EXPOSURE INTERPOLATION VIA HYBRID LEARNING

Chaobing Zheng*[1], Zhengguo Li*[2], Yi Yang[1] and Shiqian Wu[1]
[1]Wuhan University of Science and Technology, China
[2]Institute for Infocomm Research, Singapore

Motivation:

Deep learning is widely used in the field of image processing：

- *Low illumination image enhancement*
- *Super resolution*
- *Denoising*
- *Hazing removal, rain removal*

Then the question is coming:

- *Is conventional visual processing still needed?* ⟹ *It's still needed, and it's important !*

Taking exposure interpolation as an example to illustrate the importance of combining conventional and deep learning.

Exposure interpolation via hybrid learning

- Step 1: An ***virtual exposed image*** is generated from two large--ratio images of the same scene by conventional method.
- Step 2: Refinement of Intermediate Image via an lightweight residual learning (***LRL***) convolutional neural network.

Generation of Intermediate Image

Let x_1 and x_2 be two large-exposure-ratio images of the same scene. Their exposure times are $\triangle t_1$ and $\triangle t_2$, respectively. Without loss of generality, $\triangle t_1 > \triangle t_2$. A medium exposure image with exposure time as $\triangle t_3$ is supposed to be generated. $\triangle t_3$ is between $\triangle t_1$ and $\triangle t_2$, and it is defined as $\sqrt{\triangle t_1 \triangle t_2}$.

Assume the Camera Response Function (CRF) be $F()$. Let the intensity mapping functions (IMF) from x_1 to y_0 and from x_2 to y_0 be denoted as $\wedge_{13}()$ and $\wedge_{23}()$, the functions can be expressed as:

$$\wedge_{13}(z) = F(\frac{\Delta t_3}{\Delta t_1} F^{-1}(z))\,;\, \wedge_{23}(z) = F(\frac{\Delta t_3}{\Delta t_2} F^{-1}(z))$$

the intermediate image y_0 is generated:

$$y_0(p) = \frac{W_1(x_1(p)) \wedge_{13} (x_1(p)) + W_2(x_2(p)) \wedge_{23} (x_2(p))}{W_1(x_1(p)) + W_2(x_2(p))}$$

the weights are defined as:

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

$$W_2(z) = \begin{cases} 1; & if\ \ 0 \le z < 200 \\ 1 - 3h_2^2(z) + 2h_2^3(z); & \text{if } 200 \le z < \xi_U \\ 0; & \text{otherwise} \end{cases} ; \quad h_2(z) = \frac{z - 200}{\xi_U - 200}$$

Refinement of Intermediate Image via an LRL

Let the ground truth of the medium exposure image be denoted as y . $\tilde{y}(= y - y_0)$ is unmodeled information by the method.
$\tilde{y}$ is sparser than the original information y, and most values are likely to be zero or small. It can be expected that it is easier to use a neural network to approximate $\tilde{y}$ than y.

Fig.1 (a) the ground truth images y; (b) the intermediate images y_0; (c) unmodeled information (y - y_0). The unmodeled information is usually small, many pixel values are 0's.

The unmodeled information $\tilde{y}$ is learned from two images $\{y, y_0\}$; by minimizing the following loss function:

$$L = Lr + wLc$$

Reconstruction Loss L_r:

$$L_r = \|y - y_0 - \tilde{f}(y_0)\|_2^2$$

color loss L_c :

