Compressed Imaging at Long Range in SWIR

Andreas Brorsson, Carl Brännlund, David Bergström,
and David Gustafsson(✉)

FOI - Swedish Defence Research Agency, C4ISR, Linköping, Sweden
{andbro,carbra,davber,davgus}@foi.se

Abstract. In this paper, we present a single pixel camera operating in the Short Wave InfraRed (SWIR) spectral range that reconstructs high resolution images from an ensemble of compressed measurements. The SWIR spectrum provides significant benefits in many applications due to its night vision characteristics and its ability to penetrate smoke and fog. Walsh-Hadamard matrices are used for generating pseudo-random measurements which speed up the reconstruction and enables reconstruction of high resolution images. Total variation regularization is used for finding a sparse solution in the gradient space. The edge response for the single pixel camera is analysed. A large number of outdoor scenes with varying illumination has been collected using the single pixel sensor. Visual inspection of the reconstructed SWIR images indicates that most scenes and objects can be identified after a sample ratio of 3%. The reconstruction quality improves in general as the sample ratio increases, but the quality is not improved significantly after the sample ratio has reached roughly 10%. Dynamic scenes in the form of global illumination variations can be handled by temporal local average suppression of the measurements.

Keywords: Compressed sensing · Single pixel imaging · SWIR · Total variation

1 Introduction

Conventional infrared camera systems capture the scene by measuring the radiation incident at each of the thousands of pixels in a focal plane array. In compressed sensing a relatively small number of measurements from the scene is combined with sparse reconstruction procedures to recover a full resolution image. Using Compressive Sensing (CS) an image can be acquired with the content of a high resolution image captured with a high resolution focal plane array while using smaller, cheaper and lower bandwidth components. A single pixel camera in Short Wave InfraRed (SWIR), which combines a Digital Micromirror Device (DMD), an InGaAs photodiode and a compressive sensing reconstruction procedure is presented. In many applications the SWIR spectrum provides significant benefits over the visual spectrum. For example SWIR enables separation

between camouflage and vegetation and penetrate fog and smoke to some extent which enable imaging through scattering media. Furthermore, SWIR sensors can be used for passive imaging in moonless condition due to the night-glow of the sky that provides SWIR illumination.

The single pixel camera in SWIR is evaluated on realistic natural scenes with different lighting condition at long distances. Compensation methods for illumination variation in the scene while capturing the scene is proposed, implemented and evaluated.

2   Related Work

"Sparse Modeling" by Rish and Grabarnik [10] and "Sparse and redundant representation" by Elad [5] cover the fundamental principles and constraints that needs to be fulfilled in CS. "An architechture for compressive imaging" [13] and "A New compressive imaging camera architecture using Optical-Domain Compression" [12] by Wakin et al. present the first single pixel camera architecture using CS. "Single-Pixel Imaging via Compressive Sampling" [11] by Duarte et al. contains an introduction to CS, different scanning methodologies and the SPC architecture. "A high resolution SWIR camera via compressed sensing" [8] by McMackin et al., at Inview Technology (which develop and produces compressive sensing cameras), contains a brief review of Compressive Imaging (CI) followed by a presentation of Inview Technology camera architecture.

Chengbo Li:s master's thesis "An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing" [9] describes a total variation algorithm that Li constructed which solves the CS problem. The algorithm is faster and produces better results for images than previous popular algorithms.

The articles [3,6] describes why and how to implement Structurally Random Matrices (SRM). In these articles the Hadamard or DCT matrices are proposed to replace the random measurement matrix. With SRM the reconstruction time is reduced by replacing matrix multiplication with fast transforms. In addition to improving the reconstruction time, the new method does not need to store the measurement matrix in memory, which enables reconstruction of high resolution images. "An Improved Hadamard Measurement Matrix Based on Walsh Code For Compressive Sensing" [14] shows that sequency-ordered Walsh-Hadamard matrices give better reconstruction than the Hadamard matrix with the same benefits of using the Hadamard matrix. The resulting reconstructed image has near optimal reconstruction performance.

