Institutionen för systemteknik
Department of Electrical Engineering

Examensarbete

Compressed Sensing for 3D Laser Radar

Examensarbete utfört i signal- och bildbehandling
vid Tekniska högskolan vid Linköpings universitet
av

Erik Fall

LiTH-ISY-EX–14/4767–SE

Linköping 2014

Linköpings universitet
TEKNISKA HÖGSKOLAN

Compressed Sensing for 3D Laser Radar

Examensarbete utfört i signal- och bildbehandling
vid Tekniska högskolan vid Linköpings universitet
av

Erik Fall

LiTH-ISY-EX–14/4767–SE

Handledare:	**Christina Grönwall**
	FOI
	Henrik Petersson
	FOI
	Hannes Ovrén
	ISY, Linköpings Universitet
Examinator:	**Maria Magnusson**
	ISY, Linköpings Universitet

Linköping, 9 juni 2014

Titel Title	Compressed Sensing för 3D Laserradar Compressed Sensing for 3D Laser Radar

Författare Author	Erik Fall

Sammanfattning
Abstract

High resolution 3D images are of high interest in military operations, where data can be used to classify and identify targets. The Swedish defence research agency (FOI) is interested in the latest research and technologies in this area. A drawback with normal 3D-laser systems are the lack of high resolution for long range measurements. One technique for high long range resolution laser radar is based on time correlated single photon counting (TCSPC). By repetitively sending out short laser pulses and measure the time of flight (TOF) of single reflected photons, extremely accurate range measurements can be done. A drawback with this method is that it is hard to create single photon detectors with many pixels and high temporal resolution, hence a single detector is used. Scanning an entire scene with one detector is very time consuming and instead, as this thesis is all about, the entire scene can be measured with less measurements than the number of pixels. To do this a technique called compressed sensing (CS) is introduced. CS utilizes that signals normally are compressible and can be represented sparse in some basis representation. CS sets other requirements on the sampling compared to the normal Shannon-Nyquist sampling theorem. With a digital micromirror device (DMD) linear combinations of the scene can be reflected onto the single photon detector, creating scalar intensity values as measurements. This means that fewer DMD-patterns than the number of pixels can reconstruct the entire 3D-scene. In this thesis a computer model of the laser system helps to evaluate different CS reconstruction methods with different scenarios of the laser system and the scene. The results show how many measurements that are required to reconstruct scenes properly and how the DMD-patterns effect the results. CS proves to enable a great reduction, $85 - 95$ %, of the required measurements compared to pixel-by-pixel scanning system. Total variation minimization proves to be the best choice of reconstruction method.

Abstract

High resolution 3D images are of high interest in military operations, where data can be used to classify and identify targets. The Swedish defence research agency (FOI) is interested in the latest research and technologies in this area. A drawback with normal 3D-laser systems are the lack of high resolution for long range measurements. One technique for high long range resolution laser radar is based on time correlated single photon counting (TCSPC). By repetitively sending out short laser pulses and measure the time of flight (TOF) of single reflected photons, extremely accurate range measurements can be done. A drawback with this method is that it is hard to create single photon detectors with many pixels and high temporal resolution, hence a single detector is used. Scanning an entire scene with one detector is very time consuming and instead, as this thesis is all about, the entire scene can be measured with less measurements than the number of pixels. To do this a technique called compressed sensing (CS) is introduced. CS utilizes that signals normally are compressible and can be represented sparse in some basis representation. CS sets other requirements on the sampling compared to the normal Shannon-Nyquist sampling theorem. With a digital micromirror device (DMD) linear combinations of the scene can be reflected onto the single photon detector, creating scalar intensity values as measurements. This means that fewer DMD-patterns than the number of pixels can reconstruct the entire 3D-scene. In this thesis a computer model of the laser system helps to evaluate different CS reconstruction methods with different scenarios of the laser system and the scene. The results show how many measurements that are required to reconstruct scenes properly and how the DMD-patterns effect the results. CS proves to enable a great reduction, $85 - 95$ %, of the required measurements compared to pixel-by-pixel scanning system. Total variation minimization proves to be the best choice of reconstruction method.

Sammanfattning

Högupplösta 3D-bilder är väldigt intressanta i militära operationer där data kan utnyttjas för klassificering och identifiering av mål. Det är av stort intresse hos Totalförsvarets forskningsinstitut (FOI) att undersöka de senaste teknikerna inom detta område. Ett stort problem med vanliga 3D-lasersystem är att de saknar hög upplösning för långa mätavstånd. En teknik som har hög avståndsupplösning är tidskorrelerande enfotonräknare, som kan räkna enstaka fotoner med extremt bra noggrannhet. Ett sådant system belyser en scen med laserljus och mäter sedan reflektionstiden för enstaka fotoner och kan på så sätt mäta avstånd. Problemet med denna metod är att göra detektion av många pixlar när man bara kan använda en detektor. Att skanna en hel scen med en detektor tar väldigt lång tid och istället handlar det här exjobbet om att göra färre mätningar än antalet pixlar, men ändå återskapa hela 3D-scenen. För att åstadkomma detta används en ny teknik kallad Compressed Sensing (CS). CS utnyttjar att mätdata normalt är komprimerbar och skiljer sig från det traditionella Shannon-Nyquists krav på sampling. Med hjälp av ett Digital Micromirror Device (DMD) kan linjärkombinationer av scenen speglas ner på enfotondetektorn och med färre DMD-mönster än antalet pixlar kan hela 3D-scenen återskapas. Med hjälp av en egenutvecklad lasermodell evalueras olika CS rekonstruktionsmetoder och olika scenarier av lasersystemet. Arbetet visar att basrepresentationen avgör hur många mätningar som behövs och hur olika uppbyggnader av DMD-mönstren påverkar resultatet. CS visar sig möjliggöra att $85 - 95$ % färre mätningar än antalet pixlar behövs för att avbilda hela 3D-scener. Total variation minimization visar sig var det bästa valet av rekonstruktionsmetod.

Contents

1

Introduction

1.1 Background

In recent years at the Swedish Defence Research Agency (FOI) there has been research on laser technology for high resolution 3D-range images at long range distances. High resolution 3D-laser radar is of interest in target classification, especially for partly hidden targets. To achieve the best result the range resolution needs to be better, higher, than the distance between camouflage and target. High resolution at far distances is of interest since more details means better opportunities for object classification or even identification. In military and police operations classification at large distances gives more time to perform target specific actions, compared to close range classification. At the same time, short scanning times are of high value minimizing the risk of one self being detected.

One of the technologies in 3D-laser radar is based on time correlated single-photon counting (TCSPC). By repetitively sending out short laser pulses and measure the time of flight (TOF) of single reflected photons extremely accurate range measurements can be done. A drawback with this method is that it is hard to create single photon detectors with many pixels and high temporal resolution, therefore a single detector is used.

1.2 Problem formulation

To capture a scene with a single TCSPC detector a pixel-by-pixel scanning system can be used. The drawback with this method is that scanning the scene pixel-by-pixel is too slow and too mechanically sensitive. Scanning every pixel can be avoided by using compressed sensing (CS), also referred to as compressed sam-

pling. CS utilizes that measurement data from a scene normally contain redundant information. It is based on theory that differs from the traditional Shannon-Nyquist sampling theorem. In the measurements, basis functions are utilized to sparsely represent the data. With help of a digital micromirror device (DMD) and a detector element, a single-pixel camera with TCSPC technology can scan a 3D-scene with much fewer measurements than a pixel-by-pixel scanning system. The challenge is how this can be done properly and the question is how images that can be reconstructed from fewer measurements than the number of pixels. To solve this problem, knowledge about how a TCSPC laser system is characterized and how CS techniques can reconstruct laser intensity images are needed. An overview of the imaging system is shown in Figure 1.1.
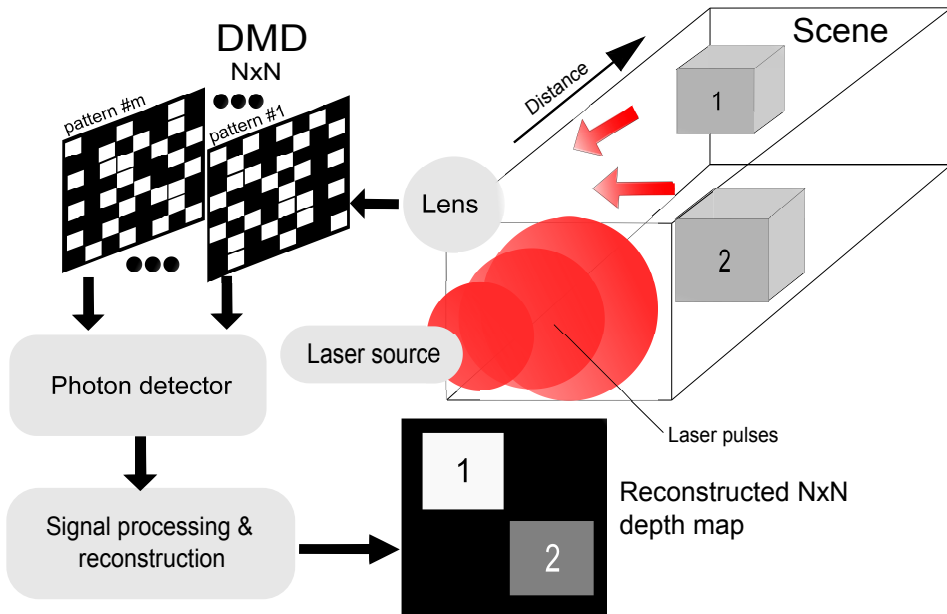

Figure 1.1: Overview of the system.

1.3   Objectives

The objectives of this thesis are:

- Develop a computer model of a TCSPC laser system, see Figure 1.1, and a CS reconstruction scheme for 3D-scenes. This model enables easy testing and evaluations of different 3D-scenes, system settings and CS algorithms.

- Determine how signal basis representation, undersampling ratio, choice of DMD and other critical factors affect the reconstructions.

- Determine how much faster CS can make the acquisition of the scene compared to a pixel-by-pixel scanning system.

1.4 Limitations

Limitations and restrictions of the simulations are:

- Static scenes. The scenes in this thesis are all static, which means that everything is assumed to be completely still during all simulated measurements.

- Objects and scene complexities are restricted to simple cases. Simulated objects are boxes, spheres, boats and trees which all have constant surface properties. Objects are only placed inside the measured depth intervals.

- System resolution is limited to 128×128 pixels. Higher resolution would have increased the required time for all simulations. Resolutions ($> 200 \times 200$) also require more system memory than available on the computers used for the simulations.

1.5 Related work

There are many articles which utilize compressed sensing when acquiring images using a single detector element, both in 2D-imaging and 3D-range imaging. Some of these are:

- *Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor* [11] by Kirmani et al., 2011. This article describes the usage of a spatial light modulator (SLM), that illuminates the scene with different binary patterns. First a fully transparent SLM is used to estimate different depth ranges. This gives information about the depth range where objects are present, which can be utilized in the 3D-reconstruction. The reconstruction uses assumptions about planar facets and the main limitation of this work is its inapplicability to scenes with curved linear objects. [3] and [12] are two articles which are shorter or slightly modified versions of this article.

- *Gated viewing laser imaging with compressive sensing* [15] by Li et al., 2012. This article presents a prototype of gated viewing imaging with compressive sensing. Total variation minimization promotes sparsity using the TVAL3 algorithm and the 3D scene is reconstructed using the mean time of flight for photons.

- *Photon-counting compressive sensing laser radar for 3D imaging* [10] by Howland et al., 2011. This article uses a TCSPC system and shows good reconstruction quality of images using Haar wavelets as the sparse basis representation of the signal.

- *Single-pixel imaging via compressive sampling: Building simpler, smaller, and less-expensive digital cameras* [5] by Duarte et al., 2008. This article describes the theory behind single-pixel camera and compressed sensing in detail. Theoretical and practical performance are evaluated compared to conventional cameras based on pixel arrays and raster scanning. Results

show that compressed sensing enables faster image acquirement than both raster scan and basis scan.

- *Photon counting compressive depth mapping* [9] by Howland et al., 2013. This article demonstrates a compressed sensing photon counting lidar system based on the single-pixel camera. The method presented recovers both depth and intensity information. The article also demonstrates the advantages of extra sparsity when reconstructing the difference between two instances of a scene.

- *A Compressive Sensing Photoacoustic Imaging System Based on Digital Micromirror Device* [16] by Naizhang et al., 2012. A photoacoustic compressed sensing imaging scheme is described in this article. Compressed sensing is used to speed up acquisition time using a digital micromirror device (DMD) as illumination mask.

- *Compressive confocal microscopy: 3D-reconstruction algorithms* [23] by Ye et al., 2009. In this article compressed sensing is used for image acquisition in Confocal Microscopy. A DMD is used together with a single photon detector to generate images. Three different 3D-reconstruction methods are evaluated, the two most interesting of them are not possible to implement in this master thesis, since their experiment setup differ too much.

1.6 Thesis Outline

In this report, theoretical backgrounds to the 3D-laser system and CS are given in Chapter 2 and Chapter 3, respectively. The analysis method together with the developed computer model are described in Chapter 4. Simulations and results are presented in Chapter 5 followed by discussion in Chapter 6 and conclusions and future work in Chapter 7.

2

The Laser System

The laser system is the center part of the entire range image system, which is illustrated in Figure 1.1. The laser system includes the laser source, noise sources affecting the laser light and the detection. The signal processing and image reconstruction is covered in Chapter 3.

2.1 TOF Laser system

Laser pulses are emitted onto a scene and by measuring the time of flight (TOF) of the reflected photons range measurements are created. Knowing the TOF t the distance to a target is calculated as

$$R = \frac{ct}{2},$$
(2.1)

where c is the speed of light. The light travels the distance ct, but since it is a reflection the range is given by half that distance [17, 19]. Generally when a laser pulse hits a surface it is reflected in different directions, depending on the surface properties and the angle of the incoming light. The returning power can be expressed by a simplified radar range equation,

$$P_R = \frac{P_T \eta_T}{\Omega_T R^2} \cdot A_\triangle \cdot \frac{A_R \eta_R}{R^2} \cdot e^{-2\alpha R},$$
(2.2)

where P_R is the received power, P_T the transmitted power, η_t and η_R are the efficiency of the transmitter and the receiver. The laser beam divergence (angular measure of the increase in beam diameter with distance) is denoted Ω_T, $A_\triangle$ is the target cross section and A_R is the aperture area of the receiver. Atmospheric

attenuation is denoted as α and R is the distance given in (2.1) [17, 19].

When measuring the 3D-contents of a scene, the task is to determine the range value for each pixel, where the round trip time is t_r, given by (2.1). By deciding when to measure after transmitting the laser pulse, the range to the measured depth interval can be chosen.

