Event-Driven Embodied Perception for Hazardous Slip Detection under Perceptual Degradation

Sylvester Kaczmarek[1]

***Abstract*— Embodied robots operating on planetary surfaces must preserve mobility when visual perception is degraded by low-angle illumination, deep shadowing, dust, and weak terrain texture. Under these conditions, vision-dominant traversability assessment can become unreliable, while continuous dense fusion is costly for power-limited rovers. We present an event-driven multi-modal perception framework for hazardous wheel slip detection using Spiking Neural Networks (SNNs). Exteroceptive visual observations and proprioceptive IMU and motor-current streams are encoded with Bayesian Spike Pattern Superposition (BSPS), producing an asynchronous representation that remains stable under visual noise, dropout, and temporal jitter. The detector is evaluated in a high-fidelity lunar surface simulation with Bekker-Wong terramechanics, controlled visual degradation, and simulator-derived wheel-terrain ground truth. It detects hazardous slip events with an F1-score of 0.86 ± 0.04. Deployed on the BrainChip Akida neuromorphic processor, the model achieves 1.2 ms latency and $48\,\mu$J per inference, showing that event-driven embodied perception can retain slip awareness under degraded vision while remaining compatible with planetary SWaP constraints.**

I. INTRODUCTION

Planetary rovers must maintain mobility under conditions that are hostile to vision-based perception. Lunar polar regions, cratered terrain, and dust-contaminated scenes can contain low-angle illumination, deep shadows, weak texture, and perceptual aliasing. These factors can degrade visual odometry and vision-based traversability assessment, which are widely used in planetary mobility pipelines [3], [4], [8], [9]. Slip is especially important because excessive wheel slip reduces traverse efficiency, corrupts dead-reckoning, and can lead to immobilization on deformable regolith.

Prior rover slip-estimation work has used visual odometry, wheel-terrain imaging, inertial estimation, and proprioceptive constraints [3]–[5]. These methods establish that slip can be inferred from both exteroceptive and proprioceptive evidence, but they also expose a deployment tension: visual methods are sensitive to degraded scenes, while continuously processing dense multi-modal streams with conventional models can be expensive under strict Size, Weight, and Power (SWaP) limits. This paper addresses that tension using event-driven embodied perception. The key contribution is a low-power event-driven visual-proprioceptive slip detector that remains responsive when visual terrain evidence is degraded. The contribution is instantiated as a BSPS-encoded SNN fusion pipeline that combines visual, IMU, and motor-current evidence for hazardous slip detection, with deployment measurements on neuromorphic hardware.

[1]Department of Computing, Imperial College London, London, United Kingdom. research@sylvesterkaczmarek.com

II. METHOD

We model mobility loss through hazardous wheel slip during traversal of deformable regolith [1], [2]. Slip s is defined as the kinematic mismatch between actual rover velocity and commanded wheel velocity,

$$s = 1 - \frac{v_a}{r\omega_c}, \tag{1}$$

where v_a is actual forward velocity, r is wheel radius, and ω_c is commanded angular velocity. Under nominal perception, v_a can be estimated from visual odometry. Under glare, shadowing, dust occlusion, or dropout, the visual estimate can become noisy or unavailable. In that case, hazardous slip must be inferred from the joint behavior of proprioceptive channels, especially IMU transients and motor-current excursions that indicate loss of traction.

Continuous sensor streams are encoded into spikes using Bayesian Spike Pattern Superposition. For a sensor channel, BSPS represents recent activity as a probability vector $\mathbf{p}(t) \in \mathbb{R}^N$ over candidate spike patterns in a sliding temporal window. Given an observation $o(t)$ and candidate pattern s_i, the softened Bayesian update is

$$p_i(t+1) = (1-\eta)p_i(t) + \eta \frac{P(o(t) \mid s_i)p_i(t)}{\sum_j P(o(t) \mid s_j)p_j(t)}, \tag{2}$$

where η controls temporal smoothing. In implementation, the likelihood includes non-zero tolerance for mismatch, reducing sensitivity to jitter, dropout, and short bursts of spurious events. Under degraded visual input, the fused spiking state avoids representational collapse and remains anchored by the more consistent proprioceptive cues from IMU and motor-current channels. Recurrent spiking layers process the fused event stream and output a slip-risk score $A(t)$. A single hazard threshold θ is selected on validation sequences and then held fixed across test traverses and visual degradation regimes.

