FPA-CS: Focal Plane Array-based Compressive Imaging in Short-wave Infrared

Anonymous CVPR submission

Paper ID 1289

Abstract

Imaging beyond the visible spectrum remains challenging and even low-resolution sensors are expensive because the sensor arrays that are sensitive to short-, mid- and long-wave infrared spectra are significantly more expensive than their counterparts for visible imaging. Over the last decade, compressive sensing (CS) has emerged as a potential means to ameliorate this bottleneck. The basic idea is that a sensor array can be replaced with a single detector that acquires coded measurements of a high-resolution image. Unfortunately, the measurement rate of a single detector-based compressive camera is insufficient to enable imaging at high spatial and temporal resolutions.

In this paper, we present a focal plane array-based compressive sensing (FPA-CS) architecture that achieves high spatial and temporal resolutions. The basic idea of FPA-CS is to use an array of single pixel cameras in parallel, thereby increasing the measurement rate and subsequently, the achievable spatio-temporal resolutions. We show that using a low-resolution sensor array, instead of a single-pixel sensor, alleviates the mismatch and results in maximizing the resultant spatio-temporal resolution of the imaging device. We develop reconstruction algorithms based on total variation regularization and present a proof-of-concept prototype in the short-wave infrared (SWIR) domain that can achieve 1 megapixel spatial resolution at 16 Hz.

1. Introduction

The cost of a high-resolution sensor in the visible spectrum has fallen dramatically over the last decade. For example, a cellphone camera module boasting a sensor with several mega-pixels costs little more than a few dollars [1]. This trend is fueled by the fact that Silicon is sensitive to the visible region of the electromagnetic spectrum and hence, the scaling trends and advances made in Silicon-based semiconductor fabrication directly benefit visible imaging technologies. Unfortunately, these scaling trends do not extend to imaging outside the visible spectrum.

Motivation. In many application domains, imaging beyond the visible spectrum provides significant benefits over traditional visible sensors [10]. For example, short-wave infrared (SWIR) penetrates fog and smog; this enables imaging through scattering media. The SWIR night-glow of the sky naturally provides illumination during the night which enables SWIR sensors to passively image even in the dark. SWIR imaging also enables a wide variety of biomedical applications [23]. Yet, SWIR imaging requires sensors made of materials like Indium Gallium Arsenide (InGaAs) or Lead Selenide (PbSe); these materials are orders of magnitude more expensive than Silicon. As a consequence, the cost of a mega-pixel sensor in the SWIR regime is still greater than a hundred thousand dollars. Hence, despite their immense potential, high-resolution SWIR sensors are beyond the reach of engineers and scientists in application domains that could most benefit from it use.

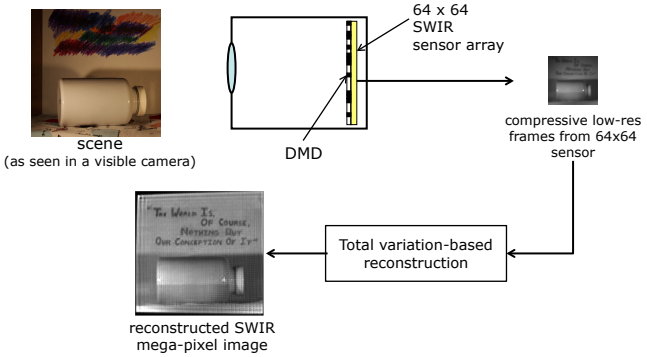

Figure 1. Focal plane array-based compressive sensing (FPA-CS) camera with a 64×64 SWIR sensor array is equivalent to 4096 single pixel cameras (SPCs) operating in parallel. This results in vastly superior spatio-temporal resolutions against what is achievable using the SPC or a traditional camera.

In this paper, we leverage the theory and practice of compressive sensing (CS) [3, 6] , a sub-Nyquist sensing paradigm, to enable high-resolution SWIR imaging from low-resolution sensor arrays. The core theoretical results of CS suggests that a signal can be sensed from a number of measurements significantly smaller than the dimensionality of the signal, provided the signal enjoys a redundant representation such as sparsity in a transform basis or sparse gradients. However, the sensing strategy envisioned by CS relies on the ability to obtain arbitrary linear measurements

of the scene; this requires a fundamental redesign of the architecture used to image the scene. The single-pixel camera (SPC) is an example of such an architecture [7]. The SPC uses a digital micro-mirror device (DMD) as a spatial light modulator and acquires coded measurements of an image onto a single photo-detector. We can build a SWIR SPC simply by employing a photo-detector sensitive to SWIR along with the appropriate choice of optical accessories. The programmable nature of the DMD enables the photo-detector to obtain the sum of any arbitrary subset of pixels which conforms perfectly to the theoretically requirements of CS. The high-resolution image can then be computationally reconstructed from a small number of measurements.

Despite the immense potential of CS, the SPC is incapable of producing high-resolution images at video rate. The measurement rate of the SPC is determined by the speed of operation of its DMD which seldom goes beyond 20kHz. At this measurement rate, conventional Nyquist-based cameras can barely support a 20 fps video at a spatial resolution of 32×32. To obtain a mega-pixel video-rate sensor — a measurement rate of 20MHz — we would need to bridge a gap of $1000\times$ in the measurement rate which is beyond the capabilities of the best CS techniques.

