Color Image Enhancement Using CNN

C.Andrews Sonia [1], C.S. Sree Thayanandeswari[2]
[1, 2] Dept of ECE
[2]Assistant Professor, Dept of ECE
[1, 2] PET Engineering College,Anna university

Abstract- *In this paper, we propose perceptually optimized enhancement of contrast and color in images using Convolution Neural Network (CNN). The existing image enhancement method is based on JND transform to estimate JND map that represents perceptual response of HVS from an image. The JND transform for contrast enhancement technique is to effectively extract detail information with much attention by HVS.The existing JND transform is perform image enhancement but does not remove noise in images. In the proposed method, the noise removal is based on two-step noise suppression. Due to the poor lighting condition and limited dynamic range of digital imaging devices, the recorded images are often under-/over-exposed and with low contrast. Most of previous single image contrast enhancement (SICE) methods adjust the tone curve to correct the contrast of an input image. Those methods, however, often fail in revealing image details because of the limited information in a single image. On the other hand, the SICE task can be better accomplished if we can learn extra information from appropriately collected training data. In this work, we propose to use the convolutional neural network (CNN) to train a SICE enhancer. Experimental results show that the proposed method achieves good performance in contrast enhancement, color reproduction, and detail enhancement which are measured in terms of LTG, CQE and FSIM .One key issue is how to construct a training dataset of low-contrast and high contrast image pairs for end-to-end CNN learning.*

Keywords- Contrast enhancement, color constancy, image enhancement, just-noticeable difference, JND transform, HVS response model, SICE, convolutional neural network (CNN).

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as twodimensional signals while applying already set signal processing methods to them. It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

The two types of methods used for Image Processing are Analog and Digital Image Processing. Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

II. DEEP SINGLE IMAGE CONTRAST ENHANCER

Reproducing the natural scene with good contrast, vivid color and rich details is an essential goal of digital photography. The acquired images, however, are often under-exposed or over-exposed because of poor lighting conditions and the limited dynamic range of imaging device. The resulting low contrast and low quality images will not only degenerate the performance of many computer vision and image analysis algorithms, but also degrade the visual aesthetics of images.

Contrast enhancement is thus an important step to improve the quality of recorded images and make the image details more visible. Traditional single image contrast enhancement (SICE) techniques include those histogram-based algorithms, which increase the contrast of an image by redistributing the luminous intensity on histogram, and retinex based algorithms which enhance the reflectance and illumination components of the image separately.

These methods, however, are difficult to reproduce a high-quality image due to the complex natural scenes and the limited information in a single low contrast image. MEF and stack-based HDR methods will produce images with better visual quality than those SICE methods since more information is available in the multi-exposure sequence. However, the acquisition of multi exposure images will complicate the imaging process, and camera shake or moving objects will lead to unpleasant fusion artifacts such as the ghosting artifacts.

The output image by a state-of-the-art deghosting MEF algorithm, which merges the multi-exposure images into a high-visibility image. It generates some ghosting artifacts due to the displacement of different frames caused by the object motion (such as human movement and ripples).

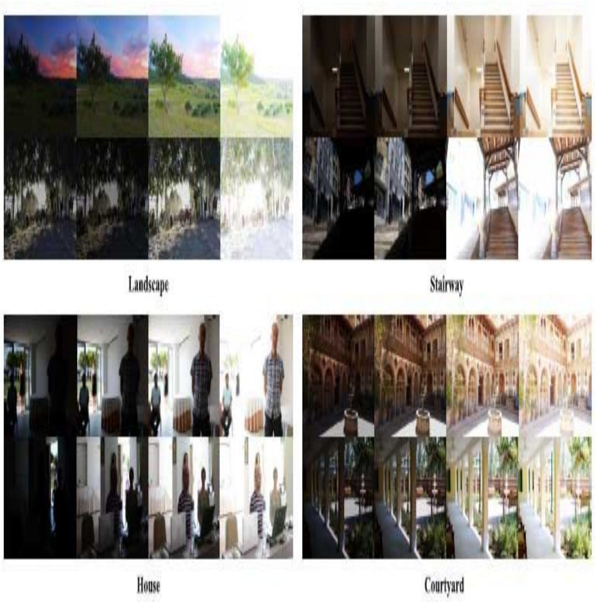


Fig: 1.Sample source image sequences with different exposure levels in our dataset.

