Heliyon 7 (2021) e08038

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Evaluation of the quality indicators in dehazed images: Color, contrast, naturalness, and visual pleasingness

Laksmita Rahadianti [a,*], Aruni Yasmin Azizah [a], Hilda Deborah [b]

[a] Faculty of Computer Science, Universitas Indonesia, Indonesia
[b] Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway

ARTICLE INFO

Keywords:
Dehazing
Image restoration
Hazy images
Image quality
Psychovisual experiment

ABSTRACT

Hazy images suffer from low quality due to blurring, veiling effects, and low contrast. To improve their visibility, dehazing methods attempt to restore them to their corresponding clear scenes, often by focusing more on obtaining an accurate estimate based on a known ground truth. The perceptual quality of dehazed images, which can be described by means of objective and subjective quality assessments, is often not considered. This paper provides a quality assessment of dehazed images, focusing on aspects, e.g., color, image structure, and naturalness. Four image dehazing methods are considered, i.e., Contrast Limited Adapted Histogram Equalization (CLAHE), Dark Channel Prior and Refinement (DCP-R), Perception Inspired Deep Dehazing Network with Refinement (PDR-Net) and Conditional Generative Adversarial Network (CGAN) Pix2pix. The dehazing results are then put through objective and subjective assessments, for a comprehensive evaluation on image quality. Overall, Pix2pix shows the best results objectively, excelling in the recovery of color and image structure. Although it is outperformed by DCP-R in terms of naturalness, our subjective assessment shows that Pix2pix is also most preferred by human observers.

1. Introduction

In a digital image acquisition, an image sensor captures light rays reflected by objects. In clear conditions, these light rays are able to travel unhindered in the environment. However, in conditions such as fog and haze, micro-particles in the surrounding media interfere with the propagating light by changing its direction and intensity. This type of media is often called *scattering* or *participating* media because it plays an active role in image capture. An image captured in scattering media, e.g., hazy images, will appear with blurring effects, low contrast, and low visibility. The appearance of hazy images makes it very difficult for observers to obtain useful information of a scene and the objects in it, making in challenging for image understanding by both human and computer vision. Thus, it is often necessary to *dehaze* a hazy image to its corresponding clear scene [1] prior to any further processing or analysis.

Image dehazing methods seek to restore a hazy image to its estimated clear counterpart, which can be achieved by, e.g., contrast or visibility enhancement techniques [2, 3]. Other approaches study the physical process of scattering model and then invert it with the help of additional constraints, e.g., statistical or spatial priors [4, 5]. Deep networks have also been used to model the translation between hazy and clear image domains [6, 7].

This paper is built on a preliminary study of the quality dehazed images by Azizah, et al. [8]. In this paper, we evaluate the image quality of dehazed images using four different dehazing methods. Contrast Limited Adaptive Histogram Equalization (CLAHE) [9] is an image enhancement technique, as opposed to a restoration technique, to improve quality and visibility. The dark channel prior (DCP) [5] uses the physical scattering model and attempts to reverse it to obtain the original clear image. Then, two deep learning based dehazing methods are also considered. Perception-inspired Single Image Dehazing Network with Refinement (PDR-Net) [10] is the state-of-the-art generative model that uses a perception-based approach to translate hazy images to their clear counterparts. Pix2pix [11] is a Conditional Generative Adversarial Network (CGAN) that has been used extensively for image to image translation. In our case, we will use it to translate hazy images to their clear counterparts.

Dehazing results are often evaluated by how close they are to their known ground truth. However, hazy images captured in natural and real

* Corresponding author.
E-mail address: laksmita@cs.ui.ac.id (L. Rahadianti).

https://doi.org/10.1016/j.heliyon.2021.e08038
Received 8 June 2021; Received in revised form 13 August 2021; Accepted 16 September 2021

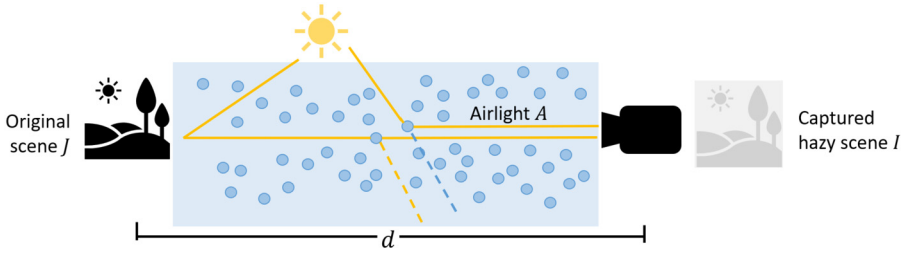


Fig. 1. Image formation model in scattering media.

settings often do not come with ground truth clear images, resulting in the lack of standardized and universal datasets. Additionally, the majority of existing works evaluate dehazed images by variations of objective quality metrics, making benchmarking of the quality of dehazed images very difficult to do [12]. In our work, we address the challenge by framing the evaluation of dehazing methods from a different standpoint. In addition to measuring accuracy with various objective quality metrics, we also carry out a subjective image quality assessment (IQA) based on *visual pleasingness*. We then study the correlation of both objective and subjective assessments, providing not only a more comprehensive IQA of dehazed images, but also analyses of their perceptual attributes. Regarding the dataset, aiming to encompass different types of hazy scenes, we use a compilation of various dehazing datasets and also generate a simulated dataset. We hope this paper will spark discussion and further research on the concept of quality in image restoration tasks such as dehazing.

2. On image dehazing

To dehaze hazy images to their clear form, it is necessary to study the physical process that captures a hazy image from a clear scene. This section provide the image formation model in scattering media, the state-of-the-art image dehazing methods, and the IQA approaches for dehazed images.

2.1. Hazy images

Digital images are formed when light rays are captured on a digital sensor grid of a camera. These light rays usually come from a primary light source, and are reflected by a scene into the lens of a camera. The light rays that are captured at different pixel locations build up an image representation of the scene. This process is quite straightforward in clear media, which allows the light rays to pass through unhindered. However, this process is not possible in scattering media environments as micro-particles contained in it may alter the direction of the traveling rays and *scatter* them in many directions. This will also result in a reduced light intensity because the media will *absorb* or divert the original light. The media actively interferes with image formation resulting in a below standard representation of the real scene. The altered intensities will thus compromise the visual features necessary for image understanding [13].

The image formation model used in this paper is based on the general atmospheric scattering model [14], see illustration in Fig. 1. This model assumes that the micro-particles in the media are very small, so the scattering can be considered homogeneous and has relatively low density. These assumptions hold in most natural hazy conditions. In these natural conditions, the appearance of objects in the image will differ based on their distance from the camera. The micro-particles in the media coupled with the amount of distance traveled, directly affect the amount of light that is able to penetrate the media [15]. The fraction of light that is able to penetrate the scattering media decay exponentially proportional to the distance from the camera d, and can be defined as:

$$t = e^{-\beta \cdot d} \tag{1}$$

where β is the scattering coefficient of the media, and d is the distance to the camera. In a homogeneous scattering media, the captured hazy image I of an original clear scene J can be defined using the transmission t as follows, where A is the *airlight*:

$$I = J \cdot t + A(1 - t) \tag{2}$$

The term airlight A refers to the color of ambient light in the scene [4]. This consists of the accumulation of scattering effects and carries no information about the original scene itself. Airlight is scattered in the scene, creating a veiling effect that subtly obscures the entire image [16]. The exact appearance and hue of the veiling effect will depend highly on the scattering media in the environment. For images captured in environments such as fog or haze, the airlight often presents a pale whitish color. As a final result, the images captured in these environments will not be able to portray the scene perfectly, due to obscured details, loss of intensity and low contrast. In this paper, these images will be referred to as *hazy images*.

2.2. Image dehazing

Image dehazing is an image restoration task, mapping hazy images to their clear counterparts with the highest possible accuracy. However, there are many cases in which a precise ground truth is not available. Thus, we restate the image dehazing process to work towards the achievable goal of obtaining a *visually pleasing* image with better quality. The end result should be just *good enough* so that observers are able to understand the scene through its visual features. A visually pleasing image does not necessarily need to be an exact match to a known standard, thus it can be achievable using image enhancement methods. These methods have the advantage of not requiring a known clear target, as their aim is to solely reduce the hazy effects such as described in Section 2.1. These methods attempt to improve the visibility of the image through methods such as color correction [3, 17], contrast correction [2, 18], contrast balancing [9, 19], and others.

