Optimal single image capture for motion deblurring

Citation: Agrawal, A., and R. Raskar. "Optimal single image capture for motion deblurring." Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. 2009. 2560-2567. ©2009 Institute of Electrical and Electronics Engineers.

As Published: http://dx.doi.org/10.1109/CVPRW.2009.5206546

Publisher: Institute of Electrical and Electronics Engineers

Persistent URL: http://hdl.handle.net/1721.1/53555

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Optimal Single Image Capture for Motion Deblurring

Amit Agrawal
Mitsubishi Electric Research Labs (MERL)
201 Broadway, Cambridge, MA, USA
agrawal@merl.com

Ramesh Raskar
MIT Media Lab
20 Ames St., Cambridge, MA, USA
http://raskar.info

Abstract

Deblurring images of moving objects captured from a traditional camera is an ill-posed problem due to the loss of high spatial frequencies in the captured images. Recent techniques have attempted to engineer the motion point spread function (PSF) by either making it invertible [16] using coded exposure, or invariant to motion [13] by moving the camera in a specific fashion.

We address the problem of optimal single image capture strategy for best deblurring performance. We formulate the problem of optimal capture as maximizing the signal to noise ratio (SNR) of the deconvolved image given a scene light level. As the exposure time increases, the sensor integrates more light, thereby increasing the SNR of the captured signal. However, for moving objects, larger exposure time also results in more blur and hence more deconvolution noise. We compare the following three single image capture strategies: (a) traditional camera, (b) coded exposure camera, and (c) motion invariant photography, as well as the best exposure time for capture by analyzing the rate of increase of deconvolution noise with exposure time. We analyze which strategy is optimal for known/unknown motion direction and speed and investigate how the performance degrades for other cases. We present real experimental results by simulating the above capture strategies using a high speed video camera.

1. Introduction

Consider the problem of capturing a sharp image of a moving object. If the exposure time can be made sufficiently small, a sharp image can be obtained. However, small exposure time integrates less amount of light, thereby increasing the noise in the captured image. As the exposure time increases, the SNR of the captured signal improves, but moving objects also result in increased motion blur. Motion deblurring attempts to obtain a sharp image by deconvolution, thereby resulting in increased deconvolution noise with exposure.

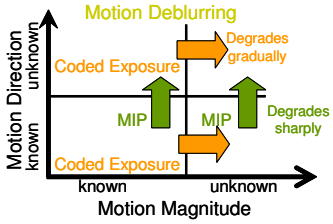
In this paper, we ask the following question: What is the best exposure time and capture strategy for capturing a single image of a moving object? We formulate the problem of optimal capture as follows: Maximize the SNR of the deconvolved image of the moving object, given a certain scene light level, while not degrading the image corresponding to the static parts of the scene[1].

To obtain the best deblurring performance, one needs to analyze the *rate of increase* of capture SNR versus deconvolution noise with the exposure time. For imaging sensors, the capture SNR increases proportional to the square root of the exposure time (sub-linear) due to the signal-dependent photon noise. It is well-known that deblurring of images obtained from a traditional camera is highly ill-posed, due to the loss of high spatial frequencies in the captured image. We first show a simple but rather non-intuitive result: the deconvolution noise for 1-D motion blur using a (static) traditional camera increases faster than capture SNR with the exposure time. Thus, increasing exposure time always *decreases* the SNR of the deconvolved moving object. We then analyze recent advances in engineering the motion PSF that dramatically improves the deconvolution performance. Two prominent methods are (a) making the PSF *invertible* using a coded exposure camera [16], and (b) making the PSF *invariant* by moving the camera with non-zero acceleration [13].

A **coded exposure camera** [16] modulates the integration pattern of light by opening and closing the shutter within the exposure time using a carefully chosen pseudo-random code. The code is chosen so as to minimize the deconvolution noise assuming a specific amount of motion blur in the image. However, coded exposure also loses light. In [16], the chosen code was 50% on/off, thereby losing half the light compared to a traditional camera with the same exposure time. While [16] analyzed the improvement in deconvolution performance, it ignored the loss of light in the image capture. We incorporate the loss of light in our analysis, and show that it is not necessary to have a 50% on/off code with signal-dependent noise; one does have the flexibility of choosing other codes. Note that the PSF is made invertible for *any* object motion direction, while mo-

[1] Otherwise a trivial capture strategy would be to move the camera with the *same* speed as the object if the motion direction is known.

978-1-4244-3991-1/09/$25.00 ©2009 IEEE



Single Image Capture	Traditional	MIP [13]	Coded Exposure [16]
Camera Motion Required?	No	Yes	No
Loss of Light	No	No	Yes
Invariant PSF	No	Yes[§]	No
PSF Invertibility	Very Bad	Good[§]	Good[‡]
Noise on Static Scene Parts (for same light level)	No	Deconvolution Noise	No

[§] Only for object motion direction same as camera motion direction & object motion magnitude in a range
[‡] For any object motion direction

Figure 1. Overview of single image capture techniques for motion deblurring. Coded exposure is optimal for deblurring for *any* motion direction, if the motion magnitude is known; but motion PSF needs to be estimated for deblurring. MIP is optimal if the motion direction is known and magnitude is within a range (could be unknown), with additional advantage that motion PSF need not be estimated (invariant). However, performance of coded exposure degrades gradually as motion magnitude differs from the desired one, while MIP performance degrades sharply as motion direction differs from camera motion direction and motion magnitude goes beyond the assumed range.

tion magnitude is required for optimal choice of code.

