

Document Classification on Relevance: A Study on Eye Gaze Patterns for Reading

Daniel Fahey, Tom Gedeon, and Dingyun Zhu

Research School of Computer Science,
College of Engineering and Computer Science,
The Australian National University,
Acton, Canberra, ACT 0200, Australia
{daniel.fahey,tom.gedeon,dingyun.zhu}@anu.edu.au

Abstract. This paper presents a study that investigates the connection between the way that people read and the way that they understand content. The experiment consisted of having participants read some information on selected documents while an eye-tracking system recorded their eye movements. They were then asked to answer some questions and complete some tasks, on the information they had read. With the intention of investigating effective analysis approaches, both statistical methods and Artificial Neural Networks (ANN) were applied to analyse the collected gaze data in terms of several defined measures regarding the relevance of the text. The results from the statistical analysis do not show any significant correlations between those measures and the relevance of the text. However, good classification results were obtained by using an Artificial Neural Network. This suggests that using advanced learning approaches may provide more insightful differentiations than simple statistical methods particularly in analysing eye gaze reading patterns.

Keywords: Document Classification, Relevance, Gaze Pattern, Reading Behavior, Statistical Analysis, Artificial Neural Networks.

1 Introduction

When people read they display some personal behaviours (usually without noticing it) that break the standard reading paradigm. These differences may be a defining factor on how well a person understands the material that they are reading, or how well they understand information in general. Is it possible to identify a pattern or a key factor, in a person's reading pattern, that can explain how well they will understand the information they are reading? If it is, then a method could be created to measure a person's understanding of some material based entirely on the way that they read that material.

With the motivation of studying eye gaze patterns particularly for reading, an experiment has been conducted to test how well a person can understand the premise for a paper when they are given paragraphs from that paper in a random order. Of the paragraphs that are given only half contain much useful information

while the other half contain much less. The experimental participants read the paragraphs with their eye gaze being tracked using a computerised eye-tracking system. Questions were asked and some other tasks referring the paragraphs were completed to score a participant's understanding of the original paper. The results of this experiment are expected to be used to try and find if there is some characteristic of a persons gaze pattern that can be attributed to having a better or worse understanding of the information. This could be used to devise a method of testing people for how well they understand information.

2 Eye Gaze for Reading

Apart from the research work on using eye gaze as an input for conventional user interfaces [2], studying human's reading behaviour in terms of their eye gaze is another field with much research effort.

Several algorithms exist to detect whether a user is reading or not based on their eye gaze. One such system is the "Pooled Evidence" system [1] which classifies a user's behaviour into either a scanning mode or a reading mode. An evidence threshold is used to determine how much evidence is required (in points) and different types of reading behaviours are given point values for how much evidence they contribute.

In [4], a thorough review of eye movements in reading and information processing has been conducted with a summary of three interesting examples of eye movement characteristics during reading, which have become important references regarding gaze parameters in reading:

1. When reading English, eye fixations last about 200-250 ms and the mean saccade size is 7-9 letter spaces.
2. Eye movements are influenced by textual and typographical variables, e.g., as text becomes conceptually more difficult, fixation duration increases and saccade length decreases. Factors such as the quality of print, line length, and letter spacing influence eye movements.
3. Eye movements differ somewhat when reading silently from reading aloud: mean fixation durations are longer when reading aloud or while listening to a voice reading the same text than in silent reading.

More recently, new methods based on advanced learning approaches have been proposed to be useful for studying gaze patterns in reading. In [8], a hybrid fuzzy approach for eye gaze pattern recognition has been introduced. This approach combines fuzzy signatures [3] with Levenberg-Marquardt optimization method for recognizing the different eye gaze patterns when a human is viewing faces or text documents. The experimental results show the effectiveness of using this method for the real world case. A further comparison with Support Vector Machines (SVM) also demonstrates that by defining the classification process

in a similar way to SVM, this hybrid approach is able to provide a comparable performance but with a more interpretable form of the learned structure.