TCSPC is a method for range profiling with high resolution. One laser pulse equals one measurement and a range profile, histogram, is created from summarizing many measurements. The TCSPC system emits short laser pulses with a high repetition rate. The detector is then able to detect single photons, which creates an electrical signal. Detection probability of a transmitted pulse is in the range of 1-10%. Thanks to that many pulse cycles photons are measured and a histogram can be generated provided that the TOF of detected photons are registered. The histogram's bins correspond to when, in time, the photons were detected [17, 20]. The time bins are then translated to range bins by using (2.1). Examples of histograms from a scene are shown in Figure 2.1.
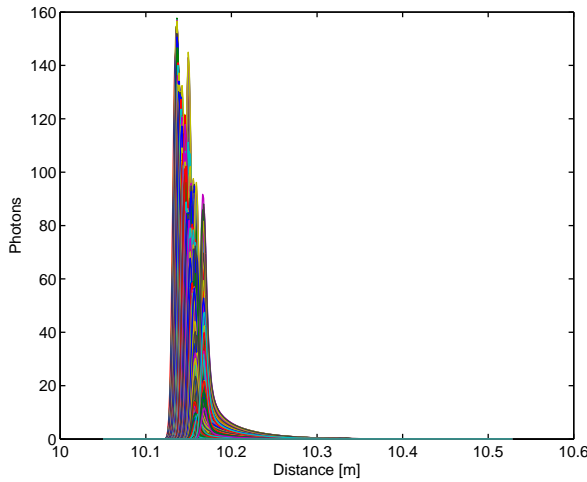

Figure 2.1: *Histograms of the received, noise free, signals from all pixels, where each colored curve corresponds to one pixel. In this case there are two objects at a distance of around* 10.15 *m.*

The detector is always active and the quantization of time intervals defines the time bins. A time bin is defined as $t_i = t_0 + i \cdot \Delta t$, where t_0 is the initial delay, which corresponds to how far away the measured depth interval is, Δt is the bin length and $i = 1, 2, ..., N_{bins}$, where N_{bins} is the number of time slices in the depth interval. The pulse repetition time determines the number of time bins, where the total time of all time bins are the time between two pulses. Having longer pulse repetition time means that the measured depth interval will be longer. In

this thesis the time bins are assumed to be 8 ps long and are in total 400, which means a bin resolution of approximately 1.2 mm and total depth range interval of approximately 0.48 m. The detector is always active which causes objects outside a measured depth range interval to be aliases into the measurements. This is caused by that transmitted laser light also hit objects further away than wanted, in the range interval, where the reflected light will not have synced TOF.

The laser system in this thesis is a combination of CS single-pixel camera technique and the described TCSPC technique.

2.2 System characterization

The histogram will, for a reflective target at a certain distance, contain photons detected in multiple time bins which is caused by random variations in time called time jitter. The temporal accuracy is determined by this uncertainty. System time jitters can be described by the following expression

$$\Delta t_{system} = \sqrt{\Delta t_{laser}^2 + \Delta t_{detector}^2 + \Delta t_{electronics}^2} \,, \tag{2.3}$$

where Δt_{laser}, $\Delta t_{detector}$ and $\Delta t_{electronics}$ are the jitter caused by the laser, detector and the electronics, respectively [19]. The characteristics of the system time jitter distribution is described by the instrumental response function (IRF), in signal theory referred to as point spread function, where the width of the IRF commonly is used as the approximation of Δt_{system}. The IRF (unpublished work; courtesy of FOI), of the system described in this thesis, is as follows

$$y = \begin{cases} e^{\frac{-t_1^2}{2s^2}} \cdot e^{\frac{t-t_1}{T_0}}, & t < t_1 \\ e^{\frac{-t^2}{2s^2}}, & t_1 < t < t_2 \\ G \cdot (a \cdot e^{\frac{-(t-t_2)}{T_1}} + b \cdot e^{\frac{-(t-t_2)}{T_2}} + q \cdot e^{\frac{-(t-t_2)}{T_3}}), & t > t_2 \end{cases} \tag{2.4}$$

where y is the IRF and G is defined as

$$G = e^{\frac{-t_2^2}{2s^2}}, \tag{2.5}$$

q is defined as

$$q = 1 - a - b \tag{2.6}$$

and T_3 is defined as

$$T_3 = \frac{q}{\frac{t_2}{s^2} - \frac{a}{T_1} - \frac{b}{T_2}}, \tag{2.7}$$

where T_0, T_1, T_2, t_1, t_2, a, b and s are all system parameters specific for the TCSPC system at FOI and t is the time interval where the IRF is calculated. The IRF in (2.4) is used in the simulated laser model. In Figure 2.2 an example of the IRF is shown. Note the exponential decreasing tail; this will cause traces of objects in the scene at distances further away from where the objects actually are, this is
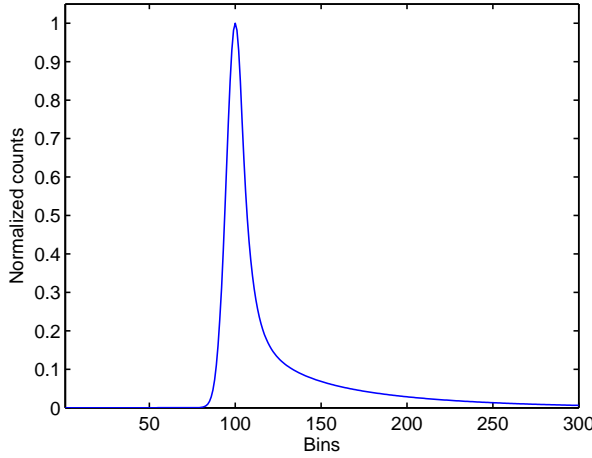
discussed later.

Figure 2.2: An example of the IRF for a single point target at time 100 ps. Here one time bin equals 4 ps.

2.3 Detection noise

There are three main noise sources that are taken into consideration in the laser system model: Shot noise, background illumination and dark counts.

The sampling of photons suffer from shot noise, which is associated with the particle nature of light. The detection is modeled as a Poisson process [7]. Measuring a certain amount, K, photons during the time interval $[t, t+\tau]$ has the following probability density function:

$$P(K; t, t+\tau) = P(K) = \frac{\bar{K}^K}{K!}e^{-\bar{K}}, \tag{2.8}$$

where K is the detected photons and $\bar{K}$ is the expected value of K in the time interval $t + \tau$. For large $\bar{K}$ the Poisson distribution approaches a normal distribution. The standard deviation is the square root of $\bar{K}$, which means that for large $\bar{K}$ the influence of the shot noise is small.

Background illumination is the detected light which does not originate from the laser transmitter, but from other illuminators such as lamps and the sun. Their illumination of the scene means that unwanted photons hit the detector. The background illumination can be simulated by adding a constant to every pixel measurement according to reference values, which later is going to be sampled according to the Poisson process in (2.8), [18].

In the detector, on its semiconductor material, electron-hole pairs are generated

from thermal energy. This causes dark counts, which add false detections of photons to the measurements.

2.4   Digital micromirror device

A digital micro mirror device (DMD) is a precise light switch which modulates light using an array of small mirrors. A single photon detector element, together with different DMD-patterns, create unique measurements of the same scene. The different array elements can be directed to reflect light towards the detector or away from it. This means that every mirror element in the array corresponds to a pixel (limiting the system resolution).

The principles of the single element detector were described in Chapter 2. Mathematically one DMD-pattern can be represented as a matrix with values 0 or 1, see example in Figure 2.3.
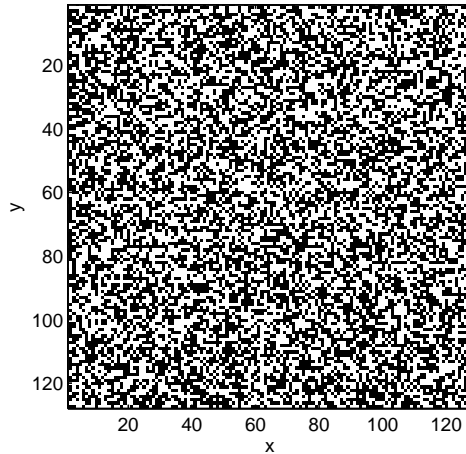

Figure 2.3: *A DMD-pattern with resolution* 128×128 *where the probability of a pixel being* 0 *or* 1 *is set to* 0.5. *Black pixels are* 1 *and white pixels are* 0.

Using compressed sensing, described in Chapter 3, the DMD-array provide the possibility of generating complete depth images, of the same resolutions as the DMD-arrays, using fewer samples than there are pixels.

3

Compressed Sensing

Compressed sensing is a method for acquiring and reconstructing signals by finding solutions to underdetermined linear systems using the knowledge of signals' compressibility and sparseness. The underdetermined systems are solved by adding the constraint of sparsity, which only allows solutions with a small number of non-zero coefficients. To perform compressed sensing the signal needs to be very sparse in some basis, while keeping its measurement matrix incoherent with the sparse transform.

3.1   Sparsity and compressible signals

A signal can often be expressed concise in some appropriate basis. Natural images can be expressed by e.g. discrete cosine transform (DCT), where few large coefficients can capture almost all information. This means that a lot of small coefficients can be discarded without losing much information and consequentially this is one of the basic ideas in image compressing. Mathematically the image can be described as a vector $f \in \mathbb{R}^n$ (where n is the number of pixels of the image). This can be represented in another basis, e.g. DCT, Ψ, as

$$f = \sum_{i=1}^{n} x_i \psi_i = \Psi x, \tag{3.1}$$

where x_i are the coefficients in the sparse basis representation of the signal f [2]. Note that certain types of images can be sparse already in the pixel basis, which means that Ψ equals the $n \times n$ identity matrix. Finding the most appropriate basis for reconstructing a scene is evaluated in Section 5.1.

By keeping the S largest values of $|x|$, in the expansion in (3.1), f can be described as $f_s = \Psi x_s$, where x_s is x with its S smallest values set to zero. If this representation can be done without much perceptual loss the signal is considered to be S-sparse in that basis representation [2].

3.1.1 Transform definitions

In image compressing it is well known that natural images have sparse representations using the DCT and wavelets, which are used in e.g. JPEG and JPEG2000. In [9] discrete Haar wavelets (DHWT) were used as sparse basis and higher order Daubechies wavelets were tested but did not significantly improve their results.

Three different transformations are used in this thesis: DCT, DHWT and the Hadamard transform. In (3.1) their inverse transform matrices corresponds to Ψ.

The orthonormal DCT of the signal f is defined as

$$F[j] = a[j] \sum_{k=0}^{n-1} f[k] \cos\left(\frac{(2k+1)j\pi}{2n}\right) = \sum_{k=0}^{n-1} f[k]c[j,k], \tag{3.2}$$

where $F[j]$ is element j of the transformed signal f and $a[j]$ is defined as

$$a[j] = \begin{cases} \sqrt{1/n} & \text{if } j = 0 \\ \sqrt{2/n} & \text{if } j = 1, 2, ..., n-1 \end{cases}. \tag{3.3}$$

The $n \times n$ cosine transform matrix C is then defined as

$$C = \begin{pmatrix} \cdots & \cdots & \cdots \\ \vdots & c[j,k] & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix}, \tag{3.4}$$

where C is orthogonal, which means that $F = Cf$ and $f = C^T F$.

The Hadamard transform is a generalized class of Fourier transforms, which computes the discrete Fourier transform of a signal, although the matrix is purely real. The Hadamard transform matrix is defined as

$$H_n = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{pmatrix}, \tag{3.5}$$

where $H_0 = 1$, from which the matrix H_n can be generated iteratively.

The Haar wavelet is the simplest possible wavelet, its definition will not be covered , instead its structure is described. The general $n \times n$ Haar wavelet transform matrix, used in this thesis, is described as

$$H_n = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{n/2} \otimes [1,1] \\ I_{n/2} \otimes [1,-1] \end{pmatrix}, \tag{3.6}$$

where $H_1 = 1$, I is the identity matrix and $\otimes$ is the Kronecker product. The normalization constant $\frac{1}{\sqrt{2}}$ ensures that $H_n^T H_n = I$. The Haar transform of the $n \times n$ signal f is then $F = H_n f$ with the inverse transform $f = H_n^T F$.

3.2   The compressed sensing problem

Consider a general linear measurement process that takes $m < n$ inner products between $f \in \mathbb{R}^n$ and a collection of vectors ϕ_j, where $j = 1, 2, ..., m$. The measurements, y_j, form the $m \times 1$ vector y,

$$y = \Phi f = \Phi \Psi x = Ax, \tag{3.7}$$

where f is substituted according to (3.1) and $A = \Phi \Psi$ is an $m \times n$ matrix. The equation in (3.7) is ill-conditioned and has an infinite number of solutions.

The CS problem consists of solving the equation in (3.7) which implies:

- Designing a reconstruction algorithm to recover x.

- Designing a stable measurement matrix Φ such that the important information in a general S-sparse or compressible signal is not damaged by reducing the dimensionality from $\mathbb{R}^n$ to $\mathbb{R}^m$.

- Find a basis where the signal is sparse, i.e designing Ψ.

3.2.1   Designing the measurement matrix

Since the number of measurements m are fewer than n the problem of recovering the signal f is ill-conditioned. If the signal f is S-sparse in some domain and if we know which elements that are non-zero the problem can be solved if $m \geq K$. For this problem to be well conditioned it is required, for all S-sparse vectors $\hat{x}$ and for each $S = 1, 2, ...$, that

$$1 - \delta_S \leq \frac{\|A\hat{x}\|_2}{\|\hat{x}\|_2} \leq 1 + \delta_S, \tag{3.8}$$

where δ_S is some small error, not too close to one, and $\|\cdot\|_2$ is the Euclidean norm. This simply means that the matrix A must preserve the lengths of the vectors, to a certain limit. This is referred to as the restricted isometries problem (RIP). Generally there is no knowledge about which elements are non-zero. The goal then is to find the matrix A that satisfies the RIP, without knowing which elements in $\hat{x}$ that are non-zero [2, 6].

Another condition that needs to be satisfied is incoherence between the measurement matrix Φ and base matrix Ψ. In other words, the rows ϕ_j of Φ must not sparsely represent the columns ψ_j of Ψ [2]. Fortunately random matrices are incoherent with any fixed fix basis matrix Ψ, which means that choosing the measurement matrices randomly is a very good idea.

Choosing a random measurement matrix Φ can be shown to fulfill the RIP with high probability using only $m \geq C \cdot S \cdot log(n/S) << n$ measurements, where C is some small constant and n is the length of the signal f [2]. This is the reason why the DMD patterns, explained in Section 2.4, will be randomly generated.

3.2.2 Designing a reconstruction algorithm

The inverse problem that needs to be solved is formulated as

$$y = \Phi\Psi x = Ax, \tag{3.9}$$

which intuitively is solved by minimizing the Euclidean norm $\|Ax - y\|_2$, which has the closed form solution $\hat{x} = A^T(AA^T)^{-1}y$. This approach is ill-conditioned and does not give sparse solutions.

Another approach of expressing the problem in (3.9) is to define the l_p norm and rewrite the expression as

$$\hat{x} = \text{argmin}\|x\|_{l_p} \text{ subject to } Ax = y, \tag{3.10}$$

where the l_p norm should promote sparsity while being convex.

