Swarm Based Optimization for Image Dehazing from Noise Filtering Perspective

Sunkavalli Jaya Prakashi¹, Manna Sheela Rani Chetty¹, Jayalakshmi Aravapalli²

¹ Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, AP, India
² Department of Computer Science and Engineering, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru 520007, Vijayawada, AP, India

Corresponding Author Email: prakashsunkavalli@sircrrengg.ac.in

https://doi.org/10.18280/isi.270416

Received: 23 March 2022
Accepted: 30 June 2022

Keywords: de-hazing, HRNFP, bat algorithm, noise filtering

ABSTRACT

Haze may readily corrupt digital photos acquired in an outside situation, degrading the information communicated. To address this issue, several studies on picture haze reduction have been done, with the technique based on the dark channel prior assumption being regarded the state-of-the-art in recent years. Consolidation of observations is a key component of this strategy. However, the suggested technique takes a theoretical approach to picture degradation, treating the degraded image as a tainted product. To mark the noise severity and ambient light, two maps are created. In terms of color change, the parameters are adjusted using Bat algorithm with a penalty function. The outcomes of the experiments were compared to seven existing methodologies, as well as an examination of algorithm complexity. These studies back up the suggested approach’s efficacy, efficiency, flexibility, and theoretical soundness.

1. INTRODUCTION

When shooting photographs outside, the quality of the images may deteriorate due to the presence of a large number of air particles, which will make it difficult to distinguish the details in the subject matter. This is because the air particles will obscure the intricacies of the subject matter. This is due to the fact that the particles in the air will make it difficult to see the subject matter. As a result of the fact that all of these different particles, which include haze, fog, and smoke, have a comparable effect on the readability of images, we refer to all of them collectively as “haze.” Alterations in color will also be produced, as a direct result of influence from the air [1]. If one wishes to make use of images in a broad variety of settings [2, 3], one must have photographs that are not only distinct but also of a good quality. In order for these photos to achieve the quality requirements that have been established, the saturation, contrast [4], and entropy levels all need to be at an acceptable level. Due to the intricacy of the issue, which needs to be addressed in order to find a solution, the removal of haze has been the topic of a substantial amount of research.

During the course of the most recent few years, a number of different strategies for dehazing photographs have been conceived of, developed, and put into practice. 2015 saw the completion of the study that led to Liu's [5] classification of the various approaches. During the same year, additional work on the review of the Dark Channel Prior (DCP) [6] was carried out. 2015 was the year when both of these pieces of work were finished. The method of eliminating haze from photographs can be broken down into three primary categories: (1) dehazing a large number of images [2, 7], (2) haze removal that requires extra information [8-11], and (3) dehazing a single shot [12-14]. On the other hand, the first two distinct kinds of approaches are not appropriate for use in real-time applications due to the high complexity of these methods and the additional resources that are required in order for them to be implemented. This is because real-time applications require a great deal of processing power in order to function properly. In recent years, a substantial number of scholars have developed an interest in a method known as single picture dehazing. This is because it is an easy and effective method for reducing haze from photographs. This is the primary reason behind this. The method that was presented by He et al. [12] in 2011 and is based on the idea of DCP is, without a doubt, the most remarkable illustration of a single picture dehazing approach that has been provided to this day. It was established on the notion of DCP. It is predicated on the concept that DCP can be applied to an image in order to lessen the amount of haze that can be observed in it. It is reported that he, age 11, was the one who first thought of using this approach.

