

Compressive 3D Scene Reconstruction Using Single-Photon Multi-spectral LIDAR Data

Rodrigo Caye Daudt

School of Engineering & Physical Sciences
Heriot-Watt University

A Thesis Submitted for the Degree of
MSc Erasmus Mundus in Vision and Robotics (VIBOT)

· 2017 ·

Abstract

Recent advances in single-photon detection technology led to the development of single-photon light detection and ranging (LIDAR) systems. With the ability to accurately count the detected photons, these systems allow for a more precise decoding of these signals to obtain reflectivity images and depth maps. With this capabilities, these systems can be pushed much further than other LIDAR systems with respect to low photon counts, but when these photon counts get extremely low more elaborate techniques are required to accurately decode these signals. In this work, convex optimization decoding methods for single-band and multi-spectral single-photon LIDAR systems are presented by proposing improvements on the state-of-the-art in single-photon LIDAR decoding. This is done by using physically accurate Poissonian measurement models for single-photon detections and by imposing low-rank priors to correlate different bands. The usage of compressive sampling is also explored as a way to optimize the acquisition strategy of scanning LIDAR systems.

Research is what I'm doing when I don't know what I'm doing...

Werner von Braun

Contents

Acknowledgments	vi
1 Introduction	1
1.1 LIDAR Technology	1
1.2 Problem Definition	2
1.3 Outline	3
2 Background	4
2.1 Single-photon LIDAR	4
2.2 Convex Optimization	8
2.3 Compressive Sensing	9
3 Methodology: Single Wavelength LIDAR Decoding	11
3.1 Main Considerations	11
3.2 Measurement Model	12
3.3 Decoding Method	13
3.3.1 Baseline Intensity	14
3.3.2 Response Intensity	15
3.3.3 Depth Estimation	16

4	Methodology: Multi-spectral LIDAR Decoding	18
4.1	Main Considerations	18
4.2	Measurement Model	19
4.3	Decoding Method	19
4.3.1	Baseline Intensity	19
4.3.2	Response Intensity	21
4.3.3	Depth Estimation	23
4.4	Comparison Against State-of-the-art	25
5	Results	26
5.1	Single-spectral Results	26
5.2	Multi-spectral Results	31
6	Conclusion	39
6.1	Main Considerations	39
6.2	Future Work	40
	Bibliography	45

List of Figures

1.1	Example of a LIDAR imaging system. The laser pulses are emitted and the impulse response is observed at each pixel.	2
2.1	Example of how the data is modelled for an example pixel whose response is composed of a main impulse response, a weaker secondary response and baseline.	5
2.2	When the acquisition time is short the observed signal becomes much harder to decode.	6
2.3	Real calibration data for a multi-spectral LIDAR system.	7
3.1	Illustration of the measurement model for a given pixel. Top: example of true response intensity per depth. Middle: expected photon count for each time bin, which depends on a , b and d . Bottom: simulated photon count y using the Poisson model.	14
5.1	Final depth estimation results obtained by our method and the ground truth depth for both images with 0.5 ppp on average for all cases.	27
5.2	Accuracy of the results considering different photon counts and subsampling ratios for both test images.	28

5.3	Baseline, reflectivity and depth results obtained from applying the proposed algorithm on real LIDAR data. The images on the left are from an acquisition with very weak baseline signal, while the one on the right was under direct sunlight, and therefore had a strong baseline signal.	29
5.4	Image of the setup used for acquiring single-spectral LIDAR images. The imaged object provides a realistic shape for evaluating the decoding algorithm.	31
5.5	Image of the setup used for acquiring multi-spectral LIDAR images. The imaged object provides realistic shape and varied spectral responses for evaluating the decoding algorithm.	32
5.6	Decoding results for real single-photon multi-spectral LIDAR data.	34
5.7	SNRs in reflectivity and depth estimation for each of the test cases.	35
5.8	Depth estimation for each test case.	36
5.9	Cumulative probability function for the absolute error in reflectivity estimation for each of the test cases.	37
5.10	Cumulative probability function for the absolute error in depth estimation for each of the test cases.	38

List of Tables

5.1	Approximate photon per pixel count and subsampling ratios	27
5.2	SNRs for the depth estimation results using the office image	30
5.3	SNRs for the depth estimation results using the cube image	30
5.4	Approximate photon per pixel per band count and subsampling ratios for multi-spectral simulations	33
5.5	SNRs for the reflectivity estimation results	33
5.6	SNRs for the depth estimation results	33

Acknowledgements

I would firstly like to thank my parents, Liliana and Geraldo, without whose support I would not be where I am.

I would like to thank my supervisors for this project, Prof. Yoann Altmann and Prof. Yves Wiaux, whose help made it possible for this work to come to fruition.

I am also thankful to the BASP group members, who helped me whenever I needed a hand and were always happy to share with me their expertise.

Finally, thank you to all friends, family, colleagues and professors who have been a part of this journey in different yet essential ways.

Chapter 1

Introduction

1.1 LIDAR Technology

Light Detection and Ranging (LIDAR) systems consist of imaging equipment capable of capturing reflectivity and depth information simultaneously. It does so by emitting streams of photons from laser sources for each pixel and building a histogram of the reflected light pulse over time. The intensity of the observed response is proportional to the reflectivity of the object being imaged at that point, and the depth is associated with the delay of the impulse response [1]. LIDAR systems can work by scanning the image pixel by pixel or by measuring all the pixels simultaneously. In this work, the considered LIDAR systems work by scanning the pixels in the image one at a time.

Advances in single-photon detection technology using single-photon avalanche diodes (SPAD) enabled single-photon LIDAR systems with enhanced photon timing and thus ranging capability, which allows 3D scenes analysis with shorter acquisition times and lower power consumption (e.g. for eye-safety purpose). These single-photon LIDAR systems usually have higher resolution and require a smaller light flux than other LIDAR systems. However, the consideration of such systems makes the decoding of LIDAR signals particularly challenging when the number of detected photons in each pixel is very low. Decoding a LIDAR signal consists of extracting depth and reflectivity information for each of the imaged pixels. A diagram of a single-photon LIDAR system can be seen in Fig. 1.1.

The applications of LIDAR systems are vast. Its use is widespread in fields such as geographical information systems (GIS) [2,3], underwater imaging [4,5] and forest monitoring [6–8], enabling fast and accurate surveying of large areas. More recently, LIDAR has been adopted by the autonomous driving industry as a complement to the collections of cameras and sensors

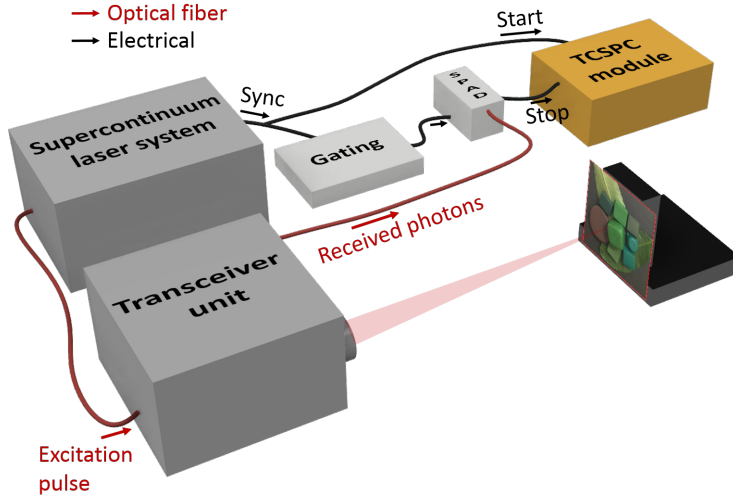


Figure 1.1: Example of a LIDAR imaging system. The laser pulses are emitted and the impulse response is observed at each pixel.

present in modern autonomous vehicles [9].

When decoding single-photon LIDAR data, the standard assumption of additive Gaussian noise that is commonly used in imaging problems is no longer adequate. The discrete nature of the photon detections when building the detection histograms for each pixel is much more accurately modelled by Poissonian statistics.

Some LIDAR systems, named multi-spectral LIDAR, are able to work with different light wavelengths. By repeating the acquisition process using a range of light wavelengths, this system is able to obtain reflectivity information for each wavelength as well as more reliable depth information. The acquisition time for multi-spectral LIDAR systems is not necessarily longer than that of single band LIDAR since it can be equipped with multiple photon detectors that acquire data on all bands simultaneously. The multi-spectral response intensity information may be used, for example, for classifying the observed material at a given pixel, which is very useful in applications such as remote sensing.

1.2 Problem Definition

The first objective of this work is to propose convex optimization methods for decoding single- and multi-spectral single-photon LIDAR data as accurately as possible for cases where the photon per pixel count is extremely low. Convex optimization methods provide a good

balance between computation speeds and flexibility for solving different types of problems. They are fast and reliable ways of dealing with some problems and are well adapted for dealing with large images.

The second objective of this work is to explore the impact of pixel subsampling and compressive sensing reconstruction to the aforementioned LIDAR systems. Given that the considered LIDAR system consists of a single laser emitter/receptor that scans the image, and that we have full control over its behaviour, the idea of sampling fewer pixels with higher quality (i.e. longer acquisition time) is explored in this work.

1.3 Outline

Chapter 2 contains a review of the state-of-the-art methods in single-photon multi-spectral LIDAR decoding technology, relevant convex optimization methods, and a quick review of compressive sensing theory. Chapter 3 details the proposed method for decoding single-spectral LIDAR. Chapter 4 expands the approach proposed in Chapter 3 to account for multi-spectral LIDAR. Chapter 5 discusses experimental results that evaluate the algorithms proposed in chapters 3 and 4. Chapter 6 concludes this work by summarising the main findings from the previous chapters and possible improvements in future work.

Chapter 2

Background

This chapter aims to discuss the state-of-the-art in the area of single and multi-spectral LIDAR decoding, as well as reviewing the relevant theory in convex optimization algorithms and compressive sensing that will be essential to the following chapters.

2.1 Single-photon LIDAR

LIDAR systems work by emitting a light pulse (from a laser or LED) in a given direction (pixel) and building a histogram of the response over time. In case there is an object in the range and field of view of the LIDAR system, the measured response should have a peak which depends on the position of the imaged object, on its reflectivity for the light's wavelength and on the angle between the emitted laser pulse and the object's surface at that point.

LIDAR decoding uses calibration data which has been obtained in a controlled environment. This data is obtained by imaging an object with a known reflectivity and at a known position relative to the LIDAR system. This calibration data contains the shape and intensity of the impulse response for that material, and can be used to calculate the depth and reflectivity in newly acquired data by comparing the magnitude and the delay of the impulse responses. It is assumed that the shape of the impulse response does not vary with the depth of the imaged surface. Another assumption that is made here is that the angle between the surface and the light beam only affects the intensity of the response and not its shape, which is an adequate assumption in case the beam of light is narrow.

In a real world environment, it is common to have other sources of light present in the scene other than the LIDAR laser emitter. These other sources of light can be detected by LIDAR systems and are called baseline response. The baseline make the decoding of the signal harder, given that baseline photon counts contain no information about either reflectivity or depth.

The baseline image is very similar to a regular passive photograph, given that it is measuring the light present in the scene that does not come from the light source in the LIDAR system.

In the case of translucent objects, edges and occlusions it is possible that the observed response at a given pixel contains impulse response from objects with different reflectivities and depths. Given that in most cases the response comes only from a single surface given the divergence of the laser beam is small, this phenomenon is ignored by many decoding algorithms.

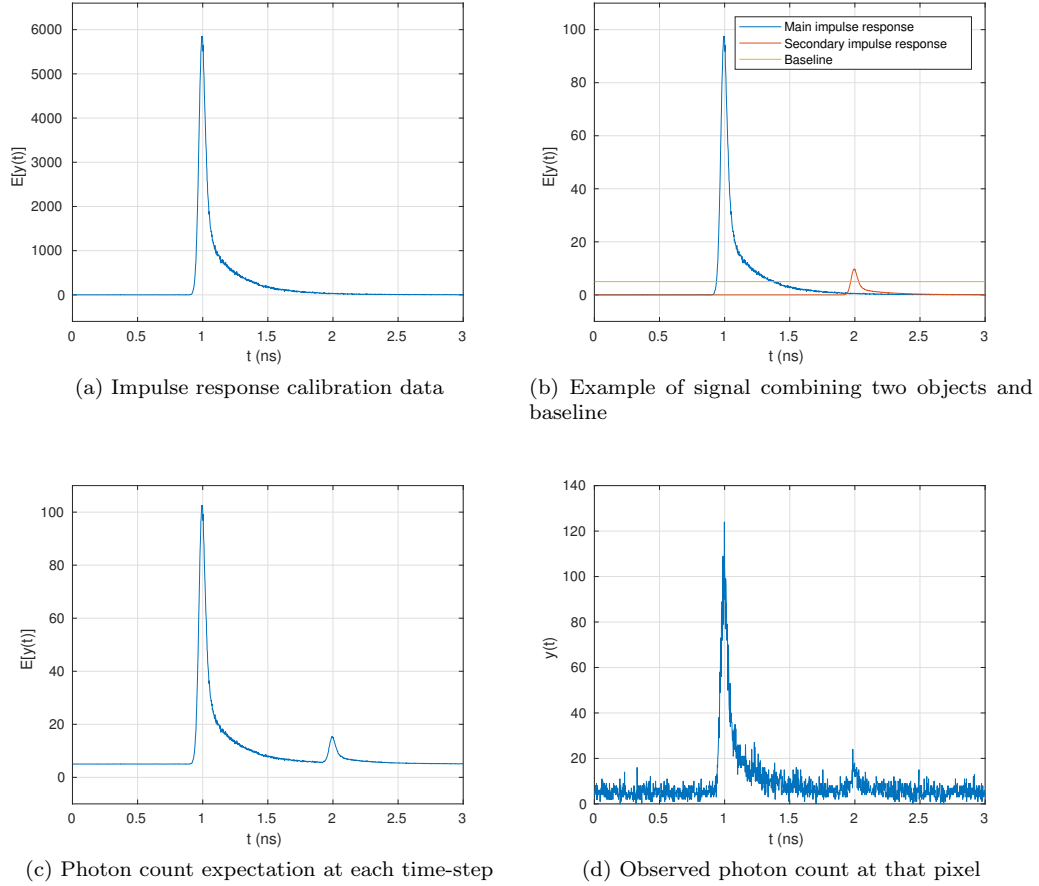


Figure 2.1: Example of how the data is modelled for an example pixel whose response is composed of a main impulse response, a weaker secondary response and baseline.

