PAPER

Computational imaging with an extended field of view

IOP Publishing Journal of Optics
J. Opt. **23** (2021) 085703 (11pp) https://doi.org/10.1088/2040-8986/ac0ff6

Computational imaging with an extended field of view

Ritika Malik[1], **Ravikrishnan Elangovan**[2] **and Kedar Khare**[1,*]

[1] Department of Physics, Indian Institute of Technology Delhi, New Delhi 110016, India
[2] Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, New Delhi 110016, India

E-mail: kedark@physics.iitd.ac.in

Received 21 April 2021, revised 8 June 2021
Accepted for publication 29 June 2021
Published 23 August 2021

Abstract

The recorded field of view of any digital imaging system is limited by the physical size of the sensor array. While it is a common knowledge that the image field exists beyond the boundaries of the sensor array, there is generally no way to measure it unless the imaging system (or the object) is translated laterally. We propose a single-shot computational imaging system with a multiple-point impulse response that is able to effectively increase the physical size of the sensor array. The image recorded on the array sensor is not visually meaningful but can be used to recover the image field beyond the native sensor boundary via a sparsity based iterative algorithm, as we demonstrate in this work. The system concept can be considered analogous to structured illumination imaging; however, the structuring is performed here in the Fourier space in order to recover an extended image field. The effective increase in the sensor size depends on the extent of the impulses in the engineered multiple point impulse response. The achievable sensor size extension is therefore limited by the resolution of the phase mask that is introduced in the Fourier plane of the imaging system. We present a simulation study where the individual impulses in the designed impulse response extend over the original array sensor size, thereby doubling the effective sensor dimensions (a four times increase in the number of pixels) without affecting the image resolution. Both binary and gray-scale objects have been considered in our study in order to illustrate that the quality of the extended field of view image depends on the sparsity of the object under consideration. The concept of extended field of view computational imaging as presented here may find a number of practical applications.

Keywords: computational imaging, multiple-point impulse response, extended field of view, coded aperture, random convolution, PSF engineering, compressive imaging

(Some figures may appear in colour only in the online journal)

1. Introduction

Digital imaging devices such as cameras (including cell phone cameras), microscopes and telescopes commonly employ a pixelated array sensor such as a charge-coupled device (CCD)/ complementary metal oxide semiconductor chip to capture image information. Most current sensor arrays have a physical size in the range of 5–7 mm, which ultimately limits the field of view (FOV) observed by the corresponding systems in a single exposure record. Intuitively, one modality that can be used to increase the FOV beyond the sensor size is to record multiple images via scanning of a single array sensor or using multiple sensors [1]. A larger FOV than what is recorded on the sensor in one exposure can then be achieved by stitching or fusing the multiple image records. One commonly encountered example of this is the panorama photography mode now available in most smart phones. Another example is the use of slide

* Author to whom any correspondence should be addressed.

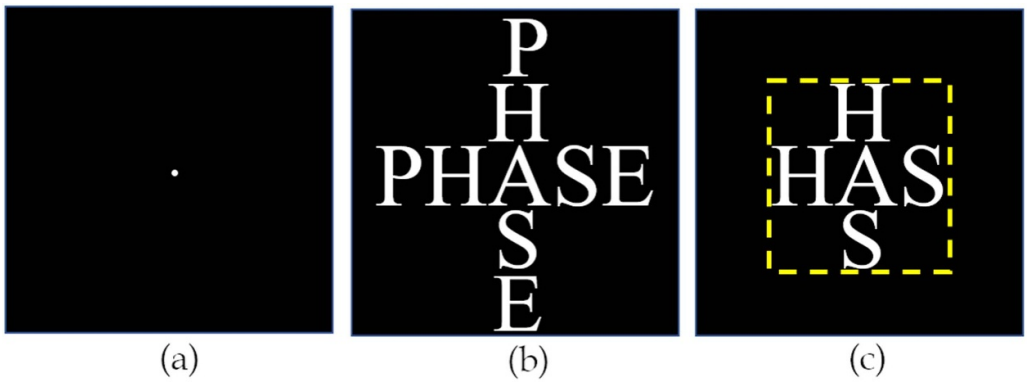

Figure 1. Pictorial representation of image recording using a conventional imaging system: (a) impulse response approximated as delta impulse. (b) Full FOV image with the binary text object. (c) Image field captured by the native sensor with size shown by the yellow dotted square. A limited FOV of the object in (b) is thus recorded on sensor.

scanners for digital pathology, where large tissue samples are scanned with a microscopy imaging system to enable whole slide imaging [2, 3]. However, this scanning and stitching process is tedious and can sometimes lead to registration artefacts due to possible issues associated with illumination and/or variable de-focus in recorded images. To overcome the FOV limitation posed by the lens in the microscopy system, a lensless imaging system for widefield microscopy was suggested in [4]. In any imaging system the resolution and FOV are coupled, and to overcome this limitation the concept of Fourier ptychography [5] allows one to obtain more resolvable pixel in the same FOV by using multiple angled illumination followed by a Gerchberg–Saxton like algorithm. A 0.5 gigapixel imaging system that uses a closed circuit television lens to relay an object onto a flatbed scanner has been reported [6]; this utilizes the scanning mechanism of a linear CCD array to get an image with a large FOV. A microlens array based system for large FOV imaging has also been studied [7]. Designing a computational imaging system that would increase the effective physical area of the native imaging sensor in a single shot can therefore offer several practical advantages in the operation of such scanning based imaging systems.

For imaging systems using spatially incoherent illumination, the forward model for the measured image intensity may be expressed as

$$I_{\text{out}}(x,y) = I_{\text{in}}(x,y) * p(x,y) + n(x,y), \tag{1}$$

where I_{out} and I_{in} represent the output and input object intensities, respectively, $p(x,y)$ is the incoherent point spread function (PSF) of the system and $n(x,y)$ is the noise in the measurement, which we assume to be additive. The symbol $*$ in the above relation represents the convolution operation. It is well known that $p(x,y) = |h(x,y)|^2$, where $h(x,y)$ represents the coherent impulse response of the system. An illustration to explain the FOV limitation posed by the limited sensor size is shown in figure 1. The PSF of a traditional imaging system is a single point impulse response (SPIR), approximated here as a delta impulse at the centre of image field as shown in figure 1(a). In a real system, the physical extent of the PSF is decided by the diffraction-limited spot size and is therefore inversely proportional to the Fourier plane aperture of the system. The whole image field is shown in figure 1(b); however, due to the limited detector size the image recorded by the system is shown in figure 1(c) where the dotted yellow square represents the active area of the detector. For example, it is a common observation while using a microscope that the image field seen through the viewfinder is larger than what is recorded on the digital camera attached to the microscope. This tells us that the image distribution often exists beyond the physical sensor boundaries, which is missed in the recording process. When the system PSF is SPIR in figure 1(a), there is essentially a pixel to pixel association between the object and the recorded image. As a result, the information existing outside the physical array sensor cannot be recorded. In this work we are interested in a computational imaging system configuration in which by suitably modifying the system PSF the information outside the physical extent of the detector can be brought into the detected intensity pattern $I_{\text{out}}(x,y)$. This may lead to a visually unrecognizable image record; however, an appropriate reconstruction algorithm applied to these data may lead to a system with an extended FOV. We wish to emphasize here that we are not simply adding more pixels to the existing image but wish to add pixels beyond the sensor boundary while maintaining the same spatial resolution. Our idea as proposed in this work has a close connection to structural illumination microscopy (SIM) [8]. SIM has become an important computational imaging modality which enables super-resolution imaging beyond the diffraction limit. This is achieved by using (typically) sinusoidal structured illumination in the object plane (figure 2(a)) which allows one to access spatial frequencies beyond the Fourier plane aperture of the microscope system, as illustrated in figure 2(b). Analogously, as seen in figure 2(c), it is conceivable that a structured phase pattern in the Fourier plane of the imaging system (which corresponds to some PSF in the coordinate domain) will allow us to access information beyond the physical dimension of the sensor (figure 2(d)). In this sense the system discussed in this paper may be considered as a conjugate domain analogue of a SIM system. The simplest PSF for accessing information

