

Final Project Proposal

EE367 Computational Imaging 2026

Topic: Foveated Imaging and Perception

Team Member: Howard Xiao

Motivation and Project Goals:

Recent advances in image sensor technology enable the capture of ultra-high-resolution images and videos. Commercial sensors beyond 200 megapixels (MP) are widely available (Choi et al., 2023) with 400 MP prototypes in development (Canon Inc. 2025). Achieving such resolution requires pixel sizes below $0.6\mu\text{m}$ and significantly increases readout bandwidth and downstream processing demands. As a result, under realistic bandwidth, latency, or power constraints, imaging systems cannot afford to acquire, transmit, or process all pixels at full resolution, making selective acquisition essential.

Ultra-high-resolution imagery is increasingly important not only for visual quality, but also for downstream video perception tasks such as object tracking, text recognition, and robotic manipulation. At the same time, the cost of acquiring and processing ultra-high-resolution video grows prohibitively large as resolution scales, particularly on edge devices such as augmented-reality glasses, drones, autonomous vehicles, and biomedical systems.

Existing approaches address this challenge through acquisition strategies such as spatial or temporal downsampling, which irrevocably discard information before task relevance can be assessed. On the other hand, an emerging class of dual-stream sensors with hundreds of millions of pixels support multiple streams of video data to be read out simultaneously, including low-resolution full-field-of-view frames as well as much smaller but full-resolution regions of interest (ROIs) with a dynamically programmable location in the image (Samsung Electronics Co., 2025). Leveraging these emerging hardware capabilities, in this work we aim to develop real-time, predictive, and task-aware foveation algorithms that address the problem of determining which pixels to acquire when, under real-world constraints.

We formulate foveated image acquisition as a sensor attention policy-learning problem, in which past observations guide actions that directly shape future measurements, closing the perception--acquisition loop for bandwidth-efficient, task-optimal sensing. We aim to demonstrate through extensive simulation that our foveation approach maintains high task performance and significantly outperforms conventional methods in pixel-limited settings across multiple perception tasks. We also aim to prototype our system using a 200MP image sensor and capture real-world videos under realistic bandwidth and latency constraints, demonstrating practical feasibility.

Related Work:

Foveated vision studies how spatial resolution can be allocated non-uniformly across the visual field in order to prioritize task-relevant regions. Recent research works including Policy-based Recurrent Attention Models (RAM) (Mnih et al., 2014) formulate foveation as a sequential

decision-making problem, selecting spatial glimpses conditioned on past observations. The AdaFocus family extends this paradigm to video by learning policies that select task-relevant regions frame by frame from full-resolution inputs (Wang et al., 2025). Our work is closely related in spirit to prior foveated vision approaches, which all post-process high-resolution image and video data, but it differs in a fundamental assumption: because full-resolution frames cannot be efficiently read out by and transferred off the sensor under real-world bandwidth constraints, our method performs foveation at acquisition time, directly determining which measurements are captured.

In foveated rendering and display systems, perceptual models guide level-of-detail and sampling decisions to allocate resources preferentially near the viewer’s gaze fixation (Krajancich et al., 2023; Fan et al., 2024). Although these graphics systems motivate foveation as a principled trade-off between fidelity and efficiency, they typically operate at rendering or display time; in contrast, our method applies foveation during image acquisition, affecting which data is captured rather than post-processing them.

Project Timeline and Milestones:

1. Method of the project would be fixed by end of January with most simulation experiments done and results ready.
2. Real-world captures with 200MP sensor prototype across a wide range of scenes both indoor and outdoor will be completed in February.
3. Final code cleanup and writing will be done before the project deadline in March.

References:

- [1] Sungsoo Choi, Seungjoon Lee, Taeheon Lee, Hochul Ji, Haeyong Park, Dongmo Im, Dongchul Lee, Jinyoung Kim, Sungyong You, Jaeho Choi, et al . 2023. World smallest 200Mp CMOS image sensor with 0.56 μm pixel equipped with novel deep trench isolation structure for better sensitivity and higher CG. In Proceedings of the Int’l Image Sensor Workshop (IISW), Crieff, UK. 22–25.
- [2] Canon Inc. 2025. Canon develops CMOS sensor with 410 megapixels, the largest number of pixels ever achieved in a 35 mm full-frame sensor. <https://global.canon/en/news/2025/20250122.html>.
- [3] Ltd. Samsung Electronics Co. 2025. ISOCELL HP2 | Mobile Image Sensor. <https://semiconductor.samsung.com/image-sensor/mobile-image-sensor/isocell-hp2/> Accessed 2025-11-11.
- [4] Volodymyr Mnih, Nicolas Heess, Alex Graves, and Koray Kavukcuoglu. 2014. Recurrent models of visual attention. *Advances in neural information processing systems* 27 (2014).
- [5] Yulin Wang, Yang Yue, Yang Yue, Huanqian Wang, Haojun Jiang, Yizeng Han, Zanlin Ni, Yifan Pu, Minglei Shi, Rui Lu, et al . 2025. Emulating human-like adaptive vision for efficient and flexible machine visual perception. *Nature Machine Intelligence* (2025), 1–19.
- [6] Brooke Krajancich, Petr Kellnhofer, and Gordon Wetzstein. 2023. Towards attention-aware foveated rendering. *ACM Transactions on Graphics (TOG)* 42, 4 (2023), 1–10.
- [7] Runze Fan, Xuehuai Shi, Kangyu Wang, Qixiang Ma, and Lili Wang. 2024. Scene-aware Foveated Rendering . *IEEE Transactions on Visualization & Computer Graphics* 30, 11 (Nov. 2024), 7097–7106. doi:10.1109/TVCG.2024.3456157