Bits from Photons: Oversampled Binary Image Acquisition

THÈSE N° 5330 (2012)
PRÉSENTÉE LE 21 MARS 2012
À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE DE COMMUNICATIONS AUDIOVISUELLES
PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Feng YANG

acceptée sur proposition du jury:
Prof. E. Telatar, président du jury
Prof. M. Vetterli, Dr L. Sbaiz, directeurs de thèse
Prof. E. Charbon, rapporteur
Prof. A. El Gamal, rapporteur
Prof. N. Kingsbury, rapporteur
Contents

Abstract 1
Résumé 3
Acknowledgements 5

1 Introduction 7
  1.1 A History of Photography ................................. 7
    1.1.1 Camera Obscura ..................................... 7
    1.1.2 Chemical Photography ............................... 8
    1.1.3 Digital Photography ................................. 11
  1.2 Trends for the Design of Image Sensors ...................... 17
    1.2.1 Image Sensors with Smaller Pixel Sizes ............ 17
    1.2.2 High Dynamic Range Image Sensors ................. 17
    1.2.3 High Sensitivity Image Sensors .................... 18
  1.3 The Human Visual System .................................. 19
  1.4 Motivation ................................................ 20
  1.5 Thesis Outline and Contributions ........................ 22

2 Oversampled Binary Image Acquisition 25
  2.1 Introduction ............................................. 25
  2.2 Imaging by Oversampled Binary Sensors ..................... 26
    2.2.1 Diffraction Limit and Linear Light Field Models .... 26
    2.2.2 Sampling the Light Intensity Field ................. 28
    2.2.3 Binary Sensing and One-Bit Poisson Statistics .... 29
    2.2.4 Multiple Exposures and Temporal Oversampling .... 31
  2.3 Performance Analysis .................................... 32
    2.3.1 The Cramér-Rao Lower Bound (CRLB) of Estimation Variances 32
    2.3.2 Asymptotic Achievability of the CRLB .............. 35
    2.3.3 Advantages over Traditional Sensors .............. 36
  2.4 Optimal Image Reconstruction and Efficient Implementations 38
    2.4.1 Image Reconstruction by MLE .......................... 39
    2.4.2 Iterative Algorithm and Efficient Implementations .. 41
  2.5 Numerical Results .................................... 44
    2.5.1 1-D Synthetic Signals ............................. 44
    2.5.2 Acquiring Scenes with High Dynamic Ranges .......... 45
    2.5.3 Results on Real Sensor Data ...................... 46
2.6 Conclusions ................................................. 47
2.A Appendix .................................................. 48
  2.A.1 Proof of Proposition 1 ................................. 48
  2.A.2 The CRLB of Binary Sensors .......................... 48
  2.A.3 The CRLB of Ideal Unquantized Sensors .............. 49
  2.A.4 Proof of Proposition 3 ................................. 49
  2.A.5 Proof of Theorem 1 .................................... 51
  2.A.6 Proof of Proposition 4 ................................. 53
  2.A.7 Proof of Lemma 2 ...................................... 54
  2.A.8 Proof of Proposition 5 ................................. 54

3 Oversampled Noisy Binary Image Acquisition 57
  3.1 Introduction ............................................... 57
  3.2 Imaging by Oversampled Noisy Binary Sensors .......... 58
  3.3 Performance Analysis ..................................... 59
    3.3.1 The Cramér-Rao Lower Bound (CRLB) of the Estimation Variance 59
    3.3.2 Closed-form MLE Solution When $q = 1$ .................. 60
    3.3.3 The Influence of the Noise on the Dynamic Range .......... 61
  3.4 Image Reconstruction ..................................... 62
    3.4.1 Image Reconstruction by the MLE ...................... 62
    3.4.2 Iterative Algorithms .................................. 65
  3.5 Numerical Results ........................................ 69
    3.5.1 1-D Synthetic Signals ................................ 69
    3.5.2 2-D Synthetic Images ................................ 70
  3.6 Conclusions .............................................. 71
  3.A Appendix ................................................ 74
    3.A.1 The CRLB of Noisy Binary Sensors .................... 74
    3.A.2 Proof of Proposition 7 ............................... 75
  4 Thresholds Values and Patterns 77
  4.1 Introduction .............................................. 77
  4.2 Influence of Thresholds .................................. 78
  4.3 Optimal Threshold Pattern Design ....................... 80
  4.4 Reconstruction under Different Thresholds .............. 84
  4.5 Numerical Results ....................................... 87
    4.5.1 1-D Synthetic Signals ............................... 87
    4.5.2 2-D Synthetic Images ............................... 87
    4.5.3 Results on Real Sensor Data ......................... 87
  4.6 Conclusions .............................................. 88
  4.A Appendix ................................................ 91
    4.A.1 The CRLB of Binary Sensors with Two Thresholds ....... 91

5 Generalized Piecewise-constant Model 93
  5.1 Introduction .............................................. 93
  5.2 Estimating 1-D Piecewise-constant Light Exposure Values 94
    5.2.1 Dynamic Programming ............................... 96
    5.2.2 Greedy Algorithm .................................. 97
    5.2.3 Pruning of Binary Trees ............................ 97
5.3 Estimating 2-D Piecewise-constant Light Exposure Values 98
  5.3.1 Dynamic Programming 98
  5.3.2 Greedy Algorithm 99
  5.3.3 Pruning of Quadtrees 99
5.4 Numerical Results 100
  5.4.1 1-D Synthetic Signals 100
  5.4.2 2-D Synthetic Images 101
  5.4.3 Results on Real Sensor Data 101
5.5 Conclusion 102

6 Conclusions 105
  6.1 Thesis Summary 105
  6.2 Future Research 106

Bibliography 109

Curriculum Vitæ 117
List of Figures

1.1 Camera obscura [1]. ................................................................. 8
1.2 View from the Window at Le Gras. The first successful permanent photograph made by Niépe at Saint-Loup-de-Varennes, France, in 1826 [2]. 9
1.3 The daguerreotype camera built by La Maison Susse Frères in 1839, with a lens by Charles Chevalier [3]. ....................................................... 9
1.4 Undeveloped black and white film [4]. ................................. 10
1.5 Tartan ribbon. The first permanent color photograph, taken by Sutton under the direction of James Clerk Maxwell in 1861 [5]. .......................... 10
1.6 Microphotography of the Autochrome trichromatic selection mosaic, made of dyed potato starch grains [6]. ........................................... 11
1.7 A silicon p-n junction in reverse bias. ...................................................... 12
1.8 The CCD [7]. (a) The MOS capacitor. (b) The first CCD device. (c) The basic CCD structure. (d) The first integrated CCD device. (e) The first CCD image. (f) The buried channel CCD. ........................................... 14
1.9 Three types of CMOS pixel architectures. (a) CMOS passive pixel sensor (PPS) [8]. RS is used as row selection switch. (b) CMOS Active Pixel Sensor (APS) [8]. The transistors RST and RS are used for resetting and selecting the pixel. (c) The simple CMOS digital pixel sensor [9]. ...................... 15
1.10 Different ways to get a color image. (a) The Bayer filter on the image sensor [10]. (b) The color separation prism [11]. (c) The color separation using micro lenses and dichroic mirrors [12]. (d) The Foveon X3 image sensor [13]. (e) The color co-site sampling method [14]. ........................................... 16
1.11 Evolution of the CMOS image sensor’s pixel size, the CMOS technology node used to fabricate the devices and the minimum dimension according to the ITRS Roadmap [8]. ......................................................... 17
1.12 A schematic of the EMCCD [15]. Differently from a conventional CCD, in the EMCCD there is a multiplication register that increasing the signal through impact ionization ................................................. 19
1.13 The human visual system [16]. (a) The human eye. (b) The human retina. 20
1.14 The model of the rod-bipolar pathway in [17]. A dim light source generates a flux of photons. The photons are absorbed by the rods, and contaminated by Gaussian rod noise. The rod-bipolar sums the rod responses after a nonlinear synapse. ............................................... 21
1.15 The model of an image sensor. The image sensor formed by a lot of pixels samples the light intensity field. Each pixel occupies a small area in the space and collects photons hitting on it during the integration time, which might be different for each pixel. ............................................... 22
2.1 The imaging model. (a) The simplified architecture of a diffraction-limited imaging system. Incident light field $\lambda_0(x)$ passes through an optical lens, which acts like a linear system with a diffraction-limited point spread function (PSF). The result is a smoothed light field $\lambda(x)$, which is subsequently captured by the image sensor. (b) The PSF (Airy disk) of an ideal lens with a circular aperture. ................................................. 27

2.2 The signal processing block diagram of the imaging model studied in this chapter. In the first step, the light exposure value $s_m$ at the $m$th pixel is related to the expansion coefficients $c_n$ through a concatenation of upsampling and filtering operations. Subsequently, the image sensor converts $\{s_m\}$ into quantized measurements $\{b_m\}$ (see Figure 2.3 and the discussions in Section 2.2.3 for details of this second step). ................................................. 29

2.3 The model of the binary image sensor. The pixels (shown as “buckets”) collect photons, the numbers of which are compared against a quantization threshold $q$. In the figure, we illustrate the case when $q = 2$. The pixel outputs are binary: $b_m = 1$ (i.e., white pixels) if there are at least two photons received by the pixel; otherwise, $b_m = 0$ (i.e., gray pixels). ... 30

2.4 Performance comparisons of three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the “IDEAL” scheme with no quantization; The solid line (in blue) corresponds to the “SAT” scheme with a saturation point set at $C_{\text{max}} = 9130$ [18]; The four dashed lines (in black) correspond to the “BIN” scheme with $q = 1$ and different oversampling factors (from left to right, $K = 2^{13}$, $2^{14}$, $2^{15}$ and $2^{16}$, respectively). ................................................. 37

2.5 The likelihood and log-likelihood functions for piecewise-constant light fields. (a) The likelihood functions $L_b(c)$, defined in (2.22), under different choices of the quantization thresholds $q = 1, 3, 5$, respectively. (b) The corresponding log-likelihood functions. In computing these functions, we set the parameters in (2.22) as follows: $K = 12$, i.e., the sensor is 12-times oversampled. The binary sensor measurements contain 10 “1”s, i.e., $K_1 = 10$. ................................................. 40

2.6 Signal processing implementations of $Ga$ and $G^Tb$. (a) The product $Ga$ can be obtained by upsampling followed by filtering. (b) The product $G^Tb$ can be obtained by filtering followed by downsampling. Note that the filter used in (b) is $g_{-m}$, i.e., the “flipped” version of $g_{m}$. (c) The polyphase domain implementation of (a). (d) The polyphase domain implementation of (b). ................................................. 43

2.7 Binary sensing and reconstructions of 1-D light fields. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernels. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with spatial oversampling factor $K = 256$. (c) An improved reconstruction result due to the use of a larger spatial oversampling factor $K = 2048$. (d) An alternative result, obtained by keeping $K = 256$ but taking $J = 8$ consecutive exposures. ... 45
2.8 High dynamic range photography using the binary sensor. (a) A sequence of images taken inside of a church with decreasing exposure times [19]. (b) The reconstructed high dynamic range radiance map (in logarithmic scales) using our MLE reconstruction algorithm. (c) The tone-mapped version of the reconstructed radiance map. 46

2.9 Image reconstruction from the binary measurements taken by a SPAD sensor [20], with a spatial resolution of 32 × 32 pixels. The final image (lower-right corner) is obtained by incorporating 4096 consecutive frames, 50 of which are shown in the figure. 46

3.1 The signal processing block diagram of the imaging model studied in this chapter. We upsample and filter the expansion coefficients $c_n$ to get the light exposure value $s_m$ at the $m$th pixel. Then, the binary image sensor converts $\{s_m\}$ into quantized measurements $\{b_m\}$. Due to noise $\{w_m\}$, we get the contaminated binary measurements $\{b'_m\}$. 58

3.2 Performance comparisons of three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the “IDEAL” scheme with no quantization; The solid line corresponds to the “SAT” scheme with a saturation point set at $c = 9130$; The four dashed lines correspond to the “BIN” schemes with $q = 1$ and $K = 2^{12}$ and different noise rates (from far right to left, $p_e = 0, 0.001, 0.005$ and 0.01, respectively). 61

3.3 Performance comparisons of three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the “IDEAL” scheme with no quantization; The solid line (in blue) corresponds to the “SAT” scheme with a saturation point set at $c = 9130$; The three dashed lines correspond to the “BIN” schemes with $q = 1$ and noise rate $p_e = 0.005$ and different oversampling factors (from far right to left, $K = 2^{14}, 2^{13},$ and $2^{12}$, respectively). 62

3.4 The likelihood and log-likelihood functions for piecewise-constant light fields under the noisy case. (a) The likelihood functions $L_E^k(c)$, defined in (3.6), under different choices of noise rate $p_e = 0, 0.005$, and 0.01, respectively. (b) The corresponding log-likelihood functions. The parameters are set as: threshold $q = 1$, oversampling factor $K = 20$, number of ones $K_1 = 4$. 64

3.5 The likelihood and log-likelihood functions for piecewise-constant light fields under the noisy case. (a) The likelihood functions $L_E^k(c)$, defined in (3.6), under different choices of noise rate $p_e = 0, 0.005$, and 0.01, respectively. (b) The corresponding log-likelihood functions. The parameters are set as: threshold $q = 2$, oversampling factor $K = 20$, number of ones $K_1 = 4$. 65

3.6 The log-likelihood function for a general light field in the noisy case. (a) The log-likelihood function. (b) The contour of the log-likelihood function. The parameters are set as: threshold $q = 5$, oversampling factor $K = 10$, noise rate $p_e = 0.05$. The binary sequence is generated with expansion coefficients $c = [100, 110]^T$. 66
3.7 Efficient signal processing operations for computing the gradient and the Hessian. (a) The fast computation of the gradient $\nabla \ell_e(c)$ specified in (3.17). (b) The signal processing block diagram for computing the product between the Hessian matrix (evaluated at $c$) and an arbitrary vector $p$.  

3.8 Binary sensing and reconstructions of 1-D light fields with different noise rates. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernels. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1, \text{and} \ 0.2$ respectively. The oversampling factor is 1024, and the threshold is $q = 1$.  

3.9 Binary sensing and reconstructions of 1-D light fields under noise with different oversampling factors. (a), (b), and (c) show the reconstruction results with different oversampling factor $K = 256, 2048, \text{and} \ 8192$ respectively. The threshold is $q = 1$, and the noise rate is $p_e = 0.1$.  

3.10 Binary sensing and reconstructions of 1-D light fields with different noise levels. (a), (b), and (c) show the reconstruction results of the light intensity field $\lambda(x)$ with different noise rates $p_e = 0, 0.1, \text{and} \ 0.2$, respectively. The original light intensity field is shown in blue solid line and the reconstructed results are shown in red dash line. The oversampling factor is 512, and the threshold is $q = 3$.  

3.11 Binary sensing and reconstructions of 2-D light fields with different noise levels. (a) The original light field. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1, \text{and} \ 0.2$, respectively. The spatial oversampling factor is $8 \times 8$, the temporal oversampling factor is 128, and the threshold is $q = 1$.  

3.12 Binary sensing and reconstructions of 2-D light fields with different noise levels. (a) The original light field. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1, \text{and} \ 0.2$ respectively. The spatial oversampling factor is $8 \times 8$, the temporal oversampling factor is 128, and the threshold is $q = 3$.  

4.1 The relation between the total light exposure value $c$ and the CRLB of “BIN” schemes with different threshold $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme. (a) Oversampling factor $K = 1024$. (b) Oversampling factor $K = 8096$.  

4.2 The relation between the oversampling factor $K$ and the CRLB of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme.
4.3 The relation between the total light exposure value $c$ and the MSE of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme. (a) Oversampling factor $K = 1024$. (b) Oversampling factor $K = 8096$. 81

4.4 The relation between the total light exposure value $c$ and the SNR of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme. The oversampling factor $K = 8096$. 82

4.5 An 2-D image sensor with interleaved thresholds $q_1$ and $q_2$. 83

4.6 Optimal threshold pattern design when the oversampling factor is $K = 1024$, the maximum threshold is $q_{\text{max}} = 20$, and the total light exposure value is $c \in [10, 10^4]$. (a) The average CRLB for different threshold patterns $\{q_1, q_2\}$. (b) The comparison of the CRLB of the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$, two single-threshold patterns $\{q_1 = q_2 = 1\}$ and $\{q_1 = q_2 = 10\}$, and the “IDEAL” scheme. 84

4.7 Optimal threshold pattern design when the oversampling factor is $K = 8096$, the maximum threshold is $q_{\text{max}} = 20$, and the total light exposure value is $c \in [10^2, 10^5]$. (a) The average CRLB for different threshold patterns $\{q_1, q_2\}$. (b) The comparison of the CRLB of the optimal threshold pattern $\{q_1 = 10, q_2 = 3\}$ two single-threshold patterns $\{q_1 = q_2 = 1\}$ and $\{q_1 = q_2 = 10\}$, and the “IDEAL” scheme. 85

4.8 Binary sensing with different threshold patterns and reconstructions of 1-D light fields. The oversampling factor $K$ is equal to 1024. The number of pixels with threshold $q_1$ is 512 and the rest are with threshold $q_2$. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernel. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with the threshold pattern $\{q_1 = q_2 = 1\}$. (c) Results with the threshold pattern $\{q_1 = q_2 = 10\}$. (d) Results using the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$ as obtained in the previous section. 88

4.9 Binary sensing with different threshold patterns and reconstructions of 2-D light fields. The spatial oversampling factor is set to $8 \times 8$ and the temporal oversampling factor is 16. So the total oversampling factor is 1024. The 2-D threshold pattern is the same as in Figure 4.5. (a) The original image. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with the threshold pattern $\{q_1 = q_2 = 1\}$. (c) Results for the threshold pattern $\{q_1 = q_2 = 10\}$. (d) Results using the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$ as obtained in the previous section. 89

4.10 The SPAD camera. 90

4.11 A binary image and the estimated light intensity for thresholds: (b) $q = 1$; (c) $q = 30$; and (d) $q = 60$. 90
5.1 Reconstructions of 1-D light exposure values. The threshold is $q = 1$, the number of pixels is $M = 1024$ and the penalized parameter is $\gamma = 5$. The red solid line is the original light exposure values and the blue dash line is the reconstructed result. The reconstruction algorithms are: (a) dynamic programming; (b) greedy algorithm; (c) pruning algorithm. .................................. 100

5.2 Reconstructions of 2-D light exposure values with resolution $512 \times 512$. The threshold is $q = 1$. (a) The original 2-D light exposure values. (b) The binary image. (c) Reconstruction using the greedy algorithm. (e) Reconstruction using the pruning of quadtrees algorithm. .......................... 101

5.3 Reconstructions of 2-D real sensor data. The resolution of the binary image is $1024 \times 1024$. The resolution of reconstructed images is $128 \times 128$. (a) The SPAD camera with resolution $32 \times 32$ is fixed on a 2-D positioning system. (b) The binary image generated by stitching 1024 binary images taken by the SPAD camera. (c) Reconstruction using the greedy algorithm. (d) Reconstruction using the pruning of quadtrees algorithm. .................................................. 103

6.1 The chip micrograph of the binary image sensor designed by Prof. Edoardo Charbon and his team. .................................................. 107
Abstract

The trends in the design of image sensors are to build sensors with low noise, high sensitivity, high dynamic range, and small pixel size. How can we benefit from pixels with small size and high sensitivity? In this dissertation, we study a new image sensor that is reminiscent of traditional photographic film. Each pixel in the sensor has a binary response, giving only a one-bit quantized measurement of the local light intensity. The response function of the image sensor is non-linear and similar to a logarithmic function, which makes the sensor suitable for high dynamic range imaging.

We first formulate the oversampled binary sensing scheme as a parameter estimation problem based on quantized Poisson statistics. We show that, with a single-photon quantization threshold and large oversampling factors, the Cramér-Rao lower bound (CRLB) of the estimation variance approaches that of an ideal unquantized sensor, that is, as if there were no quantization in the sensor measurements. Furthermore, the CRLB is shown to be asymptotically achievable by the maximum likelihood estimator (MLE). By showing that the log-likelihood function is concave, we guarantee the global optimality of iterative algorithms in finding the MLE.

We study the performance of the oversampled binary sensing scheme in presence of dark current noise. The noise model is an additive Bernoulli noise with a known parameter, and the noise only flips the binary output from “0” to “1”. We show that the binary sensor is quite robust with respect to noise and its dynamic range is only slightly reduced. The binary sensor first benefits from the increasing of the oversampling factor and then suffers in term of dynamic range. We again use the MLE to estimate the light intensity. When the threshold is a single photon, we show that the log-likelihood function is still concave. Thus, the global optimality can be achieved. But for thresholds larger than “1”, this property does not hold true. By proving that when the light intensity is piecewise-constant, the likelihood function is a strictly pseudoconcave function, we guarantee to find the optimal solution of the MLE using iterative algorithms for arbitrary thresholds. For the general linear light field model, the log-likelihood function is not even quasiconcave when thresholds are larger than “1”. In this circumstance, we find an initial solution by approximating the light intensity field with a piecewise-constant model, and then we use Newton’s method to refine the estimation using the exact model.

We then examine one of the most important parameters in the binary sensor, i.e., the threshold used to generate binary values. We prove the intuitive result that large thresholds achieve better estimation performance for strong light intensities, while small thresholds work better for low light intensities. To make a binary sensor that works in a larger range of light intensities, we propose to design a threshold array containing multiple thresholds instead of a single threshold for the binary sensing. The criterion is to minimize the average CRLB which is a good approximation of the mean squared error (MSE). The performance analysis on the new binary sensor verifies the effectiveness of
our design. Again, the MLE is used for reconstructing the light intensity field from the binary measurements. By showing that the log-likelihood function is concave for arbitrary threshold arrays, we ensure that the iterative algorithms can find the optimal solution of the MLE.

Finally, we study the reconstruction problem for the binary image sensor under a generalized piecewise-constant light intensity field model, which is quite useful when parameters like oversampling factors are unknown. We directly estimate light exposure values, *i.e.*, the number of photons hitting on each pixel. We assume that the light exposure values are piecewise-constant and we use the MLE for the reconstruction. This optimization problem is solved by iteratively working out two subproblems. The first one is to find the optimal light exposure value for each segment, given the optimal segmentation of the binary measurements. The second one is to find the optimal segmentation of the binary measurements given the optimal light exposure values for each segment. Several algorithms are provided for solving this optimization problem. Dynamic programming can obtain the optimal solution for 1-D signals, but the computation is quite heavy. To reduce the burden of computation, we propose a greedy algorithm and a method based on pruning of binary trees or quadtrees.

**Keywords:** computational photography, high dynamic range imaging, digital film sensor, photo-limited imaging, Poisson statistics, quantization, diffraction-limited imaging.
Résumé

La tendance dans la conception de capteurs d'image est de construire des capteurs à faible bruit, haute sensibilité, de gamme dynamique élevée, et à pixels de petite taille. Comment pouvons-nous bénéficier de pixels de petite taille possédant une haute sensibilité? Dans cette thèse, nous étudions un nouveau capteur d'image qui n’est pas sans rappeler les films photographiques traditionnels. Chaque pixel du capteur a une réponse binaire, produisant une mesure quantifiée sur un bit de l’intensité de la lumière locale. La fonction de réponse du capteur d’image est non-linéaire et semblable à une fonction logarithmique, ce qui rend le capteur approprié pour l’imagerie à haute gamme dynamique.

Nous formulons d’abord le schéma du capteur binaire suréchantillonné paramétrique tel un problème d’estimation basé sur les statistiques de Poisson quantifiées. Nous démontrons qu’avec un seuil de quantification à photon unique et un large facteur de suréchantillonnage, la borne inférieure de Cramèr-Rao (BICR) de la variance de l’estimation approche celle d’un capteur idéal non quantifié, à savoir, c’est comme s’il n’y avait pas de quantification dans les mesures du capteur. Par ailleurs, la BICR est asymptotiquement atteinte par l’estimateur du maximum de vraisemblance (EMV). En démontrant que la fonction de log-vraisemblance est concave, nous garantissons l’optimalité globale des algorithmes itératifs pour trouver l’EMV. Puis nous étudions les performances en présence de bruit du système de capteur binaire suréchantillonné. Le modèle de bruit est un bruit additif de Bernoulli avec un paramètre connu, et celui-ci ne renverse la sortie binaire que de “0” vers “1”. Nous démontrons que le capteur binaire est assez robuste au bruit et que sa gamme dynamique est légèrement réduite à cause du bruit. Le capteur binaire bénéficie de l’augmentation du facteur de suréchantillonnage mais souffre en terme de gamme dynamique. Nous utilisons de nouveau l’EMV pour estimer l’intensité de la lumière. Lorsque le seuil est égal à “1”, nous démontrons que la fonction de log-vraisemblance est toujours concave. Ainsi, l’optimalité globale peut être atteinte. Mais pour des seuils plus grands que “1”, cette propriété n’est pas vraie. En prouvant que lorsque l’intensité lumineuse est constante par morceaux, la fonction de vraisemblance est une fonction strictement pseudoconcave, nous produisons une solution optimale de l’EMV en utilisant des algorithmes itératifs pour des seuils arbitraires. Pour un modèle général de champ de lumière linéaire, la fonction de log-vraisemblance n’est même plus quasiconcave lorsque les seuils sont plus grands que “1”. Dans ce cas, nous obtenons une solution initiale en approximant le champ d’intensité lumineuse par un modèle constant par morceaux, et ensuite nous utilisons la méthode de Newton pour obtenir une estimation affinée utilisant le bon modèle.

Nous examinons ensuite l’un des paramètres les plus importants du capteur binaire, c’est-à-dire le seuil utilisé pour générer les valeurs binaires. Nous prouvons le résultat intuitif que les seuils élevés offrent une meilleure performance d’estimation pour des in-
Résumé

tensités de lumière forte, tandis que les petits seuils fonctionnent mieux pour de faibles intensités lumineuses. Pour concevoir un capteur binaire fonctionnant dans une grande gamme d’intensités lumineuses, nous proposons de concevoir une grille de seuils contenant plusieurs seuils au lieu d’un seul. Le critère étant de minimiser la BICR moyenne qui est une bonne approximation de l’erreur quadratique moyenne (EQM). L’analyse des performances sur le nouveau capteur binaire confirme bien l’efficacité de notre design. Encore une fois, l’EMV est utilisé pour reconstruire le champ d’intensité lumineuse à partir des mesures binaires. En démontrant que la fonction log-vraisemblance est concave pour les grilles à seuil arbitraire, nous supposons que les algorithmes itératifs peuvent trouver la solution optimale de l’EMV.

Enfin, nous étudions le problème de la reconstruction pour le capteur d’image binaire pour un modèle généralisé de champ d’intensité lumineuse qui est constant par morceaux. Ce modèle est très utile lorsque des paramètres comme le facteur de suréchantillonnage sont inconnus. Nous estimons directement les valeurs d’exposition lumineuse, c’est-à-dire, le nombre de photons frappant chaque pixel. Nous supposons les valeurs d’exposition lumineuse constantes par morceaux et utilisons l’EMV pour la reconstruction. Ce problème d’optimisation est résolu itérativement en résolvant deux sous-problèmes. Le premier est de trouver la valeur optimale de l’exposition lumineuse pour chaque segment, étant donné la segmentation optimale des mesures binaires. Le second est de trouver la segmentation optimale des mesures binaires étant donné l’exposition lumineuse pour chaque segment. Plusieurs algorithmes sont fournis pour résoudre ce problème d’optimisation. La programmation dynamique peut obtenir la solution optimale pour les signaux 1-D, mais le calcul est assez lourd. Afin de réduire la charge de calcul, nous proposons un algorithme glouton et une méthode basée sur l’élagage d’arbres binaires ou quaternaires pour résoudre le problème d’optimisation.

Mots-clés: photographie computationnelle, imagerie à grande gamme dynamique, capteur de film numérique, imagerie photo-limitée, statistiques de Poisson, quantification, imagerie à diffraction limitée.
Acknowledgements

During my PhD study, I have got help and support from many people whom I really need to thank.

First of all, I would like to thank my two advisors, Prof. Martin Vetterli and Dr. Luciano Sbaiz. Thanks a million for giving me the opportunity to study at their lab. They taught me how to solve problems as a signal processing expert. They also helped me find interesting problems to work on. They are full of ideas. I really benefit from this. They provided a lot of opportunities to me, which were just more than I could imagine. I would also like to thank Prof. Yue M. Lu. He helped me a lot after Luciano left to work at Google. My research was funded by the Swiss National Science Foundation under grant numbers: 200021-116742 and 200020-132730, which I sincerely acknowledge.

I really appreciate the support from our collaborator Prof. Edoardo Charbon and his team. Thanks a lot for lending me the single-photon avalanche diodes camera and designing image sensors to demonstrate the effectiveness of my thesis work.

I would like to take the opportunity to thank my jury members: Prof. Emre Telatar, Prof. Abbas El Gamal, Prof. Nick Kingsbury, and Prof. Edoardo Charbon. Thanks a lot for spending time on reading my thesis and accepting my work. Thanks a lot for the valuable comments.

Then, I would like to thank Dr. Kari Pulli for giving me the opportunity to do an internship at his team from Nokia Research Center, Paolo Alto. Thanks a lot for giving a chance to work on the Frankencamera project and collaborate with Prof. Marc Levoy’s team at Stanford. Thanks Marius Tico, Yingen Xiong, Timo Ahonen, and Andrew Adams. It was a great fun to work with you. I would also like to thank all my friends at Nokia. Special thanks to Weidong Shi, Frank Chun Yat Li, Jun Yang, and Du Li for their friendship and collaborations.

It is lucky that our work is recognized and acquired by Rambus Inc.. I thank David G. Stork and Thomas Vogelsang who made this happen and also support me to continuously work on this interesting project.

I would also like to thank all my LCAV colleagues and friends at EPFL for their help and support. Special thanks to Prof. Sabine Süsstrunk for the collaborations, to Jocelyne Plantefol and Jacqueline Aeberhard for organizing the lab so well, to Mihailo Kolundžija for the joy you brought to our office. And last but not least, I thank my family for their unconditional love and support.
Chapter 1

Introduction

Seeing is believing. We get most of our information from our eyes. People also like to capture what they see and share information on image and video distribution websites like YouTube, Flickr, Picasa, and Facebook. Through a historical perspective, we find that inventing tools to record what we see with high fidelity has always been an important concern. Originally, people could only use their eyes as a device to capture a scene. They encoded what they saw into words and described them to other people. Through decoding these words, other people could imagine the original scene. From pre-historic times, people have also created paintings to keep and share what they see. These two ways of recording scenes are not reliable and are labor intensive. Hence, they are not the best way to share visual information. The camera, one of the most important inventions of human kind, solves the above problems. A camera usually contains two components: a lens and an image sensor. In this dissertation, we study a new kind of image sensor, namely a sensor that has a binary response. The outline of this chapter is as follows. Section 1.1 introduces the history of photography. We present the trends in the design of image sensors in Section 1.2. Section 1.3 describes the most powerful camera, the human eye. Our motivation of designing this new image sensor is presented in Section 1.4. Section 1.5 gives a more detailed outline and the contributions of this thesis.

