An Efficient Video Desnowing and Deraining Method with a Novel Variant Dataset

Arezoo Sadeghzadeh[1], Md Baharul Islam[1,2][0000-0002-9928-5776], and Reza Zaker[3][0000-0001-8708-2077]

[1] Department of Computer Engineering Bahcesehir University, Istanbul, Turkey
arezoo.sadeghzadeh@bahcesehir.edu.tr
[2] College of Data Science & Engineering, American University of Malta
bislam.eng@gmail.com
[3] Department of Electrical Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran
zaker@azaruniv.ac.ir

Abstract. Video desnowing/deraining plays a vital role in outdoor vision systems, such as autonomous driving and surveillance systems, since the weather conditions significantly degrade their performance. Although numerous approaches have been reported for video snow/rain removal, they are limited to a few videos and did not consider the variations that occurred for the camera and background in real applications. We build a complete snow and rain dataset to overcome this limitation, consisting of 577 videos with synthetic snow and rain, quasi-snow, and real snow and rain. All possible variations of the background and the camera are considered in the dataset. Then, an efficient pixel-wise video desnowing/deraining method is proposed based on the color and temporal information in consecutive video frames. It is highly likely for a single pixel to be a background pixel rather than a snowy pixel at least once in the consecutive frames. Inspiring from this fact along with the color information of the snow pixels, we extract the background pixels from different consecutive frames by searching for the minimum gray-scale intensity. Experimental results demonstrate and validate the proposed method's robustness to illumination and high-performance static background and camera.

Keywords: Desnowing · Deraining · Temporal information · Snow and rain dataset· Static/dynamic background· Synthetic/quasi snow.

1 Introduction

Weather conditions such as rain, snow, fog, and haze have a negative effect on the perceptual quality of the videos/images captured from outdoor video/image processing or vision systems such as video/movie editing, vision-based navigation, autonomous driving, and video surveillance [4]. Consequently, the performance of the related video/image processing tasks such as object tracking and detection [17] and disparity estimation [9] is significantly degraded. Providing an example, the falling snowflakes in Fig. 1 severely distort the results of the static image-based object detection [17,5] (Fig. 1(a)) and the video-based stereo matching for disparity map estimation (Fig. 1(b)) [9]. Thus, it is vital to remove these weather conditions from the input images and videos as a pre-processing step, which has attracted much attention among researchers in the computer vision field.

Earliest rain removal approaches were mostly based on the correlation between pixels in consecutive frames [2], and physical characteristics of rain such as shape, appearance, brightness, etc. [3]. However, these methods cannot achieve good performance in complex scenes since no prior knowledge of videos is employed. Later, this limitation was overcome by considering the prior knowledge of the rain in [18]. Recently, various deep learning-based approaches have been proposed for rain removal using deep convolutional neural networks (CNNs) or deep recurrent convolutional neural networks (RCNNs).

One of the main challenges in the field of video-based snow and rain removal is the lack of appropriate dataset. In earlier works, some rainy scenes of the movies [2] or real snowy videos from YouTube [9,11] have been employed for the implementations. However, using these kinds of datasets cannot be suitable for quantitative analysis due to the lack of ground-truth information. In some other approaches, synthesis snow and rain have been added to the clean videos [11]. The characteristics of the snowflakes and rain streaks in these datasets are limited. At the same time, there can be varieties of scenarios for them with different sizes, velocities, and densities of snowflakes and rain streaks with different states (static and dynamic) for the cameras and the background. Additionally, these videos with synthetic snow and rain are not entirely the same as the real snowy, and rainy videos as their movement and scattering patterns may differ from the real ones.

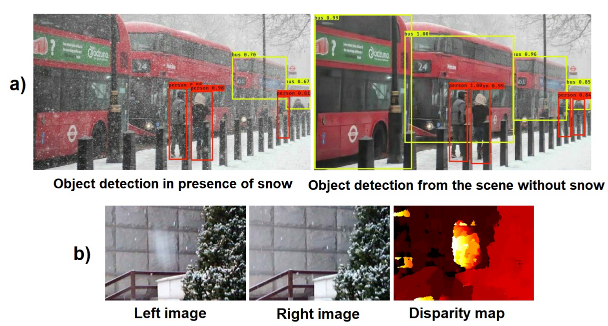


Fig. 1. The negative effects of snow on the performance of computer vision applications, a) object detection based on YOLOv3 [17,5], b) stereo disparity map estimation [9].

