ELSEVIER

Clinical Neurophysiology 115 (2004) 2181–2192

www.elsevier.com/locate/clinph

# Removal of time-varying gradient artifacts from EEG data acquired during continuous fMRI

Michiro Negishi[a,*], Mark Abildgaard[a], Terry Nixon[a], Robert Todd Constable[a,b,c]

[a]*Department of Diagnostic Radiology, Yale University School of Medicine, P.O. Box 208043, TAC Building MRRC Rm. N128, New Haven, CT 06520-8043, USA*

[b]*Department of Biomedical Engineering, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA*

[c]*Department of Neurosurgery, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA*

Accepted 8 April 2004

Available online 18 May 2004

**Abstract**

**Objective**: Recording low amplitude electroencephalography (EEG) signals in the face of large gradient artifacts generated by changing functional magnetic resonance imaging (fMRI) magnetic fields continues to be a challenge. We present a new method of removing gradient artifacts with time-varying waveforms, and evaluate it in continuous (non-interleaved) simultaneous EEG–fMRI experiments.

**Methods**: The current method consists of an analog filter, an EEG–fMRI timing error correction algorithm, and a temporal principal component analysis based gradient noise removal algorithm. We conducted a phantom experiment and a visual oddball experiment to evaluate the method.

**Results**: The results from the phantom experiment showed that the current method reduced the number of averaged samples required to obtain high correlation between injected and recovered signals, compared to a conventional average waveform subtraction method with adaptive noise canceling. For the oddball experiment, the results obtained from the two methods were very similar, except that the current method resulted in a higher P300 amplitude when the number of averaged trials was small.

**Conclusions**: The current method enabled us to obtain high quality EEGs in continuous simultaneous EEG–fMRI experiments.

**Significance**: Continuous simultaneous EEG–fMRI acquisition enables efficient use of data acquisition time and better monitoring of rare EEG events.



*Keywords:* Electroencephalography; Functional magnetic resonance imaging; Event-related potentials; P300; Artifacts

## 1. Introduction

Simultaneous recording of electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) has the potential to monitor events in the brain with higher spatial and temporal resolutions than standalone EEG or fMRI measurements. However, EEG data recorded during fMRI acquisitions are contaminated with large imaging artifacts, including gradient artifacts induced by the changing magnetic field gradients used for spatial encoding in MRI. These gradient artifacts can be avoided by EEG-triggered fMRI acquisition (Krakow et al., 1999; Seeck et al., 1998; Warach et al., 1996) or by interleaved or sparse EEG–fMRI acquisition (Baudewig et al., 2001; Goldman et al., 2000; Sommer et al., 2003), but EEG data is lost or degraded during fMRI scanning. Such loss of data can be an issue in some applications of simultaneous EEG–fMRI acquisition, including recording of spikes from epileptic patients (Lemieux et al., 2001). Simultaneous and continuous (non-interleaved) EEG–fMRI recording, on the other hand, requires gradient artifact reduction algorithms such as frequency domain processing (Garreffa et al., 2003; Hoffmann et al., 2000), average waveform subtraction (Allen et al., 2000), and spatial filtering (Bonmassar et al., 1999).

Although average waveform subtraction (Allen et al., 2000) is widely used to remove the gradient artifacts (e.g. Salek-Haddadi et al., 2002), the waveform of gradient artifact changes over time, thereby blurring the average waveform and decreasing the efficacy of this approach. Changes in the sampling of the gradient artifact waveform may be caused by (a) timing errors between fMRI scanning and EEG sampling (Cohen et al., 2001), (b) changes in electrode position and orientation over time,

* Corresponding author. Tel.: +1-203-737-5995; fax: +1-203-785-6534.
*E-mail address:* michiro.negishi@yale.edu (M. Negishi).

1388-2457/$30.00 
doi:10.1016/j.clinph.2004.04.005

and (c) mechanical vibrations that are caused by, but are not perfectly phase locked with, the switching gradient fields. This paper focuses on a methodology to remove these time-varying gradient artifacts.

One way to solve problem (a) above is to use a single clock to control EEG and fMRI acquisitions using a customized hardware (Anami et al., 2003). Another way is to record the fMRI trigger with a high temporal resolution and make the appropriate timing adjustment during average waveform subtraction (Allen et al., 2000; Cohen et al., 2001). In this paper we demonstrate an alternative approach, in which the EEG is recorded with a relatively low sampling rate after low-pass analog filtering, followed by timing error detection and correction. This approach does not require special hardware to synchronize EEG and fMRI or to record the fMRI trigger with a high temporal resolution. However, because timing errors are computed, there are possibilities for errors in the estimated timing errors. Therefore we do not claim that our solution to (a) yields a better result than using the exact fMRI trigger timing.

Adaptive noise canceling (ANC) (Allen et al., 2000) and spatial filtering (Bonmassar et al., 1999) have the ability to reduce the problem caused by (b) and (c), although spatial filtering does not reduce noises whose spatial patterns of influence on the EEG electrodes fluctuate over time. For instance, independent vibration of leads on an EEG cap may give rise to such fluctuations. We present a temporal principal component analysis (PCA) based gradient noise reduction algorithm (temporal PCA filter hereafter) that addresses these problems, and compares its gradient noise reduction ability with ANC combined with average waveform subtraction.

There have been some event-related potential (ERP) studies combining separate EEG and fMRI runs (Horovitz et al., 2002; Opitz et al., 1999) or using interleaved simultaneous EEG–fMRI (Kruggel et al., 2000; Liebenthal et al., 2003; Sommer et al., 2003). Continuous simultaneous EEG–fMRI has been successfully applied to detect relatively high amplitude EEG signals such as ictal spikes (Salek-Haddadi et al., 2002) and alpha waves (Anami et al., 2003). However, there have been very few reported ERP studies using continuous simultaneous EEG–fMRI (Bonmassar et al., 1999), indicating the difficulty of recording low amplitude EEG during fMRI. In this paper we report a continuous simultaneous EEG–fMRI recording of a P300 ERP, and compare the temporal PCA filter with average waveform subtraction followed by ANC (Allen et al., 2000) in their abilities to recover a P300 ERP.

## 2. Methods

### 2.1. Timing error correction

Because gradient artifacts contain frequency components higher than typical EEG sampling frequencies, slight de-synchronization between fMRI scanning and EEG sampling can result in a large change in the gradient artifact waveforms (Cohen et al., 2001). Hence, the first stage of the current gradient noise reduction method is analog low-pass filtering before sampling, followed by timing error detection and correction. The analog filter enables reliable timing error detection. In the timing error detection algorithm, sampled data is first interpolated using sync-interpolation and segmented into fMRI slice acquisition intervals. In this paper, each interval of data from each channel is called a frame. Acquisition of $T$ slices (number of head images × number of slices per head image) results in $T$ times $M$ frames where $M$ is the number of EEG channels. In addition to frames from all channels, $T$ summary frames are computed by taking the average of frames of all channels, after correcting for polarity of gradient artifact so that artifact waveforms of all channels have positive correlations with that of the first channel. If there is only one channel, the frames from that single channel are used as summary frames. Delay time of each summary frame with respect to the average of all summary frames is then computed. Each non-summary frame is shifted to correct for the delay of a summary frame corresponding to its slice and image, and down-sampled to the original sampling rate.

### 2.2. Gradient artifact removal

The second part of the gradient artifact removal method is a temporal PCA filter, which operates on each channel independently. First, principal components (PCs) are computed from the timing-corrected EEG signals. PCs capture component waveforms whose amplitude variation accounts for the largest variations of gradient waveforms. The basic idea behind using PCA is that if the gradient artifact waveform consists of multiple components whose amplitudes do not co-vary with each other, they would be captured in different PCs.

