An Improved Single Image Haze Removal Algorithm Based on Dark Channel Prior and Histogram Specification

Shuai Yang\textsuperscript{2}, Qingsong Zhu\textsuperscript{1}*, Jianjun Wang\textsuperscript{1}, Di Wu\textsuperscript{3}, and Yaoqin Xie\textsuperscript{1}

\textsuperscript{1}Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, \textsuperscript{2}School of Electronic Engineering, Tsinghua University, Beijing, 100084 \textsuperscript{3}School of Software, Shanghai Jiaotong University, Shanghai, 200240

\{qs.zhu,shuai.yang2,wangjianjun,di.wu,yq.xie\}@siat.ac.cn

Abstract. We introduce an improved single image haze removal algorithm, which combines dark channel prior (DCP) and histogram specification. First, the dark channel prior knowledge proposed by Kaiming He is analyzed and a conclusion is drawn that the haze removal image based on dark channel prior will have a tendency to dim and indistinct in some specific situations. Especially, when cleaning the haze in the image with large background area and low contrast, DCP result appears obvious anamorphose. Next, in order to improve the dehazing result of this kind of image, we propose an approach to change the contrast and intensity of haze removal image after DCP method by rebuilding the histogram of the image. Then, a modified approach is applied to fit general haze image. We experiment our method with a variety of outdoor haze images. The effectiveness of our method is demonstrated in comparison with DCP result when the input image contains low contrast scene and large background area, such as thick fog or dark surroundings in dusk. Our job makes up the deficiency of the dark channel model for this kind of image and enhance the contrast of the scene. Furthermore, the experimental results show that the dehazing effect on general haze image appears more close to real scene than dark channel model.

Keywords: haze removal, dark channel prior, histogram specification, intensity, contrast

1 Introduction

Images taken outdoors are usually degraded because of the light being absorbed or scattered by suspended particles in atmosphere. The degrading will arise visibly in adverse weather such as haze, fog and so on. When taking photos in haze, due to the influence of the water-droplets, on the one hand, the irradiance received by camera from the foreground will be hindered after being absorbed and scattered, on the other hand, the diffuse reflection will bring air-light into camera. So, the haze images either have lower intensity than original scene, or look greyish white due to air-light.
As we know, in computer vision applications, many techniques extract the scene feature by analyzing the intensity and contrast of the image. So the general image processing project with input of the haze image usually results in deviation. In this case, there are strong demands of haze removal techniques, especially in the applications of traffic, outdoor record, or terrain detection.

Haze removal is one of the well-studied problems in computer vision. In the literature, relate works can be divided into two directions: the one is adjusting the haze image by the traditional techniques of image processing, adding prior knowledge by analyze the physical model of haze is another. First kind of the approach focuses on contrast and intensity enhancement, include histogram equalization, gamma correction, homomorphic enhancement, wavelet method [10] and so on. Ming-Jung Seow et al. [1] combine the frequency domain with grayscale histogram transform, which use reflecting as the base of frequency domain and then enhance the contrast of the image. An effective approach proposed in [6-8] is adaptively enhancing the local image with local histogram equalization. Almost all the approaches mentioned above are compared and evaluated in [2, 9]. But the defect of these methods above is that they neglect the physical property, such as the depth information of haze or the attenuation when light passing though the haze. So, their haze removal results have limitation at the quality of images.

Recently, the research of using physical model of haze as prior knowledge has made significant progresses. Researchers devote to studying the physical characteristics of haze such as how does the light pass through the haze so as to find deep information of the haze image. Tan et al. observe that compared with haze image, the haze-free image must have higher contrast, so, based on Markov Random Field, a novel approach combined maximization of local contrast is proposed to remove haze [3]. Tan’s approach improves the contrast of the image well, but also easily lead to over-enhanced in practice. Fattal et al. find a prior knowledge that the transmission and surface shading are locally uncorrelated. Based on this prior, Fattal gets haze-free image by estimating the albedo of the scene, and then deduces the color of the whole image by MRF [4]. This method produces an impressive result in general haze image while losing effectiveness when dealing with thick haze that is lack of too much color information. In order to remove the haze drastically, Kaiming He et al. do a massive experiments of the haze-free image and propose a novel prior knowledge - dark channel prior. He points that in most of the local regions which do not cover the sky, it is very often that some pixels have very low intensity in at least one color (RGB) channel [5]. Utilizing this prior, He builds the distribution of haze thickness, and then eliminates the haze from the haze image. DCP approach is concision and graceful, and it achieves obviously dehazing result for almost all kinds of the haze images. But we observe that the contrast and intensity of haze image after DCP approach will unavoidably tend to lower than those of the real scene. We will elaborate the reason why DCP method has defects on the intensity and contrast of image in the next section. During experiment, we find that in the situation of large background (such as sky, clouds, or just thick haze) area and low contrast in the image, DCP result makes the background darker obviously, which drive the whole image to dim(Figure 1).

