

---

**NISTIR 7540**

# **Assessing Face Acquisition**

Mary Theofanos  
Brian Stanton  
Charles Sheppard  
Ross Micheals  
John Libert  
Shahram Orandi

Information Access Division  
Information Technology Laboratory

September 2008

**NIST**

**National Institute of Standards and Technology**  
Technology Administration, U.S. Department of Commerce

---

**NISTIR 7540**

**Assessing Face Acquisition**

Mary Theofanos  
Brian Stanton  
Charles Sheppard  
Ross Micheals  
John Libert  
Shahram Orandi

*U.S. Department of Commerce  
Technology Administration  
National Institute of Standards and Technology  
Information Technology Lab  
Gaithersburg, MD 20899*

September 2008



U.S. Department of Commerce  
*Carlos M. Gutierrez, Secretary*

National Institute of Standards and Technology  
*James M. Turner, Deputy Director*

---

## Table of Contents

<b>1. INTRODUCTION.....</b>	<b>7</b>
<b>2. OBSERVATION.....</b>	<b>8</b>
2.1    USERS AND CONTEXT OF USE.....	8
<b>3. METHOD.....</b>	<b>10</b>
3.1    PARTICIPANTS .....	10
3.2    EQUIPMENT .....	11
3.3    PROCEDURE .....	12
<b>4. RESULTS.....</b>	<b>12</b>
4.1    FACE IMAGE QUALITY .....	12
4.1.1    Assessing the Quality of an Image.....	14
4.1.2    Face Overlay.....	18
4.1.3    Overall Quality.....	20
4.1.4    Application of Selected Computational Conformance Metrics.....	23
<b>5. DISCUSSION.....</b>	<b>28</b>
5.1    EFFICIENCY OF IMAGE CAPTURE .....	28
5.2    QUALITY ASSESSMENT .....	30
<b>6. CONCLUSIONS AND FUTURE WORK.....</b>	<b>32</b>
<b>7. REFERENCES.....</b>	<b>33</b>
<b>APPENDIX A: SELECTED COMPUTATIONAL FACE IMAGE METRICS.....</b>	<b>34</b>
EXPOSURE FEATURES .....	35
ENTROPY .....	35
ENTROPY ADEQUACY OF FACE ROI .....	36
EYE EXPOSURE .....	37
POSE/GEOMETRY .....	37
FACE POSITION .....	38
ROLL ANGLE (IN-PLANE ROTATION).....	39
REFERENCES .....	39

---

## List of Figures

FIGURE 1 SAMPLE FACE IMAGES	8
FIGURE 2 REPRESENTATIVE BORDER ENTRY LANE	9
FIGURE 3 PARTICIPANT AGE RANGE	11
FIGURE 4 FACE IMAGE CAPTURE LAYOUT	12
FIGURE 5 FACE PICTURE ATTRIBUTES	15
FIGURE 6 EXAMPLE OF USING THE FACE OVERLAY (IMAGE ALTERED FOR PRIVACY).	18
FIGURE 7 DISTRIBUTION OF ENTROPY ADEQUACY VALUES FOR MINIMUM RATINGS OF SHADOW AND HOT-SPOTS. NO IMAGES HAVE RATING 5 FOR BOTH SHADOWS AND HOT-SPOTS, HENCE ONLY 4 RATING VALUES ARE USED.	25
FIGURE 8 DISTRIBUTIONS OF MINIMUM OF EYE ENTROPY VALUES FOR EACH OF 5 RATING CLASSES OF EYE SHADOW	26
FIGURE 10 ROLL ANGLE MEASUREMENT AND RATINGS OF VERTICAL POSITION	28
FIGURE 11 TOTAL TIME TO CAPTURE FACE	29
FIGURE 12 TIMING TREND OF LARGE-SCALE BIOMETRIC SYSTEM DATA	30
FIGURE 13 ATTRIBUTES THAT CONTRIBUTE TO POOR GEOMETRY	31
FIGURE 14 FRONTAL IMAGE FROM FERET IMAGE DATASET. ORIGINAL (A) INPUT IMAGE (385 X 256 X3 PIXELS) AND ANY NON-ZERO ROLL ANGLE REMOVED (B). THE IMAGE IS CONVERTED TO GREYSCALE (C) AND THE FACE REGION OF INTEREST ( 112 X 108 PIXELS) CROPPED AT BOUNDARIES SET AT FIXED PROPORTIONS OF THE INTER-EYE DISTANCE FROM THE INTER-EYE MIDPOINT (D). (NOTE THAT THE EXAMPLE HAS ALSO BEEN TRANSFORMED TO THE “TOKEN” FORMAT, I.E. TO A STANDARD GEOMETRY, A PRACTICE THAT IS NOT USED FOR CURRENT QUALITY MEASUREMENTS.)	36
FIGURE 15 DEPICTION OF POSE REFERENCE AXES. (FROM INCITS 385-2004 [0])	38

---

## EXECUTIVE SUMMARY

The requirements necessary for taking a successful face picture are fairly straightforward. The camera must be operational, and the subject must be illuminated sufficiently, facing the camera and in focus. Yet, a significant portion of the facial photographs taken at United States ports of entry are unusable for the purposes of automatic face recognition. In this paper, we consider the usability components of the face image capture process that contribute to the relatively high ratio of unusable images collected by United States Visitor and Immigrant Status Indicator Technology (US-VISIT). In addition, we introduce a general evaluation methodology—including the use of a simple image overlay—to quantify various characteristics of face imagery. The experimental context mimicked the point-of-entry environment, but with specific usability enhancements. The collected data suggests that these usability enhancements may be used to improve face image capture with the current equipment.

US-VISIT requested that the biometrics usability team at the National Institute of Standards and Technology (NIST) examine the current US-VISIT face image collection process to identify any usability and human factors that may improve the existing face image capture process. As such this study did not address other technologies or technology solutions. This report presents the results of a study that examined five usability and human factors enhancements to the current US-VISIT collection process:

1. the camera should resemble a traditional camera;
2. the camera should click when the picture is taken to provide feedback to the traveler that the picture is being taken;
3. the camera should be used in portrait mode;
4. the operator should be facing the traveler and the monitor while positioning the camera and
5. provide some marking on the floor (such as footprints) to indicate to the traveler where to stand for the photograph.

The study was conducted as follows: first we visited and observed a representative operational setting (Dulles Airport) in order to understand the primary users and the context of use. Based on these observations we identified the 5 usability and human factors enhancements enumerated above that may improve the face image capture process. A usability study was designed that mimicked the operational process but incorporated the 5 enhancements and face images were collected from 300 participants. A visual inspection evaluation methodology based on an image overlay was used to quantify the various characteristics of face imagery based on the face image standards. Results from the visual inspection process compared favorably with preliminary automated face image quality metrics under development.

---

This report describes three main results. The enhancements were designed to address the extreme conditions or departures in the captured images. Implementing these enhancements resulted in:

1. 100 % of the images capturing a participant's face in contrast to the current US-VISIT collection
2. all of the participants were facing the camera -- this is a significant improvement to the process currently used at the ports of entry
3. additional improvement may be realized by using the face overlay guide proactively. By incorporating the overlay into the workstations the officers could use the guide to center the camera on the participant's face.

The recommended enhancements improved the overall captured images and can be implemented relatively easily and with relatively little cost. A follow-up study incorporating the overlay into the operators' workflow is underway.

---

## 1. INTRODUCTION

The Department of Homeland Security's (DHS) United States Visitor and Immigrant Status Indicator Technology (US-VISIT) program is a biometrically-enhanced identification system primarily situated at border points of entry such as airports and seaports. The US-VISIT program's goal is to advance the security of the United States and worldwide travel through information sharing and biometric solutions to identity management. The biometrics currently captured at US-VISIT primary inspection are fingerprints and a facial image. The fingerprint component of the system uses automated matching along with manual match verification. The face image capture process does not include automated face recognition but relies on human verifiable traveler history. Currently two flat index fingerprints are collected; however, as of 2008, the migration to a 10-print "slap" has already begun.

A face image quality assessment of airport ports of entry was performed in 2004 [1]. This assessment found that key factors for face recognition included geometric properties of pose, size, cropping, fish-eye effects and photometric properties of compression, backgrounds, and saturation. An evaluation of approximately 1.5M facial images using manual inspection found that:

- The subject was frontal to the camera in only about 5 % of images, approximately 70 % of the images had a pose angle of greater than 10 degrees
- About 5 % of the images have some part of the face cropped out of the picture
- About 1 % of the images have blur (usually motion artifacts).

An automated inspection of the same facial images with a face recognition system found that 95 % of images with an interocular distance below 74 pixels. The system failed to find the eyes in 10.6 % of the images. Finally, 14 % of the images were deemed unsuitable for face recognition.

To summarize—the primary problem with the collected images is poor geometry: specifically pose, size, crop and distortion as illustrated in Figure 1 Sample Face Images.

