I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images

Cosmin Ancuti *, Codruta O. Ancuti †, Radu Timofte ‡ and Christophe De Vleeschouwer *

* ICTEAM, Universite Catholique de Louvain, Belgium
† MEO, Universitatea Politehnica Timisoara, Romania
‡ ETH Zurich, Switzerland and Merantix GmbH, Germany

Abstract. Image dehazing has become an important computational imaging topic in the recent years. However, due to the lack of ground truth images, the comparison of dehazing methods is not straightforward, nor objective. To overcome this issue we introduce I-HAZE, a new dataset that contains 35 image pairs of hazy and corresponding haze-free (ground-truth) indoor images. Different from most of the existing dehazing databases, hazy images have been generated using real haze produced by a professional haze machine. To ease color calibration and improve the assessment of dehazing algorithms, each scene includes a MacBeth color checker. Moreover, since the images are captured in a controlled environment, both haze-free and hazy images are captured under the same illumination conditions. This represents an important advantage of the I-HAZE dataset that allows us to objectively compare the existing image dehazing techniques using traditional image quality metrics such as PSNR and SSIM.

1 Introduction

Limited visibility and reduced contrast due to haze or fog conditions is a major issue that hinders the success of many outdoor computer vision and image processing algorithms. Consequently, automatic dehazing methods have been largely investigated. Oldest approaches rely on atmospheric cues [10,26], multiple images captured with polarization filters [27,30], or known depth [19,34]. Single image dehazing, meaning dehazing without side information related to the scene geometry or to the atmospheric conditions, is a complex mathematically ill-posed problem. This is because the degradation caused by haze is different for every pixel and depends on the distance between the scene point and the camera. This dependency is expressed in the transmission coefficients, that control the scene attenuation and amount of haze in every pixel. Due to lack of space, we refer the reader to previous papers for a formal description of a simplified but realistic light propagation model, combining transmission and airlight to describe how haze impacts the observed image.

Single image dehazing directly builds on this simplified model. It has been addressed recently [12,32,15,34,21,3,4,13,11,33], by considering different kinds

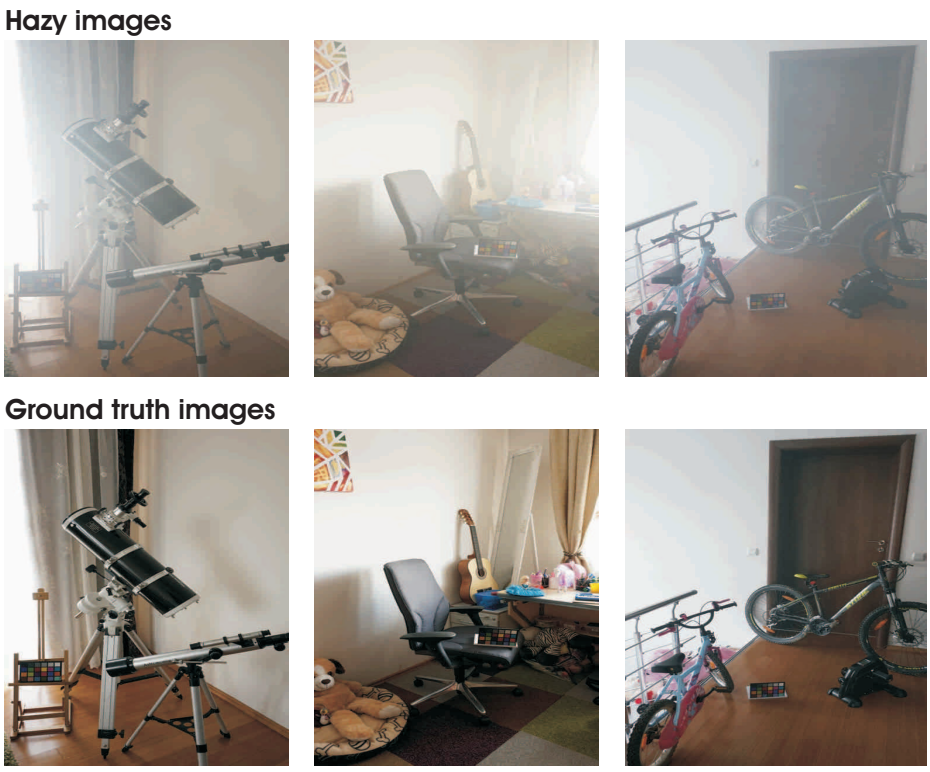

Fig. 1. *I-HAZE dataset provides 35 set of hazy indoor images and the corresponding ground truth (haze-free) images.*

