Quantifying Biometric Permanence Using Operational Data

Longitudinal Analysis of Comparison Scores

Patrick Grother, James Matey, George Quinn + Elham Tabassi
Information Access Division
National Institute of Standards and Technology

IBPC 2014 :: NIST EAB Workshop on Lifetime Biometrics
Tuesday, April 4, 2014
Ageing

Dwight D Eisenhower

ALGORITHM E20A
- 0.647
- 0.601
- 0.599
- 0.579

ALGORITHM J20A
- 0.595
- 0.578
- 0.565
- 0.548

Green indicates successful 1:1 authentication at FMR = 0.001.
Red indicates failure.
Ageing

Brad Wing

| ALGORITHM E20A | 0.617 | 0.578 | 0.532 | 0.541 |
| ALGORITHM J20A | 0.589 | 0.587 | 0.579 | 0.569 |

**Green** indicates successful 1:1 authentication at FMR = 0.001.

**Red** indicates failure.

**THE GOAL, SHARED WITH OTHER STUDIES IN THE FIELD, TO DETERMINE IF THERE’S AN ANALOG OF THIS FOR IRIS – IRREVERSIBLE CHANGE TO THE IRIS TEXTURE**
Individual iris recognition HDs over time

- Often, visually flat
- Considerable variance within eye
- Considerable variance between eyes
- Irregular sampling
- Imbalanced sampling
- Mixed effects models
  - Population part
  - Individual part

TRAJECTORIES INDICATE HETEROGENEITY – INTERCEPTS (AND GRADIENTS) VARY WITH QUALITY OF THE ENROLLMENT IMAGE cf. DODDINGTON’s ZOO
Nexus, Frequent Traveler Program

» Positive ID
  • Usually token-less
  • 1:FIRST iris

» Pop: US/CA + Perm res.
  • Motivated frequent travelers
  • US/CA air, land, sea

» Equipment
  • Operational since 2002, Daugman alg, refresh, c. 2013.
  • Panasonic BM-330ET camera
  • LG Cameras (removed c. 2007)

» As provided to NIST
  • 7.7 million log entries
  • 450K subjects
  • 680K eyes
  • Panasonic + LG cameras

» As used by NIST here
  • 1973761 log entries, from 29654 left eyes, those with 10 or more transactions, over at least 1460 days
  • Panasonic BM-330ET only
Quantifying permanence via mixed-effects regression

Model for the j-th score from the i-th eye

$$HD_{ij} = \pi_0i + \pi_1iT_{ij} + \epsilon_{ij}$$

Intercept is sum of population average term, the fixed effect, and an eye-specific random effect

$$\pi_0i = \gamma_{00} + \psi_0i$$

Slope is sum of population average term, the fixed effect, and an eye-specific random effect

$$\pi_1i = \gamma_{10} + \psi_{1i}$$

Subject to assumptions:

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

$$[\psi_0i, \psi_{1i}] \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2\end{bmatrix}\right)$$

Permanence stated by the population wide rate at which Hamming Distances are increasing.

MIXED EFFECTS MODEL RESPECT IDENTITY INFORMATION. SIMPLE LINEAR REGRESSION, IN YELLOW, DOES NOT
Accuracy vs. dilation and dilation change

The heatmap surface is a bowl, not a V-shaped valley.

The next slide introduces one model of this separable function.
Accuracy vs. dilation and dilation change

- Dilation is an explanatory variable in regression.
- Include two orthogonal terms:
  - Dilation difference $D_v$
  - Dilation magnitude $D_u$

Model this as a “quadratic bowl” $\text{POLY}(D_u, 2) + \text{POLY}(D_v, 2)$.
There are alternatives – function appears separable so $F(D_1)F(D_2)$. 
Model habituation too? It affects scores*

Habituation ... acquisition of a motor skill involving learning of an internal model of the dynamics of the task+.

**A:** For the “i”-th eye, does the mean time between captures explain observed Hamming distances

\[
\theta_i = \frac{1}{n_i} \sum_{j=2}^{n_i} T_{ij} - T_{ij-1}
\]

**B:** The model could be extended to capture learning and “muscle memory” via memory of recent experience

\[
\phi_{ij} = e^{-\frac{T_{ij} - T_{ij-1}}{\tau}}
\]

FURTHER RESEARCH NEEDED


+ Adapted from Reza Shadmehr* and Henry H. Holcomb, Neural Correlates of Motor Memory Consolidation, Science, Vol 277, 1997-AUG-07
**MODEL A: Unconditional growth model (without dilation)**

\[ HD_{ij} = \pi_0 + \pi_1 T_{ij} + \epsilon_{ij} \]

\[ \pi_1 = 3 \times 10^{-8} \quad \text{with } p = 0.8 \]

