

Compressed Sensing: Single Pixel Imaging in Short-Wave Infrared Spectrum

Examensarbete

TQET33

Andreas Brorsson, Andbr981, (911023-1215)

Linköping, 2017

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 available broadly either within the military or 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 two. The camera uses a micro mirror 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 matrixes 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.

Sammanfattning

Foton tagna i det korta infraröda spektrummet är intressanta i militära sammanhang på grund av att de är oberoende av vilken tid på dygnet det är tagna för att solen, månen, stjärnor och nattsken (night glow) lyser upp jorden med kortvägg infraröd strålning dynget runt. Ett stort problem med dagens kortvägg infraröda kameror är att de är väldigt dyra att producera och därav inte tillgängliga till en bred skara varken inom militären 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 behöver en bråkdel mätningar relativt antal pixlar som ska återskapas och reduserar kostnaden på en kortvägg infraröda kamera med en faktor av två. Kameran använder en mikrospegelmatris som används för att välja vilka speglar (pixlar) som ska mäts 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 av en enpixel-kamera. Med en enpixel-kameraarkitektur utvecklad av FOI var målet med detta examens arbete att ta fram metoder för att sampla, återskapa bilder och utvärdera deras kvalitet. Denna avhandling visar att strukturella slumpade matriser och snabstransformer måste användas för att öppna upp för högupplösta bilder och snabbar upp processen att rekonstruera bilder avsevärt. Utvärderingen av bilderna kunde göras med vanliga mått assosierade med kamerautvärdering och visade att att kameran kan återskapa högupplösta bilder med relativt hög bildkvalitet i dagsljus.

Contents

1	Introduction	5
1.1	Background	5
1.2	Compressive sensing & imaging	5
1.3	Motivation	6
1.4	Aim	7
1.5	Research questions	7
1.6	Limitations	7
1.7	Thesis outline	7
2	Related work	8
2.1	Compressive sensing	8
2.2	Compressive imaging	8
2.3	Measurement matrix & reconstruction	8
2.4	Evaluation	9
3	Method	10
3.1	Single pixel camera architecture & hardware	10
3.1.1	Newtonian telescope	10
3.1.2	DLP and DMD	11
3.1.3	Lens	12
3.1.4	Single pixel sensor	12
3.1.5	Signal spectrum	13
3.2	Compressive imaging	13
3.3	Measurement matrix & Restricted isometry property (RIP)	14
3.3.1	Permutated sequency ordered Walsh Hadamard measurement matrix	14
3.4	Reconstruction method	15
3.4.1	Total variation: TVAL3	16
3.5	Image Capturing and processing chain	17
3.5.1	Prepare the SPC	18
3.5.2	Sampling	18
3.5.3	Streaming patterns to the DMD	18
3.5.4	Signal processing	19
3.5.5	Dynamics in scene	20
3.5.6	Reconstruction	22
3.5.7	Image processing	22
3.6	Evaluation: Image quality assessment	22
3.6.1	Evaluation Using reference image	23
3.6.2	Evaluation Using no reference quality assessment	24
3.6.3	Evaluating Using Edge response	24
3.7	Method criticism	25
4	Evaluation	26
4.1	Simulated Results	26
4.1.1	Reconstruction performance Using reference image	26
4.1.2	Reconstruction performance Using no reference quality assessment	29
4.1.3	Dynamics in scene	31
4.2	SPC evaluation	37
4.2.1	Reconstructed performance Using reference image	37
4.2.2	Reconstruction performance Using no reference quality assessment	38
4.2.3	Luminance Change in scene	41

4.2.4	Edge response	43
4.2.5	Subsampling ratio	45
5	Discussion	50
5.1	Result	50
5.1.1	Reconstruction performance Using reference image	50
5.1.2	Reconstruction performance Using no reference quality assessment	50
5.1.3	Dynamics in scene	52
5.1.4	Edge response	53
5.1.5	Subsampling ratio	53
5.2	Method	54
5.2.1	Replicability, reliability & validity	54
6	Conclusion & Future Work	56
6.1	Conclusion	56
6.2	Future work	57

1 Introduction

The development and research of compressed sensing applied to a single pixel camera (SPC) is a relative new area in signal processing with the first functioning camera architecture in 2006. Since then numerous improvements and methods have been proposed how to capture images. In this section a introduction to the SPC architecture and a brief introduction of compressed imaging 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, star light and airglow thus works both by day and night. SWIR light can to some extent pass through smoke and fog which makes it 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 array (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 & 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. Fourier or wavelet 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.

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 such 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: 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 2 light from the scene enters through the aperture in the camera where the primary mirror focus the light 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 and not being measured.

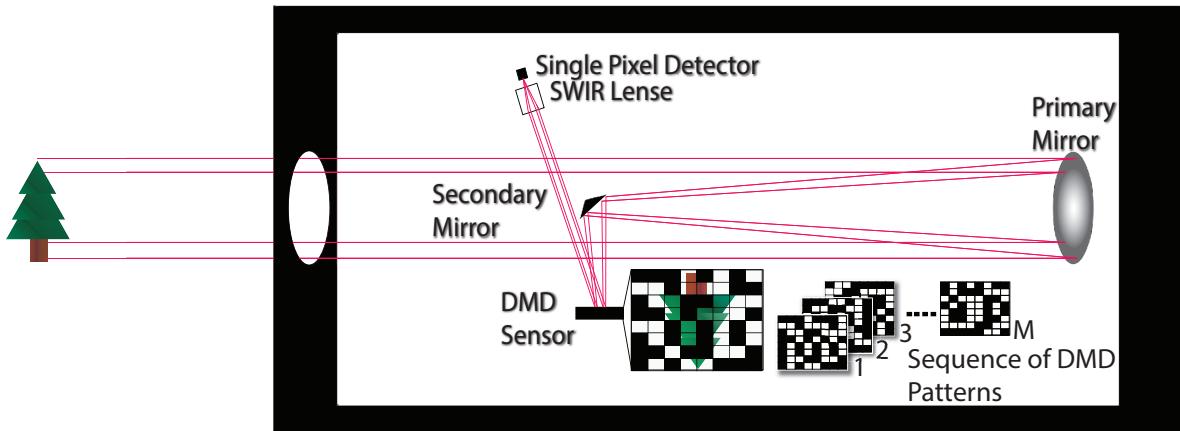


Figure 2: System overview

To connect the CS theorem 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 the performed by the next measurement matrix until all desired measurements are done, as shown if figure 2.

When all the measurements are sampled an optimization algorithm calculates the most significant coefficients of the image in an other more sparse basis, for example the wavelet basis by transforming the complete measurement matrix. Because the image is more spare or compressed in an other basis it is easier to find the solution to the equation system in that basis and 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 a 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, and the reconstruction shifts burden to the processor. There are two major reasons why a SPC is of interest, it is not to 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 spectrum of light like SWIR or Terahertz (X-ray) wavelengths where the image sensors are hard to build which 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 which can be reduced by a factor of 100 with a 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 that building.
- The reconstruction algorithm will not be developed in this thesis 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 present 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 is 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 is 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

- [1], [2] "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. The books presents different minimization algorithms and how to implement them.
- In [3] "Compressed sensing" David L. Donoho proposed the framework of compressed sensing and the application of images capturing.
- [4] "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

- [5], [6] "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.
- [7] "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.
- [8] "Single Pixel SWIR Imaging using Compressed Sensing" by C. Brännlund and D. Gustafsson. 2016, was written to show the initial results and proof of concept of the SPC architecture at FOI.
- [9] "A high resolution SWIR camera via compressed sensing" is a paper from L. McMackin et al. 2012 at Inveiw 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.
- [10] "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.
- [11] "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.

2.3 Measurement matrix & reconstruction

- [12] 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 Li constructed which solve the CS problem. The algorithm is faster and produces better results for images than previous popular algorithms.

- [13]–[15] 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 i.i.d random matrix. With SRM the reconstruction time is reduced by replacing matrix multiplication with fast transforms. In addition to improved reconstruction time the new method does not need to store the measurement matrix in memory which enables reconstruction of high resolution images.
- [16] "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

- [17] "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 PSNR and SSIM which can be used when a reference image is available.
- [18] "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.
- [19] "Prestandamått för sensorsystem" 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, which can also be used to check that the patterns are displayed correct on the DMD and focusing the image.

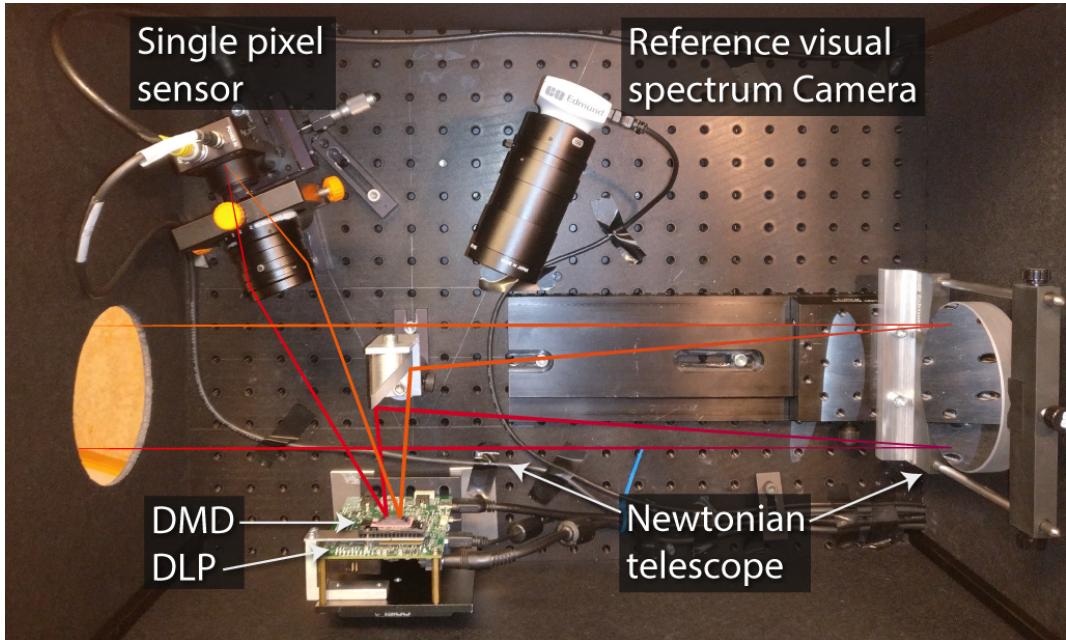


Figure 3: The single pixel camera architecture used in this thesis. In the image the aperture, reflective optics, DMD, reference camera and the single pixel sensor are shown from an areal view including red lines showing the incoming light path.

As seen in figure 3 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 an 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. 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 at $120 \times 24 = 2880$ 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.

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 4 the indexes of the mirrors column is two mirror column arrays wide while a row is a single row.

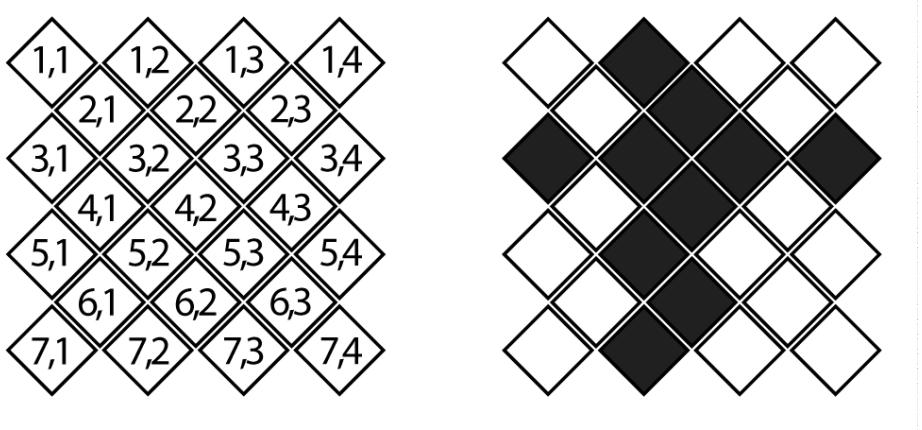


Figure 4: DMD matrix mirror index, left shows each tiles index and right shows the second row and second column in black as set by default.

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 5. 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.

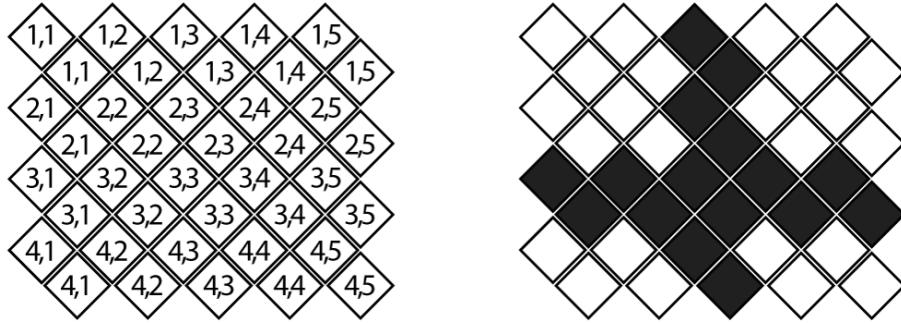


Figure 5: DMD matrix, left shows each tiles index and right shows third row and third column in black.

Mathematically the DMD is a binary operator which lets light pass or not, in figure 6 a typical pattern that could be sent to the DMD is shown.

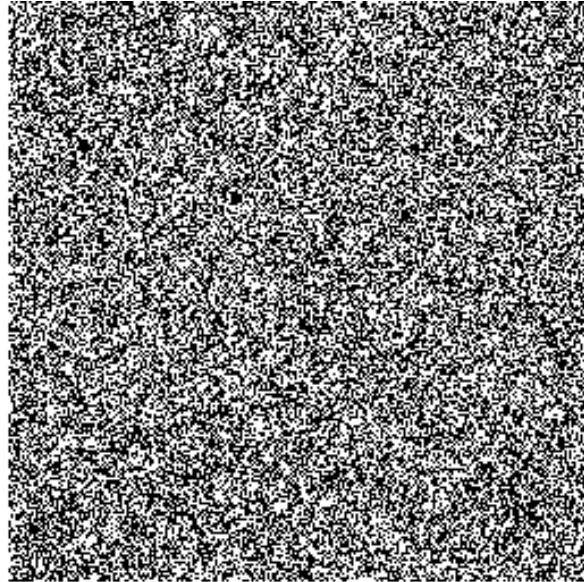


Figure 6: A typical pseudo random measurement matrix sent to the DMD with the resolution 256×256 pixels.

