Journal of Imaging Science and Technology® 62(1): 010503-1–010503-13, 2018.

A Database with Reference for Image Dehazing Evaluation

Jessica El Khoury, Jean-Baptiste Thomas▲, and Alamin Mansouri
Le2i, FRE CNRS 2005, Université Bourgogne Franche-Comté, Dijon, France
E-mail: jessica.el-khoury@u-bourgogne.fr

Abstract. *In this article, the authors introduce a new color image database, CHIC (Color Hazy Images for Comparison), devoted to haze model assessment and dehazing method evaluation. For three real scenes, they provide two illumination conditions and several densities of real fog. The main interest lies in the availability of several metadata parameters such as the distance from the camera to the objects in the scene, the image radiance and the fog density through fog transmittance. For each scene, the fog-free (ground-truth) image is also available, which allows an objective comparison of the resulting image enhancement and potential shortcomings of the model. Five different dehazing methods are benchmarked on three intermediate levels of fog using existing image quality assessment (IQA) metrics with reference to the provided fog-free image. This provides a basis for the evaluation of dehazing methods across fog densities as well as the effectiveness of existing dehazing dedicated IQA metrics. The results indicate that more attention should be given to dehazing methods and the evaluation of metrics to meet an optimal level of image quality. This database and its description are freely available at the web address http://chic.u-bourgogne.fr.* © *2018 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2018.62.1.010503]

1. INTRODUCTION

Several studies have been carried out to develop vision systems, which are basically designed to perform in good visibility conditions, such as clear weather. However, a number of undesirable visual effects are present in images acquired in a polluted area or on foggy and hazy days. In such circumstances, visibility is mainly reduced by the absorption and scattering of light by both gases and atmospheric particles. However, when it comes to modeling, the scattering of light is considered to be the main phenomenon, which lies behind visibility degradation [1]. This degradation leads not only to an attenuation of intensity but also to a change in the perception of the color of the objects in a scene. These effects are encapsulated in one of the most commonly used models that describes the visibility degradation, usually known as the Koschmieder model of haze, which is a wavelength-independent pixelwise model [2],

$$I(x) = J(x)t(x) + A_\infty(1 - t(x)), \tag{1}$$

where x denotes the pixel location, $I(x)$ is the intensity of the hazy image, $J(x)$ is the scene radiance and $t(x) = e^{-\beta z(x)}$ is the transmission factor, which represents the light that emanates from the object and is attenuated by the scattering along the line of sight. This function describes the exponential attenuation of the scene radiance with the scene depth z and the scattering coefficient of the fog β. The airlight corresponding to an object at an infinite distance is called atmospheric light, A_∞. The airlight $A(x) = A_\infty(1 - t(x))$ is the light coming from a source of light (i.e., the Sun) and scattered by the atmospheric particles toward the camera. It causes the atmospheric particles to behave like a light source. It represents, therefore, the color of the physical veil that covers the scene.

The recovery process of $J(x)$ from $I(x)$ is known as dehazing or defogging. Although weather conditions differ with respect to particle types, sizes, concentrations and the resulting veil hue and visibility [1], the same methods are used indiscriminately for all aerosols, since the haze model (Eq. (1)) is a good approximation for a wide range of weather conditions and for several situations [3].

Evaluation of the haze model and the dehazing algorithms is a major challenge for ensuring the image quality of the processed image. In this context, the lack of a fog-free reference image makes the evaluation even more difficult. We propose to address this challenge by the creation of a color image database that is characterized by a real scene covered by fog of several densities and coupled with a number of geometric and physical parameters. This database allows one to investigate the quality of a haze model according to fog densities and the efficiency of dehazing algorithms.

In the following, we briefly present the state of the art of dehazing methods and we address the challenging image quality assessment of dehazed images. Then, we review databases that have been created for similar purposes. We then introduce the CHIC (Color Hazy Images for Comparison) database by describing the material used, the scene setup and the associated metadata. We finally give a basis for the evaluation of dehazing methods across fog densities and the effectiveness of existing image quality assessment (IQA) metrics.

2. RELATED WORK

A significant number of dehazing methods have been proposed to improve the visibility of degraded images. The family of dehazing methods is presented in (Figure 1).

Two main categories are usually considered for dehazing methods. The first one requires multiple inputs such as images and additional data to break down the ill-posedness

▲ IS&T Member.

Received Nov. 23, 2016; accepted for publication July 25, 2017; published online Dec. 20, 2017. Associate Editor: Rita Hofmann-Sievert.
1062-3701/2018/62(1)/010503/13/$25.00

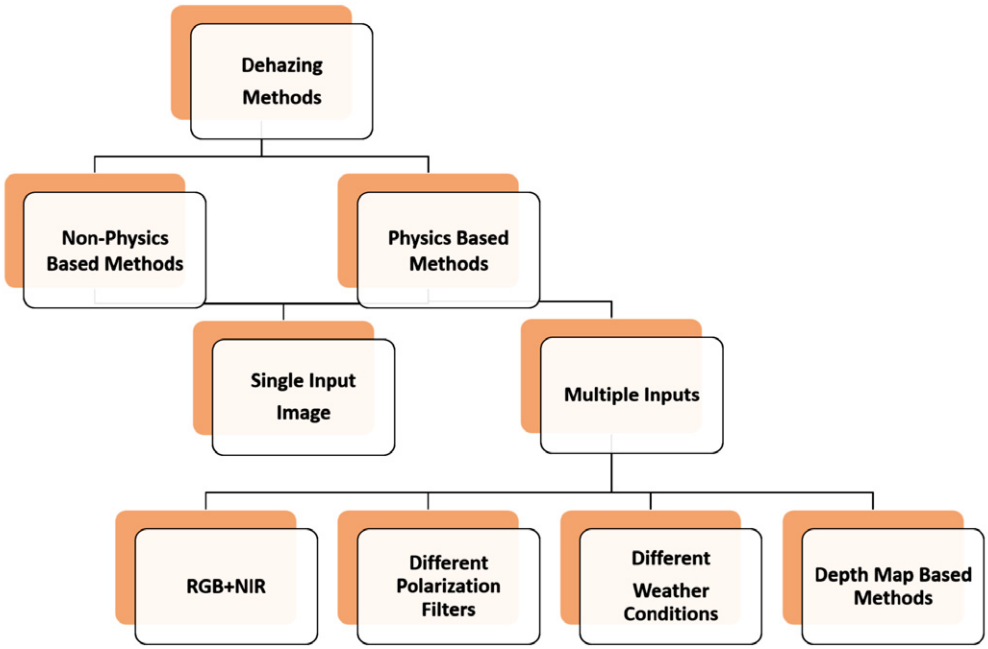

Figure 1. The categories of dehazing methods.

of the dehazing problem while estimating the haze model parameters t, A and J. The multiple images can be of different types, such as RGB and infrared images. Dehazing is then performed based on their dissimilarities [4, 5]. Different polarization filters can also be used [6] and multiple images can be taken of the same scene that have different degrees of polarization. Different images can also be taken under various weather conditions [3]. This approach can significantly improve visibility; however, it may not be convenient in practical cases due to the required setup or the time that may occur between several image captures.

Moreover, besides the hazy image, some further data such as depth and texture information may be collected. For instance, the depth map based method proposed by Kopf et al. [7] assumes that the 3D geometrical model is given by some image database such as Google Maps and the scene texture is provided by aerial or satellite photos. In order to give accurate results, the alignment of the 3D model performed by the user is required.

