Research on Single Image Dehazing Methods Based on Deep Learning

Bian Gui[1, a], Yuhua Zhu[2, b], Tong Zhen[1, c, *]

[1]College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China;

[2]Henan University of Technology, Zhengzhou 450001, China.

[a]1752011657@qq.com, [b]GB18236311520@163.com, [c]15188352425@163.com

Abstract

Under severe weather conditions such as haze, the quality of outdoor captured images is seriously degraded, which affects the further analysis, identification and utilization of image content by the monitoring system. With the continuous development of deep learning, various neural network-based methods are applied to the image dehazing field. In order to study the advantages and disadvantages of image dehazing based on deep learning methods, this paper introduces the classic DehazeNet algorithm, Multiscale_CNN algorithm, and finally summarizes the shortcomings of two classical dehazing algorithms based on deep learning indicate the direction for further research.

Keywords

Haze; deep learning; DehazeNet;Multiscale_CNN.

1. Introduction

Some researchers have tried to use deep neural networks for image dehazing by the human brain without any information to distinguish the blurred objects from the ambiguous scene and to inspire the haze thickness. For example, Cai et al. [2] proposed the DehazeNet algorithm, which used the convolutional neural network for the first time to perform image dehazing in a learning manner. The algorithm automatically learns haze-related features in a data-driven manner, avoiding the limitation of manual feature extraction by traditional methods, to get a more accurate transmission map. However, when the depth of field transition in the image is large, especially in the sky area, the restored image is not ideal. Ren et al. [3] proposed the Multiscale_CNN algorithm to design a multi-scale network for single image dehazing. First, a coarse-scale network was used to obtain a rough transmission map, which was further refined by a fine network. However, this method requires manual adjustment of parameters based on foggy image density. Tang et al. proposed a fog correlation feature algorithm. By studying various haze-related features based on the learning framework, the image features were first extracted, and then the feature vectors were input into random forests. Four fog correlation features were summarized and compared using random forest algorithm. The four kinds of fog-related features are combined to estimate the atmospheric transmittance.

2. DehazeNet Method

The human visual system can estimate the relationship between fog concentration and scene depth without relying on explicit feature transformation. DehazeNet designed a special convolutional neural network. After training, the network model can intelligently learn the haze characteristics [1] and solve the defects of manual feature extraction. The method is an end-to-end network model, and the dehazing process is as follows:

(1) Feature extraction: Different from the traditional convolutional layer, the first layer structure of the DehazeNet method is "convolution + Maxout", and Maxout has a very high nonlinear fitting ability. The layer expression is shown in Equation 1:

$$F_1^i(x) = \max_{j\in[1,k]} g^{i,j}(x), g^{i,j} = W_1^{i,j} * I + B_1^{i,j} \tag{1}$$

W1 is the inverse filter, the maximum of the channel is equivalent to the minimum of the channel, equivalent to the dark channel prior; when W1 is the ring filter, equivalent to the contrast extraction, equivalent to the maximum contrast; when W1 contains the opposite The filter and the all-pass filter are equivalent to the RGB to HSV color space conversion, which is equivalent to the color attenuation prior.

(2) Multi-scale mapping: Multi-scale features can calculate features of input images on multiple spatial scales, and have scale-invariant characteristics when performing feature extraction. Although the traditional dehazing method also improves the robustness of feature extraction at different resolutions by means of mean, minimum and median values of different scales, it has been experimentally proved that multi-scale mapping exhibits higher robustness [4]. The DehazeNet method uses three sets of different scale (3 × 3, 5 × 5, 7 × 7) filters to achieve multi-scale robustness. The expression is as follows:

$$F_2^i(x) = W_2^{[i/3],(i\backslash 3)} * F_1 + B_2^{[i/3],(i\backslash 3)} \tag{2}$$

Local extremum: This step constrains the local consistency of the transmittance and can effectively suppress the estimated noise of the transmittance. In addition, the local extrema also corresponds to the local minimum of the dark channel prior and the local maximum of the maximum contrast. The expression is as follows:

$$F_3^i(x) = \max_{y\in\Omega(x)} F_2^i(y) \tag{3}$$

Nonlinear regression: Sigmoid function is prone to gradient disappearance during neural network training, resulting in low convergence and undesired local optimal solution; RELU function is a defect in regression problem in image restoration, but RELU can suppress values when they are far away When it is far below 0, the last layer may cause a response overflow [5], because for image recovery, the output value of the last layer should be the lower limit and the upper limit in a small range. To avoid this limitation, the DehazeNet method uses BRELU as the activation function, which maintains both bilateral and local linearity.

