An Optimised Multi-Scale Fusion Method And Convolution Neural Networks for Effective Video Dehazing

B. Bhaskar Reddy
Research Scholar,
Department of Electronics,
Sri Krishnadevaraya University, Anantapuramu-515003, A.P, India

Dr. B.Ramamurthy
Professor,
Department of Instrumentation,
Sri Krishnadevaraya University, Anantapuramu-515003, A.P, India

Abstract- The videos shot in the outdoor have severe impairment in terms of visibility if those are shot in presence of fog, haze, smog etc. The degradation in visibility in the video frames not only lacks details for human vision and cause an eyesore but also may cause some of the applications in the field of computer vision like intelligent transportation systems, surveillance systems and object tracking systems to fail miserably. The winter session accidents can be reduced Suitable dehazing technique can contribute to the restoration of the hazy image frames in the video, thereby aiding in improving the video quality. In this paper, it is proposed a novel model to dehaze the hazed frames of hazed videos and compared with state of art models

Keywords – De-haze, Video, Fusion, And Multi-scale Optimal Fusion.

I. INTRODUCTION

Generally, haze in the outdoor and indoor images and videos are primarily present because of the presence fog, mist and dust particles in the atmosphere. Haze in the video frames is caused by attenuation and the air light phenomena. The attenuation is responsible for reduction of scene contrast of a video frame, whereas the air light increases the scene whiteness. The haze results the degradation of images and videos in the visibility of the scene. Degradation in visibility in the video frames not only lacks details for human vision and cause an eyesore but also may cause some of the applications in the field of computer vision like intelligent transportation systems, surveillance systems and object tracking systems to fail miserably. Suitable dehazing technique can contribute to the restoration of the hazy image frames in the video, thereby aiding in improving the video quality. The processed video through any suitable video dehazing algorithm would serve the purpose of improving the video quality thereby making the video suitable for human vision and computer vision applications even.

Single image haze removal has been a challenging problem due to the ill-posed nature of the hazy image. Poor visibility, perceptual image quality and performance of computer vision are the several challenges in image dehazing. Algorithms such as tracking, surveillance, and navigation are primarily functional and successful if the dehazing algorithm is a successful one. A simple but powerful contextual regularization algorithm for the removal of haze from single input hazy imageis proposed in [1]. With the use of the filter banks and scale invariant feature transform operator the haze can be removed. From this the boundary conditions can be estimated manually by increasing the efficiency and dehazing effect. An Image and video dehazing using view-based cluster segmentation is proposed in [2-3]. A view-based cluster segmentation method is utilized to avoid the distortion in sky regions and make the sky and white objects crystal-clear. In this the Gaussian Mixture Model is utilized to cluster the depth based on distant view to estimate the sky region and distant view to estimate is modified to reduce distortion.

It is introduced a novel non-local method for single image dehazing. The method is based on the assumption that an image can be faithfully represented with just a few hundreds of distinct colors. In RGB space, this corresponds to a few hundred tight color clusters. It is showed that in a hazy image, these tight color clusters change because of haze and form lines in RGB space that pass through the airlight coordinate. It is proposed an efficient algorithm to identify these haze-lines and estimate a per-pixel transmission based on them. It is taken into consideration the variance of our estimation in the regularization process, so only pixels that comply with the model assumptions contribute to the result. This method may fail in scenes where the airlight is significantly brighter than the scene [4]. In this paper, it is addressed the image dehazing problem via a multi-scale deep network which learns effective features to estimate the scene transmission of a single hazy image. In the proposed multi-scale model, it is first used a coarse-scale network to learn a holistic estimation of the scene transmission, and then use a fine-scale network to refine it using local information and the output from the coarse-scale network [5].

This paper presents an algorithm [6] to improve images with hazing effects. Usually, the dehazing methods based on the dark channel prior make use of two different stages to compute the transmission map of the input image. The stages are the transmission map estimation and a transmission map refinement. However, the main disadvantage of these strategies is the trade-off between accurate restoration and computational time. The proposed method uses a multilayer perceptron to compute the transmission map directly from the minimum channel and a contrast stretching technique to improve the dynamic range of the restored image. [7-8]

II. PROPOSED WORK

The hazed video is converted to corresponding frames. Each frame is having three colour channels like red, blue and green channels. The major steps to be performed on each frame in the work are followed as

i. Blind Airlight estimation
ii. Optimal Transmission Map Estimation
iii. Refined Transmission Map construction
iv. Multi Scale Transmission Map Fusion
v. Frame radiance calculation
vi. Post Processing using IRCNN

