Scene-Specific Dark Channel Prior for Single Image Fog Removal

Anparasy Sivaanpu[1], Kokul Thanikasalam[#2]

Abstract— **Extensive presence of fog in outdoor images severely alters the scene appearance and hence reduces the visibility. Image processing based defogging algorithms are used to restore the details and colour in a single foggy image. Performances of the previous defogging approaches are considerably low since they fail to consider the image-specific cues. In this paper, a novel and simple defogging approach is proposed based on the estimation of depth map by considering the density of fog in local image regions. The proposed approach uses the scene-specific depth map information to compute the dark channel and transmission. The quality of the recovered image is further improved by a post-processing technique. Experimental evaluation performed on FRIDA and FRIDA2 benchmark datasets demonstrates the proposed defogging framework outperforms state-of-the-art approaches. The results and code of this work are open-sourced for reproducibility (https://github.com/RPRO5/Defogging).**

Keywords—— **Defogging, Dark Channel Prior, Image Enhancement, Fog removal.**

I. INTRODUCTION

Poor visibility in outdoor images is one of the key challenge in many image understanding and computer vision-based applications such as traffic monitoring[1], automated driving[2], object detection[3], video surveillance[4], object tracking[5, 6] and aerial imagery[7]. Bad weather conditions, such as fog, haze, cloud and mist, can significantly reduce the visual quality of outdoor images. One of the major reasons for vehicle accidents is poor visibility due to the foggy scenes. The objective of the defogging frameworks is to increase the visual quality of foggy scenes by restoring the details and colour.

Fog is produced by the extensive presence of water droplets in the air due to the bad weather conditions. These water droplets scatter the sunlight and light reflected from other objects. Due to the scattering, the contrast of an outdoor scene will be fade, and a whiteness effect will be produced towards the observer or camera. These two effects jointly produce a poor-quality image. It is observed that[8, 9] the amount of fog in an image mainly depends on the distance between the scene and camera. Therefore, estimating the depth map of a foggy image is important for restoring the fog-free image.

Single image fog removal is a challenging task since fog is dependent on the unknown depth. Depth map prediction is a

challenging problem if the input is only a single image. Therefore, many defogging approaches have used multiple images and additional information[10, 11] to recover the details. However, without any prior information, single image fog removal in real-time is the demand of many real-world applications such as automated driving systems and vision-based surveillance systems.

Single image fog removal has attracted much research in the past[12-33] and following two major strategies[34]. First cluster of approaches[28-33] use a trainable machine learning technique to model the depth map of foggy scenes. These models recover the foggy image patches by learning the knowledge from similar foggy image samples. Although such methods showed significant recovering accuracy, their computational cost is much more expensive and therefore not suitable for many real-time applications.

The second group of single image defogging frameworks use the simple but efficient image enhancement methods, without any machine learning techniques. Most of the early works used several traditional image enhancement techniques such as median filtering[35], white balance correction[14], histogram equalization[36], edge smoothing[37], and contrast enhancement[12]. These methods are simple, fast, and can be used by most real-time applications. However, since they apply the enhancement techniques without considering the actual factors which causes poor quality in foggy images, their performances are limited and even may distort the actual colour information.

Recently, few image enhancement based approaches[15, 16] showed significant improvement in single image defogging by using stronger priors or assumptions. Based on the observations of a large number of foggy and fog-free images, the 'dark channel prior'[15] framework has been proposed and showed state-of-the-art results. This method lies in the observation of intensities of a few pixels are close to zero in at least one colour channel in a local patch of a fog-free image. Conversely, this observation does not exist in foggy images, and hence the scene can be recovered by estimating the thickness of the fog. Although the 'dark channel prior' approach is simple and effective, it applies the same defogging algorithm on all over the image without considering the depth map, which leads to distortion in recovered images, especially in dense foggy regions.

A robust single image fog removal framework should have state-of-the-art recovering performance and computationally efficient since these are the requirements of many real-world applications. To achieve this objective, we propose a novel defogging framework based on simple but efficient image enhancement techniques. The proposed approach relies on the assumption of a rough depth map of a foggy image can be observed by measuring the thickness of fog in local patches of an image. In the initial stage, the proposed approach measures the thickness of fog and then the rough depth map is obtained by using a superpixel-based segmentation technique. The captured depth map information is then used to feed the scene-

specific information in the dark channel prior algorithm. In addition, the proposed approach uses an efficient post-processing technique to enhance the quality of images. Extensive evaluation performed on two publicly available benchmark dataset demonstrates that the proposed framework outperforms state-of-the-art approaches.

The following contributions are made in the proposed single image fog removal approach:

1. In contrast to the existing image enhancement based approaches, we proposed a novel depth estimation technique for single outdoor foggy images. The proposed mechanism captures the rough depth map then feed that knowledge at different levels of the defogging process to increase the recovering performance.

2. We propose a scene-specific dark channel and transmission technique in fog removal. These techniques compute the dark channel and the amount of light that reaches the camera in the local image regions.

3. We have conducted extensive experiments on two well-known benchmark datasets to show the performances of the proposed approach. Then the recovering capability is compared with 18 state-of-the-art image enhancement based defogging approaches.

We have extended the preliminary work[38] by including a fusion based post-processing technique. In addition, an ablation study has been conducted to demonstrate the effectiveness of the proposed approach. The rest of this manuscript is organized as follows: Section II analyses the defogging and dehazing approaches. Section III introduces the background of this study. Section IV and V describes the proposed approach and experimental results, respectively. Finally, this manuscript is concluded in Section VI.

II. RELATED WORK

For several years, a large number of single image defogging and dehazing frameworks have been proposed and their performance have been evaluated on various benchmarks[19, 39, 40]. The most relevant single image defogging frameworks are stated in this section. Detail reviews on these frameworks are done in[9, 34, 41-43].

Few early frameworks are focused on image enhancement techniques without considering the fog formation model. Apurva et al.,[35] used gamma transmission and median filtering for defogging. Xu et al.,[36] used an adaptive histogram equalization technique to recover the colour foggy images. Ngo et al., [22] used a colour attenuation technique to restore the information. Also, several other enhancement methods used in the past such as wavelet transform, edge smoothing, boundary constraint[21], and Retinex theory. Few recent approaches use various filtering techniques to improve the visibility in foggy images. Tarel and Harutiere [20] used a novel filtering technique to improve the speed of the visibility restoration. Shin et al., [44] used a structured guided norm filter to enhance the image. Dhara et al., [28] used an Airlight refinement and nonliner balancing technique. These approaches tried to enhance the brightness and contrast in foggy images. However, their performances are limited since they fail to consider why the visual quality of foggy images is degraded.

Recent defogging and dehazing approaches rely on certain prior knowledge or observations. Kaiming He et al.,[15] proposed a novel single image dehazing approach, referred to as Dark Channel Prior (DCP). They found that, in most local patches of a fog-free image, intensities of a few pixels in at least one colour channel are very low. With this observation, they estimate the thickness of fog, and proposed a restoring algorithm. DCP is simple and efficient in most cases. Similar to the DCP approach, Ju et al., [26] proposed a dehazing technique, called Gamma Correction Prior (GCP). They have pre-processed the input image by using the Gamma correction technique and then used the atmospheric scattering theory to obtain the depth map of the image. Berman and Avidan proposed [23] an approach based on the observations of colours of a fog-free image are well approximated by a limited number of distinct colours.

