DEBLUR A BLURRED RGB IMAGE WITH A SHARP NIR IMAGE THROUGH LOCAL LINEAR MAPPING

Tao Yue[1], Ming-Ting Sun[2], Zhengyou Zhang[3], Jinli Suo[1] and Qionghai Dai[1]
[1] Department of Automation, Tsinghua University, Beijing 100084, China
[2] Department of Electrical Engineering, Seattle, 98105, U.S.A
[3] Microsoft Corp., Redmond, 98052, U.S.A
yuetao.thu@gmail.com, sun@ee.washington.edu, zhang@microsoft.com, {jlsuo, qionghaidai@tsinghua.edu.cn}

ABSTRACT

Image acquisition in a low light environment requires long exposure to achieve acceptable signal-to-noise ratio, which however causes blurry effect. This paper addresses this problem by using a sharp near-infrared (NIR) image when the environment has sufficient NIR light. We assume that an RGB and NIR image pair has a linear mapping in a local area and that the mapping function is valid for both the blur and sharp image pairs. Using this property, we solve the sharp RGB images from a blurred RGB image and the corresponding sharp NIR image. The effectiveness of the proposed algorithm is verified with both synthetic and real captured datasets.

Index Terms— Image acquisition, image enhancement, deblurring, RGB/NIR images, Local Linear Transform

1. INTRODUCTION

Image deblurring is a challenging problem for its ill-posed nature. Although many algorithms [1–12] have been proposed and significant progresses have been made, many aspects of the problem remain unsolved and still require further study. Existing algorithms are not sufficient for practical applications due to lack of robustness in kernel estimation and the ringing effects caused by deconvolution.

Some researches attempt to extract more powerful priors and regularizers from nature images to address these problems. For example, step-like edges are abundant in nature images, and many literatures [1-6] are based on this prior. Normalized sparsity prior, e.g. L_1 / L_2, proposed by Krishnan *et al.* [7], the L_0 sparse representation proposed by Xu *et al.* [11], and patch priors learned from nature image databases by Sun *et al.* [10] try to give more appropriate regularizers for the sharp edges. Besides, Joshi *et al.* [3] propose to solve the ill-posed problem by the aid of a color prior. Hu and Xue [8] as well as Goldstein and Fattal [9] estimate the blur kernel via the gradient-domain correlation.

Recently, an increasing number of researches focus on computational photography and try to solve the problem with additional information captured by carefully designed imaging systems. Joshi *et al.* [14] use the inertial measurements to record the camera motion and then derive the blur kernel from the motion information. Agrawal and Raskar [15] try to capture the image optimal for deblurring by using coded exposure techniques. Hybrid cameras combining traditional and other imaging systems are also applied to image deblurring. Yuan *et al.* [16] use blurred/noisy image pairs, Zhuo *et al.* [17] use blurred/flash image pairs, and Tai *et al.* [18] as well as Ben-Ezra and Nayer [19] use hybrid videos captured by high-resolution low-rate/low-resolution high-rate cameras to recover high quality sharp images or videos.

Because of the specific character of NIR and other invisible spectra, the visible/non-visible hybrid cameras are widely used in image processing applications. Fredembach and Süsstrunk [20] use NIR information to enhance the visible images. Schaul *et al.* [21] apply the NIR information on image dehazing. Krishnan and Fergus [22] propose to use dark flash, i.e., both NIR and ultra-violet (UV) light, to remove noise of pictures in captured low light environment. Yan *et al.* [27] introduce the scale map to improve the performance of the cross-channel (RGB/NIR) denoising. Similarly, Bennett *et al.* [23] use a fusion framework to perform video denoising with an RGB/NIR hybrid camera.

However, most of the existing deblurring methods recover the latent image by using deconvolution and the other modality (flashed, noisy, low frequency or NIR&UV) of the scene as gradient constraints from the edges. Although the performance is significantly improved than the deblurring algorithms for single images, they still suffer from the robustness and ringing artifact problems. Besides, most of these methods focus on the camera-shake caused blur, which can be formulated by a 2-D convolution model or a high dimensional unified blurring model. However, in practice, many pictures are deteriorated by moving objects in the scene, which cannot be described by a unified model. Most existing deblurring methods are unable to solve this kind of problems.

