Flash Photography Enhancement via Intrinsic Relighting

Elmar Eisemann and Frédo Durand
MIT / ARTIS-GRAVIR/IMAG-INRIA and MIT CSAIL

***Abstract*— We enhance photographs shot in dark environments by combining a picture taken with the available light and one taken with the flash. We preserve the ambiance of the original lighting and insert the sharpness from the flash image. We use the bilateral filter to decompose the images into detail and large scale. We reconstruct the image using the large scale of the available lighting and the detail of the flash. We detect and correct flash shadows. This combines the advantages of available illumination and flash photography.**

***Index Terms*— Computational photography, flash photography, relighting, tone mapping, bilateral filtering, image fusion**

I. Introduction

Under dark illumination, a photographer is usually faced with a frustrating dilemma: to use the flash or not. A picture relying on the available light usually has a warm atmosphere, but suffers from noise and blur (Fig. 1(a) top and (b)). On the other hand, flash photography causes three unacceptable artifacts: red eyes, flat and harsh lighting, and distracting sharp shadows at silhouettes (Fig. 1(a) bottom). While much work has addressed red-eye removal [1], [2], the harsh lighting and shadows remain a major impediment.

We propose to combine the best of the two lightings by taking two successive photographs: one with the available lighting only, and one with the flash. We then recombine the two pictures and take advantage of the main qualities of each one (Fig. 1(c)). Our central tool is a decomposition of an image into a large-scale layer that is assumed to contain the variation due to illumination, and a small-scale layer containing albedo variations.

a) Related work: Most work on flash photography has focused on red-eye removal [1], [2]. Many cameras use a pre-flash to prevent red eyes. Professional photographers rely on off-centered flash and indirect lighting to prevent harsh lighting and silhouette shadows.

Our work is related to the continuous flash by Hoppe and Toyama [3]. They use a flash and a no-flash picture and combine them linearly. The image-stack interface by Cohen et al. [4] provides additional control and the user can spatially vary the blending. Raskar et al. [5] and Akers et al. [6] fuse images taken with different illuminations to enhance context and legibility. DiCarlo et al. [7] use a flash and a no-flash photograph for white balance.

Multiple-exposure photography allows for high-dynamic-range images [8], [9]. New techniques also compensate for motion between frames [10], [11]. Note that multiple-exposure techniques are different from our flash-photography approach. They operate on the same lighting in all pictures and invert a non-linear and clamped response. In contrast, we have a quite-different lighting in the two images and try to extract the lighting ambiance from the no-flash picture and combine it with the fine detail of the flash picture.

We build on local tone-mapping techniques that decompose an image into two or more layers that correspond to small- and large-scale variations, e.g. [12], [13], [14], [15], [16], [17], [18], [19]. Only the contrast of the large scales is reduced, thereby preserving detail.

These methods can be interpreted in terms of *intrinsic images* [20], [21]. The large scale can be seen as an estimate of illumination, while the detail corresponds to albedo [22]. Although this type of decoupling is hard [21], [23], [24], tone mapping can get away with a coarse approximation because the layers are eventually recombined. We exploit the same approach to decompose our flash and no-flash images.

A wealth of efforts has been dedicated to relighting, e.g. [25], [26], [27]. Most methods use acquired geometry or a large set of input images. In contrast, we perform lighting transfer from only two images.

Simultaneously but independently from our work, Petschnigg et al. [28] presented a set of techniques based on flash/no-flash image pairs. Their decoupling approach shares many similarities with our work, in particular the use of the bilateral filter. The main difference between the two approaches lies in the treatment of flash shadows.

II. Image Decoupling for Flash Relighting

Our approach is summarized in Fig. 2. We take two photos, with and without the flash. We align the two images to compensate for camera motion between the snapshots. We detect the shadows cast by the flash and correct color using local white balance. We finally perform a non-linear decomposition of the two images into large-scale and detail layers, and we recombine them appropriately.

