A New Approach for Measuring Facial Image Quality

Robert Yen, Ph.D.
U.S. Army Biometrics Task Force
Biometric Standards Team

8 November 2007

Agenda

- Background
- FaceQM Purpose
- FaceQM Approach
- FaceQM Test Results
- FaceQM Current Status

Background

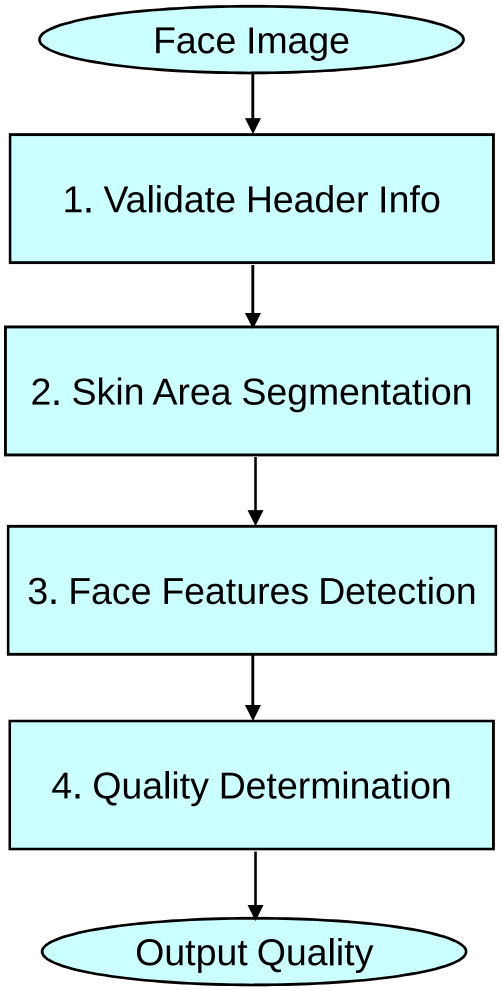
- Measuring the quality of biometric samples is a crucial step in the enrollment processes.
- Use of good-quality biometric samples increases the performance of automatic recognition systems during the matching process.
- Varying operational environments (*e.g.*, lighting, deteriorating equipment, operator training) in which samples are collected result in collection of biometric samples of varying levels of quality.
- Without consistent biometric sample quality validation, poor-quality samples contaminate databases and negatively impact the performance of automatic recognition systems.

FaceQM Purpose

- Assess quality of facial images
  - Use requirements of INCITS 385-2004 (Face Recognition Format for Data Interchange) standard and other criteria as quality parameters
  - Use FaceQM quality scores as predictor of matching performance
- Evaluate matching algorithms and systems
  - Determine "sensitivity" of individual matchers to specific quality parameters
- Support customized quality scores ("mapping")
  - For example: Acceptance-Rejection

FaceQM – Quality Measurement Flow

BIOMETRICS TASK FORCE



1. **Image Header Information Validation Module**
   - Check consistency and validate image header information
2. **Skin Area Segmentation Module**
   - Segment skin area by using Skin Color Decision Tree
   - Generate skin area mask
3. **Face Features Detection Module**
   - Locate face features – eyes, mouth, and ears
   - Measure and/or calculate face features
4. **Quality Determination Module**
   - Verify values of face features against constraints
   - Determine and output quality scores

