Recent progress on multi-domain optimization for ultra-thin cameras

Amit Ashoka\textsuperscript{a} and Mark A. Neifeld\textsuperscript{a,b}
\textsuperscript{a}Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, AZ USA 85721
\textsuperscript{b}College of Optical Sciences, University of Arizona, Tucson, AZ USA 85721

ABSTRACT

Conventional imaging systems can suffer from significant aliasing and/or blur distortions when the detector array in the focal plane under-samples the image. We propose to address this problem by engineering the optical PSF of the imaging system followed by electronic post-processing to minimize the overall distortions. The optical PSF of the candidate imaging system is modified by placing a phase-mask in its aperture-stop. We consider a particular parameterization of the phase-mask and optimize its parameters to minimize the distortions. We obtain as much as 30% improvement in the final imaging quality with the optimized optical PSF imager (SPEL) relative to the conventional imager.

Keywords: PSF engineering, aliasing, under-sampling, multiple images, imaging system, sub-pixel, phase-mask.

1. INTRODUCTION

Designing a high-resolution optical imaging system usually leads to an impulse-like optical point spread function (PSF). This is justified as the resolution of a conventional optical imager is seen to be limited by its optical PSF [1]. However, when such an optical system is followed by a semiconductor detector array, the resulting detector blur and/or aliasing can become the dominant distortion present in the overall imager. In such cases it is necessary to redesign the optical imaging system to minimize such distortions. Under these conditions the optical PSF can be seen as a method of encoding the object prior to the measurement in order to reduce the distortions introduced by the detector array. In such a framework the impulse-like PSF can be sub-optimal [2].

The distortion introduced by the detector array is illustrated in Fig. 1. Consider a point-source object. Fig. 1(a) shows the resulting image formed on the detector array. Fig. 1(b) shows the image of the same object if it was in slightly different position. For a conventional imaging system with an impulse-like optical PSF, the resulting image measurement for the two cases would be identical, if the object positions differed by less than the pixel/detector spacing. This clearly implies that the overall imager resolution is detector limited. Now consider the case where the optical PSF of the imaging system extends over several pixels. Fig. 1(c) and 1(d) show the identical case as in Fig. 1(a) and 1(b) except now for an extended PSF. In this case, the position of the point-source object can be computed to sub-pixel accuracy by simple correlation-based processing, given a high enough SNR. This example suggests that an extended PSF offers better resolution as compared to an impulse-like PSF.

We consider the optical PSF as a means of encoding the object information prior to measurement. Hence, the resulting composite imaging system can be seen as “computational” as it involves optical (encoding) elements as well as electronic post-processing (decoding) to form the final image [3]. The example in Fig. 1 suggests that although encoding the object with an extended PSF could improve the overall resolution, it may be accompanied by an SNR cost. Ref. 2 quantifies the resolution performance of such an imaging system using an alternate PSF. In this paper we will focus on quantifying the imager performance in terms of the image quality.

![Fig. 1. Example of conventional [(a) and (b)] and alternate [(c) and (d)] optical PSF.](image-url)
2. IMAGING SYSTEM MODEL

Consider a conventional imaging system with a detector array in its focal plane. The spatial bandwidth of the imaging optics determines the detector spacing of the detector array according to the Nyquist sampling criteria. Assuming diffraction-limited performance, the optical PSF of the imaging system is given as \( \text{sinc}^2(\Delta x) \), where \( \Delta \) is the optical resolution (Rayleigh) [1]. A resolution of \( \Delta \) implies that the detector spacing should be \( \Delta/2 \) according to the Nyquist criteria. If the detector spacing of the array is \( d \), then we define an under-sampling factor \( F = 2d/\Delta \). The detector array under-samples the image formed in the focal plane by a factor of \( F \geq 1 \). In order to overcome aliasing, the most popular approach is to use multiple images/frames, each obtained by a sub-pixel shift of the imager, and subsequently combine them in a post-processing step to obtain the image with reduced aliasing [4]. In principle, this approach is similar to the lenslet-array imager, which acquires multiple sub-pixel shifted images by employing an array of small identical lenses in an integrated aperture. This imager is generally referred to as the TOMBO [5]. We will use the TOMBO imager as the baseline conventional imager for comparison.

We use the following 1D linear operator formalism to model the imaging system:

\[
g = Hf + n \tag{1}
\]

where \( g \) (Mx1) is the image measurement vector, \( f \) (Nx1) is the object vector (continuous object sampled at desired resolution) and \( n \) (Mx1) is additive white Gaussian noise (AWGN). \( H \) (MxN) represents the linear system operator (optics+detector array) and is in general blockwise shift-invariant for under-sampled systems. The linear minimum-mean-squared-error (LMMSE) decoder \( W \) is used to reconstruct the object from image measurement and is expressed as:

\[
z = Wg \tag{2}
\]

This decoder minimizes the mean-square-error (MSE) defined as, \( E\{|z-f|^2\} \) where \( E\{.\} \) denotes statistical expectation taken over both object and noise. The LMMSE decoder \( W \) is defined as follows:

\[
W = R_fH^T(HR_fH^T + R_n)^{-1} \tag{3}
\]

Where \( R_f \) and \( R_n \) are the autocorrelation matrix for the input object vector and noise vector respectively. We use the power-law model for power spectral density \( (S_f(\nu) = 1/\nu^{\eta}) \); \( \eta=1.4 \) of the objects to obtain the object autocorrelation matrix [6]. The root-mean-square-error (RMSE=\( \sqrt{\text{MSE}} \)) metric (computed over a class of objects), is used to quantify the image quality performance of an imaging system.

