ORIGINAL RESEARCH PAPER

Volume-8 | Issue-1 | January-2019 | PRINT ISSN No 2277 - 8179

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

AN IMAGE DEHAZING ALGORITHM BASED ON SELF-ATTENTION GENERATIVE MODEL

Computer Science

Bowen Li

Tianjin Polytechnic University, School of Computer Science and Software Engineering, Binshuixidao 399, Tianjin, China, 30000

ABSTRACT

The traditional methods of single image dehazing are composed of the methods based on image enhancement technique and the methods based on the atmospheric scattering model, the former is simply adjusted according to the statistical global features or local features, it does not take the content of the image information into account, although the latter is strongly explainable and the result is relative better, they need extra parameters, so their generalization ability is poor. In this paper, an image dehazing algorithm based on self-attention generative model is proposed, which make dehazeing images through end-to-end training process available.

KEYWORDS

dehazing; self-attention; generative model

Introduction

With the development of automatic driving technology and intelligent security systems, clear and identifiable images and videos are becoming more and more important as the basement of securing the normal operation of the whole system. Single image haze removal technology is to suppress the influence of haze in the image captured and processed by human or machine vision. But dehazing the images taken in heavy hazed weather circumstances is still a challenging problem. It is troubled with the unknown depth which varies on the whole image. There are many approaches proposed to dealing with removing haze from a single image. Neural network based[1], atmospheric scattering model based[2][8], histogram based[3] are three main categories of them at the moment.

Related works

Histogram based methods takes the values of each pixel into consideration and calculates the statistical result of the pixel value by bins. The most typical Histogram based method is Contrast Limited Adaptive Histogram Equalization(CLAHE). CLAHE is improved a kind of the local histogram equalization (also known as Adaptive Histogram Equalization, AHE). These methods which based on histogram are time cost saving and very easy to implement with a relatively worse haze removal results. So some experts try to bring this issue under explainable physical model -- atmospheric scattering model[2][8].

The most classical method based on atmospheric scattering model is the haze removal algorithm (DCP) based on dark channel prior proposed by He et al.[4]. This method is designed on the prior knowledge obtained from the observation of natural images, namely, the pixel value of the dark channel image of the haze-free image is close to 0, while the pixel value of the dark channel image of a certain region of the haze image increases with the increase of haze concentration, and it can quickly estimate the atmospheric scattering function. The non-local image dehazing algorithm proposed by Berman et al.[5] is another classical image restoration method based on atmospheric scattering model, which has a better effect. However, the disadvantage is that some parameters should be set in advance and the generalization ability is weak. The defogging effect of this kind of methods is ideal, but the complexity of smoothing atmospheric transmission function is too high, resulting in a relatively slower dehazing speed.

Recently, influenced by the prevalent neural networks in the field of deep learning and computer vision[9], a kind of methods based on neural networks are proposed and achieve better haze removal level. Dehaze-net[1] is a pioneer and a successful trial with neural networks. Self-attention generative dehazing model we proposed is also based on deep learning and Generative Adversarial Nets(GAN)[10].

Dehazing with self-attention generative model

Because the haze may diffuse in every position of the image, is distributed uniformly and the haze concentration varies with the depth of scene. The basement of Convolutional neural network(CNN) -- convolution kernels have local receptive fields, which means that one layer could only have perceptions on a small region of last layer. Although the depth of the neural network of the multilayer network architecture can make the receptive field of the latest layers of network expanded to the whole image, but also it may cause learning disabilities of the long correlation (long - term dependencies) [7]. To handle this problem, the theory of attention mechanism is proposed. The earliest attention mechanism is proposed in the field of natural language processing[6], Then Zhang et al.[7] introduced this idea into the field of computer vision, In this method, the self-attention layer is used to better balance the long correlation and computational efficiency of the model, and the global information is used to constrain the dehazing process.

