ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 29 (2006) 938 – 947

Parallel Factor Analysis as an exploratory tool for wavelet transformed event-related EEG

Morten Mørup,[a,*] Lars Kai Hansen,[a] Christoph S. Herrmann,[b]
Josef Parnas,[c] and Sidse M. Arnfred[c]

[a]*Informatics and Mathematical Modelling, IMM, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kongens Lyngby, Denmark*
[b]*Department of Biological Psychology, Otto-Von-Guericke University, Magdeburg, Germany*
[c]*Department of Psychiatry, Hvidovre Hospital, University Hospital of Copenhagen, Denmark*

Received 5 May 2005; revised 22 July 2005; accepted 1 August 2005
Available online 26 September 2005

In the decomposition of multi-channel EEG signals, principal component analysis (PCA) and independent component analysis (ICA) have widely been used. However, as both methods are based on handling two-way data, i.e. two-dimensional matrices, multi-way methods might improve the interpretation of frequency transformed multi-channel EEG of channel × frequency × time data. The multi-way decomposition method Parallel Factor (PARAFAC), also named Canonical Decomposition (CANDECOMP), was recently used to decompose the wavelet transformed ongoing EEG of channel × frequency × time (Miwakeichi, F., Martinez-Montes, E., Valdes-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y., 2004. Decomposing EEG data into space–time–frequency components using parallel factor analysis. Neuroimage 22, 1035–1045). In this article, PARAFAC is used for the first time to decompose wavelet transformed event-related EEG given by the inter-trial phase coherence (ITPC) encompassing ANOVA analysis of differences between conditions and 5-way analysis of channel × frequency × time × subject × condition. A flow chart is presented on how to perform data exploration using the PARAFAC decomposition on multi-way arrays. This includes (A) channel × frequency × time 3-way arrays of *F* test values from a repeated measures analysis of variance (ANOVA) between two stimulus conditions; (B) subject-specific 3-way analyses; and (C) an overall 5-way analysis of channel × frequency × time × subject × condition. The PARAFAC decompositions were able to extract the expected features of a previously reported ERP paradigm: namely, a quantitative difference of coherent occipital gamma activity between conditions of a visual paradigm. Furthermore, the method revealed a qualitative difference which has not previously been reported. The PARAFAC decomposition of the 3-way array of ANOVA *F* test values clearly showed the difference of regions of interest across modalities, while the 5-way analysis enabled visualization of both quantitative and qualitative differences. Consequently, PARAFAC is a promising data exploratory tool in the analysis of the wavelets transformed event-related EEG.

Keywords: PARAFAC; CANDECOMP; Decomposition; ERP; ITPC; Gamma coherence; Multi-way arrays; Channel–time–frequency analysis; Wavelet analysis

Introduction

Electroencephalographic (EEG) recordings have traditionally been examined as indices of brain state – examining the rhythmic activity of the resting EEG – or as indices of stimulus elicited activity—examining the event-related potentials (ERP) which are averages of many trials. A long standing discussion about the interpretation of the ERP as resulting either from phase resetting of oscillatory activity also present in the ongoing EEG (Basar et al., 2001) or as a combination of several true polarized single high amplitude bursts seems to have come to a rest at the point where the ERP primarily is seen as depending on phase resetting of oscillatory activity (Gruber et al., 2004, 2005; Makeig et al., 2002). This has led to an increasing interest in examining the event-related EEG activity in time × frequency plots rather than only looking at the time series of the event-related EEG voltage (Gruber et al., 2004, 2005; Herrmann et al., 2004a,b; Jansen et al., 2004; Jones et al., 2002; Lachaux et al., 2005). Although the temporal resolution of EEG data is excellent and is the virtue of the method, the possibility of extracting spatial information from scalp recorded potentials of multiple electrodes has been a major focus in recent years, extending the data analyses in two major directions: one focusing on dipole or source localization via elaborate statistical

---

[*] Corresponding author. Fax: +45 4587 2599.
E-mail address: mm@imm.dtu.dk (M. Mørup).

doi:10.1016/j.neuroimage.2005.08.005

models trying to solve the "backward" problem inherent in the analyses (Koles, 1998) and another focusing on mathematical decomposition on the data (Dormann et al., 1987; Makeig et al., 1997; Rogers, 1991). Presently, our aim was to extend the mathematical decomposition from the traditionally handling of a channel × time ERP signal to the handling of ERP data arrays of channel × frequency × time. The work was inspired by Miwakeichi et al. (2004) who used the decomposition method Parallel Factor (PARAFAC) to decompose the power of the ongoing EEG into channel–frequency–time components. In the present report, we propose a way to use PARAFAC decomposition as an exploratory tool in the analyses of wavelet transformed event-related EEG data. In contrast to Miwakeichi et al. (2004), we analyzed the inter-trial phase coherence (ITPC) also encompassing modalities such as condition and subject in addition to an ANOVA test of difference between conditions. In order to evaluate the method, we aimed at replicating a study of visual information processing, where it had been previously documented that early evoked gamma activity (30–80 Hz) differed between two conditions, defined by stimuli having long-term memory representations and stimuli having no such representations (Herrmann et al., 2004a,b).

Methods

Theory

In the following, a description of the PARAFAC model will be given. This will be followed by a description of the inter-trial phase coherence (ITPC). Finally, the design of the multi-way array of this article will be given.

PARAFAC

The use of PARAFAC in the analysis of EEG and ERP is not new. In his original paper on PARAFAC, Harshman in 1970 suggested PARAFAC to be used to decompose the EEG (Harshman, 1970). But, it was not before 1988 when Möcks reinvented the model, naming it topographic component analysis, that PARAFAC was used to analyze the ERP of channel × time × subjects (Möcks, 1988), an idea that was further pursued by Field and Graupe (1991). In 2004, Miwakeichi suggested the use of PARAFAC on the wavelet transformed ongoing EEG in the representation channel × frequency × time. It was here shown how PARAFAC was capable of successfully identifying the theta and alpha atoms of a cognitive task and that the decomposition method could identify eye blinks (Miwakeichi et al., 2004).

Traditionally, the decomposition of the EEG into components has been based on decomposition techniques such as principal component analysis (PCA) (Collet, 1989; Dormann et al., 1987; Kayser et al., 2003; Picton et al., 2000; Rogers, 1991) and independent component analysis (ICA) (Makeig et al., 1997, 1999). When EEG data are transformed by the use of continuous wavelets for the analysis of frequency content, the increase of dimensionality gives the two-way array i.e. the matrix of channel × time, an extra modality yielding a 3-way array of channel × frequency × time. Further problems arise when modalities such as subject and condition are included, yielding a 5-way array of channel × frequency × time × subject × condition. ICA and PCA can only analyze such data by unfolding some modalities onto others, reducing the multi-way array into matrices. The unfolding makes the interpretation of the results difficult and removes the specific information endorsed by the modalities. Consequently, rather than unfolding these multi-way arrays into matrices, we analyzed the data using the multi-way model PARAFAC.

