31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Central-Tendency Estimation and Nearest-Estimate Classification of Multi-Channel Evoked Potentials

Srinivas Kota, *Member IEEE*, Phani Yarlagadda, Lalit Gupta, *Senior Member IEEE*, and D.L. Molfese

***Abstract* - By modeling evoked potentials (EPs) as random vectors in which the EP samples are random variables, a generalized strategy is introduced to determine multivariate central-tendency estimates such as the arithmetic mean, geometric mean, harmonic mean, median, tri-mean, and trimmed-mean. Additionally, a generalized strategy is introduced to develop minimum-distance classifiers based on central tendency estimates. Furthermore, procedures are developed to fuse the decisions of the nearest-estimate classifiers for multi-channel EP classification. The central-tendency estimates of real EPs are compared and it is shown that although the mathematical operations to compute the estimates are quite different, the EP estimates are similar with respect to their overall waveform shapes and latencies. It is also shown that by fusing the classifier decisions across multiple channels, the classification accuracy can be improved significantly when compared with the accuracies of individual channel classifiers.**

***Index Terms* - Central tendency estimation, EP averaging, EP classification, Evoked potentials.**

I. INTRODUCTION

The primary goal of this paper is to introduce alternatives to the arithmetic averaging for estimating evoked potentials (EPs) and to design and evaluate minimum-distance classifiers based on the alternative EP estimates. EPs are brain responses that are time-locked to the onset of an external event such as the presentation of an audio or a video stimulus. The accurate classification of EPs is of utmost importance because ERPs are used extensively in numerous human cognition studies and in clinical evaluations.

The most common estimate of an EP is obtained through time-locked ensemble averaging [1]-[4] in which the "average" is the sample arithmetic mean. The sample mean is the most often used measure of the central-tendency of a data distribution. There are other measures of central-tendency such as the geometric mean, harmonic mean, mode, median, tri-mean, and the trimmed-mean. These measures are used typically to measure the "center" of univariate data distributions and are seldom used for multivariate data such as EPs. In this paper, we introduce a generalized strategy for estimating central tendency measures for multivariate data and show how these estimates can be used to design minimum-distance classifiers for multivariate data. Six different central tendency estimates of real multichannel EPs are compared and classification experiments are designed to evaluate minimum-distance classifiers using the six EP estimates. Experiments are also designed to evaluate the combination of classifiers on single-channel and multichannel EPs.

S. Kota, P. Yarlagadda, and L. Gupta are with the Department of Electrical & Computer Engineering, Southern Illinois University, Carbondale, IL 62901.
D. Molfese is with the Birth Defects Center, University of Louisville, Louisville, KY 40292.

II. VECTOR MODEL FOR EP ESTIMATION

ERPs are multivariate signals with dimensions that are a function of the sampling rate and the duration. By modeling the EP as a *D*-dimensional random vector Z in which each element is a random variable $z(d), d = 1,2,...,D$, we introduce a unified method for estimating the central tendency of EPs from an ensemble of L single-trial EPs $Z_i, i = 1,2,...,L$ as shown in Figure 1. The random variations in each element are due to the background EEG noise and latency shifts [1]-[4]. Each row in the figure is a single-trial EP of the ensemble. The vector estimate and its elements are represented by $\overline{Z}$ and $\overline{z}(d), d = 1,2,3,...,D$ respectively. The figure shows that the vector central tendency EP estimate is obtained from the scalar central-tendency estimate of each vector element across the single-trial ensemble. For the arithmetic mean, the estimate $\overline{z}(d)$ is given by

$$\overline{z}(d) = \frac{1}{L}\sum_{i=1}^{L} z_i(d)$$

where $z_i(d), i = 1,2,...,L$ is the *d*th element of the *i*th single-trial EP. The geometric mean estimate is given by

$$\overline{z}(d) = \left[\prod_{i=1}^{L} z_i(d)\right]^{\frac{1}{L}}$$

The harmonic mean estimate is given by

$$\overline{z}(d) = \frac{1}{\sum_{i=}^{L} \frac{1}{z_i(d)}}$$

In order to determine order-statistic based central-tendency estimates, the elements $z_i(d), i = 1,2,...,L$ are ascending rank-ordered such that $z_i(d) < z_{i+1}(d)$ for $i = 1,2,..., L-1$. Let $\tilde{z}_i(d), i = 1,2,...,L$ be the rank-ordered samples, then, the median estimate is given by

$$\bar{z}(d) = \tilde{z}_{\left(\frac{D+1}{2}\right)}(d)$$

The tri-mean estimate is given by

$$\bar{z}(d) = (0.25)\tilde{z}_{\left(\frac{D+1}{4}\right)}(d) + (0.5)\tilde{z}_{\left(\frac{D+1}{2}\right)}(d) + (0.25)\tilde{z}_{\left(\frac{D+1}{4/3}\right)}(d)$$

The trimmed-mean estimate is given by

$$\bar{z}(d) = \frac{1}{L-\alpha}\sum_{i=-\alpha/2}^{\alpha/2} \tilde{z}_{\left(\frac{D+1}{2}\right)+i}(d)$$

where α is the number of samples deleted from $\tilde{z}_i(d), 1 = 1,2,...,L$. That is, the trimmed-mean estimate is given by first deleting the $\alpha/2$ lowest ranked samples and the $\alpha/2$ highest ranked samples and then computing the arithmetic mean of the remaining $(L-\alpha)$ samples.