In order to improve the defect of DCP method, we propose a method combining the dark channel prior and histogram specification. First, we do a research on the image
with large background area and low contrast, and build the histogram of this kind of haze images together with their haze removal result. Then we find that compared with original haze image, the histogram of haze removal image has a tendency of left-shifting and narrow, and some sharp points occurs in the high intensity region. This kind of change on histogram intimate that the contrast and the intensity of the whole image will get lower, and the noise on the background appears. Next, we rebuild the histogram of haze removal image by expand the low intensity area of the histogram and eliminate the sharp point at the high intensity area. We do a lot of experiment and the result demonstrate that for the haze image with much background area and low contrast, our approach achieve a much better result than DCP. Figure 1 shows the example of our work. Furthermore, in order to make our method more universality, we modify our method to fit the general haze image, which makes the haze removal result more close to the real.

Fig. 1: The example of the haze image with large background area and low contrast: (a) haze image (b) haze removal image by DCP approach (c) our result

This paper is organized as follows. In Section 2 we review DCP approach as well as the principle of histogram specification, and analysis the defects of the theory. In Section 3 we explain details of our algorithm which improve DCP result by rebuilding the histogram with the histogram of haze and haze removal image. In Section 4 we present the experimental result and analyze the limitations of our method. Finally in section 5 we summarize this paper.

2 Theory Review and Analysis

In this section we review and analyze the approach of single image haze removal using dark channel prior and the principle of histogram specification. On the one hand, we want to help the readers recall related knowledge, on the other hand, after analyzing the deficiency of DCP method, it’s easy to understand the benefit of improving DCP result by histogram specification.

2.1 Single Image Haze Removal Method based on DCP and Its Defects

In computer vision applications, a haze image $I(x)$ can be commonly present as follows:
\[ I(x) = J(x)t(x) + A[1 - t(x)] \]  

(1)

In Equation (1), \( I(x) \) is the intensity of the image caught by camera, namely, it’s the intensity of the scene mixed with haze. \( J(x) \) is the true intensity of foreground, and \( t(x) \) is the medium transmission, which describes the percentage of residual energy when the foreground irradiance passes through the medium and reaches the camera. \( A \) is global atmospheric light. The aim of haze removal is estimating \( J(x), t(x), \) and \( A \) from \( I(x) \).

He proposes a dark channel prior, this prior knowledge is summarized by a massive experiments of the haze-free image. Dark channel prior points that in most of the local regions which do not cover the sky, it is very often that some pixels have very low intensity in at least one color (RGB) channel [5]. He define \( J_{\text{dark}}(x) \) to present the dark channel at \( x \):

\[ J_{\text{dark}}(x) = \min_{c \in \{ R, G, B \}} \left( \min_{y \in \Omega(x)} J^c(y) \right) \]  

(2)

In Equation (2), \( J^c \) is one of the RGB channel of \( J \), and \( \Omega(x) \) is a square region with center \( x \). He find that for a haze-free image, if \( x \) isn’t belong to the region of sky, \( J_{\text{dark}}(x) \) is always showing a low value and tending to zero. Obviously, this is the reason \( J_{\text{dark}}(x) \) is named as dark channel.