III. EXPERIMENTAL SETUP

The framework was evaluated in a lunar surface simulation environment integrating Gazebo rigid-body dynamics with Cesium for Unreal rendering [10], [11]. The terrain model used Lunar Reconnaissance Orbiter digital elevation data from the Copernicus Crater region. A 6-DOF, 430 kg rover model was simulated over representative lunar surface traverses. Hazardous slip events were induced through controlled reductions in the effective regolith friction coefficient under Bekker-Wong terramechanics, while visual degradation was injected into the camera stream through simulated solar glare, dust occlusion, sensor noise, and dropout.

Ground-truth slip labels were obtained from simulator state and wheel-terrain interaction variables rather than from image corruption. The evaluation spans more than 100 simulated traverses and more than 500 injected anomalies in the rover scenario library. The operating threshold was validation-selected and held fixed during test evaluation; for hazardous wheel-slip declaration, the deployed rule used $A(t) > 0.8$ together with proprioceptive evidence concentrated in motor-current and wheel-encoder channels. The evaluation separates the mobility event from the perceptual degradation: degraded non-slip intervals are negative cases, and a detector that simply fires on glare or dropout is penalized. The same slip events are evaluated across nominal, degraded, and severe visual conditions to test whether the model detects kinematic slip rather than image degradation alone. The reported means and standard deviations are computed over repeated simulated traverses with fixed test conditions. For deployment evaluation, the BSPS-encoded sensor streams were executed on the BrainChip Akida AKD1000 for real-time asynchronous inference. The reported $48\,\mu$J value refers to Akida-side inference energy for the BSPS-encoded SNN, measured at deployment level; it does not include lunar rendering, simulator dynamics, or offline data generation. LSTM and Isolation Forest baselines were executed as host-CPU implementations on an Intel i7-9700K workstation in the same comparison pipeline; the energy and latency values in Table I therefore compare deployed implementations, rather than device-normalized theoretical model costs.

IV. RESULTS

The neuromorphic fusion model maintained hazardous slip detection under visual degradation. As visual noise increased and vision-only estimation deteriorated, the SNN preserved the kinematic signatures of slip through correlated proprioceptive activity in the IMU and motor-current channels. Figure 1 shows a representative event in which visual velocity becomes unreliable during a slip interval, while the BSPS-encoded proprioceptive channels remain correlated with the underlying mobility event and the risk score crosses the fixed threshold θ.

The proposed system detected hazardous slip events with an F1-score of 0.86 ± 0.04 across the evaluated degradation sweep. The non-degraded slip condition provides the sanity baseline for slip detection without image corruption, while degraded non-slip intervals verify that the detector is not merely responding to glare, dropout, or sensor noise. A corruption-only detector would fire on degraded non-slip intervals; the proposed detector instead requires visual disruption to coincide with proprioceptive evidence of traction loss. Under severe degradation, the model relies primarily on consistent IMU, wheel-encoder, and motor-current evidence, which explains the retained detection performance despite compromised visual input.

Table I reports the deployed implementation comparison. In the evaluated implementations, the SNN achieves higher F1 together with lower latency and energy than the CPU baselines. The 1.2 ms latency supports fast mobility-risk assessment in closed-loop rover operations, while the $48\,\mu$J inference cost is consistent with low-power neuromorphic deployment. The simulation pipeline also functions as a compact verification-and-validation harness: visual hazards, wheel-terrain interaction parameters, and ground-truth slip labels can be controlled independently before hardware deployment.

