

Master of Science Thesis in Electrical Engineering
Department of Electrical Engineering, Linköping University, 2018

Compressive Sensing: Single Pixel SWIR Imaging of Natural Scenes

Andreas Brorsson



Master of Science Thesis in Electrical Engineering
Compressive Sensing: Single Pixel SWIR Imaging of Natural Scenes

Andreas Brorsson

LiTH-ISY-EX--18/5108--SE

Supervisor: **Doktorand Mikael Persson**
 ISY, Linköpings universitet
 Doktor David Bergström
 FOI
 Doktor David Gustafsson
 FOI
 Master of Science Carl Brännlund
 FOI

Examiner: **Assistant professor Maria Magnusson**
 ISY, Linköpings universitet

*Computer Vision Laboratory
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden*

Copyright © 2018 Andreas Brorsson

Sammanfattning

Foton tagna i det korta infraröda spektrumet är intressanta i militära sammanhang på grund av att de är mindre beroende av vilken tid på dygnet de är tagna för att solen, månen, stjärnor och nattsken (night glow) lyser upp jorden med kortvågiga infraröd strålning dynget runt. Ett stort problem med dagens kortvågig infraröda kameror är att de är väldigt dyra att producera och därav inte tillgängliga till en bred skara, varken militärt eller civilt. Med hjälp av en relativt ny teknik kallad *compressive sensing* (CS) möjligörs en ny typ av kamera med endast en pixel i sensorn. Denna nya typ av kamera behöver bara en bråkdel mätningar relativt antal pixlar som ska återskapas och reducerar kostnaden på en kortvågig infraröd kamera med en faktor 20. Kameran använder en mikrospegelmatris som används för att välja vilka speglar (pixlar) som ska mätas i scenen och på så sätt skapa ett underbestämt ekvationssystem som kan lösas med teknikerna beskrivna i CS för att återskapa bilden. Givet den nya tekniken är det i Totalförsvarets forskningsinstituts (FOI) intresse att utvärdera potentialen hos en enpixel-kamera. Med en enpixel-kameraarkitektur utvecklad av FOI var målet med detta examensarbete att ta fram metoder för att sampla, återskapa bilder och utvärdera deras kvalitet. Detta examensarbete visar att användning av strukturella slumpade matriser och snabba transformenter öppnar upp för högupplösta bilder och snabbar upp processen att rekonstruera bilder avsevärt. Utvärderingen av bilderna kunde göras med vanliga mått associerade med kamerautvärdering och visade att kameran kan återskapa högupplösta bilder med relativt hög bildkvalitet i dagsljus. [15]

Abstract

Photos captured in the shortwave infrared (SWIR) spectrum are interesting in military applications because they are independent of what time of day the picture is captured because the sun, moon, stars and night glow illuminate the earth with short-wave infrared radiation constantly. A major problem with today's SWIR cameras is that they are very expensive to produce and hence not broadly available either within the military or to civilians. Using a relatively new technology called compressive sensing (CS), enables a new type of camera with only a single pixel sensor in the sensor (a SPC). This new type of camera only needs a fraction of measurements relative to the number of pixels to be reconstructed and reduces the cost of a short-wave infrared camera with a factor of 20. The camera uses a micromirror array (DMD) to select which mirrors (pixels) to be measured in the scene, thus creating an underdetermined linear equation system that can be solved using the techniques described in CS to reconstruct the image. Given the new technology, it is in the Swedish Defence Research Agency (FOI) interest to evaluate the potential of a single pixel camera. With a SPC architecture developed by FOI, the goal of this thesis was to develop methods for sampling, reconstructing images and evaluating their quality. This thesis shows that structured random matrices and fast transforms have to be used to enable high resolution images and speed up the process of reconstructing images significantly. The evaluation of the images could be done with standard measurements associated with camera evaluation and showed that the camera can reproduce high resolution images with relative high image quality in daylight.

Acknowledgments

HELLO

*Linköping, Januari 2018
Andreas Brorsson*

Contents

Notation	xi
1 Introduction	1
1.1 Background	1
1.2 Compressive sensing and imaging	2
1.3 Motivation	3
1.4 Aim	4
1.5 Research questions	4
1.6 Limitations	4
1.7 Thesis outline	4
2 Related work	5
2.1 Compressive sensing	5
2.2 Compressive imaging	5
2.3 Measurement matrix & reconstruction	6
2.4 Evaluation	7
3 Method	9
3.1 Single pixel camera architecture & hardware	9
3.1.1 Newtonian telescope	10
3.1.2 DLP and DMD	11
3.1.3 Lens	12
3.1.4 Single pixel sensor	13
3.1.5 Signal spectrum	13
3.2 Compressive imaging	13
3.3 Measurement matrix and Restricted isometry property (RIP)	15
3.3.1 Permutated sequency ordered Walsh Hadamard measurement matrix	15
3.4 Reconstruction method	17
3.4.1 Total variation: TVAL3	18
3.5 Image capturing and processing chain	19
3.5.1 Prepare the SPC	20
3.5.2 Sampling	20

3.5.3	Streaming patterns to the DMD	21
3.5.4	Signal processing	22
3.5.5	Dynamics in scene	23
3.5.6	Reconstruction	25
3.5.7	Image processing	26
3.6	Evaluation: Image quality assessment	26
3.6.1	Evaluation Using reference image	27
3.6.2	Evaluation Using no reference quality assessment	28
3.6.3	Evaluating Using Edge response	28
3.7	Method criticism	29
4	Results and Evaluation	31
4.1	Simulated Results	31
4.1.1	Reconstruction performance Using reference image	32
4.1.2	Reconstruction performance using no reference quality assessment	35
4.1.3	Dynamics in scene	36
4.2	SPC evaluation	43
4.2.1	Reconstructed performance using reference image	43
4.2.2	Reconstruction performance using no reference quality assessment	45
4.2.3	Luminance change in scene	48
4.2.4	Edge response	50
4.2.5	Subsampling ratio	53
5	Discussion	59
5.1	Results	59
5.1.1	Reconstruction performance using reference image	59
5.1.2	Reconstruction performance using no reference quality assessment	60
5.1.3	Dynamics in scene	62
5.1.4	Edge response	63
5.1.5	Subsampling ratio	64
5.2	Method	65
5.2.1	Replicability, reliability and validity	66
6	Conclusion & Future Work	67
6.1	Conclusion	67
6.2	Future work	69
Bibliography		71

Notation

NOMENCLATURE

Notation	Meaning
y	Measured signal
Φ	Measurement matrix
x	The spatial scene
Ψ	Basis matrix
θ	Coefficients in basis Ψ
A	Sensing matrix in new basis
N	Number of reconstructed pixels
M	Number of single pixel measurements
ϵ	Noise from measurements

ABBREVIATIONS

Abbreviations	Meaning
BRISQUE	Blind/referenceless image spatial quality evaluator
CS	Compressive sensing
CI	Compressive imaging
DLP	Digital light processor
DMD	Digital micromirror device
IID	Independent and identically distributed
FOI	Swedish Defence Research Agency
FWHT	Fast Walsh-Hadamard transform
PSNR	Peak signal-to-noise ratio
RIP	Restricted isometry property
SLM	Spatial Light Modulator
SPC	Single pixel camera
SRM	Structurally random matrix
SSIM	Structural similarity
SWIR	Short-wavelength infrared
TV	Total variation

1

Introduction

The development and research of compressed sensing applied to a single pixel camera (SPC) is a relatively new area in signal processing with the first functioning camera architecture in 2006. Since then numerous improvements and methods have been proposed on how to capture images. In this section an introduction to the SPC architecture and a brief introduction to compressed imaging (CI) is presented followed by the aim, research questions and thesis outline.

1.1 Background

Compressed sensing (CS) allows reconstruction of a sparse signal being sampled with far fewer samples required to fulfill the sampling theorem. Swedish Defence Research Agency (FOI) became interested in the subject some years ago and tests potential applications. One of the potential applications are a camera with a single pixel which can reconstruct a scene, therefore FOI built a SPC platform in the short-wave infrared (SWIR) spectrum for the purpose to study and evaluate this kind of system.

The SWIR spectrum is electromagnetic radiation with wavelengths between 700 - 2500 nm and SWIR cameras can therefore capture images illuminated by the sun, moon, starlight and airglow that works both at day and night. SWIR light can to some extent pass through smoke and fog which makes it a robust camera for day and night applications. Some camouflage that is hard to spot in visual spectrum is visible in the SWIR spectrum. The system used in this master's thesis uses a digital micromirror device (DMD) to sample the light from the scene. The system will sample less single pixel measurements than the number of pixels in the reconstructed image with the drawback that it has to capture each

measurement in consecutive order instead of all at the same time.

1.2 Compressive sensing and imaging

Compressive sensing is a new sampling strategy which reconstructs a compressible or sparse signal by finding a solution to an underdetermined linear system where the number of single pixel measurements is less than the number of pixels in the final reconstructed image. Two constraints need to be fulfilled to apply compressed sensing sampling: the sampled image needs to be sparse in some basis, e.g. the Fourier or wavelet base, and the measurement matrix must be incoherent with the sparse transform, meaning that the image needs to be compressible and the selected sampled pixels for each measurement needs to be picked at random with a 50% chance of being included in the measurement.[6]

The characteristic underdetermined linear system in CS is defined as $\mathbf{y} = \Phi\mathbf{x}$, where \mathbf{y} contains the measurements from the measurement matrix Φ sensing the image \mathbf{x} . In figure 1.1 such a linear equation system is shown.

$$\begin{matrix} \mathbf{y} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_M \end{matrix} \end{matrix} = \begin{matrix} \Phi \\ \begin{matrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \\ \vdots \\ \Phi_M \end{matrix} \end{matrix} \begin{matrix} \mathbf{x} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_N \end{matrix} \end{matrix}$$

Figure 1.1: The compressive sensing characteristic underdetermined linear system. The image \mathbf{x} represented as a vector, being sampled with a measurement matrix ϕ_m (one row of the complete measurement matrix Φ) yielding one measurement y_i .

The SPC in this master's thesis was designed with reflecting telescope optics to act as a lens to focus the scene. As seen in figure 1.2 light from the scene enters through the aperture in the camera where the primary mirror focus the light the via the secondary mirror onto the DMD. To this point, the SPC works like a conventional camera with a DMD where the image sensor would be placed in the convectional camera. The DMD resembles an image sensor, but instead of photo diodes for each pixel there is a tiny mirror which corresponds to one pixel and can individually either reflect the incoming light to the single pixel sensor being measured or reflect the light in the other direction without measuring it.

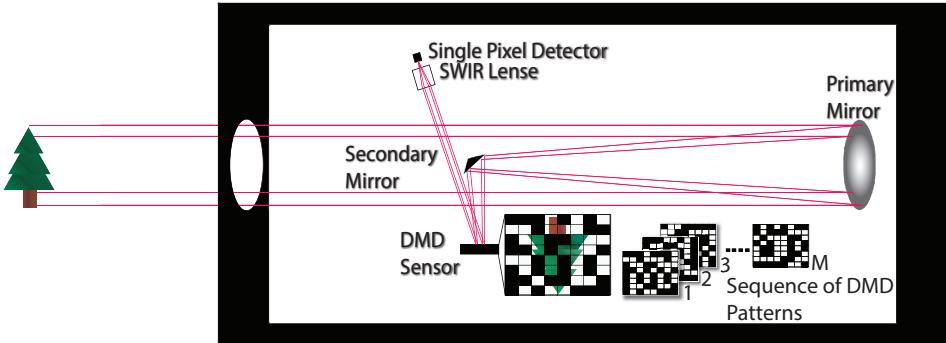


Figure 1.2: SPC system overview.

To connect the CS linear system in figure 1.1 to the single pixel camera each row of the complete measurement matrix is being re-sized to a square matrix of ones and zeros and displayed on the DMD as a pattern, which act as a filter for which pixels being sampled by the single pixel sensor. The next measurement is then performed by the next measurement matrix until all desired measurements are completed, as shown in figure 1.2.

When all the measurements are sampled, an optimization algorithm calculates the most significant coefficients of the image in another more sparse basis, for example the wavelet basis, by transforming the complete measurement matrix. Because the image is more sparse or compressed in another basis, it is easier to find the solution to the equation system in that basis. Hence the name compressive sensing and the reason why the image can be reconstructed using less measurements than number of pixels.

1.3 Motivation

Why would an SPC be beneficial to a conventional camera? The SPC has more components and several measurements have to be made over time, while a regular camera measures all pixels on the sensor at the same time. Moreover the reconstruction shifts burden to the processor. There are two major reasons why an SPC is of interest. The SPC can not compete with the conventional cameras in the visual spectrum where cameras in all price ranges and quality already exist and are relative cheap to build. The focus lies in more exotic spectra like SWIR or Terahertz (X-ray), where the image sensors are hard to build. This brings up cost and the ability to create high resolution sensors. With CS and the SPC architecture, manufacturing cost can be significantly reduced while the image resolution increases. For example, a state of the art SWIR camera cost about half a million SEK. The cost can be reduced by a factor of 20 or higher with an SPC with the same resolution.

1.4 Aim

What image quality can be achieved in natural images captured with a single pixel camera in daylight using state of the art methods?

1.5 Research questions

- How can the quality of images reconstructed by CS or a SPC be evaluated?
- What is the state of the art method to capture and reconstruct images using a SPC architecture?
- What image quality is achieved using state of the art methods applied to the SPC?

1.6 Limitations

- The SPC provided by FOI is used and only minor changes can be made.
- The SPC is stationary at FOI and images can only be captured from the building.
- The reconstruction algorithm will not be developed in this thesis and therefore free to use algorithms needs to be found.

1.7 Thesis outline

In this thesis the most important and inspirational articles will be presented with a small description in section 2 *Related work*. Section 3 *Method* presents a thorough review of the hardware, sensing- and reconstruction-method and the complete image capturing chain including pre- and post-processing. The method section includes essential compressive sensing and imaging theory, this section also presents the evaluation techniques used in the result. Section 4 *Results* is divided into two categories, *simulated results* and *SPC results*, where the same evaluation technique is performed on simulated and SPC images respectively in order to draw conclusions of the different parts of the chain. The results are followed by *Discussion and Conclusion & Future Work* in section 5 and 6 respectively.

2

Related work

In this section important, relevant and fundamental articles to this master's thesis are presented each with a summary. The articles covers compressed sensing theory applied to compressed imaging, SPC architecture and how to evaluate the images i.e. the fundamental source of information on how to build a state of the art SPC system and how to evaluate its performance.

2.1 Compressive sensing

- [9, 17] "Sparse Modeling" by G. Y. Grabarnik and I. Rish and "Sparse and redundant representation" by M Elad is two books which thoroughly presents the topic of sparse and redundant representation and modeling. The fundamental principles and constraints that needs to be fulfilled in CS are described. The books presents different minimization algorithms and how to implement them.
- In [6] by "Compressed sensing" David L. Donoho proposed the framework of compressed sensing and the application of images capturing.
- [16] "Compressive Sensing: From Theory to Applications, A survey" by S. Qaisar et al. 2013, reviews CS background, theory and mathematics and has a extensive survey of reconstruction algorithm and potential CS applications.

2.2 Compressive imaging

- [18, 23] "An architechture for compressive imaging" and "A New compressive imaging camera architecture using Optical-Domain Compression" by

M. B. Wakin, D Takhar, et al. 2006, presents the first single pixel camera architecture using CS to reconstruct the images.

- [19] "Single-Pixel Imaging via Compressive Sampling" by M. F. Duarte et al. 2008, is an introduction and summary to CS and CI including the SPC architecture. This article also compares different scanning methodologies and their conditions.
- [4] "Single Pixel SWIR Imaging using Compressed Sensing" by C. Brännlund and D. Gustafsson, 2016, shows the initial results and proof of concept of the SPC architecture at FOI.
- [13] "A high resolution SWIR camera via compressed sensing" is a paper by L. McMackin et al. 2012 at Inview Technology which develop and produces compressive sensing cameras. The paper contains a brief review of CS and CI followed by a presentation of their camera architecture.
- [11] "Compressed Sensing for 3D Laser Radar" by E. Fall, 2014, is a master's thesis where CS/CI is evaluated for a potential depth camera architecture using a one pixel sensor.
- [8] "Multi image super resolution using compressed sensing" by T. Edler et al. 2011, presents the results from using a small detector array instead of just one single sensor, but still using CS to reconstruct the images. With this technique the subsampling ratio and the exposure time is reduced compared to using a single photo diode.

2.3 Measurement matrix & reconstruction

- [14] Chengbo Li:s master's thesis "An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing" describes a total variation algorithm that Li constructed which solves the CS problem. The algorithm is faster and produces better results for images than previous popular algorithms.
- [1, 12, 22] Fast and Efficient Compressive Sensing Using Structurally Random Matrices (SRM). The articles describes why and how to implement SRM, in these articles the Hadamard or DCT matrices is proposed to replace the random measurement matrix. With SRM the reconstruction time is reduced by replacing matrix multiplication with fast transforms. In addition to improve reconstruction time, the new method does not need to store the measurement matrix in memory, which enables reconstruction of high resolution images.
- [24] "An Improved Hadamard Measurement Matrix Based on Walsh Code For Compressive Sensing" shows that sequency-ordered Walsh Hadamard matrix gives better reconstruction than the Hadamard matrix with the same benefits of using the Hadamard matrix. The resulting reconstructed image has near optimal reconstruction performance.

