

NIST Latent Fingerprint
Workshop, 19-20 Mar 2009

Vendor Panel 2. 2C: ELFT-EFS: How to test EFS for latent fingerprint matching.

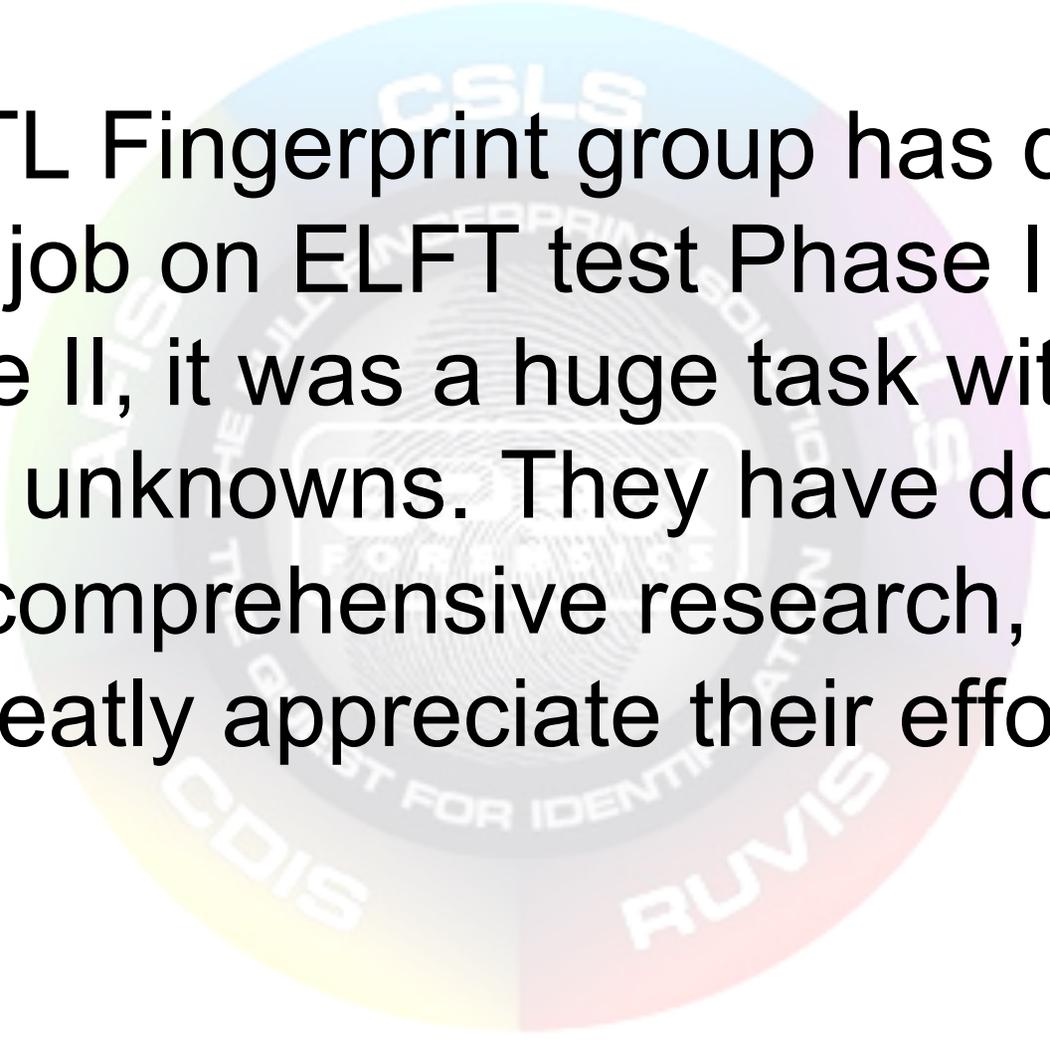
Ways to improve future NIST latent fingerprint technology evaluation tests.

Dmitry Mikhailov,
SPEX Forensics

HORIBA JOBIN YVON

Explore the future

HORIBA



NIST ITL Fingerprint group has done a great job on ELFT test Phase I and Phase II, it was a huge task with many unknowns. They have done very comprehensive research, and we greatly appreciate their efforts.

However we believe there are several things in ELFT Phase II test to be noted and accounted for in future evaluations

- The latent fingerprint test database for Phase II was taken from existing operational AFIS identifications.
- Many of the test latents were digitally pre-processed.
- All latents were extracted automatically during the test (“lights out”), that is not exactly how AFIS works in the field.
- Background database size (50K and 100K images) was not big enough.

1. The latent fingerprint test database for Phase II was created from identifications made in existing operational AFIS (in IAFIS as per NIST information)

- This approach is biased, as it benefits one technology provider (whose AFIS algorithms were used to make original identifications) and it could make negative impact to other vendors' results.
- In our opinion this is unacceptable as it calls into doubt the credibility of Phase II results.

