Paper: jc*--**-****

Estimating Writing Neatness from Online Handwritten Data

Motoki Miura and Takamichi Toda

Kyushu Institute of Technology, 1-1 Sensui, Tobata, Kitakyushu, Fukuoka, 804-8550 Japan
E-mail: miuramo@mns.kyutech.ac.jp
[Received 00/00/00; accepted 00/00/00]

Handwriting is the most fundamental expressive activity in learning. In order to utilize the nature, digital pen technology has been emerged to capture notes and transfer them. We have developed AirTransNote, a student note-sharing system that facilitates collaborative and interactive learning in conventional classrooms. With AirTransNote system, a teacher can immediately share student notes to the class by a projected screen to enhance a group learning. However, in order to improve the effectiveness of sharing notes, the teacher need to pick up an effective note for sharing. But selecting the suitable note during a lecture is difficult because it consumes much time. Also the students should change their mind to improve the presentation of their handwritten notes, because carefully written notes are more accessible for other students and they reduce irrelevant, careless mistakes.

To facilitate improvements of the learning based on note sharing, we need a method to estimate the neatness of the note automatically. If the method is established, the teacher can easily select effective notes. It can also be applied to feedback the student to write better. We examined 14 basic features from handwritten notes by considering correlation coefficients, and found that the variance of pen speed, average of angular point, and average of pen speed were the significant features for evaluating neatness of the handwritten notes.

Keywords: Consciousness, Carefulness, Improved Handwriting, Learning Attitudes, Anoto digital pen

1. Introduction

Handwriting is the most fundamental expressive activity in learning. Since the handwriting activity is common in a modern culture, handwriting skills have been trained from childhood in both drawings and writings. Similarly, the handwriting skill is sometimes regarded as a representation of the personality. A Japanese proverb says that a beautiful hand-drawing represents a person's character. Therefore, acquiring the better writing skill is important for many people.

Based on the intuitiveness and the nature of the handwriting, a digital pen technology has been emerged to capture notes and transfer them as a digitized data. Especially, using an Anoto-based pen, the handwritten note written on a paper can be immediately digitized without any burdens. By utilizing the features, we have implemented AirTransNote, a student note-sharing system that facilitates collaborative and interactive learning in conventional classrooms[7]. With AirTransNote system, a teacher can immediately share student notes to the class by a projected screen to enhance a group learning. Thanks to the simple writing interface, our system does not impose an extra burden on students who share notes.

However, in order to improve the effectiveness of sharing notes, the teacher need to pick up an effective note for sharing. We have been developing a function to classify notes by the correctness of the answer[7]. But selecting the suitable note during a lecture is still difficult because it consumes much time due to the large number of student notes. Also the students should change their mind to improve the presentation of their handwritten notes, because carefully written notes are more accessible for other students and they reduce irrelevant, careless mistakes.

To facilitate improvements of the learning based on note sharing, we need a method to estimate the neatness of the note automatically. If the method is established, the teacher can easily select effective notes for sharing. For the students, they are expected to write better by hand as well as write accurate content; nevertheless, they do not share their notes. However, the habit of writing well is not always regarded as an important skill during most conventional lectures, except for calligraphy class. When the method to estimate the neatness correctly is realized, it can be applied to make the habit a common practice.

Regarding the sharing notes, learning by teaching [4] 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 [3]. 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 [2]. 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 this paper, we investigate the method of estimating neatness from online stroke data. 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 neatness of the handwriting is independent of factors, such as the speed at which the teacher delivers the lecture or how the lecture is structured.

We focused on the neatness with which the students wrote their notes, not how beautifully the characters were written. The beauty of the character writing can also be somewhat improved by writing carefully, and it should also be improved for better presentation of the student's notes. However, beautification depends on the student's motor skills, which are generally difficult to improve in the short term. Therefore, in this study, we focused on the care with which students wrote their notes.

