Observations from ICE 2006 Quality Data

Prof. Patrick J. Flynn
University of of Notre Dame (ND)
(NIST Associate, 2007-2008)

Dr. P. Jonathon Phillips
NIST

Outline

- Motivation for this prospective analysis
- Quality measure effects on performance
- Quality measure correlation
- Subjective versus objective quality
- Conclusions and comments

Motivation

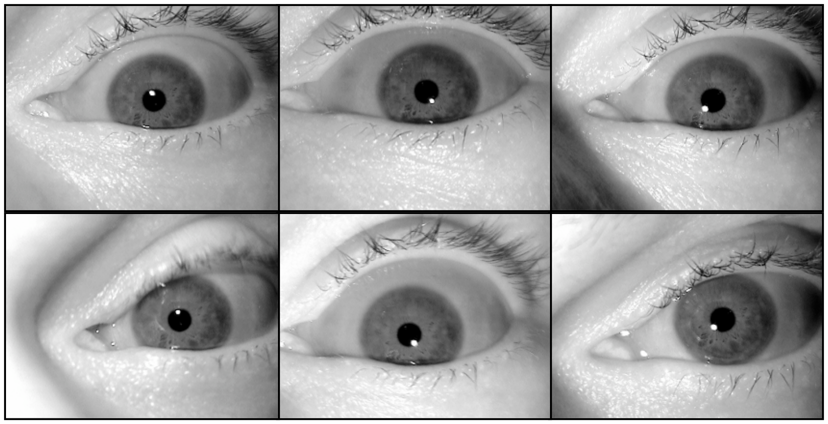
- Iris image acquisition typically expects highly controlled environment
  - Cooperative subject (minimize iris occlusion)
  - Lighting
  - Active focusing

- Strong texture contrast & focus yield subjective “good quality”
  - Strong texture filter responses
  - Reliable phase estimates

Iris Quality in the Literature

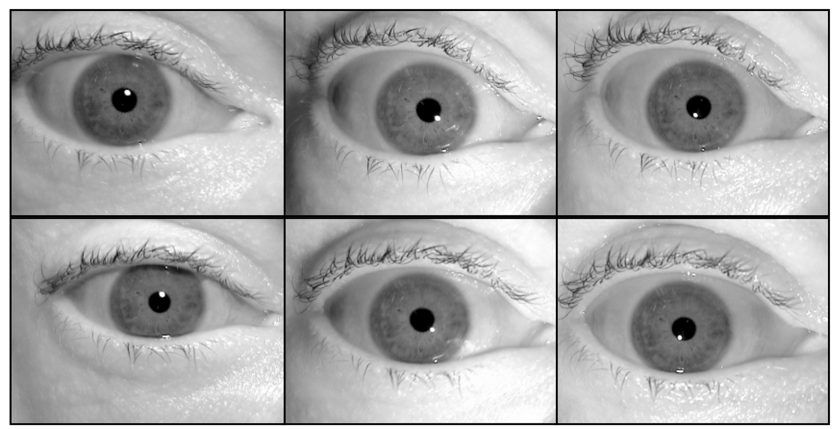
- General biometric sample quality concepts
  - Fidelity *vs.* application-specific criteria for quality
  - Methodology for quality based performance analysis (Grother and Tabassi, PAMI 2007)
  - Subject and sensor effects on quality
- Common iris recognition quality aspects
  - Focus, spectral content
  - Occlusion (e.g., % iris)
  - Averted gaze
  - Motion during acquire interval
  - Wasserman 2006 (sensor quality), Kalka 2005, Dass 2006, Valencia 2007

