Real-time high speed motion prediction using fast aperture-robust event-driven visual flow

Lounes ALLIOUI and Branis GHOUL

***Abstract*— Optical flow is an important feature for understanding dynamic scenes in computer vision, especially in cases where optical flow is low. Dynamic vision sensors are particularly suited to this task because of their asynchronous, sparse and temporally accurate representation of visual dynamics. However, many visual flow algorithms for these sensors have an aperture problem, where the estimated flow direction is influenced by the curvature of the object rather than the actual direction of motion. We propose a new visual flow algorithm based on a multiscale planar fit that is robust to the aperture problem and efficient in terms of computation time. Our algorithm performs well in a variety of scenarios, ranging from stationary cameras recording simple geometric shapes to real-world scenarios such as a camera mounted on a moving car, and can successfully perform event-by-event motion estimation of objects in the scene to allow predictions of up to 500 ms, i.e., the equivalent of 10 to 25 frames with traditional cameras.**

I. INTRODUCTION

An event-based visual stream is a type of video stream that captures data only when a significant change in brightness is detected in the image, rather than capturing continuous images at a fixed frame rate like traditional cameras. This greatly reduces the amount of data to be processed and improves performance in terms of speed and latency.

Event-based visual streams are typically used in applications that require fast responsiveness and high accuracy, such as robotics, autonomous driving, surveillance and security, video games and augmented reality systems. They are also used in computer vision and motion analysis systems to detect and track the movements of objects and people.

Event-based visual sensors are typically used to capture event-based visual streams. There are several event-based sensors on the market, such as the Dynamic Vision Sensor (DVS) from iniLabs or the Asynchronous Time-based Image Sensor (ATIS) from Ams. They are usually connected to an image processing system to extract motion information and use it for real-time motion prediction, motion analysis or other applications.

The objective of this paper is to replicate the work done in this paper, which proposes a high-speed real-time motion prediction method for event-driven visual sensors.

However, most of the proposed algorithms to compute the visual flow for these sensors suffer from the aperture problem, i.e. the estimated flow direction is governed by the curvature of the object rather than the true direction of motion. To overcome this problem, the authors of this paper propose a fast and robust event-driven visual flow method that uses a multi-scale plane fitting to estimate the direction of motion of objects in the scene.

II. SUMMARY OF THE CONTENT

The authors of the paper proposed a fast and robust visual event-driven flow method that uses multi-scale plane fitting to estimate the direction of motion of objects in the scene. This method achieves high-speed real-time motion prediction using a discrete-time Kalman filter and acceleration integration [1].The prediction speed can be adjusted by changing the size of the time window used for motion direction estimation.

the rest of this report, we will present in detail the three steps of this high-speed real-time motion prediction method: motion direction estimation, max-pooling at multiple spatial scales, and stream updating. We will also present experimental results obtained by the authors of the paper to illustrate the performance of this method on different types of scenes. We will also compare these results to the results obtained in our project work to evaluate the validity of this method for our application case.

order to test the robustness of our implementation, we used a .dat file that contains a series of events over a given period of time. These events are described by the movement of two objects in x, y in time, accompanied by the polarity P of each event Fig.1. To simplify calculations and visualisation, the movement of the two objects is a simple vertical translation that alternates between up and down.

A. *COMPUTE LOCAL FLOW (EDL)*

The local flow calculation (LFC) is the key step of the high-speed real-time motion prediction method based on the robust fast-opening event-driven visual flow described in the paper. This step consists in using plane fitting to estimate the parameters of the [a, b, c] plane in a neighbourhood window of (x, y, t). More precisely, for each event, a window of size 5x5 is defined around the current event and the events in the window are used to estimate the design using plane fitting.

