J2 Evaluation of the Language-Independent Process in the FLASH ID System for Handwriting Identification

Mark A. Walch, MBA, The Gannon Technologies Group, 1000 North Payne Street, Alexandria, VA 22134; Donald T. Gantz, PhD, George Mason University, Department of Applied Information Technology, MS 1G8, 4400 University Drive, Fairfax, VA 22030; John J. Miller, PhD*, George Mason University, Department of Statistics, MSN 4A7, 4400 University Drive, Fairfax, VA 22030; and JoAnn Buscaglia, PhD, Federal Bureau of Investigation Laboratory, CFSRU, Federal Bureau of Investigation Academy, Building 12, Quantico, VA 22135*

After attending this presentation, attendees will be exposed to some of the "inner workings" of the FLASH ID system. This will show that the FLASH ID system is not attempting to replicate the actions of a document examiner, but will present the examiner with a powerful new tool to assist in the identification of questioned documents. By understanding better the methods behind the language - independent mode, the attendees will appreciate its potential for assisting document examiners. Attendees will be exposed to a study demonstrating the effectiveness of the language-independent mode in identifying the writers of questioned documents.

This presentation will impact the forensic science community by informing them about a system that can assist forensic document examiners to identify the writer of a questioned document. This presentation will present information about the statistical methods behind the language-independent mode of operation and give results that show the excellent performance of the system.

FLASH ID was introduced to the forensics community at AAFS 2008. FLASH ID is a totally automated system for handwriting identification that can operate in either a character-based or language- independent mode. At the 2008 AAFS meeting, a quantification of handwriting and its application to handwriting identification were described. An evaluation of how the FLASH ID system performed using the character-based mode was presented. That mode is based on segmenting handwriting into alphabetic characters, which are then associated with graph structures. In the language-independent mode, the alphabetic dependency is removed through a language-independent segmentation algorithm for handwriting. The resulting segments do not necessarily correspond to any alphabetic characters. This presentation demonstrates the language-independent segmentation and gives performance results when it is the basis for identifying the writer of a questioned document. A major advantage of language-independent segmentation is that it does not require character recognition in order to segment handwriting. The language-independent technology has been implemented in FLASH ID for two years.

This presentation will contain a discussion of the language- independent segmentation algorithm and examples of its implementation. The segments of handwriting generated by the algorithm are associated with graphic entities that provide a topological classification and geometric features (just as for the character-based mode). Shape codes are simple features that are used to define sub- categories for each class of graph. The combination of a graphic class and a shape code category is called a grapheme. Handwriting is then modeled at the grapheme level. That is, each writer with known writings is characterized within each grapheme that occurs with some frequency in the known writings. A questioned document is broken down into graphemes and each writer in the database of known writers receives a score for each occurrence of a grapheme in the questioned document. Scores for all occurrences of graphemes in the questioned document are summed for writers. These total scores are used to rank writers by likely writer-ship of the questioned document. The total scores and the associated ranks are displayed by FLASH ID for evaluation by forensic document examiners. This presentation will illustrate the excellent performance results for the language-independent process in multi- language applications. **Forensic Document Examination, Automated Handwriting Identification, Statistics**

 ** Presenting Author*