# Digital Image Processing
# Project 3

——Single Image Haze Removal Using Dark Channel Prior

Name: sungx

Data: 2019-12-03

# 1. Problem

- Implement the paper "Single Image Haze Removal Using Dark Channel Prior", and apply your haze removal code to 'fog1.jpg' & 'fog2.jpg'. Implement the guide filter version instead of the matting step. (**You should display the original image, dark channel, dark channel after guide filter, processed image in a figure, and save the figure as 'result1.xxx' & 'result1.xxx'**)

- Discuss the parameters influence and limitation of dark channel combining the results in the slides and report.

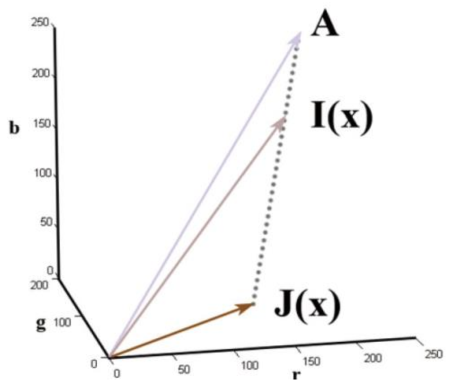
# 2. Problem Analysis

### 2.1 Background

In computer vision and computer graphics, the model widely used to describe the formulation of a haze image is as follows:

$$I(x) = J(x)t(x) + A(1 - t(x))$$

where **I** is the observed intensity, **J** is the scene radiance, **A** is the global atmospheric light, and **t** is the medium transmission describing the portion of the light that is not scattered and reaches the camera. The goal of haze removal is to recover the **J**, **A**, and **t** from **I**.




2.2.1 Image Degradation

The first degradation is the visibility reduction due to the direct attenuation. Visibility is a measure of how well the object can be discerned. In the computer vision visibility is often described by the gradient of the image. From the formula above, we have:

$$\nabla I(x) = t(\mathrm{x}) \nabla J(x)$$

where we consider t as uniform so the gradient of t is ignored. Because t is in the range [0,1], $\nabla I$ has a smaller magnitude than $\nabla J$. So the visibility is reduced and the objects are more difficult to discern. We can see that visibility reduction is due to the direct attenuation $J(x)t(x)$, or say, due to t.

The second degradation is chrominance shift due to the airlight. The chrominance describes the colorfulness regardless of the luminance, which is represented by the direction of the color vector in the RGB color space. The haze imaging equation suggests the vector in the RGB color space. The haze imaging equation suggests the vector I is a linear combination of the two vectors J and A. Due to the additive A, the vectors I and J are not in the same direction: the chrominance is shifted. Usually the atmosphere is white or gray, so a hazy image appears grayish and less vivid. See Fig below.

In sum, the multiplicative attenuation reduces the visibility, and the additive airlight changes the chrominance. Haze is troublesome in many computer vision/graphics applications. The reduced visibility impacts objection detection and recognition, lowers the reliability of outdoor surveillance systems, and obscures the satellite images. In consumer-level photography, haze changes the colors and reduces the contrast of the photos. The degradation cannot be avoided by higher level cameras or better lens, because it happens in the atmosphere before reaching the apparatus. Therefore, removing the haze effects from images is demanded in computer vision/graphics.

2.2.2 Depth

Suppose that a scene point in the position x has the distance in the position x has the distance d(x) from the observer. d is called the depth of the point. It is found that the haze transmission t is physically related to the depth d:

$$t(x) = \exp(-\int_0^{d(x)} \beta(z)dz)$$

Here, $\beta$ is the scattering coefficient of the atmosphere. $\beta$ is determined by the physical properties of the atmosphere, like particle material, size, shape, and concentration. The integral is on a line between a scene point and the observer.

If the physical properties of the atmosphere are homogenous, the scattering coefficient $\beta$ is a spatial constant. Thus, we can rewrite t as:

