Multivariate Composite Estimators: New Ways to Track Signals with Application to Human Cerebral Potentials

Wenyu Wang, Henri Begleiter, and Bernice Porjesz

Summary: In this paper, we propose new multivariate composite estimators (MCE) to track multiple channel potential waves of individual and grand averages of recorded potentials at the scalp from a number of trials in a group of subjects. The advantages of these estimators over simple averages used in the literature are that they have larger signal to noise ratios (SNR), and that they have taken into account variations and correlations of the recorded potentials at the electrode sites as well as differences among subjects. Multivariate techniques and composite concept are used for deriving the estimators. We also provide an application to human event-related potentials in a memory for faces experiment.

Key words: Multivariate composite estimator; Signal tracking; Event-related potentials.

Introduction

Tracking signals or waveforms from a number of trials in an experiment is a fundamental problem in topographic studies of event related potentials (ERP) and EEG. The quality of the tracked signals or waveforms will determine the reliability of further studies based on them. The most widely used estimators of ERP waveforms are simple averages

$$p_{ij\cdot}(t) = \frac{1}{m}\sum_{k=1}^{m} p_{ijk}(t) \tag{1.1}$$

and further grand averages

$$p_{i\cdot\cdot}(t) = \frac{1}{ms}\sum_{j=1}^{s}\sum_{k=1}^{m} p_{ijk}(t) \tag{1.2}$$

where $p_{ijk}(t)$ denotes recorded potential at time t, electrode site i on the scalp and the kth trial from subject j, $t=t_1, t_2, \ldots, t_v$, $i=1,2,\ldots,n$, $j=1,2,\ldots,s$ and $k=1,2,\ldots,m$. Mathematically the simple average $p_{ij\cdot}(t)$ and the simple grand average $p_{i\cdot\cdot}(t)$ can be thought of as the least squares estimators of $\mu_i(t)$ in the following model for subject j and for all subjects in the group, respectively.

$$\mathbf{P}_{jk}(t) = \mu(t) + \mathbf{B}_{jk}(t) \tag{1.3}$$

where

$$\mathbf{P}_{jk}(t) = (p_{1jk}(t), p_{2jk}(t), \ldots, p_{njk}(t))^{\tau} \tag{1.4}$$

$$\mu(t) = (\mu_1(t), \mu_2(t), \ldots, \mu_n(t))^{\tau} \tag{1.5}$$

$$\mathbf{B}_{jk}(t) = (\varepsilon_{1jk}(t), \varepsilon_{2jk}(t), \ldots, \varepsilon_{njk}(t))^{\tau} \tag{1.6}$$

$\varepsilon_{ijk}(t)$'s are the background EEG waves and assumed to be stationary processes, have a random relation to the stimulus, and be, for any fixed t, independent identical distributed (i.i.d.) random noise with mean zero and common variance $\sigma^2 > 0$. Such a model is based on the assumption that an event related brain signal tends to occur at about the same time, and has the same shape on each trial for every subject in the same group. Under this assumption, the SNR can be improved by taking the averages (see McGillem and Aunon 1987).

However, it is clear from the model that there are some problems with using $p_{ij\cdot}(t)$ and $p_{i\cdot\cdot}(t)$: (1) the event related signal may not occur at exactly the same time, may not have exactly the same shape, and may not be present to the same degree on every trial; (2) the background EEG waves $\{\varepsilon_{ijk}(t)\}$'s may not be stationary processes; (3) in the multiple channel case (usually 19, 31, 61, or 128), correla-

Neurodynamics Lab., Department of Psychiatry, SUNY Health Science Center at Brooklyn, New York.

Accepted for publication: May 31, 1994.

We thank David Chorlian and Arthur Stimus for providing the data. This work was supported by N.I.H. grants AA-02686 and AA-05524.

Correspondence and reprint requests should be addressed to Wenyu Wang, Ph.D., Neurodynamics Lab., Box 1203, Department of Psychiatry, SUNY Health Science Center at Brooklyn, 450 Clarkson Ave., Brooklyn, NY 11203, U.S.A.

tions among recorded potentials at different electrode sites may not be negligible, and variations of recorded potentials at different electrode sites may not be the same, and (4) the problem is even worse when using $p_{i\cdot\cdot}(t)$ than using $p_{ij\cdot}(t)$ since differences of means and variances among the subjects are also ignored in $p_{i\cdot\cdot}(t)$.

