


# Construction of computer visual dataset for autonomous driving in sand-dust weather

Hua Zhongwei<sup>1,2</sup> 

<sup>1</sup>Applied Technology College of Soochow University, Suzhou, China

<sup>2</sup>Fudan University, Shanghai, China

co-email: 19110860034@fudan.edu.cn

With the wide application of vision-based autonomous driving and mobile robots, the impact of frequent sand-dust weather on computer vision applications in landlocked countries during spring and autumn has also attracted more and more attention. Although there has been a lot of research on sand-dust image enhancement, no research has been conducted on how to improve the positioning accuracy of vision-based autonomous driving or mobile robots in sand-dust environments, especially because there is currently a lack of data sets to evaluate visual positioning in sand-dust weather. Therefore, we propose a complete set of visual positioning data set construction methods in sand-dust weather to fill the gap in the evaluation data set of application fields such as autonomous driving or mobile robot attitude estimation in sand-dust weather. At the same time, this method is also suitable for the construction of visual positioning data sets under haze and other similar weather. In addition, this paper further demonstrates to readers how to use the converted dust visual positioning data set to conduct positioning evaluation experiment of automatic driving in sand-dust weather.

**Introduction:** In the complex and changeable desert environment, due to the multipath effect caused by the harsh environment, the intermittent satellite signal cannot provide stable and reliable positioning and perception of the surrounding environment information. Stable, efficient and accurate Location-aware is the primary problem to be solved by mobile robots in unknown desert environments. Whether it is the perception or positioning of robots, the rich texture information and color that can be obtained by visual sensors similar to the human eye are always unmatched by other sensors (laser, ultrasound, satellite, inertial navigation, etc.). However, the complex and harsh desert environment poses new challenges for the extraction of visual sensor information and features, especially the frequent sand-dust weather seriously reduces the image quality as the visual SLAM input of the robot, thereby affecting the robot's visual SLAM and other computer vision applications, even the safety of the robot in the desert terrain. In addition, the sand-dust caused by the robot's movement in the desert should not be ignored.

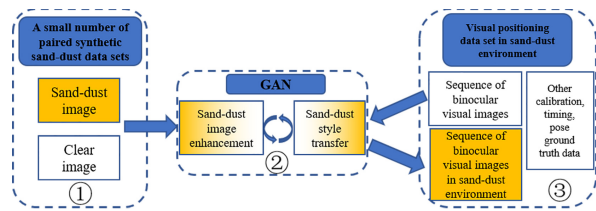
Although our previous work[1] has improved the image quality in sand-dust weather by proposing an enhanced pre-processing algorithm for sand-dust images, limited by the difficulty of collecting visual positioning data in a harsh sand-dust environment, no data sets for assessing visual positioning in sand-dust weather have been produced or published yet. Therefore, we propose a method of visual positioning dataset construction based on image style transfer to fill the gap of evaluation data sets in application fields such as mobile robot pose estimation in sand-dust weather.

The powerful image creation and style transformation capabilities of Generative Adversarial Networks (GAN) make it possible to construct visual positioning data in dust environments. The pix2pix[2] model, which consists of U-net structure as generator and PatchGAN structure as discriminator, is one of the most popular condition-based generative adversarial network models. The network model requires pairs of original images and target images as training data. Through the iterative game between discriminator and generator, the generator finally has excellent image style conversion ability. Therefore, in this section, a large number of existing visual positioning data sets can be converted into sand-dust style data through a trained generative adversarial network, so as to build a visual positioning data set in a sand-dust environment.

However, since there is no unified evaluation standard for the generated dust images, the dust images themselves are fuzzy, noisy and other characteristics, common style transformation clarity indicators are not applicable, and the training model still needs pairs of dust image training data. Therefore, these challenges are again transformed into a dust image enhancement problem based on generative adversarial network for research. As shown in the figure, the reverse application of the

enhanced model trained by small-scale data can realize the transformation of the dusty style of a large number of visual positioning data sets, and facilitate the quantitative evaluation of the enhanced clear images, so as to verify the advancement and effectiveness of the model. On the other hand, unlike traditional image enhancement methods, it is well known that training samples based on data-driven deep learning are difficult to collect, especially in real dust environments, however, artificial image datasets based on the principle of dust image formation are relatively feasible. However in order to reflect the diversity of data, the dust concentration of the composite image may be discrete, and the change is not continuous.

