

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

Examensarbete

TQET33

Andreas Brorsson, Andbr981, (911023-1215)

Linköping, 2017

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 solution to undetermined linear system where the number of measurements M is less than the number of data points N in signal. Two constraints need to be fulfilled to apply compressed sensing sampling: the sampled signal needs to be spares in some basis e.g. Fourier or gradient, the second condition is that the measurement matrix must be incoherent with the sparse transform. The characteristic undetermined 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 signal \mathbf{x} . In figure 1 such linear equation system is shown.

$$\begin{matrix} \mathbf{y} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_M \end{matrix} \end{matrix} = \boxed{\Phi} \quad \begin{matrix} \mathbf{x} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ \vdots \\ x_N \end{matrix} \end{matrix}$$

The matrix Φ is an $M \times N$ matrix where $M < N$.

Figure 1: CS undetermined linear system

Scientists at Rice university in Texas, USA realized that the new method could be used to create a new camera architecture with a single photo diode in the sensor, the single pixel camera was born and thus a new sub field of compressed sensing was created called compressive imaging.

To be able to apply CS to imaging in the first place the constraints in CS needs to hold for images as well. The first requirement is that the signal needs to be compressible or sparse in some basis which natural images is known to be because they can be compressed using for example JPEG (Discrete cosine transform), JPEG2000 (Wavelet). The second constraint is that the measurement matrix must be incoherent with the sparse transform which for example i.i.d random distribution or some structure with the same property as i.i.d random distribution.

1.3 System architecture

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 the via the secondary mirror onto the DMD. To this point, the SPC works like a conventional camera with a DMD where the image sensor would be placed in the convectional camera. The SPC has an DMD in the focal point which resemble an image sensor but instead of photo diodes for each pixel there is a tiny mirror which individually can either reflect light 12 degrees to the right or left as seen in figure 2. The incoming focused light can ether be dumped or it can be reflected into the single pixel SWIR detector through an lens.

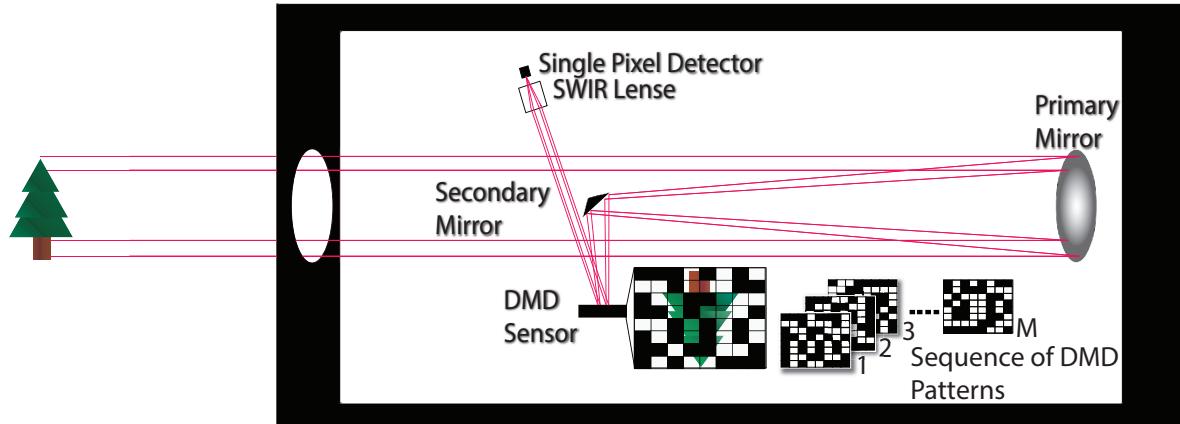


Figure 2: System overview

To connect the architecture with the math from CS it can be interpreted as, the light from the scene which is focused on the DMD is the desired signal \mathbf{x} , the image. The DMD can individually set each mirror the ether direct the light from each 'pixel' to the single pixel sensor or dump the light i.e a spatial light modulator (SLM). The DMD sets a pattern of pixel of intress which is a measurement matrix Φ_m to be summarized in the single pixel sensor y_m as a measurement. One measurement is the inner product of a measurement matrix and the signal, $\Phi_m \times x = y_m$. To complete a full measurement the process is repeated with different measurement matrices set on the DMD to the full undetermined linear system $\mathbf{y} = \Phi\mathbf{x}$.

1.4 Measurement matrix & reconstruction

How is the measurement matrix chosen? As told before the measurement matrix needs to be incoherent with the sparse transform and the DMD can only direct the light or not which mathematically is ether a zero or a one. The research tells that for example a i.i.d. Gaussian distribution with equal probability of a zero or one will with high probability be incoherent with a natural image scene. But how about the first constraint that the signal \mathbf{x} needs to be sparse or compressible in some basis? Often natural images is not sparse in the spatial domain unless the scene

is for example the night sky, well a good property of CS that the scene can be transformed to an other basis like this,

$$\mathbf{y} = \Phi \mathbf{x} \Leftrightarrow \mathbf{y} = \Phi \Psi \Theta, \quad (1)$$

where Ψ is a sparsifying basis for example to the DCT or Wavelet basis. And Θ is the coefficients vector which is more sparse then the spatial coefficient vector \mathbf{x} . And the transformation will not compromise the incoherence between the reconstruction matrix $A = \Psi \Phi$ and the coefficients Θ in the new basis. This means that the signal \mathbf{x} will be reconstructed with optimization in a more sparse basis Θ and then transformed back to the spatial domain.

What is special about CS is not just how the problem is presented but also how to solve it. It is known that an undetermined linear system has infinite many solution so how does the signal get recovered? CS exploit the characteristics of the signal x which is known to be sparse in some basis. With for example ℓ_1 optimization,

$$\hat{\Theta} = \arg \min \|\Theta\|_{\ell_1} \text{ subject to } \Phi \Psi \Theta = y, \quad (2)$$

which means that ℓ_1 optimization minimizes the non zero elements of Θ and can exactly reconstruct a K-spares vector or approximate a compressible vector. The exact recovery can be accomplished with high probability using $M \geq \mathcal{O}(K \log(N/K))$ measurements. This is why CS is powerful, it enables sub-Nyquist measurements with exact recovery in the noiseless case which can be approximated in real applications.

In the compressed imaging case where noise is present an other optimization algorithm has shown to be more successful at recovering images: total variation. Total variation regularization minimizes the magnitude of the gradient in the image and doing so it preserve edges and piece-wise constant structure in the image which is desired.

1.5 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.

1.6 Aim

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

1.7 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?

Add hypothesis that dynamics in scene can be suppressed?

1.8 Limitations

- The hardware rig provided by FOI
-

1.9 Thesis outline

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] Two books which thoroughly presents the topic sparse and redundant representation and sparse 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] David L. Donoho proposed the framework of compressed sensing and its application to images.

2.2 Compressive imaging

- [4] "Single-Pixel Imaging via Compressive Sampling"
- [5] "Compressed sensing for practical optical imaging systems: a tutorial"
- [6] "A New compressive imaging camera architecture using Optical-Domain Compression"
- [7] "An architechture for compressive imaging"
- [8] "A high resolution SWIR camera via compressed sensing"
- [5] "Compressive Sensing: From Theory to Applications, A survey"
- [9] "Compressed Sensing for 3D Laser Radar"
- [10] "Multi image super resolution using compressed sensing"

2.2.1 Measurement matrix & reconstruction

- [11] Chengbo Li:s master's thesis "An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing" describes his new total variation algorithm Li constructed which solve the CS problem. The algorithm is faster and produces better results for images than previous popular algorithms.
- [12]–[14] 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.
- [15] "An Improved Hadamard Measurement Matrix Based on Walsh Code For Compressive Sensing" Shows that sequency-ordered Walsh Hadamard matrix gives better reconstruction then the Hadamard matrix with the same benefits of using the Hadamard matrix. The resulting reconstructed image has near optimal reconstruction performance.

2.3 Evaluation

- Al Boviks book the essential guide to image processing [16] 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.
- [17] A Feature-Enriched Completely Blind Image Quality Evaluator describes how the no reference image quality assessment tool IL-NIQE works and compares to other NR-IQA algorithms. In the article a comparison with other state of the art NR-IQA is conducted which concludes that IL-NIQE is the best over all NR-IQA tool. This kind of QA is useful when there is no reference image available, which is true when taking photos with the SPC.

Change to BRISQUE

2.4 Analysis

3 Method

New method introduction depending on how the disposition will be in the final form

In order to answer the research questions stated in section 1.7 a state of the art SPC, experiments and evaluation methods needs to be set up. In this section the SPC hardware and image sensing and reconstruction scheme is described.

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 if all micro mirrors in the DMD are turned away from the single pixel sensor and towards the reference camera, it can also be used to check that the patterns are displayed correct on the DMD.

