Trilinear Modeling of Event-Related Potentials

Kongming Wang*, Henri Begleiter*, and Bernice Porjesz*

Summary: This paper describes a method for estimating a set of spatial components (brain maps) and temporal components (waveforms) of brain potentials. These components play the role of bases of a coordinate system, in the sense that the brain potentials of any subject can be represented as superpositions of these components. The representation is unique given the spatial and temporal components, and this decomposition is particularly appealing for comparing the brain potentials of different subjects (say alcoholics and controls). It can also be used for single trial modeling, clinical classification of patients, and data filtering. The method is based on the topographic component model (TCM, Möcks 1988) which models brain potentials in a trilinear form. We extend the TCM in two aspects. First, the diagonal amplitude matrix is replaced by a general loading matrix based on some neurophysiological considerations. Secondly, the number of spatial components and the number of temporal components can be different. The spatial components and temporal components are obtained respectively by performing singular value decomposition (SVD). This method is illustrated with visual P3 data.

Key words: Brain mapping; Event-related potentials; Topographic component model; Trilinear modeling.

Introduction

Brain potentials are often considered to be a combination of independent waveforms (temporal components) and each of the waveforms is produced by a spatial source distribution (spatial component). An ideal method would be able to identify these components so that the data can be interpreted neurophysiologically. Suppose that we knew *a priori* all possible waveforms and source distributions which are "common" to all existing combinations of subjects and conditions, then it is straightforward to fit a given data set to these waveforms and source distributions in order to identify the active source areas and their waveforms.

In practice we do not have the *a priori* knowledge and many methods have been proposed for estimating these components. An incomplete list includes principal component analysis (PCA; John et al. 1964; Glaser and Ruchkin 1976; Donchin and Heffley 1978), singular value decomposition (SVD; Harner 1990), independent component analysis (ICA; Bell and Sejnowski 1995; Makeig et al. 1996), topographic component model (TCM; Möcks 1988; Field and Graupe 1990 and 1991; Achim and Bouchard 1997), and wavelets (Meyer 1993; Samar et al. 1995; Samar et al. 1996; Wang et al. 1998).

PCA, SVD, and ICA estimate the individual spatial and temporal components for a given subject and a given condition, and it is not obvious how to compare the ERPs of different subjects or different conditions. The one-to-one correspondence between spatial and temporal components may not be fulfilled in many experiments, because it is unreasonable to expect the activities of individual neural "generators" to be mutually orthogonal (Field and Graupe 1991). Furthermore, these decompositions are nonunique and their spatiotemporal interpretations should be viewed with caution. Wavelet approach is a time-frequency analysis tool and it does not have a spatiotemporal interpretation in the neurophysiological context.

The method for estimating "common" spatial and temporal components is the TCM, a trilinear model. The difference between TCM and SVD is that the TCM components are linearly independent but not orthogonal. On one hand, this overcomes the drawback of SVD that orthogonal source distributions produce orthogonal waveforms. On the other hand, since the components are correlated, it is impossible to order the components according to the variance that they can explain. This makes it difficult to determine the order of the TCM model (number of the components). If the model order is given, the least squares solution of TCM modeling is unique. Therefore, a TCM solution should be neurophysiologically in-

* Neurodynamics Laboratory, Department of Psychiatry, SUNY Health Science Center at Brooklyn, Brooklyn, NY, USA.

Accepted for publication: February 21, 2000.

We thank B. Beckrich, E. Iskander, V. Kotlyarevsky, S. Valentini, and A. Stimus for providing the ERP data, D. Chorlian for discussions, and others for suggesting improvements of the manuscript. This research is supported by N.I.H. grants AA 05524 and AA 02686 to HB.

Correspondence and reprint requests should be addressed to Dr. Kongming Wang, Neurodynamics Laboratory, Department of Psychiatry, SUNY Health Science Center at Brooklyn, 450 Clarkson Avenue, Brooklyn, NY 11203, USA.

terpretable if the number of components is correct. For biophysical considerations in building the TCM model, see Möcks (1988).

The present study aims to extend the TCM model to a general trilinear modeling in order to overcome the drawbacks. First, the components are orthogonal as in SVD and they are ordered according to the variance that they can explain. This provides an easy way for selecting components. Secondly, the components are common to all subjects as in TCM and they play the role of the bases of a coordinate system, in the sense that the brain potentials of any subject can be represented as superpositions of these components. The inter-individual comparison study can be carried out on the coefficients (loading matrix) of the representations. Thirdly, the one-to-one correspondence in TCM model between spatial and temporal components is relaxed so that the activity of a spatial component (source distribution) could be captured from more than one orthogonal temporal components. This requires that the loading matrix is not diagonal. Finally, the number of temporal components and the number of spatial components are allowed to be different, because these numbers depend on the length of the signals and the number of electrodes respectively. Therefore, the new model would possess the advantages of both SVD and TCM. The components are rotated so that the average loading matrix of all subjects is diagonal. The components are then uniquely determined (up to sign change) if any two eigenvalues of the average loading matrix are different.