This paper proposes a novel scheme based on the property that there exists a *reasonable* linear mapping in a local area from the blurred NIR image to the blurred RGB image, and the mapping function is also valid for the sharp version of the NIR/RGB image set. This is a reasonable assumption because only a small local area is considered. By using this property, we compute the sharp RGB images directly from the sharp NIR image sequence. A gradient descent based algorithm is proposed to solve the local linear

This work is supported Beijing Key Laboratory of Multi-dimension & Multi-scale Computational Photography (MMCP) Tsinghua University and the Project of NSFC (No. 61035002, 61120106003 & 61171119).

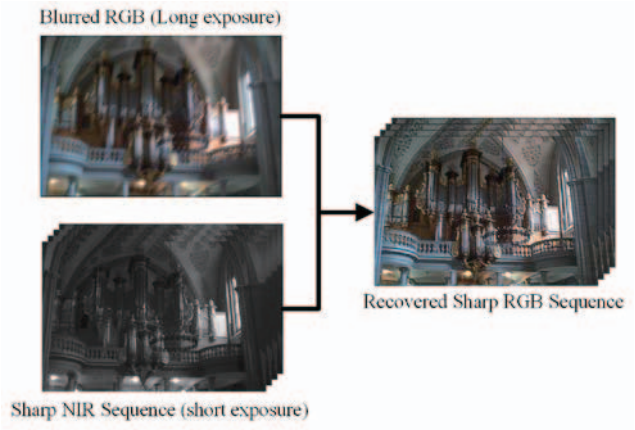

Fig. 1. Diagram of our system.

transformation mappings efficiently. After that, the sharp RGB image sequence can be computed by transforming the NIR images with these mappings. Fig. 1 illustrates our proposed system, which takes a blurred RGB image and the corresponding sharp NIR image sequence captured by a multispectrum camera with different exposure settings as the input, and recovers the sharp RGB image sequence as the output by using our proposed algorithm.

This paper contains three main contributions: (1) We propose the local linear transformation (LLT) and analyze and discuss its characteristics; (2) we propose a framework to recover the sharp RGB image from a blurred RGB image and the corresponding sharp NIR image by using local linear transformations; (3) we propose a gradient based algorithm to recover the local linear transformations efficiently.

The rest of the paper is organized as follows. Section 2 states and discusses the local linear transformation property. Section 3 describes the algorithm for recovering sharp RGB images from the RGB/NIR image pair by using the LLT in detail. The experiment results are demonstrated in Section 4. The conclusions and discussions are presented in Section 5.

2. LOCAL LINEAR TRANSFORMATION PROPERTY

In this part, we demonstrate the Local Linear Transformation (LLT) and analyze its characters.

LLT states that *in a local area with the same blur effect, the pixel values of different channels (visible and near infrared in this paper) follow a certain linear transformation, and this linear transformation is valid in the same area for pixels of the sharp version of those channels too*. Fig. 2 shows an example of the LLT. The blurred and sharp visible/near infrared channel pairs are shown in the top row. Specifically, (a) and (b) are the blurred green channel and blurred NIR channel respectively. The red and blue rectangles in (a) and (b) mark the local areas, and the pixel values between green and NIR channel in these areas are shown in (c) and (d). Similarly, the sharp green channel and sharp NIR channel are shown in (e) and (f) respectively, and the mapping transformations of the