We first present our basic technique before discussing shadow correction in Section III. We then introduce more advanced reconstruction options in Section IV and present our results in Section V.

b) Taking the photographs: The two photographs with and without the flash should be taken as rapidly as possible to avoid motion of either the photographer or subject. The response curve between the two exposures should ideally be known for better relative radiometric calibration, but this is not a strict requirement. Similarly, we obtain better results

e-mail: eisemann@graphics.csail.mit.edu
email: fredo@mit.edu

(a) (b) (c)

Fig. 1. (a) Top: Photograph taken in a dark environment, the image is noisy and/or blurry. Bottom: Flash photography provides a sharp but flat image with distracting shadows at the silhouette of objects. (b) Inset showing the noise of the available-light image. (c) Our technique merges the two images to transfer the ambiance of the available lighting. Note the shadow of the candle on the table.

when the white balance can be set to manual. In the future, we foresee that taking the two images in a row will be implemented in the firmware of the camera. To perform our experiments, we have used a tripod and a remote control (Fig. 1 and 8) and hand-held shots (Fig. 2, 5, 7). The latter in particular requires good image alignment. In the rest of this paper, we assume that the images are normalized so that the flash image is in $[0,1]$.

The registration of the two images is not trivial because the lighting conditions are dramatically different. Following Kang et al. [10], we compare the image gradients rather than the pixel values. We use a low-pass filter with a small variance (2 pixels) to smooth-out the noise. We keep only the 5% highest gradients and we reject gradients in regions that are too dark and where information is not reliable. We use a pyramidal refinement strategy similar to Ward [11] to find the transformation that minimizes the gradients that were kept. More advanced approaches could be used to compensate for subject motion, e.g. [10].

c) Bilateral decoupling: We first decouple the images into intensity and color (Fig. 2). Assume we use standard formulas, although we show in the appendix that they can be improved in our context. The color layer simply corresponds to the original pixel values divided by the intensity. In the rest of the paper, we use I^f and I^{nf} for the intensity of the flash and no-flash images.

We then want to decompose each image into layers corresponding to the illumination and the sharp detail respectively. We use the bilateral filter [29], [30] that smoothes an image but respects sharp features, thereby avoiding halos around strong edges [16].

The bilateral filter is defined as a weighted average where the weights depend on a Gaussian f on the spatial location, but also on a weight g on the pixel difference. Given an input image I, The output of the bilateral filter for a pixel s is:

$$J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s)\ g(I_p - I_s)\ \ I_p, \qquad (1)$$

where $k(s)$ is a normalization: $k(s) = \sum_{p \in \Omega} f(p-s)\ g(I_p - I_s)$. In practice, g is a Gaussian that penalizes pixels across edges that have large intensity differences. This filter was used by Oh et al. [22] for image editing and by Durand et al. for tone mapping [16].

We use the fast bilateral filter where the non-linear filter is approximated by a set of convolutions [16]. We perform computation in the $\log_{10}$ domain to respect intensity ratios. The output of the filter provides the log of the large-scale layer. The detail layer is deduced by a division of the intensity by the large-scale layer (subtraction in the log domain). We use a spatial variance σ_f of 1.5% of the images diagonal. For the intensity influence g, we use $\sigma_g = 0.4$, following Durand and Dorsey [16].

d) Reconstruction: Ignoring the issue of shadows for now, we can recombine the image (Fig. 2). We use the detail and color layer of the flash image because it is sharper and because white balance is more reliable. We use the large-scale layer of the no-flash picture in order to preserve the mood and tonal modeling of the original lighting situation. The layers are simply added in the log domain. Fig. 3 illustrates the results from our basic approach. The output combines the sharpness of the flash image with the tonal modeling of the no-flash image.

For dark scenes, the contrast of the large scale needs to be enhanced. This is the opposite of contrast reduction [16]. We set a target contrast for the large-scale layer and scale the range of log values accordingly. The low quantization from the original image does not create artifacts because the bilateral filter results in a piecewise-smooth large-scale layer.

In addition, we compute the white balance between the two images by computing the weighted average of the three channels with stronger weights for bright pixels with a white color in the flash image. We then take the ratios w_r, w_g, w_b as white-balance coefficients. This white balance can be used to preserve the warm tones of the available light. In practice, the color cast of the no-flash image is usually too strong and we only apply it partially using w^t where t is usually 0.2.

Fig. 3. Basic reconstruction and shadow correction. The flash shadow on the right of the face and below the ear need correction. In the naïve correction, note the yellowish halo on the right of the character and the red cast below its ear. See Fig. 4 for a close up.