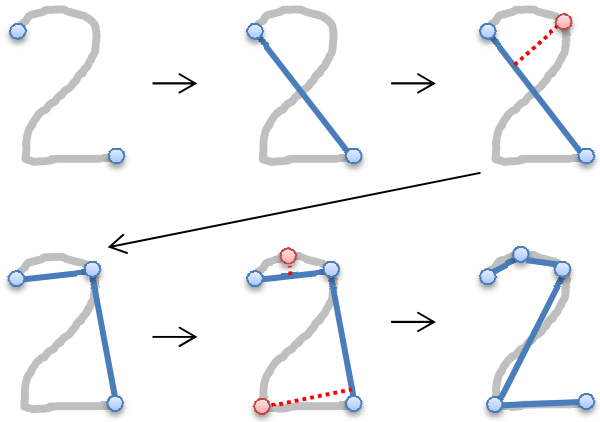

Fig. 1. Ramer-Douglas-Peucker algorithm

2. Related Works

Simard et al. [12] 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 neatness.

Julia and Faure [5] 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 [9] 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 [10] 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 neatness.

Zhu and Jin [13] 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 [1]. 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. Method

In this section, we describe our candidate features for measuring how carefully students write their notes, which we call the level of neatness. To process huge amounts of handwritten data, we needed to build a simple model of writing activity.

3.1. Presupposition

We gathered online data of handwritten notes to assess the level of neatness 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 is not 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.

Table 1. Expected relationships between character complexity and fundamental handwriting features

Metric	Complicated Characters	Simple Characters
Variance of Pen Speed	Large	Small
Average of Pen Speed	Small	Large

3.2. Basic features

Based on the coordinates, we can define distance (*dist*) and velocity (*velo*) between two coordinates P_i and P_{i-1} as follows:

$$dist_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (1 \leq i \leq n-1)$$

$$velo_i = \frac{Dist_i}{1/75} \quad (pixels/sec) \quad (1 \leq i \leq n-1)$$

We made some assumptions for estimating the neatness of handwritten letters using stroke data. We first tried to estimate the level of neatness using fundamental features obtained from the handwritten data. The fundamental features we considered were the following.

- **Variance of pen speed:** This feature is calculated by $velo_i$ of a single stroke.
- **Average of pen speed:** This feature is also calculated by $velo_i$ of a single stroke.
- **Complexity of the stroke:** This method counts the number of angular points and feature points extracted by Ramer's method. This feature is also calculated by a single stroke.

Table 1 shows our expectations of the relationship between the complexity of characters and the previously mentioned fundamental metrics. When the writer draws a complicated stroke, the variance of the pen speed will be greater than when writing a simple stroke. In addition, the average pen speed becomes slower than the average for simple strokes.

3.3. Calculation of stroke complexity (number of angular points)

To estimate the level of stroke complexity, we calculated feature points using Ramer-Douglas-Peucker algorithm [11] (hereafter, we call Ramer's method). The feature points of the algorithm 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 Ramer-Douglas-Peucker algorithm is calculated as follows. 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 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 the 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 and number of angular points.

