IMAGE DEHAZING USING CONVEX OPTIMIZATION

Aniket Ajay Sahoo
Department of ECE
Virginia Tech
Blacksburg, VA, USA
aniketsahoo@vt.edu

***Abstract*—Haze caused by pollutants affects the quality of outdoor images, causing loss of intensity and quality. Image dehazing involves recovering images by removing haze. However, even with recent approaches making great strides in quality, the complexity has remained high. This paper presents dehazing as a convex optimization problem with the goal of reducing complexity while not losing out on image quality.**

I. Introduction

Haze is commonly described as an atmospheric phenomenon caused by pollutants in the sky. Typically, it is caused by dust, smoke and accumulation of fine suspended dry particles like salt and soot. Haze obscures the sky, and in many cases, normal vision. Thus, images taken in the presence of haze seem qualitatively unclear, necessitating the process of dehazing them. Image dehazing has numerous applications in fields like computer vision, facial recognition, tomographic image processing, geographical information systems, surveillance systems, object detection, facial recognition etc. Thus some amount of research has gone into it - initially it was treated as an image processing problem, however, thanks to the recent developments in more advanced fields like artificial intelligence and the need for better computer vision tasks, several strides have been made in this department. Unfortunately, since the task is restricted to every individual image, it is complex and under constraint. Thus it struggles from the quandary of quality vs. complexity. This paper will compare existing methods of dehazing and present a technique as a result of viewing dehazing as a convex problem.

II. Existing Work

A. Feature Extraction using Regression

Proposed by Luan, Zhong et al., a learning framework is used to obtain a dehazed image[1]. The algorithm has two main parts, training and dehazing. The training involves generating numerous high quality image bits in various textures and colours while analyzing several features like contrast, histogram equalization, saturation etc. Transmission map estimation is then obtained using Support Vector Regression (SVR), which essentially combines regression with Support Vector Machine classifiers. SVR allows the authors to define the amount of acceptable error in the model, and will then determine hyperplanes to best fit the data given these error constraints. Finally, the input hazed image is categorized into smaller patches in order to reduce input dimensions and computational complexity. This method presents low complexity and strong adaptability with reasonable dehazing quality.

B. Dehazing using Scene Depth Information

Jiang, Bo et al. proposed decomposing the image into three layers of high frequency and one layer of low frequency using wavelet decomposition[2]. Different frequency layers preserve different information about the image - the low frequency layer contains lesser noise with a smooth boundary while the high frequency layers contain the boundary information of the original image. The layers are then individually enhanced and recombined. Image enhancement is done using modified guide image filtering, and by using proportion of the scene depth, haze free image is obtained as the enhancement filter is higher for scenes that are further away and lower for those nearby. The haze free image is produced after wavelet reconstruction. This method is great for retaining image features and delivers fairly high quality dehazed images, however it performs very poorly at when the focus of the image is further away.

C. Dark Channel Prior

Proposed by K. He, J. Sun and X. Tang, they observed that most local patches in haze-free images contain at least one color channel with low intensities[3].

$$J^{dark}(\mathbf{x}) = \min_{c\in\{r,g,b\}} \left(\min_{\mathbf{y}\in\Omega(\mathbf{x})} (J^c(\mathbf{y})) \right),$$

In the above equation, J^c is the color channel of J and $\Omega(x)$ is the local patch centred at x. A dark channel is the outcome of two minimum operators where the first $min_{c\in\{r,g,b\}}$ is performed on each pixel and the other $min_{y\in\Omega(x)}$ is a minimum filter. Initially, dark channel prior is computed. Then, atmospheric light is estimated using the brightest pixel. Transmission distribution is estimated after, using small patches where transmission is estimated to be constant. Finally, a refined transmission map is calculated using soft matting before resulting in a recovered dehazed image. While this method delivers high quality dehazed images, it is quite slow and complex relatively speaking to other methods, and also works very poorly when objects are similar in intensity to atmospheric light.

III. Problem Statement

Besides the work mentioned beforehand, there are even more, however, conventional methods of image dehazing are either computationally expensive and take too much time, or reduce the image quality in favour of reducing complexity. Thus, when structured as an optimization problem, the goal is a technique that solves the dehazing problem with reduced computation time and workload while maintaining the quality.

IV. Experimental Analysis

A. Hazed Image Model

Image dehazing is a non-convex problem because of the bi-linearly coupled haze free image and the atmospheric light transmission distribution.

