Chapter 1

_____

Problem Definition

1.1 Introduction

Images of outdoor scenes often contain haze, fog, smoke or other types of atmospheric degradation caused by particles which are present in the air in form of pollutants, aerosols. These particles in the atmospheric medium absorbing and scattering light as it travels from the source to the observer and result in two fundamental phenomena called 'direct attenuation' and 'airlight '. 'Direct attenuation' reduces the contrast and 'airlight' add the whiteness in the scene. As a result, images recieved by camera or our eyes, seems hazy and they lose their contrast, visual vividness and color fidelity(see Fig. 1.1a). Therefore, removing haze from images is an important and widely demamded for outdoor-vision application used for object recognition, tracking, navigation and satelite imaging. The main challenge lies in the *ambiguity* of the problem. Haze attenuates the light reflected from the scenes, and further blends it with some additive light in the atmosphere. The target of haze removal is to recover the reflected light (*i.e.*, the scene colors) from the blended light. This problem is mathematically ambiguous: there are an infinite number of solutions given the blended light. How can we know which solution is true? We need to answer this question in haze removal.

1.2 Problem Identified

In terms of mathematics, ambiguity is because the number of equations is smaller than the number of unknowns. The methods in computer vision to solve the ambiguity can roughly categorized into two strategies. The first one is to acquire more known variables, *e.g.*, some fog removal algorithms capture multiple images of the same scene under different settings (like polarizers). But it is not easy to obtain extra images in practice. The second strategy is

to impose extra constraints using some knowledge or assumptions known beforehand, namely, some "*priors*". This way is more practical since it requires as few as only one image. To this end, we focus on *single* image fog removal in this thesis. The key is to find a suitable prior.

1.3 Proposed technique

In the thesis work a new technique is proposed to remove fog from single color images keeping following aspects in mind:

- ➢ To provide high perceptual quality of dehazed image. .
- ➢ To provide less color blurriness in dehazd images.
- ➢ To preserve edges significantly.
- ➢ Method should be fast, simple and effective.

1.4 Tool used

MATLAB is used as simulator to implement the techniques. MATLAB provides highly computing environment and advanced in-built function for image processing.

Why matlab...?

Matlab is an integrated technical computing environment that combines numeric computation, advanced graphics and visualization. It is a high level programming language that can communicate with its cousins, e.g. FORTRAN and C. MATLAB allows matrix manipulation, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs in other languages.The language, tools and built-in math functions enables to explore multiple approaches and reach a solution faster than other programming languages. Besides these Matlab is used for:-

- ➢ Communications Systems
- ➢ Computational Finance
- ➢ Control Systems
- ➢ Digital Signal Processing

(a)   Input hazy image

(b) Haze removal result of our proposed approach

Figure 1.1: Haze removal from a single image by using proposed approach.

1.5 Achievements

The following achievements or we can say objectives were achieved in this work:

 ➢ Almost all the techniques other than Dark channel prior followed by soft mating are not able to provide sufficient results for image dehazing.

 ➢ Soft matting performs well in transmission map refinement but small changes of parameters from optimized value make the image worse.

 ➢ we develop a simple but effective prior called the median dark channel prior followed by guided filtering , to remove fog from a single image.

 ➢ Experiments demonstrate that our method is very successful in severe fog condition (*e.g.*, Fig. 1.1 b) and outperforms many previous approaches.

1.6 Organization of Thesis

The remaining part of the thesis is organized into following chapters:

Chapter 2: Introduce the "atmospheric dichromatic model" to describe formation of fog images and Literature survey on exixting fog removal techniques.

Chapter 3: Proposed technique is described to remove fog from a single image.

Chapter 4: Provide a review of two technique for further refining the transmission map obtained by DCP and MDCP.

Chapter 5: we described Post processing opertor and performance evaluation matrices.

Chapter 6: Simulation result and comparison with existing fog removal techniques.

Chapter 7: we conclude this thesis with a summary of findings, open question, and potential topic for future research.

Chapter 2

_____

Background

This chapter introduces the 'Atmospheric Dichromatic model' that has been widely used to describe the formation of fog ,Haze. We also investigate an overview of previous work on visibilty restoration by using both multiple-image and single image cases,with their advantages and limitations.

2.1   Atmospheric dichromatic model

In computer vision, 'Atmospheric dichromatic model' [1] or 'Fog imaging model' has been widely used to describe the formation of haze images. Pictorial description of the model is shown in **fig. 2.1**.

Formation of fog is characterized by following equation:-

$$\widehat{X}(u,v) \ = \ X(u,v)t(u,v) \ + \ a_{ir}(1 - t(u,v)) \qquad\qquad (2.1)$$

RGB, fog image $\widehat{X}$ at pixel position (u,v) is:

$$\widehat{X}(u,v) = [\ \widehat{X}(u,v,r), \widehat{X}(u,v,g), \widehat{X}(u,v,b)]\ ^{T} \qquad\qquad (2.2)$$

RGB, image X without fog at pixel position (u,v) is:

$$X(u,v) = [\ X(u,v,r), X(u,v,g), X(u,v,b)]\ ^{T} \qquad\qquad (2.3)$$

Where, (u,v) is the location of the pixel, $\widehat{X}$ is the observed fog/haze image , X is the actaul image aims to be recovered, a_{ir} is the global atmospheric light which is asumed to be constant, and t is the medium transmission coefficient.

In equation (2.1), the first term $X(u,v)t(u,v)$ is called '*direct attenuation*' which produce a multiplicative distortion of scene radiance and it reduce the contrast, and the later term

$a_{ir}(1 - t(u,v)$ is called the *'local atmospheric light'* which produce additive effect and it add the whiteness in the scene. Intuitively, the image received by the observer is the combination of the attenuated version of underlying scene with an additive atmospheric light, which represents the color of fog. The ultimate goal of fog removal is to find $\boldsymbol{X(u,v)}$ which also require knowledge of three unknown parameter $\boldsymbol{\beta}$, $\mathbf{t}$ and $\boldsymbol{a_{ir}}$. From this model, it is apparent thet fog removal is an under-constrained problem. In a grayscale image, for each pixel there is only 1 constraint but 3 unknowns;for an RGB color image, there are 3 constraint but 7 unknown (assuming t is the same for each color channel).
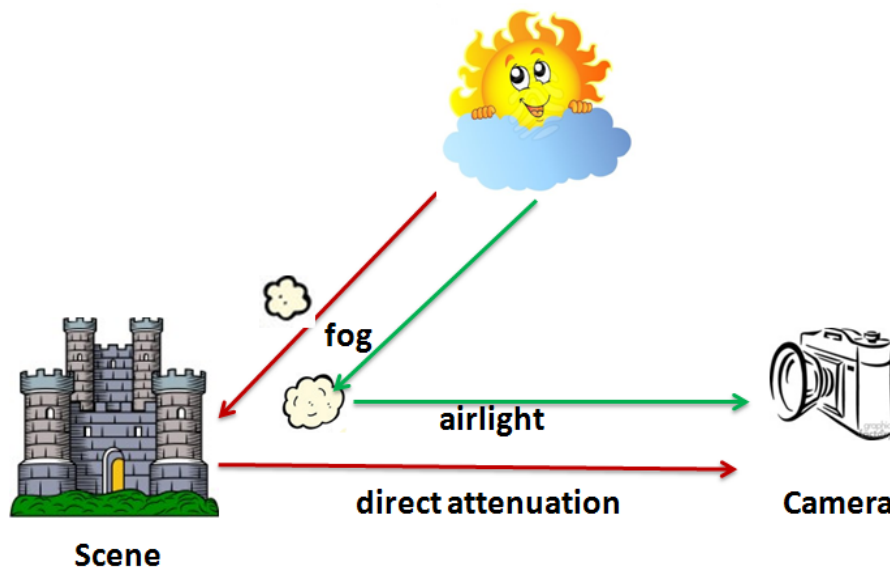

Fig.2.1 Atmospheric Dichromatic Model

In order to make the problem easier, the atmosphere is assumed to be homogenous. This has two simplifying consequence: atmospheric light is constant throughout the image: meaning it has to be estimated only once and transmission $t(u,v)$ follow the Beer-Lambert law:-

$$t(u,v) = e^{-\beta d(u,v)} \tag{2.4}$$

Where β is the atmospheric attenuation coefficient due to scattering and $\boldsymbol{d(u,v)}$ represent the distance from the observer to the scene at pixel $\boldsymbol{(u,v)}$. when we assume the atmosphere

is homgenous, it restricts β to be invariant. The medium transmission coefficient has a scalar value within $0 \leq t(u, v) \leq 1$ for each pixel which attenuates the target color.

On putting value of $t(u, v)$ in eq. (2.1)

$$\hat{X}(u, v) = X(u, v)e^{-\beta d(u,v)} + a_{ir}(1 - e^{-\beta d(u,v)}) \tag{2.5}$$

Equation (2.5) indicates that the scene radiance is attenuated exponentially with the distance(d). If we recover the transmission, we can also recover the scene depth up to an unknown scale.

Once t and a_{ir} are known, the image $X(u, v)$ can be found by using equation (2.1) as :-

$$X(u, v) = \frac{\hat{X}(u,v) - a_{ir}(1 - t(u,v))}{t(u,v)} \tag{2.6}$$
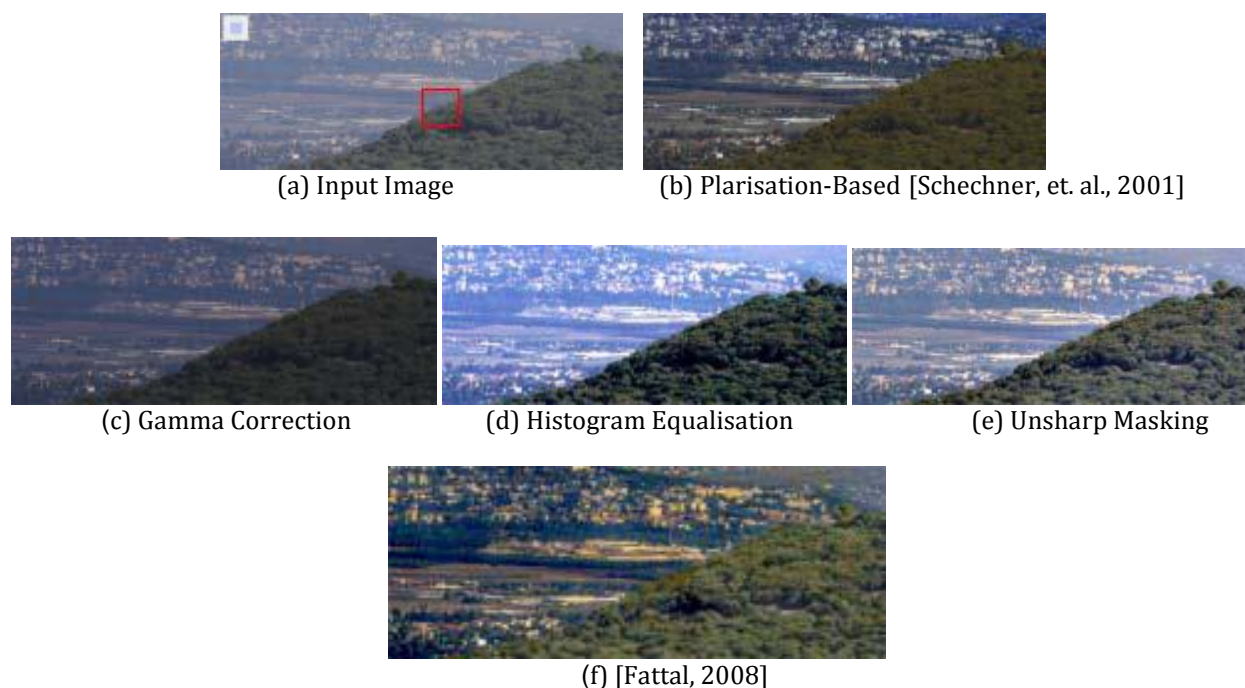
2.2 Related work

Many dehazing algorithms have been developed by different workers that were very active in this field of study in recent years. In this section, we review the previous haze removal methods. We do not discuss the technical details of these methods. Instead, we are interested in how they introduce extra constraints. All the methods are reformulated in a same framework, though they may be expressed in different forms in the original works. We believe that this is helpful to better compare these methods.

