Overview of biometric quality

Elham Tabassi – NIST

Biometric Consortium 2009

23 September 2009

outline

- Why measure quality?
- What is meant by quality?
- What are they good for?
- How to assess quality of quality?
- What are the challenges in quality computation?

Why measure quality?

Push towards zero-error biometrics

- While recognition technologies are good at what they are being used, or contemplated for use, their performance drop in difficult operational scenarios and with imperfectly controlled data.
- Although only a small fraction of input data are of poor-quality, the bulk of recognition errors can be attributed to poor-quality samples.
- Improving quality either by sensor design, by user interface design, or by standards compliance, better performance can be realized.
- For those aspects of quality that cannot be designed-in, an ability to analyze the quality of a live sample is needed.

What is meant by quality?

Predictive of performance

Quality problem: “The Last 1%”

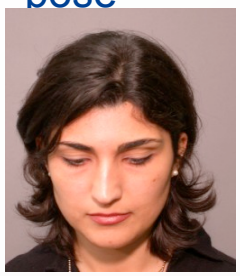
Or maybe “The Last 0.1% or 10%”

» Fraction of samples that should not be sent to the matcher

- finger, iris scanners have been designed specifically for the task, face cameras (mostly) have not
- providing constructive feedback only possible if cause of poor quality is known

character, e.g. scar

behavior, e.g pose



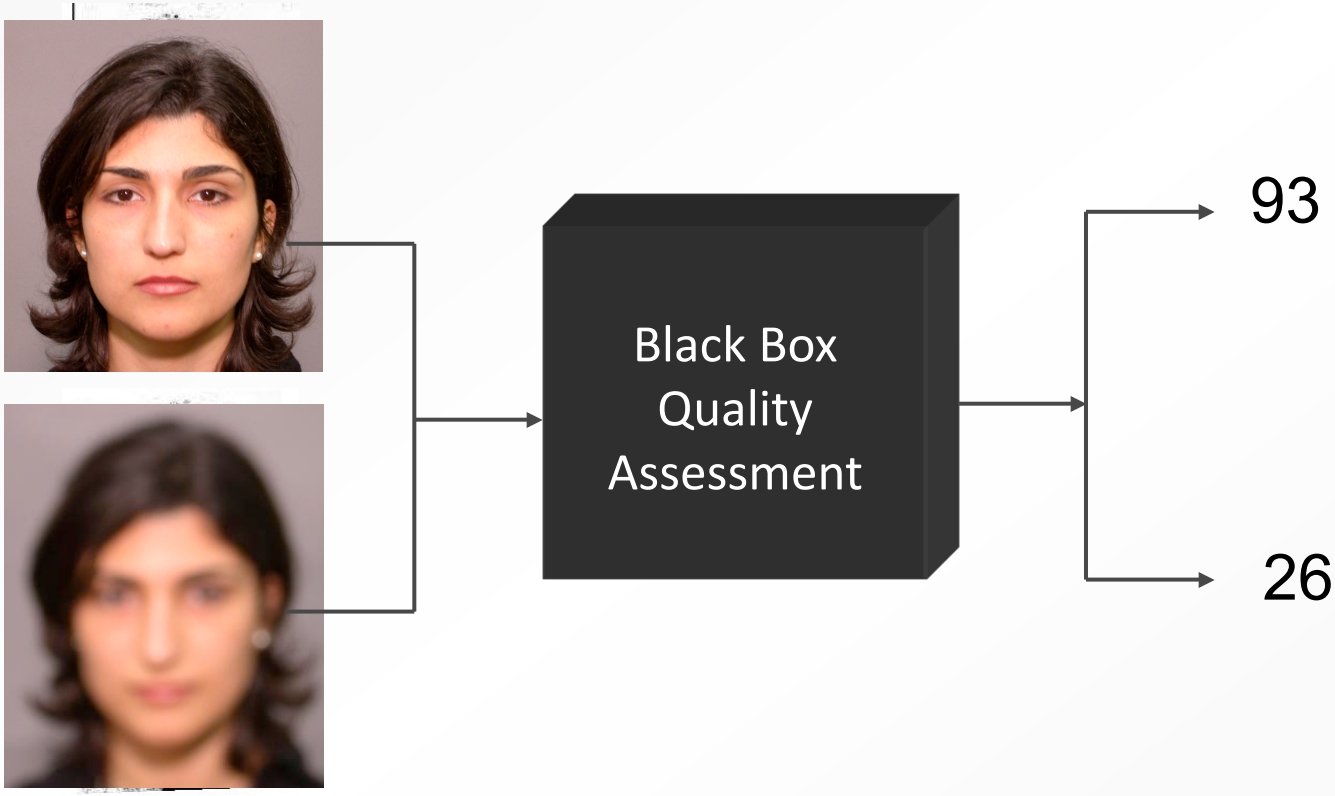
environment, e.g. shadows

imaging, e.g. focus

What are the uses of quality?

