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

Arezoo Sadeghzadeh, Md Baharul Islam, Reza Zaker

2021. 09. 23

ICVS-2021, Vienna University of Technology

BAU BAHÇEŞEHİR ÜNİVERSİTESİ
AUM AMERICAN UNIVERSITY OF MALTA
Azarbaijan Shahid Madani University
Computer Vision Systems
13th International Conference, ICVS 2021
Virtual Event, September 22–24, 2021
Proceedings

Contents

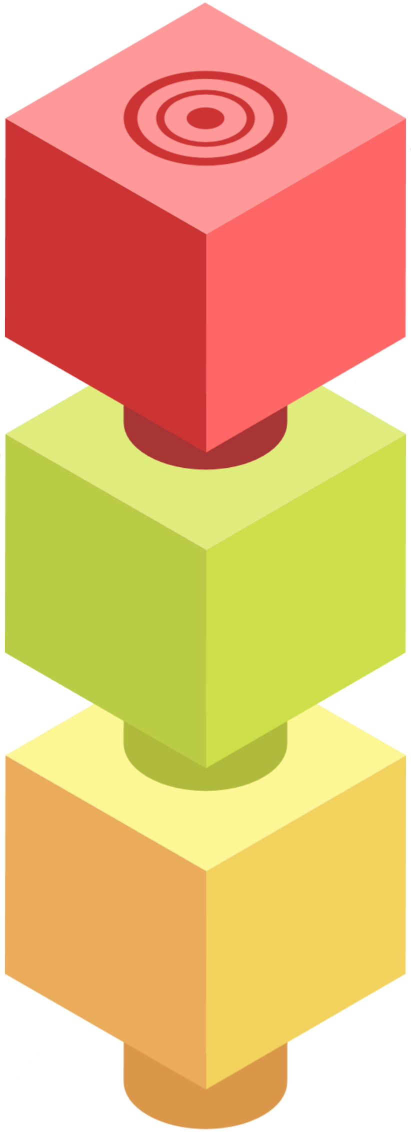
Introduction

The Importance of Snow/Rain Removal

Example For Its Importance

Object detection in presence of snow   Object detection from the scene without snow

Left image   Right image   Disparity map



Why Snow/Rain Removal?

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

Application

Video/movie editing, vision-based navigation, autonomous driving, and video surveillance

Introduction

Challenges and Limitations of Video-based Snow/Rain Removal Approaches

Lack of Appropriate Dataset

Lack of ground-truth information for real snowy/rainy scenes from videos or YouTube

Limited variations of the snowflakes, rain streaks, camera setting, and environmental conditions in synthetic snow and rain