Another challenge in the video-based snow removal approaches is the performance degradation of the video rain removal techniques applied to the snowy videos. Complicated characteristics of snow such as sparse scattering, uneven density, multi-scale shapes, and irregular transparency make video-based snow removal more challenging and inapplicable than video-based rain removal. Thus, applying rain removal techniques on snowy videos cannot remove the large bright snowflakes. Moreover, these algorithms cause severe blurring artifacts even for static backgrounds when the camera slightly shakes or moves while capturing the videos.

To address the aforementioned challenges, we build a complete dataset of snow and rain videos that are employed in our implementations. Then, a simple but efficient video desnowing/deraining method is proposed based on the temporal information of the adjacent frames and the intensity of the corresponding pixels in consecutive frames. Our dataset can be used in future research to have a complete evaluation and comparison of the proposed method. Overall, the main contributions of the this work can be summarized as follows:

- A new snow and rain videos dataset is provided for the first time. This dataset consists of 577 videos in which there are three types of particles: synthetic snow and rain, quasi snow, and real snow and rain. All variations of background (static and dynamic) and the camera (static and dynamic, e.g., translation, zooming, illumination changes, and rotation) are considered. The ground-truth information in the synthetic snow and rain videos allows the researchers to evaluate their methods in different scenarios. This dataset can be useful for all future video snow/rain removal approaches.
- A simple but very efficient video desnowing/deraining method is proposed based on the temporal information and the color of the pixels in consecutive frames of the video. This pixel-wise method can successfully remove the snow and rain for static background and camera even if there is heavy snow with high density while its computational cost is low. It is also robust to illumination changes and camera shaking.

2 Related Works

In this section, a brief review of the reported approaches in the literature for rain and snow removal is presented. These approaches can be mainly categorized into two groups: 1) image-based methods, and 2) video-based approaches. In both groups of the approaches, two major types of techniques have been applied, i.e. conventional computer vision techniques, and techniques based on deep neural networks (DNNs). Some of the main recent works are discussed in the following subsections.

2.1 Image-based Snow/Rain Removal

Considering a video in frame-by-frame manner, the image-based approaches can also achieve satisfying performance for video-based rain/snow removal. However, due to the lack of the temporal information in image-based techniques, video-based methods achieve significantly better performance for video rain/snow removal. As the first attempts in the field of image-based atmospheric-particle-removal approaches, several priors, e.g. sparsity prior, patch-rank prior, have been applied in order to detect and remove the particles. However, they suffer from the limited generalization ability which has been dealt with using the same-resolution CNNs based on the synthetic datasets.

Fu et al [1] utilized a combination of CNNs and handcrafted priors for removing rain particles from a single image. A large synthetic snowy image dataset was published in [15] by Liu et al which was called Snow100K. They presented a multistage network architecture as DesnowNet and evaluated it qualitatively and quantitavely on the Snow100K dataset. However, this model is time-consuming with bad generalization ability and a very large size which made it inapplicable to the light-weight applications.

A two-stage network was proposed by Li et al [13]. In this method, the first step was a physics-based backbone and the second step was a depth-guided generative adversarial network (GAN) refinement. Recently, Jaw et al [5] proposed a simple, efficient single-image desnowing model based on a pyramidal hierarchical design and cross-resolution lateral connections.

2.2 Video-based Snow/Rain Removal

Prior to the success of the deep learning-based methods, conventional computer vision and image processing technique-based approaches have been widely adopted for video- and image-based rain/snow removal.

Most of the video deraining approaches are based on the high correlation between the corresponding pixels in consecutive frames [2]. One of the earliest works in this field was proposed by Garg and Nayar [2] by capturing the dynamics of rain based on a correlation model and a physics-based motion blur model. Later, they proposed a method [4] to further reduce the impacts of rain before taking images/videos by adapting the parameters of the camera, e.g. exposure time. In another work, the sparsity of rain streaks along with the rain-perpendicular direction were considered by Jiang et al [7] to propose a video rain streak removal approach based on tensor. Low-rank hypothesis of the background was employed by Ren et al [18] to separate sparse and dense rain and deal with heavy rain/snow in dynamic scenes. In addition to deraining based on monocular videos, some studies have been carried out for deraining based on stereo videos [9] which can be applied widely in outdoor vision applications such as autonomous driving systems. These approaches include more information and details than the monocular videos thanks to the information provided from two cameras. However, the necessity of using two cameras results in additional cost.