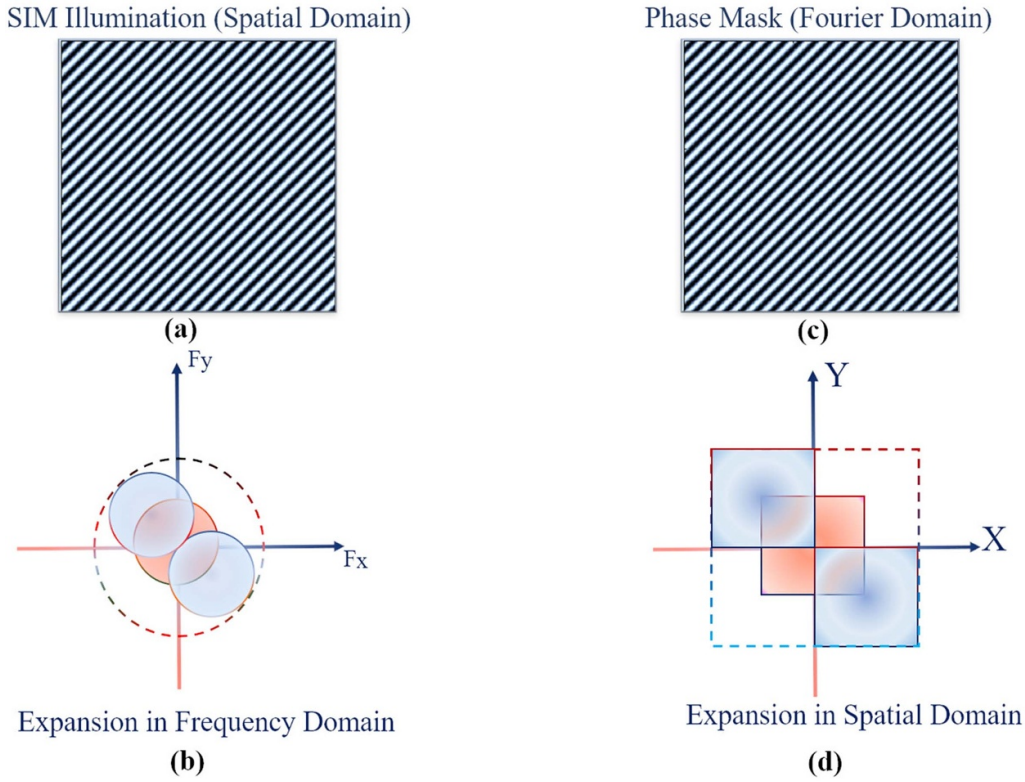

Figure 2. Analogy of the proposed extended FOV system with structure illumination microscopy (SIM). (a) In SIM, sinusoidal illumination is used in the spatial domain. (b) The extension of Fourier space occurs by shifting the spectrum from the centre (red circle), as shown by the small blue circles. By using multiple illumination directions, it is possible to fill the complete Fourier space within the dotted circle and hence extend the system resolution beyond the native diffraction limit. (c) Fourier plane phase mask in the proposed system. (d) The small red squares in the center represents the native array detector size of the traditional imaging system. Implementation of phase masking at the Fourier plane aperture of the system shifts the image FOV as shown by the blue squares. The proposed phase mask design corresponding to a multiple point impulse response fills the space shown by the dotted rectangle and therefore increases the effective sensor size.

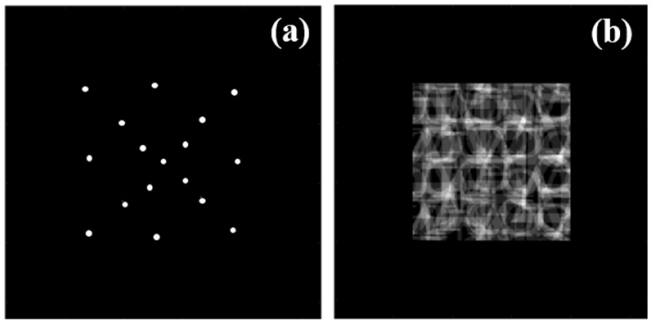

Figure 3. (a) Proposed PSF configuration with multiple point impulse response approximated as delta impulses and distributed in the central half FOV of the object field. (b) Image scrambled by the proposed PSF as recorded by a detector of limited size.

beyond the physical sensor boundary is the multiple point impulse response (MPIR), as illustrated in figure 3(a). It may be noted that the PSF now consists of multiple delta functions arranged so as to bring the missing information from outside the sensor boundary into the recorded data. When this PSF is used to image the object in figure 1(b), the resultant image becomes scrambled and captured by the limited-sized sensor as shown in figure 3(b). Although not clear visually, the information about the letters 'P' and 'E', which was completely missing in the truncated image in figure 1(c), has now been brought into the detector area in a scrambled form. In order to recover the full FOV image from this detected data, let us assume that the sensor size is $m \times n$ pixels and the full FOV that we wish to reconstruct is of size $M \times N$ with $M > m$ and $N > n$. The image reconstruction problem therefore involves MN unknown pixel values that are to be estimated from mn (possibly noisy) measurements. Based on the well-established theory of compressive imaging, this problem should be solvable if the resultant MN pixel image of the object of interest has sparse representation in some suitable transform basis. Imaging using MPIR was first suggested in [9] in the context of controlled blurring of images for encryption purposes. In order to increase the space–bandwidth product of the system, a random projection strategy was suggested in [10]. A number of random convolution based schemes have also been studied [11–13] for computational super-resolution imaging. While our focus in this work is on spatial incoherent illumination, the case of increase in the space–bandwidth product of the system via coded diffraction based phase retrieval has

also been demonstrated [14, 15]. In these studies, the gain in space–bandwidth product of the system is due to the improved resolution of the captured image. Spatially encoded PSFs in a multiplexed imaging system have been demonstrated in [16] to capture a wide FOV for spectral imaging of sparse objects. More recently, ideas on random convolution were discussed in [17], where a half masked detector was used to image in the masked area. While this work has some similarities to the main ideas proposed here, the quality of the reconstructed image in the masked detector region seems to be somewhat unsatisfactory. Moreover, when a large number of impulse response points are used in the setup, the detected pattern has a reduced contrast to noise ratio, and may lead to difficulties in extended FOV image reconstruction. In our work, we utilize a concept analogous to SIM for the distribution of impulse functions, which then reduces the number of impulses required. Based on this analogy with SIM imaging, we believe that a high degree of randomness in the location of the multiple delta functions in the impulse response is not necessary. As our simulations suggest, a small number of impulses (15–20) placed appropriately, as shown in figure 3(a), with only the intention of covering a larger FOV seem to be sufficient for obtaining an extended FOV with high image quality.