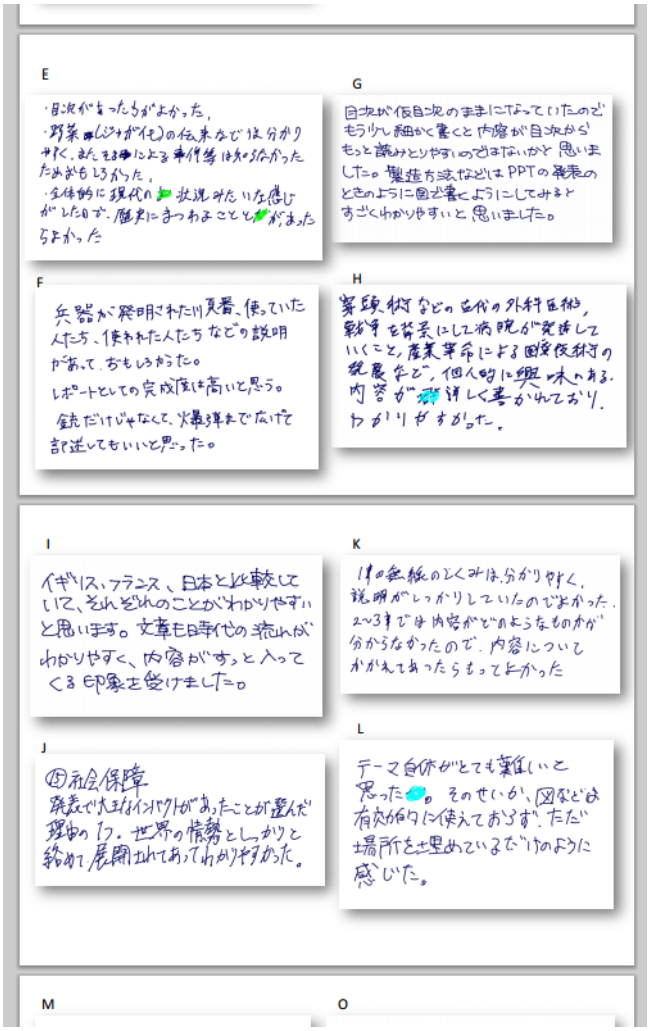

Fig. 3. PDF file for scoring (page 2 and 3 of 9)

By observing the Ramer feature with Japanese handwritten text, we found that the Ramer feature of Kanji characters are smaller than that of Hiragana characters. The reason is that the most of Kanji characters consist of simple short strokes relative to the Hiragana characters.

4. Experiment

In this section, we explain how we collected and examined the data.

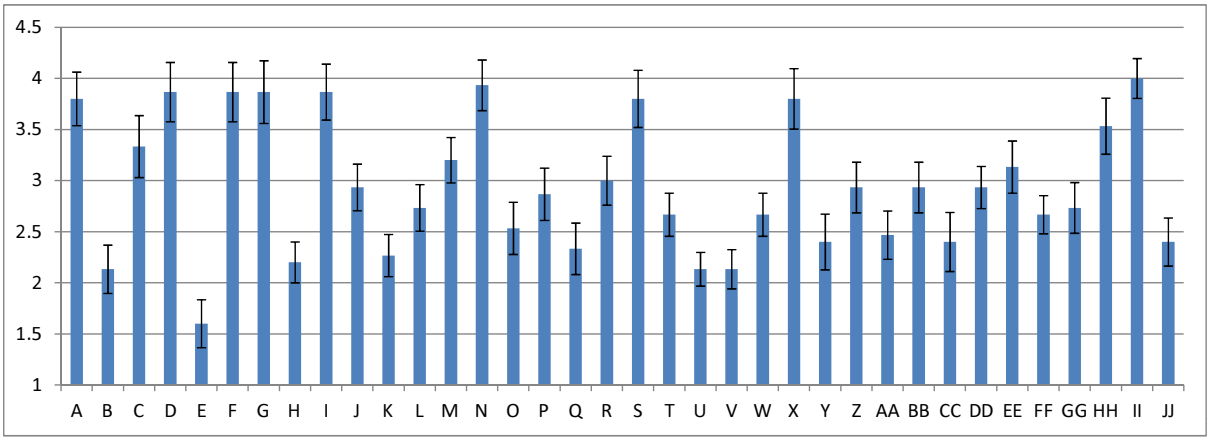

Fig. 2. Neatness (Average of 15 participants' ratings. The error bars represent standard errors.)

	avg(Point)	avg(Dist)	avg(Ramer)	var(Ramer)	avg(Vsplog)	var(Vsplog)	avg(Ppd)
r	-0.133	0.002	-0.449	-0.244	0.179	-0.007	-0.440
t(34)	-0.781	0.011	-2.929	-1.466	1.063	-0.041	-2.853
p	0.440	0.991	0.006	0.152	0.295	0.967	0.007
rank	12	14	2	8	11	13	3
	var(Ppd)	var(PpdR0)	var(PpdR1)	var(PpdR2)	var(PpdR3)	var(PpdR4)	var(PpdR5)
r	-0.398	-0.286	-0.556	-0.354	-0.248	-0.207	-0.212
t(34)	-2.531	-1.738	-3.899	-2.209	-1.490	-1.231	-1.267
p	0.016	0.091	0.0004	0.034	0.145	0.227	0.214
rank	4	6	1	5	7	10	9

Table 2. Pearson's correlation coefficients (r) between the stroke features and the reference neatness rate.

