2024 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 22–25, 2024, LONDON, UK

PHOTON-LIMITED IMAGING WITH QUANTA IMAGE SENSORS VIA AN UNSUPERVISED LEARNING FRAMEWORK

*Haosen Liu, Yunping Zhang, Edmund Y. Lam**

Department of Electrical and Electronic Engineering,
The University of Hong Kong, Pokfulam, Hong Kong SAR, China

ABSTRACT

Due to their single-photon sensitivity, quanta image sensors (QIS) are designed to complement traditional image sensors for a wide range of applications in photon-limited imaging conditions. However, the binary nature of QIS data poses compatibility challenges with existing image processing tools, necessitating the development of specialized reconstruction algorithms. While training a deep neural network with paired QIS recordings and corresponding ground truth data in a supervised manner offers superior performance compared to closed-form optimization-based solutions, collecting such a dataset can be laborious or impractical in certain scenarios. To address this issue, we propose an unsupervised framework that eliminates the reliance on clean ground truth data. Experimental results highlight the superiority of our method over other unsupervised, model-based approaches, particularly in terms of image reconstruction quality. Notably, our proposed method demonstrates competitiveness with the supervised learning method while circumventing the need for labeled training data.

***Index Terms*—** Quanta image sensor, photon-limited imaging, unsupervised learning, image reconstruction

1. INTRODUCTION

Quanta image sensor (QIS) is a single-photon imaging device envisioned to be the next generation image sensor. Different from traditional image sensors that measure and output the integration of photoelectrons as image intensity, QIS operates by detecting individual photons and providing a binary signal indicating whether the photon count exceeds a specific threshold, denoted as q. It has demonstrated extraordinary advantages in various aspects including high spatial resolution, rapid readout speed, and an extensive dynamic range [1, 2]. Owing to its novelty, QIS can be potentially applied to a wide range of scenarios such as security, night vision, and even quantum computing, and so on [3, 4, 5].

To unlock such potential, it is crucial to develop dedicated computational algorithms capable of processing its binary data, which possesses distinct properties compared to traditional images. However, this can not make full use of the great progress made by the existing computer vision community. For example, the deep neural networks trained on traditional image data could not be effectively transferred to the binary QIS data. Consequently, the objective of this study is to devise an image reconstruction method that can efficiently convert binary QIS data into traditional images, serving as a bridge between the existing computer vision toolbox and QIS technology.

One key challenge of the image reconstruction for QIS is that the 1-bit binary information is extremely limited. To alleviate this problem, it is common for existing methods to adopt the oversampling strategy, which averages a set of 1-bit frames to obtain a multi-bit frame based on the assumption that each pixel can be oversampled multiple times in the spatio-temporal domain. However, due to the practical limitation, the sampling ratio should not be too large. Consequently, the averaged multi-bit frame is still of low signal-to-noise ratio (SNR) [6].

Due to the existence of noise, directly reconstructing the image from QIS recordings with the maximum likelihood estimation (MLE) [6] results in an image that is inherently corrupted. To tackle the influence of noise, one approach is to explicitly or implicitly integrate image prior models into the image reconstruction process. For example, a pair of variance stabilizing transforms (i.e., Anscombe binomial transform [7]) and a Gaussian denoiser (i.e., BM3D [8]) have been incorporated to recover images from averaged multi-bit QIS frame in [9]. Another approach is to resort to the deep learning methods that have exhibited great success in various image processing fields. For example, an end-to-end QIS reconstruction network (QISNet) has been proposed in [10]. Generally speaking, the supervised deep learning methods can outperform the unsupervised model-based methods by a large margin in terms of the reconstruction quality. However, the performance of deep learning methods heavily relies on the collection of pairs of QIS data and high-quality images, which is not only time-consuming but also sometimes impractical under challenging conditions where high-quality images are hard to collect.

