Towards Simulating Foggy and Hazy Images and Evaluating Their Authenticity

Ning Zhang, Lin Zhang(✉), and Zaixi Cheng

School of Software Engineering, Tongji University, Shanghai, China
cslinzhang@tongji.edu.cn

Abstract. To train and evaluate fog/haze removal models, it is highly desired but burdensome to collect a large-scale dataset comprising well-aligned foggy/hazy images with their fog-free/haze-free versions. In this paper, we propose a framework, namely *Foggy and Hazy Images Simulator* (FoHIS for short), to simulate more realistic fog and haze effects at any elevation in images. What's more, no former studies have introduced objective methods to evaluate the authenticity of synthetic foggy/hazy images. We innovatively design an *Authenticity Evaluator for Synthetic foggy/hazy Images* (AuthESI for short) to objectively measure which simulation algorithm could achieve more natural-looking results. We compare FoHIS with another two state-of-the-art methods, and the subjective results show that it outperforms those competitors. Besides, the prediction on simulated image's authenticity made by AuthESI is highly consistent with subjective judgements (Source codes are publicly available at https://github.com/noahzn/FoHIS).

Keywords: Fog · Haze · Simulation · Authenticity evaluation

1 Introduction

The decline in visibility due to fog and haze seriously threatens lives of drivers. Fortunately, research on defogging or haze removal has drawn attention during the past decade [1–3]. Accompanied by the rise of deep learning [4], one of the potential topics is to design a haze removal framework using deep architecture [5], while the difficulties of collecting a great number of pixel-wise aligned foggy/fog-free and hazy/haze-free training images hinder the idea. Thus, reliable algorithms of simulating natural-looking fog and haze are urgently needed.

Rendering fog and haze has been widely concerned in the field of computer graphics, and early studies mainly focused on fog and haze effects in virtual scenes. Kazufumi et al. displayed fog effects in outdoor 3D models [6]. Zhou et al. described an analytic approximation to the airlight integral from scattering media to render inhomogeneous fog effects [7]. Anthony and Venceslas modeled the fog function in a B-Spline function basis, and rendered fog in a navigable scene [8]. Although these virtual-scene-based methods can achieve pleasing results, they need to be implemented in pre-established virtual 3D scenes. Also, they require complex settings and massive computing resources.

D. Liu et al. (Eds.): ICONIP 2017, Part III, LNCS 10636, pp. 405–415, 2017.
https://doi.org/10.1007/978-3-319-70090-8_42

Recent research has shifted attention to image-based simulation. Compared with virtual-scene-based methods, simulating fog and haze in single images is not only convenient but also not dependent on expensive computing resources. In view of these advantages, image-based simulation method is also our focus.

1.1 Related Work

Since we focus on simulating foggy and hazy images in this paper, some recent representative studies which are image-based methods will be reviewed here.

Commercial image editing software such as Adobe Photoshop and CorelDRAW Graphics Suite can be used to generate lifelike fog and haze effects. Nevertheless, the process is very complicated and time-consuming. Whether it can produce realistic effects largely depends on the user's skill level. Zhao et al. [9] took advantage of single scattering model to render fog effects into both interior and exterior photos, and they achieved realistic results. However, their method is hard to be implemented because of requiring a great deal of user assistance. Guo et al. [10] estimated the transmission map by using Markov random field model and bilateral filter, then they rendered heterogeneous foggy scenes by Perlin noise. Due to the unreliability of estimating the transmission map, they are unable to achieve satisfactory results all the time. Since depth information plays a pivotal role in simulating fog and haze effects, Liu et al. [11] estimated the depth information by stereo matching. Whereas the different density of depth-aware fog effects are manually controlled by users, which could not guarantee the authenticity of output.

1.2 Our Motivations and Contributions

Having investigated the literature, we perceive that in the field of image-based fog/haze simulation, there's still large room for further research in at least two aspects. For one thing, in reality, the visibilities of two pixels having the same depth but different elevations are not the same. However, all the aforementioned image-based methods took the depth information as the distances from objects to the camera. Thereby, they cannot simulate fog/haze at particular elevations as shown in Fig. 1(b). What' more, when evaluating the authenticity of synthetic images, subjective assessment is usually involved in the experiments, which is effective, but not efficient. It's a great pity that how to measure the authenticity of simulated foggy/hazy images has not been investigated in the literature.

In this work, we attempt to fill the aforementioned research gaps, and the main contributions of our work are summarized as follows.

