IMAGE ENHANCEMENT METHOD VIA BLUR AND NOISY IMAGE FUSION

Marius Tico, Kari Pulli

Nokia Research Center, Palo Alto, CA, USA

ABSTRACT

We present an image enhancement algorithm based on fusing the visual information present in two images of the same scene, captured with different exposure times. The main idea is to exploit the differences between the image degradations that affect the two images. On one hand the short-exposed image is less affected by motion blur, whereas the long-exposed image is less affected by noise. Different fusion rules are designed for the luminance and chrominance components to preserve the desirable properties from each input image. We also present a method for estimating the brightness transfer function between the input images. As no global blur PSF is assumed, our method can deal with blur from both camera and object motions. We demonstrate the algorithm by a series of experiments and simulations.

Index Terms— low light imaging, motion blur, noise filtering, photometric calibration, brightness transfer function

1. INTRODUCTION

Image degradation due to motion blur is caused by a relative motion between the camera and the scene during the image integration time. This degradation is particularly significant when long exposure times are needed, like for instance in low-light conditions, or in high-dynamic-range (HDR) imaging.

Several solutions to image blur degradation have been proposed. If there is a blur point spread function (PSF) that is constant over the image and that is known, the original image of the scene can be restored, to some extent, by applying an image de-convolution routine. In such a case, the key challenge in restoring the original image is caused by the zeros of the PSF in the frequency domain, resulting in lost spatial frequencies in the original image. However, in most practical situations the motion blur PSF is not known, and since it depends of the arbitrary camera motion during the exposure time, it is also different for each image. One can try using blind de-convolution approaches to restore the motion blurred images [2, 3]. Unfortunately, most such methods are computationally prohibitive for consumer devices equipped with limited computational resources, such as camera phones. Measurements of the camera motion during the exposure time could help in estimating the motion blur PSF and eventually in restoring the original image of the scene. Ezra and Nayar [4] used a secondary video camera to estimate the motion during the exposure time of the principal camera. Others have estimated the PSF [5, 6, 7] from a second image with a short exposure. Although noisy, the secondary image is much less affected by motion blur and it can be used as a reference for estimating the motion blur PSF which degraded the principal image. However, even if the camera is completely stationary during the image capture, objects such as people may move. This results in a spatially variant blur PSF making the restoration problem much more difficult if not even impossible with current techniques.

We propose a solution based on fusing visual information in two differently exposed images: a short-exposed image affected by noise, and a long-exposed image that may be affected by motion blur. Unlike the earlier solutions [5, 6, 7], our result is obtained without de-convolving the long-exposed image, which would be too computationally expensive for some devices and infeasible for a spatially varying PSF. Our method can deal both with camera and object motion blur in the long-exposed image. Our image fusion method gathers the desired visual properties from both the noisy and blurry images into a final, higher-quality picture.

2. THE PROPOSED ALGORITHM

Our algorithm can be summarized in the following three steps: (i) image registration, (ii) photometric calibration, and (iii) image fusion. Image registration should be achieved with an algorithm that is able to cope with the different degradations affecting the two images, as well as with their different brightness. We used the algorithm proposed in [9]. This paper describes the remaining two steps, namely photometric calibration and image fusion.

The photometric calibration, presented in Sec. 2.1, is a preprocessing step that compensates for different exposures, as well as for other differences between the processing of the two images. Most processing steps such as color matrix interpolation and gamma correction introduce many non-linearities, but especially auto-white balancing result often in different color shades and temperatures when the amount of light coming to the sensor varies. In our work we calibrate each color channel separately by estimating a so-called brightness transfer function (BTF) for each of them. We chose to do this operation in RGB color space because there is typically a large, yet different correlation between the two differently exposed images within each of the three color channels. In contrast, there is no similar clear correlation of the chrominance channels of CIE L*a*b*, or YUV color spaces.

