Estimating Depth and Global Atmospheric Light for Image Dehazing Using Type-2 Fuzzy Approach

Harsh Raj
Data Science and Artificial Intelligence
210150011
harsh.raj@iitg.ac.in

***Abstract*—This article introduces a new single image dehazing method based on a Type-2 membership function with a similarity function matrix. The approach starts by estimating the depth map and global atmospheric light from the hazy image. The estimated depth map is refined to create a true scene transmission map, which, combined with the global atmospheric light, enables effective dehazing via the atmospheric scattering model. To evaluate performance, quantitative comparisons are conducted on the REalistic Single Image DEhazing (RESIDE) dataset, using peak signal-to-noise ratio (PSNR) as the primary metric. Results show that the proposed method achieves higher image clarity compared to other dehazing approaches, effectively restoring natural image quality.**

Index Terms— Single image dehazing, Type-2 membership function, depth map, global atmospheric light, scene transmission, PSNR.

I. Introduction

Outdoor computer vision algorithms often face challenges due to haze or fog, which reduces image clarity and visibility. This is caused by light scattering from atmospheric particles, such as aerosols, which dim and whiten the image. As a result, haze removal or *dehazing* has become essential for enhancing image quality in such conditions.

Dehazing methods are generally divided into two main categories:

- **Multiple Image Dehazing**: Early methods used multiple images with varying views or conditions to estimate clear images. For example, Schechner et al. proposed a polarization-based approach to capture two images with different polarizations, leveraging the fact that airlight scattering is partially polarized. Although effective, these methods require multiple captures, making them unsuitable for real-time applications.
- **Single Image Dehazing**: More recent approaches work with only one image, making them more practical. These methods estimate depth or transmission maps to restore visibility. One popular method is the *Dark Channel Prior* (DCP) by He et al., which relies on the assumption that in most clear outdoor images, the darkest channel in a local region often has very low intensity. Although DCP performs well, it struggles in sky regions where this assumption fails. To address this, various deep learning approaches have been introduced, such as AOD-Net and GFN, which use Convolutional Neural Networks (CNNs) to enhance images directly. These CNN-based methods provide effective results with improved edge clarity and contrast, though they require substantial data for training.

Overall, single image dehazing, while challenging, has gained popularity for real-time applications due to its ability to work with limited information and enhance visibility effectively.

A. Motivation

Images with bad quality can really mess up the performance of vision-based algorithms, especially in hazy conditions, so enhancing these images becomes super important. Basically, how an image looks depends on the light transmission reaching the observer, but haze gets in the way. This is due to *attenuation* (when transmission drops) and *airlight* (which brightens things up) because of haze particles scattering light.

The hazy image you see is actually a mix of attenuation and airlight:

$$O_H(m,n) = O_1(m,n) + O_2(m,n) \tag{1}$$

where $O_1(m,n)$ (attenuation) and $O_2(m,n)$ (airlight) are given by:

$$O_1(m,n) = O_J(m,n)T(m,n) \tag{2}$$

$$O_2(m,n) = L(1 - T(m,n)). \tag{3}$$

Here's what each term means:

- $O_H(m,n)$: The hazy image at pixel (m,n).
- $O_J(m,n)$: The clear, haze-free image.
- $T(m,n)$: Scene transmission reaching the observer.
- L: The global atmospheric light.
- $D(m,n)$: The distance to the observer.

Transmission $T(m,n)$ is calculated as:

$$T(m,n) = \exp(-\beta \cdot D(m,n)) \tag{4}$$

where β is the scattering coefficient.

Now, looking at (3), two things happen:

- When the observer is close, $D(m,n) \to 0$, making $T(m,n) \to 1$, so $O_H(m,n) \approx O_J(m,n)$.
- When the observer is far, $D(m,n) \to \infty$, making $T(m,n) \to 0$, so $O_H(m,n) \approx L$.

So, estimating the depth map $D(m,n)$ is key for getting accurate scene transmission in dehazing.

To get the dehazed image from (1), we also need to know the global atmospheric light L. There are methods to estimate

L by finding bright pixels. For instance, Tan et al. [15] just used the brightest pixel, while He et al. [16] used the top 0.1% of the brightest pixels. But these methods can get tripped up by white objects in the sky. Kim et al. [17] improved on this with a hierarchical search, and Wang et al. [18] refined it further by considering sky region distribution.

In this work, we use a new similarity function matrix to estimate the depth map, which helps us calculate the true scene transmission $T(m,n)$ and atmospheric light L for effective dehazing.

