

*"The purpose of computing is insight, not numbers"*  
- Richard Hamming

My research vision is to establish a new paradigm at the intersection of **AI, photography, and visual computing**. I develop **AI-driven methodologies for Computational Photography, Computational Imaging, and Camera Design** to solve fundamental inverse problems in visual data acquisition. By creating differentiable models that bridge physical scene properties and machine perception, my work addresses a foundational question often overlooked in pure data-driven vision: **"Where do images come from, and how can we engineer this process to advance both science and engineering?"**

This approach, which deeply integrates physical models of light transport and sensor physics with modern AI frameworks, serves a dual purpose. First, it enables robust machine perception for autonomous systems. By providing vision algorithms with physically consistent cues—such as accurate depth, material properties, and illumination estimates derived from the imaging process itself—we can build systems that are fundamentally more reliable under challenging real-world conditions like low light, adverse weather, and sensor limitations.

Second, and more broadly, it provides the foundation for high-fidelity visual simulation and digital twins. The principles of computational imaging, when run in reverse, allow us to synthesize photorealistic and physically accurate scenes, dynamic 3D environments, and interactive world models. This capability is crucial for scalable AI training, safe validation of autonomous systems, and the creation of digital twins for applications in advanced manufacturing, computational healthcare, and urban science. Ultimately, my goal is not only to apply AI to imaging but to use the imaging pipeline as a rich domain for developing novel AI-for-Science methodologies—where learning is rigorously constrained by first principles—and to invent next-generation, intelligent imaging systems as tangible AI-for-Engineering outcomes that extend beyond conventional photography.

### Research Progress

#### AI-Driven Computational Photography and Camera Design

##### (1). AI enhanced Image Processing Pipeline

Conventional camera image signal processing (ISP) and photo post-processing techniques rely heavily on handcrafted priors and explicit assumptions about image statistics (e.g., noise patterns, local smoothness). Although effective in constrained settings, they face inherent limitations and often fail to adapt to diverse, complex real-world conditions.

The emergence of data-driven AI has transformed this paradigm, moving the field beyond fixed heuristics. My work actively explores the next frontier of this shift: building imaging systems that are not only data-driven but also efficient, interpretable, and deeply adaptive (see Fig. 1). I have demonstrated this through several key studies: developing a lightweight Transformer with only 90K parameters [1] for robust exposure correction, which proves that extreme efficiency can coexist with high performance; discovering an image-adaptive coordinate system [2] that allows processing operations to dynamically align with image content, enabling finer-grained physical control; and designing a perception-inspired color space that semantically optimizes white balance correction [3].



Figure 1: Our works [1; 2; 3] adopt AI as a tool that advances traditional image processing.

These efforts point toward a future where low-level vision systems are task-aware, resource-conscious, and seamlessly integrated with perceptual and physical constraints.

## (2). Camera-based Real-world Visual Systems

Engineering robust visual systems for real-world applications requires models that inherently understand and adapt to imperfect camera inputs. My work tackles this by co-designing perception algorithms and imaging models, treating challenging conditions as part of the learning problem. For example, in low-quality object detection [4], we reformulate image degradations as self-supervised signals rather than noise to be removed, enabling robust performance across variable resolutions and blur. Similarly, for dark object detection [5], we developed a multi-task framework that jointly optimizes for image enhancement and feature discriminability, ensuring that improving visibility directly translates to reliable perception. This approach of embedding camera-aware degradation models into the AI pipeline exemplifies my commitment to building practically viable, performance-driven vision systems.

## (3). AI for Next Generation Camera Design

The design of imaging hardware—optics, sensors, and processing pipelines—has historically prioritized **human aesthetic quality**. For AI to reliably “see” and interact with the physical world, we must co-design cameras and algorithms to produce **machine-optimal visual data**. My research pioneers this shift by using AI to bridge camera sensor and high-level computer vision, and by drawing inspiration from biological evolution to inform novel camera architectures.

