From Passive Perception to Utility-Aware Sensing: Lessons for Navigation in the Wild

Yash Turkar and Karthik Dantu

***Abstract*—Visual navigation is only as reliable as the observations it receives. In the wild, darkness, saturation, reflections, and motion blur can corrupt visual measurements before they reach downstream estimation. This paper distills lessons from three studies that treat sensing as a controllable part of the navigation loop. First, pixel-wise measurement reliability reduces visual-odometry trajectory error by 31% while using 49% fewer features. Second, image feature utility guides joint exposure and active-illumination control, improving feature detection and matching by 47–197% in low-light field experiments. Third, reasoning over counterfactual observations at candidate light levels supports trajectory-level illumination control and improves performance over fixed-light baselines across multiple runs. Together, these results show that navigation objectives can guide which measurements are used, how they are acquired, and how sensing actions are selected over time.**

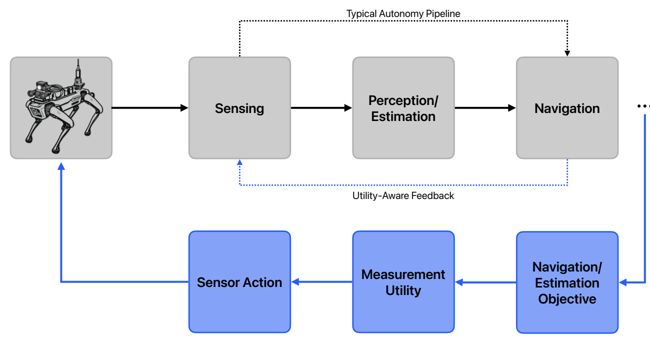

Fig. 1. Conventional autonomy pipelines propagate information from sensing through perception and estimation to navigation. We consider a complementary utility-aware feedback path in which the navigation or estimation objective defines the value of visual measurements, measurement utility informs sensing decisions, and the resulting sensor action changes the observations available to the autonomy stack.

I. Introduction

Visual navigation depends on visual measurements to support perception, state estimation, and decision-making; when those measurements degrade, failures can propagate throughout the autonomy stack. Feature-based and direct visual-odometry pipelines have made remarkable progress under controlled conditions [1], [2], and modern learned features improve repeatability and matching across a wide range of scenes [3]–[5]. Yet these methods share a fundamental dependency: they can only reason over the information that reaches the camera. In the field, a robot may fail not because its estimator cannot interpret a good observation, but because the camera never acquired one. Darkness and uneven illumination can remove scene structure, while overexposure, specular reflections, and motion blur can corrupt otherwise useful features. Such failures are especially consequential in visually difficult and subterranean environments, where appearance can change sharply along a trajectory [6], [7].

These are acquisition failures, not simply inference errors. A downstream algorithm cannot fully recover feature information that was never captured, nor is a brighter image necessarily a more useful one. Exposure-control methods adapt shutter time, gain, or other camera settings in response to image appearance [8]–[11], but photometric quality alone does not guarantee that an observation supports stable correspondence and state estimation. Longer exposures can also introduce motion sensitivity, while high gain can amplify noise. An onboard light can reveal structure that camera settings alone cannot recover, but it changes the measurement process as well: excessive illumination can produce saturation and specular artifacts, alter feature appearance across frames, and consume limited robot power [7], [12], [13].

The resulting question how navigation objectives should inform the acquisition of observations. This framing follows active perception, in which a robot acts to acquire better measurements rather than only inferring from passive ones [14]–[17]. When sensing parameters are controllable, the value of a measurement for downstream estimation can flow upstream: it can determine which current measurements are retained and how later observations should be acquired.

The feedback path in Figure 1 can be realized at several points in the navigation loop. We consider utility-aware sensing first as measurement selection, where an estimate of reliability determines which current visual features are passed to estimation; then as sensing control, where measurement utility informs the settings used to acquire the best observation; and finally as predictive sensing control, where counterfactual observations support sensing decisions over a trajectory.

II. Utility-Aware Perception

We consider utility-aware perception at three points in the navigation loop: selecting the visual measurements used by estimation, adapting the settings that acquire the next measurement, and anticipating how sensing actions will affect future measurements. In each case, the sensing decision is tied to its expected effect on downstream estimation.

Yash Turkar and Karthik Dantu are with the Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.

