A SINGLE IMAGE DEHAZING DATASET WITH LOW-LIGHT REAL-WORLD INDOOR IMAGES, DEPTH MAPS AND INFRARED IMAGES

A. Filin[a,][*]A. Kopylov[a], I. Gracheva [a]

[a] Laboratory of Cognitive Technologies and Simulation Systems, Tula State University, 92, Lenin ave., Tula, 300012, Russia - (andrewifilin, andkopylov)@gmail.com

Commission II, WG II/8

KEY WORDS: Single image dehazing, localized light sources, low-light conditions, real-world dehazing, dataset, depth map, infrared images, thermal images

ABSTRACT:

Benchmarking of haze removal methods and training related models requires appropriate datasets. The most objective metrics of assessment quality of dehazing are shown by reference metrics – i.e. those in which the reconstructed image is compared with the reference (ground-truth) image without haze. The dehazing datasets consist of pairs where haze is artificially synthesized on ground-truth images are not well suited for the assessment of the quality of dehazing methods. Accommodation of the real-world environment for take truthful pairs of hazy and haze-free images are difficult, so there are few image dehazing datasets, which consists with the real both hazy and haze-free images. The currently researcher's attention is shifting to dehazing on "more complex" images, including those that are obtained in insufficient illumination conditions and with the presence of localized light s ources. It is almost no datasets with such pairs of images, which makes it difficult of objective assessment of image dehazing m ethods. In this paper, we present extended version of our previously proposed dataset of this kind with more haze density levels and depths of scenes. It consists of images of 2 scenes at 4 lighting and 8 haze density levels - 64 frames in total. In addition to images in the visible spectrum, for each frame depth map and thermal image was captured. An experimental evaluation of state-of-the art haze removal methods was carried out on the resulting dataset. The dataset is available for free download at `https://data.mendeley.com/datasets/jjpcj7fy6t`.

1. INTRODUCTION

Currently, the use of intelligent systems based on video analysis, such as automatic navigation systems, traffic monitoring, outdoor video surveillance systems, etc., face a number of obstacles when operating under real weather conditions. In particular, the presence of haze, dust, various kinds of suspensions, rain and snow significantly complicates analysis of scenes and detection of objects. As a result, haze removal has received increasing attention from researchers, and many new techniques have been developed.

An objective comparison of haze removal methods and training-related models requires appropriate datasets for benchmarking. Different approaches to assessment the quality of haze removal put forward different requirements for the evaluating datasets. There are full-reference and non-reference metrics of image quality assessment. Full-reference metrics require a haze-free (ground truth) image for each hazy image, which is often impossible to provide, since the ground truth image is one that was obtained under constant environmental conditions, similar to those present in the corresponding hazy image, except for the absence of haze.

In cases where the dataset consists only of hazy images without corresponding haze-free (ground truth) images, benchmarking of dehazing methods can be performed using non-reference (NR) metrics for image quality assessment (IQA). An example of such a metric is the one, proposed by (Mittal et al., 2012) which employs scene statistics of locally normalized luminance coefficients to measure the loss of "naturalness" in images caused by distortions. This approach provides a comprehensive measure of quality and has low computational complexity, making it suitable for real-time applications.

In addition to NR-IQA metrics, in which image quality is evaluated using mathematical expressions, there are also machine learning-based metrics that attempt to evaluate image quality in the same manner as a person does. For example, (Talebi and Milanfar, 2018) uses a model based on a convolutional neural network architecture (CNN), which provide the possibility of prediction of both the technical and aesthetic aspects of images. To achieve higher correlation between human and model estimations of dehazing quality, rather than simply classifying or regressing images scores, the model predicts the distribution of ratings as a histogram.

The most popular and acceptable by community metrics for assessment quality of dehazing are full-reference metrics, such as PSNR and SSIM (Wang et al., 2004). These methods compare resulting images of the haze removal method and its ground truth image, but in different manner: PSNR is based on calculating direct differences between the corresponding pixels of the reference and tested images, while SSIM takes into account the relationship between pixels, which allows one to express a change in the structure of the image, and thus give a quality assessment that is closer to human perception. In addition, the range of SSIM values lies in the range [-1,1], where 1 corresponds to a comparison of identical images, which simplifies the perception and interpretation of the results.

