A Motion Sensor with On-Chip Pixel Rendering Module for Optical Flow Gradient Extraction

Jing Huang, Menghan Guo, Shizheng Wang and Shoushun Chen
VIRTUS IC Design Centre of Excellence, School of EEE
Nanyang Technological University, Singapore

Abstract—This work introduces a pixel rendering module (PRM) into an asynchronous event-based dynamic vision sensor (DVS) targeting for optical flow extraction. Optical flow using event-based cameras draws more attention since DVSs directly provide motion related information related and greatly reduce the data redundancy compared to conventional frame-based cameras. Although event-based optical flow has a high potential on real-time performance, its accuracy is limited by event sparseness and lack of intensity, especially for fast motion and highly textured areas. This paper presents a motion sensor with PRM and asynchronous gray-level events to settle these issues. The PRM enables each pixel to communicate with its neighbor pixels such that a single active pixel can force activate its neighboring inactive pixels to provide sufficient data for optical flow calculation. Furthermore, the sensor outputs asynchronous event packages including pixel position, time-stamp and its corresponding illumination. A 64 × 64 prototype was fabricated in 0.35um 2P4M Opto process. Each pixel occupies a footprint of 40 × 40 μm² with 17.7% fill factor.

I. INTRODUCTION

After many years of research, real-time optical flow computation is still a crucial requirement for vision tasks such as optical flow-based segmentation, motion detection, object tracking and obstacle avoidance for aircraft and vehicles. In real-time applications, where optical flow is a fundamental decisive element, there is a pressing need to speed up flow computation while maintaining a high degree of accuracy. Traditional optical flow methods process data from large amounts of frames captured by conventional image sensors. Redundant data is thus repeatedly generated by static background. Readout and processing of large redundancy results in enormous computational cost and limits processing speed.

Event-based motion sensors [1–3] have shown great potential to improve the real-time performance of optical flow computation. Compared with conventional cameras, the motion sensors asynchronously respond to events that represent relative intensity change. The output of the sensor is a stream of asynchronous digital events, which are not limited by exposure time or frame rate. It can detect fast moving objects which are commonly captured by expensive, high speed cameras running at thousands of frame rate, but with largely redundancy-reduced output data. Benosman et al. [4, 5] proposed event-based optical flow methods, including event-based Lucas Kanade method and Local Plane Fitting method. One important step in these event-based optical flow methods is to extract gradient based on pixel intensities in a local area. However, in conventional DVS system, the events only report the pixel position without illumination. Thus, they use the number of accumulated events during a short period to simulate the intensity of each pixel. Nevertheless, the simulate method is inaccurate because it represents the relative intensity change instead of real-time intensity level. Another issue that limits the accuracy of event-based optical flow estimation is event sparseness when detecting fast moving objects. Pixels in conventional DVS work independently and an event generated by a single activated pixel cannot provide sufficient information for optical flow estimation.

In this paper, we present a motion sensor in which each pixel contains a PRM and an intensity readout mechanism. Pixels use PRM to communicate status with neighbor pixels such that an active pixel can force its neighbor pixels to generate events even though they are not triggered. The idea is illustrated in Fig.1 where black pixels represent the motion-triggered active pixels and yellow ones represent the neighbor-triggered active pixels due to PRM. In this way, the sensor assures sufficient data to compute the optical flow of motion-triggered pixels. In addition, the sensor outputs asynchronous event packages and each of them contains the corresponding illumination besides pixel position and time-stamp. The sensor takes the output of the logarithmic detector as pixel intensity, therefore, the intensity of each event represents the absolute intensity level.
detected by the corresponding pixel. With these novel features, the designed motion sensor provides sufficient information for evaluating the optical flow of motion-triggered events. The main effort in this work is on the pixel design.

The rest of the paper is structured as follows. Section II describes the details of pixel architecture and PRM design. The VLSI implementation of the chip is discussed in Section III. Section IV reports the experimental results and Section V concludes this paper.

