ESANN 2010 proceedings, European Symposium on Artificial Neural Networks - Computational Intelligence and Machine Learning. Bruges (Belgium), 28-30 April 2010, d-side publi., ISBN 2-930307-10-2.

Identifying informative features for ERP speller systems based on RSVP paradigm

Tian Lan[1], Deniz Erdogmus[2], Lois Black[1], Jan Van Santen[1]*

1- Oregon Health & Science University - Dept of Science and Engineering
Beaverton, Oregon USA
2- Northeastern University - Dept of Electrical and Computer Engineering
Boston, Massachusetts, USA

Abstract. This preliminary study focused on identifying informative features in the frequency and spatial domains for single-trial Event Related Potential (ERP) detection for ERP spelling systems. A predefined sequence of letters was presented to subjects in a Rapid Serial Visual Presentation (RSVP) paradigm. EEG data were collected and analyzed offline. A Linear Discriminant Analysis (LDA) classifier was selected as ERP detector for its simplicity and robustness. A range of features in different frequency bands and EEG channel subsets was extracted and detection accuracies were compared for different classes of features.

1 Introduction

Children with Autism Spectrum Disorders (ASD) are heterogeneous in their verbal and nonverbal communication abilities. Augmentative and Alternative Communication (AAC) devices may help these children communicate and thereby reduce their social isolation. However, a subgroup of nonverbal children with ASD also lack the basic motor skills needed to operate keyboard based AAC devices. The communication options of these children are extremely limited. Among different AAC devices, an EEG based Brain Computer Interface (BCI) system that directly transforms brain activity into words may offer communication options for these children. Ever since Farwell and Donchin [1] showed that the visual P300 ERP can be used to select letters from a computer display, P300 spellers in different paradigms have been studied. Farwell and Donchin's P300 speller presented letters and digits in a 6×6 matrix, then highlighted rows and columns in a random sequence. The user selected the row and the column that contained target letter consecutively, thereby selecting the intended letter. Muller and Blankertz presented their P300 speller using 6 hexagons instead of matrix [2]. The user was again required to make two successive selections to find the desire letter.

Thorpe's work showed that the ERP can be used for target detection in a Rapid Serial Visual Presentation (RSVP) task [3]. An RSVP based paradigm allows fast presentation of characters, and can be integrated with statistical language models, significantly reducing sequence lengths and increasing typing speed. Our previous

* The research reported in this paper was supported in part by a grant from the Nancy Lurie Marks Family Foundation, "ERP Based Communication Device for Nonverbal Children on the Autism Spectrum", Deniz Erdogmus and Lois Black Principal Investigators.

study on a single-trial ERP detector showed that a Support Vector Machine (SVM) based classifier yielded better performance than other classifiers [4-5]. However, Krusienski compared different classifiers and found that the SVM was not necessarily the best classifier in traditional P300 speller systems [6]. Krusienski also found that given a properly selected EEG channel subset, a Linear Discriminant Analysis (LDA) based classifier outperformed SVM in target detection accuracy and robustness. Most important, compared to SVM, LDA classifiers require far less computation in training, making them more suitable for adaptive real-time systems. In Krusienski's study, pre-determined channel subsets were compared and used throughout the study. Though Krusienski claimed the selected subset exhibited session-to-session transfer ability, no explicit results demonstrated that this selection also transferred between subjects. An online feature selection approach is more desirable for real world applications. In this paper, we focused our study on feature selection in both frequency bands and EEG channels. We presented English letters in an RSVP paradigm, collected EEG data, and analyzed these offline using a LDA classifier.

