International Journal of
Scientific Engineering and Technology Research

ISSN 2319-8885
Vol.03,Issue.24
September-2014,
Pages:4808-4813

www.semargroup.org,
www.ijsetr.com

A Novel Method for Video Dehazing by Multi-Scale Fusion

ARCHA S[1], ABDUL ALI[2]

[1]PG Scholar, Dept of CSE, ICET, Muvattupuzha, Kerala, India, Email: archa1491@gmail.com.
[2]Asst Prof, Dept of CSE, ICET, Muvattupuzha, Kerala, India

Abstract: An atmospheric phenomenon like haze significantly degrades the visibility of outdoor scenes. This is mainly due to the influence of atmospheric particles such as direct attenuation and airlight. In this paper we introduce a novel dehazing technique based on hazy video. By using algorithm we remove the haze effect. Our method is multi scale fusion strategy, based on the two derived input frames and weight matrix. The visibility details of haze and non haze region will be increased adaptively by white balancing and contrast enhancing procedure. Our strategy combines the input information in a per-pixel fashion to minimize artifacts. The pyramidal decomposition of fusion strategy is also avoiding halo artifacts. Our approach is faster and suitable for real time application.

Keywords: Dehazing, Outdoor Images and Videos, Pyramidal Decomposition.

I. INTRODUCTION

Under the bad weather condition, the visibility of captured scene is significantly degraded; which means the contrast and color of the images are degraded by the presence of fog and haze. For removing these weather effects from image or video, many reliable techniques are introduced in the vision system. The bad weather effect is optically due to the presence of substantial particles in the atmosphere that absorb and scatter the light. This absorption and scattering processes are generally modulated by the direct attenuation and airlight. Restoration of image or video has taken a specific condition such as the measure of airlight particles. This is mainly used in outdoor vision application such as surveillance, terrain classification, navigation etc. Video application such as outdoor surveillance involves high degree video contrast enhancement [2]. The degradation of contrast and loss of visibility has occurred by poor video signal. Haze removal is highly applicable in computer vision. Haze, fog, smoke etc are the weather effect of atmosphere.

In this paper we focus on haze effect which is similar to others. Dehazing is the process to remove the haze effects in captured images and reconstruct the original color of natural scene. The main characteristic of haze image is the intensity of haze pixel. Background pixels are always high in all color channels [3], but in at least one of the color channel the intensity of object pixel is low. The effect of haze in the input frame of video cannot overcome by the simple image processing techniques such as contrast enhancement or adaptive histogram equalization etc. So in this paper we propose a fast dehazing method for real time image or video processing. Haze removal is the challenging problem. Since the degradation is spatial variant, it depends on the scene quality. Several existing techniques are used for dehazing purpose. The contrast restoration of the image solve the dehazing problem with one or more input image [4], that have been taken in the uniform bad weather conditions. Some additional information are used in the polarization based dehazing technique [5], it take the advantage of partially polarized airlight. The depth based method [6] is used to estimate the scene depth information from the multiple images captured under different weather condition.

To overcome the drawback of multiple images as input, many researchers focus on the single degraded image. Results of such type of cases are explained in [7] [8] [9] and [10]. Many fast single image dehazing techniques are proposed for image and video processing. The guided filter scheme [14] is used for video dehazing based on the transmission map. However based on the existing techniques, we build our proposed system, which is fusion based dehazing technique for video image. This is the best and effective fusion strategy in the multi scale level. The haze removal is performed in the video frames of single input video. In video processing, the input hazy video is divided into frames containing group of pictures; where the first frame is known as I frame and the remaining frames are known as P frames. In our technique each frame in a video is considered as a separate input. In first step, the first frame is considered as the input. Then subsequent frames are considered for same.