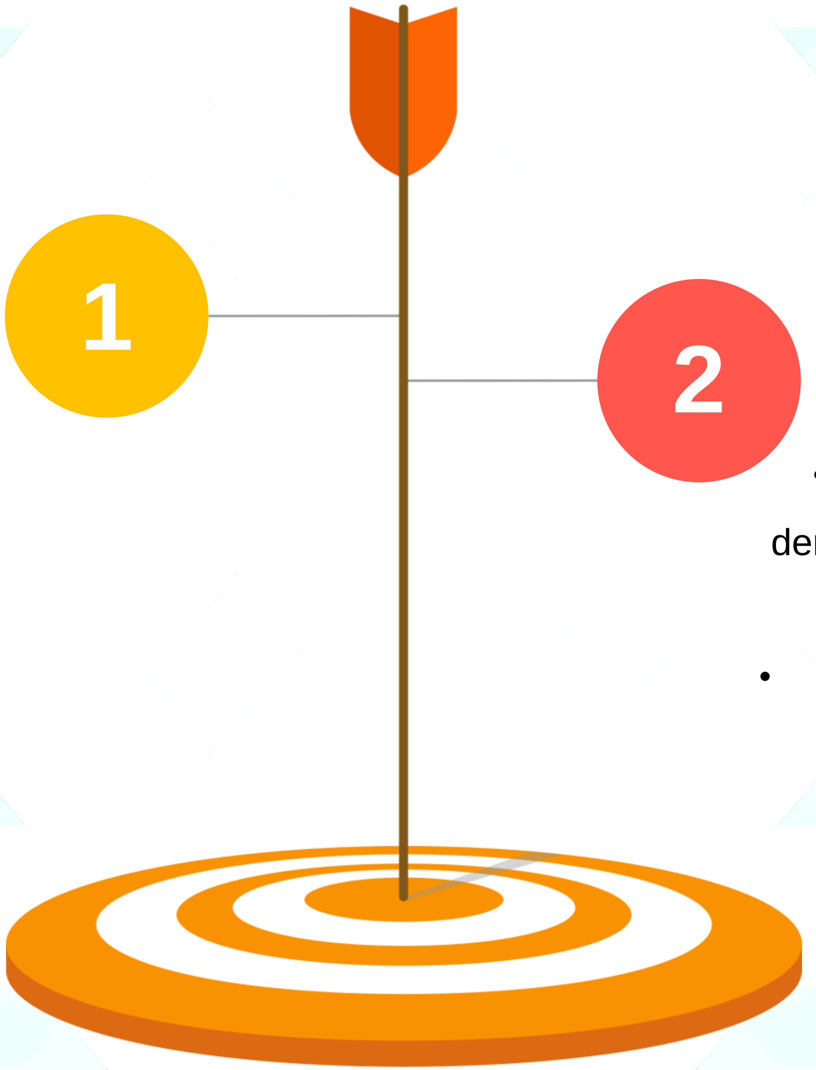
Performance of Video-based Approaches

Performance degradation of the video rain removal techniques applied to the snowy videos

Causing severe blurring artifacts even for static backgrounds when the camera slightly shakes or moves while capturing the videos

Introduction

The Main Goal of the Paper



1 Novel Dataset Development

- Providing a new snow and rain videos dataset
- Considering all variations of background (static and dynamic) and the camera (static and dynamic, e.g. translation, zooming, illumination changes, and rotation)
- Providing the ground-truth information in the synthetic snow and rain videos allows the

2 Pixel-wise Video Snow/Rain Removal Method

- A simple but very efficient video desnowing/deraining method based on the temporal information and the color of the pixels
- Removing the snow and rain for static background and camera even for heavy snow
