Summer Research ‘18-’19

Automated Workflows for Person Tracking Algorithm Evaluation

The Project

Why?

A significant challenge of academic research is developing good protocols for algorithm accuracy and speed. Additionally, the development of meaningful metrics that allow for fair comparison between algorithms is required in order to properly validate the effectiveness of a given approach.

What?

In accordance with the above and also to ensure eligibility for submission to the Multi-Object Tracking Challenge under the Duke MTMC category, it was decided that the performance of the Model Fusion Video Analytics System would be measured with that category's set of metrics.

How?

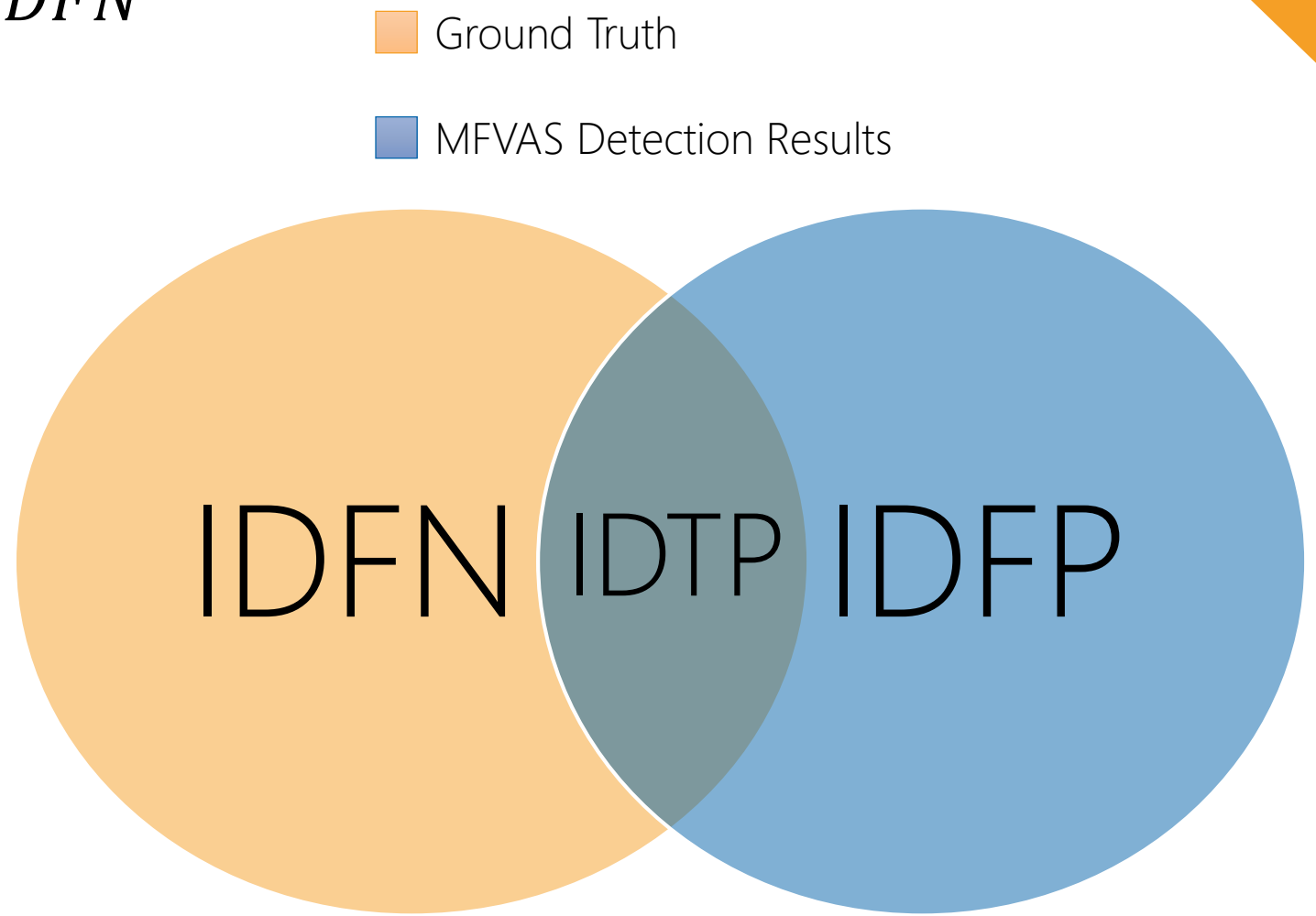
The creation of a script that compares the output of the Model Fusion Video Analytics System executed on a video sequence with the ground truth of that sequence.

