

---

# Selective averaging of cognitive evoked potentials

---

**Arthur Flexer**

The Austrian Research Institute for Artificial Intelligence  
Schottengasse 3, A-1010 Vienna, Austria  
arthur@ai.univie.ac.at

**Herbert Bauer**

Institute of Psychology, University of Vienna  
Liebiggasse 5, A-1010 Vienna, Austria

## Abstract

This work is about the development of an alternative way of averaging evoked potentials (EP) of cognitive activities. Since the main assumption of invariant waveforms time locked to the eliciting events does not hold for cognitive EPs, averaging results in distorted estimates. Our alternative selective averaging finds similar subsequences of fixed length with variable latency which are common to all EPs by transforming the multivariate time series to discrete sequences via vector quantization and applying a sequence alignment algorithm. The method yields a significant improvement over common averaging in terms of noise attenuation and is shown to be valid by comparison with results for random data. Results for EP data obtained during a spatial imagination task are reported.

## 1 Introduction

An evoked potential (EP) is the electro cortical potential measurable in the electro encephalogram (EEG) before, during and after sensoric, motoric or cognitive events. This work is about the development of a new method for the analysis of EPs of cognitive activities for which already existing methods for the simpler sensoric or motoric EPs are not suited.

An EP is defined as the combination of the brain electric activity that occurs in association with the eliciting event and ‘noise’, which is brain activity not related to the event together with inference from non-neural sources. Since the noise con-

tained in EPs is significantly stronger than the signal, the common approach is to compute an average across several EPs recorded under equal conditions to improve the signal-to-noise ratio. For common averaging of a set of  $N$  EPs, the assumed model for the  $i$ th EP (given here for only one channel of EP) is the following [Glaser & Ruchkin 76, p.178]:

$$x_i(t) = s(t) + n_i(t); \quad i = 1, 2, \dots, N; \quad 0 \leq t < T \quad (1)$$

where  $x_i(t)$  is the  $i$ th recorded EP,  $s(t)$  the underlying signal,  $n_i(t)$  the noise associated with the  $i$ th EP, and  $T$  the duration over which each EP is recorded. The average  $\hat{s}(t)$  over the sample of  $N$  EPs is used to estimate the underlying signal  $s(t)$ :

$$\hat{s}(t) = \frac{1}{N} \sum_{i=1}^N x_i(t) = s(t) + \frac{1}{N} \sum_{i=1}^N n_i(t) \quad (2)$$

Averaging will attenuate the noise  $n_i(t)$  and not the signal  $s(t)$  given that signal and noise linearly sum together to produce the recorded EPs  $x_i(t)$  and the evoked signal  $s(t)$  is the same for each recorded EP  $x_i(t)$  and the noise contributions  $n_i(t)$  can be considered to constitute statistically independent samples of a random process. Averaging is standard for the analysis of exogenous components of motoric and sensoric events like the  $N100$  waveform which lie within  $100ms$  after the eliciting event. If the above assumptions do not hold, the averaging will result in a biased, distorted estimate of the signal.

Cognitive evoked potentials are endogenous components of the EPs and start about  $100ms$  after the onset of the recording and can last several seconds. They do not elicit one specific EP waveform time locked to the onset of the recording, and their analysis is a largely unsolved problem in psychophysiology. Classical methods like [Woody 67, McGillem & Aunon 77, Weerd & Kap 81] and [Westerkamp & Aunon 87] are designed for univariate time series of simpler motoric or sensoric EPs only and can therefore not really cope with the harder problem of analysing cognitive EPs. Other more modern approaches like e.g. Independent Component Analysis [Makeig et al. 97] use EP data of short duration ( $\leq 1sec$ ) after averaging. Another approach is to average despite improper assumptions thereby loosing all information about faster components and to analyse only positive and negative shifts of very slow potentials [Bauer 98].

Our hypothesis concerning cognitive EPs is that only subsequences of the whole EPs with differing onset times can be expected to be due to the cognitive task. Our approach to discover such subsequences which are common to all recorded EPs is to replace the sequence of the original multi dimensional vectors (our data was measured via 22 electrodes) by a sequence of codebook vectors obtained via vector quantization. The trajectories across codebook vectors are univariate discrete time series to which we can apply a multiple sequence alignment procedure [Bacon & Anderson 86] that has originally been designed for molecular biology.

