Camera-Motion and Effective Spatial Resolution

Feng Xiao; Motorola Inc.; Libertyville, IL 60048
Amnon Silverstein; HP Labs; Palo Alto, CA 94304
Joyce Farrell; Stanford University; Stanford, CA 94305

Abstract

The booming mobile imaging market is aggressively pursuing CMOS imagers with higher spatial resolution and smaller form factors. Given a fixed die size, reducing pixel size seems like a straightforward way to increase spatial resolution. The reduced sensitivity of smaller pixels, however, must be compensated by longer exposure durations. Longer exposure durations result in camera motion blur when the camera is handheld, effectively reducing spatial resolution. We empirically measured how the exposure duration, camera mass and subject's skill level affect camera motion. We then used these measurements to calculate the effect that camera motion has upon the system MTF for CMOS imagers with different pixel sizes at different scene illuminance levels. We described the scene illuminance conditions for which imaging sensors with larger pixels produce sharper images than imagers with smaller pixel sizes.

Introduction

Imaging sensor vendors and camera makers have been competing on megapixels and price. The cost of the sensor is most dependent on the size of the die, so designers have crammed more pixels into smaller dice by shrinking the pixel size. Each pixel has a somewhat fixed overhead of space required for circuitry, so as the overall size is reduced, the light sensitive area of each pixel gets smaller [1,2,3,4].

As the light-sensing areas become smaller, the signal-to-noise ratios get smaller as well. To bring the signal-to-noise ratio up to an acceptable level requires a longer exposure [5]. This presents a problem for hand-held photography, since a long exposure will result in motion blur that limits the resolution of the photograph. If the loss in resolution due to motion is greater than the pixel sampling resolution, a higher resolution image could have been obtained with larger pixels and a shorter exposure.

This paper considers the tradeoffs between exposure and pixel subtense. Film photographers have long used a rule-of-thumb that a hand held 35mm camera should have an exposure in seconds that is not longer than the inverse of the focal length in millimeters. E.g. that a 50mm lens should have an exposure of 1/50sec at most. This is a very rough approximation at best, and it does not apply to digital photographs. Further, hand shake varies with camera mass and photographer. There have been very few studies of actual hand-held camera shake, so the rule-of-thumb has been difficult to formalize and generalize.

In this paper, we carefully measure camera motion as a function of exposure time. We also investigate how camera motion is affected by different photographers and camera masses. We use the measurements to calculate the effect that camera motion has on the system MTF for CMOS imagers with different pixel sizes at scene illuminance levels. We also discuss the applications to anti-shake system designs.

Experimental Setup

To measure the effects of exposure duration, camera mass and user's photography skill level on the amount of camera-shake, each subject was asked to take five pictures of a point light source (a single LED was imaged to a very small spot with a lens, and this spot was placed 9 meters from the test cameras) while holding the camera naturally at each exposure duration (twenty-one total half-stop apart from 0.01 second to 1 second). Seven unpaid subjects (five males and two females, ages from 25 to 50 years old) participated in the experiment. All of the subjects had used digital cameras prior to the experiment. Two of the subjects had used digital cameras extensively prior to the experiment, and they are regarded as "expert users" in the following data analysis. Three cameras were used in these studies (the high-mass Nikon D70, the medium-mass Canon G3 and the low-mass Canon A95) and Table 1 lists the camera-specific parameters. The room lights were turned off during the experiment.

To maximize the spatial accuracy of the measurement, all cameras were set at their maximum optical zoom position (4x Zoom for both D70 and G3 and 3x Zoom for A95). Each camera was manual focused on the light source prior to the experiment. In addition, both D70 and G3 were set to output raw image data format to minimize the effect of other in-camera image processing functions.

Table 1: Camera parameters

Cameras	Nikon D70	Canon G3	Canon A95
Mass (grams)	1251	614	335
Degree/pixel	0.00644	0.0063	0.0068
Optical zoom	4X	4X	3X
Total pixels	3008x2000	2272x1704	2592x1944
Raw output	Yes	Yes	No

Results

Data Analysis

For each picture, we extracted the standard deviation along the long axis and the short axis vector of the camera-shake pattern using the weighted principal component analysis method [6,7] as illustrated in Figure 1. We then used following Equation to fit the standard deviation along either the long axis ($_L$) or the short axis ($_S$) with the exposure duration T:

$$\begin{cases} \sigma_L(T) = a_L T^b \\ \sigma_S(T) = a_S T^b \end{cases} \quad (1)$$

We chose this form because it fits both the random-walk pattern (b=0.5) and the straight-line-walk pattern (b=1) and camera-shake pattern can be something in between these two patterns. To simplify the comparison, we chose exponent b to be the same across all cameras and subjects. An optimal value of 0.5632 gave the best fit for the averaged standard deviation across all conditions (Figure 2) and this provides further evidence that camera-shake pattern falls in between a complete random-walk pattern (b=0.5) and a straight-line-walk pattern (b=1).

