A Method of Automatic Detection of Fog Image Based on SVM Classification

Hesheng Shi*, Qiming Wang, Liming Xie

College of Information Engineering, Pingdingshan University, Pingdingshan Henan 467002, China

*shsworld@163.com

Abstract

The prerequisite of intelligent image de-fogging is to detect whether fog exists in the image. In this paper, a method to detect fog image from outdoor natural images is proposed. The method normalizes the RGB components and brightness of the image to acquire the minimal correlation between the image intensity and the external light intensity. Then, the two-dimensional discrete Fourier transform is used to obtain the spectrum signature of the image. Together with the gray-level co-occurrence matrix of the image, the spectrum signature is used as a classification feature. The author collects a large number of outdoor fog-free colored images, carries out fogging simulation, and obtains the vector feature library of foggy images through training and evaluation. The author also conducts an experiment to detect foggy and fogless natural images. The experiment proves that the said vector feature library is credible and the proposed method has good ability of detecting fog image, which provides a good precondition and feasible method for intelligent de-fogging.

Keywords: support vector machine; spectrum signature; gray-level co-occurrence matrix; fog image detection.

1. INTRODUCTION

Haze is internationally recognized as one of the ten most serious weather disasters. In recent years, the disastrous weather frequently appears in China, exerting negative impact to the transportation system, power system, and production and living (Kerr, 2007). The impact to road transport is particularly serious. Scientists have found that aerosols are the main cause of haze. As aerosols spread far and wide, way beyond the area of wind and air flow, more and more places fall victim to the haze weather (Fattal, 2008).

Due to the scattering of particles like atmospheric gels, the visibility and quality of images taken on a foggy day are more or less affected, which greatly influences the video/image effect and post-analysis and processing (Pedone and Heikkila, 2011). Therefore, many scholars have done in-depth research on image defogging, and put forward many ways to restore and enhance the degraded images taken on a foggy day (Tan, 2008). With these methods, the degraded images become clearer and more suitable for human visual perception (Li et al., 2009). However, it would only make matters worse if fogless images are defogged. Thus, it is necessary to make correct detection of whether fog exists in outdoor natural images, which provides a good precondition for intelligent image de-fogging and features high practical values (Lu and Weng, 2007).
In this paper, the author discloses the frequency-domain features of fog image by analyzing the frequency response of a large number of natural and simulated foggy images. Then, the author normalizes the RGB components and brightness of the images, establishes a vector feature library of the foggy image in light of the spectral signature of normalized images and the gray-level co-occurrence matrix of the images, and verifies the classification method and the vector feature library of foggy images with the outdoor natural images.

2. FOGGY IMAGE SIMULATION

2.1 Precondition of fog image simulation

For convenience, fogless HD outdoor images are divided into several sample images in this paper. The size of sample images is enough to assume that every sample image has approximately the same value of atmospheric light transmission function with each other (Tso and Mather, 2009).

The author collects more than 2,000 fogless HD outdoor natural images from Flickr, Baidu, Google, and Computer Vision Test Images. Since the fogless images are recognized by naked eyes, there might be a little mist or a few particles on some images. However, it is assumed in this paper that the value of atmospheric light transmission function of these fogless images \( t = 1 \) because the \( t \) value is very close to 1 (Landgrebe, 2003). The author then divides the collected fogless images into several sample images, each of which is 128x64 in size, and carries out fogging simulation of these sample images (Richards and Jia, 1999).

2.2 Implementation of fog image simulation

In computer vision and computer graphics, foggy images are typically described with the following model:

\[
I = I_0 t + A(1-t)
\]  

(1)

Where \( I \) is the foggy image, \( I_0 \) is the fogless image, \( t \) is the atmospheric light transmission function \( t = e^{-\beta d} \) (the \( t \) value describes how much of the light is not scattered and is transmitted to the camera), \( A \) is the global atmospheric light. Thus, the foggy image \( I \) is available as long as \( I_0 \), \( t \) and \( A \) are known.

According to the atmospheric light transmission theory and formula (1), the influence of fog on image quality is mainly related to the depth of field \( d \) and fog scattering coefficient \( \beta \). Since the depth of field and fog density can be integrated to the atmospheric light transfer function, it is possible to simulate foggy images with different fog densities by changing the value of atmospheric light transmission function of an image.

![Simulated fog images](image)

**Figure 1.** Simulated fog images
In fact, the human eye can barely note the fog in photos taken when \( t > 0.8 \). Therefore, the value of atmospheric light transmission function \( t \) randomly obtained to implement fog image simulation falls in the range of \([0.2, 0.8]\). See Figure 1 for the simulated foggy images with different \( t \) values.

Figures 1(a) to 1(f) are simulated fog images with the value of transmission function set as 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and 0.8 respectively, and Figure (h) is the fogless original image.

