



Questioned Documents Section – 2010

J7 Optimal Variables for Handwriting Identification

Yoko Seki, MA*, National Research Institute of Police Science, 6-3-1, Kashiwanoha, Kashiwa-shi, Chiba-ken, 277-0882, JAPAN

After attending this presentation, attendees will understand the difference between human pattern recognition and computer based superposition in handwriting identification. Attendees also will understand the availability of the human pattern recognition to handwriting identification.

This presentation will impact the forensic science community by demonstrating the variety and the flexibility of the human pattern recognition in the field of handwriting identification.

A document examiner examines handwriting mainly by a qualitative method based on his/her knowledge and experiences. The qualitative examination, compared with the quantitative examination, possesses less objectivity and is believed to be less reliable. However, an examiner's opinion is, in fact, highly reliable. This is because the examiner has much knowledge about the handwriting and chooses the strategy and variables that are most appropriate to his/her case. So, an analysis of the strategies and variables an examiner uses and the quantification of them will contribute to the establishment of the objectivity in the examination.

Classification of handwriting samples was done following the procedure below. Six people wrote two kinds of Japanese Hiragana characters six times in square style. Both characters were written in one stroke. One character has a curvature stroke and the other has a linear stroke. There were 36 handwritten samples (6 subjects x 6 times) per character and all the samples collected were 72, that is, 2 kinds of characters x 6 subjects x 6 times. Coordinates of handwritten samples were measured at 21 points such as the starting point and the stroke terminal defined beforehand. Coordinates were standardized as for the origin and the character size. Then, handwritten samples were reconstructed by connecting standardized coordinates. Three kinds of classification experiments were done. Experiment 1 - Cluster analysis: Thirty-six data sets of standardized coordinate data of thirty-six handwritten samples were classified into 6 groups using cluster analysis. Cluster analysis finished at the stage where clusters were merged into 6. Experiment 2 - Classification of the reconstructed samples by the visual examination: One subject, who was an active forensic document examiner, classified 36 reconstructed samples into 6 groups according to the similarity of the samples. The subject was instructed the samples to be geometric figures, not characters. The subject was interviewed about the variables used for the classification after the experiment. Experiment 3 - Classification of the original handwritten samples by the visual examination: This was similar to a case work. One subject, who was the same subject as the subject participated in Experiment 2, was instructed to classify 36 handwritten samples into six groups according to the similarity of the samples. Experiment 3 was done one month after Experiment 2. The subject was interviewed about the variables used for the classification after the experiment. Three experiments were done to two kinds of characters respectively. Correct classification ratio of the three experiments was calculated and compared. Correct classification was defined as follows: A cluster was defined to be equal to the subject whose samples were contained in the cluster most. That is, if a cluster had four samples of the writer No.1, one sample of the writer No.2 and one sample of the writer No.3, the cluster was defined as the writer No.1's. After labeling each cluster to the writer, correct classification ratio was calculated. Correct classification ratio was defined as the ratio of correct writer's samples to the whole samples (=36 samples).

Average correct ratio of the two characters was the highest in Experiment 3 (87.5%). Experiment 1 showed 48% and Experiment 2 showed 44%. The subject's answer showed that the hooked shape in the stroke initial was weighted in Experiment 2, while stroke initial was ignored in Experiment 3. The terminating manner in the stroke was weighted in Experiment 3. The condition of the stroke termination highly correlated to the kinetic aspect of writing. These suggested that the knowledge on characters and the kinetic information while writing were important to the correct identification.

Handwriting Identification, Cluster Analysis, Human Pattern Recognition