2.2.1 Overview of Dehazing methods

To achieve the goal of fog removal and visibility improvement several methods have been proposed from past few decades. The methods can basically be divided into two groups, non-model based methods try to enhance the contrast of a hazy image using simpler computer vision techniques such as gamma correction, unsharp masking or histogram equalization. However, these methods do not always maintain color fidelity. These algorithms are designed for images whose properties are roughly constant across the image. Model-based methods are those that need either multiple or single image for dehazing. The question of what method gives the best visibility improvement should best be answered by

the human, as he will be the ultimate judge using the dehazed image in applications such as surveillance scenarios. A comparison of a variety of methods can be seen in **figure-2.2**.

(a) Input Image

(b) Plarisation-Based [Schechner, et. al., 2001]

(c) Gamma Correction

(d) Histogram Equalisation

(e) Unsharp Masking

(f) [Fattal, 2008]

Figure 2.2: Comparison of different dehazing techniques. *Source:* [Fattal, 2008]

Figure 2.2(a) shows the input image, right next to it in 2.2(b) is the result of a polarization based defogging. The next row 2.2(c)-(e) contains images, whose visibility have been improved by simpler image processing methods such as gamma correction, histogram equalisation and unsharp masking, respectively. The last row contains the results of Fattal. Fattal's results are the best for the single image dehazing in this comparison. Recently, He et al. proposed a new method that produces comparable results to those of Fattal and it is one of the best quality avialable today.

2.3 Non model based Dehazing

The Non-model based dehazing methods try to eliminate fog by enhancing the contrast of an image, by using simpler computer vision techniques such as gamma correction, unsharp masking or histogram equalization. The middle row in fig. 2.2 shows the results of the non-

model based dehazing. The striking similarity is the blue hue that the three have in common. This can only be eliminated with model based techniques. Infect non-model based methods can improve visibility for the human eye, but these were not developed to dehaze images. So we can use these methods, to further improve an already dehazed image. Thus for this purpose, they shouldn't be used exclusively, but in combination with a dehazer. For the sake of completeness however, the three introduced concepts will now be described very briefly.

2.3.1 Unsharp Masking

The idea behind unsharp masking is to emphasize edges. An unsharp masking algorithm can detect edges and alter the levels of brightness on both sides of the edge in a way that the darker side gets even darker towards the edge and the lighter parts get even lighter towards the edge. This produces an overshoot and undershoot, respectively of the brightness curve of the pixels at the edge. This is illustrated in fig. 2.3.
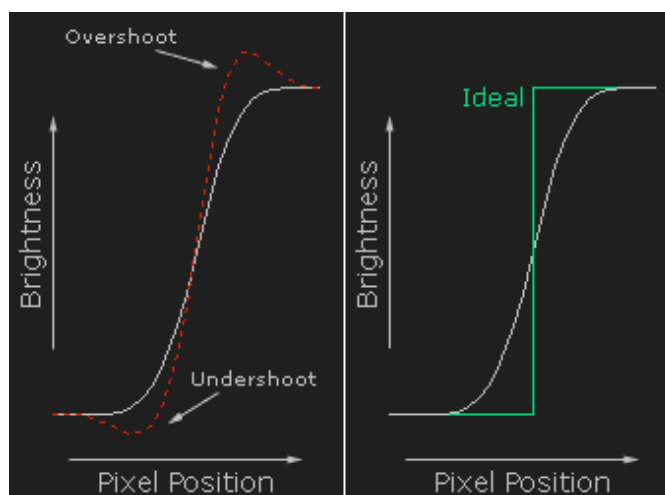
Figure 2.3.: principle of unsharp masking. *Source:* [cambridgeincolour.com,)

In fig.2.3, left plot shows the overshoot and undershoot resulting from the masking versus the actual brightness curve, right plot shows the ideal rectangular curve versus the actual curve of the input image.

2.3.2 Gamma Correction

Gamma correction refers to a nonlinear operation that amplifies or reduces the luminance intensity of an image. This operation is performed on each pixel in the same way, no matter its original value. By lowering in the power-law expression for the gamma correction:

$$B_{out} = B_{in}^{\gamma} \tag{2.7}$$

the contrast levels may be raised in dark images with low contrasts.
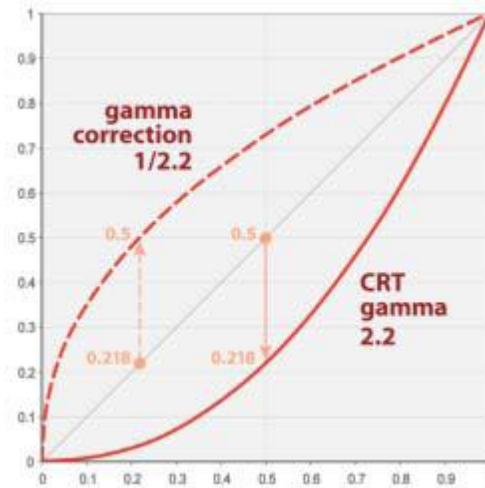
Figure 2.4.: Principle of gamma correction, *Source:* [wikipedia.org]

Here, B_{in} and B_{out} respectively denote the brightness levels before and after the gamma correction. The principle of gamma correction can be seen in the plot of **fig. 2.4**, the input brightness curve is here as an example the input of a CRT display with = 2.2, with a gamma correction curve of = 1/2.2 the actual linear curve of interest can be restored. The ordinate shows the input value(brightness) and the abscissa shows the output value(brightness).

2.3.3. Histogram Equalisation

The histogram equalisation is a method using the image's histogram in order to improve contrast. Since only so much of brightness values can be displayed, all gradations of brightnesses and therefore contrasts must be within the brightness bandwidth. Thus the

best results can be obtained, when spreading out the most frequent intensity values over the entire histogram. A disadvantage of this method is that it may increase the noise by discriminating it from the actual usable signal. However, this method is one of the more advanced methods to improve image contrast, and is of the three mentioned in this thesis the most resource intensive, but usually also the one with the best results.
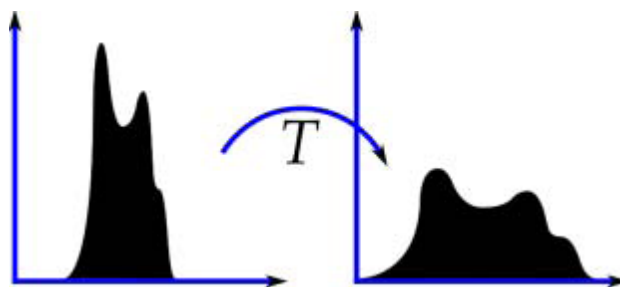
Figure 2.5.: Principle of histogram equalisation . *Source:* [wikipedia.org]

Fig. 2.5 illustrates the basic idea behind histogram equalisation. The left plot shows the input, where not all possible brightness values are used and the right plot after the transmission. This method generally improves the global contrast, locally however spots in the image with close brightness values may not be improved in some cases.
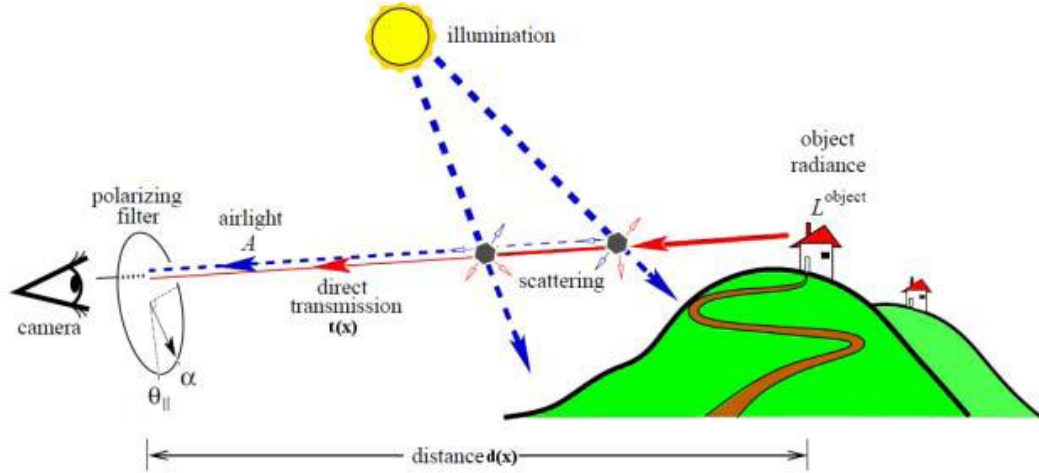
2.4 Model based dehazing

These methods use 'Atmospheric dichromatic model' to estimate the fog contribution in images and then recover the scene contrast. These methods require extra information about the imaging environment and provide better results in comparison with non-model based methods. This extra information can be in terms of estimation of transmission-map, or some heuristic assumption. Earlier dehazing techniques were based on multiple images or supplemental equipment. Some popular model-based dehazing techniques are described below:

2.4.1 Polarization Based Visibility Improvements

Polarization-based dehazing techniques proposed by Schener et. al [3] are part of the multi-image group, they usually use two input images taken with two differently polarised filters, one after another, to produce one dehazed image. Using a polarizer attached in the camera

lens, these methods take two images of the same scene under two polarization state, see **figure 2.6**.

Figure 2.6.: Model for polarization-based dehazing. *Source:* [Namer and Schechner]

Assume the direct attenuation is completely unpolarized, the haze imaging equations of the two images are:

$$\hat{X}^{.\|}(u,v) = X(u,v)t(u,v) + a_{ir}^{c\,.\|}(1 - t_1(u,v))$$

$$\hat{X}^{\perp}(u,v) = X(u,v)t(u,v) + a_{ir}^{c\,\perp}(1 - t_1(u,v)) \tag{2.8}$$

Here $.\|$ and $\perp$ denote two states.

The equations (2.8) provide $6N$ constraints, together with $3N$ unknown $\boldsymbol{X}$, N unknown $\boldsymbol{t}$, 3 unknown $A_c^{\|}$ and 3 unknown $A_c^{\perp}$: in total $4N+6$ unknown variables. The problem becomes over-constrained and can be solved. An example of polarisation-based dehazing results, is given in **fig. 2.7**.

