Performance Limits for Computational Photography

Kaushik Mitra[2], Oliver Cossairt[1], Ashok Veeraraghavan[2]

[1]Northwestern University, 2133 Sheridan Road, Evanston, IL, USA 60208

[2]Rice University, 6100 Main Street, Houston, TX, USA 77005

1 Introduction

Over the last decade, a number of Computational Imaging (CI) systems have been proposed for tasks such as motion deblurring, defocus deblurring and multispectral imaging. These techniques increase the amount of light reaching the sensor via multiplexing and then undo the deleterious effects of multiplexing by appropriate reconstruction algorithms. However, a detailed analysis of CI has proven to be a challenging problem because performance depends equally on three components: (1) the optical multiplexing, (2) the noise characteristics of the sensor, and (3) the reconstruction algorithm, which typically uses signal priors. In this paper, we utilize a recently proposed framework incorporating all three components [13]. We model signal priors using a Gaussian Mixture Model (GMM), which allows us to analytically compute Minimum Mean-Squared Error (MMSE). We analyze the specific problem of motion and defocus deblurring, showing how to find the optimal exposure time and aperture setting for defocus and motion deblurring cameras, respectively. This framework gives us the machinery to answer an open question in computational imaging: "To deblur or denoise?"

1.1 Analysis of Computational Imaging Systems

A number of camera designs have been proposed in recent years that capture different aspects of visual appearance using multiplexed measurements. Examples include defocus deblurring cameras [5,8,9,18], motion deblurring cameras [11,14], multi/hyperspectral multiplexing [7,16], and illumination multiplexing [15]. These systems use optical coding (multiplexing) to increase light throughput, which increases the SNR of captured images. The desired signal is then recovered computationally via signal processing. The quality of recovered images depends jointly on the conditioning of the optical coding and the increased light throughput.

In this work, we follow a line of research whose goal is to relate maximum performance to practical considerations (e.g. illumination conditions and sensor characteristics). We follow the convention adopted by Cossairt et al. [3] and Mitra et al. [13]. We define a conventional camera as an *impulse imaging system*, which measures the desired signal directly (e.g. without blur). CI performance is then compared against the impulse imaging system. Noise is related to the lighting level, scene properties and sensor characteristics. In this paper, we analyse defocus and motion deblurring cameras. These cameras capture blurry images, and all-focused images are then recovered computationally via deconvolution. We consider a pillbox shaped blur function for defocus blur, and a 1-D rect function for motion blur. The impulse imaging counterpart for defocus blur is a narrow aperture image. For motion blur, the impulse imaging counterpart is a short exposure image. CI techniques capture more light, but they require deconvolution, which amplifies noise. Impulse imaging doesn't require deconvolution, but captures less light. There is a parameter for both motion (exposure time) and defocus deblurring (aperture size) that can be adjusted to trade-off light gathering power and deconvolution noise. We address the problem of how to optimize this parameter to achieve the best possible performance *with signal priors taken into account*.

In this paper we analyse the special case of uncoded defocus and motion blur that has not been optically manipulated in any way to improve invertibility. However, the framework used in this paper can also be used to analyse any multiplexed imaging system. For example, defocus deblurring systems have been devised to encode defocus blur using attenuation masks [9,18], refractive masks [5], or motion [8]. In addition, motion deblurring CI systems have been devised to encode motion blur using a fluttered shutter [14] or camera motion [11]. Further analysis of these systems can be found in [3,13]. The same optimization framework we use here can also be applied to these systems.

1.1 The Importance of Signal Priors

It is well understood that multiplexing gives the greatest advantage at low light levels (where *signal-independent* read noise dominates), but this advantage diminishes with increasing light (where *signal-dependent* photon noise dominates) [7]. However, it is impractical to study the effects of multiplexing alone, since signal priors are at the heart of every state-of-the-art reconstruction algorithm (e.g. BM3D [4], GMM [12,13]). Signal priors can dramatically increase performance in problems of deblurring (multiplexed sensing) and denoising (no multiplexing). However, it has historically been very difficult to determine exactly how much of an increase in performance to expect from signal priors, making it difficult to provide a fair comparison between different cameras.

