Reading Similarity Measure Based on Comparison of Fixation Sequences

Riki Kudo
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka , JAPAN
kudo@m.cs.osakafu-u.ac.jp

Koichi Kise
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka , JAPAN
kise@cs.osakafu-u.ac.jp

Olivier Augereau
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka , JAPAN
augereau.o@gmail.com

Takuto Rou
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka , JAPAN
rou@m.cs.osakafu-u.ac.jp

Ubicomp/ISWC'15 Adjunct , September 07-11, 2015, Osaka, Japan

DOI: http://dx.doi.org/10.1145/2800835.2807935

Abstract
The eye movement is an important source of information for the reading analysis. We propose a method for computing a similarity measure between two fixation sequences. In order to estimate the effectiveness of the similarity measure, we investigate whether a high similarity is obtained when two subjects read the same document. A $F_1 score$ of 0.92 is obtained for retrieving the same document based on the reading similarity.

Author Keywords
Reading similarity; Eye tracker

ACM Classification Keywords
H.5.2 [User Interfaces]: Input devices and strategies

Introduction
Nowadays, more and more research focus on sensing and recording our daily life. For example, the food-life log provides information to the users about the calories contained in foods they eat every day [1]. The activity-life log recognizes human activities such as walking, climbing stairs, *etc.* for health monitoring [8]. In this paper, we will focus on the reading-life log, *i.e.* the analysis of reading.

Everyday, we read some documents or texts. Reading takes a very important role in our life, it is one of the major ways to get knowledge. Some researches have been done in reading analysis. For example, there are studies about recognizing the document type from eye movement with a mobile eye tracker [7]. Another application consists of estimating the number of read words [6].

The reading contains information about the contents of read documents and the reader. For example, if the documents have different layouts or contain a different number of words, the reading will be different. Furthermore, the reading also contains information about the reader such as his reading skills, comprehension or attention.

In this paper, we propose a method for computing the similarity between two reading recordings. This method does not use the document image, but only the reader's eye gaze movements. The reading similarity measure can be used to find some relationships between the readers or the documents. The similarity of reading might be correlated to the similarity of reader's English skills or to the similarity of the content of the documents. In order to compute the similarity between two readings, we will compute the similarity of two sequences of eye gazes. Several methods based on the Levenstein distance have already been proposed [2], [5]. We chose to use the ScanMatch algorithm [4] based on Needleman-Wunsch as it offer better performances than the ones based on the Levenstein distance. To prove the effectiveness of the similarity measure, we show in the experiment that if two readers read the same document, a high similarity measure is obtained. We show that for the best parameters, a precision of 100% and a recall of 83% are obtained.

The rest of this paper is organized as follows. First, we describe how to measure the reading similarity. Next, we present the experimental results. The similarity measure can be employed to identify if two subjects read the same document. Finally, we conclude and propose some future work.

Measure of the reading similarity

In this section, we explain how to measure the reading similarity through the eye movements. While reading, the eye movement is decomposed as a sequence of fixations and saccades. A fixation is obtained when the eye stares at the same place for more than a certain period of time. A saccade corresponds to a quick movement between two fixations. We use the ScanMatch method[1], an algorithm for comparing fixation sequences.

The reading similarity is computed in two major steps. First, we record the eye movement positions by using an eye tracker and we extract the fixation information from the eye gaze. Next, we measure the reading similarity by using the ScanMatch method.

1) Fixation detection

The fixations and saccades are obtained by filtering the eye gaze [3]. The fixation detection is based on two following steps (illustrated in Fig. 1):

- First, when the eye gaze is gathered closely enough (within a 30-pixel square), we regard these points as a minimum fixation.
- If the following gaze is close enough (included in a 50-pixel square), it is added to the fixation. If it is too far, it is regarded as noise. If four consecutive gazes are noise, the fixation is stopped and a new fixation detection is started.

[1] http://seis.bris.ac.uk/~psidg/ScanMatch/

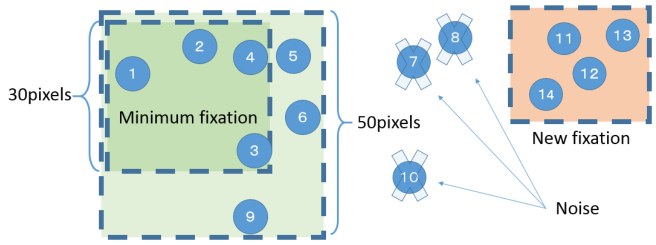

Figure 1: Fixation detection (based on Buscher *et al.* [3] method.)