Thus, a key property that datasets for the image haze removal task should provide is the inclusion of pairs of images of the same scene with and without haze, while keeping other external conditions unchanged (the location of objects on the scene, lighting, etc.). Since it is difficult to accommodate the real-world environment to ensure such property, the vast majority of dehazing datasets consist of pairs of images in which haze is artificially synthesized on ground-truth images using known depth maps and an atmospheric scattering model (usually, the Koschmieder's clas-

---
[*]Corresponding author.

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W3-2023
ISPRS Intl. Workshop "Photogrammetric and computer vision techniques for environmental and infraStructure monitoring, Biometrics and Biomedicine"
PSBB23, 24–26 April 2023, Moscow, Russia

sical model (Koschmieder, 1924)):

$$I_c(s) = J_c(s)T(s) + A_c(1{-}T(s)) \qquad (1)$$

where $S = s = (s_1, s_2) : s_1 = 1, ..., N_1, s_2 = 1, ..., N_2$ is the discrete pixels grid, $I_c(s)$ is the hazy image intensity value at the position $s\epsilon S$ and color channel $c\epsilon r, g, b$, $J_c(s)$ is the corresponding haze-free image intensity value at the same position, and the color channel c, A_c is the airlight of the color channel c and $T(s)$ is the transmission map.

It is commonly accepted to use a constant value of the airlight A_c over the entire image, a homogeneous distribution of haze particles, and independence of the scattering coefficient from the wavelength.

The last two assumptions allow us to express the medium transmission map as follows:

$$T(s) = e^{-\beta d(s)} \qquad (2)$$

where $d(s)$ is the scene depth at the position s and β is the scattering coefficient.

Thus, substituting the known depth map in d(s) of (2), as well as the corresponding haze-free (ground truth) image J(s) in expression (1), we can obtain the corresponding hazy image.

The advantage of obtaining full-reference datasets for the single image haze removal task using this method is that non-specialized datasets can be used, only ground truth images and depth maps are required. For example, the use of NYU Depth v2 (Nathan Silberman Derek Hoiem and Fergus, 2012) and Middleburry (Scharstein et al., 2014) for obtaining reference datasets for performing experiments or training models in the image haze removal task by synthesizing haze on the ground truth images using known depth map, was previously widespread. Later, to simplify and standardize the experimental evaluation and models training for image haze removal tasks, the NYU2, Middlebury, and other datasets were combined into one set, which included ready-made images with generated haze. The resulting set was named REalistic Single Image DEhazing (RESIDE) (Li et al., 2018) it contains subsets of data for both training and validation of models for image haze removal tasks. In total, the dataset has collected about 430,000 images from various sources.

As noted earlier, obtaining a reference dataset consisting of images with real haze is difficult. This is because it is necessary to obtain both images (ground truth and hazy) of the scene under unchanged environmental conditions, except for the presence/absence of haze. The appearance or dissipation of haze takes time, during which the scene often changes, and the second image from a pair can no longer be considered a reference to the first image. In this regard, all open sets of this kind have a small volume. So, the total number of pairs of hazy/ground truth images, where both images of the pair are captured in the real-world environment (Ancuti et al., 2018a, Ancuti et al., 2018b, Ancuti et al., 2019, Ancuti et al., 2020, Khoury et al., 2018), is about 190.

As we have seen, the open datasets available for single image haze removal tasks, where pairs of hazy/ground-truth images were obtained by photographing real scenes, have a small volume. In contrast, datasets where hazy images were synthesized have a large volume. Generally, the number of pairs with real images is orders of magnitude lower than those with synthesized ones.

Datasets with synthesized haze are actively employed to train models in machine learning-based haze removal methods. However, they are not well suited for the assessment of the quality of dehazing methods since, most commonly, the depth map has visible inaccuracies and glitches that affect the generated haze. Additionally, the physical model of atmospheric scattering does not fully reflect the complexity of the underlying processes that occur when light passes through haze. Figure 1 shows examples of synthesized images with noticeable misses in the overlaid haze, which occurred due to the inaccurate depth map.

(a)   (b)

Figure 1: An example of the image from SOTS (Li et al., 2018) dataset with synthesized haze (b). (a) is the original (ground truth) image. At (b) the wrong overlay of haze at the back of the front (left) chair can be seen.