$$\mathrm{t(x)} = \exp(-\beta d(\mathrm{x}))$$

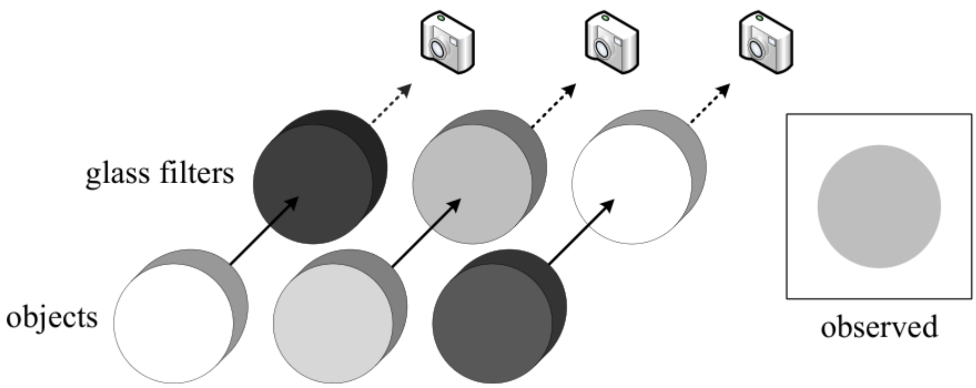
On the other hand, if we can estimate the transmission t, we can calculate the depth up to an unknown scale.

### 2.2.3 Problem Formulation and Ambiguity

Denote the number of pixels in an image as N. If the input I is an RGB color image, we have a set of 3N equation:

$$I_c(x) = J_c(x)t(x) + A_c(1-t(x))$$

where the scalars $I_c$, $J_c$ and $A_c$ are the color components in the channel c. However, we have 3N unknown reflections $J_c$, N unknown transmission t, and 3 unknown atmospheric light $A_c$. The total number of unknowns are 4N+3, much greater than the number of 3N of equations. In computer vision, we refer to this problem as ambiguous, ill-posed, or under-constrained. The ambiguity is mainly due to the spatially variant t, which contributes N variables. So we require at least one extra constraint for each pixel to solve the ambiguity.




The physical meaning of the ambiguity can be understood in the following way. The haze plays a role like a semi-transparent glass filter. The color of the filter is A, and the transparency is t. An object with a color J is seen through the filter. Objects in different colors can be observed as the same, if they are seen through filters in different colors with proper transparency. So given the observed color, how can we know the object color? We are faced with a similar question in haze removal.

**2.2 Dark Channel Prior**

Let us consider an image taken in a clear day in which no haze exists. Human beings are able to tell whether it is a haze-free image, even there is no interaction between the scene radiance and the haze at all. This motivates the author to find a prior, which concerns the scene radiance J (an haze-free image) alone. He propose the dark channel prior which is solely about haze-free image.

2.2.1 Observation

Observation is as following:

*For outdoor haze-free images, in most patches that do not cover the sky, there exist some pixels whose intensity is very low and close to zero in at least one color channel.*

He refers to the pixels "whose intensity is very slow and close to zeros in at least one color channel" as dark pixels. To understand this observation, below explain what factors contribute to the dark pixels.

First, the dark pixels can come from the shadows in the image. Outdoor images are full of shadows, e.g. the shadows of trees, buildings and cars. In most cityscape images, the windows of the buildings look dark from the outside, because the indoor illumination is often much weaker than the outdoor light. This can also be considered as a kind of shadows. See the first row in the figure below.