The Metrics

$$IDF_1 = \frac{2 * IDTP}{2 * IDTP + IDFP + IDFN}$$

$$IDP = \frac{IDTP}{IDTP + IDFP}$$

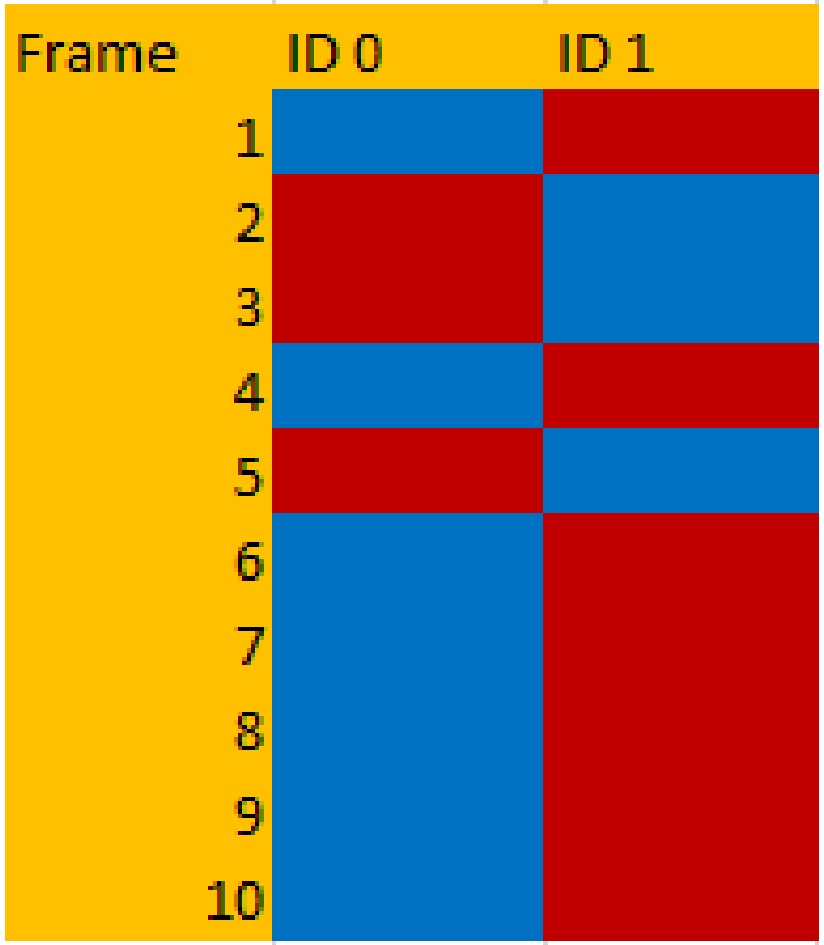
$$IDR = \frac{IDTP}{IDTP + IDFN}$$



The Metrics

Example for a Single Person Scenario

- Fragmentations: 3
- ID Switches: 4
- Mostly Tracked Targets: 1
- Mostly Lost Targets: 1