Contributions. The primary goal in this paper is to enable a novel class of CS architectures that achieve high spatial and temporal resolutions using inexpensive low-resolution sensors. The main technical contributions in this paper are:

- We characterize the spatio-temporal resolution limits of CS architectures. A key finding is that a space-bandwidth product mismatch between the DMD and the photo-detector results in sub-optimal performance.
- We propose the focal plane array-based compressive sensing (FPA-CS) camera — an imaging architecture that is optically identical to 1000s of SPCs acquiring compressive measurements in parallel (see Fig. 1). FPA-CS balances the space-bandwidth product constraints, thereby enabling the realization of CS based imaging devices with higher spatial and time resolutions.
- We develop a prototype FPA-CS camera in the SWIR and demonstrate capture of 1 megapixel SWIR images, far exceeding the capabilities of current methods.

2. Related work

Compressive sensing (CS) [3, 6] deals with the estimation of a signal $\mathbf{x} \in \mathbb{R}^{\mathbf{N}}$ from $M < N$ linear measurements $\mathbf{y} \in \mathbb{R}^M$ of the form

$$\mathbf{y} = A\mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{e}$ is the measurement noise and A is the measurement matrix. Estimating the signal $\mathbf{x}$ from the compressive measurements $\mathbf{y}$ is an ill-posed problem since the system of equations is under-determined. Nevertheless, a fundamental result from CS theory states that a robust estimate of the vector $\mathbf{x}$ can be obtained from $M \sim K\log(N/K)$ measurements if (1) the signal $\mathbf{x}$ admits a K-sparse representation; i.e., $\mathbf{x}$ has no more than K non-zero entries), and (2) the *effective* sensing matrix A satisfies the so-called restricted isometry property (RIP) [2]. For example, if the entries of the sensing matrix A are i.i.d. zero-mean Gaussian distributed or if A is formed by randomly selecting rows of "dense" orthonormal matrices, then A is known to satisfy the RIP with high probability. Furthermore, signals with sparse transform-domain coefficients or sparse gradients can be estimated stably from the noisy measurement $\mathbf{y}$ by solving a computationally tractable convex-optimization problem [3, 16]. We next look at the prior art devoted specifically to CS imaging architectures and associated signal models for recovering the signal.

Compressive imaging architectures. In the context of videos, there are two broad classes of CS architectures.

Spatial multiplexing cameras (SMCs) are imaging architectures that take a low-resolution sensor and super-resolve it based on the ideas of CS. In particular, they employ a spatial light modulator (SLM), e.g., a digital micro-mirror device (DMD) or liquid crystal on silicon (LCOS), to optically compute a series linear projections of the scene $\mathbf{x}$; these linear projections determine the rows of the sensing matrix A. Since SMCs are usually built with only a low-resolution sensors, they can operate at wavelengths where corresponding full-frame sensors are too expensive.

A prominent example of SMCs is the single pixel camera (SPC) [7]; its main feature is the ability of acquiring images using only a *single* sensor element (i.e., a single pixel) and by taking significantly fewer multiplexed measurements than the number of pixels of the scene to be recovered. In the SPC, light from the scene is focused onto a programmable DMD, which directs light from only a subset of activated micro-mirrors onto the photodetector. The programmable nature of the DMD enables us to freely direct light from each of the micro-mirror towards the photodetector or away from it. As a consequence, the voltage measured at the photodetector corresponds to an inner product of the image focused on the DMD and the activation pattern of the DMD. By changing the micro-mirror configuration we can obtain multiple linear measurements of the scene to provide the measurements in (1). FPA-CS can be viewed as an array of 4096 SPC's operating in parallel.

There have been many recovery algorithms proposed for video CS using the SPC. This includes the use of 3D wavelets [28], multi-scale wavelet lifting [17], optical flow [20], block-based models [8], sparse frame-to-frame residuals [25, 4], linear dynamical systems [24, 21].

Temporal multiplexing cameras (TMCs) use frame-rate sensors and seek to temporally super-resolve videos. Veer-

	Spatial Resolution	Temporal Resolution	Spatial Upsampling	Cost	Flexibility (spatial vs time res.)	Hardware required
Nyquist sampling	how	high	no	$$$$$	limited	high res. sensor array
Single Pixel Camera [7]	high	low	yes	$$	yes	DMD + photo-detector
CS-MUVI [20]	medium	high	yes	$$	yes	DMD + photo-detector
TMCs	high	high	no	inapplicable to SWIR	limited	DMD + high res. sensor array
Hyperspectral CS [27]	low	low	no	inapplicable to SWIR	limited	DMD + sensor array + grating
FPA-CS (this paper)	high	high	yes	$$	yes	DMD + low res. sensor array

Figure 2. Comparisons of existing CS-based SWIR cameras.

araghavan et al. [26] showed that periodic scenes could be imaged at high temporal resolutions by using glabal temporal coding. Per-pixel temporal modulation to recover higher frame-rates was demonstrated using prototypes that used LCOS for modulation [9][19]. Hitomi et al. [12] also used a per-pixel coding, that is implementable in modern CMOS sensors with per-pixel electronic shutters; however, a hallmark of their approach is the use of a highly over-complete dictionary of video patches to recovery the video at high frame rates. Llull et al. [15] propose a TMC that uses a translating mask in the sensor plane to achieve temporal multiplexing. This approach avoids the hardware complexity involved with DMDs and LCOS, and enjoys other benefits including low operational power consumption. Similarly, a low-cost compressive device for sensing color and depth was demonstrated in [29][14]. Harmany et al. [11] extend coded aperture systems by incorporating a flutter shutter [18] or a coded exposure; the resuling TMC provides immense flexibility in the choice of measurement matrix. While there has been many TMCs proposed for CS, all of them employ a high spatial resolution sensor and seek to increase the temporal resolution of the device; this makes them inapplicable for SWIR imaging where high resolution sensor arrays are prohibitively expensive. The interested reader is referred to related work in hyper-spectral CS [27].

