NEURAL AUGMENTED EXPOSURE INTERPOLATION FOR HDR IMAGING

Zhengguo Li[1], Chaobing Zheng[2], Jinghong Zheng[1], and Shiqian Wu[2]

[1] Institute for Infocomm Research, A*STAR, Singapore
[2] Wuhan University of Science and Technology, Wuhan, China

ABSTRACT

Brightness order reversal usually appears when two large-exposure-ratio images of a high dynamic range scene are directly fused together by an existing multi-scale exposure fusion algorithm. To address the problem, a novel neural augmented framework is introduced to interpolate an image with the medium exposure by integrating physics-driven and data-driven approaches. The physics-driven method infers high-frequency information while the data-driven approach learns remaining information for the interpolated image. The interpolated image and two large-exposure-ratio images are fused together. Experimental results show that the proposed framework can indeed solve the brightness order reversal problem for the fusion of of two large-exposure-ratio images.

1. INTRODUCTION

Merging differently exposed low dynamic range (LDR) images is an effective way to expand the dynamic range for high dynamic range (HDR) imaging [1]. New HDR image capturing devices are emerging to capture them simultaneously so as to avoid ghosting artifacts caused by moving objects in an HDR scene [2, 3, 4, 5]. Two large-exposure-ratio images are usually captured to reduce the cost of the new devices [4, 5, 6], and to capture information as much as possible from the HDR scene [6, 7]. Existing multi-scale exposure fusion (MEF) algorithms such as [8, 9, 10] could yield brightness order reversal in the fused image which looks unnatural if these two large-exposure-ratio images are directly fused by them. Exposure interpolation is an efficient solution to avoid the brightness order reversal for the fusion of two large-exposure-ratio LDR images. A simple physics-driven exposure interpolation algorithm was proposed in [7]. However, the limited representation capability of the physics-driven method results in an interpolated image with low quality which will affect the quality of finally fused image. An interesting neural augmentation based exposure interpolation framework was proposed in [13] in the case that camera response functions (CRFs) are available. Unfortunately, this assumption is not always true. Therefore, it is desired to develop a new exposure interpolation framework without requiring the knowledge of the CRFs for the HDR imaging on top of two large-exposure-ratio images captured simultaneously.

In this paper, a novel exposure interpolation framework is introduced for two large-exposure-ratio images Z_1 and Z_3 by fusing physics-driven and data-driven methods. The proposed physics-driven exposure interpolation is inspired by the color transfer algorithm in [19] and the instance normalization in [20] which are applicable to any pair of images. Local structures of the two images can be utilized to improve the algorithms in [19, 20] if there are neither camera movement nor moving objects in the two image. Based on this observation, a new mapping algorithm is first proposed for the two large-exposure-ratio images Z_1 and Z_3 without camera movements and moving objects. Two high-frequency images are then constructed from the pair of images (Z_1, Z_3) by using the new mapping algorithm, and finally combined together to represent the high-frequency information of the interpolated image via a content adaptive weighted averaging approach. The high-frequency information is represented well by the proposed physics-driven method. The remaining information of the interpolated image is learnt from the two large-exposure-ratio images Z_1 and Z_3 by using a data-driven exposure interpolation neural network (EI-Net). Two key components of the EI-Net are triple attention block (TAB) and short-skip-connection based recursive residual group (SRRG). The SRRG is on top of the recursive residual group (RRG) in [24], and it can retain shallow layers information and pass it into deep layers. Therefore, the SRRG can help the EI-Net learn the remaining high-frequency information better. The TAB can not only fuse all features but also adaptively learn the different weights of different level feature information so as to extract features with better representation abilities. To validate the necessity of exposure interpolation, the interpolated image and two large-exposure-ratio images are fused together via the MEF algorithm in [9] with a physics-optimization based high-frequency information enhancement component. The resultant algorithm outperforms nine state-of-the-art (SOTA) algorithms in [6, 8, 9, 10, 7, 11, 21, 22, 12]. Therefore, the exposure interpolation is demanded by the fusion of two large-exposure-ratio images.