- Low computational cost and high robust to illumination changes and camera shaking

Introduction

The Problems of Current Studies

Snow/Rain Removal

Image-based Approaches

Sparsity prior, patch-rank prior ← Suffering from the limited generalization ability

↓

CNNs and handcrafted priors ← Time-consuming and a very large size

A physics-based backbone and GAN

single-image desnowing model based on a pyramidal hierarchicaldesign

Video-based Approaches

High correlation between the corresponding pixels in consecutive frames

Stereo videos desnowing/deraining → Include more information and details than the monocular videos

↓

CNNs and RCNNs

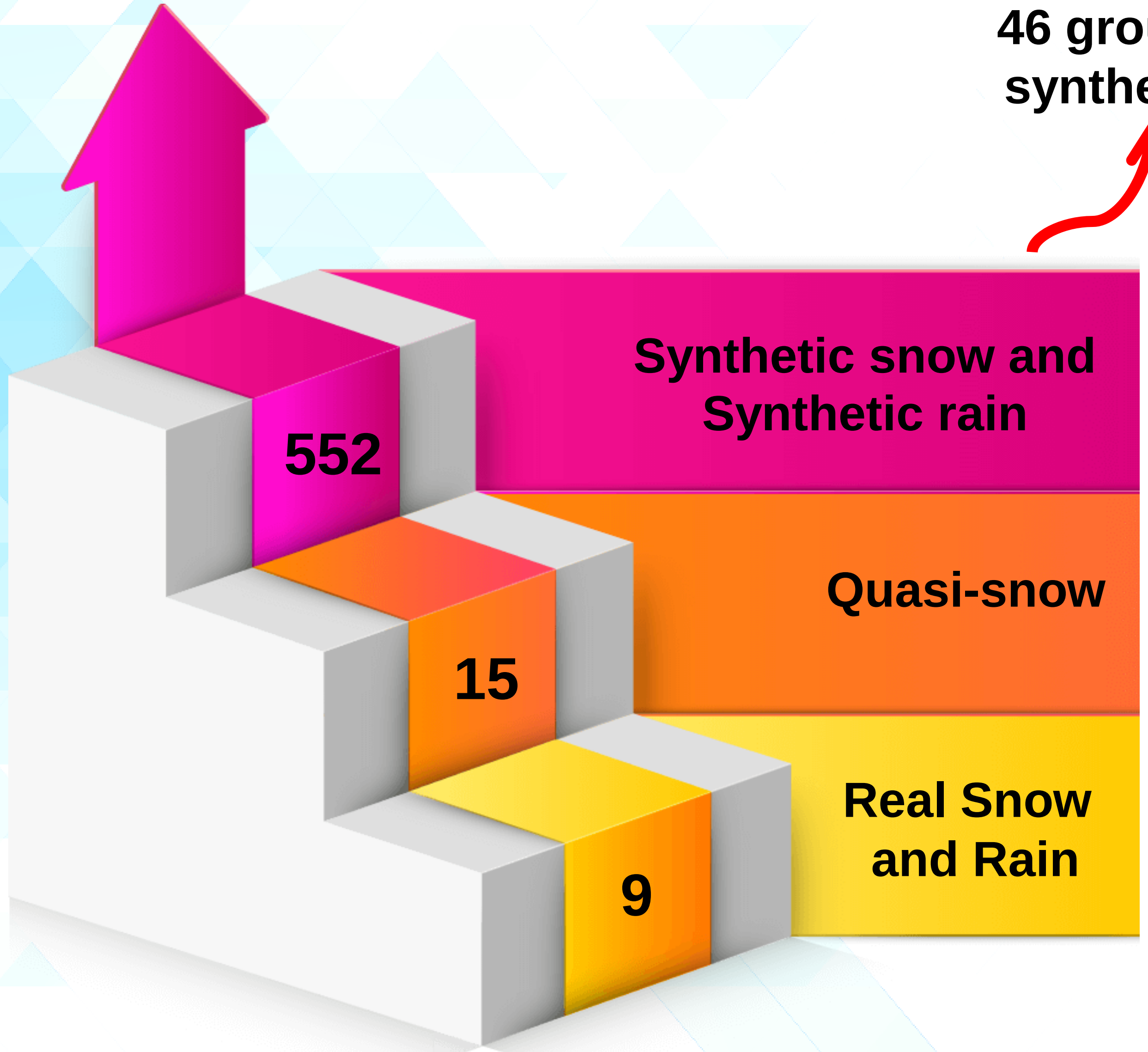
Online multi-scale convolutional sparse coding → Limitations on non-surveillance videos, moving objects, fast illumination changes, and fast moving cameras

