Image dehazing and denoising based on multiple methods

Xiao Lu
Electrical Computer Engineering department
University of California San Diego
La Jolla, US
xil078@eng.ucsd.edu

Sai Ding
Electrical Computer Engineering department
University of California San Diego
La Jolla, US
sading@eng.ucsd.edu

***Abstract*—In this project, we proposed three advanced models to implement image dehazing and denoising. The first traditional model is Open dark channel model(ODC) which aims to do image dehazing at the low frequency part and do image denoising at the high frequency part synchronously. The second model is an end to end feature fusion network(FFA-Net) which owns huge flexibility to deal with different pixels and features information, improving the skills of wild-used CNN based dehazing model. The third model is an end to end gated context network(GCANet) which can eliminate gridding artifacts in dilated convolution and use gated sub-network to fuse features of different levels. Code has been made available at Github**

***Index Terms*—Image dehazing, ODC, FFA, GCA, PSNR, SSIM**

I. Introduction

In the field of image processing, image dehazing is a very important issue, which has a great impact on the normal operation of the image in a noisy environment. Due to atmospheric absorption, such environment consist of haze, smoke and fog which lead to images with color distortion, blur, low contrast and other visible quality degradation. Thus, some visual problems will be tough to deal with, such as image tracking, classification, and object detection.

Restoring the clear image from the hazy image is very difficult, because haze is dependent on unknown depth information. Therefore, some methods using multiple image dataset can be effective, such as neural networks with training process or depths based models required rough depth information. Both of them are based on atmosphere scattering model. And a typical example is shown on Fig. 1.

Fig. 1. Image dehazing based on atmosphere scattering model

II. Literature review

The purpose of this project is to implement both traditional and advanced machine learning methods to eliminate the haze effect in the captured image and rebuild the original color of the natural scene. The original motivation came from some image contrast improvement approaches such as linear mapping, the gamma correction[1], and histogram equalization [2]. Nevertheless, such methods cannot adaptively compensate for the reduction in haze because they do not consider the difference in haze thickness, which is proportional to the depth of the object.

Therefore, how to implement image dehazing with multiple images which can provide additional information has become an important issue to many researchers. Schechner [3]proposed a polarization-based method which can eliminate the haze effect through two or more images taken with different polarization degrees. Also, Narasimhan [4] selected multiple images from different weather conditions to compare image dehazing effect. After that, they proposed a depth-based model [5] simple additional information from users, which can be used as an easy-to-use plug-in for various image processing software. However, it may be difficult to obtain interactive input specified by the user in a constantly changing scene when multiple images cannot be obtained in practice.

After that, researchers implement great progress in single image haze removal by using reasonable priors. Tan [6] proposed two strong assumptions, images with enhanced visibility (or images on sunny days) have higher contrast than images plagued by bad weather; and the illuminator whose change mainly depends on the distance from the object to the observer tends to be smooth. Thus , this method eliminates the haze by maximizing the local contrast of the restored image. Besides, R.Fattal [7] proposed a new method for estimating light transmission in hazy scenes given a single input image. Thus, scattered light can be eliminated to increase the visibility of the scene and restore the contrast of a fog-free scene according to this estimate. Although both approaches can create impressive dehazing results, the haze effect may not be eliminated if these assumptions cannot stand.

Until recently, He [8] and his group proposed a dark channel prior model to implement single image haze removal.

Although the prior may be invalid in some extreme cases, for instance, when the scene object is essentially similar to the light on a large local area and there is no shadow on the object, the prior can restore a high quality hazy-free image and produce a good depth map. With the development of state of art, X Liu [9] proposed an effective multi-scale correlation wavelet method to implement image dehazing and denoising,simultaneously in the frequency domain instead of spatial domain. Such model took the lead in using the relationship of wavelet coefficients to simultaneously eliminate haze and enhance texture.

Nowadays, with the development of machine learning, researchers proposed many neural network methods to implement image dehazing. Ren [10] and his group presented a multi-scale CNN(MSCNN) model which try to return directly to the intermediate transmission image or the final clear image, and obtain excellent performance and robustness. X Qin [11]presented an end to end feature fusion attention network (FFA-Net) to directly implement single image dehazing. In this model, not only can all features be merged, but also different weights of different levels of feature information can be learned adaptively. Also, D Chen [12] presented an end-to-end gated context aggregation network (GCANet) to restore hazy image. This model used the latest smooth dilation technology to eliminate mesh artifacts in dilated convolution, improving the performance of hazy removal.