Therefore, the method in this paper is mainly divided into three parts. First, a novel method for constructing a small amount of synthetic sand-dust image data set is proposed. Secondly, the pix2pix network is trained based on a small number of paired sand-dust images, and the corresponding performance evaluation results are presented. Finally, the classic visual positioning data set KITTI is transformed into sand-dust style, which realizes the transformation construction of visual positioning data set in sand-dust environment. In addition, this paper further demonstrates to the readers how to use the converted sand-dust visual positioning data set to carry out the positioning evaluation experiment of automatic driving in sand-dust weather.



**Fig 1** Framework for constructing visual positioning dataset in sand-dust environment

*The proposed method:*

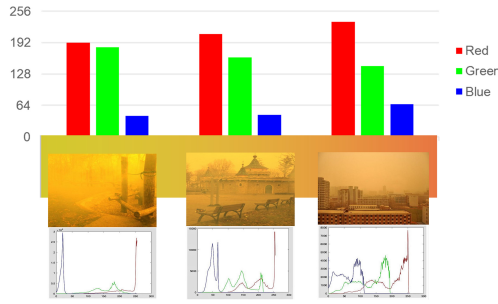
**1. Construct a small number of synthetic sand-dust image dataset:** Considering the similarity with haze images, we combined the sand-dust image forming principle to create an artificially synthesized sand-dust image data set based on the public haze data set O-HAZE[3] which contains 45 different outdoor scenes. The idea of generating sand-dust images from haze images will greatly reduce the complexity of creating artificially synthesized sand-dust image datasets.



**Fig 2** Sand-dust image dataset synthesized based on the haze dataset

However, severe sand-dust weather has a far greater impact on the image than haze. A large number of sand-dust particles are floating in the air. The absorption and scattering of light are very serious. It is found through experiments that the blue channel is most affected by absorption, and the red channel is most affected by scattering. Since the mixed light of green and red is yellow, the sand-dust image often exhibits the yellow or orange color shift. As shown in Fig.2, in the wake of the gap between red and green keeps increasing, the color shift is more towards orange. In addition, the contrast and details of the image will appear blurred. The histogram of the sand-dust image shows that the noise is approximately Gaussian.

So we adjust the red(R), green(G), and blue(B) channel information according to the principle of sand-dust image formation. In order to better simulate the impact of sand-dust in different degrees, we use a ran-



**Fig 3** Color component analysis of sand-dust image

dom function with a limited range when adjusting the color cast of the image to enhance generalization. The formula is as follows:

$$\begin{cases} R' = (1.5 \sim 1.8) \times R \\ G' = (0.9 \sim 1.3) \times G \\ B' = (0.2 \sim 0.5) \times B \end{cases} \quad (1)$$

Finally, considering the low visibility and blurred details of sand-dust images, we blur the adjusted images with different degrees of Gaussian blur and stitch them with the corresponding clear image to form the pair of sand-dust image data sets.



**Fig 4** A small number of paired synthesized sand-dust image data set

**2. Select and train the GAN model:** Then we use the synthesized sand-dust image as the original image, and send the corresponding natural clear image as the target image together to pix2pix model for training, and carry out the performance evaluation after the corresponding image style transfer. The Pix2pix network is made up of a generator and a discriminator. The generator adopts a symmetrical U-Net structure of 8 down-samples and 8 up-samples. Each sampling layer includes the fully connected convolution layer, normalization layer and activation function. The convolution kernel is  $4 \times 4$  and the step size is 2. Leaky ReLU activation function is used under each down-samples layer, the ReLU activation function is used for the up-sampling of the first 7 layers, and the Tanh activation function is used in the last layer. The down-sampling convolution layer of each layer is skip connected with the up-sampling convolution layer located in the same layer. The discriminator is based on the PatchGAN structure, that is, the judge will output a judgment matrix, and each value on the matrix corresponds to the receptive field of the original image. Through the segmentation of the original image, the detailed texture information is guaranteed. According to the test, when the Patch is set to  $70 \times 70$ , the image details and colors have better results.

