EEG SOURCE IMAGING ASSISTS DECODING IN A FACE RECOGNITION TASK

Rasmus S. Andersen, Anders U. Eliasen*, Nicolai Pedersen*,*
Michael Riis Andersen, Sofie Therese Hansen, Lars Kai Hansen

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads B324, DK-2800, Kgs. Lyngby

ABSTRACT

EEG based brain state decoding has numerous applications. State of the art decoding is based on processing of the multivariate sensor space signal, however evidence is mounting that EEG source reconstruction can assist decoding. EEG source imaging leads to high-dimensional representations and rather strong a priori information must be invoked. Recent work by Edelman et al. (2016) has demonstrated that introduction of a spatially focal source space representation can improve decoding of motor imagery. In this work we explore the generality of Edelman et al. hypothesis by considering decoding of face recognition. This task concerns the differentiation of brain responses to images of faces and scrambled faces and poses a rather difficult decoding problem at the single trial level. We implement the pipeline using spatially focused features and show that this approach is challenged and source imaging does not lead to an improved decoding. We design a distributed pipeline in which the classifier has access to brain wide features which in turn does lead to a 15% reduction in the error rate using source space features. Hence, our work presents supporting evidence for the hypothesis that source imaging improves decoding.

Index Terms— Brain state decoding, Brain computer interface, BCI, EEG source imaging, Face recognition.

1. INTRODUCTION

Brain state decoding based on electroencephalography (EEG) has numerous applications including neurofeedback [1], brain computer interfacing [2], and patient monitoring [3]. State of the art decoding is based on signal detection in the multivariate sensor space signal see e.g., [4], however evidence is mounting that EEG source reconstruction can assist decoding [5]. 'Raw' EEG sensor signals are confounded by non-brain signals such as muscle activation, eye blinks and eye motion [2] and furthermore blurred by so-called volume conduction effects. It is possible that these confounding effects may be suppressed by EEG source imaging.

Inference of a high-dimensional EEG source distribution $N = 10^3 - 10^4$, from a relatively low dimensional sensor space measurement $M = 10 - 10^2$, is mathematically ill-posed and requires prior information such as anatomical, functional or mathematical constraints to isolate a unique and plausible solution [6]. The key prior information invoked in Edelman et al. is to introduce spatially focused features relevant for detection of motor imagery [5]. The aim of our study is to test the generality of the hypothesis by applying the approach to a new experimental setting. We have therefore adapted the Edelman et al. pipeline to the present task and denote it the *focal pipeline*, see figure 1. It turns out that with this pipeline source imaging does not lead to an improvement in the present data. Instead we suggest to reduce the complexity of the source space representation using a distributed set of spatial basis functions and show that not only does this approach lead to improved decoding error rate it also performs better than a similar pipeline based on raw sensor space signals. The new pipeline is denoted the *distributed pipeline*, see figure 1. The rest of the paper is organized as follows: Section 2 presents the focal and distributed pipelines. Section 3 summarizes our results, section 4 provides a discussion of the pipelines and their respective results, and section 5 presents our conclusions.