Deep learning [7] and Generative Adverserial Networks (GAN:s) [1] have also been used for image reconstruction from an ensemble of single pixel measurements. For specific scenes, Deep learning and GAN:s often generate superior image reconstruction. Unfortunately, those machine learning approaches generalise poorly to unseen scenes.

"Single Pixel SWIR Imaging using Compressed Sensing" [2] by Brännlund and Gustafsson, shows the initial results and proof of concept of the proposed SPC architecture in SWIR.

3   Single Pixel Camera (SPC) Architecture

The SPC platform consists of a Digital Micromirror Device (DMD), acting as a Spatial Light Modulator (SLM), a Newtonian telescope and a single pixel detector. This design has a narrow field of view (22×14 μrad), which gives highly detailed scenes from a great distance. The aluminium Newtonian telescope consists of a concave primary mirror (108 mm, F4.1), and a flat secondary mirror, focusing the scene onto the DMD, see Fig. 1. The motivation to use a Newtonian telescope instead of a lens is partly that chromatic aberration is eliminated and partly that a reflective optical system works over a greater range of wavelengths. A visual spectrum reference camera is also mounted viewing the DMD to simplify setup and focusing of the system (the visual image is acquired in real time). The light on the DMD is reflected either onto the single pixel sensor or the reference camera. The lens mounted on the single pixel sensor is a 50 mm fixed focal length lens (f/1.4, 800–2000 nm spectral range). The single pixel sensor is a large area InGaAs photodiode with a built-in amplifier (Thorlabs PDA20C/M, 800–1700 nm spectral range). The DMD (Texas Instruments DLP4500NIR, 912×1140, 700–2500 nm spectral range) is a matrix of micro mirrors that can be individually tilted $\pm 12°$. The pseudo-random patterns are streamed as a video with a media player, from a computer through a video port (HDMI) to the control unit (DLPR LightCrafterTM 4500). The software (DLP LightCrafter 4500 Control Software) is setup to divide the received 24-bit color image into 1 bit planes which are displayed in consecutive order. The DLP control unit can be operated at a maximum speed of 2880 Hz (24 bit video signal at 120 Hz), but in the current configuration only 1440 Hz was achieved (24bit @60 Hz). At this rate with a subsampling ratio of 10% (number of measurements divided by number of pixels) with 512×512 pixel patterns, all samples are streamed in 17 s. The reconstructed image from the system will have the same resolution as the DMD patterns.

3.1   Signal Processing

The signal from the detector is greatly oversampled by a data acquisition device (NI USB-6210), such that a single value corresponding to each measurement matrix should be obtained. This is performed in two steps; first we omit samples when the DMD is changing pattern, then the mean is calculated and a signal $\mathbf{y}[m]$ that can be processed by the reconstruction algorithm is finally created, as seen in Fig. 2.

3.2   Dynamics in Scene

In natural outdoor scenes the sun is the primary light source. However, even on a clear day the intensity from the sun is not constant. This will reduce the reconstruction performance, because the mean intensity of the measured signal $\mathbf{y}$ is assumed to be stationary. Therefore the luminance change can be modelled as uniform additative noise.
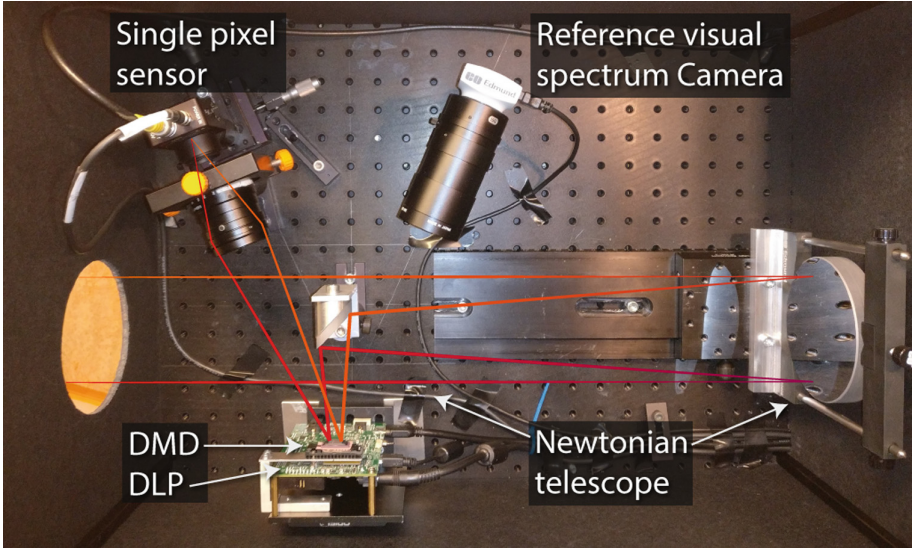
Fig. 1. The single pixel camera architecture used in this work. The optics, DMD, reference camera and the single pixel sensor are shown including red lines showing the light path. (Color figure online)