2.4 Evaluation

- [3] "The essential guide to image processing" by Al Boviks contains the majority of fundamental image processing techniques and measurements. Two image quality metrics of interest is peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) which can be used when a reference image is available.
- [2] "No-Reference Image Quality Assessment in the Spatial Domain" by M. Anish et al. 2012, is the article describing the blind/referenceless image spatial quality evaluator (BRISQUE). The BRISQUE algorithm evaluates image quality and "naturalness" based on statistics in the image. BRISQUE is used when there is no reference image available and therefore can be used to evaluate images produced by the SPC.
- [15] "Prestandamått för sensorssystem" by F. Näsström et al. 2016, describes methods and tools to evaluate sensor systems at FOI.

3

Method

In order to answer the research questions stated in section 1.5, a state of the art SPC, experiments and evaluation methods needs to be set up. In this section all the necessary hardware and software components and theory will be presented as well as the evaluation techniques.

3.1 Single pixel camera architecture & hardware

FOI designed the SWIR SPC platform using a DMD, a Newtonian telescope and a single pixel SWIR detector. The system also has a reference camera in the visual spectrum which can capture images of the scene reflected on the DMD, check that the patterns are displayed correct on the DMD and simplifying focusing of the image.

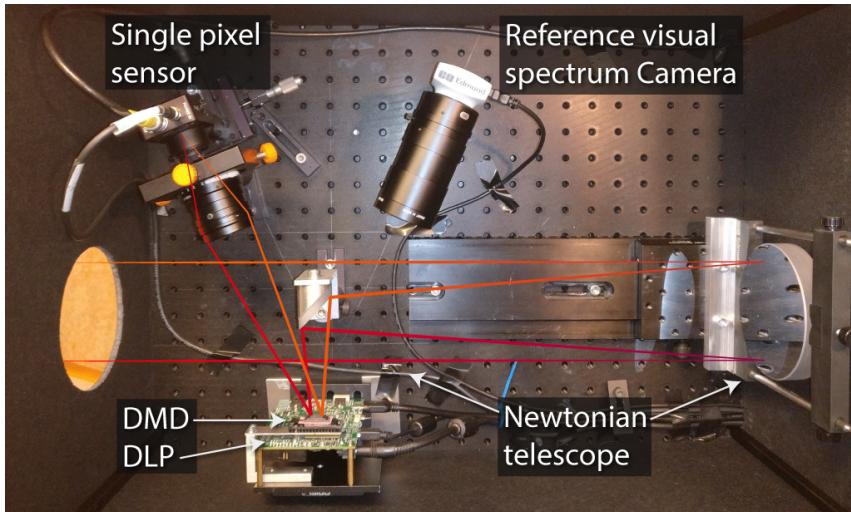


Figure 3.1: The single pixel camera architecture used in this thesis. In the image, the aperture, reflective optics, DMD, reference camera and the single pixel sensor are shown from an areal view including red lines showing the incoming light path.

As seen in figure 3.1, light from the scene is focused by the Newtonian telescope and reflected onto the DMD. The mirrors on the DMD can reflect the light individually either into the single pixel sensor or the reference camera. The DMD acts as a Spatial Light Modulator (SLM) and reflects different patterns which is 'summed up' in the single pixel sensor as a voltage intensity. The reconstructed image from the system will have the same resolution as the DMD patterns. The digital light processor (DLP) is the DMD:s control unit which controls which patterns are displayed on the DMD either by reading images from memory or the HDMI video port.

3.1.1 Newtonian telescope

A Newtonian telescope is a reflecting telescope, using a concave primary mirror and a flat diagonal secondary mirror, see figure 3.1. In this set-up the telescope act as a lens focusing the scene onto the DMD. The motivation to use a Newtonian telescope instead of a lens system is partly that chromatic aberration is eliminated and partly that a reflective optical system works over a greater range of wavelengths that includes SWIR, near infrared (NIR) and the visible spectrum. This design has a very narrow field of view which give high detailed scenes from a great distance.

3.1.2 DLP and DMD

The DMD (Texas Instruments DLP4500NIR) is a matrix of micro mirrors that can be individually tilted $\pm 12^\circ$ and reflects wavelengths in the range 700–2500 nm. The DMD is controlled by the DLP (DLP® LightCrafter™ 4500) which can be controlled either by video port (HDMI) or by the internal flash memory. The internal memory can theoretically be faster than the video port, but due to constraints in both memory and memory bandwidth, the fastest measurement matrix rate gets stuck at 270 – 300 Hz. The video port can be operated at 120 Hz and display one bit plane at the time from a 24 bit signal, which gives a maximum measurement matrix rate $120 \times 24 = 2880$ Hz, but in the current configuration only 60 Hz frame rate was achieved giving a measurement matrix rate at 1440 Hz. At this rate with subsampling ratio (the number of measurements relative number of pixels) between 20% – 30% with 512×512 pixel images, the sampling would be acquired in 36 – 52 seconds. To control the DMD the software "DLP LightCrafter 4500 Control Software" is used.[20]

The DMD used in the setup is constructed with a diamond shaped pattern instead of a regular square grid which is used in regular camera image sensors. The diamond shape causes the index of each mirror to be skewed against what a normal grid would look like. As seen in figure 3.2, the indexes of the mirrors column is two mirror column arrays wide while a row is a single row.

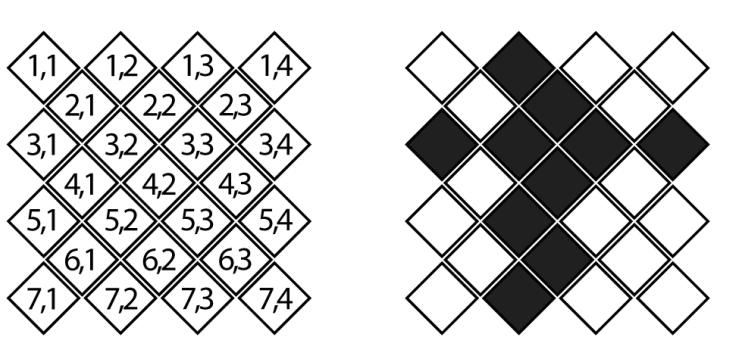


Figure 3.2: DMD matrix mirror index, left shows each tiles index and right shows the second row and second column in black as set by default from the factory.

Because the reconstruction algorithm and measurement matrix needs to be a square matrix with the side length with a power of 2, the resulting images ratio would be 2 to 1, while the image should have the ratio 1 to 1. The resulting image would need to be transformed into the real ratio where information potentially gets lost. Therefore the index of mirrors was changed so that each 'pixel' gets two mirrors as seen in figure 3.3. This will result in rows and columns gets equal amount of space and the aspect ratio will be preserved 1 to 1. By grouping

two mirrors, the amount of light from each "pixel" is doubled and thus should improve the sampled signal quality.

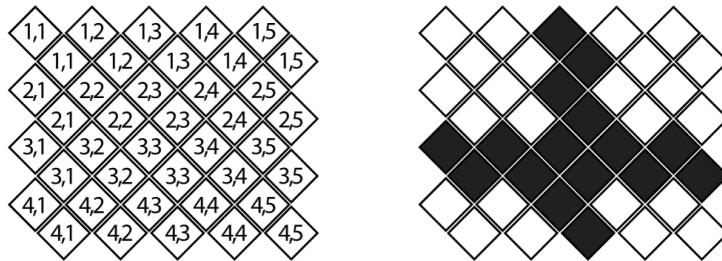


Figure 3.3: DMD matrix, left shows each tiles index and right shows third row and third column in black.

Mathematically the DMD is a binary operator which lets light pass or not, in figure 3.4 a typical pattern that could be sent to the DMD is shown.

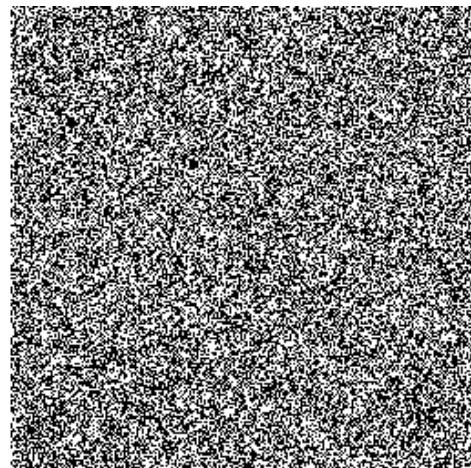


Figure 3.4: A typical pseudo random measurement matrix sent to the DMD with the resolution 256×256 pixels.

3.1.3 Lens

The lens mounted on the single pixel sensor is a 50 mm SWIR Fixed Focal Length Lens with a variable aperture from f-stop f/1.4 designed for wavelengths ranging from 800 nm in the visual spectrum to 2000 nm in the SWIR spectrum. [7]

3.1.4 Single pixel sensor

The single pixel sensor is a Thorlabs PDA20C/M and is sensitive in wavelength range 800-1700 nm which is beyond the visual spectrum (390-700 nm). The sensors built-in amplifier outputs an analog signal in volt which the sampler converts to a discrete value. [21]

3.1.5 Signal spectrum

All components characteristics assembled, the wavelengths that pass through the system and measured in the single pixel sensor is between 800-1700 nm.

3.2 Compressive imaging

Compressive imaging is the name used when sampling and reconstructing images using the compressive sensing method. CI is often realized in form of a SPC but can have different shapes. In this thesis CI is used on the SPC architecture presented in section 3.1. CI exploits the fact that natural images are compressible or approximately sparse in some basis and therefore only a few measurements relative to the image pixel resolution needs to be measured in order to reconstruct the image.

CI need to fulfill two constraints in order to utilize CS sampling, the image needs to be compressible and the complete measurement matrix need to be incoherent with the sparse transform. The first constraint is fulfilled because it is known that natural images are compressible using for example JPEG (using Discrete cosine transform) or JPEG2000 (using wavelet transform) and the second constraint is fulfilled using a measurement matrix with a random characteristic and will be explained further in section 3.3.

The single pixel sensor captures a scene by measuring the light intensity focused onto the detector reflected from the DMD pattern. The DMD pattern changes to obtain new measurements. M measurements are sampled to reconstruct an image with N pixels, where $M \ll N$. Each measurement matrix index is encoded either by a one or a zero (turning the mirror onto or away from the sensor).

The compressive imaging sampling model is defined as

$$\mathbf{y} = \Phi \mathbf{x} + \epsilon, \quad (3.1)$$

where $\mathbf{x}_{N \times 1}$ is the image rearranged as an array with N pixels, $\mathbf{y}_{M \times 1}$ is the sampled signal with M measurements, $\Phi_{M \times N}$ is the complete measurements matrix and ϵ is the noise.

In this thesis Φ is defined as the *complete measurement matrix* and mainly used in mathematical context, the rows of the complete measurement matrix contains the *measurement matrices*, where one measurement matrix is denoted ϕ_m but can also be denoted as *DMD patterns*. The complete measurement matrix thus contains M measurement matrices.

In conventional sampling the number of measurements M needs to be at least equal to the number of pixels N in the image to recover the signal, but CS states that M can be relatively small compared to N given how compressible the image is. This is because the image \mathbf{x} can be represented as

$$\Psi\theta = \mathbf{x}, \quad (3.2)$$

where, $\Psi_{N \times N}$ is some basis matrix and $\theta_{N \times 1}$ is the coefficients where θ is K -sparse. K -sparse means that the image \mathbf{x} has K non zero elements in basis Ψ , $\|\theta\|_0 = K$. Given (3.2), (3.1) can be expanded to

$$\mathbf{y} = \Phi\mathbf{x} + \epsilon = \Phi\Psi\theta + \epsilon = \mathbf{A}\theta + \epsilon, \quad (3.3)$$

where, $\mathbf{A}_{M \times N} = \Phi\Psi$ is called the reconstruction matrix.

The revelation in (3.3) is what makes CS powerful. By sampling the scene using the complete measurement matrix Φ (as (3.1)) but then in the reconstruction process transforming the complete measurement matrix Φ to the reconstruction matrix \mathbf{A} using some basis Ψ , the optimization algorithm can solve the system for the sparse coefficients θ instead of the spatial image coefficients in \mathbf{x} which are not sparse.[17]

A great advantage CI has over regular cameras, where each pixel is sampled separately, is that roughly half the pixels is sampled in one sensor, meaning that background noise of the sensor will be surpassed by the summed intensity of half the pixels making CI very robust to noise.

3.3 Measurement matrix and Restricted isometry property (RIP)

As stated in section 3.2, the complete measurement matrix needs to be incoherent with the sparse transform. In this section the most powerful constraint on a complete measurement matrix is shown, the restricted isometry property (RIP).

In the noiseless case exact recovery of the image \mathbf{x} is achievable if RIP holds for the reconstruction matrix $\Phi \Rightarrow \Phi\mathbf{\Psi} = \mathbf{A}$, the constraint is defined as,

$$(1 - \delta_K) \|\mathbf{x}\|_{\ell_2}^2 \leq \|\mathbf{Ax}\|_{\ell_2}^2 \leq (1 + \delta_K) \|\mathbf{x}\|_{\ell_2}^2, \quad (3.4)$$

where $\delta_K \in [0, 1]$ is the smallest constant to satisfy RIP for a K -sparse signal \mathbf{x} . To determine a sampling matrix is a NP-hard problem (which means that there is no feasible way of creating a optimal reconstruction matrix) and generally \mathbf{x} is not known and varies, which means that there are no general optimal reconstruction matrices for natural images. Therefore, it is desired to find a general reconstruction matrix that satisfies RIP with high probability. It has been proved that constructing the complete measurement matrix by picking independent and identically distributed (i.i.d.) random variables gives $\delta_K \ll 1$ with high probability. Constructing the measurement matrices using i.i.d random variables has showed that the number of measurements M needed to satisfy RIP with high probability is $M \geq O(K \log(N/K)) \ll N$. [9]

The problem of using random matrices is that they need to be stored in memory for the reconstruction algorithm, so when the image resolution is increased the measurement matrix increases exponentially. For images with resolution of 512×512 and larger, the data gets unfeasible for a normal computer to handle.

Fortunately, by changing the complete measurement matrix to structurally random matrices, fast transforms can be used in the reconstruction algorithm instead of vector multiplication, resulting in both faster reconstruction and no need to store the measurement matrix in memory. In this thesis, the permuted sequency ordered Walsh Hadamard measurement matrix (described in section 3.3.1) will be used with the TVAL3 reconstruction algorithm (described in section 3.4.1) to achieve higher resolution photos and faster reconstruction.

3.3.1 Permutated sequency ordered Walsh Hadamard measurement matrix

Besides from eliminating the need to store the measuring matrix in computer memory for reconstruction, the permuted sequency ordered Walsh Hadamard

matrix (PSOWHM) can be generated when sent to the DMD and thus eliminating the need to store the matrix at all. PSOWHM has approximately the same characteristics and properties as an i.i.d. random matrix but generally has a higher number of measurements for exact reconstruction of the image, $M \sim (KNs) \log^2(N)$, where s is the average number of non zero indexes in the measurement matrix [5]. Research has however shown that there is no significant loss in recovery of the image relative the i.i.d. random measurement matrix [24]. An other property of PSOWHM is that it only contains -1 and 1, which can easily be converted to 0 and 1 when sent to the DMD.

In order to construct the PSOWHM, the first step is to define the naturally ordered Hadamard matrix and then follow a few additional steps. The naturally ordered Hadamard matrix of dimension 2^k , $k \in \mathbb{N}$ is constructed by the recursive formula

$$H_0 = 1, \quad (3.5)$$

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (3.6)$$

and in general,

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = H_1 \otimes H_{k-1} \quad (3.7)$$

where \otimes denotes the Kronecker product.

To construct the permuted sequency ordered Walsh Hadamard matrix from the naturally ordered Hadamard matrix these steps are required:

- Convert row index n_H to binary.
- Convert the binary row index to Gray code.
- Apply bit reverse on the Gray code index.
- Order the rows after the bit-reverse to obtain the sequency ordered Walsh Hadamard matrix.

The procedure is illustrated with an example in table 3.1 and equation (3.8).

n_H	0	1	2	3
Binary	00	01	10	11
Gray code	00	01	11	10
Bit-reverse	00	10	11	01
n_W	0	2	3	1

Table 3.1: Example how to convert a naturally ordered Hadamard matrix to a sequency ordered Walsh Hadamard matrix by shifting row with index n_W to n_H .

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \Rightarrow W_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \quad (3.8)$$

To use the sequency ordered Walsh Hadamard matrix as a measurement matrix the first row is omitted, permutations to the columns are performed, M rows are chosen at random and the indices with -1 shifted to 0 . How the matrix are permuted and which rows are chosen in which order are stored so the reconstruction algorithm can use that information to reverse the process. This method is used to distribute the energy of the signal's sample across all measurements. [14, 22, 24].

