ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 22 (2004) 1035–1045

Decomposing EEG data into space–time–frequency components using Parallel Factor Analysis

Fumikazu Miwakeichi,[a,*] Eduardo Martínez-Montes,[b] Pedro A. Valdés-Sosa,[b] Nobuaki Nishiyama,[a] Hiroaki Mizuhara,[a] and Yoko Yamaguchi[a]

[a] *Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute, Saitama 351-0198, Japan*
[b] *Neuroscience Department, Cuban Neuroscience Center, Habana, Cuba*

Received 17 July 2003; revised 12 March 2004; accepted 17 March 2004

Finding the means to efficiently summarize electroencephalographic data has been a long-standing problem in electrophysiology. A popular approach is identification of component modes on the basis of the time-varying spectrum of multichannel EEG recordings—in other words, a space/frequency/time atomic decomposition of the time-varying EEG spectrum. Previous work has been limited to only two of these dimensions. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) have been used to create space/time decompositions; suffering an inherent lack of uniqueness that is overcome only by imposing constraints of orthogonality or independence of atoms. Conventional frequency/time decompositions ignore the spatial aspects of the EEG. Framing of the data being as a three-way array indexed by channel, frequency, and time allows the application of a unique decomposition that is known as Parallel Factor Analysis (PARAFAC). Each atom is the tri-linear decomposition into a spatial, spectral, and temporal signature. We applied this decomposition to the EEG recordings of five subjects during the resting state and during mental arithmetic. Common to all subjects were two atoms with spectral signatures whose peaks were in the theta and alpha range. These signatures were modulated by physiological state, increasing during the resting stage for alpha and during mental arithmetic for theta. Furthermore, we describe a new method (Source Spectra Imaging or SSI) to estimate the location of electric current sources from the EEG spectrum. The topography of the theta atom is frontal and the maximum of the corresponding SSI solution is in the anterior frontal cortex. The topography of the alpha atom is occipital with maximum of the SSI solution in the visual cortex. We show that the proposed decomposition can be used to search for activity with a given spectral and topographic profile in new recordings, and that the method may be useful for artifact recognition and removal.

Keywords: Parallel Factor Analysis; EEG space/frequency/time decomposition; Principal Component Analysis; Multiway analysis; Source Spectra Imaging

Introduction

The electroencephalogram (EEG) is the reflection upon the scalp of the summed synaptic potentials of millions of neurons (Lopes da Silva, 1987). Most investigators agree (Lachaux et al., 1999; Varela et al., 2001) that these neurons self-organize into transient networks (''neural masses'') that synchronize in time and space to produce a mixture of short bursts of oscillations that are observable in the EEG record. A statistical description of the oscillatory phenomena of the EEG was carried out first in the frequency domain (Lopes da Silva, 1987) by estimation of the power spectrum for quasi-stationary segments of data. More recent characterizations of transient oscillations are carried out by estimation of the time-varying (or evolutionary) spectrum in the frequency/time domain (Dahlhaus, 1997). These evolutionary spectra of EEG oscillations will have a topographic distribution on the sensors that is contingent on the spatial configuration of the neural sources that generate them as well as the properties of the head as a volume conductor (Nunez, 1993).

The purpose of the present study was to attempt the decomposition of multichannel time-varying EEG spectrum into a series of distinct components or modes. In the parlance of modern harmonic analysis (Chen and Donoho, 2001), we performed a space/frequency/time ''atomic decomposition'' of multidimensional data. In other words, we assume that each neural mass contributes a distinctive atom to the topographic frequency/time description of the EEG, so that the estimation of these atoms is possible by means of signal-processing techniques. Each atom will be defined by its topography, spectral content, and time profile; in other words, by its spatial, spectral, and temporal signatures. We expect that these extracted atoms ultimately will allow the identification of the corresponding neural masses that produce them.

There is a long history of atomic decompositions for the EEG. However, to date, atoms have not been defined by the triplet spatial, spectral, and temporal signatures but rather pairwise combinations of these components. Some of the current procedures for these analyses are reviewed below.

Space/time atoms: PCA and ICA

Space/time atoms are the basis of both Principal Component Analysis (PCA) and Independent Component Analysis (ICA) as

* Corresponding author. Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. Fax: +81-48-467-6938.
E-mail address: miwakel@brain.riken.go.jp (F. Miwakeichi).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.03.039

applied to multichannel EEG. PCA has been used for artifact removal and to extract significant activities in the EEG (Lagerlund et al., 1997; Soong and Koles, 1995). A basic problem is that atoms defined by only two signatures (space and time) are not determined uniquely. In PCA, orthogonality is therefore imposed between the corresponding signatures of different atoms. This, however, is a rather nonphysiological constraint. Even with this restriction, there is the well-known nonuniqueness of PCA that allows the arbitrary choice of rotation of axes (e.g., Varimax and Quartimax rotations). More recently, ICA has become a popular tool for space/time atomic decomposition (Cichocki and Amari, 2002; Hyvarinen et al., 2001). It avoids the arbitrary choice of rotation (Jung et al., 2001). Uniqueness, however, is achieved at the price of imposing a constraint even stronger than orthogonality, namely, statistical independence. In both PCA and ICA, the frequency information may be obtained from the temporal signature of the extracted atoms in a separate step.

Frequency/time atoms: wavelet analysis

There are many papers on the decomposition of single-channel EEG into frequency/time atoms. For this purpose, the Fast Fourier Transformation (FFT) with sliding window (Makeig, 1993) or the wavelet transformation (Bertrand et al., 1994; Tallon-Baudry et al., 1997) have been employed. In fact, any of the frequency/time atomic decompositions currently available (Chen and Donoho, 2001) could, in principle, be used for the EEG. However, these methods do not address the topographic aspects of the EEG time/frequency analysis.

Space/frequency/time atoms: PARAFAC

Gonzalez Andino et al. (2001) improved previous analyses by analyzing regions of the frequency/time plane where a single dipole model is an adequate spatial description of the signal, thus incorporating topographic information. Topographic frequency/time decomposition of the EEG was introduced by Koenig et al. (2001), which is the first work to estimate atoms characterized simultaneously by a frequency/time and spatial signature. In their analyses, it was possible to extract physiologically significant activity in the EEG. However, in order to achieve a unique decomposition, they imposed the mathematical constraints that the combined frequency/time signatures of all atoms were required to be of minimum norm and the spatial or topographic signatures were required to have maximal smoothness. These constraints have been found to be unnecessary for unique topographic time/frequency decomposition, a fact that has motivated the work described in this paper.

It has long been known, especially in the chemometrics literature, that unique multi-linear decompositions of multi-way arrays of data (more than two dimensions) are possible under very weak conditions (Sidiropoulos and Bro, 2000). In fact, this is the basic argument for Parallel Factor Analysis (PARAFAC). This technique was proposed independently by Harshman (1970, 1972) and by Carroll and Chang (1970) who named the model Canonical Decomposition, and recently has been improved by Bro (1998) who also provided a Matlab toolbox (available as of this writing at: http://www.models.kvl.dk/users/rasmus/). In PARAFAC, for three-way arrays, the data is decomposed as a sum of components (corresponding to our "atoms"), each of which is the tri-linear product of one score vector and two loading vectors (corresponding to our "signatures"). The important difference between PARAFAC and techniques such as PCA or ICA is that the decomposition of multi-way data is unique even without additional orthogonality or independence constraints.