For single channel records, considerable efforts have been devoted to dealing with the problem (1) (see above) in the literature. The methods proposed are either to average after reducing the effect of outlier trials such as averaging without outlier trials, or to use weighted averages with outliers downweighted (Gasser et al. 1983; Gevins et al. 1986), or to average latency corrected trials (Woody 1967; McGillem and Aunon 1977), or to use other estimators such as the median instead of the averages. Some low or high-pass or specially designed filters and band-pass amplifiers are also developed to improve recorded data and to reject artifacts during recording. These issues have been reviewed in Gevins (1987) and McGillem and Aunon (1987).

There has been a lack of discussion in the literature concerning problems (2)-(4), and (1) under multiple channel cases. Recently these problems have received more attention as multiple channels are currently being widely used in studies of ERP. Even if each trial can be treated as a repeated sample, the simple averages $p_{ij\cdot}(t)$ and $p_{i\cdot\cdot}(t)$ are still limited by problems (2) to (4). Therefore, it is important to find new ways to track the signal or waveform in multiple channel cases to deal with problems (1) to (4).

We propose novel ways to deal with problems (2) to (4) in multiple channel cases under an assumption that problem (1) does not exist. The proposed multivariate composite estimators (MCE) for replacing the simple averages combine information from all relevant sources such as trials, electrode sites and subjects in the same group, and are optimal estimators in the sense that they minimize the mean squared errors. We will show some application results from a memory for faces experiment.

Tracking signal by multivariate composite estimators

In order to deal with problems (2) to (4) in multiple channel cases, we consider estimators of $\vartheta_j(t)$ in the following new model

$$P_{jk}(t) = \vartheta_j(t) + B_{jk}(t) \tag{2.1}$$

as the estimators of signals or waveforms, where notations $P_{jk}(t)$ and $B_{jk}(t)$ are the same as in the introduction section, and

$$\vartheta_j(t) = (\vartheta_{1j}(t), \vartheta_{2j}(t), \ldots, \vartheta_{nj}(t))^\tau$$

We assume, for each fixed j and t, $B_{jk}(t)$, $k = 1,2,\ldots,m$, are i.i.d. random vectors with mean $E(B_{jk}(t)) = O$ and variance covariance matrix $V_j(t)$ ($V_j(t) = E(B_{jk}(t)\, B_{jk}(t)^\tau)$), and $B_{jk}(t)$ are independent for different j.

There are several advantages of this new model compared to the old model (1.3). In the new model, the background EEG waves $\{\varepsilon_{ijk}(t)\}$ need not be stationary while in the old model $\{\varepsilon_{ijk}(t)\}$ are restricted to stationary processes. In the new model, variations and correlations of recorded potentials at the electrode sites for different subjects are represented by $V_j(t)$'s, and the differences among the subjects are represented by both the mean $\vartheta_j(t)$'s and $V_j(t)$'s. In contrast, in the old model, the mean $\mu_i(t)$ is the same for any subject in the group, and $\{\varepsilon_{ijk}(t)\}$ are assumed to be i.i.d. random noises with common mean zero and common variance $\sigma^2>0$ (this means the correlations among recorded potentials at electrode sites on the scalp are zeros, and the variances are the same for recorded potentials at any electrode sites and any subject in the group). Therefore, it is expected that estimators of $\vartheta_j(t)$'s from the new model are closer to reality than the simple averages $P_{ij\cdot}(t)$'s and $P_{i\cdot\cdot}(t)$. Moreover, we will show that the new estimators also improve the simple averages in the signal to noise ratio (SNR).

We now construct estimators of ϑ_j's. For simplicity, we omit the subscript t unless it is necessary. Let us consider two extreme cases. First, if there are no similarities among the subjects, i.e., ϑ_j's are different, it is appropriate to take $P_{j\cdot} = (1/m) \sum_{k=1}^{m} P_{jk}$ as an estimator of ϑ_j, which is the simple average. On the other hand, assuming there are strong similarities among the subjects, i.e., $\vartheta_j \approx \mu$, an estimator of ϑ_j for any j should be close to

$$\mu^* \overset{def}{=} \left(\sum_{j=1}^{s} V_j^{-1}\right)^{-1} \sum_{j=1}^{s} V_j^{-1} P_{j\cdot} \tag{2.2}$$