It is also possible to model the physical transformation between hazy and clear images. The model can then be used to restore hazy images to their clear form using single image dehazing techniques. Single image dehazing considers the physical scattering model described in Section 2.1 as the process in which a scene becomes hazy. Then, single image dehazing follows the image formation model of hazy images in Eq. (2) and uses its inverse to extract the original scene J from the haze image I following Eq. (3). This equation involves many unknowns. Thus, single image dehazing methods must estimate the unknown variables in phases, i.e., estimating the ambient light or airlight $\hat{A}$, predicting transmission map $\hat{t}$, then recovering the original scene $\hat{J}$ (Fig. 2).

$$\hat{J} = \frac{I - A}{t} + A \tag{3}$$

Unfortunately, the transmission t or the distance d are rarely known. Furthermore, the scattering coefficient of the media β can be different for every scene, hindering a one-fits-all estimate for all images. To solve this issue, it is necessary to employ additional constraints to the model. One possible approach is to use statistical priors, i.e., the Dark Channel

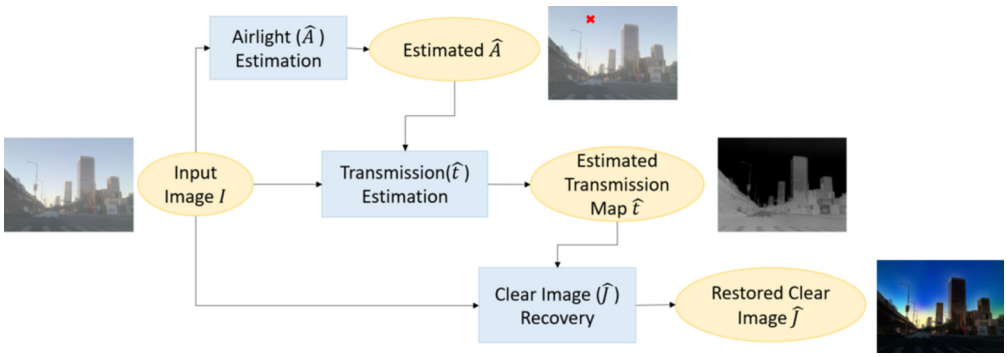


Fig. 2. General step-by-step process of single image dehazing.

Fig. 3. An example of a clear scene (**a**) and its depth map (**b**) from the NYU Depth Dataset [21], and the (**c**) synthesized hazy image generated based on Eq. (2).

Prior (DCP) [5] or Adaptive Dark and Bright Channel Prior [20]. These prior observations of natural hazy images form pre-determined information to help estimating variables involved in an image formation.

The problem of image dehazing can be modeled as an image to image translation problem, e.g., from the set of hazy images to their clear counterparts. There have been various deep networks that proposed this purpose, using various structures of convolutions [6, 22], attention cues [23] or even human perception cues [10]. In recent years, these deep network based approaches have been popular for other image to image translation problems as well. Image translation can be useful for various applications, e.g., semantic segmentation [24], domain adaptation [25], or image to depth estimation [26].

2.3. Image quality assessment

Image quality assessment (IQA) tries to measure the quality of digital images in depicting the original scene, and in turn, in relaying information. It can be divided into subjective and objective approaches. Objective IQA involves image quality metrics to compute and depict image quality quantitatively. These metrics can be divided into full-, reduced, or no-reference metrics. *Full-reference (FR) metrics* need a ground truth image, such that the difference or error from the target image can be computed [27]. *Reduced-reference (RR) metrics* can be used when only incomplete references are available, and *no-reference (NR) metrics* when there are none at all. NR metrics are also referred to as *blind image quality metrics* [28]. Subjective IQA gives the task to human observers, to evaluate the quality of a set of images based on a certain criteria by, typically, indicating their preferences. The human visual system is considered the most reliable biological image capture device able to judge visual image quality in a consistent manner [29].

The main issue for the IQA of dehazed images is the lack of standardized evaluation protocol. First of all, hazy images themselves are difficult to come by, and there is a lack of datasets that can be used for quality benchmarking. Furthermore, it is very difficult to obtain a reference set of hazy images and their exact clear counterpart. Natural haze occurs in outdoor environments where many factors are involved in image capture. Most of these factors are uncontrollable for experimental data capture and the scene is never reliable enough to capture in both hazy and clear conditions. The next best option is capturing images using synthetic haze, but it is still difficult to obtain the exact same condition to capture the scene in both conditions. This work would also highlight the need to consider IQA for dehazed images. Natural atmospheric scattering [14] occurs outdoors even in the clearest of conditions, thus human observers expect *some* haze in any natural image. Meanwhile, objective metrics would see any remaining haze as an indication of low quality. Thus, it is unrealistic to rely on accuracy alone, since it is important to consider various factors of image quality depending on the context or aim of the task.

3. Materials and methods

This section will describe the image dehazing experiments conducted for a test set of 50 images, followed by a thorough explanation of the used IQA methods.

3.1. Dataset

In our experiments, a combined dataset of hazy images with ground truth clear images from 3 different sources were used. The entire dataset is comprised of 600 hazy-clear image pairs, which then divided into a training set of 550 images to train the PDR-Net (Section 3.2.3), and a test set of 50 images. Only the test set was used to perform IQA on the dehazing results.

3.1.1. Synthetic hazy images

The New York University (NYU) depth dataset is comprised of 1449 pairs of RGB-Depth (RGBD) data, i.e., indoor scenes and their corresponding depth maps [21], see an example in Fig. 3a and 3b. They are captured in various commercial and residential buildings, with depth maps obtained using the Microsoft Kinect.

For our experiment, we generated synthetic hazy images from 210 randomly selected RGBD pairs from the NYU dataset following Eq. (2), see example in Fig. 3c. First, the pixel-wise depth map d were used to create a dense transmission map t using several scattering coefficient values $\beta \in \{0.1, 0.2, 0.3, 0.4\}$. Then, t was used to synthesize hazy images I using different airlight values $A \in [0.7, 1]$ [26, 30].

3.1.2. Hazy series dataset

Four datasets are available from The New Trends In Image Restoration And Enhancement Workshop And Challenges (NTIRE) challenge [31]. The Indoor Haze (I-Haze) dataset [32] consists of hazy-clear image pairs captured in indoor conditions, while the Outdoor Haze (O-Haze) dataset [33] was captured in outdoor conditions. In both datasets, haze was generated using two professional fog machines (LSM1500 PRI) which was dispersed using fans to ensure homogeneity. The Dense-Haze [34] and Non-Homogeneous Haze (NH-Haze) [35] datasets are both extensions of the O-Haze dataset. The haze from the fog machines was released for a longer duration to create a denser haze effect for the Dense-Haze dataset. For NH-Haze, the generated haze was not dispersed evenly through the scene ensuring a non-homogeneous haze condition. A total set of 180 hazy images and their clear counterparts was collected from all four of these datasets.

3.1.3. REalistic single-image DEhazing dataset

Li, et al. [12] created the REalistic Single-Image DEhazing (RESIDE) dataset with the intention to provide a standardized large-scale dataset of hazy-clear image pairs to facilitate the benchmarking of dehazing methods. RESIDE consists of realistically simulated hazy images and is organized into 3 subsets, i.e., the Indoor Training Set (ITS), Synthetic Outdoor Training Set (SOTS), and the Hybrid Subjective Testing Set (HSTS). Among all three subsets, SOTS is the most suitable for our needs as it has many images of urban outdoor hazy scenes. For our experiment, 210 outdoor image pairs from the SOTS subset were randomly selected.

3.2. Dehazing methods

Four methods are used in our experiments. They are an image enhancement approach, i.e., CLAHE [9], a single image dehazing model using dark channel prior (DCP) [5], and two deep learning based dehazing models. The deep learning based models selected encompass a generative model PDR-Net [10] and an adversarial model Pix2pix [11].

