latencies at SMA and M1 (1) among trials within the same subject, and (2) among subjects.

Variability of HR latencies among acquisition days within the same subject was investigated by acquiring four ER-fMRI datasets on four days between Jun. 5 and Nov. 16, 2001 for subject 1. The HR latencies distributions at SMA, M1 and the primary visual area (V1; BA17) are shown in Figure 7. In this experiment, V1 was also activated because the stimulus was given through a backlit projection screen. At each activation site, we found no significant difference across the data acquisition days using one-way analysis of variance (ANOVA) at a significance level ($p \leq 0.01$). Figure 7(c) shows the total distribution. The mean HR latencies (± SD) were 4.8 ± 0.5 s at SMA, 5.4 ± 0.5 s at M1 and 8.4 ± 0.7 s at the V1. The HR latencies among SMA, M1 and V1 were significantly different ($p \leq 0.01$). In addition, in all cases, the HR latency was in the order V1 > M1 > SMA. Miezin et al. also showed that the HR latency was greater at the V1 than at SMA [28]. However, magnetoencephalography (MEG) studies, cognitive studies and other studies have reported that the hand-gripping task with a checkerboard reversal stimulus would activate the V1, SMA and M1 in that order. The differences in the order between our results and the previous studies appear to be because the HR latency of the BOLD signal does not reflect the timing of the underlying neural activation [3].

Comparison of the HR latencies among 5 subjects whose ER-fMRI data were obtained a few times on different acquisition days is shown in Figure 8. This figure confirmed that at each activation site the inter-subject variance (i.e., the differences among subjects) was significantly larger than the intra-subject variance (temporal differences in one subject) by using ANOVA ($p \leq 0.01$). This shows that HR latencies are stable within the same subject in comparison with the variance among subjects. These results across days within a subject and across subjects agreed with the findings of Aguirre et al. that the HRF varies substantially

**Figure 6.** Time transition of activation areas (Subject 8, right-hand gripping task).

**Figure 7.** HR latency distribution at SMA, M1 and V1 in the same subject (Subject 1, right-hand gripping). (a) 1st exam (2001/6/5). (b) 2nd exam (2001/10/10). (c) 3rd exam (2001/11/8). (d) 4th exam (2001/11/16). (e) Total of (a)-(d).
among different people, but is relatively stable for a given person [16].

We next investigated the differences of the HR latencies between M1 and SMA across in 10 subjects (Figure 9). In this figure, \( t \)-value is HR latency at M1 minus the HR latency at SMA. Thus, a positive \( t \)-value means that the HR latency at M1 is greater than that at SMA. For all of the experiments except the experiment of the right-hand gripping task by subject 10, the HR latency at SMA cluster was significantly different from that of M1 cluster (\( p \leq 0.05 \)). Moreover, for all experiments except the experiment with subject 5, the HR latency at SMA cluster was shorter than that at M1 cluster. Only for subject 5, who is left-handed, was the HR latency at SMA cluster longer than the HR latency at M1 cluster.

5. CONCLUSION

We have proposed a novel method for detecting activation areas and simultaneously measuring HR latencies using ER-fMRI. The method defines the wavelet coefficient as the degree of activation, and to define the time point with the local maxima of the wavelet coefficients as the HR latencies. The novel feature is that it finds the shape and the time delay of HRF by generating various types of HRF and fitting them into the time-course data, and this methodology is easily implemented using CWT scheme. The detection accuracy of the proposed method was evaluated by comparing it with SPM99. The evaluation was done using ten subjects where the given tasks were left-hand and right-hand gripping tasks. In all experiments, HAW and SPM99 detected activation in SMA and M1 as activation areas, although the locations of activation detected by the two methods slightly differed.

We then investigated the variability of HR latencies using HAW. With respect to the variability of the HR latencies obtained by using HAW, we found:

1. HR latencies are stable at a given activation site in a given subject over a period of a few days (ANOVA; \( p \leq 0.01 \)).
2. HR latencies vary among subjects at the same activation site. HR latencies are also stable among acquisition days within a subject in comparison with the variance among subjects (ANOVA; \( p \leq 0.01 \)).
3. HR latency at SMA differs from that at M1 within a subject (T-test; \( p \leq 0.05 \)).