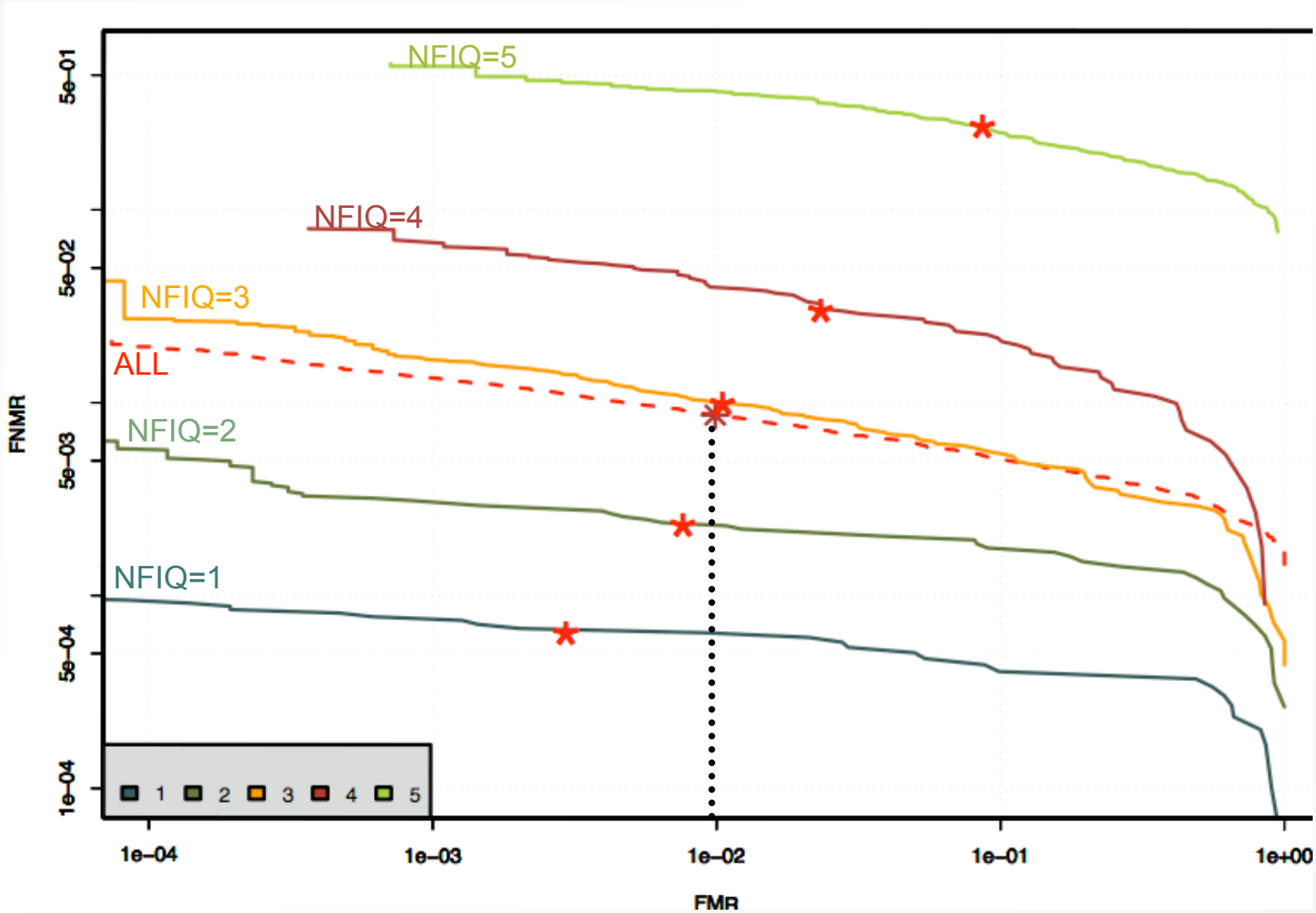
- Initiating the reacquisition from a user
  - Enrollment
    - for credential issuance (visa, passport, access card, PIV)
    - pruning the poorest quality samples (1.65% of dataset) reduced EER from .0047 t0 .0024
  - Verification
    - of the samples just captured which one to send for matching?
    - or acquire still more?
  - Identification
    - is the subject offering a poor sample deliberately?
- Selective invocation of different processing methods
  - Preprocessing phase
    - image restoration algorithms (e.g., contrast adjustment) or a different feature extraction
  - Matching phase
    - invoke a slower but more powerful matching algorithm when low-quality samples are compared
    - sending poor quality (NFIQ=4,5) to a more accurate (but perhaps costly) matcher reduced FNMR from 0.0136 to 0.0078 at FMR=0.001
  - Decision phase
    - quality directed fusion, dynamic threshold
    - performing quality based multi-algorithm contingent likelihood ratio fusion reduced FNMR from 0.0136 to 0.0068 at FMR=0.001
- Sample replacement/Template update
    - negate template aging
- Quality monitoring
    - are some biometric field locations giving low quality?
    - only in the evening?