4.1. Collection of Note Data

We collected handwritten note data of 10 undergraduate students in a lecture of our institute. The lecture was a Japanese literature and expression course. Firstly, the students were required to write their report for their own topic through investigation. The reports were shared among the students as well as the lecturer. The students were also set an assignment of reading the reports and peer-reviewing by writing short comments to each report. The short comments written by the students were digitized by Anoto digital pens. Some of the students prepared drafts of the short comments. Since the prepared drafts were stored in their personal smartphones, the students could refine the draft when they wrote to a paper sheet. We provided two A4 size paper sheet. The areas of each comment were defined and printed on the sheet. The size of the area was 10cm wide and 5cm height. We did not force the size of the handwriting character, nor length of the comment. We explained that the comments will be shared by projecting screen and the lecturer introduces the projected short comments. Regarding the neatness of their writing, we did not instruct about controlling of carefulness.

After the lecture, we grouped the digitized writing data by the short comments. In order to assess the possibility of detecting the neatness level from online stroke data and to estimate the relevance of the method, we had to collect the reference score of neatness for the digitized short comments by considering the feeling of humanity. However, we could not simply compare and score the all 116 comments. Because the number of the short comment was too much to judge one by one. Also the amount of text included in the comment was varied for each comment. The difference of amount will affect the judgment of the neatness score.

Therefore, we divided the short comments to nine groups so that the amounts of the text were almost same. After that, we eliminated the comments of the same writer. Finally, we picked up 9 groups of 4 each comment. We prepared a PDF document which contained these comment groups per page (see **Figure 3**).

We asked another 15 participants (12 male and 3 female, ages from 20 to 39) to evaluate the neatness of the writings as absolute scores at 5-point Likert scale (hasty: 1 – neat: 5). The participants browsed the 9-page PDF document on their personal PC, and inputted their scores in the form of Web pages. No time limit was specified. We instructed the participants to glance over the document, and evaluate the neatness of the writings without considering the written content itself; just see the look feel of the writings. We supplemented that the writer's consciousness level of presentation and beautification as

an alternative point of view. We also informed the difference of writing amount, and the participants should not compare writings on the different pages directly. We also noticed that the green and light blue highlighted lines (see **Figure 3**) meant scratching out on unexpected or failure characters. Since the Anoto-based digital pen could not delete these characters, the student was allowed to relieve the mistakes. We also instructed that the highlighted lines should not be considered. **Figure 2** shows the result of the evaluated neatness scores. Thanks to the grouping and limitation of 4 writings at the same time, the variance of the score was moderate. Therefore, we adopted the average scores as a reference ratings of the neatness.

4.2. Candidate Features of Handwriting Data

We extracted the following candidate features for evaluating neatness level. All features can be generated by a single stroke.

- **Point :** The number of sampling point in the stroke.
- **Dist :** The stroke length. $\sum_{i=1}^{n-1} dist_i$ (unit: pixel).
- **Ramer :** The number of angular point calculated by Ramer's method explained in the section 3.3. For example, a short straight line produces zero and 'Z'-shaped stroke produces two.
- **Vsplog :** Variance of pen speed ($dist$) per log(**Ramer**+2). The value was introduced in our former study [6]. The feature was significant under controlled situation, but it is not clear the feature is effective in uncontrolled notes. The value increased when the writer changes the speed within the single stroke.
- **Ppd :** Point per distance (**Point/Dist**). The value represents an inverse of the average pen speed. The feature was introduced in our former study [8].
- **PpdR** n **:** Point per distance of Ramer = n stroke.

In order to compute the features of the multiple handwritten text strokes, we calculated the average and variance of the above features.