3.4 Reconstruction method

To reconstruct the image \mathbf{x} , the sparsest set of coefficients in θ is desired. The optimal approach to find these coefficients would be to use ℓ_0 minimization

$$\hat{\theta} = \arg \min \|\theta\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\theta. \quad (3.9)$$

This seems simply to be minimizing nonzero indices in θ in the sparsifying basis Ψ , but this problem is known to be NP-hard. A better approach is the ℓ_1 minimization, for example Basis Pursuit denoise (BPDN),

$$\hat{\theta} = \arg \min \|\theta\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}\theta\|_2 < \epsilon. \quad (3.10)$$

In 2006 Donoho [6] for the first time guaranteed theoretical ℓ_0/ℓ_1 equivalence which holds in the CS case, which means using a ℓ_1 minimizer is guaranteed to find the sparsest solution in polynomial time in the noiseless case which can be approximated in the noisy and compressible signal case. The drawback with the ℓ_1 minimizer is that it requires more measurements than the optimal case with ℓ_0

but $M \ll N$ still holds. Since 2006 many more types of optimization algorithms have evolved which solves the problem with different methods but with the same goal: finding the largest, most significant coefficients of θ . [6, 18, 19]

3.4.1 Total variation: TVAL3

The reconstruction algorithm that was chosen in this thesis was a total variation (TV) regularization algorithm called TVAL3 and was chosen for its speed and good results in image reconstruction compared to other reconstruction algorithms created for the CS problem [14]. Natural images often contains sharp edges and piecewise smooth areas which the TV regularization algorithm is good at preserving. The main difference between TV and other reconstruction algorithms is that TV considers the gradient of signal to be sparse instead of the signal itself, thus finding the sparsest gradient.

The TV optimization problem in TVAL3 is defined as

$$\min_{\mathbf{x}} \sum_i \|D_i \mathbf{x}\|, \text{ subject to } \Phi \mathbf{x} = \mathbf{y}, \mathbf{x} \geq 0, \quad (3.11)$$

where $D_i \mathbf{x}$ is the discrete gradient of \mathbf{x} at position i .

TVAL3 stands for "Total Variation Augmented Lagrangian Alternating Direction Algorithm", where augmented Lagrangian is a method in optimization for solving constrained problems by substituting the original constrained problem with a series of unconstrained subproblems and introducing a penalty term. To solve the new subproblems the alternating direction method is used [14].

As mentioned earlier in section 3.3.1, the main reason to use the permuted sequency ordered Walsh Hadamard matrix is to eliminate the need to store the matrix in computer memory during reconstruction and to speed up the reconstruction. In TVAL3 there are two multiplications between matrix and a vector that dominates the computation time,

$$\Phi \mathbf{x}^k \text{ and } \Phi^\top (\Phi \mathbf{x}^k - \mathbf{y}). \quad (3.12)$$

The idea is to replace the multiplication with fast transforms. To explain the concept some observations and new functions need to be defined. The first observation is that the sequency ordered Walsh Hadamard matrix is a transform matrix, which also can be computed with the fast Walsh Hadamard transform (FWHT),

$$\mathbf{Wx} = \text{FWHT}(\mathbf{x}), \quad (3.13)$$

where \mathbf{W} is a sequency ordered Walsh Hadamard matrix and \mathbf{x} is the image vector. The Walsh Hadamard transform (WHT) is a generalized class of Fourier transforms, which decomposes the input vector into superposition of Walsh functions.

In section 3.3.1 it was briefly mentioned in the last paragraph that the measurement matrix columns are permuted and rows are chosen at random to create the measurement matrix from the sequency ordered Walsh Hadamard matrix. To describe the different permutations two functions are defined.

Definition 3.1. Column permutation operator $\pi(\cdot)$, permutes the order of the columns in a matrix or a vector from a random seed.

Definition 3.2. Subsampling matrix operator $\Pi_M(\cdot)$, chooses M row in a matrix at random and stacks them in a new matrix.

Now the complete measurement matrix Φ can be constructed using the sequency ordered Walsh Hadamard matrix, statement in equation 3.13, definition 3.1 and 3.2,

$$\Phi = \pi(\Pi_M(W)) = \Pi_M(\pi(W)). \quad (3.14)$$

Note that it does not matter in which order the functions are applied, it gives the same result. Also note that multiplication between a matrix and a vector where one of the variables has been permuted by $\pi(\cdot)$, the function can change variable without changing the result since,

$$\pi(\Phi)x = \Phi\pi(x). \quad (3.15)$$

With all observations combined, the matrix multiplication is replaced with the FWHT and operators in definition 3.1 and 3.2 as shown in equation 3.16,

$$y = \Phi x = \pi(\Pi_M(W))x = \Pi_M(W)\pi(x) = \Pi_M(W\pi(x)) = \Pi_M(FWHT(\pi(x))). \quad (3.16)$$

Using this method will reduce the overall computational complexity considerably and it will make the measurement matrix redundant in the reconstruction. Only the two permutation functions $\pi(\cdot)$ and $\Pi_M(\cdot)$ needs to be stored. Eliminating the need for the complete measurement matrix in the reconstruction unlocks the potential to reconstruct images with high resolution (512×512 pixels and larger). [14, 22]

3.5 Image capturing and processing chain

In figure 3.5 the whole process of capturing an image is presented with all subsystems and signal/image processing steps included.

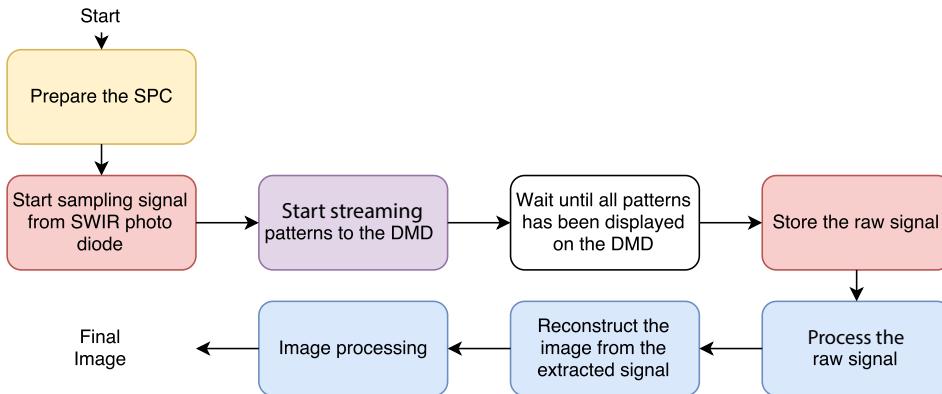


Figure 3.5: Block diagram of image capturing and processing chain, from signal acquisition to final image. Each color represents different subsystems in hardware or software (described in section 3.5.1-3.5.7).

This experimental setup is not a fully automatic system where a button can be pressed and the system produces an image. In the setup the subsystems works completely independently and needs to be operated manually in the right order at the right time. Each color in figure 3.5 represents a subsystem in hardware or software. Each subsystem is described in the following subsections.

3.5.1 Prepare the SPC

The first step in the yellow block "Prepare the SPC" (figure 3.5) is to make sure that the SPC is up and running but also to point the camera at the scene and set the correct focus. The scene is located with the aid of the reference camera (see figure 3.1), with all the mirrors in the DMD directed to that camera. The focus is adjusted manually by moving the primary mirror back or forth, this procedure may introduce some error to the focus.

3.5.2 Sampling

The red blocks subsystem "Start sampling signal from SWIR photo diode" and "Store the raw signal" (figure 3.5) is conducted in a separate software which controls the A/D converter and thus the sampling. When the SPC is prepared, the sampling of the signal is started with a sampling rate such that every measurement has several sampling points and thus over-samples the signal. The oversampling is needed because when the mirrors move from one pattern to the next, the signal is uncertain for some time. The oversampling is also used to suppress noise from the photo diode (see further section 3.5.4). After the signal is sampled the obtained signal is stored on the computer manually.

3.5.3 Streaming patterns to the DMD

The subsystem "Streaming patterns to the DMD" (figure 3.5), represented in the purple block, is controlled by two different softwares, one which manipulates the pattern-signal received by the DMD and one which sends the patterns to the DMD. The patterns are sent to the DMD through a HDMI cable where the DMD is set up such that the DMD acts as a second screen to the computer. This enables to show anything on the DMD that the screen can show. The patterns are stored as a video and played back on the DMD "screen" with a media player, which shows each pattern in consecutive order. This is the major bottleneck of the system where each measurement matrix needs to be displayed one after the other depending on how fast frame rate that can be achieved. The naive approach would be to display one pattern per frame which is linked to the frame rate of the DMD. For example, 60 frames per second (fps). For a 512×512 image subsampled at 20% which corresponds to $512 \times 512 \times 0.2 = 52429$ patterns which would take $52429/60 = 874$ seconds = 14.5 minutes to sample. This is a long exposure time for a still image with the constraint that the scene should be stationary to obtain a stationary signal.

Fortunately, with the software "DLP LightCrafter 4500 EVM GUI" controlling the DMD, the received video signal can be manipulated before displayed onto the DMD. The software includes a function which can break down the received 24-bit color image into 1 bit planes which can be displayed in consecutive order. This function improves the naive implementation by a factor of 24, which reduces the time to sample the image in the given example from 874 seconds to $874/24 = 36$ seconds. That long exposure time is of course not optimal for natural images outdoors, but acceptable for the experimental setup.

To create the video that feeds the patterns to the DMD each pattern, i.e. measurement matrix, is created as presented in section 3.3.1. Then each group of 8 unique patterns are stacked in the 8 bit planes of an 8 bit image as seen in figure 3.6.

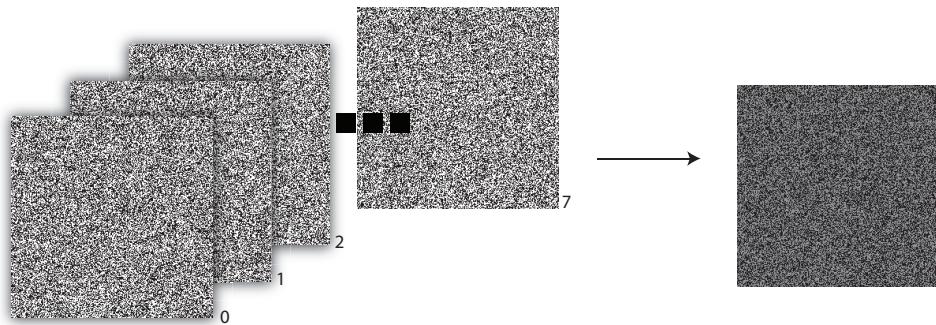


Figure 3.6: Each group of 8 measurement matrices is stored in separate bit planes in one 8 bit image.

Then for each group of three 8 bit images a 24 bit color image is constructed as seen in figure 3.7.

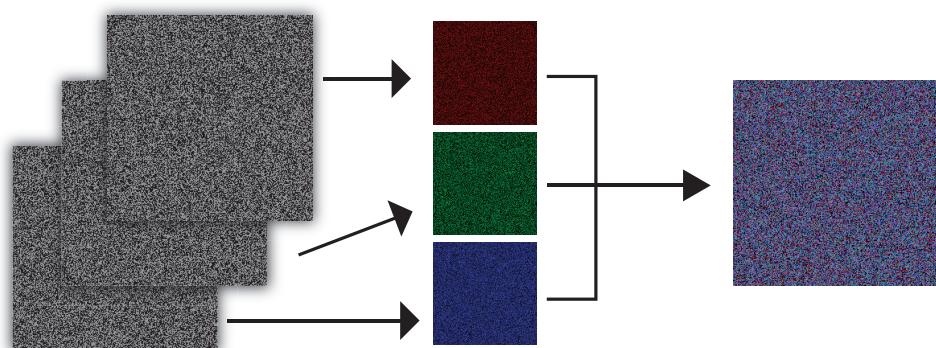


Figure 3.7: Each group of three 8 bit images is stored into one 24 bit color image. This is one frame in the video sent to the DMD.

The 24 bit color image corresponds to one frame in the video.

3.5.4 Signal processing

When the sampled signal is stored in the computer the remaining signal/image processing and reconstruction represented by blue blocks in figure 3.5 is conducted in MATLAB. In this section the signal processing of the sampled signal is described.

The first step is to refine the raw over-sampled signal so that each measurement matrix corresponds to one measurement in signal y . This is done by first finding every set of indices that correspond to every measurement matrix, see figure 3.8 where the signal indices are isolated by the magenta lines.

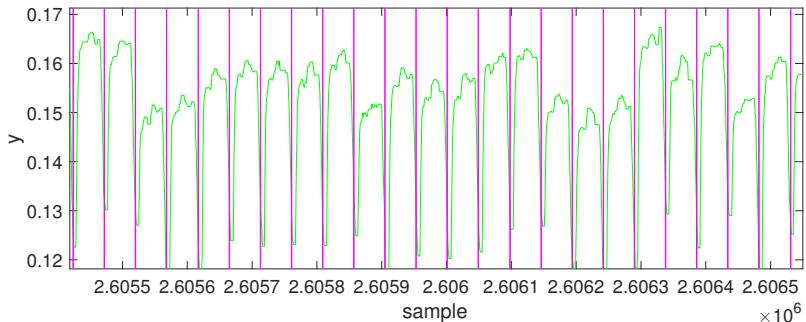


Figure 3.8: A simulated noisy over-sampled signal y where each sample in $y[m]$ is represented by multiple samples. The magenta lines separating each measurement which corresponds to one measurement matrix.

The next step is to determine one value for each measurement. This is done in two steps, the first is to omit values which corresponds to the DMD changing pattern close to the magenta lines in figure 3.8. For the remaining samples, the mean is calculated and set to the value for each sample $y[m]$, as seen in figure 3.9.

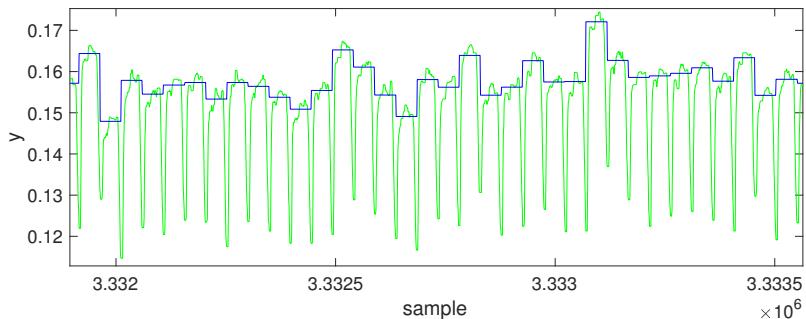


Figure 3.9: Calculated mean value for each measurement matrix with transition measurements omitted.

3.5.5 Dynamics in scene

The measured signal y should be stationary because the image (scene) is assumed to be static. When capturing images outdoors with natural light and long exposure times the image (or every pixel) may not be constant. This ambiguity of each pixel will reduce the reconstruction performance. The potential dynamics in a scene can be divided into two categories, luminance change and object movement. In this subsection the luminance change problem is modeled, and the corresponding algorithm will suppress the impact on the reconstructed image. The object movement problem will not be modeled, but will be avoided by ensuring that the scene is as static as possible.

In natural outdoor images it can be assumed that the primary source of light comes from the sun, but even on a clear day the light intensity from the sun is not constant. If the scene is assumed to be completely stationary, even the slightest intensity change will be amplified by all pixels being measured and thus changing the mean intensity of the measured signal \mathbf{y} which should be stationary. The consequence of the sampled signal \mathbf{y} not being stationary is that reconstruction performance will drop significantly. Therefore a model of light intensity change is created together with an algorithm to restore the signals stationary characteristics.

With the assumption that the scene is constant and the luminance change is uniform over the scene, the problem can be modeled.

Start with the original theorem and disregard the noise,

$$\mathbf{y} = \Phi \mathbf{x}. \quad (3.17)$$

The image \mathbf{x} can no be considered constant for all measurements, since the luminance change will change the image \mathbf{x} for every measurement matrix ϕ_i . This can be described for one measurement as,

$$y_i = \phi_i \mathbf{x}_i = \phi_i(\mathbf{x} + \mathbf{l}_i) = \phi_i \mathbf{x} + \phi_i \mathbf{l}_i, \quad (3.18)$$

where \mathbf{l}_i uniformly adds the same intensity over the whole image \mathbf{x} for measurement i . It is known from before that the measurement matrix ϕ_i contains 50% zeros and ones which gives,

$$y_i = \phi_i \mathbf{x} + \phi_i \mathbf{l}_i = \phi_i \mathbf{x} + \frac{N}{2} c_i, \quad (3.19)$$

where c_i is the uniform intensity change coefficient for measurement i . This function can be generalized for all measurements,

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{c}, \quad (3.20)$$

where \mathbf{c} is the uniform intensity change vector.

The goal is to remove the uniform intensity change vector \mathbf{c} from signal \mathbf{y} . Using the knowledge that \mathbf{y} should be stationary and assuming that the rate of change in intensity has a much lower frequency than the intensity change between individual measurement matrices, \mathbf{c} can be approximated by the moving mean and removed from \mathbf{y} . The moving mean is calculated for each sample $\mathbf{y}[m]$ by calculating the average of k samples centered around $\mathbf{y}[m]$, where k is chosen depending on the DMD pattern rate.

Moving mean is defined as

$$\mathbf{y}_{\text{MM}}[m] = \frac{1}{k} \sum_{i=m-\frac{k+1}{2}}^{m+\frac{k+1}{2}} \mathbf{y}[i], \quad (3.21)$$

where the calculation is made for each sample in \mathbf{y} and thus the algorithm to remove uniform intensity change is,

$$\mathbf{y} = \mathbf{y}_{\text{SAMPLED}} - \mathbf{y}_{\text{MM}} \approx \mathbf{y}_{\text{SAMPLED}} - \mathbf{c}. \quad (3.22)$$

The built in MATLAB function `movmean` will be used.

