Methods to Solve Salt & Pepper Noise, and Frame Dropping of Timed Address Event Representation Vision Sensor

Lu YU1, Zhonghe Chen1, Yun HAO1

1Zhonghuan Information College Tianjin University of Technology, Tianjin, 300380, China
E-mail: yuluyouxiang@sina.com

Abstract: Address event representation (AER) vision sensor which only outputs visual information of a pixel with light intensity change, eliminates redundant information radically. Compared with the traditional vision sensor, it has the advantages of high frame rate, low data volume and high dynamic range. Firstly, this paper studies the method of eliminating salt & pepper noise in AER vision sensor. By design of arbitrator, the salt and pepper noise can be removed directly from the source through 5×5-nearest neighbor algorithm and cross window-nearest neighbor algorithm. This method not only reduces the output data and preserve the details of the image, but also removes the salt and pepper noise effectively. Secondly, this paper solves the problem of "frame dropping" caused by the time delay and the non-synchronization of ON/OFF events and quantization process. The image is corrected by moving vector which is obtained by Mean Shift target tracking algorithm, and then cross window-nearest neighbor algorithm is used to remove noises. "Filling frame" is completed and a better image is obtained.

1. Introduction

Traditional vision sensor takes "frame" as the basic unit of image, which cannot meet the requirements of high frame rate, low data volume, high dynamic range and high transmission rate[1-2]. Address event representation (AER) vision sensor based on bionic vision, which only outputs visual information of a pixel with light intensity change, eliminates redundant information radically[3-4]. It uses asynchronous output and sparse representation to obtain images with high frame rate, low data volume and high dynamic range. According to advantages above, AER vision sensor is especially suitable for high-speed target shooting and tracking, target recognition, acquisition of high resolution image and other fields[5-6].

AER vision sensor is also defective. In the image follow-up processing, the image information acquired by AER vision image sensor needs to be transformed into "frame" for processing, which causes the redundancy eliminated at the source to be restored in the process of image processing. Moreover, arbitrator is needed due to the centralized output of the data conflict. This results in the shielding delay in quantization process and the "frame dropping" problem[7].

2. The structure and work flow of AER vision sensor

Figure 1 shows the structure of AER vision sensor. It is composed of pixel units, row/column control and arbitrator, address/event encoder, external image processing and data transmission control unit[8].
Figure 2 shows the structure of a pixel unit, which consists of a CD (change detector) unit and a double sampling PWM circuit unit. CD unit can sense the change of light intensity. Double sampling PWM circuit unit quantizes the light intensity[9].

2.1 Light detect
When the light intensity changes, the CD unit in a pixel judges whether the light intensity changes beyond the threshold set. If the change of light intensity exceeds threshold, time pulses will be generated. If it does not exceed the threshold, the pixel will not output any data and return to ready state. It is the foundation which AER vision image sensor generate “frameless” output based on.

2.2 Row/column control and arbitrator
The event pulse sends a request signal to the row arbitrator. When the arbitrator confirms that the row is selected, it returns the response signal. If row response signal is received, all pixels on the row emit a column request signal RH (ON event, intensity-enhanced event) or signal RL (OFF event,
intensity-weakened event) to the column arbitrator, according to different event type.

If there is only one column request signal emitted, the pixels in the column are directly selected and output quantization of light by CD unit to address/event encoder. When there are several column request signals emitted at the same time, the column arbitrator will choose only one pixel, and transmits response signal to make it output quantization of light, while the pixels in the other column are in standby mode waiting next arbitration. Then address/event encoder encodes the address of the pixel and quantization of light, and also gives time stamp on it. In this way, the address, light intensity and time stamp of a pixel is gotten and are emitted to control module.

2.3 Quantization of light in a pixel
As shown in the figure 2, when a pixel gains control of the bus, the double-sampling PWM circuit unit generates two pulses according to different light intensity. The time between the pulses is proportional to the current of light intensity. The counter quantizes the time difference and saves the quantization results in the pixel memory waiting for output.

2.4 Output and processing of quantized data
The quantized data, 8 bit, are send to control module, regarded as a whole with row/column addresses and time stamps.

In conclusion, AER vision sensor does not have the concept of "frame" to sample image information. Each pixel works independently and outputs asynchronously and serially, eliminating redundant information from the source.

However, the image processing of AER vision sensor always restores it back to the "frame" image, which does not take full advantage of its sparse representation and low data volume.