Thus, PARAFAC can be employed for a space/frequency/time atomic decomposition of the EEG. This makes use of the fact that multichannel evolutionary spectra are multi-way arrays, indexed by electrode, frequency, and time. The inherent uniqueness of the PARAFAC solution leads to a topographic time/frequency decomposition with a minimum of *a priori* assumptions.

Here, we use PARAFAC for the purpose of simultaneous space/frequency/time decompositions. Previous applications of PARAFAC to EEG data have analyzed only space/time, and some additional external dimensions provided by subject and drug dose, among other factors (Achim and Bouchard, 1997; Estienne et al., 2001; Field and Graupe, 1991). A special interpretation of this model is also the Topographic Components Model (TCM) (Möcks, 1988a,b), which gives justification for the PARAFAC model in the context of evoked potentials analysis, based on biophysical considerations (Möcks, 1988b). In this field, a relevant proof of the use of TCM over PCA using only synthetic noiseless data was given in Achim and Bouchard (1997).

To illustrate the usefulness of PARAFAC, we applied the decomposition of time-varying EEG spectrum to the comparison of resting EEG to that recorded while the subject performed mental arithmetic. Mental arithmetic produces theta activity in the frontal area and a suppression of alpha activity in the occipital area, while the converse occurs when the eyes are closed in the resting condition (Harmony et al., 1999; Ishihara and Yoshii, 1972; Sasaki et al., 1996). The PARAFAC atomic decomposition should be able to extract these components, localize them correctly, and detect the corresponding level of activity in these bands in each physiological state. Once estimated, the spatial and spectral signatures of the identified atoms may be used to search for similar types of activity in new data sets. Here, this procedure will be called "screening" for the presence of an atom.

Our focus is on space/time/frequency decompositions tailored to the description of oscillatory phenomena. These are not the only interesting phenomena in the EEG, transient activity being another example. The methods described in this paper may be generalized to this application by exchanging the basic dictionary that describes oscillations.

This paper is organized as follows. We first describe the experimental methods. Then, we consider the basic theoretical development of the space/frequency/time atomic decomposition

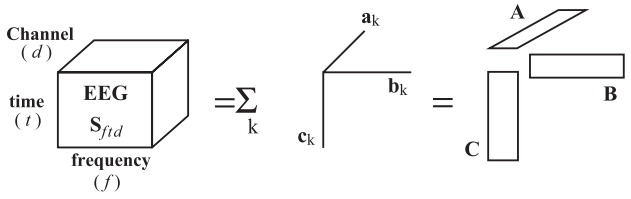

Fig. 1. Graphical explanation of the PARAFAC model. The multichannel EEG evolutionary spectrum **S** is obtained from a channel by channel wavelet transform. **S** is a three-way data array indicated by channel, frequency, and time. PARAFAC decomposes this array into the sum of "atoms". The *k*th atom is the tri-linear product of loading vectors representing spatial ($\mathbf{a}_k$), spectral ($\mathbf{b}_k$), and temporal ($\mathbf{c}_k$) "signatures". Under these conditions, PARAFAC can be summarized as finding the matrices $\mathbf{A} = \{\mathbf{a}_k\}$, $\mathbf{B} = \{\mathbf{b}_k\}$, and $\mathbf{C} = (\mathbf{c}_k)$, which explain **S** with minimal residual error.

and the use of estimated factors to screen for activity in new data segments. The results and discussion follow.

Methods

Data acquisition

Five male right-handed subjects (mean age 25.8 ± 3.96 years) that produced clear theta activities during a mental task were studied in this work. All subjects signed an informed consent form approved by the RIKEN Human Subject Protection Committee before EEG recording. All subjects were required to concentrate, for 3 min, on mental arithmetic (subtraction by serial 7 from 1000) with closed eyes. They were asked the final residual number at the end of the task. The resting EEG with closed eyes was also recorded for comparison. During the recording, we provided no visual nor auditory stimulation for the subjects.

EEG recordings were carried out with a standard 64-channel system (NeuroScan Syn Amps Model 5083) referred to linked earlobes. The EEG data were sampled at 500 Hz and bandpass filtered from 1 to 30 Hz.

Theory

In our application to EEG data, the data matrix $\mathbf{S}_{(N_d \times N_f \times N_t)}$ is the three-way time-varying EEG spectrum array obtained by using the wavelet transformation, where N_d, N_f, and N_t are the number of channels, steps of frequency, and time points, respectively. For the wavelet transformation, a complex Morlet mother function was used (Jensen and Tesche, 2002; Kronland-Martinet and Morlet, 1987; Tallon-Baudry et al., 1997). The energy $S(d,f,t)$ of the channel d at frequency f and time t is given by the squared norm of the convolution of a Morlet wavelet with the EEG signal $v(d,t)$

$$S(d,f,t) = |w(t,f)^{*}v(d,t)|^2, \tag{1}$$

where the complex Morlet wavelet, $w(t,f)$ is defined by $w(t,f) = \sqrt{\pi}\sigma_b exp\left(-\left(\frac{t}{\sigma_b}\right)^2\right) * exp(i2\pi ft)$ with σ_b being the-