3.1.3 Lens

The lens mounted on the single pixel sensor is an 50 mm SWIR Fixed Focal Length Lens with an variable appature from f1.4 designed for wavelengths ranging from the 800 nm in the visual spectrum to 2000 nm in the SWIR spectrum. [20]

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 sensor outputs an analog signal in volt which the sampler converts to a discret 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, but 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 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 more detailed explained in section 3.3.

The single pixel sensor captures a scene by measuring the light intensity focused into 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

$$\mathbf{y} = \Phi \mathbf{x} + \epsilon, \quad (1)$$

where $\mathbf{x}_{N \times 1}$ is the image arranged 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 contain 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

$$\Psi \theta = \mathbf{x}, \quad (2)$$

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 equation 2, equation 1 can be expand to

$$\mathbf{y} = \Phi \mathbf{x} + \epsilon = \Phi \Psi \theta + \epsilon = \mathbf{A} \theta + \epsilon, \quad (3)$$

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

The revelation in equation 3 is what makes CS powerful, by sampling the scene using the complete measurement matrix Φ (as equation 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 is not sparse.[1]

A great advantage CI has over regular cameras where each pixel is sampled separately is that roughly half the pixels are 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 & 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 $\Phi \Rightarrow \Phi\Phi = \mathbf{A}$, the constraint is defined as,

$$(1 - \delta_K) \|\mathbf{x}\|_{\ell_2}^2 \leq \|\mathbf{Ax}\|_{\ell_2}^2 \leq (1 + \delta_K) \|\mathbf{x}\|_{\ell_2}^2, \quad (4)$$

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 are 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 $M \geq O(K \log(N/K)) \ll N$. [2]

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 the 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 has a higher number of measurements for exact reconstruction of the image, $M \sim (K N s) \log^2(N)$, where s is the average number of non zero indexes in the measurement matrix [22]. Research has shown that there is no significant loss in recovery of the image relative the i.i.d random measurement matrix [16] and an other property of PSOWHM is that it only contains -1 and 1 which 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 ,

$k \in \mathbb{N}$ are constructed by the recursive formula

$$H_0 = 1, \quad (5)$$

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (6)$$

and in general,

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = H_1 \otimes H_{k-1} \quad (7)$$

where \otimes denotes the Kronecker product. To construct the permuted sequency ordered Walsh Hadamard matrix from the naturally ordered Hadamard matrix three steps is required:

- Convert row index to binary.
- Convert the binary row index to gray code.
- Apply bit reverse on the gray code index.
- Remove first row with only ones.
- Permute columns and choose M row at random.

then order the rows after the bit reverse to obtain the sequency ordered Walsh Hadamard matrix.

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 1: 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

for example

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \Rightarrow W_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \quad (8)$$

To use the sequency ordered Walsh Hadamard matrix as an measurement matrix the fist row is omitted, permutations to the columns is performed, M rows are choosen at random and the indices with -1 is shifted to 0 . How the matrix was permuted and which rows was choosen i which order is stored so the reconstruction algorithm can use that information to reverese the process. This method is used to distribute the energy of the signal's sample across all measurements. [12], [14], [16].

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

$$\hat{\theta} = \arg \min \|\theta\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\theta. \quad (9)$$

Simply 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),

$$\hat{\theta} = \arg \min \|\theta\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}\theta\|_2 < \epsilon. \quad (10)$$

In 2006 Donoho [3] 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 . [3], [6], [7]

3.4.1 Total variation: TVAL3

The reconstruction algorithm that was chosen in this thesis was a total variation regularization algorithm called TVAL3 and was chosen for its speed and good results in image reconstruction compared to other reconstruction algorithm created for the CS problem [12]. Natural images often contain 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 sparse instead of the signal, thus finding the sparsest gradient.

The TV optimization problem in TVAL3 is defined as

$$\min_{\mathbf{x}} \sum_i \|D_i \mathbf{x}\|, \text{ subject to } \Phi \mathbf{x} = \mathbf{y}, \mathbf{x} \geq 0, \quad (11)$$

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 introduce a penalty term. To solve the new subproblems the alternating direction method is used [12].

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,

$$\Phi \mathbf{x}^k \text{ and } \Phi^\top (\Phi \mathbf{x}^k - \mathbf{y}). \quad (12)$$

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),

$$\mathbf{W} \mathbf{x} = \text{fwht}(\mathbf{x}), \quad (13)$$

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.

From 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 1. Column permutation operator $\pi(\cdot)$, permutes the order of the columns in a matrix or a vector from a random seed.

Definition 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 13, definition 1 and 2,

$$\Phi = \pi(\Pi_M(\mathbf{W})) = \Pi_M(\pi(\mathbf{W})). \quad (14)$$

Note that it does not matter in which order the functions π 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 as showed in equation 15,

$$\pi(\Phi)\mathbf{x} = \Phi\pi(\mathbf{x}), \quad (15)$$

With all observations combined the matrix multiplication is replaced with the fwht and operators in definition 1 and 2 as shown in equation 16,

$$\mathbf{y} = \Phi\mathbf{x} = \pi(\Pi_M(\mathbf{W}))\mathbf{x} = \Pi_M(\mathbf{W})\pi(\mathbf{x}) = \Pi_M(\mathbf{W}\pi(\mathbf{x})) = \Pi_M(\text{fwht}(\pi(\mathbf{x}))). \quad (16)$$

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 of the complete measurement matrix in the reconstruction unlocks the potential to reconstruct images with high resolution (512×512 pixels and larger). [12], [14]

3.5 Image Capturing and processing chain

In figure 7 the whole process of capturing an image is presented with all subsystems and signal/image processing steps included.

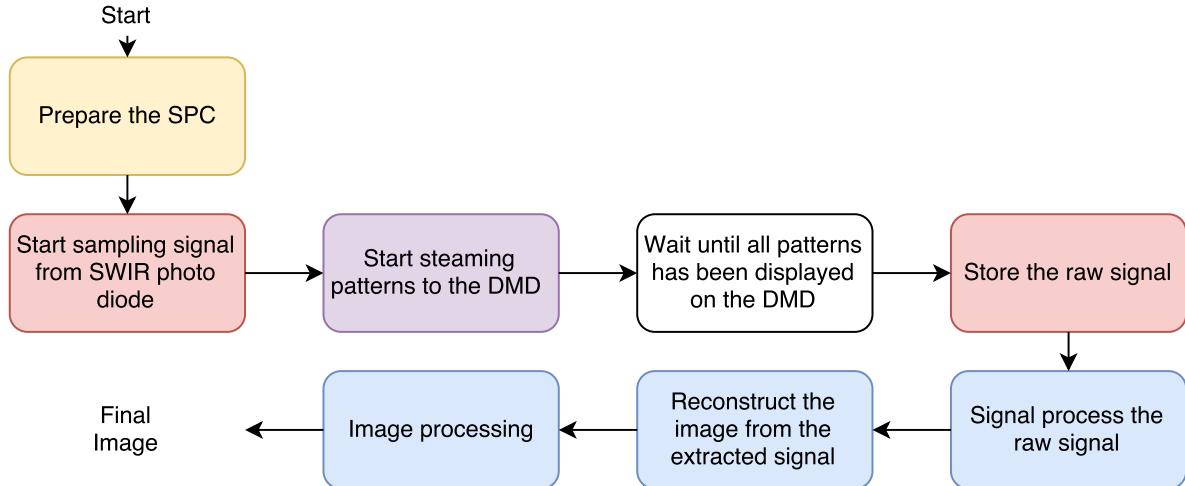


Figure 7: Block diagram of image capturing and processing chain, from signal acquisition to final image. Each color represents different subsystems in hardware or software.

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 7 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 7) is making sure 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) 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 7) 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 sampling rate such that every measurement has several sampling points thus oversampling 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, more on that in 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 7) represented in purple in the block diagram 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 a 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 can be achieved. The naive approach would be to display one pattern per frame which is linked to the frame rate of the DMD, lets say for example 60 frames per second (fps) then for a 512×512 pixel large image subsampled at 20% gives 52428 patterns which would take $52428/60 = 874$ seconds = 14.5 minutes to sample which 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, so for each frame received 24 patterns are encoded on that frame then the DMD software isolates each bit plane and displays them in consecutive order. This function improves the naive implementation by a factor of 24, which reduces the time to sample the image from the last example from 874 seconds to $874/24 = 36$ seconds. That long exposure time is of course not optimal for natural images outdoors but acceptably 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 for each group of 8 unique patterns drawn from the rows of the measurement matrix are stacked in each bit plane of a 8 bit image as seen in figure 8.

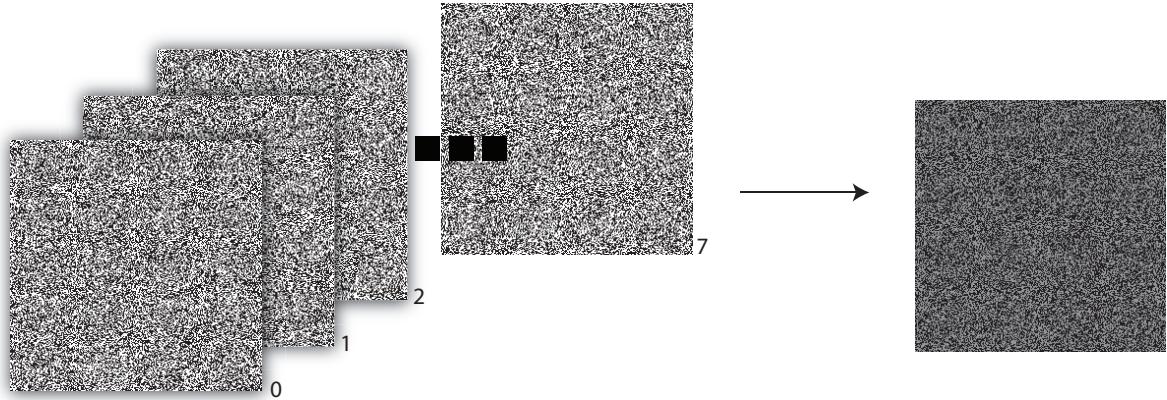


Figure 8: Stack each group of 8 measurement matrices in separate bit planes creating one 8 bit image with each matrix in one bit plane.

Then for each group of three 8 bit images a 24 bit color image is constructed as seen in figure 9.

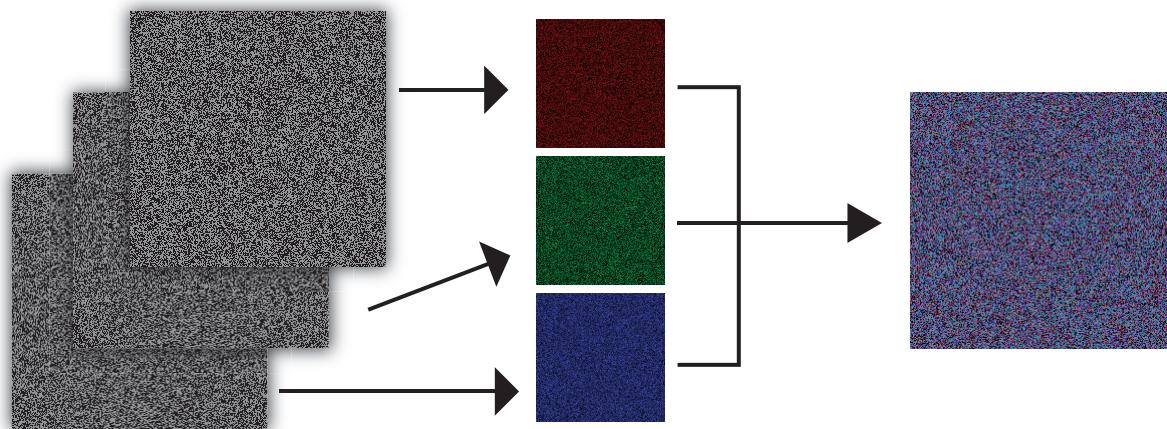


Figure 9: Stack each group of three 8 bit plane images into one 24 bit color image. This is one frame in the video sent to the DMD.

The 24 bit color image corresponds to one frame in the video, to create the video this is done so each pattern is represented in the video.

3.5.4 Signal processing

When the sampled signal is stored on the computer the remaining signal/image processing and reconstruction represented by blue blocks in figure 7 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 correspond to one measurement in signal \mathbf{y} . This is done by first find every set of indices corresponds to every measurement matrix, as seen in figure 10 where the signal indices which corresponds to one measurement matrix is isolated by the purple lines.

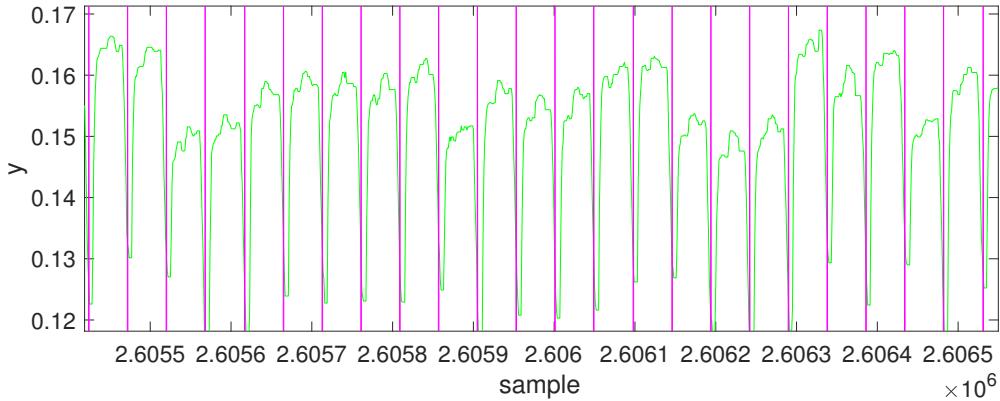


Figure 10: A simulated noisy over sampled signal \mathbf{y} where each sample in $\mathbf{y}[m]$ is represented by multiple samples. The purple lines separating each measurement which corresponds to one measurement matrix.