We will construct the spatial components and temporal components by performing singular value decomposition (SVD) on properly concatenated data of many subjects. Details are described in the next section. The number of the components depends on the size of the concatenated data. In general, more components are needed for accounting the same level of variance if the concatenated data has larger size. The question is how many subjects are needed to construct a complete set of components in the sense that the data from any subject, either used in the construction of the components or not, can be well-represented as superpositions of these components. We checked this by constructing components from the data of 10, 20, and 69 subjects respectively. It shows that the major components (the first few components which explain the most of the variance) obtained from the three runs are highly correlated (see discussion section). Therefore, a complete set of components could be constructed from the data of reasonable number of subjects (say 20).

The trilinear modeling can be used for interindividual comparison study, single trial modeling, clinical classification of patients, and data filtering. Application of this approach is illustrated with visual P3 data recorded from 30 alcoholics and 39 controls with 61 electrodes.

Method

The repeated single trials of a subject under a given experimental condition are averaged to eliminate the background EEG and to improve the signal to noise ratio (SNR). These average ERPs then go through a bandpass filter (low-pass filter of 16 Hz in our visual P3 experiment) to further improve the SNR. Let X_i be the filtered ERP of subject i, where each row of X_i is the measurements from an electrode at consecutive times and where each column is the measurements from all electrodes at a given time. If there are more than one experimental conditions (e.g., target, nontarget, and novel), then each row of X_i has more than one epoch corresponding to the experimental conditions. The trilinear modeling is to represent the data X_i as

$$X_i = BA_iC \tag{1}$$

where each column of B is a spatial component and each row of C is a temporal component. The matrix A_i is the loading matrix of subject i. Once B and C are available, the loading matrix for each subject is obtained as the least squares solution for fitting the data X_i. With orthogonal and normalized spatial and temporal components, the loading matrix $A_i=(a_i(j,k))$ of subject i is given by

$$a_i(j,k) = b_j' X_i c_k'$$

where b_j is the j^{th} column of B and c_k is the k^{th} row of C. Then formula (1) can be written as

$$X_i = \sum_j \sum_k a_i(j,k) b_j c_k. \tag{2}$$

Estimation of the temporal components

For estimating temporal components, each time point is a variable and each electrode produces an observation. The concatenation of the individual data matrix X_i by variables gives the combined data matrix

$$D = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \end{pmatrix}$$

SVD decomposition of the data D results in

$$D = USV$$

where S is a diagonal matrix and where the rows of V are orthogonal temporal components. The scaled diagonal elements of S, scaled by the total variance (the summation of the diagonal elements of S), give the ratios of the variances explained by the temporal components respectively. If each of the first K ratios are at least 0.5% and the

rest of the ratios are less than 0.5%, the first K rows of V will be chosen as the temporal components. The rest will be considered as negligible components or noise.

An alternative method for selecting temporal components is to choose enough components in order to explain a specified percentage of the total variance. If the specified percentage is higher (say 90% or 95%), some selected component explains only a very small percentage of the total variance.

Estimation of the spatial components

For estimating spatial components, each electrode is considered to be a variable, and each time point produces an observation. In this case, the concatenation of the individual data matrix X_i by variables gives the combined data matrix

$$D_s=(X_1, X_2,...).$$

Then SVD decomposition of D_s results in

$$D_s = U_s S_s V_s.$$

As in the estimation of the temporal components, we choose the first K_s columns of U_s as spatial components that each of these components explains at least 1% of the spatial variance. The rest columns of U_s will be considered as negligible components or noise. We have chosen the cut point at 1% since the number of variables (the number of electrodes) is much less in spatial case.

Rotation

With the estimated spatial and temporal components (the first K_s columns of U_s and the first K rows of V), the data matrix X_i of subject i can be decomposed as in equation 1. This representation is unique because of the orthogonality of the components. For subsequent analysis, it is desirable that the loading matrix A_i has as few significant entries as possible. This is also neurophysiologically meaningful because it leads to the interpretation that the activity of a spatial component (source distribution) is mainly captured by one temporal component, though other temporal components also capture a small part of the activity of this spatial component. To this end, we compute the average loading matrix A of all subjects $\left(A = \sum_{i=1}^{M} A_i / M \text{ if we have } M \text{ subjects}\right)$ and its SVD decomposition $A=PRQ$. Since R is diagonal, the best rotation matrix for rotating temporal components is Q and the best rotation matrix for rotating spatial components is P. Multiplying the first K_s columns of U_s by P and the first K rows of V by Q, we obtain the common spatial components B and the common temporal components C of the formula (1) respectively. The rotation idea comes from PCA.