The data was recorded with 22 electrodes with an equidistant matrix montage during a spatial imagination task (Three-dimensional Cube Test 3DC [Gittler 90]) and consists of 319 EP trials from 10 good female spatializers plus 167 EP trials from 8 poor female spatializers. After appropriate preprocessing (digital low pass filtering to frequencies below  $8Hz$  and eliminating the DC-like trend by subtracting a linear fit), each EP trial lasts  $8.5sec$  consisting of 2125 samples, each being a 22 dimensional real valued vector.

## 2 Computation of trajectories across codebook vectors

All EP time series are vector quantized together by using all the EP vectors at all the sample points as input vectors to a clustering algorithm disregarding their ordering in time. Then the sequence of original vectors  $x$  is replaced by the sequence of codebook vectors  $\hat{x}$ .  $K$ -means clustering (see e.g. [Duda & Hart 73, p.201]) is used for vector quantization using the sum of squared differences  $d(x, \hat{x}) = \sum_{i=0}^{k-1} |x_i - \hat{x}_i|^2$  as measure of distance, where both  $x$  and  $\hat{x}$  are of dimension  $k$ . Since observation of the sum of distances  $d(x, \hat{x})$  with growing size of codebooks did not indicate an optimal codebook size, we pragmatically decided to use 64 codebook vectors based on the following consideration: We vector quantized the EP data with increasing numbers of codebook vectors. We then took the original EPs and substituted at each sample point the original real valued vector  $x_l$  with the appropriate codebook vector  $\hat{x}_i$  (i.e. where  $d(x_l, \hat{x}_i) < d(x_l, \hat{x}_s)$  for all  $s$ ). We then visually inspected both the original EP time series and the coarser codebook time series as series of topographical patterns (spherical spline interpolations of the 22 values at a single point in time) and checked, whether the important features (positive and negative peaks and their development in time) of the topographies still were existent in the coarser codebook approximation. The high number of different discrete symbols (64 code book vectors) did not allow for a more principled information theoretic approach to obtain an optimal codebook size. For the 64 codebook vectors, we calculated a  $64 \times 64$  distance matrix  $D_C$ .

## 3 Sequence alignment and selective averaging

We chose a so-called *fixed length subsequence* approach for comparison of the sequences made of 64 discrete symbols (corresponding to 64 codebook vectors  $\hat{x}$ ). Given two sequences  $E$  and  $F$  of length  $m$ , all possible overlapping subsequences having a particular window length  $W$  from  $E$  are compared to all subsequences from  $F$ . For each pair of elements the score taken from the distance matrix  $D_C$  is recorded and summed up for the comparison of subsequences. The distance between two subsequences of length  $W$  from two sequences  $E$  and  $F$  is therefore:

$$D_{align}(b_e, b_f, W) = \sum_{i=0}^{W-1} d(E_{b_e+i}, F_{b_f+i}) \quad (3)$$

The indices  $b_e$  and  $b_f$  are the beginning points of the subsequences in the sequences  $E$  and  $F$  and  $E_{b_e+i}$  and  $F_{b_f+i}$  are the corresponding codebook vectors. Successive application of this pairwise methods allows for the alignment of more than two sequences. Such a *fixed subsequence* approach that is explicitly designed for *multiple sequence alignment* is given by [Bacon & Anderson 86]. It computes a multiple alignment by iteratively comparing sequences to the multiple alignment obtained so far, keeping always just the  $L$  best subsequences as an intermediate result. The succession of sequences is chosen at random. When a subsequence is compared to an intermediate “more”-way (let us say  $p$ -way) subsequence, the resulting score is computed as the sum of the  $p$  pairwise comparisons of the subsequences in the intermediate solution with the new subsequence that is to be aligned. The number of all such crosswise comparisons within the final overall alignment is given by  $P = \sum_{i=1}^{p-1} i$ . The number of all element-wise comparisons within the final overall alignment is given by  $WP$ , and its average per element, the average element-wise within alignment distance, by:

$$\bar{D}_{align} = \frac{1}{WP} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p-1} D_{align}(b_i, b_j, W) \quad (4)$$

Desired is a set of beginning points  $b_i^{min}$  for which  $\bar{D}_{align}$  is minimal. The  $b_i^{min}$  are the same for all  $d = 22$  channels of the corresponding  $i$ th EP. For each channel of EP we can compute an alternative selective average  $s'(t)$  where the duration  $T$  is equal to the length of the subsequences,  $W$ , and the beginning points of the averaging are the parameters  $b_i^{min}$ .

$$s'(t) = \frac{1}{N} \sum_{i=1}^N x_i(b_i^{min} + t); \quad 0 \leq t < W \quad (5)$$