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

The above equation has two parts - $J(x)t(x)$ represents reflected light from the object surface, while $A(1-t(x))$ represents scattering transmission, which in our case is haze. $I(x)$ is the observed intensity, $J(x)$ is the scene radiance, A is the global atmospheric light and $t(x)$ is the medium transmission. I, J and A are RBG triplets while $t(x)$ is a scalar that describes how radiance of a point is attenuated according to it's distance from the observer. Clearly the coupling term $J(x)$ and $t(x)$ with J, t and A all acting as unknowns is a typical non-convex problem. However, $J(x)t(x)$ as a single term is bi-linear. Based on this fact, this problem can be converted into a convex optimization problem.

In the RGB color space, if c represents the color channel index, then I_c and J_c become $\mathrm{M} \times \mathrm{N}$ non-negative matrices, then the model becomes:

$$I_c = J_c \cdot t + a_c(1 - t), c = 1,2,3$$

where a_c still represents the atmospheric light constant, however depending on the value of c is that of the corresponding color channel. 1 here represents a matrix as well, with all its entries being 1. Thus, the hazy image I_c is obtained from the attenuated image density J_c through the scattering transmission path, alongside the atmospheric light.

Fig.1. (a)Hazy Image, (b)Transmission, (c) Dehazed Image

Figure 1: (a)Hazy Image, (b)Transmission, (c) Dehazed Image

B. Discrete Haar Wavelet Transformation

- Thanks to Jiang, bo et al., we know that images can be decomposed into three layers of high frequency and one of low frequency[2]. Since the high frequency layers mostly contain boundary information in different directions, the haze should almost entirely be distributed in the low frequency layer alongside image details.

- Keeping this in mind, we used Discrete Haar Wavelet Transformation (DHWT) to decompose the image into it's frequency sub-bands. DHWT is used because it is very quick, which aligns well with our aim.

- DHWT transforms I_c and J_c from $\mathrm{M} \times \mathrm{N}$ matrices to having four $\frac{M}{2} \times \frac{N}{2}$ sub-band blocks, with each block representing a different frequency layer[4].

  $$\hat{I}_c = W I_c W^T = \begin{bmatrix} \hat{I}_c^a & \hat{I}_c^h \\ \hat{I}_c^v & \hat{I}_c^d \end{bmatrix}$$
  $$\hat{J}_c = W J_c W^T = \begin{bmatrix} \hat{J}_c^a & \hat{J}_c^h \\ \hat{J}_c^v & \hat{J}_c^d \end{bmatrix}$$

  where a, h, v and d represent low-frequency, horizontal, vertical and diagonal sub-bands after wavelet transformation.

- Looking at the image models on a sub-band level, we have

  $$\begin{aligned} \hat{I}_c^a &= \hat{J}_c^a \cdot \hat{t}^a + \hat{a}_c(1 - \hat{t}^a) \\ \hat{I}_c^h &= \hat{J}_c^h \cdot \hat{t}^a \\ \hat{I}_c^v &= \hat{J}_c^v \cdot \hat{t}^a \\ \hat{I}_c^d &= \hat{J}_c^d \cdot \hat{t}^a \end{aligned}$$

  since the high-frequency sub-bands are free of haze, it only applies to the low-frequency sub-band. $\hat{t}^a$ is a scalar and is constant for all bands.

- Putting all of this together, our modified image model after DHWT becomes

  $$\begin{bmatrix} \hat{I}_c^a & \hat{I}_c^h \\ \hat{I}_c^v & \hat{I}_c^d \end{bmatrix} = \begin{bmatrix} \hat{J}_c^a \cdot \hat{t}^a + 2a_c(1 - \hat{t}^a) & \hat{J}_c^h \cdot \hat{t}^a \\ \hat{J}_c^v \cdot \hat{t}^a & \hat{J}_c^d \cdot \hat{t}^a \end{bmatrix}$$

- Since our goal is to recover the haze-free image $\hat{J}_c$, we need to recover $\hat{J}_c^a, \hat{J}_c^h, \hat{J}_c^v$ and $\hat{J}_c^d$. However, we can recover the high-frequency coefficients by

  $$\begin{aligned} J_c^h &= \hat{I}_c^h \div \hat{t}^a \\ J_c^v &= \hat{I}_c^v \div \hat{t}^a \\ J_c^d &= \hat{I}_c^d \div \hat{t}^a \end{aligned}$$

  Thus, we only need to recover $\hat{J}_c^a$ and $\hat{t}^a$ to find the haze-free image $\hat{J}_c$.

- This means we only need to use our dehazing algorithm on the low frequency sub-band block, as the other coefficients are only dependent on recovering $\hat{t}^a$. This reduces the dimension of our problem from the original $M \times N$ of $\hat{J}_c$ to the dimension of the sub-band block, which we established to be $\frac{M}{2} \times \frac{N}{2}$!

- Inverse discrete haar wavelets recovers the original dimensions of $\hat{J}_c$ once $\hat{J}_c^a, \hat{J}_c^h, \hat{J}_c^v$ and $\hat{J}_c^d$ are found.
- Unfortunately, the coupling term of $\hat{J}_c^a$ and $\hat{t}^a$ is still non-convex and computationally expensive because it is a bilinearly coupled problem.

