Event-driven Robust Fitting on Neuromorphic Hardware

Tam Ngoc-Bang Nguyen[1], Anh-Dzung Doan[1], Zhipeng Cai[2], Tat-Jun Chin[1]
[1] Australian Institute for Machine Learning, The University of Adelaide,
[2] Intel Labs
{tam.nb.nguyen,dzung.doan,tat-jun.chin}@adelaide.edu.au
czptc2h@gmail.com

A. Derivation for the gradient with z

Recall the gradient descent formula for the quadratic form of least squares:

$$\nabla f(\boldsymbol{\theta}) = \mathbf{Q}\boldsymbol{\theta} + \mathbf{p} \tag{1}$$

We expand

$$\nabla f(\boldsymbol{\theta}) = \mathbf{Q}\boldsymbol{\theta} + \mathbf{p} \tag{2}$$

$$= \mathbf{Q}' \begin{bmatrix} z_1 & 0 \\ 0 & z_1 \\ z_2 & 0 \\ 0 & z_2 \\ \vdots & \vdots \\ z_N & 0 \\ 0 & z_N \end{bmatrix} \boldsymbol{\theta} + \mathbf{P}' \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \tag{3}$$

$$= \mathbf{Q}' \text{vec} \left(\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \boldsymbol{\theta}^T \right) + \mathbf{P}' \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \tag{4}$$

$$= \mathbf{Q}' \text{vec} \left(\mathbf{z}\boldsymbol{\theta}^T \right) + \mathbf{P}'\mathbf{z} \tag{5}$$

$$= \mathbf{Q}' (\mathbf{z} \otimes \boldsymbol{\theta}^T) + \mathbf{P}'\mathbf{z} \tag{6}$$

where,

$$\mathbf{Q}' = \begin{bmatrix} \mathbf{x}_1\mathbf{x}_1^T & \mathbf{x}_2\mathbf{x}_2^T & \dots & \mathbf{x}_N\mathbf{x}_N^T \end{bmatrix} \in \mathbb{R}^{d \times Nd} \tag{7}$$

$$\mathbf{P}' = \begin{bmatrix} -y_1\mathbf{x}_1 & -y_2\mathbf{x}_2 & \dots & -y_N\mathbf{x}_N \end{bmatrix} \in \mathbb{R}^{d \times N} \tag{8}$$

Let $\boldsymbol{\theta}' = \text{vec}\left(\mathbf{z}\boldsymbol{\theta}^T\right)$, Eq. (6) becomes

$$\nabla f(\boldsymbol{\theta}, \mathbf{z}) = \mathbf{Q}'\boldsymbol{\theta}' + \mathbf{P}'\mathbf{z} \tag{9}$$

B. Construct NeuroRF synaptic matrices for affine transformation

Problem definition Given a point correspondence $\langle \mathbf{x}_i', \mathbf{x}_i \rangle$, the affine transformation is defined by

$$\mathbf{x}' = \mathbf{H}_\text{A}\widetilde{\mathbf{x}} \tag{10}$$

where $\mathbf{x}' = (x', y')$ and $\widetilde{\mathbf{x}} = [\mathbf{x}^T \quad 1] = (x, y, 1)$ is $\mathbf{x}$ in homogeneous coordinates. The 2×3 affine transformation matrix $\mathbf{H}_\text{A}$ can be estimated with 3 point correspondences [1, Chap 2], which constitute a 6-equation linear system to solve 6 unknown parameters, i.e. $d = 6$

$$\mathbf{H}_\text{A} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \end{bmatrix} \tag{11}$$

Given $\mathbf{H}_\text{A}$ and Eq. (10), the residual at the point i can be defined as the "transformation error"

$$r_i\left(\mathbf{H}_\text{A}\right) = |\mathbf{x}_i' - \mathbf{H}_\text{A}\widetilde{\mathbf{x}}_i| \tag{12}$$

RANSAC is applied to find $\mathbf{H}_\text{A}$ that maximises the consensus

$$\Psi(\mathbf{H}_\text{A}) = \sum_{i=1}^{N} \mathbb{I}\left(|\mathbf{x}_i' - \mathbf{H}_\text{A}\widetilde{\mathbf{x}}_i| \leq \epsilon\right), \tag{13}$$

To apply our NeuroRF method for affine transformation, we first vectorize the affinity matrix

$$\boldsymbol{\theta} = \text{vec}(\mathbf{H}_\text{A}) \in \mathbb{R}^d \tag{14}$$

where $d = 6$.