❖ **Limiataions of polarization based methods:**

A main drawback of the polarization-based methods is the camera settings, capturing two strictly aligned polarized images is troublesome in practice. Another disadvantage, is that it require dedicated hardware for rotating the polarizer and it cannot be applied to single image.
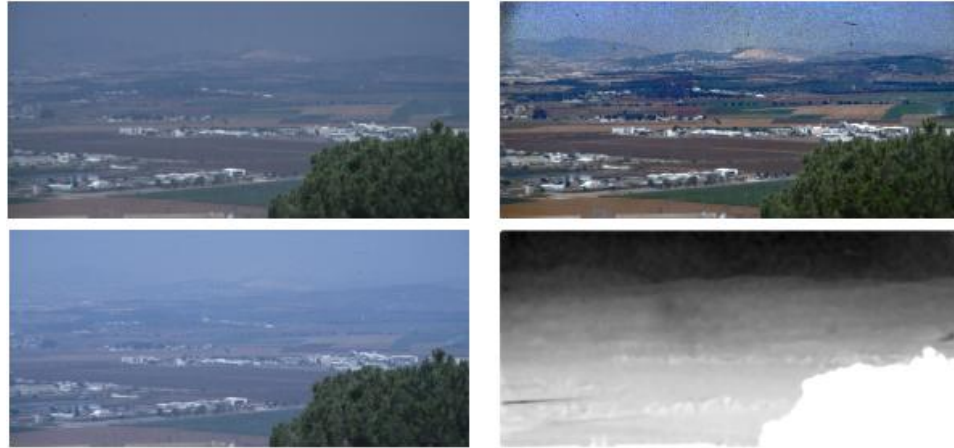
Figure 2.7: Haze removal based on Polarization method. On the left are two images taken in two polarizer states. On the right are the estimated scene radiance and depth.

Figure 2.8: Haze removal based on varying atmospheric conditions. On the left are two images taken in two hazy conditions. On the right are the estimated image and depth (the sky is ignored).

Fig. 2.9.: Haze removal based on Given Depth. From left to right: fog image, given 3D structure, defog image

2.4.2 Methods Based on Varying Atmospheric Conditions

The method proposed by Narsimhan & Nayar in [1-2] takes at least two images of the same scene, under different atmospheric conditions (Fig. 2.8 left). The two images are strictly aligned. Thus the two images shares the same depth $d(u, v)$ and the same reflectance $\rho_c(u, v)$. It further assumes that the atmospheric light a_{ir} is the only light source of the scene. Then the reflection **X** satisfies:

$$X(u, v) = \rho_c(x) a_{ir}^c \tag{2.9}$$

Under the constant β assumption, the haze imaging equations of the two images are:

$$\hat{X}^1(u, v) = \rho_c(u, v) a_{ir}^{1c} t_1(u, v) + a_{ir}^{1c}(1 - t_1(u, v))$$
$$\hat{X}^2(u, v) = \rho_c(u, v) a_{ir}^{2c} t_2(u, v) + a_{ir}^{2c}(1 - t_2(u, v)) \tag{2.10}$$

❖ **Limaitations of above method:**

Main limitation of this method is that above analysis is valid only when $t_1(u, v) \neq t_2(u, v)$. To ensure this condition, the two images must be taken under very different atmospheric conditions, *e.g.*, one in denser haze and the other in thinner haze. This is not an easy task: the weather may remain unchanged in several minutes or even hours.

2.4.3 Method based on Deep Photo System

A rather different approach was proposed by Kopf et al. in 2008 called the *Deep Photo System*[3]. Kopf developped a data-driven defogging procedure by employing a registration process to align the photograph within an existing 3D model. The 3D-model came from satellite image data obtained from google earth and bingmaps, NASA radar images. This way the method does not need to estimate the distances in the scene, but will get the exact distances right away, assuming that such kind of georeferenced digital terrain and urban models are available.

Under the constant-β assumption, the haze imaging equation becomes

$$\hat{X}(u,v) = X(u,v)e^{-\beta d(u,v)} + a_{ir}^{c}(1 - e^{-\beta d(u,v)}) \qquad (2.11)$$

We have $3N$ equations in (2.11), together with $3N$ unknown X, 3 unknown a_{ir}^{c}, and one unknown β: in total $3N + 4$ unknowns. The problem is almost well-posed.The extra unknown variables (a_{ir}^{c} and β) can be estimated by 3D model. See **Fig. 2.9 and fig. 2.10(c)** for an example. After acquiring the depth information, the method of Kopf et al. estimates the airlight and the attenuation coefficient similarly to the other haze removal methods and then basically solves haze imaging equation(2.1) .

❖ **Limitations of kopf implementation:**

Since it heavily relies on 3D model, if no 3D model of the scene is available, no dehazing can be performed. Another limitation is that when the alignment between the photograph and the model may not be completely accurate due to unprecise 3D-models, it gives poor result.

2.4.4 Summary of Multiple-image fog removal methods

The multiple-image fog removal methods share some common advantages and limitations, some of them are listed as:-

Advantages:

- ➢ These methods posses fast running time.
- ➢ The computation is often pixel-wise and no complex optimization is needed.
- ➢ Another benefit is that they may handle some special situations. For example, the dichromatic methods can handle night images, and the polarization-based or depth-based methods allow it to vary across color channels.

Limitation:-

- ➢ The common limitation of these methods is that the extra images are not easily available.

- ➢ They all require special settings with carefully calibration.
- ➢ They require dedicated hardware for rotating the polarizer .
- ➢ This is not practical in most cases, such as for hand-held cameras and outdoor surveillance systems.
- ➢ Though these methods , increases the number of known variables, at the same time it brings in more unknowns. So the setting must under certain constraints to avoid too many unknowns being introduced.
- ➢ when only a single degraded image is available as an input, these methods fails.

The limitations of these methods motivate the development of single image fog removal methods.

2.5 Single image fog removal

Single image fog removal methods have to rely on some priors. The priors can be statistical/physical properties, heuristic assumptions, and application-based rules. In recent years a significant progress has been made [4–8] that remove fog from a single image without using any other extra source of information. Some of them are:

2.5.1 Fattal's Method

Fattal[4] introduced a new technique in 2008 for single image dehazing that produces qualitatively great results on hazy images. Fattal's method is based on Independent Component Analysis (ICA). ICA is a statistical method to separate two additive components from a signal. Fattal adapts this method to remove fog from images.

Denote the luminance by $l(\mathbf{x})$. Then $X(u, v)$ is $\mathbf{R}l(\mathbf{u,v})$, and the haze imaging equation becomes:

$$\hat{X}(u, v) = R_c l(u, v) t(u, v) + A_c (1 - t(u, v)) \qquad (2.12)$$

The two scalar components $l(\mathbf{u,v})$ and $t(\mathbf{u,v})$ are to be separated. The luminance $l(\mathbf{u,v})$ depends on the illumination, object reflectance, and the scene geometry, whereas the

transmission $t(\mathbf{u,v})$ depends on the depth and the property of the haze. Fattal's assumption is that: these two components are due to unrelated sources and therefore statistically independent. Under this assumption the method can recover the reflectance of the pixels and provides extra constraints. We omit the technical details. Unlike Tan's method, this method is physically valid. The results often looks more natural and visually pleasing (see **Fig. 2.2(f)** and **Fig. 2.10(d)**).

❖ Limitations of Fattal work:-

The main limitation of this method results from the locally based statistics. To ensure the statistics is reliable, the method requires the two components $l(\mathbf{u,v})$ and $t(\mathbf{u,v})$ to vary significantly in a local patch. This condition is not always satisfied. For example, a local patch of a distant scene usually exhibit negligible variance in $t(\mathbf{u,v})$. To handle this problem, the method only apply the ICA to some reliable patches. "Moreover, as the statistics is based on color information, it is invalid for grayscale images and difficult to handle dense haze where the different component are difficult to resolve."

2.5.2 Tan's Method

Haze reduces visibility. Removing the haze will enhance the visibility of the image. In [5] Tan proposes a method to maximize the visibility. Tan's method is based on the optical model given by:

$$\hat{X}(u,v) = X(u,v)t(u,v) + a_{ir}(1 - t(u,v)) \qquad (2.13)$$

The first term in this equation is 'direction attenuation' and the second term corresponds to the 'airlight'. In this formula there are more unknowns than known. Nevertheless, there are some clues or observations that Tan makes use of in his algorithm: the transmission t is treated to be constant inside a patch, so $\nabla\hat{X}(u,v) = t\nabla X(u,v)$. The total visibility inside the patch is defined as the sum of the gradient magnitude:

$$\sum_{c,x}|\nabla X^c(u,v)| = \frac{1}{t}\sum_{c,(u,v)}|\nabla\hat{X}(u,v)| \qquad (2.14)$$

The visibility of the recovered image $\sum_{c,(u,v)} |\nabla \widehat{X}(u,v)|$ will keep increasing when the transmission t is decreasing. Tan supposes the atmospheric light is the light source of the scene, so $X^c(u,v) = \rho_c(u,v) a_{ir}^c$ with $0 <= \rho_c(u,v) <= 1$. This leads to the following constraint:

$$0 <= X^c(u,v) <= a_{ir}^c \qquad (2.15)$$

In Tan's method, the value $X^c(u,v)$ outside this range is truncated. This operation prevents the visibility of $\sum_{c,x} |\nabla X^c(u,v)|$ from increasing because the truncated values provides zero gradients. The optimal t is the value that maximizes the visibility. This computation is performed in each patch, providing a constraint for each pixel. Tan's method has the advantages of enhancing the visibility. But it has some limitations. The drawbacks of the implementation are the speed of the algorithm and the restored images often look over-saturated and unnatural as in **fig.2.10(b)**.

2.5.3 Tarel's et al. method

One of the disadvantages of the algorithm discussed in the previous section is its speed. Tarel et al.[7] proposed an algorithm which is faster and claims to have similar or in some cases better results. Tarel et al. model is based on the same model as Tan uses

$$\widehat{X}(u,v) = X(u,v)\left(1 - \frac{v(x)}{a_{ir}}\right) + v(x) \qquad (2.16)$$

$$X(u,v) = \frac{\widehat{X}(u,v) - v(x)}{1 - v(x)} \qquad (2.17)$$

As it is not possible to separate the depth and extinction coefficient the atmospheric veil is introduced $V(x) = a_{ir}(1 - t(u,v))$. The rewritten fog model is shown in Equation (2.16), $\widehat{X}(u,v)$ is the input image and $X(u,v)$ is the restored output image. Due to whitebalance in the pre-processing stage $a_{ir} = [1; 1; 1]^T$, allowing the equation to be rewritten to Equation (2.17) which isolates $X(u,v)$. When the airlight is pure white the atmospheric veil adds whiteness to the image, the amount of whiteness added depends on the depth of the object.

Because the whiteness gives information about the depth and the depth is proportional to the atmospheric veil it is possible to base the atmospheric veil on the whiteness of the image. The mentioned constraints can be formulated as Equation(2.18)

$$0 \leq V(x) \leq w(x) \qquad (2.18)$$

$$A(x) = median[w(x)] \qquad (2.19)$$

$$B(x) = A(x) - median(|w - A|)(x) \qquad (2.20)$$

$$V(x) = max(min(p\ B(x), w(x), 0) \qquad (2.21)$$

Besides the constraints as stated in Equation (2.18), the atmospheric veil is smooth except along depth discontinuities. Using a median filter, W(x) can be smoothened while keeping edges. The smoothened version of W(x) is stored in matrix A(x). Equation (2.19) shows this first step. The next step is preventing contrasted texture to be affected too much, this is done by subtracting the local standard deviation of A(x) from W(x). This operation is stored in matrix B(x) as shown in Equation (2.20). Finally V (x) is inferred using Equation (2.21). In this equation **p** is used to control the strength of the restoration. Main drawback of tarel implementation is that ir require too many parameter for adjustment.

2.5.4 Tripathi et.al implementation

In 2013, Tripathi et al. [12] proposed a fog removal algorithm that uses single image. Tripathi observed that fog has no effect on hue of the scene. Other two component saturation and intensity are affected by fog. This modification reduces calculation significantly, as processing is needed only over the saturation and intensity plane. Anisotropic diffusion is used for the refinement of transmission-map. Owing to the use of anisotropic diffusion,this technique is computationally intensive and require a number of parameter for optimal result. Fast implementation of anisotropic diffusion is a challenging task.