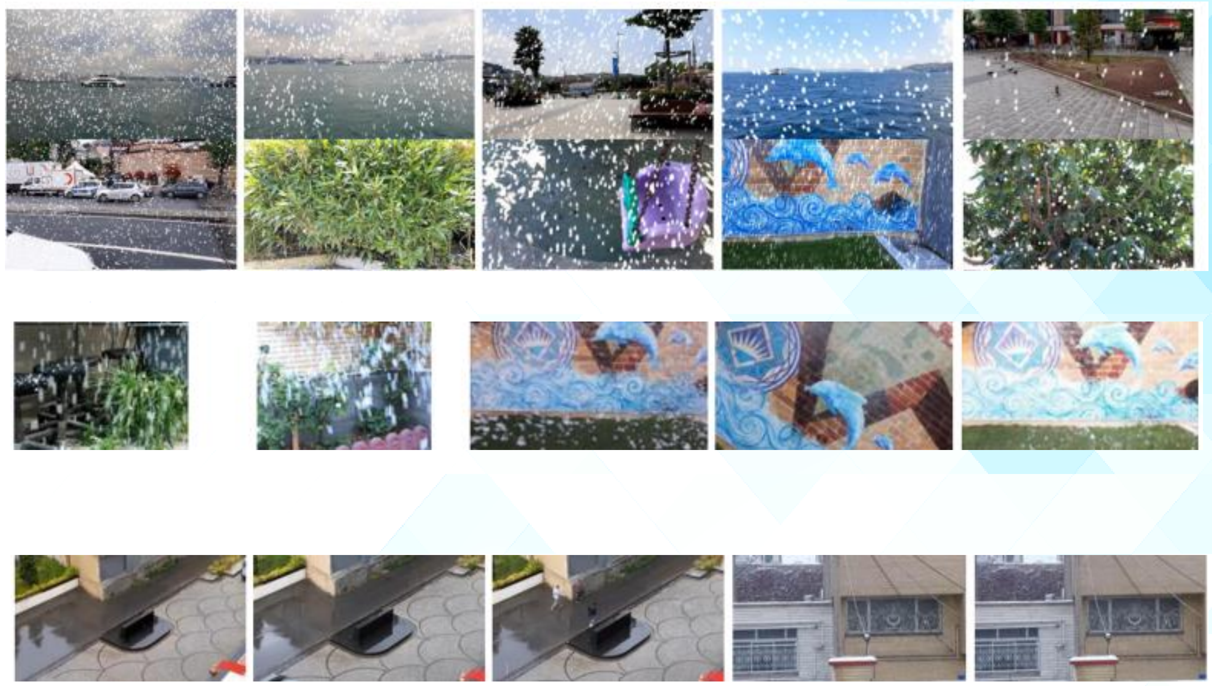
Dataset Development

Variants in Snow&Rain Videos

Totally 577 Videos

46 ground-truth videos and 506 synthetic snow/rain created by augmentation





Dataset Development

Variants in Snow&Rain Videos

Ground-truth videos for synthetic snow and rain

Camera resolution	Camera variations	Background	Number of videos	Video length
25MP 1080 × 1920 pixels fps = 30	Static	Dynamic (slow)	3	35 s, 14 s, 13 s
		Dynamic (fast)	4	8 s, 8 s, 10 s, 8 s
	Dynamic (shaking and translation)	Dynamic (slow)	1	14 s
	Static and Dynamic (slow and fast translation, zooming, rotation, illumination changes, small and large shaking, both rotation and zooming)	Dynamic (fast)	13	11 s, 11 s, 14 s, 18 s, 11 s, 10 s, 13 s, 29 s, 9 s, 12 s, 11 s, 13 s, 11s
		Dynamic (slow)	13	14 s, 21 s, 14 s, 13 s, 18 s, 10 s, 10 s, 20 s, 10 s, 34 s, 11 s, 9 s, 13 s
		Static	12	22 s, 11 s, 18 s, 11 s, 25 s, 12 s, 37 s, 11 s, 10 s, 14 s, 18 s, 11 s

Videos based on quasi-snow

8Mp 480 × 460 fps = 30	Static	Static	2	59 s, 23 s
	Dynamic (translation)		1	74 s
25MP 1080 × 1920 fps = 30	Static and Dynamic (slow and fast translation, zooming, rotation, illumination changes, small and large shaking, both rotation and zooming)		13	25 s, 19 s, 12 s, 10 s, 26 s, 16 s, 16 s, 10 s, 19 s, 28 s, 11 s, 26 s, 13 s

