Zero-Shot Image Enhancement with Renovated Laplacian Pyramid (Supplementary Material)

Anonymous ECCV submission

Paper ID 100

1 Formulation of Multiscale Laplacian Enhancement

In the main paper of the section 3.1, we formulate Multiscale Laplacian Enhancement, denoted as MLE. MLE simply, yet efficiently enhances multiscale features of an image in Laplacian pyramid domain, owing to its hierarchical image representation reflecting frequency or resolution. Detailed derivation of MLE is shown here. The part of the derivation of Laplacian pyramid is described in [1].

First, let an input image be $I(x)$, where x is the position of the image. Then, an image of Laplacian pyramid representation is expressed as follows:

$$\begin{aligned} I(x) &= I(x) - l_1\left[I(x)\right] + l_1\left[I(x)\right] \\ &:= L_1(x) + l_1\left[I(x)\right] \\ &= L_1(x) + l_1\left[I(x)\right] - l_2 l_1\left[I(x)\right] + l_2 l_1\left[I(x)\right] \\ &= L_1(x) + L_2(x) + l_2 l_1\left[I(x)\right] \\ \cdots &= \sum_{i=1}^{N} L_i(x) + \prod_{i=1}^{N} l_i\left[I(x)\right] \end{aligned}$$

Here, l_i and $L_i(x)$ mean the i-th level of low-pass filter and Laplacian component, respectively. Each component of Laplacian pyramid $L(x)$ is defined as the subtraction of adjacent output of the low-pass filter, decomposed into the sum of band-pass images [1]. In obtaining Laplacian pyramid representation of an image, we have to prepare various frequencies of low-pass filters $\{l_i\}_{i=1}^{N}$. In practical implementation, low-pass gaussian filter G combined with down-sampling and up-sampling is utilized [3]. Mathematically, Laplacian pyramid representation of an image is expressed as follows:

$$\begin{aligned} I(x) &= \sum_{i=0}^{N-1} \left\{L_i\left[I(x)\right]\right\} + (D \circ G)^N \left[I(x)\right] \\ &:= \sum_{i=0}^{N-1} \left\{(D \circ G)^i \left[I(x)\right] - G \circ (D \circ G)^i \left[I(x)\right]\right\} + (D \circ G)^N \left[I(x)\right] \end{aligned}$$

where $(D \circ G)^0 \left[I(x)\right] := I(x)$, meaning identity mapping. Above G and D mean low-pass Gaussian filter and down-sampling operation by a factor of 2, respectively. Note that each Gaussian filtering is followed by down-sampling operation in calculating the Laplacian components. The last residual element of $(D \circ G)^N [I(x)]$ is the low resolution version of the original image.

In order to reconstruct Laplacian pyramid representation of $I(x)$, recursively we have to up-sample lower level of each components and add to the one level higher Laplacian components. Mathematically, reconstructed image, the same as the input image $I(x)$, is expressed as follows:

$$
\begin{aligned}
I(x) &\equiv L_0 [I(x)] + G [I(x)] \\
&\equiv L_0 [I(x)] + (U \circ D) \circ G [I(x)] \\
&:= L_0 [I(x)] + U [R_1(x)] \\
&= L_0 [I(x)] + U [U [R_2(x)] + L_1 [I(x)]] \\
&= L_0 [I(x)] + U [L_1 [I(x)]] + U^2 [R_2(x)] \\
\cdots &= \sum_{i=0}^{N-1} \left\{ U^i [L_i [I(x)]] \right\} + U^N [R_N]
\end{aligned}
$$

where $R_N = (D \circ G)^N [I(x)]$ means down-sampled lowest resolution of the Laplacian pyramid of the image. U is defined as up-sampling operator, an inverse operator of down-sampling D, and each components are up-sampled to the original dimension of the input. The number of up-sampling is the same as down-sampling in reconstruction.

Proposed Multiscale Laplacian Enhancement (MLE) is described as enhancement in Laplacian pyramid representation, practically formulated as follows:

$$
MLE [I(x)] = \sum_{i=0}^{N-1} \left\{ U^i [L_i [I(x)] * Filter] \right\} + U^N [R_N * Filter]
$$

where pyramid level N, a convolution $Filter$, and up-sampling algorithm should be set beforehand. An input image is first divided into Laplacian pyramid representation and each pyramid elements are simply filtered followed by the reconstruction phase.

Compared to normal convolution, image size to which filters operate is different. To be specific, MLE filters various resolution of an image of each Laplacian pyramid components, while normal convolution operates only in the same size of the input, expressed as follows:

$$
I(x) * Filter = \sum_{i=0}^{N-1} \left\{ U^i [L_i [I(x)]] * Filter \right\} + U^N [R_N] * Filter
$$

The above equation is directly obtained using linearity of convolution and Laplacian pyramid representation of the input. Enlarged effective filter size is practically useful in extracting image features [2], which is naturally formulated in MLE. With the simple idea and implementation of MLE, multiscale features of an input image is efficiently enhanced. In this research, we employ basic unsharp masking filter for MLE to confirm the effectiveness of MLE. Selecting other sophisticated filters is our future work.

2 Ablation Study of MLE (Additional Results)

Qualitative results of different Gaussian kernel for constructing Laplacian pyramid of MLE in the main paper of the section 3.4 are shown in Figure 1. 1st to 3rd columns present results of $\sigma = 1, 3, 9$, and 1st to 3rd rows present results of $K = 3, 5, 9$, respectively. Pyramid level N is set to 5. The lower σ and K are, the smaller the effect of blurriness of filtering, as a result, wide range of frequency band tends to preserve also in higher pyramid level. As a consequent, result of $\sigma = 3, K = 3$ (top left) is most enhanced, while result of $\sigma = 9, K = 9$ (bottom right) is less enhanced. The degree of enhancement basically depends on contrast or sharpness of an original input as well as hyper parameters of MLE.

Fig. 1. Results of MLE with different kernel size ($K = 3, 5, 9$ from top to bottom) and standard deviation ($\sigma = 1, 3, 9$ from left to right). Pyramid level is set to 5.

References

1. Burt, P.J., Adelson, E.H.: The laplacian pyramid as a compact image code. In: Readings in computer vision, pp. 671–679. Elsevier (1987)
2. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV). pp. 801–818 (2018)
3. Paris, S., Hasinoff, S.W., Kautz, J.: Local laplacian filters: Edge-aware image processing with a laplacian pyramid. ACM Trans. Graph. **30**(4), 68 (2011)