3. Spatio-temporal resolution (STR)

3.1. STR for a Nyquist camera

Traditional cameras rely on the principle of Nyquist sampling. The STR of the camera is limited by the product of the number of pixels and the maximum frame-rate. For example, a 1 megapixel sensor at 30fps corresponds to a measurement rate (M_r) of $30 * 10^6$ pixels/sec. Thus, for a Nyquist camera, STR $= M_r$. Visible sensors have improved significantly and it is now common to obtain sensors that can achieve mega-pixel resolution at 30 fps. Unfortunately sensors outside the visible spectrum provide either a much lower spatial resolution or are expensive.

3.2. STR for compressive cameras

If we assume that CS-based reconstruction can provide a compression factor of α ($\alpha \geq 1$), then the effective STR

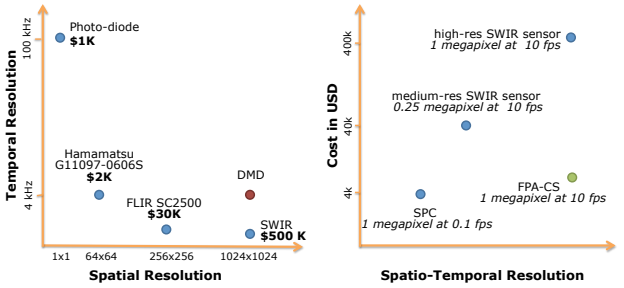

Figure 3. (left) Spatial and temporal resolutions for sensors and DMDs. The requirement of synchronization between DMD and the sensor in a CS based camera necessitates space-bandwidth product matching in order to maximize the spatio-temporal resolution of resulting imaging system. (right) Cost and resolution comparisons on existing SWIR cameras.

can be given by STR $= \alpha M_r$. In the SPC, the measurement rate is typically limited by the maximum rate at which synchronization can be achieved between the modulator and sensor. While a photo-detector can be operated at very high rates (even GHz), commercially-available DMD seldom operate faster than 10 kHz to 20 kHz. Hence, it is not possible to achieve synchronization between any of the current high resolution spatial light modulators and a photo-detector at greater than $F_{max} = 20$ kHz. This directly imposes a limit on the STR of compressive cameras based on single pixel sensors, i.e., STR $\leq \alpha M_r = \alpha F_{max}$ pixels/sec.

3.3. Maximizing the measurement rate

From the previous discussion, it is clear that in order to maximize the spatio-temporal resolution of CS-based imaging devices, one must maximize the measurement rate. Given that the DMD control circuits impose strict limits on the maximum synchronization rate, F_{max}, the only other way to increase the measurement rate is to read multiple measurements in parallel. As an example, suppose one were to design a compressive imaging system in which a $K \times K$ pixel image sensor array were being used to acquired multiplexed measurements in synchronization with a high-resolution DMD at the max operational rate of F_{max}, then such a system would result in a measurement rate of $M_r = K^2 F_{max}$ pixels/sec and consequently a spatio-temporal resolution rate of STR $= \alpha K^2 F_{max}$ pixels/sec, resulting in a spatio-temporal resolution increase of K^2.

In the SWIR prototype presented in this paper, we use a DMD modulator with a $F_{max} = 1$ kHz and a SWIR sensor array consisting of 64×64 pixels, resulting in a measurement rate of about $M_r \approx 4 \times 10^6$ pixels/sec. Further, in our experimental results, we demonstrate high-fidelity reconstructions at a compression rate of $\alpha = 4$; this enables an effective STR $= 16 \times 10^6$ pixels/sec. Since our DMD is approximately 1 mega-pixel, this corresponds to a 1 mega-pixel camera operating at approximately 16 frames/sec. Even if we conservatively do not attempt CS,

i.e., $\alpha = 1$, this still results in a 1 mega-pixel camera operating at 4 fps. In comparison, a SPC using our DMD would be only capable of operating at 1 mega-pixel at 0.01 fps, if $\alpha = 1$, and at $0.04fps$, if $\alpha = 4$. Alternatively, if we choose to use normal high resolution SWIR image sensors instead of compressive imaging devices, then an equivalent image sensor that can operate at 1 mega-pixel spatial resolution and 10 fps would cost upwards of $\$\ 500k$. In contrast, the 64×64 sensor array we use costs about \$2k, while the DMD costs about \$2k resulting in a total cost of under \$4k for our FPA-CS prototype.