Experiment

Data collection

In this visual P3 experiment, subjects are presented with 280 visual stimuli with a uniform inter-stimulus interval of 1.6 seconds. There are 210 non-target stimuli in the shape of an outline of a square, 35 target stimuli in the shape of an X, and 35 novel stimuli, each a different colored polygon or other geometrical figure. So there are three experimental conditions. Subjects are instructed to respond to target stimuli by pressing a button. The visual P3 data are recorded from 64 channels (61 channels of International 10/20 system, two channels for monitoring eye movements, and the nose channel for reference). Therefore, the data to be analyzed have 61 channels. The sampling interval is 3.906 msec and 205 sample points (800 msec signal) after each stimulus are recorded for each condition. After amplification by a factor of 10,000, artifact threshold is set at ± 73.3 microvolts. Any trial with a value above the threshold is rejected, and the other trials are averaged to obtain the averaged ERP with improved SNR. The number of trials used in the averaging is different for different subjects and experimental conditions due to the artifact thresholding. The average ERP is then filtered with a low-pass filter of 16hz.

The trilinear model

The TCM model is built up with data from all 69 subjects (30 alcoholics and 39 control subjects). As described in the last section, the SVD decomposition is performed on the concatenated data matrix D (concatenated according to record time) of size 4209 by 615, yielding 43 temporal components which account for 90% of the temporal variation. Similarly, the SVD decomposition is performed on the concatenated data matrix D_s (concatenated according to record channel) of size 61 by 42435, yielding 16 spatial components which account for 80% of the spatial variation. These components are then rotated such that the loading matrices for all subjects have as few significant entries as possible. These components will be plotted in figure 1.

The fitting error in representing the data of the 69 subjects with these components is 0.5025 microvolts, computed by

$$\text{fitting error} = \text{mean}\left\{\left|X_{ij}(t) - \hat{X}_{ij}(t)\right|\right\}$$

where $X_{ij}(t)$ is the filtered average ERP from subject i, channel j, and time t. The $\hat{X}_{ij}(t)$ is the estimated values of $X_{ij}(t)$ based on the components.

