International Biometric Group

Research • Consulting • Integration

Testing for Emerging Modalities

INTERNATIONAL BIOMETRIC PERFORMANCE CONFERENCE
3 March 2010 – NIST – Gaithersburg, MD

Michael Thieme
Director of Special Projects
International Biometric Group

Topics

- Longitudinal perspectives on performance testing for emerging biometrics
- Differences in testing emerging vs. established modalities
- Lessons learned

What Constitutes Emerging?

- A novel modality, or a modality with limited deployment / test history (e.g. keystroke dynamics)
- A technology that takes a substantially new approach to an established modality (e.g. ultrasonic fingerprint)
- Emerging is a matter of perspective
  - In 2000, iris recognition was commercialized, but still emerging
  - Palm vein and finger vein were considered emerging in North America in 2006, notwithstanding extensive deployment in Japan
  - Can be prototype, pre-commercial, or early-stage commercial
- Ambiguity on how to conduct performance testing
  - How should subjects use devices (do best practices exist)?
  - How is sample data obtained from the sensor / system?
  - How is the matcher implemented?
  - How are results analyzed?

Keystroke Dynamics

- **Net Nanny** BioPassword (2000)
  - Scenario evaluation, real-time matching
  - 200 subjects, ~2000 comparisons

Transactional	
FNMR	FMR
19.91%	1.30%

- **AuthenWare** AuthenTest (2009)
  - Scenario evaluation, real-time matching
  - Additional offline post-processing
  - 500 subjects, ~7000 comparisons

Transactional	
FNMR	FMR
3.20%	3.26%

- Observations
  - Needed regular access to a trained, controlled population
  - Ideally tested *in situ* due to emulate impact of keyboards, input devices
  - Traditional biometric terminology – samples, templates, comparison scores – not directly applicable to this technology
  - How to obtain quantity of signatures sufficient to build robust models?
  - Are trained typists more prone to false matching?
  - *Test results may substantially understate actual performance*

Palm Vein

- **Fujitsu** PalmSecure (2006)
  - Hybrid scenario/technology evaluation
  - Online capture, offline matching
  - ~650 subjects, 2 positions,
    ~22k samples, ~50m comparisons

Transactional	FNMR	FMR
Same-Visit	0.57%	0.056%
Cross-Visit	0.69%	0.063%

- Observations
  - Vendor needed tuning samples from subjects of specific ethnicities
  - Female error rates were an order of magnitude higher than male
  - Housing / cradle was a prototype designed for standalone testing
  - Most difficulties related to capture instructions
  - Offline matching took weeks – matcher never implemented for volume
  - Lack of granular thresholds reduced visibility into performance
  - *While considered emerging, technology was in many ways mature*

Finger Vein

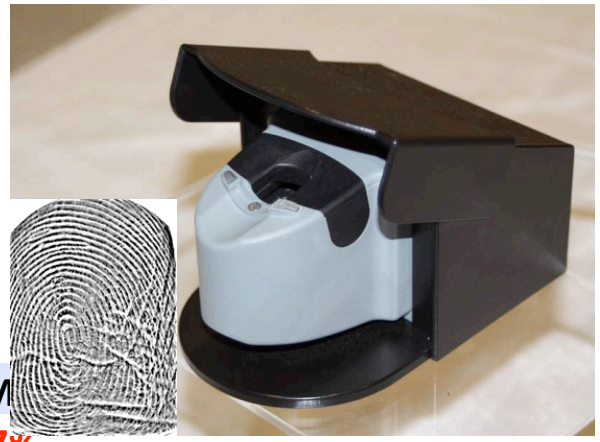
- **Hitachi** UBReader TS-E3F1 (2006)
  - Hybrid scenario/technology evaluation
  - Online capture, offline matching
  - ~650 subjects, 2 positions,
    ~22k samples, ~25m comparisons

Transactional	FNMR @ 0.10% FMR	FNMR @ 0.01% FMR
Same-Visit	0.34%	0.68%
Cross-Visit	1.94%	2.77%

- Observations
  - Vendor had to port matching capability from on-card to server-based
  - Vendor modified native recognition sample capture to acquire a 14-15 second image stream for each capture attempt
  - Capture logic relies on multi-pass quality assessments, such that typical capture behavior for genuine and impostor captures different
  - Tester implemented "Better-instance" matching logic
  - *Creative problem solving often required to test emerging modalities*

Contactless Fingerprint

- **TST Biometrics** BiRD 3 (2009)
  - Hybrid scenario/technology evaluation
  - Online capture, offline matching
  - ~500 subjects, 6 positions, ~37k samples, ~32m comparisons