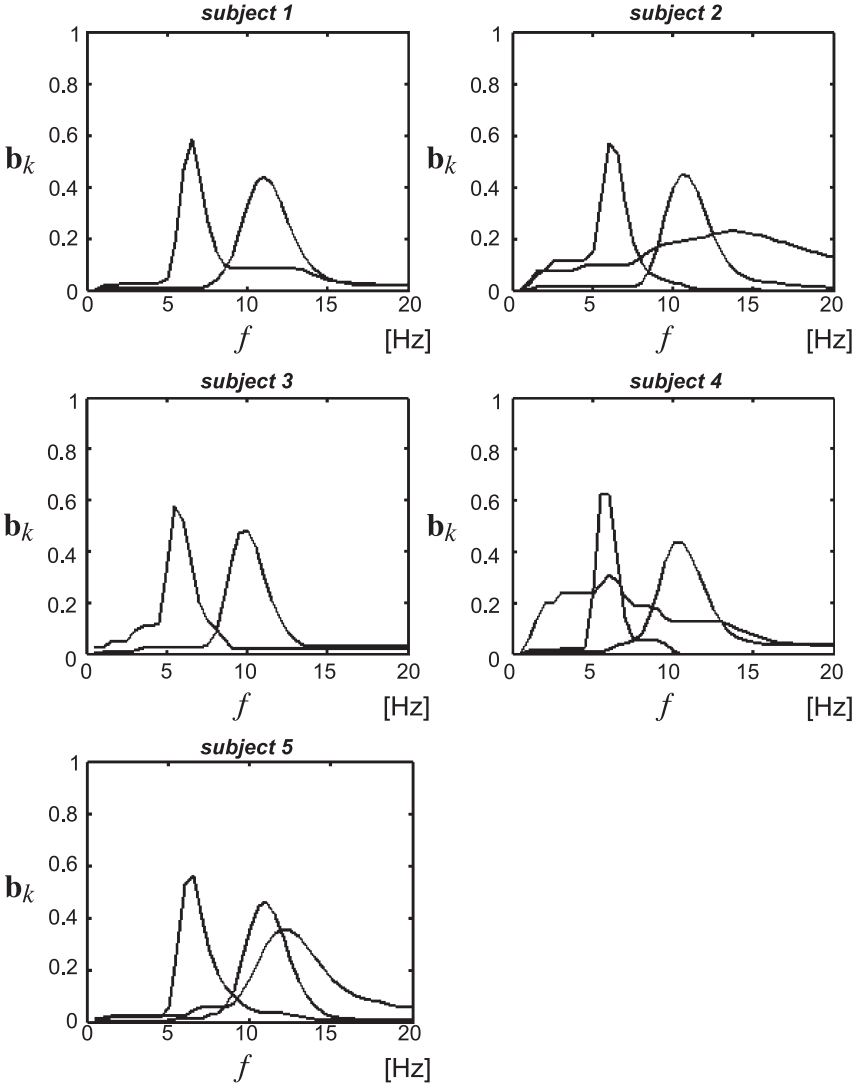

Fig. 2. Spectral signature $\mathbf{b}_k$ of atoms of Parallel Factor Analysis (PARAFAC) for each subject. Note the recurrent appearance of frequency peaks in the theta and alpha bands. The horizontal axis is frequency in Hz and the vertical axis is the normalized amplitude.

bandwidth parameter. The width of the wavelet, $m = 2\pi\sigma_b f$ is set to 7 in this study.

We closely follow here the detailed description of PARAFAC found in Bro (1998). The basic structural model for a PARAFAC decomposition of the data matrix $\mathbf{S}_{(N_d \times N_f \times N_t)}$ of elements S_{dft} is defined as:

$$\hat{S}_{dft} = \sum_{k=1}^{N_k} a_{dk} b_{fk} c_{tk} \tag{2}$$

The problem is to find the loading matrices, **A**, **B** and **C**, whose elements are a_{dk}, b_{fk}, and c_{tk}. In our application, each component k will be designated as an "atom" and the corresponding vectors $\mathbf{a}_k = \{a_{dk}\}$, $\mathbf{b}_k = \{b_{fk}\}$, $\mathbf{c}_k = \{c_{tk}\}$ will be the spatial, spectral, and temporal signatures of each atom (Fig. 1). The uniqueness of the solution is guaranteed when rank **(A)** + rank **(B)** + rank **(C)** $\geq 2N_k + 2$. As can be seen, this is a less-stringent condition than either orthogonality or statistical independence (Sidiropoulos and Bro, 2000). The decomposition (Eq. (2)) is achieved by finding $\min_{a_{dk} b_{fk} c_{tk}} \|\hat{S}_{dft} - \sum_{k=1}^{N_k} a_{dk}\ b_{fk}\ c_{tk}\|$. Since the $\hat{S}_{dft}$ are spectra, this minimization must be carried out under the non-negativity constraint for the loading vectors. This particular variant of PARAFAC has been developed by Bro (1998). PARAFAC produces the vectors $\mathbf{a}_{k(N_d \times 1)}$, which is the kth component loading vector that can be seen as topographical maps, $\mathbf{b}_{k(N_f \times 1)}$ is the spectrum for kth component and $\mathbf{c}_{k(N_t \times 1)}$ is the temporal signature for component k.

The main advantage of this method is that it provides us with a unique decomposition of the time-varying EEG spectrum corresponding to the best model in the least-squares sense. The only indeterminacy in the least-square solution is the order of the atoms and the relative scaling of the signatures. On the other hand, it has also been proved that if the data is approximately tri-linear, then the algorithm will show the true underlying phenomena, if the correct number of components is used and if the signal-to-noise ratio is appropriate (Harshman, 1972; Kruskal, 1976, 1997).

An important point is the selection of the most appropriate number N_k of components. Several methods have been developed for this purpose only (Bro, 1998). The Core Consistency Diagnostic (Corcondia) is an approach that applies especially to PARAFAC models, and has been shown to be a powerful and simple tool for determining the appropriate number of components in multiway models. In this work, we will use not only Corcondia but also the evaluation of systematic variation left in the residuals of the model.

Validation of the method

As described by Harshman (1984) and Bro (1998), the validation of a particular analysis can be seen as part of the analysis itself and can be divided into levels: zero-fit diagnostics (related to data before fitting any model, selection of proper model); one-fit diagnostics (validate the consistency of the model applied); many-fit diagnostics (comparisons between different models, use of statistical inferences on the results). Given some general considerations of the PARAFAC modeling of the time-varying EEG spectrum, we shall make a deeper analysis of the appropriateness of this procedure following these levels and the general guidelines for validating the application of multi-way models given in Bro (1998).

The usual way to assess the multiway (three-way in this case) nature of the data in study is the exploration of results provided by bi-linear analysis of the data. In particular, the application of PCA to the unfolded three-way array will provide a matrix of loadings in which a global behavior can be detected, indicating the existence of dimensional structure in the explored dimensions. For data similar to those treated here, this is clearly shown in Estienne et al. (2001), and with a more complete analysis in Field and Graupe (1991). Another way of assessing the tri-linear structure of the data is by means of the Core Consistency Diagnostic (Corcondia) (Bro, 1998; Estienne et al., 2001). This is a tool provided automatically in the implementation of PARAFAC and other related multi-way algorithms contained in the Matlab Toolbox used here. Corcondia was utilized for successfully demonstrating the presence of multiway structure in our data.

In this work, we have chosen PARAFAC among several multiway models, (e.g., PARAFAC2, PARATUCK2, TUCKER1, TUCKER3) (Bro, 1998). This is an application of Occam's razor as PARAFAC is the simplest and most restricted model. As we only consider it in our analysis, whether other versions of PARAFAC or TUCKER models can give better results in terms of explanation of the systematic variation of the data and interpretability of the results remains an open question.

On the other hand, other drawbacks of the application of PARAFAC model include the need for careful preprocessing of the data and for checking residuals, leverages and other parameters in the search of constant factors, outliers, and degeneracy. We do not detail these problems here, as such analyses appear in the literature, e.g., exhaustive ways of exploring degeneracy and model mis-specifications can be found in Field and Graupe (1991). The Matlab Toolbox used here provided us of these tools for many-fit diagnostics (residuals, leverages, Corcondia, convergence, explained variation).

What is missing in the present study is a rigorous analysis of the uniqueness of the model in our case, but it is well-known that through the use of PARAFAC, uniqueness is almost always present.

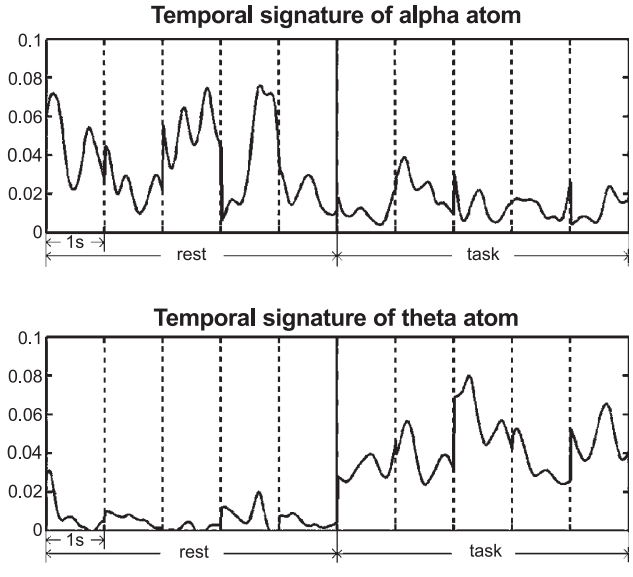

Fig. 3. Temporal signatures, $\mathbf{c}_k$, of theta and alpha atoms of Parallel Factor Analysis (PARAFAC) for a typical subject. The first five segments were chosen randomly from the rest condition; the second five segments were selected so as to contain typical theta bursts. Each segment is 1-s long, containing 100 time frames. The horizontal axis is time t, and the vertical axis is the value of $\mathbf{c}_k$, which is dimensionless.

