Paper ID: 1512

Ghost-free HDR Imaging via Unrolling Low-Rank Matrix Completion

Truong Thanh Nhat Mai†, Edmund Y. Lam‡, and Chul Lee†

†Department of Multimedia Engineering, Dongguk University, Seoul, Korea

‡Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

E-mails: mtntruong@mme.dongguk.edu, elam@eee.hku.hk, chullee@dongguk.edu

High Dynamic Range Imaging

- Multi-exposure fusion algorithms merge multiple input low dynamic range (LDR) images to produce a single high dynamic range (HDR) image. However, they suffer from ghosting artifacts caused by camera and object motions.

Figure 1:Ghosting artifacts.

Low-rank Model

- We construct the observed irradiance matrix

$$\mathbf{D} = [vec(\mathbf{H}_1), \ldots, vec(\mathbf{H}_n)] \quad (1)$$

where $\mathbf{H}_i$ is the irradiance computed from the i-th warped LDR image.
- The irradiance matrix $\mathbf{D}$ can be decomposed into two matrices,
  - $\mathbf{X}$: static background which has a rank-1 structure.
  - $\mathbf{E}$: moving objects in the foreground which resembles a sparse matrix
- Exploiting these properties, we employ the truncated nuclear norm [1] to formulate the fusion task as a rank minimization problem as

$$\begin{aligned}&\underset{\mathbf{X},\mathbf{E},\mathbf{S}}{\text{minimize}} \quad \|\mathbf{X}\|_{r=1} + \lambda\|\mathbf{E}\|_1 + f(\mathbf{X}) + g(\mathbf{E})\\ &\text{subject to } \mathbf{X}+\mathbf{E}+\mathbf{S} = \mathcal{P}_\Omega(\mathbf{D}), \|\mathcal{P}_\Omega(\mathbf{S})\|_F \le \delta,\end{aligned} \quad (2)$$

where λ is the trade-off parameter, $f(\cdot)$ and $g(\cdot)$ are adaptive priors.

Solutions

- Variable separation for (2)

$$\begin{aligned}&\underset{\mathbf{X},\mathbf{E},\mathbf{S},\mathbf{P},\mathbf{Q}}{\text{minimize}} \quad \|\mathbf{X}\|_{r=1} + \lambda\|\mathbf{E}\|_1 + f(\mathbf{P}) + g(\mathbf{Q})\\ &\text{subject to } \mathbf{P} = \mathbf{X}, \mathbf{Q} = \mathbf{E},\\ &\qquad \mathbf{X}+\mathbf{E}+\mathbf{S} = \mathcal{P}_\Omega(\mathbf{D}), \|\mathcal{P}_\Omega(\mathbf{S})\|_F \le \delta.\end{aligned} \quad (3)$$

- Closed-form solutions by an iterative technique

$$\begin{aligned}\mathbf{X}_{k+1} &= \arg\min_{\mathbf{X}} \mathcal{L}(\mathbf{X}, \mathbf{E}_k, \mathbf{S}_k, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k)\\ \mathbf{E}_{k+1} &= \arg\min_{\mathbf{E}} \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}, \mathbf{S}_k, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k)\\ \mathbf{S}_{k+1} &= \arg\min_{\|\mathcal{P}_\Omega(\mathbf{S})\|_F \le \delta} \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k)\end{aligned} \quad (4)$$

where $\mathcal{L}$ is the the augmented Lagrangian function derived from (3).
- We learn prior information from the data via CNNs

$$\begin{aligned}\mathbf{P}_{k+1} &= \arg\min_{\mathbf{P}} \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}_{k+1}, \mathbf{P}, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k)\\ &= \text{P-Net}_k\left(\mathbf{X}_{k+1} + \alpha_k^{-1}\mathbf{\Gamma}_k\right),\\ \mathbf{Q}_{k+1} &= \arg\min_{\mathbf{Q}} \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}_{k+1}, \mathbf{P}_{k+1}, \mathbf{Q}, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k)\\ &= \text{Q-Net}_k\left(\mathbf{E}_{k+1} + \beta_k^{-1}\mathbf{\Phi}_k\right)\end{aligned} \quad (5)$$