of priors to estimate the transmission. The method of Fattal [12] adopts a refined image formation model that accounts for surface shading in addition to the transmission function. This allows to regularize the transmission and haze color estimation problems by searching for a solution in which the resulting shading and transmission functions are locally statistically uncorrelated. Tan's approach [32] assumes the atmospheric airlight to be the brightest pixel in the scene, and estimate the transmission by maximizing the contrast. The maximal contrast assumption has also been exploited in [34], with a linear complexity in the number of image pixels. The dark channel priors, introduced in [15], has been proven to be a really effective solution to estimate the transmission, and has motivated many recent approaches. Meng et al. [23] extends the work on dark channel, by regularizing the transmission around boundaries to mitigate its lack of resolution. Zhu et al. [40] extend [23] by considering a color attenuation prior, assuming that the depth can be estimated from pixel saturation and intensity. The color-lines, introduced in [13], also exploit the impact of haze over the color channels distribution. Berman et al. [5] adopt a similar path. They observe that the colors of a haze-free image are well approximated by a limited set of tight

clusters in the RGB space. In presence of haze, those clusters are spread along lines in the RGB space, as a function of the distance map, which allows to recover the haze free image. Ancuti et al. [3] rely on the hue channel analysis to identify hazy regions and enhance the image. For night-time dehazing and non-uniform lighting conditions, spatially varying airlight has been considered and estimated in [1,22]. Fusion-based single image dehazing approaches [4,9] have also achieved visually pleasant results without explicit transmission estimation and, more recently, several machine learning based methods have been introduced [33,6,28]. DehazeNet [6] takes a hazy image as input, and outputs its medium transmission map that is subsequently used to recover a haze-free image via atmospheric scattering model. It resorts to synthesized training data based on the physical haze formation model. Ren et al. [28] proposed a coarse-to-fine network consisting of a cascade of CNN layers, also trained with synthesized hazy images.

Despite this prolific set of dehazing algorithms, the validation and comparison of those methods remains largely unsatisfactory. Due to the absence of corresponding pairs of hazy and haze-free ground-truth image, most of the existing evaluation methods are based on non-reference image quality assessment (NR-IQA) strategies. For example, in [14], the assessment simply relies on the gradient of the visible edges. Other non-reference image quality assessment (NR-IQA) strategies [24,25,29] have been used for dehazing assessment. A more general framework has been introduced in [8], using subjective assessment of enhanced and original images captured in bad visibility conditions. Besides, the Fog Aware Density Evaluator (FADE) introduced in [9] predicts the visibility of a hazy/foggy scene from a single image without corresponding ground-truth. Unfortunately, due to the absence of the references (haze-free), none of these quality assessment approaches has been commonly accepted by the dehazing community.

Due to the practical issues associated to the recording of reference and hazy images under identical illumination condition, all existing data-sets have been built on synthesized hazy images based on the optical model and known depth. The work [35] presents the FRIDA dataset designed for Advanced Driver Assistance Systems (ADAS) that is a synthetic image database (computer graphics generated scenes) with 66 roads synthesized scenes. In [2], a dataset of 1400+ images of real complex scenes has been derived from the *Middelbury*[1] and the *NYU-Depth V2*[2] datasets. It contains high quality real scenes, and the depth map associated to each image has been used to yield synthesized hazy images based on Koschmieder's light propagation model [20]. This dataset has been recently extended in [39].

Complementary to the existing dehazing datasets, in this paper we contribute I-HAZE, a new dataset containing pairs of real hazy and corresponding haze-free images for 35 various indoor scenes. Haze has been generated with a professional haze machine that imitates with high fidelity real hazy conditions. A similar dataset (CHIC) that contains only two quite similar scenes was intro-

[1] http://vision.middlebury.edu/stereo/data/scenes2014/

[2] http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html

duce in [18]. However, I-HAZE dataset provides significantly more scenes with a larger variety of object structures and colors. Another contribution of this paper is a comprehensive evaluation of several state-of-the-art single image dehazing methods. Interestingly, our work reveals that many of the existing dehazing techniques are not able to accurately reconstruct the original image from its hazy version. This observation is founded on SSIM [36] and CIEDE2000 [31] image quality metrics, computed using the known reference and the dehazed results produced by different dehazing techniques. This observation, combined with the release of our dataset, certainly paves the way for improved dehazing methods.