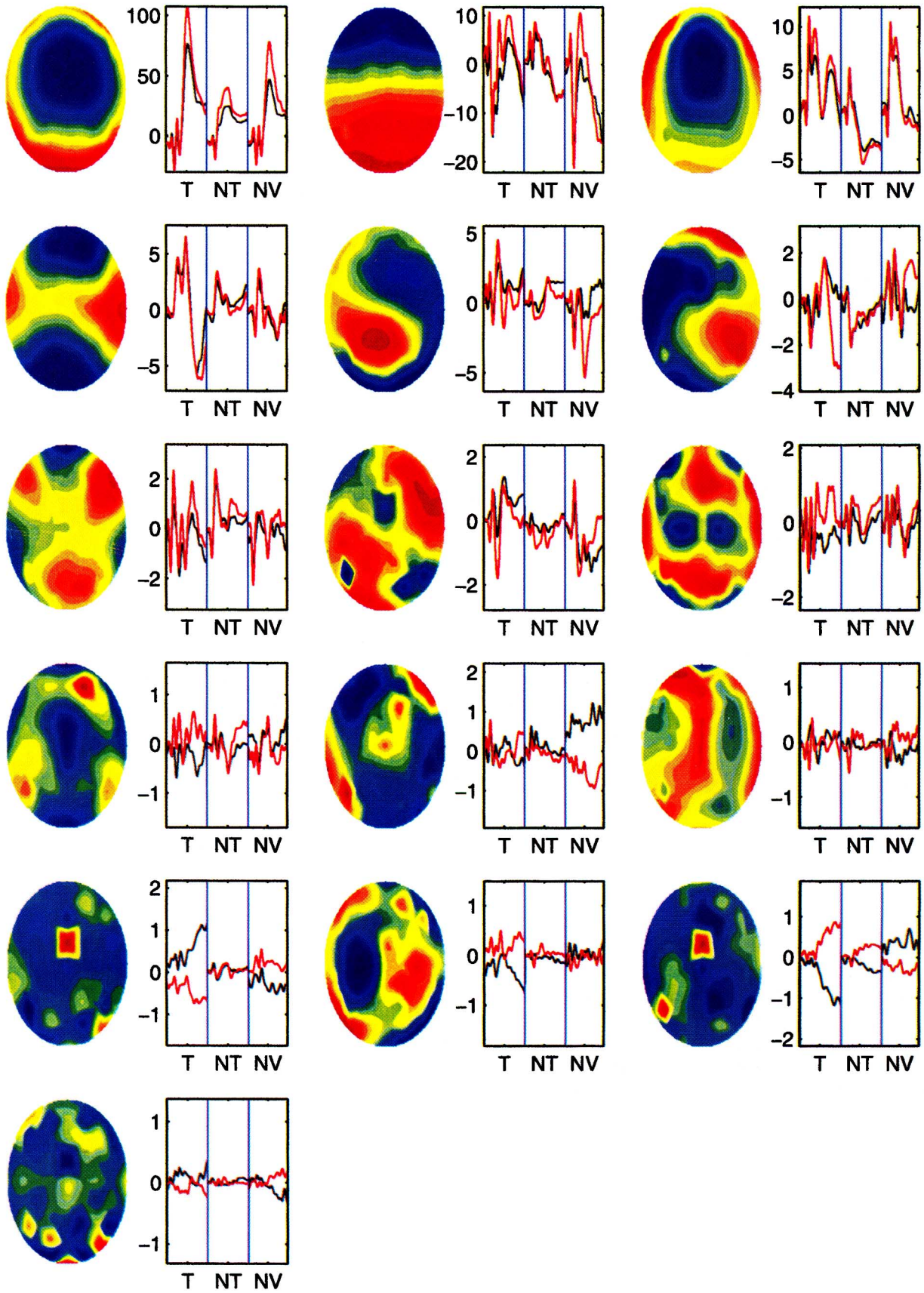

Figure 1. The 16 spatial components obtained from the data of 69 subjects (30 alcoholics and 39 controls). The average time courses (red for controls, black for alcoholics) are plotted next to their corresponding spatial components (source distributions). Three experimental conditions (T for target, NT for not-target, NV for novel) are separated by vertical lines.

Time courses of the spatial components (source distributions)

The trilinear representation (2) can be written as

$$X_i = \sum_j b_j T_{ij}$$

where $T_{ij} = \sum_k a_i(j,k)c_k$ is the time course of the source distribution b_j for subject i. Let T_{nj} denote the average time course of the source distribution b_j for alcoholic subjects and T_{cj} the average time course for control subjects. Figure 1 plots the source distributions and their time courses.

The average time courses of 30 alcoholic subjects (black line) and 39 control subjects (red line) are plotted next to the corresponding spatial component. Three experimental conditions (T for target, NT for not-target, NV for novel) are separated by vertical lines. We have 800 msec signal for each condition. Since the spatial components are all of norm one (orthogonal components), the color only indicates the source area (red) and sink area (blue). The strength of a source distribution is shown by its time course.

The first spatial component (top left) is the source distribution producing the P3 component. The P3 amplitude of control subjects is larger than the P3 amplitude of alcoholic subjects. The second spatial component (top middle) is the source distribution producing the N1 component. The third spatial component (top right) is relatively active. The strength of other spatial components is less than 5% of the strength of the first spatial component.

Figure 1 shows that the P3 component is generated mainly by the first spatial component (source distribution). For subject i, the ratio $|a_i(1,1)| / \sum_k |a_i(1,k)|$ shows how much the activity of the first spatial component contributes to the P3 component. The ratio is 21.3% for alcoholic subjects and 27.3% for control subjects. The T-test shows that this difference is significant ($p = 0.0035$). This could be a partial interpretation of the lower P3 amplitude for alcoholic subjects. The ratio is less than 30% for both groups, and it confirms that the one-to-one correspondence between orthogonal spatial and temporal components is not fulfilled in this visual P3 experiment.

Source distribution of P3 component

The representation (2) can also be written as

$$X_i = \sum_k u_{ik} c_k$$

Where $u_{ik} = \sum_j a_i(j,k)b_j$ is the source distribution generating the temporal component c_k for subject i. Let U_{n1} denote the average source distribution generating P3 component (first temporal component) for alcoholic subjects and U_{c1} the average source distribution for control subjects. The surface laplacian of U_{n1} and U_{c1} are more localized source distributions. We apply the local polynomial method (Wand and Begleiter 1999) to compute the surface laplacians, which are plotted in figure 2.

Similarly, source distributions can be computed for each temporal component.

Classification of subjects

A clinical application of ERP is to classify a new subject into one of the two groups: alcoholics and controls. The data of a new patient can be represented in the form (1) and the loading matrix can be compared to the loading matrices of the training examples. If the loading matrix is closer to the loading matrices of the alcoholic training examples than to the loading matrices of the control training examples, this patient is classified as an alcoholic patient. Otherwise the patient is classified as a control subject.

If the spatial and temporal components were constructed from 10 training examples (5 alcoholics selected randomly from the 30 alcoholics and 5 controls selected randomly from the 39 controls), the classification error is larger (3 out of the 30 alcoholics were misclassified as controls and 24 out of the 39 controls were misclassified as alcoholics). The large error of classifying controls indicates that the training examples are not representative. If we use 20 training examples (10 alcoholics and 10 controls selected randomly), the error is greatly reduced. The number of misclassified subjects is 4 for both groups. If we use all 69 subjects as training examples, the number of misclassified subjects is still 4 for both groups. It indicates that 20 training examples are enough for the classification purpose, and there are few "outliers" from each group.

