Low-Light High Dynamic Range Single Frame Image Denoising for Quanta Image Sensors

Yiheng Chi[1], Preston Rahim[1], and Stanley H. Chan[1]
[1] DeepLux Technology Inc., West Lafayette, IN 47906, USA
{yiheng.chi, preston.rahim, stanley.chan}@deeplux.tech

Abstract – Imaging low-light high dynamic range (HDR) scenes in a single capture is challenging for conventional sensors when exposure bracketing is not feasible due to application constraints. Advancements in sensor technology have narrowed the gap, as split-pixel and dual conversion gain (DCG) enables single-frame HDR capture and Quanta Image Sensors (QIS) allow counting individual photons at low light. However, removing shot noise from a single HDR image remains a difficult task due to the spatially varying nature of noise. To address this issue, we propose a learnable pipeline with a modular design for processing high bit-depth QIS raw images. Compared to existing algorithmic solutions, our approach offers superior reconstruction performance and greater robustness to variations in illuminance and noise.

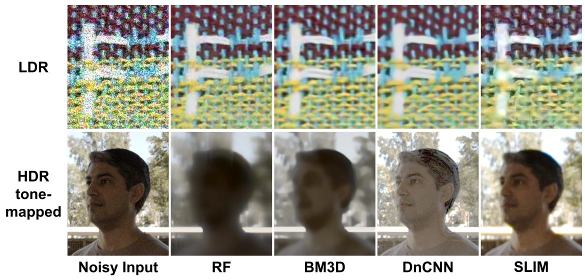

Fig. 1. Existing image denoising methods typically operate within narrow dynamic range limits. While offering satisfactory results denoising 8-bit images (top row), the conventional denoisers such as RF [4] and BM3D [5] and learning-based methods such as DnCNN [13] and SLIM learnable ISP [3] either generate blurry results or fail to remove the noise when denoising 14-bit HDR images (bottom row).

I. Introduction

Imaging a high dynamic range (HDR) scene in a single capture is particularly challenging due to the limited exposure range of most cameras, which often results in overexposed highlights or underexposed shadows. The conventional solution is to capture multiple images and fuse them with exposure bracketing; however, exposure bracketing is not applicable for most real-time applications, especially those that pose limitations on data throughput, power consumption, and computational complexity, such as high resolution HDR videos, medical imaging, and autonomous driving. Recent advances in sensor design, such as split-pixel [10], [12] and dual conversion gain (DCG) [9], attempt this issue by varying pixel exposure and gain spatially, thereby generating a single HDR image. These designs have demonstrated promising results, while post-processing is still needed when taking images at extreme low light, where the sensor read noise and photon shot noise are severe compared to the image signal. Recovering clean HDR images from the noisy single-frame observations is a critical task, as it empowers image sensors to see at a very low photon level while consuming little memory, power, and latency, thus enabling applications that have been very difficult in the past: low-light endoscopy, night-time auto-navigation, low-light object detection and tracking, etc.

To combat the low signal-to-noise ratio (SNR) issue, single-bit and multi-bit Quanta Image Sensors (QIS) are a candidate solution for low-light image capture with their sub-electron read noise. Recent research and development have also demonstrated the potential of single-bit QIS in HDR imaging applications [7], [8], [1], [2]. While one-bit QIS is already capable of imaging low-light HDR scenes, it requires capturing, transmitting, and combining 1,000 or more frames instead of a single capture. This is an alternative approach to low light HDR which comes with another set of challenges. Conversely, we use a multi-bit QIS to capture a single image instead of single-bit with multiple captures. Although QIS is at an advantage of imaging at low light, removing shot noise from a single HDR image frame is still challenging due to the spatially varying photon shot noise characteristics corresponding to the significant change in illuminance, while existing image denoising typically operate within narrow dynamic range limits. Therefore, our proposed method attempts to address this problem. Notably, multi-bit QIS pixel is similar to any CMOS pixel, except a much lower read noise, so our approach can be extended to high bit-depth CMOS sensors by its nature.

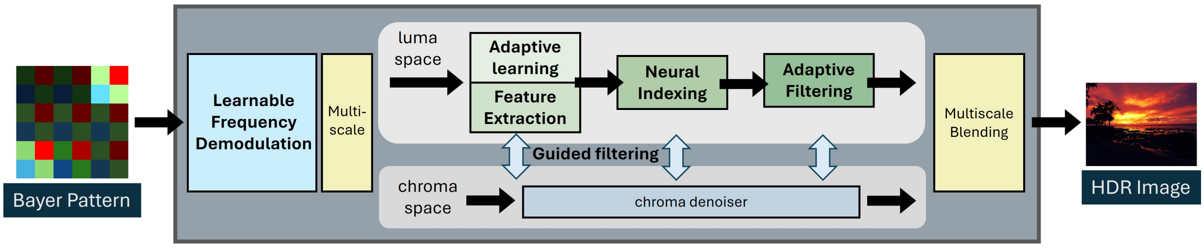

Fig. 2. Schematic diagram of the proposed method. We replace bottleneck building blocks of traditional ISP with learning-based building blocks to simultaneously handle noise and HDR.