2 I-HAZE Dataset

This section describes how the I-Haze dataset has been produced. All the 35 scenes presented in the I-Haze correspond to indoor domestic environments, with objects with different colors and specularities. Besides the domestic objects, all the scenes contain a color checker chart (Macbeth color checker). We use a classical Macbeth color checker with the size 11 by 8.25 inches with 24 squares of painted samples (4x6 grid).

After carefully setting each scene, we first record the ground truth (haze-free image) and then we immediately start the process of introducing haze in the scene. Therefore, we use two professional fog/haze machines (LSM1500 PRO 1500 W) that generate a dense vapor. The fog generators use cast or platen type aluminum heat exchangers, which causes evaporation of the water-based fog liquid. The generated particles (since are water droplets) have approximately the same diameter size of 1 - 10 microns as the atmospheric haze. Before shooting the hazy scene, we use a fan that helps to obtain in a relatively short period of time a homogenous haze distribution in the entire room (obviously that is kept isolated as much as possible by closing all the doors and windows). The entire process to generate haze took approximately 1 minute. Waiting approximately another 5-10 minutes, we obtain a homogenous distribution of the haze. The distances between the camera and the target objects range form 3 to 10 meters. The recordings were performed during the daytime in relatively short intervals (20-30 minutes per scene recording) with natural lightning and when the light remains relatively constant (either smooth cloudy days or when the sun beams did not hit directly the room windows).

To capture haze-free and hazy images, we used a setup that includes a tripod and a Sony A5000 camera that was remotely controlled (Sony RM-VPR1). We acquired JPG and ARW (RAW) 5456×3632 images, with 24 bit depth. The cameras were set on manual mode and we kept the camera still (on a tripod) over the entire shooting session of the scene. We calibrate the camera in haze-free scene, and then we kept the same parameters for the hazy scene. For each scene, the camera settings have been manually calibrated by adjusting manually the aperture (F-stop), shutter-speed (exposure-time), ISO speed and the white-balance. Setting the three parameters aperture-exposure-ISO has been realized using both the built-in light-meter of the camera and an external exponometer

Sekonic. For the white-balance we used the gray-card, targeting a middle gray (18% gray). The calibration process is straight-forward, since it just requires to set the white-balance in manual mode and to place the gray-card in front of the subject. In practice, we placed the gray-card in the center of the scene, two meters away from the camera. In addition, since each scene contains the color checker, the white-balance can be manually adjusted (a posteriori) using specialized software such as Adobe Photoshop Lightroom.

3 Quantitative Evaluation

In this section, before presenting the dehazing techniques used in our evaluation we briefly discuss the optical model assumed in most of the existing dehazing techniques.

Mathematically the dehazing optical model is expressed by the image formation model of Koschmieder [20]. Based on this model, due to the atmospheric particles that absorb and scatter light, only a certain percentage of the reflected light reaches the observer. The light intensity $\mathcal{I}$ of each pixel coordinate x, that passes a hazy medium is expressed as:

$$\mathcal{I}(x) = \mathcal{J}(x)\ T(x) + A_\infty\ [1 - T(x)] \tag{1}$$

where the haze-free image is denoted by $\mathcal{J}$, T is the *transmission* (depth map) and A_∞ is the atmospheric light.

The atmospheric light A_∞ is a color constant and represents the principal source of the additive color shifting that distorts the hazy images. Assuming an homogeneous medium [20] the transmission map T can be expressed as:

$$T(x) = e^{[-\beta\ \ d(x)]} \tag{2}$$

where β is a medium coefficient due to the light scattering and d is the distance between the sensor and the target scene.

He et al. [15,16] introduced the popular dark-channel prior, an extension of the dark object assumption [7]. Their proposed dehazing algorithm exploits the observation that the majority of local regions (except the sky, or hazy regions) include some pixels that are characterized by a very low value in at least one color channel. This helps in roughly estimating the transmission map of the hazy images. Further refinement of the transmission map can be obtained based on an alpha matting strategy [15], or by using guided filters [16]. In our assessment, we employ the dark channel prior refined based on the guided filter [17].