<sup>1</sup> Specific hardware and software products identified in this report were used in order to perform the evaluations described. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.



**Figure 1 Sample Face Images**

As the result of this assessment US-VISIT has embarked on a program for face image quality improvement. One aspect of this effort is the identification of usability and human factors issues that may impact face image capture. The National Institute of Standards and Technology's (NIST) usability and biometrics team was asked to identify any usability and human factors considerations that may improve the capture of face images at the airports.

## **2. OBSERVATION**

### **2.1 USERS AND CONTEXT OF USE**

The International Organization of Standards (ISO) defines usability as “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [2]. In order to understand the critical components of *users* and the *context of use*, the NIST usability team visited Dulles International Airport and observed the US-VISIT operational setting.

The team identified two primary users. The first is the Customs and Border Protection (CBP) Officer, the second is the traveller entering the US. We identified four components to the interaction between the officer and the traveller in the primary entry process.

1. passport and visa inspection,
2. interview questions and answers,
3. biometrics capture (fingerprint and face), and
4. processing of various paperwork.

We observed that the language barrier can be significant. During the four hour observation period the queue of travellers was constant and the officers were processing travellers as quickly as possible.



---

A representative entry lane is shown in Figure 2. The lane that we observed is approximately 208 cm (6 feet 10 inches) long. The counter height on the passenger side of the lane is 124.5 cm (49 inches). The desk or processing height for the officer is 106.7 cm (42 inches). The counter width for the passenger is 22.9 cm (9 inches) and for the officer is 55.9 cm (22 inches). The aisles are not a uniform width some are 76.2 cm (30 inches), some are 91.4 cm (36 inches), and one (accessible) aisle is 152.4 cm (60 inches). The fingerprint scanner is mounted on the 124.5 cm (49 inch) counter. A Webcam, Logitech 4000 camera, is mounted on a goose neck arm on the side of the computer monitor. In the configuration we observed, the computer monitor is positioned to the side of the officer as they faced the traveller.

We found that passengers lined up in front of the fingerprint scanner and not the camera. Generally, the officers repositioned the webcam for every passenger to accommodate the traveller's height and position in the lane. Although some officers rarely reposition the camera since positioning the camera can be trying. We observed that positioning the webcam on the traveller while verifying the image on the computer monitor and keeping an eye on the traveller required the officer's peripheral vision which contributed to the officer's difficulty in capturing acceptable images. This activity distracts the officer from directly observing the traveler during that time, thus decreasing his/her ability to observe mannerisms and behavior -- a key element in an officer's determination as to whether a person might need to be sent for secondary inspection.

Many travellers did not know the webcam was a camera. In fact several thought it was an iris scanner and moved in too close toward the camera trying to center one eye. Travellers received no indication from the device (such as a click or shutter sound) that the picture was being taken. Finally the images were taken in landscape rather than portrait mode.



**Figure 2 Representative Border Entry Lane**

---

From these observations of the users and context of use the usability team identified the following human factors enhancements:

To assist the passenger the camera should:

1. resemble a more traditional camera
2. provide some feedback that the picture is being taken.

To assist the officer:

1. Position the computer monitor so the officer can easily adjust the camera, see the image on the screen and see the traveler without using peripheral vision.
2. Use the camera in portrait mode

To assist both the officer and the traveler:

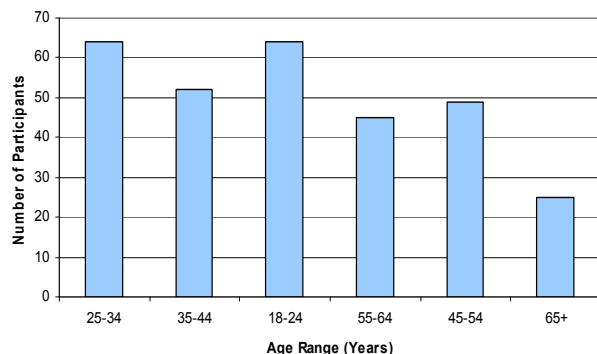
1. Provide some marking on the floor (such as footprints) to indicate to the traveler where to stand. This also assists the officer because one of the variables, the distance from the camera, is now constant. This reduces the camera adjustments for the officer, the only camera adjustment required is to accommodate the traveler's height.

From these usability observations the NIST team designed a usability experiment to examine if addressing these factors can improve face capture for US-VISIT operations.

## **3. METHOD**

### **3.1 PARTICIPANTS**

The participants were 300 adults recruited from a pool of 10,000 people who had previously agreed to participate in usability tests. There were 151 women and 149 men ranging in ages from 18 to over 65 years.



**Figure 3 Participant Age Range**

The participants ranged in self-reported height from 56 inches (142 cm) to 79 inches (201 cm). The heights were fairly normally distributed with an average height of 70.2 inches (178.3 cm) for males and 64.4 inches (163.6) for females. According to the Centers for Disease Control and Prevention (CDC) the mean individual height of men is 69 inches (175 cm) and the mean individual height of women is 63 inches (160 cm) in the US [3]. The height data collected is within 2 % of the mean in the US general population. According to the World Health Organization the worldwide mean individual male height is 5 feet 8 inches (173 cm) and the female height is 5 feet 2 inches (158 cm).

### **3.2 EQUIPMENT**

This study utilized the MBARK [4] software package for controlled capture of images from a given participant. MBARK was configured to take a single, high-resolution (1944 X 2592 pixels) image. The digital camera was mounted in a “portrait” orientation to best match the aspect ratio of the human head. The tripod was physically configured so that the MBARK system operator needed only to tilt (pitch) the camera to fill the image frame. The camera was also configured to emit a “shutter” sound upon capture.

The physical layout of the face capture station is illustrated in Figure 4. Note that the goal of using seated participants was to significantly reduced the overall image capture space. A similar reduction could be accomplished by asking participants to stand on a mark on the floor. The operators were standing and could easily observe the participant, position the camera, and view the computer monitor all at once. A live viewfinder on the rear of the camera also facilitated this multitasking.

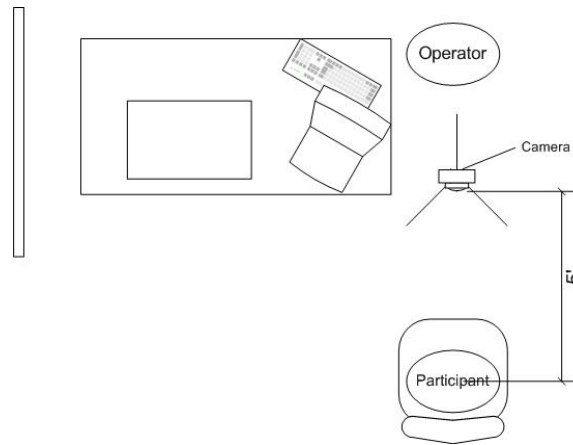


Figure 4 Face Image Capture Layout

### 3.3 PROCEDURE

Each participant completed three tasks. First participants were asked to provide fingerprint images [5]. The second task was capturing a facial image. The final task for participants was an iris scan. This paper will only address face capture. After the fingerprinting tasks, the participants were asked to sit in a chair directly facing the camera to have a picture taken. The high end digital point-and shoot camera clicked similar to a traditional single lens reflex (SLR) camera's shutter when the picture was taken.

As in the operational environment the operators were given little guidance on capturing an appropriate quality face image. The only guidance was provided in the MBARK User's Manual developed for the study. Operators were asked to position the participant in front of the camera in the chair. Participants were asked to look at the camera. Throughout the capture process the operator was afforded a continuous live preview of the cameras output. Once the participant was positioned and the camera adjusted, the operator clicked a "*Capture Now*" button to capture the image. A dialog appeared displaying a thumbnail preview of resolution sufficient for a coarse examination of the captured data. If the image was acceptable the operator clicked "Accept" otherwise he clicked "Reject", repositioned the camera and/or the subject and performed a second data capture.

## 4. RESULTS

### 4.1 FACE IMAGE QUALITY

We based the analysis for assessing the quality of the face images on the attributes identified in ANSI INCITS 385-2004 [6] and ISO/IEC 19794-5:2005(E) [7]. Those attributes are listed

---

in Table 1. Since there are no available fully automated conformance tests based on these standards, the approach used to evaluate the quality of the images was visual inspection to rate the first 14 attributes on a five-point scale where a value of one is least conducive to automatic face recognition and five is most conducive. The remaining six attributes are binary and were assigned values of 1 or 5, where 1 indicated the presence of the characteristic and 5 the absence.

