Eye Gaze-based Unknown Word Detection in Non-native Language Reading using SVMs and Random Forests

Rui Hiraoka, Hiroki Tanaka, Sakriani Sakti, Koichiro Yoshino, Graham Neubig, and Satoshi Nakamura

***Abstract*— This paper proposes machine-learning based method to detect unknown words during natural reading by using eye-tracking features. A previous approach [1] utilizes gaze duration and word rarity features for detection. However, the performance of the previous approach is not sufficient during natural reading. To improve detection performance, we try to 1) apply support vector machines (SVM) [2] and Random Forests (RF) [3], and 2) use novel eye movement features that were not considered in the previous work. The experimental results demonstrate that SVM and proposed eye movement features are capable of improving detection performance as measured by F-measure.**

I. INTRODUCTION

Eye movement has been considered as major indicator of human reading strategies [4], [5]. An unknown word detection algorithm based on total gaze duration and word rarity was proposed by Hyrskykari [1], who used a word-unit threshold function calculated using word frequency counts from a large corpus[1]. However, the system requires training of users to achieve sufficient detection performance. Therefore it is not effective for natural reading by untrained users. This paper proposes a new unknown word detection method which focuses on optimization for natural reading by untrained users.

II. PROPOSED METHOD

We examined various features including word rarity, word length, and three eye gaze features: first gaze duration, max gaze duration, and total gaze duration. The contribution of each feature was evaluated by the correlation coefficient with unknown-known label as well as the p-value on a two-way t-test between those labels. Moreover, we utilized two well-known machine learning methods, RF and SVM with radial basis function (RBF) kernels, as classifiers on the most effective eye gaze features.

III. EXPERIMENTAL EVALUATION

We recorded eye movement data from 5 Japanese speakers naturally reading English. As the eye-tracker, we used a Tobii pro x2-30. Each participant read 600-1000 words and annotated whether each word was unknown or known. 279 out of 4462 words were annotated as unknown in total. Table I shows the ranking of each feature contribution. Max gaze duration was the most effective eye gaze feature as measured by correlation coefficients and p-values. Table II shows comparison of classifiers and eye gaze features by 10-fold cross validation. As a baseline, we considered the threshold function [1]. The results shows that the proposed methods using RF and SVM with max gaze duration features improve performance as measured by F-measure. The performance achieved by SVM was significantly improved compared with the baseline threshold function and RF ($p < 0.01$).

TABLE I

FEATURE SELECTION RANKING BY CORRELATION COEFFICIENT r.

Feature	r	p-value on t-test
Word rarity: W_R	0.499	.x10^{-151}
Max gaze duration: $g_{\max}$	0.375	.x10^{-28}
Word length: W_L	0.353	.x10^{-94}
First gaze duration: g_{first}	0.336	.x10^{-20}
Total gaze duration: g_{total}	0.305	.x10^{-24}

TABLE II

10-FOLD CROSS VALIDATION CLASSIFICATION SCORES.

Method	Baseline	RF	RF	SVM	SVM
Precision	0.612	0.544	0.594	0.485	0.520
Recall	0.412	0.466	0.523	0.763	0.703
False Positive	0.016	0.026	0.024	0.050	0.041
F-measure	0.493	0.502	0.556	0.593	0.598
Features	W_R	W_R	W_R	W_R	W_R
	g_{total}	g_{total}	$g_{\max}$	g_{total}	$g_{\max}$

IV. CONCLUSION

We proposed several eye gaze features and classifiers to detect unknown words. Using the most effective eye gaze features and SVM with RBF kernel, we achieved a significant improvement in detection F-measure from the 0.493 baseline to 0.598.

*This work was supported by JSPS KAKEN 26540117 and 16K16172.

R. Hiraoka, H. Tanaka, S. Sakti, K. Yoshino, G. Neubig and S. Nakamura are with Graduate School of Information Science, Nara Institute of Science and Technology, Takayama-cho 8916-5, Ikoma-shi, Nara, Japan {hiraoka.rui.hj9, hiroki-tan, ssakti, koichiro, neubig, s-nakamura}@is.naist.jp

[1]British National Corpus (2005). Oxford University Computing Services, at http://www.natcorp.ox.ac.uk/

REFERENCES

[1] A. Hyrskykari, *Eyes in attentive interfaces: Experiences from creating iDict, a gaze-aware reading aid*. Tampereen yliopisto, 2006.

[2] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[3] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[4] P. Martínez-Gómez and A. Aizawa, "Recognition of understanding level and language skill using measurements of reading behavior," in *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 2014, pp. 95–104.

[5] K. Kunze, H. Kawaichi, K. Yoshimura, and K. Kise, "Towards inferring language expertise using eye tracking," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 217–222.