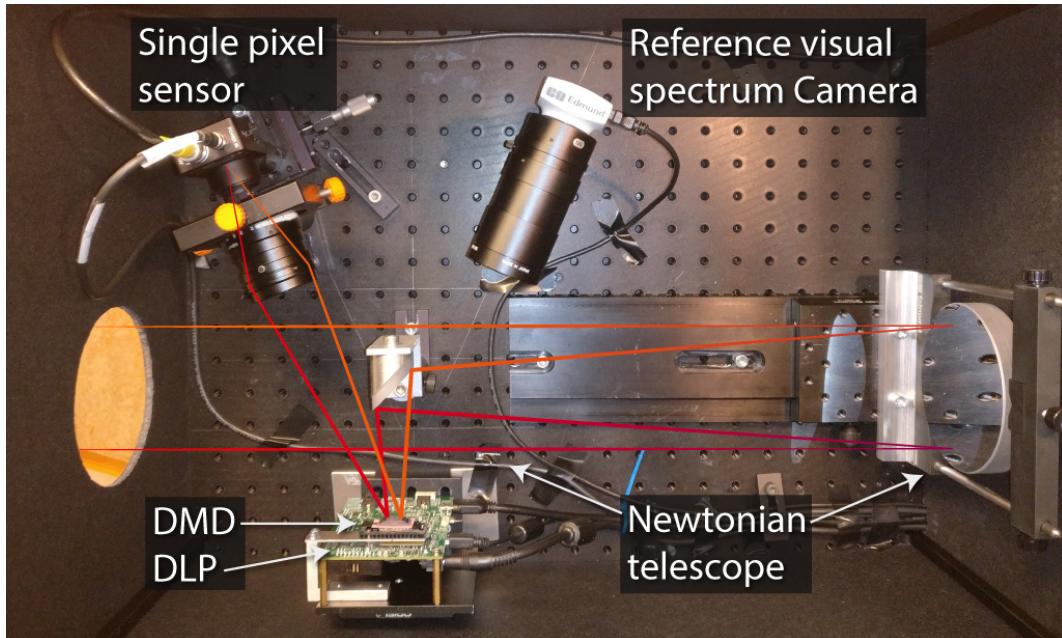


Figure 3: Single pixel imaging system (SPIS), adopted from [18].

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 turned 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 intensity. The reconstructed image from the system will have the same resolution as the DMD patterns. The DLP is the DMD control unit which controls which patterns are displayed on the DMD either by reading images from memory or the 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.

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 the number of measurements relative to number of pixels in reconstructed image between 20% – 30% a 256×256 pixel images data would be acquired in 9 – 13 seconds and for a 512×512 pixel image 36 – 53 seconds. To control the DMD the software 'DLP LightCrafter 4500 Control Software' is used.

The DMD 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.

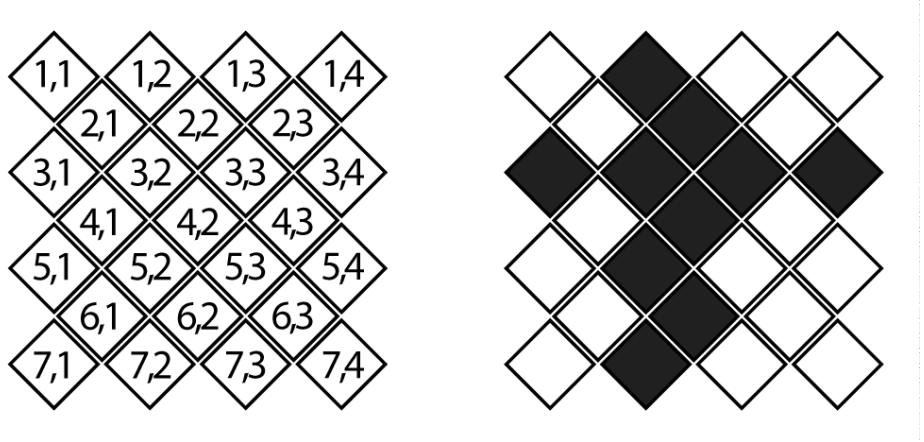


Figure 4: DMD matrix, left shows each tiles index and right shows second row and second column in black.

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 to 1 to 1.



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

Connect the DMD to CS and the physical aspect (every mirror is an pixel). I:I:D gaussian with DMD only can take values 1,0 gives 50% evenly distributed pixels measurements for every measurement matrix which looks something like figure 6.



Figure 6: A typical measurement matrix presented on the DMD with the resolution 256×256 pixels.

3.1.3 Lens

The lens mounted on the single pixel sensor is an 50mm 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. [19]

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. [20]

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

Write introduction to CS, create some intuition

The single pixel sensor captures a scene by measuring the light intensity focused into the detector reflected from the DMD matrix. The DMD sensing matrix changes to obtain new measurements, M unique sensing matrix measurements is captured to reconstruct an image with N pixels. Each sensing matrix index is encoded either by a one or a zero (turning the mirror onto or away from the sensor). The compressive imaging sampling model is defined as

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

, $\mathbf{x}_{N \times 1}$ is the signal (image) with N samples (pixels), $\mathbf{y}_{M \times 1}$ is the vector with M measurements, $\Phi_{M \times N}$ is the measurements matrix (each unique sensing matrix $\Phi_{1 \times N}$ as a row vector) and ϵ is the noise. In conventional sampling the number of measurements M needs to be at least equal to the number of samples N to recover the signal but CS states that M can be relatively small compared to N given how compressible the signal is. The signal \mathbf{x} can be represented as

$$\Psi \theta = \mathbf{x}, \quad (4)$$

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 signal \mathbf{x} has K non zero elements in basis Ψ , $\|\theta\|_0 = K$. Given equation 4, equation 3 can be expand to

$$\mathbf{y} = \Phi \mathbf{x} + \epsilon = \Phi \Psi \theta + \epsilon = \mathbf{A} \theta + \epsilon, \quad (5)$$

where, $\mathbf{A}_{M \times N} = \Phi \Psi$ is the reconstruction matrix. The last statement is what makes CS powerful, a signal which is not sparse can be sampled with measurement matrix Φ and then reconstructed with reconstruction matrix \mathbf{A} in a basis where \mathbf{x} is sparse or compressible.

Noise does not effect the measurement very much because half the intensity from the image is measured in one pixel sensor bumping the signal to noise ratio from a conv camera where each pixel has some noise half the pixels share that noise making SPC very robust to noise even in low light situations.

3.3 Measurement matrix & Restricted isometry property (RIP)

Introduce the topic.

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

$$(1 - \delta_K) \|\mathbf{x}\|_{\ell_2}^2 \leq \|\mathbf{A} \mathbf{x}\|_{\ell_2}^2 \leq (1 + \delta_K) \|\mathbf{x}\|_{\ell_2}^2, \quad (6)$$

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. The solution is to find a general

reconstruction matrix that satisfies RIP with high probability. The solution which also should be incoherent with the base matrix Ψ is to construct the measurement matrix using a i.i.d random distribution which gives $\delta_K \ll 1$ with high probability. Using random measurement matrices the number of measurements needed to satisfy RIP with high probability is $M \geq O(K \log(N/K)) \ll N$.

The problem using random matrices is that they need to be stored in memory for the reconstruction algorithm, when the image resolution is increased the measurement matrix increases exponentially. For images with resolution of 512×512 and larger the data gets infeasible for a normal computer to handle. Fortunately using fast transforms in the reconstruction algorithm can exclude using vector multiplication resulting in faster reconstruction and the need to store the measurement matrix in memory. But in order to do so special measurement matrices are used, in this master's thesis the sequency ordered Walsh Hadamard measurement matrix will be used with the TVAL3 reconstruction algorithm described in section 3.4.1.

3.3.1 Sequency ordered Walsh Hadamard measurement matrix

Besides from eliminating the need to store the measuring matrix for reconstruction the sequency ordered Walsh Hadamard (SOWH) matrix can be generated when sent to the DMD eliminating the need to store the matrix at all. SOWH has the same characteristics and properties of an i.i.d random matrix and therefore also fulfills the RIP condition with high probability and research has shown that there is no significant loss in recovery of the signal relative i.i.d random measurement matrix [15]. An other property of SOHW is that it only contains -1 and 1 which easily be converted to 0 and 1 when sent to the DMD.

The naturally ordered Hadamard matrix of dimension 2^k , $k \in \mathbb{N}$ are constructed by the recursive formula

$$H_0 = 1, \quad (7)$$

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (8)$$

and in general,

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = H_1 \oplus H_{k-1} \quad (9)$$

where \oplus denotes the Kronecker product. To construct the 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.