which is the variance covariance weighted sample mean and obtained by minimizing

$$\sum_{j=1}^{s} \sum_{k=1}^{m} (P_{jk} - \mu)\, V_j^{-1} (P_{jk} - \mu)^\tau$$

with respect to μ. When a real situation is between these two extremes one may take a composite estimator of ϑ_{ij} given by

$$\vartheta_{ij}(a_{ij}) = (1 - a_{ij})\, p_{ij\cdot} + a_{ij}\, e_i^\tau \mu^* \tag{2.3}$$

where a_{ij}'s are unknown constants ($0 \le a_{ij} \le 1$), and $\mathbf{e}_i$ is a $n*1$ column vector having 1 for the ith element and 0 for the others. Here we discuss each component of ϑ_j separately rather than for the whole ϑ_j in order to get all possible gains from the composite estimating procedure. $1\text{-}a_{ij}$ represents the weight applied to individual features (for electrode i and subject j) while a_{ij} represents the weight applied to a group feature (for all electrodes and all subjects).

Assuming that ϑ_{ij}'s and V_j's are fixed and known, we can determine a_{ij} by minimizing the mean squared error (MSE) $E[\{\vartheta_{ij}(a_{ij}) - \vartheta_{ij}\}^2]$ with respect to a_{ij}. The optimal choice is given by

$$a_{ij}^* = \frac{e_i^T(V_j - A)e_i}{e_i^T(V_j - A)\, e_i + m\left[\vartheta_{ij} - e_i^T A\, (\Sigma_{l=1}^s V_l^{-1}\vartheta_l)\right]^2} \quad (2.4)$$

where $A = (\Sigma_{l=1}^s V_l^{-1})^{-1}$

Therefore, the optimal composite estimator of ϑ_{ij} in the class described by (2.3) is given by

$$\vartheta_{ij}^* = (1 - a_{ij}^*)\, p_{ij\cdot} + a_{ij}^*\, e_i^T \mu^* \quad (2.5)$$

From a well-known matrix result that

$$V_j - (\sum_{l=1}^{s} V_l^{-1})^{-1} = V_j\,(V_j + (\sum_{l \neq j} V_l^{-1})^{-1})^{-1}\, V_j$$

which is a positive definite matrix, it is easy to see that $0 < a_{ij}^* \leq 1$. Also, when $\vartheta_j = \mu$, $a_{ij}^* = 1$ and thus $\vartheta_{ij}^* = \mu^*$. This means that if the features for every subject in a group are the same, then our procedure will take the group feature as the feature for each subject. Otherwise the value of a_{ij}^* depends on the value of

$$m\,[\vartheta_{ij} - e_i^T\,(\sum_{l=1}^{s} V_l^{-1})^{-1}\,(\sum_{l=1}^{s} V_l^{-1}\vartheta_l)\,]^2$$

The larger the m or the distance of ϑ_{ij} from $e_i^T(\Sigma_{l=1}^s V_l^{-1})^{-1}(\Sigma_{l=1}^s V_l^{-1}\vartheta_l)$, the smaller the value of a_{ij}^* is. That is to say that if m is very large or the individual feature is very different from the group feature, then our procedure will give less weight to the group feature in favor of the individual feature. We suggest the use of μ as an estimator of the grand average. In terms of the SNR (McGillem and Aunon 1987), note that

$$E[\{p_{ij\cdot} - \vartheta_{ij}\}^2] - E[\{\vartheta_{ij}^* - \vartheta_{ij}\}^2] = \frac{1}{m}\, a_{ij}^*\, e_i^T\,(V_j - (\sum_{l=1}^{s} V_l^{-1})^{-1})\, e_i > 0 \quad (2.6)$$

the SNR will be larger by using ϑ_{ij}^* than by using the simple average $p_{ij\cdot}$. This kind of composite estimator has been used in estimation of consumer expenditures for the U.S. Consumer Price Index Numbers and has been shown to be superior to other estimators (Lahiri and Wang 1992).