(a) Input Image   (b) [Tan, 2008]

(c) [Kopf et al., 2008]   (d) [Fattal, 2008]

Figure 2.10: Comparison of different model based single image dehazing techniques. *Source:* [He et al.]

(a) [He et al, 2011]

(b) [Proposed method]

Figure 2.11.: Comparison of dehazing result using the algorithm of [He et al], using 2.10.(a) as input image. Figure (b) shows the dehazing result by using the proposed algorithm.

2.5.3 He et al. implementation

 A very promising single image dehazing technique, that has been used as the basis of our thesis developed by He, Sun and Tang [8] in 2010 called the *Dark Channel Prior* .

Although every assumption limits the algorithm to specific case, the main assumption here seems to work for most outdoor scenes, except for those where "the scene object is inherently similar to the airlight over a large location and no shadow is cast on the object". The main prior in this method is, as the name lets assume, the dark channel prior, which is a statistical based assumption of fog-free outdoor images. The prior says, that in most of the local patches that aren't sky, will contain one or more pixel whose intensity is very low and may be even close to zero in atleast one of the color channels (RGB). Low value of intensity is caused by lacking of color in a channel which may be due to shadows, dark objects,colorful object.

 In the hazy image then, these dark pixels can be used to determine the airlight. The dark channel of an image **X** is defined as:-

$$\theta_D(u,v) = \min_{m,n \in \Omega(u,v)} \left(\min_{c \in (r,g,b)} \frac{\tilde{X}^c(u,v)}{a_{ir}^c} \right) \approx 0 \tag{2.16}$$

Where X^c is a color channel of X , **Ω** is a local patch centered at(u,v) . $\boldsymbol{a}_{ir}^c$ is atmospheric light of the corresponding channel .

This $\boldsymbol{\theta_D}$ can be used to create a coarse estimate of transmission given by:

$$\tilde{t}_D(u,v) = 1 - w \left(\min_{m,n \in \Omega(u,v)} \min_{c \in (r,g,b)} \frac{\tilde{X}^c(u,v)}{a_{ir}^c} \right) \tag{2.17}$$

The scaling parameter, $\boldsymbol{w}$, takes a value from 0 to 1, corresponding to the amount of fog left in the image. A typical value of $\boldsymbol{w}$ to preserve depth is 0.95. Before we move on to the smoothing step that they usesd to refine the transmission estimate, let us take a closer look at the meaning of (2.17). To simplify our exploration, suppose we only have one color channel such that image $\boldsymbol{\theta_D}$ becomes

$$\theta_D(u,v) = \min_{m,n \in \Omega(u,v)} \left(\frac{X(u,v)}{a_{ir}} \right)$$

$$= \min_{m,n \in \Omega(u,v)} \left(\frac{X(u,v)t(u,v)}{a_{ir}} + 1 - t(u,v) \right) \tag{2.18}$$

If we assume that, within patch $\Omega(u,v)$, the radiance of the target is smooth $X(u,v) = x$, then we have

$$\theta_D(u,v) = \min_{m,n \in \Omega(u,v)} \left[\left(\frac{x}{a_{ir}} - 1 \right) t(u,v) + 1 \right]$$

$$= 1 + \left(\frac{x}{a_{ir}} - 1 \right) \left(\min_{m,n \in \Omega(u,v)} t(u,v) \right) \tag{2.19}$$

In (2.19), the DCP value $\boldsymbol{\theta_D}$ is a function of transmission $\boldsymbol{t(u,v)}$ in a patch with smooth radiance (no texture). Next, if we assume that the transmission is smooth $\boldsymbol{t(u,v) = t}$, which means smooth depth within patch $\boldsymbol{\Omega(u,v)}$, then we have

$$\theta_D(u,v) = \min_{m,n \in \Omega(u,v)} \left[\left(\frac{X(u,v)}{a_{ir}} - 1 \right) t \right]$$

$$= \left(\frac{1}{a_{ir}} \left(\min_{m,n \in \Omega(u,v)} X(u,v) \right) - 1 \right) t \tag{2.20}$$

In (2.20), with smooth depth but high texture in patch $\boldsymbol{\Omega(u,v)}$, the DCP is strictly a function of the minimum value from $X(m,n)$ but yet attenuated by t. This also means that the DCP of a texture at a close distance $\boldsymbol{\Omega(u_c,v_c)}$ will be greater than the DCP of the same texture at a far distance $\boldsymbol{\Omega(u_f,v_f)}$, i.e.,

$$\theta_D(\boldsymbol{u_c,v_c}) > \theta_D(\boldsymbol{u_f,v_f}) \tag{2.21}$$

Based on our simplifications, (2.19) and (2.20) show that there is a loose relationship between the DCP value $\boldsymbol{\theta_D}$ and transmission $\boldsymbol{t}$. The relationship breaks down when the depth variation is not smooth and when the texture is high (occlusion boundaries Fig. 3.1). To account for the DCP breaking down in textured and varying depth region, He used a spectral matting method to smooth or refine the transmission estmate. The refinement method used by HE works well but require several seconds to process because of the

generation of the laplacian matrix that initialize the spectral matting method see fig 2.11(a). A detailed explanation of methods used for refinement of estimated transmission-map is described in chapter-4.

Drawback of He et al. implementation

- ➢ Processing time is slow and cannot be used in real-time systems.
- ➢ It is quite possible that some areas of the image do not fulfill the dark channel prior. As a result, these methods may suffer from one or more of the following problems:-
    - ❖ First, It usually result in color over saturation, becaues it cannot satisfy the limit requirement that the obtained dark channel should be no brighter than the minimum color channel.
    - ❖ Secondly, it usually introduce artefact into the smooth areas, where the dark channel prior is unreliable.
    - ❖ Restored image usually looks dark after haze removal see Fig. 2.11.

Chapter 3

_____

Proposed method

There are several methods for estimating the fog contribution in a single image. One of the most successful of these methods is known as the dark channel prior, which is used as the basis for fog estimation in this thesis.

3.1 Median Dark channel Prior

We propose a visibility restoration algorithm based on the utilization of median filtering operation. Our proposed method is an improvement to the Dark Channel Prior [8] by replacing the second minimum operator in equation (2.16) with a median operator. The median operator performs a non-linear filtering operation which can effectively suppress impulsive noise components while preserve edge information in detail areas and permitting dehazing in smooth areas. In the case of noise this works because the median is unaffected by the outlying values from the noise as long as there is not too much noise.

The proposed MDCP for an image X is given as:

$$\theta_M(u,v) = \underset{m,n \in \Omega(u,v)}{med} \left(\underset{c \in (r,g,b)}{min} \hat{X}^c(u,v) \right) \tag{3.1}$$

Where, $\boldsymbol{\theta_M(u,v)}$ represents the "median dark channel" of an image at pixel location (u,v).

3.2 Atmospheric light estimation

The Median dark channel prior effectively identifies the relative amount of fog in an image, and thus can aid in estimating the atmospheric light, $\boldsymbol{a_{ir}}$. In the haziest regions of the image, transmission tends toward zero, and the atmospheric light contribution dominates the scene radiance. This is apparent from examining Eq. (2.1):

$$\lim_{t(u,v)\to 0} \hat{X}(u,v) = \lim_{t(u,v)\to 0} X(u,v)t(u,v) + \lim_{t(u,v)\to 0} a_{ir}(1 - t(u,v))$$

$$= a_{ir} \tag{3.2}$$

Thus the brightest regions of $X(u, v)$ contain mostly atmospheric light. The approach adopted in this thesis, following [7], is estimating $\boldsymbol{a_{ir}}$ as the brightest intensities in each color channel chosen from the 0.1% brightest pixels in $X(u, v)$, which correspond to the top 0.1% brightest pixels in the dark channel.

After estimation of airlight normalize the MDCP :

$$\tilde{\theta}_M(u, v) = \underset{m,n\in\Omega(u,v)}{med} \left(\frac{\hat{X}^c(u,v)}{a_{ir}^c} \right) \tag{3.3}$$

3.3 Transmission-map Estimation

Assuming now that $\boldsymbol{a_{ir}}$ has been estimated, an estimate of the transmission map is obtained by normalizing the fog imaging equation (2.1) by $\boldsymbol{a_{ir}}$.

$$\frac{\widehat{X}^c(u,v)}{a_{ir}} = X^c(u,v)\frac{\tilde{t}(u,v)}{a_{ir}} + (1 - \tilde{t}(u,v)) \tag{3.4}$$

Note that we normalize each color channel c independently. Then we compute the dark channel on both sides of this equation, *i.e.*, we insert the minimum operators:

$$\underset{m,n\in\Omega(u,v)}{min} \underset{c\in(r,g,b)}{min} \frac{\widehat{X}^c(u,v)}{a_{ir}^c} = \underset{m,n\in\Omega(u,v)}{min} \underset{c\in(r,g,b)}{min} \left(X^c(u,v)\frac{\tilde{t}(u,v)}{a_{ir}^c}\right) + (1 - \tilde{t}(u,v)) \tag{3.5}$$

$$\underset{m,n\in\Omega(u,v)}{min} \underset{c\in(r,g,b)}{min} \frac{\widehat{X}^c(u,v)}{a_{ir}^c} = 0 + (1 - \tilde{t}(u,v)) \tag{3.6}$$

$$\tilde{t}(u,v) = 1 - \underset{m,n\in\Omega(u,v)}{min} \underset{c\in(r,g,b)}{min} \frac{\widehat{X}^c(u,v)}{a_{ir}^c} \tag{3.7}$$

$$\tilde{t}_M(u,v) = 1 - \tilde{\theta}_M(u,v) \tag{3.8}$$

Where, $\tilde{\theta}_M(\boldsymbol{u}, \boldsymbol{v})$ represent normalized median dark channel of fog image. Superscript c signifies a specific color channel, i.e. the red, green, or blue. Note that $\boldsymbol{a_{ir}}$ is a vector quantity containing a separate value for each color channel, and so $\boldsymbol{a_{ir}^c}$ is a scalar quantity referring to one individual value. Since $1 - \tilde{t}(u,v)$ is constant it is taken out of the **min** operation. Furthermore, since $(\boldsymbol{t})$ and $\boldsymbol{a_{ir}^c}$ are both positive, the dark channel prior implies that

$\min_{m,n \in \Omega(u,v)} \min_{c \in (r,g,b)} \cdot \left(X^c(u,v) \frac{\tilde{t}(u,v)}{a_{ir}^c} \right) \to 0$. Thus, an estimate for the transmission map is obtained by simply subtracting the dark channel of the normalized image from 1. As humans require some presence of fog to perceive depth, a scaling parameter can be introduced into Eq. (3.8). Removing the fog completely can lead to unnatural, flat looking images.

$$\tilde{t}(u,v) = 1 - w.\widetilde{\theta_M}(u,v) \tag{3.9}$$

Similar to the previous consideration for equation (2.18), equation (3.3) is simplified by assuming that there is only one color channel, i.e.,

$$\theta_M(u,v) = \underset{m,n \in \Omega(u,v)}{med} \left(\frac{\hat{X}(u,v)}{a_{ir}} \right)$$

$$= \underset{m,n \in \Omega(u,v)}{med} \left(\frac{X(u,v)t(u,v)}{a_{ir}} + 1 - t(u,v) \right) \tag{3.10}$$

We will get the same conclusions in (3.10) with assuming smooth radiance and smooth depth, as shown in (2.19) and (2.20), respectively, by replacing **min** operator with **med**.