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 seen in figure 10 where samples near the transition between DMD patterns (purple lines) is uncertain. For remaining samples that corresponds to one measurement matrix the mean is calculated and set to the value for each sample $\mathbf{y}[m]$, as seen in figure 11.

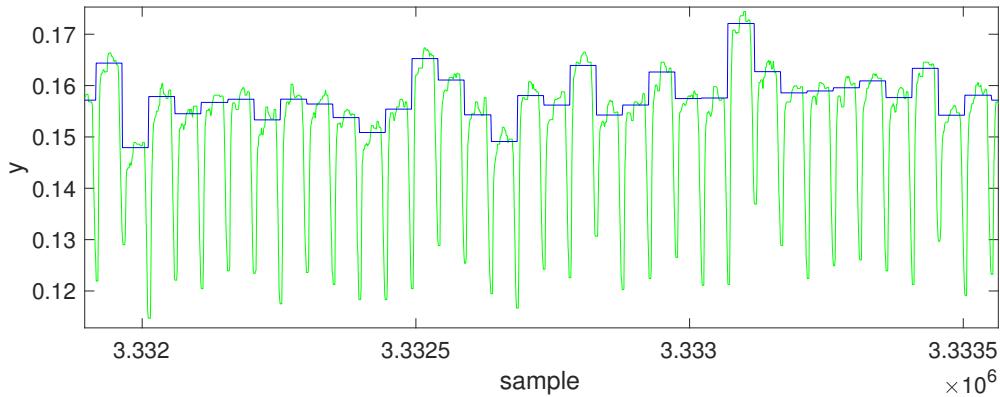


Figure 11: Calculated mean value for each measurement matrix with transition measurements omitted.

3.5.5 Dynamics in scene

So far the measurement vector \mathbf{y} has been determined, in the ideal case even with added noise, the measured signal should be stationary because the image (scene) is assumed to be static. But when capturing images outdoors with natural light in addition with long exposure time as described in section 3.5.3 it is not certain that the image (or every pixel) is constant over the exposure time which will reduce reconstruction performance because of the ambiguity of each pixel. 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 with corresponding algorithm to suppress the impact on the reconstructed image. The object movement problem will not be modeled but will be avoided when sampling the signal by making sure that the scene is as static as possible.

In natural outdoor images it can be assumed the 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. Therefor a model of light intensity change is created with a algorithm to restore the signals stationary characteristics.

With the assumption that the image is constant and the luminance change uniformly adds the same intensity to each pixel per measurement the problem can be modeled.

Start of with the original theorem and disregard the noise,

$$\mathbf{y} = \Phi \mathbf{x}, \quad (17)$$

the image \mathbf{x} can not longer be considered constant for all measurements, the luminance change will change image \mathbf{x} for every measurement matrix ϕ_i depending on the uniform luminance change. This can be described for one measurement as,

$$y_i = \phi_i \mathbf{x}_i = \phi_i(\mathbf{x} + \mathbf{l}_i) = \phi_i \mathbf{x} + \phi_i \mathbf{l}_i, \quad (18)$$

where \mathbf{l}_i uniform 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,

$$y_i = \phi_i \mathbf{x} + \phi_i \mathbf{l}_i = \phi_i \mathbf{x} + \frac{N}{2} c_i, \quad (19)$$

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

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{c}, \quad (20)$$

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 assumes that the rate of change in intensity has a much lower frequency than the intensity change between individual measurement matrices, then \mathbf{c} can be approximated by the moving mean and simply 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 i chosen depending on the DMD pattern rate.

Moving mean is defined as,

$$\mathbf{y}_{MM}[m] = \frac{1}{k} \sum_{i=m-\frac{k}{2}}^{m+\frac{k}{2}} \mathbf{y}[i] \quad (21)$$

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

$$\mathbf{y} = \mathbf{y}_{SAMPLED} - \mathbf{y}_{MM} \approx \mathbf{y}_{SAMPLED} - \mathbf{c}. \quad (22)$$

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 in measurement matrix Φ , sampled signal \mathbf{y} and algorithm settings as arguments and outputs the reconstructed image. The setting used throughout all experiments is:

- $opts.\mu = 2024$
- $opts.\beta = 64$
- $opts.\maxcnt = 10$
- $opts.\maxit = 1000$
- $opts.\text{tol_inn} = 10^{-5}$
- $opts.\text{tol} = 10^{-10}$
- $opts.\mu_0 = 2^4$
- $opts.\beta_0 = 2^0$
- $opts.\text{nonneg} = \text{true}$
- $opts.\text{isreal} = \text{true}$

Which solves for a real non negative solution shown in equation 11.

3.5.7 Image processing

From the reconstructed image some light image processing is performed. There are only two operations applied to the reconstructed image and the reason why, is that the images presented in the results should represent what can be expected from the system. Furthermore often image processing is applied on special problems or artifacts in the images and it is not desired to cover up if such artifacts exist. Therefore the two operations used is 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 is reconstructed with a high intensity peak which not correlates to 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* will be used.

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

3.6 Evaluation: Image quality assessment

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

All results is 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 two possibilities 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.

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 $\hat{\mathbf{x}}$. Because the measurement matrix and reconstruction algorithm is independent of the SPC hardware the subsystem can be evaluated independently. Two advantages of evaluation 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

$$\text{PSNR}[f(x, y), g(x, y)] = 10 \log_{10} \frac{E^2}{\text{MSE}[f(x, y), g(x, y)]} \quad (23)$$

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

$$\text{MSE}[f(x, y), g(x, y)] = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} [f(x, y) - g(x, y)]^2. \quad (24)$$

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. SSIM index is defined as

$$\text{SSIM}[f(x, y), g(x, y)] = \sum_n l[f(n), g(n)]^\alpha c[f(n), g(n)]^\beta s[f(n), g(n)]^\gamma, \quad (25)$$

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

$$\text{luminance: } l = \frac{2\mu_{f(n)}\mu_{g(n)} + C_1}{\mu_{f(n)}^2 + \mu_{g(n)}^2 + C_1}, \quad (26)$$

$$\text{contrast: } c = \frac{2\sigma_{f(n)}\sigma_{g(n)} + C_2}{\sigma_{f(n)}^2 + \sigma_{g(n)}^2 + C_2}, \quad (27)$$

$$\text{structure: } s = \frac{\sigma_{f(n)g(n)} + C_3}{\sigma_{f(n)}\sigma_{g(n)} + C_3}, \quad (28)$$

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

$$\text{SSIM}[f(x, y), g(x, y)] = \sum_n \frac{(2\mu_{f(n)}\mu_{g(n)} + C_1)(2\sigma_{f(n)g(n)} + C_2)}{(\mu_{f(n)}^2 + \mu_{g(n)}^2 + C_1)(\sigma_{f(n)}^2 + \sigma_{g(n)}^2 + C_2)}. \quad (29)$$

The SSIM index has a max value of 1 when the images are identical which makes it easy to read. [17]

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. [18]

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 thesis the distance required for the edge response to rise from 10% to 90% was chosen, in figure 12 an illustration of the definition is shown. 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.[19]

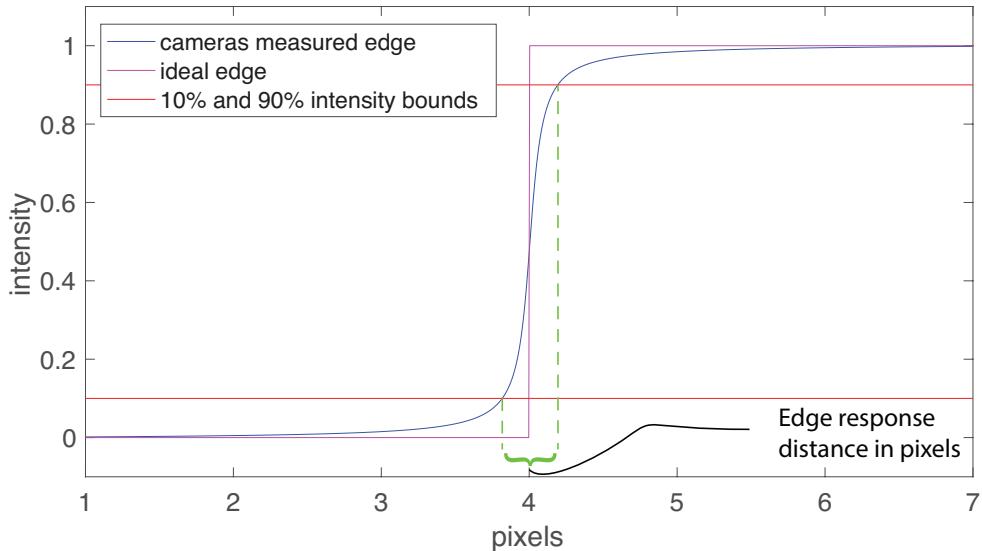


Figure 12: Definition of 10-90% Edge response.

3.7 Method criticism

- No Reference Image Quality Assessment is not designed for SWIR images or SPC:s characteristics noise therefore the results may not reflect how the QA 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 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 produced was simulated by using the reconstructing algorithm and measurement matrix described in section 3.4.1 and 3.3.1 on high quality images captured with a state of the art SWIR camera. The images captured by the SWIR camera acts as a ideal reference to the reconstructed images. By simulating the result from "ideal" images the reconstruction process gets 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 \mathbf{x} was calculated to obtain an simulated signal vector \mathbf{y}

$$\mathbf{y} = \Phi\mathbf{x} + \epsilon, \quad (30)$$

this operation was calculated for different subsampling ratios between 5-30% and different noise levels. The noise, white Gaussian noise, was added to the normalized measurement signal \mathbf{y} . The added noise represent a simple model of the noise expected in the SPC and was scaled with the standard deviation σ between 0 – 0.2.

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

To create the graphs in figure 14 and 15 this procedure was applied to all 21 images for sub sampling ratio 5% to 30% and added noise with standard deviation between 0 – 0.2. The standard deviation is not increased above 0.2 because the reconstruction fails at that point. In figure 13 a sample of reconstructed image from one of the SWIR images is presented with different amount of noise and sub sampling ratios.



(a) Reference image

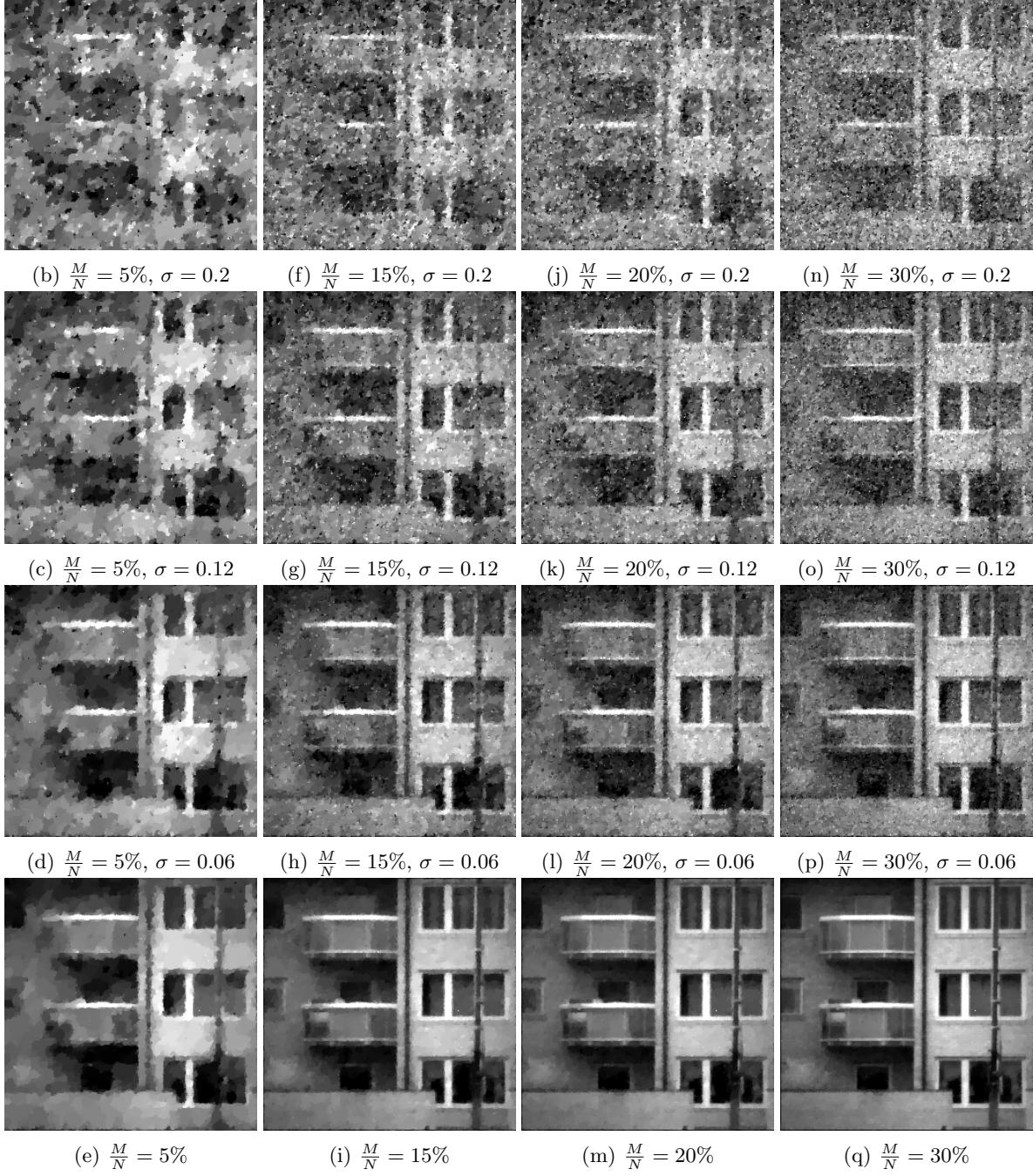


Figure 13: Example of reconstructed images with added noise at different sub sampling ratios.

