Cross-Field Joint Image Restoration via Scale Map
Supplementary File

Qiong Yan[§] Xiaoyong Shen[§] Li Xu[§] Shaojie Zhuo[†]
Xiaopeng Zhang[†] Liang Shen[†] Jiaya Jia[§]

[§]The Chinese University of Hong Kong [†]Qualcomm Incorporated
http://www.cse.cuhk.edu.hk/leojia/projects/crossfield

1. Numerical Solution

Here we explain all details on how to solve the problem we proposed in the paper in Section 3. We rewrite the original energy function defined in Eq. (14) in the paper as

$$E(s,I) = E_1(s,I) + \lambda E_2(I) + \beta E_3(\nabla s). \tag{1}$$

With the approximation described in the paper that $\rho(x) \approx \phi(x)\cdot x^2$, the first two terms can be approximated as

$$\begin{aligned}
E_1(s,I) &= \sum_i \Big(\rho(|s_i - p_{i,x}\nabla_x I_i|) + \rho(|s_i - p_{i,y}\nabla_y I_i|)\Big) \\
&\approx \sum_i \Big(\phi(s_i - p_{i,x}\nabla_x I_i)\cdot(s_i - p_{i,x}\nabla_x I_i)^2 + \phi(s_i - p_{i,y}\nabla_y I_i)\cdot(s_i - p_{i,y}\nabla_y I_i)^2\Big) \\
&= \sum_i \Big((s_i - p_{i,x}\nabla_x I_i)\cdot\phi(s_i - p_{i,x}\nabla_x I_i)\cdot(s_i - p_{i,x}\nabla_x I_i) \\
&\qquad + (s_i - p_{i,y}\nabla_y I_i)\cdot\phi(s_i - p_{i,y}\nabla_y I_i)\cdot(s_i - p_{i,y}\nabla_y I_i)\Big), \qquad (2) \\
E_2(I) &= \sum_i \rho(|I_i - I_{0,i}|) \\
&\approx \sum_i \phi(I_i - I_{0,i})\cdot(I_i - I_{0,i})^2 \\
&= \sum_i (I_i - I_{0,i})\cdot\phi(I_i - I_{0,i})\cdot(I_i - I_{0,i}). \qquad (3)
\end{aligned}$$

They can be written in vector forms respectively as

$$E_1(\mathbf{s},\mathbf{I}) = (\mathbf{s} - P_x C_x \mathbf{I})^T A_x (\mathbf{s} - P_x C_x \mathbf{I}) + (\mathbf{s} - P_y C_y \mathbf{I})^T A_y (\mathbf{s} - P_y C_y \mathbf{I}), \tag{4}$$

$$E_2(\mathbf{I}) = (\mathbf{I} - \mathbf{I}_0)^T B (\mathbf{I} - \mathbf{I}_0), \tag{5}$$

where $\mathbf{s}$, $\mathbf{I}$ and $\mathbf{I}_0$ are vector representations of s, I and I_0. C_x and C_y are discrete backward difference matrices that are used to compute image gradients in the x- and y- directions. P_x, P_y, A_x, A_y and B are diagonal matrices, whose i-th diagonal elements are defined as

$$\begin{aligned}
(P_x)_{ii} &= p_{i,x}, \quad (A_x)_{ii} = \phi(s_i - p_{i,x}\nabla_x I_i), \\
(P_y)_{ii} &= p_{i,y}, \quad (A_y)_{ii} = \phi(s_i - p_{i,y}\nabla_y I_i), \\
B_{ii} &= \phi(I_i - I_{0,i}).
\end{aligned}$$

The last term in Eq. (1) is regularization on s. Its definition is

$$E_3(\nabla s) = \sum_i \left(\mu_{i,1}(\mathbf{v}_{i,1}^T \nabla s_i)^2 + \mu_{i,2}(\mathbf{v}_{i,2}^T \nabla s_i)^2\right). \tag{6}$$

Considering one single term inside the summation, following definitions in the paper, we get

$$\mathbf{v}_{i,1} = \frac{1}{|\nabla G_i|}\begin{pmatrix}\nabla_x G_i \\ \nabla_y G_i\end{pmatrix}, \quad \mathbf{v}_{i,2} = \frac{1}{|\nabla G_i|}\begin{pmatrix}\nabla_y G_i \\ -\nabla_x G_i\end{pmatrix}, \tag{7}$$