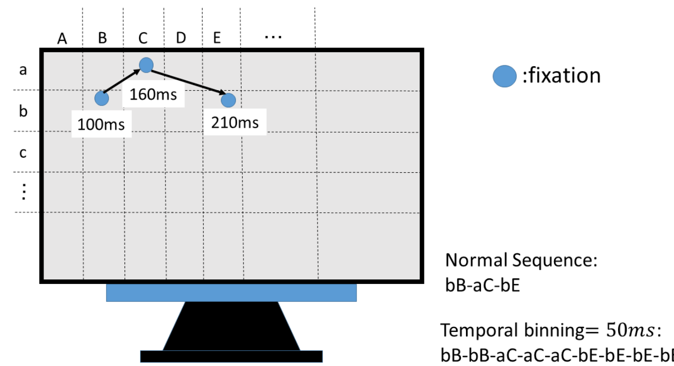

Figure 2: Creating the character sequence from the fixations (based on Cristino *et al.* [4]).

2)ScanMatch

In this section, we explain the ScanMatch algorithm. It is used for measuring the similarity between two fixation sequences. ScanMatch consists of two processes: creating a sequence of characters from the fixations, and then comparing two sequences of characters.

2.1)Creating a sequence of characters

First, the area of the monitor is divided in a grid and a character is associated to each cell of the grid. Then, the fixations are assigned to the corresponding cells, depending on their positions such as in Fig. 2. Next, for each fixation, we check which cell contains the fixation and add the corresponding character to the sequence.

A temporal binning is set up. If one fixation lasts more than a fixed threshold, it is represented as several times the same character. This process is also illustrated in the Fig. 2.

2.2)Global alignment of fixation sequences

The Needleman-Wunsch Algorithm [9] is used for determining the global alignment of two sequences. This algorithm expresses a score of similarity between two character sequences.

2.3)Normalizing the score

The score of two sequences is difficult to compare, because it depends on the length of the sequences. So the score is normalized by using the following equation:

$$N = \frac{S}{L \cdot M} \quad (1)$$

In this equation *N* is the normalized score, *S* is the score of two characters similarity, and *L* is the length of the longest sequence. *M* is the score of the two characters matching. It is a constant value fixed in Needleman-Wunsch Algorithm. The score is normalized between 0 and 1, where 1 is the highest similarity.

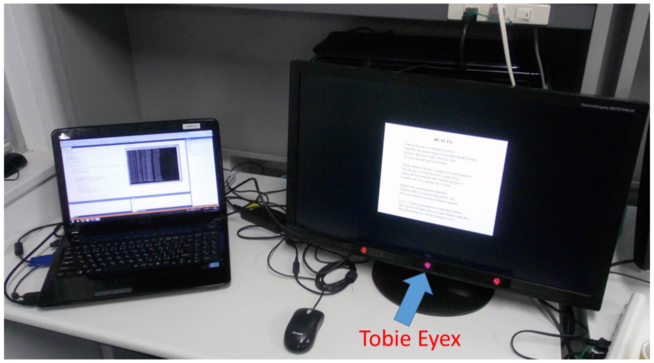

Figure 3: Experiment setup.

Experiment

The ScanMatch method was designed for visual tasks such as researching a symbol or a number of a defined color; but not specifically for a reading task. In order to prove that the reading similarity measure can be used to find if two readers read the same document, we examined how the parameters of the ScanMatch method could affect the reading similarity.

In this experiment, we asked 12 subjects to read 5 documents, being 60 reading recordings. By using the similarity measure, we wanted to show that a high reading similarity was obtained when different subjects read the same document. The Tobii EyeX[2] eye tracker is used for recording the eye gaze.

The eye tracker was attached under the monitor such as in Fig. 3. The experiment is carried out in the following process:

1. We explain the experiment to a subject.
2. The eye tracker is calibrated for the subject.
3. The subject reads the document.

The subject had been asked to read each document from the beginning to the end without rereading. This constraint had been added in order to reduce the differences between the subjects because they come from different countries and have different English skills.

For each of the 60 recordings, we compared the combinations with all the other recordings by computing the normalized score based on the ScanMatch method. A total of 1770 combinations is obtained. If the normalized score is higher than the threshold T, we considered that the two subjects read the same document. Then, this decision is compared with the ground truth in order to compute the performances of the similarity measure. We analyze the recall and the precision while changing the following 3 parameters:

1. The threshold *T*.
   A range of values from 0.3 to 0.7 is tested in order to obtain a large range of a recall and a precision.
2. The temporal binning.
   We fixed the temporal binning at 50 ms. We also tested not to apply any temporal binning.
3. The grid.
   The grid is defined as the number of columns times the number of lines. We tested the 3 different following grids: 24×12, 18×9 and 12×6.

[2]http://www.tobii.com/ja-JP/eye-experience/eyex/

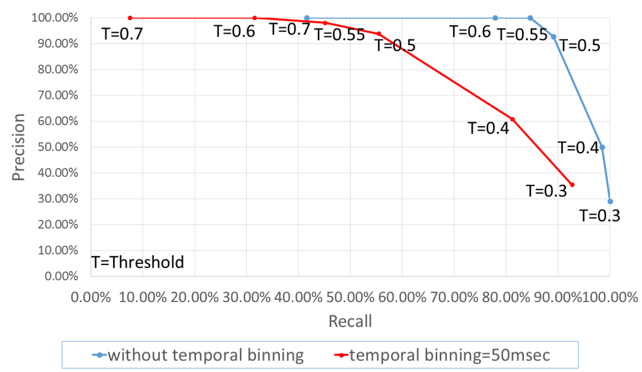

Figure 4: Influence of temporal binning.