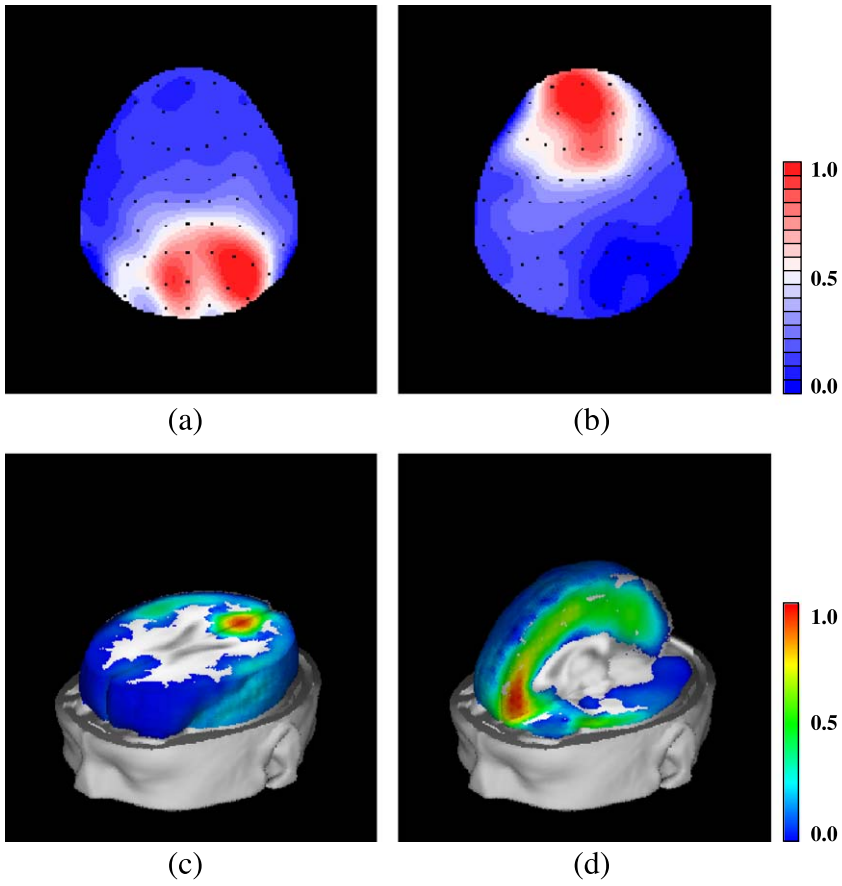

Fig. 4. Spatial signatures, $\mathbf{a}_k$, for the theta and alpha atoms of Parallel Factor Analysis (PARAFAC) of a typical subject. In (a) and (b), each signature is displayed as a topographic map; (c) and (d) are the corresponding Source Spectra Imaging solutions. The cross sections of brain were prepared for better visualization of the maximally activated regions. These are illustrated with a normalized color scale of the magnitude of $\mathbf{J}_k$.

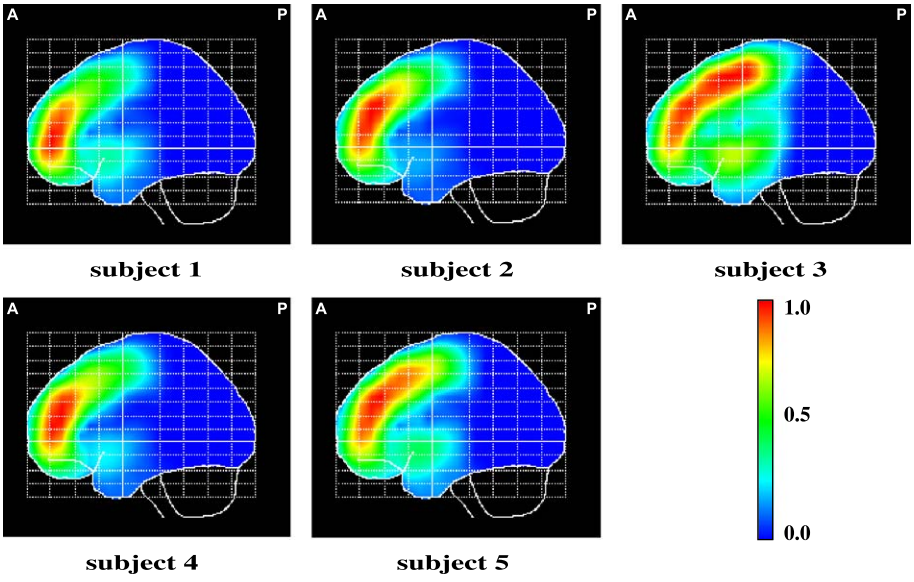

Fig. 5. Maximum-intensity projections of the Source Spectra Imaging solution of the spatial signature, $\mathbf{a}_k$, of the theta atom for each subject.

Although a sufficient condition was given above, usually, uniqueness can be assessed by checking the convergence of the algorithm and the interpretability of the results. A strong sufficient condition, but easier to verify, is that no two loading vectors are linearly dependent.

Screening and artifact detection

PARAFAC can be used not only for extracting significant activities from EEG, but also for searching for the presence of atoms in a new data set, which were not used for estimating the loadings and can be either from the same or from a different subject. If the spatial and spectral signatures of an atom are fixed, they can be used as templates for screening. Formally, after estimating atoms in a training data set, this can be reconstructed as

$$\hat{\boldsymbol{S}} = \mathbf{C}(\mathbf{B} \mid \otimes \mid \mathbf{A})^T, \tag{3}$$

here, $\mathbf{B}|\otimes|\mathbf{A} = [\mathbf{b}_1 \otimes \mathbf{a}_1 \mathbf{b}_2 \otimes \mathbf{a}_2 \ldots \mathbf{b}_{nk} \otimes \mathbf{a}_{nk}]$ is the Khatri-Rao product of $\mathbf{B}$ and $\mathbf{A}$, and represents the convolution of space and frequency. $\mathbf{X}^T$ denotes the transpose of matrix $\mathbf{X}$. For the definition of a template, not all atoms are necessary; that is, for the sake of screening, atoms that are not of interest may be eliminated. Let $\mathbf{B}'$ and $\mathbf{A}'$ be fixed spectral and spatial signatures (with some atoms possibly eliminated). The temporal signature, $\mathbf{C}'$, can then be estimated by using least squares in a new data set $\mathbf{X}$,

$$\mathbf{C}'^T = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{X}; \quad \mathbf{P} = (\mathbf{B}' \mid \otimes \mid \mathbf{A}')^T \tag{4}$$

$\mathbf{P}$ can be regarded as a template for screening for the presence of the atoms of interest. The new temporal signature, $\mathbf{C}'$, will then be an estimate of the detected activities corresponding to each atom in the new data set.