4. Specifics of the FPA-CS prototype

The imaging architecture of FPA-CS resembles the imaging architecture of many focal plane array-based TMCs [19, 12]. However, in these TMCs, the spatial resolution of the modulator and the spatial resolution of the sensor are *identical*; in contrast, in our FPA-CS imaging architecture, the modulator (DMD) has significantly higher resolution than that of the sensor. It is also related to the SPC and can be viewed as an array of 64×64 single pixel cameras operating in parallel.

In our short-wave infra-red (SWIR) prototype of FPA-CS, we utilize a Texas Instruments Light Crafter DMD as the light modulator. The DMD consists of 1140×912 micro-mirrors, each of size approximately 7.6 micron. Each of the mirrors can be independently controlled at a rate of 4kHz, and each mirror can be rotated to either by $+12^\circ$ or -12° about the optical axis. We use a customized SWIR objective lens (Edmund Optics $\#83-165$: 50 mm fixed focal length SWIR lens) to focus the scene onto the DMD. A 150mm - 150mm relay lens pair (2x Thorlabs AC254-150C 150mm Plano-Convex Lens) was placed after the SWIR objective lens to extend the original flange distance, thereby providing ample space for the light bundle coming out of the DMD. We also use a field lens (Thorlabs LB1417-50mm double convex lens) to reduce vignetting. The light incident on the DMD corresponding to pixels that are oriented at -12°, is discarded, while the light that is reflected from pixels that are oriented at $+12^\circ$, is then focused on to the SWIR sensor using the re-imaging lens. We use a 100 mm focal length and a 45 mm focal length pair (Thorlabs AC-254-100C and AC-254-045C Plano-Convex Lens) as our re-imaging lens, to provide a $1 : 2.22$ magnification, since the physical size of the sensor and that of the DMD are different. The relay lens, the field lens and the re-imaging lens were all anti-reflection coated for $1050 - 1620$ nm. We use a 64×64 SWIR sensor (Hamamatsu G11097-0606S), with 50 micron pixel size; the sensor costs about \$ 2k. The relay lens is configured to ensure that 16×16 pixels on the DMD map to each sensor pixel. Also, note that because of the 12° rotation, the DMD plane and the sensor plane are not parallel. Therefore, the relay lens and the sensor position are adjusted carefully to satisfy Scheimpflug principle, so that the entire scene plane remains in focus on the sensor. Unfortunately, the DMD is rectangular while the sensor is square and therefore, some of the DMD pixels are wasted. In practice this means that approximately only 912×912 pixels on the DMD are being used to capture our data. In our reconstructions, we spatially super-resolve each sensor pixel by a factor of 16×16. Figure 4 shows a schematic of our design and a photograph of our prototype SWIR FPA-CS system.

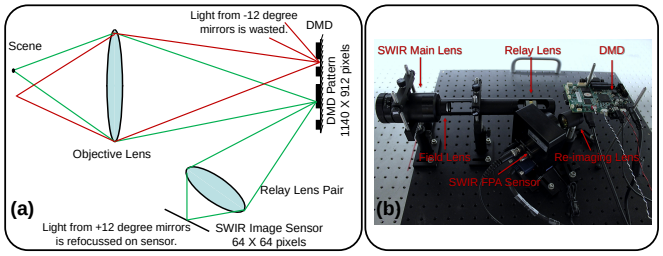

Figure 4. (left) Schematic Illustration of the optical system. (right) Photograph of our prototype.

5. Reconstruction algorithms

5.1. Forward imaging model of FPA-CS

The FPA-CS imaging architecture is equivalent to an array of 64×64 SPCs running in parallel. Behind each of these SPCs is a 16×16 array of DMD mirrors ands at each time instant, a spatial pattern is projected on those 16×16 mirrors. The measurement obtained at pixel (s_i, s_j) at time t, given by $y^t_{i,j}$, is the dot product between the spatial code $a^t_{i,j}$ projected on the mirrors behind the pixel at that time t, and the unknown scene irradiance $x_{i,j}$. If we collect all the measurements at that pixel together, then we can represent this using a vector matrix, notation as $y_{i,j} = A_{i,j}x_{i,j}$, where $x_{i,j}$ represents the unknown 256×1 vector and $y_{i,j}$ is a $T \times 1$ vector that contains the T measurements obtained at that sensor pixel. Thus, the relationship between the measurements at each sensor pixel and the corresponding unknown scene pixels is given by a multiplexing matrix $A_{i,j}$, governed by the patterns projected on the DMD. If we stack all the measurements at all pixels to create a vector $y = [y_{1,1}; y_{1,2}; ..., y_{i,j}, ...y_{N,N}]$, and similarly stack up all the unknown scene irradiances to create a vector $x = [x_{1,1}; x_{1,2}; ..., x_{i,j}, ...x_{N,N}]$, then, we obtain (1) in which the measurement matrix A is a block diagonal multiplexing matrix given by,

$$A = \begin{bmatrix} A_{1,1} & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 \\ 0 & 0 & .. & 0 & 0 \\ 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & A_{N,N} \end{bmatrix}.$$

Our goal is to reconstruct $\mathbf{x}$ from the noisy and possibly under-determined set of linear equations above.

5.2. Total Variation-based reconstruction

Over the last decade, sparse representations and compressive sensing has emerged as a popular tool to solve under-determined set of linear equations. Since the measurement equations are possibly under-determined and certainly noisy, spatial regularization should be incorporated in the reconstruction.