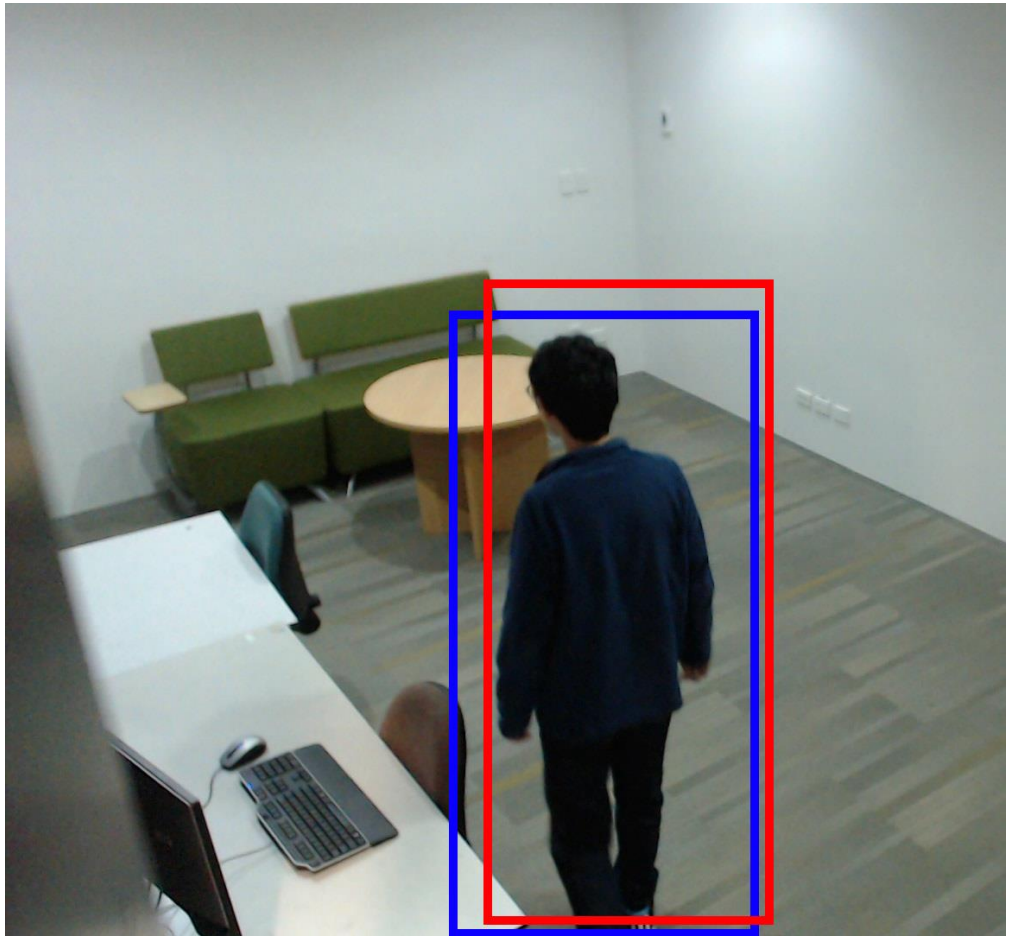
Detected

Not Detected

The Metrics

$$MOTA = 1 - \frac{IDFN + IDFP + FRAG}{IDTP}$$

$$MOTP = \frac{\sum_{n=0}^{N} IoU_n}{IDTP}$$



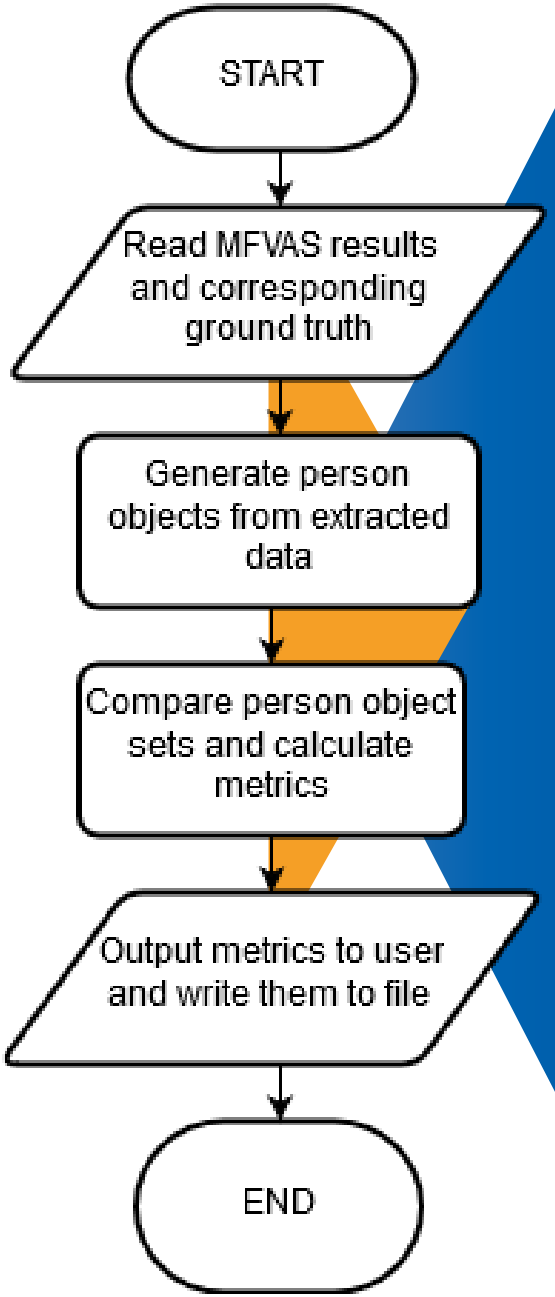
Ground Truth

MFVAS Detection Results

Program Design

At the highest level, the evaluation script works as such:

- Extract data from the results file output by MFVAS and the associated ground truth
- Process the data, creating instances of detected individuals, generating tracklets from their location data and grouping said individuals by camera
- Compare sets of person objects for each camera in the system by looking at the path of each individual and their bounding boxes in order to map MFVAS detected individuals to those in the ground truth