Sample ICE 2006 iris subject session



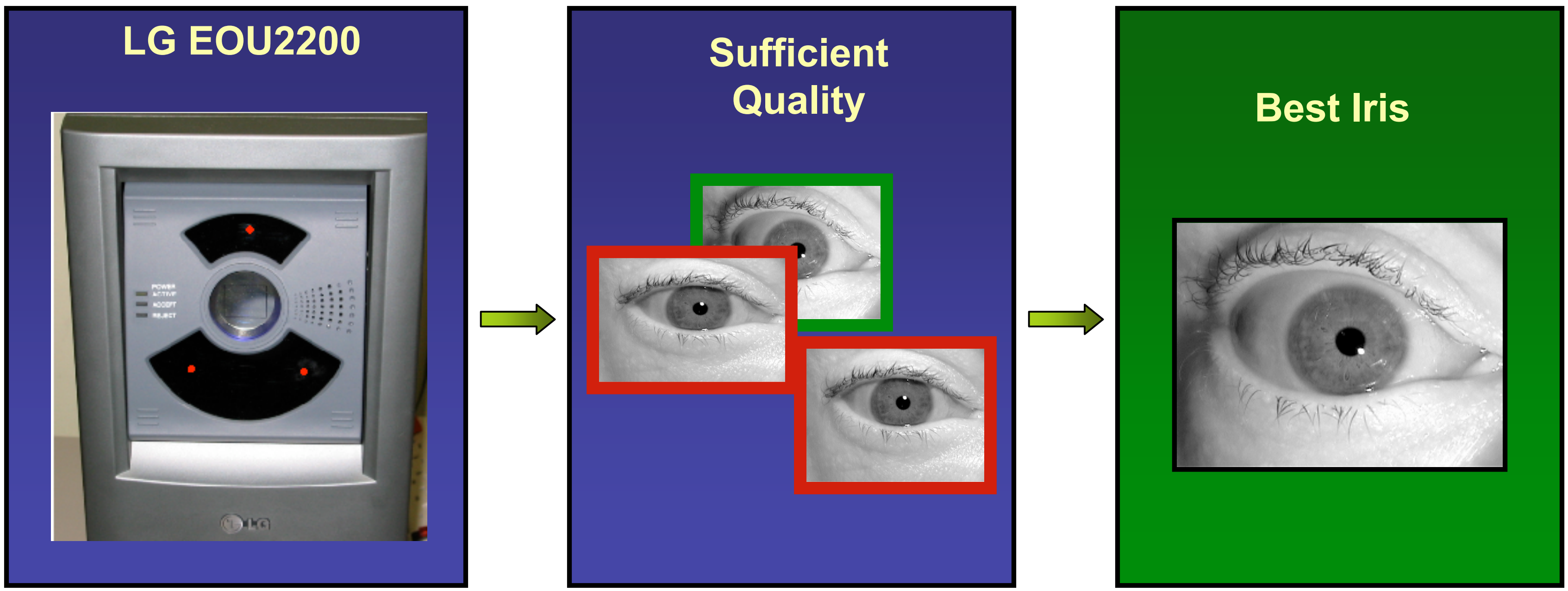
Left Eye
LG EOU 2200

Right Eye
LG EOU 2200



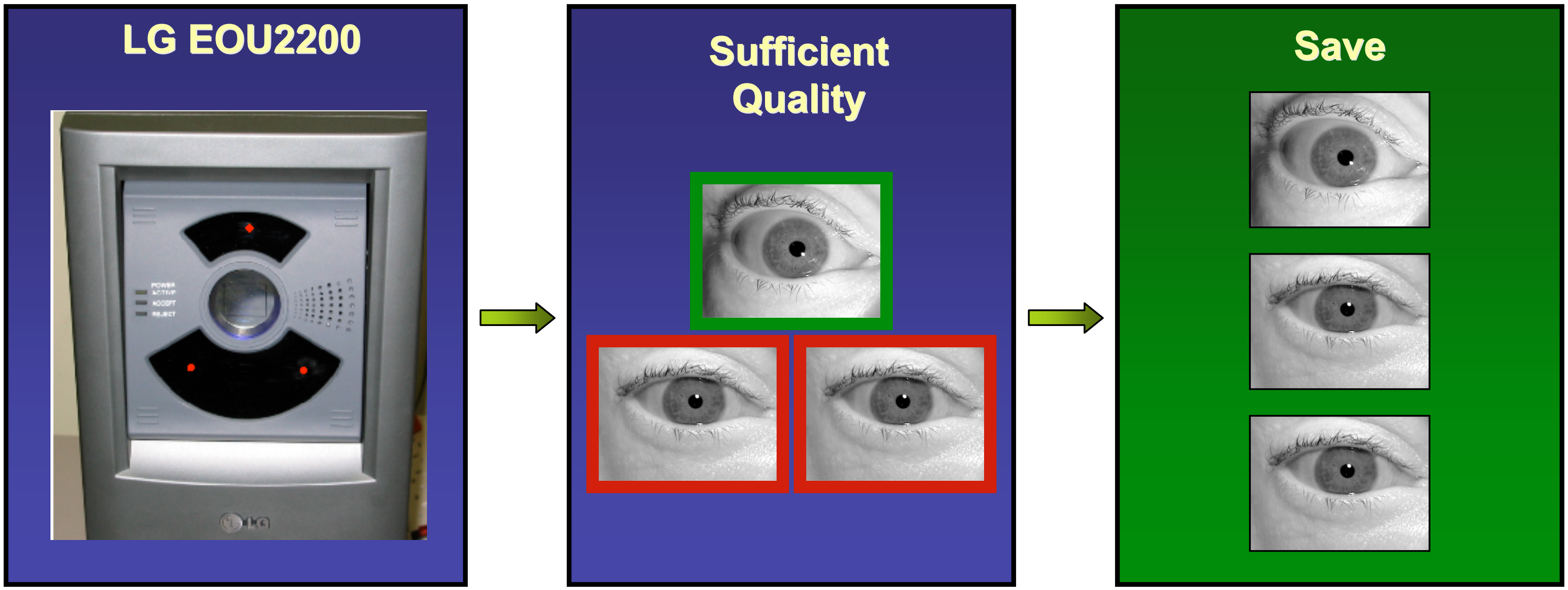
LG EOU 2200 was industry recommended at the inception of data collection.

Standard operation of Iris Imaging System



- Take 3 iris images
- Find iris of sufficient quality
- Select best if more than one
- Or retake

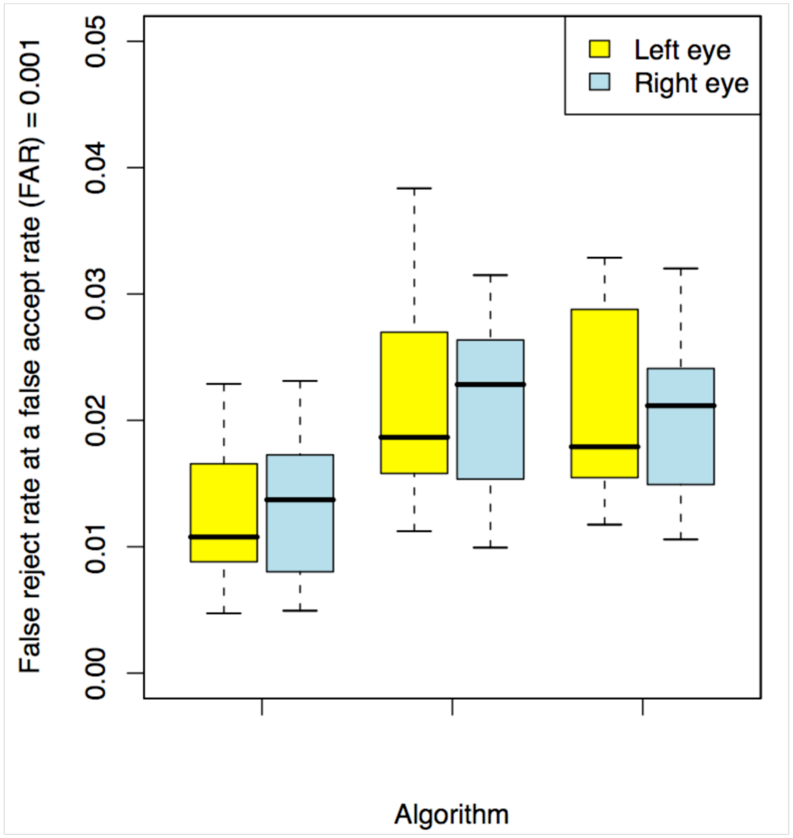
Iris image acquisition test software (CRADA between Iridian and Notre Dame)



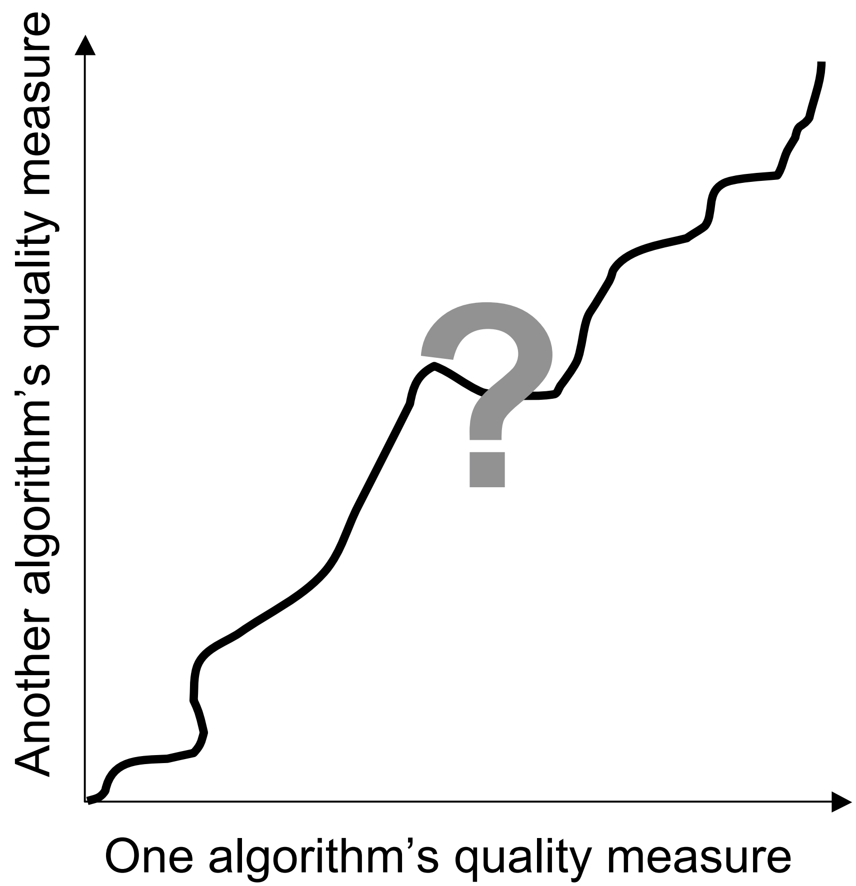
- Take 3 iris images
- One above quality threshold
- Save all three

