

Bayesian Correction of Image Intensity with Spatial Consideration^{*}

Jiaya Jia¹, Jian Sun², Chi-Keung Tang¹, and Heung-Yeung Shum²

¹ Computer Science Department,
Hong Kong University of Science and Technology,
{leo.jia, cktang}@cs.ust.hk
² Microsoft Research Asia,
{t-jiansu, hshum}@microsoft.com

Abstract. Under dimly lit condition, it is difficult to take a satisfactory image in long exposure time with a hand-held camera. Despite the use of a tripod, moving objects in the scene still generate ghosting and blurring effect. In this paper, we propose a novel approach to recover a high-quality image by exploiting the tradeoff between exposure time and motion blur, which considers color statistics and spatial constraints simultaneously, by using only two defective input images. A Bayesian framework is adopted to incorporate the factors to generate an optimal color mapping function. No estimation of PSF is performed. Our new approach can be readily extended to handle high contrast scenes to reveal fine details in saturated or highlight regions. An image acquisition system deploying off-the-shelf digital cameras and camera control softwares was built. We present our results on a variety of defective images: global and local motion blur due to camera shake or object movement, and saturation due to high contrast scenes.

1 Introduction

Taking satisfactory photos under weak lighting conditions using a hand-held camera is very difficult. In this paper, we propose a two-image approach to address the image recovery problem by performing intensity correction. In order to exploit the tradeoff between the exposure time and the blurring degree of the captured images, we take the two input images using the same camera with the following exposure settings:

- One image I_L is taken with exposure time around the safe shutter speed¹, producing an under-exposed image where motion blur is largely reduced. Since it is too dark, the colors in the image are not acceptable (Fig. 1(a)).

^{*} This work is supported by the Research Grant Council of Hong Kong Special Administration Region, China: HKUST6193/02E.

¹ In photography, the safe shutter speed is assumed to be not slower than the reciprocal of the focal length of the lens, in the unit of seconds [1]. The longer the exposure time, the blurrier the image becomes.

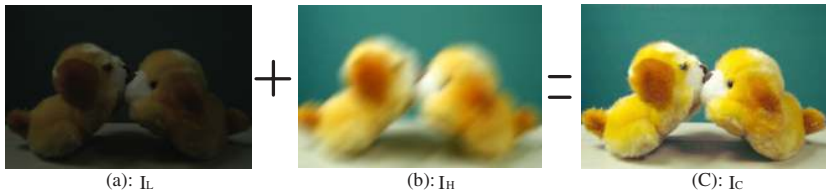


Fig. 1. We take two successive images with different exposure intervals to construct the high-quality image.

- The other image I_H is a normal image acquired under an extended exposure time. The color and brightness of this image is acceptable. However, it is motion blurred because of camera shaking or moving objects in the scene (Fig. 1(b)).

The images can be taken by a hand-held camera, and possibly in a dimly lit condition. Combining these two defective images I_L and I_H , our method automatically generates a clear and crisp image I_C , as shown in Fig. 1(c).

There are several related techniques to recover images from camera when exposure time is above the safe shutter speed. They can be roughly classified into *in-process* and *post-process* approaches, which eliminate motion blur due to long exposure and camera shake. In-process approaches are mainly hardware-based techniques, where lens stabilization is achieved by camera shake compensation [8,9]. Alternatively, CMOS cameras can perform high-speed frame captures within normal exposure time, which allows for multiple image-based motion blur restoration [11]. These methods are able to produce clear and crisp images, given a reasonable exposure time. However, they require specially designed hardware devices.

On the other hand, post-process methods are mostly motion deblurring techniques. Among them, blind deconvolution is widely adopted to enhance a single blurred image, under different assumptions on the PSF [6,15,10,14]. Alternatively, several images with different blurring directions [12] or an image sequence [2] is used, in more general situations, to estimate the PSF. In both cases, due to the discretization and quantization of images in both spatial and temporal coordinates, the PSF can not be reliably estimated, which produced a result inferior to the ground truth image if available (which is an image either taken with a camera on a tripod, or of a static scene).

Ezra and Nayar [3] proposed a hybrid imaging system consisting of a primary (high spatial resolution) detector and a secondary (high temporal resolution) detector. The secondary detector provides more accurate motion information to estimate the PSF, thus making deblurring possible even under long exposure. However, the method needs additional hardware support, and the deblurred image can still be distinguishable from the ground truth image.