In contrast, the SICE method will not have such ghosting artifacts because it takes only one single exposure image as input. Due to the above reasons, SICE is more attractive and easier to implement in practice, yet it is much more challenging because of the limited information in a single image.

Considering that multi-image based MEF and single-image based SICE methods have their pros and cons, one interesting question is: can we develop a SICE method which can approximate the contrast enhancement performance of MEF methods while being free of the ghosting artifacts? In this work, we make the first attempt to address this challenging problem. Our idea is inspired by the success of discriminative learning methods. The deep convolutional neural network (CNN) methods in image restoration. Compared with generative models which use high-quality images to learn image priors, discriminative methods utilize a set of degraded and ground-truth image pairs to learn a model to enhance the given degraded image.

As a powerful discriminative learning method, CNN has been successfully used in many low-level vision problems such as single image super-resolution and image denoising, where a large amount of paired training samples can be generated or simulated.

Down-sample a high-resolution image to generate a corresponding low-resolution image, and noise to a clean image to generate a noisy observation of it. In the application of contrast enhancement, unfortunately, it is very hard to generate such paired images due to the lack of a simple model to approximate the low-contrast image generation process. To the best of our knowledge, by far there is no dataset of paired low-contrast and high contrast images available for training a discriminative model for SICE.

A dataset of low-contrast and good-contrast image pairs, which makes the discriminative learning of SICE enhancers possible. The key idea is that we use state-of-the-art MEF and stack-based HDR methods to reconstruct the reference good-contrast image of a scene, while those under-exposure or over-exposure images of the scene can be naturally taken as the low-contrast counterparts. To build this dataset, we collect multi-exposure sequences from 2 categories of scenes (indoor and outdoor), and employ 13 recently developed MEF and HDR algorithms to generate the high-contrast images for each scene. Then, subjective experiments are conducted to select the best MEF/HDR result for each scene, and exclude those sequences which do not have satisfactory outputs (e.g., ghosting artifacts).

The multi-exposure sequences of 589 scenes and their corresponding high-quality reference images are selected in the dataset. Each sequence has 3 to 18 low-contrast images with different exposure levels, and there are 4,413 low-contrast images in total. With the constructed dataset, we design a simple yet effective CNN to learn a SICE enhancer, which is able to automatically enhance the low-contrast images with different exposure levels. The learned CNN based SICE enhancer demonstrates clear advantages over existing SICE methods, outperforming them by a large margin. The contributions of our work are summarized as follows:

1) The first time to the best of our knowledge, a large-scale multi-exposure image dataset which contains low-contrast images with different exposure levels and their corresponding high-quality reference image. The constructed dataset makes end-to-end discriminative learning of high performance SICE methods possible. It also provides a platform to quantitatively evaluate, at least to some extent, the performance of different contrast enhancement algorithms.

2) With the constructed dataset, a well designed CNN is trained for SICE, which demonstrates significant advantages over existing SICE methods. It provides a new solution to high performance SICE.

III. CONTRAST ENHANCEMENT

Contrast enhancement is a common operation in image processing. It's a useful method for processing scientific images such as X-Ray images or satellite images. And it is also useful to improve detail in photographs that are over or under-exposed.

Concentrating on the image of the tire, it would be preferable for the center of the wheel to stay at about the same brightness while enhancing the contrast in other areas of the image. In order for that to happen, a different transformation would have to be applied to different portions of the image. The Contrast-Limited Adaptive Histogram Equalization technique, implemented in can accomplish this. The algorithm analyses portions of the image and computes the appropriate transformations.

Contrast enhancement of color images is typically done by converting the image to a color space that has image luminosity as one of its components, such as the L*a*b* color space. Contrast adjustment is performed on the luminosity layer 'L*' only, and then the image is converted back to the RGB color space. Manipulating luminosity affects the intensity of the pixels, while preserving the original colors.