1.1 A History of Photography

1.1.1 Camera Obscura

The history of photography dates back to the ancient Chinese and the ancient Greeks. In the 5th and 4th centuries B.C., Chinese philosopher Mo Ti [21] and Greek mathematicians Aristotle and Euclid [22] described the idea of the camera obscura as shown in Figure 1.1. The camera obscura is a box or a room with a hole. This device projects the outside scene on a screen in the box or the room through the hole. The generated image is upside-down. The brightness of the image is decided by the size of the hole. When the hole is large, the image is bright, but blurry. To get a sharper image, we need a smaller hole, but the image becomes dimmer. To solve this problem people replaced the hole with a lens. Another problem with this system is that we can not save the image. In some sense, the rest of the history of photography is just to solve this problem.
1.1.2 Chemical Photography

Monochrome Photography

In 1727, Schulze did experiments and found that silver salts, like silver chloride and silver nitrate, darkened when exposed to light [23]. Based on this discovery, in 1816, the French inventor Niépce used a paper coated with silver chloride to take pictures [23, 24]. Unfortunately, he got negative images, i.e., the part absorbed more light would appear darker. He could not create positive prints. Also, the whole paper darkened when it absorbed more light. Therefore, he could not save the images permanently. In 1826, he captured the first permanent picture, View from the Window at Le Gras, as shown in Figure 1.2, at Saint-Loup-de-Varennes, France [24]. This picture was created on a pewter plate covered by light-sensitive medium, bitumen of Judea [23]. When light hit the bitumen, it became hard. After 8-hour exposure time, he washed away the unhardened bitumen with a dissolvant. Then the hardened region of the bitumen indicated where there was strong light, and the bare pewter region showed where there...
1.1 A History of Photography

Figure 1.2: View from the Window at Le Gras. The first successful permanent photograph made by Niépce at Saint-Loup-de-Varennes, France, in 1826 [2].

Figure 1.3: The daguerreotype camera built by La Maison Susse Frères in 1839, with a lens by Charles Chevalier [3].

was weak light. Note that this is a positive image. Since the exposure time was so long, we can see that both sides of the buildings were illuminated by sunlight.

In 1839, Niépce’s collaborator Daguerre announced the first practical photographic process, the daguerreotype process [24]. In this process, he exposed the silvered side of a silvered copper plate to iodine vapor to form a light sensitive layer of silver iodide. After this plate was exposed to light for dozens of minutes, he coated the plate with mercury vapor to form a latent image. Then, he washed away the unaltered silver iodide with sodium thiosulphate to fix the image, and the image he got was a positive image with bright areas of deposited mercury, and dark regions of silver [23]. The French government bought the patent related to this process and made it public. It led to a big commercial success. One of the daguerreotype camera made at that time is shown in Figure 1.3.

The weakness of the above process is that it is difficult to make multiple copies. In 1841, Talbot invented the calotype photographic process using paper coated with silver iodide, solving this problem [24]. The image created by this process is a negative image, from which multiple positive prints can be made. But this process was not used widely due to patent issues. The first widely used process that also generated a negative image is the collodion process developed by Archer in 1851 [24]. The advantage of this process
is that it allows a photographer to make multiple copies from a single negative, and it requires only a few seconds for exposure. The disadvantage of this process is that the negative images need to be processed in no more than 10 minutes before the plate dries. Thus, this process was overtaken by the invention of Maddox in 1871 [24], the gelatin process, in which the negatives no longer need to be developed immediately. Thanks to this process, for the first time, the camera can be made small enough for handheld usage.

When Eastman introduced the first transparent photographic film and first commercial transparent film roll to replace the photographic plate in 1885 and 1889, respectively [25], the modern history of photography began. The first transparent photographic film was actually made by putting light sensitive materials on paper. After exposure, the image-bearing layer was transferred on glass, and then printed. Later, a roll film, which was plastic and made from highly flammable celluloid, was invented. After the introduction of the 35mm film by Oskar Barnack in 1925 [26], the black and white film did not evolve much anymore. An undeveloped black and white film is shown in Figure 1.4.

**Color Photography**

The world’s first permanent color photograph shown in Figure 1.5 was taken by Sutton under the direction of Maxwell in 1861 [5]. Sutton took three images of the ribbon
1.1 A History of Photography

Figure 1.6: Microphotography of the Autochrome trichromatic selection mosaic, made of dyed potato starch grains [6].

with three different filters: red, green, and blue. Then the three images were projected on a screen with the same filters they were taken. When these three images were superimposed on the screen, a color image was obtained. In 1868, Hauron patented a subtractive assembly method for color photography [24], and published it in 1869 [27].

The first practical color photography plate, the Autochrome, was introduced by the Lumière brothers in 1906 [28]. This was an additive method. To produce an Autochrome plate, they first coated a glass support with a color screen layer, i.e., potato starch grains dyed orange-red, green, and violet-blue, as shown in Figure 1.6. After that, a light sensitive layer of silver gelatin emulsion was put on top of it. When the color plate is exposed, light comes from the other side of the glass. Therefore, the light goes through the color screen layer before reaching the light sensitive layer. The color screen layer works as a color filter. The obtained negative image is then developed to a positive image by removing the exposed silver, and exposing the remaining silver halide again. Looking from the color screen layer side, we can see the color image. The main disadvantage of the technique is that we need a longer exposure time for taking pictures, and strong light conditions for viewing images.

In 1935, Mannes and Godowsky, Jr. invented the first widely used color film, the Kodachrome [29,30]. This was a subtractive method. The Kodachrome film is made by putting three layers of emulsion to a single base. Each layer is sensitive to one of three colors: red, green and blue. The processing of the Kodachrome film is quite complex for amateurs. So users usually sent their exposed film to Kodak for processing. In 2009, after a 74-year run as a photography icon, this Kodachrome color film retired [31] due to insufficient demand. This reminds us that we are now in the digital era.

1.1.3 Digital Photography

The main difference between digital photography and chemical photography is that in digital photography, light sensitive sensors like charge-coupled devices (CCDs), or a complementary metal-oxide-semiconductor (CMOS) sensors are used to capture images instead of light sensitive plates or film in chemical photography. There are many advantages of digital photography compared to chemical photography. In digital photography,
we do not need to replace the image sensor for each picture. We have more flexibility to edit images. As is clear today, digital photography has taken the place of chemical photography, since far more images are taken using digital cameras than film cameras.

**Photodetector**

The foundation of digital photography is the photoelectric effect, *i.e.*, when certain light hits matter, electrons are generated. This effect was first observed by Hertz in 1887. Einstein described the law of the photoelectric effect in 1905 [32], which won him the Nobel Prize in 1921 [33]. In his theory, light consists of photons. The energy of a photon is equal to the frequency multiplied with a constant. When the energy is larger than a threshold, the photoelectric effect is observed.

In most image sensors, the reverse biased p-n junction photodiode, shown in Figure 1.7 serves as a photodetector. The p-type semiconductor is generated by adding certain atoms to the semiconductor to create an abundance of holes. The n-type semiconductor is generated by adding certain atoms to the semiconductor to create an abundance of electrons. When p-type and n-type semiconductor are in contact with each other, holes will move from the p-type semiconductor to the n-type semiconductor and electrons will move in the reverse direction due to diffusion. Then a depletion region will form near the junction. In this region, there are no charge carriers like holes or electrons and an electric field is created. With the reverse bias, the width of the depletion region becomes larger. This just acts like a capacitor. When light hits the semiconductor, some electrons will be generated in the depletion region due to the photoelectric effect. The electrons will go to the n-type side and holes will go to the p-type side due to the electric field force. This acts like discharging the capacitor. The stronger the light intensity, the larger the photocurrent.

**CCD Sensors**

In 1969, Boyle and Smith at AT&T Bell Labs invented the CCD [7, 34]. They got the Nobel Prize in Physics for the invention of an imaging semiconductor circuit — the CCD sensor in 2009 [35]. At that time, their lab wanted to build a semiconductor bubble memory. On October 17, 1969, after a discussion lasting no more than an hour, they sketched out the basic structure of the CDD, how it worked, and its applications [7]. In 1970, Boyle and Smith described the concept of the new semiconductor device, the CCD,
1.1 A History of Photography

in [36]. The basic element of a CCD is the metal-oxide-semiconductor (MOS) capacitor, which is made by concatenating a p-type semiconductor layer, an oxide layer, and a metal layer, shown in Figure 1.8(a). When light hits this device, hole-electron pairs will be generated and electrons will be collected at the surface between the oxide and the p-type semiconductor. To solve the problem of shifting the charge, they just put many MOS capacitors closely together in a row connected to a three-phase voltage source, shown in Figure 1.8(c). The charge can be moved through changing the voltage for each MOS capacitor. Figure 1.8(b) shows the first CCD made to verify their charge transfer idea [37]. Then in [38], they built the first integrated device, an 8-bit shift register, to demonstrate the concept of the CCD and show the feasibility of using CCDs for shift registers, logic operations, and optical imaging, shown in Figure 1.8(d). They used this device to form a simple line scan image, shown in Figure 1.8(e). To solve the problem of charge transfer inefficiency and the inability to transfer all the charges from one element to the next of the original CCD, Boyle and Smith created the buried channel CCD [39], shown in Figure 1.8(f).

CMOS Sensors

Good surveys of CMOS image sensors can be found in [40, 41, 8]. The main difference between the CCD and CMOS image sensors is the readout method. In a CMOS sensor, we read the charges or the voltages of pixels by providing row and column addresses. According to the pixel architecture, there are mainly three types of CMOS image sensors: passive pixel sensor (PPS), active pixel sensor (APS), and digital pixel sensor (DPS).

The history of the PPS dates back to 1960s. Weckler and Dyck proposed the PPS in [42,43]. As shown in Figure 1.9(a), the PPS contains a photodiode and a row selection switch. After reset, the photodiode is reverse biased to a high voltage. Then during the exposure, the photodiode absorbs photons and the voltage across the photodiode is decreased. After the exposure, when we set the row selection switch on, the voltage on the photodiode is read out by a charge integrating amplifier readout circuit. The advantage of the PPS is that the fill factor is high. The disadvantage is the large noise level introduced by the bus capacitance.

To solve the noise problem, the APS concept was first suggested in [44]. The main difference between the PPS and the APS is that in the APS, each pixel has its own in-pixel amplifier. As shown in Figure 1.9(b), the pixel has a reset transistor, a source follower transistor, and a row select transistor. The source follower isolates the photodiode from the large bus capacitance.

The first DPS was proposed in [45]. From Figure 1.9(c), we can see that the pixel of the DPS contains a photodiode, an analog-to-digital converter (ADC), and a digital memory. The advantage of the DPS is: better scaling with CMOS technology, the elimination of read-related column fixed-pattern noise and column readout noise, and unlimited potential for high-speed “snap-shot” digital imaging [9]. The drawbacks of the DPS are: too many transistors in each pixel and low fill factors.

Color Sensors

Since most of the previous mentioned photodetectors do not have color preference, we need color filters or other methods to get the color information of the incident light.

The most popular way is to put a color filter array on top of the image sensor, shown in Figure 1.10(a). In 1976, Bayer invented the Bayer filter [46], also called the RGBG,
Introduction

The CCD [7]. (a) The MOS capacitor. (b) The first CCD device. (c) The basic CCD structure. (d) The first integrated CCD device. (e) The first CCD image. (f) The buried channel CCD.

Figure 1.8: The CCD [7]. (a) The MOS capacitor. (b) The first CCD device. (c) The basic CCD structure. (d) The first integrated CCD device. (e) The first CCD image. (f) The buried channel CCD.

the GRGB, or the RGGB filter. It is the replication of $2 \times 2$ unit containing 2 green, 1 red and 1 blue, shown in Figure 1.10(a). The output image of the sensor is a Bayer pattern image. Due to the spatial multiplexing of the color filter, we do not have the
1.1 A History of Photography

Figure 1.9: Three types of CMOS pixel architectures. (a) CMOS passive pixel sensor (PPS) [8]. RS is used as row selection switch. (b) CMOS Active Pixel Sensor (APS) [8]. The transistors RST and RS are used for resetting and selecting the pixel. (c) The simple CMOS digital pixel sensor [9].

R, G, and B values for each pixel. Then demosaicing algorithms have to be applied to compute the missing color information for each pixel. Other color filter arrays like the CYGM filter, the RGBE filter, and the RGBW filter were also proposed later.

Another way to generate color images is using a color separation prism, shown in Figure 1.10(b) to separate the light into three primary colors: red, green and blue, and then take a picture of each color [47] with three sensors (e.g., CCDs). The advantage of this method is that no color information is lost compared to the color filter array method. The disadvantage is that we need multiple sensors.

In 2006, Hoshuyama proposed a method based on placing a micro lens over a triplet of photoreceptors, and then using specially designed dichroic mirrors to separate the light [12]. As shown in Figure 1.10(c), the first dichroic mirror reflects red and green light and allows blue light to go through it. The first light receiving surface absorbs the blue light. The second dichroic mirror reflects green light and allows red light going through it. The second light receiving surface gets the green light. Then the red light is reflected by the third dichroic mirror and absorbed by the third light receiving surface.

Researchers also tried to use a multi-layer silicon sensor, also called the vertical color filter, to get the full resolution color information [48–53]. One of these sensors is the Foveon X3 sensor, shown in Figure 1.10(d). We can see that in this kind of sensors, the
three primary colors are absorbed at different layers.

In the color co-site sampling method, we move the sensor by exactly one pixel to capture R, G, and B values for each pixel [54], shown in Figure 1.10(e).
1.2 Trends for the Design of Image Sensors

1.2.1 Image Sensors with Smaller Pixel Sizes

The scaling effect in CMOS technology has also influenced the CMOS image sensors’ pixel sizes. In [8], Theuwissen gave an overview of the CMOS image sensor’s pixel pitch according to the data in International Electron Devices Meeting (IEDM) and International Solid-State Circuits Conference from 1993 to 2009, shown in Figure 1.11. The bottom, blue line shows the technology nodes, defined as the half-pitch of first-level interconnect dense lines, given by the ITRS. The middle, green line plots the technology nodes used to design the pixels. The top, red line gives the pixel sizes of the CMOS image sensors. From Figure 1.11, we can see that the technology nodes used for the design of image sensors are larger than the technology nodes described by the International Technology Roadmap for Semiconductors (ITRS), and the pixel sizes almost decrease at the same pace as the technology nodes. As shown in [8], the advantage of shrinking the pixels is that we can reduce pixel area, chip area, chip cost, energy to read the sensor, lens volume, camera volume, and camera weight. But when the pixel size become small, the full well capacity of the pixel will be reduced. Thus, the signal-to-noise ratio (SNR), and the dynamic range become worse [8].

1.2.2 High Dynamic Range Image Sensors

When taking a picture with a person in front of a window, we might have the problem that if we want to have a reasonable brightness of the person, the outside scene will be saturated, and if we have an acceptable brightness of the outside scene, the person will become underexposed. The reason for this problem is that due to limitation of the full well capacity of the pixels, the camera usually has a limited dynamic range which
Introduction

is defined as the ratio between the largest nonsaturated input light intensity and the smallest detectable light intensity. Many methods for solving this problem have been proposed. Spivak et al. divided all the sensors into seven categories in [55].

The first category is that of companding sensors. In this kind of sensors, the relation between the input light intensity and the output voltage is logarithmic and thus can represent a wide range of light intensities using a small voltage swing [56]. The problem of this sensor is the large fixed pattern noise (FPN), which is a general term that identifies a temporally constant lateral non-uniformity in the image sensor, and reduced sensitivity at high light intensities. The second category is that of multiple mode sensors. These sensors combine multiple modes like linear and logarithmic modes to take the advantage of both modes [57]. The problem of this method is the reduced sensitivity at high light intensities. The third category is that of capacitance adjustment sensors. During the integration time, the integrating capacitance of a pixel is increased [58]. This makes the pixel hold more photons and thus increases the dynamic range. The fourth category is that of frequency-based sensors. In these sensors, the pixel converts the light intensity to pulses and the light intensity is proportional to the frequency of the pulses [59]. The disadvantages of this sensor is that the pixel is large and power consumption is high. The fifth category is that of time-to-saturation sensors. In these sensors, the time for the pixel to become saturated is encoded [60]. A global varying reference voltage generated by the voltage ramp is connected to each pixel. When the integrated voltage exceeds this voltage, the pixel comparator fires a pulse, which is used to encode the time of the saturation. The final image is obtained by knowing the time until the first firing event and the reference voltage at the time of the event [55]. These sensors can achieve very high dynamic range. The sixth category is that of global-control-over-the-integration-time sensors. These sensors take several images with different exposure times, and they reconstruct a high dynamic range image using these images [61]. The drawback of these sensors is the limited frame rate. The last category is that of the sensors with autonomous control over the integration time. By using a conditional reset circuitry, these sensors can automatically adjust the integration times of pixels according to the light intensity [62].

1.2.3 High Sensitivity Image Sensors

Another trend for the design of image sensors is to build high sensitivity image sensors that allow us to capture low-light level images. In 1983, Madan showed the first experimental observation of avalanche multiplication in a CCD under moderate clock voltage conditions [63]. In 2001, Jerram et al. [64] and Hynecek [65] proposed the first practical devices using a multiplication register. These CCDs are also called electron-multiplying CCDs (EMCCDs). A schematic of the EMCCD is shown in Figure 1.12. Differently from a conventional CCD, in the EMCCD there is a multiplication register before the output amplifier. With this multiplication register, the signal can be boosted above the noise floor of the amplifier. The EMCCD is sensitive to a single photon and can be used as a photon counting device [66]. One example of EMCCD sensors sold by Andor Technology is with resolution $1004 \times 1002$, and pixel size $8 \times 8 \text{µm}$. A CMOS single-photon avalanche diode (SPAD) was built by Rochas et al. [67] and Carrara et al. [20].

From the above, we can see that nowadays, people can build a high sensitivity image sensor with a large resolution and a small pixel size. Along with the trend of the shrinking pixel size, we can get an oversampled photon counting device. So theoretical results on the performance of this oversampled photon counting device are quite useful for the
1.3 The Human Visual System

Although the camera history is quite long, it is still hard for us to beat the camera from nature, namely the human eye. As suggested by Suzuki, the challenge of image sensor design in the future is to exceed the human vision [68].

Figure 1.13(a) shows the cross section of a human eye. The cornea and the lens act like a camera lens. The iris is used to control the size of the pupil, which determines the amount of light hitting the retina, acting like an image sensor. After collecting photons, it will send the signal to the brain through the optical nerve. Figure 1.13(b) plots a human retina. In the retina, there are two types of photoreceptors: rods and cones. Rods are responsible for achromatic vision at low-light conditions. Evidence shows that rods can detect a single photon [69]. Each retina has about 100 to 120 million rods [70]. There are no rods in the central fovea, a small area near the center of the retina. Cones are responsible for color vision at high light levels. There are three types of cones: long- (L-), middle- (M-), and short- (S-) wavelength-sensitive cones. The retina has about 7 to 8 million cones [70] and are concentrated in the fovea. The signal detected by rods and cones are transmitted through horizontal cells and bipolar cells to the human brain. A model of the rod-bipolar pathway is given in [17], shown in Figure 1.14.

Despite the advances in image sensors, the human eye is still better than them in terms of dynamic range and noise. So it is useful to know the mechanics of the human eye in order to apply, if possible, similar techniques in the design of new sensors. The

**Figure 1.12:** A schematic of the EMCCD [15]. Differently from a conventional CCD, in the EMCCD there is a multiplication register that increasing the signal through impact ionization.
Introduction

Figure 1.13: The human visual system [16]. (a) The human eye. (b) The human retina.

The model shown in Figure 1.14 is similar to the model of photographic film, in which pixels are also binary and the light intensity information is encoded in the local density of opaque silver grains. We know that the film had been better than the digital image sensor in term of dynamic range for a long time. Thus, the digital image sensor design can benefit from the analysis of this model.

1.4 Motivation

Figure 1.15 shows the simple model of a one-dimensional image sensor. From the above, we know that the image sensor is just a device to collect photons during a given time period and on a certain area. The ideal image sensor is a sensor that can record the
1.4 Motivation

The model of the rod-bipolar pathway in [17]. A dim light source generates a flux of photons. The photons are absorbed by the rods, and contaminated by Gaussian rod noise. The rod-bipolar sums the rod responses after a nonlinear synapse.

**Figure 1.14:** The model of the rod-bipolar pathway in [17]. A dim light source generates a flux of photons. The photons are absorbed by the rods, and contaminated by Gaussian rod noise. The rod-bipolar sums the rod responses after a nonlinear synapse.

position and time stamp of a photon hitting on the sensor. Then the output is a spatial-temporal binary image with “1” indicating that there is one or more photons during a certain time and in a certain area, and “0” indicating no photon. We then try to reconstruct the light intensity field with high precision.

The emergence of the EMCCD and the CMOS SPAD ensures us to have a fine temporal resolution. The reduction of the pixel size allows to have more spatial details. But the problem with smaller pixels is that we have potentially lower SNR. From the model of the human retina in low-light conditions, we know that thresholding can reduce noise. It is also well-known that the photographic film which has binary pixel values achieves high dynamic range. Motivated by these facts, in this dissertation, we study a new digital image sensor that is reminiscent of the human retina and the photographic film. Each pixel in the sensor has a binary response, giving only a one-bit quantized measurement of the local light intensity. At the start of the exposure period, all pixels are set to “0”. A pixel is then set to “1” if the number of photons reaching it during the exposure time is at least equal to a given threshold $q$. One way to build such binary sensors is to modify the standard memory chip technology, where each memory cell is designed to be sensitive to visible light [71]. With current CMOS technology, the level of integration of such systems can exceed $10^9 \sim 10^{10}$ (i.e., 1 giga to 10 giga) pixels per chip. In this case, the corresponding pixel sizes (around 50 nm [72]) are below the diffraction limit of light, and thus the image sensor is oversampling the optical resolution of the light field. Intuitively, one can exploit this spatial redundancy to compensate for the information loss due to one-bit quantizations, as is classically done in oversampled analog-to-digital (A/D) conversions [73–76].

Building a binary sensor that emulates the photographic film process was first envisioned by Fossum [77], who coined the name “digital film sensor”. The original motivation was mainly out of technical necessity. The miniaturization of camera systems calls for the continuous shrinking of pixel sizes. At a certain point, however, the limited full well capacity (i.e., the maximum photon-electrons a pixel can hold) of small pixels
becomes a bottleneck, yielding very low (SNRs) and poor dynamic ranges. In contrast, a binary sensor whose pixels only need to detect a few photon-electrons around a small threshold \( q \) has much less requirement for full well capacities, allowing pixel sizes to shrink further. However, such small pixels with oversampled and highly quantized (typically to 1-bit) value require new reconstruction techniques, which will be studied in this thesis.

1.5 Thesis Outline and Contributions

In Chapter 2, we explain the model of our image sensor, and then we present the performance analysis of this image sensor. We show that, with a single-photon quantization threshold and large oversampling factors, the Cramér-Rao lower bound (CRLB) of the estimation variance approaches that of an ideal unquantized sensor, \( i.e. \), as if there were no quantization in the sensor measurements. Furthermore, the CRLB is shown to be asymptotically achievable by the maximum likelihood estimator (MLE). By showing that the log-likelihood function is concave, we guarantee the global optimality of iterative algorithms in finding the MLE. We also show that the new image sensor has a potential application in high dynamic range photography.

In Chapter 3, we study the model of our image sensor in the noisy case. The noise is an additive Bernoulli noise with a known parameter, and it only flips the binary output from “0” to “1”. We derive the CRLB of the estimation variance and show that due to noise, the dynamic range of the sensor is reduced slightly. With increasing oversampling factors, the dynamic range will first become higher before becoming smaller. We again use an MLE to estimate the light intensity. We show that with a single-photon threshold, the log-likelihood function is still concave. We ensure that iterative algorithms can find the optimal solution. When thresholds are larger than “1”, we show that under a piecewise-constant light intensity model, the likelihood function and the log-likelihood function for each constant light intensity are strictly pseudoconcave functions, which
1.5 Thesis Outline and Contributions

guarantee to achieve the optimal solution of the MLE using iterative algorithms. A bisection algorithm is used to solve the optimization problem. For a general linear light field model, we first solve the MLE by assuming that the light intensity field is piecewise-constant, and then use Newton’s method to get a refined estimation based on the correct model.

In Chapter 4, we study the influence of the threshold in the new image sensor. We verify that for high light intensities, a larger threshold works better in the sense of estimation performance, and for low light intensities, a smaller threshold is preferred. The dynamic range for large thresholds is worse than that for smaller thresholds of a given oversampling factor. Then, we design an optimal threshold pattern by minimizing the average CRLB. We show that the performance of the designed optimal threshold pattern is better than a single threshold scheme over a given range of light intensities. By showing that the log-likelihood function is concave under arbitrary threshold patterns, we guarantee to find the optimal solution of the MLE.

In Chapter 5, we extend our piecewise-constant light intensity field model proposed in Chapter 2 to the case of piecewise-constant light exposure values, i.e., average photons received by each pixel. Under this model, we do not need to know oversampling factors as required in the previous model. In addition, the estimation performance can be improved by taking into account the dependence of multiple light exposure values. An MLE is used for the reconstruction. The solution of the MLE can be computed by iteratively solving two subproblems. The first is to find the optimal light exposure value for each segment, given the optimal segmentation of the binary measurements. The second is to find the optimal segmentation given the estimated light exposure value for each segment. We propose several algorithms to solve this optimization problem. Dynamic programming can find the optimal solution for 1-D signals, but has a relatively high complexity. We provide a greedy algorithm and an algorithm based on pruning of binary trees or quadtrees to find a solution (not necessarily optimal) with lower complexity.

Chapter 6 concludes the dissertation and shows some future research directions.
Chapter 2

Oversampled Binary Image Acquisition

2.1 Introduction

In this chapter, we study the new digital image sensor described in Chapter 1. Each pixel in the sensor has a binary response, giving only a one-bit quantized measurement of the local light intensity. We present a theoretical analysis of the performance of the binary sensor, and propose an efficient and optimal algorithm to reconstruct images from the measurements. Our analysis and numerical simulations demonstrate that the dynamic range of the binary sensor can be orders of magnitude higher than those of a conventional image sensor, thus providing one more motivation for considering this binary sensing scheme.

Since photon arrival at each pixel can be well-approximated by a Poisson random process whose rate is determined by the local light intensity, we formulate the binary sensing and subsequent image reconstruction as a parameter estimation problem based on quantized Poisson statistics. Image estimation from Poisson statistics has been extensively studied in the past, with applications in biomedical and astrophysical imaging. Previous work in the literature has used linear models [78], multiscale models [79,80], and nonlinear piecewise smooth models [81,82] to describe the underlying images, leading to different (penalized) maximum likelihood and/or Bayesian reconstruction algorithms. The main difference between our work and previous works is that we have only access to one-bit quantized Poisson statistics. The binary quantization and spatial oversampling in the sensing scheme add interesting dimensions to the original problem. As we will show in Section 2.3, the performance of the binary sensor depends on the intricate interplay of three parameters: the average light intensity, the quantization threshold $q$, and the oversampling factor.

The binary sensing scheme studied in this chapter also bears resemblance to oversampled A/D conversion schemes with quantizations (see, e.g., [73–76]). Previous work on one-bit A/D conversions considers bandlimited signals or, in general, signals living in the range space of some overcomplete representations. The effect of quantization is often approximated by additive noise, which is then mitigated through noise shaping [73,75], or dithering [76], followed by linear reconstruction. In our work, the binary sensor measurements are modeled as one-bit quantized versions of correlated Poisson
random variables (instead of deterministic signals), and we directly solve the statistical inverse problem by using maximum likelihood estimation, without any additive noise approximation.

The rest of the chapter is organized as follows. After a precise description of the binary sensing model in Section 2.2, we present three main contributions:

1. **Estimation performance**: In Section 2.3, we analyze the performance of the proposed binary sensor in estimating a piecewise constant light intensity function. In what might be viewed as a surprising result, we show that, when the quantization threshold \( q = 1 \) and with large oversampling factors, the Cramér-Rao lower bound (CRLB) [83] of the estimation variance approaches that of unquantized Poisson intensity estimation, that is, as if there were no quantization in the sensor measurements. Furthermore, the CRLB can be asymptotically achieved by a maximum likelihood estimator (MLE), for large oversampling factors. Combined, these two results establish the feasibility of trading spatial resolutions for higher quantization bit depth.

2. **Advantage over traditional sensors**: We compare the oversampled binary sensing scheme with traditional image sensors in Section 2.3.3. Our analysis shows that, with sufficiently large oversampling factors, the new binary sensor can have higher dynamic ranges, making it particularly attractive in acquiring scenes containing both bright and dark regions.

3. **Image reconstruction**: Section 2.4 presents an MLE-based algorithm to reconstruct the light intensity field from the binary sensor measurements. As an important result in this chapter, we show that the log-likelihood function in our problem is always concave for arbitrary linear field models, thus ensuring the achievement of global optimal solutions by iterative algorithms. For numerically solving the MLE, we present a gradient method, and derive efficient implementations based on fast signal processing algorithms in the polyphase domain [84, 85]. This attention to computational efficiency is important in practice, due to extremely large spatial resolutions of the binary sensors.

Section 2.5 presents numerical results on both synthetic data and images taken by a prototype device [20]. These results verify our theoretical analysis on the binary sensing scheme, demonstrate the effectiveness of our image reconstruction algorithm, and showcase the benefit of using the new binary sensor in acquiring scenes with high dynamic ranges.

To simplify the presentation we base our discussions on a one-dimensional (1-D) sensor array, but all the results can be easily extended to the 2-D case.

### 2.2 Imaging by Oversampled Binary Sensors

#### 2.2.1 Diffraction Limit and Linear Light Field Models

In this section, we describe the binary imaging scheme studied in this chapter. Consider a simplified camera model shown in Figure 2.1(a). We denote by \( \lambda_0(x) \) the incoming light intensity field. By assuming that light intensities remain constant within a short exposure period, we model the field as only a function of the spatial variable \( x \). Without loss of generality, we assume that the dimension of the sensor array is of one spatial unit, i.e., \( 0 \leq x \leq 1 \).

After passing through the optical system, the original light field \( \lambda_0(x) \) gets filtered by the lens, which acts like a linear system with a given impulse response. Due to imperfections (e.g., aberrations) in the lens, the impulse response, a.k.a. the point
2.2 Imaging by Oversampled Binary Sensors

Figure 2.1: The imaging model. (a) The simplified architecture of a diffraction-limited imaging system. Incident light field \( \lambda_0(x) \) passes through an optical lens, which acts like a linear system with a diffraction-limited point spread function (PSF). The result is a smoothed light field \( \lambda(x) \), which is subsequently captured by the image sensor. (b) The PSF (Airy disk) of an ideal lens with a circular aperture.

spread function (PSF) of the optical system, cannot be a Dirac delta, thus, imposing a limit on the resolution of the observable light field. However, a more fundamental physical limit is due to light diffraction \[86\]. As a result, even if the lens is ideal, the PSF is still unavoidably a small blurry spot [see, for example, Figure 2.1(b)]. In optics, such diffraction-limited spot is often called the Airy disk \[86\], whose radius \( R_a \) can be computed as

\[
R_a = 1.22w f,
\]

where \( w \) is the wavelength of the light and \( f \) is the F-number of the optical system.

**Example 1.** At wavelength \( w = 420 \text{ nm} \) (i.e., for blue visible light) and \( f = 2.8 \), the radius of the Airy disk is 1.43 \( \mu m \). Two objects with distance smaller than \( R_a \) cannot be clearly separated by the imaging system as their Airy disks on the image sensor start blurring together. Current CMOS technology can already make standard pixels smaller than \( R_a \), reaching sizes ranging from 0.5 \( \mu m \) to 0.7 \( \mu m \) [87]. In the case of binary sensors, the simplicity of each pixel allows the feature size to be further reduced. For example, based on standard memory technology, each memory bit-cell (i.e., pixel) can have sizes around 50 nm [72], making it possible to substantially oversample the light field.