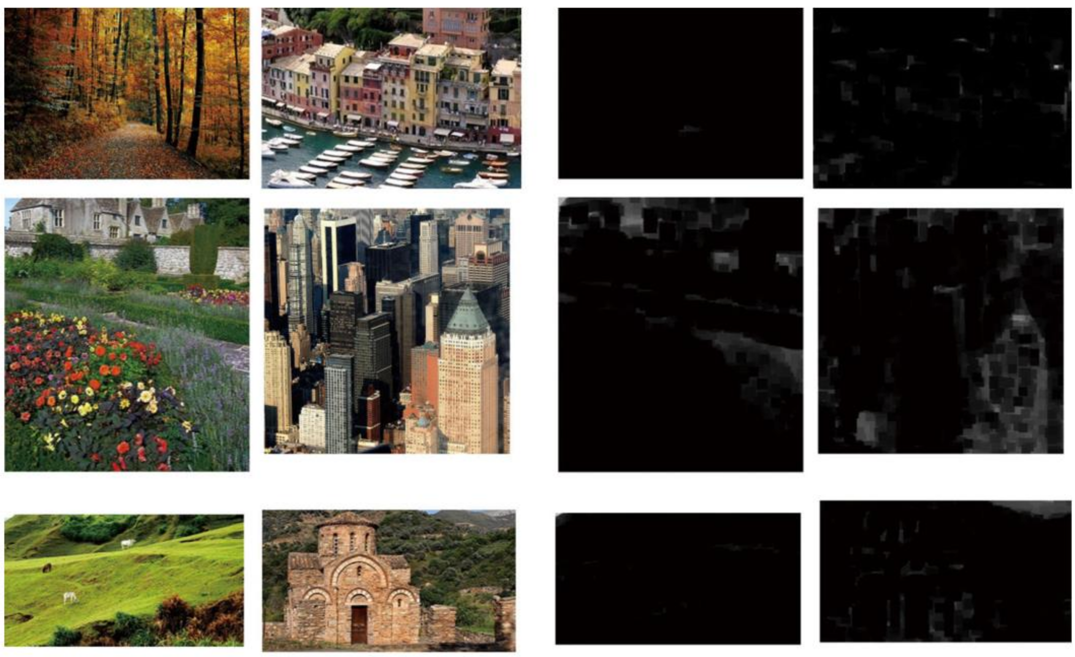
Second, the dark pixels can come from colorful objects. Any object with low reflectance in any color channel will result in dark pixels. For example, a green color has low intensity in its red and blue channels, and a yellow color has low intensity in its blue channels. Outdoor images often contain objects in various colors, like flowers, leaves, cars, buildings, road signs, or pedestrians. See the second row in the figure below. The colorfulness of these objects generates many dark pixels. Notice that by our definition a dark pixel is not necessarily dark in terms of its total intensity; it is sufficient to be dark in only one color channel. So a bright red pixel can be a dark pixel if only its green/blue component is dark.

Third, the dark pixels can come from black objects, like vehicles tyres, road signs, and tree trunks. See the third row in the figure below. These dark pixels are particularly useful for in-vehicle camera which oversees the road conditions.

- Shadows
- Colorful objects
- Black objects


2.2.2 Experimental Verification

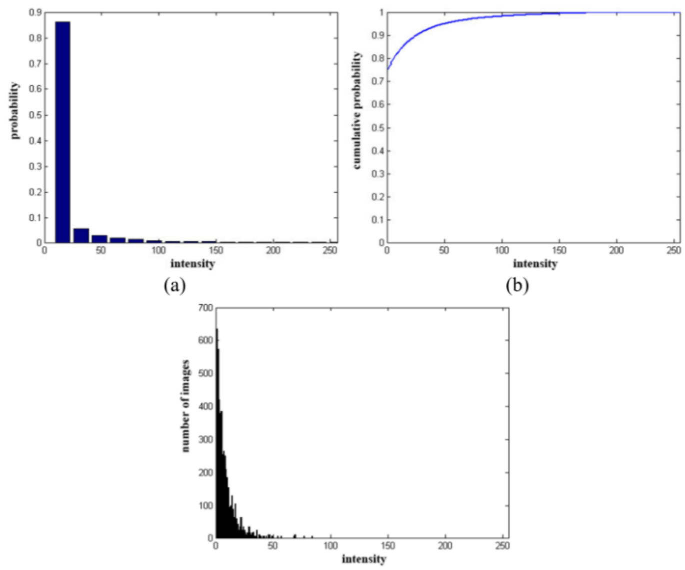
According to the prior, the dark channels of haze-free images should be mostly dark. Below figure shows some examples. Carefully inspecting these images, we can



find out the sources of the dark pixels: shallows, colorful objects, and black objects. They exist almost everywhere in the image.

To verify how good the dark channel prior is, He tests a large collection of images and studies their statistical properties. He downloads over 300,000 images from popular image search engines using 150 most popular tags annotated by the Flickr users. He is interested in outdoor landscape and cityscape scenes because they are the potential victims of haze. He only focuses on daytime images. Among them, he manually labels the haze-free ones. He randomly chooses 5,000 images and cuts out the sky regions. The images are resized so that the maximum of width and height is 500-pixel.

To compute the dark channels, he sets the patch size. As mentioned before, a larger patch has a better chance to contain a dark pixel. But we expect the prior to provide local constraints instead of more global ones. In this experiment, He sets the patch as a rectangle of 15 × 15 pixels. He computes the dark channels of all the 5,000 images.




Left top in the figure above is the distribution of the pixel intensity of all the 5,000 dark channels. Each bin contains 16 intensity levels. We can find that about 86% of the pixels fall in the first bin. Right top is the corresponding cumulative distribution. We can see that about 75% of the pixels in the dark channels have zero intensity, 90% of them is below the intensity 25.

