

# Compressive Sampling Using a Pushframe Camera

Stuart Bennett, Yoann Noblet, Paul F. Griffin, Paul Murray, Stephen Marshall, *Senior Member, IEEE*, John Jeffers, and Daniel Oi

**Abstract**—The recently described pushframe imager, a parallelized single pixel camera capturing with a pushbroom-like motion, is intrinsically suited to both remote-sensing and compressive sampling. It optically applies a 2D mask to the imaged scene, before performing light integration along a single spatial axis, but previous work has not made use of the architecture’s potential for taking measurements sparsely. In this paper we develop a strongly performing static binarized noiselet compressive sampling mask design, tailored to pushframe hardware, allowing both a single exposure per motion time-step, and retention of 2D correlations in the scene. Results from simulated and real-world captures are presented, with performance shown to be similar to that of immobile — and hence inappropriate for satellite use — whole-scene imagers. A particular feature of our sampling approach is that the degree of compression can be varied without altering the pattern, and we demonstrate the utility of this for efficiently storing and transmitting multi-spectral images.

**Index Terms**—Compressive sampling, pushframe imaging, columnar block compressed sensing (BCS), parallel single pixel camera (SPC).

## I. INTRODUCTION

THE ‘pushframe’ camera concept, described by Noblet et al. in [1], may be viewed as a parallelized single pixel camera (SPC). As in common SPC designs, a spatial light modulator (SLM) optically imposes a variable 2D mask on the incoming image, but rather than concentrating all unmasked light on a single photodiode, the pushframe architecture optically sums the light along one axis only, focusing each masked image column on to a separate photodiode, and so the image as a whole on to a 1D sensor array. This is depicted in the Fig. 1 schematic, using an LCD for the SLM to show a simple transmissive optical path. If there is relative motion between the camera and the scene, similar to that required by a pushbroom sensor, but with the sensor array moving along- rather than across-track, a single appropriately designed 2D mask pattern can be constantly applied while a faithful image of the scene is stored, a coefficient being recorded for every scene column at every mask column. Indeed, if the SLM pattern is a diagonal line, with the elements not forming the line being a blocking mask, the camera is equivalent to a pushbroom imager. The cited application is use on a satellite

This work was supported by the UK Space Agency under the CEOI-11 grant NSTP3-PF-031. (*Corresponding author: Stuart Bennett.*)

S. Bennett, P. Murray and S. Marshall are with the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, G1 1XQ, UK (email: stuart.bennett.100@strath.ac.uk; paul.murray@strath.ac.uk; stephen.marshall@strath.ac.uk)

Y. Noblet, P. F. Griffin, J. Jeffers and D. Oi are with the Department of Physics, University of Strathclyde, Glasgow, G4 0NG, UK (email: yoann.noblet@strath.ac.uk; paul.griffin@strath.ac.uk; john.jeffers@strath.ac.uk; daniel.oi@strath.ac.uk)

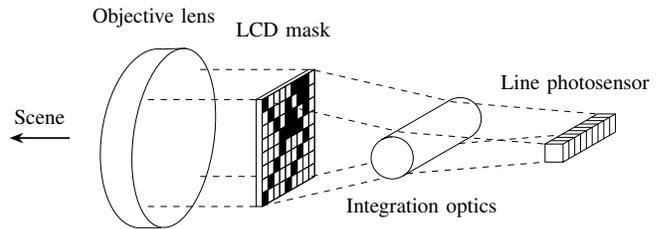


Fig. 1. Basic pushframe optical path

platform for Earth observation, where a scanning pushbroom motion is easily achieved, and the rapid movement over the Earth makes a static pattern a near-necessity if any sort of competitive ground sampling distance is to be obtained, as there is insufficient dwell time to apply multiple masks. The main benefit of the imager, relative to a pushbroom sensor, and demonstrated below in our simulated results, is improved signal to noise ratio (SNR), as multiple exposures may be merged to form the reconstructed image: while with a pushbroom sensor a ground resolution cell is captured only once per pass, a pushframe imager pass captures modulated data for the cell in each pixel of the array, and for read-noise limited sensors (common at non-visible wavelengths) combining these multiple measurements significantly improves the output SNR. This optical summation, 1D sensor array, and consequent SNR boost makes the pushframe architecture quite distinct from other binary coded aperture imagers, such as CASSI devices ([2]) where significant work has already gone into optimal pattern design (e.g. [3], [4]).

Noblet et al.’s architecture paper introduced a simple sampling scheme, based on a Walsh-Hadamard matrix, using a complete set of linearly independent one dimensional patterns, which achieved good reconstruction of the scene. However, the described approach requires  $n$  samples to reconstruct an  $n$  pixel image: while there are SNR advantages, compared to the use of a pushbroom sensor, there would be no reduction in the data storage and transmission requirements. This paper exists to advance the pushframe literature by addressing this particular shortcoming. The 2006 papers of Candès et al. ([5]) and Donoho ([6]) may be seen as establishing the field of compressive sampling (CS), where a signal is undersampled in a specific way, but then reconstructed near-perfectly through the use of prior knowledge of the signal’s properties. In particular, it is required that in some basis the signal has a sparse representation, and the samples have been taken in a different basis, preferably with these two bases having high mutual incoherence. In the case of image capture, the prior knowledge takes the form of being aware that natural

scenes have high compressibility — i.e. sparseness — in the wavelet domain. Yuan and Haimi-Cohen’s recent study ([7]) compares image storage sizes, both compressed by CS (taking fewer samples) and the JPEG algorithm (quantized discrete cosine transform), and shows that, at 10:1 compression ratio, a CS reconstruction may have similar quality to that achieved by JPEG decompression, while both methods maintain 80–90% similarity to the original image. CS has been used in combination with SPCs ([8]) to reduce the number of masks and exposures necessary to adequately reconstruct a scene, and hence reduce both capture time, and storage requirements.

In this paper we describe a bespoke pushframe CS scheme tailored to the hardware architecture described in [1]. The design of our algorithm is partially dictated by the combination of pushframe hardware and an Earth-observation (EO) application: for a given scene, the rapid scanning motion of the camera over the planet means there is insufficient time to significantly vary the sampling pattern. While an LCD was used in Fig. 1, these SLMs are not broadband, whereas a digital micromirror device (DMD) is, making a DMD the more attractive SLM for a multi-spectral EO device where all wavelengths are identically modulated and thereby intrinsically co-registered. However, a DMD can only implement a binary mask. Discounting the use of a randomly generated mask (and other designs without fast-computable transforms), on the ground that the associated large sensing matrix can pose unacceptable storage challenges during reconstruction, the constraint of the sensing matrix being binary-valued restricts the range of candidate deterministic constructed sensing bases, though several options exist (e.g. [9] and [10]). Ruling out binary schemes requiring pattern changes, such as Zhang et al.’s Fourier approach ([11]), reduces the field further still. However, Pastuszczyk et al.’s binary representation of discrete noiselets ([12]), a family of functions first described by Coifman et al. in [13], which have fast  $\mathcal{O}(n \log n)$  transforms, is ideal for DMD CS imaging. Apart from being binary, their translation from noiselets’ native complex representation is efficient, requiring only  $m + 1$  binary samples to determine  $m$  complex noiselet coefficients, and noiselets are an excellent CS sensing basis for natural scenes, as exemplified by Candès and Romberg’s use of them in [14]. Noiselets’ suitability for sampling images stems from them having provably minimal coherence with Haar wavelets (demonstrated by Tuma and Hurley in [15]), which in turn are known to permit sparse representations of natural scenes. This perfect incoherence distinguishes noiselets from other sensing bases whose use for pushframe CS would be equally valid, such as the Structurally Random Matrices (SRMs) described by Do et al. in [16], where despite the SRM approach providing for fast transforms and binary mask elements, the Haar incoherence property is not as strong. Note that the Walsh-Hadamard basis, used for a sensing matrix in [1], has maximal coherence with Haar wavelets ([17]), making such a matrix unsuitable for CS even if the reconstruction algorithm used by Noblet et al. were exchanged for an optimization approach similar to that described below.

