Single-Trial Event-Related Potential Based Rapid Image Triage System

Ke YU
Division of Bioengineering, National University of Singapore
Singapore, 117576, Singapore

and

Kaiquan SHEN, Shiyun SHAO, Wu Chun NG, Xiaoping LI
Department of Mechanical Engineering, National University of Singapore
Singapore, 117576, Singapore

and

Kenneth Kwok
Temasek Laboratories and DSO National Laboratories
Singapore, 117411 and 117510, Singapore

ABSTRACT

Searching for points of interest (POI) in large-volume imagery is a challenging problem with few good solutions. In this work, a neural engineering approach called rapid image triage (RIT) which could offer about a ten-fold speed up in POI searching is developed. It is essentially a cortically-coupled computer vision technique, whereby the user is presented bursts of images at a speed of 6–15 images per second and then neural signals called event-related potential (ERP) is used as the 'cue' for user seeing images of high relevance likelihood. Compared to past efforts, the implemented system has several unique features: (1) it applies overlapping frames in image chip preparation, to ensure rapid image triage performance; (2) a novel common spatial-temporal pattern (CSTP) algorithm that makes use of both spatial and temporal patterns of ERP topography is proposed for high-accuracy single-trial ERP detection; (3) a weighted version of probabilistic support-vector-machine (SVM) is used to address the inherent unbalanced nature of single-trial ERP detection for RIT. High accuracy, fast learning, and real-time capability of the developed system shown on 20 subjects demonstrate the feasibility of a brain-machine integrated rapid image triage system for fast detection of POI from large-volume imagery.

Keywords: Rapid Image Triage, Real-Time, ERP, Single-Trial and Common Spatio-Temporal Pattern.

1. INTRODUCTION

Like in many other process optimization problems, triage techniques can be extremely effective to improve the efficiency of POI searching in large volume imagery. Recently, some pilot work has been done to explore the feasibility of neurophysiologically-driven rapid image triage methods that leverage split-second human perceptual judgment capability [1]. In these RIT methods, the user, an image analyst or trained personnel, is presented, using rapid serial visual presentation (RSVP) paradigm (which is essentially a visual oddball paradigm [1]), a sequence of image chips, some of which contain POI to be identified. Then, an unique brain signal recorded non-invasively from the scalp known as event-related potential (ERP), represented by P300, a large positive voltage deflection (5 μV or greater) occurring at approximately 300 ms after the onset of a image chip containing POI, is used to determine whether or not the user sees a POI image chip within the high speed sequences of image chips via RSVP.

In those pilot studies, a trainable classifier was used to learn the ERP pattern in relation to brain responses to POI images so as to classify automatically whether the user sees a POI or a non-POI image. Such a cortically-coupled computer vision technique represents a promising performance augmentation for image analysts. Unfortunately, the challenging problem of the single-trial ERP detection has not been carefully addressed and therefore very little evidence exists on the efficacy of incorporating the single-trial ERP detection technique into an effective RIT system. In particular, though surprisingly, no feature extraction was done and only the raw electroencephalogram (EEG) data, congregated from all channels, were to form the feature vector subjected to the classifier. This resulted in an extremely sparse problem: there were thousands of features but only very limited training samples. For such a sparse problem, it is well known in the domain of machine learning that, if no dimensionality reduction is provided, a standard classifier will certainly perform poorly due to the limitation called "curse of dimensionality" [2]. Secondly, the problem of single-trial ERP detection (POI vs non-POI) in the context of RIT is typically a highly unbalanced problem due to the fact that, in a large volume imagery, only very few images may contain POI to be identified. For such an unbalanced problem, a standard classifier used in the past work would likely fail to perform satisfactorily.

This paper reported a RIT system based on novel robust single-trial ERP detection. By searching for both spatial and temporal projections of multi-channel ERP signals, very small number of tempo-spatial features that provide the best separation between ERPs induced by POI and non-POI images can be extracted to aid the classification. Also, a weighted SVM classifier, a special version of SVM tailored to unbalanced classification problems, is used for the first time to solve the unbalance problem of the single-trial ERP detection in RIT. The developed RIT system also comes with a real-time automatic artifact removal functionality that is effective to increase the signal-to-noise ratio for high-accuracy single-trial ERP detection.