3.5.6 Reconstruction

Reconstruction is performed using the TVAL3 algorithm described in section 3.4.1. The algorithm takes the measurement matrix Φ , the sampled signal \mathbf{y} and the algorithm settings as input arguments and outputs the reconstructed image. The settings used throughout all experiments is:

- $opts.mu = 2024$
- $opts.beta = 64$
- $opts.maxcnt = 10$
- $opts.maxit = 1000$
- $opts.tol_inn = 10^{-5}$
- $opts.tol = 10^{-10}$
- $opts.mu0 = 16$
- $opts.beta0 = 1$
- $opts.nonneg = \text{true}$
- $opts.isreal = \text{true}$

This solves for a real non-negative solution as shown in equation 3.11. The variables was derived from TVAL3 default image reconstruction settings and then tweaked by changing each variable independently and inspecting the results on a set of test images.

3.5.7 Image processing

After reconstruction of the image some simple image processing is performed. There are only two operations applied to the reconstructed image and the reason is that the presented image results should represent what can be expected from the system. Furthermore image processing is often applied on special problems or artifacts in the images and it is not desired to cover up if such artifacts exist. Therefore the only two operations used are the median filter and adjusting the intensity for higher contrast.

The reconstructed image has a high dynamic range and if only a small set of neighboring pixels are reconstructed with a high intensity peak, which not correlates with the rest of the image, these pixels will drop the contrast in the rest of the image. To remove these peaks the median filter is used. The median filter will also remove "salt and pepper" noise while edges are preserved. The built in MATLAB function `medfilt2` is used.

The second operation is an intensity transform to maximize the contrast in the image, the built in MATLAB function `imadjust` is used.

3.6 Evaluation: Image quality assessment

The evaluation will be divided into two categories: reconstructed images from synthetic data and images reconstructed from data acquired by the SPC.

All results are produced with subsampling ratios ranging from 5-30% and evaluated. The upper limit was set to 30% partly because of the hardware limitations with long exposure time and partly because the main advantage of CS/CI is to minimize the required subsampling ratio.

The evaluation on synthetic data is focused on evaluating the performance of the measurement matrix and reconstruction algorithm. Evaluating synthetic data gives advantages that can not be achieved with images reconstructed using the SPC which is that there is a reference image which the resulting image can be compared to.

A reconstructed image from synthetic data is acquired by creating a signal $\mathbf{y}_{M \times 1}$ taking the inner product of $\mathbf{y} = \Phi\mathbf{x} + \epsilon$ where, \mathbf{x} is the synthetic image reshaped to a vector, Φ is the measurement matrix with the desired amount of measurements M and synthetic noise ϵ which can be regulated to simulate different conditions, then using the reconstruction algorithm on the signal \mathbf{y} to obtain the reconstructed image $\hat{\mathbf{x}}$. Since the measurement matrix and the reconstruction algorithm are independent of the SPC hardware the subsystem can be evaluated

independently. Two advantages of evaluating the sensing and reconstruction independently of the SPC is that parameters such as number of measurements and noise can be regulated easy and the second advantage is that a reference image is available for comparison.

3.6.1 Evaluation Using reference image

With a reference image available, two image quality assessments are performed on the result from the simulation: Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index. PSNR is defined as

$$\text{PSNR}[f(x, y), g(x, y)] = 10 \log_{10} \frac{E^2}{\text{MSE}[f(x, y), g(x, y)]} \quad (3.23)$$

where, $f(x, y)$ and $g(x, y)$ are the intensity in pixel (x, y) , E is the maximum possible pixel value and MSE is the mean square error between the images defined as

$$\text{MSE}[f(x, y), g(x, y)] = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} [f(x, y) - g(x, y)]^2. \quad (3.24)$$

The SSIM algorithm is not focused on pixel to pixel differences like PSNR, but instead of the structure of the image in small windows. SSIM separates luminance, contrast and structure and calculates the difference in each category in a small window to calculate the similarity of the images. The SSIM index is defined as

$$\text{SSIM}[f(x, y), g(x, y)] = \sum_n l[f(n), g(n)]^\alpha c[f(n), g(n)]^\beta s[f(n), g(n)]^\gamma, \quad (3.25)$$

where n is the window, $\alpha = \beta = \gamma = 1$ and

$$\text{luminance: } l = \frac{2\mu_{f(n)}\mu_{g(n)} + C_1}{\mu_{f(n)}^2 + \mu_{g(n)}^2 + C_1}, \quad (3.26)$$

$$\text{contrast: } c = \frac{2\sigma_{f(n)}\sigma_{g(n)} + C_2}{\sigma_{f(n)}^2 + \sigma_{g(n)}^2 + C_2}, \quad (3.27)$$

$$\text{structure: } s = \frac{\sigma_{f(n)g(n)} + C_3}{\sigma_{f(n)}\sigma_{g(n)} + C_3}, \quad (3.28)$$

where,

- $\mu_{f(n)}$ and $\mu_{g(n)}$ is window mean.
- $\sigma_{f(n)}$ and $\sigma_{g(n)}$ is window standard deviation.

- $\sigma_{f(n)g(n)}$ is window cross covariance.

- $C_1 = 0.01 * 255$, (MATLAB default).

- $C_2 = 0.03 * 255$, (MATLAB default).

- $C_1 = C_2/2$, (MATLAB default).

Which summarizes to

$$\text{SSIM}[f(x, y), g(x, y)] = \sum_n \frac{(2\mu_{f(n)}\mu_{g(n)} + C_1)(2\sigma_{f(n)g(n)} + C_2)}{(\mu_{f(n)}^2 + \mu_{g(n)}^2 + C_1)(\sigma_{f(n)}^2 + \sigma_{g(n)}^2 + C_2)}. \quad (3.29)$$

The SSIM index has a max value of 1 when the images are identical which makes it easy to read. [3]

3.6.2 Evaluation Using no reference quality assessment

In order to evaluate image quality when there is no reference image to compare against, the BRISQUE algorithm is used as a complement. BRISQUE is a no reference image quality assessment model which is based on natural scene statistics and quantifies the "naturalness" of the image. [2]

3.6.3 Evaluating Using Edge response

The edge response measures the sharpness of the image by calculating the distance in pixels required for an edge in the image to rise. In this master's thesis, the distance required for the edge response to rise from 10% to 90% was chosen, see figure 3.10. This evaluation is performed on static images captured in constant light indoors for consistent results. Furthermore, the motive of the image is slanted geometric objects printed on a sheet of paper.[15]

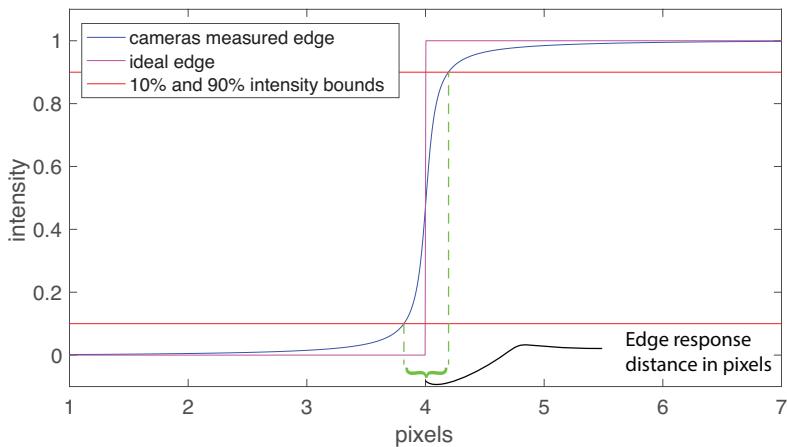


Figure 3.10: Definition of 10-90% Edge response.

3.7 Method criticism

- The BRISQUE algorithm is not designed for SWIR images or SPC:s characteristics noise. Therefore the results may not reflect how the quality assessment would answer to visual wavelength cameras. BRISQUE definition of "naturalness" may not reflect the images captured by the SPC.
- The effect on the reconstructed images caused by the DMD mirrors alignment and pairing is not known.

4

Results and Evaluation

This section is divided into two categories, simulated results and SPC results. In both sections the same evaluation is performed in order to draw conclusions about the performances of the different parts of the SPC chain and to be able to answer the questions in section 1.5 and 1.4.

4.1 Simulated Results

In this section the results were simulated by using the reconstructing algorithm and measurement matrix described in section 3.3.1 and 3.4.1 on high quality images, captured with a state of the art SWIR camera. These images act as ideal references to the reconstructed images. By simulating the result from "ideal" images, the reconstruction process got a benchmark independent of the SPC hardware.

To generate the *simulated reconstructed images*, the inner product between the complete measurement matrix Φ and the reshaped "ideal" image vector x was calculated to obtain a simulated signal vector y

$$y = \Phi x + \epsilon. \quad (4.1)$$

This operation were calculated for different subsampling ratios between 5-30% and different noise levels. White Gaussian noise was added to the normalized measurement signal y . The added noise represents a simple model of the noise expected in the SPC and was scaled with the standard deviation σ between 0 – 0.2. The standard deviation was not increased above 0.2 because the reconstruction failed at that point.

Then the simulated images were produced by the reconstruction algorithm using signal vector \mathbf{y} . 21 images were simulated in 6 different subsampling ratios and 10 different noise levels yielding 1260 simulated images as foundation for this evaluation.

4.1.1 Reconstruction performance Using reference image

The performance of the reconstruction was calculated using PSNR and SSIM for different degree of noise and subsampling ratios.

To create the graphs in figure 4.2 and 4.3 this procedure was applied to all 21 images for subsampling ratio 5% to 30% and noise was added with standard deviation between 0–0.2. In figure 4.1 a sample of reconstructed images from one of the SWIR images is presented with different amount of noise and subsampling ratios.

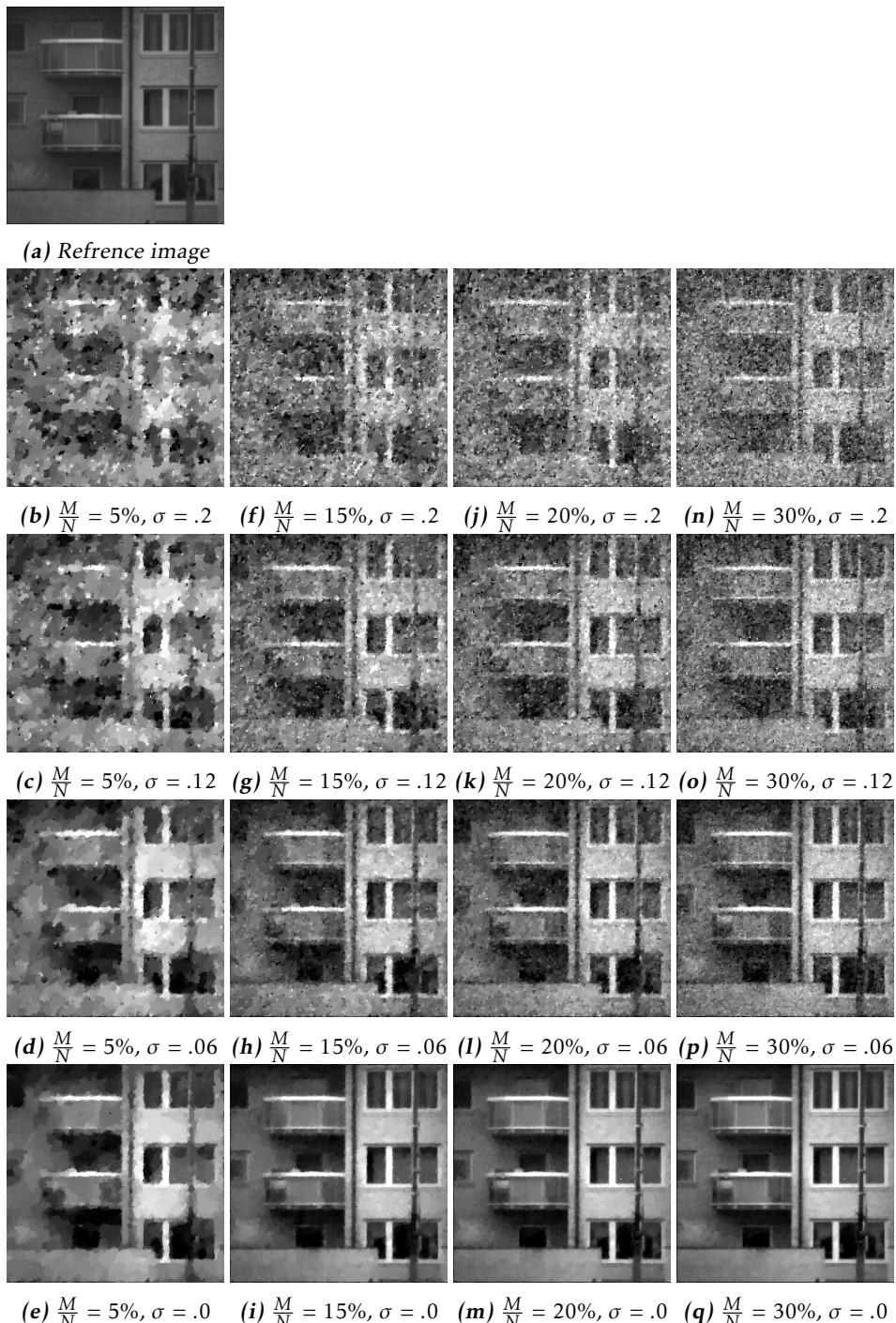


Figure 4.1: Example of reconstructed images with added noise at different subsampling ratios.

As seen in figure 4.1 the reconstructed image quality increased with more measurements and lower noise levels. This observation is confirmed in the graphs in figure 4.2 and 4.3 where PSNR and SSIM respectively have been calculated and interpolated for all 21 reconstructed images.

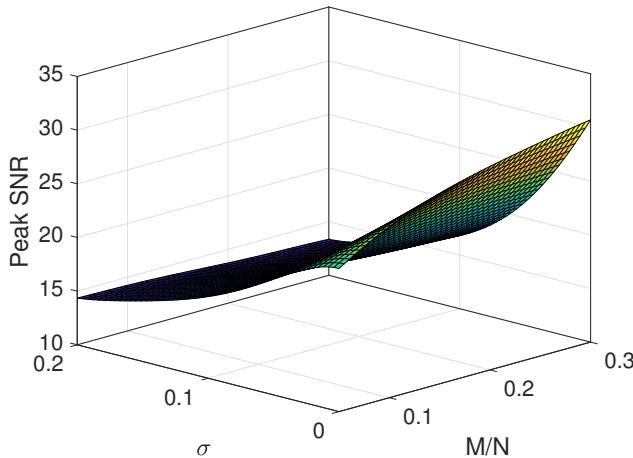


Figure 4.2: Peak SNR result depending on number of measurements and simulated noise level. $\frac{M}{N}$ is the subsampling ratio and σ is the standard deviation added to \mathbf{y} .

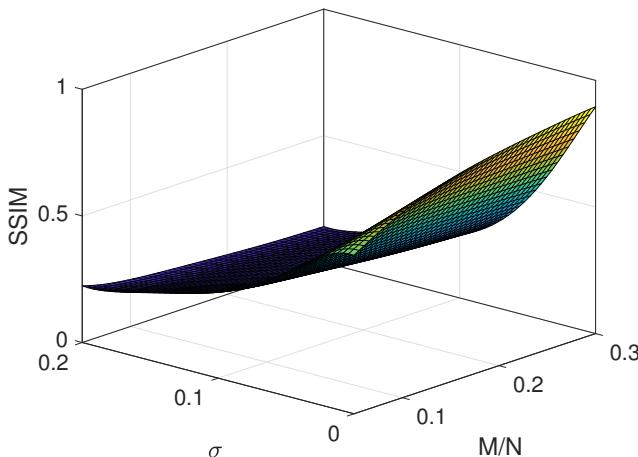


Figure 4.3: SSIM result depending on number of measurements and simulated noise level. $\frac{M}{N}$ is the subsampling ratio and σ is the standard deviation added to \mathbf{y} .

In both figure 4.2 and 4.3, it can be seen that when the noise increases the reconstructed image quality is not improved at the same rate as in the noiseless case, when the subsampling ratio is increased.

4.1.2 Reconstruction performance using no reference quality assessment

In this sub section the same reconstructed image set from section 4.1.1 is used to calculate the *no reference image quality* with the BRISQUE algorithm.

The results displayed in figure 4.4 show that less noise and more samples yield better performance in the reconstruction. The figure also contain the mean results from the "ideal" SWIR images as the flat blue surface, which has scored a far greater score than the reconstructed images.

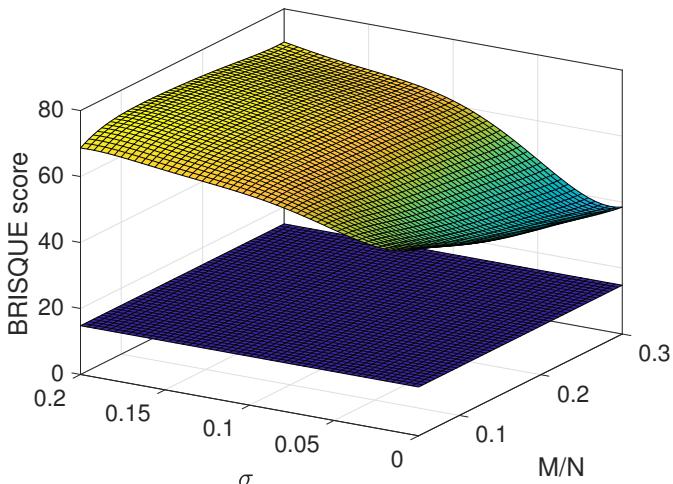


Figure 4.4: BRISQUE result depending on number of measurements and simulated noise level. Lower surface is reference image score. $\frac{M}{N}$ is the subsampling ratio and σ is the standard deviation added to \mathbf{y} , lower BRISQUE scores are better.

