Proceedings of the Int'l. Conf. on Forum Forum for Interdisciplinary Mathematics
[FIM2013-920-00319]

Extraction of Writing Carefulness from Online Handwritten Data

Motoki Miura†, Takamichi Toda††

†Faculty of Basic Sciences, Kyushu Institute of Technology,
Sensui, Tobata, Kitakyushu 804-8550, Japan
email: miuramo@mns.kyutech.ac.jp

††Graduate School of Engineering, Kyushu Institute of Technology
Sensui, Tobata, Kitakyushu 804-8550, Japan
email: n350917t@tobata.isc.kyutech.ac.jp

Abstract. In order to improve the effectiveness of sharing notes on peer learning, students need to improve the presentation of their handwritten notes, because carefully written notes are more accessible for other students and reduce irrelevant, careless mistakes. To facilitate improvements in note-taking, we considered a simple metric that determine how carefully the notes were written, and checked the metric by using students note taken at a real lecture. The result showed some correlations among the metric and the writing time.

Keywords: Consciousness, Neatness, Improved Handwriting, Attitudes, Anoto digital pen

1 Introduction Learning by teaching [1] is one of the primary strategies for effective learning. Bielaczyc et al. examined the impact of self-explanation and self-regulation strategies on student explanations and performance [2]. The results indicated that particular self-explanation and self-regulation strategies contributed to learning and problem-solving performance. Barnard reported peer-tutoring interactions and their interpretation from a socio-cultural perspective [3]. Therefore, attitudes and strategies for explaining learning content are necessary, and they can be improved by efforts to improve the way explanations are made.

In our previous study, we developed AirTransNote, a student note-sharing system that facilitates collaborative and interactive learning during regular lectures in conventional classrooms [4]. Because taking notes on paper is a regular activity, our system does not impose an extra burden on students who share notes. However, in order to improve the effectiveness of sharing notes in peer learning, students need to improve the presentation of their handwritten notes, because carefully written notes are more accessible for other students and reduce irrelevant, careless mistakes.

In this study, we consider how to determine from online stroke data the level of care that a student takes to write notes. Our target is to examine the writing activity of students during lectures—not the accuracy of the content of their notes compared to the teacher's lecture. Teachers usually check whether students understand the lecture content by asking questions soon after a topic is introduced, and students are expected to answer within a sufficient time. Therefore, our evaluation of the carefulness of the handwriting is independent of factors, such as the speed at which the teacher delivers the lecture or how the lecture is structured.

2 Related Works Simard et al. [5] proposed a warping algorithm for ink normalization and beautification. They concentrated on the preprocessing of the recognition of handwritten text; therefore, their final goal was to reduce recognition errors. The concept of ink normalization could be applied to our research in terms of presenting beautified notes, but instead we focused on giving feedback based on metrics of carefulness.

Julia and Faure [6] presented an algorithm of recognition and beautification for graphical design applications on a pen-based computer. Their method recognizes tables, gestures, geometric figures, or diagram networks, and it beautifies the drawings for each drawing category. Miyao and Maruyama [7] proposed a method to segment and recognize online handwritten flowchart symbols by SVM technique. They also proved the effectiveness of their method and implemented a system that beautifies handwritten flowcharts. Paulson and Hammond also proposed a new low-level recognition and beautification system called PaleoSketch [8] that can recognize eight primitive shapes as well as combinations of these primitives. The concepts of interactivity in handwritten drawings and demand for beautification are commonly researched; however, our goal is to provide a method of diagnosis that finds metrics of carefulness.

Zhu and Jin [9] proposed a method for beautifying online handwritten Chinese-character calligraphy. They first applied a speed-based calligraphy simulation to produce a paint-brush style stroke. Afterward, the method matched strokes with template characters. Part of the transfiguration technique in their method can be applied to beautify our students' notes. However, our aim is to make the students improve their attitude about writing carefully while thinking.

Aşıcıoğlu and Turan examined the quality of the handwriting of subjects under the influence of alcohol [10]. The aim of the research was to learn how alcohol and alcohol-related neurological deterioration affected handwriting. The results revealed that the handwriting parameters, such as the length of words, the height of uppercase and lowercase letters, the height of ascending letters, the height of descending letters, the spacing between words, the amount of angularity, the amount of tremor, and the number of tapered ends, were all significantly increased under the effect of alcohol. Some of their metrics regarding handwriting are attractive for examining quality, but most of their metrics were evaluated by human examiners.