Furthermore, a similar method has been introduced in [6] by which detecting the level of engagement in reading based on a person's gaze pattern becomes possible. Through their experimental results, they demonstrate the feasibility of the applying this approach in real-life systems.

3 The Experiment

In order to analyse different reading patterns an experiment was designed. The experiment involved reading a series of paragraphs and then answering some questions about those paragraphs.

3.1 Experiment Design

In all there were ten paragraphs for the participants to read. Seven of the paragraphs were taken from a selected paper [7]. The remaining three paragraphs were written by students who were required to write about the paper for course work. Five of the paragraphs from the paper were chosen for the amount of useful information that was contained within. The other two paragraphs from the paper and the three student paragraphs were chosen because of their generality and lack of useful information. Care was taken to make sure that this fact was not obvious.

The paragraphs were presented to different participants in different orders to prevent any specific paragraph ordering from affecting the results. The paragraphs all come from different places in the paper or from a completely different source altogether (the student's paragraphs). As well as being presented in different orders, the overall composition of the paragraphs became very convoluted. This was an experiment design choice to help show which participants could look at the bigger picture even when the information is out of place and scattered. The participants were given 90 seconds to read each paragraph.

After reading the ten paragraphs, the participants were asked to answer five multiple choice questions on the material. These questions asked about the content of the five paragraphs that contained the most relevant information. Furthermore, they were asked to write describe the paper in one sentence. Only one sentence was asked for, to not inundate the participant with a writing task. Then they were asked to rank the paragraphs from the one with the most useful information for completing the questions, as number one, and the one with the least information, as number ten. All the data were used to analyse how well they had understood the material that was presented to them. Then the utility of their reading patterns and characteristics could be assessed.

3.2 Experimental Setup

During the experiments the participants read all the paragraphs off a screen which was connected to the same computer that was recording their eye movements.

The computer was a standard desktop machine that was running Windows XP. The eye tracking system that was connected to the computer was provided by Seeingsmachines with FaceLab V4.5 software [5].

As shown in Fig. 1, the computer had two screens connected to it, one for controlling and monitoring the experiment and a 19 inch screen with a resolution of 1280 by 1024 for the participants to read the paragraphs and questions off. Before the experiment could begin, the system was calibrated for each participant. All the paragraphs and questions were set to the same resolution so no scaling was required. The entire system was housed on a cart that had a mounted chin rest to help the participants keep their head still. Although the chin rest helped to keep the participants head still there were still times when the gaze tracking system would lose its target, usually if the participant started to squint when reading the bottom of the screen (when tracking was lost no data points were recorded and so it can be identified where this happened and is taken into account in the analysis).



Fig. 1. The Setup for the Reading Experiment

3.3 Participants

Altogether 18 volunteers from a local university participated in the experiment, of them 3 were removed because of the poor results, i.e. the gaze tracker recorded only noise or nothing.

4 Analysis and Results

4.1 Gaze Points to Fixations

A person's gaze is characterised by two behaviours, fixations and saccades [2]. Fixations being the time when a person focuses on an object and they move that object into view of their fovea (the part of the eye with the most photosensitive cells). A saccade is the high-speed, ballistic movement of the eye when it is between fixations. It is reasonable to display everything in fixations (and saccades, although saccades are not really displayed because they are just the movement between the really meaningful data).

To break the gaze points into fixations an approximate method was used. As shown in Fig.2, the fixations are represented as circles that are centred at the average position of all the gaze points that are contained within them and their radius is determined by the length of time that the participant spent in that fixation. Thin lines are drawn between the fixations and could be considered saccades although, they are only there to show an observer which fixation comes next and they do not take into account any of the gaze points in the saccades. The gaze points in the saccades are essentially omitted. The same colouring scheme applies on the fixations as did on the gaze points, the colour gets lighter as time passes.

