Eye Blink Artifact Removal in EEG Using Tensor Decomposition

Dimitrios Triantafyllopoulos and Vasileios Megalooikonomou

Multidimensional Data Analysis and Knowledge Management Laboratory
Dept. of Computer Engineering and Informatics, University of Patras
26500 Rion-Patras, Greece
{dtriantafy,vasilis}@ceid.upatras.gr

Abstract. EEG data are usually contaminated with signals related to subject's activities, the so called artifacts, which degrade the information contained in recordings. The removal of this additional information is essential to the improvement of EEG signals' interpretation. The proposed method is based on the analysis, using Tucker decomposition, of a tensor constructed using continuous wavelet transform. Our contribution is an automatic method which processes simultaneously spatial, temporal and frequency information contained in EEG recordings in order to remove eye blink related information. The proposed method is compared with a matrix based removal method and shows promising results regarding reconstruction error and retaining the texture of the artifact free signal.

Keywords: eye blink, Tucker decomposition, wavelet transform, EEG.

1 Introduction

Electroencephalographic (EEG) data are used in several cases where the brain's functionality needs to be analyzed. Their relative ease of use, their low power consumption, compact size and convenience as well as their temporal resolution, are some of the reasons that EEG recordings are preferred in several cases from their alternatives, such as functional Magnetic Resonance Imaging. Some of the applications where EEG data are used are Brain Computer Interaction (BCI) applications as well as medical tasks that require the supervision of a patient's brain functionality, like seizure monitoring.

Unfortunately, during EEG recordings except the underlying brain related activity, there are several signals captured also that degrade the information gathered. These signals are either related to external or internal reasons and are generally referred to as artifacts. External artifacts are often occurring due to technology's shortcomings such as electrical lines introducing a 50 Hz component, affecting the data recorded. The surrounding environment, such as walls and electrical devices, may cause additional noise. Those artifacts are usually counteracted with filtering techniques, for example the application of a Notch filter at 50 Hz.

Internal artifacts are related to the signals captured due to the subject's physical actions. For example eye blinking causes capturing of additional information in the

L. Iliadis et al. (Eds.): AIAI 2014 Workshops, IFIP AICT 437, pp. 155–164, 2014.

frontal electrodes. Eye movement, as well as muscle movement, generate also undesired signals that are, of course, recorded. Avoiding these artifacts is one solution, but as showed in several studies [1,2] , instructing subjects to avoid movement or any other activity that can cause an artifact affects the recorded data. On the other hand, the treatment of these artifacts is a challenging procedure and requires two steps. First the detection of their occurrence and then the handling of the detected events aim in the manipulation of these phenomena.

The detection of artifacts can be performed with visual inspection from a medical expert. But this procedure is, as expected, time consuming. Alternatively there are several automatic techniques that can be applied during or after recordings and detect the occurring artifacts. There are techniques, which are based in thresholds applied in time or frequency domain like in [3]. Some others are based in a supervised learning method using statistical [4] or autoregressive (AR) [5] features. Other methods, like [6], use a combination of feature extraction and data driven thresholds.

Once the artifacted epochs are detected, there are two ways of manipulating these periods. The first one is rejecting the epoch that contains the artifacted signals. This procedure may solve the problem of artifacted epochs but results in substantial information loss. The second methodology is the removal of artifact related information and conservation of brain activity.

Removal methods are separated in two main categories. The first one aims in the removal of the EOG captured signal, which is captured with a separate sensor, with the use of linear combination and regression techniques [6]. The second category of removal methods aims in the clarification of the components-sources that compose the recorded data and afterwards the identification of the artifact related ones. The most used decompositions are Principal Component Analysis (PCA), Canonical Correlation Analysis (CCA) and Independent Component Analysis (ICA). In [7] and [8] CCA is performed in order to remove muscle artifacts. Several methods report that ICA based removal methods outperform the methods based in other decompositions. In [9] an ICA based method discriminates the artifact related independent components, using an SVM classifier. In [10] using high order statistics as criteria, wavelet decomposition in order to divide the recorded signal in frequency bands and ICA as the decomposition method, the proposed technique removes eye blink or muscle movement related artifacts from a pre specified epoch. In [11] the proposed ICA based method by using correlation coefficients between independent sources, specifies the independent component that most likely contains the artifact's signature. In [12] a combination of ICA and Wiener filtering supress eye blink related artifacts.

