

# An Averaging Technique for the P300 Spatial Distribution

A Tahirovic<sup>1</sup>, M Matteucci<sup>2</sup>, L Mainardi<sup>3</sup>,

<sup>1</sup>Faculty of Electrical Engineering, University of Sarajevo, Bosnia and Herzegovina;

<sup>2</sup>Department of Electronics and Informatics, Politecnico di Milano, Italy;

<sup>3</sup>Department of Biomedical Engineering, Politecnico di Milano, Italy;

## Abstract

*This paper proposes a novel averaging technique for the spatial distribution of P300 evoked potentials using independent component analysis. The averaging procedure is performed using the P300 signals obtained from different target epochs providing an averaged P300 spatial distribution for a subject of interest. An improved spatial selection of independent components or more accurate P300 extraction are some of the possible applications where the proposed approach might be used.*

**Keywords** ICA, BCI Competition, P300

## 1 Introduction

The P300 is an event-related potential (ERP) that is visible in an EEG recording as a positive peak at approximately 300 ms from the appearance of the event (stimulus). It follows unexpected, rare, or particularly informative stimuli. Both P300 amplitude and latency change across the scalp by decreasing over frontal and increasing over parietal areas [1], [2]. Topographic research has been presented in [3] [4] and an overview of P300 theory can be found in [5].

The P300 has been widely used for Brain Computer Interface (BCI), a communication interface which bypasses any muscle or nerve mediation and connects a computer directly with the brain by picking up signals generated by the brain activity. Application of P300 in BCI has many variants, but in all cases the paradigm is the same: the BCI system presents the user with some choices, one at a time; when it detects a P300 potential, the associated choice is selected. A well known BCI that exploits the P300 evoked potential is certainly the BCI P3 speller device [6], [7]. The BCI P300 speller device presents to a user a matrix of letters while its columns and rows flash randomly and one at a time. The user focuses the attention on a letter of interest to be spell out. After the letter is intensified, the brain generates a P300 potential natively. As the P300 is an innate response, it does not require training on part of the user.

Several signal processing techniques have been used to investigate and recognize the P300 and among these a quite natural tool seems to be Independent Component Analysis (ICA) which has been introduced in the BCI

paradigm with significant results [8]. The authors were winners of the BCI Competition 2003 using carefully found fixed ICA mixing matrix. Moreover, they have used temporal and spatial selection of generated independent components enhancing P300 signal and obtaining 100 accuracy in the testing data for the BCI P3 speller device. In [9], the authors have also used spatial selections of independent components, where they compared the spatial distribution of each independent component with the most-likely spatial distribution of the P300 signal obtained by training data from BCI Competition 2003 dataset. In [10] the fuzzy logic combined with ICA was employed to obtain high accuracy with one single intensification of raw and columns of a BCI-speller.

In this paper, we aim at analyzing the P300 amplitude spatial distribution over a scalp and a spatial distribution averaging technique based on the statistical properties of independent components is proposed. The main purpose in the recovery of the averaged topographic distribution of the P300 amplitude over a scalp is the "optimization" of P300-based BCIs through the selection of the optimal electrode set and the individuation of more informative area on the scalp for extracting P300 potentials.

The paper is organized as follows. The proposed technique for the P300 spatial averaging across a scalp is presented in Section II. Sections III shows the results obtained by the proposed technique using the data from the BCI Competition 2003 dataset. Conclusion is outlined in Section IV.

## 2 Spatial distribution from selected sources

In this section, we introduce a novel technique to estimate the information about P300 spatial distribution. Unlike the technique based on P300 data enhancement, this approach **does not require the building of an artificial signal**, requiring only few target epochs. Additionally, it is not affected by the occasional inconsistency of the solution obtained by the ICA algorithm.

Before we proceed, it should be stressed that it is an ideal assumption that the ICA algorithm can isolate a P300 signal as a single component. Instead, one should consider P300 possibly spread over more components. In order to select components that most likely carry the P300 pattern a temporal and spatial selection are performed.