Figure 2.1 illustrates the decomposition of the observed photon counts. Figure 2.1a contains the impulse response that was obtained during the calibration process in a controlled environment. Figure 2.1b illustrates a possible composition for the response of a given pixel: main impulse response from the object covering most of the area of the pixel, secondary response from

an object covering a small area of the pixel, and baseline response constant over time. These three components are combined into a single expectation of number of observed photons, which is depicted in Fig. 2.1c. Figure 2.1d shows the actual observed photon count, which follows a Poisson distribution with the expectation shown in Fig. 2.1c for each bin of the histogram.

Ideally, the decoding process will be able to use the calibration data to infer the information present in Fig. 2.1b from the observed photon count in Fig. 2.1d, which is the only data available for that pixel. When the photon count is very low, decoding the data accurately becomes more and more challenging. Low photon counts may happen due to low laser power or fast acquisition times. An illustration of such an acquisition can be seen in Fig. 2.2. It is clear that in a case with such few observed photons the shape of the photon expectation is much less clear in the photon count histogram. Given that in this work the aim is to decode signals with an average photon per pixel count as low as 0.5, this must be taken into account.

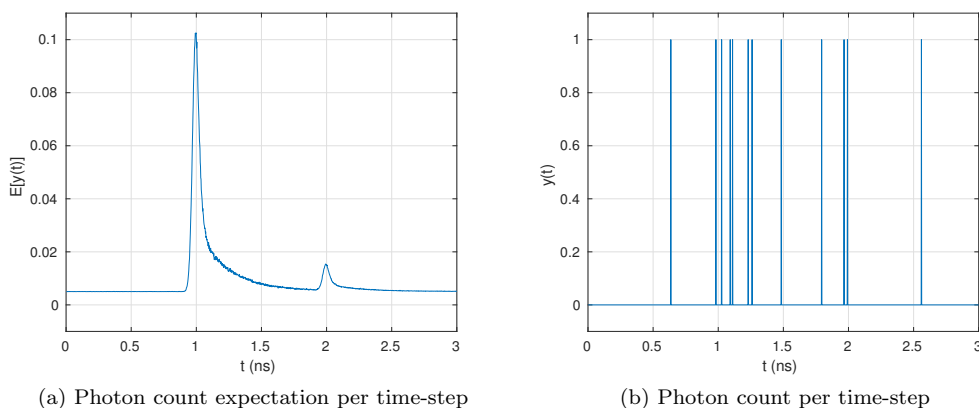


Figure 2.2: When the acquisition time is short the observed signal becomes much harder to decode.

One recent approach for decoding single-photon LIDAR was proposed by Shin et al. in [10, 11]. The main assumption for this method is that the impulse response at a given pixel is the sum of a limited number of objects with different reflectivities and different depths. This is a valid assumption, and it is able to model translucent objects and pixels that contain edges, for example. A union-of-subspaces model is proposed to impose a restriction to the solution space for each pixel, which mathematically models the assumption of limited responses in a pixel. The problem is formulated as an optimization problem and is solved using a modified version of the CoSaMP algorithm [12].

This method was used when the photon count was not as low as in some of the other methods in the state-of-the-art. This is not surprising, as it does not exploit the potential spatial

correlation between neighbouring pixels and therefore needs to observe a sufficient number of photons in each pixel to produce good results. This method has also no obvious extension to the case of multi-spectral LIDAR that allows correlation between different wavelengths.

Shin et al. proposed a different decoding technique in [13,14] for decoding single-photon LIDAR signals. It assumes a baseline noise that is ergodic for all the measured pixels and that this baseline noise is known before the acquisition and decoding of the signal. Neither of these assumptions is adequate for several applications. The intensity of the baseline signal is pixel dependent, as will be shown in data presented in the next chapter. Furthermore, the assumption that the baseline signal can be estimated during the calibration process does not hold for any use case where the LIDAR system should be mobile.

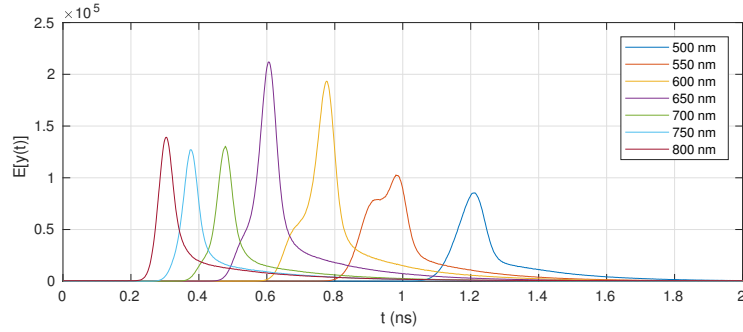


Figure 2.3: Real calibration data for a multi-spectral LIDAR system.

Despite these assumptions, this method is able to achieve good results for extremely low photon counts. This is achieved by correlating the pixels spatially using total variation (TV) regularisation during the decoding process. This method also assumes a known analytic form for the impulse response, and uses this knowledge to model depth estimation as a convex estimation problem. This assumption can be restrictive, as in the case for multi-spectral LIDAR the shape of the impulse responses may vary significantly and modelling it as a log-concave function does not adequately model impulse responses in different wavelengths. Figure 2.3 contains real calibration data for different wavelengths of a multi-spectral LIDAR system. It is clear that these impulse responses have no simple analytical form that accurately describes them, and the authors have not disclosed the log-concave function used on their model in the paper.

The decoding method proposed in this work uses this method as a starting point, given its efficiency in the case of low photon counts, and the proposed convex optimization problems can be extended to the multi-spectral case in a meaningful way. Each of the challenged assumptions is analysed and replaced by more flexible assumptions that lead to a more accurate and general algorithm.

In [15], Altmann et al. proposed a method for analysis of multispectral single-photon LIDAR data. In a Bayesian context, this algorithm propose simultaneous decoding and spectral unmixing using the Markov chain Monte Carlo (MCMC) method. This method uses previous knowledge about the spectral response of the materials in the scene to help in the decoding process by estimating the abundances of each material in each pixel. In [16], convex optimization is used along the known spectral responses of the materials in the scene aiming to minimize the number of materials in each pixel while maintaining the data fidelity and total variation terms low. In this case, previous knowledge about the spectral response of all the materials in the scene is also necessary for the decoding to take place.

2.2 Convex Optimization

Many problems in modern signal processing and computer vision are solved by solving an optimization problem. A cost function is proposed, and an optimization algorithm is used to find its global minimum. Therefore, when dealing with problems in this way there are two main parts to focus on: the design of the cost function, and the choice of algorithm used to find its global minimum.

The design of the cost function depends on the problem at hand. More complex cost functions may lead to better results but may be harder to optimize. One such case is in the convexity of the proposed cost function. Non-convex cost functions are able to model a wider range of problems, but they are significantly harder to optimize.

If the proposed cost function is convex, a class of optimization algorithms called convex optimization algorithms can be used to find the minimum point of that function. The knowledge that the cost function is convex can be used to build optimization algorithms that are faster than general optimization algorithms based on properties of the cost function, e.g. a local minimum must be a global minimum.

One algorithm used for convex optimization problems is the Alternating Direction Method of Multipliers (ADMM) [17]. ADMM is a simple yet powerful convex optimization algorithm. While the algorithm itself is not recent, it is still widely used due to its versatility and simplicity. Assume we have a problem with the following form:

$$\begin{aligned} &\text{Minimize: } f(x) + g(z) \\ &\text{Subject to: } \mathbf{A}x + \mathbf{B}z = c \end{aligned} \tag{2.1}$$

Where $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^m \rightarrow \mathbb{R}$, $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{B} \in \mathbb{R}^{p \times m}$, and $c \in \mathbb{R}^p$. This problem has the following augmented Lagrangian form:

$$L_\rho(x, z, y) = f(x) + g(z) + y^T(\mathbf{A}x + \mathbf{B}z - c) + (\rho/2)\|\mathbf{A}x + \mathbf{B}z - c\|_2^2 \quad (2.2)$$

ADMM solves the problem defined in Eq. 2.1 iteratively using its augmented Lagrangian form shown in Eq. 2.2. The iterations for the scaled form of the ADMM algorithm are:

$$\begin{aligned} x^{k+1} &:= \arg \min_x (f(x) + (\rho/2)\|\mathbf{A}x + \mathbf{B}z^k - c + u^k\|_2^2) \\ z^{k+1} &:= \arg \min_z (g(z) + (\rho/2)\|\mathbf{A}x^{k+1} + \mathbf{B}z - c + u^k\|_2^2) \\ u^{k+1} &:= u^k + \mathbf{A}x^{k+1} + \mathbf{B}z^{k+1} - c \end{aligned} \quad (2.3)$$

The convergence of the ADMM algorithm is guaranteed given a few weak assumptions on f and g . For more information on the derivations, proofs and other properties of the ADMM the reader should refer to the review by Boyd et al. [17].

Figueiredo et al. have developed an algorithm named PIDAL which is used to denoise images corrupted by Poissonian noise [18]. This algorithm is an extension of the scaled ADMM algorithm to deal with a specific problem formulation involving a Poissonian negative log-likelihood. The problem solved by the proposed PIDAL algorithm is very similar to the problems proposed for single-photon LIDAR decoding described in [13, 14]. Moreover, Figueiredo et al. describe in detail the process for adapting the scaled form of the ADMM algorithm into three forms of the PIDAL algorithm. Following the same steps proposed in the paper, in Chapters 3 and 4 convex optimization algorithms will be proposed to solve the problems proposed to our specific applications.

2.3 Compressive Sensing

Compressive sensing is the study of techniques that allow the recovery of signals from underdetermined linear systems [19, 20]. These techniques allow for the acquisition of signals with sub-Nyquist sampling in some cases, and are extremely useful in cases where it is desirable to subsample the signal for convenience (i.e. reducing acquisition time) or to reduce costs, as in the cases of magnetic resonance images [21–23] and radio astronomy [24, 25].

These underdetermined systems are combined with prior knowledge about the nature of the signal to be recovered in order to define a tractable minimization problem. This prior knowledge is usually described by the sparsity of the signal in a known basis. Examples of such priors include total variation (TV) [26–28], which is related to sparsity in the image gradient, and sparsity in different wavelet bases [25, 29].

Compressive sensing problems are generally formulated as an optimization problem based

on the measurement model of the problem at hand. These problems can be written as Tikhonov regularization problems [13, 14, 18, 30], Morozov regularization problems [23], or Ivanov regularization problems [24, 25, 29, 31].

Consider a complex vector $\mathbf{x} \in \mathbb{C}^N$, an orthogonal basis $\Psi \in \mathbb{C}^{N \times N}$. Vector \mathbf{x} is said to be K -sparse in base Ψ if $\Psi^\dagger \mathbf{x}$ has K or fewer non-zero elements, i.e. $\|\Psi^\dagger \mathbf{x}\|_0 \leq K$. Assume that this vector \mathbf{x} is probed by a linear measurement operator $\Phi \in \mathbb{C}^{M \times N}$ with additive noise, where $M \ll N$. In other words, $\mathbf{y} = \Phi \mathbf{x} + \mathbf{n}$, where $\mathbf{y} \in \mathbb{C}^M$.

