Learning People Detectors for Tracking in Crowded Scenes

Siyu Tang1  Mykhaylo Andriluka2  Anton Milan3  Konrad Schindler3  Stefan Roth2  Bernt Schiele3
1Max Planck Institute for Informatics 2TU Darmstadt 3ETH Zürich

Goal
• Detect and track all the people in the crowded street scenes

Joint detection

Structural learning for joint detection
• Given training images, learning the parameters of the joint detection model is formulated as the optimization problem [2]:
  \[
  \min_{\beta, \xi \geq 0} \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i, \\
  \text{subject to} \quad \max(\beta, \phi(I, y_i, b_i)) - \max(\beta, \phi(I, y, b)) \\
  \geq \Delta_{\text{voc}(y, y')} - \xi_i, \quad \forall y \in Y
  \]

Loss functions
• The detection with larger overlap with the ground truth bounding box has higher score than the detection with lower overlap with the ground truth bounding box
  \[
  \Delta_{\text{voc}(y, y')} = \begin{cases} 
  0, & \text{if } y' = y' = 1 \\
  1 - |y' = y'| \Delta_{\text{voc}(y, y')} \left( \frac{\Delta_{y} + \Delta_{b}}{\Delta_{y} + \Delta_{b}} \right), & \text{otherwise}
  \end{cases}
  \]

Experimental result
The same experiment setup with [1]

Learn people detectors for tracking

Design occlusion patterns
• Manually design regular occlusion combinations that appear frequently due to long-term occlusions and therefore most relevant for tracking

Experiments

Method | Roll Pose | MOTAP | MOTAP | MOTAP
--- | --- | --- | ---
Baseline detector | 81.7 91.3 79.3 % 74.1 %
Multi-target tracker | 81.7 91.3 79.3 % 74.1 %
Synthetic training image pool | 81.7 91.3 79.3 % 74.1 %

Mining sequence
Output:
• Joint detector optimized for multi-target tracking

1: run baseline detector on mining sequence
2: run target tracker on mining sequence, based on the detection result from baseline detector
3: repeat
4: collect missing recall from the tracking result
5: cluster occlusion patterns
6: generate training images for mined patterns
7: train a joint detector with new training images
8: run the joint detector on mining sequence
9: run the target tracker on mining sequence
10: until tracking results converge

Reference