II. Proposed Methodology

This section goes over our single image dehazing method using a Type-2 Membership Function (MF). As shown in Fig. 1, the approach starts by getting the minimum channel of the hazy input image. Then, this minimum channel is processed with a Type-2 MF-based similarity function matrix. This step helps in estimating the depth map, which is essential for calculating the true scene transmission and global atmospheric light. Finally, using these estimates, we get the dehazed image via the atmospheric scattering model. Each step in Fig. 1 is explained below.

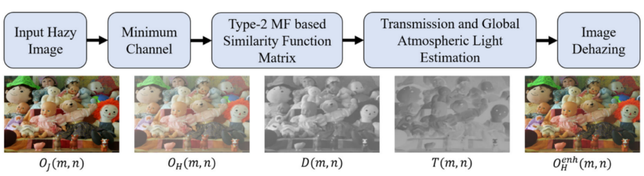

Fig. 1: Block diagram of the proposed method. Here, $O_J(m,n)$, $O_H(m,n)$, $D(m,n)$, $T(m,n)$, and $O_{\text{enh}H}(m,n)$ are the reference haze-free image, observed hazy image, depth map, transmission map, and dehazed image, respectively. The observed hazy image is synthetically generated with haze concentration $HC = 40\%$.

A. Minimum Channel

As mentioned earlier, haze affects distant pixels more than nearby ones. This means that an initial depth map estimate can be done by calculating the minimum intensity across the R, G, and B channels at each pixel:

$$O_H^{\text{min}}(m,n) = \min_{\tau \in \{R,G,B\}} O_H^{\tau}(m,n) \tag{5}$$

where τ represents the R, G, and B channels in the hazy image $O_H^{\tau}(m,n)$.

B. Dark Channel Prior (DCP)

According to Dark Channel Prior (DCP) theory [16], at least one color channel will have close to zero intensity in a haze-free image, a concept known as the *dark channel*. For each pixel in a local patch Ω_k centered at (m,n) with size $\sqrt{k}\times\sqrt{k}$, the dark channel is defined as:

$$O_H^{\Omega_k}(m,n) = \min_{\tau \in \{R,G,B\}} \left(\min_{(i,j)\in\Omega_k} O_H^{\tau}(i,j) \right) \tag{6}$$

Patch size plays a big role in this:

- **Larger patch sizes** darken the dark channel but can lead to halo effects near edges.
- **Smaller patch sizes** make the recovered scene over-saturated.

To address these issues, our proposed Type-2 MF-based similarity function considers neighborhood info and spatial ambiguities, allowing us to avoid halo effects and maintain sharpness around edges.

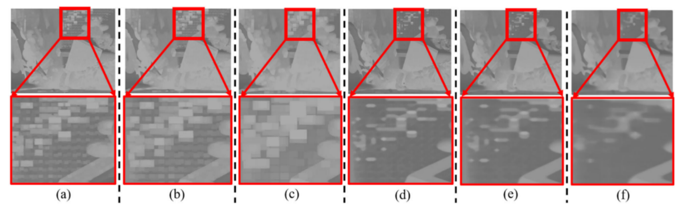

Fig. 2: Comparison of scene transmission. Top row: Original image with marked regions. Bottom row: Zoomed view of marked regions. (a)-(c) show DCP with patch sizes 5x5, 7x7, and 11x11, respectively. (d)-(f) show results using the proposed method with the same patch sizes, maintaining edge clarity.

Fig. 2 shows a comparison between scene transmission estimated using DCP and our proposed method across different patch sizes. With DCP, scene transmission varies around edges as patch size changes. However, our method keeps the edges sharp and avoids halo effects, even as patch size increases.

In the next section, we explain how our Type-2 MF-based similarity function matrix is designed to improve these results.

C. Type-2 MF Based Similarity Function Matrix

So, let's dive into the Type-2 MF (Membership Function) Based Similarity Function Matrix. Normally, a Type-1 MF just has a single membership value, which is alright for some basic uncertainties in our elements. But Type-2 MF, as referenced in [**?**], takes things a step further by allowing the membership values themselves to be uncertain (like a range of values instead of a fixed number).