The second category requires a single input image and it is based on prior assumptions. Thus, it has a fully automatic performance. The latest works mostly focus on single-image methods. The first type, physics based methods, consider some assumptions to estimate the unknown physical parameters of the visibility degradation model (Eq. (1)). Some of them are based on statistical information such as the dark channel prior **DCP** [8]; some others perform white balance and contrast enhancement weight fusion, such as **FUSION** [9]. The **FAST** algorithm estimates the unknown parameters of the visibility degradation model through a median filtering and adjusts after processing the image intensity to the visible range [10]. The second type, image enhancement methods, adopt a subject process that aims to improve the quality of the image according to the visual experience by improving contrast and visibility, like **CLAHE** (Contrast Limited Adaptive Histogram Equalization) [11], or adopt a variational way through an iterative process conditioned by the convergence toward an optimal image quality [12] (**VAR**).

Methods of both categories are usually evaluated using non-standardized images in a typically qualitative way. Moreover, the lack of a fog-free reference image in the evaluation of dehazing methods is a considerable problem, since it is hard to quantify in subjective and objective manners, what an image of good quality is. Moreover, the metrics that have already been used to evaluate dehazed images examine the dissimilarity between the dehazed and the hazy image [13–15], which does not necessarily denote an enhancement in the image quality when its value increases. Hence, it is important to conceive a database of real foggy images with the fog-free one, which represents with its additional data a suitable tool for quality evaluation.

3. EXISTING DATABASES

A collection of single outdoor images taken in bad weather without any other considerations is commonly used to test dehazing algorithms and for comparative studies of methods. The lack of common conditions of acquisition such as illuminant, viewing geometry and weather conditions usually leads to a qualitative evaluation according to the subjective opinion of the authors.

There have been for a while no commonly accepted databases, metrics or procedures that allow an objective evaluation of dehazing quality. A suitable database should contain at least benchmarked methods with an objective

evaluation procedure. This may be created by having a reference fog-free image, which is supposed to have an optimal image quality, and foggy images containing real surfaces and covered with fog.

Among existing databases, we can find FRIDA (Foggy Road Image DAtabase) [16, 17] and FRIDA2 [17, 18], which are two synthetic databases with references. They comprise a number of synthetic images of urban road scenes and diverse road scenes, respectively. They were mainly designed for fog detection and visibility enhancement through the use of onboard cameras [19, 20]. The software *SiVIC*^{TM} was used to build physical rendering of road environments and to generate a moving vehicle with a physically driven model of its dynamic behavior, and virtual embedded sensors. Images were produced from a realistic complex urban model and a virtual camera onboard a simulated vehicle moving on a road path. To each image without fog, four foggy images and a depth map were associated. The depth map was required to be able to add fog properly to the images. Different types of fog were added to each of the four associated images: uniform fog, heterogeneous fog, cloudy fog and cloudy heterogeneous fog. These four types of fog were inserted by applying Koschmieder's law [2] via weighting the attenuation coefficient and/or the atmospheric light differently with respect to the pixel position. In order to increase the realism of the scenes, Perlin's noise was embedded into the images. However, the simulated images fail to represent natural phenomenon effects accurately, which implies some inaccuracies after dehazing at high fog densities [21, 22]. The physical interaction between light and atmospheric particles modifies the perceived colors, while, in the simulated image, colors maintain their hue information and only their saturation component shifts between the original color (saturated) and the fog color (unsaturated). This database also provides a single type of images with the same scene composition of unnatural surfaces on similar backgrounds.

WILD (Weather and Illumination Database) [23, 24] is an outdoor urban scene database, where images are acquired automatically every hour over several months. These images are taken under different weather and illumination conditions. Atmospheric conditions, scene distances and temporal data are also associated to images. This database with uncontrolled illumination and regulated haze/fog densities remains far from correctly evaluating dehazing processes. Moreover, small changes that could occur over time would bias a pixel by pixel image comparison, such as full-reference image quality metrics, even if the illumination changes between images are subtracted.

In addition to these databases, there are hazy image databases, without references. They are usually of small size with various contexts [25, 26].

All of these limitations were reported in our previous work [27] and recently in [28], where the authors provide a dataset of synchronous and registered RGB/NIR image pairs captured simultaneously of seven outdoor hazy images without the ground-truth images and four indoor hazy scenes available with the ground-truth haze-free images. Haze was generated using a haze machine with a water based haze fluid. Since it provides a couple of RGB and NIR images, besides the evaluation of single image dehazing algorithms, it serves to evaluate multiple spectral input algorithms. However, only one haze level is provided, without any other additional information such as the depth. Consequently, the investigation of the performance of dehazing methods across fog levels as well as the haze model assessment are not feasible.

4. CHIC DATABASE

Based on the shortcomings of previous databases reported earlier, the CHIC database is characterized by the presence of the ground-truth reference images. Several densities of fog are provided in a well-controlled environment with various lighting conditions. Indeed, several pieces of equipment were used in order to provide a full package that allows estimation of the haze model parameters, including depth, and subsequent reconstruction of the clear images.

All of the provided data make the CHIC database suitable for evaluation and comparison of the quality of dehazed images in both qualitative and quantitative studies and the assessment of IQA metrics. An example is given in Section 6. The data also offer the ability to evaluate the accuracy of the haze model across fog densities. An example of this is proposed in our previous article [27], where the description of the database was not exhaustive and the methods and IQA metric evaluation were not addressed. Accordingly, the haze model does not hold at high fog densities: while color saturation is mainly affected at low fog density, the color hue becomes unrecoverable beyond fog level 5 [29].

5. MATERIAL AND SETUP

1. Nikon D7100 color camera [30]. This camera was used to obtain RGB images of the scenes when the fog density changed. As the white balance option cannot be disabled, the cloudy option was selected. ISO sensitivity remained tuned to the default value of 100. The auto-focus was disabled. Two image formats were recorded: NEF (RAW) and fine-quality JPEG images, whose dimensions are 6000×4000 (the recorded JPEG images are compressed at a compression ratio of roughly 1:4).

During the shooting session for each scene, the camera was maintained at the same position with the same selected options. The capture conditions such as the field of view of the camera, the illumination and the object colors remained the same. The only variable element between two shots was the density of the fog layer. This allowed us to follow up the haze model parameter evolution according to its value.

2. Konica Minolta CS-2000 spectroradiometer [31]. This spectroradiometer was used to measure the transmission of the white patch of the gretag macbeth color checker placed at the back of the scene for each fog density. It

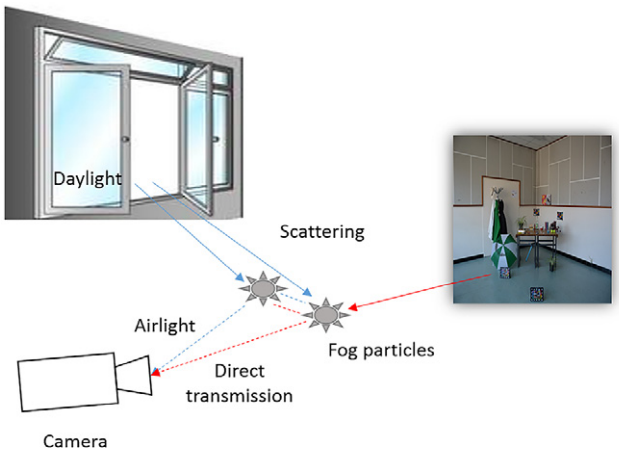

Figure 2. The light reaching the camera is the sum of the light coming from the object and attenuated by the scattering and the light coming from outside and also scattered by fog particles toward the camera.

was calibrated focusing on the same patch without fog. The selected measuring angle was 1° with the standard observer 2° and the standard illuminant D65.

3. Fog machine. We used a FOGBURST 1500, whose flow rate was 566 m^3/min and spraying distance was 12 m, which emitted a dense vapor that appeared similar to fog. A large quantity of fog was initially emitted until it was evenly distributed in the room and formed an opaque layer. Fog was then progressively evacuated through the window. Thereby, images of different levels of fog were captured. This machine operates by evaporation. The vaporization of the water based liquid mixed with glycol therein is achieved by heating. The difference is in the quality of the desired smoke. Fog is diffuse and not very dense. Smoke is denser and less diffuse. The particles of the ejected fog are water droplets whose radius is close to the radius size of atmospheric fog ($1 - 10$ μm) [29]. Similarly to what was evoked in [28], unlike synthetic images, when a fog machine is used, we get a realistic scattering of the atmospheric light.