3.2.3 The l_p- norm

The l_p norm of the vector x, for $p \geq 1$, is defined as

$$\|x\|_{l_p} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \tag{3.11}$$

where n is the dimension of x. The l_0 norm counts the number of non-zero elements of a vector, but minimizing using the l_0 norm is a non-convex problem, as is all l_p norms for $p < 1$, and therefore not computationally tractable or numerically stable. This is the reason why the l_0 norm is not used in the compressed sensing problem.

We can get a intuition why the l_1 norm is a great substitute for sparsity by studying the sketches in Figure 3.1, where $x \in \mathbb{R}^2$. The red lines represents the subspace $H = \{x : \Phi\Psi x = y\}$, which corresponds to the noiseless case in (3.15). The minimal norm where each l-ball (all points where $\|x\|_{l_p}$ is the same constant) intersects H is marked as x^*.

Figure 3.1b shows the l_1-ball when it intersects the subspace H. Note that it is spiky along the axes as the non-convex $l_{1/2}$-ball in Figure 3.1a, whereas the l_2-ball in Figure 3.1c is not. This is commonly denoted as that the l_1 and $l_{1/2}$-ball are anisotropic whereas the standard Euclidean l_2-ball is isotropic. Larger p spread the norms more evenly along two coefficients, whereas smaller p give norms where the coefficients are more unevenly distributed and sparse. This also generalizes to higher signal dimensions n, where $x \in \mathbb{R}^n$ [4].
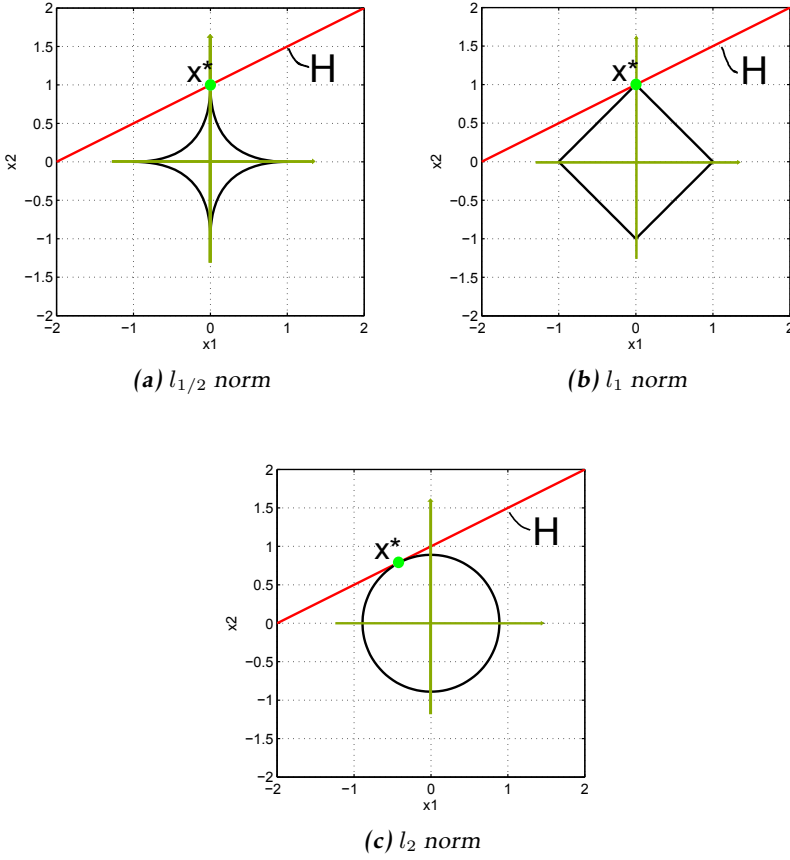
(a) $l_{1/2}$ *norm*

(b) l_1 *norm*

(c) l_2 *norm*

Figure 3.1: *Three different l_p norms plotted at minimal intersection, $\boldsymbol{x}^*$, with the subspace $H = \{x : \Phi\Psi x = y\}$. Coordinate axes are marked in green.*

3.2.4   Total variation minimization

Total variation (TV) minimization is another approach compared to l_1 minimization. Instead of assuming that the signal is sparse in some basis it considers the gradient of the signal being sparse. In other words, TV minimization seeks the solution with the sparsest gradient. Advantages of using TV minimization are good preservation of edges and good recovery of dense staircase signals or piecewise constant images [14]. The TV compressed sensing problem is defined as

$$\operatorname{argmin} \sum_{i,j} \|D_{i,j}f\|_{l_2} \quad \text{subject to} \quad Af = y, \tag{3.12}$$

where f is the signal, in this case image, $D_{i,j}f \in \mathbb{R}^2$ is the gradient in pixel (i,j), A the measurement matrix ($A = \Phi$) and y the measurements [14].

3.2.5 The single-pixel camera approach

To understand the sampling process of the TCSPC in combination with CS the simple 2D single-pixel camera is first described. The single-pixel camera is a digital camera with one single sensor element, a photodiode. Using a DMD, the single-detector element can samples the entire image of a scene. A simplified illustration of the single-pixel camera is shown in Figure 3.2.
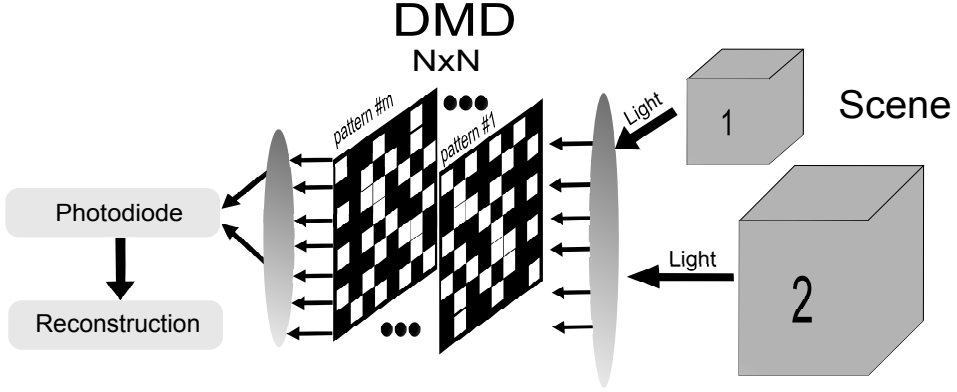
Figure 3.2: *Simplified illustration of the single-pixel camera, where $N \times N$ is the number of pixels n. The objects marked as 1 and 2 are placed somewhere in a 3D-scene reflecting light from the surroundings onto a lens which focuses the light to the current DMD-array, which creates a linear combination of the light which focused onto the single photodiode by a lens. The CS scheme then reconstructs the image from the measured data.*

Sampling the original image signal f, with a DMD-matrix can be described as a summation of the inner product of f and the measurement function $\phi_i \in \mathbb{R}^n$, generating a single measurement $y_i \in \mathbb{R}$. Taking detection error, ϵ which for the TCSPC system is described in Section 2.3, into consideration y_i can be formulated as

$$y_i = \langle f, \phi_i \rangle + \epsilon_i. \tag{3.13}$$

If not using CS the image is sampled using one active mirror element for each DMD-pattern, which requires as many DMD-patterns as there are pixels, n. Undersampling the signal f is represented by taking $m < n$ measurements, writing Φ as the $m \times n$ measurement matrix, where each row ϕ_i corresponds to one DMD pattern, and expressing the measurements as

$$y = \Phi f + \epsilon = \Phi \Psi x + \epsilon, \tag{3.14}$$

where Ψx is the sparse representation of f. Assume that the signal is compressible in some domain and x is K-sparse. Then the problem can be formulated as

(3.10), but with noise taken into consideration, as

$$\arg\min\|x\|_{l_1} \text{ subject to } \|y - \Phi\Psi x\|_{l_2} \leq \epsilon, \tag{3.15}$$

This particular formulation is referred to as basis pursuit denoise (BPDN) [2]. TV minimization, described in Section 3.2.4, is another approach.

3.2.6   3D-compressed sensing

The single pixel camera gives a good intuition of how one intensity image, which corresponds to a specific time bin, can be reconstructed. Imagine the single pixel camera taking images of 2D-slices out of the 3D-scene in Figure 3.2 and by combining all 2-D slices an image of the scene can be created. Using this approach, for intensity images from laser, a range image can be created.

When measuring using the TCSPC system each slice corresponds to an intensity image, where the intensity values are how many photons that hit the detector in that specific TOF interval. This means that if all the intensity images can be reconstructed a depth map can be created.

Mathematically the measurements can be described as

$$Y = \Phi F, \tag{3.16}$$

where Y is the $m \times bins$ matrix containing all the measurements and F is the $n \times bins$ signal matrix.
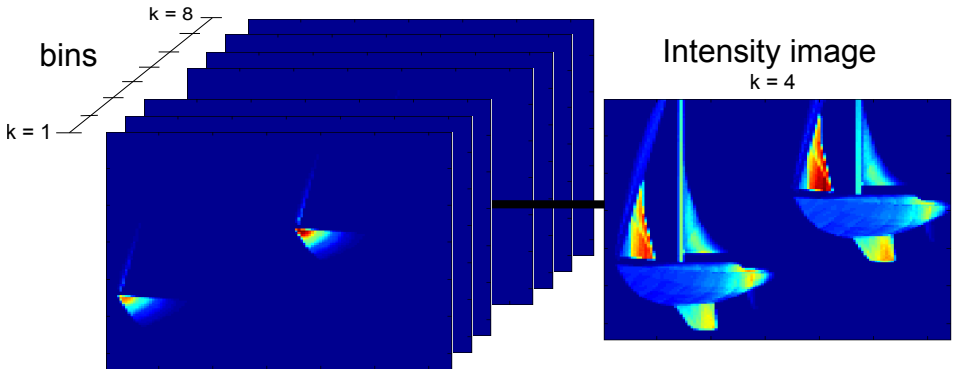
Figure 3.3: *Illustration of how the intensity images are stacked in 3D-space. In this figure, the entire scene is considered to be covered in 8 bins. These intensity images from different bins generate the range image. Here, the front sails have their maximum intensity values in bin 4. This means that their range image corresponding pixels will have range corresponding to bin 4.*

The measurement process is done for one DMD-array at the time, for the entire range, filling one row of Y for each DMD-array. The data used for reconstructing

slice j is then in column j in Y. Reconstructing the total signal F is performed using e.g. (3.15) or TV minimization for each slice f_j.

A range image is created by setting the range in each pixel corresponding to its bin with the maximum intensity value, i.e. finding the maximum intensity in all the intensity images. An illustration of how the intensity images are stacked in 3D-space is shown in Figure 3.3.

Noise is often present in the reconstructed images, normally taking negative or smaller values than those origin from the laser. This means that the simple operation of thresholding often can remove very much or all noise. Other image filtering techniques e.g. median filtering can be used to remove noise, but like thresholding it comes with the risk of removing information.

4

Method

4.1 Feasibility study

A pre-study was initially performed, finding related work, covering compressed sensing and laser systems. Related articles and publications were found using web search and article database services provided by LiTH and FOI. Another pre-study, already done by FOI, was also used in the early stages [8] as inspiration. In Section 1.5 related works are summarized. Inspiration and ideas, which were good and possible to implement, were taken from these related works.

4.2 Computer model

To evaluate different compressed sensing and 3D-reconstruction schemes, a computer model of the laser system was needed. The developed model simulates the TCSPC laser system, described in Chapter 2. The computer model was implemented in MATLAB®. It is based on a general laser simulator developed by FOI. In Figure 4.1 the processing steps, implemented in the computer model, are described using a block scheme.
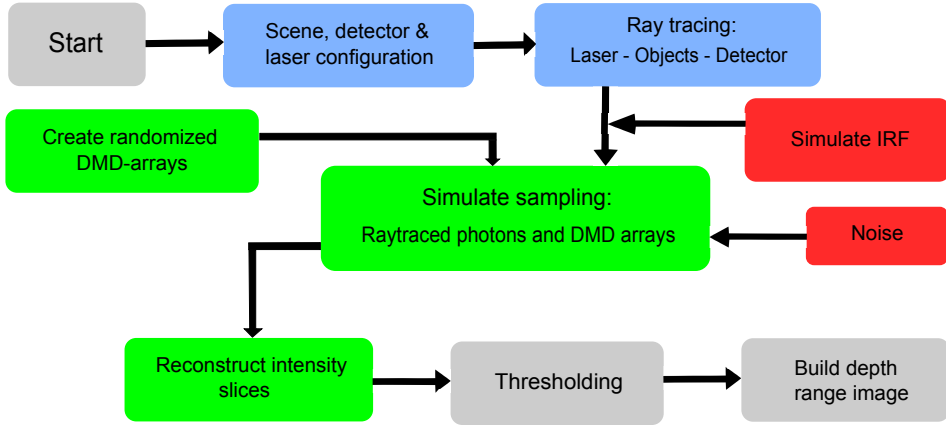
Figure 4.1: *Block scheme of the simulation process steps. The blue blocks are implemented based on the laser simulator developed by FOI. The green blocks are directly related to CS. The red blocks are related to the characteristics of the laser detector. The gray blocks are neither directly related to the CS nor to the laser simulations.*

A scene is configured with objects of different sizes, shapes and surface properties. Ray tracing, of the laser light, is performed and returns the amount of laser light that hits the detector. A model of the detector is applied on the light, that hits the detector, and in the end gives a measurement of photons in different time bins and DMD-patterns. Intensity images of specific bins is then reconstructed using compressed sensing and finally a depth image is created using the intensity images.

4.2.1 Configuration

The configuration of the laser model and scene include different type of objects with different surface parameters, laser settings and detector settings. The objects settings are parameters controlling diffuse and specular reflectance. The laser source settings are pulse energy, pulse width, focus range and divergence. The only parameter that is varied here is the pulse energy, which results in different received signal strengths. The aperture settings are resolution, detector size, distance, height, and orientation. The detector also has settings, the used ones were; the number of bins, time resolution, system efficiency, bandwidth and aperture radius.

4.2.2 Laser calculations

The ray tracing calculates how much of the transmitted laser light that hits the detector and when it hits the detector. In Figure 4.2 the concept of ray tracing is illustrated. The laser model calculates reflections depending on the surfaces' diffuse and specular properties. Specularity makes the reflected light dependent of the reflection angle to the observer.
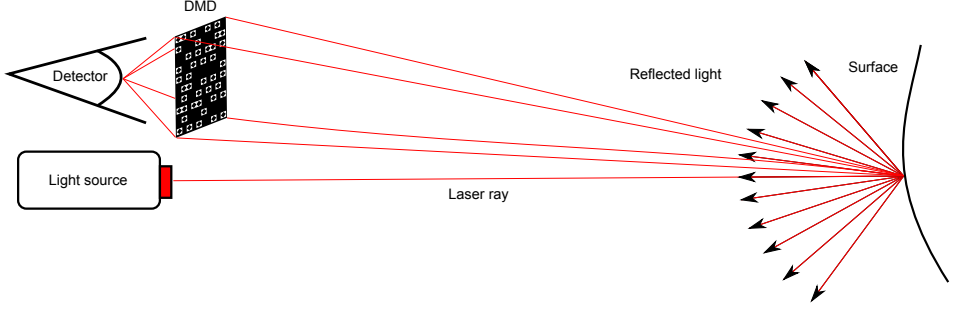
Figure 4.2: *Illustration of light being scattered in different directions when hitting a surface. The figure illustrates diffuse reflections, which reflects equal amounts of light in all direction, independent on the angle of the incoming light.*

The instrumental response function, described in Section 2.2, is combined with the ray traced data using 2D convolution. This creates the characteristics of the TCSPC laser detector. Background illumination noise is added here.

