# A Framework to Combine Multi-Object Tracking and Segmentation

Sehr Nadeem<sup>1</sup>, Dr. Anis ur Rahman<sup>1</sup>, Dr. Asad Anwar Butt<sup>1,2</sup>

<sup>1</sup>School of Electrical Engineering and Computer Science (SEECS), National University of Science and Technology (NUST), Islamabad, Pakistan

<sup>2</sup> National Institute of Standards and Technology (NIST), Gaithersburg, USA

### Correlation of Tracking and Segmentation



Multi-Object Tracking

Multi-Object Video Segmentation

To implement a framework for multi-object tracking and segmentation in which their results are jointly processed leading to improved accuracy in both.

#### Proposed Methodology Overview



# Multi-Object Tracking Module

- Tracking-by-detection framework
- Pre-trained pedestrian tracker
- Higher-order constraints like constant velocity
- Lagrangian relaxation

5 [13] Butt, A.A. & Collins, R.T. Multi-target tracking by Lagrangian relaxation to min-cost network flow. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, **2013**, 1846-1853

# Multi-Object Video Segmentation

- Negative and positive samples are clustered using k-means.
- Linear SVM is trained using Lab color channel
- Color and optical flow are used to assign the super-pixels to target

[24] Milan, A.; Leal-Taixé, L.; Schindler, K. & Reid, I. Joint tracking and segmentation of multiple

6 targets. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, **2015**, 5397-5406

Joint Processing: Updating Tracking Results

The multi-object tracking results are updated in two aspects:

- 1. Locations of bounding boxes
- 2. Adding missed targets

# Updating Tracking Results I: Location

- Correspondence established between segmentation and tracking bounding boxes
- Based on minimum difference in coordinates.

$$Penalty(B_i) = \sum_p ((p \in B_i, p \neq S_i) + (p \notin B_i, p = S_i))$$

 Only x coordinates and width updated as height is more accurate in tracking results.

## Updating Tracking Results I: Location cont.

Bounding Boxes extracted from segmentation results.







(Only detections which have corresponding results in the other module are considered here).

# Multi-Object Tracking Precision (MOTP)

	Initial Tracking Results	After Joint Processing
MOTP	61.31	64.88

## Updating Tracking Results II: Adding Missed Targets

- Targets with no corresponding detections in tracking results are added to the updated results.
- D ID is assigned of the spatially closest target in window of 10 frames.



The added target is correctly assigned ID 19 which can be verified from the next frame in which that target has been detected by the tracking module.







The added target is correctly assigned ID 17 in frame 7 which can be verified from frame 9 in which that target appears in the tracking results.



# Multi-object Tracking Accuracy (MOTA)

	Initial Tracking Results	After Joint Processing
<b>Correct Detections</b>	796	862
Number of Missed	194	95
False Positives	0	10
Number of Mismatches	18	41
ΜΟΤΑ	78.96	85.51

MOTA = I - <u>(false Positives + misses + mismatch)</u> \* 100 total number

## Joint Processing: Updating Segmentation Results

□ The segmentation results are updated in two aspects:

- 1. **ID Correction**
- 2. Adding Missed Targets

# Updating Segmentation Results I: ID correction

- Tracking module performs better in aspect of ID assignment
- Segmentation IDs are replaced with the IDs of their corresponding tracking results.



From frame 9 to 10 ID-2 gets mislabelled as ID-5. After updating this ID is corrected.



From frame 45 to 46 three targets get mislabelled. After updating these are corrected.

## Updating Segmentation Results II: Adding Missed Targets

- Missed targets are identified
- Edges of motion boundaries are computed
- Morphological closing is used to connect them and inside region is filled.
- Which is added to the segmentation results.
- The tracking IDs of the added targets are retained.



Top left is the original motion boundaries image and top right shows the edges extracted from it. The bottom right image shows final results. Frame 12: Frame 12:

## Adding Missed Targets to Segmentation Results

#### □ Frame 12:



#### **Initial Results**

#### After Updating

# Adding Missed Targets cont.

#### □ Frame **46**:



#### **Initial Results**

After Updating

Multi-Object Segmentation Evaluation

Evaluation Metrics Used:

- 1. Average percentage of misclassified foreground pixels per frame
- 2. Average number of false negatives per frame

# Multi-Object Segmentation Results

	Initial Segmentation	After Joint Processing
Average % of misclassified pixels	7.59	1.90
Average number of false negatives	14122	12045

### Comparison of Our Approach with Other Segmentation and Tracking Techniques

		<b>Tracking Results</b>			Segmentation Results			
		MOTA	MOTP	Precision	Recall	Precision	Recall	% of misclassified pixels
Approaches which incorporate both tracking and segmentation	Ours	85.5	64.9	0.98	0.90	0.98	0.62	1.90
	Y.Tian et al., 2016	91.7	62.4	0.96	0.94	-	-	3.88
	M. Keuper et al., 2016	83.3	77.3	0.97	0.85	0.62	0.64	-
Multi-Object Tracking Approaches	Nguyen et al., 2016	72.0	67.0	-	-			
	M.Wang et al., 2016	67.2	74.1	-	-	-	-	-
	S.Tang et al., 2016	82.3	76.3	0.88	0.94			

# Conclusion

- If one module is weak in one aspect, the other module has a better accuracy in that aspect.
- The tracking module was able to improve bounding box locations
- The segmentation module was able to improve ID assignment
- Missed detections were added to both the modules' results.

# Thank you!