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 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 then 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 $\mathbf{y} = \mathbf{\Phi x}$ where $\mathbf{y}$ contains the measurements from the measurement matrix $\mathbf{\Phi}$ sensing the signal $\mathbf{x}$. In figure 1 such linear equation system is shown.

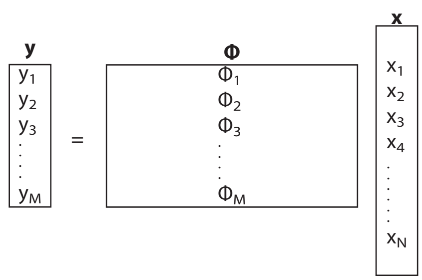


Figure 1: CS undetermined linear system

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 (Descrete cosine transform), JPEG2000 (Wavelet). The second constraint is that the measurement matrix must be incoherent with the sparse transform which for example white noise or some structure with the same property as white noise.

1.2 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. So the SPC has an digital micromirror array 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 ether reflect light 12 degrees to the right or left as seen in figure 2. The incomming focused light can ether be dumped or it can be reflected into the single pixel detector through an lens.

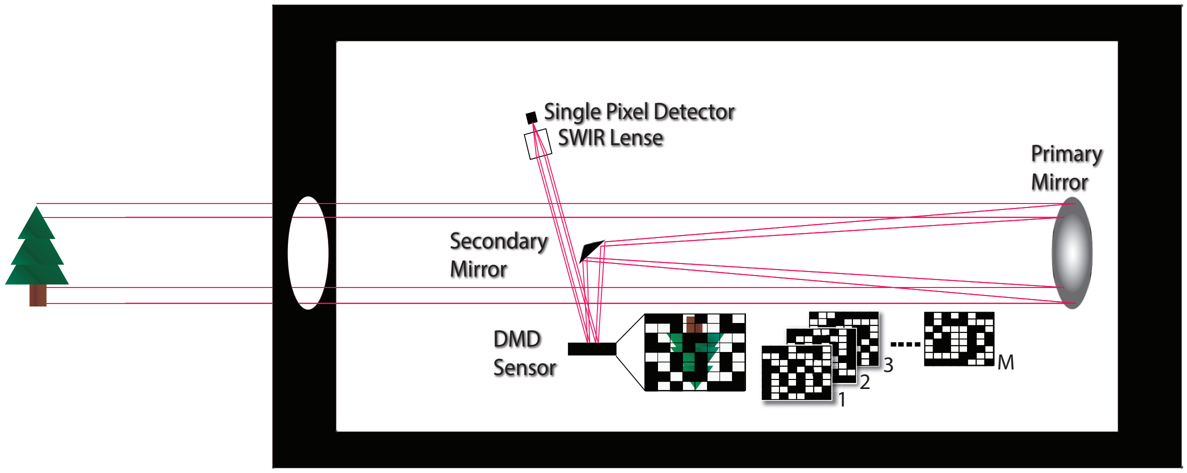


Figure 2: System overview

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 $\mathbf{y} = \mathbf{\Phi x}$.

1.3 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 he 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,

$$\mathbf{y} = \mathbf{\Phi x} \Leftrightarrow \mathbf{y} = \mathbf{\Phi \Psi \Theta}, \tag{1}$$

where $\mathbf{\Psi}$ is a sparsitfying basis for example to the DCT or Wavelet basis. And $\mathbf{\Theta}$ is the coefficients vector which is more sparse then the spatial coefficent vector $\mathbf{x}$. And the transformation will not comprimize the incoherence between the reconstruction matrix $A = \Psi\Phi$ and the coeficents 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 i 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 as 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,

$$\widehat{\mathbf{\Theta}} = \arg\,\min ||\mathbf{\Theta}||_{\ell_1} \text{ subject to } \mathbf{\Psi \Theta}\Theta = y, \tag{2}$$

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.

1.4 Motivation

Why would a SPC be beneficial to a conventional camera? The SPC has more components to work 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.

- Why?
- Compress before the images is taken
- Cheaper and in some cases only possible case to go to larger resolutions
- It is more effective then scanning each pixel one by one because the one need to do N samples.

1.5 Aim

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

1.6 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?

1.7 Limitations

- The hardware rig provided by FOI
-

1.8 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 Compressed imaging

- In **article:CS˙donoho1** David L. Donoho proposed the framework of compressed sensing and its application to images.
-
- **article:SRM** Fast and Efficient Compressive Sensing Using Structurally Random Matrices. Describes that SRM works and why.

2.2 Evaluation

- Al Boviks book the essential guide to image processing **book:image˙processing** 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.
- **article:il˙niqe** 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.

2.3 Analysis

3 Method

In order to answer the research questions stated in section 1.6 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 as well as the different evaluation metrics and methods.

3.1 Single pixel camera architecture and hardware

FOI designed the SWIR SPC platform using a DMD, a Newtonian telescope and a single pixel detector which are further described in section 3.1.1. 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.