2. EXPERIMENTAL DESIGN AND SETUP

While the study of Edelman et al. concerned motor imagery, we here investigate a response known to be distributed with several foci, namely face recognition [7]. The benchmark EEG data set analyzed in this study was collected by D. Wakeman and R. Henson and is distributed with the Statistical Parametric Mapping (SPM) toolbox. The data set contains two runs of a subject viewing 86 face images and 86 scrambled face images producing a total of 344 trials. [7, 8, 9]. The Henson and Wakeman data set was acquired using a 128-channel ActiveTwo system, sampling at 2048 Hz [7, 8]. Two bipolar channels measuring Horizontal and Vertical Electrooculography (HEOG and VEOG) were removed. The data was preprocessed according to the suggestions from the SPM8 manual [9]. This involved a downsampling to 200
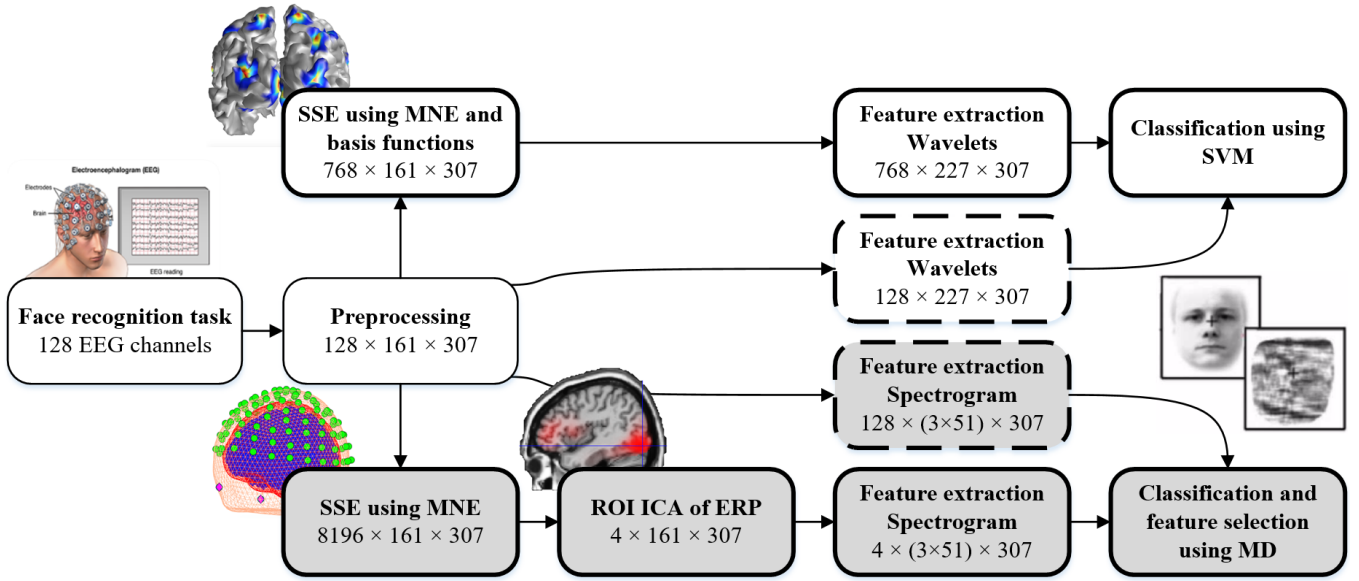
Fig. 1: Flowchart for the investigated pipelines in sensor- and source spaces. Source space estimates (SSE) are obtained using the minimum norm estimator (MNE) [10]. The grey path illustrates the approach for the focal pipeline, and the white path indicates the process of the distributed pipeline. The dimensions in each square are organized as follows $M, N \times T \times Q$. M and N represent the number of sensor channels and dipoles/basis functions respectively, T is the length of the epoch time series, and Q is the sample size of 'epochs' [9].