2. SYSTEM AND EXPERIMENTAL SETUP

Overview

RIT comprised of two sub-systems: stimulation & acquisition subsystem (SAS), and data analysis subsystem (DAS). Each subsystem ran on a separate laptop, and was linked with each other in the local network following TCP/IP protocol. SAS was responsible for delivering the visual stimuli to the subject, and at the same time acquiring EEG signals through an EEG amplifier. The analog signals from the response button in the system were recorded simultaneously to an auxiliary channel of EEG amplifier. DAS continuously received raw signals transmitted by SAS through the network. The signals were evaluated in real-time and also stored for the possible off-line analysis.

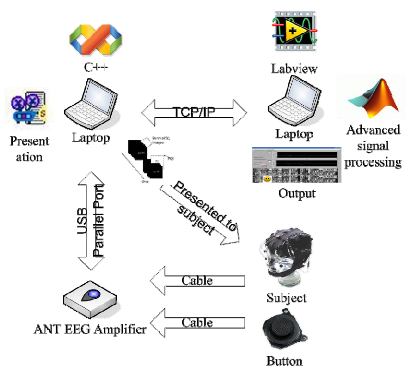


Figure 1. The scheme of RIT system.

The EEG system used was a 64-channel ANT amplifier (ANT B.V., Enschede, Netherlands) with its compatible 64-channel Waveguard Cap, and the two laptops for SAS and DAS respectively have the same specifications (Intel Core 2 Duo T9600 and 2GB DDR2 memory). Software integration includes Presentation (Neurobehavioral Systems Inc., Albany, USA) for stimulus delivery, C++ program for data acquisition and transmission, Labview (National Instrument, Inc., Austin, Texas, USA) for user interface, and Matlab (Mathworks, Inc., Natick, MA, USA) for signal analysis.

Participants

RIT experiments were done on twenty right-handed subjects (21-30 years old, 15 males and 5 females), recruited from local tertiary institutions, who fulfilled the inclusion criteria of not being on any medication, having no history of neurological or psychiatric problems, with normal or corrected-to-normal sight. Recruitment of human subjects for this study was reviewed and approved by the National University of Singapore Institutional Review Board (NUS-IRB).

Data Acquisition and Preprocessing

EEG signals were collected from 62 channels (excluding M1 and M2) over the scalp at a 250-Hz sampling rate, with the reference set to the linked ears and the grounding electrode on the forehead. EEG signals went through a digital band-pass filter of 1 Hz to 25 Hz.

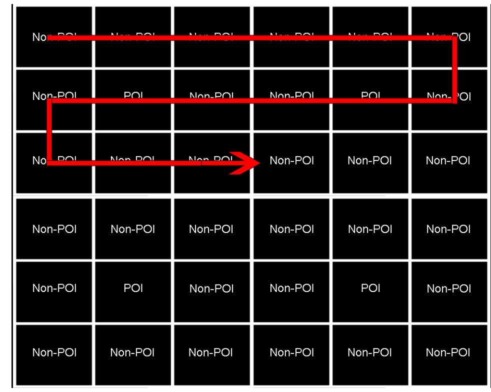
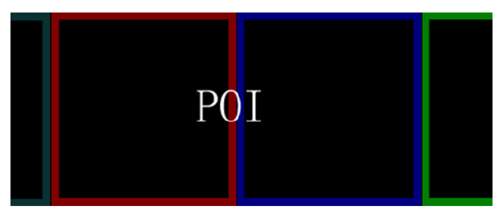


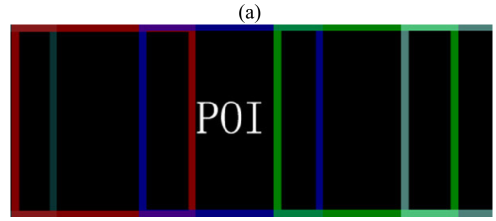
Figure 2. Chips were chopped following the raster scan order. The red arrow demonstrated the track of the raster scan order.

Stimuli Preparation

For the ease of RSVP, the original large imagery was chopped to a sequence of much smaller images (chips). The size of chips was 500×500 pixels except those in the boundary which might be smaller. Spatial information regarding the original location of chips in the imagery was preserved, as it would be useful for registering POI ERP or non-POI ERP to the imagery in the final stage. Considering that human vision is used to consecutive images without sudden changes in context, raster scan order was adopted in chopping (Fig. 2) in the reported experiments.



