Color Image Defogging in Dark Channel Prior Based Image Local Features

Rui-qiang MA$^{1,2}$ and Shan-jun ZHANG$^{2,*}$

$^{1}$Inner Mongolia University of Technology, Huhhot, Inner Mongolia, 010080, China
$^{2}$Department of Information Science, Faculty of Science, Kanagawa University, Hiratsuka, Kanagawa 259-1293, Japan

*Corresponding author

Keywords: Defogging, Dark channel prior, Color image, Image local feature.

Abstract. Bad weather, for example fog, haze, smog, etc. Nice to advanced video surveillance will be enhanced, so the needs of pattern recognition AI and related research will increase. The problem is that the contrast of the photographed image decreases, the distant visibility deteriorates greatly, the components, size, and shape of fogs are different but the adverse effects are similar. This study is based on a large amount of analysis of predecessors on the Dark channel prior defogging algorithm based on the research results, especially the He team's research results, depth of field processing, to abandon the He team in order to enhance the sense of depth to set the range of depth of field approach. In this research, we proposed a method to remove fog from uneven defogging image with fog density change, using dark channel prior and local gradient key point extraction technology. Estimate the transmission map by using the image local information and shaping the coarse map using the detailed map. In this way, a clear image is obtained while maintaining the fog removing quality.

Introduction

In recent years, there has been a growing demand for high-definition surveillance image clarity due to the widespread use of surveillance video lock-in and tracking off criminals, driverless, PM 2.5 response and the like. Whether it is user browsing or smart identification, high-definition video is the premise and basis of pattern recognition and object tracking accuracy.

The formation of a color Fog image is as shown in Figure 1, which makes it harder to use color image defogging, using a given fog image without any additional information, restoring it as a sharp image (Fog-free image), was once considered underdetermined problem.

![Figure 1. The model of light in fog state.](image)

I(x)=J(x)t(x)+A(1−t(x)) (1)

1.1 In the fields of computer vision and digital image processing, the atmospheric scattering model (1) based on the principle of foggy image formation is often used. In a foggy environment, a scene accepted by a camera sensor can be sourced in two parts, one from an attenuated scene and the other from a light source in the atmosphere to The light intensity of the camera. According to the law of conservation of energy, part of the intensity of light scattered by attenuation is equal to the intensity of light scattered by other light sources. Therefore, a foggy image can be represented as a linear model as follows:
In the formula (1), I(x) is a foggy image, J(x) is an object scene radiance, A is a global atmospheric light, t(x) is a medium transmission. The first term J(x)t(x) on the right side of the equation can also be regarded as the direct attenuation of the image. The second term A(1-t(x)). A can be approximated as a position-independent constant coefficient and the transmittance t(x) is position-dependent, and it is a two-dimensional transmissive image.

The atmosphere can be assumed to be homogenous, and the transmissivity is inversely exponentially related to the distance of the object from the camera:

$$t(x) = \exp(-\beta d(x))$$  \hspace{1cm} (2)

In equation (2), d(x) is the depth map, and $\beta$ is the atmospheric scattering coefficient, we can estimate the global illumination of the atmosphere by using the property that the distance x tends to infinity and obtain $A = I(x)$, that is, finding the maximum light intensity with the minimum transmissivity, As shown in equation (3):

$$A = \max_{y \in \{x | t(x) \leq t_0\}} I(y)$$  \hspace{1cm} (3)

Based on I(x), t(x), and A are known, the atmospheric scattering model formula is used to reduce and solve J(x) as shown in equation (4):

$$J(x) = I(x) - A t(x) + A$$  \hspace{1cm} (4)

However, for the color image information, solving the three quantities of J(x), t(x), and A at the same time really becomes a problem of under-definition. For each pixel, corresponding to three different color channels, you can get three equations. To sum up, researchers often estimate the transmissivity t(x) through various prior means such as dark channel a prior, and thus solve the remaining unknowns.