The Parallel Factor (PARAFAC) model was independently proposed by Harshman (1970) and by Carrol and Chang (1970), the latter naming it Canonical Decomposition (CANDECOMP). The model is a parsimonious extension of the factor analysis to higher orders as revealed in Fig. 1. In matrix notation, factor analysis can be expressed as:

$$\mathbf{X}^{I \times J} = \mathbf{A}^{I \times F} \mathbf{S}^{J \times F^{T}} + \mathbf{E}^{I \times J} \tag{1}$$

Where $\mathbf{A}$ is the factor loading, $\mathbf{S}$ the factor score, $\mathbf{E}$ the error, and F the number of factors.Similarly, the PARAFAC of the 3-way array $\mathbf{X}^{I \times J \times K}$ can be expressed by unfolding one modality to another, i.e.

$$\mathbf{X}^{I \times JK} = \mathbf{A}^{I \times F} \left(\mathbf{S}^{K \times F} \odot \mathbf{D}^{J \times F}\right)^{T} + \mathbf{E}^{I \times JK} \tag{2}$$

Where $\mathbf{D}$ is the factor score corresponding to the second modality. $\mathbf{S} \odot \mathbf{D} = [\mathbf{s}_1 \otimes \mathbf{d}_1 \; \mathbf{s}_2 \otimes \mathbf{d}_2 \ldots \mathbf{s}_F \otimes \mathbf{d}_F]$ is the Khatri–Rao product (Bro, 1998). Equivalently, the jth matrix corre-

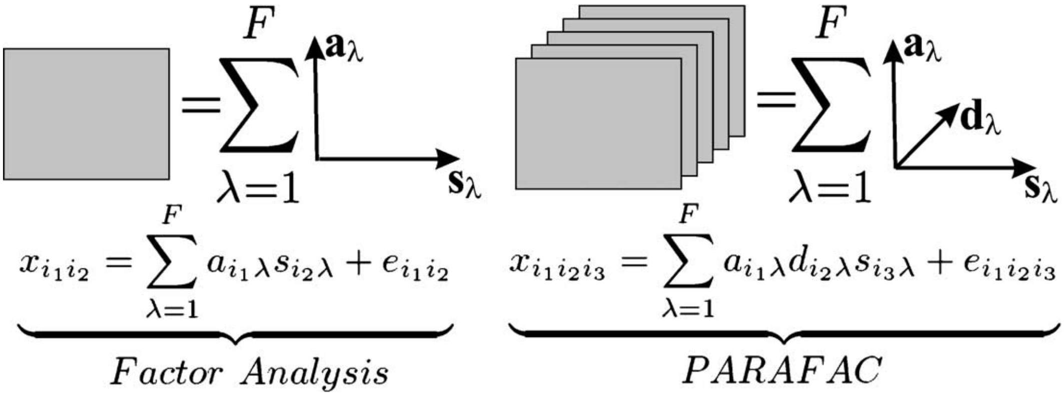

Fig. 1. Graphical representation of the factor analysis to the left and the PARAFAC decomposition of a 3-way array to the right. Like the factor analysis, PARAFAC decomposes the data into factor effects pertaining to each modality. F denotes the number of factors.

sponding to the jth slice of the second modality of the 3-way array can be expressed as:

$$\mathbf{X}^{I \times j \times K} = \mathbf{A}^{I \times F} \mathbf{D}_j^{F \times F} \mathbf{S}^{K \times F^T} + \mathbf{E}^{I \times j \times K} \tag{3}$$

Where $\mathbf{D}_j$ is a diagonal matrix having the jth row of $\mathbf{D}$ along the diagonal. In a similar fashion for multi-way arrays $\boldsymbol{X}^{I_1 \times I_2 \times \ldots \times I_N}$ of higher order than 3, the PARAFAC model is given by the following three equivalent expressions:

$$x_{i_1 i_2 \cdots i_N} = \sum_{\lambda=1}^{F} a_{i_1 \lambda}^{(1)} a_{i_2 \lambda}^{(2)} \cdots a_{i_N \lambda}^{(N)} + e_{i_1 i_2 \cdots i_N}$$

$$\mathbf{X}^{I_1 \times I_2 \cdots I_{N-1} I_N} = \mathbf{A}^{(1)} \left(\mathbf{A}^{(N)} \odot \mathbf{A}^{(N-1)} \odot \cdots \odot \mathbf{A}^{(2)} \right)^T + \mathbf{E}^{I_1 \times I_2 \cdots I_{N-1} I_N}$$

$$\begin{aligned} \mathbf{X}^{I_1 \times i_2 \times i_3 \times \cdots \times i_{N-1} \times I_N} &= \mathbf{A}^{(1)} \mathbf{A}_{i_2}^{(2)} \mathbf{A}_{i_3}^{(3)} \cdots \mathbf{A}_{i_{n-1}}^{(N-1)} \mathbf{A}^{(N)^T} \\ &\quad + \mathbf{E}^{I_1 \times i_2 \times i_3 \times \cdots \times i_{N-1} \times I_N} \end{aligned} \tag{4}$$

Consider again the 3-way array. If this array was to be analyzed using ICA or PCA one of the modalities, i.e. the third modality would have to be unfolded onto another modality, i.e. the second in order to have an analyzable matrix. This would require a much larger number of free parameters than the PARAFAC model, as in general the following holds:

$$\underbrace{F(I + J \cdot K)}_{\text{PCA/ICA}} >> \underbrace{F(I + J + K)}_{\text{PARAFAC}} \tag{5}$$

Furthermore, factor analysis suffers from rotational freedom as a rotation of the solution of $\mathbf{A}$ by the matrix $\mathbf{P}$ and a counter rotation of $\mathbf{S}^T$ by $\mathbf{P}^{-1}$ yields an equivalent model:

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T = (\mathbf{A}\mathbf{P})\left(\mathbf{P}^{-1}\mathbf{S}^T\right) = \tilde{\mathbf{A}}\tilde{\mathbf{S}}^T \tag{6}$$

Consequently, constraints in the form of orthogonality and VARIMAX rotation for PCA and the stronger statistical independence for ICA have to be imposed in order to insure unique solutions. Rotational indeterminacy in PARAFAC on the other hand would require:

$$\begin{aligned} \mathbf{X}^{I \times j \times K} &= \mathbf{A}\mathbf{D}_j\mathbf{S}^T \Rightarrow \mathbf{X}^{I \times j \times K} = (\mathbf{A}\mathbf{P})\left(\mathbf{P}^{-1}\mathbf{D}_j\mathbf{Q}\right)\left(\mathbf{Q}^{-1}\mathbf{S}^T\right) \\ &= \tilde{\mathbf{A}}\tilde{\mathbf{D}}_j\tilde{\mathbf{S}}^T \end{aligned} \tag{7}$$

However, since $\tilde{\mathbf{D}}_j$ has to be a diagonal matrix for all j, $\mathbf{P}^{-1}$ $\mathbf{D}_j\mathbf{Q}$ has to be diagonal which in practice restricts $\mathbf{P}$ and $\mathbf{Q}$ to only be each a scaling and permutation matrix. As a result, the PARAFAC model is in general unique apart from scaling and permutation indeterminacies. Kruskal gave in 1977 a very mild uniqueness condition. The result makes use of the k-rank $k_\mathbf{B}$ which is given by the smallest subset of columns of the matrix $\mathbf{B}$ that are linearly independent (Kruskal, 1977):

$$k_\mathbf{A} + k_\mathbf{D} + k_\mathbf{S} \geq 2F + 2 \tag{8}$$

This condition was extended by Sidiropoulos and Bro to higher orders, N (Sidiropoulos and Bro, 2000), i.e.

$$\sum_{i=1}^{N} k_{\mathbf{A}^{(i)}} \geq 2F + (N-1) \tag{9}$$

Consequently, the main advantage of PARAFAC over PCA or ICA is that uniqueness is ensured under mild conditions, making it unnecessary to impose constraints in the form of orthogonality or statistical independence. This is achieved through a very restricted but also very easy interpretable model.