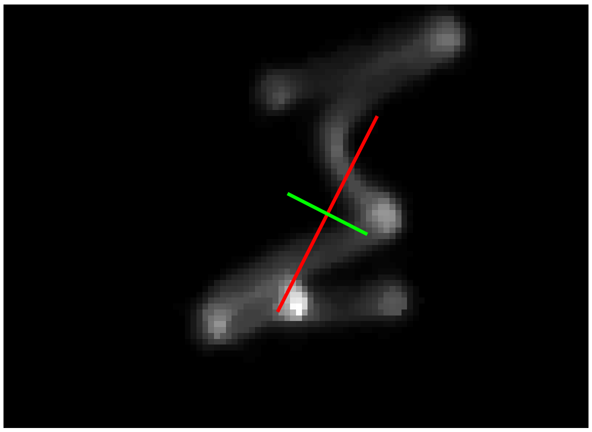

***Figure 1**. An example showing the extracted standard deviation along the long axis and short axis on the original image of captured camera-shake pattern at exposure duration of 0.5 second.*

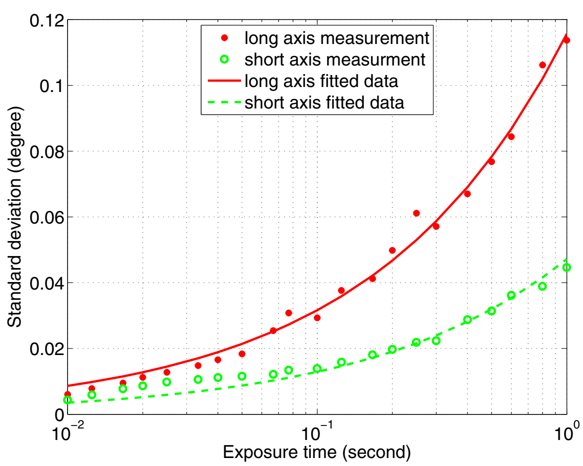

***Figure 2**. Comparison between the measured standard deviation along long axis (star dot) and short axis (circle dot) of camera shake pattern and the best fitted value from Equation 1 (where a = 0.1157 for long axis and a=0.0472 for short axis).*

Camera Variation

To evaluate the effect of camera mass on the amount of camera-shake, we repeated the same fitting procedure (Equation 1) for each camera separately. Table 2 shows the estimated parameter a along long axis and short axis for each camera. There is a clear trend that the less massive the camera, the greater the amplitude of the camera motion (along long axis). This trend is also illustrated in Figure 3. Since most camera phones are less massive than the least massive camera used in this study (335 grams), we would expect camera-shake to be an even more serious problem for camera phones.

Table 2: Estimated values of parameter a along the long and short axis cameras with different mass.

Cameras	D70	G3	A95
Mass (grams)	1251	614	335
Long-axis	0.092	0.1218	0.1333
Short-axis	0.0453	0.0445	0.0520