No detectable increase in Hamming distance

---

**MODEL B: Unconditional growth model with quadratic dilation terms**

\[ HD_{ij} = \pi_0 + \pi_1 T_{ij} + \pi_2 D_u + \pi_3 D_u^2 + \pi_4 D_v + \pi_4 D_v^2 + \epsilon_{ij} \]

\[ \pi_1 = 1 \times 10^{-6} \quad \text{with } p = 0 \]

Hamming distance increases by 0.004 per decade

---

**MODEL C: Unconditional growth model with habituation terms**

\[ HD_{ij} = \pi_0 + \pi_1 T_{ij} + \pi_2 \theta_{ij} + \pi_3 \phi_{ij} + \epsilon_{ij} \]

\[ \pi_1 = -3 \times 10^{-7} \quad \text{with } p = 0.01 \]

Hamming distance decreases!

Further modelling needed
Model Validation: Satisfying Normality Assumptions

\[ \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2) \]

\[ [\psi_{0i}, \psi_{1i}] \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{10}^2 & \sigma_1^2 \end{bmatrix} \right) \]

Level 1 :: Intra-eye, \( \varepsilon_{ij} \)

Distribution of the residuals

Distribution of the BLUPs

Intercepts, \( \psi_{0i} \)

Gradients, \( \psi_{1i} \)

Distribution of the BLUPs
Dilation = Pupil radius / Iris Radius

- Pupil size decreases by 0.2 mm per decade under office like illumination [WINN94]
- Iris size decreases by 0.13mm per decade [HALL04]

OVER 7 YEARS, DILATION IS VISUALLY FLAT, THE DOWNWARD TREND OF PREVIOUS SLIDE IS DOMINATED BY “NOISE” ASSOCIATED WITH AMBIENT LIGHT, PHYSIOLOGY, MOOD, ETC.
Longitudinal Dilation Change

Eye-specific dilation change trajectory

\[ D_{ij} = \pi_{0i} + \pi_{1i} T_{ij} + \epsilon_{ij} \]

And the coefficients are “fixed effects” + “random effects”

\[ \pi_{0i} = \gamma_{00} + \psi_{0i} \]
\[ \pi_{1i} = \gamma_{10} + \psi_{1i} \]

Population average intercept “fixed effect”

Individual eye intercept “random effect”

Population average growth rate “fixed effect”

Individual eye growth rate “random effect”

\[ \epsilon_{ij} \sim N(0, \sigma_{\epsilon}^2) \]
\[ \begin{bmatrix} \psi_{0i} \\ \psi_{1i} \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{0}^2 & \sigma_{01}^2 \\ \sigma_{10}^2 & \sigma_{1}^2 \end{bmatrix} \right) \]

MIXED EFFECTS MODEL QUANTIFIES “NOISE” AND TREND
Seasonal Dilation Variation

Dataset includes integer month of capture.

\[ D_{ij} = \pi_{0i} + \pi_{1i} T_{ij} + \pi_{2i} \cos \left( \frac{2\pi (m_{ij} - 1)}{12} \right) + \epsilon_{ij} \]

\[ \pi_{0i} = \gamma_{00} + \psi_{0i} \]

\[ \pi_{1i} = \gamma_{10} + \psi_{1i} \]

\[ \pi_{2i} = \gamma_{20} \]

Likely cause: Length of day effects, travelers exposed to outdoor lighting.
Latitude ~45 degrees N.
Magnitudes of NEXUS Dilation Variation

- Population average
  - Fixed effect $\pi_{0i} = 0.4344 \pm 0.0003$

- Inter-eye variation (between people)
  - Standard deviation of $\psi_{0i} = 0.066$

- Intra-eye variation (within eye)
  - Standard deviation of $\varepsilon_{ij} = 0.047$

- Seasonal term
  - Magnitude of $\pi_{2i} = 0.00667 \pm 0.00005$

- Elapsed time
  - Magnitude of $\pi_{1i} = -0.00143 \pm 0.00004$ per year

\[ x_{32} \text{ so about 3 decades for magnitude of permanent constriction to be comparable with short term changes} \]

\[ x_{5} \]
An Explanation For the Notre Dame Results

1. Observed genuine comparison score distributions shift with time.
2. Observed dilation distributions shift also.
3. Dilation differences degrade comparison scores.
Dilation change due to long term pupil constriction would be approx. -0.004

Source: NIST application of three commercial iris algorithms to ND images used in Fenker et al.
Adjusting ND Scores for Dilation

\[ d'_{ij} = d_{ij} - \beta_{i2} \Delta D_{ij} \]

- **Adjusted dis-similarity score**
- **Dis-similarity score from comparison of the j-th pair of images from the i-th individual eye**
- **Pupil dilation difference in the j-th pair of images**
- **Coefficient applicable to all scores from the i-th eye**

