ABSTRACT

There is an increasing number of methods for removing haze and fog from a single image. One of such methods is Dark Channel Prior (DCP). The goal of this paper is to develop a mathematical explanation on why DCP works well by using principal component analysis, and minimum volume ellipsoid approximations.

Index Terms— Contrast enhancement, Dehazing, Minimum volume ellipsoid

1. INTRODUCTION

The introduction of the dark channel prior (DCP) by He [1] was a novel approach to solving the difficult problem of removing haze from a single image because the model for representing fog or haze is represented by one equation but at minimum 3 unknowns.

Until a few years ago, multiple images of a static scene with differing amount of haze was required to estimate the scene for the purpose of removing haze or increasing the contrast [2]. Then Fattal [3], Tan [4] both presented unique methods for removing haze using a single image. Each of which used a prior or assumption of the scene to overcome the ill-posed problem. Following that, He [1] used the dark channel prior and Tarel [5] inferred a veiling for single image haze removal (color or grayscale). Then a modification of the DCP was presented in [6] which used a median filtering operation to improve the speed of the method for the purpose of video dehazing.

The DCP does require a refinement step to estimate the scene parameters which requires a large amount of memory and is complex [1]. However the method for generating the DCP is itself simple and fast. The authors have also observed more papers being in submission now that also use the DCP (perhaps due to the simplicity) but unwittingly know exactly why it works. Therefore we present in this paper a geometrical explanation for how the DCP behaves in hazy images and as a by-product present another prior that is similar to the DCP but is based on a geometric understanding.

1.1. Dichromatic Model and RGB Cluster

In order to represent a hazy image \( \hat{x} \) and the non-hazy image \( x \), we will use

\[
\hat{x}(m, n) = t(m, n)x(m, n) + (1 - t(m, n))a,
\]

which is a dichromatic model to represent the presence of fog or haze in an image which is used in \([1, 2, 3, 4, 5, 6]\).

For scenes where fog and haze are present, (1) is used frequently for modeling the contrast degradation caused by the scattering particles in the air. The transmission, \( t(m, n) \), is an exponentially decaying function parametrized by wavelength, distance, position, and time. The common assumptions made are that scattering is homogeneous, independent of color, and stationary in time [2]. Therefore we have the transmission to be \( t(m, n) = e^{-\beta d(m, n)} \), where \( \beta \) is a function of particle size and \( d(m, n) \) is the distance from the camera to the radiant object. The airlight color vector \( a \) is the color of any radiant object viewed over-the-horizon at an infinite distance when scattering is present. The term \( a(1 - t) \) is also called the veiling [5].

To observe the statistical nature of patches of a scene, we can plot each pixel as a 3-D point in the RGB cube [7]. Observe in Fig. 1 the RGB cluster is scaled and shifted when haze is present. This suggests a relation that can be inferred from the position and size of the RGB cluster with regard to the amount of scattering \( t(m, n) \).

2. ANALYSIS OF DCP USING ELLIPSOID APPROXIMATIONS

When viewing a cluster of points within the RGB Histogram we see a geometric relationship to scattering and airlight. In this section, principal component analysis (PCA) or Karhunen-Loeve (K-L) transform will be used to describe how the parameters of the cluster, approximated as an ellipsoid, are affected when haze is present (similar to [7]). The ellipsoid is then formally characterized as a minimum volume ellipsoid [8] based on the cluster of points. Using the properties of the ellipsoid, a new prior called Ellipsoid Prior (EP) is presented. EP is then compared to DCP mathematically and visually.
Fig. 1. Examples of RGB histogram of a hazy image. (a) Original image \( x \) (b) Histogram of grass indicated by the lower yellow square. (c) Histogram of far away trees indicated by the upper red square.

### 2.1. Analysis of Cluster using PCA

Suppose we slide an \( M \times N \) size patch \( \Omega_x(m, n) \) at pixel location \( (m, n) \) across an image of a natural scene and collect a real valued sequence of \( p \times 1 \) color vectors \( x \) within that patch. Let us also assume that this discrete sequence is random with autocorrelation matrix \( R_x \) which is of size \( p \times p \). The K-L transform [9] is then

\[
y = \Phi^T x
\]

\[E[yy^T] = \Phi E[xx^T]\Phi^T = \Phi R_x\Phi^T = \Lambda.\]  \hfill (3)

The elements of the diagonal matrix \( \Lambda \) are in decreasing order and are the eigenvalues of \( R_x \). (\( \Lambda = \text{diag}({\lambda_1, \ldots, \lambda_p}) \).) The \( p \times p \) matrix \( \Phi \) is orthonormal.