Hz and a re-referencing to 'average reference'. The EEG data was then segmented into epochs of 0.8 s (0.2 s before to 0.6 s after onset) thus including the entire length of the visual stimuli. Lastly artifact rejection was performed using a threshold approach, which excluded epochs containing channels with amplitudes larger than $200 \ \mu V$. The resulting data set contains 307 epochs each consisting of 161 samples, of which 153 are trials with face stimuli and 154 epochs are trials with scrambled face stimuli. Projecting the data from sensor- to source space requires knowledge about the electrical conductance between the scalp sampling sites and the modelled dipole sources of the brain, see figure 2. Assuming a linear relationship between a given source distribution inside the brain and the electric potentials measured at the scalp [11], this relationship - 'the EEG forward problem' - can be formulated as $\mathbf{B} = \mathbf{AX} + \mathbf{E}$, where $\mathbf{B}$ is the $M \times T$ sensor space signal, $\mathbf{A}$ is the $M \times N$ lead field matrix containing the head volume conductor model, $\mathbf{X}$ is the source $N \times T$ space and $\mathbf{E}$ is additive noise. Because the number of dipoles is typically much larger than the number of electrodes the inverse problem is mathematically ill-posed and strong assumptions are needed to regularize the solution, e.g., of the form $\min_X ||\mathbf{AX} - \mathbf{B}||^2 + \lambda ||\mathbf{WX}||^2$, where λ is a regularization parameter and $\mathbf{W}$ is a spatial matrix that can be designed to promote spatial smoothness of the solution. The solution is given by $\hat{\mathbf{X}} = (\mathbf{A}^T \mathbf{A} + \lambda \cdot \mathbf{W})^{-1} \mathbf{A}^T \mathbf{B}$. When $\mathbf{W}$ is the identity matrix, the solution is often referred to as the 'minimum norm estimator' (MNE) [10].
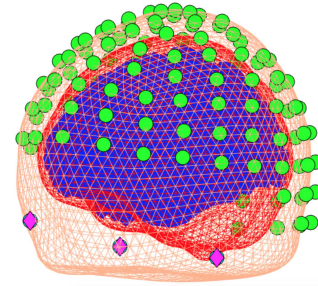

Fig. 2: Overview of the geometry of source space estimation. The green spheres on the skull represent the EEG electrodes, the purple diamond shaped objects are fiducials (nasion left and right preauricular) for cap alignment. The inner pyramidal mesh illustrates the dipole model of the brain [9].

Following [5] a spatial region of interest (ROI) was derived using a data-driven approach. To identify the ROI for the face recognition task independent component analysis was performed on the event related potential (ERP) for each of the two classes (mean response to faces or scrambled faces) using the extended InfoMax algorithm [12]. The independent components were ranked according to how well they separated the two classes based on the difference in their respective spectrogram. The single independent component that displayed the largest separation of the two classes in the time-frequency space was selected. The topography of the selected component was then mapped onto a template brain
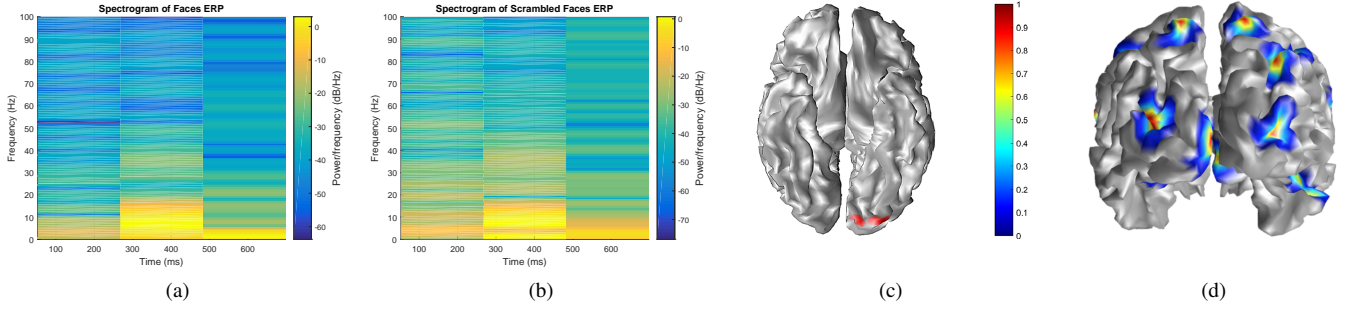
Fig. 3: (a) shows the spectrogram of the average responses (ERP) for the face trials. (b) spectrogram for the ERP of the trials consisting of scrambled faces. Time and frequency bins of the spectrograms correspond as used in the focal pipeline. (c) is the focal pipeline's ROI scalp map, with the selected dipoles highlighted. (d) displays ten (of 768) basis functions and their distribution across the cortical surface