3. Eliminating salt & pepper noise
Salt & pepper noise is a form of noise sometimes seen on images. It is also known as impulse noise. This noise can be caused by sharp and sudden disturbances in the image signal. It presents itself as sparsely occurring white and black pixels. An effective noise reduction method for this type of noise is a median filter or a morphological filter. For reducing either salt noise or pepper noise, but not both, a contra harmonic mean filter can be effective.

Rely on the ON/OFF event is generated by CD unit in AER visual image sensor, the arbitrator can judge whether there is salt & pepper noise in the pixel. If yes, double-sampling PWM circuit unit is not needed to quantize and there is no output. And the quantization of light intensity is still the value before. If no, double-sampling PWM circuit unit quantizes light intensity and outputs it. And the quantization of light intensity is current data.

In this paper, two methods are used to determine whether the pixels of the event output by arbitrator to remove salt and pepper noise from the source.

3.1 5×5-nearest neighbor algorithm
When one pixel (i, j) sends a ON/OFF event request signal to row and column arbitrators, the arbitrators statistics the number of events of 16 pixels in the periphery of the 5×5 pixels, which is yellow pixels in Figure 3.

If the number of pixels occurred ON/OFF event is less than or equal to 8, no matter how much ON/OFF events occur in the 3×3 nearest neighbor, which is green pixels in Figure 3, the pixels in the 3×3 nearest neighbor and itself don’t output data. The quantization of light is still as same as before in the external register. In other cases, the pixels in the 3×3 nearest neighbor and itself output data, and then the quantization of light is replaced by output data of CD unit at this moment.
3.2 Cross window-nearest neighbor algorithm

When one pixel \((I, j)\) sends a ON/OFF event request signal to row and column arbitrators, the arbitrators statistics the events’ number of up, down, lift, and right of the pixel. If the number of pixels occurred ON/OFF event is less than or equal to 3, the pixels in cross window-nearest neighbor and itself don’t output data. The quantization of light is still as same as before in the external register. In other cases, the pixels in the cross window-nearest neighbor and itself output data, and then the quantization of light is replaced by output data of CD unit at this moment.

3.3 Simulation result

Figure 4 shows the simulation result of elimination of salt and pepper noise. Fig.4(a) shows the first frame without noise. Fig.4(b) shows the second frame with salt and pepper noise. In order to simulate the process of AER vision image sensor, the image of second frame is obtained by translating the first frame along - 45 degrees, and adding salt and pepper noise with a noise coefficient of 0.1. Fig.4(c) shows the image after median filter from the second frame. Fig.4(d) shows ON/OFF event output of the second frame in AER mode, which represented as the difference between the first frame and the second frame. When there is no event to output, zero is stored. When there is ON/OFF event to output, one is stored. Fig.4(e) shows the second frame gotten by cross window-nearest neighbor algorithm. Fig.4(f) shows the second frame gotten by 5×5-nearest neighbor algorithm.

When ON/OFF event rate is 46.19%, the output rate in median filter is 100%, in cross window-nearest neighbor algorithm is 0.93%, in 5×5-nearest neighbor algorithm is 12.85%.
Compared with median filter for image acquisition only processing a single image, the two methods proposed in this paper are based on two consecutive frames. As shown in Fig.4, the two methods provide better image quality, especially in maintaining the details of the image. Furthermore, cross window-nearest neighbor algorithm depresses noise better than 5×5-nearest neighbor algorithm, and outputs less data.

4. "Frame dropping" caused by the time delay and its solution

Part 2 of this paper describes the work flow of AER visual image sensor. It can be found that a pixel can’t quantize light after it sent ON/OFF event signal until response signal returns by the arbitrator. When the intensity of light changes strongly, there are much more pixels sending request signal. The beginning time of quantizing process is much later than ON/OFF event, caused by waiting response signal returned by the arbitrator. In this way, the quantification of light intensity isn’t in the time ON/OFF event occurring, but in the time response signal returning. As a result, there will be "losing frames" problems caused by time delay. This problem is particularly prominent in high-speed moving objects shooting and large pixel array vision image sensors.

AER vision image sensor only outputs visual information of a pixel with light intensity change, so there is no "frame" concept."Frame dropping" refers to the value of the next light intensity quantization cycle replacing the value in present moment.

Figure5 shows "frame dropping". Fig.5(a)-(c) are c images without time delay in which the car is moving down the right side unknown its angle and speed. When "frame dropping" happens in the second frame, some light intensity of its pixels is replaced by the intensity of the third frame. As shown in Fig.5(d), it is evident that there are distinct noise points similar to salt and pepper noise on moving objects, and these points are mainly concentrated in areas with more details.
Because of the correlation between the frames of AER vision sensor, the method of video repair can be used to solve the problem of frame dropping. The algorithm is shown in Figure 6.