Second, activations of PCs (or PC scores) are computed, along with the moving average and the associated standard deviation of PC activation. Also, the mean and the standard deviation of PC activations are computed from standalone (in-magnet but no-scanning) EEG data.

Third, PC activations associated with gradient artifacts are estimated using the PC activations and statistics computed in the previous step. An estimated gradient artifact at each frame is then computed as a sum of PCs weighted by estimated PC activations for each frame, and is subtracted from the original frame. Finally, an 80 Hz low-pass filtering is applied.

If it is assumed that the number of PCs is large enough to account for most of the variance of the signals, then using the average PC activation regardless of the PC activation at each frame is similar to average waveform subtraction. On the other hand, using current activation at each frame, allowing unlimited deviation from the mean, results in the removal of all signals. The current algorithm estimates the PC activation corresponding to the gradient artifact in

a statistically optimal way, allowing some deviations from the mean (Eqs. (2) and (3)).

### 2.3. Key equations

A frame and a summary frame are denoted by a vector $\mathbf{x}_{ij}$ and a vector $\mathbf{f}_j$ ($i$, channel number; $j$, serial slice number starting from the first slice of the first image), respectively. Each element of these vectors represents an instantaneous voltage and is denoted by $\mathbf{x}_{ij}(k)$ (or $\mathbf{f}_j(k)$) where $k$ is a digitized sample number within a frame.

Delay time of slice $j$ with respect to the average of all slices is computed by

$$\Delta T_j \approx \frac{\Delta t}{N} \sum_{k=1}^{N} \frac{\mathbf{f}_j(k) - \mathbf{f}^{\mathrm{av}}(k)}{\mathbf{f}'^{\mathrm{av}}(k)} \tag{1}$$

where $\Delta$t is the sampling interval, $N$ is the number of digitized samples in a frame, $\mathbf{f}^{\mathrm{av}}$ is the average vector of the summary frames of all slices, and $\mathbf{f}'^{\mathrm{av}}$ is the first-order difference vector of $\mathbf{f}^{\mathrm{av}}$, computed as $\mathbf{f}'^{\mathrm{av}}(1) = \mathbf{f}^{\mathrm{av}}(2) - \mathbf{f}^{\mathrm{av}}(1)$, $\mathbf{f}'^{\mathrm{av}}(k) = (\mathbf{f}^{\mathrm{av}}(k+1) - \mathbf{f}^{\mathrm{av}}(k-1))/2$ for $k = 2, 3, 4, \ldots, N-1$, and $\mathbf{f}'^{\mathrm{av}}(N) = \mathbf{f}^{\mathrm{av}}(N) - \mathbf{f}^{\mathrm{av}}(N-1)$. Then each frame $\mathbf{x}_{ij}$ is shifted to correct for the delay $\Delta T_j$ and down-sampled to the original sampling rate. The resultant frame is denoted by $\mathbf{y}_{ij}$ below.

It can be shown that an optimal subtraction of PCs can be written as follows, assuming that the activation of PCs found in the gradient noise and in the interested EEG signal follow independent normal distributions

$$\mathbf{z}_{ij} = \mathbf{y}_{ij} - \sum_{k=1}^{N^{\mathrm{c}}} [(1 - \alpha_{ijk})(\mathbf{y}_{ij} \cdot \mathbf{c}_{ik}) + \alpha_{ijk}\mu_{ijk} - \mu^{\mathrm{s}}_{ik}]\boldsymbol{c}_{ik} \tag{2}$$

In this formula, $\mathbf{z}_{ij}$ is a frame $j$ of channel $i$ after gradient noise removal; $N_{\mathrm{c}}$ is number of PCs; $\boldsymbol{c}_{ik}$ is a $k$-th PC; $\mu_{ijk}$ is the moving average of the PC activation $\mathbf{y}_{ij} \cdot \boldsymbol{c}_{ik}$ of the EEG signals during fMRI scanning; $\mu^{\mathrm{s}}_{ik}$ is the average PC activation $\mathbf{y}^{\mathrm{s}}_{ij} \cdot \mathrm{c}_{ik}$ (where $\mathbf{y}^{\mathrm{s}}_{ij}$ is a frame segmented from standalone signal) of the whole standalone EEG signal, and $\alpha_{ijk}$ is a weighting factor of PC activations computed as follows

$$\alpha_{ijk} = (\sigma^{\mathrm{s}}_{ik}/\sigma_{ijk})^2 \tag{3}$$

In Eq. (3), $\sigma_{ijk}$ is the standard deviation of the PC activation of the EEG signals during fMRI scanning around frame $j$ and $\sigma^{\mathrm{s}}_{ik}$ is the standard deviation of the PC activation of the whole standalone signal. As was described in words in Section 2.2, if the PCs account for most of the variances in the EEG signal, setting $\alpha_{ijk}$ to 1 in Eq. (2) results in average waveform subtraction using moving average, whereas setting $\alpha_{ijk}$ to 0 corresponds to removal of most signals. (note: when there are large movement artifacts, we find it empirically useful to limit the values of the PC activation so that they vary only from $[\mu_{ijk}, -\sigma_{ijk}]$ to $[\mu_{ijk}, +\sigma_{ijk}]$).

### 2.4. Cardiac artifact removal

Although the main purpose of this paper is to evaluate the temporal PCA filter for gradient artifact removal, we also explored the possibility that the same algorithm can be used to remove another major error in the in-magnet EEG signal, namely cardiac artifact (or pulse artifact, Allen et al., 1998). Cardiac artifacts in the EEG signals are caused by small movements associated with heartbeat and blood flow. Cardiac artifacts can become more than 50 microvolts (μV) in 3 Tesla (T) fMRI, making it difficult to recover contaminated EEG waveforms without averaging a large number of samples. We applied cardiac artifact removal to the data acquired from the alpha wave detection experiment, but not to the data from the oddball experiment where a large number of trials were averaged. If we assume that the cardiac artifact reduces with averaging as random noise does, a 50 μV noise would be reduced to 7.5 μV with averaging of 45 trials and to 2.6 μV with 384 trials. Thus, cardiac artifact may account for much of the noise that are seen in Fig. 5. Allen et al. (1998) used a variant of average waveform subtraction, allowing the waveform to lengthen or shorten according to heart rate fluctuations. In an alpha wave detection experiment, we used the temporal PCA-based noise removal algorithm to remove cardiac artifacts. For cardiac artifact removal, 'cardiac trigger pulses' (a vector with one at the onsets of heartbeats and zero elsewhere) have to be computed. It is accomplished by the following steps that operate on the cardiac EEG channel after it has been processed by the temporal PCA filtering.

(1) *Conditioning*. Apply a 5 Hz fourth-order Butterworth low-pass filter to the cardiac signal that was recovered by the temporal PCA-based gradient noise removal.
(2) *Standardization*. Subtract moving average with a 1 s time window and divide by standard deviation computed with a 4 s time window.
(3) *Thresholding*. Find sections of contiguous points where the standardized values are above 0.7.
(4) *Finding centers*. Find the centers of above-threshold sections. Remove center points whose immediate predecessors are less than 0.4 times the average interval between centers.
(5) *Computing average waveform*. Compute an average cardiac waveform segmented at the center points.
(6) *Computing match*. Compute inner product of the average cardiac waveform and a sliding window with the same length moving on the time series computed by the step (2).
(7) *Finding peaks*. Find points where the inner products are local maxima and above a quarter of the standard deviation of the inner products.