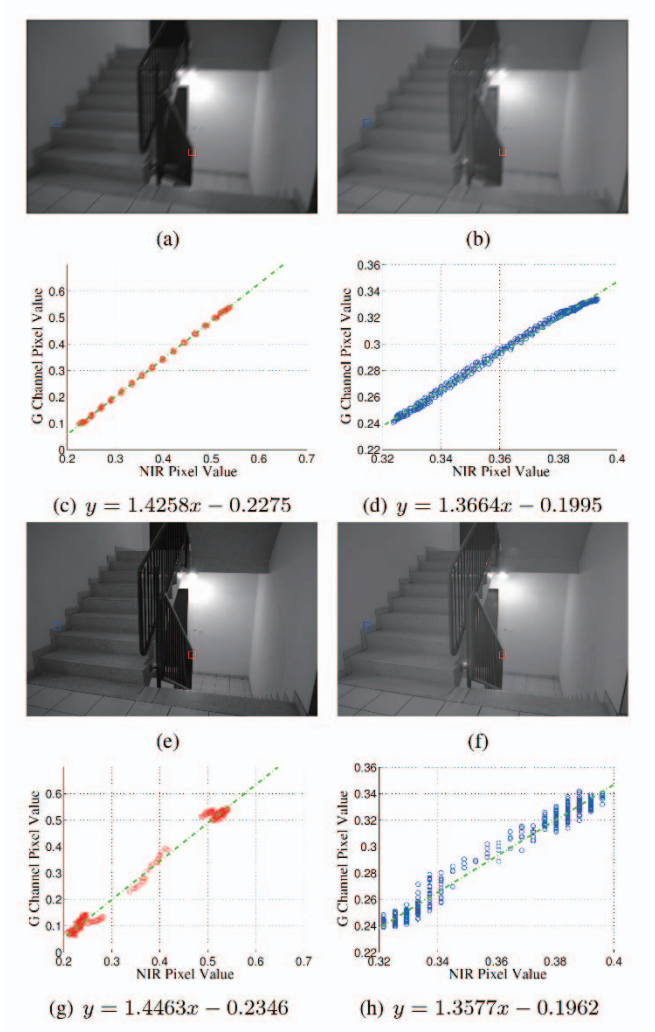

Fig. 2. Example of Local Linear Transformation (LLT). (a) Blurred Green channel. (b) Blurred NIR image. (c) Scattering map for blurred Green vs. NIR pixel values in red rectangle. (d) Scattering map for blurred Green vs. NIR pixel values in blue rectangle. (e) Sharp Green channel. (f) Sharp NIR image. (g) Scattering map for sharp Green vs. NIR pixel values in red rectangle. (h) Scattering map for sharp Green vs. NIR pixel values in blue rectangle.

pixels in the same areas are shown in (g) and (h) of Fig. 2. We fit the linear coefficients of (c)(d)(g)(h), and give the fitted functions in their subtitles. It can be seen that the pixel values in these local areas follow linear transformations reasonably well. In addition, the linear transformation may be totally different in different areas, but they are the same for both blurred and sharp image pairs in the same area.

We have shown an example of LLT between different channels (green and NIR) for both blurred and sharp images, but is it always valid for every area of the image, or is it only valid in some regions? In the following, we show that the LLT are not valid for all the areas, but almost always valid for the area with intensity variation, i.e., edge or texture

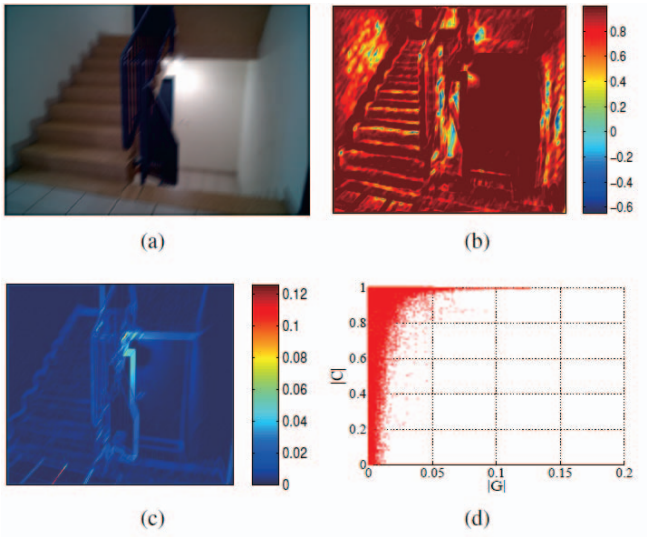

Fig. 3. Distribution of large correlation coefficients between the blurred NIR image and the Green channel of the visible image. (a) The original blurred RGB image. (b) The correlation coefficient map between the blurred NIR and the Green channel. Each coefficient of a certain pixel is computed from a local (21×21) patch. (c) The magnitude of gradients of the NIR image. (d) The scatter points of magnitude of gradient |G| and the corresponding absolute correlation coefficient |C|.

regions. As shown in Fig. 3, (a) is the blurred RGB image, (b) shows the correlation coefficient maps between the pixel values in the local area in the blurred green and NIR channels. Each coefficient of a pixel is computed from its neighborhood patch with the 21×21 size. (c) is the magnitude of gradients of the NIR image. It is evident that a region with large gradients also has a large correlation coefficient, i.e., the pixels in this region are highly linearly correlated. Fig. 3(d) shows the scatter points of magnitude of gradients and its corresponding absolute correlation coefficients. We can see that almost all the pixels with gradients larger than 0.02 are highly linearly correlated, with a correlation coefficient larger than 0.8. Considering only the regions with large gradients are salient and significant for the restoration task, the LLT can be used for our sharp image recovering scenario.

