Document Title:	**Evaluation of an Emerging Automated Searching Technology to Improve the Efficiency and Reliability of Latent Print Comparisons**
Author(s):	**Jessica J. Davis, Mary M. Hood**
Document Number:	**300372**
Date Received:	**March 2021**
Award Number:	**2016-DN-BX-K004**

Final Summary Overview

Federal Agency: U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Award Number: 2016-DN-BX-K004

Project Title: Evaluation of an Emerging Automated Searching Technology to Improve the Efficiency and Reliability of Latent Print Comparisons

Authors: Jessica J. Davis, Latent Print Section Supervisor, Mary M. Hood, Forensic Scientist, Virginia Department of Forensic Science

Submission Date: June 25, 2019

Submitting Official: Alka B. Lohmann

Recipient: Virginia Department of Forensic Science
700 North Fifth Street
Richmond, VA 23219-1416

Project Period: January 1, 2017 – September 30, 2019

Signature of Submitting Official

Date

Purpose

The goal of this research was to determine the accuracy and reliability of the LatentSleuth technology and determine if integrating LatentSleuth into current comparison workflow for complex comparisons improves efficiency and reproducibility in relation to existing methods.

The objectives for this project were:

1. Validate the LatentSleuth software using ground truth latent prints.
2. Integrate LatentSleuth into the current comparison workflow to determine efficiency.
3. Disseminate the results and methods to the latent print community.

Project Design and Methods

Phase 1: Validation

Accuracy of the LatentSleuth software was evaluated based on latent print complexity, determined by the quality of the latent print, and comparison complexity, determined by the number of comparisons conducted. Latent prints, with corresponding true-mate reference images, were purchased from Mideo Systems, Inc. through the TrainingWorks software.

To ensure consistent characterization of latent print quality, two qualified examiners independently evaluated the quality of each latent print. Quality assessment levels of high, medium high, medium low and low were based on a modified version of the Scientific Working Group of Friction Ridge Analysis, Study and Technology (SWGFAST) guidelines[1] to improve the consistency of assessments between the examiners. The validation study only included those latent prints with a consensus opinion of the quality. Over seven hundred latent print images were analyzed independently; fifty latent prints of each quality level were used in the validation. The validation included only latent fingerprints for the validation; joints and palm prints were not included in the research. Using fifty data points per quality level resulted in an estimated error rate of at most +/- 15%. The choice of n=50 per level was selected for efficiency of the study to minimize the impact on case productivity at the Virginia Department of Forensic Science (DFS) and to have a reasonable estimated error rate from this study.

Comparison complexity was stratified into three categories: low, medium, and high. A comparison was defined as one latent print against one finger, e.g., one latent print compared against one ten-print exemplar equaled ten comparisons. For this validation, low comparison complexity was defined as thirty comparisons (three exemplars), medium as fifty comparisons (five exemplars), and high as one hundred comparisons (ten exemplars), Table 1.

Comparison Complexity	Quality of Latent Print			
	High	Medium High	Medium Low	Low
Low	50	50	50	50
Medium	50	50	50	50
High	50	50	50	50

Table 1: Number of searches completed for various qualities and comparison complexities

The latent print and corresponding true-mate reference print images were exported from Mideo TrainingWorks and imported into the LatentSleuth software at 1000 pixels per inch. Six hundred LatentSleuth searches were conducted: fifty latent prints of each quality level at each comparison complexity. Initially, all of the latent prints were processed automatically only by the software. This included the software's creation of a high contrast overlay image of the latent print based on ridgelines from the short segments of the ridges. A quality map, also automatically produced, masks areas of limited quality ridge flow**,** preventing the areas from being included in the search. Preliminary results indicated manual processing of medium low and low quality latent prints was necessary. Manual processing, examiner editing of ridgelines and the quality map, was conducted on each medium low and low quality latent print. The results for the automatic and manual processing searches are included below. A trained technician conducted data collection, including importation of all images, selection of the quality region (the area of latent print to be searched), manual processing of the latent prints, and review of results against ground truth. A qualified examiner reviewed the results as necessary.

Phase II: Incorporation into casework: efficiency determination

All latent print examiners in the state worked actual cases as the Original Examiner. Specific examiners in each of the four DFS testing laboratories completed competency tests on the LatentSleuth software and used the software as the verifying examiner. Every latent print analyzed as part of routine casework, now included a quality assessment in the analysis phase based on the modified version of the SWGFAST quality chart. Latent prints (only fingerprints – lower joints and palm prints were not used) deemed medium low or low quality required use of the LatentSleuth software by the verifying examiner. Time spent searching for corresponding ridge detail was recorded by both the original examiner, using traditional manual searching methods, and the verifying examiner, using LatentSleuth.

