Image Defogging based on combined Sparse Gradient Minimization and CNN Architecture

Anisha Jana
Department of Electronics and Communication Engineering
National Institute of Technology, Rourkela
Odisha, India
anishajana4@gmail.com

Upendra Kumar Sahoo
Department of Electronics and Communication Engineering
National Institute of Technology, Rourkela
Odisha, India
sahooupen@nitrkl.ac.in

***Abstract*—Some environmental factors like haze or fog degrades the quality of the image. These factors affect some real time processes such as object detection and recognition, automated vehicles and remote sensing which needs clear visible images for making critical decisions. Therefore, restoring the true image from the foggy image becomes significant. Now with the advancement in image processing, many image defogging and dehazing algorithms has been developed to improve the quality of the image. Many standard filtering techniques such as high boost filter, homomorphic filter can be used for image defogging but it fails to restore the foggy images completely so some advanced techniques like dark channel prior, decomposition techniques, convolution network-based algorithms are used. Image quality assessment (IQA) is done to measure the quality of the defogged image. These performance metrics mainly includes mean squared error (MSE), structural similarity index metric (SSIM),peak signal to noise ratio (PSNR).**

***Index Terms*—defogging, image enhancement, dark channel prior (DCP), convolution neural network (CNN), sparse gradient minimization (SGM), atmospheric scattering model (ASM).**

I. INTRODUCTION

The most frequent atmospheric phenomena are haze and fog. They both absorb and disperse the reflected light as well as the ambient light that is directed towards the camera [11]. Physical or non-physical are the most widely used types of model to improve the image quality. Three sorts of image dehazing algorithms exist: image enhancement, image restoration and learning-based. Image quality is mostly improved through image enhancement techniques in accordance with the image information. It is based on non-physical model. It uses a variety of methods, including CLAHE, histogram analysis, and image filtering. Image restoration is a representation of a physical model where the problem is solved by applying an atmospheric light scattering model after taking into account the cause of the degradation. A single image haze removal method utilising the DCP algorithm was proposed by He et al [8]. Using the atmospheric scattering model, they calculated the ambient light and the transmittance map and finally produced the recovered image. This approach is efficient, but the recovered image will exhibit severe colour distortion for images with huge white spaces, like sky or snow. KeYan Wang et al [7] employed a sky detection and segmentation technique to partially defog the sky region. However colour distortion persisted in some regions of the defogged image due to border restrictions and regularisation of the transmission map. Although the aforementioned conventional image defogging techniques have advanced greatly and produced excellent results, many of them depend on different prior knowledge. There is a high chance of getting inaccurate results. These issues are addressed utilising a range of deep learning-assisted techniques that recreate high-quality de-hazed images using end-to-end learnable networks. For a single foggy image reconstruction, Li et al. [8] offered a learning-based method. They collectively approximated the ambient light and the transmission map by using a combined CNN and an atmospheric scattering model. Multi scale convolution neural networks were used by Cai et al [14] to determine the transmittance pattern of the hazed image and learn the hazy features. They then did the inverse operation on the fog-free image. Since the data-set was a localized image after cutting and the network's characteristics lacked global representation, the defogged image only showed partial defogging. AOD network is a learning-based defogging technique that builds defogging maps from fog maps in a single process while also using CNN to estimate ambient light and transmittance of the hazy images, as proposed by Li et al. [8]. The quality of the recovered pictures was hampered by the network model's shallow topology and inability to effectively grasp the attributes of the hazy images. Overall, these learning-based methods demonstrated exceptional performance for image dehazing issues. However, some of them were developed using artificial hazy data, which has drawbacks when applied to real-world hazy pictures and some of them also raises the overall computing cost and complexity. Our suggested technique differs from the current deep learning-based approaches. The first component is a sparse gradient minimization (SGM) module, which boosts contrast and maintains edges in the recovered image. This is combined with the ASM and CNN architecture which approximates both the transmittance map and the dark channel. This lowers the overall computational cost and complexity. The remaining of the paper is structured as follows. In section II, SGM is employed as a pre-processing step, and we then go over the atmospheric scattering model and the CNN-based architecture that were both used in this study. Section III provides a description of the data-set used, the results, and a comparison of performance metrics. We give our final conclusions in section IV.