*Corresponding author: elam@eee.hku.hk

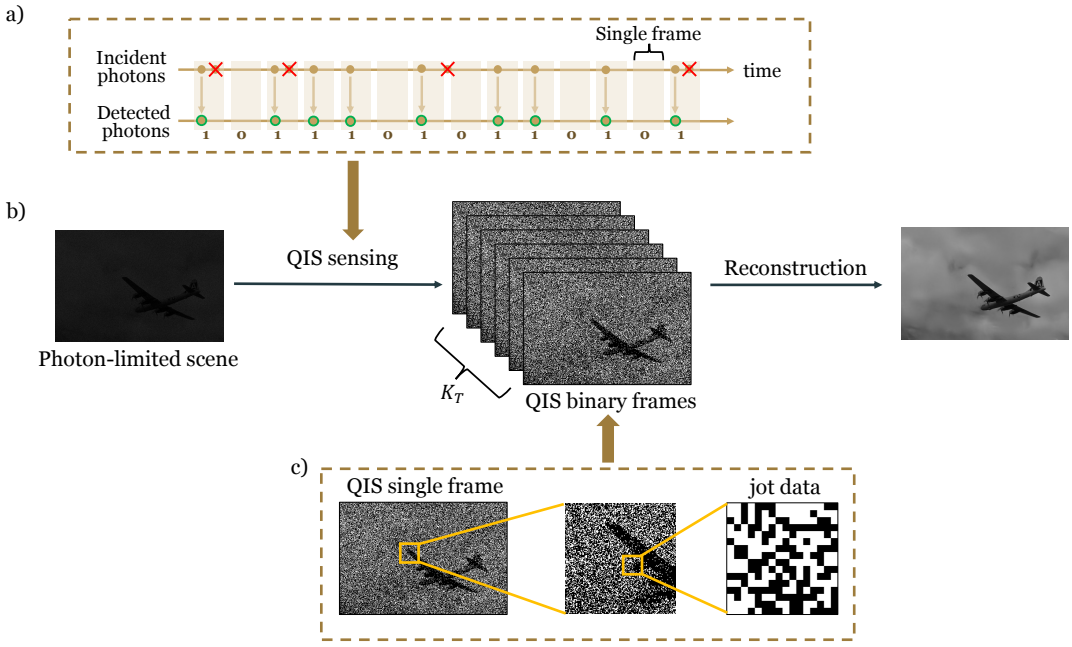

Fig. 1. Illustration of image formation and reconstruction process of QIS. (a) The incident photons are detected by QIS with a rapid sampling rate to produce binary signals. (b) In the context of photon-limited imaging, the primary goal of the reconstruction process is to estimate the underlying gray-scale image from the binary QIS frames. (c) The conceptual illustration of a single QIS frame (left), an enlarged region-of-interest (middle) with a closer view, and (c) raw jot data (right).

Inspired by Noise2Noise (N2N) [11], which is originally designed for training deep neural networks without clean images for the image denoising task, an unsupervised learning framework is proposed for reconstructing images from QIS frames. This work points out that the N2N can be applied to training a network to estimate the probability of photon arrivals. The proposed unsupervised QIS reconstruction learning framework utilizes pairs of low-SNR QIS frames as training data, without access to the corresponding ground truth information. Through this approach, the network trained within the unsupervised framework surpasses traditional model-based methods and achieves comparable performance to networks trained using supervised learning frameworks, all without the need for clean ground truth images.

The rest of this paper is organized as follows. Section 2 introduces QIS imaging model. Section 3 introduces the newly proposed unsupervised QIS reconstruction learning framework. Section 4 provides the experimental setting and comparison results. Section 5 summarizes this work.

2. BACKGROUND

In this section, we present an overview of the imaging modality employed by QIS, which has been extensively analyzed and validated through real experimental tests conducted in prior research [1, 2, 12].