Due to extremely fast improvements in deep learning models, e.g. CNNs and RCNNs, and their ability in investigating complex patterns, they have been extensively applied in variety of applications [6,17] as well as rain/snow removal systems. Liu et al [14] proposed a deep RCNN for rain removal based on spatial texture appearances. In this method, the background detail was reconstructed by considering temporal coherence. A two-stage recurrent network was presented by Yang et al [20] through which the inverse recovery process of the main synthesis model was performed for deraining. Additionally, they provided a new rain synthesis model in order to create more videos for training and evaluation.

In some recent studies [10,11,12], multi-scale convolutional sparse coding has been applied for rain/snow removal from both dynamic and static background scenes. Li et al [11] proposed a method for dynamic background, in which the rain/snow was encoded based on an online multi-scale convolutional sparse coding (OMS-CSC) model. However, it still suffers from some limitations on handling the videos from non-surveillance cameras and videos with extensive moving objects, fast illumination changes, and fast moving cameras.

3 Dataset Development

Recently, a video dataset including synthetic snow and rain has been proposed in [11] which mainly focused on rain rather than snow. Additionally, the variations on the camera and background settings were not considered in this dataset. An image-based snow removal dataset, namely Snow100K [15] is available. Recently, the Canadian Adverse Driving Conditions (CADC) video dataset has been proposed in [16] which was collected during winter within the Region of Waterloo. It was specifically designed for autonomous driving in a variety of winter weather conditions. However, as it is the same as real snowy videos, it suffers from the lack of ground truth information. To the authors' best knowledge, there is no complete unique dataset for rain/snow removal from videos, especially for snow. We build a complete snow and rain dataset with distinguishable characteristics, called **Variants in Snow&Rain Videos**[1]. This dataset comprises three main sections: 1) videos with synthetic snow and rain, 2) videos with quasi-snow, and 3) videos with real snow and rain.

3.1 Videos with Synthetic snow and Synthetic rain

Ground-truth Videos Only two scenarios are considered for videos with synthetic snow/rain in the available datasets, i.e., static background and dynamic background. In our dataset, different background and camera settings and variations are considered in 46 recorded videos. We use a 25MP camera to record these videos with a frame rate of 30. These frames are saved in "JPEG" format as images with a size of 1080x1920 pixels. For the dynamic background, different scenarios are considered, such as slow movements (i.e., the tree leaves are swinging), fast movements (i.e., cars on the street or the movements of a swing in the park). The camera is also either dynamic or static. The dynamic camera refers to the scenarios as slow and fast translation, slow and fast rotation, slow and fast zooming, small and large shaking, slow and fast illumination changes, and simultaneous zooming and rotation. Different possible combinations of background and camera variations are also utilized, and the videos are recorded as ground truth. Additionally, some more videos are recorded from other locations such as the sea, roads, and parks to make the dataset rich in various locations

[1] The dataset is available at `https://bit.ly/3BHKeRo`. For any issues regarding downloading the dataset, please contact `arezoo.sadeghzadeh@bahcesehir.edu.tr` or `bislam.eng@gmail.com`.

Table 1. Total characteristics of the videos and the considered scenarios in the developed dataset.

<table>
<tr><td colspan="5">Ground-truth videos for synthetic snow and rain</td></tr>
<tr><td>Camera Resolution</td><td>Camera Variations</td><td>Background</td><td>Number of Videos</td><td>Video Length</td></tr>
<tr><td rowspan="6">25MP
1080x1920 pixels
fps=30</td><td rowspan="2">Static</td><td>Dynamic (slow)</td><td>3</td><td>35s, 14s, 13s</td></tr>
<tr><td>Dynamic (fast)</td><td>4</td><td>8s, 8s, 10s, 8s</td></tr>
<tr><td>Dynamic (shaking and translation)</td><td>Dynamic (slow)</td><td>1</td><td>14s</td></tr>
<tr><td rowspan="3">Static and Dynamic (slow and fast translation, zooming, rotation, illumination changes, small and large shaking, both rotation and zooming)</td><td>Dynamic (fast)</td><td>13</td><td>11s, 11s, 14s, 18s, 11s, 10s, 13s, 29s, 9s, 12s, 11s, 13s, 11s</td></tr>
<tr><td>Dynamic (slow)</td><td>13</td><td>14s, 21s, 14s, 13s, 18s, 10s, 10s, 20s, 10s, 34s, 11s, 9s, 13s</td></tr>
<tr><td>Static</td><td>12</td><td>22s, 11s, 18s, 11s, 25s, 12s, 37s, 11s, 10s, 14s, 18s, 11s</td></tr>
<tr><td colspan="5">Videos based on quasi-snow</td></tr>
<tr><td rowspan="2">8Mp
480x460
fps=30</td><td>Static</td><td rowspan="3">Static</td><td>2</td><td>59s, 23s</td></tr>
<tr><td>Dynamic (translation)</td><td>1</td><td>74s</td></tr>
<tr><td>25MP
1080x1920
fps=30</td><td>Static and Dynamic (slow and fast translation, zooming, rotation, illumination changes, small and large shaking, both rotation and zooming)</td><td>13</td><td>25s, 19s, 12s, 10s, 26s, 16s, 16s, 10s, 19s, 28s, 11s, 26s, 13s</td></tr>
<tr><td colspan="5">Videos with real snow and rain</td></tr>
<tr><td>13MP
1080x1920
fps=30</td><td>Static</td><td>Static</td><td>1 (snow)</td><td>8s</td></tr>
<tr><td>25MP
1080x1920
fps=30</td><td>Static</td><td>Static and Dynamic</td><td>8 (light and heavy rain)</td><td>21s, 7s, 21s, 11s, 12s, 42s, 11s, 18s</td></tr>
</table>