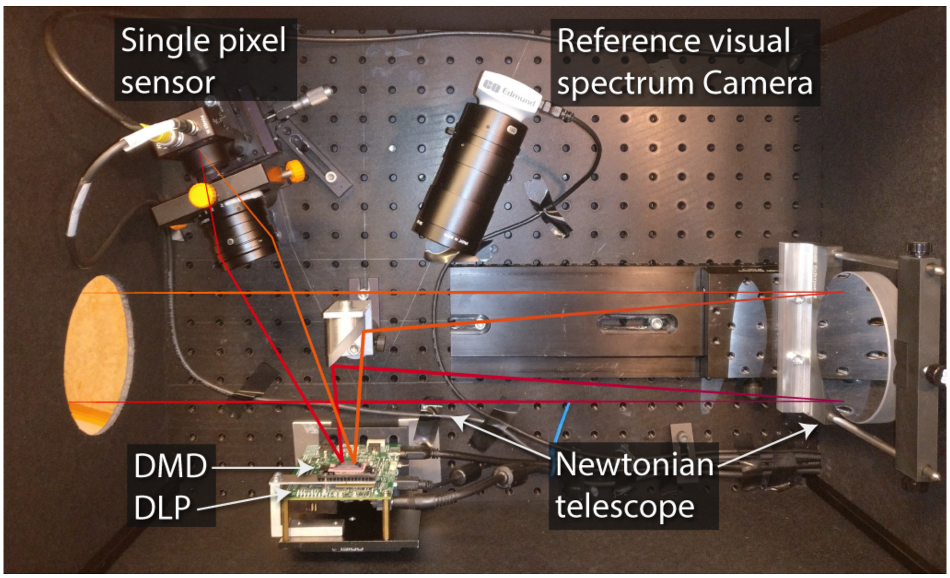


Figure 3: Single pixel imaging system (SPIS), adopted from **article:foiSPIS**

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 that are displayed on the DMD either by reading images from memory or the video port.

3.1.1 Hardware

3.1.2 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.

3.1.3 DLP and DMD

The DMD (Texas Instruments DLP4500NIR) is a matrix of micro mirrors that can be individually tilted $\pm 12°$ 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 then 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, which gives a maximum measurement matrix rate at $120 * 24 = 2880$ Hz, but in the current configuration only 60 Hz frame rate was achieved giving a measurement matrix rate at 1440 Hz. Because the diamond shape of the mirrors and how their index is defined in the DLP where one column is two mirror column arrays wide, see figure 4, the reconstructed image needs to be reshaped to a Cartesian grid. To control the DMD the software 'DLP LightCrafter 4500 Control Software' is used. **manual:DLP**

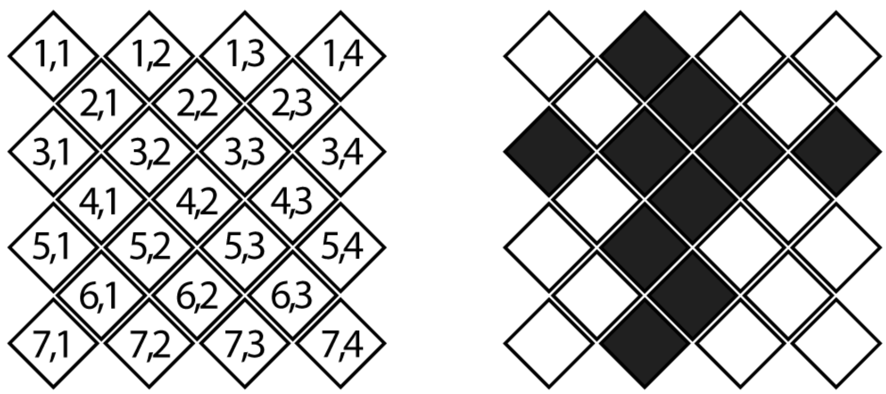


Figure 4: DMD matrix, left shows each tiles index and right shows second row and second column in black.

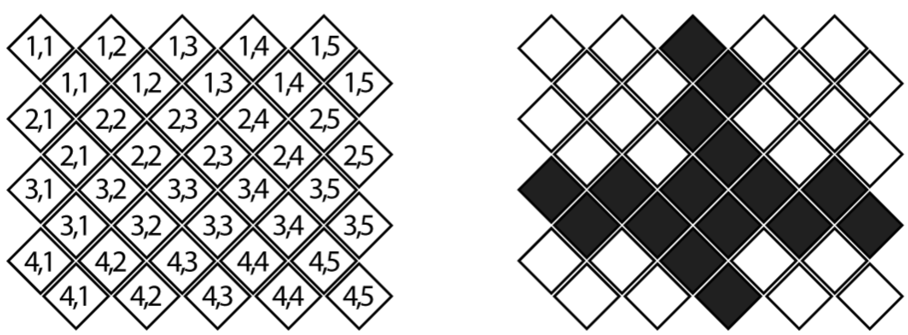


Figure 5: DMD matrix, left shows each tiles index and right shows second row and second column in black.

How does this effect the result?

3.1.4 Lens

The lens mounted on the single pixel sensor is an 50mm SWIR Fixed Focal Length Lens f1.4 designed for wavelengths raging from the 800 nm in the visual spectrum to 2000 nm in the SWIR spectrum. **website:SWIR˙objective**

3.1.5 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). **manual:PDA**

3.1.6 Signal spectrum

All components characteristics assembled should capture signals in 800-1700 nm spectrum.

3.2 Compressed imaging

The CS sampling model is defined as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \epsilon, \tag{3}$$

where, $\mathbf{x}_{N\times 1}$ is the signal with N samples, $\mathbf{y}_{M\times 1}$ is the vector with M measurements, $\mathbf{\Phi}_{M\times N}$ is the measurements matrix (each unique sensing matrix as a column vector $1 \times N$) 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

$$\mathbf{\Psi}\theta = \mathbf{x}, \tag{4}$$

where, $\mathbf{\Psi}_{N\times N}$ 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 $\mathbf{\Psi}$, $||\theta||_0 = K$. Given equation 4, equation 3 can be expand to

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

where, $A_{M\times N} = \Phi\Psi$ 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 the reconstructed with reconstruction matrix A in a basis where x is sparse.