2 Experiment and data collection

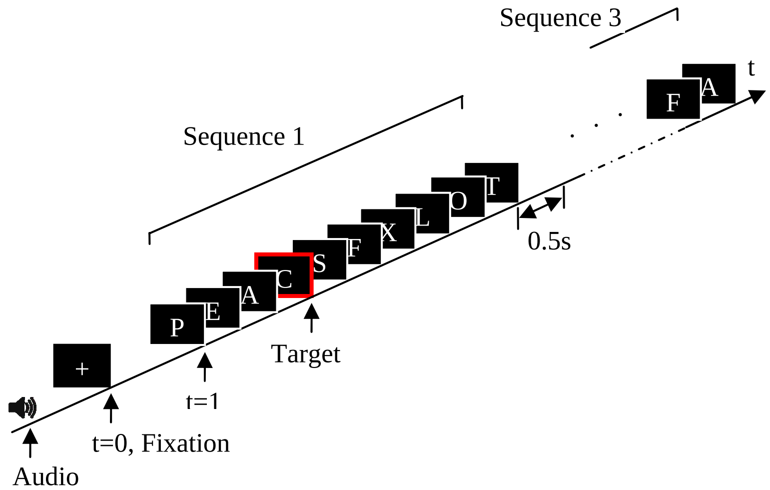


Fig. 1: An example of a RSVP trial

Three adult subjects were recruited for the study under an approved IRB protocol for RSVP and EEG acquisition. Each subject finished 2 sessions in two days. Each session contained 100, 10-second trials. A trial started with an audio presentation of the target letter. At time stamp 0, a one-second fixation screen was presented, followed by 3 sequences of English letters. Each sequence contained 10 images (one letter per image) at 167ms/image in a random order. Within each sequence, there was only one target letter, all others were distracters. There were 0.5 second intervals between two sequences. EEG recordings were made synchronously. An example of an RSVP trial is shown in Figure 1.

We used two computers to acquire data, one for image display and the other for data collection. The EEG data were collected using a 32-channel Biosemi ActiveTwo system at sampling rate 256Hz. Presentation™ (Neurobehavioral Systems, Albany, CA) software was used to present images with a high degree of temporal precision and to output pulses or triggers to mark the onset of the target and distracter stimuli. The triggers were received by the Biosemi system over a parallel port and recorded concurrently with the EEG signals.

3 Data analysis and results

3.1 Frequency bands analysis

We first filtered EEG signals using a bandpass filter. Previously[4-5], we used a 1-45Hz passband filter. In the current study, we compared ERP energy in narrow bands ranging from 0 to 44Hz, and determined that ERP energy was mainly concentrated in the 0-1Hz and 10-20Hz bands. Since the P300 has limited energy in the 10-20Hz band, this indicates that it is not the sole contributor to ERP based detection. Based on the same analyses, we determined that the 0-20Hz (no DC) passband yielded best performance; consequently, this filter was used throughout.

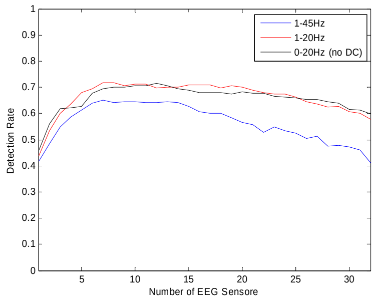
The thus filtered data were truncated using a [0, 500ms] window following each image stimulus (called an "epoch" in what follows) and normalized with the [-100ms, 0] pre-stimulus window. We concatenated data within one epoch by channels, and to obtain a data point with 32×129=4128 dimensions (each channel contained 129 samples).

3.2 EEG channel ranking and selection

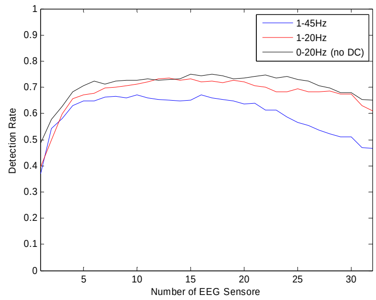
EEG channels were ranked using the wrapper approach (error based approach) with a greedy search strategy. For a given channel subset, all samples from these channels were concatenated to form a new data point for each epoch. We used epochs from the first 50 trials as training set, used the remaining data as testing set, applying the LDA classifier on three sequences separately, and fusing results using a majority vote for the final decision. The accuracy of this final decision was used as criterion for ranking the EEG channels. The channel ranking results are shown in Figure 2 (a-f). Each figure corresponds to a subject/session. Three curves of different colors denote features in three different frequency bands. As mentioned, a frequency range 0-20Hz, no DC (black curve) yielded best performance. The horizontal axis denotes the number of optimal channels used for ERP detection; the vertical axis denotes detection accuracy.