2. THE PROPOSED EXPOSURE INTERPOLATION

2.1. An Exposure Interpolation Dataset

The exposure interpolation dataset is constructed from two existing datasets on HDR imaging in [4, 26]. All the differently LDR images are captured by two different cameras which are VETHDR-Nikon and VETHDR-Canon. Each set inclues three differently exposed LDR images, Z_1, Z_2, and Z_3. There are neither camera movements nor moving objects among the three images. Their exposure times are Δt_1, Δt_2 and Δt_3, respectively. Without loss of generality, it is assumed that

$$\Delta t_1 \ll \Delta t_3 \; ; \; \Delta t_2^2 = \Delta t_1 \Delta t_3. \tag{1}$$

The inputs to the proposed algorithm are the two large-exposure-ratio images such as those in Figs. 1(a) and 1(c). The images in Fig. 1(b) are the ground-truth ones of the interpolated images. The resolution of the differently LDR images is 720×1280. The interval between the two adjacently exposed images is 3-EV. Thus, the interval of two large-exposure-ratio inputs is 6-EV. *The dataset is randomly split into three parts: 650 sequences for training, 50 ones for verifying, and the rest 100 ones for testing.*

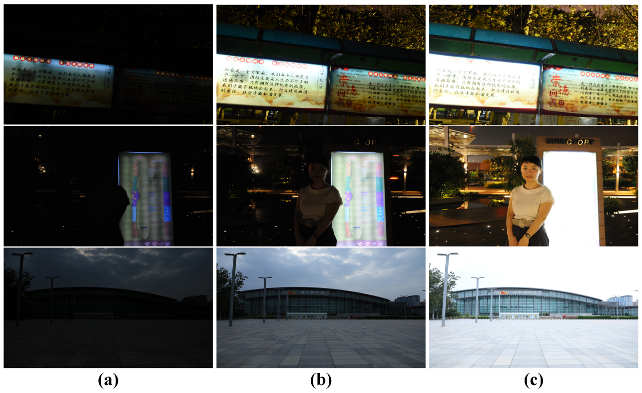

Fig. 1: *Three sets of (a) small-exposure, (b) medium-exposure, and (c) large-exposure images. There are neither camera movements nor moving objects in each set of images.*

2.2. The Proposed Framework

Given a pair of two large-exposure-ratio LDR images Z_1 and Z_3 with N pixels, a new neural augmented framework will be proposed in this section to to represent Z_2 as

$$f(Z_1, Z_3) = f_h(Z_1, Z_3) + f_r(Z_1, Z_3), \quad (2)$$

where $f_h(Z_1, Z_3)$ and $f_r(Z_1, Z_3)$ are high-frequency information and remaining information, respectively. The high-frequency information $f_h(Z_1, Z_3)$ is first estimated by using a physics-driven method. The remaining information $f_r(Z_1, Z_3)$ is then learnt by a data-driven deep neural network (DNN). Figs. 2 and 3 summarize the proposed neural augmented exposure interpolation framework.

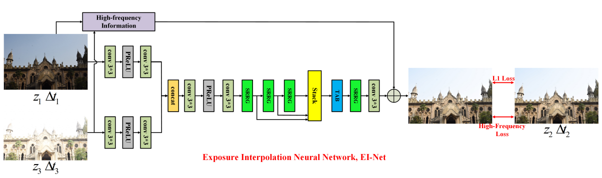

Fig. 2: *The proposed neural augmentation based exposure interpolation framework for the two large-exposure-ratio images Z_1 and Z_3.*

2.3. Physics-Driven Representation of $f_h(Z_1, Z_3)$

The proposed physics-driven method is inspired by the following normalization algorithm [19, 20]:

$$\tilde{Z}^{i\to 2}(p) = \frac{\sigma_{Z_2}}{\sigma_{Z_i}}(Z_i(p) - \mu_{Z_i}) + \mu_{Z_2}\ ;\ i \in \{1, 3\}, \quad (3)$$

where μ_{Z_i} and $\sigma_{Z_i} (i = 1, 2, 3)$ are the mean and variance of all the pixels in the images Z_i, respectively.

