I-HAZE: A DEHAZING BENCHMARK WITH REAL HAZY AND HAZE-FREE INDOOR IMAGES

Codruta O. Ancuti *, Cosmin Ancuti *, Radu Timofte \ddagger and Christophe De Vleeschouwer \dagger

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

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 include 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.

Index Terms- dehazing, ground truth, quantitative evaluation

I. 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 [1], [2], multiple images captured with polarization filters [3], [4], or known depth [5], [6]. 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 [7], [8], [9], [6], [10], [11], [12], [13], [14], [15], [16], [17], by considering different kinds of priors to estimate the transmission. The method of Fattal [7] 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 [8] assumes the atmospheric airlight to be the brightest pixel in the Hazy images



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

scene, and estimate the transmission by maximizing the contrast. The maximal contrast assumption has also been exploited in [6], with a linear complexity in the number of image pixels. The dark channel priors, introduced in [9], has been proven to be a really effective solution to estimate the transmission, and has motivated many recent approaches. Meng et al. [18] extends the work on dark channel, by regularizing the transmission around boundaries to mitigate its lack of resolution. Zhu et al. [19] extend [18] by considering a color attenuation prior, assuming that the depth can be estimated from pixel saturation and intensity. The color-lines, introduced in [14], also exploit the impact of haze over the color channels distribution. Berman et al. [20] 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. [11] 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 [21], [22]. Fusion-based single image dehazing approaches [13], [23] have also achieved visually pleasant results without explicit transmission estimation and, more recently, several machine learning based methods have been introduced [17], [24], [25]. DehazeNet [24] 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 resort to synthesized training data based on the physical haze formation model. Ren et al. [25] 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 groundtruth image, most of the existing evaluation methods are based on non-reference image quality assessment (NR-IQA) strategies. For example, in [26], the assessment simply relies on the gradient of the visible edges. Other non-reference image quality assessment (NR-IQA) strategies [27], [28], [29] have been used for dehazing assessment. A more general framework has been introduced in [30], using subjective assessment of enhanced and original images captured in bad visibility conditions. Besides, the Fog Aware Density Evaluator (FADE) introduced in [23] 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 [31] 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 [32], a dataset of 1400+ images of real complex scenes has been derived from the *Middleburry*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 [33]. This dataset has been recently extended in [34].

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 introduce in [35]. 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 [37] 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.

II. 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 contains 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 has 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.

III. QUANTITATIVE EVALUATION

In this section, the I-HAZE dataset is used to perform a comprehensive quantitative evaluation of several state-of-the-art single image based dehazing techniques. Before presenting the 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 [33]. 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:

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

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

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



Fig. 2. Comparative results. The first row shows the hazy images and the last row shows the ground truth. The other rows from left to right show the results of He et al. [9], Meng et al. [18], Fattal [14], Cai et al. [24], Ancuti et al. [21], Berman et al. [20] and Ren et al. [25].

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

He et al. [9], [38] introduced the popular dark-channel prior, an extension of the dark object assumption [39]. 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 [9], or by using guided filters [38]. In our assessment, we employ the dark channel prior refined based on the guided filter [40].

Meng et al. [18] extend the dark channel prior [9]. 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 [14] 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. [24] 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. [21] 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 nightime hazy scenes. The method, however, generalizes to day-time dehazing.