But, if the scene is merging with haze, because \( I(x) \) is the intensity of the foreground mixed with atmospheric light, \( I(x) \) is usually brighter than \( J(x) \), namely, the medium transmission \( t(x) \) usually has a low value. So, the dark channel of \( I(x) \) will have a high value compare with \( J(x) \). In intuition, the distinction on dark channel between \( I(x) \) and \( J(x) \) can help to remove haze.

Using dark channel prior, we can easily estimate global atmospheric light \( A \), and then estimate the medium transmission \( \tilde{t} \). The way to estimate \( \tilde{t} \) is shown below.

We use atmospheric light \( A \) to divide Equation (1), then calculate the dark channel, so we obtain the equation below:

\[ \min_{c \in \{ R, G, B \}} \left( \min_{y \in \Omega(x)} \frac{J^c(y)}{A^c} \right) = \tilde{t}(x) \min_{c \in \{ R, G, B \}} \left( \min_{y \in \Omega(x)} \frac{J^c(y)}{A^c} \right) + [1 - \tilde{t}(x)] \]  

(3)

According to the dark channel prior we mentioned above, the dark channel of image without haze tend to zero, so we allege the following equation:

\[ \min_{c \in \{ R, G, B \}} \left( \min_{y \in \Omega(x)} \frac{J^c(x)}{A^c} \right) = 0 \]  

(4)

Combine Equation (3) with Equation (4), we obtain the present of \( \tilde{t} \):

\[ \tilde{t}(x) = 1 - \min_{c \in \{ R, G, B \}} \left( \min_{y \in \Omega(x)} \frac{J^c(y)}{A^c} \right) \]  

(5)

Furthermore, He introduces the parameter \( w \) into Equation (5) to keep a small amount of haze in order that people can perceive the depth of image. Finally, the medium transmission \( \tilde{t} \) is:
\[
\tilde{t}(x) = 1 - \omega \min_{c \in \{R,G,B\}} \left( \min_{x \in \Omega} \frac{f^c(x)}{A^c} \right)
\]

Then, he applies a soft matting algorithm to refine the medium transmission \( \tilde{t} \), and obtain the accurate medium transmission \( t \). Now we can estimate the foreground image \( J(x) \) using following equation with \( I(x) \), \( t(x) \), and \( A \) (in Equation (7), \( t_0 \) is applied to avoid a low value of denominator):

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

As we can see, DCP approach is concise and graceful, but there still exist three defects in DCP theory as elaborated below.

Firstly, considering the second term of Equation (1), we find that \( 1 - t(x) \) is equivalent to the thickness of haze, then we can conclude that in DCP theory, when different intensity of foreground irradiance passing through same thickness haze, they have the same degree of attenuation. Now, it is not hard to understand why the image contrast after DCP method become lower. But in practice, distortion of the light will more evident on the higher intensity region, so the foreground in haze-free image usually have higher contrast than haze removal image by DCP method.

Secondly, we think that the first term of Equation (1), which is called direct attenuation, is not suitable to be presented as linear relation with image intensity, namely, we think that DCP method underestimates the attenuation of the foreground irradiance though haze. So, the haze removal image should be brighter than DCP result.

Thirdly, if the haze image has large background area or low contrast, DCP result will cause bad result on dark background and low contrast. This is because the background is regarded as the scene merging with thick haze and on the other hand, DCP method reduce the contrast of the foreground. In vision, this kind of image exposes the deficiency of dark channel prior, and corresponding haze removal image is not an optimized result which is not close to the real scene.

Obviously, the defects mentioned above have a little effect on haze removal result. But if the haze image has additional characteristic, just like the haze image with large background area and low contrast, the deficiency of DCP method will result obviously anamorphose. Figure 1 shows an example. We can see that after DCP method, the intensity of the background becomes dark, and some details, such as the wheels of plane and the write line on the ground, is indistinct. For these situation, we need to improve the deficiency of DCP method by adding histogram specification to the result. Next we will review the principle of histogram.

### 2.2 Principle of Histogram

The intensity histogram, which is usually the only information for a grayscale image, indicates the probability of every gray value of the pixel. It also connects the image with its probability distribution. What we are interested in intensity histogram is the corresponding probability of each gray value, but not the coordinate position of pixel. We can conveniently obtain the contrast and intensity information of one image by
analyzing the histogram. In addition, a lot of traditional image processing techniques are based on histogram, such as object segmentation[11], gray-scale transformation, Contrast transformation, etc.

