DHS SCIENCE AND TECHNOLOGY

Estimating relative skin reflectance and measuring its effect on recognition.

November 2018

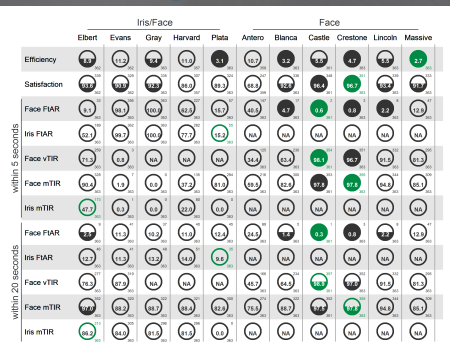
Homeland Security
Science and Technology

Yevgeniy B. Sirotin, PhD.
SAIC Identity and Data Sciences Laboratory
Maryland Test Facility

Arun Vemury
Biometrics and Identity Technology Center,
Department of Homeland Security,
Science and Technology Directorate

2018 Biometric Technology Rally

- Large-scale biometric scenario test.
  - High-throughput security/screening process.
  - Unstaffed operation.
  - 11 face acquisition systems from different commercial organizations.
  - 363 paid volunteers:
    - Demographically diverse.
    - Used each acquisition system.
- The Rally evaluated the efficiency, effectiveness, and satisfaction with the technologies.
- The effect of subject demographics, especially Race and Gender, on biometric performance has become a topic of interest.
- This presentation will describe the effects of relative skin-reflectance on biometric performance.



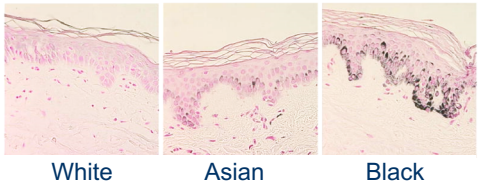
http://mdtf.org

Overview

- Skin Reflectance Variation
- Estimating Skin Reflectance using Rally Enrollment Images
- Reflectance for Different Self-Reported Race Groups
- Statistical Modeling of Biometric Performance
  - Acquisition System-independent effects
  - Acquisition System-specific effects
- Conclusions

Skin Reflectance Variation

- The absorption of light by human skin is governed by chromophores:
  - Melanin and Blood are normally the major absorbers in the visible range.
  - Bilirubin creates a jaundiced appearance in pathological cases.
- Melanin concentration varies among individuals:
  - Produced by melanocytes and prevents DNA damage by absorbing UVR (<400 nm).
    - The amount and type of melanin produced differs between lighter and darker skin.
  - Melanin is a spectrally broad absorber and changes the net diffuse reflectance of skin.
- Skin with higher constitutive melanin content is less susceptible to sunburn.



Brenner and Hearing. The Protective Role of Melanin Against UV Damage in Human Skin. *Photochem. Photobiol.* 2008:84(3).

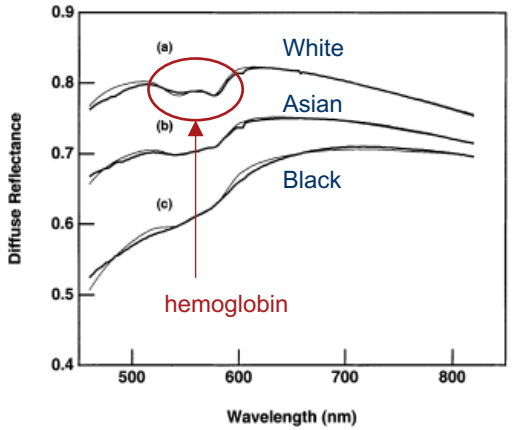
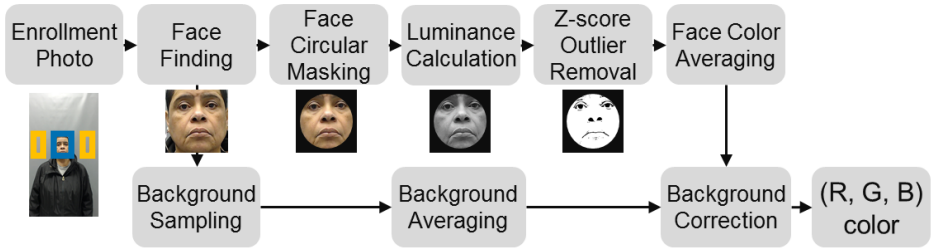