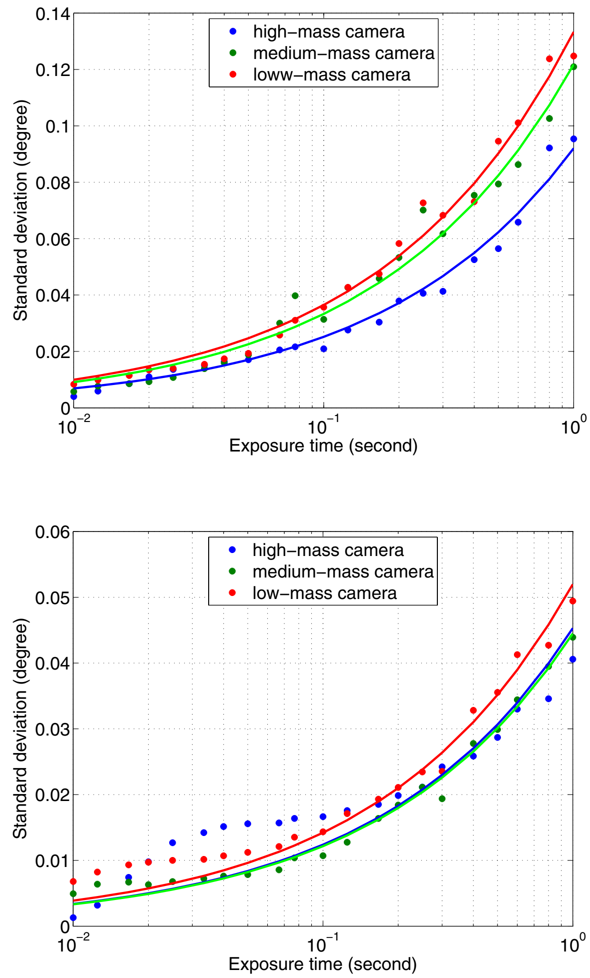

***Figure 3**. The effect of camera mass on the standard deviation of camera-shake along: (top) long axis and (bottom) short axis.*

Expert vs. Non-Expert

To evaluate the effect of subject's photography skill level on the amount of camera-shake, we repeated the same fitting procedure for the expert group (two subjects) and the non-expert group (five subjects). The optimal values of parameter a for these two groups and each camera are shown in Table 3. Again, there is a clear trend that skilled photographer tends to hold the camera more stable than less skilled photographer and the difference grows when the camera becomes lighter. It is also interesting to note the difference also grows with exposure duration while the difference is ignorable for exposure duration shorter than 0.04 second (Figure 4).

Table 3: Estimated long axis parameter a for subjects with different photography skill levels.

Cameras	Nikon D70	Canon G3	Canon A95
Expert	0.0544	0.0727	0.0785
Non-expert	0.1071	0.1414	0.1553

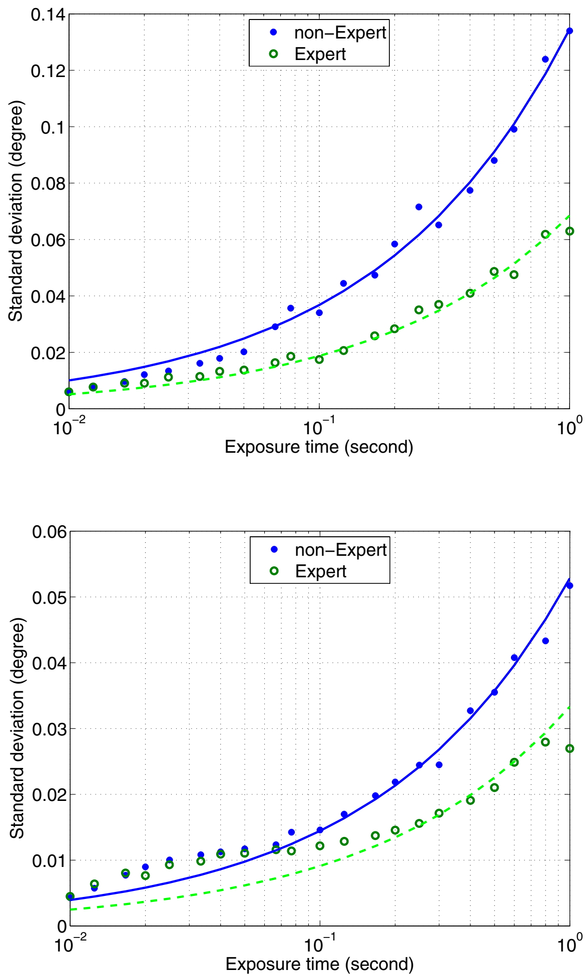

Figure 4. Effect of subject's photography skill level on the standard deviation of camera-shake along: (top) the long axis and (bottom) the short axis.

Applications

In this section, we show the applications of the camera-shake measurements into the pixel size tradeoff analysis and the potential use of these data for anti-shake algorithm design.