In this paper, we present an algorithmic solution for processing high bit-depth QIS raw images. The input to our solution is multi-bit QIS data from a Bayer color filter array pattern, e.g., 20-bit images which are already achievable as reported using the 1280×960 2.8μm DCG and split-pixel combined sensor [11], after standard pre-processing steps, e.g., gray-level offset, pixel response non-uniformity calibration, dead pixel removal, etc. Our work focuses on **single image** low-light HDR demosaicking and denoising. This problem remains unsolved as existing HDR methods either fail to handle heavy noise or require multiple captures of the same scene. To address this, we propose a learnable pipeline with a modular design that enhances flexibility and simplifies debugging. In contrast to traditional rule-based image signal processing pipelines, our approach dynamically generates and applies filters to local image patches using learned operators, offering superior reconstruction performance and greater robustness to variations in illuminance and noise.

II. High Bit-Depth QIS Image Denoising

Our goal is to solve the problem of recovering the signal from noise by using only a single (HDR) capture. A common misconception about HDR imaging is that the noise in an HDR image can be handled just like any conventional image. Unfortunately, this is not true because the signal-to-noise ratio (SNR) of a pixel is proportional to the number of photons it receives. In the presence of very bright and very dark content, the noise distribution will be non-uniform within the single image. Figure 1 presents an example of an 8-bit image and a 14-bit image. The 8-bit image is an LDR image. In this example, we simulate the photon shot noise as if the image were captured at a light level of on average 10 photons per pixel. Although the noise is heavy, existing image denoisers are capable of handling the noise. However, when we switch the problem to 14-bit (HDR), the same algorithms fail to produce any meaningful result. This shows the limitations of existing solutions and motivates the need for **single-image** low-light HDR denoising for high bit-depth sensors.

Figure 2 illustrates the schematic of our proposed solution. Key components include: learnable frequency demodulation for Bayer color filter array demosaicing, content-aware feature extraction via adaptive learning to capture spatial information from demodulated color channels, learned continuous neural indexing with learned adaptive filtering to adaptively compose filters suited to the characteristics of image patches, and a multiscale blending design. Our proposed solution expands on our learning-based ISP [3] with additional designs to manage the wide dynamic range. In the proposed pipeline, we replace key steps in traditional denoising processes by learning-based methods.

Learnable Frequency Demodulation. Frequency demodulation is the concept we developed previously [6] and succeeded in [3]. The main idea is to convert the raw Bayer pattern into the luma and chroma channels. On HDR image denoising for QIS, we extend this idea to tackle the low SNR problem at low light regime of the dynamic range. This is because, by applying frequency demodulation, we obtain the luma signal which carries a SNR that is 3x that of individual Bayer pixels, making the denoising easier. We then use the features extracted from the luma signal to guide the filtering of the two chroma channels. We design this module learnable so that it can be co-optimized with other components. Another idea we integrate to mitigate low SNR is larger receptive fields. While the image is noisy, the underlying signal still contains local structures, such as edges, surfaces, and textures. By accumulating more information about these structures from a larger receptive field, it is more likely to recover the image signal.

Adaptive Learning and Adaptive Filtering. Adaptive learning and adaptive filtering are a new concept we introduce to account for the spatially-varying SNR. In [3] and almost all other convolution-based network designs, the operations in the image and the feature space are spatially invariant and fixed after training. We recognize that this is a big limitation if we want to handle tasks such as removing spatially-varying noise and generalizing it to a wide range of testing scenarios.

In this work, we first apply a small feature encoding which can be interpreted as an illumination or noise level estimation. We use this result to control the subsequent feature processing. Such guidance is realized by using separated mini processing branches and feedback signals. A similar estimation and control strategy is also applied to adjust the strength of filtering. This concept enables our proposed method to adapt to changes in SNR within a single image, and it improves the generalization capabilities across a wide range of environmental factors as well, such as lighting conditions, scene content, and color spectra.