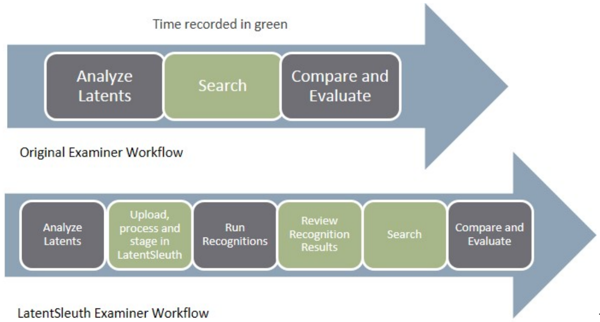

Figure 1: Comparison workflow

A search was considered finished, and the recording of time stopped, when an examiner was content to not look for another source for the latent print, i.e. when an examiner was no longer searching for ridge correspondence. Time spent completing the comparison, making an evaluation and documenting was not included in the time recorded (Figure 1). This distinction was made in an effort to eliminate examiner variability on documentation and time spent on decision-making. In doing so, a more accurate evaluation of the efficiency of LatentSleuth could be determined. Data collected for the verifying examiners using LatentSleuth included the time spent uploading images (latent prints and exemplars) into the software, processing the latent prints and exemplars, and reviewing the prioritized list, as well as any time spent conducting manual searches (i.e. if an identification was not made from review of the prioritized list).

Data Analysis

Phase I: Validation

A result was deemed accurate if the true-mate reference image was located within the top five positions of the prioritized list and if the overlay was located on the correct corresponding ridge detail. Accuracy results for fully automated processed latent prints across all comparison complexities were 98% (high quality), 97.3% (medium high quality), 70.7% (medium low quality), and 52.0% (low quality), Table 2. Accuracy results for manually processed medium low and low quality latent prints across all comparison complexities were 98.7% and 86.7%, respectively.

Manual processing of lower quality prints is in line with the LatentSleuth software suggested guidelines.

		Quality of Latent Print					
		High	Medium High	Medium Low		Low	
		Fully automated	Fully automated	Fully automated	Manual processing	Fully automated	Manual processing
Comparison Complexity	Low	49/50, 98%	49/50, 98%	37/50, 74%	50/50, 100%	28/50, 56%	44/50, 88%
	Medium	49/50, 98%	49/50, 98%	35/50, 70%	50/50, 100%	26/50, 52%	44/50, 88%
	High	49/50, 98%	48/50, 96%	34/50, 68%	48/50, 96%	24/50, 48%	42/50, 84%
	Overall	98.0%	97.3%	70.7%	98.7%	52.0%	86.7%

Table 2: Accuracy results based on quality of latent print and comparison complexity

Table 3 is the result of an in-depth review of the medium low and low quality latent print searches (manual processed) to determine how often the true-mate reference print was in the number one position on the prioritized list.

		Quality of Latent Print			
		Medium Low		Low	
		Accurate Results	#1 Candidate	Accurate Results	#1 Candidate
Comparison Complexity	Low	50/50, 100%	48/50, 96%	44/50, 88%	41/50, 82%
	Medium	50/50, 100%	47/50, 94%	44/50, 88%	35/50, 70%
	High	48/50, 96%	46/50, 92%	42/50, 84%	31/50., 62%

Table 3: Percent of number 1 candidates in result for Low and Medium Low searches

Phase II: Incorporation into casework: efficiency determination

In one hundred and thirteen cases, the Original Examiners completed comparisons using traditional manual searching methods for locating the source of a latent print, while the LatentSleuth Examiners, serving as verifiers of the cases, completed the cases using the software to assist in searching. Cases ranged in size from one to eight latent prints combined with one to six exemplars. The number of latent prints recorded reflected only the low and medium low latent prints in the case. Additional latent prints of higher value may have been present in the case as well, but were not included in the data. On average, the Original Examiner spent 44.9 minutes (standard deviation 106.0) per case searching. The LatentSleuth Examiner spent an average of 51.6 minutes (standard deviation 119.58) minutes. The mean difference between examiner's times per case is 6.7 minutes; the standard error of that sample mean is 3.5. The lower and upper quartiles also are close to zero (-6 and 13), and three-fourths (75%) of the 113 times are between -20 and 20. The formal student's t-test renders a "p-value" of 0.056, which suggests that the Original Examiner, not using LatentSleuth, took slightly longer to complete the cases than the LatentSleuth Examiner.

