Noise Analysis for Sec.5.2 in Coded Rolling Shutter Photography

Jinwei Gu, Columbia University
Yasunobu Hitomi, Sony Corporation
Tomoo Mitsunaga, Sony Corporation
Shree Nayar, Columbia University

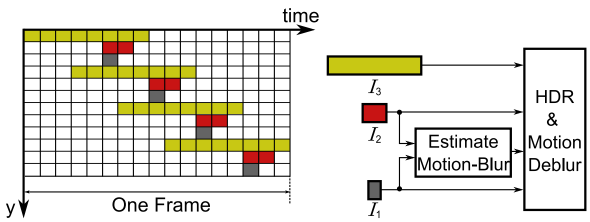

Figure 1. Staggered readout and multiple exposure coding for HDR imaging with hand-held cameras.

In Sec.5.2, we proposed to use staggered readout and coded exposure for HDR Imaging for hand-held cameras. The idea is summarized in the diagram in Fig. 1. The pixel array of a CMOS image sensor is coded with staggered readout ($K = 3$) and three exposures, Δt_{e1}, Δt_{e2}, and Δt_{e3}. Thus, from a single input image, I, we can extract three sub-images, I_1, I_2, and I_3. These sub-images are resized vertically to full resolution using cubic interpolation. For static scenes/cameras, these sub-images can be directly used to compose a HDR image. For hand-held cameras, however, we should consider the motion blur caused by camera shake.

In order to analyze the noise amplification due to motion deblurring in the proposed method, we make the following assumptions. (1) We assume no pixel is saturated in the input image, and thus the dynamic range is essentially determined by noise. (2) Camera shake is assumed to be a uniform translational motion, with v pixels/second. (3) Image noise is Gaussian additive noise σ.

For a conventional rolling shutter camera, without introducing any motion blur in the captured image (*i.e.*, the motion has to be less than 1 pixel), the maximum Signal-to-Noise Ratio (SNR) we can achieve is

$$\text{SNR}_0 = \frac{I_0 \cdot (1/v)}{\sigma} = \frac{I_0}{v\sigma}.$$

With a coded rolling shutter, multiple exposures can be embedded into one single image. Suppose we have three exposures — the shortest exposure Δt_{e1} introduces no motion blur, and the medial exposure Δt_{e2} causes the motion blur of r_2 pixels wide, and the long exposure Δt_{e1} causes the blur r_3 pixels wide, $1 \le r_2 \le r_3$, we have

$$I_1 = I_0 + \eta_1, \tag{1}$$
$$I_2 = (I_0 \cdot r_2) * K_2 + \eta_2, \tag{2}$$
$$I_3 = (I_0 \cdot r_3) * K_3 + \eta_3, \tag{3}$$

where I_i is the i-th sub-image, and $\eta_i \sim \mathcal{N}(0, \sigma)$ is noise, $i = 1, 2, 3$. K_2 and K_3 are the two normalized motion blur kernels.

As shown in [1], deblurring each of the two combined images (*i.e.*, $I_1 \oplus I_2$ and $I_1 \oplus I_2 \oplus I_3$) using the $1/f$ prior on natural images will amplify image noises:

$$\sigma_{b1}^2 = \sigma^2 \cdot \frac{1}{N} \sum_{\zeta} \frac{1}{\| \frac{\mathbf{K}_2(\zeta)+1}{r_2+1} \|^2 + \frac{\sigma^2}{\mathbf{A}(\zeta)}}, \tag{4}$$
$$\sigma_{b2}^2 = \sigma^2 \cdot \frac{1}{N} \sum_{\zeta} \frac{1}{\| \frac{\mathbf{K}_3(\zeta)+\mathbf{K}_2(\zeta)+1}{r_3+r_2+1} \|^2 + \frac{\sigma^2}{\mathbf{A}(\zeta)}}, \tag{5}$$

where ζ is the frequency, N is the number of pixels in the captured image, $\mathbf{K}_i(\zeta)$ is the Fourier transform of the motion blur kernel of the i-th image (*i.e.*, of length r_i pixels), and $\mathbf{A}(\zeta)$ is the averaged power spectrum of natural images. Based on Eq. (8) in the original paper, the noise in the recovered image I_r is given by:

$$\sigma_r^2 = \left(\sigma^2 + \sigma_{b1}^2 + \sigma_{b2}^2\right)/9. \tag{6}$$

Therefore the decrease in SNR (compared to the SNR_0) can be evaluated as

$$\frac{\text{SNR}_0}{\text{SNR}} = \frac{\frac{I_0 \cdot (1/v)}{\sigma}}{\frac{I_0 \cdot (1/v)}{\sigma_r}} = \frac{\sigma_r}{\sigma}, \tag{7}$$

for any exposure ratios (r_2, r_3). Theoretically, the optimal (r_2^*, r_3^*) minimizes SNR_0/SNR. Note that SNR_0/SNR is a function of noise, and thus (r_2^*, r_3^*) will change at different noise levels. In our experiments, the ratios are set to 2:8 to meet the dynamic range requirement and to calculate optical flow — potentially useful for handling more general camera motion.

The same analysis can be extended to multiple exposures other than three. Moreover, the analysis can be extended for deblurring multiple sub-images simultaneously.

References

[1] C. Zhou and S. K. Nayar. What are good apertures for defocus deblurring? In *Proceedings of IEEE International Conference on Computational Photography (ICCP)*, 2009.