Suggestion for future tests:

- Every participating vendor will provide equal number of latent prints (for example 150-200) and their ten-print mates.
- NIST will make conditions for distribution of those latents by fingerprint pattern type, finger position and number of minutiae if necessary.
- If possible, NIST can also add latent matches that did not come from AFIS environment (manually matched)
- The combination of the above will form a test set that is not biased in favor of any vendor.

This idea was proposed for the first time during the NIST Latent Testing Workshop in April of 2005, and we are surprised that it did not receive any response.

HORIBA JOBIN YVON

2. Many of the test latents were digitally pre-processed. In our opinion this is unacceptable, due to:

- Digital processing permanently changes an original image.
- Digital processing (for example, use of image filters) eliminates some image details (degrades an image) that can affect extraction by AFIS
- Image enhancement could be done to benefit a specific AFIS extraction algorithm to the detriment of the other participating vendors algorithms

- Using the enhanced images, new factor was added to the test, which influence is hard to measure.
- A paradox in the test arises: test organizers used “lights out” concept to “decouple the performance of the software from the performance of Latent Examiners”, but at the same time they allowed somebody (Latent Examiners?) to modify (pre-process) latent images to be tested.

Suggestion for future tests:

All latents used for the test should not to be pre processed, they should be as they are lifted from a crime scene.

3. *“Evaluation of Latent Fingerprint Technology (ELFT) is a NIST project for evaluating automated one-to-many latent fingerprint search technology”*

NIST, “Summary of the Results of Phase I ELFT Testing”

We believe that to evaluate a technology, tests are to be performed in as close to “real world” conditions as possible in the environment and within conditions the technology was designed for.

What is the real life scenario for Latent Fingerprint Search Technology (AFIS)?

AFIS in the field deals with two types of data, very different from each other by source, number and quality

1. Fingerprints of known suspects (ten-print cards):
2. Latent fingerprints, found on crime scenes

Fingerprints of known suspects:

- Are represented in large quantities, thousands and millions, normally exceeding number of latents by the order of 10 or even 100
- On average have acceptable image quality
- For many individuals multiple ten-print records are available (two or more)
- Manual extraction of fingerprint features is not feasible because of large volume.

Conclusion: Processing of ten-print cards in AFIS should be automatic.

Latent fingerprints from crime scenes :

- Are typically 10-100 times smaller in quantity than ten-print cards
- Are partial, fragmented fingerprints
- Normally have low image quality
- May have geometrical distortions
- May have a background that complicates extraction
- May overlap with other fingerprints
- Price of identification mistake (missed hit) is high.

Conclusion: preferred way of processing latent fingerprints in AFIS is automatic extraction followed by human editing and verification

Suggestion for future tests:

- NIST will send all latents to participating vendors for feature extraction and verification (latent images only, not their fingerprint mates).
- Vendors will return latents with verified proprietary feature sets, or proprietary + CDEFFS, or proprietary + CDEFFS + Minutiae Only, as required by test designers.
- All ten-print cards will be enrolled automatically.
- This test will NOT take the place of the “lights out” test, but will be run IN ADDITION to it, thus allowing the results to be compared.

4. Background database size (50K and 100K IMAGES) was not sufficient to observe AFIS technology bounds and limitations.

- Normally AFIS systems in the field deal with larger amount of data that was used for ELFT testing.
- 1% percent drop of performance between searching a gallery of 50K fingerprint images and 100K fingerprint images is not significant. Also this is an average decrease in rank 1 identification, that does not show how every participating SDK did between 50K and 100K images.

What is a good size for latent fingerprint technology evaluation test?

- There is no simple answer. For one AFIS vendor it's 10 thousand ten-print cards, for another it's 10 million.
- For example we have three fingerprint data sets, A, B and C. Gallery B is 10 times bigger than gallery A. Gallery C is 10 times bigger than B.
- For example some latent print has a mate in gallery A with rank 1. Can we predict what rank this mate is going to be in galleries B and C? No, we can not.
- Now let's assume that some latent has rank 2 in gallery A. In this case we can say that in average in gallery B it will have rank 11, and in gallery C the rank will be 101, because non-mates appear in a candidate list randomly in large data sets.

Suggestion for future tests:

- To increase the background database size 10 times and use 100K TEN-PRINT RECORDS (1M IMAGES).
- In order to decrease evaluation time and cost, NIST can establish enrollment time and search time limitations.
- In order to support ELFT-EFS test participating vendors can provide PC equipment for testing as per NIST specifications.



Thanks for your attention!

Dmitry Mikhailov
AFIS Product Manager
Horiba Jobin Yvon Inc
SPEX Forensics Division
3880 Park Ave Edison, NJ 08820
Phone: 732-623-8312
Fax: 732-623-8273
dmitry_mikhailov@usa.com
dmikhailov@mail.crimescope.com

HORIBA JOBIN YVON

Explore the future

HORIBA