Meng et al. [23] extend the dark channel prior [15]. It estimates the transmission map by formulating an optimization problem that embeds the constraints imposed by the scene radiance and a weighted L1-norm based contextual regularization, to avoid halo artifacts around sharp edges.

Fattal [13] introduced a method that exploits the observation that pixels of

small image patches typically exhibit a one-dimensional distribution in RGB color space, named color lines. Since the haze tends to move color lines away from the RGB origin, an initial estimation of the transmission map is obtained by computing the lines offset from the origin. The final transmission map is refined by applying a Markov random field, which filters the noise and other artifacts resulting from the scattering.

Cai et al. [6] introduced an end-to-end system built on CNN that learns the mapping relations between hazy and haze free corresponding patches. To train the network they syntheses hazy images based on Middlebury stereo dataset. The method employs a non linear activation function that uses a bilateral restraint to improve convergence.

Ancuti et al. [1] introduced a local airlight estimation applied in a multi-scale fusion strategy for single-image dehazing. The method has been designed to solve the problems associated to the scattering effect, which is especially significant in the nigh-time hazy scenes. The method, however, generalizes to day-time dehazing.

Berman et al. [5] introduced an algorithm that also uses local airlight estimation, and extends the color consistency observation of Fattal [13]. The algorithm builds on the observation that the colors of a haze-free image can be approximated by a few distinct tight color clusters. Since the pixels of a cluster are spread on the whole image, they are affected differently by haze, to be spread along an elongated line, named haze-line. Those lines convey information about the transmission in different regions of the image, and are thus used for transmission estimation.

Ren et al. [28] adopts a multi-scale convolutional neural network to learn the mapping between hazy images and their corresponding transmission maps. The network is trained based on synthetic hazy images, generated by applying a simplified light propagation model to haze-free images for which corresponding depth maps are known.

4 Results and Discussion

In Fig. 2 are shown seven hazy images (first column), the corresponding ground truth (last column) and the dehazing results yielded by the specialised techniques of He et al. [15], Meng et al. [23], Fattal [13], Cai et al. [6], Ancuti et al. [1], Berman et al. [5] and Ren et al. [28].

Qualitatively, the well-known method of He et al. [15] yields results that visually seems to recover the structure, but suffers from color shifting in the hazy regions due to the poor airlight estimation. This happens when the scenes con-