4.3. Result

Table 2 shows the result of Pearson's Correlation Coefficients (r) between the 14 stroke features and the reference neatness rate. The most significant feature was the variance of point per distance on Ramer = 1 ($r = -.556, t(34) = -3.899, p = .0004$). **Figure 4** shows the scatter plot with a regression line. The second significant feature was the average of Ramer ($r = -.449, t(34) = -2.929, p = .006$). The third significant feature was the average of point per distance ($r = -.440, t(34) = -2.853, p = .007$). **Figure 5** and **Figure 6** also show the scatter plot and a regression line of the two features. The following two features (the variance of point per distance

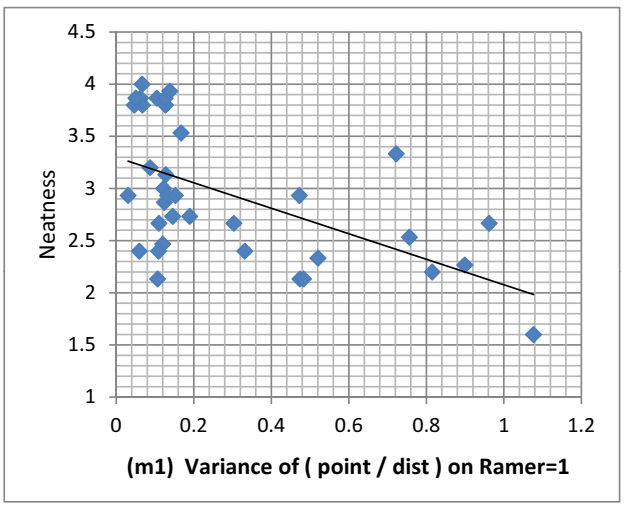

Fig. 4. (m1) Variance of (point/dist) on Ramer=1, $r = -.556$

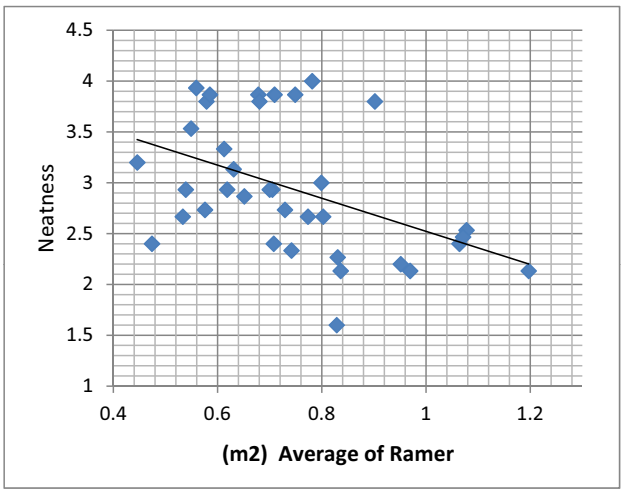

Fig. 5. (m2) Average of Ramer, $r = -.449$

and the variance of point per distance in Ramer = 2) were also 5% significant.

Based on the above results, we tried to evaluate two kinds of mixed features. The first one (mf_1) is a simple average:

$$mf_1 = \frac{m1 + m2 + m3}{3}$$

and the second one (mf_2) is a weighted average.

$$mf_2 = \frac{m1 \times 3 + m2 \times 2 + m3}{6}$$

Both mixed features were also significant on the correlation coefficients (regarding the mf_1 feature, $r = -.598, t(34) = -4.350, p < .0001$, and for the mf_2 feature, $r = -.607, t(34) = -4.450, p < .0001$). **Figure 7** and **Figure 8** also show the scatter plot and a regression line of the two mixed features, respectively. Since the mf_2 was the most significant in the above features, we could conclude that the mf_2 feature was the most effective met-