The number of photons, denoted as Y, incident on the unit sensing area follows a Poisson distribution. This statistical distribution can be mathematically modeled as

$$p_Y(k) = \frac{\varphi^k e^{-\varphi}}{k!}, \tag{1}$$

where φ is the exposure which is assumed to be constant over the unit space. In a photosensitive device, the process of converting accumulated photons into electrical signals involves a quantization mechanism. Traditional sensors, such as CMOS, employ an analog-to-digital converter to generate digital signals consisting of 8-14 bits. In contrast, QIS utilizes an array of specialized photodetectors called "jots" that exhibit the capability to detect individual photon arrivals. Notably, QIS adopts a binary quantization scheme, operating on a 1-bit basis. Each jot within QIS array produces a binary output B, representing a truncated form of the signal Y, determined by a threshold value q:

$$B = \begin{cases} 1, & Y \geq q, \\ 0, & Y < q. \end{cases} \tag{2}$$

The conceptual illustration of image formation and reconstruction of QIS is demonstrated in Fig. 1. The probability distribution of the Bernoulli random variable B is characterized by the random variable Y, which is sampled from the distribution described in Eq. 1. Hence, the probability of ob-

serving B is

$$\Pr[B=1] = \sum_{k=q}^{\infty} \frac{\varphi^k e^{-\varphi}}{k!}, \tag{3}$$

$$\Pr[B=0] = 1 - \sum_{k=q}^{\infty} \frac{\varphi^k e^{-\varphi}}{k!}. \tag{4}$$

In this study, we investigate the system with a uniform quantization threshold $q = 1$. While the design of an optimal threshold is an emerging area of research that holds potential for improving the performance of QIS devices, particularly in high dynamic range imaging [13], we choose to adopt a standard configuration for the purpose of consistency and comparability with existing literature [14, 15]. Consequently, the probabilities of Eq. 3 become

$$\Pr[B=1] = \sum_{k=1}^{\infty} \frac{\varphi^k e^{-\varphi}}{k!} = 1 - e^{-\varphi}, \tag{5}$$

$$\Pr[B=0] = 1 - \Pr[B=1] = e^{-\varphi}. \tag{6}$$

which gives the probability mass function $p_B(b)$ of the one-bit QIS output for each jot as

$$p_B(b) = \Pr[B=b] = (1-b)e^{-\varphi} + b(1-e^{-\varphi}). \tag{7}$$

Due to its compact jot size and rapid temporal response, QIS is recognized as an oversampling device both in temporal and spatial dimensions with sampling ratio K_T and K_S, respectively [6]. Specifically, given an object field that is described with N degree of freedom $\boldsymbol{\varphi} = [\varphi_0, \varphi_1, \ldots, \varphi_{N-1}]$, QIS outputs a signal $\mathbf{b} = [b_0, b_1, \ldots, b_{NK-1}]$ with an oversampling factor $K = K_T K_S$. The joint distribution is given as:

$$\begin{aligned} f(\mathbf{b}|\boldsymbol{\varphi}) &= \prod_{n=0}^{N-1} \prod_{k=0}^{K-1} \big((1-b_{nK+k}) e^{-\frac{\varphi_n}{K}} \\ &\qquad + b_{nK+k}(1-e^{-\frac{\varphi_n}{K}})\big), \\ &= \prod_{n=0}^{N-1} (e^{-\frac{\varphi_n}{K}})^{K(1-\bar{b}_n)} (1-e^{-\frac{\varphi_n}{K}})^{K\bar{b}_n}, \end{aligned} \tag{8}$$

where $\bar{b}_n = \frac{1}{K}\sum_{k=0}^{K-1} b_{nK+k}$ for notation simplicity.

The objective of QIS reconstruction is to recover the underlying image, which represents the intensity of flux over the exposure time, from the 1-bit signal observation streams. Due to the novelty of QIS technology and its limited commercial availability [14, 16], the number of reconstruction methods specifically designed for QIS devices is currently limited. Several approaches have been explored in this context, such as MLE [6], alternating direction method of multipliers (ADMM) [17], Transform-denoise method (TD) [9] and neural networks [10]. In this particular research paper, our main objective is to propose an innovative unsupervised learning framework for QIS image reconstruction. Within this framework, we intend to develop a methodology that does not rely on labeled training data or prior knowledge about the specific image content.

