Spatio-Temporal Detection of Divided Attention in Reading Applications Using EEG and Eye Tracking

Mathieu Rodrigue,[1] **Jungah Son,**[2] **Barry Giesbrecht,**[3] **Matthew Turk,**[1,2] **Tobias Höllerer**[1,2]
[1]Computer Science, University of California, Santa Barbara
[2]Media Arts and Technology, University of California, Santa Barbara
[3]Psychological and Brain Sciences, University of California, Santa Barbara
{mathieu, mturk, holl}@cs.ucsb.edu, jungah@umail.ucsb.edu, barry.giesbrecht@psych.ucsb.edu

ABSTRACT
Reading is central to learning and communicating, however, divided attention in the form of distraction may be present in learning environments, resulting in a limited understanding of the reading material. This paper presents a novel system that can spatio-temporally detect divided attention in users during two different reading applications: typical document reading and speed reading. Eye tracking and electroencephalography (EEG) monitor the user during reading and provide a classifier with data to decide the user's attention state. The multimodal data informs the system where the user was distracted spatially in the user interface and when the user was distracted. Classification was evaluated with two exploratory experiments. The first experiment was designed to divide the user's attention with a multitasking scenario. The second experiment was designed to divide the users attention by simulating a real-world scenario where the reader is interrupted by unpredictable audio distractions. Results from both experiments show that divided attention may be detected spatio-temporally well above chance on a single-trial basis.

Author Keywords
Attention classification; EEG; eye tracking; learning tools;

ACM Classification Keywords
H.1.2 Information Systems: User/Machine Systems

INTRODUCTION
It is often easy to become distracted while reading, leaving the reader to not fully understand or acquire knowledge from portions of text. Typically, this happens when distracting sounds are heard or when dividing attention while attempting to multitask. When learning, divided attention creates interference during the encoding process of the brain, thus affecting memory and retrieval of the information [7]. This could have a large, negative impact in fields where sustained attention and understanding is necessary. For example, in an educational or workplace environment, one might be distracted by surrounding peers, causing a decrease in task performance.

IUI'15, March 29–April 1, 2015, Atlanta, GA, USA.

http://dx.doi.org/10.1145/2678025.2701382

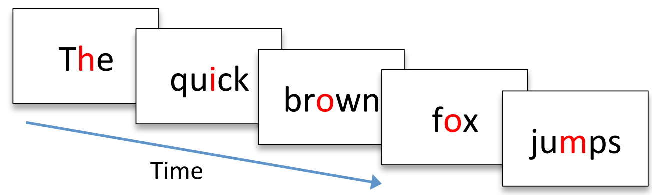

Figure 1. Users focus on the sequential red letters during speed reading.

Interest in reading applications is larger than ever before, given mobile devices with high resolution screens and digital documents such as PDFs and e-books. Digital reading materials are increasingly popular these days due to the fact that they are instantaneously and easily accessible through the internet, and typically cheaper than their physical counterparts. The change from physical to digital reading modalities makes it possible to more easily employ perceptual devices [14] that monitor one's attention state. Here, we discuss two types of reading applications and introduce a novel system to detect divided attention during reading.

Ordinary reading applications are the most common as they are analogous to reading a physical book. In this style of reading application, the user reads from left to right, and top to bottom. Pages are turned by button press or touch gesture. Speed reading has gained recent interest with mobile applications such as Spritz [3]. A speed reading application shows individual, consecutive words to a user (Figure 1), at a rate set by the user. By eliminating saccades and other potentially-distracting words in the periphery, one can read at much higher speeds. Speed reading applications are particularly useful on mobile devices, with screens too small to provide a comfortable reading experience of full-sized documents (e.g. mobile phones, smart watches). By allowing the device to rapidly display consecutive words, the user must be fully attentive at all times. Once a word, sentence, or paragraph is missed, the user may not be able to easily navigate back, or may not be aware of the precise point where attention was lost. This presents a problem to users who are, for example, multitasking or easily distracted by audio within their vicinity.

Novel Attention Monitoring

In this paper, we present a novel system, employing consumer-grade EEG and eye tracking devices, to detect spatially (in the reading application) and temporally if a user is either attending to a reading application or dividing their attention. The contributions of this paper are two-fold: 1) an algorithm for accurately detecting divided attention in reading applications, and 2) the coupling of consumer-grade EEG and eye tracking modalities to improve detection.