In some cases the sparsity promoting convex relaxation ℓ_1 norm is used instead of the ℓ_0 norm so that convex optimization algorithms may be used to find a solution to the problem. The recovery of \mathbf{x} from \mathbf{y} in this case can be formulated in three ways:

$$\begin{aligned}
 \text{Tikhonov formulation: } \hat{\mathbf{x}} &= \arg \min_x \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda \|\Psi^\dagger \mathbf{x}\|_1 \\
 \text{Morozov formulation: } \hat{\mathbf{x}} &= \arg \min_x \|\mathbf{y} - \Phi \mathbf{x}\|_2^2, \text{ s.t. } \|\Psi^\dagger \mathbf{x}\|_0 < K \\
 \text{Ivanov formulation: } \arg \min_x &\|\Psi^\dagger \mathbf{x}\|_1, \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{x}\|_2 < \epsilon
 \end{aligned} \tag{2.4}$$

In cases where the noise statistics are well known, the Ivanov formulation, also known as constrained formulation, is generally preferred. In this work, however, the Tikhonov formulation, also known as unconstrained formulation, was adopted since the Euclidean distance was not an appropriate measure, as will be discussed in the next chapter. Once the appropriate minimization problem has been proposed, an optimization algorithm can be chosen to find its solution accordingly.

Chapter 3

Methodology: Single Wavelength LIDAR Decoding

In this chapter, the proposed method for decoding single-photon LIDAR signals on one wavelength is presented. The method is described along with its motivations, assumptions, measurement model and the necessary mathematical derivations.

3.1 Main Considerations

In order to reduce overall acquisition times, it is interesting to analyse whether it is preferable to reduce the number of measurements (number of pixels sensed for each wavelength) and recover the signal using compressive sensing techniques, rather than reducing the per-pixel acquisition time uniformly, as is commonly done [15, 16, 32]. In this chapter we consider a single-photon LIDAR system with only one wavelength and reduce the number of sensed pixels. A decoding pipeline is proposed that estimates the depth of a scene with an unknown baseline signal (which is related to the level of ambient illumination) that is pixel-dependent.

This method builds on the algorithm proposed by Shin et al. in [13, 14]. As mentioned in Chapter 2, the method proposed by Shin et al. makes assumptions that are not always adequate and that should be addressed in order to improve the performance and versatility of the algorithm. These three assumptions are:

1. Shin’s algorithm assumes ergodicity in the baseline signal, i.e. that the baseline signal has the same intensity on all imaged pixels. This assumption is often incorrect, as will be seen in real data results shown later in this chapter. Therefore, it is necessary to extend the mathematical model to be able to model different baseline intensities at different pixels.

2. The second incorrect assumption is that the baseline intensity can be measured during calibration procedure and that value can be used on all subsequent measurements. Although that may be the case in controlled static environments, the moment a LIDAR system is moved this assumption becomes invalid. It is essential then to estimate the baseline value from each measurement individually in applications such as the ones described in [2, 3, 6, 7]. This assumption is replaced here by the assumption that we can define a minimum distance between the LIDAR system and any imaged object.
3. The final assumption that is contested here is that we can formulate an analytic expression that accurately describes the impulse response function. As was shown in Section 2.1, this assumption does not generalise well to the multi-spectral case, and therefore is not adequate for this work. This assumption is made in order to allow for depth calculation using convex optimization algorithms and achieving fast calculation results by avoiding direct likelihood calculations. The algorithm proposed in this chapter achieves fast calculation speeds for likelihoods by exploiting the low number of measured photons instead of making approximations, which can be used directly with any impulse response obtained during the calibration process regardless of its shape or whether it has a log-concave analytical form.

The decoding is performed in three sequential steps: 1) baseline intensity estimation, 2) response intensity estimation and 3) depth estimation. Each step is modelled as a Tikhonov regularization problem. The order in which these steps are performed is important, as the result of each step is necessary for the following step.

To account for the Poisson noise model, an extension of the PIDAL [18] algorithm is used for background and target intensity estimation. Moreover, the regularization parameters, which control the smoothness of the estimated background and target intensity profiles, can automatically adjusted using approaches such as the one described in [33] or SURE based methods [34].

The final estimated depth profile is finally obtained by refining the depth estimated via maximum likelihood estimation (conditioned on the previously estimated background and target intensities) using convex optimization and compressive sensing signal recovery techniques.

3.2 Measurement Model

In this work we focus on decoding signals composed of extremely low photon per pixel (ppp) count (~ 0.5 ppp) measurements. Thus, Gaussian noise models are not well adapted to model observation noise. As described in [14], the Poisson noise assumption, which is adopted here, is much more relevant considering the discrete nature of the measurements (photon detections).

The observation model can be described as follows. We consider a measurement matrix $\mathbf{F}_{\text{calib}}$ which is constructed during the calibration of the equipment. Each column of this matrix contains the instrumental response associated with each admissible depth, and is obtained by imaging a reference object with unitary reflectivity over a long acquisition time (100s here) to reduce the impact of noise. For each scenario considered in Section 5.1, this matrix then is scaled according to the per-pixel acquisition time considered.

$$F = F_{\text{calib}} \left(\frac{\text{acquisition time in seconds}}{100} \right) \quad (3.1)$$

Assuming the response in each pixel is based on the reflection of photons onto a single surface assumed orthogonal to the incident beam, we can define for pixel p a vector $\mathbf{x}_p = [0 \dots a_p \dots 0]^T$, where a_p is the material response intensity (which is wavelength dependent), and its position d_p is related to the detected depth z_p by

$$z_p = \frac{d_p T_{\text{bin}} c}{2}$$

where T_{bin} is the duration of each time bin used for creating the photon histogram and c is the speed of light in the medium (e.g. water or air). For the experiments conducted in the following sections, $T_{\text{bin}} = 2$ ps.

We then model the detected photon count $y_{p,t}$ for pixel p and time bin t as a random variable drawn from a inhomogeneous (time dependent) Poisson distribution according to the following model

$$\mathbf{y}_p \sim Pr(\mathbf{F}\mathbf{x}_p + b_p)$$

where $Pr(\lambda)$ denotes a Poisson distribution with mean λ and $\mathbf{y}_p = [y_{p,1}, \dots, y_{p,N_{\text{bins}}}]^T$. An illustration of such model using a real calibration matrix \mathbf{F} is displayed in Fig. 3.1. In this case, $b_p = 0.1$, $a_p = 1$, $d_p = 125$.

3.3 Decoding Method

The proposed decoding method is separated into three sequential steps: baseline estimation, response intensity estimation and depth estimation. This method assumes we can separate the data into two parts. The red dotted lines in Fig. 3.1 represent the depth/range below which we assume that no object is present, and thus the signal corresponds solely to the baseline response.

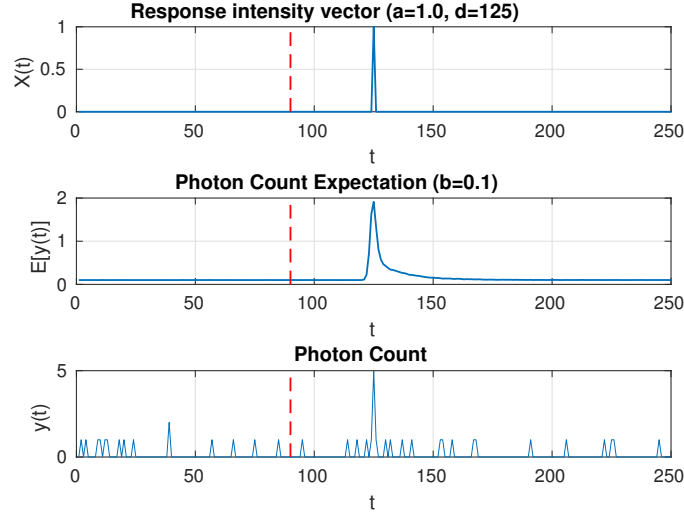


Figure 3.1: Illustration of the measurement model for a given pixel. Top: example of true response intensity per depth. Middle: expected photon count for each time bin, which depends on a , b and d . Bottom: simulated photon count y using the Poisson model.

3.3.1 Baseline Intensity

Starting from the assumption that before a time T_b only baseline photons are detected, we select for the baseline estimation only the bins with $t \leq T_b$ from the photon count histograms. For these measurements, we can model the photon count $y_{p,t}$ for a given pixel p and bin t as a random variable

$$y_{p,t} \sim Pr(b_p)$$

where the baseline (or background) level b_p is pixel dependent. This is equivalent to writing the following probability mass function:

$$P(y_{p,t} = k) = \frac{b_p^k e^{-b_p}}{k!}.$$

Considering all the measurements $\{y_{p,t}\}$ in a single matrix $\mathbf{Y} \in \mathbb{N}^{T_b \times P}$, the background levels $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_P]^T$ could be estimated by minimizing the following negative log-likelihood

$$-\log(P(\mathbf{Y}|\mathbf{b})) = \sum_{p=1}^P \left[-\log(b_p) \sum_{t=1}^{T_b} y_{p,t} + T_b b_p + \sum_{t=1}^{T_b} \log(y_{p,t}!) \right] \quad (3.2)$$

but when the number of photons is extremely small, this pixel-wise approach yields poor results (likelihood weakly informative). Moreover, Eq. (3.2) assumes that all the P pixels are observed.

Consider S_α a set of pixels randomly chosen out of the P pixels of the grid, where the subsampling ratio $0 < \alpha \leq 1$ corresponds to the proportion of spatial locations sampled. In such a case, the data fidelity term (3.2) becomes

$$\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{b}) = \sum_{p \in S_\alpha} [-\log(b_p)\sigma_p + T_b b_p], \quad (3.3)$$

where $\sigma_p = \sum_{t=1}^{T_b} y_{p,t}$. Given that (3.3) is convex, we can then formulate the following convex optimization problem

$$\hat{\mathbf{b}} = \arg \min_{\mathbf{b}} [\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{b}) + \tau p_b(\mathbf{b}) + i_{\mathbb{R}^+}(\mathbf{b})] \quad (3.4)$$

that aims to minimize the negative log-likelihood along with an appropriate regularization while ensuring the positivity of the estimated background profiles $\hat{\mathbf{b}}$ (through an indicator function $i_{\mathbb{R}^+}(\cdot)$).

Various convex regularizations $p_b(\mathbf{b})$ can be used, including total variation [26–28], Laplacian spatial filters [35], or ℓ_1 -based sparsity promoting regularizations (e.g. of wavelet coefficients [25, 29, 36, 37]). The parameter $\tau > 0$ controls the impact of the regularization on the final solution. Using convex regularisers, the problem (3.4) can be solved for instance using the PIDAL-TV or PIDAL-FA algorithms proposed in [18] or SPIRAL [38].

The hyperparameter τ may be automatically chosen by iteratively solving the minimization problem and updating τ using $\tau = P/(p_b(\hat{\mathbf{b}}) + 1)$ as was proposed in [33]. This increases the necessary time to reach a solution to the problem, but avoids the manual choice of optimization parameters.

3.3.2 Response Intensity

The method for estimating the response intensity per pixel is similar to the method used for baseline estimation, but we need to take into account our baseline estimate $\hat{\mathbf{b}}$ for estimating the response intensity $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_P]^T$. Considering now the second part of our split data of duration T_a , we need to consider simultaneously the baseline and the impulse response to model the photon reception at pixel p and bin t as

$$y_{p,t} \sim Pr(a_p f_p(t) + b_p) \quad (3.5)$$

where $f_p(t)$ is the column of \mathbf{F} relative to the depth of the observed object in pixel p .

We can then define a new variable $\boldsymbol{\sigma} = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_P]^T$ where $\sigma_p = \sum_{t=1}^{T_a} y_{p,t}$. Since σ_p

is a sum of independent Poisson random variables, σ_p follows also a Poisson distribution. We assume here that the impulse response is never truncated, and therefore the columns of \mathbf{F} have the same ℓ_1 -norm (same sum as the entries of \mathbf{F} are positive), regardless of the observed depth. From (3.5), we obtain

$$\sigma_p \sim \text{Pr}(a_p \sigma_f + T_a b_p), \text{ where } \sigma_f = \sum_t f_p(t). \quad (3.6)$$

Following a similar procedure as in Section 3.3.1, the data fidelity term (with respect to \mathbf{a}) can be expressed as

$$\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{a}|\mathbf{b}) = \sum_{p \in S_\alpha} [a_p \sigma_f - \sigma_p \log(a_p \sigma_f + T_a b_p)], \quad (3.7)$$

which, after replacing \mathbf{b} by its estimates $\hat{\mathbf{b}}$, leads to the following convex optimization problem

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} [\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{a}|\hat{\mathbf{b}}) + \tau p_a(\mathbf{a}) + i_{\mathbb{R}^+}(\mathbf{a})] \quad (3.8)$$

since (3.7) is convex and by choosing an appropriate convex prior $p_a(\mathbf{a})$.