The measurements y can be decomposed according to

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{c}, \tag{1}$$

where $\mathbf{\Phi}$ is the measurement matrix, $\mathbf{x}$ is the scene viewed as an vectorized image, and c is the uniform intensity change vector. The goal is to remove the uniform intensity change vector c from signal y. c can be approximated by the moving average computed over k measurements and removed from y.
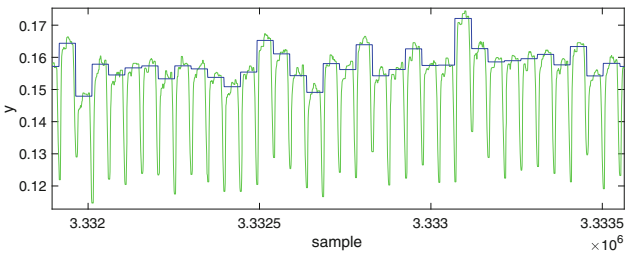

Fig. 2. Calculated mean values (blue) for each measurement matrix with transition measurements (green) omitted. (Color figure online)

4   Sparse Reconstrution Using Compressed Sensing

4.1   Sparse Reconstruction

Compressive Imaging (CI) exploits the fact that natural images are compressible or approximately sparse in some basis and therefore only a few measurements relative to the image pixel resolution needs to be measured in order to reconstruct the image. Two constraints must be fulfilled in order to utilize CS sampling: the image needs to be compressible and the 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 or JPEG2000 (using wavelet transform) and the second constraint is fulfilled using a measurement matrix with a random characteristic.

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

The compressive imaging sampling model is defined as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \epsilon, \tag{2}$$

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

$$\mathbf{\Psi}\theta = \mathbf{x}, \tag{3}$$

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 image $\mathbf{x}$ has K non-zero elements in basis $\mathbf{\Psi}$, $||\theta||_0 = K$. Given (3), (2) can be expanded to

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \epsilon = \mathbf{\Phi}\mathbf{\Psi}\theta + \epsilon = \mathbf{A}\theta + \epsilon, \tag{4}$$

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

The revelation in (4) is what makes CS powerful. By sampling the scene using the measurement matrix $\mathbf{\Phi}$ (as in (2)), but then in the reconstruction process transforming the measurement matrix $\mathbf{\Phi}$ to the reconstruction matrix $\mathbf{A}$ using some basis $\mathbf{\Psi}$, the optimization algorithm can solve the system for the sparse coefficients θ instead of the dense spatial image coefficients in $\mathbf{x}$ [10].

