Multiple temporal resolution imaging and processing for obtaining HDR image with low motion blur and noise

Takuro Kosaka, Takayuki Hamamoto
Graduate School of Engineering, Tokyo University of Science
6-3-1 Niijuku, Katsushika, Tokyo, 125-8585 Japan
E-mail: {kosaka, hamamoto}@isl.ee.kagu.tus.ac.jp

Abstract In this paper, we propose a new method for obtaining high dynamic range images with low motion blur and noise called "multiple temporal resolution imaging and processing" by controlling the exposure pattern of a 4×4 pixel block unit in a single image sensor. We demonstrate the effectiveness of our method through a simulation.

Keywords: high dynamic range, denoising, deblurring, multiple temporal resolution, exposure control

1. Introduction

Over the years, the demand for high dynamic range (HDR) imaging has increased in the domains of automobiles and security. An HDR image is obtained by synthesizing multiple exposure images. However, motion artifacts, such as ghosts, are known to be caused if multiple exposure images are captured sequentially [1].

Alternatively, Nayer et al. [2] proposed the spatially varying exposure (SVE) technique. In this method, multiple exposure images are captured simultaneously using an on-chip neutral-density filter. However, this optical SVE method has a trade-off relationship between signal to noise ratio (S/N) and motion blur because the exposure time is equal in all pixels. For example, we can obtain a high S/N image using low frame rate images, however, this leads to an increase in motion blur. Similarly, it is possible to obtain a low motion blur image using high frame rate images; however, this leads to an increase in noise.

In this paper, we propose a multiple temporal resolution imaging and processing method using a single image sensor that can be assigned an exposure pattern of a 4×4 pixel block unit. Multiple frame rate images can be obtained simultaneously. High frame rate images have low motion blur information, whereas low frame rate images have high S/N information. We reconstruct HDR images with low motion blur and noise from multiple temporal resolution images.

2. Multiple temporal resolution imaging and processing using a single image sensor

2.1 Imaging method

Fig. 1 shows an overview of our imaging and processing method. We have used an interleave-exposure-control image sensor [3]. In this sensor, we can assign an exposure pattern of a 4×4 pixel block unit and obtain multiple frame rate images simultaneously. In this paper, we have considered the case of a color filter array (CFA) mounted on the prototype image sensor.

2.2 HDR image reconstruction

2.2.1 Bayer and SVE demosaicing

Let the captured raw images taken with high, middle, and low frame rates be $\boldsymbol{y}_{\mathrm{HR}}$, $\boldsymbol{y}_{\mathrm{MR}}$, and $\boldsymbol{y}_{\mathrm{LR}}$, respectively. Consider the following:

$$\boldsymbol{y}_{\mathrm{HR}} = \left(\boldsymbol{y}_{\mathrm{HR}}^{(1)}, \dots, \boldsymbol{y}_{\mathrm{HR}}^{(N_{\mathrm{HR}})}\right), \quad (1)$$

$$\boldsymbol{y}_{\mathrm{MR}} = \left(\boldsymbol{y}_{\mathrm{MR}}^{(1)}, \dots, \boldsymbol{y}_{\mathrm{MR}}^{(N_{\mathrm{MR}})}\right), \quad (2)$$

$$\boldsymbol{y}_{\mathrm{LR}} = \left(\boldsymbol{y}_{\mathrm{LR}}^{(1)}\right), \quad (3)$$

where N_{HR} and N_{MR} are scale factors of high and middle frame rates against a low frame rate. In this step, we construct color images $\boldsymbol{x}_{\mathrm{HR}}$, $\boldsymbol{x}_{\mathrm{MR}}$, and $\boldsymbol{x}_{\mathrm{LR}}$ from $\boldsymbol{y}_{\mathrm{HR}}$, $\boldsymbol{y}_{\mathrm{MR}}$, and $\boldsymbol{y}_{\mathrm{LR}}$ using a bicubic method with color-differential interpolation.

