World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:17, No:07, 2023

Event Data Representation Based on Time Stamp for Pedestrian Detection

Authors : Yuta Nakano, Kozo Kajiwara, Atsushi Hori, Takeshi Fujita

Abstract : In association with the wave of electric vehicles (EV), low energy consumption systems have become more and more important. One of the key technologies to realize low energy consumption is a dynamic vision sensor (DVS), or we can call it an event sensor, neuromorphic vision sensor and so on. This sensor has several features, such as high temporal resolution, which can achieve 1 Mframe/s, and a high dynamic range (120 DB). However, the point that can contribute to low energy consumption the most is its sparsity; to be more specific, this sensor only captures the pixels that have intensity change. In other words, there is no signal in the area that does not have any intensity change. That is to say, this sensor is more energy efficient than conventional sensors such as RGB cameras because we can remove redundant data. On the other side of the advantages, it is difficult to handle the data because the data format is completely different from RGB image; for example, acquired signals are asynchronous and sparse, and each signal is composed of x-y coordinate, polarity (two values: +1 or -1) and time stamp, it does not include intensity such as RGB values. Therefore, as we cannot use existing algorithms straightforwardly, we have to design a new processing algorithm to cope with DVS data. In order to solve difficulties caused by data format differences, most of the prior arts make a frame data and feed it to deep learning such as Convolutional Neural Networks (CNN) for object detection and recognition purposes. However, even though we can feed the data, it is still difficult to achieve good performance due to a lack of intensity information. Although polarity is often used as intensity instead of RGB pixel value, it is apparent that polarity information is not rich enough. Considering this context, we proposed to use the timestamp information as a data representation that is fed to deep learning. Concretely, at first, we also make frame data divided by a certain time period, then give intensity value in response to the timestamp in each frame; for example, a high value is given on a recent signal. We expected that this data representation could capture the features, especially of moving objects, because timestamp represents the movement direction and speed. By using this proposal method, we made our own dataset by DVS fixed on a parked car to develop an application for a surveillance system that can detect persons around the car. We think DVS is one of the ideal sensors for surveillance purposes because this sensor can run for a long time with low energy consumption in a NOT dynamic situation. For comparison purposes, we reproduced state of the art method as a benchmark, which makes frames the same as us and feeds polarity information to CNN. Then, we measured the object detection performances of the benchmark and ours on the same dataset. As a result, our method achieved a maximum of 7 points greater than the benchmark in the F1 score.

Keywords : event camera, dynamic vision sensor, deep learning, data representation, object recognition, low energy consumption
Conference Title : ICIVA 2023 : International Conference on Intelligent Vehicles and Applications
Conference Location : Prague, Czechia
Conference Dates : July 03-04, 2023