3 Proposed Metric In this section, we describe our proposed metric that representing how carefully students write their notes, which we call the level of carefulness.

3.1 Presupposition We gathered online data of handwritten notes to assess the level of carefulness of note-taking. The online data could be captured by tablet or smartphones, but we employed Anoto-based digital pens in this study. The Anoto-based digital pen has the capability to store and send handwritten notes written on a specific dotted paper sheet. Using the Anoto-based digital pens, we collected accurate and stable student notes.

The Anoto-based digital pen generates (1) the coordinates of the pen-tip (x,y) in a frequency of 75 times per second, and (2) the start time of the writing. Although the end time of the drawing cannot be captured, it can be estimated using the start time and the number of coordinates that represent a drawing. Therefore, a one-stroke drawing contains n coordinates $P_i(x_i, y_i)$ $(0 \leq i \leq n-1)$ and has a start time T_0 in milliseconds.

3.2 Hypotheses We made some assumptions for estimating the carefulness of handwritten letters using variance of pen speed and stroke complexity[11]. However, the metric was not straightforward as well as low stability. Therefore, we try to estimate the level of carefulness using fundamental metrics obtained from the handwritten data. We considered the following points:

Captured Points per Distance simply represents carefulness. If the student wrote quickly, the number of captured points decreases since the sampling rate is same. Therefore, the simple metric (captured points per stroke distance) may describe the students' mood and attitude.

Complexity of the stroke may affect the carefulness. If the stroke contains many angular points, the above metric will be biased. Since such curvy stroke require the students to care, the speed of pen should decrease.

3.3 Calculation of stroke complexity To estimate the level of stroke complexity, we calculated feature points using Ramer's method [12]. The feature points of Ramer is often utilized for handwritten recognition, that reduce the original points and pick up the significant points. We utilize the number of feature points as a metric of stroke complexity.

The following is the Ramer's algorithm. First, the start and end points of every stroke were captured as feature points (**Figure 1**, top-left). Then, the most distant point from the straight

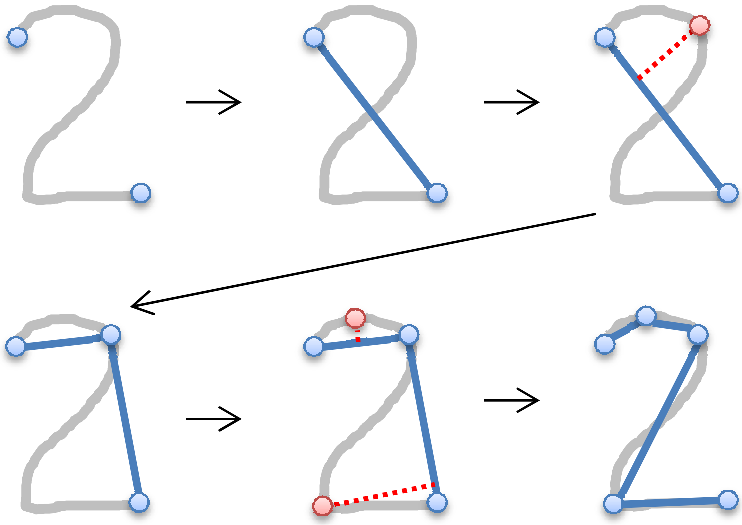

Figure 1: Ramer's method

line between adjacent feature points was selected as a feature point if the distance to the straight line was greater than a threshold value (**Figure 1**, top-right). This selection was done recursively until no more feature points were selected.

We set the threshold value of Ramer's method as one fifth of the stroke height or width, which is larger than others. The number of feature points found using Ramer's method represents the ratio of curves to angular points, and it is somewhat larger than the number of the angular points we defined.