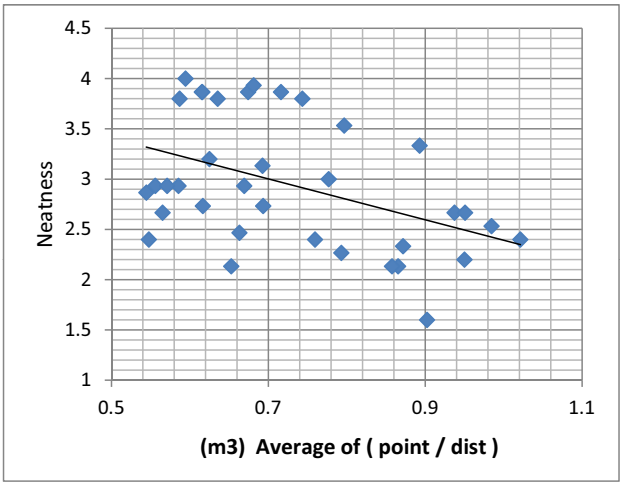

Fig. 6. (m3) Average of (point/dist), $r = -.440$

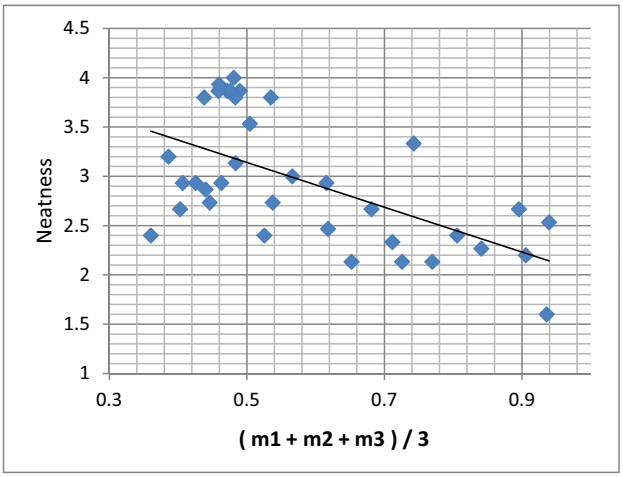

Fig. 7. Simple mixed feature. $(m1+m2+m3)/3$, $r = -.598$

ric for estimating neatness of handwritten notes.

4.4. Discussion

Here we consider the reasons of the relationship between the significant features and the neatness rate. Basically the variance of point per distance, which was the important feature, was reduced when the student kept the writing speed constant. Especially the strokes with Ramer=1 were simple, but curvy. Hasty students tend to write such strokes without keeping consciousness. As a result, note of careful students produces low variance of point per distance.

The second feature, average of Ramer, was also straightforward for reasoning. Careful students securely split their strokes even if the stroke was simple and short especially in Kanji characters. Hasty students tend to concatenate these short strokes, which increased the percentage of higher Ramer value.

The third feature, average of point per distance, was

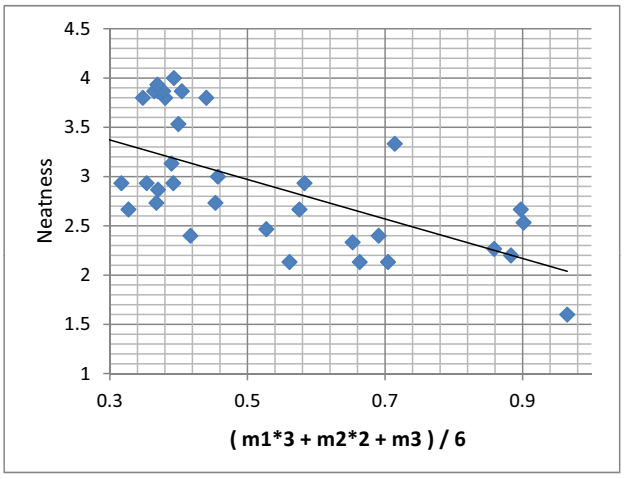

Fig. 8. Weighted mixed feature. $(m1 \times 3 + m2 \times 2 + m3)/6$, $r = -.607$

difficult to find the reasons. Intuitively, the value increased when students wrote slowly. From the result, however, when the average of point per distance becomes higher, the neatness rate becomes lower. By observing the largest average note, we found the student wrote smaller characters by many shorter strokes. Since the distance of the shorter stroke was smaller, it increases the average value of point per distance. Incidentally, the student wrote smaller character did not keep the pen speed constant.