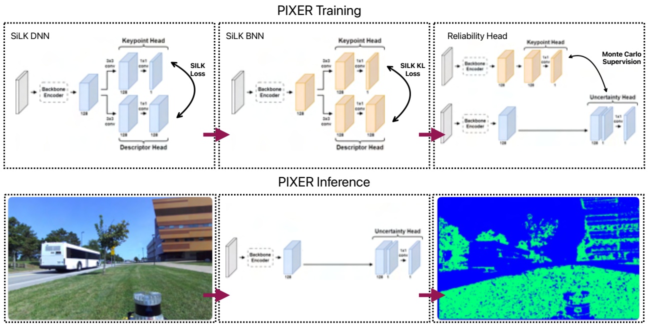

Fig. 2. PIXER learns a detector-agnostic reliability mask in three stages. A deterministic SiLK network is first trained to model feature utility; its Bayesian counterpart provides Monte Carlo variance targets; a lightweight reliability head then predicts pixel-wise uncertainty in a single forward pass. At inference, this prediction suppresses low-utility image regions before feature matching.

A. *Measurement Selection*

At the first stage, a system can improve navigation by deciding which of its current measurements deserve to be used. PIXER [18] estimates dense pixel-wise reliability from an image and uses the prediction to suppress features that are unlikely to support reliable matching and visual odometry. The key contribution is not a new feature detector, but a utilty umask that can be used with multiple feature extractors.

PIXER learns this reliability estimate offline in three stages. A deterministic SiLK backbone [4] is first trained to produce dense feature representations. A Bayesian version of that network then provides variance-based supervision through multiple stochastic forward passes [19]. Finally, a lightweight reliability head is trained to predict that uncertainty directly from the deterministic backbone's feature maps. At inference, this produces a dense uncertainty map in one forward pass; thresholding the map yields a mask that removes low-utility regions before a feature detector and matcher pass their measurements to visual odometry. This separation between reliability prediction and the downstream detector allows the same mask to be used with multiple feature pipelines.

Across three datasets and eight feature algorithms, reliability-based filtering reduces visual-odometry trajectory error by an average of 31% while using 49% fewer features. Figure 3 makes the selection effect visible at the trajectory level. In both examples, filtering produces an estimate that follows the reference trajectory more closely. The filtered Shi–Tomasi result also improves on the unfiltered SIFT baseline, illustrating that a lightweight detector combined with reliability filtering can outperform a more expensive unfiltered feature pipeline. These results establish the first lesson: measurement utility is not uniformly distributed across an image, and selecting fewer but more reliable measurements can improve both estimation quality and efficiency.

B. *Reactive Sensing Control*

Once utility can be estimated, it can also guide how the next measurement is acquired. NightHawk [20] uses feature-based

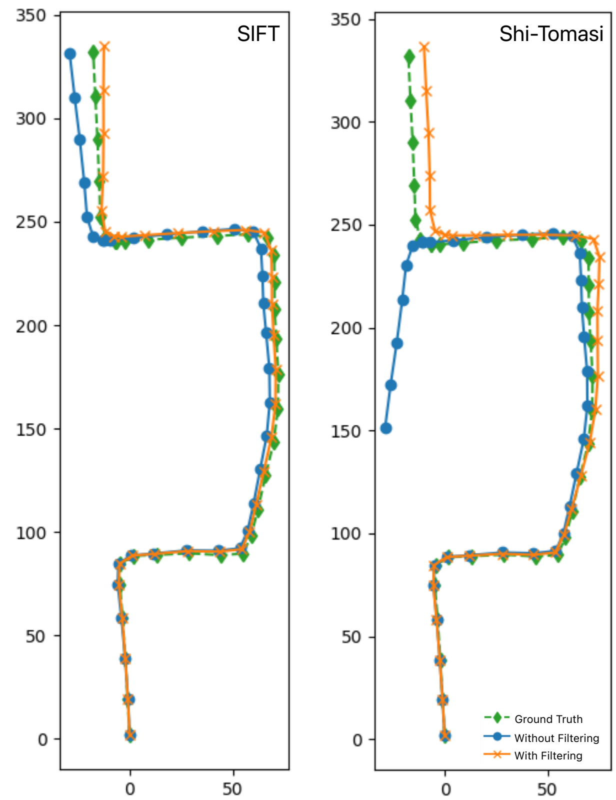

Fig. 3. Visual-odometry trajectories on KITTI using SIFT (left) and Shi–Tomasi (right) features. Reliability-aware filtering brings the estimated trajectories closer to ground truth in both cases; the filtered Shi–Tomasi result improves on the unfiltered SIFT baseline.

utility to jointly choose camera exposure-time and the intensity of an onboard co-located light in low-light robot navigation. This is a direct intervention in measurement acquisition.