(1) We propose a framework based on atmospheric scattering model which can simulate both fog and haze effects at any elevation in image, namely *Foggy and Hazy Images Simulator* (FoHIS). Since the distances from objects through specific particle layers to camera are very important, instead of using the depth map as distance values, we estimate the elevation of each pixel in the image using perspective projection transformation and then compute

the distances interacted with the particle layer. Moreover, to simulate the heterogeneous fog effects, 3D Perlin noise is used to present more natural-looking results.

(2) As far as we know, we are the first to design an efficient *Authenticity Evaluator for Synthetic foggy/hazy Images* (AuthESI for short) to objectively measure the simulation results. We get inspiration from no-reference image quality assessment [12,13]. A collection of typical fog and haze features based on natural scene statistic (NSS) are selected and fit them to a multivariate Gaussian(MVG) model. Given a synthetic image, the authenticity is measured by computing a modified Bhattacharyya distance.

The remainder of this paper is organized as follows. Section 2 introduces the procedures of FoHIS. Section 3 presents our novel AuthESI. The experimental results and discussions are arranged in Sect. 4. Section 5 concludes the paper. Source codes are publicly available at https://github.com/noahzn/FoHIS.

2 FoHIS: Foggy and Hazy Images Simulator

2.1 Homogeneous Fog/Haze Simulation

According to [14], the fractional change of radiance I_o passing through thickness of $\mathrm{d}x$ in the direction ω due to absorption and scattering can be described as:

$$\frac{\mathrm{d}I_o(x,\omega)}{I_o(x,\omega)} = (-\beta_a(\omega) - \beta_s(\omega))\mathrm{d}x, \tag{1}$$

where β_a and β_s are absorption and scattering coefficients, respectively. Seeing that the result of absorption and scattering are the loss of visibility, we can combine the two coefficients as β_{ex}, *i.e.,* the extinction coefficient which is related to the visibility range of the atmosphere, V [15]:

$$\beta_{ex} = \frac{3.912}{V} \quad (m^{-1}) \ . \tag{2}$$

Then, in terms of Beer-Lambert law of attenuation [16], the intensity of a beam with the original radiance I_o after traveling a distance d through the particle layer can be expressed as:

$$I = I_o exp(-\beta_{ex} d) \ , \tag{3}$$

also, the opacity can be defined as:

$$O = 1 - exp(-\beta_{ex} d) \ . \tag{4}$$

From Eq. 3 we can notice that when the extinction coefficient is constant, the loss of the radiance intensity is directly related to the distance d. All the previous image-based methods we mentioned in Sect. 1.1 took depth values of the scene as the distances d. Supposing in a plan-parallel atmospheric layer with

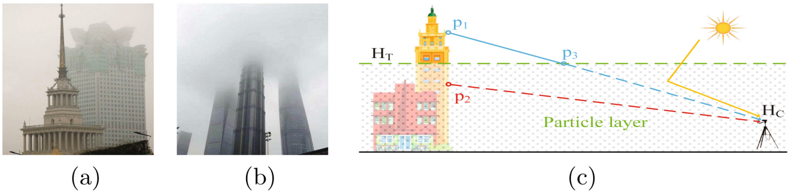


Fig. 1. (a) and (b) are images taken in reality. (a): Different elevations will result in different visibilities. (b): Fog/haze may occur at particular elevation. (c): Suppose the green dotted line denotes the top of the haze. Although p_1 and p_2 have equal depth value to the camera, the distances through the haze and reaching the camera (blue/red dotted line) are not equal. (Color figure online)

thickness H_T, and the elevation of the camera is H_C. As shown in Fig. 1(c), obviously, some pixels in images have large differences in elevations but share the same depth value. In order to obtain more accurate distances from objects through particle layer to camera, we first estimate pixels' elevations in images.

The field of view (FOV) of a camera is a pyramid, and the camera is located on the top of the cone. The pyramid is truncated by front clipping plane (FCP) and back clipping plane (BCP), thus forming a frustum which is called view frustum. Only those objects inside the view frustum are visible. 3D objects can be displayed onto 2D images by perspective projection transformation [17].

Given an RGB image, we segment main objects using GrabCut algorithm [18] and their depth values are set manually. With the depth map and vertical FOV of the camera, we can estimate all the elevations of pixels in the image.