Let $\Omega_\rho(p)$ be a square window centered at the pixel p of a radius ρ. $\mu_{Z_i,\rho}(p)$ is the mean value of the image Z_i in the window $\Omega_\rho(p)$. Since there are neither camera movements nor moving objects in the images Z_1, Z_2 and Z_3, the mapping algorithm (3) can be replaced by

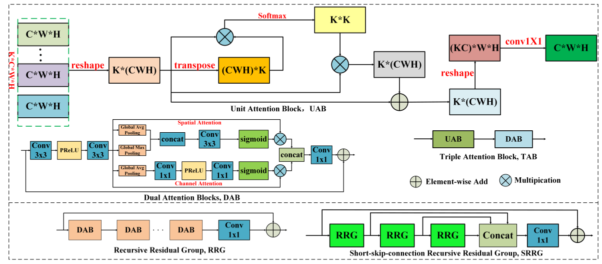

Fig. 3: *Illustration of the SRRG, DAB, UAB, and TAB. he short-skip-connection structure can help the proposed EI-Net preserve the remaining high-frequency information. The TAB can make the network pay more attention to information that is more important for the restoration of the remaining information $f_r(Z_1, Z_3)$.*

$$\begin{aligned}\tilde{Z}^{i\to 2}(p) =& \frac{\sigma_{Z_2}}{\sigma_{Z_i}}(Z_i(p) - \mu_{Z_i,1}(p)) + \mu_{Z_2,1}(p) \\ \doteq& \frac{\sigma_{Z_2}}{\sigma_{Z_i}}\varphi_{h,1}(Z_i(p)) + \mu_{Z_2,1}(p).\end{aligned} \quad (4)$$

The proposed physics-driven representation of $f_h(Z_1, Z_3)$ is on top of the above mapping algorithm *which is also useful to address ghost removal [23]*. The high frequency information $\varphi_{h,1}(Z_i(p))$ can be easily computed from the image $Z_i (i = 1, 3)$. However, the image Z_2 is not available, and it is impossible to compute the scaling factor $\sigma_{Z_2}/\sigma_{Z_i}$ directly. Fortunately, the scaling factors $\sigma_{Z_2}/\sigma_{Z_1}$ and $\sigma_{Z_2}/\sigma_{Z_3}$ can be estimated from the images Z_1 and Z_3 as $\sqrt{\sigma_{Z_3}/\sigma_{Z_1}}$ and $\sqrt{\sigma_{Z_1}/\sigma_{Z_3}}$, respectively. Two high-frequency images are then generated from the images Z_1 and Z_3 as

$$\begin{cases} \tilde{f}_{h,1}^{1\to 2}(p) = \sqrt{\frac{\sigma_{Z_3}}{\sigma_{Z_1}}}\varphi_{h,1}(Z_1(p)) \\ \tilde{f}_{h,1}^{3\to 2}(p) = \sqrt{\frac{\sigma_{Z_1}}{\sigma_{Z_3}}}\varphi_{h,1}(Z_3(p)) \end{cases}. \quad (5)$$

The high-frequency information $f_h(Z_1, Z_3)$ is finally obtained via the following weighted averaging formula:

$$f_h(Z_1(p), Z_3(p)) = \frac{\sum_{i\in\{1,3\}} W_i(Z_i(p))\tilde{f}_{h,1}^{i\to 2}(p)}{\sum_{i\in\{1,3\}} W_i(Z_i(p))}, \quad (6)$$

where $W_1(z)$ and $W_3(z)$ are defined as [27]:

$$W_1(z) = \begin{cases} 0; & if\ 0 \le z < 10 \\ 1 - 3h_1^2(z) + 2h_1^3(z); & if\ 10 \le z < 55 \\ 1; & otherwise. \end{cases}, \quad (7)$$

$$W_3(z) = \begin{cases} 1; & if\ 0 \le z < 200 \\ 1 - 3h_3^2(z) + 2h_3^3(z); & if\ 200 \le z < 245 \\ 0; & otherwise. \end{cases}, \quad (8)$$

and $h_1(z)$ and $h_3(z)$ are defined as $\frac{55-z}{45}$ and $\frac{z-200}{45}$, respectively.

2.4. Data-Driven Representation of $f_r(Z_1, Z_3)$

As shown in Figs. 2 and 3, the images Z_1 and Z_3 are passed into shallow feature extraction parts, and the shallow features of