Simulation of the DMD is done by sampling the returned light from the ray tracing and IRF with different randomized matrices, see example in Figure 2.3. The matrices consists of values 0 or 1, each with a user-selected probability for being 1. $p = 0.1$ is used in the simulations in Chapter 5. Elements in the matrix are multiplied with corresponding pixel coordinates and the returning measure is a scalar value. The returned value equals the total amount of photons hitting the DMD-pixel elements which have value one.

The detector's shot noise is added after the simulated DMD-sampling, according to (2.8). Finally dark counts are added, where the total amount of dark counts increase linearly with the number of measurements.

4.2.3   Image reconstruction

The CS solvers were downloaded from the web and were configured, tuned and tweaked to work for multiple intensity images. The image reconstructions are done by using one of the software packages called ProPPa, SPGL1 or TVAL3.

ProPPa solves a basis pursuit problem or a l_1-regularized optimization problem. The basis pursuit solved by ProPPa is formulated as

$$\text{argmin} \|x\|_{l_1} \text{ subject to } Ax = y \ , x \in \mathbb{R}^n, \tag{4.1}$$

where, x is the sparse representation of the signal, $A = \Phi\Psi$, y the measurements and n the number of pixels.

The l_1-regularized least squares solved by ProPPa is formulated as

$$\text{argmin} \left(\|Ax - y\|_{l_2} + \beta \|x\|_{l_1} \right) \ , x \in \mathbb{R}^n, \tag{4.2}$$

where β is some constant and the rest of the variables are as in (4.1). Although noise is not included in the optimization formulations in (4.1) and (4.2) ProPPa works for noisy signals too.

SPGL1 solves the BPDN problem described in (3.15) or the LASSO problem defined as

$$\text{argmin}\|Ax - y\|_{l_2} \text{ subject to } |x| \leq \tau, \tag{4.3}$$

where τ is some constant limiting the norm of x [22].

TVAL3 is a software package, which solves the TV minimization problem described in Section 3.2.4 [14].

4.2.4   3D-reconstruction

The range image construction is done by first reconstructing all intensity images, one for each range bin. After that, the intensity images are thresholded to remove noise. The threshold level is chosen to remove intensities that are so low that they most likely are noise.

After the thresholding the range image is constructed by for each pixel finding its bin with the maximum intensity intensity value. Each found maximum corresponds to certain depth which becomes the range value of that specific pixel. If no maximum is found the thresholding has zeroed all bins in that particular pixel. That range pixel is then assigned to the maximum range in the measured depth interval.

4.3   Evaluation

Evaluations of the different methods were done by running simulations in MAT-LAB® applying different parameter settings and scene setups. Every test was specifically developed for the specific type of evaluation and generally basic parameters for solvers and settings were set to chosen default values. The reason for this was to have consistency in the results, making them comparable across different evaluations.

All simulations were performed with MATLAB® 2012b, running on a 64 bit PC with the hardware specs; Intel Xeon W3530 CPU and 12 GB of RAM.

When deciding the best CS solvers the choices were based on the following criteria:

- The solver should work on the provided computer hardware.

- The solver should handle large problems. The length of the 1D vector representation of images scales quadratically with the image side lengths. This means the signal vectors have the dimensions $f \in \mathbb{R}^{N_1 \cdot N_2}$, where N_1 and N_2 are the image side lengths. Hence f becomes very large when N_1 and N_2 grow.

- The solver should be able to solve the optimization problems within reasonable times (minutes instead of hours).

- The solver should work for different types of intensity images (different scene setups and signal strengths) without individual pre-tuning of parameters.

- The solver should preserve geometry, shapes and distances.

- Equally fast solvers are differentiated by their quality of the signal reconstruction, which is compared with reference data.

Simulated image sizes were 128×128 pixels with a total of 400 range bins per simulated scene. Mainly three reference scene were used in the evaluations, which are described in Figure 5.1. Another scene containing trees was also used, see Figure 5.22a.

The measures when evaluating different scenarios are either based on visual inspection or SNR of the reconstruction. The SNR, given in dB, is defined as

$$\text{SNR} = 20 \log \frac{\|X\|_F}{\|X - \hat{X}\|_F}, \tag{4.4}$$

where X is the reference image, $\hat{X}$ is the reconstructed image and $\| \cdot \|_F$ the Frobenius norm defined as

$$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}, \tag{4.5}$$

where a_{ij} is the matrix element i, j.

Undersampling ratio, q, is defined as the number of measurements m divided by the number of pixels n, $q = m/n$. This means that the undersampling ratio is a direct measure of how much faster the measuring of the scene is, compared to a pixel-by-pixel scanning system.

The evaluation was divided into the following test scheme:

- Evaluation of signal basis representation determine the best basis or bases when reconstructing the intensity images.

- This/these basis/bases are then used in the following evaluations which start with testing the importance of signal strength and noise.

  – The probability of the DMD-elements being 1 is evaluated.

  – Evaluation of optimization methods is performed, where reconstruction speed and reconstruction quality are tested.

  – Reconstruction of occluded objects and trees are tested, with focus on geometry preserving properties.

- Other approaches than the slice-by-slice intensity image reconstruction are briefly tested.

All the results are presented in Chapter 5.

5

Simulations and Results

In this chapter parameters and methods related to the CS reconstruction are evaluated. The objectives of the simulations is to give answers to the questions in Section 1.3. An important limitation is that most simulations were very time consuming, therefore limiting the number of data sets and resolution in plots.

Four different simulated scenes were used as test scenes. The scenes were chosen with unique settings covering the cases; objects at different distances, objects at the same distance, different object types and obscured objects. The reason for this was to get as reliable results as possible, since in reality a scene can contain almost anything. The four scenes are shown in Figure 5.1 and Figure 5.22a.

Figure 5.2 shows the variances in the scenes intensity images, which gives a intuition of how the pixel values are spread in each intensity image. The variance σ^2 is defined as $\sigma^2 = \frac{1}{n} \sum_i \sum_j (x_{i,j} - \bar{x})^2$, where, n is the number of pixels, i and j are the pixels coordinates of x and $\bar{x}$ its mean. Here, $x_{i,j}$ has values distributed from 0 to approximately 160. The plots show that the boat scene have more bins with object content than the other scenes.

5.1 Sparse signal representation

Representing the sampled signal in a sparse basis is essential in compressed sensing and the choice of basis representation should be the one that represents the signal as sparse as possible. As mentioned in Section 3.1, a signal can be transformed to another basis using a matrix multiplication with Ψ and this discrete representation of the transform is required when using the chosen solvers.

The acquired signals, which should be represented in a sparse basis, are intensity
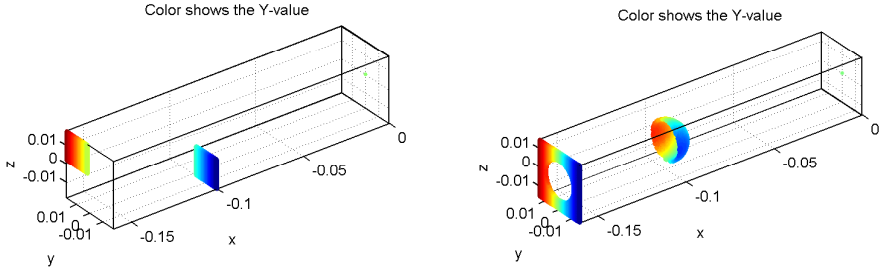
(a) Simulated scene with two squares at different distances, Reference scene 1.

(b) Simulated scene with a sphere in front of a square background, Reference scene 2.
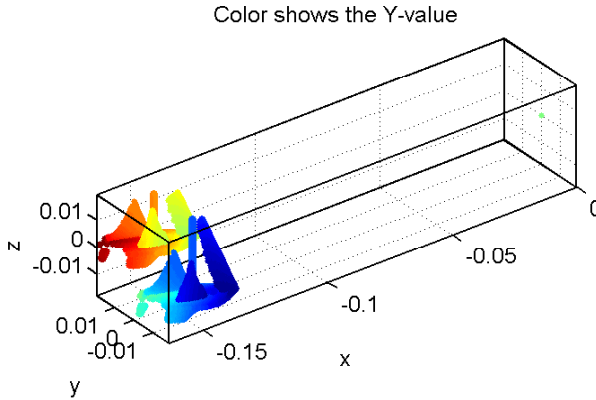
(c) Simulated scene with two boats at the same distance, Reference scene 3.

Figure 5.1: *The three different test scenes where x, y and z are in meters and the detector is placed at $(x, y, z) = (0, 0, 0)$.*

images from different depth bins. These intensity images may contain a lot of information with many non-zero pixel values or less information with many pixel values that are zero. Generally, the intensity images closest to the detector contain more zero-valued pixel elements than measures further away. This is because of traces in the measurements due to the tail of the IRF, described in Section 2.2.
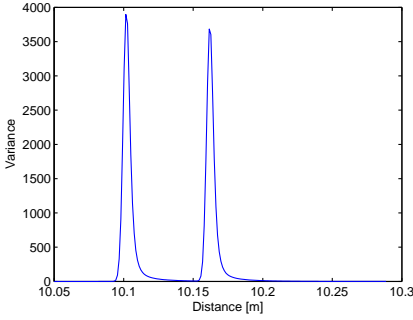
(a) *Intensity variance of Reference scene 1 in Figure 5.1a.*
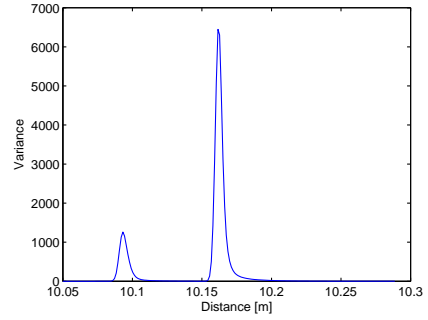
(b) *Intensity variance of Reference scene 2 in Figure 5.1b.*
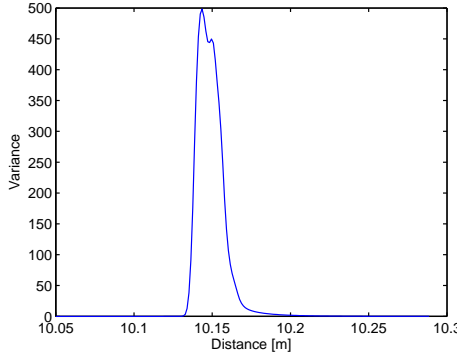
(c) *Intensity variance of Reference scene 3 in Figure 5.1c.*

Figure 5.2: *The variances, for different distances from the detector, of the reference scenes in Figure 5.1.*

This means that intensity images can be sparse in pixel basis in a measured depth interval close to the detector, while not in the other depth intervals. To be able to reconstruct the signal in all depths, a good basis that represents the signal sparse in all depths needs to be determined.

5.1.1   Intensity image reconstruction results

Four different basis representations were tested by reconstructing intensity images from the three different reference scenes in Figure 5.1. The signals were reconstructed and evaluated using discrete cosine transform (DCT), Haar wavelets (DHWT), Hadamard transform, normal pixel basis representation and TV minimization. TV is not a specific basis transformation, but is evaluated together with the different bases because it is used a lot in various CS literature and the question is if it suits this case. The following results are the SNR from reconstructing

slices out of the three test scenes with four different bases and TV minimization for different undersampling ratios.

The plots in Figure 5.3 and 5.4 show that the Reference scene 1 was best reconstructed using TV minimization. When considering basis choice, both the DCT and the DHWT outperformed the pixel base representation and the Hadamard transform.

Results for Reference scene 2 are shown in Figures 5.5 and 5.6. It is clear that TV minimization performed best and it is notable that, together with the DCT case, it performed well even for very low undersampling ratios. The pixel basis performed much worse than all other basis representations for all undersampling ratios, which makes it a bad choice for Reference scene 2. This is expected, since in this case the intensity images are not sparse in pixel basis.

Reconstruction results for Reference scene 3 are shown in Figures 5.7-5.9. The best reconstruction performances were achieved using TV minimization or when choosing basis representation DHWT or DCT. TV minimization handled all three investigated slices of the scene well. The scene had an increasing level of non-zero pixel values for larger distances. As a result of the increase of non-zero pixel values in range, the pixel basis representation performed worse compared to others. This was expected.
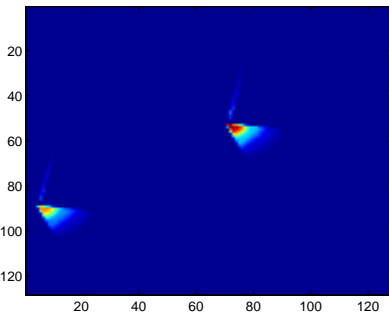