Motion invariant photography (MIP) [13] moves the camera with a constant acceleration while capturing the image. The key idea is to make the motion PSF *invariant* to object speed within a certain range. Thus, objects moving with different speeds within that range would result in same motion PSF. Note that MIP needs to know the *direction* of the object motion, since the camera should be moved accordingly, but knowledge of motion magnitude is not required. Another disadvantage is that the static parts of the scene are also blurred during capture, leading to deconvolution noise on those scene parts. We compare the three techniques in terms of SNR of the deconvolved image and obtain optimal parameters given a scene light level and object velocity (or range of velocities). Given capture parameters for a scenario, we investigate how the performance degrades for different motion magnitudes and directions. An overview is shown in Figure 1.

1.1. Contributions

- We formulate the problem of optimal single image capture of a moving object as maximizing the SNR of the deconvolved image of the moving object.
- We show that for a traditional image capture using a static camera, SNR of the deblurred moving object decreases with the increase in exposure time.
- We investigate which capture strategy to choose, choice of exposure time and associated parameters and analyze its performance for different operating conditions such as known/unknown motion magnitude and direction.

1.2. Related work

Motion deblurring has been an active area of research over last few decades. Blind deconvolution [9, 5] attempts to estimate the PSF from the given image itself. Since deblurring is typically ill-posed, regularization algorithms such as Richardson-Lucy [14, 18] are used to reduce noise. Recent papers [22, 20, 11, 10, 2] have shown promising results for PSF estimation and deblurring.

Manipulating PSF: By coding the exposure, [16] made PSF invertible and easy to solve. Wavefront coding [3] modifies the defocus blur to become depth-independent using a cubic phase plate with lens, while Nagahara *et al.* [15] move the sensor in the lateral direction during image capture to achieve the same. MIP [13] makes motion PSF invariant for a range of speeds by moving the camera. Coded apertures [4] has been used in astronomy using MURA codes for low deconvolution noise, using broadband codes for digital refocusing [21] and for depth estimation in [8, 12].

Improved capture strategies: In [6], optimal exposures were obtained to combine images for high dynamic range imaging. Hadamard multiplexing was used in [19] to increase the capture SNR in presence of multiple light sources. The effect of photon noise and saturation was further included in [17] to obtain better codes.

2. Optimal single image capture

Consider an object moving with a velocity v m/sec. For simplicity, we assume that the object is moving horizontally in a single plane parallel to the image plane and the object motion results in an image blur of v_i pixels/ms. Let i denote the captured blurred image at exposure time of 1 ms and $\bar{i_o}$ and $\bar{i_b}$ be the average image intensity of the moving object and the static background in the captured image respectively. Define a *baseline exposure time* $t_0 = \frac{1}{v_i}$ for which the blur is 1 pixel in the capture image. Let m be the size of the object in the image along the motion direction if it was static and let $\mathtt{SNR}_0$ be the minimum acceptable SNR for the object.

Image noise model: We use the affine noise model [17, 1, 7], where the noise η is described as the sum of a signal independent term and a signal dependent term. The signal-independent term is due to dark current, amplifier noise and the A/D quantizer. Let the gray level variance of this term be σ^2_{gray}. Signal-dependent noise is related to photon flux and the uncertainty of the electron-photon conversion process. The variance of the photon generated electrons linearly increases with the measured signal, and hence with

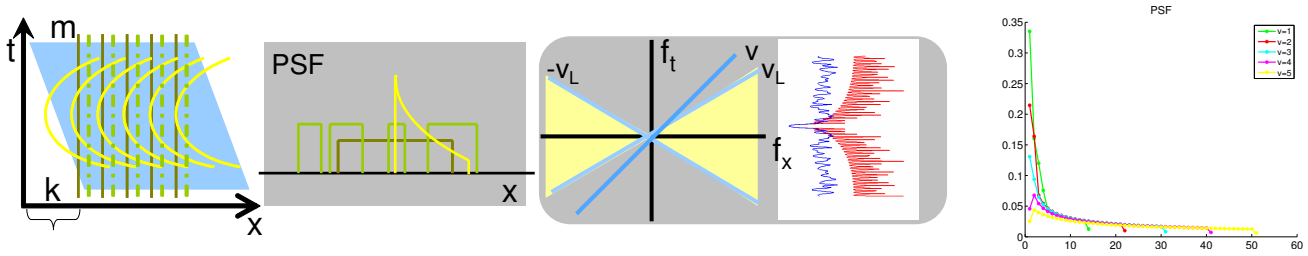


Figure 2. Comparison of capture strategies. (Left) A 1D object (blue) of length m blurs by k in x-t space with integration lines corresponding to traditional camera (solid brown), coded exposure (dotted) and MIP (yellow) and the resulting PSF's. Objects moving with speed v have energy along a single line in frequency domain f_x-f_t space. For traditional & coded exposure (static cameras), the captured image corresponds to the $f_t = 0$ slice after modulation by sinc (red) and broadband (blue) filter respectively. Thus, for coded exposure, any velocity v results in *non-zero* energy on $f_t = 0$ slice for all spatial frequencies. MIP optimally captures energy within the wedge given by $[-v_r, v_r]$ [13] but performs poorly for v outside this range. (Right) Motion PSF for MIP becomes similar to box function as speed increases beyond the desired range ($v_r = 3$).