This method treats exposure-time and illumination power as a joint sensing action. For the current scene, it evaluates candidate camera–light configurations with a feature-based utility metric and uses Bayesian optimization [21] to search for a configuration that improves this metric. The selected settings are executed by the vision system, and the utility of the resulting image is monitored as the robot moves. When the score falls sufficiently below the value associated with the current configuration, the optimizer is triggered again; otherwise, the previous setting is retained. This event-triggered loop makes the controller responsive to meaningful scene changes without continually re-running optimization.

In subterranean culvert experiments [22], relying on auto-exposure without external light produced an average feature-track length of 2.33. A fixed maximum light increased this value to 4.70, but introduced reflection-related artifacts. The utility-aware controller achieved an average track length of 6.94 while reducing exposure time to 7.77 ms. The feature tracks in Figure 6 show the tradeoff directly: neither no external light nor a fixed maximum light provides the performance

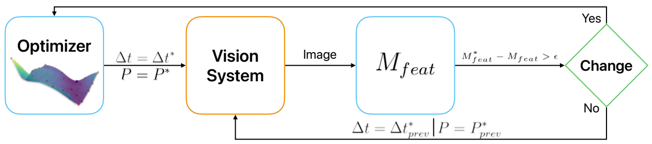

Fig. 4. NightHawk's [20] event-triggered sensing-control loop. Bayesian optimization selects illumination power and exposure time to maximize the feature-based utility metric. The selected configuration is retained until a sufficient decline in image utility triggers re-optimization.

TABLE I

ONLINE ACTIVE-ILLUMINATION PERFORMANCE ACROSS THREE AUTONOMOUS SEQUENCES. C IS TRAJECTORY COMPLETION RATIO AND WRMSE IS WEIGHTED TRAJECTORY ERROR; THE FINAL COLUMNS REPORT THE LEARNED POLICY'S MEAN ILLUMINATION AND POWER USE.

	Light Off		Lightning		*Light On*		Policy	
Sequence	C	WRMSE	C	WRMSE	C	WRMSE	μ (%)	W
`113dark`	0.26	2.756	**0.89**	1.405	0.44	**0.641**	48.28	19.86
`kitchenloop`	0.48	0.744	**0.91**	**0.393**	0.88	0.425	60.42	22.45
`start2corr`	0.95	0.220	**0.99**	**0.123**	0.98	0.249	47.33	19.66

achieved by jointly adapting illumination and exposure. These results show that better sensing is not equivalent to more light or longer exposure; the useful setting depends on the scene and on the visual-estimation task.

C. Predictive Sensing Control

Reactive adaptation is useful, but sensing decisions also accumulate along a trajectory. Lightning [23] extends utility-aware sensing from a frame-wise decision to illumination control over a trajectory.

In this work, we first addresses the counterfactual-data problem: a robot normally observes each point on a trajectory under only the illumination it actually chose. The co-located illumination decomposition model separates the ambient appearance of a captured scene from the contribution of the robot's light, allowing the observation to be synthesized at candidate light intensities. These relit observations form an expanded sequence on which an offline oracle computes an optimal intensity schedule. The schedule balances SLAM-relevant utility and inter-frame matchability against illumination power and changes in intensity over time. Because this offline solution depends on future frames, it cannot be used directly on the robot. A policy is therefore trained by imitation learning to predict the next discrete light intensity from the current image and previous commanded level. Figure 5 summarizes the resulting three-stage pipeline.

Across the reported robot sequences, the oracle achieved the lowest weighted trajectory error on every sequence, while the deployed policy improved over baselines on all three autonomous runs. Table I shows that the online policy attains the highest completion ratio on every run. In the challenging `113dark` run, it reaches a completion ratio of 0.89, compared with 0.26 with the light off and 0.44 with the light fixed on. Constant illumination can yield a lower trajectory error on part of a run, but it does not provide the same robustness to changing visual conditions.

Figure 7 illustrates the resulting behavior over time. The policy brightens the scene when image utility would otherwise fall, dims the light when reflection becomes harmful, and uses intermediate illumination where neither extreme is appropriate. This result motivates prediction as a way to choose sensing actions before a poor observation is acquired.