Attribute	Scoring
Pose	1 to 5
Expression	1 to 5
Shoulders	1 to 5
Background	1 to 5
Subject and scene lighting	1 to 5
Shadows over the face	1 to 5
Shadows in eye-sockets	1 to 5
Hot spots	1 to 5
Eye glasses	1 to 5
Horizontal Face	1 to 5
Vertical Face Position	1 to 5
Width of head	1 to 5
Length of head	1 to 5
Obstruction*	1 to 5
Eye Color	1: undefined or 5: Color
Hair Color	1: undefined or 5: Color
Assistance in Position	1: Yes or 5: No
Eye patches	1: Yes or 5: No
Facial hairs	1: Yes or 5: No
Radical Lens Distortion	1: Yes or 5: No

**Table 1 Face Image Attributes**

As described in [8] inspecting the images manually is reliable. A human observer is capable of identifying a particular problem even in the presence of other problems, and can distinguish between failure modes. Consider an image where the facial region is saturated and also cropped at the left eye, an automated quality assessment tool may not find the face at all and report nothing. A drawback of the approach is that it is subjective. Thus, when

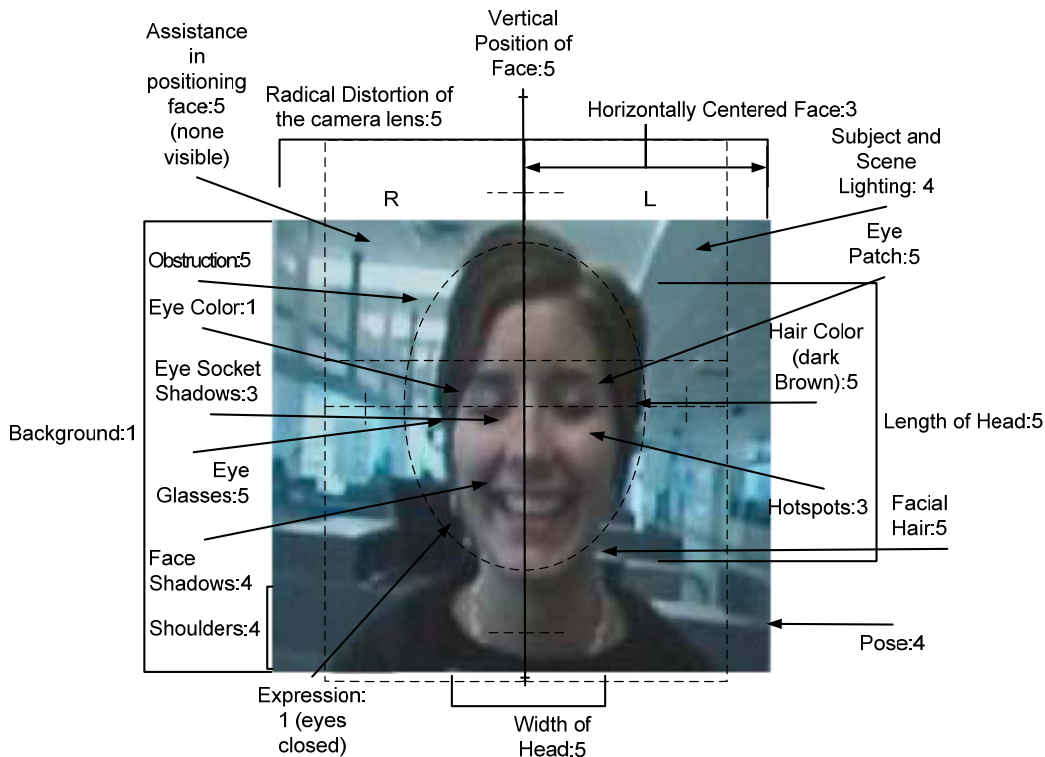
---

categorizing an attribute such as saturation, there is an inherent judgment to be made. Thus, for consistency, all images were analyzed by the same individual.

#### **4.1.1 Assessing the Quality of an Image**

To illustrate the approach used to analyze the face images consider Figure 5 and each individual attribute identified in Table 1. The binary attributes were assigned a 1 or a 5 as indicated in Table 1. Consider the six binary attributes of the image in Figure 5.

1. eye color was indeterminable and therefore was assigned a 1.
2. hair color was identifiable, assigned a 5
3. there was no facial hair, assigned a 5
4. no eye patches, assigned a 5
5. no obvious assistance in positioning, assigned a 5
6. no radical lens distortion, assigned a 5.



**Figure 5 Face Picture Attributes**

A 5-point scale was used to assess the remaining 14 attributes as described in the following sections.

*Obstruction*

The standard specifies that there shall be no head gear except in the case of religion, and in such cases, the head gear shall not cause any shadow to be cast over the face. The scale was defined as:

1. Some form of head gear;
2. Sun glasses on top of head or head band;
3. Eye glasses on top of head;
4. Large earrings;
5. No obstructions.

Using this scale the image in Figure 5 was assigned a 5.

*Expression*

Expression is an attribute of the full-face frontal pose that is known to strongly affect the performance of automated face recognition systems. It is recommended that the expression

---

should be neutral (non-smiling) with both eyes open normally (i.e., not wide-open), and mouth closed, and a smile with closed jaw is not recommended. The scale was defined as

1. Eyes closed;
2. Eyes looking away from camera and/or squinting;
3. Smile where mouth opened and teeth expose but not full frontal;
4. Closed jaw smile; and
5. Neutral with both eyes open and mouth closed.

Since the eyes were closed, this image was assigned a 1.

#### *Subject and Scene Lighting*

The standard indicates that lighting shall be equally distributed on the face. There shall be no significant direction of the light from the point of view of the photographer. The scale was defined as :

1. Excessive shadows caused by head gear
2. Excessive shadows caused by poor lighting;
3. Few shadows;
4. Lighting is fairly equally distributed;
5. Lighting is distributed equally.

The lighting is fairly equally distributed in Figure 5; the image was assigned a 4.

#### *Shadows over the Face*

The standard specifies that the region of the face from the crown of the head to the base of the chin and from ear-to-ear shall be clearly visible and free of shadows. Special care should be exercised in cases when veils, scarves or headdresses cannot be removed for religious reasons to ensure these coverings do not obscure any facial features and do not generate shadow. In all other cases head coverings shall be absent. In light of the standard specifications the following scale was developed:

1. Excessive shadows caused by head gear;
2. Large areas of shadows by poor lighting;
3. Few shadows;
4. Shadows only in eye-sockets; and
5. No shadows.

There were shadows only in the eye-sockets resulting in a score of 4.

#### *Shadows in Eye-Sockets*

The standard specifies that there shall be no dark shadows in the eye-sockets due to brow. The iris and pupil of the eyes shall be clearly visible. Our grading scale looked at the degree of shadows in the eye-sockets where we defined:

1. Shadows in both eye-sockets (caused by head gear);
2. Shadows in both eye-sockets (caused by poor sighting);
3. Shadows in only one eye socket;



- 
4. Little shadow; and
  5. No shadows.

Since there are shadows in only one eye-socket the image received a 3.

#### *Hot Spots*

It is specified in the standard that care shall be taken to avoid “hot spots” (bright areas of light shining on the face). The use of a single bare “point” light source is not acceptable. The scale was:

1. Extensive glaring from hot spots;
2. Multiple areas of hot spots (3 or more);
3. One or two hot spots;
4. One softly lighted hot spot; and
5. No hot spots.

This image was assigned a 3.

#### *Eye Glasses*

The standard specifies that eye glasses should be worn only if the individual normally wears them, and in these cases, care shall be taken that the glasses frames do not obscure the eyes. The glasses shall be clear glass and transparent so the eye pupils and irises are clearly visible. Heavily tinted glasses or sunglasses are acceptable only for medical reasons. And, there shall be no lighting artifacts or flash reflections on glasses. Using these guidelines the following scale was defined:

1. Extensive glare in both lenses;
2. Extensive glare in one lens;
3. Glare from lenses and shadows from rims;
4. Small amount of glare;
5. No glasses.

Since the person had no glasses, the image received a 5.

#### *Background*

The first step in the computer face recognition process is the segmentation of the face from the background for the purpose of registration (landmark determination). The standard recommends that the background should be plain, and shall contain no texture containing lines or curves that could cause computer face finding algorithms to become confused. Therefore, the background should be a uniform color or a single color pattern with gradual changes from light to dark luminosity in a single direction. The scale included

1. Many visible objects in background;
2. Three objects in background;
3. Two objects in background;
4. One object in background; or
5. Uniformity with no objects in background.

---

This image has at least 2 images in the background and was assigned a 1.