In fact, the LLT can be inferred from the two-color model [3, 24] which is based on the piecewise smoothness of images and states that any pixel color can be represented as a linear combination of two basic colors within the local neighborhood. Considering the piecewise smoothness is valid for both 3-channel (RGB) and 4-channel (RGBN) images, the two-color model is also valid for RGBN images, thus for a local area, we have

$$I^i = p^i I^1 + (1-p^i) I^2, \tag{1}$$

where $I = [I_r\ I_g\ I_b\ I_n]^T$ is the color vector, I^i is the color of pixel i in the local area, I^1 and I^2 are the two basic colors, and p^i is the combination weight. For a visible channel like I_r, we can derive a linear mapping (A, B) to map the NIR channel to the red channel of basic colors,

$$\begin{aligned} I_r^1 &= AI_n^1 + B \\ I_r^2 &= AI_n^2 + B \end{aligned}. \tag{2}$$

Then, the red channel intensity of pixel i becomes,

$$\begin{aligned} I_r^i &= p^i I_r^1 + (1-p^i) I_r^2 \\ &= p^i (AI_n^1 + B) + (1-p^i)(AI_n^2 + B) \\ &= A(p^i I_n^1 + (1-p^i) I_n^2) + B(p^i + 1 - p^i) \\ &= AI_n^i + B \end{aligned}. \tag{3}$$

Since the blurring process is a linear combination of local pixels, the two basic colors do not change. The LLT parameters A and B are constants for both blur and sharp images. Thus, we can infer A and B from the blurred images, and compute the sharp RGB image from the sharp NIR image. However, the LLT is different from that color model in three aspects: (1) The local linearity in the LLT is between different color channels; (2) the LLT is valid for arbitrary blurred channel pairs; and (3) the linearity is valid and approximately the same for sharp and blurred channels.

Based on the above characteristics of the LLT, we propose a novel framework to remove the blurry effect by using an RGB/NIR hybrid camera.

3. PROPOSED FRAMEWORK

3.1. Proposed Approach

The proposed method has no assumption or constraint on the blur types of the image. That means either uniform or non-uniform, camera-shake caused or moving-object caused, depth dependent or independent, can all be solved by our proposed method. Instead of convolution, a general blur model is applied,

$$I_{\{v,n\}} = HL_{\{v,n\}} + N_{\{v,n\}} \tag{4}$$

where I is the blurred image, L is the latent sharp image, N is additive noise and H is the blur matrix, subscript v denotes visible channel and $v \in \{r, g, b\}$, n indicates the near infra-red channel.

According to the LLT, we know that there exist local linear transforms between visible channels and the near infra-red channel. That means for a local patch centered at pixel i in the visible channels and the n channel, the pixel values inside the patches follow,

$$I_v(j) = A_v(i) I_n(j) + B_v(i), \text{ for } j \in G(i) \tag{5a}$$

$$L_v(j) = A_v(i) L_n(j) + B_v(i), \text{ for } j \in G(i) \tag{5b}$$

where A and B are the local linear transformation maps, $G(.)$ denotes the neighborhood set operation and j is the pixel index for the neighborhood set of i.

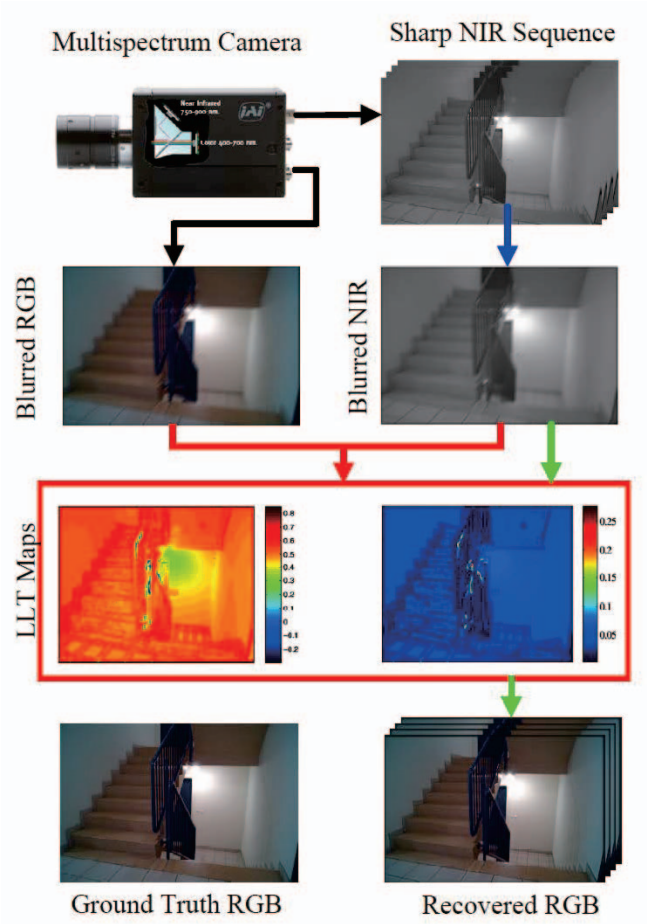

Fig. 4. Flowchart of our hybrid camera based sharp image recovering.