3.2.1. Contrast limited adaptive histogram equalization (CLAHE)

One of the effects that occur in hazy images is low contrast and loss of detail. This is why hazy images are so difficult to perceive, because contrast carries significant information about image structure [36]. Thus, to improve visibility, we can attempt to improve or correct the contrast. Contrast correction can be done by balancing the distribution of intensity values in an image. The histogram of a low contrast image will be concentrated in a narrow range of values, making differentiating intensity values difficult.

We can improve image contrast by mapping the intensity values to a wider range through histogram equalization. The increase of difference between individual intensities will amplify edges and image structures, hereby improving visibility. There have been various proposals towards histogram equalization, and in this paper, we select the well-known and easily implemented CLAHE [9]. CLAHE separates an image into different contextual regions, and performs an adaptive histogram equalization on each region separately. Additionally, it also clips the intensity distribution at a certain limit to prevent over-saturation in homogeneous areas. Our experiments involve color images, so we first transformed the RGB color image to the CIELAB color space [37] and CLAHE was performed on the brightness channel L^* [9].

3.2.2. Dark channel prior and refinement (DCP-R)

The image formation model in Eq. (2) shows the transformation from a clear image J to a hazy image I, which we would like to reverse. However, this equation involves many unknown variables, making it difficult to solve. Additional constraints are needed to allow reversing the equation to estimate I from J. In this paper, we select the Dark Channel Prior (DCP) [5] which is widely acknowledged and used as a basis of many dehazing works.

DCP is a statistical prior based on the observation of natural outdoor images. It stipulates that for local patches of non-sky areas in natural images, there is always one color channel that has a very low value or $\min(I^s) \approx 0$, $I^s \in [R, G, B]$. These *dark pixels* occur due to natural dark objects, lighting, and distinctly colored objects. In such cases, there is often one channel with low intensity, while other channels contain the color information, thus creating the aforementioned *dark channel*.

Dark channel typically occurs in natural clear images. In hazy images, a dark channel becomes saturated with the effect of haze, which can be denoted as a proportion of airlight intensity A from the image formation model in Eq. (2). Based on the concept of transmission in Eq. (1), it is known that airlight will aggregate in the dark channel, with higher quantity indicating further distance [5, 38]. Consequently, the distance value will also grow proportionally with the value of DCP, allowing it to be used as an indicator of relative depth. He, et al. [5] proposed a method to estimate transmission based on DCP as described in Eq. (4). Assuming an RGB hazy image I in the RGB color space ($s \in [R, G, B]$), the transmission $\hat{t}$ can be estimated for every pixel x in the hazy image I considering every pixel y in the local area $\Omega(x)$ surrounding x.

$$\hat{t}_{DCP}(x) = 1 - \min_{s \in \{R,G,B\}} \left(\min_{y \in \Omega(x)} \left(\frac{I^s(y)}{A^s} \right) \right) \tag{4}$$

Note that the value of airlight A in the RGB color space ($s \in [R, G, B]$) or its estimation is needed to compute $\hat{t}$. There are various approaches that can be used to estimate A, the simplest of which is by using the DCP value itself. He, et al. [5] set the value of airlight with the pixel value at the location of the maximum DCP. Airlight, or ambient light, is often equated with the color of the sky, which can be found at a maximum distance $d \approx \infty$. Since a larger DCP value would indicate a further distance, the maximum value of DCP should indicate the furthest point as well, at which point can be used as an estimate of $\hat{A}$. With the estimated transmission $\hat{t}$ based on Eq. (4), the original scene J can be recovered for every pixel x based on the principles of DCP according to Eq. (5). To avoid a division by zero, it is necessary to limit the values of $\hat{t}$ by a minimum value of t_0.

$$\hat{J}^s(x) = \frac{I^s(x) - A^s}{\max\left(\hat{t}(x), t_0\right)} + A^s \tag{5}$$

In our experiments, image dehazing using DCP was performed by a two-step approach. The first step results in an approximated clear image of the scene. However, since DCP assumes constant transmission in local areas, these images suffer from the loss of details or appearing to be patchy. Thus, a refinement step was added using the guided filter [39], aiming to further improve and smooth the dehazed images. With an edge-preserving smoothing property, the guided filter uses a reference image to guide the refinement process, resulting in a linear transform of the guidance image. Finally, the complete single image dehazing approach in this paper is a combination of DCP and refinement using guided filter, further referred to as DCP-R, and it was used to dehaze the 50 hazy images from the test set.

3.2.3. Perception-inspired dehazing network with refinement (PDR-Net)

Li, et al. [10] proposed PDR-Net, a perception-based deep learning architecture to model end-to-end image to image translation between hazy images and their clear counterparts. Similar to DCP-R, it treats the process of dehazing as two steps, i.e., haze removal and refinement, each implemented as two separate sub-networks. PDR-Net employs a perceptual loss in the haze removal sub-network, to describe higher-level perceptual cues. Meanwhile, the refinement sub-network employs a multi-term loss to recover color distortion and enhance the visual quality of the dehazed image. Through the sub-networks, PDR-Net dehazes images in two steps, resulting in not only an accurately dehazed image, but also a visually pleasing one.

In our experiments, PDR-Net was trained using the training set of 550 hazy-clear image pairs. Since deep architectures often benefit from

more samples, the training set was augmented by flipping each image horizontally, creating a total of 1100 training images. Following the procedure detailed in [10], each sub-network was trained separately, with 700 training images to train the haze removal sub-network, and 400 for the refinement sub-network. Each sub-network was trained for 200 epochs. As with our DCP-R, PDR-Net was also used to dehaze 50 hazy images of the test set.

3.2.4. Pix2pix

Common generative networks may experience difficulty in training image to image translations due to the large number of variables and probabilistic computations to estimate. Thus, the Generative Adversarial Networks (GAN) [40] framework may be used to assist with training these generative networks. A GAN subjects a generative model to a discriminative model, trained to work as opponents. The generator is trained to generate realistic renderings of the intended outputs, while the discriminator aims to detect these reconstructions from the ground truth output [40]. A Conditional Generative Adversarial Network (cGAN) extends the conventional GAN framework, i.e. the generator and discriminator are conditioned with extra information to guide the generative process [41].

Pix2pix is a cGAN model proposed for general non-specific image to image translation task [11]. Pix2pix uses a modified U-Net [24] as its generative model, guided by a PatchGAN discriminator [42]. Pix2pix was thoroughly examined using various image to image translation tasks, such as grayscale to color images or day to night images. The results show its ability to reproduce a visually similar output to the target for seemingly arbitrary image pairs. Considering a dehazing task, we are optimistic that the discriminator component can improve the visual pleasingness of the resulting image. Since the discriminator judges the generated output based on overall similarity, not only on a minimized error, we believe this may mimic the observation or judgement of a human observer. In our experiments, we also used Pix2pix to train image to image translation from hazy images to their clear counterparts. Following the training and testing setup of PDR-Net, Pix2pix was also trained for 200 epochs using the same augmented hazy-clear image pairs, with 1100 training images. The final trained model was then used to dehaze 50 hazy images of the test set.

3.3. IQA of dehazed images

The dehazing results were then assessed by means of objective and subjective IQAs. The image quality metrics and the design of the psychovisual experiment for the subjective IQA are detailed in this section.

3.3.1. Objective metrics

In this study, four IQA metrics were selected to evaluate various quality aspects of an image. Three metrics are full-reference metrics that assign a quantitative value to the accuracy of dehazing compared to the ground truth, based on the quality cues focused on in this paper, namely color, image structure, and pixel intensity.