3.3 Measurement matrix

3.3.1 Random

3.3.2 Walsh Hadamard

3.4 Reconstruction methods

3.4.1 ℓ_1 minimization

3.4.2 Greedy algorithms

3.4.3 Total variation

3.5 Evaluation

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 $\mathbf{y} = \mathbf{\Phi x} + \epsilon$ where, $\mathbf{x}$ is the synthetic image reshaped to a vector, $\mathbf{\Phi}$ 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-nise ratio (PSNR) and SSIM. PSNR is defined as

$$\text{PSNR}[f(x,y), g(x,y)] = 10log_{10}\frac{E^2}{\text{MSE}[f(x,y), g(x,y)]} \tag{6}$$

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

3.6 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

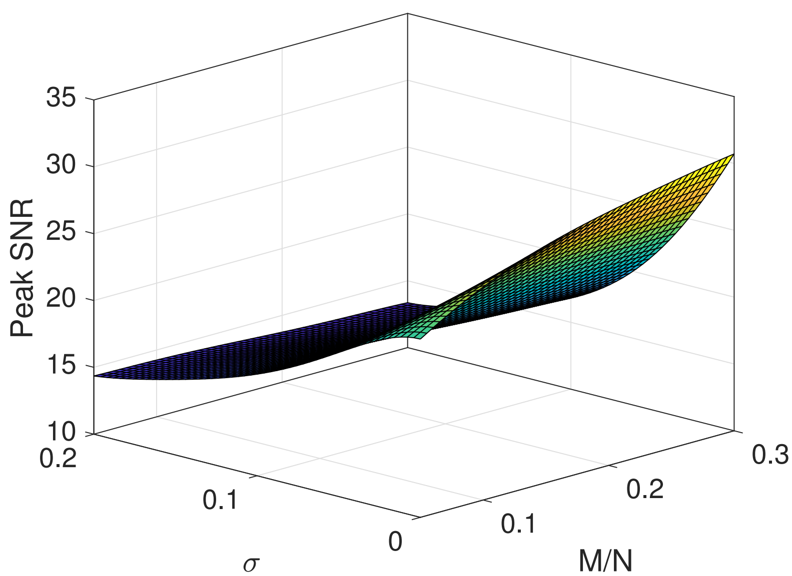


Figure 6: Peak SNR result depending on number of measurements and simulated noise level.

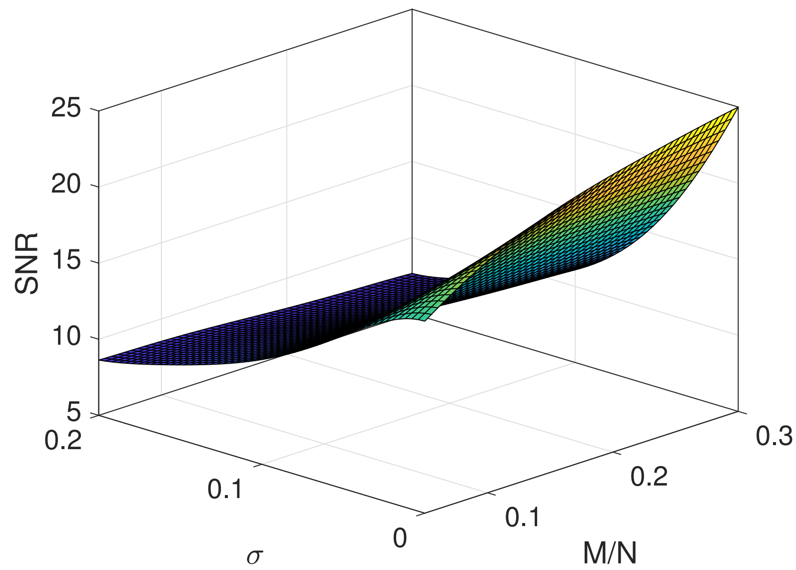


Figure 7: SNR result depending on number of measurements and simulated noise level.

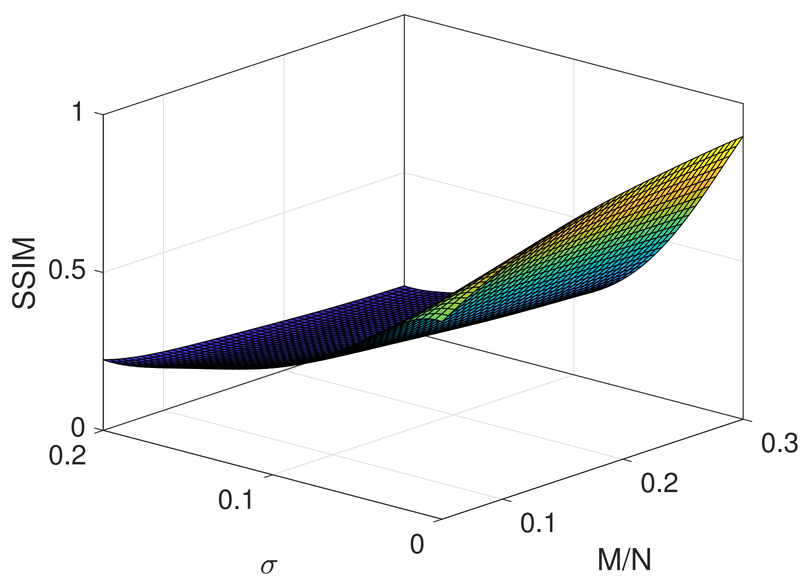


Figure 8: SSIM result depending on number of measurements and simulated noise level.