Fig. 4. Enlargement of Fig. 3. Correction of smooth shadows. From left to right: no flash, flash, naïve white balance, our color correction

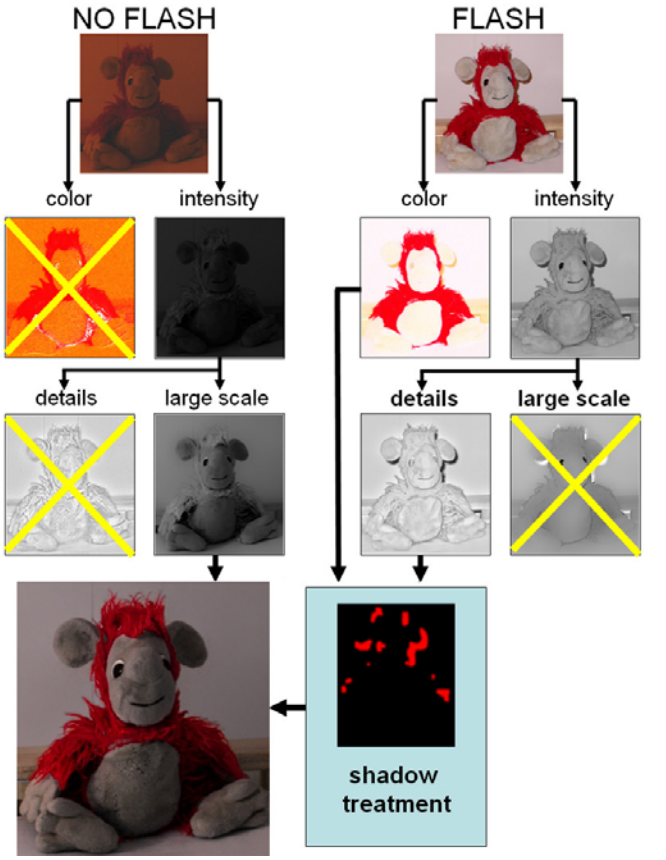


Fig. 2. We take two images with the available light and the flash respectively. We decouple their color, large-scale and detail intensity. We correct flash shadows. We re-combine the appropriate layers to preserve the available lighting but gain the sharpness and detail from the flash image.

We must still improve the output in the flash shadow. While their intensity is increased to match the large scale of the no-flash image, there is a distinct color cast and noise. This is because, by definition, these areas did not receive light from the flash and inherit from the artifacts of the no-flash image. A ring flash might reduce these artifacts, but for most cameras, we must perform additional processing to alleviate them.

III. Shadow treatment

In order to correct the aforementioned artifacts, we must detect the pixels that lie in shadow. Pixels in the umbra and penumbra have different characteristics and require different treatments. After detection, we correct color and noise in the shadows. The correction applied in shadow is robust to false positives; Potential detection errors at shadow boundaries do not create visible artifacts.

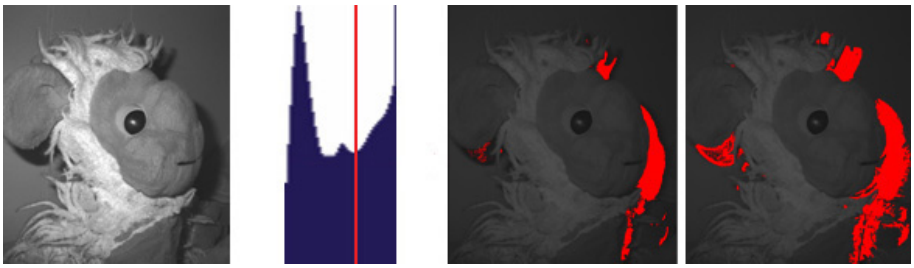
e) Umbra detection: We expect the difference image ΔI between flash and no-flash to tell how much additional light was received from the flash. When the images are radiometrically calibrated, ΔI is exactly the light received from the flash. However, shadows do not always correspond to $\Delta I = 0$ because of indirect lighting. While shadow pixels always correspond to the lowest values of ΔI, the exact cutoff is scene-dependent.

We use histogram analysis to compute a threshold $t_{\Delta I}$ that determines umbra pixels. Shadows correspond to a well-marked mode in the histogram of ΔI. While the additional light received by parts of the scene lit by the flash varies with albedo, distance and normal, the parts in shadow are only indirectly illuminated and receive a more uniform and very low amount of light.

