Computer vision techniques for video surveillance

Huiyu Zhou, Ph.D.

January, 2016
London riots: Tottenham violence, 5 August, 2011
- Motivation
- Age classification
- Gender classification
- Behaviour analysis
- Summary
• >4,000,000 cameras, UK, 2014.
• >4,000,000 cameras, UK, 2014.
• Major concern: crime in public places.
• >4,000,000 cameras, UK, 2014.
• Major concern: crime in public places.
• \(~70\%\) of offenders are young adolescent males [1].

• >4,000,000 cameras, UK, 2014.
• Major concern: crime in public places.
• ~70% of offenders are young adolescent males [1].
• Our research focus: what is the age/gender of the target? What is s/he doing (behaviour)?

• Motivation
• Age classification
• Gender classification
• Behaviour analysis
• Summary
Challenges

• Intrapersonal variation: anatomical changes on faces.

• Interpersonal variation: individual evolution of faces.

Tony Blair: 10+, 30+, 50+ (left to right)

Bill Gates: 10+, 20+, 50+ (left to right)
Whole picture of our system

- Original images
- Adaptive Difference of Gaussian (DoG)
- Radon Transform (RT): $x$ – intensity, $y$ – bins
- Feature selection/SVM classification
Feature extraction: Adaptive DoG

- **Benefits**
  - To reduce the effects of rapid intensity changes on faces

- **Adaptive DoG filtering:**
  - Subtracting two convolutions: \( \sigma_1 = \sigma_0/8 \), \( \sigma_2 = \sigma_0/16 \)
  - Gamma correction
  - Contrast equalisation

Contrast equalisation (x - greyscale, y - pixel no.)
Feature extraction: why Radon Transform?

- In-plane rotation invariant
- Detecting facial curves (e.g. wrinkles)

1-D illustration of Radon transform at different rotations (x – angle/deg, y – projection displacement).
Feature extraction: how I use Radon Transform?

- Similarity measured by Radon projection correlation distance [2].

2-D Radon transform of different images (x – angle/deg, y – projection displacement)

What is scaling?
- A scheme to select the hyper-parameters (SVM) for the least generalisation error

Scaling SVM
- Continuously update kernel $K$ and weight $w$

Classification results of parameter set 1
Classification results of parameter set 2
Illustration of scaling SVM
Experimental work: set-up

**Objective:** to separate teenagers and adults

**Comparisons:** our system (DRTP) against 5-fold SVM with

- a) PCA (principal component analysis)
- b) LBP (linear binary pattern)
- c) HOG (histogram of oriented gradients)
- d) DRT (DoG/RT/no feature selection)
- e) DRTC (DoG/RT/feature selection)
- f) HOGSS (HOG with feature selection)

**Test databases:** FG-NET and MORTH

Examples from the two databases
Experimental work: MORTH dataset

Images of different ages

PCA reconstruction of 50 eigenvectors

LBP (x – bins, y – numbers)

HOG (x - feature index, y – gradient values)

Proposed (x - feature index, y – intensity pixels)
Experimental work: MORTH dataset

Feature selection outcomes

Classification by seven algorithms
• Motivation
• Age classification
• Gender classification
• Behaviour analysis
• Summary
Challenges

• Research categories: Face and full body based

• Face based: require frontal faces and affected by occlusions [3]

Challenges – demo of walking patterns

- Full body based: gaits
- Side-view problem

Courtesy of Biomotion Lab, Canada
Our approach

1) Combination of facial and full body measurements
1) Combination of facial and full body measurements

2) Face channel: face detection $\rightarrow$ PCA features
Face detection and PCA
Our approach

1) Combination of face and full body measurements

2) Face channel: face detection $\rightarrow$ PCA features

3) Full body channel: background subtraction $\rightarrow$ PiHOG features
Background subtraction and PiHOG
Our approach

1) Combination of face and full body measurements

2) Face channel: face detection $\rightarrow$ PCA features

3) Full body channel: background subtraction $\rightarrow$ PiHOG features

4) “EntropyBoost” classifier $\rightarrow$ probability estimate in each channel
Our approach

1) Combination of face and full body measurements
2) Face channel: face detection $\rightarrow$ PCA features
3) Full body channel: background subtraction $\rightarrow$ PiHOG features
4) “EntropyBoost” classifier $\rightarrow$ probability estimate in each channel
5) Fusing two channels: score integration [4]

Demo video: gender classification
Gender classification errors of different systems: “CF” – face/body HOG features + SVM; “FP” – face PCA features + SVM; “BH” – body HOG features + SVM; “EF” – our system.