4 Experiment In this section, we describe how we collected and examined the data.

4.1 Collection of Data We collected handwritten note data of 10 undergraduate students on a lecture. The lecture involved peer-reviewing, and short comments written by the students were digitized by Anoto digital pens. Each students were expected to write comments of other 13 reports. All students read these reports, and some of the students prepared drafts of the short comments. Since the prepared drafts were stored in their personal smartphones, the students refined the draft while writing to a paper sheet. We provided two A4 size sheet. The sheet specified areas of each comments. The size of area was 10×5 cm. Though the size was limited, we did not force the size of the handwriting character nor length of the comment. We explained that the comments will be shared by projection, but we did not instructed about controlling of carefulness.

After the lecture, we classified the comments into 5 levels by considering the time stamp for each student. **Figure 2** show the result of the classification. Each color represents neat5 (firstly written) to neat1 (lastly written). Note that the lecture 1 (Top-left) wrote only 3 comments, we set neat5, neat3, and neat1. Since the students wrote these comments by their favorite order, the position of the comment on the sheet does not indicate the time stamp. However, four students (ex. Student 6) wrote the position order.

We expect that the "neat5" wrote carefully because the students could pay attention to the first writings. Similarly, neat1 would be the hastiest writings because of their fatigue.

Consequently we gathered 48 comments (10 persons $\times$ 5 levels - 2). The data contains *Ramer* (ratio of curves), *point* (number of sample points) and *dist* (length) for each stroke.

4.2 Result Firstly we consider the *Ramer* as a metrics of classifying *point*/*dist* by graph visualization. After that, we discuss the *point*/*dist* for each participants.

Figure 2: We classified the comments into 5 levels (neat5–neat1) by considering the time stamp.

4.2.1 Point/Dist for each Ramer **Figure 3** shows the result of Point/Dist metrics. The top-left shows Ramer 0, and the top-right shows Ramer 1, and the bottom-left shows Ramer 2, and the bottom-right shows Ramer 3, respectively. The value of Ramer is obtained by subtracting 2 from the number of original feature points (more than 1).

From the graph, we can see that the Point/Dist metric somewhat explains the neatness (or fatigue) when the Ramer was small. Therefore, the Point/Dist metric can be applied for simpler stroke only.

4.2.2 Differences among participants **Figure 4** shows the transitions of Point/Dist metrics for each participants. The x-axis represents neat5 to neat1 (less fatigue to more fatigue). Most of the regression lines shows the relation of Point/Dist metric and the fatigue.

5 Conclusion and Future Work We considered the relationships between Point/Dist of handwritten stroke and a fatigue. The results supported our first hypothesis that the Point/Dist tends to increase when the strokes are written carefully under the condition of simple stroke.

In this study, we assumed that the level of carefulness is affected by the time of writing, caused by the students' fatigue. However, the relationship between the time and fatigue was not clear. We will ensure the condition of fatigue and carefulness in our future research. Also, we will explore a method that considers the shape and beautification of the writing.

Acknowledgment Our research is partly supported by Japan Society for the Promotion of Science(JSPS) KAKENHI Grant-in-Aid for Young Scientists (A): Grant Number 23680078.

References

[1] Alan Gartner, Mary Conway Kohler, and Frank Riessman. *Children teach children: Learning by teaching*. Harper & Row, 1971.

[2] Katerine Bielaczyc, Peter L. Pirolli, and Ann L. Brown. Training in Self-Explanation and Self-Regulation Strategies: Investigating the Effects of Knowledge Acquisition Activities on Problem Solving. *Cognition and Instruction*, 13(2):221–252, 1995.

[3] Roger Barnard. Peer tutoring in the Primary Classroom: A Sociocultural Interpretation of Classroom Interaction. *New Zealand Journal of Educational Studies*, 37(1):57–72, 2002.