3. METHODOLOGY

The objective of QIS reconstruction is to recover an object field that is described with N degree of freedom $\boldsymbol{\varphi} = [\varphi_0, \varphi_1, \ldots, \varphi_{N-1}]$, given the binary QIS outputs $\mathbf{b} = [b_0, b_1, \ldots, b_{NK-1}]$. Since QIS is regarded as an oversampling device with an oversampling factor K, we define $\bar{\mathbf{b}} = [\bar{b}_0, \bar{b}_1, \ldots, \bar{b}_{N-1}]$, where each component is calculated as

$$\bar{b}_n = \frac{1}{K} \sum_{k=0}^{K-1} b_{nK+k}. \tag{9}$$

Following QIS model introduced in Section 2, the binary output b_n from QIS follows the i.i.d Bernoulli distribution as defined in Eq. 7. Consequently, the expectation of $\bar{b}_n$ is

$$\mathbb{E}\big(\bar{b}_n\big) = \frac{1}{K} \sum_{k=0}^{K-1} \Pr[b_{nK+k} = 1] = 1 - e^{-\varphi_n}. \tag{10}$$

It suggests that with the available expectation of $\bar{\mathbf{b}}$, the image intensity (or photon flux) $\boldsymbol{\varphi}$ can be directly calculated using

$$\boldsymbol{\varphi} = -\log\left(1 - \mathbb{E}\big(\bar{\mathbf{b}}\big)\right). \tag{11}$$

Thus, the problem of image reconstruction is reformulated as a probability estimation task, in which an estimation network $f_\theta(\bar{\mathbf{b}})$ is trained to approximate the expectation:

$$\hat{\boldsymbol{\varphi}} = -\log\left(1 - f_\theta(\bar{\mathbf{b}})\right), \tag{12}$$

where θ stands for the trainable parameters. In supervised learning using a large number of pairs $\big(\bar{\mathbf{b}}_i, \mathbb{E}\big(\bar{\mathbf{b}}_i\big)\big)$, θ is optimized by minimizing the loss function as

$$\theta = \arg\min_\theta \sum_i \left\| f_\theta(\bar{\mathbf{b}}_i) - \mathbb{E}\big(\bar{\mathbf{b}}_i\big) \right\|_2^2. \tag{13}$$

However, to get the promised results, a large volume of high-quality ground truth data of the actual photon flux $\boldsymbol{\varphi}_i$ is necessary to calculate $\mathbb{E}\big(\bar{\mathbf{b}}_i\big)$ using Eq. 11. This requirement can be labor-intensive or even infeasible under photon-limited imaging scenarios.

To overcome this challenge, inspired by N2N [11], a key observation in this work is that, given pairs of independent QIS data resulting from the same underlying exposure $\boldsymbol{\varphi}_i$, i.e., $\bar{\mathbf{b}}_{i,1}$ and $\bar{\mathbf{b}}_{i,2}$, the probability estimation network can be trained in an unsupervised manner with the loss function as

$$\theta = \arg\min_\theta \sum_i \left\| f_\theta(\bar{\mathbf{b}}_{i,1}) - \bar{\mathbf{b}}_{i,2} \right\|_2^2. \tag{14}$$

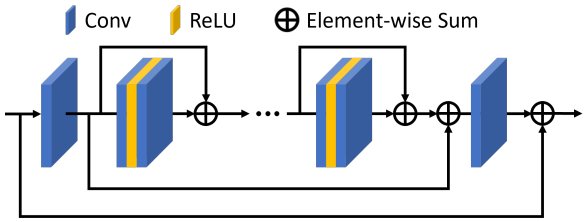

Fig. 2. Network architecture. This network consists of 1 convolutional layer, 16 residual blocks, and 1 convolutional layer. A long skip connection from head to tail is adopted to enable the residual learning.