the exposure time t. Thus, the photon noise variance can be written as Ct, where C is a camera dependent constant. Thus, $\sigma_\eta^2 = \sigma_{gray}^2 + Ct$. Given this noise model, the SNR of the captured image is given by

$$\mathrm{SNR}_{capture} = \frac{\bar{i_o} t}{\sqrt{\sigma_{gray}^2 + Ct}}. \quad (1)$$

For long exposures, $Ct \gg \sigma_{gray}^2$, the photon noise dominates and $\mathrm{SNR}_{capture} \approx \frac{\bar{i_o}\sqrt{t}}{\sqrt{C}}$ increases as the square root of the exposure time. When $Ct \ll \sigma_{gray}^2$, $\mathrm{SNR}_{capture}$ increases linearly with t.

Deconvolution noise: At exposure time t, the amount of blur $k = tv_i$. The captured image $i(x, y)$ is modeled as a convolution of the sharp image of the object $s(x, y)$ with the motion PSF $h(x)$, along with added noise

$$i(x, y) = s(x, y) * h(x) + \eta(x, y), \quad (2)$$

where $*$ denotes convolution. For 1D motion, the discrete equation for each motion line is given by $\mathbf{i} = A\mathbf{s} + \mathbf{n}$, where $A_{(m+k-1)\times m}$ denotes the 1D circulant motion smear matrix, and $\mathbf{s}$, $\mathbf{i}$ and $\mathbf{n}$ denote the vector of sharp object, blurred object and noise intensities along each motion line. The estimated deblurred image is then given by

$$\widehat{\mathbf{s}} = (A^T A)^{-1} A^T \mathbf{i} = \mathbf{s} + (A^T A)^{-1} A^T \mathbf{n}. \quad (3)$$

The covariance matrix of the noise in the estimate $\widehat{\mathbf{s}} - \mathbf{s}$ is equal to

$$\Sigma = (A^T A)^{-1} A^T \sigma_\eta^2 A (A^T A)^{-T} = \sigma_\eta^2 (A^T A)^{-1}. \quad (4)$$

The root mean square error (RMSE) increases by a factor $f = \sqrt{\mathtt{trace}(A^T A)^{-1}/m}$. Thus, the SNR[2] of the deconvolved object at exposure time t is given by

$$\mathrm{SNR}_d = \frac{\bar{i_o} t}{f\sqrt{\sigma_{gray}^2 + Ct}}, \quad (5)$$

where f denotes the deconvolution noise factor (DNF).

[2]We use $20\log_{10}(.)$ for decibels.

2.1. Traditional camera

For a traditional capture, motion PSF is a box function whose width is equal to the blur size k

$$h(x) = 1/k \quad \text{if } 0 < x < k, \quad 0 \quad \text{o.w.} \quad (6)$$

Figure 3 (left) show the plots of $\frac{\sqrt{t}}{f}$ which is proportional to SNR_d at high signal dependent noise ($Ct \gg \sigma_{gray}^2$). Plots are shown for different object velocities assuming $m = 300$. Note that the SNR decreases as exposure time is increased. Thus, for traditional capture, increasing the exposure time decreases the SNR of the deconvolved object. For a specific camera, the minimum exposure time that satisfies $\mathrm{SNR}_d > \mathrm{SNR}_0$ would be optimal, if this condition could be satisfied.

Trivial capture: If the SNR at baseline exposure t_0 is greater than SNR_0, then the optimal exposure time is t_0. For example, if there is enough light in the scene (bright daylight), a short exposure image will capture a sharp image of a moving object with good SNR.

3. Reducing deconvolution noise

Now we consider the following two approaches: (a) coded exposure camera, and (b) MIP for reducing the deconvolution noise and analyze the optimal capture strategy for known/unknown motion magnitudes and directions.

3.1. Coded exposure camera

In coded exposure, the PSF h is modified by modulating the integration pattern of the light without camera motion. Instead of keeping the shutter open for the entire exposure time, a coded exposure camera 'flutters' the shutter open and close using a carefully chosen binary code. Let n be the code length and s be the number of ones in the code. Light is integrated when the code is 1 and is blocked when it is 0. This preserves high spatial frequencies in the captured blurred image at the expense of losing light. Note that $s = 1$ is equivalent to the short exposure image and $s = n$ is