If certain atoms obtained by PARAFAC decomposition contain artifact (e.g., eye movements, eye blinking, electromyogram, etc. . .) their space/frequency reconstructions can be used as templates for an artifact detector. The reconstruction, obtained by eliminating the component that corresponds to artifacts, will be an artifact removal method.

Inverse solution for the spatial signatures of atoms: source spectra imaging

Each column $\mathbf{a}_k$ of matrix $\mathbf{A}$ can be seen as the topography of atom k. Thus, it would be desirable to obtain the sources inside the brain that can produce these topographies to highlight more precise anatomical details. The difficulty here is that the spatial signatures are all positive values, as they are the differential topographic

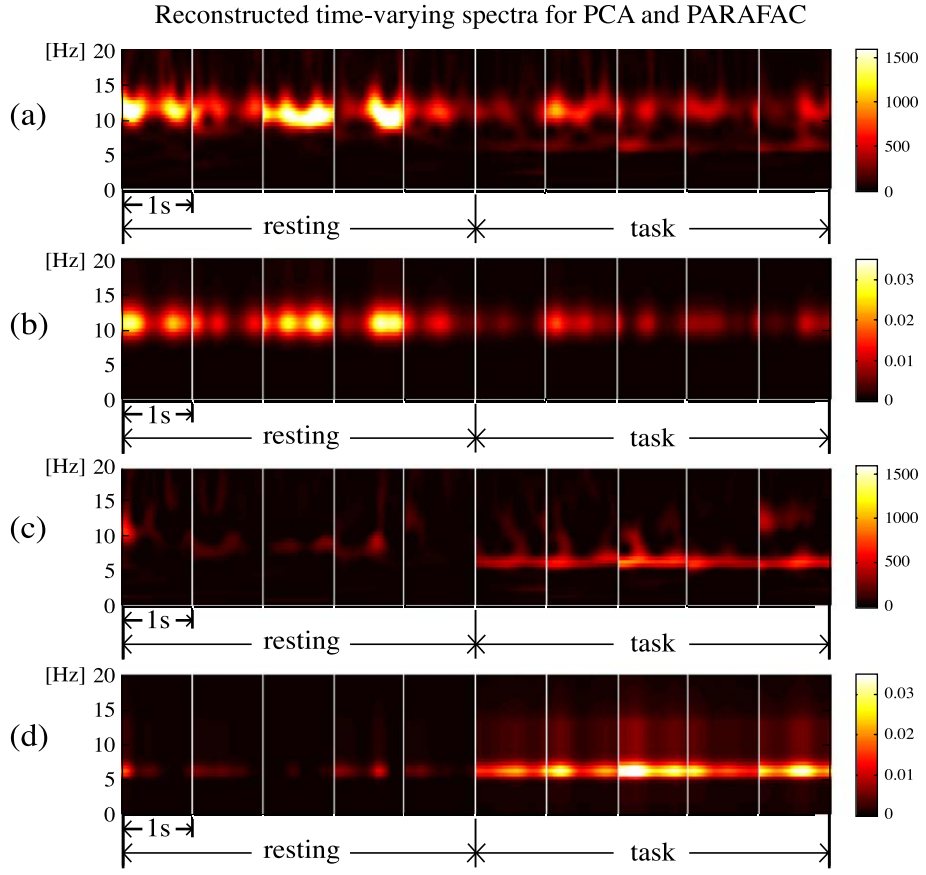

Fig. 6. Panels (a) and (b) illustrate the reconstructed decomposed component by Varimax rotated Principal Component Analysis (PCA) corresponding to the two largest eigenvalues. Panels (c) and (d) are the reconstructed alpha and theta atoms of PARAFAC decomposition in the frequency/time plane, respectively. The first five segments were randomly selected from the resting condition; the second five segments were selected so as to contain typical theta bursts. Each segment is 1-s long and contains 100 time frames.

profiles of EEG spectra, i.e., the variances of complex Fourier coefficients, and therefore the inverse solutions in this case are not the ordinary estimates for the current densities. However, a simple exploratory analysis can be performed in which, under certain assumptions and simplifications, the underlying sources for these topographies can be obtained by an inverse solution. Furthermore, these sources will be shown to be the spectra of electric current densities.

The well-known relation between the electric current densities inside the brain and the electric potentials measured by a set of electrodes on the scalp is:

$$\mathbf{V}_f = \mathbf{K}\mathbf{J}_f \tag{5}$$

Here we have written Eq. (5) directly in the frequency domain, i.e., $\mathbf{V}_{f(N_d \times 1)}$ and $\mathbf{J}_{f(3N_v \times 1)}$ are the vectors of Fourier coefficients of the voltages and electric current density time series, respectively. N_v is the number of voxels of a regular grid inside the brain. The matrix $\mathbf{K}_{(N_d \times 3N_v)}$ is the electric lead field, which is unaffected by the Fourier transformation. As the absolute value of the electric potential has no physical meaning, the average value of voltages was taken as the reference. From Eq. (5), we can find the spectra of voltages as $\alpha = \mathrm{diag}(\mathbf{V}_f\mathbf{V}_f^*)$:

$$\alpha = \mathrm{diag}(\mathbf{K}\mathbf{J}_f\mathbf{J}_f^*\mathbf{K}^T) \tag{6}$$

If we assume that there is no correlation between the current densities in different voxels, i.e., $\mathbf{J}_f\mathbf{J}_f^*$ is a diagonal matrix, we can obtain their spectra as $\gamma = \mathrm{diag}(\mathbf{J}_f\mathbf{J}_f^*)$. Eq. (6) then becomes:

$$\boldsymbol{\alpha} = \boldsymbol{K}^{\wedge 2}\boldsymbol{\gamma} \tag{7}$$

where $\mathbf{K}^{\wedge 2}$ indicates the operation of squaring each element of the matrix, $\mathbf{K}$. This represents a linear relation between α (spatial signatures obtained by PARAFAC decomposition), and the spectra of current sources that generate the scalp voltages. For the sake of simplicity, it may be assumed that the spectrum vector of the current density has the same magnitude in all directions for each voxel. Therefore, Eq. (7) may be rewritten as:

$$\boldsymbol{\alpha} = \mathbf{M}\boldsymbol{\mu} \tag{8}$$

Here, matrix $\mathbf{M}_{(N_d \times N_v)}$ was obtained by averaging every three columns of matrix $\mathbf{K}^{\wedge 2}$, and $\mu_{(N_x \times 1)}$ is the spectrum of current densities for each voxel.

From Eq. (8), the spectra of current sources can be found by an inverse solution procedure. Note here that spectra α and μ are non-negative vectors, allowing us to solve Eq. (8) as a minimum least squares problem under the non-negativity constraint for μ. Eq. (8) is undetermined; thus, we shall constrain the solution to be the smoothest one. In this case, the underlying sources (μ_k) for the topographic signature ($\mathbf{a}_k$) of atom k, can be obtained from

$$\boldsymbol{\mu}_k = \underset{\mu_k \geq 0}{\arg\min} \; |\mathbf{a}_k - \mathbf{M}\boldsymbol{\mu}_k|^2 + \lambda |\mathbf{L}\boldsymbol{\mu}_k|^2 \tag{9}$$

where $\mathbf{L}$ is the discrete Laplacian operator as described in Pascual-Marqui et al. (1994) and λ is a regularization parameter.