C. *Formulating Convex Optimization Problem*

- To recover the dehazed image, we must estimate $\hat{J}_c^a$ and $\hat{t}^a$. This means that in the coupling term $\hat{J}_c^a \cdot \hat{t}^a$, both are unknowns - this is a typical non-convex problem.
- To formulate this problem as a convex problem, we first make the following substitutions:

$$\hat{Y}_c^a = \hat{I}_c^a - \hat{a}_c 1$$
$$\hat{Q}_c^a = \hat{J}_c^a \cdot \hat{t}^a$$

- After considering these substitutions, our sub-band model now becomes

$$\hat{Y}_c^a = \hat{Q}_c^a - \hat{a}_c \hat{t}^a, c = 1,2,3$$

This is now linear if we solve for $\hat{Q}_c^a$ and $\hat{t}^a$.

- In the above equation, for estimated airlight $\hat{a}_c$, we will simply assume that it is equivalent to the brightest pixel in I_c after filtering like in [5]. While there are more accurate methods to do this, it will further slow down the dehazing for little gain, so we will go with the simple method. We will use a 3x3 min-filter for this. Thus we have

$$\hat{a}_c = \max\min I_c(m,n), c = 1,2,3$$

- Now that we know $\hat{a}_c$, we also know $\hat{Y}_c^a$ since we already know $\hat{I}_c^a$. Thus, we only need to estimate $\hat{Q}_c^a$ and $\hat{t}^a$.
- Once we successfully estimate $\hat{Q}_c^a$ and $\hat{t}^a$, we can further estimate the following:

$$J_c^a = \hat{Q}_c^a \div \hat{t}^a$$

- Once J_c^a is obtained, we can also obtain $\hat{J}_c^h, \hat{J}_c^v$ and $\hat{J}_c^d$ as mentioned previously thanks to estimating $\hat{t}^a$ and then use inverse DHWT to obtain $\hat{J}_c$.
- Thus, convex optimization on the reformulated model estimating $\hat{Q}_c^a$ and $\hat{t}^a$:

$$min_{\hat{Q}_c^a,\hat{t}^a} \sum_{c=1,2,3} \left(\left\| \hat{Y}_c^a - \hat{Q}_c^a + \hat{a}_c \hat{t}^a \right\|_2^2 \right) + R(\hat{t}^a, \widehat{Q}_c^a, c = 1,2,3),$$
$$\text{s.t } 0 < \hat{t}^a \leq 1, 0 \leq \hat{Q}_c^a, c = 1,2,3$$

where $R(\hat{t}^a, \widehat{Q}_c^a, c = 1,2,3)$ denotes a convex regularization function.

- For R, we want a regularization function that will reduce haze. Thus we want to affect a quantity in such a way that haze is reduced. Contrast is a good quantity to do this as it can be quantified[6], haze reduces contrast of images and dehazing actively involves enhancing contrast.
- From [6], we can represent mean square contrast of J_c by

$$C_{ms} = \sum_{c=1,2,3;m,n\epsilon\Omega_a} (\hat{J}_c^a(m,n) - \bar{J}_c^a)^2 / N_{\Omega_a}$$
$$C_{ms} = \sum_{c=1,2,3;m,n\epsilon\Omega_a} \frac{(\hat{I}_c^a(m,n) - \bar{I}_c^a)}{(\hat{t}^a(m,n))^2 N_{\Omega_a}}$$

where Ω_a is 2D pixels index domain, $\bar{J}_c^a$ and $\bar{I}_c^a$ are average pixel of $\hat{J}_c^a$ and $\hat{I}_c^a$ respectively and N_{Ω_a} is the total number of pixels.

- Clearly, C_{ms} is inversely proportional to $(\hat{t}^a(m,n))^2$. As such, we can use $\left\|\hat{t}^a\right\|_2^2$ as our regularization function.

$$R(\hat{t}^a, \widehat{Q}_c^a, c = 1,2,3) = \left\|\hat{t}^a\right\|_2^2$$

- After adding this regularization function our problem becomes

$$min_{\hat{Q}_c^a,\hat{t}^a} \sum_{c=1,2,3} \left(\left\| \hat{Y}_c^a - \hat{Q}_c^a + \hat{a}_c \hat{t}^a \right\|_2^2 \right) + \left\|\hat{t}^a\right\|_2^2$$

V. METHODOLOGY

- On the input image I_c, we first perform DHWT.
- Then, we estimate atmospheric light constant $\hat{a}_c$.
- Reformulated convex optimized haze image model is solved.
- Inverse DHWT is applied to recover hazefree image.