III. DISCUSSION AND LESSONS LEARNED

A. Optimize Task Utility, Not Image Appearance

The three projects begin with the same premise: a visually pleasing image is not necessarily the measurement that best supports the downstream application. PIXER makes this distinction explicit at the feature level, where the useful quantity is whether a measurement remains reliable for matching and pose estimation. NightHawk uses a feature-based utility objective rather than brightness alone when selecting camera and lighting settings. Lightning takes the same idea into a sequence-level decision, selecting illumination according to its expected effect on SLAM.

B. More Is Not Always Better

We show that more aggressive sensing is not necessarily better. In NightHawk, additional illumination can reveal structure but can also create saturation and specular artifacts; longer exposure can brighten an image but may increase motion sensitivity. Lightning likewise finds that the useful illumination level depends on the scene: a controller may brighten dark regions, reduce light near reflective surfaces, or select an intermediate level. Sensing actions should therefore be evaluated by their benefit to estimation as well as their costs in power, latency, blur, and degraded measurements.

C. Control Becomes Sequential When the Robot Moves

The appropriate sensing decision changes as geometry, reflectance, and ambient illumination change along a trajectory. Reactive utility estimates are therefore valuable, but they are not sufficient for every setting. Lightning shows the value of modeling how candidate light actions would alter future observations, then using those counterfactual observations to select a schedule and train a policy for deployment. This progression connects measurement quality to motion: a useful navigation system must reason not only about the observation available now, but also about the observations it can acquire after moving and sensing differently.

D. Open Directions

Two needs emerge from these studies. First, common benchmarks should compare measurement selection, reactive sensing control, and predictive sensing under controlled changes in illumination, reflectance, motion, and sensing action, while evaluating their effect on downstream tracking, localization,

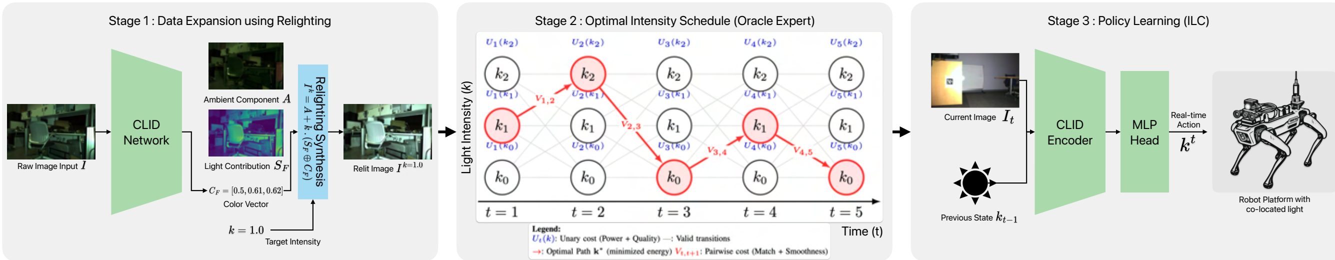

Fig. 5. Lightning's three-stage predictive sensing pipeline. First, CLID expands a recorded sequence with counterfactual observations at candidate illumination levels. Second, an optimal intensity schedule is computed offline from these alternatives. Third, imitation learning distills that schedule into a controller that selects illumination online.

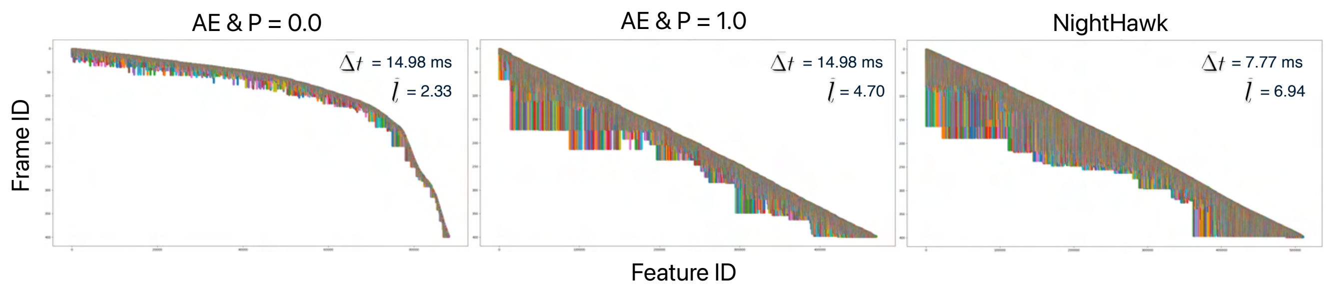

Fig. 6. Feature persistence under auto-exposure without external illumination, auto-exposure with fixed maximum illumination, and NightHawk. NightHawk obtains a mean track length of 6.94 at a 7.77 ms exposure, compared with 2.33 and 4.70 at 14.98 ms for the two baselines.

