

Eigenbrains: the Free Vibrational Modes of the Brain as a New Representation for EEG

Riccardo Poli, Luca Citi, Mathew Salvaris, Caterina Cinel, and Francisco Sepulveda

Abstract—We present a new transform for EEG signals whose basis functions are well suited to represent the large-scale dynamics associated with event related potentials. The method involves instantiating an approximate model of the electrical properties of the brain as a conductor medium and then studying the free vibrational modes of the model. These form a set of basis functions, which we call *eigenbrains*, that can be used to meaningfully re-represent the brain’s electrical activity. Eigenbrains are compared to principal component analysis and independent component analysis to highlight differences and similarities.

I. INTRODUCTION

Various linear transformations and basis changes on EEG signals have been proposed in the literature to combine different channels and to extract meaningful components. The most widespread ones are Principal Component Analysis (PCA) and Independent Component Analysis (ICA).

PCA has been used as a tool for the analysis of EEG and Event Related Potentials (ERPs) since the mid sixties [1]–[3]. PCA is based on the idea that the data are a linear combination of “principal components” which need to be identified. PCA components are orthogonal and they maximally account for the variance present in the data. Because of this, it is often possible to accurately represent the original data with a small set of components. Spatial PCA is used in ERP analysis to find components that represent the covariance in the measurements taken at different electrodes, typically measured over multiple epochs. These components are obtained by extracting the eigenvalues and eigenvectors of the covariance matrix.

More recently ICA [4] has seen considerable popularity in EEG and ERP analysis [5]–[7]. If a set of signals is the result of linearly superimposing statistically independent sources, ICA can decompose the signals into their primitive sources or “independent components”. When ICA is applied to the signals recorded at different electrodes on the scalp, it can separate important sources of EEG and ERP variability. This can then be exploited, for example, for removing artifacts.

Both PCA and ICA rely on some assumptions. PCA is linear (it assumes that channels do not interact), it assumes that the major sources of variance (the principal components) are orthogonal, and it implicitly assumes that only the amplitudes of brain activity varies, not its spatial position. ICA, too, assumes the linearity of the brain as a conduction medium. Furthermore it assumes the statistical independence

of the sources of electrical activity, the fixed position of such sources, the absence of conduction delays in the brain, and the non-Gaussianity of the statistical distributions of the sources.

In this paper we introduce a new representation for EEG signals based on a set of functions, which we call *eigenbrains*, that are particularly suitable to represent the large-scale dynamics associated with ERPs. The method has some similarity with PCA in that eigenbrains are the eigenvectors of a matrix, they are orthogonal and we assume linearity. However, unlike for PCA, we formulate an approximate model of the brain as a collection of coupled harmonic oscillators, and the matrix is its representation. The eigenbrains represent the *free vibrational modes* of the model.

The paper is organised as follows. In Section II we present the eigenbrains technique. We apply it to experimental data and compare it with PCA and ICA in Section III. We draw some conclusions in Section IV.

II. EIGENBRAINS

PCA assesses how two variables are related via their covariance. However, covariance is not the only way to represent interactions. In EEG, variables represent the voltage measured at different electrodes. So, it would make sense to assess the degree of interaction between variables in terms of similarity or dissimilarity between voltages. This is, rather obviously, naturally represented by voltage differences. When the magnitude of the difference between the voltages measured at two electrodes is consistently small, we can infer that the portion of the brain being traversed by currents flowing from one electrode to the other has very high conductance.¹ Conversely, when the magnitude of the difference in voltage between a pair of electrodes is consistently large, we can infer that there is a low conductance between them.

In the eigenbrains technique, we assess the strength of the interaction between variables in exactly these terms (with a minor correction to avoid divisions by zero). More precisely, let $s_i(t)$ be the time-varying signal at channel i and $s_{ij}(t) = s_i(t) - s_j(t)$. Also, let $W = (w_{ij})$ be the matrix of pairwise interaction strengths associated with the model. We set

$$w_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 1/(0.001 + f(s_{ij})) & \text{otherwise,} \end{cases} \quad (1)$$

¹Note that this is a model of behaviour, i.e., we do not try to distinguish between conductance due to the physical proximity between the electrodes and conductance due to long distance connections between brain regions.

