SenseBrain Technology

Massachusetts Institute of Technology

Overexposure Mask Fusion: Generalizable Reverse ISP Multi-Step Refinement

Jinha Kim, Jun Jiang, and Jinwei Gu
Massachusetts Institute of Technology, SenseBrain Technology

ECCV TEL AVIV 2022

Background

Image signal processor (ISP) is a collection of operations integrated in today's digital cameras that maps camera sensor readings into visually pleasing RGB images. Unlike RGB images, RAW data holds a linear relationship with scene irradiance, which has led to improved performance in various computer vision tasks notably for image denoising

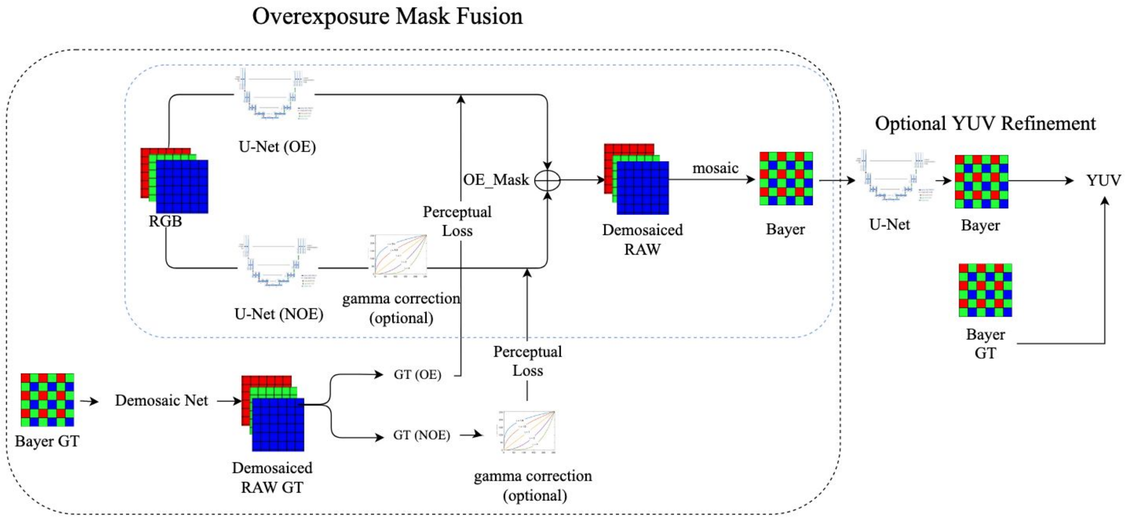
Limitations to the task of RAW reconstruction:

- Degradation in performance with overexposure in large regions in input RGB images
- Less availability of training methodologies for mapping RGB to RAW compared to RGB to RGB (e.g. perceptual losses)

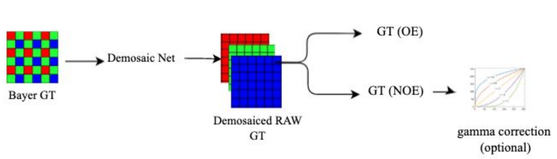
Contribution

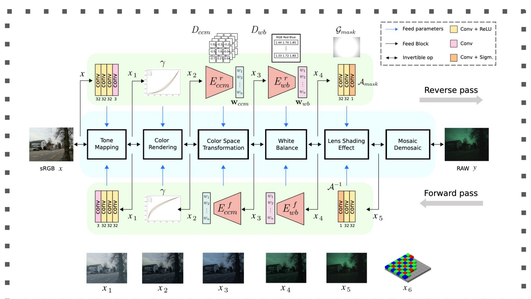
- Multi-step pipeline that redefines the task to the easier mapping from RGB to RGB, allowing use of perceptual loss by mapping from RGB images to demosaiced RAW then mosaicing to Bayer
- Overexposure mask fusion to address the major limitation of overexposure by mapping overexposed and non-overexposed pixels separately and fusing the pixels using the overexposure mask
- Significant improvement from the baseline U-Net performance in fidelity measures and perceptual quality of reconstructed RAW through augmentations in pipeline
- Pipeline capable of incorporating and enhancing other solutions that support end-to-end learning and modifiable to map from RGB to demosaiced RAW.
- Generalizable, multi-step refinement process for enhanced performance for reversed ISPs by allowing use of perceptual losses

Method

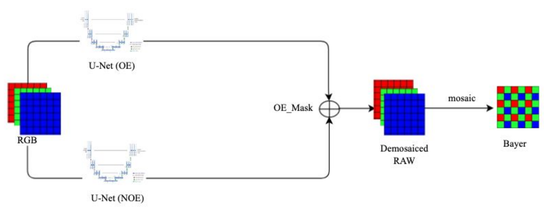


Schematic representation of the overall pipeline. Neural networks were all unified to be U-Nets for objective comparisons. U-Net (overexposure or OE) and U-Net (non-overexposure or NOE) both take in input RGB images and outputs demosaiced RAW. For inference, the overexposed and non-overexposed pixels are blended using the overexposure mask.





- Generates two separate groundtruth images after passing the bayer groundtruth into a pre-trained Demosaic Net
- GT (OE) and GT (NOE) can be used to separately compute loss for overexposed and non-overexposed pixels
- Capable of integrating other methodologies instead of U-Net by making small modifications



Result

Team name	Track 1 (Samsung S7) Test1 PSNR ↑	Test1 SSIM ↑	Test2 PSNR ↑	Test2 SSIM ↑	Track 2 (Huawei P20) Test1 PSNR ↑	Test1 SSIM ↑	Test2 PSNR ↑	Test2 SSIM ↑
NOAHTCV	31.86	0.83	32.69	0.88	38.38	0.93	35.77	0.92
MiAlgo	31.39	0.82	30.73	0.80	40.06	0.93	37.09	0.92
CASIA LCVG (*)	30.19	0.81	31.47	0.86	37.58	0.93	33.99	0.92
HIT-IIL	29.12	0.80	30.22	0.87	36.53	0.91	34.07	0.90
SenseBrains (Ours)	28.36	0.80	30.08	0.86	35.47	0.92	32.63	0.91
CS˜2U (*)	29.13	0.79	29.95	0.84	-	-	-	-
HiImage	27.96	0.79	-	-	34.40	0.94	32.13	0.90
0noise	27.67	0.79	29.81	0.87	33.68	0.90	31.83	0.89
OzU VGL	27.89	0.79	28.83	0.83	32.72	0.87	30.69	0.86
PixelJump	28.15	0.80	-	-	-	-	-	-
CVIP	27.85	0.80	29.50	0.86	-	-	-	-
CycleISP [13]	26.75	0.78	-	-	32.70	0.85	-	-
UPI [2]	26.90	0.78	-	-	-	-	-	-
U-Net Base	26.30	0.77	-	-	30.01	0.80	-	-

(a) Input Full Resolution Image (b) Visualized Output