Two terms must be considered if we want to decide the intensity of pixel p in image: 1) I_{ex}, the light reflected by p toward the direction of camera, and attenuated by particles. 2) I_{al}, the light scattered toward the direction of camera by particles. The former term can be expressed as:

$$I_{ex} = I_p exp(-\beta_{ex} r_p), \tag{5}$$

where I_p denotes the color of p in RGB channels, and r_p is the distance through the particle layer and reaching the camera. Since we have acquired the elevations of all the pixels, r_p can be easily derived following the properties of similar triangles. The latter term I_{al} is the color of fog or haze. We combine them as:

$$I = I_{ex} + O_p * I_{al}, \tag{6}$$

where O_p denotes the opacity of p calculated by Eq. 4. We can handle each pixel in three color channels of fog-free/haze-free images by Eq. 6 to simulate fog/haze effects.

2.2 Heterogeneous Fog

Perlin noise [19] is usually used to simulate inhomogeneous fog caused by atmospheric turbulence. [8,10,20] all used precomputed 2D Perlin noise to simulate the inhomogeneous foggy scenes. Due to the different depth values of pixels, it's well-founded to adopt 3D Perlin noise to achieve more natural effect. Given a 2D image and its corresponding depth map, we generate three 3D perlin noise with different amplitude and frequency of parameters, and combine them by:

$$noise = \frac{1}{3} \sum_{i=1}^{3} \frac{noise_i}{2^{i-1}} \ . \tag{7}$$

The computed noise coefficient is multiplied by β_{ex} in Eq. 5 when simulating heterogeneous fog.

3 AuthESI: Authenticity Evaluator for Synthetic Foggy/Hazy Images

In this section, we give an elaborated description of AuthESI, which is designed to objectively measure the authenticity of simulated foggy/hazy images. AuthESI is based on constructing a collection of typical natural scene statistic (NSS) features of fog/haze and fitting them to a multivariate Gaussian(MVG) model. The overall flowchart of AuthESI is shown in Fig. 2.

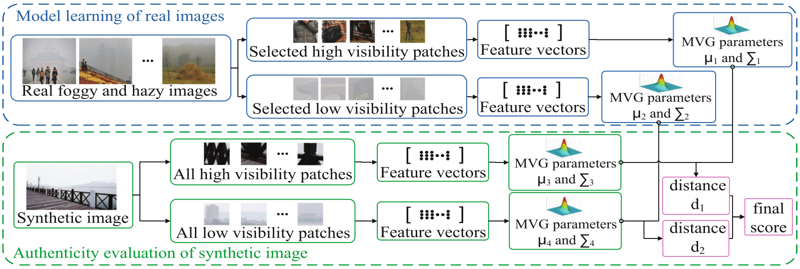


Fig. 2. Process flow of AuthESI.

3.1 NSS Features

Ruderman [21] has pointed out that the luminance of an input gray-scale image I conforms to a Gaussian distribution. The mean subtraction and divisive normalization operators (MSCN) can be computed as:

$$\hat{I} = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}, \tag{8}$$

where i and j are spatial coordinates, and

$$\mu(i,j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} \omega_{k,l} I(i+k, j+l), \tag{9}$$

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K}\sum_{l=-L}^{L}\omega_{k,l}[l(i+k,j+l)-\mu(i,j)]^2} \tag{10}$$

are the local image mean and contrast, where $\omega = \{\omega_{k,l}|k = -K,...,K, l = -L,...,L\}$ defines a unit-volume Gaussian window.

Zhang et al. [22] have found that the log-derivative statistics are effective for analyzing natural images. After logarithmically converting the MSCN, we get:

$$J(i,j) = log[\hat{I}(i,j) + C], \tag{11}$$

where C is a small constant added to avoid numerical instabilities. Then, we compute the seven types of neighbours: $J(i,j+1) - J(i,j)$, $J(i+1,j) - J(i,j)$, $J(i+1,j+1) - J(i,j)$, $J(i+1,j-1) - J(i,j)$, $J(i-1,j) + J(i+1,j) - J(i,j-1) - J(i,j+1)$, $J(i,j) + J(i+1,j+1) - J(i,j+1) - J(i+1,j)$ and $J(i-1,j-1) + J(i+1,j+1) - J(i-1,j+1) - J(i+1,j-1)$.

Both MSCN values and log-derivative values can be modeled using a Generalized Gaussian Distribution (GGD), which is presented by: $f(x : \alpha, \beta) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} exp(-(\frac{|x|}{\beta})^{\alpha})$, where $\Gamma(x) = \int_0^{\infty} t^{(x-1)}e^{-t}\,dt, x > 0$ denotes the gamma function. The variables α and β are shape and scale parameters which can be effectively used to describe the authenticity of fog and haze. Thereby, a 16-features vector is computed.