The image fusion procedure should preserve the edge sharpness from the short-exposed image and the reduced noise characteristics of the long-exposed image, both reflected mainly in the luminance data of the two images. Colors, however, are much better sampled in the long-exposed image, and because of the way the human visual system works, the blur is less perceived in the chrominance data than in the luminance. We propose thereby different rules for fusing the luminance and chrominance data as detailed in Sections 2.2 and 2.3.

2.1. Photometric calibration

For each color channel, we estimate a brightness transfer function (BTF) by analyzing the joint histogram of pixel values in the two images, also called *comparagram* [12]. The advantages of doing so, instead converting both images in the radiance domain (e.g., [10, 11]) are that we do not need precise knowledge of the individual image exposures or prior knowledge about the imaging system response

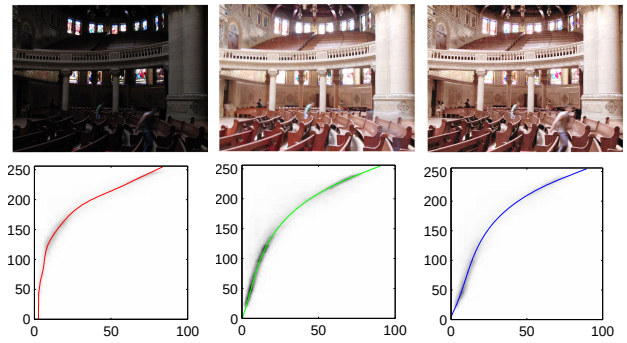

Fig. 1. Photometric calibration example.

function. In addition, the separate BTF estimation for each color channel automatically compensates for color balance differences between the two images, which we observed that often occurs on images of the same scene with different exposure times.

We expect the BTF of a single color channel between the two images to be a smooth monotonically increasing function between the limits of saturation and under-exposure. We estimate the BTF from the comparagram that has been cleaned from saturated pixels by fitting a low-order polynomial to the data. Fig. 1 shows an example of photometric calibration achieved with the algorithm presented in this section. On the top row, from left to right, the figure represents: the short and long-exposed images, followed by the photometrically calibrated short-exposed image. On the bottom row are represented the estimated BTF curves of the three color channels overlapped over their corresponding comparagrams. The horizontal and vertical axes in these plots correspond to short- and long-exposed image respectively.

The following steps summarize our BTF estimation algorithm:

1. Estimate the comparagram between the two images as a $K \times K$ array, where K denotes the number of intensity levels (typically 256). On large images it is sufficient to sub-sample the images (e.g., every 4^{th} row and column).

2. Calculate a smooth version of the comparagram by applying a spatial low-pass filter onto the original $K \times K$ joint histogram.

3. Truncate small entries (threshold = 1% of the maximum value) as they have the most noise, and large entries (> 245) as they are getting saturated, from the comparagram. 10.

4. Identify a set of points that are likely to belong to the BTF. We interpret the comparagram as a joint probability density function (pdf) between two random variables X and Y, i.e., $p(x, y)$ denotes the comparagram entry at $x, y \in \{0, \ldots, K-1\}$, where Y represents the brightness in the long-exposed image, and X represents the brightness in the short exposed image. Now, as Y has larger entropy, for each value y_i of Y the most probable value x_i of X is the minimum mean squared error estimate (MMSE) $y_i = E[y_i|x_i]$, which is the weighted average $x_i = \sum_{x=0}^{K-1} x\, p(x, y_i) \,/\, \sum_{x=0}^{K-1} p(x, y_i)$.

5. Fit a smooth curve over the set of selected points. We first calculate the principal component axes of the selected points $\{(x_i, y_i)\}$, and transform their coordinates into the system defined by their principal axes. Next, a least-squares polynomial curve is fitted in the new system to approximate the selected points. In our work we used a 5th order polynomial for this approximation.