The most common way of estimating the PARAFAC model is by alternating least squares. In this approach, a cost function (normally the squared error) is minimized in order to explain most of the variation in the data, i.e. $\left[\hat{\mathbf{A}}, \hat{\mathbf{S}}, \hat{\mathbf{D}}\right] = \operatorname{argmin}_{\mathbf{A},\mathbf{S},\mathbf{D}} \|\mathbf{X}^{I \times JK} - \mathbf{A}(\mathbf{S} \odot \mathbf{D})^T\|^2$, which corresponds to optimizing the maximum likelihood of a Gaussian noise model. This is done by alternating between re-estimating each parameter given the estimation of the other parameters. The algorithm can be initialized in several ways, i.e. by randomly defining all parameters and stopped when all parameters have converged. For a description of this simple but popular algorithm, confer Bro (1998). For recent alternative algorithms, see Beckmann and Smith (2005) and Cao et al. (2000, 2003) (Fig. 2).

Inter-trial phase coherence (ITPC)

Although the wavelet's estimate of the coefficient at a given time–frequency is local and the whole analysis is influenced by the choice of wavelet, wavelet analysis is considered a powerful tool in the analysis of the temporal development of the frequency of the event-related EEG (Herrmann et al., 2005). For the analysis of the frequency changes in the ERP signal, the complex Morlet wavelet as

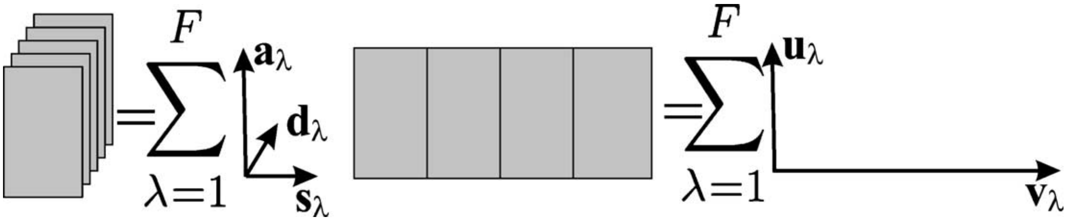

Fig. 2. The PARAFAC decomposition of the 3-way array to the left compared to the corresponding factor analysis decomposition of the 3-way array by unfolding the third mode to the second. Clearly, the number of free parameters is larger in the factor analysis compared to PARAFAC (compare the length of $\boldsymbol{d}_\lambda$ and $\boldsymbol{s}_\lambda$ with $\boldsymbol{v}_\lambda$). Consequently, PARAFAC is a more restricted model ensuring uniqueness under very mild conditions. PARAFAC is easy to interpret as the effect of each modality on the λ^{th} factor is given by the component vector of that modality. In contrast to this, no simple interpretation of each modality can be extracted in the factor analysis since $\boldsymbol{v}_\lambda$ is a vector with contributions from several, i.e. in the figure, two modalities.

described by Herrmann et al. (1999) and Miwakeichi et al. (2004) was used:

$$\tilde{\varphi}(t) = \frac{1}{\sqrt{\pi \Omega_b}} \exp(i2\pi\Omega_c t)\exp\left(-\frac{t^2}{\Omega_b}\right) \tag{10}$$

Ω_c is the center frequency and Ω_b is a bandwidth parameter.

Let $\mathbf{X}_e$ (c, f, t) be the coefficient of the wavelet transform at channel c at frequency f and time t for epoch e, and let there be a total of n epochs. When analyzing the wavelet transformed event-related EEG, the following measures are useful (Delorme and Makeig, 2004):

$$\mathrm{ERSP}(c,f,t) = \frac{1}{n}\sum_{e=1}^{n} |\mathbf{X}_e(c,f,t)|^2$$

$$\mathrm{ITPC}\ (c,f,t) = \left| \frac{1}{n}\sum_{e=1}^{n} \frac{\mathbf{X}_e(c,f,t)}{|\mathbf{X}_e(c,f,t)|} \right| \tag{11}$$

While the event-related spectral perturbation (ERSP) is a measure of the power in the signal, the inter-trial phase coherence (ITPC), also referred to as phase locking factor, is a measure of evoked activity. An ITPC value of one indicates perfect phase coherence in all epochs, while random noise on the average has a coherence of $\frac{1}{\sqrt{n}}$ (Nunez et al., 1997). In the following analysis, the ITPC is of interest. Since the ITPC is computed as the mean of n unit vectors in the complex plane and as the variance of the distribution is finite, as the ITPC takes values between 0 and 1, we can appeal to the central limit theorem and expect that the distribution of the ITPC is asymptotically normal for $n \to \infty$ (see also Mardia and Jupp, 1999). By bootstrapping, it was indeed found that for n = 104, the ITPC could be considered normally distributed (data not shown).

Design of the multi-way arrays

In the present article, three types of multi-way decompositions are performed: (A) a multi-way of F test values from an analysis of variance (ANOVA), (B) a 3-way array of ITPC values given by a subjects channel × frequency × time, and (C) a 5-way array of ITPC values also having subject and condition as modalities. Each multi-way array is described below.

Since the ITPC is approximately normal distributed (n = 104 epochs), an ANOVA will be used to investigate the difference between the evoked activities given by the ITPC of two conditions. Let ITPC(c, f, t, s, k) be the ITPC at channel c at frequency f at time t for subject s in condition k. Let there be a total of K conditions and S subjects. It is assumed that the ITPC of each subject is drawn from the same normal-distribution. The ANOVA F test value $Z(c, f, t)$ of difference between the two conditions at channel c at frequency f at time t is then given by:

$$Z(c,f,t) = \frac{\sum_{k=1}^{K} S(I(c,f,t,k) - I(c,f,t))^2/(K-1)}{\sum_{k=1}^{K}\sum_{s=1}^{S} (\mathrm{ITPC}(c,f,t,s,k) - I(c,f,t,k))^2/(S-K)}$$

where

$$I(c,f,t,k) = \frac{1}{S}\sum_{s=1}^{S} \mathrm{ITPC}\,(c,f,t,s,k)$$

$$I(c,f,t) = \frac{1}{KS}\sum_{k=1}^{K}\sum_{s=1}^{S} \mathrm{ITPC}\,(c,f,t,s,k) \tag{12}$$

This multi-way array of ANOVA test values given by $Z(c, f, t)$ will be decomposed using the PARAFAC model to indicate where the largest differences between the two conditions are present.