3.2 Patch Selection

Given an image from the dataset which consists of a total number of 180 real foggy and hazy images. Let the $M \times M$ (M is fixed to 48 in this paper) sized patches be indexed as $P_1, P_2, ..., P_n$. Specific NSS features are then computed from each patch. Since we want to select the features to best express the characteristics of real fog and haze, only a subset of the patches are used. We take the strategy of collecting two sets of patches, S_1 has the highest visibilities and S_2 has the lowest visibilities.

The dark channel prior proposed by He et al. [1] can be defined as $I_{dark}(i,j) = \min_{c \in R,G,B}[I_c(i,j)]$, where $c \in R, G, B$ represents the RGB color channels. Sky, fog or haze regions in photos usually have high value in I_{dark}, On the contrary, regions with high visibilities represent low value in I_{dark}. For each patch P, We first compute the average of dark channel values:

$$A_p = \frac{\sum_{i_p=1}^{M}\sum_{j_p=1}^{M} I_{dark}(i_p, j_p)}{M \times M}, \tag{12}$$

where $I_{dark}(i_p, j_p)$ is the dark channel values in patch P. Next, the binarization is performed in the patch as:

$$B(i,j) = \begin{cases} 0 & G(i,j) < \delta \\ 255 & G(i,j) \geq \delta \end{cases}, \tag{13}$$

where $G(i,j)$ is the gradient magnitude (GM) of $P(i,j)$ computed by *Sobel* operator, and δ is a threshold, which is fixed to 20 in this paper. For each patch P, we record the richness R_p by counting the numbers of 255 in P. If A_p is less than 30 and R_p is more than 400, this patch will be put into S_1. And if A_p is more than 100 and R_p is less than 80, this patch will be put into S_2.

3.3 Evaluate the Authenticity of Synthetic Images

The average of 16-features vector for all patches in S_1 and S_2 are computed separately. Then we use the MVG probability density to fit them:

$$MVG(f) = \frac{1}{(2\pi)^{d/2}|\sum|^{1/2}} exp[-\frac{1}{2}(f-\mu)^T \sum{}^{-1}(f-\mu)], \tag{14}$$

where f is the set of NSS features, μ and $\sum$ denotes the mean and covariance matrix of the MVG model.

Similarly, any synthetic image with fog or haze effect is treated like this and fit their features with two MVG model. But this time, all patches are used and if A_p of patch in Eq. 12 is less than 30, the patch will be put into the set S_3, otherwise, we put them into S_4. The authenticity of the synthetic image is expressed as the sum of modified Bhattacharyya distances:

$$D=\sqrt{(\mu_1-\mu_3)^T(\frac{\sum_1+\sum_3}{2})^{-1}(\mu_1-\mu_3)}+\sqrt{(\mu_2-\mu_4)^T(\frac{\sum_2+\sum_4}{2})^{-1}(\mu_2-\mu_4)}, \tag{15}$$

where $\mu_1, \mu_3, \mu_2, \mu_4$ and $\sum_1, \sum_3, \sum_2, \sum_4$ are the mean vectors and covariance matrices of the natural MVG model and the simulated image's MVG model.

4 Experiments and Discussions

The performance of our proposed FoHIS and AuthESI are evaluated in this section. The experiments were conducted on a PC equipped with an Intel(R) Core(TM) i7-4790 3.60 GHz CPU, 16 GB RAM.

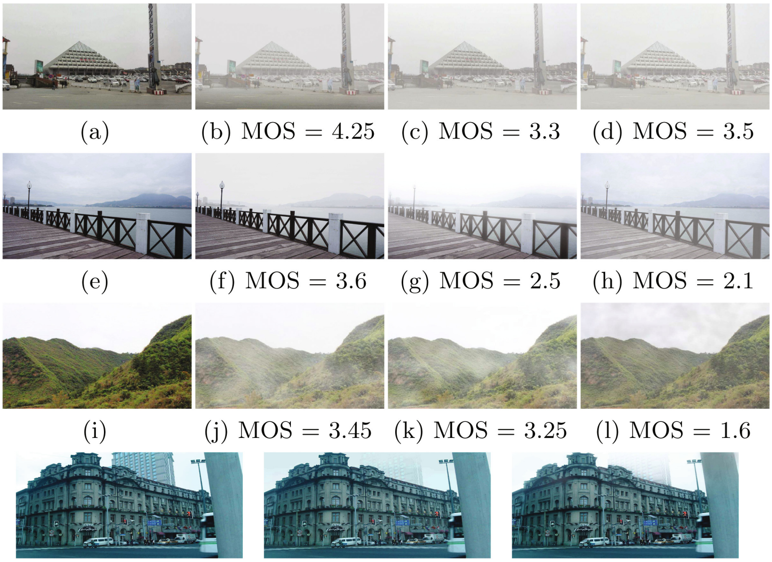
Dataset: We select 18 fog-free/haze-free images with different estimated maximum depth, and each image is used to simulate fog or haze with FoHIS, Adobe Photoshop [23] and Guo's method [10]. FoHIS and Guo's method are implemented with Python. Since Guo's method cannot simulate fog/haze at particular elevations, two of the foggy images were only simulated by FoHIS and Photoshop. Thus, our dataset consists of 52 simulated images. Only 11 of these simulations are displayed in this paper, and their details are listed in Table 1. Please visit our source code page to check the complete dataset.