The main concept behind the fusion strategy is that we derive two input frames from the first frame of video that is I frame, with the aim of improving the visibility and contrast of the frame. Additionally some weight matrixes are used for the derived frames to improve the color cast and visibility. The first derived frame ensures the visibility of the output

mainly in the hazy-free region by using white balancing algorithm. The second derived input frame ensures the chromatic cast by using the contrast enhancing approach. But it mainly focuses on the hazy region of input frame. However by employing these two techniques, the derived outputs have poor visibility. Therefore we mix together the two derived inputs, and filter the important features with the measure of weight. Mainly three normalized weight matrices are used for preserving the region with good visibility. Luminance, chromaticity and saliency are the three normalized weight matrices. These are used for improving the visibility of derived inputs. Finally, to minimize the artifacts introduced with the weight matrix, multi scale fusion pyramidal decompositions is applied. Image fusion is the well known technique that fuses one or more inputs into single one by keeping specific features of output image. Here Laplacian pyramid [15] of the inputs combined with Gaussian pyramid of normalized weights. The per-pixel computation of fusion technique reduces the amount of artifacts. The complexity of our approach is lower when compared to other strategies; because our approach does not estimate the transmission map. So our technique is more reliable than the existing techniques of dehazing.

II. RELATED WORKS

The vision algorithms play an important role in haze removal in image or video. The conventional space invariant filtering techniques fail to recover the degraded images. Restoring hazy images requires specific strategies and therefore an important variety of methods have to be introduced to solve this problem. Several dehazing techniques are developed by employing multiple images or supplemental equipments as input. Such detection is based on the uniform bad weather condition. It is valid for visible and near IR spectra; and weather conditions such as mist, haze, fog and other aerosols. This type introduces the haze optimized transformation with contrast restoration [4]. Such restoration method produces pleasing results. The medium properties give hazy image information. But the main drawback of such contrast restoration is that, it may require number of steps for producing good result. So this method is time consuming and hard to carry out. Another method is the polarization method. It also takes multiple images as input under the different degree of polarization. It takes advantage of the fact that the airlight is partially polarized [5]. In this paper, to estimate the haze effects of scene, different angles of polarized filter are used. The magnitude of polarized haze light component is also estimated. But this polarization method has shown less robustness for scenes with dense haze, because the scene depth is the major degradation factor rather than the angle of polarization light.

Generally depth based method is used for the estimation of scene depth from the multiple images captured under different weather conditions. The Deep photo [6] system is applicable in many situations which are related to the depth value of foggy images. The depth information is obtained by the 3D alignment of the outdoor images. But this method is impractical due to the fact that requirements are changed with the scene quality. The layer estimation of outdoor photographs is also done with this method. To overcome the drawbacks of multiple image dehazing techniques many researchers focus on the single degraded image as input. Tan [7] propose a method which enhances the visibility of degraded image by maintaining their contrast. This contrast maximization technique yield too strong emphasis of the gradients. Fattal [8] presents a new method to estimate the fragmentation of airlight in hazy scenes in a single input image. This paper employs a graphical model of uncorrelated path radiance.

The novel channel prior method is proposed by He et al [9] for single image haze removal. The new prior model is the dark channel prior model, which is based on the key observation; that is, local patches in the haze-free image contain very low intensity pixels in at least one of the color channel. The fast and efficient video enhancement algorithm based on the motion estimation [3] is a new technique, which enhances the scene and object; while decreases additional airlight noise all over the image. This paper states that inverted low lighting video is much like the haze video because in both condition the airlight content will be constant along the entire image. The degraded haze image model proposed by Koschmieder [11] in 1924 describes the degradation of inverted low lighting video. This degradation model proposes that,

$$R(x) = J(x)\, t(x) + A\,(1 - t(x)) \qquad (1)$$

Where, R(x) is the intensity of pixel ‘x’, J(x) is the intensity of original object or scene, t(x) is the transmission medium and the airlight ‘A’, which is present in the input image. The effect of fog in the input video cannot overcome simply by using image processing techniques such as contrast enhancement, adaptive histogram equalization etc. But [2] paper explains the contrast limited adaptive histogram equalization (CLAHE) technique for enhancing the foggy video sequences. This CLAHE limits the pixel intensity to user determined maximum. Before applying CLAHE algorithm the foreground and background images are extracted from the video. The defogged foreground and background videos are fused into a new frame. Finally, the enhanced video sequence is obtained. This technique particularly uses the model and nonmodel methods for their contrast enhancement. And it focuses on the wavelet decomposition, for the fusion purpose.