Histogram specification, whose core idea is to rebuild the histogram of image, is one of the classical method to adjust the gray intensity value. It is very convenient to change gray value of image by altering histogram, due to the property of the histogram which has little relation with the scene shape. The benefit of the histogram specification is that we can enhance the only intensity region which we are interested in, but the other region is invariant. Meanwhile, the weakness of the histogram specification is also obvious. During histogram specification, it will occur grayscale degeneracy on image, which can lead to serious anamorphose. So, it is appropriate that we use histogram specification in a small intensity area, such as doing some posterior processing. In this paper, we will use the histogram specification to improve the image after DCP method.

We can utilize gray value of every pixel in grayscale image to build the corresponding histogram, but in this paper, we need a kind of color image processing. Relate to dark channel prior, we know that for each colorful pixel, its intensity of RGB channel has big distinction, which indicate that doing the histogram specification of RGB channel respectively and then merging them is an impractical opinion. Instead, we change the image representation method from RGB matrix to HSI matrix, and then use intensity matrix $I$ to build histogram.

HSI matrix uses hue matrix, saturation matrix, and intensity matrix to describe a color image. In virtue of separating the color information and intensity information in an image, HSI matrix is practical for image processing. The transformation formula between RGB matrix and HSI matrix is:

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \quad (8)$$

$$S = 1 - \frac{3}{R + G + B} \min(R, G, B) \quad (9)$$

$$I = \frac{1}{3}(R + G + B) \quad (10)$$

The parameter $\theta$ in Equation (8) is defined below:

$$\theta = \arccos \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \quad (11)$$

Now, we can build intensity histogram using $I$ matrix.

### 3 Our approach

#### 3.1 Improvement on Haze Removal Image with Large Background Area and Low Contrast

In section 2.1, we analyze the defects of DCP method, and we point that, if the image has the feature of low contrast and owning large area with background, DCP method
will obtain an indistinct and dim result. In this section, we demonstrate our point with image and its histogram, and improve DCP result with histogram specification.

We have observed a lot of image histogram, finding that if the image has large area of atmospheric light, there will be a distinct peak at the region of high intensity in histogram, which represents the atmospheric light area in image. The interesting thing is that, for the most of images which have bad result using DCP method, their histogram has the same peak in the region of high intensity. The reason is evident, it is because for the image with large background area and low contrast, its background pixels contribute to the peak on histogram. Figure 2 shows one of the results.

Comparing the histogram of haze and haze removal image, we find some interesting characteristics, which we will show below.

Fig. 2: The haze removal example of the haze image with large background area and low contrast: (a) haze image (b) haze removal image by DCP approach (c) our result (d) histogram of haze image (e) histogram of haze removal image by DCP approach (f) histogram that we rebuild (g) histogram of our result

Firstly, we find in the haze image that, it is obviously in the histogram that there exist a minimum point between foreground and background distribution, namely, we can easily distinguish whether the intensity present the foreground or background from the figure of histogram. But after haze removal, the distribution of foreground and background have tendency to merging, so we can not find the critical point between them. This kind of merging indicates that the contrast of the scene become lower.

Secondly, we find that after haze removal, the peak in the region of high intensity disappeared. Instead, in the high intensity region of histogram, there occurs a lot of
sharp points. This is because the intensity of the background is similar to global atmospheric light $A$, so the area will be removed as the result of being regarded as thick haze merged with scene. Also, the dehazing on the background will leave some bad point. As the result, background area on the haze removal image will become dark, and occurs noise inescapably. The appearance above has a bad effect on the qualities of image, but this characteristic also benefits us to distinguishing the foreground and background distribution on histogram. In intuition, we can utilize the start point of the sharp point in histogram as the boundary of foreground and background.

Thirdly, compared with haze image, the histogram of image after haze removal have the tendency of left-shifting, that is to say, the image after haze removal is darker than the haze image.