Note that this problem can not be directly solved by PIDAL since $\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{a}|\hat{\mathbf{b}})$ depends on $\hat{\mathbf{b}}$. This can be solved by calculating the Moreau proximity operator of $\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{a}|\hat{\mathbf{b}})$ and changing the corresponding step of the PIDAL algorithm. In Eq. (3.9), the symbol \odot represents the Hadamard or element-wise product.

$$\begin{aligned} \text{prox}_{\mathcal{L}_{\mathbf{Y},\alpha}/\mu}(\boldsymbol{\nu}) &= \arg \min_{\mathbf{a}} \left[\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{a}|\hat{\mathbf{b}}) + \frac{\mu}{2} \|\mathbf{a} - \boldsymbol{\nu}\|_2^2 \right] = \\ &= \frac{1}{2} \left(\boldsymbol{\nu} - \frac{\sigma_f}{\mu} - \frac{T_a \hat{\mathbf{b}}}{\sigma_f} + \sqrt{\left(\boldsymbol{\nu} - \frac{\sigma_f}{\mu} - \frac{T_a \hat{\mathbf{b}}}{\sigma_f} \right)^2 + \frac{4\sigma}{\mu} + \frac{4T_a \hat{\mathbf{b}} \odot \boldsymbol{\nu}}{\sigma_f} - \frac{4T_a \hat{\mathbf{b}}}{\mu}} \right) \end{aligned} \quad (3.9)$$

3.3.3 Depth Estimation

Once the estimates $\hat{\mathbf{b}}$ and $\hat{\mathbf{a}}$ have been calculated (by solving (3.4) and (3.8), respectively), the last step is the depth estimation using the available data. For each pixel, the log-likelihood is not convex with respect to d_n unless approximations on the shape of the impulse response are used. Here, the log-likelihood defined by

$$\log(P(\mathbf{Y}|\hat{\mathbf{b}}, \hat{\mathbf{a}}, \mathbf{d})) = \sum_{p \in S_\alpha} \sum_{t=1}^{T_b+T_a} \left[y_{p,t} \log(\hat{a}_p f_{d_p}(t) + \hat{b}_p) - (\hat{a}_p f_{d_p}(t) + \hat{b}_p) \right] + C, \quad (3.10)$$

where C is a constant, is computed without any approximation.

By assuming that $\sum_{t=1}^P f_{d_p}(t)$ does not depend on d_p we can further simplify the log-likelihood computation. Furthermore, computations can be sped up by taking advantage of the fact that $y_{p,t}$ is extremely sparse when the photon count is low (the sum in $\log(P(\mathbf{Y}|\hat{\mathbf{b}}, \hat{\mathbf{a}}, \mathbf{d}))$ reduces to a much smaller number of terms). Therefore, the calculations for the log-likelihood are given by:

$$\log(P(\mathbf{Y}|\hat{\mathbf{b}}, \hat{\mathbf{a}}, \mathbf{d})) = \sum_{p \in S_\alpha} \sum_{t, y_{p,t} \neq 0} [y_{p,t} \log(\hat{a}_p f_d(t) + \hat{b}_p)] + C \quad (3.11)$$

Note that this formulation avoids using approximations (as in [13, 14, 39]) and directly uses the calibration data from each LIDAR system, which will be more scalable when considering multispectral LIDAR systems. For each observed pixel, the maximum likelihood estimator of d_p , denoted by $d_{p,ML}$, is obtained by exhaustive search, i.e., by finding for each pixel the column of \mathbf{F} which maximizes $\log(P(\mathbf{Y}|\hat{\mathbf{b}}, \hat{\mathbf{a}}, \mathbf{d}))$. Then, we can then solve a convex optimization problem to estimate the complete depth vector $\hat{\mathbf{d}}$ as an inpainting problem. Here we use a weighted Euclidean distance as data fidelity term along a convex regularization $p_d(\mathbf{d})$ to build a minimization problem that aims to estimate $\hat{\mathbf{d}}$. Several techniques exist to solve such problems [40]. The proposed minimization problem can be once solved by using the ADMM algorithm once again, or other convex optimizations such as ISTA [41] and FISTA [28].

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d}} [(d_{ML} - \mathbf{M}_\alpha \mathbf{d})^T \mathbf{W} (d_{ML} - \mathbf{M}_\alpha \mathbf{d}) + \tau p_d(\mathbf{d})] \quad (3.12)$$

In (3.12), \mathbf{M}_α is the selection matrix relative to S_α . Moreover, a diagonal weighting matrix \mathbf{W} is introduced to attribute higher confidence values to the pixels where more information is available, i.e. where more photons have been observed. \mathbf{W} is defined as:

$$\mathbf{W} = \text{diag}(\log(1 + \boldsymbol{\sigma}))$$

Chapter 4

Methodology: Multi-spectral LIDAR Decoding

An extension of the decoding algorithm presented in Chapter 3 to the case of multi-spectral single-photon LIDAR signals is presented in this Chapter. The algorithm presented previously could be applied to each wavelength of a given multi-spectral LIDAR signal individually, given all calibration data is provided. In this chapter we aim to correlate information in different wavelengths during the decoding process to push further the efficacy of the algorithm regarding the observed number of photons per pixel.

4.1 Main Considerations

The algorithm presented in this chapter for single-photon multi-spectral decoding starts from the same assumptions made in Section 3.1. The decoding algorithm still works in three sequential steps: baseline estimation, reflectivity estimation, and depth estimation. The measurement model is updated to cope with the new multi-spectral data and some associated phenomena, and the appropriate modifications are made to the algorithm.

In many cases, the wavelengths imaged by the multi-spectral system are very close to one another. In the case of the data used at the end of this chapter, for example, the wavelengths of the used lasers were between 500nm and 820nm in steps of 10nm. Furthermore, the surface responsible for the main response in each pixel is the same for all wavelengths. Thus, it is sensible to assume there is a high correlation between the baseline images and reflectivity images in different wavelengths.

It is important to note that in the single-band case, the accurate estimation of reflectivity α

was important because it led to accurate depth estimation results. In the case of multi-spectral signals, it is also the case but the way the reflectivity changes with the wavelength of the laser can also be used to classify the imaged material in several applications. Although material classification is outside the scope of this work, we acknowledge that the accurate estimation of reflectivity images is of even more importance in the multi-spectral case.

4.2 Measurement Model

The modelling of the acquisition process for the multi-spectral LIDAR is very similar to the single-spectral case, but it is important to address some differences. The first and most important difference is that the number of baseline and reflectivity variables scales with the number of wavelengths L that are imaged by the LIDAR system, while \mathbf{d} does not. In mathematical terms, $\mathbf{B} \in \mathbb{R}^{N \times L}$, $\mathbf{A} \in \mathbb{R}^{N \times L}$, and $\mathbf{d} \in \mathbb{R}^N$.

Another important difference is that the impulse response is different for each of the considered wavelengths and is calibrated separately. That is to say that the calibration matrix \mathbf{F} considered in Section 3.2 is now a collection of L matrices \mathbf{F}_ℓ , each of which containing the impulse response for surfaces at different depths for each wavelength.

The final difference between the single-spectral and the multi-spectral cases considered here is that there is a pixel and wavelength dependent vignetting effect that takes place during the measurement, not unlike the phenomenon that happens with regular cameras, which takes place due to physical limitations of the system (e.g. lenses) [15]. This vignetting effect is modelled using a multiplicative factor $r_{p,\ell}$ at pixel p and wavelength ℓ that is applied to the impulse response intensity. This vignetting factor is estimated during the calibration process and does not change after that.

Given these considerations, the measurement model used here for the photon count histogram $\mathbf{y}_{p,\ell}$ for pixel p and wavelength ℓ , for the single-photon multi-spectral LIDAR system can be written as in Equation (4.1), where $\mathbf{x}_{p,\ell}$ is the one-hot vector $\mathbf{x}_{p,\ell} = [0 \dots a_{p,\ell} \dots 0]^T$.

$$\mathbf{y}_{p,\ell} \sim Pr(r_{p,\ell} \mathbf{F}_\ell \mathbf{x}_{p,\ell} + b_{p,\ell}) \quad (4.1)$$

4.3 Decoding Method

4.3.1 Baseline Intensity

Just as for the single-spectral case, the first step in decoding single-photon multi-spectral LIDAR signals is to estimate the baseline. The baseline is now defined as $\mathbf{B} \in \mathbb{R}^{P \times L}$, where P is the number of imaged pixels and L is the number of wavelengths that the LIDAR system is

able to image. \mathbf{B} is the concatenation $[\mathbf{b}_1 \dots \mathbf{b}_\ell \dots \mathbf{b}_L]$ of the baselines at each wavelength. The measurements for the time-steps where no response is observed yet can then be modelled as:

$$y_{p,t,\ell} \sim Pr(b_{p,\ell})$$

The same steps taken in Section 3.3.1 can then be taken with this new formulation. Assume $\Sigma \in \mathbb{R}^{P \times L}$ such that $\sigma_{p,\ell} = \sum_{t=1}^{T_b} y_{p,t,\ell}$. Consider also once again $S_{\alpha,\ell}$ a set of pixels randomly chosen out of the P pixels of the grid for wavelength ℓ , where the subsampling ratio $0 < \alpha \leq 1$ corresponds to the proportion of spatial locations sampled. Note that the locations of the sampled pixels are randomized for each wavelength ℓ . The data fidelity term in this case can be defined as the negative log-likelihood as follows:

$$\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{B}) = \sum_{\ell=1}^L \sum_{p \in S_{\alpha,\ell}} [-\log(b_{p,\ell})\sigma_{p,\ell} + T_b b_{p,\ell}]. \quad (4.2)$$

Next, it is necessary to determine what regularization will be used for building the minimization problem. We propose here a regularization term that is the sum of the total variation in each of the columns of \mathbf{B} , corresponding to the sum of the total variations of each image \mathbf{b}_ℓ . For the sake of simplicity, this regularization will be called spectral total variation, or STV. The Moreau proximity operator for the STV regularization function is equivalent to the Moreau proximity operator for the TV regularization at each wavelength, which in turn must be calculated iteratively.

$$\text{STV}(\mathbf{B}) = \sum_{\ell=1}^L \text{TV}(\mathbf{b}_\ell) \quad (4.3)$$

Additionally, a technique used in compressive sensing settings for multi-spectral radio astronomy aiming to correlate different bands during signal reconstruction consists of using the nuclear norm, i.e. the sum of absolute singular values, as additional regularization in the minimization problem [24, 42]. The nuclear norm is a convex regularization that promotes low-rank matrices, which correlates the columns of \mathbf{B} and helps the decoding of a given wavelength by using information present in the other wavelengths. The nuclear norm of a matrix \mathbf{X} is here represented by $\|\mathbf{X}\|_* = \|\Sigma\|_{1,1} = \sum_i |\sigma_i|$ where $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$ is the SVD decomposition of \mathbf{X} . The Moreau proximity operator for the nuclear norm is a soft thresholding operation on the singular values of the matrix. Therefore, this regularization requires the computation of an SVD decomposition at each iteration, which may be costly depending on the dimensions of the considered data and the computing power available.

Putting all these elements together, the minimization problem can then be written as:

$$\hat{\mathbf{B}} = \arg \min_{\mathbf{B}} [\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{B}) + \tau_1 \text{STV}(\mathbf{B}) + \tau_2 \|\mathbf{B}\|_* + i_{\mathbb{R}^+}(\mathbf{B})] \quad (4.4)$$

Note that since the problem now includes two hyperparameters, the method mentioned in Section 3.3.1 for automatic parameter tuning can no longer be used, and thus τ_1 and τ_2 must be tuned manually.

Let us define $\boldsymbol{\sigma}_f = [\sigma_{f,1} \dots \sigma_{f,L}]$, where $\sigma_{f,\ell}$ is the sum of expected number of observed photons for wavelength ℓ , i.e. $\sigma_{f,\ell} = \sum_{t=1}^{T_a} f_{d,\ell}(t)$, which depends only on ℓ and not on d as long as the impulse responses are not truncated, which is assumed. Following the same method described in [18] I proposed an algorithm able to solve (4.4). The resulting algorithm is Algorithm 1 described below.

4.3.2 Response Intensity

Once the baseline estimation $\hat{\mathbf{B}}$ has been calculated, the multi-spectral reflectivity estimation can be performed. Using the complete measurement model described by Equation (4.1) we can once again calculate the data fidelity term as a function of $\mathbf{A} \in \mathbb{R}^{P \times L}$ using the negative log-likelihood of our measurement model. The data fidelity term is then given by

$$\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{A}|\hat{\mathbf{B}}) = \sum_{\ell=1}^L \sum_{p \in S_{\alpha,\ell}} [r_{p,\ell} \sigma_{f,\ell} a_{p,\ell} + T_a b_{p,\ell} - \sigma_{p,\ell} \log(r_{p,\ell} \sigma_{f,\ell} a_{p,\ell} + T_a b_{p,\ell})] \quad (4.5)$$

where T_a is the total number of time-steps in the considered photon count histogram, S_{α} is the set of measured pixels, and $\sigma_{p,\ell} = \sum_{t=1}^{T_a} y_{p,t,\ell}$.

Similarly to the case of multi-spectral baseline estimation, the STV and nuclear norm regularization terms are used, along with the positivity enforcing indicator function, to build the minimization problem that yields the estimate $\hat{\mathbf{A}}$, defined as

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} [\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{A}|\hat{\mathbf{B}}) + \tau_1 \text{STV}(\mathbf{A}) + \tau_2 \|\mathbf{A}\|_* + i_{\mathbb{R}^+}(\mathbf{A})], \quad (4.6)$$

where \mathbf{B} has been replaced by $\hat{\mathbf{B}}$.

