A Novel Event Based Image Sensor Architecture

M. AKRARAI$^1$  L. FESQUET$^2$  G. SICARD$^3$

$^1$University Grenoble Alpes, TIMA, Grenoble, France

$^2$CNRS, TIMA, Grenoble, France

$^3$CEA LETI, Grenoble, France

IPSOC 2019
Image Sensors

- Image sensors are present in various aspects of our life.
- Today we face the following challenges:

Power consumption, increasing resolution and frame rate. The last two, increase the resources required for image and video processing, a good example is **Automotive**.

How can we solve these engineering challenges?
Outline

1. State of the art
2. Proposed novel architecture
3. Simulation results
Introduction of **frameless event based** image sensors and processing architectures:

- Compression of temporal redundancy (Tobi Delbruck [2]).
- Compression of spacial redundancy (Amani Darwish [1]).
- Event-based Object Classification (Prophesee [3]).
In both [2] and [1], the output of the image sensors is a **frameless** series of row and column addresses of active pixels:

(a) Block schematic of the IS in [Tobi Delbruck [1]](https://example.com)

(b) Block schematic of the IS in [Amani Darwish [2]](https://example.com)

(c) Event based image sensor output
(a) Principle of a Dynamic Vision Sensor pixel in [1]

(b) Principle of a Time to First Spike pixel in [2]
Can we combine spatial and temporal redundancies suppression and generate an event based processable output?
The combination of the two pixel (DVS, TFS) to harvest the benefits of both.

The TFS will measure absolute luminance only if the DVS pixel detects a change.
Figure: Suggested image sensor kernels

(a) Architecture A depicting 1 DVS pixel per 1 TFS
(b) Architecture B depicting 1 DVS per 3 TFS
(c) Architecture C depicting 1 DVS per 5 TFS
(d) Architecture D depicting 1 DVS per 8 TFS
(e) Architecture E depicting 1 DVS per 24 TFS
To form the image sensor matrix, the kernel is repeated until we reach the desired resolution.
### Simulation Results

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Dimensions</th>
<th>Frame rate</th>
<th>DVS thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway</td>
<td>1200 × 600</td>
<td>30 fps</td>
<td>1,5,10 %</td>
</tr>
<tr>
<td>Parking</td>
<td>1200 × 600</td>
<td>30 fps</td>
<td>1,5,10 %</td>
</tr>
</tbody>
</table>

Worst case, will be detecting all the events from the input videos for 10 seconds:

\[
1200 \times 600 \times 30 \times 10 = 216000000 \text{Events}
\]

(a) Highway scenario: car on the road  
(b) Parking scenario: a man walks in the front of a car

**Figure**: The two simulation scenarios
State of the art

Proposed novel architecture

Simulation results

Highway scene

(a) Original highway scene

Highway.mp4

Parking scene

(b) Original parking scene

parking.mp4

(c) Simulation output video for architecture B, 5 % DVS threshold, Highway scenario

(d) Simulation output video for architecture B, 5 % DVS threshold, Parking scenario
(a) Highway test case  
(b) Parking test case

Figure: The number of generated events per architecture and DVS threshold
Figure: Comparison of architectures activity per test case for 5% DVS threshold
Conclusion

The simulation results confirm that achieving complete redundancies suppression is possible, while maintaining the relevant information of the scene.
Upcoming work

One of the presented architectures will be implemented in 28 nm FDSOI technology to provide the first proof of concept.
References I


This work has been partially supported by the OCEAN12 project (H2020-EU.2.1.1.7 no78327)
THANK YOU