To explore the feasibility of classifying between attention and divided attention, two experiments were conducted for each reading task. The first experiment explores to what degree divided attention can be detected in a multitasking scenario. The second experiment explores to what degree divided attention can be detected in the presence of real-world unpredictable distractions.

RELATED WORK

There have been a handful of attempts to accurately measure states of attention using a wide variety of EEG devices. Berka et al. developed one of the first commercial systems, known as B-Alert [6], to detect states of attention (high-engagement, low-engagement, relaxed wakefulness, and sleepy) in real time by a medical grade EEG device, but were not configured to detect divided attention. Hamadicharef et al. were the first to apply the filter bank common spatial pattern (FBCSP) [8] algorithm to an attention task with EEG data. They found that the FBCSP method classified up to 89.4% between states of attention and relaxation using a 15-channel medical grade EEG device. Liu et al. [10] used a consumer-grade EEG device with one electrode (Fp1) to create an attention classifier from pooled subject data. The participants listened to English phrases then answered related questions on a quiz under attentive and distracted conditions. Classification was reported to be on average 76% accurate (55%-60% for the inattention class, 87%-90% for attention, depending on the cost function used). Martínez-Gómez et al. [11] found characteristic features from eye tracking data that describe a subject's level of understanding and English language skill. Putze et al. [13] used the combination of EEG and eye tracking to facilitate video surveillance event selection. They demonstrated accurate event detection in an abstract event selection task, using a set of five electrodes.

The majority of the above attempts were performed unimodally, with devices far too expensive for the average consumer. Similarly, the previously stated work has not attempted classifying divided attention from attention. Hence, we explore the possibility of combining low-cost EEG and eye tracking to detect states of attention from divided attention in popular reading applications.

BRAIN-COMPUTER INTERFACE

The Emotiv Epoc [1], a consumer-grade EEG device, was used for signal acquisition. The device samples at 128hz and is equipped with 14 electrodes (AF3, AF4, F3, F4, F7, F8, FC5, FC6, O1, O2, P7, P8, T7, T8) which are positioned according to the international 10-20 system. The brain-computer interface developed for this system detects changes in power from the frequency bands α (8-13hz), β (14-30hz), and θ (4-7hz) between divided and non-divided attention states.

EEG Preprocessing

EEG data was first low-pass filtered with a cutoff frequency of 50hz and high-pass filtered with a cutoff frequency of 0.16hz, both using a third-order butterworth filter. The data was then segmented into 1.5-second epochs, overlapping each previous epoch by 50%.

Epochs that contained eye blinks were detected using an eye tracker, and filtered according to [15]. An artifact removal algorithm was implemented based on Adaptive SWT-based Denoising (ASWTD). For the detected EEG segments, the wavelet coefficients at level 6 were obtained using the stationary wavelet transform (SWT) and thresholded with the adaptive thresholding mechanism using a hard thresholding function. Denoised signals without artifacts were obtained by computing an inverse SWT on the thresholded wavelet coefficients. Since we were removing ocular artefacts, only the wavelet coefficients from levels 3 to 6 were thresholded. Epochs with other types of artifacts above a predetermined threshold were discarded.

Filter-Bank Common Spatial Pattern

EEG data was spatially filtered according to an adapted FBCSP algorithm [5], which involved IIR filtering the data into frequency bins, then performing the common spatial pattern algorithm on each frequency bin to find the source location of the signal on the scalp. CSP spatially filters the EEG channels by determining a matrix W such that $W\Sigma_1 W^\top = D_1$ and $W\Sigma_2 W^\top = D_2$ where Σ_1 and Σ_2 are the estimated class covariance matrices, and D_1 and D_2 are diagonal matrices such that $D_1 + D_2 = I$. The first and last n ($n = 6$, in practice) columns of W contain spatial filters that maximally discriminate between the two classes.

Each preprocessed EEG epoch was first band pass filtered, using a third-order butterworth filter, in the α, β, and θ frequency bands. Literature [9] suggests that changes in the chosen frequency bands are correlated with attention. A dedicated portion of the EEG data was deployed to train a corresponding spatial filter W_α, W_β, and W_θ using the common spatial pattern algorithm (CSP). The remaining band-pass filtered epochs were then spatially filtered by their corresponding spatial filter, using the $n = 6$ most discriminative spatial filters as reported by the CSP algorithm.