Decomposing the ITPC of a subject given by the 3-way array of channel × frequency × time using PARAFAC, i.e.

$$x_{i_1 i_2 i_3} = \sum_{\lambda=1}^{F} a_{i_1\lambda} d_{i_2\lambda} s_{i_3\lambda} + e_{i_1 i_2 i_3} \tag{13}$$

corresponds to the assumption that the underlying factors consist of a given frequency signature $\boldsymbol{d}_\lambda$ of strength in time given by $\boldsymbol{s}_\lambda$ that has been mixed in the channels by $\boldsymbol{a}_\lambda$.

Extending the PARAFAC analysis to a 5-way array of channel × frequency × time × subject × condition, i.e. into the PARAFAC model,

$$x_{i_1 i_2 i_3 i_4 i_5} = \sum_{\lambda=1}^{F} a_{i_1\lambda} d_{i_2\lambda} s_{i_3\lambda} p_{i_4\lambda} q_{i_5\lambda} + e_{i_1 i_2 i_3 i_4 i_5} \tag{14}$$

amounts to the additional assumptions that the underlying factors are the same in all subjects but with a subject-dependent strength given by the subject score vector $\boldsymbol{p}_\lambda$ and furthermore with a variable strength in the two conditions given by the score $\boldsymbol{q}_\lambda$.

Although the above assumptions are very strong, a main benefit is the reduction in the number of parameters $F\sum_{i=1}^{N} N_i \ll \prod_{i=1}^{N} N_i$, thus the PARAFAC analysis has the great advantage of providing unique and interpretable components. While variation between subjects and conditions is certainly present, we believe that the variation is not too large for the model to extract a weighted average of the underlying process. This is in line with the conventional assumption of event-related analysis, assuming that, although the temporal signature of the event-related EEG varies from epoch to epoch, the averaged ERP is still believed to depict the true (average) temporal signature.

While the PARAFAC analysis of the ANOVA can give an indication where the difference between the ITPC of the two conditions are present, analyzing the ITPC given by the 5-way array enables the visualization not only of quantitative differences but also of qualitative, cross-modal, differences between conditions. If a factor is clearly expressed in both conditions, but with different strengths, it is an indication of a quantitative difference. If a factor is mainly present in one of the two conditions, it is an indication of a qualitative difference.

Experimental details

Eleven healthy male subjects participated in our study. They gave informed consent as approved by the Ethics Committee, and they were paid to participate in the experiment. The mean age of the sample was 25 years (SD 1.4), and the mean length of education was 15.7 years (SD 1.8).

EEG was recorded with 64 scalp electrodes (BioSemi Active electrodes system) arranged according to the International 10–10 system. Additional recordings were obtained from earlobes and at lateral canthus/brow for each eye. The grounding electrodes for the active electrodes (CMS and DRL) were placed centrally, close to POz. Data were recorded continuously at 2048 Hz/channel, band pass 0.1–160 Hz, by a LabView© application (ActivView©) on a Windows©-based PC. After down sampling to 512 Hz/channel, the further processing was performed in EEGLAB for MatLab©

(Delorme and Makeig, 2004). The data were referenced to digitally linked earlobes and cut into epochs (−250 to +500 ms).

The stimulus paradigm has been described in detail previously (Herrmann et al., 2004a,b). Briefly, it consists of two types of black and white drawings: 1. objects (Ob), which are easily recognizable everyday type of objects like a chair, a number or a pipe, and 2. non-objects (Nob), which are chaotic re-arrangements of the Ob drawings. No stimulus is shown more than once. Each stimulus category included 104 events. Stimulus delivery was controlled by the Presentation© software. Each stimulus was presented for 1 s, and randomized inter-stimulus interval had durations of 1.3–1.7 s. The stimulus presentation monitor was placed 75 cm in front of the comfortably seated subject. To keep stimuli in focus yet keep the object/non-object dichotomy unaware, the subjects were instructed to respond with a mouse button press depending on their judgment of the drawings as primarily having round contours or primarily having edges. This response was not included in the data analysis, which is focused on the difference between Ob and Nob un-related to conceptual categorization. In the prior report of this experiment in healthy volunteers, evoked gamma activity was elicited bilaterally in the occipital region at approximately 100 ms post-stimulus. The evoked activity was significantly stronger in the Ob condition as compared to the Nob condition (Herrmann et al., 2004a,b).

Pre-processing

The data were wavelet transformed using a complex Morlet wavelet from MatLab© Wavelet Toolbox with center frequency $\Omega_c = 1$ and bandwidth parameter $\Omega_b = 2$ with frequencies represented from 1 to 80 Hz with 1 Hz interval between each frequency. As a measure of the evoked activity the inter-trial phase coherence (ITPC), also named the phase locking factor (see Eq. (11)), was calculated for each subject and each condition. This yielded a multi-way array of channel × frequency × time as shown in Fig. 3. Baseline subtraction was not performed prior to wavelet transformation since the wavelet transform is invariant of additions or subtractions of constants. The ITPC is a measure of phase consistency, where each epoch contributes with a vector of the same (unit) length giving all epochs equal influence on the ITPC. Since even very noisy signals might include the correct phase information, no epochs were rejected. This enabled the ITPC to be calculated as an average across all trials, improving signal to noise ratio (SNR). Furthermore, to avoid reduction of SNR, the data were not normalized across subjects. Normalizing would increase the influence of subjects having less coherence compared to random activity in the analysis. In the PARAFAC analyses, background activity was removed by subtracting the mean activity of each channel and frequency from −250 to −100 ms.

PARAFAC analysis

Since ITPC data approximately represent a normal distribution and assuming that the PARAFAC model can capture the systematic variation in the data, the noise can be assumed normal. Consequently, the PARAFAC model was calculated using the alternating least square algorithm. Both the ANOVA and ITPC yield non-negative values. As a result, neither $\boldsymbol{a}_\lambda$, $\boldsymbol{d}_\lambda$, $\boldsymbol{s}_\lambda$ nor $\boldsymbol{p}_\lambda$ or $\boldsymbol{q}_\lambda$ can take negative values. Thus, the model was calculated with non-negativity constraints on all modalities (Bro and Jong, 1997). Although the alternating least squares PARAFAC algorithm has been proven to suffer from degeneration and slow convergence (Beckmann and Smith, 2005; Cao et al., 2000, 2003; Paatero, 2000), imposing non-negativity constraints makes degeneration of solutions hard since no factor can counteract the effect of another factor and improves convergence since the search space is greatly reduced. While the PARAFAC model theoretically is unique, in practice, unique solutions are not always achieved (Bro, 1998). However, it has been shown that in general imposing non-negativity also improves the uniqueness (Bro, 1998; Donoho and