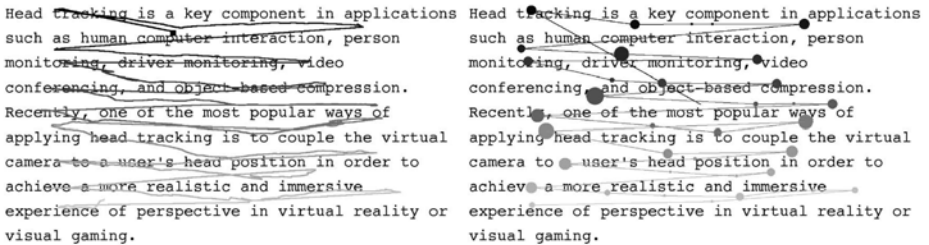


Fig. 2. Gaze Points / Lines (left) vs Fixations (right) Generated from the Collected Gaze Data

4.2 Scoring the Participants

The evaluation of the participants was a step that was inherent in the experiment. It was the purpose of asking the questions, and having the participants write a sentence and rank the paragraphs. The experiment was designed so that the participants could be scored using the following guidelines:

Paragraph Ranking: ten paragraphs and one point would be awarded for each paragraph that was correctly ranked in the correct half.

Multiple Choice: one point would be awarded to each correct answer.

Sentence Writing: a possible three points awarded for the sentence regarding whether participants have mentioned the key content of the paper.

The participant could have received a score of up to 18 points. These scores would allow the participants understanding to be quantified so that the ones that understood better could be identified. In the end, the highest scoring participant received a score of 16, the lowest scoring participant received a score of 4, the mean score was 9.6 with a standard deviation of 3.29.

4.3 Statistical Analysis

Before the statistical analysis, a few measurements were taken about the way that the participants read. These measurements were taken as averages across entire slides. The measurements that were taken were:

1. Time taken to read a slide.
2. Horizontal distance between fixations.
3. Vertical distance between fixations.
4. Number of gaze points per slide.
5. Number of fixations per slide.
6. Length of fixations.

These measurements were plotted against scores to try and find trends. There were some slight trends although none of them were statistically significant. It seems that simple statistical analysis did not show any real correlation between the simple measurements and the scores.

4.4 Further Analysis by ANN

To look into the merits of using more advanced analysis techniques on the data, a neural network was trained to determine whether a given paragraph was relevant or irrelevant. To do this only the data from the gaze patterns of the paragraphs was used. The neural network was back propagation trained and its inputs consisted of the measurements that were taken above except on the individual paragraphs. The neural network had six hidden nodes and one output, which was the class for that given paragraph, that the inputs corresponded to, was relevant or irrelevant. The neural network was trained with 60% of the data while 20% was used to generalise the network and prevent over-fitting and the last 20% was used as the test data.

The neural network produced good results (see Fig.3) with a correct classification rate of approximately 86% (assuming that there is no undecided class, so all points that are on the correct side of 0.5 are considered correct). Training this neural network was only an example of how learning algorithms can be used to analyse this data.