In conventional SPC applications, as described by Pastuszczyk et al., a scene may be sampled by one basis function at a time, shaping the basis to a 2D mask covering the entire scene, and quickly iterating through a sufficient number of functions to

obtain enough samples to achieve reconstruction. While Noblet et al. note the possibility of adapting the pushframe mask mid-capture for the detection of specific targets, in general iterative patterning is not possible in the orbital pushframe context, because of time constraints. Instead, at a basic level, one can apply a function to a column of the scene, and by having the adjacent mask column represent a different function, apply this other function to the same scene column at the next time step, and so on across the mask as the scene moves before the camera. The mask’s width is then dictated by the number of basis functions one wishes to use when collecting samples. The disadvantage of this approach is that one is compressing, and later reconstructing, each scene-column independently: there is no spatial, 2D, image context, and this lack of constraint hinders good reconstruction.

One way of exploiting the correlations between adjacent columns is Ouyang et al.’s compressive line sensing (CLS), described in [18]. CLS features a scanning imager *projecting* a cross-track line of varying structured illumination, whose return is captured by a single photomultiplier tube; while such a system is not directly applicable to an Earth observation application, the captured data are comparable to those obtained via SLM masking of an externally illuminated scene. Their sensing matrix models 2D blocks of recorded lines as featuring a common component with per-line unique components, and applies random per-line masks. Later work from the same group ([19]) augments their previous approach to 2D reconstruction by integrating a Bayesian inference concept.

Block compressed sensing (BCS) by Gan [20], is another departure from the ‘whole scene’ SPC approach, and conceptually more like the conventional block approach used in JPEG etc. In BCS a (small) sampling basis is fitted to a square 2D sampling window,  $32 \times 32$  pixels being suggested, which is then tiled over the scene. The tiled windows are independently sampled with a variety of patterns, allowing each tile to be stored and reconstructed separately, while capturing some 2D spatial structure. Subsequent global optimization reduces blocking artefacts, with this stage being a major focus of later BCS works, such as [21] and [22]. Since the pushframe architecture’s minimum sampling unit is the sum of a single column, small square sampling blocks are not feasible, but it is possible to construct a column-height window which is several columns wide. Summing the integrated coefficients of each window’s columns would broadly reproduce the BCS technique, but is unnecessarily wasteful of the available spatial information. Merging columnar sampling with aspects of BCS would allow 2D structure to be captured, and present as a constraint during reconstruction, giving superior results.

We believe the combination of the scanning pushframe device with a block-based compressive sampling strategy to be a unique contribution. Other ‘SPC array’ works do not feature important aspects of our approach. Arnob et al.’s system ([23]) is static, and uses the linear detector to have a set of SPCs where each handles a different spectral range, a diffraction grating having been added just before the detector in an otherwise conventional SPC design, with the whole scene sampled at once. Fowler’s theoretical paper ([24]) is more relevant, with outlines for compressive pushbroom and whiskbroom devices, but again

the aim is hyper-spectral capture. His pushbroom approach requires multiple exposures per step, as a number of masks must be displayed, with the sensors of the array independently capturing the spectrum of one spatial pixel each per pushbroom step. The whiskbroom variant uses a static mask, with the whole array sampling one spatial pixel in each exposure. Henriksson’s patent ([25]) details a hardware architecture equivalent to that of the pushframe device, but regards the avoidance of physical scanning as one of the invention’s advantages. The mask is described as consisting of varying vertical stripes, the rows displayed on the SLM being identical, with no mention of block-based compression. In [26] Wang *et al.*, and in [27] Li *et al.*, use similarly restrained one-dimensional pattern designs.

Looking more widely at the state of the art in other computational imagers, coded masks are also used in several light field camera works — while pushframe imagers do not aim to capture light fields, it is informative to compare some details to emphasize important aspects of our pushframe approach. The compressive sampling design of Babacan *et al.* ([28]) uses a random mask and multiple exposures to achieve light field image acquisition. The device requires a 2D sensor array, displays multiple amplitudes in the mask pattern, and while a high SNR is highlighted, it is still less than that achieved using a traditional 2D camera. As is the case for our CS pushframe scheme, in [29] Marwah *et al.* emphasize the highly integrated optical and algorithmic design of their architecture. It only requires a single (2D) exposure, relying on an expensive training step to form an overcomplete dictionary subsequently used in reconstruction, an approach which the authors acknowledge depends on the similarity of the training data to the captured scene: a potential issue when making measurements of an uncontrolled environment, such as in an Earth observation application. The paper of Miandji *et al.* ([30]) describes how they varied a coloured random mask between exposures by physically translating it before the sensor using a piezo system — this is not dissimilar to how a pushframe imager exploits its scanning motion to apply different masks to portions of a scene while the overall mask remains constant.

In the following Sections we shall first describe our adaptation of binarized noiselets to column-based sensing, and then detail the enhancements of our ‘columnar BCS’. Section II explains the implementation of these points, while Section III shows results obtained under both approaches, with simulated data quantifying the effects of block-size and compression parameters, and real-world data illustrating observed performance using a prototype pushframe device. Section IV discusses an extension of our methods to a multi-spectral setting, using pan-sharpening techniques. Finally, we discuss our results in Section V.

## II. METHOD

CS is typically expressed as the recovery of some original signal  $\mathbf{x}$ , from a set of measurements  $\mathbf{y}$ , where  $\mathbf{y} = \Phi\mathbf{x}$ , and  $\Phi$  is a (known) sensing matrix, of size  $m \times n$ ,  $m < n$  — this inequality leading to the compression. There are infinite solutions for  $\mathbf{x}$ , as the system is underdetermined, but the constraint that  $\mathbf{x}$  has a sparse representation in some basis

$\Psi$ , i.e.  $\mathbf{x} = \Psi\mathbf{f}$ , where  $\mathbf{f}$  is a vector of mostly zero or near-zero coefficients, makes the problem tractable. Since, by substitution,  $\mathbf{y} = \Phi\Psi\mathbf{f}$ , then if  $\Psi$  is known,  $\mathbf{f}$  can be obtained through optimization, in turn allowing the reconstruction of  $\mathbf{x}$ . In the imaging case it is common that  $\mathbf{x} = \text{vec}(\mathbf{X})$ , where  $\text{vec}(\mathbf{X})$  denotes the vectorization of the 2D image  $\mathbf{X}$  (of width  $w$  and height  $h$ ) by vertically stacking all of  $\mathbf{X}$ ’s columns into a single vector.

As Pastuszczyk *et al.* explain, fewer measurements,  $m$ , are required as the mutual coherence of  $\Phi$  and  $\Psi$  decreases; natural scenes compress well in the Haar wavelet domain, making it a good candidate basis for  $\Psi$ ; and noiselet functions have minimal coherence with the Haar basis. Noiselet matrices are unitary, with orders of the form  $2^q$ , and a  $\Phi$  matrix may be simply obtained by choosing random rows from an appropriately sized noiselet matrix (assuming  $n$  is a power of two) in proportion to the desired compression ratio. Therefore the noiselet family of functions are a good CS imaging measurement basis. But noiselets are complex valued, and can have -1, 0 and +1 magnitudes in both their real and imaginary parts. In [12] Pastuszczyk *et al.* detail transformations such that an  $m \times n$  noiselet sensing matrix  $\Phi$  can be replaced by a  $(m + 1) \times n$  matrix  $\mathbf{P}$ , where the elements of  $\mathbf{P}$  are binary valued — suitable for DMD patterning — and samples  $\tilde{\mathbf{y}}$  may be taken by

$$\tilde{\mathbf{y}} = \mathbf{P}\mathbf{x}. \quad (1)$$

A reverse transformation converts  $\tilde{\mathbf{y}}$  to complex-valued  $\mathbf{y}$ , suitable for standard CS reconstruction algorithms, as long as a measurement of  $\mathbf{X}$ ’s mean intensity is available for scaling. Such a measurement can be achieved by adding an all-ones row to  $\mathbf{P}$ , hence the  $m+1$  rows stated above.