Theoretically, detection accuracy should increase as more EEG channels are used since more information is available. However, due to noise and the finite amount of training data, using more channels beyond a critical dimensionality causes over-fitting and hence reduces performance. This explains why in Figure 2 optimal performance is obtained when only a subset of the channels is selected.

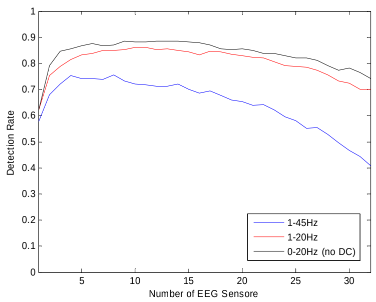
We note that determining an optimal set in the original high dimensional feature space is time-consuming and can only be done offline. If session-to-session transfer and subject-to-subject transfer were unproblematic, one could find the optimal subset of channels offline and then use them in the online system. However, transfer appears to be an issue. For example, the optimal channel subset for subject 1, session 1 was: F7, O1, Fp2, P7, Pz, PO3, AF4, CP2, P3, T8, FC1, FC2, while the optimal channel subset for subject 2, session 1 was: PO3, T7, Oz, CP2, Cz, F8, AF3, O2, P8, C4, P7, F4. This points out the need for fast feature selection algorithms in real-time BCIs.



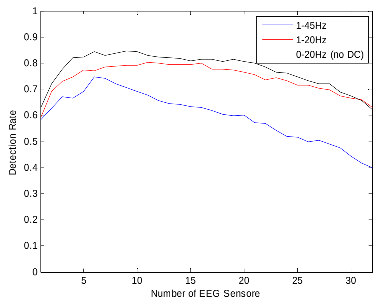
(a) Subject 1, session 1



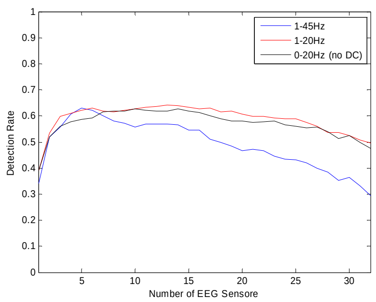
(b) Subject 1, session 2



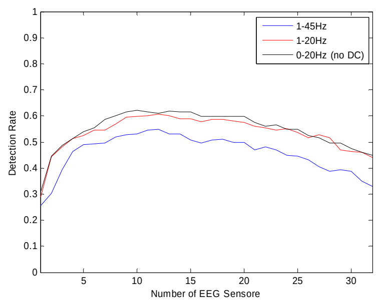
(c) Subject 2, session 1



(d) Subject 2, session 2



(e) Subject 3, session 1



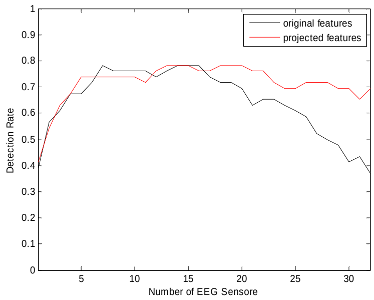
(f) Subject 3, session 2

Fig. 2: Detection Accuracies for all subjects/sessions using features with different frequency passbands and different EEG channel subsets.

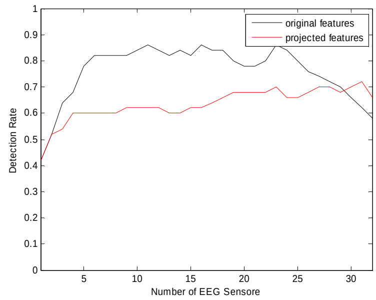
3.3 Channel-wise dimensionality reduction

Online channel ranking and subset channel selection were not feasible in high dimension due to the computational cost, making dimension reduction critical.