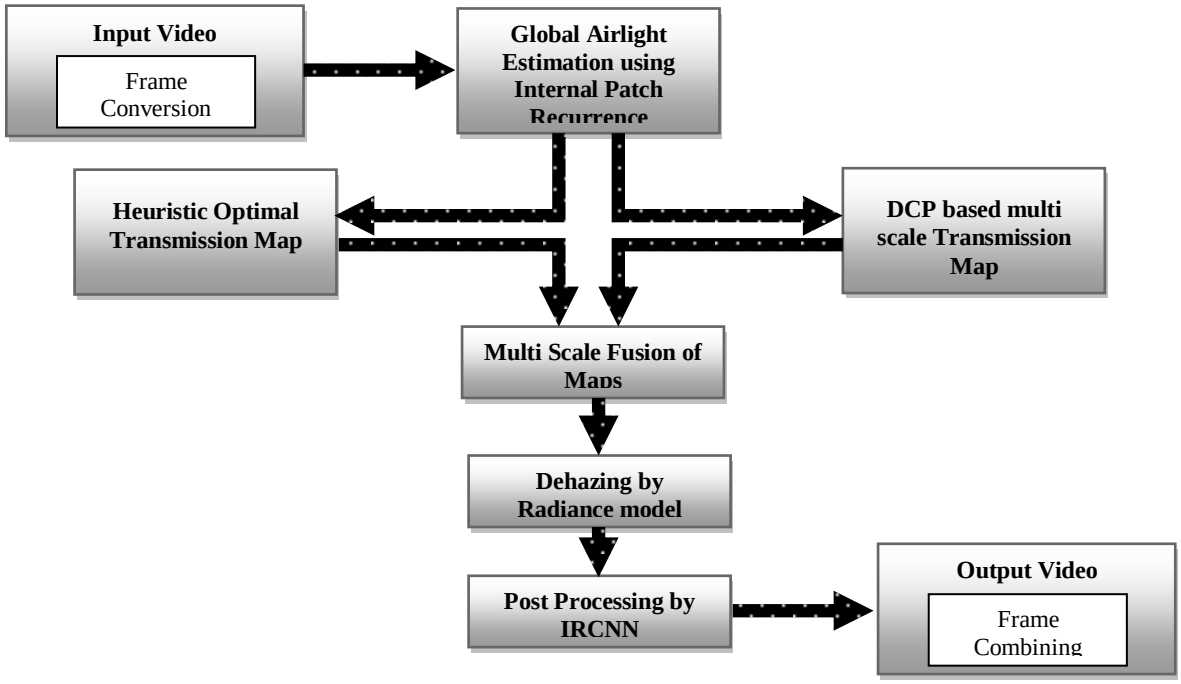

Figure 1. Block Diagram of Proposed System

The complete block diagram of implementation is illustrated in Fig. 1. For every video to be process, we divide the videos into number of frames. Generally fog or haze is considered as brighter pixel, which covers some useful information in the nature. It is very much essential remove or reduces the effect. Patch recurrence airlight estimation is done on each channel of the frame. Frames are nothing but the data structure of images which is captured on corresponding frame duration. So, this data structure is again converted to image format of matrix dimension with their resolution. After this, airlight estimation using internal patch recurrence is performed for any one of the extracted frame image. To reduce the time computation in the system we are performing the airlight estimation only for one frame that may be any of the video. A patch of 5*5 is taken to search for nearest neighbour window to generate optimal transmission map estimation. Any odd patch size window (3*3, 5*5, 7*7,..) can be considered. Average of three channels is used to get optimal transmission map estimation.

The optimised Transmission Map Estimation is done for the last frame of the video, instead of average of all frames. It reduces the computation time. Once this atmospheric light is computed, we estimated the optimal transmission map using transmission heuristic procedure. For the same frame we calculated the number of transmission maps with different number of scales using DCP. A 1*1 patch is used for measuring the gradient map. Gradient map is used to calculate the texture from the taken frame that is used calculate the smoothness. In general, multi scale transmission map, patch wise map and pixel wise maps are fused. Refined transmission maps are calculated from normalized airline values estimated from red, blue and green channels.

These maps are fused with optimized transmission map. Different window sizes are used in refined transmission maps of multi scale transmission map fusion. Totally three refined transmission maps and one optimized map are fused in multi scale transmission map fusion. Here it is considered the sizes of 5*5, 11*11, 17*17 and 23*23 window sizes for refined maps. In multi scale fusion the optimal map, 5*5 sized refined map, 11*11 sized refined map, 17*17 sized refined map, and 23*23 sized refined map. These number of scaled maps can be increased, but at cost of time. It is better to maintain some distance between the window sizes to get better results. Un-even weight delay is considered as 0.138.