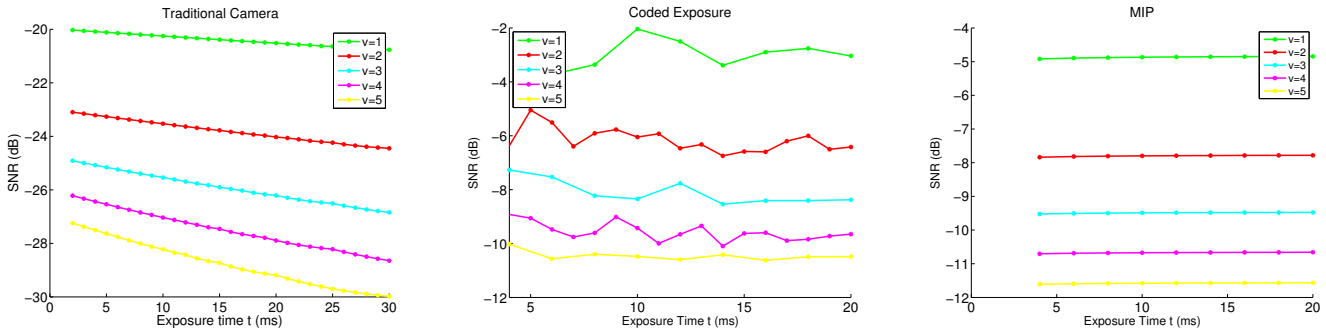


Figure 3. Key idea: At large signal-dependent noise ($Ct \gg \sigma^2_{gray}$), SNR_d decreases as the exposure time is increased for traditional camera but not for coded exposure and MIP. Plots show the decrease in SNR for different object speeds. For these plots, parameters depending on exposure time and object speed were used for both coded exposure and MIP.

equivalent to the traditional camera. Thus, coded exposure provides tradeoff between the amount of light and amount of deconvolution noise. The SNR of the deconvolved image for the coded exposure camera is given by

$$\mathrm{SNR}_d^{CE} = \frac{\bar{i_o} ts/n}{f_{CE}\sqrt{\sigma^2_{gray} + Cts/n}}, \tag{7}$$

since both the signal and signal dependent noise will be reduced by a factor of s/n.

In [16], the light loss was kept equal to 50% ($s = n/2$). Note that [16] ignores the loss of light in their analysis of deconvolution noise and thus only minimizes f_{CE} for finding the best code, while one should maximize SNR_d^{CE}. We first evaluate the relationship between n and s for optimal code design incorporating the loss of light.

Code selection incorporating light loss: First, we analyze choice of n for fix amount of light (same s). The same amount of light ensures that capture noise is similar and one can directly compare DNF's for different n. Figure 4 (left) show plots of DNF versus n, for several values of s. For each plot, n is in the range $[s, 3s]$. Note that DNF decreases sharply as n is increased, as expected. However, the 'knee' in the curves shows that increasing n beyond a certain point leads to similar performance. Since the knee occurs before $2s$, this implies that a smaller code length can be used. Small codes are easier to search for and also lead to larger on/off switching time in a practical implementation.

Next, we plot the SNR as s is increased from 1 to n for a fixed n. At low light levels, $\mathrm{SNR} \propto \frac{s/n}{f_{CE}}$. At low s, the increase in noise due to low light overwhelms the reduction in deconvolution noise. At high s, deconvolution noise dominates. Thus, there is a distinct valley for each curve and $s = n/2$ is a good choice as shown in Figure 4 (middle).

However, now consider the effect of signal-dependent noise at high light levels. In this case, $\mathrm{SNR} \propto \frac{\sqrt{s/n}}{f_{CE}}$ and the plots are shown in Figure 4 (right). Notice that for a given n, the performance is good for a range of s and not just for $s = n/2$. Thus, it means that a code with smaller s can be used. Since the size of search space is of the order of $\binom{n}{s}$, it leads to a faster search for small s. For example, for $n = 40$, $\binom{n}{20} = 1.37 \times 10^{11}$, while $\binom{n}{8} = 76.9 \times 10^6$.

Choice of t: Now we analyze the performance of coded exposure for different exposure time t by considering a code depending on the blur size. However, in practice, the object speed (blur size) is not known a-priori. The analysis of how a code generalizes for different object velocities is done in Section 4. Figure 3 (middle) plots SNR ($\propto \frac{\sqrt{ts/n}}{f_{CE}}$) for coded exposure versus t for different velocities v_i, where for every blur size $k = tv_i$, the best code was used. Note that with signal dependent noise, SNR does not decrease with exposure time. Thus, the exposure time could be increased and is useful as static parts of the scene could then be captured with higher SNR.

3.2. Motion invariant photography

In MIP, the camera is moved with a constant acceleration so that objects moving with speed within $[-v_r, v_r]$ and in the *same* camera motion direction result is same motion PSF. Intuitively, for every object velocity $v \in [-v_r, v_r]$, the moving camera spends an equal amount of time moving with v, thus capturing it sharply for that time period. The PSF is thus peaked at zero and has low deconvolution noise compared to a box (flat) PSF for the traditional capture. Let the acceleration of the camera be a and let $T = t/2$. For velocity range $[-v_r, v_r]$ and exposure time t, $a \geq v_r/2T = v_r/t$ [13] for good performance. SNR of the deconvolved image for MIP is given by

$$\mathrm{SNR}_d^{MIP} = \frac{\bar{i_o} t}{f_{MIP}\sqrt{\sigma^2_{gray} + Ct}}, \tag{8}$$

where f_{MIP} depends on the modified PSF due to camera motion.

Choice of t: We first analyze the performance of MIP for a given velocity v and exposure time t using $a = v/t$.