In what follows, we denote by \( \lambda(x) \) the diffraction-limited (i.e., “observable”) light intensity field, which is the outcome of passing the original light field \( \lambda_0(x) \) through the lens. Due to the lowpass (smoothing) nature of the PSF, the resulting \( \lambda(x) \) has a finite spatial-resolution, i.e., it has a finite number of degrees of freedom per unit space.

**Definition 1** (Linear field model). *In this chapter, we model the diffraction-limited light intensity field as

\[
\lambda(x) = \frac{N}{\tau} \sum_{n=0}^{N-1} c_n \varphi(Nx - n),
\]

where \( \varphi(x) \) is a nonnegative interpolation kernel, \( N \) is a given integer, \( \tau \) is the exposure time, and \( \{c_n : c_n \geq 0\} \) is a set of free variables.*
Remark. The constant \(N/\tau\) in front of the summation is not essential, but its inclusion here leads to simpler expressions in our later analysis.

The function \(\lambda(x)\) as defined in (2.1) has \(N\) degrees of freedom. To guarantee that the resulting light fields are physically meaningful, we require both the interpolation kernel \(\varphi(x)\) and the expansion coefficients \(\{c_n\}\) to be nonnegative. Some examples of the interpolation kernels \(\varphi(x)\) include the box function,

\[
\beta(x) \overset{\text{def}}{=} \begin{cases} 1, & \text{if } 0 \leq x \leq 1; \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}
\]
cardinal B-splines [88],

\[
\beta_k(x) = \left( \beta \ast \ldots \ast \beta \right) \left( x + \frac{k}{2} \right), \tag{2.3}
\]
and the squared sinc function, \(\sin^2 \left( \pi \left( x - \frac{1}{T} \right) \right) / \left( \pi \left( x - \frac{1}{T} \right) \right)^2\).

2.2.2 Sampling the Light Intensity Field

The image sensor in Figure 2.1(a) works as a sampling device of the light intensity field \(\lambda(x)\). Suppose that the sensor consists of \(M\) pixels per unit space, and that the \(m\)th pixel covers the area between \([m/M, (m+1)/M]\), for \(0 \leq m < M\). We denote by \(s_m\) the total light exposure accumulated on the surface area of the \(m\)th pixel within an exposure time period \([0, \tau]\). Then,

\[
s_m \overset{\text{def}}{=} \int_0^\tau \int_{m/M}^{(m+1)/M} \lambda(x) \, dx \, dt = \tau \langle \lambda(x), \beta(Mx - m) \rangle, \tag{2.4}
\]
where \(\beta(x)\) is the box function defined in (2.2) and \(\langle \cdot, \cdot \rangle\) represents the standard \(L^2\)-inner product. Substitute the light field model (2.1) into the above equality,

\[
s_m = \tau \left( \frac{N}{\tau} \sum_n c_n \varphi(Nx - n), \beta(Mx - m) \right)
= \sum_n c_n \langle N \varphi(Nx - n), \beta(Mx - m) \rangle
= \sum_n c_n \langle \varphi(x), \beta \left( \frac{M(x + n)}{N} - m \right) \rangle, \tag{2.5}
\]
where (2.5) is obtained through a change of variables \((Nx - n) \to x\).

**Definition 2.** The spatial oversampling factor, denoted by \(K\), is the ratio between the number of pixels per unit space and the number of degrees of freedom needed to specify the light field \(\lambda(x)\) in (2.1), i.e.,

\[
K \overset{\text{def}}{=} \frac{M}{N}. \tag{2.6}
\]

In this chapter, we are interested in the “oversampled” case where \(K > 1\). Furthermore, we assume that \(K\) is an integer for simplicity of notation. Using (2.6), and by introducing a discrete filter

\[
g_m \overset{\text{def}}{=} \langle \varphi(x), \beta(Kx - m) \rangle, \quad m \in \mathbb{Z}, \tag{2.7}
\]
2.2 Imaging by Oversampled Binary Sensors

Photon Counting

Binary sensing

Light exposure

\( c_n \)

\( s_m \)

\( y_m \)

\( b_m \)

\( K \)

\( g_m \)

\( m \)

\( n \)

\( k \)

\( s_m \)

\( y_m \)

\( b_m \)

\( y_m \)

\( \sum \)

\( c_n g_{m-Kn} \)

\( s_m = \sum c_n g_{m-Kn} \)

\( (2.8) \)

\( S(z) = \sum c_n z^{-Kn} G(z) = C(z^K) G(z) \)

\( (2.9) \)

\[ g_m = \begin{cases} 1/K, & \text{for } 0 \leq m < K; \\ 0, & \text{otherwise}. \end{cases} \]

\( (2.10) \)

2.2.3 Binary Sensing and One-Bit Poisson Statistics

Figure 2.3 illustrates the binary sensor model. Recall from (2.4) that \( \{s_m\} \) denote the exposure values accumulated by the sensor pixels. Depending on the local values of \( \{s_m\} \), each pixel (depicted as “buckets” in the figure) collects a different number of photons hitting on its surface. In what follows, we denote by \( y_m \) the number of photons impinging on the surface of the \( m \)th pixel during an exposure period \([0, \tau]\). The relation between \( s_m \) and the photon count \( y_m \) is stochastic. More specifically, \( y_m \) can be modeled as realizations of a Poisson random variable \( Y_m \), whose intensity parameter is equal to \( s_m \), i.e.,

\[ \mathbb{P}(Y_m = y_m; s_m) = \frac{s_m^{y_m} e^{-s_m}}{y_m!}, \quad \text{for } y_m \in \mathbb{Z}^+ \cup \{0\}. \]

\( (2.11) \)
It is a well-known property of the Poisson process that $E[Y_m] = s_m$. Thus, the average number of photons captured by a given pixel is equal to the local light exposure $s_m$.

As a photosensitive device, each pixel in the image sensor converts photons to electrical signals, whose amplitude is proportional to the number of photons impinging on that pixel. In a conventional sensor design, the analog electrical signals are then quantized by an A/D converter into 8 to 14 bits (usually the more bits the better). In this chapter, we study a new sensor design using the following binary (i.e., one-bit) quantization scheme.

**Definition 3** (Binary Quantization). Let $q \geq 1$ be an integer threshold. A binary quantizer is a mapping $Q: \mathbb{Z}^+ \cup \{0\} \rightarrow \{0, 1\}$, such that

$$Q(y) = \begin{cases} 
1, & \text{if } y \geq q; \\
0, & \text{otherwise.}
\end{cases}$$

In Figure 2.3, we illustrate the binary quantization scheme. White pixels in the figure show $Q(y_m) = 1$ and gray pixels show $Q(y_m) = 0$. We denote by $b_m \overset{\text{def}}{=} Q(y_m)$, $b_m \in \{0, 1\}$, the quantized output of the $m$th pixel. Since the photon counts $\{y_m\}$ are drawn from random variables $\{Y_m\}$, so are the binary sensor output $\{b_m\}$, from the

---

\footnote{The exact ratio between these two quantities is determined by the quantum efficiency of the sensor.}
random variables \( \{ B_m \equiv Q(Y_m) \} \). Introducing two functions

\[
p_0(s) \equiv \sum_{k=0}^{q-1} \frac{s^k}{k!} e^{-s}, \quad p_1(s) \equiv 1 - \sum_{k=0}^{q-1} \frac{s^k}{k!} e^{-s}
\]

we can write

\[
P(B_m = b_m; s_m) = p_{b_m}(s_m), \quad b_m \in \{0, 1\}.
\]

Remark. The noise model considered in this chapter is that of Poisson noise. In practice, the performance of image sensors is also influenced by thermal noise, which in our case can be modeled as random bit-flipping in the binary sensor measurements. The effect of this additional noise source and its impact on reconstruction performance is discussed in the following chapter.

### 2.2.4 Multiple Exposures and Temporal Oversampling

Our previous discussions focus on the case of acquiring a single frame of quantized measurements during the exposure time \([0, \tau]\). As an extension, we can consider multiple exposures and acquire \( J \) consecutive and independent frames. The exposure time for each frame is set to \( \tau/J \), so that the total acquisition time remains the same as the single exposure case. In what follows, we call \( J \) the **temporal oversampling factor**.

As before, we assume that \( \tau \ll 1 \) and thus light intensities \( \lambda(x) \) stay constant within the entire acquisition time \([0, \tau]\). For the \( j \)th frame \((0 \leq j < J)\), we denote by \( s_{j,m} \) the light exposure at the \( m \)th pixel. Following the same derivations as in Section 2.2.2, we can show that

\[
s_{j,m} = \frac{1}{J} \sum_n c_n g_{m-K_n}, \quad \text{for all } j,
\]

where \( \{c_n\} \) are the expansion coefficients of the light field \( \lambda(x) \), and \( g_m \) is the discrete filter defined in (2.7). The only difference between (2.14) and (2.8) is the extra factor of \( 1/J \), due to the change of exposure time from \( \tau \) to \( \tau/J \). In the \( z \)-domain, similar to (2.9),

\[
S_j(z) = \frac{1}{J} C(z^K) G(z).
\]

In what follows, we establish an equivalence between temporal oversampling and spatial oversampling. More precisely, we will show that an \( M \)-pixel sensor taking \( J \) independent exposures (i.e., with \( J \)-times oversampling in time) is mathematically equivalent to a single sensor consisting of \( MJ \) pixels.

First, we introduce a new sequence \( \tilde{s}_m, 0 \leq m < MJ \), constructed by interlacing the \( J \) exposure sequences \( \{s_{j,m}\} \). For example, when \( J = 2 \), the new sequence is

\[
\begin{align*}
s_{0,0}, & \quad s_{1,0}, \quad s_{0,1}, \quad s_{1,1}, \ldots, \quad s_{0,m}, \quad s_{1,m}, \ldots, \quad s_{0,M-1}, \quad s_{1,M-1},
\end{align*}
\]

where \( \{s_{0,m}\} \) and \( \{s_{1,m}\} \) alternate. In general, \( \tilde{s}_m \) can be obtained as

\[
\tilde{s}_m \equiv s_{j,m}, \quad m = Jn + j, \quad 0 \leq j < J, \quad 0 \leq n < M.
\]

In multirate signal processing, the above construction is called the **polyphase representation** [84,85], and its alternating subsequences \( \{s_{j,m}\}_{j=0}^{J-1} \) the polyphase components.
Proposition 1. Let \( \tilde{g}_m \) be a filter whose \( z \)-transform

\[
\tilde{G}(z) \overset{\text{def}}{=} G(z^J)(1 + z^{-1} + \ldots + z^{-(J-1)})/J,
\]

where \( G(z) \) is the \( z \)-transform of the filter \( g_m \) defined in (2.7). Then,

\[
\tilde{s}_m = \sum_n c_n \tilde{g}_{m-KJn}.
\]

Proof. See Appendix 2.A.1.

Remark. Proposition 1 formally establishes the equivalence between spatial and temporal oversampling. We note that (2.18) has exactly the same form as (2.8), and thus the mapping from \( \{c_n\} \) to \( \{\tilde{s}_m\} \) can be implemented by the same signal processing operations shown in Figure 2.2—we only need to change the upsampling factor from \( K \) to \( KJ \) and the filter from \( g_m \) to \( \tilde{g}_m \). In essence, by taking \( J \) consecutive exposures with an \( M \)-pixel sensor, we get the same light exposure values \( \{\tilde{s}_m\} \), as if we had used a more densely packed sensor with \( MJ \) pixels.

Remark. Taking multiple exposures is a very effective way to increase the total oversampling factor of the binary sensing scheme. The key assumption in our analysis is that, during the \( J \) consecutive exposures, the light field remains constant over time. To make sure this assumption holds for arbitrary values of \( J \), we set the exposure time for each frame to \( \tau/J \), for a fixed and small \( \tau \). Consequently, the maximum temporal oversampling factor we can achieve in practice will be limited by the readout speed of the binary sensor.

Thanks to the equivalence between spatial and temporal oversampling, we only need to focus on the single exposure case in our following discussions on the performance of the binary sensor and image reconstruction algorithms. All the results we obtain extend directly to the multiple exposure case.

2.3 Performance Analysis

In this section, we study the performance of the binary image sensor in estimating light intensity information, analyze the influence of the quantization threshold and oversampling factors, and demonstrate the new sensor’s advantage over traditional sensors in terms of higher dynamic ranges. In our analysis, we assume that the light field is piece-wise constant, i.e., the interpolation kernel \( \varphi(x) \) in (2.1) is the box function \( \beta(x) \). This simplifying assumption allows us to derive closed-form expressions for several important performance measures of interest. Numerical results in Section 2.5 suggest that the results and conclusions we obtain in this section applies to the general linear field model in (2.1) with different interpolation kernels.

2.3.1 The Cramér-Rao Lower Bound (CRLB) of Estimation Variances

From Definition 1, reconstructing the light intensity field \( \lambda(x) \) boils down to estimating the unknown deterministic parameters \( \{c_n\} \). Input to our estimation problem is a sequence of binary sensor measurements \( \{b_m\} \), which are realizations of Bernoulli random variables \( \{B_m\} \). The probability distributions of \( \{B_m\} \) depend on the light exposure
values \( \{s_m\} \), as in (2.13). Finally, the exposure values \( \{s_m\} \) are linked to the light intensity parameters \( \{c_n\} \) in the form of (2.8).

Assume that the light field \( \lambda(x) \) is piecewise constant. We have computed in Example 2 that, under this case, the discrete filter \( g_m \) used in (2.8) is a constant supported within \([0, K-1]\), as in (2.10). The mapping (2.8) between \( \{c_n\} \) and \( \{s_m\} \) can now be simplified as

\[
s_m = c_n / K, \quad \text{for} \ nK \leq m < (n+1)K. \tag{2.19}
\]

We see that the parameters \( \{c_n\} \) have disjoint regions of influence, meaning, \( c_0 \) can only be sensed by a group of pixels \( \{s_0, \ldots, s_{K-1}\} \), \( c_1 \) by \( \{s_K, \ldots, s_{2K-1}\} \), and so on. Consequently, the parameters \( \{c_n\} \) can be estimated one-by-one, independently of each other.

In what follows, and without loss of generality, we focus on estimating \( c_0 \) from the block of binary measurements \( b \overset{\text{def}}{=} [b_0, \ldots, b_{K-1}]^T \). For notational simplicity, we will drop the subscript in \( c_0 \) and use \( c \) instead. To analyze the performance of the binary sensing scheme, we first compute the CRLB [83], which provides a theoretical lower bound on the variance of any unbiased estimator.

Denote by \( L_b(c) \) the likelihood function of observing \( K \) binary sensor measurement \( b \). Then,

\[
L_b(c) \overset{\text{def}}{=} P(B_m = b_m, 0 \leq m < K; c),
\]

\[
= \prod_{m=0}^{K-1} P(B_m = b_m; c), \tag{2.20}
\]

\[
= \prod_{m=0}^{K-1} p_{b_m}(c/K), \tag{2.21}
\]

where (2.20) is due to the independence of the photon counting processes at different pixel locations, and (2.21) follows from (2.13) and (2.19). Defining \( K_1 \) \((0 \leq K_1 < K)\) to be the number of “1”s in the binary sequence \( b \), we can simplify (2.21) as

\[
L_b(c) = (p_1(c/K))^{K_1} (p_0(c/K))^{K-K_1}. \tag{2.22}
\]

**Proposition 2.** The CRLB of estimating the light intensity \( c \) from \( K \) binary sensor measurements with threshold \( q \geq 1 \) is

\[
\text{CRLB}_{\text{bin}}(K, q) = c \left( \sum_{j=0}^{q-1} \frac{(q-1)!(c/K)^{-j}}{(q-1-j)!} \right) \left( \sum_{j=0}^{\infty} \frac{(q-1)!(c/K)^{j}}{(q+j)!} \right), \quad \text{for} \ c > 0. \tag{2.23}
\]

**Proof.** See Appendix 2.A.2. \( \square \)

It will be interesting to compare the performance of our binary image sensor with that of an ideal sensor which does not use quantization at all. To that end, consider the same situation as before, where we use \( K \) pixels to observe a constant light intensity value \( c \). The light exposure \( s_m \) at each pixel is equal to \( c/K \), as in (2.19). Now, unlike the binary sensor which only takes one-bit measurements, consider an ideal sensor that can perfectly record the number of photon arrivals at each pixel. By referring to Figure 2.3,
the sensor measurements in this case will be \( \{y_m\} \), whose probability distributions are given in (2.11).

In Appendix 2.A.3, we compute the CRLB of this unquantized sensing scheme as

\[
\text{CRLB}_{\text{ideal}}(K) = c,
\]

which is natural and reflects the fact that the variance of a Poisson random variable is equal to its mean (i.e., \( c \), in our case).

To be sure, we always have \( \text{CRLB}_{\text{bin}}(K, q) > \text{CRLB}_{\text{ideal}}(K) \), for arbitrary oversampling factor \( K \) and quantization threshold \( q \). This is not surprising, as we lose information by one-bit quantizations. In practice, the ratio between the two CRLBs provides a measure of performance degradations incurred by the binary sensors. What is surprising is that the two quantities can be made arbitrarily close, when \( q = 1 \) and \( K \) is large, as shown by the following proposition.

**Proposition 3.** For \( q = 1 \),

\[
\text{CRLB}_{\text{bin}}(K, q) = c + \frac{c^2}{2K} + \mathcal{O}\left(\frac{1}{K^2}\right),
\]

which converges to \( \text{CRLB}_{\text{ideal}}(K) \) as the oversampling factor \( K \) goes to infinity. For \( q \geq 2 \),

\[
\frac{\text{CRLB}_{\text{bin}}(K, q)}{\text{CRLB}_{\text{ideal}}(K)} > 1.31,
\]

and \( \lim_{K \to \infty} \frac{\text{CRLB}_{\text{bin}}(K, q)}{\text{CRLB}_{\text{ideal}}(K)} = \infty. \)

**Proof.** Specializing the expression (2.23) for \( q = 1 \), we get

\[
\text{CRLB}_{\text{bin}}(K, 1) = c \left(1 + \frac{c}{2K} + \frac{c^2}{3!K^2} + \frac{c^3}{4!K^3} + \ldots\right),
\]

and thus (2.25). The statements for cases when \( q \geq 2 \) are shown in Appendix 2.A.4. \( \square \)

Proposition 3 indicates that it is feasible to use oversampling to compensate for information loss due to binary quantizations. It follows from (2.25) that, with large oversampling factors, the binary sensor operates as if there were no quantization in its measurements. It is also important to note that this desirable tradeoff between spatial resolution and estimation variance only works for a single-photon threshold (i.e., \( q = 1 \)). For other choices of the quantization threshold, the “gap” between \( \text{CRLB}_{\text{bin}}(K, q) \) and \( \text{CRLB}_{\text{ideal}}(K) \), measured in terms of their ratio, cannot be made arbitrarily small, as shown in (2.26). In fact, it quickly tends to infinity as the oversampling factor \( K \) increases.

The results of Proposition 3 can be intuitively understood as follows: The expected number of photons collected by each pixel during light exposure is equal to \( s_m = c/K \). As the oversampling factor \( K \) goes to infinity, the mean value of the Poisson distribution tends to zero. Consequently, most pixels on the sensor will only get zero or one photon, with the probability of receiving two or more photons at a pixel close to zero. In this case, with high probability, a binary quantization scheme with threshold \( q = 1 \) does not lose information. In contrast, if \( q \geq 2 \), the binary sensor measurements will be almost uniformly zero, making it nearly impossible to differentiate between different light intensities.
2.3 Performance Analysis

2.3.2 Asymptotic Achievability of the CRLB

In what follows, we show that, when \( q = 1 \), the CRLB derived in (2.23) can be asymptotically achieved by a simple maximum likelihood estimator (MLE). Given a sequence of \( K \) binary measurements \( b \), the MLE we seek is the parameter that maximizes the likelihood function \( L_b(c) \) in (2.22). More specifically,

\[
\hat{c}_{ML}(b) = \arg \max_{0 \leq c \leq S} L_b(c) = \arg \max_{0 \leq c \leq S} (1 - p_0(c/K))^{K_1} (p_0(c/K))^{K - K_1}, \tag{2.27}
\]

where we substitute \( p_1(c/K) \) in (2.22) by its equivalent form \( 1 - p_0(c/K) \). The lower bound of the search domain \( c \geq 0 \) is chosen according to physical constraints, i.e., the light field can not take negative values. The upper bound \( c \leq S \) becomes necessary when \( K_1 = K \), in which case the likelihood function \( L_b(c) = p_1(c/K)^K \) is monotonically increasing with respect to the light intensity level \( c \).

**Lemma 1.** The MLE solution to (2.27) is

\[
\hat{c}_{ML}(b) = \begin{cases} 
K \cdot p_0^{-1}(1 - K_1/K), & \text{if } 0 \leq K_1 \leq K(1 - p_0(S/K)), \\
S, & \text{otherwise},
\end{cases} \tag{2.28}
\]

where \( p_0^{-1}(x) \) is the inverse function of \( p_0(x) \).

**Remark.** From the definition in (2.12), we can easily verify that \( \frac{d}{dx} p_0(x) < 0 \) for all \( x > 0 \). It follows that the function \( p_0(x) \) is strictly decreasing for \( x > 0 \) and that the inverse \( p_0^{-1}(x) \) is well-defined. For example, when \( q = 1 \), we have \( p_0(x) = e^{-x} \) and thus \( p_0^{-1}(x) = -\log(x) \). In this particular case, and for \( K_1 \ll K \), we have \( \hat{c}_{ML}(b) = -K \log(1 - K_1/K) \approx K_1 \). It follows that we can use the sum of the \( K \) binary measurements as a first-order approximation of the light intensity estimation.

**Proof.** At the two extreme cases, when \( K_1 = 0 \) or \( K_1 = K \), it is easy to see that (2.28) is indeed the solution to (2.27). Next, we assume that \( 0 < K_1 < K \).

Computing the derivative of \( L_b(c) \) and setting it to zero, we can verify that the equation \( \frac{d}{dc} L_b(c) = 0 \) has a single solution at

\[
\hat{c}_{\max} = K \cdot p_0^{-1}(1 - K_1/K).
\]

Since \( L_b(c) \geq 0 \) and \( L_b(0) = \lim_{c \to \infty} L_b(c) = 0 \), we conclude that the likelihood function \( L_b(c) \) achieves its maximum value at \( \hat{c}_{\max} \). Finally, the MLE solution \( \hat{c}_{ML} = \min \{ \hat{c}_{\max}, S \} \), and thus, we have (2.28).

**Theorem 1.** When \( q = 1 \), we have

\[
E[\hat{c}_{ML}(b)] = c + \varepsilon_1 + \mathcal{O}(1/K), \quad \text{for } c < S - 2, \tag{2.29}
\]

where \( |\varepsilon_1| \leq 2c e^{1-c} \left( \frac{e}{S-2} \right)^{S-1} \). Meanwhile, the mean squared error (MSE) of the estimator approaches \( \text{CRLB}_{\text{ideal}} \), i.e.,

\[
E[(\hat{c}_{ML}(b) - c)^2] = c + \varepsilon_2 + \mathcal{O}(1/K), \quad \text{for } c < S - 2, \tag{2.30}
\]

where \( |\varepsilon_2| \leq 2c(c+1)e^{1-c} \left( \frac{e}{S-2} \right)^{S-2} \).
Remark. It is easy to verify that, for fixed $c$, the two terms $\varepsilon_1$ and $\varepsilon_2$ converge (very quickly) to 0 as $S$ tends to infinity. It then follows from (2.29) and (2.30) that the MLE is asymptotically unbiased and efficient, in the sense that

$$\lim_{S \to \infty} \lim_{K \to \infty} \mathbb{E}[\hat{c}_{\text{ML}}(b)] = c \quad \text{and} \quad \lim_{S \to \infty} \lim_{K \to \infty} \mathbb{E}[(\hat{c}_{\text{ML}}(b) - c)^2] = c.$$

We leave the formal proof of this theorem to Appendix 2.A.5. Its main idea can be summarized as follows. As $K$ goes to infinity, the area of each pixel tends to zero, so does the average number of photons arriving at that pixel. As a result, most pixels on the sensor will get only zero or one photon during exposure. A single-photon binary quantization scheme can record perfectly the patterns of “0”s and “1”s on the sensor. It loses information only when a pixel receives two or more photons, but the probability of such events tends to zero as $K$ increases.

Suppose, now, that we use a quantization threshold $q \geq 2$. In this case, as $K$ tends to infinity, the binary responses of different pixels will almost always be “0”, essentially obfuscating the actual light intensity values. This problem leads to poor performance in the MLE. As stated in the following proposition, the asymptotic MSE for $q \geq 2$ becomes $c^2$ instead of $c$.

**Proposition 4.** When $q \geq 2$, the MLE is asymptotically biased, that is, for any fixed $c$ and $S$,

$$\lim_{K \to \infty} \mathbb{E}[\hat{c}_{\text{ML}}(b)] = 0.$$  \hspace{1cm} (2.31)

Meanwhile, the MSE becomes

$$\lim_{K \to \infty} \mathbb{E}[(\hat{c}_{\text{ML}}(b) - c)^2] = c^2.$$  \hspace{1cm} (2.32)


### 2.3.3 Advantages over Traditional Sensors

In what follows, we demonstrate the advantage of the oversampled binary sensing scheme, denoted by “BIN”, in achieving higher dynamic ranges. We focus on the case where the quantization threshold is set to $q = 1$. For comparisons, we also consider the following two alternative sensing schemes: The first, denoted by “IDEAL”, uses a single pixel to estimate the light exposure parameter (i.e., nonoversampled), but that pixel can record perfectly the number of photon arrivals during exposure; The second scheme, denoted by “SAT”, is very similar to the first, with the addition of a saturation point $C_{\text{max}}$, beyond which the pixel can hold no more photons. Note that in our discussions, the “SAT” scheme serves as an idealized model of conventional image sensors, for which the saturation is caused by the limited full well capacity of the semiconductor device. The general trend of conventional image sensor design has been to pack more pixels per chip by reducing pixel sizes, leading to lower full well capacities and thus lower saturation values.

Figure 2.4 compares performances of the three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values. We measure the performances in terms of signal-to-noise ratios (SNRs), defined as

$$\text{SNR} = 10 \log_{10} \frac{c^2}{\mathbb{E}[(\hat{c} - c)^2]},$$  \hspace{1cm} (2.33)
2.3 Performance Analysis

Figure 2.4: Performance comparisons of three different sensing schemes ("BIN", "IDEAL", and "SAT") over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the "IDEAL" scheme with no quantization; The solid line (in blue) corresponds to the “SAT” scheme with a saturation point set at $C_{\text{max}} = 9130$ [18]; The four dashed lines (in black) correspond to the “BIN” scheme with $q = 1$ and different oversampling factors (from left to right, $K = 2^{13}$, $2^{14}$, $2^{15}$ and $2^{16}$, respectively).

where $\hat{c}$ is the estimation of the light exposure value we obtain from each of the sensing schemes.

We observe that the “IDEAL” scheme (the red dash-dot line in the figure) represents an upper-bound of the estimation performance. To see this, denote by $y$ the number of photons that arrive at the pixel during exposure. Then $y$ is a realization of a Poisson random variable $Y$ with intensity equal to the light exposure value $c$, i.e.,

$$\mathbb{P}(Y = y; c) = \frac{c^y e^{-c}}{y!}.$$  

Maximizing this function over $c$, we can compute the MLE for the “IDEAL” scheme as $\hat{c}_{\text{IDEAL}}(y) = y$. It is easy to verify that this estimator is unbiased, i.e., $\mathbb{E}[\hat{c}_{\text{IDEAL}}(Y)] = \mathbb{E}[Y] = c$, and that it achieves the ideal CRLB in (2.24), i.e., $\text{var}(\hat{c}_{\text{IDEAL}}(Y)) = \text{var}(Y) = c$. Accordingly, we can compute the SNR as

$$\text{SNR}_{\text{IDEAL}} = 10 \log_{10}(c^2/c) = 10 \log_{10}(c),$$

which appears as a straight line in our figure with the light exposure values $c$ shown in logarithmic scale.

The solid line in the figure corresponds to the “SAT” scheme, with a saturation point set at $C_{\text{max}} = 9130$, which is the full well capacity of the image sensor in [18]. The sensor measurement in this case is $y_{\text{SAT}} = \min\{y, C_{\text{max}}\}$, and the estimator we use is

$$\hat{c}_{\text{SAT}}(y_{\text{SAT}}) = y_{\text{SAT}}.$$  

(2.34)
We can see that the “SAT” scheme initially has the same performance as “IDEAL”. It remains this way until the light exposure value \( c \) approaches the saturation point \( C_{\text{max}} \), after which there is a drastic drop\(^2\) in SNR. Denoting by \( \text{SNR}_{\text{min}} \) the minimum acceptable SNR in a given application, we can then define the dynamic range of a sensor as the range of \( c \) for which the sensor achieves at least \( \text{SNR}_{\text{min}} \). For example, if we choose \( \text{SNR}_{\text{min}} = 20 \) dB, then, as shown in the figure, the SAT scheme has a dynamic range from \( c = 10^2 \) to \( c \approx 10^4 \), or, if measured in terms of ratios, \( 100 : 1 \).

Finally, the three dashed lines represent the “BIN” scheme with \( q = 1 \) and increasing oversampling factors (from left to right: \( K = 2^{13}, 2^{14}, 2^{15} \) and \( 2^{16} \), respectively). We use the MLE given in (2.28) and plot the corresponding estimation SNRs. We see that, within a large range of \( c \), the performance of the “BIN” scheme is very close to that of the “IDEAL” scheme that does not use quantization. This verifies our analysis in Theorem 1, which states that the “BIN” scheme with a single-photon threshold can approach the ideal unquantized CRLB when the oversampling factor is large enough. Furthermore, when compared with the “SAT” scheme, the “BIN” scheme has a more gradual decrease in SNR when the light exposure values increase, and has a higher dynamic range. For example, when \( K = 2^{16} \), the dynamic range of the “BIN” scheme spans from \( c = 10^2 \) to \( c = 10^{5.8} \), about two orders of magnitude higher than that of “SAT”. In Section 2.5, we will present a numerical experiment that points to a potential application of the binary sensor in high dynamic range photography.

**Remark.** Note that \( K \) is the product of the spatial oversampling factor and the temporal oversampling factor. For example, the pixel pitch of the image sensor reported in [18] is \( 1.65\mu m \). If the binary sensor is built on memory chip technology, with a pitch size of 50 nm [72], then the maximum spatial oversampling factor is about 1089. To achieve \( K = 2^{13}, 2^{14}, 2^{15} \) and \( 2^{16} \), respectively, as required in Figure 2.4, we then need to have temporal oversampling factors ranging from 8 to 60. Unlike traditional sensors which require multi-bit quantizers, the binary sensors only need one-bit comparators. This simplicity in hardware can potentially lead to faster readout speeds, making it practical to apply temporal oversampling.

### 2.4 Optimal Image Reconstruction and Efficient Implementations

In the previous section, we studied the performance of the binary image sensor, and derived the MLE for a piecewise-constant light field model. Our analysis establishes the optimality of the MLE, showing that, with single-photon thresholding and large oversampling factors, the MLE approaches the performance of an ideal sensing scheme without quantization. In this section, we extend the MLE to the general linear field model in (2.1), with arbitrary interpolation kernels. As a main result of this chapter, we show that the log-likelihood function is always concave. This desirable property guarantees the global convergence of iterative numerical algorithms in solving the MLE.