To simplify notation, let $v_{i,x} = \nabla_x G_i/|\nabla G_i|$ and $v_{i,y} = \nabla_y G_i/|\nabla G_i|$. It makes $\mathbf{v}_{i,1} = (v_{i,x}, v_{i,y})^T$ and $\mathbf{v}_{i,2} = (v_{i,y}, -v_{i,x})^T$. Now we expand the term $E_3(\nabla s_i)$ as

$$\begin{aligned}
E_3(\nabla s_i) &= \mu_{i,1}\Big(v_{i,x}\nabla_x s_i + v_{i,y}\nabla_y s_i\Big)^2 + \mu_{i,2}\Big(v_{i,y}\nabla_x s_i - v_{i,x}\nabla_y s_i\Big)^2 \\
&= \mu_{i,1}\Big(v_{i,x}^2(\nabla_x s_i)^2 + 2v_{i,x}v_{i,y}(\nabla_x s_i)(\nabla_y s_i) + v_{i,y}^2(\nabla_y s_i)^2\Big) \\
&\quad + \mu_{i,2}\Big(v_{i,y}^2(\nabla_x s_i)^2 - 2v_{i,x}v_{i,y}(\nabla_x s_i)(\nabla_y s_i) + v_{i,x}^2(\nabla_y s_i)^2\Big) \\
&= \nabla_x s_i\Big(\mu_{i,1}v_{i,x}^2 + \mu_{i,2}v_{i,y}^2\Big)\nabla_x s_i + \nabla_y s_i\Big(\mu_{i,1}v_{i,y}^2 + \mu_{i,2}v_{i,x}^2\Big)\nabla_y s_i \\
&\quad + 2\nabla_y s_i\Big(\mu_{i,1} - \mu_{i,2}\Big)v_{i,x}v_{i,y}\nabla_x s_i.
\end{aligned} \tag{8}$$

With diagonal matrices, Σ_1, Σ_2, V_x and V_y are with main-diagonal elements

$$\begin{aligned}
(\Sigma_1)_{ii} &= \mu_{i,1}, \quad (V_x)_{ii} = \nabla_x G_i / \max(|\nabla G_i|, \varepsilon), \\
(\Sigma_2)_{ii} &= \mu_{i,2}, \quad (V_y)_{ii} = \nabla_y G_i / \max(|\nabla G_i|, \varepsilon),
\end{aligned}$$

where ε is introduced to avoid division by zero. The regularization term can be rewritten in a vector form as

$$E_3(\mathbf{s}) = (C_x\mathbf{s})^T(\Sigma_1 V_x^2 + \Sigma_2 V_y^2)(C_x\mathbf{s}) + (C_y\mathbf{s})^T(\Sigma_2 V_x^2 + \Sigma_1 V_y^2)(C_y\mathbf{s}) + 2(C_y\mathbf{s})^T(\Sigma_1 - \Sigma_2)V_x V_y(C_x\mathbf{s}). \tag{9}$$

Let

$$L = C_x^T(\Sigma_1 V_x^2 + \Sigma_2 V_y^2)C_x + C_y^T(\Sigma_2 V_x^2 + \Sigma_1 V_y^2)C_y + 2C_y^T(\Sigma_1 - \Sigma_2)V_x V_y C_x. \tag{10}$$

The final term is expressed as

$$E_3(\mathbf{s}) = \mathbf{s}^T L\mathbf{s}. \tag{11}$$

Finally, the overall energy function is

$$\begin{aligned}
E(\mathbf{s}, \mathbf{I}) =& (\mathbf{s} - P_x C_x \mathbf{I})^T A_x(\mathbf{s} - P_x C_x \mathbf{I}) + (\mathbf{s} - P_y C_y \mathbf{I})^T A_y(\mathbf{s} - P_y C_y \mathbf{I}) \\
&+ \lambda(\mathbf{I} - \mathbf{I}_0)^T B(\mathbf{I} - \mathbf{I}_0) + \beta\mathbf{s}^T L\mathbf{s}.
\end{aligned} \tag{12}$$

In this function, A_x and A_y depends on I and s; B depends on I only. To solve it, we adopt a two-step iterative solver, fully described in the paper in section 3.1.

