Chapter 31
A Haze Removal Method Based on Additional Depth Information and Image Fusion

Tian Tian and Bin Zhang

Abstract   To address the problem of image degradation in foggy days, we propose a haze removal method based on additional depth information and image fusion. With recent advances in depth-sensing technology, it has been realized that sensing devices can produce depth images in which the depth value are quite accurate. We adopt the depth estimation dataset of Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) which contains images collected from different real-world environments. The additional information includes the LiDaR scanning points and original depth images which can be used to estimate the optical depth of each point in the scene. In this paper, we investigate how to use additional depth information to remove haze for a single image. Our method focuses on LiDaR depth imaging, image fusion, and the atmospheric scattering model. We use LiDaR scanning points as input and then deduce a rough depth image with prominent features. The rough depth image is then combined with original depth image to improve reliability of depth estimation by image fusion. Using the atmospheric scattering model, we can remove haze for a single image. Experimental results show that our proposed approach provides better performance of dehazing under different fog conditions and holding the details of remote sensing images than current research methods.

31.1   Introduction

Due to low visibility in foggy days, the reflecting and scattering of impurities in the air seriously reduce contrast and clarity of outdoor images causing the visual effect getting awful. As a result, haze removing has become pervasive in applications such as target detection, autonomous driving, and scene recognition. With recent advances in depth sensing technology, it has been realized that sensing devices can produce depth images very easily. Unfortunately, depth-sensing devices often miss data when capturing images which cause the depth value are accurate but incomplete.

T. Tian · B. Zhang (✉)
Xi'an Jiaotong University, Xi'an, China
e-mail: bzhang82@xjtu.edu.cn

The goal of our work is to use additional depth information captured by LiDaR and original depth image to estimate depth of scene and then remove haze for a single image. Though image dehazing has received lots of attention during the recent years, it has generally been solved by different types of methods that recover image by image enhancement or auxiliary information [1–6]. Newer methods have been proposed to remove haze from color images, for example, machine learning [7].

Depth image is widely used as a tool to represent 3D information of scene. Nowadays, with the continuous improvement of depth-sensing technology, additional depth information such as scene depth and multiple images are easy to obtain for practical applications.

According to the different depth sensing devices, collecting scene depth information can be divided into two categories: passive ranging sensing and active depth sensing. [8] The most commonly used method of passive ranging is Binocular Stereoscopic Vision [9]. In this method, two cameras at a certain distance are used to shoot at the same time, so as to obtain two images of the same scene from different perspectives. The corresponding pixel points in the two images are found by the stereo matching algorithm, and the parallax is calculated by the triangle similarity principle and then converted into the scene depth information. However, this method has some limitations on the range and accuracy of parallax map, so the reliability of the scene depth is low. In active depth sensing, the acquisition of depth image is independent of the acquisition of color image. Depth-sensing devices capture 3D information by emitting energy. For example, LiDaR ranging technology can calculate the distance by firing lasers into the space and recording the time interval between the starting point to the surface of objects in the scene and reflecting point back to the LiDaR. It has been widely used in outdoor 3D space sensing system because of its wide range and high accuracy.

31.2   Related Work

The current dehazing algorithms are mainly divided into three categories.

The first category adopts image enhancement to highlight details and improve contrast so as to improve the visual effect of images. There exist algorithms using generalizations of histogram equalization [1] and Retinex [2]. Retinex algorithms have advantages in improving image color constancy and enhancing image details, but it is extremely easy to produce halos when processing images under a condition with a strong light and dark contrast.

The second category adopts image restoration based on auxiliary information, such as using partial differential equation [3], prior information [4] and depth information [5]. Using depth information makes it feasible to deduce the medium transmission and the global atmospheric light which is essential to recover the haze-free images by the atmospheric scattering model [6]. Nonetheless, it has not been used widely because of limitation of depth-sensing devices.

The third category is based on machine learning which focuses on training models for atmospheric conditions [7]. It is worth noting that this method is only suitable for image degradation caused by fog and results in image distortion. In addition, this approach is costly and time-consuming.

Currently, the most simple but effective method is the dark channel priori proposed by He et al. [10]. Since this method cannot deal with the sky region and halo phenomenon very well, it may fail when the haze imaging model is invalid. Moreover, it has a large amount of computation. From the perspective of additional information of scene, Narasimhan et al. proposed a method for scene depth evaluation by discussing the influence law of atmospheric scattering on the contrast of different depths [11]. However, this method only considers the grayscale or color information which causes that the evaluation of depth is not reliable.