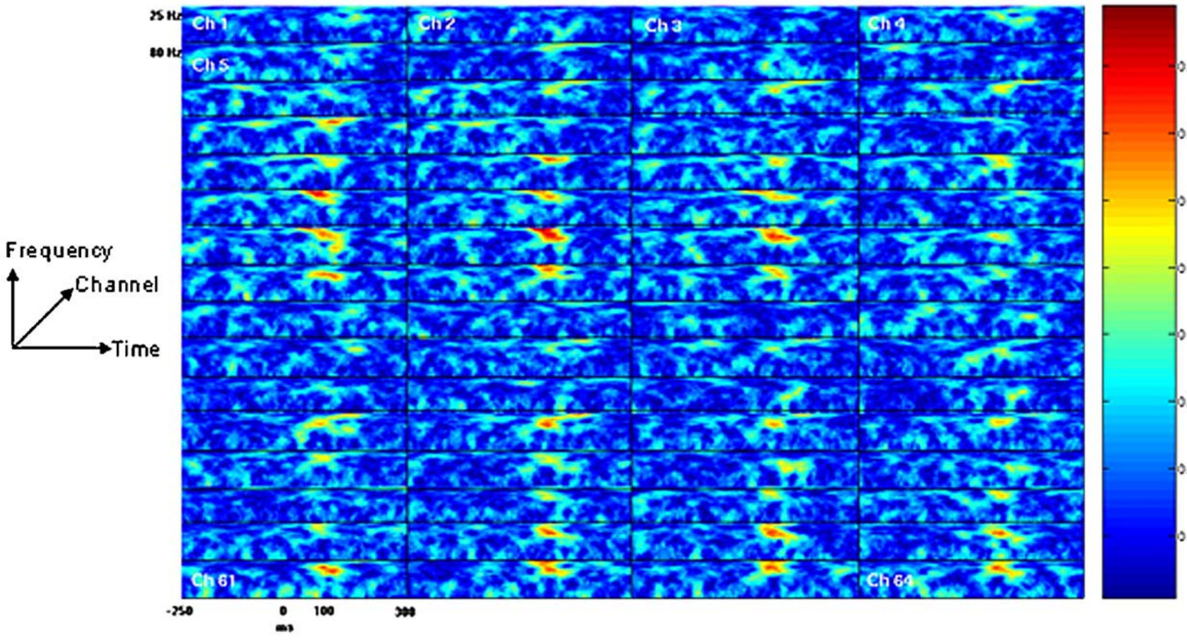

Fig. 3. Example ITPC of one subject given by the multi-way array of channel × frequency × time. The 64 channels are shown in a 16 × 4 array with their respective time–frequency plots from −250 to 300 ms and 25 to 80 Hz.

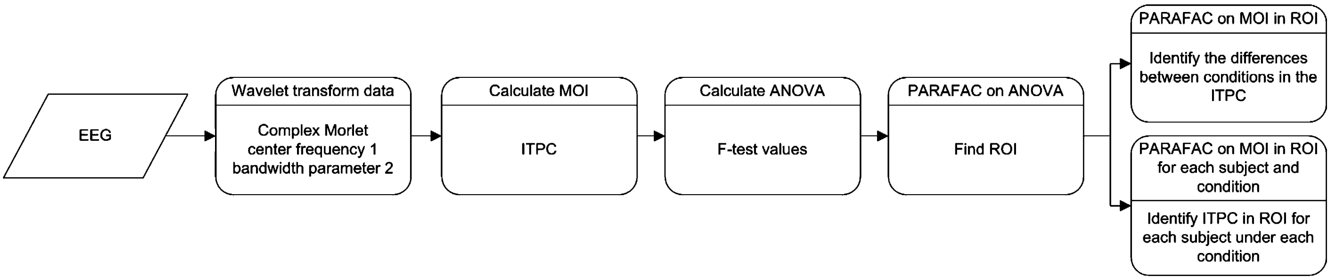

Fig. 4. Flow chart of the analysis. The ERP was wavelet transformed using a complex Morlet wavelet with center frequency 1 and bandwidth parameter 2. The measure of interest (MOI), i.e. the ITPC, was calculated for each subject and each condition according to Eq. (11) and a 3-way array of the ANOVA *F* test value calculated according to Eq. (12). The *F* test array was analyzed using PARAFAC and the region of interest (ROI, here in the time–frequency domain) of most difference between the two conditions identified. The ITPC in the ROI was then analyzed using PARAFAC on the 5-way array given by channel × frequency × time × subject × condition of ITPC values. Finally, the ITPC of the ROI of each subject under each condition given by the 3-way array of channel × frequency × time was also analyzed by PARAFAC and the individual peaks of the ITPC identified under each condition.

Stodden, 2004). Consequently, the ALS algorithm with non-negativity constraints was found adequate to model the data. The model was calculated using the “N-way toolbox” for MatLab© kindly provided by Rasmus Bro (http://www.models.kvl.dk/source).

Accessing the correct number of factors, *F*, is important since over- or underestimating the number of factors does impact the decomposition (Beckmann and Smith, 2005; Bro, 1998). The number of factors to include in the PARAFAC model was estimated using Core Consistency Diagnostic (Bro, 1998) and the Bayesian Information Criterion. Furthermore, the explained variation as well as convergence of the algorithm was taken into consideration. The algorithm was run three times for each analysis to ensure by visual inspection that stable solutions were reached. For most of the decompositions, a clear indication of the number of factors to include was given. We suspect that imposing non-negativity reduces the impact of the choice of *F*, but this needs further investigation. No problems of convergence were encountered (Fig. 4).

Results

The 3-way array analysis of the ANOVA of difference between the ITPC of the Ob and Nob conditions taken over the 11 subjects (Eq. (12)) yielded a good fit in a one-component PARAFAC model (cf. Fig. 5) where it is evident that the main difference in the evoked activity is in the gamma band between 40 and 80 Hz around 100 ms.

Due to the successful fit in this frequency region at that time, subsequent analyses were performed in that time–frequency range (30–80 Hz; 0–200 ms).The number of channels was not restricted. The ITPC 5-way array (channel × frequency × time × subject × condition) could only entail a two-component PARAFAC model. As seen in Fig. 6, the first component encompassed occipital activity at approximately 30 Hz and 100 ms. This first component was present in both conditions, but it was attenuated in the Nob condition (condition 2). The activity was present in all subjects, but it was not strong in subjects 3, 4, and 5. The second component was localized more anteriorly, and it revealed a higher frequency, while peak latency was similar to the first component. This second component was almost totally limited to the Nob condition. The specific Nob activity was not present in subjects 1, 4, 10, and 11.

14.8% of the total variation was described by the 5-way analysis. This might not seem very much, but one has to bear in mind that ITPC of EEG data is rather noisy even with $n = 104$ epochs and that the PARAFAC model is very restricted. To evaluate the appropriateness of PARAFAC to model the data, we reconstructed the ITPC arrays of each subject from the factors found in Fig. 6 and calculated the residuals to look for systematic variation, as revealed in Fig. 7.

When the ITPC was decomposed in each condition of every subject (the 3-way array of channel × frequency × time), a one-

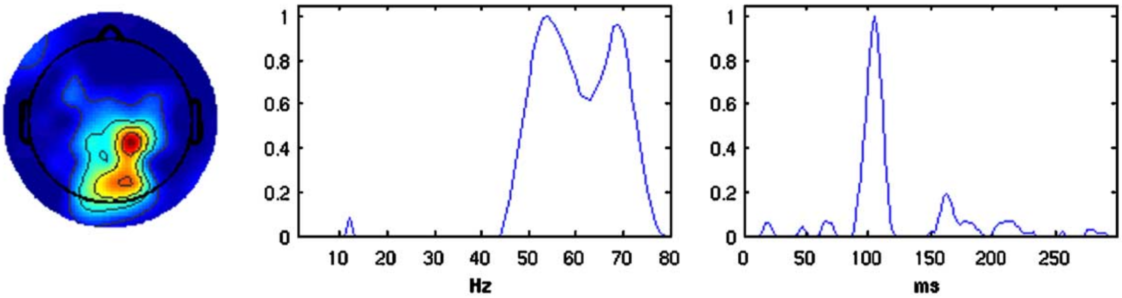

Fig. 5. The PARAFAC decomposition of the ANOVA *F* values multi-way array testing the difference between the two conditions in the 11 subjects. The difference between Ob and Nob condition is mainly in the gamma band around 40–80 Hz of the occipital region at about 100 ms (due to scale indeterminacy, the frequency and time components have been given a maximal value of 1).