1. Mean Shift target tracking algorithm[10] is used to process the first and second frames to obtain the moving vectors including the direction and distance of the moving object.

2. Moving and clipping the first frame image according to the moving vector obtained in the previous step, the first frame image is corrected.

3. Cross window-nearest neighbor algorithm is used to the new first and second frames, and "frame dropping" reduces just like the way eliminating salt & pepper noise.
Figure 7 shows the simulation result in reducing frame dropping. Fig. 7(a) shows the second frame with frame dropping, where there are several blurred target contour and multiple noise points. Fig. 7(b) shows the simulation result of mean filter. The new light intensity in the second frame is average value of the initial second frame, the new first frame and the new third frame corrected by Mean Shift target tracking algorithm. It can be clearly seen that, because the moving vector is not necessarily 100% accurate, there will be a problem of ghosting image (part of dotted line is obvious). At the same time, the noise point has been improved, but the improvement effect is not obvious. Fig. 7(c) shows the simulation result of the new second frame by cross window-nearest neighbor algorithm after the new first frame corrected by Mean Shift target tracking algorithm. The problem of ghosting image is eliminated effectively, further more frame dropping is corrected well.

In addition, the algorithm to get moving vectors in this paper is Mean Shift target tracking algorithm. Other algorithm can also be used to obtain moving vectors. The key is acquiring exact moving vectors to correct the first frame.
5. Conclusion

This paper is based on the research of AER vision image sensor. Firstly, this paper studies the method of eliminating salt & pepper noise in AER vision sensor. By design of arbitrator, the salt and pepper noise can be removed directly from the source through 5x5-nearest neighbor algorithm and cross window-nearest neighbor algorithm. This method not only reduces the output data and preserves the details of the image, but also removes the salt and pepper noise effectively. Secondly, this paper solves the problem of "frame dropping" caused by the time delay and the non-synchronization of ON/OFF events and quantization process. The image is corrected by moving vector which is obtained by Mean Shift target tracking algorithm, and then cross window-nearest neighbor algorithm is used to remove noises. "Filling frame" is completed and a better image is obtained.

References

[1] Belbachir A N, Hofstatter M, Litzenberger M, et al. High-Speed Embedded-Object Analysis Using a Dual-Line Timed-Address-Event Temporal-Contrast Vision Sensor[J]. IEEE Transactions on Industrial Electronics, 2011, 58(3):770-783.
[2] Belbachir A N, Hofstatter M, Litzenberger M, et al. High-precision shape representation using a neuromorphic vision sensor with synchronous address-event communication interface[J]. Measurement Science & Technology, 2009, 20(10):104007.
[3] Posch C, Matolin D, Wohlgenannt R. A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS [J]. IEEE Journal of Solid-State Circuits, 2011, 46(1):259-275.
[4] Posch C, Matolin D, Wohlgenannt R. High-DR frame-free PWM imaging with asynchronous AER intensity encoding and focal-plane temporal redundancy suppression [C]. IEEE International Symposium on Circuits and Systems, Paris, 2010: 2430 - 2433.
[5] Posch C, Matolin D, Wohlgenannt R. High-DR frame-free PWM imaging with asynchronous AER intensity encoding and focal-plane temporal redundancy suppression [C]. IEEE International Symposium on Circuits and Systems, Paris, 2010: 2430 - 2433.
[6] Matolin D, Wohlgenannt R, Litzenberger M, et al. A load-balancing readout method for large event-based PWM imaging arrays [C]. Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS), Paris, 2010: 361 – 364.
[7] J Xu, D Li, S Yao. A Time Error Correction Method Applied to High-Precision AER Asynchronous CMOS Image Sensor [J]. Journal of Signal Processing Systems, 2014, 75(1):1-13.
[8] Yu Lu, Yao Suying, Xu Jianguo. An Implementation Method of Real-Time Vision Sensor Based on Address Event Representation [J]. Acta Optica Sinica, 2013(1):251-257.
[9] Yu L, Hao Y, Chen Z, et al. A Time Error Model for Correlated Double Sampling PWM Pixel[C]/
International Conference in Communications, Signal Processing, and Systems. Springer, Singapore, 2016:669-679.

[10] Jiangtao Xu, Mengxing Zhang, Shi Yan, et al. A Method to Solve the Side Effects of Dual-Line Timed Address Event Vision System [J]. Journal of Circuits Systems & Computers, 2015, 24(03):1550028.