II. SENSOR DESIGN

In the designed motion sensor, pixels are connected with each other. Each pixel communicates with its neighbor pixels instead of working individually. Assume a single pixel \(P(x, y)\) becomes active triggered by motion while its neighbor pixels \(P(x-1, y), P(x+1, y), P(x, y-1)\) and \(P(x, y+1)\) are inactive, the active pixel \(P(x, y)\) will send signals simultaneously to the neighbor pixels to force activating them, as shown in Fig. 2 (a). Active pixels, both motion triggered or neighbor triggered, generate event packages including pixel location, time-stamp and the corresponding intensity. Therefore, for a single motion triggered pixel, the sensor outputs five event packages used to extract gradient parameters for optical flow estimation. Gradient-based optical flow is estimated based on an important constraint, i.e. Brightness Constraint, which is expressed as \(I_xu + I_yv = -I_t\), where \(I_x, I_y\) indicate the gradients along horizontal and vertical directions. The collected five event packages can thus be calculated to extract \(I_x = [I(P(x+1,y)) - I(P(x-1,y))]/2\) and \(I_y = [I(P(x,y+1)) - I(P(x,y-1))]/2\). According to the theorem of Markov Random Field (MRF) [6], additional information from four neighbors are sufficient enough to extract optical flow gradient parameters. Therefore, each pixel is connected to its four neighbors instead of eight neighbors or even more such that it improves the accuracy of optical flow estimation while the size of data is still acceptable for real-time processing.

A. PIXEL ARCHITECTURE

The pixel architecture is illustrated in Fig. 2 (b). The pixel is designed by introducing the PRM to the previously designed pixel structure in [3]. The new pixel can be simplified into several blocks such as photo-receptor, event generator, logarithmic pixel readout, AER handshaking logic and pixel rendering module. The photo-receptor consists of one NMOS transistor with photo-diode and a negative feedback amplifier. Its logarithmic response provides high dynamic range, at the cost of sacrificing temporal contrast sensitivity at high illumination. The logarithmic photo-receptor transforms the intensity to a voltage signal in real-time. The transformed intensity signal is fed into a delta modulator followed by two comparators. An OR gate merges the VH and VL events which corresponds to light changes from either dark to bright or vice versa. An external forcefire signal is also merged to unconditionally force all pixels to set their respective status latch 1 (SL-P) and generate requests through the AER handshaking circuits. SL-P can be set by either the output of the VH or VL comparator or the external forcefire signal. The output signal of SL-P (EP) is active high and connected with its neighbor pixels. An inter-pixel handshaking logic, PRM module, is introduced to monitor primary event status (EP) of its four neighbor pixels (Left, Right, up and down). If one or more neighbor pixels become active, i.e. signal EP switches from low to high, the inter-pixel handshake generates a signal EN, indicating that at least one of the neighbor events is active, and initial the AER handshaking circuits. The AER handshaking logic can thus be initialized by intensity variation detection, forcefire control or PRM module. The pixel reports the intensity in real-time, which is read out synchronously with event position, without the limitation of exposure time, thus it takes no effort to match the event position and intensity information, in both spatial and temporal domain.
The switch transistor (Msw) functions as a source follower which buffers the output voltage of logarithmic detector onto the shared bus (Col_analog_bus). The switch transistor (Msw) is controlled by RA such that only pixels on the acknowledged row would buffer out the real-time intensity. The readout analog value is quantized using the same structure as in [3]. In this way, each pixel event packet consists of three elements: pixel address, time-stamp and its illumination. Taking use of the spatial-temporal character of events, our sensor can be used to compute flow using real-time event-based flow methods [8]. The introduced asynchronous intensity brings motion sensor to wider application using algorithms based on feature extraction [9].