The points thus selected are used as cardiac trigger pulses to segment EEG data after gradient artifact removal into cardiac frames, which will be processed by the temporal

PCA-based (cardiac) artifact removal algorithm. Note that in both gradient artifact removal and cardiac artifact removal, frame lengths may differ from one to another. In the actual implementation, all frames are extended to the maximal frame length, allowing overlaps between adjacent frames.

### 2.5. Experiments

The current method was tested in a phantom experiment and two in vivo experiments using human subjects, namely a visual oddball experiment and an alpha wave detection experiment. For the phantom experiment and the visual oddball experiment, we built an analog single channel Chebyshev type I low-pass filter (roll off frequency = 125 Hz). We used a sampling factor of 128 for sync-interpolation for the timing error correction and a sampling window of 50 for the moving average $\mu_{ijk}$.

The experimental system for the phantom experiment consisted of an EEG signal emulator (PocketTrace, Neuroscan Inc., Sterling, VA, attenuated to 15 μV peak to peak), an EEG cap and carbon fiber cables (Neuroscan QuickCap/MagLink), an analog low-pass filter, and an EEG recorder (Neuroscan NuAmps, parameters: 1 kHz sampling, 32 bit DC recording, no low-pass filtering. After gradient noise removal, a fourth-order 80 Hz low-pass Butterworth filter was applied). Note that the built-in filters in NuAmps are digital filters and thus are not suitable for signal conditioning before timing correction. Typically, biological signals of interest are below 70 Hz. For EEG recording in the fMRI scanner, various sampling frequencies from 200 Hz (Goldman et al., 2000) to 5 kHz (Allen et al., 2000) are used. We chose the 1 kHz sampling rate to demonstrate that the combination of analog low-pass filtering and timing error correction can reliably correct timing errors in EEG signals that are collected at a moderate sampling rate. The EEG cap was placed on a spherical phantom in a 3 T MR imager (Magnetom Trio, Siemens Medical Systems, Erlangen, Germany), from which the EEG signal was collected from one channel (F7). The F7 and the ground electrodes were short-circuited at the cap to form a loop, and a 5 kohm register was inserted serially at the filter input to mimic the skin conductance. The fMRI parameters were: echo-planar blood oxygen level dependent (EPI BOLD) sequence, repetition-time (TR) = 2220 milliseconds (ms), 30 slices per image, echo time (TE) = 30 ms, flip angle (FA) = 70, 30 × 6 mm slices, field of view (FOV) = 220 mm, matrix size = 64 × 64.

For the oddball experiment, the EEG was recorded from linked CP3–CP4 electrodes against a linked ear reference, and the EEG was recorded at 500 Hz. The fMRI parameters were the same as the phantom experiment except TR = 1240 ms, TE = 30 ms. The fMRI data was analyzed using SPM99 software (Welcome Department of Cognitive Neurology, University College, London. http://www.fil.ion.ucl.ac.uk). Visual stimuli consisted of frequent stimuli 'oooooo' and rare stimuli 'xxxxxx', which were generated using the shareware software PsyScope (Department of Psychology, Carnegie Mellon University) and projected as white letters on a dark screen. Stimulus onset asynchrony was 1 s, stimulus duration was 0.25 s, and the frequency of the rare stimuli was 6.7%. Four healthy subjects (average age = 38 years, range = 25–48 years, one female) participated in the study with written consent (Yale HIC#11970).

The reasons for recording EEG from only one channel in both the phantom and the oddball experiments were: (1) the timing correction algorithm and the temporal PCA filtering both operate on each channel independently with an exception of the computation of the summary frames, and thus can be evaluated with only one channel; and (2) so far we had built only a one-channel analog low-pass filter. To test the temporal PCA filtering on multiple-channel data collected from an EEG cap, and to test the applicability of the temporal PCA-based noise removal algorithm to cardiac artifacts, we conducted an alpha wave detection experiment, without the analog filter and without timing error correction. An alpha wave is an oscillatory EEG wave at around 10 Hz that appears when normal subjects close their eyes. Since the phase of an alpha wave cannot be controlled, it cannot be averaged. Hence, it is difficult to detect in the fMRI environment without resorting to a frequency domain analysis. Anami et al. (2003) detected alpha waves from EEG signals whose acquisition was completely synchronized with a special fMRI imaging sequence (see Section 4.2 for a discussion). In our experiment, we detected alpha waves by the temporal PCA-based gradient artifact removal followed by a temporal PCA-based cardiac artifact removal. To reduce the gradient noise, we built a cap that had 19 bipolar outputs (PF1, PF2, F3, Fz, F7, F4, F8, T7, C3, Cz, C4, T4, P7, P3, Pz, P4, P8, O1, O2). Reference electrodes from the ear references were bifurcated into 19 carbon wires, each of which served as a reference in a twisted pair. The 19 twisted pairs were connected to 38 independent channels in the NuAmps and were combined back into 19 bipolar recordings using a montage editor in the Scan software (NeuroScan). One additional electrode was placed on the subject's chest to detect the heartbeat. One 48-year-old male participated in the study after a written consent (Yale HIC#11970). The subject opened and closed his eyes during the fMRI scanning every 30 s, guided by a voice command controlled by PsyScope. The fMRI and EEG recording parameters were the same as the visual oddball experiment.

## 3. Results

### 3.1. Phantom experiment

Fig. 1 shows the timing errors that were detected and corrected using the timing correction algorithm described

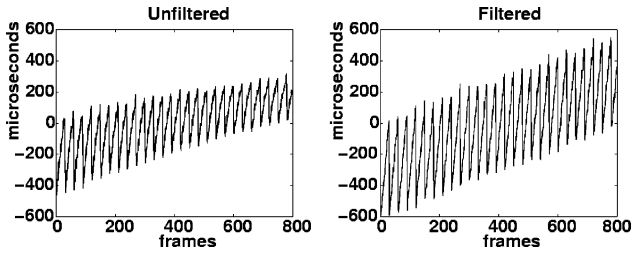


Fig. 1. Timing errors detected in unfiltered (left) and analog filtered (right) signals from the phantom experiment. Timing errors (in ms) computed by the algorithm presented in this paper is shown against number of frames (or number of slice acquisitions). A partial data set consisting of 800 frames (30 frames/volume) was used for these plots.

in Section 2.1. The abscissa corresponds to frames and the number on the ordinate shows the computed delay (if positive) or advancement (if negative) of the waveform in each frame compared to the averaged frame. If there are no timing errors, the line should be horizontal with ordinate equaling zero. It can be seen from Fig. 1 that there are two kinds of timing errors at different timescales. At a shorter time scale, there are zig-zag patterned lines with a period of 30 frames. This is because the fMRI slice acquisitions (30 per image) were not evenly distributed within an image acquisition. That is, slice acquisition interval is slightly shorter than an image acquisition time (or repetition time) divided by number of slices per image (2220 ms/30 = 75 ms), and there is a small timing gap after the acquisition of the last slice in an image and the beginning of the next image acquisition. This could have been predicted and adjusted during the timing error correction, but it has been left unadjusted to show the capability of the algorithm. At a longer scale, it can be seen that there is a positive linear trend in the baseline of the zig-zag trend in Fig. 1, indicating timing errors up to 10 μs/s, or 0.6 ms/min, between the fMRI acquisition clock and EEG acquisition clock. We also confirmed this timing shift by plotting and visually examining the frame data through time (not shown). Fig. 2 shows initial parts of the averaged gradient artifact waveform at different slice positions (1, 8, 15, and 22) before and after timing correction. It can be seen that timing error correction reduces the timing error caused by unevenly distributed slice acquisitions in each image.