For an observed hazy image O_H, if p_{ij} represents the pixel intensity at position (i,j), we define a Type-1 fuzzy set like this:

$$M_{ij} = \{(p_{ij}, \mu(p_{ij})) \mid \forall p_{ij} \in O_H\}$$

where $0 \le \mu_{ij} \le 1$ indicates the membership values associated with each pixel intensity p_{ij}. Now, a Type-2 fuzzy set is defined by:

$$\tilde{M}_{ij} = \{(M_{ij}, \mu(M_{ij}))\}.$$

The Type-2 MF similarity function we're using is inspired by Singh et al. [**?**] for image denoising, based on fuzzy sets from L. A. Zadeh [**?**]. First, we extract the minimum channel of the observed hazy image using equation (6). Then, we build a similarity function matrix for each pixel position in this minimum channel, using Type-2 MF and considering its neighborhood pixels within a small patch. This patch, Ω_k, is centered at (i,j) and has a size of $\sqrt{k} \times \sqrt{k}$, defined as:

$$\Omega_k = \{O_H^{\min}(i+u, j+v) \mid \forall (u,v) \in [-\text{FHW}, \text{FHW}]\}$$

where FHW is the half-filter width around the pixel, given by $\text{FHW} = \frac{\sqrt{k}-1}{2}$.

Remark 2: The patch size $\sqrt{k}$ can be {3, 5, 7, 9, ...}. If $k = 9$, the patch size is 3×3, and thus, $\text{FHW} = \frac{3-1}{2} = 1$.

Now, for each patch Ω_k, we define the similarity function as:

$$\eta_k^h = f(\Omega_k, H_h) = \exp\left(-\frac{1}{2}\left(\frac{\Omega_k - H_h}{\sigma}\right)^2\right)$$

where σ is the standard deviation and $H_h = \{l_1, l_2, \dots, l_h\}$ represents h-middle means for the patch. For odd-sized patches, these h-middle means are calculated as:

$$H_h = \frac{1}{2h-1} \sum_{i=q-h+1}^{q+h-1} \Omega_i$$

where $q = \frac{k+1}{2}$ and $h = 1, 2, 3, \dots, q$.

Remark 3: In (12), h is the number of means for the patch. For a patch size $\sqrt{k} = 9$, or 3×3, the neighboring pixels will be $\{\Omega_1, \Omega_2, \Omega_3, \Omega_4, \Omega_5, \Omega_6, \Omega_7, \Omega_8, \Omega_9\}$. With $q = \frac{9+1}{2} = 5$, we have $h = 1, 2, 3, 4, 5$. Using equation (12), the h-middle means are:

$$H_1 = \Omega_5, \quad H_2 = \frac{1}{3}(\Omega_4 + \Omega_5 + \Omega_6),$$

$$\text{H}_3 = \tfrac{1}{5}(\Omega_3 + \Omega_4 + \Omega_5 + \Omega_6 + \Omega_7),$$

$$H_4 = \frac{1}{7}(\Omega_2 + \Omega_3 + \Omega_4 + \Omega_5 + \Omega_6 + \Omega_7 + \Omega_8),$$

$$\text{H}_5 = \tfrac{1}{9}(\Omega_1 + \Omega_2 + \Omega_3 + \Omega_4 + \Omega_5 + \Omega_6 + \Omega_7 + \Omega_8 + \Omega_9).$$

The Type-2 MF based similarity function matrix from (11) can also be expressed as:

$$\eta_k^h = \begin{bmatrix} f(\Omega_1, l_1) & f(\Omega_2, l_1) & \dots & f(\Omega_k, l_1) \\ f(\Omega_1, l_2) & f(\Omega_2, l_2) & \dots & f(\Omega_k, l_2) \\ \vdots & \vdots & \ddots & \vdots \\ f(\Omega_1, l_h) & f(\Omega_2, l_h) & \dots & f(\Omega_k, l_h) \end{bmatrix}_{h \times k}.$$

or, simply as:

$$\eta_k^h = \begin{bmatrix} \eta_1^1 & \eta_2^1 & \dots & \eta_k^1 \\ \eta_1^2 & \eta_2^2 & \dots & \eta_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^h & \eta_2^h & \dots & \eta_k^h \end{bmatrix}_{h \times k}.$$

We'll use this similarity matrix to estimate the depth map and global atmospheric light of the hazy image. Next up, we'll dive into how we estimate scene transmission using this depth map.