VI. RESULTS

A. *Expected Results*

There are really only two factors to evaluate when it comes to image dehazing - complexity and quality. While complexity can be estimated and compared via runtime, quality of dehazing is trickier to evaluate. After all, it is subjective since ground truth is often unobtainable for hazed images. However, we can make this an objective comparison by comparing the pixel differences of a dehazed image with the ground truth. We will use Mean Square Error for this purpose alongside an image that we will purposely haze and use as a case study. In the equation below, $M \times N$ represent image dimensions, $f(i,j)$ is the reference image and $f'(i,j)$ is the dehazed image.

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f(i,j) - f'(i,j)]$$

We will run this same image through the Dark Channel Prior setup[3] and then compare the results, both subjectively and objectively. Seeing as our method uses DHWT and uses a simplistic approach to estimate airlight, with the purpose to save runtime, I fully expect our method to be faster - however, dark channel prior is likely to be qualitatively better.

B. Evaluation

1) Quality

As seen in fig.2 with brighter colors, the dark channel prior subjectively looks slightly more vibrant, and just better. With darker colors, as seen in fig.3, the results are actually extremely comparable, with the former only looking slightly better. The MSE comparison in table 1 simply affirms this. However, one major flaw in the proposed method is the pixels don't look as continuous as DCP's results. This is likely a result of the simplistic regularization function.

Fig. 2. (a) Original Image, (b) Hazed Image, (c) Dark Prior Result, (d) Our method

Fig. 1		Fig. 2	
DCP	Our Method	DCP	Our Method
0.54	0.67	0.88	0.96

Table 1. Mean Square Error Comparison

2) Speed

The runtimes in table 2 are the average of the time it took to dehaze 8 images. The results were extremely consistent. As you can see, our method, although simplistic, pays dividends as far as speed goes, and is almost fifteen times faster than DCP.

Runtime	DCP	Our Method
(in seconds)	1.1	14.9

Table 2. Runtime comparison

Fig. 3. (a) Original Image, (b) Hazed Image, (c) Dark Prior Result, (d) Our method

VII. CONCLUSION

While quite a few algorithms can now be used for dehazing, they all struggle with the same quandary of quality vs. complexity. Treating dehazing as a convex optimization problem allowed us to build a systematic, efficient and extremely quick algorithm - however, while the images are dehazed successfully, they could definitely be better qualitatively. This showed both objectively and subjectively when this algorithm was compared to DCP. However, the speed of the algorithm also really shone through, running at about fifteen times as fast as DCP. As such, for the purpose of time consuming tasks, for example, dehazing multiple images, our method is likely superior. However, if quality is the focus, then the DCP definitely comes through.

VIII. FUTURE WORK

Given the results, it is clear that our algorithm works as fast as expected - however, the results, while acceptable, aren't perfect. As such, improving quality would be the focus of future work. Two things come to mind immediately to improve quality - better methods to estimate airlight $\hat{a}_c$ and a more complex regularization function to display better continuity between related pixels. Finally, a parameter to control dehazing intensity could be great so that images with lesser haze would see even faster results while images with heavy haze could get qualitatively better dehazing, at the cost of more time.

IX. REFERENCES

- [1] Zhong Luan, Yuanyuan Shang, Xiuzhuang Zhou, Zhuhong Shao, Guodong Guo, and Xiaoming Liu. 2017. Fast single image dehazing based on a regression model. Neurocomput. 245, C (July 2017), 10–22. DOI:https://doi.org/10.1016/j.neucom.2017.03.024
- [2] Jiang, Bo & Wanxu, Zhang & Zhao, Jian & Ru, Yi & Liu, Min & Ma, Xiaolei & Chen, Xiaoxuan & Meng, Hongqi. (2016). Gray-Scale Image Dehazing Guided by Scene Depth Information. Mathematical Problems in Engineering. 2016. 1-10. 10.1155/2016/7809214.
- [3] Kaiming He, Jian Sun and Xiaoou Tang, "Single image haze removal using dark channel prior," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1956-1963, doi: 10.1109/CVPR.2009.5206515.
- [4] H. Kanagaraj and V. Muneeswaran, "Image compression using HAAR discrete wavelet transform," 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), 2020, pp. 271-274, doi: 10.1109/ICDCS48716.2020.243596
- [5] R. Tan: Visibility in bad weather from a single image[C]. CVPR, 2008:1-8
- [6] Kim, J. H., Jang, W. D., Sim, J. Y., & Kim, C-S. (2013). Optimized contrast enhancement for real-time image and video dehazing. Journal of Visual Communication and Image Representation, 24(3), 410-425. https://doi.org/10.1016/j.jvcir.2013.02.004
- [7] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: Nonlinear Phenomena 60 (1) (1992) 259–268