Because the first and third findings are in agreement with the findings of Aguirre et al. [16], we conclude that HAW can successfully estimate HR latencies from ER-fMRI data.
This makes it easier to observe differences of HR latencies among activation sites or among subjects because the method automatically detects adequate activation areas whose HRFs are unknown.

Based on our findings concerning the variability of the HR latencies, we hypothesize that subjects have an inherent HRF at each activation site but that the type and the load level of the stimulus does not contribute to the HRF. This is probably because the form of HRF is strongly affected by the structure of the capillary bed, which is different in different subjects [29], especially in subjects with cerebrovascular diseases [30]. This raises the possibility that ER-fMRI using HAW can be used to diagnose cerebrovascular diseases related to the capillary bed by investigating the change of the HR latency over time.

The present method has some limitations. One limitation is that CWT filters out the high frequency components of the BOLD signal, which are important for understanding the characteristics of the HRF. Another limitation is that inter-trial variations cannot be observed. Third limitation is that HAW requires the assumption that HRF is a Gaussian function in order to compute a value of signal change from the resultant wavelet coefficient. To avoid this assumption, in the future, we need to develop a methodology in which the resultant wavelet coefficients are significantly different between activated voxels and unactivated voxels. Then we need to apply another stimulus that activates SMA, M1, V1 and other areas to investigate the variability of HR latencies among different stimuli. This will test our hypothesis that there is an inherent HR at each activation site. That is, it will test our hypothesis that the type of stimulus does not contribute to the shape and the latency of HRF. In addition, making HR latency maps of the whole brain by applying different stimuli that activate different sites should provide new insights into the functioning of the brain.

ACKNOWLEDGEMENTS
This work was supported in part by KAKENHI (18613007, 19209022), a grant from Ishikawa Hospital Grant, and the Berkeley Initiative in Soft Computing (BISC) Program of the University of California at Berkeley.

REFERENCES


AUTHOR INFORMATION

Syoji Kobashi is an associate professor in Graduate School of Engineering, University of Hyogo, Japan. His research interests include soft computing approach to medical signal/image processing and human brain functions. He received the Joseph F. Engelberger Best Paper Award at the 2nd World Automation Congress in 2000, the IEEE EMBS Japan Young Investigators Competition from IEEE EMBS Japan Chapter in 2003.

Yuri T. Kitamura is the associate professor of Social and Environmental Medicine, Osaka University Graduate School of Medicine, Japan. She’s got the license of medical practitioner since 1990 and become a specialist for Anesthesiology and Psychiatry. She belongs to the following academic societies, Japan Society of Psychiatry and Neurology, Japan Neuroscience Society, Japan Human Brain Mapping Society, etc. Recently, main research field is the Neuroimaging with fMRI, PET, and NIRS, targeting on Psychiatric disorders in children such as Autism, ADHD, and Asperger’s syndrome.

Toshio Yanagida received the B.Eng. Sci. degree in 1969, the M.Eng. Sci. degree in 1971, and the D.E. degree in 1976, all from Osaka University, Osaka, Japan. He is currently a Professor at Osaka University’s Department of Physiology and Biosignaling in the Graduate School of Medicine and Laboratories for Nanobiology in the Graduate School of Frontier Biosciences. He is currently the research supervisor of Novel Measuring and Analytical Technology Contributions to the Elucidation and Application of Life Phenomena, CREST, JST and Vice-director of WPI Immunology Frontier Research Center, Osaka University. He received the 25th Naito Memorial Science Promotion Award in 1994, a Japan Academy Award, the Imperial Award in 1998, and an Asahi Award in 1999.
Yutaka Hata is a professor in Graduate School of Engineering, University of Hyogo, Japan. He spent one year in BISC Group, University of California at Berkeley from 1995 to 1996 as a visiting scholar. He is the Founding Editor-in-Chief of the International Journal of Intelligent Computing in Medical Sciences and Image Processing, and the regional editor of Intelligent Automation & Soft Computing. He received Joseph F. Engelberger Best Paper Award and Best Paper Award at the WAC2000, USA, WAC Contribution Award at 2002, 2004, 2006, and Life Time Achievement Award at 2008.