31.3   Method

In this paper, we use the additional information of KITTI dataset [12] which includes LiDaR scanning points and original depth images to estimate the scene depth. The process of our method includes LiDaR depth imaging and depth information fusing. First, use scanning points captured by LiDaR to obtain the rough depth image. We project the scanning point onto the original color image. The grayscale pixel value of the projected image can characterize the distance from the camera to surface of the object in the scene. Then, we can obtain a rough depth image with prominent features. Second, aiming at such phenomena as uneven LiDaR scanning points and fuzzy features of objects in original depth images (see Fig. 31.1), we propose to fuse the rough depth image with the original depth image so as to obtain an exact depth image. Finally, calculate the medium transmission from the exact depth image and recover haze-free image using the atmospheric scattering model.
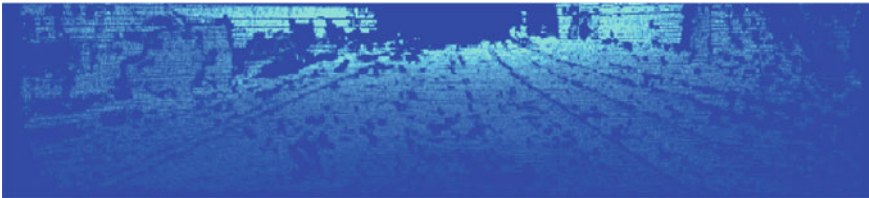

Fig. 31.1   An additional depth image in the scene of KITTI dataset

31.3.1 Haze Imaging Model

In computer vision, aiming at the problem of low visibility in foggy days, Narasimhan [6] explained the process and key elements of imaging by establishing a physical model and proposed that the reasons for image quality reduction include two aspects. On the one hand, it is the energy attenuation caused by the absorption and scattering of reflected light by atmospheric suspended particles. On the other hand, it is the blurring caused by the scattering of ambient light. This model can be described as:

$$E(d, \lambda) = E_0(\lambda)e^{-\beta(\lambda)d} + E_\infty(\lambda)\big(1 - e^{-\beta(\lambda)d}\big) \tag{31.1}$$

where E is the irradiance, d is the distance that the light travels, λ is the wavelength of light, $E_0(\lambda)$ is the illuminance of the light source ($d = 0$), $\beta(\lambda)$ is the total scattering coefficient, and $E_\infty(\lambda)$ is the maximum radiation of atmospheric light. The first term on the right side of (31.1) is the attenuation model of incident light, and the second term is the atmospheric light model.

The atmospheric scattering model is [10]:

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{31.2}$$

where x is the special coordinate of a pixel, I measures the observed intensity, J measures the scene radiance, A measures the global atmospheric light, and t measures the medium transmission describing the portion of the light that is not scattered and reaches the camera.

It can be inferred from (31.2) that the goal of haze removal is to recover J from I:

$$J = \frac{I - A(1 - t)}{t} \tag{31.3}$$

When the fog is uniform, the transmission t is [10]:

$$t(x) = e^{-\beta d(x)} \tag{31.4}$$

where β is the scattering coefficient of the atmosphere and d measures the scene depth. When the transmission $t(x)$ is close to zero, the scene radiance J is on the high side which will cause the image to transition to a white field. Therefore, we shrink the scene depth to a range of 0–1 when using Eq. (31.3) to recover J.

Therefore, the key of dehazing is to estimate scene depth and then substitute the medium transmission into the haze imaging model to recover the target image.

31.3.2   LiDaR Depth Imaging

LiDaR depth imaging is a process in which the scene scanning points projected onto the coordinate system of color image to obtain the rough depth image. The scanning points are stored in the form of point cloud in KITTI dataset.