FaceQM
1. Header Information Validation

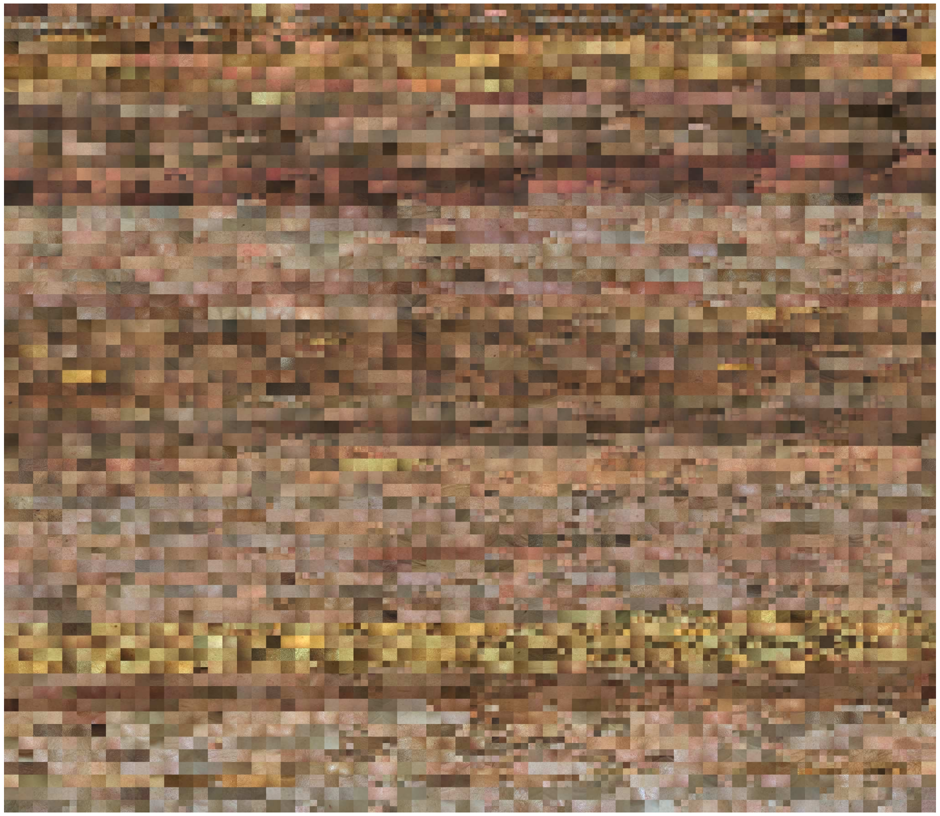
BIOMETRICS TASK FORCE

- The module verifies the following header information
  - Scanned resolution
  - Width and height values
  - Color bytes: Red-Green-Blue (RGB) bytes for each pixel
  - File size: consistency checking

FaceQM
2. Skin Area Segmentation

- Considerations: Accuracy and Efficiency
  - Our approach: Non-Uniform Binary Splitting Algorithm
- Uses trained skin color features as reference
- Based on the training process; establishes Decision Tree (DT)
- Uses the DT to segment face area

Skin Color Training Dataset



- All pixels belong to human skin color
- 2,048 x 2,048 = 4,194,304 pixels in total
- Collected from FERET Database

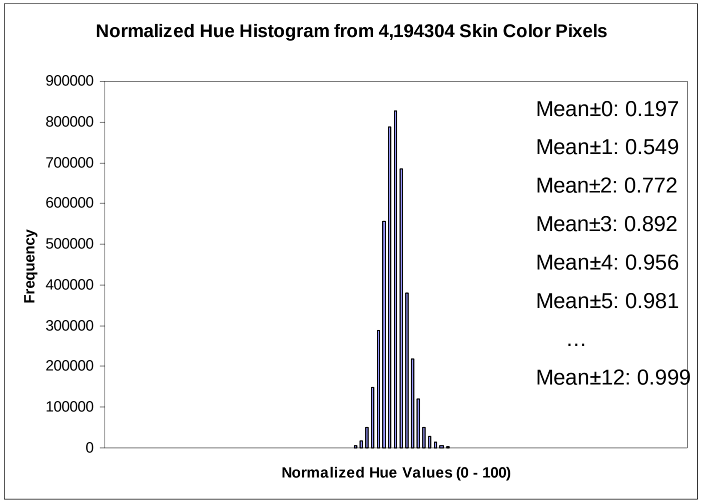
Skin Color Space

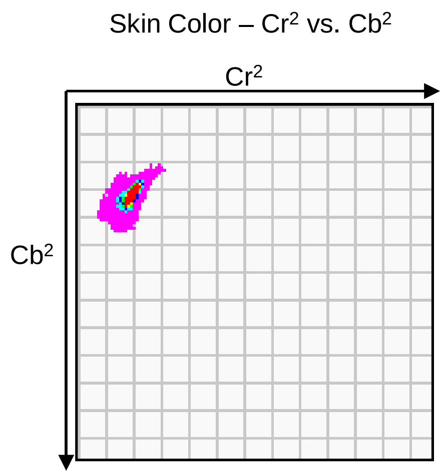
- Total combination of RGB bytes: $2^8 \times 2^8 \times 2^8 = 16,777,216$
- Skin color spectrum is narrow
- Skin color is a combination of red, yellow, and brown colors
- Several color spaces related to skin are available:

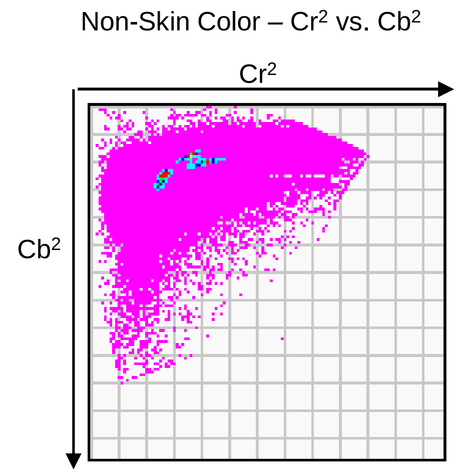
Color Space	Advantage	Disadvantage
RGB (Red-Green-Blue)	Easy to form from current capture devices	Not a good color space to present skin-tone value
YUV (Luminance-Hue-Saturation)	Good for broadcast television and compression/decompression	Has no chrominance components
YCrCb (Luminance-Chrominance)	Good for handling video information	Has no hue component
HIS (Hue-Intensity-Saturation)	Good for providing Intensity and Saturation information of color	Has no chrominance components