The proof supporting this observation is provided as follows:

$$\begin{aligned}
&\mathbb{E}\Big(\big\|f_\theta(\bar{\mathbf{b}}_{i,1}) - \bar{\mathbf{b}}_{i,2}\big\|_2^2\Big) \\
=&\mathbb{E}\big(f_\theta^T(\bar{\mathbf{b}}_{i,1})f_\theta(\bar{\mathbf{b}}_{i,1}) - 2f_\theta^T(\bar{\mathbf{b}}_{i,1})\bar{\mathbf{b}}_{i,2} + \bar{\mathbf{b}}_{i,2}^T\bar{\mathbf{b}}_{i,2}\big)\,, \\
=&f_\theta^T(\bar{\mathbf{b}}_{i,1})f_\theta(\bar{\mathbf{b}}_{i,1}) - 2f_\theta^T(\bar{\mathbf{b}}_{i,1})\mathbb{E}\big(\bar{\mathbf{b}}_{i,2}\big) + const, \\
=&\big\|f_\theta(\bar{\mathbf{b}}_{i,1}) - \mathbb{E}\big(\bar{\mathbf{b}}_{i,2}\big)\big\|_2^2 + const, \qquad (15)
\end{aligned}$$

where $const$ denotes constant numbers that do not affect the parameter learning process. Given that $\bar{\mathbf{b}}_{i,1}$ and $\bar{\mathbf{b}}_{i,2}$ are paired measurements obtained from the same underlying exposure $\boldsymbol{\varphi}_i$, it can be assumed that they share the same expectation. As such, Eq. 15 indicates that training with the loss functions given by Eq. 13 and Eq. 14 are statistically equivalent. In other words, one can train a network to estimate the probability of photon arrival directly from the paired QIS data, without the need of clean images. In practice, one way to collect such paired QIS data is to use a beam splitter and two paired QIS devices. Another approach is to divide the binary stream collected by one QIS device into two parts, where one serves as the input and the other as the target.

The network architecture used in this research is illustrated in Fig. 2. This network is a modified version of the EDSR [18], which is originally designed for image super-resolution task. Since we are not aimed at improving the image resolution, the up-sampling block in the original architecture is removed. Besides, to enable the residual learning suggested by [19], a long skip connection is added between the input and output, which are the averaged multi-bit QIS frame $\bar{\mathbf{b}}$ and the estimation of its corresponding expectation $\mathbb{E}(\bar{\mathbf{b}})$, respectively. It is also worth noting that our proposed learning framework is compatible with arbitrary network architectures.

4. EXPERIMENTAL RESULTS

4.1. Experimental setting

In our synthetic experiments, QIS data are generated from the source images in publicly available datasets. During the training phase, 800 4K-resolution source images from the DIV2K [20] are converted to the gray images with the intensity range of $(0, 1)$. For each source image, only one pair of multi-bit QIS frames is synthesized. To reduce the training time spent on the data loading, these QIS frames are cropped into 22988 patches of size 480×480 as a pre-processing step. The optimizer is the Adam algorithm, with the learning rate initially set as 10^{-4} and halved every 2×10^5 training steps. There are 10^6 training steps in total. During the testing phase, 68 gray images from BSD68 [21] are employed. All experiments are performed on AMD Ryzen Threadripper PRO 3955WX CPU and NVIDIA GeForce RTX 3090 GPU.

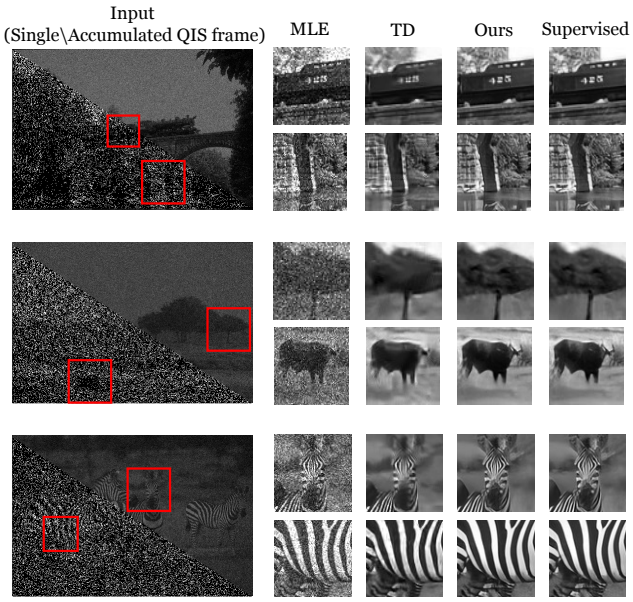

Fig. 3. Qualitative evaluation of QIS reconstruction. The reconstructed images are demonstrated in sequence using MLE [6], the Transform-Denoise method (TD) [9], and ours. Additionally, reconstruction results from supervised learning, which utilizes pairs of accumulated QIS and its corresponding clean ground truth (supervised), are included for reference. The comparison aims to provide insights into the performance and effectiveness of the unsupervised methods in the context of QIS technology. In each case, two selected region-of-interest are enlarged, providing a closer examination of the reconstructed details.