III. NEAREST-ESTIMATE EP CLASSIFICATION

EPs are difficult to classify because they are embedded in the ongoing background EEG with signal-to-noise ratios (SNRs) typically less than –5 dB. In general, classification is conducted on ERPs averaged over a large number of single-trials because signal averaging improves the signal-to-noise ratios (SNRs) of ERPs. The improvement in classification accuracy through signal averaging has been shown systematically in [2],[3].

The nearest mean classifier is often used in practice because it is relatively simple to implement given that it requires only an estimate of the mean. Furthermore, it is optimal if the feature vectors are Gaussian, the features are independent and each feature has the same variance, and the prior probabilities of each class are equal. The nearest mean classifier also tends to give good results even if the conditions for optimality are not satisfied. For a *C*-class problem, the discriminant function of the nearest mean classifier may be modified to give the general expression for a nearest-estimate classifier as

$$g_c(Z) = Z^T \bar{Z}_c - (1/2)\bar{Z}_c^T \bar{Z}_c$$

where $\bar{Z}_c$ is a central-tendency vector of the patterns in class $c = 1,2,...,C$. A test vector Z is assigned to the category c^* given by

$$c^* = \arg\max_c \{g_c(Z)\}.$$

IV. FUSION CLASSIFIERS

Classifiers may be combined in different ways in attempts to improve the performance [5]. For EP classification, the outputs of the different nearest-estimate classifiers operating on the EPs of a given channel can be combined. Alternatively, the outputs of one type of nearest-estimate classifier operating on multiple channels can be combined. Yet another possibility is to combine the outputs of different nearest-estimate classifiers operating on multiple channels. We consider all possibilities through the development of a generalized decision fusion strategy. The decisions of the nearest-estimate classifiers are combined into a decision fusion vector and the final decision of the EP brain activity class is then made by classifying the decision fusion vector. If b_j, $j = 1,2,...,B$ is the decision of the j^{th} classifier and $Y_B = (b_1, b_2, ..., b_B)^T$ is the decision fusion vector formed by concatenating the independent decisions, the Bayes discriminant function for class *c* can be written as [3]

$$g_c(Y_B) = \sum_{j=1}^{B}\Big[\delta(b_j - 1)\ln(p_{j,1/c}) + \delta(b_j - 2)\ln(p_{j,2/c}) + \cdots + \delta(b_j - C)\ln(p_{j,C/c})\Big] + \ln P_c$$

where $\delta(x-a) = 1 \; if \; x = a$, $\delta(x-a) = 0 \; if \; x \neq a$, $p_{j,a/c} = P(b_j = a/c), \; j = 1,2,...,B$ is the probability that $b_j = a$ when the true class is *c*, and P_c is the *a priori* probability of class *c*. The final decision c_B^* resulting from fusing *B* decisions is given by

$$c_B^* = \arg\max_c [g_c(Y_B)]$$

V. EP Data

The data selected to demonstrate the estimation and nearest-estimate classification of EPs were borrowed from a previous study in which a subject made explicit match/mismatch comparisons between 2 sequentially presented stimuli [2]. The goal of the study was to show that EPs can reliably identify when a match occurs between what a subject thinks and sees. EP data were collected from 6 electrodes placed on the scalp over frontal (F3, F4), temporal (T3, T4), and parietal (P3, P4) regions over the left and right hemispheres. A total of 280 single-trial match and 280 single-trial mismatch responses were collected for each ensemble. Further details of the data collection and preprocessing operations can be found in [2].