(a) Intensity image.   *(b) Reconstruction results.*
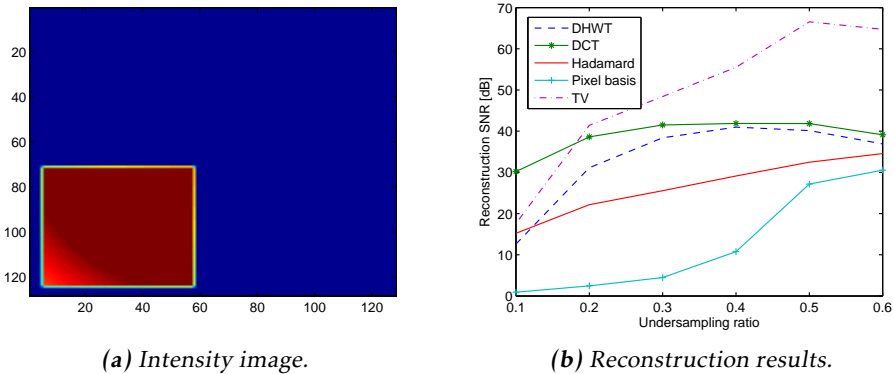
Figure 5.3: *Intensity image and reconstruction results from Reference scene 1 for the approximate depth* 10.1 *m (first peak in Figure 5.2a).*
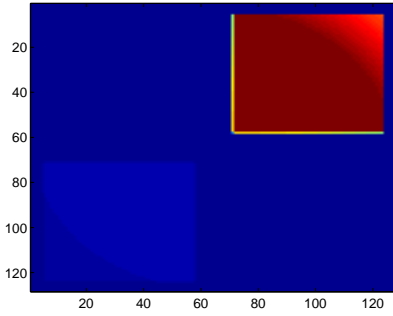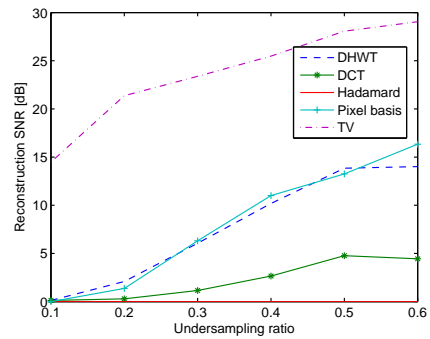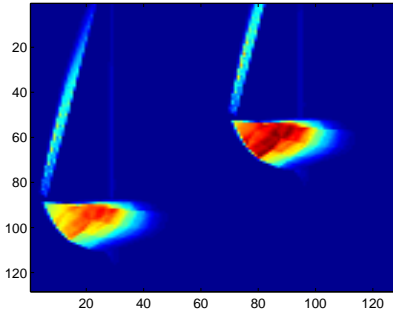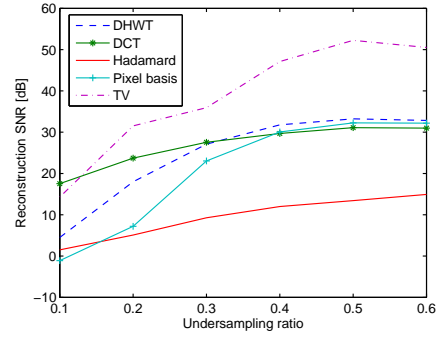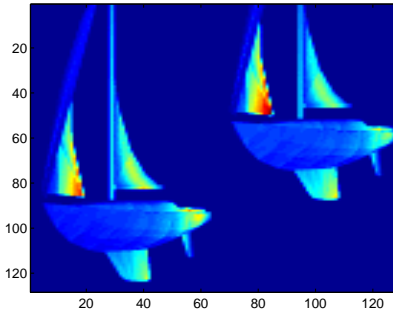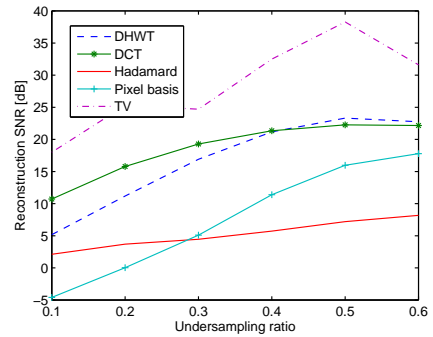
(a) Intensity image.
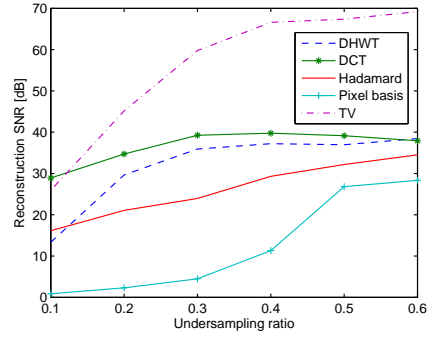
(b) Reconstruction results.

Figure 5.4: *Intensity image and reconstruction results from Reference scene 1 for the approximate depth* 10.16 *m (second peak in Figure 5.2a).*

The use of TV minimization generally performed better or much better compared to the DCT and DHWT. There is no clear winner when deciding which basis that has the best overall performance. The DCT and the DHWT perform good in almost all cases, which separates them from the Hadamard transform and pixel basis. Therefore the bases used in the remaining evaluations are the DCT and the DHWT.
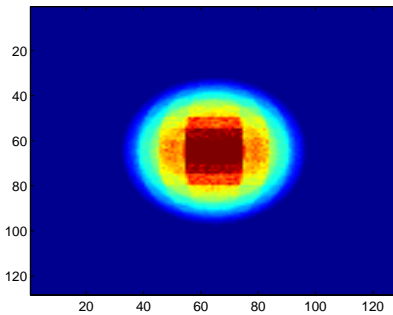
(a) Intensity image.
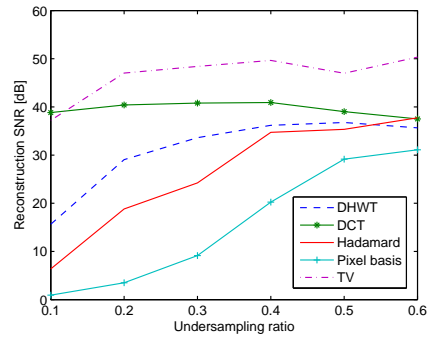
(b) Reconstruction results.

Figure 5.5: *Intensity image and reconstruction results from Reference scene 2 for the approximate depth* 10.09 *m (first peak in Figure 5.2b).*
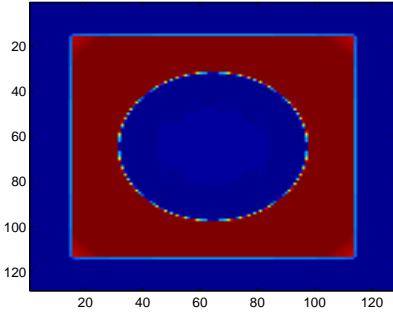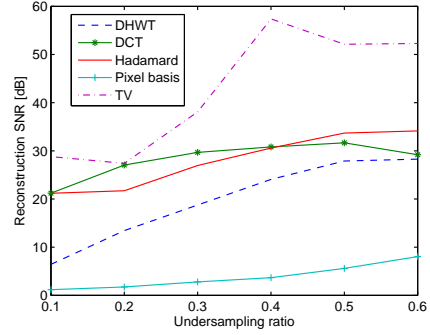
(a) Intensity image.

(b) Reconstruction results.

Figure 5.6: *Intensity image and reconstruction results from Reference scene 2 for the approximate depth* 10.16 *m (second peak in Figure 5.2b).*

(a) Intensity image.

(b) Reconstruction results.

Figure 5.7: *Intensity image and reconstruction results from Reference scene 3 for the approximate depth* 10.13 *m (beginning of the peak in Figure 5.2c).*

(a) *Intensity image.*   **(b)** *Reconstruction results.*

Figure 5.8: Intensity image and reconstruction results from Reference scene 3 for the approximate depth 10.14 *m (middle of the peak in Figure 5.2c).*

(a) *Intensity image.*   **(b)** *Reconstruction results.*

Figure 5.9: Intensity image and reconstruction results from Reference scene 3 for the approximate depth 10.16 *m (end of the top level of the peak in Figure 5.2c).*

5.1.2   Sparsity comparison

I this section the correlation between sparse signal representation and reconstruction results are evaluated. The reconstruction results of two different intensity images are correlated with the number of non-zero (or close to zero) coefficients for different basis representations. The histograms in this section are normalized, meaning that values have been scaled so they vary between min 0 and max 1.

The histograms in Figure 5.10a show that the image in Figure 5.7 represented in DCT or Hadamard transform have more non-zero coefficients compared to

DHWT or pixel basis. Note that reconstructing using DCT or Hadamard give
bad results for this image, see Figure 5.7b.

The histograms in Figure 5.10b show that the image in Figure 5.5 represented
in different bases have approximately the same amount of non-zero coefficients.
Note that when reconstructing this image the results were more similar, between
bases, than when reconstructing Figure 5.7b.
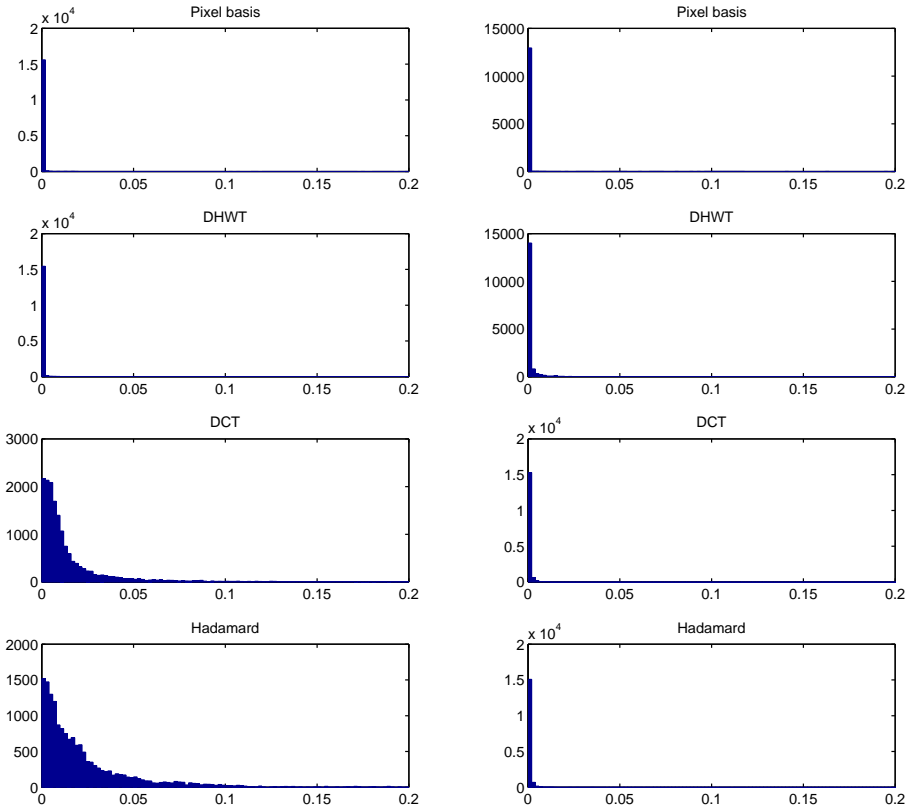
(a) *Normalized histogram of Figure 5.7 represented in different bases.*

(b) *Normalized histogram of Figure 5.5 in different bases.*

Figure 5.10: *The longitudinal axes shows the number of values of a certain image value (the horizontal axis). The entire interval of values $(0 - 1)$ is not included.*

5.2 Signal strength and noise

To evaluate different noise levels in the measurements, signal strengths were varied in five different simulations. The Poisson noise has a bigger impact on small values, because the standard deviation of the Poisson noise is the square root of the expected value. The amount of dark counts is dependent on the number of measurements and will have a bigger impact on small signal values. In these simulations there are about 50 dark counts per intensity image.

Five different signal strengths were evaluated, each case reconstructing an intensity image, using the DCT or DHWT as sparse basis or the TV minimization. The different signal levels are found in Table 5.1. The reconstructions were performed for different undersampling ratios. Results are shown in Figure 5.11.
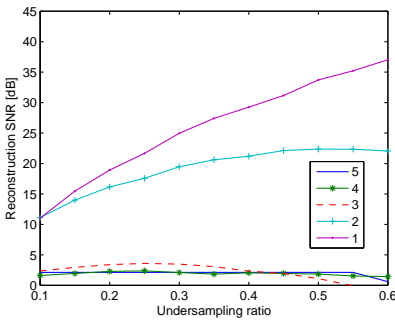
(a) *Reconstruction performance using DCT as sparse basis.*
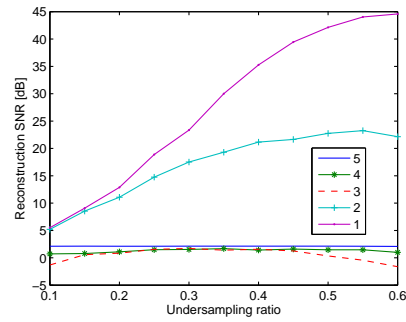
(b) *Reconstruction performance using DHWT as sparse basis.*
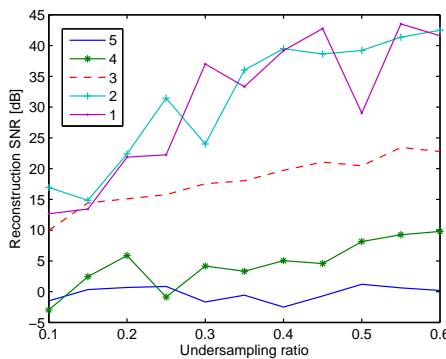
(c) Reconstruction performance using TV.

Figure 5.11: *The reconstruction performance for the five different signal strengths, according to Table 5.1. The reconstruction was done using DCT and DHWT as sparse basis and using TV minimization.*

Table 5.1: *Signal cases with their approximate maximum received signal strength. Signal case 2 is the same signal strength as used in Section 5.1.*

Signal	Approx. max photons/pixel	Signal SNR [dB]
1	2000	102
2	160	75
3	10	48
4	2	4
5	0.2	9

The reconstruction SNR drops significantly for the signal cases 3, 4 and 5 when using DCT or DHWT, whereas the TV minimization handles case 3 better. The reconstruction performances in signal case 4 and 5, which has the lowest signal strengths, are almost constant for all three reconstruction methods. The others signal cases have improved reconstruction results for higher undersampling ratios. In Figure 5.13 reconstructed images are shown. Note that even with reconstruction SNR close to 0, as in the case of Figure 5.13b and Figure 5.13d, the boats are visible, although not clearly.

The images reconstructed from Signal case 3 are very noisy and the question is if noisy images like those can reconstruct accurate depth maps. The most important thing is that high intensities are at correct positions in the reconstructed images, whereas other lower intensities can be seen as traces left from the IRF(which can be thresholded). These two signal cases 2 and 3 are therefore tested for reconstruction of the entire scene (generating depth maps), see reference in Figure 5.12. The results are shown in Figure 5.14.