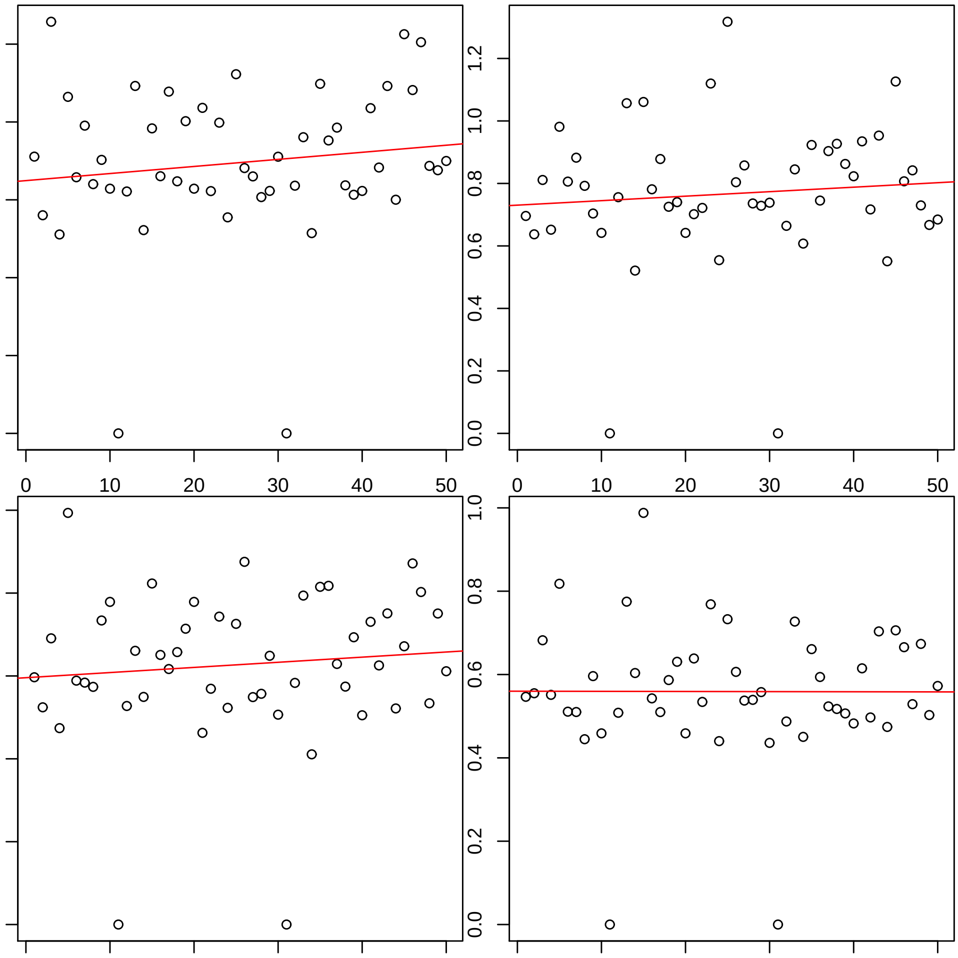

Figure 3: Point/Dist for each Ramer. The x-axis represents neat×10+StudentID.

[4] Motoki Miura, Taro Sugihara, and Susumu Kunifuji. Improvement of Digital Pen Learning System for Daily Use in Classrooms. *Educational Technology Research*, 34:49–57, October 2011.

[5] Patrice Y. Simard, Dave Steinkraus, and Maneesh Agrawala. Ink Normalization and Beautification. In *Proc. of 2005 Int. Conf. on Document Analysis and Recognition (ICDAR'05)*, pages 1182–1187, August 2005.

[6] Luc Julia and Claudie Faure. Pattern Recognition and Beautification for a Pen Based Interface. In *Proc. of the 3rd Int. Conf. on Document Analysis and Recognition*, volume 1, pages 58–63, August 1995.

[7] Hidetoshi Miyao and Rei Maruyama. On-Line Handwritten Flowchart Recognition, Beautification and Editing System. In *Int. Conf. on Frontiers in Handwriting Recognition (ICFHR)*, pages 83–88, 2012.

[8] Brandon Paulson and Tracy Hammond. PaleoSketch: Accurate Primitive Sketch Recognition and Beautification. In *Proc. of the 13th Int. Conf. on Intelligent user interfaces*, IUI '08, pages 1–10, New York, NY, USA, 2008. ACM.

[9] Xinghua Zhu and Lianwen Jin. Calligraphic Beautification of Handwritten Chinese Characters: A Patternized Approach to Handwriting Transfiguration. In *Proc. of ICFHR2008*, pages 135–140, 2008.