In figure 4.5 the result has been flattened to a 2D graph with fewer selected data points for clarity. In the noiseless case the score will not be better than approximately 40 for the reconstructed images, while the SWIR images have a mean value of 15.

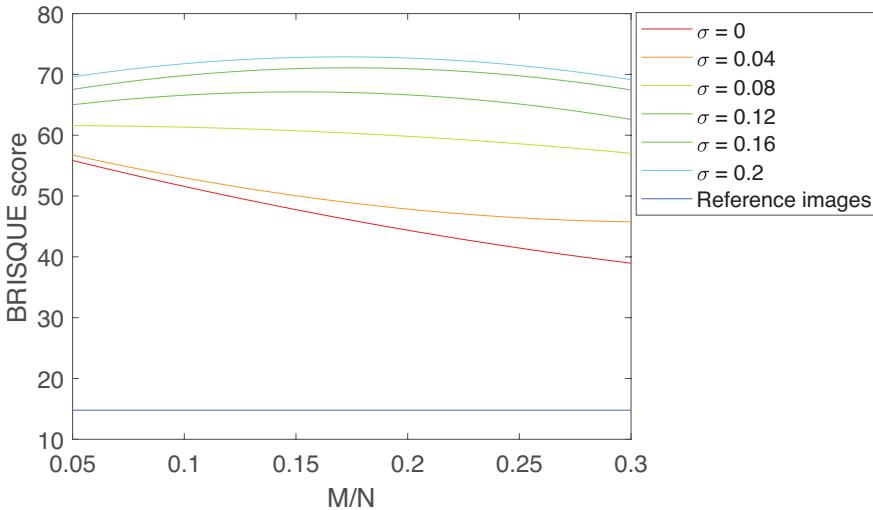


Figure 4.5: BRISQUE result depending on number of measurements for different simulated noise levels. $\frac{M}{N}$ is the subsampling ratio and σ is the standard deviation added to \mathbf{y} , lower BRISQUE scores are better.

In the simulated images with $\sigma > 0.08$ the BRISQUE score start to get unexpected results, first yielding a worse BRISQUE score when increasing the subsampling ratio up to 15-20% but then gets better after 20%, as seen in figure 4.5.

4.1.3 Dynamics in scene

In the SPC setup, the exposure time was between 10 and 50 seconds, which increased the risk of dynamics in the scene. Dynamics in the scene reduces the reconstruction performance because the scene is assumed to be constant. By simulating dynamic scenes in a controlled environment, their individual effects to the sampled signal \mathbf{y} could be identified and evaluated. As mentioned in section 3.5.5 dynamics in the scene can roughly be divided into two separate categories, luminance change and movement. In this section, global luminance change and two kinds of movement are simulated. The goal was to see how the signal changes when dynamics are introduced in the scene. In the case of luminance change, the moving mean algorithm presented in section 3.5.5 was evaluated.

To generate a simulated measurement representing a dynamic scene each sample $\mathbf{y}[m]$ is constructed using a unique image \mathbf{x}_m , which has been changed from the previous image,

$$\mathbf{y}[m] = \phi_m \mathbf{x}_m. \quad (4.2)$$

In the first scenario an object was placed in an image, but for each measurement the location of the object was moved in a small bounded area of the image. Consequently, this model represents a scene where the background is static with a person moving in a small area.

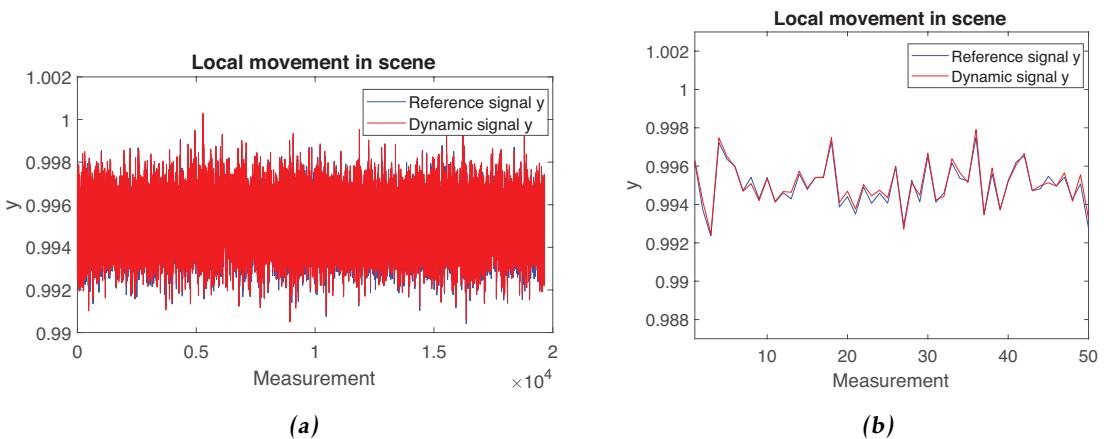


Figure 4.6: (a) Perturbed signal from local movement on top of reference signal. (b) Zoomed in view of some samples from figure (a).

As seen in figure 4.6a there was no obvious difference between the non perturbed reference signal and the distorted signal. Neither in the zoomed in view in figure 4.6b, any large difference can be seen.

The reconstructed images from the reference signal and the perturbed signal are displayed in figure 4.7b and 4.7c, respectively. The difference between the images are visible to the naked eye. Not only does the moving object get blurry and noisy, but the whole image globally.

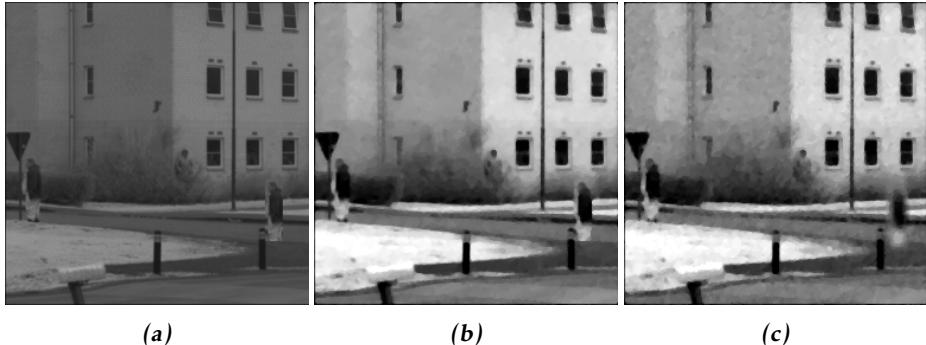


Figure 4.7: The results of local movement in a reconstructed image, subsampled at 30%. (a) Original reference image. (b) Reference image reconstructed from the original image without movement. (c) Reconstructed image from a scene with local movement.

In table 4.1 the results from calculating PSNR and SSIM of the the reconstructed images are presented. It can be observed that the dynamic test image (figure 4.7c) has been affected to some degree by the movement compared to the unperturbed image in figure 4.7b.

Peak SNR	SSIM
29	0.91

Table 4.1: Evaluation comparing unperturbed reconstructed images against reconstructed images with local movement.

The second scenario is an object passing through the whole scene. The problem is modeled with a static background and a simulated object crossing the whole scene, like a car, human or animal might do when using the SPC. The object will cross the scene in 1000 measurements of approximately 19000 in total which corresponds to approximately 0.7 seconds, when sampling with the SPC in its current setup.

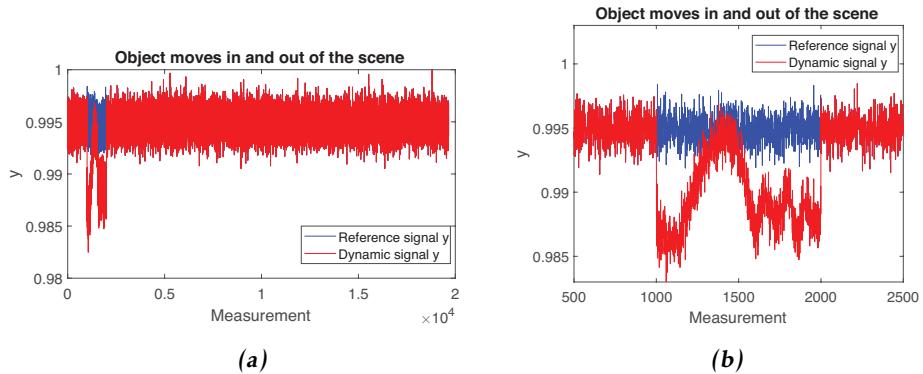


Figure 4.8: (a) Perturbed signal from large movement on top of the reference signal. (b) Zoomed in view of some samples from figure (a).

As seen in figure 4.8, at measurement 1000 the exact moment the object enters the scene, the signal changes. This is because a completely new structure has entered the scene and therefore the DC level changes. It can also be noted that after a while the object passed something which has approximately the same intensity as the background and therefore the DC signal almost returns to its original value for a brief moment.

In figure 4.9 the effect of the moving object can be seen in the reconstructed image, which has gained a lot of global noise. Note that the object passing through can not be seen because there is more measurements of the background than of the moving object. Nevertheless, the object is creating uncertainty in the whole image, resulting in global noise.

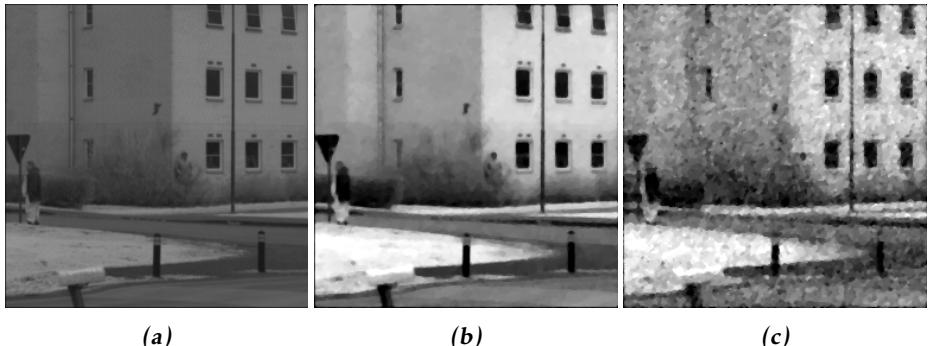


Figure 4.9: The results of large movement on a reconstructed image, subsampled at 30%. (a) Original reference image. (b) Reference image reconstructed from the original image without movement. (c) Reconstructed image from a scene with an object passing trough.

In table 4.2 the results from calculating PSNR and SSIM of the reconstructed images are presented. It can be observed that the image has been effected heavily by the movement, lowering the SSIM index to 0.58.

Peak SNR	SSIM
23	0.58

Table 4.2: Evaluation comparing unperturbed reconstructed image against reconstructed image with movement.

The third scenario is luminance change in the scene caused by inconsistency of light intensity from the source. Outdoors this means that the light intensity from the sun will vary over time, the most obvious being clouds occluding the sun but for example even change in air density can change the intensity. This scenario is modeled by adding or subtracting the global intensity in the image over the measurements.

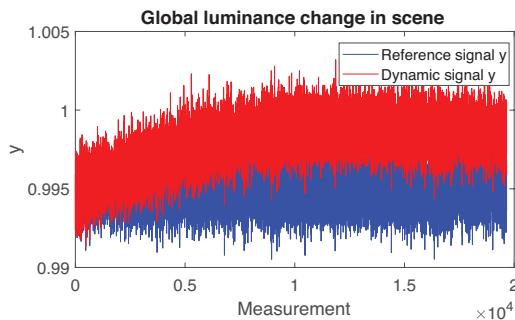


Figure 4.10: Signal effected by light intensity change on top of reference signal.

As seen in figure 4.10, the DC level of the signal will slowly change, but the structure of the signal stay the same. In figure 4.11 the reconstructed images from the perturbed signal and the reference signal are displayed. The reconstructed image from the dynamic signal has gained a lot of global noise even though the structure in the image has not been changed over the measurements.

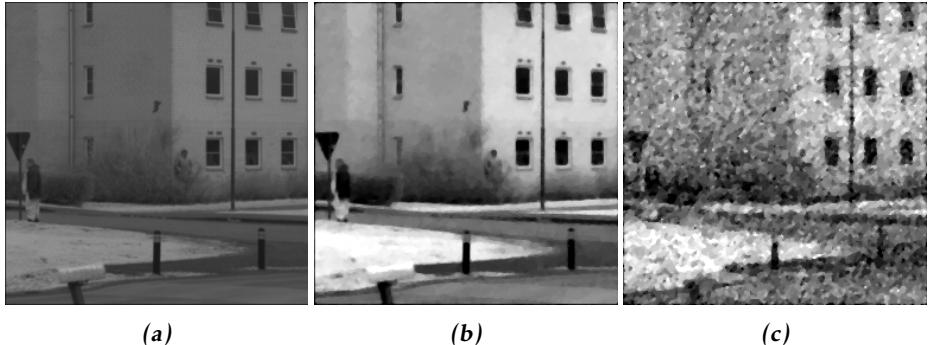


Figure 4.11: The result of global light intensity change on a reconstructed image subsampled at 30% (a) Original reference image. (b) Reference image reconstructed from the original image without light intensity change. (c) Reconstructed image from a scene with global light intensity change over the measurements.

In section 3.5.5, a model of this problem was proposed along with an algorithm to suppress the impact of global luminance change. The algorithm is applied to this experiment to evaluate its performance. The moving mean subtraction method is applied and in figure 4.12a the resulting signal is plotted over the dynamic signal. Note that the processed signal is stationary again. In figure 4.12b and 4.12c, where the processed signal is plotted over the reference signal, it can be seen that the processed signal has gained its original structure and almost fit exactly to the original.

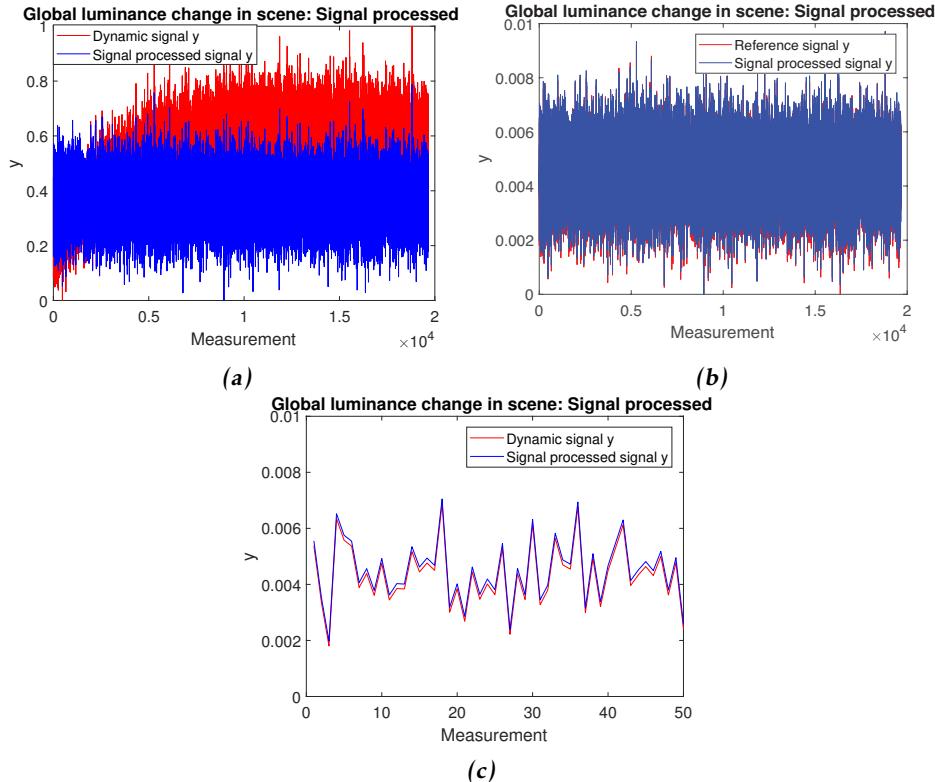


Figure 4.12: Post-processed signal using moving mean subtraction. (a) Post-processed signal on top of the dynamic signal. (b) Post-processed signal on top of the reference signal. (c) Zoomed in view of (b).

In figure 4.13, the processed signals reconstructed image is displayed between the reference and perturbed signals reconstructed images. The moving mean algorithm improve the reconstruction significantly, the image has gained some noise compared to the reference image, but over all there is not much difference between them.