4.2. Results

For the visual comparison, several representative reconstructed methods have been provided in Fig. 3. It is worth noting that the field of QIS technology is still relatively under research, resulting in limited commercial availability and a scarcity of reconstruction methods specifically designed for QIS devices. In order to ensure a fair and unbiased evaluation, our analysis focuses exclusively on unsupervised model-based reconstruction methods that employ a single-

Table 1. Quantitative evaluation of QIS reconstruction. The arrow direction beside each metric indicates better reconstruction quality.

	Unsupervised			
Metrics	MLE [6]	TD [9]	Ours	Supervised
PSNR(dB) ↑	19.17	24.51	28.17	28.20
SSIM ↑	0.51	0.81	0.86	0.86
RMSE ↓	0.012	0.004	0.003	0.003

shot inference process. Consequently, we select two methods for our approach: MLE [6] and TD [9]. TD [9] employs a variance stabilizing transform to stabilize its variance before employing the standard Gaussian denoiser. Furthermore, we have included reconstruction results from supervised learning, which directly utilizes the clean image as the training target, as a benchmark to gauge the performance of the unsupervised methods. This choice is based on the prevailing consensus that the performance of unsupervised methods generally does not surpass that of supervised learning methods. By incorporating this comparison, we aim to provide a comprehensive assessment of the capabilities and limitations of our chosen unsupervised methods. As observed in the Fig. 3, it is evident that TD [9] exhibits noticeable improvements compared to MLE [6]. However, our proposed method goes beyond TD by introducing further enhancements, particularly in terms of sharp edges and finer details. Notably, our method achieves these improvements without relying on the availability of clean ground-truth data. It attains competitive performance levels comparable to those achieved by the supervised network, despite operating in an unsupervised manner. To conduct a quantitative comparison, we have calculated and presented the evaluation indicators, namely Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Root Mean Square Error (RMSE), for each method in Table 1. The results clearly demonstrate that our approach achieves the highest scores across all metrics compared to other unsupervised methods, while maintaining a comparable level of performance to the supervised method. Notably, there is only a slight decrease in the PSNR value ($28.17\,\mathrm{dB}$ vs $28.20\,\mathrm{dB}$). These findings highlight the efficacy of our unsupervised approach in achieving competitive results, closely approaching the performance of supervised methods, while offering distinct advantages in terms of simplicity and independence from clean ground-truth data.

To provide a comprehensive and intuitive demonstration of the comparable performance between our unsupervised method and the supervised method, we present the convergence curves of both frameworks in Fig. 4. The convergence curves depict the progress of our proposed learning framework alongside the supervised learning framework. As illustrated in Fig. 4, initially, there is a discernible gap between the two frameworks during the early stages of learning. However, it becomes evident that they eventually converge to closely aligned points. This convergence signifies that training the image reconstruction network for QIS utilizing our proposed unsupervised framework is equivalent to training in a supervised manner. This finding further reinforces the effectiveness and viability of our unsupervised methodology in achieving comparable results to the supervised approach.

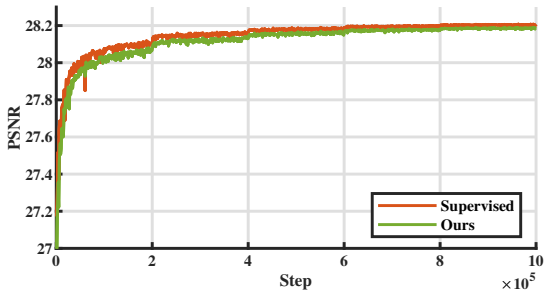

Fig. 4. PSNR (dB) on BSD68 [21] as a function of training step.