The objective of this article is to introduce a technique known as picture BAT-Haze Removal from the Noise Filtering Perspective. This name was given to the technology by its developers (BAT-HRNP). This approach is predicated on a theory that was initially introduced by Lee [13] in the year 1980. This method, which employs a noise filtering strategy, is intended to improve the clarity of photographs that have been muddled by haze. In most instances, images that have been destroyed by haze in the form of noise have a high intensity level but a low saturation level. This is because haze tends to make the colors in the shot appear washed out. This is due to the fact that haze has a tendency to water down colors. These are the qualities that can be seen in these photographs and immediately grab your attention. As a consequence of this, the amount of haze that is actually there can be interpreted as the weighted total of the picture's intensity and saturation values that were input.
It is feasible to make use of the same procedure in order to estimate the amount of light that is present in the surroundings; nevertheless, a small modification will be required anytime images contain an excessive number of bright things. The process of filtering out picture noise begins with the building of two weighted maps, which is then followed by the utilization of the severity map's local statistics. Finally, the procedure concludes with the elimination of any remaining image noise. The removal of any leftover picture noise brings the operation to a successful conclusion. The bat algorithm is used in order to optimise all four of these parameters in order to attain the most beneficial results that may be achieved. The end result of this undertaking should be a picture whose degree of saturation is as close to that of the utmost that can be achieved given the constraints of the medium. It is expected that this will be the case with the finished product. In addition, a penalty function is executed whenever an overall evaluation of the subject's fitness is being finished in order to minimize the color shift. This is done with the intention of ensuring that the color change is as small as possible. This is done with the goal of avoiding the color change that will occur as a result. The reason for this is to ensure that the shift does not become extremely dramatic, therefore this is why it is done. Studies that were carried out on a total of one hundred images, some of which featured haze and some of which did not, revealed that the shift-scale parameters have a tendency to cluster around constant values with relatively modest changes. These findings were discovered as a result of the findings of the studies that were carried out. The experiments were carried out so that an accurate determination could be made regarding the shift-scale characteristics. The results of the experiments offered undeniable proof for the statement that this was the case. It has been discovered that selecting the parameters supplied by the Bat Algorithm and their mean values for the shift-scale factors is the most effective strategy for maximizing the time efficiency of HRNFP. This was discovered after extensive research was conducted. It has been found that this is the most effective approach. This is the conclusion that we reached after a significant amount of research and consideration.

2. IMAGE DEHAZING

Image Dehazing Handcraft priors, such as dark channel prior (DCP) [5], color attenuation prior (CAP) [7], color-lines [6], and hazelines [8], were typically utilized in the early single-image dehazing algorithms. The aforementioned techniques, which are based on priors, frequently result in the production of visual representations that exhibit a high degree of visibility. On the other hand, when the scenes do not satisfy these priors, various methods of dehazing have a tendency to produce outcomes that are unrealistic. This is owing to the fact that these priors are dependent on the data that has been empirically collected. The approach of dehazing that has gained the most popularity in recent years is one that is based on learning. This is because there has been an accelerated expansion of deep learning in recent years. It is believed that DehazeNet [9] and MSCNN [10] were the first to explore with the use of CNNs for the purpose of image dehazing. They learn how to make an estimate of $t$, and after that, they acquire the result together with an estimate of $A$ that was made using the standard approach. After this stage, both the DCPDN [11] and the GFN [13] use two sub-networks to estimate $t$ and $A$, respectively, whereas the GFN [13] predicts the fusion coefficient maps for the three picture operations that have been specified. AOD-Net [12] rewrites Eq. (1) in such a way that the network only needs to estimate a single component. This reduces the amount of work that needs to be done by the network. This is a different approach than the one that was used before. Because of this, the amount of work that needs to be done by the network is going to be cut down significantly. According to GridDehazeNet [14], it is recommended that one first have an understanding of how to restore the image before attempting to estimate $t$. This is owing to the fact that attempting to estimate $t$ will lead to subpar solutions. The reason for this is due to the fact that any attempt to estimate $t$ will result in inadequate solutions. In addition, the vast majority of the most recent studies either estimates the picture that is clear of haze or the residual that is present between the picture that is clear of haze and the picture that is foggy. Because the effectiveness of the learning-based approaches for dehazing is highly dependent on the quality and quantity of the dataset, many different datasets have been presented. This is because of the fact that there have been many different datasets proposed. As a direct consequence of this, a wide variety of distinct datasets have been suggested. These dehazing datasets can be loosely sorted into two categories: natural datasets and artificial datasets. Both types of datasets contain natural phenomena. Genuine datasets generate real hazy photos by employing real haze, which is produced by haze machines designed specifically for that purpose by trained professionals. This is done in order to ensure the highest level of precision feasible. The bulk of the time, synthetic datasets make use of the Eq. (1) in order to synthesize the necessary hazy photos with haze-free photographs and depth maps. This is done in order to fulfill the purpose of the synthetic dataset. It is difficult to collect enough image pairs, and the distribution of the haze produced by the haze machine is still noticeably different from the actual haze, despite the fact that real datasets appear to have more appeal. As a direct result of this, the majority of approaches typically involve the utilization of synthetic datasets in order to fulfill the requirements of both training and testing. This study introduces a new synthetic remote sensing picture dehazing dataset that has been given the name RS-Haze. The purpose of this work is to analyze the capability of the method to remove highly nonhomogeneous haze, and so the dataset was created with that goal in mind. The phrase "remote sensing picture dehazing" serves as the inspiration for the name of this particular dataset. This dataset offers a different perspective than the datasets that were covered before in the discussion. When compared to prior datasets, RS-Haze is more comprehensive and accurate since it takes into consideration sensor characteristics, haze distribution and particle size, wavelengths of light, and other aspects of the environment that are not always taken into account. Image dehazing is the act of restoring an image to its original state after it has been distorted due to the interference of atmospheric light ($A$) and the physical distance that separates the camera from the subject of the photograph ($t$). The technical definition of a pixel in an image is as follows:

$$P(i) = f(i) \times t(i) + A(1 - t(i))$$  \hspace{1cm} (1)$$

where, $P$ refers the actual intensity recorded in the form of RGB values, $J$ refers the actual scenic radiance value at the distance $t$ and $A$ refers the atmospheric light. On the other hand, an image is nothing more than a collection of pixels organised
into either one layer (in the case of a grayscale image) or three layers (RGB) depending on the colour of the image. Therefore, the most important task is to derive the radiance from the pixel values by making an estimate of the intensity of the ambient light and the distance that separates the camera and the subject being photographed.

According to a recent study, one strategy that might be taken demonstrates that the radiance of the scene can be retrieved from the collected pixels by making use of noise filtering; this method is referred to as Hazal Removal from Noise Filtering Perspective (HRNFP). In this method, the traditional method of representing images has been rebuilt in the form that is described below:

\[ P(i) - A = (J(i) - A) \times n(i) \] (2)

where, the transmission distance \( t \) is treated as noise disturbance \( n \).

From the studies, it is observed that the pixel value affected due to haze has high brightness value \( B \) with low saturation \( S \). Based on these parameters \( B \) and \( S \), the noise on an image can be reframed as follows:

\[ M_n = \rho(1 - B) + (1 - \rho)S \] (3)

where, \( B \) is the average of R, G, B colors of image and \( S \) is the difference between the numeric value 1 and minimum pixel value \( P \) to the ratio of maximum pixel value. The parameter \( \rho \) represents the weight factor that restricts the involvement of brightness value and saturation based on the scenic radiance.

Based on the shift-scale factor the noise map has been further refined as:

\[ M_s = \alpha + \beta M_n \] (4)

Now the entire noise has been computed and the final formation of pixel in the actual picture (i.e. Eq. (2)) can be reframed as:

\[ P(i) - A = (J(i) - A) \times M_s(i) \] (5)

Apart from the noise due to transmission distance another factor that affects the scenic radiance is atmospheric light. And this is also related to the brightness and saturation value and hence another constriction factor namely \( M_{st} \) is derived from \( B \) and \( S \) as follows:

\[ M_{st} = \varphi B + (1 - \varphi)S \] (6)

where, \( \varphi \) is the weight factor. And thus the value of \( A \) can be identified as:

\[ A = \max(M_{st}(i)) \] (7)

The final \( A \) can be cumulated as:

\[ A = A + \gamma \] (8)

where, \( \gamma \) is a correlation scalar value related to the atmospheric light.

On addressing all the values of \( \rho, \alpha, \beta, \gamma \) as weight factors, these parameter values should be optimal such that it should not be deviating the actual radiance value of the scene. Hence a swarm based optimization model is employed to optimize the parameter values.