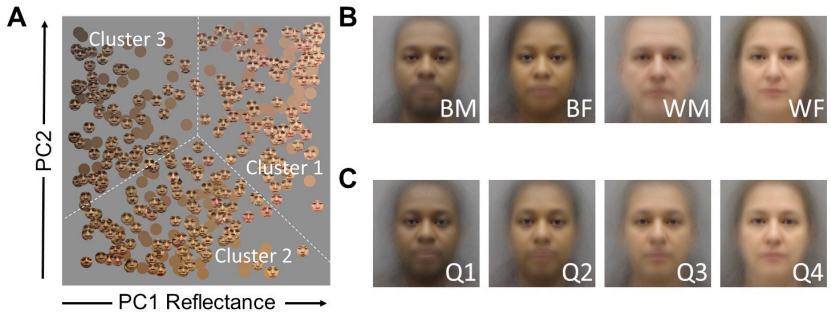
Figure 1. Diffuse reflectance spectra from skin sites on the back of three volunteers with different skin types. (*a*) White Caucasian (type II), (*b*) Japanese (type III), and (*c*) black African (type V). The presence of melanin is most evident by the larger intensity variation in the 460–560 nm region of the spectrum. Reflectance spectra (*thick lines*) and model fits (*thin lines*).

Zonios, Bykowski, and Kollias. Skin Melanin, Hemoglobin, and Light Scattering Properties can be Quantitatively Assessed *In Vivo* Using Diffuse Reflectance Spectroscopy. *Journal of Investigative Dermatology*. 2001:117(6).

Estimating Reflectance from Enrollment Images: Correcting for Variation in Exposure/Gain

- Correction of Rally Enrollment Images for incident light and camera gain.
  - Face pixel intensity is the amount of incident light I_o reflected at the camera after some transformation f based on skin chromophores and scattering and camera gain g:
    $$I_{face,\lambda} = gI_{o,\lambda}f(\mu_{HbO,\lambda}, \mu_{HbR,\lambda}, \mu_{M,\lambda}, G_\lambda)$$
  - Grey reference pixel intensity is just a constant wavelength independent attenuation k:
    $$I_{ref,\lambda} = gI_{o,\lambda}k$$
  - To correct for $gI_{o,\lambda}$:
    $$I_{face,\lambda}/I_{ref,\lambda} = k^{-1}f(\mu_{HbO,\lambda}, \mu_{HbR,\lambda}, \mu_{M,\lambda}, G_\lambda)$$
- PCA on obtained (R,G,B) values shows a face color space.
  - PC1 and PC2 explained 96.1% and 3.4% of the variance in (R,G,B) values, respectively.
- We used PC1 as the estimate of face *Reflectance*.
  - PC1 Loading: {R: 0.7, G: 0.5, B: 0.4}





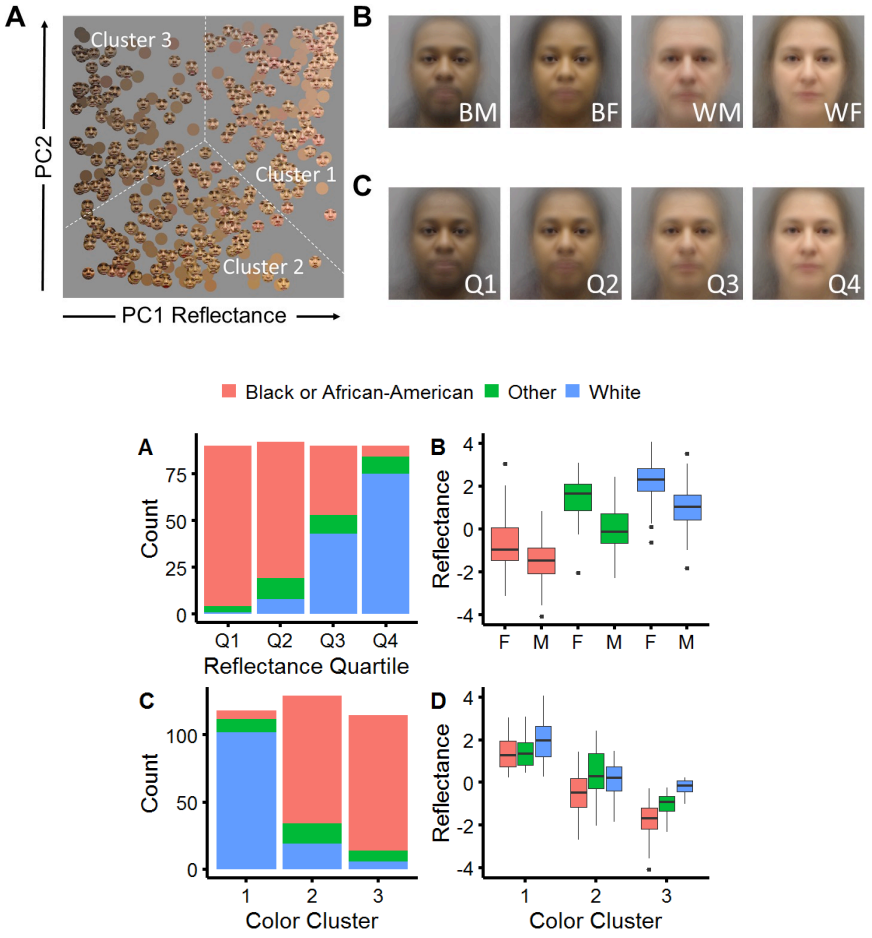
Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Reflectance vs. Self-reported Race