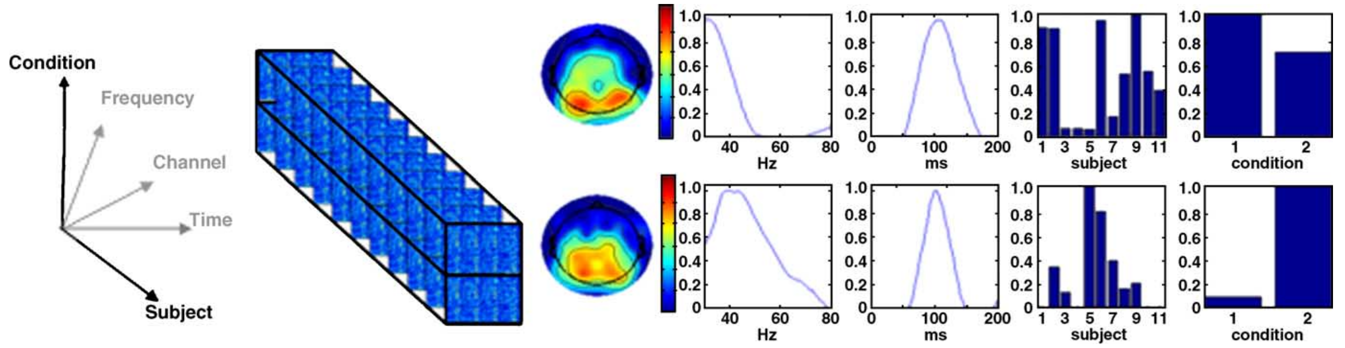

Fig. 6. Left panel: the analyzed 5-way array of the ITPC given by channel × frequency × time × subject × condition in the ROI, i.e. 30–80 Hz and 0–200 ms. Each subjects ITPC of channel × frequency × time similar to the one shown in Fig. 3 is given by the Ob condition (upper ITPCs) and Nob condition (Lower ITPCs) for 11 subjects yielding a total of 22 ITPCs. Right panel: the corresponding two-component PARAFAC model. Whereas the first component indicates a quantitative difference between the coherence in the occipital region across the two conditions, the second component indicates a qualitative difference where a slightly more anterior higher frequency activity prevails in the non-object condition. The color range of the topographic plots is the same as the color range given in Figs. 8 and 9. The first factor explains 11.2% of the total variation, whereas the second factor only explains 3.6%. Notice that the factors do not signify statistical significance but explain most of the variation in the data given the model. Although only 14.8% of the variation is explained, it is the most systematic variation being the most common through the subjects which is also the variation of most interest in the data.

component PARAFAC model yielded a good fit. This resulted in 22 decomposition plots (11 subjects · 2 conditions) as illustrated in Fig. 8, showing the Ob condition of subject 6.

In these decomposition plots, the time and frequency point of the peak of the gamma activity was visually identified. In four plots, it was not possible to identify a peak. Here, the peaks were defined at 40 Hz and 100 ms. In some subjects, more than one peak was found in the frequency signature. Here, the peak of highest frequency was chosen. The ITPC topography corresponding to the individual peak time and frequency was obtained. The grand average of these topographies for each condition is shown in Fig. 9, illustrating that the peak frequency/moment evoked activity is attenuated in the Nob condition.

Discussion

Although the PARAFAC model seemed to capture the systematic variation in the data as revealed in Fig. 7, it is an open question whether other models are more appropriate. The PARAFAC analysis of the 5-way array might have been more accurate if modalities were combined. The PARAFAC model was chosen

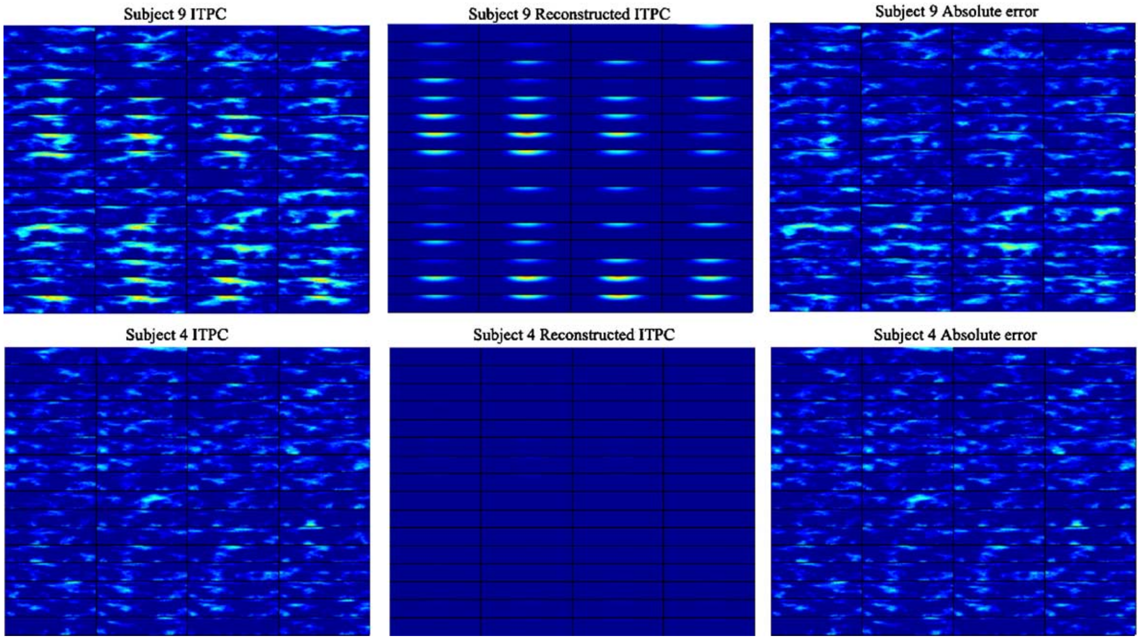

Fig. 7. ITPC arrays as described in Fig. 3 here shown from 30 to 80 Hz and 0–200 ms for the Ob condition. Data are for subject 9 (top images) having a high subject score in Fig. 6 and subject 4 (bottom images) having a low subject score. The original ITPC is shown in the left images, the reconstructed ITPC calculated from the factors found in Fig. 6 is depicted in the middle images, and the absolute value of the residuals is given in the right images. While subject 9 seems to have systematic variation in the ITPC, the variation of subject 4 appears to be much more random. The reconstructed ITPC seems to capture most of the systematic variation of subject 9, whereas almost none of the variation in subject 4 is described. The residuals of subject 9 and 4 however look both quite unsystematic.