6. Returning into the original coordinate system, we need to ensure the monotonicity of the estimated polynomial curve segment, and to extrapolate it to cover the entire range $(0, \ldots, K-1)$ of Y that corresponds to the longer exposed image. Defining x_k to correspond to each value $k \in \{0, \ldots, K-1\}$ such that the point (x_k, k) is a point on the estimated curve segment, we analyze the set of values x_k starting from the middle, i.e., $x_{K/2}$, in both directions, towards x_{K-1} and towards x_0. Whenever we encounter a value that breaks the monotonicity requirement, i.e., $x_\ell < x_{\ell-1}$, the value is re-estimated by fitting a line onto a number of previously analyzed points. In our experiments we used up to 10 points to perform this linear fitting.

7. Finally, the BTF is defined by the modified set of points $\{(x_k, k)\}$, which are converted into a look-up table for faster computation.

2.2. Luminance fusion

Once the two input images are geometrically and photometrically aligned, the differences between their intensities are due to the presence of noise in the short-exposed image, and blur in the long-exposed image. The absolute difference between the two images is typically larger close to the image edges than in smooth image areas, as illustrated in Fig. 2. Therefore we aim for an estimator that emphasizes more the short-exposed image where the absolute difference between the two images is larger, and the long-exposed image otherwise.

Formally, denoting the two images as g_l and g_s, we can write the following model:

$$g_l(\mathbf{x}) = f_b(\mathbf{x}), \text{ and } g_s(\mathbf{x}) = f(\mathbf{x}) + n(\mathbf{x}), \tag{1}$$

where $\mathbf{x} = (x, y)$ denotes the coordinates of an image pixel, f denotes the original image, f_b denotes a blurred version of the original image, and $n(\mathbf{x})$ denotes a zero-mean additive Gaussian noise term.

In order to achieve a better separation between signal and noise we derive an image estimator in the wavelet domain. The edge locations (i.e., large values in the difference signal), are emphasized at specific scales, whereas the noise variance is evenly distributed across the scale space. Considering an orthonormal wavelet transform of the two images, and denoting by $G_i(k)$ the k-th wavelet coefficient of an image, the model (1) becomes

$$G_l(k) = F_b(k), \text{ and } G_s(k) = F(k) + N(k), \tag{2}$$

where $F_b(k)$ and $F(k)$ denote the k-th wavelet coefficients of the blurred and original images, and N is an additive white Gaussian noise of variance σ^2.

We can now fuse the images together using different weights at different scales. Taking advantage of the de-correlation in the wavelet domain, we propose an MMSE diagonal estimator of the original image in the form of a linear combination between the wavelet coefficients of the two images

$$\hat{F}(k) = G_s(k) + W(k)D(k), \tag{3}$$

where $\hat{F}(k)$ stands for the wavelet coefficients of the restored image, $D(k) = G_l(k) - G_s(k)$ denotes the difference signal between the wavelet coefficients of the two observed images, and $W(k)$ is a weight value. We can estimate the best weight $W(k)$ for each wavelet coefficient by minimizing the mean squared error

$$E[||\hat{F}(k) - F(k)||_2^2] = E[||G_s(k) - F(k) + W(k)D(k)||_2^2], \tag{4}$$

whose derivative with respect to $W(k)$, equated with zero gives

$$W(k)E[|D(k)|^2] = E[D(k)N(k)]. \tag{5}$$

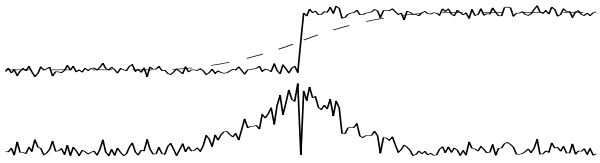

Fig. 2. Absolute difference (bottom plot) between the noisy and blurred versions of a step function (upper plot).

Fig. 3. Color weighting: weight function, short-exposed image, weight image, and a long-exposed image.