The assumption that PC activations corresponding to gradient artifacts and interested signals follow normal distribution (a basis of Eq. (2)) was examined by plotting the distributions of PC activations and corresponding normal probability plot (Fig. 3). If the points in a normal probability plot align to form a straight line, the distribution is close to normal. The first row shows the distribution and normal probability plot of the first PC activation corresponding to the gradient noise. The second row shows the same plots for the second PC activation corresponding to the gradient noise. The third row shows the same plots for the first PC activation corresponding to the injected signal. The fourth row shows the same plots for the second PC activation corresponding to the injected signal.

Fig. 4 shows the correlation of injected and recovered data in the phantom experiment plotted against the number of averaged epochs. Starting from the curve with the lowest correlations (average subtraction with ANC), one can see the effect of adding an analog filter, further adding timing correction as described in Section 2.1, and finally replacing average waveform subtraction and ANC with the temporal PCA filtering. For this comparison, we implemented average waveform subtraction and ANC (Allen et al., 2000; Chen et al., 1989; Widrow et al., 1975) and applied it to the same data. We modified the ANC method to improve the result, as will be described in Section 4.2. As the formula given for the weight adaptation coefficient (called $\mu$) given in Allen et al. (2000) was not optimal for our data, we varied $\mu$ from 0.05 to 0.2 with a step of 0.01 and settled with 0.12, which resulted in the maximum correlation between the injected and recovered signals.

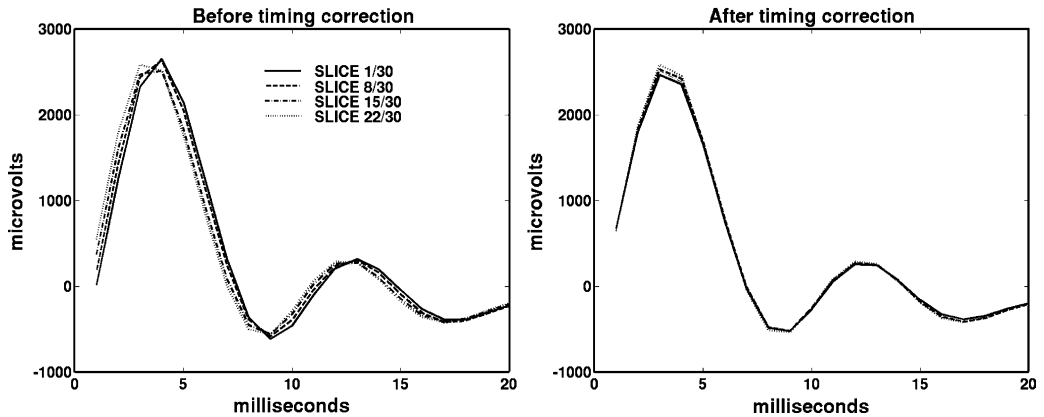


Fig. 2. An example of the effect of timing error correction. Left: first 20 ms of averaged artifact waveforms at different slice positions within images before timing correction. Right: the same after timing correction.

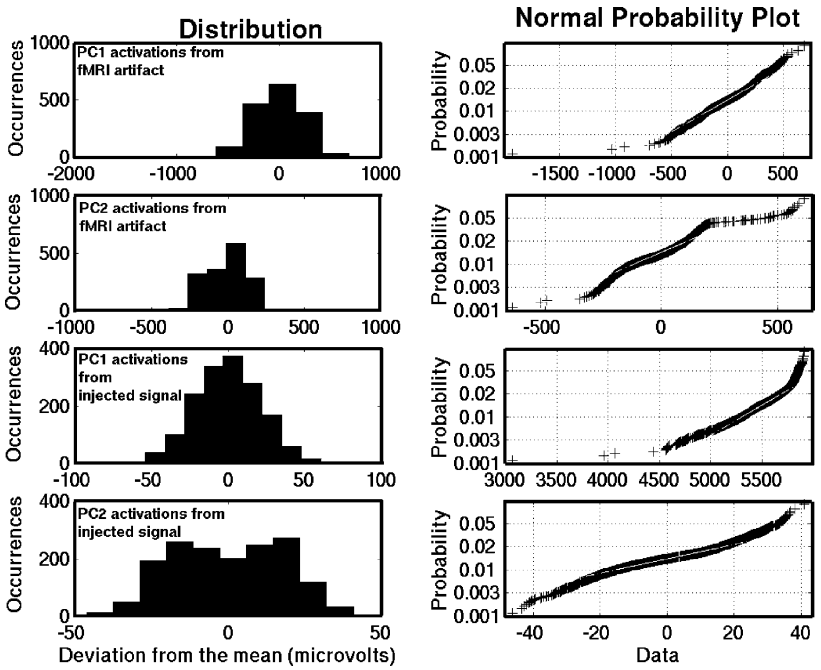


Fig. 3. Test of 'normal distribution' assumptions of PC activations (or scores). The top row shows the distribution of the first PC activation corresponding to the gradient noise (left) and its normal probability plot (right). If the distribution is normal, the points in the normal probability plot aligns on a straight line. The second row shows the same set of plots for the second PC activation corresponding to the gradient noise. The third row shows the same set plots for the first PC activation corresponding to the injected signal. The third row shows the same set of plots for the second PC activation corresponding to the injected signal.

### 3.2. Visual oddball experiment

The ERPs computed from the visual oddball experiment are shown in Fig. 5. The upper half of Fig. 5 shows ERPs computed from the results of timing error correction (as described in Section 2.1) and the temporal PCA filtering, whereas the lower half of Fig. 5 shows the ERPs computed from the results of timing error correction (as described in Section 2.1) and average waveform subtraction with ANC (Allen et al., 2000). In both the upper and the lower halves of Fig. 5, there are two pairs of plots, one on the right and one on the left. Pairs on the left half are computed from 45 non-rejected trials (see below for the rejection criterion) that were randomly selected from the full set. Pairs on the right half are computed from all non-rejected trials. The upper plot in each pair is computed from EEG data acquired without analog filtering, whereas the lower plot is computed from EEG data acquired with an analog filter.

To exclude trials with large movement artifacts, trials whose peak-to-peak amplitudes exceeded mean peak-to-peak amplitude plus 80% of the standard deviation of peak-to-peak amplitudes were excluded from averaging. This artifact rejection was determined using all of the analog filtered EEG data that were processed by average noise subtraction with ANC, and was applied to all other cases.

A composite fMRI map of the four subjects (Fig. 6) showed a significant activation on the left superior temporal gyrus (STG) (Brodmann's area 22, corrected $P < 0.05$) and the medial prefrontal gyrus (Brodmann's area 11, corrected $P < 0.05$) corresponding to oddball visual stimuli. Information from the EEG was not utilized in obtaining this fMRI map.

### 3.3. Alpha wave detection

Fig. 7 shows the result of alpha wave detection. The top plot shows the alpha wave recorded outside the fMRI scanner for reference. The middle plot shows the EEG

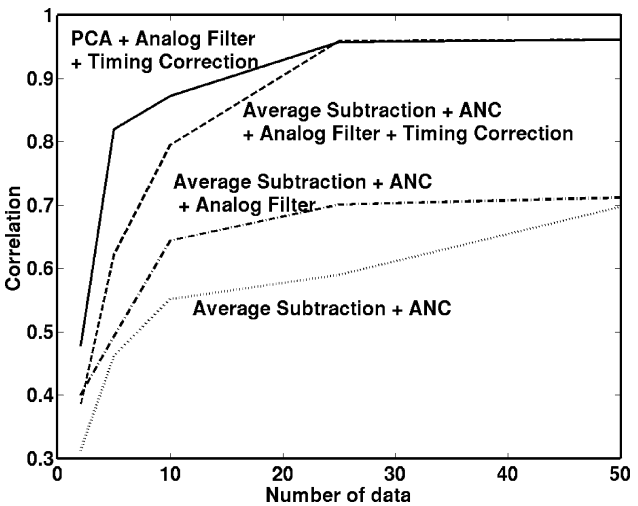


Fig. 4. Correlation of injected and recovered data in a phantom experiment plotted against number of averaged data. Starting from the curve with the lowest correlations (average subtraction + adaptive noise canceling (ANC)), one can see the effect of adding an analog filter, further adding timing correction (described in Section 2.1), and finally replacing average subtraction + ANC with the temporal PCA filtering.