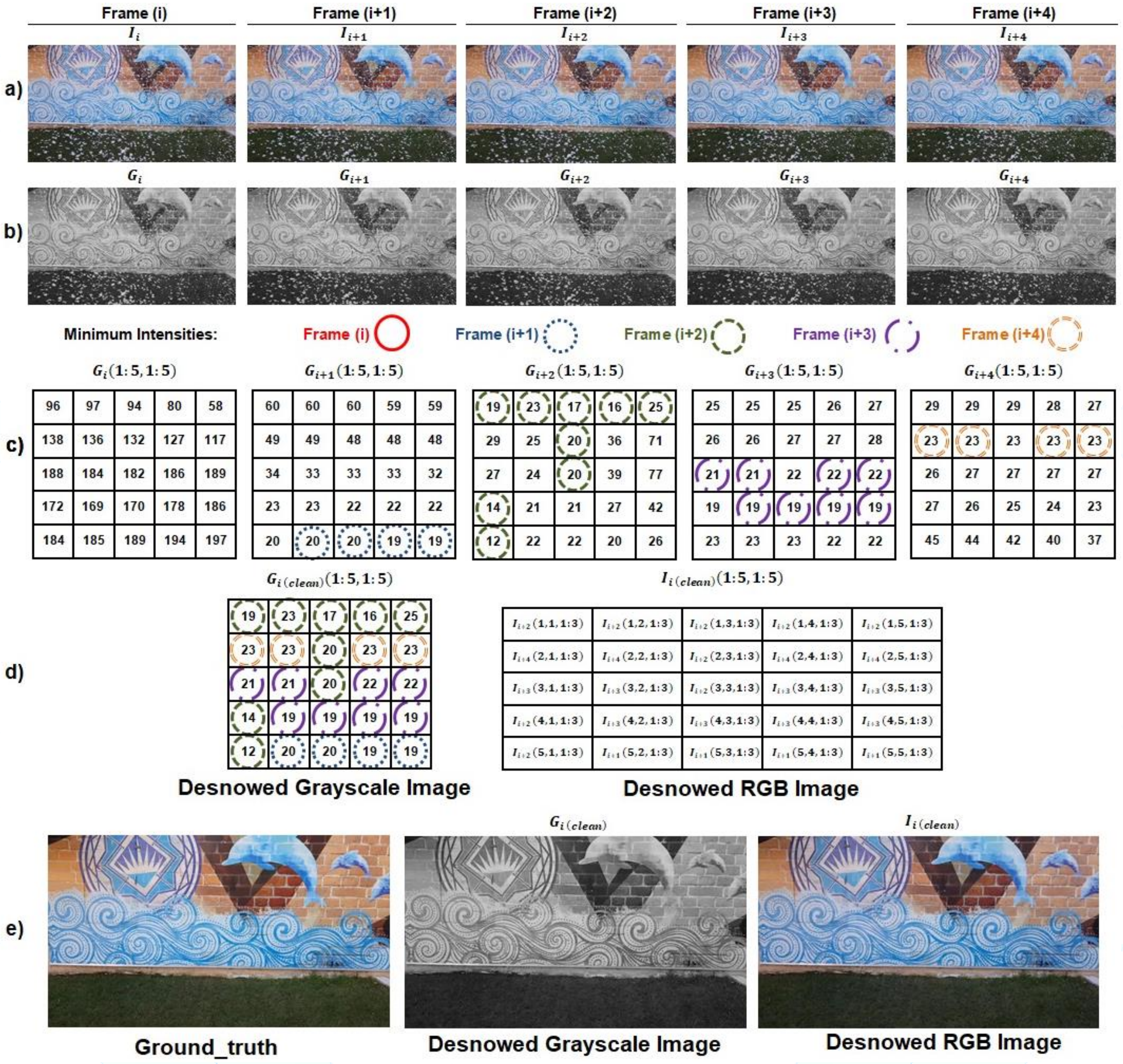
Videos with real snow and rain

13MP 1080 × 1920 fps = 30	Static	Static	1 (snow)	8 s
25MP 1080 × 1920 fps = 30	Static	Static and Dynamic	8 (light and heavy rain)	21 s, 7 s, 21 s, 11 s, 12 s, 42 s, 11 s, 18 s

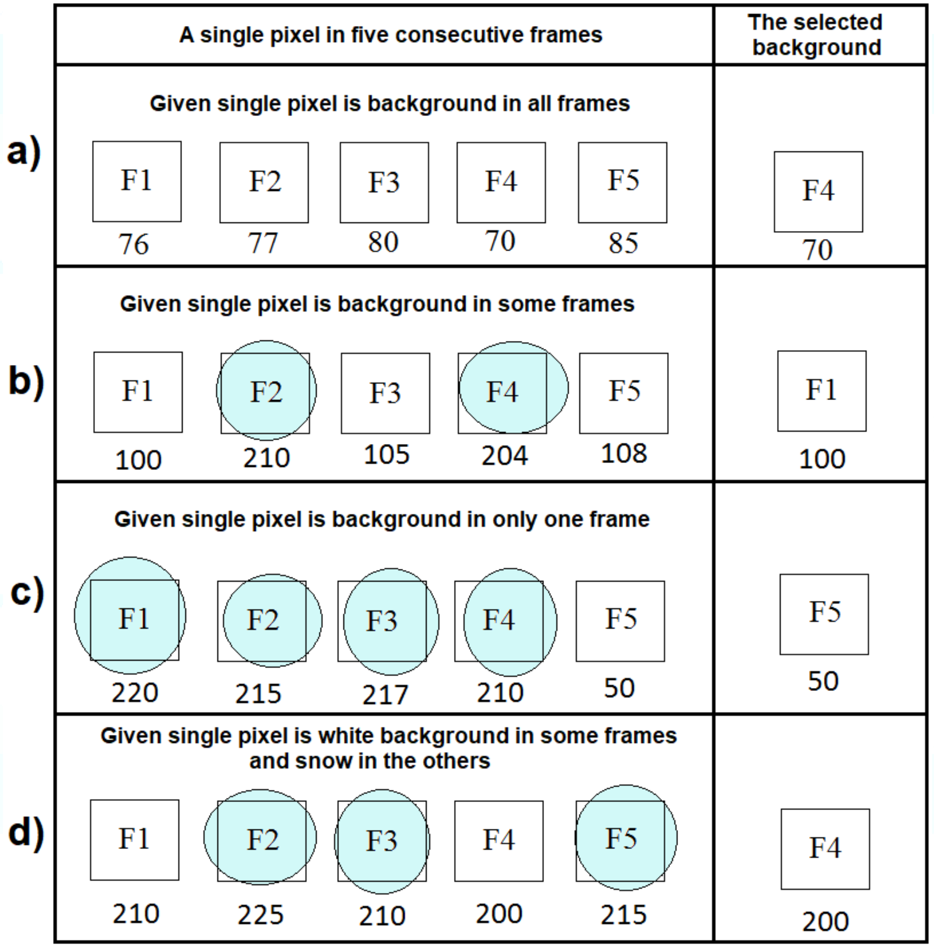
Total characteristics of the videos and the considered scenarios in the developed dataset.