$$L_c = \sum_p \angle(y(p), y_{0(p)} + \tilde{f}(y_0(p)))$$

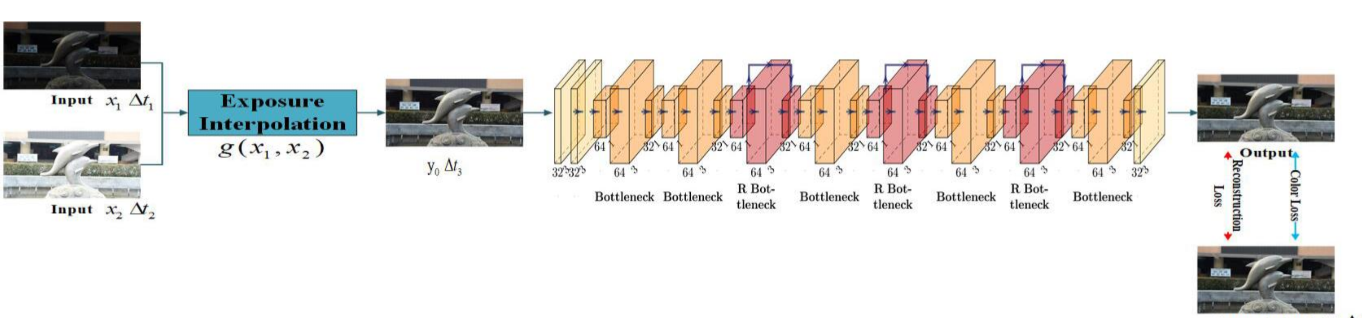


Fig. 2.The proposed hybrid learning framework for exposure interpolation.

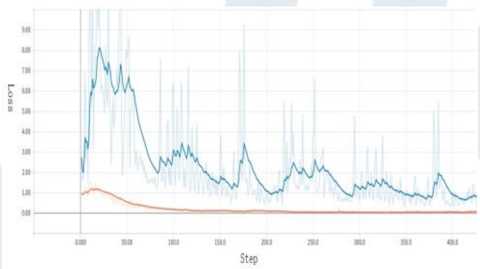
EXPOSURE INTERPOLATION VIA HYBRID LEARNING

Chaobing Zheng*[1], Zhengguo Li*[2], Yi Yang[1] and Shiqian Wu[1]
[1]Wuhan University of Science and Technology, China
[2]Institute for Infocomm Research, Singapore

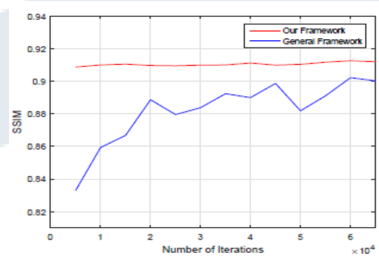
(a)

(b)

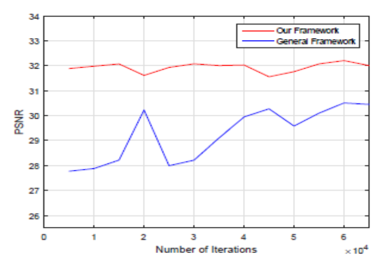
Fig. 3.(a) The residual image (y-y_0); (b) The residual image (y-y_0-$\tilde{f}(y_0)$); The residual image (y-y_0) includes much more visible information than the residual image (y -y_0 - $\tilde{f}(y_0)$).



(a)



(b)



(c)

Fig. 4. (a) Comparison of training, the red is our hybrid learning framework, the blue is existing deep learning method; Comparison of (b) SSIM and (c) PSNR between our hybrid learning and existing deep learning method.

(a) (b) (c)

Fig. 5.(a) The ground truth images; (b) Results of our framework; and (c) Results of existing deep learning method.

The proposed framework is firstly compared with the conventional method in [12] by using both structural similarity index (SSIM) and Peak Signal to Noise Ratio (PSNR).

Table 2. SSIM and PSNR of Two Algorithm

	SSIM	PSNR
Method in [12]	0.8289	22.1127
Proposed	**0.9125**	**32.2080**

The proposed framework is adopted to improve multiscale exposure fusion.

Table 3. MEF-SSIM Of Six Different Algorithms

	[13]	[15]	[14]	[16]	[12]	Ours
Set1	0.9658	0.9486	0.9425	0.9411	0.9629	**0.9822**
Set2	0.9681	0.9736	0.9671	0.9723	0.9782	**0.9812**
Set3	0.9816	0.9537	0.9588	0.9578	0.9673	**0.9860**
Set4	0.9385	0.9099	0.9148	0.9109	0.9449	**0.9866**
Set5	0.9340	0.9353	0.9374	0.9314	0.9436	**0.9506**
Set6	0.9347	0.8926	0.8981	0.9000	0.9410	**0.9728**
Set7	0.9458	0.9328	0.9287	0.9271	0.9548	**0.9810**
Set8	0.9588	0.9447	0.9610	0.9590	0.9706	**0.9879**
Avg	0.9534	0.9364	0.9385	0.9374	0.9579	**0.9785**