DCP approach has drawn a grate interest in the image enhancement group and a large number of follow-up approaches have been proposed[8, 12-14, 16, 18, 27, 45-47]. A review on DCP based approach can be found in [48]. Few approaches [13, 46] improve the recovering performance and speed of DCP by introducing a new smoothing filter, called as guided filter. Renjie He et al.,[14] proposed a white balance correction technique to refine the DCP algorithm. Shunyuan Yu et al.,[45] included a multiple transmission layer fusion technique to enrich the performance of DCP. Chunlin Chen et al.,[47] replaced the global parameters of DCP by a location based local parameter setting. Qingsong et al.,[16] proposed a colour attenuation prior model based on the inspirations of DCP. Jin-Hwan et al.,[12] proposed a contrast enhancement technique for defogging in images and video. Anwar and Arun[8] proposed a novel post-processing technique in DCP. Salazar et al., [27] used a multi-layer perception to reduce the computation time of DCP. Wang et al., [49] used Markov random field (MRF) to measure the depth map of a haze image and then revise the DCP accordingly.

Few recent frameworks are proposed based on certain fusion techniques. These approaches rely on two or more separate image enhancement techniques and then combine the results together at the end. Zhu et al., [24] proposed a haze removal technique by fusing the DCP and Luminance. Thulasika and Ramanan [18] fused the contrast based method and statistical based approaches in fog removal. Galdran [25] artificially generated multiple images by using a sequence of Gamma correction operations and then merged them together for image dehazing. Although fusion based techniques showed significant recovering performance, they are relying on the fusion parameter. Therefore, their performances are different from one environment to another.

Although DCP and fusion based single image defogging approaches are simple and computationally efficient, there is still a considerable recovering performance difference observed when these approaches are evaluated on benchmark datasets. Our intention is to reduce this performance difference by feeding the depth map information in DCP algorithm without using any machine learning model as they are computationally expensive.

III. BACKGROUND

In this section, formation of a foggy image is described in detail. In addition, since we use the Dark Channel Prior (DCP)[15] algorithm as the baseline defogging framework in this work, it is reviewed in detail.

A. Atmospheric Scattering Model for Fog formation

In computer graphics, the formation of fog is described by an atmospheric scattering model[50]. In this model, fog is

Fig. 1 An illustration of dark channel in a 3 ×3 image patch. (a). Red Channel values (b) Green Channel values (c) Blue Channel values (d) Obtained dark channel values.

treated as a combination of two components: Direct attenuation and Airlight. Attenuation diminishes the contrast and Airlight adds a whiteness effect in a foggy image. Based on that, the formation of fog is expressed as:

$$\mathbf{I}(x) = I_{Attenuation}(x) + I_{Airlight}(x) \qquad (1)$$

Where x is the location of a pixel within the image, $\mathbf{I}$ is the observed foggy image, and $I_{Attenuation}$ and $I_{Airlight}$ are the attenuation and airlight on that location, respectively. The attenuation describes the scene radiance (i.e. the fog-free image that we want to recover) and medium transmission. Airlight is the scattered light that leads to the shift of the scene colours. The both components can be expressed by the following equations:

$$I_{Attenuation}(x) = \mathbf{J}(x)\, t(x), \qquad (2)$$

$$I_{Airlight}(x) = \mathbf{A}\,(1 - t(x)), \qquad (3)$$

Where $\mathbf{J}$ is the scene radiance, $\mathbf{A}$ is the atmospheric light, and t is the transmission parameter, which indicates the part of the light that penetrates through the fog. The transmission parameter t depends on the distance between the camera and scene, and it can be described as:

$$t(x) = e^{-\beta d(x)}, \qquad (4)$$

Where β is the scattering coefficient and d is the depth of the scene. The objective of any defogging framework is to obtain $\mathbf{J}, \mathbf{A}$ and t from $\mathbf{I}$. To obtain these values, several assumptions and prior knowledge are utilised in many single image defogging frameworks.

B. Dark Channel Prior

Dark Channel Prior (DCP) is a computationally efficient technique, and hence many of the recent defogging and dehazing frameworks make use of it as their baseline.

DCP works only for colour images and is built based on the concept of a dark channel. In a colour image, its dark channel $\mathbf{J^{dark}}$ can be defined as:

$$J^{dark}(x) = \underset{y \in \Omega(x)}{Min}\left(\underset{c \in \{r, g, b\}}{Min} J^c(y)\right), \qquad (5)$$

Where $\Omega(x)$ is a local image patch centred at x and $\mathbf{J^c}$ is the colour channel of $\mathbf{J}$. The authors of DCP assume that the transmission is constant in the local patch $\Omega(x)$. Fig.1 demonstrates the concept of dark channel in a 3× 3 image patch. Based on their observation on a large number of fog-free images, the intensity of dark channel is tends to be zero, except the sky regions as:
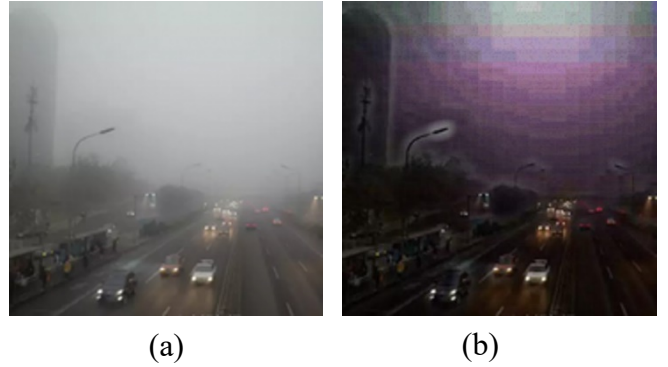

Fig. 2 DCP approach [15] shows poor performance (unnatural effect) in brighter regions. (a) Foggy image. (b) Recovered image.

$$J^{dark} \rightarrow 0, \qquad (6)$$

This observation is called as *dark channel prior*. Based on that, the transmission can be derived as:

$$t(x) = 1 - \omega \underset{y \in \Omega(x)}{Min}\left(\underset{c}{Min}\frac{I^c(y)}{A^c}\right), \qquad (7)$$

Where $\omega\ (0 \leq \omega \leq 1)$ is the defogging parameter that will control the degree of fog removal. The DCP algorithm assumes that the brightest pixel in $\mathbf{I}$ is approximately equals to $\mathbf{A}$. Based on these assumptions and equations, the scene radiance $\mathbf{J}$ is described as:

$$\mathbf{J}(x) = \frac{\mathbf{I}(x) - A}{\max(t(x),\ 0.1)} + A, \qquad (8)$$

The performance of DCP defogging framework is mainly depends on its key parameters: Patch size (Ω) and defogging parameter ω. The dark channel concept becomes stronger for larger values of Ω because possibility of a large patch contains a dark pixel is increased. On the other side, DCP built on the assumption of transmission is consent in a patch, and hence large Ω produces halos in near depth scenes. Therefore, based on the experiments, patch size Ω is fixed to 15 ×15 in DCP and other follow-up approaches. We have observed that Ω should be as much as large in constant transmission regions to get the better dark channel. Based on equation (4), amount of transmission is same in equal depth regions. Therefore, patch size Ω should be fixed locally, based on the depth map of a foggy image.