4.2   Permutated Sequency Ordered Walsh-Hadamard Measurement Matrix

Besides from eliminating the need to store the measuring matrix in computer memory for reconstruction, the Permutated Sequency Ordered Walsh-Hadamard

matrix (PSOWHM) can be generated when sent to the DMD and thus eliminating the need to store the matrix. PSOWHM has approximately the same characteristics and properties as an independent and identically distributed (i.i.d.) random matrix but generally has a higher number of measurements for exact reconstruction of the image, $M \sim (KNs)\log^2(N)$, where s is the average number of non-zero indices in the measurement matrix. Research has however shown that there is no significant loss in recovery of the image relative to the i.i.d. random measurement matrix [14]. In order to construct the PSOWHM, the fist 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 is constructed by the recursive formula

$$H_0 = 1, \qquad H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and in general,

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = H_1 \otimes H_{k-1} \tag{5}$$

where $\otimes$ denotes the Kronecker product. The (natural ordered) Walsh-Hadamard matrix is converted to permutated sequence ordered Walsh-Hadamard matrix.

To use the sequence ordered Walsh-Hadamard matrix as a measurement matrix the first row is omitted, permutations to the columns are performed, M rows are choosen at random and the indices with -1 are shifted to 0. The permutations of the matrix is stored and used in the reconstruction [3,9,14].

4.3   Total Variation Regularization

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

$$\hat{\theta} = \arg \min ||\theta||_0 \text{ subject to } \mathbf{y} = \mathbf{A}\theta. \tag{6}$$

Unfortunately, minimizing the non-zero indices in θ in the sparsifying basis $\mathbf{\Psi}$ is NP-hard. A better approach is the ℓ_1 minimization, for example Basis Pursuit denoise (BPDN),

$$\hat{\theta} = \arg \min ||\theta||_1 \text{ subject to } ||\mathbf{y} - \mathbf{A}\theta||_2 < \epsilon. \tag{7}$$

Donoho [4] proved the ℓ_0/ℓ_1 equivalence holds in the CS case, which implies a ℓ_1-minimizer is guaranteed to find the most spare 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 [4,11,12].

The total variation (TV) based TVAL3 (Total Variation Augmented Lagrangian Alternating Direction Algorithm) is used for image reconstruction

[9]. 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 to be sparse instead of the signal itself, thus finding the most sparse gradient. The TV optimization problem in TVAL3 is defined as

$$\min_{\mathbf{x}} \Sigma_i ||D_i\mathbf{x}||, \text{ subject to } \mathbf{\Phi}\mathbf{x} = y, \mathbf{x} \geq 0, \tag{8}$$

where $D_i\mathbf{x}$ is the discrete gradient of $\mathbf{x}$ at position i. TVAL3 is an optimization method for solving constrained problems by substituting the original constrained problem with a series of unconstrained subproblems and introducing a penalty term. To solve the new subproblems the alternating direction method is used [9].

The main reason to use the PSOWH 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 a matrix and a vector that dominates the computation time, $\mathbf{\Phi}\mathbf{x}^k$ and $\mathbf{\Phi}^\top(\mathbf{\Phi}\mathbf{x} - y)$. The multiplication can be replaced with a fast transform. The sequency ordered Walsh-Hadamard matrix is a transform matrix, which also can be computed with the Fast Walsh-Hadamard Transform (FWHT), $\mathbf{W}\mathbf{x} = \text{FWHT}(\mathbf{x})$, where 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 a superposition of Walsh functions.

5   Experiments

The performance of the SPC is analyzed by evaluating a range of methods. All images captured by the SPC of natural scenes were taken at a distance between 200 to 900 m on sunny days with some clouds. Due to the DMD's matrix indexing and diamond shape, the rows of the patterns were repeated before it was sent to the DMD (two mirrors per pixel binned, 1024×512). That indexing made the pattern's ratio on the DMD to correspond to the correct ratio of 1 to 1. In the cases where the resolution was lower than 512×512, the pattern was scaled up so that it filled the maximum area of the DMD.