As seen in figure 13 the reconstructed image quality increases with more measurements and lower noise levels. This observation is confirmed in the graphs i figure 14 and 15 where PSNR and SSIM respectively has been calculated for all 21 reconstructed images and interpolated.

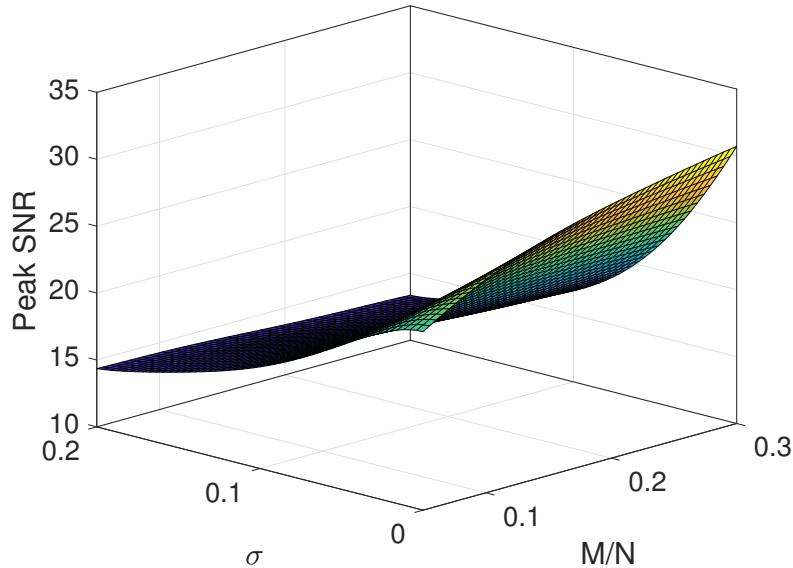


Figure 14: Peak SNR result depending on number of measurements and simulated noise level.

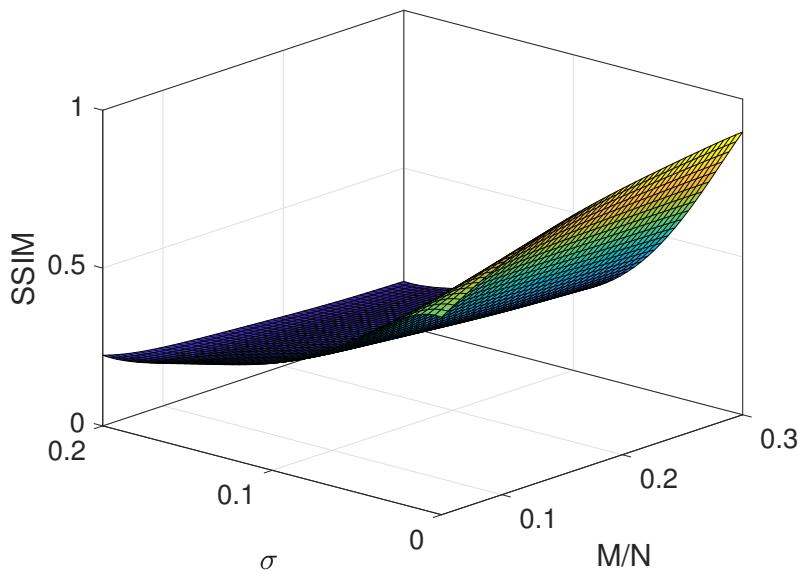


Figure 15: SSIM result depending on number of measurements and simulated noise level.

In both figures 14 and 15, it can be seen that when the noise increases the reconstructed image quality is not improved at the same rate as 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 the graph in figure 16 shows less noise and more samples yields 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.

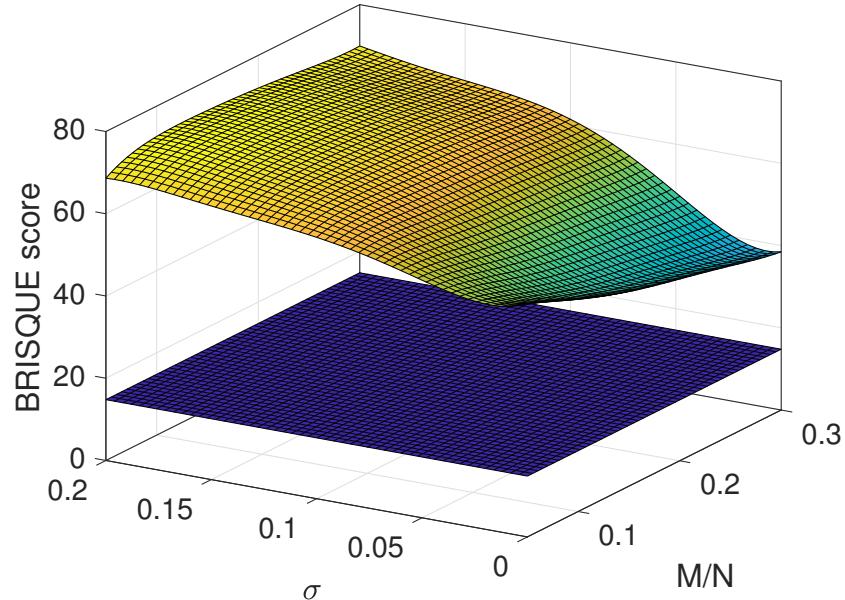


Figure 16: BRISQUE result depending on number of measurements and simulated noise level. Lower surface is reference image score. (Lower score are better)

In figure 17 the result has been flatten to a 2D graph with fewer selected data points for clarity. In the noiseless case the score will not be better then approximately 40 for the reconstructed images while the SWIR images has a mean value of 15.

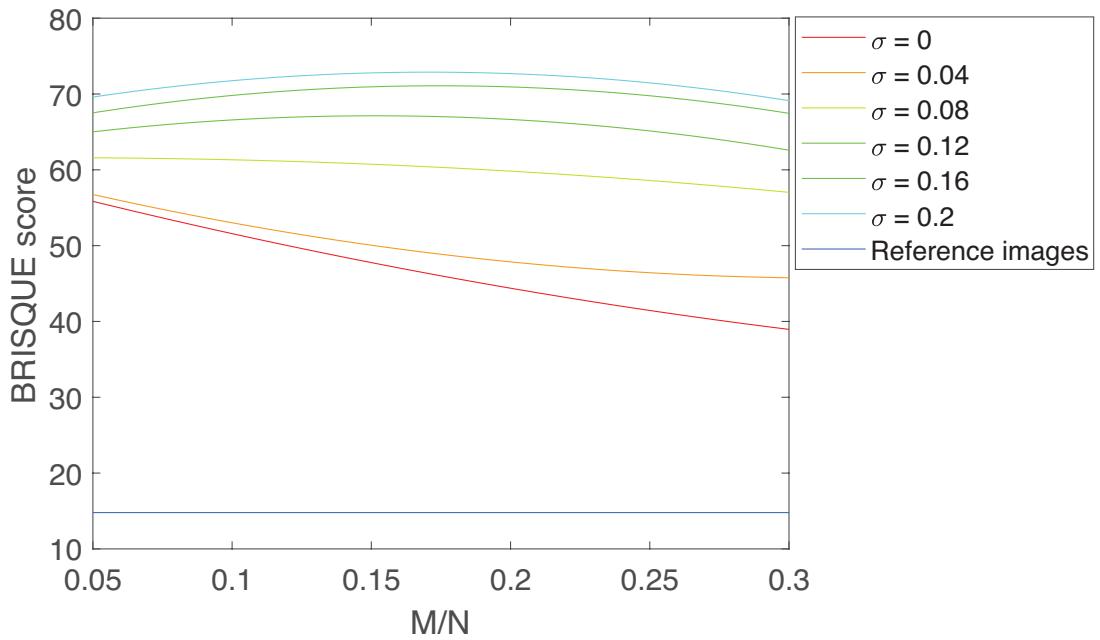


Figure 17: BRISQUE result depending on number of measurements for different simulated noise levels.

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 get better after, as seen in figure 17.

4.1.3 Dynamics in scene

In the current experimental SPC setup the exposure time is between 10–50 seconds which increases the chance of dynamics in the scene. Dynamics in the scene will reduce the reconstruction performance because the scene is assumed to be constant. By simulating dynamics scenes in a controlled environment their individual effects to the sampled signal \mathbf{y} can be identified and evaluated. 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 is simulated. The goal is to see how the signal change when dynamics is introduced in the scene. In the case of luminance change the moving mean algorithm presented in section 3.5.5 is evaluated.

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

$$\mathbf{y}[m] = \phi_m \mathbf{x}_m. \quad (31)$$

In the first scenario an object is placed in an image but for each measurement the location of the object will be moved in a small bounded area of the image. This model represents a scene where the background is static but a person is standing in the same area but moves around.

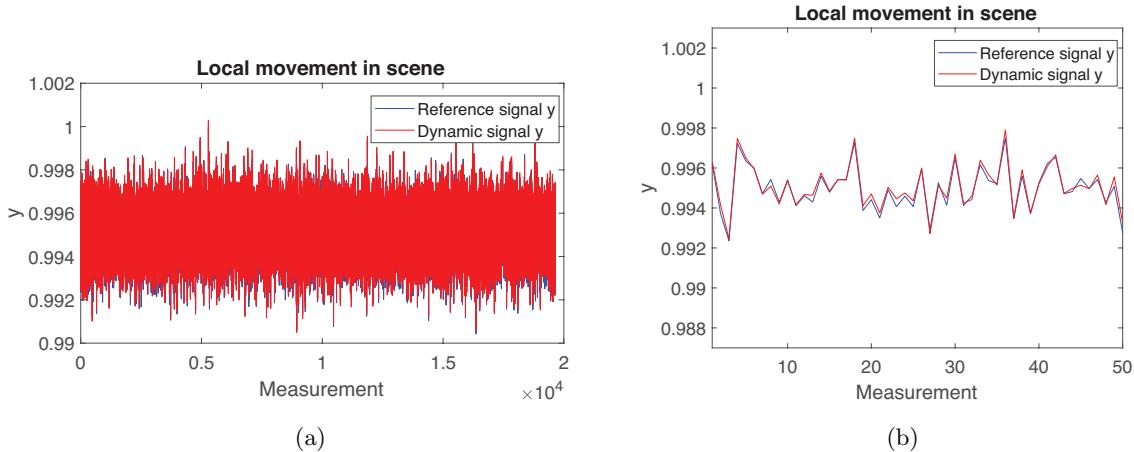
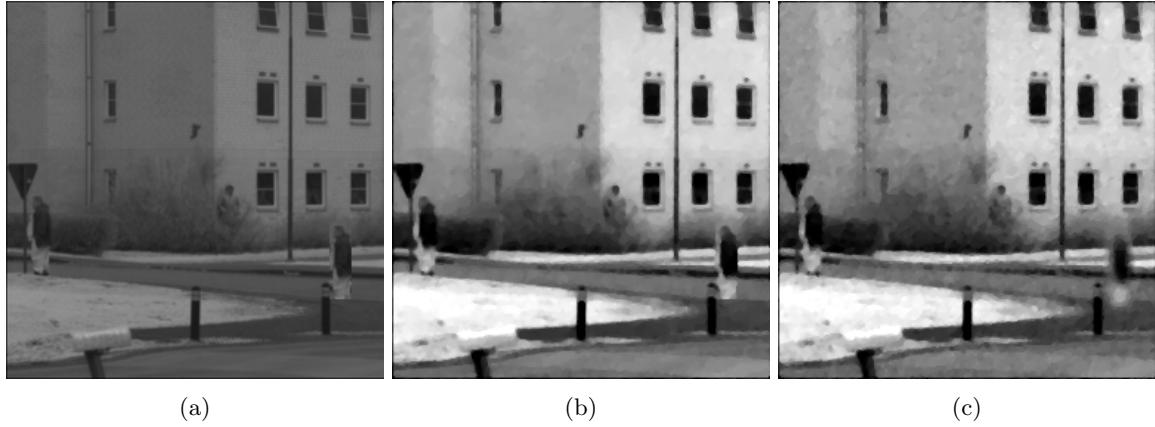


Figure 18: (a) Perturbed signal from local movement on top of reference signal. (b) Zoomed in view of some samples from figure (a).

As seen in figure 18a there is no obvious difference between the non perturbed reference signal and the distorted signal. In figure 18b where some of the samples is displayed no large difference can be seen either.

The reconstructed images from the reference signal and the perturbed signal is displayed in figure 19b and 19c respectively. The difference between the reconstructed images are visible to the naked eye, not only does the object moving around get blurry and noisy but the whole image globally.



(a)

(b)

(c)

Figure 19: The results of local movement on a reconstructed image, sub sampled 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 2 the results from calculating PSNR and SSIM of the the reconstructed images is presented. It can be observed that the dynamic test image has been effected to some degree by the movement.

Peak SNR	SSIM
29	91

Table 2: Evaluation comparing non perturbed reconstructed image against reconstructed image with local movement

The second scenario is an object passing through the whole scene to an other place in the scene far from the original place. The problem is modeled with a static background then as the simulated measurement is acquired the object will cross the 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 which corresponds to approximately 0.7 seconds when sampling with the SPC in its current setup.

