Universe Scientific Publishing

Unmanned multi target tracking based on boundary IoU target association

Ke Xu[1], Linxi Hu[2], Wei Long[2], Linhua Jiang[1,2*]

1.School ofInformation Engineering, Huzhou University, Huzhou 313000, China.
2.Jinhua Institute of Zhejiang University, Jinhua 321000, China.

Abstract: Aiming at the defects of target matching measurement and fixed life cycle management in traditional multi-target tracking, a target association and adaptive life cycle management strategy based on boundary IoU is proposed. Boundary IoU takes into account the advantages of Euclidean distance and IoU, which can improve the accuracy of target matching. The adaptive life cycle management correlates the target trajectory confidence with the life cycle, which significantly reduces the target loss and trajectory false detection during the tracking process. Experiments on KITTI multi-target tracking dataset demonstrate the effectiveness of the proposed method.

Keywords: Unmanned Driving; Lidar; 3D Target Detection; Iou Loss; Multi Target Tracking

The existing multi-target tracking technology can be divided into two categories: end-to-end tracking and detection-based tracking. The former treats object detection and tracking as a unified process, and obtains an object detection box with unique ID tag through the input of a single image or point cloud data. The latter algorithm treats target tracking as a progressive process. Firstly, the location of the detection frame is obtained by using the compound YOLO, PointRCNN and other target detection networks. Secondly, multi-target tracking is realized by using the time-space correlation between multi-frame images.

1. The overall process of 3D target tracking

Multi-target tracking refers to the discovery of objects to be detected from a group of images in a specific image sequence, and then the movement information of the objects is obtained through correlation analysis, so as to provide an accurate identification for each object. On this basis, this project proposes a method based on Track-by-Detection. Based on the detection results of PointRCNN, the 3D object tracking is realized by Kalman filter, Hungarian matching and fixed lifetime management.

As shown in Figure 1, the overall process of 3D target tracking proposed in this paper can be divided into the following parts:

(1) Using 3D target detector to process the original point cloud to get 3D target border in the environment;

(2) 3D Kalman filter predicts the target state of the current frame to get the predicted border;

(3) Hungarian matching is performed between the detected target border and the predicted border estimated by Kalman filter;

(4) Kalman filter updates the target state according to the matching result;

(5) Manage and control the life cycle of unmatched targets;

(6) Get the target border with a unique ID flag.

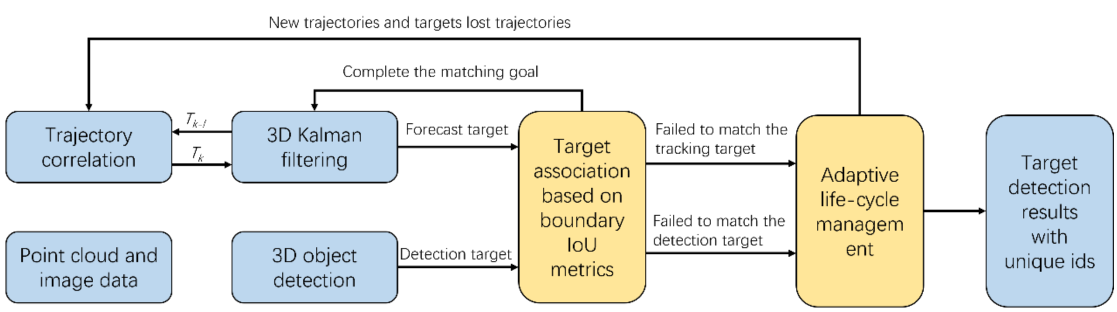

Figure 1 3D target tracking based on Kalman filter

2. Point cloud raw image processing

2.1 Data characteristics of LiDAR

Lidar is a kind of precision optical sensor, its ranging accuracy can reach ±2cm, according to the different built-in laser beam can be divided into 4 lines, 8 lines, 16 lines, 32 lines, 64 lines and 128 lines and other models. The data of point cloud image is a data set composed of points, each point can be expressed as a four-dimensional tensor composed of spatial coordinates and intensity, and the JTH point of point cloud can be expressed as:

$$P_j = [x, y, z]^T$$

Among them, x,y and z are the three-dimensional coordinates of the point in space respectively, and i is the reflection intensity of the laser beam at the point. The reflection intensity is affected by factors such as light and the surface material of the reflector, which is an important information.