The estimated LatentSleuth (LS) Examiner time was calculated by plotting the LatentSleuth Examiner search times against the Original Examiner search times, removing the one highest and one lowest outlier, (Figure 2: LS time ~ 4.56(Examiner time)$^{0.6}$). The exponent being lower than one suggests that use of the LatentSleuth software does, on average, utilizes less examiner time.

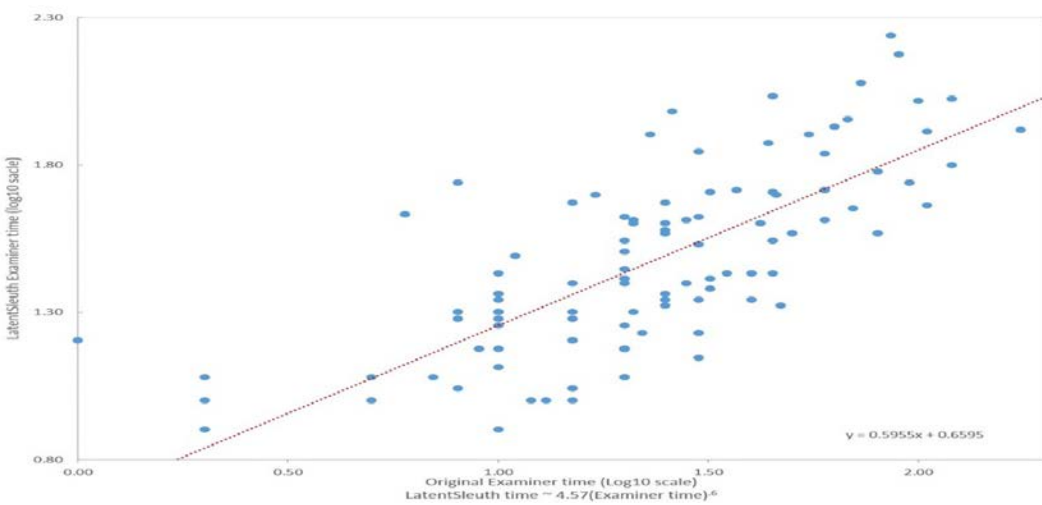

Figure 2: Scatter graph illustrating the relationship between LatentSleuth examiner and Original examiner search time per case

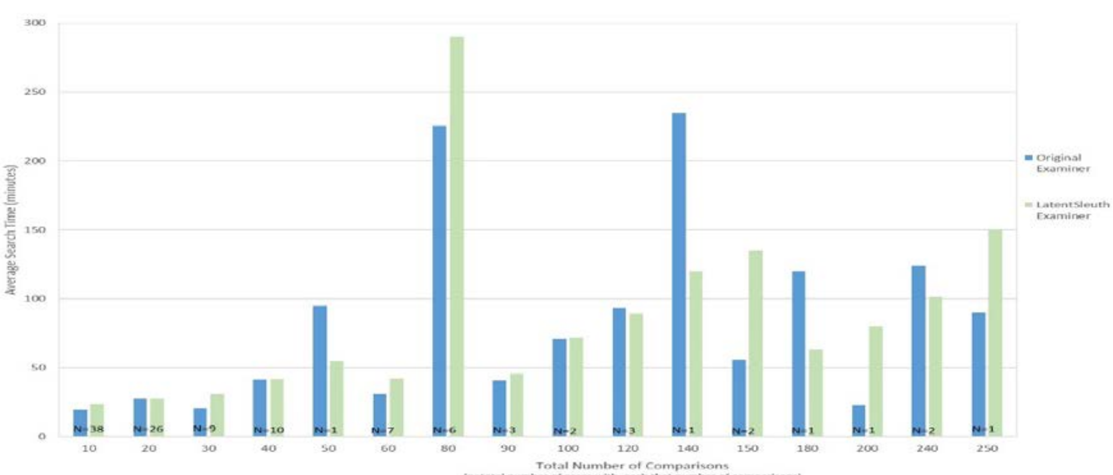

Figure 3: Average search time of LatentSleuth and Original examiners versus total number of comparisons in a case

Search time alone did not tell the complete story of the effectiveness of the LatentSleuth software. Assessing the software's usefulness as related to the number of comparisons, Figure 3 shows the relationship between each examiner's average search times per total comparisons of medium low and low quality latent prints in a given case. For simplicity in data collection, the total number of comparisons (calculated by the number of latent prints multiplied by the number of exemplars multiplied by ten) represents the total number of potential comparisons, regardless of whether an examiner conducted comparisons with each card (e.g. an identification was made to the first exemplar examined, so additional cards were not considered).