2. More Results

As discussed in our paper, images captured in different fields can have large variation in gradient magnitudes and directions. Directly applying joint filtering for image restoration could blur structures that are weak in their clean counterpart. In the meantime, strong edges caused by highlight/shadow in the guidance will also be introduced to the output, making the result implausible. A result of using the guided filtering [5] is shown in Fig. 1(b). Another way to utilize filtering is to transfer details of the clean image to the noisy one. In [7], joint bilateral filters are used to denoise and separate the detail layer. We show a result via transferring detail layer using this method in Fig. 1(c). This method finds similar problems as guided filtering, and additionally results in blurred edges where gradient direction reverts in the two inputs. Our method, by

handling all these discrepancies in the scale map s, is able to restore a sharp and clean image as shown in Fig. 1(d). Another example with flash/non-flash image pairs is shown in Fig. 2 with comparison to the result of [7].

Fig. 3 shows the whole-image inputs and results, corresponding to the close-ups in Fig. 8 in the paper. Our method only needs the IR image guidance, and the result is comparable to that of [6], which are produced with both UV (ultraviolet) and IR guidance images. Our edges and details are very clear.

In Fig. 4, we show the whole-image inputs and results with close-ups already presented in Figs. 6 and 7 in our paper. It is for comparison to the state-of-the-art BM3D single-image denoising method [1], and the detail-transfer approach [9] utilizing the WLS filter [2]. The two input images are respectively RGB and NIR images taken from the same location. Our result contains many readable small characters and complicated texture. One more result produced by our method is shown in Fig. 5.

In Fig. 6, we show the result by applying our guided restoration method to haze removal, which is an important application and unfortunately still suffers from significant noise and JPEG visual artifacts after haze suppression. These noise and compression artifacts are caused by strong contrast enhancement in local regions, which originally in the hazed image are almost visually unnoticeable. Our method greatly improves the result quality after restoring the dehazed image with the NIR image guidance.

Besides the flash/non-flash and RGB/NIR image settings illustrated in previous examples, our method is also applicable to other cross-field restoration problems. For example, images captured at night are usually dark with many structures being eliminated. An example is shown in Fig. 7(a). Brightening it by increasing color contrast in PhotoShop improves the image, and in the meantime boosts image noise, as shown in (c). With a reference day image as described in [8], our method is able to remove the annoying noise with high-quality edge preservation. The output (d) contains nearly all fine details and keeps, as well, strong and clean structures with the day image information.

Additionally, depth images from Kinect or other realtime depth capture devices contain strong noise caused by missing or erroneous data for many pixels. We similarly apply our restoration method with the corresponding RGB visible images as guidance to raw depth maps and produce much improved results, as shown in Fig. 8.

References

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095, 2007.

[2] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.*, 27(3), 2008.

[3] C. Fredembach and S. Ssstrunk. Colouring the near infrared. In *Proceedings of the IS&T/SID 16th Color Imaging Conference*, pages 176–182, 2008.

[4] K. He, J. Sun, and X. Tang. Guided image filtering. In *ECCV (1)*, pages 1–14, 2010.

[5] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(12):2341–2353, 2011.

[6] D. Krishnan and R. Fergus. Dark flash photography. *ACM Trans. Graph.*, 28(3), 2009.

[7] G. Petschnigg, R. Szeliski, M. Agrawala, M. F. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. *ACM Trans. Graph.*, 23(3):664–672, 2004.

[8] R. Raskar, A. Ilie, and J. Yu. Image fusion for context enhancement and video surrealism. In *NPAR*, pages 85–152, 2004.

[9] S. Zhuo, X. Zhang, X. Miao, and T. Sim. Enhancing low light images using near infrared flash images. In *ICIP*, pages 2537–2540, 2010.