#### 4.1.2 Face Overlay

In order to assist in evaluating the additional attributes a face overlay was developed (Figure 6). The overlay was scaled to match the size of the images on the screen. The oval shape, vertical, and horizontal lines when placed over a face image assist in the visual inspection of the geometric attributes of a face image. The overlay is not visible in the captured images. According to the ANSI INCITS 385-2004 Standard [6], the approximate horizontal midpoints of the mouth and of the bridge of the nose shall lie on an imaginary vertical line AA at the horizontal center of the image. The upper tick-mark represents the height of the crown of the head and the distance from the edge of the picture. The lower tick-mark represents the base of the shoulder-line to the bottom edge of the picture. Line BB represents the imaginary horizontal line passing through the center of both eyes of an individual's face image. Line CC helps to line-up the horizontal midpoint of the bridge of the nose with the horizontal center of the image. The tick-marks on line CC helps with grading the horizontal midpoint between the ear lobe and the outer edge of the image. The R and L represent the right and left from the vantage point of the individual in the image looking at the camera. The oval helps with centering the face. Figure 6 illustrates the use of the face overlay. The face overlay was used to assess the following six attributes.



Figure 6 Example of using the face overlay (image altered for privacy).

#### *Pose*

The pose is known to strongly affect the performance of automated face recognition systems. The standard specifications for the full-face frontal pose were used to guide the quality assessments in this study. The standard requires that the entire head is in the image, as well as the outline of the shoulders. Also, the rotation of the head shall be less than +/- 5 degrees

---

(i.e., roll, pitch and yaw should be close to the coordinates (0,0,0). The five grading points incorporate this guidance:

1. Wearing head gear, top of head chopped off;
2. Chopped off of a shoulder;
3. Head turned at angle;
4. Full frontal not centered and;
5. Full frontal centered.

Since the face was full frontal but not fully centered, the example in Figure 5 was given a score of 4.

### *Shoulders*

According to the standard, the shoulders in a full-face frontal pose shall be “square on” to the camera. In our assessments, the outline of both shoulders was required to determine the squareness of the individual’s shoulders. Thus, the grading scale of the shoulders was based on the available outline of both of the individual’s shoulders in an image. The scale was defined as:

1. Excessive chopping of shoulder;
2. Not enough of shoulders visible;
3. Square shoulders but uneven chopping of shoulders;
4. Square shoulders, almost even chopping of shoulders; and
5. Square shoulders, even chopping of shoulders.

The shoulders were square and almost even, resulting in a score of 4.

### *Horizontally Centered Face*

The standard specifies that the approximate horizontal midpoints of the mouth and of the bridge of the nose shall lie on an imaginary vertical line positioned at the horizontal center of the image. In our experience of assessing the images, we found that severe chopping of the image on either the individual’s right or left prevented an image from being horizontally centered. The scale was defined as:

1. Severe chopping of image or excess amount of space on one side;
2. Larger amount of space on one side than the other;
3. Small difference in the spacing of the 2 sides;
4. Very small difference in the spacing; and
5. Perfect centering.

The example in Figure 5 was graded a 3.

### *Vertical Position of Face*

The standard specifies that the face shall be positioned such that the distance between the imaginary horizontal line passing through the center of the eyes and the bottom edge of the image is 50 % to 70 % of the total vertical length of the image. The scale was defined as:

1. Large amount of head tilting and looking away from camera;

- 
2. A large amount of head tilting;
  3. Small amount of tilting eyes are off the horizontal;
  4. Small amount of tilting, eyes are slightly off horizontal; and
  5. Eyes are perfectly horizontal.

This image was graded as 5.

#### *Width of Head*

The standard specifies that the width of a head is the horizontal distance between the midpoints of two imaginary vertical lines drawn between the upper and lower lobes of each ear and shall be positioned where the external ear connects the head. The scale was defined as:

1. Part of the head is chopped off;
2. image is chopped too closed to the head;
3. part of the hair is chopped off;
4. the head is turned slightly causing an ear to be out of sight; and
5. Adequate head width.

The example was assigned a 5.

#### *Length of Head*

The standard specifies that the length of a head is defined as the vertical distance between the base of the chin and the crown (top of the hair). The crown to the chin portion of the Full Frontal Image pose shall be no more than 80 % of the vertical length of the image. The scale was defined as:

1. Head gear or the chopping off of the top of the head;
2. Image is chopped very close to top of the head;
3. Sunglasses on top of head ;
4. Too much head length ;
5. Adequate head length.

Using this scale the example was assessed as a 5.

### **4.1.3 Overall Quality**

Due to hardware and software failures during the data collection process, data for several participants was found to be corrupt. Face images were available for 267 of the 300 participants. The following tables show the counts and percentages for each of the attributes described in the previous sections.

Tables 2, 3 and 4 identify the counts and percentages of the 267 images for the binary attributes. Note that a participant may have multiple attributes in Table 2.

---

Attribute	Count	Percentage
Eye Patches	0	0.00 %
Facial Hair	84	31.00 %
Radical distortion of the camera lens	2	0.75 %
Obstruction	19	7.00 %

**Table 2 Binary Attributes**

Eye Color	Count	Percentage
Indeterminate	35	13.00 %
Black	47	18.00 %
Blue	49	18.00 %
Brown	50	19.00 %
Dark Brown	73	27.00 %
Dark Green	1	0.40 %
Gray	1	0.40 %
Green	11	4.00 %

**Table 3 Counts of Eye Color**

---

Hair Color	Count	Percentage
Indeterminate	18	7.00 %
Brown	68	25.00 %
Dark Brown	89	33.00 %
Graying Brown	10	4.00 %
Graying Dark Brown	6	2.00 %
Blond	21	8.00 %
Gray Blond	4	1.00 %
Black	36	13.00 %
Black with reddish coloring	3	1.00 %
Graying Black	4	1.00 %
Gray	3	1.00 %
Red	3	1.00 %
Light Brown	2	0.70 %

**Table 4 Counts of hair color**

Tables 5 and 6 provide the number of images for each element of the rating scale for the attributes. Table 5 includes the attributes that were rated using the face overlay. Table 6 includes the remainder of the attributes.

Rating	Background		Pose		Shoulders		Horizontally Centered		Vertical Position		Width of Head		Length of Head	
	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
1	2	75.00	32	12.00	28	10.50	2	0.75	0	0.00	9	3.37	28	10.49
2	158	59.00	31	11.60	3	1.00	18	6.74	3	1.12	7	2.62	13	4.87
3	77	29.00	38	14.00	149	56.00	184	68.91	75	28.09	9	3.37	8	3.00
4	30	11.00	162	60.70	11	4.00	23	8.61	61	22.85	7	2.62	0	0.00
5	0	0	4	1.50	76	28.46	40	14.98	128	47.94	235	88.01	218	81.65

Table 5 Face Image Data Summary

Rating	Expression		Subject/Scene Lighting		Shadows/Face		Shadows/Eyes		Hot Spots		Glasses		Obstruction	
	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
1	0	0.00	3	1.00	5	1.87	8	3.00	0	0.00	0	0.00	11	4.12
2	2	0.70	259	97.00	252	94.38	52	19.48	56	20.97	0	0.00	6	2.75
3	37	14.00	4	1.49	8	3.00	154	57.68	147	55.06	28	10.49	1	0.37
4	82	31.00	1	0.51	1	0.37	40	14.98	58	21.72	39	14.61	0	0.00
5	146	55.00	0	0.00	1	0.37	13	4.87	6	2.25	200	74.91	249	93.26

Table 6 Face Image Data Summary

#### 4.1.4 Application of Selected Computational Conformance Metrics

Several of the subjective visual assessments were found to be procedurally consistent with some of the computational face conformance testing metrics [9] developed recently at NIST.

---

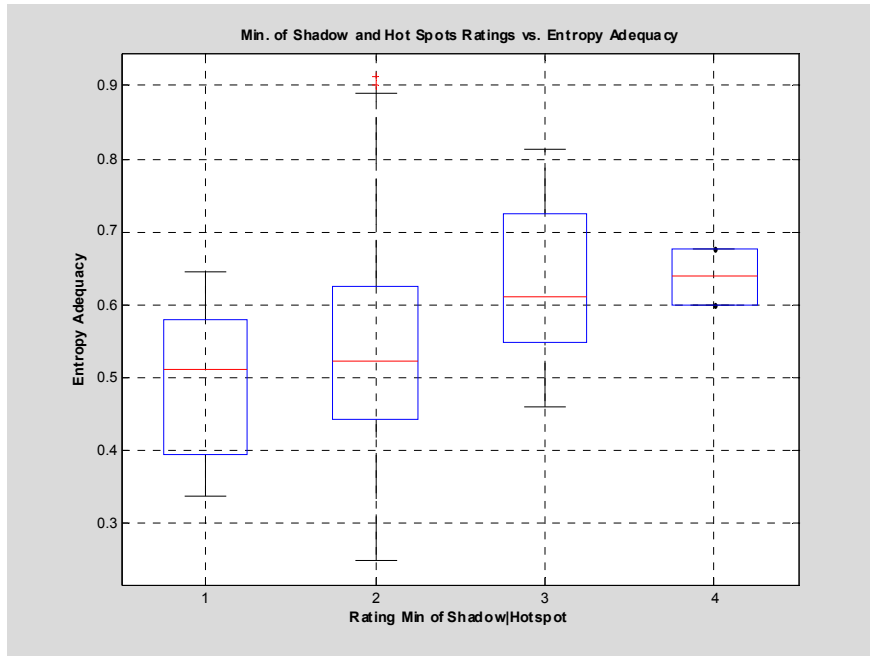
In several cases it is possible to cross-validate the two forms of quality assessment by comparing the 5-point ratings with the numerical scores of the objective measures to the extent that the measurements apply to specific image attributes.