---

\(^2\)The estimator in (2.34) is biased around \( c = C_{\text{max}} \). For a very narrow range of light intensity values centered around \( C_{\text{max}} \), the MSE of this biased estimator is lower than the ideal CRLB. Thus, there is actually a short “spike” in SNR right before the drop.
2.4 Optimal Image Reconstruction and Efficient Implementations

2.4.1 Image Reconstruction by MLE

Under the linear field model introduced in Definition 1, reconstructing an image \(i.e.,\) the light field \(\lambda(x)\) is equivalent to estimating the parameters \(\{c_n\}\) in (2.1). As shown in (2.8), the light exposure values \(\{s_m\}\) at different sensors are related to \(\{c_n\}\) through a linear mapping, implemented as upsampling followed by filtering as in Figure 2.2. Since it is linear, the mapping (2.8) can be written as a matrix-vector multiplication

\[
s = Gc,
\]

where \(s \equiv [s_0, s_1, \ldots, s_{M-1}]^T, c \equiv [c_0, c_1, \ldots, c_{N-1}]^T,\) and \(G\) is an \(M \times N\) matrix representing the combination of upsampling (by \(K\)) and filtering (by \(g_m\)). Each element of \(s\) can then be written as

\[
s_m = e_m^T Gc,
\]

where \(e_m\) is the \(m\)th standard Euclidean basis vector.³

Remark. In using the above notations, we do not distinguish between single exposure and multiple exposures, whose equivalence has been established by Proposition 1 in Section 2.2.4. In the case of multiple exposures, the essential structure of \(G\)—upsampling followed by filtering—remains the same. All we need to do is to replace \(s\) by the interlaced sequence \(\{\tilde{s}_m\}\) constructed in (2.16), the oversampling factor \(K\) by \(KJ\), and the filter \(g_m\) by \(\tilde{g}_m\) in (2.17).

Similar to our derivations in (2.20) and (2.21), the likelihood function given \(M\) binary measurements \(b \equiv [b_0, b_1, \ldots, b_{M-1}]^T\) can be computed as

\[
\mathcal{L}_b(c) = \prod_{m=0}^{M-1} \mathbb{P}(B_m = b_m; s_m) = \prod_{m=0}^{M-1} p_{b_m}(e_m^T Gc),
\]

where (2.37) follows from (2.12) and (2.36). In our subsequent discussions, it is more convenient to work with the log-likelihood function, defined as

\[
\ell_b(c) \equiv \log \mathcal{L}_b(c) = \sum_{m=0}^{M-1} \log p_{b_m}(e_m^T Gc).
\]

For any given observation \(b\), the MLE we seek is the parameter that maximizes \(\mathcal{L}_b(c)\), or equivalently, \(\ell_b(c)\). Specifically,

\[
\hat{c}_{\text{ML}}(b) \equiv \arg \max_{c \in [0, S]^N} \ell_b(c) = \arg \max_{c \in [0, S]^N} \sum_{m=0}^{M-1} \log p_{b_m}(e_m^T Gc).
\]

The constraint \(c \in [0, S]^N\) means that every parameter \(c_n\) should satisfy \(0 \leq c_n \leq S\), for some preset maximum value \(S\).

³Here we use zero-based indexing. Thus, \(e_0 \equiv [1, 0, \ldots, 0]^T, e_1 \equiv [0, 1, \ldots, 0]^T,\) and so on.
Figure 2.5: The likelihood and log-likelihood functions for piecewise-constant light fields. (a) The likelihood functions $L_b(c)$, defined in (2.22), under different choices of the quantization thresholds $q = 1, 3, 5$, respectively. (b) The corresponding log-likelihood functions. In computing these functions, we set the parameters in (2.22) as follows: $K = 12$, i.e., the sensor is 12-times oversampled. The binary sensor measurements contain 10 “1’s”, i.e., $K_1 = 10$.

Example 3. As discussed in Section 2.3, when the light field is piecewise-constant, different light field parameters $\{c_n\}$ can be estimated independently. In that case, the likelihood function has only one variable [see (2.22)] and can be easily visualized. In Figure 2.5, we plot $L_b(c)$ in (2.22) and the corresponding log-likelihood function $\ell_b(c)$, under different choices of the quantization thresholds. We observe from the figures that the likelihood functions are not concave, but the log-likelihood functions indeed are. In what follows, we will show that this result is general, namely, the log-likelihood functions in the form of (2.38) are always concave.

Lemma 2. For any two integers $i, j$ such that $0 \leq i \leq j < \infty$ or $0 \leq i < j \leq \infty$, the function
\[
\log \sum_{k=0}^{j} \frac{x^k e^{-x}}{k!}
\]
is concave on the interval $x \in [0, \infty)$.

Proof. See Appendix 2.A.7.

Theorem 2. For arbitrary binary sensor measurements $b$, the log-likelihood function $\ell_b(c)$ defined in (2.38) is concave on the domain $c \in [0, S]^N$.

Proof. It follows from the definition in (2.12) that, for any $b_m \in \{0, 1\}$, the function $\log p_{b_m}(s)$ is either
\[
\log \sum_{k=0}^{q-1} \frac{s^k e^{-s}}{k!} \quad \text{or} \quad \log \sum_{k=q}^{\infty} \frac{s^k e^{-s}}{k!}.
\]
We can apply Lemma 2 in both cases, and show that $\{\log p_{b_m}(s)\}$ are concave functions for $s \geq 0$. Since the sum of concave functions is still concave and the composition of a
concave function with a linear mapping \((s_m = e_m^T Gc)\) is still concave, we conclude that the log-likelihood function defined in (2.38) is concave.

In general, there is no closed-form solution to the maximization problem in (2.39). An MLE solution has to be found through numerical algorithms. Theorem 2 guarantees the global convergence of these iterative numerical methods.

### 2.4.2 Iterative Algorithm and Efficient Implementations

We compute the numerical solution of the MLE by using a standard gradient ascent method. Denote by \(c^{(k)}\) the estimation of the unknown parameter \(c\) at the \(k\)th step. The estimation \(c^{(k+1)}\) at the next step is obtained by

\[
e^{(k+1)} = P_D \left( e^{(k)} + \gamma_k \nabla \ell_b(e^{(k)}) \right),
\]

where \(\nabla \ell_b(e^{(k)})\) is the gradient of the log-likelihood function evaluated at \(e^{(k)}\). \(\gamma_k\) is the step-size at the current iteration, and \(P_D\) is the projection onto the search domain \(D \equiv [0, S]^N\). We apply \(P_D\) to ensure that all estimations of \(c\) lie in the search domain.

Taking the derivative of the log-likelihood function \(\ell_b(e)\) in (2.38), we can compute the gradient as

\[
\nabla \ell_b(e^{(k)}) = G^T \left[ D_{b_0}(s_0^{(k)}), D_{b_1}(s_1^{(k)}), \ldots, D_{b_{M-1}}(s_{M-1}^{(k)}) \right]^T,
\]

where \(s^{(k)} \equiv [s_0^{(k)}, \ldots, s_{M-1}^{(k)}]^T = Ge^{(k)}\) is the current estimation of the light exposure values, and

\[
D_b(s) \equiv \frac{d}{ds} \log p_b(s) \quad \text{for } b = 0, 1.
\]

For example, when \(q = 1\), we have \(p_0(s) = e^{-s}\) and \(p_1(s) = 1 - e^{-s}\). In this case, \(D_0(s) = -1\) and \(D_1(s) = 1/(1 - e^{-s})\), respectively.

The choice of the step size \(\gamma_k\) has significant influence over the speed of convergence of the above iterative algorithm. We follow [78] by choosing, at each step, a \(\gamma_k\) so that the gradient vectors at the current and the next iterations are approximately orthogonal to each other. By assuming that the estimates \(s^{(k+1)}\) and \(s^{(k)}\) at consecutive iterations are close to each other, we can use the following first-order approximation

\[
D_b(s_m^{(k+1)}) \approx D_b(s_m^{(k)}) + H_b(s_m^{(k)})(s_m^{(k+1)} - s_m^{(k)}),
\]

where

\[
H_b(s) \equiv \frac{d}{ds} D_b(s) = \frac{d^2}{ds^2} \log p_b(s) , \quad \text{for } b = 0, 1.
\]

It follows that

\[
\nabla \ell_b(e^{(k+1)}) \\
= G^T \left[ D_{b_0}(s_0^{(k+1)}), D_{b_1}(s_1^{(k+1)}), \ldots, D_{b_{M-1}}(s_{M-1}^{(k+1)}) \right] \\
\approx \nabla \ell_b(e^{(k)}) + G \text{diag} \left\{ H_{b_0}(s_0^{(k)}), H_{b_1}(s_1^{(k)}), \ldots, H_{b_{M-1}}(s_{M-1}^{(k)}) \right\} (s^{(k+1)} - s^{(k)}).
\]
Assuming that the gradient update $c^{(k)} + \gamma_k \nabla \ell_b(c^{(k)})$ is inside of the constraint set $\mathcal{D}$, we can neglect the projection operator $P_\mathcal{D}$ in (2.41), and write

$$s^{(k+1)} - s^{(k)} = G(c^{(k+1)} - c^{(k)}) = \gamma_k G \nabla \ell_b(c^{(k)}).$$

Substituting the above equality into (2.43), we get

$$\nabla \ell_b(c^{(k+1)}) \\
\approx \nabla \ell_b(c^{(k)}) + \gamma_k G^T \text{diag} \left\{ H_{b_0}(s_0^{(k)}), H_{b_1}(s_1^{(k)}), \ldots, H_{b_{M-1}}(s_{M-1}^{(k)}) \right\} G \nabla \ell_b(c^{(k)}).$$

Finally, by requiring that $\nabla \ell_b(c^{(k+1)})$ be orthogonal to $\nabla \ell_b(c^{(k)})$, we compute the optimal step size as

$$\gamma_k = \frac{\|\nabla \ell_b(c^{(k)})\|^2}{\|\text{diag} \left\{ \sqrt{-H_{b_0}(s_0^{(k)}), \ldots, \sqrt{-H_{b_{M-1}}(s_{M-1}^{(k)})} \right\} G \nabla \ell_b(c^{(k)})\|^2}. \quad (2.44)$$

**Remark.** By definition, $H_b(s)$ (for $b = 0, 1$) are the second-order derivatives of concave functions (see Lemma 2), and are thus nonpositive. Consequently, the terms $\sqrt{-H_b(s)}$ in the denominator of (2.44) are well-defined.

At every iteration of the gradient algorithm, we need to update the gradient and the step size $\gamma_k$. We see from (2.42) and (2.44) that the computations always involve matrix-vector products in the form of $Ga$ and $G^Tb$, for some vectors $a, b$. The matrix $G$ is of size $M \times N$, where $M$ is the total number of pixels. In practice, $M$ will be in the range of $10^9 \sim 10^{10}$ (i.e., gigapixels per chip), making it impossible to directly implement the matrix operations. Fortunately, the matrix $G$ used in both formulae is highly structured, and can be implemented as upsampling followed by filtering (see our discussions in Section 2.2.2 and the expression (2.8) for details). Similarly, the transpose $G^T$ can be implemented by filtering (by $g_{-m}$) followed by downsampling, essentially “flipping” all the operations in $G$. Figure 2.6(a) and Figure 2.6(b) summarizes these operations.

We note that the implementations illustrated in Figure 2.6(a) and Figure 2.6(b) are not yet optimized: For example, the input to the filter $g_m$ in Figure 2.6(a) is an upsampled sequence, containing mostly zero elements; In Figure 2.6(b), we compute a full filtering operation (by $g_{-m}$), only to discard most of the filtering results in the subsequent downsampling step. All these deficiencies can be eliminated by using the tool of polyphase representations from multirate signal processing [84, 85], as follows.

First, we split the filter $g_m$ into $K$ non-overlapping polyphase components $g_{0,m}, g_{1,m}, \ldots, g_{K-1,m}$, defined as

$$g_{k,m} = g_{Km+k}, \quad \text{for } 0 \leq k < K. \quad (2.45)$$

Intuitively, the polyphase components specified in (2.45) are simply downsampled versions of the original filter $g_m$, with the sampling locations of all these polyphase components forming a complete partition. The mapping between the filter $g_m$ and its polyphase components is one-to-one. To reconstruct $g_m$, we can easily verify that, in the $z$-domain,

$$G(z) = G_0(z^K) + z^{-1}G_1(z^K) + \ldots + z^{-(K-1)}G_{K-1}(z^K). \quad (2.46)$$

Following the same steps as above, we can also split the sequences $u_m$ and $b_m$ in Figure 2.6 into their respective polyphase components $u_{0,m}, u_{1,m}, \ldots, u_{K-1,m}$ and $b_{0,m}, b_{1,m}, \ldots, b_{K-1,m}$. 


2.4 Optimal Image Reconstruction and Efficient Implementations

Figure 2.6: Signal processing implementations of $Ga$ and $G^Tb$. (a) The product $Ga$ can be obtained by upsampling followed by filtering. (b) The product $G^Tb$ can be obtained by filtering followed by downsampling. Note that the filter used in (b) is $g_{-m}$, i.e., the “flipped” version of $g_m$. (c) The polyphase domain implementation of (a). (d) The polyphase domain implementation of (b).
Proposition 5. Denote by $U_k(z)$ and $B_k(z)$ (for $0 \leq k < K$) the $z$-transforms of $u_{k,m}$ and $b_{k,m}$, respectively. Then,

$$U_k(z) = A(z)G_k(z), \quad \text{for } 0 \leq k < K, \quad (2.47)$$

and

$$V(z) = \sum_{k=0}^{K-1} B_k(z)G_k(z^{-1}). \quad (2.48)$$

Proof. See Appendix 2.A.8.

The results of Proposition 5 require some further explanations. What equation (2.47) suggests is an alternative implementation of $Ga$, as shown in Figure 2.6(c). We compute $K$ parallel convolutions between the input $a_m$ and the polyphase filters $\{g_{k,m}\}$. The channel outputs are the polyphase components $\{u_{k,m}\}$, which can be combined to form the desired output $u_m$. Similarly, it follows from (2.48) that $G^Tb$ can be implemented by the parallel filtering scheme in Figure 2.6(d).

The new implementations in Figure 2.6(c) and Figure 2.6(d) are significantly faster than their respective counterparts. To see this, suppose that the filter $g_m$ has $L$ coefficients. It is easy to see that the original implementation in Figure 2.6(a) requires $O(KL)$ arithmetic operations for every pixel in $a_m$. In contrast, each individual channel in Figure 2.6(c) requires only $O(L/K)$ arithmetic operations (due to the shorter supports of the polyphase filters), and thus the total cost of Figure 2.6(c) stays at $O(L)$ operations per pixel. This represents a $K$-fold reduction in computational complexities. A similar analysis also shows that Figure 2.6(d) needs $K$-times fewer operations than Figure 2.6(b). Recall that $K$ is the oversampling factor of our image sensor. As we operate in highly oversampled regimes (e.g., $K = 1024$) to compensate for information loss due to one-bit quantizations, the above improvements make our algorithms orders of magnitude faster.

2.5 Numerical Results

We present several numerical results in this section to verify our theoretical analysis and the effectiveness of the proposed image reconstruction algorithm.

2.5.1 1-D Synthetic Signals

Consider a 1-D light field $\lambda(x)$ shown in Figure 2.7(a). The interpolation filter $\varphi(x)$ we use is the cubic B-spline function $\beta_3(x)$ defined in (2.3). We can see that $\lambda(x)$ is a linear combination of the shifted kernels, with the expansion coefficients $\{c_n\}$ shown as blue dots in the figure.

We simulate a binary sensor with threshold $q = 1$ and oversampling factor $K = 256$. Applying the proposed MLE-based algorithm in Section 2.4, we obtain a reconstructed light field (the red dashed curve) shown in Figure 2.7(b), together with the original "ground truth" (the blue solid curve). We observe that the low-light regions are well-reconstructed but there exist large "overshoots" in the high-light regions.

We can substantially improve the reconstruction quality by increasing the oversampling factor of the sensor. Figure 2.7(c) shows the result obtained by increasing the spatial oversampling factor to $K = 2048$. Alternatively, we show in Figure 2.7(d) a
Figure 2.7: Binary sensing and reconstructions of 1-D light fields. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernels. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with spatial oversampling factor $K = 256$. (c) An improved reconstruction result due to the use of a larger spatial oversampling factor $K = 2048$. (d) An alternative result, obtained by keeping $K = 256$ but taking $J = 8$ consecutive exposures.

different reconstruction result obtained by keeping the original spatial oversampling factor at $K = 256$, but taking $J = 8$ consecutive exposures. Visually, the two sensor configurations, i.e., $K = 2048, J = 1$ and $K = 256, J = 8$, lead to very similar reconstruction performances. This observation agrees with our previous theoretical analysis in Section 2.2.4 on the equivalence between spatial and temporal oversampling schemes.

2.5.2 Acquiring Scenes with High Dynamic Ranges

A well-known difficulty in photography is the limited dynamic ranges of the image sensors. Capturing both very bright and very dark regions faithfully in a single image is difficult. For example, Figure 2.8(a) shows several images taken inside of a church with different exposure times [19]. The scene contains both sun-lit areas and shadow regions, with the former over a thousand times brighter than the latter. Such high dynamic ranges are well-beyond the capabilities of conventional image sensors. As a result, these images are either overexposed or underexposed, with no single image rendering details in both areas. In light of this problem, an active area of research in computational photography is to reconstruct a high dynamic range radiance map by combining multiple images with different exposure settings (see, e.g., [19, 89]). While producing successful results, such multi-exposure approaches can be time-consuming.

In Section 2.3.3, we have shown that the binary sensor studied in this chapter can achieve higher dynamic ranges than conventional image sensors. To demonstrate this advantage, we use the high dynamic range radiance map obtained in [19] as the ground truth data [i.e., the light field $\lambda(x)$ as defined in (2.1)], and simulate the acquisition
Figure 2.8: High dynamic range photography using the binary sensor. (a) A sequence of images taken inside of a church with decreasing exposure times [19]. (b) The reconstructed high dynamic range radiance map (in logarithmic scales) using our MLE reconstruction algorithm. (c) The tone-mapped version of the reconstructed radiance map.

Figure 2.9: Image reconstruction from the binary measurements taken by a SPAD sensor [20], with a spatial resolution of $32 \times 32$ pixels. The final image (lower-right corner) is obtained by incorporating 4096 consecutive frames, 50 of which are shown in the figure.

of this scene by using a binary sensor with a single photon threshold. The spatial oversampling factor of the binary sensor is set to $32 \times 32$, and the temporal oversampling factor is 256 (i.e., 256 independent frames). Similar to our previous experiment on 1-D signals, we use a cubic B-spline kernel [i.e., $\varphi(x) = \beta_3(x)$] along each of the spatial dimensions. Figure 2.8(b) shows the reconstructed radiance map using our algorithm described in Section 2.4. Since the radiance map has a dynamic range of $3.3 \times 10^5 : 1$, the image is shown in logarithmic scale. To have a visually more pleasing result, we also shown in Figure 2.8(c) a tone-mapped [89] version of the reconstruction. We can see from Figure 2.8(b) and Figure 2.8(c) that details in both light and shadow regions have been faithfully preserved in the reconstructed radiance map, suggesting the potential application of the binary sensor in high dynamic range photography.

2.5.3 Results on Real Sensor Data

We have also applied our reconstruction algorithm to images taken by an experimental sensor based on single photon avalanche diodes (SPADs) [20]. The sensor has binary-valued pixels with single-photon sensitivities, i.e., the quantization threshold is $q = 1$. Due to its experimental nature, the sensor has limited spatial resolution, containing an array of only $32 \times 32$ detectors. To emulate the effect of spatial oversampling, we apply
temporal oversampling and acquire 4096 independent binary frames of a static scene. In this case, we can estimate the light intensity at each pixel independently by using the closed-form MLE solution in (28). Figure 2.9 shows 50 such binary images, together with the final reconstruction result (at the lower-right corner). The quality of reconstruction verifies our theoretical model and analysis.

2.6 Conclusions

We presented a theoretical study of a new image sensor that acquires light information using one-bit pixels—a scheme reminiscent of traditional photographic film. By formulating the binary sensing scheme as a parameter estimation problem based on quantized Poisson statistics, we analyzed the performance of the binary sensor in acquiring light intensity information. Our analysis shows that, with a single-photon quantization threshold and large oversampling factors, the binary sensor performs much like an ideal sensor, as if there were no quantization. To recover the light field from binary sensor measurements, we proposed an MLE-based image reconstruction algorithm. We showed that the corresponding log-likelihood function is always concave, thus guaranteeing the global convergence of numerical solutions. To solve for the MLE, we adopt a standard gradient method, and derive efficient implementations using fast signal processing algorithms in the polyphase domain. Finally, we presented numerical results on both synthetic data and images taken by a prototype sensor. These results verify our theoretical analysis and demonstrate the effectiveness of our image reconstruction algorithm. They also point to the potential of the new binary sensor in high dynamic range photography applications.
2. A Appendix

2. A. 1 Proof of Proposition 1

The sequence $\tilde{s}_m$ in (2.16) can be written, equivalently, as

$$\tilde{s}_m = \sum_{j=0}^{J-1} \sum_{n=0}^{M-1} s_{j,n} \delta_{m-Jn-j},$$

where $\delta_i$ is the Kronecker delta function. Taking $z$-transforms on both sides of the equality leads to

$$\tilde{S}(z) = \sum_{j=0}^{J-1} \sum_{n=0}^{M-1} s_{j,n} z^{-J} = \sum_{j=0}^{J-1} S_j(z^{-j}).$$

(2.49)

By substituting (2.15) into (2.49) and using the definition (2.17), we can simplify (2.49) as

$$\tilde{S}(z) = C(z^{KJ})\tilde{G}(z).$$

(2.50)

Finally, since $C(z^{KJ})$ is the $z$-transform of the sequence $\sum_n c_n \delta_{m-KJn}$, it follows from (2.50) that $\tilde{s}_m = (\sum_n c_n \delta_{m-KJn}) \ast \tilde{g}_m$, and thus (2.18).

2. A. 2 The CRLB of Binary Sensors

We first compute the Fisher information, defined as $I(c) = \mathbb{E}[-\frac{\partial^2}{\partial c^2} \log \mathcal{L}_b(c)]$. Using (2.22), we get

$$I(c) = \mathbb{E} \left[ -\frac{\partial^2}{\partial c^2} \left( K_1 \log p_1(c/K) + (K - K_1) \log p_0(c/K) \right) \right]$$

$$= \mathbb{E} \left[ \frac{K_1 (p''_0(c/K) p_1(c/K) + p'_0(c/K))^2}{K^2 p_1(c/K)^2} - \frac{(K - K_1) (p''_0(c/K) p_0(c/K) - p'_0(c/K))^2}{K^2 p_0(c/K)^2} \right],$$

(2.51)

where $p'_0(x) = \frac{d}{dx} p_0(x)$ and $p''_0(x) = \frac{d^2}{dx^2} p_0(x)$ are the first and second order derivative of $p_0(x)$, respectively. In reaching (2.51), we have also used the fact that $p_1(x) = 1 - p_0(x)$ and thus $p'_0(x) = -\frac{d}{dx} p_1(x)$ and $p''_0(x) = -\frac{d^2}{dx^2} p_1(x)$.

Note that $K_1 = \sum_{0 \leq m < K} b_m$ is a binomial random variable, and thus its mean can be computed as

$$\mathbb{E}[K_1] = K p_1(c/K) = K (1 - p_0(c/K)).$$

On substituting the above expression into (2.51), the Fisher information can be simplified as

$$I(c) = \frac{p''_0(c/K) p_1(c/K) + p'_0(c/K)^2}{K p_1(c/K)} - \frac{p''_0(c/K) p_0(c/K) - p'_0(c/K)^2}{K p_0(c/K)},$$

$$= \frac{p'_0(c/K)^2}{K p_0(c/K) p_1(c/K)}.$$  

(2.52)
Using the definition of $p_0(x)$ in (2.12), the derivative in the numerator of (2.52) can be computed as
\[
p_0'(x) = -e^{-x} \frac{x^{q-1}}{(q-1)!}.
\] (2.53)

Finally, since $\text{CRLB}_{\text{bin}}(K, q) = 1/I(c)$, we reach (2.23) by substituting (2.12) and (2.53) into (2.52), and after some straightforward manipulations.

### 2.A.3 The CRLB of Ideal Unquantized Sensors

Without quantization, the sensor measurements are $y \overset{\text{def}}{=} [y_0, y_1, \ldots, y_{K-1}]^T$, i.e., the number of photons collected at each pixel. The likelihood function in this case is
\[
L_y(c) \overset{\text{def}}{=} \mathbb{P}(Y_m = y_m, 0 \leq m < K; c),
\]
\[
= \prod_{m=0}^{K-1} \frac{(c/K)^{y_m} e^{-c/K}}{y_m!},
\] (2.54)

where (2.54) follows from the independence of $\{Y_m\}$ and the expressions (2.11) and (2.19).

Computing the Fisher information $I(c) = \mathbb{E}[-\frac{\partial^2}{\partial c^2} \log L_y(c)]$ in this case, we get
\[
I(c) = \mathbb{E} \left[ -\frac{\partial^2}{\partial c^2} \sum_{m=0}^{K-1} \left( y_m \log(c/K) - c/K - \log(y_m!) \right) \right]
\]
\[
= \mathbb{E} \left[ \sum_{m=0}^{K-1} y_m \right] / c^2.
\] (2.55)

Since $\{y_m\}$ are drawn from Poisson distributions as in (2.11), we have $\mathbb{E}[y_m] = s_m = c/K$ for all $m$. It then follows from (2.55) that $I(c) = K(c/K)/c^2 = 1/c$, and therefore $\text{CRLB}_{\text{ideal}}(K) = 1/I(c) = c$.

### 2.A.4 Proof of Proposition 3

Using (2.24), (2.52) and (2.53), and through a change of variables $c/K \rightarrow x$, we have
\[
\text{CRLB}_{\text{bin}}(K, q)/\text{CRLB}_{\text{ideal}}(K) = \frac{p_0(x)p_1(x)}{x^{q-1}e^{-2x}/(q-1)!^2}.
\] (2.56)

It follows from the properties of incomplete gamma functions that
\[
p_0(x) = \frac{1}{(q-1)!} \int_x^\infty t^{q-1}e^{-t} dt \quad \text{and} \quad p_1(x) = \frac{1}{(q-1)!} \int_0^x t^{q-1}e^{-t} dt.
\]

Using a change of variables $t \rightarrow \frac{x^2}{t}$, we can further rewrite $p_0(x)$ as
\[
p_0(x) = \frac{1}{(q-1)!} \int_0^x \left( \frac{x^2}{t} \right)^q e^{-\frac{x^2}{t}} \frac{dt}{t}.
\]
It follows that
\[
\frac{p_0(x)p_1(x)}{x^{2q-1}e^{-2x}/(q-1)!^2} = \left(\frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\frac{2}{t}\right)^q e^{-t^2/2} dt\right) \left(\frac{1}{\sqrt{2\pi}} \int_0^\infty t^{q-1}e^{-t} dt\right) \\
= x e^{2x} \int_0^\infty t^{q-1}e^{-t^2/2} dt \int_0^\infty t^{q-1}e^{-t} dt \\
\geq x e^{2x} \left(\int_0^\infty t^{-1}e^{-t^2/2} dt\right)^2,
\] (2.57)
where in reaching (2.57) we have used the Cauchy-Schwarz inequality.

It is easy to verify (through a change of variables \( t \to \frac{x}{t} \)) that \( \int_0^\infty t^{1}e^{-\frac{x^2}{2t^2} - \frac{t}{x}} dt = \int_0^\infty t^{-1}e^{-\frac{x^2}{2t^2} - \frac{t}{x}} dt \). Consequently, the term on the right-hand side of (2.57) is equal to \( xe^{2x} \left(\frac{1}{2} \int_0^\infty t^{-1}e^{-\frac{x^2}{2t^2} - \frac{t}{x}} dt\right)^2 \). Through a change of variables \( t \to t^2 \), we have,
\[
x e^{2x} \left(\frac{1}{2} \int_0^\infty t^{-1}e^{-\frac{x^2}{2t^2} - \frac{t}{x}} dt\right)^2 = xe^{2x} \left(\int_0^\infty t^{-1}e^{-\frac{x^2}{2t^2} - \frac{t}{x}} dt\right)^2 \\
= \left(\sqrt{x} \int_0^\infty t^{-1}e^{-\frac{1}{2}(x-t)^2} dt\right)^2 \\
= \left(\int_{-\infty}^\infty e^{-\frac{1}{2}u^2} / \sqrt{4 + u^2} du\right)^2,
\] (2.58)
where (2.58) is obtained through another change of variables \( t \to \frac{t}{2} + \sqrt{x + \frac{1}{4}} \).