5.1   Image Reconstruction - Dynamic Scenes

Varying light conditions in the scene during image readout result in poor reconstruction performance, especially during long exposure times. The sensor sums up half the scene's light - 50% of the mirrors are oriented toward the sensor -, which make the intensity change for each sensor readout a global change. In Fig. 3, signals and a reconstructed image (captured in good conditions with strong lighting) with and without the moving mean algorithm are shown. The moving mean is corresponding to light intensity change. The number of neighboring samples to calculate the average was set to $k = 75$, which corresponds to a window of 50 ms. This method increases the image quality significantly for outdoor scenes.
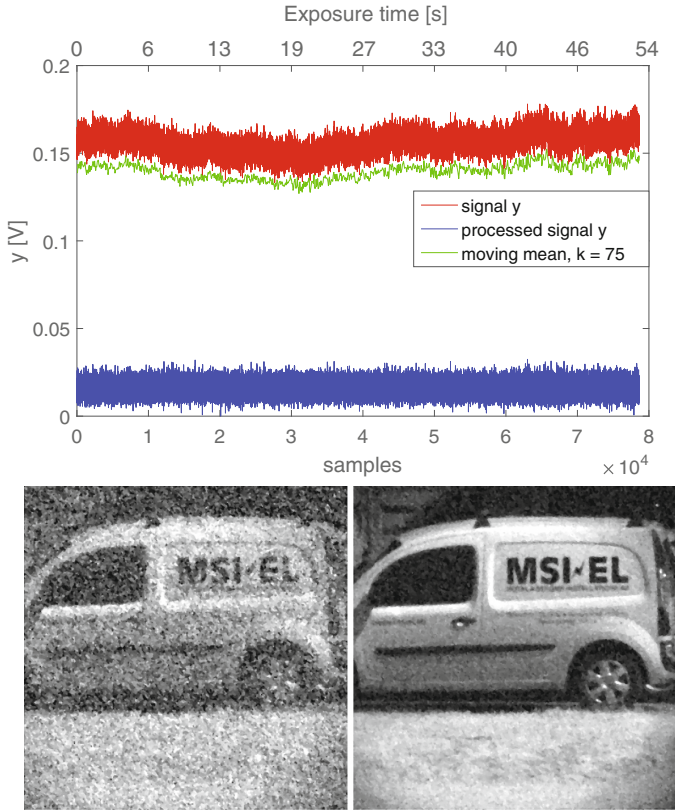
Fig. 3. Top: Sampled signal from SPC with light intensity change distortion and the signal processed by the moving mean algorithm. Bottom: Reconstructed images before (left) and after (right) applying moving mean algorithm on the sampled signal
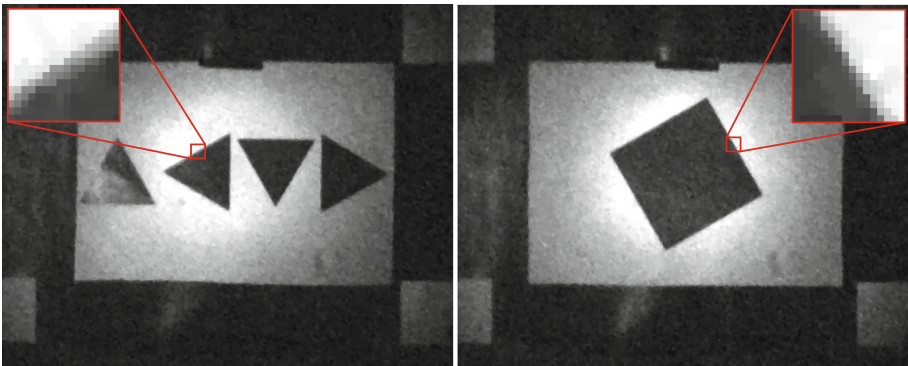

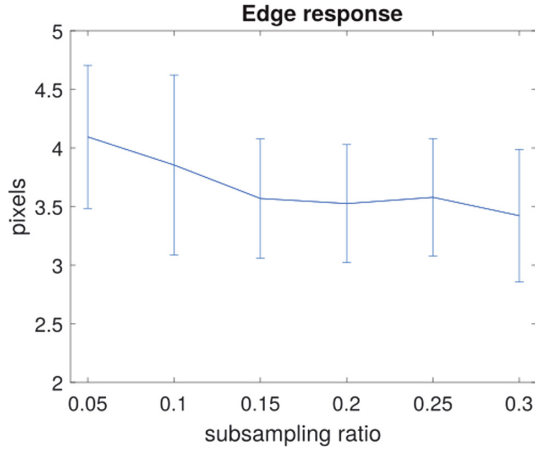**Fig. 4.** Shows images captured with the SPC with a zoomed-in view of a sample edge.

