WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 2
CII 2023

Comparative Study on BEV Vision and LiDAR Point Cloud Data Fusion Methods

Junyu Zhou

Southwest Jiaotong University, Chengdu 610097, China

Abstract. With the gradual maturity of autonomous driving technology, efficient fusion and processing of multimodal sensor data has become an important research direction. This study mainly explores the strategy of integrating BEV (Bird's Eye View) vision with LiDAR point cloud data. We evaluated the performance and applicability of each of the three main data fusion methods through in-depth comparison: early fusion, mid-term fusion, and late fusion. First of all, we summarize the working principle and data characteristics of BEV vision and LiDAR, and emphasize their key roles in auto drive system. Then, the theoretical basis and implementation methods of the three fusion strategies were described in detail. The experimental results show that different fusion strategies exhibit their own advantages in different application scenarios and requirements. For example, early fusion performs well in high-precision tasks, but has a high demand for computing resources. And mid-term fusion is more suitable in scenarios with high real-time requirements. Overall, this study provides in-depth insights and practical suggestions on the fusion of BEV vision and LiDAR data in the field of autonomous driving, laying a solid foundation for future research and applications.

Keywords: BEV Vision; LiDAR Point Cloud; Data Fusion; Comparative Research.

1. Introduction

With the rapid development of autonomous driving technology, how to efficiently process and analyze a large amount of data from different sensors has become a key issue. Among many sensors, BEV vision camera and LiDAR are the most important sensing components in autonomous driving system, which provide abundant environmental information for vehicles[1]. However, due to their different working principles and data characteristics, it is a challenge to directly integrate their data. In order to solve this problem, this study focuses on the fusion method of BEV vision and LiDAR point cloud data. The importance of sensors in autonomous driving: autonomous vehicles need to have accurate perception of their surroundings. Therefore, vehicles are equipped with a series of sensors, such as cameras, LiDAR, radar and ultrasonic sensors, to detect and identify objects on the road, such as other vehicles, pedestrians and traffic signs [2]. BEV vision provides a bird's-eye image around the vehicle, which is helpful to determine the position and shape of the object, but it may be difficult to calculate the depth information. LiDAR measures the distance of an object by emitting laser pulses and measuring its return time, thus generating high-precision three-dimensional point cloud data, but it does not contain color information and may be disturbed under some environmental conditions [3]. Because of the different data characteristics and working principles of BEV vision and LiDAR, direct fusion of their data may lead to redundancy or omission of information. In addition, the computational burden of processing and parsing these two kinds of data is also quite heavy. Therefore, finding an effective data fusion method can not only make full use of the advantages of the two sensors, but also avoid their limitations, which is very important for improving the performance and safety of the automatic driving system [4-5]. This study deeply understands the characteristics and challenges of BEV vision and LiDAR data. Explore and compare different data fusion methods to determine their advantages and limitations. The performance of different fusion methods is evaluated by experimental verification to provide reference for practical application [6]. In the field of autonomous driving, many researchers have tried various data fusion methods, which rely on predefined rules to integrate data, but may lack flexibility [7]. For example, deep learning can automatically learn the characteristics of data, but it needs a lot of labeled data. Such as early integration, mid-term integration and late integration, each strategy has its own applicable scenarios and challenges. The following parts of this paper will introduce the theoretical background, three main fusion strategies, experimental verification and analysis of the results, and finally summarize the research results, which is a detailed description of the introduction, providing readers with the research background, importance, purpose and structure, and laying the foundation for the following papers [8].

2. Fundamentals of Data Fusion

2.1. BEV Visual Overview

2.1.1. Definition and Application Fields

BEV, also known as bird's-eye view, is a view of an object from the top or vertical direction. In the field of robotics and autonomous driving, BEV vision is often used for environmental perception, obstacle detection, and path planning. Due to its ability to provide an algorithm with a view without perspective distortion, BEV is particularly popular in specific applications.