4. Macbeth Color Checker (MCC). Several color calibration targets were distributed in the scene at different distances from the camera. The known values of the spectral reflectances of the patches allowed us to follow color visibility variation with distance and fog density. The matte-colored patches of the MCCs allow them to be captured in any orientation and line of sight [32]. This can be useful for following up the color alteration when fog covers a scene, and how far the color fidelity can be preserved when the haze model formula is simply reversed without considering any prior hypothesis.

5.1 *Scenes*

Scenes were set up in a closed rectangular room, which was large enough to simulate the effect of the distance and

Figure 3. Scene A. The distance of different MCCs from the camera.

Figure 4. Scene B. The distance of different MCCs from the camera.

the fog density on the object radiance (length = 6.35 m, width = 6.29 m, height = 3.20 m), with a large window (length = 5.54 m, width = 1.5 m) that allowed a large amount of outdoor light to get in, on a sunny day. The photo session of each scene lasted for 20 minutes. During this limited time, the daylight was considered to remain steady. This experiment was set up in February, from 1:00 to 4:00 pm, when sunlight was not directly coming in through the window (Figure 2). Thus, the global light was close to the airlight on a cloudy day. Several MCCs were placed at various distances in the scene to assess color recovery. The most distant MCC served to estimate the atmospheric light. Moreover, in the captured scenes, there were various colored natural surfaces (reflective and glossy surfaces, rough surfaces, etc.).

The three scenes are grouped into two categories: static and dynamic scenes.

5.2 *Static Scenes*

In the static scenes, the camera position is fixed and only the fog density changes during imaging of the same scene.

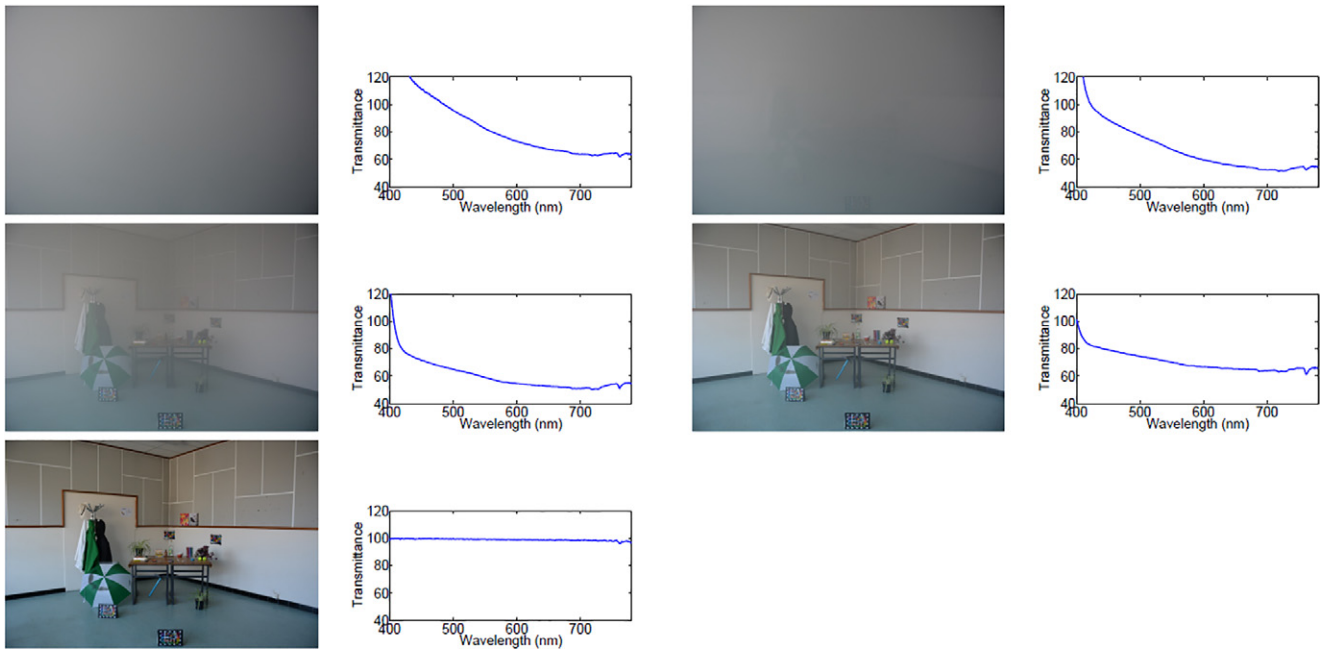

Figure 5. Scene A. Scene images taken under different fog levels and their corresponding transmission curves, which are measured over the most distant white patch from the camera. From left to right, row 1: level 1, level 3; row 2: level 5, level 8 and row 3: the clear image. The image levels are obtained as fog is progressively evacuated. The calibration process is performed on the clear scene. These spectra show the airlight behavior. The transmission of short wavelengths exceeds 100. This is due to the incoming short-wavelength light scattered by the particles before reaching the scene objects. The nonlinearity of the curves shows that different visible wavelengths are not identically handled in such a foggy medium.

- Scene A. The shooting session for this scene was carried out at around 2:00 pm (see Figure 3). This scene shows a typical indoor view. We put on the table placed in the middle a number of items with different characteristics such as shapes, colors, positions, surface types (smooth or rough surfaces) and textures. The wall behind the scene is half white, and the top half with the white lines and the black holes represents distinctive elements to study algorithm performance near edges.
- Scene B. The shooting session for scene B was carried out two hours later (see Figure 4). During this time, the correlated color temperature of the illuminant changed significantly. Compared with the first scene, the distance from the camera to the most distant point is shorter. It contains larger geometric shapes. The fog densities that are randomly chosen are characterized by the transmittance spectrum of each fog level (see Figures 5 and 6).

For each fog density, scenes were captured under two illuminants: outdoor daylight and a compound light (outdoor light + ceiling lamp light (fluorescent tube)).

The illuminant as well as its intensity are not the same for both scenes, nor are the distance from the camera and the density of fog of the corresponding levels (Table I).

Table I shows the values of the relative transmittance of the fog, which are calculated through the fog levels over the black patch of the MCC placed at the back of the scene of the original hazy images, as follows:

$$T = 1 - \frac{S_{\text{Levelx}} - S_{\text{airlight}}}{S_{\text{fog-free}} - S_{\text{airlight}}}, \tag{2}$$

where S_{Levelx}, S_{airlight} and $S_{\text{fog-free}}$ are the spectral values of the green channel of the image at fog level x, of the airlight image of our database where the scene is completely covered by fog and the fog-free image, respectively.

Such scenes with the fog-free reference images are suitable to evaluate the haze model across fog levels and to compare the performance of different dehazing methods with regard to the fog-free images. They could be particularly useful for the design of metrics dedicated to dehazing image assessment quality.