5. Conclusion and Future Work

In this paper, we examined significant handwritten features for estimating neatness rate from online note data. We calculated 14 basic features from handwritten notes, and checked the correlation coefficients with a reference neatness rate. We found that (1) variance of pen speed, (2) average of angular point, and (3) average of pen speed were the significant features for evaluating neatness of the handwritten notes. We also clarified that the weighted average of the above three features produces significant correlation coefficients ($r = -.607$). The metrics can be effectively used for estimating neatness rate from the huge amount of handwritten stroke data consists of natural text written in Japanese.

We have not confirmed the effectiveness of the metrics for other language such as English, Spanish, and Chinese. For future work, we will evaluate the applicability of the metrics. Also, we will employ the metrics for our student note sharing system to assess how the proposed method relieves the teacher's burden and encourages the students' improvement of carefully writings.

Acknowledgements

Our research is partly supported by Japan Society for the Promotion of Science(JSPS) KAKENHI Grant-in-Aid for Young Sci-

entists (A): Grant Number 23680078.

References:

[1] F. Aşıcıoğlu and N. Turan, "Handwriting changes under the effect of alcohol", Forensic Science International, 132(3):201 – 210, 2003.

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

[3] K. Bielaczyc, P. L. Pirolli, and A. 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.

[4] A. Gartner, M. C. Kohler, and F. Riessman, "Children teach children: Learning by teaching", Harper & Row, 1971.

[5] L. Julia and C. Faure, "Pattern Recognition and Beautification for a Pen Based Interface", In Proc. of the 3rd Int. Conf. on Document Analysis and Recognition, volume 1, pp. 58–63, August 1995.

[6] M. Miura, S. Akase, and T. 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, pp. 195–203, June 2013.

[7] M. Miura, T. Sugihara, and S. Kunifuji, "Improvement of Digital Pen Learning System for Daily Use in Classrooms", Educational Technology Research, 34:49–57, October 2011.

[8] M. Miura and T. Toda, "Extraction of Writing Carefulness from Online Handwritten Data", In 22nd FIM Int. Conference on Interdisciplinary Mathematics, Statistics and Computational Techniques, pp. 113–118, November 2013.

[9] H. Miyao and R. Maruyama, "On-Line Handwritten Flowchart Recognition, Beautification and Editing System", In Int. Conf. on Frontiers in Handwriting Recognition (ICFHR), pp. 83–88, 2012.

[10] B. Paulson and T. Hammond, "PaleoSketch: Accurate Primitive Sketch Recognition and Beautification", In Proc. of the 13th Int. Conf. on Intelligent user interfaces, IUI '08, pp. 1–10, New York, NY, USA, 2008, ACM.

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

[12] P. Y. Simard, D. Steinkraus, and M. Agrawala, "Ink Normalization and Beautification", In Proc. of 2005 Int. Conf. on Document Analysis and Recognition (ICDAR'05), pp. 1182–1187, August 2005.

[13] X. Zhu and L. Jin, "Calligraphic Beautification of Handwritten Chinese Characters: A Patternized Approach to Handwriting Transfiguration", In Proc. of ICFHR2008, pp. 135–140, 2008.

Name:
Motoki Miura

Affiliation:
Faculty of Basic Sciences, Kyushu Institute of Technology

Address:
1-1 Sensui, Tobata, Kitakyushu, Fukuoka, 804-8550, JAPAN

Brief Biographical History:
He received B.S., M. E., and D. E. degrees in electronics engineering from the University of Tsukuba in 1997, 1999, and 2001, respectively. From August 2001 to March 2004, he worked as a research associate at TARA Center, University of Tsukuba. From April 2004 to March 2009, he worked as an assistant professor in the School of Knowledge Science, Japan Advanced Institute of Science and Technology.

Main Works:
• Human-Computer Interaction, Learning Environments, and Creativity Support Systems

Membership in Learned Societies:
• IPSJ, JSSST, IEICE, JSAI, ACM, JSET