Predictive of performance



A biometric quality assessment method derives a numerical quality value from an input biometric sample. The quality value is related to the biometric error rates that are likely to be realized when the sample is matched.

Quality: rank statistic for performance



NFIQ::(fmr,fnmr) at fixed threshold

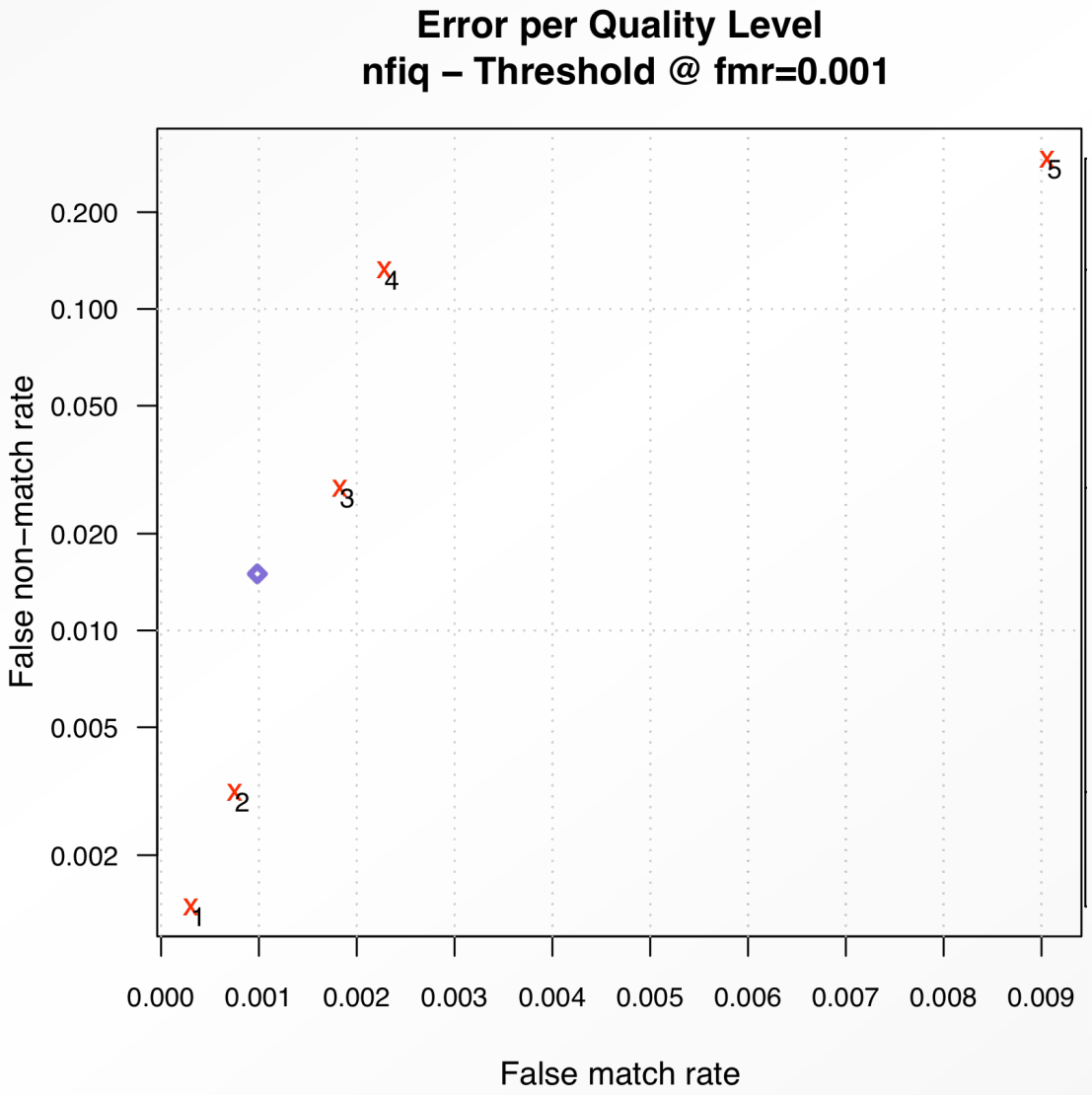
NIST fingerprint image quality

- NIST developed NFIQ in 2004
  - Open source
- Key innovation: quality as a rank statistic for performance
- NFIQ is a machine learning algorithm
  - Exploratory variables: image properties (minutiae, ridge density and clarity)
  - Response variable: separation of genuine and impostor comparison

NFIQ

- **feature extraction**: computes appropriate signal or image fidelity characteristics and results in an 11-dimensional feature vector.
- **neural network**:classifies feature vectors into five classes of quality based on various quantiles of the normalized match score distribution.
- **quality number**: an integer value between 1(highest) and 5 (poorest).