5.3 *Dynamic Scenes*

In dynamic scenes, for a given density of fog, the distance between the scene and the camera changes. We perceive across images that the fog density changes as the distance is varied. In an approximately uniform smoke-filled closed room, we took photos of an outdoor landscape while moving toward the window separating the foggy indoor from the clear outdoor medium from two points of view. At each point of view, nine levels of fog density were considered by moving in steps, each of 70 cm. Image level 1 is 5.6 m from the window and image level 9 is taken when this distance is null (Figure 7). The corresponding foggy and fog-free images that are taken at the same distance are available. The moving objects within the scenes make them unsuitable

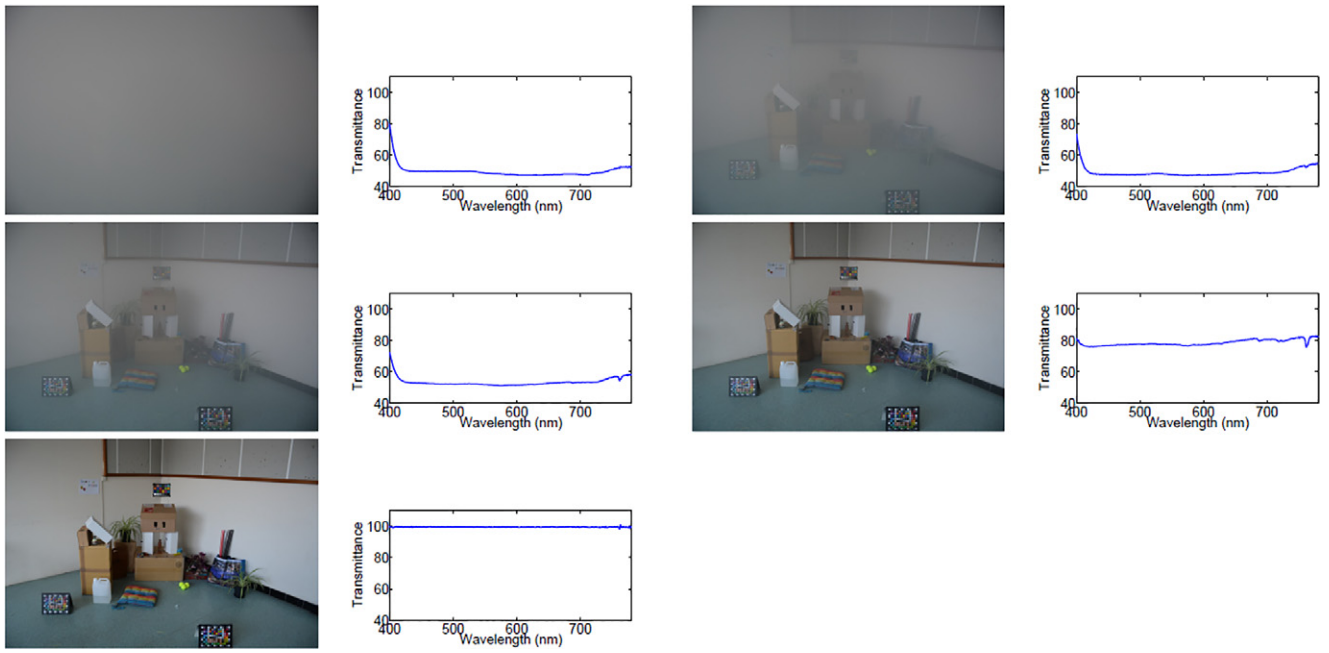

Figure 6. Scene B. Scene images taken under different fog levels and their corresponding transmission curves, which are measured over the most distant white patch from the camera. From left to right, row 1: level 1, level 3; row 2: level 5, level 8 and row 3: the clear image. The calibration process is performed on the clear scene. As the dimensions of the scene are smaller, the distance traveled by light is shorter; hence, the weak short wavelengths of the incoming light are weakly scattered; therefore, transmission does not exceed 100, and the transmission values across wavelengths are almost the same.

Table I. Relative transmittance T of fog in the original foggy images of scenes A and B. Level 1: highest fog density. Level 9: lowest fog density.

T	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9
Scene A	100 (%)	92(%)	91(%)	90 (%)	84 (%)	75 (%)	71 (%)	52 (%)	28 (%)
Scene B	100 (%)	97(%)	91(%)	89 (%)	74 (%)	64 (%)	55 (%)	30 (%)	15 (%)

for evaluation using full-reference IQA. However, other evaluation modes such as visual judgment are possible. In these conditions, the fog is an additive veil disconnected from the scene. These scenes could be used to investigate dehazing method performance on non-classical foggy scenes.

5.4 *Global Meta Information*

According to Eq. (1), I is the captured image with a given fog density; J is the clear image under the same illuminant as the foggy image. The atmospheric light A_∞ comprises the R, G and B values of the fog layer calculated from image level 1, which is completely covered by fog. For each level of fog, the airlight $A(x) = A_\infty(1 - t(x))$ is calculated over the black patch of the MCC placed at the back of the scene. Once A_∞ and $A(x)$ have been calculated, the transmittance $t(x)$ of light coming from this MCC is then deduced, as its distance from the camera is known. At this distance, β is then deduced from $t(x) = e^{-\beta z(x)}$. For the rest of the image, $t(x)$ is calculated based on the provided object distances. In addition to the image set, a descriptive image shows the distance of some specific objects from the camera, especially the ones of the MCCs (see Figs. 2 and 3).

Although these scenes are acquired using the same techniques, they show different characteristics. This is due to the change of the illuminant and the reduction of the distance between the camera and the scene. When we move away from the window, the light intensity decreases. This fact is accentuated in the presence of fog, and it is clearer in scene A when the intensity of the incoming light is higher. Small wavelengths are easily scattered by fog particles. Thus, the transmission through fog increases with wavelength, although, according to the transmission curves in Fig. 5, the transmission of the low wavelengths exceeds 100. This is due to the incoming short wavelengths being scattered by fog particles before reaching the scene objects (see Fig. 4).

Since the light intensity is higher in scene A than in scene B, and the camera is closer to the window, the incoming light hits the fog particles and it is redirected toward the camera before reaching the scene. This effect is shown in the curves of Fig. 6, and it is clear at the short wavelengths, which have a lower penetration capability than long wavelengths.

In all of this, scenes A and B could be employed for outdoor dehazing method evaluation. Scene A, in particular, could be used as well for the evaluation of indoor smoky

Figure 7. Dynamic scene. The left column shows images with fog taken at 5.6 m (level 1), 3.5 m (level 4) and 0 m (level 9) from the window, respectively. The right column shows the fog-free images taken at the same distances.

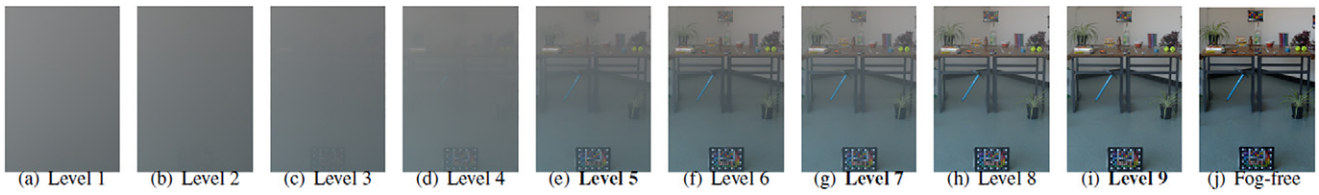

Figure 8. Scene A. From (a)–(j): level 1 (highest fog level), level 2, level 3, level 4, level 5, level 6, level 7, level 8, level 9 (lowest fog level) and fog-free. The image size is 1537 × 2049. These images are successively acquired within 20 minutes. The cropped areas are assumed to have an even intensity to correlate better with most dehazing method assumptions.

scene algorithms. Dynamic scenes could be useful for cross validation of a given method and confirming its ability to handle different scenarios of hazy images. Moreover, this database prompts researchers to address the uneven illuminant issue, which has been often neglected [33]. The known levels of fog make this database useful for testing the efficiency of algorithms dedicated to hazy image classification [34].