Our video dehazing multiscale decomposition uses the pyramidal representation of images. It is a video enhancement technique. The main goal of pyramidal representation is that, it develops filter-based representation to decompose images into information at multiple scales, to extract structures/features of interest and attenuate noise. Gaussian pyramid and Laplacian pyramid are the examples of pyramidal decomposition. Several video enhancement composition methods are explained in the [16]. Overview of video enhancement category is included in the figure 1. The Laplacian pyramid transformation [15] uses for

enhancing the encoded and decoded images with its scaling operators. It represents each image as quasi band passed image and each sampled at successive sparer densities. The resulting image has similar structure, and is localized in both space and spatial frequencies. The Laplacian pyramid is a sequence of error images, i.e., L_0, L_1, L_2….. L_N. These are quantized to yield compressed code for pyramidal representation.

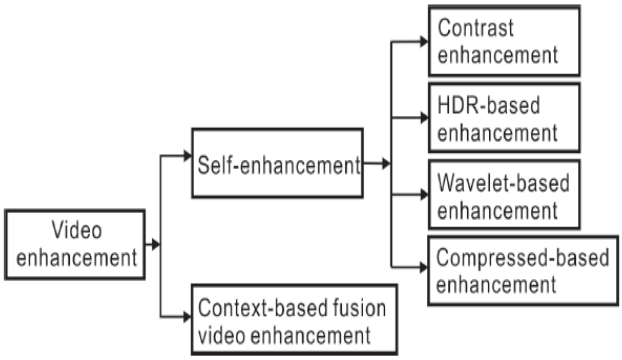

Figure 1. Video Enhancement Category.

III. PROPOSED SYSTEM

The proposed system is the fusion based video dehazing method that employs only the inputs and weights derived from the original hazy video frames (fig 2). Video is a sequence of images (frames). In our technique each frame in a video is considered as a separate input. As a first step, first frame in the video is considered as the input and performed the operation. Similarly subsequent frames are also considered as inputs. The main concept behind our fusion based technique is that; derive two input frames from the original input frame, with the aim of recovering the visibility of video. The first derived input ensures the natural property of output by using an algorithm. The second derived input ensures the contrast properties of output. These two derived input are weighted with some measures, because of the derived input still suffer from low visibility mainly in those region which have dense haze and low light conditions. Hence we introduce three measures (weight maps); luminance, chromaticity and saliency. These weight matrices are designed in a per-pixel fashion, due to define the spatial relations of degraded region.

One of the main effects of our weight maps is to increase the contrast in highlighted and shadowed region. By processing each frame of degraded video, we observed that, the impact of the measure is equally imported in all frames. The first weight matrix mainly focus on the impact on the visibility, the second matrix focus with the colorfulness of the frame and finally the saliency matrix lead to the global contrast of the frame. To yield consistent result, multiply the processed three weight maps. Then the multi scale fusion process employs the inputs, where which are weighted by specific computer measures in order to conserve the most important detected features. In this paper the multi scale pyramid strategy is opted for overcome the degradation problem of native blending solution of fusion process. Generally to yield the fusion result the Laplacian pyramid decomposition of inputs and Gaussian pyramid decomposition of weight maps are combined and performed successively for each 5 pyramidal level.