**THIS REMOVES AN AMOUNT PROPORTIONAL TO SOLELY THE DILATION DIFFERENCE, OVER TWO YEARS, THE PUPIL CONSTRUCTION IS NEGLIGIBLE**
THIS ADVERSE SHIFT IN THE GENUINE DISTRIBUTION THRU TIME LARGELY DISAPPEARS ONCE DILATION DIFFERENCES ARE ACCOUNTED FOR
Raw, Adjusted Mate Score Distributions vs. $\tau$

False Non-match Rate FNMR($\tau$)

Images Notre Dame 08-10

THIS RESULT, PUBLISHED PREVIOUSLY IN IREX VI, HOLDS FOR OTHER ALGORITHMS TOO
1. Iris identification algorithms short circuit distance computations for speed.
2. Distances above 0.27 are not computed and not returned – the search returns no enrolled identity.
But the NEXUS data is thresholded at HD = 0.27

- The system doesn’t return scores above $\tau \sim 0.27$
  - when unenrolled eyes presented
  - when poor images are collected
  - if the camera is defective
  - if the iris has changed, or the cornea
  - This is done to expedite search

- Regression is potentially flawed by this system-specific feature:
  - Suppose scores were samples from a non-stationary distribution
Notes on why truncation is not that influential

1. Estimating the center of the distribution is not that sensitive to tail truncation
   - Analytic results
   - Empirical results

2. Nature of what goes in the tail
   - Mostly due to: motion blur, occlusion, gaze angle, specular on boundary, dilation

p.76 IBG’s exhaustive report on the ITIRT iris recognition trial for the camera used in NEXUS
   - 3% best of three attempts T-FRR
   - 8% single image FNMR

*M. Thieme, Independent Testing of Iris Recognition Technology Final Report, IBG, May ’05*
For the whole population, HD degrades by $1 \times 10^{-6}$ per day

- For 15457 “high achievers”, eyes that have $0.075 \leq \text{mean HD} \leq 0.135$, the HD degrades by $0.3 \times 10^{-6} \text{ day}^{-1}$

- For 7518 eyes that never produce a HD above 0.21, HD degrades by $0.2 \times 10^{-6} \text{ day}^{-1}$

MIXED EFFECTS MODELS HEED IDENTITY:

ANALOGY: USING A 1.8M TAPE MEASURE DOESN’T PRECLUDE MEASURING HEIGHT OF SHORTER INDIVIDUALS
Face Ageing

Longitudinal Analysis of FRVT Scores Derived from Mugshot Images
Comparative evaluation of FR algorithms: Resistance to time lapse

» Vanilla mixed-effects model, without additional explanatory variables
» Problem: Scores exist on proprietary ranges, with little ability to interpret.
  • Option 1: Express growth rate as number of years before the mean genuine distribution would increase to, say, the FNMR = 0.1 threshold.
  • Option 2: Use z-norm on scores, express growth as “number of standard deviations per year”.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>$\gamma_{10} = \text{fixed effect growth rate (yr}^{-1})$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A30A</td>
<td>-0.127 ± 0.002</td>
<td></td>
</tr>
<tr>
<td>B30A</td>
<td>-0.172 ± 0.002</td>
<td></td>
</tr>
<tr>
<td>E30A</td>
<td>-0.129 ± 0.002</td>
<td></td>
</tr>
<tr>
<td>D30A</td>
<td>-0.162 ± 0.002</td>
<td>Always zero</td>
</tr>
<tr>
<td>J30A</td>
<td>-0.149 ± 0.002</td>
<td></td>
</tr>
</tbody>
</table>

!! Work in progress !!

MIXED EFFECT MODELS CAN INVOLVE:
Time invariant covariates:
- > Sex, race
Time varying covariates:
- > Pose, age
Conclusions

» For longitudinal analysis, mixed effects regression
  • Is appropriate for longitudinal analysis of imbalanced, irregular, auto-correlated data, from individuals (eyes) with heterogeneous responses.
  • has been developed independently by (at least) NIST (IREX VI) and MSU (Soweon Yoon)

» Permanence
  • For iris: No detectable population-wide shifts in scores in NEXUS data
  • For face: Provisional work implies measureable shifts in genuine scores.

» Habituation
  • For iris, frequency of use and time-since-last-use give improved scores

» Operational data
  • It’s volume affords excellent opportunities to detect and quantify effects
  • Is useful even in “logged” form, without images.

» Dilation change
  • Natural short term variance > seasonal-related variation > the first longitudinal estimate of dilation change (assoc. with pupil constriction)
  • Trend is barely observable in iris-ageing studies, so does not present a co-linearity hazard in regression analyses
Thanks