At the web page http://chic.u-bourgogne.fr, the main directory of the database contains sub-directories for each of the aforementioned scenes. Each of them contains full images and images of the cropped color checker placed at the back of the scenes. They also include cropped images, which we used to compare methods (see Figures 8 and 9), and reduced-size directories: the cropped areas are assumed to have an even intensity to correlate better with most dehazing

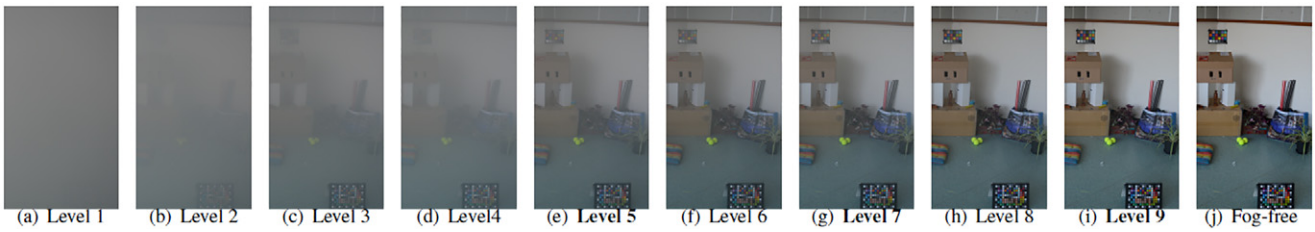

Figure 9. Scene B. From (a)–(j): level 1 (highest fog level), level 2, level 3, level 4, level 5, level 6, level 7, level 8, level 9 (lowest fog level) and fog-free. The image size is 1537 × 3073. These images are successively acquired within 20 minutes. The cropped areas are assumed to have an even intensity to correlate better with most dehazing method assumptions.

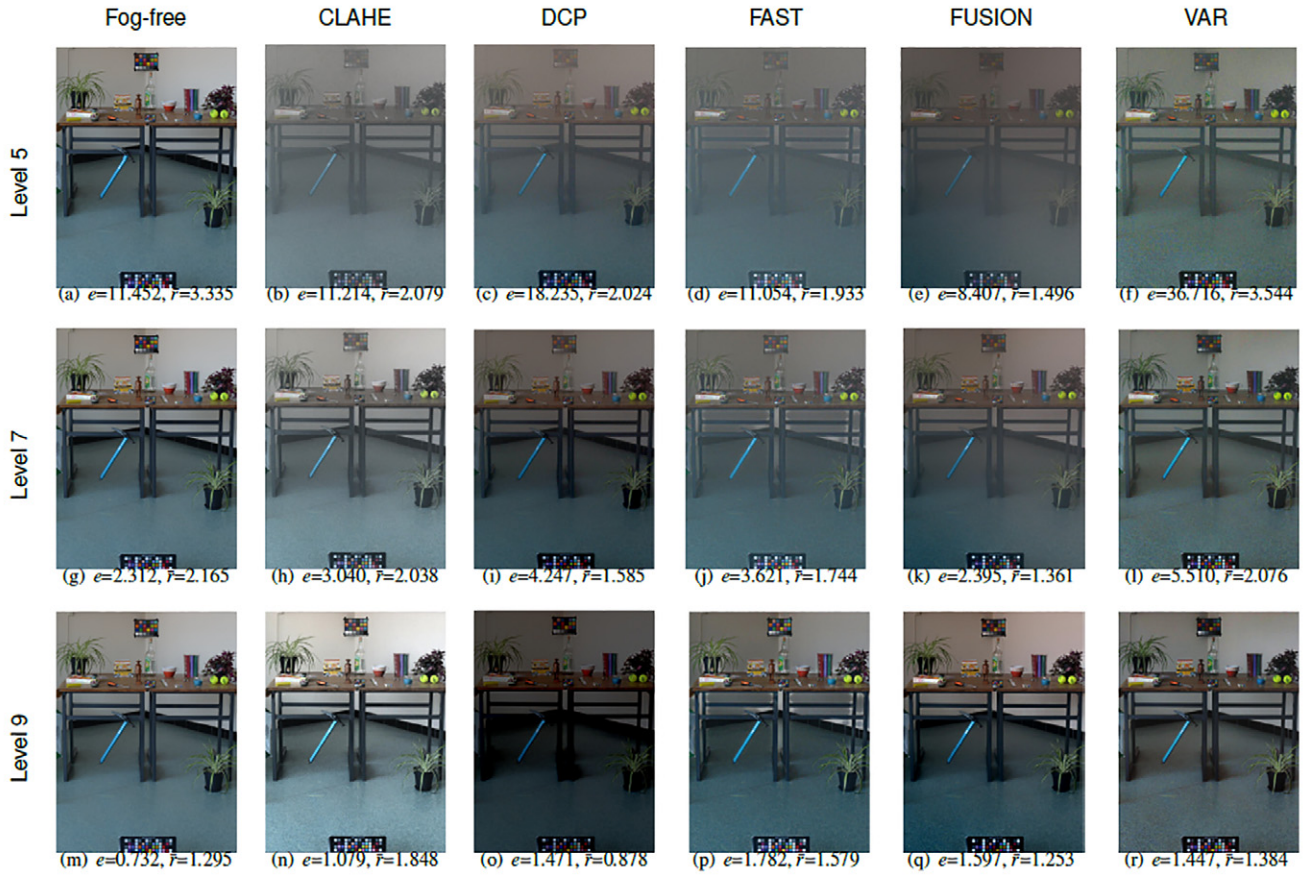

Figure 10. The foggy images of scene A processed by the selected dehazing methods at fog levels 5 (row 1), 7 (row 2) and 9 (row 3). Column 1, fog-free image; column 2, CLAHE result; column 3, DCP result; column 4, FAST result; column 5, FUSION result; column 6, VAR result. The values of e and $\bar{r}$, which are close to the values computed on the fog-free image, underline, respectively, fidelity in the structure recovery and the contrast enhancement compared with the fog-free image.

method assumptions, and the reduced-size images can be used to reduce the prohibitive computational time of several dehazing methods.

6. APPLICATION OF EXPERIMENTAL DATA: EVALUATION OF DEHAZING METHODS AND QUALITY METRICS

6.1 *Data and Tools*

Our data can be used to evaluate dehazing algorithms and IQA metrics to show the advantage of a reference based objective assessment across several fog densities. According to Narasimhan et al. [3], the haze model is more suitable for short ranges of distance. Similarly, it fails to correctly estimate the light attenuation of scenes covered by a thick veil of fog. According to our previous work [27], from level 5, dehazing methods produce visually clear images that can be compared with the free-fog images. As the foggy images of adjacent fog levels are very close to each other (see Figs. 8 and 9 and Table I), we only selected levels 5, 7 and 9 for this benchmark. Five of the most commonly used dehazing methods were applied to the selected images of scenes A and B: CLAHE, DCP, FAST, FUSION and VAR, which we introduced in Section 2. The corresponding dehazed images of both scenes are shown in Figures 10 and 11.