A cornerstone of this effort is our work RAW-Adapter [6; 7] (Fig. 2), which enables vision models to directly interpret RAW sensor data. By learning a lightweight adaptation module, we bypass the handcrafted, human-centric image signal processor (ISP), preserving crucial photometric and geometric information often lost during conventional development. This provides a differentiable pathway from the raw physical signal to semantic understanding, establishing a foundation for task-aware camera optimization. Building on this, Dr.RAW [8] further explores efficient task conditioning from RAW data, moving toward a general vision system that can dynamically align the imaging pipeline with diverse downstream objectives.

Moreover, our work Paleoinspired Vision [9] investigates how principles of biological vision evolution can inspire better artificial cameras. By modeling the co-evolution of color vision and natural environments, we derive spectral sensitivities optimized for scene understanding rather than colorimetry, illustrating how AI can translate biological insights into novel imaging hardware. Overall, this line of research reframes the camera from a fixed capture device into a learnable, task-aware component of an intelligent visual system, enabling cameras co-designed with—and for—AI.

## Applying AI to Understand and Reconstruct the 3D Physical World

Traditionally, a gap has persisted between computational photography, which models the physical world from captured data, and computer graphics, which synthesizes believable visual experiences. The former seeks to faithfully understand reality, while the latter aims to create programmable digital representations of it. With the advent of **AI-driven neural scene representations**, most notably

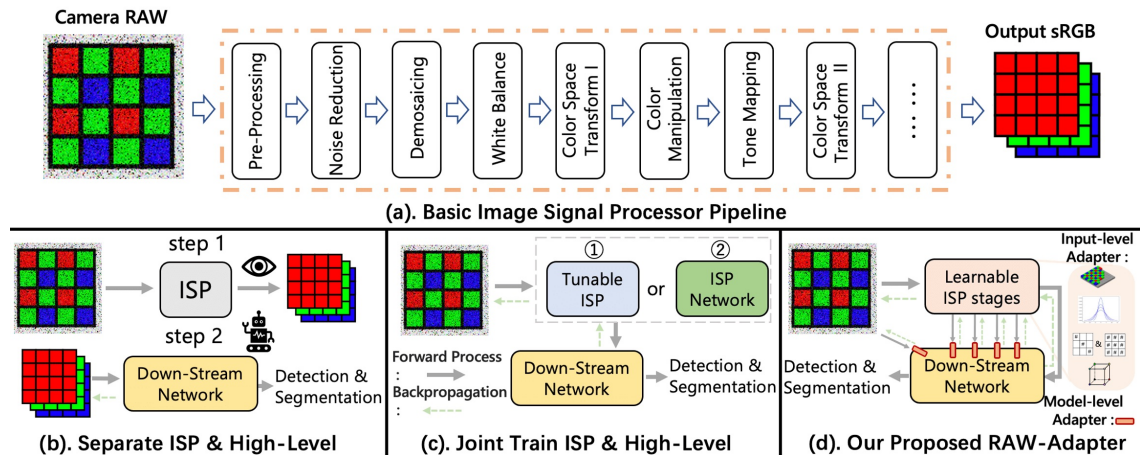


Figure 2: Our work [6; 7; 8] optimize camera image signal processor (ISP) for downstream vision.

the Neural Radiance Field (NeRF) [10], this divide has been fundamentally bridged. By learning continuous scene representations directly from images, NeRF provide a unifying framework that bridges physical image capture with editable 3D synthesis. This shift enables not only photorealistic view generation but also the extraction of implicit physical properties, paving the way for AI to build dynamic and comprehensible digital models of the world.

My research builds directly upon this paradigm, extending neural representations to model more complex and challenging physical phenomena in real-world scenes. In Aleth-NeRF [11], we address the critical problem of variable illumination by introducing a disentangled framework that separately models scene reflectance and transient "concealing field", enabling robust 3D reconstruction under changing lighting conditions. Further advancing physical grounding, I2-NeRF [12] explicitly simulates the effects of participating media (e.g., dark, fog, underwater) within the volume rendering integral, allowing for accurate scene recovery and clear-view synthesis in degraded visibility. Finally, our work also explores next-generation representations, as seen in Luminance-GS [13], which adapts 3D Gaussian Splatting [14] for challenging lighting via a view-adaptive tone adjustment mechanism. Collectively, these efforts demonstrate how AI-driven neural scene representations can evolve from purely visual models into physics-aware digital twins, capable of interpreting and reconstructing the 3D world under the complex conditions in which autonomous systems must operate.