using the MNE with λ tuned by the L-curve method [13]. The source space representation was thresholded at 75% of the peak value to identify dipole locations on the cortex to be included in the ROI. Also, following [5] the reference sensor space analysis features were extracted from all electrode time series while for the source space analysis time-series from dipoles within the ROI were included. A low resolution time frequency representation of each time-series was obtained based on three time windows and 51 frequency bins in the range of $0 - 100$ Hz, see figure 3. The most discriminating features were found using the Mahalanobis distance (MD). Using 10-fold cross validation the data was split into a training- and a test set. A forward feature selection scheme was then applied on the training set to optimize the MD. Figure 4 illustrates the error rate as a function of the number
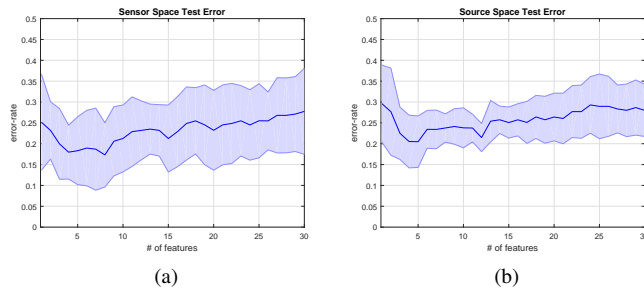

Fig. 4: (a) illustrates the result of the feature selection for the sensor space, with the optimal number of features being 8. (b) shows the result of the source space feature selection, where the optimal number of features were 5.

of features. The lowest error rate in the sensor space was achieved with 8 features, whereas the lowest error rate in source space was found using 5 features. Having found the optimal feature set, the test set was classified using an epoch based MD-classifier. The mean and covariance matrices were calculated for both classes based on the training set and used to compute the MD to the class means for each test epoch [5].

For more complex distributed tasks the focal hypothesis may be too restrictive. We therefore suggest an alternative route to simplify the source space representation. We will apply the basis function approach of Friston et al. [14], see figure 3(d). The result of this procedure is a reduced and smoothed representation of the dipole activity. The advantage of this approach is that the entire cortical mesh is spanned by a smaller feature set, in case, a source space representation reduced from 8196 dipole sources to 768 basis functions. We used the MNE estimator with the basis function representation and parameter λ was optimized in cross-validation. Dynamical features were obtained using a wavelet decomposition in the distributed pipeline. A Daubechies 4 (db4) mother wavelet was used, see e.g., Bahram et al. [15]. Feature extraction was performed for each epoch, viz., ten detail bands were extracted for each of the 768 basis functions and concatenated into one feature vector of 227 entries. All individual feature vectors $\mathbf{f}$ were further combined into a feature matrix, consisting of 174336 features per epoch, i.e., the final data matrix for the distributed pipeline was 174336x307. The support vector machine (SVM) is efficient for high-dimensional binary classification tasks [16] and computes $y(\mathbf{f}) = \sum_{q=1}^{Q} \alpha_q t_q k(\mathbf{f}, \mathbf{f}_q) + b$, where $k(\mathbf{f}, \mathbf{f}_q) = \mathbf{f}^T \cdot \mathbf{f_q}$ is the linear kernel and α_q are learned SVM parameters for Q data points with targets $t_q = \pm 1$ for face/scrambled face trials. The class label of test observations are obtained as $\text{sign}(y(\mathbf{f}))$. The performance of the SVM classifier is evaluated using 10-fold cross validation.

3. RESULTS

Two different sensor space and two different source space approaches - the focal and the distributed pipelines - were applied for decoding of the face recognition tasks. The results

obtained are summarized and compared in table 1.

Pipeline	Sensor space	Source space
Focal	17.3%	20.5%
Distributed	10.7%	9.1%

Table 1: Sensor- and source space test error rate for both pipelines. All the results were obtained using 10-fold cross-validation.

