TIME-DEPENDENT DEMIXING OF TASK-RELEVANT EEG SOURCES

N. J. Hill, J. Farquhar, T. N. Lal, B. Schölkopf

Max Planck Institute for Biological Cybernetics, Tübingen, Germany
e-mail: jez@tuebingen.mpg.de

SUMMARY: Given a spatial filtering algorithm that has allowed us to identify task-relevant EEG sources, we present a simple approach for monitoring the activity of these sources while remaining relatively robust to changes in other (task-irrelevant) brain activity. The idea is to keep spatial *patterns* fixed rather than spatial filters, when transferring from training to test sessions or from one time window to another. We show that a fixed spatial pattern (FSP) approach, using a moving-window estimate of signal covariances, can be more robust to non-stationarity than a fixed spatial filter (FSF) approach.

INTRODUCTION

Since EEG data are highly spatially blurred, it is often beneficial to apply a *spatial filtering* algorithm such as Independent Component Analysis (ICA) or the Common Spatial Pattern (CSP) method. Either method may return a (let's assume, square) matrix W such that sources S are estimated from data matrix X (s sensors by t time samples) by premultiplication $\mathrm{S} = \mathrm{WX}$. Each row of W gives us a *spatial filter*, i.e. a vector of sensor weightings for estimating one source signal. We refer to the columns of the mixing matrix $\mathrm{A} = \mathrm{W}^{-1}$ as *spatial patterns*: each one shows, for a given source, that source's relative amplitude as received at the s different sensors.

It is common practice to obtain spatial filters on one set of data X_1 (computing W from the training trials only, using ICA or CSP), infer which sources are relevant to the task, and then apply the corresponding rows of W to some new test data X_2 (perhaps from a subsequent feedback session). A potential drawback is that an optimized spatial filter can only be guaranteed to remain optimal for estimating a given source as long as the spatial patterns of the *other* sources remain constant: changing any column of A may easily affect *all* rows of A^{-1}. So a spatial filter optimized for listening to a particular part of the motor cortex in the presence of, say, prominent frontal-cortex activity may look different from a spatial filter optimized for listening to the same source in the presence of prominent occipital activity. It seems reasonable to hypothesize that, over the course of a motor-imagery BCI experiment, the spatial patterns for relevant sources will change relatively little (we will assume that the positions of the sources in the motor cortex, and the spectral content of the signals they generate, are relatively constant). By contrast we might expect the spatial patterns in the rest of the decomposition to change more significantly, particularly in transfer between training and feedback sessions (when conditions of visual stimulation and general arousal change), but also perhaps with regard to other factors like tiredness, hunger, thirst or cognitive activity. For this reason, we outline a simple approach based on fixed spatial patterns (FSP) rather than fixed spatial filters (FSF).

FIXED SPATIAL PATTERN (FSP) DEMIXING

Both ICA and CSP can be seen as performing a *whitening* or decorrelation, followed by a rotation, in the s-dimensional space of sensors:

$$\begin{aligned} \mathrm{S} &= \mathrm{WX} = \mathrm{RPX} \\ \mathrm{X} &= \mathrm{AS} = \mathrm{P}^{-1}\mathrm{R}^{-1}\mathrm{S}. \end{aligned}$$

The whitening matrix P can be any matrix such that $\mathrm{P}^\top \mathrm{P} = \Sigma^{-1}$, where Σ is the sensor covariance matrix. The rotation matrix R is optimized according to some criterion (class difference in projected variance for CSP, independence of outputs for ICA). Let us assume that we have used one of these methods to estimate P_1 and R_1 from training data X_1, and have partitioned the mixing matrix A_1 into two sets of columns, $\mathrm{A}_1 = [\mathrm{A}_1^{[r]} : \mathrm{A}_1^{[i]}]$ corresponding to the task-relevant and irrelevant sources respectively. We then observe test data X_2 and estimate a new P_2 from it. We now want a new R_2 that will best separate our sources, but under the constraint that the relevant columns of the resulting A_2 be the same as they were in A_1.

As in the spatially constrained ICA (SCICA) approach described in [1], we partition R_2^{-1} into constrained and unconstrained columns, [C : U]. The FSP constraint gives us $\mathrm{C} = \mathrm{P}_2\mathrm{A}_1^{[r]}$. Since it is unlikely that $\mathrm{P}_1 = \mathrm{P}_2$, we cannot assume that columns C are orthonormal. However, like [1] we will assume that C and U occupy orthogonal subspaces. This allows us to write R_2 as a vertical concatenation of pseudoinverses, to obtain:

$$\mathrm{W}_2 = \mathrm{R}_2\mathrm{P}_2 = [\mathrm{C} : \mathrm{U}]^{-1}\mathrm{P}_2 = \begin{bmatrix} (\mathrm{C}^\top\mathrm{C})^{-1}\mathrm{C}^\top \\ (\mathrm{U}^\top\mathrm{U})^{-1}\mathrm{U}^\top \end{bmatrix} \mathrm{P}_2.$$

If, like [1, 2], we were using this technique to correct the EEG for *artifacts* with known spatial patterns $\mathrm{A}_1^{[r]}$, we would then have to proceed to optimize the U (making the further assumption that columns U are orthonormal) in order to estimate the remaining sources. However, since we have already decided that the remaining sources are irrelevant, we can ignore the lower rows of W_2 and hence U. Substituting for C, we obtain:

$$W_2^{[r]} = (A_1^{[r]\top} \Sigma_2^{-1} A_1^{[r]})^{-1} A_1^{[r]\top} \Sigma_2^{-1} \quad (1)$$

This simple formula requires only the fixed spatial patterns $A_1^{[r]}$ and a new estimate Σ_2 of the covariance of the sensor signals from which we want to extract the corresponding sources.

