RoadTrack: Tracking Road Agents in Dense and Heterogeneous Environments

Proof of Theorem 1

March 3, 2019

1 Reduced Probability of Track Loss

We define $\mathcal{T}_t = \{\Psi_{1:t}\}$ to be the set of positively identified states for p_i until time t. We denote the time since the last update to a track ID as μ. We denote the ID of p_i as α and we represent the correct assignment of an ID to p_i as $\Gamma(\alpha)$. The threshold for the Cosine metric is $\lambda \underset{\text{i.i.d.}}{\sim} \mathbb{U}[0,1]$. The threshold for the track age, *i.e.*, the number of frames before which track is destroyed, is ξ. We denote the probability of an event that uses Mask R-CNN as the primary object detection algorithm with $\mathbb{P}^M(\cdot)$ and the probability of an event that uses a standard Faster R-CNN as the primary object detection algorithm (*i.e.*, outputs bounding boxes without boundary subtraction) with $\mathbb{P}^F(\cdot)$. Finally, $\mathcal{T}_t \leftarrow \{\phi\}$ represents the loss of $\mathcal{T}_t$ by occlusion.

Theorem 1.1. *With high probability $1-\frac{B}{A}$ binary tensors extracted from background subtracted (BGS) representations decrease the number of road agent tracks lost, thereby reducing the number of false negatives, in comparison to regular bounding boxes.*

Proof. In our approach, we use Mask R-CNN for pedestrian detection, which outputs bounding boxes and their corresponding masks. We use the mask and bounding box pair to generate a BGS-Representations (Section 4A). The correct assignment of an ID depends on successful feature matching between the predicted measurement feature and the optimal detection feature. In other words,

$$l(f_{p_i}, f_{h_j^*}) > \lambda \Leftrightarrow (\alpha = \phi) \tag{1}$$

Using Lemma 1 and the fact that $\lambda \underset{i.i.d.}{\sim} \mathbb{U}[0,1]$,

$$\mathbb{P}(l(f_{p_i}, f^M_{h^*_{j,p_i}}) > \lambda) < \mathbb{P}(l(f_{p_i}, f^F_{h^*_{j,p_i}}) > \lambda)$$

Using Eq. 1, it directly follows that,

$$\mathbb{P}^M(\alpha = \phi) < \mathbb{P}^F(\alpha = \phi) \tag{2}$$

In our approach, we set

$$(\mu > \xi) \wedge (\alpha = \phi) \Leftrightarrow \mathcal{T}_t \leftarrow \{\phi\}$$

Using Eq. 2, it follows that,

$$\mathbb{P}^M(\mathcal{T}_t \leftarrow \{\phi\}) < \mathbb{P}^F(\mathcal{T}_t \leftarrow \{\phi\}) \tag{3}$$

We define the total number of false negatives (FN) as

$$FN = \sum_{t=1}^{T} \sum_{p_g \in \mathcal{G}} \delta_{\mathcal{T}_t} \tag{4}$$

where $p_g \in \mathcal{G}$ denotes a ground truth pedestrian in the set of all ground truth pedestrians at current time t and $\delta_z = 1$ for $z = 0$ and 0 elsewhere. This is a variation of the Kronecker delta function. Using Eq. 3 and Eq. 4, we can say that fewer lost tracks ($\mathcal{T}_t \leftarrow \{\phi\}$) indicate a smaller number of false negatives. ■

□

The upper bound, $\mathbb{P}^F(\mathcal{T}_t)$, in Eq. 3 depends on the amount of padding done to f_{p_i} and f_{h_j}. A general observed trend is that a higher amount of padding results in a larger upper bound in Eq. 3.

Lemma 1. *For every pair of binary tensors* $(f^M_{h_j}, f^F_{h_j})$ *generated from a BGS-Representation and a bounding box respectively, if* $\|f^M_{h_j}\|_0 > \|f^F_{h_j}\|_0$*, then* $l(f_{p_i}, f^M_{h_j}) < l(f_{h_j}, f^F_{h_j})$ *with probability* $1 - \frac{B}{A}$*, where* A *and* B *are positive integers and* $A > B$*.*

Proof. Using the definition of the Cosine metric, the lemma reduces to proving the following,

$$f^T_{p_i}(f^M_{h_j} - f^F_{h_j}) > 0 \tag{5}$$

We pad both $f^M_{h_j}$ and $f^F_{h_j}$ such that $\|f^M_{h_j}\|_0 > ||f^F_{h_j}||_0$.

We reduce $f^T_{p_i}$, $f^M_{h_j}$, and $f^F_{h_j}$ to binary vectors, *i.e.*, vectors composed of 0s and 1s. Let $\Delta f = f^M_{h_j} - f^F_{h_j}$. We denote the number of 1s and -1s in Δf as A and B, respectively. Now, let x and y denote the L_0 norm of $f^M_{h_j}$ and $f^F_{h_j}$, respectively. From our padding procedure, we have $x > y$. Then, if $x = A$, and $y = B$, we trivially have $A > B$. But if $y > B$, then $A = x - (y - B) \implies A - B = x - y$. From $x > y$, it again follows that $A > B$. Thus, $x > y \implies A > B$.

Next, we define a $(1, 1)$ coordinate in an ordered pair of vectors as the coordinate where both vectors contain 1s. Similarly, a $(1, -1)$ coordinate in an ordered pair of vectors is the coordinate where the first vector contains 1 and the second vector contains -1. Then, let p_a and p_b respectively denote the number of $(1, 1)$ coordinates and $(1, -1)$ coordinates in the pair $(f^T_{p_i}, \Delta f)$. By definition, we have $0 < p_a < A$ and $0 < p_b < B$. Thus, if we assume p_a and p_b to be uniformly distributed, it directly follows that $\mathbb{P}(p_a > p_b) = 1 - \frac{B}{A}$. ■

□