4.1 Evaluation of FoHIS

In this experiment, we organized the subjective assessment to evaluate the performance of FoHIS. A total number of 20 subjects who were 20 to 25 years old

Table 1. Details about the selected 11 simulated foggy/hazy images.

Input/Output	Maximum depth	Effect	Homogeneous	Particular elevation
(A)/(b), (c), (d)	150 m	Haze	Yes	No
(E)/(f), (g), (h)	800 m	Haze	Yes	No
(I)/(j), (k), (l)	30 m	Fog	No	No
(M)/(n), (o)	150 m	Fog	No	Yes



(a) (b) MOS = 4.25 (c) MOS = 3.3 (d) MOS = 3.5

(e) (f) MOS = 3.6 (g) MOS = 2.5 (h) MOS = 2.1

(i) (j) MOS = 3.45 (k) MOS = 3.25 (l) MOS = 1.6

(m) (n) MOS = 3.4 (o) MOS = 2.3

Fig. 3. For each row, from left to right: source image, FoHIS, Photoshop method and Guo's method [10]. The last row only contains ours and Photoshop method.

were involved in this experiment. According to the recommendations of [24], a LCD monitor with the resolution of 1920 × 1080 pixel was used to display one source image and corresponding synthetic images at the same time. We told the subjects that the first image was taken in real world, and the others (out of order) were simulated by different algorithms. The subjective score ranges from 1 to 5. 5 denotes the highest level of authenticity while 1 denotes the lowest. Each complete scoring operation (52 images) was limited to one hour so as to minimize the impact of fatigue. The benefits of doing so are mainly in three aspects. First, it enables the subjects to evaluate all images by using the same scoring strategy. Second, the relative ranking of the images simulated by different methods can be easily obtained. At last, authenticity comparison across different methods is meaningful in the evaluation of our AuthESI.

The mean opinion score (MOS) is computed by averaging all subjects' subjective scores. Pearson linear correlation coefficient (PLCC) and Spearman's rank-order correlation coefficient (SRCC) are computed as the evaluation criteria. Both coefficients range from [0, 1] with a higher value standing for better performance. The average values of PLCC and SRCC between the subjective scores of individual subjects and MOSs across the dataset are **0.7236** and **0.7014**, respectively. We can discover that the subjects reach a consensus on the authenticity of the synthetic images in the dataset. The MOSs of simulated images are shown in Fig. 3, which demonstrate that our FoHIS outperforms another two methods. Although Photoshop can generate reasonable effects, the results are not stable by reason of depending on user's skill level. Guo's method couldn't accurately estimate the transmission map all the time, hence their algorithm may fail in some images. Table 2 lists time cost of three simulation methods, from which we can see that although FoHIS needs preprocessing step, it runs fastest. When using the same source image to simulate dozens of foggy/hazy images with different extinction coefficients, FoHIS will show its high efficiency.

Table 2. Time cost of different methods for simulating an image. (640×480)

Method	Preprocessing	Homogeneous fog/haze	Heterogeneous fog
Photoshop	—	5 min	20 min
Guo's method	—	50 s	55 s
FoHIS	5 min	1.5 s	6 s

4.2 Evaluation of AuthESI

In this experiment, the performance of AuthESI was evaluated. We first computed the authenticity values (AV) of each simulated image in the dataset using Eq. 15. Then, we computed the PLCC and SRCC values of MOSs against AV across the dataset, which are **0.8124** and **0.8414**, respectively. The high values of these correlations indicate high consistency between prediction on simulated image's authenticity made by AuthESI and subjective judgements. In other words, our proposed innovative AuthESI can effectively evaluate the authenticity of simulated foggy/hazy images.