Natural images have been shown to have very sparse gradients. For such signals, one can solve an optimization problem of the form [5, 16]

$$\text{(TV)} \quad \widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \text{TV}(\mathbf{x}) \quad \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon,$$

where the term $\text{TV}(\mathbf{x})$ refers to the total-variation of $\mathbf{x}$. In the context of images where $\mathbf{x}$ denotes a 2D signal, the operator $\text{TV}(\mathbf{x})$ can be defined as

$$\text{TV}_{\text{iso}}(\mathbf{x}) = \sum_{\mathbf{i}} \sqrt{(\mathbf{D_x x(i)})^2 + (\mathbf{D_y x(i)})^2},$$

where $D_x\mathbf{x}$ and $D_y\mathbf{x}$ are the spatial gradients in x- and y-direction of the 2-dimensional image $\mathbf{x}$, respectively. This definition can easily be extended to extended to higher-dimensional signals videos where the 3rd dimension is time. In all our experimental results, we use TVAL3 [13] to solve (TV) and the estimate of ϵ is obtained from the measurement noise level.

6. Experiments

To demonstrate the fidelity of our device, we show results on several real scenes captured using our prototype SWIR FPA-CS camera.

Resolution Chart. To study the spatial resolution characteristics of our device and demonstrate the increase in spatial resolution, we first captured images of a standard Air Force test target using our prototype device. During each measurement, the DMD projected a random binary spatial pattern. The exposure duration for each of the acquired images was less than $1ms$. Each sensor image was of 64×64 pixels, and we acquired a total of 512 images with varying patterns on the DMD. We then reconstructed the scene at the resolution of the DMD using the reconstruction algorithm described in Section 5. To study the impact of the number of measurements used, we performed reconstructions using $64, 128, 256$ and 512 measurements. These correspond to compression rates $\alpha = 4, 2, 1,$ and 0.5, respectively. Figure 5, shows the results obtained at various compression rates. In addition, to demonstrate the improvement in spatial resolution over that of the 64×64 sensor images, we also show a bicubic interpolated version of the sensor image with full-field illumination. (A single sensor image corresponds to a compression rate of approximately 256 over the DMD resolution.) To evaluate the best-case performance of our prototype, limited only by the imaging optics, we also show a reconstruction that used approximately 3000 images that independently scan each of the mirrors in the DMD without any multiplexing. Such a scan would correspond to slow frame-rate of operation but demonstrates the imaging fidelity limited only by the imaging optics. Our results demonstrate that high resolution reconstructions can be obtained from a small number of multiplexed images.

Compression and measurement rates. One important fact to note is that while devices such as the SPC can typically achieve a higher compression rate, FPA-CS only achieves smaller compression rates. This is because, the global nature of the multiplexing in the SPC allows for greater reductions in measurements compared to the block-diagonal structure of the multiplexing matrix in FPA-CS. Although the raw compression rate for FPA-CS is slightly lower, the three order of magnitude improvement in measurement rate achieved due to 4096 parallel measurements, compensates for the small reduction in compression rate. This is because both FPA-CS and SPC require synchronization between DMD and sensor and this is restricted to be at a rate no greater than F_{max}. As a consequence, the practical temporal resolution that can be achieved by FPA-CS far exceeds that of an SPC. In particular, SPC can only match the temporal resolution of FPA-CS if it can achieve a compression rate of 4000, which is impractical. Existing SPC systems typically achieve only compression rates of the order of $\alpha \approx 10 - 20$, resulting in FPA-CS achieving a temporal resolution that is at-least an order of magnitude better.

Modulation transfer function (MTF). We study the spatial frequency response of our prototype FPA-CS system via its MTF. Figure 6 shows the MTF of our system at varying compression rates. First, note the dramatic improvement in MTF of multiplexed reconstruction (i.e., $\alpha = 0.5, 1, 2, 4$) compared to the raw sensor MTF (i.e., $\alpha = 256$). Secondly, note the gradual improvement in MTF with additional measurements, allowing the possibility for adaptive systems where the number of measurements used is adaptively chosen based on the estimated motion.

Applications of SWIR Imaging. There are a large number of applications that are difficult or impossible to perform using the visible spectrum, but become much simpler due to the characteristics of the SWIR spectrum. SWIR imaging is currently used in a host of applications including automotive, electronic board inspection, solar cell inspection, produce inspection, identification and sorting, surveillance, anti-counterfeiting, process quality control, and much more. Some unique properties of SWIR that enable many of these applications include: (a) SWIR penetrates through fog and haze allowing long-distance imaging in haze and fog, (b) SWIR penetrates deeper into tissue enabling many biomedical applications, (c) many packaging materials are transparent in SWIR allowing liquid-level measurements from the

