VIDEO SURVEILLANCE AT AN INDUSTRIAL ENVIRONMENT USING AN ADDRESS EVENT VISION SENSOR

Comparative between Two Different Video Sensor based on a Bioinspired Retina

Fernando Perez-Peña, Arturo Morgado-Estevez, Rafael J. Montero-Gonzalez
Applied Robotics Research Lab, University of Cadiz, C/Chile 1, Cadiz, Spain

Alejandro Linares-Barranco, Gabriel Jimenez-Moreno
Robotic and Technology of Computers Lab, University of Seville, Av. Reina Mercedes, Seville, Spain

Keywords: Bio-inspired, Video, Industrial Surveillance, Spike, Retinomorphic Systems, Address Event Representation.

Abstract: Nowadays we live in very industrialization world that turns worried about surveillance and with lots of occupational hazards. The aim of this paper is to supply a surveillance video system to use at ultra fast industrial environments. We present an exhaustive timing analysis and comparative between two different Address Event Representation (AER) retinas, one with 64x64 pixel and the other one with 128x128 pixel in order to know the limits of them. Both are spike based image sensors that mimic the human retina and designed and manufactured by Delbruck's lab. Two different scenarios are presented in order to achieve the maximum frequency of light changes for a pixel sensor and the maximum frequency of requested pixel addresses on the AER output. Results obtained are 100 Hz and 1.88 MHz at each case for the 64x64 retina and peaks of 1.3 KHz and 8.33 MHz for the 128x128 retina. We have tested the upper spin limit of an ultra fast industrial machine and found it to be approximately 6000 rpm for the first retina and no limit achieve at top rpm for the second retina. It has been tested that in cases with high light contrast no AER data is lost.

1 INTRODUCTION

It is easy to find a good surveillance or monitoring system for an industrial environment but not at all for ultra fast industrial machinery. The first system could be formed by a network of commercial cameras and complex software for tracking objects and humans that even includes an intelligent procedure as the one showed by Fookes, Denman, Lakemond, Ryan, Sridharan, and Piccardi (2010). At Messinger and Goldberg (2006) a large study of different techniques and critical components of this type of networks is presented.

But in this paper we propose a surveillance system based on an AER retina. This visual sensor mimics the human retina and thus, it produces events instead of frames with a quick response, what definitely implies live time at ultra fast industrial machinery. Our purpose could be complemented with any of the previous one; an AER retina inside or next to industrial machinery. It could be also very interesting in this kind of surveillance systems an ultra fast face detection like the one describe on He, Papakonstantinou and Chen (2009) based on a novel SoC (System on Chip) architecture on FPGA. The detection speed reaches 625 fps.

It is necessary to know the human retina behavior to perform the results and to understand the real time vision system presented in this paper.

The human retina is made up of several layers. The first one is based on rods and cones that capture light. The following three additional layers of neurons are composed of different types of cells (Linsenmeier, 2005). Horizontal cells implement a previous filter, the bipolar cells are responsible for the graded potentials generation. There are two different types of bipolar cells, ON cells and OFF cells. The last type of cells of this layer are the amacrines, they connect distant bipolar cells with ganglion cells. The last layer of the retina is composed by ganglion cells. They are responsible for the action potentials or spikes generation.

Perez-Peña F., Morgado-Estevez A., J. Montero-Gonzalez R., Linares-Barranco A. and Jimenez-Moreno G..
VIDEO SURVEILLANCE AT AN INDUSTRIAL ENVIRONMENT USING AN ADDRESS EVENT VISION SENSOR - Comparative between Two Different Video Sensor based on a Bioinspired Retina.
DOI: 10.5220/0003521701310134
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications* (SIGMAP-2011), pages 131-134
ISBN: 978-989-8425-72-0

AER retina was firstly proposed at 1988 by Mead and Mahowald (1988) with an analog model of a pixel. But it was in 1996 when Kwabena Boahen presented his work (Boahen, 1996) that established the basis for the silicon retinas and their communication protocol. After them, Culurciello, Etienne-Cummings and Boahen (2003) described a gray level retina with 80x60 pixel and a high level of response with AER output. The most important fact in all these works is the design of the spikes generator.