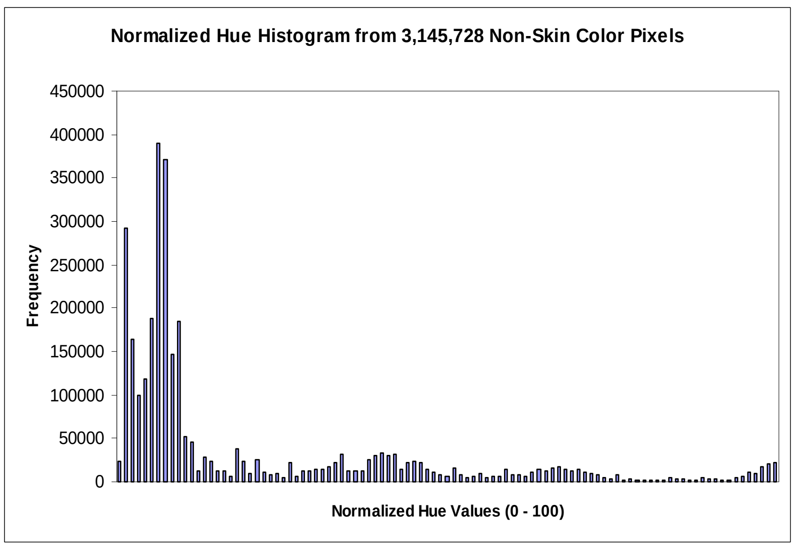
- Our approach: Combinations of Y, Cr, Cb, Hue, and Saturation