Error rates per NFIQ level



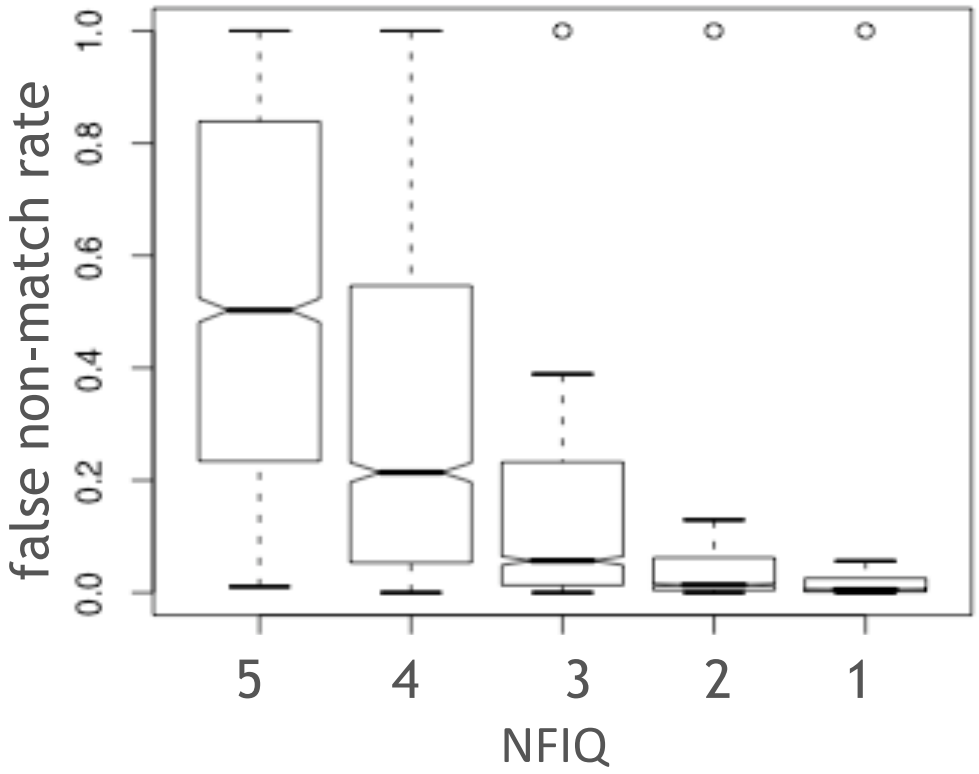
Quality of quality

- Biometric quality shall be evaluated based on its ability to predict performance.
  - e.g. error vs. reject curves
- Comparison of quality algorithms shall compare their effectiveness in predicting performance.

Quality challenges

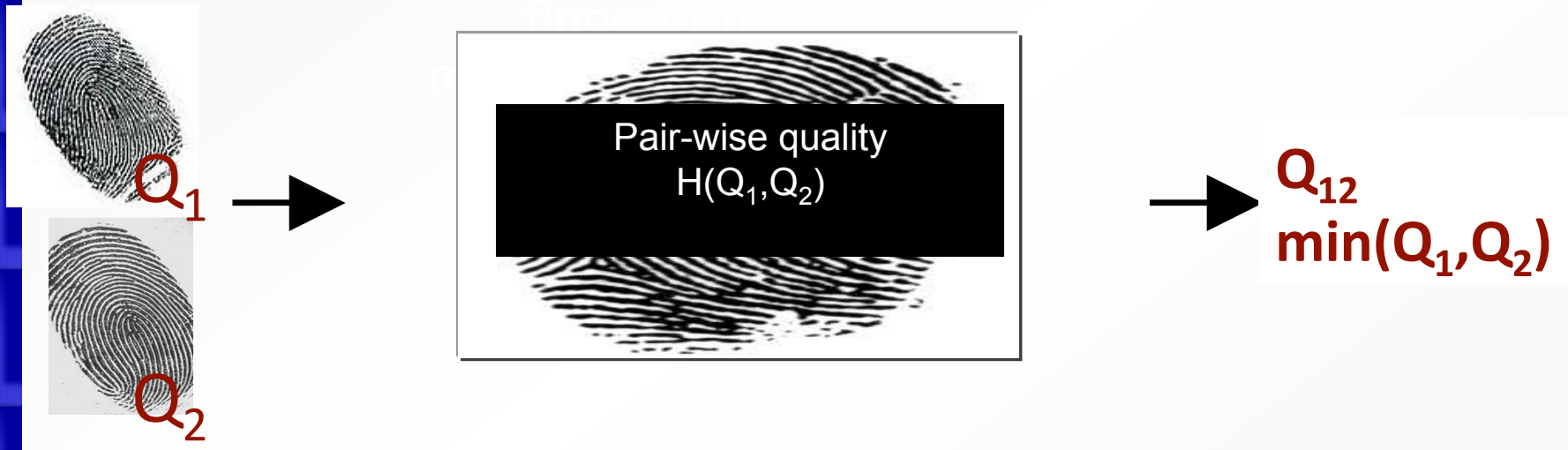
- Scalar vs vector
- Matcher dependency
- How many levels?
- Pair-wise quality
- Calibration