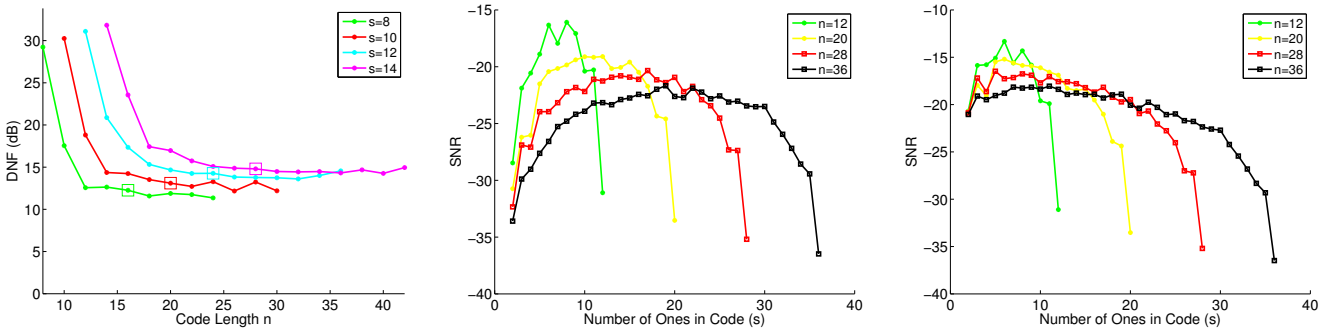


Figure 4. Choice of optimal n and s for coded exposure. (Left) For a given s, DNF decreases as n is increased and the 'knee' in each curve happens before $n = 2s$ (marked with square). This indicates that $s < n/2$ could be used. (Middle) For signal-independent noise, SNR maximizes around $s = n/2$. (Right) However, for signal-dependent noise, smaller s can be used.

Note that in practice, the speed and direction of the object is not known a-priori and how the performance generalizes is described in Section 4. Figure 3 (right) plots SNR at high signal dependent noise ($\propto \frac{\sqrt{t}}{f_{MIP}}$) with respect to t for various speeds, assuming known motion direction. Note that the SNR does not decrease with increase in exposure time. Thus, exposure time for MIP can be increased similar to coded exposure camera for capture.

4. Comparisons and performance analysis

First we compare different capture strategies for the *same* amount of captured light. This ensures that the capture noise is similar for all three capture strategies, allowing directly comparisons of the DNF's. Note that to keep the same light level, t is decreased by a factor of n/s for MIP and traditional camera. This will lead to more blur in coded exposure image by the same factor. In [13], coded exposure deblurring was visually compared with MIP using synthetic data, but [13] does not cite the code and blur size used for comparisons. Thus, it is difficult to fairly evaluate the performance in [13]. In addition, the captured light level is not same for comparisons in [13].

DNF comparison: Figure 5 compares DNF's with t for different velocities. For coded exposure, the motion direction is not known but speed was assumed to be known for computing the optimal code. In contrast, for MIP, motion direction was assumed to be known and maximum speed was set to $v_r = 3$. In addition, $a = v_r/t$ was used separately for each t for best performance. While at lower speeds ($v < v_r$), MIP gives low deconvolution noise, as the speed approaches v_r, coded exposure performs better than MIP.

Performance generalization for motion magnitude: The acceleration parameter a for MIP is set based on the desired velocity range $[-v_r, v_r]$. We first analyze how a particular choice of a performs when the object velocity is outside this range. Note that we assume that the camera can still choose a different a based on the exposure time.

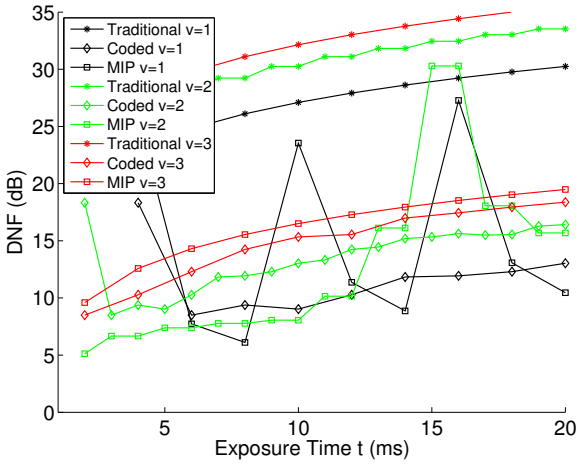


Figure 5. Comparison of DNF for various capture strategies for the *same* light level. t represents the exposure used for coded exposure camera. At lower speeds ($v < v_r$), MIP gives low deconvolution noise than compared to coded exposure. But coded exposure becomes better as v approaches v_r.

Figure 6 (middle) shows DNF for velocity range $[0, 2v_r]$ with t, using $v_r = 3$. Within the velocity range $[0, v_r]$, the deconvolution noise is low, but it increases as the object speed increases. Intuitively, the reason for good performance of MIP is that for some amount of time within the exposure, the camera moves with the *same* speed as the object. Thus, PSF for all object speeds between $[-v_r, v_r]$ is highly peaked. However, when the speed of object lies outside $[-v_r, v_r]$, this does not happen, and PSF becomes more like a box function as shown in Figure 2. Using the frequency domain analysis in [13], the image due to MIP correspond to a parabola shaped slice in the f_x-f_t space. The parabola lies within the wedge given by the velocity range $[-v_r, v_r]$ and is optimal for $[-v_r, v_r]$. However, object speed v outside $[-v_r, v_r]$ corresponds to a line outside the wedge and the *parabolic* slice does not capture high frequency information for those speeds (Figure 2). Thus, the performance degrades rapidly.