We can easily verify that (2.58) is a monotonically increasing function with respect to \( x > 0 \). So, for \( x \geq 1 \), (2.58) is greater than
\[
\left(\int_{-\infty}^\infty e^{-\frac{1}{2}u^2} / \sqrt{4 + u^2} du\right)^2 \approx 1.31.
\] (2.59)

For \( 0 \leq x < 1 \), we can obtain the following inequalities by keeping the first two terms in \( p_0(x) \) and the first term in \( p_1(x) \):
\[
\frac{p_0(x)p_1(x)}{x^{2q-1}e^{-2x}/(q-1)!^2} \geq \frac{(1 + x)e^{-2x}x^q/q!}{x^{2q-1}e^{-2x}/(q-1)!^2} = \frac{(q-1)!(x^{1-q} + x^{2-q})}{q} \geq \frac{2(q-1)!}{q}.
\]
It is easy to see that \( \frac{2(q-1)!}{q} \) is a monotonically increasing function with respect to \( q \geq 2 \). Therefore,
\[
\frac{2(q-1)!}{q} \geq 4 \geq 1.33 \quad \text{for} \quad q \geq 3.
\] (2.60)

Finally, for \( q = 2 \), we keep the first two terms in \( p_0(x) \) and \( p_1(x) \) and get
\[
\frac{p_0(x)p_1(x)}{x^{2q-1}e^{-2x}/(q-1)!^2} \geq \frac{(1 + x)e^{-2x}(x^2/2 + x^3/6)}{x^3e^{-2x}} = x/6 + 2/3 + 1/2e \geq 1.33.
\] (2.61)
The second inequality is due to \( 0 < x < 1 \).
Combining (2.59), (2.60) and (2.61), we reach the desired result in (2.26).
Finally, we show that, for \( q \geq 2 \), the “gap” between CRLB_{bin}(K, q) and CRLB_{ideal}(K) will only get bigger as the oversampling factor \( K \) grows. To that end, we notice that when \( K \to \infty \), the variable \( x = c/K \to 0 \). Keeping the first terms in \( p_0(x) \) and \( p_1(x) \), we have
\[
\frac{p_0(x)p_1(x)}{x^{q-1}e^{-2x}/(q-1)!} \geq \frac{e^{-2x}x^q/q!}{x^{2q-1}e^{-2x}/(q-1)!} = \frac{(q-1)!x^{1-q}}{q}.
\]
For \( q \geq 2 \), the above quantity goes to infinity as \( x \to 0 \). Therefore,
\[
\lim_{K \to \infty} \frac{\text{CRLB}_{bin}(K, q)}{\text{CRLB}_{ideal}(K)} = \infty.
\]

### 2.A.5 Proof of Theorem 1

When \( q = 1 \), we have \( p_0(x) = e^{-x} \) and thus \( p_0[^{-1}](x) = -\log(x) \). In this case, the MLE solution in (2.28) can be rewritten as
\[
\hat{c}_{\text{ML}}(b) = \begin{cases} -K \log(1 - K_1/K), & \text{if } 0 \leq K_1 \leq K(1 - e^{-S/K}), \\ S, & \text{otherwise.} \end{cases}
\]
We note that \(-K \log(1 - K_1/K) = K_1 + \frac{K_1^2}{2K} + \frac{K_1^3}{3K^2} + \ldots \) and that \( \lim_{K \to \infty} K(1 - e^{-S/K}) = S \). Thus, for sufficiently large \( K \), the above MLE solution can be further simplified as
\[
\hat{c}_{\text{ML}}(b) = \begin{cases} K_1 + \mathcal{O}(1/K), & \text{if } 0 \leq K_1 < S, \\ S, & \text{otherwise.} \end{cases}
\]
(2.62)
Without loss of generality, we assume that \( S \) is an integer in what follows. The expected value of the MLE then becomes
\[
\mathbb{E}[\hat{c}_{\text{ML}}(b)] = \sum_{n=0}^{S-1} n \mathbb{P}(K_1 = n) + S \sum_{n=S}^{K} \mathbb{P}(K_1 = n) + \mathcal{O}(1/K).
\]
Using the following identity \( c = \sum_{n=0}^{\infty} n e^{n-c} / n! \) about the mean of a Poisson random variable, we have
\[
|\mathbb{E}[\hat{c}_{\text{ML}}(b)] - c| = \sum_{n=0}^{S-1} n \left( \mathbb{P}(K_1 = n) - \frac{e^n e^{-c}}{n!} \right) + S \sum_{n=S}^{K} \mathbb{P}(K_1 = n) - \sum_{n=S}^{\infty} \frac{e^n e^{-c}}{n!} + \mathcal{O}(1/K)
\leq S \sum_{n=0}^{S-1} \left( \mathbb{P}(K_1 = n) - \frac{e^n e^{-c}}{n!} \right) + S \sum_{n=S}^{K} \mathbb{P}(K_1 = n) + \sum_{n=S}^{\infty} \frac{e^n e^{-c}}{(n-1)!} + \mathcal{O}(1/K). \quad (2.63)
\]
In what follows, we derive bounds for the quantities on the right-hand side of the above inequality. First, consider the probability \( \mathbb{P}(K_1 = n) \). Since \( K_1 \) is a binomial random variable, we have
\[
\mathbb{P}(K_1 = n) = \binom{K}{n} (1 - p_0(c/K))^n p_0(c/K)^{(K-n)}
= \frac{K(K-1)\ldots(K-n+1)}{n!K^n} (K(1 - e^{-c/K}))^n e^{-c(K-n)/K}. \quad (2.64)
\]
For every $n < S$, it is easy to verify that 
\[
\frac{K(K-1) \ldots (K-n+1)}{K^n} = 1 + \mathcal{O}(1/K), \quad (K-1)^n e^{-c} = c^n + \mathcal{O}(1/K) \quad \text{and} \quad e^{-c(K-n)/K} = e^{-c} + \mathcal{O}(1/K).
\]
Thus, for any $n < S$, we can simplify (2.64) as
\[
P(K_1 = n) = \frac{c^n}{n!} e^{-c} + \mathcal{O}(1/K).
\] (2.65)

It follows that
\[
S \left| \sum_{n=0}^{S-1} \left( P(K_1 = n) - \frac{c^n e^{-c}}{n!} \right) \right| = \mathcal{O}(1/K). \quad (2.66)
\]

Next, consider the second term on the right-hand side of (2.63).
\[
S \sum_{n=S}^{K} P(K_1 = n) = S(1 - \sum_{n=0}^{S-1} P(K_1 = n))
\]
\[
= S(1 - \sum_{n=0}^{S-1} \frac{c^n}{n!} e^{-c}) + \mathcal{O}(1/K) \quad (2.67)
\]
\[
= S \sum_{n=S}^{\infty} \frac{c^n}{n!} e^{-c} + \mathcal{O}(1/K)
\]
\[
\leq S e^{-c} \left( \frac{ec}{S} \right)^S + \mathcal{O}(1/K), \quad (2.68)
\]

for all $c < S$, where (2.67) follows from (2.65) and the inequality (2.68) is due to the Chernoff bound on the tail of Poisson distributions [90]. Similarly, the third term on the right-hand side of (2.63) can be rewritten as
\[
\sum_{n=S}^{\infty} \frac{c^n e^{-c}}{(n-1)!} = c \sum_{n=S-1}^{\infty} \frac{c^n e^{-c}}{n!} \leq c e^{-c} \left( \frac{ec}{S-1} \right)^{S-1}, \quad (2.69)
\]

where the inequality is again an application of the Chernoff bound. Finally, on substituting (2.66), (2.68) and (2.69) into (2.63), and after some simple manipulations, we reach (2.29).

The proof for the mean-squared error formula (2.30) is similar. Using (2.62), we have
\[
E \left[ (\hat{c}_{ML}(b) - c)^2 \right] = \sum_{n=0}^{S-1} (n-c)^2 P(K_1 = n) + (S-c)^2 \sum_{n=S}^{K} P(K_1 = n) + \mathcal{O}(1/K)
\]
\[
= \sum_{n=0}^{S-1} (n-c)^2 \frac{c^n e^{-c}}{n!} + (S-c)^2 \sum_{n=S}^{\infty} \frac{c^n e^{-c}}{n!} + \mathcal{O}(1/K), \quad (2.70)
\]

where in reaching (2.70), we have used the estimation (2.65) of the Binomial probabilities. We note that the variance of a Poisson random variable is equal to its mean. Thus,
\( c = \sum_{n=0}^{\infty} (n - c)^2 \frac{e^{-c}}{n!} \). On combining this identity with (2.70),

\[
\left| E \left[ \left( \hat{c}_{\text{ML}}(b) - c \right)^2 \right] - c \right| = \left| (S - c)^2 \sum_{n=S}^{\infty} \frac{e^{-c}}{n!} - \sum_{n=S}^{\infty} (n - c)^2 \frac{e^{-c}}{n!} + O(1/K) \right| \\
\leq 2 \sum_{n=S}^{\infty} \frac{n^2 e^{-c}}{n!} + O(1/K), \quad \text{for } c \leq S \\
\leq 2c(c + 1) \sum_{n=S-2}^{\infty} \frac{e^{-c}}{n!} + O(1/K), \quad \text{for } c < S - 2.
\]

Applying the Chernoff bound to the above inequality, we get (2.30).

### 2.A.6 Proof of Proposition 4

We have \( p_0(x) = e^{-x} \sum_{n=0}^{q-1} \frac{x^n}{n!} \), and thus \( 1 - p_0(x) = e^{-x} \sum_{n=q}^{\infty} \frac{x^n}{n!} \). It follows that

\[
K(1 - p_0(S/K)) = Ke^{-S/K} \left( \frac{S^q}{K^q q!} + \frac{S^{q+1}}{K^{q+1} (q+1)!} + \ldots \right) = e^{-S/K} \left( \frac{S^q}{K^{q-1} q!} + \frac{S^{q+1}}{K^q (q+1)!} + \ldots \right).
\]

For \( q \geq 2 \) and any fixed constant \( S \), the above quantity converges to 0 as \( K \) tends to infinity. As a result, for sufficiently large \( K \), the MLE solution in (2.28) can be simplified as

\[
\hat{c}_{\text{ML}}(b) = \begin{cases} 
0, & \text{if } K_1 = 0, \\
S, & \text{otherwise},
\end{cases} \tag{2.71}
\]

where we have also used the fact that \( p_0(0) = 1 \) and thus \( p_0^{-1}(1) = 0 \). Using (2.71), we can compute the expected value of the MLE as

\[
E[\hat{c}_{\text{ML}}(b)] = 0 P(K_1 = 0) + S(1 - P(K_1 = 0)) = S(1 - P(K_1 = 0)). \tag{2.72}
\]

We have \( K_1 = 0 \) when all the pixel responses are uniformly 0. The probability of seeing such an event is

\[
P(K_1 = 0) = p_0(c/K)^K = e^{-c} \left( 1 + \frac{c}{K} + \ldots + \frac{c^{q-1}}{K^{q-1} (q-1)!} \right)^K,
\]

which converges to 1 as \( K \) tends to infinity, i.e.,

\[
\lim_{K \to \infty} P(K_1 = 0) = 1. \tag{2.73}
\]

Substituting (2.73) into (2.72), we get (2.31).

Next, we compute the MSE as

\[
E \left[ \left( \hat{c}_{\text{ML}}(b) - c \right)^2 \right] = c^2 P(K_1 = 0) + (S - c)^2 (1 - P(K_1 = 0)),
\]

which, upon taking the limit as \( K \to \infty \), leads to (2.32).
2.A.7 Proof of Lemma 2

The function \( h(x) \defeq \log \sum_{k=1}^{j} \frac{x^{k-x}}{k!} \) is continuously differentiable on the interval \((0, \infty)\). Therefore, to establish its concavity, we just need to show that its second derivative is nonpositive. To that end, we first introduce a sequence of functions \( \{r_k(x)\}_{k \in \mathbb{Z} \cup \{\infty\}} \) defined as

\[
  r_k(x) = \begin{cases} 
    x^k/k!, & \text{if } 0 \leq k < \infty; \\
    0, & \text{if } k < 0 \text{ or } k = \infty.
  \end{cases}
\] (2.74)

It is straightforward to verify that \( \frac{d}{dx} r_k(x) = r_{k-1}(x) \) for all \( k \in \mathbb{Z} \cup \{\infty\} \). Now, rewriting \( h(x) \) as \( \log \sum_{k=i}^{j} r_k(x) - x \) and computing its second derivative, we get

\[
  \frac{d^2}{dx^2} h(x) = \frac{\left( \sum_{k=i}^{j} r_{k-2} \right) \left( \sum_{k=i}^{j} r_{k-1} \right) - \left( \sum_{k=i}^{j} r_{k-1} \right)^2}{\left( \sum_{k=i}^{j} r_k \right)^2},
\] (2.75)

where we have omitted the function argument \( x \) in \( r_k(x) \), \( r_{k-1}(x) \) and \( r_{k-2}(x) \) for notational simplicity.

Recall that our goal is to show that \( \frac{d^2}{dx^2} h(x) \leq 0 \), for \( x > 0 \). Since the denominator of (2.75) is always positive, we just need to focus on its numerator. Using the identities \( \sum_{i \leq k \leq j} r_k = \sum_{i < k \leq j} r_{k-1} + r_j - r_i \) and \( \sum_{i \leq k \leq j} r_{k-1} = \sum_{i \leq k \leq j} r_{k-2} + r_{j-1} - r_{i-1} \), we can simplify the numerator of (2.75) as follows:

\[
  \left( \sum_{i \leq k \leq j} r_{k-2} \right) \left( \sum_{i \leq k \leq j} r_{k-1} + r_j - r_i \right) - \left( \sum_{i \leq k \leq j} r_{k-1} \right) \left( \sum_{i \leq k \leq j} r_{k-2} + r_{j-1} - r_{i-1} \right)
\]

\[
  = \sum_{i \leq k \leq j} \left( (r_{k-2}r_j - r_{k-1}r_{j-1}) + (r_{k-1}r_{i-1} - r_{k-2}r_i) \right).
\] (2.76)

In what follows, we show that

\[
  r_{k-2}(x)r_j(x) - r_{k-1}(x)r_{j-1}(x) \leq 0
\] (2.77)

for arbitrary choices of \( x \geq 0 \) and \( i \leq k \leq j \), where \( 0 \leq i \leq j < \infty \) or \( 0 \leq i < j = \infty \). Note that, when \( k < 2 \) or \( j = \infty \), the left-hand side of (2.77) becomes \( -r_{k-1}(x)r_{j-1}(x) \) and thus (2.77) automatically holds. Now, assume that \( k \geq 2 \) and \( j < \infty \). From the definition in (2.74), the left-hand side of (2.77) is

\[
  \frac{x^{k-2}x^j}{(k-2)!j!} - \frac{x^{k-1}x^{j-1}}{(k-1)!(j-1)!} = \frac{x^{k+j-2}}{(k-2)!(j-1)!} \left( \frac{1}{j} - \frac{1}{k-1} \right) \leq 0
\]

for \( i \leq k \leq j \). Using similar arguments, we can also show that

\[
  r_{k-1}(x)r_{i-1}(x) - r_{k-2}(x)r_i(x) \leq 0, \quad \text{for } x \geq 0.
\] (2.78)

On substituting the inequalities (2.77) and (2.78) into (2.76), we verify that the numerator of (2.75) is nonpositive, and therefore \( \frac{d^2}{dx^2} h(x) \leq 0 \), for all \( x > 0 \).

2.A.8 Proof of Proposition 5

Expressing the signal processing operations in Figure 2.6 in the \( z \)-domain, we have

\[
  U(z) = A(z^K)G(z)
  = A(z^K)G_0(z^K) + z^{-1}A(z^K)G_1(z^K) + \ldots + z^{-(K-1)}A(z^K)G_{K-1}(z^K),
\] (2.79)
where $A(z^K)$ in the first equality is the $z$-transform of the $K$-times upsampled version of $a_m$, and (2.79) follows from (2.46). Similar to (2.46), we can expand $U(z)$ in terms of the $z$-transforms of its polyphase components, as

$$U(z) = U_0(z^K) + z^{-1}U_1(z^K) + \ldots + z^{-(K-1)}U_{K-1}(z^K).$$  

(2.80)

Comparing (2.79) against (2.80) and using the uniqueness of the polyphase decomposition, we conclude that $U_k(z) = A(z)G_k(z)$, for all $0 \leq k < K$.

Now, consider Figure 2.6(b). We note that the $z$-transform of $g_{-m}$ is $G(z^{-1})$. Denote by $d_m$ the output of the filtering operation. Then, its $z$-transform can be computed as

$$D(z) = B(z)G(z^{-1})$$

$$= \left( \sum_{k=0}^{K-1} z^{-k}B_k(z^K) \right) \left( \sum_{k=0}^{K-1} z^kG_k(z^{-K}) \right)$$

$$= \sum_{k=0}^{K-1} B_k(z^K)G_k(z^{-K}) + \sum_{0 \leq i \neq j < K} z^{j-i}B_i(z^K)G_j(z^K).$$  

(2.81)

When downsampling $d_m$ by $K$, only the first term on the right-hand side of (2.81) will be retained; the second term is “invisible” to the sampling operation due to mismatched supports. Therefore, we have, after downsampling, $V(z) = \sum_{k=0}^{K-1} B_k(z)G_k(z^{-1})$. 
Chapter 3

Oversampled Noisy Binary Image Acquisition

3.1 Introduction

In this chapter, we will study the noise performance of the proposed binary image sensor. Since the shot noise, i.e., the fluctuation of the photons received by the pixel, has already been considered in the previous chapter, we will focus on circuit noise here. The circuit noise is related to the pixel architecture and the technologies that are used to design the binary sensor.

If we want to build a binary sensor with small thresholds, then we need techniques such as those used in CMOS single-photon avalanche diode (SPAD) [67,20] and electron-multiplying charge-coupled device (EMCCD) sensors [64–66] described in Chapter 1. In these sensors, a single photon may generate multiple electrons. The gain in the SPAD and the EMCCD is defined as the number of electrons divided by the number of photons. Electrons are accumulated in order to obtain a voltage that is compared with a threshold. Therefore, the threshold defined in our binary image sensor could be converted to a voltage threshold in the SPAD or the EMCCD. Due to the large gain in the SPAD and the EMCCD, even if the voltage threshold has noise, the influence of the noise on the threshold in the model of our binary image sensor can be ignored. As indicated in [91,92], the main noise in these sensors are shot noise and dark current noise. Here we consider the dark current noise. Due to the positive property of the dark current noise, it can only turn a pixel from “0” to “1”. We can model this noise as an additive Bernoulli noise with a known parameter $p_e$, which is called the noise rate. This can be easily estimated by covering the lens, taking some pictures, and computing the percentage of the “1”s in the binary image.

In this chapter, we present a theoretical analysis of the performance of the binary sensor under noise. We show that the binary sensor is quite robust to noise and the performance worsens only slightly. The dynamic range remains much higher than that of a conventional sensor.

In Section 3.2, we describe the noisy binary sensing model. Following that are the two main contributions of this chapter:

1. *Influence of the noise:* In Section 3.3, we study the performance of the noisy binary sensor for estimating a piecewise-constant light intensity field. We derive the
3.2 Imaging by Oversampled Noisy Binary Sensors

We consider the same imaging model as shown in Figure 2.1 in the previous chapter and we modify it as Figure 3.1 to take into account the presence of noise. In the new model, the binary output $b_m$ is contaminated by the additive noise $w_m$, $w_m \in \{0, 1\}$.

We model $w_m$ as the realization of a Bernoulli random variable $W_m$ with parameter $p_e$, called noise rate, and then we can write

$$\mathbb{P}(W_m = w_m; p_e) = \begin{cases} p_e, & \text{if } w_m = 1, \\ 1 - p_e, & \text{otherwise}, \end{cases} \quad (3.1)$$

Figure 3.1: The signal processing block diagram of the imaging model studied in this chapter. We upsample and filter the expansion coefficients $c_n$ to get the light exposure value $s_m$ at the $m$th pixel. Then, the binary image sensor converts $\{s_m\}$ into quantized measurements $\{b_m\}$. Due to noise $\{w_m\}$, we get the contaminated binary measurements $\{b'_m\}$.

Cramér-Rao lower bound (CRLB) of the estimation variance, and show that unlike the noiseless case, when the oversampling factor is too large, it will deteriorate the estimation performance. We compare the noisy binary sensor with the noiseless sensor and a conventional image sensor in Section 3.3.3. The performance of the noisy binary sensor is just slightly worse than that of the noiseless sensor, which shows that the binary sensor is quite robust to noise. Larger noise rates will result in a narrower dynamic range. But, with suitable oversampling factors, the binary sensor is still better than the conventional sensor in terms of dynamic range.

2. Image reconstruction: In Section 3.4, we propose to use the maximum likelihood estimator (MLE) for estimating the light intensity field. For a single-photon threshold, we prove that the log-likelihood function is still concave. Thus, we can find the optimal solution using iterative methods. But for thresholds larger than “1”, this property does not hold true. We show that when the light intensity is piecewise-constant, the expansion coefficients can be estimated separately and both the likelihood function and the log-likelihood function are strictly pseudoconcave, which ensures us to find the optimal solution using a bisection method. For the general light intensity field, we show that the log-likelihood function is not even quasiconcave. But by assuming the piecewise-constant light field model, we can get an initial estimation using the bisection method. Then a refined result can be obtained using Newton’s method.

Section 3.5 gives numerical results on both synthesized data and real images. The results demonstrate the correctness of our theoretical analysis and the efficacy of our reconstruction algorithm.

Similarly to the previous chapter, to simplify the notation, we focus our discussion on a one-dimensional (1-D) sensor array. All the results can be easily extended to the 2-D case.
We define the final output for the \( m \)th pixel \( b_m^e \) as \( b_m \lor w_m \), where \( b_m \in \{0,1\} \) and \( \lor \) is the disjunction binary operator. It is the realization of the random variable \( B_m^e \) as \( B_m \lor W_m \). Introducing two functions
\[
p_0^e(s) \triangleq p_0(s)(1 - p_e), \quad p_1^e(s) \triangleq p_0(s)p_e + p_1(s) ,
\]
where \( p_0(s) \) and \( p_1(s) \) are the functions defined in (2.12), and \( p_e \) is the noise rate.

From (2.13), (3.1), and (3.2), we have
\[
P(B_m^e = b_m^e; s_m, p_e) = p_0^e(s_m), \quad b_m^e \in \{0,1\}.
\]

### 3.3 Performance Analysis

In this section, we study the performance of the noisy binary image sensor for estimating the light intensity and analyze the influence of the noise. We show that the noisy binary sensor is still better than the traditional sensor in terms of dynamic range with a reasonable noise rate. To simplify our analysis and derive closed-form solutions, we assume that the light intensity field is piecewise-constant. Numerical results in Section 3.5 show that the conclusions hold for general linear models.

#### 3.3.1 The Cramér-Rao Lower Bound (CRLB) of the Estimation Variance

Same to Section 2.3.1 in the previous chapter, we assume that the light intensity field \( \lambda(x) \) is piecewise-constant, i.e., the interpolation kernel \( \phi(x) \) in (2.1) is the box function \( \beta(x) \). Thus (2.19) holds true here, and \( \{c_n\} \) can be estimated independently. In the following, we will focus on estimating \( c_0 \) from noisy binary measurements \( b^e \) as \([b_0^e, \ldots, b_{K-1}^e]^T\) and write \( c \) instead of \( c_0 \) for simplicity.

Let \( L^e_b(c) \) be the likelihood function of observing \( K \) noisy binary sensor measurement \( b^e \). Then,
\[
L^e_b(c) \triangleq P(B_m^e = b_m^e, 0 \leq m < K; c, p_e),
\]
\[
= \prod_{m=0}^{K-1} P(B_m^e = b_m^e; c, p_e), \quad (3.4)
\]
\[
= \prod_{m=0}^{K-1} p_0^e(c/K), \quad (3.5)
\]
where (3.4) is because each pixel counts the photons independently, and (3.5) is derived from (3.3) and (2.19). Denote by \( K_1 \) \((0 \leq K_1 \leq K)\) the number of “1”s in the noisy binary sequence \( b^e \). Then (3.5) becomes
\[
L^e_b(c) = \left( p_1^e(c/K) \right)^{K_1} \left( p_0^e(c/K) \right)^{K-K_1}. \quad (3.6)
\]

**Proposition 6.** The CRLB for estimating the light exposure value \( c \) from \( K \) noisy binary sensor measurements with threshold \( q \geq 1 \) and Bernoulli noise with parameter \( p_e \) is
\[
CRLB^e_{bin}(K, q, p_e) = \frac{Kp_0^e(c/K)p_1^e(c/K)}{(1 - p_e)^2 e^{-2c/K} (c/K)^{p_e - q}}. \quad (3.7)
\]

Proposition 7. For any $q$, if $p_e > 0$, $\lim_{K \to \infty} \frac{CRLB_{bin}(K, q, p_e)}{CRLB_{ideal}(K)} = \infty$.


### 3.3.2 Closed-form MLE Solution When $q = 1$

In what follows, we derive the closed-form MLE solution when the threshold is $q = 1$.

Lemma 3. The functions $p_0(s)$ in (2.12) and $p_0'(s)$ in (3.2) are strictly decreasing with respect to $s$.

Proof. When $q = 1$, $p_0(s) = e^{-s}$, so $\frac{d}{ds}p_0(s) = -e^{-s} \leq 0$. When $q > 1$,

$$\frac{d}{ds}p_0(s) = -\sum_{k=0}^{q-2} \frac{s^k}{k!} e^{-s} - \sum_{k=0}^{q-2} \frac{s^k}{k!} e^{-s} = -\frac{s^{q-1}}{(q-1)!} e^{-s} \leq 0.$$  

Thus for any $q$, we have $\frac{d}{ds}p_0(s) \leq 0$ and the equality is achieved at the boundary point $s = 0$. Thus, the function $p_0(s)$ is strictly decreasing. From (3.2), we have $p_0'(s) = (1 - p_e)p_0(s)$, therefore, $p_0'(s)$ is also strictly decreasing.

Given $K$ noisy binary measurements $b^e$, the MLE is to find the parameter $c$ which can maximize the likelihood function $L_b^e(c)$ in (3.6), i.e.,

$$\hat{c}_{ML}(b^e) \overset{\text{def}}{=} \arg \max_{0 \leq c \leq S} L_b^e(c) = \arg \max_{0 \leq c \leq S} \left( p_1^e(c/K) \right)^{K_1} \left( p_0^e(c/K) \right)^{K - K_1},$$

where the upper and lower bound are used to make the solution physically meaningful, i.e., the light exposure value cannot take negative value, and when the likelihood function $L_b^e(c) = (p_1^e(c/K))^K = (1 - p_0^e(c/K))^K$, i.e., under the case that $K_1 = K$, is monotonically increasing with respect to $c$, we cannot set the light exposure value to be $\infty$.

Lemma 4. When the threshold is $q = 1$, the solution to (3.8) is

$$\hat{c}_{ML}(b^e) = \begin{cases} 
-K \ln \frac{K_1}{K_1 - p_e}, & \text{if } 0 < K_1 < K, p_e < \min\{\frac{K_1}{K}, \frac{1}{2}\}, \\
0, & \text{if } K_1 = 0 \text{ or } 0 < K_1 < K, \frac{K_1}{K} \leq p_e < \frac{1}{2}, \\
S, & \text{if } K_1 = K
\end{cases}$$

where $K_1$ is the number of “1”s, and $K$ is the total number of pixels.

Proof. This lemma is a special case of Theorem 4 in Section 3.4 in this chapter, by setting threshold $q = 1$. The proof of Theorem 4 is in Appendix 3.A.3.
3.3 Performance Analysis

Figure 3.2: Performance comparisons of three different sensing schemes ("BIN", "IDEAL", and "SAT") over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the "IDEAL" scheme with no quantization; The solid line corresponds to the "SAT" scheme with a saturation point set at $c = 9130$; The four dashed lines correspond to the "BIN" schemes with $q = 1$ and $K = 2^{12}$ and different noise rates (from far right to left, $p_e = 0, 0.001, 0.005$ and $0.01$, respectively).

3.3.3 The Influence of the Noise on the Dynamic Range

We will focus our analysis on the case that the threshold is $q = 1$. We denote our binary sensing scheme as "BIN". We also compare our scheme with two other methods "IDEAL" and "SAT". In the "IDEAL", the pixel counts all the photons hitting on the pixel. The estimated light exposure value is just the number of the photons received by the pixel. The "SAT" scheme is similar to "IDEAL", except that it has a saturation point $C_{max}$. We use signal-to-noise ratios (SNRs) to measure the performance. The SNR is defined as

$$\text{SNR} = 10 \log_{10} \frac{c^2}{\mathbb{E}[(\hat{c} - c)^2]},$$

where $\hat{c}$ is the estimated light exposure value.

Let $y$ be the number of photons impinging on a pixel. Then for the "IDEAL" scheme, as shown in the previous chapter, the MLE is $\hat{c}_{\text{IDEAL}}(y) = y$, and $\text{SNR}_{\text{IDEAL}} = 10 \log_{10}(c)$. For the "SAT" method, the sensor measurement is $y_{\text{SAT}} \overset{\text{def}}{=} \min\{y, C_{\text{max}}\}$, and the estimator is $\hat{c}_{\text{SAT}}(y_{\text{SAT}}) = y_{\text{SAT}}$.

In Figure 3.2, we show the SNR performance for "IDEAL", "SAT", and "BIN" with different noise rates. The dash-dot line in the figure corresponds to the "IDEAL" scheme. The solid line is for the "SAT" method. The four dashed lines represent the "BIN" scheme with fixed oversampling factor $K = 2^{12}$, and different noise rates (from far right to left $p_e = 0, 0.001, 0.005$, and $0.01$, respectively). We can see that the larger the noise rate $p_e$, the worse the SNR performance of the "BIN" scheme. We can also notice that the noise has more influence on lower light intensities. For the large light intensities, the SNR is almost the same for the noiseless and noisy cases. This indicates that our binary sensing method is quite robust to noise.

We show the influence of the oversampling factor $K$ in the presence of noise. In Figure 3.3, the three "BIN" schemes with fixed noise rate $p_e = 0.005$, and different oversampling factors $K = 2^{14}, 2^{13},$ and $2^{12}$ are shown in dashed line from far right to left.
Figure 3.3: Performance comparisons of three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values $c$ (shown in logarithmic scale). The dash-dot line (in red) represents the “IDEAL” scheme with no quantization; The solid line (in blue) corresponds to the “SAT” scheme with a saturation point set at $c = 9130$; The three dashed lines correspond to the “BIN” schemes with $q = 1$ and noise rate $p_e = 0.005$ and different oversampling factors (from far right to left, $K = 2^{14}, 2^{13},$ and $2^{12}$, respectively).

respectively. We can see that when the light exposure value is small, increasing oversampling factor will deteriorate the performance. The intuition is that when the light exposure value is small, large oversampling factors will make the number of photons hitting on each pixel small. Thus, we can not distinguish the noise between the signal and get worse reconstruction results. But when the light exposure value is large enough, even if the oversampling factor is large, there are still a lot of photons absorbed by each pixel. Due to the high sensitivity of threshold $q = 1$ in low light intensities, we can still benefit from large oversampling factors.

3.4 Image Reconstruction

In the previous section, we derived the closed-form solution of the MLE for the piecewise-constant light intensity field model when $q = 1$. We extend the MLE to the general linear field model in (2.1) with arbitrary interpolation kernels and thresholds. We show that for general light field model, the log-likelihood function is concave when $q = 1$ even if there is noise. Thus we guarantee to find the optimal solution using iterative algorithms. For arbitrary thresholds, we show that when there is noise, under the piecewise-constant light field model, the likelihood function is strictly pseudoconcave. This guarantees that there is a unique global maximum that can be found using iterative algorithms. For the general light intensity field model, we first assume the light field is piecewise-constant and get an initial result. Then the reconstructed image can be refined using Newton’s method.

3.4.1 Image Reconstruction by the MLE

From (2.1), we know that the estimation of the light intensity field $\lambda(x)$ is equivalent to the reconstruction of the parameters $\{c_n\}$. In (2.8), we showed the relation between the
light exposure values \( \{s_m\} \) and expansion coefficients \( \{c_n\} \). As done in Section 2.4.1 in the previous chapter, we use the matrix form

\[
s = Gc,
\]

where \( s = [s_0, s_1, \ldots, s_{M-1}]^T, \quad c = [c_0, c_1, \ldots, c_{N-1}]^T, \) and \( G \) is an \( M \times N \) matrix denoting the combination of upsampling (by \( K \)) and filtering (by \( g_m \)). Each entry of \( s \) is

\[
s_m = e_m^T Gc,
\]

where \( e_m \) is the \( m \)th standard Euclidean basis vector.\(^1\)

Similar to our derivations in (3.4) and (3.5), the likelihood function given \( M \) noisy binary measurements \( \hat{b} \) can be written as

\[
\mathcal{L}_b^e(c) = \prod_{m=0}^{M-1} P(B_m^e = b_m^e; s_m) = \prod_{m=0}^{M-1} p_{s_m}^e(e_m^T Gc), \tag{3.12}
\]

where (3.12) follows from (3.3) and (3.11). We also define the log-likelihood function as

\[
\ell_b^e(c) \overset{\text{def}}{=} \log \mathcal{L}_b^e(c) = \sum_{m=0}^{M-1} \log p_{s_m}^e(e_m^T Gc). \tag{3.13}
\]

Given \( \hat{b} \), the MLE is the parameter that maximizes \( \mathcal{L}_b^e(c) \), or \( \ell_b^e(c) \). Specifically,

\[
\hat{c}_{\text{ML}}(\hat{b}) \overset{\text{def}}{=} \arg \max_{c \in [0,S]^N} \mathcal{L}_b^e(c) = \arg \max_{c \in [0,S]^N} \ell_b^e(c) \tag{3.14}
\]

The constraint \( c \in [0,S]^N \) means that every parameter \( c_n \) should satisfy \( 0 \leq c_n \leq S \), for some preset maximum value \( S \).