In this project, we will use the open dark channel prior(ODC), FFA-Net, and GCANet to implement image hazy removal. We will compare the performance in the experiment and adopt all the advantages from such three models.

III. Method

This section will overview total three methods we used to implement haze removal and one is traditional but efficient model and the other two are based on deep learning. Also, some evaluation standards such as PSNR, SSIM will be mentioned.

A. Open dark channel model

The first model we selected is open dark channel prior model, which can implement image dehazing in low frequency part and image denoising in high frequency part, synchronously.

1) Haze removal in low frequency part: First of all, a basic atmospheric scattering model [13] will be introduced which is wildly used in computer vision:

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

Where I is the blurred image observed, J is the radiance of the scene to be restored, x indicates the pixel position, t(x) is medium transmission, which describes the part of light that is not scattered and reaches the camera, and A is the global atmospheric light. For haze removal, we aim to obtain the scene radiance J(x) from I(x). The J(x) will be well restored from the hazy image once we estimate the medium transmission t(x) and global atmospheric light A:

$$J(\mathrm{x}) = (I(\mathrm{x}) - A)/\max(t(\mathrm{x}), \sigma) + A \tag{2}$$

Where σ is the lower limit used to limit the transmission whose default value is 0.1. Assuming that the overall atmospheric light A is given, we can get the following constraints on t(x):

$$0 \leqslant t_b(\mathrm{x}) \leqslant t(\mathrm{x}) \leqslant 1 \tag{3}$$

$$t_b(\mathrm{x}) = \min\left\{1 - \min_c\left\{\frac{I^c(\mathrm{x})}{A^c}, \frac{\tilde{I}^c(\mathrm{x})}{\tilde{A}^c}\right\}, 1\right\} \tag{4}$$

where $t_b(x)$ is the lower bound of t(x), $\tilde{I}^c = 1 - I^c$ is the inverse image of I^c. Thus, the haze removal problem can be transferred as such equation:

$$\arg\min_t \Phi(J(\mathrm{x})), \quad \text{s.t.} \begin{cases} t_b(\mathrm{x}) \leq t(\mathrm{x}) \leq 1 \\ J(\mathrm{x}) = (I(\mathrm{x}) - A)/t(\mathrm{x}) + A \end{cases} \tag{5}$$

However, such method(DCP) without inherent boundary constraints usually overestimates the thickness of the haze, which cannot handle very bright areas, such as the sky area. Also, the best global atmospheric light A is darker than the brightest pixels in the image in most cases. In order to solve these problems, inherent boundary constraints and adopted morphological opening operations are added in this model to accurately estimate the transmission map:

$$\tilde{t}(\mathrm{x}) = 1 - \underset{B,\mathrm{y}\in\Omega_\mathrm{x}}{\mathrm{open}}\left(\min_c\left\{\frac{I^c(\mathrm{y})}{A^c}, \frac{\tilde{I}^c(\mathrm{y})}{\tilde{A}^c}\right\}\right) \tag{6}$$

$$\underset{B}{\mathrm{open}}(I) = \max\{\min\{I(x-s, y-t) - B(s,t)\} + B(s,t)\} \tag{7}$$

where B(s, t) is the structuring element whose size is s × t. With open operation, the model will be able to delete small bright details and impose spatial continuity which will produce an accurate scene transmission. Typical example is shown in Fig. 2.

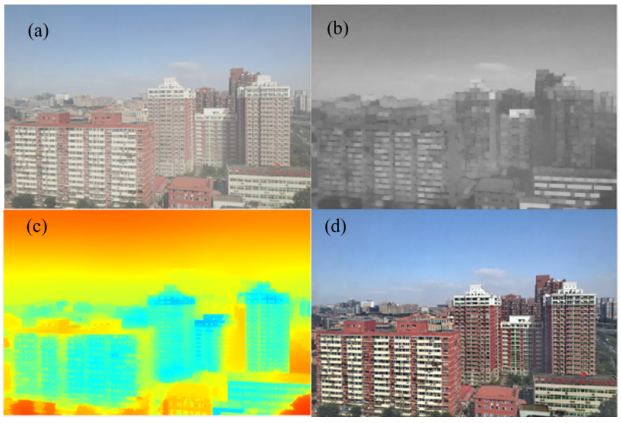


Fig. 2. (a) Input (b) Dark channel (c) Transmission map (d) Output

2) Texture enhancement in high frequency part: This model decompose the hazy image into low frequency part and high frequency part and the image noise is always left at the high frequency. Thus, this noise can be effectively eliminated by directly increasing the soft threshold operation [14] $\mathcal{S}_\tau$ in the high frequency part:

$$\mathcal{S}_\tau(\mathrm{x}) = \mathrm{sign}(\mathrm{x}) \cdot (|\mathrm{x}| - \tau) \tag{8}$$