These both the results are applied into MOF and obtained the combined and refined transmission map. Radiance image from refined transmission map is estimated to get the dehazed frame output. Finally we can combine all frames and get dehazed video output. Denoising is done with exposure convolution neural network model. The denoising and enhancement can also be done at this level. Convolution neural network improves the visible quality of the frames or video. This method is producing measurable and appreciable results that can be utilized in hazy environments to avoid accidents, monitor critical traffic issues, control shoplifting the vehicles, monitor crimal activities, etc.

III. RESULTS AND DISCUSSIONS

The Proposed method has been tested on different hazed videos with possible parameters and is compared with state of art techniques of the same area. Table 1 illustrates the comparison of quality parameters of video that contains 90 frames. The gamma value is considered here as 0.9. This value is related to brightness of the image. Figure 2 shows the input hazed frame or image of the taken video. Airline estimated values are used to get optimal transmission maps. Figure 3 explains optimized transmission and radiance maps. Figure 3 gives different scaled transmission maps. Figure 5 shows optimal and refined transmission maps that are used in multi scale fusion. Figure 6 illustrates hazed, dehazed and enhanced frames. Figure 7 shows hazed frame and enhanced frame.

TABLE I. COMPARISON OF QUALITY PARAMETERS

Frame No	PSNR	MSE	SSIM	CNR	Gain Visibility
1	58.735	0.087	0.803	1.252	107.443
2	58.845	0.085	0.787	1.288	91.437
3	58.906	0.084	0.800	1.288	75.487
4	58.899	0.084	0.800	1.257	104.116
5	59.061	0.081	0.839	1.406	145.428
6	59.124	0.080	0.816	1.421	124.733
7	59.535	0.072	0.825	1.859	143.280
8	58.825	0.085	0.825	1.205	92.688

9	59.232	0.078	0.806	1.564	166.050
10	59.307	0.076	0.811	1.608	163.516
11	59.055	0.081	0.817	1.402	164.043
12	58.767	0.086	0.830	1.180	138.690
13	58.857	0.085	0.830	1.187	105.545
14	58.898	0.084	0.824	1.281	114.155
15	58.922	0.083	0.841	1.362	169.381
16	58.660	0.089	0.842	1.258	104.292
17	58.883	0.084	0.835	1.402	142.108
18	59.235	0.078	0.848	1.633	70.451
19	59.151	0.079	0.851	1.484	75.583
20	59.212	0.078	0.830	1.540	112.466
21	59.327	0.076	0.873	1.737	70.285
22	59.422	0.074	0.868	1.787	77.311
23	59.301	0.076	0.849	1.610	76.471
24	59.649	0.070	0.886	1.924	73.439
25	59.710	0.070	0.878	2.053	116.445
26	60.067	0.064	0.893	2.665	91.537
27	59.684	0.070	0.876	1.945	84.645
28	60.325	0.060	0.910	2.799	94.730
29	60.294	0.061	0.912	2.462	107.133
30	60.424	0.059	0.919	2.692	92.595
31	60.456	0.059	0.919	2.769	125.343
32	60.366	0.060	0.892	2.607	147.912
33	60.357	0.060	0.919	2.572	89.500
34	60.581	0.057	0.910	2.955	85.400
35	60.573	0.057	0.926	2.667	95.565
36	60.586	0.057	0.922	2.639	102.635
37	60.713	0.055	0.927	2.842	156.283
38	60.624	0.056	0.914	2.557	116.891
39	61.025	0.051	0.909	2.968	120.786
40	60.850	0.053	0.926	2.760	94.735
41	60.745	0.055	0.928	2.578	101.513
42	60.809	0.054	0.931	2.981	74.938
43	60.610	0.057	0.919	2.891	66.717
44	60.635	0.056	0.928	3.020	60.030
45	60.859	0.053	0.942	3.644	78.539
46	60.489	0.058	0.932	2.991	88.458
47	60.434	0.059	0.932	2.974	165.997
48	60.477	0.058	0.918	3.376	113.818
49	60.395	0.059	0.905	3.276	117.271
50	60.349	0.060	0.919	3.223	74.884
51	60.269	0.061	0.925	3.026	89.699
52	60.384	0.060	0.929	3.107	81.325
53	60.427	0.059	0.909	3.103	129.309
54	60.535	0.057	0.930	3.183	109.172
55	60.618	0.056	0.917	3.183	81.174
56	60.469	0.058	0.910	2.695	69.715
57	60.610	0.057	0.920	2.991	80.965