Fig. 4 shows the flowchart of the proposed algorithm. The camera shown in left top of Fig. 4 is a JAI AD-080CL camera [26] which can capture both the visible and non-visible spectrum of the same scene and output the signals by two connectors separately. As shown in Fig. 4, our system capture a single blurred RGB image and a series of corresponding sharp NIR image sequence simultaneously. The black line in Fig. 4 shows the capturing procedure of our system. Given the sharp NIR image sequence during exposure, we can derive the blurred NIR image by averaging the images, as shown with the blue lines.

The LLT maps in Eqs. (5a), (5b) will be computed from that estimated blurry NIR image and the captured blurry RGB image according to Eq. (5a), as shown with red lines in Fig. 4. Finally, the green lines in Fig. 4 show the process of recovering the sharp RGB image from the sharp NIR image using the computed LLT maps by Eq. (5b).

3.2. Objective Function

Based on Eq.(5a), we can define the data term of our objective function,

$$E_d = \sum_i (A_v(i) I_n(i) + B_v(i) - I_v(i))^2 \quad (6)$$

However, the LLT maps cannot be determined by Eq. (6) because many possible combinations of A_v and B_v can meet the constraints, but not all of them meet Eq. (5b). To make the problem tractable, additional constraints have to be introduced. In fact, we can derive additional constraints of the LLT map A_v by applying the gradient operation on Eq. (5a),

$$\nabla I_n(i) = A_v(i) \nabla I_n(j) \quad (7)$$

According to Eq. 7, the constraints of the gradient terms can be defined by,

$$E_g = \sum_i (A_v(i) \nabla I_n(i) - \nabla I_v(i))^2 \quad (8)$$

Besides, as discussed above, the linear transformation is valid in local areas. In other words, the LLT maps are constant in the corresponding area or the LLT maps should be smooth. Therefore, the smoothness constraints on the LLT maps A_v and B_v are introduced as,

$$E_s = \sum_i (\nabla A_v(i))^2 + \sum_i (\nabla B_v(i))^2 \quad (9)$$

Finally, we have the following objective function

$$E = E_d + \lambda E_g + \beta E_s \quad (10)$$

where λ and β are the weight coefficients and are set to 1and 0.1 empirically in this paper.

3.3. Gradient Based LLT Maps Estimation

In this paper we use a gradient descent based method to optimized cost function in Eq. (10). Considering there are two parameter sets A_v and B_v, we compute the derivatives of energy in Eq. (10) with respect to A_v and B_v, respectively, and iteratively optimize them to speed up the convergence. All the terms in Eq. (10) are square terms, so that the derivatives can be computed by,

$$g_{A_v}(i) = 2I_n(i)(A_v(i)I_n(i) + B_v(i) - I_v(i)) + 2\lambda \nabla I_n(i)(A_v(i)\nabla I_n(i) - \nabla I_v(i)) + 2\beta \Delta A_v(i) \quad (11a)$$

$$g_{B_v}(i) = 2(A_v(i)I_n(i) + B_v(i) - I_v(i)) + 2\beta \Delta B_v(i) \quad (11b)$$

where Δ is the Laplacian operator.

Given the derivatives of the cost function with respect to A_v and B_v, we can update them iteratively by setting the descent direction as $-g_{A_v}$ and $-g_{B_v}$. In order to derive the optimal step lengths l_{A_v} and l_{A_v} for update, we substitute A_v and B_v in Eq.(7) by $A_v - l_{A_v} g_{A_v}$ and $B_v - l_{B_v} g_{B_v}$, and set the derivatives to 0. This gives

$$l_{A_v} = -\frac{\sum_i g_{A_v}(i)^2}{\sum_i g_{A_v}(i)^2 (I_n(i)^2 + \lambda \nabla I_n(i)^2 + \beta (\Delta g_{A_v}(i))^2)} \quad (12a)$$