To solve this minimization problem using the same strategy that was used in the previous chapter it is necessary to calculate the Moreau proximity operator of the proposed data fidelity

Input: $\mathbf{M}_\alpha, \mu, \tau_1, \tau_2, \Sigma, \sigma_f$
Result: $\hat{\mathbf{B}}$
 Initialise $\mathbf{u}_0^{(1)}, \mathbf{u}_0^{(2)}, \mathbf{u}_0^{(3)}, \mathbf{u}_0^{(4)}, \mathbf{d}_0^{(1)}, \mathbf{d}_0^{(2)}, \mathbf{d}_0^{(3)}, \mathbf{d}_0^{(4)}$
 Precompute $(\mathbf{M}_\alpha^T \mathbf{M}_\alpha + 3\mathbf{I})^{-1}$
 $k \leftarrow 0$
while $\|\mathbf{z}_k - \mathbf{z}_{k-1}\|_2 / \|\mathbf{z}_k\|_2 > \epsilon$ *AND* $k \leq k_{max_iter}$ **do**
 $\zeta_k^{(1)} \leftarrow \mathbf{u}_k^{(1)} + \mathbf{d}_k^{(1)}$
 $\zeta_k^{(2)} \leftarrow \mathbf{u}_k^{(2)} + \mathbf{d}_k^{(2)}$
 $\zeta_k^{(3)} \leftarrow \mathbf{u}_k^{(3)} + \mathbf{d}_k^{(3)}$
 $\zeta_k^{(4)} \leftarrow \mathbf{u}_k^{(4)} + \mathbf{d}_k^{(4)}$
 $\gamma_k \leftarrow \mathbf{M}_\alpha^T \zeta_k^{(1)} + \zeta_k^{(2)} + \zeta_k^{(3)} + \zeta_k^{(4)}$
 $\mathbf{z}_{k+1} \leftarrow (\mathbf{M}_\alpha^T \mathbf{M}_\alpha + 3\mathbf{I})^{-1} \gamma_k$
 $\nu_k^{(1)} \leftarrow \mathbf{M}_\alpha \mathbf{z}_{k+1} - \mathbf{d}_k^{(1)}$
 $\mathbf{u}_{k+1}^{(1)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(1)}\|_2^2 + \mathcal{L}_{Y,\alpha}(\mathbf{v})$
 $\nu_k^{(2)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(2)}$
 $\mathbf{u}_{k+1}^{(2)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(2)}\|_2^2 + \tau_1 \text{STV}(\mathbf{v})$
 $\nu_k^{(3)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(3)}$
 $\mathbf{u}_{k+1}^{(3)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(3)}\|_2^2 + \tau_2 \|\mathbf{v}\|_*$
 $\nu_k^{(4)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(4)}$
 $\mathbf{u}_{k+1}^{(2)} \leftarrow \max(0, \nu_k^{(4)})$
 $\mathbf{d}_{k+1}^{(1)} \leftarrow \mathbf{d}_k^{(1)} - (\mathbf{M}_\alpha \mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(1)})$
 $\mathbf{d}_{k+1}^{(2)} \leftarrow \mathbf{d}_k^{(2)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(2)})$
 $\mathbf{d}_{k+1}^{(3)} \leftarrow \mathbf{d}_k^{(3)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(3)})$
 $\mathbf{d}_{k+1}^{(4)} \leftarrow \mathbf{d}_k^{(4)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(4)})$
 $k \leftarrow k + 1$
end
 $\hat{\mathbf{B}} \leftarrow \mathbf{z}_k$

Algorithm 1: Adapted PIDAL algorithm to estimate $\hat{\mathbf{B}}$

term. Following the definition of the Moreau proximity operator

$$\begin{aligned} \text{prox}_{\mathcal{L}_{\mathbf{Y},\alpha}/\mu}(\mathbf{N}) &= \arg \min_{\mathbf{A}} \left[\mathcal{L}_{\mathbf{Y},\alpha}(\mathbf{A}|\hat{\mathbf{B}}) + \frac{\mu}{2} \|\mathbf{A} - \mathbf{N}\|_{2,2}^2 \right] = \\ &= \frac{1}{2} \left(\mathbf{N} - \frac{\sigma_f \mathbf{R}}{\mu} - \frac{T_a \hat{\mathbf{B}}}{\sigma_f \mathbf{R}} + \sqrt{\left(\mathbf{N} - \frac{\sigma_f \mathbf{R}}{\mu} - \frac{T_a \hat{\mathbf{B}}}{\sigma_f \mathbf{R}} \right)^2 + \frac{4\mathbf{\Sigma}}{\mu} + \frac{4T_a \hat{\mathbf{B}} \odot \mathbf{N}}{\sigma_f \mathbf{R}} - \frac{4T_a \hat{\mathbf{B}}}{\mu}} \right) \end{aligned} \quad (4.7)$$

where \odot represents the element-wise product and \div represents element-wise division.

Following once again the method described in [18] I proposed an algorithm able to solve (4.6). The resulting algorithm is Algorithm 2 described below. Since the SVD calculations used for calculating the Moreau proximity operator of the nuclear norm can be very costly, one strategy to speed up the convergence of the algorithm is to first run the algorithm as if the nuclear norm was not included in the minimization problem, which will result in reconstruction using only the STV regularization. Upon convergence, the solution can be used as initial guess for the algorithm, which is then run in its complete form. This can be seen as using a simpler algorithm to approach the solution and provide a smart initial guess for the full algorithm.

4.3.3 Depth Estimation

The final step of the decoding process is once again the depth estimation. The log-likelihood for the multi-spectral model in use here with respect to the depth d is given by Eq. (4.8). Note that unlike the baseline and the reflectivity estimation, the depth estimation does not scale with the number of wavelengths considered in the problem. Once again the calculations can be sped up by reducing the summation over t only to the terms where $y_{p,t,\ell} \neq 0$. This leads to much faster likelihood calculations when few photons are observed since $y_{p,t,\ell}$ becomes very sparse. This is ideal since that is exactly the case that this work focuses on. The joint log-likelihood then becomes

$$\log(P(\mathbf{Y}|\hat{\mathbf{B}}, \hat{\mathbf{A}}, \mathbf{d})) = \sum_{\ell=1}^L \sum_{p \in S_{\alpha,\ell}} \sum_{t, y_{p,t,\ell} \neq 0} [y_{p,t,\ell} \log(\hat{a}_{p,\ell} f_d(t) + \hat{b}_{p,\ell})] + C. \quad (4.8)$$

We can then define a sampling matrix \mathbf{M}_{α} as a selection matrix for all the pixels that have been observed using at least one wavelength. We also define the weights matrix \mathbf{W} as a function of the number of photons that have been observed at pixel p on all wavelengths.

$$\mathbf{W} = \text{diag}(\log(1 + \boldsymbol{\sigma})), \text{ where } \boldsymbol{\sigma} = \sum_{\ell=1}^L \sum_t^{T_a} y_{p,t,\ell} \quad (4.9)$$

The minimization problem can then be built similarly to the single-spectral case using a

Input: $\mathbf{M}_\alpha, \mu, \tau_1, \tau_2, \Sigma, \sigma_f, \hat{\mathbf{B}}$
Result: $\hat{\mathbf{A}}$
 Initialise $\mathbf{u}_0^{(1)}, \mathbf{u}_0^{(2)}, \mathbf{u}_0^{(3)}, \mathbf{u}_0^{(4)}, \mathbf{d}_0^{(1)}, \mathbf{d}_0^{(2)}, \mathbf{d}_0^{(3)}, \mathbf{d}_0^{(4)}$
 Precompute $(\mathbf{M}_\alpha^T \mathbf{M}_\alpha + 3\mathbf{I})^{-1}$
 $k \leftarrow 0$
while $\|\mathbf{z}_k - \mathbf{z}_{k-1}\|_2 / \|\mathbf{z}_k\|_2 > \epsilon$ *AND* $k \leq k_{max_iter}$ **do**
 $\zeta_k^{(1)} \leftarrow \mathbf{u}_k^{(1)} + \mathbf{d}_k^{(1)}$
 $\zeta_k^{(2)} \leftarrow \mathbf{u}_k^{(2)} + \mathbf{d}_k^{(2)}$
 $\zeta_k^{(3)} \leftarrow \mathbf{u}_k^{(3)} + \mathbf{d}_k^{(3)}$
 $\zeta_k^{(4)} \leftarrow \mathbf{u}_k^{(4)} + \mathbf{d}_k^{(4)}$
 $\gamma_k \leftarrow \mathbf{M}_\alpha^T \zeta_k^{(1)} + \zeta_k^{(2)} + \zeta_k^{(3)} + \zeta_k^{(4)}$
 $\mathbf{z}_{k+1} \leftarrow (\mathbf{M}_\alpha^T \mathbf{M}_\alpha + 3\mathbf{I})^{-1} \gamma_k$
 $\nu_k^{(1)} \leftarrow \mathbf{M}_\alpha \mathbf{z}_{k+1} - \mathbf{d}_k^{(1)}$
 $\mathbf{u}_{k+1}^{(1)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(1)}\|_2^2 + \mathcal{L}_{Y,\alpha}(\mathbf{v} | \hat{\mathbf{B}})$
 $\nu_k^{(2)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(2)}$
 $\mathbf{u}_{k+1}^{(2)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(2)}\|_2^2 + \tau_1 \text{STV}(\mathbf{v})$
 $\nu_k^{(3)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(3)}$
 $\mathbf{u}_{k+1}^{(3)} \leftarrow \arg \min_v \frac{\mu}{2} \|\mathbf{v} - \nu_k^{(3)}\|_2^2 + \tau_2 \|\mathbf{v}\|_*$
 $\nu_k^{(4)} \leftarrow \mathbf{z}_{k+1} - \mathbf{d}_k^{(4)}$
 $\mathbf{u}_{k+1}^{(2)} \leftarrow \max(0, \nu_k^{(4)})$
 $\mathbf{d}_{k+1}^{(1)} \leftarrow \mathbf{d}_k^{(1)} - (\mathbf{M}_\alpha \mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(1)})$
 $\mathbf{d}_{k+1}^{(2)} \leftarrow \mathbf{d}_k^{(2)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(2)})$
 $\mathbf{d}_{k+1}^{(3)} \leftarrow \mathbf{d}_k^{(3)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(3)})$
 $\mathbf{d}_{k+1}^{(4)} \leftarrow \mathbf{d}_k^{(4)} - (\mathbf{z}_{k+1} - \mathbf{u}_{k+1}^{(4)})$
 $k \leftarrow k + 1$
end
 $\hat{\mathbf{A}} \leftarrow \mathbf{z}_k$

Algorithm 2: Adapted PIDAL algorithm to estimate $\hat{\mathbf{A}}$

weighted Euclidean distance as data fidelity term and a convex regularization term $p_d \mathbf{d}$. Once this minimization problem is solved the decoding of the single-photon multi-spectral LIDAR signal is complete.

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d}} [(\mathbf{d}_{ML} - \mathbf{M}_\alpha \mathbf{d})^T \mathbf{W} (\mathbf{d}_{ML} - \mathbf{M}_\alpha \mathbf{d}) + \tau TV(\mathbf{d}) + i_{\mathbb{R}^+}(\mathbf{d})] \quad (4.10)$$

4.4 Comparison Against State-of-the-art

The decoding method proposed in the previous sections can be seen as an improvement on the method proposed in [13, 14]. The main differences are:

1. The method proposed here has been extended to decode single-photon multi-spectral LIDAR signals, as opposed to only single-spectral ones.
2. The baseline model has been extended to estimate a baseline value for each pixel, which will be shown to be important in Chapter 5.
3. The algorithm presented here infers the baseline intensity from the measured data instead of doing so during calibration, which enables mobile LIDAR systems.
4. The present method uses directly calibration data for depth estimation, instead of proposing approximations in the shape of the impulse response.
5. The weighting matrix in the data fidelity term based on the photon count that has been proposed in this work has not been proposed by any other method in the state-of-the-art.

The proposed method also differs from the ones proposed in [15, 16] mainly in the fact that it is able to decode the signal without any previous knowledge of the spectral responses of the materials in the scene. The present method also uses the nuclear norm to correlate the different bands during the decoding process to improve photon efficiency by exploiting the redundancy between the baseline and reflectivity images in different bands.

Chapter 5

Results

5.1 Single-spectral Results

A set of test images of size 200×200 pixels was generated from an RGB image and its depth map with different amounts of measured photons. A scaled version of the R channel was used as response intensity a , a scaled version of channel B was used as baseline response b , and a scaled version of the depth map was used as the depth ground truth d to simulate a realistic scene where baseline, reflectivity and depth are correlated. The expectation of the observed number of photons at each pixel and within each time bin was calculated and a random sample was drawn from a Poisson distribution accordingly. Two test images were used: one (office) is a real photograph and its estimated depth map [43], and the other (cube) is a 3D rendering of a simulated scene. Total variation regularization was used for the optimization problems given its proven success in the recovery and denoising of natural images.

The signals were simulated with several different numbers of photons per pixel and sub-sampling ratios, as can be seen in Table 5.1. The values in the table are for response intensity photons, and an equal amount of baseline photons was present in each simulation, thus only half the photons contained depth information. The values were chosen to maintain constant average per-pixel photon counts for each group of scenes. This was done to address the question proposed earlier regarding generalization to multispectral LIDAR.