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

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \Rightarrow W_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \quad (10)$$

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 a -1 is shifted to 0. This last step is required to convert the measurement matrix so it gets the characteristics of an i.i.d random matrix and thus fulfill the RIP condition [13]. How the matrix was permuted and which rows was choosen i which order is stored so the reconstruction algorithm can use that information. [11], [13], [15].

3.4 Reconstruction method

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

$$\hat{\theta} = \arg \min \|\theta\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\theta. \quad (11)$$

Simply minimizing nonzero indices 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),

$$\hat{\theta} = \arg \min \|\theta\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}\theta\|_2 < \epsilon. \quad (12)$$

In 2006 Donoho [3] for the fist time guarantied 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 require more measurements than the optimal case with ℓ_0 but it is still $M \ll N$. Since 2006 many more types of optimization algorithms has evolved which solves the problem with different methods but with the same goal: finding the largest most significant coefficients of θ .

3.4.1 Total variation: TVAL3

The reconstruction algorithm that was chosen in this Master's thesis was a total variation regularization algorithm. Natural images often contains sharp edges and piecewise smooth areas which the TV regularization algorithm is good at preserving. The main difference between TV an 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

$$\min_{\mathbf{x}} \sum_i \|D_i \mathbf{x}\|, \text{ subject to } \Phi \mathbf{x} = \mathbf{y}, \mathbf{x} \geq 0, \quad (13)$$

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", accordingly is a TV regularization algorithm which uses augmented Lagrangian and alternating direction methods, where augmented Lagrangian is a method in optimization for solving constrained problems by substitute the original constrained problem with a series unconstrained subproblems and introduce a penalty term. To solve the new subproblems the alterning direction method is used [11].

As mentioned earlier in section 3.3.1 the main reason why the sequency ordered Walsh Hadamard matrix is used is to eliminate the need to store the matrix in memory during reconstruction and a

promise to speed up the reconstruction. In TVAL3 there are two multiplications between matrix and a vector that dominates the computation time,

$$\Phi \mathbf{x}^k \text{ and } \Phi^\top (\Phi \mathbf{x}^k - \mathbf{y}). \quad (14)$$

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),

$$\mathbf{W}\mathbf{x} = \text{fwht}(\mathbf{x}), \quad (15)$$

where \mathbf{W} is a sequency ordered Walsh Hadamard matrix and \mathbf{x} is the signal vector. The wht is a generalized class of Fourier transforms which decomposes input vector into superposition of Walsh functions.

From section 3.3.1 it was briefly mention in the last paragraph that in order for the measurement matrix to fulfill RIP the columns is permuted and rows are chosen in random to create the measurement matrix from the sequency ordered Walsh Hadamard matrix, two functions is created to carry out does operations. First the permutation function $\pi(\cdot)$, which from a random seed permute the order of the columns in a matrix or the order of a vector. The second function $\Pi_M(\cdot)$ chooses M row in a matrix at random and stacks them in a new matrix. Then the definition of the measurement matrix Φ constructed from the sequency ordered Walsh Hadamard matrix \mathbf{W} leads to observation 2

$$\Phi = \pi(\Pi_M(\mathbf{W})) = \Pi_M(\pi(\mathbf{W})). \quad (16)$$

It does not matter in which order the functions i applied, it gives the same result. With matrix \mathbf{A} and vector \mathbf{u} observation 3 is formulated as,

$$\pi(\mathbf{A})\mathbf{u} = \mathbf{A}\pi(\mathbf{u}), \quad (17)$$

which shows that there is no difference between multiply a column-permuted matrix with a vector and multiply the same matrix with the vector permuted.

With all observations combined the matrix multiplication is replaced with the fwht in observation 4

$$\mathbf{y} = \Phi \mathbf{x} = \pi(\Pi_M(\mathbf{W}))\mathbf{x} = \Pi_M(\mathbf{W})\pi(\mathbf{x}) = \Pi_M(\mathbf{W}\pi(\mathbf{x})) = \Pi_M(\text{fwht}(\pi(\mathbf{x}))), \quad (18)$$

with the conclusion that the multiplication between the measurement matrix constructed using the permuted sequency ordered Walsh Hadamard matrix and the signal can be performed with the signal permuted, fast transformed using fwht and choosing rows, both permutations using the same functions $\pi(\cdot)$ and $\Pi_M(\cdot)$ and random seed as when the measurement matrix was created.

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. Excluding the measurement matrix in the reconstruction results in larger resolution images (512×512 pixels and larger) can be reconstructed.

3.5 Implementation

3.6 Image capturing and processing chain

To capture an image the SPC setup is only a subsystem in the whole process from acquiring the signal to reconstructing the image. 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 has no full 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 subsection.

3.6.1 Prepare the SPC

The first step in the yellow block "Prepare the SPC" 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 visual 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.

3.6.2 Sampling

The red blocks subsystem "Start sampling signal from SWIR photo diode" and "Store the raw signal" 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 (measurement matrix) 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 3.6.4. After the signal is sampled the obtained signal needs to be stored on the computer manually.

3.6.3 Streaming patterns to the DMD

The subsystem "Streaming patterns to the DMD" 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 sampled with $M/N = 20\%$ measurement matrices gives 52428 patterns which would take

$$52428/60 = 874 \text{ seconds} = 14.5 \text{ minutes to sample which is a long exposure time for a still image with the constraint that the scene should be stationary to obtain a stationary signal.}$$

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 is 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. This 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.

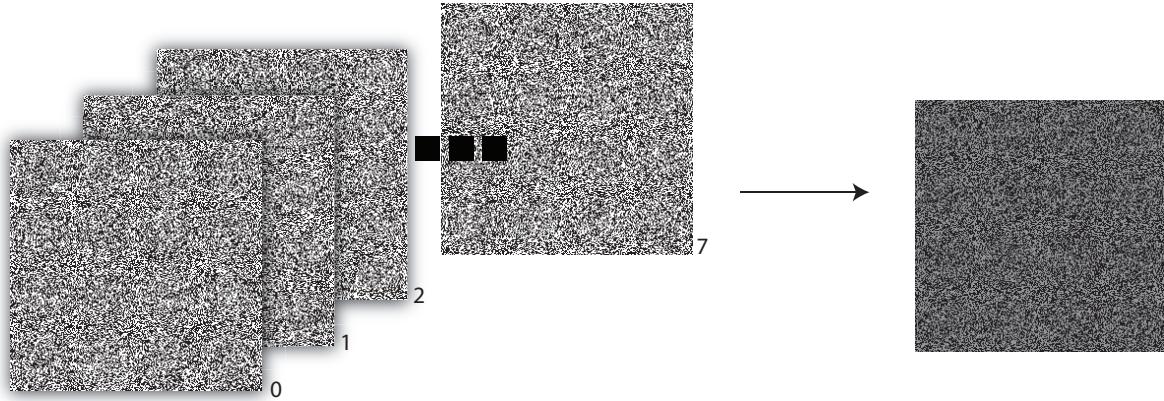


Figure 8: Stack each group of 8 measurement matrices in separate bit planes creating one 8 bit image with each matrix in one bit plane.

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.

This 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.

3.6.4 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.

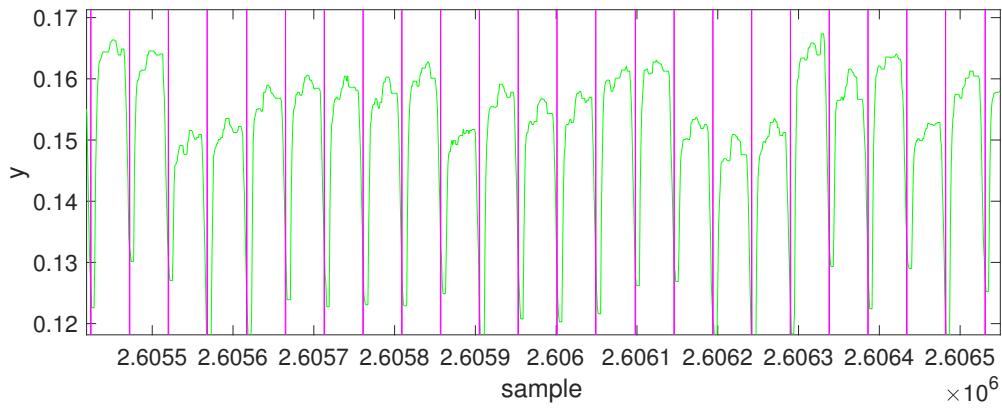


Figure 10: —

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 the purple line divides the measurements the DMD is changing pattern which gives an uncertain signal. With

the omitted parts of the signal corresponding to one measurement the mean is calculated and set to the value for each measurement, as seen in figure 11.