In this study the effectiveness of a different decomposition method is examined. The decompositions in the methods mentioned so far are based on two dimensional matrices, whose dimensions correspond to electrodes and time samples respectively. In order to examine, when necessary, the frequency information encapsulated in recordings the aforementioned methods, first extract spectral features and then classify the components as artifacts or not. By constructing a three-way array (tensor) whose first two modes correspond to electrode and time domain, while the third encloses the frequency information, it is possible to examine the three domains simultaneously. This concept was used previously in EEG signals in several studies [13, 14, 15].

In [16] the same model was used in order to examine it's abilities in seizure localization as well as in artifact extraction. Regarding the latter the proposed method suggests multilinear subspace analysis with applying Tucker decomposition in the tensor created, in favor of removing the information that is correlated with the artifact. The main aim of our work is to examine the capability of tensors and more specifically Tucker decomposition's as a model to discriminate artifact related information from brain related activity. Our contribution is an automatic artifact removal method based on a similar tensor model with the one in [13] and [16]. The proposed method is compared with an ICA method, similar to the one proposed in [10]. The result of the proposed modeling is a procedure which automatically removes artifact related information, while encapsulating simultaneously the spatial, temporal and frequency underlying structure of the recorded data.

In the following sections the theoretical background, the methodology followed as well as the experimental scheme is presented. In the end the conclusions and future work are discussed.

2 Background

2.1 Multilinear Arrays

Multilinear arrays or tensors, as more than often are referred to, are a multidimensional generalization of vectors. A N-th order tensor is a product on N vector spaces having their own coordinate system [17]. A matrix is a special case of tensors of order 2. It should be noted that the order of a tensor is the number of its dimensions. A general N-th order real tensor is notated as $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$. As in the 2-dimensional case, where there exist several decomposition methods, like SVD, QR or NMF, there are high order generalizations of those methods that can be applied in high order tensors.

The most important of those decompositions are PARAFAC and Tucker decomposition. The former, introduced in [18] and in [19] independently, decomposes the tensor in its rank-1 tensor components. The latter proposed in [20] is considered as the higher order generalization of SVD.

2.2 High Order Decomposition

PARAFAC. Introduced in [18] and in [19] independently, is based in the proposition of rank-1 decomposition of tensors in [21]. In [22] the same model was augmented, by introducing the parallel proportional profiles, whose aim was the estimation of sources ("source traits") that fit simultaneously in many profiles. After its third introduction, in 1970 by the two teams, the model is considered a very important tool in multilinear analysis. Its computation in [23], the toolbox used in this research, is based in an Alternating least squares (ALS) algorithm. The 3-rd order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ using PARAFAC is decomposed in R rank-1 tensors as follows:

$$\mathcal{X} = \sum_{i=1}^{R} \mathrm{a_i} \circ \mathrm{b_i} \circ \mathrm{c_i} + \mathcal{E} \tag{1}$$

where $\mathrm{a_i} \in \mathbb{R}^{I_1}$, $\mathrm{b_i} \in \mathbb{R}^{I_2}$, $\mathrm{c_i} \in \mathbb{R}^{I_3}$, $\mathcal{E} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is the residual tensor and $\circ$ denotes the outer product.

Tucker. It is considered as the generalization of SVD, hence there are several cases where it is referred to as High Order SVD (HOSVD). A three way example of this decomposition is parted from three projection matrixes corresponding in each of the modes of the tensor, and one core tensor capturing the correlation of the different components. Since this interaction is allowed Tucker is considered as a more flexible model than PARAFAC. In [23] Tucker is computed using a truncated SVD like method as an initialization step and then an ALS algorithm is followed. The Tucker decomposition of a 3-rd order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, with R components in every mode is considered as follows:

$$\mathcal{X} \approx \mathcal{C} \times_1 A \times_2 B \times_3 C \tag{2}$$

where $\mathcal{C} \in \mathbb{R}^{R \times R \times R}$ is the core tensor, $\mathrm{A} \in \mathbb{R}^{I_1 \times R}$, $\mathrm{B} \in \mathbb{R}^{I_2 \times R}$, $\mathrm{C} \in \mathbb{R}^{I_3 \times R}$ are the projection matrices and $\times_i$ denotes the i-mode product (for more detail refer to [17]).