3. Multiscale_CNN Method

In order to avoid the limitation of manual feature extraction in the traditional dehazing method, Ren et al. proposed a multi-scale convolutional neural network for image dehazing. The network model learns the mapping relationship between the hazy image and its corresponding transmission map during the training process. Firstly, the coarse transmission map is predicted based on the coarse scale network, and then the coarse transmission map is refined by the fine scale network. Finally, the clear image is restored according to the atmospheric scattering model [6]. The method of defogging is as follows:

(1) Coarse-scale network: This network model is used to predict the overall transmission map of an image. It consists of four parts: convolution, pooling, upsampling, and linear combination. Input a color image, the filter is convolved with the feature image of the input image, and the response of each convolution layer is given by Equation 4.

$$f_n^{l+1} = \sigma(\sum_m (f_m^l * k_{m,n}^{l+1}) + b_n^{l+1}) \tag{4}$$

f_n^l is the feature map of the current convolutional layer, and n is the feature map of the next volume. In addition, k is a convolution kernel, and (m, n) represents a current feature map from the m layer to the next n feature map.

(2) Pooling layer: In order to maintain the image (scale, translation, rotation) invariance and reduce parameters, reduce the amount of calculation to prevent over-fitting, improve the generalization ability of the model, Multiscale_CNN method uses maximum pooling operation, each convolution. The layer's downsampling factor is set to 2. After this step, the image feature map size is halved.

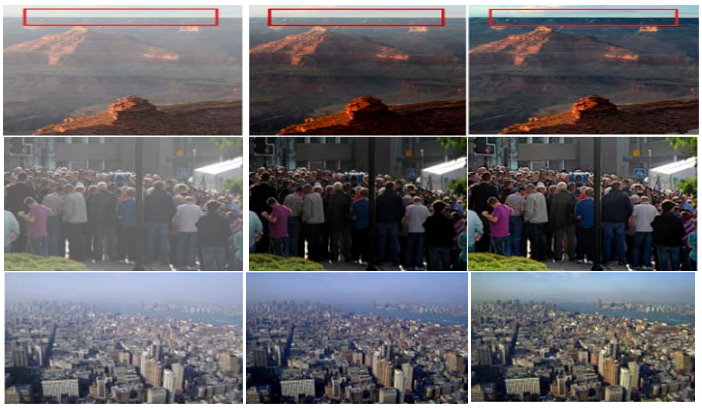
(3) Upsampling: Since the image size is halved after the second step operation, the purpose of increasing the upsampling layer is to keep the output image size consistent with the input image size, and retain the nonlinear characteristics of the network.

(4) Linear combination: In the penultimate layer of the coarse-scale network model, the image features have multiple channels in the layer, so a linear binder is required to combine the image feature channels.

(5) Fine-scale network: In order to further refine the coarse transmission map obtained from the coarse-scale network, the Multiscale_CNN method designs a fine-scale network similar to the coarse-scale network structure, so that the obtained transmission map is more accurate.

4. Comparative Study of Methods

In order to visually compare the post-fog effects of the above two algorithms, this paper selects the real-world images and compares them with qualitative and quantitative indicators. It can be seen from Fig. 1 that the DehazeNet method is more natural for the sky region of three images, close to the real image [7], but the method deals with the overall brightness of the image is dark; the color image processed by the Multiscale_CNN method is more integrated with the DehazeNet method. Brighter, closer to the real image, but not ideal for sky area processing.



(a) input images (b)DehazeNet Method (b)Multiscale_CNN Method

Fig 1. Comparisons of dehazing methods

5. Summary

The DehazeNet method and the Multiscale_CNN method are used to estimate the transmission parameters. Compared with the prior method, the dehazing effect of the algorithm is more accurate, the probability of over-enhancement is lower, the universality is strong, and the robustness is high. The disadvantage is that the effect is not ideal for the heavier part of the image, and the traditional method is used to estimate the sky brightness parameter. When dealing with images with strong brightness and darkness, the details of the dark part of the image are lost, which is not suitable for processing bright areas with large areas. Image [8]; in the defogging process, the network model only predicts the transmission map, and the estimation of atmospheric light requires another algorithm, so in the future research, we can consider designing an integrated network model. The transmission map and atmospheric light are predicted to avoid errors in the calculation process.

References

[1] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior，” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 12,pp. 2341–2353, Dec. 2011.

[2] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, “DehazeNet: An end-to-end system for single image haze removal,” IEEE Trans. Image Process.,vol. 25, no. 11, pp. 5187–5198, Nov. 2016.

[3] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang，“Single image dehazing via multi-scale convolutional neural networks，” In European.

[4] J.-P. Tarel and N. Hautière, “Fast visibility restoration from a single color or gray level image,” in Proc. IEEE Conf. Comput. Vis., Sep./Oct. 2010,pp. 2201–2208.

[5] Y.-S. Lai, Y.-L. Chen, and C.-T. Hsu, “Single image dehazing with optimal transmission map,” in Proc. IEEE Conf. Pattern Recognit., Nov. 2012,pp. 388–391.

[6] B. Li, X. Peng, and Z. Wang, J. Xu, and D. Feng, “AOD-Net: All-inOne dehazing network,” in Proc. IEEE Conf. Comput. Vis., Oct. 2017,pp. 4780–4788.

[7] C. Chen, M. N. Do, and J. Wang, “Robust image and video dehazing with visual artifact suppression via gradient residual minimization,” in Proc.Eur. Conf. Comput. Vis., 2016, pp. 576–591.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A feature similarity index for image quality assessment,” IEEE Trans. Image Process., vol. 20,no. 8, pp. 2378–2386, Aug. 2011.