2.2 Point cloud feature coding

In order to make up for the information loss of point cloud in the process of projection dimensionality reduction, and to reduce the data redundancy of manual features and improve efficiency, this chapter proposes the main feature input with the height variance feature graph. Variance is a measure of the degree of data dispersion, and the height variance map can effectively highlight obstacles, which has positive significance for subsequent feature extraction and regression. Finally, we select five manual features, namely maximum height, minimum height, height variance, maximum intensity and density, for feature coding of the aerial view, and define these feature channels as follows:

Among them, Zmax and Zmin represent maximum and minimum height channels, Zvar represents the height variance channel, Zi denotes maximum intensity channel, Zd representing density channels, m*n is the size of the rasterized feature map.

Finally, the sparsity and coverage area of point cloud are considered comprehensively. According to the imaging principle of Lidar, the point cloud farther away from the center of the radar is more sparse, and the distant point cloud will be more difficult to extract features.

Therefore, by setting the distance limit, cutting the effective point cloud region Ω has a positive significance for improving efficiency and performance.

3. Hungarian matching algorithm based on boundary IoU metric

This project intends to apply Hungarian matching technology to a series of SORT multi-target tracking, and transform it into a kind of matching problem, that is, matching the detected nk pedestrian information with the existing mk target data.

Due to the influence of point cloud data characteristics and model generalization ability, the existing Lidar 3D target detection algorithm has some inaccurate Angle estimation when measuring object Angle, IoU measurement method cannot distinguish 0 Angle and −pi Angle, and point cloud data features are difficult to analyze the orientation of the head and tail, etc., which will affect the matching result.

To solve the above problems, this project selects the edge processing method of aerial images in TBIoU, and takes the standardized distance of the left and right upper and right edges as the penalty condition to reduce the influence of edge rotation on image matching results. This method combines the advantages of Euclidean distance and IoU measurement to achieve high precision.

4. Adaptive life-cycle management

In practical applications, due to the influence of generalization ability, illumination, occlusion and other factors, the 3D target detector often appears false detection or missing detection phenomenon. If the above simple lifetime management method is still used, a large number of false tracks will be generated. This is mainly because, in the lifetime, the number of fixed frames is taken as the unit, and the lifetime management strategy of the fixed number of frames cannot make full use of the confidence of the target, and for the tracking track with a large number of fixed frames, a certain number of frames must be retained before it can be eliminated. For tracking with high confidence, when there are omissions or occlusions in the detection process, enough frame length should be retained as far as possible, while tracking with low confidence is expected to be removed as soon as possible to prevent false detection.

On this basis, this paper proposes an adaptive survival management strategy, which dynamically adjusts the maximum life cycle $F_{\max}$, and its formula is as follows:

$$F_{A\max} = sig\,\mathrm{mod}(a \cdot s + b) \times F_{\max}$$

Where, s is the confidence score of the current tracking trajectory, a and b are the scale coefficients of the adjusted sigmod function distribution, Fmax is the maximum life cycle, FAmax is the life cycle calculated according to the tracking trajectory score, and better tracking effect can be achieved by appropriate α and b values.

5. Experimental details and results

5.1 Data sets and evaluation indicators

5.1.1 KITTI multi-target tracking dataset

On this basis, based on KITTI database, 21 sets of training sequences and 29 sets of test sequences are constructed and evaluated. Lidar point clouds, RGB images and calibration files are available for each series. The number of data frames used for training was 8008, and 11095 for testing. To test the collection, KITTI does not give the user any tags, but saves them on the server for MOT calculations. In the training use case, 30601 targets and 636 tracks are included, as well as the vehicle class, Pedestrain and Cyclist class.

5.1.2 Evaluation index

For a broad monitoring approach, the ideal assessment indicator should be able to meet three requirements at once: first, any new targets can be identified in a timely manner. Second, the target location found should be as consistent as possible with the actual target location. Third, the consistency of monitoring targets must be maintained to avoid repeated exchange of target labels.

According to the above three requirements, the evaluation indicators of traditional multi-target monitoring Settings are as follows: Multi-target monitoring accuracy (MOTA) is used to determine the number of targets and cumulative errors during the monitoring process, and multi-target tracking accuracy (MOTP) is used to measure the accuracy of the target's position and the tracking trajectory of most targets (MostlyTracked, MT).

5.2 Experiment and result analysis

The limit value IoU is used as the cost index to construct the cost matrix, and the observation result is consistent with the prediction target of Kalman filter. The fault target is placed in the lifecycle management module and adaptively adjusts the maximum lifetime according to the confidence point of the target trajectory. Finally, unique identifiers are used to obtain the results of target observations and trajectories. The minimum detection frame rate Fmin=3 in the life cycle management module means that it is considered as a new target trajectory before at least three consecutive frames are detected. The maximum survival frame rate Fmax and scale parameters in the adaptive life cycle setting are obtained by tuning.