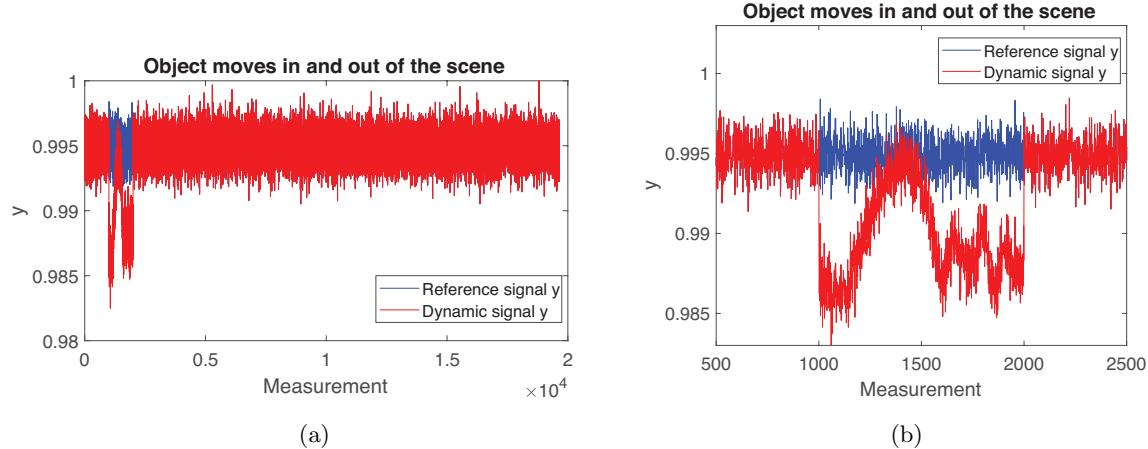


Figure 20: (a) Pertubated signal from large movement on top of reference signal. (b) Zoomed in view of some samples from figure (a).

As seen in figure 20 the exact moment the object enters the scene the signal changes. This is because a completely new structure has entered the scene and therefore changing the DC level. It can also be noted that the object passed something which has approximately the same intensity as the background and therefore the DC signal almost comes back up to its original state for a brief moment.

In figure 21 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 object moving, but the object is creating uncertainty in the whole image, resulting in global noise.

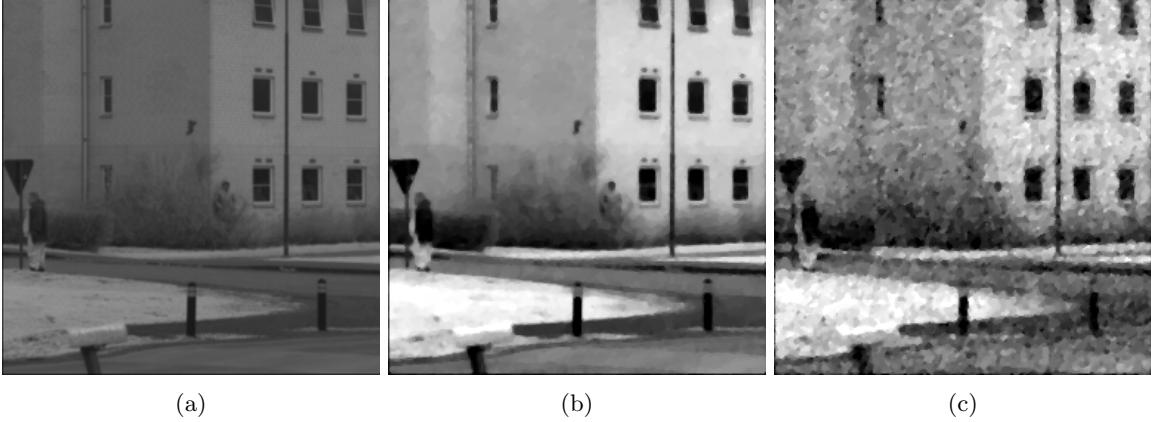


Figure 21: The results of large movement on a reconstructed image, sub sampled at 30%. (a) Original reference image. (b) Reference image reconstructed from the original image without movement. (c) Reconstructed image from a scene with object passing through.

In table 3 the results from calculating PSNR and SSIM of the reconstructed images is presented. It can be observed that the image has been effected heavily by the movement, lowering the SSIM index to 58.

Peak SNR	SSIM
23	58

Table 3: Evaluation comparing non perturbed reconstructed image against reconstructed image with local movement

The third scenario is luminance change in the scene which is 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.

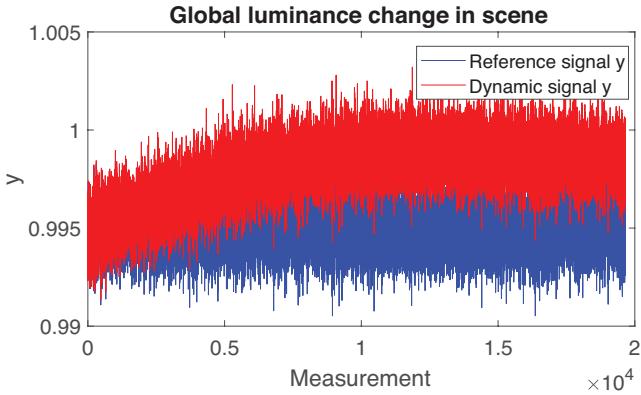


Figure 22: Signal effected by light intensity change on top of reference signal.

As seen in figure 22 the DC level of the signal will slowly change but the stricture of the signal stay the same. In figure 23 the reconstructed images from the perturbed signal and the reference signal is 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.

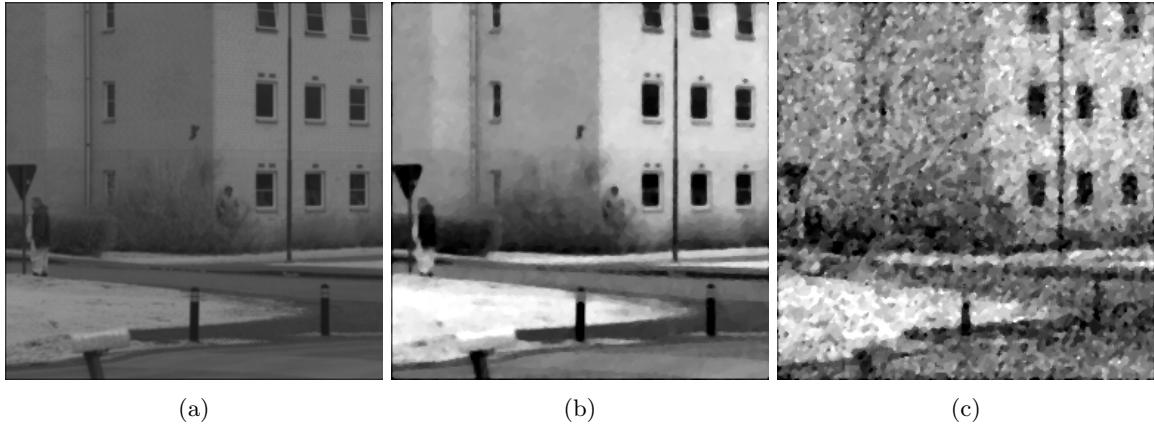


Figure 23: The result of global light intensity change on a reconstructed image sub sampled at 30%
(a) Original reference image. (b) Reference image reconstructed from the original image without light intensity change. (c) Reconstructed image from a scene with global light intensity change over the measurements.

In section 3.5.5 a model of this problem was proposed with a 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 24a the resulting signal is plotted over the dynamic signal, the processed signal is stationary again. In figure 24b and 24c 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.

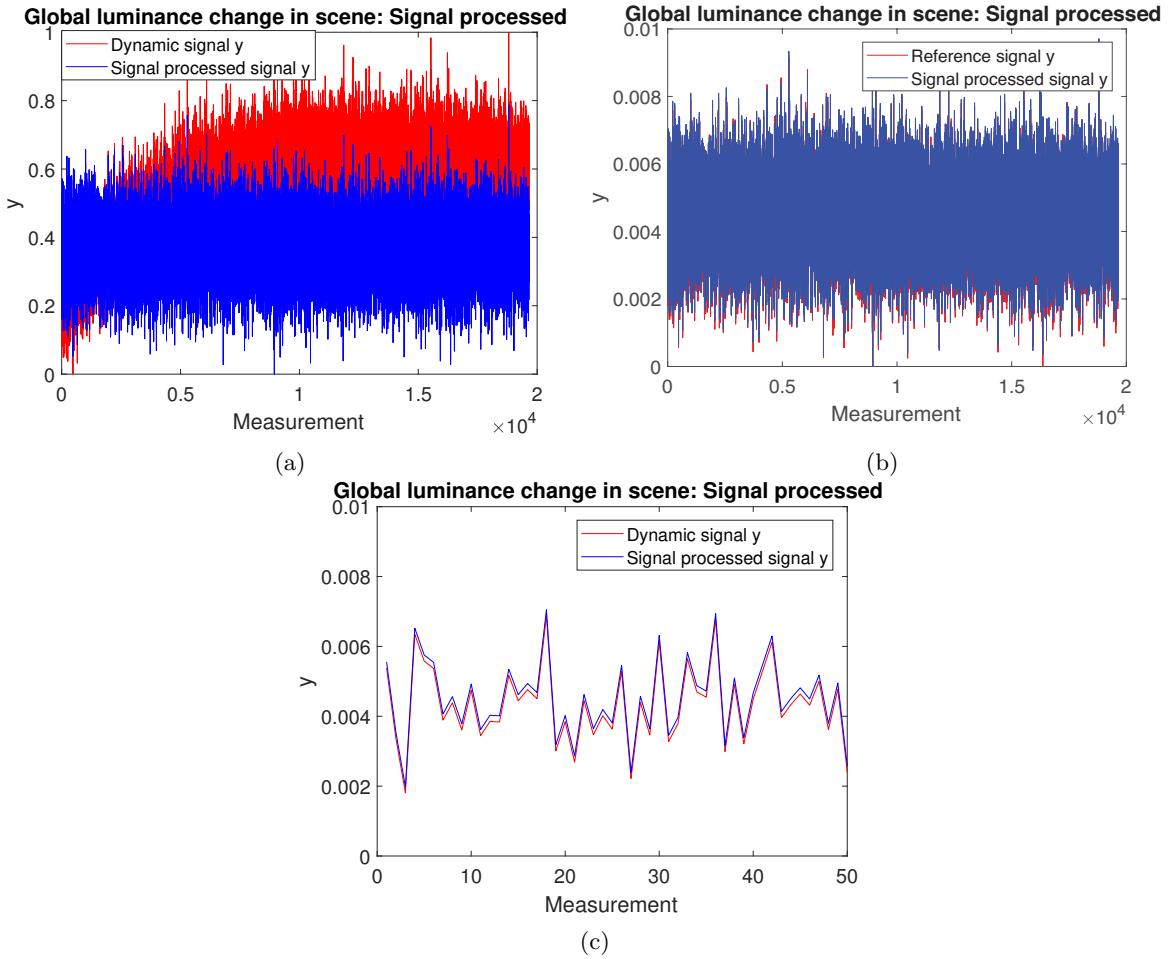


Figure 24: Post processed signal using moving mean subtraction. (a) Post processed signal on top of the dynamic signal. (b) Post processed signal on top of the reference signal. (c) Zoomed in view of (b).

In figure 25 the processed signals reconstructed image is displayed between the reference and perturbed signals reconstructed images. The signal processing 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.

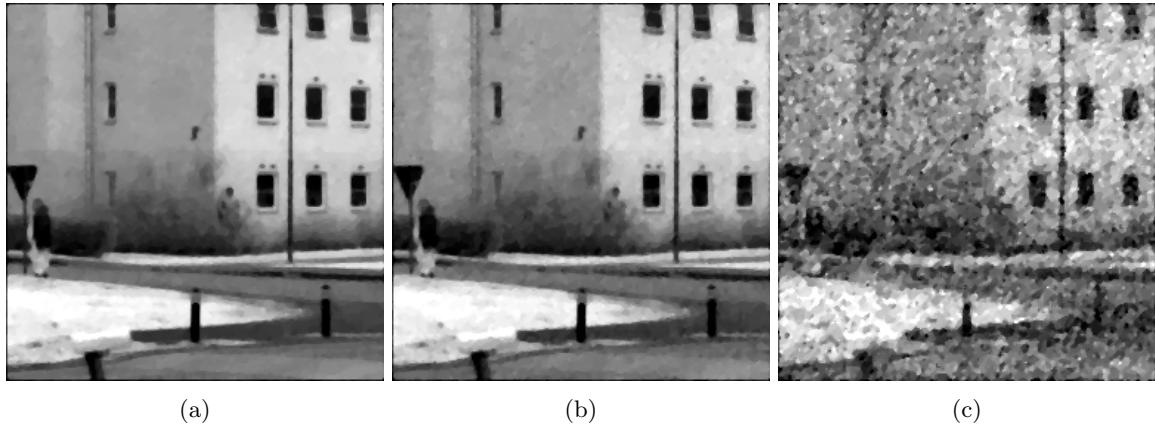


Figure 25: The result of processed signal perturbed by light intensity change on a reconstructed image sub sampled at 30% (a) Reference image reconstructed from the non perturbed signal without light intensity change. (b) Reconstructed image from a scene with global light intensity change and post processed by moving mean subtraction. (c) Reconstructed image from a scene with global light intensity change over the measurements.

In table 4 the results from calculating PSNR and SSIM of the reconstructed images is presented. Both PSNR and SSIM is increased for the reconstructed image using moving mean subtraction.

Signal	Peak SNR	SSIM
Perturbed signal	19	38
Mean subtracted signal	33	93

Table 4: Evaluation comparing non perturbed reconstructed image against global luminance change reconstructed image and 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 evaluation, edge response and subsampling ratio. All images captured by the SPC of natural scenes was 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 to the images produced by the SPC. One solution to obtain a reference image is to calculate a homography between images. But there are two problems of just performing a homography to obtain a reference image, the first one is that, 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 for example high contrast edges in the images will not match and produce large errors in the performance measurements. To solve these two problems the scene is a flat surface with a printed pattern and to avoid the error from sharp edges the pattern is constructed with smooth contrast changes. To not complicate this experiment the reference image is the computer generated optimal images which was being transformed to fit the image captured by the SPC.

The reference image which was created using sine functions to avoid edges can be seen in figure 26 and the reconstructed images with different subsampling ratio.