with a quiet or crowded background in nature and daily life. These scenarios and characteristics of the ground-truth videos are summarized in Table. 1.

Augmentation for Synthetic Rain and Snow Once the ground-truth videos have been recorded, the synthetic snow and synthetic rain are added to those videos using an augmentation library in Python as "image" [8], i.e."iaa.Snowflakes" for synthetic snow and "iaa.Rain" for synthetic rain. Ten different scenarios for snowflakes form the synthetic snow dataset: 9 combinations of snowflake sizes from small to big (i.e., 0.7, 0.85, and 0.95) and speed from slow to fast (i.e., 0.001, 0.01, and 0.03), and a scenario with randomly chosen snowflake size and speed in the range of [0.7, 0.95] and [0.001, 0.03], respectively. As in reality, rain has very random scattering. The synthetic-rain dataset is formed by one scenario of raindrops, randomly choosing raindrop size and speed in the range of [0.1, 0.2] and [0.1, 0.3], respectively. Consequently, in this section, in addition to 46 ground-truth videos, 46x10=460, and 46x1=46 videos have been created for synthetic-snow and synthetic-rain, respectively. Some samples of the synthetic-snow dataset are illustrated in Fig. 2 (a).

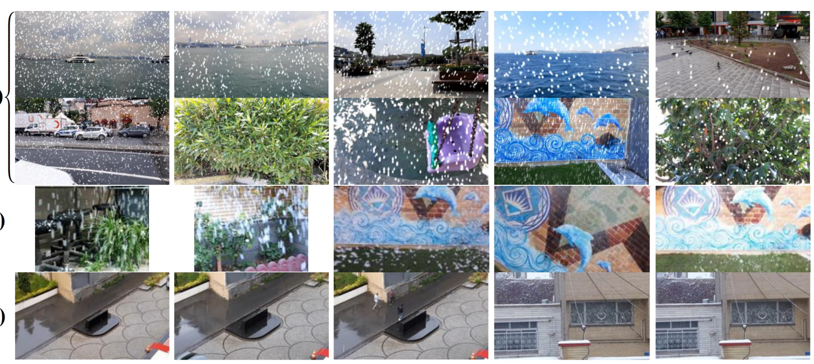
a)
b)
c)


Fig. 2. The sample videos of the developed dataset with a) synthetic-snow, b) quasi-snow, and c) real snow and rain for different scenarios.

3.2 Videos with Quasi-snow

Although the videos with synthetic snow are really rich in the possible variations for background and camera status, the shape and scattering pattern of the snowflakes cannot be as similar as those of the natural snowflakes. Therefore, in this part of the dataset, we simulate the real snowflakes by using the snow spray, which produces snowflakes with the most similar shape and scattering pattern and velocity to those of real snowflakes. A total of 16 videos was created with a static background and dynamic camera. All available scenarios are summerized in Table. 1 and some sample videos are illustrated in Fig. 2 (b). For the static background and camera, the first frames can be considered as ground truth.

3.3 Videos with Real Snow and Rain

There are many available videos for real rain and snow, only 9 videos are included in our dataset, 8 for real rain, and 1 for real snow. Different scenarios of real videos are listed in Table. 1 with some sample videos illustrated in Fig. 2 (c).

