Report

Introduction

This report details the development of a Multi-Object Tracking system. The project progressed through four stages: Single Object Tracking, IoU-based, Kalman-Guided, and Appearance-Aware. The system was tested on the `ADL-Rundle-6` sequence using YOLOv5 detections.

TP 1: Single Object Tracking

Objective: Implement a Kalman Filter for a single object.

- Implementation: A `KalmanFilter` class was created to estimate the state (x, y, v_x, v_y) of a centroid.
- Results: The filter successfully smoothed the trajectory of a detected object, predicting its position even when detection was noisy.

TP 2: IoU-based Tracking

Objective: Extend to Multiple Object Tracking using IoU and Hungarian Algorithm.

- Methodology:
  - Detections were loaded from `Yolov5s` results.
  - A similarity matrix was constructed using Intersection over Union (IoU).
  - The Hungarian algorithm (`scipy.optimize.linear_sum_assignment`) was used to assign detections to tracks optimally.
  - Track management included initializing new tracks for unmatched detections and deleting tracks after 5 missed frames.
- Observations: The tracker worked well for distinct objects but failed during occlusions, often assigning new IDs when objects reappeared.

TP 3: Kalman-Guided IoU Tracking

Objective: Improve association using Kalman Filter predictions.

- Methodology:
  - Integrated the `KalmanFilter` from TP1 into the tracker.
  - In each frame, the filter predicts the new centroid of each track.
  - The bounding box is shifted to this predicted position before calculating IoU with new detections.
- Matched tracks update the Kalman filter with the measured centroid.
- Improvements: The prediction step allows the tracker to "look ahead", maintaining association even if the object moves significantly or detection is slightly displaced. It reduced fragmentation of trajectories.

TP 4: Appearance-Aware Tracking (ReID)

Objective: Integrate Deep Learning-based Re-Identification.

- Methodology:
  - Used `reid_osnet_x025_market1501.onnx` for feature extraction.
  - Preprocessing: Detection patches were resized to 64×128, converted to RGB, and normalized (mean/std subtraction).
  - Association: A combined score S was defined:

    $$S = \alpha \cdot \text{IoU} + \beta \cdot \text{CosineSimilarity}$$

    with $\alpha = 0.5, \beta = 0.5$.
  - This score was maximized using the Hungarian algorithm.
- Results: The inclusion of visual features significantly improved robustness. The tracker could distinguish between spatially close objects if they looked different and recover identities after longer occlusions.

Challenges and Conclusion

- Format Parsing: Handling different `det.txt` formats (space vs comma separated) required robust parsing logic.
- ReID Preprocessing: ensuring the input to the ONNX model matched the training data (Market1501) was critical for meaningful feature vectors.
- Conclusion: The final system represents a robust MOT pipeline. While IoU provides fast geometric association, Kalman filtering adds temporal smoothness, and ReID provides identity persistence, resulting in a comprehensive tracking solution.