Because of the weakness of the deblurring methods, we do not directly perform deblurring on I_H . Instead, an image color correction approach is adopted. By incorporating the color statistics and the spatial structures of I_H and I_L , we

propose a Bayesian framework, and maximize the *a posterior* (MAP) of the color mapping function $f(\cdot)$ from I_L to I_H in the color space so that the under-exposed I_L is enhanced to a normally exposed image I_C .

Our method can deal with camera shake and object movement at the same time, and in an unified framework. Moreover, change of object topology or object deformation can also be naturally handled, which is difficult for most deblurring methods, since different parts of the object have different PSFs. Besides, by slightly modifying one constraint, our method can be extended to deal with high contrast scenes, and automatically produce images which capture fine details in highlight or saturated area.

The rest of this paper is organized as follows: we describe our image acquisition system in Section 2. Section 3 defines the relationship between I_L and I_H . In Section 4, we state and define our problem, propose our probabilistic model, and infer the color mapping function in the Bayesian framework. Section 5 presents our results. Finally, we conclude our paper in Section 6.

2 Image Acquisition

To correctly relate two images, we require that I_L be taken almost immediately after I_H is taken. This is to minimize the difference between the two images and to maximize the regional match of the positions of each pixel if the time lapse is kept as short as possible, as illustrated in Fig. 2(a). In other words, the under-exposed image I_L can be regarded as a sensing component in the normally exposed image I_H in the temporal coordinates. This requirement makes it possible to reasonably model the camera movement during the exposure time, and constrain the mapping process.

Our image acquisition system and its configuration is in Fig. 2(b). The digital camera is connected to the computer. The two successive exposures with different shutter speeds are controlled by the corresponding camera software. This setup

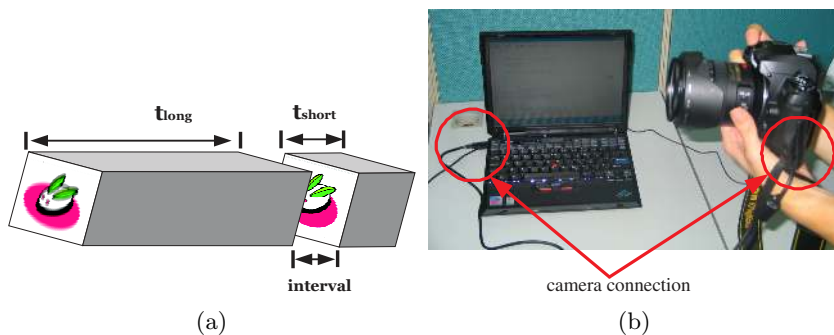


Fig. 2. (a) Two successive exposures guarantee that the center of the images do not vary by too much. (b) The configuration of our camera system.

frees the photographer from manually changing the camera parameters between shots. So that s/he can focus on shooting the best pictures.

A similar functionality, called *Exposure Bracketing*, has already been built in many digital cameras, e.g., Canon G-model and some Nikon Coolpix model digital cameras. With one shutter pressing, two or three successive images are taken with different shutter speeds under the same configurations. However, using the built-in camera functionality has some limitations: it does not operate in manual mode, and the difference of shutter speeds is limited.

In the next section, we analyze the relationship between I_L and I_H , and propose the constraints that relate these two images.

3 Relationship between I_L and I_H

I_L and I_H are two images of the same scene taken successively with different exposures. Therefore, they are related not only by the color statistics, but also by the corresponding spatial coherence. In this section, we describe their relationship, which are translated into constraints for inferring a color mapping function in our Bayesian framework, which will be described in the next section.

3.1 Color Statistics

In RGB color space, important color statistics can often be revealed through the shape of a color histogram. Thus, the histogram can be used to establish explicate connection between I_H and I_L . Moreover, since high irradiance always generates brighter pixels [7], the color statistics in I_L and I_H can be matched in order from lower to higher in pixel intensity values. Accordingly, we want to reshape the histogram of I_L , say, h_{I_L} , such that:

$$g(h_{I_L}) \doteq h_{I_H} \quad (1)$$

where $g(\cdot)$ is the transformation function performed on each color value in histogram, and h_{I_H} is the histogram of I_H . A common method to estimate $g(\cdot)$ is *adaptive histogram equalization*, which normally modifies the dynamic range and contrasts of a image according to a destination curve.