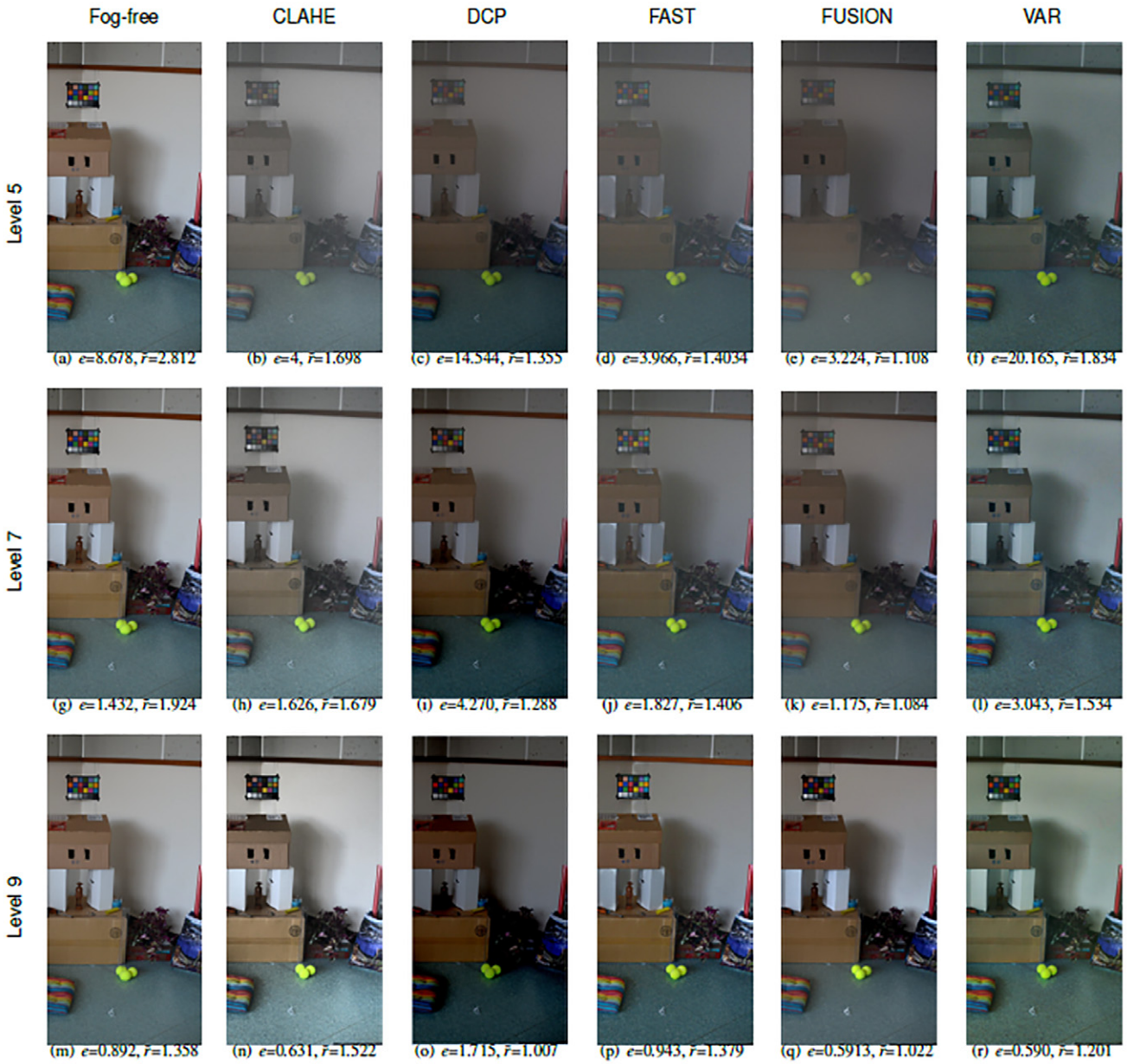

Figure 11. The foggy images of scene B processed by the selected dehazing methods at fog levels 5 (row 1), 7 (row 2) and 9 (row 3). Column 1, fog-free image; column 2, CLAHE result; column 3, DCP result; column 4, FAST result; column 5, FUSION result; column 6, VAR result. The values of e and $\bar{r}$, which are close to the values computed on the fog-free image, underline, respectively, fidelity in the structure recovery and the contrast enhancement compared with the fog-free image.

Dehazed images are usually assessed visually. The lack of a reference image of an outdoor foggy image makes the assessment a challenging task. To evaluate the performance of a new proposed method, authors often use the blind measures introduced in [13] and used in a large number of works, such as the ones presented in [10, 35–37]. These measures account for the visible edges in the restored image compared with the foggy image. The index e calculates the rate of the new visible edges. This value underlines the ability of dehazing methods to recover edges hidden by fog. The index $\bar{r}$ reveals the contrast enhancement as the ratio between the gradient of the visible edges before and after dehazing. High visual quality of dehazed images is described by high values of e and $\bar{r}$.

Since the ground-truth images of both scenes are available, we use a set of metrics with reference to the fog-free image to benchmark the dehazing methods through metrological based measures. From this category, we select metrics that mainly consider the similarity of image structure and contrast, such as MS-SSIM [38], VIF [39] and VSI [40], and one that also takes into account color degradation, such as MS-iCID [41] and FSIMc [42]. It should be noted that a good reconstruction is indicated when the scores of MS-SSIM, VIF, VSI and FSIMc are close to one and the

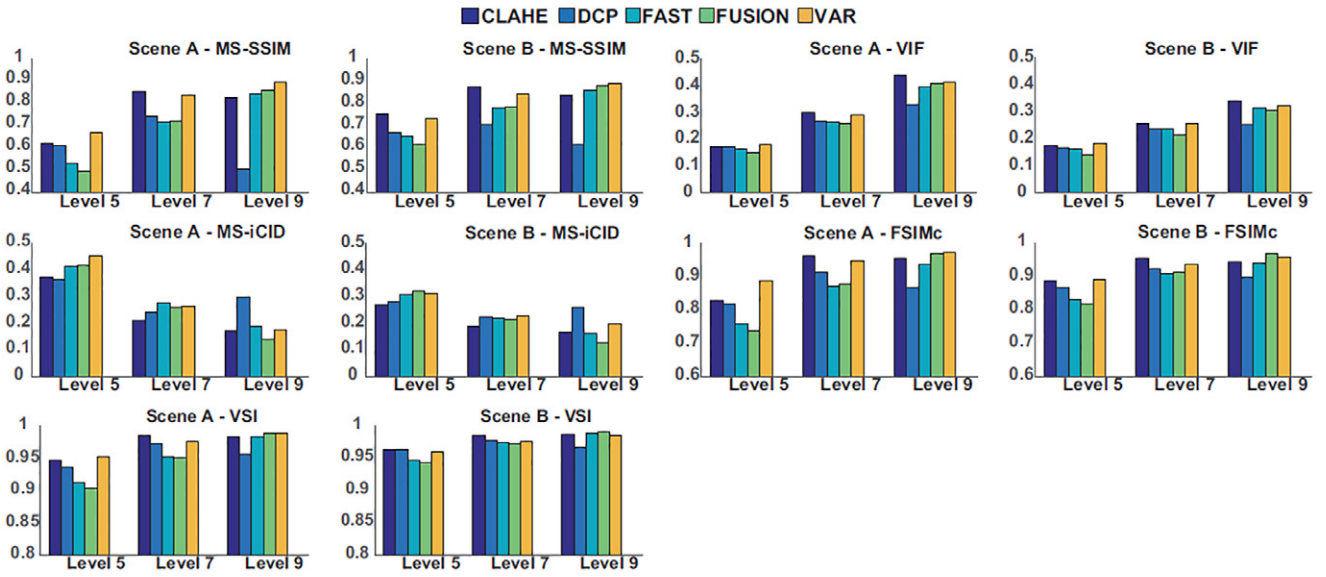

Figure 12. The scores of the metrics with reference to the fog-free image, which are obtained for each scene at each fog level. These metrics (MS-SSIM, VIF, MS-iCID, FSIMc and VSI) evaluate the quality of the dehazed image as being the similarity rate of specified features between the foggy and the fog-free images. All of these metrics underline a higher similarity rate at low fog densities [43].

Figure 13. Typical benchmark images for dehazing algorithms. The foggy images are processed by the selected dehazing methods. Column 1, foggy image; column 2, CLAHE result; column 3, DCP result; column 4, FAST result; column 5, FUSION result; column 6, VAR result. As the reference images are not available, if we consider the highest combination scores of e and $\bar{r}$, we may assume that FUSION followed by CLAHE, and VAR followed by DCP are the best performing methods for the New York and Train images, respectively.

score of MS-iCID tends toward zero. In Figure 12, we present the corresponding values of these metrics, which are already provided in reference [43].