Where τ is the threshold, usually estimated from the median of the first-level decomposition. We can obtain a local linear model from atmospheric scattering model:

$$\nabla I_\Omega = \{\partial_x I_\Omega, \partial_y I_\Omega, \partial_y \partial_x I_\Omega\} = t_\Omega \cdot \{\partial_x J_\Omega, \partial_y J_\Omega, \partial_y \partial_x J_\Omega\} \tag{9}$$

$$\begin{cases} \partial_x I_\Omega = t_\Omega \cdot \partial_x J_\Omega & \text{horizontal} \\ \partial_y I_\Omega = t_\Omega \cdot \partial_y J_\Omega & \text{vertical} \\ \partial_y \partial_x I_\Omega = t_\Omega \cdot \partial_y \partial_x J_\Omega & \text{diagonal.} \end{cases} \tag{10}$$

where Ω is a local path with small size. Since using wavelet decomposition here, the high frequency elements can be regarded as horizontal, vertical and diagonal gradients in the equation. Using the above gradient, we can adaptively correlate the wavelet coefficients between the recovered low-frequency part and the high-frequency part. Typical example is shown in Fig. 3.

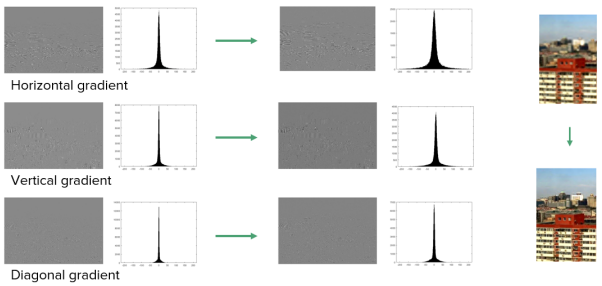


Fig. 3. Correlation wavelet coefficients for adaptive high-frequency texture enhancement

B. GCANet

GCANet is an end-to-end gated context aggregation network for image dehazing.[12] It first encodes the hazy input image into feature maps by the encoder, and then aggregates the context information and fuses the different-level features using smoothed dilated convolution and an extra gate sub-network. Next, the enhanced feature maps will be decoded into output dehazed image. The overall architecture is shown in Fig. 4.

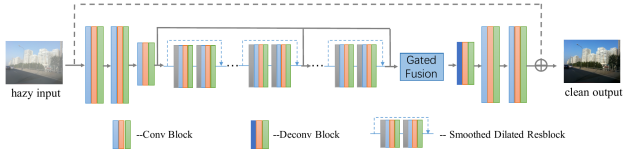


Fig. 4. GCANet structure

1) Smoothed Dilated Convolution: The dilated convolution layer can be regarded as a convolution with a dilated filter. The traditional convolution layer w with kernel k can been written as:

$$(f \otimes w)(i) = \sum_{j=1}^{k} f[i+j]w[j] \tag{11}$$

However, the dilated convolution layer can be written as :

$$(f \otimes_r w)(i) = \sum_{j=1}^{k} f[i+r*j]w[j] \tag{12}$$

where r is the dilation rate. In this way, the dilated convolution elevates the receptive field from k to $r(k-1)+1$ while preserving the resolution.

2) Gated Fusion Sub-network: In GCANet, the feature maps are extracted from three different levels and fed into gated fusion sub-network respectively. Three output importance weights are related to those feature levels. Finally, the feature maps are combined with those importance weights. The combination can be written as:

$$F_o = \mathcal{M}_l * F_l + \mathcal{M}_m * F_m + \mathcal{M}_h * F_h \tag{13}$$

where $\mathcal{M}_l, \mathcal{M}_m, \mathcal{M}_h$ are the regressed importance weights,and F_l, F_m, F_h are the feature maps in different levels. The F_o is the output feature map, which is fed into the decoder.