He also concerns whether some images have bright dark channels. He computes the average intensity of each dark channel. The distribution is show in the below of the above figure. We can find that most dark channels have very low average intensity, indicating that most images obey this prior.

**2.3 Haze Removal Using Dark Channel Prior**

2.3.1 Estimating The Transmission

Here, we first assume that the atmospheric light A is given. We will present an automatic method to estimate the atmospheric light in section 2.3.4. We further assume that the transmission in a local patch $\Omega(x)$ is constant. We denote the patch's transmission as $\tilde{t}(x)$. Taking the min operation in the local patch on the haze imaging Equation, we have:

$$\min_{y\in\Omega(x)}(I^c(y)) = \tilde{t}(x)\min_{y\in\Omega(x)}(J^c(y)) + (1-\tilde{t}(x))A^c$$

Notice that the min operation is performed on three color channels independently. This equation is equivalent to:

$$\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c}) = \tilde{t}(x)\min_{y\in\Omega(x)}(\frac{J^c(y)}{A^c}) + (1-\tilde{t}(x))$$

Then, we take the min operation among three color channels on the above equation and obtain:

$$\min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c})) = \tilde{t}(x)\min_c(\min_{y\in\Omega(x)}(\frac{J^c(y)}{A^c})) + (1-\tilde{t}(x))$$

According to the dark channel prior, the dark channel $J^{dark}$ of the haze-free radiance J should tend to be zero:

$$J^{dark}(x) = \min_c(\min_{y\in\Omega(x)}(J^c(y))) = 0$$

As $A^c$ is always positive, this leads to:

$$d = \min_c(\min_{y\in\Omega(x)}(\frac{J^c(y)}{A^c})) = 0$$

Putting the above equation to the dark channel prior one, we can estimate the transmission $\tilde{t}$ simply by:

$$\tilde{t}(x) = 1 - \min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c}))$$

In fact, d is the dark channel of the normalized haze image $I^c(y)/A^c$. It directly provides the estimation of the transmission.

As we mentioned before, the dark channel prior is not a good prior for the sky regions. Fortunately, the color of the sky is usually very similar to the atmospheric light A in a haze image and we have:

$$\min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c})) \to 1, \text{ and } \tilde{t}(x) \to 0$$

in the sky regions. Since the sky is at infinite and tends to has zero transmission, the equation above gracefully handles both sky regions and non-sky regions. We do not need to separate the sky regions beforehand.

In practice, even in clear days the atmosphere is not absolutely free of any particle. So, the haze still exists when we look at distant objects. Moreover, the presence of haze is a fundamental cue for human to perceive depth. This phenomenon is called aerial perspective. If we remove the haze thoroughly, the image may seem unnatural and the feeling of depth may lost. So we can optionally keep a very small amount of haze for the distant objects by introducing a constant parameter w ranges in 0 to 1:

$$\tilde{t}(x) = 1 - w\min_c(\min_{y\in\Omega(x)}(\frac{I^c(y)}{A^c}))$$

The nice property of this modification is that we adaptively keep more haze for the distant objects. The value of w is application-based.

(a) (b) (c) (d)

Figure above's (b) is the estimated transmission map using the patch size 15*15. It is roughly good but contains some block effects since the transmission is not always constant in a patch. In the next subsection, we refine this map using guided filter.

### 2.3.2 Guide Filter

We first define a general linear translation-variant filtering process, which involves a guidance image I, an input image p, and an output image q. Both I and p are given beforehand according to the application, and they can be identical. The filtering output at a pixel i is expressed as a weighted average:

$$q_i = \sum_j W_{ij}(I) p_j$$

where i and j are pixel indexes. The filter kernel $W_{ij}$ is a function of the guidance image I and independent of p. This filter is linear with respect to p.

A concrete example of such a filter is the joint bilateral filter. The bilateral filtering kernel is given by:

$$W_{ij}^{bf}(I) = \frac{1}{K_i} \exp(-\frac{|x_i - x_j|^2}{\sigma_s^2}) \exp(-\frac{|I_i - I_j|^2}{\sigma_r^2})$$

where x is the pixel coordinate, K is a normalizing parameter to ensure that $\sum W_{ij}^{bf} = 1$. The parameter $\sigma_s$ and $\sigma_r$ adjust the spatial similarity and the range (intensity/color) similarity respectively. The joint bilateral filter degrades to the original bilateral filter when I and p are identical.