EEG Classification

Offline EEG data was determined to be representative of attention or divided attention by extracting features from the CSP source signals, and classifying test examples with a trained classifier. Log-variance features were extracted from the spatially filtered data $f_\alpha = \log(\mathrm{var}(W_\alpha X))$, $f_\beta = \log(\mathrm{var}(W_\beta X))$, $f_\theta = \log(\mathrm{var}(W_\theta X))$. 10x10-fold cross validation with a support vector machine (SVM) classifier was employed using the MATLAB implementation with an RBF kernel, to decide if an EEG epoch was in the attention or divided attention state. A grid search was performed to optimize σ for all participants, the remaining parameters were left as default.

EYE TRACKING

The Eye Tribe [4], a consumer-grade eye tracker was used for this work. Raw data was recorded at 60hz and divided into four-second epochs. Each EEG epoch had a corresponding,

Table 1. Average classification accuracy and per-class classification accuracy for EEG, eye tracking and combined modalities.

		EEG			Eye Tracking			EEG + Eye Tracking		
		Average	Att.	Div. Att.	Average	Att	Div. Att.	Average	Att.	Div. Att.
Exp. I_{SR}	Subject 1	95.93%	97.45%	94.40%	63.04%	71.43%	54.05%	96.54%	98.43%	93.74%
	Subject 2	97.68%	97.22%	98.13%	65.99%	54.7%	77.28%	97.79%	96.92%	98.66%
	Subject 3	87.99%	88.21%	87.77%	67.81%	65.88%	69.75%	90.06%	90.97%	89.15%
Exp. I_R	Subject 1	98.16%	99.59%	96.73%	72.26%	69.53%	74.09%	97.8%	99.28%	95.09%
	Subject 2	99.97%	99.95%	100%	74.01%	71.6%	76.42%	99.19%	99.82%	98.56%
	Subject 3	92.41%	93.02%	91.81%	63.14%	51.98%	74.31%	91.25%	93.51%	88.99%
Exp. II_{SR}	Subject 1	84.52%	85.71%	83.33%	66.31%	68.77%	64.02%	83.51%	82.35%	84.60%
	Subject 2	74.82%	73.76%	75.89%	72.34%	75.18%	69.5%	79.43%	75.18%	83.69%
	Subject 3	69.07%	65.56%	73.33%	73.61%	71.48%	75.74%	82.41%	80.0%	84.81%
Exp. II_R	Subject 1	85.71%	88.10%	83.33%	78.57%	73.81%	83.33%	90.48%	95.24%	85.71%
	Subject 2	77.14%	80.0%	74.29%	69.29%	57.14%	81.43%	84.29%	81.43%	87.14%
	Subject 3	75.93%	74.63%	77.22%	71.39%	70.37%	72.41%	80.93%	79.26%	82.59 %

four-second eye tracking epoch, aligned at the last sample. The Eye Tribe eye tracking algorithms are currently in active development. We used version 0.9.41 of the SDK, and a chin rest was used to minimize noise.

Fixation and Saccade Detection

An offline saccade and fixation detection algorithm adapted from [12] was employed. Saccades were found by finding the distance between the mean of two consecutive sliding windows,

$$m_{before}(n) = \left[\frac{1}{r}\sum_{k=1}^{r} s_x(n-k), \frac{1}{r}\sum_{k=1}^{r} s_y(n-k)\right]$$

$$m_{after}(n) = \left[\frac{1}{r}\sum_{k=1}^{r} s_x(n+k), \frac{1}{r}\sum_{k=1}^{r} s_y(n+k)\right]$$

where n is the sample of interest, and r is the window size. The distance d is calculated at every sample in eye tracking data by $d = \sqrt{(m_{after} - m_{before}) \cdot (m_{after} - m_{before})^\top}$ which forms a sequence of peaks, where the max of each peak represents a saccade. Once saccades are detected in the epoch, the median of the samples between saccades are marked as fixations. This algorithm may be applied to real-time system where data is processed in epochs.

Eye Tracking Classification

Four of the most discriminative eye tracking features from [11] were used for in our classifier: number of fixations, median saccade length, mean saccade velocity x, and mean saccade velocity y. Among other tested features, the chosen features were found to best characterize between states of attention in this experiment. Median was chosed instead of mean saccade length so that outliers would not affect classification. Additionally, log-variance features for saccade length and mean pupil size were employed after observing discriminative differences in the data. The features were added to their corresponding EEG-epoch feature vector and then classified with the SVM classifier to determine attention states.