This paper is organized as follows. In section 2 we will provide details on the design of the multiple point PSF and additionally suggest methodologies that may be used to realize this PSF experimentally. We also provide details of a sparse iterative reconstruction algorithm that can achieve a good quality extended FOV image. A description of the algorithm is followed in section 3 by illustrations of extended FOV image reconstruction for a binary text object and a gray-scale object. In section 4, we show the performance of the proposed methodology with different levels of noise and object sparsity. Finally in section 5, we provide concluding remarks and future directions.

2. Extended FOV image reconstruction from scrambled and truncated data

2.1. Problem overview

In order to explain the feasibility of an extended FOV system, in this section we assume that the PSF of the proposed imaging system is in the idealized form and consists of a collection of delta impulses at locations (x_k, y_k) as given by

$$p(x,y) = \sum_{k=1}^{N} a_k \delta(x - x_k, y - y_k). \tag{2}$$

Here a_k is the weighting factor for the delta impulses, which is determined by the normalization of the PSF. In our example, all values of a_k are equal to the normalized constant. For the simulation, we use a 256×256 computational window and assume that the native detector has an active area over the central 128×128 pixels. The delta impulses in figure 3(a) are all placed within the detector extent. The finite size of the Fourier plane aperture in a real imaging system will imply that the individual delta spikes will have a finite extent equal to the diffraction limit. These multiple point impulse responses can be realized by using an appropriate computer-generated hologram (CGH) or phase mask corresponding to the PSF shown in figure 3(a) in the Fourier plane of the standard 4f system shown in figure 4. The design of the phase mask or appropriate CGH for multi-spot patterns can be realized using iterative phase retrieval algorithms [18–21]. The coherent impulse response of the system is proportional to the Fourier transform of the phase mask in the Fourier plane aperture of the imaging system. The phase mask in the Fourier plane may be displayed on a spatial light modulator (SLM) or fabricated on a transparent glass substrate using a lithographic process. The physical coordinates (u, v) of the SLM correspond to spatial frequencies (f_x, f_y), where

$$f_x = \frac{u}{\lambda f}, \quad f_y = \frac{v}{\lambda f}. \tag{3}$$

Here λ and f are the illumination wavelength and focal length of the lens used in the 4f system, respectively. The impulse function at position (x_0, y_0) can be considered to result from of a plane wave $\exp[i2\pi(f_x x_0 + f_y y_0)]$ in the phase mask, which has to be represented in a sampled form on the SLM pixels. The extent to which (x_0, y_0) can be placed away from the optical axis of the system thus depends on the Nyquist sampling requirements for the SLM pixels. If the physical pixel size in the SLM plane is $\Delta u = \Delta v$, the spatial frequency sampling interval is given by

$$\Delta f_x = \frac{\Delta u}{\lambda f}, \quad \Delta f_y = \frac{\Delta v}{\lambda f}. \tag{4}$$

For a plane wave represented as $W(f_x, f_y) = \exp[i2\pi(f_x x_0 + f_y y_0)]$, the Nyquist sampling criterion along the u-axis requires that

$$2\pi x_0 \, \Delta f_x \leq \pi. \tag{5}$$

This leads to the condition

$$x_0 \leq \frac{1}{2\Delta f_x} = \frac{\lambda f}{2\,\Delta u}. \tag{6}$$

Thus the maximum value of x_0 (or the extent of impulses) that could be possible with the 4f configuration is a function of wavelength, the focal length of the lens and the pixel pitch of the SLM. For a nominal tabletop 4f system (as in figure 4) with a lens with focal length $f = 5$ cm, SLM pixel size of 6.4 μm and wavelength $\lambda = 650$ nm, the largest allowed x_0 is approximately equal to 2.5 mm (note that x_0 can be either positive or negative). For a nominal sensor size of 5 mm, this amounts to doubling the effective sensor size if an effective reconstruction algorithm can be designed. The relation in equation (6) allows us to design the required extent of the multiple impulse response relative to a detector size by the use of suitable hardware components.

As explained above, our aim is to select a small number of delta impulses that allow us to cover the desired FOV beyond the physical sensor size. As explained in [9], a uniform periodic arrangement of the delta impulses may result in lines of zeros in the optical transfer function (OTF) of the system

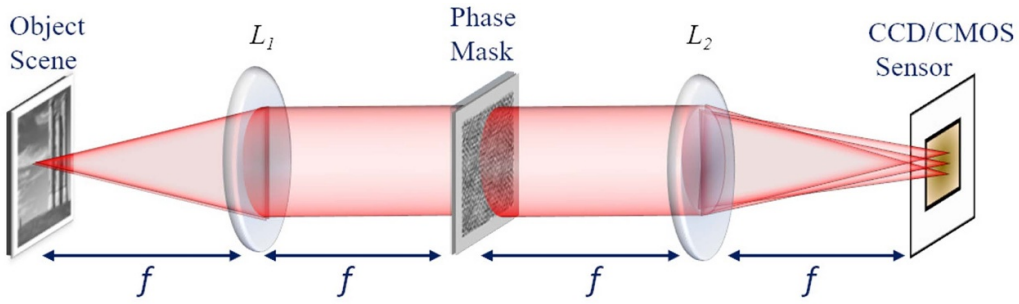

Figure 4. Schematic experimental setup of extended FOV imaging with a phase mask at the Fourier plane. The phase mask scrambles the object field recorded by the detector shown by a yellow square in the image plane. The larger blank square in the image plane is the expected total FOV.

and may eventually lead to artefacts in the recovered extended FOV image. As shown in figure 3(a), we therefore arrange the delta impulses symmetrically around the central delta impulse and then dither their positions randomly by 2–5 pixels. The nature of the data recorded on the truncated sensor is shown in figure 3(b). For completeness we state here that after the convolution with the multiple point impulse response, a Poisson noise equivalent to 10^5 photons/pixel on an average has been added in the data.

2.2. Description of the extended FOV algorithm

In order to reconstruct the extended FOV image $I_{\text{in}}(x,y)$ we aim to solve an optimization problem which may be nominally formulated as minimization of a cost function of the form

$$C(I_{\text{in}}) = C_1 + \alpha C_2 = ||S \odot [I_{\text{out}} - p * I_{\text{in}}]||^2 + \alpha\, H(I_{\text{in}}). \quad (7)$$

In the above cost function, the first term C_1 represents the L2-norm squared data consistency error. The symbol S denotes the truncated detector support and the symbol $\odot$ denotes that this error is computed only over S. The second term C_2 is a regularization term for which we use the modified Huber functions which may be defined as

$$H(I_{\text{in}}) = \sum_{\text{all pixels}} \left[\sqrt{1 + \frac{|\nabla I_{\text{in}}|^2}{\delta^2}} - 1 \right]. \quad (8)$$

The Huber function contains a parameter δ. This parameter controls the local behaviour of the sparsity function. For a large value of gradient (i.e. at the edges) the Huber function behaves like the total variation (TV) penalty which has an edge-preserving property. For a small gradient magnitude the Huber function is quadratic in gradient magnitude and acts like a smoothing penalty. Therefore it is suitable for the reconstruction of images having both edge and grey scale features. In the iterative algorithm when the Huber penalty is used, the parameter δ may be made proportional to the median of the image gradient magnitude at all pixels of the guess solution at any iteration. In this way the penalty function automatically decides whether a given pixel has edge-like features or gray-scale and operates on it in an adaptive manner.