Unfortunately, this histogram equalization does not produce satisfactory results. The quantized 256 (single byte accuracy) colors in each channel are not sufficient to accurately model the variety of histogram shapes. Hence, we adopt the following method to optimally estimate the transformation function:

1. Convert the image from RGB space to a perception-based color space $l\alpha\beta$ [4], where the l is the achromatic channel and α and β contain the chromaticity value. In this way, the image is transformed to a more discrete space with known phosphor chromaticity.
2. Accordingly, we cluster the color distributions in the new color space into 65536 (double byte precision) bins, and perform histogram equalization.
3. Finally, we transform the result back to the RGB space.

By performing this transformed *histogram equalization*, we relate the two images entirely in their color space.

3.2 Color Statistics in High Contrast Scene

In situations that the images are taken in a high contrast scene, bright regions will become saturated in I_H . Histogram equalization can not faithfully transfer colors from I_L to I_H , especially in the saturated area, which not only degrades the structured detail in the highlight region, but also generates abrupt changes in the image color space. To solve this problem, the color mapping function $g(\cdot)$ described in section 3.1 needs to be modified to cover a larger range. In our experiments, we adopt the color transfer technique in [13] in this situation. It also operates on image histogram, which transfers the color from the source image to the target by matching the mean and standard deviation for each channel. It has no limit on the maximum value of the transferred color since the process is actually a Gaussian matching. In our method, all non-saturated pixels in I_H are used for color transfer to I_L . After applying [13], the mapping result of I_L exceeds the color depth (that is, above 255), and extends the saturated pixels to larger color values. Hence, we construct a higher intensity range² image to reveal details in both bright and dark regions.

3.3 Spatial Constraint

The statistics depicted above does not consider any temporal coherence between I_H and I_L . However, since the two images are taken successively, there is a strong spatial constraint between I_H and I_L .

Let us consider the situation that a region contains similar color pixels, Fig. 3(a) shows a region from the original image, while Fig. 3(b) shows the same region taken with motion blur. The yellow dots mark the region centers. The lower curves show pixel colors along one direction. From this figure, we can observe that the color toward the center of the region is less affected by blurring, given that the region area is sufficient large and homogeneous. Additionally, the consistency of colors in the region also guarantees that the color of central pixels can be matched. Therefore, we adopt the following region matching method to robustly select matching seeds in I_H and I_L :

1. Over-segment I_H such that each region $R_m(I_H)$ contains similar colors (Fig. 4(a)).
2. To sort all regions according to the homogeneity and size, we perform the same morphological eroding operation for each region $R_m(I_H)$, and record the number of iterations to completely erode it and the region centers which are the last few pixels in the eroding process for each region. Fig. 4(b) shows an intermediate image in the eroding process.
3. We sort all iteration numbers in descending order, and select the first M regions as the most possible candidates. As a result, the positions of these region centers are selected as matching positions. Finally, we pick out pixel pairs $\{c_L^m, c_H^m\}$ in I_H and I_L in the matching position and calculate the value for each c^m as a Gaussian average of the colors of neighboring pixels,

² We do not construct HDR since we do not perform radiometric calibration

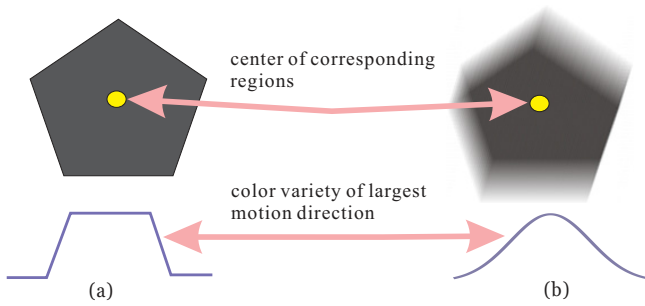


Fig. 3. Matching homogeneous region in blurred situation. (a) original homogeneous region. (b) blurred region. Color towards the center is less influenced by blurring.