We characterize the performance of CI systems under a GMM prior which has two unique properties: Firstly, GMM satisfies the universal approximation property which says that any probability density function can be approximated to any fi-

delity using a GMM with an appropriate number of mixtures [13]. Secondly, a GMM prior lends itself to analytical tractability so that we can use MMSE as a metric to characterize the performance of both impulse and CI systems.

2 Related Work

Several recent works have analyzed the performance of multiplexed acquisition techniques, including the seminal work by Harwit and Sloan [7], as well as more recent techniques that incorporate signal-dependent noise into the analysis [1,6,15,17]. Recently, Cossairt et al. [3] have analyzed CI systems taking into consideration the application (e.g. defocus deblurring or motion deblurring), lighting condition (e.g. moonlit night or sunny day), scene properties (e.g. albedo, object velocity) and sensor characteristics (size of pixels). However, all the above works, do not analyze the performance of CI systems when a signal prior is used for demultiplexing. A few recent papers have analyzed the fundamental performance limits of image denoising in the presence of image priors [2,10]. A similar approach was used by Mitra et al. [13] to extend this analysis to general framework for analyzing computational imaging systems. We have adopted this framework for analyzing motion and defocus deblurring cameras. We analyse only single image CI techniques. Multi-image capture techniques have been analyzed by Hasinoff et al. [6] (defocus deblurring), and Zhang et al. [17] (motion blurring).

3 Imaging Model

We use assume a linear imaging model that takes into account signal-dependent noise and models signals using a GMM prior.

3.1 Image Formation Model

We consider linear multiplexed imaging systems that can be represented as

$$y = Hx + n \, , \tag{1}$$

where y is the measurement vector, $x \in \mathbb{R}^N$ is the unknown signal we want to capture, H is the $N \times N$ multiplexing matrix and n is the observation noise. In this paper, we analyze motion and defocus deblurring systems which produce shift-invariant blur. For the case of 1D motion blur, the vectors x and y represent a

scanline in a sharp and blurred image patch, respectively. The multiplexing matrix H is a Toeplitz matrix where the rows contain the system point spread function. For the case of 2D defocus blur, the vectors x and y represent lexicographically reordered image patches, and the multiplexing matrix H is block Toeplitz.

3.1 Noise Model

To enable tractable analysis, we use an affine noise model [1,6,15,17]. We model signal independent noise as a Gaussian random variable with variance σ_r^2. Signal dependent photon noise is Poisson distributed with parameter λ equal to the number of photons collected at a pixel (J). We approximate photon noise by a Gaussian distribution with mean and variance λ. This is a good approximation when λ is greater than 10. We also drop the pixel-wise dependence of photon noise and instead assume that the noise variance at every pixel is equal to the average signal intensity. We relate the number of photons collected by an *impulse imaging system[1]* to scene and sensor characteristics using the equation [3,13]:

$$J \approx 10^{15} \cdot (F/\#)^2 \cdot q \cdot R \cdot t \cdot \delta^2 , \qquad \textbf{(2)}$$

Which depends on the scene illumination level I_{src} (in lux), the average scene reflectivity (R) and the camera parameters such as the f-number ($F/\#$), exposure time (t), sensor quantum efficiency (q), and pixel size (δ). For our experiments, we assume a reflectivity of $R = .5$, quantum efficiency $q = .5$, aperture setting $F/20$, exposure time $t = 1/50$ secs, and pixel size $\delta = 5$ microns.

3.1 Signal Prior Model

We learn GMM patch priors from a large collection of about *50* million training patches. For learning we use a variant of the Expectation Maximization approach to ascertain the model parameters. We also test that the learned model is an adequate approximation of the real image prior by performing rigorous statistical analysis and comparing performance of the learned prior with state of the art image denoising methods [4]. For defocus deblurring, we learn a GMM patch prior, of patch size 16×16, with 1770 Gaussian mixtures. For motion deblurring, we learn a GMM patch prior of patch size 4×256, with 1900 Gaussian mixtures. We

---

[1] The increase in light collection for CI systems is encoded in the multiplexing matrix H.

use the analytic lower bound for MMSE derived in [13] to compute the optimal parameter setting for motion and defocus deblurring.