The other key parameter ω is used to control the degree of fog removal in the DCP approach. The value of ω is fixed to 0.95 in DCP and other approaches, since if the entire fog is removed from a foggy image, the recovered image may seem unnatural. However, based on our observations, ω should be
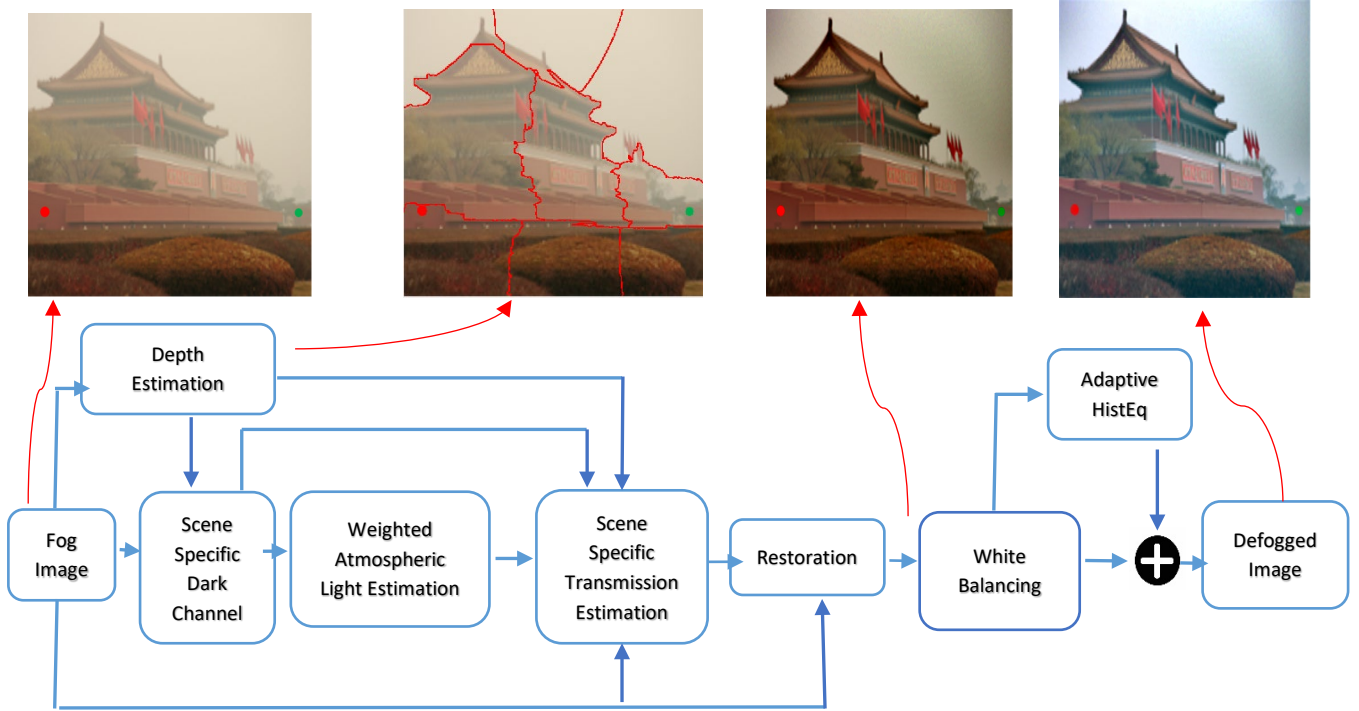
Fig. 3 Flow diagram of the proposed framework. It feeds the scene-specific information into the dark channel prior based algorithm by estimating the rough depth map at the initial stage, and then used it to obtain the scene-specific dark channel and the adaptive transmission. In post-processing, white balancing techniques are used with adaptive histogram equalization.

large in dense foggy regions and be small thin foggy regions. Otherwise, as shown in Fig.2, it causes an unnatural effect in the recovered image. Since the density of the fog depends on the depth map, values of ω should be fixed locally, based on the depth map of a foggy image.

We have also noticed that DCP and follow-up approaches show poor performance in brighter regions because they apply fixed values for the both key parameters. To overcome the limitations of DCP and other dark channel prior based approaches, we propose a novel single image defogging framework by capturing and including the scene-specific information.

IV. METHODOLOGY

We proposed a robust framework for single image fog removal. The key objective of this framework is to include the scene-specific knowledge to the dark channel prior based defogging to increase the restoring performance. In the initial stage, the proposed approach estimates the rough depth map of the foggy image based on the density of the fog in local regions.

Then the obtained rough depth map is used to calculate the scene-specific patch sizes (Ω) and defogging parameters (ω) for corresponding individual local regions. Based on the local values of these parameters, scene-specific dark channel is obtained and then the scene-specific transmission is estimated. In the post-processing, we use a colour balancing and white balancing mechanism with an adaptive histogram equalization technique to enhance the quality of recovered image. The proposed approach shows state-of-the-art recovering performance in homogeneous and heterogeneous fogs. This framework is evaluated on two benchmark datasets and the results are compared with 18 state-of-the-art approaches. The overview of the proposed framework is shown in Fig.3. The details of each steps are explained in the following subsections.

A. Pre-processing

In the initial stage of the proposed approach, all the foggy images are resized to a static size. This resizing process is used to conduct the methodology and evaluation without any bias.
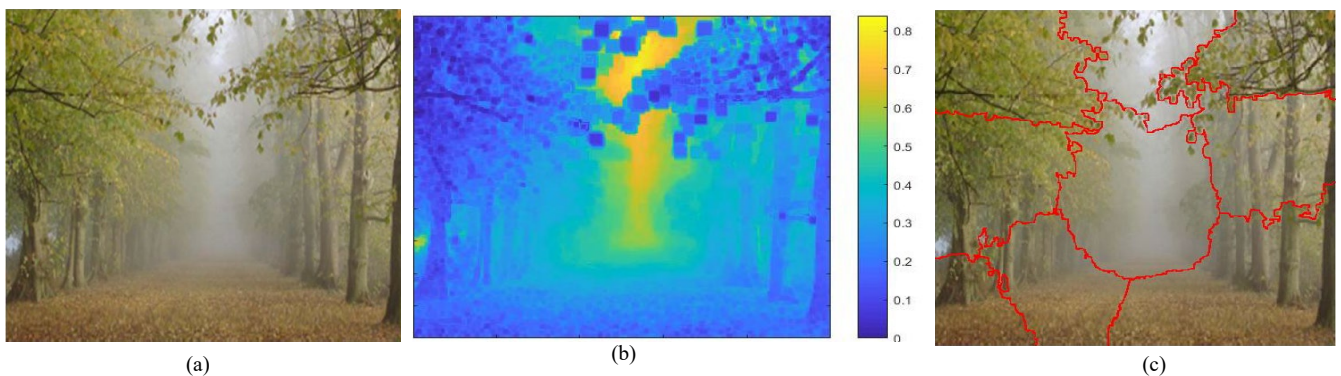

Fig.4. Demonstration of proposed depth estimation technique in a foggy image. (a) Foggy image. (b) Intensity values of dark channel patches. (c) Identified equal depth regions by using the superpixel based segmentation technique. Boundaries of the regions are shown in red colour.
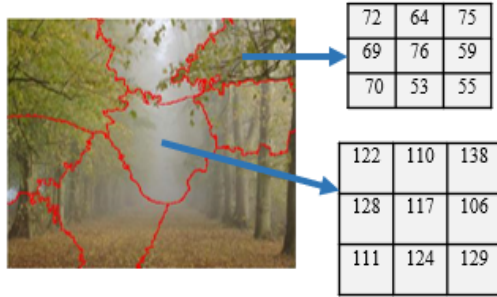
Fig.5: Relationship between dark channel and density of fog in local regions. Dark channel is high in high fog regions and low in low fog regions.

B. Depth Map Estimation

Estimating the depth or geometry of a scene from a single image is one of the major problems in several computer vision based applications. Without a machine learning technique, proposing a depth estimation model is a challenging task, since the appearance information is insufficient to resolve depth ambiguities. However, machine learning based depth estimation techniques are computationally expensive and hence not suitable for real-time single image fog removal. In this background, we focused to develop a simple but efficient depth map estimation technique based on some prior assumptions and observations in foggy images. The proposed technique identifies the distinct depth regions in a foggy image and then that knowledge is used to improve recovering performance of the proposed defogging framework.

We have observed that density of fog is peak in depth scenes and low in nearby scenes in outdoor foggy images. Also, it is observed[15] that the intensity values of dark channel patches are proportional to the thickness of the fog. As illustrated in Fig.5, we have observed that intensity values of dark channel patches are proportional to the density of fogs in that patches. Based on these observations, we assume that rough depth map of a foggy image can be estimated by the intensity of the dark channel.