The majority of the computational metrics of face image quality are applied to a rectangular *face region of interest* (ROI) sampled from the full-frame image. Details of the extraction of the face ROI are described in another NIST publication currently in preparation. For the present purposes it may suffice that the vertical and horizontal limits of the sampling window are defined in relation to the image coordinates of the eyes (Appendix A: ). Thus, in order to implement the computational metrics, the eye coordinates were extracted using an interactive graphics tool enabling the sampling of the eye positions as specified in the face standards.

As described previously, subjective ratings are assigned to images for assessments of both shadows and hot-spots. For purposes of computational analysis of these image defects, it is possible to consider both of these defects in terms of information loss from the image rendering, suggesting the use of an entropy measure. For gray scale images, entropy translates directly to the number of bits used in the image rendering. Its calculation shown in Appendix A: is a function of the number of gray levels used in the rendering. Thus, the typical grayscale image consisting of values 0 to 255, or 256 discrete levels of gray. Assuming the image contains at least 1 pixel at each of the 256 levels, the number of bits required would be 8, i.e.,  $256 = 2^8$ .

As a measure of overall exposure of the face region, one may look simply at the entropy, or use of the available grayscale as an indicator. In order to assess shadows and hotspots, which may be more locally placed, we might define a metric that evaluates entropy on a more local scale. Such a metric, referred to as *entropy adequacy*, has been defined so as to evaluate the entropy at each position of a moving window covering the face ROI. Currently the dimension of the square window is  $1/20^{\text{th}}$  of the width of the face ROI. The entropy adequacy is computed as the proportion of window positions of the total having entropy values of at least 50 % of the maximum entropy possible for the number of pixels contained in the sample window.

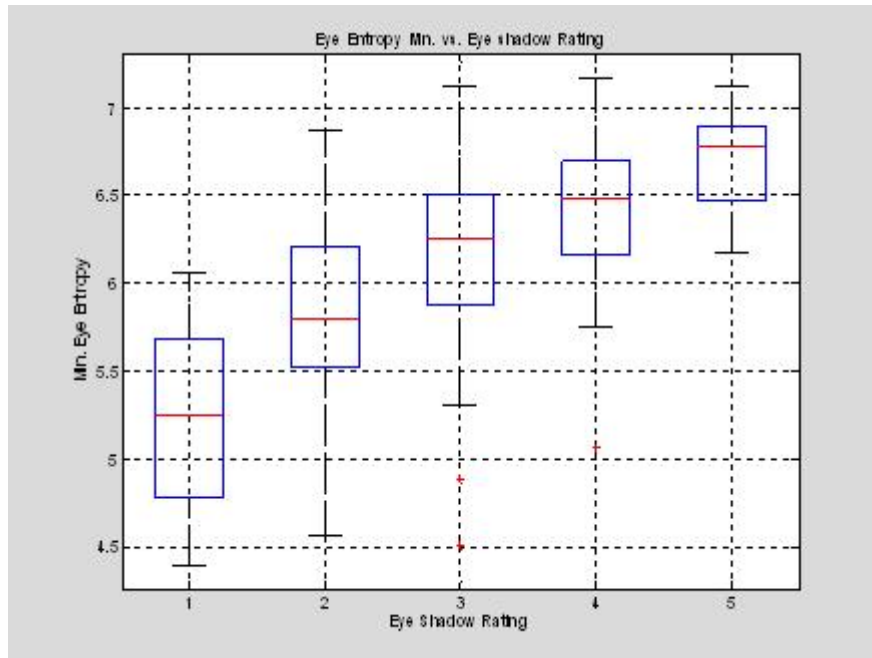




**Figure 7 Distribution of entropy adequacy values for minimum ratings of shadow and hot-spots. No images have rating 5 for both shadows and hot-spots, hence only 4 rating values are used.**

As entropy adequacy makes no distinction between shadows and hot-spots, comparison was made between entropy adequacy and the minimum of the ratings assigned the images for shadows and hotspots. Figure 7 shows the distributions of entropy adequacy for rating values 1 to 4. Note that there were no cases in which a rating of 5 occurred for both shadows and hotspots. Moreover, the absence of whiskers for the distribution box of rating 4 indicates that only a few cases had minimum rating of 4 for both shadows and hotspots. Regardless, however, the trend of the subjective ratings is consistent with increasing entropy adequacy for reduced prominence of either shadows or hotspots.

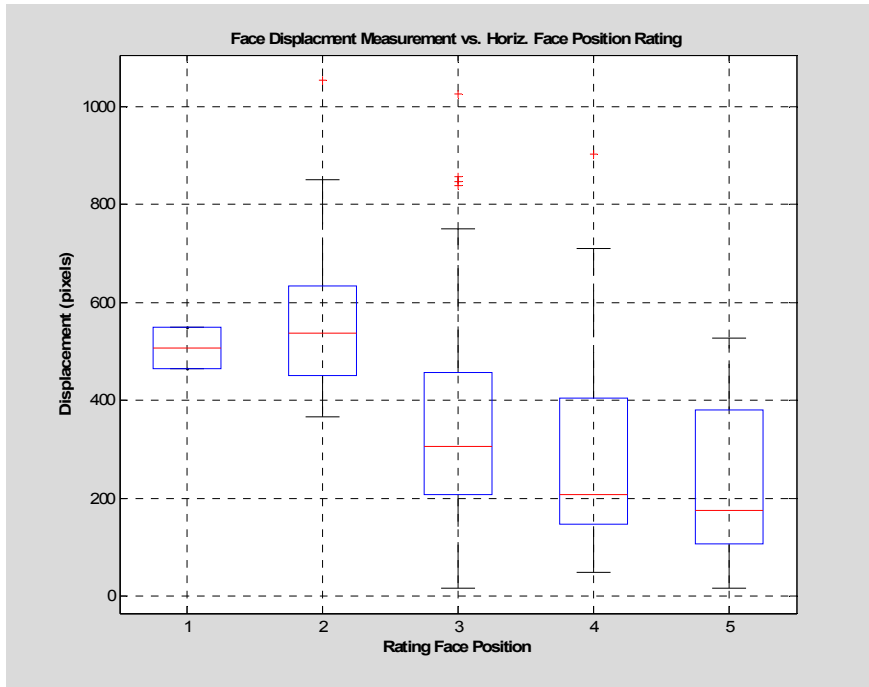
Another of the subjective measures having corresponding computational metrics is the evaluation of eye socket shadowing. The NIST suite of computational quality metrics includes entropy measurements of rectangular regions just containing the eyes. In general the detail present in the eye regions should yield relatively high entropy values unless shadowed or the gray levels reduced by eyes being closed or by specular reflection from eyeglasses. Inasmuch as the eye shadow rating does not differentiate between right or left locations of shadows, we might compare the ratings to the minimum of the entropy calculations for the left and right eye. This comparison is shown in Figure 8 and suggests strong agreement between the subjective assessment of eye shadowing and the eye entropy metric.



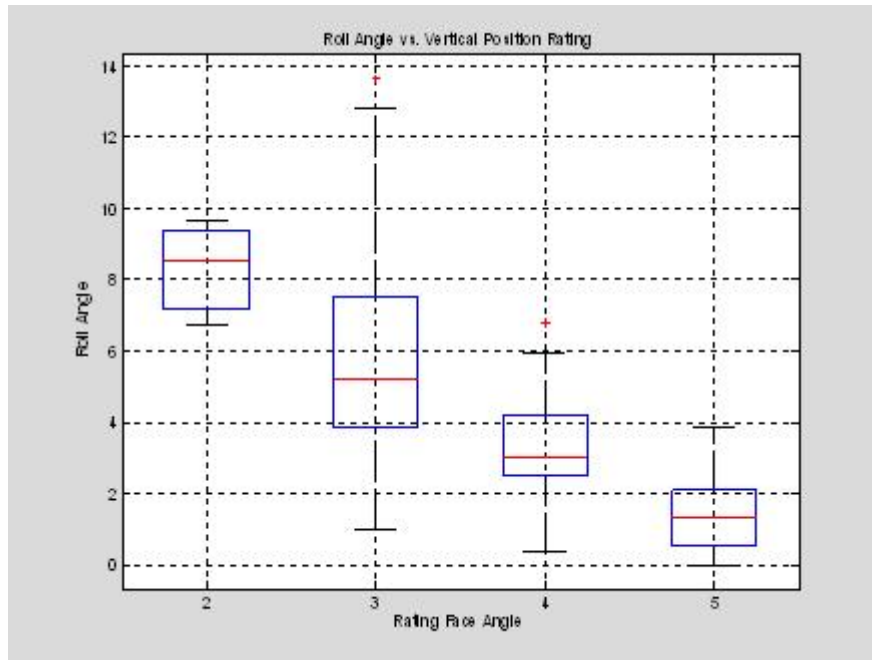
**Figure 8 Distributions of minimum of eye entropy values for each of 5 rating classes of eye shadow**

The subjective evaluation of face position was consistent with the description of the standard regarding vertical and horizontal placement of the head in the image frame, but differed from the computational approach. The subjective measurements examined the head position with regard to horizontal displacement from the ideal as indicated by equality of right – to – left margins from the vertical line passing through the interpupillary midpoint. Vertical placement ideal was defined as having the horizontal line passing through the eye coordinates at between 50 % to 70 % from the bottom of the image frame.