EXPERIMENTS

In this section, we'll discuss two experiments to study the effects of distraction from multitasking, and external stimuli on classification. Both experiments required the subject to complete two tasks, both split up into attention and divided attention conditions. The first task involved speed reading a passage of text. Each session lasted approximately two minutes. An open source speed reading application, Spray [2], was used for this task. The second task required the subject to read a normal passage of text, one page at a time. Each session lasted the necessary amount of time for the subject to finish two pages, which was approximately two minutes. Participants were not allowed to go back to a previous page, but were encouraged to read as they normally would. Both tasks were split up into two conditions: attention and divided attention, and required the subjects to wear headphones. After each task, the participant reported a summary of what they had read to ensure the subject was participating in the task. EEG electrodes were checked after each task to ensure proper impedance levels, and eye tracking calibration was checked between experiments, but no recalibration was ever needed.

Experiment I

The first experiment was designed to simulate internal divided attention, such as multitasking. Three participants (one female, two males, ages 25-27) volunteered for the experiment. All had normal or corrected to normal vision. A total of eight sessions were performed, within subjects, for each task. The names of nine different colors were spoken through the headphones in a random order at one-second intervals. The volume of the headphones were set to be as low as possible while still allowing the subject to hear the colors clearly. During the attention condition, the participant was asked to ignore the colors and completely focus on the speed reading task. During the divided attention condition, the participant was told to focus on the reading material but also count the number of times a target color was spoken during the session. The participant reported the number of colors heard after each session, which was compared to ground truth data to ensure the user was dividing attention between both stimuli.

Experiment II

The second experiment was designed to simulate an environment where a reader is distracted by external stimuli. Three participants (three male, ages 22-34) volunteered for the experiment. All had normal or corrected to normal vision. A total of four sessions were performed, within subjects, for each task. The headphones were set to a moderate volume as determined by the participant. During the reading task, the participant heard multiple, unique sound clips that lasted for 30 seconds each and would play randomly, but never overlap.

The sound clips contained distracting noises one might typically hear in their environment while reading, such as energetic music, television shows, movies, dogs barking, or conversations from other people. After each session, the participant reported how distracted they were during audio playback (1-9 Likert scale, 9 being most distracted). Distractions rated below 4 were not used for classification. Three detachable EEG electrodes were found to be damaged and thus excluded from data acquisition: (F3, FC6, P7), (AF3, F3, FC5), and (F7, P7, T7), for subjects one, two and three, respectively.

RESULTS AND DISCUSSION

The objective of this work was to develop algorithms which allow EEG and eye tracking consumer devices to accurately detect divided attention from attention during reading. Additionally, we wanted to provide a proof of concept for mapping human attention into a user interface, so software could intelligently react to and provide useful information about a reader's experience. A user's eye fixations are seen in (Figure 3) which shows a clear discrimination between attention and divided attention conditions.

To avoid any bias in the data, a 10x10-fold cross-validation was used on the feature data for both experiments. Table 1 shows that classification accuracy of EEG and combined modalities in experiment I exceeds those for experiment II for both tasks. Figure 2 shows the higher dimensional EEG features from the highest classification accuracy obtained in experiment I. The feature data was projected onto the two largest principal components in order to visualize the higher dimensional feature space. It is clear from the visualization that the two classes are highly separable, resulting in accurate classification. The first experiment simulated a multitasking scenario, where the reader may be engaging in other activities while they are reading, e.g. listening to a breaking news story on the television while reading. This type of task was designed to keep the subject's attention divided among two sets of stimuli. The EEG classification accuracy is high, most likely due to the fact that the multitasking distraction effect was stronger, resulting in different power in the frequency bands between conditions. This experiment ensured that classification was based on true divided attention and not driven by auditory perception signals in the brain, given that an identical auditory signal (a spoken set of colors in pseudo-random order) was present across both conditions.

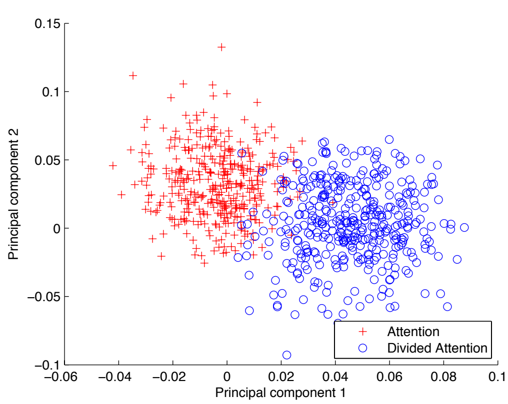

Figure 2. First two principal components of EEG features from subject two during the ordinary reading condition for experiment I.