Figure 1 shows the Tucker decomposition of the above scheme:

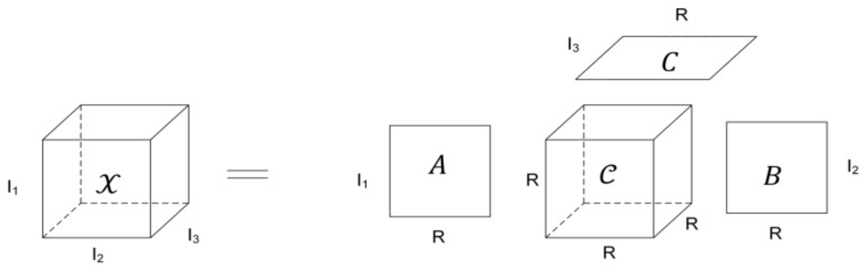

Fig. 1. Tucker decomposition. PARAFAC is a special case of the Tucker decomposition where the core tensor is diagonal and the number of components in each mode are the same.

3 Methodology

In the proposed method the higher order decomposition used is Tucker decomposition. As described in [16] and verified by our experiments, PARAFAC was unable to distinguish the artifact related information without removing also a large amount of useful information. This is due to the fact that generally artifact-related components correspond to components with large variance, therefore the removed information contains also great amounts of useful information. The more relaxed model of Tucker decomposition, due to the core tensor's non-diagonal structure, is more flexible to distinguish the artifact from clean signals' components.

The tensor constructed in the proposed method is the outcome of the application of continuous wavelet transform (CWT) in an EEG matrix with a variety of scales. The resulting tensor's modes are therefore $electrode \times samples \times scales$, capturing the spatial, temporal and frequency information of the recorded data. The mother wavelet used in our experiments is Morlet wavelet due to its ability to reconstruct better an EEG signal. Before the application of CWT each channel is normalized using min-max normalization.

Once the decomposition of the tensor takes place the projection matrix corresponding to the second mode, which is the time domain, is examined for artifact related information. The criteria used in the identification of the artifact related components are the same as in [10]. For each temporal signature kurtosis and entropy is extracted. The kurtosis criteria aims in the identification of peaky distributions, expected from an eye blink signature, while the entropy aims in artifacts of a noisy background. Then these features are separately normalized in order to have zero mean and unit variance. A threshold is applied in order to decide which signatures are artifact related. The latter ones are then removed and the reconstruction of the signal follows. Fig. 2 summarizes the process followed by the proposed method.