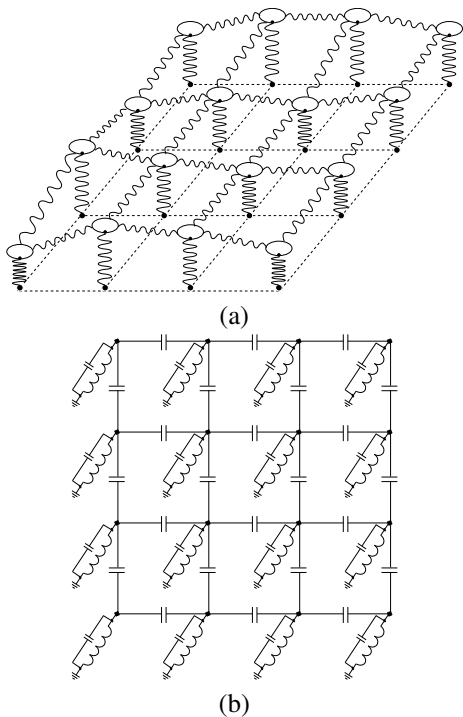


Fig. 1. Mechanical model of the brain (a) and its electrical analogue (b). For clarity, only connections between neighbouring nodes are shown here. However, in general, the models are fully connected networks.

where $f(s_{ij})$ is the mean absolute value or the root mean squared value of s_{ij} , although, for space limitations, in this paper we will only report results for the former, i.e.,

$$f(s_{ij}) = \frac{1}{T} \sum_{t=0}^{T-1} |s_{ij}(t)|. \quad (2)$$

We can interpret this with either a mechanical or an electrical analogy. Embracing the former, we can imagine the EEG electrodes as masses organised in a mesh. Each mass can slide up or down along the vertical direction (i.e., orthogonally to the mesh) and is connected to other masses (electrodes) by springs as illustrated in Fig. 1(a). The matrix W represents the stiffness of the springs. Using an electrical analogy, instead, we can imagine that the EEG electrodes are the nodes in an LC electrical circuit such as the one shown in Fig. 1(b). The nodes are connected to ground via a capacitor and an inductor and between them via capacitors.

With either analogy, the eigenbrains, i.e., the eigenvectors of the matrix W , are the free vibrational models of the corresponding model of the brain.

III. EXPERIMENTATION

We applied our approach to data collected in an experiment related to the development of a brain-computer interface. In this section we describe the experiment and our results.

A. Participants

We used data from 3 participants: A aged 28 (male), B aged 25 (male) and C aged 21 (female). Participants were right handed and with normal or corrected to normal vision.

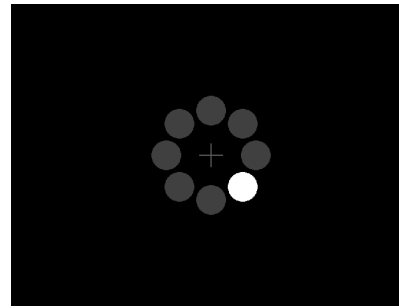


Fig. 2. Display used in our experiments. The eight circles flash in random order. The picture shows the display when a stimulus is presented. Subjects focus their attention on the flashes of one of the circles.

B. Procedure and Apparatus

Eight grey circles and a fixation cross are shown on an LCD screen. At 200 ms intervals, one of the circles is flashed, as shown in Fig. 2. The flash lasts for 100 ms. Circles are randomly flashed. However, after flashing, a circle is not allowed to flash again until all other circles have flashed. We term this a *cycle*. To limit the risks of perceptual errors (such as repetition blindness) occurring, the last circle to flash in a cycle is not allowed to be the first in the following cycle.

Participants were comfortably seated on an armchair with their eyes at approximately 80 cm from the screen. Data were collected from 64 electrode sites using a BioSemi ActiveTwo EEG system. The EEG channels were referenced to the mean of the electrodes placed on either earlobe. The data were initially sampled at 2048 Hz and then sub-sampled at 256 Hz.

Participants were instructed to focus their attention on a particular circle on the screen. To facilitate this, they were asked to count how many times that particular circle flashed during a sequence of stimulus presentations (a block). Each run of our experiments involved 8 blocks, one for each of the eight possible target circles. Each block contained between 20 and 24 cycles. Each participant performed two runs, i.e., a total of approximately 2,800 stimulus flashes.