We next compute $\mathbf{Q}'$ and $\mathbf{P}'$ from $\mathbf{X}' \in \mathbb{R}^{N \times 2}$ and $\widetilde{\mathbf{X}} \in \mathbb{R}^{N \times 3}$ to apply the gradient descent of the ModelHypothesis (Eq. (9))

$$\mathbf{Q}' = \begin{bmatrix} \widetilde{\mathbf{x}}_1\widetilde{\mathbf{x}}_1^T & \mathbf{0}_{3\times3} & \dots & \widetilde{\mathbf{x}}_N\widetilde{\mathbf{x}}_N^T & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \widetilde{\mathbf{x}}_1\widetilde{\mathbf{x}}_1^T & \dots & \mathbf{0}_{3\times3} & \widetilde{\mathbf{x}}_N\widetilde{\mathbf{x}}_N^T \end{bmatrix} \in \mathbb{R}^{d \times Nd} \tag{15}$$

$$\boldsymbol{\theta}' = \text{vec}\left(\mathbf{z}\boldsymbol{\theta}^T\right) = \mathbf{z} \otimes \boldsymbol{\theta} \in \mathbb{R}^{Nd} \tag{16}$$

$$\mathbf{P}' = \begin{bmatrix} -x_1'\widetilde{\mathbf{x}}_1 & -x_2'\widetilde{\mathbf{x}}_2 & \dots & -x_N'\widetilde{\mathbf{x}}_N \\ -y_1'\widetilde{\mathbf{x}}_1 & -y_2'\widetilde{\mathbf{x}}_2 & \dots & -y_N'\widetilde{\mathbf{x}}_N \end{bmatrix} \in \mathbb{R}^{d \times N} \tag{17}$$

Note that in Eq. (17), x_i' and y_i' are broadcast over $\widetilde{\mathbf{x}}_i$.

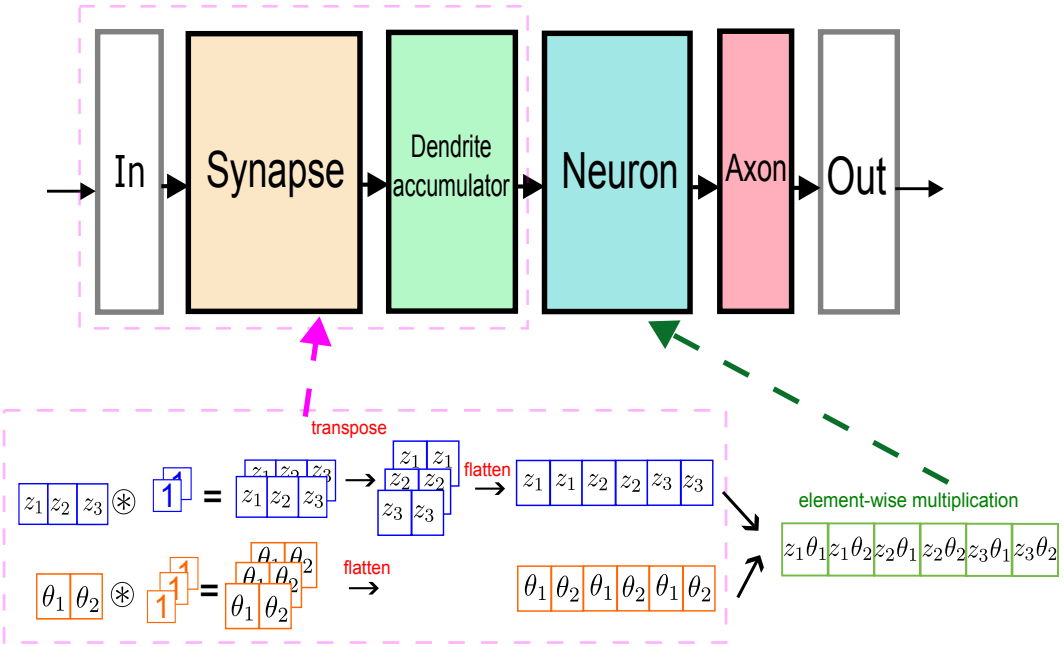


Figure AA. The illustration for using multiple 1×1 convolution filters to simulate the Kronecker product with 3 Sampling neurons $\{z_i\}_{i=1}^{3}$ and 2 GD neurons $\{\theta_j\}_{j=1}^{2}$. The 1×1 convolutions are defined in Synapse, while the $\mathrm{transpose}(\cdot)$ and $\mathrm{flatten}(\cdot)$ are inherently supported by Lava-Loihi framework. The element-wise multiplication is computed in Auxiliary neuron.