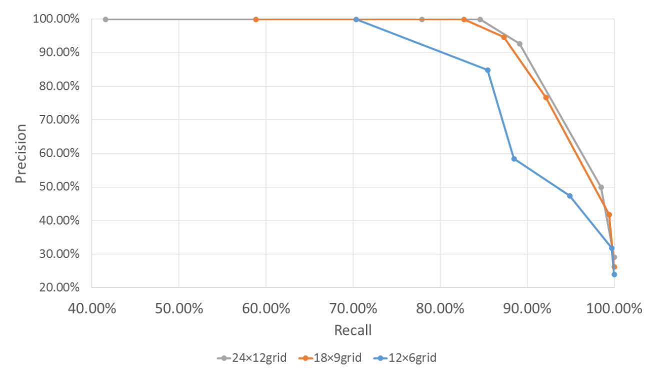

Figure 5: Influence of the grid.

Experiment results
We show the experiment results about the threshold *T* and the temporal binning in Fig. 4. For the threshold, the precision sharply decreases when it is less than 0.5 because the similarity measure of most of the same documents is more than 0.5. We also observe that the performances are better without using the temporal binning. In the reading context, the differences of English reading skills between the subjects have an impact on the duration of the fixations. So, if we use a temporal binning, when two subjects with different English skills read the same document, the similarity will be lower.

Next, we show the impact of the grid on the performances in Fig. 5. This figure shows the performances of the 24×12 and 18×9 grid are similar, but the 12×6 grid has a lower precision. If the cell size is too large, many fixations are merged in the same cell, which causes a low precision.

In order to obtain the best $F_1 score$, the optimal parameters are the following: T=0.55, a 24×12 grid and no temporal binning. With this parameters, the precision is 100%, the recall is 85% and the $F_1 score$ is 0.92. In other words, if the similarity measure is higher than 0.55, the two recordings come from the same document. And, if the recordings came from the same document, the similarity is higher than 0.55 85% of the times.

Conclusion and Future Work

We explained how to compute the reading similarity between two fixation sequences. Our experiment shows that 3 parameters have a great impact on the reading similarity. As a result, we show that the similarity measure can be used for recognizing if two readers read the same document with 100% precision and 85% recall.

We proved that the reading similarity measure proposed in this paper can be used to find some relationships between the documents. One of the next steps will be to use the reading similarity measure to find relationships between the readers.

We fixed the constraints of reading all the text from the beginning to the end without rereading. This is just a first step towards computing a reading similarity. The next step will be to analyze the reading similarity on recordings with no constraint. However, in this case, some readers might skip or reread different parts of the document. So, a global alignment is not suitable anymore, the algorithm must be local in order to align only some parts of the reading sequences.

REFERENCES

1. Kiyoharu Aizawa, Gamhewage C De Silva, Makoto Ogawa, and Yohei Sato. 2010. Food log by snapping and processing images. In *Virtual Systems and Multimedia (VSMM), 2010 16th International Conference on*. IEEE, 71–74.

2. Stephan A Brandt and Lawrence W Stark. 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of cognitive neuroscience* 9, 1 (1997), 27–38.

3. Georg Buscher, Andreas Dengel, and Ludger van Elst. 2008. Eye movements as implicit relevance feedback. In *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2991–2996.

4. Filipe Cristino, Sebastiaan Mathôt, Jan Theeuwes, and Iain D Gilchrist. 2010. ScanMatch: A novel method for comparing fixation sequences. *Behavior research methods* 42, 3 (2010), 692–700.

5. Tom Foulsham and Geoffrey Underwood. 2008. What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision* 8, 2 (2008), 6.

6. Kai Kunze, Hitoshi Kawaichi, Kazuki Yoshimura, and Kenji Kise. 2013a. The wordometer–estimating the number of words read using document image retrieval and mobile eye tracking. In *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 25–29.

7. Kai Kunze, Yuzuko Utsumi, Yuki Shiga, Koichi Kise, and Andreas Bulling. 2013b. I know what you are reading: recognition of document types using mobile eye tracking. In *Proceedings of the 2013 International Symposium on Wearable Computers*. ACM, 113–116.

8. Myong-Woo Lee, Adil Mehmood Khan, and Tae-Seong Kim. 2011. A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation. *Personal and Ubiquitous Computing* 15, 8 (2011), 887–898.

9. W John Wilbur and David J Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proceedings of the National Academy of Sciences* 80, 3 (1983), 726–730.