The parameter α in equation (7) decides the weight between the two terms of the cost function. Determining an appropriate value of α usually requires empirical tuning. In order to avoid this somewhat tedious tuning process, we follow an alternating minimization scheme inspired by the adaptive steepest descent projection onto convex sets (ASD-POCS) algorithm, which has been demonstrated successfully in the context of computed tomography image reconstruction [22] as well as in digital holography [23–25]. The main idea in ASD-POCS is that, in one iteration, the change in the solution due to reduction of C_1 is made approximately equal to the change in the solution due to reduction of C_2 in an adaptive manner. This procedure is known to adjust the solution automatically (depending on the nature of the data), without the requirement to have any free parameter such as α in the algorithm. The details of the iteration are explained below.

Given the truncated data over $m \times n$ pixels, one may initiate the algorithm with an image consisting of $M \times N$ pixels over which we wish to recover the image with the truncated data in the centre. In our work we initiate the algorithm with the truncated data itself with zero padding to make the total guess image size equal to $M \times N$. The following steps are carried out in the $(k+1)$th iteration:

(a) Compute the intermediate solution by progressing the guess image in the direction of the negative gradient of the cost function C_1 as

$$I_{\text{int}}^{(k)} = I^{(k)} - t\,(\nabla_I C_1)_{I=I^{(k)}}. \quad (9)$$

Here, the step size t is calculated using the standard back-tracking line search method [26] to make sure that the numerical value of C_1 is reduced. The functional gradient $(\nabla_I C_1)$ is computed as

$$(\nabla_I C_1) = -2p_- * [S \odot (I_{\text{out}} - p * I)], \quad (10)$$

where $p_-(x,y) = p(-x,-y)$ is the PSF inverted along both x and y dimensions. Some extra constraints such as positivity can be incorporated after this step.

(b) The change in the solution due to this error reduction step is given by the distance

$$d_1 = ||I_{\text{int}}^{(k)} - I^{(k)}||_2. \quad (11)$$

The notation $||..||_2$ represents the L_2 norm.

(c) In order to impose the Huber penalty, the intermediate solution is updated in small steps in the direction of $-\nabla_I C_2$ with a step size proportional to d_1. This process may be described as

$$I_{\text{int}}^{(k,l+1)} = I_{\text{int}}^{(k,l)} - \beta d_1 \left[\frac{\nabla_I C_2}{||\nabla_I C_2||_2}\right]_{I=I_{\text{int}}^{(k,l)}}. \quad (12)$$

Here the initial guess for the sub-iteration $I_{\text{int}}^{(k,l=0)}$ is equal to the intermediate solution $I_{\text{int}}^{(k)}$. The parameter β is in $(0, 1)$ and we use an initial value of $\beta = 0.2$. A fixed number L of sub-iterations of this form were performed ($L = 20$ in our case) in each outer iteration. For completeness we mention that the functional gradient $\nabla_I C_2$ corresponding to the Huber penalty function may be evaluated as [27]

$$\nabla_I C_2 = -\frac{1}{\delta^2}\nabla \cdot \left[\frac{\nabla I}{\sqrt{1 + \frac{|\nabla I|^2}{\delta^2}}}\right]. \quad (13)$$

(d) The change in the intermediate solution due to the C_2 sub-iterations above is computed as

$$d_2 = ||I_{\text{int}}^{(k,L-1)} - I_{\text{int}}^{(k)}||_2. \quad (14)$$

(e) Our aim is to achieve $d_1 \approx d_2$ eventually and, as a result, if $d_1 < d_2$, the parameter β is reduced by a fixed fraction $\beta \to 0.95\beta$ for use in the next outer iteration. We wish to remark here that the algorithm is not very sensitive to the initial value of the parameter β and so its initial choice is not critical to the eventual solution.

(f) At this point, the $(k+1)$th iteration is complete and we set

$$I_{\text{in}}^{(k+1)} = I_{\text{int}}^{(k,L-1)}. \quad (15)$$

This alternating minimization procedure does not require a critical parameter such as α and in addition has an interesting feature that as the iterations progress the angle between the two directions $-\nabla_I C_1$ and $-\nabla_I C_2$ starts to increase and goes to a large obtuse value. This suggests an equilibrium between the two terms of the cost function and the iterative process may be stopped when this angle becomes sufficiently large (e.g. greater than 140°) or when change in the solution is less than a pre-determined threshold value. In order to accelerate the iterative process, we used the Nestrov accelerated gradient descent [28, 29] with backtracking line search for the reduction of the error term C_1. This involves computing the functional gradient $-\nabla_I C_1$ at the intermediate point $[I^{(k)} + \gamma_k(I^{(k)} - I^{(k-1)})]$ instead of at $I^{(k)}$ as in simple gradient descent. The coefficient γ_k for the added momentum term is changed as per the sequence [29]

$$\eta_{k+1}^2 = (1 - \eta_{k+1})\eta_k^2 \quad (16)$$

$$\gamma_k = \frac{\eta_k(1 - \eta_k)}{\eta_k^2 + \eta_{k+1}}. \quad (17)$$

The initial value of η_k for $k = 0$ is selected as 1 in the above sequence.

3. Extended FOV reconstruction results

In order to demonstrate the proposed concept of extended FOV imaging, simulation was carried out on two types of test object. One is a binary object which is sparse in the gradient domain and the other one is the standard 'cameraman' image with both edge-like and gray-scale features (and hence less sparse). Both the images contain 256×256 pixels. Multiple point impulse responses are distributed in the central 128×128 pixel computational window. Since in a practical system the impulse functions have a finite extent depending on the diffraction limit, we present a simulation study that includes both cases, one with the ideal delta impulses and another with the Gaussian impulse, as shown in figure 6(g). The finite size of Gaussian impulses arises due to finite size of the SLM, which may be used as a phase mask in the Fourier plane. For a square area of a SLM of dimensions $d \times d$ acting as a Fourier plane aperture, the incoherent PSF is proportional to a distribution $\text{sinc}^2(\frac{xd}{\lambda f})\,\text{sinc}^2(\frac{yd}{\lambda f})$ at the location of the individual impulse. Here λ and f represent the wavelength of illumination and the focal length of the lens in a 4f setup (see figure 4). Using the same parameters as before in section 2.1 ($\lambda = 650$ nm, $f = 5$ cm) and assuming $d = 1$ cm, it may be observed that the full width at half maximum (FWHM) of the $\text{sinc}^2(...)$ function is equal to a pixel width assuming an array sensor with square pixels of size 3 μm. For numerical illustration we approximate the $\text{sinc}^2(...)$ PSF with a Gaussian with the same FWHM as shown in figure 6(g).