Now suppose the image is degraded by haze and let us assume the scattering \( t(m, n) \) within the patch \( \Omega_x \) (dropping \( (m, n) \) for simplicity) is deterministic (but unknown) and constant within the patch, \( t(m, n) = t, \forall (m, n) \in \Omega_x \). Using (1) we have the autocorrelation matrix of the hazy patch,

\[
R_{\hat{x}} = E[(\hat{x} - E[\hat{x}]) (\hat{x} - E[\hat{x}])^T] \\
= E[\hat{x}\hat{x}^T] - E[\hat{x}]E[\hat{x}]^T \\
= t^2 (E[xx^T] - E[x]E[x]^T) = t^2 R_x,
\]

which is a scaled version of the non-hazy patch autocorrelation matrix. The eigenvalues of the hazy patch are also scaled versions of the non-hazy patch,

\[
t^2 R_x = t^2 (\Phi^T \Lambda \Phi) = \Phi^T (t^2 \Lambda) \Phi = \Phi^T \hat{\Lambda} \Phi,
\]

with \( t^2 \Lambda = \hat{\Lambda} \). (\( \hat{\Lambda} = \text{diag}(t^2 \lambda_1 = \hat{\lambda}_1, \ldots, t^2 \lambda_p = \hat{\lambda}_p) \).) Note that the transmission \( t \) acts as an attenuator,

\[
t = e^{-\beta d} \Rightarrow 0 \leq t \leq 1.
\]

The expected value of the hazy patch is a vector sum of the veiling and the scaled non-hazy mean color vector,

\[E[\hat{x}] = t E[x] + (1 - t) a.\]  \hfill (7)

This shows the mean of the hazy patch, \( E[\hat{x}] \), shifts from \( E[x] \) to the airlight vector \( a \) as \( t \rightarrow 0 \), or \( d \rightarrow \infty \).

Based on (4, 5, 6, 7) we can expect to see two characteristics of the ellipsoids: As \( t \rightarrow 0 \), the mean vector or centroid moves closer to the airlight vector \( a \) and the principal eigenvalues decrease therefore shrink the size of the ellipsoid.

### 2.2. RGB Cluster Approximation

Instead of manipulating the estimated PCA parameters to grow the ellipsoid so that the sample set \( \Omega_x \) is covered as mentioned in [10], we will use the minimum volume ellipsoid or Löwner-John ellipsoid of this set [11]

\[E_{Ij} = \{ x \mid \hat{x} T A^{-1} (x - b) \leq 1 \} \]

where the finite set \( \Omega_x(m, n) = \{ x_1, \ldots, x_p \} \subseteq \mathbb{R}^p \) contains the color vectors \( x \) within an \( M \times N \) neighborhood of pixel location \( (m, n) \). The symmetric, positive definite matrix \( A \in \mathbb{R}^p \times p \) and \( p \) offset vector \( b \) are used to parameterize \( E_{Ij} \). Therefore we desire an ellipsoid \( E_{Ij} \) that covers the convex hull of the set (or patch selected from image) \( \Omega_x(m, n) \). The objective and constraints to enforce a minimum volume ellipsoid are

\[
\text{minimize } \log \det A^{-1} \\
\text{subject to } \sup_{x \in \Omega_x(m, n)} \|Ax + b\|_2 \leq 1.
\]

We then solve (9) using Khachiyan’s algorithm [8]. Due to the properties of \( A \), we can decompose and obtain the eigenvalue/eigenvector pairs of its inverse

\[A^{-1} = V^T D V\]

with \( D \) as the diagonal matrix with decreasing eigenvalues \( diag(d_1, \ldots, d_p) \) and eigenvectors \( V = (v_1, \ldots, v_p) \).

Let’s take a step back and look at an example of approximating an ellipsoid using the R-G plane so that illustration is easier. In Fig. 2 there are three clusters in the R-G plane. The upper right cluster is the color points from the distant trees indicated by the upper red box in Fig. 1(a). This cluster is very compact and the centroid is almost on top of the airlight vector \( a \) indicated by the blue arrow. The next cluster is from the yellow tree in Fig. 1(a) in the left half of the image indicated by the dark green square. This is a narrower but longer ellipse and the centroid is farther away from \( a \). And finally the grass cluster is the farthest from \( a \) (closest to camera) and the size of the ellipse is larger. This illustrates what we expected to see as the distance increases 1) the centroid should be closer to airlight vector 2) the size of the ellipse should condense.
2.3. Proposed Transmission Estimation using a Löwner-John Ellipsoid