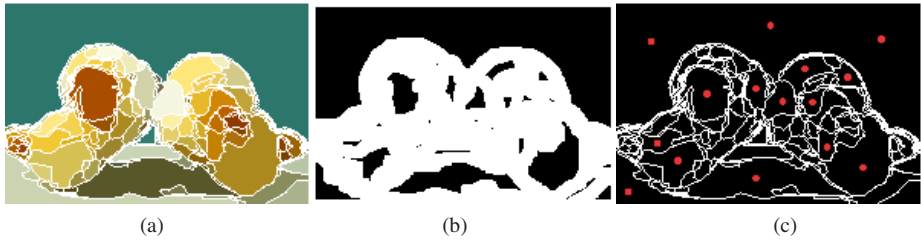


Fig. 4. Region matching process. (a) Initial segmentation. (b) In the eroding process, small regions are filled quickly. (c) The final selected regions, in which the red dots represent the selected region centers after eroding.

where the variance is proportional to the iteration numbers. We illustrate the selected region centers as red dots in Fig. 4(c), which are in the largest and most homogeneous M regions.

The matching process implies that an ideal color mapping function should robustly transform some matching seeds colors in I_L to those in I_H . In the next section, we propose our Bayesian framework which incorporates the two constraints, color and spatial, into consideration, so as to infer a constrained mapping function.

4 Constrained Mapping Function

We define the color mapping function $f(\ell_i) = \ell'_i$, where ℓ_i and ℓ'_i are color values in the two sets respectively. Accordingly, the resulting image I_C is built by applying $f(\cdot)$ to the under-exposed image I_L : $I_C(x, y) = f(I_L(x, y))$, where $I_k(x, y)$ is pixel values in image I_k . Note that the form of $f(\cdot)$ is constrained by both I_L and I_H .

In Bayesian framework, we maximize the *a posterior* probability (MAP) to infer f^* given the observations from I_L and I_H :

$$f^* = \arg \max_f p(f | I_L, I_H) \quad (2)$$

In section 3, we observe two kinds of connections between I_L and I_H . One is color statistics which can be described by two histograms h_{I_L} and h_{I_H} of I_L and I_H respectively. The other is region matching constraint which can be represented by a number of M corresponding matching color seeds $\{c_L^m, c_H^m\}_{m=1}^M$ between I_L and I_H . In our formulation, we regard them as our constraints and rewrite (2) as:

$$\begin{aligned} f^* &= \arg \max_f p(f|h_{I_L}, h_{I_H}, \{c_L^m, c_H^m\}_{m=1}^M) \\ &= \arg \max_f p(h_{I_L}, h_{I_H}, \{c_L^m, c_H^m\}_{m=1}^M | f) p(f) \end{aligned} \quad (3)$$

Next, we define the likelihood $p(h_{I_L}, h_{I_H}, \{c_L^m, c_H^m\}_{m=1}^M | f)$, and the prior $p(f)$.

4.1 Likelihood

Since we perform global matching in discrete color space, f is approximated by a set of discrete values $f = \{f_1, f_2, \dots, f_i, \dots, f_N\}$, where N is the total number of bins in color space. Hence, the likelihood in Eqn. (3) can be factorized under the i.i.d. assumption:

$$p(h_{I_L}, h_{I_H}, \{c_L^m, c_H^m\}_{m=1}^M | f) = \prod_{i=1}^N p(g(\ell_i), \{\bar{c}_L^i, \bar{c}_H^i\} | f_i) \quad (4)$$

where $g(\ell_i)$ is a function to transform h_{I_L} to h_{I_H} at color value ℓ_i . The \bar{c}_L^i is the most similar color to ℓ_i in color seeds set $\{c_L^m\}_{m=1}^M$, and \bar{c}_H^i is the corresponding color of \bar{c}_L^i in color seed pairs.

According to the analysis in section 3, $g(\ell_i)$ and $\{\bar{c}_L^i, \bar{c}_H^i\}$ are two constraint factors for each f_i . Both of their properties should be maintained on the mapping function. As a consequence, we balance the two constraints and model the likelihood as follows:

$$p(g(\ell_i), \{\bar{c}_L^i, \bar{c}_H^i\} | f_i) \propto \exp\left(-\frac{\|f_i - (\alpha g(\ell_i) + (1 - \alpha)\bar{c}_L^i)\|^2}{2\sigma_f^2}\right) \quad (5)$$

where the scale α weights these two constraints, and σ_f^2 is a variance to model the uncertainty of two kinds of constraints. The larger the value of α is, the smaller the confidence of the matching seed pairs. We relate α to the following factors:

- The distance $\|\ell_i - \bar{c}_L^i\|$. Large distance indicates weak region matching constraint, which makes α approach to 1. Hence, the α is inversely proportional to this distance.
- The uncertainty of correspondence in matching color pair $\{\bar{c}_L^i, \bar{c}_H^i\}$. As depicted in section 3.3, the larger the matching region size is, the larger confidence we can get from the region center for the matching colors. Hence, we define uncertainty σ_c to be proportional to the region size for each matching color.