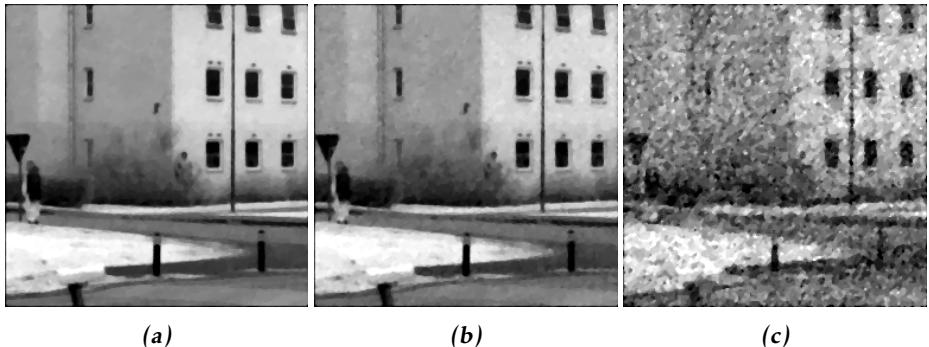


Figure 4.13: The result of processed signal perturbed by light intensity change on a reconstructed image subsampled at 30% (a) Reference image reconstructed from the non perturbed signal without light intensity change. (b) Reconstructed image from a scene with global light intensity change and post processed by moving mean subtraction. (c) Reconstructed image from a scene with global light intensity change over the measurements.

In table 4.3 the results from calculating PSNR and SSIM of the reconstructed images are presented. Both PSNR and SSIM are increased for the reconstructed image using moving mean subtraction.

Signal	Peak SNR	SSIM
Perturbed signal	19	0.38
Mean subtracted signal	33	0.93

Table 4.3: Evaluation comparing unperturbed reconstructed image against global luminance change reconstructed image and moving mean subtracted signal processed reconstructed image.

4.2 SPC evaluation

The evaluation section is now shifted to examine the images produced by the SPC. The images will be analyzed using a range of methods to examine the performance of the SPC. This section will start with the same evaluations performed in the previous section and will end with two new evaluations, edge response and subsampling ratio. All images captured by the SPC of natural scenes were taken at a distance between 200 to 900 meter on sunny days with some clouds.

4.2.1 Reconstructed performance using reference image

This evaluation is designed to get a measurement of expected image quality with the same metrics used in the synthetic case. As stated before, it is hard to obtain a reference image for the images produced by the SPC. One solution to obtain

a reference image is to adjust a homography between two images taken by the SPC and a high quality SWIR camera respectively. But there are two problems of just performing a homography to obtain a reference image. The first one is that a homography estimates the transformation between flat surfaces, which excludes most natural images, and the second is that the estimated homography will not be perfect and thus high contrast edges in the images will not match and produce large errors in the performance measurements. To solve the first problem, the scene is a flat surface with a printed pattern. To avoid the second problem, error from sharp edges, the pattern is constructed with smooth contrast changes. To not complicate this experiment, the reference image is a computer generated optimal image which was transformed via a homography to fit the image captured by the SPC.

The reference image which was created using sine functions to avoid edges and the reconstructed images with different subsampling ratio can be seen in figure 4.14.

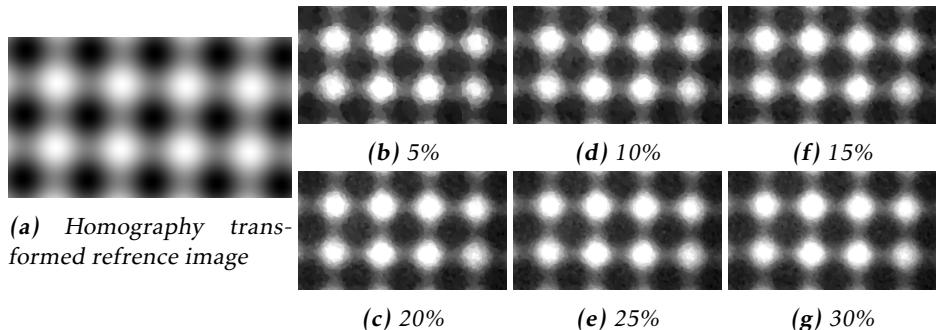


Figure 4.14: The reconstructed images with subsampling ratios between 5-30% and the reference image transformed to fit the SPC images using homography.

Before the results from the evaluation is presented, it is worth noting that the reference image is a perfectly simulated reference image, which was not affected by any uneven light source or quality of the print as the reconstructed image from the SPC, which for example can be seen in the edges of the reconstructed images in figure 4.14b to 4.14g.

In figure 4.15, PSNR and SSIM has been calculated for each subsampling ratio.

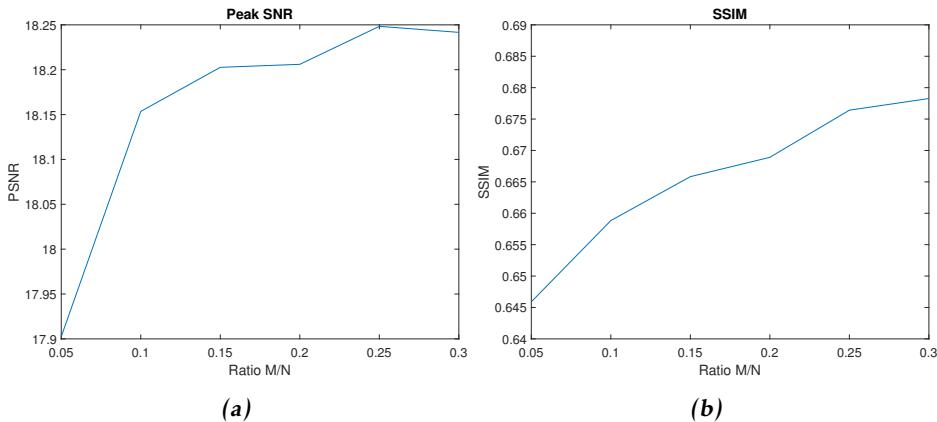


Figure 4.15: Signal quality of SPC images compared to reference image. (a) Peak SNR for reconstructed images against reference image. (b) SSIM score for reconstructed images against reference image.

The results show increased quality with increased subsample ratio. In the case of PSNR the image quality rapidly improves when increasing the subsample ratio up to 15% and then flattens out.

4.2.2 Reconstruction performance using no reference quality assessment

In this section the blind quality assessment tool BRISQUE is used to evaluate the reconstructed images from the SPC, examples of these images are shown in figure 4.17 to 4.19. In addition to presenting the BRISQUE score for the images, each image has been classified into one of three distinct groups in order to link the BRISQUE scores to each image and their sampled signal.

In figure 4.16 each image is evaluated using BRISQUE with subsampling rate from 5 – 30%.

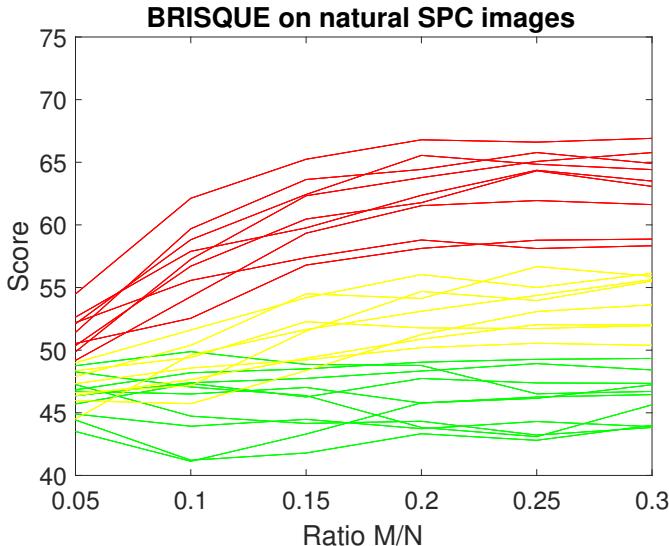


Figure 4.16: BRISQUE score for images reconstructed from the SPC with subsampling ratios from 5% to 30%. Each line represent one image and is classified in three different groups indicated in different colors depending on the initial score at subsampling ratio 5% and the general trend for higher subsampling ratios. Lower score is better. Examples of these images are shown in figure 4.17 to 4.19.

As seen in figure 4.16, each image has been plotted separately due the high variance in the image set and the distinct different trends. Note that lower BRISQUE score is better.

The images have been classified into three different classes depending on the initial score at 5% subsampling ratio and the trend when increasing the subsampling ratio. The classes have been color coded where:

- the *red image set* has a higher initial BRISQUE score than the rest of the set and got worse BRISQUE score at a relative high rate when increasing the subsampling ratio.
- The *yellow image set* is distinguish by gradual increase in BRISQUE score when the subsampling ratio is increased.
- The *green image set* has a plain trend or decreased BRISQUE score when the subsampling ratio is increased.

In figure 4.17-4.19 a subset of four images from each class is presented. All images are reconstructed using 30% subsampling ratio.

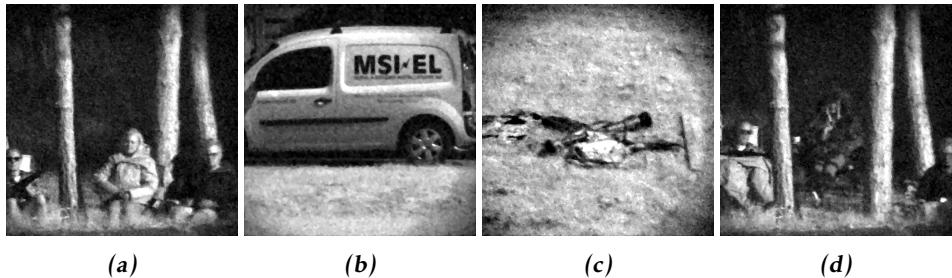


Figure 4.17: Sample images from the green image set. (a) and (d) People sitting in the edge of a forest. (b) Stationary car. (c) Camouflage jackets and a AT-4 anti-tank weapon on the ground.

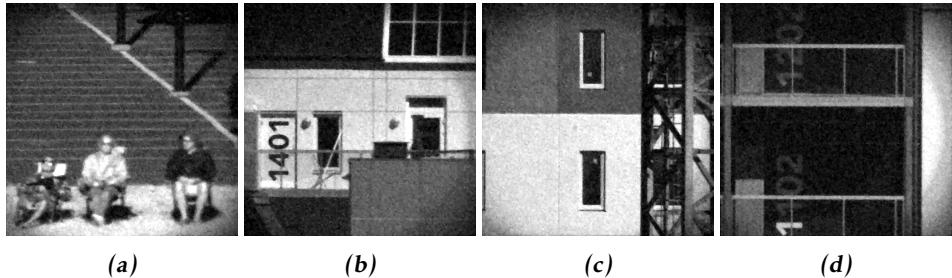


Figure 4.18: Sample images from the yellow image set. (a) People sitting next to a parking lot. (b) - (d) House facades.

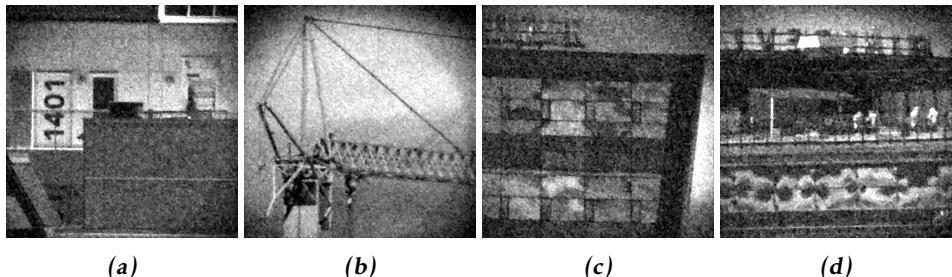


Figure 4.19: Sample images from the red image set. (a) House facade. (b) Construction crane (Moving clouds in background). (c) Mjärdevi Center facade (Moving clouds reflected in the windows). (d) Mjärdevi Center balcony with people having a break (Moving clouds reflected in the windows).

As seen in the images in figure 4.17 - 4.19, the motive in each class is quite different, the images from the green image set contains forest, people and objects, images from the yellow image set contains buildings and structures with horizon-

tal and vertical lines and images from the red image set includes the sky but are also a bit darker and overall noisier.

The last result in this section is the mean signal strength plotted against the normalized signal variance of the background noise where each signal has the same corresponding color classification as given before.

The background noise normalized signal variance was calculated by normalizing the background noise relative the desired signal and then calculating the variance. The mean signal intensity is the DC level of the measurement i.e the signal strength of approximately half the pixels of the scene.

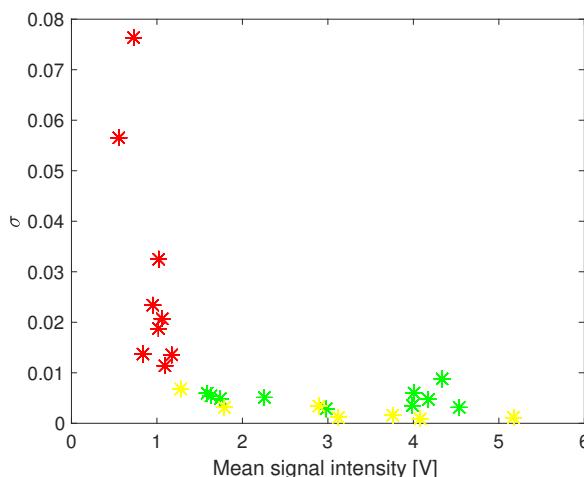


Figure 4.20: Normalized background noise, measured in standard deviation σ plotted against mean sampled signal intensity, each signal has the same color classification as given before and shown in figure 4.16.

From the plot in figure 4.20, it can be seen that for a mean signal intensity over approximately 1.2 volt, there are no images from the red image set, while images from both the yellow and green image set are mixed over 1.2 volt. For a mean signal intensity under 1.2 volt, the noise standard deviation rapidly increases.

4.2.3 Luminance change in scene

As predicted in section 3.5.5, dynamics in the scene could result in poor reconstruction performance and an algorithm to suppress the distortion caused by luminance change was tested in the simulated evaluation in section 4.1.3. With an exposure time of just under one minute for the SPC, this problem turned out to be constantly present when taking photos of natural scenes outdoors. Moreover the luminance change in the signal was more complex than in the simulated test case and the sensor was highly sensitive to luminance change. This observation

should be anticipated, the sensor sums up half the scenes light, which make the tiniest intensity change for each pixel a large global change.

In this section the impact of luminance change on the signal and the results of using the moving mean algorithm (defined in section 3.5.5) is presented. In figure 4.21 the sampled signal effected by luminance change is plotted in red and the signal processed by the moving mean algorithm is shown in blue.

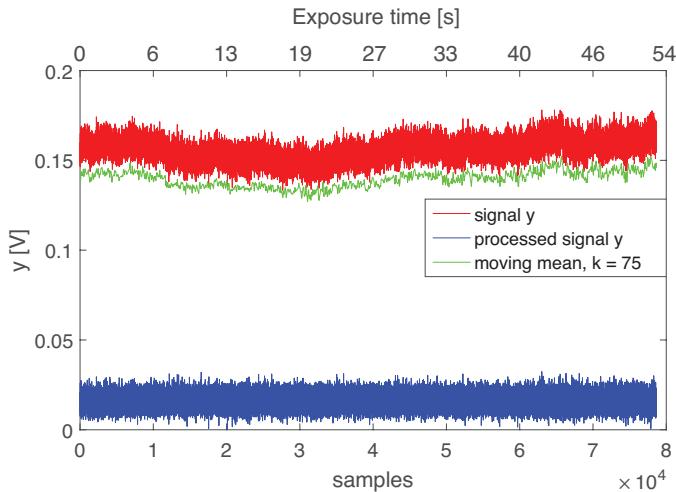


Figure 4.21: Sampled signal from SPC with light intensity change distortion and the signal processed by the moving mean algorithm.

As can be seen in figure 4.21, the moving mean (plotted in green) and corresponding to light intensity change, change with a high frequency. The number of neighboring samples to calculate the average was set to $k = 75$ to match the fast change which corresponds to a window of 50 milliseconds. In figure 4.22, a reconstructed image without the moving mean algorithm used and one with the method used is displayed.

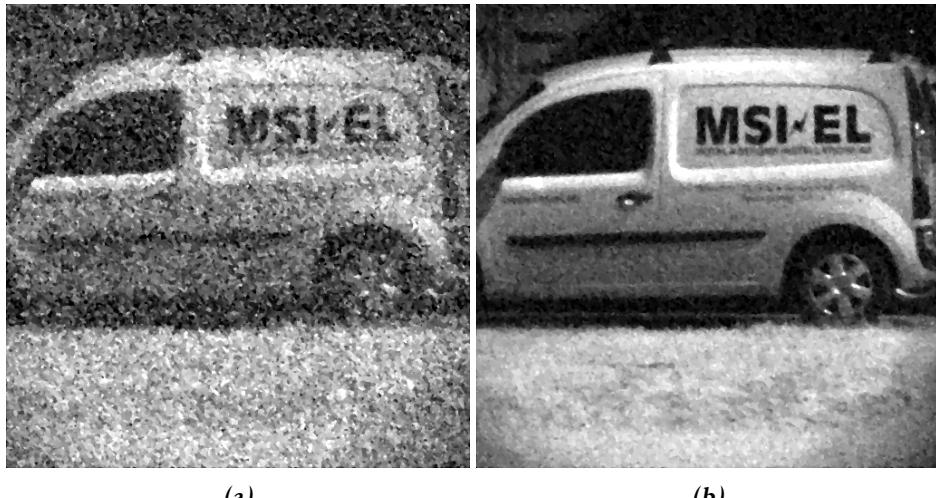


Figure 4.22: Reconstructed images, (a) before and (b) after applying moving mean algorithm on the sampled signal.