By contrast, the computational approach described head position displacement as a distance of the interpupillary midpoint from the ideal position, without consideration of whether the displacement was vertical, horizontal, or some combination. While the measurements are not strictly comparable, Figure 9 suggests some degree of consistency between the rating measure of horizontal face displacement and the computational displacement measure, at least for the rating values 2 – 5. The distribution of displacement distances for rating level 1 is degenerate, consisting of only a few values.



**Figure 9 Computational face displacement (distance from reference position) vs. ratings of horizontal position**



**Figure 10 Roll angle measurement and ratings of vertical position**

What is somewhat surprising is the apparent correlation between the subjective measure of vertical displacement and the roll angle metric. Figure 10 shows a rather strong correspondence between vertical displacement and roll angle. The rather strong agreement is explained in examining the criteria used for vertical placement rating. In making the rating for vertical position, head tilt figures heavily in the subjective assessment.

## 5. DISCUSSION

### 5.1 EFFICIENCY OF IMAGE CAPTURE

We also examined the time required to capture the face images or efficiency. For many biometric capture processes, the measure of efficiency is unexpectedly ill-defined. Consider a requirement that specifies that the capture process will be less than 15 seconds. This requirement does not specify when the process starts or ends. Measuring task time requires precise and easily measurable start and stop events. For many modalities, it can be difficult to establish a definitive, yet common event that delimits the task. For example, what defines the start of a face photo?

It is necessary to identify the sequential steps that are performed in order to capture a face image. As such efficiency depends on more than just the camera shutter speed, one must also

---

take into account the arrival and departure times of the different individuals. For this study, the arrival time included both the time for the photographer to provide instructions and the time required for the individual to sit. As individuals entered, they were instructed where to sit. After sitting down, there was an occasional verbal exchange between the individual and the photographer about their pose. In addition there may have been adjustments to the camera, or repositioning of the individual, and finally the actual taking of the photograph or image capture. Thus, the definitive start and stop events that everyone performed were the subject sitting down and the image capture.

Figure 11 shows the time required to capture the face images. The median time was 12s se with an average time of 15s.

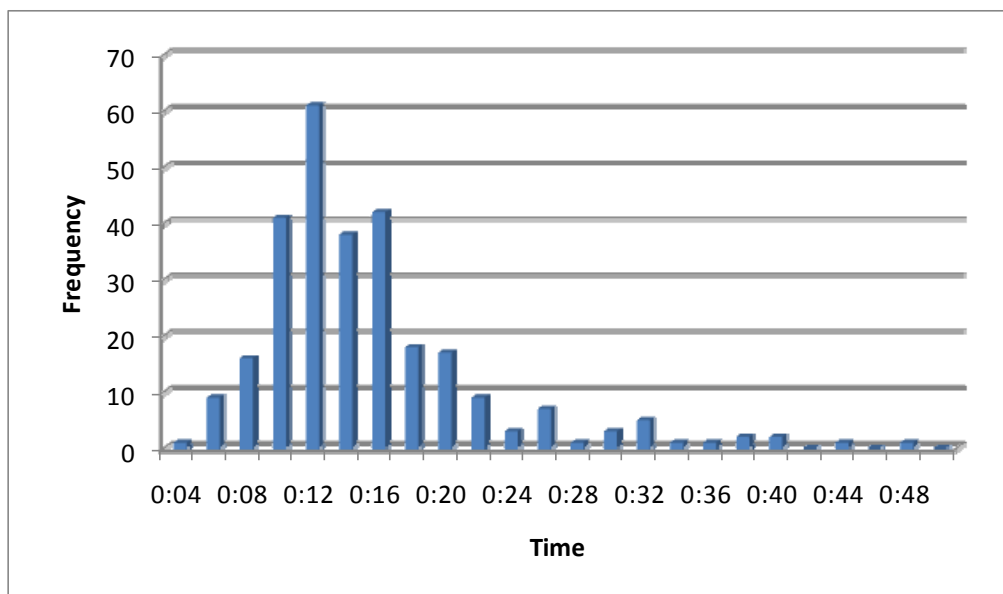


Figure 11 Total time (min:sec) to capture face

The trend of the total process time is similar to the trend that has been found in the collections of data for large-scale biometric systems as illustrated in Figure 12.

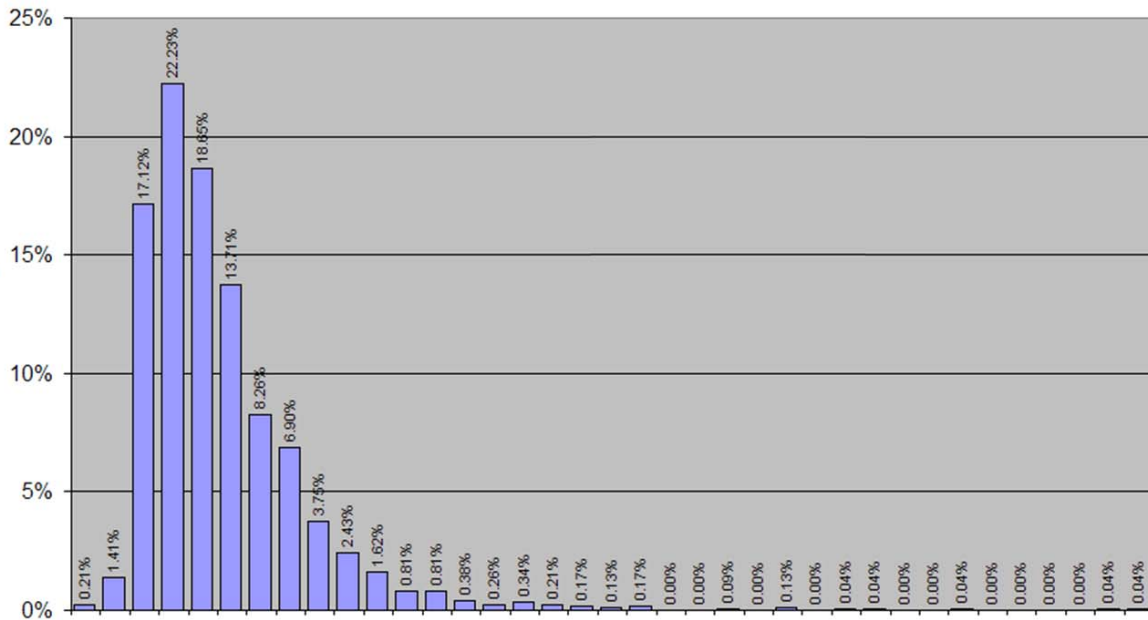


Figure 12 Timing trend of large-scale biometric system data

## 5.2 QUALITY ASSESSMENT

This usability study was designed to determine if addressing human factors could improve the current face image capture process for CBP operators without introducing additional tasks for the operators. Since operators currently use visual inspection and their judgment to determine the acceptability of the image, we developed a systematic methodology based on a similar, but highly regimented, visual inspection method for evaluating the attributes of the image quality from the ANSI [6] and ISO [7] standards.

According to a more systematic survey of the 2004 POE images described in [8] three specific defects that are easily detected using manual inspection are known to exist in the POE images: cropped faces, over-exposed faces, and non-frontal head poses. Eleven percent of the images were cropped primarily due to the camera not pointing at the subject, or the subject standing too close to the camera.

Using the methodology described in this report to evaluate the data we did not observe significant pose problems. All (100 %) of the images captured a participant's face in contrast to the US-VISIT collection. In addition, all of the participants were facing the camera. This is a significant improvement to the process currently used at the ports of entry. To further evaluate the attributes characterizing the pose and positioning of the images, we developed a face overlay template. Using this overlay we found that the majority of the images were centered and had an appropriate pose as illustrated in Figure 13. Six attributes

---

contribute to the poor geometry of the facial images collected by US-VISIT. Those attributes include pose, shoulders, horizontally centered, vertical position, width of head, and length of head. Figure 13 provides an overview of our attributes which contribute to pose.