To generate the data for simulations, the original object was convolved with the proposed PSF followed by truncation of the blurred image to the limited detector size. Further, we added Poisson noise in the blurred image corresponding a mean light level of 10^5 photons/pixel in the truncated detector pixels. Both illustrations used 1000 iterations of the algorithm, as explained in section 2.2. The algorithm was written in MATLAB/GNU Octave. Figures 5(a) and (b) show the reconstructed image with the proposed iterative approach with delta impulses and Gaussian impulses in the PSF definition, respectively. Figure 5(c) shows the relative mean squared error (RMSE) plot. The solid line represents the error when the PSF with delta impulses is used and the dotted line corresponds to the PSF with Gaussian impulses. The RMSE is calculated with respect to the ground truth test object as

$$\text{RMSE} = \frac{||I_{\text{ground truth}} - I_{\text{in}}||_2}{||I_{\text{ground truth}}||_2}. \quad (18)$$

$I_{\text{ground}-\text{truth}}$ and I_{in} represent the pixel values corresponding to the original and restored images, respectively. As expected, the RMSE value with the Gaussian impulse is somewhat higher than that with the delta impulses, since the Gaussian impulse leads to some additional blurring in the data. Other metrics for comparing the quality of the reconstructed image, such as the peak signal to noise ratio (PSNR) and the structure similarity index measure (SSIM), were also employed and their numerical values for various cases studied here are provided in table 1. These values suggest high quality recovery for the binary object with the same resolution as that of the ground truth

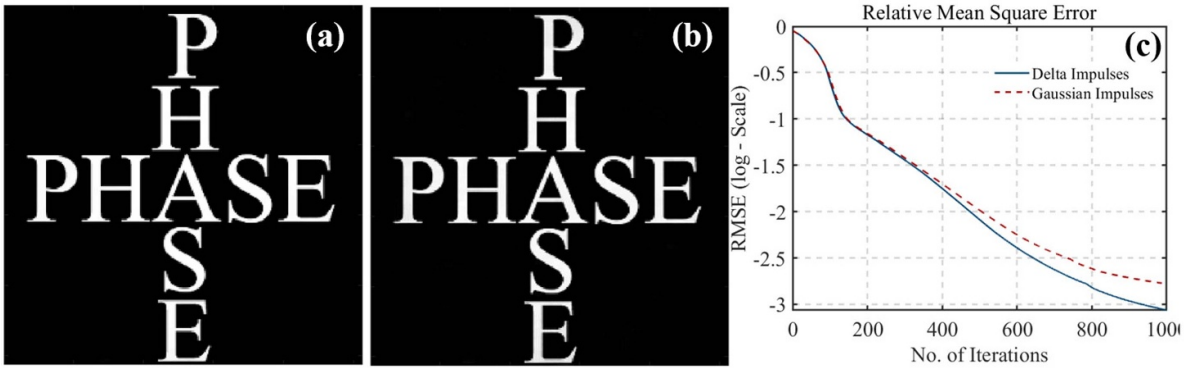

Figure 5. (a), (b) Reconstructed images from the scramble data in figure 3(b). (a) Recovered image using a PSF with delta impulses. (b) Recovered image using a PSF with Gaussian impulses. In both cases the information beyond the native sensor size, i.e. the letters '*P*' and '*E*', gets recovered completely. (c) Relative mean square error (RMSE) as a function of iteration numbers with the solid blue line and dotted red line representing the error (on log-scale) when a PSF with delta impulses and Gaussian impulses is used, respectively.

Table 1. The error performance of the reconstructed images with two different levels of sparsity as measured by the Tamura coefficient (TC) of the gradient magnitude image corresponding to the test objects.

		Delta impulses			With Gaussian impulses		
	TC	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
Binary object	1.84	8.6×10^{-4}	41.6	0.95	1.2×10^{-3}	40.0	0.92
Cameraman 128×128 sensor size	1.34	2.9×10^{-2}	30.2	0.84	4.0×10^{-2}	29.1	0.82
Cameraman 150×150 sensor size	1.34	1.6×10^{-2}	32.9	0.88	2.1×10^{-2}	32.0	0.87

image (figure 1(b)). Moreover, reconstruction of the image, and hence the effective increase in sensor size, is a function of the sparsity of the object. To be more specific, we use the Tamura coefficient (TC), a measure of sparsity, in this work to evaluate the sparsity of the test object [30]. The TC evaluates the gradient sparsity of the image $I_{\text{in}}(x, y)$ and is defined as

$$\text{TC} = \sqrt{\frac{\sigma_G}{\langle G \rangle}}. \tag{19}$$

Here σ_G and $\langle G \rangle$ represent the standard deviation and mean of the gradient magnitude image $G = |\nabla I_{\text{in}}|$, respectively. A higher numerical value of TC suggests higher gradient sparsity of the test image. More details on the accuracy of image recovery as a function of image sparsity will be provided in the next section. Thus, the gray scale object such as a cameraman is less sparse than the binary object. Therefore, it is expected that the number of samples required for the reconstruction is larger in this case. We reconstruct the cameraman image with two sensor sizes: (a) 128×128 pixels, which corresponds to sampling gain of 4, and (b) 150×150 pixels, which corresponds to an approximate sampling gain of a factor of about 3. The two truncated detector sizes are shown by dotted yellow and white squares in figure 6(b). Figure 6(c) shows the data collected at the sensor using the delta impulses as the PSF, with the dotted yellow square showing a sensor size of 128×128 pixels and the white dotted square showing a sensor size of 150×150 pixels. The reconstructed images with two different sensor sizes are shown in figures 6(d) and (e), respectively. The reconstructed images now show an extended FOV beyond the truncated detector boundaries. However, the image quality with 150×150 pixels is better than with 128×128 pixels, as expected. The same numerical experiment is repeated with the Gaussian impulses in figure 6(f). The individual Gaussian impulse is shown in the inset figure 6(g). Visual inspection of the recovered extended FOV images in figures 6(i) and (j) are observed to be similar to those with a PSF of ideal delta impulses. However, the quality of the image gets degraded, as confirmed by the numerical values of the metrics in table 1 and the error plots in figure 6(k). The blue and red curves in this plot represent the cases corresponding to the 128×128 and 150×150 pixel sensors, respectively. The solid and dotted curves correspond to the PSF with ideal delta impulses and Gaussian impulses, respectively. Along with the extended FOV these results also suggest another possibility for increasing the frame rate of existing sensors. For example, a multiple point impulse response may be used to code information captured by an existing sensor into a smaller area. A smaller number of pixels may then be read out, thereby improving the frame rate of the existing sensor without losing a number of pixels in the final image.