[10] Faruk Aşıcıoğlu and Nurten Turan. Handwriting changes under the effect of alcohol. *Forensic Science International*, 132(3):201 – 210, 2003.

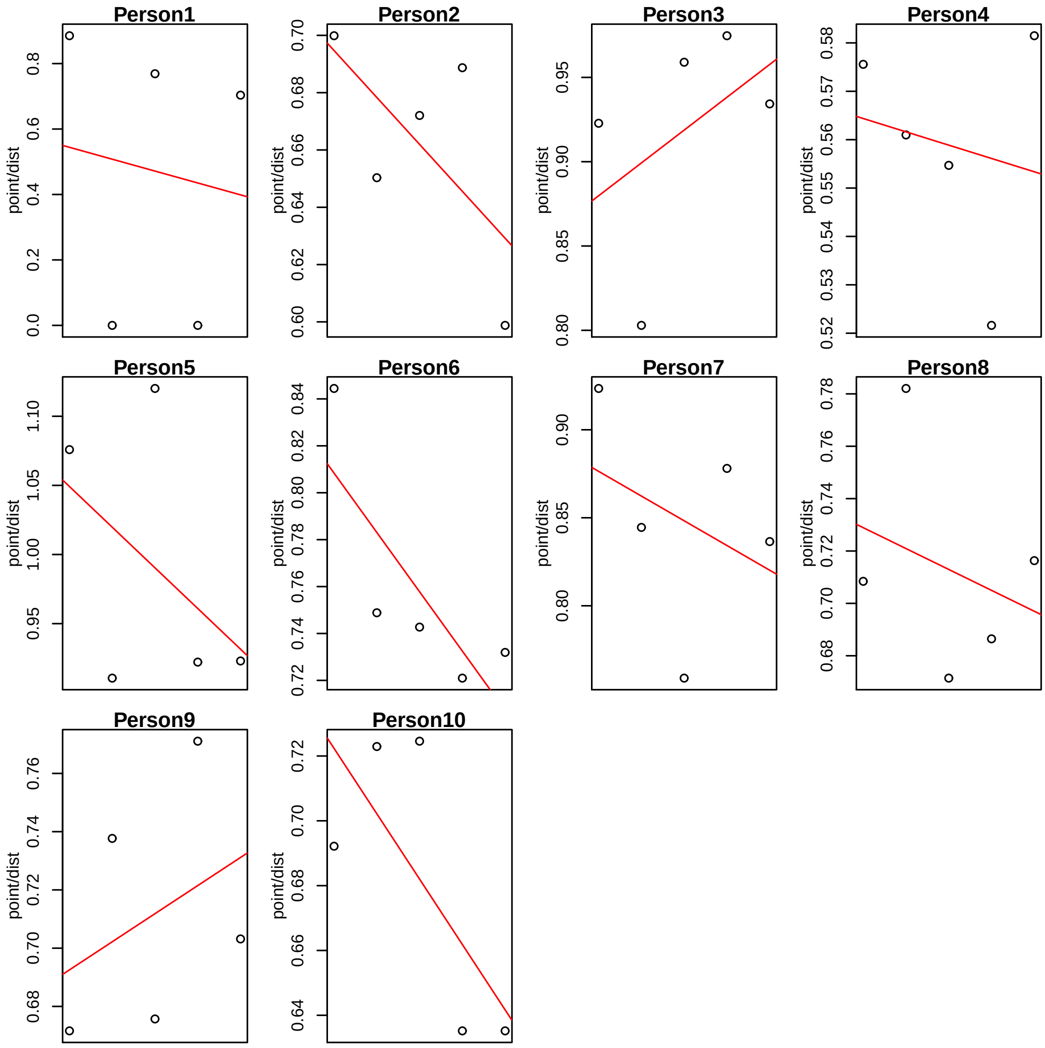

Figure 4: Metrics transition for each participants. (The x-axis represents neat5 to neat1)

[11] Motoki Miura, Shun Akase, and Takamichi Toda. Estimating the care with which notes are written from online handwritten character data. In *Proceedings of the 6th International Conference on Intelligent Interactive Multimedia Systems and Services*, pages 195–203, June 2013.

[12] Urs Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1(3):244–256, 1972.