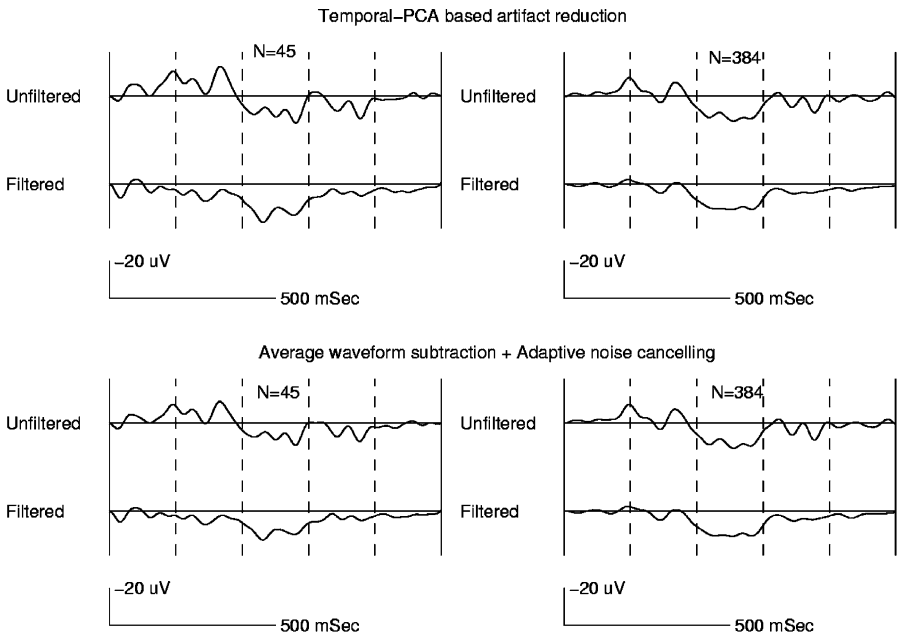


Fig. 5. Average ERPs of four subjects from a visual oddball experiment. The upper half of the figure shows ERPs computed from the results of timing error correction (as described in Section 2.1) and temporal PCA filtering, whereas the lower half of the figure shows the ERPs computed from the results of timing error correction (as described in Section 2.1) and average waveform subtraction with adaptive noise canceling. In both the upper and the lower halves of figure, there are two pairs of plots, one on the right and one on the left. Pairs on the left half are computed from 45 data that were randomly selected from the full set. Pairs on the right half are computed from the full data set. The upper plot in each pair is computed from EEG data acquired without analog filtering, whereas the lower plot is computed from EEG data acquired with an analog filter.

recorded during fMRI scanning and recovered by the temporal PCA-based gradient noise removal, before cardiac artifact removal. The bottom plot shows the data recovered by a cascade of temporal PCA-based gradient artifact removal and temporal PCA-based cardiac artifact removal.

## 4. Discussion

### 4.1. Timing error correction

Fig. 1 demonstrates that the 125 Hz low-pass filtering prior to sampling improves the accuracy of timing error detection. This can be seen from the fact that the width of the zig-zag line stays constant and the longer timescale slope forms a straight line in the plot from the analog filtered EEG, correctly reflecting the fact that timing errors within each brain image acquisition are the same and that the difference between the speed of the EEG and fMRI clocks stays constant. One cycle of a 125 Hz signal yields eight digitized samples at 1 kHz, which provide enough points to compute the phase of the signal. In the gradient artifact removal method described in Allen et al. (2000), timing error is precisely corrected using slice acquisition triggers that are recorded with a high temporal resolution. In our timing error detection algorithm, the delay time of each frame is computed with respect to an averaged frame. As was mentioned in Section 1, because timing errors are computed, there are possibilities for errors in the estimated timing errors. For instance, an estimation error may occur because an averaged frame is not the same as a frame with the average time delay. The averaged frame has a more blurred waveform than individual frames and this can lead to slight errors in the shift measured. However, blurring of sharp edges in the waveforms does not cause a big problem because timing error detection on analog filtered signals depends more on shallow slopes in the gradient artifact waveform. It can be

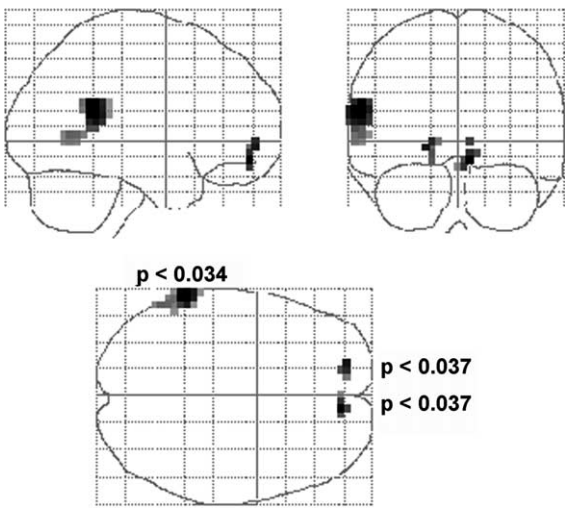


Fig. 6. A composite statistical parametric map of four subjects corresponding to visual oddball stimuli. Non-white voxels had above threshold (corrected $P < 0.005$) significance. A $t$ value shown beside each cluster in the bottom figure shows the minimum $P$ value within that cluster. The right side in the figure corresponds to the right side of the brain.