ICE2006 Quality data

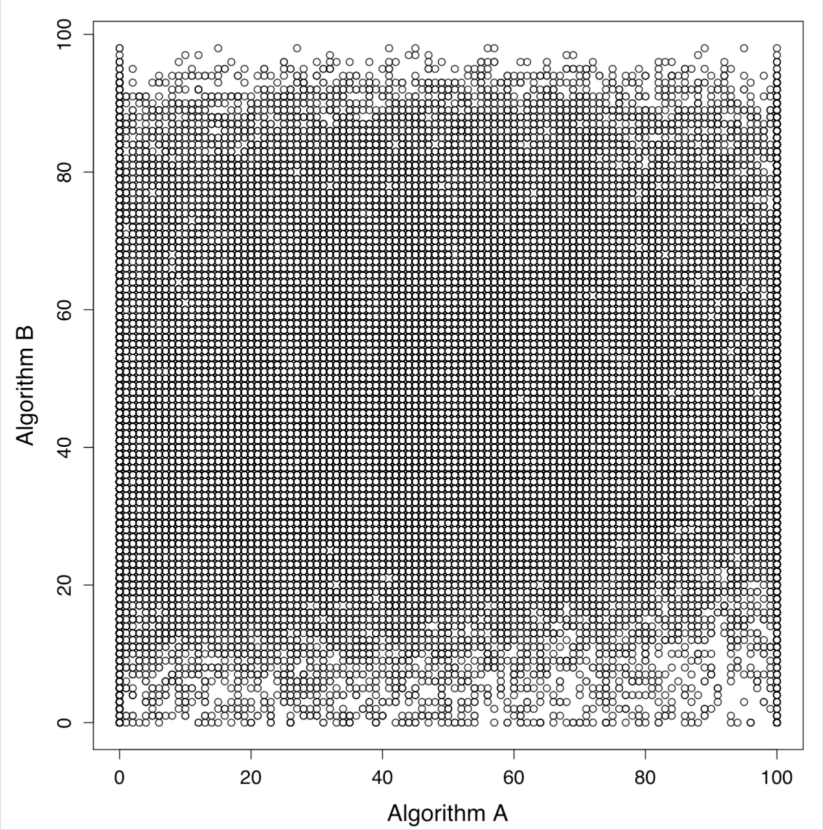
- Three competitive ICE 2006 performers (labeled "A", "B" and "C")
- 59558 iris images
- Each image has three quality scores (Q_A, Q_B, Q_C)

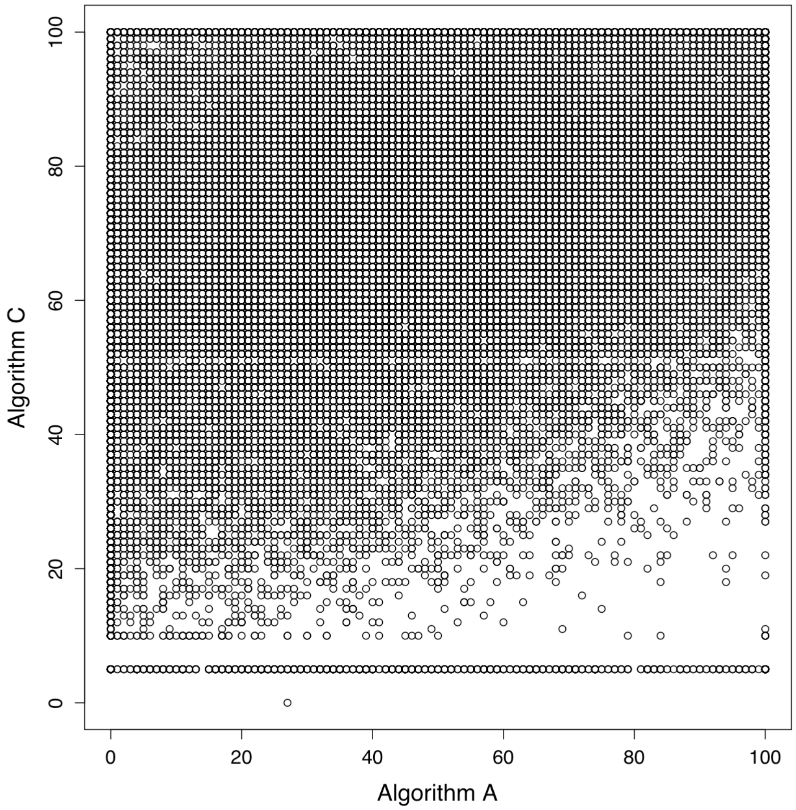


Should quality measures produced by different algorithms be correlated?



Quality measure scatter plots





Correlation of Quality scores table

Algorithms	Pearson's r	Spearman's ρ
A vs. B	0.122	0.131
A vs. C	0.349	0.348
B vs. C	0.120	0.108

Subjective interpretations of good and poor quality

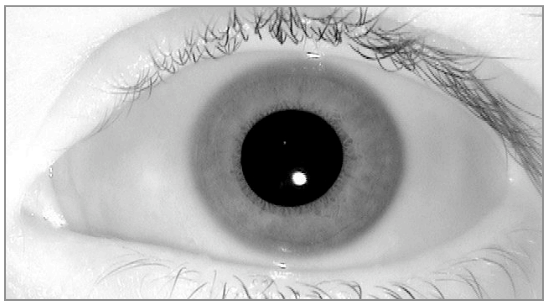
- Animations run from “good” to “poor” quality
- One image from each distinct fused quality value (captioned by average rank)