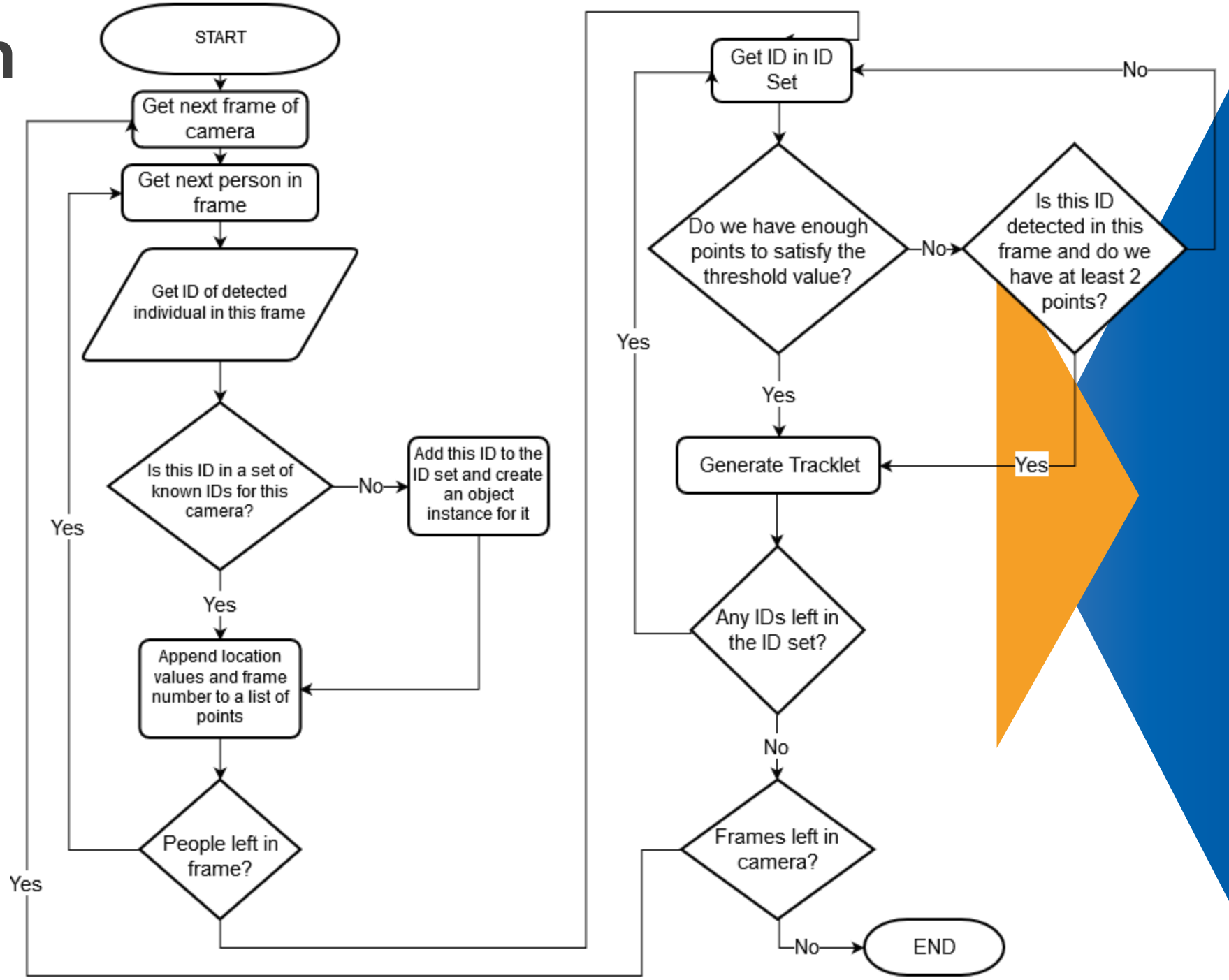


Program Design

Processing the data

Once data is extracted, it needs to be processed and turned into person objects.

- Each ID is added to a set of known IDs
- Points comprised of space-time data are accrued
- After a certain number of points is accrued for an ID, a tracklet is generated