In this paper we use the Delbruck's retinas developed under the EU project CAVIAR (IST-2001-34124). These retinas use the AER communication strategy. If any pixel of the retina needs to communicate a spike, an encoder assigns a unique address to it and then this address will be put onto the bus using a handshake protocol. AER was proposed by Mead lab in 1991 (Sivilotti, 1991) as an asynchronous communication protocol for inter neuromorphic chips transmissions.

2 AER RETINA CHIP

We have used two silicon bio-inspired retinas. Both designed by P. Lichtsteiner and T. Delbruck at Neuroinformatics Institute at Zurich (Lichtsteiner, 2005) and (Lichtsteiner, 2008). These retinas generate events corresponding to the sign of the derivative of the light evolution respect to the time, so static scenes do not produce any output. For this reason, each pixel has two outputs, ON and OFF events or two directions if we look through AER. If a positive change of light intensity within a configurable period of time appears, a positive event is transmitted and the opposite for a negative change.

2.1 Frequency, Tests and Standards of AER Retina

At Delbruck's papers there are several tests to characterize the retinas but we need to know the behavior at the worst condition in order to use the retinas with an industrial manufacturing machinery as a target. It is very important to know exactly the maximum detected change of pixel light in the AER retina in order to determine the maximum frequency of rotation for a particular object. It is also important to know if there is any lose of events at those frequencies.

At CAVIAR project (2009) a standard for the AER protocol was defined by Häfliger. This standard defines a 4-step asynchronous handshake protocol. It stablishes several time parameters defined as follow: t1 is the establishment time for data, t2 from data requested to data acknowledged, t3 goes from the acknowledge data to the disappeared of valid data at the bus; these three times could take any time. Times t4 and t5 are defined from the edge of the acknowledge signal to the deactivation of the request signal and from this point to acknowledge deactivation respectively; they could take just 100ns length. The last time, t6 goes from the final of t5 until a new request is presented and also could take any time.

3 EXPERIMENTAL METHODOLOGY

In this section we present and describe two different methods in order to extract the bandwidth limit and the percent of lost events.

We have used the jAER viewer and Matlab functions, available at the jAER wiki (http://jaer.wiki.sourceforge.net/). Furthermore, a logic analyzer from manufacturer Digiview (Model DVS3100) (Figure 1) has been used.

3.1 Environment

The first one is splitted depends on with retina is the target of the test.

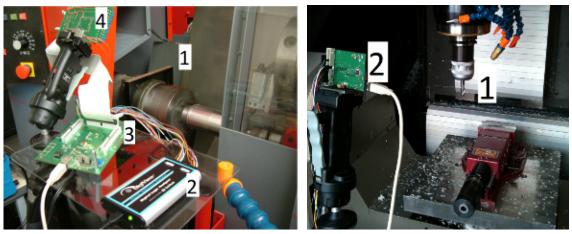

Figure 1: Left: assembly prepared to proceed with the first test for 64x64 retina. The components are 1. Lathe, 2. Logic Analyzer, 3. Sequencer Monitor AER and 4. 64x64 pixel retina, right: assembly prepared to proceed with the first test for 128x128 retina. The components are 1. CNC Machine and 2. 128x128 pixel retina.

The reason to use this type of mechanical tools is because they provided a huge margin of spin frequency. This fact allows us to compare the spin frequency and the maximum frequency of one pixel.

For the second test, we have taken advantage of the fluorescent tubes. Because they change their luminosity with the power network frequency (50 Hz at Spain) it is possible to achieve that all the pixel spiking by focusing the retina on the tubes. With this scenario, the logic analyzer will show the proper times of each spike and the Häfliger times could be extracted.

The Sequencer/Monitor AER board called USB2AER is described by Berner, Delbruck, Civit-Balcells and Linares-Barranco (2007).

3.2 Maximum Spike Frequency

In order to determine the frequency it is necessary to focus on a few pixel of the retina. To obtain this response at both retinas we have stimulated it with a high range of frequency allowed by machinery tools. Both assemblies are showed at Figure 1. Once the retinas have been placed, the Lathe and CNC machine are stimulating just a few pixel of the retinas.