In contrast, coded exposure camera is optimized for a

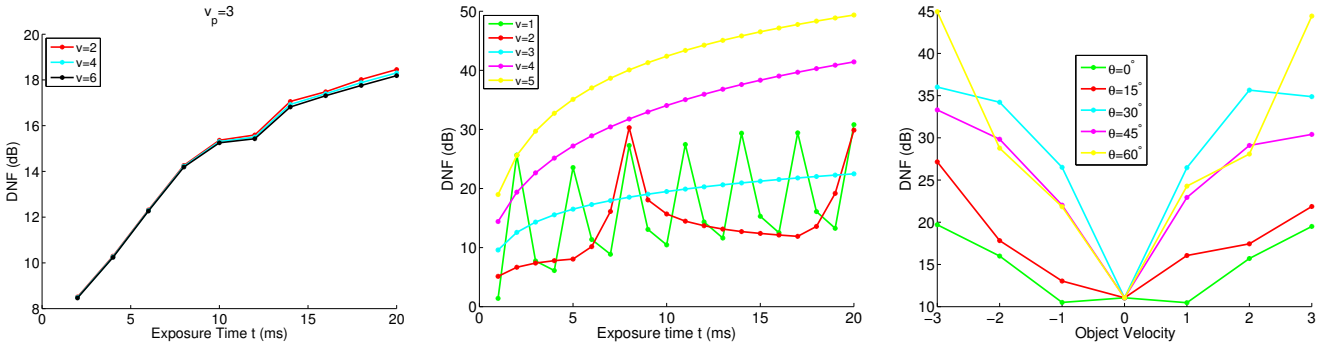


Figure 6. Performance generalization: (Left) As the object speed increases beyond assumed, the deconvolution noise does not increase for coded exposure, but the minimum resolvable feature size increases. For MIP, performance degrades rapidly as the motion magnitude v increases beyond $v_r = 3$ (middle) and as motion direction differs from camera motion direction (right).

particular velocity v_p instead of a velocity range. Similar to above, we assume that the camera can still choose a different code based on the exposure time t. Thus, for each t, a code with length $k = v_p t$ is chosen, while the actual object velocity v could lead to a different amount of blur $k = vt$. Figure 6 (left) plots DNF versus t for different object speeds assuming $v_p = 3$. Note that the deconvolution noise does not increase as v increases beyond v_p. However, since the blur could only be resolved within one 'chop' (single 1) of the code, the minimum resolvable feature size increases with v. In frequency domain, coded exposure modulates along the f_t direction using the frequency transform of the chosen code and the image corresponds to the horizontal slice ($f_t = 0$). Even if the object velocity v differs from v_p, broadband modulation along f_t allows high frequency information to be captured in the horizontal slice, leading to good performance.

Another interesting observation is that MIP optimizes the capture bandwidth for *all* speeds with $[-v_r, v_r]$. In a practical situation, however, the speed of the object may not vary from 0 to v_r, but rather in a small range, centered around a speed greater than zero. In such cases, the MIP capture process does not remain optimal.

Performance generalization for motion direction: Coded exposure makes PSF invertible for *any* motion direction, but the direction needs to be known for deblurring as shown in [16]. For MIP, the camera needs to be moved along the object motion direction, while the magnitude of motion is not required for deblurring as PSF becomes invariant. However, as the object motion direction differs from the camera motion direction, performance degrades sharply for MIP, as the PSF does not remain invertible or invariant. Let θ denote the difference in camera and object motion directions. Figure 6 (right) plots DNF for θ ranging from 0 to 90°. Note that noise increase sharply with θ and all curves meet at $v = 0$ (static scene).

Static scene parts: For coded exposure and traditional camera, the static parts of the scene are captured without any degradation (for the same light level). For MIP, PSF estimation is not required for static scene parts, but they are also blurred due to camera motion, leading to SNR degradation.

In conclusions, if the motion direction is known exactly and the motion magnitude is (unknown) within a range, MIP solution should be used for capture. However, if motion direction is unknown, coded exposure is the optimal choice. Moreover, the performance degrades slowly for coded exposure as the object speed/direction differs from assumed, but degrades sharply for MIP.

5. Implementation and results

We capture a high speed video and simulate the various capture strategies for comparisons. A traditional camera image can be obtained by simply averaging the frames in the high speed video and that corresponding to a coded exposure camera can be obtained by averaging frames corresponding to the 1 of the code. MIP can be simulated by shifting the images according to the camera motion before averaging. For these experiments, individual high speed camera images are sufficiently above the noise bed of the camera (high signal-dependent noise). Thus, averaging N high speed camera images has similar noise characteristics as a single N times longer exposure image, since both increases capture SNR by $\sqrt{N}$.

Using a high speed video enables us to evaluate the performance of all three cameras on the same data, which otherwise would require complicated hardware setup. In addition, images corresponding to different exposure times and camera motions can be easily obtained for the same data. In general, there could be an integration gap between the frames of a high speed camera. We use the Photron FASTCAM-X high speed camera, with a mode that allows frame integration time to be equal to the inverse of frame rate. Thus, this gap will not have any significant effect. Moreover, any such effect will be identical for all three techniques. For all techniques, we deblur simply by solving the linear system (without any regularization) to analyze the effects of deconvolution noise clearly.