Videos

Quality effects on matching performance

Step 1: Compute Image Quality

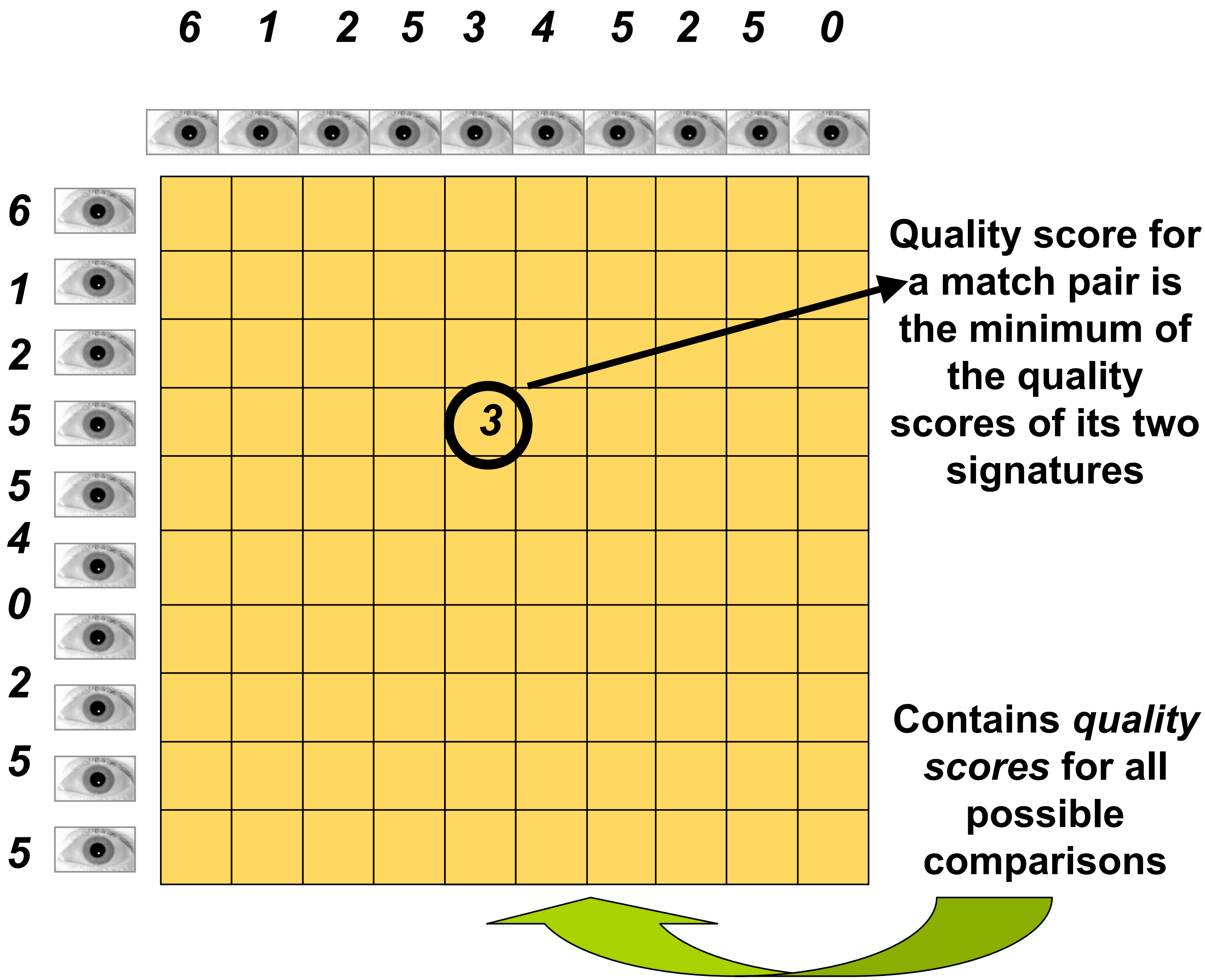
Input Iris Image



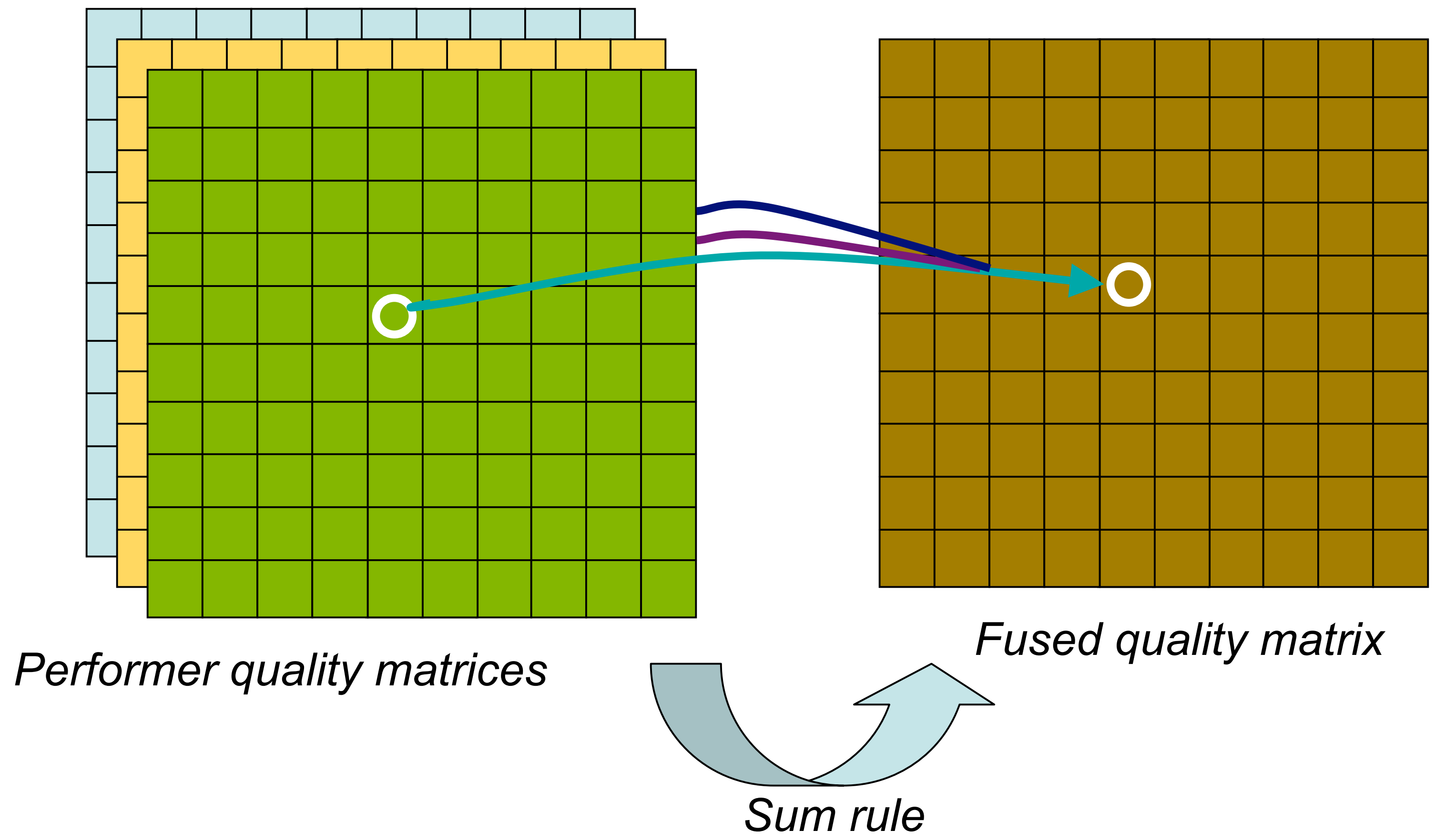
Performer's Image Quality Module

Integer {0..100}