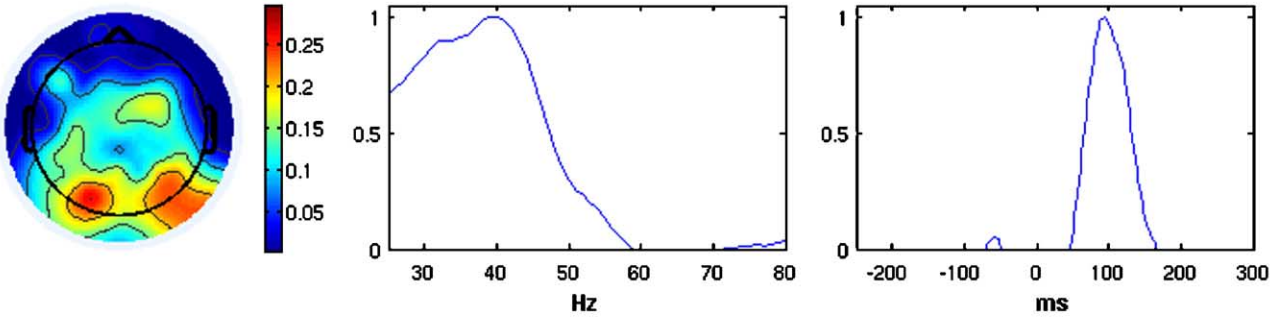

Fig. 8. Example of a one-component PARAFAC model calculated from a subject's channel × frequency × time multi-way array of the ITPC. The component indicates bilateral gamma activity in the occipital region around 40 Hz approximately 100 ms after stimuli onset.

from a variety of multi-way models, e.g., PARAFAC2, TUCKER1, TUCKER2, and TUCKER3 among others (Bro, 1998; Smilde et al., 2004). This was mainly due to the uniqueness properties of PARAFAC and the fact that the components were very easy to interpret contrary to most of the other multi-way models (Bro, 1998; Miwakeichi et al., 2004; Smilde et al., 2004). However, the very strong assumption of PARAFAC that identical activity is to be found across subjects is questionable. This was also one of the reasons why some subjects lacked the evoked gamma activity in Fig. 6—simply because their activity deviated too much from the mean activity.

From the PARAFAC analysis of the ANOVA F test values, it was evident that the difference between the Ob and Nob condition was in the gamma band of the occipital region around 100 ms. This matched previous findings in this paradigm, where the difference was statistically confirmed based on an explicit a priori hypothesis of difference in early gamma activity in this region (Herrmann et al., 2004a,b). It would seem that the explorative application of PARAFAC on the ANOVA F test values could give a solid idea if and where in the three-dimensional array a region of difference between two conditions is present. Presently, the F test values represent a means for exploration. If PARAFAC was considered for hypothesis testing, the channel, frequency, and time peak found from the ANOVA 3-way or ITPC 5-way analysis could be used to test for difference between conditions on a new dataset. Another approach would be to use the channel, frequency, and time signatures of a factor found by PARAFAC. These signatures could be used on new datasets to find subject scores p_{ik}, where i is the subject number and k is the condition. Hypothesis testing could then be achieved through statistical analysis of these scores.

Having identified the region of interest, the PARAFAC decomposition of the 5-way array of ITPC values (Fig. 6) revealed two effects: 1. a quantitative difference between the two conditions as the Ob condition entailed more evoked activity in the occipital region at the lower gamma band than the Nob condition. Again, this was a replication of previous findings, even though the measures of evoked activity and the procedures for the detection of difference were entirely different (Herrmann et al., 2004a,b). 2. A qualitative difference where evoked activity of a higher frequency, located more anteriorly, was more pronounced in the Nob than in the Ob condition. This qualitative difference outside the peak frequency region might easily have been overlooked with traditional data processing. Before it is attempted to interpret this finding, it is essential to verify it in another dataset as the procedure here was exploratory.

Both the 3-way and 5-way PARAFAC decompositions performed successfully. Admittedly, the 3-way array could have been investigated using factor analysis models by unfolding one modality to the other. But, it is noteworthy that the results of the 5-way PARAFAC analysis could not be obtained by unfolding the array to a matrix. In this case, at least one of the factor analysis' modalities would represent three of the original modalities which would make the factor components hard to interpret. The 5-way PARAFAC endorses novel and easy applicable comprehensive

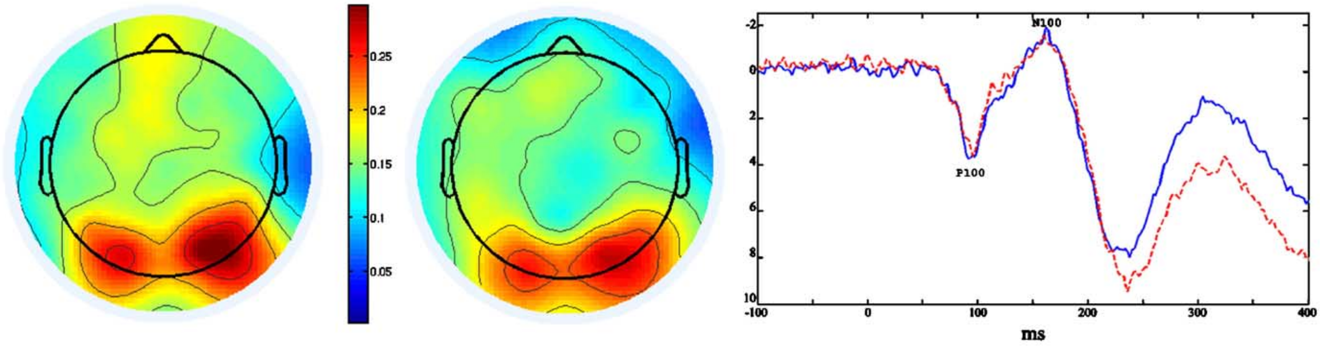

Fig. 9. Left panel: the grand average ITPC topography (11 subjects) at the individual peak of the gamma activity separately averaged for each condition (Ob to the left and Nob to the right). Evoked activity is attenuated in the Nob condition compared to the Ob condition (since a few subjects lacked coherent gamma activity this difference is not significant). Right panel: the grand average ERP (11 subjects) of the two conditions at channel O2 (Ob: solid, Nob: dashed). No digital filtering was applied, baseline was subtracted from −250 to 0 ms. Artifact rejection was based on variance: 20% rejected. No difference is observed between the two conditions at the mid-latency components (P100 and N100), while the later processing diverges, un-related to the task of discriminating between rounded and edgy drawings (this difference between the ERPs is not significant).

view of data, which has not previously been seen. Finally, the PARAFAC decomposition plot of the ITPC of each individual subject in each condition enabled easy read off, a procedure which here enabled the construction of a cross modality grand average topography again substantiating the previously reported condition difference. Following the results of exploratory PARAFAC analyses, this individual type of PARAFAC decomposition could be applied in even narrower windows of interest in another subject sample, examining the quantitative and qualitative differences in separate windows. This could be used in the statistical confirmation of the differences.

In this report on evoked activity, the measure of interest (MOI) was the ITPC. Yet, the PARAFAC is expected to perform just as successful on other MOI like the ERSP. Consequently, the flow chart of the present exploration as shown in Fig. 4 is considered applicable to a wide range of analyses of the wavelet transformed event-related EEG. Employing PARAFAC on ANOVA F values to identify the region of interest throughout the sample and exploring the region, identified in this manner, by PARAFAC on the MOI is a promising route in data exploration. For confirmatory studies, other routes as suggested above would be more appropriate.