The results obtained for each of the test images in each of the two sets can be seen in Figs. 5.1 and 5.2. Each of the lines refers to the cases in which the average photon count considering all the pixels in the image is constant. For these test cases, the measurements were divided in baseline and response intensity halves using $T_b = 1200$ and $T_a = 2500$. The numerical values for the SNR of each result can be found in Tables 5.2 and 5.3.

Table 5.1: Approximate photon per pixel count and subsampling ratios

ppp (sampled pixels)	α	ppp (average)
0.5, 1, 2, 4, 8	1, 1/2, 1/4, 1/8, 1/16	0.5
1, 2, 4, 8, 16	1, 1/2, 1/4, 1/8, 1/16	1
10, 20, 40, 80, 160	1, 1/2, 1/4, 1/8, 1/16	10
100, 200, 400, 800, 1600	1, 1/2, 1/4, 1/8, 1/16	100
1000, 2000, 4000, 8000, 16000	1, 1/2, 1/4, 1/8, 1/16	1000

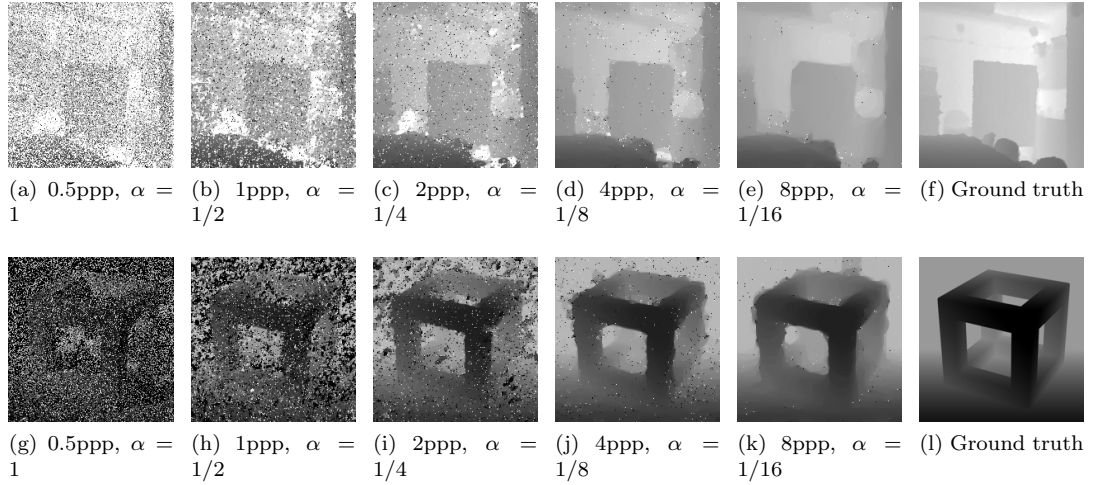
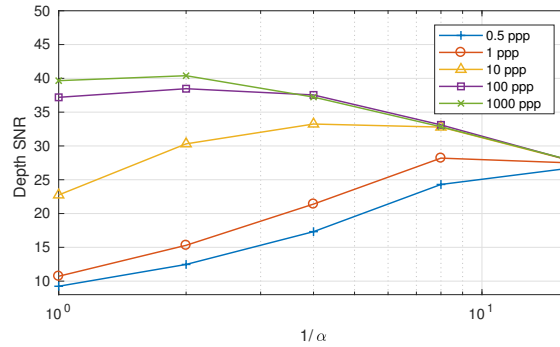


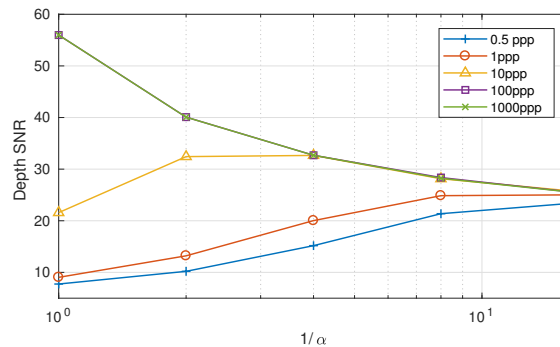
Figure 5.1: Final depth estimation results obtained by our method and the ground truth depth for both images with 0.5 ppp on average for all cases.

The results displayed in Figs. 5.1 and 5.2 were obtained by manually tuning the optimization parameters instead of using the method described in [33] to achieve faster optimizations. The SNRs obtained by manually tuning the optimization parameters were also better than the ones obtained by automatic parameter tuning. The method proposed in [33] is most useful when no previous information is known about the measured images, which was not the case in our tests.

Observing the 0.5 ppp and 1 ppp lines on plots 5.2a and 5.2b we can see that when the acquisition time is very low, and therefore the photon count is low, it is preferable to acquire more data on a random selection of pixels than to acquire a uniform amount of data for all pixels. This is not the case when a large amount of photons is observed, in which case it seems better to sample all the pixels. This can also be clearly seen in the images displayed in Fig. 5.1. These images show the estimated depths for the cases where the averagenumber of photons per pixel over all pixels is 0.5 ppp.



(a) Results for office image



(b) Results for cube image

Figure 5.2: Accuracy of the results considering different photon counts and subsampling ratios for both test images.

We can clearly see that the quality of the results improves as we sample less pixels with more accuracy. We also observe that the lines for 100 ppp and 1000 ppp are very close for both test images, which indicates that for large amounts of photons the quality of the depth estimation is no longer strongly affected by the amount of measured photons. These conclusions are corroborated by the locations of the maximum of each line in Tables 5.2 and 5.3, marked in bold. Finally, we can also conclude that since the lines converge when the subsampling ratio grows, there is a limit to the maximum subsampling ratio in order to enable an adequate depth estimation.

While these tests show the performance of the algorithm under different conditions, it is important to note that they were executed using data generated using the exact measurement model that was proposed. It is important then to verify the accuracy of the proposed measurement model by decoding real single-photon LIDAR data.

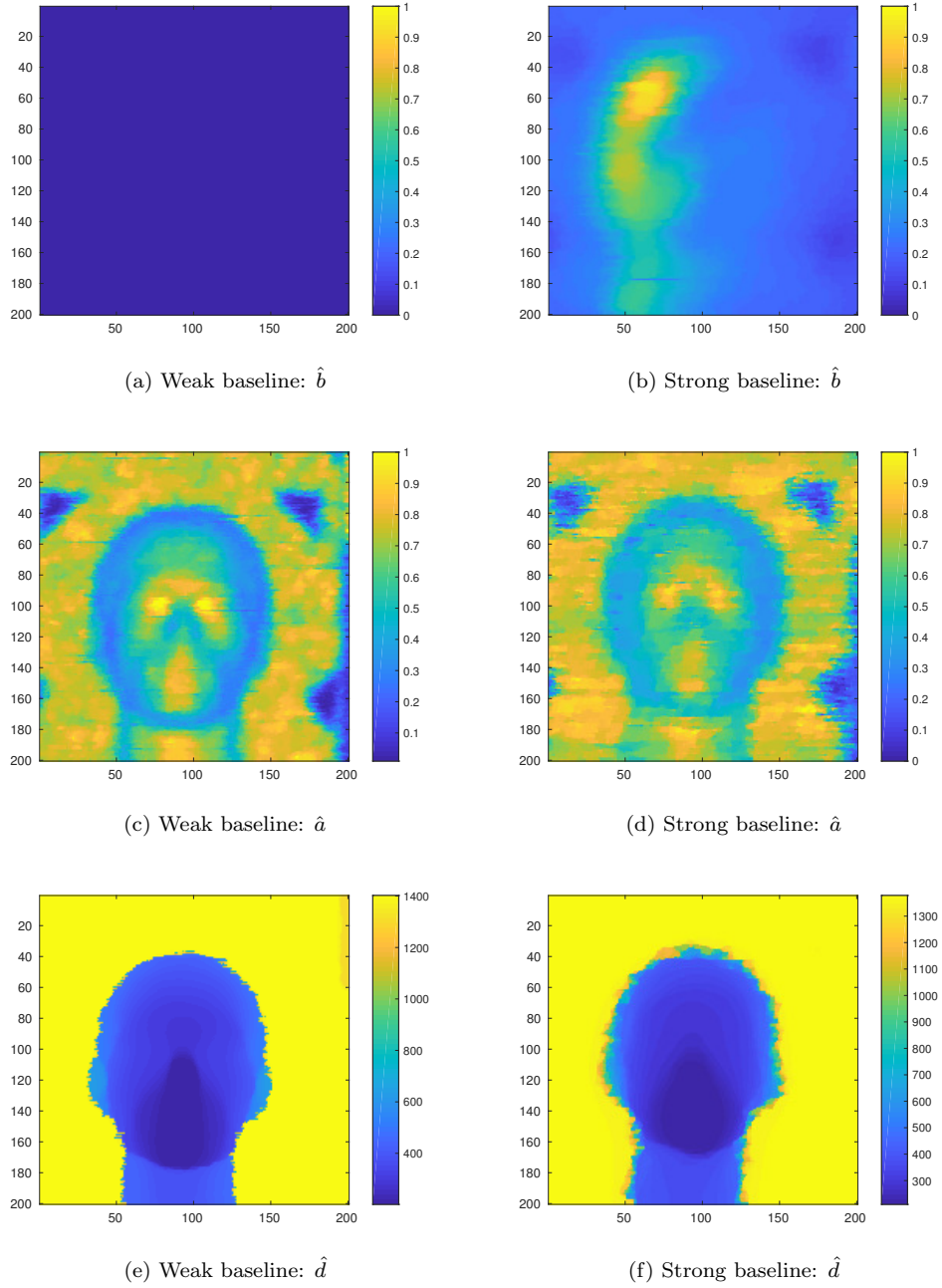


Figure 5.3: Baseline, reflectivity and depth results obtained from applying the proposed algorithm on real LIDAR data. The images on the left are from an acquisition with very weak baseline signal, while the one on the right was under direct sunlight, and therefore had a strong baseline signal.

Table 5.2: SNRs for the depth estimation results using the office image

Mean ppp	$\alpha = 1$	$\alpha = 1/2$	$\alpha = 1/4$	$\alpha = 1/8$	$\alpha = 1/16$
0.5	9.2417	12.468	17.3194	24.3011	26.6858
1	10.7236	15.2998	21.3991	28.1975	27.5057
10	22.7464	30.3103	33.241	32.79	27.9123
100	37.1936	38.4678	37.5441	33.1134	27.9258
1000	39.658	40.3907	37.2302	32.84	27.9777

Table 5.3: SNRs for the depth estimation results using the cube image

Mean ppp	$\alpha = 1$	$\alpha = 1/2$	$\alpha = 1/4$	$\alpha = 1/8$	$\alpha = 1/16$
0.5	7.7425	10.2014	15.1753	21.3681	23.3367
1	9.0366	13.2075	20.0279	24.8601	25.015
10	21.5767	32.4258	32.659	28.1395	25.8196
100	55.9804	40.0808	32.6872	28.3709	25.6323
1000	55.9827	40.069	32.6852	28.2385	25.6

Figure 5.3 shows the results from using this decoding method on two different acquisitions of the setup displayed in Fig. 5.4. The triangles that can be seen in Figs. 5.3c, 5.3d and 5.4 are markers used for scale calibration. The first one was acquired when ambient illumination was very weak, and therefore the majority of the observed photons were reflections from the emitted laser pulse. The second one was acquired with the imaged object under direct sunlight, and therefore there is a strong baseline signal present. This being the case, there is no way of calculating the SNR for these results since there is no ground truth available for comparison. Nevertheless, several conclusions can be drawn from the qualitative analysis of these results.

The first thing to note is that the algorithm has been able to infer the baseline intensity, reflectivity and depth for all the pixels in the image. This verifies the assumptions that were made, the measurement model, and the proposed decoding algorithm.

Another important observation is that, not surprisingly, the presence of a strong baseline signal reduces the quality of the reconstruction. Ideally all measured photons would come from the emitted laser pulse, and the baseline signal is, in this case, noise that complicates the extraction of information from the measurements.

Finally, here it can clearly be seen in Fig. 5.3b that the assumption made in [13, 14] that the baseline intensity is equal for all pixels in the image is incorrect. In this case, the baseline signal is significantly stronger on the parts of the image where the object was under direct sunlight illumination. This therefore justifies the more precise pixel-wise modelling of baseline intensity



Figure 5.4: Image of the setup used for acquiring single-spectral LIDAR images. The imaged object provides a realistic shape for evaluating the decoding algorithm.

that was proposed in this chapter.

5.2 Multi-spectral Results

To test the multi-spectral extension of the decoding method proposed in Chapter 4 real and simulated LIDAR data was used. To begin with, the scene shown in Fig. 5.5 was used. This decoding shows that the decoding method is able to deal with a variety of materials with different spectral responses. It also shows that the proposed measurement model is accurate, including the wavelength dependent vignetting effect. This scene was imaged using 190×190 pixels over 33 laser wavelengths, from 500nm to 820nm in steps of 10nm.

Some results from the decoding of this scene can be seen in Fig. 5.6, which contains the estimated reflectivity of the scene for five of the measured wavelengths. Figure 5.6f shows the depth estimation results for this acquisition. Although there is no ground truth for this data

since it is a real acquisition, the qualitative analysis of the results indicate that the proposed method succeeded in accurately decoding the LIDAR signal.