The process of transforming from point cloud to depth image involves four coordinate systems: world coordinate system (X_w, Y_w, Z_w), laser LiDaR coordinate system (X_v, Y_v, Z_v), on-board camera coordinate system (X_c, Y_c, Z_c), and color image coordinate system (u, v). The same point has the same depth value in the camera coordinate system and the world coordinate system:

$$Z_w = Z_c \tag{31.5}$$

According to the sensor configuration status in KITTI acquisition platform, the installation distance between LiDaR and camera is 0.27 m. The transformation correlation between the coordinate system of laser LiDaR and the vehicle camera is:

$$X_v = Z_c + 0.27 \tag{31.6}$$

The depth value of points in world coordinate system can be deduced by Eqs. (31.5) and (31.6):

$$Z_w = X_v - 0.27 \tag{31.7}$$

The depth value (Z_w) of each point in scene can characterize the gray value of the depth image. The mapping correlation between point $\boldsymbol{M}$ in LiDaR coordinate system and point $\boldsymbol{N}$ in image coordinate system is [13]:

$$\boldsymbol{N} = \boldsymbol{P} \times \boldsymbol{R} \times \boldsymbol{T} \times \boldsymbol{M} \tag{31.8}$$

where $\boldsymbol{P}$ is the corrected projection matrix, $\boldsymbol{R}$ is the corrected rotation matrix, and $\boldsymbol{T}$ is the translation matrix between LiDaR and camera.

Algorithm 1. LiDaR Depth Imaging

Input: LiDaR scanning points
Output: A rough depth image

(continued)

Algorithm 1. LiDaR Depth Imaging

begin

1: Calculate the depth value of points in the world coordinate system:

$Z_w = X_v - 0.27$

2: Calculate the mapping correlation of points between LiDaR and image:

$N = P \times R \times T \times M$

3: Use the pixel value limited in [0, 255] to characterize the depth value of points in image
    coordinate:

cols = gray;

4: Get a rough depth image.

end

31.3.3 Image Fusion

Even though the rough depth images can obtain relatively prominent object features, depth values are sparse. We notice that the original depth images have more scene depth values, while the object features are not prominent. Aiming at the complementary phenomenon, we propose to fuse the additional depth image and the rough depth image based on image fusion algorithms so as to obtain an exact depth image with prominent features.

Considered that pixel-level image fusion algorithms can retain as much detail information as possible which is conducive to further image analysis and understanding, we adopt Haar wavelet transform [14] in pixel-level image fusion. The steps of image fusion are as follows [15]:

- Wavelet decomposition
  Use Haar wavelet transform for original depth image and rough depth image, respectively, to establish multi-scale two-dimensional wavelet decomposition.
- Build a wavelet-pyramid
  Select different coefficients for fusion processing of each decomposition layer to build wavelet-pyramid.
  For the high-frequency part, select the wavelet coefficients with large absolute value as the coefficient because the wavelet coefficients with large absolute value correspond to the detail characteristic of the objects with significant changes in the gray value of depth image.
  For the low-frequency part, select the average value of wavelet coefficients as the coefficient.
- Wavelet reconstruction
  Execute wavelet reconstruction for the wavelet pyramid. The reconstructed image is the fused image which includes more accurate depth value.

Algorithm 2. Image Fusion

Input: Two images (M1 and M2)
Output: A fused image (Y)

---

begin
1: Establish multi-scale two-dimensional wavelet decomposition using Haar wavelet transform:
[c0, s0] = wavedec2(M1, 3, 'Haar');
[c1, s1] = wavedec2(M2, 3, 'Haar');
2: Build a wavelet-pyramid based on different coefficients.
For the high-frequency part, compare the absolute values and then take the larger one as
coefficient:
mm = (abs(MM1)) > (abs(MM2));
Y = (mm.*MM1) + ((~ mm).*MM2);
Coef_Fusion(s1(1,1) + 1:KK(2)) = Y;
For the low-frequency part, take the average value as coefficient:
Coef_Fusion(1:s1(1,1)) = (c0(1:s1(1,1)) + c1(1:s1(1,1)))/2;
3: The image fusion is achieved through wavelet reconstruction:
Y = (mm.*MM1) + ((~ mm).*MM2);
end

31.4   Experimental Results

Our experiments are implemented on a Windows 7 desktop computer with 3.4 GHz
Intel Core i7-6700 CPU, 8 GB (space) RAM, and MATLAB R2013a (64 bit).

In this section, we conduct several experiments to verify the advantages of our
method in dehazing performance on KITTI dataset. Each scene of KITTI consists of
approximately 100 images as an image sequence. We added different concentrations
of fog by photoshop to synthesize hazy image and tested on different images in
respective sequence. In addition, we compared the results with algorithms of He
et al. [10], Zhu et al. [16] (CAP), and Meng et al. [17] (BCCR).

