Multi-Camera People Tracking and Re-Identification within a Crowd

Students: Ofer Geva & Matan Sela
Supervisor: Yair Moshe
In Association with: Mango DSP
Abstract

• Tracking systems are necessary for surveillance, traffic control and multiple computer vision applications
Problem Definition

Sparse Crowd

Medium Density Crowd

Dense Crowd
Part A - Single Camera Tracking Scheme

- Video Frames
  - Background Modeling & Pedestrian Segmentation
  - Occlusion/Reidentification Detector

- Appearance Tracking
- Motion Tracking (Appearance Learning)

- Marked Target
Background Modeling and Pedestrian Segmentation

Frames

Median

Last Frame

HoG Detector

$C_1 \times$

Motion Foreground

$C_2 \times$

HoG Confidence Function
Background Modeling and Pedestrian Segmentation

Improved Foreground Image
Background Modeling and Pedestrian Segmentation

Foreground Detection in scenes with massive moving background and shades
Single Camera Tracking
Motion Mode – Kalman Filter
Tracker Data – Appearance Model

- Sampling the target uniformly

- Extracting features from samples
  - Color features
  - Texture features
  - Spatial features

\[
\text{LBP} = 1 + 2 + 4 + 8 + 128 = 143
\]
Tracker Data – Appearance Model

- **Multi-Variable Kernel Density Estimation** – Estimating probability density function in the feature space:

\[
\hat{p}(z) = \frac{1}{N_p \sigma_1 \ldots \sigma_n} \sum_{i=1}^{N_p} \prod_{j=1}^{n} \kappa \left( \frac{z_j - s_{ij}}{\sigma_j} \right)
\]

- **KL Distance** – Calculating the similarity between a new detection and the model:

\[
D_{KL}(\hat{p}_b \mid \hat{p}_a) = \int \hat{p}_b(z) \log \frac{\hat{p}_b(z)}{\hat{p}_a(z)} \, dz
\]
Part A - Single Camera Appearance Mode

Blobs Search
Part A - Tracking Demonstration
Part B - Multiple Cameras Tracking Scheme
Part B - Multiple Cameras Tracking Scheme
3D Pedestrian Detector

- Blobs Analysis:
  - In each view, **image coordinates** of each blob’s lowest point are converted to **world coordinates**.
  - blobs’ lowest point is assumed to be in $Z = 0$ plane.
Camera Projection

- **Goal:** Estimating the 3D geometry of the scene from the 2D images
3D Pedestrian Detector

Camera Projection

\[
\begin{bmatrix}
X_{\text{im}} \\
Y_{\text{im}} \\
1
\end{bmatrix} = K \begin{bmatrix} R | t \end{bmatrix} \times \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\]

Intrinsic:
- Focal lengths
- Sensor axes skew
- Optical center

Extrinsic:
- Rotation matrix
- Translation vector
3D Pedestrian Detector
3D Pedestrian Detector

Starting with random blob from view with least occlusions
3D Pedestrian Detector

Nearest blob from next view is labeled as a part of the detection (group average location is calculated)
3D Pedestrian Detector

Nearest blob from next view is too far from current average location – not labeled
Labeled blobs are signed as a detection and deleted.
Labeled blobs are signed as a detection and deleted
3D Pedestrian Detector

Detection #1

View 1

View 2

View 3
3D Pedestrian Detector

Detection #1

View 1

View 2

View 3
3D Pedestrian Detector

Detection #1
3D Pedestrian Detector

Detection #1

Detection #2

View 1

View 2

View 3
Part B - Multiple Cameras Tracking Scheme
Greedy Data Association

Example:

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.70</td>
<td>0.20</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>0.65</td>
<td>0.80</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td>0.70</td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

World distance and KL-Distance are weighted for creating a score function for each Detection-Tracker pair.
## Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection 2</td>
<td>0.70</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td>Detection 3</td>
<td>0.65</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>Detection 3</td>
<td><strong>0.15</strong></td>
<td>0.70</td>
<td>0.55</td>
</tr>
</tbody>
</table>
# Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 3</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.70</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>0.15</strong></td>
<td>0.70</td>
<td></td>
<td>0.55</td>
</tr>
</tbody>
</table>
## Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>∞</td>
<td>0.20</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>∞</td>
<td>0.80</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
<td></td>
</tr>
</tbody>
</table>

The diagram illustrates the association between detections and trackers, with arrows indicating the assignment of detections to trackers based on the associated values.
## Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>∞</td>
<td>0.20</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>∞</td>
<td>0.80</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
<td></td>
</tr>
</tbody>
</table>

![Diagram showing Data Association](chart.png)
# Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Detection 2</th>
<th>Detection 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracker 1</td>
<td>Tracker 2</td>
<td>Tracker 3</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>

![Diagram showing data association](image_url)
### Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>Detection 2</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>0.75</td>
</tr>
<tr>
<td>Detection 3</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>

- Detection 1 is associated with Tracker 2
- Detection 3 is associated with Tracker 1
- Detection 3 also has a 0.75 matching score with Tracker 2

**Note:** The infinite symbols represent a non-assignable or non-matching status.
Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 3</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>0.75</td>
<td>∞</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
</tbody>
</table>
### Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Tracker 1</th>
<th>Tracker 2</th>
<th>Tracker 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td></td>
<td>∞</td>
</tr>
</tbody>
</table>

- Above Threshold
# Greedy Data Association

<table>
<thead>
<tr>
<th>Detection 1</th>
<th>Detection 2</th>
<th>Detection 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracker 1</td>
<td>Tracker 2</td>
<td>Tracker 3</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>0.75 (Above Threshold)</td>
</tr>
<tr>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
</tbody>
</table>

- Detection 1 is associated with Tracker 1.
- Detection 2 is associated with Tracker 2.
- Detection 3 is associated with Tracker 1 and Tracker 2.
- New Tracker is created.
Occlusions Segmentation

Frame #n – Real Foreground

Frame #n+1 – Estimated Foreground

Frame #n+1 – Real Foreground

Frame #n+1 – Occlusion Segmentation
Tracking Demonstration
Part B - Tracking Demonstration
Experimental Results

• System was mainly tested on PETS 2010 S2.L1 dataset

• Results were evaluated against single view (view #1) ground truth data

• The evaluation used the following metrics[8]:
  o MOTA (Multiple Object Tracking Accuracy)
  o MOTP (Multiple Object Tracking Precision)

Evaluation Results:
Conclusion

- A **multiple targets tracking system** has been developed and implemented
- **Novel pedestrian detection method** has been developed
- **Novel Tracking-by-Detection method** was developed – Trackers’ data is used for next frames pedestrian detection
- The solution involved **multi camera** input handling, using **homography** tools
- The system includes a re-identification capability, using a **probabilistic appearance model**
- The Algorithm was tested and evaluated using ground truth data according to CLEAR-MOT metrics