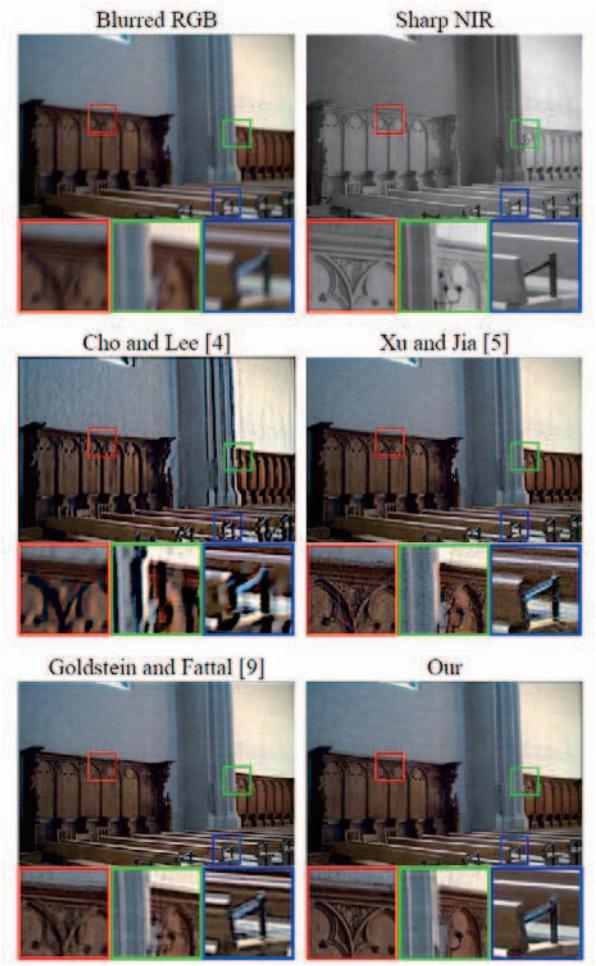

Fig. 5. Subjective comparisons of the proposed and state-of-the-art algorithms on two examples. The close-ups are shown under each image to show the details.

$$l_{B_v} = -\frac{\sum_i g_{B_v}(i)^2}{\sum_i g_{B_v}(i)(g_{B_v}(i) + \beta \Delta g_{b_v}(i))} \tag{12b}$$

We iteratively optimize Eq. (7) by computing Eqs. (11a), (11b) to get the optimal LLT maps A_v^* and B_v^*. The algorithm will be terminated when the optimal step lengths are both less than 1e-7, or the iteration number is larger than 100. The sharp RGB image can then be recovered by Eq. (5b).

4. EXPERIMENT RESULTS

In this section, we apply our algorithm on both synthetic and real captured datasets, and compare with the state-of-the-art deblurring algorithms both subjectively and qualitatively to verify the efficiency and effectiveness of the proposed method.

Comparison with state-of-the-art single image deblurring algorithms on synthetic image:

We test our algorithm on the blurred image simulated

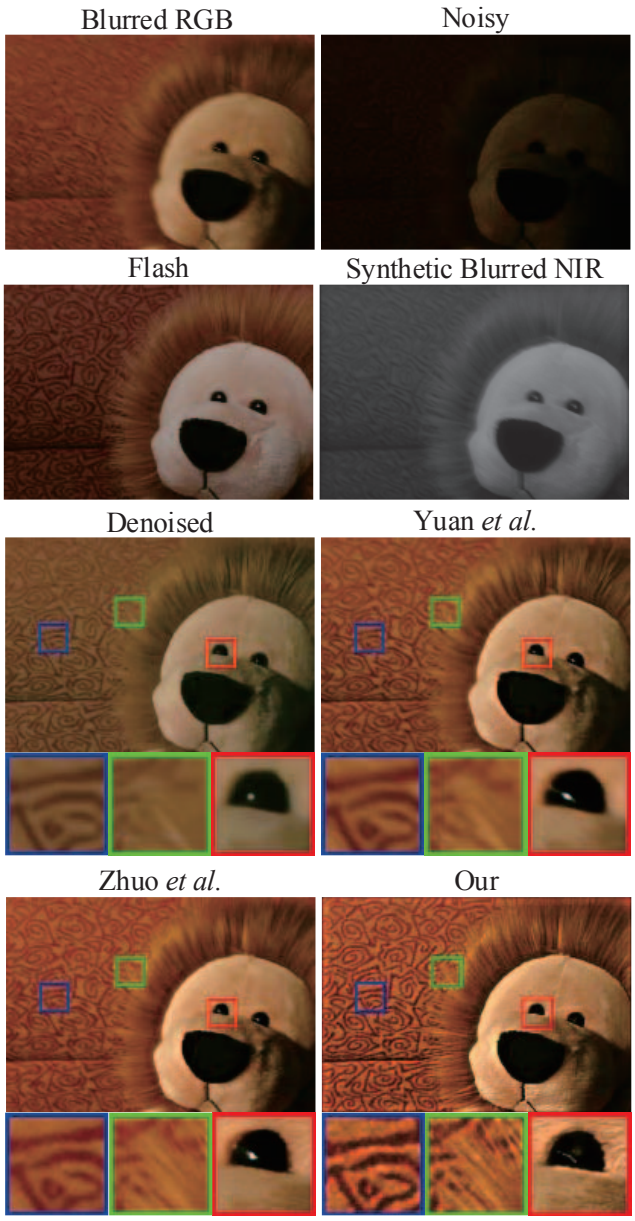

Fig. 6. Comparison with other multi-input methods.

gives the comparisons between our method and state-of-the-art deblurring algorithms. It is obviously that our results is much better than the deblurred ones which estimate the kernel without taking additional information into consideration.

Comparison with state-of-the-art multi-input deblurring methods on synthetic images:

We also compare our algorithm with other multi-input methods, as shown in Fig. 6. Because the additional input besides the blurred image is different for each method, we use the data from Zhuo *et al.* [17] and synthesize a pseudo NIR image from the additional flashed image by NIR = 0.1B +0.3G+0.6R, and the blurred NIR image is generated by blurring the pseudo NIR image with the kernel estimated by Zhuo *et al.* [17], as shown in top row of Fig. 6. The results of

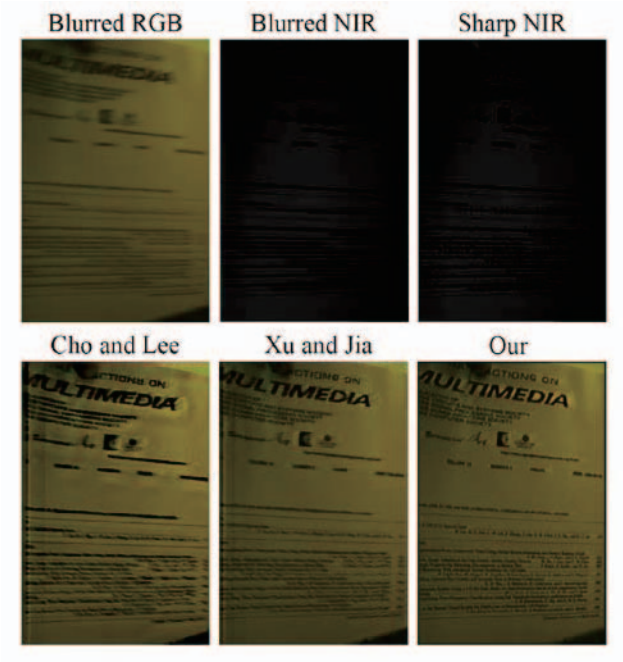

Fig.7. Results on real captured data.

denoising using Yuan *et al.* [16], Zhuo *et al.* [17] and our proposed method are shown at the bottom of Fig. 6. Our results appear sharper and have better contrasts.

It is worth noting that the red region is not consistent between the blurred RGB/NIR image pair because the pseudo NIR image was derived from flash images and specularity is strong in that region under flash. Therefore, the proposed algorithm cannot handle the specular effect in this region, but the result is still reasonable.

Comparison with state-of-the-art deblurring algorithms on real images:

We test our algorithm and compare the results with the state-of-the-art blind image deblurring algorithms on real captured data as well. Fig. 8 shows the input set and results of the proposed algorithm, Cho and Lee [4], Goldstein and Fattal [9], Xu and Jia [5] as well as Krishnan and Fergus [7] respectively. As in most real cases, the blur is slightly non-uniform (spatially varying). Thus, the other methods cannot achieve good performance in all the regions. Conversely, our algorithm gives promising results in most regions.

The proposed algorithm achieves promising results while other traditional methods almost fail on the second example for its non-uniformity. Our algorithm still has problems in the some regions, mainly in the over/under exposure regions in the NIR channel, which are caused by specific NIR illuminations. Therefore, we recommend to use the uniform NIR light in practice to solve this problem.