6.2 *Objective Evaluation*

According to the scores of the metrics with reference to the foggy image presented in Fig. 12, we can globally notice that these metrics underline an increasing reconstruction quality of both structure and color when the fog density decreases. This means that dehazing methods lead to a good matching between the dehazed and reference images when the fog density is basically low. Although global quality is recovered better at low fog densities, it is still far from being perfectly recovered. While all dehazing algorithms show better performance at low fog levels, the physics based DCP method, according to its prior assumptions [8], seems to overestimate the fog when its density is basically very low. This produces gloomy images, as we see at level 9 in Figs. 10 and 11. This is mainly revealed by the metrics MS-SSIM, MS-iCID, FSIMc and VSI. Another deduction one can make over all of these metrics is that the ranking of dehazing methods is different between fog levels, in particular between distant levels (levels 5 and 9). For instance, considering MS-SSIM for scene A, from being ranked as the worst

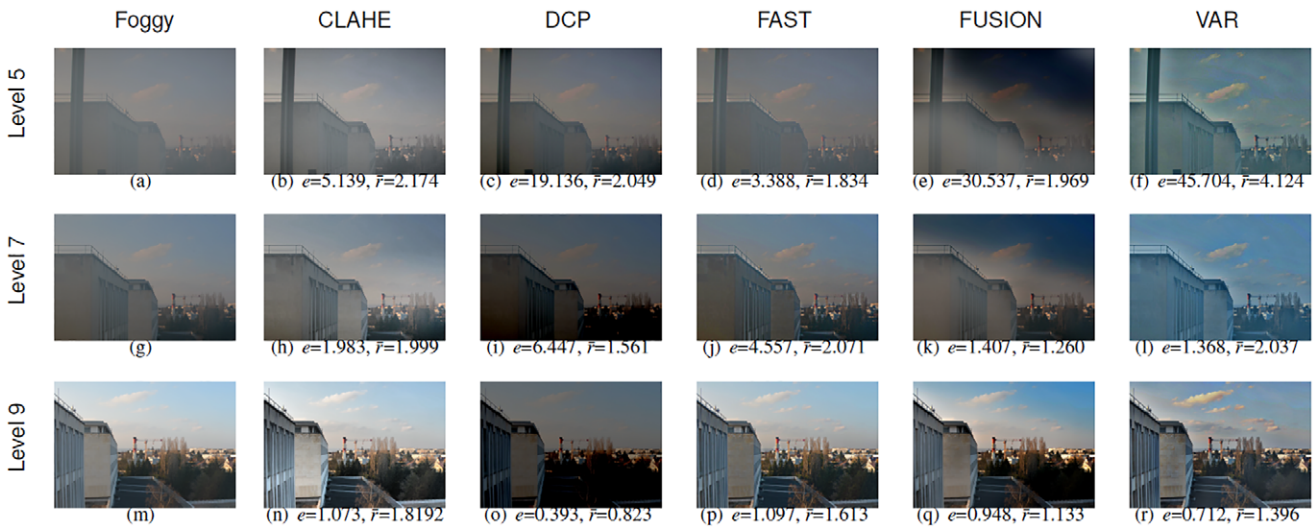

Figure 14. The foggy images of the dynamic scene processed by the selected dehazing methods at fog levels 5 (row 1), 7 (row 2) and 9 (row 3). Column 1, foggy image; column 2, CLAHE result; column 3, DCP result; column 4, FAST result; column 5, FUSION result; column 6, VAR result. The high values of e at high fog densities do not necessarily reflect a good performance. At such densities, the values of $\bar{r}$ are more significant. At level 9, the highest combination scores of e and $\bar{r}$ meet the visual judgment: the best restoration quality in terms of structure and contrast recovery is provided by CLAHE and then FAST.

performing dehazing method at level 5, the FUSION method switches to the second best method at level 9.

While all dehazing methods fail to achieve a perfect match between the reference fog-free and dehazed images, one can notice the particularly good behavior shown by CLAHE and VAR at all fog densities. It should be noted that both of these are image enhancement methods.

Considering e and $\bar{r}$, we calculate the corresponding scores for the processed images presented in Figs. 10 and 11. We also provide the scores for the fog-free image considered to have the best natural perceived quality. The e values at fog level 5 diverge remarkably from the value computed on the fog-free image. As e denotes the rate of the new visible edges after dehazing, it does not distinguish between edges in the scene and edges created by noise, which is likely to occur after dehazing process [44]. Relying on this, a high value, which is remarkably higher than the value computed on the fog-free image, does not denote a better quality but rather an image with undesirable artifacts. When the value is lower than the value of the reference image, this denotes a bad recovery of the scene edges hidden by the fog. This gap noticeably decreases when the fog is lighter and the artificial edges are fewer. When it comes to $\bar{r}$, at fog levels 5 and 7, the contrast stretching performed by any dehazing method is not better than the fog-free image. However, at level 9, the contrast is sometimes more enhanced than that of the fog-free image. This meets the earlier deductions on metrics with reference to the fog-free image about maintenance of fidelity at light fog levels. However, such deductions on e and $\bar{r}$ would never have been made if the reference image was not there, since the range of values cannot be easily associated to a perceived image quality. This is where the importance of such a database with reference images and several fog densities lies.

Moreover, we consider two typical benchmark images for dehazing algorithms, which are usually provided with a single fog density and no reference image. Their calculated e and $\bar{r}$ values are presented under the corresponding dehazed images in Figure 13. Since the reference images of these two scenes are not available, it seems to be extremely difficult to evaluate the image quality only according to the e and $\bar{r}$ values. We may consider that the highest combination scores of e and $\bar{r}$ indicate the best methods to recover structure and enhance the contrast: FUSION followed by CLAHE, and VAR followed by DCP for the New York and Train images, respectively.

Another important aspect provided by dynamic scenes involves investigation of dehazing method performance when handling different fog scenarios. Like the typical images presented in Fig. 13, dynamic images have no reference images. Figure 14 shows the dehazed images of the foggy dynamic scenes taken at 2.8 m (level 5), 1.4 m (level 7) and 0 m (level 9). Considering the large dimensions of the images, 6000×4000, these images have been cropped to 3329×2305. While all methods provide visually unpleasing results at level 5, we single out the particularly good behavior of all methods except VAR, as they do not accentuate the chromatic distortions induced by the window. Although it adds a color cast to the image, VAR improves the contrast (as can be seen by looking at the area with trees). This is denoted by the highest score of $\bar{r}$. However, the highest score for e, for the image processed by the FUSION method, does not indicate good image recovery. Using FUSION and FAST, the images remain partially foggy. All methods maintain the same handling with reduced unpleasant effects when the fog density is lower. At level 9, when all methods perform well, the highest combination scores of e and $\bar{r}$ meet the visual

judgment: the best restoration quality in terms of structure and contrast recovery is provided by CLAHE and then FAST.

All of the above results suggest the following.

- The performance of dehazing methods changes with fog scenario, fog density and image context.
- State-of-the-art blind quality metrics demonstrate a limited accuracy in dehazing assessment, especially for heavy fog. The reference image allows quantification of scores, which it is difficult to do otherwise.
- Structure fidelity and contrast enhancement are not similarly handled by dehazing methods, especially for heavy fog.

7. CONCLUSION

We have introduced the new CHIC database dedicated to dehazing and haze model evaluation. This database, which provides a set of images of different fog densities with the corresponding haze-free images, could be used for various purposes like haze model assessment, dehazing method evaluation across fog levels, and probably for the development of dehazing detection and classification methods. Compared with existing databases used for the assessment of visibility enhancement algorithms, CHIC provides the data required for reliable enhancement and comparison studies. Five different dehazing methods are benchmarked on three intermediate levels of fog using existing IQA metrics with reference to the provided fog-free images. Further focus should be given to the enhancement and development of dehazing methods able to maintain natural color fidelity with minimum noise amplification. For dehazing quality metrics, much work needs to be done to standardize reliable metrics.

REFERENCES

[1] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," *Proc. Seventh IEEE Int'l. Conf. on Computer Vision, 1999* (IEEE, Piscataway, NJ, 1999), Vol. 2, pp. 820–827.