2.1.2. Generation of BEV Data

BEV images are usually obtained through perspective transformation of the original image obtained by the camera. This requires a known relationship between the camera and the ground, and through this transformation, the size and position of the object in the BEV image can directly correspond to the actual physical size and position [9].

2.1.3. Characteristics of BEV Data

The size of objects in BEV images is directly proportional to their size in the actual environment, and is not affected by the camera's perspective or depth. The relative positions between objects are clearly visible in BEV, facilitating spatial analysis. Due to being viewed from the top perspective, certain areas often obscured by other objects may become visible in BEV.

2.2. Overview of LiDAR Point Cloud Data

2.2.1. Working Principle

LiDAR (Light Detection And Ranging) is a remote sensing technology based on optics, which can determine the distance of the target object by emitting a laser beam and measuring the time difference of the reflected laser signal. Each measured data point contains the information of spatial position, thus forming a point cloud data in a three-dimensional space.

2.2.2. Main Application Fields

LiDAR technology is widely used in topographic mapping, forestry, urban planning, mining and autonomous driving. In autonomous driving, LiDAR provides high-precision and high-resolution environmental awareness, which plays a key role in detecting obstacles, buildings and other road features.

2.2.3. Characteristics of Point Cloud Data

The characteristics of point cloud data mainly have the following four characteristics:

(1) High precision

LiDAR point cloud data can provide millimeter-level measurement accuracy.

(2) Variable density

Different LiDAR systems have different scanning frequencies and densities, ranging from low-density point clouds to high-density point clouds with thousands of points per square meter.

(3) Robustness

Compared with visual sensors, LiDAR is more robust to changes in lighting conditions.

(4) No texture information

Unlike RGB images, LiDAR data does not contain color or texture information, but it can be obtained by fusing with cameras.

3. Fusion Method

Data fusion plays a vital role in many applications, especially in the field of autonomous driving. By synthesizing information from multiple sensors, a more complete and accurate environmental representation can be obtained. This chapter will discuss three main data fusion methods in detail: early fusion, middle fusion and late fusion.

3.1. Early Fusion

Early fusion, also known as sensor-level fusion or raw data fusion, involves merging raw data of multiple sensors directly before feature extraction. Early fusion is achieved by directly combining the raw data from different sensors. It shows a high degree of accuracy when processing a large number of raw data, especially in applications that require high resolution and accuracy, such as fine object detection and depth estimation. However, it has a high demand for computing resources and may not be suitable for real-time applications. For example, the BEV image and LiDAR point cloud data are superimposed to produce a deep image, which requires that the time synchronization and geometric calibration between sensors are very accurate, the data structure is simple and easy to process [10]. All the information of the sensor can be directly used. And potential errors caused by subsequent processing are reduced. The disadvantage is that it requires a high degree of time synchronization and accurate geometric calibration. The amount of data is huge, and the demand for calculation and storage is high. Different sensors may have different noise and error characteristics, and direct fusion may introduce errors.

3.2. Mid term Integration

Mid term fusion is carried out at the feature level, which involves extracting features from various sensors before fusion. In this method, features are first extracted independently from the data of each sensor, and then combined together. This may involve mapping features into a common representation space and then combining or comparing them. The advantage is that the data dimension is low, reducing computational and storage requirements. Feature extractors can be designed specifically for each sensor to maximize the utilization of information from each sensor. This method provides a balanced solution that allows for fusion at the feature level. It provides a compromise between computational efficiency and accuracy, suitable for most autonomous driving scenarios. Reduced noise and errors caused by direct data fusion. The disadvantage is that some sensors may lose their original information. We need to design and optimize feature extractors for each sensor.