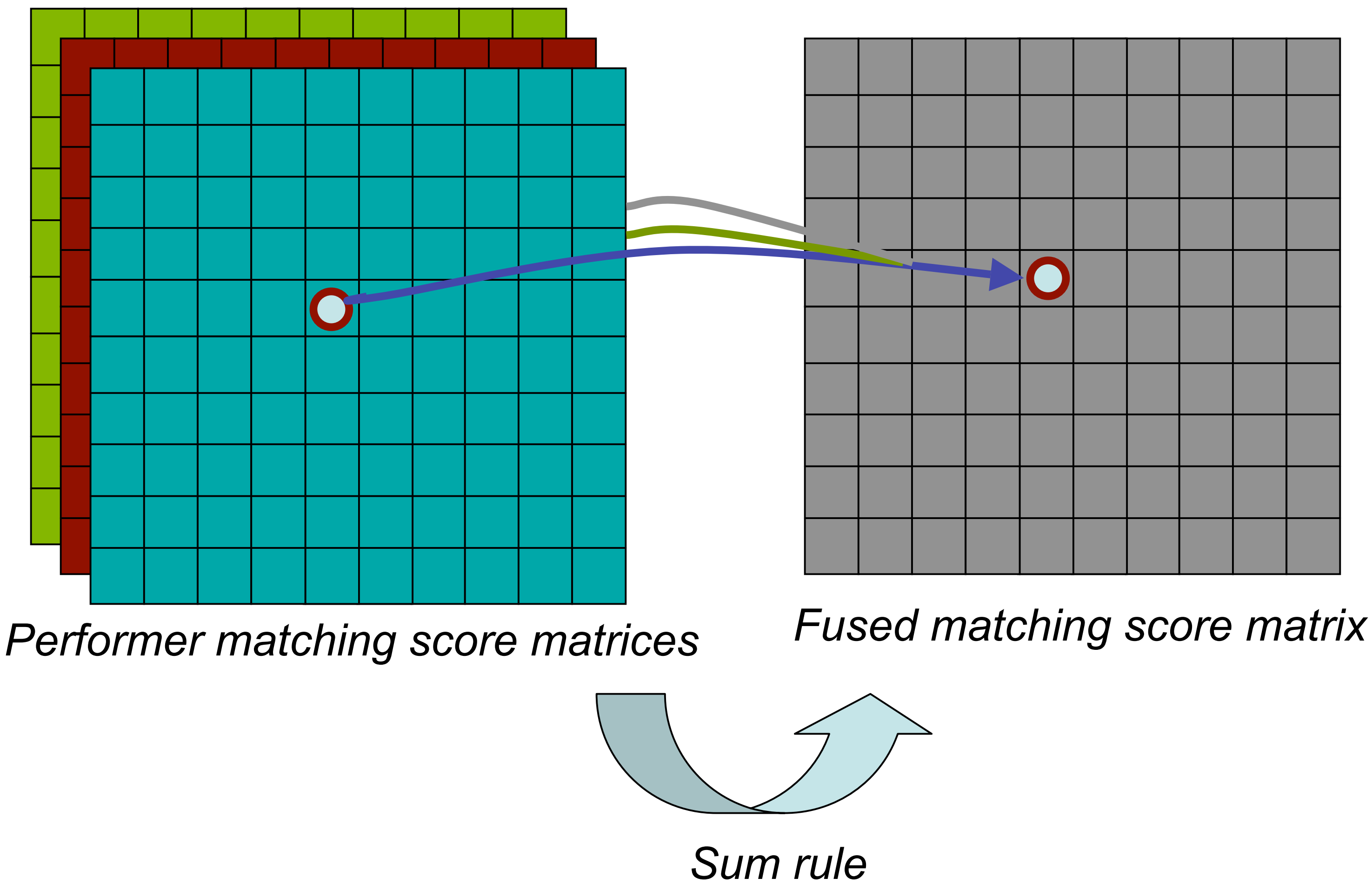
Step 2: Compute Quality Matrix



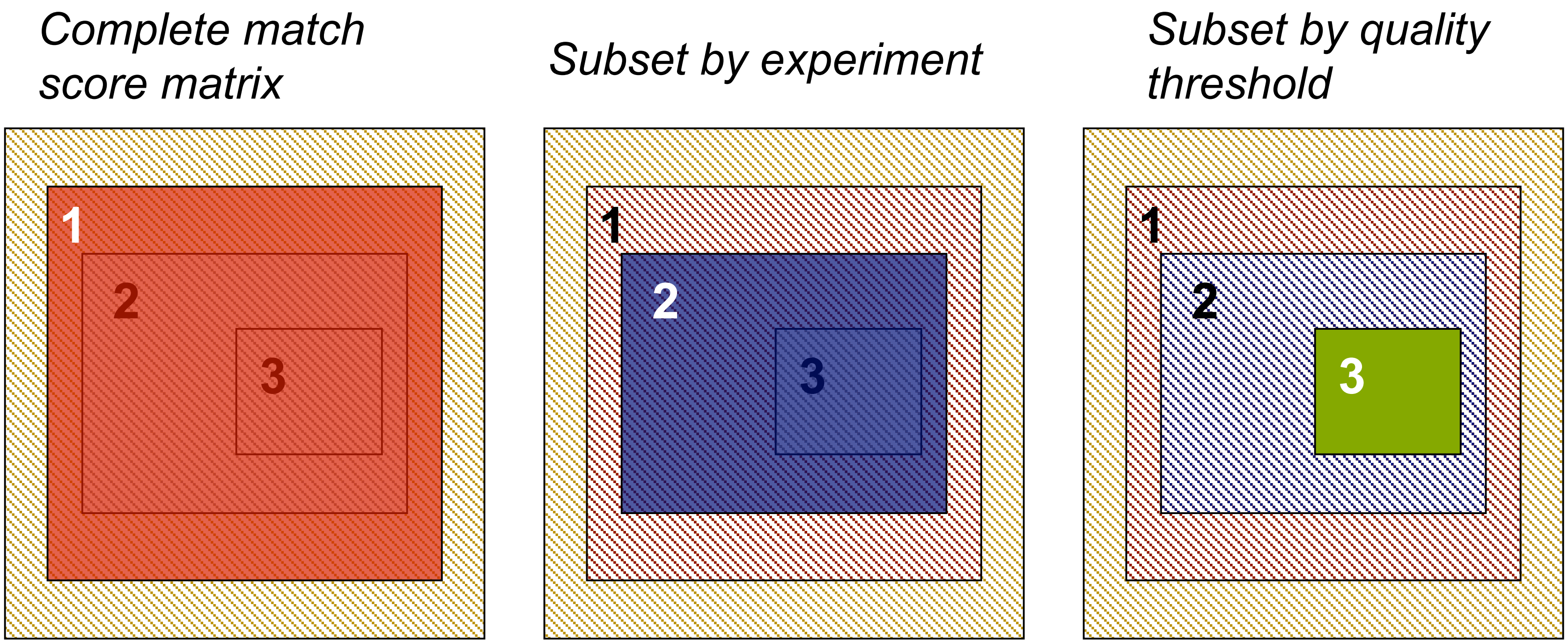
Step 3: Fuse quality matrices



Step 4: Fuse matching score matrices



Step 5: Prune matching scores by quality

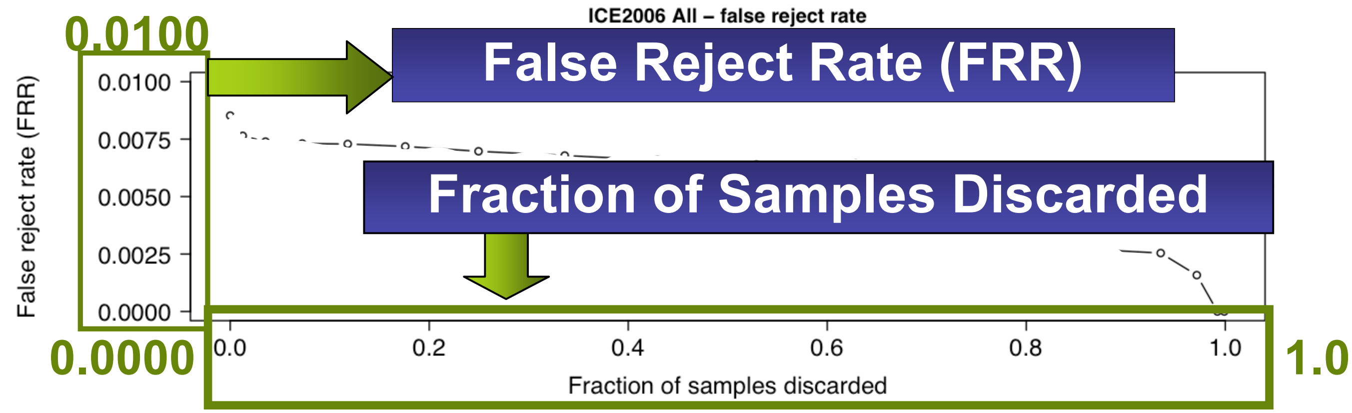