This approach guarantees that the obtained multiple alignments contain subsequences that are part of all the original sequences. The number of single element-wise comparisons is  $LW(m + 1 - W)P$ . For a given  $L$  and  $m$ , this function is proportional to  $p^2$ , in contrast with  $m^p$  comparisons in “brute force” searching where not just the  $L$  best but all possible alignments are considered. As experiments with  $L$  equal 100, 1000 and 10000 showed, it is sufficient to keep 100 intermediate results to avoid the omission of good alignments that are weak in the first few sequences but strong in the later ones. Experiments varying the window length  $W$  from 31 to 62, 125 and 187 showed that  $W = 125$  (corresponding to 500ms of EP) is short enough to yield alignments of satisfactory quality which are still long enough to be significant in terms of their psychophysiological interpretation.

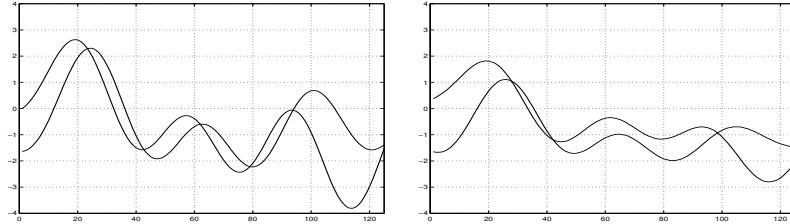
## 4 Results

To verify that our procedure yields different results for real human EPs and for unstructured random input, we compared results obtained from 21 EPs of one test subject with time-shuffled EPs and artificial EPs. The latter consisted of random Gaussian sequences whose power spectrum was changed appropriately to resemble the characteristics of real EPs. For each of the three types of EP, five runs of the fixed subsequence algorithm have been computed (to account for the variability due to its stochastic nature) and the means of the average element-wise within alignment distances (see Equ. 4) are  $\bar{D}_{align}(real) = 3.97 \pm .020$ ,  $\bar{D}_{align}(timeshuffled) = 4.38 \pm .010$  and  $\bar{D}_{align}(artificial) = 6.33 \pm .017$ . A one-way analysis of variance yielding a value  $F = 29070 > 18 = F_{99}(df = 2 \text{ and } 4)$  plus additional Duncan t-Tests allow us to rank the result for real EP as being significantly better than the result for time-shuffled EP, which is again significantly better than the result for random Gaussian EP, both with a probability of 99%.

To compare the gain in noise attenuation of the common average and of our selective average, the respective estimated standard deviations of the background noise,  $\hat{\sigma}(t)$  and  $\hat{\sigma}'(t)$ , are being compared.

$$\hat{\sigma}(t) = \left[ \frac{\sum_{i=1}^N [x_i(t) - \hat{s}(t)]^2}{N-1} \right]^{\frac{1}{2}} \quad (6) \quad \hat{\sigma}'(t) = \left[ \frac{\sum_{i=1}^N [x_i(b_i^{min} + t) - \hat{s}'(t)]^2}{N-1} \right]^{\frac{1}{2}} \quad (7)$$

Since the  $\hat{\sigma}(t)$  and  $\hat{\sigma}'(t)$  are given for each of the  $d = 22$  channels and for the duration of  $t = m$  or  $t = W$  respectively, the following average estimates of the standard deviations of the background noise are being computed:



(a)

(b)

Figure 1: Comparison of selective averages computed via beginning points  $b_i^{min5}$  obtained by minimizing the sum of  $\bar{D}_{align}$  from five repeated runs each. (a) is for electrode F1, (b) for P2, both for poor spatializers.

$$\hat{S} = \frac{1}{dm} \sum_{j=1}^d \sum_{t=0}^{m-1} \hat{\sigma}_j(t) \quad (8) \quad \hat{S}' = \frac{1}{dW} \sum_{j=1}^d \sum_{t=0}^{W-1} \hat{\sigma}'_j(t) \quad (9)$$

$\hat{S}$  is the estimate for the common averaging and  $\hat{S}'$  for the selective averaging. An  $\hat{\sigma}_j(t)$  is the  $\hat{\sigma}(t)$  for channel  $j$  given by Equ. 6. An  $\hat{\sigma}'_j(t)$  is the  $\hat{\sigma}'(t)$  for channel  $j$  given by Equ. 7. For all EPs of good and poor spatializers the common average  $\hat{s}$  as well as five selective averages  $\hat{s}'$  have been computed. Results for the good spatializers were  $\hat{S} = 7.68$  vs. mean  $\hat{s}' = 4.35 \pm .068$  and for the poor spatializers  $\hat{S} = 7.84$  vs. mean  $\hat{s}' = 4.37 \pm .048$ . Computing Z-values shows the differences in noise attenuation to be significant:  $Z_{good} = |(4.35 - 7.68)/(.068/\sqrt{5})| = |-109.5| > Z_{99} = 2.58$ ;  $Z_{bad} = |(4.37 - 7.84)/(.048/\sqrt{5})| = |-161.6| > Z_{99} = 2.58$ .