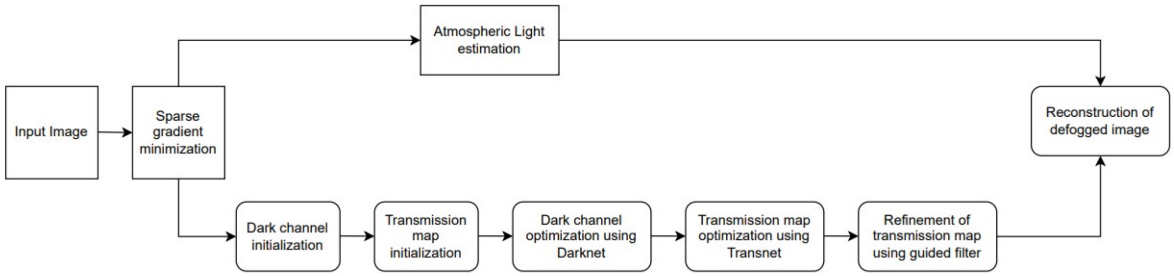

Fig. 1. Block diagram of the proposed learning based method.

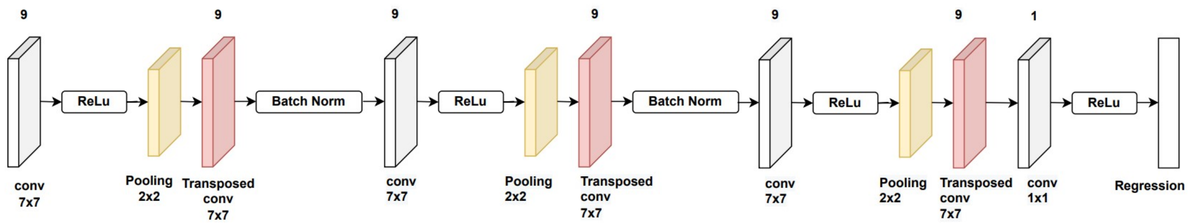

Fig. 2. Darknet architecture

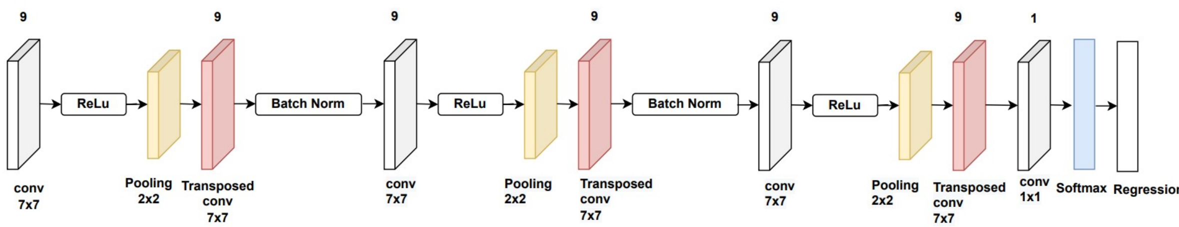

Fig. 3. Transnet architecture

II. PROPOSED METHOD

In this section, first a sparse gradient module is explained followed by the scattering model and finally the CNN based architecture used for dark channel and transmission map optimization is described. Figure 1 displays the suggested method's overall block diagram.