We performed experiments based on the above setup and compared performance with an improved Kalman filter-based target tracking method and two typical deep learning target tracking algorithms, ComplexerYOLO and FANTrack. The experimental results of the above methods are from the literature, and the comparison results of CAR-level performance are shown in Table 1. It should be noted that this method only needs to run the processor in real time and does not require prior training, while the other two methods are based on deep learning models, which require high computing power for the computing platform.

Table 1 Comparison of Car class multi-target tracking performance on KITTI validation set

method	$F_{\max}$	MOTA(%)↑	MOTP(%)↑	MT(%)↑	ML(%)↓	IDS↓	FRAG↓
Complexer-YOLO	-	72.23	73.42	61.09	4.96	985	1583
FANTrack	-	72.18	75.62	72.75	8.86	13	70
	7	77.87	77.90	80.00	4.86	0	10
AB3DMOT	5	81.26	77.94	80.00	4.86	0	10
	3	84.75	78.21	77.84	4.86	0	13

	7	81.86	77.82	80.00	4.86	2	12
Ours	5	83.69	78.26	77.30	5.41	2	14
	3	86.60	78.27	75.68	4.32	2	16

In addition, the results of multi-target monitoring are compared. Since ComplexerYOLO and FANTrack did not produce experimental Pedstrain results, only the AB3DMOT comparison algorithms were compared.

Table 2 Comparison of multi-target tracking performance in Pedestrian category in KITTI verification set.

method	Fmax	MOTA(%)↑	MOTP(%)↑	MT(%)↑	ML(%)↓	IDS↓	FRAG↓
	7	68.17	66.91	58.45	19.01	1	40
AB3DMOT	5	69.17	66.97	59.15	16.90	2	46
	3	69.52	67.02	53.28	22.54	2	74
	7	70.58	66.91	61.27	18.31	3	48
Ours	5	71.32	67.17	58.45	16.20	3	61
	3	70.34	67.55	55.63	16.90	3	76

The experimental results in Table 1 and 2 show that the adaptive life cycle management strategy proposed in this chapter and the Hungarian matching algorithm based on the limit IoU meter can effectively improve most of the performance indicators of multi-target monitoring, especially its key performance indicators MOTA and MOTP. Compared to the improved AB3DMOT system, IDS and FRAG have slightly improved performance metrics for downtime, while other performance aspects are improving. The table shows that different kinematic models for vehicles and pedestrians allow the pedestrian class to achieve better results with relatively high life cycle parameters. For the vehicle class, for MOTA's key multi-function indicators, the approach described in this chapter achieves a relative improvement of more than 2% compared to several typical life cycle parameters. For pedestrians, the relative improvement is about 1 percent or more. The above tests prove the effectiveness of the strategy proposed in this chapter. The Hungarian matching algorithm based on boundary IoU measurement can effectively improve the accuracy of target compliance, and the adaptive life cycle management strategy can improve the sustainability of multi-target monitoring.

References

[1] Liu Yuchen, Tang Xing, Su Yan. Real-time target 3D detection method of unmanned platform based on camera and LiDAR :CN201911244310.7[P].CN110879401A[2023-11-02].

[2] Cheng Xiao-qiang. Unmanned 3D target detection and multi-target tracking based on LiDAR [D]. Tianjin University,2020.

[3] Cao Weihao. Ground Target Detection method and Application based on 3D LiDAR [J].[2023-11-02].

[4] Liu Mingyu, Yan Jun, GE Wancheng. Review of 3D target detection technology based on LiDAR point cloud [J]. Television Technology,2022(001):046.

[5] Zhang Pu, LIU JQ, XIAO Jchao, et al. Target location and tracking based on fusion of camera and LiDAR [J]. Advances in Laser and Optoelectronics, 2019,61(08).

[6] Zhang Xiangli. Research on surface vessel detection and recognition based on LiDAR and visual information fusion [D]. Shanghai Jiao Tong University,2017.

[7] Xiong Zhenkai, Cheng Xiaoqiang, Wu Youdong, et al. Based on unmanned 3 d laser radar multi-target tracking [J]. Journal of automation, 2022, : 1-11. DOI: 10.16383 / j.a as c210783.

[8] Jiang Jingwen, Zhu Bing, Ma Tianfei, et al. A 3D target detection and tracking method based on camera and LiDAR :CN202011337882.2[P].CN112487919A[2023-11-02].

[9] Wen Long. Research on obstacle detection and tracking technology of unmanned vehicle based on vehicle Lidar [D]. Jilin University,2020.

[10] Wang Qing-Lin, Li Hui, Xie Li-Zhi, et al. Research on Improvement of vehicle target detection Algorithm based on LiDAR point Cloud [J]. Electronic Measurement Technology,2023,46(1):120-126.