In general, we shall call this approximation to the source reconstruction for spectra of voltages the "Source Spectra Imaging" (SSI) solution. Despite the assumption of independence of electric current densities, finding the SSI solution for the spatial signatures implies the a priori assumption of spatial smoothness of the spectrum of current densities. Moreover, for this work, the SSI solution for each atom was obtained by imposing anatomical constraints using the Montreal Neurological Institute Probabilistic Brain Atlas as described in Casanova et al. (2000). This set of

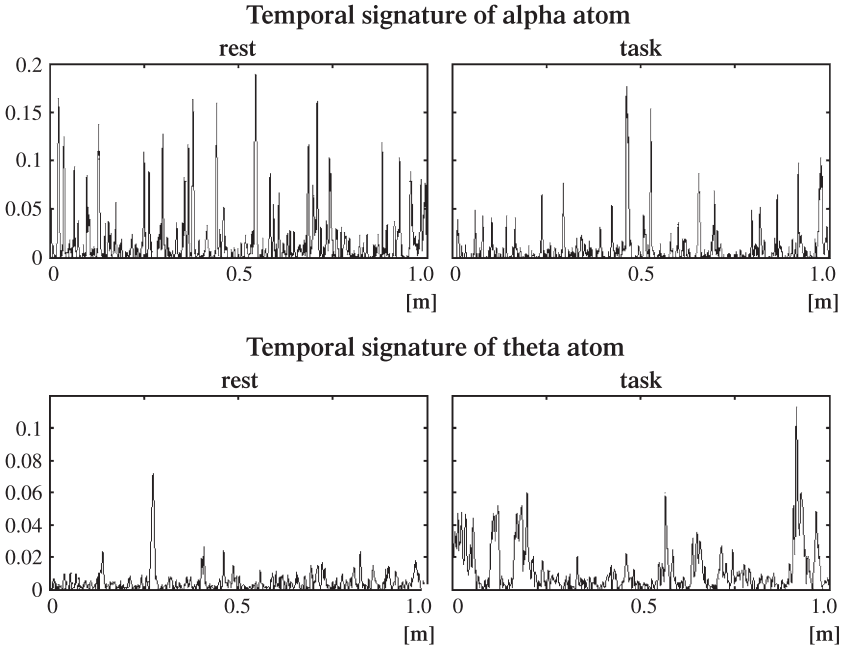

Fig. 7. Theta and alpha activities detected by the screening procedure, based on previously identified spatial and spectral signatures for each component. The screening was applied for 1 min of continuous data sets in the resting and task condition of a new data set of Subject 2. There are more theta bursts and less alpha bursts in the task state, while the converse relationship holds in the resting state.

assumptions has been amply used widely in the so-called distributed inverse solutions, which find the best applicability in the reconstruction of activation of wide areas in the brain. Advantages and shortcomings of these methods have been discussed extensively in the literature (Fuchs et al., 1999; Pascual-Marqui 1995, 1999; Pascual-Marqui et al., 1994).

Results

Parallel Factor Analysis

To evaluate the performance of PARAFAC for extracting alpha and theta activities in EEG, two different states were prepared in a benchmark data set. For this purpose, 10 segments of 1 s each were selected from the wavelet-transformed data (after wavelet transformation the time-varying EEG spectrum data set was subsampled to 100 Hz to reduce the computational cost of PARAFAC). Clear alpha activity is observed continuously during resting and task condition; however, strong theta activity appears only intermittently during task condition. Therefore, five segments were selected randomly from the resting condition and the other five segments were selected from the portions that contain typical theta bursts during the task. The segments were concatenated into the benchmark data set consecutively to form the three dimensional matrix $\mathbf{S}_{(N_d \times N_f \times N_t)}$.

In the PARAFAC decomposition of **S**, two atoms appeared for all subjects with spectral signature peaking in the alpha and theta range (Fig. 2).

The analyzed frequency range was 0.5–20 Hz step by 0.5 Hz, which is sufficient to extract theta and alpha activity. The use of the Corcondia index suggested that in three subjects, these two atoms were sufficient to explain the data set. In two subjects, an additional atom was needed. The Corcondia was more than 90% in all cases (optimally it should be 100%). The alpha and theta peaks were around 11 and 7 Hz, respectively. Subject 1 is typical of those who showed strong alpha and little theta activity during rest conditions. Temporal signatures (Fig. 3) show that during the task condition, this subject produced strong theta activities and reduced alpha activity.

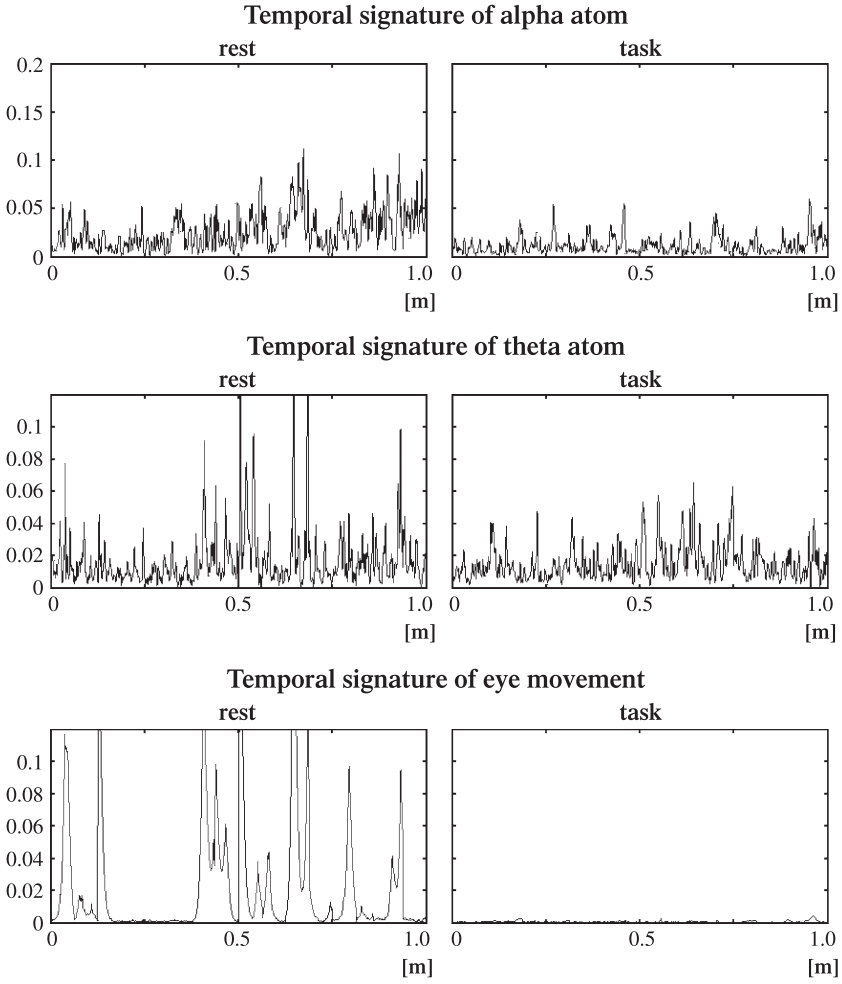

Fig. 8. To train the screening algorithm for the detection of eye movement artifacts, PARAFAC was applied to a data set containing typical theta and alpha activity, as well as eye blinks. Using the spatial and spectral signatures of these three components as a template, we screened 1 min of continuous data sets obtained in the resting condition, which was not used for estimating factors. The data set recorded from Subject 1 contained many eye blink artifacts in the resting condition and fewer in the task condition. On visual inspection of the raw data, it appears that there are more theta bursts during rest than during the task condition. The PARAFAC atomic decomposition showed that this is due to eye blink, as the many peaks in the temporal signature of theta and eye blink atom coincide.