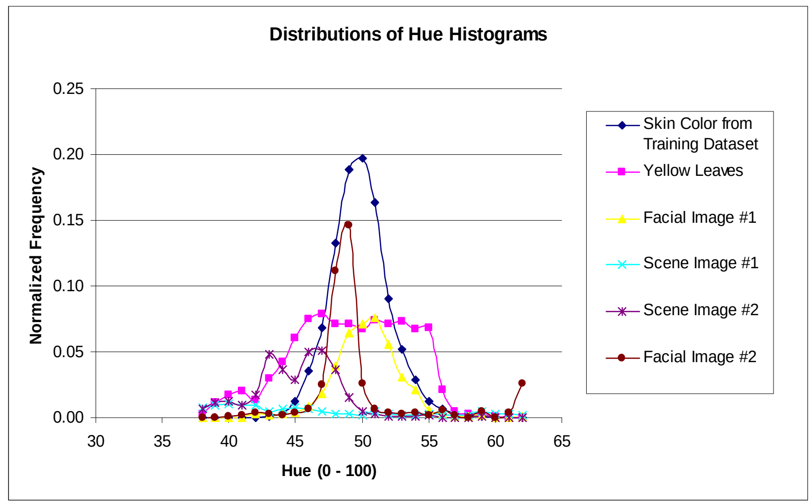
Skin Color Space - Statistic



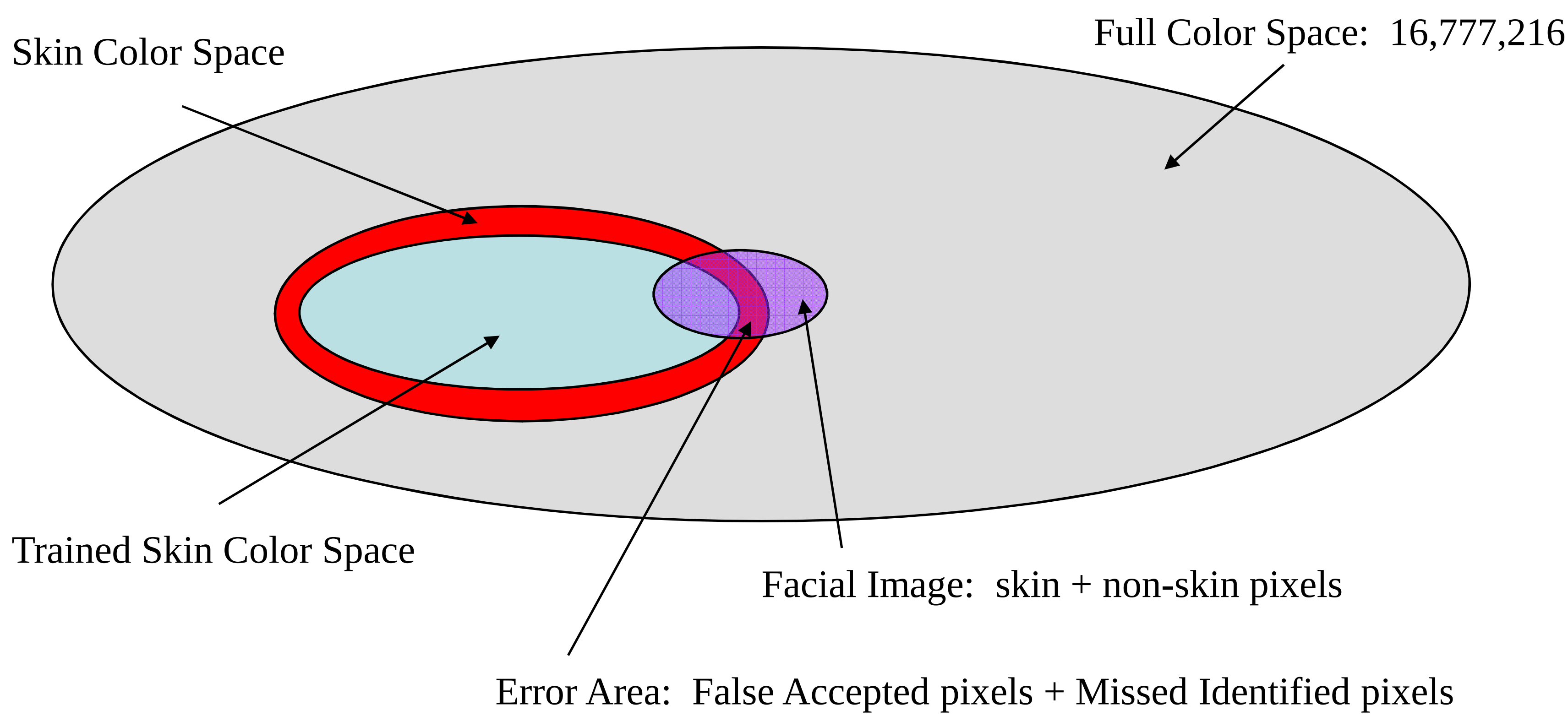








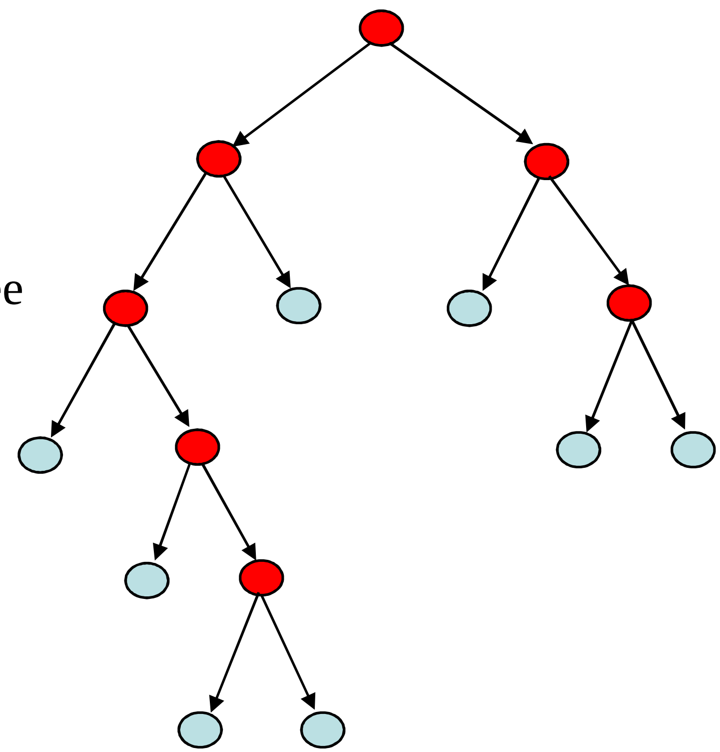
Skin Color Segmentation Analysis



Non-Uniform Binary Splitting

Mean vector of whole training dataset

NUBS Tree



Parent Node

Child Node

Each node has (1) mean vector, (2) node variance, and (3) number of pixels.

Node Splitting Rule:

(1) node’s variance > pre-defined value

(2) number of pixels > pre-defined value

(3) each parent node only splits to two child nodes

NUBS Tree:

For each parent node, need to record the child node’s number and mean vector.

Skin Color Training Process

Step 1: Set the ε (variance threshold)

Step 2: Convert each pixel's RGB bytes to a skin feature vector, v

Step 3: Calculate the mean vector and variance of whole training dataset

Step 4: Split current node all pixels into two sub-classes by using NUBS algorithm

Step 5: Calculate the mean vectors and variances for both sub-classes

Step 6: Register the nodes, mean vectors, and variance into the Decision Tree (DT)

Step 7: Repeat Step 4 to Step 6 for all nodes that need to split

Skin Color Detection Process

Step 1: Convert each pixel's RGB bytes to a skin feature vector, v

Step 2: Calculate the mean vector and variance of whole image

Step 3: Split current node all pixels into two sub-classes by using NUBS algorithm

Step 4: Calculate the mean vectors and variances for both sub-classes

Step 5: Label the nodes, mean vectors, and variance into the Tree

Step 6: Repeat Step 3 to Step 5 for all nodes that need to split

Step 7: Use Trained Decision Tree (DT) to verify each region

Skin Color Segmentation Example

Original Facial Image

Segmented Image

FaceQM
3. Face Features Detection

Feature	Constraints
24-bit color	To be examined
Red-eye	To be examined
Blur	To be identified
Near/Far*	Image width : head width = 7:4
Centered image*	Δ(mid of detected eyes or detected face width – mid of image width) < 5% of half image width
Position of eyes*	50% - 70% of the vertical distance up from the bottom edge of the captured image
Pose angle – roll*	$\pm 5^0$
Pose angle – yaw*	$\pm 5^0$
Color contrast	0.45 < average contrast value < 0.95
Color saturation*	Top half : bottom half area and Left half : right half area should have close to 1 in both saturation distributions
Luminance density*	The dynamic range of the image should have at least 7 bits of intensity variation in the facial region of the image
Eyes' locations	To be detected
Mouth location	To be detected
Ears' locations	To be detected

* INCITS 385 full frontal or token facial image's parameters tested by FaceQM

FaceQM
4. Quality Score Determination

- Quality Score is measured by the Face Features Detection and Quality Determination modules
  - Quality score indicates the “GOOD IMAGE” or “NEED TO RE-SCAN” from the validation results of all face features
- The value of the quality score can be presented in two forms:
  - D-Score: the total number of facial features that satisfy the constraints
  - C-Score: the minimum value from the quality levels of 12 facial features

Notes

1: New Quality measurement features can be increased if necessary.