We compute the histogram of pixels ΔI. We use 128 bins and smooth it with a Gaussian blur of variance two bins. We start with a coarse threshold of 0.2 and discard all pixels where ΔI is above this value. We then use the first local minimum of the histogram before 0.2 as our threshold for shadows detection (Fig. 5). This successfully detects pixels in the umbra. However, pixels in the penumbra correspond to a smoother gradation and cannot be detected with our histogram technique. This is why we use a complementary detection based on the gradient at shadow boundaries.

f) Penumbra detection: Shadow boundaries create strong gradients in the flash image that do not correspond to gradients in the no-flash image. We detect these pixels using two criteria: the gradients difference, and connectedness to umbra pixels.

We compute the magnitude of the gradient ∇I^f and ∇I^{nf} and smooth it with a Gaussian of variance 2 pixels to remove



ΔI histogram of ΔI detected umbra with penumbra

Fig. 5. Shadow detection

noise. We identify candidate penumbra pixels as pixels where the gradient is stronger in the flash image. We then keep only pixels that are "close" to umbra pixels, that is, such that at least one of their neighbors is in umbra. In practice, we use a square neighborhood of size 1% of the photo's diagonal. This computation can be performed efficiently by convolving the binary umbra map with a box filter.

We also must account for shadows cast by tiny objects such as pieces of fur, since these might have a pure penumbra without umbra. We use a similar strategy and consider as shadow pixels that have a large number of neighbors with higher gradient in the flash image. We use a threshold of 80% on a square neighborhood of size 0.7% of the photo's diagonal.

We have observed that the parameters concerning the penumbra are robust with respect to the scene. The image-space size of the penumbra does not vary much in the case of flash photography because the distance to the light is the same as the distance to the image plane. The variation of penumbra size (ratio of blocker-receiver distances) and perspective projection mostly cancel each other.

g) Flash detail computation: Now that we have detected shadows, we can refine the decoupling of the flash image. We exploit the shadow mask to exclude shadow pixels from the bilateral filtering. This results in a higher-quality detail layer for the flash image because it is not affected by shadow variation.

h) Color and noise correction: Color in the shadow cannot simply be corrected using white balance [7] for two reasons. First, shadow areas receive different amounts of indirect light from the flash, which results in hybrid color cast affected by the ambient lighting and color bleeding from objects. Second, the no-flash image often lacks information in the blue channel due to the yellowish lighting and poor sensitivity of sensors in the small wavelengths. Fig. 3 illustrates the artifacts caused by a global white balance of the shadow pixels.

In order to address these issues, we use a *local* color correction that copies colors from illuminated regions in the flash image. For example, in Fig. 3, a shadow falls on the wall, sofa frame and jacket. For all these objects, we have pixels with the same intrinsic color in the shadow and in the illuminated region.

Inspired by the bilateral filter, we compute the color of a shadow pixel as a weighted average of its neighbors in the flash image $\mathbf{I}^f$ (with full color information). The weight depends on three terms: a spatial Gaussian, a Gaussian on the color similarity in $\mathbf{I}^{nf}$, and a binary term that excludes pixels in shadow (Fig. 6). We perform computation only on the color layer (see Fig. 2) in Luv. We use σ_f of 2.5% of the photo's diagonal for the spatial Gaussian and $\sigma_g = 0.01$ for the color similarity. As described by Durand and Dorsey [16] we use the sum of the weights k as a measure of pixel uncertainty. We discard color correction if k is below a threshold. In practice, we use a smooth feathering between 0.02 and 0.002 to avoid discontinuities.

Recall that the large-scale layer of intensity is obtained from the no-flash image and is not affected by shadows. In the shadow, we do not use the detail layer of the flash image because it could be affected by high-frequencies due to shadow boundary. Instead, we copy the detail layer of the no-flash image, but we correct its noise level. For this we scale the no-flash detail to match the variance of the flash detail outside shadow regions.

In order to ensure continuity of the shadow correction, we use feathering at the boundary of the detected shadow: We follow a linear ramp and update pixels as a linear combination of the original and shadow-corrected value. Fig. 3 and 4 show the results of our shadow correction. It is robust to false shadow positives because it simply copies colors from the image. If a pixel is wrongly classified in shadow, its color and noise are preserved as long as there are other pixels with similar color that were not classified in shadow.