58	60.520	0.058	0.906	2.474	98.993
59	60.703	0.055	0.836	2.801	86.196
60	60.582	0.057	0.882	2.415	71.159
61	60.686	0.056	0.814	2.572	105.571
62	60.781	0.054	0.809	2.609	65.087
63	60.371	0.060	0.806	2.049	79.578
64	60.798	0.054	0.822	2.537	87.787
65	61.024	0.051	0.815	2.873	80.976
66	60.957	0.052	0.841	2.565	104.527
67	60.981	0.052	0.832	2.322	81.120
68	61.073	0.051	0.830	2.361	64.715
69	61.093	0.051	0.827	2.362	71.959
70	61.150	0.050	0.873	2.523	78.330
71	61.106	0.050	0.841	2.362	68.261
72	61.218	0.049	0.854	2.654	77.797
73	61.128	0.050	0.875	2.573	57.534
74	60.921	0.053	0.864	2.383	97.071
75	60.655	0.056	0.869	2.227	126.673
76	60.696	0.055	0.840	2.399	132.261
77	60.507	0.058	0.868	2.087	66.559
78	60.545	0.057	0.838	2.203	72.036
79	60.481	0.058	0.857	2.055	87.320
80	60.839	0.054	0.880	3.023	73.603
81	60.521	0.058	0.862	2.620	59.983
82	60.277	0.061	0.838	2.297	60.234
83	60.022	0.065	0.841	2.042	79.257
84	59.873	0.067	0.849	1.977	83.363
85	59.697	0.070	0.849	1.884	74.905
86	59.672	0.070	0.844	1.868	122.059
87	59.760	0.069	0.843	2.148	71.415
88	59.707	0.070	0.831	2.203	84.303
89	59.637	0.071	0.809	2.200	83.179
90	59.382	0.075	0.817	2.093	71.303
Average	60.121	0.064	0.867	2.299	97.102

