A Comparison of Temporal Windowing Schemes for Single-trial ERP Detection

Tian Lan[1], Catherine, Huang[1], Deniz Erdogmus[2,1]
[1] Department of Science and Engineering, Oregon Health & Science University, Beaverton, Oregon, USA
[2] Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts, USA

***Abstract* – Single trial ERP detection is critical for stimulus-synchronous brain computer interfaces. This paper presents a comparison of three different algorithmic schemes for single-trial ERP detection: SVM (baseline), hierarchical SVM-(naïve) Bayes, selected temporal windows-based SVM-(naïve) Bayes. An ERP-based image search system, including experimental setup, data collection, pre-processing, and three ERP detection schemes is described and utilized as the framework for comparison. We apply three schemes on EEG data from four subjects acquired on four days (eight sessions) each. Results indicate that a properly trained SVM operating on data from the post-stimulus [0,500]ms interval and SVMs trained on 50ms nonoverlapping windows spanning the poststimulus [200,450]ms interval (where P300 is expected) whose binary decisions are fused via the naïve-Bayes approach perform similarly in terms of area under the ROC curve measure, while the latter fusion approach applied to all ten nonoverlapping windows spanning the [0,500]ms poststimulus interval is inferior. The poststimulus time limit of 500ms is imposed on all data that goes into the ERP detector because in our experimental setup the subjects are asked to press a button when they recognize the event of interest, which creates motor responses in the brain typically in the [600,800]ms interval and around.**

***Keywords* – Electroencephalography (EEG), Brain Computer Interface (BCI), Event Related Potential (ERP), Rapid Serial Visual Presentation (RSVP), Support Vector Machine (SVM), Classifier Fusion**

I. INTRODUCTION

EEG has been widely used as a non-invasive approach for clinic diagnostics and brain computer interfaces. Thorpe's work showed that event related potentials (ERP) can be used for target detection on a rapid serial visual presentation (RSVP) task [1]: a positive perturbation in EEG appears after around 300ms of target stimulus, also referred to as P300. The basic idea of ERP based target recognition is to monitor and collect brain waves and analyze brain activities using machine learning techniques.

ERP detection is challenging due to the limited signal-to-noise ratio (SNR) of the non-invasive measurement. Eye blink, facial muscle movement and environment noise can contaminate ERP signals. Conventionally, ERP is studied by averaging stimulus-locked responses from multiple trials. However, the average process not only eliminates useful information about brain dynamics, but also compromises bandwidth of communication in a BCI setup. Recently, single trial ERP detection over multi-channel EEG collection received increasing interests due to its numerous potential applications, such as object recognition in brain computer interfaces, and information search from a large image database [2-7]. Huang et al. investigated several machine learning techniques for single trail ERP detection, including linear and nonlinear detectors [4], a boosting algorithm [5], and an SVM detector [7]. Among all ERP detectors designed and reported, SVM (with Gaussian kernels) yields better performance, with particular classification accuracy from 75% to 95% depending on subject and session when 32-channel EEG is utilized.

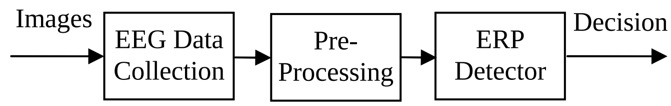

Figure. 1 ERP-based image recognition system.

Parra and his colleagues investigated a series of linear algorithms for ERP detection [8]. They also proposed hierarchical discriminant component analysis based on 50ms sliding non-overlapping windows [9-10] and reported 92% accuracy across five subjects (to our knowledge, the Parra-Sajda group prefers using 64 electrodes, which makes a few percent addition to performance in our own experience).

In this work, we attempt to exploit the positive aspects of both our own and Parra et al's work by utilizing SVMs in a hierarchical window-classifier scheme followed by naïve Bayesian fusion.

II. METHOD

Figure 1 shows a typical ERP based stimulus detection system.

Data Collection: Four subjects were recruited for the study under an approved IRB protocol for RSVP and EEG acquisition. Each subject finished eight sessions in four days (one session in the morning, one session in the afternoon). Each session contained 200 trials, each of which lasted 5 seconds. A trial contained one second fixation followed by 40 images (512x512) displayed at 100ms/image. A trial could contain no target or one target, the chance of a target appearing in a trial was set to 75% (making the total target prior <2%). The subjects indicated target detection by clicking on a button as soon as they saw a target (ERP signatures corresponding to motor activity occur typically in the 500-1000ms post-stimulus interval and depends on the speed of response). At the same time, we monitored their brain activity via EEG and recorded all data for subsequent analysis.