4.1.2 No Reference quality assessment

BRISQUE lower score is better.

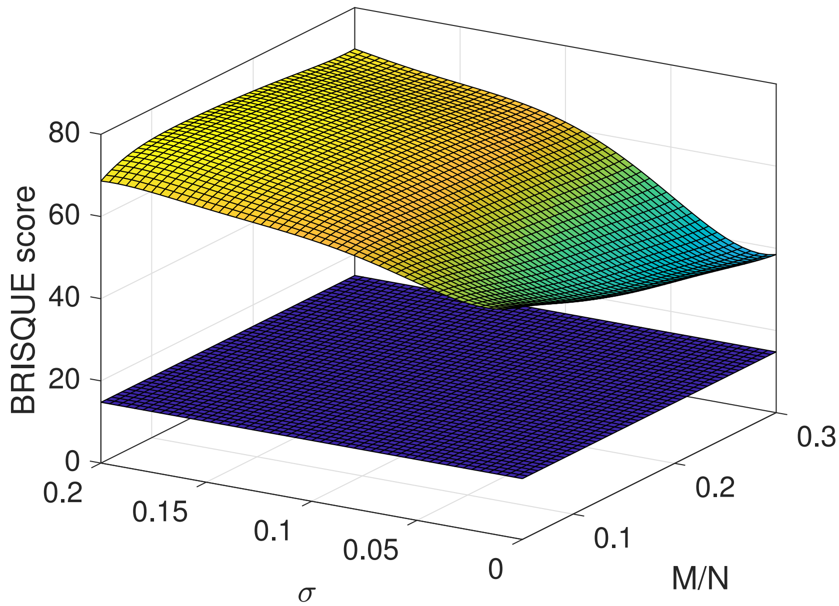


Figure 9: BRISQUE result depending on number of measurements and simulated noise level. Lower surface is reference image score.

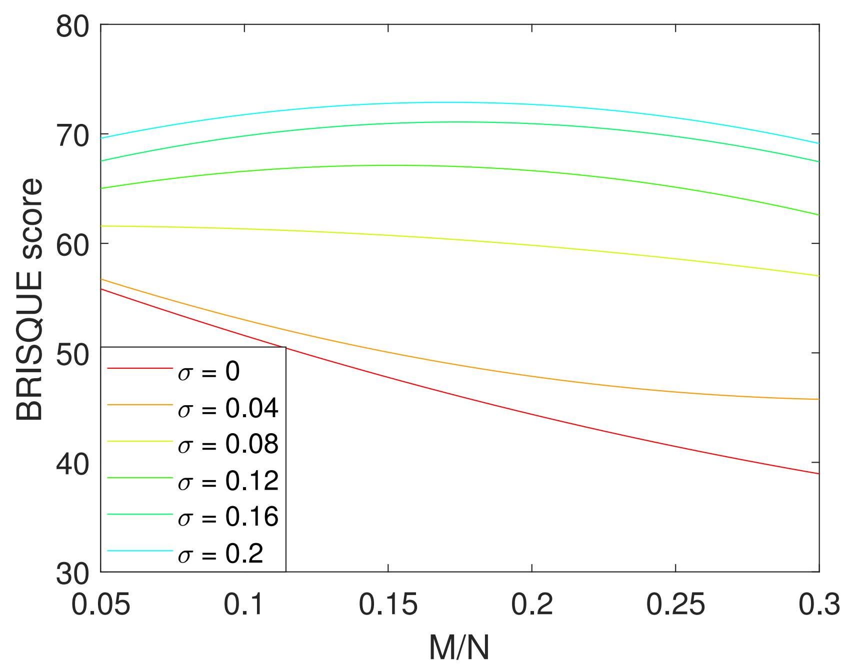


Figure 10: BRISQUE result depending on number of measurements for different simulated noise levels.

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 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

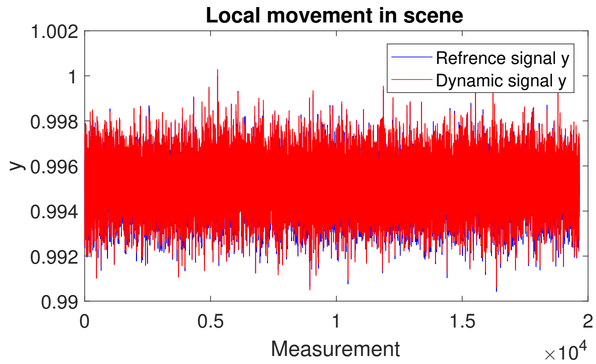
Figure 11: Local movement

The difference between figure 11b and 11c is visible with the naked eye, not only does the object moving around get blurry and noisy but the whole image globally. In table 1...