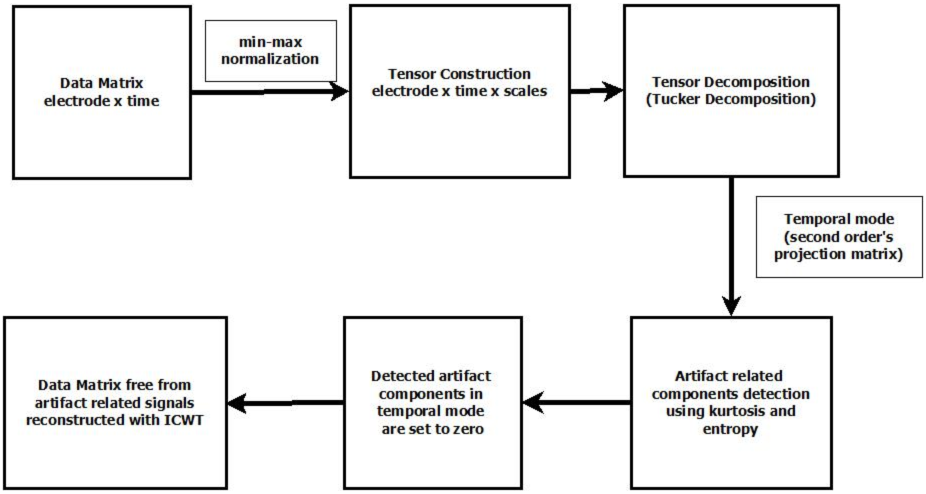

Fig. 2. Flowchart of the proposed method

4 Experiments

The data used in our experiments are from [7]. The recorded data contain intentionally generated artifacts from a healthy 27-year-old male. From the EOG related recording an unmarked epoch was extracted as reference point of artifact free EEG period. A marked as an eye blink period was used as the reference point for artifact related information. The whole set of artifacted period's channels were added to the clean epoch's corresponding ones in order to create a test set with a ground truth data set.

Afterwards the ICA based method, as well as the tensor based one were used in order to remove the artifact related information. It should be mentioned that it is not the same as performing the removal methods in the artifact's period on it's own, since the artifact data contain their clean information also, which will be added to the clean epoch used as references. For better comparison of the two methods we plan to apply the two methods on the artifact epoch alone and then with the help of a medical expert having the results examined.

Once the clean epoch was extracted, one of the marked artifacts was added to the clean epoch's center. The clean epoch's duration was twice the size of the one of the artifact's epoch. The two methods were compared with different thresholds.

Fig.3 shows the average mean square error for both of these techniques. For every method and threshold value combination (thresholds belong in the set (1.2,1.3,1.4...1.8) five experiments of the removal procedure were performed. This error evaluation procedure was followed due to the fact that both methods rely on a random initialization. In more detail the ICA initializes the unmixing matrix randomly, while the Tucker decomposition initializes randomly the projection matrices. The mean square error is calculated based on the equation:

$$MSE = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(CleanEpoch(i,j) - ResultingEpoch(i,j))^2}{n*m} \tag{3}$$

where CleanEpoch is the matrix of the clean epoch and ResultingEpoch is the matrix corresponding to the result of every method, n and m are the number of rows and columns of each matrix respectively.

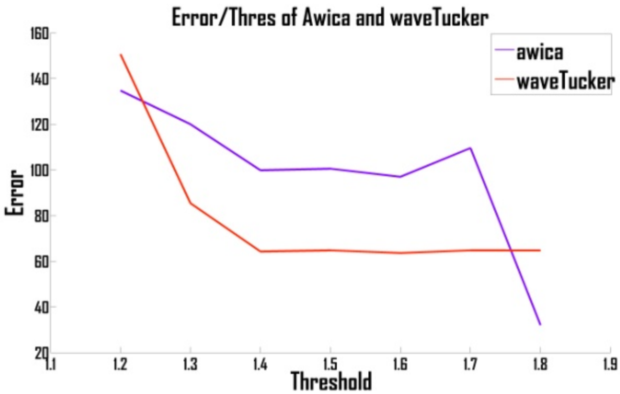

Fig. 3. Tensor based (waveTucker) in comparison with ICA based (awica) for several threshold values.

It is shown that the tensor based technique is more stable than the ICA based and has the lowest error for most of the thresholds. After a certain value, the ICA based method is clearly a better performer but the artifact signal in those thresholds is not cleared completely. It should be noted that for thresholds greater than 1.9 the suspected, for artifact containment, independent components were less than two. Therefore it was not possible to apply the ICA algorithm, leading in not removing the artifact at all. As noted in [10] the preferred value of the threshold for the ICA method is 1.5. Applying the same threshold in the tensor based method had satisfactory results, as proven by our experinments, since the artifact was sufficiently removed without removing a great amount of the rest information in the signal.

Fig.4 and Fig.5 show the two best results of the two methods. Each electrode's signal was normalized using min-max normalization, in order to show the relative amount removed from each method. The channels shown in this figure are part of the whole set of electrodes and are chosen to show each methods drawbacks.

