

Research Challenges in Large Scale Biometrics Systems

Test & Evaluation



Michael Garris (mgarris@nist.gov)
Biometrics Senior Scientist
Information Technology Laboratory
National Institute of Standards and Technology

<http://biometrics.nist.gov>

Categories

- Data, Data, Data
- Experimental Design & Methodology
- Analysis and Interpretation of Results

Data, Data, Data

- Quantity
 - Statistically significant and operational relevant
 - Ground truthed identities
- Character
 - Representative Samples
 - Source, quality, unbiased?
 - Controlled Samples
 - Single factor, multi-factor, mixed effects?
 - Difficult to scale
- Development vs. Sequestered
 - Developers may not have operationally collected samples
 - Public releasable data often from a different source

Experimental Design & Methodology

(1 of 3)

- Operations vs. Robustness
 - Performance on average vs. how far till it “breaks”?
 - Fundamentally different sample data populations
 - Expectations must be managed
 - Clearly communicated, designed, and reported
- Operational vs. Offline
 - Offline Testing - can assess false rejects and false positives
 - Operational Testing – cannot automatically assess misses
- Acceptance vs. Comparison
 - Evaluating within one system?
 - Comparing and contrasting across systems?

Experimental Design & Methodology

(2 of 3)

- Methodology for Longitudinal Studies
 - Critically Relevant
 - Our biometric systems are “maturing”
 - Definition and Hypothesis Testing
 - pattern + behavior + environment?
 - Model Alternatives
 - Each with assumptions and limitations

Experimental Design & Methodology

(3 of 3)

- Hardware vs. Software
 - Hardware
 - Pro: Exact compute platform/architecture
 - Con: Logistically complex and expensive
 - Software
 - Pro: More flexible and less expensive
 - Con: Common (non-optimized) hardware platform
- Open Testing API and Framework
 - Control and Specificity
 - For fair and unbiased comparisons
 - Abstraction
 - To not stifle ingenuity , innovation, and operational relevance
 - Specialized Computation & Functionality
 - Typical: create template, enroll subject, search gallery, ...
 - + Bulk enrollment management - parallelized gallery CRUD tools
 - + Logging, auditing, and graceful error recovery

Analysis and Interpretation of Results

- Measuring Uncertainty
 - Significance Tests and Confidence Intervals
 - Alternatives: Parametric vs. Bootstrap?
- Projecting Biometric Performance
 - Conduct one expensive well-designed large-scale one-to-many test
 - Characterize the sample population along with the resulting genuine and imposter distributions
 - Characterize a new sample population
 - Project/predict performance on the new sample population without conducting a new test