As shown in Section 3.3, when the filter \( g_m \) is a box filter, \( e.g. \), the light intensity is piecewise-constant, we can estimate the \( \{c_n\} \) separately. For each \( c_n \), the likelihood function \( \mathcal{L}_b^e(c) \) and log-likelihood function \( \ell_b^e(c) \) only have one variable and are shown in Figure 3.4. We can see that when the threshold is \( q = 1 \), the presence of noise affects the log-likelihood function only slightly. In particular, the log-likelihood function is still concave as in the noiseless case. In the following, we will show that the log-likelihood function defined in (3.13) is always concave, when the threshold is \( q = 1 \).

**Theorem 3.** When the threshold is \( q = 1 \), for arbitrary noisy binary sensor measurements \( \hat{b} \), the log-likelihood function \( \ell_b^e(c) \) defined in (3.13) is concave on the domain \( c \in [0,S]^N \).

**Proof.** When \( q = 1 \), \( b_m^e = 0 \), according to (3.2) and (2.12),

\[
\log p_{s_m}^e(s) = \log((1 - p_e)e^{-s}) = \log(1 - p_e) - s,
\]

which is a concave function.

When \( q = 1 \), \( b_m^e = 1 \),

\[
\log p_{s_m}^e(s) = \log(1 - (1 - p_e)e^{-s}) = \log(e^s + p_e - 1) - s,
\]

\(^1\)Here we use zero-based indexing. Thus, \( e_0 \overset{\text{def}}{=} [1,0,\ldots,0]^T, \quad e_1 \overset{\text{def}}{=} [0,1,\ldots,0]^T, \) and so on.
do concave function with respect to $c$ is concave.

The second derivative of this function is still concave and the composition of a concave function with a linear mapping defined in (3.8) is still concave, we conclude that the log-likelihood function $\ell_b(c)$ defined in (3.13) is concave.

According to Theorem 3, we can find the optimal solution using iterative numerical methods like the reconstruction method in Chapter 2, even if there is noise.

Unfortunately, Theorem 3 does not hold true for threshold $q > 1$. Figure 3.5 shows the likelihood function $L_b(c)$ and log-likelihood function $\ell_b(c)$ for one of the $c_n$ in a piecewise-constant light intensity when the threshold is $q = 2$. We can see that neither the likelihood function nor the log-likelihood function is concave. In the following, we show that the likelihood function and log-likelihood function for the piecewise-constant light intensity are strictly pseudoconcave with respect to $c$ for an arbitrary threshold $q$. Therefore, we ensure that the local optimal solution of (3.14) is the global optimal solution.

**Lemma 5.** For a differentiable function $h(x)$ on domain $D$, if the first derivative $\frac{dh(x)}{dx} < 0$ or $\frac{dh(x)}{dx} > 0$ holds for all $x \in D$, then the function $h(x)$ is strictly pseudoconcave.

**Proof.** From the definition of strictly pseudoconcave function in Appendix B [93], we need to prove that for all $x_1, x_2 \in D$, $x_1 \neq x_2$, $h(x_1) \leq h(x_2) \Rightarrow \frac{dh(x_1)}{dx} > 0$.

When the first derivative $\frac{dh(x)}{dx} < 0$ holds for all $x \in D$, for any $h(x_1) \leq h(x_2)$ and $x_1 \neq x_2$, we will have $x_1 > x_2$. Since $\frac{dh(x_1)}{dx} < 0$, $(x_2 - x_1) < 0$, we have $\frac{dh(x_1)}{dx}(x_2 - x_1) > 0$. The case $\frac{dh(x)}{dx} > 0$ can be proved in the same way.

**Theorem 4.** For the piecewise-constant light intensity field, the likelihood function $L_b(c)$ defined in (3.8), and the log-likelihood function $\ell_b(c)$ defined in (3.13) are strictly pseudoconcave function with respect to $c \in (0, S]$ for an arbitrary threshold $q$. The solution
Figure 3.5: The likelihood and log-likelihood functions for piecewise-constant light fields under the noisy case. (a) The likelihood functions $L_b(c)$, defined in (3.6), under different choices of noise rate $p_e = 0, 0.005,$ and $0.01$, respectively. (b) The corresponding log-likelihood functions. The parameters are set as: threshold $q = 2$, oversampling factor $K = 20$, number of ones $K_1 = 4$.

for (3.8) is

$$\hat{c}_{ML}(b^T) = \begin{cases} c^*, & \text{if } 0 < K_1 < K, p_e < \min\{K_1/K, 1/2\}, \\ 0, & \text{if } K_1 = 0 \text{ or } 0 < K_1 < K, K_1/K \leq p_e < 1/2, \\ S, & \text{if } K_1 = K, \end{cases}$$

(3.15)

where $c^*$ satisfies $\frac{dL_b(c^*)}{dc} = 0$.


In the case of a light intensity field that is not piecewise-constant, the log-likelihood function is not pseudoconcave and not even quasiconcave (See Appendix B of [93]). For instance, Figure 3.6 shows the log-likelihood function for a general light intensity field under the noisy case. The expansion coefficients are $c = [100, 110]^T$. The matrix $G$ is a randomly generated $20 \times 2$ matrix, the oversampling factor is $K = 10$, and the threshold is $q = 5$. We can see that the superlevel set of the $\ell_b^0(c)$ is not convex set, which shows the log-likelihood function is not even quasiconcave according to the definition of quasiconcave function given in Appendix B of [93].

3.4.2 Iterative Algorithms

When $q = 1$, the log-likelihood function is concave and we can use the reconstruction algorithm in Chapter 2. So in this subsection, we will mainly focus on the reconstruction problem for the binary sensor with threshold $q > 1$. But algorithms proposed in this subsection can also find the optimal solution for threshold $q = 1$.

Piecewise-constant Light Intensity Field

We first propose an algorithm to find the optimal solution for the piecewise-constant light intensity field. As shown in the previous section, we can estimate each expansion
coefficient $c_n$ independently. From (3.15), to find the solution of the MLE, we need to compute $c^*$, such that $\frac{dL_b(c^*)}{dc} = 0$. According to (3.23), this is equal to find $c^*$ that satisfies $(1 - p_e)p_0(c^*/K) - \frac{K - K_b}{K} = 0$. Since the likelihood function is strictly pseudoconcave, there is only a unique solution. We can get $c^*$ using the bisection method, shown in Algorithm 1.

**Algorithm 1** Solve MLE for piecewise-constant light field by bisection method

Initialize: set $l := 0$, $u := S$, $\epsilon := 10^{-5}$, $f(c) \overset{\text{def}}{=} (1 - p_e)p_0(c/K) - \frac{K - K_b}{K}$

Loop: while $|u - l| > \epsilon$

If $f(l) > 0$ return $c^* = l$, else if $f(u) < 0$ return $c^* = u$

Else $\{h := (u + l)/2, \text{ if } f(h) > 0 \text{ then } u = h \text{ else } l = h\}$

**General Linear Light Intensity Field**

For the general light field model, as shown in the previous section, the likelihood function or the log-likelihood function is not even quasiconcave; therefore, we can not ensure that the bisection method converges to the optimal solution. We present an algorithm which can still achieve good estimation performance. We first assume that the light intensity field is piecewise-constant and find an initial solution using the bisection algorithm. Then we use a projected Newton’s method to obtain a refined solution.

Denote by $c^{(k)}$ the estimate of the unknown parameter $c$ at the $k$th step. The proposed Newton’s method updates $c^{(k)}$ as follows. We first compute

$$\nabla L_b^c(c^{(k)}) \text{ and } H^c L_b^c(c^{(k)}),$$

which are respectively the gradient and the Hessian matrix of the log-likelihood function defined in (3.13).
Algorithm 2 The projected Newton’s method for finding the MLE solution.

Require: A set of quantized sensor measurements \( \mathbf{b} \in \mathbb{R}^M \).
Ensure: Estimated expansion coefficients \( \hat{\mathbf{c}} \in \mathbb{R}^N \) of the light intensity field \( \lambda(x) \).

Estimated a starting point of \( \mathbf{c}^{(0)} \) by assuming that the light intensity field is piecewise-constant using Algorithm 1.
Initialize the iteration number: \( k := 0 \)
repeat
    Compute the gradient \( \nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}) \), evaluated at \( \mathbf{c}^{(k)} \).
    Use the conjugate gradient algorithm [94] to compute the update direction
    \[ d \equiv [H_{\mathbf{b}}(\mathbf{c}^{(k)})]^{-1}\nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}). \]
    Use backtracking line search to determine the optimal steps ize
    \[ \gamma^* = \arg \max_{0 \leq \gamma \leq 1} \ell_{\mathbf{b}}(P_{\mathcal{D}}(\mathbf{c}^{(k)} + \gamma d)). \]
    Update the estimate as \( \mathbf{c}^{(k+1)} = P_{\mathcal{D}}(\mathbf{c}^{(k)} + \gamma^* d) \).
    \( k = k + 1 \)
until \( ||\nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)})|| \leq \varepsilon \) or \( d^T \nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}) \leq 2\varepsilon \), where \( \varepsilon = 10^{-5} \) is a predetermined precision threshold.
return \( \hat{\mathbf{c}} = \mathbf{c}^{(k)} \).

We then construct a one-parameter function
\[ \hat{\mathbf{c}}(\gamma) = P_{\mathcal{D}} \left( \mathbf{c}^{(k)} + \gamma [H_{\mathbf{b}}(\mathbf{c}^{(k)})]^{-1}\nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}) \right), \]
where \( 0 \leq \gamma \leq 1 \) is a small stepsize, and \( P_{\mathcal{D}} \) is the projection onto the search domain \( \mathcal{D} \equiv [0, S]^N \) of the unknown parameters \( \mathbf{c} \). We apply \( P_{\mathcal{D}} \) to ensure that all estimates of \( \mathbf{c} \) lie in the search domain.
Finally, we choose the stepsiz e \( \gamma \) to maximize the log-likelihood function, i.e.,
\[ \gamma^* = \arg \max_{0 \leq \gamma \leq 1} \ell_{\mathbf{b}}(\hat{\mathbf{c}}(\gamma)), \]
(3.16)
and
\[ \mathbf{c}^{(k+1)} = \hat{\mathbf{c}}(\gamma^*). \]
In practice, the 1-D maximization problem in (3.16) can be solved by backtracking line search.

We summarize in Algorithm 2 the main steps of the projected Newton’s method as well as its stoping criteria.
At every step of the Newton’s method, the computationally most expensive part is to find the update direction
\[ [H_{\mathbf{b}}(\mathbf{c}^{(k)})]^{-1}\nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}), \]
which, by using the conjugate gradient algorithm [94], boils down to evaluating the gradient \( \nabla \ell_{\mathbf{b}}(\mathbf{c}^{(k)}) \) and the matrix-vector products \( H_{\mathbf{b}}(\mathbf{c}^{(k)}) \mathbf{p}_k \) for some vectors \( \{\mathbf{p}_k\} \).
In what follows, we describe how to obtain these quantities efficiently using well-known signal processing operations.
Lemma 6. The gradient and the Hessian of the log-likelihood function are

\[ \nabla \ell_b(c) = G^T [D_{b_0}'(s_0), D_{b_1}'(s_1), \ldots, D_{b_{M-1}}'(s_{M-1})]^T \]  \tag{3.17} 

and

\[ H\ell_b(c) = G^T \text{diag} \{ D_{b_m}''(s_m) \} G, \]  \tag{3.18} 

respectively. In the expressions above, \( G \) is the matrix defined in (3.10), \( s_m \) is defined in (3.11), and \( D_{b}^c(s), D_{b}''(s) \) denote the first- and second-order derivatives of \( \log p_{eb}(s) \), respectively, i.e.,

\[ D_{b}^c(s) = \frac{\partial^2}{\partial s^2} \log p_{eb}(s) \quad \text{and} \quad D_{b}''(s) = \frac{\partial^2}{\partial s^2} \log p_{eb}(s) - (\frac{\partial}{\partial s} \log p_{eb}(s))^2. \]

In practice, the size of our image reconstruction problem is very large. For example, current memory chip technology allows us to build binary-quantized image sensors packing \( 10^9 \sim 10^{10} \) pixels (i.e., gigapixels) per chip [95]. Such large problem dimensions make it infeasible to directly implement the matrix operations in (3.17) and (3.18). Fortunately, the matrix \( G \) used in both formulae are highly structured, and can be implemented as upsampling followed by filtering and scalar multiplication (see the optics part in Figure 2.2 for details). It follows that we can compute the gradient and the Hessian by using efficient signal processing operations as shown in Figure 3.7. Note that \( G^T \), the transpose of \( G \), is implemented by “flipping” all the operations in \( G \).
3.5 Numerical Results

This section shows some numerical results on synthesized 1-D signals and 2-D images. The results validate the theoretical analysis and the effectiveness of our proposed image reconstruction algorithm.

3.5.1 1-D Synthetic Signals

Given expansion coefficients \( \{ c_n \} \) shown as blue dots in the Figure 3.8(a), and the interpolation filter \( \varphi(x) \) which is the cubic B-spline function \( \beta_3(x) \), we generate a 1-D light field \( \lambda(x) \), i.e., the blue line. As shown in Figure 3.8(a), \( \lambda(x) \) is a linear combination of the shifted kernel.

We first set the threshold \( q = 1 \) and the oversampling factor \( K = 1024 \). Applying the proposed MLE-based algorithm in Section 3.4.2, the reconstructed light intensity fields with the values of noise rate \( p_n \) are 0, 0.1, and 0.2 are shown in Figure 3.8(a), Figure 3.8(b), and Figure 3.8(c) (in red), respectively. For comparison, the ground truth is given by the blue solid curve. We can see that when the noise rate increases, the performance becomes slightly worse. This obeys the performance analysis of Section 3.3.3, and shows the robustness of our proposed binary sensing scheme.

Figure 3.9 shows the results with different oversampling factors. The threshold is still 1, and the noise rate is \( p_n = 0.1 \). The oversampling factors of Figure 3.9(a), Figure 3.9(b), and Figure 3.9(c) are 256, 2048, and 8192, respectively. We can see that increasing the oversampling factor the performance initially improves and then worsens.
Figure 3.9: Binary sensing and reconstructions of 1-D light fields under noise with different oversampling factors. (a), (b), and (c) show the reconstruction results with different oversampling factor $K = 256, 2048, \text{and } 8192$ respectively. The threshold is $q = 1$, and the noise rate is $p_e = 0.1$.

validating the performance analysis of Section 3.3.3.

Then, in Figure 3.10, we show the reconstruction results for threshold $q = 3$, under different noise rates $p_e = 0, 0.1, 0.2$. We can see that the noise has a big influence on the estimation performance when the light intensity is low. The intuition is that due to the large threshold and noise rate, when the light intensity is low, the sensor can not distinguish the noise and the signal.

3.5.2 2-D Synthetic Images

Consider a 2-D light intensity field as shown in Figure 3.11(a). The values of the light intensity are in the range $[500, 2.5 \times 10^4]$. We simulate the acquisition of this light intensity field using different noisy binary sensors. We consider the values of thresholds set to $q = 1$ and $q = 3$. For the noise rate, we consider the cases $p_e = 0, p_e = 0.1,$ and $p_e = 0.2$. The spatial oversampling factor of the binary sensor is set to $8 \times 8$, and the temporal oversampling factor is 128 (i.e., 128 independent frames). Similarly to our previous experiment on 1-D signals, we use a cubic B-spline kernel along each of the spatial dimensions. The reconstruction results for thresholds $q = 1$ and $q = 3$ for the different noise cases are shown in Figure 3.11 and Figure 3.12. The MSE of the reconstruction results are shown in Table 3.1. We can see that for $q = 1$, the MSE of the noise case $p_e = 0.1$ is better than that of noiseless case. Since the noise can not change the binary measurements from “1” to “0”, the influence of noise when the light intensity is large is small. For a single experiment, there is a chance that the noise improves the estimation of large light intensity values. From the figures, we can see that our binary sensing scheme is quite robust to noise. When threshold $q = 1$, we can hardly notice the
3.6 Conclusions

We presented a theoretical analysis of the binary image sensor under the noise scenario. The noise is modeled as an additive Bernoulli noise with a known parameter, and it can only change the binary output from “0” to “1”. We showed that the noise has limited influence on the performance of the sensor and would slightly deteriorate the dynamic range. Under the noisy case, increasing the oversampling factor will first improve and then worsen the performance of the reconstruction. However, the algorithm is still able to reconstruct a high dynamic range image (the effect of noise remains slightly visible). We used the MLE to estimate the light intensities. When the threshold is a single photon, the log-likelihood function is still concave and the optimal solution can

\[
\text{Table 3.1: The MSE for different thresholds under different noise rates.}
\]

<table>
<thead>
<tr>
<th>MSE</th>
<th>( p_e = 0 )</th>
<th>( p_e = 0.1 )</th>
<th>( p_e = 0.2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q = 1 )</td>
<td>( 2.716 \times 10^5 )</td>
<td>( 2.61 \times 10^5 )</td>
<td>( 2.767 \times 10^5 )</td>
</tr>
<tr>
<td>( q = 3 )</td>
<td>( 3.406 \times 10^5 )</td>
<td>( 5.848 \times 10^5 )</td>
<td>( 7.727 \times 10^5 )</td>
</tr>
</tbody>
</table>

The presence of noise, although 10% or 20% of the binary measurements are contaminated by the noise. When threshold \( q = 3 \), we can see noise in the dark regions. And this follows the analysis in the previous section.
Figure 3.11: Binary sensing and reconstructions of 2-D light fields with different noise levels. (a) The original light field. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0.0, 0.1, \text{and } 0.2$, respectively. The spatial oversampling factor is $8 \times 8$, the temporal oversampling factor is $128$, and the threshold is $q = 1$.

be computed using iterative algorithms. But this does not hold true for thresholds larger than “1”. For the piecewise-constant light field, we showed that we could estimate each expansion coefficient independently, and the likelihood and log-likelihood functions are strictly pseudoconcave for arbitrary thresholds. Thus, the optimal solution can be found using the bisection method. For the general light intensity field, the likelihood function is not even quasiconcave. We solved this case by first assuming that the light field is piecewise-constant, and then using Newton’s method to refine the approximated solution. The experimental results demonstrate our theoretical analysis and prove the effectiveness of our proposed algorithm.
3.6 Conclusions

Figure 3.12: Binary sensing and reconstructions of 2-D light fields with different noise levels. (a) The original light field. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1, \text{and } 0.2$ respectively. The spatial oversampling factor is $8 \times 8$, the temporal oversampling factor is 128, and the threshold is $q = 3$. 

(a)

(b)  (c)
3.A Appendix

3.A.1 The CRLB of Noisy Binary Sensors

We first compute the Fisher information, defined as

\[ I(c) = \mathbb{E} \left[ -\frac{\partial^2}{\partial c^2} \log L_b(c) \right] \]

Using (3.6), we get

\[
I(c) = \mathbb{E} \left[ -\frac{\partial^2}{\partial c^2} \left( K_1 \log p_1(c/K) + (K - K_1) \log p_0(c/K) \right) \right] = \\
\mathbb{E} \left[ \frac{K_1 (p_0''(c/K)p_1(c/K) + p_0'(c/K)^2)}{K^2 p_1(c/K)^2} - \frac{(K - K_1) (p_0''(c/K)p_0(c/K) - p_0'(c/K)^2)}{K^2 p_0(c/K)^2} \right],
\]

(3.19)

where \( p_0'(x) = \frac{d}{dx} p_0(x) \) and \( p_0''(x) = \frac{d^2}{dx^2} p_0(x) \) are the first and second order derivative of \( p_0(x) \), respectively. In reaching (3.19), we have also used the fact that \( p_1(x) = 1 - p_0(x) \) and thus \( p_0''(x) = -\frac{d}{dx} p_1(x) \) and \( p_0''(x) = -\frac{d^2}{dx^2} p_1(x) \).

Note that \( K_1 = \sum_{0 \leq m < K} b_m \) is a binomial random variable, and thus its mean can be computed as

\[
\mathbb{E}[K_1] = K p_1(c/K) = K(1 - p_0(c/K)).
\]

On substituting the above expression into (2.51), the Fisher information can be simplified as

\[
I(c) = \frac{p_0''(c/K)p_1(c/K) + p_0'(c/K)^2}{K p_1(c/K)} - \frac{p_0''(c/K)p_0(c/K) - p_0'(c/K)^2}{K p_0(c/K)} \\
= \frac{p_0'(c/K)^2}{K p_0(c/K)p_1(c/K)}.
\]

(3.20)

Using the definition of \( p_0(x) \) in (2.12) and \( p_0'(x) \) in (3.2), the derivative in the numerator of (3.20) can be computed as

\[
p_0' = (1 - p_e)p_0(x) = -(1 - p_e)e^{-x} \frac{x^{q-1}}{(q - 1)!}.
\]

(3.21)

Finally, since \( \text{CRLB}_{\text{bin}}(K, q) = 1/I(c) \), we reach (3.7) by substituting (3.21) into (2.52), and after some straightforward manipulations.
3.A.2 Proof of Proposition 7

We define a function \( h_q(x) \) as, \( h_q(x) \equiv \frac{x^{q-1}e^{-x}}{(q-1)!} \) and using the definition of \( p_0^*(x) \) in (3.3), we have

\[
\text{CRLB}_\text{bin}^e(K, q, p_e) = \frac{K p_0^*(c/K)p^*_c(c/K)}{(1 - p_e)^2 e^{-2c/K} (c/K)^2p_0^*(c/K)}
= \frac{K((1 - p_e)p_0(c/K)(1 - (1 - p_e)p_0(c/K))}{(1 - p_e)^2 h^2_q(c/K)}
= \frac{K((1 - p_e)p_0(c/K) - (1 - p_e)^2 p_0^*(c/K))}{(1 - p_e)^2 h^2_q(c/K)}
\geq \frac{K}{h^2_q(c/K)} (p_0(c/K) - p_0^*(c/K))
= \text{CRLB}_\text{bin}^e(K, q, 0)
\]

As in Chapter 2, we know that when \( q > 2 \), \( \text{lim}_{K \to \infty} \text{CRLB}_\text{bin}^e(K, q, 0)/\text{CRLB}_{\text{ideal}}(K) = \infty \). Then we have when \( q > 2 \), \( \text{lim}_{K \to \infty} \text{CRLB}_\text{bin}^e(K, q, p_e)/\text{CRLB}_{\text{ideal}}(K) = \infty \).

When \( q = 1 \), \( h_q(c/K) = e^{-c/K}, p_0(c/K) = e^{-c}/K \), then

\[
\text{CRLB}_\text{bin}^e(K, q, p_e) = \frac{K}{(1 - p_e)^2 e^{-2c/K}} \left( (1 - p_e)e^{-c/K} - (1 - p_e)^2 e^{-2c/K} \right)
= \frac{K}{1 - p_e} \left( e^{c/K} - 1 \right) + \frac{p_e}{1 - p_e}
\]

when \( K \) goes to \( \infty \), the first term converges to \( c/(1 - p_e) \), the second term goes to \( \infty \), so when \( q = 1 \), we also have \( \text{lim}_{K \to \infty} \text{CRLB}_\text{bin}^e(K, q, p_e)/\text{CRLB}_{\text{ideal}}(K) = \infty \).

3.A.3 Proof of Theorem 4

From Appendix B in [93] and Lemma 5, we know that to prove the likelihood function \( \mathcal{L}_b^c(c) \) is strictly pseudoconcave, we just need to prove that the derivative of \( \mathcal{L}_b^c(c) \) is always positive or negative, or \( \mathcal{L}_b^c(c) \) is quasiconcave and every critical point is a strict local maximum for \( \mathcal{L}_b^c(c) \). To prove \( \mathcal{L}_b^c(c) \) is quasiconcave, as shown in [96], it is sufficient to show that \( \frac{d\mathcal{L}_b^c(c)}{dc} = 0 \) when \( c = c^* \), \( \frac{d^2\mathcal{L}_b^c(c)}{dc^2} > 0 \) when \( c < c^* \), and \( \frac{d^2\mathcal{L}_b^c(c)}{dc^2} < 0 \) when \( c > c^* \).

We define a sequence of functions \( \{r_q(x)\}_{q \in \mathbb{Z} \cup \{\infty\}} \) as

\[
r_q(x) = \begin{cases} \frac{x^q}{q!}, & \text{if } 0 \leq q < \infty; \\ 0, & \text{if } q < 0 \text{ or } q = \infty. \end{cases}
\]

Then \( \frac{dr_q(x)}{dx} = r_{q-1}(x) \) for all \( q \in \mathbb{Z} \cup \{\infty\} \). According to (2.12), (3.2), and (3.22), we have

\[
\frac{dp_1^*(s)}{ds} = \frac{d(p_1(s) + p_0(s)p_e)}{ds} = (1 - p_e)^{-1}e^{-s},
\]

and

\[
\frac{dp_0^*(s)}{ds} = \frac{d(p_0(s)(1 - p_e))}{ds} = -(1 - p_e)^{-1}e^{-s},
\]
where we ignore the function argument \( s \) in \( r_{q-1}(s) \) for simplicity.

If \( K_1 = 0 \), following from (3.8),

\[
\frac{dL_b^*(c)}{dc} = \frac{d(p_0^*(c/K))^K}{dc} = -(p_0^*(c/K))^{K-1}(1-p_e)r_{q-1}(c/K)e^{-c/K} < 0,
\]

for \( c \in (0,S] \). So \( L_b^*(c) \) is pseudoconcave, and \( \hat{c}_{ML}(b^*) = 0 \). Although the domain \( c \) is \( (0,S] \) that does not contain 0. But here the optimal solution is something quite close to 0, so we set the estimation value as 0.

If \( K_1 = K \),

\[
\frac{dL_b^*(c)}{dc} = \frac{d((1-p_0^*(c/K)))^{K_1}(p_0^*(c/K))^{K-K_1}}{dc} = (1-p_0^*(c/K))^{K-1}(1-p_e)r_{q-1}(c/K)e^{-c/K} > 0
\]

for \( c \in (0,S] \). Thus \( L_b^*(c) \) is quasiconcave, and the \( \hat{c}_{ML}(b^*) = S \).

If \( 0 < K_1 < K \),

\[
\frac{dL_b^*(c)}{dc} = \frac{d((1-p_0^*(c/K)))^{K_1}(p_0^*(c/K))^{K-K_1}}{dc} = (K_1/K)(1-p_0^*(c/K))^{K_1-1}(1-p_e)r_{q-1}(c/K)e^{-c/K}(p_0^*(c/K))^{K-K_1}
\]

\[
- (1-K_1/K)(1-p_0^*(c/K))^{K_1-1}(p_0^*(c/K))^{K-K_1-1}(1-p_e)r_{q-1}(c/K)e^{-c/K}
\]

\[
= (1-p_e)(1-p_0^*(c/K))^{K_1-1}(p_0^*(c/K))^{K-K_1-1}(1-p_e)r_{q-1}(c/K)e^{-c/K}(p_0^*(c/K))^{K-K_1-1}(1-p_e)r_{q-1}(c/K)e^{-c/K}(1-p_0^*(c/K))^{K_1}/((1-p_e)p_0 - (K-K_1)/K), \tag{3.23}
\]

where we have omitted the function argument \( c/K \) in \( p_0^*(c/K) \), \( p_0(c/K) \), and \( r_{q-1}(c/K) \) for notational simplicity.

If \( p_e < \min\{K_1/K, 1/2\} \), since \( p_0(c/K) \) is strictly decreasing as shown in Lemma 3, it has a single intersection with \((1-K_1/K)/(1-p_e)\) at \( c^* \). Then when \( c < c^* \), \( p_0(c/K) > K-K_1/K(1-p_e) \), so \( \frac{dL_b^*(c)}{dc} > 0 \), when \( c > c^* \), \( p_0(c/K) = K-K_1/K(1-p_e) \), so \( \frac{dL_b^*(c)}{dc} < 0 \). From this we know that \( L_b^*(c) \) is quasiconcave [96]. Also \( c^* \) is a strict local maximum point. Thus \( L_b^*(c) \) is pseudoconcave according to [93]. Then the \( \hat{c}_{ML}(b^*) = c^* \).

If \( K_1/K \leq p_e < 1/2 \), \( \frac{dL_b^*(c)}{dc} < 0 \), so \( L_b^*(c) \) is pseudoconcave and \( \hat{c}_{ML}(b^*) = 0 \).

Since \( l(x) = \log(x) \) is a differentiable function and \( \frac{d(l(x))}{dx} = 1/x > 0 \), and \( L_b^*(c) \) is pseudoconcave, according to Appendix B [93], the log-likelihood function \( L_b^*(c) \) is also strictly pseudoconcave.
Chapter 4

Thresholds Values and Patterns

4.1 Introduction

One of the important parameters in our proposed binary image sensor is the threshold $q$. If the number of photons received by the pixel is larger or equal to $q$, the pixel value will be “1”, otherwise it will be “0”. In Chapter 2, we did some theoretical analysis on thresholds. We studied the asymptotic behavior of the binary image sensor with different thresholds. We showed that when $q = 1$, with a large oversampling factor, the binary image sensor acted like an ideal photon counting image sensor. But if $q > 1$ and the oversampling factor is very large, the performance of the binary image sensor will be far from the photon counting image sensor.

In this chapter, we will give a more detailed study of thresholds’ influence on the performance of the binary image sensor. The three main contributions in this chapter are as follows:

1. **Influence of the threshold:** In Section 4.2, we study the influence of the threshold. We show that smaller thresholds work better for small light intensities in terms of the mean squared error (MSE) or the Cramér-Rao lower bound (CRLB), and larger thresholds work better for large light intensities. The dynamic ranges of large thresholds are worse than those of smaller thresholds for given oversampling factors. We also observe that the MSE of our maximum likelihood estimator (MLE) is quite close to the CRLB.

2. **Optimal threshold pattern design:** In Section 4.3, using the fact that the MSE of our MLE is quite close to the CRLB, we design the optimal threshold pattern through minimizing the average CRLB instead of the MSE, due to the fact that we have a closed-form formula for the CRLB. We show that the performance of the designed optimal threshold pattern is better than a single threshold scheme given the range of light intensity values.

2. **Image reconstruction:** In Section 4.4, we propose to use MLE for estimating the light intensity field. We prove that the log-likelihood function is concave under arbitrary threshold patterns. So it is guaranteed that we can find the optimal solution of the MLE using iterative numerical algorithms.

Section 4.5 shows numerical results on both synthesized data and real images. The results verify our analysis on the influence of thresholds and efficacy of our optimal threshold pattern design and reconstruction algorithms.

Similar to previous chapters, to simplify the notation in this chapter, we focus our
discussion on a one-dimensional (1-D) sensor array. All the results can be easily extended
to the 2-D case.

4.2 Influence of Thresholds

In this section, we study the influence of thresholds on the performance of the binary
image sensor in estimating the light intensity. We have the same setup as in Section 2.3
of Chapter 2. We assume that the light intensity field is piecewise-constant, i.e., the
interpolation kernel \( \varphi(x) \) in (2.1) is the box function \( \beta(x) \). Under this model, the relation
between the light exposure values \( \{s_m\} \) and parameters \( \{c_n\} \) is

\[
s_m = \frac{c_n}{K}, \quad \text{for } nK \leq m < (n+1)K. \tag{4.1}
\]

From (4.1), we can see that \( \{c_n\} \) can be also considered as the total light exposure
value for \( K \) binary pixels. As shown in Section 2.3, in this case, we can estimate the
parameters \( \{c_n\} \) independently. Here we also focus on estimating \( c_0 \) from the block of
binary measurement \( \mathbf{b} \overset{\text{def}}{=} [b_0, \ldots, b_{K-1}]^T \). For notational simplicity, we will drop the
subscript in \( c_0 \) and use \( c \) instead.