First, we evaluate color using ΔE^*_{ab}, a color difference function defined for the CIELAB color space [37], where perceptual non-uniformities are accounted for. A hazy image I and its dehazed version J will be converted to the CIELAB color space before the computation of their color difference as in Eq. (6). Images in the CIELAB color space have three components or channels, i.e., L^* (lightness), a^* (green-red chromaticity), and b^* (blue-yellow chromaticity). The range of color difference values from ΔE^*_{ab} is $[0, 100]$, which we rescale to $[0, 1]$, with a lower value indicating better quality.

$$\Delta E^*_{ab}(I,J) = \sqrt{(L^*_J - L^*_I)^2 + (a^*_J - a^*_I)^2 + (b^*_J - b^*_I)^2} \tag{6}$$

To evaluate the accuracy of pixelwise image intensities, we use the root mean square error (RMSE), shown in Eq. (7). RMSE computes the differences in image intensity or brightness levels [43]. Assuming pixel values in the range of $[0, 1]$, the RMSE of the entire image is averaged over the total number of pixels in the image (n). Thus, the RMSE will also range between 0 and 1, with lower value indicating a better estimation of the clear image.

$$\mathrm{RMSE}(I,J) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(I_i - J_i)^2} \tag{7}$$

The accuracy of the pixel intensities themselves, does not ensure a clear distinct structure in an image. Thus, we also used the structural similarity index measure (SSIM) [44] to evaluate the image structure recovered. SSIM is used to define the perceptual difference of two images by means of their luminance (l), contrast (c), and structural (s) components. The formula of SSIM is provided in Eq. (8), with μ, σ^2_I, and σ_{IJ} as the measure of average, variance, and covariance between I and J, respectively. Other parameters in the formula are $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$, $c_3 = c_2/2$, $L = 2^b - 1$, $k_1 = 0.01$, $k_2 = 0.03$, (α, β, γ) as adaptable weights, and b as the number of bits per pixel in an image which will typically be 8.

$$\begin{aligned} &\mathrm{SSIM}(I,J) = [l(I,J)^\alpha \cdot c(I,J)^\beta \cdot s(I,J)^\gamma], \text{ where} \\ &l(I,J) = \frac{2\mu_I\mu_J + c_1}{\mu_I^2 + \mu_J^2 + c_1}, c(I,J) = \frac{2\sigma_I\sigma_J + c_2}{\sigma_I^2 + \sigma_J^2 + c_2}, s(I,J) = \frac{\sigma_{IJ} + c_3}{\sigma_i\sigma_J + c_3}. \end{aligned} \tag{8}$$

The range of values that the SSIM can take on is $[-1, 1]$, with a larger value indicating better quality. Meanwhile, with all of the other metrics used in this study, better quality is indicated with a lower value. Furthermore, SSIM is also the only metric that can be represented with a negative number. Thus, the SSIM is converted into an inverted SSIM which will be denoted $SSIM^{-1}$ with a modified range of $[0, 1]$, with a lower value indicating better quality based on Eq. (9).

$$\mathrm{SSIM}^{-1}(I,J) = 1 - \left(\frac{\mathrm{SSIM}(I,J) + 1}{2}\right) \tag{9}$$

Belonging to the category of FR metrics, ΔE^*_{ab}, RMSE, and SSIM^{-1} all require two images as input, i.e., ground truth or reference image and the target of evaluation. Note that the image dimension of both images have to be exactly the same. These three metrics measure the accuracy of the dehazed image compared to the known ground truth.

We have already suggested that often, accurate images do not necessarily have better visual quality. Frequently, human perception will consider abstract concepts such as naturalness. Thus, we also used a metric that was proposed to measure naturalness, The Naturalness Image Quality Evaluator (NIQE) [45]. NIQE is an NR metric that computes the score of an image against a model of natural scene images. This model is represented as multidimensional Gaussian distributions. The range of values that NIQE can take on is $[0, \infty)$, with a lower value indicating better image quality.

3.3.2. Psychovisual experiment

Based on the objective evaluation result in Section 3.3.1, we could naively infer that a dehazed image with a good objective result indicates that it is visually pleasing to the human eye. The same principle should intuitively apply to the opposite case. However, subjective and objective measures do not always align [44]. Thus, subjective IQA by means of a psychovisual experiment was also carried out for the test set of 50 images.

Dehazing results of the four dehazing methods were sent out in a survey in which observers were asked to identify their preferred result. The display of the psychovisual experiment can be seen in Fig. 4. Each image is shown on an individual slide, with a neutral gray background and the slide number written at the top-left of each slide indicating which image is currently displayed. On each slide, five images are displayed, i.e. the original hazy image in the first row and four dehazed images on the second row, each obtained using CLAHE, DCP-R, Pix2pix, and PDR-Net, respectively. Observers were then asked to rank the four dehazed images from the image that is most visually pleasing to the

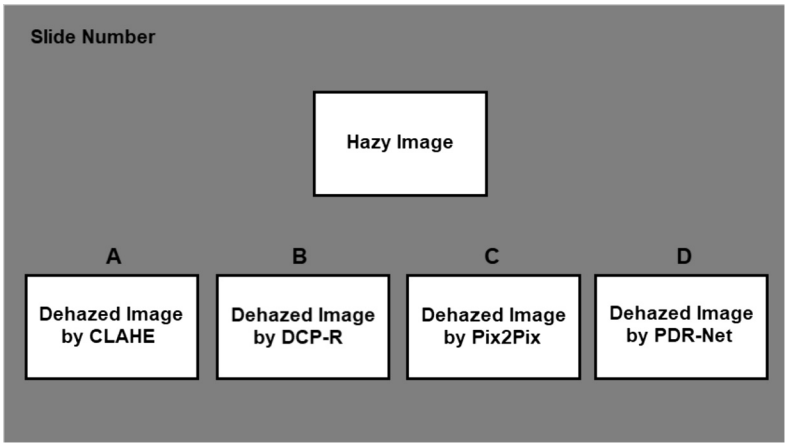


Fig. 4. Image arrangement for the psychovisual experiment.

least. Considering observer fatigue, each observer was only shown 25 randomly selected images from the entire test set of 50 images. Observers were free to use any criteria they deem important to determine their preference.

Responses from observers were then populated using the Mean Opinion Score (MOS) [46]. The MOS of image I is determined by the scores S assigned to it by N observers, which can be computed based on Eq. (10).

$$\text{MOS}(I) = \frac{1}{N}\sum_{n=1}^{N} S_n \tag{10}$$

In our experiment, the scores were given based on the rank order given by human observers. For each response, a score of 1 would be assigned to the most preferred result, 2 to the next, 3 to the third, and 4 to the least preferred result for that particular image I_i. Thus, a lower score will indicate better perceived quality, which is consistent with the objective metrics in Section 3.3.1. The score for each image was then averaged over the number of its observers N. The average score was then further converted to a modified range of $[0, 1]$. This was computed for all 50 test set images I_i, for $i \in [1, 50]$. At the end of the experiment, observers were also asked to give open-ended comments to elaborate further about their observations about the dehazed images. These comments may give an indication about the factors that influenced their preference in terms of color, structure, clarity, or naturalness.

4. Results and discussion

This section will display, evaluate, and analyze the results using the four dehazing methods in our experiments on the test set of 50 images described in Section 3.2. A randomly selected subset of the results are shown in Fig. 5. From the results, it is visible that all methods performed well in light and simulated haze such as in the O-Haze, Synthetic, and RESIDE images. However, it is more challenging to handle images from the Dense-Haze and NH-Haze datasets.

In the case of test images with dense haze, none of the four dehazing methods were able to recover the scene correctly, as shown in the Dense-Haze column in Fig. 5. Fig. 6 shows more details of the recovered Denze-Haze images, which are still shrouded in a veiling haze. CLAHE was the least successful, due to the over-saturated haze that dominates the image and hence its histogram. DCP-R was more successful in recovering details of the scene and the results shows more saturated colors. However, these colors are far from colors in the ground truth image. PDR-Net produces a darker image with less saturated colors. It also appears to have a lot of texture similar to that of the tree object in the ground truth image. However, not much information was successfully restored aside from the white structures at top right corner of the image. The dehazed image using Pix2pix was able to reconstruct the area of the leaves, but also produced color artefacts in the haze areas of the image.

Table 1. Average objective and subjective IQA metric results on the test set using 4 dehazing methods. All metrics have been converted and scaled so that the lower value will indicate better quality.