Now we define the guided filter and its kernel. The key assumption of the guided filter is a local linear model between the guidance I and the filter output q. We assume that q is a linear transform of I in a window $w_k$ centered at the pixel k:

$$q_i = a_k I_i + b_k, \forall i \in w_k$$

where $a_k$, $b_k$ are some linear coefficient assumed to be constant in $w_k$. We use a square window of a radius r. This local linear model ensures that q has an edge only if I has an edge, because $\nabla q = a \nabla I$. This model has been proven useful in image matting, image super-resolution, and haze removal.

To determine the linear coefficients, we seek a solution to minimizes the difference between q and the filter input p. Specifically, we minimize the following cost function in the window:

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

Here $\varepsilon$ is a regularization parameter preventing $a_k$ from being too large. We will investigate its significance below. The solution can be given by linear regression:

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

$$b_k = \overline{p}_k - a_k \mu_k$$

Here, $\mu_k$ and $\sigma_k^2$ are the mean and variance of I in $w_k$, $|w|$ is the number of pixels in $w_k$, and $\overline{p}_k = 1/|w| * \sum_{i\in w_k} p_i$ is the mean of p in $w_k$.

Next we apply the linear model to all local windows in the entire image. However, a pixel i is involved in all the windows that contain i, so the value is not the same when it is computed in different windows. A simple strategy is to average all the possible values of $q_i$. So after computing $(a_k, b_k)$ for all patches $w_k$ in the image, we compute the filter output by:

$$q_i = \overline{a_k} I_i + \overline{b_k}$$

With this modification $\nabla q$ is no longer scaling of $\nabla I$, because the linear coefficients $(\overline{a_k}, \overline{b_k})$ vary spatially. But since $(\overline{a_k}, \overline{b_k})$ are the output of an average filter, their gradients should be much smaller than that of I near strong edges. In this situation we can still have $\nabla q = \overline{a} \nabla I$, meaning that abrupt intensity changes in I can be mostly maintained in q.

We point out that the relationship among I, p and q are indeed in the form of image filtering. In fact, $a_k$ can be rewritten as a weighted sum of p: $a_k = \sum_j A_{kj}(I) p_j$, where $A_{ij}$ are the weights only dependent on I. For the same reason, we also have $b_k = \sum_j B_{kj}(I) p_j$ and $q_i = \sum_j W_{ij}(I) p_j$. It can be proven that the kernel weights can be explicitly expressed by:

$$W_{ij}(I) = \frac{1}{w^2} \sum_{w_k} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon}\right)$$

Some further computations show that $\sum_j W_{ij}(I) = 1$. No extra effort is needed to normalize the weights.

The criterion of a “flat patch” or a “high variance” is given by the parameter $\varepsilon$. The patches with variance much smaller than $\varepsilon$ are smoothed, whereas those with variance much larger than $\varepsilon$ are preserved. The effect of $\varepsilon$ in the guided filter is similar with the range variance $\sigma_r^2$ in the bilateral filter. Both parameters determine what is an edge/a high variance patch that should be preserved".

We apply guided filter to $\tilde{t}(x)$ and obtain $t(x)$.

2.3.3 Recovering the Scene Radiance

With the transmission map, we can recover the scene radiance according to Equation (1). But the direct attenuation term J(x)t(x) can be very close to zero when the transmission t(x) is close to zero. The directly recovered scene radiance J is prone to noise. Therefore, we restrict the transmission t(x) to a lower bound $t_0$, which means that a small certain amount of haze are preserved in very dense haze regions. The final scene radiance $J(x)$ is recovered by:

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A$$

A typical value of $t_0$ is 0.1. Since the scene radiance is usually not as bright as the atmospheric light, the image after haze removal looks dim. So we increase the exposure of $J(x)$ for display.