## Ongoing and Future Directions

My future research program builds upon this foundation, aiming to create a virtuous cycle between physical measurement and visual intelligence, which will be pursued along three synergistic thrusts.

### Thrust 1: Co-designing Intelligent Cameras and Robotic Perception

Robotic perception relies heavily on cameras, yet existing camera designs are still not systematically optimized for robotics. This thrust aims to bridge that gap through holistic co-design, establishing a tight feedback loop across the camera lens, CMOS hardware, ISP pipeline, and downstream perception algorithms. The goal is to develop robust, task-specific imaging systems that maintain reliable performance under challenging real-world conditions such as low light, water, and fog, as well as in resource-constrained environments, including Mars or polar regions.

### Thrust 2: Solving Inverse Problems for Scientific and Industrial Imaging

My research will leverage computational photography to solve the fundamental inverse problem of recovering a scene's intrinsic physical properties from images. This involves inverting the imaging

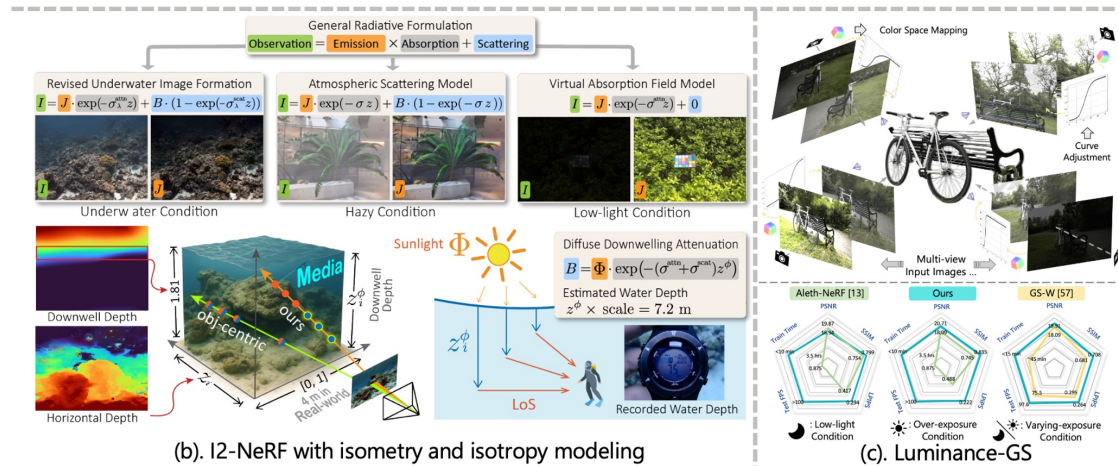
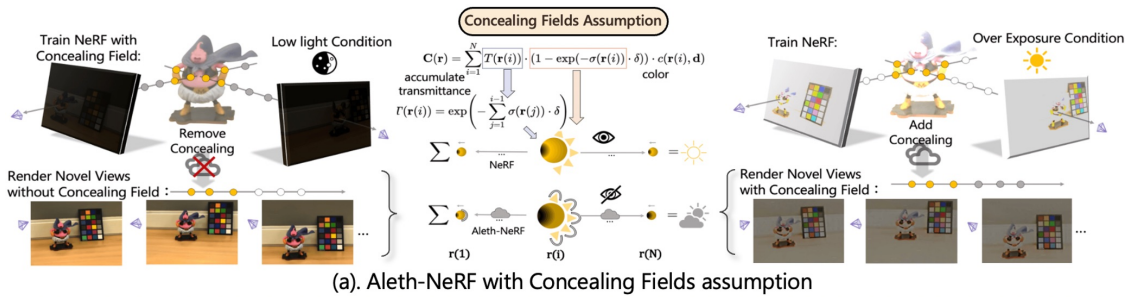


Figure 3: Our works [11; 13; 12] incorporate physics into 3D AI vision tools, aiming to bridge computer graphics and computational photography to better represent the 3D physical world.

pipeline—from RGB back to raw sensor data and ultimately to the underlying photometric and geometric scene attributes. By developing physics-aware inverse models that decompose images into factors like albedo, shading, and material properties, this work enables quantitative, interpretable scene understanding. Such capabilities are essential for applications requiring high accuracy and physical consistency, from scientific imaging (e.g., biomedical or remote sensing) to industrial tasks (e.g., precision inspection and augmented reality).