2.2.2 Motion area detection

An inter-frame difference calculation detects a motion area. The motion area binary masks $\boldsymbol{M}_{\mathrm{HR}}$, $\boldsymbol{M}_{\mathrm{MR}}$, and $\boldsymbol{M}$ are constructed as:

$$\boldsymbol{M}_{\mathrm{HR},p} = \begin{cases} 1 & \text{if } \exists n < N_{\mathrm{HR}}, \\ & \left|\mathrm{GF}\left(\boldsymbol{x}_{\mathrm{HR},p}^{(n)}, \boldsymbol{x}_{\mathrm{LR},p}^{(1)}\right) - \mathrm{GF}\left(\boldsymbol{x}_{\mathrm{HR},p}^{(n+1)}, \boldsymbol{x}_{\mathrm{LR},p}^{(1)}\right)\right| \geq th_{\mathrm{HR}} \\ 0 & \text{otherwise}, \end{cases} \quad (4)$$

$$\boldsymbol{M}_{\mathrm{MR},p} = \begin{cases} 1 & \text{if } \exists n < N_{\mathrm{MR}}, \\ & \left|\mathrm{GF}\left(\boldsymbol{x}_{\mathrm{MR},p}^{(n)}, \boldsymbol{x}_{\mathrm{LR},p}^{(1)}\right) - \mathrm{GF}\left(\boldsymbol{x}_{\mathrm{MR},p}^{(n+1)}, \boldsymbol{x}_{\mathrm{LR},p}^{(1)}\right)\right| \geq th_{\mathrm{MR}} \\ 0 & \text{otherwise}, \end{cases} \quad (5)$$

$$\boldsymbol{M} = \boldsymbol{M}_{\mathrm{HR}} + \boldsymbol{M}_{\mathrm{MR}}, \quad (6)$$

where p is the position of a pixel, and n is a natural number. th_{HR} and th_{MR} are thresholds of inter-frame difference calculation. $\mathrm{GF}(\boldsymbol{i}, \boldsymbol{j})$ denotes the operation of a guided filter [4] with input image $\boldsymbol{i}$ and guidance image $\boldsymbol{j}$. We use the guided filter to improve motion area detection accuracy. When $\boldsymbol{M}_p = 1$,

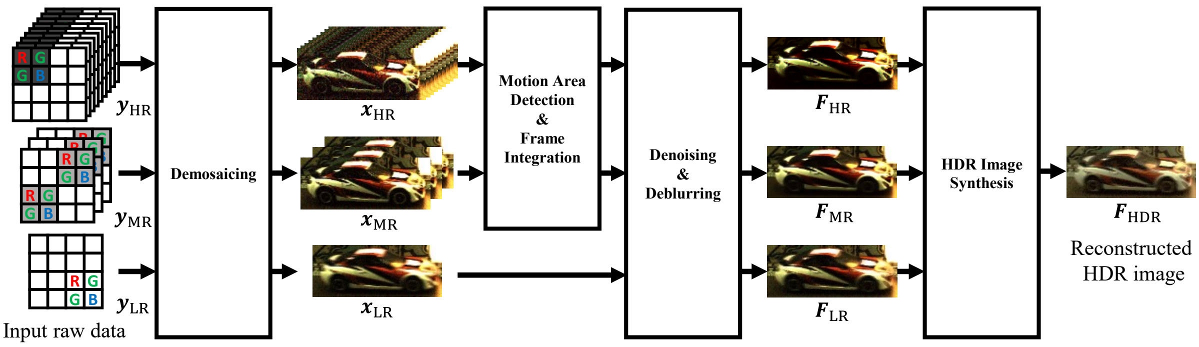


Fig. 1: Overview of our HDR image reconstruction.

the pixel p is the motion area.

2.2.3 Frame integration

$\boldsymbol{x}_{\mathrm{HR}}$ and $\boldsymbol{x}_{\mathrm{MR}}$ contain multiple frame information. They are integrated with single images $\boldsymbol{x}'_{\mathrm{HR}}$ and $\boldsymbol{x}'_{\mathrm{MR}}$ as follows:

$$\boldsymbol{x}'_{\mathrm{HR}} = \frac{1}{N_{\mathrm{HR}}}\overline{\boldsymbol{M}}\sum_{n=1}^{N_{\mathrm{HR}}} \boldsymbol{x}_{\mathrm{HR}}^{(n)} + \boldsymbol{M}\mathrm{GF}\left(\boldsymbol{x}_{\mathrm{HR}}^{(\mathrm{center})}, \boldsymbol{x}_{\mathrm{HR}}^{(\mathrm{center})}\right), \quad (7)$$

$$\boldsymbol{x}'_{\mathrm{MR}} = \frac{1}{N_{\mathrm{MR}}}\overline{\boldsymbol{M}}\sum_{n=1}^{N_{\mathrm{MR}}} \boldsymbol{x}_{\mathrm{MR}}^{(n)} + \boldsymbol{M}\mathrm{GF}\left(\boldsymbol{x}_{\mathrm{MR}}^{(\mathrm{center})}, \boldsymbol{x}_{\mathrm{MR}}^{(\mathrm{center})}\right), \quad (8)$$

where $\boldsymbol{x}_{\mathrm{HR}}^{(\mathrm{center})}$ and $\boldsymbol{x}_{\mathrm{MR}}^{(\mathrm{center})}$ represent the center frame of $\boldsymbol{x}_{\mathrm{HR}}$ and $\boldsymbol{x}_{\mathrm{MR}}$ respectively. In this way, we obtain denoised images $\boldsymbol{x}'_{\mathrm{HR}}$ and $\boldsymbol{x}'_{\mathrm{MR}}$.