Suppose a vector  $\mathbf{x}$  of  $m$  measurements is obtained from labeled data (i.e., one single target epoch), where  $m$

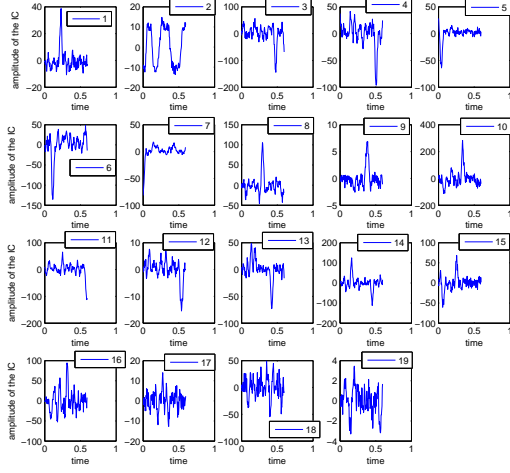


Figure 1: Independent components generated within one target epoch at the site of electrode Cz

is the number of electrodes used for the measurements, and the ICA algorithm is applied to  $\mathbf{x}$ . An example of obtained independent components are given in Fig. 1 for electrode Cz, which is selected according to the international 10/20 electrode allocation system.

Temporal selection of an independent component is based on the knowledge that a P300 component averagely appears 300ms after the excitation of the external stimulus and most likely has a peak around 310-350ms. From Fig. 1, one can select independent components  $s_8$ ,  $s_{10}$ ,  $s_{15}$  and  $s_{16}$  being most likely the P300 signal or a part of the P300 pattern.

Spatial selection additionally exploits the prior knowledge on the P300 signal. It is known that the P300 signal appears most intensively across the vertex region, eg. at Cz, Fz, Pz, C3 and C4 electrode allocations, and it is usually the most intensive at Cz.

Considering the spatial distribution of each component given in Fig. 2, one can select  $s_{10}$  as the most likely P300 independent component. When only one component is selected then its spatial distribution is equal to the one obtained by the values of the component correspondent column of the mixing matrix  $\mathbf{A}$ .

The component  $s_{16}$  is also included and combined with  $s_{10}$  as given in Fig. 3 since it fits both temporal and spatial criteria (see Figs. 1 and 2). When two P300-like independent components are selected, the question is how to combine the information on the spatial distribution of each selected vector to calculate the spatial distribution of their linear combination (eq. 1).

$$P = \mathbf{c}s_{P300} = \mathbf{a}_{10}s_{10} + \mathbf{a}_{16}s_{16}, \quad (1)$$

The vector  $\mathbf{c}$  represents the desired spatial distribution of P300 signal within the current epoch, and  $s_{P300}$  is the unique component which the ICA algorithm might extract after the first run in an ideal case.

In order to derive the final expression for the spatial distribution averaging of two components, we have to re-

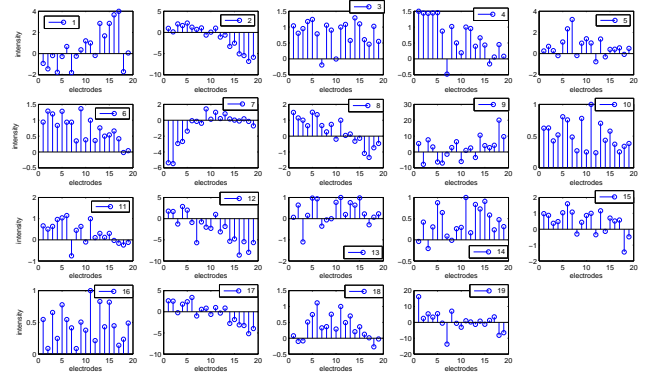


Figure 2: Spatial distribution of each IC generated within one target epoch

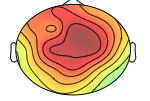
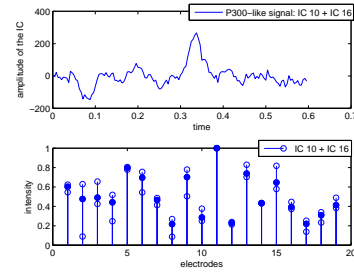