The repeated runs of the fixed subsequence algorithms showed a problem with the reproducability of the selective averages. Allthough a majority of the obtained beginning points  $b_i^{min}$  are identical, there is a difference in the selective averages. The mean of all possible pairwise average differences (average taken across all channels and points of time) between the five selective averages of the poor spatializer EPs is  $1.95 \pm .314$ . This is high considering that the mean absolute value of all selective averages is  $1.72 \pm 1.194$ . Sliding all five alignments across the sequences of codebook vectors and fixing beginning points  $b_i^{min5}$  to the points in time where the sums of the five  $\bar{D}_{align}$  (cf. Equ. 4) are minimal solved this problem. The average difference (average taken across all channels and points of time) between two selective averages computed via  $b_i^{min5}$  is  $0.81 \pm .687$ . The pairs of selective averages for poor spatializers are now very similar which is corroborated by visual inspection (Fig. 1).

The results of computing selective averages for both good and poor spatializers via beginning points  $b_i^{min5}$  are given in Fig. 2 as sequences of topographical patterns. Each topography is a spherical spline interpolation of the 22 values at a single point in time of the selective averaging window. Given are topographies at  $40, 80, \dots, 440, 480\text{ msec}$  of the window for poor spatializers (top two rows) and good spatializers (lower two rows). The poor spatializers show a peak of frontal activity which goes from positivity to negativity in 3 oscillating steps. The last topography shows wide spread negativity with emphasis in occipito-parietal regions. This is most prominent on the left side at F1 (see also Fig. 1(a)). Good spatializers show a similar but much weaker pattern of frontal activity. It is also rather symmetrical and the oscillation is not so clear. The final wide spread negativity is more

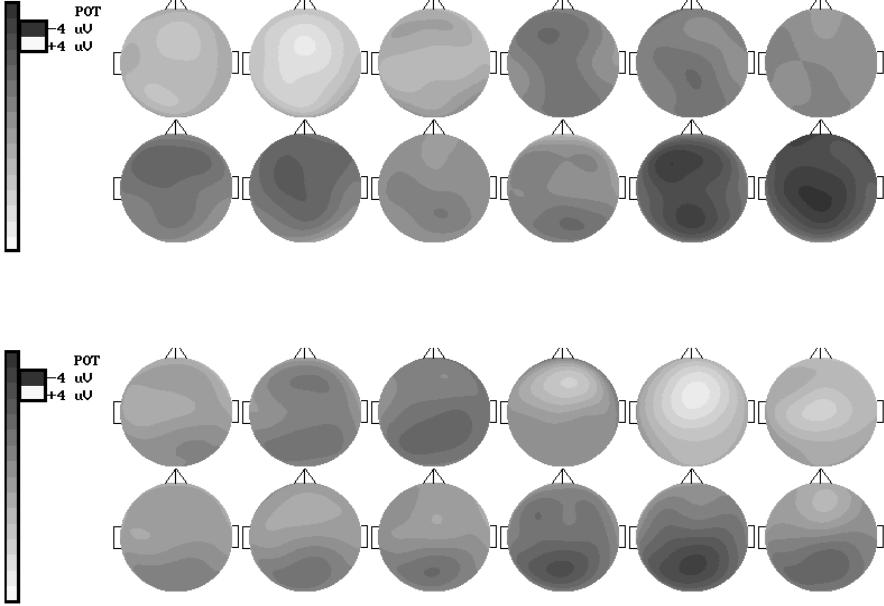


Figure 2: Sequences of topographies for poor spatializers (top two rows) and good spatializers (lower two rows). Scale is from  $-4$  to  $+4mV$ .

pronounced in right parietal regions.

## 5 Discussion

Our selective averaging finds common subsequences of a set of EPs which have fixed length but variable latencies and are sufficiently similar across all EP channels. The comparison with artificial random EPs shows that the obtained subsequences are not an artefact of our algorithm. The selective averaging also almost doubles the noise attenuation compared to common averaging.