A. Sparse Gradient Minimization

Sparse gradient minimization's key contribution is to limit the discrete number of intensity variations between nearby pixels, which logically connects to the L0 norm for information sparsity search. It mainly preserves the major edges and removes the low amplitude structures [15]. The input discrete signal is represented as i and the smoothed output as y. The approach counts discrete amplitude changes, denoted as

$$c(y) = \#\{x| \; |y_x - y_{x+1}| \neq 0\} \tag{1}$$

where x and x + 1 are nearby samples. A gradient with respect to x is represented by the forward difference $|y_x - y_{x+1}|$. The counting operator # results the number of x that satisfy $|y_x - y_{x+1}| \neq 0$. The related objective function is expressed as:

$$min_y \sum_x (y_x - i_x)^2 \; subject \; to \; c(y) = q \tag{2}$$

The expression c(y) = q denotes that the result has q non-zero gradients. Equation (2) has a great deal of capability for structural information abstraction. A bigger q results in a closer approximation that yet captures the most pronounced contrast. In real-world situations, the value of q in equation (2) could vary from tens to thousands. In order to manage it, we use a generic form shown in equation (3) to try to strike a compromise between flattening the structure and producing results that are similar to the input.

$$min_y \sum_x (y_x - i_x)^2 + \lambda.c(y) \tag{3}$$

where c(y), which is actually a smoothing parameter, is directly controlled by the weight λ. The result has extremely few edges when λ is large. This optimization problem is solved using a unique optimization technique that involves the introduction of auxiliary variables [15].

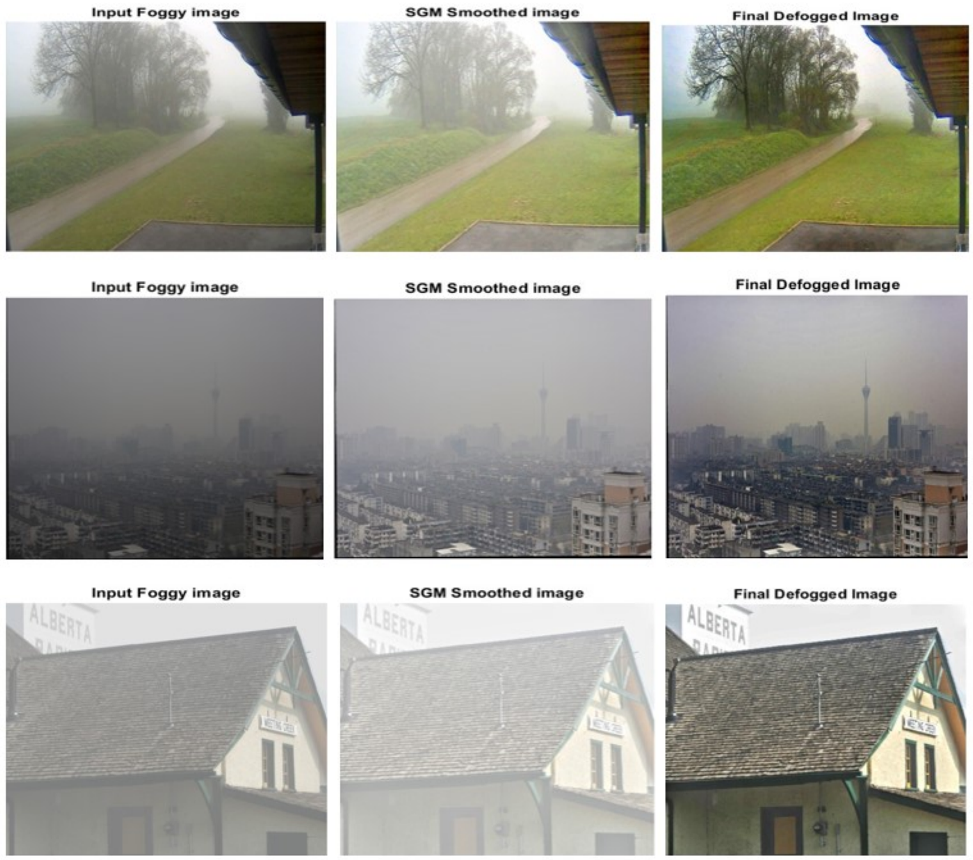

Fig. 4. Sample images each from MRFID, BeDDe and HazeRD data-set and their SGM smoothed image and final defogged image respectively.