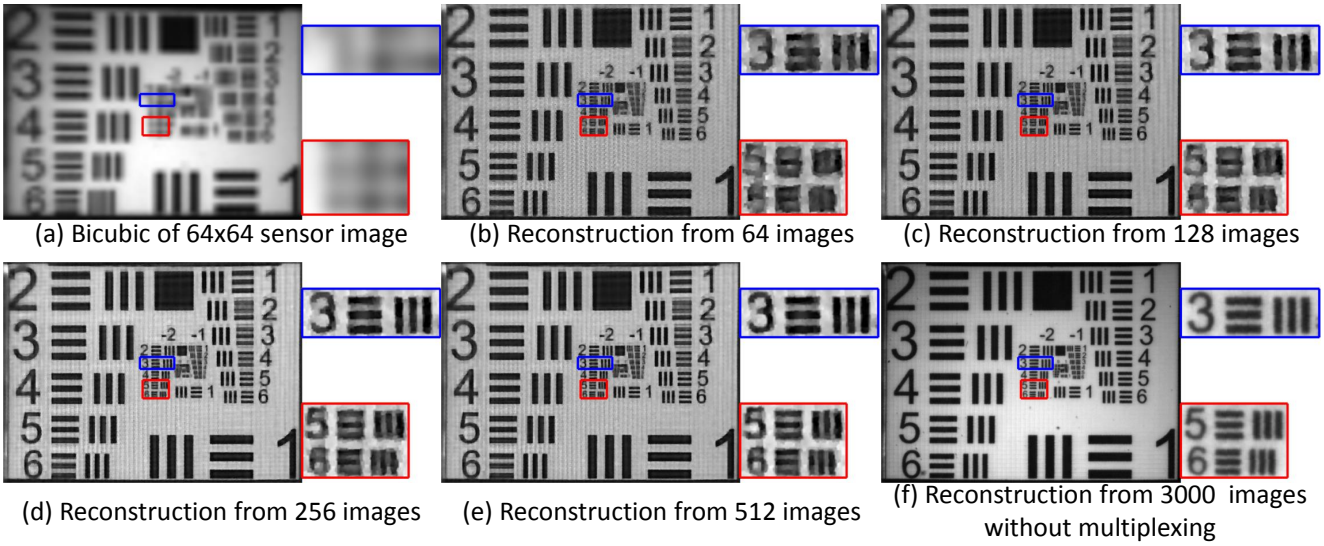

Figure 5. Real results on images of a Air Force resolution chart using our SWIR FPA-CS prototype. Comparison with the captured sensor image show dramatic improvement in resolution. Reconstructions also show gradual improvement with increasing number of measurements being used in the reconstruction.

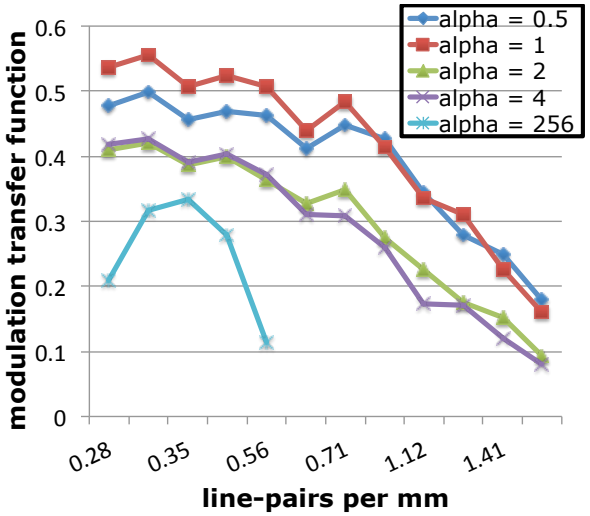

Figure 6. Modulation Transfer Function: The MTF of all imaging system at various compression rates α demonstrates the gradual degradation of higher spatial frequencies as the compression rate is increased.

outside, (d) SWIR can be used to detect quality of fruits and (e) the night-glow of the sky provides sufficient illumination even on a moon-less night to perform long-distance SWIR imaging without the need for extra illumination sources that could compromise reconnaissance.

While the variety of SWIR aplications are almost endless, we present here a simple experiment highlighting two attributes of SWIR. Consider the scene in Figure 7, which consists of an opaque bottle and a crayon scribble in the background. While the visible image in un-remarkable, the corresponding reconstrcuted SWIR image shows two interesting differences. First, note that the crayon is transparent in SWIR, which allows us to read the text behind the scribbles. Second, note that the plastic bottle is opaque in visible, and the visible image contains no information about the contents of the bottle. But the bottle is transparent in SWIR, which allows us to see through the bottle and determine the water level inside the bottle. Figure 7 also shows 1D plots of the mean intensity of columns inside the highlighted color boxes; SWIR intensity changes because of the water inside the bottle, but the visible light intensity remains unchanged. Finally, to evaluate the best-case performance of our prototype, limited only by the imaging optics, we also show a reconstruction that used approximately 3000 images that independently scan each of the mirrors in the DMD without any multiplexing.

FPA-CS video. Figure 8 shows two different frames taken at two different instants of a human hand moving within the field-of-view. SWIR imaging has immense potential for biomedical applications since the penetration depth of SWIR is larger than that of visible light.