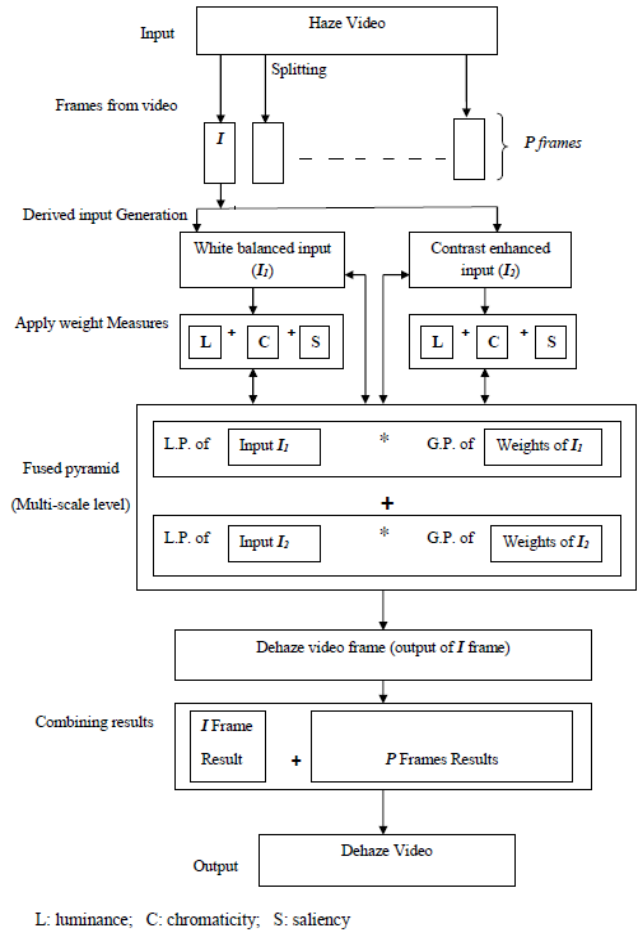

Figure 2. The Proposed Architecture.

The final haze free output is obtained by summing all resulting inputs with the sampling operator. The dehazing technique is applied with all the frames of video. Finally combine these frames into single one, which is the final dehaze video output.

IV. SOLUTION METHODOLOGY

In this section we present the details of fusion techniques that employ only the inputs and weights derived from the original hazy video. In order to derive the video that fulfill the visibility assumption required for the fusion process, we derive an optical model for this type of degradation. By processing appropriate weight maps and inputs, we demonstrate that our fusion-based model is able to dehaze the video effectively.

A. Scattered light propagation model

The exact nature of scattering is slightly complex and depends on the size, orientation, types and media particle distribution; as well as wavelength, polarization states and incident light direction. But in the normal condition the size of air molecules is relatively small compared with the

wavelength of visible light. In this paper we deal with the special atmospheric conditions such as haze. Haze reduces visibility of distant regions by yielding gray hue particles. Haze is traditionally an atmospheric phenomenon. Haze is directly influenced by two atmospheric conditions such as direct attenuation and airlight.

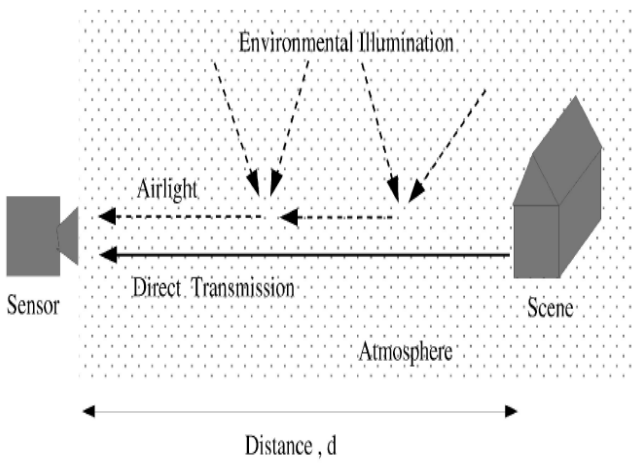

Figure 3. Scattering of light by atmospheric particles can be described by two models: - direct transmission (or attenuation) and airlight.