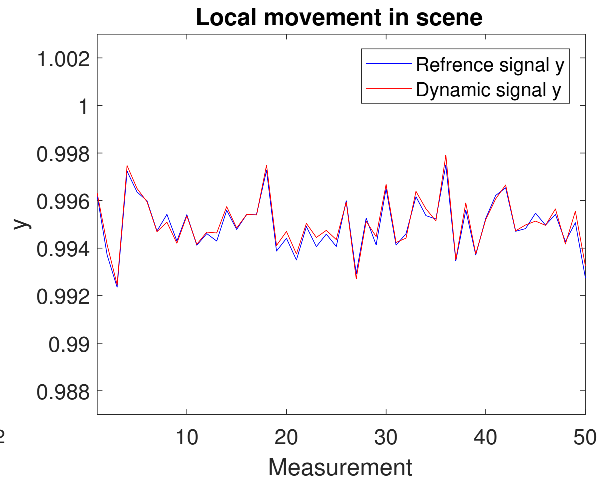
Peak SNR	SNR	SSIM
29	25	91

Table 1: Effects comparing non perturbed reconstructed image against reconstructed image with local movement

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



(a) Signal.



(b) Zoomed in view of the signal.

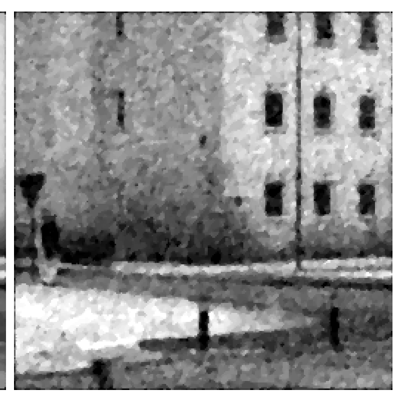
Figure 12: Local movement, acquired signal

As seen in figure 14a there is no obvious difference between the non perturbed reference signal and the distorted signal. In figure 11b 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.

(a) Original reference image

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



(c) Reconstructed 30% image with object passing trough

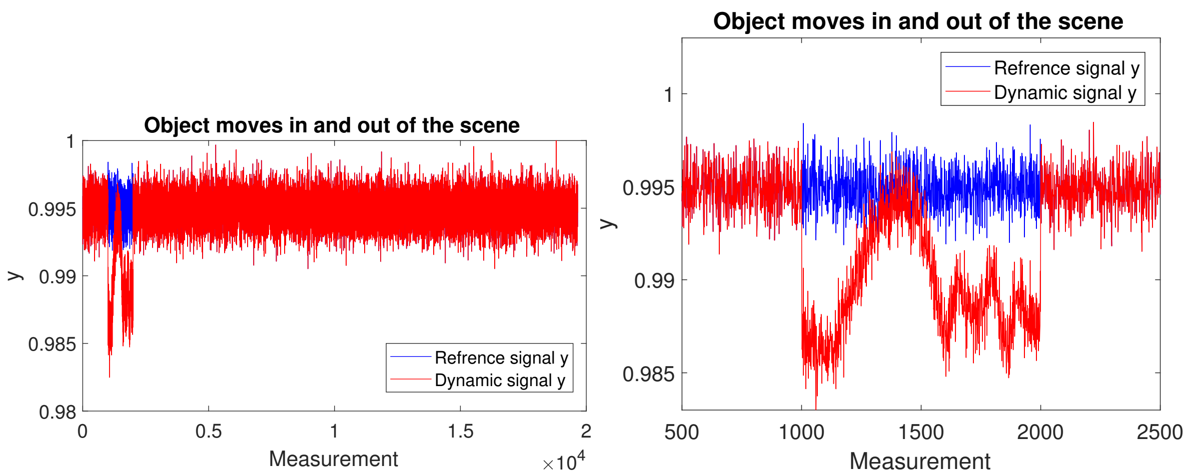
Figure 13: Object passing trough scene.

The difference between figure 13b and 13c is visible with the naked eye, A global noise arises in the image and the object cant be seen. In table 2...

Peak SNR	SNR	SSIM
23	18	58

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

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



(a) Signal.

(b) Zoomed in view of the signal.

Figure 14: Global movement, acquired signal

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

The third scenario i 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.

(a) Original reference image

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

(c) Reconstructed 30% image with global luminance change

(d) Reconstructed 30% image with mean subtraction

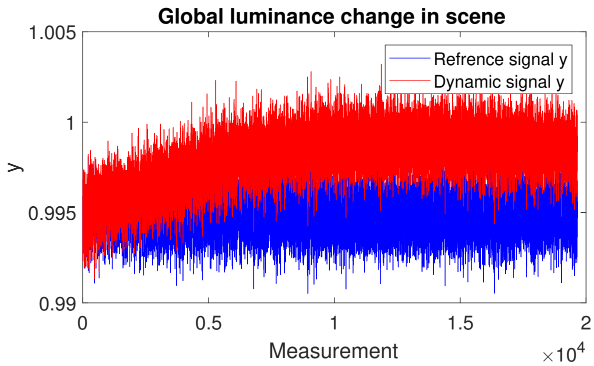
Figure 15: Global luminance change in scene.

The difference between figure 15b and 15c is visible with the naked eye, A global noise arises in the image, but as seen in figure 15d the effect can be suppressed explained under figure 16. In table 2...