(a) Flash/Non-flash Images

(b) Result of [4]

(c) Result of [7]

(d) Our Result

(e) Close-ups

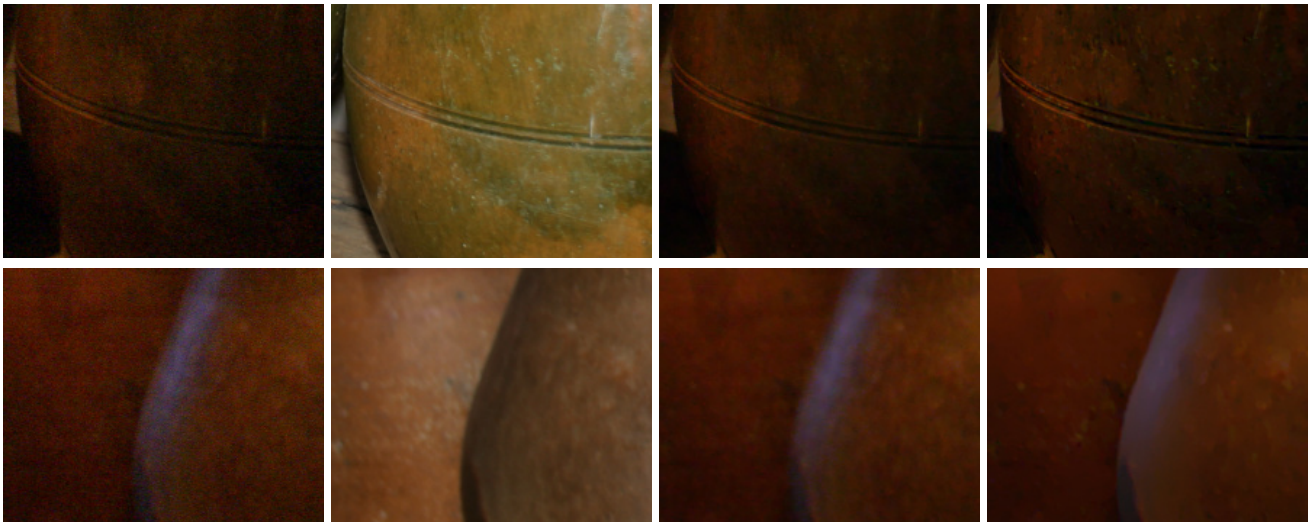
Figure 1. Flash/non-flash restoration. The input images and result of [7] are obtained from the original paper.

(a) Non-flash Image

(b) Flash Image

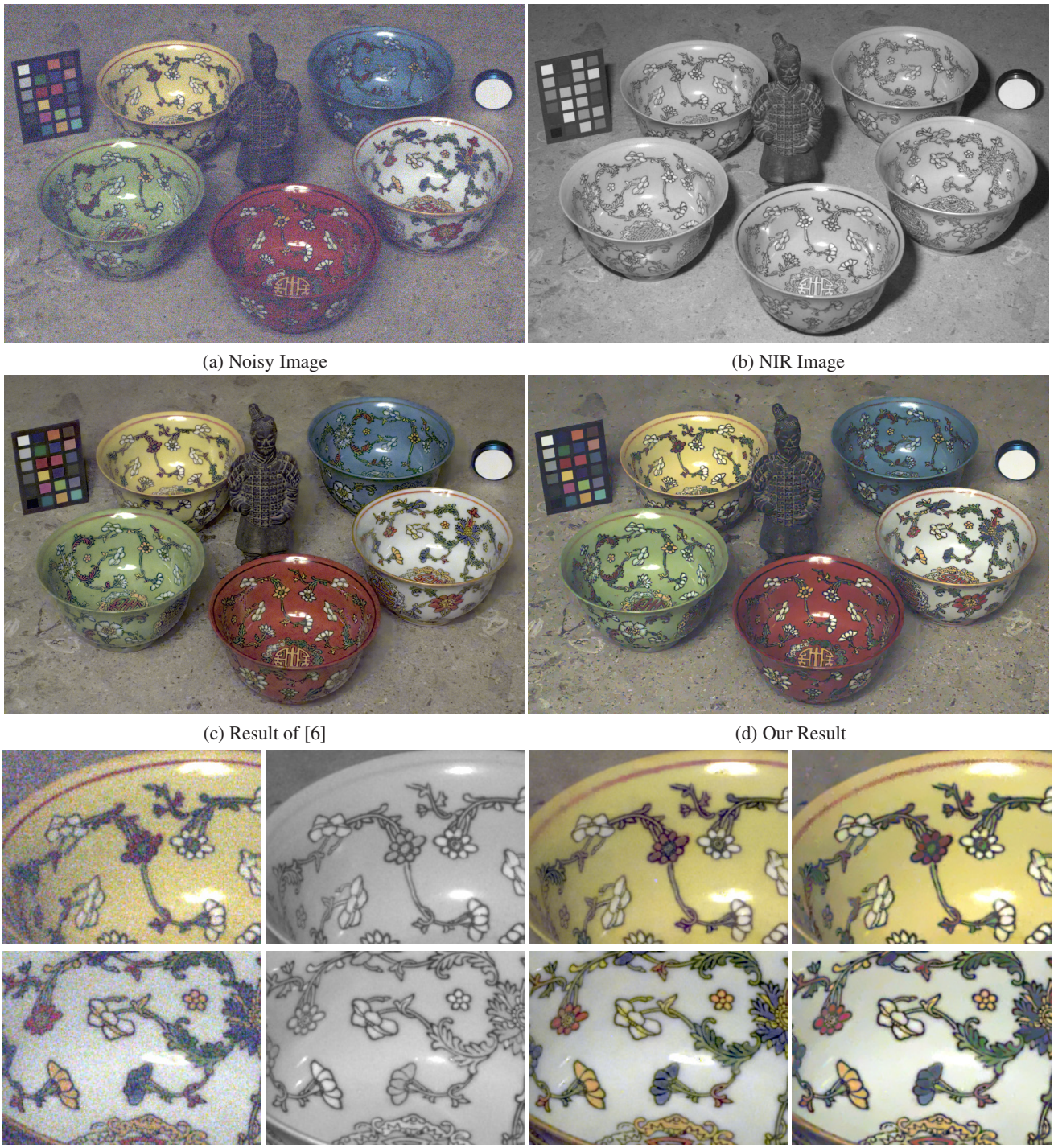
(c) Result of [7]