Feature Extraction and Neural Indexing. In [3], we replaced the conventional image denoising idea of using local gradients to select filters by using features extracted from a shallow neural network. We then mapped these features to an array of weights indicating the relative emphasis of the filter at each index in a filter bank. Expanding this idea, we extract more complex features with dedicated network branches to represent more information about local image contents, such as brightness and noise level. These features will guide the processing in later stages within the pipeline and determine the final filter at each local image patch. We further propose a new indexing scheme, called neural indexing, that allows dense representations of filter weights and adaptive filter construction on the fly. Instead of directly predicting filter weights, which can be very sparse since most filters in the filter bank are irrelevant to any given specific image patch, the proposed module only outputs a few parameters and activates only the relevant filters. Then, through a computationally cheap mapping, these parameters and active filters are converted into the final filter to be applied to denoise the image patch.

III. Results and Conclusion

We test our method on single-frame 20-bit low-light HDR images. We first capture LDR images with a Sony ILCE-7M2 camera and a Sigma Art 24-70mm F2.8 DG DN lens. At each scene, twelve images are captured at a fixed ISO of 100 and an aperture of f/5. The exposure times are 1/80, 1/40, 1/20, 1/10, 1/5, 0.4, 0.8, 1.6, 3.2, 6, 13, and 25 seconds. We then combine these frames to create HDR images of 60 dB dynamic range using exposure bracketing, which only serve to verify our method and do not represent the deployment scenarios. The input low-light QIS images for experiments are then simulated from these HDR images assuming a Bayer color filter array and using parameters of 0.19e- read noise, 0.02e-/s dark current, 20-bit analog-digital converter, 50% quantum efficiency, and a uniform sensor response.

Figure 3 presents low-light HDR scenes and the photon shot noise with varying strength across single-frame high bit-depth images. It also compares existing image denoising methods with our proposed solution. Traditional denoising techniques operate within a narrow dynamic range, leading to either over-smoothed or noisy results, whereas our method produces HDR image reconstruction that adapt to the local photon shot noise strengths.

References

[1] Stanley H Chan. What does a one-bit quanta image sensor offer? *IEEE Transactions on Computational Imaging*, 8:770–783, 2022.

[2] Stanley H Chan. On the insensitivity of bit density to read noise in one-bit quanta image sensors. *IEEE Sensors Journal*, 23(4):3666–3674, 2023.

[3] Stanley H Chan, Yiheng Chi, and Preston Rahim. SLIM: Small and learnable image signal processing module for CMOS and quanta image sensors. In *International Image Sensor Workshop*, 2023.

[4] Yiheng Chi and Stanley H Chan. Fast and robust recursive filter for image denoising. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 1708–1712. IEEE, 2018.

[5] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.*, 16(8):2080–2095, 2007.

[6] Omar A Elgendy, Abhiram Gnanasambandam, Stanley H Chan, and Jiaju Ma. Low-light demosaicking and denoising for small pixels using learned frequency selection. *IEEE Transactions on Computational Imaging*, 7:137–150, 2021.

[7] Abhiram Gnanasambandam and Stanley H Chan. HDR imaging with quanta image sensors: Theoretical limits and optimal reconstruction. *IEEE transactions on computational imaging*, 6:1571–1585, 2020.

[8] Jiaju Ma, Dexue Zhang, Dakota Robledo, Leo Anzagira, and Saleh Masoodian. Ultra-high-resolution quanta image sensor with reliable photon-number-resolving and high dynamic range capabilities. *Scientific Reports*, 12(1):13869, 2022.

[9] Johannes Solhusvik, Sam Hu, Robert Johansson, Zhiqiang Lin, Siguang Ma, Keiji Mabuchi, Sohei Manabe, Duli Mao, Bill Phan, Howard Rhodes, et al. A 1392x976 2.8 μm 120db CIS with per-pixel controlled conversion gain. In *Proceedings of the 2017 International Image Sensor Workshop*, pages 298–301, 2017.

[10] Johannes Solhusvik, Jiangtao Kuang, Zhiqiang Lin, Sohei Manabe, Jeong-Ho Lyu, Howard Rhodes, et al. A comparison of high dynamic range CIS technologies for automotive applications. In *Proceedings of the 2013 International Image Sensor Workshop*, pages 12–16, 2013.

[11] Johannes Solhusvik, T Willassent, Sindre Mikkelsen, Mathias Wilhelmsen, Sohei Manabe, Duli Mao, Zhaoyu He, Keiji Mabuchi, and TA Hasegawa. 1280×960 2.8 μm HDR CIS with DCG and split-pixel combined. In *Proceedings of the International Image Sensor Workshop (IISW), Snowbird, UT, USA*, pages 23–27, 2019.