How many levels?



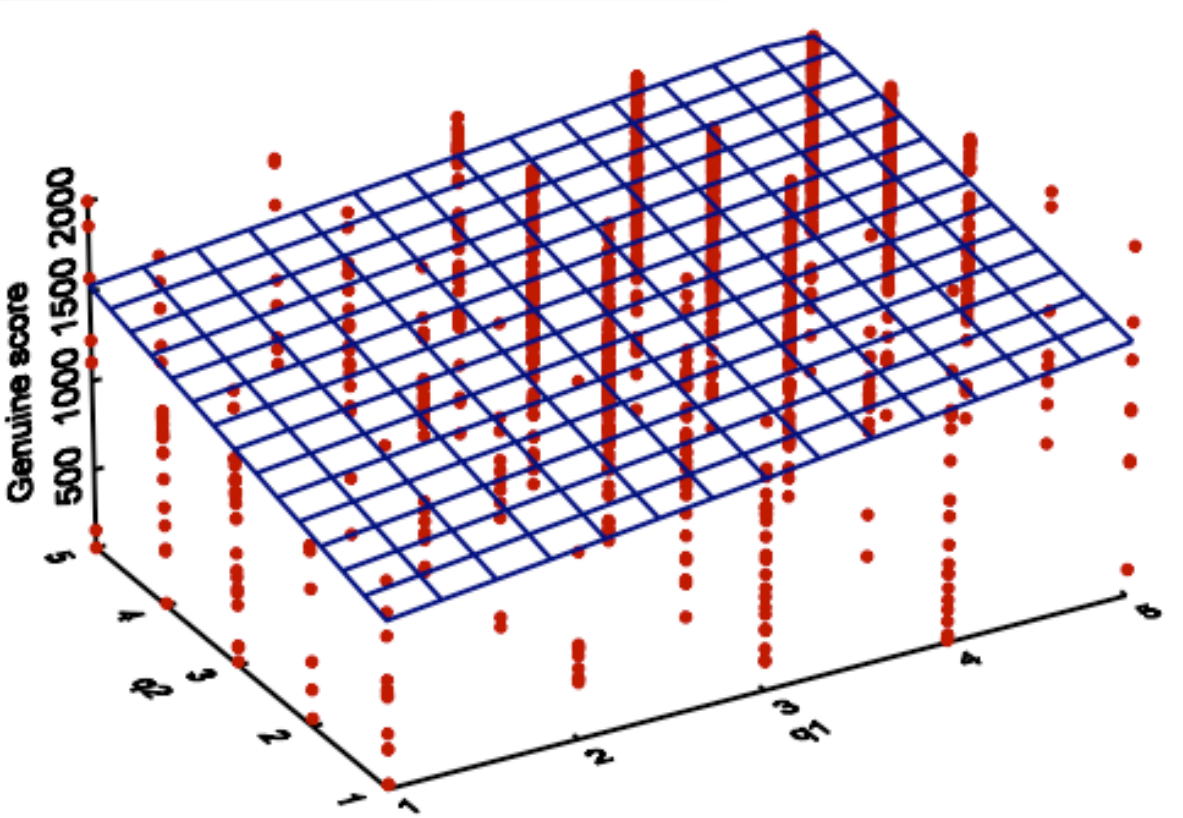
Statistically different level of performance

Pair-wise quality - 1



when the enrollment sample is of good quality and better than that of the authentication sample, the authentication sample's quality is sufficient to predict performance.