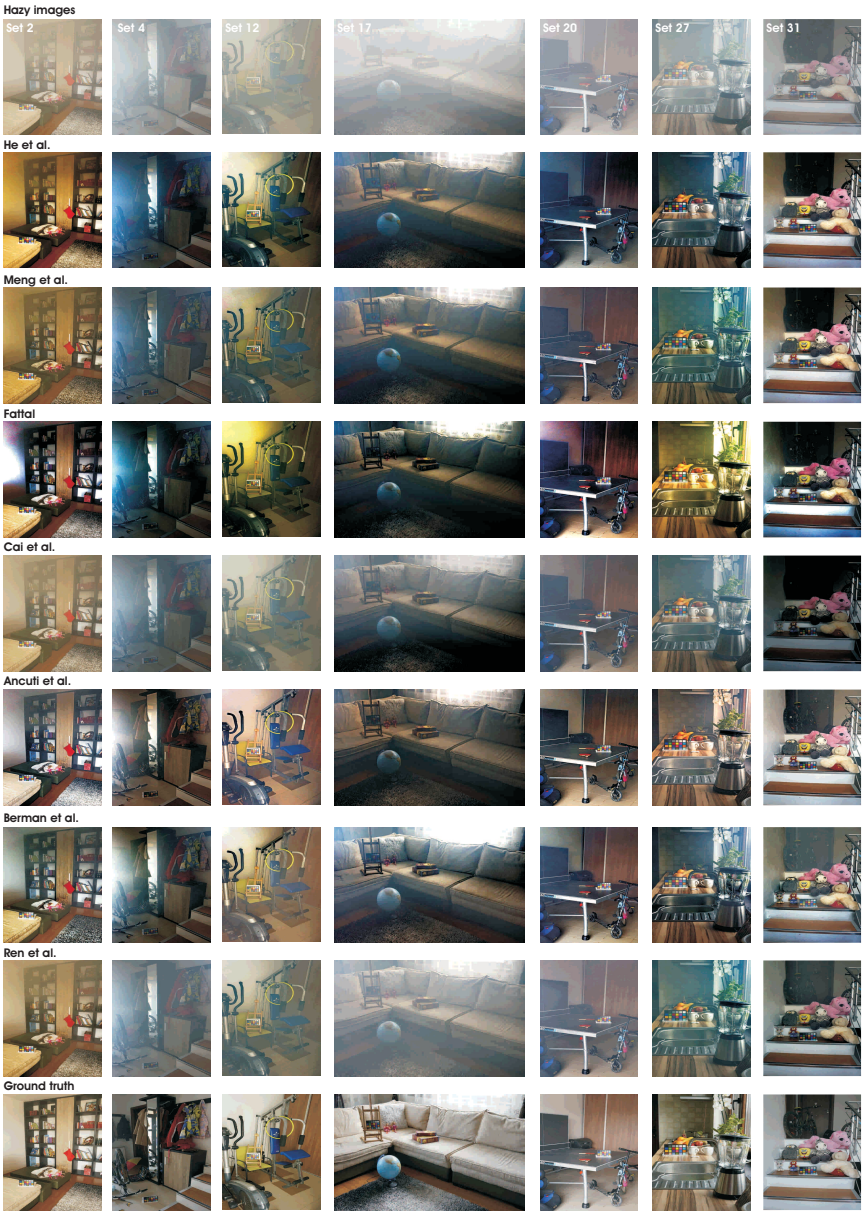

Fig. 2. ***Comparative results for seven I-HAZE sets of images.*** *The first row shows the hazy images and the last row shows the ground truth. The other rows from top to bottom show the results yielded by the methods of He et al. [15], Meng et al. [23], Fattal [13], Cai et al. [6], Ancuti et al. [1], Berman et al. [5] and Ren et al. [28].*

		Set 2	**Set 4**	**Set 12**	**Set 17**	**Set 20**	**Set 27**	**Set 31**
He et al.[15]	CIEDE2000	15.66	20.77	15.79	20.28	24.84	18.03	10.45
	SSIM	0.78	0.63	0.74	0.58	0.62	0.78	0.82
Meng et al[23]	CIEDE2000	13.58	16.58	16.39	20.01	19.57	17.25	16.12
	SSIM	0.82	0.70	0.78	0.66	0.79	0.82	0.79
Fattal[13]	CIEDE2000	19.85	20.92	18.97	26.77	23.43	15.88	16.24
	SSIM	0.61	0.57	0.70	0.44	0.54	0.82	0.66
Cai et al.[6]	CIEDE2000	9.81	19.97	17.32	18.91	24.04	20.86	24.66
	SSIM	0.85	0.59	0.76	0.56	0.61	0.63	0.41
Ancuti et al.[1]	CIEDE2000	14.07	14.47	13.23	21.90	15.76	10.52	7.95
	SSIM	0.82	0.74	0.84	0.62	0.76	0.88	0.88
Berman et al.[5]	CIEDE2000	14.13	17.60	9.57	17.51	16.01	16.87	11.24
	SSIM	0.80	0.69	0.83	0.71	0.81	0.79	0.81
Ren et al.[28]	CIEDE2000	9.91	18.06	11.81	12.15	10.74	13.78	11.12
	SSIM	0.85	0.64	0.83	0.67	0.90	0.87	0.87

Table 1. ***Quantitative evaluation.*** *We randomly selected seven set of images from our I-HAZE dataset and we compute the CIEDE2000 and SSIM indexes between the ground truth images and the enhanced results of the evaluated techniques. The hazy images, ground truth and the results are shown in Fig.2.*

tains lighter color patches in the close-up regions or small reflections. The Meng et al. [23] approach, built also on dark channel prior, as expected, generates similar results as He et al. [15]. It presents an improved alternative filtering of the transmission for artifacts reduction and a more precise airlight estimation. The results of color-lines method of [13] suffers from unpleasing color shifting while the results of Ancuti et al. [1] are more accurate due to the local airlight estimation strategy. This approach also is less prone to introduce structural artifacts due to the multi-scale fusion strategy. The results of Berman et al [5] are yielding visually compelling results and although it is built on a haze-line strategy due to its locally estimation of the airlight, shown to introduce less color shifting than [13]. In addition to the algorithms built on priors, deep learning techniques produce less visual artifacts. The results of Ren et al [28] generates visually more compelling results in comparison with the ones generated by DehazeNet of Cai et al [6].