(d) Our Result



(e) Close-ups

Figure 2. Flash/non-flash restoration. The input images and result of [7] are obtained from the original paper.



(a) Noisy Image

(b) NIR Image

(c) Result of [6]

(d) Our Result

(e) Close-ups

Figure 3. RGB/flashed-NIR restoration. The input images and result of [6] are obtained from the original paper. Note the result of [6] is outputed from two guidance images while ours is with only one IR image as guidance.

(a) Noisy Image (b) NIR Image

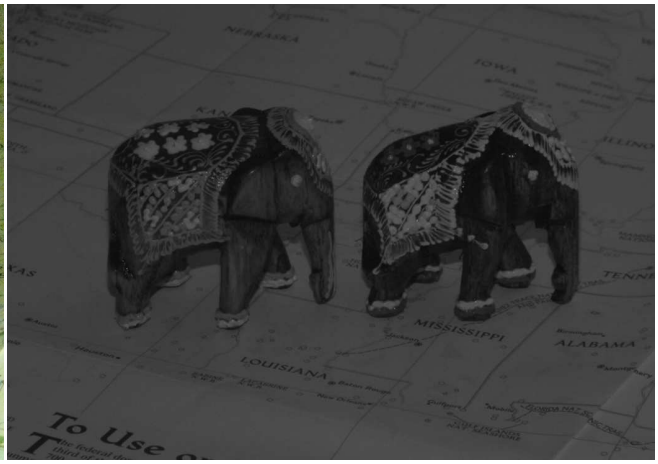
(c) BM3D [1] (d) Result of [9]

(e) Our Result (f) Our scale map

(g) Close-ups of (c)-(f)

Figure 4. RGB/flashed-NIR restoration. Results of different methods are compared.