C. Results

Fig. 3(a) shows the six eigenbrains (i.e., the eigenvectors of W) corresponding to the largest (in magnitude) eigenvalues of the simplified brain model obtained using Eqs. (1) and (2). Let us briefly analyse them. E0 represents a patterns of activation where the anterior pole frontal region and the centro-parietal regions are both active, but one is positive while the other is negative. E1 represents the same counter-activation pattern but for the frontal electrodes and the parieto-occipital region. E2 represents the activation of the central region. E3 represents counter-lateralised activation. E4 and E5 represent second order vibrational modes where multiple areas are jointly active or inactive at the same time.

We can project the EEG channels along some of the eigenbrains (via scalar product) to see which vibrational modes are active at different times in the ERP produced by a stimulus. In Fig. 3(b) we show the mean of the projections of our epochs along E0–E3 for target (target0–target3) and non-target stimuli (ntarget0–ntarget3). Note that

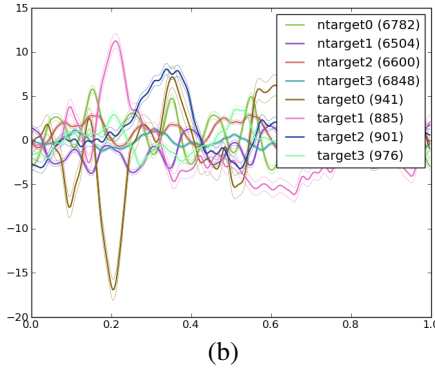
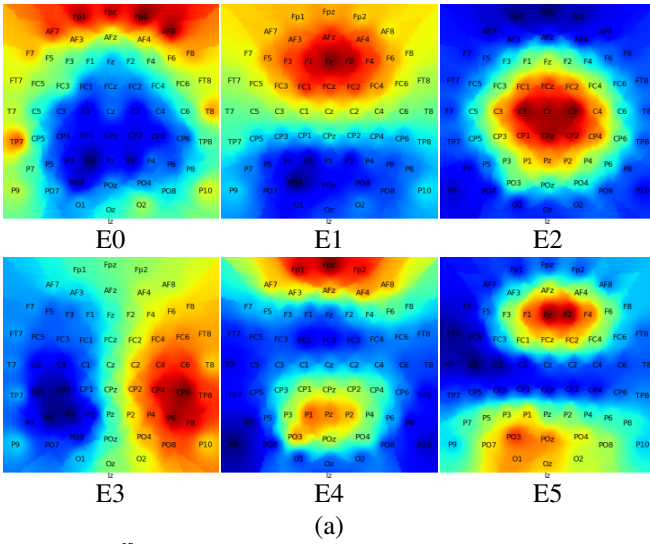


Fig. 3. Eigenbrains: (a) eigenbrains corresponding to the six eigenvalues of W with the largest magnitude and (b) temporal variations of the projections of the EEG along the first 4 eigenbrains averaged across targets and non-targets.

we see significant differences in the projections for targets and non-targets corresponding to the N1, N2 and P3 ERP components. In particular, E2 appears to be a P3 detector, which was expected given its on-centre off-surround shape.

For comparison we also ran ICA (using JadeR) on the same data set. Fig. 4(a) shows the 6 most important independent components found by ICA. These appear to be much less symmetric and large-scale than the eigenbrains. Also, they are much harder to explain in terms of their associated brain activity than the eigenbrains. As shown in Fig. 4(b), there is a good separation between the projections for targets and non-targets in some time windows. However, projections represent multiple potentials more often than the eigenbrains (e.g., I2 picks up N1, N2, P3 and a late action-planning potential).

For further comparison, we have also applied PCA to the same data. In Fig. 5(a) we show the six most representative eigenvectors produced via PCA and in Fig. 5(b) we show the projection of the EEG channels along the first four of these vectors. As one can see, while a couple of the components are reasonably symmetric and similar to our eigenbrains, the other components seem to be focused on specific electrodes and look more similar to those of ICA. The projections in