5. DISCUSSION AND CONCLUSION

We propose a novel LLT property which is valid for both blurred and sharp RGB/NIR image pairs. Based on this property, a multispectrum camera based framework is proposed to recover the sharp RGB images from a blurred RGB image and the corresponding NIR image sequence. According to experiments, the proposed method achieves much better performance on both effectiveness and efficiency aspects than the state-of-the-art debluring algorithms.

REFERENCES

[1] J. Jia, "Single Image Motion Deblurring Using Transparency," in CVPR, 2007, pp. 1–8.
[2] N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," in CVPR, 2008, pp. 1–8.
[3] N. Joshi, C.L. Zitnick, R. Szeliski, and D. Kriegman, "Image deblurring and denoising using color priors," in CVPR, 2009, pp. 1550–1557.
[4] S. Cho and S. Lee, "Fast motion deblurring," ACM Trans. on Graphics, vol. 28, no. 5, pp. 1, Dec. 2009.
[5] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in ECCV, 2010, pp. 157–170.
[6] T. S. Cho, S. Paris, B. KP Horn, and W. T. Freeman, "Blur kernel estimation using the radon transform," in CVPR, 2011, pp. 241–248.
[7] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in CVPR, 2011, pp. 233–240.
[8] W. Hu and J. Xue, "PSF Estimation via Gradient Domain Correlation," IEEE Trans. on Image Processing, vol. 21, no. 1, pp. 386–392, 2012.
[9] A. Goldstein and R. Fattal, "Blur-Kernel Estimation from Spectral," in ECCV, 2012, pp. 1–8.
[10] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," in ICCP, 2013.
[11] L. Xu, S. Zheng, and J. Jia, "Unnatural L 0 Sparse Representation for Natural Image Deblurring," in CVPR, 2013, pp. 1838–1857.
[12] Y. HaCohen, E. Shechtman, and D. Lischinski, "Deblurring by example using dense correspondence," in ICCV, 2013, pp. 1–8.
[13] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in CVPR, 2009.
[14] N. Joshi, S.B. Kang, C.L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," ACM Trans. on Graphics, vol.29, no. 4, Jul. 2010, pp. 2560–2567.
[15] A. Agrawal and R. Raskar, "Optimal single image capture for motion deblurring," in CVPR, 2009.
[16] L. Yuan, J. Sun, L. Quan, and H.Y. Shum, "Image deblurring with blurred/noisy image pairs," ACM Trans. on Graphics, vol. 26, no. 3, Jul. 2007.
[17] S. Zhuo, D. Guo, and T. Sim, "Robust flash deblurring," in CVPR, 2010, pp. 2440–2447.
[18] Y.W. Tai, H. Du, M.S. Brown, and S. Lin, "Image/video deblurring using a hybrid camera," in CVPR, 2008, pp. 1–8.
[19] M. Ben-Ezra and S.K. Nayar, "Motion-based motion deblurring," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 26, no. 6, pp. 689–698, Jun. 2004.
[20] C. Fredembach and S. Süsstrunk, "Colouring the near-infrared," in Color and Imaging Conference. Society for Imaging Science and Technology, 2008. pp. 176–182.
[21] L. Schaul, C. Fredembach, and S. Süsstrunk, "Color image dehazing using the near-infrared," in ICIP, 2009, pp. 1629–1632.
[22] D. Krishnan and R. Fergus, "Dark flash photography," in ACM Transactions on Graphics, SIGGRAPH 2009 Conference Proceedings, 2009.
[23] Eric P Bennett, John L Mason, and Leonard McMillan, "Multispectral bilateral video fusion," IEEE Trans. on Image Processing, vol. 16, no. 5, pp. 1185–1194, 2007.
[24] E. P Bennett, M. Uyttendaele, C L. Zitnick, R. Szeliski, and S.B. Kang, "Video and image bayesian demosaicing with a two color image prior," in ECCV, 2006, pp. 508–521.
[25] M. Brown and S. Süsstrunk, "Multi-spectral sift for scene category recognition," in CVPR, 2011, pp. 177–184.
[26] JAI Inc., "AD-080CL Multispectral camera," http://www.jai.com/en/products/ad-080cl/, [Online; accessed 4-Nov.-2013].
[27] Q. Yan, X. Shen, L. Xu, S. Zhuo, X. Zhang, L. Shen and J. Jia, "Cross-field joint image restoration via scale map", in ICCP, 2013.