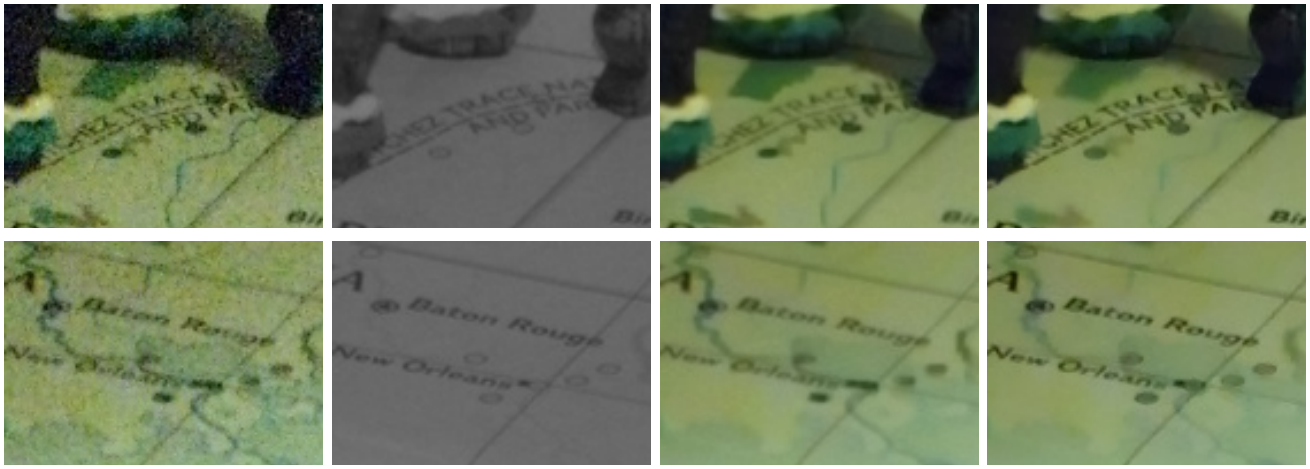
(a) Noisy Image



(b) NIR Image

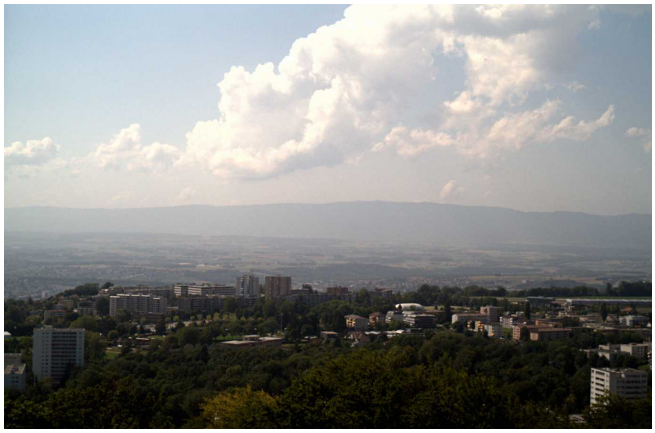
(c) Result of [9]

(d) Our Result



(e) Close-ups

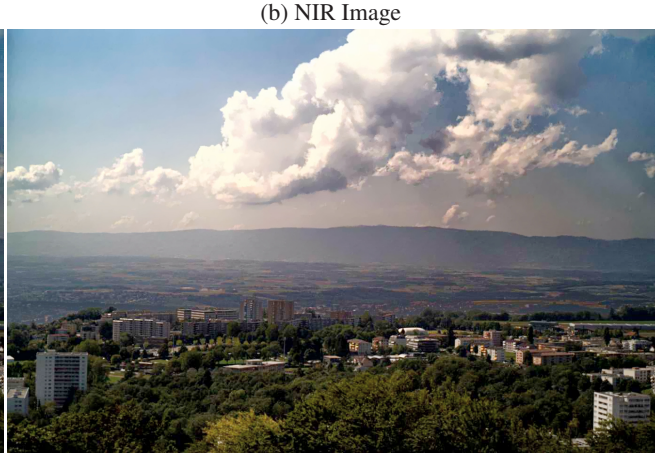
Figure 5. One more RGB/flashed-NIR restoration example with result comparison.



(a) Haze Image in [3]

(b) NIR Image

(c) Result of [5]

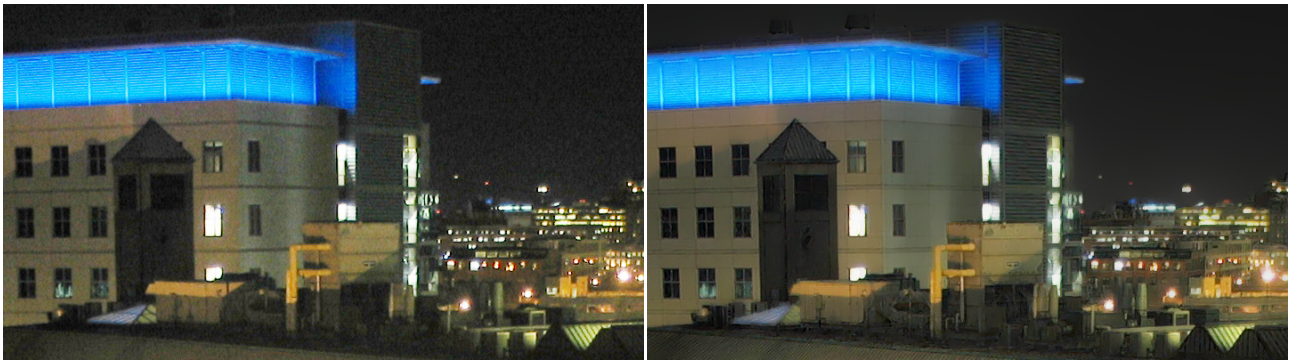


(d) Our Result from (b) and (c)