Some remarks regarding difference in quality of the extended FOV image recovery for the binary text object and the cameraman object are in order here. We wish to highlight again that our aim in this work is to recover an MN pixel image from its mn noisy measurements (with $M > m$ and $N > n$). For a given sampling gain of MN/mn the quality of recovery is expected to be better for objects that have higher sparsity

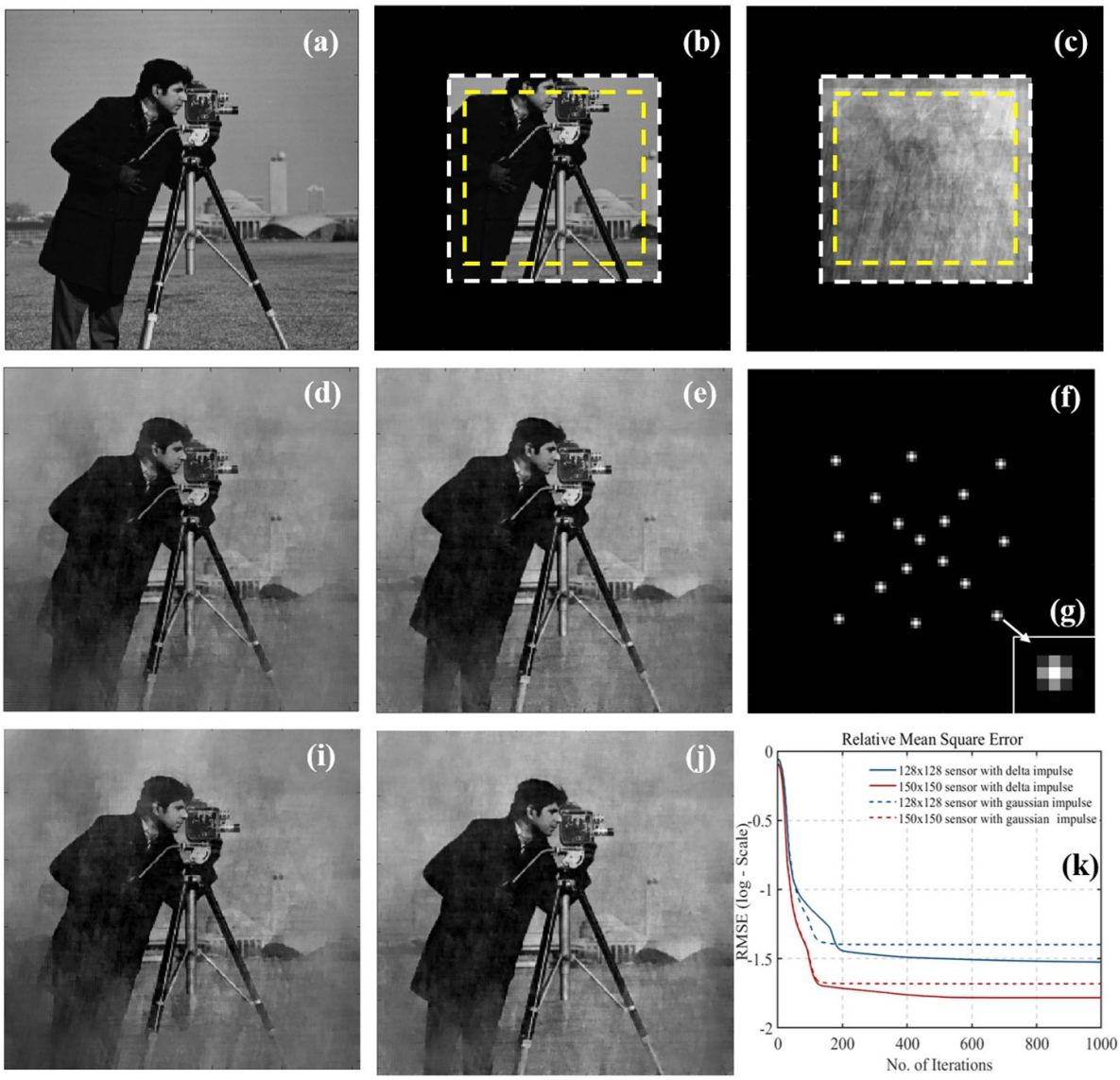

Figure 6. Extended FOV simulation results for a grey-scale object. (a) Ground truth image with 256 × 256 pixels. (b) Image captured with the native sensor of size 128 × 128 pixels is shown by a dotted yellow square and that with a detector size of 150 × 150 pixels is shown by a white dotted square. The information outside the respective sensor sizes is lost in the traditional measurement process. (c) Captured data with the proposed PSF, with yellow and white dotted squares representing the respective sensor sizes. (d), (i) Restored images with the 128 × 128 pixel sensor using a PSF with delta and Gaussian impulses, respectively. (e), (j) Restored images with the 150 × 150 pixel sensor using a PSF with delta and Gaussian impulses, respectively, (f) PSF with Gaussian impulses with the zoomed in version of the individual impulse shown in (g). (k) The relative mean square error plot for the two sensor sizes is shown by the solid lines with blue and red curves, corresponding to the 128 × 128 pixel sensor and 150 × 150 pixel sensor, respectively, when a PSF with delta impulses is used. The respective error plots for the Gaussian impulses are shown by the dotted lines.

in a suitable transform basis. Among the two objects used in our illustrations, the binary text object has higher degree of sparsity (TC = 1.84) and therefore a near perfect recovery of the extended FOV is possible. On the other hand the cameraman object is not as sparse (TC = 1.34) since it has a number of sharp and grey-scale features as well as textured regions. As a result the quality of recovery as measured by the parameters in table 1 is comparatively lower for the same sampling gain. This observation is in line with what is expected from the compressive sensing formalism [31] and is important for implementing practical extended FOV imaging systems. Note that, here, TC is used as a measure of sparsity as it is simple to evaluate. Other measures of sparsity may be used [32] if needed.

4. Effect of noise and object sparsity on extended FOV reconstruction

In this section, we examine the effect of noise in the truncated detector data and the sparsity of the object under consideration on the proposed reconstruction methodology. In order to understand the effect of noise, we simulated the truncated detector data (as illustrated in figure 3(b)) with Poisson

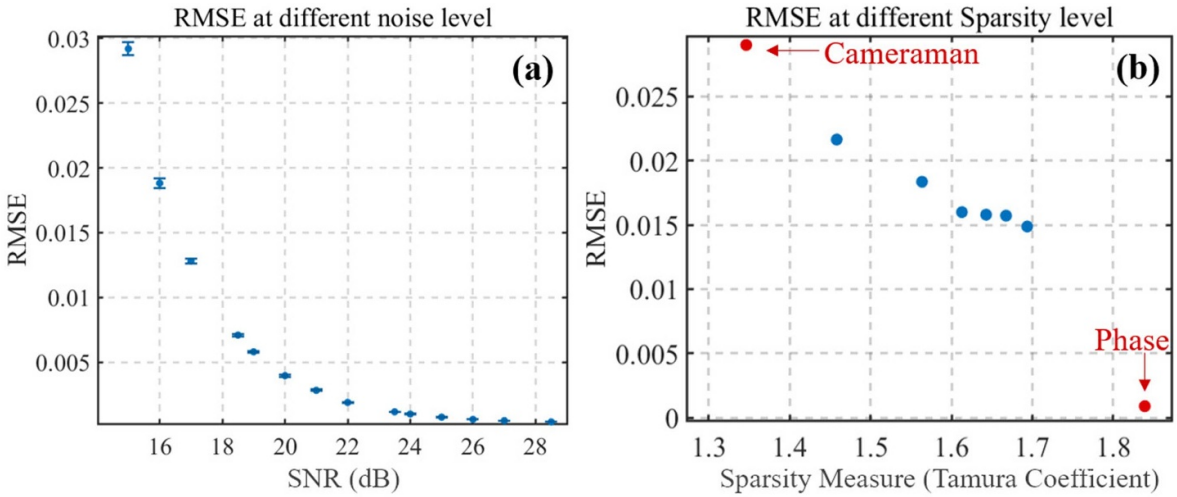

Figure 7. RMSE performance of the restored images after 1000 iterations. (a) RMSE as a function of different SNR levels (in dB). (b) RMSE as a function of different sparsity level for the test object as measured by the Tamura coefficient.