We used two computers to acquire data, one for image display and one for data collection. The EEG data were collected using a 32-channel Biosemi ActiveTwo system. Presentation™ (Neurobehavioral Systems, Albany, CA)

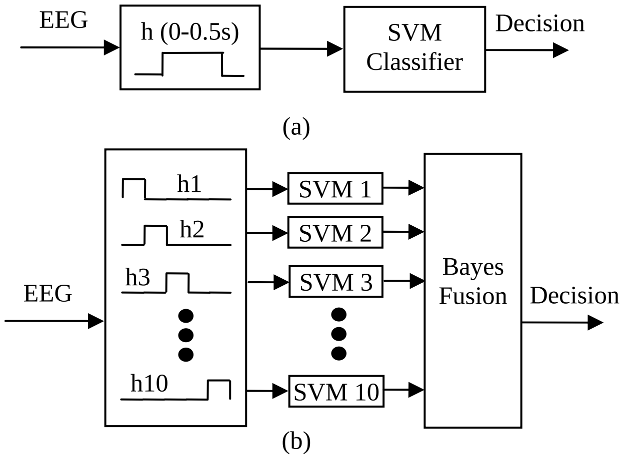

Figure. 2 Three single ERP detector schemes: (a) SVM classifier on 0-0.5 second window EEG data after image triggers. (b) 10 SVM classifiers on 10 50-ms non-overlapping windows after image triggers, and then fused by Bayes approach. The third scheme uses only SVMs on windows 5-9 during naïve Bayes fusion.

software was used to present images with a high degree of temporal precision and to output pulses or triggers to mark the onset of target and distractor stimuli. The triggers were received by the Biosemi system over a parallel port and recorded concurrently with the EEG signals. The user's button presses of indicating the response to target presence were recorded by the Biosemi system as well.

Data Pre-processing: By investigating EEG data after locking time origins of each response to stimulus onset, one can observe the P300after averaging a relatively small number of target trial responses (with subsequent button responses which confirm subject's recognition). After 600ms, motor response (due to button clicks) are also evident. These signals are not present in responses to non-target (distractor) stimuli. In order to avoid ERP detection based on motor activity, we truncated each response to 500ms post-stimulus. Each such truncated response (called an epoch in the following) represents the spatiotemporal electrical activity across brain regions associated with novel visual stimulus recognition, including early response from the visual cortex, and perhaps some weak signals corresponding to premotor planning activity towards the end of our window. We filtered the EEG signals using a bandpass filter in the range 1-45 Hz and normalized the data from each epoch to the unit-interval using the statistics of the 100ms-prestimulus window. If there was no button click following a target, we assumed there was no ERP and removed the epoch. If there was more than one click in one trial, we only retained the earlier one. The 32-channel data with [0,500]ms window in each epoch were used as the raw input to the ERP detector.

SVM ERP Detector: Our goal in classification is to build an ERP detector to accurately detect the ERPs associated with target stimuli. We adopt SVM [11,12] as our baseline ERP detector. A radial basis (Gaussian) kernel SVM is used. The kernel size σ and the cost parameter C can be set using cross-validation or chosen by the designer. To avoid overfitting, in

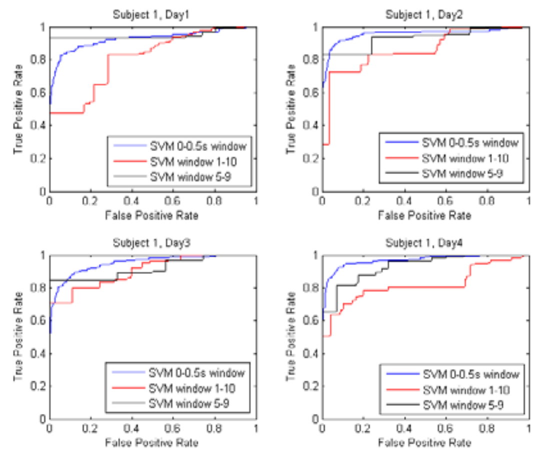

Figure 3. ROC curves for four days data using three different ERP detection schemes from subject 1.