Monitoring single trials

If the spatial and temporal components are obtained based on enough subjects, they form a complete basis for decomposing ERPs in the sense that data from new subjects under the same experimental modality can also be represented in this system. In particular, we can use this system to monitor single trials.

Figure 3 illustrates the visual P3 data recorded from an alcoholic subject whose data was not used in obtaining the spatial and temporal components. There are 25 trials under the target condition. Recall that the temporal components have three epochs corresponding to three experimental conditions. The first epoch, which corresponds to the target condition, can be used to monitor the single trials under the target condition. Figure 3 plots two single trials, the modeled ERP data, and noise which represents background EEG. It shows that single trials can be filtered effectively with the temporal and spatial components.

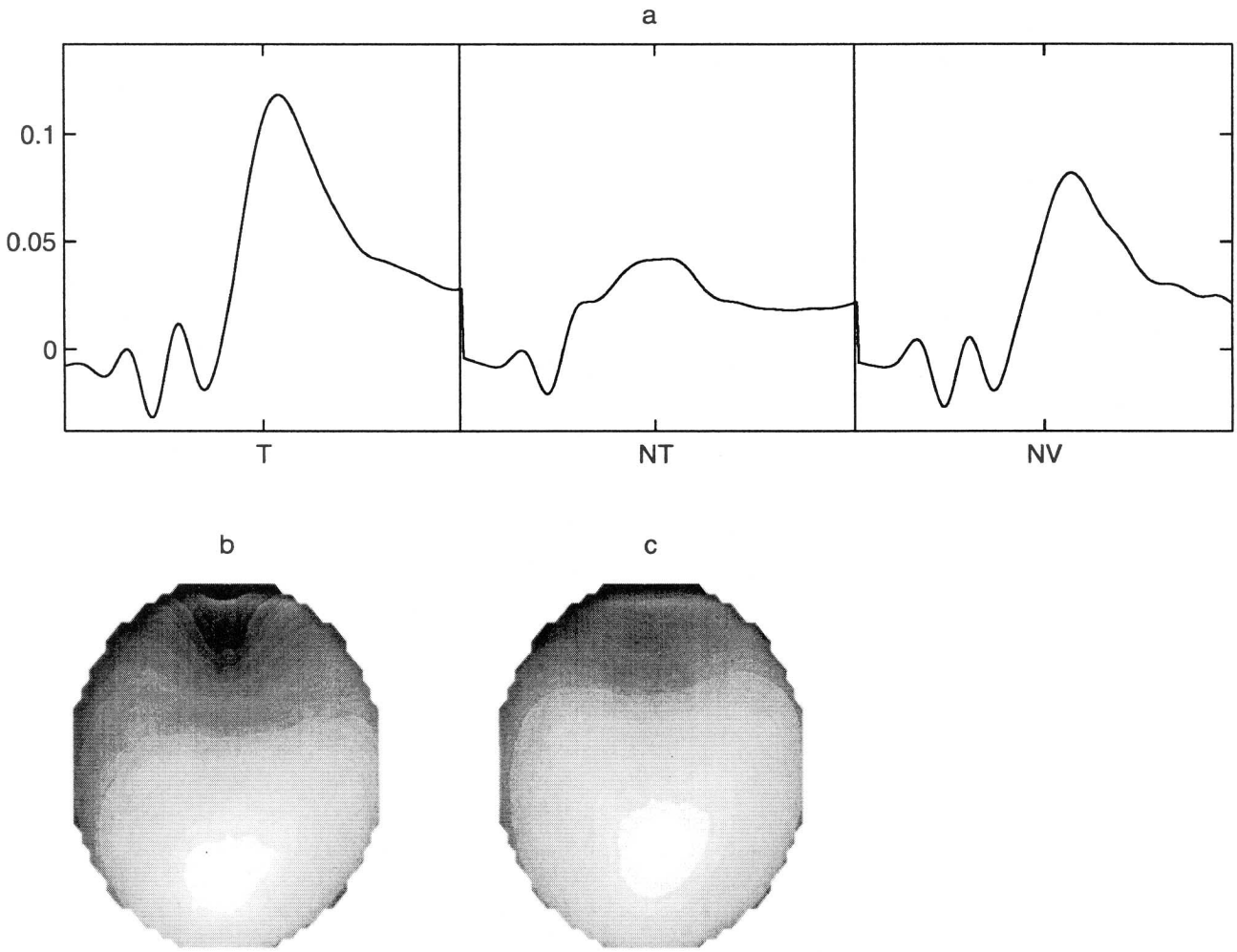

Figure 2. (a) The P3 component (first temporal component). (b) Source distribution (surface Laplacian) generating P3 component for alcoholic subjects. (c) Source distribution (surface Laplacian) generating P3 component for control subjects.