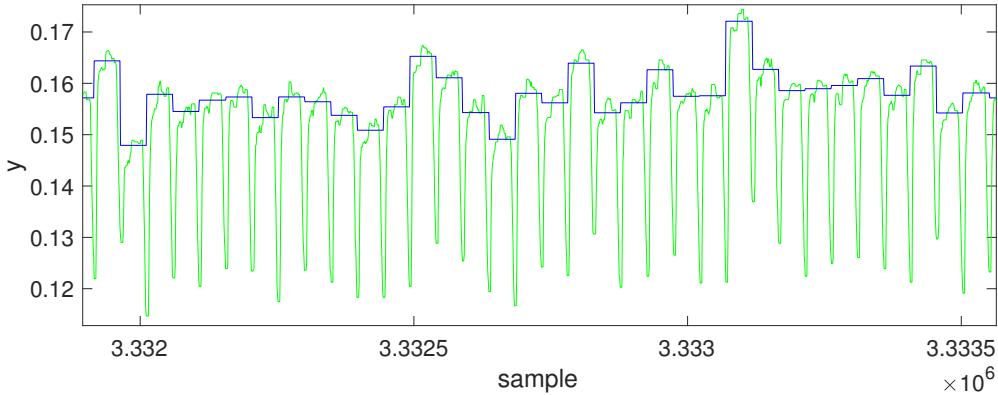


Figure 11: —

So far the 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 is assumed to be constant with half the pixels measured at random uniformed spread across the image. Each pixel in the image has the same value in each measurement with some noise added which is divided by all pixels being measured i.e. half the pixels. But when capturing images outdoors with natural light and scenes in addition with long exposure time as described in section 3.6.3 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 structure movement.

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 added for all pixels being measured changing the mean intensity of the measured signal \mathbf{y} which should be stationary. can change the mean intensity of the measurements over time which reduce the reconstruction performance.

Due to the long exposure time two types of situation can arise where the image Scenes where a lot of movement or the risk of movement is high can be avoided

3.6.5 Reconstruction

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

- $opts.mu = 2^8$
- $opts.beta = 2^5$
- $opts.maxcnt = 10$
- $opts.tol_inn = 10^{-5}$
- $opts.tol = 10^{-10}$

- $opts.maxit = 1000$
- $opts.mu0 = 2^4$
- $opts.beta0 = 2^0$
- $opts.nonneg = \text{true}$
- $opts.isreal = \text{true}$

Which solves for a real non negative solution.

3.6.6 Image processing

From the reconstructed image some light image processing is done. There are only two operations applied to the reconstructed image and the reason why is that the images presented in the evaluation 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 there exist such problems. 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.

3.7 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.

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.

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

$$\text{PSNR}[f(x, y), g(x, y)] = 10\log_{10} \frac{E^2}{\text{MSE}[f(x, y), g(x, y)]} \quad (19)$$

where, $f(x, y)$ and $g(x, y)$ is intensity in pixel $(x, y)...$

Large syntetic test of SWIR images, Whar result can be expected

Synthetic test of the methods presented in section 3.6.4 will be conducted to measure the performance and validity.

Number of measurements needed to get a good image

Homography test

Brisque

MTF which measures Edge response

3.8 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.

4 Evaluation

This section is structured as, for each experiment and setup a detailed explanation and motivation on how and why the experiment is needed followed by the results of that experiment. The experiments are motivated by gathering as much information and results as possible to answer the research questions. The first subsection 4.1 will present the results from experiments with synthetic data where a reference image is available. The second subsection 4.2 will present the result from images reconstructed from the SPC. No perfect reference image is available in those experiments therefore the images will be evaluated against near optimal image, no reference QA and against a state of the art SWIR camera.

4.1 Synthetic data

4.1.1 PSNR, SNR, SSIM

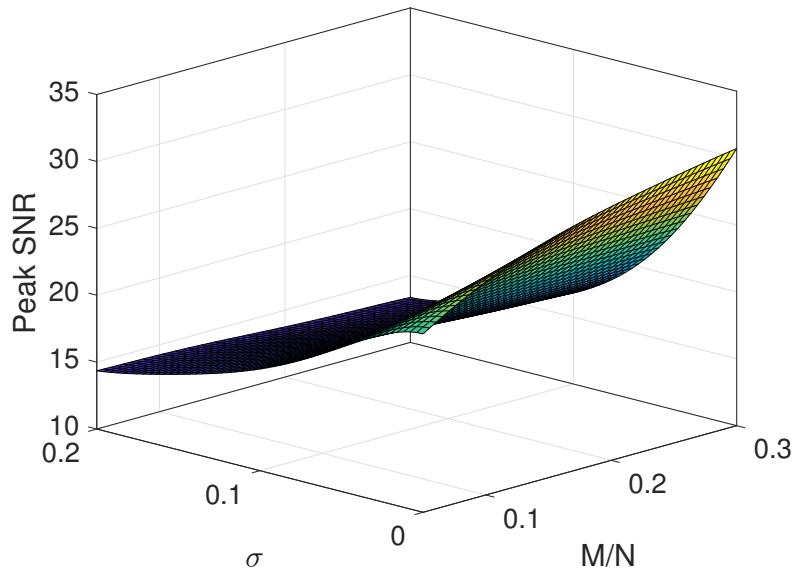


Figure 12: Peak SNR result depending on number of measurements and simulated noise level.

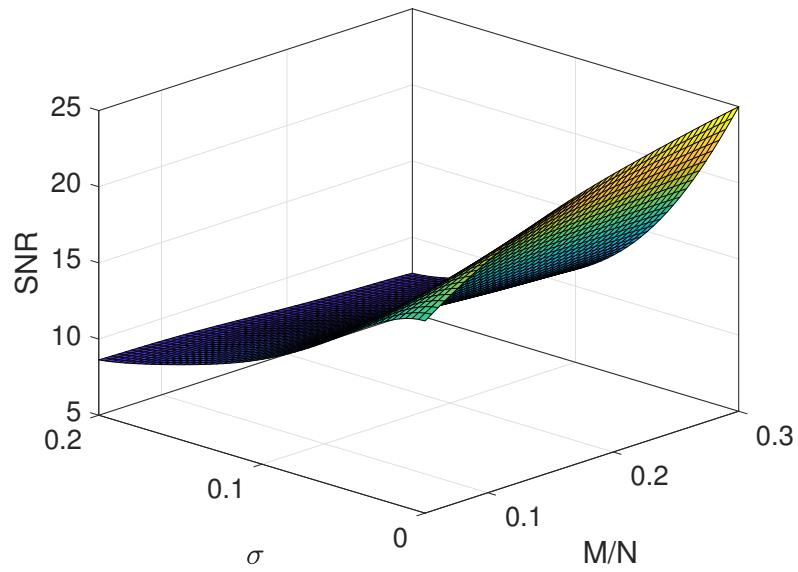


Figure 13: SNR result depending on number of measurements and simulated noise level.

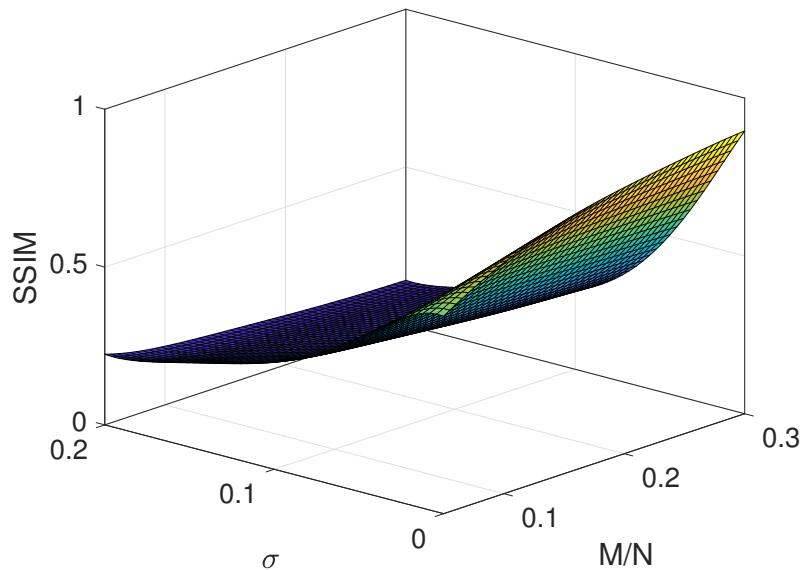


Figure 14: SSIM result depending on number of measurements and simulated noise level.

4.1.2 No Reference quality assessment

BRISQUE lower score is better.