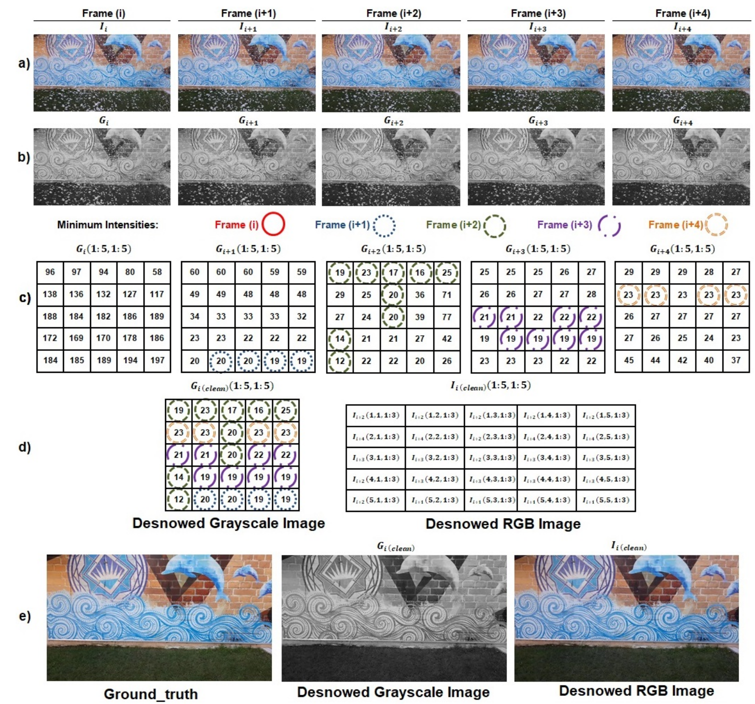


Fig. 3. The overall flowchart of the proposed desnowing/deraining method.

4 Proposed Method

Based on the studies, the rain streaks and snowflakes have sparse scattering, randomly distributed in the atmosphere. It is observed that rain streaks and snowflakes appear at different locations in temporally adjacent frames. Hence, it is unlikely for a single pixel to contain rain/snow in several consecutive video frames. Based on this fact, a simple but efficient method is proposed for snow removal using the temporal information of the adjacent frames and the color of the snowflakes. Generally, rainfall and snowfall on a digital image or a video frame have very similar characteristics, which causes desnowing techniques to be applied for deraining scenarios simultaneously and vice versa. The overall flowchart of the proposed method is illustrated in Fig. 3. In this method, after dividing the input video into frames, for ith frame of the video (Ii), a few consecutive frames (e.g. four frames in Fig. 3 (a)) after that frame are selected and converted to grayscale images (Fig. 3(b)). In this step, the grayscale intensities of every single pixel in consecutive frames are compared. As the color of snowflakes and rain streaks are almost white, their intensity value is 255 or near 255. Therefore, for a single pixel in consecutive frames, the background pixel has an intensity less than the pixel's intensity including snowflakes, or at most near to the intensity of that pixel if the background is white.

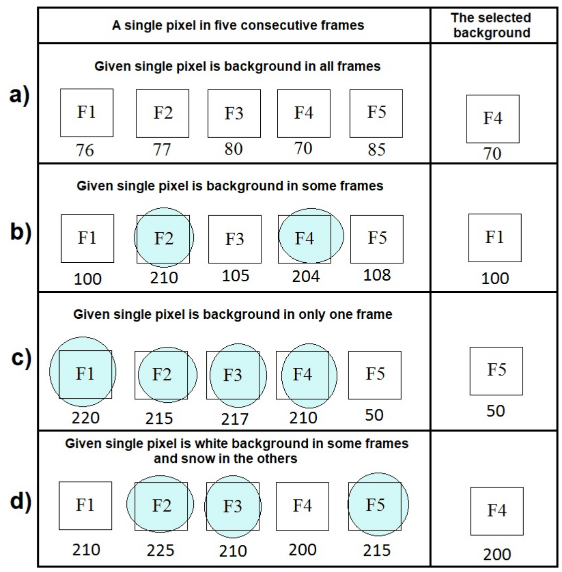


Fig. 4. The possible conditions for a single pixel in five consecutive frames.

On the other hand, based on the temporal information of the snowy videos, it is considered that for a given single pixel, it is clean of snow at least in one frame among a few consecutive frames. Accordingly, the pixel with the minimum intensity is selected as the background. This idea is investigated for different scenarios for a single pixel in five consecutive frames of a video in Fig. 4. In Fig. 4 (a), the same pixel in all five consecutive frames is assumed to be without any snowflakes. So, all can be considered as the background pixel, and selecting the minimum intensity is also a background pixel. Another scenario is illustrated in Fig. 4 (b), in which the same pixel in some frames contains snowflakes, and in some others, it does not. As illustrated, the snowy pixels have grayscale intensities near 255, and the clean pixels have intensities less than 200. Hence, selecting the pixel with minimum intensity results in choosing the background pixel.

