Adding value to CCTV: issues in testing of facial recognition systems

Tony Mansfield on behalf of Marek Rejman-Greene
2 April 2014
Final published version
We know AFR works
... but would it work here?
... or here?
### CCTV scales of operation - manual

<table>
<thead>
<tr>
<th></th>
<th>Monitor</th>
<th>Detect</th>
<th>Recognise</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>5%</td>
<td>10%</td>
<td>50%</td>
</tr>
</tbody>
</table>

- **Identify**: 120%
- **Inspect**: 250%

25 pixels

50 pixels
Back in the late 1990s

Plate 1 Mandrake face recognition system in use by the London Borough of Newham
BKA tests

Der Erfassungsbereich der Kameras
(Foto: BKA)

Die Kameras im Mainzer HBF
(Foto: BKA)
Quality issues in AFR/CCTV footage and watchlist imagery
... sometimes video helps

Frame number
### Not just the technology …

<table>
<thead>
<tr>
<th>Technology</th>
<th>Systems &amp; Processes</th>
<th>Subjects</th>
<th>Society</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Capture Subsystem</td>
<td>Data Storage Subsystem</td>
<td>Comparison Subsystem</td>
<td>Decision Subsystem</td>
</tr>
<tr>
<td>Biometric Characteristics</td>
<td>Enrollment Database</td>
<td>Reference</td>
<td>Comparison</td>
</tr>
<tr>
<td>Data Storage</td>
<td>Enrolment</td>
<td>Reference</td>
<td>Match/Non-match</td>
</tr>
<tr>
<td>Decision Subsystem</td>
<td>Candidate List</td>
<td>Identified</td>
<td>Possible matches found!</td>
</tr>
</tbody>
</table>

**Possible matches found!**

---

**Technology Systems & Processes Society**

**Biometric**

**Characteristics**

**Data Storage**

**Data Subsystem**

**Data Capture**

**Subsystem**

**Comparison**

**Score(s)**

**Match?**

**Verified?**

**Captured Biometric Sample**

**Reference**

**Decision**

**Policy**

**Features**

**Re-capture**

**Threshold**

**Comparison Subsystem**

**Segmentation**

**Feature Extraction**

**Enrolment**

**Database**

**Decision Subsystem**

**Candidate?**

**Identified?**

**Reference Creation**

**Biometric Claim**

**Enrolment Verification Identification**

**Outcome**

**Outcome**

**Presentation Sensor**

**New HD CCTV puts human rights at risk**
CAST contribution

• Datasets representing ‘typical’ situations in a crowded environment – public transport hub
• Aims:
  – how well could ‘standard’ AFR algorithms work in unconstrained environments?
  – understand the role of other system components
    • e.g. the operator working with unfamiliar people
  – provide technology developers and researchers with resources to innovate
• FRL2011 dataset available on license
  – 3 technology suppliers have been assessed/reported
## Facial Recognition Library 2011

### Cameras

<table>
<thead>
<tr>
<th>HD #</th>
<th>SD</th>
<th>HD Camera</th>
<th>HD fps</th>
<th>Description</th>
<th>Lighting Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>2 Mp mono</td>
<td>5</td>
<td>Entrance to station</td>
<td>Highly backlit</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>2 Mp colour</td>
<td>5</td>
<td>Up escalator</td>
<td>250% view</td>
</tr>
<tr>
<td>3</td>
<td>Y</td>
<td>4 Mp mono</td>
<td>7</td>
<td>Walkway past shops</td>
<td>Walking at angle</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>2 Mp mono</td>
<td>6</td>
<td>At ticket barriers</td>
<td>Highly backlit</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>2 Mp mono</td>
<td>13</td>
<td>Walking past shops, camera at an angle</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>2 Mp mono</td>
<td>18</td>
<td>Ramp up from platform</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>2 Mp colour</td>
<td>20</td>
<td>Ramp down</td>
<td>200% view</td>
</tr>
<tr>
<td>8</td>
<td>Y</td>
<td>2 Mp colour</td>
<td>10</td>
<td>Train platform</td>
<td>50% view</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>11 Mp colour</td>
<td>1.5</td>
<td>Station concourse</td>
<td>20% view</td>
</tr>
<tr>
<td>10</td>
<td>Y</td>
<td>2 Mp mono</td>
<td>6</td>
<td>Taxi drop-off</td>
<td>At 7.5-40 m distance</td>
</tr>
</tbody>
</table>
FRL2011 specification

• 100 watchlist subjects
  – of whom 50 walk through field of view of SD/HD CCTV cameras
  – M:F ratio: 49:51
  – Ethnicity:
    • approximate balance of light and dark Caucasian, Afro-Caribbean, Asian, Oriental, Arab and mixed race
  – Age: 15-64
    • with some preponderance of 20-29 (34%) and 30-39 (23%)

• Additional imagery (number of subjects)
  – 81 multi-pose images (96, 78 in IR)
  – User provided photos (97)
  – Informal photos (94)
  – Autocue speech (25)

• 10 x 1hr HD footage, with 7 x 1hr SD footage

Home Office
Multi-pose images
Learnings so far

• Manageable watchlist
  – Hundreds
  – Quality – we’ve used FIND compliant (1 MP)
• Need for clear definitions
  – What constitutes a missed recognition?
    • Detection Opportunity
  – Recognition and false alert rates
    • On a frame by frame basis, traverse through FOV, per hour?
• False Alert Rate
  – For one supplier, many of the false alerts from one individual on watchlist
Thank you for your attention

marek.rejman-greene@homeoffice.gsi.gov.uk