(a)



(b)

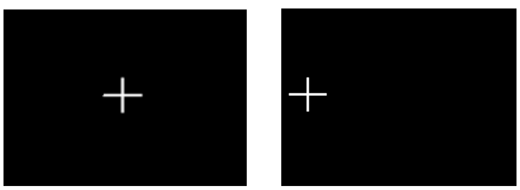
Figure 3. (a) Ordinary chopping. (b) Chopping with overlapping frames.

It is worth pointing out that POI might fall into the boundary of image chips, or even be split by two neighboring chips, as can be seen in Fig. 2(a). POI not lying in the center of the chip could be hard to be detected by the subject, especially in the present case of rapid image triage where image chips were presented to the subject in high speed (6 image chips per second). To address this problem, an overlapping method was adopted instead. In Fig. 2(b), each pair of neighboring chips shared some parts, which ensured that POI would be at least allocated in the centre of one chip. Hence all the POI would have chance to appear in the focal center of subject's visual field. The drawback is that it reduced the efficiency of RIT system as the total number of chips increased.

Experimental Paradigm

Sitting in an adjustable, comfortable chair, each subject underwent a 2-h RIT experiment in a temperature-controlled laboratory in the ambience of silence. Prior to each experiment, a detailed orientation was given to the subject by the operator. Each experiment consisted of one eye calibration session, one training session and one testing session. There were 10-min breaks among sessions.

Eye Calibration: Electrooculogram (EOG) artifact can obscure ERP which are of small amplitudes. In this session the artifact models of eye blinking and eye movement were modeled. The subject was instructed to: 1) blink eyes with repeated flashes of a white cross on a black screen; 2) make horizontal eye movements by following the white cross which alternatively appeared on the left and right of the screen; 3) make vertical eye movements by following the white cross which alternatively appears at the top and bottom of the screen.



(a) (b)

Figure 4. Eye calibration. (a) Subjects blinked upon the disappearance of the white cross in the centre of the screen. (b) Subjects made eye movements while the white cross alternated repeatedly from left to right, up to down.

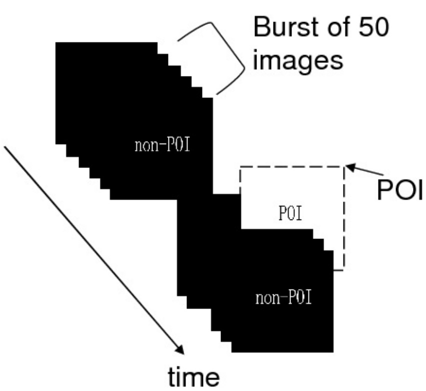


Figure 5. Rapid serial visual presentation in RIT. Each stimulus lasted for 150 ms.

Training and Testing: Following the standard RSVP paradigm [3], in both training session and testing session, subjects were presented bursts of chips. Each burst contained 50 chips, each of which lasted for 150 milliseconds on the screen (Fig. 5). Every burst was separated by a fixation screen (a black screen with a fixation cross in the center) for a subject-controlled duration (up to 10 seconds) to break the monotony and to minimize possible eye strain. In training session, POI chips were randomly inserted into bursts, in a manner that each burst contained at most one POI chip, and the POI chip was not among the first and the last 10 chips of the burst. The duration of training session was determined by the subject's performance. If the subject had correctly responded to POI chips by pressing the button to pre-defined times, training session would automatically terminate. This ensured adequate data for building robust classification model. In testing session, chips were arranged and presented according to the raster scan order (Fig. 2).

3. DATA PROCESSING

Artifact Removal

A linear modeling approach [4] is used to deal with the typical artifacts that contaminated EEG activity, i.e. EOG artifacts. This method assumes a linear model between the underlying sources and the scalp potentials. After deriving a forward model for eye blinking, horizontal and vertical eye movements based on the data collected in eye calibration session, the artifact-free EEG signals were reconstructed by subtracting the estimated contribution of EOG artifacts to the observed raw EEG in the training and testing sessions.