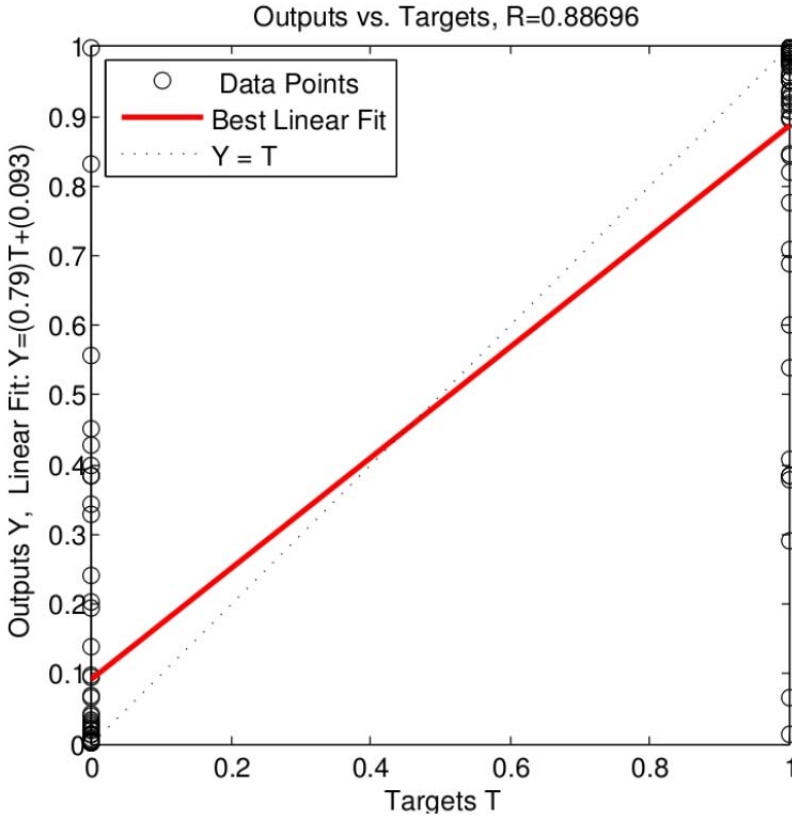


Fig. 3. A graph of the results for the neural network. The dots down the sides correspond to the paragraphs. The ones on the left are the irrelevant paragraphs and the ones on the right are the relevant paragraphs. Their location shows their class as one or the other according to this model. So, the irrelevant paragraphs should be near the bottom and the relevant ones should be near the top. The solid line that runs across the graph is the line of best fit between all the dots. The dotted line that runs across the graph is the ideal solution (where every paragraph is correctly classed).

5 Discussion

From the results, it shows that using classical statistical methods, we could hardly find any significant correlations between the measures we defined in terms of the gaze data and the scores of the participants in the reading experiment. However, good classification results were generated for discriminating between relevant and irrelevant paragraphs by training a simple artificial neural network with the same input data from the defined measures. This implies the potential advantages of using advanced learning approaches especially for analysing eye gaze patterns in reading. These approaches might be more useful in studying more detailed information within the gaze data than applying traditional methods, which also requires further investigations and comparisons.

Future studies could include using the same method to see if the learning algorithms could determine which questions a participant will get right or wrong, or perhaps even predict which ordering a participant will order their paragraphs in. But what would be much more useful for trying to quantify a participants understanding would be to train a learning algorithm on the values of the gaze points, or fixations themselves.

References

1. Compbell, C.S., Maglio, P.P.: A Rbust Algorithm for Reading Detection. In: 2001 Workshop on Perceptive User Interfaces, vol. 15, pp. 1–7. ACM (2001)
2. Jacob, R.J.K.: The Use of Eye Movements in Human-computer Interaction Techniques: What You Look at is What You Get. *ACM Transactions on Information Systems* 9(2), 152–169 (1991)
3. Koczy, L.T., Vamos, T., Biro, G.: Fuzzy Signatures. In: Proceedings of the 4th Meeting of the Euro Working Group on Fuzzy Sets and the 2nd International Conference on Soft and Intelligent Computing (EUROPUSE-SIC 1999), Budapest, Hungary, pp. 210–217 (1999)
4. Rayner, K.: Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin* 124(3), 372–422 (1998)
5. Seeingmachines, Inc: FaceLAB (2011), <http://www.seeingmachines.com/faceLAB.html>
6. Vo, T., Mendis, B.S.U., Gedeon, T.: Gaze Pattern and Reading Comprehension. In: Wong, K.W., Mendis, B.S.U., Bouzerdoum, A. (eds.) *ICONIP 2010 Part II*. LNCS, vol. 6444, pp. 124–131. Springer, Heidelberg (2010)
7. Zhu, D., Gedeon, T., Taylor, K.: Keyboard before Head Tracking Depresses User Success in Remote Camera Control. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) *INTERACT 2009*. LNCS, vol. 5727, pp. 319–331. Springer, Heidelberg (2009)
8. Zhu, D., Mendis, B.S.U., Gedeon, T., Asthana, A., Goecke, R.: A Hybrid Fuzzy Approach for Human Eye Gaze Pattern Recognition. In: Köppen, M., Kasabov, N., Coghill, G. (eds.) *ICONIP 2008*. LNCS, vol. 5507, pp. 655–662. Springer, Heidelberg (2009)