Fig. 5. Edge response, as the distance (pixels) to rise from 10% to 90% in average for an edge.

5.2 Edge Response Anlaysis

The edge response is used to compare and characterize the image quality of camera systems (including lenses) and in this section the edge response for SPC is analysed. Two scenes were captured by the SPC, containing printed sheets of paper with tilted edges lit by a 135 W halogen lamp placed two meters from the sheet.

Fig. 6. Varying minimum subsampling ratios (2%, 3% and 3%) to reconstruct sample images captured by the SPC.

The mean and the standard deviation are estimated for the SPC, for all edges in the images. For the SPC, images reconstructed from 5% to 30% were tested in order to see if the subsampling ratio affected the edge response result. The images are presented in Fig. 4. The edge response is measured as the distance in (pixels) required for an edge to rise from 10% to 90% intensity change. The result from the experiment is presented in Fig. 5.

Fig. 7. Varying subsampling ratios - first row: visual image, second row: 5%, third row: 10% and fourth row 15% - to reconstruct sample images captured by the SPC.
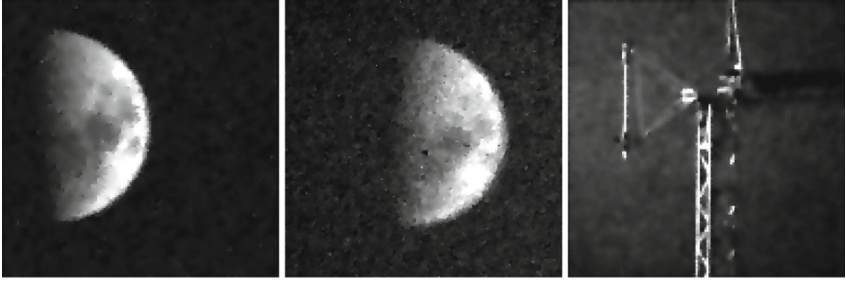
Fig. 8. (left) The moon captured at night at resolution 128×128, subsampling ratio 0.3 and sampling time 1 s, (middle) the moon captured at resolution 256×256, subsampling ratio 0.2 and sampling time 9 s. The noise is due to the motion of the moon during sampling. (right) A construction crane captured at resolution 128×128, subsampling ratio 0.3 and sampling time 1 s.

Some improvement is seen when the subsample ratio is increased, but the standard deviation is almost constant.

5.3   Image Reconstruction

The number of measurements needed to reconstruct an image depends on the sparsity or compressibility of the image. It is hard to give a good estimate of the subsampling ratio needed to obtain a desired quality of the reconstruction. In addition, using an SPC where noise contaminates 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.

Reconstruction of an image using a minimum of sample results in a noisy image. A subjective investigation of the minimum subsampling ratio was performed on the images captured by the SPC, where the quality of each image was studied in 1% subsample ratio increment. A recognizable image was obtained when the ratio was in the range from 2% to 4%.

In the sample images in Fig. 6, the minimum subsampling ratio varied between 2% to 3%. However, as can be seen in the images the motive is merely recognizable, large structures can be identified but details are lost. In general a subsampling ratio of 5% has always succeeded to reconstruct an identifiable image with some fine details.

In Fig. 7 three scenes are reconstructed using different subsample ratios ranging 5%, 10% and 15%. This result can give a perception of which subsampling ratio is good enough for a given purpose. As can be seen, the increase of the subsampling ratio is not linearly proportional to the increase in perceived image quality.

From the images in Fig. 7, it can be observed that already at 5% subsampling ratio the scene can be properly identified. At 10% the finer details appear, like the gap between the panels in the facade. The outline of the text on the car and

door are getting sharper and the structures in the faces start to appear. After 15% the reconstitution quality does not improve much (not shown in the figure).