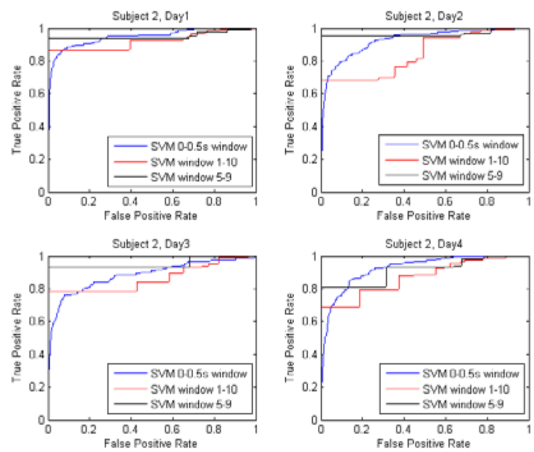

Figure 4. ROC curves from four days data using three different ERP detection schemes for subject 2.

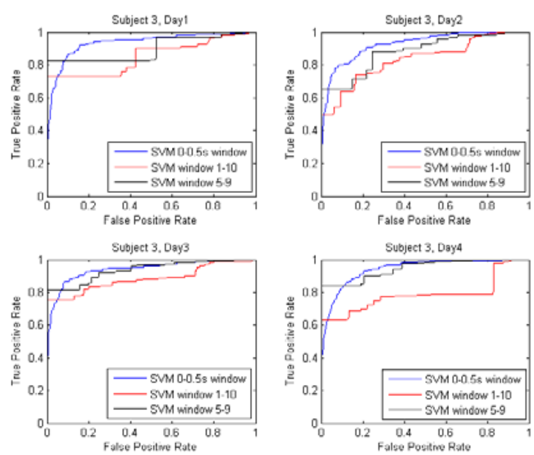

Figure 5. ROC curves from four days data using three different ERP detection schemes for subject 3.

previous studies we adopted 10-fold cross-validation [11] to adjust these model and regularization parameters. However, this process takes a long time when a brute-force grid-search approach is taken to find the global optimal in a preset domain.

Instead, based on our experience, for our particular data, we have proceeded with the selections σ=100 and C=10.

Hierarchical SVM-Bayes ERP Detector: As an alternative to classifying the vectorized raw data using the Gaussian-SVM, ten 50ms non-overlapping windows were used to partition the feature vector (in time, the interval 0,500[ms]) that are then individually classifier by Gaussian-SVMs of their own. The decisions of these first-layer SVMs are fused using the naïve-Bayes classifier approach, assuming that each binary decision is conditionally independent (given true label). Figure 2 (b) and (c) illustrate the structure of this hierarchical SVM-Bayes ERP detector. In (b), all windows were used and in (c), only component 5 – 9 were used due to the prior knowledge and observation of ERP.

For m-windows, let the first layer SVM decisions be denoted by d_j (j=1,…,m) and let c be the true ERP label:

$$p_{j,kl} = p(d_j = k \mid c = l), \quad k, l = 0,1 \tag{1}$$

Employing Bayes' rule and invoking the conditional independence assumption for the decision:

$$p(c \mid d_1 ... d_m) \propto p(d_1 ... d_m \mid c) p(c) = p(c) \prod_{j=1}^{m} p(d_j \mid c) \tag{2}$$

For equal risks, defining the discriminant threshold as th=p(c=0)/p(c=1), according to Bayes rule, the naïve Bayes decision-level fusion rule becomes:

$$\prod_{j=1}^{m} p(d_j \mid c = 1) \Big/ \prod_{j=1}^{m} p(d_j \mid c = 0) \overset{c=1}{>} th \tag{3}$$

Combining (1) and (3) and noting that the threshold can be modified for different risk-ratios for miss and false detections:

$$\prod_{j=1}^{m} p_{j01}{}^{(1-d_j)} p_{j11}{}^{d_j} \Big/ \prod_{j=1}^{m} p_{j00}{}^{(1-d_j)} p_{j10}{}^{d_j} \overset{c=1}{>} th \tag{4}$$

where we estimate p_{j00}, p_{j01}, p_{j10}, p_{j11} from training set (via validation), and obtain d_j from layer-one SVM classifiers for each test sample.

The fact that P300 happens around 300ms after stimulus onset implies that not all windows carry useful information. For example, window 1 ([0,50]ms) carries much stronger background activity than any potentially present stimulus-related response or even transients from stimulus switching boundaries, which is expected to compromise ERP detection accuracy. Eliminating irrelevant windows (features) is expected to improve performance and make the detector robust. Based on these observations, windows 5-9 ([200,450]ms) are expected to carry most discriminative energy. Two variations of the SVM-Bayes detector described above and compared in our study use (i) all 10 windows, and (ii) only windows 5-9 during fusion (see Figure 2).