In practice, V_j, μ^* and a_{ij}^* in (2.4) and (2.5) are unknown and need to be estimated. They and ϑ_{ij}^* estimators are constructed in Appendix. To compare the performance of ϑ_{ij}^* with the simple average $p_{ij\cdot}$, we use the relative improvement in MSE ($RIMSE_{ij}$) of ϑ_{ij}^* to the simple average $p_{ij\cdot}$,

$$RIMSE_{ij} = \frac{E[\{p_{ij\cdot} - \vartheta_{ij}\}^2] - E[\{\vartheta_{ij}^* - \vartheta_{ij}\}^2]}{E[\{p_{ij\cdot} - \vartheta_{ij}\}^2]} \quad (2.7)$$

and the improvement in SNR ($ISNRij$) of ϑ_{ij}^* to the simple average $pij.$,

$$ISNR_{ij} = \frac{E[\{p_{ij\cdot} - \vartheta_{ij}\}^2]}{E[\{\vartheta_{ij}^* - \vartheta_{ij}\}^2]} = \frac{1}{1 - RIMSE_{ij}}\,. \quad (2.8)$$

Their estimators are also given in Appendix.

Application

In this section, we apply the new model to track human brain signals or waveforms to an event-related potential (ERP) experiment dealing with memory for faces. Twenty two right-handed males (mean age = 20.7) volunteered for this experiment and were paid for their participation. All individuals were fitted with an electrode cap (ECI Electrocap International). For half of the subjects we recorded from 31 electrodes, for the other subjects we recorded from 62 electrodes. All scalp electrodes were referred to Cz according to methods described previously (Begleiter et al. 1993). Subjects were grounded with a forehead electrode, and all impedances were kept below 5 kOhms. Vertical eye movements were monitored with electrodes placed directly above and below the right eye, and horizontal eye movements were monitored with electrodes placed at the outer canthi of the eyes. Trials with excessive eye movements (> 73.2 μV) were eliminated. The electrical activity recorded at each electrode was fed to a set of amplifiers (Sensorium 2000) with a 10,000 gain and a bandpass of 0.02-100 Hz. The amplified activity was sampled at a rate of 256 Hz during an epoch of 100 msec. preceding, and 1 sec. following each stimulus presentation.