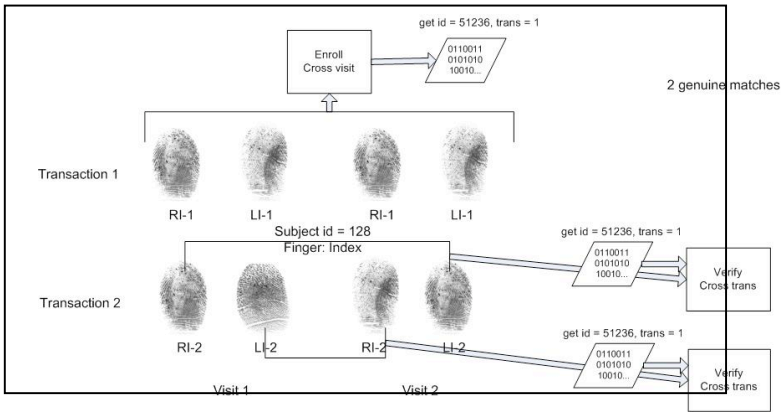


Transactional	FNMR @ 0.10% FMR	FNMR @ 0.01% FM
Same-Visit	0.067%	0.067%
Cross-Visit	0.372%	0.661%

- Observations
  - Example of an emerging approach to an established modality
  - Vendor used results to support engineering decisions (e.g. default thresholds, interoperability with contact sensors)
  - Presentation duration considerably longer than for contact devices
  - Interoperability with contact systems essential
  - *When appropriate, emulate governing principles from mainstream modalities*

Fingerprints with Crypto Keys

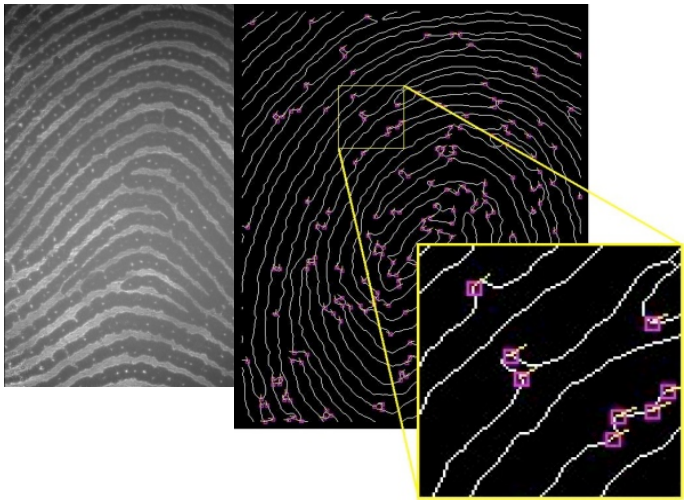
- **GenKey** (2007)
  - Technology evaluation
  - Offline matching (and a separate online test)
  - ~1200 subjects, ~20k samples, ~4.5m comparisons
  - Numerous exploratory tests conducted



- Observations
  - Testing intended to assess the viability of (paper) token issuance from a workflow perspective; matching secondary to the concept
  - Experimentation with different thresholds necessary to get into the vicinity of sensible match scores
  - Vendor primarily interested from a marketing perspective
  - *Testing was meant to validate what the vendor already knew*

3700dpi Fingerprint (Pores, Ridge Contour)

- **Aprilis** Holosensor (2005-6)
  - Hybrid scenario/technology evaluation
  - Online capture, offline matching
  - ~650 subjects, 4 positions,
    ~9k samples, ~32m comparisons
  - EER ~20%



- Observations
  - Company disintegrated, more or less, over the course of the evaluation
  - Thus, no matching technology (used NBIS Bozorth) and no support
  - Technology on the extreme side of prototype-to-commercial spectrum
  - Had to develop best practices for presentation on the fly
  - In retrospect, we needed a way to evaluate pure imaging capabilities (e.g. fidelity) without reliance on matching
  - *Sometimes modalities are emerging for a reason*

General Observations

- Testing typically assumes a consultative aspect
- Vendors expect feedback on how to improve their technology
  - Tester may be identifying and solving problems that the vendor has not seen or anticipated
    - Bugs in development software, libraries
- Results may be generated solely for the vendor
- While test approaches should reflect relevant best practices, flexibility and creativity may be required to accommodate novel technology aspects
- Allow for trial and error, more dry run testing, further exploration of parameters that may impact quality