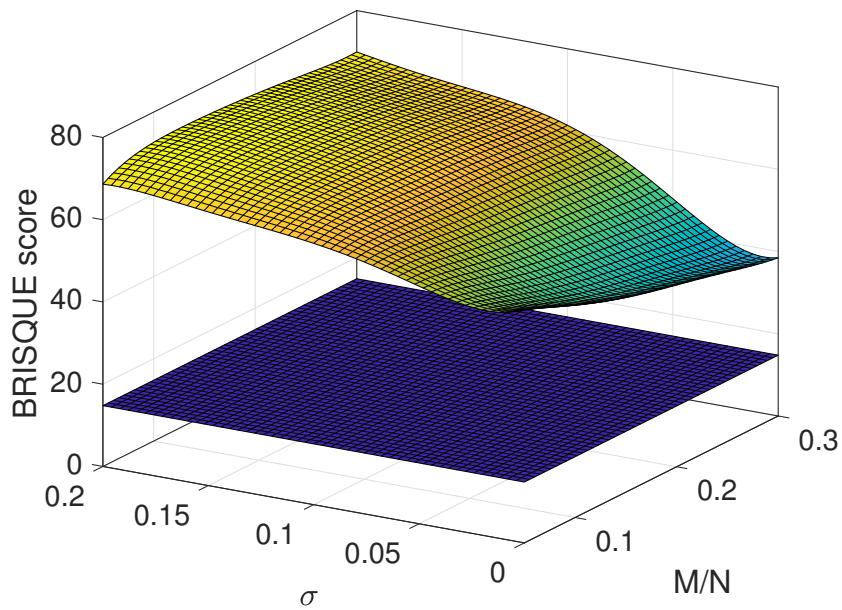


Figure 15: BRISQUE result depending on number of measurements and simulated noise level. Lower surface is reference image score.

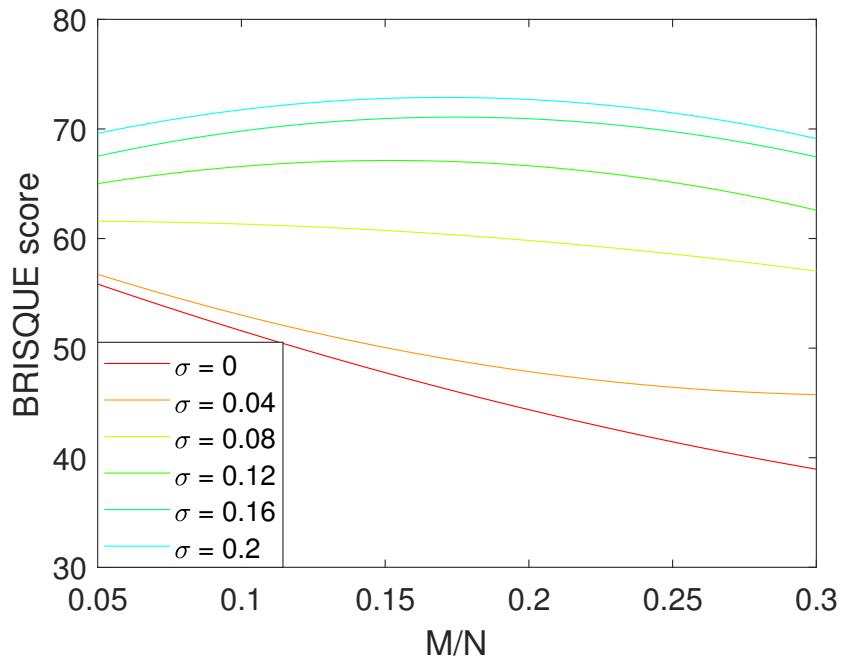


Figure 16: BRISQUE result depending on number of measurements for different simulated noise levels.

Add "optimal performance into the graph"

4.1.3 Dynamics in scene

Dynamics in the scene can roughly be divided into three separate scenarios, in this section each of them will be tested in a controlled environment with each scenario isolated to show how the signal and the reconstructed image is effected.

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



(a) Original reference image

(b) Reconstructed 30% image from reference image without movement

(c) Reconstructed 30% image with local movement

Figure 17: Local movement

The difference between figure 17b and 17c is visible with the naked eye, not only does the object moving around get blurry and noisy but the whole image globally. In table 2...

Peak SNR	SNR	SSIM
29	25	91

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

Commenting the result from the table... In figure 18 the effects of the movement is shown plotted against the non perturbed signal.

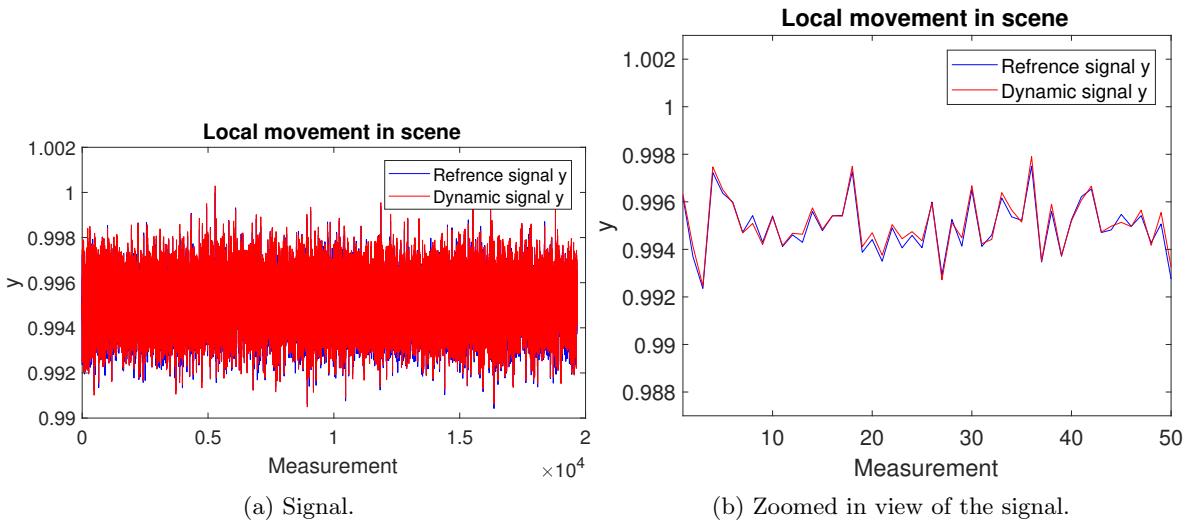


Figure 18: Local movement, acquired signal

As seen in figure 18a there is no obvious difference between the non perturbed reference signal and the distorted signal. In figure 17b where some of the samples is displayed no large difference can be seen ether, the conclusion of this test implies that local movement in a scene will cause noise in the image globally and especially locally where the movement occurred. It also implies that local movement is very hard to detect on the signal even if a reference signal is available.

The second scenario is an object is passing through, moves out or moves to an other place in the scene far from the original place. In other words, large global movement in the scene. The problem is modeled with a static background then as the simulated measurement is acquired the same object as in the first experiment 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, corresponding to approximately 0.7 seconds when capturing with the SPC in its current setup.



Figure 19: Object passing trough scene.

The difference between figure 19b and 19c is visible with the naked eye, A global noise arises in the image and the object cant be seen. In table 3...

Peak SNR	SNR	SSIM
23	18	58

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

Commenting the result from the table... In figure 20 the effects of the movement is shown plotted against the non perturbed signal.

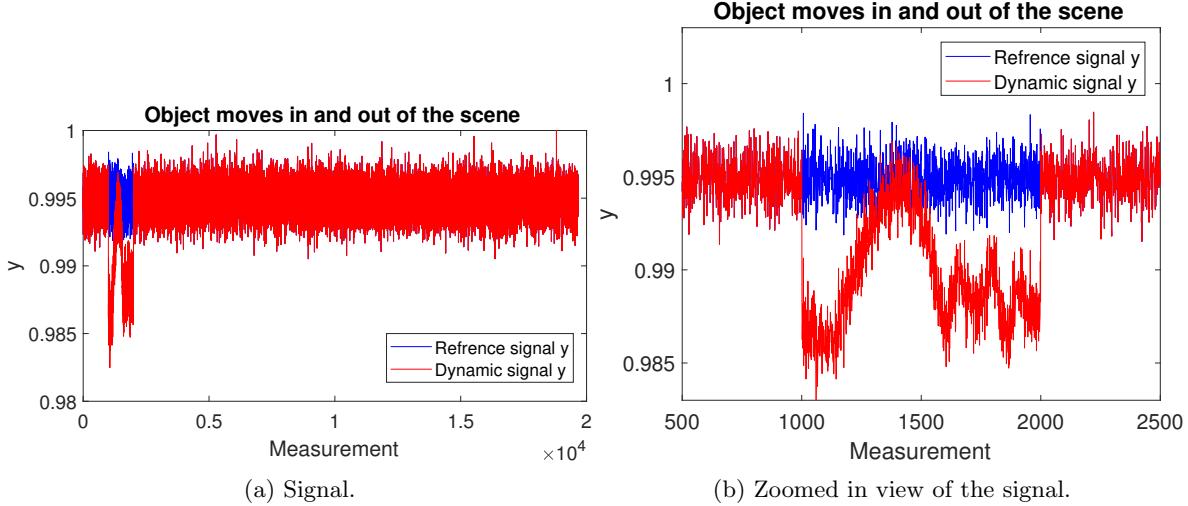


Figure 20: Global movement, acquired signal

- Large changes in the scene can be detected
- Remove the identified measurements to get a good signal

The third scenario is luminance change in the scene caused by clouds occludes the sun or the light intensity from the lights is not constant. This scenario is modeled by adding or subtracting the global intensity in the image over the measurements.