C. FFA-Net

FFA-Net is an end-to-end feature fusion attention network for single image dehazing.[11] In the FFA-Net, the hazy image is passed into a shallow feature extraction net, followed by a N Group Architectures with multiple skip connections. Then the output features are fused together through the Feature Attention Module, which is a new attention structure. Next the fused features are reconstructed by a set of layers similar to the feature extraction part, followed by a global residual learning structure. The overall architecture is shown in Fig. 5.

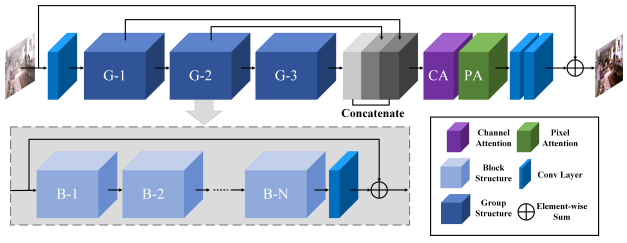


Fig. 5. FFA-Net structure

The feature attention module combines the channel attention and pixel attention to get a more effective fusion of features.

Firstly, the channel attention take a channel-wise global spatial information into a channel descriptor with a global average pooling:

$$g_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_c(i,j) \tag{14}$$

where $X_c(i,j)$ is the value of c-th channel at position (i,j). H_P stands for global pooling function that changes the shape of feature map from $C \times H \times W$ into $C \times 1 \times 1$. The the features are passed through a convolution layer followed a ReLu activation layer, then another convolution layer followed by a sigmoid activation layer to get the weights on each channel. The weights will be element-wise multiplied to the input, and the output shape is still $C \times H \times W$.

Next, the pixel attention is aimed to handle the unevenly distributed haze. It plays a role similar to the transmission matrix in the open dark channel method. The structure is identical to the rear part of channel attention, which is convolution-ReLu-convolution-softmax. Pixel attention receives the output of channel attention whose shape is $C \times H \times W$, and generate a weight matrix of $1 \times H \times W$. Like the channel attention, the weight matrix is element-wise multiplied to the input, and the output shape is still $C \times H \times W$.

IV. Results

A. Dataset

We consider the following two datasets:

1) SOTS: Synthetic Objective Testing Set (SOTS) is part of RESIDE testing set proposed in 2019. It consists of 50 indoor synthesized hazing images and 500 outdoor synthesized hazing images.

2) I-HAZE and O-HAZE: I-HAZE is a new dataset that contains 35 image pairs of hazy and corresponding haze-free (ground-truth) indoor images. O-HAZE contains 45 different outdoor scenes depicting the same visual content recorded in haze-free and hazy conditions, under the same illumination parameters. The haze in the image was generated by a professional haze machine while the lighting condition are controlled to be identical. They were first used in NTIRE 2018 image dehazing challenge.

B. Filter Selection for ODC

We have tested different filters and different levels for open dark channel method. The filters we used include: Haar, db10, db45, coif1, coif2, sym2, sym4, fk4, dmey, bior1.1 and rbio1.1. We also tried different filter levels from 2 to 4, and evaluated the PSNR and SSIM for each combination. The average result over all datasets is shown in Table I, and the best 10 combinations are shown in Fig. 6.
We find that the 3-level sym4 filter got the relative higher average SSIM over other combinations. Although dmey has slightly higher PSNR, the SSIM is a more comprehensive metric, so we preserve the setup with highest SSIM.

C. Model Comparison

We have tested all three models, open-dark-channel, GCA and FFA, over I-HAZE, O-HAZE, SOTS-indoor and SOTS-oudoor. The result is shown in Table II. The distribution of PSNR and SSIM on different datasets are shown in Fig. 7. We also showed several sample images to evaluate the intuitive dehazing performance in Fig. 8.