Cases with ten comparisons consisted of one latent print and one exemplar; all other total number of comparisons consisted of combinations of numbers of latent prints and exemplars (i.e. twenty comparisons may be comprised of two latent prints and one exemplar or one latent print and two exemplars). Therefore, either the number of latent prints or the number of exemplars may drive the efficiency of LatentSleuth. Figure 4 and Figure 5 depict normalized data to explore if either variable had an impact.

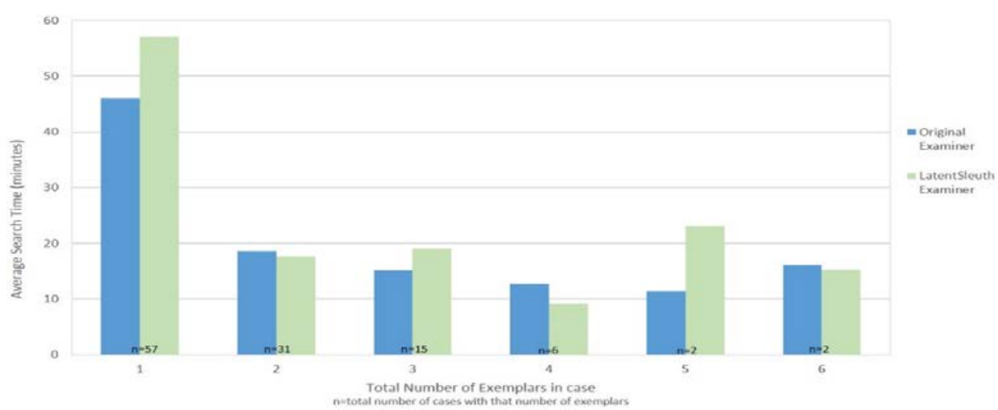

Figure 4: Average search time per exemplar in relation to total number of exemplars in a case

There is no apparent trend to indicate that the number of exemplars or latent prints had an effect on the LatentSleuth Examiner. However, the use of LatentSleuth demonstrated the most beneficial impact on time in cases involving the comparison of six and seven latent prints (Figure 5). While the sample sets for the cases concerning six and seven latent prints are low, just one case at each amount, it can be surmised that examiner fatigue could occur with this number of manual comparisons of lower quality latent prints, demonstrating that using LatentSleuth is more efficient. The two cases with eight latent prints do not support this theory; however, it may be of note that the same LatentSleuth Examiner was involved in both cases.

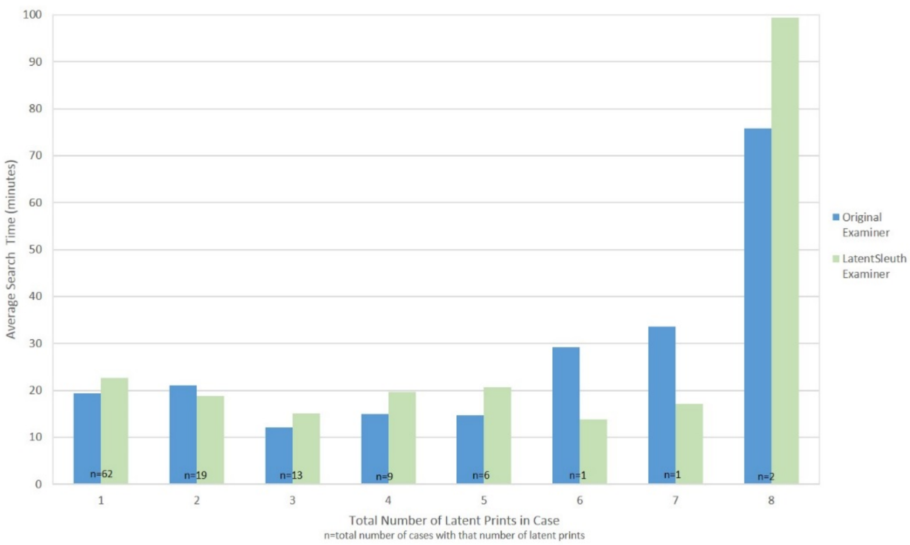

Figure 5: Average search time per latent print in relation to total number of latent prints in a case