Fig 3: The case when two P300-like ICs are selected. Filled stems are the result of spatial distribution averaging, while empty stems represent the spatial distribution of selected ICs individually. The right hand-side figure gives corresponding topographic representation of the spatial distribution

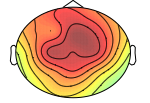
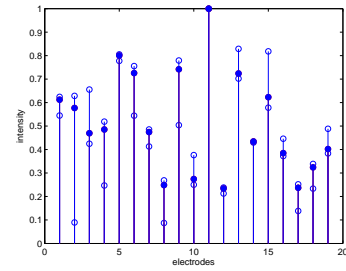


Fig 4: Spatial distribution obtained by two selected P300-like ICs using columns of mixing matrix  $\mathbf{A}$ . Filled stems are the result of spatial distribution averaging, while empty stems represent the spatial distribution of P300-like selected ICs individually. The right hand-side figure gives corresponding topographic representation of the spatial distribution

call the constraints imposed on the ICA algorithm result, the unit variance

$$E\{\mathbf{s}\mathbf{s}^T\} = \mathbf{I}, \quad (2)$$

and the mutual orthogonality

$$E[s_i s_k] = 0, i \neq j. \quad (3)$$

We can reformulated the goal of finding the spatial distribution of P300 as the computation of the coefficients of the new vector  $\mathbf{c}$ . After the variance operator is performed to the left and the right hand side of equation (1), recalling the constraint (3) we obtain

$$\mathbf{c}\mathbf{c}^T = \mathbf{a}_{10}\mathbf{a}_{10}^T + \mathbf{a}_{16}\mathbf{a}_{16}^T, \quad (4)$$

where  $\mathbf{c}\mathbf{c}^T$  is a diagonal matrix containing squares of the values hidden in the vector  $\mathbf{c}$ . For the sake of clarity, we consider only one electrode  $i$  among all measurement allocations, where  $i = 1, \dots, m$  and  $m$  is the number of electrode sites. By doing this, the matrix equation (4) becomes a scalar equation

$$c_i^2 = a_{i,10}^2 + a_{i,16}^2. \quad (5)$$

The problem occurs when at least one of the projection coefficients  $a_{i,10}, a_{i,16}$  is negative since we can not easily determine the sign of  $c_i$ . In order to cover this critical case, each column that projects the selected P300-like IC and has at least one negative value is shifted so to have all values greater or equal to zero; the linear combination extracted from (1) and the scalar equation (5) becomes now

$$(c_i - c_i^{min})s_{P300} = (a_{i,10} - a_{min,10})s_{10} + (a_{i,16} - a_{min,16})s_{16}, \quad (6)$$

$$(c_i - c_i^{min})^2 = (a_{i,10} - a_{min,10})^2 + (a_{i,16} - a_{min,16})^2, \quad (7)$$

where  $a_{min,10}, a_{min,16}$  are the minimal values of the  $\mathbf{a}_{10}$  and  $\mathbf{a}_{16}$  selected projection columns, if they are negative, respectively and  $c_i^{min}$  is an unknown shifting coefficient for each element of the required column  $\mathbf{c}$ . Please note that, if the value  $a_{min,i}$  is positive or equal to zero, the shifting of the  $i^{th}$  column is not necessary and we can set  $a_{min,i} = 0$ . A first relation among such coefficient  $c_i$  and its shifting coefficients  $c_i^{min}$  is derived from eq. (7).

$$c_i = \sqrt{(a_{i,10} - a_{min,10})^2 + (a_{i,16} - a_{min,16})^2} + c_i^{min}. \quad (8)$$

The second relation is obtained from eq. (1) for electrode  $i$  and it is given by

$$c_i s_{P300} = a_{i,10} s_{10} + a_{i,16} s_{16} = P_i. \quad (9)$$

Eqs. (6) and (9) imply

$$c_i^{min} s_{P300} = a_{min,10} s_{10} + a_{min,16} s_{16} = P_i^{min}. \quad (10)$$