IV. Advanced Decoupling

The wealth of information provided by the pair of images can be further exploited to enhance results for very dark situations and more advanced lighting transfer.

When the no-flash picture is too dark, the edge-preserving property of the bilateral filter is not reliable, because noise level is in the range of the signal level. Similar to the technique we use for color correction, we can use the flash image as a similarity measure between pixels. We propose a *cross-bilateral filter*[1] where we modify Eq. 1 for the no-flash image and compute the edge-preserving term g as a function of the flash-image values:

$$J_s^{nf} = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s)\ g(I_p^f - I_s^f)\ \ I_p^{nf}, \qquad (2)$$

This preserves edges although they are not really present in the no-flash image. Shadow correction can however not

[1] Petschnigg et al. [28] propose a similar approach that they call *joint* bilateral filter.

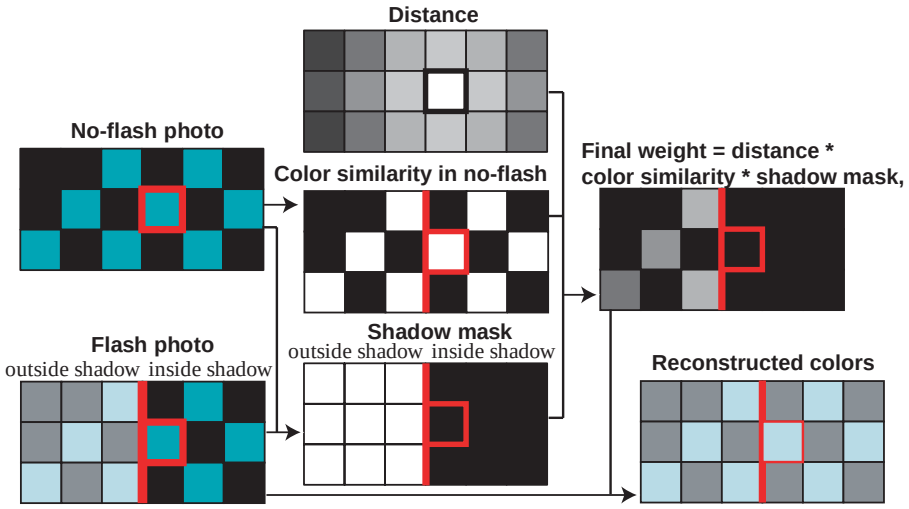


Fig. 6. For a pixel in the flash shadow, the color layer is computed as a weighted average of non-shadow colors. The weights depend on three terms: distance, similarity in the no-flash image and a binary shadow mask.

be performed because the shadow edges of the flash picture are transferred by the g term. Fig. 1 exploits cross-bilateral decomposition.

The large-scale layer of the flash image can also be exploited to drive the reconstruction. The distance falloff makes objects closer to the camera brighter. We use this pseudo-distance to emphasize the main object. We use a shadow-corrected version of ΔI as our pseudo-distance. Pixels in shadow are assigned a pseudo-distance using a bilateral-weighted average of their neighbors where similarity is defined in the no-flash image. The principle is to multiply the large scale of the no-flash image by the pseudo-distance. This can be performed using a user-provided parameter. Pseudo-distance was used in Fig. 8.

V. Results and discussion

Our technique takes about 50 seconds on a 866 MHz Pentium 3 for a 1280x960 image. The majority of the time is spent in the color correction, because this bilateral filter cannot be efficiently piecewise-linearized [16] since it operates on the three channels. Images such as Fig. 8 that do not include shadow correction take about 10 seconds.

Fig 1, 3, 7 and 8 illustrate our results. The ambiance of the available light is preserved and the color, sharpness and detail of the flash picture is gained. In our experience, the main cause of failure of our technique is poor quality (not quantity) of available lighting. For example, if the light is behind the subject, the relighting results in an under-exposed subject. We found, however, that it is not hard to outperform the poor lighting of the flash. It is well known that lighting along the optical axis does not result in good tonal modeling. In contrast, Fig. 2 and 8 present a nice 3/4 side lighting. We received conflicting feedback on Fig. 7, which shows that image quality is a subjective question. In this image, the light is coming from the 3/4 back, which is an unusual lighting for a photograph. Some viewers appreciate the strong sense of light it provides, while others object to the lack of tonal modeling.