5 Conclusion

In this paper, a simulation framework FoHIS which can simulate natural-looking fog/haze effects in images is proposed. Meanwhile, an innovative objective evaluator AuthESI is designed and shows effectiveness in evaluating the authenticity of synthetic foggy/hazy images. Our future work may focus on constructing a large-scale dataset of foggy/hazy images and applying it to fog/haze removal.

Acknowledgments. This work was supported in part by the Natural Science Foundation of China under grant no. 61672380 and in part by the ZTE Industry-Academia-Research Cooperation Funds under grant no. CON1608310007.

References

1. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. PAMI **33**(12), 2341–2353 (2011)
2. Li, Y., Tan, R., Brown, M.S.: Nighttime haze removal with glow and multiple light colors. In: IEEE International Conference on Computer Vision, pp. 226–234 (2015)
3. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. IP **24**(11), 3522–3533 (2015)
4. Zhang, D., Zhao, R., Shen, L., et al.: Action recognition in surveillance videos with combined deep network models. ZTE Commun. **14**(S1), 48–54 (2016)
5. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: Dehazenet: an end-to-end system for single image haze removal. IEEE Trans. IP **25**(11), 5187–5198 (2016)
6. Kazufumi, K., Takashi, O., Eihachiro, N., Tomoyuki, N.: Photo realistic image synthesis for outdoor scenery under various atmospheric conditions. Vis. Comput. **7**(5), 247–258 (1991)
7. Zhou, K., Hou, Q., Gong, M., Snyder, J., Guo, B., Shum, H.Y.: Fogshop: real-time design and rendering of inhomogeneous, single-scattering media. In: Pacific Conference on Computer Graphics and Applications, pp. 116–125 (2007)
8. Anthony, G., Venceslas, B.: Modeling and rendering heterogeneous fog in real-time using B-spline wavelets. WSCG **2010**, 145–152 (2010)
9. Zhao, F., Zeng, M., Jiang, B., Liu, X.: Render synthetic fog into interior and exterior photographs. In: 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, pp. 157–166 (2013)
10. Guo, F., Tang, J., Xiao, X.: Foggy scene rendering based on transmission map estimation. IJCGT **2014**, 10 (2014)
11. Liu, D., Klette, R.: Fog effect for photography using stereo vision. Vis. Comput. **32**(1), 1–11 (2016)
12. Zhang, L., Zhang, L., Bovik, A.C.: A feature-enriched completely blind image quality evaluator. IEEE Trans. IP **24**(8), 2579–2591 (2015)
13. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a completely blind image quality analyzer. IEEE SPL **20**(3), 209–212 (2013)
14. Cox, L.J.: Optics of the atmosphere-scattering by molecules and particles. J. Mod. Optics **28**(7), 521–521 (1977)
15. Mahalati, R.N., Kahn, J.M.: Effect of fog on free-space optical links employing imaging receivers. Optics Express **20**(2), 1649 (2012)
16. Lykos, P.: The beer-lambert law revisited: a development without calculus. J. Chem. Educ. **69**(9), 730 (1992)
17. Carlbom, I., Paciorek, J.: Planar geometric projections and viewing transformations. ACM Comput. Surv. **10**(10), 465–502 (1978)
18. Rother, C., Kolmogorov, V., Blake, A.: "GrabCut": interactive foreground extraction using iterated graph cuts. ACM TOG **23**(3), 309–314 (2004)
19. Ken, P.: An image synthesizer. ACM SIGGRAPH Comput. Graph. **19**(3), 287–296 (1985)
20. Zdrojewska, D.: Real time rendering of heterogeneous fog based on the graphics hardware acceleration. Proc. CESCG **4**, 95–101 (2004)

21. Ruderman, D.L.: The statistics of natural images. Netw. Comput. Neural Syst. **5**(4), 517–548 (2009)
22. Zhang, Y., Chandler, D.M.: No-reference image quality assessment based on log-derivative statistics of natural scenes. J. Electron. Imaging **22**(4), 451–459 (2013)
23. TrickyPhotoshop.: Trickyphotoshop - How to Create Mist (Fog) using Photoshop CS6.http://www.youtube.com/watch?v=F9NTddMLSvM
24. ITU Radiocommunication Assembly: Methodology for the Subjective Assessment of the Quality of Television Pictures. International Telecommunication Union (2003)