PSNR and SSIM are used as image quality evaluation indexes. The calculation of
PSNR is as follows [18]:

$$PSNR = 10 \times \log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right) \tag{31.9}$$

where MSE is the mean square error between the original image and the processed
image. The larger the PSNR value, the better the defogging effect.

SSIM is used to evaluate the retention degree of image structure information
which can be calculated as follows [18]:

$$SSIM\,(x,\,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{31.10}$$

(a)
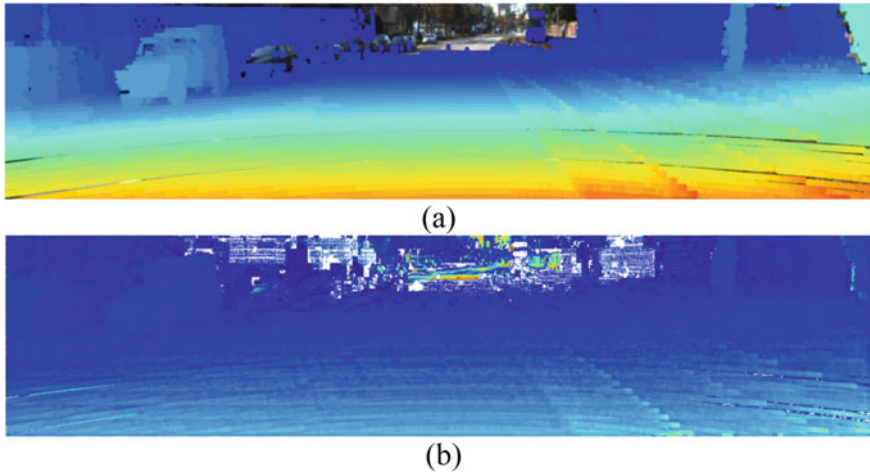

(b)

Fig. 31.2 Results of LiDaR imaging and image fusion

where μ_x and μ_y are mean values of image x and image y, σ_x^2 and σ_y^2 are variances of x and y, σ_{xy} is the covariance of x and y, c_1 and c_2 are both constants. The higher the SSIM value, the better the defogging effect.

In the process of LiDaR depth imaging, we use (31.7) and (31.8) to obtain the rough depth image, as shown in Fig. 31.2a. Due to the limited measuring range of LiDaR, the imaging process cannot be realized in the distant region (120 m away). We use image fusion to obtain the accurate depth image, as shown in Fig. 31.2b. In order to appreciate the accurate depth image, we convert the rough depth image and the accurate image to jet colormap.

We use the accurate depth image to deduce the medium transmission and recover a haze-free image based on the atmospheric scattering model, as shown in Figs. 31.3, 31.4, and 31.5. Figure 31.3 shows results of different dehazing methods at fog condition of 80%, and Fig. 31.4 shows results at fog condition of 50%. Figure 31.5 shows partial results in the same sequence which concludes approximately 100 images at fog condition of 50%. In Fig. 31.3, 31.4, and 31.5, (a) shows the original RGB-image with no haze. (b) shows the responding accurate depth image. (c) shows the hazy image with 80% fog concentration. (d) shows results of our method. (e) shows results using method of He et al. [10]. (f) shows results using method of Zhu et al. [15] (CAP). (g) shows results using method of Meng et al. [16] (BCCR). Table 31.1 shows the image quality evaluation index of different algorithms.