D. Depth Map and Scene Transmission

Alright, let's talk about the depth map and scene transmission estimation. For each pixel in a local patch Ω_k, the MF value is estimated using equation (14) as:

$$\eta_k = \begin{bmatrix} \eta_1 & \eta_2 & \dots & \eta_k \end{bmatrix}$$

where η_k is the average MF value of h-middle means for each pixel k in Ω_k. This can be broken down as:

$$\eta_1 = \frac{1}{h} \sum_{i=1}^{h} \eta_i^1 = \eta_1^1 + \eta_2^1 + \eta_3^1 + \dots + \eta_h^1$$

$$\eta_2 = \frac{1}{h} \sum_{i=1}^{h} \eta_i^2 = \eta_1^2 + \eta_2^2 + \eta_3^2 + \dots + \eta_h^2$$

$$\vdots$$

$$\eta_k = \frac{1}{h} \sum_{i=1}^{h} \eta_i^k = \eta_1^k + \eta_2^k + \eta_3^k + \dots + \eta_h^k$$

So, we get a depth map $D(m,n)$ for the hazy image at each pixel (m,n) by taking a weighted average of pixel values in Ω_k around it, using the MF values η_k from equation (16):

$$D(m,n) = \frac{\sum_{i=1}^{k} \eta_i(m,n)\, \Omega_i(m,n)}{\sum_{i=1}^{k} \eta_i(m,n)}$$

Now, the true scene transmission $T(m,n)$ can be estimated as:

$$T(m,n) = \exp(-\beta \cdot D(m,n))$$

where β is set to 1 (found empirically) to represent the atmosphere's angular scattering coefficient.

Remark 4: For visible light, the scattering coefficient β relates to wavelength λ using an inverse power law:

$$\beta(\lambda) = \frac{\text{Constant}}{\lambda^\gamma}$$

where $\gamma \in [0, 4]$. For fog or dense haze, $\gamma = 0$, meaning β doesn't vary much with wavelength. Narasimhan and Nayar [**?**] noted that for visible light, the scattering coefficient remains mostly constant. Through experiments, $\beta = 1$ gave the best results here. If $\beta < 1$, bright areas get brighter, while $\beta > 1$ makes dark regions darker, which can lead to over-saturation.

Next, we'll go into estimating the global atmospheric light using our Type-2 MF similarity function matrix from (14).

E. Global Atmospheric Light

Alright, let's get into how we find the global atmospheric light. Within a local patch, the h-middle means tend to be smoothly increasing or decreasing in homogeneous regions but get non-monotonic when there's any kind of break or edge in the image [**?**], [**?**]. This behavior also impacts the similarity function matrix from equation (14). So, using these observations, we compute the atmospheric light at each pixel location with:

$$L(m,n) = \min\left(\eta_h^k(m,n)\right)$$

Typically, we use the most haze-opaque regions to estimate atmospheric light. But here's the catch: the really bright objects in these regions are ignored. This is because bright spots tend to have more variation in means (less smoothness) compared to hazy areas. So, we define the global atmospheric light as the average:

$$L = \text{mean}\,(L(m,n))$$

The global atmospheric light, in essence, represents the intensity of the farthest hazy pixels, often found in the sky regions. This approach is designed to minimize interference from bright or white objects, thanks to how the means vary in the Type-2 MF similarity function.

Remark 5: If a patch contains a bright or white object in a haze-free area, our similarity function matrix will find h different means. The weighted average of MFs for these means (combined with the patch's pixel intensity values) will be led by the darker pixels within that patch. This way, the global atmospheric light avoids getting affected by those bright or white objects.

F. Image Dehazing

Now, let's put it all together and get the final dehazed image! Using our observed hazy image $O_H(m,n)$, the estimated scene transmission $T(m,n)$ from equation (18), and the global atmospheric light L from equation (21), we get the dehazed output image by rearranging equation (1) as:

$$O_H^{\text{enh}}(m,n) = \frac{O_H(m,n) - L}{T(m,n)} + L$$

Algorithm 1: Proposed Method for Single Image Dehazing.

Patch Size (Input): $k = 9 \Rightarrow \sqrt{k} \times \sqrt{k} = 3 \times 3$
Min Channel: $O_H^{min}(m,n) = \min_{\tau \in \{\text{R,G,B}\}} O_H^{\tau}(m,n)$
for $O_H^{min}(i,j)$ **do**
  $\Omega_k(i,j) = \{O_H^{min}(i+u, j+v)\}\, \forall u,v \in [-F_{HW}, F_{HW}]$
  Similarity Function Matrix: $\eta_k^h(i,j) = f(\Omega_k, H^h)$
  MF Value: $\eta_k'(i,j)$ by (15)
  Depth map: $D(i,j)$ by (17)
  Scene transmission: $T(i,j)$ by (18)
end for
Global Atmospheric Light: L by (21)
Dehazed Image: $O_H^{enh}(m,n)$ by (22)

In this result, $O_H^{\text{enh}}(m,n)$ represents our dehazed image, which aims to resemble a clear, haze-free version $O_J(m,n)$.