1V. CONVOLUTIONAL NEURAL NETWORK

In machine learning, a convolutional neural network (CNN or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to require minimal pre-processing. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared weights architecture and translation invariance characteristics.

Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field.
CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. They have applications in image and video recognition, recommender systems and natural language processing.

Stride Convolution and Deconvolution:

The convolutional operations will reduce the size of feature maps. To ensure that the output image will have the same size as the input one, methods have been proposed to pad zero before convolutions. However, for the luminance enhancement network, we experimentally found that padding zeros would lead to artifacts around the boundary of the output image. Therefore, instead of padding zero, we apply deconvolutions to keep the size of the output unchanged. The convolutional and deconvolutional strategy not only avoid artifacts in the boundary area, but also reduce computational burden with stride filters.

Parametric Rectified Linear Unit

In many CNN based image restoration methods, rectified linear unit (ReLU) is adopted as the activation function. However, since both the positive and negative coefficients contain important local structural information of the input image, simply setting the negative responses to zero may not be a good choice. In this paper, we adopt the PReLU as the activation function, which could improve model fitting with nearly zero extra computational cost and little over-fitting risk. Without ignoring negative coefficients, PReLU is able to generate high quality estimation with less filters.

V. PROPOSED METHOD

The system design of the propose method is shown below and the modules are explained briefly.

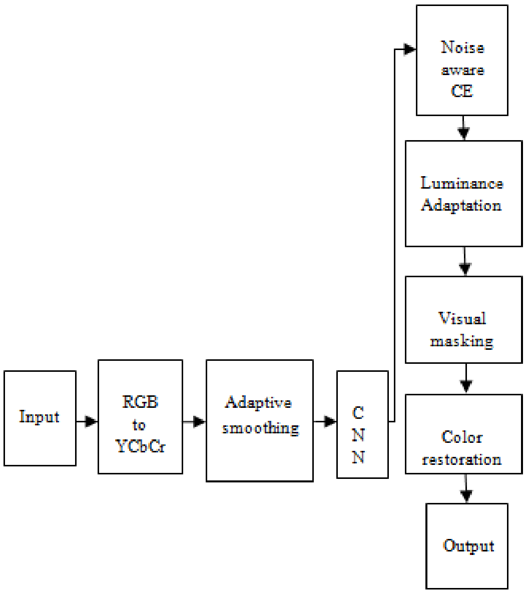


Fig.2. Block Diagram

A. COLOR SPACE CONVERSION:

The input image is RGB image. RGB is a color space. RGB color space uses Red- Green-Blue component of the color and use those component to display the color. The RGB image is converted to YCbCr color space. YCbCr.Cb and Cr is the blue component and red component related to the chroma component. That means "Cb is the blue component relative to the green component. Cr is the red component relative to the green component."These components are less sensitive to the human eyes. Since the Y component is more sensitive to the human eye, it needs to be more correct and Cb and Cr is less sensitive to the human eye. It uses these sensitivities of the human eye and eliminates the unnecessary details of the image.