DEMONSTRATION

We present a preliminary illustration that this approach can make motor-imagery BCI classification more robust to changes in task-irrelevant brain activity. We use 7 two-class data sets. The first is the 118-channel EEG dataset IVc from BCI Competition III: we took the 500–1500 msec interval of each trial in both training and test set, with the 0 class removed, resulting in 210 training trials and 280 test trials of left-hand/foot motor imagery. The other 6 are imagined left/right hand movement data sets from our lab, each consisting of 400 trials of 39-channel EEG. We use the first 200 as training and the second 200 as test points.

First, we perform ordinary CSP on the training trials with a wide-band (7–30 Hz) temporal filter. We invert the full s-by-s filter matrix and keep the first 4 and last 4 spatial patterns as our $A_1^{[r]}$. Next, we track the activity of the sources associated with these 8 patterns throughout the whole data set (training and test trials). For each trial i, we obtain spatial filters W_i using equation (1) with a moving estimate of the covariance: each Σ_i is obtained from the last n trials including the current one, i.e. trials $(i - n + 1) \ldots i$. After applying the spatial filter, we compute the log amplitude spectrum using the Welch's short-time Fourier transform method. We then normalize the vector of amplitude features for each trial and source. Using this feature set, we then classify using a linear Support Vector Machine, finding the regularization parameter by 10-fold cross-validation within the training set.

The one hyperparameter that needs to be set is the size of the moving window, n. In practice this could be found by cross-validation, or perhaps by adding a known artificial signal to the data, in a known artificial spatial pattern, and empirically determining the value of n that allows it to be recovered most accurately. Here, we simply present the results for each of a range of values, to see its effect on test set performance.

In Figure 1, filled symbols show the results for data set IVc (circles) and for the average of the 6 subjects in the other study (triangles)—the individual subject results were broadly consistent with the average, but we do not have space to show them individually here. The dashed lines show, for comparison, the performance of a fixed spatial *filter* approach analogous to ordinary CSP-based methods: the W_i were simply the original filters found using the training trials, held constant for all trials. We can see that, for a sufficient window size, say $n \geq 20$, the moving-window FSP approach does not perform significantly better or worse than the FSF approach.

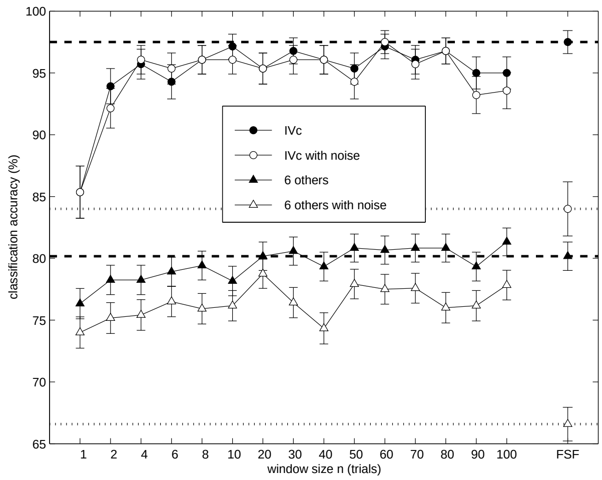

FIGURE 1: Performance of FSP and FSF methods.

At this preliminary stage we cannot say whether the algorithm is not sensitive enough to non-stationarities in the data to improve performance, or whether these particular data sets do not suffer from a significant non-stationarity problem in the first place (our six data sets were all single-session without feedback). However, we can demonstrate the moving-window FSP approach's robustness to non-stationarity by *introducing* non-stationarity into the data. In a second set of tests, we added two Gaussian noise sources to the test trials only. This introduces a difference in the training and test distributions, resulting in clear problems for the FSF method (dotted lines). Both artificial noise sources had fixed spatial patterns (chosen randomly), but their amplitudes drifted over time: one increased linearly from $a/2$ to $2a$ over the course of the entire test set, and the other decreased from $2a$ down to $a/2$, with a chosen such that the FSF method suffered about a 10–15% degradation in performance. Open symbols show performance on the noisy data. Comparison of the filled and open symbols shows that the introduction of non-stationary noise into the test set did not greatly affect the moving-window FSP method's performance (hardly at all for some subjects, like IVc), and hence it performed better than the FSF approach for nearly all values of n. This suggests that it is a promising candidate for dealing with non-stationarities in EEG data, although a wider range of data sets will be required in order to see whether it is effective at coping with the kind of non-stationarities that occur in reality.

REFERENCES

[1] Hesse C. W and James C. J. The FastICA algorithm with spatial constraints. *IEEE Sig. Proc. Letters* **12**, 792–795 (2005).

[2] Ille N, Berg P, and Scherg M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *J. Clin. Neurophysiol.* **19**(2), 113–124 (2002).