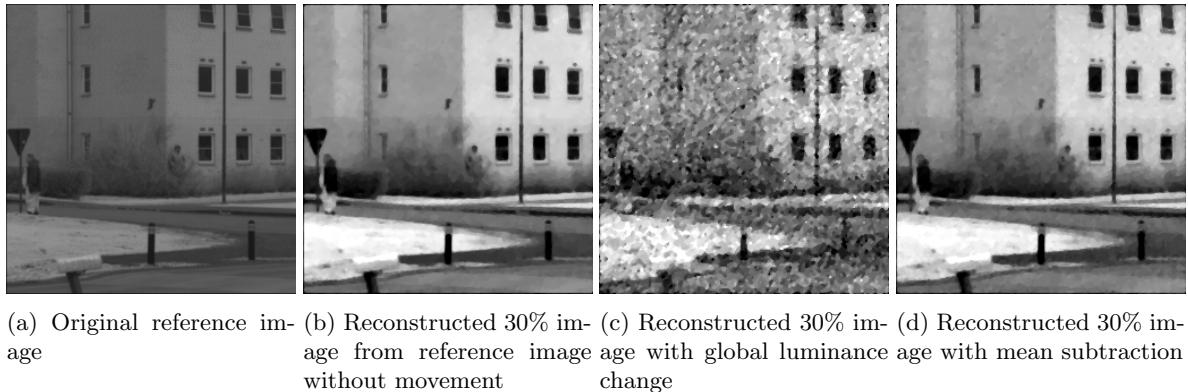


Figure 21: Global luminance change in scene.

The difference between figure 21b and 21c is visible with the naked eye, A global noise arises in the image, but as seen in figure 21d the effect can be suppressed explained under figure 22. In table 3...

	Peak SNR	SNR	SSIM
Perturbed signal	19	14	38
Mean subtracted signal	33	29	93

Table 4: Effects comparing non perturbed reconstructed image against reconstructed image with global luminance change

Commenting the result from the table... In figure 22 the effects of global luminance is shown plotted against the non perturbed signal.

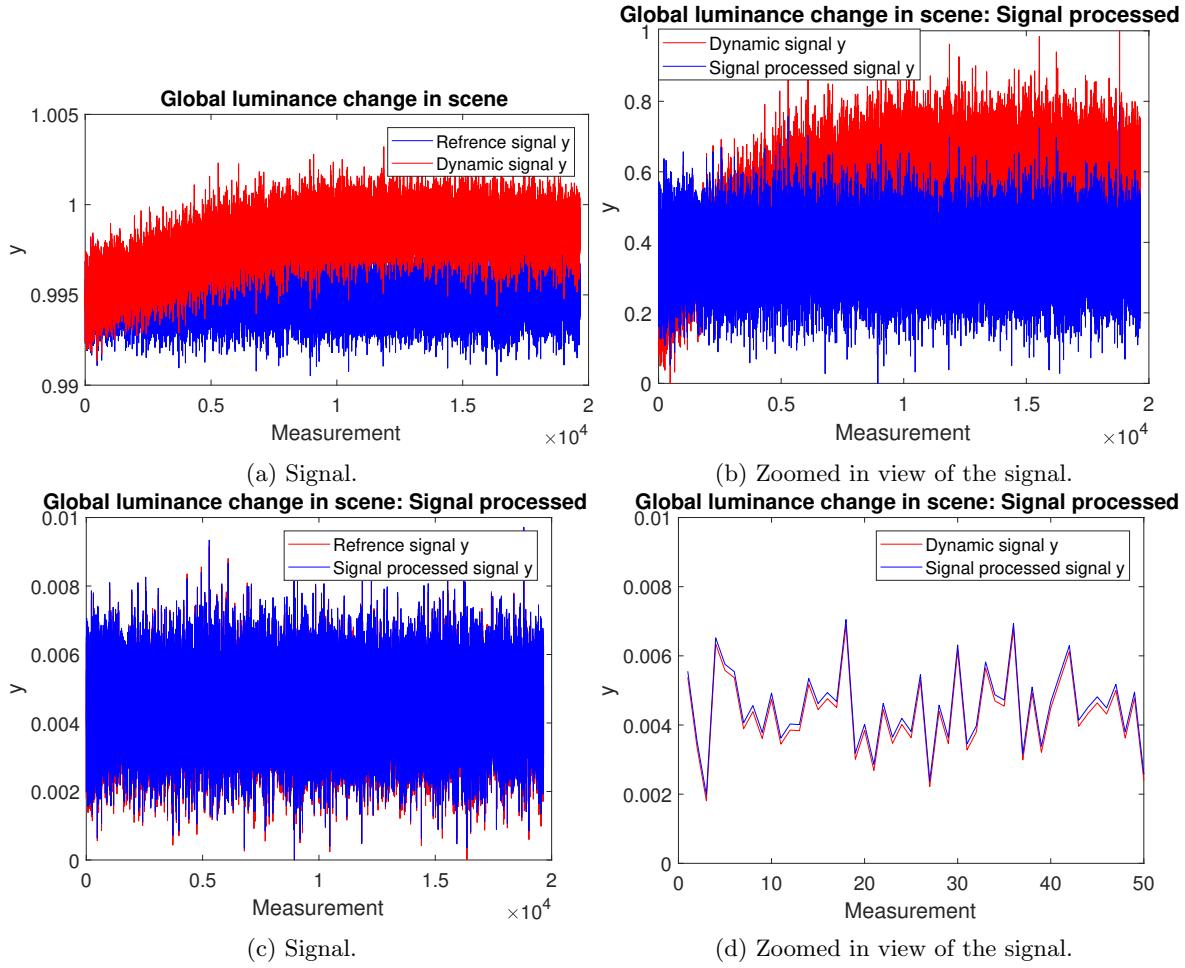


Figure 22: Global movement, acquired signal

- Dynamic signal v. Reference signal
- Dynamic signal v. Mean subtracted signal
- Reference signal v. Mean subtracted signal
- Comment on the window, pretty good.
- Can be detected with the knowledge that the signal should be stationary. Signal process the signal to look like a stationary signal.

4.2 SPC evaluation

4.2.1 Number of measurements



(a) m15

(b) m10

(c) m10



Figure 24: Images reconstructed using $M/N = 5\%$ to 30% measurements from top down.

4.2.2 Soft chessboard

Todo: Skapa rekonstruerade bilder från homographin och jämför de rekonstruerade med referensbilden

This evaluation is designed to confirm that the images reconstructed by the SPC follows the same characteristics as the reconstruction of the synthetic data.

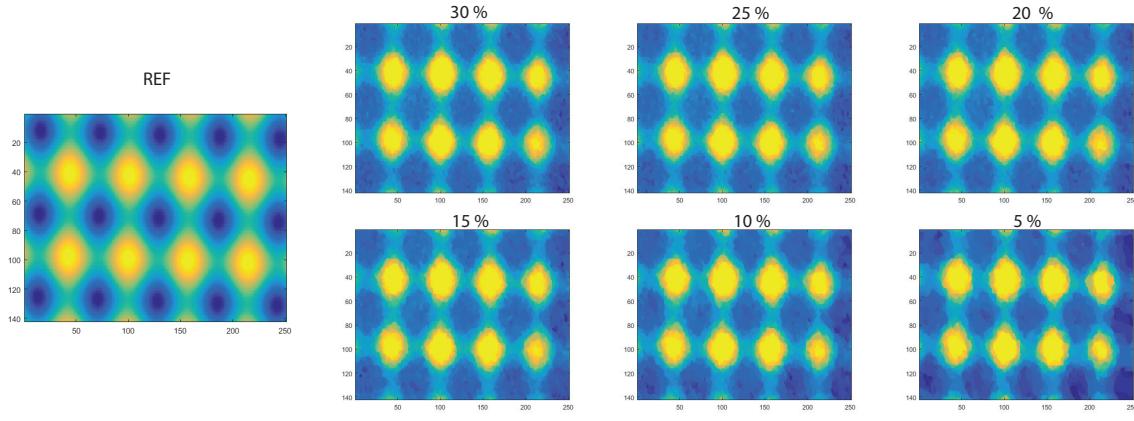


Figure 25: The reconstructed images with different number of measurements and the reference image transformed to fit the SPC images using homography.