5. CONCLUSION

In this work, we have proposed an unsupervised learning framework for reconstructing images from QIS data. Our framework leverages pairs of independent and noisy QIS data as training examples to estimate the probability of photon arrivals. This approach eliminates the need for collecting labeled image data during training. Once the network is trained, the image can be reconstructed by relating the image intensity to the estimated probability. Experimental results demonstrate that our proposed method achieves performance close to that of supervised learning approaches and significantly outperforms model-based methods. This advancement contributes to the practical implementation and widespread adoption of QIS technology in various imaging applications.

6. ACKNOWLEDGEMENT

This work is supported by the Research Grants Council of Hong Kong (GRF 17201822).

7. REFERENCES

[1] Jiaju Ma, Stanley Chan, and Eric R. Fossum, "Review of quanta image sensors for ultralow-light imaging," *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2824–2839, June 2022.

[2] Eric R Fossum, Jiaju Ma, Saleh Masoodian, Leo Anzagira, and Rachel Zizza, "The quanta image sensor:

Every photon counts," *Sensors*, vol. 16, no. 8, pp. 1260, August 2016.

[3] Abhiram Gnanasambandam and Stanley H Chan, "HDR imaging with quanta image sensors: Theoretical limits and optimal reconstruction," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1571–1585, November 2020.

[4] Yash Sanghvi, Abhiram Gnanasambandam, and Stanley H Chan, "Photon limited non-blind deblurring using algorithm unrolling," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 851–864, September 2022.

[5] Yunping Zhang, Stanley H Chan, and Edmund Y Lam, "Photon-starved snapshot holography," *APL Photonics*, vol. 8, no. 5, pp. 056106, May 2023.

[6] Feng Yang, Yue M Lu, Luciano Sbaiz, and Martin Vetterli, "Bits from photons: Oversampled image acquisition using binary poisson statistics," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1421–1436, December 2011.

[7] Francis J Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, December 1948.

[8] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, July 2007.

[9] Stanley H Chan, Omar A Elgendy, and Xiran Wang, "Images from bits: Non-iterative image reconstruction for quanta image sensors," *Sensors*, vol. 16, no. 11, pp. 1961, November 2016.

[10] Joon Hee Choi, Omar A Elgendy, and Stanley H Chan, "Image reconstruction for quanta image sensors using deep neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, April 2018, pp. 6543–6547.

[11] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila, "Noise2Noise: Learning image restoration without clean data," in *International Conference on Machine Learning*. PMLR, April 2018, pp. 2965–2974.

[12] Stanley H Chan, "On the insensitivity of bit density to read noise in one-bit quanta image sensors," *arXiv preprint arXiv:2203.06086*, March 2022.

[13] Omar A Elgendy and Stanley H Chan, "Optimal threshold design for quanta image sensor," *IEEE Transactions on Computational Imaging*, vol. 4, no. 1, pp. 99–111, December 2017.

[14] Jiaju Ma and Eric R Fossum, "Quanta image sensor jot with sub 0.3 e-rms read noise and photon counting capability," *IEEE Electron Device Letters*, vol. 36, no. 9, pp. 926–928, July 2015.

[15] Zhaoyang Yin, Jiaju Ma, Saleh Masoodian, and Eric R Fossum, "Threshold uniformity improvement in 1b quanta image sensor readout circuit," *Sensors*, vol. 22, no. 7, pp. 2578, March 2022.

[16] 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*, vol. 12, no. 1, pp. 13869, August 2022.

[17] Stanley H Chan and Yue M Lu, "Efficient image reconstruction for gigapixel quantum image sensors," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, December 2014, pp. 312–316.

[18] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, July 2017, pp. 136–144.

[19] 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*, vol. 26, no. 7, pp. 3142–3155, Feburary 2017.

[20] Eirikur Agustsson and Radu Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, July 2017, pp. 126–135.

[21] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *IEEE International Conference on Computer Vision*. IEEE, July 2001, vol. 2, pp. 416–423.