Setup: A high speed video of a moving resolution chart is captured at 1000 fps. The speed is determined manually to be $v = 0.28$ pixels/frame and is fairly low for accurate simulation of various strategies.

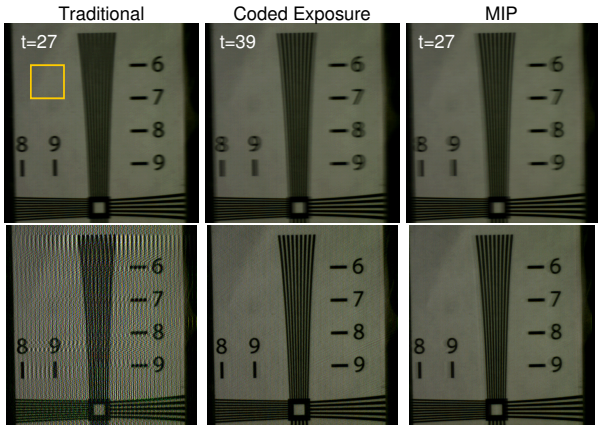


Figure 7. Comparison of three approaches for the same light level. (Top row) Blurred images. (Bottom Row) Corresponding deblurred images. DNF were empirically estimated to be 19.8, 2.41 and 1.5 dB for traditional, coded exposure and MIP respectively. Visually, the coded exposure deblurring is sharper than MIP deblurring.

Comparisons: Figure 7 show comparisons of the three techniques. The blurred image for coded exposure camera was generated using the code 1010110011111 ($n = 13, s = 9$) and exposure time of 39 ms (chop time of 3 ms). For traditional camera (box) and MIP, the exposure time t was reduced to $39 * 9/13 = 27$ ms to have the same amount of light level. For MIP, $a = 2v/t$ was used to get good PSF ($v_r = 2v$). As expected, both MIP and coded exposure results in good deblurring. DNF was calculated empirically using a 100×100 homogeneous region (shown in yellow box) as the ratio of the variance in the deblurred and blurred images. DNF values were 19.8, 2.41 and 1.5 dB for traditional, coded exposure and MIP respectively. Visually, the coded exposure deblurring result is sharper than MIP deblurring.

Coded exposure performance: An easy way to simulate faster moving object for coded exposure is to increase the object displacement within each '1' of the code by adding consecutive frames. We simulate the resolution chart motion with speeds of 1.07, 1.61 and 2.1 pixels/ms by adding 4, 6 and 8 consecutive images for each chop respectively using the same $n = 13$ code (optimal for $v_p = 1$). Figure 8 shows that as the speed of object increases, deblurring results do not show deconvolution artifacts, but the size of minimum resolvable feature increases (the vertical lines at the bottom are not resolved clearly). However, the deconvolution noise remains almost constant: values were 2.61, 2.65 and 2.69 dB respectively. This is a very useful property since only the 'effective resolution' on the deblurred object is decreased without any deconvolution artifacts, leading to a gradual performance decay.

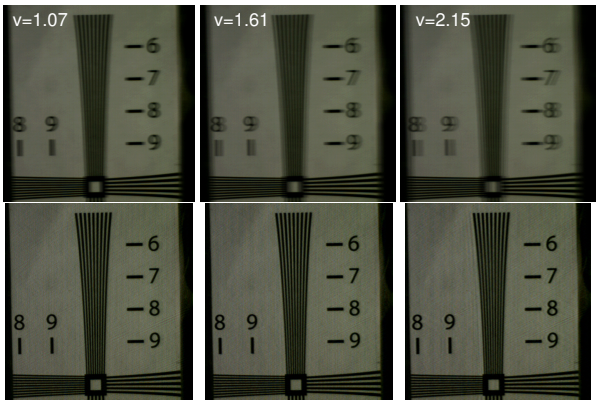


Figure 8. As object speed v increases beyond desired speed v_p, performance of coded exposure does not degrade in terms of deconvolution noise. But the size of minimum resolvable feature increases. Deconvolution noise increase is 2.61, 2.65 and 2.69 dB respectively.

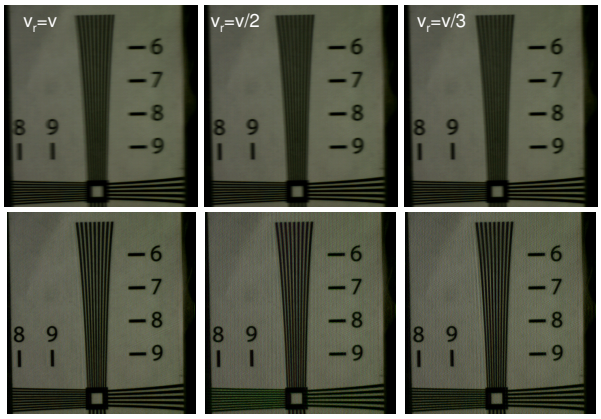


Figure 9. Performance of MIP degrades sharply as the object velocity increases beyond assumed limit. Results are shown for $v = 0.28$ and $v_r = 0.28$, 0.14 and 0.09 pixels/ms respectively. Note that the deblurring shows increased noise when v is greater than v_r. Corresponding DNF's are 2.99, 10.7 and 13.1 dB respectively.