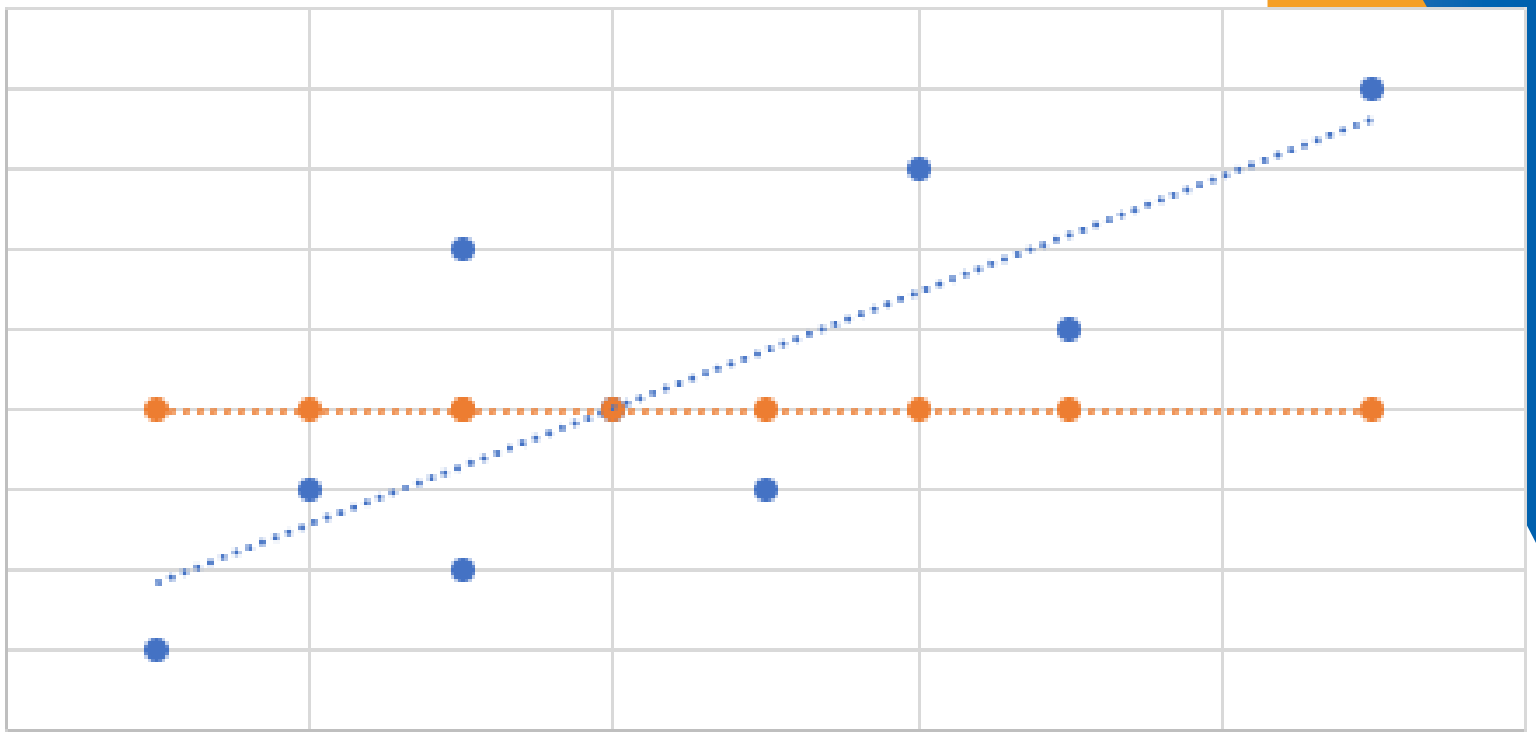


Program Design

Tracklet Definition and Generation

Tracklets are defined as ((x1,y1,t1),(x2,y2,t2)).
Generated using a best fit line equation.
The resultant line is then constrained in time in a way that maximises length.

$$m = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{X})^2}$$



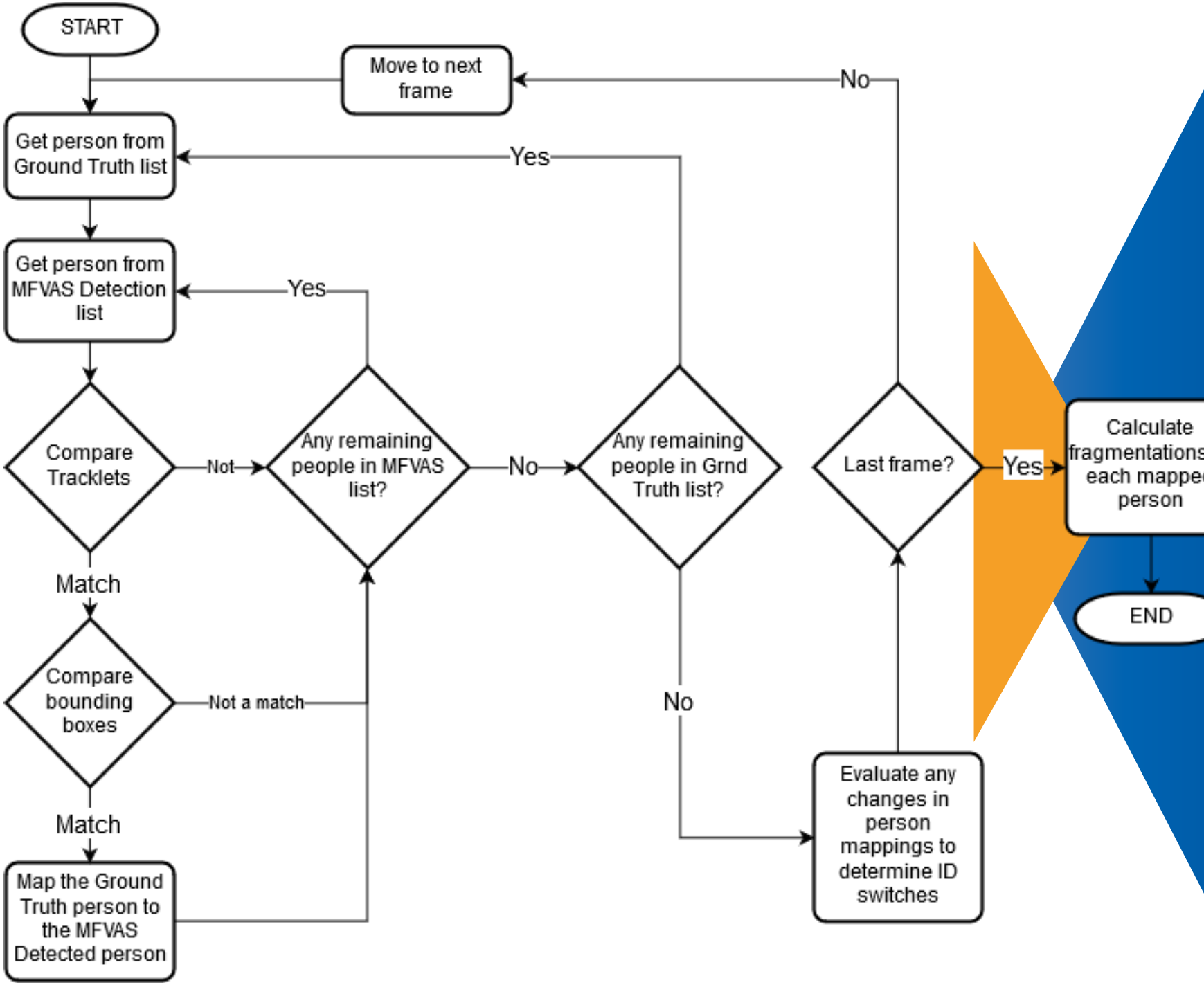
Typical Case

Stationary Case

Program Design

Comparing person objects

Mid-level overview of the process to determine if a ground truth identity matches a MFVAS detected identity