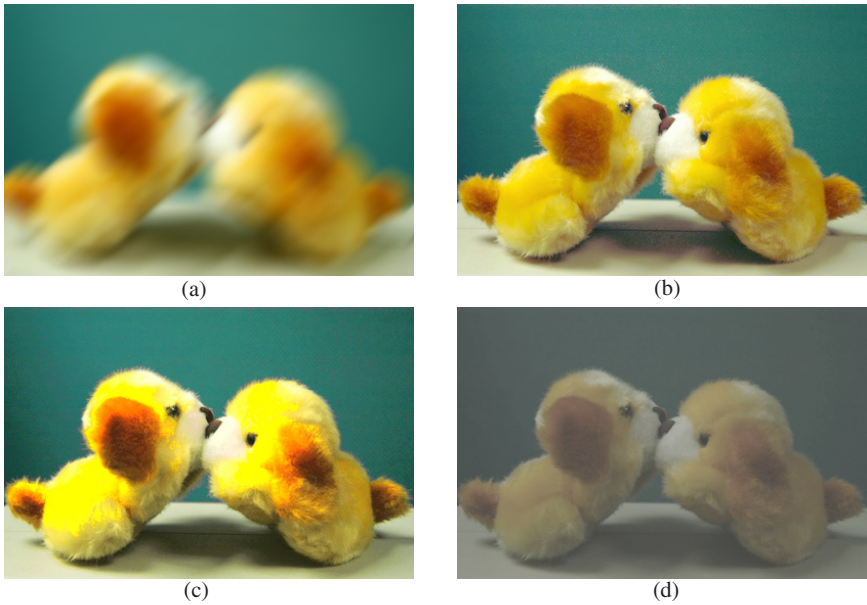


Fig. 5. Puppies. (a) Input blurred image. (b) Our result. (c) Color transfer result [13]. (d) Result of Gamma correction by 2.5. Better visual quality and more details are achieved by using spatial constraint in our framework.

Combining these two factors, we define α as:

$$\alpha = \exp\left(-\frac{\sigma_c^2 \|\ell_i - \vec{c}_L^i\|^2}{2\beta^2}\right) \quad (6)$$

where β is the scale parameter to control the influence of α .

4.2 Prior

As a prior, we enforce the monotonic constraint on $f(\cdot)$, which maintains the structural details in I_L . In addition, to avoid abrupt change of the color mapping for neighboring colors, we require that $f(\cdot)$ be smooth in its shape. In this paper, we minimize the second derivative of f :

$$\begin{aligned} p(f) &\propto \exp\left(-\frac{1}{2\sigma_f^2} \int (f'')^2\right) \\ &\propto \exp\left(-\frac{1}{2\sigma_f^2} \sum_i (f_{i-1} - 2f_i + f_{i+1})^2\right) \end{aligned} \quad (7)$$

where σ_f^2 is the variance to control the smoothness of f .