To quantitatively evaluate the dehazing methods described in the previous section, we compare directly their outcome with the ground-truth (haze free) images. Besides the well known PSNR, we compute also the structure similarity index SSIM [37] that compares local patterns of pixel intensities that have been normalized for luminance and contrast. The structure similarity index yields values in the range [-1,1] with maximum value 1 for two identical images. Additionally, to evaluate also the color restoration we employ the CIEDE2000 [31,38]. Different than the earlier measures (e.g. CIE76 and CIE94) that shown shown important limitations to resolve the perceptual uniformity issue, CIEDE2000 defines a more complex, yet most accurate color difference algorithm. CIEDE2000

	SSIM	PSNR	CIEDE2000
He et al. [15]	0.711	15.285	17.171
Meng et al. [23]	0.750	14.574	16.834
Fattal [13]	0.574	12.421	21.385
Cai et al. [6]	0.771	16.983	12.991
Ancuti et al. [1]	0.770	16.632	14.428
Berman et al. [5]	0.767	15.942	14.629
Ren et al. [28]	0.791	17.280	12.736

Table 2. Quantitative evaluation of all the 35 set of images. In this table are shown the average values of the SSIM, PSNR and CIEDE2000 indexes over the entire dataset.

yields values in the range [0,100] with smaller values indicating better color preservation.

Table 1 presents a detailed validation based on SSIM and CIEDE2000 for seven randomly selected images of the I-HAZE dataset that are shown in the Fig. 2. In Table 4 are presented the average values over the entire dataset of the SSIM, PSNR and CIEDE indexes.

From these tables, we can conclude that the methods of Berman et al. [5], Ancuti et al. [1] and Ren et al. [28] performs the best in average when considering the SSIM, PSNR and CIEDE indexes. A second group of methods including Meng et al. [23] and He et al. [15], perform relatively well both in terms of structure and color restoration.

In general, all the tested methods introduce structural distortions such as halo artifacts close to the edges, that are amplified in the faraway regions. Moreover, due to the poor estimation of the airlight and transmission map from the hazy image, some color distortions may create some unnatural appearance of the restored images.

5 Conclusions

In this paper we introduced a novel dehazing dataset named I-HAZY. It consists of 35 image pairs of hazy and corresponding haze-free (ground-truth) indoor images. I-HAZE has been generated using real haze produced by a professional haze machine with both haze-free and hazy images captured under the same illumination conditions. Compared with previous dehazing datasets I-HAZE has an important advantage since it allows to objectively compare the existing image dehazing techniques using traditional image quality metrics. Additionally, we perform a comprehensive evaluation of several state-of-the-art single image dehazing methods. In summary, there is not a single technique that performs the best for all images. The relatively low values of SSIM, PSNR and CIEDE2000

measures prove once again the difficulty of single image dehazing task and the fact there is still much room for improvement.