Visual impairments of varying degrees are exhibited by both Autobots and Decepticons. In the beginning, transformers were proposed as a possible solution to the problems that may arise when carrying out natural language processing (NLP). These issues need the employment of numerous heads of self-attention and feed-forward MLP layers that are piled on top of one another in order to recognize non-local connections between individual words. This is necessary in order to do so. In addition, Dosovitskiy and his coworkers developed the very first pure Transformer model for use in picture recognition. They are also the ones who thought up the phrase “Vision Transformer,” which is often referred to by its acronym "ViT." In addition, Dosovitskiy and his coworkers developed the very first pure Transformer model for use in picture recognition. Recent research, such as the one that resulted in the development of the pioneering pretrained image processing Transformer, has focused on the application of Transformers in order to investigate the possibility of treating low-level vision issues. This is being done in order to investigate the prospect of curing low-level vision issues (IPT). IPT is a picture patching and image processing tool that works very similarly to ViT and applies vanilla Transformers to picture patches in a straightforward manner. IPT is also capable of processing photos. IPT was developed by Adobe Systems, the company that owns IPT. In order to attain video super-resolution, the authors of suggested developing a spatial-temporal convolutional self-attention network. This network would make use of the information that is specific to the area. The researchers are the ones who are in charge of putting together this network. SwinIR and UFormer are two relatively new computer programmes that are capable of applying effective window-based local attention models to a wide range of image restoration activities. These programmes were developed in response to the growing demand for image restoration tools. These applications are available for use on personal computers running either Windows or Mac OS. MLP vision models. Recently, a number of researchers have asserted that when a patch-based architecture is utilized, such as the one that is utilized in ViT, it raises the question of whether or not extensive self-attention mechanisms are necessary or not. This assertion was made in light of the fact that ViT utilizes a patch-based architecture. This claim was made taking into consideration the fact that the architecture of ViT is one that is patch-based. [There is probably more than one citation for this] [There is probably more than one citation for this] For instance, MLP-Mixer gets rid of the need for self-attention in ViT by utilizing a straightforward token-mixing MLP, which, in the end, leads to an all-MLP design. This is accomplished by applying an uncomplicated token-mixing MLP. The authors of proposed a model that they dubbed the gMLP, and they referred to it as their contribution to the field. In this model, the model's aims are accomplished by the collaborative efforts of a spatial gating unit and visual tokens. An Affine transformation, as opposed to the more traditional Layer Normalization method, is utilized in the processing of ResMLP in order to speed the procedure as much as possible. This is done in order to get the desired result in a more timely manner. Recent algorithms such as FNet and GFNet have shown that the fundamental Fourier Transform can be utilized as a workable alternative to either self-attention or MLPs. These algorithms were developed relatively recently. These methods have demonstrated that the fundamental Fourier Transform can be applied in a variety of contexts. The results
of these algorithms provided unmistakable evidence that this assertion is correct.

3. SWARM BASED IMAGE DEHAZING

Yang and Gandomi [14] first introduced the Bat methodology as a method for addressing problems involving non-linear continuous optimization. Bat is an acronym for "battlement algorithm." It was designed with the intention of addressing issues connected to the optimization of a singular objective, which was the driving motivation behind its creation.

The following constituents make up the core of the bat algorithm, which can be decomposed as follows: Bats may make use of echolocation, which is analogous to sonar and enables them to detect prey, obstacles, and roosting niches in their surroundings. Bats may also utilize this ability to communicate with one another. In order to locate the source of a loud noise, bats will often make a loud noise first, and then listen for the echo of that sound. The presentation of the bat algorithm, which is based on bat behavior, includes the following aspects that are incorporated into the presentation.

1. Bats use reflecting echolocation to measure distance and can tell the difference between food, prey, and barriers.
2. Bats fly at a random velocity $v_i$ and loudness $A_0$ from point $x_i$ in the wavelength. Each bat's wavelength may be modified dynamically depending on the distance to the target. The loudness is a dynamic value that varies from $A_0$ to $A_{min}$ (highest to lowest).