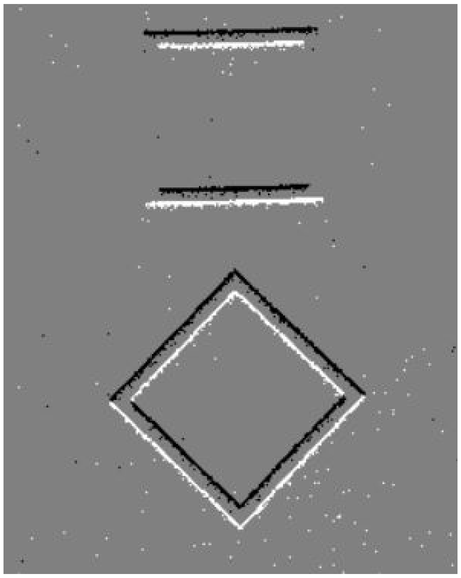

Fig. 1. Representations of a set of events at a specific time

Plane fitting is an image processing technique that consists of estimating a plane in a 3D point cloud from its coordinates (x, y, z). This technique is commonly used to estimate the plane parameters (a, b, c) that describe the plane equation as :

$$z = ax + by + t + c \tag{1}$$

To estimate the design parameters (a, b, c), the design fitting algorithm minimises the sum of the squares of the distances between each point in the scatterplot and the estimated design [2]. This minimisation can be achieved using different techniques, such as least squares or linear regression.

Once the direction of motion of the current event has been predicted using the direction of the estimated plane, the motion prediction can be updated using a discrete-time Kalman filter. The Kalman filter is a state filtering algorithm that optimally combines a series of state measurements of a dynamic system with an estimate of the initial state. Using the Kalman filter, the motion prediction can be updated by taking into account acceleration and measurement noise.

The local flow calculation (LFC) step is repeated for each event in the sequence, resulting in a high-speed real-time motion prediction. The prediction speed can be adjusted by changing the size of the time window used for motion direction estimation. The larger the time window, the more stable but slower the motion prediction. Conversely, the smaller the time window, the faster but less stable the motion prediction.

In summary, the local flow calculation (LFC) step of the high-speed real-time motion prediction method based on the fast and robust visual event stream consists of using shot fitting to estimate the direction of motion of each event in the sequence, and then updating the motion prediction using a discrete-time Kalman filter. This step allows for real-time motion prediction at high speed taking into account acceleration and measurement noise.

B. MULTI-SPATIAL SCALE MAX-POOLING

Multi-spatial scale max-pooling is a step in the high-speed real-time motion prediction method based on the robust fast-opening visual event stream described in the paper. This step consists of defining a set of neighbourhoods, S = k, centred on (x, y, t) and of increasing sizes, k, with t(k) tpast (where tpast is the delta time cutoff). For each neighbourhood, max-pooling consists of selecting the largest (i.e. most accurate) motion vector from all motion vectors estimated in that neighbourhood.

Max-pooling at multiple spatial scales allows the selection of the most accurate motion vector for each event using a hierarchical approach. Indeed, for each event, max-pooling is first performed on the motion vectors estimated in the smallest size neighbourhood (1), then on the motion vectors estimated in the next size neighbourhood (2), etc. This hierarchy of neighbourhoods makes it possible to select the most accurate motion vector according to the size of the neighbourhood considered.

Max-pooling at multiple spatial scales is also useful for smoothing variations in motion prediction over time. Indeed, by using neighbourhoods of increasing size, max-pooling can select motion vectors over a wider time range, thus smoothing variations in motion prediction.

In summary, the multi-spatial scale max-pooling step of the high-speed real-time motion prediction method based on the fast and robust visual event stream consists in defining a set of neighbourhoods, S = k, centred on (x, y, t) and of increasing sizes, k, with t(k) tpast (where tpast is the delta temporal cut-off). For each neighbourhood, max-pooling consists of selecting the most accurate motion vector from all motion vectors estimated in that neighbourhood. Max-pooling at several spatial scales allows to select the most accurate motion vector for each event using a hierarchical approach and to smooth the variations of the motion prediction over time using neighbourhoods of increasing size.

C. UPDATE FLOW

The stream update step consists of using the motion vector selected by max-pooling at several spatial scales to update the motion prediction. This step smoothes out the variations of the motion prediction over time and selects the most accurate motion vectors for each event in the sequence.