Of the 113 cases, five (4.4%) involved changes of conclusion opinions. A change of opinion occured when the Original and LatentSleuth Examiners reached different conclusion decisions for a comparison. After consultation, a discussion of the information each examiner used to render an opinion was had and a consensus decision on the final conclusion of the comparison was determined. Typically, differences in latent print enhancements, orientation, and assessments of distortion play a role in conclusion differences.

Case	Number of latent prints	Number of Exemplars	Conclusion Change
1	8	1	Inconclusive to Identification
2	5	2	Exclusion to Identification
3	1	1	Inconclusive to Identification
4	1	1	Inconclusive to Identification
5	1	1	Inconclusive to Identification

Table 4: Total number of low and medium low latent prints and exemplars in cases with latent prints resulting in changes of conclusion opinion

The changes of opinions in the study occurred in the comparisons of six latent prints (two of which were from the same case) out of the total 234 low and medium low quality latent prints involved, or 2.6%. Furthermore, five of the six conclusions changed from inconclusive (either due to the quality of the latent print, exemplar, or both) to identification, and one conclusion changed from exclusion to identification. In each of these comparisons, the use of the LatentSleuth software assisted the LatentSleuth Examiner to locate the ridge correspondence and make the identification conclusion.

Three of the cases consisted of one exemplar and one latent print (Table 4). Two latent prints were from the same case (Case 1), which contained a total of eight latent prints and one exemplar. Case 2 involved two exemplars and five medium low and low quality latent prints. Figure 6 depicts the latent print from this case that was initially excluded by the Original Examiner, but was ultimately identified due to LatentSleuth locating corresponding ridge detail not considered by the Original Examiner. In Case 2, there were two exemplars and sixteen total latent prints (only five of which were of low and medium low quality and included in the study).

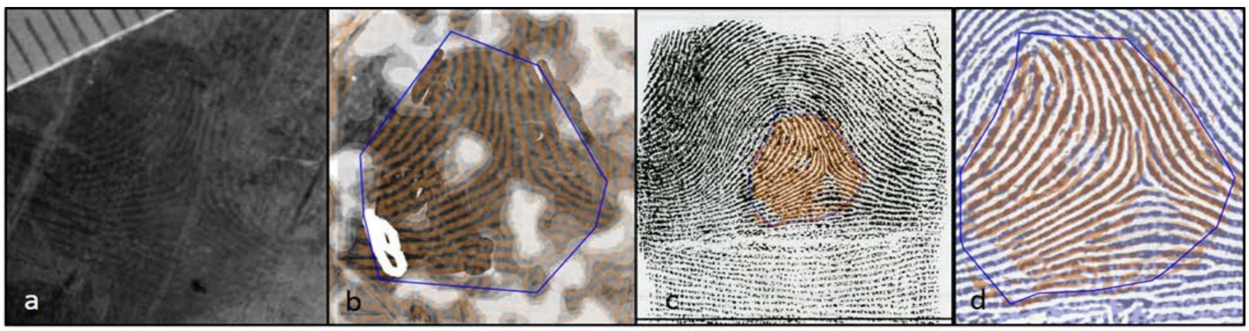

Figure 6. Original examiner excluded, identification was made to the LatentSleuth, candidate 1, a) original image, b) processed image with quality area, c) latent overlay on corresponding detail of exemplar, d) latent overlay on processed exemplar

Findings

The LatentSleuth software provided accurate results across all quality levels of latent fingerprints. Of the six hundred automated searches conducted within this validation study, accurate results (the true-mate reference image located within the top five positions and on the correct corresponding ridge detail) were produced in 571 (95.2%) searches with the true-mate reference image ranked in the number one position in 533 of the 600 (88.8%) searches conducted. Not surprisingly, the LatentSleuth software was less effective with low quality latent prints in high comparison complexity. The software was validated for use in casework.

The LatentSleuth software exhibited effectiveness in casework. Applying the formula achieved from the scatter graph in Figure 2: LS time ~ 4.56 (Examiner time)$^{0.6}$, the benefit of an examiner incorporating LatentSleuth into comparisons is substantial in instances where examiner time is greater than two hours, Figure 7.