In the initial stage of proposed depth map estimation technique, dark channel of a foggy image is obtained by using a small image patch. Then the intensity values (**i**) of these patches are obtained as,

$$i = [i_1, i_2, ..., i_m] , \qquad (9)$$

Where m is the number of dark channel patches in an image. In the next step of proposed depth map estimation technique, adjacent patches with equal or similar intensity values are identified and then grouped since they are from equal depth regions. To do this task, we have used an image segmentation technique.

Several image segmentation techniques are proposed in the past[51] and they are different one to another based on their techniques and performances. In general, superpixel based segmentation techniques are more suitable to segment the irregular shapes. Since the depth map regions are in irregular shapes, we have used a superpixel based image segmentation algorithm. We have followed the 2D superpixel based segmentation algorithm of Achanta et al.,[52] to obtain the distinct depth map regions. This algorithm identifies the similar depth regions (**R**) as,

$$R = [r_1, r_2, ..., r_n] , \qquad (10)$$

Where n is the number of equal or similar depth regions and $n \le m$. Fig.4. visualizes the proposed depth map estimation technique in a foggy image. In the next step of proposed approach, the obtained depth map information is used to feed the scene-specific knowledge in defogging.

C. Scene-Specific Dark Channel

Defogging performance of DCP and all follow-up approaches are mainly depending on the dark channel estimation. In all these approaches, dark channel of a foggy image is computed (based on equation 5) by assuming the transmission is constant within a local patch. All approaches are used a 15×15 patch (Ω) without considering any depth information of the image.
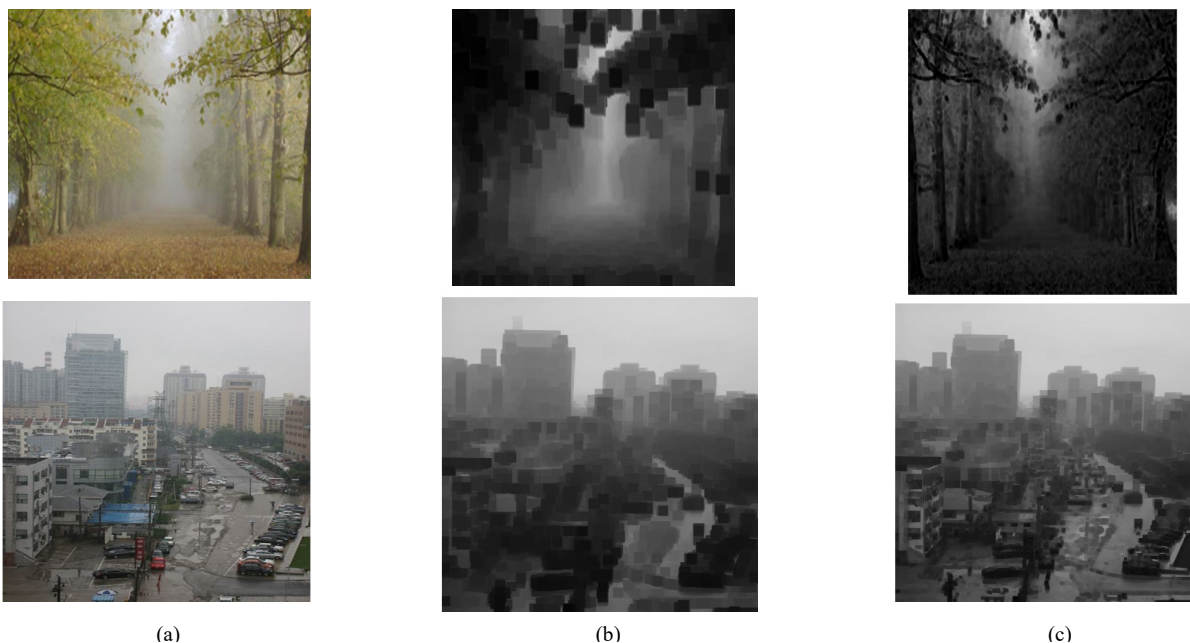

| (a) | (b) | (c) |

Fig.6. Effectiveness of the proposed scene-specific dark channel. (a) Foggy image. (b) Dark channel of DCP [15] approach. (c) Proposed scene-specific dark channel.
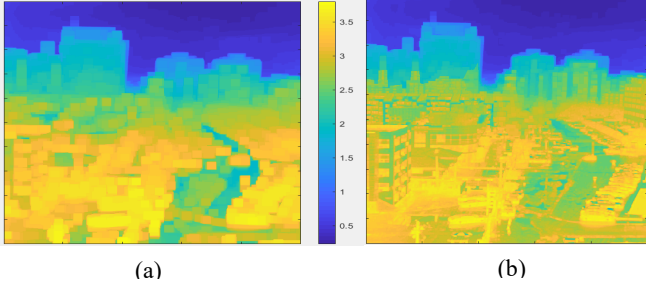
Fig.7. Effectiveness of the proposed scene-specific transmission estimate. (a) Transmission of the DCP [15] approach. (b) Proposed scene-specific transmission.

It is observed that dense fog regions are brighter than thin fog regions due to the additive airlight. As shown in Fig. 5, We have also notice that dark channel pixel values are lower in brighter regions. Based on these observation and fact, we have proposed a novel technique to get the robust dark channel, which is called as scene-specific dark channel. This proposed technique lies on the assumption of large patch sizes are desirable in dense fog regions since these regions are much brighter and hence the probability of a dark pixel is much lower.

From the initial step of the proposed framework, depth regions were identified. We have used that regions to feed scene-specific knowledge in dark channel computation. Firstly, average intensity (r_i) values of each regions are calculated, and then these values are considered as the density of fog in that regions. Based on the average intensity values, we have fixed the patch size for each region. Large patch size is fixed for high density regions while small size is fixed for low density regions. The details of this assignment are described in Table I.

TABLE I: Average intensities of regions and corresponding patch size Assignment.

Average Intensity of Regions (r_i)	Patch Size (Ω)
$r_i \geq 240$	15×15
$240 > r_i \geq 140$	11×11
$140 > r_i \geq 90$	7×7
$r_i < 90$	3×3

Based on these patch size assignments, dark channel of a local region J_r^{dark} is calculated as

$$J_r^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} J^c(y) \right), \quad (11)$$

Where $\Omega(x)$ is a local image patch centred at x in region r, and $\mathbf{J^c}$ is the colour channel of $\mathbf{J}$. Then, the overall dark channel of an image is calculated as:

$$J^{dark} = J_1^{dark} \cup J_2^{dark} \dots \cup J_n^{dark}, \quad (12)$$

Where n is the number of depth regions in that image. The proposed technique computes more robust scene-specific dark channel based on the depth map of corresponding local regions. This proposed technique is more robust than the dark channel computation of previous approaches. Fig. 6 visualizes the effectiveness of our technique by comparing the scene-specific dark channel with dark channels of DCP and follow-up approaches. It is clearly seen that the proposed scene-specific dark channel produces more robust results based on the depth map information.

D. Weighted Atmospheric Light Estimation

Fig.8. Effectiveness of the white balancing technique. (a) Recovered image (b) post-processed image.

Atmospheric light (denoted as $\mathbf{A}$ in equations) estimation is an important step in all defogging frameworks since it is used to recover scenes. DCP and follow-up approaches estimate the $\mathbf{A}$ by assuming the intensity of the brightest pixel in the dark channel as the $\mathbf{A}$ in a foggy image. We have used the scene-specific dark channel to estimate the atmospheric light estimation.

In the proposed approach, a set of brightest pixels of the scene-specific dark channel are identified. Among them top 0.1% brightest pixels are selected, and their average intensity value is considered as the $\mathbf{A}$. Since our scene-specific dark channel is better than the previous dark channel computations, more accurate $\mathbf{A}$ is obtained.