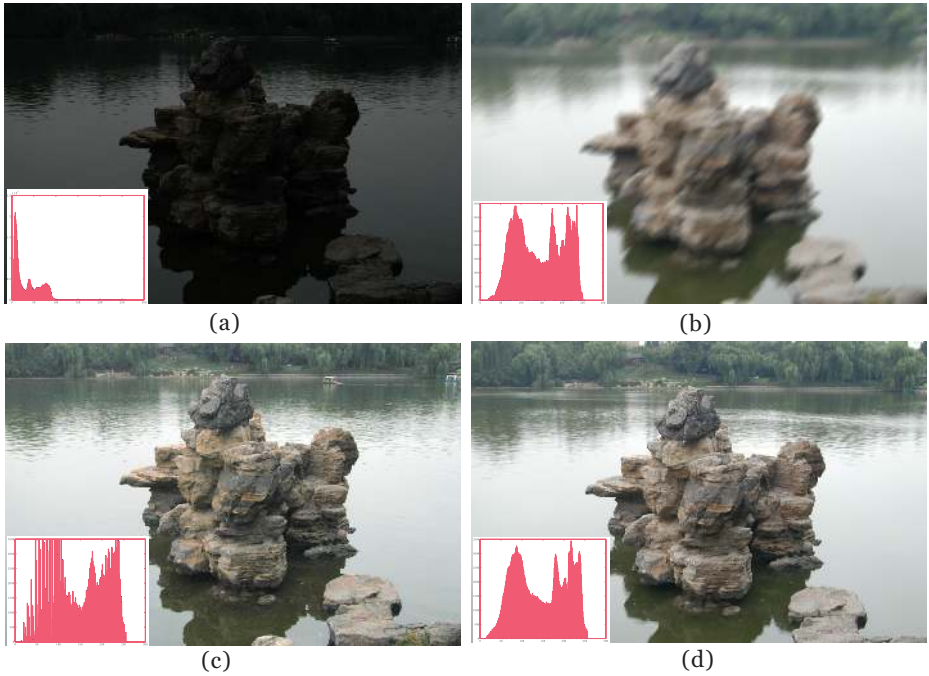


Fig. 6. Rock example of image correction. The upper two images are input defective images. (c) is our result. (d) is the ground truth. Note the histograms in (c) and (d) are much closer than those in (a) and (b). However, because of the quantization error and large exposure difference between I_L and I_H , they can not be identical in shapes.

4.3 MAP Solution

Combining the log likelihood of Eqn. (4) and the log prior in Eqn. (7), we solve the optimization problem by minimizing the following log posterior function:

$$E(f) = - \sum_i \log p(g(\ell_i), \{\bar{c}_L^i, \bar{c}_H^i\} | f_i) - \log p(f) \quad (8)$$

where $E(f)$ is a quadratic objective function. Therefore, the global optimal mapping function $f(\cdot)$ can be obtained by the singular value decomposition (SVD).

Although the monotonic constraint is not enforced explicitly in Eqn. (7), we find the smoothness constraint is sufficient to construct the final monotonic f in our experiments.

5 Results

We evaluate our method in difficult scenarios to show the efficacy of our approach. The results are classified into 4 different groups as follows, all of them are illustrated in color:

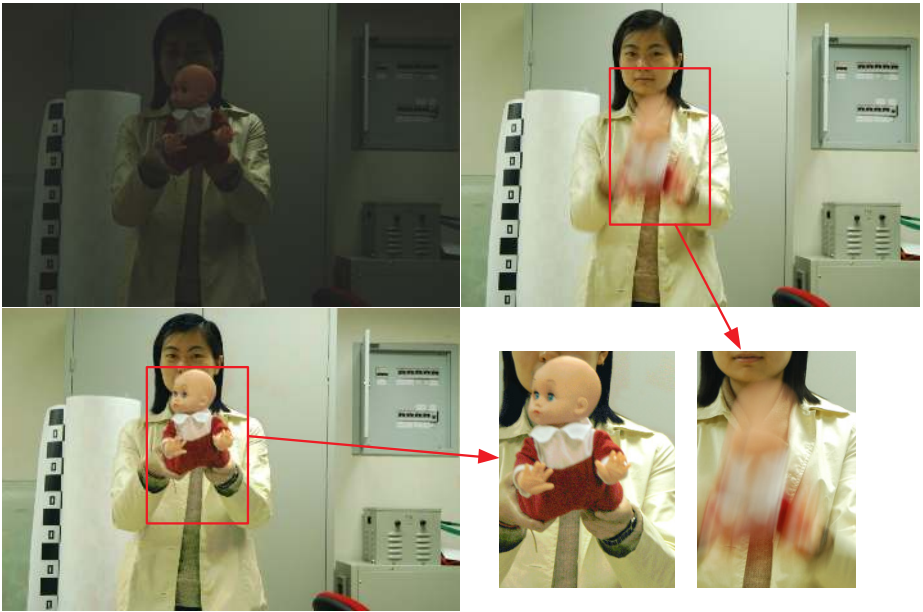


Fig. 7. Doll example. The upper two images are our input. Our result is the left bottom image, which indicates that local blurring in images can be naturally handled.