Camera-Motion and Pixel Size Tradeoff

Based on the camera-shake model (Equation 1), we used the ISET toolbox [8] to evaluate how the amount of camera-shake affects the effective 50% system MTF cutoff frequency of sensors with the same die size (1/4 inch for example) but different pixel size (from 7.4 um to 1.7 um) at different illuminance levels (from 10 Lux to 10000 Lux). To simplify the analysis, monochromatic sensors with a diffraction limited lens (F#=2.8) were assumed and

Table 4 shows other sensor parameters used in the simulation. For each condition, the exposure duration is set to achieve the same SNR level for a 20% reflectance gray Lambertian surface in the scene for sensors with different pixel. Therefore, sensors with smaller pixel size will need longer exposure duration and thus introduce higher amount of camera-shake. Two targeted SNR levels (30dB and 20dB) were used in this simulation where noise is barely visible at 30dB SNR [5] and is visible but not very objectionable at 20dB SNR. For the camera-shake model, we chose the estimated parameters for the non-expert group using the lightest A95 camera for the simulation.

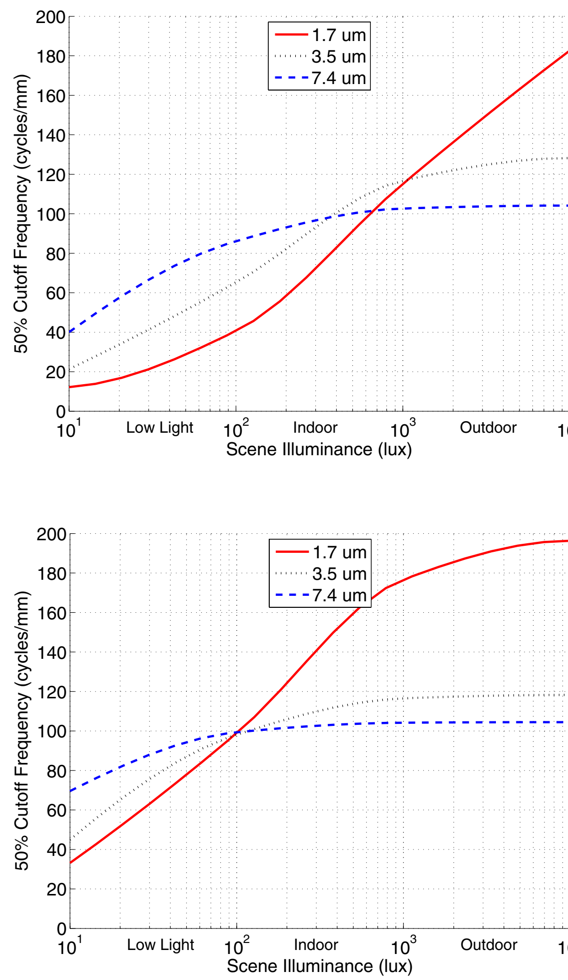

Figure 5. Camera-motion effected system spatial resolution at different light levels for sensors with the same die size but different pixel size at targeted SNR level of: (top) 30dB and (bottom) 20dB.

For targeted SNR level of 30dB, the simulation (Figure 5) shows that 1.7um sensor records sharper pictures than 3.5um sensor only when scene illuminance level is higher than 1000 Lux (outdoor daylight). In other words, a 5M pixel 1.7um camera phone could record a more blurred pictures than a 1.3M pixel 3.5 um camera phone for most indoor light conditions (less than 1000 Lux) unless the amount of camera-shake can be greatly suppressed through electrical/optical stabilization mechanism or shooting with tripod. However, if we lower the SNR requirement to 20dB, then the spatial resolution advantage reverses at scene illuminance level around 100 Lux. With the addition of other noise sources and the amplification of existing noise by the image processing pipeline, too low level of targeted SNR at the sensor output should be avoided.

Table 4: Basic sensor parameters used in the ISET simulation.

Pixel size (um)	1.7	3.5	7.4
Read noise (e^-)	10	30	43
Voltage swing (volts)	0.7	1.0	1.2
Effective fill factor	50%	50%	50%
Peak QE (at 550nm)	0.6	0.6	0.6
Conversion gain (10^{-6} v/e-)	60	25	13
Dark voltage (e^-/second)	80	150	240

Anti-Shake Algorithms

Figure 6 shows how the camera-shake pattern changes with the exposure duration from 0.01 second to 0.8 second. For exposure duration shorter than 0.125 second, the camera motion is close to a straight line. Anti-shake algorithms can take advantage of this through the estimation of the direction and length of this line and then undo part of the motion.