**3. Transfer into sand-dust style:** The trained pix2pix dust image enhancement model is reversely applied to the classic KITTI[4] dataset. Specifically, the binocular vision image sequence in KITTI data set is sanded style transfer, each image sequence is composed of thousands of images corresponding to the left and right binocular, while retaining the original truth reference track, time series and calibration parameters in

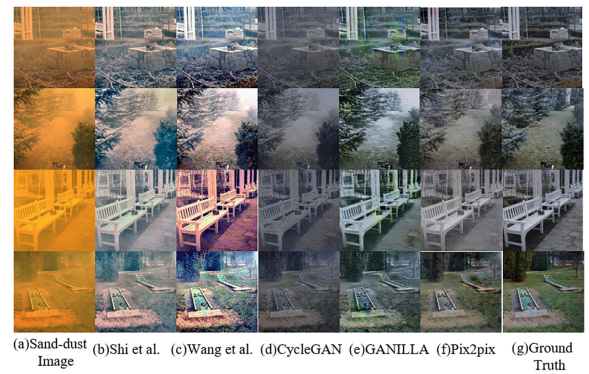
KITTI data set to provide evaluation reference for further visual positioning tests. The following figure shows the effect of the data set after migration and conversion. It can be found that, different from the diverse sand and dust images synthesized in the previous section, the overall style of the converted sand and dust images is consistent and more similar to the characteristics of the images collected in the same period of time under the actual sand and dust environment.



**Fig 5** KITTI dataset transferred by sand-dust style

### Experiment and Evaluation:

**Evaluation of sand-dust image enhancement:** In order to demonstrate the effectiveness and advancement of the style transfer model selected in this paper, the experiments in this section are compared not only with two popular GANs(CycleGAN[5] and GANILLA[6]), but also with two advanced traditional sand-dust image enhancement algorithms(Shi et al.[7] and Wang et al[8]), and through the peak signal-to-noise ratio (PSNR[9]) and structural similarity (SSIM[9]) to quantify the effect. PSNR is one of the most widely full-reference image quality assessment methods based on the difference between corresponding pixels. The larger the value of PSNR, the higher the image quality and the less distortion between the two images. SSIM is another important full-reference image quality assessment method, which evaluates the similarity between two images based on the perspective of brightness, contrast, and structure. The larger the value of SSIM, the higher the similarity between the two images.



**Fig 6** Comparison of the effects of sand-dust image enhancement methods

As shown in the data in Table 1 and Fig 6, the images enhanced by pix2pix model selected in this paper has the best performance in both PSNR and SSIM, and have high image quality and reducibility.

**Evaluation of visual positioning enhancement:** The visual positioning data set constructed in this paper under sand and dust weather can be used to verify the effectiveness of the algorithm in enhancing feature extraction and improving visual positioning accuracy. In this section, the 07 sequence of the sand-dusted KITTI data set is taken as an example, and the classic point feature visual SLAM algorithm ORBSLAM3[10] is taken as a unified visual positioning test algorithm. The sand-dust image sequence, the image sequence after style transfer conversion and the image sequence after sand-dust image enhancement are compared.



Table 1. Corresponding to the experimental results of different methods in Fig. 6

Method	The first row		The second row		The third row		The fourth row	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Shi et al.	14.61	0.78	12.53	0.66	13.82	0.73	12.67	0.71
Wang et al.	12.53	0.66	10.85	0.66	14.96	0.69	11.68	0.56
CycleGAN	19.7	0.55	18.79	0.5	16.69	0.57	18.79	0.56
GANILLA	19.02	0.67	15.22	0.57	19.5	0.68	19.09	0.62
Pix2pix	<b>21.59</b>	<b>0.62</b>	<b>19.12</b>	<b>0.6</b>	<b>19.95</b>	<b>0.64</b>	<b>20.91</b>	<b>0.66</b>

Table 2. The result of the extracted feature points of the corresponding frame in Fig. 7

	(a)Original	(b)Sand-dust	(c)enhanced[2]	(d)enhanced[1]
Left image	<b>268</b>	164	149	247
Right image	98	76	125	<b>155</b>

Table 3. Visual Enhancement Positioning Absolute Error (RMSE) in sand-dust environment in Fig. 8

Seq	(a)Original		(b)Sand-dust		(c)enhanced[2]		(d)enhanced[1]	
	Trans	Rot	Trans	Rot	Trans	Rot	Trans	Rot
07	<b>0.45</b>	0.47	0.85	0.51	1.27	0.96	0.69	<b>0.41</b>

The following figure shows the image frames captured during the experiment and the corresponding ORB feature point extraction results.

ORB-SLAM3 is a feature point extraction that converts color images directly into gray images. Although the great influence of dust color bias is removed, it can still be seen from the data in the table that the influence of sand-dust environment on image extraction features is still significant. It can be seen from the data in Table 2 that after our previous enhanced[1] preprocessing, the feature extraction capability of the image is restored and even exceeds that of the original image.