### A. Pushframe binarized noiselets

Equation (1) may be decomposed into  $\tilde{y}_i = \langle \mathbf{p}_i, \mathbf{x} \rangle$ , where  $\tilde{y}_i$  ( $1 \leq i \leq m$ ) is a single measurement coefficient, and  $\mathbf{p}_i$  is the  $i$ th row of  $\mathbf{P}$ . In SPC sensing the whole of the scene  $\mathbf{X}$  is sampled by one  $\mathbf{p}_i$  at any time, and a single  $\tilde{y}_i$  recorded. That is,  $n$ , the number of columns in  $\Phi$  and  $\mathbf{P}$ , is equal to  $wh$ , the number of elements in  $\mathbf{X}$ . The dot product summation is physically achieved by all the light transmitted through the  $\mathbf{p}_i$  mask being integrated by a single photodiode.

In pushframe sampling the summation occurs separately for each column of  $\mathbf{X}$ , each sum only capturing a part of the effect of the overall mask, and each generating a distinct coefficient. A simple way of exploiting this flexibility is to generate  $\Phi$  and  $\mathbf{P}$  such that  $n$  is equal to  $h$ , similar to BCS having block-sized sensing matrices; in the above equations  $\mathbf{x}$  is then no longer a vectorized version of  $\mathbf{X}$ , but a single column of it. By displaying all  $m + 1$  sensing row vectors *as columns* adjacent to each other on the SLM (pattern ‘one’ values transmit/reflect, pattern ‘zero’ values block), the dot product of each column of  $\mathbf{X}$  with all  $m + 1$  vectors will be recorded after the capturing device has taken  $m + w$  exposures as it steps in a pushbroom-like manner across the scene (see Fig. 2). Each column of  $\mathbf{X}$  can then be reconstructed independently, once the coefficients are transformed to complex numbers. In a theoretical realization, such as that used in simulation, the complex coefficients used



to sample a vectorized four columns-wide block

$$\Phi = \begin{bmatrix} \Phi_1 & & & \\ & \Phi_1 & & \\ & & \Phi_1 & \\ & & & \Phi_1 \end{bmatrix}. \quad (3)$$

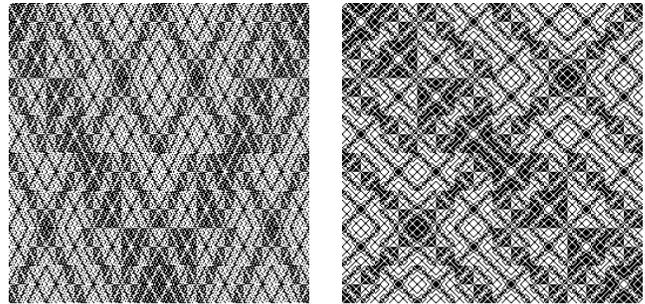
It may be seen that if  $\Phi$  were not sparse, using such a sensing matrix would require the summation of coefficients from all the columns in the block, as in conventional BCS. By specializing our  $\Phi$ , so each sub-matrix is applied to one scene column at a time, we save these additions, and retain more spatial information for reconstruction, resulting in better performance than if we had filled the matrix with  $n = bh$  noiselets.

Our more sophisticated proposal however does not repeat  $\Phi_1$ . As described at the start of this Section,  $\Phi_1$  was created by drawing  $m$  rows from an  $n \times n$  noiselet matrix  $\mathbf{N}$ , where  $m < n$ . There are therefore other possible length  $n$  noiselet rows which are not present in  $\Phi_1$ . As different noiselet rows capture different aspects of a sampled scene column, and, on average, adjacent scene columns have high correlation, sampling scene column  $j + 1$  with a noiselet row not used in sampling column  $j$  may still provide information relevant to the reconstruction of column  $j$ , as long as the image recovery algorithm does not process the columns independently. So treating the  $n$  available rows of  $\mathbf{N}$  as a pool, we form  $\Phi_2$  using rows not used in  $\Phi_1$ . Clearly the number of unused rows depends on the chosen compression ratio, i.e. how much smaller  $m$  is than  $n$ : for small  $m$  there may still be unused rows to be drawn when forming  $\Phi_3$ , otherwise, at some point we start redrawing from the pool those least recently used, until we have constructed  $\Phi_b$ . These sensing sub-matrices are then placed in our block-diagonal  $\Phi$ :

$$\Phi = \begin{bmatrix} \Phi_1 & & & \\ & \Phi_2 & & \\ & & \ddots & \\ & & & \Phi_b \end{bmatrix}. \quad (4)$$

It is apparent that  $\Phi$  grows quickly with  $b$ : if  $b$  were 16,  $n$  256, and a compression ratio of 40% (i.e. 2.5:1) gave an  $m$  of 102, the sensing matrix would have over six million elements, albeit with high sparsity. Fortunately, no such impractically large numbers need apply to a hardware sensing implementation. We simply display the entire pool of (binarized) noiselet rows, together with an all-ones column, on the SLM as an  $n \times (n + 1)$  array, and then map the relevant measured coefficients from each exposure to those rows allocated to each of the  $b$  columns in each block. To achieve image recovery,  $b$  rows of the cropped sample matrix  $\mathbf{S}$  are passed as a single vector to the minimization algorithm, and the output vector from this is reshaped to a  $h \times b$  image block, and assembled with adjacent blocks to form the image.

In displaying the entire set of binarized noiselet rows on an SLM, there is free choice for which order the rows appear in — as described above the coefficients are going to be remapped in any case. Pastuszcak et al. translate noiselets to binarized noiselets by having the  $j$ th row of  $\mathbf{N}$  map to rows  $2j - 1$  and  $2j$  of the pattern matrix. Since  $\mathbf{N}$  has a canonical row ordering, resulting from the definition of noiselet matrices, a full pattern



(a) A full  $n = 256$  binarized noiselet matrix, following Pastuszcak et al.'s mapping  
(b) A full  $n = 256$  binarized noiselet matrix, with mirrored mapping

Fig. 4. Alternative SLM patterns. (b) is more optically robust, empirically.

matrix constructed following Pastuszcak et al.'s mapping can be generated, as shown in Fig. 4a. However, we prefer to use a mirrored mapping, as illustrated in Fig. 4b, where the  $j$ th row of  $\mathbf{N}$  maps to rows  $j$  and  $n + 1 - j$  of the pattern matrix. The main benefit of a matrix formed using this mapping is its superior real-world performance: it seems that the larger contiguous black or white regions mean the degrading effects of optical crosstalk are less severe in experimental data. A secondary advantage is that Pastuszcak et al.'s binarization scheme requires that where the noiselet row  $j$  is used, so is the row  $n + 1 - j$ , so our scheme means the same mapping can be employed for tracking the row indices used both with  $\mathbf{P}$  for sampling, and with  $\Phi$  for reconstruction.