Figure 2. Input Haze frame

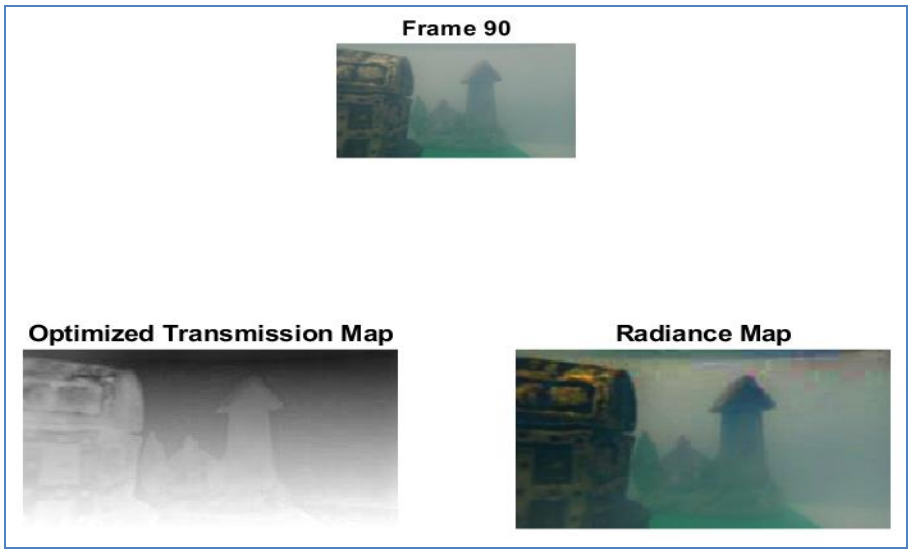

Figure 3. Optimized Transmission and Radiance Maps

Figure 4. Different scaled Transmission Maps

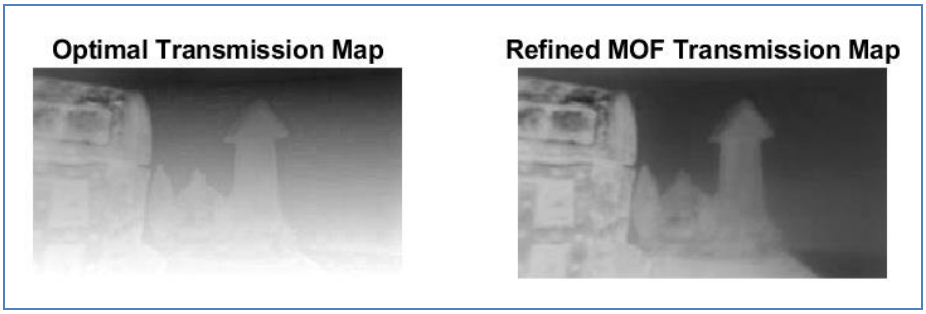

Figure 5. Optimal and Refined Transmission Maps

Figure 6. Haze, Dehazed and Enhanced Frames

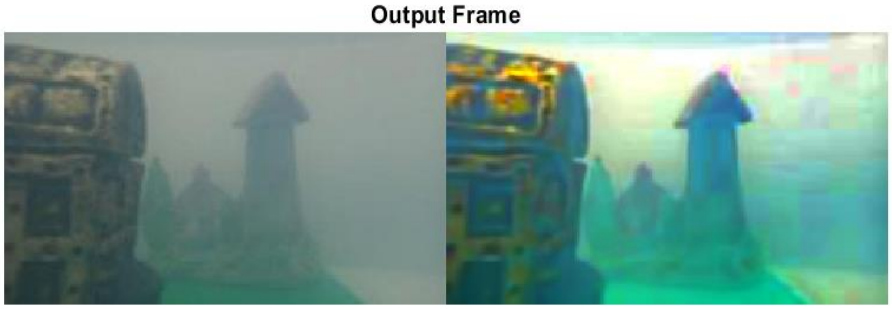

Figure 7. Hazed Frame and Enhanced Frame

Table 2 illustrates the comparison of proposed method with state of art dehazed models. The gamma value is considered here as 0.7. Figure 8 shows Comparison of PSNR value with different dehazed techniques. Figure 9 shows Comparison of SSIM value with different dehazed techniques. Figure 10 shows Comparison of gain visibility value with different dehazed techniques. Figure 11 shows Comparison of CNR value with different dehazed techniques.

TABLE II. COMPARISON OF DEHAZED TECHNIQUES

	PSNR (dB)	SSIM	Haze Level (%)	Haze Intensity (0-255)	CNR	Gain Visibility
Existing Method 1[4]	16.17	0.66	38.59	98.41	1.45	14.10
Existing Method 2 [5]	16.38	0.78	49.17	125.39	1.00	6.25
Existing Method 3 [6]	15.51	0.75	33.69	85.92	1.78	21.85
Proposed Method	63.72	0.92	52.03	132.67	2.21	70.39