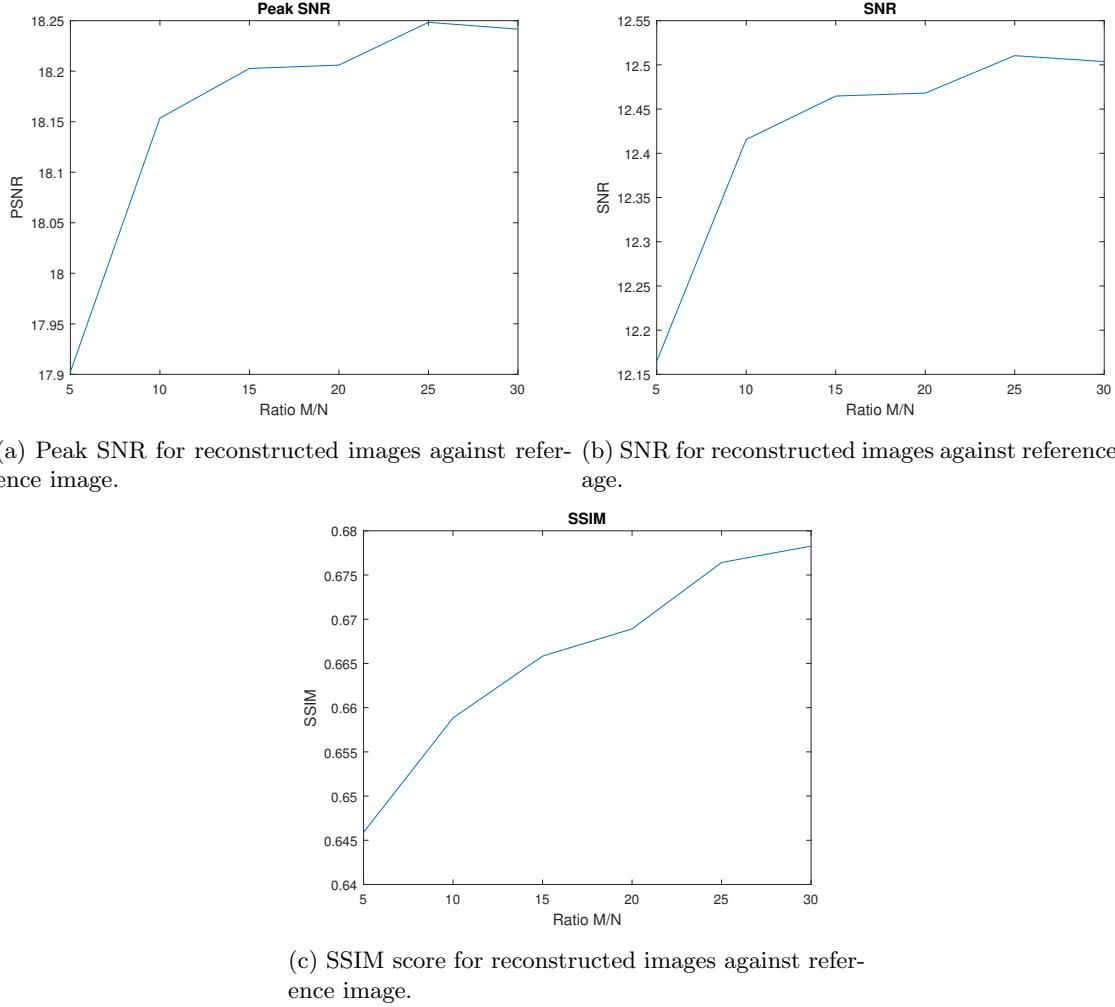


Figure 26: Signal quality of SPC images compared to reference image

4.2.3 No reference quality assessment

Using the no reference quality assessment measurement BRISQUE to evaluate the SPC images. Each image is evaluated at reconstruction rate 5% to 30%.

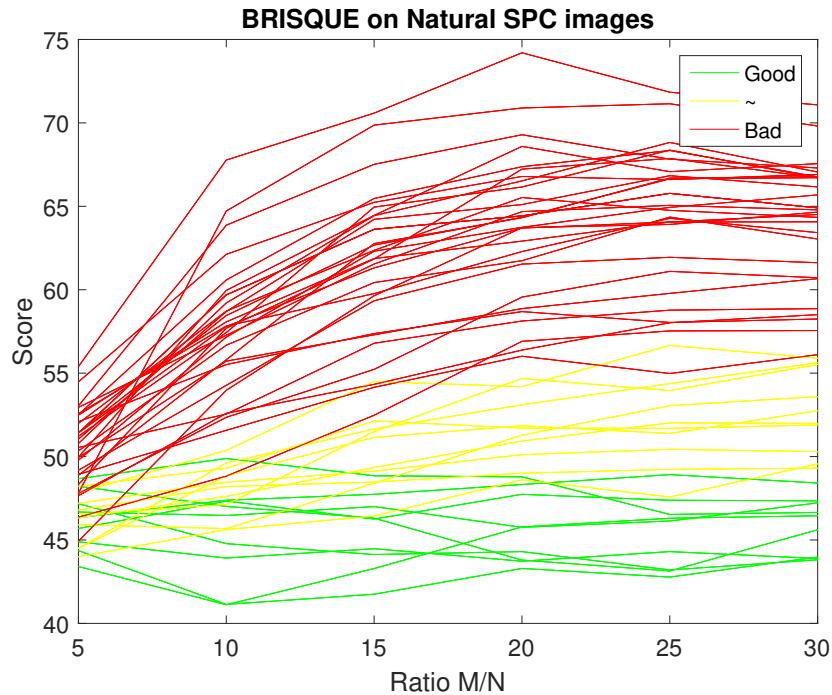


Figure 27: BRISQUE result.



Figure 28: Example of 'good' images corresponding to the green lines in figure 27.

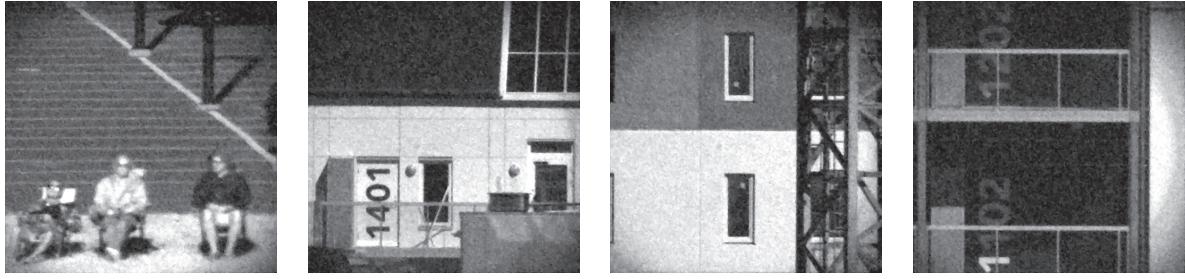


Figure 29: Example of 'medium good' images corresponding to the yellow lines in figure 27.

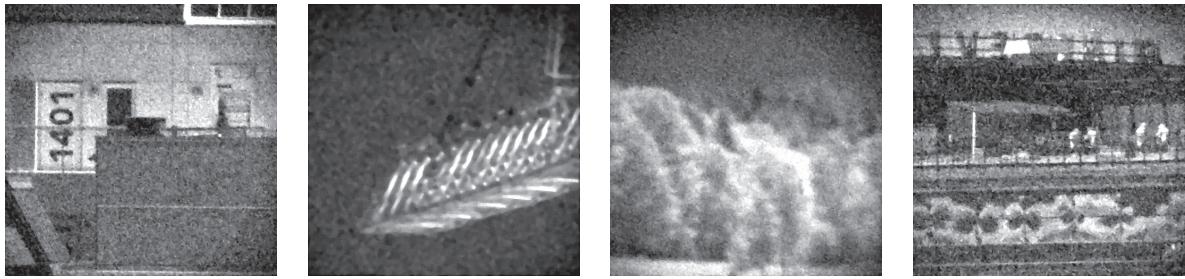


Figure 30: Example of 'bad' images corresponding to the red lines in figure 27.

- Good images are:
- Medium good images are:
- Bad images are:

4.2.4 Modulation Transfer Function

The MTF is used to comparing the sharpness of cameras and lenses.

The MTF from the SPC is compared to a state of the art SWIR camera. Two scenes was captured by the SPC and a conventional SWIR camera containing printed sheath of paper with simple tilted shapes on them, see figure 31.