We have used the Java application jAER viewer to take a sequence, MATLAB to processed it in order to know which pixel are spiking and logic analyzer to study the sampled frequency for these pixel for each spin frequency of both machines.

3.3 Maximum Frequency of Requested Addresses

For this test we cannot use the AER monitor board for 64x64 retina because its USB interface will limit the bandwidth peak of events to the size of the buffer and clock speed.

To determine the maximum frequency on the output AER bus of the retina it is necessary to light all pixel with a high frequency changes, in order to study the limit of the arbiter inside the retina that is managing the writing operation of events on the AER bus. The procedure is described by Pérez-Peña, Morgado-Estevez, Linares-Barranco, Montero-Gonzalez and Jimenez-Moreno (2011).

4 RESULTS AND DISCUSSIONS

The results show the evolution of spike frequency for the most repetitive direction calculated in front of the spin frequency of the manufacturing machinery expressed in rpm.

When the spin frequency is increased, the spiking frequency of a fixed pixel increased up to 100Hz which is the saturation level for the first retina.

For the second analyzed retina it is not possible to determine a saturation level for the spike frequency because at the top limit of the CNC machine, 10000 rpm, the retina still catch the movement. But it is possible to enounce that there are peaks around 1.3KHz. These frequency peaks comes up because a fact at those kind of industrial machinery that is the stability of the head.

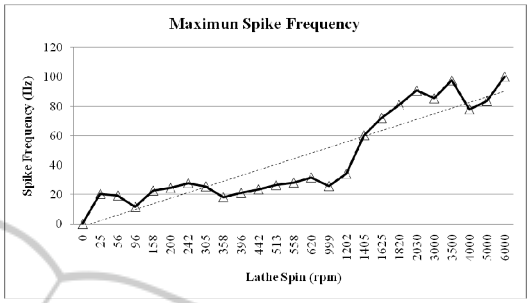

Figure 2: Maximun spike frequency evolution for the 64x64 retina spike frequency in front of the spin of the lathe expressed at rpm.

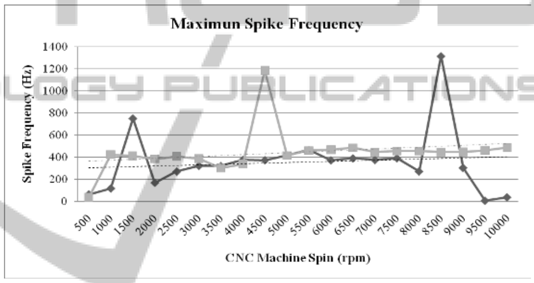

Figure 3: Maximun spike frequency evolution for the 128x128 retina spike frequency in front of the spin of the CNC Machine expressed at rpm.

Working with 64x64 retina, when the spin of the Lathe goes from 6000 rpm to 7000 rpm the target began to disappear from the retina view. This is the empiric limit for this retina.

For the 128x128 retina is not possible to reach an empiric limit because we have no margin to increase the spin frequency beyond 10k rpm.

Another result of this analysis for these retinas should be highlighted: if the maximum frequency is 100 Hz and if we considered the peak of 1.3 KHz, it is necessary to fit the 4096 and 16384 addresses within 10 ms and 769.23 us respectively, in order to aim no miss events.

In both trials, the times by Häfliger standard have been obtained as it is shown in table 1:

Table 1: Timing table obtained at trials.

Times	Lathe Trial (64x64) (ns)	Tube Trial (64x64) (ns)	CNC Machine Trial (ns)	Tube Trial (128x128) (ns)
t1	10	200	770	120
t4	990	60	140	30

At the tube trial we were looking for the maximum frequency of any requested address and it results on 1.88 MHz for the 64x64 retina and 8.33 MHz for the 128x128 retina.

For the 64x64 pixel retina, if we join together the 10 ms obtained at the Lathe scenario between two consecutive events of the same pixel, that could be called frame time, and t2+t4+t5+t6 obtained on the tubes scenario between any two consecutive events, a maximum to 18867 addresses could be placed on the AER bus. If we had considered an address space of 4096 pixel, it would have confirmed the fact of no lost events.