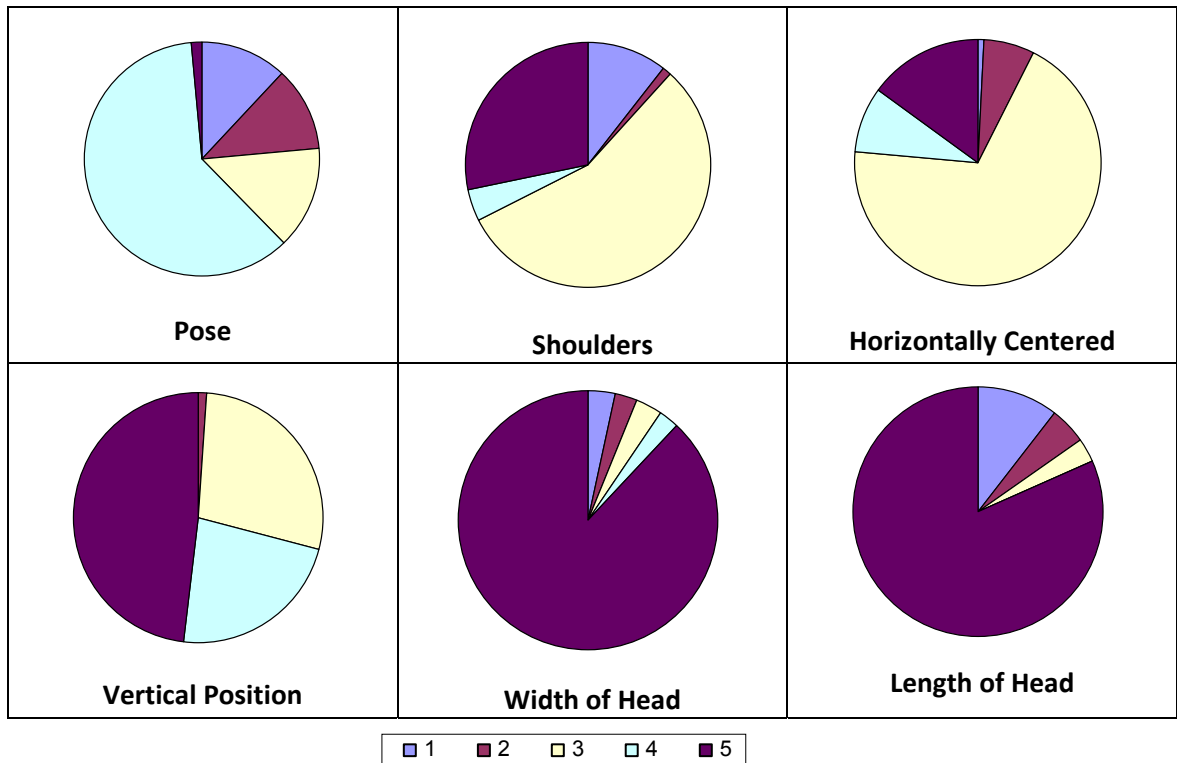


Figure 13 Attributes that contribute to poor geometry (where 1 is poor and 5 is good)

Grother and Quinn [8] also applied a commercial image quality analysis tool to the 2004 POE data and found that:

- Yaw had large variance; about half the faces had more than 5 degrees of yaw.
- Eye distances were very small; the median eye distance was only about 50 pixels.
- The background was very cluttered in most of the images and frequently included partially visible faces of other people waiting in line.
- Faces were poorly centered or un-centered – generally the result of the camera operator not correctly pointing the camera at the individual.
- Many of the images were blurry, possibly due to the face or camera moving at the time of capture.

---

Again, applying the usability enhancements and the inspection methodology we did not observe these problems in our image data. As illustrated in Figure 13, the faces were generally centered with little variance in yaw. Eye distances were in the 210 to 575 pixel range. None of the images had any significant blurriness.

## **6. CONCLUSIONS AND FUTURE WORK**

Currently US-VISIT Customs and Border Patrol Officers have little guidance on face image quality and use visual inspection to determine the acceptability of the image. Moreover, officers must process passengers quickly and face long queues of passengers. In this study we examined five usability and human factors enhancements to the current collection process:

1. the camera resembled a traditional camera;
2. the camera clicked as the picture was taken providing participants feedback on the process;
3. the camera was used in portrait mode;
4. the operator was facing the participant and the monitor while positioning the camera.;
5. the participants were seated at a fixed distance from the camera limiting the camera adjustments required by the operator.

Since a goal of US-VISIT is to reduce the operators workload, these enhancements were designed to have minimum impact on the operators. We specifically did not introduce any new technology or requirements on the operator.

The enhancements were designed to address the extreme conditions or departures in the captured images. For example, we found that implementing these enhancements resulted in 100 % of the images capturing a participant's face in contrast to the current US-VISIT collection and all of the participants were facing the camera with an appropriate pose (centered) and no observed distortion or blurriness. This is a significant improvement to the process currently used at the ports of entry. Thus the recommended enhancements improved the overall captured images and can be implemented easily and with little cost.

However, additional improvement may be realized by using the face overlay guide proactively. By incorporating the overlay into the workstations the officers could use the guide to center the camera on the participant's face. A follow-up study incorporating the overlay into the operators' workflow is underway.



---

## 7. REFERENCES

- [1] L. Nadel, "Approaches to Face Image Capture at US-VISIT Ports of Entry," NIST Biometric Quality Workshop II, Nov. 2007, retrieved from <http://www.itl.nist.gov/iad/894.03/quality/workshop07/presentations.html>.
- [2] International Organization for Standards. *ISO 9241-11 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: guidance on usability* Geneva, Switzerland: Author, (1998).
- [3] C. L. Ogden, C.D. Fryar., M. D. Carroll, and K.M. Flegal, "Mean Body Weight, Height, and Body Mass Index, United States 1960 – 2002" [Electronic Version]. *Advanced Data From Vital and Health Statistics*, pp. 347, Oct. 2004. Retrieved from <http://www.cdc.gov/nchs/data/ad/ad347.pdf>
- [4] National Institute of Standards and Technology. *Multimodal Biometric Application Resource Kit*. Gaithersburg, MD. Retrieved from <http://www.itl.nist.gov/iad/894.03/nigos/mbark.html>
- [5] M.F. Theofanos, B. C. Stanton, S. Orandi, R. Micheals, and N. F. Zhang, *Usability Testing of Ten-Print Fingerprint Capture (NIST IR 7403)*, 2006, retrieved from <http://zing.ncsl.nist.gov/biiousa/>
- [6] ANSI INCITS 385-2004, Face Recognition Format for Data Interchange. American National Standards Institute, Inc.
- [7] ISO/IEC 19794-5:2005(E) Information Technology - Biometric Data Interchange Formats - Part 5: Face image data. JTC1 : SC37, International Standard Edition, 2005. <http://isotc.iso.org/isotcportal>.
- [8] P. Grother and G. Quinn, "Baseline Quality of US VISIT POE Facial Images", NIST Deliverable to DHS US-VISIT Face Image Quality Improvement Project, April 20, 2008.
- [9] J. Libert, "NIST Facial Image Conformance Testing Guidelines", in progress.

---

**APPENDIX A: SELECTED COMPUTATIONAL FACE IMAGE METRICS**

---

---

## Exposure Features

(excerpt from ISO/IEC 19794-5:2005(E))

### 7.4.2.1 Greyscale density

The dynamic range of the image should have at least 7 bits of intensity variation (span a range of at least 128 unique values) in the facial region of the image. The facial region is defined as the region from crown to chin and from the left ear to the right ear. This recommendation may require camera, video digitizer, or scanner settings to be changed on an individual basis when the skin tone is excessively lighter or darker than the average (preset) population.

Standards for face image data [0,0] specify several attributes which bear upon the exposure of the digital face image. These include the greyscale density of the image, i.e., the degree to which the image has used the available greyscale, 0 – 255 in an 8-bit/pixel image, the existence of shadows (under-exposure) or “hot spots” (over-exposure) in the image, and shadows in the eye regions of the image. The approach taken by NIST researcher, Libert [0] is to address each of these quality defects in terms of information loss with respect to the ideally exposed face rendering. Hence, it is appropriate to examine the information content of the image via measurements of entropy.

### Entropy

First, color images are converted to greyscale (luminance) as a weighted average of the three color channels [0]

$$I_{greyscale} = 0.2989 I_{red} + 0.5870 I_{green} + 0.1140 I_{blue} \quad (1)$$

Then entropy [0] of the greyscale image may be calculated according to the expression

$$H = - \sum_{i=1}^n p_i \log_2 p_i, \quad i = 1 \dots 256; \quad p_i > 0 \quad (2)$$

where  $p_i$  = proportion of pixels in  $i^{th}$  bin of the histogram of the image region of interest (ROI) under examination.