Berman et al. [20] introduced an algorithm that also use local airlight estimation, and extends the color consistency observation of Fattal [14]. 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

	He et al.[9]		Meng et al.[18]		Fattal[14]		Cai et al.[24]		Ancuti et al.[21]		Berman et al.[20]		Ren et al.[25]	
	SSIM	CIEDE2000	SSIM	CIEDE2000	SSIM	CIEDE2000	SSIM	CIEDE2000	SSIM	CIEDE2000	SSIM	CIEDE2000	SSIM	CIEDE2000
Set 2	0.783	15.665	0.815	13.577	0.607	19.849	0.849	9.806	0.824	14.073	0.802	14.127	0.848	9.909
Set 4	0.633	20.767	0.700	16.579	0.568	20.920	0.588	19.969	0.742	14.472	0.691	17.597	0.643	18.063
Set 12	0.735	15.788	0.781	16.392	0.703	18.972	0.758	17.322	0.844	13.229	0.827	9.574	0.829	11.810
Set 20	0.617	24.836	0.790	19.568	0.539	23.428	0.608	24.042	0.763	15.763	0.806	16.010	0.899	10.737
Set 27	0.778	18.028	0.819	17.251	0.818	15.876	0.631	20.863	0.883	10.518	0.788	16.867	0.867	13.780

Table I. Quantitative evaluation. We randomly selected five set of images from our I-HAZE dataset and we compute the SSIM and CIEDE2000 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.

	He et al. [9]	Meng et al. [18]	Fattal [14]	Cai et al. [24]	Ancuti et al. [21]	Berman et al. [20]	Ren et al. [25]
SSIM	0.711	0.750	0.574	0.6397	0.770	0.767	0.791
PSNR	15.285	14.574	12.421	14.329	16.632	15.942	17.280
CIEDE2000	17.171	16.834	21.385	17.114	14.428	14.629	12.736

Table II. 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.

thus used for transmission estimation.

Ren et al. [25] 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.

IV. RESULTS AND DISCUSSION

In Fig. 2 are shown five hazy images (first column), the corresponding ground truth (last column) and the dehazing results yielded by the specialised techniques of He et al. [9], Meng et al. [18], Fattal [14], Cai et al. [24], Ancuti et al. [21], Berman et al. [20] and Ren et al. [25].

Qualitatively, the well-known method of He et al. [9] 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 contains lighter color patches in the close-up regions or small reflections. The Meng et al. [18] approach, built also on dark channel prior, as expected, generates similar results as He et al. [9]. 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 [14] suffers from unpleasing color shifting while the results of Ancuti et al. [21] 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 [20] 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 [14]. In addition to the algorithms built on priors, deep learning techniques produce less visual artifacts. The results of Ren et al [25] generates visually more compelling results in comparison with the ones generated by DehazeNet of Cai et al [24].

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 [41] 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 [37], [42]. 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 yields values in the range [0,100] with smaller values indicating better color preservation.

Table I presents a detailed validation based on SSIM and CIEDE2000 for five randomly selected images of the I-HAZE dataset that are shown in the Fig. 2. In Table III 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. [20], Ancuti et al. [21] and Ren et al. [25] performs the best in average when considering the SSIM, PSNR and CIEDE indexes. A second group of methods including Meng et al. [18] and He et al. [9], 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. 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.

V. REFERENCES

- [1] F. Cozman and E. Krotkov, "Depth from scattering," *IEEE Conf. Computer Vision and Pattern Recognition*, 1997.
- [2] S.G. Narasimhan and S.K. Nayar, "Vision and the atmosphere," Int. J. Computer Vision, 2002.
- [3] S.G. Narasimhan and S.K. Nayar, "Contrast restoration of

weather degraded images," IEEE Trans. on Pattern Analysis and Machine Intell., 2003.