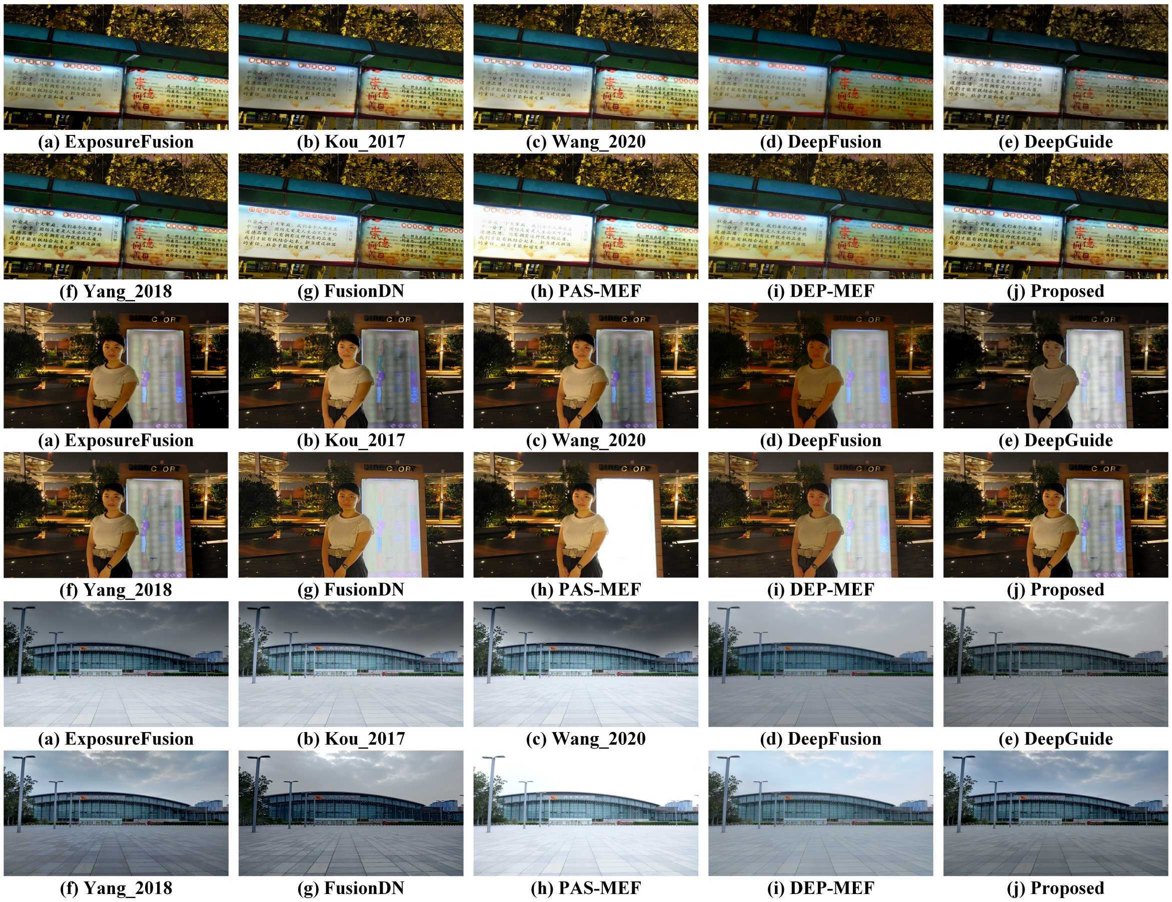

Fig. 4: *Results of ten different algorithms for three pairs of large-exposure-ratio images in Figures 1(a) and 1(c). Fused images by using (a) ExposureFusion [8]; (b) Kou_2017 [9]; (c) WQT_2020 [10]; (d) DeepFusion [6]; (e) DeepGuide [11]; (f) Yang_2018 [7]; (g) FusionDN [21]; (h) PAS-MEF [22] ; (i) DEP-MEF [12]; and (j) the proposed one.*

the images Z_1 and Z_3 are then fused together. Subsequently, the fused features are fed into the SRRGs with multiple short-skip-connections, and the output features of the SRRGs are fused together through the TAB module. Finally, the features will be passed to the reconstruction part and global residual learning structure, thereby restoring the remaining information $f_r(Z_1, Z_3)$ efficiently.

The key components of the EI-Net are SRRG and TAB. The SRRG is on top of the RRG in [24]. Each RRG consists of multiple dual-attention blocks (DABs). With the short-skip-connection structure, the proposed SRRG can retain shallow layers information and pass it into deep layers. As such, the remaining high-frequency information could be well preserved, facilitating the efficient learning of the remaining information $f_r(Z_1, Z_3)$. The TAB is composed of a unit attention block (UAB) and a DAB [24]. The DAB combines the channel attention block and spatial attention block, and it can be applied to suppress those less useful features and only allow the propagation of more informative ones [24]. The UAB can learn the interdependencies of different layers among the SRRGs, and assign different attention weights to features of different layers so as to extract features with better representation abilities.