In Fig. 8 the moon captured at night at resolution 128×128 and 256×256 is captured. The noise in the 256×256 reconstruction is due to motion of the moon during sampling. A construction crane captured at resolution 128×128 is also shown in Fig. 8 .

6   Conclusions

The main attractiveness of compressive imaging, in comparison to conventional imaging using focal plane arrays, lies in the relaxation of the size, weight, power and cost of components for both image collection and transmission. The single pixel camera (SPC) developed and here presented has demonstrated the possibility to capture high quality SWIR images at resolutions up to 512×512 pixels, with about the tenth of the cost a conventional camera with the same resolution. A large number of outdoor scenes with varying illumination and over relatively long distances has been collected. We also capture images of the moon at 128×128 after sunset.

Visual inspection of the reconstructed images show that image quality generally increases up to about 15–20% subsampling ratios, but that often an adequate quality can be achieved already at 5–10% and that the minimum subsampling ratio is about 2–4 % for most of the collected scenes. Reconstruction speed of the SPC and image resolution has been increased by introducing the Walsh-Hadamard pseudo-random matrices presented in the paper, but the collection time for a full resolution 512×512 image at a 10% subsampling ratio still takes about 17 s with the current system setup. Such long collection times naturally introduces sensitivities to scene and light changes, mainly manifested as noise in the reconstructed images, but it has been shown that the latter can be somewhat mitigated by introducing a moving average in the signal processing chain. It is however still a high priority to reduce both collection and reconstruction time, in order to make the system more useful for more dynamic scenes and real-time operation. This will be the focus of future work.

References

1. Bora, A., Jalal, A., Price, E., Dimakis, A.G.: Compressed sensing using generative models. In: International Conference on Machine Learning (ICML) (2017)
2. Brännlund, C., Gustafsson, D.: Single pixel SWIR imaging using compressed sensing. In: Symposium on Image Analysis, 14–15 March 2017, Linköping, Sweden (2017)
3. Do, T.T., Gan, L., Nguyen, N., Tran, T.: Fast and efficient compressive sensing using structurally random matrices. IEEE Trans. Signal Process. **60**(1), 139–154 (2012)
4. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)

5. Elad, M.: Sparse and Redundant Representations. Springer, New York (2010). https://doi.org/10.1007/978-1-4419-7011-4
6. Gan, L., Do, T.T., Tran, T.D.: Fast compressive imaging using scrambled block Hadamard ensemble. In: 16th European Signal Processing Conference (2008)
7. Higham, C.F., Murray-Smith, R., Padgett, M.J., Edgar, M.P.: Deep learning for real-time single-pixel video. Sci. Rep. **8**, 2369 (2018)
8. McMackin, L., Herman, M.A., Chatterjee, B., Weldon, M.: A high-resolution SWIR camera via compressed sensing. In: SPIE, vol. 8353 (2012)
9. Li, C.: An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing. Master's thesis, Rice University (2009)
10. Rish, I., Grabarnik, G.Y.: Sparse Modeling. CRC Press, Boca Raton (2015)
11. Takhar, D., Laska, J.N., Duarte, M.F., Kelly, K.F., Baraniuk, R.G., Davenport, M.A.: Single-pixel imaging via compressive sampling. IEEE Signal Process. Mag. **25**, 83–91 (2008)
12. Takhar, D., et al.: A new compressive imaging camera architecture using optical-domain compression. In: IS&T/SPIE Computational Imaging, vol. 5 (2006)
13. Wakin, M.B., et al.: An architecture for compressive imaging. In: IEEE International Conference on Image Processing (ICIP), 8–11 October 2006, pp. 1273–1276 (2006)
14. Zhuoran, C., Honglin, Z., Min, J., Gang, W., Jingshi, S.: An improved Hadamard measurement matrix based on Walsh code for compressive sensing. In: 9th International Conference on Information, Communications Signal Processing (2013)