

Review of NIST Offline Biometric Testing

Toward Automated Latent Fingerprint Identification Evaluation

Patrick Grother, Image Group, NIST

Latent Workshop

April 5 and 6, 2006

Overview

- NIST Offline Tests
 - Characteristics
 - Review

- Latent Testing Specifics
 - Data
 - Procedures

Offline “Technology” Testing

Definition: Execution of many 1:1 comparisons or 1:N searches on archived biometric samples. Outputs are estimates of fundamental error rates.

- Assess core biometric power
 - Identification, verification, throughput
- Large Scale
 - Can be run on very large populations
 - Statistically Robust
- Repeatable
 - (auditable, also new technologies can be compared)
- Fairly Administered
 - (e.g. competitive evaluations)
- Can be used for multimodal or multi-sample
 - (e.g. Use 2,4,6 fingers)
- Suitable for experimentation
 - (e.g. Latent image restoration studies)
- Inexpensive
 - Compared to scenario tests (which are suited to other things)

Roles

- Black box evaluation
 - Test organization supplies imagery over standard interface
 - Vendor is solely responsible for internal algorithmic application

Recognition Performance Testing

- FRVT 2000 (face, 5 vendors)
- FRVT 2002 (face, 10 vendors)
- FpVTE 2003 (fingerprint, 12+)

Vendor attends with own hardware, software. Costly, inconvenient, insecure

- SDK 2003 - (fingerprint, ongoing)
- SDK 2003 - (face, 4 vendors)
- Minutiae Interoperability Exchange Test (2005-2006)
- FRVT 2006 - (face, started early 06)

Vendor supplies software that implements NIST specified API providing core template generation and matching functions

- Face Recognition Grand Challenge (2004-2005)
- Iris Challenge (2005-2006)

Sample data is used to drive technological development. Testing Component measures improvement.

Testing Timeline

1. Collect Data
2. Supply K% of Data to Capable Organizations for Development
3. Retain 100-K% as Sequestered Data
4. Development phase
5. Post vendor-reported Results
6. Host Workshop to present Analysis and Define Next Steps
7. Loop to Step 1 as needed

8. Acquire (software) implementations for formal evaluation
9. Execute SDK-based Independent Evaluation using Sequestered Data
10. Report

Essentially the approach of ARPA/NIST Grand Challenges

- Iterative phase of Tester-Supplier Cooperative R+D
- A final large scale, independent, evaluation

Dedicated Latent Collection

- Acquire Enrollment Images from N people
 - Ink (to represent legacy)
 - Live scan (current practice)
- Acquire Latent Images from same N people
 - From K surfaces
 - glass,
 - plastic,
 - paper,
 - Using L “imaging” methods
 - Luminescent (Arc, UV etc) + DFO
 - Ninhydrin
 - Powder
 - Cyanoacrylate
- Control / Measure Relevant Factors
 - Sex, Age,
 - Ambient Humidity
 - Positive, Negative, Blood, No-blood
 - 500 dpi vs. 1000 dpi enrollment images
 - Time between latent impression and latent acquisition

Reflect contents of CMF

Target forensically most common Surfaces

Target Best Practices

Possible Testing Goals

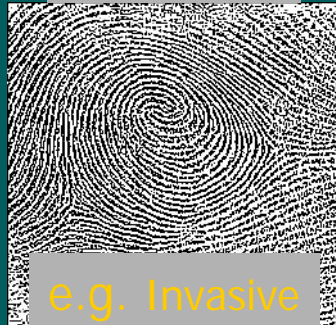
- Primary Goal: Measure latent search performance.
 - Measure this for various matchers
 - Measure it for various types of imagery
 - Metrics must quantify “misses” (reliability, FNMR) and false matches (incorrect members returned in a candidate list)
- Secondary:
 - If latent quality can be summarized numerically, then test extent to which quality algorithm is indicative of ultimate matching performance.
 - If automated minutiae annotation is practical, then which annotator improves matcher performance the most?

Possible Inputs to SDKs

Latent Image

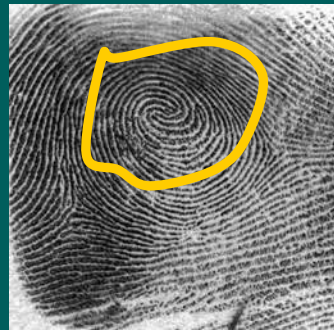


Bimodal Latent Image



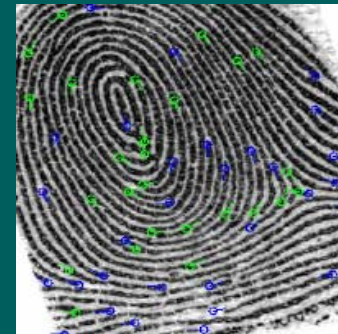
Multimodal

Latent Image + ROI



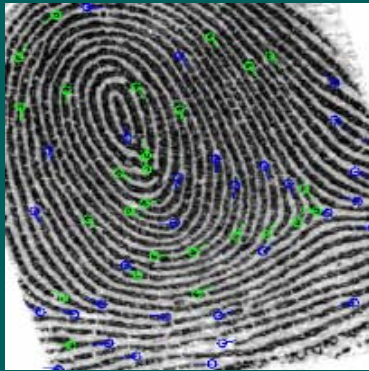
Operator assisted

Minutiae Annotated Latent Image



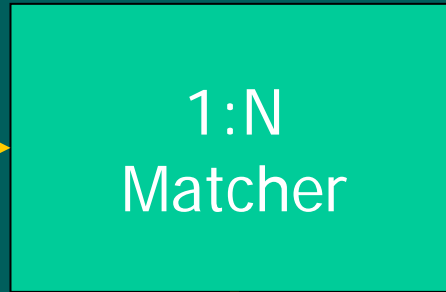
Matching

Enrollment Images



Latent Image
+ Minutiae

SDK might take:
Types 4 + 9, or
INCITS 381 + 378



1:N
Matcher

Candidate List



CMF
IDENT
W+W
ABIS

Conclusions

- SDK style testing is appropriate
 - Identification mode, not verification
 - Hardware searching would be more difficult
- Fully lights-out is fairest, most easily automated.
- Software-only test seems possible
- Improvement is possible with marking of ROIs or features is likely to be matcher-specific. Best performance and fairest if annotation is directed by matcher vendor.