E. Scene-Specific Transmission Estimation

Transmission (denoted as t in equations) indicates the amount of light that penetrates through the fog. Robust transmission estimation is important in all defogging frameworks since it is used to recover the fog-free image through equation 8. Based on equation 7, the transmission estimation process is mainly dependent on the defogging parameter ω since it controls the degree of fog removal in an image.

In DCP and follow-up approaches, ω is kept to a fixed value (0.95), and hence the fog removal process is conducted with equal probability in all the pixels of an image. However, we have notice that conducting fog removal with equal probability in all pixels will create an unnatural effect in the recovered image. The degree of fog removal should be large in dense fog regions and be small in thin fog regions. Based on that concept, we have used a novel technique, which is called scene-specific transmission estimate.

TABLE II: Average intensities of regions and corresponding ω values Assignment

Average Intensity of Regions (r_i)	defogging parameter (ω_r)
$r_i \geq 230$	0.95
$230 > r_i \geq 140$	0.9
$140 > r_i \geq 100$	0.8
$100 > r_i \geq 80$	0.7
$r_i < 80$	0.6

We have utilised the obtained depth map regions, which were identified at the initial stage, to fix the density based local ω values. The average intensity of each depth regions is computed and then corresponding local ω values are assigned based on the experiment. High values are assigned for high

density fog regions and low values are assigned for low density regions. The details of this process are described in Table II.

Based on the corresponding local defogging parameter, the scene-specific transmission is calculated as:

$$t(x) = 1 - \omega_r \min_{y \in \Omega(x)} \left(\min_c \frac{I^c(y)}{A^c} \right), \qquad (13)$$

Where ω_r is the region specific defogging parameter. The proposed technique removes more fog in high density regions while keeps few amounts of fog in thin density regions. Therefore, the recovered fog-free image becomes more natural than the previous approaches. Fig.7 compares the heatmap of DCP[15] and proposed scene-specific transmission estimate. As illustrate in this picture, proposed transmission technique is more accurate than the previous approaches.

F. Recovering Scene Radiance

Similar to other approaches, scene radiance **J** is recovered by using the equation 8 in the proposed framework. The estimated scene-specific transmission and inputted foggy images are used to recover the fog-free image.

G. Post processing: White Balancing

In the proposed approach, quality of recovered image is enhanced further by using few post processing techniques. In the initial state of post processing, a white balancing algorithm is used, which is similar to [14]. Let **I** be the recovered image and it is normalized between 0 and 1. Then the white balancing technique is applied as:

$$I'_c = W . I_c, \qquad (14)$$

Where I_c and I'_c are the colour channel of recovered image and post-processed image, respectively. $c \in \{R, G, B\}$ is a colour channel. W is a three channel matrix and it is obtained as:

$$W_c = \frac{A^c}{max^c}, \qquad (15)$$

Where A^c is the Atmospheric light. This white balancing technique provides more natural effect in the recovered image. Fig.8 shows the effectiveness of the white balancing technique.

H. Post processing: Fusion based contrast enhancement

In the next step of post processing, a fusion based contrast enhancement technique is used to improve the quality of recovered image. This proposed technique computes two inputs by using two different mechanisms and then combine the results by using an efficient fusion algorithm.

(a)   (b)

Fig.9. Fusion based contrast enhancement (a) recovered fog-free image. (b) Enhanced post-processed image.

The first input (I_1) of the fusion mechanism is the white balanced recovered image from previous step. The second input (I_2) of the fusion process is the colour corrected version of the recovered image. We have used the local adaptive histogram equalization technique[53] to obtain the colour corrected image. Similar to [54], we have used a multiscale fusion technique to obtain the restored image. In this process, the enhanced recovered image **F** is obtained at every pixel location (x,y) as:

$$F(x, y) = W_1(x, y) . I_1(x, y) + W_2(x, y) . I_2(x, y), \quad (16)$$

Where W_1 and W_2 are the gaussian normalized weights. This post processing technique is used to produce more natural effects in the recovered image. Fig.9 compares the recovered image and fused image.

V. EXPERIMENTAL RESULTS

In this section, we present the experimental evaluation of the proposed defogging framework. Initially, implementation details and evaluation criteria are described. Then the details of the dataset are presented. Then the effectiveness of the proposed scene-specific defogging approach is illustrated through an ablation study. Finally, the evaluation results are matched with the state-of-the-art defogging frameworks.

A. Implementation Details

The architecture of the baseline defogging framework is built from the DCP[15] approach. In the pre-processing of the proposed approach, all foggy images are resized to 500×500. In the proposed depth estimation technique, a 3×3 patch is used to obtain the density of the fog. In the next step, each image is segmented into five to seven regions by using the superpixel based technique. The proposed approach is implemented in MATLAB. All experiments are conducted on an Intel core i7-8550U CPU.

B. Evaluation Protocols

Mean-Squared Error (MSE), Peak Signal-To-Noise Ratio (PSNR) and Structural SImilarity Index (SSIM) are the well-known metrices to evaluate the performance of defogging and dehazing frameworks. MSE calculates the average of the squares of the errors between ground truth (fog-free) image (**G**) and recovered image (**F**) by the following equation:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [F(i, j) - G(i, j)]^2, \qquad (17)$$

Where M and N are the width and height of the image, respectively. PSNR measures the peak error as follow:

$$PSNR = 10 \log_{10} \left(\frac{MAX_G^2}{MSE} \right), \qquad (18)$$

Where MAX_G is the maximum pixel value in the ground truth image. Since PSNR is depending on MSE, both metrices are producing similar results. The SSIM[55] index is used to measure the quality of an image based on a reference image. SSIM compares the luminance, contrast, and structure between the ground truth image and recovered fog-free image. The SSIM index is measured as:

$$SSIM = \frac{(2\mu_F \mu_G + C_1)(2\sigma_{FG} + C_2)}{(\mu_F^2 + \mu_G^2 + C_1)(\sigma_F^2 + \sigma_G^2 + C_2)} \qquad (19)$$

Where μ_F and μ_G are the average of fog-free image and recovered image, respectively. σ_F^2 and σ_G^2 are the variance of

Fig.10 Ground truth – foggy image pairs of benchmarks. (first column) FRIDA benchmark (second column) FRIDA2 benchmark

fog-free image and recovered image, respectively. σ_{FG} is the covariance of $\mathbf{F}$ and $\mathbf{G}$. C_1 and C_2 are two variables and used to stabilize the division. These variables are computed as:

$$C_1 = (K_1 L)^2, \qquad (20)$$

$$C_2 = (K_2 L)^2. \qquad (21)$$

Where K_1 and K_2 are small constants. L is the dynamic range of pixels values. A high SSIM index represents high similarity between the ground truth image and recovered fog-free image.

In this evaluation, we have used the PSNR and SSIM to measure the distance error and the similarity between the ground truth image and the recovered image, respectively.

C. Benchmark Datasets

In general, foggy images and corresponding ground truth (fog-free) images are important to evaluate the performance of the defogging frameworks. However, it is difficult to obtain such kind of pairs of images and hence synthetic images are used to evaluate the recovering performance in many approaches. In this background, the proposed approach is evaluated on two benchmark datasets, called Foggy Road Image DAtaset (FRIDA)[19] and Foggy Road Image Dataset version 2 (FRIDA2)[40].

FRIDA benchmark has 90 synthetic foggy images of the outdoor road scenes and the same number of fog-free images. All the images of FRIDA are in the size of 640×480. It has four types of heterogeneous foggy images, which were generated by using different parameter settings.

FRIDA2 benchmark has 330 synthetic outdoor foggy images of the road scenes including the fog-free (ground truth) images. It has homogeneous foggy images and three types of heterogeneous foggy images. All the images are the size of 640×480. A set of ground truth image and foggy image pairs of both datasets are shown in Fig.10.