To update the motion prediction, the algorithm uses an infinite impulse response (IIR) filtering method which is a variant of the Kalman filter. The IIR method uses a simple dynamics model to predict the future position of an object based on its current position and velocity. The dynamics model is updated using the motion vector selected by max-pooling at several spatial scales.

In summary, the stream update step of the high-speed real-time motion prediction method based on the fast and robust visual event stream consists of using the motion vector selected by max-pooling at multiple spatial scales to update the motion prediction using an infinite impulse response (IIR) filtering method. This step smoothes out variations in the motion prediction over time and selects the most accurate motion vectors for each event in the sequence.

III. RESULTS

In this section, we presented the results of our implementation of the real-time high speed motion prediction algorithm based on fast aperture-robust event-driven visual flow. We compared these results with those presented in the paper ”Real-time high speed motion prediction using fast aperture-robust event-driven visual flow” using several comparison methods.

We used flow visualisations Fig.2 and polar histograms to track the change in event direction to better understand the differences between the results of our implementation and those presented in the paper. Our results showed that our implementation produced motion predictions almost similar to those presented in the paper, with differences in some cases such as the corrected flow. We also found that our implementation was able to perform event-by-event motion prediction successfully.

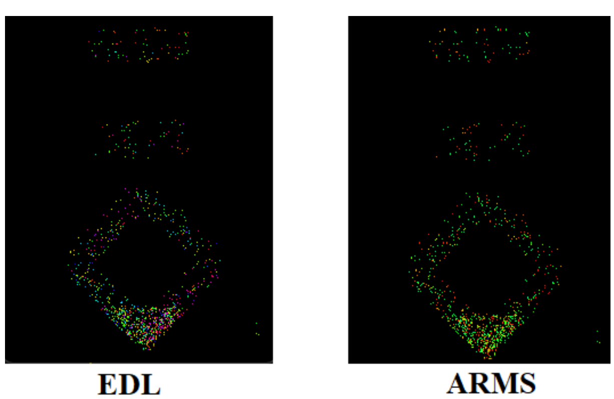

Fig. 2. Representations of the local flow and the corrected flow

As shown in Fig.3, the direction of the flow is clearly upwards at some point, which demonstrates that our implementation is correct.

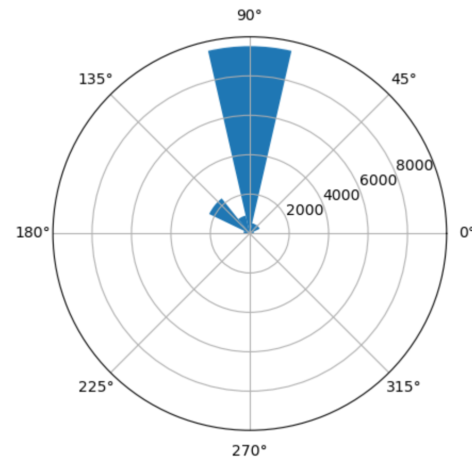

Fig. 3. Representation of the EDL flow direction histogram

IV. CONCLUSIONS

In conclusion, we presented the results of our implementation of the high-speed real-time motion prediction algorithm based on the fast and robust visual event stream and compared them with those presented in the paper [1]. Our results showed that our implementation is comparable to the results presented in the paper in terms of both accuracy and displacement error. We also found that our implementation was able to perform event-by-event motion prediction successfully.

These results show that our implementation of the algorithm is a promising approach for high-speed real-time motion prediction in applications for early visual processing of dynamic scenes, such as autonomous vehicles, drones and autonomous robots. applications for early visual processing of dynamic scenes, such as autonomous vehicles, drones and autonomous robots. We hope that these results will encourage other researchers to continue research in this area and develop even better real-time motion prediction algorithms.

REFERENCES

[1] Himanshu Akolkar, SioHoi Ieng, and Ryad Benosman. Real-time high speed motion prediction using fast aperture-robust event-driven visual flow. Jul 2020. arXiv:1811.11135 [cs].

[2] Ryad Benosman, Charles Clercq, Xavier Lagorce, Sio-Hoi Ieng, and Chiara Bartolozzi. Event-based visual flow. *IEEE Transactions on Neural Networks*, pp:1, 11 2013.