Objective IQA					
Metric	Range↓	CLAHE	DCP-R	PDR-Net	Pix2pix
RMSE	[0, 1]	0.4221	0.3361	0.3729	0.2367
$SSIM^{-1}$	[0, 1]	0.2206	0.2425	0.2292	0.1919
ΔE^*_{ab}	[0, 1]	0.2391	0.2291	0.2194	0.1514
NIQE	$[0, \infty)$	3.1521	2.8290	4.7216	2.8626
Subjective IQA					
Metric	Range↓	CLAHE	DCP-R	PDR-Net	Pix2pix
MOS	[0, 1]	0.4461	0.4877	0.8447	0.2220

Section 2.1 shows that the basic hazy image formation model itself is built on the assumption that scattering occurs homogeneously. Thus, images with non-homogeneous haze pose a challenge to common dehazing methods. See NH-Haze column in Fig. 5, where CLAHE, DCP-R, and Pix2pix were all unable to fully remove the haze, and the dehazed images still contain haze in areas where the haze is thicker. Meanwhile, the result of PDR-Net was successfully void of haze effects, albeit with low color saturation. Detailed observations of the results on NH-Haze can be seen in Fig. 7.

Objective evaluation of the dehazing results for the four dehazing methods is presented in Table 1. Note that all objective metrics have been converted and scaled so that a lower value indicates a better reconstruction. Overall, Pix2pix is able to obtain the best score in RMSE, $SSIM^{-1}$, and ΔE^*_{ab}. The best NIQE score is achieved by DCP-R, showing its superiority in terms of naturalness. Furthermore, all other methods come in second in different aspects, where CLAHE has the second best $SSIM^{-1}$ score, DCP-R in RMSE, and PDR-Net in ΔE^*_{ab}. CLAHE is a method that modifies the image very minimally, explaining the good $SSIM^{-1}$ score. DCP-R has a good RMSE, indicating a superior pixel intensity recovery. However, DCP-R does suffer from color distortion, explaining a lower ΔE^*_{ab} score compared to PDR-Net.

The psychovisual experiment of the 50 test images was conducted with a total of 40 human observers. The average MOS for each method is shown in the bottom row of Table 1, indicating a preference for dehazing by Pix2pix reflected in its lower average MOS. CLAHE and DCP-R come in second and third, with PDR-Net as the least preferred dehazing method. The psychovisual experiment also provided an opportunity for observers to add open-ended comments, which may further provide insights on the possible perceptual attributes for the perception of dehazed images. They will be valuable for a deeper analysis of visual cues affecting the perceived image quality and potential improvements

Fig. 5. Examples of dehazing results of test images from various datasets, i.e., Dense-Haze, NH-Haze, O-Haze, Synthetic, and RESIDE datasets. The original hazy image is shown in the first row, followed by the ground truth clear image in the second. The dehazed images obtained using CLAHE, DCP-R, PDR-Net, and Pix2pix are shown in row 3-6.

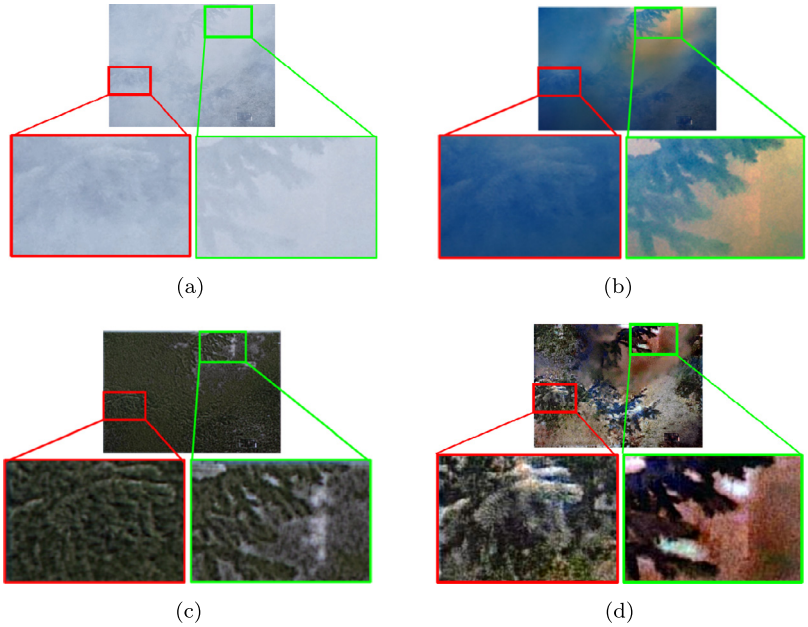


Fig. 6. Detailed comparison of a result of **(a)** CLAHE, **(b)** DCP-R, **(c)** PDR-Net, and **(d)** Pix2pix for an image with dense haze. See the corresponding original and ground truth images in Dense-Haze column in Fig. 5.

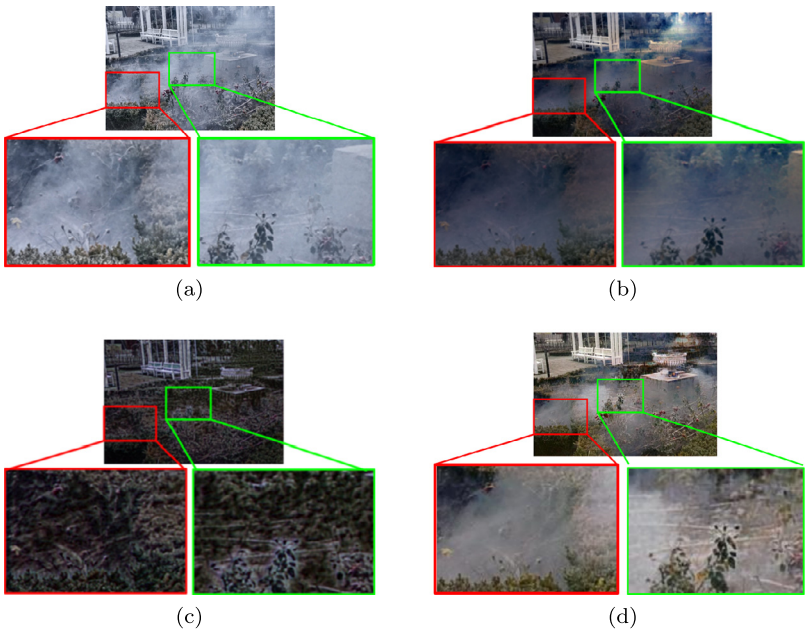


Fig. 7. Detailed comparison of a result of **(a)** CLAHE, **(b)** DCP-R, **(c)** PDR-Net, and **(d)** Pix2pix for an image with non-homogeneous haze.

Table 2. Insights from the open-ended comments provided by observers in the psychovisual experiment.

No.	Comment
1	Images dehazed with CLAHE still contain a very visible haze, indicating a less successful dehazing process.
2	Images dehazed with CLAHE are still more visually pleasing regardless of the haze, because they do not contain noise or artifacts.
3	The leftover haze in CLAHE images actually contributes to the naturalness of the image.
4	Images dehazed with DCP-R are generally bright, smooth and sharp. DCP-R seems to perform very well when the haze in the original image is light.
5	Images dehazed with DCP-R often have color distortions making it less visually pleasing. This is especially visible in outdoor scenes with have unnatural sky color.
6	Images dehazed with DCP-R sometimes appear dull and dark with some scenes.
7	Several images dehazed with PDR-Net have spots, artifacts, or noise. Although some produce colors that are similar to real objects in the scene but the artifacts makes PDR-Net the least visually pleasing.
8	PDR-Net results seem to have lost the sense of depth of the scene, due to a very high contrast.
9	PDR-Net and Pix2pix both are able to eliminate more haze, but are visibly unnatural, making observers uncomfortable.
10	Images dehazed with Pix2pix tend to have softer colors which make it more visually pleasing.
11	Images dehazed with Pix2pix are most often more visually pleasing compared to the other methods, aside from some unnatural colors.
12	Clear objects and minimal artifacts/distortion are a large factor in evaluating visually pleasing images. That is why CLAHE and Pix2pix images were frequently placed at the top rank.

for dehazing methods. Some interesting points about the observations that can be concluded from these comments are shown in Table 2.

