
第 1 章 “MCTrack: A Unified 3D Multi-Object Tracking Framework for Autonomous Driving”^[1]阅读

1.1 主要贡献

1. 提出统一的数据结构。
2. 提出新的数据关联方法，包括一个新的相似度矩阵和新的匹配方法。
3. 各种状态进行解耦，提高计算效率。
4. 提出新的评价指标，讲速度等状态纳入考虑，更加贴近实际情况。

1.2 具体算法

1.2.1 新的相似度函数

当两个目标重合时，过去的相似度函数难以描述相似度，故引入新的相似度函数。

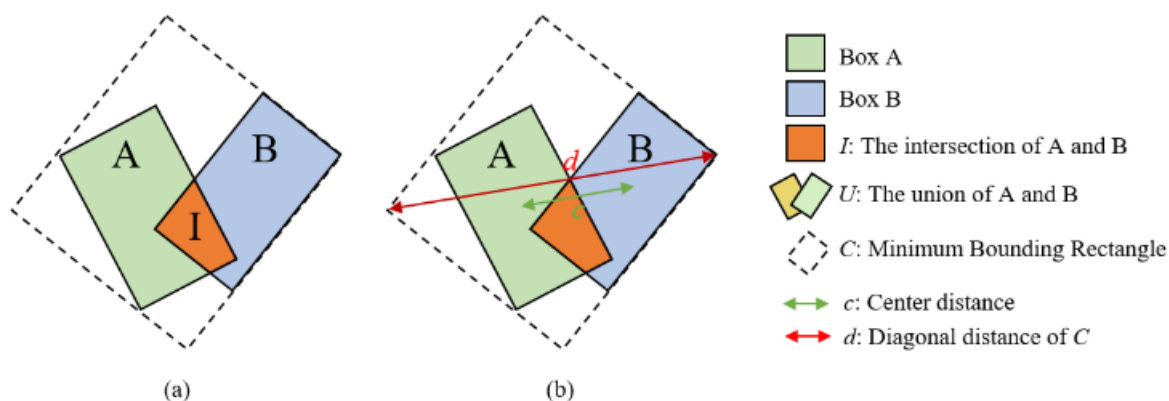


图 1-1 新的相似度函数示意图

输入:

检测结果 $B^d = (x^d, y^d, z^d, l^d, w^d, h^d, \theta^d)$

已有轨迹 $B^t = (x^t, y^t, z^t, l^t, w^t, h^t, \theta^t)$

相似度函数

$$\text{Ro_GDIoU} = \text{Ro_IoU} - w_1 \cdot \frac{C-U}{C} - w_2 \cdot \frac{c^2}{d^2}$$

$$\text{Ro_IoU} = I/U$$

表 1-1 Pseudo-code of Ro_GDIoU Algorithm

Algorithm 1: Ro_GDIoU 伪代码

Input:

检测框, $B_d = (x_d, y_d, z_d, l_d, w_d, h_d, o_d)$

轨迹框, $B_t = (x_t, y_t, t, l_t, w_t, h_t, \theta_t)$

Output:

Ro_GDIoU

1. 转换到 BEV 空间 $B_{\text{bev}}^d, B_{\text{bev}}^t = F_{\text{global} \rightarrow \text{bev}}(B_d, B_t)$
 2. 计算交集 $I = F_{\text{inter}}(B_{\text{bev}}^d, B_{\text{bev}}^t)$
 3. 计算并集 $U = F_{\text{union}}(B_{\text{bev}}^d, B_{\text{bev}}^t)$
 4. 计算最小外框矩形 $C = F_{\text{rect}}(B_{\text{bev}}^d, B_{\text{bev}}^t)$
 5. 中心点距离 $c = F_{\text{dist}}(B_{\text{bev}}^d, B_{\text{bev}}^t)$
 6. 外框对角线 $d = F_{\text{dist2}}(B_{\text{bev}}^d, B_{\text{bev}}^t)$
 7. $\text{Ro-IoU} = I/U$
 8. $\text{Ro_GDIoU} = \text{Ro-IoU} - w_1 \cdot \frac{c-u}{w_2 \cdot d^2}$
(w_1 和 w_2 代表权重)
-

1.2.2 二阶段的匹配方法

本文采用了一种从多角度进行匹配的方法, 先从 BEV 空间进行匹配, 再从 RV 空间进行匹配。

坐标变换:

$$C = R \cdot P + T \quad (1-1)$$

$$P = \begin{bmatrix} \frac{l}{2} & \frac{l}{2} & \frac{l}{2} & \frac{l}{2} & -\frac{l}{2} & -\frac{l}{2} & -\frac{l}{2} & -\frac{l}{2} \\ \frac{w}{2} & -\frac{w}{2} & -\frac{w}{2} & \frac{w}{2} & \frac{w}{2} & -\frac{w}{2} & -\frac{w}{2} & \frac{w}{2} \\ \frac{h}{2} & \frac{h}{2} & -\frac{h}{2} & -\frac{h}{2} & \frac{h}{2} & \frac{h}{2} & -\frac{h}{2} & -\frac{h}{2} \end{bmatrix} \quad (1-2)$$

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (1-3)$$

表 1-2 二阶段匹配算法的伪代码

Input: Trajectory boxes T at time $T - 1$, detection boxes D at time T

Output: Matching indices M

First matching: BEV Plane

$T_{\text{bev}}, D_{\text{bev}} = F_{3d \rightarrow \text{bev}}(T, D)$

Compute cost $L_{\text{bev}} = \text{Ro-GDIOU}(D_{\text{bev}}, T_{\text{bev}})$

Matching pairs $M_{\text{bev}} = \text{Hungarian}(L_{\text{bev}}, \text{threshold}_{\text{bev}})$

Second matching: RV Plane

For each $d_{\text{bev}} \in D_{\text{bev}}$:

 If $d_{\text{bev}} \notin M_{\text{bev}}[0]$:

$d_{\text{bev}} \rightarrow D_{\text{res}}$

For each $t_{\text{bev}} \in T_{\text{bev}}$:

 If $t_{\text{bev}} \notin M_{\text{bev}}[1]$:

$t_{\text{bev}} \rightarrow T_{\text{res}}$

Matching pairs $M_{\text{rv}} = \text{Greedy}(L_{\text{rv}}, \text{threshold}_{\text{rv}})$

Obtain the final matching pairs $M = M_{\text{bev}} \cup M_{\text{rv}}$

参考文献

- [1] Wang X, Qi S, Zhao J, et al. MCTrack: A Unified 3D Multi-Object Tracking Framework for Autonomous Driving[J]. arXiv preprint arXiv:2409.16149, 2024.