The attenuation model describes the way in which light get attenuated as it traverses from a scene point to observer. Due to atmospheric scattering, a fraction of light flux is removed from the incident beam. The un-scattered flux called direct transmission is transmitted to the observer. The airlight model explains how a column of atmosphere acts as a light source by reflecting media illumination towards an observer. The observed brightness of captured image in the presence of haze can be modeled based on atmospheric optics. i.e.,

$$I(x) = J(x)\, t(x) + A\,(1 - t(x)) \qquad (2)$$

Where, I(x) is haze image, J(x) is scene irradiance, A is the airlight, $t(x) \in [0, 1]$ is the transmission of the light reflected by the object. The combination of J(x) t(x) is called direct attenuation which describes the scene radiance and its decay in the medium. The second term A (1 – t(x)) is the atmospheric veil, which causes color shifting and distortion in the scene. The goal of haze removal is to find A, J(x) and t(x) from I(x).

B. Inputs

The input generation process seeks to recover optimal region visibility in the frames of video. In practice, there is no enhancing approach that is able to remove the entire haze effect of degraded video inputs.

1. Frames Generation

Video means motion of many images (frames). Our proposed system employs hazed video as input. Our aim is to remove the haze effect from video. So firstly we convert the video into a no of frames based on its length or duration. A frame is assumed as one image of the video. If a particular video have long duration the frame count is high. In our approach each frame is considered as inputs. In the first step, the first frame is considered as input and performs the dehazing operation. In the next step, second frame is considered as input and operate the same function. Similarly each consecutive frame is considered for the same with the aim of dehaze the video.

2. Two inputs creation

Consider the algorithms which generate two inputs from one of the original video frame, which recover the visibility of the image. The first derived input depicts the haze-free region while the second derived input increases visible details of the hazy region. Our first input I_1 frame is obtained by white balancing the original hazy frame I. The main objective of white balancing algorithms is to identify the illuminant color e (λ) or its projection on the RGB. In this step we use Gray World algorithm [34] with the aim of natural improvements of images, by eliminating chromatic casts. The Gray World algorithm estimates that:-

- The average color of a scene is gray.
- The deviation of the average of the image intensities from the expected gray color is due to the illuminant.

For a given image f, the intensity measured can be modeled as,

$$F(x) = \int_{\Omega} e(\lambda,x)\, s(\lambda,x)\, c(\lambda) \qquad (3)$$

Where, Ω denotes the spectrum of visible light, λ is the wavelength of the light, e(λ,x) is spectrum of the illuminant, s(λ,x) is the surface reflectance of an object and c(λ) is color sensitivities of the camera. Gray world hypothesis incorporate with 3 parameters:

- Derivative order, n
- Minkowski norm, p
- Gaussian smoothing, σ

Putting these aspects together, Van de Weijer et al. proposed an algorithm to estimate the color of illuminant 'e' as,

$$k\, e^{n,p,\sigma} = \left(\int |\partial^{n} f^{\sigma}(x) / \partial x^{n}|^{p}\, dx \right)^{1/p} \qquad (4)$$

The integral is computed over all pixels of the image . Where x is a particular position (pixel coordinate), k is a scaling factor, $|\cdot|$ is the absolute value, ∂ is the differential operator and $f^{\sigma}(x)$ is the observed intensities at position 'x', smoothed with a Gaussian kernel σ. Even though the first input frame I_1 shows good visibility in non-haze regions, the second input I_2 frame enhances the contrast in hazy regions. Steps for generating second input frame is:-

- Luminance plane 'l' detection based on color space conversion.
- Calculation of the average Luminance value, $\tilde{l}$.
- The average Luminance value of the entire image I is subtracted from the original image I.

Mathematically, the second input computed for each pixel x is obtained by applying the following expression,

$$I_2(x) = \gamma (I(x) - \tilde{I}) \quad (5)$$

Where γ is a factor that increase linearly the luminance in the recovered hazy regions (default value is $\gamma = 2.5$). These operations improve the visibility in regions degraded by haze, but have some degradation in the rest of the images. However the degradation is detected and solved in our fusion technique by defining proper weight matrices.