pair-wise quality



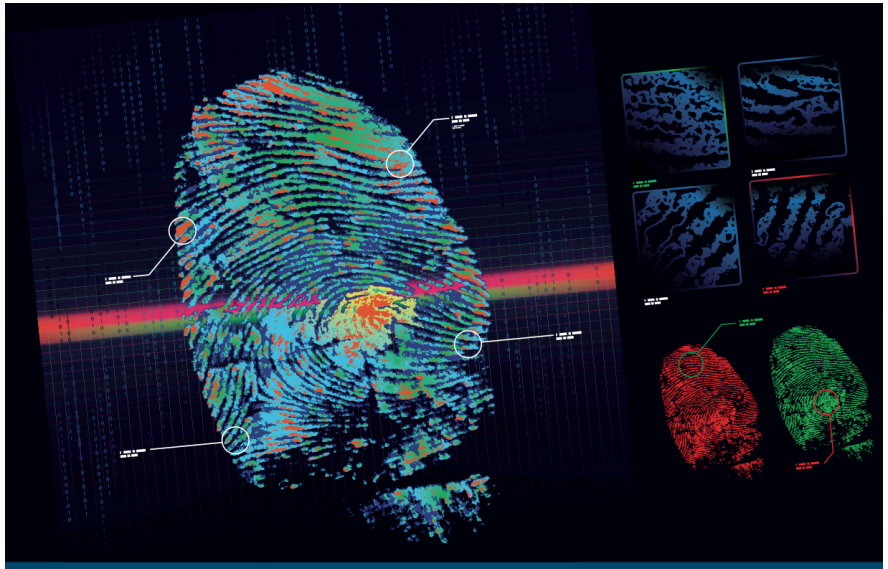
Quality

WHAT AFFECTS QUALITY (IMAGE LEVEL)

- devise metrics for quantifying specific aspects of quality to
  - distinguish cause of poor quality (provide feedback)
  - Introduce tolerance (quality-by-design)
- design framework for assessing effect of quality
  - sensitivity analysis to each metric
  - statistical method for combining effects (neural net, svm, etc)
- perform analysis on large datasets of images
  - preferably on subset of images with specific defect (focus, pose, …)

WHAT IS THE EFFECT OF QUALITY (APPLICATION LEVEL)

- devise metrics for quantifying the dependence of the accuracy of the core algorithms on
  - the quality of biometric samples (error vs reject), and
  - systematic quality variation (quality summarization procedure)
- design framework for assessing effect of quality on accuracy and security
  - how quality of enrolled samples affects probability of false accept?
  - How about probability of false reject?
- perform analysis on large datasets of images
  - examine the role of quality in biometric zoo