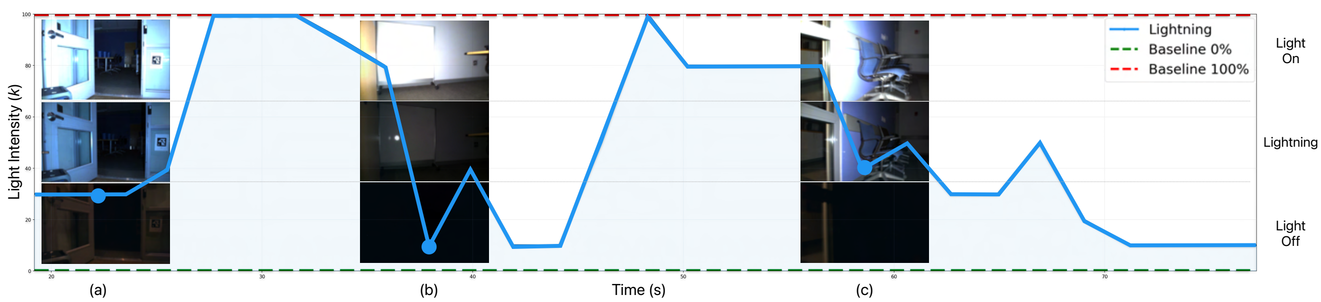

Fig. 7. Learned illumination schedule during an autonomous rollout. The policy increases illumination in low-light portions of the route, reduces it near a reflective whiteboard to limit saturation, and selects intermediate values where dark regions and reflections compete. Fixed light-off and light-on baselines are shown for reference.

and task completion. Second, useful sensing decisions will ultimately need to be coupled with robot motion. This requires predictive models that remain reliable as viewpoint and scene geometry change, and planners that can trade immediate estimation quality against information available from future observations.

IV. Conclusion

Utility-aware sensing reframes robust navigation as more than the processing of whatever measurements happen to arrive. Across three works the central pattern is to estimate the value of visual measurements for a downstream task and use that estimate to select, improve, or anticipate observations. The results show that reliable navigation can benefit from rejecting low-utility features, adapting camera and illumination settings, and reasoning over future sensing actions.

The broader lesson is that sensing belongs inside the navigation loop. When a robot has authority over how it observes the world, it can use that authority to acquire measurements that better support estimation and control. Developing the predictive models, joint controllers, and benchmarks needed to realize this capability is an important direction for navigation in the wild.

References

[1] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An Accurate Open-Source Library for

Visual, Visual–Inertial, and Multimap SLAM," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890, Dec. 2021, conference Name: IEEE Transactions on Robotics.

[2] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," Oct. 2016, arXiv:1607.02565 [cs]. [Online]. Available: http://arxiv.org/abs/1607.02565

[3] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-Supervised Interest Point Detection and Description," Apr. 2018, arXiv:1712.07629 [cs]. [Online]. Available: http://arxiv.org/abs/1712.07629

[4] P. Gleize, W. Wang, and M. Feiszli, "SiLK – Simple Learned Keypoints," Apr. 2023, arXiv:2304.06194 [cs]. [Online]. Available: http://arxiv.org/abs/2304.06194

[5] J. Revaud, P. Weinzaepfel, C. D. Souza, N. Pion, G. Csurka, Y. Cabon, and M. Humenberger, "R2D2: Repeatable and Reliable Detector and Descriptor," Jun. 2019, arXiv:1906.06195 [cs]. [Online]. Available: http://arxiv.org/abs/1906.06195