### III. RESULTS

#### A. Simulated

We begin this section by showing simulated, noise-free, results, as simulation allows us to know exactly how the recovered image should appear, and provide quantitative measures of our algorithm's performance in achieving that aim. Our simulated results feature the capture and reconstruction of the two images below. The first, 'Blue Jay', seen in Fig. 5a, features a variety of textures, areas of high and low detail, and variations in contrast, making a reasonably challenging subject. The second, 'Grézac'<sup>1</sup>, seen in Fig. 5b, possesses some similar features, but is more representative of Earth-observation satellite imagery, the target application for a pushframe imager. Both images are  $256 \times 256$  pixels in size, this being the highest resolution achieved on an experimental pushframe imager by Noblet et al. From this it follows that all sensing is performed with  $n = 256$ , unless otherwise noted.

In Fig. 6, the comparative performance of three sampling methods from the previous Section can be seen. The three methods are:

- 1) the basic scheme described in Section II-A;
- 2) the 'naïve' scheme mentioned in Section II-B; as we developed towards
- 3) our enhanced proposal, using all available noiselet rows, in the same subsection.

<sup>1</sup>cropped from an aerial photograph by Jacques Dassié, published under the CC BY 2.5 license



(a) 'Blue Jay'



(b) 'Grézac'

Fig. 5. Test images used in reconstruction simulations

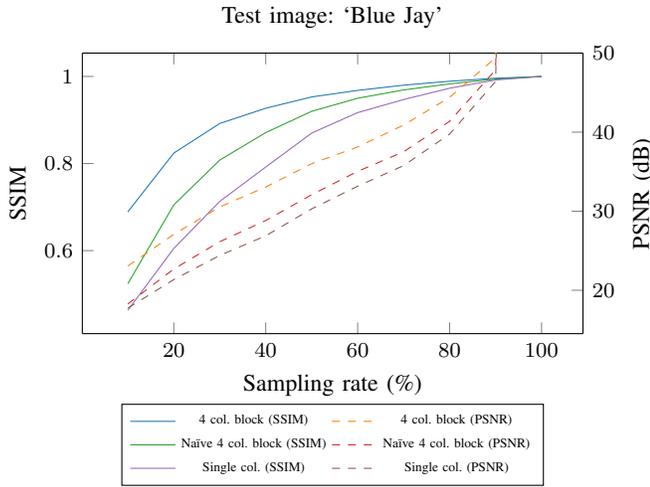


Fig. 6. Comparison of image recovery performance for the three methods described in Section II, for varying levels of data compression

For the later two methods a modest block width of four columns is used.  $m$  is varied to change the number of samples taken, which manifests as the level of data compression: 10% being a reduction to a tenth of the uncompressed image's sample count, and 100% being no compression. For each tested configuration the reconstructed image is compared to the uncompressed original, and we measure the peak signal to noise ratio (PSNR) and the more perceptually based structural similarity index measure (SSIM), presented in [37]. PSNR reaches infinity for an identical image, whereas SSIM takes values between -1 and 1, with 0 indicating no similarity, and 1 perfect similarity.

It is easily seen that by both metrics, the block-based methods outperform the single column approach, and the more advanced block sensing algorithm beats the more basic one, for all tested sampling rates. Fig. 6 only shows data from the 'Blue Jay' image, but the plots from the 'Grézac' image are very similar.

Our columnar block sensing method having been shown to perform well, in Fig. 7 we can see the effect of varying the block width, using the same metrics as before. The clear finding is that wider blocks perform better. Intuitively this makes sense:

the more blocks that are independently reconstructed, the more errors are likely to occur at the boundaries between blocks, as the optimization process has no visibility over a block's border (hence the research interest in reducing BCS blocking artefacts). Nonetheless, beyond a certain block size the increase in returns diminishes; the two test images behave a little differently, but the improvement beyond a block width of 16 or 32 pixels is slight.

Fig. 7 shows a maximal block width of 256 pixels — the size of the sampled image. For comparison, at the  $x = 256$  point of the graphs, crosses have been plotted to show the comparable (binarized noiselet) performance of doing a whole-frame sampling reconstruction (i.e.  $n = 65536$ ), as one might use with a conventional SPC, using the same sampling rates as tested for the pushframe imager. For the 'Blue Jay' image the pushframe block sensing achieves near identical SSIM performance to the whole-frame approach, whereas the whole-frame approach consistently outperforms the other on PSNR by around 2 dB. The situation differs for the 'Grézac' image however, with the pushframe narrowly beating the conventional SPC approach on all metrics. While finding that the pushframe has a natural affinity to Earth-observation images would be pleasing, it is more likely that the better performer of pushframe and conventional SPC depends highly on the details of any particular scene.

The above simulations assume perfect optical and sensor behaviour, but in Fig. 8 the effects on image recovery zero-mean Gaussian noise being additively introduced to the sampled coefficients are shown for the 'Blue Jay' image (the 'Grézac' image results are similar). This gives an indication of expected performance in more realistic circumstances, particularly that likely under low-light conditions, where sensor noise can be significant. The source image is normalized such that the maximal intensity passing through any mask pixel is one. Three compressive pushframe trials are included, all using a 256 px 2D block width, but varying the sampling rate between 20 and 80%. All three show gradual degradation as noise increases, but at higher noise levels the reconstructions using fewer samples perform better. This counter-intuitive behaviour is believed to result from the optimization stage having fewer contradictory

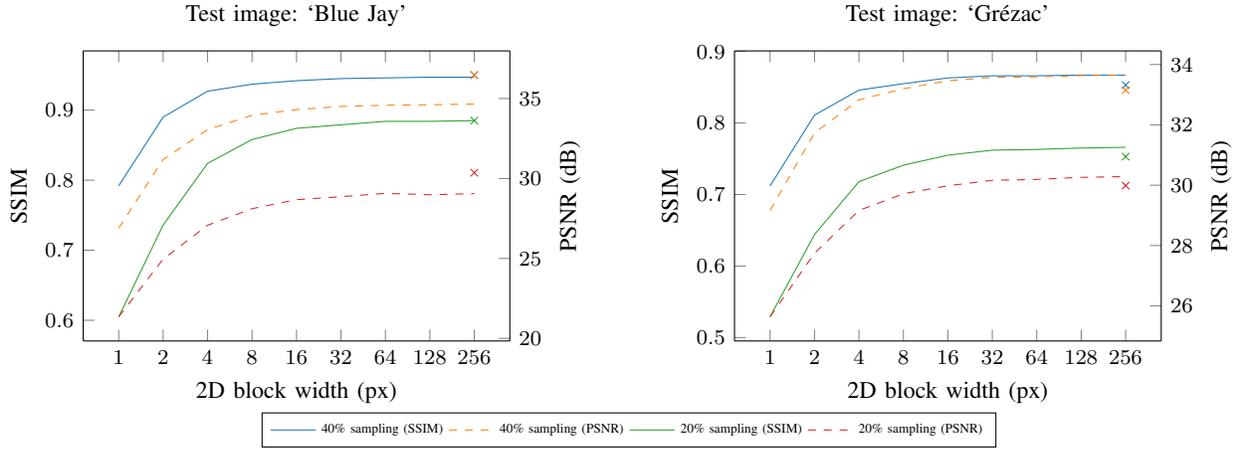


Fig. 7. The effect of different 2D block widths on image recovery performance. The cross ( $\times$ ) marks at the right of the plots are the comparable performance obtained when simulating an SPC-like whole-frame sensing approach.