1.2 The past ten years, about the color image to fog about the situation is Tan and Toda team used MRF (Markov Random Field) to model the transmittance and maximize the local contrast, achieving the purpose of de-fogging color images [1][2]. The algorithm assumes relatively low fog contrast, the transmittance is related to the depth of the scene, and the local area is approximately constant, the drawback of this algorithm is simply restores the scene from the perspective of image enhancement rather than the physically handicapped problem. Therefore, the images obtained by this method are often supersaturated, and the statistical characteristics of the old images.

That year, proposed an image de-fog method based on the physical model[3]. It takes the simplified scene radiation J(x), using local albedo constant to achieve. Since the value of shading the transmittance t(x) depends on the scene depth d(x) and the concentration of the fog, that is, the scattering degree $\beta$. The paper proposes the hypothesis that the surface shading And transmittance of the scene are statistically independent. However, due to the limitations of the "statistical independence" assumption, such as the statistical characteristics obtained in the case of low SNR (dense fog scene), the accuracy of the statistics will be inaccurate. Moreover, the algorithm is based on color information and does not apply to Grayscale, or fog scene caused by the colorless.

1.3 A Chinese research team proposed a very simple and straightforward algorithm-dark channel prior[4], deftly cleansing process, with epoch-making significance. Based on the dark channel a prior de-fogging algorithm is actually a statistical algorithm, the author summarizes a large number of outdoor fog-free images, it found in its local area there are some pixels, these pixels at least A color channel has a very low brightness value. Dark channel prior proof is from 5000 Flickr.com no fog pictures, manually cut off the sky area, and adjusted to 500×500 size. After doing the dark channel calculation, the brightness histogram is calculated and found that 75% of the pixels have a brightness of 0 and 90% of the pixels have a brightness value of less than 25. The main cause of low brightness is the shadow of an object, a colored object or surface, and a black object or surface. Based on the transformation (1), the transmittance can be obtained. Based on the transmittance and the estimated atmospheric light component, a restored image:
\[
J_{\text{dark}(x)} = \min_{c \in \{r, g, b\}} \left( \min_{y \in \Omega} (J_c(y)) \right)
\]

(5)

However, there are also some drawbacks: it is generally difficult to obtain a satisfactory result for similar images in the target scene and atmospheric light, such as snow, white background wall, sea, and the like.

1.4 In the subsequent research, some basic improvements are based on this algorithm. The various prior knowledge in the work of predecessors play a complementary role to make up for the problem that a single prior feature does not apply in some scenarios. The Tang team proposed constructing a algorithm that takes various prior characteristics as input, and estimates the transmissibility map using a random forest model. The machine learning algorithm is applied to the de-fogging algorithm[5], a certain amount of training data is used to automatically find the suitable model to estimate the transmission map, which proves that the dark channel a prior features are the most informative features. At the same time, this method also affects many subsequent algorithms that use machine learning to defog. However, in practice, it is often difficult to train a large number of pairs of foggy and foggy images to train the network. The authors use the images synthesized by the atmospheric scattering model to conduct experiments as training data. From the Internet to collect part of the fog-free photos, arbitrarily sampled 16*16 image blocks, and then the image blocks into a foggy image block. The synthesis is based on two assumptions: the image content is not related to transmittance verification, that is, the same image content may have different depths, the transmittance is constant in a local area, that is, pixels in a small image block have the same depth. Therefore, we can arbitrarily select a value of 0 to 1 as a block of image t, and then use this to compose a foggy photo. To simplify the learning process, set A to constant 1.

The features of prior knowledge extracted from the training model tend to be highly correlated with the image content, which is contrary to the above assumption that "image content is not related to transmittance". In order to eliminate the influence of this correlation, the article first sorts the extracted features during training. As the location information is disrupted, the relevance of the content is also broken.

The team use the regression model to estimate the scene depth. The difference is presents a new a prior knowledge: color attenuation a prior. The a prior is based on the finding that in the fog-free region the brightness and saturation are similar; however, the brightness and color saturation of the fog are quite different[6].