- Volunteers were asked to self-report their *Race* and *Gender*.
- Estimated Reflectance was related to self-reported *Race* and *Gender*.
  - Lower reflectance for Volunteers self-identifying as “Black or African-American” than as “White”
  - Females of each *Race* had higher reflectance, on average, than males

Firooz et. al, Variation of Biophysical Parameters of the Skin with Age, Gender, and Body Region. ScientificWorldJournal. 2012: 386936.

- Estimated *Reflectance* values for different race groups showed strong overlap.



Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Statistical Modeling Approach – Average Response Variables

- Fitted three response variables (i.e. three models):
  - **Average transaction time:**
    - Average amount of time volunteers spent with the systems.
  - **Average same-day gallery accuracy**:
    - Average mated similarity score between probe images from the systems relative to the enrollment photos.
  - **Average historic gallery accuracy**:
    - Average mated similarity score between probe images from the systems relative to images taken at MdTF over 4 years.

Full Model
(11 covariates)

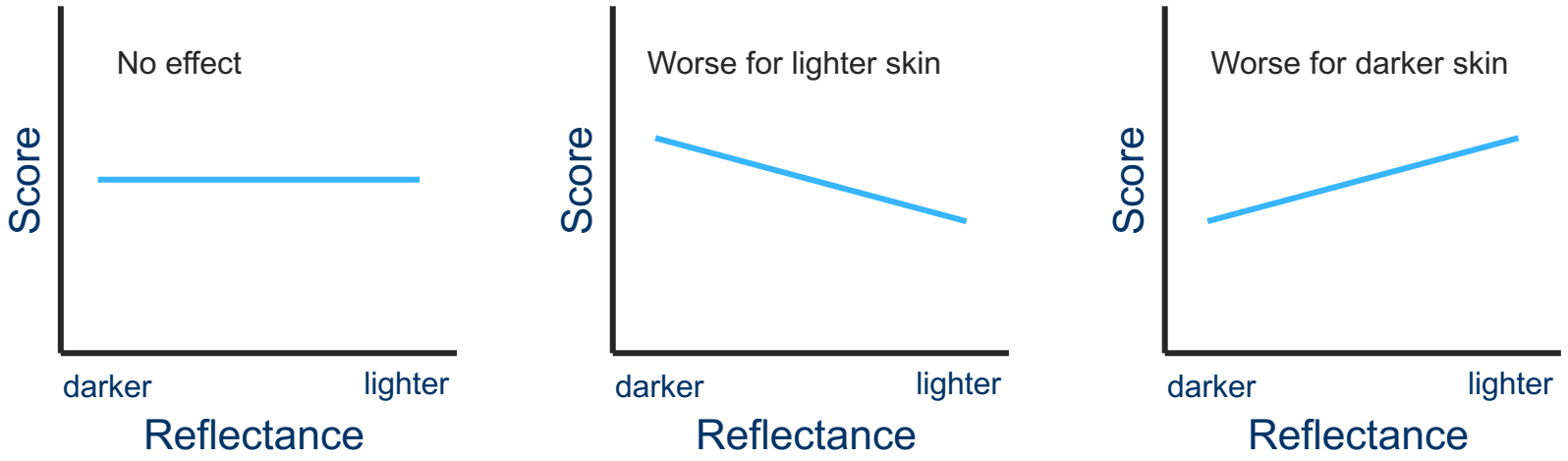
$$\Theta_j = \beta_0 + \beta_1 gender_j + \beta_2 eyewear_j + \beta_3 race_j + \beta_4 age_j + \beta_5 age_j^2 + \beta_6 height_j + \beta_7 height_j^2 + \beta_8 weight_j + \beta_9 weight_j^2 + \beta_{10} reflectance_j + \beta_{11} reflectance_j^2 + \epsilon_j$$