2: Each feature’s weighting coefficient is adjustable – all 1s for equal weighting.

3: Each feature’s quality level is converted from measured value with it’s piece-wise mapping function that is formed by the constraint.

FaceQM – Example Results

Sight Features

24-Bit Color:	Yes	
No Red-Eye:	Yes	
Non-Blurred Degree:	Yes	-0.94 (<= 0.0)
Near/Far:	Passed	1.33 (1.17 - 1.75)
Centered Image:	Passed	4% (< 10%)
High/Low:	Passed	59% (50 - 70%)

Orientation Features

Roll Angle:	Passed	0 (< +- 5 Deg.)
Yaw Angle:	Passed	-1 (< +- 5 Deg.)

Lighting Features

Contrast:	Passed	57% (>= 45%)
Vertical Saturation Ratio:	Passed	0.13 (< +- 0.25)
Horizontal Saturation Ratio:	Passed	-0.01 (< +- 0.25)
Luminance Dynamic Range:	Passed	146 (>= 128)

Face Area, Eyes, Mouth, and Ears Detection

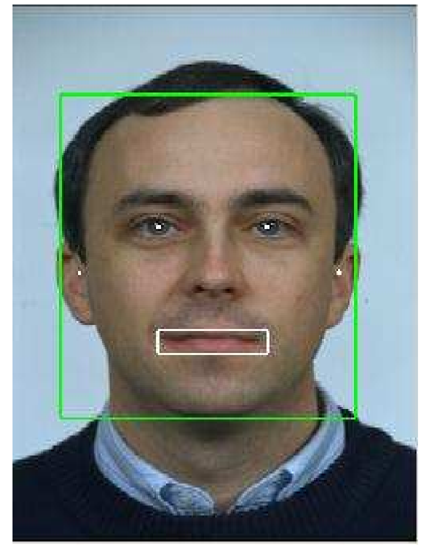
Face Area:	Detected
Eyes (Right - Left):	Y & Y
Mouth:	Y
Ears (Right - Left):	Y & Y

Quality Examined Result

D-Score:	GOOD IMAGE (D-QS = 12/12)
C-Score:	GOOD IMAGE (C-QS = 84/100)

Header Information

Image Type:	BMP
Color Bits:	24
Image Height:	768
Image Width:	512
Image Size:	1179702
Consistency Check:	PASSED