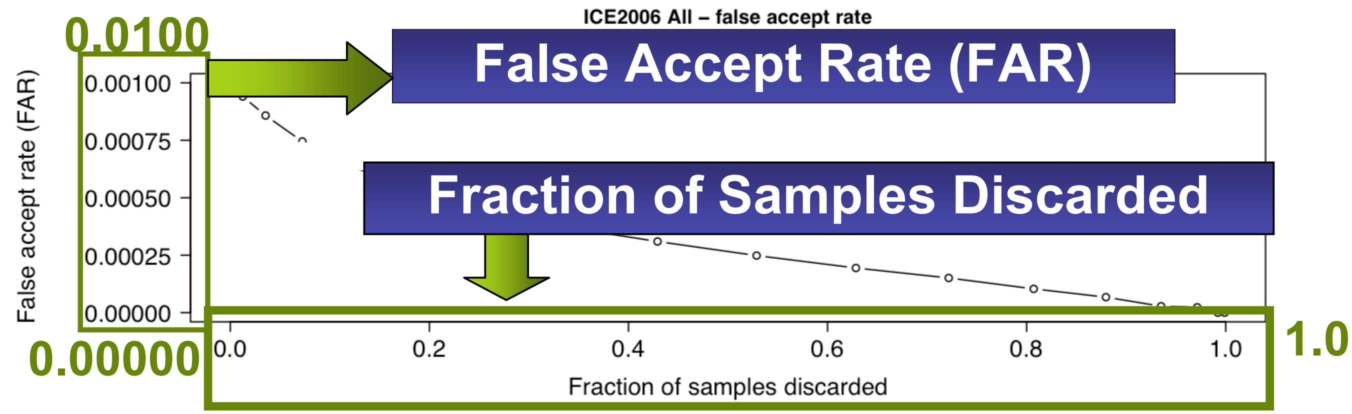
Fused quality threshold values: 5, 10, … 90, 95, 100

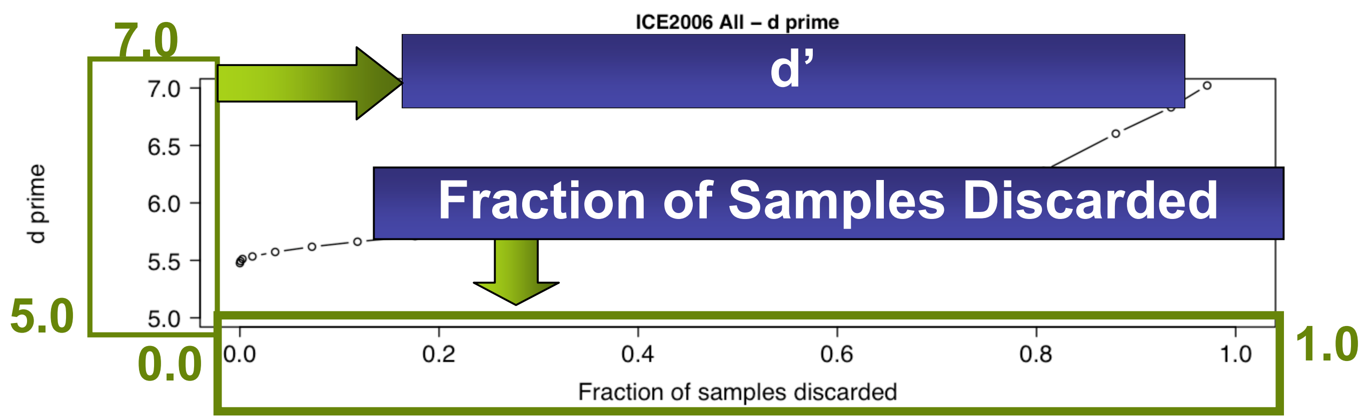
- 20 sub-experiments with nested sets of matching scores)
- Compute FAR, FRR from fixed threshold