MIP performance: Figure 9 shows that MIP deblurring performance degrades rapidly as the object velocity increases beyond v_r. For these results, blurred images were generated using $a = v/t$, $a = v/2t$ and $a = v/3t$, effectively setting v_r to v, $v/2$ and $v/3$ respectively. For $v_r = v$, DNF was low (2.99 dB) as expected, but increases to 10.7 and 13.1 dB for $v_r = 2v$ and $v_r = 3v$ respectively.

Figure 10 shows that as object motion direction differs from camera motion, deblurring performance degrades sharply for MIP due to the vertical component of motion blur. Even though the vertical blur is smaller than 4 pixels for all three cases, deblurring results shows artifacts.

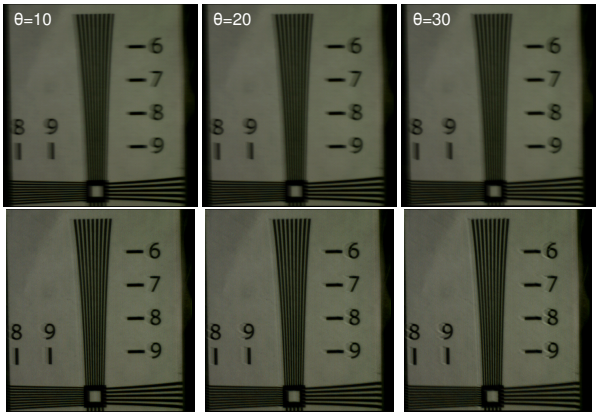


Figure 10. MIP performance degrades as the camera motion direction differs from object motion direction by angle θ. Although the vertical blur is small (4 pixels), deblurring shows artifacts since the resulting PSF does not remain invertible.

6. Conclusions

We posed the problem of optimal single image capture for motion deblurring as maximizing the SNR of the deconvolved object, taking into account capture noise, light level and deconvolution noise. We showed that increasing exposure time to gain more light is not beneficial for a traditional camera in presence of motion blur and signal dependent noise. For both coded exposure and MIP, exposure time could be increased without SNR degradation on moving objects. Coded exposure is optimal for any unknown motion direction with known motion magnitude, and its performance degrades gradually as motion magnitude differs from desired. MIP is optimal if the motion direction is known and the motion magnitude is within a known range, but its performance degrades rapidly as the motion magnitude and direction differs, along with increased noise on the static scene parts. We showed that optimal codes for coded exposure need not be 50% on-off if signal-dependent noise is taken into account. We presented evaluation on real datasets, allowing the design of optimal capture strategy for single image motion deblurring. Our analysis could also be extended for comparing other capture strategies for motion/defocus blur using single/multiple images and more complicated blur functions.

Acknowledgements We thank the anonymous reviewers and several members of MERL for their suggestions. We also thank Jay Thornton, Keisuke Kojima, and Haruhisa Okuda, Mitsubishi Electric, Japan, for help and support.

References

[1] R. N. Clark. Digital camera sensor performance summary. http://clarkvision.com, 2008.

[2] S. Dai and Y. Wu. Motion from blur. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 1–8, June 2008.

[3] E. R. Dowski and W. Cathey. Extended depth of field through wavefront coding. *Appl. Optics*, 34(11):1859–1866, Apr. 1995.

[4] E. Fenimore and T. Cannon. Coded aperture imaging with uniformly redundant arrays. *Appl. Optics*, 17:337–347, 1978.

[5] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3):787–794, 2006.

[6] M. Grossberg and S. Nayar. High Dynamic Range from Multiple Images: Which Exposures to Combine? In *ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV)*, Oct 2003.

[7] G. E. Healey and R. Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(3):267–276, 1994.

[8] S. Hiura and T. Matsuyama. Depth measurement by the multi-focus camera. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 953–961, 1998.

[9] P. Jansson. *Deconvolution of Image and Spectra*. Academic Press, 2nd edition, 1997.

[10] J. Jia. Single image motion deblurring using transparency. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 1–8, June 2007.

[11] N. Joshi, R. Szeliski, and D. Kriegman. PSF estimation using sharp edge prediction. In *Proc. Conf. Computer Vision and Pattern Recognition*, June 2008.

[12] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3):70, 2007.

[13] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman. Motion-invariant photography. *ACM Trans. Graph.*, 27(3):1–9, 2008.

[14] L. Lucy. An iterative technique for the rectification of observed distributions. *J. Astronomy*, 79:745–754, 1974.

[15] H. Nagahara, S. Kuthirummal, C. Zhou, and S. Nayar. Flexible Depth of Field Photography. In *Proc. European Conf. Computer Vision*, Oct 2008.

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

[17] N. Ratner and Y. Y. Schechner. Illumination multiplexing within fundamental limits. In *Proc. Conf. Computer Vision and Pattern Recognition*, June 2007.

[18] W. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. of America*, 62(1):55–59, 1972.

[19] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. A theory of multiplexed illumination. In *Proc. Int'l Conf. Computer Vision*, volume 2, pages 808–815, 2003.

[20] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graph.*, 27(3):1–10, 2008.

[21] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.*, 26(3):69, 2007.

[22] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph.*, 27(3):1–10, 2008.