C. Weight matrices

The derived inputs then perform three weight measures (weight maps). These matrices are designed in a per-pixel fashion to define the spatial relations of degraded regions in a better manner. Our weight maps balance the contribution of each input and ensure those regions with high contrast or more saliency from a derived input.

1. Luminance Weight Map

Measures the visibility of each pixel and assigns high values to regions with good visibility and small values to the rest. Since hazy frames present low saturation, an effective way to evaluate this property is to measure the loss of colorfulness. This weights evaluation is based on the channel information of RGB color. This weight map is simply computed as the deviation between R, G and B color channels and luminance 'L' from the input:

$$W^k_L=\sqrt{(1/3[(R^k-L^k)^2+(G^k-L^k)^2+(B^k-L^k)^2])} \quad (6)$$

The 'k' indexes the derived inputs. The luminance L is computed by averaging the RGB channels. The luminance weight acts as an identifier of the degradation induced in I_2 in the haze-free regions. Also this measure tends to reduce the global contrast and colorfulness. To overcome the drawback of this technique we define two additional weight maps: a chromatic map (colorfulness) and a saliency map (global contrast).

2. Chromatic Weight Map

This weight map characterizes the frames by a high level of saturation. That is, it controls the saturation gain in the output image. To obtain this measure, for each pixel the distance between its saturation value S and the maximum of the saturation range is computed based on the following equation:

$$W^k_C (x) = \exp (- (S^k (x) - S^k_{max})^2 / 2\sigma^2) \quad (7)$$

Where 'k' indexes the inputs, and we use standard deviation value is, $\sigma = 0.3$ and the constant S_{max} is depends on the HIS color space employed ($S_{max} = 1$). Therefore, small values are assigned to small pixels and saturated pixels get high value. As a result, this map ensures that the initial saturated regions will be better depicted in the final result.

3. Saliency Weight Map

This measure identifies the degree of visible regions with respect to the neighborhood regions. In general, saliency estimates the contrast of image regions relative to their surroundings (based on different image features). For this measure, we use the recent saliency algorithm:-

- Saliency algorithm uses the biological concept of center-surround contrast.
- The saliency weight at pixel position (x, y) of input I^k is defined as:

$$W^k_S (x) = \| I^{\omega hc}_k (x) - I^{\mu}_k \| \quad (8)$$

Where I^{μ}_k represents the arithmetic mean pixel value of the input I_k (a constant value) while $I^{\omega hc}_k$ is the blurred version of the same input that aims to remove high frequency such as noise. $I^{\omega hc}_k$ is obtained by employing a small binomial kernel using the Gaussian filter. After computing the parameters, the saliency is obtained in a per pixel fashion. This map is to produce well-defined boundaries and uniformly highlighted salient regions. These details of saliency map prevent introducing unwanted artifacts in the result image yielded by our fusion technique since neighboring comparable values are assigned similarly on the saliency map. As a result, the effect of this gain is to enhance the global and local contrast appearance.

4. Resulted Weight Maps

By processing a large video the impact of these three measures is equally important. The resulted weights W^k are obtained by multiplying the processed weight maps W^k_L, W^k_C, W^k_S . To yield better results, we normalize the resulted weight maps:-

$$(W^k(x)\tilde{} = W^k (x) / \sum_k W^k (x)) \quad (9)$$

D. Multi-scale fusion

In the fusion process, the inputs are weighted by specific computed maps. Generally the output frame F is obtained by summing the inputs I_k weighted by corresponding normalized weight maps $W_k\tilde{}$. i.e.,

$$F (x) = \sum_k W_k\tilde{}(x) I_k(x) \quad (10)$$

But when this equation introduces strong halos artifacts, mostly in the region of weight maps transitions. To prevent such degradation problems, we use classical multi-scale pyramidal fusion strategy. In our case, each input I_k is decomposed into a pyramid by applying Laplacian operator at different scales. Similarly, a Gaussian pyramid is computed for each weight map $W_k\tilde{}$,. In both Gaussian and Laplacian pyramids have the same number of levels (5 levels), the mixing between the Laplacian inputs and Gaussian normalized weights is performed at each level independently, and gets the fused pyramid:

$$F_l(x) = \sum_k G_l \{ W_k\tilde{}(x)\} L_l\{ I_k (x)\} \quad (11)$$

Where 'l' is the number of the pyramid levels (here l=5) and L {I} is the Laplacian version of the input I while G{W˜} represents the Gaussian version of the normalized weight map of the W˜. This step is performed successively for each pyramid layer, based on the bottom-up manner. The

final haze-free image R is obtained by summing the contribution of the resulting inputs (levels of pyramid):

$$R(x) = \sum_{l} F_l (x) \uparrow^{d} \quad (12)$$

Where $\uparrow^{d}$ is the up sampling operator with factor $d = 2^{l-1}$.

E. Dehaze video creation

The final output dehaze video creation is based on the number of frames, which are operated in the multi-scale fusion strategy. Based on the frames analysis, we obtain the data information and color maps of particular hazed video. By combining each dehaze frames; the final haze- free output is obtained.

V. CONCLUSION

In this paper we introduce a fusion based dehazing technique, which can effectively remove the haze and foggy effects from videos. Our approach removes the haze by simply blending the two derived inputs weighted by several measures. This paper combines the input information in per pixel manner and minimizing the loss of the image structure by a multi-scale strategy. The method is faster than existing video dehazing strategies and yields accurate results.

VI. ACKNOWLEDGMENT

The authors wish to thank the Management and Principal and Head of the Department (CSE) of Ilahia College of Engineering and Technology for their support and help in completing this work.

VII. REFERENCES

[1] Codruta Orniana Ancuti and Cosmin Ancuti "Single Image Dehazing by Multi-Scale Fusion" IEEE Transactions on Image Processing, vol. 22, no. 8, August 2013.

[2] C.Ramya and Dr.S.Subha Rani" Contrast Enhancement for Fog Degraded Video Sequences Using BPDFHE" International Journal of Computer Science and Information Technologies, Vol. 3 (2), 2012, 3463-3468.

[3]Xuan Dong, et al." A Fast Efficient Algorithm for Enhancement of Low Lighting Video" Journal of Information & Computational Science 7: 10 (2010) 2021–2030 Available at http://www.joics.com.

[4] S. Narasimhan and S. Nayar, "Contrast restoration of weather degraded images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 6, pp. 713–724, Jun. 2003.

[5] T. Treibitz and Y. Y. Schechner, "Polarization: Beneficial for visibility enhancement?" in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 525–532.

[6] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," ACM Trans. Graph., vol. 27, no. 5, p. 116, 2008.

[7] R. T. Tan, "Visibility in bad weather from a single image," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2008, pp. 1–8.

[8] R. Fattal, "Single image dehazing," ACM Trans. Graph., SIGGRAPH, vol. 27, no. 3, p. 72, 2008.

[9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1956–1963.

[10] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in Proc. IEEE Int. Conf. Comput. Vis., Sep.–Oct. 2009, pp. 2201–2208.

[11] H. Koschmieder, "Theorie der horizontalen sichtweite," in Beitrage zur Physik der Freien Atmosphare. Munich, Germany: Keim & Nemnich, 1924.

[12] G. Buchsbaum, "A spatial processor model for object colour perception," J. Franklin Inst., vol. 310, no. 1, pp. 1–26, 1980.

[13] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in Proc. 12th Color Imag. Conf., 2004, pp. 37–41.

[14] Zheqi Lin and Xuansheng Wang "Dehazing for Image and Video Using Guided Filter" Open Journal of Applied Sciences Supplement: 2012 world Congress on Engineering and Technology.

[15] P. Burt and T. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Commun., vol. 31, no. 4, pp. 532–540, Apr. 1983.

[16] Yunbo Rao, and Leiting Chen "A Survey of Video Enhancement Techniques" Journal of Information Hiding and Multimedia Signal Processing 2012 ISSN 2073-4212 Ubiquitous International Volume 3, Number 1, January 2012.