Optimal Model
(Select k covariates minimizing AIC)

$$\mathbf{x}_j = [x_{1,j}, x_{2,j}, \ldots x_{k-1,j}]$$
$$\boldsymbol{\beta} = [\beta_1 \ , \beta_2 \ , \ldots \beta_{k-1} \ ]$$
$$\bar{\Theta}_j = \beta_0 + \boldsymbol{\beta}^T \mathbf{x}_j + \epsilon_j$$

Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Visualizing Effects of Reflectance

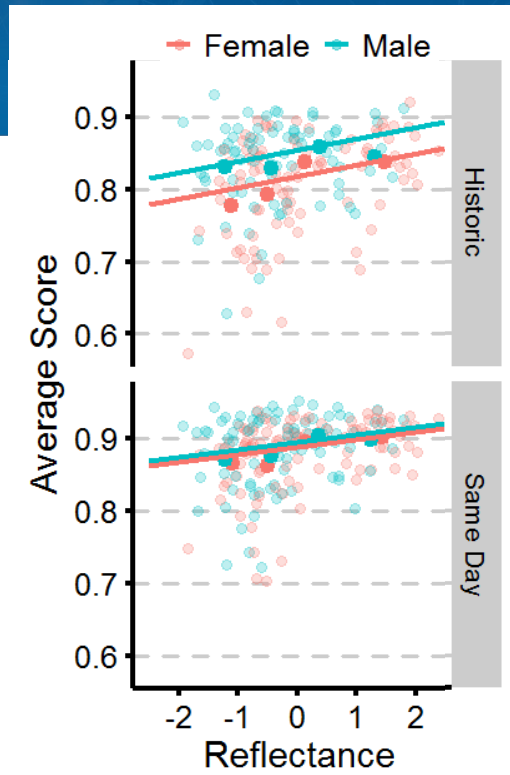


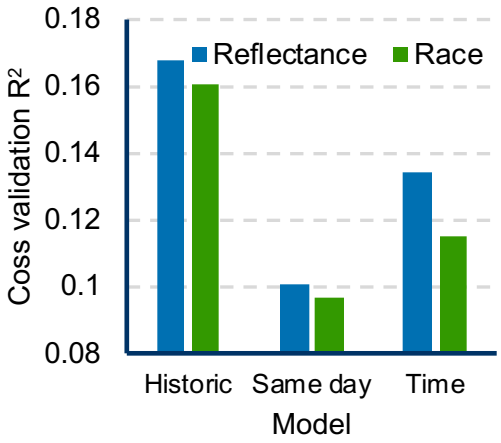
Statistical Modeling Results

- *Reflectance* was included in all three optimal models.
  - Performance was worse for volunteers with darker skin:
    - Lower similarity scores.
    - Slower transaction times.

    Note slope of regression lines.
- Unlike reflectance, *Gender* was NOT included in the Same Day model.

  Note overlapping regression lines in Same Day panel.
- Reflectance was a better predictor of performance than self-reported Race labels for all three models.
  - Confirmed with R^2 values from 10-fold cross-validation.
- Other demographic covariates were also selected in all optimal models, e.g. *Eyewear*.
  - Performance was worse for volunteers reporting *Eyewear*:
    - Lower similarity scores.
    - Slower transaction times.





Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Statistical Modeling Approach – Mixed Effects Model

Start with optimal model for average response variables

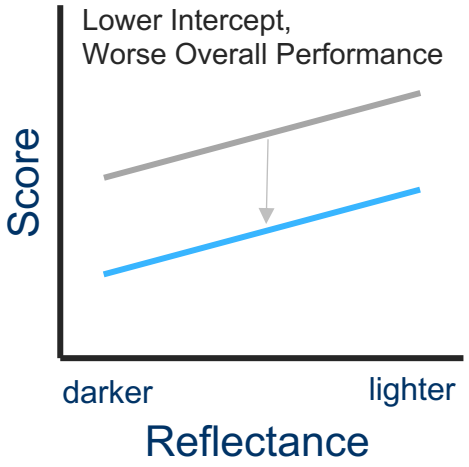
$$\mathbf{x}_j = [x_{1,j}, x_{2,j}, \ldots x_{k-1,j}]$$
$$\boldsymbol{\beta} = [\beta_1 \ , \beta_2 \ , \ldots \beta_{k-1} \ ]$$
$$\bar{\Theta}_j = \boxed{\beta_0 + \boldsymbol{\beta}^T \mathbf{x}_j} + \epsilon_j$$