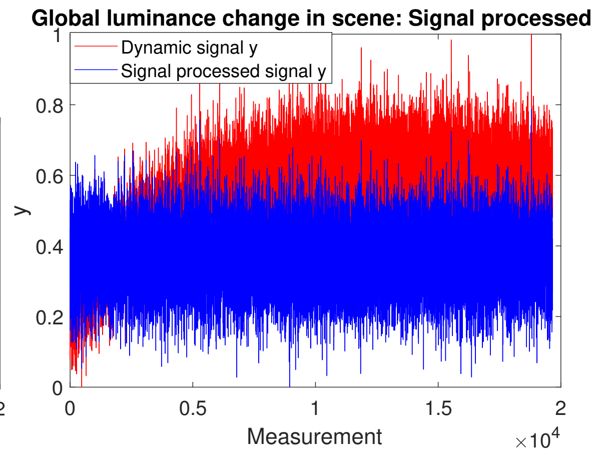
	Peak SNR	SNR	SSIM
Perturbed signal	19	14	38
Mean subtracted signal	33	29	93

Table 3: Effects comparing non perturbed reconstructed image against reconstructed image with global luminance change

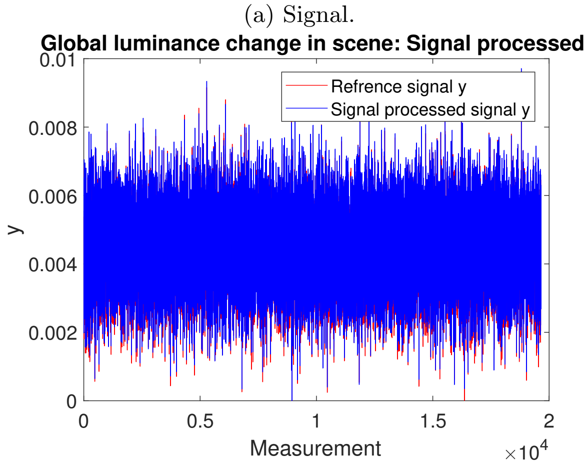
Commenting the result from the table... In figure **??** the effects of global luminance is shown plotted against the non perturbed signal.



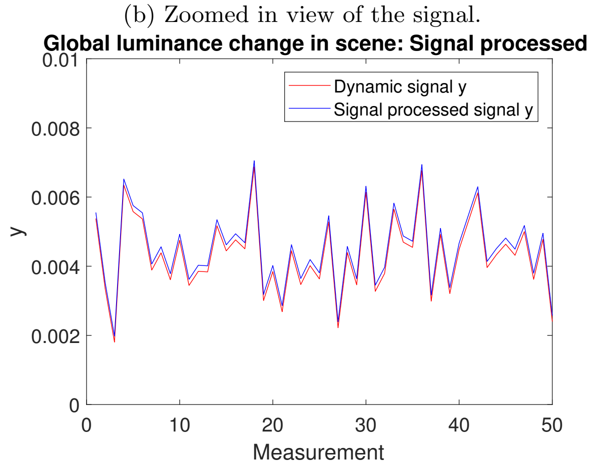
(a) Signal.



(b) Zoomed in view of the signal.



(c) Signal.



(d) Zoomed in view of the signal.

Figure 16: 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 Soft chessboard

Todo: Skapa rekonstruerade bilder från homagraphin och jämnför de rekonstruerade med referens-bilden

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

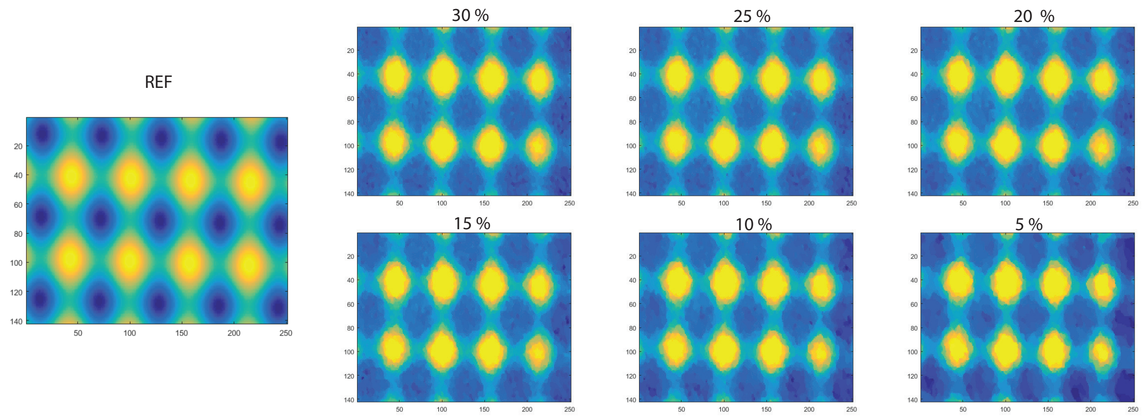
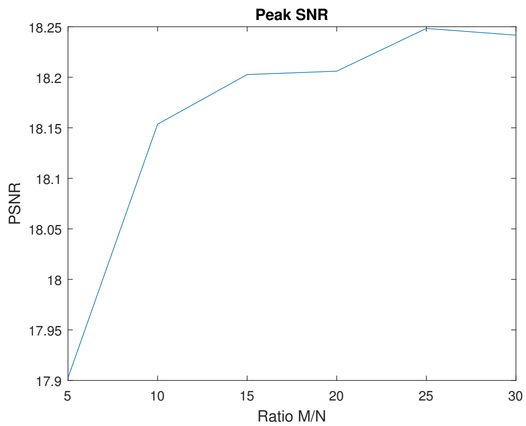
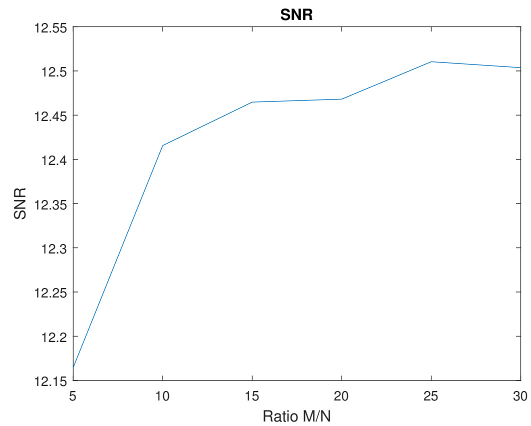