5.1 Bayesian Color Mapping versus Other Adjustment Techniques

The two constraints described in section 3 are both essential in our method. They optimize the solution in two different aspects cooperatively. Therefore, the combination and balance of these constraints guarantee the visual correctness of our method. Fig. 5 compare our result with that from pure color transfer method [13] and adaptive histogram equalization. We take the first two images in Fig. 1 as input. They are taken with shutter speed $\frac{1}{30}$ s and $\frac{1}{1.6}$ s respectively. Fig. 5(b) is generated with our method. Fig. 5(c) and (d) are the results of pure color transfer and gamma correction. Clearly, Fig. 5(b) has higher visual quality, and the colors are closest to the input image in Fig. 5(a).

5.2 Motion Blur Caused by Hand-Held Camera

The rock example in Fig. 6 shows the ability of our method to optimally combine the color information of the two input images. Unlike other deblurring methods, the resulting edges are very crisp and clear. The two input images (a) and (b) are taken with shutter speeds $\frac{1}{40}$ s and $\frac{1}{3}$ s respectively. (c) and (d) are our color mapped image I_C and ground truth with their corresponding histograms. The ground truth is taken by using a tripod. Note that colors are visually and statistically close.

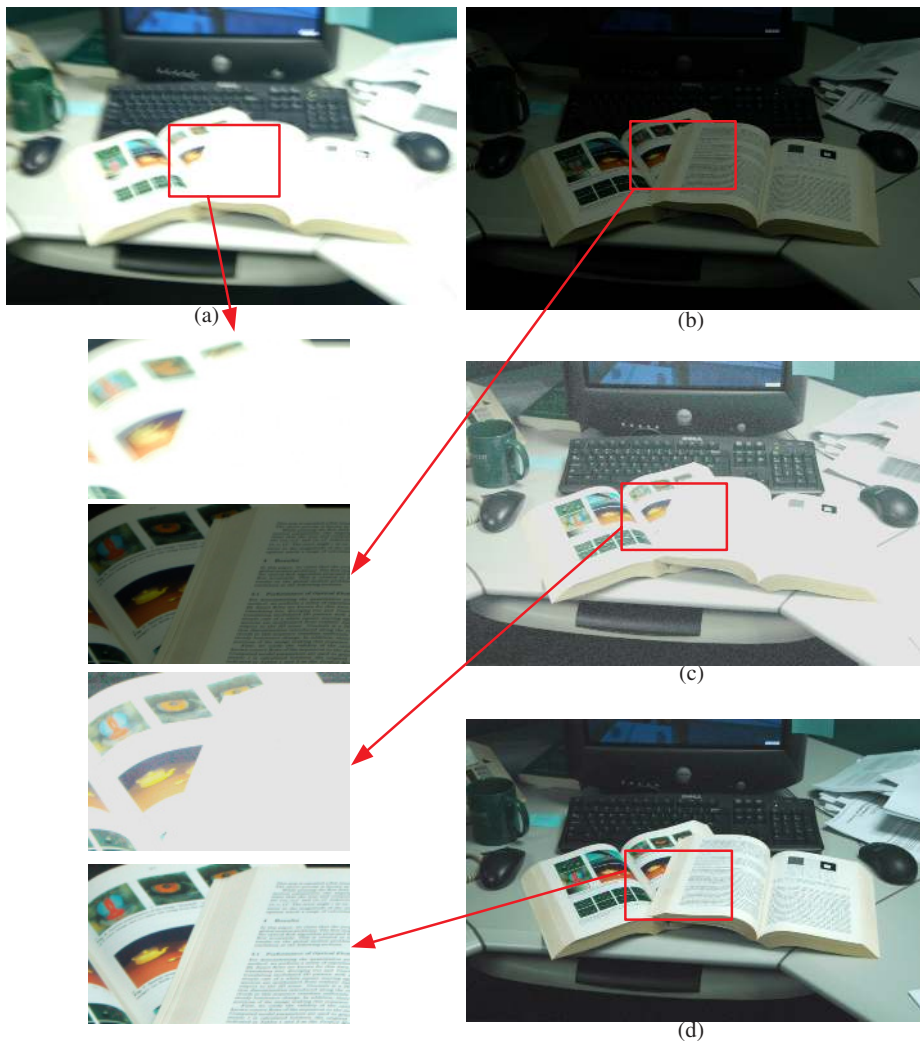


Fig. 8. Image correction for high contrast scene. (a) I_H , which has a large saturated area. (b) I_L has clear structure information. (c) The result produced by applying original histogram equalization. (d) Our final result I_C where pixel intensity values are enhanced and fine details are maintained. The bottom left images are selected enlarged portions of I_C .