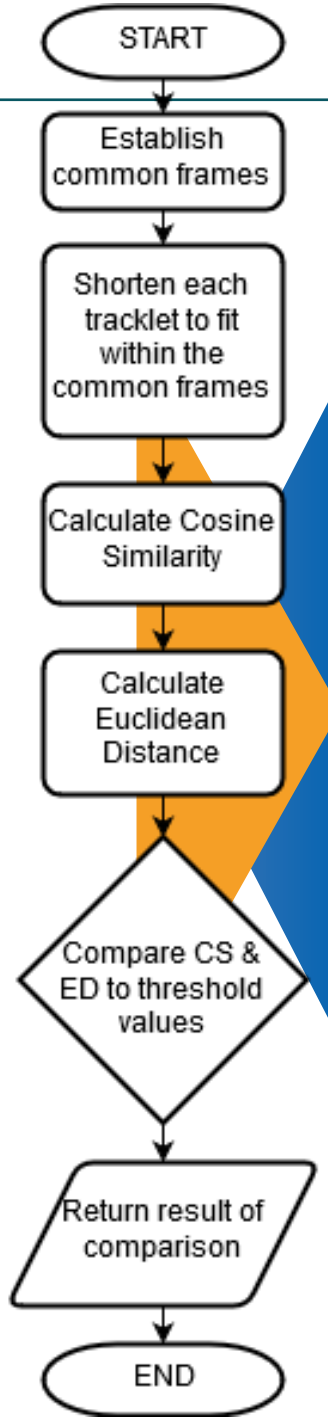
Program Design

Tracklet Comparison

In order to determine the similarity of vectors in a 2D image space, both the proximity (EDAvg) and alignment (Cosθ) need to be considered

$$Cos\theta = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

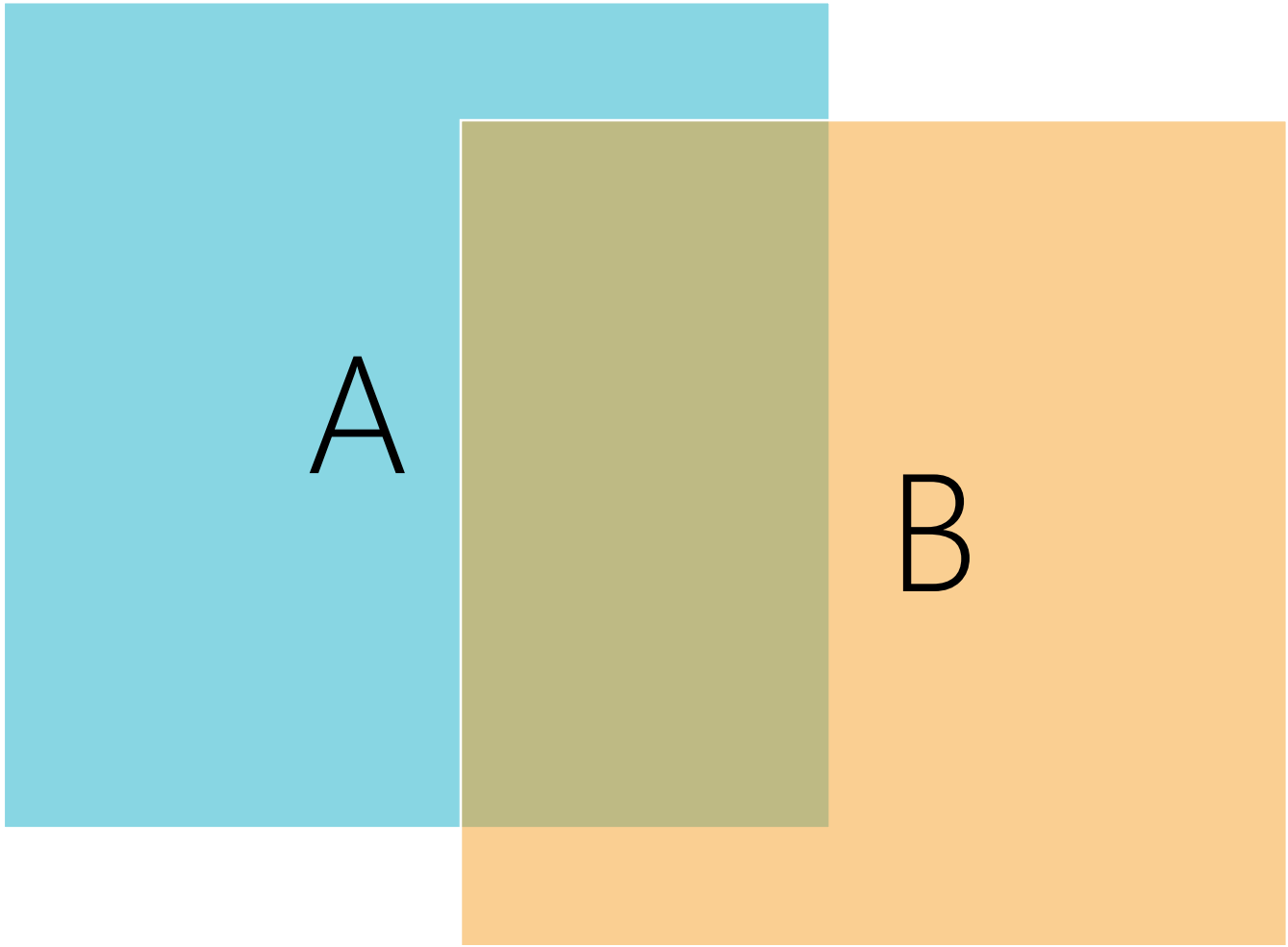
$$ED_{Avg} = \frac{\sum_{i=1}^{n} \sqrt{\left(x_{B_i} - x_{A_i}\right)^2 + \left(y_{B_i} - y_{A_i}\right)^2}}{n}$$