3.3. Post Fusion

Post-fusion, or decision-level fusion, involves making decisions independently on the data of each sensor, and then combining these decisions. In post-fusion, the data of each sensor is processed independently to produce a decision or output. These decisions are then combined to produce a final output or decision. This usually involves voting, averaging or some other form of combination strategy. The advantage is that the data of each sensor can be processed by a special processing chain. Sensors can be easily added or deleted without redesigning the whole fusion process. It is suitable for those situations where each sensor can produce high-quality decisions. The disadvantage is that the correlation information between multi-sensor data is lost. If a sensor makes an inaccurate decision, it may affect the final output. Which fusion method to choose depends on the specific application and available resources. Early fusion provides abundant data, but the computing demand is high. As the most efficient method, late integration ensures rapid response, especially in resource-constrained environments. However, its accuracy may be affected by the decision of a single sensor, and it is suitable for applications that require high real-time performance but relatively low accuracy. Post-fusion is carried out at the feature level, which provides a compromise solution, while post-fusion is simpler and more flexible, but some information may be lost.

4. Method Comparison

In the context of data fusion, comparative analysis is the key to evaluating the impact of different fusion methods on system performance. This chapter will compare early fusion, mid-term fusion, and late fusion in terms of data processing speed and accuracy, as shown in Figure 1-2.

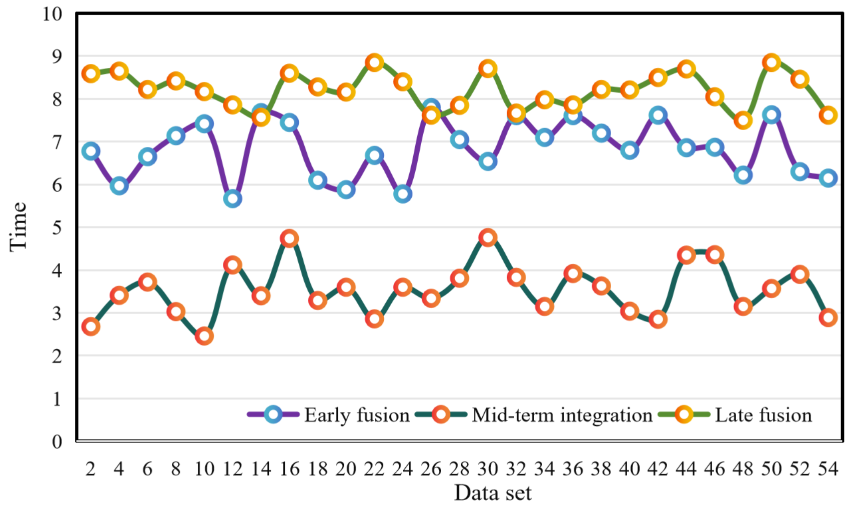

Figure 1. Speed Comparison

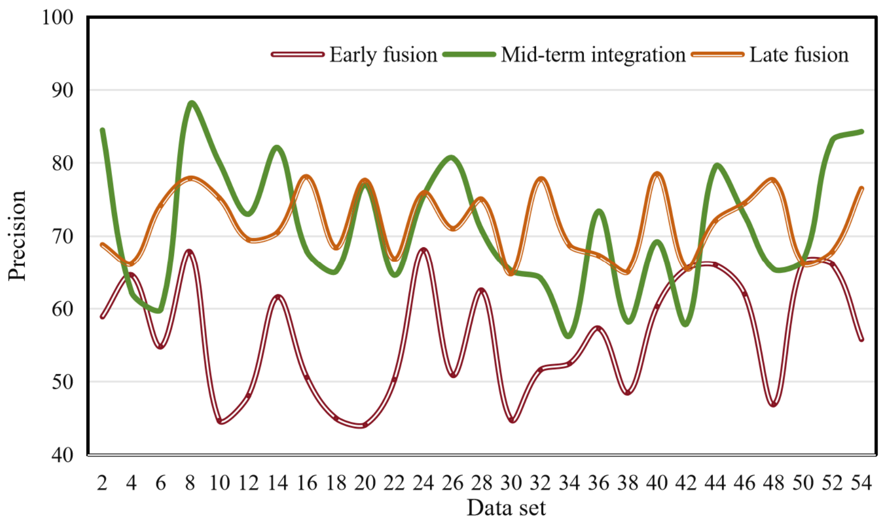

Figure 2. Precision Comparison

Through the analysis of the above experimental results, it can be found that through the comparison of three different data fusion methods, mid-term fusion performs the best among the three methods, and is superior to the other two methods in terms of speed and accuracy.