As can be seen in figure 4.22, the method increase image quality significantly and the image reconstructed without the moving mean algorithm has a severe global noise due to the non-stationary sampled signal. This image was captured in good conditions with strong lighting and relative mild intensity change compared to the majority of the sampled signals.

4.2.4 Edge response

The edge response is used to compare the sharpness of cameras and lenses and in this section the SPC is compared to a state of the art SWIR camera. Two scenes were captured by the SPC and the conventional SWIR camera respectively, containing printed sheet of paper with simple tilted shapes on them, see figure 4.23. The scene was lit by a 135 Watt halogen lamp placed two meters from the sheet.

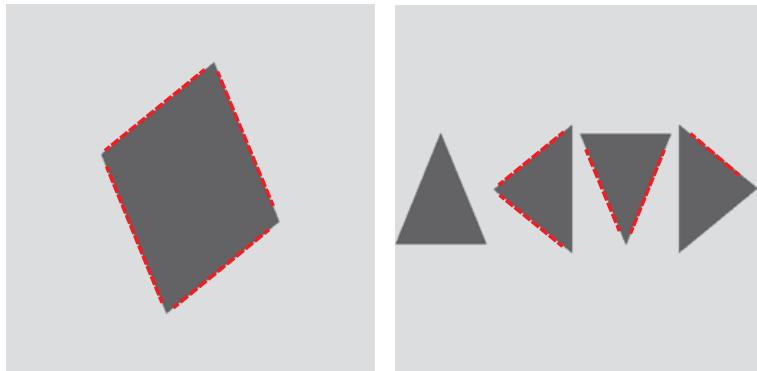


Figure 4.23: Printed targets with markings where the edge response measurements were performed.

In the resulting images, edge response measurements were gathered from the specified edges in figure 4.23. With the result from all edges and both images for each camera, a mean and standard deviation is calculated. For the SPC, images reconstructed from 5% to 30% was tested in order to see if the subsampling ratio affected the edge response result. The images from the SWIR camera and SPC are presented in figure 4.24.

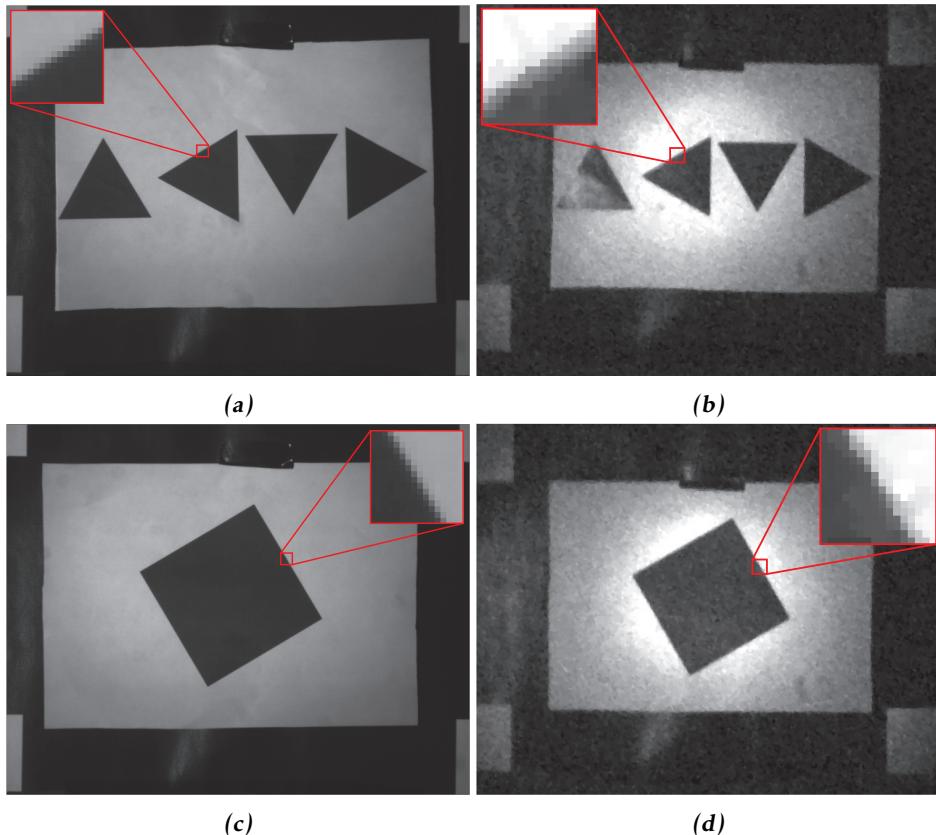


Figure 4.24: SPC and state of the art SWIR camera images. (a) and (c) shows images from the conventional SWIR camera and (b) and (d) shows images captured with the SPC. The images contain a zoomed in view of a sample edge.

The edge response is measured as the distance in (pixels) required for an edge to rise from 10% to 90% intensity change, see section 3.6.3. The result from the experiment in presented in figure 4.25.

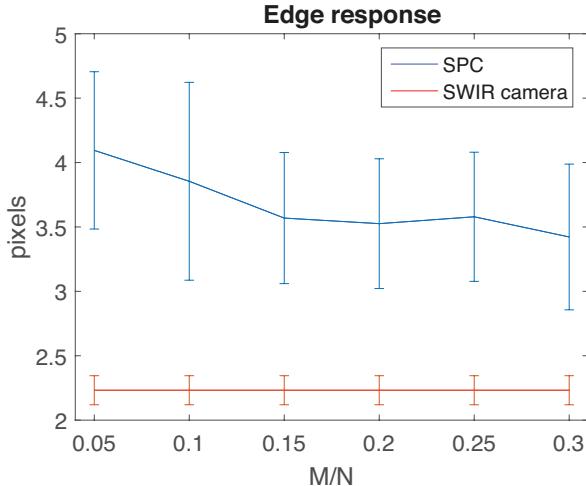


Figure 4.25: Edge response, distance (pixels) to rise from 10-90% in average for an edge.

From the plot in figure 4.25, a clear difference between the SPC and state of the art SWIR camera can be seen, where the conventional SWIR camera has in average half the edge distance compared to the SPC images. Some improvement is seen when the subsample ratio is increased, but the standard deviation is almost constant, meaning that the difference between the state of the art SWIR camera and the SPC, in best case only differ about 0.5 pixels, but in the worst case differ about 1.7 pixels and in average 1.2 pixels.

4.2.5 Subsampling ratio

From the theory of compressive sensing, the number of measurements needed to reconstruct an image is correlated with the sparsity or compressibility of the image. Therefore it is hard to give a good estimate of the subsampling ratio needed to obtain a desired quality of the reconstruction. In addition, using an SPC where noise contaminate the signal and the scene may not be completely stationary, the number of measurements needed will increase in proportion to the noise and the change in the scene. In this subsection the minimum subsampling ratio will be presented followed by how the reconstructed image quality is affected by an increase of subsampling ratio.

The minimum number of measurements to reconstruct an image where the motive can be recognized is also affected by the factors described in the previous paragraph. Trying to reconstruct an image under the minimum subsampling ratio results in an image with noise. A subjective investigation of the minimum subsampling ratio was performed on all images captured by the SPC in this thesis, the quality of each image was studied in 1% subsample ratio increment. The

result ranged from 2% to 4% to obtain a recognizable image. In figure 4.26, a sample of three images with varying minimum subsampling ratios are displayed.

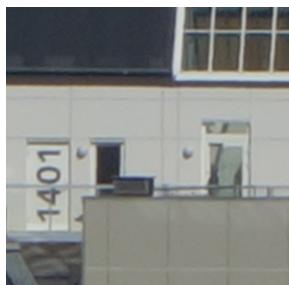


(a) Subsampling ratio 2%. (b) Subsampling ratio 3%. (c) Subsampling ratio 3%.

Figure 4.26: Varying minimum subsampling ratios to reconstruct sample images captured by the SPC.

In the sample images in figure 4.26, the minimum subsampling ratio varied between 2% to 3%. However, as can be seen in the images, the motive is merely recognizable, large structures can be identified but details are lost. In general a subsampling ratio of 5% has always succeeded to reconstruct a identifiable image with some fine details.

In figure 4.27 three scenes are reconstructed using different subsample ratios ranging from 5% to 30%. For each row the subsampling ratio is increased by 5% and at the top a reference image from a visual camera is presented. This result can give a perception of which subsampling ratio is good enough for a given purpose, and as can be seen, the increase of the subsampling ratio is not linearly proportional to the increase in perceived image quality, like in regular image compression.



(a) Visual camera image.



(b) Subsampling ratio 5%



(c) Subsampling ratio 10%



(d) Subsampling ratio 15%



Figure 4.27: Reconstructed images captured by the SPC with increasing subsampling ratios. The first row shows a reference visual spectrum image and the following rows show SPC images reconstructed from 5% up to 30% subsampling ratio.

From the images in figure 4.27, it can be observed that:

- The reconstructed images behave as expected, when the subsample ratio increases, the image quality increases.
- Already at 5% subsampling ratio the scene can be properly identified unlike the case in minimal subsamples ratio, shown in figure 4.26. The images have some artifacts, it look like they are made of small patches.
- Between 10-15% the finer details start to appear, like the gap between the

panels in the facade. The outline of the text on the car and door are getting sharper and the structures in the faces start to appear.

- As stated before, the image quality does not increase at the same rate as the subsampling ratio. This is most noticeable when increasing the subsampling ratio above 15%. The improvement after 15% is not as easy to spot as before, but can be found in the details.

5

Discussion

In this section the results and the method are analyzed and discussed. When discussing results, a focus on consistency, relation to theory and real world applicability is held. In the discussion of method, an analysis of replicability, reliability and validity is held as well as the method used to capture images with the SPC.

5.1 Results

Overall, the results obtained in this thesis reflect what has been stated in theory and research before. However no reference of using the SPC to capture natural scene has been found and thus this thesis will present the link between real world applications and theory/lab results and the challenges that come with it.

5.1.1 Reconstruction performance using reference image

In the simulated reconstruction, the results behaved in most parts as expected given the CS theory i.e. with increased subsampling ratio, the performance increased. However the interesting part of the results is what happens when increasing the noise. Not only does the general performance drop for all subsample ratios, but also the improvement rate by increasing the subsample ratio drops, which figure 4.2 and 4.3 shows. This result tells us that if the signal contains a high degree of noise, a higher subsampling ratio may not improve the reconstructed image as much as expected. However, the result could be questioned in images with a lot of noise, where both PSNR and SSIM shows worse results with higher subsampling ratio, but the images subjectively looks better as seen in figure 4.1.

When performing similar measurements with the SPC, one low frequency image was captured and reconstructed. The images were captured in a near optimal environment, so that a homography could be established between the reconstructed image and the reference image with good precision. In figure 4.15a, we can see that the PSNR and SSIM of the SPC images compared to reference images act as theory predicted and confirmed in simulations. The performance increases non linearly when the subsampling ratio increases. If we look closer at the PSNR plot, we can see that the largest increase in performance is up to 15% subsampling ratio, which can be confirmed when inspecting the images in figure 4.14, where the image quality rapidly improves when increasing subsample ratio up to 15%, then the improvement rate stagnates.

5.1.2 Reconstruction performance using no reference quality assessment

The graph of the simulated reconstruction results in figure 4.4 looks like an inverted version of the PSNR and SSIM graphs in figure 4.2 and 4.3. This results alone are positive for this thesis, because it was unknown if the BRISQUE evaluation method would work well for SWIR images and reconstructed images from the SPC. However, the result is not perfect for images with a lot of noise, where the BRISQUE score gets worse with higher subsampling ratio while the corresponding images subjectively looks better as seen in figure 4.1. We can also see that the reference images score about 20 BRISQUE points better than the best reconstructed images. Consequently even in the best case scenario without noise and a high subsampling ratio, the SPC will not yield as good results as a conventional camera. This evaluation was performed with the specific motivation to evaluate the measurement matrix and reconstruction method, and these results shows that an SPC will not be able to reconstruct images with the same quality as a conventional SWIR camera. Besides these results give a hint of which BRISQUE score can be expected from the SPC in the optimal case. Consequently a BRISQUE score equal to 40 can be seen as the optimum given that noise and post processing have affected the sampled signal.

When studying BRISQUE scores from natural images reconstructed by the SPC in figure 4.16, we can see that the best images score just over 40 BRISQUE points, which is in agreement with the results for simulated images with small or no noise added. This means that the SPC can compare to the benchmark set by the simulation and thus gives a theoretical optimal reconstruction given the measurement matrix and reconstruction algorithm. Furthermore we can see that the trend of the images follows the same characteristics as the simulation in figure 4.5 for different noise levels. Thus we can conclude that simulations give a good indication of where the real images will score, given a certain noise level.

In figure 4.17 to 4.19, we see sample images from three classes given their BRISQUE score and trend as described in section 4.2.2. As the BRISQUE score

tells, the quality of the images should vary a lot, and when taking a closer look the red image set in figure 4.19 it stands out the most. My analysis of why the BRISQUE score and image quality differ is summarized in the following three points.

- The images from the green and yellow sets in figure 4.17 and 4.18, have a similar image quality and lighting, but yet they differ in the BRISQUE score. This might be a property of the BRISQUE classifier, which is built to assess image quality in natural images. If we take a look at the main difference between these two data sets, we can see that one contains pictures of a car, humans, forest and clothing and the other mainly of buildings and large structures with smooth surfaces and low frequencies. These are not so "natural", which can affect the score.
- The major difference between the images in the green and yellow image set compared to the red image set is that the latter appears to contain a lot more global noise. The increase in global noise arises from two separate sources, the first one being the light intensity, we can see that the images in figure 4.18b and 4.19a are practically the same motive, but the latter is darker and noisier. The darker scene was shot in morning when the sun did not illuminate the facade directly. Thus the sampled signal was weaker and the resulting reconstruction was effected more by the sensors background noise and gave rise to global noise in the produced image.
- The second reason for worse BRISQUE score in the red set is large movement in the scene. Most of the images in the red set had movement mainly from clouds when sampled, which definitely increased global noise in the reconstructed images as seen in the results in section 3.5.5.

In the last part of section 4.2.2, the results from plotting normalized noise standard deviation against mean signal intensity in figure 4.20, was presented. The plot give more information on why the BRISQUE score differed between the sets. From the plot it becomes clear at which mean signal intensity we can expect to produce good images given that the background noise becomes insignificant. But in the plots there are only two signals with higher standard deviation than 0.04, which is the threshold where the simulated images started to get both worse initial BRISQUE score and worse trend when increasing the subsampling ratio in figure 4.5. This implies that there probably must be at least one additional factor at play to reduce the image quality in the red image set.

In figure 4.20, we can see that there are a subset of images from the red image set with almost the same normalized noise standard deviation and mean signal intensity as from the yellow and green image sets but yields a worse BRISQUE score anyway. This strengthens the statement the there probably is at least one more factor that reduces reconstruction performance. Furthermore as stated in the third point above, this is probably due to motion in the scene when sampling the signal. Unfortunately for this experiment, it seems like the images containing

motion also had a low mean signal intensity, otherwise we would probably also have "bad" images for stronger mean signals.

The last observation in these plots is the mix of images from the green and yellow image set in the whole mean intensity span, which tells that a strong signal will not yield a good BRISUQE score, which implies that the motive in the images affecting the BRISQUE score as suspected when inspecting the reconstructed image sets.

5.1.3 Dynamics in scene

In this category there are results both from the simulated images and from the SPC, where the results was divided into three characteristic dynamics: small local changes in the scene, large global changes and luminance change.

The effect of local movements on the sampled signal is shown in figure 4.7. We can see that there is no significant difference between the unperturbed reference signal and the distorted signal. The effect on the reconstructed image, seen in figure 4.7c, looks like global noise is added and where the object is moving the image gets extra blurry. The test implies that local movement in a scene will slightly distort the reconstructed image globally and especially locally where the movement occurred. It also tells that local movement is very hard to detect in the signal even if a reference signal is available.

When increasing the movement by letting an object passing through the whole scene, the samples with movement was very easy to spot, which figure 4.8 shows. The effect in the reconstructed image is much worse than for local movement, with a global distortion, as seen in figure 4.9b. In this simple isolated case the image could be saved by removing the measurements when the object was moving and reconstructing an image with fewer measurements. The resulting image would not be as good as the image in figure 4.9b but it would not contain the noise present in figure 4.9c.

In the case of luminance change, the effect on the reconstructed image is even worse than for scenes containing movement, which can be seen in figure 4.11 and by comparing table 4.1, 4.2 and 4.3. Because this problem can not be avoided in natural scenes, method to suppress it was developed and tested with good result. However as can be seen in figure 4.13a, 4.13c and table 4.3, the method will not suppress the effect completely even on a simulation and thus add some global noise in the same form as local movement or signal noise.

When capturing images using the SPC, the luminance change became a larger problem than anticipated. All images captured in natural lighting had luminance

change and it changed at a higher frequency and larger amplitude even in scenes where the intensity seemed stationary. This is of course due to the fact that the intensity change from every mirror in the DMD is summed in the sensor, so even for small changes the sum will change the signal significantly, as seen in figure 4.21. However as seen in figure 4.22, the moving mean method worked well despite the more complex changes to the signal. Considering that this problem was present for all natural images, this method became essential to produce any good result at all. As stated before, this method is a model of global luminance change in the image, and therefore it is hard to know which side effects this method have on image quality. But as the test showed, the method is essential and was used for all images captured by the SPC and presented and evaluated in this thesis.