References

1. Ancuti, C., Ancuti, C.O., Bovik, A., Vleeschouwer, C.D.: Night time dehazing by fusion. IEEE ICIP (2016)
2. Ancuti, C., Ancuti, C.O., Vleeschouwer, C.D.: D-hazy: A dataset to evaluate quantitatively dehazing algorithms. IEEE ICIP (2016)
3. Ancuti, C.O., Ancuti, C., Hermans, C., Bekaert, P.: A fast semi-inverse approach to detect and remove the haze from a single image. ACCV (2010)
4. Ancuti, C., Ancuti, C.: Single image dehazing by multi-scale fusion. IEEE Transactions on Image Processing **22(8)**, 3271–3282 (2013)
5. Berman, D., Treibitz, T., Avidan, S.: Non-local image dehazing. IEEE Intl. Conf. Comp. Vision, and Pattern Recog (2016)
6. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing (2016)
7. Chavez, P.: An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sensing of Environment (1988)
8. Chen, Z., Jiang, T., Tian, Y.: Quality assessment for comparing image enhancement algorithms. In IEEE Conference on Computer Vision and Pattern Recognition (2014)
9. Choi, L.K., You, J., Bovik, A.C.: Referenceless prediction of perceptual fog density and perceptual image defogging. In IEEE Trans. on Image Processing (2015)
10. Cozman, F., Krotkov, E.: Depth from scattering. IEEE Conf. Computer Vision and Pattern Recognition (1997)
11. Emberton, S., Chittka, L., Cavallaro, A.: Hierarchical rank-based veiling light estimation for underwater dehazing. Proc. of British Machine Vision Conference (BMVC) (2015)
12. Fattal, R.: Single image dehazing. SIGGRAPH (2008)
13. Fattal, R.: Dehazing using color-lines. ACM Trans. on Graph. (2014)
14. Hautiere, N., Tarel, J.P., Aubert, D., Dumont, E.: Blind contrast enhancement assessment by gradient ratioing at visible edges. Journal of Image Analysis and Stereology (2008)
15. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. In IEEE CVPR (2009)
16. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. on Pattern Analysis and Machine Intell. (2011)
17. He, K., Sun, J., Tang, X.: Guided image filtering. In IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2013)
18. Khoury(B), J.E., Thomas, J.B., Mansouri, A.: A color image database for haze model and dehazing methods evaluation. ICISP (2016)
19. Kopf, J., Neubert, B., Chen, B., Cohen, M., Cohen-Or, D., Deussen, O., Uyttendaele, M., Lischinski, D.: Deep photo: Model-based photograph enhancement and viewing. In: Siggraph ASIA, ACM Trans. on Graph. (2008)
20. Koschmieder, H.: Theorie der horizontalen sichtweite. In: Beitrage zur Physik der freien Atmosphare (1924)
21. Kratz, L., Nishino, K.: Factorizing scene albedo and depth from a single foggy image. ICCV (2009)

22. Li, Y., Tan, R.T., Brown, M.S.: Nighttime haze removal with glow and multiple light colors. In IEEE Int. Conf. on Computer Vision (2015)
23. Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In IEEE Int. Conf. on Computer Vision (2013)
24. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. In IEEE Trans. on Image Processing (2012)
25. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a completely blind image quality analyzer. In IEEE Signal Processing Letters (2013)
26. Narasimhan, S., Nayar, S.: Vision and the atmosphere. Int. J. Computer Vision, (2002)
27. Narasimhan, S., Nayar, S.: Contrast restoration of weather degraded images. IEEE Trans. on Pattern Analysis and Machine Intell. (2003)
28. Ren, W., Liu, S., Zhang, H., J. Pan, X.C., Yang, M.H.: Single image dehazing via multi-scale convolutional neural networks. Proc. European Conf. Computer Vision (2016)
29. Saad, M.A., Bovik, A.C., Charrier, C.: Blind image quality assessment: A natural scene statistics approach in the dct domain. In IEEE Trans. on Image Processing (2012)
30. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Polarization-based vision through haze. Applied Optics (2003)
31. Sharma, G., Wu, W., Dalal, E.: The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color Research and Applications (2005)
32. Tan, R.T.: Visibility in bad weather from a single image. In IEEE Conference on Computer Vision and Pattern Recognition (2008)
33. Tang, K., Yang, J., Wang, J.: Investigating haze-relevant features in a learning framework for image dehazing. In IEEE Conference on Computer Vision and Pattern Recognition (2014)
34. Tarel, J.P., Hautiere, N.: Fast visibility restoration from a single color or gray level image. In IEEE ICCV (2009)
35. Tarel, J.P., Hautire, N., Caraffa, L., Cord, A., Halmaoui, H., Gruyer, D.: Vision enhancement in homogeneous and heterogeneous fog. IEEE Intelligent Transportation Systems Magazine (2012)
36. Wang, Z., Bovik, A.C.: Modern image quality assessment. Morgan and Claypool Publishers (2006)
37. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing (2004)
38. Westland, S., Ripamonti, C., Cheung, V.: Computational colour science using matlab, 2nd edition. Wiley (2005)
39. Zhang, Y., Ding, L., Sharma, G.: Hazerd: an outdoor scene dataset and benchmark for single image dehazing. IEEE ICIP (2017)
40. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE IEEE Trans. Image Proc. (2015)