3. FREQUENCY-DOMAIN FEATURES OF FOG IMAGE

3.1 Brightness normalization

Due to the intensity difference of external lights, the brightness often varies greatly from image to image even if the images are taken outdoors in daytime. If the images are directly analyzed by RGB color space, the results are definitely affected by light intensity (Jia and Richards, 1999). So, this paper pre-treats the RGB color space of the images by formula (2) and obtains \( r, g, \) and \( b \), respectively the equivalent value of \( R, G \) and \( B \).

\[
\begin{align*}
    r &= R / (R + G + B) \\
    g &= G / (R + G + B) \\
    b &= B / (R + G + B)
\end{align*}
\]  

(2)

The features extracted in this paper are more independent because \( r, g, \) and \( b \) are obviously irrelevant to the changing light intensity of the scene. Then, the author normalizes the obtained rgb space by formula (3).

\[
I(x) = \frac{\max(y(i(y)) - i(y))}{\max(y(i(y)) - \min(y(i(y))) \times 255}
\]  

(3)

Where, \( I(x) \) is the image in the \((r', g', b')\) space after the normalization.

3.2 Frequency-domain feature clustering

3.2.1 Frequency-domain features of outdoor natural images

![Sample images of different atmospheric light transmissions](image)

**Figure 2.** Sample images of different atmospheric light transmissions
Outdoor images are stored as two-dimensional matrices in the spatial domain in computer. As target information, background information, haze information and noise information are entangled with each other, it is extremely difficult to isolate such information directly from the spatial domain for analysis (Boardman and Kruse, 1987). As a result, the author carries out frequency-domain spectrum analysis for the 2,000 fogless, simulated foggy and natural foggy images gathered for this paper (Green et al., 1988). Figure 2 contains a series of sample images divided from a natural foggy image taken outdoors. It can be distinguished with naked eye that the value of atmospheric light transmission gradually increases from the left to the right. To verify the observation, the author obtains the mean value of atmospheric light transmission by the method in the reference. It is calculated that the mean atmospheric light transmission of Figures 2(a) to 2(e) are 0.1422, 0.1640, 0.1765, 0.1313, and 0.5531, respectively.

![Sample Images](image)

**Figure 3.** The corresponding spectrum signatures of the sample images

Figure 3 is the corresponding spectrogram of Figure 2. Comparing the two figures, the author discovers the spectrum signature pattern of fog images, i.e. the lower the value of atmospheric light transmission, the more frequently the frequency-domain feature is expressed, the lower the spectral energy, and vice versa.

![Spectrograms](image)

**Figure 4.** Spectrograms of simulated images with different fog densities

The said pattern also applies to simulated fog images. The author also studies the frequency-domain features of simulated fog images with different values of atmospheric light transmission. It is found that the pattern of frequency-domain features is consistent with that of outdoor natural images. Figure 4 is the spectrogram of simulated fog images with different fog densities, which corresponds to Figure 1. Figure 4(h) is the spectrogram of a fogless image, while Figures 4(a) to 4(h) are the spectrograms of fog images in which the fog density is successively decreased.

In short, the smaller the value of the atmospheric light transmission function, the denser the fog, the higher the frequency of the image, and the lower the spectral energy, and
vice versa. Therefore, frequency-domain features can be used to judge if there is haze in outdoor images.

Fourier transform is the most commonly used method to extract frequency-domain features from images. In this paper, two-dimensional discrete Fourier transform is used to extract the frequency features of target images. For an $M \times N$ image, the two-dimensional discrete Fourier transform is performed according to formula (4).

$$I(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} i(x, y) e^{-j2\pi \frac{ux}{M} \frac{vy}{N}}$$

(4)

Where $u$ and $v$ are image transformation frequency variables, $x$ and $y$ are spatial variables of the image, and $I(u, v)$ is the Fourier transform of the image. In order to reduce the influence of the visible edges of the image on the results, the author filters the low frequency signals in the image with the Gaussian high-pass filter.

**3.2.2 Frequency-domain feature clustering**

If the frequency-domain information of the Fourier transform results is taken as the classification feature, the performance of the classifier is under the direct influence of the high feature dimension. Besides, there are many elements with close or zero eigenvalues, which have very little influence on this method (Benediktsson et al., 2003). Thus, the feature dimension can be reduced by clustering. In this paper, an improved ISODATA clustering method is used to obtain the spectral density feature (Schechner et al., 2003).

The author takes the frequency-domain eigenvalues obtained by formula (4) as a finite sample group $F = \{f_1, f_2, ..., f_N\}$, where each sample $f_n$ is an eigenvalue of the image frequency-domain, and uses the following fuzzy classification matrix to divide the $N$ samples into $K$ classes ($2 < K < N$):

$$U_k = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_K \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{K1} & \mu_{K2} & \cdots & \mu_{KN} \end{bmatrix}$$

(5)

The spectral density feature is:

$$Z_{ij} = \frac{\sum_{k=1}^{N} (\mu_{ik})^y F_{kj}}{\sum_{k=1}^{N} (\mu_{ik})^y}$$

(6)

Where $i \in [1, K]$, $j \in [1, S]$, and $S$ is the number of samples with mean index value of a class of samples.