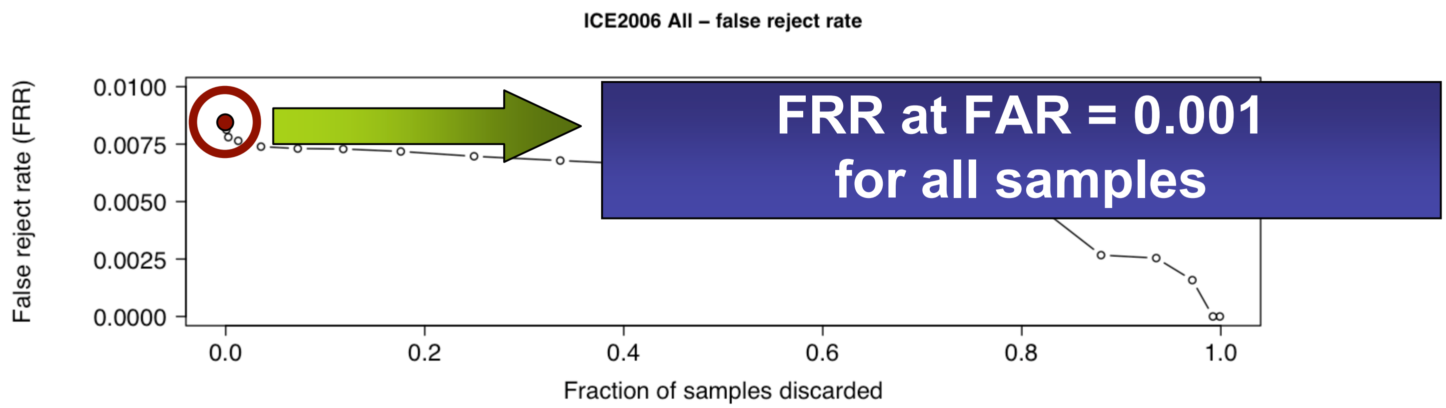
Calculation of FAR, FRR and d'

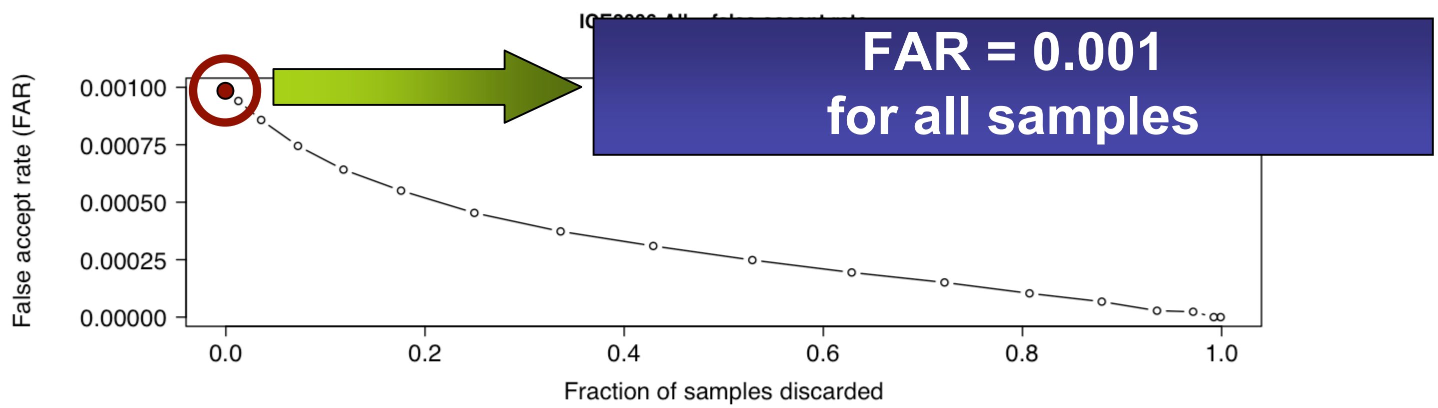
- From unpruned set, identify threshold T that yields FAR = 0.001 (ICE 2006 operating point)
- Let $Q_F(g)$ and $Q_F(p)$ be the qualities of gallery and probe samples g and p
- At a fixed quality point q, calculate FAR, FRR, and d' from all match pairs (g', p') with $\min\{Q_F(g'),Q_F(p')\} >= q$

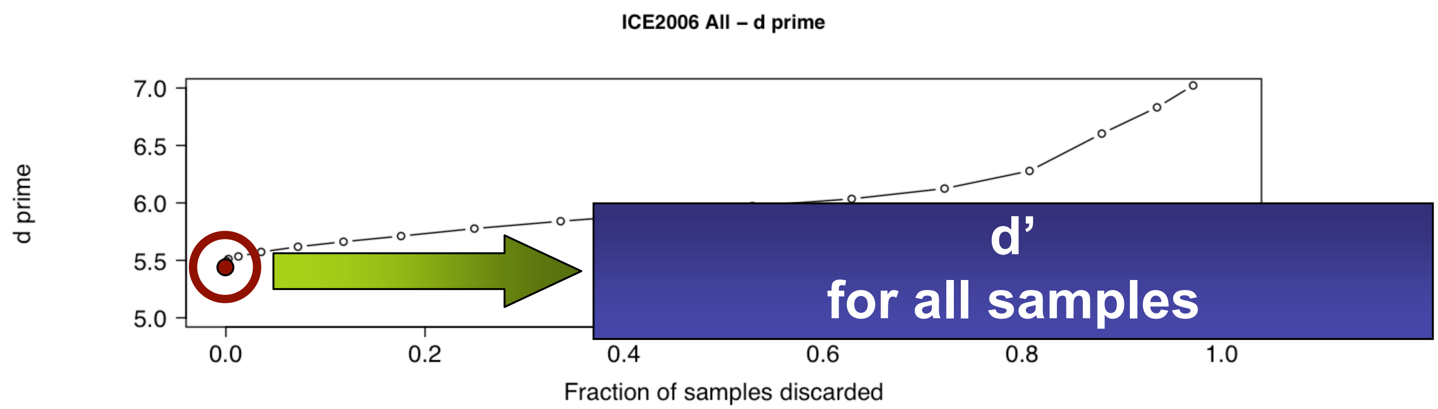

ICE2006 All – false reject rate
False Reject Rate (FRR)
Fraction of Samples Discarded
0.0100
0.0000
1.0
False reject rate (FRR)
0.0100
0.0075
0.0050
0.0025
0.0000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded