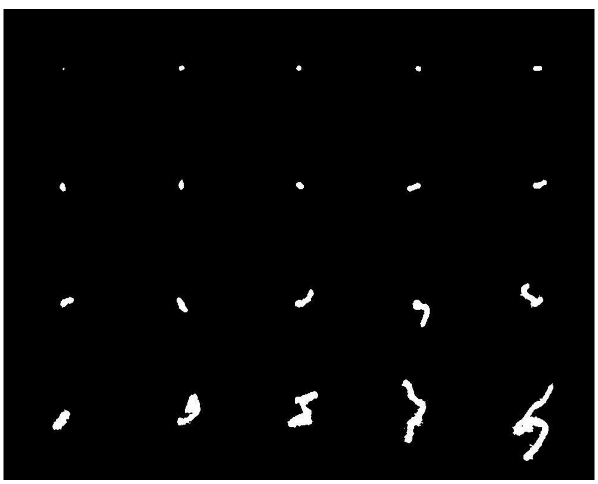

Figure 6*. Example showing how the camera-shake pattern changes with exposure duration (from left to right and from top to down, exposure duration increases from 0.01 second to 0.8 second).*

Summary

The camera-shake measurement results show that the camera motion pattern is somewhat in between the random-walk pattern and straight-walk pattern. The relationship between the standard deviation of camera motion along the long axis (or short axis) and the exposure duration can be efficiently described by Equation 1. The lighter the camera, the larger the camera motion tends to be. Subject's photography skill also has a significant impact on the amount of camera-shake, especially for exposure duration longer than 0.04 second. The difference between expert subjects and non-expert subjects is also amplified for lighter camera. Based on the camera-shake model, we demonstrated (using the ISET toolbox) that a 5M pixel 1.7 um camera phone could record a more blurred pictures than a 1.3M pixel 3.5 um camera phone for most indoor light conditions (less than 1000 Lux) unless the amount of camera-shake can be greatly suppressed through either electrical or optical stabilization mechanism or shooting with tripod. The close-to-a-line camera motion pattern for exposure duration shorter than 0.125 second also implies that anti-shake algorithms can take advantage of this through the estimation of the direction and length of this line and then undo part of the motion. Future improvement can be done by using a more complicated mathematical model than Equation 1 and the camera-shake measurement for shorter exposure durations might also need to be improved through better experimental setup. It is also worthwhile to extend this study to color sensors.

Acknowledgements

We would like to thank Prof. Brian Wandell and Dr. Xuemei Zhang for insightful discussions and all subjects who participated in this study.

References

[1] T. Chen, P. Catrysse, A. El Gamal & B. Wandell, How Small Should Pixel Size Be? Proc. SPIE, Vol. 3965, pg. 451-461. (2000).

[2] P. Catrysse & B. Wandell, Roadmap for CMOS Image Sensors: Moore Meets Planck and Sommerfeld, Proc. SPIE, Vol. 5678, pg. 1-13. (2005).

[3] H. Rhodes, G. Agranov, C. Hong, U. Boettiger, R. Mauritzson, J. Ladd, I. Karasev, J. McKee, E. Jenkins, W. Quinlin, I. Patrick, J. Li, X. Fan, R. Panicacci, S. Smith, C. Mouli & J. Bruce, CMOS Imager Technology Shrinks and Image Performance, IEEE Workshop on Microelectronics and Electron Devices, pg. 7–18. (2004).

[4] J. Farrell, F. Xiao & S. Kavusi, Resolution and light sensitivity tradeoff with pixel size, Proc. SPIE, Vol. 6069, (2006).

[5] F. Xiao, J. E. Farrell & B. Wandell, Psychophysical Thresholds and Digital Camera Sensitivity: The Thousand Photon Limit, Proc. SPIE, Vol. 5678, pg. 75-84. (2005).

[6] D. Skocaj & A. Leonardis, Weighted Incremental Subspace Learning, Proc. Workshop on Cognitive Vision, (2002).

[7] K. Forbes & E. Fiume, An Efficient Search Algorithm for Motion Data Using Weighted PCA, ACM SIGGRAPH Symposium on Computer Animation, (2005).

[8] J. E. Farrell, F. Xiao, P. Catrysse & B. Wandell, A Simulation Tool for Evaluating Digital Camera Image Quality, Proc. SPIE, Vol. 5294, pg. 124-131. (2004).

Author Biography

Feng Xiao is currently a principal staff electrical engineer with Motorola Inc. Before joining Motorola, he worked as a color scientist with Agilent Technologies and an imaging software engineer with Texas Instrument Inc. Feng Xiao received his MS in Electrical Engineering and PhD in Psychology, both from Stanford University. He is researching topics in the system modeling of digital imaging devices. He is a member of IS&T and IEEE.