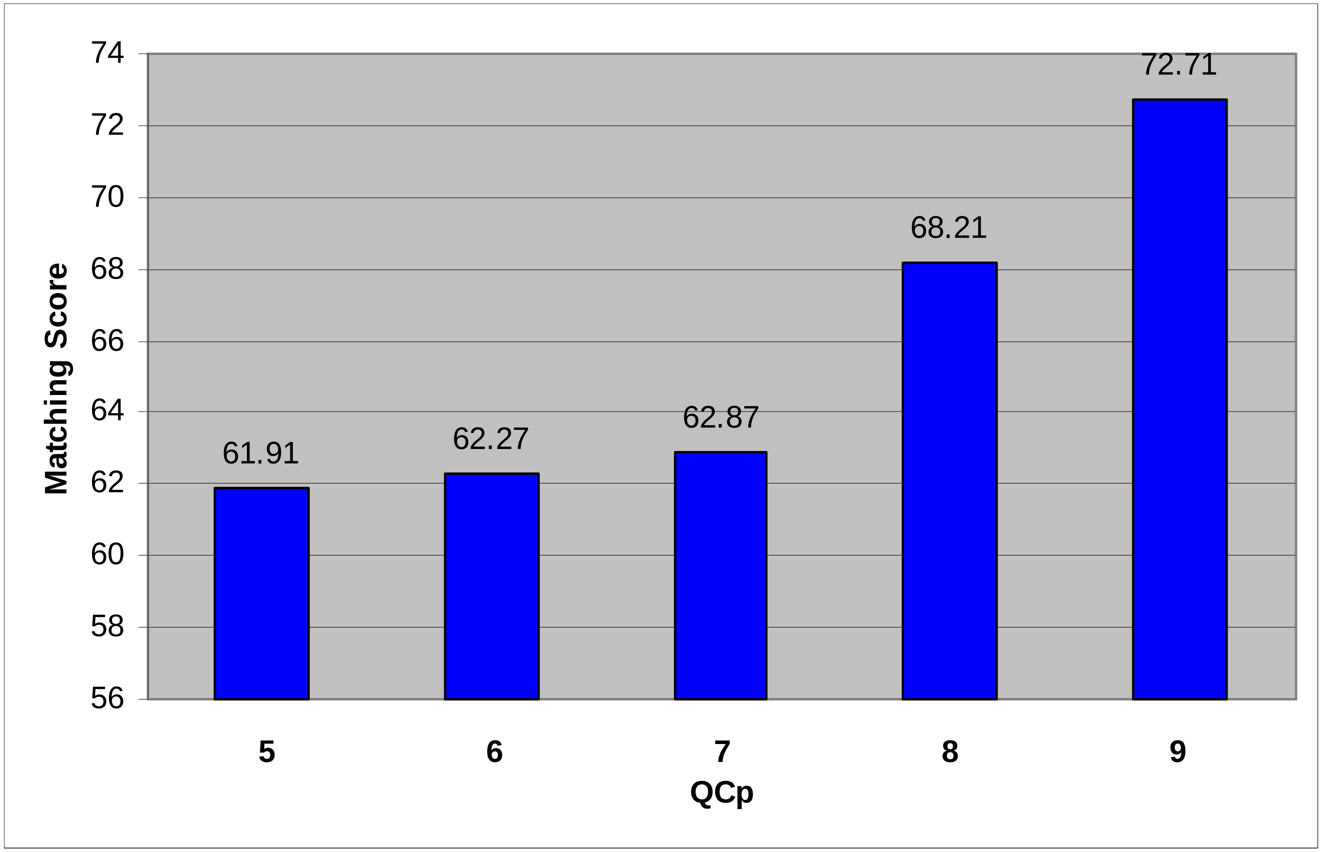
BIOMETRICS TASK FORCE

Preliminary Test Results

Training & Testing

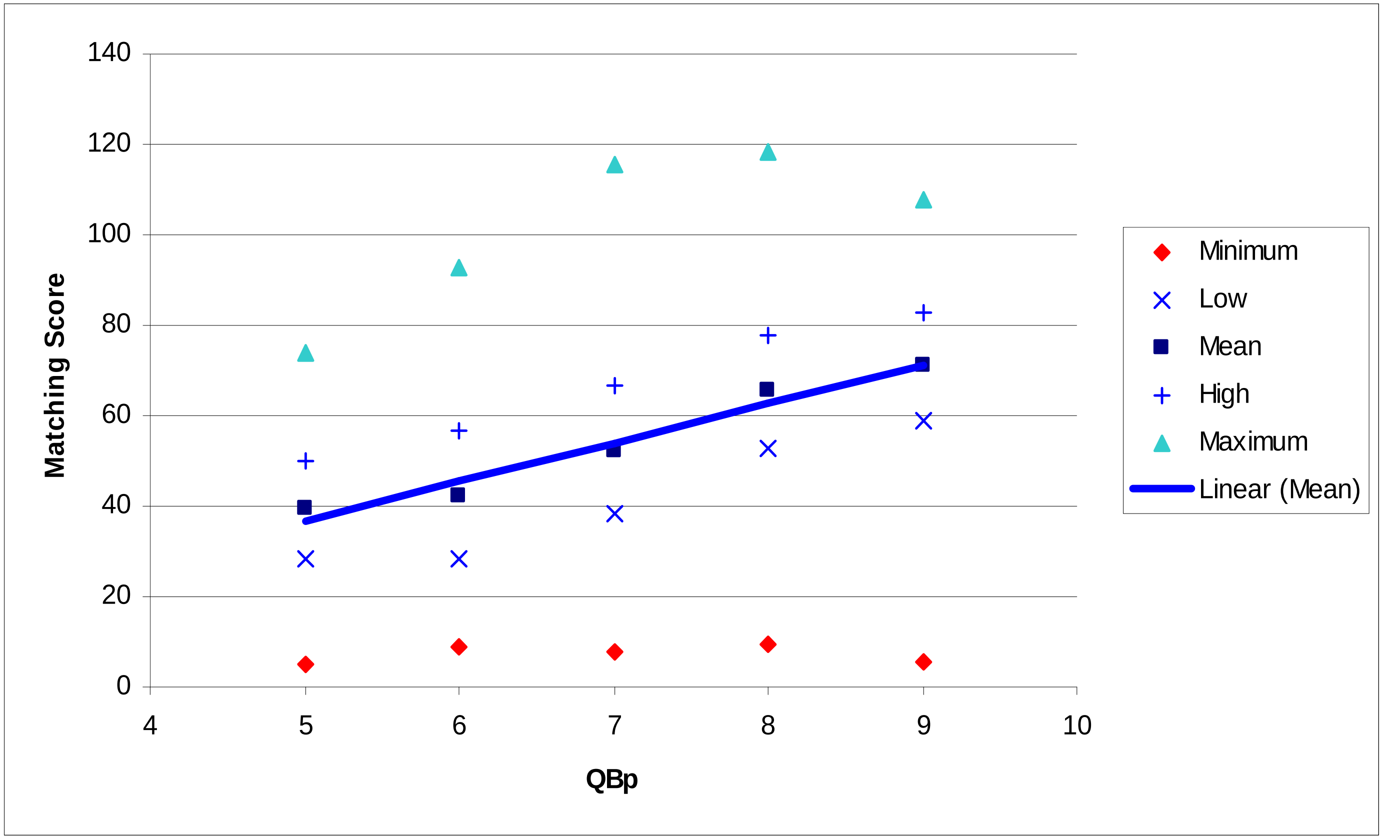
- Training Data:
  - BTF skin color training dataset – created from Color FERET Facial Image Database
- Test Data:
  - 1,806 (903 pairs) facial images
    - Color FERET Facial Image Database
    - AR (Aleix Martinez & Robert Benavente) Facial Image Database
    - No 24-bit color, Red-eyes, and Blurred detection
    - Matching scores analysis
  - 500 plus individual facial images
    - DoD ABIS Facial Image Database
    - University of Notre Dame Biometrics Database
    - Face Features testing
- Matcher:
  - One of the leading Face Recognition products in the market