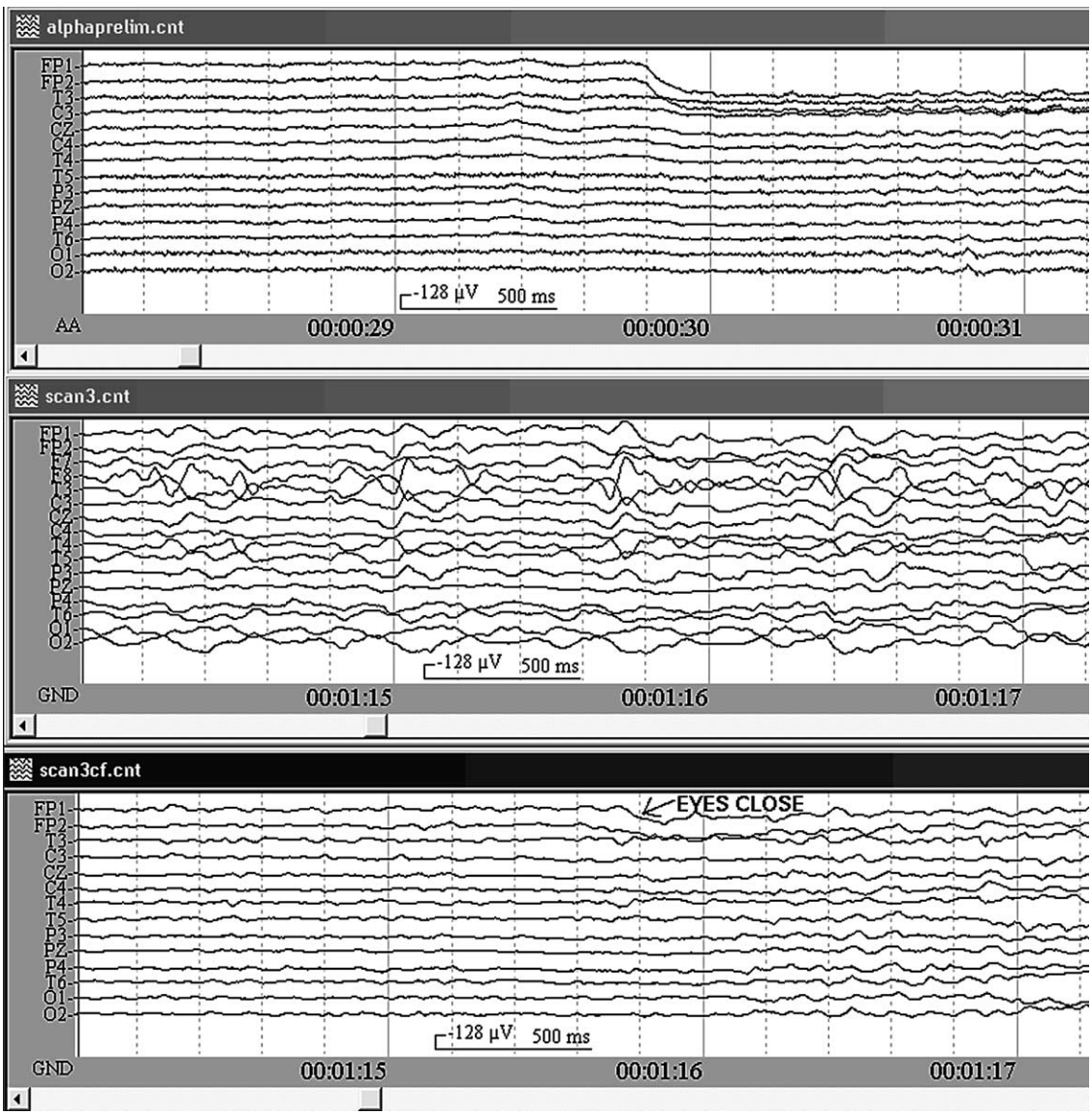


Fig. 7. Alpha wave detection. The top plot shows EEG data recorded outside the fMRI scanner. The subject closed his eyes at around 29.8 s. The middle plot shows EEG data recorded during scanning and recovered by the temporal PCA-based gradient noise removal, but before cardiac artifact removal. The bottom plot shows EEG after temporal PCA-based cardiac artifact removal. It can be seen that the subject closed his eyes at around 1 min and 15.8 s. Some channels are omitted for visibility of the plots.

seen from Fig. 2 that the algorithm can operate on the timing shift between shallow slopes on the waveform and correctly rectify timing errors due to uneven slice acquisitions within a head image.

Instead of computing and correcting timing errors, Bénar et al. (2003) divided frames with different delays into small bins and then computed and subtracted average waveforms within each bins. At least theoretically, timing error detection and correction would result in a better result, because there will be no quantization error of delay times.

As a summary of the evaluation of the timing error correction algorithm described in this paper, it was demonstrated that the algorithm was able to detect and correct timing errors between fMRI slice acquisitions and EEG sampling caused by uneven slice acquisition times within one head image acquisition as well as slight clock speed difference between the fMRI and EEG systems. However, because the current algorithm computes timing errors instead of measuring them, the current algorithm cannot outperform Allen et al.'s (2000) method.

### 4.2. Gradient artifact removal

Before we examine the results of the temporal PCA filtering, an assumption that was made in the formulation of the algorithm, namely the assumption that the distributions of PC activations corresponding to the gradient noise and the interested signal are normal, has to be confirmed. Fig. 3 shows that assumption is largely confirmed for the first PC (the first row for the gradient noise and the third row for

the injected signal), as a majority of points in the normal probability plots are close to a straight line. Admittedly, the higher end of the normal probability plot in the third row shows a deviation towards positive, indicating that the first PC activation corresponding to the injected signal has a longer tail. This means that the temporal PCA filter may attenuate the strong interested signal more than it should (because the variance due to the signal is underestimated), but for the majority of frames its gradient error estimate is close to statistically optimal. The correctness of the normal distribution assumption for first PC is especially important, as 80% of variance in the gradient artifact in the phantom experiment is accounted for by the first component (the second component accounts for 4%).

Next, we compare the performance of the temporal PCA filter to that of average waveform subtraction with ANC (Allen et al., 2000). Fig. 4 shows the effect of adding analog filtering and timing error correction to a baseline performance measured with average waveform subtraction and ANC alone. It also shows that the temporal PCA filter achieved higher correlation between the injected and recovered signals with same numbers of averaged samples, compared to average waveform subtraction with ANC (note: 'a sample' in this context refers to a section of EEG data corresponding to one wave generated by the EEG signal emulator, which is different from the term 'a digitized sample' used in Section 2.4). It is important to measure the quality of the output signals with different number of averaged samples, because as long as the residual error of an artifact reduction is uncorrelated with the interested signal, averaging of the samples reduces the residual error amplitude by a factor of the square root of the number of samples. To examine the amount of noise quantitatively, let us assume that the total variance of the recovered signal consists of three orthogonal components: signal variance $w_s$, residual noise variance that cannot be removed by averaging $w_n$, and noise variance that reduces with averaging $w_g$. It can be assumed that $w_s$ and $w_n$ do not change with averaging. The signal variance can be known from the injected signal, and in this case it was $w_s = 111\ \mu V^2$ (squared microvolts). From the asymptotic correlation $r_{max}$, $w_n$ can be computed as $w_s(1 - r_{max}^2)/r_{max}^2 = 9.2\ \mu V^2$ for both the temporal PCA filter and average waveform subtraction with adaptive noise filtering (both with analog filtering and timing error correction). From the correlation $r$ at each measurement points in Fig. 4, error variance that reduces with averaging can be computed as $w_g = w_s(1 - r^2)/r^2 - w_n$. When 10 samples were averaged, $w_g = 25.7\ \mu V^2$ or $5.1\ \mu V$ root mean square (RMS) for the temporal PCA filtering and $w_g = 55.438\ \mu V^2$, or $7.4\ \mu V$ RMS for average waveform subtraction with adaptive noise filtering.

The main goal of the temporal PCA filter is to remove components in the signal that are not perfectly repetitive, but are time-locked to the image acquisition. An indication of the ability of the temporal PCA filter to remove gradient artifacts with rapidly fluctuating waveforms was observed when timing error correction stage was bypassed. In that case, PCs that had not previously been seen emerged that corresponded to small timing differences among slices within head volumes (not shown). These PCs contained narrow peaks reflecting differences among gradient artifact waveforms at different slice positions within a head image seen in Fig. 2. The temporal PCA filter relies on the statistical characteristics of the variance of PC amplitudes in the EEG during scanning and those without scanning. The ability of adaptive noise cancellation (Allen et al., 2000) to remove time-varying noise crucially depends on the weight adaptation parameter. If the weights do not adapt, the mechanism always predicts the same noise waveform, computed from low-pass filtered trigger pulses. Adaptive weights allow cancellation of time-varying noise, at some risk of attenuating the interested signal as will be described below.