D. Ablation Studies

The proposed scene-specific defogging framework has several step-by-step process. In addition, the recovered image is further corrected by using a set of post processing techniques. In the proposed approach, several parameters have been used in each step. Therefore, an ablation study is performed to validate the methodology and to verify the parameters. The FRIDA benchmark is used in this study.

1) *Verification of the parameters:* In the proposed depth estimation technique, a small patch is used to obtain the amount of fog in an image. Table III compares the defogging performance for different patch sizes.

TABLE III: COMPARISON OF PATCH SIZE FOR DEPTH ESTIMATION

Patch Size	Average SSIM
3×3	**0.8290**
5×5	0.8288
7×7	0.8283
9×9	0.8278

Based on these results, patch size of 3×3 is used in this study.

In the scene-specific dark channel step, patch sizes are assigned based on the average intensity of local regions (Table II). We have tested various patch sizes (in descending order) and picked the best performing one. Table IV compares this assignment for different patch sizes.

TABLE IV: COMPARISON OF PATCH SIZES FOR SCENE-SPECIFIC DARK CHANNEL

Patch Sizes	Average SSIM
29×29,21×21,15×15,9×9	0.8271
25×25,17×17,13×13,7×7	0.8285
21×21,15×15,11×11,5×5	0.8290
15×15,11×11,7×7,3×3	**0.8295**

Based on these results, 15×15,11×11,7×7 and 3×3 patch sizes are used to obtain the scene-specific dark channel.

2) *Effectiveness of Steps:* Performance of the proposed approach is depending on the depth estimation technique and post processing steps. Table V compares the performance of these steps based on their individual and joint contribution in defogging.

TABLE V: COMPARISON BETWEEN INDIVIDUAL AND JOINT STEPS

Baseline (DCP)	Depth Estimation	White Balancing	Contrast Enhancement	Average SSIM
√				0.6549
√	√			0.7921
√	√	√		0.8251
√	√	√	√	0.8296

Based on the results in Table V, it is clearly seen that feeding depth information in scene-specific dark channel and scene-specific transmission estimation brings more gain than the post processing steps. Approximately, the depth estimation technique increases the defogging performance by 14%.

TABLE VI: PERFORMANCE COMPARISON ON FRIDA AND FRIDA2 BENCHMARK DATASETS. THE BEST AND SECOND SCORES ARE SHOWN IN RED AND BLUE COLOURS, RESPECTIVELY.

Defogging Approach	FRIDA		FRIDA2	
	PSNR	SSIM	PSNR	SSIM
Proposed Approach	15.3365	0.8296	12.7029	0.7530
Our Preliminary Work[38] (2020)	14.5531	0.8225	12.2428	0.7427
Galdran [25] (2018)	14.8589	0.8207	12.3126	0.7462
Thulasika and Ramanan [18] (2015)	14.7063	0.8126	12.3097	0.7404
Qingsong et al.,[16] (2015)	13.9140	0.8051	12.3959	0.7463
Jin-Hwan et al.,[12] (2013)	14.3187	0.8043	12.3786	0.7404
Meng et al.,[21] (2013)	14.1193	0.7940	12.4388	0.7318
Tarel et al.,[19] (2010)	12.1701	0.7921	10.4687	0.7354
Ngo et al.,[22] (2019)	13.6022	0.7790	12.3109	0.7312
Berman and Avidan[23] (2016)	13.6960	0.7778	11.9464	0.7200
Tarel and Hautiere[20] (2019)	13.8208	0.7747	12.4109	0.7262
Shin et al.,[44] (2020)	14.2896	0.7394	12.4476	0.7231
Zhu et al.,[24] (2017)	13.3543	0.7323	11.7765	0.6646
Dhara et al.,[28] (2010)	13.1461	0.7249	11.3341	0.6107
Salazar et al.,[27] (2018)	13.1038	0.7246	11.9212	0.6718
He et al., [13] (2010)	12.7739	0.7182	11.1081	0.6096
DCP[15] (2009)	12.9382	0.6192	11.6093	0.5947
He et al.,[14] (2012)	12.4027	0.4273	10.7532	0.3935
Gundawar et al.,[17] (2014)	11.0105	0.3758	10.8823	0.4154
Ju et al.,[26] (2020)	9.7032	0.3626	8.7933	0.2983

E. Comparison with state-of-the-art defogging approaches

The performance of the proposed defogging framework is compared with state-of-the-art single image defogging approaches on FRIDA and FRIDA2 benchmarks. We compare the proposed approach with the works of Kokul and Anparasy (our preliminary work)[38], Galdran [25], Thulasika and Ramanan [18], Qingsong et al.,[16], Jin-Hwan et al.,[12], Meng et al.,[21], Tarel et al.,[19], Ngo et al.,[22], Berman and Avidan[23], Tarel and Hautiere[20], Shin et al.,[44], Zhu et al.,[24], Dhara et al.,[28], Salazar et al.,[27], He et al., [13], DCP[15], He et al.,[14], Gundawar et al.,[17], and Ju et al.,[26]. In this comparison study, source codes of all approaches are obtained and then evaluated on FRIDA and FRIDA2 benchmarks with the same parameter settings.

Table VI compares the performance of proposed single image defogging framework with other state-of-the-art approaches. Fig.11 compares the MSE and SSIM in FRIDA benchmark. Similarly, Fig.12 compares these values in FRIDA2 benchmark. Based on these tables and graphs, it is clearly seen that the proposed framework produces excellent defogging performance in both benchmarks for both evaluation metrices.
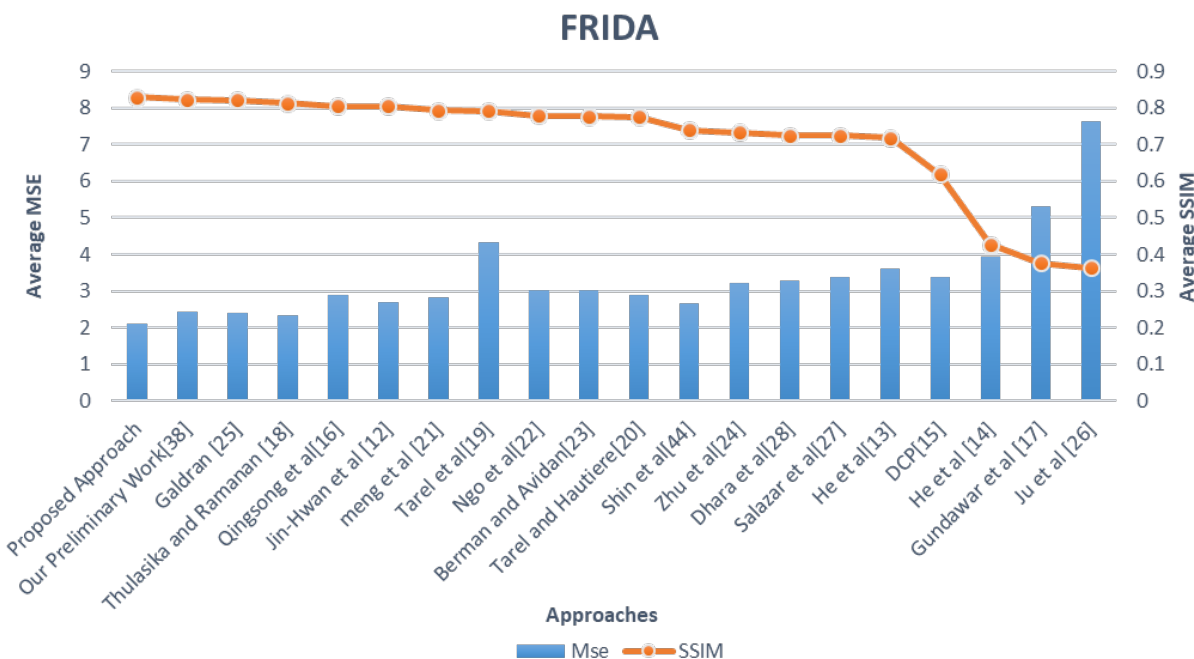

