J6   How Does a Document Examiner Differentiate Authors?

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 in the pattern recognition between human and the computer based superposition. Attendees also will understand the availability of the human pattern recognition to the handwriting identification.

This presentation will impact the forensic science community by demonstrating the superiority of the human pattern recognition to the computer based superposition in the handwriting identification under a certain condition.

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.

Six subjects wrote 24 kinds of Japanese characters six times in square style. Six out of 24 kinds were Hiragana characters and 18 out of 24 kinds were Kanji characters. Each kind of character had 36 samples (6 subjects x 6 times) and all the samples collected were eight hundred and sixty-four characters, that is, 24 kinds of characters x six subjects x six times. Then, 36 samples of the same character were classified into six groups following the procedures below: (1) Cluster analysis: Each handwritten sample was measured its x-y coordinates at 21 measuring points defined beforehand. Coordinates were standardized as for the origin and the size. Cluster analysis was done using standardized coordinates. Cluster analysis finished at the stage where clusters were merged into 6, (2) Classification by the visual examination: This was similar to a case work. A document examiner observed 36 samples and classified them into six groups according to the similarity of the samples, and (3) Cluster analysis based on the variables used for the visual examination: The examiner was interviewed about the variables used for the classification after the trial. Coordinates of characters were measured and calculated based on the variables used by the examiner and then cluster analysis was done.

Correct classification ratios of the three classifications methods were 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. If a cluster had four samples of the subject No.1, one sample of the subject No.2 and one sample of the subject No.3, the cluster was defined as subject No.1's. After labeling each cluster to the subject, correct classification ratio was calculated. Correct classification ratio was defined as the ratio of correct subject's samples to the whole samples (=36 samples).

Classification by the visual examination showed the highest correct classification ratio in the three methods. Correct classification ratio was higher in the cluster analysis using coordinates than the cluster analysis using the observation variables. Variables used for the visual examination were the relationship between the components in space and the size. Thinking that the number of the variables is much smaller in variables used for the visual examination than the coordinates (=42 variables), human recognition is more effective than pattern matching using the superposition of the patterns under the condition that the task was the classification of small number of data into several groups. These suggest that the human recognizes a pattern roughly and focuses the relationship between the components, different from the local pattern matching.

Handwriting Identification, Cluster Analysis, Visual Examination