B. ADAPTIVE SMOOTHING

In next step adaptive smoothing filter is used for YCbCr color space. It attenuates the noise by smoothing while preserving the edges. These algorithms are applied in order to reduce noise and/or to prepare images for further processing. Smoothing and enhancement are important image processing operations. Smoothing operation is necessary to reduce noises and to blur the false/ stray contour fragments in order to enhance the overall visual quality of the degraded image. In order to clean an image and enhance its features, either the spatial or the frequency domain techniques can be used. The frequency domain smoothing, uses filtering in the Fourier domain. Spatial domain techniques, on the other hand, normally employ linear or nonlinear spatial operations. Many efficient techniques have been developed in spatial domain. The simplest smoothing technique uses (unweighted) averaging over a predefined neighborhood. This reduces noise significantly but at the same time it blurs the edges of objects. So the overall image quality deteriorates. With the increase of the neighborhood size blurring becomes more prominent. Some of the weighted averaging techniques which have been proposed to reduce blurring, can be found in . Assigning weights play a significant role in the smoothing operation and hence their determination is an important task. One of the techniques of selecting weights is to use the local mean and variance. Wang et al. have published a good 3urvey on weighted averaging and enhancement techniques. One of the drawbacks of these fixed weighted methods is that they cannot remove noise as efficiently as the unweighted averaging schemes. To make the smoothing schemes more efficient, iterative weighted techniques have been reported, The weighting coefficients are proportional to the inverse gradient between the central point and its neighbors. The convergence of these methods is not known. Nagao and Matsuyama used a simple technique to smooth images, preserving edges. They rotated a mask inside a 5 x 5 window about the center pixel. For every position of the mask, there may exist two regions. They calculated the variances of all such regions due to all possible rotations of the mask and replaced the gray value of the center pixel by the average gray value of the region having the minimum variance. The process is repeated iteratively until all the gray levels in the image remain almost unchanged. Recently, Marc et al have reported an iterative weighted averaging scheme which does both sharpening and smoothing. The method implements anisotropic diffusion, and its iterative behavior has also been discussed. The algorithm considers only the step edges, preservation of roof edges has not been taken into account. Since the weighting coefficients are based on gradients at all 3 x 3 neighborhood points, the averaging result is influenced not only by the neighborhood points but also by some other points beyond it and that deteriorates the image quality. Furthermore, the number of iterations required for different operations e.g., edge detection, is heuristic in nature. Human intervention is needed for termination ofthe algorithm and to judge the image quality. Without such an intervention, the noise cleaning may become insufficient, or there may be excessive useless iterations.

C. CONVOLUTIONAL NEURAL NETWORK

Convolution Neural Network are deep learning algorithms that are particularly powerful for analysis of images. Convolution Neural Network use the data that is represented in images to learn Convolutions are the fundamental building blocks of convolutional Neural Network. With the constructed dataset, we can design a CNN based SICE enhancer to learn a mapping function between the low contrast input image and its corresponding reference image. The proposed CNN has 5 types of layers which are shown in Figure 8 with 5 different colors. i) Conv+PReLU: 64 filters of size 3×3, 5×5 and 9×9 with strides 1 and 2 are used to generate 64 feature maps, and PReLU (parametric rectified linear unit) is utilized for the nonlinearity. ii) Deconv+PReLU: 64 filters of size 9 × 9, 5 × 5 and 3 × 3 with strides 2 and 1 are used to generate 64 feature maps, and PReLU is utilized as the activation function. iii) Conv+BN+PReLU: 64 filters of size 3 × 3 are used, and batch normalization is added between convolution and PReLU. iv) Conv: 3 filters of size 1 × 1 are used to reconstruct the output. v) Skip connection: the add operation is used to connect the feature maps of two layers.

D. LUMINANCE ADAPTATION AND VISUAL MASKING

Second, we perform noise aware contrast enhancement using a noise aware histogram to consider both local contrast and noise distribution. Noise aware contrast enhancement prevents contrast overstretching and noise amplication in large at regions with dark intensity and severe noise. Third, we perform perceptual noise suppression in the detail layer based on the luminance adaptation and visual masking effect. Luminance adaptation measures the noise visibility caused by luminance enhancement, while visual masking measures the noise visibility on smooth and texture regions. The combination of two models achieves noise reduction while preserving details in enhanced results.

E. COLOR RESTORATION

The final step of the proposed method restores the color component by maintaining the ratio between the three color components. The Final output [Rout ,Gout, Bout]T from the original RGB components [R,G,B]T as follows,

$$\begin{pmatrix} \mathrm{Rout} \\ \mathrm{Gout} \\ \mathrm{Bout} \end{pmatrix} = \begin{bmatrix} \frac{YB}{YB} & 0 & 0 \\ 0 & \frac{YB}{YB} & 0 \\ 0 & 0 & \frac{YB}{YB} \end{bmatrix}$$

VI. EXPERIMENTAL REULTS

For quantitative measurements, we evaluate the results in terms of three measures: Feature similarity index (FSIM) , local tuned global (LTG) , and color quality enhancement (CQE)

1. Feature Similarity Index(FSIM):

FSIM reflects the overall similarity between the output image and reference image as follows

$$\mathrm{FSIM}=\frac{\sum_{x\in\Omega} SL(x)PCm(x)}{\sum_{x\in\Omega} PCm(x)}$$

Where Ω means the whole image spatial domain and PC is perceived with a maximizing Fourier component in phase and is used to weight the importance of SL(x)PCm(x) in the overall similarity between the output and reference.