As described before, the concept of image quality is very complex. The quantitative metrics that are commonly used to describe accuracy of dehazed images, are not always adequate to denote its quality. This is apparent in the objective and subjective results presented in this section. Fig. 8 shows the scatterplot of MOS against all 4 objective metrics, i.e. RMSE, $SSIM^{-1}$, ΔE^*_{ab}, and NIQE. For visualization purposes, each objective metric is scaled to the range of $[0, 1]$. We then attempt to infer the correlation between MOS and each objective metric, using curve fitting with a polynomial regression of order 1. The plot shows weak correlation of MOS with ΔE^*_{ab} and $SSIM^{-1}$, whose line plots are almost horizontal, indicating no relation. Meanwhile, the MOS scores

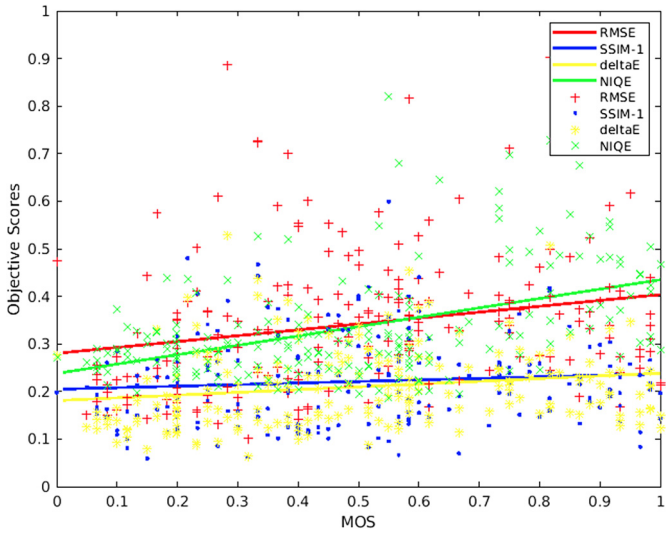


Fig. 8. Scatterplot between the subjective MOS and RMSE in red, $SSIM^{-1}$ in blue, ΔE^*_{ab} in yellow, and NIQE in green. The correlation line for each metric is obtained through curve fitting with polynomial regression order 1.

with RMSE and NIQE show a stronger correlation through lines that incline visibly.

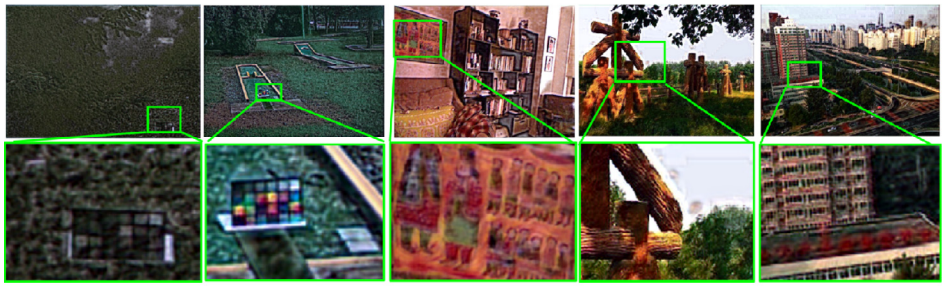
To measure the correlation quantitatively, we also computed the Pearson r correlation coefficient between MOS and each objective metric [47]. The Pearson r correlation coefficient is commonly used to quantitatively represent the relationship between variables [47]. The Pearson r coefficient of MOS with RMSE, $SSIM^{-1}$, ΔE^*_{ab}, and NIQE are shown in Table 3. From the results, we are able to confirm the insights we obtained from Fig. 8. All 4 objective metrics are positively correlated with MOS, to different degrees of strength. Among them, NIQE is the objective metric that has the strongest correlation with subjective MOS with a Pearson r coefficient of 0.4486.

4.1. Color analysis

The color quality of the dehazed images can be objectively measured using the color difference metric ΔE^*_{ab}. From Table 1, Pix2pix is clearly

(a)

(b)



(c)

(d)

Fig. 9. Detailed areas of the dehazed images using **(a)** CLAHE, **(b)** DCP-R, **(c)** PDR-Net, and **(d)** Pix2pix in terms of color and color recovery, contrast, and smoothness.

Table 3. Pearson r correlation coefficient between MOS and 4 objective metrics.

	Metric paired with MOS			
	RMSE	$SSIM^{-1}$	ΔE^*_{ab}	NIQE
Pearson r	0.2284	0.0914	0.1856	0.4486

ahead with an average ΔE^*_{ab} of 0.1514 followed by PDR-Net, DCP-R, and CLAHE, in that order. This can also be confirmed by visually observing the dehazing results shown in Fig. 9. The two deep learning methods are in the lead in this criterion, as they both learn complex mappings for pixel-based regressions from input to output images with little external factors.

CLAHE enhances the image by re-distributing the intensities without correcting the color which is apparent Fig. 9a and Fig. 10a. In the two first images of Fig. 9d, where Macbeth ColorCheckers are present, Pix2pix shows its capability to recover the colors on the color checkers. DCP-R and PDR-Net also do well restoring the vividness of the color in the color checkers in the first two images. For the selected areas in images in columns 3-5 of Fig. 9b, DCP-R shows that its recovery of colors results in higher saturation while maintaining smoothness. However, DCP-R is highly reliant on the airlight of the scene, which is frequently inaccurately estimated. This error manifests in the unnatural sky colors, which we highlight in the first two images of Fig. 10b For the simulated dataset involving indoor images in column 3, these color distortions result in color artifacts such as in the

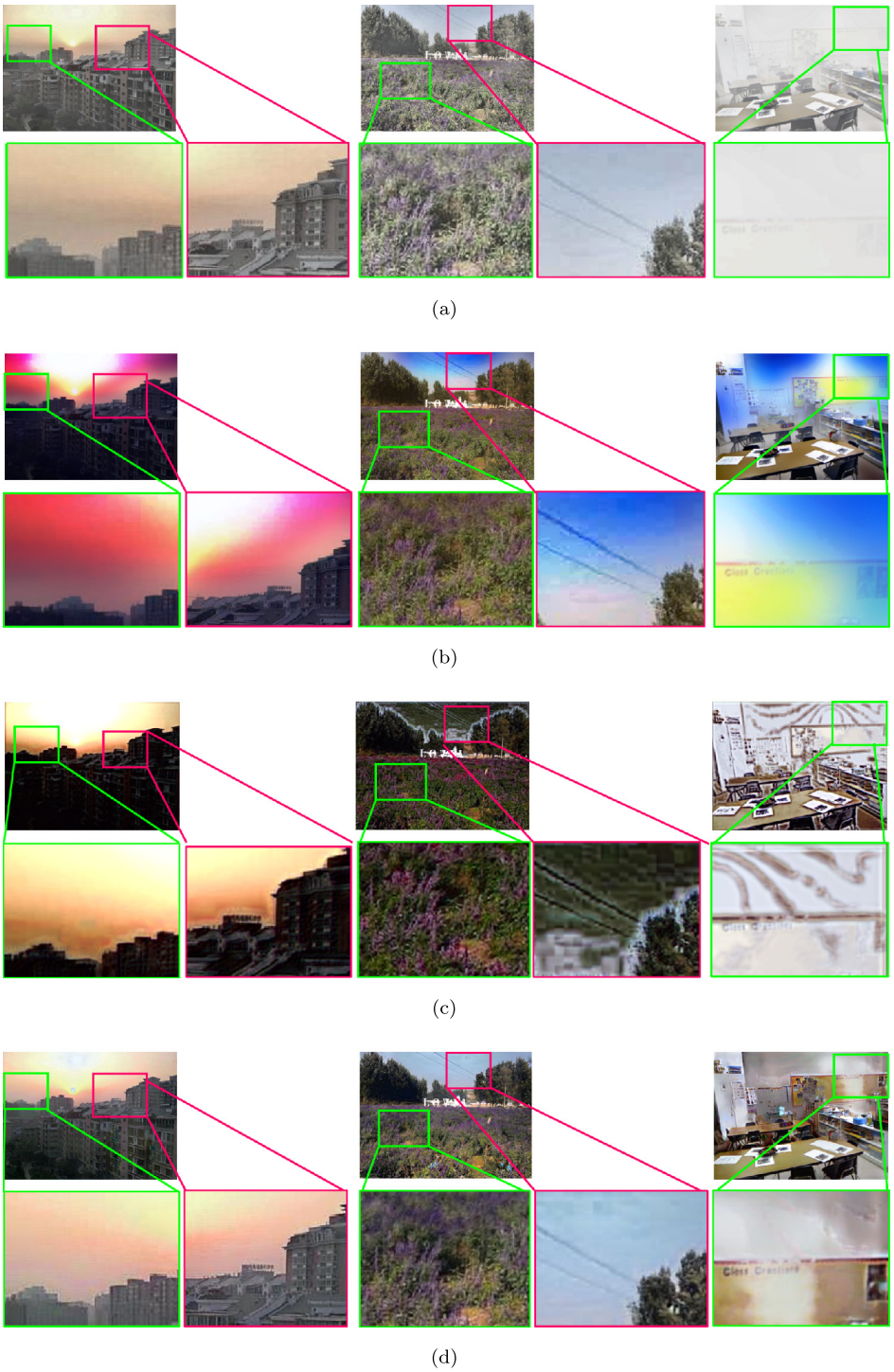


Fig. 10. Detailed areas of the dehazed images using **(a)** CLAHE, **(b)** DCP-R, **(c)** PDR-Net, and **(d)** Pix2pix, specifically the artifacts and color distortion.