- The HDR image is computed by weighted average of columns in $\widehat{\mathbf{X}}$

$$\mathbf{R} = \frac{1}{n}\sum_{i=1}^{n} \omega_i \widehat{\mathbf{X}}_i$$

The Network

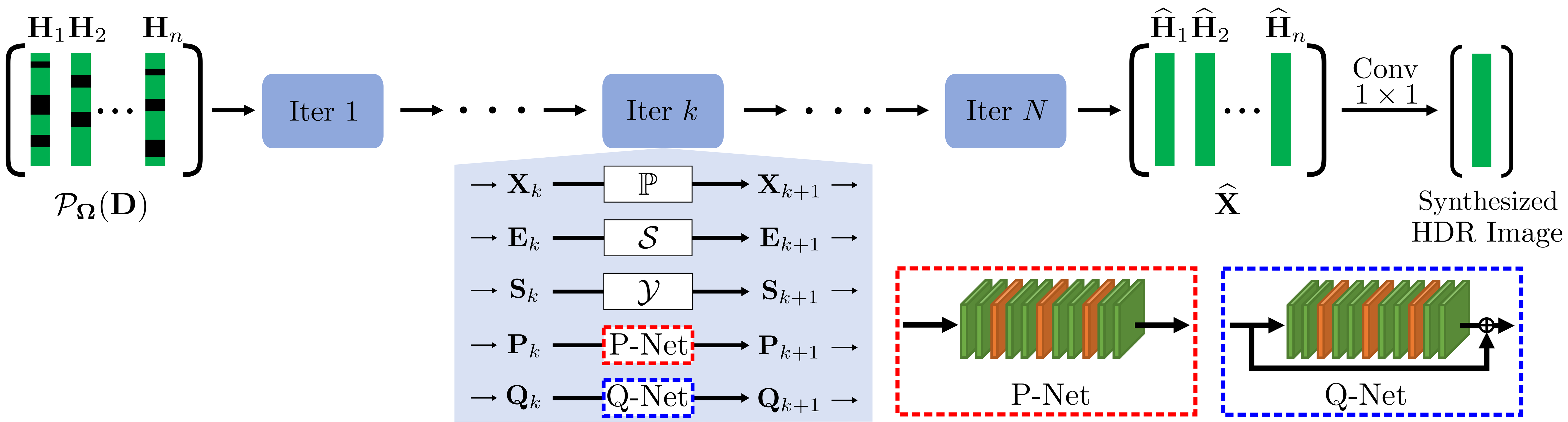

Figure 2:The network takes $\mathbf{D}$ in (1) as input. For initialization, $\mathbf{X}_1 = \mathcal{P}_\Omega(\mathbf{D})$ and others are zero matrices. Operations in each iteration correspond to solutions in (4) and (5). The green and orange blocks in P-Net and Q-Net indicate the convolutional layers and ReLU layers, respectively. The network process image channels separately and the weights are shared.

Experimental Results

- Conventional algorithms produce unfaithful textures since end-to-end learning algorithms infers them from learned features, which may be different from those in the ground-truth.
- In contrast, the proposed algorithm faithfully synthesizes HDR images without texture losses and color artifacts by strictly constraining low-rank priors with learned regularizers.