For the 128x128 pixel retina, it is possible to pick up a similar case. If we considered the peak of 1.3KHz (769.23 us) like the top spike frequency for a pixel and join it together with the sum of t2, t4, t5 and t6 obtained on the tubes scenario, it is possible to place 6410 addresses within the frame time. If we had considered an address space of 16384 pixel, it noticed that lost events could appear with the peak frequency selected, but this situation is not actually very real because we have considered the peaks of frequency spikes. If we have taken a frequency of 500 Hz, which is the saturation level for our test, no lost events appears.

5 CONCLUSIONS

We have presented a study of two different retinomorphic systems to use them in a visual surveillance at any industrial environment. We have checked the upper limit of the first system with a Lathe, approximately 6000 rpm and no limits for the second system at usual manufacturing machinery. It has been tested at an ultra fast CNC Machine up to 10000rpm with an excellent result. Also, the results reveal that in the worst condition of luminosity change for our retinas there will be no lost of events for the first one and at the second one could be some very improbable lost event. Therefore, these AER retinas can be used for a visual surveillance system at any high speed industrial manufacturing machinery.

ACKNOWLEDGEMENTS

This work was supported by the Spanish grant VULCANO (TEC2009-10639-C04-02).

Also thanks to group Engineering Materials and Manufacturing Technologies, TEP-027.

REFERENCES

Fookes, C., Denman, S., Lakemond, et al. 2010. Semi-Supervised Intelligent Surveillance System for Secure Environments.

Messinger, G. and G. Goldberg. 2006. Autonomous Vision Networking - Miniature Wireless Sensor Networks with Imaging Technology. *Proceedings of SPIE - The International Society for Optical Engineering.*

He, C., Papakonstantinou, A. and Chen, D. 2009. A Novel SoC Architecture on FPGA for Ultra Fast Face Detection.

Linsenmeier, Robert A. 2005. Retinal Bioengineering. 421-84.

Lichtsteiner, P. and Delbruck, T. 2005. A 64×64 AER Logarithmic Temporal Derivative Silicon Retina. *Paper presented at 2005 PhD Research in Microelectronics and Electronics Conference*. Lausanne. 25 July 2005 through 28 July 2005.

Lichtsteiner, P., C. Posch and T. Delbruck. 2008. A 128x128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor. *Solid-State Circuits, IEEE Journal of* 43:566-76.

Mead, C. A. and M. A. Mahowald. 1988. A Silicon Model of Early Visual Processing. *Neural Networks* 1:91-7.

Boahen, K. 1996. Retinomorphic Vision Systems. Paper presented at Microelectronics for Neural Networks, 1996., *Proceedings of Fifth International Conference on.*

Culurciello, E., R. Etienne-Cummings and K. A. Boahen. 2003. A Biomorphic Digital Image Sensor. *Solid-State Circuits, IEEE Journal of* 38:281-94.

Sivilotti, M.: Wiring Considerations in Analog VLSI Systems with Application to Field-Programmable Networks, Ph.D. Thesis, *California Institute of Technology*, Pasadena CA (1991).

Serrano-Gotarredona, R., Oster, M., et al. 2009. CAVIAR: A 45k Neuron, 5M Synapse, 12G connects/s AER Hardware Sensory-Processing-Learning-Actuating System for High-Speed Visual Object Recognition and Tracking. *IEEE Transactions on Neural Networks* 20:1417-38.

Berner, R., Delbruck, T., et al. 2007. A 5 Meps $100 USB2.0 Address-Event Monitor-Sequencer Interface. *Paper presented at 2007 IEEE International Symposium on Circuits and Systems*, ISCAS 2007. New Orleans, LA. 27 May 2007 through 30 May 2007.

Pérez-Peña F., Morgado-Estevez A., Linares-Barranco A. et al. 2011. Frequency Analysis of a 64x64 pixel Retinomorphic System with AER output to estimate the limits to apply onto specific mechanical environment. *Proceedings presented at International Work Conference on Artificial Neural Networks*, IWANN'11.