Fig. 11. Comparison of MSE (lowest is best) and SSIM(highest is best) in FRIDA benchmark.
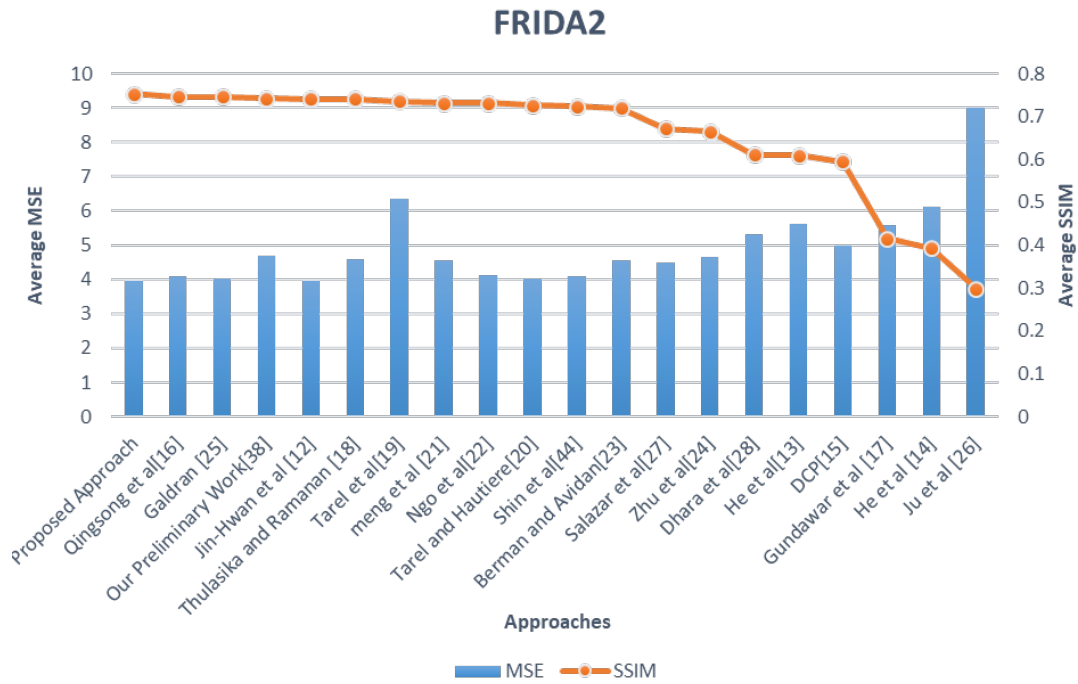
Fig.12 Comparison of MSE and SSIM in the FRIDA2 benchmark

F. Qualitative Comparison on Real-World Scenes

Although, proposed approach showed excellent performance on FRIAD and FRIDA2 benchmarks, images of these datasets are artificially created (synthetic) images. Therefore, we have conducted an additional comparison on some challenging real-world foggy images. These images are collected from multiple internet sources. Since these images have not any ground truths (fog-free images), it is difficult to rank them based on their recovering performance. However, qualitative comparisons of these images are widely accepted by many researchers.

Fig.13. Qualitative comparison of five best performing methods on real-world images. (a) Foggy images (b) Galdran [25] 's result (c) Thulasika & Ramanan[18]'s results (d) Qingsong et al. [16]'s results (e) Jin-Hwan et al. [12]'s results (f) Our results.

Fig.13 shows the qualitative comparison of the proposed approach with five top performing defogging frameworks. As shown in this figure, the proposed approach successfully resorted the details of scenes and objects. In addition, it is observed that while outputs of other's approaches are oversaturated and fail to balance the colour and contrast, our approach produces more natural outputs. Both comparison studies proved that the proposed approach can restore the fog-free images from homogeneous and heterogeneous fogs.

VI. Conclusion

In this study, an image processing based simple but efficient framework is proposed for single image defogging. Based on the density of fog in local regions, we have obtained the rough depth map of a foggy image, and then used that depth map information to feed the scene-specific knowledge to defogging. In addition, we have proposed a scene-specific dark channel estimation technique and a transmission estimation technique, which are used to improve the recovering performance. Also, the quality of recovered images is further improved by using a set of post processing techniques. The proposed defogging method is evaluated on FRIDA and FRIDA2 benchmark datasets and obtained 15.3365 and 12.7029 of PSNR values and 0.8296 and 0.7530 of SSIM indexes, respectively.

As the future work, we plan to train a Convolutional Neural Network (CNN) to learn the depth of foggy images based on the availability of a large number of training samples. In addition, we plan to use a Generative Adversarial Network (GAN) to generate additional number of samples to manage the data deficiency.

References

[1] K. Sangkyoon, P. Soonyoung, and C. Kyoungho, "A system architecture for real time traffic monitoring in foggy video," in Proceedings of the 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), pp. 1-4, 2015.

[2] L. Tyron and M.Natasha, "Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation," Transportation Research Part C: Emerging Technologies, vol. 76, pp. 35-50, 2017.

[3] W. Chengjia, D. Shizhou, Z. Xiaofeng, P. Giorgos, Z. Heye, and Y. Guang, "Saliencygan: Deep learning semi-supervised salient object detection in the fog of iot," IEEE Transactions on Industrial Informatics, pp. 2667 - 2676, 2019.

[4] H.-M. Hu, Q. Guo, J. Zheng, H. Wang, and B. Li, "Single image defogging based on illumination decomposition for visual maritime surveillance," IEEE Transactions on Image Processing, vol. 28, pp. 2882-2897, 2019.

[5] T. Kokul, C. Fookes, S. Sridharan, A. Ramanan, and U. A. J. Pinidiyaarachchi, "Target-specific siamese attention network for real-time object tracking," IEEE Transactions on Information Forensics and Security vol. 15, pp. 1276-1289, 2019.

[6] G. Liu, S. Liu, K. Muhammad, A. K. Sangaiah, and F. J. I. A. Doctor, "Object tracking in vary lighting conditions for fog based intelligent surveillance of public spaces," IEEE Access, vol. 6, pp. 29283-29296, 2018.

[7] G. Woodell, D. J. Jobson, Z.-u. Rahman, and G. Hines, "Advanced image processing of aerial imagery," in Proceedings of the Visual Information Processing XV, p. 62460E, 2006.

[8] M. I. Anwar and A. Khosla, "Vision enhancement through single image fog removal," Engineering science technology, vol. 20, pp. 1075-1083, 2017.

[9] A. K. Tripathi and S. Mukhopadhyay, "Removal of fog from images: A review," IETE Technical Review, vol. 29, pp. 148-156, 2012.

[10] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. , pp. I-I, 2001.

[11] X. Li and Q. Huang, "Polarization filtering for automatic image dehazing based on contrast enhancement," in Proceedings of the IEEE 9th International Conference on Communication Software and Networks (ICCSN), pp. 1266-1271, 2017.

[12] J.-H. Kim, W.-D. Jang, J.-Y. Sim, and C.-S. Kim, "Optimized contrast enhancement for real-time image and video dehazing," Journal of Visual Communication and Image Representation, vol. 24, pp. 410-425, 2013.

[13] K. He, J. Sun, and X. Tang, "Guided image filtering," in Proceedings of the European conference on computer vision, pp. 1-14, 2010.

[14] R. He, Z. Wang, H. Xiong, and D. D. Feng, "Single image dehazing with white balance correction and image decomposition," in Proceedings of the International Conference on Digital Image Computing Techniques and Applications (DICTA), pp. 1-7, 2012.

[15] H. Kaiming, S. Jian, and T. Xiaoou, "Single image haze removal using dark channel prior," IEEE transactions on pattern analysis and machine intelligence, vol. 33, pp. 2341-2353, 2010.