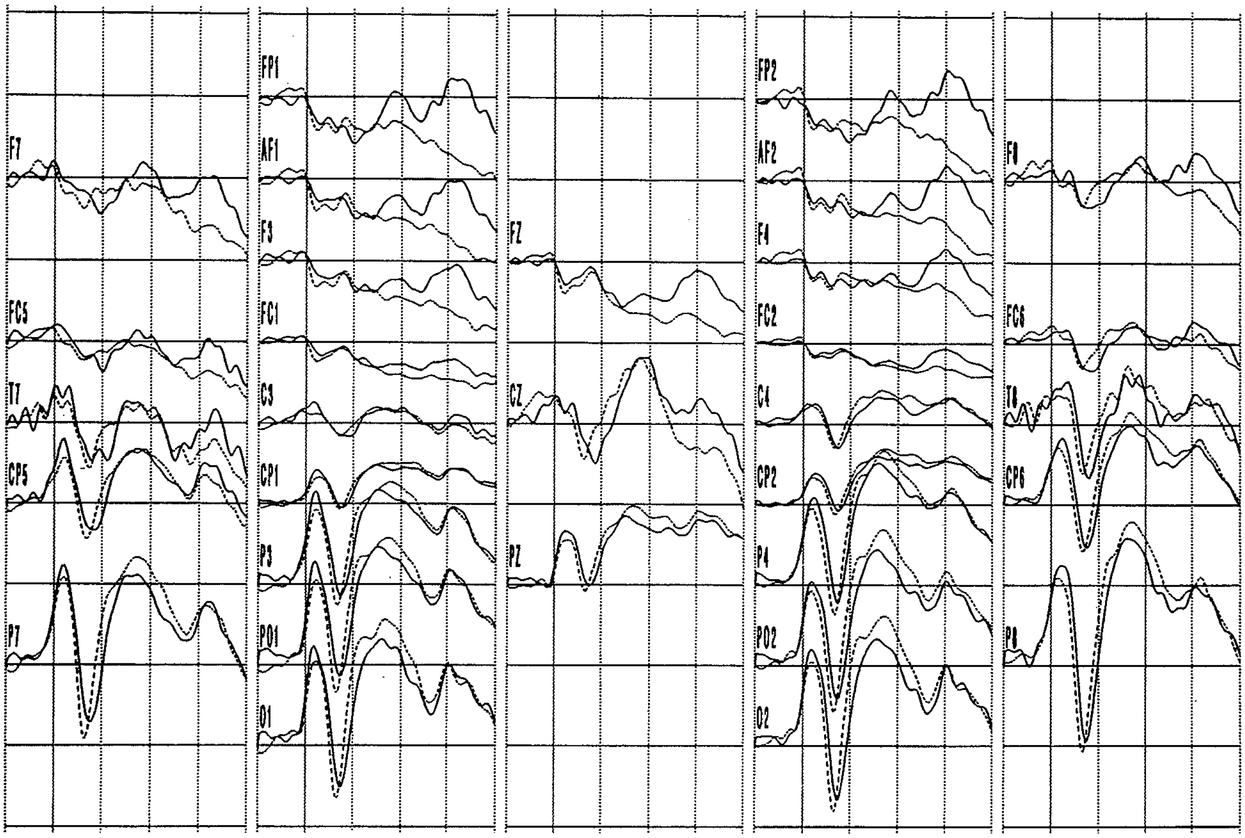

Figure 1. The grand average waves of potentials recorded at 31 electrode sites in the unprimed famous face (solid lines) and the unprimed non-famous face (dashed lines) cases. The interval between two adjacent reference lines on the time-axis (horizontal axis) is 100 ms. The interval between two adjacent reference lines on the vertical axis is 4 micro volts.

The subject was seated in a reclining chair located in a sound- attenuated RF shielded room (IAC) and fixated a point in the center of a computer display located 1 m away from his eyes. A series of 240 faces were presented with an interstimulus interval (ISI) of 1.6 seconds. The face stimuli consisted of 120 male and female famous faces. These famous faces were selected from high quality photographs of well known personalities in the entertainment world, the world of politics, and the sports world. The other 120 faces were carefully selected males and females from medical school yearbooks. Each face was presented in black and white on a high resolution computer screen as a frontal view without shoulders. The experimental paradigm was designed to examine the effects of repetition priming for both famous and non-famous faces. For famous faces, 40 trials consisted of the immediate repetition of the same face, 40 trials were preceded by a different famous face, and 40 trials were preceded by a non-famous face. Similarly, for the non-famous faces, 40 trials consisted of the immediate repetition of the same face, 40 trials were preceded by different faces, and 40 trials were preceded by famous faces. These conditions yielded a total of 240 trials which were equally likely to be presented with the restriction that no more than five famous or non-famous faces would be displayed in a row. At the beginning of the experiment, the subject was only told that two types of stimuli would be presented: famous faces and non-famous faces; he was asked to press a button in one hand as quickly as possible if he recognized the person, and to press the button in the other hand if he did not recognize the individual. The order of the buttons was randomized across subjects. The ERPs were recorded for six cases: primed famous faces, unprimed famous faces, primed non-famous faces, unprimed non-famous faces, famous faces followed by non-famous faces, and non-famous faces followed by famous faces. Figure 1 shows the grand average waves of potentials recorded at 31 electrode sites in the unprimed famous faces case (solid lines) and the unprimed non-famous faces case (dashed lines).

For brevity, we only used the raw data (consisting of 8 subjects, 24 trials and 127 samples at the rate 3.906 ms. per sample in a 500 ms interval for each trial) recorded for the unprimed famous face case. For each sample point, the data at the point, the previous point and the next point are used in calculation of $\hat{V}_j$ to ensure nonsingularity. Denote $r_{il}(j,t)$ the sample correlation coefficient of recorded potentials between electrode i and l at time t

Table I. Statistics of $\{r_{il}(j,t), i, l = 1,\ldots,31;\ t = 1,\ldots,127\}$ for each subject

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	-0.818	0.370	0.972	0.113	0.406	0.665
2	-0.853	0.399	0.986	0.154	0.446	0.683
3	-0.808	0.278	0.975	-0.018	0.305	0.609
4	-0.683	0.376	0.984	0.119	0.427	0.683
5	-0.783	0.356	0.993	0.052	0.402	0.697
6	-0.836	0.330	0.984	0.021	0.368	0.679
7	-0.862	0.383	0.992	0.131	0.428	0.683
8	-0.905	0.362	0.990	0.069	0.404	0.714

1. $r_{il}(j,t)$ - The sample correlation coefficient of recorded potentials between electrode i and l at the t-th sample for subject j.

2. Here and hereafter P25, P50 and P75 denote the twenty-fifth, the fiftieth and the seventy-fifth percentiles, respectively.