We found that for this particular signal, addition of ANC reduced the correlation between the injected signal and the recovered signal over average waveform removal with analog filtering and timing correction, regardless of the value of the weight adaptation constant ($\mu$). We hypothesized that this is because the injected signal is perfectly periodic: a periodic signal may be easily predicted and removed by the adaptive noise cancellation. In fact, we found that if we use the pure injected waveform (with no gradient noise or white noise) as the primary input, and use the fMRI trigger as the reference signal to the ANC, then the correlation of the output to the primary input was only 0.86. This does not mean that the interested signal is degraded by 14% while the gradient noise is being processed by the ANC, but it does mean that ANC can interfere with a signal that is not synchronized with the reference signal. We found two ways to reduce the alternation of injected signals caused by ANC. One way was to set the bias, or the weight that multiplies to a constant input in the adaptive noise canceller, to zero. This reduces the tendency of the adaptive noise canceller to remove slow changes in the primary signal by changing the bias. Another way, which resulted in a better correlation (0.99), was to high pass filter the estimated noise by a fourth-order Butterworth filter with a cutoff frequency equal to the slice acquisition frequency before it will be subtracted from the primary signal. However, even with this modification to the ANC, addition of ANC did not improve the correlation between the injected and recovered signals over average waveform removal with analog filtering and timing correction. For all measurement points used in Fig. 4, addition of ANC changed the correlation only within the range from $-0.015$ to $+0.012$.

Because processing of a perfectly periodic signal would not provide a fair comparison between the temporal PCA filter and average waveform removal with ANC, we performed another experiment using a signal that is closer in its characteristics to that used in Allen et al. (2000), namely simulated epileptic spikes generated at random timings (Poisson distribution with mean spike interval = 2 s).

Spikes were simulated by a Gaussian function with standard deviation of 10 ms and a peak voltage of 150 μV, and the signal was added to 2 min of phantom data (but without the injected periodic signals). In this case, addition of ANC (with the high pass filter mentioned above) to the average waveform subtraction did improve the correlation between simulated spikes and the recovered signal from 0.22 to 0.26. The temporal PCA filter again outperformed the combination of average waveform subtraction and ANC, yielding a correlation of 0.31.

Compared to the gradient artifact removal method presented in Allen et al. (2000), the temporal PCA filter also has an advantage of utilizing the standalone EEG signal. This means that the temporal filter can utilize some apriori knowledge about the interested signal. On the other hand, the standalone signal has to be recorded carefully so that the amplitude and/or the frequency of interested signal (epileptic spikes for instance) are representative under the experimental condition. Otherwise the detection of interested signals can be positively or negatively biased. If the standalone signal is unavailable, the algorithm still works, by using a white noise of small amplitude as the standalone signal for instance, although the result tends to be negatively biased.

Gradient artifact removal based on frequency domain analysis (Garreffa et al., 2003; Hoffmann et al., 2000) are more effective in removing time-varying gradient artifacts than average waveform subtraction, because they can remove artifacts regardless of phase fluctuation. However, for the same reason, it is possible that they will degrade the interested signal that happen to have frequency components that overlap with gradient noise. The temporal PCA filter may also possibly degrade interested signals: if some frequency component consistently occurs with random phase, then sine and cosine components of this frequency may be included in different PCs, thereby attenuating that frequency component. However, at least the temporal PCA filter takes into account the occurrence of such PC activities in the standalone EEG, and the temporal PCA filter can also be sensitive to the timing of the amplitude change of the frequency component within a frame. As a thought experiment, suppose that the gradient artifact caused by a mechanical vibration is large at the beginning of each slice acquisition and decays towards the end of a slice acquisition, but the phase of the vibration is not the same for all slices. Then two decaying sinusoidal waves with 90° difference would serve as PCs corresponding to this artifact, and so the temporal PCA filter can remove noises with this particular amplitude profile.

Anami et al. (2003) presented a technique called stepping stone sampling, where the EEG is sampled at the exact moments when the gradients are held constant for a very short moment. Thus, the gradient artifacts are extremely small, and because they use a specialized hardware so that EEG and fMRI are driven by a single clock, there is no need for timing error correction. We do not claim that the temporal PCA filtering outperforms stepping stone sampling. In fact, a temporal PCA filter could be used in the place of an average waveform subtraction algorithm used in Anami et al. (2003) after stepping stone sampling, possibly resulting in a better result, considering that stepping stone sampling itself does not reduce time-varying artifacts due to mechanical vibrations.

To summarize the evaluation of the temporal PCA filtering using the data from the phantom experiment, it was found that the current method required less number of samples to be averaged to obtain the same correlation between the injected and recovered signals, compared to average waveform subtraction with ANC. We also argued that the current method reduces noise that are caused by, but are not perfectly synchronized with, slice acquisitions, partly on the basis of the characteristics of the algorithm and partly from an observation of PCs that emerged when timing error correction was bypassed.

### 4.3. Visual oddball experiment

A typical visual P300 latency peaking around 400 ms after stimulus presentation can be seen in all plots in Fig. 5. In both the result from the temporal PCA filtering (the upper half of Fig. 5) and the results of average waveform subtraction with adaptive noise filtering (the lower half), plots that are computed from 45 trials have higher frequency components of around 15 Hz than the plots that are computed from the full set, suggesting that such frequency components are reflection of noises. If that is the case, the filtered data in all filtered–unfiltered pairs show low noise, indicating the effect of analog-filtering.

There was no apparent difference in waveforms obtained by the temporal PCA filtering and by average waveform subtraction with ANC. However, a closer look on the ERPs computed from 45 unfiltered samples reveals that there was a small extra peak (at 830 ms) and a very small valley (at 620 ms) in the plot computed by average waveform subtraction with ANC compared to the plot computed by the temporal PCA filter. Moreover, amplitudes of the P300 computed from 45 samples were higher when processed by the temporal PCA filtering. For instance, the positive peak voltages were 20.7 (temporal PCA filtering) versus 15.4 μV (average waveform subtraction with ANC). Given that there were no amplitude differences between these two methods when all samples were used, it is unlikely that the difference was caused by attenuation of signals by the ANC. Rather, it suggests, although not conclusively, that the residual noise level was lower when the temporal PCA filter is used.

Our fMRI analysis found left temporal and medial frontal activations during the oddball experiment (Fig. 6). As our study used data from only four subjects and our single-channel ERP did not allow us to perform dipole localization, it is impossible for us to tell whether these fMRI activities truly corresponds to P300 found in our

ERPs. However, in light of other studies, left temporal activation may correspond to the visual oddball stimuli. Other EEG–fMRI studies of visual oddball found right (Kruggel et al., 2000) or bilateral (Opitz et al., 1999) STG activations. However, Stevens et al. (2000) compared response to auditory and visual oddball tasks and found strongest activations in the middle temporal gyrus for visual oddball stimuli and activations in the transverse temporal gyrus for auditory oddball stimuli. Identification of the exact loci of activation and their dependency on experimental parameters requires more extensive studies. Near infrared optical topography (OT), like fMRI, is a non-invasive technique that measures hemodynamic changes. Being an optical recording, NIRS does not interfere with EEG recordings. Using simultaneous OT and EEG recording, Kennan et al. (2002) found supramarginal gyrus activation (Brodmann's area 40) corresponding to auditory oddball stimuli.

Our oddball study is only preliminary and is meant to demonstrate the applicability of the temporal PCA filtering to in vivo EEG data. The limitation includes number of subjects (four) and number of EEG channels (only one channel). In a future study, we plan to use more than 15 subjects and at least 19 electrodes, in order to compare EEG source localization with fMRI activation maps.

### 4.4. Cardiac artifact removal

Fig. 7 demonstrates the ability of the temporal PCA-based cardiac artifact removal in reducing the cardiac artifacts. After cardiac noise removal, the artifact due to eye-closing and the alpha rhythm following that became apparent. However, the average peak-to-peak amplitude of the residual noise was 25 μV before eye-closing, indicating that further efforts including improvements to the current cardiac artifact removal method are required to recover lower amplitude signals.