Experiment Results (1/4)



The FaceQM tool has a very high confidence in predicting the higher matching score on the Gq ≥ 8 and Pq ≥ 8 paired combination on both gallery and probe datasets.

Experiment Results (2/4)



Gallery Quality Scores: 8 or 9

Probe Quality Scores: 5, 6, 7, 8, and 9

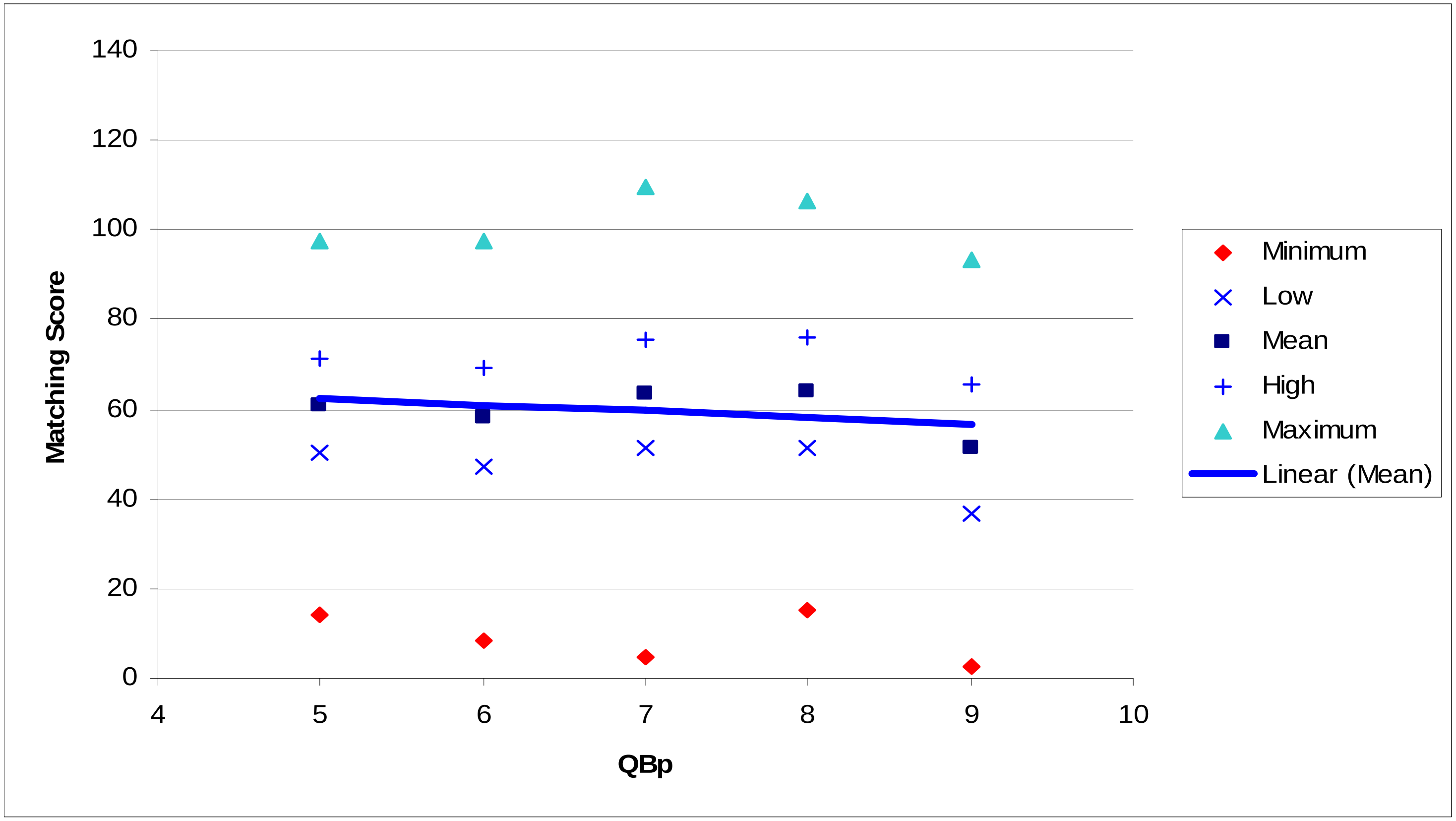
Where QB_p is Quality Scores bin of Probe Group

ms_e= -6.39 + 8.64 × QBp

rms = 2.22

Correlation Coeff. = 0.984

Experiment Results (3/4)