Another cause of failure is overexposure of the flash, leading to a flat detail layer. In this situation, the detail information is neither in the no-flash (due to noise) nor in the flash photo (due to saturation).

Shadow detection works best when the depth range is limited. Distant objects do not receive light from the flash and are detected in shadow. While this is technically correct, this kind of shadow due to falloff does not necessitate the same treatment as cast shadow. Fortunately, our color correction is robust to false positives and degrades to identity in these cases (although transition areas could potentially create problems). Similarly, black objects can be detected as shadows, but this does not affect quality since they are black in the two images and remain black in the output. Light flares can cause artifacts by brightening shadow pixels. The method by Ward [11] could alleviate this problem.

We have used our algorithms with images from a variety of cameras including a Sony Mavica MVC-CD400 (Fig. 1), a Nikon Coolpix 4500 (all other images), a Nikon D1 and a Kodak DC4800 (not shown in the paper). The choice of the camera was usually dictated by availability at the time of the shot. The specifications that affected our approach are the noise level, the flexibility of control, the accuracy of flash white balance, and compression quality. For example, the Kodak DC4800 exhibited strong JPEG artifacts for dark images, which required the use of the cross-bilateral filter.

The need for the cross-bilateral filter was primarily driven by the SNR in the no-flash picture. The Kodak DC4800 has higher noise levels because it is old. Despite its age, the size of its photosites allows the Nikon D1 to take images in dark conditions. In addition, the use of the RAW format with 12 bits/channel allows for higher precision in the flash image (the lower bits of the no-flash image are dominated by noise). However, with the sensitivity at 1600 equivalent ISO, structured noise makes cross-bilateral filtering necessary.

VI. Conclusions and future work

We have presented a method that improves the lighting and ambiance of flash photography by combining a picture taken with the flash and one using the available lighting. Using a feature-preserving filter, we estimate what can be seen as

no-flash flash result

Fig. 7. The flash lighting results in a flat image. In our result, light seems to be coming from the window to the right.

Fig. 8. The tonal modeling on the cloth and face are accurately transferred from the available lighting. The main subject is more visible in the result than he was in the original image.

intrinsic layers of the image and use them to transfer the available illumination to the flash picture. We detect shadows cast by the flash and correct their color balance and noise level. Even when the no-flash picture is extremely noisy, our method successfully transfers lighting due to the use of the flash image to perform edge-preserving filtering.

The method could be tailored to particular cameras by fine-tuning parameters such as σ_g based on a sensor-noise model. Traditional red-eye removal could benefit from the additional information provided by the pair of images. Texture synthesis and in-painting could be used to further improve shadow correction. Ideally, we want to alleviate the disturbance of the flash and we are considering the use of infrared illumination. This is however challenging because it requires different sensors and these wavelengths provide limited resolution and color information.

The difference of the flash and no-flash images contains much information about the 3D scene. Although a fundamental ambiguity remains between albedo, distance and normal direction, this additional information could greatly expand the range and power of picture enhancement such as tone mapping, super-resolution, photo editing, and image based-modeling.

i) Acknowledgments: We acknowledge support from an NSF CISE Research Infrastructure Award (EIA-9802220) and a Deshpande Center grant. Elmar Eisemann's stay at MIT was supported by MIT-France and ENS Paris. Many thanks to the reviewers, Joëlle Thollot, Marc Lapierre, Ray Jones, Eric Chan, Martin Eisemann, Almuth Biard, Shelly Levy-Tzedek, Andrea Pater and Adel Hanna.

APPENDIX

a) Appendix: Intensity-Color decoupling: Traditional approaches rely on linear weighted combinations of R, G, and B for intensity estimation. While these formulae are valid from a color-theory point of view, they can be improved for illumination-albedo decoupling. Under the same illumination, a linear intensity computation results in lower values for primary-color albedo (in particular blue) than for white objects. As a result, the intensity transfer might overcompensate as shown in Fig. 9(left) where the red fur becomes too bright. To alleviate this, we use the channels themselves as weights in the linear combination:
$I = \frac{R}{R+G+B}R + \frac{G}{R+G+B}G + \frac{B}{R+G+B}B.$

In practice, we use the channels of the flash image as weight for both pictures to ensure consistency between the two decoupling operations. The formula can also be used with tone mapping operators for higher color fidelity.

