At CVPR 2010, Sam Hasinoff made an observation that further refines the comparison between Eq. (6) and Eq. (7) in Section 3.2.3, and reconciles our conclusions with [Hasinoff et al. 2009].

In the paper we observed that for sufficiently high N, averaging N frames with unit gain (i.e., fixed ISO) has lower noise variance than a single frame whose exposure time is N times longer and whose gain is adjusted to achieve the same intensity range (i.e., by decreasing ISO to compensate).

Let ρ_0^2 and ρ_1^2 be the read noise variance before and after the (unit) amplifier gain. Consider N frames, each captured with τ exposure time. Let J be the mean number of photons collected during τ. After averaging N frames, the noise variance is

$$\frac{1}{N}(J + \rho_0^2 + \rho_1^2) \qquad (6)$$

If we capture one frame during the $\Delta = N\tau$ time interval, the noise variance of this frame is

$$\frac{1}{N}J + \frac{1}{N^2}\rho_0^2 + \rho_1^2 \qquad (7)$$

and Eq. (6) is less than Eq. (7) if $N > \frac{\rho_0^2}{\rho_1^2}$.

While this analysis may suggest that we should capture as many frames as possible, it rests on the comparison to a single frame with a very low corresponding ISO setting, which may not be realizable in practice.

For a complementary view, let us consider a fixed time interval $\Delta = N\tau$ and examine how this interval should be divided to minimize the noise variance [Hasinoff et al. 2009]. As the following analysis shows, the optimal number of frames to capture is generally not equal to N, even when varying ISO is taken into account.

Let $\phi = NJ$ be the total number of photons collected during Δ. If we capture M frames with $\frac{N\tau}{M}$ exposure time each and then average them, the noise variance is

$$\frac{1}{M}\left(\left(\frac{N}{M}\right)^{-2}\left(\frac{1}{M}\phi + \rho_0^2\right) + \rho_1^2\right) = \frac{1}{N}J + \frac{M}{N^2}\rho_0^2 + \frac{1}{M}\rho_1^2\,,$$

assuming that an amplifier gain of $\frac{N}{M}$ was used for each shot, to match the intensity range above. In this context, we find that variance is minimized when $M = N\frac{\rho_1}{\rho_0}$.

The above analysis assumes that the overhead (dead time between frames) is ignored and the scene is static during the image capture. The optimal number of frames in the presence of motion is still open.