Supplementary Material of Paper ID: 0616

Burst Reflection Removal using Reflection Motion Aggregation Cues

Contents

- Illustration of Reflection Motion Aggregation
- More Comparisons on Our Burst Image Reflection Removal (BIRR) dataset
- Dataset Samples
- Analysis on Failure Case

Section 1: Reflection Motion Aggregation - Illustration

(a) Burst Input | (b) RMA Cue | (c) Final Output | (d) Multi-Image Input (Large Baseline) | (e) RMA Cue | (f) Final Output

The RMA cue emphasizes the transmission layer more than the reflection layer as it is evident in (b) and (e) above.

Section 2: More Extensive comparisons on Our Burst Image Reflection Removal (BIRR) Dataset

We compare our method against the following state of the art methods.
(Short form names for these papers are mentioned in the brackets)

Single Image Based Methods (We use the reference frame (middle frame of burst) for evaluation)

[3] Dong, Z., Xu, K., Yang, Y., Bao, H., Xu, W., Lau, R.W.: Location-aware single image reflection removal. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5017-5026 (2021) **(Loc. Aware)**

[14] Chao Li, Yixiao Yang, Kun He, Stephen Lin, and John E Hopcroft. Single image reflection removal through cascaded refinement. In proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 35653574, 2020. 1, 2, 5, 6 **(IBCLN)**

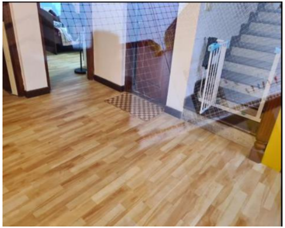
[30] Kaixuan Wei, Jiaolong Yang, Ying Fu, David Wipf, and Hua Huang. Single image reflection removal exploiting misaligned training data and network enhancements. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 81788187, 2019. 1, 2, 3, 5, 6 **(ERRNet)**

Multi Image Based Methods (We use 5 frames for evaluation)

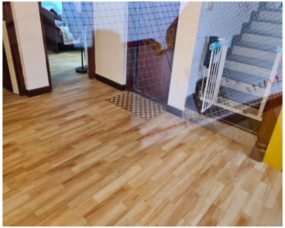
[1] Alayrac, J.B., Carreira, J., Zisserman, A.: The visual centrifuge: Model-free layered video representations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2457-2466 (2019) **(Alayrac)**

[18] Liu, Y.L., Lai, W.S., Yang, M.H., Chuang, Y.Y., Huang, J.B.: Learning to see through obstructions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14215-14224 (2020) **(LTSO)**

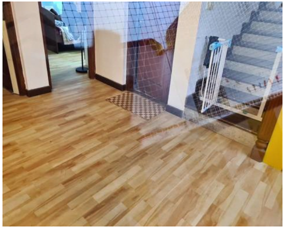
Set 1



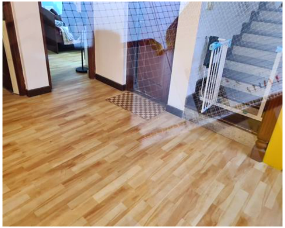
(a) Burst Input Reference



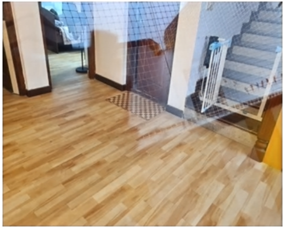
(b) ERRNet



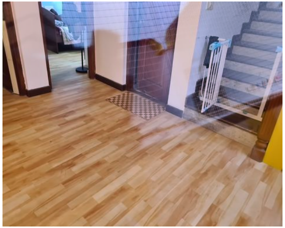
(c) IBCLN



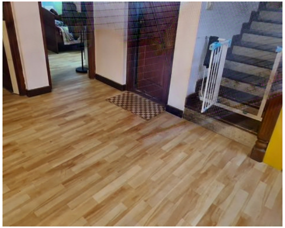
(d) Loc.Aware



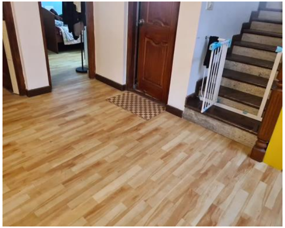
(e) Alayrac



(f) LTSO



(g) Ours



(h) Ground Truth

Set 2




(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware




(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Set 3

(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Set 4

(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Set 5

(a) Burst Input Reference (b) ERRNet (c) IBCLN (d) Loc.Aware

(e) Alayrac (f) LTSO (g) Ours (h) Ground Truth

Set 6

(a) Burst Input Reference

(b) ERRNet

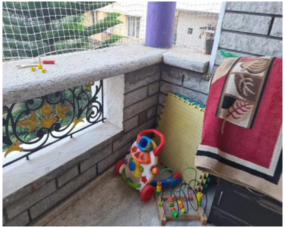
(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO

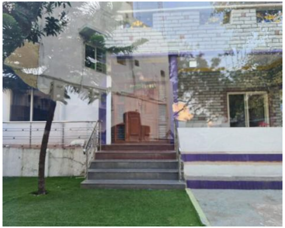
(g) Ours



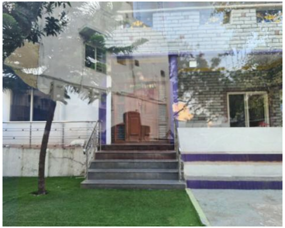
(h) Ground Truth

Set 7

(a) Burst Input Reference



(b) ERRNet

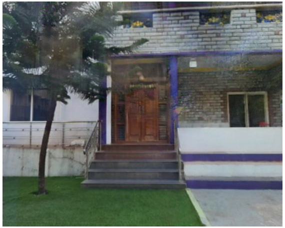


(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO



(g) Ours

(h) Ground Truth

Set 8

(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Set 9

(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware

(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Set 10

(a) Burst Input Reference

(b) ERRNet

(c) IBCLN

(d) Loc.Aware

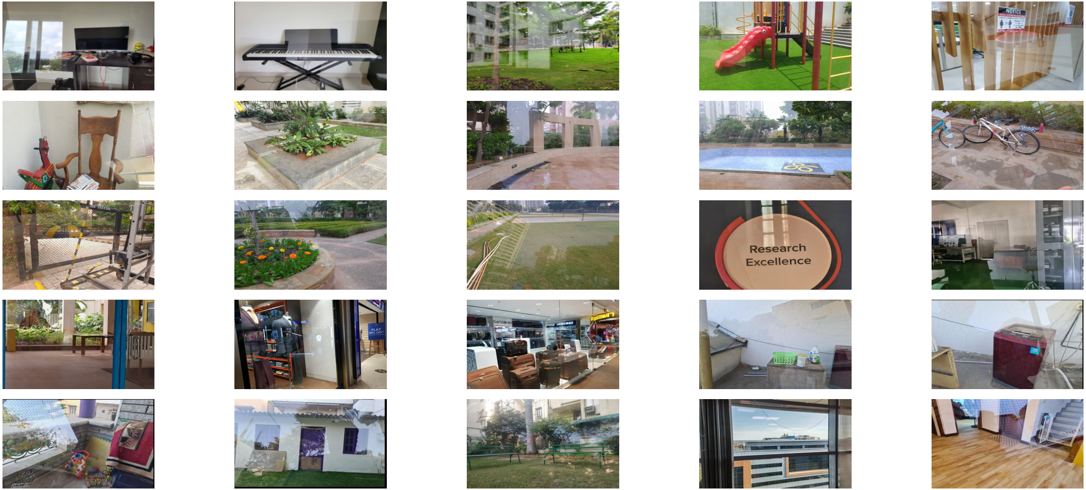
(e) Alayrac

(f) LTSO

(g) Ours

(h) Ground Truth

Dataset Samples*



*showing one frame of the burst in each sample

Analysis of Failure case

Discussion:

The proposed method requires an initial reflection suppression to be able to align the burst of images using homography. However in the presence of reflections that occupy a majority of the image pixels, the homography estimation fails, leading to inferior reflection removal. Figure (b-c) demonstrate this failure scenario where our method is unable to generate a reflection free output in the presence of reflections that occlude the majority of transmission layer. We also show in Figure (d-e), the result of our method with an improved homography alignment by manually rejecting transmission pixels for homography estimation. This indicates that a further improvement of transmission alignment can help overcome this particular failure case. One possibility is to perform reflection detection to identify the reflection pixels and exclude them in homography estimation.

(a) (b) (c) (d) (e) (f)

Figure: Analysis of Failure case. (a) Input image (b) RMA using failed homography (c) Output for failed RMA (d) RMA using corrected homography (e) Output for corrected RMA (f) Ground Truth

Thank You