III. SENSOR IMPLEMENTATION

A 64 × 64 test prototype of dynamic vision sensor with PRM was implemented in AMS 0.35μm 2P4M process. Fig.4(a) shows the layout of the sensor die, of size 4.2 × 5.9 mm². Pixel area is 40 × 40 μm², as shown in Fig.4(b). Analog circuitry and digital circuitry are placed separately as far as possible to reduce crosstalk and signal coupling. The capacitors are placed next to each other for better matching. In addition, metal shielding layer is also used.

IV. EXPERIMENTAL RESULTS

To test the performance of the sensor, a FPGA-based testing platform was developed based on an Opal Kelly FPGA board. The board communicates with a local host PC through USB link and is programmed for sensor control signal configuration, temporary data storage and data transfer. Each received event packet consists of three elements: pixel address, time-stamp and its illumination. The positions of received events can describe the contour of a moving object while the intensity information is used to implement partial of optical flow algorithms. Full frame pictures can be generated by issuing a forcefire signal, all the pixels on the sensor will be activated and read out one by one. During the period of a full-frame readout, the motion sensor can operate motion detection as normal. Table 1 summarizes the sensor specifications and its main features. The sensor is powered at 3.3V and clocked at 40MHz.
TABLE I: Chip Characteristics

<table>
<thead>
<tr>
<th>Process Technology</th>
<th>AMS 0.35 μm OPTO process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Supply</td>
<td>3.3V</td>
</tr>
<tr>
<td>Chip Size</td>
<td>4.2 × 5.9 mm²</td>
</tr>
<tr>
<td>Array Size</td>
<td>64 × 64 Pixels</td>
</tr>
<tr>
<td>Pixel Size</td>
<td>40 × 40 μm²</td>
</tr>
<tr>
<td>Fill Factor</td>
<td>~17.7%</td>
</tr>
<tr>
<td>Fixed Pattern Noise</td>
<td>0.79%</td>
</tr>
<tr>
<td>Maximum Clock Speed</td>
<td>50 MHz</td>
</tr>
<tr>
<td>Dynamic Range</td>
<td>1-100k lux</td>
</tr>
<tr>
<td>Power Consumption</td>
<td>200 mW at 20 MHz</td>
</tr>
</tbody>
</table>

Fig. 5: Setup platform for testing.

Fig. 5 shows the test platform which is placed on an optical table in a dark room. The sensor was mounted on a camera holder fixed to a linear motor [10] and the target pattern is fixed on a stationary black board. The high precision linear motor provides uniform motion with a speed of 0.5 meter per second. When the camera moves along the motor track, the object is observed to move towards left or right.

Fig.6 presents sample images with flow estimation. The full frame image in Fig.6(a) is generated by the force fire control signal, which forces all pixels to become active and capture intensity. During motion, event packages are generated continuously in spatial and temporal domain. To estimate the spatial gradients, slices are generated by accumulating events during last 100μs. The accumulation time period is based on experimental results, such that the accumulated events after noise filtering is able to construct clear features. Fig.6(b) shows binary value of a sample event slice. The binary event image is made up of crosses due to the PRM. Fig.6(c) presents the same event slice containing active pixels with gray-level intensity that are used to compute optical flow. In this experiment, we use the classic+NL-fast method [11], which is a classical gradient-based method with high performance, to evaluate the optical flow of events. The flow result in vector is shown in Fig.6(d), indicating clearly that the object is moving to the left. The improvement on accuracy of optical flow estimation and the incremental size of readout data compared to the flow estimation using traditional DVSs is currently under evaluation.

V. CONCLUSION

In this paper, we present a smart motion sensor with PRM and a capability to generate events with asynchronous absolute intensity. With the PRM and asynchronous intensities, the designed motion sensor provides sufficient pixel information to extract gradient parameters in optical flow estimation, which settles the accuracy issues caused by conventional DVS. With an estimation of 2.5 incremental size of readout data, the speed of this motion sensor can still be equivalent to a high speed camera with several hundred frame rate. The new features will favor the optical flow estimation with real-time performance and higher accuracy.

REFERENCES