In one of the most challenging scenarios, in which only one pixel is not snowy, the background can be easily detected based on the proposed method and selecting the pixel with the least intensity (Fig. 4 (c)). Another essential scenario occurs when the background is white, as illustrated in Fig. 4 (d). It is the most straightforward scenario because the snowy pixel is almost white. In this case, the background can be detected by selecting the pixel with minimum intensity because the intensities of all pixels are nearly 255, representing the white color. This pixel-wise selection is illustrated in Fig. 3 (c) for the first 5x5 pixels of the input images (left top part of the images) in five consecutive frames of a video. As it is presented, for example, for the pixel of (1,2), the minimum value belongs to the third frame. So, the grayscale intensity value of that pixel in the final image equals the grayscale intensity of that pixel in the third frame. As shown in Fig. 3 (d), this process is repeated for all pixels, which results in the desnowed grayscale image. Since the RGB desnowed video is required, the final desnowed image is extended to three RGB channels, in which, for example, for the given pixel of (1,2), the intensities of R, G, and B channels in desnowed image equal to those the same pixel in the third frame.

Fig. 5. Qualitative performance of the proposed method for static background and camera on quasi-snow videos, real-snow and rain videos, and synthetic-snow and rain videos.

5 Experimental Results and Comparison

5.1 Performance Evaluation

In this section, all the experiments have been implemented on a PC with 32G RAM and i7 CPU. As was mentioned above, one of the critical parameters is the number of consecutive frames considered in the analysis. This factor is essential because it defines the least number of consecutive frames among which all pixels appear at least once as background. Implementing different values for this parameter, we observe that 3, 7, and 10 frames are sufficient for light, heavy, and weighty snow/rain. In the following experiments, seven consecutive frames are considered for completely removing the snow/rain. The proposed method

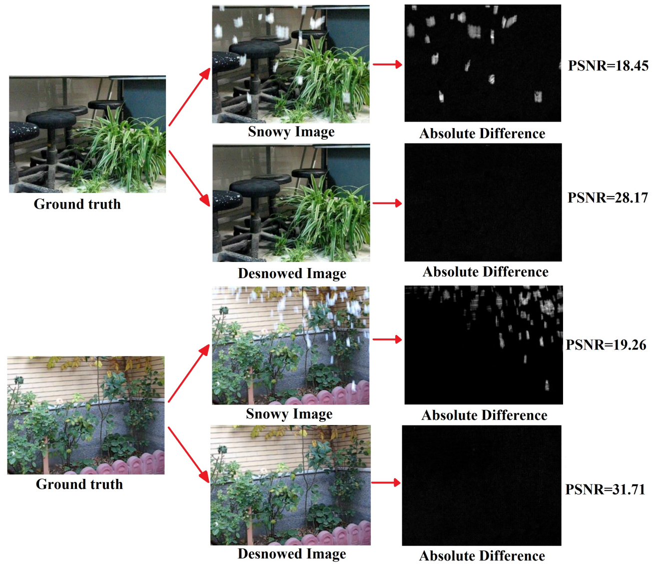


Fig. 6. Quantitive performance of the proposed method for static background and camera on quasi-snow videos.

is tested and evaluated qualitatively for different scenarios in Fig. 5. It can efficiently remove the snow/rain for static background and camera. It can also remove the snow/rain when there is slight camera shaking while recording videos and is robust to illumination changes.

The quantitative evaluation of the proposed method is carried out on two different quasi snow frame sequences considering an image quality assessment metric called peak signal-to-noise ratio (PSNR) and the absolute difference between the desnowed frame and the ground truth frame. PSNR (in dB) is defined as $PSNR(x, y) = 10 \log_{10} \frac{max(x)}{\sqrt{MSE}}$ in which MSE is as $MSE = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - y_{ij})^2$, where x and y are the $m \times n$ ground truth image and its corresponding desnowed image, respectively. In Fig. 6, the PSNR is significantly improved after removing the snow (from 18.45 to 28.17 and from 19.26 to 31.71). Additionally, the visual results of the absolute difference between the desnowed image and the ground truth image, which is almost zero represented with black color, prove the effective performance of the proposed method in removing snow.