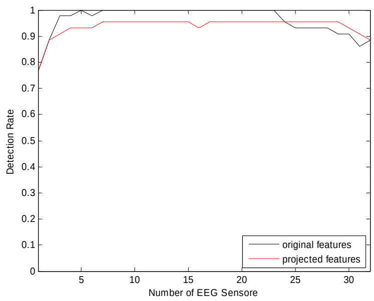
Subspace feature projection approaches are widely used for dimensionality reduction, such as PCA, ICA, and LDA. We compared different subspace projection methods and found that LDA yielded the best performance, as expected. Thus, prior to channel ranking, we applied dimensionality reduction on each channel separately using LDA, which reduced the dimension from 4128 to 32, with one feature per channel. This resulted in it taking only seconds to rank channels in the projected feature space.



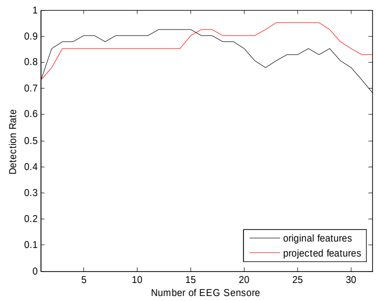
(a) Subject 1, session 1



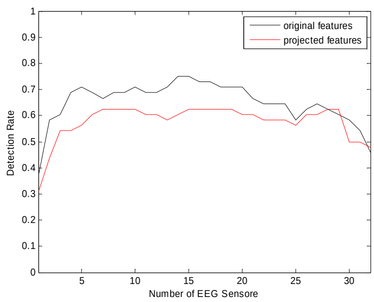
(b) Subject 1, session 2



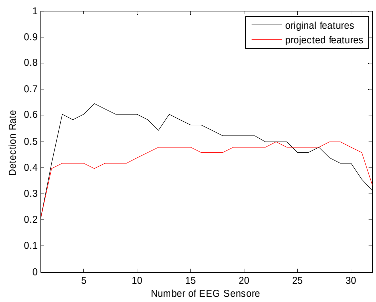
(c) Subject 2, session 1



(d) Subject 2, session 2



(e) Subject 3, session 1



(f) Subject 3, session 2

Fig. 3: Detection Accuracies for different EEG channel ranking methods for all subjects/sessions.

The performances of channel ranking with original features and projected features were compared and shown in Figure 3 (a-f). Each figure contains two curves with different colors, denoted different channel ranking methods. Unsurprisingly, the overall performances of using original features were better than using projected features, since the projection from 129 dimensions to 1 dimension on each channel also eliminated useful information. However, we observe that the best performances of a subset channels ranked from projected features were generally better than using all channels on original features. This indicates that for online real-time BCI systems, dimensionality reduction combined with channel ranking could present a better tradeoff when online channel selection using the original feature space is not feasible.

4 Conclusion and discussion

This preliminary study focused on identifying informative features for ERP speller systems based on the RSVP paradigm. We compared different features extracted from a range of frequency bands and EEG channel subsets.

Experimental results indicated that ERP energy is concentrated in the 0-1Hz and 10-20Hz bands. This indicates that not only P300, but also other ERPs are relevant for ERP based communication. The results were consistent across subjects and sessions although we did not have enough subjects to employ statistical confidence tests.

The performance of the error based feature ranking approach indicated that not all EEG channels were needed for ERP detection. Using a subset of EEG channels yielded better performance than using all channels. Furthermore, this subset was differed across people. Ranking channels using the original features gave accurate results, but it was computation-intensive and not suitable for real-time system. Channel-wise dimensionality reduction followed by ranking channels using the projected features dramatically reduces computational needs. Although this approach sacrifices accuracy, the results were still better when channel selection was performed. We may improve the approach by employing other subspace projection methods instead of using LDA. This constitutes future work.

References

[1] Farwell, L.A., & Donchin, E, Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology, 70, page 512-523, 1988.

[2] K.R.Muller, B.Blankertz, Toward non-invasive Brain Computer Interfaces, IEEE Signal Processing Magazine, vol.23, no.5, pp.126-128, 2006.

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

[4] 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.

[5] T. Lan, Y. Huang, D. Erdogmus, “A Comparison of Temporal Windowing Schemes for Single-trial ERP Detection,” Proceedings of NER 2009, Antalya, Turkey, Apr 2009.

[6] Krusienski,D.J., Sellers,E.W., Cabestaing,F. "A comparison of classification techniques for the P300 Speller" Journal of Neural Engineering vol.3, pp.299-305 (2006).