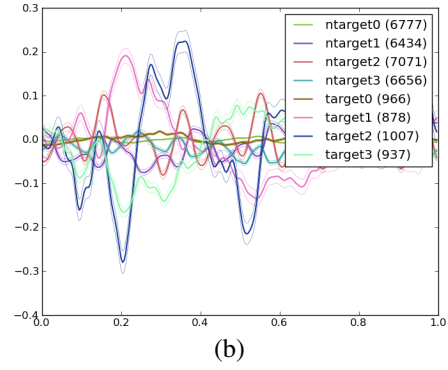
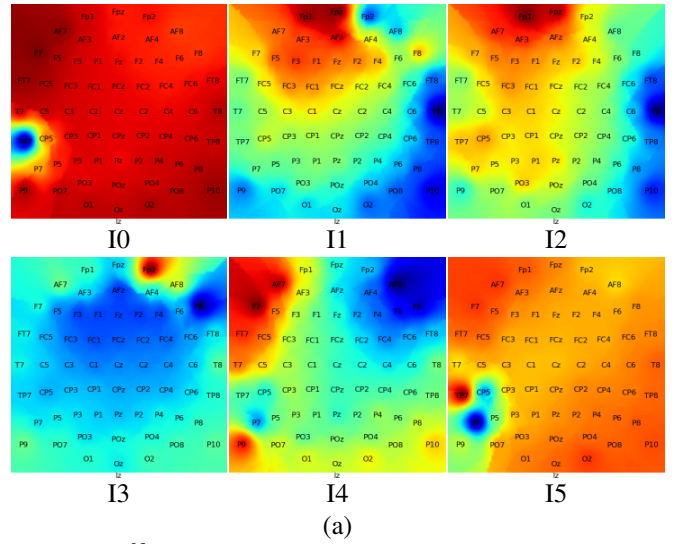


Fig. 4. ICA: (a) components corresponding to the six largest sources and (b) temporal variations of the projections of the EEG along the first 4 components averaged across targets and non-targets.

Fig. 5(b) are somehow in between those produced by the eigenbrains and those resulting from ICA.

While the visual inspection of components and their use to re-represent EEG are useful and suggest that eigenbrains are meaningful, we also want to see whether the use of eigenbrains can aid the automated single-trial classification of ERPs.

To study this aspect we used the EEG projections produced by the 6 most important eigenbrains, PCA components, and ICA components, respectively, and sampled them at 7 time steps: 100ms, 200ms, ..., 700ms (at each time step we averaged a window of 150ms of signal). This produced a vector of 42 features for each epoch in our data set. We then used these data with the corresponding labels (target or non-target) to train a Fisher's linear discriminant classifier. We used 5-fold cross-validation recording the classification error obtained with each fold. The mean and standard deviation of these errors for different EEG representations are reported in Tab. I. As one can see, eigenbrains compare well with both ICA and PCA.

To verify the statistical significance of these results we computed the p values for the one-sided, two-sample Kolmogorov-Smirnov test. As shown in Tab. II, Eigenbrains are statistically superior to ICA in this application and are

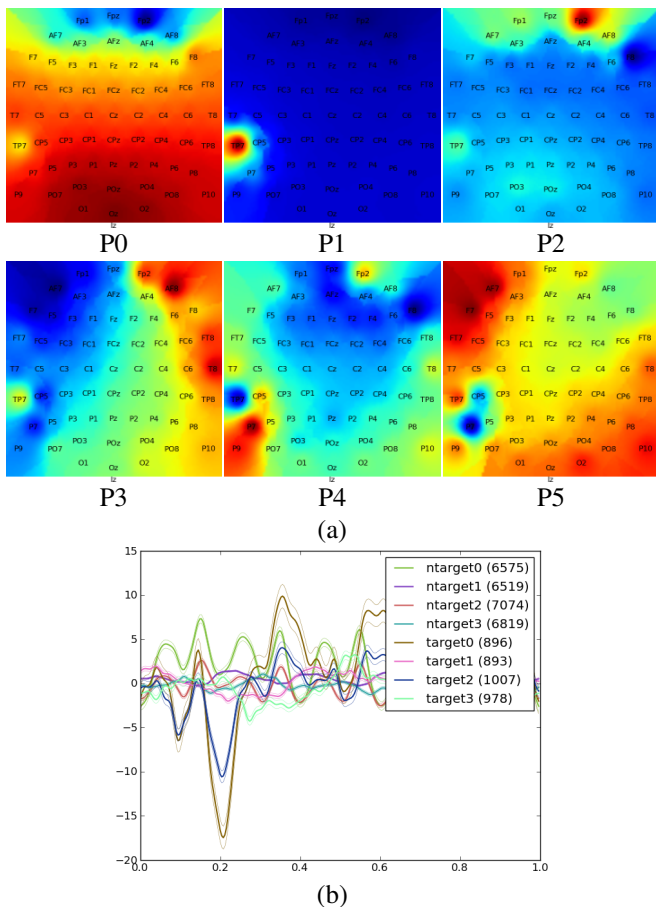


Fig. 5. Spatial PCA: (a) principal components corresponding to the six eigenvalues of the covariance matrix with the largest magnitude and (b) temporal variations of the projections of the EEG along the first 4 components averaged across targets and non-targets.