Discussion

The TCM of Möcks (1988) is extended to general trilinear modeling of ERPs. The goal is to estimate "common" spatial and temporal components of ERPs from clinical data. These components should form a complete system in the sense that the ERPs of any subject can be well-represented as superpositions on the common components. An important question is how many subjects should be measured in order to achieve this completeness. We have constructed the components based on the data of 10 (5 alcoholics and 5 controls, randomly selected), 20 (10 alcoholics and 10 controls, randomly selected), and 69 (30 alcoholics and 39 controls) subjects respectively. Based on the data of the 10 selected subjects, we find 16 spatial components that each accounts for at least 1% of the total spatial variance and 30 temporal components that each accounts for at least 0.5% of the total temporal variance. Based on the data of the 20 selected subjects, we find 15 spatial components and 36 temporal components. Based on the data of all the 69 subjects, we find 16 spatial components and 43 temporal components. If the components are stable in the sense that the first few components are about the same no matter how many subjects are used in constructing the components, then these three runs with 10, 20, and 69 subjects indicate that 10 subjects should be enough for fulfilling the completeness. To be on the safe side, we would suggest that data from at least 20 subjects should be used in constructing the components in order to offset the effects of outliers, as indicated by the classifications in a previous section. Figure 4 plots the pairwise correlations between the corresponding components of different runs. The high correlations show that the components are indeed stable.

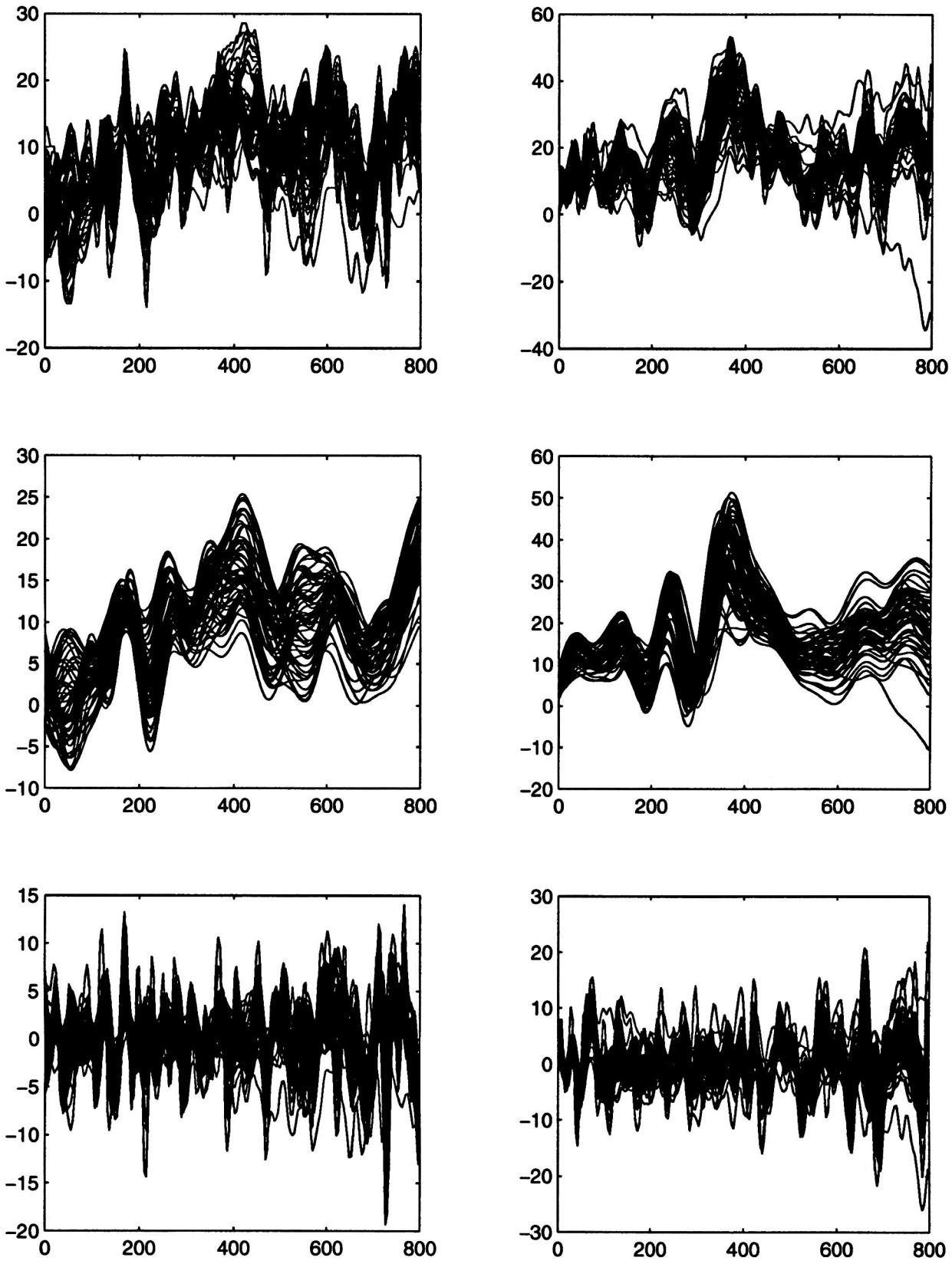

Figure 3. Top row: two single trials evoked by target stimulus. Middle row: the two trials modeled with the 16 spatial components and the 43 temporal components. Bottom row: noise or background EEG. Each subgraph plots 61 signals from 61 channels. The horizontal unit is in msec and the vertical unit is in microvolts.