3.1 Motion of bats

Each bat travels closer to the bats that have superior answers. In the meanwhile, each bat's frequency and velocity are updated throughout the number of iterations. For the following iteration ($t+1$), the frequency, velocity, and position of each bat are updated as follows:

$$
\begin{align*}
\dot{f}_i &= f_{min} + (f_{max} - f_{min})\beta \\
\dot{v}_i &= \dot{v}_i + (x_i - x_s)e^{-\gamma t}f_i \\
x_i^{t+1} &= x_i^t + \dot{x}_i \\
\end{align*}
$$

where, $\beta$ is a random number ranges from 0 to 1. $x_s$ represents the global best solution which is attained from iteration 1 to $t$.

In bat algorithm, the neighborhood search accomplished from the so far obtained optimal solution with a random walk which is represented as:

$$
x_{new} = x_{old} + \varepsilon A^t
$$

where, $A^t$ is the average loudness of all bats and $\varepsilon$ is a vector with values ranging from -1 to +1.

3.2 Loudness and pulse emission

The loudness of a bat and the rate at which it emits pulses are inversely associated; as the bat gets closer to finding food, the loudness of the bat will decrease, but the rate at which it emits pulses will increase, and vice versa. Both the decibel level and the pulse emission can be expressed numerically as:

$$
\begin{align*}
A_i^{t+1} &= \alpha A_i^t \\
r_i^t &= r_i^t[1 - \exp(-\gamma t)]
\end{align*}
$$

where, $\alpha$ and $\gamma$ are constants.

In this Bat optimization algorithm, the weight factors of $\rho$, $\alpha$, $\beta$, $\gamma$ will be given as input for extracting the optimal values. The output picture saturation is the goal function that must be maximized. However, increasing picture saturation without limits will produce an artificial output that fails to express the genuine information contained in the original image. As a consequence, a penalty factor ($N$) is used to exclude particles that have seen a significant color shift as a result of the dehazing procedure.

$$
N = 1 - \varepsilon
$$

where, $\varepsilon$ is the value of $t$-test for every individual in the population. The fitness value for every individual can be computed using the fitness function defined mathematically as follows:

$$
f(x_i) = (1 - N)S + N(1 - S)
$$

where, $S$ is the average saturation rate. The algorithm for identifying optimal parameter values is given in Algorithm 1.

**Algorithm 1: Bat algorithm for image Dehazing (BAT-HRNFP)**

**Input:** The parameters $\rho$, $\alpha$, $\beta$, $\gamma$ upper and lower bound Objective Function $f$

Set the parameters Pulse Frequency $PF_i$, Pulse Rates $r_i$ and Loudness $A_i$ 
Initialize $N$- Number of Bats, $Max_{iterations}$, $t$=1

**repeat**

**Gbest** $\leftarrow$ $x$(min(Fit))

**repeat**

**end for**

$Gbest$ $\leftarrow$ Initializing Random Solution

if $(rand<A_i && f(x_i) < f(Gbest))$ then

$A_i$ $\leftarrow$ $y_i$

Increase $r_i$

Deduce $A_i$

end if

$t$ $\leftarrow$ $t$ + 1

until $(Max_{iterations} \geq t)$

**OUTPUT:** Gbest
4. EXPERIMENTAL ANALYSIS

One hundred color pictures that had been captured in a variety of environments were put through a series of tests in order to illustrate how practical and adaptable the method that had been given was. These tests were performed on the photographs. There are images here that suffer from issues such as haze, low lighting, and an absence of contrast are shown in the Table 1. For archival purposes, these pictures are preserved in the JPEG format with 8 bits of depth. The computer programmes were created with Matlab 2018a, and they were run on a personal computer that has a Core i7 CPU from the 10th generation, 8 gigabytes of random access memory (RAM), and a 1 terabyte (TB) hard drive.

In Table 2, the performance metrices values of BAT-HRNFP such as colourfulness, saturation, contrast, sharpness, mean brightness, and entropy are all compared with the other existing algorithms.

When compared to the outcomes of other methods, the BAT-HRNFP method performs significantly better than the other methods in a variety of different metric comparisons.