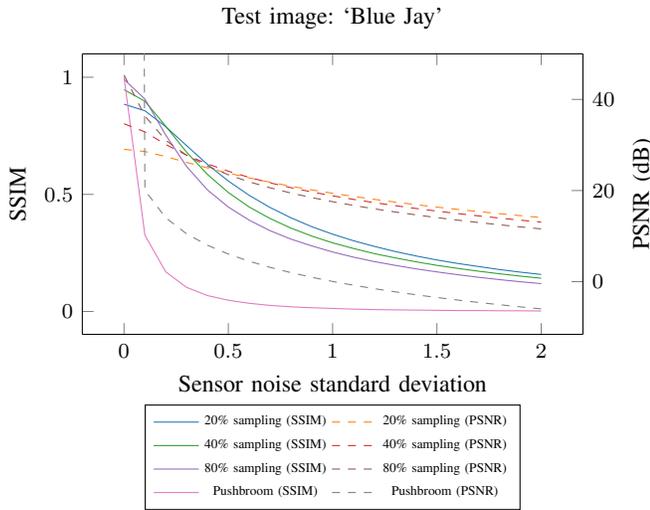


Fig. 8. The effect on image recovery performance of adding different strengths of Gaussian noise to the linear detector samples, at three compressive pushframe sampling rates. A pushbroom simulation is included for comparison.

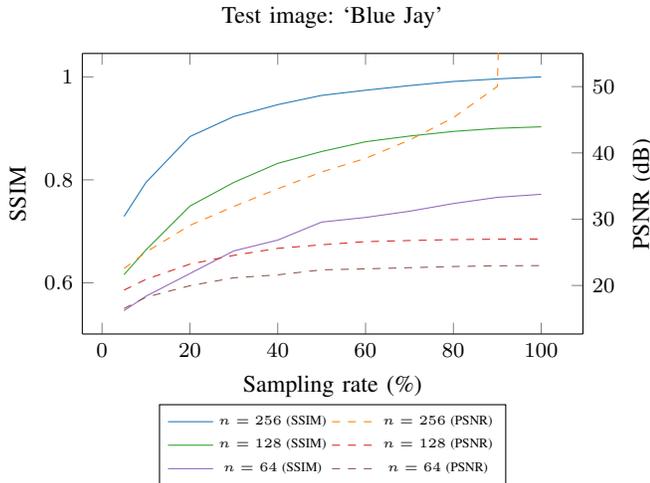


Fig. 9. Trade-off analysis of using different mask resolutions (varying the pattern density) at different sampling rates

noisy samples to reconcile, instead being able to rely on the image’s Haar-domain sparseness, and TV smoothness, to counter the introduced noise. A fourth pair of plots allows comparison of the recovery performance when the image is sampled using a pushbroom approach, with the sensor subject to the same noise/light levels, and it is clear that the pushframe arrangement is vastly more robust to such noise. The reason for this is clear: the pushframe architecture sums the intensity of each masked pixel with  $n$  other intensities along the whole column, so for  $n = 256$ , the intensity observed by the sensor is likely to be two orders of magnitude greater than in the pushbroom case, meaning the addition of a certain strength of noise will have a significantly greater corrupting effect in the pushbroom case.

The final simulation for this section examines the trade-off between mask resolution and sampling rate. The test image is downsampled to match the height of the mask, with the  $n$  of the mask varied between 64, 128 and 256 px (a 2D block size of 64 px was used throughout). Following reconstruction the recovered image is upscaled to the original test image size before comparison. Essentially this simulates varying the mask pattern’s density. The results for ‘Blue Jay’ are shown in Fig. 9, with ‘Grézac’ again having no substantial differences. For all tested resolutions, performance deteriorated as the mask’s resolution decreased.

### B. Real-world

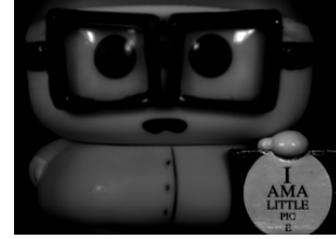
For our experimental data, we mount our pushframe imager to a translation stage, causing the scene, approximately 2.5 m away, to move sideways from the imager’s viewpoint. We use the same arrangement as in Noblet *et al.*’s experiments, emulating the optical integration stage by summing the columns captured by a 2D imaging sensor. This eases experimental alignment and debugging, allows capture of comparison data, and avoids the cost of manufacturing the unusual optics required. A reference image of the capture subject (a doll of *South Park*’s Eric Cartman character), recorded with the same optics and camera module as used to record the pushframe data, but using the pushbroom-emulation mask mentioned in

the introduction, is shown in Fig. 10a. This reference is *not* expected to match the recovered images exactly, as the different masks will sample a given part of the scene with the imager at different positions relative to the imaged subject, so perspective and lighting differences are certain. Our entire optical system uses commercial off-the-shelf components: a Rubinar 300 mm f/4.5 telescope, an Epson L3C07U-85G13 LCD module as the SLM, and a FLIR Blackfly S BFS-U3-200S6M camera with a Kowa LM50JC10M 50 mm lens mounted via a 22 mm extension tube.

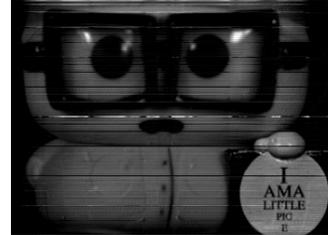
The transmissive LCD is used to achieve high-quality imaging: in a pushframe system the incoming scene must be focused on the SLM, and the image of the mask must also be well-focused on the sensor, but a DMD cannot be simultaneously perpendicular to the optical axes of both the objective lens and the integration optics, making achieving uniform sharp focus across the entire frame challenging. On other relevant measures, the DMD/LCD differences are in this case insignificant: while a DMD mirror exhibits theoretically infinite contrast, the extinction ratio of over 1000 measured for our LCD is more than adequate when compared to the sensor’s noise and intensity resolution performance. Spatial contrast for the pattern is instead limited by the modulation transfer function (MTF) of the optics between the SLM and sensor — if optical crosstalk between adjacent white and black pattern pixels results in a loss of contrast at the sensor, this is not caused by the SLM, but nonetheless will degrade reconstruction performance — hence the use of the high resolution LM50JC10M lens. A DMD does have a high fill factor (FF), over 90%, transmitting more light through the system than an LCD’s near 50% FF, but this loss is uniform across the whole frame, and so does not affect image reconstruction, as long as sensor exposure times are adjusted to normalize the incoming light levels. The experiments are not conducted at the read-noise limited wavelengths which benefit most from pushframe imagers (and might necessitate the use of a broadband DMD), as the design of the sensing masks is independent of the imaged wavelengths.

As in Noblet et al.’s paper, we derive a partial flat field correction by recording the pushframe coefficients obtained when imaging an all-white scene. These values are then input as a normalizing calibration curve, weighting subsequently measured coefficients, to compensate for non-uniform optical transmission (vignetting etc.) resulting from the telescope. This spatial intensity variation is naturally a 2D phenomenon, while the pushframe architecture constrains our correction to only be 1D, so pushframe reconstructions will always be limited by imperfect optics.