Fig. 9. The computation of intensity from RGB can greatly affect the final image. Left: with linear weights, the red pixels of the fur become too bright. Right: using our non-linear formula.

REFERENCES

[1] Zhang and Lenders, "Knowledge-based eye detection for human face recognition,," in *Conf. on Knowledge-based Intelligent Systems and Allied Technologies,*, 2000.

[2] Gaubatz and Ulichney, "Automatic red-eye detection and correction," in *IEEE Int. Conf. on Image Processing*, 2002.

[3] Hoppe and Toyama, "Continuous flash," MSR, Tech. Rep. 63, 2003.

[4] Cohen, Colburn, and Drucker, "Image stacks," MSR, Tech. Rep. 40, 2003.

[5] Raskar, Ilie, and Yu, "Image fusion for context enhancement," in *Proc. NPAR*, 2004.

[6] Akers, Losasso, Klingner, Agrawala, Rick, and Hanrahan, "Conveying shape and features with image-based relighting," in *Visualization*, 2003.

[7] J. M. DiCarlo, F. Xiao, and B. A. Wandell, "Illuminating illumination," in *9th Color Imaging Conference*, 2001, pp. 27–34.

[8] Mann and Picard, "Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *Proc. IS&T 46th ann. conference*, 1995.

[9] Debevec and Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. SIGGRAPH*, 1997.

[10] Kang, Uyttendaele, Winder, and Szeliski, "High dynamic range video," *ACM Trans. on Graphics*, vol. 22, no. 3, 2003.

[11] Ward, "Fast, robust image registration for compositing high dynamic range photographs from handheld exposures," *J. of Graphics Tools*, vol. 8, no. 2, 2004.

[12] Chiu, Herf, Shirley, Swamy, Wang, and Zimmerman, "Spatially nonuniform scaling functions for high contrast images," in *Graphics Interface*, 1993.

[13] Jobson, Rahman, and Woodell, "A multi-scale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. on Image Processing*, vol. 6, pp. 965–976, 1997.

[14] Tumblin and Turk, "Lcis: A boundary hierarchy for detail-preserving contrast reduction," in *Proc. SIGGRAPH*, 1999.

[15] J. DiCarlo and B. Wandell, "Rendering high dynamic range images," *Proc. SPIE: Image Sensors*, vol. 3965, pp. 392–401, 2000.

[16] Durand and Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. on Graphics*, vol. 21, no. 3, 2002.

[17] Reinhard, Stark, Shirley, and Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. on Graphics*, vol. 21, no. 3, 2002.

[18] Ashikhmin, "A tone mapping algorithm for high contrast images," in *Eurographics Workshop on Rendering*, June 2002, pp. 145–156.

[19] Choudhury and Tumblin, "The trilateral filter for high contrast images and meshes," in *Eurographics Symposium on Rendering*, 2003.

[20] Tumblin, Hodgins, and Guenter, "Two methods for display of high contrast images," *ACM Trans. on Graphics*, vol. 18, no. 1, 1999.

[21] Barrow and Tenenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*. Academic Press, 1978.

[22] Oh, Chen, Dorsey, and Durand, "Image-based modeling and photo editing," in *Proc. SIGGRAPH*, 2001.

[23] Weiss, "Deriving intrinsic images from image sequences," in *ICCV*, 2001.

[24] M. F. Tappen, W. T. Freeman, and E. H. Adelson, "Recovering intrinsic images from a single image," in *NIPS*, 2003.

[25] Marschner and Greenberg, "Inverse lighting for photography," in *Proc. IS&T/SID 5th Color Imaging Conference*, 1997.

[26] Sato, Sato, and Ikeuchi, "Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions," in *ICCV*, 1999. [Online]. Available: citeseer.nj.nec.com/sato99illumination.html

[27] Yu, Debevec, Malik, and Hawkins, "Inverse global illumination: Recovering reflectance models of real scenes from photographs from," in *Proc. SIGGRAPH*, 1999. [Online]. Available: citeseer.nj.nec.com/yu99inverse.html

[28] Petschnigg, Agrawala, Hoppe, Szeliski, Cohen, and Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. on Graphics*, vol. in this volume, 2004.

[29] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV*, 1998, pp. 836–846.

[30] S. M. Smith and J. M. Brady, "SUSAN - a new approach to low level image processing," *IJCV*, vol. 23, pp. 45–78, 1997.