2.2.4 Denoising and deblurring

In this step, we reconstruct denoised and deblurred multiple exposure images $\boldsymbol{F}_{\mathrm{HR}}$, $\boldsymbol{F}_{\mathrm{MR}}$, and $\boldsymbol{F}_{\mathrm{LR}}$ using the guided filter. Thus,

$$\begin{aligned}\boldsymbol{F}_{\mathrm{HR}} = {} & \overline{\boldsymbol{M}}\boldsymbol{U}_{\mathrm{LR}}\mathrm{GF}(\boldsymbol{x}'_{\mathrm{HR}}, \boldsymbol{x}_{\mathrm{LR}}) + \overline{\boldsymbol{M}}\overline{\boldsymbol{U}}_{\mathrm{LR}}\boldsymbol{U}_{\mathrm{MR}}\mathrm{GF}(\boldsymbol{x}'_{\mathrm{HR}}, \boldsymbol{x}'_{\mathrm{MR}}) \\ & + \overline{\boldsymbol{M}}\overline{\boldsymbol{U}}_{\mathrm{LR}}\overline{\boldsymbol{U}}_{\mathrm{MR}}\boldsymbol{x}'_{\mathrm{HR}} + \boldsymbol{M}\boldsymbol{x}'_{\mathrm{HR}} ,\end{aligned} \quad (9)$$

$$\begin{aligned}\boldsymbol{F}_{\mathrm{MR}} = {} & \overline{\boldsymbol{M}}\boldsymbol{U}_{\mathrm{LR}}\mathrm{GF}(\boldsymbol{x}'_{\mathrm{MR}}, \boldsymbol{x}_{\mathrm{LR}}) + \boldsymbol{M}\boldsymbol{U}_{\mathrm{HR}}\mathrm{GF}(\boldsymbol{x}'_{\mathrm{MR}}, \boldsymbol{x}'_{\mathrm{HR}}) \\ & + \overline{\boldsymbol{M}}\overline{\boldsymbol{U}}_{\mathrm{LR}}\boldsymbol{x}'_{\mathrm{MR}} + \boldsymbol{M}\overline{\boldsymbol{U}}_{\mathrm{HR}}\boldsymbol{x}'_{\mathrm{MR}} ,\end{aligned} \quad (10)$$

$$\begin{aligned}\boldsymbol{F}_{\mathrm{LR}} = {} & \boldsymbol{M}\boldsymbol{U}_{\mathrm{HR}}\mathrm{GF}(\boldsymbol{x}_{\mathrm{LR}}, \boldsymbol{x}_{\mathrm{LR}}) + \boldsymbol{M}\overline{\boldsymbol{U}}_{\mathrm{HR}}\boldsymbol{U}_{\mathrm{MR}}\mathrm{GF}(\boldsymbol{x}_{\mathrm{LR}}, \boldsymbol{x}'_{\mathrm{MR}}) \\ & + \boldsymbol{M}\overline{\boldsymbol{U}}_{\mathrm{HR}}\overline{\boldsymbol{U}}_{\mathrm{MR}}\boldsymbol{x}_{\mathrm{LR}} + \overline{\boldsymbol{M}}\boldsymbol{x}_{\mathrm{LR}} ,\end{aligned} \quad (11)$$

where $\boldsymbol{U}_{\mathrm{HR}}$, $\boldsymbol{U}_{\mathrm{MR}}$, and $\boldsymbol{U}_{\mathrm{LR}}$ are binary masks indicating that pixel p is effective pixel.

2.2.5 HDR image synthesis

Finally, we synthesize the HDR image $\boldsymbol{F}_{\mathrm{HDR}}$ as:

$$\boldsymbol{F}_{\mathrm{HDR}} = \frac{1}{3}(\boldsymbol{F}_{\mathrm{HR}} + \boldsymbol{F}_{\mathrm{MR}} + \boldsymbol{F}_{\mathrm{LR}}) . \quad (12)$$

In this way, we obtain an HDR image with low motion blur and noise.