However, let us take a look at how the MDCP method compares with the DCP method at occlusion boundaries. First, let vector $\boldsymbol{w_{u,v}}$ contain the sorted arrangement of pixels $\hat{X} \in \Omega(u,v)$ such that the lowest value is at $\boldsymbol{w_{u,v}(0)}$ and the highest value is at $\boldsymbol{w_{u,v}(N^2 - 1)}$ for window of size $\boldsymbol{N \times N}$, i.e.,

$$\boldsymbol{w_{u,v}} = \begin{bmatrix} min_{\ m,n \in \Omega(u,v)} \hat{X}(u,v) \\ \vdots \\ max_{\ m,n \in \Omega(u,v)} \hat{X}(u,v) \end{bmatrix} \tag{3.11}$$

We now have the nonsmoothed DCP and MDCP as:

$$\theta_M(u,v) = \left. w_{u,v}(0) \middle/ a_{ir} \right. \tag{3.12}$$

$$\theta_M(u,v) = \left. w_{u,v}(\tfrac{N^2-1}{2}) \middle/ a_{ir} \right. \tag{3.13}$$

for N odd.

Suppose, at an occlusion boundary, there are two different types of pixels $\hat{x}$ in patch $\Omega(u,v)$ that are either foreground pixels $\hat{x}_f$ or background pixels $\hat{x}_b$ in the sorted vector , $w_{u,v}$ i.e.,

$$w_{u,v} = \begin{bmatrix} min_{\ m,n\in\Omega(u,v)}\hat{x}_f(u,v) \\ \vdots \\ max_{\ m,n\in\Omega(u,v)}\hat{x}_f(u,v) \\ min_{\ m,n\in\Omega(u,v)}\hat{x}_b(u,v) \\ \vdots \\ max_{\ m,n\in\Omega(u,v)}\hat{x}_b(u,v) \end{bmatrix} \tag{3.14}$$

With w set to 0.95 for our experiments. The recovered dehazed image using the non-smoothed MDCP is:

$$X_m(u,v) = \frac{\hat{X}(u,v) - a_{ir}}{max(t(u,v),\varepsilon)} + a_{ir} \tag{3.15}$$

fog removal using , direct result from the dark channel prior leads to halo-artifacts or occulusion in the final image see f**ig 3.1**. As shown in **Fig. 3.2(a)**, the closer pixels are darker than hazier background pixels that are farther away. Notice that, comparing (3.12) with (3.13), even if there was one foreground pixel at the extreme edge of neighborhood $\Omega(u,v)$, the DCP will choose the foreground pixel value $\hat{x}_f$ at $\Omega(u,v)$, whereas the MDCP method will choose the background pixel $\hat{x}_b$. If no refinement is used, then this will cause *halo* artifacts at occlusion boundaries for the case of using the DCP. In **fig. 3.2** we also compare the *halo* effects when using the DCP, Tarel's method and MDCP. Notice the reduction of the *halo* effect with the MDCP method.

(a) Input fog image   (b) Estimated transmission map

(c) Recovered image
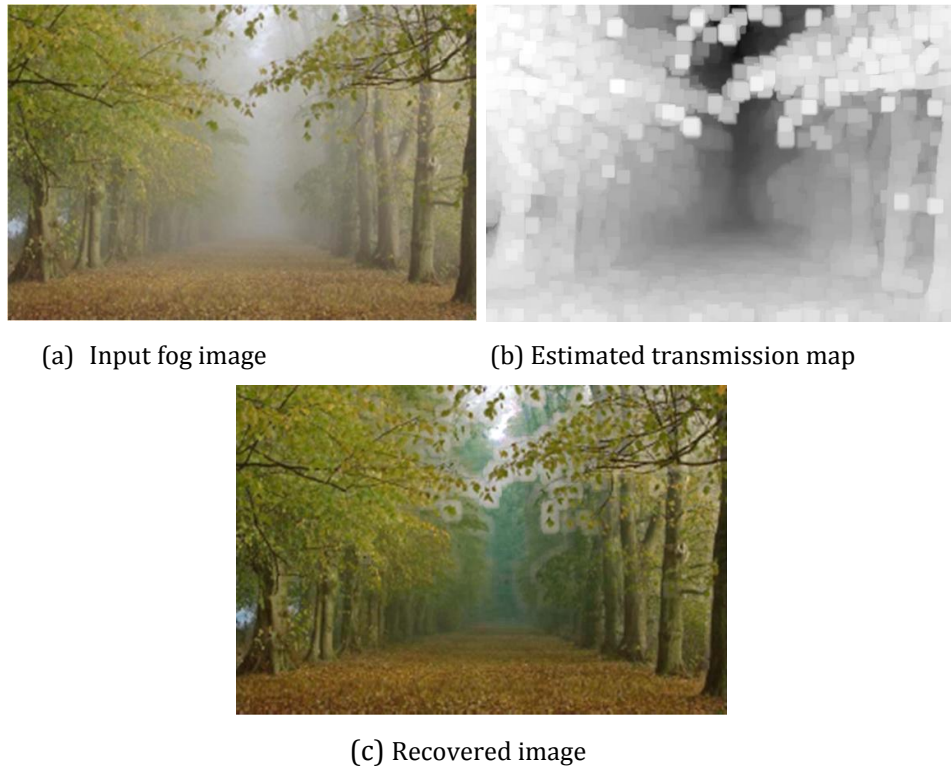
Figure 3.1: Recovering the scene radiance using the transmission map obtained directly from the Dark channel prior results in undesirable artifacts.

(a) Hazy image   **(b) θ_D**   **(c) θ_M**

(d) θ_T
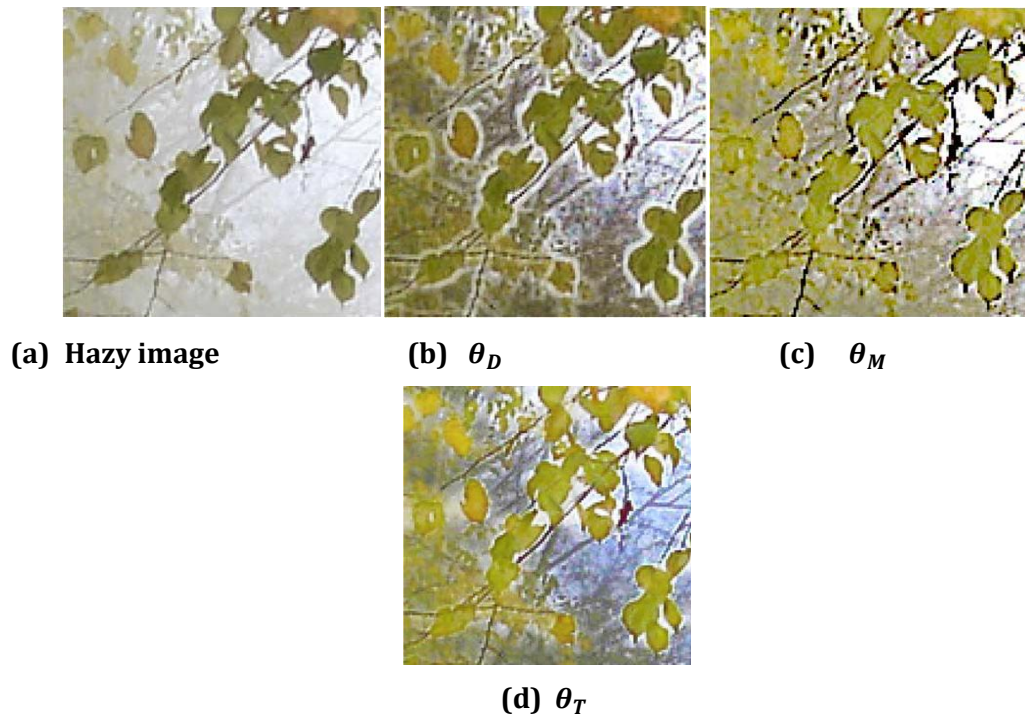
Figure 3.2 : Examples of occlusion boundaries. (a) Original image. (b) Dehazed using unrefined DCP method (c) Dehazed using our proposed MDCP method (d) Dehazed using method from Tarel. Notice in (b) and (d) that there are *halo* artifacts around the leaf edges. The *halo* artifacts are not as dominant in (c).

3.4 Flowchart of proposed method

Input image

Median Dark channel

MDCP for each patch

Estimate airlight

Create Transmission map

Transmission map

Guide Image

Guided filter

Recovered image

Refined transmission map

Recover scene radiance
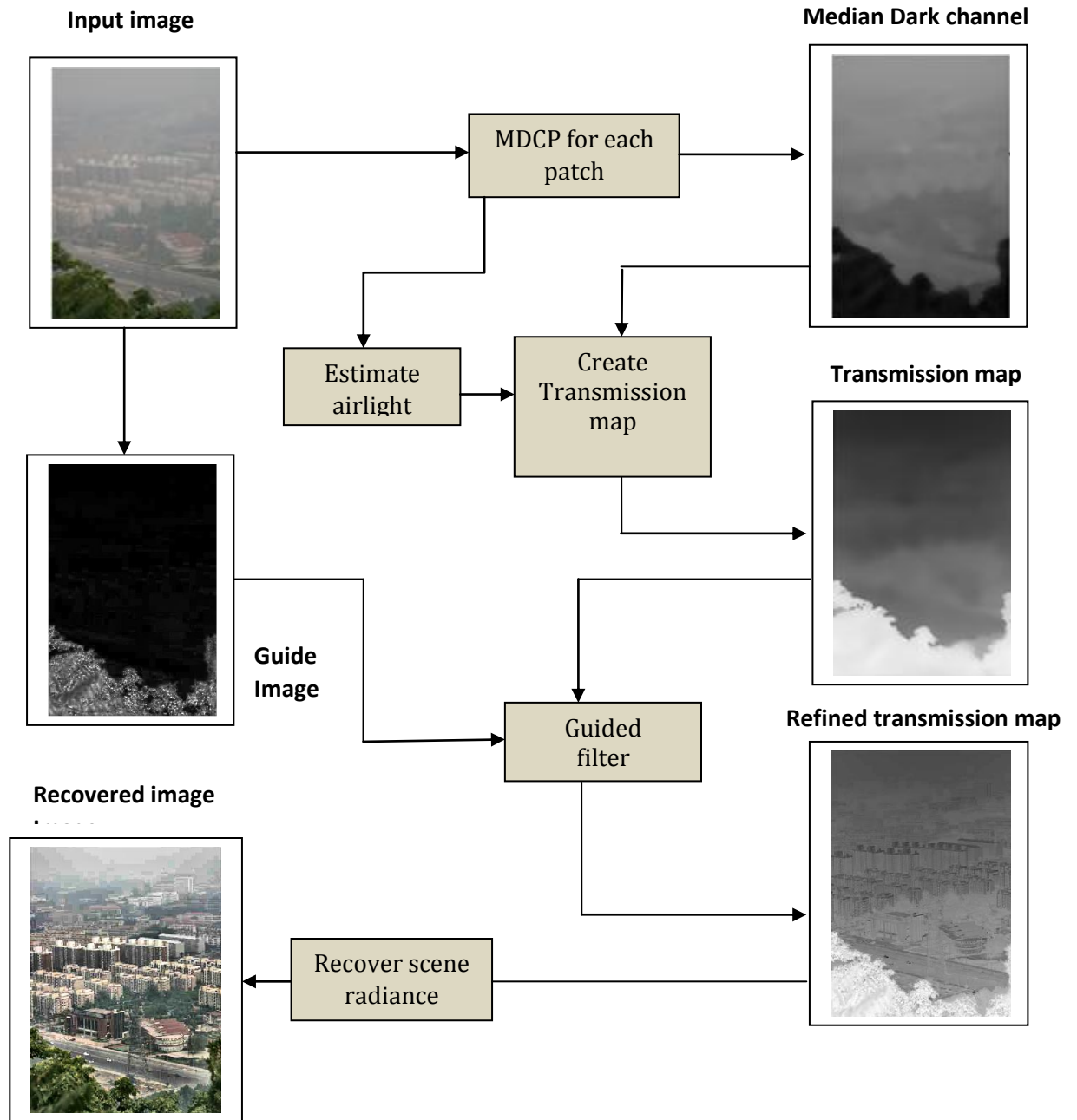
Fig 3.3 Process flowchart of proposed dehazing method. The patchsize in this example is set to 15 x 15.