TABLE I
AVERAGE PERFORMANCE AMONG DATASETS

Combination		Average Metrics	
Level	**Filter**	***PSNR***	***SSIM***
3	sym4	16.43753	0.75636
4	sym4	16.31351	0.75419
2	sym4	16.26881	0.75386
3	coif1	16.40684	0.75353
3	sym2	16.38614	0.75353
3	dmey	16.44058	0.75348
2	coif2	16.24264	0.75268
2	dmey	16.28358	0.75248
3	coif2	16.37335	0.75244
2	coif1	16.22535	0.75185
2	sym2	16.23191	0.75182
4	sym2	16.28019	0.75092
2	db10	16.22561	0.75041
4	coif1	16.27202	0.74967
2	fk4	16.17851	0.74867
3	fk4	16.28869	0.74839
3	rbio1.1	16.35290	0.74760
3	haar	16.35290	0.74760
3	bior1.1	16.35290	0.74760
2	rbio1.1	16.15401	0.74688
2	haar	16.15401	0.74688
2	bior1.1	16.15401	0.74688
3	db10	16.25327	0.74687
4	coif2	16.20523	0.74458
4	rbio1.1	16.19419	0.74362
4	haar	16.19419	0.74362
4	bior1.1	16.19419	0.74362
4	fk4	16.09905	0.74297
4	dmey	16.09799	0.73889
4	db10	15.92963	0.72970
2	db45	16.00228	0.72869
3	db45	15.94791	0.71352
4	db45	15.61004	0.70019

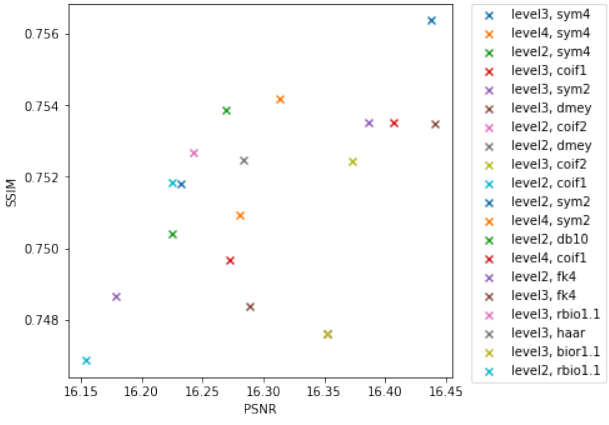
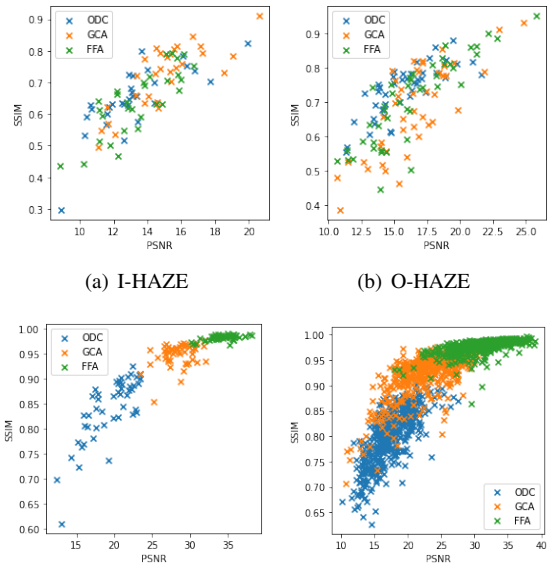


Fig. 6. Average Metrics for top 16 combinations

TABLE II
AVERAGE PERFORMANCE AMONG SOTS-INDOOR

Dataset	Models	Metrics	
		PSNR	*SSIM*
SOTS-indoor	ODC	19.48819	0.84175
	GCA	23.43391	0.90934
	FFA	30.37756	0.96920
SOTS-outdoor	ODC	19.48819	0.84175
	GCA	19.71648	0.93665
	FFA	30.39279	0.98214
I-HAZE	ODC	13.22065	0.65035
	GCA	15.34912	0.79141
	FFA	12.20358	0.67127
O-HAZE	ODC	15.65713	0.73335
	GCA	18.76480	0.80525
	FFA	16.40884	0.78238