[16] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," IEEE transactions on image processing, vol. 24, pp. 3522-3533, 2015.

[17] N. Gundawar and V. J. I. J. R. E. T. Baru, "Improved single image dehazing by fusion," International Journal of Research in Engineering and Technology, vol. 3, pp. 432-437, 2014.

[18] V. Thulasika and A. Ramanan, "Single Image Fog Removal based on Fusion Strategy," Digital Image Processing and Pattern Recognition (DPPR), pp. 115-123, 2015 2015.

[19] J.-P. Tarel, N. Hautiere, A. Cord, D. Gruyer, and H. Halmaoui, "Improved visibility of road scene images under heterogeneous fog," in Proceedings of the IEEE Intelligent Vehicles Symposium, pp. 478-485, 2010.

[20] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in IEEE 12th International Conference on Computer Vision, pp. 2201-2208, 2009.

[21] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in Proceedings of the IEEE international conference on computer vision (CVPR), pp. 617-624, 2013.

[22] D. Ngo, G.-D. Lee, and B. J. A. S. Kang, "Improved Color Attenuation Prior for Single-Image Haze Removal," Applied Sciences, vol. 9, p. 4011, 2019.

[23] D. Berman and S. Avidan, "Non-local image dehazing," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1674-1682, 2016.

[24] Y. Zhu, G. Tang, X. Zhang, J. Jiang, and Q. J. N. Tian, "Haze removal method for natural restoration of images with sky," neurocomputing, vol. 275, pp. 499-510, 2018.

[25] A. J. S. P. Galdran, "Image dehazing by artificial multiple-exposure image fusion," Signal Processing, vol. 149, pp. 135-147, 2018.

[26] M. Ju, C. Ding, Y. J. Guo, and D. J. I. T. o. I. P. Zhang, "Idgcp: Image dehazing based on gamma correction prior," IEEE Transactions on Image Processing, vol. 29, pp. 3104-3118, 2019.

[27] S. Salazar-Colores, I. Cruz-Aceves, and J.-M. J. J. o. E. I. Ramos-Arreguin, "Single image dehazing using a multilayer perceptron," Journal of electronic imaging, vol. 27, p. 043022, 2018.

[28] S. K. Dhara, M. Roy, D. Sen, P. K. J. I. T. o. C. Biswas, and S. f. V. Technology, "Color Cast Dependent Image Dehazing via Adaptive Airlight Refinement and Non-linear Color Balancing," IEEE Transactions on Circuits and Systems for Video Technology, 2020.

[29] Z. Ling, X. Li, W. Zou, and M. Liu, "Joint Haze-relevant Features Selection and Transmission Estimation via Deep Belief Network for Efficient Single Image Dehazing," in Proceedings of the 24th International Conference on Pattern Recognition (ICPR), pp. 133-139, 2018.

[30] Z. Li, B. Gui, T. Zhen, and Y. Zhu, "Grain depot image dehazing via quadtree decomposition and convolutional neural networks," Alexandria Engineering Journal, vol.59, pp. 1463-1472, 2020.

[31] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," IEEE Transactions on Image Processing, vol. 25, pp. 5187-5198, 2016.

[32] N. Bharath Raj and N. Venkateswaran, "Single Image Haze Removal using a Generative Adversarial Network," arXiv preprint arXiv:.09479, 2018.

[33] W. Liu, X. Hou, J. Duan, and G. Qiu, "End-to-end single image fog removal using enhanced cycle consistent adversarial networks," IEEE Transactions on Image Processing, vol. 29, pp. 7819-7833, 2020.

[34] B. Das, J. P. Ebenezer, and S. J. T. V. C. Mukhopadhyay, "A comparative study of single image fog removal methods," The Visual Computer, pp. 1-17, 2020.

[35] A. Kumari, P. J. Thomas, and S. Sahoo, "Single image fog removal using gamma transformation and median filtering," in Proceedings of the Annual IEEE India conference (INDICON), pp. 1-5, 2014.

[36] Z. Xu, X. Liu, and N. Ji, "Fog removal from color images using contrast limited adaptive histogram equalization," in Proceedings of the International Congress on Image and Signal Processing, pp. 1-5, 2009.

[37] J. Yu and Q. Liao, "Fast single image fog removal using edge-preserving smoothing," in Proceedings of the IEEE International Conference on ICASSP, pp. 1245-1248, 2011.

[38] T. Kokul and S. Anparasy, "Single Image Defogging using Depth estimation and Scene-Specific Dark Channel Prior," in IEEE International Conference on Advances in ICT for Emerging Regions (ICTer), pp.190-195, 2020.

[39] N. Hautière, J.-P. Tarel, H. Halmaoui, R. Brémond, and D. Aubert, "Enhanced fog detection and free-space segmentation for car navigation," Machine vision applications, vol. 25, pp. 667-679, 2014.

[40] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. J. I. I. T. S. M. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," IEEE Intelligent Transportation Systems Magazine, vol. 4, pp. 6-20, 2012.

[41] R. Sharma and V. Chopra, "A review on different image dehazing methods," International Journal of Computer Engineering Applications, vol. 6, p. 11, 2014.

[42] M. K. Saggu and S. Singh, "A review on various haze removal techniques for image processing," International journal of current engineering technology, vol. 5, pp. 1500-1505, 2015.

[43] G. Yadav, S. Maheshwari, and A. Agarwal, "Fog removal techniques from images: A comparative review and future directions," in Proceedings of the International Conference on Signal Propagation and Computer Technology, pp. 44-52, 2014.

[44] J. Shin, M. Kim, J. Paik, and S. J. I. T. O. M. Lee, "Radiance-Reflectance Combined Optimization and Structure-Guided l 0-Norm for Single Image Dehazing," IEEE Transactions on Multimedia, vol. 22, pp. 30-44, 2020.

[45] S. Yu, H. Zhu, Z. Fu, and J. Wang, "Single image dehazing using multiple transmission layer fusion," Journal of Modern Optics, vol. 63, pp. 519-535, 2016.

[46] J. Pang, O. C. Au, and Z. J. P. A. A. Guo, "Improved single image dehazing using guided filter," APSIPA ASC, pp. 1-4, 2011.

[47] C. Chen, J. Li, S. Deng, F. Li, and Q. Ling, "An adaptive image dehazing algorithm based on dark channel prior," in Proceedings of the 29th Chinese Control And Decision Conference (CCDC), pp. 7472-7477, 2017.

[48] S. Lee, S. Yun, J.-H. Nam, C. S. Won, S.-W. J. E. J. o. I. Jung, and V. Processing, "A review on dark channel prior based image dehazing algorithms," EURASIP Journal on Image Video Processing, vol. 2016, pp. 1-23, 2016.

[49] Y.-K. Wang, C.-T. Fan, and C.-W. Chang, "Accurate depth estimation for image defogging using Markov Random Field," in International Conference on Graphic and Image Processing (ICGIP 2012), p. 87681Q, 2013.

[50] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," International journal of computer vision, vol. 48, pp. 233-254, 2002.

[51] Y. J. Zhang, "A review of recent evaluation methods for image segmentation," in Proceedings of the Sixth International Symposium on Signal Processing and its Applications pp. 148-151, 2001.

[52] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE transactions on pattern analysis machine intelligence, vol. 34, pp. 2274-2282, 2012.

[53] J. B. Zimmerman, S. M. Pizer, E. V. Staab, J. R. Perry, W. McCartney, and B. C. J. I. T. o. M. I. Brenton, "An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement," IEEE Transactions on Medical Imaging, vol. 7, pp. 304-312, 1988.

[54] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in IEEE Conference on Computer Vision and Pattern Recognition, pp. 81-88, 2012.

[55] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing, vol. 13, pp. 600-612, 2004.