Another important fact is that modern haze removal methods demonstrate high-quality dehazing of images with sufficient (daylight) illumination, so researchers' attention is shifting to dehazing "more complex" images, including those obtained in insufficient illumination conditions and with the presence of localized light sources - i.e., conditions that simulate night-time. Among the publicly available datasets, we did not find any that consist of pairs of real images obtained in low light conditions and with the presence of localized light sources. The proposed dataset has specific properties and allows for benchmarking of dehazing methods in terms of performance in night-time conditions.

Section 2 describes the equipment and methodology used for acquiring the dataset. Afterwards, we will present the parameters of the collected dataset and some examples from it.

Section 3 provides a description of the haze removal methods used in the experiments, as well as the results of dehazing on the proposed dataset and some other datasets.

Section 4 draws conclusions about the collected dataset and experimental results.

2. THE PROPOSED DATASET

Previously (Filin et al., 2022), we proposed the single image haze removal dataset that consists entirely of real images, including images taken at low light conditions and with the presence of localized light sources.

Two scenes were prepared with 4 degrees of illumination and 4 degrees of haze density, 32 frame variations in total. One of the scenes includes localized light sources. Several shortcomings were revealed during the analysis of the experimental research on this dataset including the small depth of scenes, a small number of haze density levels, and varying camera exposure during shooting.

In this work, we have approximately doubled the depth of the scene. The number of variations of haze density levels was increased from 4 to 8, and 4 illumination levels were adjusted evenly

such that objects on the scene were visible both with maximum and minimum illumination without changing the camera's settings.

We keep here the same principles of dataset collection as before:
• The most straightforward way to achieve a hazy environment is by using a smoke machine. It generates particles, closed by its size to the particles of atmospheric haze.
• A fan can be used to speed up a homogeneous distribution of haze particles. But it should be turned off for a few minutes before capturing images to make the particles slow down.
• ColorChecker should be placed in front of the camera so it will fully fit the image. Optionally, other camera calibration and support tools can be placed similarly.
• The camera parameters (aperture, shutter speed, ISO) should be controlled.

In this study, as in the previous work, 2 scenes were set up with objects of varying sizes, shapes, materials, and the presence or absence of localized light sources. Equipment for setting up camera settings and post-processing images, such as SpyderLensCal for precise focusing, Datacolor SpyderCube for color correction, and ISO 12233 test chart, were also placed in front of the camera.

For each scene, 32 frames were captured - with 4 degrees of illumination and 8 degrees of haze density. The illumination was adjusted by regulating the number of lighting lamps. The minimum illumination was set so that the objects in the scene remained visible without changing camera settings, such as ISO, aperture, and shutter speed. The accepted settings were ISO = 800, aperture = f/5, and shutter speed = 1/30s.

To generate haze, a haze machine Involight FM900 was used. After placing the objects and setting up the equipment, 4 shots were taken with varying degrees of illumination to all available equipment one by one, which took about 10 seconds in total for each cycle.

After the ground truth images were taken, the haze was generated for 15 minutes, and the fan was turned on to distribute it evenly. Then, the fan and hazer were then turned off, and a waiting period of 30 seconds was made for the haze particles slowed down. After the waiting period, a series of shots were taken with varying degrees of illumination, using all available equipment.

To obtain varying degrees of haze intensity, a waiting period of 3 minutes was made for the haze to dissipate before the next level of haze intensity was captured in a series of four shots. 6 more similar cycles were carried out to obtain a total of 8 degrees of haze variation for each scene, including one haze-free (ground truth) and 7 hazy images with varying intensity.

After capturing images for all combinations of illumination and haze intensity for the first scene, the remaining haze dissipated, and the next scene was formed. The same process of capturing images with varying degrees of illumination and haze intensity was repeated for the second scene.

A total of 64 frames were obtained - 2 scenes with 4 light and 8 haze density levels (1 ground-truth and 7 with haze). Each frame was shot with the camera (both on Canon 2000d and Intel RealSense d435i), the depth camera (Intel RealSense d435i), and the thermal imager (Flir C2).