Clinical findings as well as related EEG studies point to the significance of occipito-parieto-temporal regions for spatial processing, with an asymmetry to the right for males. In [Bauer et al. 98] the same data as in this work plus EPs obtained during an additional verbal task are analysed using common averaging to analyse positive and negative shifts of very slow potentials. In [Vitouch et al. 97] the same experimental design is used with male test subjects. Both studies confirmed the importance of occipito-parieto-temporal regions by showing significant DC-negativation during the cube rotation task. In our work, the focus is still on the slower parts of the EPs, but on those that are superimposed on the baseline shifts. Contrary to the above studies, we found mainly frontal activities at the left side which could be explained by the long standing hypotheses that especially poor spatializers invoke a strategy of verbalization during spatial tasks.

Both [Vitouch et al. 97] and [Bauer et al. 98] observed a significantly higher investment of cortical effort in both spatial and verbal tasks with poorer performers

investing more activity (i.e. showing more pronounced topographical patterns). The same holds for our results, the pattern peaked at frontal regions is more prominent in poor than in bad spatializers.

The final topographies of our selective averages show wide spread negativity with emphasis in occipito-parietal regions which is well in line with both clinical findings and with the results obtained in [Bauer et al. 98].

Having been able to show the soundness of our approach by comparison with results for artificial data as well as with common averaging and by discussing the consistency of our results for spatial EPs with those known from literature, future work will have to prove whether it is possible to obtain and discriminate task specific topographies with selective averaging of e.g. EPs from a spatial and a verbal task.

**Acknowledgements:** The EEG recordings have been made by R. Gstättner, Dept. of Psychology, University of Vienna. Parts of this work were done within the BIOMED-2 BMH4-CT97-2040 project SIESTA, funded by the EC DG XII. The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry of Science and Transport. The author was supported by a doctoral grant of the Austrian Academy of Sciences.

## References

- [Bacon & Anderson 86] Bacon D.J., Anderson W.F.: Multiple Sequence Alignment, *Journal of Molecular Biology*, 191, 153-161, 1986.
- [Bauer 98] Bauer H.: Slow potential topography, *Behavior Research Methods, Instruments and Computers*, 30(1), 20-33, 1998.
- [Bauer et al. 98] Bauer H., Vitouch O., Gstättner R., Lamm C., Leodolter M.: Investigation of cortical effort and sex differences with cognitive processing of a homogeneous spatial task, submitted, 1998.
- [Duda & Hart 73] Duda R.O., Hart P.E.: *Pattern Classification and Scene Analysis*, John Wiley & Sons, N.Y., 1973.
- [Gittler 90] Gittler G.: Dreidimensionaler Wuerfeltest (3DW). Ein Rasch-skalierter Test zur Messung des raeumlichen Vorstellungsvermoegens. *Theoretische Grundlagen und Manual*, Beltz Test, Weinheim, 1990.
- [Glaser & Ruchkin 76] Glaser E.M., Ruchkin D.S.: *Principles of Neurobiological Signal Processing*, Academic Press, 1976.
- [Makeig et al. 97] Makeig S., Jung T-P., Bell A.J., Ghahremani D., Sejnowski T.J.: Blind Separation of Auditory Event-related Brain Responses into Independent Components, *Proc. Natl. Acad. Sci. USA*, 94:10979-10984, 1997.
- [McGillem & Aunon 77] McGillem C.D., Aunon J.I.: Measurements of signal components in single visually evoked brain potentials, *IEEE Transactions on Biomedical Engineering*, 24, 232-241, 1977.
- [Vitouch et al. 97] Vitouch O., Bauer H., Gittler G., Leodolter M., Leodolter U.: Cortical activity of good and poor spatial test performers during spatial and verbal processing studied with Slow Potential Topography, *International Journal of Psychophysiology*, Volume 27, Issue 3, p.183-199, 1997.
- [Weerd & Kap 81] Weerd J.P.C.de, Kap J.I.: A Posteriori Time-Varying Filtering of Averaged Evoked Potentials, *Biological Cybernetics*, 41, 223-234, 1981.
- [Westerkamp & Aunon 87] Westerkamp J., Aunon J.I.: Optimum multielectrode a posteriori estimates of single-response evoked potentials, *IEEE Transactions on Biomedical Engineering*, 34, 13-22, 1987.
- [Woody 67] Woody C.D.: Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals, *Medical and Biological Engineering*, 5, 539-553, 1967.