B. Dark Channel Prior

Dark channel prior algorithm is the basic mathematical model for image defogging. In DCP algorithm, a minimum filter of window size 15 is used to obtain the DCP image.

$$I_{Dark} = \min_{w \in \Omega(x)} (min_{C \in (R,G,B)} (I^C)) \tag{4}$$

I(dark) is the obtained dark channel, I(c) is the pixel intensity value of individual channels of image I, c is the individual color channel of RGB. The greatest intensity value in each colour channel is individually determined from the locations of the 0.1% pixels of highest intensity for the purpose of calculating ambient light. These three RGB channel intensity values are used to represent ambient light A. The initial transmittance map is calculated by normalising the scattering model equation and applying the minimum filter as shown by the following equation.

$$t(x) = 1 - f\frac{1}{3}(\sum_{C=1}^{3} \frac{I^C}{A^C}) \tag{5}$$

Where f is the correction factor set between [0,1] which makes the image look more natural, I is the input blurry image, while A is the ambient light. For the refinement of the transmittance map, a guided filter has been used. This dark channel and transmission map is optimized using CNN architecture which is discussed in the next subsection. Finally, the reconstructed image is estimated by the DCP model equation.

$$I(x) = J(x)t(x) + (1 - t(x))A \tag{6}$$

Where I(x), A, t(x) and J(x) denotes the input foggy image, ambient light, transmission map and defogged image respectively.

C. CNN based Architecture

Most of the CNN based architecture developed in this field tries to optimize the combined atmospheric light and transmission map, however here we have developed two CNN based architectures called Darknet and Transnet shown in figure 2 and 3 to optimize initial dark channel and transmittance map obtained from atmospheric scattering model.

Convolution layers:The filters or the kernels that are convoluted with the input feature maps make up this layer. Only one filter of size 1x1 is present in the final convolution layer; all other layers have nine filters of size 7x7.

Pooling Layer: With a stride of 2, the pool size used for this layer is 2x2. This layer largely keeps the crucial data while lowering the parameters, dimensions and resolution of the feature maps.

Transposed Convolution Layer: It slides the input over the kernel and conducts element-wise multiplication and summation in place of sliding the kernel over the input and conducting these operations. All these layers contain 9 filters of size 7x7.

Batch Normalization Layer: It normalizes a small batch of all the data. Between convolution layers and non-linear activation functions, this layer speeds up convolution neural network training.

Regression layer: The regression layer computes the half mean squared error loss for regression problems.

III. RESULTS

Several data-sets have been used to test the proposed approaches. To start, foggy images from the Multiple Real-World Foggy Image data-set (MRFID) [3] are used. The clear and hazy MRFID images cover 200 outdoor scenarios. Benchmark Data-set (BeDDe) for dehazing Assessment [6] is another natural data-set included in this study. The BeDDE has 208 pairs of natural pictures, each of which mainly consists of a naturally blurred image and a clear reference image. Lastly, we used a synthetic data-set named HazeRD to test the suggested method. To test dehazing techniques, fifteen different real-world outdoor images with artificial haze in five different weather settings are included in the Haze Realistic Data-set [9]. The input foggy image from each of these three data-sets is shown in Figure 4, along with the smoothed output produced by applying SGM and the corresponding defogged result. It can be observed that the suggested technique recovers a haze-free image while maintaining every semantic of the original image.

A. *Quantitative Analysis*

Various authors adopted various criteria to assess their findings. Table. 1 shows the quantitative analysis with different parameters measured on images from HazeRD synthetic data-set and Table.2 shows the same for MRFID natural data-set. It can be observed from the tables that our proposed SGM with CNN architecture method gives better PSNR, SSIM and RMS Contrast as compared to various other methods listed. Here higher values of RMS contrast corresponds to more contrast. A popular objective statistic for evaluating images is PSNR. Less distortion occurs as the PSNR value increases. A reference image that is initially distortion-free serves as the basis for the SSIM index. The general principle is that as the SSIM value increases, the image becomes less distorted.