Loss functions also play an important role in the proposed data-driven approach. The information $f_r(Z_1, Z_3)$ is learnt by minimizing the following loss function:

$$L_d = L_r + L_h, \tag{9}$$

where L_r is defined as $\|Z_2 - (f_h(Z_1, Z_3) + f_r(Z_1, Z_3))\|_1$, and L_h is defined by using the high-frequency information as $\|\varphi_{h,1}(Z_2) - \varphi_{h,1}(f_h(Z_1, Z_3) + f_r(Z_1, Z_3))\|_1$, and the function $\varphi_{h,1}(\cdot)$ is defined in the equation (4).

The final image is generated via fusing the two large-exposure-ratio images and the interpolated image by using the MEF algorithm in [9]. As shown in [14], the DNNs are biased toward low-frequency information. In addition, fine details are further lost by the MEF algorithm [17] as the total number of pyramid layers increases. Fortunately, the physics-optimization based method in [10] inherits a superiority to achieve an *instance adaptation* through exploiting the inner statistics of each instance, and it can be adopted to preserve the real high-frequency information in the fused image for each set of input images.

3. EXPERIMENTAL RESULTS

The proposed algorithm is compared with five conventional MEF algorithms: ExposureFusion [8], Kou_2017 [9], WQT_2020 [10],

Table 1: *MEF-SSIM of Ten Different Algorithms (↑: larger is better)*

set	[8]	[9]	[10]	[6]	[11]	[7]	[21]	[22]	[12]	Ours
1	0.9019	0.9014	0.9204	0.8062	0.6814	0.9178	0.9098	0.9190	0.9098	**0.9219**
2	0.8857	0.8831	0.8503	0.8736	0.7834	0.9618	0.9078	0.9074	0.9247	**0.9738**
3	0.8704	**0.9433**	0.7725	0.8558	0.7340	0.9233	0.8664	0.6543	0.8993	0.9400
4	0.7796	0.7870	0.7881	0.8094	0.6541	**0.8825**	0.7021	0.8083	0.8098	0.8721
5	0.7675	0.7821	0.7249	0.8003	0.7438	0.8466	0.8050	0.7970	0.8117	**0.8536**
6	0.8179	0.8190	0.8240	0.7885	0.7355	0.8549	0.8248	0.8244	0.8173	**0.8663**
7	0.8500	0.8547	0.8714	0.8528	0.7142	0.9107	0.8672	0.8723	**0.9424**	0.9196
8	0.8564	0.8556	0.8116	0.8324	0.6844	0.9300	0.8572	0.8497	0.8698	**0.9520**
9	0.8375	0.8210	0.7407	0.8054	0.7158	0.8982	0.8750	0.8547	0.8703	**0.9158**
10	0.9177	0.9132	0.8975	0.9244	0.8552	0.9472	0.9115	0.9034	0.9233	**0.9615**
11	0.8580	0.8639	0.7990	0.8767	0.7445	0.9296	0.8981	0.8833	0.9113	**0.9453**
12	0.8425	0.8614	0.8847	0.8267	0.7192	0.9468	0.8745	0.8812	0.9164	**0.9681**
13	0.8735	0.9007	0.8418	0.8222	0.7241	0.9182	0.8680	0.8705	0.8879	**0.9306**
14	0.8553	0.8554	0.7855	0.7982	0.7177	0.9450	0.8395	0.6717	0.9021	**0.9584**
15	0.9059	0.9173	0.9391	0.8999	0.7234	0.9220	0.9122	**0.9388**	0.9377	0.9334
avg	0.8547	0.8640	0.8301	0.8382	0.7287	0.9156	0.8612	0.8424	0.8889	**0.9275**
Δ	-0.0728	-0.0635	-0.0974	-0.0893	-0.1988	-0.0119	-0.0663	-0.086	-0.0386	**0**

Yang_2018 [7], and PAS-MEF [22] as well as four deep-learning based exposure fusion algorithms: DeepFusion [6], DeepGuide [11], FusionDN [21], and DEP-MEF [12]. The inputs to all the algorithms are the two large-exposure-ratio images Z_1 and Z_3. Our implementation is on top of a PyTorch framework with 4 NVIDIA GP100 GPUs. Randomly cropping samples $128 * 128$ patch from each input is employed to augment training data. The Network is trained using the proposed loss functions and an Adam optimizer with the batch size as 8. The learning rate is initially set to 10^{-4}, then decreased using a cosine annealing schedule.