Based on the characteristics mentioned above and the defects of DCP method we have analyzed, we propose an algorithm to enhance the dehazing image by rebuilding the histogram of image after DCP method. The rebuilt histogram strengthen the contrast as well as remove the sharp points. Then, we do the histogram specification with rebuilt histogram to obtain the better dehazing result. The details of our approach are stated below: (relevant flow chart is showed in Figure 3)

![Flow Chart](image)

**Fig. 3:** The flow chart of our improvement on haze removal image with large background area and low contrast

1. Build the histogram $A$ of haze image, and find the intensity $x_A$ corresponding to the peak in the high intensity region of histogram.

2. Remove the haze in image using DCP method, then build the histogram $B$ of haze removal image, and next find the start point $x_B$ of the sharp points in histogram $B$. 
3. Divide the histogram $B$ into two parts by the point $x_B$. We define that the low intensity part $B_1$ presents foreground distribution and the high intensity part $B_2$ presents background distribution.

4. Rebuilt $B_2$. We use the increase convex expression $B_2'$ in place of $B_2$, in order to strengthen the contrast as well as eliminate the sharp points. In this paper, we use quadratic function to replace $B_2$. We set $x_B$ and $x_A$ as start point and end point respectively, and the function value is adaptively determined by histogram $B$.

5. Combine $B_1$ and $B_2'$, and we get the rebuilt histogram $B'$. Finally, do the histogram specification with rebuilt histogram $B'$, and then we get the result image.

Note that we don’t rebuild foreground part $B_1$, this is because when rebuilding $B_2$, we strengthen the probability of high intensity, and at the same time, we reduce the probability of low intensity, which essentially strengthen the contrast of the foreground.

3.2 Improve on General Haze Image

As we mentioned in section 2.1, in practice, the foreground in haze-free image usually has higher contrast than haze image. On the contrary, the image contrast after DCP method becomes lower. Although this influence of the image contrast is not strong, it

![Fig. 4: The haze removal example of the general haze image: (a) haze image (b) DCP result with $w = 0.85m = 1$ (c) DCP result with $w = 0.85m = 2$ (d) our result (e) histogram of (a) (f) histogram of (b) (c) histogram of (c) (f) histogram of (d)
does decrease the image quality. So, based on the approach stated in section 3.1, we propose an alternative method for general haze image by rebuilding the histogram of haze removal image. This method strengthens the contrast of the haze remove image, and makes the improve image more close to the real image without haze.

By experimenting a lot of haze image, we find that, the cleaner haze is removed, the lower contrast of the scene will become. In order to show the relative between the effective of haze remove and the contrast of foreground, we apply two parameters to present the degree of haze removal. The one parameter is $\omega$ in DCP method, the other parameter is the times of haze removal $n$. Obviously, the degree of haze removal and the parameter has positive correlation, namely, when we increase the parameter $\omega$ or $n$, the haze in the image will have the tendency to become rarefaction. Figure 4 shows an example.

Comparing the histogram of haze removal in different degree, we have two conclusion which is similar to the observation in section 3.1.

Firstly, the histogram of haze removal image tend to left, in other words, tend to low intensity region with the degree of haze removal increasing.

Secondly, the histogram of haze removal image tend to shrinking with the degree of haze removal increasing, namely, the edge of the histogram tend to cliffy.

Based on the observation above, we proposed an approach to improve the haze removal image: (relevant flow chart is showed in Figure 5)

![Flow Chart](image)

**Fig. 5:** The flow chart of our improvement on general haze removal image

1. Choose an image with high degree of haze removal. This is because the intensity of haze is higher than foreground in practice, so the haze area will be enlarged while
the contrast of high intensity region in histogram is strengthen. If we choose an image with some haze remained, the improve image will be merged with more haze.

2. Select the right edge of the foreground distribution in histogram, and stretch the edge to high intensity region (See Figure 4.). So we obtain the rebuilt histogram and then use it to do the histogram specification.

There are two key points that we have to explain below as the additional remark.

Firstly, the method which we mentioned above need to extract the foreground distribution from whole histogram of haze removal image. If the image has small area of background, the foreground distribution is easy to be detected (See Figure 4(e-g)), but if there is large area of background occurring on image, distribution of the foreground and the background may not easy to be distinguished. In this case, as we know that after DCP method, there will appears sharp points on the histogram of image, we can use the start position of the sharp points to detect the distribution of foreground.