2. Local Tuned Global (LTG):

LTG approaches the process of human visual perception to image quality, which is an effective color image quality assessment (IQA) algorithm as follows:

$$\mathrm{LTE}(x,y) = \frac{\varphi(G_S^{\theta 1})}{\varphi(G_S^{\theta 2})} \cdot \varphi(I_m^{\theta 3} \cdot Q_m^{\theta 3})$$

where Gm the difference of gradient magnitude (GM) maps of the original image x and its contaminated version y, Gs indicates the highest s% values in Gm, Im and Qm to measure the distinction of chrominance between the original and distorted images, $\theta 1$, $\theta 2$, $\theta 3$ are model parameters.

3. Color Quality Enhancement (CQE):

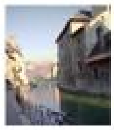
CQE is composed of sharpness, colorfulness and contrast attributes, and has the advantage of being applicable to a wider variety of distorted images. After the colorfulness, sharpness and contrast metrics are obtained, multiple linear regression (MLR) is applied to obtain the three coefficients as follows:

CEQ=C1xcolorfulness+C2Xsharpness +C3Xcontrast

Where c1, c2 and c3 are constants.

The experimental results of the existing and the proposed methods are tabulated below for comparing the performance of both the methods.

Table1. Comparison between both methods

METHOD	INPUT	OUTPUT
EXISTING (USING JND TRANSFORM)		
PROPOSED (USING CNN)		

The performance of the methods are measured in terms of the parameters namely FSIM, LTG, CQE and they are tabulated below

Table2. Performance Measure

METHOD	FSIM	LTG	CQE
EXISTING (USING JND TRANSFORM)	0.8417	0.2363	0.1106
PROPOSED (USING CNN)	0.9653	0.3381	0.8765

VII. CONCLUSION AND FUTURE WORK

Thus the proposed method improving image quality using JND transform and Convolutional Neural Network. The availability of low-contrast images and their high-quality reference images in our dataset allows the end-to end learning of high performance SICE methods. As a demonstration, we developed a simple yet powerful CNN-based SICE enhancer, which is capable of adaptively generating high quality enhancement result for a single over-exposed or underexposed input image. Our experimental results showed that the developed SICE enhancer significantly outperforms state-of the- art SICE methods, and even outperforms MEF and stack based HDR methods for dynamic scenes.

Video enhancement is another important application. To apply the proposed methods to videos, we could consider enlarging our dataset and learning an LSTM (long short term memory) based CNN enhancer to convert the conventional videos to HDR videos.

REFERENCES

[1] KeGu, GuangtaoZhai, Xiaokang Yang, Wenjun Zhang, “An efficient color image quality metric with local-tuned-global model, ” IEEE International Conference on Image Processing., 29 January 2015.

[2] J. Wu, G. Shi, W. Lin, and C. C. J. Kuo, ``Enhanced just noticeabledifference model with visual regularity consideration,'' in Proc.IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2016,pp. 1581_1585.

[3] H. Xu, G. Zhai, X.Wu, and X. Yang, ``Generalized equalization model forimage enhancement,'' IEEE Trans. Multimedia, vol. 16, no. 1, pp. 68_82,Jan. 2014.

[4] X. Wu, ``A linear programming approach for optimal contrast-tonemapping,'' IEEE Trans. Image Process., vol. 20, no. 5, pp. 1262_1272,May 2011.

[5] Z. Farbman, D. Lischinski, and R. Szeliski, ``Edge-preserving decompositionsfor multi-scale tone and detail manipulation,'' Trans. Graph., vol. 27,no. 3, p. 67, Aug. 2008.