According to the dehazing results in Fig. 31.3, 31.4, and 31.5 and the image quality evaluation index value in Table 31.1, we can see that the average PSNR value (average of 100 images in the same sequence) of our method is significantly higher than others. It demonstrated that our method can effectively improve image clarity and avoid color distortion.
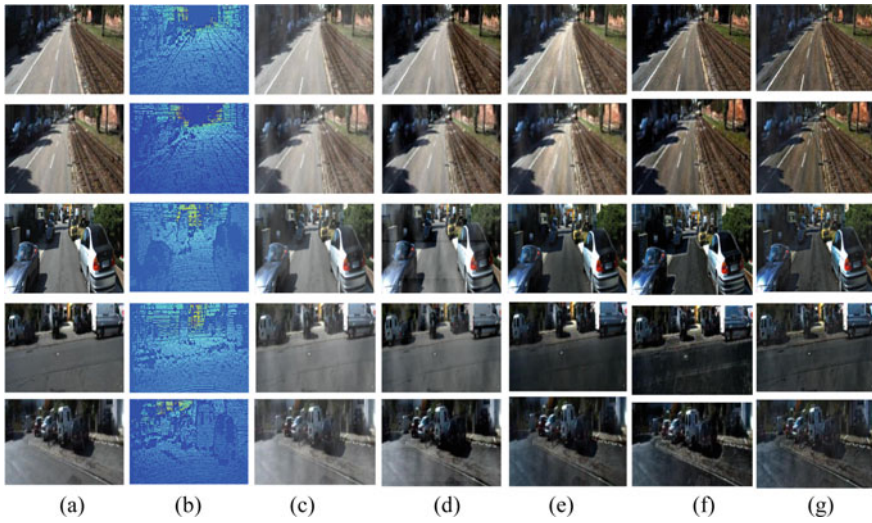
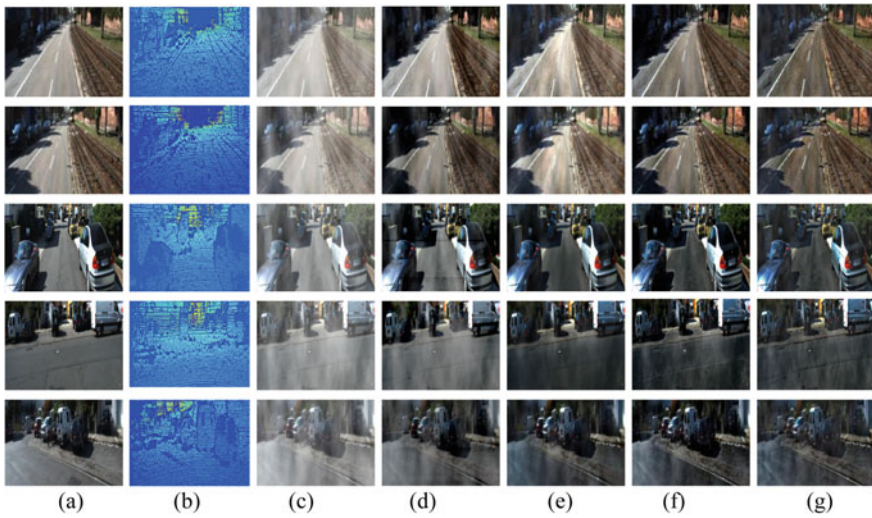**Fig. 31.3** Comparison of dehazing results for images in different scenes with 80% fog concentration

Fig. 31.4 Comparison of dehazing results for images in different scenes with 50% fog concentration

All of the algorithms we mentioned can properly remove the fog, but one main concern is the color distortion (such as surface of road) caused by He et al. [10], Zhu et al. [15] (CAP), and Meng et al. [16] (BCCR). By comparison, the method proposed in this paper can not only effectively remove fog but also avoid image distortion to some extent.
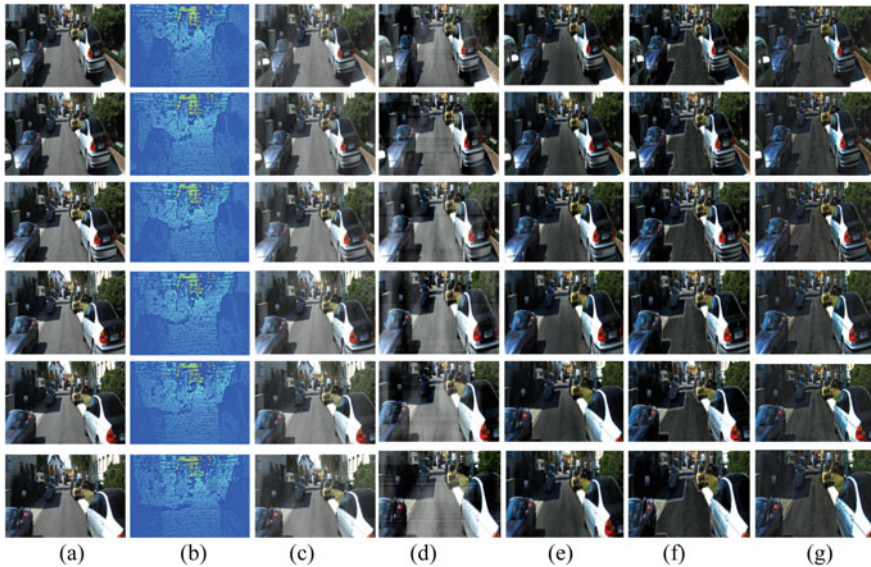
Fig. 31.5 Comparison of dehazing results of an image sequence with 50% fog concentration