Figure 5 shows the corresponding spectral density features of Figure 4. The discrete spectral energy points in Figure 4 are converted to continuous spectral densities in Figure 5. It is meaningless to seek the value of discrete spectral density at a point is meaningless. Instead, it is meaningful to seek the area below the spectral density
function within a small segment near the point for the area indicates the amplitude of haze signal at the frequency point, i.e. the haze density.

Figure 5. Spectral density feature

4. GRAY-LEVEL CO-OCCURRENCE MATRIX

The images of the sky, the ground and glass curtain walls on buildings feature slow-changing texture and similar frequency-domain feature with fog images. As detection errors may occur due to the over-fitting of the classifier, this paper uses the gray-level co-occurrence matrix to enhance the feature significance.

Proposed by Haralck in 1973, the gray level co-occurrence matrix (GLCM) is a matrix function of pixel distance and angle. With fixed pixel distance and angle direction, the function can be used to determine the correlation between the gray-level of two points in an image. The correlation reflects the general information of the image, ranging from the direction, spacing, amplitude of variation and speed. The properties of the GLCM include contrast, angular second moment (ASM), entropy (ENT) and inverse differential moment (IDM).

Considering that the sample pool of this paper is established by dividing an outdoor natural image into 128x64 image blocks, the author takes 16x8 pixel region as a cell unit and 8 pixels as the moving step length during the extraction of gray-level feature of the image blocks. The checker performs a traversal scan on the image and each cell constructs a sub-feature, which is expressed in formula (7).

\[
\begin{align*}
\mathbf{f}_2 &= \{c, asm, ent, idm\} \\
\end{align*}
\]

Where \( c \) is the contrast attribute, \( asm \) is the angular second-order moment attribute, \( ent \) is the entropy attribute, and \( idm \) is the inverse differential matrix.

In this paper, the author calculates the GLCM by a horizontal spacing of 3 pixels in each cell and obtains relevant attributes. Through the calculation, the author gets 4 attributes of each cell’s GLCM, namely, 4 dimension eigenvectors. On each sample graph, all cells form \( 4 \times (128/8) \times (64/8) = 512 \) dimensional vectors. Using the CLCM to represent the
spatial distribution of texture, the author makes up the deficiency of frequency-domain feature of fog images. Coupled with non-texture features (frequency-domain features of fog images), the texture features of the image are summarized and used to classify the regions in the image. The combined method both fulfills the purpose of detecting haze and reduces the detection error.

5. EXPERIMENT AND RESULTS

The experimental environment consists of Dell CS24-TY, an 8G memory industrial machine running on Intel® Xeon®@2.27GHz CPU and Windows 2008 Enterprise Editor (64-bit). In order to strike a balance between the detection speed and accuracy of the image, the author attempts to experiment with different number of samples. Group 1 has 2,000+ positive and 2,000+ negative samples, Group 2 has 4,000+ positive and 4,000+ negative samples, and Group 3 has 8,059 positive and 8,059 negative samples. Then, the author extracts the features of the three groups of samples and conducts training with the method and tools proposed by Chih-Jen Lin et al. After that, the author classifies the test images with similar results, obtaining 1,967 positive sample images and 1,967 negative sample images.

To compare the performance between different feature sample sizes, the ROC curves of the three sample groups are used to describe the relationship between the true haze image rate and the false haze image rate. The former is denoted by the vertical axis and the latter by the horizontal axis. The two rates change with the feature sample size. Figure 6 is generated by plotting the points with a smooth curve. The area under the curve (AUC) indicates the performance of the corresponding feature sample size. In fog image detection, the closer the AUC value is to 1, the better the detection performance.

Figure 6. ROC curves of fog image detection

As shown in Figure 6, when there are 2,000 positive and 2,000 negative samples, the AUC is about 0.8712; when there are 4,000 positive and 4,000 negative samples, the AUC is about 0.8921; when there are 8,059 positive and 8,059 negative samples, the AUC is about 0.9086. Thus, the more samples in the feature library, the more accurate the judgment if the image contains fog. However, the experiment also shows that: when the number of positive and negative samples both rises from 4,000 to 8,059, the AUC only grows by 0.0374, indicating that the accuracy increases much slower after the sample size reaches a certain point.
Moreover, the author finds that the sample size of the feature library has a very small effect on speed. Table 1 displays the detection speeds of different sample sizes.

6. CONCLUSION

Without relying on any prior condition, this paper judges whether an image or a video frame has fog by LibSVM. Specifically, the author normalizes the RGB components and brightness of the image, uses spectrum signature of the normalized image and the gray-level co-occurrence matrix of the image as classification features, and obtains the vector feature library of the foggy image. The experiment proves that the said vector feature library is credible and the proposed method has good ability of detecting fog image, which provides a good precondition and feasible method for intelligent de-fogging.

7. REFERENCES