Figs. 4(a) and (b) show the spatial signatures of the alpha and theta atoms as topographic maps for this subject.

The alpha and theta atoms appear in occipital and frontal area, respectively. Figs. 4(c) and (d) are the corresponding SSI solutions for these spatial signatures. The sources for the alpha and theta atoms are in the calcarine sulcus and in the anterior middle frontal cortex, respectively. These spatial distributions were relatively stable for all subjects. Fig. 5 shows the estimated SSI solutions for the spatial signatures of all subjects, corresponding to the theta atom. The activated region had a predominantly frontal distribution in all subjects.

Principal Component Analysis

To compare it with PARAFAC, we also carried out PCA of **S**. For this purpose, the data were transformed into a matrix by unfolding the three dimensional array. Results from our PARAFAC decomposition of **S** were matched with corresponding results obtained by applying PCA to the unfolded data set. Figs. 6(a) and (c) show the reconstructed components in the frequency/time plane that correspond to the two largest eigenvalues of PCA. The first component showed strong alpha activities during the resting condition. The second component shows strong theta activities and reduced alpha activity during the task condition. These components had a marked resemblance to the frequency/time reconstructed plane of the alpha and theta atoms of PARAFAC (Figs. 6(b) and (d)).

The peaks of these activities, as well as the order of appearance of the atoms, were the same in PARAFAC and PCA decompositions. The topographies of the PCA components were also very similar to the spatial signature of the PARAFAC atoms (Figs. 4(a) and (b)).

Screening

Using the screening procedure described above, it was possible to use PARAFAC to search for the presence of atoms in a new data set, which were not used for estimating the loadings. In this study, we consider new data from the same subject and only theta and alpha atoms were of interest. The spatial and spectral signatures for the templates of theta and alpha atoms were estimated by using Subject 2 data as a benchmark. Reconstruction of the temporal signature for new data was carried out by screening 1 min of continuous data in the resting and task conditions. Fig. 7 shows the appearance of pronounced theta bursts and the decrease of alpha bursts in the task state. In the resting state, the theta burst disappeared and the alpha bursts increased.

Artifact detection

If PARAFAC is applied to a data set that contains artifact, some of the atoms will correspond to such activity (e.g., eye blink, eye movement, EMG, etc...). Using these atoms as templates, artifact detection can be carried out by the screening procedure. As an example, PARAFAC was applied first to a training data set from Subject 1 that contained theta and alpha oscillations as well as eye movement artifact (this was assessed empirically by an experienced electrophysiologist). The number of atoms was chosen such they could be identified easily as theta, alpha, and eye movement artifact. Using the spatial and spectral signatures of these three atoms as templates, 1 min of continuous data in resting and task conditions were screened. Fig. 8 shows the corresponding temporal signature of the three atoms for Subject 1.

The data set recorded from this subject contained many eye movement artifacts in the resting condition and far fewer in the task condition. A superficial analysis would lead to the conclusion that there are more theta bursts during the resting than the task condition. However, these are probably due to the presence of artifacts, because there are many coincident peaks in the temporal signatures of the theta and artifact atoms.

Discussion

This paper introduces a new type of space/frequency/time atomic decomposition of the EEG. It takes advantage of the fact that three-way arrays of data may be decomposed into a sum of atoms of which is a trilinear combination of factors or signatures. This decomposition will be unique if the number of atoms is less than half the sum of the ranks of the three matrices formed by concatenating the signatures. The application of this concept to obtain unique space/frequency/time decomposition for the EEG is possible by arranging the multichannel evolutionary spectrum of the EEG in a three-way data array with dimensions indexed by channel, frequency, and time. The underlying theoretical requirement is that of a moderate amount of linear independence for atom topographies, spectra, and time courses. This is a much milder requirement than previous models underlying space/time atomic decompositions (PCA or ICA). This is the first intrinsically unique space/frequency/time atomic decomposition proposed in the literature.

A physiological interpretation of the model presented here is intuitively appealing. It assumes neural sources with a fixed geometrical relation to the sensors that produce oscillatory activity with a fixed spectral whose amplitude is temporally modulated. This model is not a completely general; for example, a frequency-modulated chirp would require a large number of components, such that the rank condition would be violated.

On the other hand, at most three space/frequency/time atoms are necessary for an adequate description of the EEG data analyzed in this paper. The use of the Corcondia index facilitates the selection of the number of components, an issue that is still difficult for most decomposition methods including PCA and ICA. Also, for the data set analyzed in this paper, two of the spectral signatures had a clear and common interpretation as theta and alpha oscillatory activity. Other components were not so constant and were sometimes difficult to interpret. It may be that more a priori information must be built into the model to avoid identification ambiguity. In this regard, PARAFAC shares with ICA the lack of inherent ordering of the extracted components. In the case of ICA, clustering techniques have been applied to identify common modes (Makeig et al., 2002). In the future, this approach might be used also for the space/frequency/time decomposition.

Our work also shows that the temporal signatures of the theta and alpha atoms may be used as indicators of physiological states. A comparison with a PCA-Varimax analysis shows that the results of the latter may sometimes be similar to those of PARAFAC in terms of description of space and frequency/time profiles. PARAFAC, however, provides a more parsimonious description of the data in a qualitatively simpler manner.

An important application of the space/frequency/time atomic decomposition is the screening of new data sets for the presence of particular atoms. In other words, PARAFAC offers the opportunity to screen recordings for bursts of oscillatory activity with a given

topographic and spectral content. The results shown here demonstrate the feasibility of this technique, not only to detect physiological activity but also for the ever-present problem of artifact removal.

One limitation of the implementation of the method presented here is the estimation by the least-squares techniques. Embedding the model in a Bayesian framework would allow more flexibility in incorporating a priori knowledge and a principled testing of different hypotheses about signatures within and between subjects.

Are these results 'real'?

As noted above, it has been proven that if the data is approximately trilinear, if the correct number of components is used, and if the signal-to-noise ratio is appropriate, then the true underlying phenomena will be found with PARAFAC (Harshman 1972; Kruskal, 1976, 1997). Also, there have been examples in which the PARAFAC model coincides with a physical model, e.g., fluorescence excitation–emission, chromatography with spectral detection, and spatiotemporal analysis of multichannel-evoked potentials (Field and Graupe, 1991).

The usual and stronger way to validate the truthfulness of results given by the application of multiway models is by split-half analysis (Harshman, 1984; Harshman and De Sarbo, 1984). Due to the uniqueness of the PARAFAC model, the same loadings must be obtained in the non-split modes from models of any suitable subset of the data. This analysis was not accomplished in this work. Although we do not have definitive proof that our results reflect exactly the real physical phenomena underlying the data, there are some aspects of the method we can lean upon for assessing the robustness of the model.