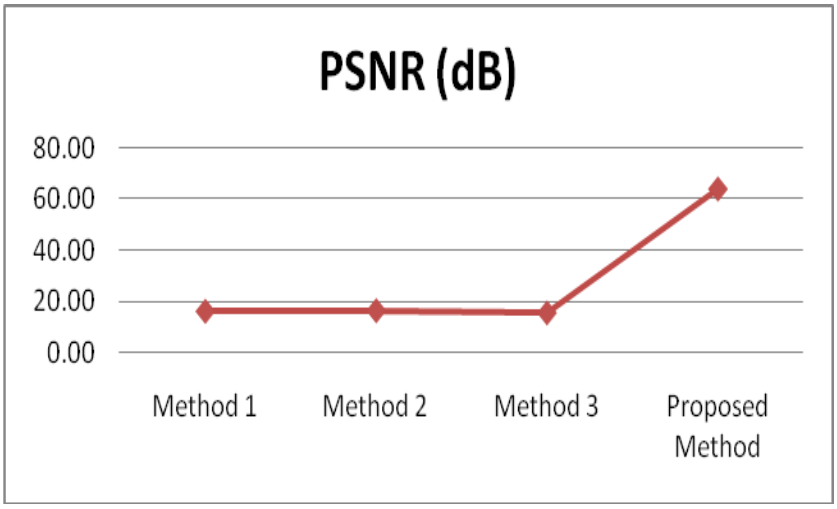

Figure 8. Comparison of PSNR value with different dehazed techniques

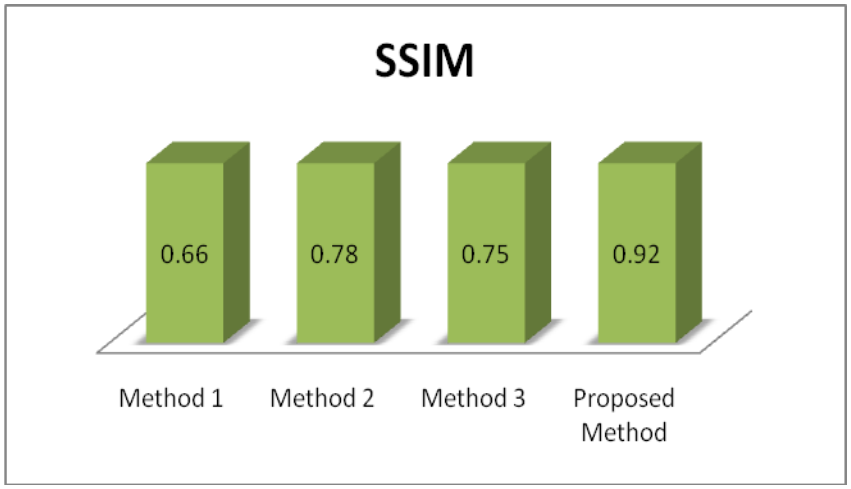

Figure 9. Comparison of SSIM value with different dehazed techniques

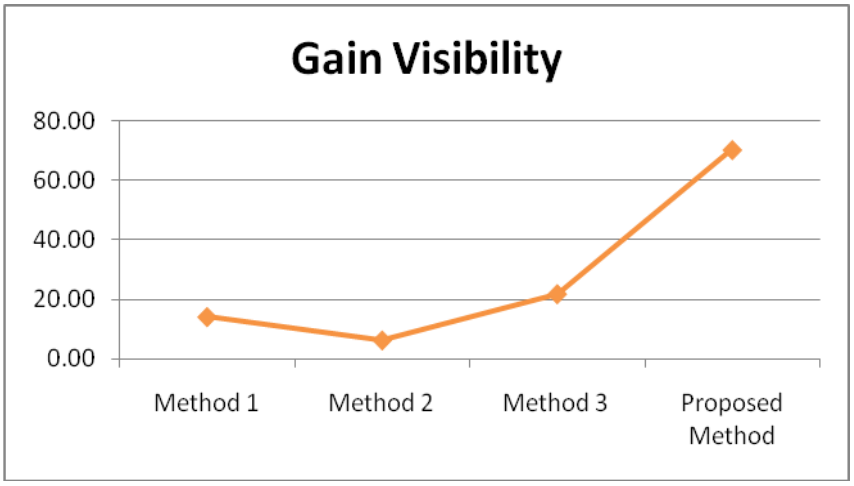

Figure 10. Comparison of Gain Visibility value with different dehazed techniques

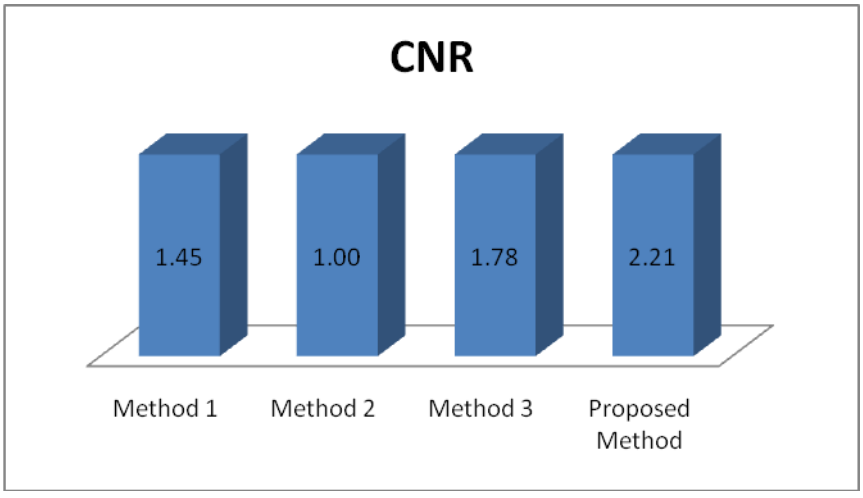

Figure 11. Comparison of CNR value with different dehazed techniques

Table 3 demonstrates the Comparison of quality parameters for different gamma values. Figure 12 shows Comparison of PSNR values with different gamma values. Figure 13 shows Comparison of Haze levels with different gamma values. Figure 14 shows Comparison of Haze intensity values with different gamma values. Figure 15 shows Comparison of Gain Visibility values with different gamma values. The gamma value doesn't give the conclusion of having high or low is good. The gamma value adjustment is required for monitor all types of haze environments (High haze and low haze). It the haze content is low, it is better to operate the equipment with low gamma values and vice versa.