Figure 2 shows examples of images in the visible spectrum with changes in illumination and haze densities. Figure 3 shows examples of depth maps. Figure 4 shows examples of images taken with a thermal imager.
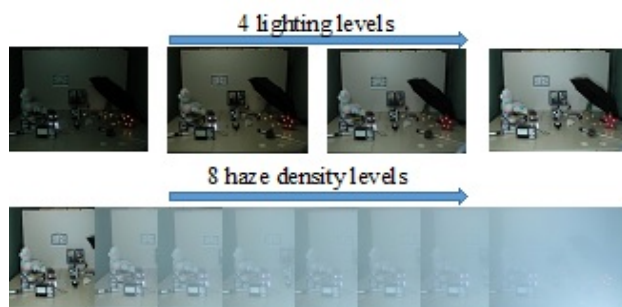

Figure 2: Examples of images from the proposed dataset. Shown changing illumination and haze density levels.
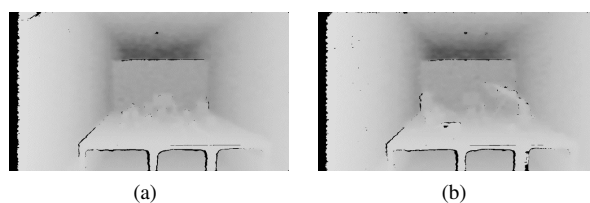

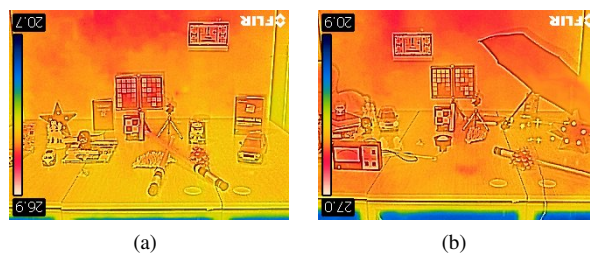Figure 3: Examples of depth maps from the proposed dataset for scenes 1 (a) and 2 (b).

Figure 4: Infrared images from the proposed dataset for scenes 1 (a) and 2 (b).

3. EXPERIMENTAL RESEARCH

Experimental research was performed using several state of the art image haze removal methods. Some methods attempt to remove haze from images by evaluating the transmission map and atmospheric light and applying them in an atmospheric scattering model. He et al. (He et al., 2011) has formulated a widely used dark channel prior method that allows direct haze quantification for reconstruction of the haze-free image. The method is based on the discovered pattern that in local areas where there is no haze, at least one channel in the RGB color space contains pixels with low intensity. Later, Berman et al. (Berman et al., 2016) discovered that the colors of an image without haze can be well approximated by several hundred different colors that form dense clusters in RGB space. In a hazy image, each color cluster forms a line in RGB space that can be used to reconstruct the image. Zhu et al. (Zhu et al., 2015) proposed a linear model for depth estimation in a hazy image using a method based on the attenuation color prior effect. The model parameters were obtained by supervised learning. The method proposed by Dhara et al. (Dhara et al., 2020) separates hazy images into those with some hue and those without it based on an estimate of the range of color tones in the image pixels. Color correction is then performed using a non-linear transformation, followed by tone-dependent atmospheric lighting refinement.

Recently, considerable research attention has been directed to

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W3-2023
ISPRS Intl. Workshop "Photogrammetric and computer vision techniques for environmental and infraStructure monitoring, Biometrics and Biomedicine"
PSBB23, 24–26 April 2023, Moscow, Russia

machine learning-based haze removal methods. In (Gui et al., 2022) an extensive review of such methods is given. According to this article, the Qin method (Qin et al., 2020) shows a leading position in the quality of haze removal in terms of PSNR and SSIM metrics. This method employs the attention-based deep neural network architecture, which was improved by passing shallow layers' information into deep layers – it is assumed that high and low-level features will allow the core network to find more patterns in the data, which will improve the quality of dehazing.

In this research, to perform experiments with the Qin method on the indoor and outdoor datasets, we used corresponding trained models, available at the following link: `https://github.com/zhilin007/FFA-Net/tree/master/net/trained_models`.