Table 31.1 Quantitative evaluation index value of different methods

Metric	Images	HE	CAP	BCCR	Ours
PSNR	No. 006	24.8198	25.2075	20.4942	**27.6063**
	No. 027	24.6488	**24.7741**	21.7419	24.3777
	No. 069	21.6596	22.0607	24.1417	**25.4896**
	No. 112	22.5323	22.2851	21.9657	**25.3335**
	Average	23.4151	23.5751	22.0859	**26.2021**
SSIM	No. 006	0.6827	0.9464	0.9013	**0.9504**
	No. 027	0.7547	**0.9348**	0.9258	0.9286
	No. 069	0.8371	0.8783	0.9256	**0.9347**
	No. 112	0.9259	0.8947	0.9086	**0.9268**

31.5 Conclusion

In this paper, we propose a haze removal method using depth information provided by the LiDaR and original depth image to obtain a more accurate depth image so as to recover haze-free images based on hazing image model. Experimental results show that our method can not only deal with the close-range and foggy scene but also have good performance to improve the contrast and color saturation of the image. Compared with other algorithms, our method has better performance of enhancing the veins feature and holding the details of remote sensing image. Due to the limitation

of the LiDaR scanning mechanism, our method needs to be improved for operating speed and scene images which are long-range or under non-uniform fog conditions.

References

1. Stark, J.A.: Adaptive image contrast enhancement using generalizations of histogram equalization. IEEE Trans. Image Process. **9**(5), 889–896 (2000)
2. McCann, J.: Lessons learned from mondrians applied to real images and color gamuts. In: 7th Color and Imaging Conference, pp. 1–8. (1999)
3. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Instant dehazing of images using polarization. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 325–332. IEEE, Los Alamitos (2001)
4. Thiebaut, E., Conart, J.M.: Strict a priori constrains for maximum—likelihood blind deconvolution. J. Opt. Soc. Am. A-Opt. Image Sci. Vision **12**(3), 485–492 (1995)
5. Kopf, J., Neubert, B., Chen, B., Cohen, M., Cohen-Or, D., Deussen, O., Uyttendaele, M., Lischinski, D.: Deep photo: model-based photograph enhancement and viewing. ACM Trans. Graph. **27**(5), 1–10 (2008)
6. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vision **48**(3), 233–254 (2002)
7. Chen, J., Chau, L.: Heavy haze removal in a learning framework. In: IEEE International Symposium on Circuits and Systems, 1590–1593 (2015)
8. CSDN Homepage. https://blog.csdn.net/zuochao_2013/article/details/69904758. Last accessed 11 July 2019
9. Huang, P.C., Jiang, J.Y., Yang, B.: Research status and progress of binocular stereo vision. J. Opt. Instrum. **40**(4), 81–86 (2018)
10. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)
11. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. IEEE Trans. Pattern Anal. Mach. Intell. **25**(6), 713–724 (2003)
12. Uhrig, J., Schneider, N., Schneider, L., Franke, U., Brox, T., Geiger, A.: Sparsity invariant CNNs. In: International Conference on 3D Vision. IEEE Conference. Qingdao (2017)
13. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: the KITTI dataset. Int. J. Rob. Res. **32**(11), 1231–1237 (2013)
14. Li, H., Manjunath, B.S., Mitra, S.K.: Multi-sensor image fusion using the wavelet transform. Graphical Models Image Process. **13**(16), 51–55 (1995)
15. CSDN Homepage. https://blog.csdn.net/Chaolei3/article/details/80961941. Last accessed 21 July 2019
16. Zhu, Q., Mai, J., Shao, L.: A fast-single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)
17. Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: IEEE International Conference on Computer Vision, pp. 617–624. ICCV, Sydney (2013)
18. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)