The moving mean method was mainly constructed because it was known that the SPC would have a long exposure time, but even if the exposure time is reduced to a few seconds or less, there is some indication that the luminance change will still affect the result. In this thesis, the moving mean window corresponded to 50 milliseconds which indicates that the luminance change is so fast that even reducing exposure time significantly could benefit to this method.

Basically all scenes in natural environments contained both dynamics from local movement and luminance change. Local movement often arose from vegetation, objects or clouds moving in the wind but also from turbulence which not move the object but how it is perceived on the DMD. Because of all these dynamic movement existing even in a "static" scene, it was decided not to photograph scenes where large movement occurred i.e. from a car, object or human. Even though such movements could be detected, they could also be mistaken for a luminance change.

As stated, even "static" scenes will with high probability contain both movement and luminance change which will affect the reconstructed images. Therefore it can be concluded that all reconstructed images in this thesis have to some degree, added global noise from local movement and the moving mean algorithm.

5.1.4 Edge response

When comparing the edge response between the conventional camera against the SPC the results were very clear, the conventional camera outperformed the SPC with one to two pixels in distance. I think that there are multiple factors why the results from the SPC differed so much from the conventional camera, and they are listed below,

- The largest impact on image quality is probably the reconstruction algorithm which produces "patches", which can be seen in the SPC images in figure 4.24, especially in the contrast of the white background where the light intensity drops. The "patch" artifact from the reconstruction algorithm

can affect the sharpness of the image. We can also see from previous tests that even from synthetic data, the BRISQUE score is significantly worse for the reconstructed image than for the original image.

- The pixel grid setup in the DMD has two problems that could effect the sharpness. The DMD mirrors are aligned in the diamond shape and in the current setup to fix the ratio and index, two mirrors are merged to form one pixel. The reconstruction algorithm will still interpret the measurement as a regular square pixel, which can distort the image.
- The focus in the DMD is set manually and is possibly not optimal.
- In this thesis no significant image improvement from post processing such as denoising or sharpening was performed unlike in the conventional camera.
- As stated before, with the long exposure, vibrations and light intensity change affected the results, which contribute to global noise in the reconstruction. For example, the SPC could detect significant light intensity change from the halogen lamp powered by a DC-unit.

Even though the image quality is not in favor to the SPC, I think that the results looks promising and can be improved both by changing to more suitable hardware, such as a DMD with "regular" square mirrors and by using more advanced image denoising algorithms.

5.1.5 Subsampling ratio

The first results from section 4.2.5 was the minimum subsampling ratio required to reconstruct a merely recognizable image. It turned out that for the whole image set, the results varied between 2-4%. The difference in value could depend on several factors such as image complexity, SNR and dynamics in the scene, which everyone contribute to add uncertainty to the equation system.

One possible application that take advantage of sampling an image with minimum subsampling ratio is to construct a test image and calibrate the camera for the longer exposure high quality image. The calibration could for example be used to set focus, determine which pixels to include in the image and calculate which subsampling ratio to use given the complexity of the image. The test image is sampled with fewer samples and thus reconstructed faster relative to the high quality images.

In the second part of section 4.2.5, a set of images reconstructed with different subsampling ratio was presented. The result is presented for the reader to obtain a more concrete perception of the generated image quality and a supplement to the numerical results given subsampling ratio, but also overall expected image quality.

The results show that a stable reconstruction was obtained at 5% subsampling ratio for all images and that a higher subsampling ratio increased the image quality. With greater subsampling ratio than 15%, the images start to reach their highest potential where details start to show up and the images get sharp. Some global noise is present in all images. This noise is probably mainly due to the long exposure time, where dynamics in the scene play a big role.

5.2 Method

The methods used in thesis can be divided into four categories, the SPC hardware, the sampling matrix and reconstruction, post processing and the method used to capture the images. Two of these categories, the hardware setup and the sampling matrix and reconstruction were heavily influenced and implemented by widely accepted methods from articles and experiments. The other two categories, the post-processing and the image capturing, depended more on the hardware limitation and competence achieved from the university.

The first method in the chain was feeding the DMD measurement matrices and sampling the signal. Because of the interface to the DMD, which acted as a second monitor to the computer, the method to stream the measurement matrices as a video was thought to be a good method because it was easy to implement. In the early stages of this thesis work a much smaller target resolution of 128×128 pixels was set and thus significantly fewer measurement matrices needed to be streamed to the DMD while having the same exposure time. With that initial goal in mind, the video player streaming method would have worked well. But when the target resolution was pushed, the video players frame rate had to be pushed to its limits in an application it was not designed to handle. This resulted in a higher probability that the video player would go out of sync and thus ruin the reconstruction and there was no way of knowing if this had happened during the sampling. By the time this issue was discovered there was no time left to implement a new method, although it was much needed.

The sampling matrix chosen in this thesis was the permuted sequency ordered Walsh Hadamard measurement matrix. This sampling matrix together with the implemented fast transform in the reconstruction algorithm enabled the huge increase in image resolution. The method of using structurally random matrices is the only feasible method today to enable high resolution and fast reconstruction with low memory usage in the computer. Thus, if implemented optimally both the feeding of measurement matrices to the DMD and reconstruction could be calculated in run time in an agile and memory efficient program.

The reconstruction algorithm TVAL3 was used throughout this masters thesis work and was chosen after the literature study where several articles mentioned

total variation as a good optimization algorithm for compressive imaging. The algorithm worked as described and according to the developer it is one of the fastest and best reconstruction algorithms that are free to use and available. The only negative criticism of using this algorithm is that it seems to produce a dark tint in the edges of the reconstructed images.

In the post-processing quite basic signal and image processing was performed and was intentionally designed that way in order to present the result as true as possible. One algorithm that was developed specific for the SPC, was the moving mean algorithm to suppress the impact of illumination change in the sampled signal. The algorithm showed great results and became essential when taking photos of scenes outdoor. Both hardware and software signal denoising was implemented to the architecture, but had to be changed every time the DMD update rate was changed, which happened a couple of times. Finding a new solution every time was a time consuming task and dropped completely in the end when good results were achieved anyway. If there had been more time or a target DMD pattern rate had been set earlier, a good signal denoising implementation could have improved the result.

5.2.1 Replicability, reliability and validity

Given the methods in this thesis, I think it would be quite straight forward to replicate this experimental setup and architecture. The setup is quite simple and the software developed is not so heavy, therefore I think the experiments have good replicability.

In the case of compressive imaging, the reliability and validity go hand in hand, if what's being measured is not the correct data, the reconstruction will fail and if the reconstruction succeeds, the measurement must be correct. Therefore compressive sensing is very binary, either you get it right or you fail and thus I think my results have both high validity and reliability.

6

Conclusion & Future Work

6.1 Conclusion

In this thesis a complete compressive imaging SWIR SPC architecture was implemented and evaluated. The aim was to find out which image quality could be achieved in natural images captured in daylight. The results produced in this thesis both presented evaluation from simulated images and images captured by the SPC to show how the chosen sampling method and the reconstruction algorithm performed and how the whole architecture performed in unison.

The sampling strategy using the structural random matrices method with sequence ordered Walsh Hadamard measurement matrix solved the problems of scaling the reconstructed images to high definition photos and enabled the image resolution 512×512 pixels, which is, the same resolution as the state of the art reference SWIR camera.

The total variation solver TVAL3 was used as reconstruction algorithm which took advantage of SRM with FWHT to reconstruct the images fast and with good preservation of edges. Using the chosen sampling and reconstruction method could potentially make for a lightweight sampling and reconstruction procedure with few variables stored in memory and calculations made in real time.

Feeding the measurement matrices to the DMD through a HDMI cable and video software was an easy method to start with. The first set of measurements was performed with low frame rate and blank frames for control. But when maximizing the frame rate the risk of duplicate or frame drops increased which made it hard to pair each measurement to the correct measurement matrix. This

method need to be replaced for more control and faster sampling rates.

For post-processing, an algorithm to correct the signal affected by luminance change was implemented. This algorithm showed to be of utmost importance to this thesis. Long exposure time in natural scenes always gave a significant DC change in the sampled signal which should be stationary. Results showed in both simulations of luminance change and real scenes a significant improvement of image quality when using this method. The analysis also indicate that the algorithm may be relevant even if the exposure time gets reduced to near a second.

The resulting images produced by the SPC showed that high quality and high resolution images can be acquired in natural daylight scenes. In good conditions with sufficient intensity to overcome the sensors background noise and relatively stationary scenes, detailed images where even people could be identified, could be reproduced. Given the result and further work and improvement of the hardware, compressive imaging and the SPC architecture have potential of real world applications.

The resulting reconstructed images from both simulations and the SPC was evaluated with a range of methods. When a reference image was available, standard image evaluation techniques, PSNR and SSIM was used. Experiments in the subsampling range 5-30% that was used throughout the thesis, showed that image quality increased with increasing subsampling ratio, but stagnated around a subsampling ratio between 15-20%. The BRISQUE no reference image quality assessment algorithm based on statistics of "naturalness" in the images proved useful. It could be seen that the SPC could get the same results as the best results in the simulations. This indicates that the sampling and reconstruction is the main source of image degradation and that the SPC hardware in right conditions does not affect the resulting image quality negatively.

The edge response algorithm calculates the sharpness of the image and therefore comparison between cameras and image processing methods is easily performed. This evaluation showed that the reference SWIR camera performs a bit better than the SPC. Because the SWIR SPC is a completely different camera than a conventional camera and has a different purpose and application than a regular camera a good evaluation is also to present the produced images to get a subjective view of the quality and what is good enough.

To summarize the whole thesis, the research questions are answered:

- The image quality of the produced images can be evaluated with the same methods used when evaluating a regular camera. In the case of the SPC, the sampling matrix and reconstruction algorithm can be evaluated inde-

pendently of the SPC hardware as well.

- The state of the art method to capture and reconstruct images using an SPC architecture is to sample all the measurement matrices as fast as possible using structurally random matrices and fast transforms in the reconstruction algorithm.
- The image quality achieved using state of the art methods applied to the SPC is high resolution images with good enough quality to be used in real world applications.

6.2 Future work

This thesis shows that there is possible to use the SPC architecture to capture and reconstruct natural scenes in daylight, but for the technology to be used in a realistic application some improvements should be made. I think that the most crucial problems will disappear by using more sophisticated hardware.

As identified multiple times in this thesis, the largest contributors to decreased image quality are exposure time and noise. The exposure time can be decreased by a faster DMD, today there exist DMD:s with 8 times the maximum pattern rate of the DMD used in this thesis (which was operated at half the maximum speed), which would enable an exposure time of 1.125 seconds at 10% subsampling ratio. This upgrade would however require a new approach to feeding the measurements matrices to the DMD because of the limitations of the HDMI cables transfer speed.

Either a new transfer approach is implemented or not, a more sophisticated method of generating each measurement matrix should be implemented. The current method of generating all measurement matrices offline to a video file and then playbacked by a third party video player, has a limitation of FPS but also a reliability problem. The software was not designed to necessarily show every frame in the video file, which is a problem for the SPC, where each measurement needs to be paired with the correct measurement matrix. My suggestion would be a program which would calculate all matrices at start up and upload them to video memory ready to display at the DMD. This solution would also be more agile where, subsampling ratio, type of measurement matrix and exposure time could easily be adjusted prior to sampling.

The second hardware upgrade with the aim to reduce noise, would be to replace the photo diode used in this thesis to a more appropriate sensor for the application. The new sensor should have less background noise and the surface area of the diode should be analyzed to match the rest of the architecture to maximize the incoming light onto the sensor through the focusing lens. This upgrade

could also unlock the true potential of the SWIR spectrum where images could be captured in dark environments illuminated by the moon, stars and night glow.

New research shows promising results by adding a second sensor which measures the intensity of the mirrors representing zero or turned away from the sensor and being dumped in the current architecture. The compliment of each unique measurement matrix could also be seen as a unique measurement matrix and thus more information can be collected in each measurement and reduce the subsampling ratio needed to reconstruct a corresponding single sensor image.[10]

The last future work proposition will not increase image quality but would simplify research or usage and time consumption taking pictures with the SPC. The second and third most time consuming task after exposure time is the complexity of sampling the scene and the reconstruction algorithm in the post-process. A fully integrated capturing and reconstruction chain with a GPU accelerated reconstruction algorithm should ease the work and save a lot of time for the user.

Bibliography

- [1] Fast compressive sampling with structurally random matrices. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3369–3372, 3 2008. doi: 10.1109/ICASSP.2008.4518373. Cited on page 6.
- [2] Mittal Anish, Moorthy Anush Krishna, and Bovik Alan Conrad. No-reference image quality assessment in the spatial domain. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 21:4695 – 4708, 12 2012. ISSN 1057?7149. Cited on pages 7 and 28.
- [3] Al Bovik. *The Essential Guide to Image Processing*. Elsevier, 2009. ISBN 978-0-12-374457-9. Cited on pages 7 and 28.
- [4] C. Brännlund and D. Gustafsson. Single pixel swir imaging using compressed sensing. *Symposium on Image Analysis, 14-15 March 2017, Linköping, Sweden*, 2017. Cited on page 6.
- [5] T. Do. Fast compressive sampling using fast compressive sampling using structurally random matrices, 4 2008. Cited on page 16.
- [6] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 4 2006. ISSN 0018-9448. doi: 10.1109/TIT.2006.871582. Cited on pages 2, 5, 17, and 18.
- [7] Laurent Duval. 50mm fl swir fixed focal length lens, 2017. URL <https://www.edmundoptics.com/imaging-lenses/fixed-focal-length-lenses/50mm-fl-swir-fixed-focal-length-lens/>. Cited on page 12.
- [8] T. Edeler, K. Olinger, S. Hussmann, and A. Mertins. Multi image super resolution using compressed sensing. *ICASSP*, pages 2868–2871, 2011. Cited on page 6.
- [9] Michel Elad. *Sparse and Redundant Representations*. Springer, New York, New York, 2010. Cited on pages 5 and 15.
- [10] F. Soldevila et al. Computational imaging with a balanced detector. *Nature: Scientific Reports*, 6, 2016. Cited on page 70.

- [11] Erik Fall. Compressed sensing for 3d laser radar, 2014. Cited on page 6.
- [12] L. Gan, T.T.Do, and T.D Tran. Fast compressive imaging using scrambled block hadamard ensemble. In *2008 16th European Signal Processing Conference*, pages 1–5, 8 2008. Cited on page 6.
- [13] B. Chatterjee L. McMackin, M. A. Herman and Matt Weldon. A high-resolution swir camera via compressed sensing. *SPIE*, 8353, 2012. Cited on page 6.
- [14] Chengbo Li. An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing. 2009. Cited on pages 6, 17, 18, and 19.
- [15] F. Näsström, D. Bergström, F. Bissmarck, P. Grahn, D. Gustafsson, and J. Karlholm. Prestandardmått för sensorsystem. 11 2015. ISSN 1650-1942. Cited on pages iii, 7, and 28.
- [16] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee. Compressive sensing: From theory to applications, a survey. *Journal of Communications and Networks*, 15(5), 2013. Cited on page 5.
- [17] Irina Rish and Genady Ya. Grabarnik. *Sparse Modeling*. CRC Press, Boca Raton, 2015. Cited on pages 5 and 14.
- [18] D. Takhar, J. N. Laska, M. B. Walking, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk. A new compressive imaging camera architecture using optical-domain compression. *IS&T/SPIE Computational Imaging*, 5, 2006. Cited on pages 5 and 18.
- [19] D. Takhar, J. N. Laska, M. F. Duarte, K. F. Kelly, R. G. Baraniuk, and M. A. Davenport. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 83, 2008. Cited on pages 6 and 18.
- [20] *DLP LightCrafter 4500 Evaluation Module, User's Guide*. Texas Instruments Incorporated, 9 2015. URL <http://www.ti.com/lit/ug/dlpu011e/dlpu011e.pdf>. Cited on page 11.
- [21] *Large Area InGaAs Amplified Photodetectors PDA20C Operation Manual*. Thorlabs GmbH, 1.1 edition, 9 2016. URL https://www.thorlabs.com/drawings/e22761fd83df1eab-B1D1D2AA-C921-6D8E-D2B14E11544D8225/PDA20C_M-Manual.pdf. Cited on page 13.
- [22] T.T.Do, L. Gan, N.H. Nguyen, and T.D Tran. Fast and efficient compressive sensing using structurally random matrices. *IEEE Transactions on Signal Processing*, 60(1):139–154, 01 2012. ISSN 1053-587X. doi: 10.1109/TSP.2011.2170977. Cited on pages 6, 17, and 19.

- [23] Michael B. Wakin, Jason N. Laska, Marco F. Duarte, Dror Baron, Shriram Sarvotham, Dharmpal Takhar, Kevin F. Kelly, and Richard G. Baraniuk. An architecture for compressive imaging. In *IEEE International Conference on Image Processing (ICIP), October 8-11, 2006*, pages 1273–1276, 2006. Cited on page 5.
- [24] C. Zhuoran, Z. Honglin, J. Min, W. Gang, and S. Jingshi. An improved hadamard measurement matrix based on walsh code for compressive sensing. In *2013 9th International Conference on Information, Communications Signal Processing*, pages 1–4, 12 2013. doi: 10.1109/ICICS.2013.6782833. Cited on pages 6, 16, and 17.