Results for three pairs of large-exposure-ratio images in Figs. 1(a) and 1(c) are shown in Figure 4. To evaluate the robustness of the proposed framework, the middle pair is not from these two datasets in [4, 26]. Readers are invited to view to electronic version of figures and zoom in them so as to better appreciate differences among all images. The left display board is brighter than the right one and the tree in the first pair of large-exposure-ratio images. However, it is darker than the right display board and the tree in the fused images by ExposureFusion [8], Kou_2017 [9], and WQT_2020 [10]. Information in the brightest regions is also serious lost by PAS-MEF [22]. Even though the relative brightness order is preserved better by DeepGuide [11], PAS-MEF [22], Yang_2018 [7], DeepFusion [6], and DEP-MEF [12], global contrast and local contrast (for example, those Chinese characters) are not preserved well. There is also visible color distortion by FusionDN [21], DeepGuide [11] and DeepFusion [6]. Similar observations can be observed in the fused images for the second and third pairs of large-exposure-ratio images. For example, the floor looks brighter than the sky in the fused images by the algorithms in [8, 9, 10] for the third pair even though it is actually darker than the sky. The local contrast except for those of the directory board in the second set and the sky in the third set is usually preserved well by the FusionDN [21], however it suffers from seriously brightness order reversal as highlighted by the red box in the fused image for the second pair. It can also be verified by the experimental results that DNNs are indeed biased towards learning low-frequency functions [14, 15]. All these problems are overcome by the proposed algorithm.

Besides the subjective evaluation in Fig. 4, the MEF-SSIM in [25] which is widely utilized in the field of HDR imaging is also adopted to compare all these algorithms. The MEF-SSIM is computed by using the fused image with the reference images as the three captured differently exposed images Z_1, Z_2 and Z_3 [25]. As shown in Table 1, the proposed algorithm also significantly outperforms all the nine state-of-the-art MEF algorithms in terms of the MEF-SSIM. It can be shown that DEP-MEF [12], Yang_2018 [7], and the proposed algorithm outperform other seven algorithms in Fig. 4 from subjective quality point of view and outperform them from the MEF-SSIM point of view in Table 1. This implies that the MEF-SSIM is coincided with the subjective evaluation. It is worth noting that the exposure interpolation based algorithm in [7] and the proposed algorithm outperform the deep-learning based algorithms in [6, 12, 21] from both the subjective quality and the MEF-SSIM points of view. Clearly, the exposure interpolation is highly demanded by the HDR imaging on top of two large-exposure-ratio images captured simultaneously by the emerging HDR imaging devices.

4. CONCLUSION REMARKS

Data-driven approaches have a fundamental limitation on preserving real high-frequency information for each instance while physics-driven methods have a superiority to achieve the *instance adaptation* in the sense that they can infer the high-frequency information for each instance independently. The physics-driven and data-driven approaches can thus be seamless integrated together to form a neural augmentation. As an illustration, a neural augmented exposure interpolation framework is proposed for two large-exposure-ratio images in this paper. Experimental results validate that the physics-driven and data-driven methods indeed compensate well in the proposed neural augmented framework.

Acknowledgement

This work was supported in part by Nature Science Foundation of Hubei Province, China (Grant No. 2022CFB676).

References

[1] Paul E. Debevec and J. Malik. "Recovering high dynamic range radiance maps from photographs", in *Special Interest Group on Graphics and Interactive Techniques (SIGGRAPH)*, pp. 369-378 1997.

[2] Michael D. Tocci, C. Kiser, N. Tocci, and P. Sen, "A versatile HDR video production system", in *Special Interest Group on Graphics and Interactive Techniques (SIGGRAPH)*, pp. 1-9, 2011.

[3] J. Gu, Y. Hitomi, T. Mitsunaga, and S. Nayar, "Coded rolling shutter photography: flexible space-time sampling", in *IEEE International Conference on Computational Photography*, pp. 1-8, 2010.

[4] Y. Xu, Z. Liu, X. Wu, W. Chen, C. Wen, and Z. Li, "Deep joint demosaicing and high dynamic range imaging within a single shot", IEEE Trans. on Circuits and Systems for Video Technology, vol. 32, no. 8, pp. 4255-4270, Aug. 2022.

[5] S. Tanaka, T. Otaka, et al., "Single exposure type wide dynamic range CMOS image sensor with enhanced nir sensitivity", ITE Trans. on Media Technology and Applications, vol. 6, no. 7, pp. 195-201, Jul. 2018.