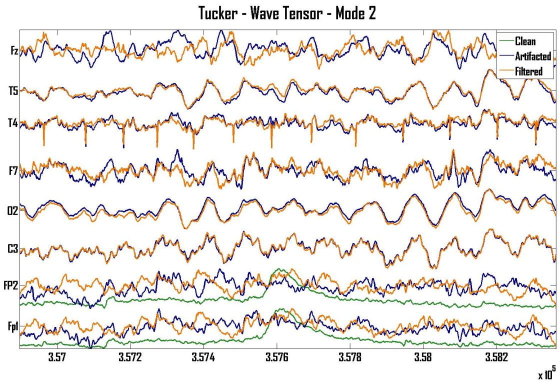

Fig. 4. The result of the tensor-based method. Every channel is affected, since the method is in essence a higher order PCA, but the texture of the reference signal is retained.

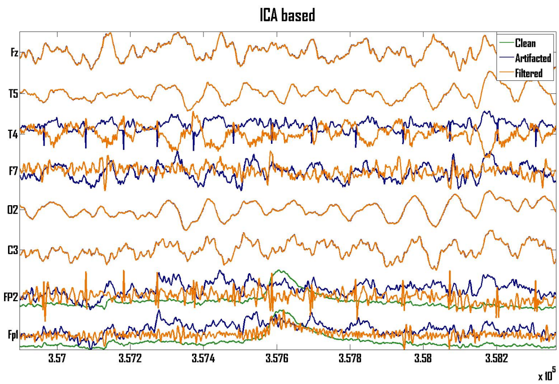

Fig. 5. The ICA-based method. Most of the electrode signals correctly remain intact but the artifact related electrodes, as well as the ones containing some sort of components that fulfilled the removal criteria, have lost the texture of the reference signal.

It should be noted that the artifact exists only in the FP1 and FP2 channels, in the figures showed above. Any other information loss is unacceptable. The ICA based method retains most of the channels unfiltered. However, in the ones where a component was rejected, no texture was retained. By texture we mean the morphology of the clean signals. On the other hand the Tucker based method rejected information from all channels. But the texture of the signals is retained with a certain detail in almost every channel.

5 Discussion

Each method's usage has to be chosen depending on the application where the cleared data will be used on, and it's results. It should be noted that the criteria in the two methods were the same (kurtosis and entropy) and that these criteria play an important role in the results. Improved detection of artifact related components will certainly result in boosting the performance of both methods. For an application requiring the texture of the clean signal to be as less influenced as possible, the usage of the tensor based method is advised. In the case where the artifact suspected independent components are expected to be of small number compared to the amount of electrodes, ICA based methods should be preferred.

Applying spatial and frequency criteria could improve the detection of artifact related components. The creation of spatial, temporal or frequency profiles derived from tensor analysis could be also an interesting direction to be explored.

The goal of this work was to extend the tensor-based techniques for artifact removal by introducing automatic selection of artifact-related components. The results were promising and show that tensors can form the basis for an automatic removal method obtaining comparable results to the ICA based methods.

For future directions we aim to explore the possibilities of success of these methods on removing several other types of artifacts such as muscle artifacts. Using a medical expert's expertise it would be interesting also to see the result of applying these methods directly in an epoch, containing an artifact. The result will in this case be subjective but it will give a guideline of the ability of these methods to work efficiently in an artifact removing procedure during the recording.

Acknowledgements. The reported research was partially supported by the ARMOR Project (FP7-ICT-2011-5.1 - 287720) "Advanced multi-paRametric Monitoring and analysis for diagnosis and Optimal management of epilepsy and Related brain disorders".

References

1. Ochoa, C.J., Polich, J.: P300 and blink instructions. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 111(1), 93–98 (2000), http://www.ncbi.nlm.nih.gov/pubmed/10656515 (retrieved)
2. Verleger, R.: The instruction to refrain from blinking affects auditory P3 and N1 amplitudes. Electroencephalography and Clinical Neurophysiology 78(3), 240–251 (1991), http://www.ncbi.nlm.nih.gov/pubmed/1707797 (retrieved)
3. Delorme, A., Sejnowski, T., Makeig, S.: Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 34(4), 1443–1449 (2007), doi:10.1016/j.neuroimage.2006.11.004