4 Computational Imaging Examples

The optimal parameters for motion deblurring (exposure time) and defocus deblurring (aperture setting) obviously depend on the scene light level. Hence, we compute the optimal parameters as a function of light level.

4.1 Motion Deblurring Cameras

We first set the exposure time of the impulse imaging systemso that the motion blur for the desired range of velocities is less than a pixel. We then analytically compute the expected increase in SNR for different exposure settings (PSF kernel lengths) with respect to the impulse imaging system (of PSF kernel length 1) at various light levels, see Fig. 1(a). For light levels less than 150 lux capturing the image with a longer exposure time and then deblurring is a better option, whereas for light levels greater than 150 *lux* we should capture the impulse image and then denoise. Fig. 1(b) shows the optimal blur PSF length at different light levels. At a light level of 1 *lux*, the optimal PSF length is 23 (corresponding to an exposure time of 460 milliseconds), whereas for light levels greater than or equal to 150 lux the optimal length is 1, i.e., the impulse image setting. Figs. 1(c-e) show the simulated results with different PSF kernel lengths at a few lighting levels.

4.2 Defocus Deblurring Cameras

We first fix the aperture size of the impulse imaging system so that the defocus deblur is less than one pixel. We then analytically compute the increase in SNR for a larger aperture setting (and PSF kernel size) at various light levels (see Fig. 2(a)), again computed with respect to impulse imaging. For light levels less than 400 *lux*, capturing the image with a larger aperture and then deblurring is a better option, whereas for light levels greater than 400 *lux*, we should capture the impulse image and then denoise. Fig. 1(b) shows the optimal