Program Design

Intersection Over Union

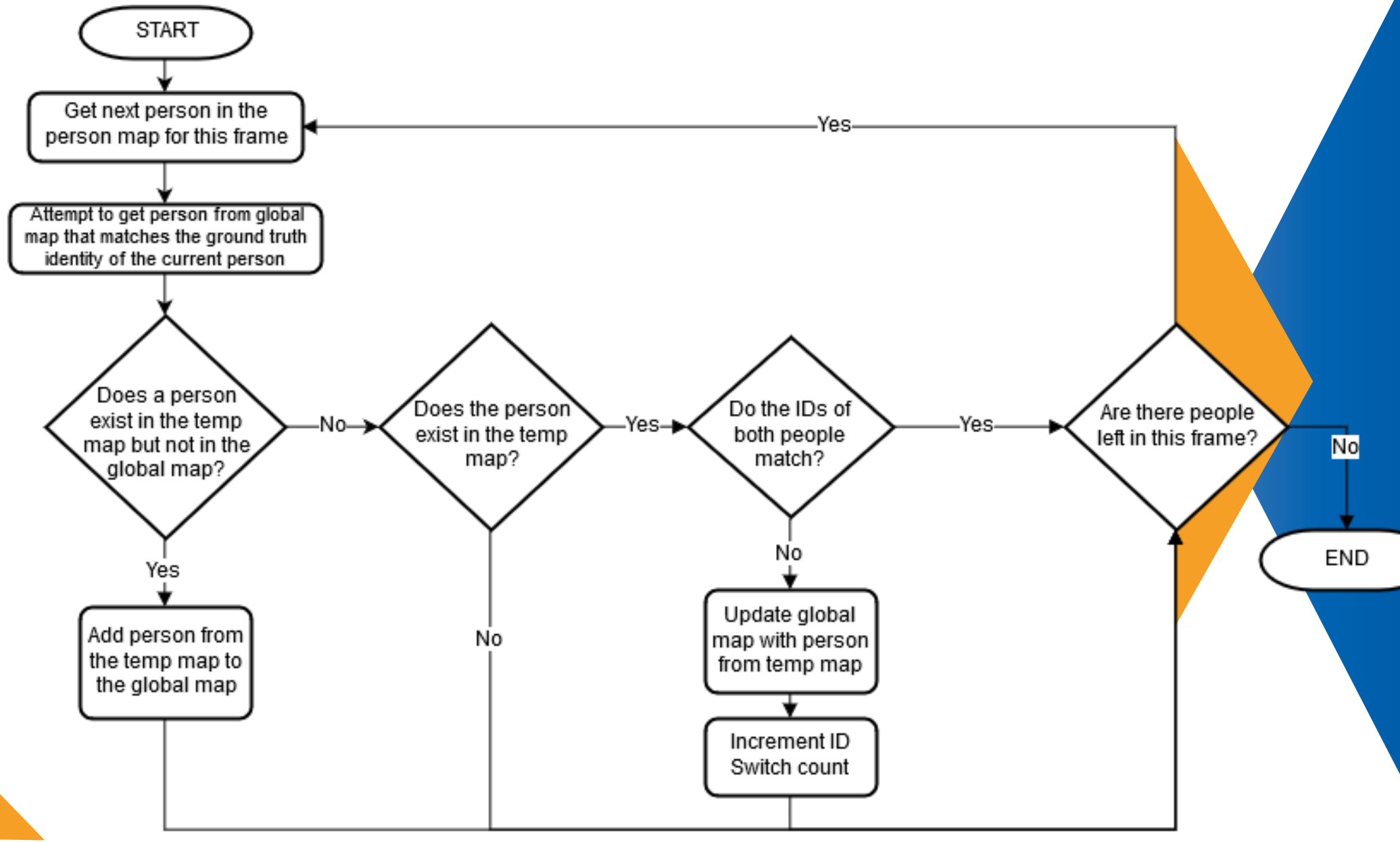
In the context of bounding boxes, IoU is the ratio of shared area to non-shared area