7. Discussions

Choice of modulation masks. There are two requirements on modulation masks to obtain high fidelity reconstruction. First, the measurement matrix achieved within each of the 4096 parallel single-pixel cameras should be well-conditioned so that the noise amplification of linear inversion remains low. Second, the spatial code should have high light-throughput, to maximize the signal-to-noise ratio in sensor images. In this paper, we tested two mask patterns — Hadamard and random binary — both of which satisfy

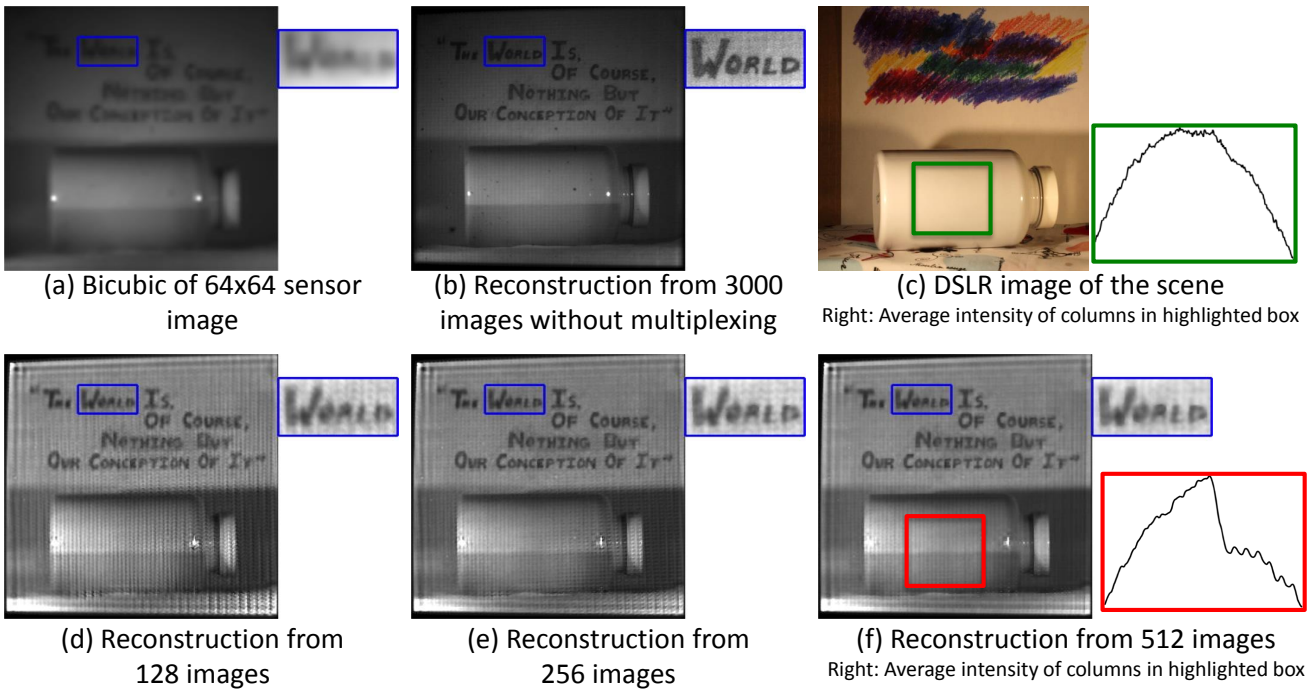

Figure 7. **The invisible scene.** Real results on images of an opaque bottle in the foreground and a crayon scribble in the background. Notice that the crayons are transparent in SWIR allowing us to read the text behind the crayon scribbles. The bottle in the foreground is opaque in visible but transparent in SWIR. Highlighted boxes display mean intensity of columns in the corresponding regions of the visible and SWIR images, showing that one can estimate the water level inside the bottle from the SWIR image, but not from the visible image.

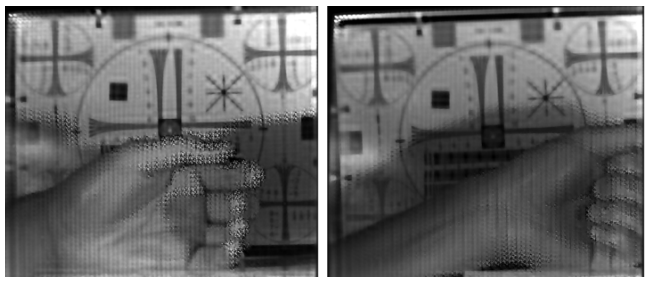

Figure 8. Two images of a moving hand. Each image is reconstructed using 236 random, binary patterns.

these characteristics. Hadamard matrices are known to be optimal in terms of linear multiplexing [22]. The results shown in Figure 7 correspond to Hadamard measurements. We also used random binary patterns since they are known to satisfy the restricted isometry [2], and therefore lead to robust inversion when used along with sparse regularizers such as total variation. The results shown in Figure 5 correspond to binary random measurements. Notice that in both cases 50% of the light reaches the sensor after modulation. In practice, we observed that the reconstructions obtained from the two modulation patterns were near-identical.

Benefits. FPA-CS provides three advantages over conventional imaging. First, since high-resolution image sensors are very expensive in the SWIR regime, our CS inspired FPA-CS system provides an inexpensive alternative to achieve high-resolution SWIR imaging. Second, compared to traditional single pixel-based compressive cameras, FPA-CS simultaneously records data from 4096 parallel, compressive systems, thereby significantly improving the measurement rate of compressive imaging systems. As a consequence, the achieved spatio-temporal resolution of our device is an order of magnitude better than the single-pixel camera.