The mean value for colorfulness BAT-HRNFP is 0.108, which is quite close to the HRNFP value of 0.115. The least amount of hue shift was seen in samples Fattal08 and Zhu15, with mean values of 0.099 and 0.125, respectively. Other methods, such as Tarel09, Tarel10, He11, and Meng13, have significant color deviations, with mean values of 0.167, 0.153, 0.163, and 0.140 respectively.

It is possible to observe the result of the saturation. The algorithm He11 comes in first with a mean value of 0.539, followed by BAT-HRNFP with a mean value of 0.5212 and Tarel10 with a mean value of 0.502. The HRNFP has a mean score of 0.446 and a median score of 0.5, which places it in fourth position among the other variables. The outcome of He11's method, on the other hand, is prone to color distortion, according to quantitative study. Tarel10's method, on the other hand, fails in the majority of scenarios and has poor adaptability in picture de-hazing.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Haze affected Image</th>
<th>De-hazed image using BAT-HRNFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>3</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Table 2. Comparison of Mean values of BAT-HRNFP vs other existing algorithms

<table>
<thead>
<tr>
<th></th>
<th>Colourfulness</th>
<th>Saturation</th>
<th>contrast</th>
<th>Sharpness</th>
<th>Mean-brightness</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fattal 08</td>
<td>0.099</td>
<td>0.316</td>
<td>0.023</td>
<td>0.020</td>
<td>0.363</td>
<td>6.267</td>
</tr>
<tr>
<td>Tarel 09</td>
<td>0.167</td>
<td>0.364</td>
<td>0.046</td>
<td>0.048</td>
<td>0.397</td>
<td>6.927</td>
</tr>
<tr>
<td>Tarel 10</td>
<td>0.153</td>
<td>0.502</td>
<td>0.032</td>
<td>0.041</td>
<td>0.226</td>
<td>6.618</td>
</tr>
<tr>
<td>He11</td>
<td>0.163</td>
<td>0.539</td>
<td>0.031</td>
<td>0.024</td>
<td>0.205</td>
<td>6.425</td>
</tr>
<tr>
<td>Meng13</td>
<td>0.140</td>
<td>0.425</td>
<td>0.042</td>
<td>0.036</td>
<td>0.250</td>
<td>6.706</td>
</tr>
<tr>
<td>Zhu 15</td>
<td>0.125</td>
<td>0.286</td>
<td>0.030</td>
<td>0.030</td>
<td>0.370</td>
<td>6.959</td>
</tr>
<tr>
<td>HRNFP</td>
<td>0.115</td>
<td>0.446</td>
<td>0.041</td>
<td>0.044</td>
<td>0.280</td>
<td>6.582</td>
</tr>
<tr>
<td>BAT-HRNFP</td>
<td>0.108</td>
<td>0.512</td>
<td>0.038</td>
<td>0.031</td>
<td>0.200</td>
<td>6.266</td>
</tr>
</tbody>
</table>
5. CONCLUSIONS

This study proposes a novel technique for removing haze from photographs by making use of a noise filtering perspective in order to obtain shots that are free of haze. The method was developed as part of this research. When it comes to regulating the depth shift near picture edges, the quantitative analysis indicates that pixel-wise noise estimate is more accurate than techniques based on local patches. This is the case when comparing the two methods. One of the numerous advantages of the BAT-HRNFP technique that has been presented is that it is resistant to images that have highly bright things in them. This is only one of the many advantages. This objective can be accomplished by making a modification to a rough estimate of the atmospheric value, which will make it possible to do so. The implementation of the Bat Algorithm, which enables the parameter settings to be automatically adjusted based on the picture being read in, is an additional way in which the adaptability of the recommended method can be further improved. The purpose of this optimization technique is to achieve the highest feasible amount of saturation in the final image while adhering to the limitations imposed by a hue change restriction. The qualitative analysis demonstrates that our method is superior to others in terms of its ability to produce images with enhanced saturation and contrast while simultaneously reducing the number of viewing artefacts that are introduced into the mix. This was demonstrated by the fact that our method received the highest score.

REFERENCES