Classifier-Based Source Localisation in Independent Component Space: Progress Report

Laurens R. Krol[1,*], Mahta Mousavi[2], and Thorsten O. Zander[3]

[1]Biological Psychology and Neuroergonomics, Technische Universität Berlin
[2]Electrical and Computer Engineering, University of California San Diego
[3]Zander Laboratories B.V., Amsterdam, The Netherlands
[*]Correspondence: `lrkrol@gmail.com`

Introduction:

A brain-computer interface (BCI) ideally works solely on *brain* activity, i.e. signals originating from the cortex or other brain structures. Because of its high temporal resolution, relative cost-effectiveness, and portability, many BCI systems use electroencephalography (EEG) to monitor brain activity. However, with EEG in general and in particular with increasingly popular dry and mobile EEG systems (Zander et al., 2017), the recordings invariably also include artefacts, i.e. non-brain signals. It is important to validate BCI classifiers in order to see to what extent its performance relies on artefacts rather than brain activity.

Earlier we proposed a method to source-localise the signals isolated by a spatio-temporal linear discriminant analysis (LDA)-based classifier (Zander, Krol, Birbaumer, & Gramann, 2016). Here, we present work-in-progress that focuses on validating that approach using simulated EEG data. We furthermore propose and discuss an extension of this method for classifiers based on common spatial patterns (CSP).

Material, Methods and Results:

The general approach uses independent component analysis (ICA; Makeig, Jung, Bell, & Sejnowski, 1996) to first obtain a source model of the EEG recording, independent of the classifier. The result of this decomposition is an unmixing matrix that transforms sensor activity into source activity. Following Haufe et al., 2014, the filter weights of a spatio-temporal LDA classifier can be transformed into a forward model. These LDA pattern weights can then be distributed to the independent components (ICs) via the ICA's unmixing matrix. As such, LDA filter weights can be transformed into "relevance weights" in source space, weighting the ICs with their relative contribution to classification (Zander et al., 2016). We apply two additional weighting factors to compensate for 1) LDA's amplitude alignment, and 2) noise representations in the LDA weights.

We are currently simulating EEG data with different numbers of known signal sources among noise, and applying the proposed method to evaluate its accuracy. Figure 1 shows the ground-truth and reconstructed locations of two sources in one time window. The results are based on the simulation of 20 "participants" with 62 randomly chosen noise sources, and 2 relevant sources in the illustrated locations. Visual inspection shows that these sources are accurately identified.

We are testing how many different sources can be detected in the same time window, and what factors influence this number. Furthermore, we are evaluating the influence of the two compensatory weights that we apply, as well as the method's dependence on ICA.

We are also working to extend this method to CSP-based classifiers. After estimating the CSP filters, selected patterns (i.e. those sets of patterns of the eigenvalue decomposition that best describe the classes' variance changes) will be mapped to IC space via the ICA's unmixing

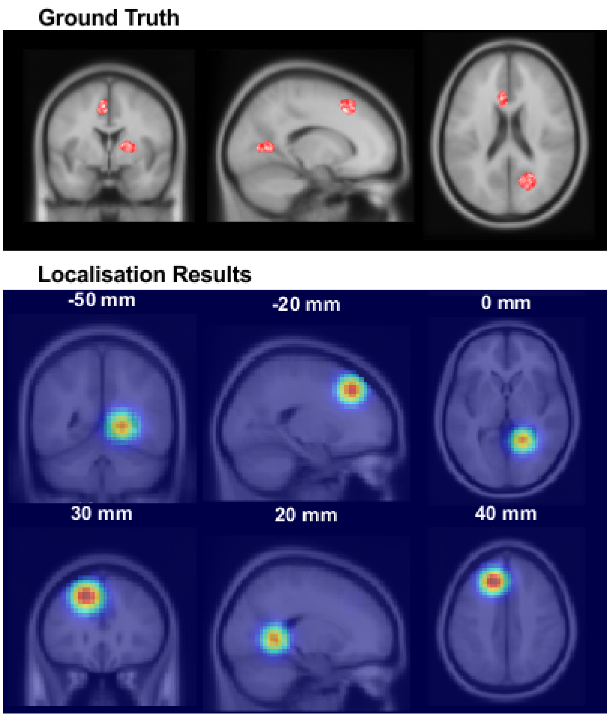

Figure 1: Top: Ground-truth location of two sources with class-dependent activations. Bottom: Reconstructed source locations. Coordinates are in Talairach.

matrix. Since in CSP-based methods, an LDA is trained on the CSP-filtered signals, the weights in source space will be corrected through multiplication by the corresponding LDA weight to account for their relevance to the classifier. As above, potential additional corrections will be investigated. We will verify our methodology through simulations and by using real EEG data from a motor imagery task (Mousavi, Koerner, Zhang, Noh, & de Sa, 2017).

Discussion:

It is valuable to be able to identify the origin of the signals taken into account by a classifier. The proposed method has already produced reliable source localisation results, but given that it uses a number of parameters external to the classifier itself, it must still be determined to what extent these results reflect the exact signals isolated by the classifier. This is ongoing work. The CSP-based method is being developed to generalise our approach to CSP-based classifiers.

Significance:

A source localisation method that identifies the sources used by the exact same classifier as the one used online is valuable to evaluate BCI system performance. Our proposed approach also enables BCI methodology to be used as a novel source-localisation tool for general neuroscientific experiments.

Acknowledgements:

Parts of this work were supported by the Deutsche Forschungsgemeinschaft (ZA 821/3-1), and a DAAD Short-Term Research Grant awarded to MM.

References:

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*(0), 96–110. doi: 10.1016/j.neuroimage.2013.10.067

Makeig, S., Jung, T.-P., Bell, A. J., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in neural information processing systems* (Vol. 8, p. 145-151). Cambridge, MA: MIT Press.

Mousavi, M., Koerner, A. S., Zhang, Q., Noh, E., & de Sa, V. R. (2017). Improving motor imagery BCI with user response to feedback. *Brain-Computer Interfaces*, *4*(1-2), 74-86. doi: 10.1080/2326263X.2017.1303253

Zander, T. O., Andreessen, L. M., Berg, A., Bleuel, M., Pawlitzki, J., Zawallich, L., ... Gramann, K. (2017). Evaluation of a dry EEG system for application of passive brain-computer interfaces in autonomous driving. *Frontiers in Human Neuroscience*, *11*, 78. doi: 10.3389/fnhum.2017.00078

Zander, T. O., Krol, L. R., Birbaumer, N. P., & Gramann, K. (2016). Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences*, *113*(52), 14898–14903. doi: 10.1073/pnas.1605155114