[12] Trygve Willassen, Johannes Solhusvik, Robert Johansson, Sohrab Yaghmai, Howard Rhodes, Sohei Manabe, Duli Mao, Zhiqiang Lin, Dajiang Yang, Orkun Cellek, et al. A 1280x1080 4.2 μm split-diode pixel HDR sensor in 110 nm BSI CMOS process. In *Proceedings of the International Image Sensor Workshop*, pages 8–11, 2015.

[13] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7):3142–3155, 2017.

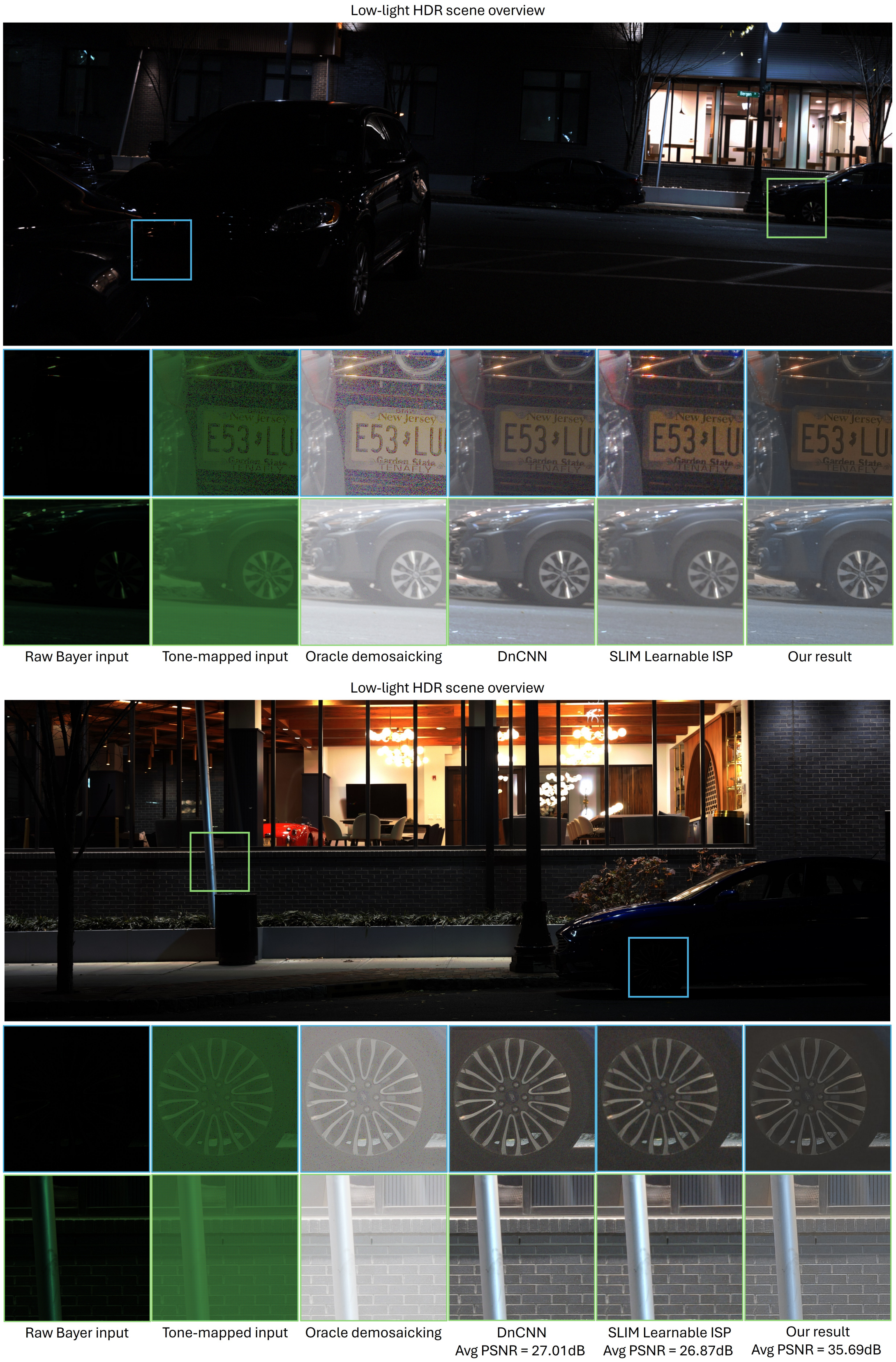

Fig. 3. Denoising and demosaicking results. The oracle demosaicing represents the best achievable demosaicing result, assuming that the capture is not Bayer-filtered, i.e., every pixel already has the full RGB values, instead of having just one due to the color filter array. This oracle case serves as a reference to illustrate the amount of noise we are handling here. Compared to existing learning-based image denoising [13] and learnable ISP [3], our method produces HDR image reconstruction that adapt to the local photon shot noise strengths.