![Fig. 2. SPEL imaging system layout.](image)

As proposed in the previous section, the overall performance of an imager can be improved by modifying the imaging optics to obtain a suitable extended optical PSF. In this paper, we consider a class of extended optical PSFs which can be obtained by placing a phase-mask in the aperture-stop of a conventional imaging system as shown in Fig. 2.
We consider a particular parameterization of the phase-mask comprising several sinusoidal gratings. Here the phase \( \phi(x) \), of the mask is expressed as follows:

\[
\phi(x) = \sum_{k=1}^{J} \alpha_k \sin(\beta_k x - \gamma_k)
\]  

(4)

Here \( J \) is the number of gratings, \( \alpha_k \) is the amplitude, \( \beta_k \) is the frequency and \( \gamma_k \) is the phase offset of the \( k^{th} \) grating element. Each grating element is defined by three parameters resulting in total \( 3J \) parameters that define the phase mask. This phase mask augmented imager will be referred to as a sinusoidal phase enhanced lens (SPEL) imager.

3. SIMULATION

To quantify the relative performance of the SPEL and TOMBO imagers we conduct a simulation study. The detector array used in the simulation has a pixel spacing \( d=7.5\mu m \), with a 100% fill factor and a full well capacity (FWC) of 45,000 photo-electrons. We assume that the imaging lens has diffraction-limited performance and operates at \( F/# =1.8 \). This choice of optics and detector spacing results in an under-sampling factor \( F=16 \). The measurement SNR \( (10\log_{10}(\text{FWC}):\text{shot-noise limited}) \) is 46.5dB for the given FWC. The number of frames/image measurements is denoted by \( K \). To achieve an overall resolution equal to the optical resolution, the number of frames required is \( K=F \) for a conventional imaging system. However, with the extended PSF we strive to achieve nearly equivalent performance for \( K<F \). In this simulation study we distribute the total number of photons equally between each of the \( K \) frames. The total photon budget is set to \( 16\times \text{FWC} \) so that when \( K=16 \), each frame has the measurement SNR of 46.5dB corresponding to the FWC. Note that as \( K \) increases the measurement SNR in each frame decreases. For example, when \( K=1 \) the measurement SNR=58.5dB and for \( K=16 \) the SNR=46.5dB. Each frame is acquired by shifting the imager by a sub-pixel amount. The optimal choice of the sub-pixel shifts is: \( \{ \delta_L=Lxd/K : L=(0,1,\ldots,K-1) \} \) [7]. We select the sub-pixel shifts according to this method.

The solid curve in Fig. 3 shows reconstruction RMSE for the conventional imaging system (TOMBO) for increasing values of \( K \). The main observation here is that the RMSE decreases with increasing values \( K \). This is expected as the number of image measurements increases the LMMSE reconstruction reduces the aliasing distortion resulting in a smaller RMSE. In case of the SPEL imager, the parameters of the phase-mask are optimized to minimize the reconstruction RMSE for each value of \( K \). Thus for \( K=1 \) we might obtain a different optimized optical PSF compared to...
the one obtained for say $K=4$. Note that we use the same optimized optical PSF for each sub-pixel shifted image measurement for a given value of $K$. The reconstruction RMSE for the SPEL imager is plotted with dotted line in the Fig. 3. We observe the same trend of decreasing RMSE with increasing $K$, as we did for the TOMBO imager. However, the main observation here is that the RMSE for the SPEL imager is consistently lower than that for TOMBO, especially for small values of $K$. In fact, a RMSE performance of 3% is achieved at $K=3$ for SPEL imager as compared to $K=5$ for TOMBO imager. It is clear that the SPEL imager achieves a significant advantage over TOMBO imager in terms of image quality as measured by the RMSE metric. Note that for high values of $K$, SPEL and TOMBO imager performance is similar as the optimal optical PSF of SPEL approaches the conventional lens PSF. This is expected as this particular choice of optical PSF, in case of SPEL imager, minimizes the SNR cost of extended PSF, which is crucial for relatively lower measurement SNR when value of $K$ is high.

4. CONCLUSIONS

We have shown that for imaging systems with inherent aliasing/blur distortions due to under-sampling detector arrays, engineering the optical PSF to minimize these distortions improves the final image quality. The proposed SPEL imager uses a phase-mask element in the aperture-stop to obtain the modified optical PSF. It is shown that by optimizing the optical PSF of the SPEL imager, it out-performed the conventional imaging system in terms of image quality, as measured with the RMSE metric, by as much as 30% for $K=1$. In conclusion, we have demonstrated the viability of the approach of engineering the optical PSF of an imager to overcome distortions, aliasing/blur in this case, with suitable post-processing.

REFERENCES