blur PSF size at different light levels. At a light level of 1 *lux*, the optimal PSF size is 9×9 (corresponding to an aperture setting of $F/2.2$), whereas, for light levels greater than 400 *lux* the optimal size is 1×1, i.e., the impulse image setting. Figs. 1(c-d) show the simulated results with different PSF sizes and lighting levels.
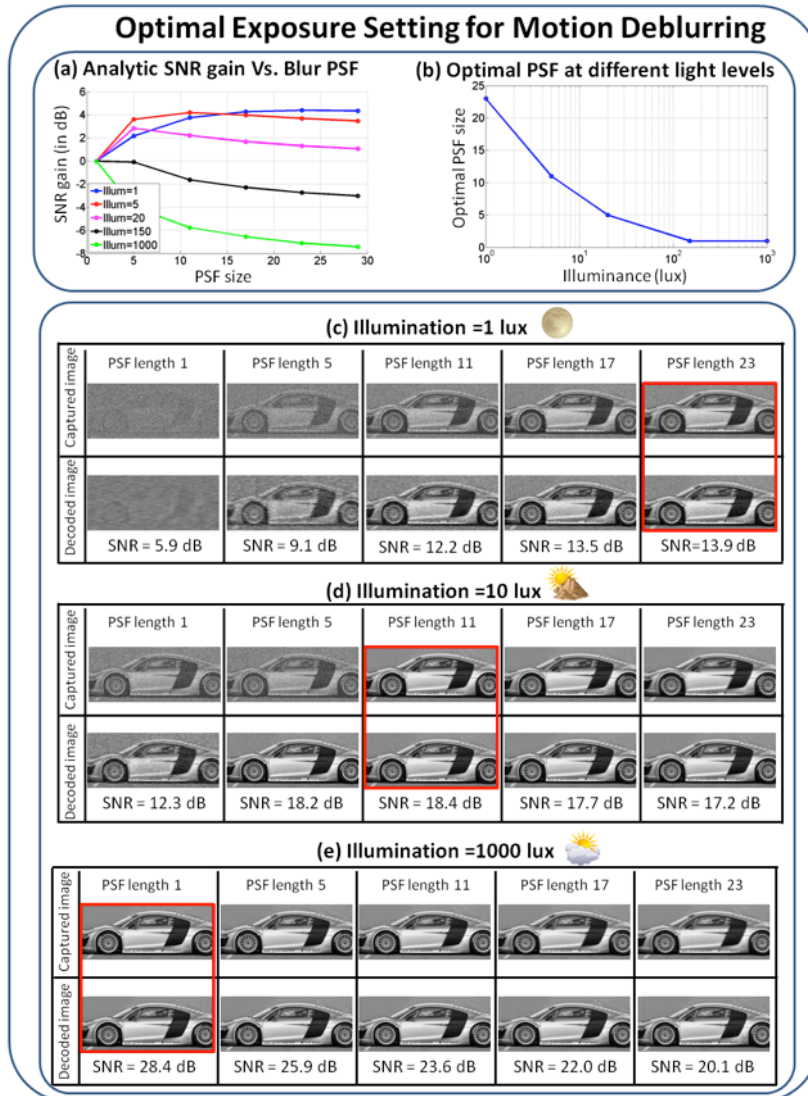
Fig. 1. Optimal exposure setting for motion deblurring: In subplot (a) we analytically compute the expected increase in SNR for different exposure settings (PSF kernel lengths), with respect to the impulse imaging system (of PSF kernel length 1), at various light levels. The impulse imaging system has an exposure time of 20 milliseconds. Subplot (b) shows the optimal blur PSF length at different light levels. At light level of 1 *lux*, the optimal PSF size is 23 (corresponding to an exposure time of 460 milliseconds), whereas for light levels greater than or equal to 150 *lux* the optimal size is 1, i.e. the impulse image setting. Subplots (c-e) show the simulated results with different PSF kernel lengths at a few different lighting levels.
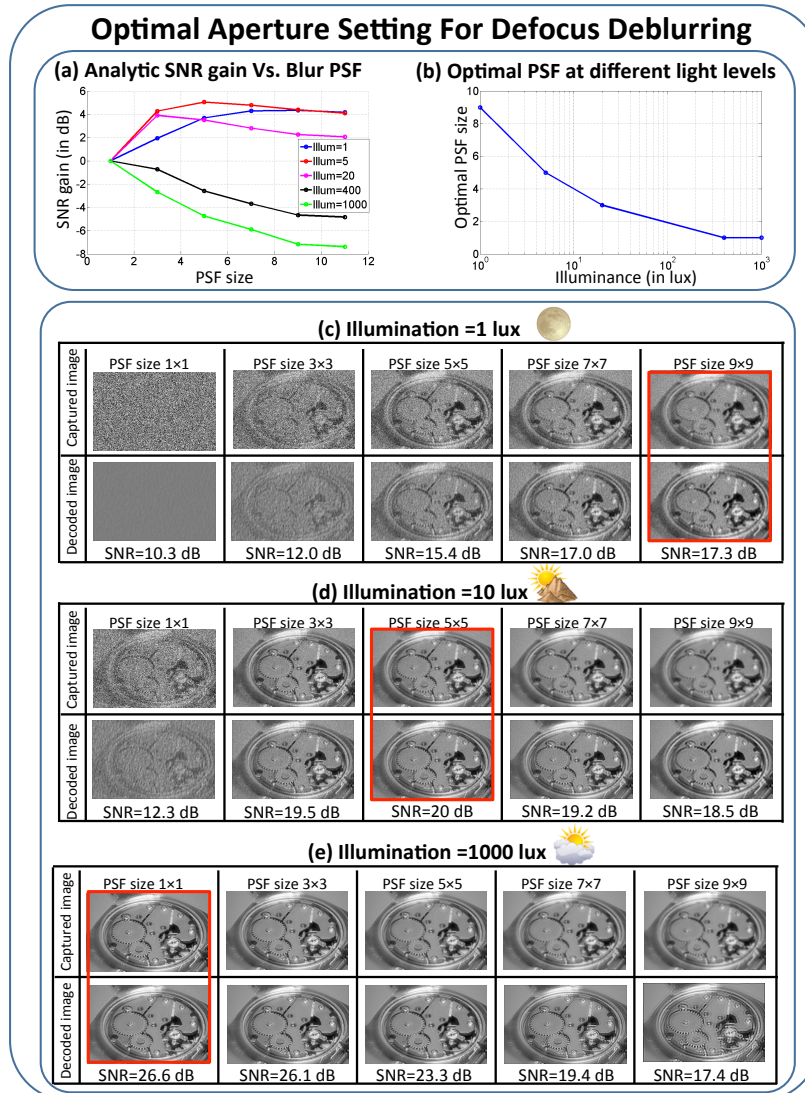
Fig. 2. Optimal exposure setting for defocus deblurring: In subplot (a) we analytically compute the increase in SNR for different aperture settings (PSF kernel size), with respect to the impulse imaging system of PSF kernel size 1×1 (corresponding to an aperture setting of $F/20$), for various light levels. In subplot (b) we show the optimal blur PSF size at different light levels. At light level of 1 *lux*, the optimal PSF is 9×9 (corresponding to an aperture setting of $F/2.2$, whereas for light levels greater than 400 *lux* the optimal is 1×1, i.e. the impulse image setting. Subplots (c-d) show the simulated results with different PSF size at a few different lighting levels.