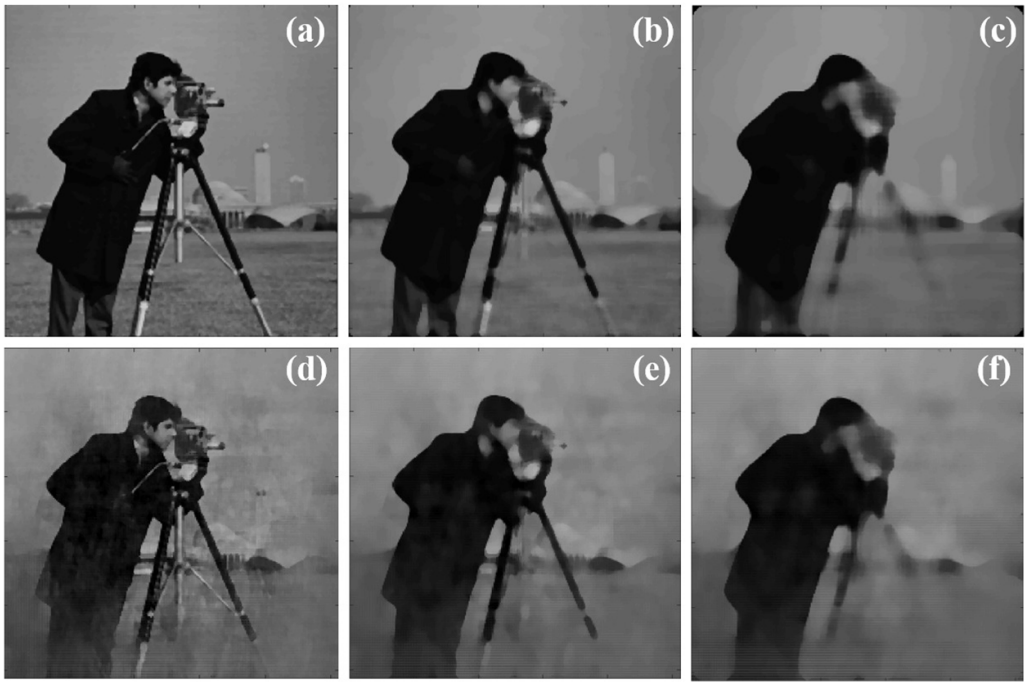

Figure 8. Cameraman image with increase in the level of sparsity. (a)–(c) Ground truth images with median filters of sizes $3 \times 3, 7 \times 7$ and 13×13 with TC values of 1.46, 1.61 and 1.7, respectively. (d)–(f) Restored images after 1000 iterations. For the reconstruction of images, the extent of MPIR and the sensor size lie within the central 128×128 pixels window.

noise corresponding to an average light level in the range 5×10^5 photons/pixel. A total of 14 light levels in this range were used so that the signal to noise ratio (SNR) in dB ranged from 15 to 28. The RMSE of the reconstructed extended FOV image with respect to the ground truth as a function of detection SNR (dB) is plotted in figure 7(a). The error bars in this plot represent the standard deviation in RMSE over five independent noise realizations at each of the SNR levels. As expected, the RMSE is seen to fall off with increasing detection SNR. Analysis of the behaviour of the reconstructions with varying levels of object sparsity is a more involved problem. For a given level of sparsity, the quality of extended FOV reconstruction may depend on the distribution of features in the test image that are within and outside the truncated detector region. In order to keep the analysis simple, we artificially constructed images with increasing levels of sparsity by using a median filter of varying size applied to the camera image. As the size of the filter increases, finer textural details in the image are washed out, which then increases the image sparsity. The sparsity of these median filtered images was once again measured by means of the TC corresponding to their gradient magnitude image as described in the previous section. The RMSE for the six reconstructed test images is plotted in figure 7(b), corresponding to median filter of sizes varying from 3×3 to

13×13, respectively. In the previous section we used two test objects—the binary text PHASE object and the grayscale cameraman object. In terms of the TC sparsity measure, these two cases lie beyond the range of sparsity values covered by the six test objects studied here. Three of these test objects correspond to median filters of sizes $(3 \times 3), (7 \times 7)$ and (13×13) and are shown in figures 8(a)–(c) along with the extended FOV reconstructed image in figures 8(d)–(f), respectively. The sensor size used for the reconstruction of these images consists of 128×128 pixels and the PSF of the system was the ideal delta impulses. This corresponds to a sampling gain of 4. A noise level corresponding to 25 dB SNR was used in all six cases. In general we observe that the extended FOV reconstruction performance improves with increasing sparsity. From the RMSE plot in figure 7(b), we further observe that the RMSE of the restored images may also depend on the nature of the test object and does not have a simple relationship with a single sparsity measure such as TC (or probably any other choice of sparsity measure [32]). For example, the RMSE for the binary PHASE object with the highest sparsity does not necessarily fit the trend of RMSE performance for various cameraman test objects (generated by median filtering). The RMSE for the PHASE object is seen to be much lower than for the six cameraman test objects, thus indicating that the feature distribution in the test object also plays an important role. The topic of performance of the compressive sensing algorithms with object sparsity was studied earlier in [33] for Gaussian random measurements. However, in the context of imaging problems such as those discussed here that use sparse reconstructions, this topic needs more investigation that is not within the scope of the present article.

5. Conclusion

We proposed a computational optical imaging system for increasing the FOV beyond the physical extent of the image sensor. Our method involves engineering the PSF of the traditional imaging system such that it consists of a number of impulses separated from each other spatially. The design of the proposed PSF is analogous to structured illumination imaging. The structured pattern in the Fourier plane of the imaging system allows us to extend the FOV in the detector space. Our numerical experiments indicate effective extended FOV image recovery with a two-fold larger FOV without the loss of image resolution for realistic experimental situations. The extent to which the effective size of the sensor may be increased depends on the smallest realizable feature size in the Fourier plane phase mask, and the number and distribution of impulses in the engineered PSF. The sparsity of the object under consideration is also important, and the quality of the recovered image for a given sampling gain is expected to be better for sparser objects. The exact sampling gain that will provide a satisfactory extended FOV reconstruction will, however, depend on the nature of the object or specific application, as seen in our illustrations. The concept described here may also be utilized to improve the frame rate of existing image sensors without the loss of pixel number in the recovered image. The proposed extended FOV system concept is likely to find multiple applications which we will explore with experiments in future.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

R M acknowledges fellowship support from The Prime Minister's Research Fellowship. K K wishes to acknowledge partial support from DST India (Award No. 2016/MED/34/2016).