If we perform the expected value operator  $\mathbf{E}$  on both sides of equations (9) and (10), we have

$$c_i \mathbf{E}[s_{P300}] = \mathbf{E}[P_i], \quad (11)$$

$$c_i^{min} \mathbf{E}[s_{P300}] = \mathbf{E}[P_i^{min}]. \quad (12)$$

If we divide left and righthand sides of these last two equations, the second relation among  $c_i$  and  $c_i^{min}$  is found and it is given by

$$c_i^{min} = c_i \frac{\mathbf{E}[P_i^{min}]}{\mathbf{E}[P_i]}. \quad (13)$$

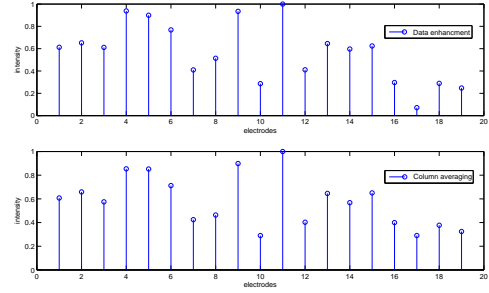


Fig 5: Comparison of the spatial distribution obtained by data enhancement (top) and averaging of the columns of mixing matrix  $\mathbf{A}$  (bottom)

It should be stressed that if  $\mathbf{E}[P_i] = 0$  then  $c_i = 0$  follows from (11), since the time domain shape of a P300 signal implies  $\mathbf{E}[s_{P300}] \neq 0$ . Recalling first relation (8), the final expression for  $c_i$  becomes

$$c_i = \sqrt{(a_{i,10} - a_{min,10})^2 + (a_{i,16} - a_{min,16})^2} \frac{\mathbf{E}[P_i]}{\mathbf{E}[P_i] - \mathbf{E}[P_i^{min}]} \quad (14)$$

A singular case arises when  $\mathbf{E}[P_i] = \mathbf{E}[P_i^{min}]$ . This case is possible if there exists such  $i$  that all  $i^{th}$  projection coefficients of columns  $\mathbf{a}_{10}$  and  $\mathbf{a}_{16}$  are  $a_{i,10} = a_{min,10}$  and  $a_{i,16} = a_{min,16}$  simultaneously. This is the case when the  $i^{th}$  projection coefficients of two vectors are negative and the correspondent coefficient  $c_i$  can be easily obtained by

$$c_i = -\sqrt{a_{i,10}^2 + a_{i,16}^2}. \quad (15)$$

The result can be generalized in accordance to the presented example. Suppose that within the data there is a P300 signal, and the ICA algorithm is performed separating  $m$  independent components  $s_i$ . If it is possible to select  $\kappa$  P300-like signals among those components ( $\kappa \in \{1, 2, \dots, m-1\}$  where  $m$  is usually equal to 2, 3 or 4) then the spatial distribution of the appropriate linear combination is hidden in the vector  $\mathbf{c}$  as follows.

**Proposition 1:** The spatial distribution of the evoked P300 signal within a target epoch could be obtained by

$$c_i = \sqrt{\sum_{j=1}^{\kappa} (a_{i,j} - a_j^{min})^2} \frac{\mathbf{E}[P_i]}{\mathbf{E}[P_i] - \mathbf{E}[P_i^{min}]}$$

if  $\mathbf{E}[P_i] \neq \mathbf{E}[P_i^{min}]$ , otherwise

$$c_i = -\sqrt{\sum_{j=1}^{\kappa} a_{i,j}^2},$$

where  $P_i = \mathbf{E}[\sum_{j=1}^{\kappa} a_{i,j} s_j]$  and  $P_i^{min} = \mathbf{E}[\sum_{j=1}^{\kappa} a_j^{min} s_j]$ .

Fig. 4 shows an example of the proposed spatial averaging for two selected P300-like components.