The focal pipeline achieved an error rate of 17.32% in the sensor space, using 128 electrodes and the 8 most prominent features extracted from the spectrogram and using the MD classifier. Classification in source space achieved an error rate of 20.48% using 4 dipole sources and the 5 most prominent features. The source space representation did not improve the classification performance for the focal pipeline for the face recognition task. The distributed pipeline did see a reduced error rate from 10.7% in sensor space to 9.1% in the source space. The sensor space was comprised of 128 electrodes, where ten db4 detail bands were extracted from each electrode. The source space was spanned by 768 basis functions, also with ten db4 detail bands extracted for each basis function.

4. DISCUSSION

Two different classification pipelines were applied in our study, and for both we compared performances for their sensor space and source space variants. The first approach was based on the *focal* spatial approach proposed in Edelman et al. [5]. The pipeline uses a Mahalanobis distance classifier and a small subset of the available dipole sources for classification. Even based on a small number of dipoles feature selection is still critical. Using a forward feature selection the best results were obtained using five features for the source space version and eight features for the sensor space version. This data-driven ROI approach succeeded in detecting the main activity in the occipital region, a region that is known to participate in face recognition [17]. Our alternative *distributed* pipeline employs a much larger feature space, spanning the entire cortical mesh using 768 basis functions. The distributed pipeline used a linear support vector machine for decoding, which by margin optimization is somewhat resilient to over-fitting. While the use of the SVM avoided feature selection we did find the typical signature of variance inflation, see figure 5, however due the balanced classes in the present task this has limited effects on the classification performance [18]. An analysis of the misclassified epochs in the distributed pipeline showed that 74 % of the misclassified epochs in the source space representation were also misclassified in the sensor

space, indicating that these epochs are indeed hard to decode, presumably due to the generally poor signal to noise conditions of EEG.
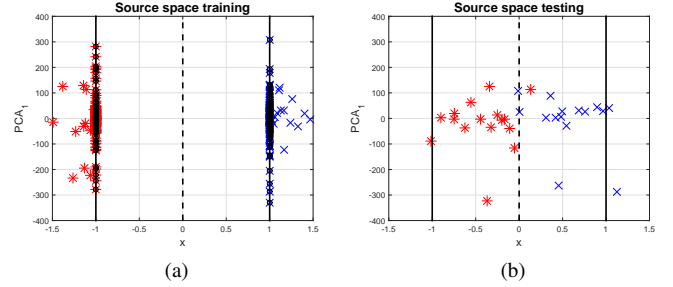

(a)   (b)

Fig. 5: Illustration of the classification problem in the source space. Both plots show one fold in a 10-fold cross-validation. The x-axis is the perpendicular distance to the decision boundary and the y-axis is the first principal component. (a) shows perfect separation of classes in the training set. (b) shows the classification of the resulting test set. With the present high-dimensional feature representation and relatively limited sample size, *variance inflation* is pronounced in the training set [18], however with the face recognition task's balanced classes this only leads to limited performance loss.

Figure 5 shows the training and test classification problem for one of the cross-validation folds in the source space. A perfect separation in the training set is achieved, although the test set is less separated, which is likely due to variance inflation. Despite the variance inflation, which can cause generalization problems for the SVM, the achieved test error of 9.1 % suggests that the classifier avoids some of these problems in this study.

5. CONCLUSION

Brain state decoding is important to many applications and new evidence was recently put forward by Edelman et al. that EEG source imaging can support decoding [5]. The proposed pipeline involves the use of a spatially focal source space representation. The aim of our study was test the generality of this approach. It was not possible to show source space improvements for the decoding in the face recognition task with the focal pipeline. An alternative distributed pipeline was proposed, which did achieve a lower error rate using source decoding reducing the error rate by 15%: from 10.7% error in the sensor space to 9.1% in the source space representation. Hence, our result supports the overall hypothesis of Edelman et al. namely that a carefully crafted source space representations can improve decoding performance.