TABLE I

FIVE-FOLD CROSS-VALIDATION RESULTS FOR FISHER'S LINEAR DISCRIMINANT CLASSIFICATION WHEN USING 7 WINDOW AVERAGES OF THE SIX MOST PROMINENT COMPONENTS OBTAINED USING ICA, PCA AND TWO VERSIONS OF EIGENBRAINS. OUR DATA SET CONTAINS 8,246 EXAMPLES (1029 TARGETS AND 7217 NON-TARGETS).

Classifier	Average Prediction Error	Standard Deviation of Prediction Error
PCA	0.264	0.028
ICA	0.306	0.027
Eigenbrains (Mean Abs. Diff.)	0.236	0.017

almost statistically superior to PCA as well.

IV. CONCLUSION

We have proposed a new representation for EEG signals which is based on the an approximate model of the brain as a collection of coupled harmonic oscillators. The parameters of the model are stored in a matrix. Its entries represent the strength of the interactions between oscillators. These can be interpreted mechanically as spring stiffnesses or electrically as conductances. The eigenvectors of the matrix—the eigenbrains—are the free vibrational modes of the system.

As we have experimentally illustrated, the low-frequency eigenbrains are particularly suited to represent the large-

TABLE II

P-VALUES OF THE KOLMOGOROV-SMIRNOV ONE-TAILED SIGNIFICANCE TEST FOR THE RESULTS REPORTED IN TAB. I.

	worse	better	PCA	ICA	Eigenbrains (Mean Abs. Diff.)
PCA			1.000	1.000	0.165
ICA			0.165	1.000	0.007
Eigenbrains (Mean Abs. Diff.)			1.000	1.000	1.000

scale dynamics associated with ERPs. Specifically, they seem to capture, perhaps better than ICA and PCA, spatial and temporal activation patterns typically used by human experts to describe and characterise ERPs. We have also shown that the application of eigenbrains to the automated single-trial classification of ERPs has significant promise.

Naturally, since the sources of brain activity are not known in our experiments, we cannot claim that Eigenbrains are generally more accurate representations of such sources of brain activity than ICA and PCA. In the future, we hope to be able to establish this more firmly by either using data sets in which the putative brain sources are known in advance (e.g., those generated by the stimulation of the primary sensory cerebral areas) or using simulations in which the characteristics of the sources are known in advance. Also, in future research we want to extend the eigenbrains technique and to deploy it in both brain-computer interfaces and in psycho-physiological studies.

ACKNOWLEDGEMENTS

The UK Engineering and Physical Sciences Research Council (grant EP/F033818/1) is thanked for financial support.

REFERENCES

- [1] D. N. Streeter and J. Raviv, "Research on advanced computer methods for biological data processing," *AMRL-TR-66-24 Aerospace medical research laboratories (6570th)*, pp. 1–52, Mar 1966.
- [2] E. Donchin, "A multivariate approach to the analysis of average evoked potentials." *IEEE transactions on bio-medical engineering*, vol. 13, no. 3, pp. 131–139, Jul 1966.
- [3] E. Donchin and E. Heffley, "Multivariate analysis of event-related potential data: A tutorial review," in *Multidisciplinary perspectives in event-related brain potential research*, D. Otto, Ed. Washington, D.C.: U.S. Government Printing Office, 1978, no. EPA-600/9-77-043, pp. 555–572.
- [4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley-IEEE, 2001.
- [5] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," in *Advances in neural information processing systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds., vol. 8. The MIT Press, 1996, pp. 145–151.
- [6] S. Makeig, M. Westerfield, T. P. Jung, J. Covington, J. Townsend, T. J. Sejnowski, and E. Courchesne, "Functionally independent components of the late positive event-related potential during visual spatial attention." *The journal of neuroscience*, vol. 19, no. 7, pp. 2665–2680, Apr 1999.
- [7] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials." *Human brain mapping*, vol. 14, no. 3, pp. 166–185, Nov 2001.