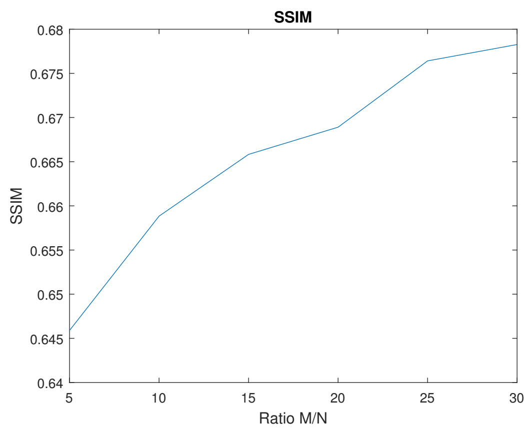
Figure 17: The reconstructed images with different number of measurements and the reference image transformed to fit the SPC images using homography.



(a) Peak SNR for reconstructed images against reference image.



(b) SNR for reconstructed images against reference image.



(c) SSIM score for reconstructed images against reference image.

Figure 18: Signal quality of SPC images compared to reference image

4.2.2 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%.

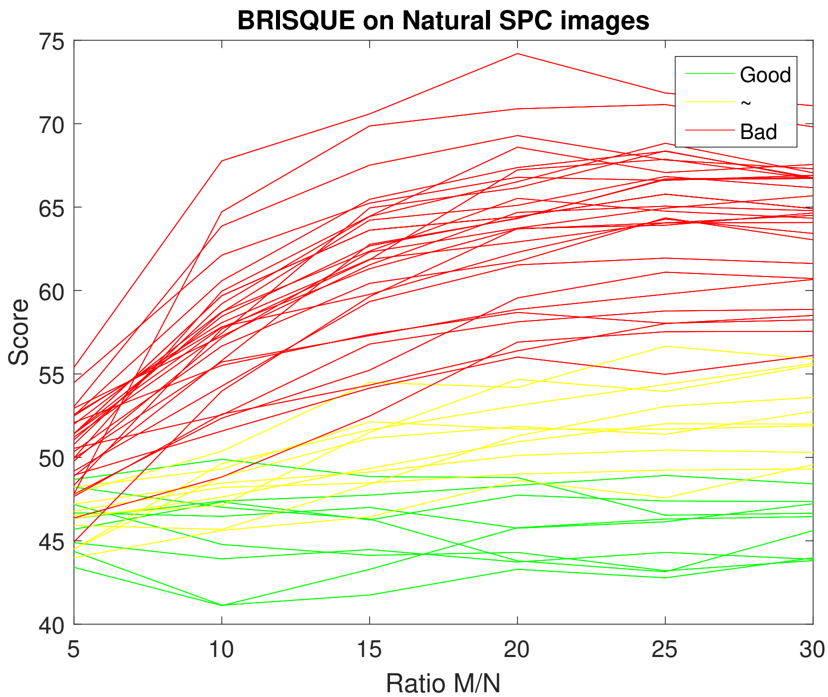


Figure 19: BRISQUE result.

Figure 20: Example of ’good’ images corresponding to the green lines in figure 19.

Figure 21: Example of ’medium good’ images corresponding to the yellow lines in figure 19.

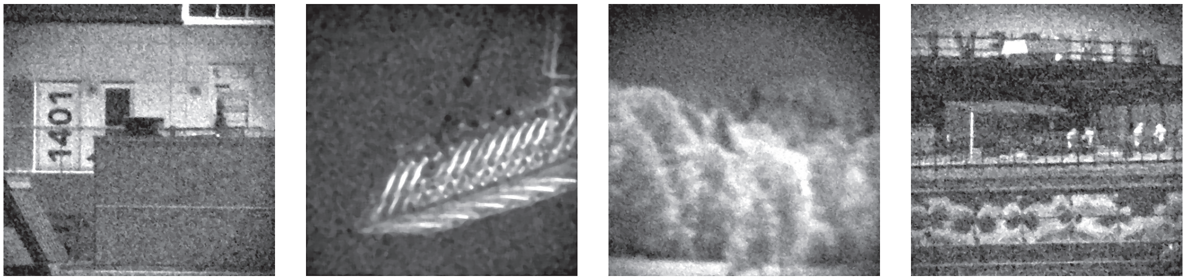


Figure 22: Example of 'bad' images corresponding to the red lines in figure 19.

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

4.2.3 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 23.

