Learned Spike Encoding of the Channel Response for Low-Power Environment Sensing

Eleonora Cicciarella, Riccardo Mazzieri, Jacopo Pegoraro, Michele Rossi

March 15, 2024
Radio Frequency sensing

Transmitter

Target (reflector)

Change in frequency, phase, or amplitude

Receiver

Why radio signals instead of cameras?

• Radio Detection and Ranging (RADAR) principle

Privacy
Edge computing

Complex Deep Learning models are run in the cloud

NEW PARADIGM:
Bring computation here

Accuracy-efficiency trade-off
Spiking Neural Networks

1. Biological neurons communicate via action potentials, or spikes
2. Biological neurons spend most of their time at rest
3. Event-based processing
The Leaky Integrate and Fire (LIF) neuron

\[
\begin{align*}
\frac{dU(t)}{dt} &= -U(t) + I_{in}(t) R \\
\tau &\quad \text{time constant of the circuit} \\
U(t) &\quad \text{membrane potential} \\
I_{in}(t) &\quad \text{input current}
\end{align*}
\]

\[
\beta = e^{-1/\tau}
\]

\[
U[t] = \beta U[t-1] + WX[t] - S[t-1]\theta
\]

\[
S[t] = \begin{cases} 
1, & \text{if } U[t] > \theta \\
0, & \text{otherwise}
\end{cases}
\]
Spike encoding

Consider also negative spikes

GOAL
- Sparsity
- Preserve spectral content

Temporal Contrast encoding
- Keep track of temporal changes in the signal
- Inaccurate and dense encoding

What if we learned the spike encoding?
Signal model

\[ x[k] \triangleq x(kT) = \sum_{m=1}^{M} a_m e^{j(2\pi f_m kT + \phi_m)} + w(kT), \quad k = 0, \ldots, K - 1 \]

- \( T \): sampling time
- \( M \): # of sinusoids
- \( K \): window length
- Each sinusoidal component accounts for one moving reflector
- Dataset: 3,000 windows for each \( M=1, \ldots, 5 \)
Network architecture

\[ \mathcal{L}_1 = \text{MSE}(X, \hat{X}) \]

Encoder

\[ \mathcal{Y} \xrightarrow{\mathcal{H}_\tau(\cdot)} \mathcal{Z} \]

Spike encoding

\[ \lambda (\Omega(Z)) \]

Decoder

\[ \mathcal{L}_2 = \text{MSE}(f, \hat{f}) \]
\[ \mathcal{L}_3 = \text{MSE}(a, \hat{a}) \]

Spiking Neural Network

Total number of spikes

\[ \Omega(Z) \triangleq \frac{1}{2K} \sum_{c=1}^{2} \sum_{k=0}^{K-1} |Z_c[k]| \]

\[ \mathcal{H}_\tau(y) = \begin{cases} \text{sign}(y) & \text{if } |y| \geq \tau \\ 0 & \text{otherwise} \end{cases} \]
Comparison with Temporal Contrast methods

- Threshold-based representation (TBR)
- Step-forward (SF)
- Moving-window (MW)

- channel reconstruction
- spectral components
- sparsity
- robustness to noise

\[
\begin{align*}
   x[k + 1] - x[k] &> \text{thr.} \quad \rightarrow \text{spike } = 1 \\
   x[k + 1] - x[k] &< -\text{thr.} \quad \rightarrow \text{spike } = -1
\end{align*}
\]
Channel reconstruction

**Metric:** Per-window Root Mean Squared Error

<table>
<thead>
<tr>
<th>Method</th>
<th>Recon. RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBR</td>
<td>0.374 ± 0.075</td>
</tr>
<tr>
<td>SF</td>
<td>0.222 ± 0.044</td>
</tr>
<tr>
<td>MW</td>
<td>0.262 ± 0.059</td>
</tr>
<tr>
<td>LSE</td>
<td>0.133 ± 0.028</td>
</tr>
</tbody>
</table>

LSE = Learned Spike Encoding
DFT magnitude reconstruction

**Metric:** Per-window Root Mean Squared Error

| Method | $|DFT|^2$ RMSE |
|--------|---------------|
| TBR    | 0.043 ± 0.010 |
| SF     | 0.029 ± 0.009 |
| MW     | 0.039 ± 0.011 |
| LSE    | 0.017 ± 0.004 |
Sparsity of the encoding

**Sparsity:** Fraction of zeros in the encoding

| Method | $|\text{DFT}|^2$ | RMSE |
|--------|---------------|------|
| TBR    | 0.015         |      |
| SF     | 0.433         |      |
| MW     | 0.202         |      |
| LSE    | 0.736         |      |

70% higher sparsity than SF

Direct control of the sparsity-accuracy trade-off
Robustness to noise

Channel response reconstruction

|DFT|² reconstruction
Concluding remarks

- Learn a **tailored** spike encoding for RF channel responses
- CAE for encoding + SNN for reconstructing amplitudes and frequencies
- **Lightweight** neural network: <120K parameters, ~2MB of size
- Direct control of the performance-sparsity trade-off
Thank you

eleonora.cicciarella@phd.unipd.it