Supplementary Material for

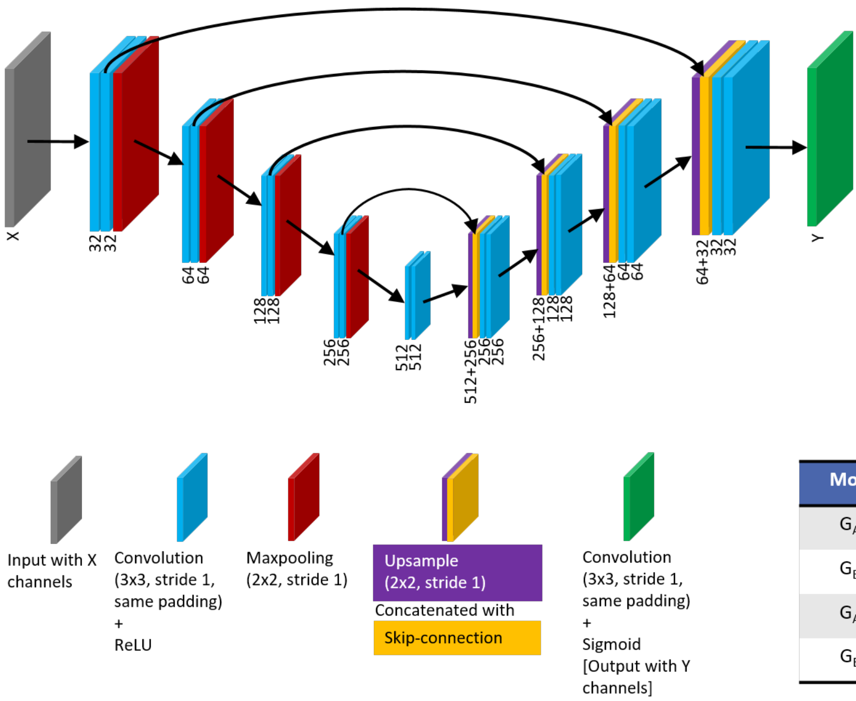
Multi-illumination fusion with crack enhancement using cycle-consistent losses

Milind G. Padalkar* Carlos Beltrán-González* Alessio Del Bue†

* Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia, Genova, Italy

†Visual Geometry and Modelling (VGM), Istituto Italiano di Tecnologia, Genova, Italy

1 Network Architecture



Model	X	Y
G_{A2B}	N	1
G_{B2A}	1	N
G_{A2C}	N	1
G_{B2C}	1	1

Figure 1: Architecture used for training G_{A2B}, G_{B2A}, G_{A2C} and G_{B2C}.

For an input sequence having N images ($N = 65$ in our case), the architecture for models G_{A2B}, G_{B2A}, G_{A2C} and G_{B2C} is shown in Fig. 1, which is inspired from the UNet architecture [1]. The weights are initialized using Glorot Uniform initializer [2].

2 Initial Estimate using Modified Exposure Fusion

The exposure fusion method [3] does not enhance the image in case of pixels being underexposed in most of the images of the input sequence. This is a common problem with multi-exposure fusion (MEF) methods. In order to generate the initial estimate for training the image generators, we modified the exposure fusion method by including a preprocessing step and additional criteria for calculating the weights for fusion. These changes are discussed below.

2.1 Preprocessing

Every image in the input sequence is processed independently. We start the preprocessing step by generating 18 kernels obtained by rotating a 3×3 identity kernel in steps of 10 degrees. These kernels are used to perform morphological closing on the input image to obtain 18 feature maps. These feature maps are combined into a single map by considering the maximum value at pixel across the 18 channels. The resulting feature map is added to the input image and this operation is repeated to generate the output I_{out} as indicated by Eq. 1.

$$\begin{aligned} I_1 =& I_{in} + f(I_{in}), \\ I_2 =& I_{in} + f(I_1), \\ I_{out} =& I_{in} + f(I_2), \\ f(I) =& \max_{i=1}^{18}(I \bullet k_i), \end{aligned} \tag{1}$$

where $\bullet$ denotes the closing operation and I_{in} is the input image.

2.2 Additional Criteria

2.2.1 Change across subsequently acquired images

The images are acquired in an anti-clockwise order of placement of the illumination source. Therefore, a large change in the intensity across subsequently acquired images can indicate presence of interesting details, the weightage for an image for a pixel $I_k(x, y)$ in the image $I_k, i \in 0, 1, ..., N$ of the acquired sequence, the corresponding weight $w_k(x, y)$ for this criteria is calculated as:

$$\begin{aligned} w_k(x, y) =& \sum_{j=max(0,k-1)}^{min(k+1,N)} |I_k(x, y) - I_j(x, y)|, \\ w_k(x, y) =& \frac{w_k(x, y)}{\sum_{j=0}^{N-1} w_j(x, y)}. \end{aligned} \tag{2}$$

2.2.2 Mean based prior

If the mean intensity of the pixels in an image is close to that of a well exposed image then it should get higher weightage. Thus, the weightage used for mean based prior is

$$\begin{aligned} w_k =& exp\left(-\frac{(avg(I_k) - \mu)^2}{2\sigma^2}\right), \\ w_k =& \frac{w_k}{\sum_{j=0}^{N-1} w_j}. \end{aligned} \tag{3}$$

where μ is the mean and σ is the standard deviation of an expected well exposed image. In our implementation we have used $\mu = 0.5$ and $\sigma = 0.2$ with image intensities in the range $[0, 1]$.

2.2.3 Contrast

In addition to the two additional criteria described above, we calculate the contrast weight using magnitude of the Sobel operators' output over a 9 neighbourhood. In the exposure fusion [3] method, this is done using the absolute value of the Laplacian operator's output over a 3×3 neighbourhood.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241. 1

[2] Xavier Glorot and Yoshua Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Yee Whye Teh and Mike Titterington, Eds., Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010, vol. 9 of *Proceedings of Machine Learning Research*, pp. 249–256, JMLR Workshop and Conference Proceedings. 1

[3] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *15th Pacific Conference on Computer Graphics and Applications (PG'07)*, 2007, pp. 382–390. 2, 3