Overview of the Method

The steps outlined in Algorithm 1 make up our single-image dehazing approach based on the Type-2 MF. This method produces a dehazed image that closely matches the original haze-free version. It also adds some illumination to make the enhanced image visually appealing, avoiding issues like over-saturation in haze-free and dark areas. The proposed similarity function helps improve visibility in darker regions by adding balanced illumination.

III. EXPERIMENTAL RESULTS AND VALIDATIONS

This section presents the experimental analysis of the proposed method as well as comparison with other state-of-theart methods. The experiments have been performed on Ubuntu 64-bit Operating System, x64-based processor of Intel(R) Core i5 .

A. Datasets

The Dataset used here to generate the result is Outdoor REalistic Single Image DEhazing (RESIDE). Due to low power of CPU, using only top 30 hazy images and it's ground truth image.

B. Performance Metrics for Validations

The performance metrics used for quantitative comparison are Peak Signal to Noise Ratio (PSNR).

PSNR Formula

The PSNR is calculated using the Mean Squared Error (MSE) as follows:

$$\text{PSNR} = 20 \cdot \log_{10}\left(\frac{\text{MAXI}}{\sqrt{\text{MSE}}}\right)$$

where:

- MAXI is the maximum pixel intensity in the image (255 for 8-bit images),
- MSE is the Mean Squared Error between the dehazed and reference images.

Calculating MSE

The MSE is computed by averaging the squared differences between corresponding pixels in the dehazed (I) and reference (K) images:

$$\text{MSE} = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} (I(i,j) - K(i,j))^2$$

where m and n are the image dimensions, and $I(i,j)$ and $K(i,j)$ are pixel values in the dehazed and reference images.

Interpretation: Higher PSNR values indicate better quality. Generally, values above 30 dB are considered high quality for most image processing tasks.

C. Results

Following Fig. 3 histogram shows the effectiveness of our algorithm. As we can see that most of the dehazed images have 28 to 30 PSNR value which is a good sign means it's been dehazed as expected and we can also see that some of images have approx 36 PSNR means these images are dehazed almost it's becomes clear images.

In Fig. 4 we have some images of hazed and dehazed one after applying our algorithm. We can see that, the hazes which are at low depth are being dehazd but hazes in background or nearly at infinite distances are still present.

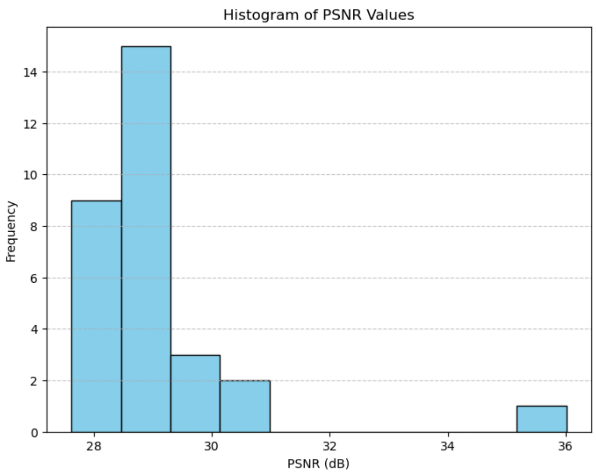

Fig. 3: Histogram of PSNR values for 30 images comparing hazed and dehazed images. This histogram shows the distribution of PSNR values obtained from our proposed dehazing algorithm. Higher PSNR values indicate that the dehazed images closely resemble clear, haze-free images, demonstrating the effectiveness of the algorithm.

(a) Hazed Image 1

(b) Dehazed Image 1

(c) Hazed Image 2

(d) Dehazed Image 2

(e) Hazed Image 3

(f) Dehazed Image 3

Fig. 4: Comparison of Hazed and Dehazed Images

D. Variation in Patch Size

On increasing patch size image are becoming more dehazed but computaion are increasing very high so there is tradeoff between quality and computaions. While the results presented are using patch size of 3.

IV. Future works

As previously mentioned that, the hazes which are at low depth are being dehazd but hazes in background or nearly at infinite distances are still present. So, there is scope of future work for managing the pixels which are at large distance. This algorithm also doesn't work well when so many pixels are white in cluster it simply output all white so identifying these regions using gradient method or something more better can help this algorithm for better accuracy.

V. Conclusion

This paper proposes a novel Type-2 membership function based similarity function matrix to estimate the depth map and global atmospheric light for single image dehazing. The proposed method outperforms others qualitatively and quantitatively. The dehazed images produced using the proposed method manage to retain the structure near edges which avoids halo artifacts in the sky region as well. This results in enhanced contrast with naturalness close to human perception.