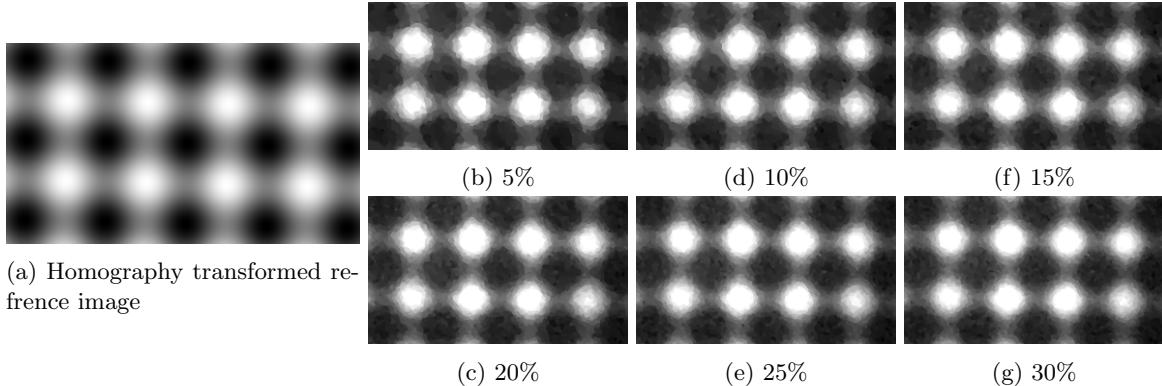


Figure 26: The reconstructed images with subsampling ratios between 5-30% and the reference image transformed to fit the SPC images using homography.

Before the results from the evaluation is presented it is worth noting that the reference image is a perfect simulated reference image, which was not effected by the uneven light source and 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 26b to 26g. In figure 27, PSNR and SSIM has been calculated for each subsampling ratio.

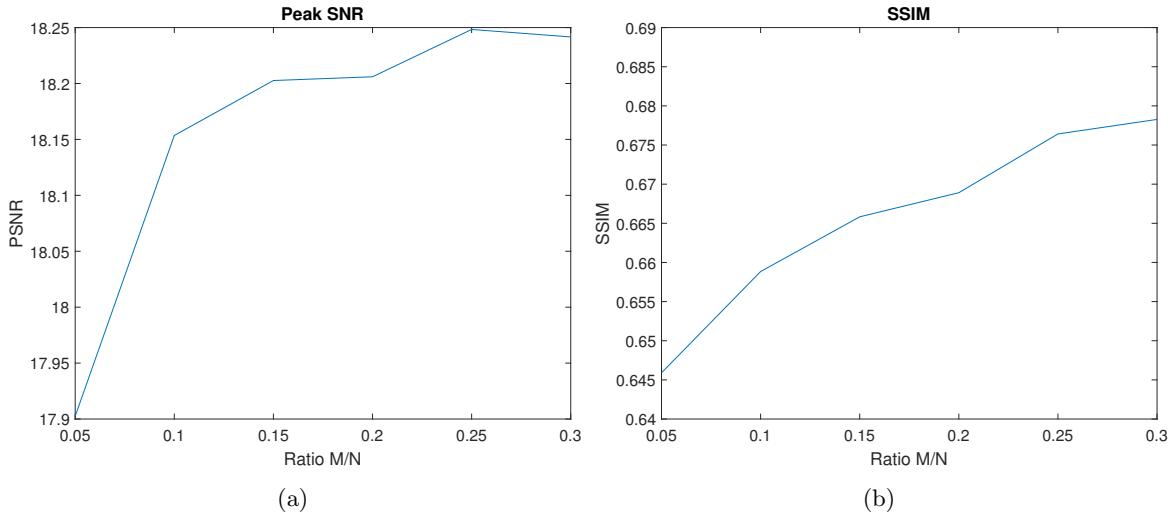


Figure 27: Signal quality of SPC images compared to reference image. (a) Peak SNR for reconstructed images against reference image. (b) SSIM score for reconstructed images against reference image.

The result shows increased quality with increased subsample ratio. In the case of PSNR the image quality rapidly improves when increasing subsample ratio up to 15%.

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. In addition to presenting the BRISQUE score for the images, each image has been classified into three distinct groups in order to link the BRISQUE scores to each image and their sampled signal.

In figure 28 each image is evaluated using BRISQUE with subsampling rate from 5 – 30%.

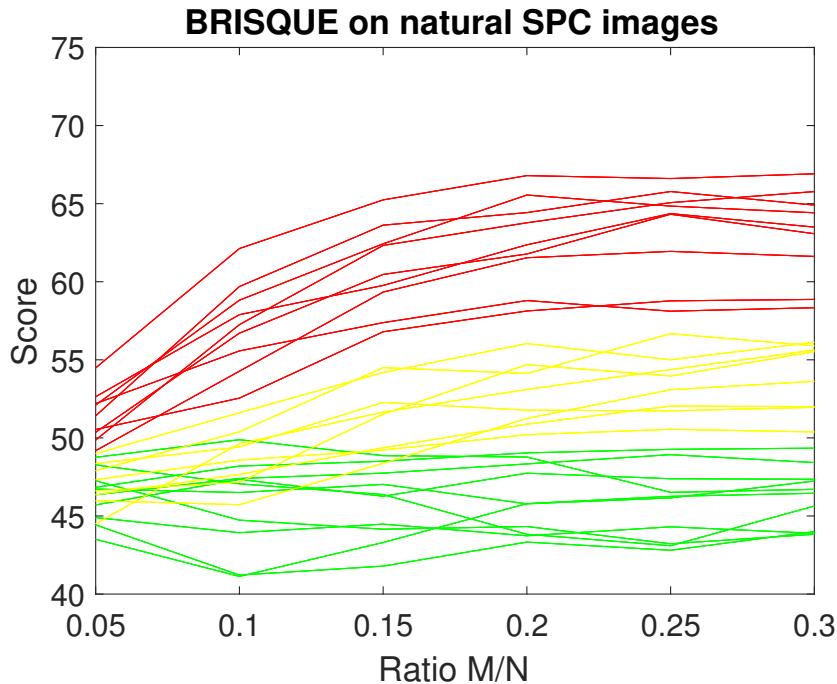


Figure 28: BRISQUE score for images reconstructed from the SPC with subsampling ratios from 5% to 30%. Each line represent one image and is classified with different colors depending on the initial score at subsampling ratio 5% and the general trend for higher subsampling ratios. Lower score is better.

As seen in figure 28 each image has been plotted separately due the high variance in the image set and the distinct different trends.

The images has 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 has been color coded where:

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

In figure 29-31 a subset of four images from each class is presented. All images is reconstructed using 30% subsampling ratio.

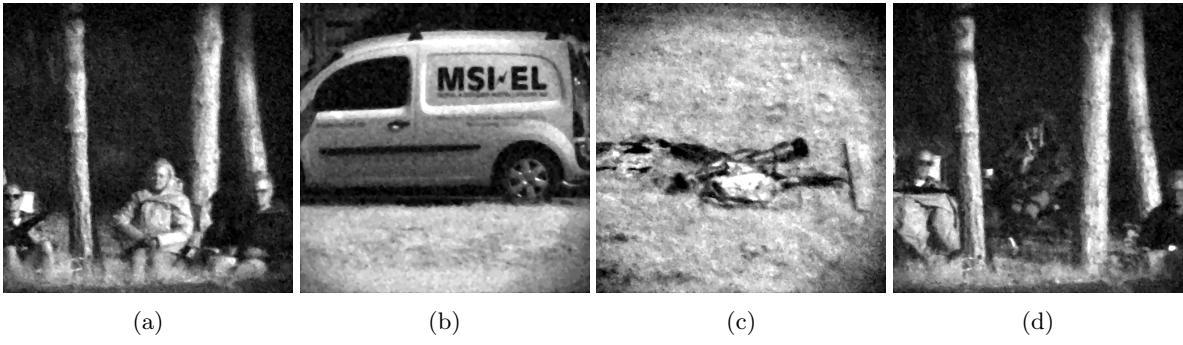


Figure 29: Sample images from the green image set. (a) and (d) People sitting in the edge of a forest. (b) Stationary car. (c) Camouflage jackets and a AT-4 anti-tank weapon on the ground.

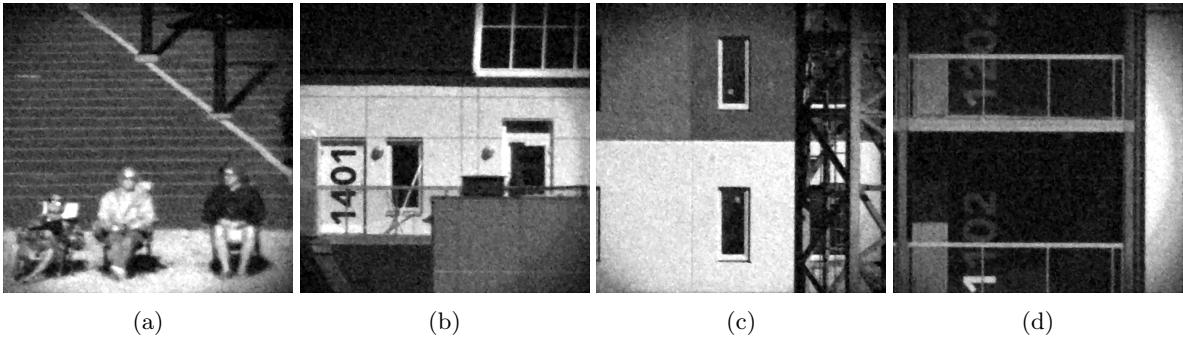


Figure 30: Sample images from the yellow image set. (a) People sitting next to a parking lot. (b) - (d) House facades.

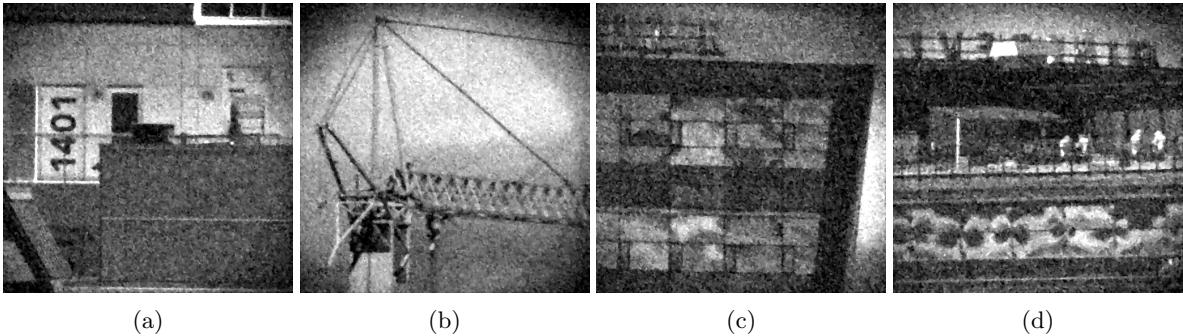


Figure 31: Sample images from the red image set. (a) House facade. (b) Construction crane (Moving clouds in background). (c) Mjärdevi Center facade (Moving clouds reflected in the windows). (d) Mjärdevi Center balcony with people having a break (Moving clouds reflected in the windows).

As seen in the images in figure 29 - 31, the motive in each class is quite different, the images from the green image set contains forest, people and things, images from the yellow image set contains buildings and structures with horizontal 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 scene.

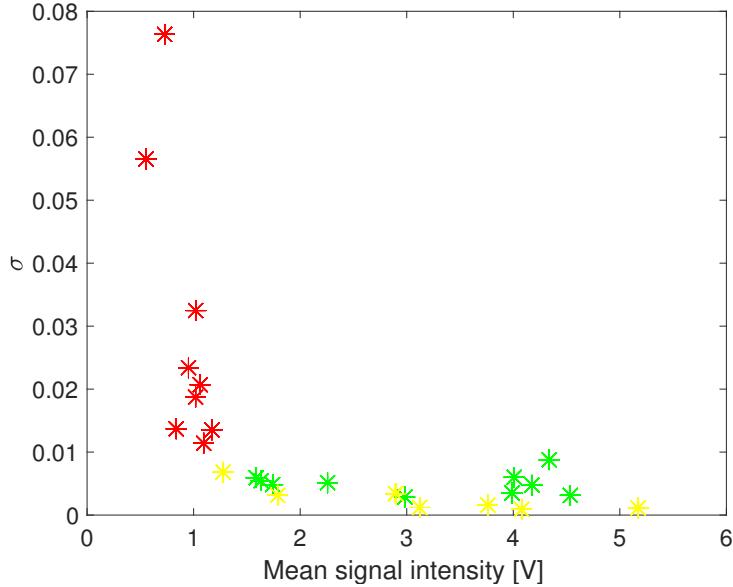


Figure 32: Mean sampled signal intensity compared to the normalized background noise variance, each signal has the same corresponding color classification as given before, shown in figure 28.

From the plot in figure 32, it can be seen that over approximately 1.2 volt mean signal intensity there are no images from the red image set while images from both the yellow and green image set is mixed over 1.2 volt. Under 1.2 volt mean signal intensity the noise variance to desired signal 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 case 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 and the luminance change in the signal was more complex then the simulated test case, the sensor was highly sensitive to luminance change. This observation should not be unanticipated, 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 33 the sampled signal effected by luminance change and the signal proceeded by the moving mean algorithm in blue is shown.