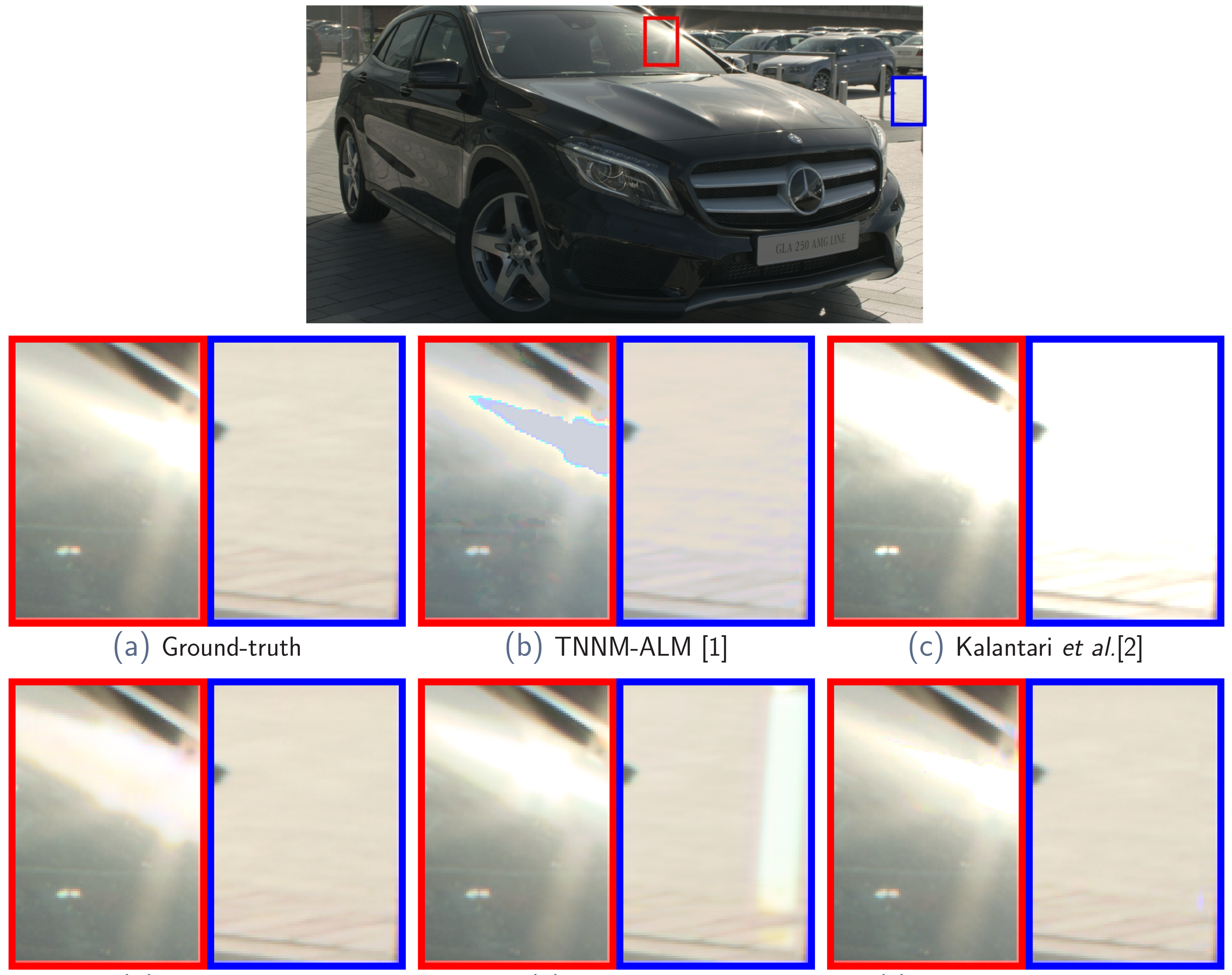

Figure 3:Comparison of the HDR synthesis results.

References

[1] C. Lee and E. Y. Lam, "Computationally efficient truncated nuclear norm minimization for high dynamic range imaging," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4145-4157, Sep. 2016.

[2] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 144:1-144:12, Jul. 2017.

[3] S. Wu, J. Xu, Y.-W. Tai, and C.-K. Tang, "Deep high dynamic range imaging with large foreground motions," in *Proc. Euro- pean Conf. Comput. Vis.*, Sep. 2018, pp. 120-135.

[4] Q. Yan, D. Gong, Q. Shi, A. van den Hengel, C. Shen, I. Reid, and Y. Zhang, "Attention-guided network for ghost-free high dynamic range imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 1751-1760.