We denote our oversampled binary sensing scheme as the “BIN” scheme. Then the
CRLB of estimating the total light exposure value \( c \) from \( K \) binary sensor measurements
with the threshold \( q \), \( \text{CRLB}_{\text{bin}}(K, q) \), is given in (2.23). We call the sensing scheme that
can perfectly record the number of photons hitting on the pixel, the “IDEAL” scheme.
The CRLB of this scheme is \( \text{CRLB}_{\text{ideal}}(K) = c \) given in (2.24). Figure 4.1 shows
the relation between the total light exposure value \( c \) and the CRLB of “BIN” schemes
with different thresholds \( q = 1, 3, 5 \) and the “IDEAL” scheme. We can see that the
CRLB of “IDEAL” scheme is the lower bound of “BIN” schemes. When \( q = 1 \), the
smaller the \( c \), the better the CRLB. If \( q > 1 \), and we increase the total light exposure
value \( c \), the CRLB first becomes small, then grows. When the threshold \( q \) is small,
it has a low CRLB in the low light intensities. When \( q \) is large, it has low CRLB in
the high light intensities. From Figure 4.1(a) and Figure 4.1(b), we can say that the
above conclusion holds for different oversampling factors \( K \). Then Figure 4.2 plots the
relationship between the oversampling factor \( K \) and the CRLB for “BIN” schemes and
the “IDEAL” scheme. The curves obey the theoretical analysis in Section 2.3. When
\( q = 1 \), the larger oversampling factor \( K \), the better the performance. When \( q > 1 \), larger
oversampling factors deteriorate the performance.

We define the (MSE) of the estimators as,

\[
\text{MSE} \overset{\text{def}}{=} \mathbb{E}[ (\hat{c} - c)^2], \tag{4.2}
\]

where \( \hat{c} \) is the estimation of the total light exposure value we obtain from each of the
sensing schemes. Then the SNR defined in (2.33) is

\[
\text{SNR} \overset{\text{def}}{=} 10 \log_{10} \frac{\hat{c}^2}{\text{MSE}}.
\]

As in Section 2.3, the MSE of the “IDEAL” scheme is \( \text{MSE}_{\text{ideal}} = c \). For “BIN”
schemes, we first use the MLE given in (2.28) to get \( \hat{c} \), then compute the MSE according
to (4.2). Figure 4.3 shows the relationship between the total light exposure value \( c \)
and the MSE of “BIN” schemes with different thresholds \( q \) and the “IDEAL” scheme.
4.2 Influence of Thresholds

Figure 4.1: The relation between the total light exposure value $c$ and the CRLB of “BIN” schemes with different threshold $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme. (a) Oversampling factor $K = 1024$. (b) Oversampling factor $K = 8096$.

From the figure, we can see that all the conclusions made for the CRLB in the previous paragraph also hold for the MSE. When we compare the value of the CRLB and the corresponding MSE, they are quite close, so we can use the CRLB as an approximation for the MSE. This also indicates that our MLE almost achieves the CRLB.
Figure 4.2: The relation between the oversampling factor $K$ and the CRLB of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme.

We plot the relationship between the total light exposure value $c$ and the SNR of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme when oversampling by a factor $K = 8096$ in Figure 4.4. Like in Section 2.3, we denote $\text{SNR}_{\text{min}}$ the minimum acceptable SNR in a given application. Then the dynamic range (DR) of a sensor is the ratio between the largest value and small value of $c$ for which the sensor achieves at least $\text{SNR}_{\text{min}}$. When we set $\text{SNR}_{\text{min}} = 20$ dB, the DR of the “BIN” schemes with $q = 1, 3,$ and $5$ are shown in Table 4.1. We can see that when we increase the threshold $q$, although the range becomes larger, the ratio actually decreases, which means that the dynamic range becomes smaller.

<table>
<thead>
<tr>
<th>$q$</th>
<th>$c_{\min}$</th>
<th>$c_{\max}$</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$10^2$</td>
<td>$10^{1.7}$</td>
<td>$10^{1.7} : 1$</td>
</tr>
<tr>
<td>3</td>
<td>$10^{3.26}$</td>
<td>$10^{4.93}$</td>
<td>$10^{1.67} : 1$</td>
</tr>
<tr>
<td>5</td>
<td>$10^{5.76}$</td>
<td>$10^{5.06}$</td>
<td>$10^{1.3} : 1$</td>
</tr>
</tbody>
</table>

4.3 Optimal Threshold Pattern Design

The previous section shows that a single threshold can not work for the whole range of light intensities. Usually in an image sensor, the oversampling factor is fixed. So we want to design an optimal threshold pattern, which works for a reasonable range
of light intensities. The goal of the optimal pattern is to minimize the MSE. But as in previous section, we know that the CRLB is a good approximation of the MSE. So we will design the optimal threshold pattern based on the CRLB. In this section, we consider the optimal threshold pattern containing only two thresholds $q_1$ and $q_2$. But
Thresholds Values and Patterns

Figure 4.4: The relation between the total light exposure value $c$ and the SNR of “BIN” schemes with different thresholds $q$ and the “IDEAL” scheme. The dashed line (in black) is the “BIN” scheme with threshold $q = 1$. The dotted line (in blue) represents the “BIN” scheme with threshold $q = 3$. The dash-dot line corresponds to the “BIN” scheme with threshold $q = 5$. The solid line (in green) shows the “IDEAL” scheme. The oversampling factor $K = 8096$.

The threshold pattern can have multiple thresholds $\{q_1, q_2, \ldots, q_K\}$.

We consider again the problem that we want to estimate a piecewise-constant light intensity field. As previously, we focus on estimating the parameter $c$ from the block of binary measurement $\mathbf{b} = [b_0, \ldots, b_{K-1}]^T$. The difference is that we now have two thresholds $q_1$ and $q_2$ instead of a single threshold $q$ for all the pixels. Let $K_{q_1}$ be the number of pixels with threshold $q_1$, $K^{(i)}_{q_1}$ be the number of “$i$’s, $i \in \{0, 1\}$, in the binary sequence $\mathbf{b}$ with threshold $q_1$, $K_{q_2}$ be the number of pixels with threshold $q_2$, $K^{(i)}_{q_2}$ be the number of “$i$’s, $i \in \{0, 1\}$, in the binary sequence $\mathbf{b}$ with threshold $q_2$. Thus,

$$K_{q_1} + K_{q_2} = K, \quad K^{(1)}_{q_1} + K^{(0)}_{q_1} = K_{q_1}, \quad \text{and} \quad K^{(1)}_{q_2} + K^{(0)}_{q_2} = K_{q_2}.$$

Introducing four functions

$$p_{q_1}^{(0)}(s) \overset{\text{def}}{=} \sum_{k=0}^{q_1-1} \frac{s^k}{k!} e^{-s}, \quad p_{q_1}^{(1)}(s) \overset{\text{def}}{=} 1 - \sum_{k=0}^{q_1-1} \frac{s^k}{k!} e^{-s}, \quad (4.3)$$

and

$$p_{q_2}^{(0)}(s) \overset{\text{def}}{=} \sum_{k=0}^{q_2-1} \frac{s^k}{k!} e^{-s}, \quad p_{q_2}^{(1)}(s) \overset{\text{def}}{=} 1 - \sum_{k=0}^{q_2-1} \frac{s^k}{k!} e^{-s}. \quad (4.4)$$

Denote by $\mathcal{L}_b(q_1, q_2, c)$ the likelihood function of observing $K$ binary sensor measurement $\mathbf{b}$ with thresholds $q_1$ and $q_2$. Similarly to the single threshold case, due to the independence of the photon counting processes at different pixel locations, we have

$$\mathcal{L}_b(q_1, q_2, c) = \left(p_{q_1}^{(1)}(s)\right)^{K^{(1)}_{q_1}} \left(p_{q_1}^{(0)}(s)\right)^{K^{(0)}_{q_1}} \left(p_{q_2}^{(1)}(s)\right)^{K^{(1)}_{q_2}} \left(p_{q_2}^{(0)}(s)\right)^{K^{(0)}_{q_2}}. \quad (4.5)$$
4.3 Optimal Threshold Pattern Design

\[q_1 \quad q_2 \quad q_1 \quad q_2\]

\[q_2 \quad q_1 \quad q_2 \quad q_1\]

\[q_1 \quad q_2 \quad q_1 \quad q_2\]

\[q_2 \quad q_1 \quad q_2 \quad q_1\]

\[\text{Figure 4.5: An 2-D image sensor with interleaved thresholds } q_1 \text{ and } q_2.\]

**Proposition 8.** The CRLB of estimating the total light exposure value \(c\) from \(K_{q_1}\) binary sensor measurements with threshold \(q_1 \geq 1\) and \(K_{q_2}\) binary sensor measurements with threshold \(q_2 \geq 1\) is

\[
\text{CRLB}_{\text{bin2}}(K_{q_1}, K_{q_2}, q_1, q_2, c) = \frac{1}{\sum_{i=1,2} \left( \frac{K_{q_i}(p_{q_i}^{(0)} - p_{q_i}^{(0)} c/K - p_{q_i}^{(0)}e/K - p_{q_i}^{(0)} c/K)_{c_{\min}}^{c_{\max}}}{K_{q_i}} \right)}. \tag{4.6}
\]

where \(K = K_{q_1} + K_{q_2}\) is the total number of pixels and we omit the function argument \(c/K\) in \(p_{q_i}^{(0)}(c/K)\) and \(p_{q_i}^{(0)}(c/K)\) for notational simplicity.

**Proof.** See Appendix 4.A.1. \(\Box\)

Note that here we only consider two thresholds, but our method can be easily generalized to multiple thresholds.

In this section, we want to find the optimal threshold pattern \(\{q_i\}\), and the optimal number of pixels \(\{K_{q_i}\}\), \(i = 1, 2\), given the oversampling factor \(K\), the maximum threshold \(q_{\max}\), and a uniformly distributed total light exposure value \(c\) in \([c_{\min}, c_{\max}]\). The optimal criterion is to find the parameters which have the smallest average MSE. In the previous section, we have shown that we can use the CRLB as an approximation of the MSE. So we propose the following optimality criterion for finding the optimal threshold pattern based on the CRLB.

**Optimality Criterion.** The optimal thresholds \(\{q_{i,\text{opt}}\}\), and number of pixels \(\{K_{q_i,\text{opt}}\}\), \(i = 1, 2\) are obtained as,

\[
(q_{1,\text{opt}}, q_{2,\text{opt}}, K_{q_1,\text{opt}}, K_{q_2,\text{opt}}) = \arg \min_{1 \leq q_i \leq q_{\max}, 0 \leq K_i \leq K_{q_i,\text{opt}}, i = 1, 2} c_{\max} - c_{\min} \int_{c_{\min}}^{c_{\max}} \text{CRLB}_{\text{bin2}}(K_{q_1}, K_{q_2}, q_1, q_2, c) \, dc. \tag{4.7}
\]

For a practical image sensor design, the number of pixels with specified thresholds should be fixed. So we set \(K_{q_1} = K_{q_2} = K/2\) and we interleave the two thresholds as shown in Figure 4.5. Also the term \(\frac{1}{c_{\max} - c_{\min}}\) does not effect the optimization results, so the optimization problem becomes

\[
(q_{1,\text{opt}}, q_{2,\text{opt}}) = \arg \min_{1 \leq q_1, q_2 \leq q_{\max}} \int_{c_{\min}}^{c_{\max}} \text{CRLB}_{\text{bin2}}(K/2, K/2, q_1, q_2, c) \, dc \tag{4.8}
\]
Figure 4.6: Optimal threshold pattern design when the oversampling factor is $K = 1024$, the maximum threshold is $q_{\text{max}} = 20$, and the total light exposure value is $c \in [10, 10^4]$. (a) The average CRLB for different threshold patterns $\{q_1, q_2\}$. (b) The comparison of the CRLB of the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$, two single-threshold patterns $\{q_1 = q_2 = 1\}$ and $\{q_1 = q_2 = 10\}$, and the “IDEAL” scheme.

Example 4. In this example, the oversampling factor is $K = 1024$, the maximum threshold is $q_{\text{max}} = 20$, and the range of the total light exposure value $c$ is $[10, 10^4]$, by solving the optimization problem in (4.8), we find the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$. Figure 4.6(a) shows the average CRLB for different threshold patterns $\{q_1, q_2\}$. Then Figure 4.6(b) shows the comparison of the CRLB of the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$, two single-threshold patterns $\{q_1 = q_2 = 1\}$ and $\{q_1 = q_2 = 10\}$, and the “IDEAL” scheme. We can see that although the optimal threshold pattern is not the optimal for each $c$, it is better than other threshold patterns by considering the average performance.

Example 5. In this example, the oversampling factor is $K = 8096$, the maximum threshold is $q_{\text{max}} = 20$, and the range of the total light exposure value $c$ is $[10^2, 10^5]$, by solving the optimization problem in (4.8), we find the optimal threshold pattern $\{q_1 = 10, q_2 = 3\}$. Figure 4.7(a) shows the average CRLB for different threshold patterns $\{q_1, q_2\}$. Then Figure 4.7(b) shows the comparison of the CRLB of the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$, two single-threshold patterns $\{q_1 = q_2 = 1\}$ and $\{q_1 = q_2 = 10\}$, and “IDEAL” scheme.

4.4 Reconstruction under Different Thresholds

In the previous section, we proposed how to design an optimal threshold pattern. In this section, we will show that for an arbitrary threshold pattern and a general linear light
4.4 Reconstruction under Different Thresholds

Figure 4.7: Optimal threshold pattern design when the oversampling factor is \( K = 8096 \), the maximum threshold is \( q_{\text{max}} = 20 \), and the total light exposure value is \( c \in [10^2, 10^5] \). (a) The average CRLB for different threshold patterns \( \{q_1, q_2\} \). (b) The comparison of the CRLB of the optimal threshold pattern \( \{q_1 = 10, q_2 = 3\} \) two single-threshold patterns \( \{q_1 = q_2 = 1\} \) and \( \{q_1 = q_2 = 10\} \), and the “IDEAL” scheme.

Since we only change the threshold for each pixel, the linear field model used here is the same as in Chapter 2. The MLE for our image sensor under arbitrary threshold pattern is almost the same as that in Chapter 2, except the probability that each pixel is “1” or “0” is calculated using different thresholds. To make this chapter self-contained, we will describe the image reconstruction by the MLE again as showed in Chapter 2.

To reconstruct an image (i.e., the light field \( \lambda(x) \)) is equivalent to estimating the parameters \( \{c_n\} \) in (2.1) due to the linear field model introduced in Definition 1 in Chapter 2. As shown in (2.8), the light exposure values \( \{s_m\} \) at different sensors are related to \( \{c_n\} \) through a linear mapping, implemented as upsampling followed by filtering as in Figure 2.2. Since it is linear, the mapping (2.8) can be written as a matrix-vector multiplication

\[
s = Gc,
\]

where \( s = [s_0, s_1, \ldots, s_{M-1}]^T \), \( c = [c_0, c_1, \ldots, c_{N-1}]^T \), and \( G \) is an \( M \times N \) matrix representing the combination of upsampling (by \( K \)) and filtering (by \( g_m \)). Each element of \( s \) can then be written as

\[
s_m = e_m^T Gc,
\]

where \( e_m \) is the \( m \)th standard Euclidean basis vector.\(^1\)

As shown in (2.37), the likelihood function given \( M \) binary measurements \( b \) \( \overset{\text{def}}{=} \)

\(^1\)Here we use zero-based indexing. Thus, \( e_0 \overset{\text{def}}{=} [1, 0, \ldots, 0]^T \), \( e_1 \overset{\text{def}}{=} [0, 1, \ldots, 0]^T \), and so on.
[b_0, b_1, \ldots, b_{M-1}]^T$ can be computed as

$$
\mathcal{L}_b(c) = \prod_{m=0}^{M-1} \mathbb{P}(B_m = b_m; s_m)
= \prod_{m=0}^{M-1} p_m(e_m^T Gc),
$$

(4.11)

where (4.11) follows from (2.13) and (4.10). The log-likelihood function, as shown in (2.38), is defined as

$$
\ell_b(c) \overset{\text{def}}{=} \log \mathcal{L}_b(c) = \sum_{m=0}^{M-1} \log p_m(e_m^T Gc).
$$

(4.12)

For any given observation $b$, the MLE we seek is the parameter that maximizes $\mathcal{L}_b(c)$, or equivalently, $\ell_b(c)$. Specifically,

$$
\hat{c}_\text{ML}(b) \overset{\text{def}}{=} \arg \max_{c \in [0, S]^N} \ell_b(c) = \arg \max_{c \in [0, S]^N} \sum_{m=0}^{M-1} \log p_m(e_m^T Gc).
$$

(4.13)

The constraint $c \in [0, S]^N$ means that every parameter $c_n$ should satisfy $0 \leq c_n \leq S$, for some preset maximum value $S$.

The only difference in calculating the likelihood function defined (4.11) and the log-likelihood function defined in (4.12) for a binary image sensor with different thresholds and a single threshold is that in the different thresholds case we need to compute the function $p_m(e_m^T Gc)$ using the threshold $q(m)$ for the $m$th pixel instead of a single threshold $q$ for all the pixels.

**Corollary 1.** For arbitrary binary sensor measurements $b$ and arbitrary threshold patterns, i.e., an arbitrary threshold $q(m)$ for the $m$th pixel, the log-likelihood function $\ell_b(c)$ defined in (4.12) is concave in the domain $c \in [0, C]^N$.

**Proof.** The function $\log f_{b_m}(s)$ is either

$$
\log \sum_{k=0}^{q(m)-1} \frac{s^k e^{-s}}{k!} \quad \text{or} \quad \log \sum_{k=q(m)}^{\infty} \frac{s^k e^{-s}}{k!},
$$

(4.14)

where $q(m)$ is the threshold for $m$th pixel. We can apply Lemma 2 of Chapter 2 for both cases, and show that $\{\log f_{b_m}(s)\}$ are concave functions on $s \geq 0$. Since the sum of concave functions is still concave and the composition of a concave function with a linear mapping ($s_m = e_m^T Gc$) is still concave, we conclude that the log-likelihood function defined in (2.38) is concave.

In general, we can not get a closed-form solution to the maximization problem in (4.13). Corollary 1 guarantees that we can find the optimal solution using iterative numerical methods.
4.5 Numerical Results

In this section, we give several numerical results to verify our analysis about the influence of thresholds and show the effectiveness of our optimal threshold pattern design.

4.5.1 1-D Synthetic Signals

Consider a 1-D light intensity field $\lambda(x)$ shown in Figure 4.8(a). The interpolation filter $\varphi(x)$ we use is the cubic B-spline function $\beta_3(x)$ defined in (2.3). The light intensity field $\lambda(x)$ is a linear combination of the shifted kernel, with the expansion coefficients $\{c_n\}$ shown as blue dots in the figure.

We simulate a binary sensor with threshold pattern $\{q_1 = q_2 = 1\}$, oversampling factor $K = 1024$, and $Kq_1 = Kq_2 = 512$. Applying the proposed MLE-based algorithm in Section 2.4, we obtain a reconstructed light intensity field (the red dashed curve) shown in Figure 4.8(b), together with the original “ground truth” (the blue solid curve). Due to the small thresholds, we can not estimate the high intensity value precisely.

Then we change the threshold pattern to $\{q_1 = q_2 = 10\}$. Figure 4.8(c) shows the result. Due to the large threshold, the estimation for the high light intensities is improved, but the estimation for the low light intensity value becomes worse. Then we use the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$ and the result is shown in Figure 4.8(d). Using this optimal threshold pattern, we get much better estimation results for both low and high light intensities.

4.5.2 2-D Synthetic Images

Consider a 2-D image as shown in Figure 4.9(a). The value of each element in the image is in $[500, 8000]$. The interpolation filter $\varphi(x)$ is the cubic B-spline function $\beta_3(x)$ defined in (2.3). We simulate a binary sensor with a threshold pattern that is interleaving two thresholds $q_1$ and $q_2$ as in Figure 4.5. The spatial oversampling factor is set to $8 \times 8$, and the temporal oversampling factor is 16. The reconstruction result when the threshold pattern is $\{q_1 = q_2 = 1\}$ is shown in Figure 4.9(b). We can notice that there is noise in the region where the light intensity values are large, because of the small thresholds. Figure 4.9(c) shows the results when the threshold pattern is $\{q_1 = q_2 = 10\}$. In this case, noise is visible when the light intensity is low. Then we use the optimal threshold pattern designed in Example 4 of Section 4.3 (assuming $c$ in the range $[10, 10^4]$ and $K = 1024$) i.e., $\{q_1 = 8, q_2 = 3\}$. The result is shown in Figure 4.9(d). The MSE for the reconstruction results of the three threshold patterns $\{q_1 = q_2 = 1\}$, $\{q_1 = q_2 = 10\}$, and $\{q_1 = 8, q_2 = 3\}$ are $3.293 \times 10^6$, $1.647 \times 10^6$, and $4.601 \times 10^4$, respectively. The optimal threshold pattern works better. This obeys our theoretical analysis of Section 4.3.

4.5.3 Results on Real Sensor Data

We also did some experiments with the SPAD camera [20], shown in Figure 4.10. The pixel value of the SPAD camera is “1” and “0”. If there is at least one photon hitting on the pixel, its value will become “1”, otherwise “0”. The resolution of the SPAD camera is $32 \times 32$. Due to the small resolution, we can not obtain an image with many gray levels by only using the spatial binary samples. Multiple exposures are used to compensate for the low spatial resolution. One picture is taken during each exposure. The exposure time for each picture is $4\mu s$. For the threshold $q = 1$, we can directly use the binary
Figure 4.8: Binary sensing with different threshold patterns and reconstructions of 1-D light fields. The oversampling factor $K$ is equal to 1024. The number of pixels with threshold $q_1$ is 512 and the rest are with threshold $q_2$. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernel. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with the threshold pattern $\{q_1 = q_2 = 1\}$. (c) Results with the threshold pattern $\{q_1 = q_2 = 10\}$. (d) Results using the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$ as obtained in the previous section.

images taken by the SPAD camera to estimate the light intensity using (2.28). For $q > 1$, we need to make the following assumption: during each exposure time, the probability that there are photons hitting on the pixel is very small. Through adding $J$ binary images, the SPAD camera becomes a photon-counting device. A binary image can then be obtained by thresholding the outcome of this photon-counting device. We choose $J = 45$ for $q = 30$, and $J = 95$ for $q = 60$. For each threshold, 256 binary images are taken. One of the binary images taken by the SPAD camera, and the estimated light intensity for $q = 1$, $q = 30$ and $q = 60$ are shown in Figure 4.11(a),(b),(c), and (d) respectively. We can see that the estimated light intensities for $q = 30$ and $q = 60$ are noisier than that for $q = 1$. This is because a small threshold is more sensitive for the low light intensities.

4.6 Conclusions

We studied the influence of the threshold in the proposed binary image sensor. We showed that the CRLB and MSE of small thresholds for smaller light exposure values were better than those of large thresholds while large thresholds worked better for larger light exposure values. Large thresholds decreased the dynamic range of the sensor. We observed that the CRLB was a good approximation for the MSE of our MLE estimator.

We designed an optimal threshold pattern which worked for a large range of light
4.6 Conclusions

Figure 4.9: Binary sensing with different threshold patterns and reconstructions of 2-D light fields. The spatial oversampling factor is set to $8 \times 8$ and the temporal oversampling factor is 16. So the total oversampling factor is 1024. The 2-D threshold pattern is the same as in Figure 4.5. (a) The original image. (b) The reconstruction result obtained by the proposed MLE-based algorithm, using measurements taken by a sensor with the threshold pattern $\{q_1 = q_2 = 1\}$. (c) Results for the threshold pattern $\{q_1 = q_2 = 10\}$. (d) Results using the optimal threshold pattern $\{q_1 = 8, q_2 = 3\}$ as obtained in the previous section.

exposure values based on minimization the average MSE. Through showing that the log-likelihood function is concave under arbitrary thresholds pattern, we ensured that we can find the global optimal solution when estimating the light intensity field using the MLE. The experimental results verify our analysis on the influence of the thresholds and prove the effectiveness of our proposed optimal threshold pattern design algorithm.
Figure 4.10: The SPAD camera

Figure 4.11: A binary image and the estimated light intensity for thresholds: (b) $q = 1$; (c) $q = 30$; and (d) $q = 60$. 
4.A Appendix

4.A.1 The CRLB of Binary Sensors with Two Thresholds

We first compute the Fisher information, defined as $I(c) = \mathbb{E}[-\frac{\partial^2}{\partial c^2} \log \mathcal{L}_b(c)]$. Using (4.5), we get

\[
I(c) = \mathbb{E} \left[ -\frac{\partial^2}{\partial c^2} \left( K q_1 \log p_{q_1}^1 + K q_1 \log p_{q_1}^0 + K q_2 \log p_{q_2}^0 \right) \right] \\
= \sum_{i=1,2} \mathbb{E} \left[ \frac{K q_i (p_{q_i}^0 q_{q_i}^1 + (p_{q_i}^0 q_{q_i}^0)^2)}{K^2 (p_{q_i}^0)^2} - \frac{K q_i (p_{q_i}^0 q_{q_i}^0 - (p_{q_i}^0 q_{q_i}^0)^2)}{K^2 (p_{q_i}^0)^2} \right] 
\]

(4.15)

where we omit the function argument $c/K$ in $p_{q_i}^0(c/K)$, $p_{q_i}^1(c/K)$, $p_{q_i}^0 c/K)$, and $p_{q_i}^0 c/K)$, $i = 1, 2$, for notational simplicity, $p_{q_i}^i(x) = \frac{\partial}{\partial x} q_{q_i}^i(x)$, and $p_{q_i}^ii(x) = \frac{\partial^2}{\partial x^2} q_{q_i}^i(x)$ are the first and second order derivative of $q_{q_i}^i(x)$, $i = 1, 2$, respectively. In reaching (4.15), we have also used the facts that $q_{q_i}^i(x) = 1 - q_{q_i}^i(x)$, and thus $q_{q_i}^i(x) = -\frac{\partial}{\partial x} q_{q_i}^i(x)$ and $p_{q_i}^ii(x) = -\frac{\partial^2}{\partial x^2} q_{q_i}^i(x)$, $i = 1, 2$.

Note that $K q_i = \sum_{0 \leq m < K q_i} b_{m,q_i}$, $i = 1, 2$, where $b_{m,q_i}$ is the pixels with threshold $q_i$, is a binomial random variable, and thus its mean can be computed as

\[
\mathbb{E}[K q_i] = K q_i p_{q_i}^1 = K q_i (1 - p_{q_i}^0).
\]

(4.16)

On substituting the above expression into (4.15), the Fisher information can be simplified as

\[
I(c) = \sum_{i=1,2} K q_i \left( \frac{p_{q_i}^0 q_{q_i}^0 + (p_{q_i}^0 q_{q_i}^0)^2}{K^2 q_{q_i}^i} - \frac{p_{q_i}^0 q_{q_i}^0 - (p_{q_i}^0 q_{q_i}^0)^2}{K^2 q_{q_i}^i} \right) \\
= \sum_{i=1,2} \frac{K q_i (p_{q_i}^0 q_{q_i}^0)^2}{K^2 q_{q_i}^i (p_{q_i}^0 q_{q_i}^i)} 
\]

(4.17)

Using the definition of $p_{q_i}^0$ in (4.3) and (4.4), the derivative in the numerator of (4.17) can be computed as

\[
p_{q_i}^0 q_i(x) = -e^{-x} \frac{x^{q_i-1}}{(q_i-1)!} = p_{q_i-1}^0(x) - p_{q_i}^0(x).
\]

(4.18)

Finally, since $\text{CRLB}_{\text{bin2}}(K q_1, K q_2, q_1, q_2, c) = 1/I(c)$, we reach (4.6) by substituting (4.3), (4.4), and (4.18) into (4.17), and after some straightforward manipulations.
Chapter 5

Generalized Piecewise-constant Model

5.1 Introduction

As we have shown in Chapter 2, when assuming that the light intensity field $\lambda(x)$ is piecewise-constant, we can estimate each of the expansion coefficients $\{c_n\}$ independently. But if we have more prior information of $\{c_n\}$, according to the theoretical analysis in Section 2.3 of Chapter 2, we can improve the estimation performance. Moreover, the assumptions about the piecewise-constant model limit its application. Namely, these are: the oversampling factor $K$ is known; the starting point of each piece is a multiple of the oversampling factor $K$ in the light exposure value sequence $s$; the length of each piece is a multiple of $K$.

To solve these problems and improve the estimation performance, in this chapter, we directly consider the problem of estimating the light exposure values $\{s_n\}$ for each pixel. We model $\{s_n\}$ as piecewise-constant. The starting point and the length of each piece are arbitrary. The oversampling factor $K$ is also unknown. We propose to use the maximum likelihood estimator (MLE) for reconstructing the set of segments and their corresponding light exposure values. To find the solution, we iteratively solve two subproblems: estimating the light exposure values given the optimal segmentation; finding the optimal segmentation given the estimated light exposure values.

We use dynamic programming, a greedy algorithm, and methods based on pruning of binary trees or quadtrees to find the optimal segmentation. Dynamic programming can find the optimal solution for 1-D signals, but with high complexity, $O(M^3)$ for 1-D signals with $M$ pixels and $O(M^9)$ for 2-D images with a resolution $M \times M$. The greedy algorithm and the pruning algorithms can achieve a suboptimal solution, but with lower complexity. The complexity of the greedy algorithm is $O(M^2)$ for 1-D signals with $M$ pixels and is $O(M^3)$ for 2-D images with a resolution $M \times M$. The pruning algorithms have complexity $O(M)$ for 1-D signals with $M$ pixels and $O(M^2)$ for 2-D images with a resolution $M \times M$.

This chapter is organized as follows. Section 5.2 focuses on the estimating light exposure values $\{s_n\}$ from binary observations when the light exposure values is modeled as generalized 1-D piecewise constant. Section 5.3 deals with the 2-D case. Numerical results for synthesized data and real images are in Section 5.4, and the conclusion is
5.2 Estimating 1-D Piecewise-constant Light Exposure Values

In Chapter 2, we modeled the light intensity field $\lambda(x)$ using the $N$ expansion coefficients $\{c_n\}$ as in (2.1), i.e., we assume that the light exposure values $\{s_n\}$ are not independent. By setting the interpolation kernel $\varphi(x)$ in (2.1) as the box function $\beta(x)$, we have a piecewise-constant light intensity field and the light exposure values $s$ are also piecewise-constant. In that case, all the starting indexes of the constant pieces of $s$ are a multiple of the oversampling factor $K$. Each piece has also a length which is a multiple of $K$ pixels. The first limitation of this model is that sometimes we do not know exactly the oversampling factor $K$. The second is that we cannot have an arbitrary starting index and an arbitrary length for each segment. These limitations reduce the potential signal-to-noise ratio (SNR) improvement given by the result of the MLE. For example, if $c_0 = c_1$, all the $2K$ measurements can be used to estimate $c_0$ and $c_1$, instead of estimating separately, obtaining better SNR performance.