TABLE III. COMPARISON OF QUALITY PARAMETERS WITH DIFFERENT GAMMA VALUES

	PSNR (dB)	**SSIM**	**Haze Level (%)**	**Haze Intensity (0-255)**	**CNR**	**Gain Visibility**
Gamma-0.3	63.67	0.92	51.03	130.12	1.43	39.10
Gamma-0.5	65.61	0.95	50.18	127.96	1.81	51.14
Gamma-0.7	63.72	0.92	52.03	132.67	2.21	70.39
Gamma-0.9	60.12	0.87	57.72	147.20	2.30	97.10

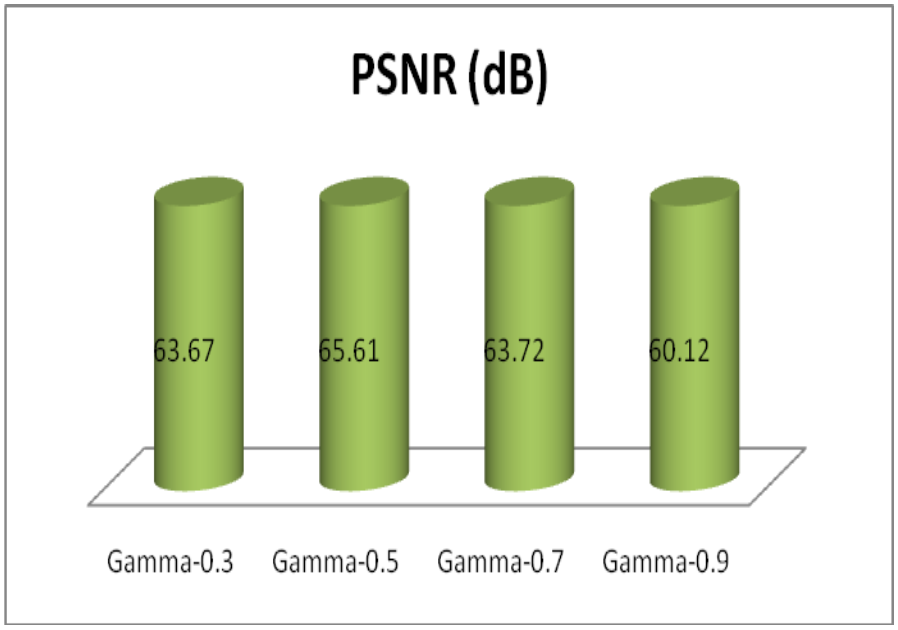

Figure 12. Comparison of PSNR values with different gamma values

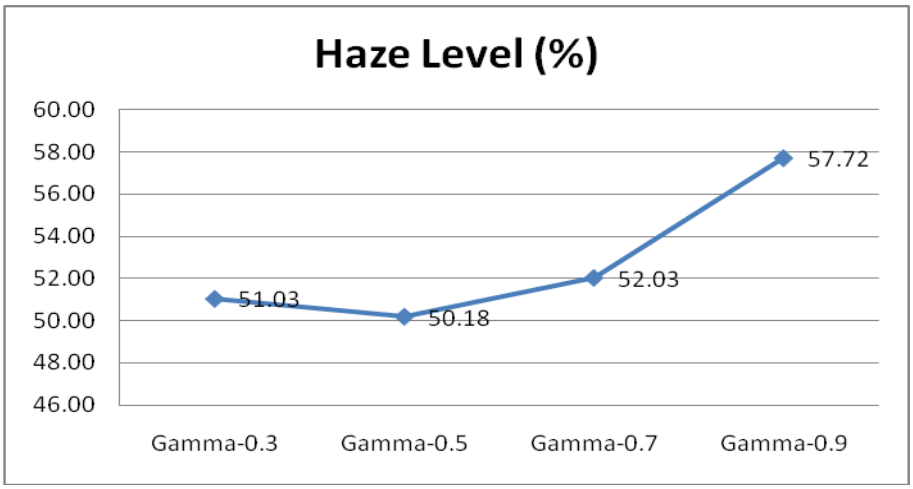

Figure 13. Comparison of Haze levels with different gamma values

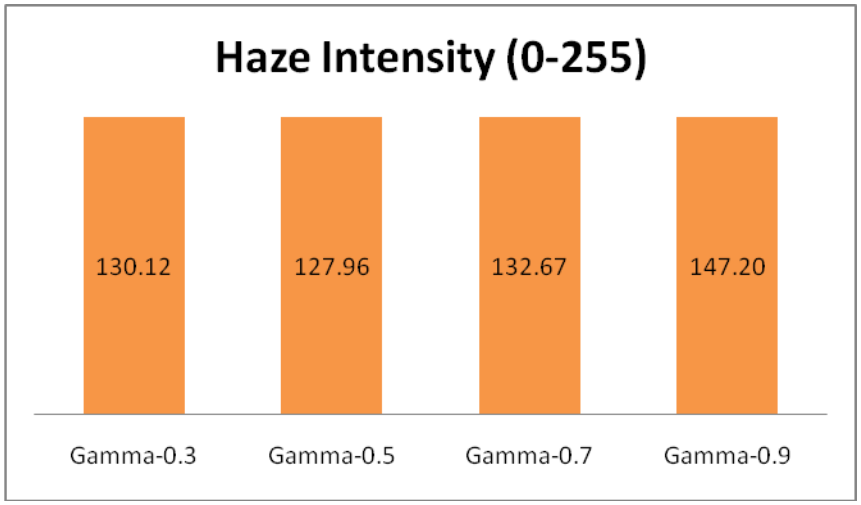

Figure 14. Comparison of Haze intensity values with different gamma values

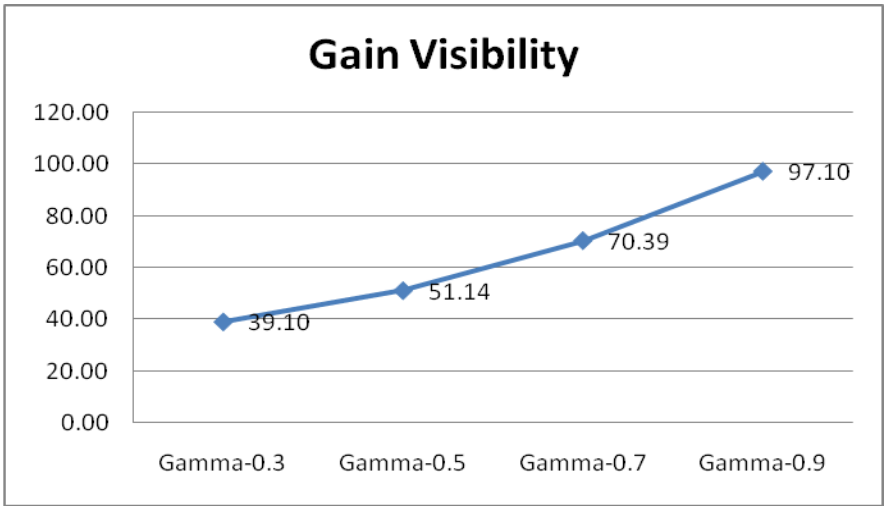

Figure 15. Comparison of Gain Visibility values with different gamma values

IV. CONCLUSION

In this paper, it is proposed a novel method to dehaze the haze component from all frames of a video that is affected. It is shown that the proposed method produced better results among all. The window size plays major role in generation of refined transmission map, which is very much crucial area to in the multi scale fusion. This window size may be changed in accordance with the video environment. The last frame is considered for calculation of optimal transmission map. The first or intermediate frame can also be considered. The gamma value can be changed to get better output enhanced image. The denoising can be added in the neural network stage to denoise the frames, if it is affected. It is final concluded that the proposed method shows a flexible nature that adjust to all type of haze environments that gives remarkable results.