Expanding the right hand side of this equation we have

$$E[D(k)N(k)] = E[(F_b(k) - F(k))N(k)] + E[|N(k)|^2], \quad (6)$$

where using the assumption of white Gaussian noise model for $N(k)$, we can approximate the first term with zero and the second term with σ^2. Finally, replacing this result in (Eq. 5) we obtain

$$W(k) = \sigma^2 / E[|D(k)|^2]. \quad (7)$$

The computation of the weight $W(k)$ requires an estimate of the noise variance in the short-exposed image, and an estimate of the expectation $E[|D(k)|^2]$. In order to estimate the noise variance in the short-exposed image we extend the approach presented by Mallat [8] (p. 459). Noting that in practice the noise is spatially variant over the image, our extension consists of applying the wavelet-based noise estimate in the pixel neighborhood (e.g., 7×7). Finally, we approximate $E[|D(k)|^2]$ with $\max(\sigma^2(k), avg(|D(k)|^2))$, where avg stands for local spatial average, and $\sigma^2(k)$ is the noise variance at the spatial location that corresponds to the k-th wavelet coefficient.

The weight $W(k)$ emphasizes the short-exposed image in areas where the absolute difference signal is larger than the noise variance. Consequently, in accordance to our initial observation, the proposed estimator acts by emphasizing the short-exposed image close to the image edges, and the long-exposed image in the smooth areas.

2.3. Chrominance fusion

As discussed before, we want to fuse colors in a color space that de-correlates chrominance from luminance. We experimented with both YUV and CIE L*a*b* color spaces without noticing much difference, our experiments shown in this paper use the CIE L*a*b* color space.

We fuse the two color channels based on a weighted average between their values in the two images, where the weight assigned to the short-exposed image is a function of the fused luminance value: $w_s(L) = (L/255)^6$, and the weight of the long-exposed image is $w_l(L) = 1 - w_s(L)$, for $L \in [0, 255]$. This results in taking almost all the color from the long-exposed image, except the areas where this image is almost saturated. Figure 3 shows the weight for the short-exposed image colors as a function of the estimated luminance, the boosted short-exposed image, the per-pixel weight for the short-exposed image, and the long-exposed image.

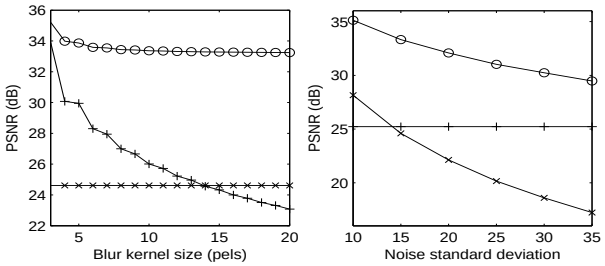

Fig. 4. Performance for different degradation levels: restored image (o), noisy image (x), and blurry image (+).

3. EXPERIMENTAL RESULTS

A first set of experiments has been conducted in order to verify the performance of the estimate (Eq. 3), proposed for luminance fusion. We simulated pairs of noisy and blur images by adding white Gaussian noise and smoothing the original image with uniform blur kernels. Figure 4 shows the performance achieved at different levels of degradations in the two input images. We note that the restoration result is significantly improved with respect to both input images. We remark also a lower sensitivity to the amount of blur, which suggests that after some level of blur, the longer exposed image has less contribution to the result than the shorter exposed image. Next, we carried out a series of comparisons with respect to different noise filtering approaches applied onto the short-exposed image. The results of these comparisons are shown in Table 1, and they reveal the effectiveness of the proposed algorithm compared to de-noising the short-exposed image independent of the long-exposed image. In this case we used a 10×10 uniform blur kernel for simulating the blurry image. The de-noising approaches used for comparison are Matlab's spatial local Wiener filtering, hard thresholding of wavelet coefficients [14], and hard thresholding of curvelet coefficients [15].