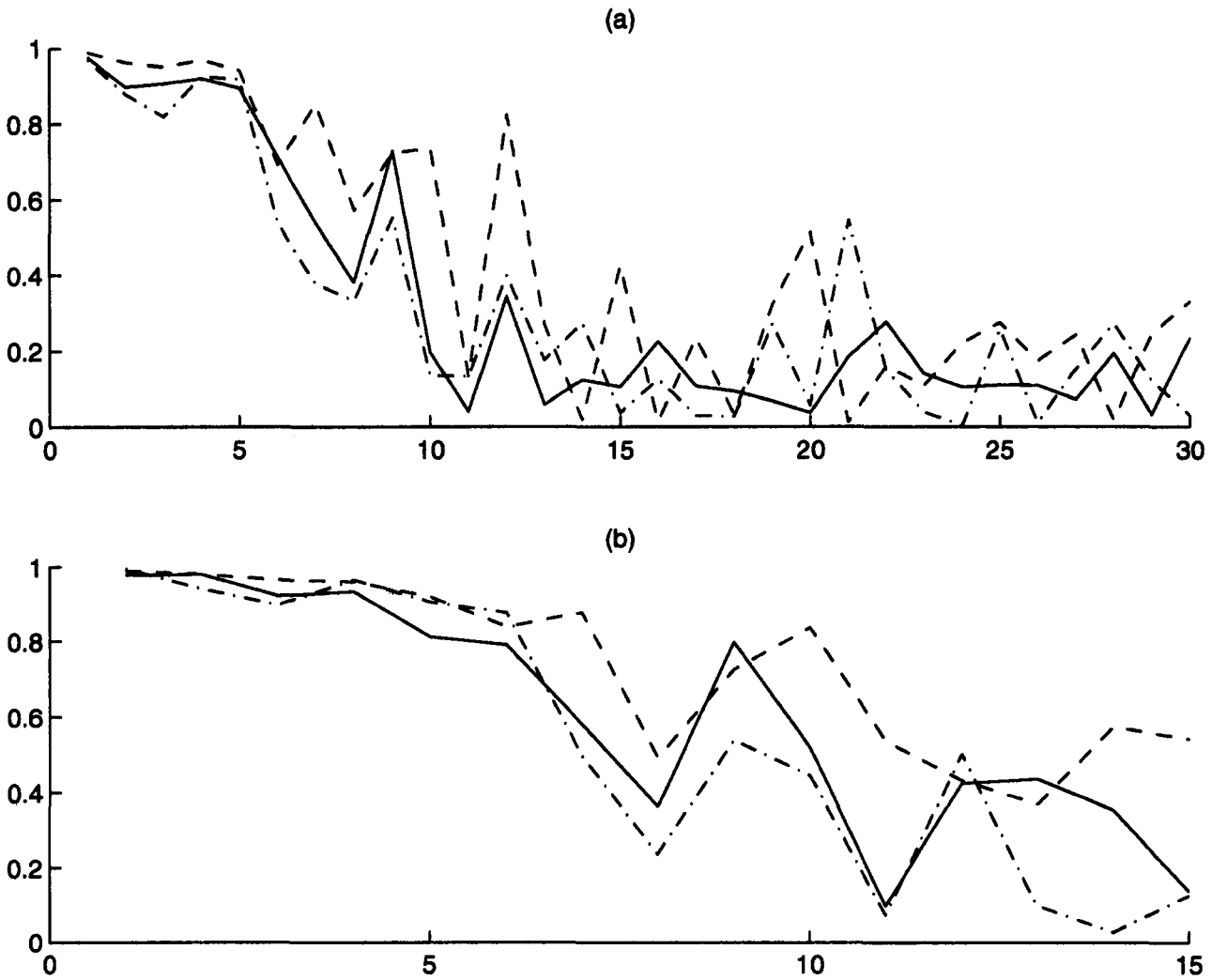

Figure 4. (a) The pairwise correlations between the temporal components obtained by using the data of different subjects. Solid line: correlations between the temporal components obtained by using 10 and 69 subjects; dashed line: correlations between the temporal components obtained by using 20 and 69 subjects; dash-dotted line: correlations between the temporal components obtained by using 10 and 20 subjects. (b) The pairwise correlations between the spatial components obtained by using the data of different subjects. Same interpretation of line types as in (a).

