Temporal Pulses Driven Spiking Neural Network for Time and Power Efficient Object Recognition in Autonomous Driving

Wei Wang 1, Shibo Zhou 2, Jingli Li 2, Xiaohua Li 2, Junsong Yuan 1, Zhanpeng Jin 1
wwang49@buffalo.edu, {szhou19, xli}@binghamton.edu, {wei.wang, zjin}@buffalo.edu

1 University at Buffalo (SUNY), 2 Binghamton University (SUNY)

Abstract

• We proposed an SNN model with temporal coding to directly process LiDAR pulses for object recognition.
• We created a comprehensive temporal pulses dataset, “Sim LiDAR”, which simulated LiDAR reflection of different road conditions and target objects in diverse noise environments.
• Demonstrates novel contribution in time and computational efficiency for real-time deep learning applications.

Motivation & Rationale

Why SNN: [See Figure 1]
A. Spikes have inherent temporal information.
B. Integrate and fire, more biological plausible.
C. Event driven, asynchronism, energy efficient.

Why raw temporal pulses (rather than point clouds):
A. Eliminates the restrictions of frames.
B. Can achieve better time efficiency with less computational overhead.

Methods

Neuronal model: non-leaky integrate and fire (n-LIF) neuron [1]

\[
\frac{d v(t)}{dt} = \sum_{c} w_{c} u_{c}(t - t_{c})
\]

\[u(t) = \begin{cases} 1, & \text{if } t \geq 0 \\ 0, & \text{otherwise} \end{cases}\]

Temporal coding:

\[e_{t}^{i} = \sum_{c} w_{i}^{c} u_{c}(t - t_{c}) \]

Network structure: Spiking CNN

Object recognition

Datasets

Sim LiDAR (32 categories)

Objects:
car, pedestrian, truck
Road conditions:
tunnel, open road
tower/upper bridge
road (walls on one side)
road (lamps on one side)

Figure 3. Sim LiDAR covers different objects and road conditions

KITTl (truncated, 8 categories)

Dynamic vision sensor (DVS) dataset (36 categories)

Figure 5. DVS barrel dataset [2]

Results

DVS Dataset (compare with existing models)

Table 1. Comparison of DVS datasets

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy (%)</th>
<th>T_vom (ms)</th>
<th>R_data (%)</th>
<th>Power Consumption (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCNn</td>
<td>96.62</td>
<td>2.02</td>
<td>76</td>
<td>29.83 (n/3.63 \mu)</td>
</tr>
<tr>
<td>CNN</td>
<td>88.22</td>
<td>2.58</td>
<td>100</td>
<td>0.67 J</td>
</tr>
<tr>
<td>VGG-16</td>
<td>92.72</td>
<td>11.34</td>
<td>100</td>
<td>2.95 J</td>
</tr>
<tr>
<td>ResNet-50</td>
<td>92.84</td>
<td>71.30</td>
<td>100</td>
<td>18.54 J</td>
</tr>
</tbody>
</table>

Figure 4. Truncation of KITTl with examples

Figure 6. Efficiency of proposed model

Conclusion

• The proposed SNN has remarkable accuracy
• Extraordinary time and energy efficiency
• Great potential in resource- and/or time-constrained applications.
• Calls for the combination with neuromorphic hardware.

References