Limitations. FPA-CS exploits spatio-temporal redundancy in the reconstruction and so, extremely complex scenes such as a bursting balloon cannot be directly handled by the camera. Since the spatio-temporal redundancy exploited by traditional compression algorithms and our imaging architecture are very similar, one can assume that scenes that can be compressed efficiently, can be captured well using our method. Our prototype uses a binary per-pixel shutter and this causes a 50% reduction in light throughput, while half of the light is wasted. In future, a separate, synchronized 64×64 image sensor can be used in the other arm, thereby doubling the measurement rate and further increasing the spatio-temporal resolution that can be achieved. The algorithm is currently not real-time and this precludes the direct-view capability.

Conclusion. We presented focal plane array-compressive sensing (FPA-CS), a new imaging architecture for massively parallel compressive measurement acquisition and demonstrated optical super-resolution for imaging in the SWIR. The architecture proposed here is generic and can be adapted to other spectral regimes such as mid-wave infrared and thermal imaging, where, much like SWIR, sensors are prohibitively expensive.

References

[1] http://www.amazon.com/isight-camera-module-iphone-replacement/dp/b009z5wp1i.

[2] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes rendus-Mathématique*, 346(9-10):589–592, 2008.

[3] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, Feb. 2006.

[4] V. Cevher, A. C. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa. Compressive sensing for background subtraction. In *Euro. Conf. Comp. Vision*, Marseille, France, Oct. 2008.

[5] A. Chambolle. An algorithm for total variation minimization and applications. *J. Mathematical Imaging and Vision*, 20(1-2):89–97, 2004.

[6] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, Apr. 2006.

[7] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.*, 25(2):83–91, Mar. 2008.

[8] J. E. Fowler, S. Mun, E. W. Tramel, M. R. Gupta, Y. Chen, T. Wiegand, and H. Schwarz. Block-based compressed sensing of images and video. *Foundations and Trends in Signal Processing*, 4(4):297–416, 2010.

[9] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. Narasimhan. Flexible voxels for motion-aware videography. In *Euro. Conf. Comp. Vision*, Crete, Greece, Sep. 2010.

[10] M. P. Hansen and D. S. Malchow. Overview of SWIR detectors, cameras, and applications. In *SPIE Defense and Security Symposium*, pages 69390I–69390I, 2008.

[11] Z. T. Harmany, R. F. Marcia, and R. M. Willett. Compressive coded aperture keyed exposure imaging with optical flow reconstruction. *arXiv preprint arXiv:1306.6281*, 2013.

[12] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar. Video from a single coded exposure photograph using a learned over-complete dictionary. In *IEEE Intl. Conf. Comp. Vision*, Barcelona, Spain, Nov. 2011.

[13] C. Li, W. Yin, and Y. Zhang. TVAL3: TV minimization by augmented lagrangian and alternating direction algorithms, 2009.

[14] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. Brady. Compressive sensing for video using a passive coding element. In *Computational Optical Sensing and Imaging*, pages CM1C–3. Optical Society of America, 2013.

[15] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. J. Brady. Coded aperture compressive temporal imaging. *Optics express*, 21(9):10526–10545, 2013.

[16] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling and Simulation*, 4(2):460–489, 2005.

[17] J. Y. Park and M. B. Wakin. Multiscale algorithm for reconstructing videos from streaming compressive measurements. *Journal of Electronic Imaging*, 22(2):021001–021001, 2013.

[18] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graphics*, 25(3):795–804, 2006.

[19] D. Reddy, A. Veeraraghavan, and R. Chellappa. P2C2: Programmable pixel compressive camera for high speed imaging. In *IEEE Conf. Comp. Vision and Pattern Recog*, Colorado Springs, CO, USA, June 2011.

[20] A. C. Sankaranarayanan, C. Studer, and R. G. Baraniuk. CS-MUVI: Video compressive sensing for spatial-multiplexing cameras. In *IEEE Intl. Conf. Computational Photography*, 2012.

[21] A. C. Sankaranarayanan, P. K. Turaga, R. Chellappa, and R. G. Baraniuk. Compressive acquisition of linear dynamical systems. *SIAM Journal on Imaging Sciences*, 6(4):2109–2133, 2013.

[22] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. Multiplexing for optimal lighting. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(8):1339–1354, Aug. 2007.

[23] J. M. Schmitt. Optical coherence tomography (oct): a review. *IEEE J. Selected Topics in Quantum Electronics*, 5(4):1205–1215, 1999.

[24] N. Vaswani. Kalman filtered compressed sensing. In *IEEE Conf. Image Process.*, San Diego, CA, USA, Oct. 2008.

[25] N. Vaswani and W. Lu. Modified-cs: Modifying compressive sensing for problems with partially known support. *IEEE Trans. Signal Processing,*, 58(9):4595–4607, 2010.

[26] A. Veeraraghavan, D. Reddy, and R. Raskar. Coded strobing photography: Compressive sensing of high speed periodic events. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):671–686, Apr. 2011.

[27] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *App. Optics*, 47(10):44–51, 2008.

[28] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk. Compressive imaging for video representation and coding. In *Pict. Coding Symp.*, Beijing, China, Apr. 2006.

[29] X. Yuan, P. Llull, X. Liao, J. Yang, G. Sapiro, D. J. Brady, and L. Carin. Low-cost compressive sensing for color video and depth. *arXiv preprint arXiv:1402.6932*, 2014.