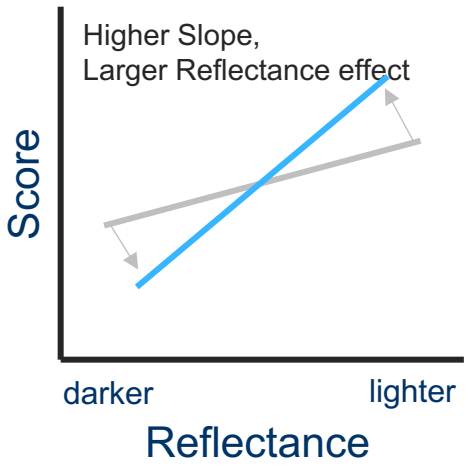
Add system-specific intercepts/slopes that reduce AIC fitting system-specific response variables

$$\mathbf{y}_j = [y_{1,j}, y_{2,j}, \ldots y_{m,j}]$$
$$\boldsymbol{\beta}_s = [\beta_{1,s}, \beta_{2,s}, \ldots \beta_{m,s}]$$
$$\Theta_{j,s} = \boxed{\beta_0 + \boldsymbol{\beta}^T \mathbf{x}_j} + \boxed{\beta_{0,s} + \boldsymbol{\beta}_s^T \mathbf{y}_j} + \epsilon_j + \gamma_s$$

Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Visualizing System-specific Slope and Intercept

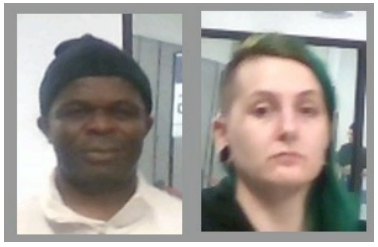




Acquisition System-Specific Effects

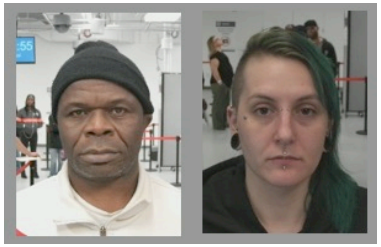
- Similarity scores varied significantly across acquisition systems.
  - Significant system-specific *Intercept*.
  - Significant system-specific *Reflectance* slope.
- Net effect of *Reflectance* was higher for systems with lower average similarity scores (intercepts).
- The magnitude of score variation with reflectance was similar to the magnitude of score variation across acquisition systems.