Table II. Statistics of $\{v_i(j,t), i = 1,\ldots,31;\ t = 1,\ldots,127\}$ for each subject.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	1.504	66.529	411.696	24.847	52.545	84.138
2	1.312	33.794	268.328	16.735	29.786	43.814
3	0.240	11.763	87.243	3.907	8.713	15.641
4	0.332	33.448	303.846	11.982	25.707	45.204
5	0.615	43.921	395.140	13.845	33.805	58.551
6	0.708	26.301	129.221	8.367	20.991	38.339
7	3.805	62.491	277.631	33.825	57.684	85.157
8	3.622	83.368	434.102	36.987	66.253	114.406

1. $v_i(j,t)$ - The sample variance of recorded potentials at electrode i at the t-th sample for subject j.

for subject j, and $v_i(j,t)$ the sample variance of recorded potentials at electrode i at the t-th sample for subject j, i, l=1,...,31; t=1,...,127; j=1,...,8.

Table I shows the statistics of $\{r_{il}(j,t),\ i,\ l = 1,\ldots,31;\ t = 1,\ldots,127\}$ for each subject j. Comparing with the critical value of Pearson's Correlation Coefficient $r_{n\text{-}2,\alpha}$, where n denotes the sample size (here, n=24x3=72) and α the significant level, $r_{70,0.01}$=.3017, we see that more than fifty percent of the correlation coefficients of recorded potentials among electrode sites are significantly different from zero. In other words, very strong linear correlations exist among recorded potentials at different electrode sites. Therefore, the assumptions in the model (1.3) that correlations among recorded potentials at different electrode sites are negligible and variations of recorded potentials at different electrode sites are the same, which the simple averages are based on, are not suitable in multiple channel cases. Such impropriety of the assumptions can also be seen in Table II. Table II shows strong varibility of the variances of recorded potentials at different electrode sites. In contrast, the correlations and variability are considered in our proposed estimators. Table III shows the statistics of $\{\hat{a}_{ij}(t), i = 1,\ldots,31;\ t = 1,\ldots,127\}$ for each subject j, where $\hat{a}_{ij}(t)$ is the estimator (defined in Appendix) of the composite factor $a^*_{ij}(t)$. Since $a^*_{ij}(t)$ represents the weight or the percentage of information borrowed from all relative sources, we can see from the table the contribution of all relative sources is about 38 percents in average. The relative improvement in mean squared error $RIMSE_{ij}(t)$ of the proposed $\vartheta^*_{ij}(t)$ to the simple average $p_{ij\cdot}(t)$ is shown in table IV, where $\widehat{RIMSE}_{ij}(t)$ is the estimator (defined in Appendix) of the $RIMSE_{ij}(t)$. From this table, we see that the proposed estimator improves the simple averages in MSE about 38 percent in average. Table V shows how the signal to noise ratio is improved by using the proposed $\hat{\vartheta}^*_{ij}(t)$ compared to using the simple average $p_{ij\cdot}(t)$, where $\widehat{ISNR}_{ij}(t)$ is the estimator (defined in Appendix) of the improvement in the signal to noise ratio $ISNR_{ij}(t)$ of the

Table III. Statistics of $\{\hat{a}_{ij}(t), i = 1,\ldots,31;\ t = 1,\ldots,127\}$ for each subject.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	0.017	0.498	1.000	0.210	0.461	0.796
2	0.011	0.407	1.000	0.124	0.303	0.679
3	0.009	0.380	1.000	0.112	0.269	0.636
4	0.010	0.349	1.000	0.076	0.216	0.589
5	0.003	0.250	1.000	0.035	0.098	0.365
6	0.009	0.414	1.000	0.126	0.317	0.699
7	0.011	0.369	1.000	0.079	0.241	0.646
8	0.011	0.368	1.000	0.074	0.230	0.671
Average		0.379				

1. $\hat{a}_{ij}(t)$ - The estimator (defined in Appendix) of the composite factor $a^*_{ij}(t)$ at electrode i at the t-th sample for subject j.