**Figure 14 Frontal image from FERET image dataset. Original (a) input image (385 x 256 x3 pixels) and any non-zero roll angle removed (b). The image is converted to greyscale (c) and the face region of interest ( 112 x 108 pixels) cropped at boundaries set at fixed proportions of the inter-eye distance from the inter-eye midpoint (d). (Note that the example has also been transformed to the “token” format, i.e. to a standard geometry, a practice that is not used for current quality measurements.)**

In a typical quality evaluation, the entropy measurement is applied to a rectangular subset of the image that includes most of the face. This ROI (see Figure 144)dimension with respect to the midpoint between the eye coordinates and the measurement provides a direct assessment of the “grey-level density” of the face image. Entropy measurements are made for each of small rectangular regions containing the left right eyes

### **Entropy Adequacy of Face ROI**

**(excerpt from ISO/IEC 19794-5:2005(E))**

#### **7.2.8 Shadows over the face**

The region of the face, from the crown (as defined in section 4.6) to the base of the chin, and from ear-to-ear, shall be clearly visible and free of shadows. Special care shall be taken in cases when veils, scarves or headdresses cannot be removed for religious reasons to ensure these coverings do not obscure any facial features and do not generate shadow. In all other cases head coverings shall be absent.

#### **7.2.10 Hot spots**

Care shall be taken to avoid "hot spots" (bright areas of light shining on the face). These artefacts are typically caused when one, high intensity, focused light source is used for illumination. Instead, diffused lighting, multiple balanced sources or other lighting methods shall be used. A single bare “point” light source is not acceptable for imaging. Instead, the illumination should be accomplished using other methods that meet requirements specified in this clause.

This feature attempts to detect local regions of over- or underexposure of the face ROI. The measure is defined as the proportion of  $n \times n$  blocks of the face ROI that attain at least 50 %

---

of the maximum possible entropy value for a block of a selected size. The metric is computed as follows:

1. The size of a square sample window is taken as  $1/20^{\text{th}}$  of the width of the face ROI;
2. The procedure requires that the length of the window be an odd number, so if the computed length is even, it is enlarged by 1 pixel;
3. To accommodate application of the window to edges of the ROI, padding is added by symmetric mirroring of the image pixel values across the edge boundaries;
4. Centering the window on each pixel of the original face ROI, the entropy is calculated as described above and entered into an array corresponding to the dimensions of the face ROI.;
5. Given the block size, the maximum entropy possible if computed as

$$E_{\max} = \frac{\log_{10} n \times m}{\log_{10} 2} = \frac{\log_{10} 81}{\log_{10} 2} = 6.3399 \quad (3)$$

6. A count is made of the number of values in the entropy array having values greater than or equal to a threshold,  $0.5 E_{\max}$ . The count is normalized by dividing by the total number of elements in the entropy array to form a single proportion, i.e. value between 0 and 1. A value of 1.0 would indicate that all sampled blocks attain at least 50 % of the maximum entropy possible. Areas of “hot spots” or shadows as well as low contrast should yield lower values of this metric. The 50 % threshold and the size of the sample window are somewhat arbitrary at this point, supported only by a few tests. Further testing is planned, but the intent of the measure is to evaluate image exposure on a local level.

## Eye Exposure

**(excerpt from ISO/IEC 19794-5:2005(E))**

### **7.2.9 Shadows in eye-sockets**

There shall be no dark shadows in the eye-sockets due to the brow. The iris and pupil of the eyes shall be clearly visible.

The following features form the beginning of analysis of exposure in eye regions. The eyes tend to be the most detailed regions of the face image and if properly exposed should have relatively high entropy values, if not the highest of any subregion of the face. Moreover, one would expect entropy values should be about equal values for both eyes unless illumination is uneven across the face.

For this measurement. A rectangular ROI is sampled about each image eye coordinate pair, sized proportional to the image width (currently  $W_{\text{roi}} = 0.1 \times W_{\text{face}}$  and  $H_{\text{roi}} = 0.05 \times W_{\text{face}}$ , selected by inspection of images). Entropy as specified above is then computed for each eye.

## Pose/Geometry

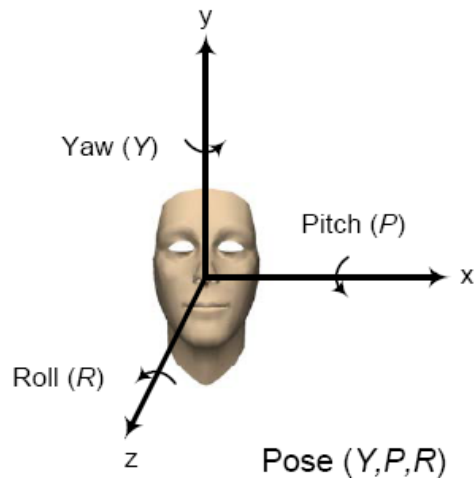


Figure 15 Depiction of pose reference axes. (from INCITS 385-2004 [0])

Face image geometry, especially yaw and pitch orientation of the face in the image frame, may have the greatest single effect upon matcher performance. If not excessive, roll may be corrected via 2-D affine transformation, commonly applied by face matchers. Pitch and yaw are much more difficult to measure precisely much less to rectify. But unless the matcher incorporates some reasonably accurate rectification scheme able to recover a frontal face image pose from some non-frontal viewpoint, most comparison methods will be challenged. Face standards such as [0, 0] for subject (or face) pose to be described relative to three axes as shown in Figure 15 Optimum or frontal pose would have  $0^\circ$  departure for all three measures yaw, pitch, and roll (Y, P, R). Assuming that a matcher is able to find critical reference points such as eye coordinates, roll should be removed during the geometric normalization (or tokenization) step or in the face ROI sampling procedure as described. Non-zero yaw and pitch angles are more troublesome.

The present feature set includes measures for face position and roll (also known as in-plane rotation). Additional features developed at NIST include a set of candidate features aimed at sensing departures from the frontal,  $Y = 0^\circ$ , yaw condition by measuring symmetry of the face ROI. These features are not included here as this was not evaluated in the subjective quality assessment.

### Face Position

Face position is taken as the distance (in pixels) of the inter-eye midpoint from the optimal placement relative to its position in a token image as specified in ISO/IEC and INCITS face standards [0, 0]. The midpoint of actual eye positions is the mean of the eye coordinates (column, row) values. According to the standard, the appropriate vertical placement of the inter-eye midpoint is given as  $0.6W$  ( $W$ =image width) from the top of the token format image in which height is specified to be  $W / 0.75$ . The untokenized image may not conform to a

---

standard height. Yet an “optimum” position proportional to that of the tokenized frame might be calculated using the top of the the acquired image as the reference. In this case, the position may be computed as  $0.45H$ , where  $H$  is the height of the acquired image. Distance between actual and optimal eye midpoint position is then computed using the distance formula

$$d_{pixels} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} , \quad (4)$$

where  $(x_1, y_1), (x_2, y_2)$  are coordinates of actual and optimal inter-eye midpoints respectively. This metric does not resolve horizontal and vertical components of the displacement vector, but the additional values could be provided if the need is justified. In this case, however, it might be more appropriate to describe the displacement using conventional vector notation of a length and an angle.

### **Roll Angle (In-Plane Rotation)**

Roll or in-plane rotation angle (in degrees) is measured with respect to offset of eye coordinates from horizontal,  $R = 0^\circ$ . Here the measurement is implemented by the function *interocular.m* Roll angle is generally corrected during geometric normalization performed by the matcher. However, Grother [0] reports that mean differences of greater than  $8^\circ$  between probe and gallery image pairs adversely affects matcher verification performance in spite of apparent geometric transformation to  $0^\circ$  roll angle.

Given  $x, y$  coordinates of right and left eye, roll angle,  $R$ , can be computed via

$$R_{deg} = \tan^{-1}\left(\frac{y_r - y_l}{x_r - x_l}\right) \times \frac{180^\circ}{\pi} \quad (5)$$

where  $(x_r, y_r)$  and  $(x_l, y_l)$  are right and left eye coordinates.

### **References**

- ANSI-INCITS 385-2004 Face Recognition Format for Data Interchange. American National Standards Institute, Inc.
- ISO/IEC 19794-5:2005(E) Information Technology - Biometric Data Interchange Formats - Part 5: Face image data. JTC1 : SC37, international standard edition, 2005. <http://isotc.iso.org/isoportal>.
- J. M Libert Face Image Quality: Part 1 Algorithms [in preparation for publication as a NIST Internal Report].
- C. A. Poynton, A Technical Introduction to Digital Video, NY, John Wiley & Sons, 1996.
- R. C. Gonzalez, R.E. Woods, S.L. Eddins, Digital Image Processing Using MATLAB, New Jersey, Prentice Hall, 2003.
- P. Grother. Face Recognition Vendor Test 2002: Supplemental Report, NISTIR 7083, February 2, 2004.