(a) I-HAZE (b) O-HAZE

(c) SOTS-indoor (d) SOTS-outdoor

Fig. 7. PSNR and SSIM distribution for datasets

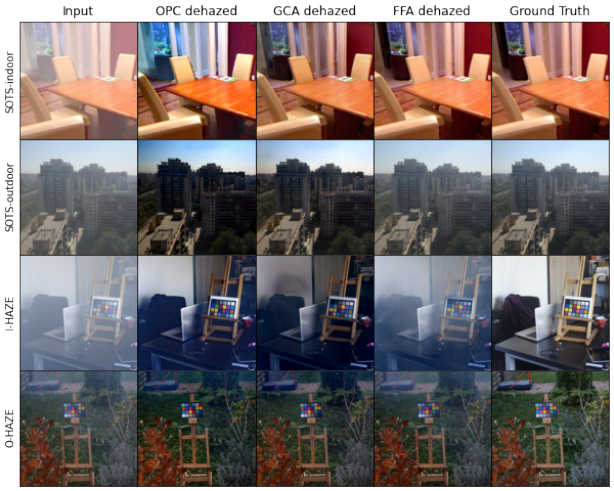


Fig. 8. Intuitive comparison

V. CONCLUSION

In this project, we selected three typical dehazing models to implment haze removal and made comparison based on their performance. ODC can achieve low frequency haze removal in the low frequency part, and remove the noise and enhance the texture detail adaptively in the high frequency parts, synchronously. GCA net can both preserve the original brightness and remove the haze as much as possible from the input, which exhibit best PSNR and SSIM performance when dealing with I-HAZE and O-HAZE dataset. FFA is obviously superior in the performance of image details and fidelity, which exhibit best PSNR and SSIM performance when dealing with SOTS dataset. However the performance in real-world datasets are much worse, which shows the drawback of fine-tuned Neural-Net's overfitting problems. To sum up, GCA exhibits best performance dealing with images with unevenly distributed haze.

VI. TEAM MEMBER CONTRIBUTION

Xiao Lu : implement GCANet and FFA-Net model
Sai Ding: implement ODC model
Percentage of contribution: 50%&50%

REFERENCES

[1] Juha Katajamäki. "Methods for gamma invariant colour image processing". In: *Image and Vision Computing* 21.6 (2003), pp. 527–542. ISSN: 0262-8856. DOI: https://doi.org/10.1016/S0262-8856(03)00033-7. URL: http://www.sciencedirect.com/science/article/pii/S0262885603000337.

[2] Hui Zhu, Francis H.Y. Chan, and F.K. Lam. "Image Contrast Enhancement by Constrained Local Histogram Equalization". In: *Computer Vision and Image Understanding* 73.2 (1999), pp. 281–290. ISSN: 1077-3142. DOI: https://doi.org/10.1006/cviu.1998.0723. URL: http://www.sciencedirect.com/science/article/pii/S1077314298907238.

[3] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. "Instant dehazing of images using polarization". In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. Vol. 1. 2001, pp. I–I. DOI: 10.1109/CVPR.2001.990493.

[4] S. G. Narasimhan and S. K. Nayar. "Chromatic framework for vision in bad weather". In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*. Vol. 1. 2000, 598–605 vol.1. DOI: 10.1109/CVPR.2000.855874.

[5] S. Narasimhan and S. Nayar. "Interactive (De) Weathering of an Image using Physical Models ". In: 2003.

[6] R. T. Tan. "Visibility in bad weather from a single image". In: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. 2008, pp. 1–8. DOI: 10.1109/CVPR.2008.4587643.

[7] Raanan Fattal. “Single Image Dehazing”. In: *ACM Trans. Graph.* 27.3 (Aug. 2008), pp. 1–9. ISSN: 0730-0301. DOI: 10.1145/1360612.1360671. URL: https://doi.org/10.1145/1360612.1360671.

[8] K. He, J. Sun, and X. Tang. “Single Image Haze Removal Using Dark Channel Prior”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.12 (2011), pp. 2341–2353. DOI: 10.1109/TPAMI.2010.168.

[9] Xin Liu et al. “Efficient single image dehazing and denoising: An efficient multi-scale correlated wavelet approach”. In: *Computer Vision and Image Understanding* 162 (2017), pp. 23–33. ISSN: 1077-3142. DOI: https://doi.org/10.1016/j.cviu.2017.08.002. URL: http://www.sciencedirect.com/science/article/pii/S1077314217301431.

[10] W. Ren et al. “Gated Fusion Network for Single Image Dehazing”. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018, pp. 3253–3261. DOI: 10.1109/CVPR.2018.00343.

[11] Xu Qin et al. *FFA-Net: Feature Fusion Attention Network for Single Image Dehazing*. Nov. 2019.

[12] D. Chen et al. “Gated Context Aggregation Network for Image Dehazing and Deraining”. In: *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2019, pp. 1375–1383. DOI: 10.1109/WACV.2019.00151.

[13] S. G. Narasimhan and S. K. Nayar. “Contrast restoration of weather degraded images”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.6 (2003), pp. 713–724. DOI: 10.1109/TPAMI.2003.1201821.

[14] D. L. Donoho. “De-noising by soft-thresholding”. In: *IEEE Transactions on Information Theory* 41.3 (1995), pp. 613–627. DOI: 10.1109/18.382009.