Secondly, we only stretch the right edge of the foreground distribution, but not operate the whole foreground distribution. This is because that the operating of whole foreground distribution will cause the serious anamorphose. On the other hand, similar to what we have explained in section 3.1, the stretching transformation either strengthen the contrast of the foreground, or increase the intensity of whole image, so the operation of one edge is enough to improve the haze removal image. We can see in Figure 4 that our improve is better than DCP result, and more close to the real image without haze.

4 Experiment Results

4.1 Experimental on Image with Much Background Area and Low Contrast

Figure 6 shows part of our experimental results. We can observe that when the haze image has low contrast, DCP result will become indistinct or even anamorphose. The middle column shows the haze removal image by DCP method. We can see on the image of tree, the intensive difference between the branch and the background is low and the whole image is dark; on the image of ice-mountain, the color becomes deeper, and there appears some black noise on the snow; on the image of city, the color in the sky is discontinuous. Our haze remove result showed on the right column demonstrate that our method improves the anamorphose above. On the image of tree, our result enhance the foreground by increase the intensity of the background; on the image of ice-mountain, our image strengthen the contrast compared with DCP result, and decrease the noise on the ice; on the image of city, our result brighten the image, moreover, smooth the discontinuous of the sky color. In a word, compared with DCP haze removal image, our result brighten the background and smooth the noise, which make the image more close to the real one.

Meanwhile, our method also has limitations. When the foreground and the background have high contrast, our method can not improve DCP result well. See Figure 7, because of the background has high intensity, our result lead to serious anamorphose.

4.2 Experimental on general haze image

Figure 8 shows part of our experimental results on the general haze image removal. The images contained in second and third row are haze removal image by DCP method with
different parameters. We can see that the images in second row have a good contrast but remain some haze; the images in third row have the better work on haze removal, but lose some contrast, so the whole images look dark and not real. Our result in fourth row is based on rebuilding the histogram of the images in third row, and we can see that compared with the image in third row, our result strengthen the contrast.

On the image of the mountain, our result strengthen the texture of the rock; on the image of the haystack, our result makes the whole image more bright, and the details of the haystack are more distinct; on the image of forest, the leaves are more vivid, and the color is beautiful. All in all, comparing with DCP result, our result is more bright and shows more details in scene, which makes the whole image more close to real.

But if the haze in the image is not removed cleanly, our method will increase the thick of the haze. See the forest image in the third and fourth row, our method enlarge the haze on the left.

Fig. 6: Some results of our experiment. The left column contains haze image, the middle column contains haze removal image by DCP method, and the right column contains our result.
**Fig. 7:** Failure case. The left is haze image, the middle is haze removal image by DCP result, and the column is our result.

**Fig. 8:** Some results of our experiment. The first row is haze image. The second row is haze removal image by DCP method(parameters: (a) $\omega = 0.65, n = 1$; (b) $\omega = 0.88, n = 1$; (c) $\omega = 0.83, n = 1$). The third row is haze removal image by DCP method(parameters: (a) $\omega = 0.85, n = 1$; (b) $\omega = 0.85, n = 2$; (c) $\omega = 0.83, n = 2$). The fourth row is our result.
5 Conclusion and Further work

We summarize our contributions as follows. We analyze DCP theory of haze removal, and find that DCP method has the deficiency of the theory in dealing with the contrast and intensity of haze image. Then a lot of observation on histogram of haze and haze removal images demonstrate our analysis. First, aiming at the images with much background area and low contrast, which have not good result on DCP result, we propose a single image haze removal by combining with dark channel prior and histogram specification, which strengthen the contrast and intensity of image after DCP method. Then, we improve our approach to fit general haze image, and experiment result demonstrate that our work make the scene more close to the real scene.

As further work, we would like to exploit our work in two directions:
1. Consider that the light with different frequency will suffer different attenuation when passing through haze, we could rebuild histogram on RGB channel respectively;
2. We know that matrix S (saturation) in HSI presents how much the white light merge into the pure color. In intuition, it is similar to Equation (1), and we will use this similarity to do some work.

References