This phenomenon can be explained by the atmospheric scattering model: in regions of high fog concentration, the proportion of light that is scattered by the atmosphere increases, which results in the colors becoming white, decreasing in saturation and increasing in brightness. Since the fog concentration is related to the scene depth, we can conclude that the scene depth is directly proportional to the "difference in brightness and saturation." In order to express this relationship concretely, the article uses a simple linear regression model to predict the scene depth with the brightness and saturation as variables.

1.5 In 2016, Cai team is using convolution neural network to estimate the transmission map [7], to achieve the de-fog effect. In terms of model usage, the convolutional neural network in the dissertation can be divided into three layers, which are the "feature extraction" layer based on Maxout unit respectively, the "multi-scale mapping" layer using different size convolution kernels, Maximum "layer, and a" non-linear regression "layer using a bilateral ReLU as an activation function.

1.6 Ren team proposed a multi-scale convolutional neural network structure [8]. By using rough networks to generate coarse-grained transmission maps and then finer networks to get more detailed images. The Li team thoroughly solved the problem of de-fogging into an end-to-end problem [9]. In order to achieve end-to-end training and avoid the use of additional methods to estimate the global atmospheric lighting A, the author made some simple mathematical transformations. In
addition, the author also applied the part of defogging to the target detection as a kind of pretreatment.

Our Method

Based on the research results of He term, the usage of Soft Matting has not been described in detail yet. Using the local gradient key point extraction technique, it consists of two stages, detection of keypoint detection and description of keypoint amount. Each process is as follows:
1. Scale and key point detection, 2. Localization of key points, 3. Calculation of orientation, 4. Description of key point amount.
1. In scaling and key point extraction, scale and key points are detected by DoG processing, and in 2. localization of key points, points not pointed as key point points are deleted from the key points detected in 1, Perform pixel estimation. 3. In the calculation of the orientation, obtain the orientation of the keypoint in order to obtain the invariant key points for the rotation. 4. In the description of the keypoint amount, in 3, describe the key point amount of the keypoint based on the obtained orientation. Details of each process will be described below.

Dark channel processing is based on the observation of the outdoor fog-free images of He et al. In most of the image areas other than the sky, at least one color channel has very low intensity in some pixels.

In the case where the image has few texture patterns or the light from the object is blocked by dark fog, it becomes difficult to see the difference only with the dispersion maximum map creation process, and select an image with noises There is also the possibility of doing it. Therefore, when the variance value of the color image patch falls below a certain threshold value, an image that takes the maximum value in transmission is selected. This is because there is a high possibility that good results are obtained because direct light J from the object becomes large where transmission value is high.

For the dispersion maximum map creation process and the transmission maximum map creation process, a fog removal effect can be expected in the case where there is unevenness even by single process alone. However, there is a possibility of selecting an image including noise only by the dispersion maximum map creation processing, and when the fog becomes dark by only the transmission maximum value calculation processing, a dark portion in the image is found And it is possible that the situation that defogging processing using the dark channel prior is difficult to be established is possible. Therefore, by looking at the threshold value of the dispersion value, it is selected whether to use the dispersion maximum map or the transmission maximum map, and combines these two processes to create an integrated image (Figure 2).

Results and Discussion

In this research, we proposed a method to remove fog from uneven defogging image with fog density change, using dark channel prior and local gradient key point extraction technology. Conventional fog removal is based on the assumption that the fog is uniform in the space, and it was not able to cope with the state with the concentration change. In the proposed method, the dispersion and transmission are combined using the color image, and images with less fog shading
are connected well, and the image visibility is improved as a whole. As a result, it was possible to remove fog with shading which was difficult to take with only one piece. Estimate the transmission map by using the image local information and shaping the coarse map using the detailed map. In this way, a clear image is obtained while maintaining the fog removing quality.

Acknowledgment
The research was supported by the funds of the following. Inner Mongolia University of Technology Scientific Research Fund (ZD201417), Inner Mongolia Autonomous Region College of Science and Technology Research Project (NJZY16094) and China Scholarship Council (201608155034) and the National Natural Science Foundation of China (61762070).

References