Figure 5.5: Image of the setup used for acquiring multi-spectral LIDAR images. The imaged object provides realistic shape and varied spectral responses for evaluating the decoding algorithm.

To quantitatively evaluate the proposed method, a set of twenty scenes were simulated using as ground truth the reflectivity and depth estimations from the decoding of the real scene to ensure the simulations were as realistic as possible. These scenes also contained 190×190 pixels ($P = 36100$) and thirty three wavelengths ($L = 33$). The impulse responses were assumed to come from a single imaged surface at each pixel with the depth and reflectivity as given by the ground truth signals. The acquisition times were scaled so that the expected mean photon counts per band in each simulated signal assumed the values described in Table 5.4 to cover a range of situations and evaluate once again the effects of compressive sampling of the pixels.

The results from these simulation can be seen in Tables 5.5 and 5.6, as well as in the plots in Figure 5.7. It is interesting to note from these plot that, unlike in the single-spectral case, observing fewer pixels for longer periods of time does not seem to be a good strategy in the

Table 5.4: Approximate photon per pixel per band count and subsampling ratios for multi-spectral simulations

ppp (sampled pixels)	α	ppp (average)
0.5, 1, 2, 4, 8	1, 1/2, 1/4, 1/8, 1/16	0.5
1, 2, 4, 8, 16	1, 1/2, 1/4, 1/8, 1/16	1
10, 20, 40, 80, 160	1, 1/2, 1/4, 1/8, 1/16	10
100, 200, 400, 800, 1600	1, 1/2, 1/4, 1/8, 1/16	100

Table 5.5: SNRs for the reflectivity estimation results

Mean ppp	$\alpha = 1$	$\alpha = 1/2$	$\alpha = 1/4$	$\alpha = 1/8$	$\alpha = 1/16$
0.5	14.1998	14.1922	14.1642	13.9935	12.9723
1	16.5688	16.7516	16.631	16.1428	14.7065
10	20.1400	21.8927	21.6714	19.46	16.3172
100	25.0042	25.8788	23.5377	20.1337	16.5169

Table 5.6: SNRs for the depth estimation results

Mean ppp	$\alpha = 1$	$\alpha = 1/2$	$\alpha = 1/4$	$\alpha = 1/8$	$\alpha = 1/16$
0.5	33.4773	32.416	30.0932	27.6252	27.2448
1	47.3494	39.0025	37.5994	32.9088	30.5927
10	62.0842	62.1623	40.9781	41.7838	38.7348
100	67.6337	67.5869	45.8926	45.4588	41.0779

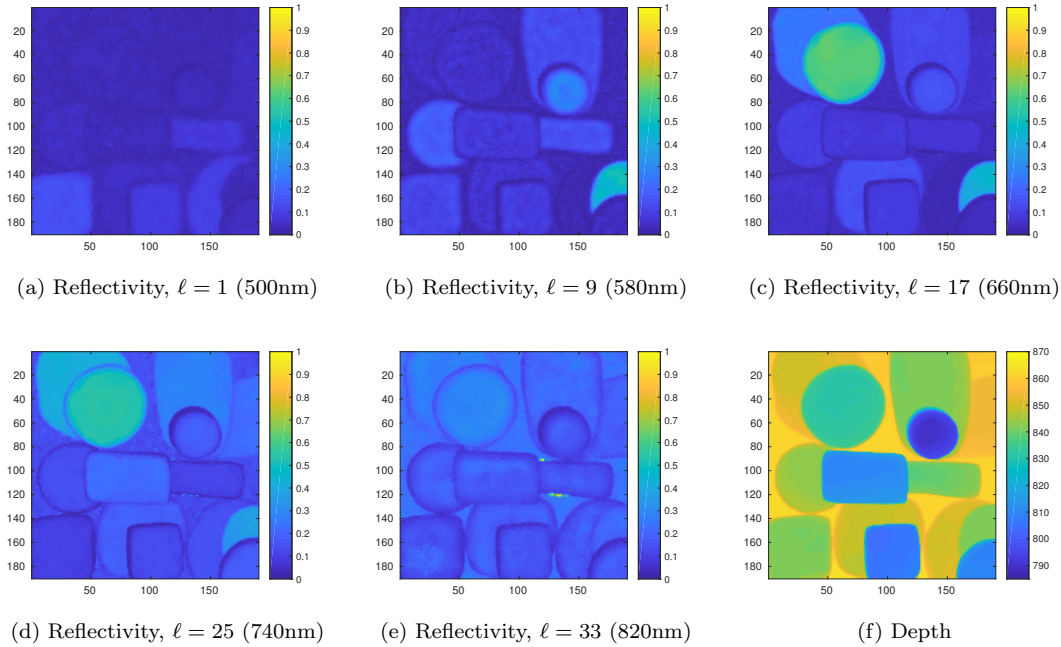


Figure 5.6: Decoding results for real single-photon multi-spectral LIDAR data.

multi-spectral case for either reflectivity or depth estimation, except for a few cases where a subsampling rate of $\alpha = 1/2$ yields the best reflectivity results.

The resulting depth estimations for each of the test cases can be seen in Fig. 5.8. As expected, we can observe that more photons yield better depth estimations. Also, in accordance to the SNRs previously mentioned, these results indicate that lower values of α lead to worse depth estimations. It is also interesting to note that the errors in these depth estimations are very localised, and these errors seem to increase in number and not in magnitude as the results degrade. This is likely to occur due to noise photons that have been observed far from the peak of the impulse response. In a case where very few photons are measured, a single noise photon has a strong weight and is able to disturb the accuracy of the depth estimation. These results could be improved with denoising techniques, but the scope of this work is to obtain the best decoding results using state-of-the-art LIDAR modelling and convex optimization methods, and the comparison of denoising techniques is a well studied area.

Another way to evaluate the results that may give more information than simply calculating the SNRs for each case is to analyse the cumulative probability function of the absolute error in each reflectivity and depth estimation. This allows us to see the distribution of the error in the performed estimations. The cumulative probability functions from the absolute errors

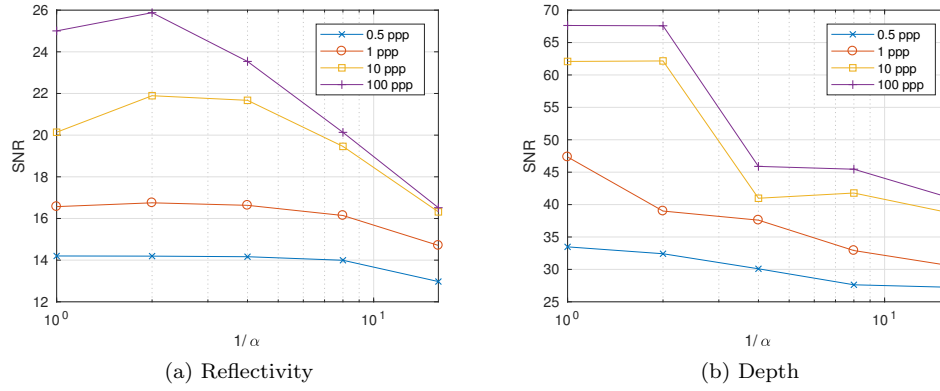


Figure 5.7: SNRs in reflectivity and depth estimation for each of the test cases.

for reflectivity estimation can be seen in Fig.5.9, and for depth estimation in Fig. 5.10. The cumulative error distribution for the reflectivity estimation follows what is expected, which is a smooth curve that approaches a unitary step function as the results improve. On the other hand, the cumulative error distribution for depth estimation shows strange behaviours in the curves, which indicates that few large errors seem to be having a strong impact in the SNR calculations.

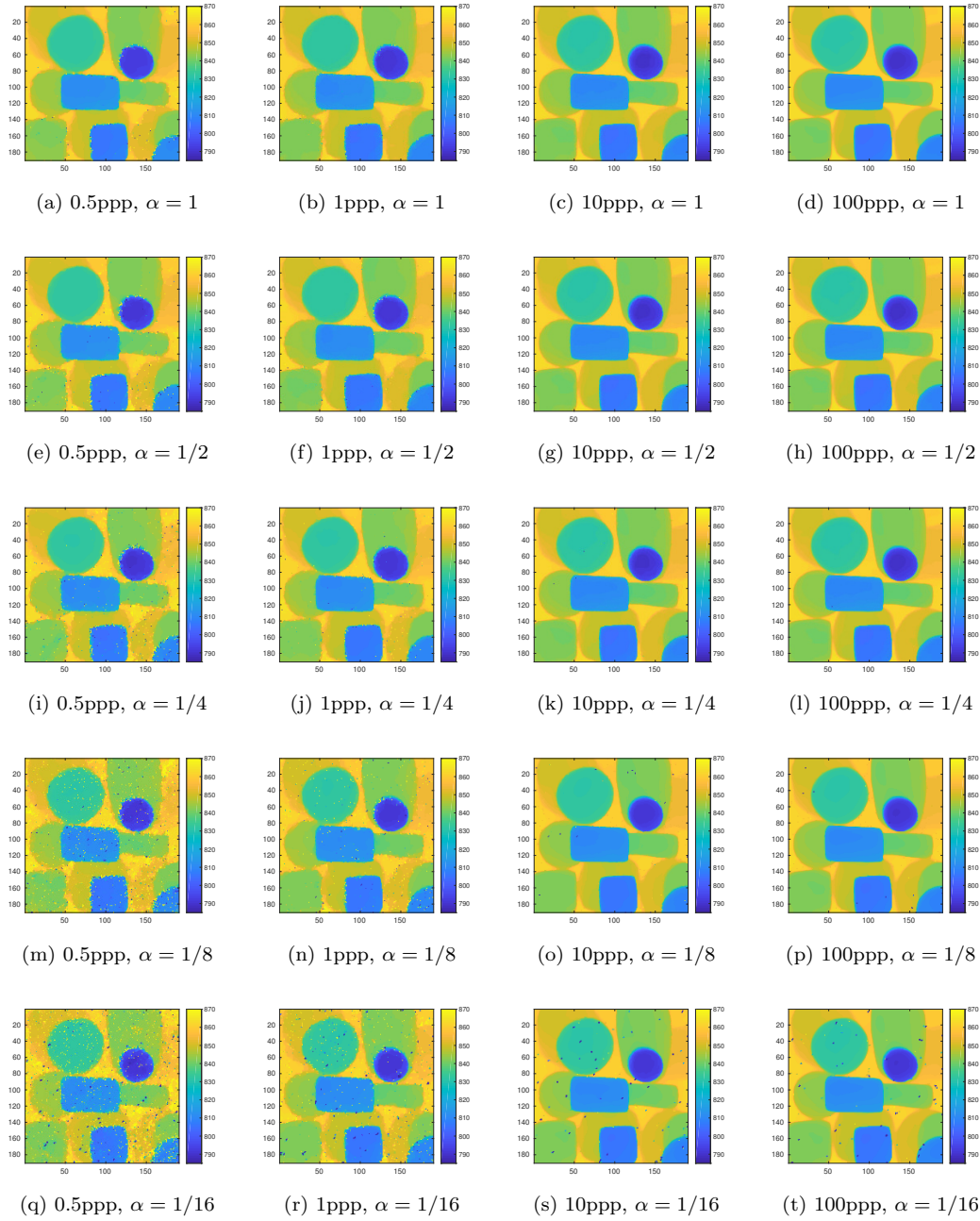
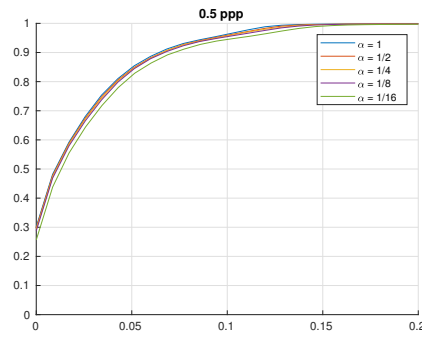
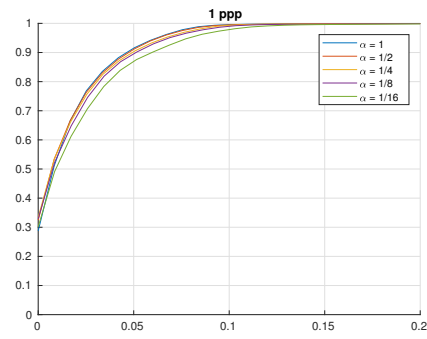


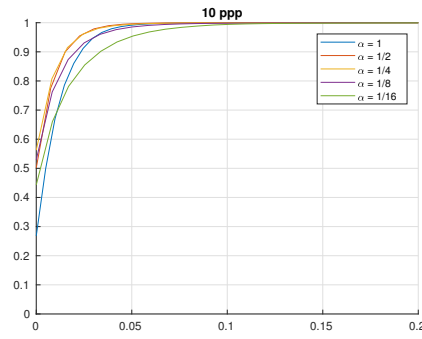
Figure 5.8: Depth estimation for each test case.