Figure 3. Classified user fixation data visualized from experiment I. Red circles represent divided attention, blue circles represent attention.

Whereas experiement I was highly controlled to rule out the possibility that auditory brain signals are driving classification results, experiment II was designed to test a more realistic scenario in which a user's attention is divided by unexepected and intrusive auditory distractions. Table 1 shows a lower classification percentage using both modalities for experiment II, but still well above chance. One possibility for lower classification accuracy was the exclusion of three EEG electrodes for each participant. Also, the second experiment did not guarantee that a subject would be distracted during the distraction condition. Given that some people may not have an issue with background noise as they read, e.g. people who commonly read with a running television in the room, this experiment may not have distracted all of the participants equally. This was also observed from the post-session questionnaire, which confirmed that the external stimuli were not fully distracting to all participants. Mean scores for distraction level during the two tasks were (6, 6.25), (7.25,4.75), and (4, 4.5), with their respective standard errors (0.629, 0.722), (0.913, 0.479), (0.408,0.654) for the three participants. Additionally, EEG classification from this experiment may have been partially influenced by the user's perception of hearing different, intermittent audio clips.

CONCLUSION

We proposed a novel approach to detecting attention and divided attention during reading applications, using EEG and eye tracking consumer devices. Results from two experiments show that EEG classification is highly accurate during controlled multitasking scenarios, and still well above chance during unpredictable distraction scenarios. On the other hand, eye tracking data while not as accurate as EEG data, may be effective as a single mode of input for certain types of applications. Additionally, by combining EEG and eye tracking features, classification generally performs better than either modality on its own. Future work will focus on classifying attention from divided attention in real-time. For example, the system could make note of when the user was distracted during speed reading, providing an easy way to navigate to areas the user missed, or warn the user, to regain focus. Similarly, during ordinary reading, the system could make note of spatio-temporal data received from the EEG and eye tracker, and remind the user to go back and review those areas. Other future work involves detection and multiclass classification among other states of attention (e.g. mind wandering), and attention detection in other evironment scenarios.

ACKNOWLEDGEMENTS

This work was partially supported by ONR grants N00014-14-1-0133 and N00014-13-1-0872, as well as NSF grant IIS-0747520.

REFERENCES

1. Emotiv EPOC. `http://www.emotiv.com/`.
2. Spray Open Source Speed Reader. `https://github.com/chaimpeck/spray/`.
3. Spritz. `http://www.spritzinc.com/`.
4. The Eye Tribe Eye Tracker. `http://www.theeyetribe.com/`.
5. Ang, K. K., Chin, Z. Y., Zhang, H., and Guan, C. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *IEEE International Joint Conference on Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence)* (June 2008), 2390–2397.
6. Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., and Olmstead, R. Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human-Computer Interaction 17*, 2 (June 2004), 151–170.
7. Fernandes, M. A., and Moscovitch, M. Divided attention and memory: evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology. General 129*, 2 (June 2000), 155–176.
8. Hamadicharef, B., Zhang, H., Guan, C., Wang, C., Phua, K. S., Tee, K. P., and Ang, K. K. Learning EEG-based spectral-spatial patterns for attention level measurement. In *IEEE International Symposium on Circuits and Systems, 2009. ISCAS 2009* (May 2009), 1465–1468.
9. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews 29*, 2–3 (Apr. 1999), 169–195.
10. Liu, N.-H., Chiang, C.-Y., and Chu, H.-C. Recognizing the degree of human attention using EEG signals from mobile. *Sensors 13*, 8 (Aug. 2013), 10273–10286.
11. Martínez-Gómez, P., and Aizawa, A. Recognition of understanding level and language skill using measurements of reading behavior. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, ACM (New York, NY, USA, 2014), 95–104.
12. Olsson, P. Real-time and offline filters for eye tracking. *Master Thesis* (2007).
13. Putze, F., Hild, J., Kärgel, R., Herff, C., Redmann, A., Beyerer, J., and Schultz, T. Locating user attention using eye tracking and EEG for spatio-temporal event selection. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, IUI '13, ACM (New York, NY, USA, 2013), 129–136.
14. Turk, M., and Robertson, G. Perceptual user interfaces (introduction). *Commun. ACM 43*, 3 (Mar. 2000), 32–34.
15. Yong, X., Fatourechi, M., Ward, R. K., and Birch, G. E. Automatic artefact removal in a self-paced hybrid brain-computer interface system. *Journal of NeuroEngineering and Rehabilitation 9*, 1 (July 2012), 50.