<table>
<thead>
<tr>
<th>Seqs.</th>
<th>Image No.</th>
<th>Ped. No.</th>
<th>Gender</th>
<th>PD</th>
<th>CF</th>
<th>FP</th>
<th>BH</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>763</td>
<td>1</td>
<td>F</td>
<td>672</td>
<td>16.4%</td>
<td>67.6%</td>
<td>17.8%</td>
<td>8.2%</td>
</tr>
<tr>
<td>2</td>
<td>805</td>
<td>2</td>
<td>F/M</td>
<td>730</td>
<td>15.4%</td>
<td>95.2%</td>
<td>16.3%</td>
<td>11.6%</td>
</tr>
<tr>
<td>3</td>
<td>623</td>
<td>2</td>
<td>M/M</td>
<td>571</td>
<td>13.3%</td>
<td>95.9%</td>
<td>14.2%</td>
<td>9.0%</td>
</tr>
<tr>
<td>4</td>
<td>108</td>
<td>1</td>
<td>F</td>
<td>101</td>
<td>15.9%</td>
<td>96.2%</td>
<td>16.5%</td>
<td>13.4%</td>
</tr>
<tr>
<td>5</td>
<td>2000</td>
<td>&gt;3</td>
<td>M/F</td>
<td>1554</td>
<td>17.4%</td>
<td>99.2%</td>
<td>19.1%</td>
<td>15.1%</td>
</tr>
<tr>
<td>6</td>
<td>1626</td>
<td>&gt;3</td>
<td>M/F</td>
<td>1019</td>
<td>19.8%</td>
<td>76.3%</td>
<td>19.6%</td>
<td>14.7%</td>
</tr>
</tbody>
</table>
• Motivation
• Age classification
• Gender classification
• Behaviour analysis
  – Human tracking (single and multiple cameras)
  – Trajectory clustering
  – Event reasoning
• Summary
• **Challenges**
  – Occlusions/pose or light changes

---

Single-camera human tracking
Challenges
– Occlusions/pose or light changes

Heterogeneous sensors
– Kalman filter based audio/visual data association scheme [5]

Demo video can be found at: http://sites.google.com/site/huiyujoe/

Particle filter
Graph matching
Audio Detection (TOA)
Our system
• Challenges
  – Occlusions/pose or light changes

• Heterogeneous sensors
  – Kalman filter based audio/visual data association scheme [5]

• Kernel estimation and local features
  – Effective combination of mean shift and SIFT features [6]

More results can be found at:
http://sites.google.com/site/huiyujoe/
Demo: Multi-camera human tracking

Simulated Annealing Particle Filter
Trajectory clustering – walking

Walking trajectories to be clustered
Clustering using individual features

(a) Actual walking trajectories  
(b) Distance difference features  
(c) Direction deviation features
Markov Chain Monte Carlo based clustering

(a) Ground truthed trajectories
(b) Proposed approach
• Motivation
• Age classification
• Gender classification
• Behaviour analysis
• Summary
• Automatic feature extraction and selection for age classification.

• Combining facial and full body measurements for gender classification.

• Behaviour analysis (ongoing): human tracking, trajectory clustering and event reasoning.
Acknowledgments

• Collaborators
  – Internal: colleagues in ECIT/CSIT…
  – External: BAE, Thales, Microsoft, IBM, Google, NIH, U. of London…

• Funding agencies
  – EPSRC
  – Invest NI
  – EU ICT
Thank you very much!

Q & A