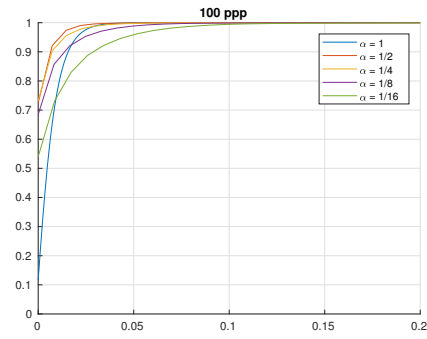
(a) 0.5 photons per pixel



(b) 1 photons per pixel

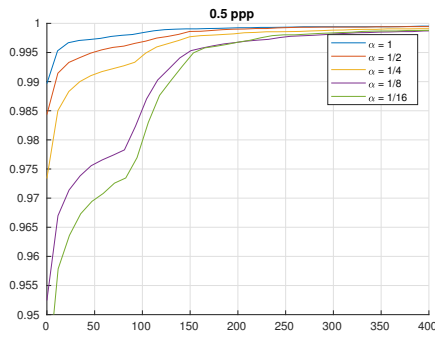


(c) 10 photons per pixel

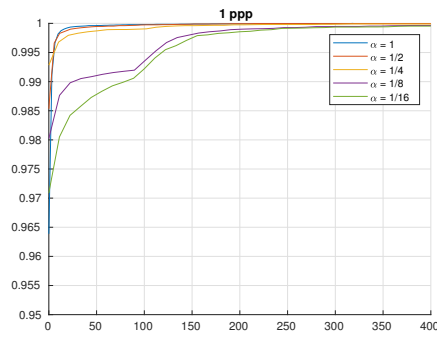


(d) 100 photons per pixel

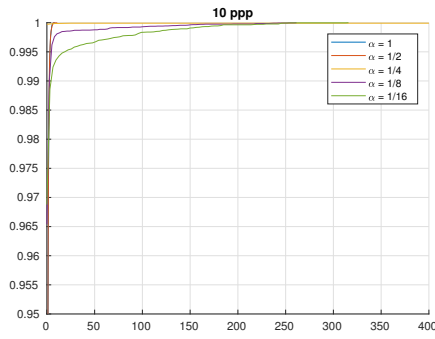
Figure 5.9: Cumulative probability function for the absolute error in reflectivity estimation for each of the test cases.



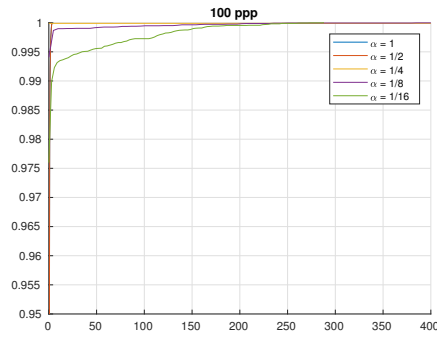
(a) 0.5 photons per pixel



(b) 1 photons per pixel



(c) 10 photons per pixel



(d) 100 photons per pixel

Figure 5.10: Cumulative probability function for the absolute error in depth estimation for each of the test cases.

Chapter 6

Conclusion

6.1 Main Considerations

In the previous chapters a decoding method for single-spectral and multi-spectral single-photon LIDAR signals was proposed based on the state-of-the-art single-photon LIDAR decoding algorithms, and using compressive sensing and convex optimization ideas. In both cases, the method was proven to be able to successfully decode real LIDAR signals, validating the proposed measurement models and decoding methods.

The usage of compressive sensing to mitigate the effects of the lack of photons in the LIDAR signals was tested. The results suggest that in the single-spectral case it is interesting to sample fewer pixels for longer periods of time using a scanning LIDAR system in order to acquire higher quality data in those pixels, and infer the rest of the values using compressive sensing methods. In the multi-spectral case, the results indicate that this strategy is not advisable as it seems to reduce the decoding quality, and it is better to use a uniform scanning of the pixels in the image.

Nevertheless, the proposed decoding method improves on the state-of-the-art in different ways. It provides a more complete modelling of the baseline noise, it enables baseline estimation from the measurement itself instead of during calibration, it uses directly the impulse response shape from calibration to calculate the depth likelihoods at each pixel. Furthermore, the nuclear norm was added in the multi-spectral case to correlate the different bands, and a novel weighting of the data fidelity term based on the number of observed photons was used for depth estimation.

6.2 Future Work

The choice of parameters for the proposed optimization problems remains a challenge, especially for the multispectral cases where τ_1 and τ_2 must be tuned simultaneously. It would be interesting to explore techniques for tuning these optimization parameters, perhaps based on the physics of the observation model.

It would be interesting to explore the usage of other regularizations that correlate the different wavelengths during the decoding process. While the nuclear norm regularization is useful, it is expensive to calculate and other priors such as joint sparsity models could be explored in this context.

It is also important to evaluate the reflectivity results from the multi-spectral decoding algorithm with existing spectral unmixing methods for material classification to ensure that the method is precise enough for these applications in low photon count settings.

Another possible improvement to the contributions in this work would be to find a way to move the minimization problem to a constrained formulation. This is challenging since the data fidelity term here uses the Poissonian negative log-likelihood instead of the Euclidean distance, but finding a way of constraining the solution space could lead to improvements in the quality of the results.

Finally, it would be interesting to explore the usage of other convex optimization algorithms that converge faster, demand less computational power or are able to run in a distributed manner. Such algorithms already exist, such as the Primal-Dual family of optimization algorithms [44].

Bibliography

- [1] Brent Schwarz. Lidar: Mapping the world in 3d. *Nature Photonics*, 4(7):429, 2010.
- [2] Guoqing Zhou, C Song, John Simmers, and Penggen Cheng. Urban 3d gis from lidar and digital aerial images. *Computers & Geosciences*, 30(4):345–353, 2004.
- [3] Keqi Zhang, Jianhua Yan, and S-C Chen. Automatic construction of building footprints from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2523–2533, 2006.
- [4] Aurora Maccarone, Aongus McCarthy, Ximing Ren, Ryan E Warburton, Andy M Wallace, James Moffat, Yvan Petillot, and Gerald S Buller. Underwater depth imaging using time-correlated single-photon counting. *Optics express*, 23(26):33911–33926, 2015.
- [5] Nail Cadalli, Peter J Shargo, David C Munson Jr, and Andrew C Singer. Three-dimensional tomographic imaging of ocean mines from real and simulated lidar returns. In *International Symposium on Optical Science and Technology*, pages 155–166. International Society for Optics and Photonics, 2002.
- [6] Michael A Lefsky, Warren B Cohen, Geoffrey G Parker, and David J Harding. Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience*, 52(1):19–30, 2002.
- [7] Michael A Wulder, Christopher W Bater, Nicholas C Coops, Thomas Hilker, and Joanne C White. The role of lidar in sustainable forest management. *The Forestry Chronicle*, 84(6):807–826, 2008.
- [8] Daniel Sergio Martinez Ramirez, Gerald Stuart Buller, Aongus McCarthy, Simone Morak, Caroline J. Nichol, Ximing Ren, Andrew Michael Wallace, and Iain H. Woodhouse. *Devel-*

- oping Hyperspectral LiDAR for Structural and Biochemical Analysis of Forest Data*, pages 59–60. 5 2012.
- [9] Ryan W Wolcott and Ryan M Eustice. Visual localization within lidar maps for automated urban driving. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 176–183. IEEE, 2014.
- [10] Dongeek Shin, Jeffrey H Shapiro, and Vivek K Goyal. Single-photon depth imaging using a union-of-subspaces model. *IEEE Signal Processing Letters*, 22(12):2254–2258, 2015.
- [11] Dongeek Shin, Jeffrey H Shapiro, and Vivek K Goyal. Computational single-photon depth imaging without transverse regularization. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 973–977. IEEE, 2016.
- [12] Deanna Needell and Joel A Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.
- [13] Dongeek Shin, Ahmed Kirmani, Vivek K Goyal, and Jeffrey H Shapiro. Computational 3d and reflectivity imaging with high photon efficiency. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 46–50. IEEE, 2014.
- [14] Dongeek Shin, Ahmed Kirmani, Vivek K Goyal, and Jeffrey H Shapiro. Photon-efficient computational 3-d and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging*, 1(2):112–125, 2015.
- [15] Yoann Altmann, Aurora Maccarone, Aongus McCarthy, Gregory Newstadt, Gerald S Buller, Steve McLaughlin, and Alfred Hero. Robust spectral unmixing of sparse multispectral lidar waveforms using gamma markov random fields. *To appear in IEEE Transactions on Computational Imaging*, 2016.
- [16] Yoann Altmann, Aurora Maccarone, Abderrahim Halimi, Aongus McCarthy, Gerald Buller, and Steve McLaughlin. Efficient range estimation and material quantification from multispectral lidar waveforms. In *Sensor Signal Processing for Defence (SSPD), 2016*, pages 1–5. IEEE, 2016.
- [17] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.
- [18] Mário A.T. Figueiredo and José M. Bioucas-Dias. Restoration of poissonian images using alternating direction optimization. *IEEE transactions on Image Processing*, 19(12):3133–3145, 2010.

-
- [19] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.
 - [20] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.
 - [21] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing mri. *IEEE signal processing magazine*, 25(2):72–82, 2008.
 - [22] Gilles Puy, Jose P Marques, Rolf Gruetter, Jean-Philippe Thiran, Dimitri Van De Ville, Pierre Vanderghenst, and Yves Wiaux. Spread spectrum magnetic resonance imaging. *IEEE transactions on medical imaging*, 31(3):586–598, 2012.
 - [23] Anna Auría, Alessandro Daducci, J-P Thiran, and Yves Wiaux. Structured sparsity for spatially coherent fibre orientation estimation in diffusion mri. *NeuroImage*, 115:245–255, 2015.
 - [24] Abdullah Abdulaziz, Arwa Dabbech, Alexandru Onose, and Yves Wiaux. A low-rank and joint-sparsity model for hyper-spectral radio-interferometric imaging. In *Signal Processing Conference (EUSIPCO), 2016 24th European*, pages 388–392. IEEE, 2016.
 - [25] Rafael E Carrillo, JD McEwen, and Yves Wiaux. Sparsity averaging reweighted analysis (sara): a novel algorithm for radio-interferometric imaging. *Monthly Notices of the Royal Astronomical Society*, 426(2):1223–1234, 2012.
 - [26] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.
 - [27] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1):89–97, 2004.
 - [28] Amir Beck and Marc Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 18(11):2419–2434, 2009.
 - [29] Rafael E Carrillo, Jason D McEwen, Dimitri Van De Ville, Jean-Philippe Thiran, and Yves Wiaux. Sparsity averaging for compressive imaging. *IEEE Signal Processing Letters*, 20(6):591–594, 2013.
 - [30] Patrick L Combettes and J-C Pesquet. Image restoration subject to a total variation constraint. *IEEE transactions on image processing*, 13(9):1213–1222, 2004.

-
- [31] Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier analysis and applications*, 14(5):877–905, 2008.
 - [32] Yoann Altmann, Ximing Ren, Aongus McCarthy, Gerald S Buller, and Steve McLaughlin. Lidar waveform-based analysis of depth images constructed using sparse single-photon data. *IEEE Transactions on Image Processing*, 25(5):1935–1946, 2016.
 - [33] Marcelo Pereyra, José M. Bioucas-Dias, and Mário A.T. Figueiredo. Maximum-a-posteriori estimation with unknown regularisation parameters. In *Signal Processing Conference (EU-SIPCO), 2015 23rd European*, pages 230–234. IEEE, 2015.
 - [34] Charles M Stein. Estimation of the mean of a multivariate normal distribution. *The annals of Statistics*, pages 1135–1151, 1981.
 - [35] Yoann Altmann, Reuben Aspden, Miles Padgett, and Steve McLaughlin. A bayesian approach to denoising of single-photon binary images. *arXiv preprint arXiv:1609.04862*, 2016.
 - [36] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on pure and applied mathematics*, 57(11):1413–1457, 2004.
 - [37] Emmanuel Candes and Justin Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23(3):969, 2007.
 - [38] Zachary T Harmany, Roummel F Marcia, and Rebecca M Willett. This is spiral-tap: Sparse poisson intensity reconstruction algorithmstheory and practice. *IEEE Transactions on Image Processing*, 21(3):1084–1096, 2012.
 - [39] Abderrahim Halimi, Aurora Maccarone, Aongus McCarthy, Stephen McLaughlin, and Gerald Buller. Object depth profile and reflectivity restoration from sparse single-photon data acquired in underwater environments. *IEEE Transactions on Computational Imaging*, 2017.
 - [40] Christine Guillemot and Olivier Le Meur. Image inpainting: Overview and recent advances. *IEEE signal processing magazine*, 31(1):127–144, 2014.
 - [41] Patrick L Combettes and Valérie R Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.
 - [42] Mohammad Golbabaee and Pierre Vandergheynst. Compressed sensing of simultaneous low-rank and joint-sparse matrices. *arXiv preprint arXiv:1211.5058*, 2012.

- [43] Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012.
- [44] Nikos Komodakis and Jean-Christophe Pesquet. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. *IEEE Signal Processing Magazine*, 32(6):31–54, 2015.