Segmentation

The preprocessed EEG signals were segmented into single-trial ERP epochs according to event markers. Each epoch was an EEG segment falling into the event-locked window of [-200ms 500ms], i.e. from 200 ms before to 500 ms after the onset of each image. For each epoch, the baseline mean was calculated by using the data within the window [-200ms 0ms] and subsequently subtracted from each channel.

Feature Extraction

The developed RIT system uses a novel feature selection method, termed common spatio-temporal patterns (CSTP). Unlike past methods such as conventional common spatial pattern (CSP) method whereby only spatial patterns of ERP are considered [5-7], this method exploits both spatial and temporal patterns of ERP, providing complementary spatial and temporal features for high-accuracy single-trial ERP detection.

Let X_c be a single epoch matrix (channel$\times$time) in condition *c,* where *c* is either POI condition (+) or non-POI condition (-). The normalized spatial covariance R_c and the normalized temporal covariance $\overline{R_c}$ can be obtained from:

$$\begin{cases} R_c = \dfrac{X_c X_c^T}{trace(X_c X_c^T)} \\ \overline{R_c} = \dfrac{X_c^T X_c}{trace(X_c^T X_c)} \end{cases} \tag{1}$$

where $(\bullet)^T$ stands for the transpose operator and $Trace(\bullet)$ is the summation of diagonal elements. By averaging R_c and $\overline{R_c}$ over the trials in condition *c*, the spatial covariance Σ_c and the temporal covariance $\overline{\Sigma_c}$ can be archived, respectively.

The motivation of the CSTP method is to find both spatial filters v and temporal filters $\bar{v}$ that maximize the variance of filtered signals of one condition and at the same time minimize the variance of filtered signals of another condition. Since the spatial filters v and temporal filters $\bar{v}$ that provide best separation between two conditions are independent, and can be obtained by the same technique used in conventional common spatial pattern method [5-7].

Let's assume P be the whitening transformation of Σ_{cm} which satisfies

$$P\Sigma_{cm}P^T = P\Sigma_+P^T + P\Sigma_-P^T = I \quad (2)$$

where I refers to the identity matrix. Let S_+ and S_- denote $P\Sigma_+P^T$ and $P\Sigma_-P^T$, respectively. Since $S_+ + S_- = I$ from (3), it follows from spectral theorem for matrices that S_+ and S_- share common eigenvectors, i.e., suppose $S_+ = B^T \Lambda_+ B$, then $S_- = B^T \Lambda_- B$ and $\Lambda_+ + \Lambda_- = I$. Λ_+ and Λ_- are the diagonal matrices of corresponding eigenvalues for S_+ and S_-, respectively. Therefore a bigger eigenvalue in one condition will have a corresponding smaller eigenvalue in the other condition, and vice versa. This attribute grants eigenvectors B the ability to discriminate two conditions. Integrated with the whitening transformation, the project matrix is written as $V = (BP)^T$, with each column being a spatial filter v.

Similarly, the desired temporal filters $\overline{V}$ for optimal separation of two conditions can be found by simultaneous diagonalization of $\overline{\Sigma}_+$ and $\overline{\Sigma}_-$.

Finally, the mapping of a single epoch matrix X is

$$\begin{cases} Y = V^T X \\ \overline{Y} = \overline{V}^T X^T \end{cases} \quad (4)$$

Typically, only the filters corresponding to the biggest difference in eigenvalues between two conditions are used. Features for classification are variances of projected signals on the chosen filters.

Classification

The classification of POI ERP vs. non-POI ERP is an extremely unbalanced problem due to the fact that, in many practical image triage problems, only very few images may contain POI to be identified. Handling such an extremely unbalanced problem is still an ongoing research topic. This work uses a modified version of probabilistic SVM, called weighted probabilistic SVM [8, 9], to accommodate the unbalance nature of the problem. The weighted probabilistic SVM uses real unbalanced data for training and compensates the bias of prior class probabilities by penalizing more on the classification errors produced by the samples from the minority class. By doing so, it offers a better tradeoff among classification performance on each class by greatly increasing classification accuracy on minority class at the cost of a relatively minor decrease in classification accuracy on majority class, making it a preferable solution to solving unbalanced problems. In addition, unlike the standard SVM which only gives hard decision, the weighted probabilistic SVM provides a useful confidence estimate of each classification that it makes, through a elegant mapping from the SVM outputs to posterior probabilities [10]. The detailed algorithm of the weighted probabilistic SVM can be found in [8-10].

