Blind Source Seperation

July 2019

1 Blind Source Seperation

Blind source separation (BSS) methods are used in signal processing to recover the independent sources from observations of a linear mixture of the sources [3]. BSS use the observations $X(t)$ to generate the unmixing matrix W, which can be used to obtain the independent component signals. With the original sources being the matrix representations of the true signal B and the artifact O, X is a linear combination $X = A\,[B + O]$ with mixing matrix A. Given the observation matrix X, the task is to estimate both the mixing matrix A and the original sources $B + O$. Once the mixing matrix A is known, sources can be recovered from X with the unmixing matrix $W = A^{-1}$ with $B + O = WX$. With the original sources of $X(t)$ having been identified, they can be selected and removed, and the signal is reconstructed without the artifacts to produce the corrected signal $C(t)$ (see figure **??**).

BSS techniques generally require more electrodes than expected signal sources, and more time points than the square of the number of electrodes. Either the length of the recording or the sampling frequency can be increased to satisfy this requirement. BSS techniques fall in the class of unsupervised machine learning and do not need training data sets. In principle, the methods do not automatically identify artifact components, but the analysis can be automated with the use reference channels.

1.1 Second order blind inference

Second order blind inference (SOBI) uses decorrelation across several time points as its main computational step [1]. SOBI considers the relationship between components at different time lags and insists that these are decorrelated as much as possible. First the data is pre-whitened and a set of cross-correlation matrices are calculated at different delays. Using the joint diagonalizer of the cross-correlation matrices, the sources can be recovered. SOBIs ability to resolve correlated activity is essential for ocular artifact detection since artifact signals coming from the two eyes are highly correlated [6].

The matrix of cross-correlations of the measured (whitened) signal $X(t)$ at time-lag τ is defined as

$$R(\tau) = \mathbb{E}[X(t)X(t-\tau)] = \int_{-\infty}^{\infty} X(t)\overline{X}(t-\tau)\,\mathrm{d}t \tag{1}$$

where $\overline{X}$ represents the complex conjugate. The unmixing matrix W is then computed as the matrix that jointly diagonalizes (rotates) a set of p whitened cross-correlation matrices

$$\{R(\tau_i)|i = 1, ..., p\} \tag{2}$$

The diagonalization R_W of a matrix R by matrix W is defined as

$$R_W = WRW^{-1} \tag{3}$$

and is equivalent to the eigen-decomposition of R, where W is the matrix containing the eigen-vectors and R_W the diagonal matrix containing the eigenvalues. For a set of matrices, the joint diagonalizer is therefore a form of an average eigenstructure. Computation time for calculating the joint diagonalizer is directly related to the number p of cross-correlation matrices.

The advantage of SOBI is that the time-delay of artifacts propagating over the scalp is considered in the identification of the components. SOBI is able to identify correlated activity over time if the right time-lags are chosen for computing the cross-correlation matrices. Determining the right lags is related to the expected delay of propagating signals from the front to the back of the scalp. Studies on effects of the chosen delays suggest a set of lags from 1 to 300 ms works best [5] [4]:

$$\begin{aligned}\tau \in \{&1,2,3,4,5,6,7,8,9,10,12,14,16,18,20,25,30,35,40,45,50,55,60,65,\\&70,75,80,85,90,95,100,120,140,160,180,200,220,240,260,280,300\}\end{aligned}$$

This implies that the signals needs to be sampled at $1kHz$ or higher.

SOBI can be automated [2] by identifying components that correlate with components from EOG channel recordings. First, the EEG/EOG data is decomposed into a number of components equal to the number of sensors. Second, the sign on all lower and horizontal EOG channels is inverted (i.e. multiplied by -1) and the new data is also decomposed into components. The new components that invert compared to the old are flagged. In this step the components corresponding to EOG signals that do not propagate far from the EOG recording site are identified. Third, the components that correlate above a certain level with the lower and horizontal EOG data and components with high power in the low frequency band are flagged as well. This step identifies the components containing larger eye-movements and blinks that propagate far and are not inverted in the second step. The idea is that the components containing eye activity will correlate more strongly with the lower and horizontal EOG channels than with non-ocular components, since these EOG channels reflect primarily ocular motions. The correlation threshold is derived from relations between components originating near the eyes, as identified by geometric relationships contained in mixing matrix W^{-1}. Finally, the flagged components are removed from the data by setting the corresponding rows of the source matrix to zero before multiplying with mixing matrix W^{-1}.

References

[1] A. Belouchrani, K. Abed-meraim, J. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–445, 1997.

[2] C.A. Joyce, I. Gorodnitsky, and M. Kutas. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, 41:313–325, 2004.

[3] A. Moreaux and G.D. Iannetti. Across-trial averaging of event-related EEG responses and beyond. *Magnetic Resonance Imaging*, 26:1041–1054, 2008.

[4] M.T. Sutherland, J.Y. Liu, and A.C. Tang. Temporal delays in blind identification of primary sometosensory cortex. *Proceedings of the third international conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004.*, page 4222, 2004.

[5] A.C. Tang, M.T. Sutherland, and C.J. McKinney. Validation of SOBI components from high-density EEG. *NeuroImage*, 25:539–553, 2004.

[6] L. Zhang, J. Kwok, and B.L. Lu. Applications of second order blind identification to high-density EEG-based brain imaging: a review. *ISNN 2010: Advances in Neural Networks*, 2:368–377, 2010.