In addition, the EVO evaluation tool can also be used to calculate the RMSE (root mean square error) of absolute position translation error and absolute attitude rotation error on all key frames. The smaller the RMSE, the higher the positioning accuracy of the algorithm. As shown in Fig 8, the trajectories of various colors represent the positioning trajectories in different states, and the dotted line is GroundTruth, which shows that some of the red trajectories (our previous work) are even closer to the true value than the blue ones(original data).

**Conclusion:** we propose a complete set of visual positioning data set construction methods in sand-dust weather and demonstrate to readers how to use the converted dust visual positioning data set to conduct positioning evaluation experiment of automatic driving in sand-dust weather.

**Acknowledgments:** This research was funded by Natural Science Foundation of the Higher Education Institutions of Jiangsu Province, China. (NO.22KJB520034).

© 2023 The Authors. *Electronics Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## References

1. Zhongwei Hua, Lizhe Qi, Min Guan, Hao Su, and Yunquan Sun. Colour balance and contrast stretching for sand-dust image enhancement. *IET Image Processing*, 16(14):3768–3780, 2022.
2. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

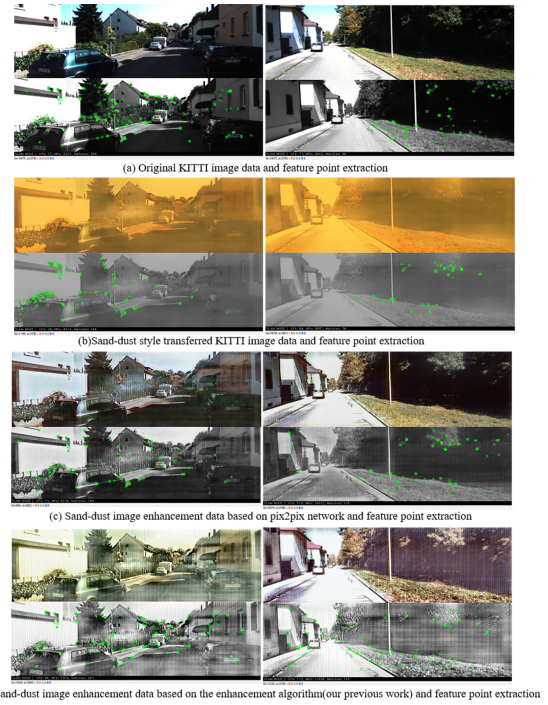


Fig 7 The effect of image extraction feature points after enhancement

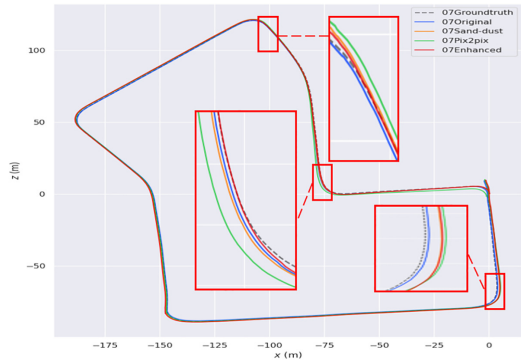


Fig 8 Comparison of visual positioning trajectories after enhancement

3. Codruta O Ancuti, Cosmin Ancuti, and Christophe DeVleeschouwer. O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 754–762, 2018.
4. A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *IEEE Conference on Computer Vision & Pattern Recognition*, 2012.
5. Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.
6. Samet Hicsonmez, Nermin Samet, Emre Akbas, and Pinar Duygulu. Ganilla: Generative adversarial networks for image to illustration translation. *Image and Vision Computing*, 95:103886, 2020.
7. Zhenghao Shi, Yaning Feng, Minghua Zhao, Erhu Zhang, and Lifeng He. Normalised gamma transformation-based contrast-limited adaptive histogram equalisation with colour correction for sand-dust image enhancement. *IET Image Processing*, 14(4):747–756, 2020.
8. Bo Wang, Bowen Wei, Zitong Kang, Li Hu, and Chongyi Li. Fast color balance and multi-path fusion for sandstorm image enhancement. *Signal, Image and Video Processing*, 15(3):637–644, 2021.
9. Z. Wang. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004.
10. Carlos Campos, Richard Elvira, Juan J Gómez Rodríguez, José MM Montiel, and Juan D Tardós. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. *IEEE Transactions on Robotics*, 37(6):1874–1890, 2021.