TABLE I
AVERAGE PSNR, SSIM AND RMS CONTRAST VALUES OF VARIOUS METHODS ON HAZERD DATASET.

Method	PSNR	SSIM	RMS Contrast
DCP [29]	18.53	0.8337	18.42
AOD-Net [29]	19.69	0.8478	19.84
VMD [1]	23.24	0.9172	16.97
Proposed method	30.85	0.9285	26.04

TABLE II
AVERAGE PSNR, SSIM AND RMS CONTRAST VALUES OF VARIOUS METHODS ON MRFID DATASET.

Method	PSNR	SSIM	RMS Contrast
DCP	12.36	0.344	9.43
AOD-Net	16.84	0.6253	15.27
VMD	14.16	0.5923	14.25
Proposed method	22.77	0.8197	19.69

IV. CONCLUSION AND FUTURE SCOPE

The proposed work produces better and faster results than traditional DCP and some other methods like VMD. The performance metrics like PSNR and SSIM and RMS contrast has also improved. Once the network is trained, it produces the defogged output in 8.07s as compared to 99s from DCP based method and 86s from variational mode decomposition method [1]. The method can be further optimized more to produce result in lesser time.

REFERENCES

[1] QIN Chaoxuan and GU Xiaohui “A single image dehazing method based on decomposition strategy.” Journal of Systems Engineering and Electronics Vol. 33, No. 2, 2022.

[2] Shuping Li, Qianhao Yuan, Yeming Zhang, Baozhan Lv and Feng Wei “Image dehazing algorithm based on deep learning coupled local and global features” MDPI journal Appl. Science 2022.

[3] Wei Liu, Fie Zhou, Tao Lu, Jiang Duan and Guoping Qiu “Image defogging quality assessment: Real world database and Method” IEEE Transactions on image processing, vol. 30, 2021.

[4] Weilin Luo, Shunqiang Duan and Jiwen Zheng “Underwater image restoration and enhancement based on a fusion algorithm with color balance, contrast optimization and histogram stretching” IEEE Access vol.9, 2021.

[5] Wenhan Yang, Shiqi Wang and Jiaying Liu, ”Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement”. IEEE transactions on image processing, vol. 30, 2021

[6] Shiyu Zhao, Lin Zhang, Ying Shen and Shengjie Zhao “Dehazing evaluation: real world benchmark datasets, criteria and baselines” IEEE transaction on image processing vol.29, 2020.

[7] Wang, Y.K.; Hu, Y.; Wang, H.; Li, Y.S. ”Image dehazing algorithm by sky segmentation and superpixel-level dark channel”. J. Jilin Univ. 2019, 49, 1377–1384

[8] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “AOD-Net: All-in-one dehazing network,” in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2017, pp. 4770–4778.

[9] Yangfu Zhang, Li Dhing and Gaurav Sharma ”HazeRD:An outdoor scene Dataset and Benchmark for Single Image Dehazing”. 978-1-5090-2175-8/17/ 2017 IEEE

[10] Cai, B.; Xu, X.; Jia, K. ”An End-to-End System for Single Image Haze Removal”. IEEE Trans. Image Process. 2016, 25, 5187–5198.

[11] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao and Ming-Hsuan Yang “Single image dehazing via Multi-scale convolutional neural networks” . Springer International Publishing AG 2016.

[12] Yong –Sik Shin, Younggun Cho, Gaurav Pandey and Ayoung Kim “Estimation of ambient light and transmission map with common convolutional architecture” . IEEE conference, 16518024, 2016.

[13] Meng, G.; Wang, Y.; Duan, J. ”Efficient Image Dehazing with Boundary Constraint and Contextual Regularization”.IEEE International Conference on Computer Vision, Sydney, Australia, December 2013

[14] 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.

[15] Li XU, Cewu Lu, Yi Xu and Jiaya Jia, ”Image Smoothing via L0 Gradient Minimization”. ACM Transactions on Graphics, Vol. 30, No. 6, Article 174, December 2011