Since PARAFAC performed so excellently decomposing the various multi-way arrays, it is surprising that the model has been employed so limited within EEG research. Probably, the main reasons are the current focus on matrix decomposition and source localization. Another reason could be the increased load on computer running memory implemented by the multi-way data. The wavelet analysis increases the size of the data immensely, requiring 2 GB of computer RAM for the 5-way array decomposition. RAM of this size has only recently become available on PC. Thirdly, the earlier attempts of using PARAFAC on EEG measures might have been aborted due to problems of degeneration in the factors when applying PARAFAC on the voltage data (Field and Graupe, 1991).

Finally, PARAFAC might become an important tool in the analysis of a wide range of brain data. The scheme developed here is obviously applicable also to MEG data, and we think it is worthwhile to mention that PARAFAC has previously been applied to fMRI (Beckmann and Smith, 2005). Consequently, great potentials lie ahead in terms of the analysis of brain data using PARAFAC.

Conclusion

PARAFAC decomposition is a promising data exploratory tool in the analysis of wavelet transformed event-related EEG. The method was able to extract the expected features of a previously reported ERP paradigm also incorporating subject and condition modalities. The PARAFAC decomposition of the 3-way array of ANOVA F test values clearly shows the difference regions of interest across modalities, while the 5-way ITPC analysis enabled visualization of both quantitative and qualitative differences.

Acknowledgments

We wish to thank the reviewers for excellent, insightful critique. The study was economically supported by the Lundbeck Foundation, the Gangsted Foundation, the Novo Nordisk Foundation, and the Danish Research Council.

References

Basar, E., Schurmann, M., Demiralp, T., Basar-Eroglu, C., Ademoglu, A., 2001. Event-related oscillations are ‘real brain responses’—Wavelet analysis and new strategies. Int. J. Psychophysiol. 39, 91–127.

Beckmann, C.F., Smith, S.M., 2005. Tensorial extensions of independent component analysis for multisubject FMRI analysis. NeuroImage 25, 294–311.

Bro, R., 1998. Multi-way Analysis in the Food Industry: Models, Algorithms and Applications. Amsterdam, Copenhagen.

Bro, R., Jong, S.D., 1997. A fast non-negativity-constrained least squares algorithm. J. Chemom. 11, 393–401.

Cao, Y.Z., Chen, Z.P., Mo, C.Y., Wu, H.L., Yu, R.Q., 2000. A PARAFAC algorithm using penalty diagonalization error (PDE) for three-way data array resolution. Analyst 125, 2303–2310.

Cao, Y.Z., Chen, H., Wu, H.L., Yu, R.Q., 2003. An algorithm for three-way data analysis that alternatively minimizes coupled vector (COV) resolution error and PARAFAC error. Anal. Sci. 19, 753–756.

Carrol, 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.

Collet, W., 1989. Doubts on the adequacy of the principal component varimax analysis for the identification of event-related brain potential components: a commentary on Glaser and Ruchkin, and Donchin and Heffley. Biol. Psychol. 28, 163–172.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Donoho, D., Stodden, V., 2004. When does non-negative matrix factorization give a correct decomposition into parts? Proc. NIPS* 2003. Adv. Neural Inf. Process, vol. 16. MIT Press.

Dormann, W.U., Mundlos, S., Haschke, R., 1987. Comparison of principal components computed with principal factor analysis on the basis of averaged and single-trial ERPs using the Fischer–Roppert procedure. Int. J. Psychophysiol. 4, 319–323.

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.

Gruber, T., Malinowski, P., Muller, M.M., 2004. Modulation of oscillatory brain activity and evoked potentials in a repetition priming task in the human EEG. Eur. J. Neurosci. 19, 1073–1082.

Gruber, W.R., Klimesch, W., Sauseng, P., Doppelmayr, M., 2005. Alpha phase synchronization predicts P1 and N1 latency and amplitude size. Cereb Cortex. 371–377.

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

Herrmann, C.S., Mecklinger, A., Pfeifer, E., 1999. Gamma responses and ERPs in a visual classification task. Clin. Neurophysiol. 110, 636–642.

Herrmann, C.S., Lenz, D., Junge, S., Busch, N.A., Maess, B., 2004a. Memory-matches evoke human gamma-responses. BMC Neurosci. 5, 13.

Herrmann, C.S., Munk, M.H.J., Engel, A.K., 2004b. Cognitive functions of gamma-band activity: memory match and utilization. Trends Cogn. Sci. 8 (8), 347–355.

Herrmann, C.S., Grigutsch, M., Busch, N.A., 2005. EEG oscillations and wavelet analysis. In: Handy, T. (Ed.), Event-Related Potentials: A Methods Handbook. MIT Press, Cambridge, pp. 229–259.

Jansen, B.H., Hegde, A., Boutros, N.N., 2004. Contribution of different EEG frequencies to auditory evoked potential abnormalities in schizophrenia. Clin. Neurophysiol. 115, 523–533.

Jones, K., Begleiter, H., Porjesz, B., Wang, K., Chorlian, D., 2002. Complexity measures of event related potential surface Laplacian data calculated using the wavelet packet transform. Brain Topogr. 14, 333–344.

Kayser, J., Fong, R., Tenke, C.E., Bruder, G.E., 2003. Event-related brain potentials during auditory and visual word recognition memory tasks. Brain Res. Cogn. Brain Res. 16, 11–25.

Koles, Z.J., 1998. Trends in EEG source localization. Electroencephalogr. Clin. Neurophysiol. 106, 127–137.

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., George, N., Tallon-Baudry, C., Martinerie, J., Hugueville, L., Minotti, L., Kahane, P., Renault, B., 2005. The many faces of the gamma band response to complex visual stimuli. NeuroImage 25, 491–501.

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. U. S. A. 94, 10979–10984.

Makeig, S., Westerfield, M., Townsend, J., Jung, T.P., Courchesne, E., Sejnowski, T.J., 1999. Functionally independent components of early event-related potentials in a visual spatial attention task. Philos. Trans. R. Soc. Lond. B Biol. Sci. 354, 1135–1144.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Mardia, K.V., Jupp, P.E., 1999. Directional Statistics. Wiley and sons, pp. 76–77.

Miwakeichi, F., Martinez-Montes, E., Valdes-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y., 2004. Decomposing EEG data into space–time–frequency components using parallel factor analysis. NeuroImage 22, 1035–1045.

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

Nunez, P.L., Srinivasan, R., Westdorp, A.F., Wijesinghe, R.S., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1997. EEG coherency. I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. Electroencephalogr. Clin. Neurophysiol. 103, 499–515.

Paatero, P., 2000. Construction and analysis of degenerate PARAFAC models. J. Chemom. 14, 285–299.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 127–152.

Rogers, L.J., 1991. Determination of the number and waveshapes of event related potential components using comparative factor analysis. Int. J. Neurosci. 56, 219–246.

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

Smilde, A., Bro, R., Geladi, P., 2004. Multi-way analysis: applications in the chemical sciences. Multi-way Analysis. John Wiley & Sons (ISBN 0-471-98691-7).