The average running times of the proposed method for removing the snow from a single frame with quasi-snow are presented in Table 2 for two different sizes of the image and three different numbers of the consecutive frames. From this table, the speed advantage of the proposed method is evident especially for the small size of the images. As the process is pixel-wise, increasing the size of the image can increase the computational time. However, even for the very big size of the image as 1080x1920 it takes at most 4.40s to remove snow. Although increasing the number of the frames can slightly increase the running time, it significantly increases the accuracy of the final result.

The performance of the proposed method is further evaluated on two real snowy videos with dark and light backgrounds downloaded from YouTube[2] and compared with the other methods in the literature in Figs. 7 and 8, respectively. As the video in Fig. 7 contains heavy rain, its implementation is carried out using ten consecutive frames while the video in Fig. 8 is desnowed only with 7 frames. It is worth mentioning that the video of Fig. 8 composed of scenarios with both stable and dynamic cameras, though only the frames with static background and camera (i.e., $1480-1486^{th}$ frames) are selected for our experiments. Comparing the results in Figs. 7 and 8, it is clear to see that most of the other competing techniques cannot entirely remove the snowflakes and recover the original texture information with details, and some others cause blurring artifacts. In contrast, our proposed method can successfully remove the snowflakes and recover the background texture (even small details in Fig. 8) while the image's contrast is still kept high. No blurriness occurred in the desnowed image, no matter if the background is dark or light. Additionally, it is worth mentioning that this efficient performance is achieved by our method while its computational complexity is significantly less than that of other methods.

Table 2. Average running time of the proposed method for desnowing a single frame with two different sizes of images for 3 different numbers of consecutive frames.

Image size	**3 consecutive frames**	**7 consecutive frames**	**10 consecutive frames**
480x640	0.61s	0.65s	0.75s
1080x1920	3.82s	4.10s	4.40s

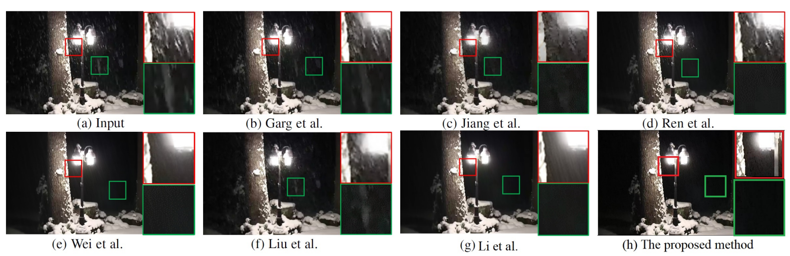


Fig. 7. Qualitative comparison with the methods in the literature on a video with dark background, (a) input, (b) Garg et al [4], (c) Jiang et al [7], (d) Ren et al [18], (e) Wei et al [19], (f) Liu et al [14], (g) Li et al [12], (h) the proposed method.

5.2 Failure Cases

As it was presented, the proposed method can successfully deal with the snowy video sequences with static background and camera containing light and heavy snow/rain. However, it still suffers from limitations when there are extensive

[2] `https://www.youtube.com/watch?{v=kNTYEKjXqzs,v=HbgoKKj7TNA}`

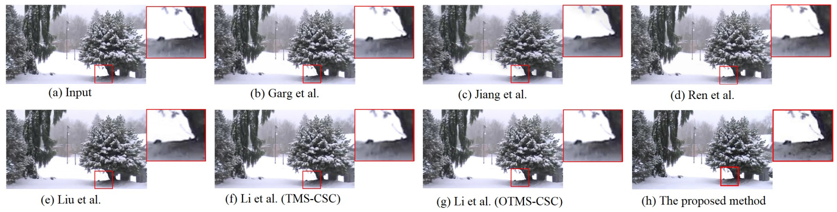


Fig. 8. Qualitative comparison between the proposed method and the methods in the literature on a video with light background, (a) input, (b) Garg et al [4], (c) Jiang et al [7], (d) Ren et al [18], (e) Liu et al [14], (f and g) Li et al [12], (h) the proposed method.

movements in the background and camera. As our method is a pixel-wise method based on the temporal information of the frame sequences, variations in spatial information can lead to slight blurriness, especially for heavy rain/snow when more frames are required for completely removing the snow. In our future work, we will attempt to improve the method by adding spatial information to make it more robust to background movements and camera translation, rotation, and zooming in the presence of heavy snow/rain.