Table IV. Statistics of $\{\widehat{RIMSE}_{ij}(t), i = 1,...,31; t = 1,...,127\}$ for each subject.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	0.0171	0.497	0.998	0.209	0.460	0.794
2	0.0110	0.405	0.997	0.124	0.302	0.676
3	0.0096	0.374	0.993	0.110	0.265	0.628
4	0.0108	0.346	0.997	0.075	0.214	0.585
5	0.0037	0.249	0.998	0.035	0.098	0.364
6	0.0093	0.411	0.997	0.125	0.315	0.694
7	0.0117	0.368	0.999	0.079	0.241	0.644
8	0.0115	0.367	0.999	0.074	0.229	0.670
Average		0.378				

1. $\widehat{RIMSE}_{ij}(t)$ - The estimator (defined in Appendix) of the relative improvement in mean squared error $RIMSE_{ij}(t)$ of the proposed $\vartheta^*_{ij}(t)$ to the simple average $p_{ij\cdot}(t)$ at electrode *i* at the *t*-th sample for subject *j*.

Table V. Statistics of $\{\widehat{ISNR}_{ij}(t), i = 1,..., 31; t = 1,..., 127\}$ for each subject.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	1.017	16.068	822.42	1.265	1.854	4.865
2	1.011	8.221	360.70	1.142	1.433	3.091
3	1.009	5.023	162.37	1.124	1.361	2.688
4	1.011	7.289	378.57	1.082	1.272	2.410
5	1.003	5.625	524.03	1.036	1.109	1.573
6	1.009	8.375	390.15	1.143	1.461	3.274
7	1.011	14.070	1627.04	1.085	1.318	2.810
8	1.011	13.611	2392.20	1.080	1.297	3.032
Average		9.786				

1. $\widehat{ISNR}_{ij}(t)$ - The estimator (defined in Appendix) of the improvement in the signal to noise ratio $ISNR_{ij}(t)$ of the proposed $\vartheta^*_{ij}(t)$ to the simple average $p_{ij\cdot}(t)$ at electrode *i* at the *t*-th sample for subject *j*.

proposed estimator to the simple average. The improvement in SNR is about 10 times on average.

Tables VI to VIII show another advantage of the proposed estimator when the trial number is small. Ten trials are used to create the statistics there. Comparing with the corresponding tables III to V where 24 trials are used, we see that the percentage of information ($a^*_{ij}(t)$) borrowed from all relative sources is increased by 5 percent in average (tables VI), the relative improvement in mean squared error $RIMSE_{ij}(t)$ of the proposed $\vartheta^*_{ij}(t)$ to the simple average $p_{ij\cdot}(t)$ is increased to 43 percent (tables VII), and the improvement in SNR of the proposed $\vartheta^*_{ij}(t)$ to the simple average $p_{ij\cdot}(t)$ is increased to 43 times in average (Tables VIII). This represents one of the basic motivations for using the proposed multivariate composite estimator, namely, that we can improve the quality of the estimation of the features for each subject by borrowing information from all relative sources when the sample size is small for each subject.

Table VI. Statistics of $\{\hat{a}_{ij}(t), i = 1,..., 31; t = 1,..., 127\}$ for each subject by using 10 trials.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	0	0.504	1.000	0.207	0.462	0.820
2	0	0.534	1.000	0.259	0.517	0.832
3	0	0.416	1.000	0.133	0.323	0.687
4	0	0.402	0.999	0.147	0.306	0.629
5	0	0.323	0.999	0.050	0.179	0.563
6	0	0.430	1.000	0.141	0.350	0.713
7	0	0.395	1.000	0.111	0.298	0.666
8	0	0.417	1.000	0.146	0.324	0.670
Average		0.428				

Table VII. Statistics of $\{\widehat{RIMSE}_{ij}(t), i = 1,..., 31; t = 1,..., 127\}$ for each subject by using 10 trials.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	0	0.504	0.999	0.207	0.461	0.819
2	0	0.533	0.999	0.259	0.516	0.832
3	0	0.415	0.999	0.133	0.322	0.686
4	0	0.402	0.999	0.147	0.306	0.629
5	0	0.323	0.999	0.050	0.179	0.562
6	0	0.429	0.999	0.141	0.350	0.712
7	0	0.394	0.999	0.111	0.298	0.665
8	0	0.417	0.999	0.146	0.324	0.670
Average		0.428				

Table VIII. Statistics of $\{\hat{ISNR}_{ij}(t),\ i = 1,\ldots, 31;\ t = 1,\ldots, 127\}$ for each subject by using 10 trials.