- [4] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," *Applied Optics*, 2003.
- [5] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," in *Siggraph ASIA, ACM Trans. on Graph.*, 2008.
- [6] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," *In IEEE ICCV*, 2009.
- [7] Raanan Fattal, "Single image dehazing," SIGGRAPH, 2008.
- [8] Robby T. Tan, "Visibility in bad weather from a single image," In IEEE Conference on Computer Vision and Pattern Recognition, 2008.
- [9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *In IEEE CVPR*, 2009.
- [10] L. Kratz and K. Nishino, "Factorizing scene albedo and depth from a single foggy image," *ICCV*, 2009.
- [11] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert, "A fast semi-inverse approach to detect and remove the haze from a single image," ACCV, 2010.
- [12] C. O. Ancuti, C. Ancuti, and P. Bekaert, "Effective singleimage dehazing by fusion," in *In IEEE ICIP*, 2010.
- [13] C.O. Ancuti and C. Ancuti, "Single image dehazing by multiscale fusion," *IEEE Transactions on Image Processing*, vol. 22(8), pp. 3271–3282, 2013.
- [14] Raanan Fattal, "Dehazing using color-lines," ACM Trans. on Graph., 2014.
- [15] C. Ancuti and C. O. Ancuti, "Effective contrast-based dehazing for robust image matching," *IEEE Geoscience and Remote Sensing Letters*, 2014.
- [16] S. Emberton, L. Chittka, and A. Cavallaro, "Hierarchical rankbased veiling light estimation for underwater dehazing," *Proc.* of British Machine Vision Conference (BMVC), 2015.
- [17] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," *In IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [18] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," *In IEEE Int. Conf. on Computer Vision*, 2013.
- [19] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE IEEE Trans. Image Proc.*, 2015.
- [20] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," *IEEE Intl. Conf. Comp. Vision, and Pattern Recog*, 2016.
- [21] C. Ancuti, C. O. Ancuti, A.C. Bovik, and Christophe De Vleeschouwer, "Night time dehazing by fusion," *IEEE ICIP*, 2016.
- [22] Y. Li, R. T. Tan, and M. S. Brown, "Nighttime haze removal with glow and multiple light colors," *In IEEE Int. Conf. on Computer Vision*, 2015.
- [23] L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," *In IEEE Trans. on Image Processing*, 2015.
- [24] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet:

An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, 2016.

- [25] W. Ren, S. Liu, H. Zhang, X. Cao J. Pan, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," *Proc. European Conf. Computer Vision*, 2016.
- [26] N. Hautiere, J.-P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Journal of Image Analysis and Stereology*, 2008.
- [27] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *In IEEE Trans. on Image Processing*, 2012.
- [28] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *In IEEE Signal Processing Letters*, 2013.
- [29] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *In IEEE Trans. on Image Processing*, 2012.
- [30] Z. Chen, T. Jiang, and Y. Tian, "Quality assessment for comparing image enhancement algorithms," *In IEEE Conference* on Computer Vision and Pattern Recognition, 2014.
- [31] J.-P. Tarel, N. Hautire, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," *IEEE Intelligent Transportation Systems Magazine*, 2012.
- [32] C. Ancuti, C. O. Ancuti, and Christophe De Vleeschouwer, "D-hazy: A dataset to evaluate quantitatively dehazing algorithms," *IEEE ICIP*, 2016.
- [33] H. Koschmieder, "Theorie der horizontalen sichtweite," in *Beitrage zur Physik der freien Atmosphare*, 1924.
- [34] Yanfu Zhang, Li Ding, and Gaurav Sharma, "Hazerd: an outdoor scene dataset and benchmark for single image dehazing," *IEEE ICIP*, 2017.
- [35] Jessica El Khoury(B), Jean-Baptiste Thomas, and Alamin Mansouri, "A color image database for haze model and dehazing methods evaluation," *ICISP*, 2016.
- [36] Z. Wang and A. C. Bovik, "Modern image quality assessment," Morgan and Claypool Publishers, 2006.
- [37] G. Sharma, W. Wu, and E. Dalal, "The ciede2000 colordifference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Research and Applications*, 2005.
- [38] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. on Pattern Analysis* and Machine Intell., 2011.
- [39] P.S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sensing of Environment*, 1988.
- [40] K. He, J. Sun, and X. Tang, "Guided image filtering," In IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2013.
- [41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, 2004.
- [42] S. Westland, C. Ripamonti, and V. Cheung, "Computational colour science using matlab, 2nd edition," *Wiley*, 2005.