Biometric Quality

The last 1% Biometric Quality Assessment for Error Suppression

September 2009

Homeland Security

Science and Technology

Thank You

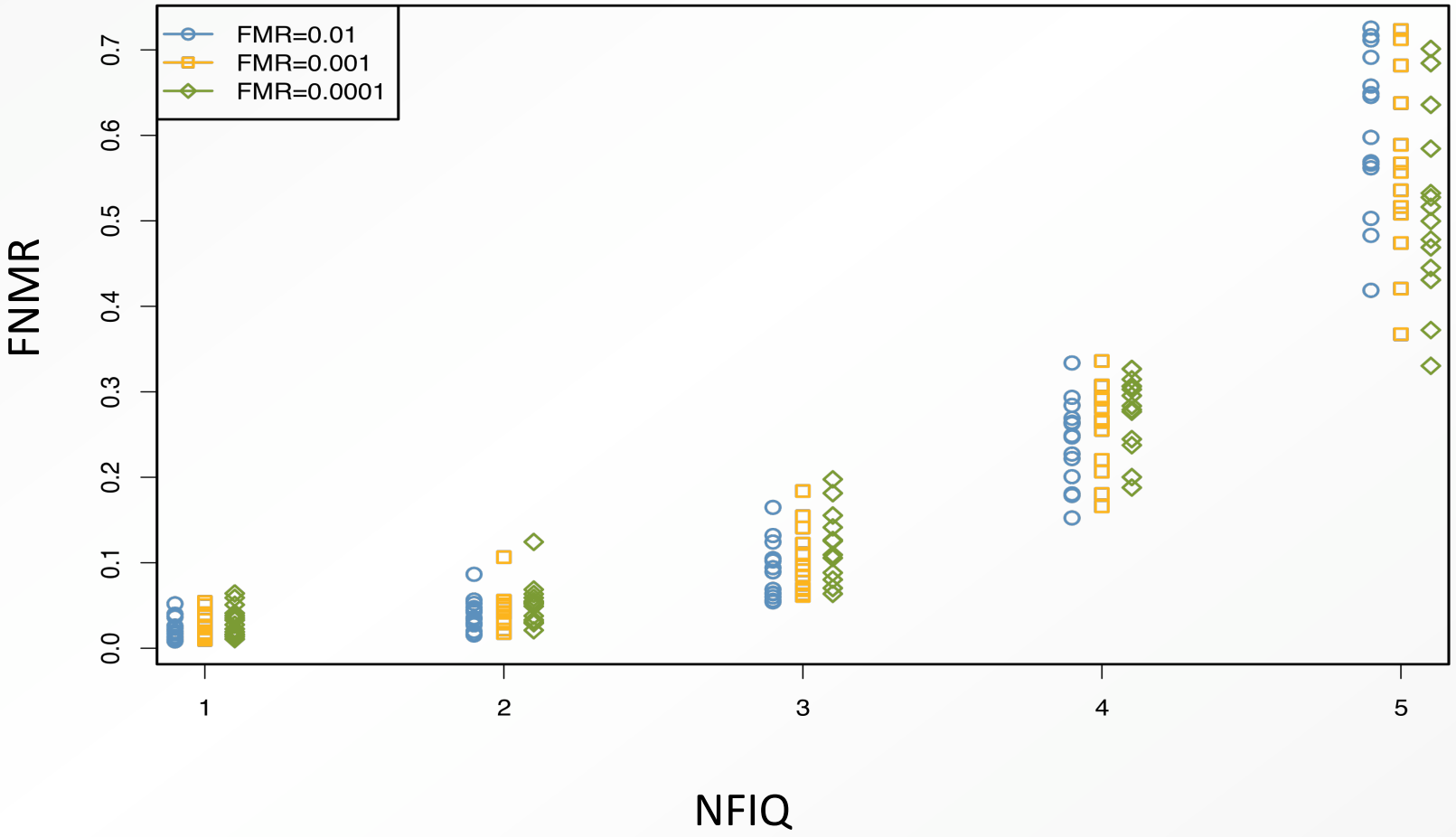
Elham Tabassi

tabassi@nist.gov

301 975 5292

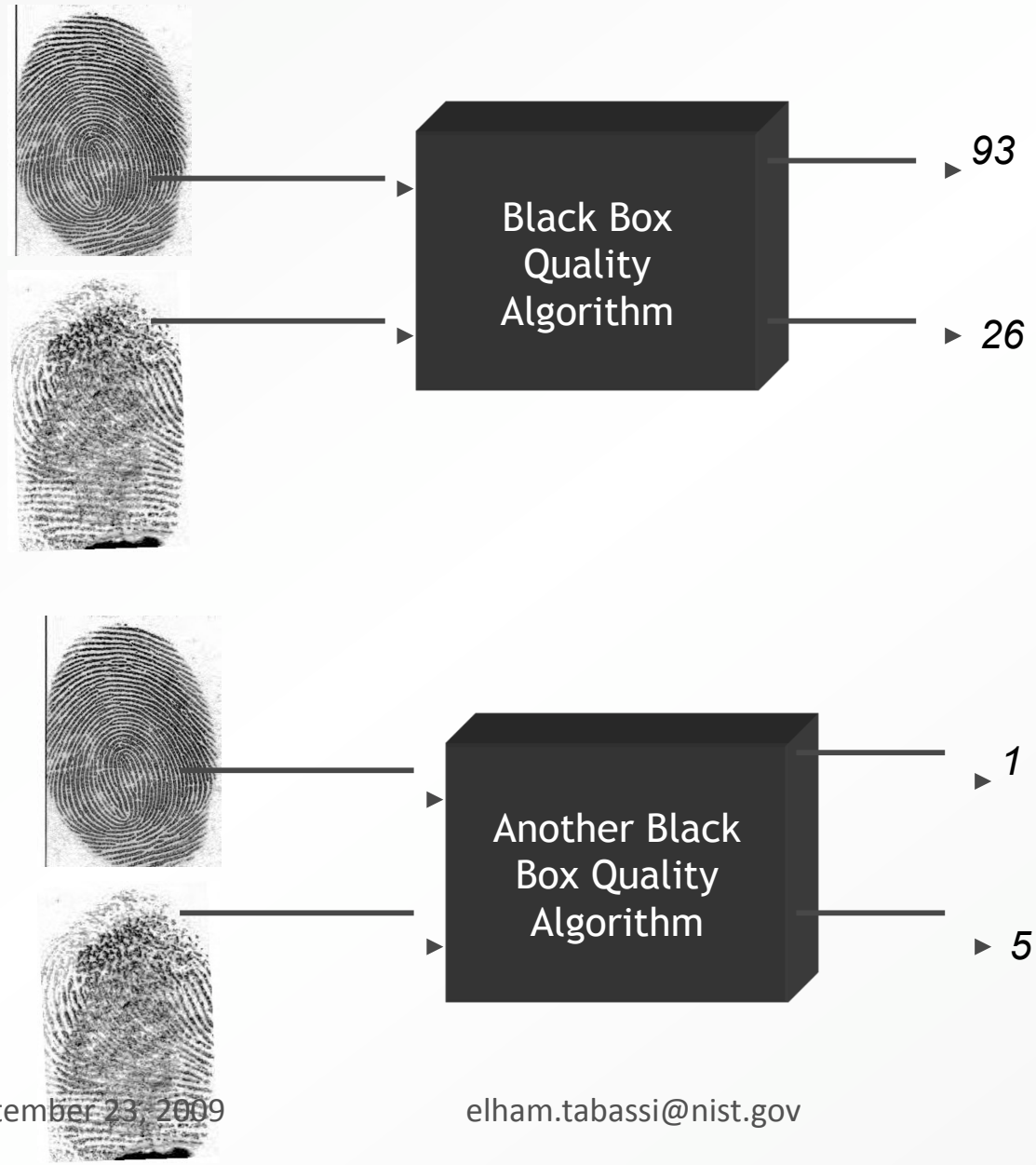