### 4.5. Computational requirements

We implemented timing error correction, the temporal PCA filtering, and temporal PCA-based cardiac artifact removal algorithms using MATLAB (The MathWorks Inc., Natick, MA). In the current implementation with non-compiled MATLAB code, to process 1 s of EEG data sampled at 1 kHz it takes 13 ms for timing error detection, 26 ms per channel for timing error correction, 7 ms per channel for temporal PCA filtering, 57 ms for cardiac pulse detection, and 10 ms per channel for temporal PCA-based cardiac artifact removal, on a 2.8 GHz Intel Xeon™ based Dell Precision™ desktop computer (Dell computer corporation, Austin, Texas) with 512 kilobytes cache and 4 gigabytes main memory. Note that for temporal PCA, computation time increases only linearly with EEG acquisition time increases, because the size of the co-variance matrix does not change with the acquisition time.

Our current algorithms work offline on stored EEG data. Online removal of fMRI artifacts (e.g. Garreffa et al., 2003) would be a desirable feature in applications such as continuous EEG–fMRI acquisition of epileptic spikes and sleep state monitoring during fMRI scanning. In the current implementation, EEG data cannot be processed online because temporal PCA requires all the EEG data beforehand. However, there are PCA algorithms that continuously update PC's as more data are added (Oja, 1992; Sanger, 1989). Using such algorithms, it is possible to realize online processing of EEG data. Moreover, unlike spatial PCA, temporal PCA algorithm operates independently on each channel. Therefore computation time increases only linearly as more channels are added, and it is also possible to take advantage of multiple processor architectures.

## 5. Conclusions

We have proposed an EEG–fMRI timing error correction algorithm and temporal PCA filtering, which can be used in combination to recover EEG contaminated with time-varying gradient artifacts. In a phantom experiment, the temporal PCA filter required less samples to be averaged, compared to the average waveform subtraction method with ANC, to achieve the same correlation between the injected and recovered signals, indicating a better gradient noise reduction capability. Using these algorithms, we were also able to obtain a P300 ERP in a continuous simultaneous EEG–fMRI experiment.

## Acknowledgements

This research was supported by NIH R01-NS40497 and NIH R01-NS38467. The authors would like to thank Jed Meltzer for reading drafts of the paper and providing useful discussions.

## References

Allen PJ, Polizzi G, Krakow K, Fish DR, Lemieux L. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. NeuroImage 1998;8:229–39.

Allen PJ, Josephs O, Turner R. A method for removing imaging artifact from continuous EEG recorded during functional MRI. NeuroImage 2000;12:230–9.

Anami K, Mori T, Tanaka F, Kawagoe Y, Okamoto J, Yarita M, Ohnishi T, Yumoto M, Matsuda H, Saitoh O. Stepping stone sampling for retrieving artifact-free electroencephalogram during functional magnetic resonance imaging. NeuroImage 2003;19:281–95.

Baudewig J, Bittermann HJ, Paulus W, Frahm J. Simultaneous EEG and functional MRI of epileptic activity: a case report. Clin Neurophysiol 2001;112(7):1196–200.

Bénar CG, Aghakhani Y, Wang Y, Izenberg A, Al-Asmi A, Dubeau F, Gotman J. Quality of EEG in simultaneous EEG–fMRI for epilepsy. Clin Neurophysiol 2003;114:569–80.

Bonmassar G, Anami K, Ives J, Belliveau JW. Visual evoked potential (VEP) measured by simultaneous 64-channel EEG and 3T fMRI. NeuroReport 1999;10:1893–7.

Chen J, Vandewalle J, Sansen W, Vantrappen G, Janssens J. Adaptive method for cancellation of respiratory artifact in electrogastric measurements. Med Biol Eng Comput 1989;27: 57–63.

Cohen MS, Goldman R, Jerome Jr E. Simultaneous EEG and fMRI made easy. Proc Org Hum Brain Mapping, Brighton, UK 2001;6.

Garreffa G, Carnì M, Gualniera G, Ricci GB, Bozzao L, De Carli D, Morasso P, Pantano P, Colonnese C, Roma V, Maraviglia B. Real-time MR artifacts filtering during continuous EEG/fMRI acquisition. Magn Reson Imaging 2003;21:1175–89.

Goldman RI, Stem JM, Engel Jr J, Cohen MS. Acquiring simultaneous EEG and functional MRI. Clin Neurophysiol 2000; 111:1974–80.

Hoffmann A, Jäger L, Werhahn KJ, Jaschke M, Noachtar S, Reiser M. Electroencephalography during functional echo-planer imaging: detection of epileptic spikes using post-processing methods. Magn Reson Med 2000;44:791–8.

Horovitz SG, Skudlarski P, Gore JC. Correlations and dissociations between BOLD signal and P300 amplitude in an auditory oddball task: a parametric approach to combining fMRI and ERP. Magn Reson Imaging 2002;20:319–25.

Kennan RP, Horovitz SG, Maki A, Yamashita Y, Koizumi H, Gore JC. Simultaneous recording of event-related auditory oddball response using transcranial near infrared optical topography and surface EEG. NeuroImage 2002;16:587–92.

Krakow K, Woermann FG, Symms MR, Allen PJ, Lemieux L, Barker GJ, Duncan JS, Fish DR. EEG-triggered functional MRI of interictal epileptiform activity in patients with partial seizures. Brain 1999; 122:1679–88.

Kruggel F, Wiggins CJ, Hermann CS, von Cramon DY. Recording of the event-related potentials during functional MRI at 3.0 Tesla field strength. Magn Reson Med 2000;44:277–82.

Lemieux L, Salek-Haddadi A, Josephs O, Allen P, Toms N, Scott C, Krakow K, Turner R, Fish DR. Event-related fMRI with simultaneous and continuous EEG: description of the method and initial case report. NeuroImage 2001;14:780–7.

Liebenthal E, Ellingson ML, Spanaki MV, Prieto TE, Ropella KM, Binder JR. Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. NeuroImage 2003;19:1395–404.

Oja E. Principal components, minor components, and linear neural networks. Neural Netw 1992;5:927–35.

Opitz B, Mecklinger A, von Cramon DY, Kruggel F. Combining electrophysiological and hemodynamic measures of the auditory oddball. Psychophysiology 1999;36:142–7.

Salek-Haddadi A, Mershhemke M, Lemieux L, Fish DR. Simultaneous EEG-correlated ictal fMRI. NeuroImage 2002;16:32–40.

Sanger TD. Optimal unsupervised learning in a single-layer linear feedforward neural network. Neural Netw 1989;2:459–73.

Seeck M, Lazeyras F, Michel CM, Blanke O, Gericke CA, Ives J, Delavelle J, Golay X, Haenggeli CA, de Tribolet N, Landis T. Non-invasive epileptic focus localization using EEG-triggered functional MRI and electromagnetic tomography. Electroencephalogr Clin Neurophysiol 1998;106:508–12.

Sommer M, Meinhardt J, Volz HP. Combined measurement of event-related potentials (ERPs) and fMRI. Acta Biol Exp 2003;63:49–53.

Stevens AA, Skudlarski P, Gatenby JC, Gore JC. Event-related fMRI of auditory and visual oddball tasks. Magn Res Imaging 2000; 18:495–502.

Warach S, Ives JR, Schlaug G, Patel MD, Darby DG, Thangaraj V, Edelman RR, Schomer DL. EEG-triggered echo-planer functional MRI in epilepsy. Neurology 1996;47:89–93.

Widrow B, Glover JR, McCool JM, Kaunitz J, Williams CS, Hearn RH, Zeidler JR, Dong Jr EFJ, Goodlin RC. Adaptive noise cancellation: principles and applications. Proc IEEE 1975;63:1692–716.