	Noise standard deviation					
	10	15	20	25	30	35
Lena (blurred image: 26 dB)						
Noisy	28.13	24.62	22.12	20.18	18.61	17.24
Proposed	**35.19**	**33.42**	**32.18**	**31.15**	**31.00**	**30.58**
Wiener	33.71	31.21	29.11	27.31	25.81	24.48
wavelet	29.92	28.38	27.24	26.44	25.73	25.17
curvelet	33.70	32.32	31.27	30.41	29.68	29.07
Barbara (blurred image: 22 dB)						
Noisy	28.13	24.62	22.09	20.17	18.58	17.26
Proposed	**33.29**	**30.98**	**29.40**	**28.25**	**27.35**	**27.00**
Wiener	29.79	28.22	26.73	25.43	24.26	23.22
wavelet	26.12	24.51	23.51	22.80	22.33	21.89
curvelet	29.15	26.62	25.30	24.68	24.31	24.00

Table 1. PSNR (dB) results achieved with different approaches.

Two results of the proposed algorithm on real examples are shown in Figs. 5 and 6. In both figures we may note that the result image is highly improved in comparison with both inputs. Our algorithm can also deal with spatially varying blur in the long-exposed image, as seen in Fig. 6 (top-right image), where the blur caused by moving people in the lower-left corner of the image is very different from the blur process in the rest of the image. This is a major advantage compared to earlier methods for blur/noisy image combination (e.g., [7]). These methods cannot solve blur caused by moving objects because they are based on the assumption that the blur process is spatially invariant.

Fig. 5. Image enhancement result: short- and long-exposed images (upper row), the photometrically calibrated short-exposed image (bottom left), and the result (bottom right).

Fig. 6. See the caption of Fig. 5.

4. CONCLUSIONS

We introduced an approach to image fusion that relies on preserving the desirable properties found in two differently exposed images. The luminance channels are fused using rules emphasizing the sharpness of the short-exposed image while denoising using information from the long-exposed image, and the color channels are fused by mostly taking the better color data of the long-exposed image, except where its colors are about to saturate, in which case we get the colors mostly from the photometrically calibrated colors of the short-exposed image. In contrast to earlier approaches of blurry/noisy image pair combination we neither apply an image de-convolution procedure nor assume a spatially invariant uniform blur model for the long-exposed image, allowing us to deal with both camera and object motion blur simultaneously. The proposed algorithm has been demonstrated through several experiments and comparisons.

5. REFERENCES

[1] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

[2] Tony F. Chan and Chiu-Kwong Wong, "Total Variation Blind Deconvolution," *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998.

[3] Yu-Li You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," *IEEE Trans. on Image Processing*, vol. 5, no. 3, pp. 416–428, Mar. 1996.

[4] Moshe Ben-Ezra and Shree K. Nayar, "Motion-Based Motion Deblurring," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, 2004.

[5] Marius Tico, Mejdi Trimeche, and Markku Vehviläinen, "Motion blur identification based on differently exposed images," in *ICIP*, 2006, pp. 2021–2024.

[6] Marius Tico and Markku Vehviläinen, "Image stabilization based on fusing the visual information in differently exposed images," in *ICIP*, 2007, vol. 1, pp. 117–120.

[7] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum, "Image deblurring with blurred/noisy image pairs," *ACM Trans. on Graphics*, vol. 26, no. 3, 2007.

[8] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, USA, 1990.

[9] Marius Tico, Sakari Alenius, and Markku Vehviläinen, "Method of motion estimation for image stabilization," in *ICASSP*, 2006, vol. 2, pp. 277–280.

[10] Paul E. Debevec and Jitendra Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," in *SIGGRAPH*, 1997.

[11] Tomoo Mitsunaga and Shree K. Nayar, "Radiometric self calibration," in *CVPR*, 1999.

[12] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Trans. Image Processing*, vol. 9, no. 8, pp. 1389–1406, 2000.

[13] Seon Joo Kim and Marc Pollefeys, "Robust radiometric calibration and vignetting correction," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 562–576, 2008.

[14] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[15] Jean-Luc Starck, Emmanuel J. Candes, and David L. Donoho, "The Curvelet Transform for Image Denoising," *IEEE Trans. on Image Processing*, vol. 11, no. 6, pp. 670–684, 2002.