### 3 Experiment

Fig. 5 and Fig. 6 give the comparison between the two averaged spatial distributions obtained by data enhancement and the proposed method using 20 target

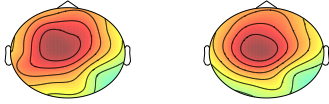


Fig 6: Topographic representation of the spatial distribution obtained by data enhancement (left) and averaging of the columns of mixing matrix **A** (right)

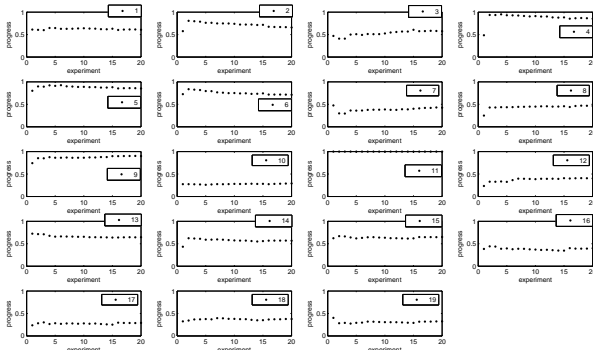


Fig 7: Progress of the spatial averaging based on the columns of the mixing matrix **A**

epochs. The results show that the distributions are quite similar, in which the correlation coefficients is  $\rho_{I,II} = 0.986571$ . The obtained similarity suggests that the proposed method can be used for the purpose of spatial selection of the independent components as in [9]. The experiments also show that the values computed by the proposed technique for P300 spatial distribution averaging converge after only several target epochs for each electrode allocation as shown in Fig. (7).

## 4 Conclusion

The work presents a novel technique for the understanding of the P300 spatial distribution. The P300 data enhancement combined with the ICA has already been used for the spatial selection of independent components. The proposed approach extracts a spatial distribution output and the obtained results are similar to the data enhancement technique. Unlike to the P300 data enhancement technique, the proposed approach gives a more explicit way to provide information on the averaged P300 spatial distribution of a subject of interest. Moreover, the P300 selection within a target epoch can be made manually throughout the experiment avoiding all unclear cases where a strict separation of the P300 signal is not possible to obtain. This gives additional reliability to the proposed method making it almost independent from the ICA algorithm despite the fact that ICA is used to separate source signals. An additionally acquired information on the spatial distribution of the P300 signal for a particular subject can be used to improve the technique aiming at the extraction of not only P300 signal but also other evoked potentials such as VEP.

## References

- [1] R. Mertens and J. Polich. P300 from a single-stimulus paradigm: passive versus active tasks and stimulus modality. *Electroencephalography and clinical neurophysiology evoked potentials*, 104(6):488, 1997.
- [2] J. Polich. On the relationship between eec and p300: Individual differences, aging, and ultradian rhythms. *International Journal of Psychophysiology*, 26(1-3):299, 1997.
- [3] R. Croft, C. Gonsalvez, C. Gabriel, and R. Barry. Target-to-target interval versus probability effects on p300 in one- and two-tone tasks. *Psychophysiology*, 40(3):322, 2003.
- [4] I. Tarkka and D. Stokic. Source localization of p300 from oddball, single stimulus, and omitted-stimulus paradigms. *Brain topography*, 11(2):141, 1998.
- [5] J. Polich. Updating p300: An integrative theory of p3a and p3b. *Clinical Neurophysiology*, 118(10):2128, 2007.
- [6] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical neurophysiology*, 70(6):510–523, 12 1988.
- [7] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a p300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2):174–179, 2000.
- [8] X. Gao, B. Hong, X. Miao, S. Gao, F. Yang, and N. Xu. Bci competition 2003 - data set iib: Enhancing p300 wave detection using ica-based subspace projections for bci applications. *IEEE transactions on biomedical engineering*, 51(6):1067, 2004.
- [9] S. Wang and C. J. James. Enhancing evoked responses for bci through advanced ica techniques. *Trials*, 10(1):4.
- [10] F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni, G. Palmas, and F. Beverina. Bci competition 2003 - data set iib: Enhancing p300 wave detection using ica-based subspace projections for bci applications. *Clinical Neurophysiology*, 117(3):531–537, 2005.

Address for correspondence:

Adnan Tahirovic  
University of Sarajevo  
Faculty of Electrical Engineering  
Department for Automatic Control  
atahrovic@etf.unsa.ba