SUBJECT	MIN	MEAN	MAX	P25	P50	P75
1	1.000	75.233	15959.88	1.262	1.858	5.554
2	1.000	61.001	11437.71	1.350	2.070	5.964
3	1.000	25.056	6873.56	1.154	1.475	3.193
4	1.000	26.756	9174.36	1.172	1.441	2.700
5	1.000	20.994	6149.07	1.053	1.218	2.288
6	1.000	39.804	8492.28	1.164	1.540	3.480
7	1.000	34.412	8922.55	1.125	1.425	2.992
8	1.000	63.633	16180.99	1.172	1.479	3.036
Average		43.361				

Appendix

We construct the estimators of V_j, μ^* and a_{ij}^* as follows. Define

$$\hat{V}_j = \frac{1}{m-1} \sum_{k=1}^{m} (P_{jk} - P_{j\cdot})(P_{jk} - P_{j\cdot})^{\tau}$$

$$\hat{\mu} = \Big(\sum_{j=1}^{s} \hat{V}_j^{-1}\Big)^{-1} \sum_{j=1}^{s} \hat{V}_j^{-1} P_{j\cdot}$$

$$\hat{a}_{ij} = \frac{e_i^{\tau}\big(\hat{V}_j - (\sum_{l=1}^{s} \hat{V}_l^{-1})^{-1}\big)\, e_i}{e_i^{\tau}\big(\hat{V}_j - (\sum_{l=1}^{s} \hat{V}_l^{-1})^{-1}\big)\, e_i + m\,(p_{ij\cdot} - e_i^{\tau}\hat{\mu})^2}$$

It is easy to see that $\hat{V}_j$, $\hat{\mu}$ and $\hat{a}_{ij}$ converge to V_j, μ^* and a_{ij}^*, respectively, as m is large enough. We propose the following estimator of ϑ_{ij}^*:

$$\hat{\vartheta}_{ij} = (1 - \hat{a}_{ij})\, p_{ij\cdot} + \hat{a}_{ij}\, e_i^{\tau} \hat{\mu}$$

and suggest the use of $\hat{\mu}$ as an estimator of the grand average. For the relative improvement in MSE ($RIMSE_{ij}$) and the improvement in SNR ($ISNR_{ij}$) of ϑ_{ij}^* to the simple average $p_{ij\cdot}$, we have from (2.6)- (2.8) their corresponding estimators

$$\hat{RIMSE}_{ij} = \frac{\hat{a}_{ij}\, e_i^{\tau}\big(\hat{V}_j - (\sum_{l=1}^{s} \hat{V}_l^{-1})^{-1}\big)\, e_i}{e_i^{\tau} \hat{V}_j\, e_i}$$

and

$$\hat{ISNR}_{ij} = \frac{1}{1 - \hat{RIMSE}_{ij}}$$

References

Begleiter, H., Porjesz, B. and Wang, W. A neurophysiologic correlate of visual short-term memory in humans. Electroenceph. Clin. Neurophysiol., 1993, 87: 46-53.

Gevins, A.S., Morgan, N., Bressler, S., Doyle, J. and Cutillo, B. Improved ERP estimation via statistical pattern recognition, Electroenceph. Clin. Neurophysiol., 1986, 64: 177-186.

Gevins, A.S. Correlation analysis. In: A.S. Gevins and A. Remond (Eds.), Methods of Analysis of Brain Electrical and Magnetic Signals. Elsevier, Amsterdam - New York - Oxford, 1987: 171-193.

Gasser, T., Mocks, J. and Verleger, R. A method to deal with trial- to-trial variability of evoked potentials. Electroenceph. Clin. Neurophysiol., 1983, 55: 717-723.

Lahiri, P. and Wang, Wenyu. A multivariate procedure towards composite estimation of consumer expenditures for the U.S. Consumer Price Index Numbers, Survey Methodology. 1992, 18: 2, 279-292.

McGillem, C.D. and Aunon, J.I. Measurements of signal components in single visually evoked brain potentials, IEEE Trans. Biomed. Engng., 1977, 24: 232-241.

McGillem, C.D. and Aunon, J.I. Analysis of event-related potentials. In: A.S. Gevins and A. Remond (Eds.), Methods of Analysis of Brain Electrical and Magnetic Signals. Elsevier, Amsterdam - New York - Oxford, 1987: 131-169.

Woody, C. Characterization of an adaptive filter for the analysis of variable components. Med. Biol. Engng., 1967, 5: 539-553.