[6] Kaiming He, Jian Sun, Xiaoou Tang, “Guided Image Filtering,” IEEE Transactions on Pattern Analysis and Machine Intelligence., Volume: 35 , Issue: 6 , June 2013.

[7] Yi-FeiPu, Ji-Liu Zhou, Xiao Yuan, “Fractional Differential Mask: A Fractional Differential-Based Approach for Multiscale Texture Enhancement, ” IEEE Transactions on Image Processing., Volume: 19 , Issue: 2 , Feb. 2010.

[8] Chul Lee, Jin-Hwan Kim, Chulwoo Lee, “Optimized Brightness Compensation and Contrast Enhancement for Transmissive Liquid Crystal Displays,”IEEE Transactions on Circuits and Systems for Video Technology., Volume: 24 , Issue: 4 , April 2014.

[9] SnehanShourya, Sumit Kumar, Rajib Kumar Jha, “Adaptive fractional differential approach to enhance underwater images, ”nternational Symposium on Embedded Computing and System Design., 15-17 Dec. 2016.

[10] C. Tomasi, “Bilateral filtering for gray and color images,”Sixth International Conference on Computer Vision., 06 August 2002.

[11] A. Buades, B. Coll, and J.-M.Morel, ``Nonlocal image and movie denoising,'' Int. J. Comput. Vis., vol. 76, no. 2, pp. 123_139, Jul. 2007.

[12] K. He, J. Sun, and X. Tang, ``Guided image _ltering,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1397_1409, Jun. 2013.

[13] Z. Farbman, D. Lischinski, and R. Szeliski, ``Edge-preserving decompositions for multi-scale tone and detail manipulation," Trans. Graph., vol. 27, no. 3, p. 67, Aug. 2008.
[14] Y.-F. Pu, J.-L. Zhou, and X. Yuan, ``Fractional differential mask: A fractional differential-based approach for multiscale texture enhancement," IEEE Trans. Image Process., vol. 19, no. 2, pp. 491_511, Feb. 2010.
[15] E. D. Pisano et al., ``Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," J. Digit. Imag., vol. 11, no. 4, pp. 193_200, 1998.
[16] J. A. Stark, ``Adaptive image contrast enhancement using generalizations of histogram equalization," IEEE Trans. Image Process., vol. 9, no. 5, pp. 889_896, May 2000.
[17] D. Coltuc, P. Bolon, and J.-M. Chassery, ``Exact histogram speci_cation," IEEE Trans. Image Process., vol. 15, no. 5, pp. 1143_1152, May 2006.
[18] X. Wu, ``A linear programming approach for optimal contrast-tone mapping," IEEE Trans. Image Process., vol. 20, no. 5, pp. 1262_1272, May 2011.
[19] H. Xu, G. Zhai, X.Wu, and X. Yang, ``Generalized equalization model for image enhancement," IEEE Trans. Multimedia, vol. 16, no. 1, pp. 68_82, Jan. 2014.
[20] A. K. Jain, Fundamentals of Digital Image Processing. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.
[21] C.-H. Chou and Y.-C. Li, ``A perceptually tuned subband image coder based on the measure of just-noticeable-distortion pro_le," IEEE Trans. Circuits Syst. Video Technol., vol. 5, no. 6, pp. 467_476, Dec. 1995.
[22] G. D. Finlayson and E. Trezzi, ``Shades of gray and colourconstancy,"in Proc. Color Imag. Conf. Spring_eld, VA, USA: Society for Imaging Science and Technology, 2004, pp. 37_41.
[23] R. C. Gonzalez and R. E.Woods, Digital Image Processing. Reading, MA, USA: Addison-Wesley, 1992.
[24] J. Wu, G. Shi, W. Lin, and C. C. J. Kuo, ``Enhanced just noticeable difference model with visual regularity consideration," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2016, pp. 1581_1585.
[25] H. Zhang, Q. Zhao, L. Li, Y.-C.Li, and Y.-H. You, ``Muti-scale image enhancement based on properties of human visual system," in Proc. 4th Int. Congr.Image Signal Process.(CISP), vol. 2, Oct. 2011, pp. 704_708.