We have demonstrated that the trilinear modeling can be used to analyze ERPs in many ways. Data is reduced to a small portion of the original size (2% in our example). The inter-groups comparison analysis and patient classification can be performed on this reduced data set because the components are common to all subjects. It also helps to interpret the ERPs. The loading matrix indicates which source distributions are active and which wave forms capture their activities. In addition, it can be used as a filter to extract the ERPs from single trials. Signals that can not be interpreted by these components, even low frequency signals, are considered as noise or background EEG. Single trial modeling is important and difficult.

We have applied SVD to properly concatenated data to construct the components. This means that the components are uncorrelated but not statistically independent. A better way of finding common components would decompose the concatenated data into independent components in a unique way. ICA would be a good alternative if we knew the exact number of spatial components and the number of the temporal components. When applying ICA, one also runs into the problem of not having enough data for obtaining spatially independent components. Even one has data of 128 channels, it is still insufficient to apply any existing method to obtain more than 10 spatially independent components. This means that it is very difficult to separate 10 independent distributions from their mixture if one has only 128 observations of the mixture.

Some parametric methods have been proposed to implement the TCM of Möcks (1988). A dipole component model was proposed by Turetsky at al. (1990). The temporal components are modeled by damped sinusoids which tightly constrain the temporal morphology of source activity. The spatial components are determined by the positions and orientations of the underlying sources which are estimated from the data. They modeled the head as a homogeneous spherical conductor. Several issues make it difficult to apply the dipole component model. First, locating the underlying sources with a realistic head model is itself a very difficult task. Secondly, fitting a dipole component model needs to solve a high dimensional minimization problem. Achim and Bouchard (1997) proposed a dynamic version of the TCM which further admits component modulation in time scale to accommodate possible changes in the component expression across conditions. They tested the method with simulated data of only two components and it is not clear how it performs on real data sets.

References

Achim, A. and Bouchard, S. Toward a dynamic topographic components model. Electroenceph. Clin. Neurophys., 1997, 103: 381-385.

Bell, A. and Sejnowski, T. An information-maximization approach to blind separation and blind deconvolution. Neural Comp., 1995, 7: 1129-1159.

Donchin, E. and Heffley, E. Multivariate analysis of event-related potential data: A tutorial review. In: D. Otto (Ed.), Multidisciplinary perspectives in event-related brain potential research, U.S. Gov. Printing Office, Washington DC, 1978: 555-572.

Field, S. and Graupe, D. Topographic components analysis of evoked potentials: estimation of model parameters and evaluation of parameter uniqueness. J. Biomed. Eng., 1990, 12: 287-300.

Field, S. and Graupe, D. Topographic component (parallel factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decomposition. Brain Topography, 1991, 3: 407-423.

Glaser, E. and Ruchkin, D. Principles of neurobiological signal analysis. Academic Press, New York, 1976: 233:290.

Harner, R. Singular value decomposition - a general linear model for analysis of multivariate structure in the electroencephalogram. Brain Topography, 1990, 3: 43-47.

John, E., Ruchkin, D. and Villegas, J. Experimental background: signal analysis and behavioral correlations of evoked potential configuration in cats. Ann. N.Y. Acad. Sci., 1964, 112: 362-420.

Makeig, S., Bell, A., Jung, T. and Sejnowski, T. Independent component analysis of electroencephalographic data. In: D. Touretzky, M. Mozer and M. Hasselmo (Eds.),Advances in neural info. proc. sys. 8, MIT Press, Cambridge MA, 1996: 145-151.

Meyer, Y. Wavelets Algorithms & Applications. SIAM, Philadelphia, 1993.

Möcks, J. Topographic components model for event-related potentials and some biophysical considerations. IEEE Trans. Biomed. Eng., 1988, 35: 482-484.

Samar, V., Begleiter, H., Chapa, J., Raghuveer, M., Orlando, M. and Chorlian, D. Matched Meyer neural wavelets for clinical and experimental analysis of auditory and visual evoked potentials. Proceeding of the VIII European signal processing conference (EUSIPCO-96), Trieste, Sept. 1996: 10-13.

Samar, V., Swartz, K. and Raghuveer, M. Multiresolution analysis of event-related potentials by wavelet decomposition. Brain and Cognition, 1995, 27: 398-438.

Turetsky, B., Raz, J. and Fein, G. Representation of multichannel evoked potential data using a dipole component model of intracranial generators: application to the auditory P300. Electroenceph. Clin. Neurophys., 1990, 76: 540-556.

Wang, K. and Begleiter, H. Local polynomial estimate of surface Laplacian. Brain Topography, 1999, 12: 19-29.

Wang, K., Begleiter, H. and Porjesz, B. Spatial enhancement of event-related potentials using multiresolution analysis. Brain Topography, 1998, 10: 191-200.