[2] H. Koschmieder, *Theorie der horizontalen Sichtweite: Kontrast und Sichtweite* (Keim & Nemnich, Munich, 1925).

[3] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," IEEE Trans. Pattern Anal. Mach. Intell. **25**, 713–724 (2003).

[4] L. Schaul, C. Fredembach, and S. Süsstrunk, "Color image dehazing using the near-infrared," *Proc. IEEE Int'l. Conf. on Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2009), number LCAV-CONF- 2009-026.

[5] C. Feng, S. Zhuo, X. Zhang, L. Shen, and S. Süsstrunk, "Near-infrared guided color image dehazing," *20th IEEE Int'l. Conf. on Image Processing (ICIP), 2013* (IEEE, Piscataway, NJ, 2013), pp. 2363–2367.

[6] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proc. 2001 Conf. on IEEE Computer Society* (IEEE, Piscataway, NJ, 2001), Vol 1, pp. I–325.

[7] 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," *ACM Transactions on Graphics (TOG)* (ACM, New York, USA, 2008), Vol. 27, p. 116.

[8] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE Trans. Pattern Anal. Mach. Intell. **33**, 2341–2353 (2011).

[9] C. O. Ancuti and C. Ancuti, "Single image dehazing by multi-scale fusion," IEEE Trans. Image Process. **22**, 3271–3282 (2013).

[10] J.-P. Tarel and N. Hautière, "Fast visibility restoration from a single color or gray level image," *IEEE 12th Int'l. Conf. Computer Vision, 2009* (IEEE, Piscataway, NJ, 2009), pp. 2201–2208.

[11] Z. Xu, X. Liu, and N. Ji, "Fog removal from color images using contrast limited adaptive histogram equalization," *2nd Int'l. Congress on Image and Signal Processing, 2009. CISP'09* (IEEE, Piscataway, NJ, 2009), pp. 1–5.

[12] A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, "Enhanced variational image dehazing," SIAM J. Imaging Sci. **8**, 1519–1546 (2015).

[13] N. Hautière, J.-P. Tarel, D. Aubert, and É. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," Image Analysis and Stereology Journal **27**, 87–95 (2008).

[14] S. Fang, J. Yang, Z. Jiqing, H. Yuan, and R. Rao, "Image quality assessment on image haze removal," *Control and Decision Conf. (CCDC), 2011 Chinese* (IEEE, Piscataway, NJ, 2011), pp. 610–614.

[15] F. Guo, J. Tang, and Z.-x. Cai, "Objective measurement for image defogging algorithms," J. Central South University **21**, 272–286 (2014).

[16] J.-P. Tarel, N. Hautière, A. Cord, D. Gruyer, and H. Halmaoui, "Improved visibility of road scene images under heterogeneous fog," *2010 IEEE on Intelligent Vehicles Symposium (IV)* (IEEE, Piscataway, NJ, 2010), pp. 478–485.

[17] IFSTTAR. Frida (foggy road image database) http://www.lcpc.fr/english/products/image-databases/article/frida-foggy-road-image-database.

[18] J.-P. Tarel, N. Hautière, L. Caraffa, Aurélien Cord, H. Halmaoui, and D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," IEEE Intelligent Transportation Systems Magazine **4**, 6–20 (2012).

[19] N. Hautiére, J.-P. Tarel, J. Lavenant, and D. Aubert, "Automatic fog detection and estimation of visibility distance through use of an onboard camera," Mach. Vis. Appl. **17**, 8–20 (2006).

[20] N. Hautière, J.-P. Tarel, H. Halmaoui, R. Brémond, and D. Aubert, "Enhanced fog detection and free-space segmentation for car navigation," Mach. Vis. Appl. **25**, 667–679 (2014).

[21] J. El Khoury, J.-B. Thomas, and A. Mansouri, "Haze and convergence models: Experimental comparison," *In AIC 2015* (Tokyo, Japan, 2015).

[22] J. E. Khoury, J.-B. Thomas, and A. Mansouri, "Does dehazing model preserve color information?," *2014 Tenth Int'l. Conf. on Signal-Image Technology and Internet-Based Systems (SITIS)* (IEEE, Piscataway, NJ, 2014), pp. 606–613.

[23] S. G. Narasimhan, C. Wang, and S. K. Nayar, "All the images of an outdoor scene," *Computer Vision—ECCV 2002* (Springer, Copenhagen, 2002), pp. 148–162.

[24] WILD (Weather and Illumination Database). http://www.cs.columbia.edu/CAVE/software/wild/index.php.

[25] Image dehazing. http://ivrl.epfl.ch/supplementary_material/SFS--ICIP09.

[26] Waterloo ivc dehazed image database. http://ivc.uwaterloo.ca/database/Dehaze/Dehaze-Database.php.

[27] J. El Khoury, J.-B. Thomas, and A. Mansouri, "A color image database for haze model and dehazing methods evaluation," *International Conference on Image and Signal Processing* (Springer International Publishing, Trois-Rivières, Canada, 2016), pp. 109–117.

[28] J. Lüthen, J. Wörmann, M. Kleinsteuber, and J. Steurer, "A rgb/nir data set for evaluating dehazing algorithms," Electron. Imaging **2017**, 79–87 (2017).

[29] J. E. Khoury, Model and quality assessment of single image dehazing. http://www.theses.fr/s98153, 2016.

[30] User's manual, nikon d7100. http://cdn-10.nikoncdn.com/pdf/manuals/dslr/D7100EN.pdf.

[31] Konica minolta, cs-2000 spectroradiometer. http://sensing.konicaminolta.us/products/cs-2000-spectroradiometer/.

[32] C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," J. Appl. Photogr. Eng. **2**, 95–99 (1976).

[33] H. Kaur and R. Mahajan, "A review on various visibility restoration techniques," Int. J. Advanced Research in Computer and Communication Engineering **3** (2014).

[34] Y. Xu, J. Wen, L. Fei, and Z. Zhang, "Review of video and image defogging algorithms and related studies on image restoration and enhancement," IEEE Access **4**, 165–188.

[35] V. Jacob, W. de Dravo, and J. Y. Hardeberg, "Multiscale approach for dehazing using the stress framework," Electron. Imaging **2016**, 1–13 (2016).
[36] M. Qi, Q. Hao, Q. Guan, J. Kong, and Y. Zhang, "Image dehazing based on structure preserving," Optik-Intl J. Light and Electron Optics **126**, 3400–3406 (2015).
[37] Z. Gu, M. Ju, and D. Zhang, "A single image dehazing method using average saturation prior," Mathematical Problems in Engineering **2017** (2017).
[38] Z. Wang, E. P. Simoncelli, and A. C Bovik, "Multiscale structural similarity for image quality assessment," *Conf. on Signals, Systems and Computers, 2004. Conf. Record of the Thirty-Seventh Asilomar* (IEEE, Piscataway, NJ, 2003), Vol. 2, pp. 1398–1402.
[39] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Transactions on Image Processing **15**, 430–444 (2006).
[40] L. Zhang, Y. Shen, and H. Li, "Vsi: A visual saliencyinduced index for perceptual image quality assessment," IEEE Transactions on Image Processing **23**, 4270–4281 (2014).
[41] I. Lissner, J. Preiss, P. Urban, M. S. Lichtenauer, and P. Zolliker, "Image-difference prediction: From grayscale to color," IEEE Transactions on Image Processing **22**, 435–446 (2013).
[42] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," IEEE Transactions on Image Processing **20**, 2378–2386 (2011).
[43] J. El Khoury, S. Thomas, J. B. Thomas, and A. Mansouri, "Color and sharpness assessment of single image dehazing," Multimedia Tools Appl. 1–22 (2017).
[44] K. Ma, W. Liu, and Z. Wang, "Perceptual evaluation of single image dehazing algorithms," *2015 IEEE Int'l. Conf. on Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2015), pp. 3600–3604.