VI. ESTIMATION RESULTS

Figure 2 shows the six estimates of the match and mismatch EPs of obtained from the single-trial EPs of the match and mismatch ensembles. Only the estimates of channel F3 are shown for brevity. For the trimmed mean, alpha was chosen to be 56 samples, that is, 10% of the samples were dropped from each end of the rank-ordered samples. Although the mathematical operations used to compute the estimates are quite different, it is interesting to observe the similarity between the EP estimates with respect to the shapes and latencies. Also quite interesting to note is that that the individual samples of median and the tri-mean estimates are selected, very likely, from different single-trial EPs. For example, $\bar{z}(n)$ may be $z_i(n)$ whereas $\bar{z}(m), m \neq n$, may be $z_j(m)$. That is $\bar{z}(n)$ is the *n*th sample from the *i*th single-trial EP and $\bar{z}(m)$ is the *m*th sample from the *j*th single-trial EP. This is also true for the trimmed-mean in the sense that the arithmetic mean of the individual samples is not necessarily computed using the samples of the same $(L-\alpha)$ single-trial EPs. It is also likely that samples from different single-trials are selected for the tri-mean. Yet, the shape and latencies are maintained in these cases. Similar observations were made for the EP estimates of the five other channels.

VII. CLASSIFICATION EXPERIMENTS

The first set of experiments was designed to classify the match and mismatch EPs of the six channels using the six estimates. Consequently, a total of 36 nearest mean classifiers were implemented. The second set of experiments were designed to optimally fuse the decisions of the six classifiers for each channel. The third set of experiments optimally fused the decisions of across the six channels for each classifier type. Finally, the fourth set of experiments optimally fused the decisions of the six classifiers from the six channels.

Through random partitioning of the match and mismatch ensembles into training and test sets [2],[3], each classification experiment was repeated 200 times and the classification accuracy was estimated as the average across the 200 trials. The classification accuracies, expressed as percentages, are summarized in Table I and Table II for EPs averaged across r=2 and 4 single-trials, respectively. The results in the tables can be interpreted as follows: in Table I, the row labeled F3 shows the classification accuracies obtained for the EPs of channel F3 using the six different estimates. The last column labeled Classifier Fusion contains the results of fusing the nearest-estimate decisions for each channel. Therefore, the last element in the F3 row shows the classification accuracy when the six decisions were fused and classified using a discrete Bayes classifier. The remaining rows labeled F4 to T4 can be interpreted in a similar fashion. The last row labeled Channel Fusion shows the results of fusing the decisions of each classifier type across the six channels. For example, the last element in the column labeled Arithmetic Mean is the result of fusing the six decisions of the nearest arithmetic mean classifiers for the six channels. Finally, the last element in the lower right corner shows the results of fusing the decisions of all 36 nearest-estimate classifiers.

Based on the results in the tables, the following conclusions can be drawn:

1. For the EPs of each channel, the six nearest-estimate classifiers yield different classification accuracies, however, the differences are quite small. This is not unexpected because the six central tendency estimates, as shown in Figure 2, are not very different from each other.
2. Fusing the decisions across channels improves the performance when compared with the individual channel accuracies. However, fusing the decisions across the classifiers for each channel did not result in an improvement. This again is due to the fact that the estimates are quite similar. The performance improved by fusing all 36 classifier decisions. However, the improvement was slightly less than that obtained by fusing the decisions of the nearest mean classifiers across the six channels.
3. Pairwise t-tests were also performed on the classification accuracies obtained by fusing the decisions of the nearest mean classifiers across the six channels and the best individual channel classification accuracies. The following results were obtained for the differences in the classification accuracies: (79.01 vs. 72.19, p=2.05e-108 in Table I); (86.64 vs. 80.41, p=7.9945e-75 in Table II). These very small p-values suggest that the differences in the classification accuracies are statistically significant.

VIII. CONCLUSIONS

A unique study was conducted to compare different central-tendency estimates of EPs, to develop and compare the performances of minimum-distance classifiers based on different central-tendency estimates, and to combine the decisions of the nearest-estimate classifiers in order to improve the classification accuracies over the individual nearest-estimate classifiers. The estimation results quite interestingly showed that the although the mathematical operations to compute the central tendency estimates were quite different, the estimates of the EPs were similar in the sense that the shapes and latencies were similar across the six estimates. It was also shown that the improvement in performance by fusing classifier decisions across channels is statistically significant when compared with the performance of the best individual channel classifiers.

REFERENCES

[1] C.E. Davila & R. Srebro, "Subspace averaging of steady-state visual evoked potentials," IEEE Trans. Biomed. Eng., vol. 47, No. 6, 720-728, 2000.

[2] L. Gupta, J. Phegley, & D.L. Molfese, "Parametric classification of multichannel evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 8, pp. 905-911, Aug. 2002 (vol. 49, no. 9, 1070, September 2002).

[3] L. Gupta, B. Chung, M.D. Srinath, D.L. Molfese, & H. Kook "Multi-channel fusion models for the parametric classification of differential brain activity," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 11, pp. 1869-1881, 2005.

[4] S.J. Luck An Introduction to the Event-Related Potential Technique, MIT Press, 2005.

[5] D.M.J. Tax, M. V. Breukelen, R.R.W. Duin, & J. Kittler, "Combining multiple classifiers by averaging or by multiplying?" *Pattern Recognition* 33, 1475-1485, 2000.