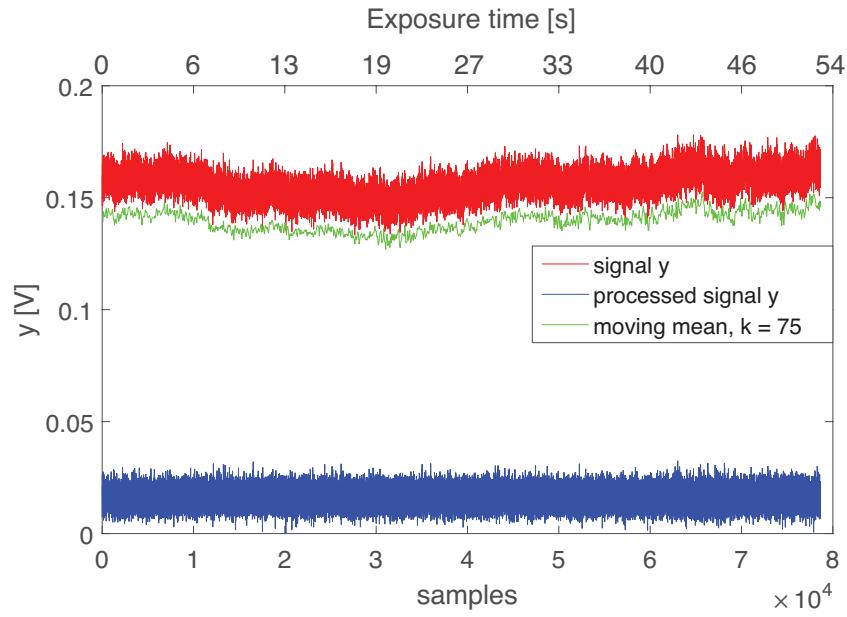
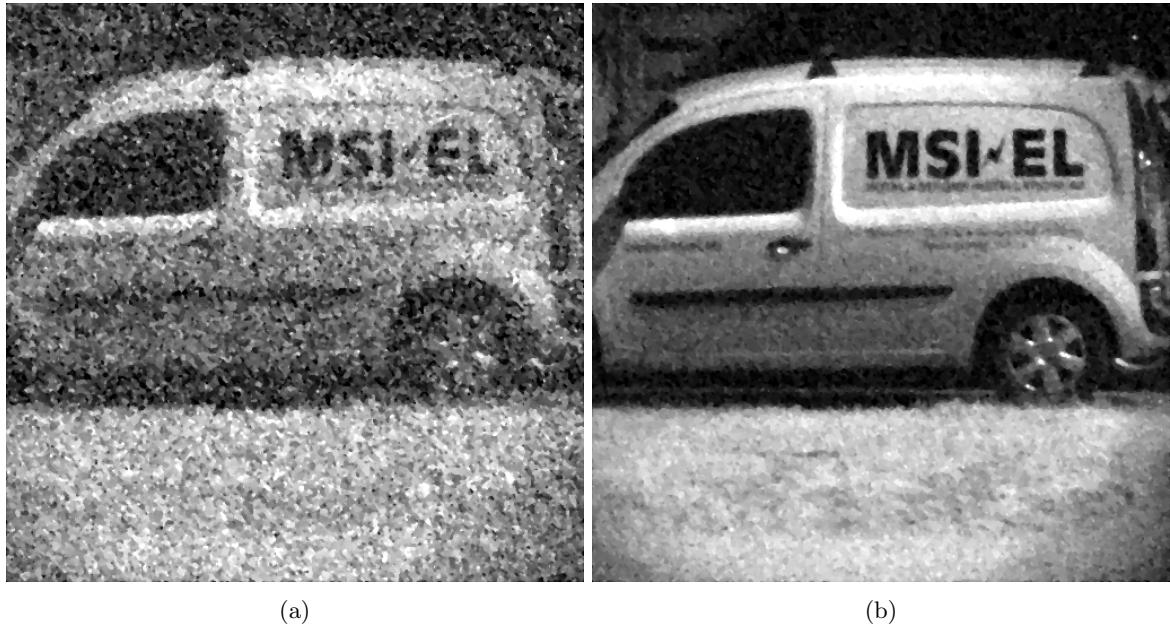


Figure 33: Sampled signal from SPC with light intensity change distortion and the signal processed by the moving mean algorithm.

As can be seen In figure 33 it can bee seen that the moving mean plotted in green which corresponds to light intensity change, change with a high frequency. "k" which is the number of neighboring samples to calculate the average was set to 75 to match the fast change which corresponds to a window of 50 milliseconds.

In figure 34 a reconstructed image without the moving mean algorithm used and one with the method used is displayed.



(a)

(b)

Figure 34: Reconstructed images, (a) before and (b) after applying moving mean mean algorithm on the sampled signal.

As can be seen in the figure 34, the method increments image quality significantly and the image reconstructed without the signal processing 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 signals sampled.

4.2.4 Edge response

The edge response is used to comparing the sharpness of cameras and lenses and in this section the SPC is compared to a state of the art SWIR camera. Two scenes was captured by the SPC and the conventional SWIR camera containing printed sheath of paper with simple tilted shapes on them, see figure 35. The scene was lit by a 135 Watt halogen lamp placed two meters from the sheath.

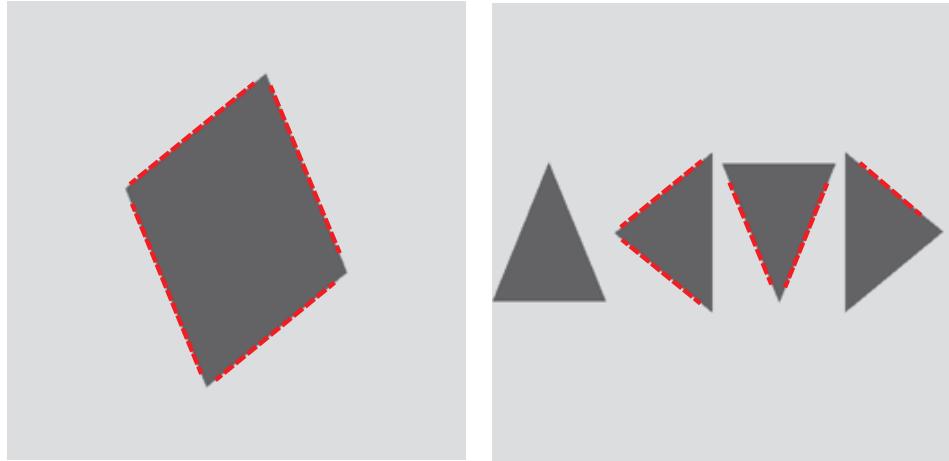


Figure 35: Printed targets with markings where the edge response measurements was performed

In the resulting images, edge response measurements were gathered from the specified edges in figure 35, 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 effected the edge response result. In figure 36 the images from the SWIR camera and SPC are presented.

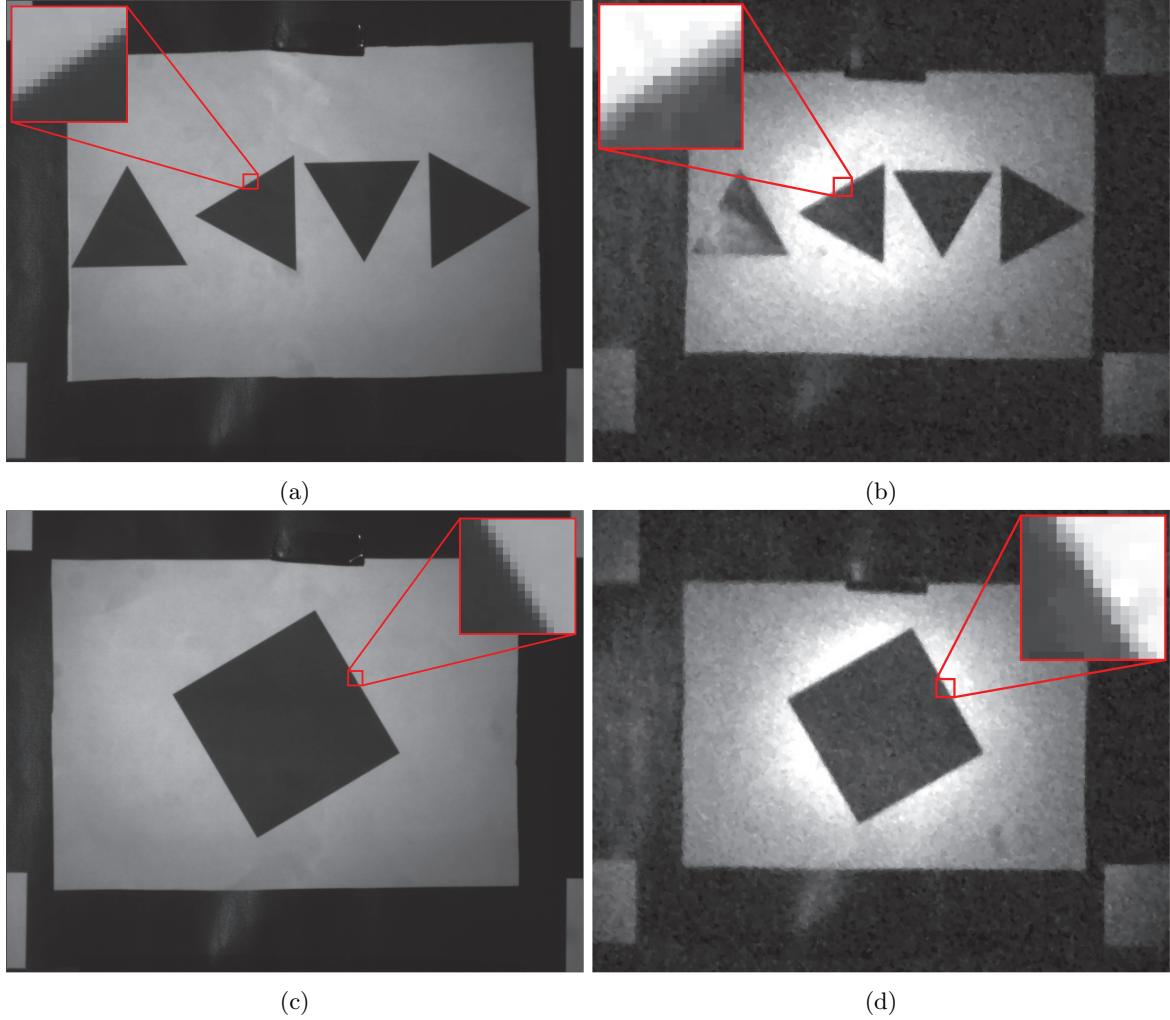


Figure 36: SPC and state of the art SWIR camera images. (a) and (c) from the conventional SWIR camera and (b) and (c) captured with the SPC. The images contain a zoomed in view of an sample edge.

The edge response is measured in the distance (pixels) required for a edge to rise from 10% to 90% intensity change. In figure 37 the result from the experiment is presented.

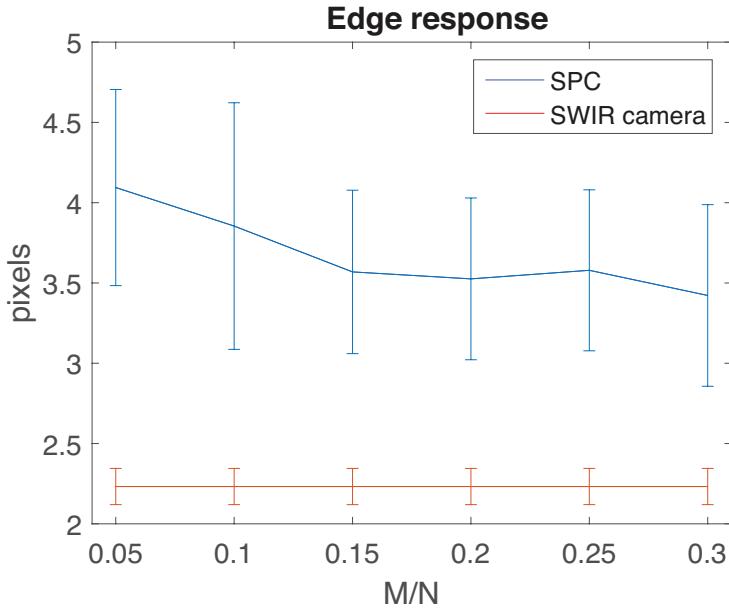


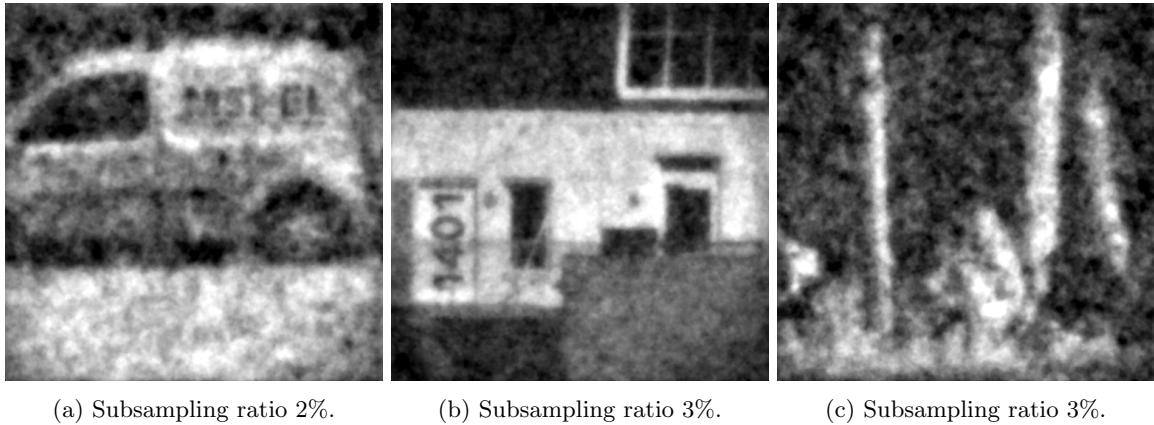
Figure 37: Edge response, distance (pixels) to rise from 10-90% in average for an edge.

From the plot in figure 37, 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 distance against 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 pixel.

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 a subsampling ratio needed to obtain a desired quality of the reconstruction. In addition, using a 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 effected by an increase of subsampling ratio.

The minimum number of measurements to reconstruct a image where the motive can be recognized is also effected 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 survey of the minimum subsampling ratio was performed on all images captured by the SPC in this thesis, and the result ranged from 2% to 4% to obtain a recognizable image, in figure 38 a sample of three images with varying minimum subsampling ratios is displayed.



(a) Subsampling ratio 2%.

(b) Subsampling ratio 3%.

(c) Subsampling ratio 3%.

Figure 38: Varying minimum subsampling ratios to reconstruct sample images captured by the SPC.

In the sample images in figure 38 the minimum subsampling ratio varied between 2% to 3%, but as can be seen, this is only the minimum subsampling ratio where the motive is merely recognizable, large structure can be identified but detail is lost. In general a subsampling ratio of 5% has always succeeded to reconstruct a identifiable image with some fine details.