4. RESULTS

For each subject, there was a different set of 4,800 non-POI images versus 70 POI images for each of the training and testing sessions. The hit rate of the trained RIT system on twenty subject was 81±11% of about 70 POI in the testing session, while the false alarm rate was about 12%.

The triage performance can be visualized by highlighting POI regions in the original imagery with posterior probability mapping. Posterior probabilities representing the likelihood at which each chip belongs to POI category are estimated based on single-trial ERP detection, and are further interpolated and converted to a color-coded hotspot layer (Fig. 6) which can be overlaid on the original imagery. Those chips with relatively higher posterior probability are marked in red or yellow, whilst chips with lower posterior probability are marked in blue.

The developed RIT system, in essential a cortically-coupled computer vision technique, offers significant performance enhancement on POI searching in original imagery by leveraging split-second human perceptual judgment capability. The image chips were presented to the user at a speed of 6 images per second, which represents multi-fold speed up in POI searching [1] though we did not conduct a control experiment for accurate quantification of the performance enhancement. This certainly warrants further study, preferably through field tests by intended users.

Like all other ERP-based brain-computer interface systems, the developed RIT system requires pre-learning/calibration so that the system can function correctly for a new user. It's worth noting that the developed system did offer fast learning. It only took about 15 mins to go through a training session for system pre-learning/calibration (including model selection), with minimal user intervention.

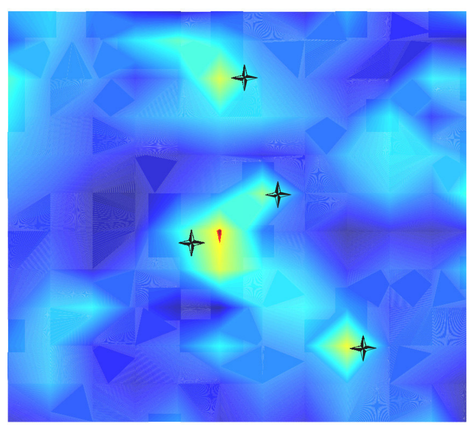


Figure 6. The plot of the posterior probabilities of subject seeing POI images given by the developed RIT system. illustrates the actual position of POI.

The developed system also has real-time capability. Following each chip shown to the user, the triage result in terms of the posterior probability of that chip being a POI chip was determined in a near real-time fashion. Therefore, feedback of triage performance to the user is possible with the developed RIT system. The influence of feedback on triage performance and even the ERP morphology is unknown, but it is another interest area for future work.

5. REFERENCES

[1] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 14, no. 2, pp. 174-179, 2006.

[2] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed., New York: Wiley, 2000.

[3] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature,* vol. 381, no. 6582, pp. 520-2, Jun 6, 1996.

[4] L. C. Parra, C. D. Spence, A. D. Gerson *et al.*, "Recipes for the linear analysis of EEG," *Neuroimage,* vol. 28, no. 2, pp. 326-41, Nov 1, 2005.

[5] K. Fukumizu, "Active learning in multilayer perceptrons," *Advances in Neural Information Processing Systems 8. Proceedings of the 1995 Conference.* pp. 295-301.

[6] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *Rehabilitation Engineering, IEEE Transactions on,* vol. 8, no. 4, pp. 441-446, 2000.

[7] K.-R. Muler, M. Tangermann, G. Dornhege *et al.*, "Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring," *Journal of Neuroscience Methods,* vol. 167, no. 1, pp. 82-90, 2008.

[8] E. E. Osuna, R. Freund, and F. Girosi, *Support Vector Machines: Training and Applications,* Technical Report AI Memo 1602, MIT, 1997.

[9] S. Y. Shao, K. Q. Shen, C. J. Ong *et al.*, "Automatic EEG artifact removal: a weighted support vector machine approach with error correction," *IEEE Trans Biomed Eng,* vol. 56, no. 2, pp. 336-44, Feb, 2009.

[10] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Advances in Large Margin Classifiers*, A. Smola, P. Bartlett, B. Scholkopf *et al.*, eds., Cambridge, MA: MIT Press, 2000.