This chapter analyzes the best fusion methods based on different application scenarios. Different data fusion methods have their specific advantages and limitations. It can be seen that when selecting fusion methods, specific application requirements, available computing resources, and required accuracy and robustness should be considered. With the advancement of sensor technology and the enhancement of computing power, new data fusion methods and strategies may emerge. In addition, deep learning and other advanced machine learning technologies have brought new possibilities for data fusion. Future research can explore the application of these methods in the fusion of BEV vision and LiDAR data.

5. Conclusion

In this study, three main methods of data fusion between BEV vision and LiDAR point cloud are discussed: early fusion, middle fusion and late fusion. Each method has its unique advantages and limitations and is suitable for different application scenarios and requirements. Early fusion shows high accuracy when dealing with a large number of raw data, especially in applications that require high resolution and accuracy, such as fine object detection and depth estimation. However, it has a high demand for computing resources and may not be suitable for real-time applications. Mid-term integration is the most efficient method, and late integration ensures rapid response, especially in the resource-limited environment. However, its accuracy may be affected by the decision of a single sensor, and it is suitable for applications that require high real-time performance but relatively low accuracy. Post-fusion provides a balanced solution, allowing fusion at the feature level. It provides a compromise between calculation efficiency and accuracy, and is suitable for most autopilot scenarios. The experimental verification part of this study clarifies the actual performance of various fusion methods in different scenarios and conditions. The fusion of BEV vision and LiDAR point cloud data is a key problem in autonomous driving technology. Through this study, we not only deepen our understanding of various fusion methods, but also provide valuable reference for practical application. It is hoped that with the development of technology, data fusion can further promote the safety and efficiency of autonomous driving system.

References

[1] Wang H, Lou X, Cai Y, et al. Real-Time Vehicle Detection Algorithm Based on Vision and Lidar Point Cloud Fusion[J]. Journal of Sensors, 2019, 20(5):1-9.

[2] Li R, Li J, Wang J, et al. Dual-view 3D object recognition and detection via Lidar point cloud and camera image[J]. Robotics and Autonomous Systems, 2022, 25(150):150-158.

[3] Li J, Li R, Wang J, et al. Obstacle information detection method based on multiframe three-dimensional lidar point cloud fusion[J]. Optical Engineering, 2019, 58(11):1-7.

[4] Kim K, Kim C, Jang C, et al. Deep learning-based dynamic object classification using LiDAR point cloud augmented by layer-based accumulation for intelligent vehicles[J]. Expert Systems with Applications, 2020, 15(4):113861-113869.

[5] Huaping L. Fusion Method of Airborne LiDAR Point Cloud and Imagery Data[J]. Geospatial Information, 2018, 5 (4):15-20.

[6] Mukherjee A, Dipta Das S, Ghosh J, et al. Semantic Segmentation of Surface from Lidar Point Cloud[J]. arXiv e-prints, 2020, 15(4):18-23.

[7] Zhanhua X U, Zejun X, Fenghai L. A fusion method of LiDAR point cloud and ladybug panoramic image[J]. Bulletin of Surveying and Mapping, 2019, 26(5):9-14.

[8] Peng L, Liu F, Yu Z, et al. Supplementary Material: Lidar Point Cloud Guided Monocular 3D Object Detection[J]. Forest ecosystems: English version, 2022, 22(5):16-20.

[9] Fekry R, Yao W, Cao L, et al. Ground-based/UAV-LiDAR data fusion for quantitative structure modeling and tree parameter retrieval in subtropical planted forest[J]. Forest ecosystems: English version, 2022, 9(5):674-691.

[10] Gandi N M L. Classification Airbrone LiDAR Point cloud Data[J]. International Journal for Modern Trends in Science and Technology, 2021, 7(1):36-39.