REFERENCES

[1] M. Gokila, Dr. E. Gomathi, and Mrs. J. Rama, "Filter Bank Technique for efficient dehazing in images and videos", International Journal of Applied Engineering Research, Volume 12, Number 1, pp. 242-248, 2017.

[2] Feng Yu, Chunmei Qing, Xiangmin, Xu, Bolun Cai, "Image and video dehazing using view-based cluster segmentation", Visual Communications and Image Processing (VCIP) 2016.

[3] Lucas T. Gonalves, Joel O. Gaya, Paulo Drews-Jr, Silvia, and S. C. Botelho, "Deep Dive: An End-to-End Dehazing Method Using Deep Learning", SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI) 2017.

[4] Dana Berman, Tali Treibitz, and Shai Avidan, "Non-Local Image Dehazing", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp: 1-6, 2016.

[5] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang, "Single Image Dehazing via Multi-scale Convolutional Neural Networks", LNCS, pp. 154–169, 2016.

[6] Sebastian Salazar-Colores, Ivan Cruz-Aceves, Juan-Manuel Ramos-Arreguin, "Single image dehazing using a multilayer perceptron", Journal of Electronic Imaging, Volume 27, Number 4, 2018.

[7] K.Radhika, S. Varadarajan & Y. MMBabu, "Multi spectral classification using cluster ensemble technique", International Journal of Intelligent Systems Technologies and Applications, Volume 17, Issue 1/2, 55-69, April 2018.

[8] Y. M Mohan Babu, M.V. Subramanyam & M.N. Giriprasad, "Despeckling of Medium Resolution Scan SAR data", Advances in Intelligent Systems and Computing, Volume 467, 187-195, October 2016.