Program Design

Monitoring ID Switches

Post-processing block that runs each frame to track changes in mappings between frames

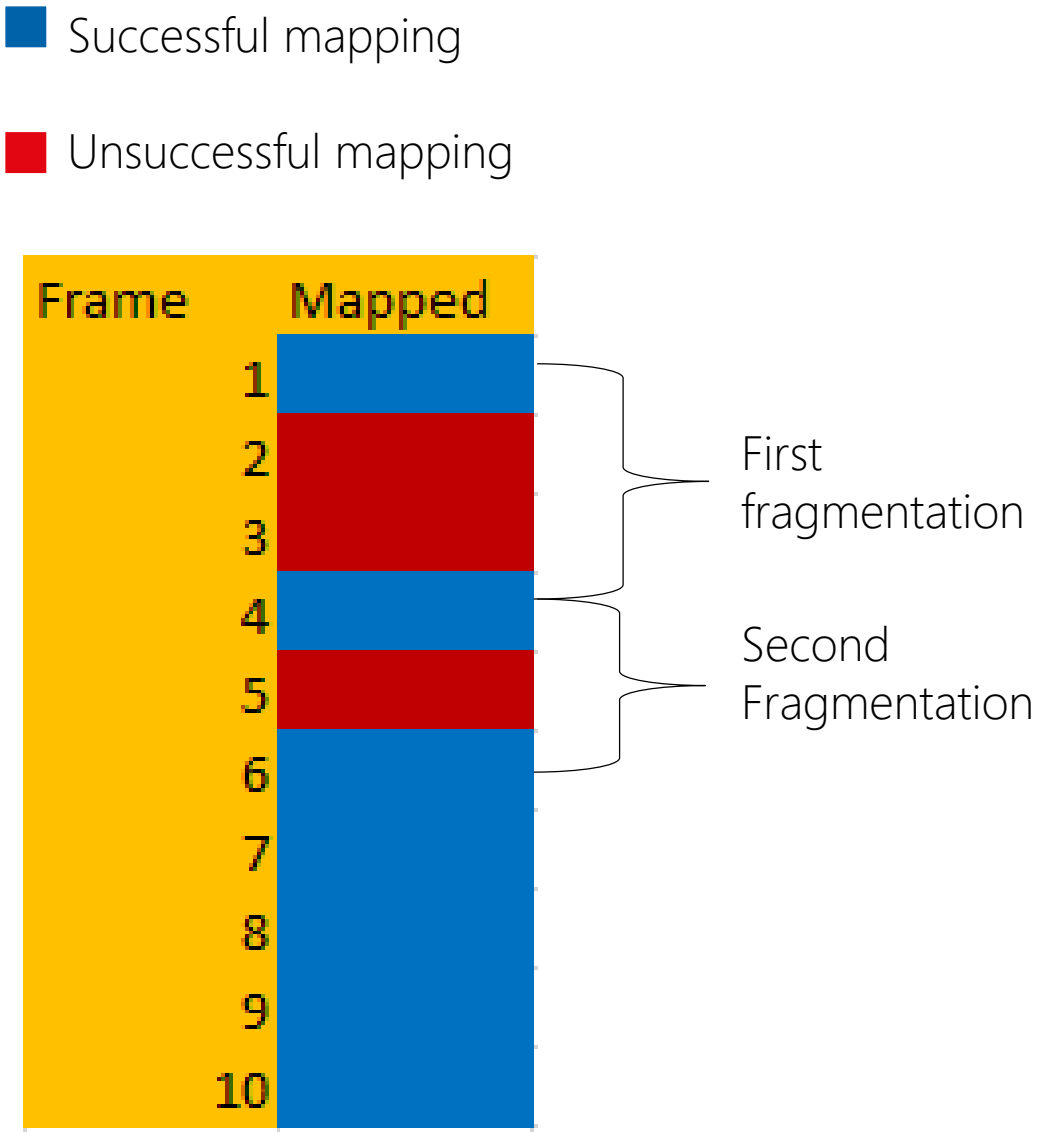


Program Design

Calculating Fragmentations

A fragmentation is represented by successful mappings separated by a gap made up of unsuccessful mappings

To determine if a target has been mostly tracked/lost involves simply counting successful vs unsuccessful mappings



Thank You

The Metrics

The Model Fusion Video Analytics Systems has been evaluated according to the following set of metrics:

- **ID F1 Score [IDF1]:** Harmonic mean of IDP and IDR
- **ID Precision [IDP]**
- **ID Recall [IDR]**
- **Multi-Object Tracking Accuracy [MOTA]**
- **Multi-Object Tracking Precision [MOTP]**
- **False Alarms per Frame [FAF]:** False positives per frame
- **Mostly Tracked/Lost Targets [MT/ML]**
- **False Positives [FP]**
- **False Negatives [FN]**
- **ID Switches [IDSW]**
- **Fragmentations [FRAG]:** Total number of fragmented trajectories

Q&A