First, the algorithm is implemented such that we can select different initial values. We obtain the same results by applying the method with initial values given by direct trilinear decomposition of the data as we do by random guesses. Second, changing the convergence criterion over four orders of magnitude did not affect the results. Finally, the interpretability of the results, their agreement with previous studies of this kind of electrophysiological experiment, and their robustness with constraints to the loadings like non-negativity and orthogonality; as well as the small variability among subjects, all give additional evidence in this regard.

From this perspective, we think that PARAFAC space/frequency/time atomic decomposition of multichannel evolutionary spectrum of the EEG can reliably and uniquely extract meaningful and significant physiological activities, although this does not ensure that the results correspond to the physical sources that generated the data. Furthermore, the application of this technique requires careful preprocessing of the data, exploration of outliers and degenerate solutions, use of constraints, selection of appropriate model order, and validation of the results as this cannot be accomplished easily without prior knowledge of, or a theoretical basis for, of the expected results. PARAFAC should be simply considered another promising addition to the Neuroimaging analysis toolkit.

Acknowledgments

The authors want to thank Prof. Mark S. Cohen, Director of Functional MR Imaging, Ahmanson-Lovelace Brain Mapping Center, UCLA School of Medicine, for his very helpful advice and suggestion for this work.

References

Achim, A., Bouchard, S., 1997. Toward a dynamic topographic components model. Electroencephalogr. Clin. Neurophysiol. 103, 381–385.

Bertrand, O., Bohorquez, J., Pernier, J., 1994. Time–frequency digital filtering based on an invertible wavelet transform: an application to evoked potential. IEEE Trans. Biomed. Eng. 41, 77–88.

Bro, R., 1998. Multi-way Analysis in the Food Industry: Models, Algorithms and Applications. PhD Thesis. University of Amsterdam (NL) and Royal Veterinary and Agricultural University (DK).

Carroll, J.D., Chang, J., 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition. Psychometrika 35, 283–319.

Casanova, R., Valdes-Sosa, P., Garcia, F.M., 2000. Frequency domain distributed inverse solution. In: Aine, C.J., Okada, Y., Stroink, G., Swithenby, S.J., Wood, C.C. (Eds.), Biomag 96: Proceedings of the Tenth International Conference on Biomagnetism. Springer Verlag.

Chen, S., Donoho, D., 2001. Atomic decomposition by basis pursuit. SIAM Rev. 43, 129–159.

Cichocki, A., Amari, S., 2002. Adaptive Blind Signal and Image Processing. John Wiley & Sons, Ltd.

Dahlhaus, R., 1997. Fitting time series models to non-stationary processes. Ann. Stat. 25, 1–37.

Estienne, F., Matthijs, N., Massart, D.L., Ricoux, P., Leibovici, D., 2001. Multi-way modelling of high-dimensionality electroencephalographic data. Chemom. Intell. Lab. Syst. 58, 59–71.

Field, A.S., Graupe, D., 1991. Topographic component (Parallel Factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decomposition. Brain Topogr. 3, 407–423.

Fuchs, M., Wagner, M., Kohler, T., Wischman, H.A., 1999. Linear and nonlinear current density reconstructions. J. Clin. Neurophysiol. 16, 267–295.

Gonzalez Andino, S.L., Grave de Peralta Menendez, R., Lantz, C.M., Blank, O., Michel, C.M., Landis, T., 2001. Non-stationary distributed source approximation: an alternative to improve localization procedures. Hum. Brain Mapp. 14, 81–95.

Harmony, T., Fernandez, T., Silva, J., et al., 1999. Do specific EEG frequencies indicate different processes during mental calculation? Neurosci. Lett. 266, 25–28.

Harshman, R.A., 1970. Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multi-modal factor analysis. UCLA Work. Pap. Phon. 16, 1–84.

Harshman, R.A., 1972. Determination and proof of minimum uniqueness conditions for PARAFAC1. UCLA Work. Pap. Phon. 22, 111–117.

Harshman, R.A., 1984. "How can I know if it's 'real'?" A catalog of diagnostics for use with three-mode factor analysis and multidimensional scaling. In: Law, H.G., Snyder, C.W., Hattie, J.A., McDonald, R.P. (Eds.), Research Methods for Multimode Data Analysis. Praeger, New York, pp. 566–591.

Harshman, R.A., De Sarbo, W.S., 1984. An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques. In: Law, H.G., Snyder, J.A., Hattie, J.A., McDonald, R.P. (Eds.), Research Methods for Multimode Data Analysis. Praeger, New York, pp. 602–642.

Hyvarinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis. John Wiley & Sons, Inc.

Ishihara, T., Yoshii, N., 1972. Multivariate analytic study of EEG and mental activity in juvenile delinquents. Electroencephalogr. Clin. Neurophysiol. 33, 71–80.

Jensen, O., Tesche, C.D., 2002. Frontal theta activity in human increased with memory load in a working memory task. Eur. J. Neurosci. 15, 1395–1399.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2001. Analysis and visualization of single-trial event-related potential. Hum. Brain Mapp. 14, 166–185.

Koenig, T., Marti-Lopez, F., Valdes-Sosa, P., 2001. Topographic time-frequency decomposition of the EEG. NeuroImage 14, 383–390.

Kronland-Martinet, R., Morlet, J., 1987. Analysis of sound patterns through wavelet transforms. Int. J. Pattern Recogn. Artif. Intell. 1, 273–302.

Kruskal, J.B., 1976. More factors than subjects, test and treatments: an indeterminacy theorem for canonical decomposition and individual differences scaling. Psychometrika 41, 281–293.

Kruskal, J.B., 1977. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra Appl. 18, 95–138.

Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208.

Lagerlund, T.D., Sharbrough, F.W., Busacker, N.E., 1997. Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. J. Clin. Neurophysiol. 14, 73–83.

Lopes da Silva, F., 1987. EEG analysis: theory and practice. In: Neidermeyer, E., Lopes da Silva, F. (Eds.), Electroencephalography. Urban and Schwartzenberg.

Makeig, S., 1993. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr. Clin. Neurophysiol. 86, 283–293.

Makeig, S., Westerfield, M., Jung, T.-P., et al., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Möcks, J., 1988a. Decomposing event-related potential: a new topographic components model. Biol. Psychol. 26, 199–215.

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

Nunez, P.L., 1993. Electric Fields of the Brian: The Neurophysics of EEG. Oxford Univ. Press.

Pascual-Marqui, R.D., 1995. Reply to comments by Hamalainen, Ilmoniemi and Nunez. In source localization: continuing discussion of the inverse problem. Skrandies W ISBET Newsletter, vol. 6, pp. 16–28. ISSN 0947-5133.

Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. Int. J. Bioelectromagn. 1, 1.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.

Sasaki, K., Tsujimoto, T., Nishikawa, S., Nishitani, N., Ishihara, T., 1996. Frontal mental theta wave recorded simultaneously with megnetroencephalography and electroencephalography. Neuroscience 26, 79–81.

Sidiropoulos, N.D., Bro, R., 2000. On the uniqueness of multilinear decomposition of *N*-way arrays. J. Chemom. 14, 229–239.

Soong, A.C., Koles, Z.J., 1995. Principal-component localization of the sources of the background EEG. IEEE Trans. Biomed. Eng. 42, 59–67.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1997. Oscillatory g-band (30–70 Hz) activity induced by a visual search task in humans. J. Neurosci. 17, 722–734.

Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. Nat. Rev. Neurosci. 2, 229–239.