www.itl.nist.gov/iad/894.03/quality

Dependence on matcher

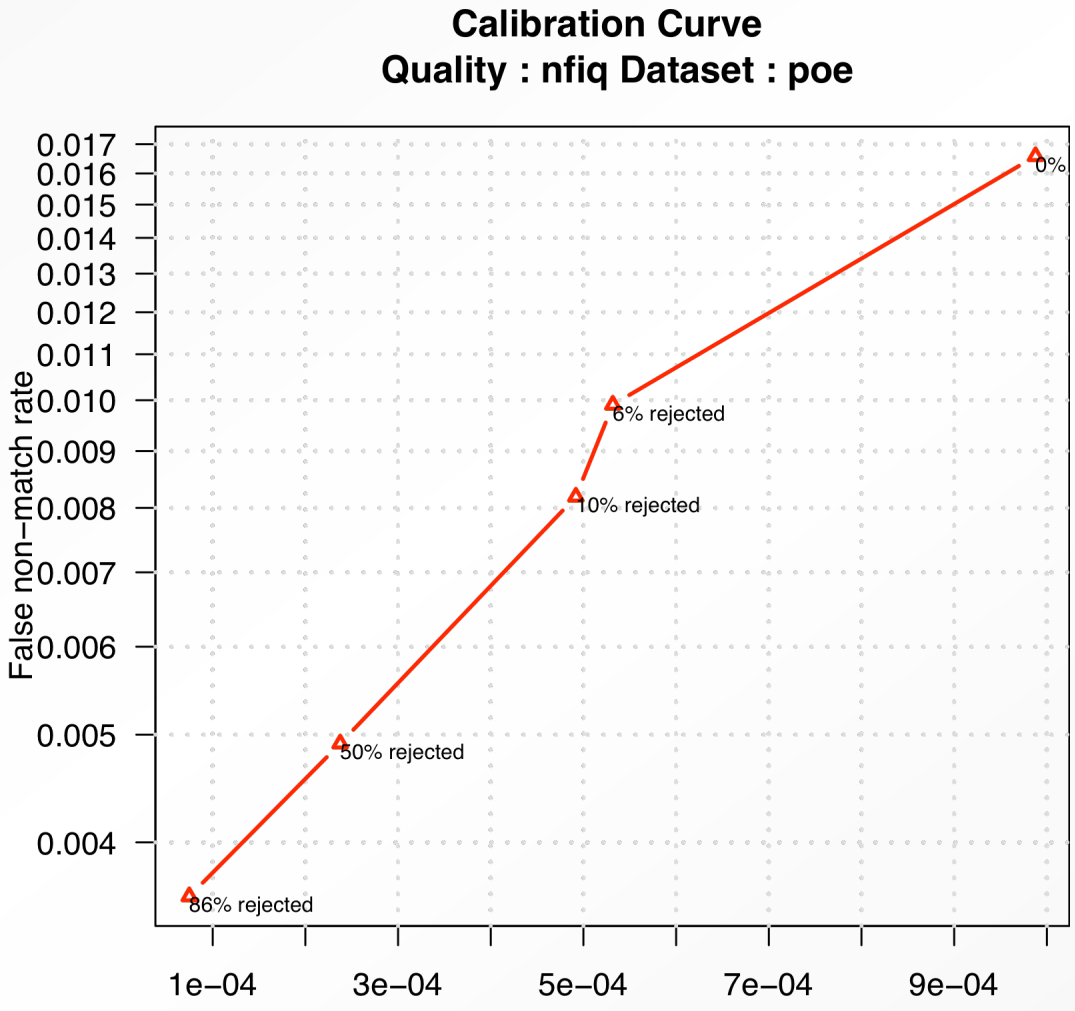


Each point corresponds to one algorithm.

Quality score interpretation



Calibration Curve: Error vs reject : NFIQ



elham.tabassi@nist.gov