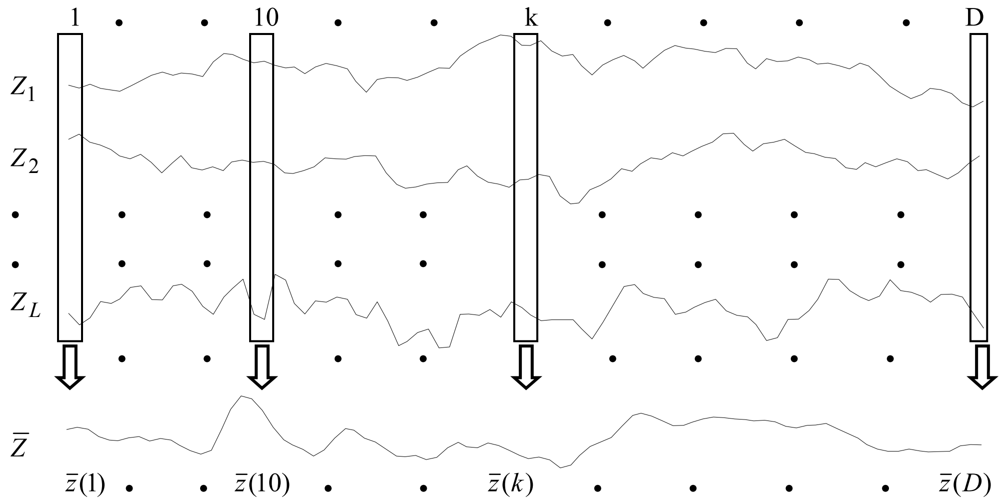

Fig. 1 Central-tendency estimation of EPs from an ensemble of single-trial EPs

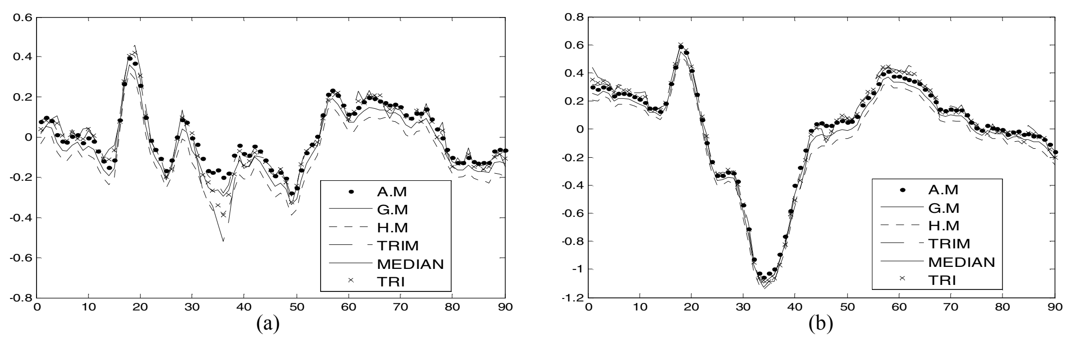

Fig. 2 Central-tendency estimates (a) Match EPs (b) Mismatch EPs

TABLE I Classification accuracies for r=2

r = 2	Arithmetic Mean	Geometric Mean	Harmonic Mean	Trimmed Mean	Median	Tri Mean	Classifier Fusion
F3	**72.19**	71.96	71.73	71.82	70.57	71.35	71.69
F4	**69.39**	69.21	69.06	69.15	68.67	68.95	69.16
P3	69.67	69.39	69.13	**69.77**	69.59	69.7	69.68
P4	65.00	64.92	64.86	65.20	65.05	**65.23**	65.07
T3	**69.60**	69.3	69.13	69.44	69.16	69.31	69.39
T4	67.47	67.59	**67.71**	67.61	67.15	67.54	67.62
Channel Fusion	**79.01**	78.66	78.39	78.63	77.65	78.35	79.00

TABLE II Classification accuracies for r=4

r = 4	Arithmetic Mean	Geometric Mean	Harmonic Mean	Trimmed Mean	Median	Tri Mean	Classifier Fusion
F3	**80.41**	80.24	80.02	80.14	78.68	79.67	80.00
F4	**76.38**	76.22	76.03	76.12	75.73	75.98	76.11
P3	76.30	76.16	75.89	**76.39**	76.05	76.30	76.26
P4	69.92	69.83	69.68	70.14	69.98	**70.16**	70.00
T3	**76.38**	76.17	75.96	76.16	75.74	75.99	76.14
T4	74.24	74.42	**74.51**	74.33	73.92	74.33	74.33
Channel Fusion	**86.64**	86.44	86.19	86.28	85.39	85.93	86.57