System 07

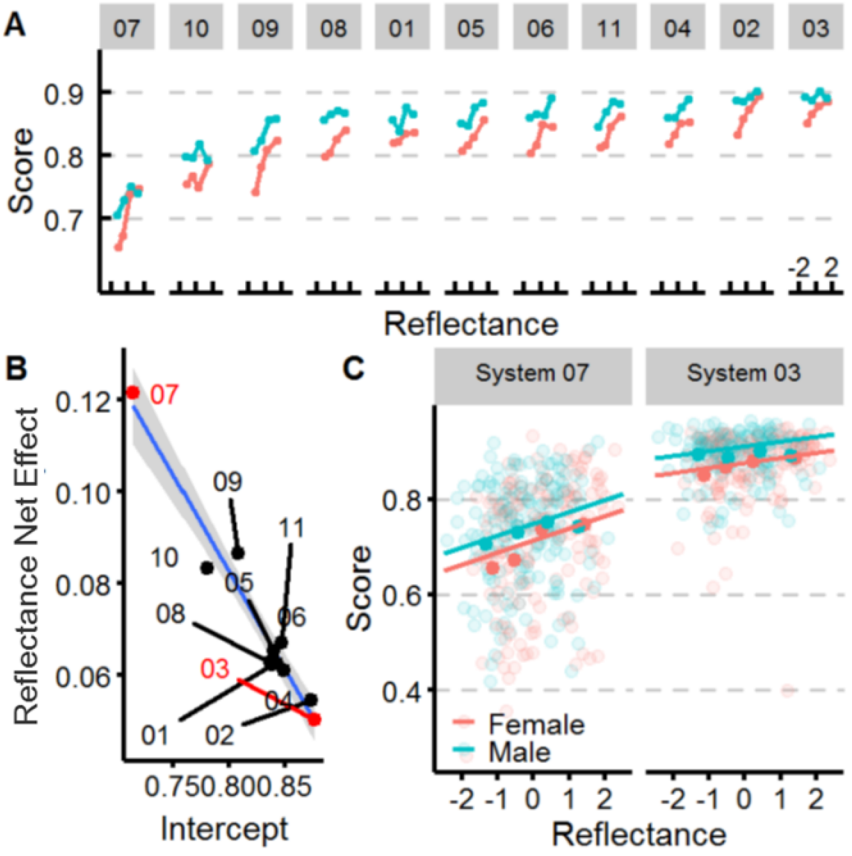


0.658 0.802

System 03



0.930 0.882



Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Brief Note on Fitzpatrick Skin Types

- Fitzpatrick Skin Types (FST).
  - Developed primarily to assess risk of *skin-cancer*.
  - Subjective determination via *survey instrument* or *structured interview*.

- Physician-reported FST correlates strongly with *Race*
  - $r = 0.55$
- Skin reflectance measures correlate weakly with *Race*
  - $r = 0.22$
- Raises the possibility of subjectivity in physicians' assessment
  - Tendency to classify African Americans into skin types IV-VI.

- Relating FST to biometric performance is problematic:
  - May give different results depending on how FST is obtained (self-reported or physician-reported).

FST I
Always burns, never tans (palest; freckles).
FST II
Usually burns, tans minimally.
FST III
Sometimes mild burn, tans uniformly.
FST IV
Burns minimally, always tans well (moderate brown).
FST V
Very rarely burns, tans very easily (dark brown).
FST VI
Never burns (deeply pigmented dark brown to darkest brown).

Pichon et al. Measuring Skin Cancer Risk in African Americans: Is the Fitzpatrick Skin Type Classification Scale Culturally Sensitive. Ethnicity and Disease. 2010: 20.

Chan et al. Assessing the role of race in quantitative measures of skin pigmentation and clinical assessments of photosensitivity. J Am Acad Dermatol. 2005: 52(4).

Conclusions

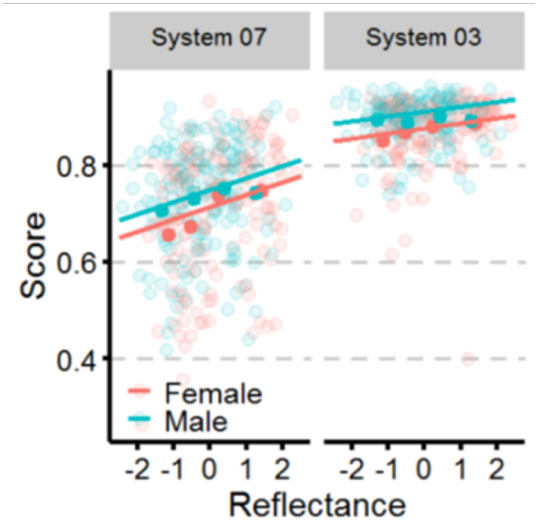
- Demographics influence both the accuracy and throughput of commercial biometric face acquisition systems.
  - Skin reflectance has the greatest effect on average biometric performance.

- The choice of acquisition systems matters:
  - Women with darker skin using a superior system are more likely to match than men with lighter skin using an inferior system.
  - Superior systems are more accommodating for people with lower skin *Reflectance*.
  - Variation in pose, motion blur, and lower contrast of probe images may contribute.

- Not all demographic effects have the same underlying cause:
  - Effects of Reflectance likely due to intrinsic properties of the face.
  - Effects of Gender may be related to differences in behavior.

- Advantages of using skin *Reflectance* as a covariate:
  - It is a *phenotype* and not subjective.
  - A better predictor of biometric performance than self-reported *Race*.
  - Can be obtained automatically from enrollment images.

- Future directions:
  - Examine additional matching algorithms.
  - Examine the imposter distribution.
  - What is the right phenotype or behavior for *Gender*?



Cook, Howard, Sirotin, Tipton, and Vemury. Fixed and Varying Effects of Demographic Factors on the Performance of Eleven Commercial Facial Recognition Systems. T-BIOM. *Under Review.*

Thank You!

We are hiring:

Email: info@mdtf.org

For more information about Biometric Technology Rallies:

Email: peoplescreening@hq.dhs.gov

Visit: http://mdtf.org