5 Conclusion

We have analysed the problem of parameter optimization for motion and defocus deblurring cameras, answering the question: "To deblur or denoise?". We use a GMM signal prior and compute performance for both CI (deblurring) and impulse imaging (denoising) techniques. We relate performance to illumination and sensor characteristics. We showed that for a typical camera specification, denoising is preferrable when illumination is greater than 400 *lux* for motion deblurring, and 150 *lux* for defocus deblurring. In addition, we optimized parameters (exposure time for motion blur, aperture size for defocus blur) for darker illumination.

6 References

1. Agarwal and R. Raskar. Optimal single image capture for motion deblurring. In *CVPR*, 2009.
2. P. Chatterjee and P. Milanfar. Is denoising dead? *Image Processing, IEEE Transactions on*, 19(4):895–911, 2010.
3. O. Cossairt, M. Gupta, and S. K. Nayar. When does computational imaging improve performance? *IEEE transactions on image processing*, 22(1-2):447–458, 2013.
4. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *TIP*, 16(8), 2007.
5. E. Dowski Jr and W. Cathey. Extended depth of field through wave- front coding. *Applied Optics*, 34(11):1859–1866, 1995.
6. S. Hasinoff, K. Kutulakos, F. Durand, and W. Freeman. Time- constrained photography. In *ICCV*, pages 1–8, 2009.
7. M. Harwit and N. Sloane. Hadamard transform optics. *New York: Academic Press*, 1979.
8. S. Kuthirummal, H. Nagahara, C. Zhou, and S. K. Nayar. Flexible Depth of Field Photography. In *PAMI*, 2010.
9. A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. In *SIGGRAPH*. ACM, 2007.
10. A. Levin, B. Nadler, F. Durand, and W. T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. In *ECCV (5)*, pages 73–86, 2012.
11. A. Levin, P. Sand, T. Cho, F. Durand, and W. Freeman. Motion- invariant photography. In *SIGGRAPH*, 2008. C. Liang, T. Lin, B. Wong, C. Liu, and H. Chen. Programmable aperture photography: multiplexed light field acquisition. In *SIGGRAPH*, 2008.
12. K. Mitra and A. Veeraraghavan. Light field denoising, light field superresolution and stereo camera based refocusing using a gmm light field patch prior. In *CVPR Workshops*, 2012.
13. K. Mitra, A. Veeraraghavan, and O. Cossairt. Performance Analysis of Computational Imaging Systems and its Practical Implications. *submitted to PAMI*, May 2013.
14. R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH*, 2006.
15. N. Ratner and Y. Schechner. Illumination multiplexing within fundamental limits. In *CVPR*, 2007.
16. A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008.
17. L. Zhang, A. Deshpande, and X. Chen. Denoising versus Deblurring: HDR techniques using moving cameras. In *CVPR*, 2010.
18. C. Zhou and S. Nayar. What are Good Apertures for Defocus Deblurring, in ICCP 2009.