To overcome these problems, in this chapter, we directly estimate the light exposure values $\{s_n\}$ instead of the expansion coefficients $\{c_n\}$. We propose a generalized piecewise-constant light exposure value model which allows an arbitrary starting index and an arbitrary length for each segment. Let the total number of pixels in the binary sensor be $M$, and the indexes of the pixels be from 1 to $M$. Let $S_i$ be the $i$th segment with $i = 1, 2, \ldots, p$, where $p$ is the number of segments, $1 \leq p \leq M$. $S_i$ can be denoted as $[I_{s,i}, I_{e,i}]$, where $I_{s,i}$ is the starting index and $I_{e,i}$ is the ending index of the $i$th segment. Note that the union of all the segments is $[1, M]$, and there is no overlap between any two segments. In each segment, the light exposure value is the same for each pixel and denoted as $s_i$. Denote by the segment set $G \overset{\text{def}}{=} \{S_1, S_2, \ldots, S_i, \ldots, S_p\}$. We define $P$ as the set containing all the segments and their corresponding light exposure values $s_i$, i.e.,

$$P \overset{\text{def}}{=} \{(S_1, s_1), (S_2, s_2), \ldots, (S_i, s_i), \ldots, (S_p, s_p)\}; \quad (5.1)$$

Then our goal is to estimate $P$ from the $M$ binary measurements $b \overset{\text{def}}{=} [b_0, b_1, \ldots, b_{M-1}]^T$. The unknown parameters are the number of segments $p$, the light exposure value for each segment $s_i$, and the segments $S_i$, $i = 1, 2, \ldots, p$.

We assume that the number of segments $p$ can be modeled as a realization of the random variable $P$, which follows geometric distribution with parameter $\theta$, i.e.,

$$P(P = p) = \theta^{p-1}(1 - \theta), \quad \text{for } p \in \mathbb{Z}^+, 0 < \theta < 1 \quad (5.2)$$

According to this distribution, a small number of segments is more likely than a large number.

Denote by $L_b(P)$ the likelihood function of observing $M$ binary sensor measurements...
5.2 Estimating 1-D Piecewise-constant Light Exposure Values

b. Then,

\[ L_b(\mathcal{P}) = \mathbb{P}(B_m = b_m, 0 \leq m \leq M, P = \mathcal{P}) \]
\[ = \mathbb{P}(B_m = b_m, 0 \leq m \leq M | P = \mathcal{P}) \mathbb{P}(P = p) \]
\[ = \prod_{i=0}^{p} \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, s_i)) \theta^{p-1}(1 - \theta), \tag{5.3} \]

where (5.3) is due to the independence of the photon counting processes in each segment and (5.2).

Let \( K_i \) be the number of pixels in the \( i \)th segment, and \( K_i^{(1)} \) be the number pixels with values equal to “1” in this segment. Then

\[ \mathbb{P}(B_m = b_m, m \in \mathcal{S}_i; (\mathcal{S}_i, s_i)) = \prod_{m \in \mathcal{S}_i} \mathbb{P}(B_m = b_m; (\mathcal{S}_i, s_i)), \tag{5.4} \]
\[ = (p_1(s_i))^{K_i^{(1)}} (p_0(s_i))^{K_i - K_i^{(1)}} \tag{5.5} \]

where (5.4) is due to the independence of the photon counting processes at different pixel locations, and (5.5) follows from (2.13). Plug (5.5) into (5.3), we have

\[ L_b(\mathcal{P}) = \prod_{i=0}^{p} (p_1(s_i))^{K_i^{(1)}} (p_0(s_i))^{K_i - K_i^{(1)}} \theta^{p-1}(1 - \theta). \tag{5.6} \]

The log-likelihood function is defined as

\[ \ell_b(\mathcal{P}) \overset{\text{def}}{=} \log L_b(\mathcal{P}) \]
\[ = \sum_{i=1}^{p} \left( K_i^{(1)} \log p_1(s_i) + (K_i - K_i^{(1)}) \log p_0(s_i) \right) + (p - 1) \log \theta + \log(1 - \theta). \tag{5.7} \]

Then the MLE we seek is the parameters \( \mathcal{P} \) that maximizes \( L_b(\mathcal{P}) \), or \( \ell_b(\mathcal{P}) \). Specifically,

\[ \hat{\mathcal{P}} \overset{\text{def}}{=} \arg \max_{\mathcal{P}, s_i \in [0, S]} \ell_b(\mathcal{P}) \]
\[ = \arg \max_{\mathcal{P}, s_i \in [0, S]} \sum_{i=1}^{p} \left( K_i^{(1)} \log p_1(s_i) + (K_i - K_i^{(1)}) \log p_0(s_i) \right) - \gamma p, \tag{5.8} \]

where \( \gamma = - \log \theta \).

In this optimization problem, we need to jointly estimate each of the segment \( \mathcal{S}_i \) and its corresponding light exposure value \( s_i \), \( i = 1, 2, \cdots, p \). The above optimization problem can be solved by iteratively working out two optimization problems.

The first optimization problem is to estimate the light exposure value \( s_i \) for each segment in \( \mathcal{P} \), given the estimated segment set \( \hat{\mathcal{G}} \). Therefore (5.8) becomes

\[ \hat{s}_i \overset{\text{def}}{=} \arg \max_{s_i \in [0, S]} \left( K_i^{(1)} \log p_1(s_i) + (K_i - K_i^{(1)}) \log p_0(s_i) \right). \tag{5.9} \]

Similarly to Lemma 1 in Chapter 2, we have a closed-form formula for the above MLE.
Lemma 7. The MLE solution to (5.9) is

\[ \hat{s}_i = \begin{cases} p_0^{-1}(1 - K_i^{(1)} / K_i), & \text{if } 0 \leq K_i^{(1)} \leq K_i (1 - p_0(S)) / S, \\ S, & \text{otherwise} \end{cases} \]  

where \( p_0^{-1}(x) \) is the inverse function of \( p_0(x) \).

If we ignore the saturation effect, i.e., the constraint of the maximum value \( S \), we have

\[ p_0(\hat{s}_i) = 1 - K_i^{(1)} / K_i \text{ and } p_1(\hat{s}_i) = K_i^{(1)} / K_i. \]  

The second optimization problem is to find the optimal segment set \( \hat{G} \) in \( P \), given the estimated light exposure value \( \hat{s}_i \). In this case, (5.8) becomes

\[
\hat{G} = \arg\max_{\mathcal{G}} \sum_{i=1}^{p} \left( K_i^{(1)} \log p_1(\hat{s}_i) + (K_i - K_i^{(1)}) \log p_0(\hat{s}_i) \right) - \gamma p \\
= \arg\max_{\mathcal{G}} \sum_{i=1}^{p} \left( K_i^{(1)} \log (K_i^{(1)} / K_i) + (K_i - K_i^{(1)}) \log (1 - K_i^{(1)} / K_i) \right) - \gamma p,
\]

where (5.12) is obtained by plugging in the values of \( p_0(\hat{s}_i) \) and \( p_1(\hat{s}_i) \) as in (5.11).

Dynamic programming, a greedy algorithm, and a method based on pruning of binary trees are proposed to solve (5.12) and find the optimal segmentation.

5.2.1 Dynamic Programming

To find a recursive expression of the cost function, we consider the segments in the region \([1, k]\). Let \( C(k, l), 1 \leq k \leq l \leq M \) be the cost when there is only one segment in the region \([k, l]\) and \( F(t, k), 1 \leq t \leq k \leq M \) be the maximum total cost when there are \( t \) segments in \([1, k]\). We define \( C(k, l) \) as

\[ C(k, l) \overset{\text{def}}{=} K_{kl}^{(1)} \log (K_{kl}^{(1)} / K_{kl}) + (K_{kl} - K_{kl}^{(1)}) \log (1 - K_{kl}^{(1)} / K_{kl}) - \gamma, \]

where \( K_{kl} \) is the number of pixels in the segment \([k, l]\), \( K_{kl}^{(1)} \) is the number of “1”s, and \( b_m \) is the pixel value of the \( m \)th pixel. Then, \( F(t, k), 1 \leq t \leq k \leq M \) is computed using the following iteration equations.

\[
\begin{cases}
F(1, k) = C(1, k), & 1 \leq k \leq M \\
F(t, k) = \max_{2 \leq j \leq k} \{ F(t - 1, j - 1) + C(j, k) \}, & 2 \leq t \leq k, 1 \leq k \leq M
\end{cases}
\]

The optimal segmentation

\[ \hat{G} = \arg\max_{\mathcal{G}} F(p, M), \quad 1 \leq p \leq M. \]

The complexity of this algorithm is \( \mathcal{O}(M^3) \), where \( M \) is the number of pixels.
5.2 Estimating 1-D Piecewise-constant Light Exposure Values

Algorithm 3 The 1-D greedy algorithm.

**Initialize:** \( U := \{ [1, N] \} \), \( U_n := \text{null} \). \( U \) is the set which contains the segments to be divided, \( U_n \) is the segments which can not be divided.

**Loop:** while \( \# U \neq 0 \), \( \# U \) is the number of elements in \( U \)

for \( i = 1 \) to \( \# U \)

suppose the \( i \)th element of \( U \) is \([k, l]\),

\[ t = \arg \max_{k+1 \leq t \leq l} C(k, t-1) + C(t, l) \]

if \( C(k, t-1) + C(t, l) > C(k, l) \)

put \([k, t-1]\) and \([t, l]\) into a set \( U_t \)

delete \([k, l]\) from the set \( U \)

else {put \([k, l]\) into the set \( U_n \)}

\( U = U_t \)

5.2.2 Greedy Algorithm

Since the dynamic programming’s complexity is \( O(M^3) \), a simple greedy algorithm can be employed to increase the speed. We define the same cost function \( C(k, l), 1 \leq k \leq l \leq M \) for a segment \([k, l]\) as in dynamic programming. In the greedy algorithm, the total cost is first set to be \( C(1, M) \). Then, we decide whether to divide the segment \([1, M]\) into two segments \([1, t-1]\) and \([t, M]\), where

\[ t = \arg \max_{2 \leq t \leq M} \{ C(1, t-1) + C(t, M) \} \]

If \( C(1, M) < C(1, t-1) + C(t, M) \), we make the division, otherwise not. If the segment is cut into two segments, then we consider if the two segments can still be cut to gain in the total cost. This is done iteratively until no segment can be cut. The procedure gives a sub-optimal solution. The complexity for this algorithm is \( O(M^2) \), which is faster than dynamic programming. The pseudocode for this algorithm is shown in Algorithm 3.

5.2.3 Pruning of Binary Trees

To further reduce complexity, a binary tree is first constructed. The tree is denoted as \( T_{k, l}, 0 \leq k \leq \log_2 M, l = 1, \cdots, 2^k \). Each node \( T_{k, l} \) means that the light exposure value \( s_j, j \in [(k-1)2^{\log_2 M-1} + 1, k \times 2^{\log_2 M-1}] \) is constant. We remark that this choice restricts the number of allowed partitions with respect to the previous models. The cost function \( C(m, n), 0 \leq n \leq m \leq M \) is the same as in the greedy algorithm. Each node has a cost value \( C_n(T_{k, l}) = C \left( (k-1)2^{\log_2 M-1} + 1, k \times 2^{\log_2 M-1} \right) \). We prune this tree from the leaf nodes. If the cost value of the parent node is larger than the summation of two child nodes, then we prune two child nodes. If the parent node has a child node that has unpruned child nodes, then the parent will not be considered for pruning. The pruning process is implemented iteratively until no node can be pruned. All the leaf nodes in the pruned tree denote the segmentation of the \( M \) pixels. The time complexity of this algorithm is \( O(M) \). The pseudocode for this algorithm is shown in Algorithm 4.
5.3 Estimating 2-D Piecewise-constant Light Exposure Values

In this section, we consider the 2-D generalized piecewise-constant model for the light exposure values. Differently from the previous section, we want to reconstruct the 2-D light exposure values $s$ from the 2-D binary measurements $b$, both of which are $M \times M$ matrix, $M^2$ pixels.

Denote by $\mathcal{B}$ as a set containing all the indexes of the pixels, i.e., $[1, M] \times [1, M]$. In our model, $\mathcal{B}$ contains several blocks $\mathcal{B}_i$, $i = 1, 2, \cdots, p$, where $p$ is the number of blocks, $1 \leq p \leq M^2$, and $\mathcal{B}_i$ is a set containing pixels’ indexes of the $i$th block. $\mathcal{B}_i$ can be denoted as $[I_{lv,i}, I_{rbv,i}] \times [I_{lth,i}, I_{rbh,i}]$, where $I_{lv,i}, I_{rbv,i}, I_{lth,i},$ and $I_{rbh,i}$ are the left upper vertical index, the right bottom vertical index, the left upper horizontal index, and the right bottom horizontal index of the $i$th block, respectively. Note that the union of all the blocks is $\mathcal{B}$, and there is no overlap between any two blocks. In each block, the light exposure value is the same for all the pixels and denoted as $s_i$. In the 2-D case, $\mathcal{P}$ is defined as a set of 2-D blocks and their corresponding light exposure value $s_i$, i.e.,

$$\mathcal{P} \equiv \{(\mathcal{B}_1, s_1), (\mathcal{B}_2, s_2), \ldots, (\mathcal{B}_i, s_i), \ldots, (\mathcal{B}_p, s_p)\}.$$ 

The block set $\mathcal{G}$ is defined as

$$\mathcal{G} \equiv \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_i, \ldots, \mathcal{B}_p\}.$$ 

We can use the MLE to estimate $\mathcal{P}$. The MLE has the same form as the 1-D case as shown in (5.8), and we only need to use the corresponding $K_i$, $K_i^{(1)}$, namely, the number of pixels and the number of “1”s in the block $\mathcal{B}_i$. Similarly to the 1-D case, we iterate two optimization problems (5.9) and (5.12) to find the optimal solution $\hat{\mathcal{P}}$.

Dynamic programming, a greedy algorithm and a method based on pruning of quadtrees are proposed to solve the 2-D case of (5.12) and find the optimal blocks.

5.3.1 Dynamic Programming

Let $C(k, l, m, n), 1 \leq k \leq m \leq M, 1 \leq l \leq n \leq M$ be the cost when there is only one block in the region $[k, m] \times [l, n]$ and $F(i, k, l, m, n), 1 \leq k \leq m \leq M, 1 \leq l \leq n \leq M$ is the maximum total cost when there are $i$ blocks in the $[k, m] \times [l, n]$. We denote by $b_{uv}$ the binary pixel value at position $(u, v)$, $k \leq u \leq m, l \leq v \leq n$, $K$ the number of pixels, and $K^{(1)}$ is the number of pixels with values “1” in the region $[k, m] \times [l, n]$. Then the cost $C(k, l, m, n)$ is defined as

$$K^{1} \log(K^{1}/K) + (1 - K^{(1)}/K) \log(1 - K^{(1)}/K) - \gamma.$$
5.3 Estimating 2-D Piecewise-constant Light Exposure Values

The maximum total cost \( F(i, k, l, m, n), 1 \leq k \leq m \leq M, 1 \leq l \leq n \leq M \) is computed using the following iteration equations.

\[
\begin{align*}
F(i, k, l, m, n) &= C(k, l, m, n), \; 1 \leq k \leq m \leq M, 1 \leq l \leq n \leq M \\
F(i, k, l, m, n) &= \max \{ \mathcal{H}, \mathcal{V} \}, \; 2 \leq i \leq K, \; 1 \leq k \leq m \leq M, 1 \leq l \leq n \leq M \\
\mathcal{H} &= \max_{\{k+1 \leq a \leq m\}, \{1 \leq u \leq i-1\}} \{ F(a, k, l, u-1, n) + F(i-a, a, u, l, m, n) \}, \\
\mathcal{V} &= \max_{\{i+1 \leq v \leq n\}, \{1 \leq u \leq i-1\}} \{ F(a, k, l, m, v-1) + F(i-a, a, k, v, m, n) \},
\end{align*}
\]

The optimal block set is \( \hat{G} = \arg \max_{\hat{G}} F(p, 1, 1, M, M), \; 1 \leq p \leq M^2 \).

The time complexity of this algorithm is \( O(M^9) \).

5.3.2 Greedy Algorithm

The same cost function \( C(k, l, m, n) \) as in the dynamic programming is used here. In the greedy algorithm, the total cost is first set to \( C(1, 1, M, M) \). After that, we decide whether to divide the block \([1, M] \times [1, M]\) into two blocks \([1, M] \times [1, u-1]\) and \([1, M] \times [u, M]\) or \([1, v-1] \times [1, M]\) and \([v, M] \times [1, M]\), through solving the following optimization problem.

\[
\begin{align*}
u &= \arg \max_{\{2 \leq u \leq M\}} \{ C(1, 1, M, u-1) + C(1, u, M, M) \}, \\
v &= \arg \max_{\{2 \leq v \leq M\}} \{ C(1, 1, v-1, M) + C(v, 1, M, M) \}.
\end{align*}
\]

If \( C(1, 1, M, M) \) is smaller than \( \max\{ C(1, 1, M, u-1) + C(1, u, M, M), C(1, 1, v-1, M) + C(v, 1, M, M) \} \), we make the division, otherwise not. If the block is cut into two blocks, then we consider whether the two blocks can still be cut to gain in the total cost. This is done iteratively, until a sub-optimal solution is reached. The complexity of this algorithm is \( O(M^3) \).

5.3.3 Pruning of Quadtrees

A quadtree is constructed first. The tree is denoted as \( T_{i,m,n}, 0 \leq i \leq \log_4 M, m = 1, \cdots, 4^i, n = 1, \cdots, 4^i \). Each node \( T_{i,m,n} \) means that the light intensity \( s_{uv} \in [(m-1)a + 1, ma] \times [(n-1)a + 1, na] \), where \( a = 4^{(\log_4 M - i)} \), is constant. The cost function \( C(k, l, m, n) \) is the same as in the greedy algorithm. Each node has a cost value defined as

\[
C_n(T_{i,m,n}) = C((m-1)a + 1, (n-1)a + 1, ma, na),
\]

where \( a = 4^{(\log_4 M - i)} \).

We prune this tree from the leaf nodes. If the cost value of the parent node is larger than the summation of four child nodes, we prune all the four child nodes. If the parent node has a child node that has at least an unpruned child node, the parent node will not be considered for pruning. The pruning process is implemented iteratively until no node can be pruned. All the leaf nodes in the pruned tree denote the segmentation of the \( M \times M \) pixels. The complexity of this algorithm is \( O(M^2) \).
Figure 5.1: Reconstructions of 1-D light exposure values. The threshold is $q = 1$, the number of pixels is $M = 1024$ and the penalized parameter is $\gamma = 5$. The red solid line is the original light exposure values and the blue dash line is the reconstructed result. The reconstruction algorithms are: (a) dynamic programming; (b) greedy algorithm; (c) pruning algorithm.

5.4 Numerical Results

5.4.1 1-D Synthetic Signals

A 1-D piecewise-constant signal is generated. A binary sensor with threshold $q = 1$ is simulated to take images of this signal and the number of pixels is $M = 1024$. We choose the parameter $\gamma$ to be 5 based on experiments. The results are shown in Figure 5.1(a), Figure 5.1(b), and Figure 5.1(c) respectively. The mean squared error (MSE) for dynamic programming, the greedy algorithm, and the pruning algorithm are 0.21, 1.0504, and 0.1583, respectively. The MSE of the dynamic programming is larger than the pruning algorithm. The reason is that there is no one-to-one correspondence between the maximum value of the cost function and a minimum value of the MSE. Even if dynamic programming gets the optimal solution for the objective function, the MSE may be bigger than that of the pruning of binary trees algorithm. The performance of the two algorithms also depends on $\gamma$. We need to choose this parameter based on a prior assumption on the number of segments. If the signal has many segments, $\gamma$ is small, otherwise large.
5.4 Numerical Results

Figure 5.2: Reconstructions of 2-D light exposure values with resolution 512 × 512. The threshold is $q = 1$. (a) The original 2-D light exposure values. (b) The binary image. (c) Reconstruction using the greedy algorithm. (c) Reconstruction using the pruning of quadtrees algorithm.

5.4.2 2-D Synthetic Images

We synthesize a 2-D piecewise-constant light exposure values as shown in Figure 5.2(a). Again, we use a binary sensor with threshold $q = 1$. Figure 5.2(b) shows the captured binary image. The reconstruction results using the greedy algorithm and the pruning of quadtrees algorithm with $\gamma = 5$ are shown in Figure 5.2(c) and Figure 5.2(d). The MSEs for these two algorithms are $4.294 \times 10^{-4}$ and 0.156, respectively.

5.4.3 Results on Real Sensor Data

We did some experiments using the SPAD camera presented in [20]. The device is shown in Figure 5.3(a). The resolution of the SPAD camera is $32 \times 32$. The pixel value of the image taken by the SPAD camera is “1” (at least one photon hitting the pixel) or “0” (no photon hitting the pixel).

To get a larger resolution binary image with the SPAD camera, it is fixed on the 2-D positioning system shown in Figure 5.3(a), which can move the SPAD camera on a 2-D plane. The camera is moved on a uniform grid of $32 \times 32$ positions and 1024 images are taken at each position. By stitching the $32 \times 32$ binary images, a binary image with resolution $1024 \times 1024$ is obtained. We use our proposed reconstruction algorithms to estimate the light exposure values from the binary image. Here we will show the light
intensity field instead of the light exposure values. We assume that the oversampling
factor is $8 \times 8$. Thus, the resolution of the reconstructed image is $128 \times 128$. The large
resolution image generated by stitching the binary images is shown in Figure 5.3(b).
The greedy algorithm and the pruning of quadtrees algorithm are used to reconstruct
the image. The results are shown in Figure 5.3(c) and Figure 5.3(d), respectively (For
visualization purpose, the values are rescaled and clipped).

5.5 Conclusion

In this chapter, we extend our piecewise-constant model of the light intensity field $\lambda(x)$. We
directly model the light exposure values $\{s_n\}$ for each pixel as piecewise-constant. This avoids the need to know the oversampling factor $K$ for the reconstruction. We propose to use an MLE for the estimation and find the optimal segmentation of the binary measurements and the light exposure values for each segment iteratively. Dynamic programming is used to solve the optimization problem. To increase the speed of reconstruction, two other methods, the greedy algorithm and methods based on pruning of binary trees or quadtrees are also proposed. Numerical results on both synthesized data and real images demonstrate the effectiveness of our methods.
5.5 Conclusion

Figure 5.3: Reconstructions of 2-D real sensor data. The resolution of the binary image is $1024 \times 1024$. The resolution of reconstructed images is $128 \times 128$. (a) The SPAD camera with resolution $32 \times 32$ is fixed on a 2-D positioning system. (b) The binary image generated by stitching 1024 binary images taken by the SPAD camera. (c) Reconstruction using the greedy algorithm. (d) Reconstruction using the pruning of quadtrees algorithm.
Generalized Piecewise-constant Model
Chapter 6

Conclusions

6.1 Thesis Summary

The trends in the design of image sensors are to build pixels with small pixel size, high sensitivity, and high dynamic range. Unfortunately, when we reduce the pixel size, for a single pixel, the dynamic range will become lower and the noise performance will be worse. But if we consider a group of small pixels, we can achieve better performance like that of film, where the pixel value is “0” or “1” and the light intensity is encoded as the local density of the binary values. Due to the lowpass effect of the lens, when the pixel becomes small, the image sensor is an oversampling device. In this thesis, we proposed a new image sensor that captures the light intensity field using oversampled one-bit pixels. How about the performance of this new binary image sensor compared with a conventional image sensor? Is this sensor robust to noise or not? How to generate the binary pixels and what are the influences of different parameters? How can we get a continuous-tone image from the binary image? In this thesis, we answered all these questions.

Due to the lowpass effect of the lens, the captured light intensity field has a finite number of degrees of freedom per unit space. Therefore, we can model it as a weighted summation of shifted and scaled versions of a nonnegative interpolation kernel. We formulated the oversampled binary sensing scheme as a parameter estimation problem based on quantized Poisson statistics. We showed that with a single-photon quantization threshold and large oversampling factors, the Cramér-Rao lower bound (CRLB) of the estimation variance approaches that of an ideal unquantized sensor, i.e., as if there were no quantization in the sensor measurements. We also proved that the maximum likelihood estimator (MLE) can asymptotically achieve the CRLB if the threshold is a single photon. Through showing that the log-likelihood function is concave, we guaranteed to find the optimal solution in the MLE. We showed that our sensor has the potential application in capturing high dynamic ranging scenes.

Then we studied the noise performance of our image sensor. We modeled the noise as an additive Bernoulli noise with a known parameter, i.e., noise rate. The noise can only convert a pixel from “0” to “1”. We showed that our image sensor is quite robust to noise. The noise would effect the estimation performance when the light intensity is low, and has limited influence on large light intensities. The dynamic range of the sensor decreases only slightly. In the noisy case, oversampling factors that are too large can
also deteriorate the performance. We also used an MLE to estimate the light intensity. When the threshold is single-photon, we have proved that the log-likelihood function is concave. Therefore, optimal solution can be achieved using iterative algorithms. Unfortunately this does not hold true for thresholds larger than “1”. For the piecewise-constant light intensity field, each of the expansion coefficients can be estimated independently. We proved that the likelihood function and the log-likelihood function are strictly pseudoconcave for arbitrary thresholds. Thus the optimal solution can be obtained using iterative algorithms. For a general light intensity field, the log-likelihood function is not even quasiconcave, and there is no guarantee that an iterative method converges to the optimal solution. But we can first assume the light intensity is piecewise-constant to find an approximated solution and then refine it with an iterative algorithm that uses the exact model.

We then studied the influence of one of the most important parameters of our image sensor, the threshold. As to be expected, we showed that lower thresholds work better for the low light intensity field, and larger thresholds work better for the strong light intensity field. To make an image sensor that works for a wide range of light intensities, we proposed to use the combination of several thresholds in the same sensor. We designed an optimal threshold pattern based on the criterion that minimized the average CRLB, which is a good approximation of the mean squared error (MSE) of our MLE. Through showing that the log-likelihood function is concave under arbitrary threshold patterns, we ensured to find the global optimal solution for estimating the light intensity field using the MLE.

Finally, we extended our light intensity model. We wanted to directly estimate the light exposure value, i.e., the average number of photons hitting on each pixel. We modeled the light exposure values as piecewise constant. We proposed to use an MLE for the reconstruction. This optimization problem is solved by iteratively working out two subproblems. The first one is to find the optimal light exposure value for each segment, given the optimal segmentation of the binary measurements. The second one is to find the optimal segmentation of the binary measurements given the optimal light exposure value for each segment. We used several algorithms to solve the optimization problem. Dynamic programming can achieve the optimal solution for 1-D signals, but it is quite complex. We proposed a greedy algorithm and a method based on pruning of binary trees or quadtrees to increase the reconstruction speed.

### 6.2 Future Research

So far, we have performed a lot of theoretical analysis on the performance of the over-sampled binary sensor. Actually, our collaborators Prof. Edoardo Charbon and his team have also designed a binary image sensor, as shown in Figure 6.1. This sensor is designed using standard CMOS process with 90nm technology node. The chip size is 2mm × 2mm. The resolution of the image is 1600 × 1200 with pixel size 0.75µm × 0.75µm. Our goal on the image sensor design is to have as small pixel size as possible so that the oversampling factor is large. One of future research goals is to build the binary image sensor with smaller CMOS technology nodes to achieve smaller pixel size.

In this thesis, we use Poisson statistics to model the input signals. When the light intensity is large and the threshold is also quite large, we can use Gaussian statistics to model the input signals. To implement large thresholds, we might be not able to first amplify the signal then implement the thresholding process like in the CMOS
6.2 Future Research

Figure 6.1: The chip micrograph of the binary image sensor designed by Prof. Edoardo Charbon and his team.

single-photon avalanche diode (SPAD) and electron-multiplying charge-coupled device (EMCCD). If we have to implement the thresholds without the amplification process, the noise model need to be changed to Gaussian noise model and we also need to consider the fluctuation of the thresholds.

In our current approach to image reconstruction, we only assume that the unknown image has finite degrees of freedom, i.e., bandlimited. However, natural images have in most of the cases a certain geometrical structure, and can be sparsely represented in certain basis (e.g. the wavelets). To further improve the performance of the MLE algorithm, we can incorporate the additional sparsity prior of natural images in the reconstruction step.

As we explained in Chapter 2 if the spatial oversampling factor of the binary sensor is not large enough, we could use the temporal oversampling to compensate for that. In the analysis, it was assumed that all these binary images are aligned. There are no shifts and rotations between different binary images. But in practice, there may be slight movements of the camera. These movements can be estimated before the reconstruction of the light intensity field. We can model this as a multi-channel binary sampling problem. Researchers have already used the multi-channel sampling methods modeling the image super-resolution problem, but the case of binary sampling not been considered so far.

How to use the binary sensor to capture color images is not explored in this thesis. We can use the same type of methods used in conventional image sensors, i.e., putting a color filter array (CFA) on top of the sensor. The color filter can be placed randomly like in the human retina or regularly like the Bayer filter array. Demosaicing algorithms are needed, if the sampling rate is not large enough. In this context, an interesting problem is the joint design of the threshold pattern and the color filter array to achieve optimal performance.
Bibliography


Curriculum Vitæ

Feng Yang

Audiovisual Communications Laboratory (LCAV)
School of Computer and Communication Sciences (I&C)
École Polytechnique Fédérale de Lausanne (EPFL)
1015 Lausanne, Switzerland

Personal

Date of birth: January 25, 1983
Nationality: Chinese

Education

2007–2012 Ph.D. in School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
2004–2007 M.Sc. in Automatic Control, Tsinghua University, Beijing, China
2000–2004 B.Eng. in Automatic Control, Tsinghua University, Beijing, China

Professional Experience

09/2007–02/2012 Research and teaching assistant, Audiovisual Communications Laboratory (LCAV), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
08/2010–02/2011 Research intern, Visual Computing and Ubiquitous Imaging Team, Nokia Research Center, Palo Alto, California, USA
09/2004–07/2007 Research and teaching assistant, Broadband Networks & Digital Media Lab, Department of Automation, Tsinghua University, Beijing, China
12/2006–06/2007 Research intern, Corporate Technology Group, Intel China Research Center, Beijing, China
Patents

2. Weidong Shi, Jun Yang, Feng Yang, and Yingen Xiong. Apparatus and methods of passive user identification using multiple sensing devices on a mobile computing system. (Application number: 12/979608)
3. Feng Yang, Yue M. Lu, and Martin Vetterli. A method, apparatus and system for image acquisition and conversion. (Filing number: US12/831,712)
5. Qionghai Dai, Feng Yang, and Guiguang Ding. A method for robust image compression based on distributed source coding. (Patent number: CN1988670)

Publications

Journal Papers


Conference Papers


**Awards and Honors**

- 2007 Outstanding Master Thesis of Tsinghua University
- 2005 Outstanding Student Scholarship of Tsinghua University (Rockwell scholarship)
- 2003 Outstanding Student Scholarship of Tsinghua University (GE scholarship)
- 2002 Second Prize in the 4th Electronic Design Competition, Tsinghua University
- 2001 Third Prize in the 18th College Physics Competition, Beijing
- 2000 Scholarship for Freshman, Tsinghua University (ranked 6th in Shaanxi Province, about 90,000 students)

**Academic Service**

- **Publicity**  *International Workshop on Emerging Mobile Sensing Technologies, Systems, and Applications* in conjunction with *Pervasive 2011*, San Francisco, USA
- **Reviewer**  *IEEE Transactions on Image Processing, IEEE Transactions on Circuits and Systems for Video Technology*