J28   Adding an Objective Approach to Questioned Document Examination Using Principal Component Analysis (PCA) and Mahalanobis Distance

Loren M. Williams, BS*, Bethlehem, PA 18020; Lawrence Quarino, PhD, Cedar Crest College, Allentown, PA 18104; Morgan Mills, MSFS, Burlington, NC 27215

Learning Overview: After attending this presentation, attendees will better understand the potential role that statistical analysis can play in helping to support subjective determinations of authorship in questioned document examination.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by demonstrating a simplified way to incorporate statistics into questioned document examination.

Document examiners have a skill that lay people do not. Studies have shown that professional examiners make fewer mistakes and are more accurate than laypeople when determining authorship of writing.[1,2] This skill, which is based on experience and expertise, is essential in helping to determine fraudulent writing and to correctly assign the writing origin of documents.

However, after the release of the National Academy of Science's Report on strengthening forensic science in 2009, it was made clear that objective measures of comparison are needed to augment and strengthen conclusions that are essentially subjective in nature.[3] One possible approach is to couple the skill of questioned document examiners with statistics to create a more objective methodology, allowing stronger conclusions to be made. Since questioned documents offer measurable characteristics such as the height of letters, the ratio of heights of adjacent letters, and the distance between words, the potential to use some form of multivariate statistical analysis in questioned document examination exists. To this end, this study incorporated the use of PCA and Mahalanobis distance to classify groups of known writing. In addition, the potential of using this method to compare an unknown with exemplar writing was explored.

To test this, 20 handwriting samples were collected from 50 individuals. The handwriting samples were all the same and taken on four different days. Writing was in script and the template used was a bank check. Various known features, such as height of capital and lowercase letters, distances between words, and ratios of letter heights, were measured using Adobe® Photoshop® 2017. Measurements were placed into a spreadsheet and used to perform PCA to attempt discrimination of handwriting between individuals. Of the 50 individuals, 32 were able to be distinguished visually when examining the principal components in a 3D plot. Unknown writing samples taken from 10 participants in the original 50 were collected and compared to the dataset. Using PCA, it was determined that each unknown could be visually placed into the correct data set. In addition, Mahalanobis distances were also calculated. Results from this statistical test showed that 8 of the 10 unknowns were able to be correctly classified into their respective data sets. An additional 10 unknowns were collected from individuals not in the original group of 50 and were compared to the dataset. The visual analysis using PCA indicated 9 of the 10 unknowns were not classified into any of the original 50 datasets, although the Mahalanobis distances placed several of the unknowns into one of the original 50 datasets based on criteria used in the study.

Results from the study seem to indicate that PCA offers the potential of objectively confirming conclusions reached in traditional questioned document examinations. Going forward, results reached in this study should be confirmed by the subjective examination of a trained document examiner. If results in this manner are consistent with the statistical analysis, it will lend credence to the use of statistical analysis (PCA in particular) to be used in tandem with traditional handwriting examination. This could help alleviate the criticisms often levied on handwriting comparisons as being too reliant on subjective analysis. Finally, PCA could be a tremendous tool in showing the variability in handwriting exemplars taken from individuals. Objective determination of this variability can be helpful in assessing the validity of particular exemplars used in writing comparisons.

Reference(s):
[1.] Kam M., Fielding G., Conn R. Writer Identification by Professional Document Examiners. *J Forensic Sci* 1997;42:778–86.
[2.] Sita J., Found B., Rogers D. Forensic Handwriting Examiners' Expertise for Signature Comparison. *J Forensic Sci* 2002;47:1117–24.
[3.] National Research Council (NRC), Committee on Identifying the Needs of the Forensic Science Community. (2009). *Strengthening Forensic Science in the United States: A Path Forward.* Washington, DC: The National Academies Press. 1-328.

Principal Component Analysis, Mahalanobis Distance, Handwriting Comparisons