[6] K. Ram Prabhakar, V. Sai Srikar, and R. Venkatesh Babu, "Deepfuse: a deep unsupervised approach for exposure fusion with extreme exposure image pairs", in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4724-4732, 2017.

[7] Y. Yang, W. Cao, S. Wu, and Z. Li, "Multi-scale fusion of two large-exposure-ratio images", IEEE Signal Processing Letters, vol. 25, no. 12, pp. 1885–1889, Dec. 2018.

[8] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography", Computer Graphics Forum, vol. 28, no. 1, pp. 161-171, Jan. 2009.

[9] F. Kou, Z. Li, C. Wen, and W. Chen, "Multi-scale exposure fusion via gradient domain guided image filtering", in *IEEE International Conference on Multimedia and Expo*, pp. 1105-1110, 2017.

[10] Q. Wang, W. Chen, X. Wu and Z. Li, "Detail-enhanced multi-scale exposure fusion in YUV color space", IEEE Trans. on Circuits and Systems for Video Technology, vol. 30, no. 8, pp. 2418-2429, Aug. 2020.

[11] K. Ma, Z. Duanmu, H. Zhu, Y. Fang, and Z. Wang, "Deep guided learning for fast multi-exposure image fusion", IEEE Trans. on Image Processing, vol. 29, pp. 2808-2819, 2019.

[12] D. Han, L. Li, X. Guo, and J. Ma, "Multi-exposure image fusion via deep perceptual enhancement", Information Fusion, vol. 79, pp. 248-262, 2022.

[13] C. Zheng, W. Jia, S. Wu, and Z. Li, "Neural augmented exposure interpolation for two large-exposure-ratio images", IEEE Trans. on Consumer Electronics, 2023.

[14] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, "On the spectral bias of neural networks", in *International Conference on Machine Learning*, pp. 5301-5310, 2019.

[15] R. Basri, D. Jacobs, Y. Kasten, and S. Kritchman, "The convergence rate of neural networks for learned functions of different frequencies", in *Annual Conference on Neural Information Processing Systems*, 2019.

[16] A. Jacot, F. Gabriel, and C. Hongler, "Neural tangent kernel: convergence and generalization in neural networks", in *Annual Conference on Neural Information Processing Systems*, 2018.

[17] Z. Li, Z. Wei, C. Wen, and J. Zheng, "Detail enhanced multi-scale exposure fusion", IEEE Trans. on Image Processing, vol. 26, no. 3, pp. 1243-1252, Mar. 2017.

[18] S. Wang, O. Wang, R. Zhang, A. Owens, and Alexei A. Efros, "CNN-generated images are surprisingly easy to spot... for now", in *2020 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8692-8701, 2020.

[19] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images", IEEE Computer Graphics and Applications, vol. 21, no. 5, pp. 34–41, May 2001.

[20] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[21] H. Xu, J. Ma, Z. Le, J. Jiang, and X. Guo, "FusionDN: a unified densely connected network for image fusion", in *Thirty-Fourth AAAI Conference on Artifical Intelligence*, 2020.

[22] D. Karakaya, O. Ulucan and M. Turkan, "PAS-MEF: multi-exposure image fusion based on principal component analysis, adaptive well-exposedness and saliency map", in *the International Conference on Acoustics, Speech, and Signal Processing*, pp. 2345-2349, 2022.

[23] Z. Li, J. Zheng, Z. Zhu, and S. Wu, "Selectively detail-enhanced fusion of differently exposed images with moving objects", IEEE Trans. on Image Processing, vol. 23, no. 10, pp. 4372-4382, Oct. 2014.

[24] S. Waqas Zamir, A. Arora, et al. "CycleISP: real image restoration via improved data synthesis", in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2693-2702, 2020.

[25] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multiexposure image fusion", IEEE Trans. on Image Processing, vol. 24, no. 11, pp. 3345-3356, Nov. 2015.

[26] C. Zheng, Z. Li, Y. Yang, and S. Wu, "Single image brightening via multi-scale exposure fusion with hybrid learning", IEEE Trans. on Circuits and Systems for Video Technology, vol. 31, no. 4, pp. 1425-1435, Apr. 2021.

[27] A. Akyuz and E. Rinhard, "Noise reduction in high dynamic range imaging", Journal of Visual Communication and Image Representation, vol. 18, pp. 366-376, 2007.