In the previous section, we mentioned that the size and position of the approximated ellipsoid gives a queue to the value of transmission. Therefore we will generate an Ellipsoid Prior (EP) \( \theta_E : \mathbb{R}^{M \times N \times p} \to \mathbb{R} \), that is parameterized by the properties of the Löwner-John ellipsoid from the set \( \Omega_x(m, n) \):

\[
\theta_E(m, n) = ||g|| / ||f||, \quad g = b - \sqrt{d_1}v_1 \tag{11}
\]

with \( b, d_1, \) and \( v_1 \) defined in (9) and (10). The term \( ||f|| \) is a scaling value where the vector \( f \) represents the largest possible vector value within the \( p \) channel color cube. For \( x(m, n, \lambda) \in [0, 1], ||f|| = \sqrt{p} \). This enforces \( \theta_E \in [0, 1] \). The vector \( g \) in (11) points to the closest endpoint (on the principal axis) of the ellipsoid from the origin. For example, the \( g \) vector for the grass cluster is indicated by a red arrow in Fig. 2 labeled \( g_1 \). With \( p \times 1 \) vector of all ones \( 1_p \), note that if \( b \times 1_p = v_p \times 1_p = 0, g \) is the minimal vector to the ellipsoid from the origin.

3. DARK CHANNEL PRIOR AND LÖwner-John ELLIPSOID RELATIONSHIP

The concept of the DCP used for single image dehazing methods is constructed as [1]

\[
\theta_D(m, n) = \min_{k, l \in \Omega_x(m, n)} \left( \min_{c \in \{r, g, b\}} \frac{\hat{d}(k, l, c)}{a(c)} \right). \tag{12}
\]

Thus the DCP is approximately the minimum distance to the Löwner-John ellipsoid either from the R-G, G-B, or R-B planes,

\[
\theta_D(m, n) \approx \begin{cases} \arg\min_{c\neq z} ||z_c - y||_2 \text{, subject to} \\ z_c^T e_c = 0, \\ y^T \mathbf{A}^{-1} y = 1, \end{cases} \tag{13}
\]

with equivalency when \( z_c \in \Omega_x(m, n) \). The unit vector \( e_c \) represents the normal to one of the three color planes within the RGB cube and matrix parameter \( \mathbf{A} \) is from the Löwner-John ellipsoid (9) for the set in \( \Omega_x(m, n) \).

Now if the airlight is colorless, \( a = \alpha 1_p \) for \( 0 < \alpha \leq 1 \), then we have

\[
\theta_D(m, n) = \frac{1}{\alpha} \min_{k, l \in \Omega_x(m, n)} \left( \min_{c \in \{r, g, b\}} \hat{d}(k, l, c) \right). \tag{14}
\]

If we also consider that the principal eigenvector \( v_1 \) and the center \( b \) of the ellipsoid \( E_{ij} \) are both parallel to \( a \) or grayish [7], then the EP is a scaled version of the DCP,

\[
\theta_D(m, n) = \alpha \theta_E(m, n) \tag{15}
\]

More generally, EP is the upper bound of the DCP

\[
\theta_D(m, n) \leq \theta_E(m, n). \tag{16}
\]

We have a lower bound if we compute the transmission

\[
t_D(m, n) = 1 - w \theta_D(m, n) \geq 1 - w \theta_E(m, n) = t_E(m, n), \tag{17}
\]

where the transmission functions \( t_E \) and \( t_D \) are respectively the EP and DCP transmission estimates constructed in the same fashion as in [1] with the weighting value set to \( w = 0.95 \).

3.1. Experimental Comparisons

Fig. 3 demonstrates the relationship between EP and DCP with two hazy images processed with a patch \( \Omega_x \) size of \( 11 \times 11 \). Fig. 3(a) has the airlight very close to gray and results in EP and DCP being very similar in appearance. The image in 3(d) was chosen because the airlight is more orange hence farther away from the gray line. Therefore the upper-bound relationship (16) is evident by viewing regions in the EP image are brighter than in the same regions in the DCP image.

The idea that the EP is the upper bound breaks down when the patch \( \Omega_x \) is near an occlusion boundary because the transmission function of depth is smooth is invalid. This is evident by observing that the objects in EP image Fig. 3(f) appear fatter than the objects in the DCP image Fig. 3(e).

4. CONCLUSION

We presented a new prior, EP, that is based on approximating an RGB cluster with a minimum volume or Löwner-John...
Fig. 3. (a) Original image. (b) DCP of image in (a). (c) Image prior using $\theta_E$ (11). (d) Image from [1]. (e) DCP of image in (d). (f) Image prior using $\theta_E$ (11). The DCP and EP images are very similar where the upper bound nature of $\theta_E$ can be seen at the left of the yellow trees in the top row image set and on top of the bushes close to the camera in the bottom row image set.

5. REFERENCES