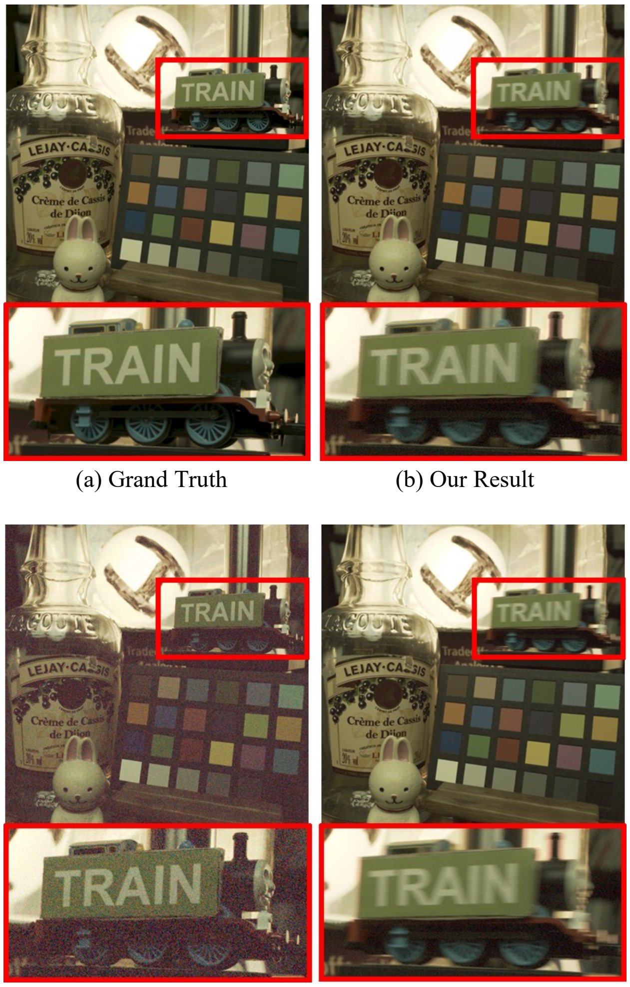
3. Simulation results

In this section, we present simulation results to demonstrate the effectiveness of our method using raw data captured by a Canon EOS 50D camera. In this simulation, we set the high, middle, and low frame rates to 270, 90, and 30 fps, respectively. We acquired images with exposure times as 1/270, 1/90, and 1/30 of a second. We gradually moved a stationary object in each frame. Furthermore, we appended motion blur to the high, middle and low frame rate images by calculating the average of 5, 15, and 45 frames, respectively. Finally, we appended random noise to blurred images using the noise model [5] that considers the read out noise, photon shot noise, and camera gain. The standard deviation of the read out noise, denoted by σ, is 3.

Fig. 2 and Table 1 show the results compared with the optical SVE method [2]. Fig. 2 (c) and (d) show the trade-off relationship between S/N and motion blur, as the exposure time is equal in all pixels. On the other hand, our result presented in Fig. 2 (b) shows reduction of motion blur and noise.

4. Conclusion

This paper proposed a novel imaging and processing method using a single sensor with multiple temporal resolution imaging in order to reduce motion blur and noise in HDR images. We reconstructed clear HDR images using low motion blur information of high frame rate images and high S/N information of low frame rate images. Through a simulation, we demonstrated the effectiveness of the proposed method.



(a) Grand Truth (b) Our Result

(c) Result of [2] (i) (d) Result of [2] (ii)

Fig. 2: Results compared with the optical SVE method [2] (gamma 2.2 enhanced).

Table 1: Results of PSNR compared with the optical SVE method [2].

	PSNR [dB]	
	Whole Region	Red Frame Region
Our Result	**35.00**	**30.49**
Result of [2] (i)	22.41	22.44
Result of [2] (ii)	32.20	26.25

References

[1] Takao Jinno, Masahiro Okuda, "Multiple Exposure Fusion for High Dynamic Range Image Acquisition," IEEE Transactions on Image Processing, Vol21, No.1, pp.358-365 (2012)

[2] Shree K. Nayar, Tomoo Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," CVPR, Vol.1, pp.472-479 (2000)

[3] Takuro Kosaka, Takashi Kobayashi, Takayuki Hamamoto, "Light-adaptive imaging method using interleave-exposure-control image sensor," ITE Tech. Rep., Vol.40, No.15, IST2016-26, pp.21-24 (2016)

[4] Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," IEEE transactions on pattern analysis and machine intelligence, Vol.35, No.6, pp.1397-1409 (2013)

[5] C. Aguerrebere, A. Almansa, Y. Gousseau, J. Delon, P. Muse, "Single shot high dynamic range imaging using piecewise linear estimators," ICCP, pp.1-10 (2014)