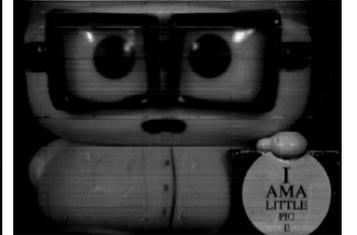
To give a comparison of the sampling and recovery techniques *without data compression*, Fig. 10b–10e feature a previously described technique, that same technique with post-processing, and two variants of the enhanced approach from Section II-B. Fig. 10b shows a reconstruction of the scene using the Walsh-Hadamard sensing matrix approach previously described in Noblet et al.’s paper. Fig. 10c shows the same data post-processed with a  $3 \times 1$  (i.e. purely vertical) median filter to remove the characteristic impulsive rows; these rows correlate with the sensing matrix structure, but also depend on the scene content, so more specific filters are not appropriate,



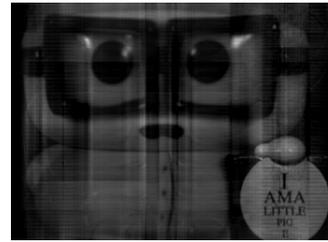
(a) Pushbroom-like reference capture



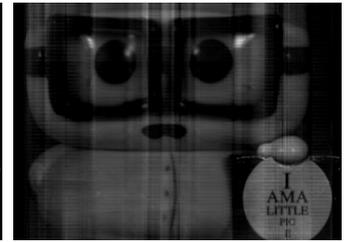
(b) Hadamard matrix capture



(c) Median-filtered Hadamard capture



(d) Columnar BCS, 100% sampled with Pastuszczyk’s P



(e) Columnar BCS, 100% sampled with ‘mirrored’ P

Fig. 10. Comparison of fully sampled scene images

TABLE I  
PERFORMANCE OF FULLY SAMPLED IMAGE RECOVERY, RELATIVE TO REFERENCE IMAGE

Variant	PSNR (dB)	SSIM
Walsh-Hadamard	21.6	0.629
Walsh-Hadamard, median filtered	26.3	0.826
CS, Pastuszczyk’s ordering	22.1	0.550
CS, mirrored ordering	22.9	0.644

and a larger kernel size would suppress detail. Fig. 10d and Fig. 10e were captured using the Fig. 4a and Fig. 4b sensing patterns respectively, with  $b = n = 256$ . All four recovered images have good qualitative similarity to the reference image, but, of the two CS variants, the superiority of the ‘mirrored’ one is clear, with less horizontal banding in the reconstruction, while the Walsh-Hadamard images suffer some ‘ghost’ artefacts, such as the outline of the character’s eyes being present in its chest. In all cases, reconstruction artefacts may be attributed to nonidealities (mask crosstalk, telescope vignetting, etc.) in the optical path. Table I gives PSNR and SSIM values for the reconstructions relative to the reference image, but, as noted above, the comparison is not expected to be perfect.

On PSNR and SSIM the mirrored CS mask outperforms both the unfiltered Walsh-Hadamard result and the other CS variant, but the filtered Walsh-Hadamard image has the best qualitative performance tabulated — though the reconstruction

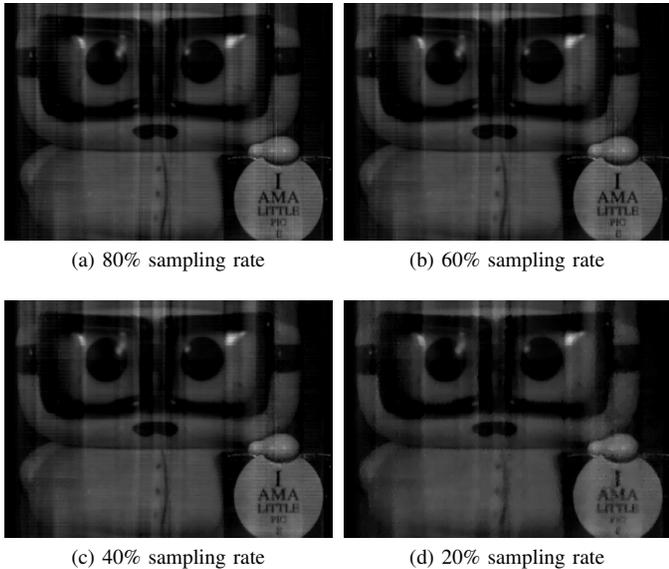


Fig. 11. Effects of reducing sampling rate on columnar-BCS reconstructions

method does not accommodate less than 100% sampling. With all other experimental parameters unchanged, Fig. 11 shows the effects of reducing the sampling rate, when a mirrored-ordering sensing pattern is used with the Section II-B algorithm. There is little degradation apparent until the sampling rate is below 60%, even at 40% the smallest text in the image is clearly legible, and at 20% the qualitative recovery is still good.

#### IV. MULTI-SPECTRAL EXTENSION

The pushframe architecture elaborated by Noblet *et al.* can simultaneously capture *intrinsically co-registered* images in multiple wavelengths, by following the SLM with dichroic mirror beamsplitting, diffraction of the integrated light on to an area sensor, or a colour (Bayer filtered) camera. The sampling technique described in Section II can obviously treat wavelengths separately, and reconstruct them independently, but then the data storage and transmission requirement will scale linearly with the number of colour channels. Images captured at different wavelengths will usually exhibit strong spectral correlation, which is indicative that independent sampling is suboptimal, and some sort of inter-band sensor fusion would yield improved results.

Pushframe multi-spectral data automatically have the same sensing matrix applied in all channels, the incoming scene having interacted with the single SLM before reaching the sensors. Using the refined block-based sensing approach of Section II-B, where the SLM is patterned with all possible binarized noiselet rows, each scene column is in fact sampled with every sensing pattern, and compression ratio is a matter of which samples are discarded. For multi-spectral samples, the question then is can this complete availability of samples be exploited, by intelligently choosing which coefficients from each wavelength are to be discarded or combined with those from other wavelengths, in order to optimize image recovery following storage or onward transmission of the reduced dataset. It is possible to retain different coefficients for different

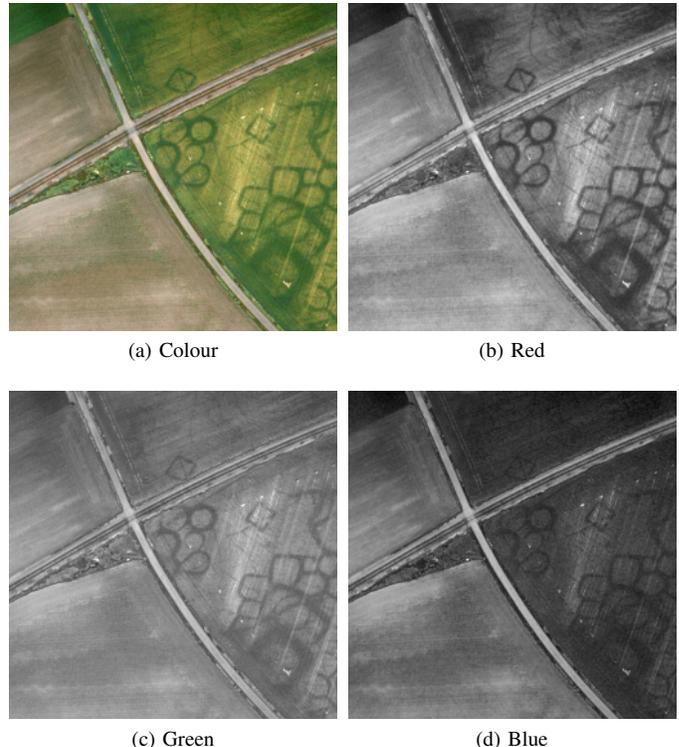


Fig. 12. Colour Grézac image, along with its RGB decomposition

bands, capturing a richer variety of spatial data for a given scene column, and then either fuse all the information in a multi-spectral CS reconstruction algorithm (for instance [38], [39] or [40]), or use a multi-spectral super-resolution technique (such as [41]) on independently reconstructed channels. However, to more directly demonstrate the possibilities resulting from our fully patterned SLM, we instead show a simple way of achieving superior multi-spectral compression below, inspired by the existing satellite-sensing practice of pan-sharpening. This is typically used where satellites have a high-resolution broadband ‘panchromatic’ camera, and a number of narrowband cameras of lower resolution, and it is desired to enhance the apparent resolution of the latter.