5.3 Motion Blur Caused by Objects Movement

Another strength of our method is that we can easily solve the object movement or deformation problem if the object movement is too fast in normal exposure interval. Fig. 7 illustrates one experiment. The input normal exposure image is locally blurred, i.e., PSF has no uniform representation in the whole image,

which easily makes deconvolving methods fail. In our method, by reducing the camera shutter speed by 4 stops, we produce I_C with largely reduced blurring effect.

5.4 High Contrast Scene

As described in section 3.2, for high contrast scene, we modify the statistical color mapping function from adaptive histogram equalization to the color transfer function [13] in the framework. We present our results in Fig. 8. (a) and (b) are input I_H and I_L , respectively, and (c) is reconstructed by setting $g(\cdot)$ as the original histogram equalization function. (d) is our final result with enhanced colors and details by modifying $g(\cdot)$ to use the color transfer method in [13]. Tone mapping [5] is performed to display the image we constructed in (d).

6 Conclusion

In this paper, we propose a Bayesian approach to combine two defective images to construct a high quality image of the scene, which may contain moving objects. No special hardware is built to compensate camera shake. Instead, a color mapping approach is adopted. Yet our color mapping is constrained by spatial details given by the under-exposed image, and thus differs from and improves on previous pure color transfer techniques. By properly formulating color statistics and spatial constraints, and incorporating them into our Bayesian framework, the MAP solution produces an optimal color mapping function that preserves structural details while enhancing pixel colors simultaneously. Using only two images in all our experiments, we produce a high quality image, and largely reduce the shutter speed by 3 to 4 stops to enhance the image quality in dim light.

However, the color statistics is largely dependent the image quality of the camera. If the dark image contains a large amount of noise, the contaminated information needs to be treated first. One solution is taking more under-exposed images to reduce noise level. Another issue is the search for spatial correspondence in the presence of fast movement of camera or objects. These issues will be investigated in future work.

References

1. *Complete Digital Photography (2nd Edition)*. Charles River Media Press, 2002.
2. B. Bascle, Andrew Blake, and Andrew Zisserman. Motion deblurring and super-resolution from an image sequence. In *ECCV*, pages 573–582, 1996.
3. Moshe Ben-Ezra and Shree K. Nayar. Motion deblurring using hybrid imaging. *Processings of CVPR*, 2003.
4. T.W.Cornin D.L.Rudeman and C.C.Chiao. Statistics of cone responses to natural images: Implications for visual coding. In *J. Optical Soc. of America*, number 8, pages 2036–2045, 1998.

5. Peter Shirley Erik Reinhard, Mike Stark and Jim Ferwerda. Photographic tone reproduction for digital images. In *Siggraph 2002*, pages 267–276, 2002.
6. R. Fabian and D. Malah. Robust identification of motion and out-of-focus blur parameters from blurred and noisy images. *CVGIP: Graphical Models and Image Processing.*, 1991.
7. M. D. Grossberg and S. K. Nayar. What can be known about the radiometric response function from images? In *ECCV*, May 2002.
8. Canon Inc. http://www.canon.com.my/techno/optical/optical_b.htm.
9. Nikon Inc. http://www.nikon.co.jp/main/eng/society/tec-rep/tr8-vr_e.htm.
10. D. Kundur and D. Hatzinakos. A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing.*, pages 46(2):375–390, February 1998.
11. X. Liu and A. Gamal. Simultaneous image formation and motion blur restoration via multiple capture. *Proc. Int. Conf. Acoustics, Speech, Signal Processing.*, 2001.
12. A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in different directions. *IEEE Workshop on Applications of Computer Vision (WACV)*, 2000.
13. M. Gooch B. Reinhard, E. Ashikhmin and P. Shirley. Color transfer between images. In *IEEE Computer Graphics and Applications*, pages 34–40, 2001.
14. A. Lantzman Y. Yitzhaky, I. Mor and N. S. Kopeika. Direct method for restoration of motion-blurred images. *J. Opt. Soc. Am. A.*, pages 15(6):1512–1519, June 1998.
15. Y. Levy Y. Yitzhaky, G. Boshusha and N.S. Kopeika. Restoration of an image degraded by vibrations using only a single frame. *Optical Engineering*, 2002.