The self-attention layer calculates the response using a weighted sum of the output values of the positions used in the input feature map. Suppose output of layer just before the layer of the attention layer as $x \in R^{CXN}$ Firstly, calculate the feature map $f(x)$ and $g(x)$ with:

$$f(x)=Wfx,\ g(x)=W_g x \qquad (1)$$

Then use $\beta j,i$ to indicate the attention weights of the model on the position i when generating the region j. There is the formula to calculate $\beta j,i$ as follow:

$$\beta_{j,i} = \frac{e^{s_{ij}}}{\sum_{i=1}^{N} e^{s_{ij}}}, \text{with } s_{ij} = f(x_i)^T g(x_j) \qquad (2)$$

So the output of self-attention layer could be expressed as $o=(o_1,o_2,...,o_j,...,o_N) \in R^{C\times N}$, then

$$o_j = \sum_{i=1}^{N} \beta_{j,i} h(x_i), \text{with } h(x_i) = W_i x_i \qquad (3)$$

In the formula mentioned above, the learnable weights $Wg \in R^{CXC}$, $W_f \in R^{CXC}$, $W_h \in R^{CXC}$ are implemented with convolution filters of size 1*1. In the following experiments, we use C =C/8.

In addition, in spired by the tricks used in the work[1], Bilateral Rectified Linear Unit(BRELU), is also used in our model to replace the common RELU activation function, and BRELU corresponds to the edge suppression operation (such as DCP) in the traditional defrosting method. Bilateral constraints introduce prior information to narrow the search space of parameters and make the network easier to train. The following are our experiments.

Experiments

In this section, we compare the model we proposed with CLAHE、DCP、Non local and Dehaze-net. Firstly we compare the time cost of each method. We take 100 images with haze as their input, and calculate the average time cost per image. The results are demonstrated as the following table.

Table 1. The average time cost of different methods

Method	CLAHE	DCP	Non local	Dehaze -net	Our method
Time Cost (s/per image	0.363	2.819	6.333	3.827	1.211

In table 1, our method can achieve a high computation efficiency with full use of GPU, a kind of parallel computing device. The time cost of CLAHE is much smaller than ours because the methods is very simple and the effect of dehazing is poor. The haze removal results are shown below.



Figure 1. The dehazing results by different models.

As we can see in figure 1, the effect of CLAHE algorithm is limited. DCP trends to make the image darker, so a post process if often needed and adjusted by hands. Non local is relative better than the former but saturation is too high to behave like a real image. Dehaze-net is quite good, but our method is much better obviously.

Conclusion

In this paper, the image dehazing algorithm based on self-attention generative model we proposed is compared with the image based on the processing of classical enhancement method and the classical physical model -- atmospheric scattering model. The experimental results show that the effect of deep network based on a large number of samples end-to-end training is prior to the traditional dehazing algorithm. The self-attention mechanism we introduced into our model does take significant effects. The algorithm proposed in this paper can significantly improve the contrast of images, restore real information, and then directly improve the effect of the following visual tasks (such as image segmentation, target recognition, etc.).

The physical model has incomparable interpretability, and it is difficult to guarantee the correctness and extensibility of the model simply because of the nonlinear relationship between the degradation of the scene and the depth of the scene. Even so, it is still of great value for the research and improvement of the algorithm of haze day image restoration based on the physical model. Research on how to build a universal physical model that is more consistent with the complex real environment to explain the influence of haze on images is also the research direction of the intersection of computer and physics. In addition, it is another interesting research direction to infer depth information based on the concentration of haze.

REFERENCES

[1] Cai, B., Xu, X., Jia, K., Qing, C., & Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing, 25(11), 5187-5198.
[2] Narasimhan, S. G., & Nayar, S. K. (2000). Chromatic framework for vision in bad weather. In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on (Vol. 1, pp. 598-605). IEEE.
[3] Kim, T. K., Paik, J. K., & Kang, B. S. (1998). Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering. IEEE Transactions on Consumer Electronics, 44(1), 82-87.
[4] He, K., Sun, J., & Tang, X. (2011). Single image haze removal using dark channel prior. IEEE transactions on pattern analysis and machine intelligence, 33(12), 2341-2353.
[5] Berman, D., & Avidan, S. (2016). Non-local image dehazing. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1674-1682).
[6] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998-6008).
[7] Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2018). Self-Attention Generative Adversarial Networks. arXiv preprint arXiv:1805.08318.
[8] Narasimhan, S. G., & Nayar, S. K. (2002). Vision and the atmosphere. International Journal of Computer Vision, 48(3), 233-254.
[9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436.
[10] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).
[11] Reza, A. M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. Journal of VLSI signal processing systems for signal, image and video technology, 38(1), 35-44.