The results in Figure 5.14 show that both signal cases generate depth maps with different noise levels. In Figure 5.14b and Figure 5.14d it is clear that depth information is missing, specially if considering the sails, and there is many small artifacts. The reconstructed depth map in Figure 5.14a and Figure 5.14c are looking much better with only small range deviations on the boats and very few single pixel artifacts. DCT shows better results, in both signal cases, than the DHWT. TV minimization reconstructs the range image from Signal case 2 almost perfect with nearly no artifacts at all and the range image from Signal case 3 is much better reconstructed than using the other two approaches. TV minimization is clearly best at handling low signal SNR.
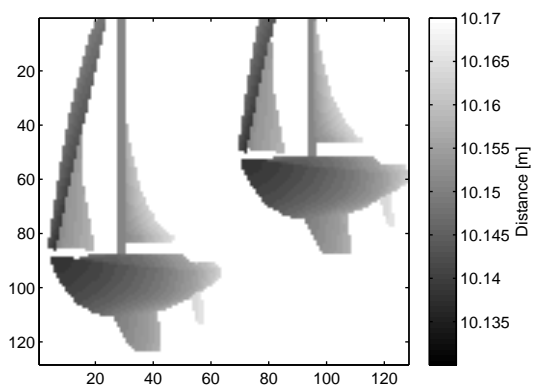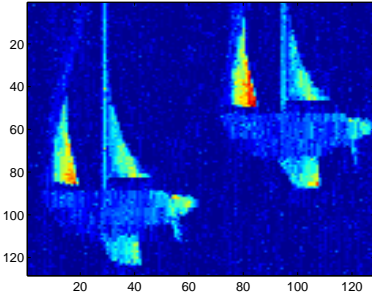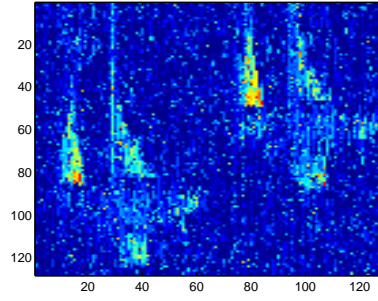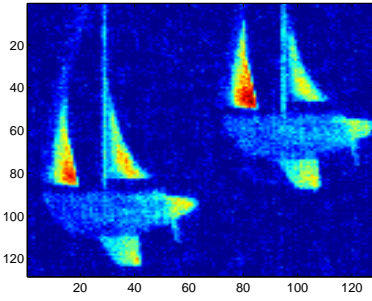
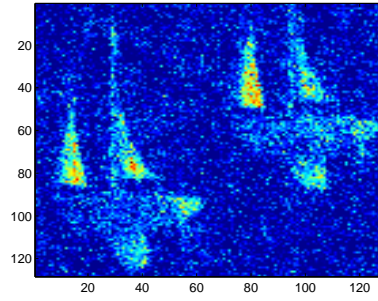**Figure 5.12:** *Reference depth map.*

(a) *Reconstructed intensity image from Signal case 2 using DHWT as sparse basis.*
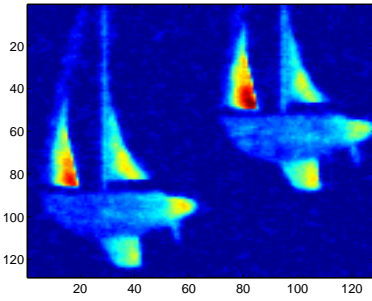
(b) *Reconstructed intensity image from Signal case 3 using DHWT as sparse basis.*
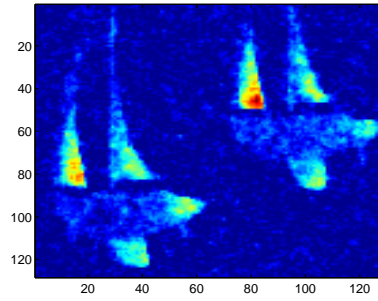
(c) *Reconstructed intensity image from Signal case 2 using DCT as sparse basis.*

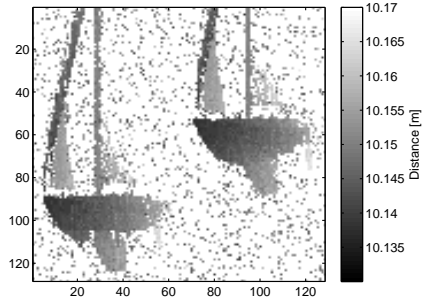**(d)** *Reconstructed intensity image from Signal case 3 using DCT as sparse basis.*

(e) *Reconstructed intensity image from Signal case 2 using TV minimization.*

(f) *Reconstructed intensity image from Signal case 3 using TV minimization.*

Figure 5.13: *Reconstructed intensity images. All images are reconstructed using an undersampling ratio of 0.4. A reference intensity image is shown in Figure 5.9a. The reconstruction SNR of the intensity images are found in Figure 5.11 and their signal strengths in Table 5.1.*
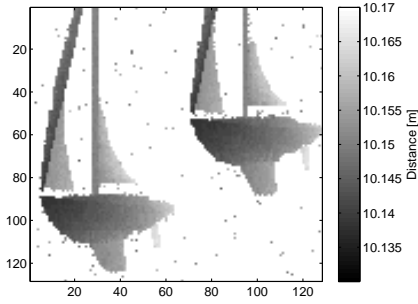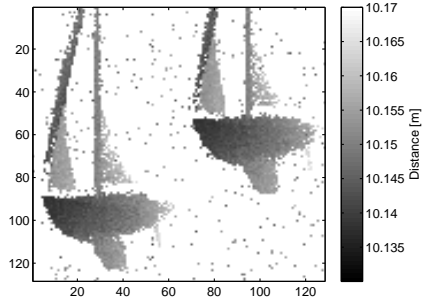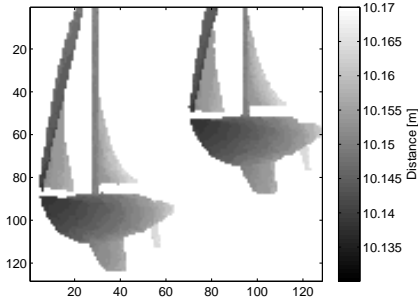
(a) *Reconstructed depth map from Signal case 2 using DHWT.*

(b) *Reconstructed depth map from Signal case 3 using DHWT.*
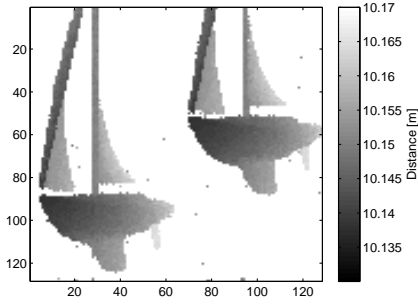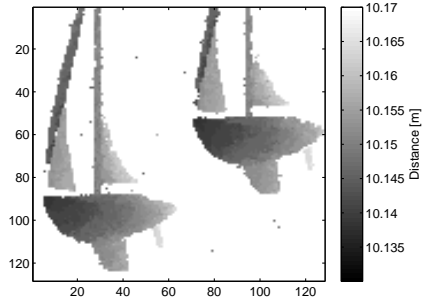

(c) *Reconstructed depth map from Signal case 2 using DCT.*

(d) *Reconstructed depth map from Signal case 3 using DCT.*

(e) *Reconstructed depth map from Signal case 2 using TV minimization.*

**(f)** *Reconstructed depth map from Signal case 3 using TV minimization.*

Figure 5.14: *Four reconstructed depth maps and their reference depth map. The shown depth interval is set to maximize the contrast on the boats. The depth images are reconstructed using DHWT, DCT or TV minimization with an undersampling ratio of* 0.4. *No image filtering, apart from hard thresholding, was used when generating the depth maps from the intensity images.*

5.3   DMD-array choice

The choice of how to form the DMD-arrays have limitations. The array elements can either reflect photons towards the detector or away from it, which mathematically means that its elements can only have the integer values 0 or 1.

When designing the DMD-array patterns, the incoherence with the signal basis representation is important, as discussed in Subsection 3.2.1. In [2] they present the idea of having a noiselet as measurement basis and states that it gives very low coherence with the Haar wavelet basis. Unfortunately, noiselets have complex values and such measurement bases are not implementable for this sensor system.

Uniformly random measurement matrices are incoherent with almost any basis, but an unanswered question is how the randomness itself affect the reconstruction results. In [3], the uniformly random DMD-arrays are chosen with elements being 0 or 1 with equal probability $p = 0.5$. p is the probability for each DMD-array element being 1, i.e. the ratio of mirror elements which reflect light to the detector. By changing this probability the reconstruction results may change. This is studied here and results are shown in Figure 5.15.
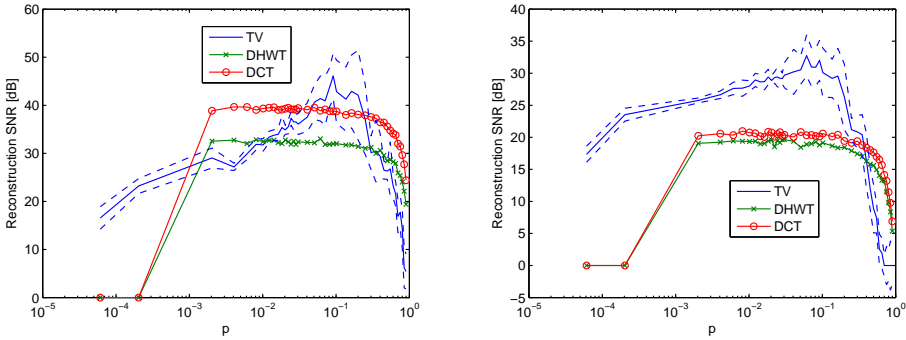
(a) *Reconstructions for Reference scene 2, same slice as in Figure 5.6.*

(b) *Reconstructions for Reference scene 3, same slice as in Figure 5.9.*

Figure 5.15: *Reconstruction performance for different probabilities p of two different intensity images. The undersampling ratio was 0.35. The two blue lines around the TV curve marks the standard deviation of the TV results and the solid line marks the mean of the results (average of 10 evaluations).*

From the results in Figure 5.15 it is clear that it is better to chose a p around 0.1, rather than 0.5 as in [3]. When p is close to zero, the performance drops significantly, which makes it inappropriate to pick a too small p. Patterns generated by very small p, each with around $1 - 3$ numbers of non-zero elements per pattern, were not able to reconstruct the intensity images at all using DCT and DHWT. TV could reconstruct the images with less than 1 (in average over all patterns)

non-zero elements per DMD-array. TV failed when p was larger than 0.55 and performed best around $p = 0.1$. This could be a result of that the tuning of the TV-solver was done for settings which had $p = 0.1$, however the tuning of the solver using DCT and DHWT was definitely not dependent on a specific p. In all other simulations in Chapter 5 p was set to 0.1.

5.4 Optimization methods

This section present results regarding properties of different optimization methods and solvers. Three solvers were used. The solvers SPGL1 and ProPPa utilizes the l_1 norm to solve the CS problem, while TVAL3 uses TV minimization.

5.4.1 Reconstruction speed

The methods where tested by varying the number of allowed iterations, measuring the elapsed times and calculate the corresponding SNR.
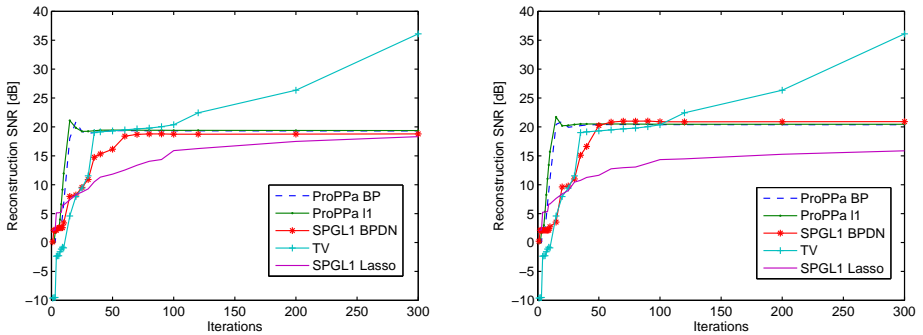
(a) *Reconstruction performance using DHWT, together with the performance using TV.*

(b) *Reconstruction performance using DCT, together with the performance using TV.*

Figure 5.16: *SNR plotted against the number of allowed iterations for DHWT and DCT using different solvers. The performance using TV is also included. The evaluation is performed on the same slice as in Figure 5.9a and the undersampling ratio was set to 0.35.*
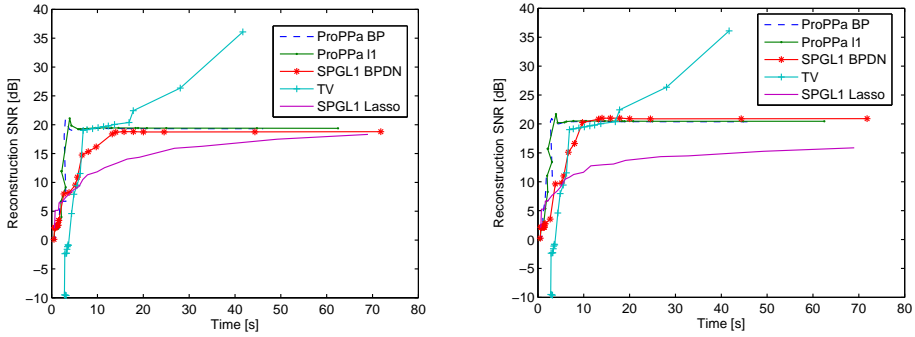
(a) *Reconstruction performance using DHWT, together with the performance using TV.*

(b) *Reconstruction performance using DCT, together with the performance using TV.*

Figure 5.17: *SNR plotted against the reconstruction time for DHWT and DCT using different solvers. The performance using TV is also included. The evaluation is performed on the same slice as in Figure 5.9a and the undersampling ratio was set to* 0.35.
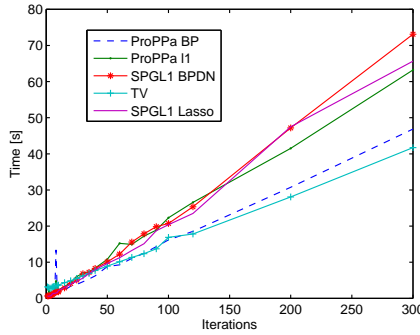

Figure 5.18: *Reconstruction time plotted against the number of allowed iterations.*

In Figure 5.16a and Figure 5.16b it is clear that both ProPPa solvers reaches their maximum reconstruction SNR after very few iterations(≈ 20). The BP solution and the l_1 regression converges to the same SNR and most likely to the same solution. The SPGL1 BPDN solver converges after approximately 60 iterations and reaches the same SNR as the ProPPa solvers. The SPGL1 LASSO solver needs more iterations for convergence in the DHWT case and when using the DCT it does not reach 20 dB, even when using 300 iterations.

The TV minimization solver reached 20 dB after approximately 35 iterations.

Apart from the other solvers it does not converge at 20 dB and reaches a maximum of approximately 36 dB using 300 iterations. Increasing the iterations to more than 300 does not give reconstruction SNR above 36 dB. This is not shown.

All solvers' reconstruction times, shown in Figure 5.18, increase linearly with the number of allowed iterations, where the ProPPa BP solver and the TV solver demands least time per iteration. TV performs best if allowing it to do 300 iterations, which takes about 40 seconds (Figure 5.17), where it reaches an unmatched SNR level. If an SNR of 20 dB is good enough the ProPPa solvers are the fastest ones and requires approximately 20 iterations, which only takes a couple of seconds per reconstruction.