Appendix A. Algorithm pseudocode

Pseudo code for extended FOV

Inputs: I_{out} (data image with zero padding), t (step size parameter in main loop), β (step size parameter for penalty), L (number of iterations for the penalty loop), p (multiple point impulse response of the imaging system (MPIR)), S (FOV support: a binary mask equal to the sensor size), *maxiter* (maximum number of iterations in the main loop (1000))
Initialization:
$\beta \leftarrow 0.2$,
$L \leftarrow 20$,
$k \leftarrow 0$,
$I^{(k=0)} \leftarrow I_{out}$
$\eta \leftarrow 1$,
$\gamma \leftarrow 0$
for $k = 0$ to *maxiter* **do**
    $I_{prev} = I^{(k)}$
    $t \leftarrow 0.01||I^{(k)}||_2$
    $C_1 \leftarrow ||S \odot [I_{out} - I^{(k)} * p]||_2^2$
    Nestrov Acceleration Loop with Backtracking Line search
    Initialization: *flag* $\leftarrow 1$
    $\Delta I \leftarrow (I^{(k)} - I^{(k-1)})$ Momentum term, set to zero for $k = 0$.
    while *flag* AND $t > 10^{-30}$ **do**
        $I_{next} \leftarrow I^{(k)} + \gamma \Delta I$ [**Nestrov Update**]
        $dC_{1\,next} \leftarrow \nabla_{Inext} C_1$ [as given in equation (10).]
        $dC_{1\,next} \leftarrow \frac{dC_{1\,next}}{||dC_{1\,next}||_2}$
        $I_{try} \leftarrow I_{next} - t\,dC_{1\,next}$
        $C_{1\,try} \leftarrow ||S \odot [I_{out} - I_{try} * p]||_2^2$
        if $C_{1\,try} < C_1$ **then**
            flag $= 0$,
        else
            $t = t/2$
        end if
    end while
    Update η [as given in equation (16)]
    Update γ [as given in equation (17)]
    $I_{int}^{(k)} \leftarrow I_{try}$ [intermediate solution]
    Enforce positivity constraint on $I_{int}^{(k)}$,
    Calculate $d_1 \leftarrow ||I_{int}^{(k)} - I_{prev}||_2$
    $I_{int}^{(k,l=0)} \leftarrow I_{int}^{(k)}$
    Steepest descent loop for enforcing Sparsity
    for $l = 0$ to $(L-1)$ **do**
        $dC_2 \leftarrow [\nabla_I C_2]_{I=I_{int}^{(k,l)}}$ [as given in equation (13)]
        $I_{int}^{(k,l+1)} \leftarrow I_{int}^{(k,l)} - \beta d_1 \left[\frac{dC_2}{||dC_2||_2}\right]_{I_{int}^{(k,l)}}$
    end for
    $I^{(k+1)} \leftarrow I_{int}^{(k,L)}$ [next solution update]
    $d_2 \leftarrow ||I_{int}^{(k)} - I^{(k+1)}||_2$
    if $d_2 > d_1$ **then**
        $\beta \leftarrow 0.95\beta$
    end if
end for
Output: $I^{(k=maxiter)}$.

ORCID iDs

Ravikrishnan Elangovan https://orcid.org/0000-0002-0068-7944
Kedar Khare https://orcid.org/0000-0002-3104-3850

References

[1] Brady D J, Gehm M E, Stack R A, Marks D L, Kittle D S, Golish D R, Vera E M and Feller S D 2012 *Nature* **486** 386–9
[2] Gilbertson J R, Ho J, Anthony L, Jukic D M, Yagi Y and Parwani A V 2006 *BMC Clin. Pathol.* **6** 4
[3] Lu Q, Liu G, Xiao C, Hu C, Zhang S, Xu R X, Chu K, Xu Q and Smith Z J 2018 *PLoS One* **13** e0194063
[4] Greenbaum A, Luo W, Su T W, Göröcs Z, Xue L, Isikman S O, Coskun A F, Mudanyali O and Ozcan A 2012 *Nat. Methods* **9** 889–95
[5] Zheng G, Horstmeyer R and Yang C 2013 *Nat. Photon.* **7** 739–45
[6] Zheng G, Ou X and Yang C 2014 *Biomed. Opt. Express* **5** 1–8
[7] Orth A and Crozier K 2012 *Opt. Express* **20** 13522–31
[8] Gao P, Ma Y, Wen K, Liu M, Zheng J, Chu K, Smith Z J and Liu L 2021 *J. Phys. Photon.* **3** 024009
[9] Stossel B J and George N 1995 *Opt. Commun.* **121** 156–65
[10] Stern A and Javidi B 2007 *J. Display Technol.* **3** 315–20
[11] Romberg J 2009 *SIAM J. Imaging Sci.* **2** 1098–128
[12] Marcia R F and Willett R M 2008 *IEEE Int. Conf. on Acoustics, Speech and Signal Processing* vol 2 pp 833–6
[13] Thapa D, Raahemifar K and Lakshminarayanan V 2015 *J. Mod. Opt.* **62** 415–29
[14] Shi B, Lian Q and Chang H 2020 *Signal Process.* **168** 107350
[15] Katkovnik V, Shevkunov I, Petrov N V and Egiazarian K 2017 *Optica* **4** 786–94
[16] Horisaki R and Tanida J 2010 *Opt. Express* **18** 23041–53
[17] Marcos D, Lasser T, López A and Bourquard A 2016 *Opt. Express* **24** 1269–90
[18] Gerchberg R and Saxton W 1972 *Optik* **35** 237–46
[19] Ripoll O, Kettunen V and Herzig H P 2004 *Opt. Eng.* **43** 2549–2548
[20] Waller E H and von Freymann G 2013 *Opt. Express* **21** 21708–13
[21] Matsumoto N, Inoue T, Ando T, Takiguchi Y, Ohtake Y and Toyoda H 2012 *Opt. Lett.* **37** 3135–7
[22] Sidky E Y and Pan X 2008 *Phys. Med. Biol.* **53** 4777–807
[23] Singh M and Khare K 2017 *J. Opt. Soc. Am.* A **34** 349–55
[24] Singh M and Khare K 2018 *J. Mod. Opt.* **65** 1127–34
[25] Mangal J, Monga R, Sandeep R, Dinda A K, Joseph J, Ahlawat S and Khare K 2019 *J. Biophoton.* **12** e201800409
[26] Armijo L 1966 *Pacific J. Math.* **16** 1–3
[27] Gaur C, Mohan B and Khare K 2015 *J. Opt. Soc. Am.* A **32** 1922–7
[28] Nesterov Y E 1983 *Sov. Math. Dokl.* **27** 372–6
[29] O'Donoghue B and Candés E 2015 *Found. Comput. Math.* **15** 715–32
[30] Zhang Y, Wang H, Wu Y, Tamamitsu M and Ozcan A 2017 *Opt. Lett.* **42** 3824–7
[31] Candes E J, Romberg J K and Tao T 2006 *Commun. Pure Appl. Math.* **59** 1207–23
[32] Hurley N and Rickard S 2009 *IEEE Trans. Inf. Theor.* **55** 4723–41
[33] Donoho D and Tanner J 2009 *Phil. Trans. R. Soc.* A **367** 4273–93