The experimental research was performed on several datasets, including images with synthesized haze (Li et al., 2018); images with real haze (Ancuti et al., 2018a, Ancuti et al., 2018b), and the proposed dataset. As full-reference metrics was used PSNR and SSIM. Quantitative dehazing results are shown in Table 1. Examples of dehazing results are show in Figure 5.

	Databases	Berman et al.	Dhara et al.	He et al.	Qin et al.	Zhu et al.
PSNR	i-haze+o-haze	15.76	14.85	13.51	15.16	16.58
	SOTS-ITS	17.28	19.61	16.56	29.58	19.05
	SOTS-OTS	17.96	16.62	14.40	19.48	22.05
	night-haze	15.76	18.59	17.42	19.37	17.65
	night-haze-ext	14.60	19.28	17.39	18.99	19.74
SSIM	i-haze+o-haze	0.74	0.67	0.62	0.65	0.70
	SOTS-ITS	0.78	0.86	0.80	0.97	0.81
	SOTS-OTS	0.83	0.81	0.75	0.84	0.89
	night-haze	0.73	0.71	0.49	0.74	0.62
	night-haze-ext	0.56	0.63	0.60	0.63	0.69

Table 1: Quantitative experimental results of the proposed and some other datasets.

As can be seen from the quantitative results from Table 1, the PSNR metric shows a greater variety of results depending on the dataset, which complicates their interpretation. The SSIM metric is more stable - it can be seen that on most methods the value of the metrics is lower on datasets consisting of images with real haze. In addition, on the presented dataset, the SSIM metric shows the lowest value on most methods. For the rest, the lowest metrics were obtained on the night-haze dataset (Filin et al.,

(a)

(b)

(c)

(d)

(e)

(f)
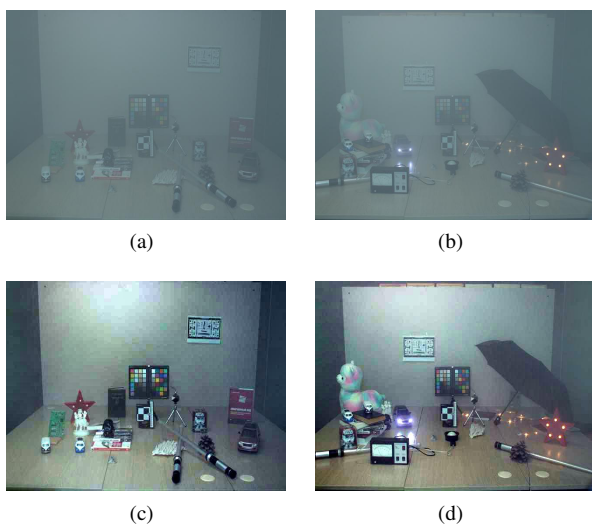

(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

Figure 5: Examples of experimental results utilizing image dehazing methods on the images from the proposed dataset (night-haze-ext) captured at the 4th level of haze density and the 2nd level of lighting. The left column shows images from the first scene, the right column displays images from the second scene. Input hazy images (a, b), dehazing results by methods Berman et al. (Berman et al., 2016) (c,d), Dhara et al. (Dhara et al., 2020) (e, f), He et al. (He et al., 2011) (g, h), Qin et al. (Qin et al., 2020) (i, j) and Zhu et al. (Zhu et al., 2015) (k, l), and also ground truth images (m, n) for scenes are provided at the corresponding columns.

2022), which is similar to the one presented. Such results were obtained because of the peculiarities of calculating the metrics, which can be explained by the results of the He et al. method. Probably, this method demonstrates high PSNR results on the night-haze and night-haze-ext (proposed) datasets, because they contain images obtained in low light conditions and with the pres-

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W3-2023
ISPRS Intl. Workshop "Photogrammetric and computer vision techniques for environmental and infraStructure monitoring, Biometrics and Biomedicine"
PSBB23, 24–26 April 2023, Moscow, Russia

ence of localized light sources. The method He et al. is not designed to work with such images because localized light sources are mistaken for bright atmospheric illumination, which leads to an overall darkening of the resulting image. Since the original image was obtained in low light conditions, the difference between the corresponding pixels of the resulting image and its ground truth will not be large, which will lead to a high PSNR metric. The value of the SSIM metric for such images will be small because the overall structure of the image has changed a lot as a result of such darkening.