5.4.2   Reconstruction quality

To determine how much faster the acquisition of a scene can be done, different scenes need to be reconstructed for different undersampling ratios. The undersampling ratio is directly related to how much faster the measuring of the scene is done, compared to a full scan. An undersampling ratio of 0.2 equals a measuring time of only 20 % of a full scan.
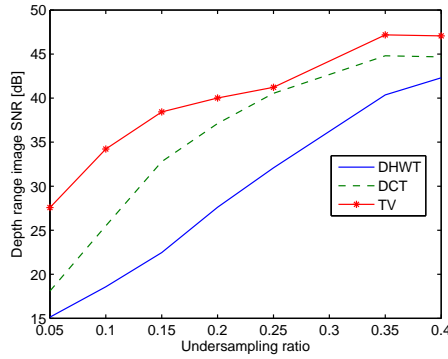
The SNR of the reconstructed depth images are calculated as in (4.4), but with the value 10.05 m subtracted from them, since that is the start of the range interval.
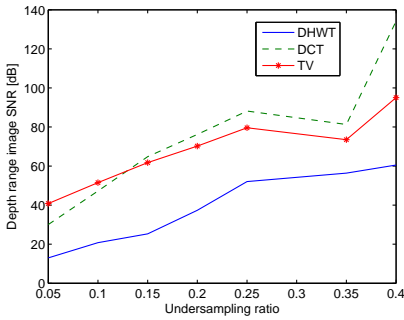
**(a)** *Range image reconstruction performance of Reference scene 1.*
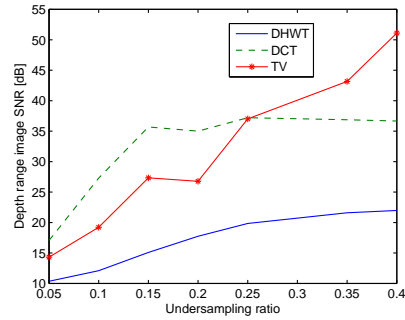
(b) *Range image reconstruction performance of Reference scene 2.*


(c) *Range image reconstruction performance of Reference scene 3.*

Figure 5.19: *Range image reconstruction performance for different levels of undersampling ratio. The three reference scenes are reconstructed using TV, DHWT or DCT.*

The results in Figure 5.19 show that the best reconstruction is achieved when using either DCT or TV minimization. DHWT is worse than both the other methods for all undersampling ratios. According to the SNR, the difference between using DCT or TV minimization are small. For low undersampling ratios, where reconstruction artifacts are present, the characteristics of the noise are important. The noise should preferably not disable operations such as object characterization. In Figure 5.20 range images are shown for low undersampling ratios, which gives an intuition of the noise characteristics.
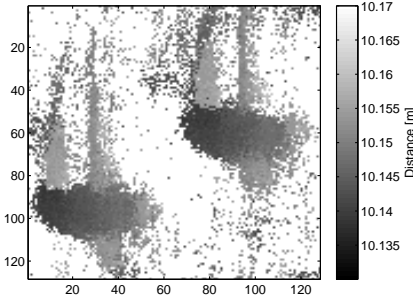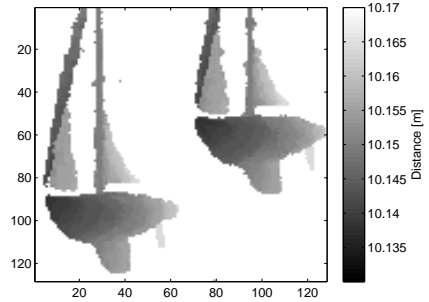
(a) *Reconstructed range image using DCT with an undersampling ratio of 0.05.*
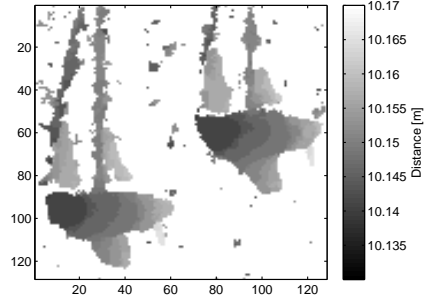
(b) *Reconstructed range image using TV minimization with an undersampling ratio of 0.05.*
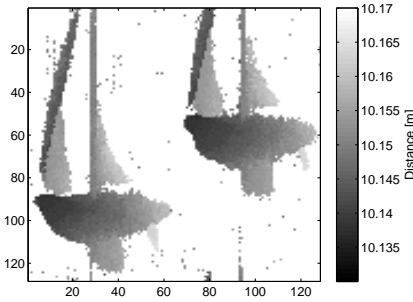
(c) *Reconstructed range image using DCT with an undersampling ratio of 0.15.*

(d) *Reconstructed range image using TV minimization with an undersampling ratio of 0.15.*

**Figure 5.20:** *Depth range images of Reference scene 3 for different levels of undersampling ratio.*

In Figure 5.20a and Figure 5.20b the type of noise differ a lot. When using the DCT approach there is more noise around the boats, compared to the TV case, and the boats' shapes are not that clear. The TV minimization results in a discontinuous look of the ranges on the boats and the artifacts are more in clusters.

Using an undersampling ratio of 0.15 gives range images with small amount of noise, which is shown in Figure 5.20c and Figure 5.20d. Using DCT gives somewhat higher noise levels on and around the boats than TV minimization.

5.4.3   Occlusion

An important property in target recognition is how well the reconstruction handles occlusion, e.g. when scanning for vehicles hidden behind vegetation. TV minimization and l_1 minimization using DCT and DHWT were tested for reconstruction of a scene occluded by a checkered pattern. The checkered pattern

increases the TV of the intensity images by a mean of 273 %. The range image reconstruction results are shown in Figure 5.21c and two reconstructed example images are shown in Figure 5.21a and Figure 5.21b.
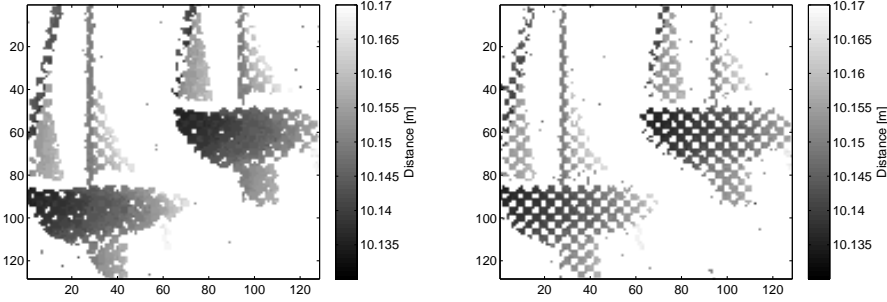
(a) Reconstructed range image using TV minimization with an undersampling ratio of 0.25.

*(b) Reconstructed range image using* l_1 *minimization with DCT and an undersampling ratio of* 0.25.
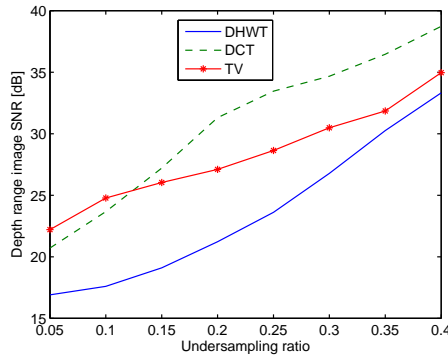
(c) Range image reconstruction performance.

Figure 5.21: *Results showing reconstruction performance using TV minimization,* l_1 *minimization using DCT and DHWT. The SNR is calculated with a reference range image, with the same checkered pattern.*

The SNR curves for TV and DCT develop differently for increasing undersampling ratio. TV minimization has worse performance compared to DCT, according to the SNR. DHWT performed worst of all, but did not fill the holes. The difference between the reconstruction results in Figure 5.21a and Figure 5.21b are that TV minimization fills some of the holes created by the checkered pattern. Increasing the undersampling ratio causes less holes to be filled by the TV minimization.

5.4.4 Reconstructing trees

An attempt to simulate a more natural scene, than the three reference scenes, was done using three tree models. The scene is illustrated in Figure 5.22. The trees were placed at different distances and reconstructed using TV minimization or l_1 minimization with DCT and DHWT. The reconstruction results are shown in Figure 5.23.
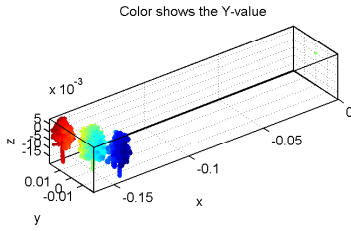

(a) The simulated scene with the three trees. x, y and z are in meters and the detector is placed at $(x, y, z) = (0, 0, 0)$.
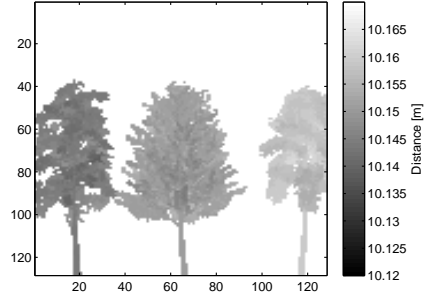
(b) Reference range image.

Figure 5.22: *The simulated scene with the three trees and the corresponding reference range image.*

TV minimization reconstructs the trees scene best. All three methods have problems reconstructing the small details of the trees, like leaves and branches, even for higher undersampling ratios. DCT performed better than DHWT and both methods again show the same reconstruction noise characteristics as before. The reconstructed range images, for higher under sampling ratios, are not included in this report.
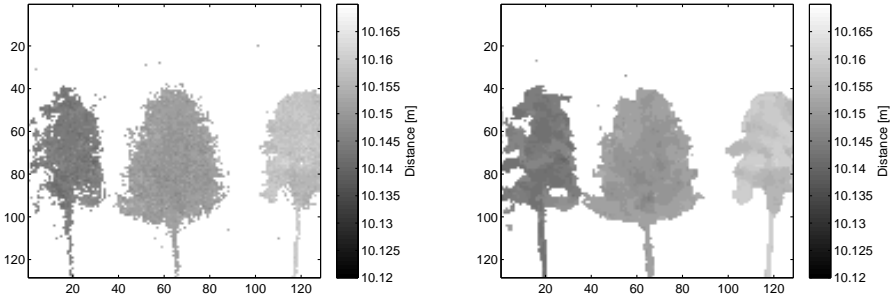
(a) *Reconstructed range image using DCT with an undersampling ratio of 0.15.*

(b) *Reconstructed range image using TV minimization with an undersampling ratio of 0.15.*
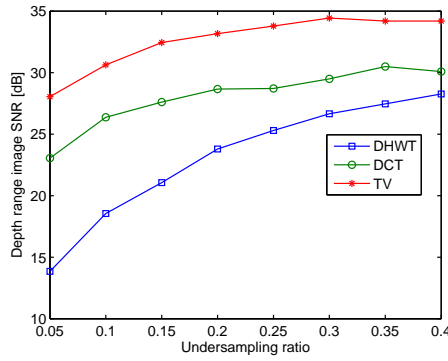
(c) *Depth range image reconstruction performance.*

Figure 5.23: *Reconstructed range images of the scene Figure 5.22 and the reconstruction results for different undersampling ratios.*

5.5   Other 3D-reconstruction methods

In all results in earlier sections the 3D-range data was reconstructed using intensity images corresponding to certain bins. The fact that the data is all part of a common 3D-volume is not utilized in the CS reconstruction. Here different approaches are presented and studied.

By considering the entire 3D-volume as one signal, which could be more sparse than the individual intensity images, the reconstruction results could be improved. This approach has an obvious problem, the dimension of the 3D-volume signal is very big. Tests were performed by folding out four images and reconstructing them together as one signal. MATLAB® could not handle the dimension increase, hence giving no reconstruction results.

Another approach utilizes that intensity images for neighboring bins are similar and hence can be iteratively reconstructed the difference between intensity image slices. This is modeled according to

$$y^k - y^{k-1} = A(x^k - x^{k-1}) = Ax^{\triangle_k},\qquad(5.1)$$

where x^k is the intensity image corresponding to bin k. Each intensity image is given by $x^k = x^{\triangle_k} + x^{k-1}$. These differences, $x^{\triangle_k}$, are possibly more sparse than the intensity images themselves. The problem with this method is that it requires interlacing images(normally reconstructed). The noise from each reconstruction is not removed in each iterative steps, which means that noise propagates. Figure 5.24 shows reconstruction of the tree scene using no interlacing images, interlacing images every fourth x^k, and normal reconstruction. The result in Figure 5.24a shows that more noise is present using this method compared to the normal method shown in Figure 5.24c. Using an interlacing image every fourth reconstruction, see Figure 5.24b, has similar results as the normal method. In this test there was 400 bins, which means that about 300 differences and 100 normal intensity images were reconstructed.
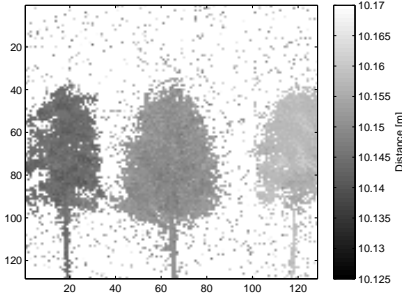
(a) *Reconstructed range image using no interlacing images. l_1 minimization with DCT was used with an undersampling ratio of* 0.3.
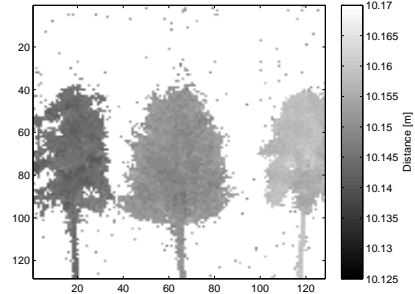
(b) *Reconstructed range image using interlacing images every fourth reconstruction. l_1 minimization with DCT was used with an undersampling ratio of* 0.3.
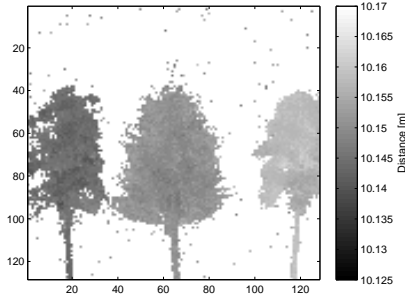
(c) *Reconstructed range image using the normal method. l_1 minimization with DCT was used with an undersampling ratio of* 0.3.

Figure 5.24: *Reconstructed range images of the scene Figure 5.22 using the method in (5.1).*

6

Discussion

In this chapter the results are discussed and related to the theory. Then the analysis method is discussed, especially the reliability and replicability of the chosen methods.

6.1 Results

6.1.1 Sparse signal representation

The results show that the DCT gives the best reconstruction results of the different basis representations when using the l_1 norm minimization. The DHWT did not perform as well as the DCT, but maybe other wavelets will perform better. Sparsity in a basis is dependent on the content of the image. The simulated images may not look like real images at all, which makes the simulation results a bit uncertain.

The sparsity and reconstruction result are closely related, which Figure 5.10a and Figure 5.10b shows. The image in Figure 5.7 were better reconstructed using the two basis representations (pixel basis and DHWT) which clearly represents the image more sparse. These results are two examples which happens to correlate nicely to the theory of CS and sparseness. Generating the same sort of histogram plots from the other images gave equal correspondences between sparseness and reconstruction quality, but are not included in the report.

TV minimization outperforms all different basis representations, which means that the total variation is the best sparsity promoter for the simulated intensity images. A weakness regarding TV is the results shown in Section 5.4.3, where TV minimization proves to have decreased performance for increased variation

in the intensity images.

6.1.2 Signal strength and noise

The best results, this time regarding noise, are again received using TV minimization, which handles low signal SNR better than DCT and DHWT. TV minimization seems to give more stepwise blurred type of noise while the two others produces spiky noise in areas where there should not be any intensities. TV minimization does not produce the spiky noise, which is a great advantage for TV when reconstruction the range images. High valued spikes are not removed by thresholding.

6.1.3 DMD-array choice

The results for designing the DMD suggest to pick a probability p, of a DMD-element being 1, to about $0.05 - 0.1$. This is quite interesting. The CS theory only states that a random sensing matrix will be incoherent with any basis representation and therefore work, not how the randomness itself effect the result.

When the reconstruction does not work, for very small p, it is probably because of that the DMD-arrays need to, at least once, sample information (DMD-element = 1) from every pixel. If p is low enough, e.g. $\frac{1}{n}$, where n is the number of pixels, the $m < n$ DMD-arrays will not reflect photons from all pixels. This means that there would be pixels where no information was sampled and therefore resulting in bad results in the reconstruction.