III. Experiments And Results

We applied the three ERP detector schemes described above to EEG data collected from four subjects. Each subject finished eight sessions of experiments in four days, with two sessions per day. We used the data from the morning sessions as training set and the data from afternoon sessions on the same day as testing set. The results are represented using ROC curves, area under ROC curves, and MFAR (minimum false alarm rate at zero miss).

The results are comprehensively reported in Figures 3-6 for each of the four subjects in the form of ROC curves and in Tables 1-8 in the form of area under the ROC curve (AUC) and MFAR broken down by each subject and each day's experimental sessions. While we don't observe that one method is clearly superior to the other two in all the instances,

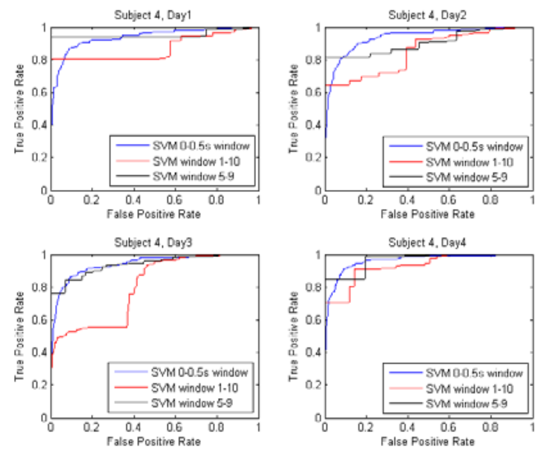

Figure 6. ROC curves from four days data using three different ERP detection schemes for subject 4.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.92	0.96	0.95	0.97
window 1-10	0.81	0.86	0.91	0.82
window 5-9	0.95	0.93	0.92	0.93

Table 1. Area under curves for four days data using three different ERP detection schemes from subject 1

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.95	0.96	0.81	0.77
window 1-10	0.91	0.88	0.68	0.98
window 5-9	0.95	0.99	0.87	0.81

Table 2. MFAR for four days data using three different ERP detection schemes from subject 1

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.95	0.93	0.88	0.92
window 1-10	0.92	0.84	0.86	0.87
window 5-9	0.95	0.96	0.95	0.91

Table 3. Area under curves for four days data using three different ERP detection schemes from subject 2.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.71	0.93	1.00	0.64
window 1-10	0.99	0.93	0.97	0.89
window 5-9	1.00	0.95	0.94	0.82

Table 4. MFAR for four days data using three different ERP detection schemes from subject 2.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.93	0.93	0.94	0.94
window 1-10	0.85	0.84	0.88	0.79
window 5-9	0.90	0.88	0.94	0.95

Table 5. Area under curves for four days data using three different ERP detection schemes from subject 3.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.96	0.66	0.85	0.87
window 1-10	0.97	0.85	0.99	0.93
window 5-9	0.99	0.89	0.91	0.75

Table 6. MFAR for four days data using three different ERP detection schemes from subject 3.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.94	0.94	0.94	0.96
window 1-10	0.87	0.85	0.80	0.93
window 5-9	0.95	0.91	0.94	0.97

Table 7. Area under curves for four days data using three different ERP detection schemes from subject 4.

	Day 1	Day 2	Day 3	Day 4
0-0.5s window	0.87	0.83	0.81	0.82
window 1-10	0.97	0.93	0.82	0.62
window 5-9	0.98	0.90	0.82	0.69

Table 8. MFAR for four days data using three different ERP detection schemes from subject 4.

the following qualitative observations emerge: (i) In general, using windows 5-9 in SVM classification, followed by naïve Bayes fusion of their decisions performs best more often than using an SVM on the whole [0,500]ms window, and these two are better almost all the time from fusing SVM decisions of all 10 windows; (ii) the naïve fusion of decisions on windows 5-9 approaches zero-false alarm rate at a higher detection rate than SVM on [0,500]ms, while the latter approaches 100% detection rate (0-miss) at a lower MFAR, based on the ROC curves on finite amount of trials used to obtain these results.