Gallery Quality Scores: 5, 6, or 7

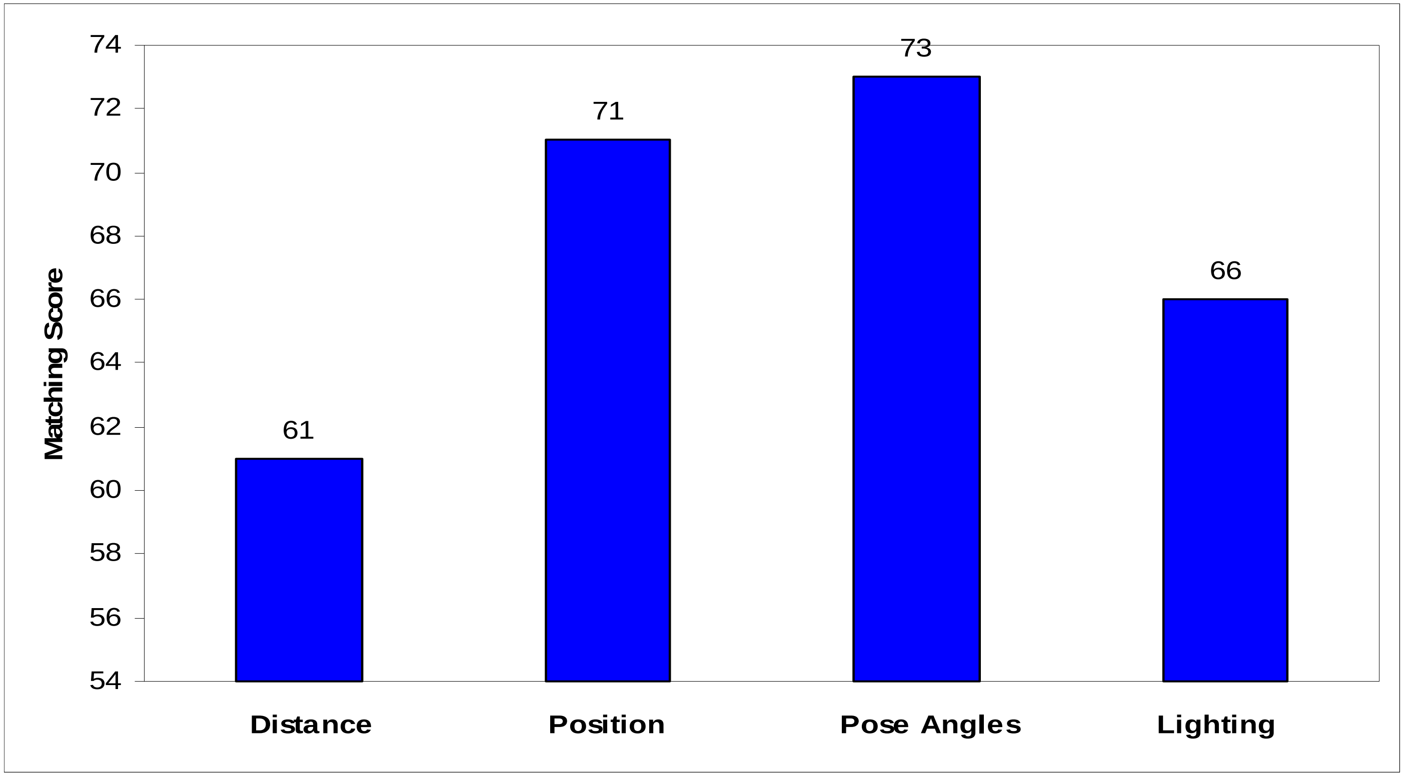
Probe Quality Scores: 5, 6, 7, 8, and 9

$ms_e = 69.3 - 1.392 \times QBp$

rms = 4.19

Correlation Coeff. = -0.425

Experiment Results (4/4)



"Sensitivity" measurement of the tested matcher to specific quality parameters

Distance category: Near/Far
Position category: Centered image and Position of eyes
Pose Angles category: Rolled and Yaw
Lighting category: Contrast, Horizontal Saturation, Vertical Saturation, Luminance density

FaceQM Current Status

- Algorithm development completed, preparing for publication
- Prototype development completed, tested with limited number of images
- Production version development in final stages, testing and debugging is underway
- Large-scale testing is underway and being conducted by NIST

Contact Information

Dr. Bob Yen
Biometrics Task Force
Booz Allen Hamilton
703-984-0434
yen_robert@bah.com

Mr. Gregory Zektser
Biometrics Task Force
Booz Allen Hamilton
703-984-0432
zektser_gregory@bah.com

www.biometrics.dod.mil