4. CONCLUSIONS

This paper presents a dataset that has features that allow a more objective assessment of single image haze removal methods relative to real-life conditions because it contains images obtained in low light conditions and with the presence of localized light sources.

The experimental results show noticeably better values of PSNR and SSIM metrics on datasets in which image have synthesized haze than on datasets that consist of images with real haze. This may indicate that the atmospheric scattering model that was used to generate haze also underlies the methods used in the experiment, so the haze removal quality metrics obtained on sets consisting of real images are more objective.

In addition to images in the visible spectrum, the resulting dataset also includes infrared images and depth maps of scenes. The presence of additional modalities allows for the expansion of the scope of the dataset - for example, the presence of a depth map makes it possible to evaluate the accuracy of the calculated depth map and investigate transmission map and depth map relations.

ACKNOWLEDGEMENTS

REFERENCES

Ancuti, C., Ancuti, C. O., Timofte, R. and De Vleeschouwer, C., 2018a. I-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Indoor Images. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 11182 LNCS, Springer, Cham, pp. 620–631.

Ancuti, C. O., Ancuti, C. and Timofte, R., 2020. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE CVPR 2020, Vol. 2020-June, pp. 1798–1805.

Ancuti, C. O., Ancuti, C., Sbert, M. and Timofte, R., 2019. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In: 2019 IEEE international conference on image processing (ICIP), IEEE, pp. 1014–1018.

Ancuti, C. O., Ancuti, C., Timofte, R. and De Vleeschouwer, C., 2018b. O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Vol. 2018-June, pp. 867–875.

Berman, D., Avidan, S. and Others, 2016. Non-local image dehazing. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1674–1682.

Dhara, S. K., Roy, M., Sen, D. and Biswas, P. K., 2020. Color Cast Dependent Image Dehazing via Adaptive Airlight Refinement and Non-linear Color Balancing. IEEE Transactions on Circuits and Systems for Video Technology 8215(c), pp. 1–1.

Filin, A., Kopylov, A., Seredin, O. and Gracheva, I., 2022. Hazy images dataset with localized light sources for experimental evaluation of dehazing methods. In: The 6th International Workshop on Deep Learning in Computational Physics, p. 19.

Gui, J., Cong, X., Cao, Y., Ren, W., Zhang, J., Zhang, J., Cao, J. and Tao, D., 2022. A Comprehensive Survey and Taxonomy on Single Image Dehazing Based on Deep Learning. ACM Computing Surveys.

He, K., Sun, J. and Tang, X., 2011. Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(12), pp. 2341–2353.

Khoury, J. E., Thomas, J., Technol, A. M. J. I. S. and undefined 2018, 2018. A Database with Reference for Image Dehazing Evaluation. jbthomas.org 62(1), pp. 10503–10504.

Koschmieder, H., 1924. Theorie der horizontalen Sichtweite. Beitr. Phys. freie Atmos. 12, pp. 33–55.

Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W. and Wang, Z., 2018. Benchmarking single-image dehazing and beyond. IEEE Transactions on Image Processing 28(1), pp. 492–505.

Mittal, A., Moorthy, A. K. and Bovik, A. C., 2012. No-reference image quality assessment in the spatial domain. IEEE Transactions on image processing 21(12), pp. 4695–4708.

Nathan Silberman Derek Hoiem, P. K. and Fergus, R., 2012. Indoor Segmentation and Support Inference from RGBD Images. In: ECCV.

Qin, X., Wang, Z., Bai, Y., Xie, X. and Jia, H., 2020. FFA-Net: Feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34number 07, pp. 11908–11915.

Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X. and Westling, P., 2014. High-resolution stereo datasets with subpixel-accurate ground truth. In: Pattern Recognition: 36th German Conference, GCPR 2014, Münster, Germany, September 2-5, 2014, Proceedings 36, Springer, pp. 31–42.

Talebi, H. and Milanfar, P., 2018. NIMA: Neural image assessment. IEEE transactions on image processing 27(8), pp. 3998–4011.

Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13(4), pp. 600–612.

Zhu, Q., Mai, J. and Shao, L., 2015. A fast single image haze removal algorithm using color attenuation prior. IEEE transactions on image processing 24(11), pp. 3522–3533.