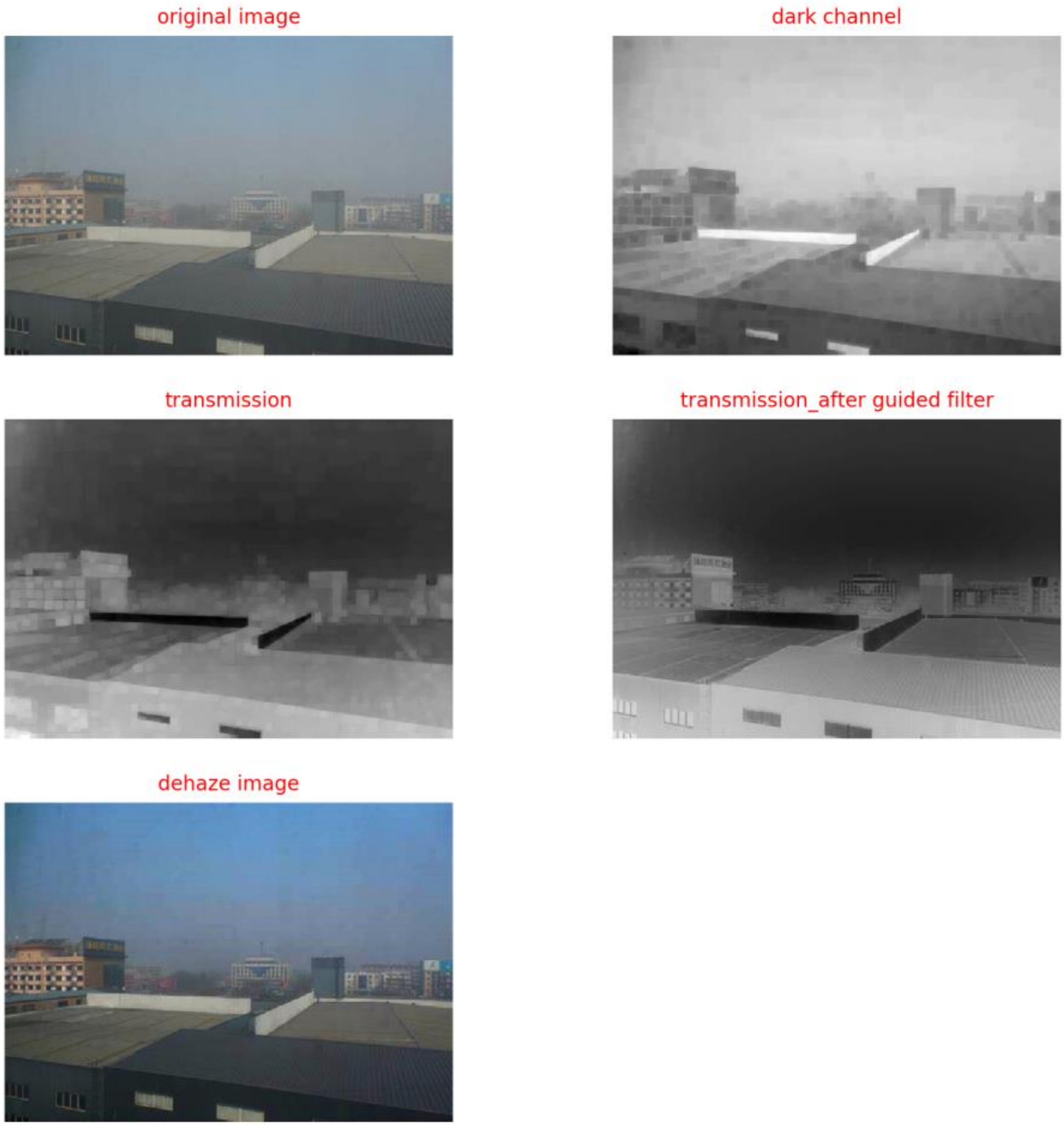
2.3.4 Estimating the Atmospheric Light

In most of the previous single image methods, the atmospheric light A is estimated from the most haze-opaque pixel. For example, the pixel with highest intensity is used as the atmospheric light and is furthered refined in another paper. But in real images, the brightest pixel could on a white car or a white building.

The dark channel of a haze image approximates the haze denseness well. We can use the dark channel to improve the atmospheric light estimation. We first pick the top 0.1% brightest pixels in the dark channel. These pixels are most haze-opaque. Among these pixels, the pixels with highest intensity in the input image I is selected as the atmospheric light. Note that these pixels may not be brightest in the whole image. This simple method based on the dark channel prior is more robust than the "brightest pixel" method. We use it to automatically estimate the atmospheric lights for all images shown in the paper.