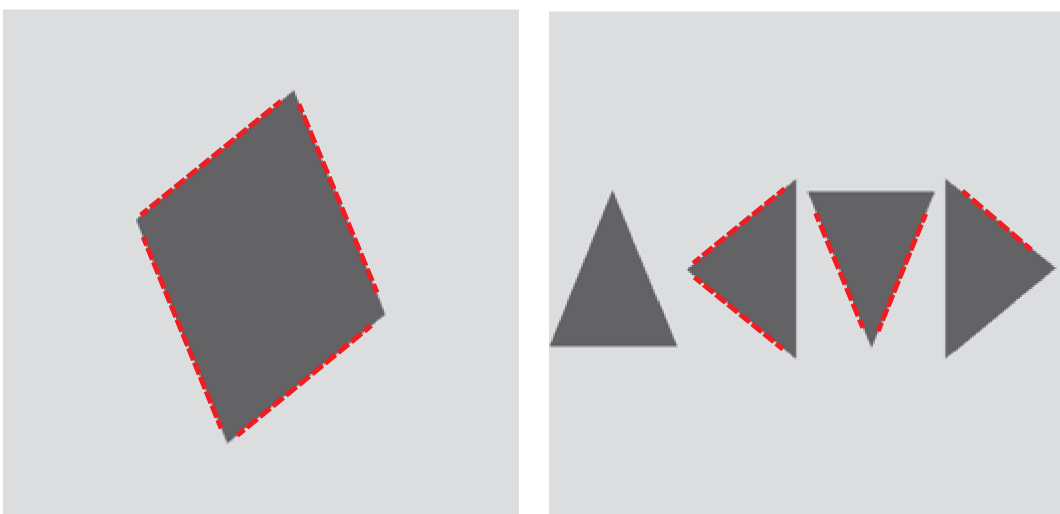


Figure 23: Printed targets with markings where the MTF measurements was performed

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 24 the images from the SWIR camera and SPC are presented.

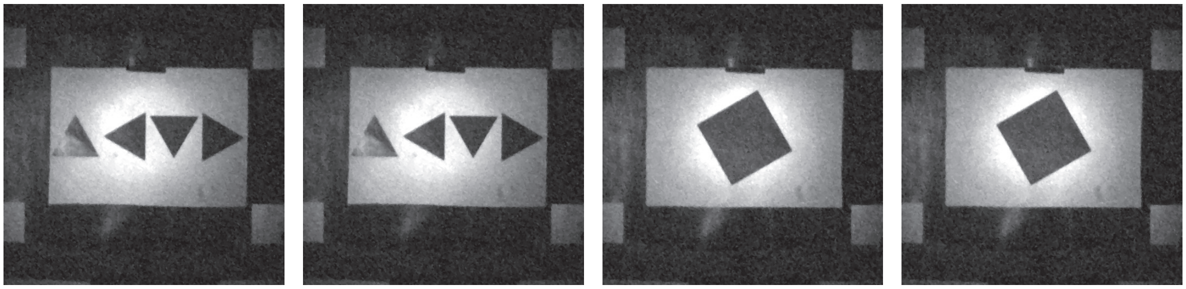
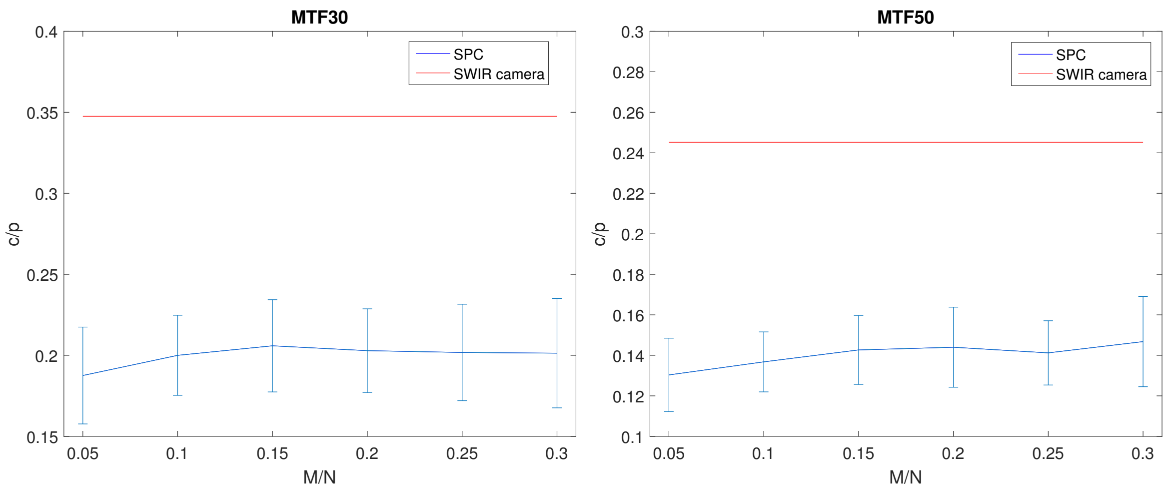


Figure 24: SPC and state of the art SWIR camera output images. (OBS! Bilder från Raptorn ska läggas till)

4.3 MTF



(a) MTF30 result.

(b) MTF50 result.

Figure 25: 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 26 the result from the experiment in presented.

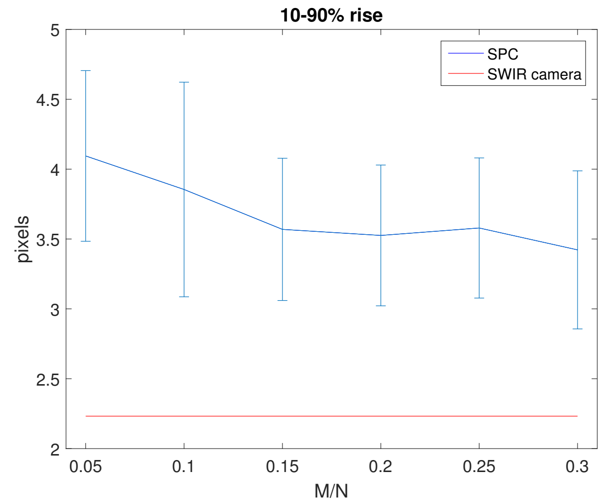


Figure 26: 10-90% rise in pixels. (OBS! inte rätt figur)