rightmost image of Fig. 10b. While this color distortion does not occur in the result of PDR-Net, artifacts are introduced, such as shown in Fig. 10c.

Finally, the psychovisual experiment results also show the observers' general preference towards the dehazed images obtained by Pix2pix. Based on the MOS in Table 1, the order of preference after Pix2pix is CLAHE, then DCP-R, and lastly PDR-Net. Most observers take note of the color recovery as an major factor to determine visual pleasingness. DCP-R especially was singled out frequently about its colors, due to the color distortions that occur. Many observers commented particularly on the sky areas, that the colors were "uncomfortable", "unnerving", and "unnatural". Based on MOS, we can see that CLAHE is thus more preferred, although it clearly does not recover color correctly.

4.2. Contrast and clarity

The contrast and clarity of dehazed images can be inferred from RMSE or $SSIM^{-1}$ in Table 1. Based on those two metrics, once again Pix2pix is superior to the other metrics. In terms of image structure, CLAHE is able to keep it intact since it only modifies image intensities without changing the structure at all. This is shown by the CLAHE's

$SSIM^{-1}$ score that is second best to Pix2pix. Lastly, we have PDR-Net followed by DCP-R. The recovery of structure and detail in an image is difficult, particularly if pixel intensities are saturated with haze.

However, in dense haze images, such as in the first column of Fig. 9, the results of DCP-R show ability to recover contrast and detail, partially due to the help of the guided filter in the refinement step, which guarantees a smooth image. In these conditions, CLAHE is only able to enhance the edges that were already there, but is incapable of handling over-saturated haze areas. Meanwhile, dehazing by PDR-Net and Pix2pix is deep learning based, ensuring a highly accurate reconstruction as projected by a lower RMSE. Despite being the best-performing method in terms of RMSE or $SSIM^{-1}$, Pix2pix also struggles to recover images in dense haze. Additionally, since deep learning methods use a global accuracy over the entire image, local artifacts are still present, e.g., the first and third images of Fig. 10.

The results of the subjective evaluation show that observers also gravitate towards Pix2pix dehazing. Many observers take note of edges, visibility, and distinction of objects when evaluating the dehazed images. CLAHE is often noted for being unable to handle various hazy images, especially in dense haze. Observers have taken notice about the ability of DCP-R and Pix2pix to recover objects clearly with distinct edges. Next, images dehazed with CLAHE often still contain haze obscuring the objects and edges of the scene. Although edges and contrast is important in a scene, some observers mentioned that the contrast and edges in images dehazed by PDR-Net were extreme. PDR-Net created unnecessary edges and artifacts taht were pointed out explicitly as reasons to reject the dehazing results of PDR-Net.

4.3. Naturalness

Naturalness is a difficult concept to quantify because it is a largely subjective concept. This involves many immeasurable factors such as preference, perception, and familiarity. This paper attempts to measure naturalness quantitatively using NIQE. Based on the results in Table 1, DCP-R obtained the best NIQE of 2.8290, followed by Pix2pix, CLAHE, and lastly PDR-Net. Among the top two, DCP-R and Pix2pix, the subjective evaluation shows a preference towards Pix2pix dehazed images. CLAHE, on the other hand, may obtain natural results but is unable to complete the task of dehazing. Finally, PDR-Net performs the worst in reconstructing a natural dehazed image. A visual observation of some of the results can be done through Fig. 10.

PDR-Net dehazed images are generated by a deep-learning network through convolutions, allowing a high-level abstraction that does not necessarily match with how the human visual system or perception works. This process is computed based on accuracy and error, both of which are objective measures that do not reflect or account for cues that are important for human perception. PDR-Net was designed to use a modified error that considers both accuracy and visual perception, but the results here show that it still has this limitation. Meanwhile, while Pix2pix is also a deep generative network, it is trained using an adversarial discriminator, not just by a simple error or loss computation. The discriminator component of Pix2pix acts as if it is perceiving the image, trying to identify the generated image from the ground truth. Thus, the generator tries to generate an output that can *trick* the discriminator, resulting in a visually pleasing and perhaps more natural image.

Images dehazed with Pix2pix are most preferred by observers based on the MOS in Table 1. However, although the comments in Table 2 are generally positive towards the Pix2pix results, they still point out Pix2pix images to be unnatural. In terms of naturalness, we infer that the color distortions created by DCP-R skewed the subjective scores in favor of Pix2pix, even though DCP-R has sharper edges and smoother regions. Some comments also favored CLAHE in terms of naturalness, although observers acknowledge the haze was not successfully removed. In fact, many commented that the presence of haze actually adds to the naturalness of the image, as it is common to see such scenes in the real world. Many observers agree that the dehazed images by PDR-Net are the least visually pleasing, consistent with the NIQE scores. Many cite the rough edges, artifacts, noise, and loss of depth as the reason for this judgement.

5. Conclusion

In our experiments, we conducted image dehazing on a uniform set images using four methods, i.e., CLAHE, DCP-R, PDR-Net and Pix2pix. The two deep learning approaches, PDR-Net and Pix2pix were both trained using the same set of training images. In this work, we evaluate dehazing results not only by how similar they are to the ground truth, but also by image quality using objective and subjective IQA. The objective metrics were selected to measure color difference (ΔE^*_{ab}), contrast and image structure (RMSE, SSIM), and naturalness (NIQE). The subjective result is measured with the mean opinion scores (MOS). Among all four methods, Pix2pix is superior in terms of color difference, contrast and image structure. The MOS also shows that Pix2pix is the most preferred by human observers. Although Pix2pix is surpassed by DCP-R in terms of objective naturalness, DCP-R dehazing obtained the second to worst MOS due to the color distortion that sometimes occurs. CLAHE is unable to remove haze at all as reflected in its bad objective scores. Interestingly, CLAHE obtains the second best MOS after Pix2pix, indicating that the presence of haze is not a problem for human observers. Finally, PDR-Net dehazing does not excel at any of the objective criteria, although it is second place for color recovery. PDR-Net dehazing also falls in last place for naturalness and MOS.

The context of the dehazing becomes very relevant to the discussion. In cases where the images are intended to be viewed by humans, CLAHE is simple and easy to implement. However, although CLAHE is able to enhance the visibility of a hazy image, it would be misleading to claim CLAHE as a dehazing method since the results are often still hazy. For a fully automated machine based approach, PDR-Net is straightforward and succeeds to obtain good reconstruction with minimal human intervention. Alternatively, DCP-R is able to dehaze an image most naturally, but in order to avoid the color distortions, it is necessary to improve airlight estimation. It is necessary to look into the relevant quality cues of each dehazing method, and determine which ones are more important in the context of a certain task. This is particularly relevant for larger computer vision applications, in which higher-level image understanding is necessary. Often, these applications assume clear images as inputs, so any hazy inputs must be dehazed prior to processing. In this paper, we provided a general evaluation of image quality, focusing on comprehensive analysis and discussion on the color, image structure, and naturalness of the dehazed images.