Chapter 4

_____

Transmission Map Refinement

once the transmission map, $\tilde{t}$, and atmospheric light, a_{ir} , are known, the scene radiance can be solved for directly using Eq. (2.6), which is repeated here for convenience:

$$X(u, v) = \frac{\hat{X}(u,v) - a_{ir}(1 - t(u,v))}{t(u,v)}$$

$$= \frac{\hat{X}(u,v) - a_{ir}}{t(u,v)} + a_{ir} \tag{4.1}$$

where (u, v) is a pixel location, $\hat{X}$ is the input image, X is the recovered image, a_{ir} is the atmospheric light, and t is the transmission. When using Eq. (4.1), t is typically lower bounded to a small number, such as 0.1, to avoid instability. Chapter 3 provided a simple method for estimating the transmission and atmospheric light. However, if the image $X(u, v)$ is recovered directly from this transmission estimate, the result contains block artifacts and halos around depth discontinuities, due to the nature of the estimation method (see **Fig. 3.1 and Fig.3.2**).

In order to remove these artifacts, it is necessary to further refine the transmission map. This chapter first provides a review of two existing methods that are used to refine the transmission map. The first is a result from Levin et al. [10] originally intended for image matting that was adopted by He et al.**[8]** for refinement of transmission-map. The second is a much faster method called the "Guided Image Filter" [9]. Although in [9], the Guided Filter was briefly mentioned as applicable to transmission map refinement, here it is examined more closely for this application, with a direct comparison to the results using the first approach.

4.1 Matting Laplacian

Interactive image matting is a problem in which one must separate a foreground object from a given image based on limited user input. For example, one may wish to separate a person from the surrounding scenery. User input generally involves the creation of a coarse image mask, called a trimap [10], where white pixels definitely belong to the foreground object, black pixels definitely belong to the background object, and gray pixels can belong to either. The job of an image matting algorithm is then to refine this coarse image mask, resolving whether the gray pixels belong to the foreground or background. Formally, the image matting problem can be described by the following equation:

$$X(u, v) = F(u, v)\alpha(u, v) + B(u, v)(1 - \alpha(u, v)) \tag{4.2}$$

where (u, v) is a pixel location, X is the input image, **F** is the foreground image, **B** is the background image, and α is called the alpha matte, describing the opacity of the fore ground. Looking closely, this is exactly the fog formation equation first presented in Chapter 1, repeated here for convenience:

$$\hat{X}(u, v) = X(u, v)t(u, v) + a_{ir}(1 - t(u, v)) \tag{4.3}$$

where **F** corresponds to the underlying scene image, **X**, **B** corresponds to the atmospheric light, a_{ir}, and α corresponds to the transmission, t. Furthermore, the coarse transmission map estimated in Chapter 2 can be interpreted as the user-input trimap in interactive image matting.

In [10], Levin et al. proposed a state-of-the-art, closed-form solution to this problem. Their solution is based on the assumption that within a small window, pixels belonging to a single object are a linear mixture of two colors. From this assumption, Levin et. al derive an affinity matrix, L, known as the "Matting Laplacian."

The (i, j) element of L is defined as:

$$L(i,j) = \sum_{(k|i,j)\in w_k} \left(\delta_{ij} - \frac{1}{|w_k|}\right)\left(1 + \left(X_i - \mu_k\right)^T \left(\sum k + \frac{\varepsilon}{|w_k|}U_3\right)^{-1}\left(X_j - \mu_k\right)\right) \tag{4.4}$$

Where, $\boldsymbol{\delta_{ij}}$ is the Kronecker delta, $\boldsymbol{\mu_k}$ and $\sum k$ are the mean and covariance of the colors in window ($\boldsymbol{w_k}$)centered around k, $|\boldsymbol{w_k}|$ is the number of pixels in each window, and $\boldsymbol{U_3}$ is a 3x3 identity matrix. $\boldsymbol{\varepsilon}$ is a small regularization parameter. The summation is for all windows shared by pixels $\boldsymbol{i}$ and $\boldsymbol{j}$ in the image $\mathbf{X}$. Note that for an image of size MxN, $\boldsymbol{L}$ is a symmetric matrix with size MN x MN, and that i and j refer both to a location in L and to the ith and jth pixels in the vectorized image, $\mathbf{X}$. Since window sizes that are too large may violate the color line model, the typical size used, and the size used for all examples shown here, is 3x3.

4.1.1 Transmission Map Refinement using Matting Laplacian

Given the coarse transmission map estimate, the Matting Laplacian is used to find a refined transmission map by minimizing the following quadratic cost function:

$$\text{E}(\mathbf{t}) = \mathbf{t}^\text{T}L\mathbf{t} + \lambda(\mathbf{t} - \tilde{\mathbf{t}})^\text{T}(\mathbf{t} - \tilde{\mathbf{t}}) \tag{4.5}$$

where $\mathbf{t}$ is the vector form of t, $\tilde{\mathbf{t}}$ is the vector form of the coarse transmission estimate, and λ is a regularization parameter. The solution minimizing Eq. (4.5) is found by solving for $\mathbf{t}$ from the following:

$$(L + \lambda U)t = \lambda\tilde{\mathbf{t}} \tag{4.6}$$

where U is an identity matrix with the same size as L, and λ is a small value (10^{-3} to 10^{-4}) so that $\mathbf{t}$ is softly constrained by $\tilde{\mathbf{t}}$. Although L is dimensionally large (MNxMN), due to the window constraint on i and j, it is sparse. This method seems very elegant and shows very good results, as one can see from **Fig.4.2(c).** In Fig. 4.1 we shows the results of applying the Matting Laplacian to the results from Fig. **3.1**.

Summary

Since the computational cost in soft mating is quite high, we can't used it for refinement of large image and it is suitable for small images only.

(a) Estimated transmission map from DCP

(b) Refined transmission map using Soft-mating

(c) Recovered image from refined transmission map

Figure 4.1.: Refining the DCP transmission-map using the Matting Laplacian results in a much smoother recovered scene radiance. For this example, $\lambda = 10^{-3}$, and $\varepsilon = 10^{-4}$

4.2 Guided Filter

The Matting Laplacian which was used for refinement of transmission map in **[8]** produces visually satisfying result, but the computational cost is very high as it involves solving a large linear system, by inverting a very large matrix. Therefore to speed up the defogging process, the transmission map obtained in chapter-2 is refined by using guided filter [9]. The guided filter is a fast, accurate, non-iterative, edge-preserving smoothing operator, which filter the input image under the guidance of another image. Denote the input image as **p** the guidance image as **I** and the filtering output as **q**. The local linear model of guided filter assumes that **q** is a linear transform of the guidance **I** in a window w_k centered at pixel k, so that mathematically we have:

$$q_i = a_k I_i + b_k, \forall_i \in w_k \tag{4.7}$$

Where, q_i are the filter output, $(\boldsymbol{a_k b_k})$ are some linear cofficients assumed to be constant in a window w_k. We also denote the radius of the window w_k as r, which is the pixel distance from the centre pixel to the outer pixel. Since square window are used ,the total window size is therefor$(2r + 1) * (2r + 1)$.

The guided filter seeks for coefficient(a_k, b_k) that minimize the difference between the output **(q)** and input **(p)**, by using following cost function.

$$E(a_k, b_k) = \sum_{i \in w_k}((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2) \tag{4.8}$$

Where, ε is regularization parameter to prevent a_k from being too large. The solution of equation (4.8) is found to be

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \varepsilon} \tag{4.9}$$

$$b_k = \overline{p_k} - a_k \mu_k \tag{4.10}$$

Here, μ_k and σ_k^2 are the mean and variance of **I** in a window w_k ,and $|w|$ is the no. of pixel in w_k. $\overline{p_k}$ is the mean of **p** in w_k. $\sum_k$ *is* the 3x3 covaiance matrix of **I** in w_k . So after computing all filter coefficients (a_k, b_k), in the image, the final output is :-

$$q_i = \frac{1}{|w|}\sum_{k:(i)\in w_k}(a_k I_i + b_k) = \bar{a}_\iota I_i + \bar{b}_\iota \tag{4.11}$$

Where $\bar{a}_\iota$ and $\bar{b}_\iota$ are obtained with average filter: $\bar{a}_\iota = \frac{1}{|w|}\sum_{k\in w_i} a_k$, $\bar{b}_\iota = \frac{1}{|w|}\sum_{k\in w_i} b_k$. Since $\bar{a}_\iota$ and $\bar{b}_\iota$ are outputs of an average filter, their gradients are quite small, so $\nabla p \approx \bar{a}\,\nabla I$. The Guided Filter is also extended to color guidance images by rewriting Eq. (4.11) as:

$$q_i = a_k^T I_i + b_k, \qquad \forall_i \in w_k \tag{4.12}$$

The guided filter than become

$$a_k = (\textstyle\sum_k + \varepsilon U)^{-1}\left(\frac{1}{|w|}\sum_{i\in w_k} I_i p_i - \mu_k \overline{p_k}\right) \tag{4.13a}$$

$$b_k = \overline{p_k} - a_k^T \mu_k \tag{4.13b}$$

$$q_i = \bar{a}_i^t I_i + \bar{b}_\iota \tag{4.13c}$$

4.2.1 Transmission-map refinement using Guided filter

For guided filter, the guidance image I, the input p and the filtering output q play similar roles as the input image, the trimap and the alpha matte in the closed-form matting framework [9]. An illustration of guided filter is shown in **Fig.4.2**.

transmission map

Filtering output **t**
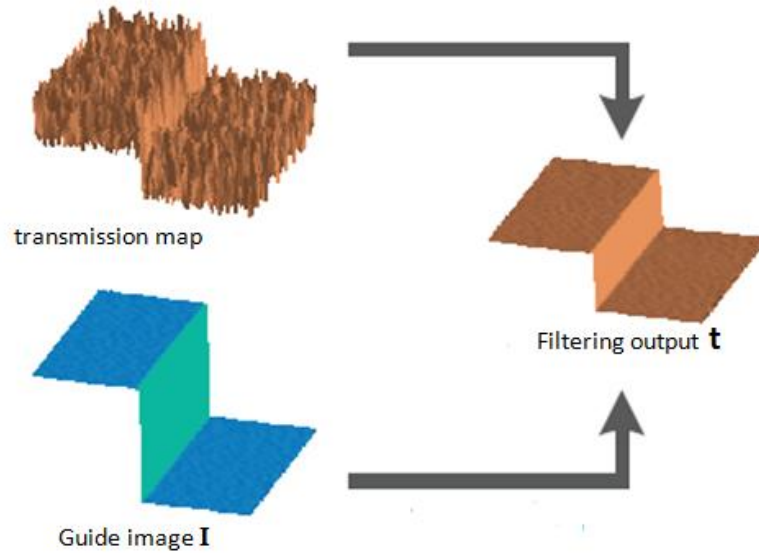
Guide image **I**

Fig. 4.2 Illustration of guided filter. After filtering the edges of the guide image are transferred to the filtering image

4.3 Algorithm:- Guided Filter

Input: transmission map $\tilde{t}$,guidance image I ,regularization parameter ε.