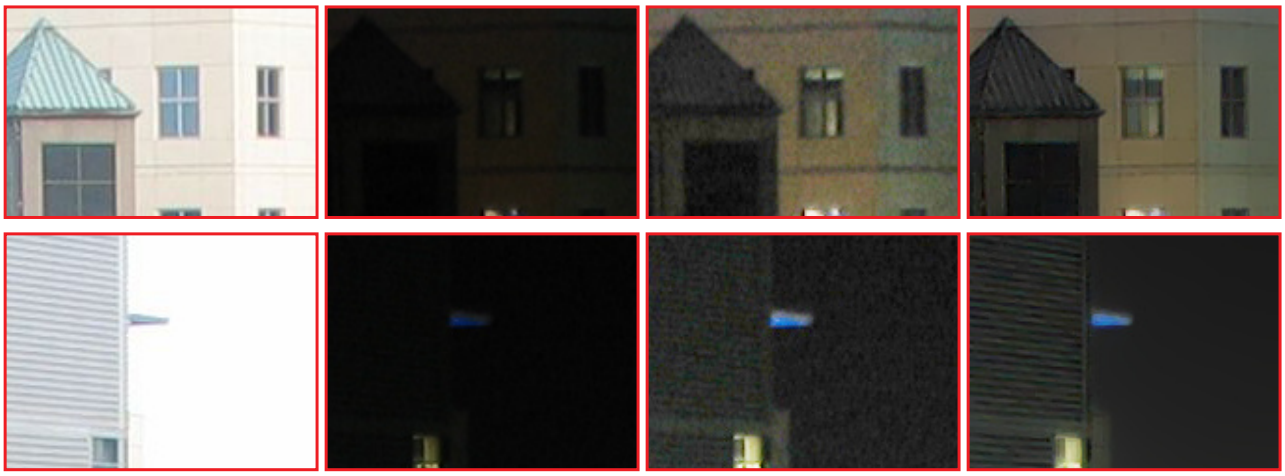
(e) Close-ups

Figure 6. Guided image restoration from haze images. Note our method takes the input of the dehazed image and the NIR guidance image shown in (c) and (b) to improve the result quality.

(a) Night Image from [8] (b) Day Image from [8]

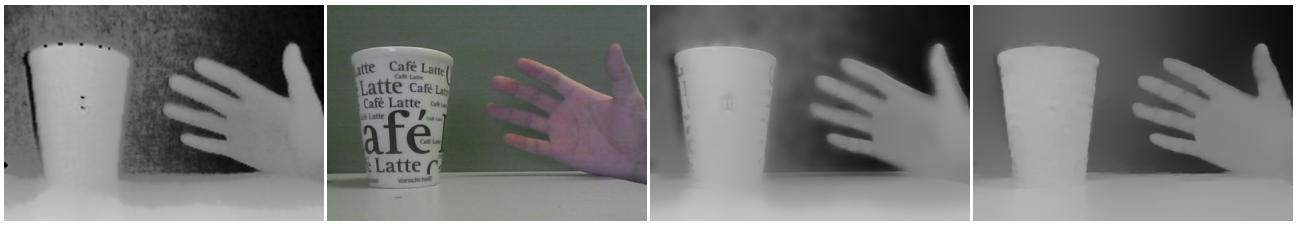


(c) Enhanced Night Image (d) Our Result from (b) and (c)



(e) Close-ups

Figure 7. Image restoration from day and night images. Our method takes the input of the day and enhanced night images in (b) and (c) to suppress noise and other visual artifacts.



(a) Raw Depth Map (b) RGB Image (c) Result by [4] (d) Our Result

Figure 8. Depth map restoration with RGB image as guidance. Compared with [4], our method completely removes the background noise and does not blur edges.