C. Runtime results for NeuroRF-CPU vs RS-CPU

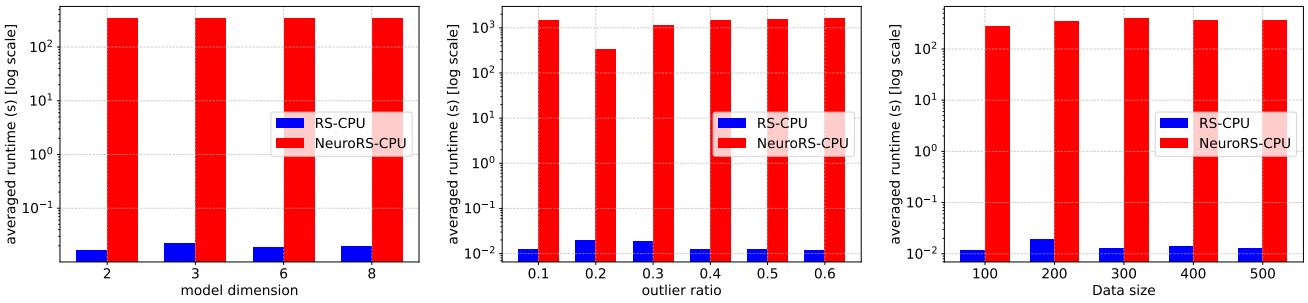
Fig. BB reported the average runtime between NeuroRF-CPU and RS-CPU of Sec 5.2. in the Main paper. The runtime of NeuroRF-CPU was from a CPU simulation of our NeuroRF, hence not reflective of NeuroRF-Loihi.

D. Performance on neuromorphic hardware on N = 10

We include additional results for the problem size of $N = 10, d = 2$ for Sec. 5.3 in the Main paper. Fig CCa demonstrates the equivalent solution quality between NeuroRF-Loihi and RS-CPU. Fig CCb shows that NeuroRF-Loihi consumes approximately 13% dynamic energy compared to its competitor, which again confirms the superiority of energy efficiency of our NeuroRF-Loihi. Fig CCc indicates that the runtime of NeuroRF-Loihi was higher than RS-CPU , which was probably because NeuroRF-Loihi solved least squares with GD, while RS-CPU computed least squares with analytical solutions.

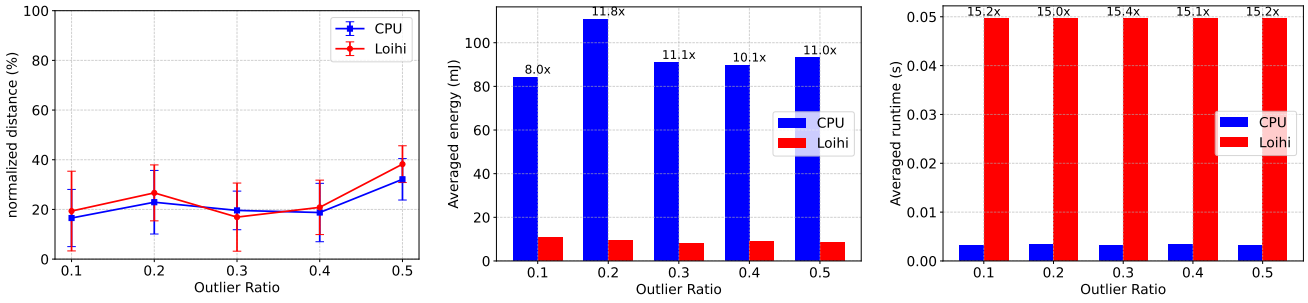
References

[1] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 1



(a) Vary d with $(N = 200, \eta = 0.2)$ (b) Effect of outlier ratio η $(N = 200, d = 8)$ (c) Vary N with $(\eta = 0.2, d = 8)$

Figure BB. Averaged runtime across various levels of difficulty. Results were averaged over 10 trials for each method.



(a) (b) (c)

Figure CC. Performance on neuromorphic hardware. (a) Average consensus size, (b) average dynamic energy consumption and (c) average runtime of NeuroRF-Loihi and RS-CPU on synthetic line fitting instances with $N = 10$ points, plotted against outlier rate.