Output: filtered transmission map t.

Step 1.
$$mean_I = f_{mean}(I)$$
$$mean_{\tilde{t}} = f_{mean}(\tilde{t})$$
$$corr_I = f_{mean}(I.*I)$$
$$corr_{I\tilde{t}} = f_{mean}(I.*\tilde{t})$$

Step 2.
$$var_I = corr_I - mean_I.*mean_I$$
$$cov_{I\tilde{t}} = corr_{I\tilde{t}} - mean_I.*mean_{\tilde{t}}$$

Step 3.
$$a = cov_{I\tilde{t}}./var_I + \varepsilon$$
$$b = mean_{\tilde{t}} - a.*mean_I$$

Step 4. $\qquad\qquad\qquad\qquad t = f_{mean}(a).*I + f_{mean}(b)$

The results of transmission refinement using guided filter is shown in Fig. 4.3.

4.4 Summary

- ❖ Guided filter reduce the problem of solving a large linear system of equation to a simple filtering process that can be computed in O(n) time. This property is important for single image dehazing because it provides a chance to improve the efficiency of transmission refinement in [8].
- ❖ Guided filter has the ability that it can transfer the structure of guidance image to the filtering output so that it has good behaviour near edges.
- ❖ complexity of our method is quite low, and it is capable of processing the large image with short running time.

(a) Estimated transmiission map from MDCP

(b) refined transmission map using Guided filter

(c)Recovered image from refined transmission map

Figure 4.3: Refining the MDCP transmission-map using the Guided Filter($\mathcal{E}=10^{-4}$, r=15)

Chapter 5

Post processing & Performance Evaluation matrices

5.1 Introduction

It is found that image after removal of fog loose some contrast and appear dim. Hence there is a requirement of some contrast-enhancement technique which can enhance the visibility. The most common method for visibility enhancement are histogram equalization and histogram stretching. Histogram equalization is the most popular method and it is used to enhance the contrast of the image by gray-scale transformation but,its major disadvantage is that it over-enhances the image and shift its mean brightness and consequently it creates an unnatural look. While in histogram stretching we have to be careful of clipping otherwise it eliminate visual information in very bright and in very dark region. Therefor instead of using HE and histogtram stretching which affect the whole image CLAHE is used to enhance the contrast of low-contrast regions.

5.2 Contrst Limited Adaptive Histogram Equalisation

Contrast limited adaptive histogram equaization is a visibility enhancement algorithm which can provide optimal equalization and enhance the local contrast of the image. It can overcome the problem of standard histogram equalization technique. As proposed in [15] CLAHE algorithm divide the images into small blocks [8x8], and apply histrogram equalization to each block for contrast-enhancement and to combine the neighboring block in an image bilinear interpolation is used which eliminate the artificially induced boundaries.There are two parameter which are used in CLAHE to control image quality.These are block-size and clip-limit. Block-size specify the size of contextual-region and clip-limit is a scaler parameter in range [0 1] specify the contrast enhancment limit and prevents over-saturation. A representation of CLAHE algorithm is shown in **fig.(5.1)**

The CLAHE method can be derived into following steps :

$$N_{avg} = \frac{N_x \ X \ N_y}{N_{Gray}}$$

(5.1)

Where , N_{avg} = average no. of pixel.

N_x = No. of pixel in x-direction of the contextual-region.

N_y = No. of pixel in y-direction of the contextual-region.

N_{Gray} = No. of gray level in the contextual-region.

Based on equation (4.1) N_{CL} can be calculated by:

$$N_{CL} = N_{clip} \ X \ N_{avg}$$

(5.2)

Where, N_{CL} is the actual clip-limit. N_{clip} is the maximum multiple of average pixels in each gray level of the contextual-region.

(a) Original Histogram   (b) Clipped Histogram
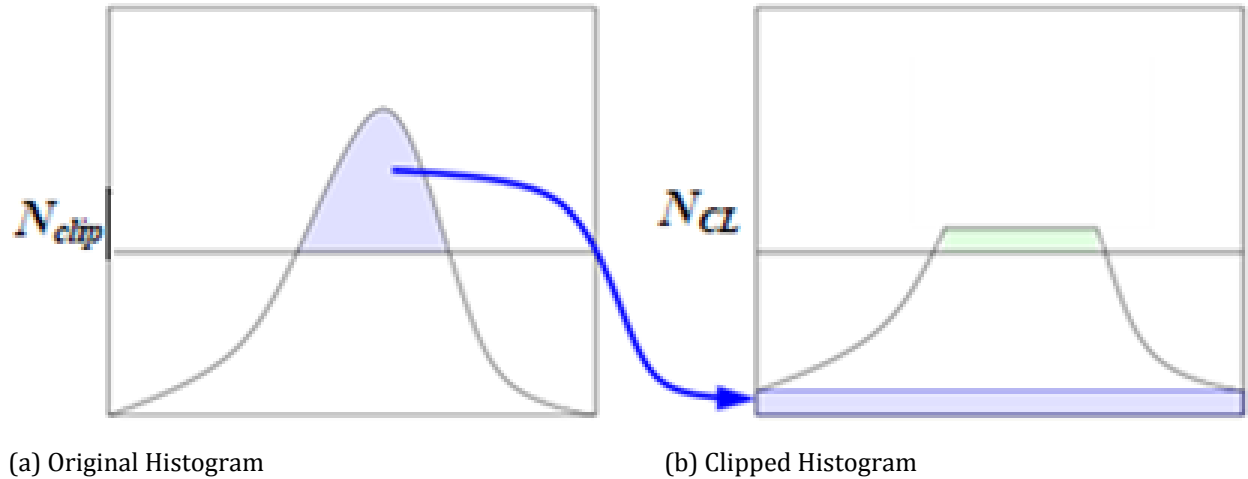
Fig. 5.1 Representation of CLAHE

From **Fig. 5.1**, we can observed when number of pixels is greater than N_{clip}, the pixels will be clipped. We could defines the total number of clipped pixels as $N_{\Sigma \, clip}$, then the number of pixels distributed averagely into each gray level is given by:-

$$N_{ap} = \frac{N_{\Sigma \, clip}}{N_{gray}}$$

(5.3)

Rules for calcalating the CLAHE of the contextual region :

Rule 1. $if \quad N_x(i) > N_{CL} \quad, \quad N_x(i) = N_{CL}$

Rule 2. $else \ if \quad N_x(i) + N_{acp} \geq N_{CL} \quad, N_x(i) = N_{CL}$

Rule 3. $else \quad N_x(i) = \ N_x(i) + N_{ap}$

After the distributed pixel are given by:

$$S = \frac{N_{Gray}}{N_{LP}} \tag{5.4}$$

Where, N_{LP} are the remaining number of clipped pixel after following the above distribution.

5.3 Performance Evaluation Matrices

Evaluation of images, after processing, is an important step for determining how well the images are being restored. Quality of image is usually assessed using image quality metrics. Many metrices has been proposed over the past few decades for measuring the visual quality of enhanced image. Each one has its advantage and disadvantage in terms of accuracy,speed and application considered. Infact there is no universal measure, which specify both the subjective and objective validity of enhancement for all type of images. In this paper performance of our proposed method is measured in terms of Absolute mean brightness error**(AMBE)**, visibility metric**(VM)**, PSNR and run time **(t_{comp})**.

5.3.1 Absolute mean brightness error (AMBE)

A qualitative measure of visibility enhancement is checked in terms of AMBE. To examine how the apearance of image has changed after removal of fog, the deviation of the mean intensity of the enhanced image from the mean intensity of the reference image is computed.

$$AMBE = \left| \mu_r - \mu_e \right| \tag{5.5}$$

AMBE should be as small as possible for better similarity between two images.

5.3.2 Visibility metric (VM)

The visuality enhancement performance of proposed method are defined by visibilty metric. Visibility is a measure of image quality and used to tell how well an observer can view a texture and color of image.Visibility metric gives an objective measure of detail enhancement. It is useful for measuring the improvement in contrast as well as sharpness.The visibility metric calculate the contrast to noise ratio (CNR). CNR is similar to SNR but subtract off a term such as fog before taking a ratio.

$$\text{VM}=CNR(r,e) = \frac{\mu_r - \mu_n}{\sigma_n} \qquad (5.6)$$

Where,

$$\mu_r = \frac{1}{M \text{ x } N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} r(i,j) \qquad (5.7)$$

$$\mu_n = \frac{1}{M \text{ x } N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} n(i,j) \qquad (5.8)$$

$$n(i,j) = r(i,j) - e(i,j) \qquad (5.9)$$

$$\sigma_n^2 = \frac{1}{(M \text{ x } N)-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (n(i,j) - \mu_n)^2 \qquad (5.10)$$

Here, r and e denote the reffernece and enhanced image. The effectiveness are described by high value of visibility metric.

5.3.3 Peak signal to noise ratio

There are many version of signal–to-noise ratio,but the PSNR is simple and widely used parameter for fidelity measurement. PSNR is calculated in decibals units, which measure the ratio of the peak signal and the mean-square-error of the two images. PSNR is defined by:

$$\text{PSNR} = 10 log_{10} \left(\frac{(L-1)^2}{MSE}\right) \text{db} \qquad (5.11)$$

Where,

$$\text{MSE} = \frac{1}{M \text{ x } N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (r(i,j) - e(i,j))^2 \qquad (5.12)$$

Where r and e denote the reffernece and enhanced image respectively . M x N represent the size of image and L is dymanic range of pixel value(256 for 8-bit grayscale image). A high value of PSNR yield good result.

5.3.4 Run time (t_{run})

For a 600 x 400 color image by using guided filter run time is decreased by a factor of 10. For the run time comparison ,we have chosen three different algorithm from the literature. A comparison of run time (t_{run}) is summarized in table 1.

Chapter 6

_____

Result and Evaluation

This section presents an assesment of the proposed method. The proposed method is implemented in MATLAB R2011b, 64-bit Intel Core-i3 processor,RAM-3GB. To compare the performance of our method, images used in our bechmarking are derived from several sources such as the 'google database' or private collection.

6.1 Analyses and Parameter Settings

An important parameter for computing the dark channel is patch size, we denote its radius as r^{dark}. Since the computational cost in soft mating is quite high, we can't used it for refinement of large image and it is suitable for small images only. So in [8] the patch size is fixed as 15x15, which is relatively small. However the complexity of our method is quite low, and it is capable of processing the large image with short running time. So it is not appropriate to use fixed patch size in our case. Therefor, we adapt a simple stratgey to adjust r^{dark} according to area of image . When the no. of pixel in the image exceed $5x10^5$ the radius is fixed as r^{dark}= 30, preventing the patch size from growing too large; When the no. of pixel in the image is less than $2x10^5$, the patch radius is fixed as r^{dark}=7, preventing the patch size become too small. For images with pixel numbers in between, we interpolate the patch radius linearly. Notice that there exist a tradeoff for the patch size. When it decreases, the blocky artifacts in the coarse transmission map reduce, making it easier for the guided filter to refine the transmission; however, a smaller patch size makes the **dark channel prior** less appropriate, the transmission t(x) is underestimated and the recovered image still have poor visibility. However, a large radius r implies that the filtering output is linear to the guidance image in a large range, which helps to reduce the block artifacts in the transmission map, so that the halo effect in the recovered image can be eliminated. However, if r is too large, the transmission map will capture too much details from the guidance (the input hazy image), making the recovered image over-saturated. Therefore the

patch size should neither be too large nor too small. For the guided filter, our experiments find that the refined transmission map is not sensitive to the regularizer ϵ, and we fix it to be 10^{-3}. But for the filtering radius r, its behavior is quite different. **Fig. 6.1** shows an image (with size of 896×672) recovered with different radii r. In this example, the patch size to compute the dark channel is deliberately set as 15 × 15. As can be seen, $r = 8$ (corresponds to the window of 17×17, which is of similar scale as the patch size 15×15), in **Fig.6.1(d)**;the halo effect is quite severe , when $r = 800$, the halos are suppressed **Fig.6.1(f)**, but notice the leaves in the image become saturated; when $r = 80$, there're still some halos **(Fig. 6.1(e))**, but it compromises between the two extremes.