6 Conclusion

This paper first built a snow and rain dataset by considering all possible background and camera scenarios. In this dataset, three kinds of particles were used, i.e., synthetic snow and rain with different sizes and speeds, quasi-snow, and real snow and rain. Additionally, a simple but very efficient desnowing/deraining method was proposed based on the temporal information of consecutive frames and the color of pixels for static background and camera. Experimental results and comparisons with the related works proved that our method could efficiently remove the snow/rain without causing any blurriness. It performed faithfully even for heavy snow scenarios and illumination changes and could recover the original background information. We will improve the method to deal with the dynamic background and moving cameras in our future research work.

References

1. Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J.: Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing **26**(6), 2944–2956 (2017)
2. Garg, K., Nayar, S.K.: Detection and removal of rain from videos. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. vol. 1, pp. I–I. IEEE (2004)
3. Garg, K., Nayar, S.K.: When does a camera see rain? In: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. vol. 2, pp. 1067–1074. IEEE (2005)
4. Garg, K., Nayar, S.K.: Vision and rain. International Journal of Computer Vision **75**(1), 3–27 (2007)

5. Jaw, D.W., Huang, S.C., Kuo, S.Y.: Desnowgan: An efficient single image snow removal framework using cross-resolution lateral connection and gans. IEEE Transactions on Circuits and Systems for Video Technology **31**(4), 1342–1350 (2020)
6. Jeny, A.A., Sakib, A.N.M., Junayed, M.S., Lima, K.A., Ahmed, I., Islam, M.B.: Sknet: A convolutional neural networks based classification approach for skin cancer classes. In: 2020 23rd International Conference on Computer and Information Technology (ICCIT). pp. 1–6. IEEE (2020)
7. Jiang, T.X., Huang, T.Z., Zhao, X.L., Deng, L.J., Wang, Y.: A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors. In: Proceedings of the ieee conference on computer vision and pattern recognition. pp. 4057–4066 (2017)
8. Jung, A.B., Wada, K., Crall, J., Tanaka, S., Graving, J., Reinders, C., Yadav, S., Banerjee, J., Vecsei, G., Kraft, A., Rui, Z., Borovec, J., Vallentin, C., Zhydenko, S., Pfeiffer, K., Cook, B., Fernández, I., De Rainville, F.M., Weng, C.H., Ayala-Acevedo, A., Meudec, R., Laporte, M., et al.: imgaug. `https://github.com/aleju/imgaug` (2020), online; accessed 01-Feb-2020
9. Kim, J.H., Sim, J.Y., Kim, C.S.: Stereo video deraining and desnowing based on spatiotemporal frame warping. In: 2014 IEEE International Conference on Image Processing (ICIP). pp. 5432–5436. IEEE (2014)
10. Li, M., Cao, X., Zhao, Q., Zhang, L., Gao, C., Meng, D.: Video rain/snow removal by transformed online multiscale convolutional sparse coding. arXiv preprint arXiv:1909.06148 (2019)
11. Li, M., Cao, X., Zhao, Q., Zhang, L., Meng, D.: Online rain/snow removal from surveillance videos. IEEE Transactions on Image Processing **30**, 2029–2044 (2021)
12. Li, M., Xie, Q., Zhao, Q., Wei, W., Gu, S., Tao, J., Meng, D.: Video rain streak removal by multiscale convolutional sparse coding. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6644–6653 (2018)
13. Li, R., Cheong, L.F., Tan, R.T.: Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1633–1642 (2019)
14. Liu, J., Yang, W., Yang, S., Guo, Z.: Erase or fill? deep joint recurrent rain removal and reconstruction in videos. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3233–3242 (2018)
15. Liu, Y.F., Jaw, D.W., Huang, S.C., Hwang, J.N.: Desnownet: Context-aware deep network for snow removal. IEEE Transactions on Image Processing **27**(6), 3064–3073 (2018)
16. Pitropov, M., Garcia, D.E., Rebello, J., Smart, M., Wang, C., Czarnecki, K., Waslander, S.: Canadian adverse driving conditions dataset. The International Journal of Robotics Research **40**(4-5), 681–690 (2021)
17. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
18. Ren, W., Tian, J., Han, Z., Chan, A., Tang, Y.: Video desnowing and deraining based on matrix decomposition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4210–4219 (2017)
19. Wei, W., Yi, L., Xie, Q., Zhao, Q., Meng, D., Xu, Z.: Should we encode rain streaks in video as deterministic or stochastic? In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2516–2525 (2017)
20. Yang, W., Liu, J., Feng, J.: Frame-consistent recurrent video deraining with dual-level flow. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1661–1670 (2019)