In closing, the absence of a standardized hazy dataset is a common obstacle in hazy images. As such, our experiment was conducted on a fairly limited dataset of hazy images. Thus, there is much to be desired in regards to the generalization ability of dehazing methods. For future work, a general approach would be instrumental for dehazing and its subsequent applications. In line with this target, the establishment of a large standardized dataset with hazy, clear, and depth information will be very beneficial to scattering media image understanding in general, e.g., for underwater environments.

Declarations

Author contribution statement

L. Rahadianti, A. Y. Azizah: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper. H. Deborah: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This research was funded by Universitas Indonesia through Hibah Publikasi Terindeks Internasional (PUTI) Q3 grant number NKB-

1822/UN2.RST/HKP.05.00/2020. NTNU collaboration is supported by FRIPRO FRINATEK Metrological texture analysis for hyperspectral images (project nr. 274881) funded by the Research Council of Norway.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] R. Fattal, Single image dehazing, ACM Trans. Graph. 27 (3) (2008) 1–9.
[2] J.-H. Kim, J.-Y. Sim, C.-S. Kim, Single image dehazing based on contrast enhancement, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp. 1273–1276.
[3] K. Tan, J.P. Oakley, Enhancement of color images in poor visibility conditions, in: IEEE International Conference on Image Processing, vol. 2, 2000, pp. 788–791.
[4] C.O. Ancuti, C. Ancuti, C. De Vleeschouwer, Effective local airlight estimation for image dehazing, in: IEEE International Conference on Image Processing, 2018, pp. 2850–2854.
[5] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2341–2353.
[6] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, Single image dehazing via multi-scale convolutional neural networks, in: European Conference on Computer Vision, Springer, 2016, pp. 154–169.
[7] R. Li, J. Pan, Z. Li, J. Tang, Single image dehazing via conditional generative adversarial network, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 8202–8211.
[8] A.Y. Azizah, L. Rahadianti, H. Deborah, An introductory study on image quality of dehazed images, in: International Conference on Advanced Computer Science and Information Systems (ICACSIS), IEEE, 2020, pp. 301–308.
[9] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J.B. Zimmerman, K. Zuiderveld, Adaptive histogram equalization and its variations, Comput. Vis. Graph. Image Process. 39 (3) (1987) 355–368.
[10] C. Li, C. Guo, J. Guo, P. Han, H. Fu, R. Cong, Pdr-net: perception-inspired single image dehazing network with refinement, IEEE Trans. Multimed. 22 (3) (2019) 704–716.
[11] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1–17.
[12] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, Z. Wang, Benchmarking single-image dehazing and beyond, IEEE Trans. Image Process. 28 (1) (2018) 492–505.
[13] S.G. Narasimhan, S.K. Nayar, Interactive deweathering of an image using physical models, in: IEEE Workshop on Color and Photometric Methods in Computer Vision, 2003, pp. 1–8.
[14] S.G. Narasimhan, S.K. Nayar, Vision and the atmosphere, Int. J. Comput. Vis. 48 (3) (2002) 233–254.
[15] F. Cozman, E. Krotkov, Depth from scattering, in: IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 801–806.
[16] C. Tsiotsios, M.E. Angelopoulou, T.-K. Kim, A.J. Davison, Backscatter compensated photometric stereo with 3 sources, in: IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2259–2266.
[17] C.O. Ancuti, C. Ancuti, C. De Vleeschouwer, M. Sbert, Color channel compensation (3c): a fundamental pre-processing step for image enhancement, IEEE Trans. Image Process. 29 (2020) 2653–2665.
[18] C. Ancuti, C.O. Ancuti, Effective contrast-based dehazing for robust image matching, IEEE Geosci. Remote Sens. Lett. 11 (11) (2014) 1871–1875.
[19] C. Wang, B. Zhu, Image segmentation and adaptive contrast enhancement for haze removal, in: 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), 2020, pp. 1036–1039.
[20] V. Whannou de Dravo, J.E. Khoury, J.B. Thomas, A. Mansouri, J.Y. Hardeberg, An adaptive combination of dark and bright channel priors for single image dehazing, J. Imaging Sci. Technol. 61 (4) (2017) 40408-1.
[21] N. Silberman, D. Hoiem, P. Kohli, R. Fergus, Indoor segmentation and support inference from RGBD images, in: European Conference on Computer Vision, 2012, pp. 746–760.
[22] J. Zhang, D. Tao, Famed-net: a fast and accurate multi-scale end-to-end dehazing network, IEEE Trans. Image Process. 29 (2020) 72–84.
[23] X. Liu, Y. Ma, Z. Shi, J. Chen, GridDehazeNet: attention-based multi-scale network for image dehazing, in: IEEE/CVF International Conference on Computer Vision, 2019, pp. 7313–7322.
[24] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
[25] Z. Murez, S. Kolouri, D. Kriegman, R. Ramamoorthi, K. Kim, Image to image translation for domain adaptation, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4500–4509.
[26] L. Rahadianti, F. Sakaue, J. Sato, Depth estimation from single hazy images with 2-phase training, in: International Conference on Advanced Computer Science and Information Systems (ICACSIS), IEEE, 2020, pp. 309–316.
[27] M. Pedersen, J.Y. Hardeberg, Full-reference image quality metrics: classification and evaluation, Found. Trends Comput. Graph. Vis. 7 (1) (2012) 1–80.
[28] X. Li, Blind image quality assessment, in: International Conference on Image Processing, vol. 1, IEEE, 2002, pp. 449–452.
[29] D.M. Chandler, Seven challenges in image quality assessment: past, present, and future research, ISRN Signal Process. 2013 (1-53) (2013) 905685.
[30] C. Li, J. Guo, F. Porikli, H. Fu, Y. Pang, A cascaded convolutional neural network for single image dehazing, IEEE Access 6 (2018) 24877–24887.
[31] C.O. Ancuti, C. Ancuti, R. Timofte, L. Van Gool, L. Zhang, M.-H. Yang, Ntire 2019 image dehazing challenge report, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 2241–2253.
[32] C. Ancuti, C.O. Ancuti, R. Timofte, C. De Vleeschouwer, I-haze: a dehazing benchmark with real hazy and haze-free indoor images, in: Advanced Concepts for Intelligent Vision Systems, Springer International Publishing, 2018, pp. 620–631.
[33] C.O. Ancuti, C. Ancuti, R. Timofte, C. De Vleeschouwer, O-haze: a dehazing benchmark with real hazy and haze-free outdoor images, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 867–8678.
[34] C.O. Ancuti, C. Ancuti, M. Sbert, R. Timofte, Dense-haze: a benchmark for image dehazing with dense-haze and haze-free images, in: IEEE International Conference on Image Processing, 2019, pp. 1014–1018.
[35] C.O. Ancuti, C. Ancuti, R. Timofte, Nh-haze: an image dehazing benchmark with non-homogeneous hazy and haze-free images, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 1798–1805.
[36] E. Peli, Contrast in complex images, JOSA A 7 (10) (1990) 2032–2040.
[37] A.R. Robertson, The CIE 1976 color-difference formulae, Color Res. Appl. 2 (1) (1977) 7–11.
[38] P. Drews Jr, E. do Nascimento, F. Moraes, S. Botelho, M. Campos, Transmission estimation in underwater single images, in: IEEE International Conference on Computer Vision Workshops, 2013, pp. 825–830.
[39] K. He, J. Sun, X. Tang, Guided image filtering, in: European Conference on Computer Vision, Springer, 2010, pp. 1–14.
[40] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.
[41] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint, arXiv: 1411.1784, 2014.
[42] C. Li, M. Wand, Precomputed real-time texture synthesis with Markovian generative adversarial networks, in: European Conference on Computer Vision, Springer, 2016, pp. 702–716.
[43] H. Deborah, N. Richard, J.Y. Hardeberg, A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (6) (2015) 3224–3234.
[44] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[45] A. Mittal, R. Soundararajan, A.C. Bovik, Making a “completely blind” image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2013) 209–212.
[46] R.C. Streijl, S. Winkler, D.S. Hands, Mean opinion score (mos) revisited: methods and applications, limitations and alternatives, Multimed. Syst. 22 (2) (2016) 213–227.
[47] K. Pearson, Correlation coefficient, in: Royal Society Proceedings, vol. 58, 1895, p. 214.