Rather than have a high resolution panchromatic band separately captured from lower resolution colour bands, we synthesize a panchromatic band from the colour band samples before they are compressed further. More formally, if sample sets  $\mathbf{S}_R$ ,  $\mathbf{S}_G$ , and  $\mathbf{S}_B$  have been collected with  $m = n+1$  in red, green and blue bands respectively,  $\mathbf{S}_{pan} = \frac{1}{3}(\mathbf{S}_R + \mathbf{S}_G + \mathbf{S}_B)$ , and subsequently only  $m_{pan}$  columns of  $\mathbf{S}_{pan}$ , and  $m_{band}$  columns ( $m_{band} \ll m_{pan}$ ) of  $\mathbf{S}_R$ ,  $\mathbf{S}_G$  and  $\mathbf{S}_B$ , are retained. Standard pan-sharpening techniques (such as those in [42]) can then be used for recovering high resolution colour bands. In the following results we use an efficient IHS method, as derived in [43], where the high frequency differences between the reconstructed panchromatic image and a panchromatic image formed from the reconstructed colour bands are applied straightforwardly to each of the reconstructed colour bands, i.e.  $\Delta = \mathbf{I}_{pan} - \frac{1}{3}(\mathbf{I}_R + \mathbf{I}_G + \mathbf{I}_B)$ ,  $\mathbf{I}'_R = \mathbf{I}_R + \Delta$ , and similarly for  $\mathbf{I}'_G$  and  $\mathbf{I}'_B$ .  $\mathbf{I}_{pan}$ ,  $\mathbf{I}_R$ ,  $\mathbf{I}_G$  and  $\mathbf{I}_B$  are the recovered panchromatic

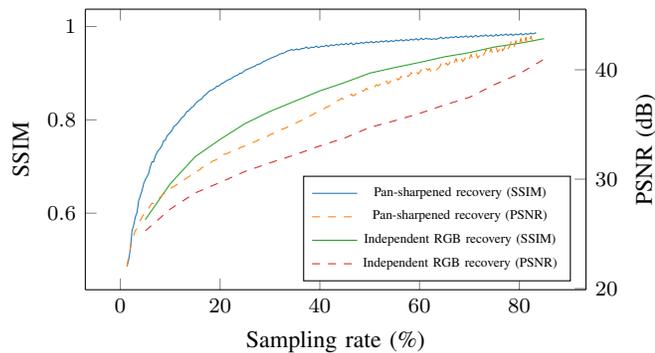


Fig. 13. Comparison of image recovery performance using heterogeneous (overall sampling ratio  $\bar{m}$ ) and homogeneous ( $m$ ) band sampling, using pan-sharpening and independent RGB reconstruction respectively

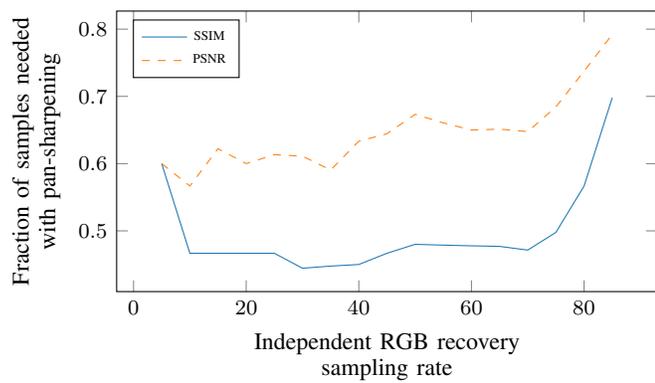


Fig. 14. Illustration of sample count savings made possible by using non-homogeneous band sampling

and band images respectively, while the  $\mathbf{I}'$  images are the high resolution versions. For our simulated capture we use a colour version of the ‘Grézac’ image (Fig. 12a), whose similarities and differences between the red, green and blue channels can be seen in Fig. 12b–12d.

The ability to vary  $m_{pan}$  and  $m_{band}$  independently gives a 2D parameter space for sampling rate, but these may be combined into a single effective value  $\bar{m}$ , comparable to having a single  $m$  applied to independent band sampling/reconstruction:  $\bar{m} = \frac{m_{pan}}{3} + m_{band}$ . In Fig. 13 the PSNR and SSIM are plotted for the  $\bar{m}$  having the highest PSNR or SSIM respectively for a variety of  $m_{pan}$  and  $m_{band}$  combining to give that  $\bar{m}$ , along with the PSNR and SSIM resulting from varying  $m$  when the bands are independently recovered.

For all tested sampling rates the pan-sharpened approach achieves substantial improvements in both metrics, demonstrating the usefulness of a block-compressing pushframe imager’s capability of choosing which samples to retain. The improvement is further illustrated in Fig. 14, where the proportion of pan-sharpening samples needed to achieve the PSNR/SSIM obtained with independent channel reconstruction is plotted. For the majority of compression ratios a half to two thirds of the data are required.

## V. CONCLUSION

We have described a way of achieving compressive sampling on a pushframe imager. By using a block-based reconstruction approach, with column-high blocks, our technique delivers good performance, both in simulation and on real-world data, exploiting 2D spatial structure in an imaged scene. At the same time, it is capable of efficient hardware implementation, through our observation that it is straightforward to sample with a complete binarized noiselet matrix, and then only retain desired combinations of coefficients for image recovery. Our method of arranging the sampling basis functions obtains superior results to an alternative arrangement on optically captured data.

Our simulations demonstrate that drawing rows from the noiselet basis matrix *uniformly* when sampling each block gives better results than repeated use of a subset of rows, or not using blocks at all. It may be possible to further optimize the drawing of rows, perhaps following the blue noise approach used in the more mature CASSI literature ([3]), but this remains to be studied in future work. We observed that performance increases as blocks become wider, but gains become more and more marginal. For offline image recovery, reconstructing very large blocks to gain the best recovery, or using optimization techniques from the BCS literature to remove blocking artefacts, may not be problematic, but in an online situation intermediate block-widths might be more suitable to allow a reconstructed image, still of high quality, to be built up over time. In a comparison of pushframe and conventional SPC sensing of different scenes there was no clear winner, performance being scene-dependent. It is fair however from these results to conclude that pushframe block sensing is not necessarily worse than whole-frame SPC sensing, and can be better. Our noise simulation illustrates the major performance advantage of pushframe capture over pushbroom capture for read-noise limited sensors, as are common at non-visible wavelengths.

Our experiments showed little deterioration in recovery performance at low compression ratios, though the reconstructions were consistently imperfect. This performance limit at higher sampling rates is likely to arise from optical nonidealities — the quality of recovery is very sensitive to the 1D flat field correction curve for instance. This is also the case for other pushframe reconstruction techniques however, so we can conclude that compressive sampling is especially beneficial when using a pushframe imager: higher data rates confer no advantage in terms of recovered image quality, so to not compress would be wasteful. Our results show that good recovery is still achieved at lower sampling rates, with the image quality decreasing gracefully as compression increases.

Finally, we have outlined a method for getting even greater data compression when multi-spectral samples are captured, as may easily be achieved via dichroic mirror beamsplitting or a colour filter array. In itself, the pan-sharpening recovery technique is not sophisticated, but the  $\sim 40\%$  data reduction nonetheless attained is a powerful demonstration of the utility of our enabling pushframe CS approach — without changing the common mask pattern, variation of different bands’ compression levels is very computationally cheap.

## ACKNOWLEDGMENT

The authors would like to thank Wideblue Limited, Prof. M. Macdonald and Dr. S. R. Owens for useful discussions.