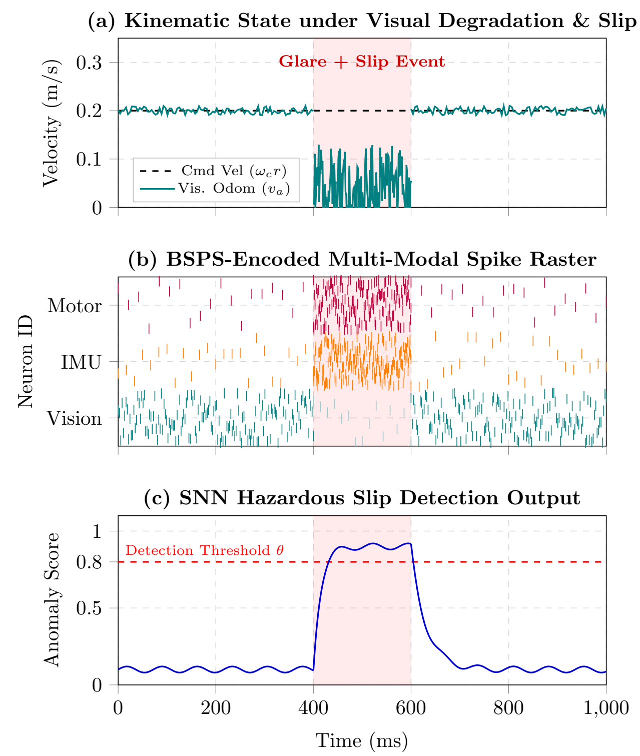

Fig. 1. Representative hazardous slip event under degraded vision. Top: commanded and visually estimated velocity during loss of traction. Middle: BSPS-encoded visual and proprioceptive spiking activity. Bottom: slip-risk score $A(t)$ crossing the fixed hazard threshold θ.

TABLE I
DEPLOYED IMPLEMENTATION COMPARISON FOR HAZARDOUS SLIP DETECTION. SNN RUNS ON AKIDA; LSTM AND ISOLATION FOREST ARE HOST-CPU BASELINES.

Model	F1-score	Latency (ms)	Energy (μJ)
Proposed SNN (Akida)	$\mathbf{0.86 \pm 0.04}$	$\mathbf{1.2 \pm 0.3}$	$\mathbf{48 \pm 9}$
LSTM (CPU)	0.80 ± 0.05	35.2 ± 5.1	8500 ± 1300
Isolation Forest (CPU)	0.68 ± 0.04	12.8 ± 1.8	2500 ± 400

V. CONCLUSION

This paper presented an event-driven embodied perception framework for hazardous slip detection under perceptual degradation. By combining BSPS spike encoding, visual-proprioceptive fusion, and neuromorphic deployment, the method retained slip detection under degraded vision while operating at 1.2 ms and $48\,\mu$J per inference. The results indicate that SNN-based multi-modal fusion is a practical candidate for mobility awareness in power-constrained planetary rovers and motivate transfer to physical regolith testbeds.

REFERENCES

[1] M. G. Bekker, *Introduction to Terrain-Vehicle Systems*. Ann Arbor, MI, USA: University of Michigan Press, 1969.

[2] J. Y. Wong, *Theory of Ground Vehicles*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, 2008.

[3] D. M. Helmick, Y. Cheng, D. S. Clouse, L. H. Matthies, and S. I. Roumeliotis, "Path following using visual odometry for a Mars rover in high-slip environments," in *Proc. IEEE Aerospace Conf.*, 2004.

[4] A. Angelova, L. Matthies, D. Helmick, and P. Perona, "Slip prediction using visual information," in *Proc. Robotics: Science and Systems*, 2006.

[5] C. Kilic, Y. Gu, and J. N. Gross, "Proprioceptive slip detection for planetary rovers in perceptually degraded extraterrestrial environments," *arXiv:2207.13629*, 2022.

[6] W. Maass, "Networks of spiking neurons: The third generation of neural network models," *Neural Networks*, vol. 10, no. 9, pp. 1659–1671, 1997.

[7] M. Davies *et al.*, "Advancing neuromorphic computing with Loihi: A survey of results and outlook," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 911–934, 2021.

[8] NASA, "Lunar Terrain Vehicle." [Online]. Available: `https://www.nasa.gov/suits-and-rovers/lunar-terrain-vehicle/`

[9] NASA Jet Propulsion Laboratory, "CADRE." [Online]. Available: `https://www.jpl.nasa.gov/missions/cadre/`

[10] N. Koenig and A. Howard, "Design and use paradigms for Gazebo, an open-source multi-robot simulator," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2004, pp. 2149–2154.

[11] P. Cozzi and K. Ring, *3D Engine Design for Virtual Globes*. Boca Raton, FL, USA: CRC Press, 2011.