In figure 39 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 an perception of which subsampling ratio is good enough for a given purpose, and as can be seen, the increase of subsampling ratio is not linear proportional to the increase in perceived image quality, like regular image compression.



(a) Visual camera image.



(b) Subsampling ratio 5%



(c) Subsampling ratio 10%



(d) Subsampling ratio 15%



(e) Subsampling ratio 20%



(f) Subsampling ratio 25%



(g) Subsampling ratio 30%

Figure 39: Reconstructed images captured by the SPC with increasing subsampling ratios. First row showing a reference visual spectrum image followed by SPC images reconstructed with 5% from top then increased subsampling ratio by 5% for each row to 30% on the last row.

As seen in the images in figure 39,

- 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. The images has some artifacts, the images looks 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 text on the car and door are getting sharper and the structure of 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 from the increments before, but they can be found in the details.

5 Discussion

In this section the results and method is analyzed and discussed. When discussing result, 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 capturing images with the SPC.

5.1 Result

Overall the results obtained in this thesis reflects what has been stated in the theory and research, but no reference of using the SPC to capture natural scene could be found and thus this thesis will present the link between real world application 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 part as expected given the CS theory, with increased subsampling ratio the performance increased. But the interesting part of this results is whats 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 14 and 15 shows. This result tells us that if the signal contains a high degree of noise, a larger subsampling ratio may not improve the reconstructed image as mush as expected.

When obtaining the same measurements with the SPC one low frequency image was captured and reconstructed, the images was captured in near optimal environment such that a homography between the reconstructed image and reference image could be made with good precision. In figure 27a and 27b we can see the PSNR and SSIM of the image given the reference image, as theory predict and confirmed in the simulated case, the performance increases when the subsampling ratio increases. But 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 26b-26g, 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 simulated reconstruction with results in figure 16, it can be see that the graph looks like a inverted version of the PSNR and SSIM graphs of the simulated images. This results alone is positive for this thesis because it was unknown if BRISQUE as a evaluation method would work well for SWIR images and reconstructed images from the SPC. We can also see that the reference images score about 20 BRISQUE points better then the best reconstructed images and thus can say that the sampling and reconstruction even in the best case scenario without noise, will not yield as good results as an conventional camera. This evaluation was perform with the specific motivation to evaluate the measurement matrix and reconstruction method, and these results shows that they will not be able to reconstruct images with the same quality as an conventional SWIR camera. But these results also give a hint of which BRISQUE score we can expect from the SPC in the optimal case, where 40 can be seen as an optimum given that noise and post processing will effect the sampled signal.

When studying the plot with BRISQUE score given by the results from images reconstructed from the SPC in figure 28, we can see that the best images score just over 40 BRISQUE points, which is the same score as score as simulated images with small or no noise added, which 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 17 for different noise levels, thus we can conclude that simulations gives a good indication of where the real images will score given a noise level.

In figure 29 to 31, we see the images divided into 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 31 it stands out the most. My analysis of why the BRISQUE score and image quality differ is,

- first if we take a look at the green and yellow image sets in figure 29 and 30, where the image quality and lighting look quiet the same but yet differ in the BRISQUE score, it might be a property of the BRISQUE classifier. The BRISQUE classifier is built to assess image quality in natural images, and if we take a look at the main difference between these two data sets we can see that one is pictures of a car, humans, forest and clothing and the other mainly of buildings and large structure in the images with large low frequency sections i.e. not so natural, which can effect the score.
- The major difference between the images in the green and yellow image set and 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 30a and 31a are practically the same motive, but the latter is darker. The darker scene was shot in morning and did not luminate the facade directly and 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 is large movement in the scene, most of the images in the "bad" image 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 and therefore decreased the BRISQUE score significantly.

In the last part of section 4.2.2 the results from plotting normalized noise standard deviation against mean signal intensity in figure 32, was presented. The plot give more information on why the BRISQUE score was so deviated. 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 variance 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 17. 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 the plot, we can see that there are a subset of images from the red image set with almost the same normalized noise variance and mean signal intensity as from the yellow and green image sets but yields a worse BRISQUE score anyway, this strengthening the statement the there probably is at least one more factor that reduces reconstruction performance. And as stated in the last paragraph 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 are 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 BRISQUE score, which implies that the motive in the images effecting 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 19, we can see that there is no significant difference between the non perturbed reference signal and the distorted signal. It can even be interpreted as added noise to the signal and it is barely detectable even if the signal is known.

The effect on the reconstructed image, seen in figure 19a, looks like global noise is added. The test implies that local movement in a scene will 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 and modeled an unseen object pass through, the samples with movement was very easy to spot, which figure 18b shows. The effect in the reconstructed image is as expected even worse than local movement, with a global distortion, seen in figure 19b. In this simple isolated case the image could be saved by removing the measurements when the object was moving, reconstructing an image with fewer measurements. The resulting image would not be as good as the image in figure 21b but it would not contain the noise present in figure 21c.

In the case of luminance change, the effect on the reconstructed image is even worse than scenes containing movement, which figure 19c and comparing table 2, 3 and 4. Because we know that this problem is real and can not be avoided in natural scenes, a model to suppress this issue was tested with good result, but as can be seen in figure 24b, 23 and table 4, 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 image 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 33. But as seen in the results in figure 34, the moving mean method worked despite the more complex changes to the signal. Considering that this problem was consistent for all natural images this method became essential for this thesis 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 show, the method is essential and was used for all images captured by the SPC and presented and evaluated in this thesis.

This method was mainly constructed because I knew 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 effect 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 environment 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 the dynamics presented that is persistent even in a "static" scene, I decided to not photograph scenes where large movement occurred as a car, object or human, even though it could be detected, because it could also be detected as luminance change.

As stated, even "static" scenes will with high probability contain both movement and luminance change which will effect the reconstructed images. Therefor I can conclude that all reconstructed

images in this thesis has to some degree added global noise from local movement and the signal processing to counter luminance change.

5.1.4 Edge response

When comparing the edge response between the conventional camera against the SPC the results was very clear, the conventional camera outperformed the SPC with one to two pixels in distance. I think there are multiple factors why the results from the SPC differed so much from the conventional camera, and have listed them 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 36, especially in the contrast of the white background where the light intensity drops. The "patch" artifact from the reconstruction algorithm can effect the sharpness of the image. We can also see from previous test that even from synthetic data the BRISQUE score is significantly worse then the original image.
- The pixel grid setup in the DMD has two problems that could effect the sharpness, the DMD mirrors is aligned by the diamond shape and in the current setup to fix the ratio and index, two mirrors is 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 possibly not optimal.
- In this thesis no significant image improvement from post processing such as denoise or sharpening was performed unlike the conventional camera.
- And as stated before, with the long exposure, vibrations and light intensity change effected the results (the SPC could detect significant light intensity change from the halogen lamp powered by a DC-unit), which contribute to global noise in the reconstruction.

Even though the result 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 with a DMD with "regular" square mirrors and by more advance 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, for the whole image set the results varied between 2-4%. The variance could be effected by several factors such as image complexity, SNR and dynamics in the scene, which 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 a stable reconstruction at 5% subsampling ratio for all images and for higher subsampling ratio increased image quality. With greater subsampling ratio than 15% the images start to reach their highest potential where details start to show and the images gets sharp. Some global noise is present in all images, the noise is probably in large 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 was heavily influenced and implemented by widely accepted methods from articles and experiments. While the two other categories, the post process and the image capturing dependent 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 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 happen during the sampling. By the time this issue was discovered there where no time to implement a new method, but it was much needed.

The sampling matrix chosen in this thesis was the permuted sequency ordered Walsh Hadamard measurement matrix, this sampling matrix in unison with 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 a agile and memory efficient program.

The reconstruction algorithm TVAL3 was used throughout this masters thesis 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 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 process quite basic signal and image processing was performed and was intentional designed that way in order to present the result as true as possible. One algorithm was developed specific for the SPC which was the moving mean algorithm to suppress the impact of lumination change in the sampled signal. The algorithm showed great result 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 was achieved anyway. If I had more time or a target DMD pattern rate would have been set earlier a good signal denoising implementation could have improve the result.

5.2.1 Replicability, reliability & validity

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

In the case of compressive imaging the reliability and validity goes hand in hand, if whats being measured is not the correct data the reconstruction will fail and if the reconstruction succeed the

measurement must be correct. Therefore compressive sensing is very binary, ether you get it right or you fail and thus I think my result 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 thesis both presented evaluation from simulated images and images captured by the SPC to show how the chosen sampling method and reconstruction algorithm performed and how the whole architecture performed in unison.

The sampling strategy using the structural random matrices method with sequency 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 in this thesis, 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.

In post processing an algorithm to correct the signal effected 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. Result showed in both simulation of luminance change and in real scenes a significant improvement of image quality when using this method and the analysis 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 shows that high quality and high resolution images can be acquired in natural daylight scenes. In good conditions with enough 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. In the subsampling range 5-30% as was used through out 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 indexed based on statistics of "naturalness" in the images and it could be seen that the SPC could get the same results as the best in the simulations indicating that the sampling and reconstruction is the main source of image degradation and that the SPC hardware in the right conditions do not effects 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.

In this last paragraph a summarize of the whole thesis is made by answering the research questions:

- The image quality of the produced images can be evaluated with the same methods used when evaluating a regular camera. But in the case of the SPC the sampling matrix and reconstruction algorithm can be evaluated independently of the camera as well.
- The state of the art method to capture and reconstruct images using a 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 can be made. And I think the most crucial problems could be "bought out" by more sophisticated hardware.

As identified multiple times in this thesis, the largest contributor to decreased image quality is exposure time and noise. The exposure time can be decreased by a faster DMD, today there exist a 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.

Ether if 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 which 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 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 for to display at the DMD. This solution would also be more agile where subsampling ratio, which type of measurement matrix and exposure time could easily be adjusted prior to sampling.

The second hardware upgrade, 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 doid 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 is collected in each measurement and thus reduce the subsampling ratio needed to reconstruct a corresponding single sensor image.[25]

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.

References

- [1] I. Rish and G. Y. Grabarnik, *Sparse Modeling*. Boca Raton: CRC Press, 2015.
- [2] M. Elad, *Sparse and Redundant Representations*. New York, New York: Springer, 2010.
- [3] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006, ISSN: 0018-9448. DOI: 10.1109/TIT.2006.871582.
- [4] 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*, vol. 15, no. 5, 2013.
- [5] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, “An architecture for compressive imaging,” in *IEEE International Conference on Image Processing (ICIP), October 8-11, 2006*, 2006, pp. 1273–1276.
- [6] 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*, vol. 5, 2006.
- [7] 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*, vol. 83, 2008.
- [8] C. Brännlund and D. Gustafsson, “Single pixel swirl imaging using compressed sensing,” *Symposium on Image Analysis, 14-15 March 2017, Linköping, Sweden*, 2017.
- [9] B. C. L. McMackin M. A. Herman and M. Weldon, “A high-resolution swirl camera via compressed sensing,” *SPIE*, vol. 8353, 2012.
- [10] E. Fall, “Compressed sensing for 3d laser radar,” Linköpings University, Master’s thesis, 2014, ISRN: LiTH-ISY-EX-14/4767-SE.
- [11] T. Edeler, K. Olinger, S. Hussmann, and A. Mertins, “Multi image super resolution using compressed sensing,” *ICASSP*, pp. 2868–2871, 2011.
- [12] C. Li, “An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing.,” 2009.
- [13] “Fast compressive sampling with structurally random matrices,” in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2008, pp. 3369–3372. DOI: 10.1109/ICASSP.2008.4518373.
- [14] T.T.Do, L. Gan, N. Nguyen, and T. Tran, “Fast and efficient compressive sensing using structurally random matrices,” *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 139–154, Jan. 2012, ISSN: 1053-587X. DOI: 10.1109/TSP.2011.2170977.
- [15] L. Gan, T.T.Do, and T. Tran, “Fast compressive imaging using scrambled block hadamard ensemble,” in *2008 16th European Signal Processing Conference*, Aug. 2008, pp. 1–5.
- [16] 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*, Dec. 2013, pp. 1–4. DOI: 10.1109/ICICS.2013.6782833.
- [17] A. Bovik, *The Essential Guide to Image Processing*. Elsevier, 2009, ISBN: 978-0-12-374457-9.
- [18] M. Anish, M. Anush Krishna, and B. Alan Conrad, “No-reference image quality assessment in the spatial domain.,” *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 21, pp. 4695–4708, Dec. 2012, ISSN: 1057-7149.
- [19] F. Näsström, D. Bergström, F. Bissmarck, P. Grahn, D. Gustafsson, and J. Karlholm, “Prestandamått för sensorsystem,” Nov. 2015, ISSN: 1650-1942.

- [20] L. Duval, *50mm fl swir fixed focal length lens*, Edmund Optics Inc, 2017. [Online]. Available: <https://www.edmundoptics.com/imaging-lenses/fixed-focal-length-lenses/50mm-fl-swir-fixed-focal-length-lens/> (visited on 05/17/2017).
- [21] *Large area ingaas amplified photodetectors pda20c operation manual*, 1.1, Thorlabs GmbH, Sep. 2016. [Online]. Available: https://www.thorlabs.com/drawings/e22761fd83df1eab-B1D1D2AA-C921-6D8E-D2B14E11544D8225/PDA20C_M-Manual.pdf (visited on 05/16/2017).
- [22] T. Do, *Fast compressive sampling using fast compressive sampling using structurally random matrices*, Apr. 2008.
- [23] R. Haimi-Cohen and Y. Mark Lai, “Compressive measurements generated by structurally random matrices: Asymptotic normality and quantization.,” *Signal Processing*, vol. 120, pp. 71–87, 2016, ISSN: 0165-1684.
- [24] S. L. Shishkin, “Fast and robust compressive sensing method using mixed hadamard sensing matrix,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 353–361, Sep. 2012, ISSN: 2156-3357. DOI: [10.1109/JETCAS.2012.2214616](https://doi.org/10.1109/JETCAS.2012.2214616).
- [25] F. S. et al, “Computational imaging with a balanced detector,” *Nature: Scientific Reports*, vol. 6, 2016.