[6] K. Ebadi, L. Bernreiter, H. Biggie, G. Catt, Y. Chang, A. Chatterjee, C. E. Denniston, S.-P. Deschênes, K. Harlow, S. Khattak, L. Nogueira, M. Palieri, P. Petráček, M. Petrlík, A. Reinke, V. Krátký, S. Zhao, A.-a. Agha-mohammadi, K. Alexis, C. Heckman, K. Khosoussi, N. Kottege, B. Morrell, M. Hutter, F. Pauling, F. Pomerleau, M. Saska, S. Scherer, R. Siegwart, J. L. Williams, and L. Carlone, "Present and Future of SLAM in Extreme Environments: The DARPA SubT Challenge," *IEEE Transactions on Robotics*, vol. 40, pp. 936–959, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10286080/

[7] F. Crocetti, R. Brilli, A. Dionigi, M. L. Fravolini, G. Costante, and P. Valigi, "Comparison of DSO and ORB-SLAM3 in Low-Light Environments With Auxiliary Lighting and Deep Learning Based Image Enhancing," *Journal of Field Robotics*, p. rob.22595, May 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/rob.22595

[8] I. Shim, J.-Y. Lee, and I. S. Kweon, "Auto-adjusting camera exposure for outdoor robotics using gradient information," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Chicago, IL, USA: IEEE, Sep. 2014, pp. 1011–1017. [Online]. Available: http://ieeexplore.ieee.org/document/6942682/

[9] J. Kim, Y. Cho, and A. Kim, "Exposure Control Using Bayesian Optimization Based on Entropy Weighted Image Gradient," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane, QLD: IEEE, May 2018, pp. 857–864. [Online]. Available: https://ieeexplore.ieee.org/document/8462881/

[10] R. Gomez-Ojeda, Z. Zhang, J. Gonzalez-Jimenez, and D. Scaramuzza, "Learning-Based Image Enhancement for Visual Odometry in Challenging HDR Environments," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane, QLD: IEEE, May 2018, pp. 805–811. [Online]. Available: https://ieeexplore.ieee.org/document/8462876/

[11] B. Han, Y. Lin, Y. Dong, H. Wang, T. Zhang, and C. Liang, "Camera Attributes Control for Visual Odometry With Motion Blur Awareness," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 4, pp. 2225–2235, Aug. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10040760/

[12] Z. Zhang, C. Forster, and D. Scaramuzza, "Active exposure control for robust visual odometry in HDR environments," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. Singapore, Singapore: IEEE, May 2017, pp. 3894–3901. [Online]. Available: http://ieeexplore.ieee.org/document/7989449/

[13] J. Kim, Y. Cho, and A. Kim, "Proactive Camera Attribute Control Using Bayesian Optimization for Illumination-Resilient Visual Navigation," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1256–1271, Aug. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9098963/

[14] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, no. 8, pp. 966–1005, Aug. 1988. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5968

[15] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 333–356, Jan. 1988. [Online]. Available: https://doi.org/10.1007/BF00133571

[16] J. Bohg, K. Hausman, B. Sankaran, O. Brock, D. Kragic, S. Schaal, and G. S. Sukhatme, "Interactive Perception: Leveraging Action in Perception and Perception in Action," *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1273–1291, Dec. 2017. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8007233

[17] S. Isler, R. Sabzevari, J. Delmerico, and D. Scaramuzza, "An information gain formulation for active volumetric 3D reconstruction," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 3477–3484.

[18] Y. Turkar, T. Chase, C. Aluckal, and K. K. Dantu, "Enhancing visual odometry with reliable pixel masking," in *23rd Conference on Robots and Vision*, 2026. [Online]. Available: https://droneslab.github.io/PIXER/

[19] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," Oct. 2016, arXiv:1506.02142 [cs, stat]. [Online]. Available: http://arxiv.org/abs/1506.02142

[20] Y. Turkar, Y. Kim, and K. Dantu, "Active illumination control in low-light environments," in *International Symposium on Experimental Robotics (ISER)*, 2025, to appear. Preprint available at https://arxiv.org/abs/2506.06394.

[21] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," Aug. 2012, arXiv:1206.2944 [stat]. [Online]. Available: http://arxiv.org/abs/1206.2944

[22] K. Singh, Y. Turkar, Y. Kim, M. Lengel, and K. Dantu, "Autonomous culvert inspection on the erie canal using legged robots," in *Workshop on Field Robotics (WFR), International Conference on Robotics and Automation (ICRA) 2025*, 2025. [Online]. Available: https://droneslab.github.io/INSPECTOR/

[23] Y. Turkar, S. Sadeghi, and K. Dantu, "Adaptive illumination control for robot perception," 2026. [Online]. Available: https://arxiv.org/abs/2602.15900