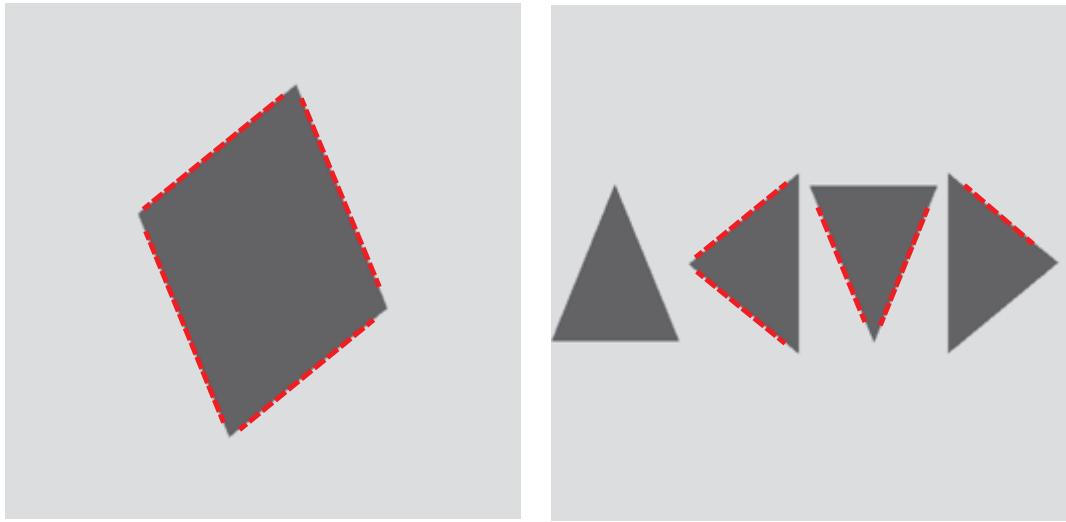


Figure 31: Printed targets with markings where the MTF measurements was performed

In the resulting images MTF measurements was performed on the specified edges to gather a mean and standard deviation for each camera. For the SPC, images reconstructed from 5% to 30% was tested in order to see if the number of measurements effected the MTF result. In figure 32 the images from the SWIR camera and SPC are presented.

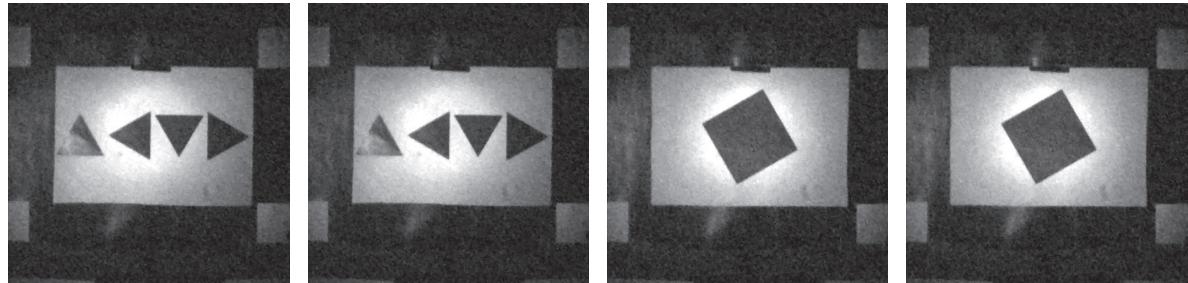


Figure 32: SPC and state of the art SWIR camera output images. (OBS! Bilder från Raptorn ska läggas till)

4.3 MTF

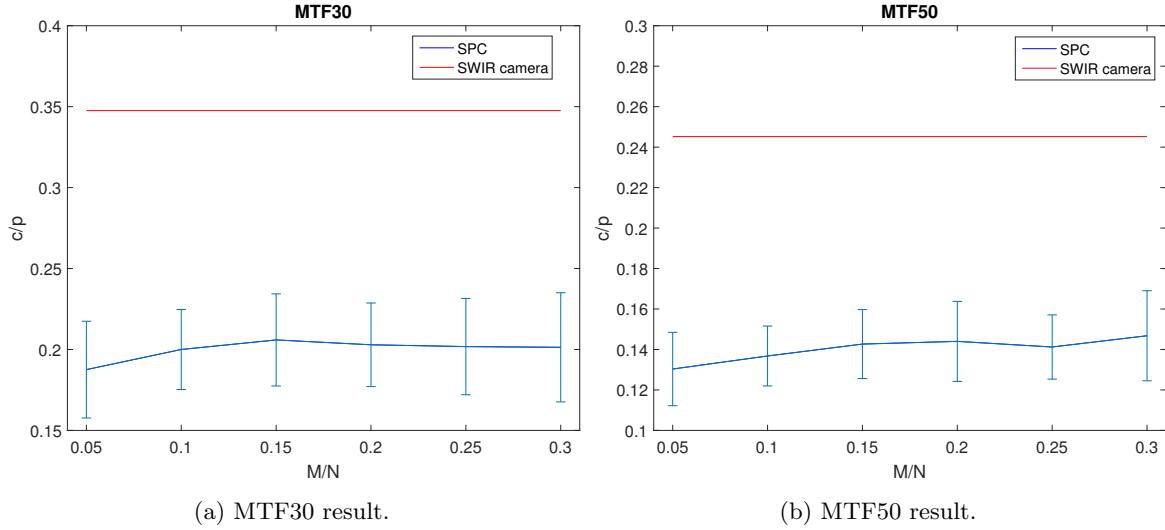


Figure 33: MTF results. (OBS! inte rätt figurer)

4.4 Edge response

The edge response is measured in the distance (pixels) required for the edge to rise from 10% to 90%. In figure 34 the result from the experiment is presented.

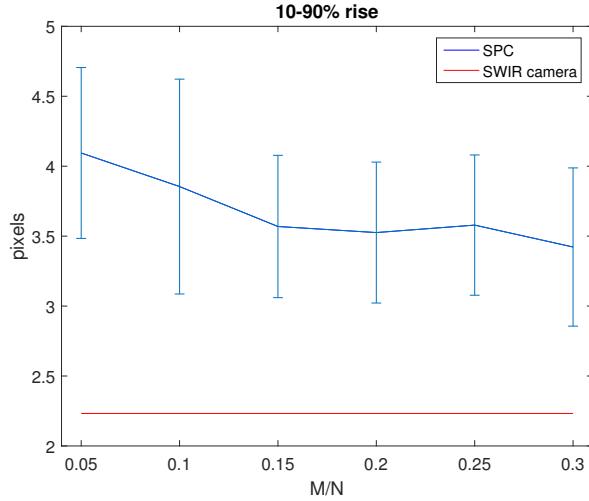


Figure 34: 10-90% rise in pixels. (OBS! inte rätt figur)

References

- [1] G. Y. G. Irina Rish, *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] 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.
- [5] R. M. Willett, R. F. Marcia, and J. M. Nichols, “Compressed sensing for optical imaging systems: A tutorial,” *SPIE Optical Engineering*, vol. 50(7), 2011.
- [6] D. Takhar, J. N. Laska, M. B. Wakin, 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] 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.
- [8] B. C. L. McMackin M. A. Herman and M. Weldon, “A high-resolution swir camera via compressed sensing,” *SPIE*, vol. 8353, 2012.
- [9] E. Fall, “Compressed sensing for 3d laser radar,” Linköpings University, Master’s thesis, 2014, ISRN: LiTH-ISY-EX-14/4767-SE.
- [10] T. Edeler, K. Olinger, S. Hussmann, and A. Mertins, “Multi image super resolution using compressed sensing,” *ICASSP*, pp. 2868–2871, 2011.
- [11] C. Li, “An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing.,” 2009.
- [12] “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.
- [13] T.T.Do, L. Gan, N. Nguyen, and T. Tran, “Fast and efficent 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.
- [14] 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.
- [15] 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.
- [16] A. Bovik, *The Essential Guide to Image Processing*. Elsevier, 2009, ISBN: 978-0-12-374457-9.
- [17] L. Zhang, L. Zhang, and A. C. Bovik, “A feature-enriched completely blind image quality evaluator,” *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015, ISSN: 1057-7149. DOI: 10.1109/TIP.2015.2426416.
- [18] C. Brännlund and D. Gustafsson, “Single pixel swir imaging using compressed sensing,” *Symposium on Image Analysis, 14-15 March 2017, Linköping, Sweden*, 2017.
- [19] L. Duval, *50mm f1 swir fixed focal length lens*, Edmund Optics Inc, 2017. [Online]. Available: <https://www.edmundoptics.com/imaging-lenses/fixed-focal-length-lenses/50mm-f1-swir-fixed-focal-length-lens/> (visited on 05/17/2017).

- [20] *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).
- [21] “Maximum length sequence encoded hadamard measurement paradigm for compressed sensing.,” *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Advanced Intelligent Mechatronics (AIM) Besancon, France, July 8-11, 2014*, pp. 1151–1156, 2014, ISSN: 978-1-4799-5736-1.