## REFERENCES

- [1] Y. Noblet, S. Bennett, P. F. Griffin, P. Murray, S. Marshall, W. Roga, J. Jeffers, and D. Oi, "Compact multispectral pushframe camera for nanosatellites," *Applied Optics*, vol. 59, no. 27, pp. 8511–8518, Sep. 2020.
- [2] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics*, vol. 47, no. 10, pp. B44–B51, Apr. 2008.
- [3] C. V. Correa, H. Arguello, and G. R. Arce, "Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging," *Journal of the Optical Society of America*, vol. 33, no. 12, pp. 2312–2322, Dec. 2016.
- [4] H. Zhang, X. Ma, D. L. Lau, J. Zhu, and G. R. Arce, "Compressive spectral imaging based on hexagonal blue noise coded apertures," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 749–763, 2020.
- [5] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [7] X. Yuan and R. Haimi-Cohen, "Image compression based on compressive sensing: End-to-end comparison with JPEG," *IEEE Trans. Multimedia*, vol. 22, no. 11, pp. 2889–2904, 2020.
- [8] 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.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.
- [9] A. Amini and F. Marvasti, "Deterministic construction of binary, bipolar, and ternary compressed sensing matrices," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2360–2370, Apr. 2011.
- [10] W. Lu, K. Kpalma, and J. Ronsin, "Sparse binary matrices of LDPC codes for compressed sensing," in *2012 Data Compression Conference*, ser. DCC 2012, Snowbird, UT, USA, Apr. 2012, p. 405.
- [11] Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Fast Fourier single-pixel imaging via binary illumination," *Scientific Reports*, vol. 7, no. 1, p. 12029, Sep. 2017.
- [12] A. Pastuszczak, B. Szczygiel, M. Mikołajczyk, and R. Kotyński, "Efficient adaptation of complex-valued noiselet sensing matrices for compressed single-pixel imaging," *Applied Optics*, vol. 55, no. 19, pp. 5141–5148, Jul. 2016.
- [13] R. Coifman, F. Geshwind, and Y. Meyer, "Noiselets," *Applied and Computational Harmonic Analysis*, vol. 10, no. 1, pp. 27–44, 2001.
- [14] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, Apr. 2007.
- [15] T. Tuma and P. Hurley, "On the incoherence of noiselet and Haar bases," in *SAMPLing Theory and Applications*, ser. SAMPTA'09, Marseille, France, May 2009, pp. 243–246.
- [16] T. T. Do, L. Gan, N. H. Nguyen, and T. D. Tran, "Fast and efficient compressive sensing using structurally random matrices," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 139–154, Jan. 2012.
- [17] A. Moshtaghpour, J. M. Bioucas-Dias, and L. Jacques, "Close encounters of the binary kind: Signal reconstruction guarantees for compressive Hadamard sampling with Haar wavelet basis," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 7253–7273, Nov. 2020.
- [18] B. Ouyang, F. R. Dagleish, F. M. Caimi, T. E. Giddings, W. Britton, A. K. Vuorenkoski, and G. Nootz, "Compressive line sensing underwater imaging system," *Optical Engineering*, vol. 53, no. 5, p. 051409, May 2014.
- [19] B. Ouyang, W. W. Hou, F. M. Caimi, F. R. Dagleish, A. K. Vuorenkoski, and C. Gong, "Integrating dynamic and distributed compressive sensing techniques to enhance image quality of the compressive line sensing system for unmanned aerial vehicles application," *Journal of Applied Remote Sensing*, vol. 11, no. 3, p. 032407, Jul. 2017.
- [20] L. Gan, "Block compressed sensing of natural images," in *2007 15th International Conference on Digital Signal Processing*, ser. DSP 2007, Cardiff, Wales, UK, Jul. 2007, pp. 403–406.
- [21] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *2009 IEEE International Conference on Image Processing*, ser. ICIP 2009, Cairo, Egypt, Nov. 2009, pp. 3021–3024.
- [22] T. V. Chien, K. Q. Dinh, B. Jeon, and M. Burger, "Block compressive sensing of image and video with nonlocal Lagrangian multiplier and patch-based sparse representation," *Signal Processing: Image Communication*, vol. 54, pp. 93–106, May 2017.
- [23] M. M. P. Arnob, H. Nguyen, Z. Han, and W.-C. Shih, "Compressed sensing hyperspectral imaging in the 0.9–2.5  $\mu\text{m}$  shortwave infrared wavelength range using a digital micromirror device and InGaAs linear array detector," *Applied Optics*, vol. 57, no. 18, pp. 5019–5024, Jun. 2018.
- [24] J. E. Fowler, "Compressive pushbroom and whiskbroom sensing for hyperspectral remote-sensing imaging," in *2014 IEEE International Conference on Image Processing*, ser. ICIP 2014, Paris, France, Oct. 2014, pp. 684–688.
- [25] M. Henriksson, "An imaging system parallelizing compressive sensing imaging," International Patent WO 2016/028 200 A1, Feb., 2016.
- [26] J. Wang, M. Gupta, and A. C. Sankaranarayanan, "LiSens — a scalable architecture for video compressive sensing," in *2015 IEEE International Conference on Computational Photography*, ser. ICCP 2015, Houston, TX, USA, Apr. 2015, pp. 1–9.
- [27] Y.-H. Li, X.-D. Wang, and Z. Wang, "Compressed sensing imaging system based on improved theoretical model and its weighted iterative strategy," *Optics Communications*, vol. 439, pp. 76–84, May 2019.
- [28] S. D. Babacan, R. Ansorge, M. Luessi, P. R. Mataran, R. Molina, and A. K. Katsaggelos, "Compressive light field sensing," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4746–4757, Dec. 2012.
- [29] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar, "Compressive light field photography using overcomplete dictionaries and optimized projections," *ACM Transactions on Graphics*, vol. 32, no. 4, pp. 46:1–12, Jul. 2013.
- [30] E. Miandji, J. Unger, and C. Guillemot, "Multi-shot single sensor light field camera using a color coded mask," in *26th European Signal Processing Conference*, ser. EUSIPCO 2018, Rome, Italy, Sep. 2018, pp. 226–230.
- [31] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 1, Pacific Grove, CA, USA, Nov. 1993, pp. 40–44.
- [32] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, May 2009.
- [33] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Scientific Reports*, vol. 8, no. 1, p. 2369, Feb. 2018.
- [34] D. Needell and R. Ward, "Stable image reconstruction using total variation minimization," *SIAM Journal on Imaging Sciences*, vol. 6, no. 2, pp. 1035–1058, 2013.
- [35] S. Becker, J. Bobin, and E. J. Candès, "NESTA: A fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.
- [36] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical Programming Series A*, vol. 103, no. 1, pp. 127–152, May 2005.
- [37] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [38] P. Nagesh and B. Li, "Compressive imaging of color images," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2009, Taipei, Taiwan, Apr. 2009, pp. 1261–1264.
- [39] A. Majumdar and R. K. Ward, "Compressed sensing of color images," *Signal Processing*, vol. 90, no. 12, pp. 3122–3127, Dec. 2010.
- [40] D. Sugimura, M. Tomabechi, T. Hosaka, and T. Hamamoto, "Compressive multi-spectral imaging using self-correlations of images based on hierarchical joint sparsity models," *Machine Vision and Applications*, vol. 27, no. 4, pp. 499–510, May 2016.
- [41] A. Zomet and S. Peleg, "Multi-sensor super-resolution," in *Sixth IEEE Workshop on Applications of Computer Vision*, ser. WACV 2002, Orlando, FL, USA, Dec. 2002, pp. 27–31.
- [42] G. Vivone, L. Alparone, J. Chanussot, M. D. Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.
- [43] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "Efficient intensity-hue-saturation-based image fusion with saturation compensation," *Optical Engineering*, vol. 40, no. 5, pp. 720–728, May 2001.