IV. Discussion and Conclusion

In this paper, we described three schemes for single ERP detection. Based on our previous work, SVM yielded better performance than other methods we utilized. Using the Gaussian-SVM on temporal window [0,500]ms as a baseline, we evaluated two decision level Bayesian fusion approaches that utilize short-window SVM decisions of the ERP waveform following stimulus onset (inspired by the success reported by Parra and Sajda in their various papers). Experimental results on four subjects across sessions illustrate that SVM on [0,500]ms (scheme 1) and SVMs on windows at [200-450]ms fused at decision level via naïve Bayes method (scheme 3) yield similar performances, while fusing the decisions of all ten windows in the [0,500]ms post-stimulus interval (scheme 2) is inferior; expectedly so since this approach corrupts accuracy by introducing decisions from temporal data that do not contain sufficiently powerful evidence about the presence or lack of ERP waveforms.

Although scheme 3 does not exhibit better performance than scheme 1 in all instances we analyzed, it has several advantages. First, it utilizes a smaller dimensionality feature vector, thus is expected to be more robust over long-term BCI training; this also contributes to computational efficiency for real-time implementation. Second, by breaking the raw data into multiple discriminant temporal components, EEG channel selection becomes feasible for each short window; this is important because at different phases of the response, different regions of the brain become active, therefore for each window, appropriate channels can be retained, contributing to further reduction of feature dimensionality, hence classifier robustness. Future work will include utilizing mutual information based EEG channel selection for each short temporal window in order to achieve this goal [13]. Third, scheme 3 allows us to further improve the Bayesian fusion model utilizing more elaborate graphical models of dependencies between the temporal windows. Even with the naïve Bayesian fusion, this approach is competitive with our baseline approach in the experiments performed.

Acknowledgement

The authors would like to acknowledge funding from the following sources that contributed to the completion of the reported work: NSF grants ECS-0524835, ECS-0622239, and IIS-0713690; and a grant from the Nancy Lurie Marks Family Foundation.

References

[1] S. Thorpe, D. Fize, C. Marlot, "Speed of processing in the human visual system," *Nature*, 381, 520-522, 1996.

[2] S. Lemm, B. Blankertz, G. Curio, and K.-R. M¨uller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, 52(9):1541–1548, 2005.

[3] A. Gerson, L. Parra and P. Sajda, "Single-trial analysis of EEG for enabling cognitive user interfaces," Ch. 40 in *Handbook of Neural Engineering*, M. Akay (Ed), Wiley/IEEE Press, 2006.

[4] Y. Huang, D. Erdogmus, S. Mathan, M. Pavel, "Comparison of linear and nonlinear approaches in single trial ERP detection in rapid serial visual presentation tasks," in *Proceedings of IEEE International Joint Conference on Neural Networks*, Vancouver, Canada, 2006.

[5] Y. Huang, D. Erdogmus, S. Mathan, M. Pavel, "Boosting linear logistic regression for single trial ERP detection in rapid serial visual presentation tasks," in *Proceedings of the 28th International Conference of IEEE EMBS*, New York, 2006.

[6] P Sajda, A. D. Gerson, M.G. Philiastides and L.C. Parra, "Single-trial analysis of EEG during rapid visual discrimination: Enabling cortically-coupled computer vision," in *Towards Brain-Computer Interfacing*, Eds. G. Dornhege, J. R. Millan, T. Hinterberger, D.J. McFarland and K.R. Muller. MIT Press, invited, in press, 2007.

[7] Y. Huang, D. Erdogmus, S. Mathan, M. Pavel, "Large-scale image database triage via EEG evoked responses," in *Proceedings of the 2008 IEEE International Conference of Acoustics, Speech, and Signal Processing*, Las Vegas, 2008.

[8] L.C. Parra, C.D. Spence, A.D. Gerson, P. Sajda, "Recipes for the linear analysis of EEG," *Neuroimage*, 28(2), 326-341, 2005.

[9] A.D.Gerson, L.C. Parra, P. Sajda, "Cortically coupled computer vision for rapid image search", in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. Vol. 14, No.2. pp. 174-179

[10]L.C.Parra, C.Christoforou, A.D.Gerson, M. Dyrholm, A. Luo, M. Wagner, M. G. Philiastides, P. Sajda, "Spatiotemporal Linear Decoding of Brain State", in *IEEE Signal Processing Magazine*, Jan. 2008, pp.107-115

[11] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol 2, pp.121-167, 1998.

[12] R.O. Duda, R.E. Hart, D.G. Stork, *Pattern Classification*, 2nd Edition, John Wiley & Sons, New York, 2001.

[13] T. Lan, D. Erdogmus, A. Adami, S. Mathan, M. Pavel, "Feature and Channel Selection for Cognitive State Estimation Using Ambulatory EEG," Computational Intelligence and Neuroscience, vol. 2007, Sep 2007.