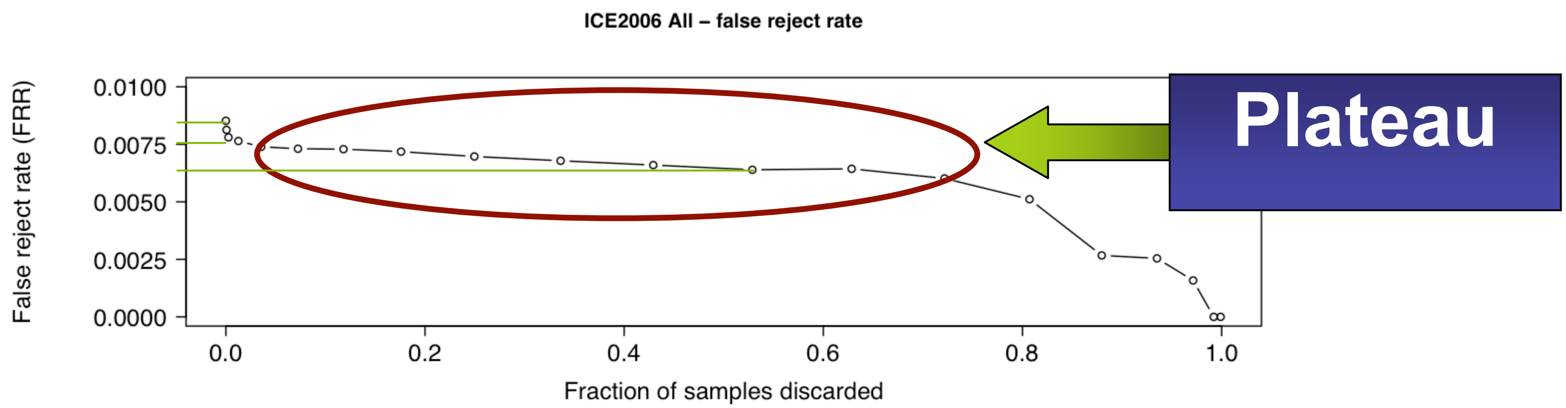

ICE2006 All – false accept rate
False Accept Rate (FAR)
Fraction of Samples Discarded
0.0100
0.00000
1.0
False accept rate (FAR)
0.00100
0.00075
0.00050
0.00025
0.00000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded


ICE2006 All – d prime
d’
Fraction of Samples Discarded
7.0
5.0
0.0
1.0
d prime
7.0
6.5
6.0
5.5
5.0
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded

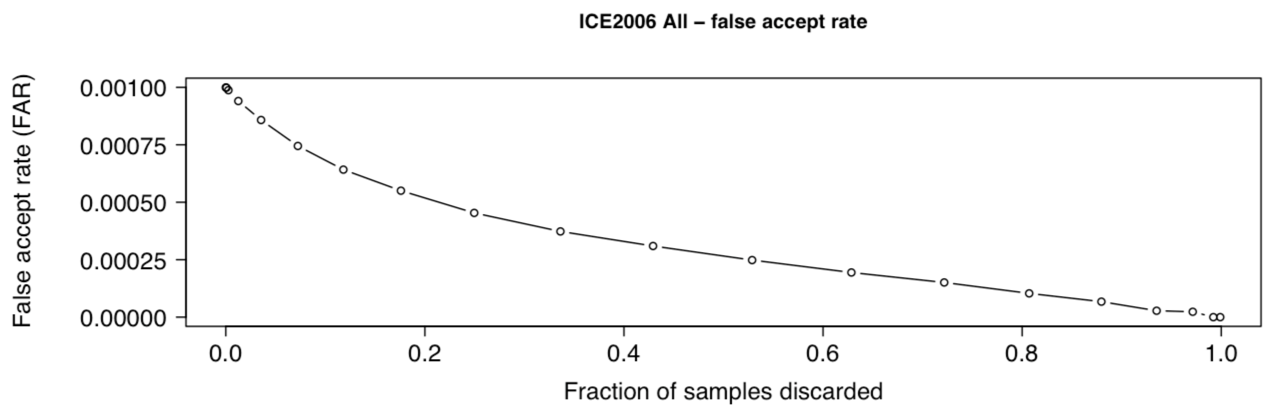

ICE2006 All – false reject rate
FRR at FAR = 0.001
for all samples
False reject rate (FRR)
0.0100
0.0075
0.0050
0.0025
0.0000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded


FAR = 0.001
for all samples
False accept rate (FAR)
0.00100
0.00075
0.00050
0.00025
0.00000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded

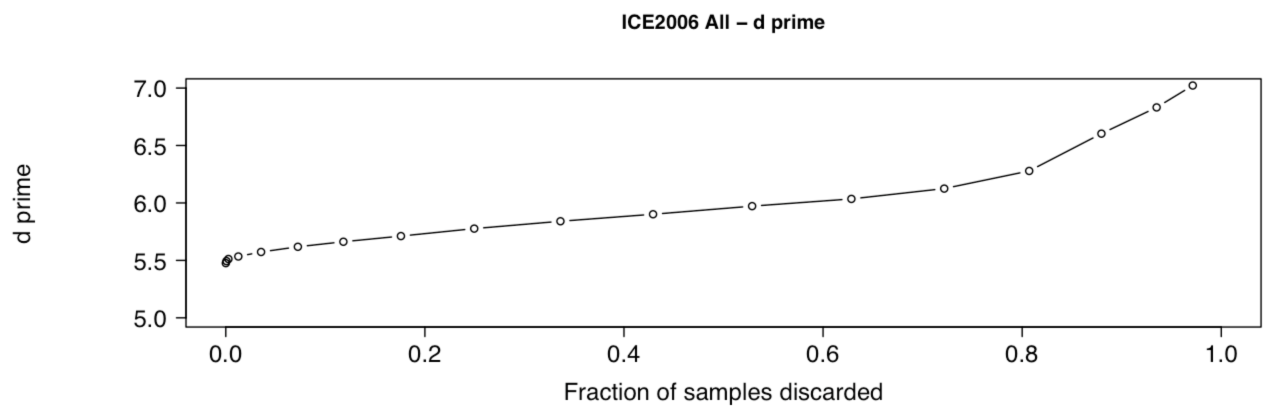

ICE2006 All – d prime
d'
for all samples
d prime
7.0
6.5
6.0
5.5
5.0
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded



ICE2006 All – false reject rate
False reject rate (FRR)
0.0100
0.0075
0.0050
0.0025
0.0000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded
Plateau



ICE2006 All – false accept rate
False accept rate (FAR)
0.00100
0.00075
0.00050
0.00025
0.00000
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded



ICE2006 All – d prime
d prime
7.0
6.5
6.0
5.5
5.0
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of samples discarded

Observations and Conclusions

- Prospective study (quality measures were required as part of the protocol)
- Iris image quality affects performance (general trends, from aggregated ICE2006 performance data)
  - d’ improves with restrictive pruning of samples by quality
  - FAR decays with quality @ fixed FRR
  - FRR nearly invariant for a range of quality ranks after an initial drop, at fixed FAR

Conclusions (contd.)

- Iris image quality measurement needs more research and thorough testing
  - Lack of correlation between three ICE2006 responders suggests that they were measuring different aspects of quality, or measuring them with different degrees of accuracy
  - Opportunities:
    - for further research
    - fusion

Acknowledgment

- Patrick Flynn thanks ND and NIST for sabbatical leave support

- ND biometrics research is supported by:
  – IARPA (formerly ITIC)
  – CIA
  – UNISYS
  – National Institute of Justice

Thank You