For large p the reconstruction also fails. Having a large p means that when sampling information with a DMD-pattern, the certainty about where the information comes from will be smaller than for a smaller valued p. Larger p could also mean larger coherence between measurement matrix and basis matrix, since the different DMD-patterns become more alike.

If $p = 0.5$, we know that the produced scalar value comes from a combination of 50 % of the pixels. Some of these pixels does not add intensity to the scalar value while some do and its impossible to know which does and which does not (with only one measurement). If the scalar value turns out to be e.g. 0, we know that all pixels corresponding to the DMD-elements which are one (reflecting to the sensor), should equal 0. By lowering p to e.g. $2/n$, where n is the number of pixels, the scalar value will be produced by an average of two mirror elements per DMD-pattern. The scalar value produced now will more likely be 0 than for larger p which means that if so, two pixel elements can be connected to the value 0. Likewise if the scalar value is e.g. 1 there are only two possibilities; either one of the pixel elements is 1 and the other 0 or vice versa. Imagine getting the scalar value 1 but having $p = 4/n$. This means that in average there are four active mirror elements and therefore four combinations of pixels giving the scalar value 1 for that specific pattern. This means that more measurements (more patterns), are needed to receive more certainty of the information for larger p.

The best choice for p proved to be approximately $0.05 - 0.1$, which can be seen as

some kind of equilibrium where there is a high probability of always getting information from all pixels while not loosing track of where the information comes from.

6.1.4   Optimization methods

The speed of the reconstruction is somewhat important but the measuring time is considered much more important and costly. All calculations where done in MATLAB® which may be slower compared to reconstruction schemes implemented in e.g. C++ and GPU [1]. All three solvers used in this thesis, together with a few others, fulfilled the criteria of reasonable computation times.

The solver ProPPa was used when evaluating different scenarios because the creators of ProPPa claims that their package is faster and better than other well known, fast, optimization packages such as SPGL1, YALL, NESTA and FISTA [13]. Their claims, together with some confirming tests, were the reasons why this package was chosen. SPGL1 and ProPPa were tested against each other to confirm these statements.

ProPPa proved to be very fast, but TV minimization using the TVAL3 software package gave better reconstruction results and better looking range images overall. TVAL3 required more time to converge to its optimal solutions which is its only drawback. If neglecting the speed of the reconstruction TV minimization handles lower undersampling ratios, lower signal SNR and creates less noise in the reconstructed signals. In Figure 5.20 TV minimization shows better performance and produces images more suitable for automatic detection and classification. Figure 5.20a has noise that interferes with the shapes of the boats. This noise would make it, with no prior knowledge about the objects, almost impossible to classify the boats for a human observer or image processing algorithm. However TV minimization performs better and those results show that accurate range images can be obtained using undersampling ratios somewhere between 0.05 and 0.15.

The question is if the scenes used to evaluate these optimization methods suit TV minimization better than real world intensity images. The simulated scenes have quite few edges, which means low total variation, and maybe a scene with more edginess would suit l_1 minimization using DCT better. What contravenes this is the results [15] where TV minimization performs very well on natural scenes with e.g. trees, which have a lot of edges. The results in Section 5.4.4 also show that TV minimization handles trees even better than the other methods. There is a loss of small details on the reconstructed trees, but that is common for all three tested methods.

The results in Section 5.4.3 show that TV minimization does not perform as well as using the l_1 minimization with DCT when the TV increases for the intensity images. In this case the checkered pattern increases the TV by 273 % compared to no checkered pattern. A very interesting result is that the TV minimization fills holes created by the checkered pattern. This is most probably a result of the TV minimization algorithm forcing the solution to have low TV. The question

is if this is a preferred behavior or not. If the scene has natural holes that are considered important information this is a very bad behavior.

6.1.5   Other 3D-reconstruction methods

Reconstructing each intensity image slice separately and then building the range image is not the only method. Reconstruction differences between neighboring intensity image slices shows promising results. Similar results compared to the normal method are achieved using a interlacing image every forth reconstruction. Without interlacing images noise propagates in the reconstructions. If some intensity images needs to be reconstructed fully the benefit of the method is questionable. The difference between images could be more sparse and hence giving better result, but they become dependent of the reconstruction result from the interlacing images. One benefit could be to let the interlacing images be reconstructed with more iterations than the differences, thus saving reconstruction time. This works only if the reconstructions of the differences require less iterations than the normal intensity images, for equal quality.

Too large signal dimensions limits the approach considering the entire 3D-scene as one signal. Increasing the signal dimension makes the transform matrices too big. As an example, reconstructing a 256×256 image requires $256^2 \times 256^2$ transform matrices, using the problem setup as in Chapter 3. Other approaches using functions handles in the CS solvers could work. Designated hardware could also be a solution for handling larger problems.

6.2   Analysis method

Gathering all the results required at least two different solvers. TV minimization required one software package whereas the l_1 minimization required another. This means that there are uncertainties in the evaluations since the different results could be a consequence of not the methods themselves, but tuning parameters or different skill levels of the programmers behind the solvers. Possibly this means that better results could have been achieved using l_1 minimization with another software package than ProPPa or SPGL1.

Using SNR ratio as a measure gives a mathematical estimation of the quality of reconstruction. The image quality is often closely related to the SNR, lower SNR means worse image quality. Two reconstructed images of the same scene can however have the same SNR while containing different kinds of noises. Examples of this is seen in Figure 5.13b and Figure 5.13d. The SNR are the same for both images, but the images have different types of noise. The results in Figure 5.13d suits range image construction better.

The evaluations could have been done for more varying of settings. This would give more general results. What limited this was that a lot of simulations were very time consuming and required a lot of development time. With more computational power, simulations would have been faster and maybe larger images

could have been tested. Higher resolution images were not evaluated since the signal dimension is $N_1 \cdot N_2$, where N_1 is the width and N_2 the height of the image. This means that a small increase in image size largely increases the problem complexity and computational cost.

Originally the simulated model was supposed to be compared to a real TCSPC lab system setup at FOI. This would most probably have meant some changes in the simulation settings and would of course increased the reliability of the simulated laser data. A real functional system would also have meant that real objects and scenes of nature easily could have been tested. The simulated laser model also lacks in some of its theoretical implementation, where e.g. atmospheric turbulence is not taken into consideration. There is also a question whether the background illumination effects the results or not, since it just adds a constant intensity value to all pixels.

The replicability of the results are dependent on how well similar laser data can be generated. The simulated laser data are based on a software provided by FOI, which is not open to public. The CS implementations, used in the simulations, are all based on open source packages and should therefore be replicable, provided proper laser data.

Related work and other references are all articles or publications with detailed explanations and high reliability. Most articles and publications have many and reliable references to others. The methods are well described theoretically, but most articles lack repeatability in implementation because important information lacks description. For example p for the DMD-arrays are almost never specified. It is not always clear that the specific method will work, or work as good, for other cases or scenes than the specific cases in the articles.

7

Conclusions and Future Work

7.1 Conclusions

A laser model of a TCSPC laser system was developed, which enabled easy simulations. The computer model has the advantage over a real lab system because it can be configured in many ways, which simplifies testing and development of image reconstruction algorithms. The model also gives an great intuition of how to configure the real system for optimal results.

Results show that CS enables laser intensity image acquisition using less measurements than the number of pixels. Accurate range images can be obtained using undersampling ratios somewhere between 0.05 and 0.15, which is shown in Figure 5.20. Neglecting the reconstruction time, the undersampling ratio directly translates to how much faster CS can make the image acquisition. This means that using CS, $85 - 95$ % fewer measurements are required compared to a pixel-by-pixel scanning TCSPC laser system.

Some critical factors have been identified. Too low signal levels results in too low signal SNR, which proves to be critical factors affecting the 3D-reconstruction. Spiky noise are present in the reconstructed intensity images, using l_1 minimization, which also affected the reconstructed range images. For TV minimizations a critical factor is when putting a checkered pattern in front of the scene, which basically is an increase of the total variation of the intensity images.

TV minimization gives best reconstruction results in most cases, but requires more reconstruction time to reach its optimal solutions. Computational time is not considered a significant cost, which makes TV minimization the best reconstruction method. Using l_1 minimization, DCT was the best basis.

Designing the DMD-arrays are not straight forward and assuming that the best p is 0.5 is not correct. Results show that an optimal choice of p is probably in the interval $0.05 - 0.1$. Further evaluations are needed.

There are other methods reconstruction the 3D-data. Increasing the signal dimensions causes computational problems. Reconstructing differences of intensity images works and could speed up reconstruction times.

7.2 Future work

Based on discoveries during this thesis work, the following problem areas have been identified for future work:

- A real lab system should be developed, which would make the simulated model more realistic by enabling model validation.

- More tests of the systems capability of handling noise needs to be tested. This is related to having a working lab system, where the real noise levels easily can be set in the simulated model. Simulations with object aliasing should also be performed.

- The current CS solvers should be implemented in e.g. C++ or on GPU, which will give faster reconstruction performance.

- More studies and theoretical analysis are needed to determine if the best choice of p generally is $0.05 - 0.1$ or if it is scene dependent. More scenarios need testing, including other scene setups for different undersampling ratios.

- Find and test other basis representations, for example other wavelets. Better basis representations could mean that that l_1 minimization suddenly outperforms the TV minimization.

- More studies of how the entire 3D-scene can be reconstructed smarter needs to be done. Instead of reconstructing 2D-slices separately the whole 3D-volume should be considered.

- Deconvolution could increase sparsity in the measured data. The tail of the IRF could be removed in the measurements by deconvolution of the measurements y. This should be investigated further.

Bibliography

[1] J.D. Blanchard and J. Tanner. Gpu accelerated greedy algorithms for compressed sensing. *Mathematical Programming Computation*, 5(3):267–304, 2013. ISSN 18672949. Cited on page 51.

[2] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, March 2008. ISSN 1053-5888. doi: 10.1109/MSP.2007.914731. Cited on pages 11, 12, 13, 14, 17, and 38.

[3] A. Colaco, A. Kirmani, G.A. Howland, J.C. Howell, and V.K. Goyal. Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 96–102, June 2012. doi: 10.1109/CVPR.2012.6247663. Cited on pages 3 and 38.

[4] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Compressed sensing: Theory and applications: Introduction to compressed sensing. *Cambridge University Press*, pages 1–55, 2011. Cited on page 14.

[5] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling: Building simpler, smaller, and less-expensive digital cameras. *IEEE Signal Processing Magazine*, 25(2):83–91, 2008. URL www.scopus.com. Cited on page 3.

[6] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser, 2013. ISBN 9780817649470. Cited on page 13.

[7] J. W. Goodman. *Statistical Optics*, volume 1. Wiley, 2000. Cited on page 8.

[8] Markus Henriksson, Lars Sjöqvist, and Lars Allard. Compressive sensing 3d laser radar. literature study and model experiments. *FOI-D–0561–SE*, 2013. Cited on page 19.

[9] G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell. Photon counting compressive depth mapping. *Opt. Express*, 21(20):23822–23837, Oct 2013. doi: 10.1364/OE.21.023822. URL http://www.opticsexpress.org/abstract.cfm?URI=oe-21-20-23822. Cited on pages 4 and 12.

[10] G.A. Howland, P.B. Dixon, and J.C. Howell. Photon-counting compressive sensing laser radar for 3d imaging. *Applied Optics*, 50(31):5917–5920, 2011. ISSN 00036935. Cited on page 3.

[11] A. Kirmani, A. Colaçdo, F.N.C. Wong, and V.K. Goyal. Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor. *Opt. Express*, 19(22):21485–21507, Oct 2011. doi: 10.1364/OE. 19.021485. URL http://www.opticsexpress.org/abstract.cfm? URI=oe-19-22-21485. Cited on page 3.

[12] A. Kirmani, A. Colaçdo, F.N.C. Wong, and V.K. Goyal. Codac: A compressive depth acquisition camera framework. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, number 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012 - Proceedings, pages 5425–5428, Research Laboratory of Electronics, Massachusetts Institute of Technology, 2012. Cited on page 3.

[13] Ranch Y. Q. Lai and Pong C. Yuen. Proppa: A fast algorithm for l1 minimization and low-rank matrix completion. *CoRR*, abs/1205.0088, 2012. Cited on page 51.

[14] C Li. Compressive sensing for 3d data processing tasks: Applications, models and algorithms. 2011. URL http://www.caam.rice.edu/ ~optimization/L1/TVAL3/. Cited on pages 15 and 22.

[15] L. Li, L. Wu, X. Wang, and E. Dang. Gated viewing laser imaging with compressive sensing. *Appl. Opt.*, 51(14):2706–2712, May 2012. doi: 10. 1364/AO.51.002706. URL http://ao.osa.org/abstract.cfm?URI= ao-51-14-2706. Cited on pages 3 and 51.

[16] F. Naizhang, S. Mingjian, and M. Liyong. A compressive sensing photoacoustic imaging system based on digital micromirror device. *International Proceedings of Computer Science and Information Technology*, 48:23, 2012. ISSN 2010460X. Cited on page 4.

[17] T. Neimert-Andersson. 3d imaging using time-correlated single photon counting. Master's thesis, University of Uppsala, 2010. Cited on pages 5 and 6.

[18] R. D. Richmond and S. C. Cain. *Direct-detection LADAR Systems*. SPIE, 2010. ISBN 978-0-8194-8072-9. Cited on page 8.

[19] L. Sjöqvist, M. Henriksson, P. Jonsson, and O. Steinvall. Time-of-flight range profiling using time-correlated single-photon counting. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 6738, Dept. of Optronic Systems, Swedish Defence Research Agency, FOI, 2007. URL https://lt.ltag.bibl.liu.se/login?url=http: //search.ebscohost.com/login.aspx?direct=true&db=

`edselc&AN=edselc.2-52.0-42149159479&site=eds-live`. Cited on pages 5, 6, and 7.

[20] O. Steinvall, L. Sjöqvist, and M. Henriksson. Photon counting ladar work at foi, sweden. *Proc. SPIE*, 8375:83750C–83750C–14, 2012. doi: 10.1117/12. 920294. Cited on page 6.

[21] E. Van Den Berg and M. Friedlander. Sparse optimization with least-squares constraints. *SIAM Journal on Optimization*, 21(4):1201–1229, 2011. doi: 10. 1137/100785028. URL `http://epubs.siam.org/doi/abs/10.1137/` `100785028`. Not cited.

[22] E. Van Den Berg and M. P. Friedlander. SPGL1: A solver for large-scale sparse reconstruction, June 2007. http://www.cs.ubc.ca/labs/scl/spgl1. Cited on page 22.

[23] P. Ye, J.L. Paredes, Y. Wu, C. Chen, G.R. Arce, and D.W. Prather. Compressive confocal microscopy: 3d reconstruction algorithms. In *Proceedings of the SPIE - The International Society for Optical Engineering*, volume 7210, page (12 pp.), University of Delaware, Department of Electrical and Computer Engineering, Newark, DE, 19716, USA, 2009. Cited on page 4.