### Thrust 3: Physics-Guided Generative Models for Data Simulation

This thrust focuses on integrating the physical principles of image formation into generative AI. The core idea is to build physics-embedded generative models where the forward processes of optics, geometry, and material are internalized within the network’s architecture or training objective. Such models go beyond traditional data-driven synthesis—they can generate visually realistic and physically consistent scenes, objects, and phenomena by construction. This capability enables high-quality synthetic data generation for training robust systems, forms the foundation for building realistic, interactive simulators and digital twins for robotics, autonomous testing, and virtual prototyping.

### Publications

- [1] **Ziteng Cui**, Kunchang Li, Lin Gu, Shenghan Su, Peng Gao, ZhengKai Jiang, Yu Qiao, and Tatsuya Harada. You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction. In *33rd British Machine Vision Conference 2022, BMVC 2022, London, UK, November 21-24, 2022*. BMVA Press.
- [2] **Ziteng Cui**, Lin Gu, and Tatsuya Harada. Discovering an image-adaptive coordinate system for photography processing. In *35rd British Machine Vision Conference 2022, BMVC 2024*. BMVA Press.

- [3] Yang Cheng, **Ziteng Cui**, Lin Gu, Shenghan Su, and Zenghui Zhang. Perception-inspired color space design for photo white balance editing. In *WACV, 2026*.
- [4] **Ziteng Cui**, Yingying Zhu, Lin Gu, Guo-Jun Qi, Xiaoxiao Li, Renrui Zhang, Zenghui Zhang, and Tatsuya Harada. Exploring resolution and degradation clues as self-supervised signal for low quality object detection. In *European Conference on Computer Vision (ECCV)*, pages 473–491. Springer, 2022.
- [5] **Ziteng Cui**, Guo-Jun Qi, Lin Gu, Shaodi You, Zenghui Zhang, and Tatsuya Harada. Multitask aet with orthogonal tangent regularity for dark object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2553–2562, October 2021.
- [6] **Ziteng Cui** and Tatsuya Harada. Raw-adapter: Adapting pretrained visual model to camera raw images. In *European Conference on Computer Vision (ECCV)*. Springer, 2024.
- [7] **Ziteng Cui**, Jianfei Yang, and Tatsuya Harada. Raw-adapter: Adapting pre-trained visual model to camera raw images and a benchmark. *arXiv preprint arXiv:2503.17027*, 2025.
- [8] Wenjun Huang\*, **Ziteng Cui\*** (co first), Yinqiang Zheng, Yihui He, Tatsuya Harada, and Mohsen Imani. Dr. RAW: Towards general high-level vision from RAW with efficient task conditioning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- [9] Junjie Zhang, Zhimin Zong, Lin Gu, Shenghan Su, **Ziteng Cui**, Yan Pu, Zirui Chen, Jing Lu, Daisuke Kojima, Tatsuya Harada, et al. Paleoinspired vision: From exploring colour vision evolution to inspiring camera design. *arXiv preprint arXiv:2412.19439*, 2024.
- [10] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020.
- [11] **Ziteng Cui**, Lin Gu, Xiao Sun, Xianzheng Ma, Yu Qiao, and Tatsuya Harada. Aleth-nerf: Illumination adaptive nerf with concealing field assumption. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2024.
- [12] Shuhong Liu, Lin Gu, **Ziteng Cui**, Xuangeng Chu, and Tatsuya Harada. l2-nerf: Learning neural radiance fields under physically-grounded media interactions. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- [13] **Ziteng Cui**, Xuangeng Chu, and Tatsuya Harada. Luminance-gs: Adapting 3d gaussian splatting to challenging lighting conditions with view-adaptive curve adjustment. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.
- [14] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), July 2023.