4. Chadwick, N., McMeekin, D., Tele, T.: Classifying eye and head movement artifacts in EEG Signals. In: Digital Ecosystems and Technologies Conference (DEST), vol. 5, pp. 285–291 (2011), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5936640 (retrieved)
5. Lawhern, V., Hairston, W.D., McDowell, K., Westerfield, M., Robbins, K.: Detection and classification of subject-generated artifacts in EEG signals using autoregressive models. Journal of Neuroscience Methods 208(2), 181–189 (2012), doi:10.1016/j.jneumeth(2012).05.017
6. Wu, J., Zhang, J., Yao, L.: An automated detection and correction method of EOG artifacts in EEG-based BCI. Complex Medical Engineering 100875, 1–5 (2009), doi:10.1109/ICCME(2009).4906624
7. De Clercq, W., Vergult, A.: Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. Biomedical Engineering 53(5), 2583–2587 (2006), doi:10.1109/TBME.2006.879459
8. Vergult, A., De Clercq, W., Palmini, A., Vanrumste, B., Dupont, P., Van Huffel, S., Van Paesschen, W.: Improving the interpretation of ictal scalp EEG: BSS-CCA algorithm for muscle artifact removal. Epilepsia 48(5), 950–958 (2007), doi:10.1111/j.1528-1167.2007.01031.x
9. Halder, S., Bensch, M., Mellinger, J., Bogdan, M., Kübler, A., Birbaumer, N., Rosenstiel, W.: Online artifact removal for brain-computer interfaces using support vector machines and blind source separation. Computational Intelligence and Neuroscience 10 (2007), doi:10.1155/2007/82069
10. Mammone, N., la Foresta, F., Marabito, F.C.: Automatic artifact rejection from multichannel scalp EEG by wavelet ICA. IEEE Sensors Journal 12(3), 533–542 (2012), doi:10.1109/JSEN.2011.2115236
11. Sheng, H., Tian, H.: A Novel Method to Remove Eye-Blink Artifacts Based on Correlation Using ICA. In: Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, pp. 1946–1948. IEEE Computer Society Press (2013), doi:10.1109/GreenCom-iThings-CPSCom.2013.362
12. Heute, U., Guzman, A.: Removing “cleaned” eye-blinking artifacts from EEG measurements. Signal Processing and Integrated ..., 576–580 (2014), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6777020 (retrieved)
13. Miwakeichi, F., Martínez-Montes, E., Valdés-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y.: Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis. NeuroImage 22(3), 1035–1045 (2004), doi:10.1016/j.neuroimage.2004.03.039
14. Mørup, M., Hansen, L.K., Herrmann, C.S., Parnas, J., Arnfred, S.M.: Parallel Factor Analysis as an exploratory tool for wavelet transformed event-related EEG. NeuroImage 29(3), 938–947 (2006), doi:10.1016/j.neuroimage.2005.08.005
15. Acar, E., Bingöl, C.A., Bingöl, H., Yener, B.: Computational analysis of epileptic focus localization. In: Ruggiero, C. (ed.) Proceedings of the 24th IASTED International Conference on Biomedical Engineering (BioMed 2006), pp. 317–322. ACTA Press, Anaheim (2006)
16. Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., Yener, B.: Multiway analysis of epilepsy tensors. Bioinformatics s 23(13), i10-8 (2007), doi:10.1093/bioinformatics/btm210

17. Kolda, T.G., Bader, B.W.: Tensor Decompositions and Applications. SIAM Review 51(3), 455–500 (2009), doi:10.1137/07070111X
18. Harshman, R.: Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis (10), 1–84 (1970)
19. Carroll, J., Chang, J.: Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika 35 (1970), http://link.springer.com/article/10.1007/BF02310791 (retreived)
20. Tucker, L.R.: Implications of factor analysis of three-way matrices for measurement of change. In: Harris, C.W. (ed.) Problems in Measuring Change, pp. 122–137. University of Wisconsin Press, Madison (1963)
21. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products (1927)
22. Cattell, R.B.: Parallel proportional profiles and other principles for determining the choice of factors by rotation. Psychometrika 9(4), 267–283 (1944)
23. Bader, B.W., Kolda, T.G.: MATLAB Tensor Toolbox Version 2.5., http://www.sandia.gov/~tgkolda/TensorToolbox/ (accessed)