6. REFERENCES

[1] Martijn Arns, Hartmut Heinrich, and Ute Strehl, "Evaluation of neurofeedback in adhd: the long and winding road," *Biological psychology*, vol. 95, pp. 108–115, 2014.

[2] Mikhail A Lebedev and Miguel AL Nicolelis, "Brain–machine interfaces: past, present and future," *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.

[3] Jan Claassen, Fabio S Taccone, Peter Horn, Martin Holtkamp, Nino Stocchetti, and Mauro Oddo, "Recommendations on the use of eeg monitoring in critically ill patients: consensus statement from the neurointensive care section of the esicm," *Intensive care medicine*, vol. 39, no. 8, pp. 1337–1351, 2013.

[4] Elisa Mira Holz, Loic Botrel, Tobias Kaufmann, and Andrea Kübler, "Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study," *Archives of physical medicine and rehabilitation*, vol. 96, no. 3, pp. S16–S26, 2015.

[5] Bradley J Edelman, Bryan Baxter, and Bin He, "Eeg source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 4–14, 2016.

[6] Sylvain Baillet, John C Mosher, and Richard M Leahy, "Electromagnetic brain mapping," *Ieee Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.

[7] Richard N Henson, Daniel G Wakeman, Vladimir Litvak, and Karl J Friston, "A parametric empirical bayesian framework for the eeg/meg inverse problem: generative models for multi-subject and multi-modal integration," *Frontiers in human neuroscience*, vol. 5, pp. 76, 2011.

[8] Daniel G Wakeman and Richard N Henson, "A multi-subject, multi-modal human neuroimaging dataset," *Scientific data*, vol. 2, 2015.

[9] John Ashburner et al, "The fil methods group at ucl (2013) smp8 manual," *Insitute of Neurology, University College London, London. Available online at http://www.fil.ion.ucl.ac.uk/spm/*, 2013.

[10] Matti S Hämäläinen and Risto J Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical & biological engineering & computing*, vol. 32, no. 1, pp. 35–42, 1994.

[11] Olaf Hauk, "Keep it simple: A case for using classical minimum norm estimation in the analysis of EEG and MEG data," *NeuroImage*, vol. 21, no. 4, pp. 1612–1621, 2004.

[12] Jean-Francois Cardoso, "Infomax and maximum likelihood for blind source separation," *Ieee Signal Processing Magazine*, vol. 4, no. 4, pp. 112–114, 1997.

[13] Per Christian Hansen, *The L-curve and its use in the numerical treatment of inverse problems*, IMM, Department of Mathematical Modelling, Technical Universityof Denmark, 1999.

[14] Karl Friston, Lee Harrison, Jean Daunizeau, Stefan Kiebel, Christophe Phillips, Nelson Trujillo-Barreto, Richard Henson, Guillaume Flandin, and Jérémie Mattout, "Multiple sparse priors for the m/eeg inverse problem," *NeuroImage*, vol. 39, no. 3, pp. 1104–1120, 2008.

[15] Bahram Perseh and Ahmad R Sharafat, "An efficient p300-based bci using wavelet features and ibpso-based channel selection," *Journal of medical signals and sensors*, vol. 2, no. 3, pp. 128, 2012.

[16] Christopher M Bishop, *Pattern recognition and machine learning*, Springer-Verlag New York, 2006.

[17] Nancy Kanwisher and Galit Yovel, "The fusiform face area: a cortical region specialized for the perception of faces," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 361, no. 1476, pp. 2109–2128, 2006.

[18] Trine Julie Abrahamsen and Lars Kai Hansen, "Restoring the generalizability of svm based decoding in high dimensional neuroimage data," in *Machine Learning and Interpretation in Neuroimaging*, pp. 256–263. Springer, 2012.