Digital Photography with Flash and No-Flash Image Pairs

Philippe Lamoureux

Georg Petschnigg, Richard Szeliski, Maneesh Agrawala, Michael Cohen, Hugues Hoppe, and Kentaro Toyama. Digital photography with flash and no-flash image pairs. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers*, pages 664–672, New York, NY, USA, 2004. ACM

Petschnigg et al's *Digital Photography with Flash and No-Flash Image Pairs* aims primarily at improving point-and-shoot photography in low-light environments by combining the strengths of pictures taken with and without flash. The authors attempt to extract the fine details from the flash picture and combine them with the lighting ambiance of a denoised version of the no-flash picture. They also present variations on existing image correction techniques exploiting such image pairs, including white balancing, continous flash interpolation, and red-eye correction.

The first step in enhancing the ambient image A is to denoise it using an edge-preserving smoothing filter. The technique presented in the paper builds on the *bilateral filter*, a simple non-iterative technique that replaces each pixel by a weighted average of neighbours having similar intensity values. The weights are determined by the product of two, usually Gaussian, functions g_d (distance) and g_r (intensity difference).

$$A_p^{Base} = \frac{1}{k(p)} \sum_{p' \in \Omega} g_d(p' - p) g_r(A_p - A_{p'}) A_{p'}, \text{ where } k(p) = \sum_{p' \in \Omega} g_d(p' - p) g_r(A_p - A_{p'})$$

Because the flash image F is believed to be a much better estimate of the true high-frequency information, more detail can be preserved using a slight variation on the bilateral filter, the *joint bilateral filter*, which uses the intensity differences in F. To solve the over-blur and under-blur problems arising inside and on the boundaries of flash shadow and specularity regions, the presented solution falls back to the basic bilateral filter around these regions.

With its higher signal-to-noise ratio, the flash image can also be used to recover some of the subtle nuances overwhelmed by noise in the ambient image and lost in the smoothing process. To transfer additional or accentuate existing detail, the noise-reduced ambient image A^{NR} is multiplied by a *quotient* or *ratio image*, $F^{Detail} = \frac{F+\epsilon}{F^{Base}+\epsilon}$, where F^{Base} is computed using the basic bilateral filter to reduce haloing. Because very little detail can be recovered from flash shadows and specularities, these regions are typically ignored.

Other image correction techniques such as white balancing can also benefit from flash/no-flash image pairs. Not surprisingly, the white balance of flash pictures are usually more reliable, and a "pure flash" image Δ obtained as the difference of minimally processed (raw) images $F^{Lin} - A^{Lin}$ proves to be useful in determining the colour of the ambient illumination. By averaging the relative gain $C_p = \frac{\Delta_p}{A_p}$ of each channel over high-confidence pixels, colour correction ratios for each channel of the ambient image can be estimated. The intensity of the flash can also be adjusted by linearly interpolating the image pair over the YCbCr space. Extrapolation must however be restricted to the Y channel in order to prevent excessive distortion of the hue. Furthermore, the authors note that change in pupil colour between the ambient and flash images can be detected in the red-difference chroma components Cr. By looking for pixels that are highly saturated in this component and relatively dark in the ambient image, potential red-eye regions are identified and later validated by comparing their size and evaluating their eccentricity.

While the ideas presented in this paper are relatively simple, their application has led to interesting results and could benefit most casual photographers. The concepts, although not new, are well introduced and defined, with the possible exception of white balancing which the authors could have elaborated further. Justification could have also been provided for certain design decisions such as the choice of Gaussian functions over other robust weight functions for the (joint) bilateral filter. It would be interesting to see how much the denoising algorithm could benefit from knowing the inherent noise level of the camera, or from simply using a different width σ_r for the edge-stopping function of the blue channel which is generally more subject to noise in yellowish ambient images.

Though the authors mention Fourier techniques could be used to rapidly approximate their joint bilateral filter, they fail to address some important practical considerations. For example, the questionable reliability of the shadow/specularity detection algorithm, the lack of a robust registration technique, and the heavy reliance on user interaction might limit the practicality of the presented solution. Also, the lack of raw image support of most consumer-grade cameras makes the white balancing and shadow detection algorithms unworkable, at least for now.