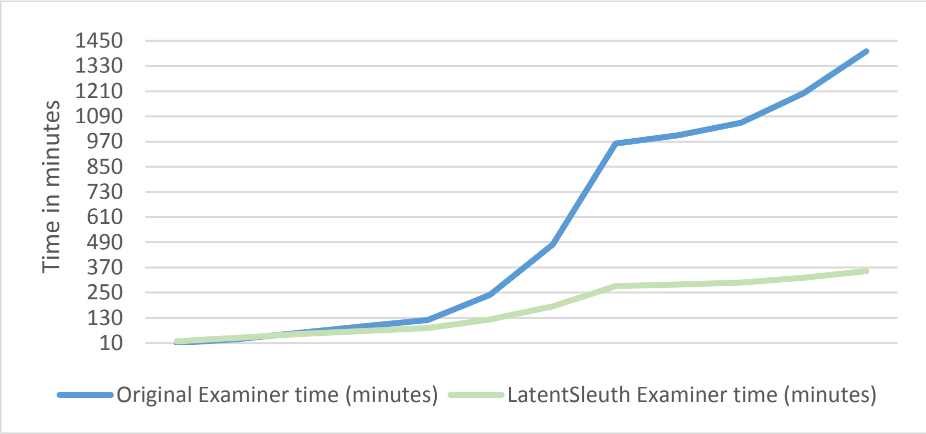

Figure 7: Demonstrating LatentSleuth time based on LS time ~ 4.56 (Examiner time)$^{0.6}$

Moreover, taking into account the number of comparisons alone does not depict the best use of LatentSleuth. Overall, the data does not clearly support a single type of comparison case where LatentSleuth is most effective or efficient. Our hypothesis was that in large comparison cases, specifically in cases with a large number of exemplars, use of LatentSleuth in the workflow would demonstrate a significant reduction in the time an examiner spent searching for ridge detail correspondence. The data, however, suggests the opposite - the benefits of the software were more apparent based on the number of latent prints. The more latent prints in a case, specifically lower quality prints, the more helpful LatentSleuth appeared to be, suggesting a possible remedy for examiner fatigue.

Cases involving ten comparisons were the most common (mode) in the study (39 of the 113 cases or 11%). Three of these cases resulted in a change of opinion; thus, 7.9% of the ten comparison cases had a change of conclusion opinion as a result of the LatentSleuth software locating ridge detail correspondence further assessed by the examiner. Consensus was reached in every conclusion after the involved examiners discussed the information considered in each comparison. The most common reasons identified for the changes in these conclusions were unknown orientation of the latent prints and limited ridge detail present in the latent prints. In hindsight, requiring a third examiner to verify the Original Examiner's conclusions without LatentSleuth could have verified that LatentSleuth had truly found missed identifications.

Implication for Criminal Justice Policy and Practice in the United States

The validation of the software and demonstrated effectiveness in casework has several implications for criminal justice practices. Traditional manual comparison workflows can be arduous, particularly in searching for low quality latent prints and latent prints with unknown orientation and/or anatomical aspect. Automated searching not only assists in limiting the amount of time spent searching for corresponding ridge detail, the tool provides an effective means to locate corresponding ridge details in the more challenging latent prints. Furthermore, incorporating the LatentSleuth software into casework can help overcome examiner fatigue, which has been shown to affect search accuracy[2].

The information obtained within this research project is being disseminated in several different forums. A lecture was given at the Chesapeake Bay Division International Association of Identification (IAI) Spring Education Conference in Richmond, Virginia in April 2019; a presentation is scheduled for the parent body IAI Educational Conference in Reno, Nevada in August 2019. Two papers, one on each phase of the project, are being prepared for publication in the Journal of Forensic Identification. In addition, the criteria for use of LatentSleuth, published in the DFS Latent Print Procedures Manual, is available for public access on the DFS website at https://www.dfs.virginia.gov/documentation-publications/manuals/.

References

1. SWGFAST. Standard for Examining Friction Ridge Impressions and Resulting Conclusions. www.swgfast.org 2013, ver. 2.0, (accessed Mar 2016).
2. Busey, T., Swofford, H.J., Vanderkolk, J., Emerick, B. The impact of fatigue on latent print examinations as revealed by behavioral and eye gaze testing. Forensic Science International, 2015, Vol 251, p202-208.