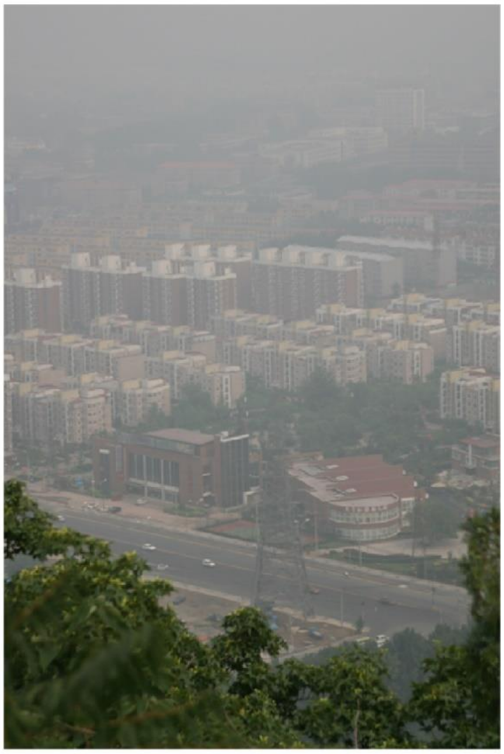
## 3. Experiment

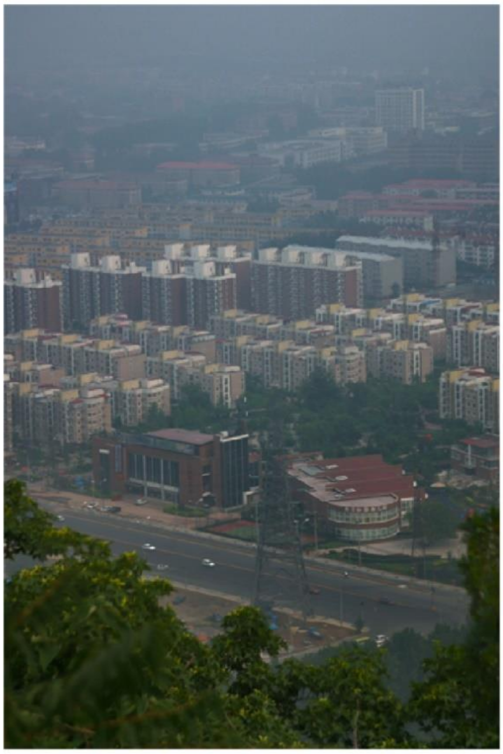
We first show basic experiment about the fog1.jpg and fgo2.jpg

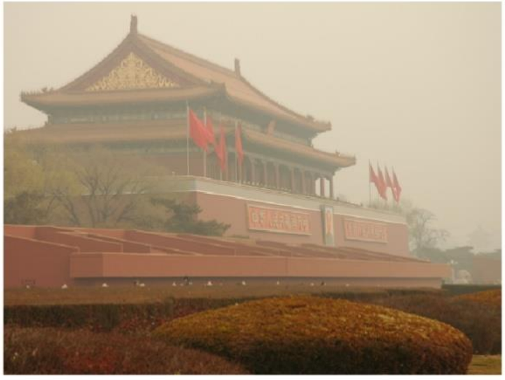

original image
dark channel
transmission
transmission_after guided filter
dehaze image

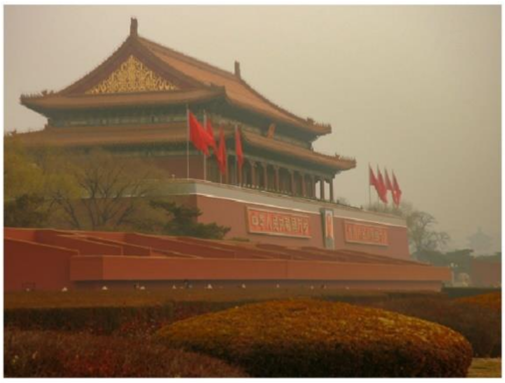
As we can see above the dark prior is successful in haze removal problem.

We also test some other images, and this algorithm still works.









## 4. Discuss the parameters and limitation

dark_window_size=7

dark_window_size =15


dark_window_size =31


We can find that as dark window size increases, dark channel is getting large, the saturation is getting small. The noise number increases.

w=0.95


w=0.75


w=0.55


We can find that as w increases, the haze removal ability decreases.

t0=0.1


t0=0.3


t0=0.5


We can find that as t0 increases, the haze removal ability decreases and the image is getting dark, due to t0 affects the explosure time.

Limitation:

The limitation of this algorithm is that, it is a statistic based method. It will fail in some region where colors is small to the atmospheric color. Besides, there are still 25% pixels not zeros in the dark channel, so a modified solution is to adaptively find a close to zero value instead of using zero directly in the computation.