6.2 Qualitative analysis

Qualitative analysis of our work are compared with recent state-of–the-art methods in Fig. 6.2 and in Fig. 6.3. Experimental result show that proposed algorithm produces high quality fog free images without scarificing the fidelity of the colors. **Figure 6.4.** Defog result of our work and comparison with He et al.'s work.

6.3 Quantative Analysis

Quantitative perforance of proposed method are measured in terms of Absolute mean brightness error (AMBE), visibility metric(VM), PSNR while run time (t_{run}) is also used to compare the efficiency of our method. Compared to the work in [8], the main advantage of combining the median dark channel prior and guided filter to defog the images lies in its low computational cost. Our Matlab implementation takes about 4 seconds to process a 1-mega pixel image, but in [8], it takes about 10-20 seconds to process the same image. By virtue of its exact O(N) time complexity, the running time of our algorithm becomes tolerable for many applications. To check the effetiveness of our proposed method, we done the visibility metric comparison with other state-of-art methods in Table-I. In Table-I, 2nd column represent the actual visibility of fog images, wheres the restored visibilty are depicted in last column of table II. We observed that actual visibility of image 'mountain-01' is **70.59** which is increased to **115.62** after fog removal. Similarly in Table II and in Table III we compar e

the Peak signal to noise ratio and absolute mean brightness error with other state-of-art methods.

Fig. 6.1. Scene radiance recovery using different filtering radii *r* (patch size of dark channel is 15x15, $\epsilon = 10^{-3}$). (a) Input hazy image. (b) Dark channel. (c) Coarse transmission map. (d)(e)(f) Scene radiance recovered with *r* equals to 8, 80, and 800, respectively.

Fig.6.2 Defog result comparison with state-of-the-art methods.
(a) image 'mountain 01'.
(b) Kopf et al.
(c) Fattal's
(d) Tan et al.
(e) He et al.
(f) Tarel et al.
(g) our work

Fig.6.3 Defog result comparison with recent state-of-the-art methods. Source[He et al]
a) original foggy image 'tower 02'.
b) Kopf et al.
c) Fattal's
d) Tan et al.
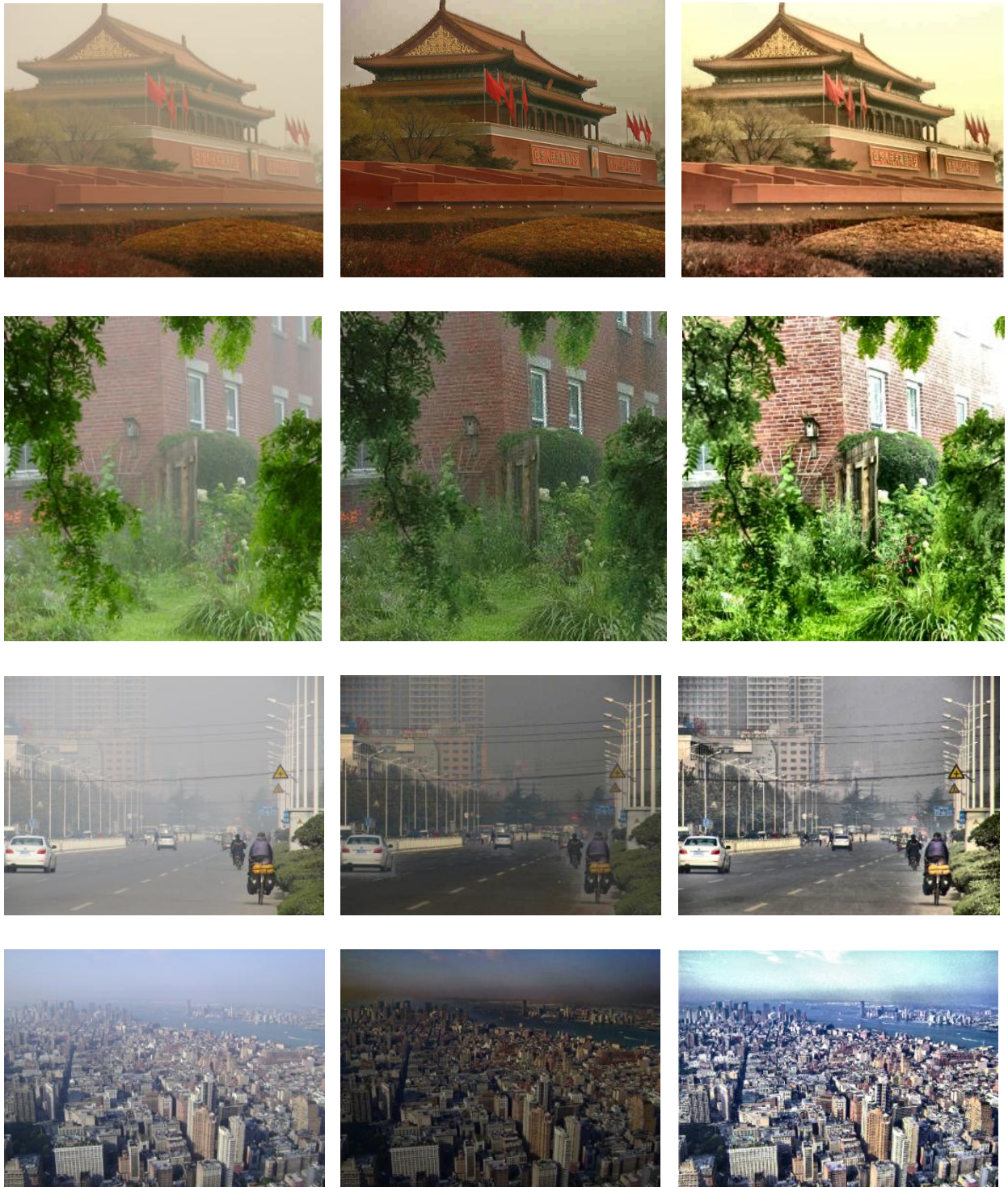e) He et al.
f) Tarel et al.
g) our work

(a) Input images   (b) He's Result   (c) our Result

Fig. 6.4 Defog result comparison for some popular fog images with He et al. work

Table-I

Visibility metric comparison with various state-of–art method

Image	VM	Kopf	Fattal	Tan	He	Our
Mountain'01'	70.59	80.14	78.18	76.62	87.95	115.62
Tower '02'	65.92	99.61	99.26	97.48	103.41	150.63
Sweden	62.99	100.87	85.21	114.15	113.24	115.98

Table II

PSNR comparison with various state-of–art method

Image	Kopf	Fattal	Tan	He	Our
Mountain'01'	34.83	45.74	30.07	32.16	48.40
Tower '02'	31.47	36.82	28.62	33.87	42.83
Sweden	22.67	22.71	22.50	24.60	28.74

Table III

AMBE comparison of various state-of–art method

Image	Kopf	Fattal	Tan	He	Our
Mountain'01'	22.61	17.86	47.97	48.27	10.92
Tower '02'	16.54	34.46	47.80	34.88	5.51
Sweden	45.00	76.38	75.86	71.06	43.58

Table IV
run time (seconds) comparision with He et al. work

He	Our
56.32	15.71

Chapter 7

Conclusion

In this paper ,we presented a new fast algorithm to improve the visibility of fog degraded images. This new approach is physically sound and restore the visibility upto maximum extent. The proposed algorithm uses a combination of Guided filter and CLAHE. Guided filter is a edge-preserving filter which is used for quickly refinement of transmission map and CLAHE is used to improve the local contrast of image by partioning the image into small boxes. It is observed that in restored image, no oversaturated region exist due to which there is a negligible apearance of halo-artifact. The advantage of proposed method are multiple.

1. It is simple in principle and hence easy to implement.
2. It provide good result in most cases(homogenous fog), without introducing artifact.
3. It is efficient for various type of fog images.

Finally we have experimentally compared the result of proposed algorithm against competing techniqes, by showing that our method achieve state-of-the-art performance by producing   high quality fog free images and it is expected to be used in real-time outdoor vision system owing to its fast exucution speed. In   conclusion, with respect to above mentioned quality criteria , our method out-perform the all other esisting method by enhance the detail in fog degraded images. However, guided image filtering is actually an approximation of soft mating; this method fail when the input image contain abrupt depth changes. Failure of transmission refinement using guided filter is shown in **Fig.7.1.** Firstly, the dehazed result Fig. 7.1(b) contains noticeable halos, this happens because the depth change at the object edge is too abrupt, the guided filter needs a larger filtering radius r to suppress the halos. As a result, when the input hazy image contains discontinuities that are too abrupt, it is difficult to find an appropriate filtering radius that compromises between the halo effect and over-saturation.

(a) Input fog image

(b) Dehazing result

Figure 7.1 Failure of transmission refinement using guided filter.

Therefore the low complexity of this dehazing algorithm comes with the price of some failures. To address the mentioned problem, future improvements may focus on it.

References

1. S.G. Narasimhan and S.K. Nayar, "Vision and the Atmosphere," Int'l J. Computer Vision, vol. 48, pp. 233-254, 2002.

2. S.G. Narasimhan and S.K. Nayar, "Chromatic Framework for Vision in Bad Weather," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 598-605, June 2000.

3. Y.Y. Schechner, S.G. Narasimhan, and S.K. Nayar, "Instant Dehazing of Images Using Polarization," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 325-332, 2001.

4. R. Fattal, "Single Image Dehazing," Proc. ACM SIGGRAPH '08, 2008.

5. R. Tan, "Visibility in Bad Weather from a Single Image," Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2008.

6. J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep Photo: Model-Based Photograph Enhancement and Viewing," ACM Trans. Graphics, vol. 27, no. 5, pp. 116:1-116:10, 2008.

7. J.P. Tarel, and N. Hautiere, "Fast visibility restoration from a single color or gray level image", IEEE International Conference on Computer Vision, pp. 2201-8, 2009.

8. K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior", IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1956-63, 2009.

9. K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. Europ. Conf. on Comp. Vis.*, Sep. 2010.

10. A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," in Proc. IEEE Conf. on Comp. Vis. Pat. Rec., June 2006.

11. Tripathi, Abhishek Kumar, and Sudipta Mukhopadhyay. "Single image fog removal using bilateral filter", 2012 IEEE International Conference on Signal Processing Computing and Control, 2012

12. Tripathi, Abhishek Kumar, and Sudipta Mukhopadhyay. "Single image fog removal using anistropic diffusion" IET image process.

13. R.C. Gonzalez, and R.E. Woods, "Digital Image Processing", Addison-Wesley, Reading, Mass., 1992.

14 S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie,A. Geselowitz, T. Geer, B. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," Computer Vision, Graphics, and Image Processing, 39(3): pp. 355-368, September 1987.

15. E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, S. M. Pizer (1998), "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," J Digit Imaging, Vol. 11, pp. 193-2000.