Proposed Method

The Overall Flowchart



Proposed Method



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

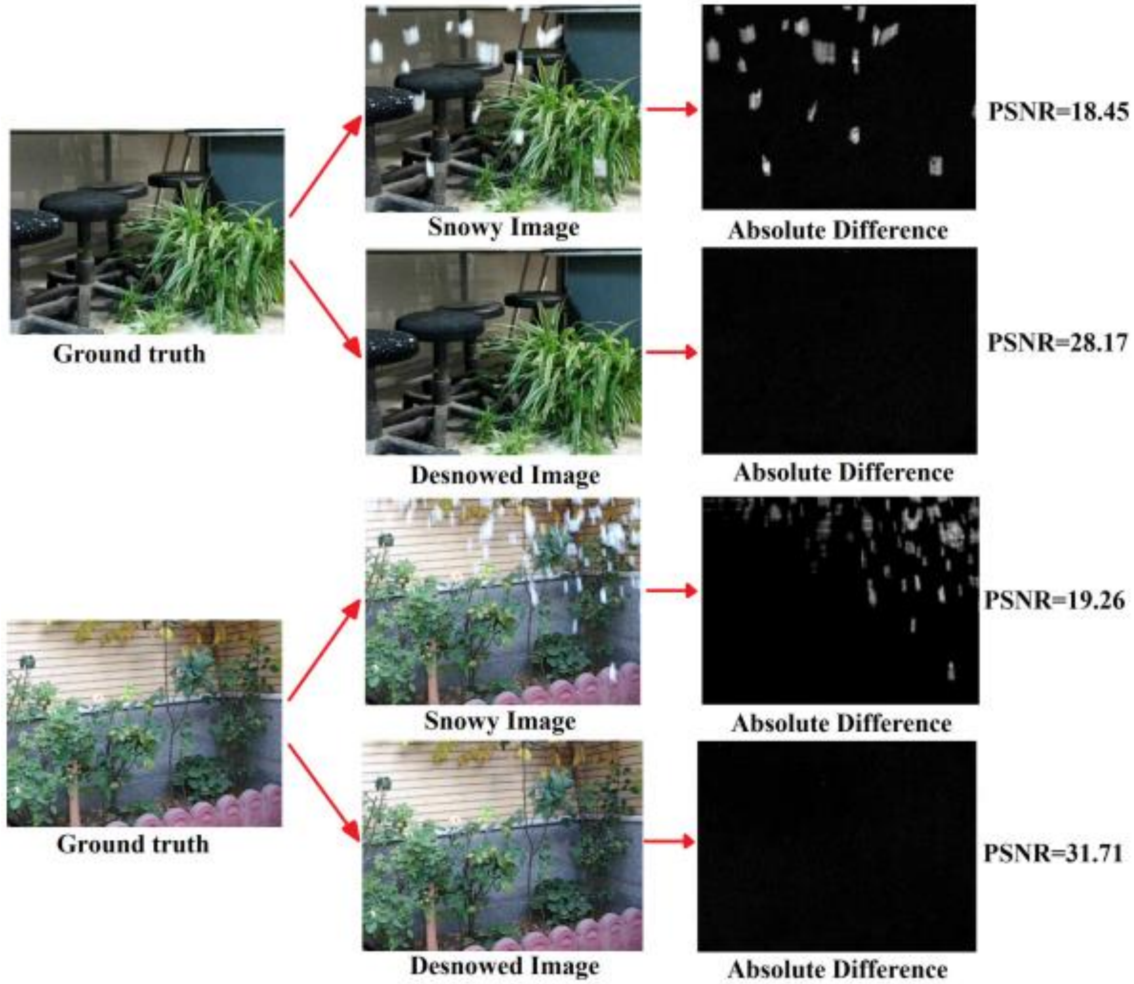
Experimental Results

Qualitative and Quantitative Results and Computational time

Qualitative Results

Image size	3 consecutive frames	7 consecutive frames	10 consecutive frames
480 × 640	0.61 s	0.65 s	0.75 s
1080 × 1920	3.82 s	4.10 s	4.40 s

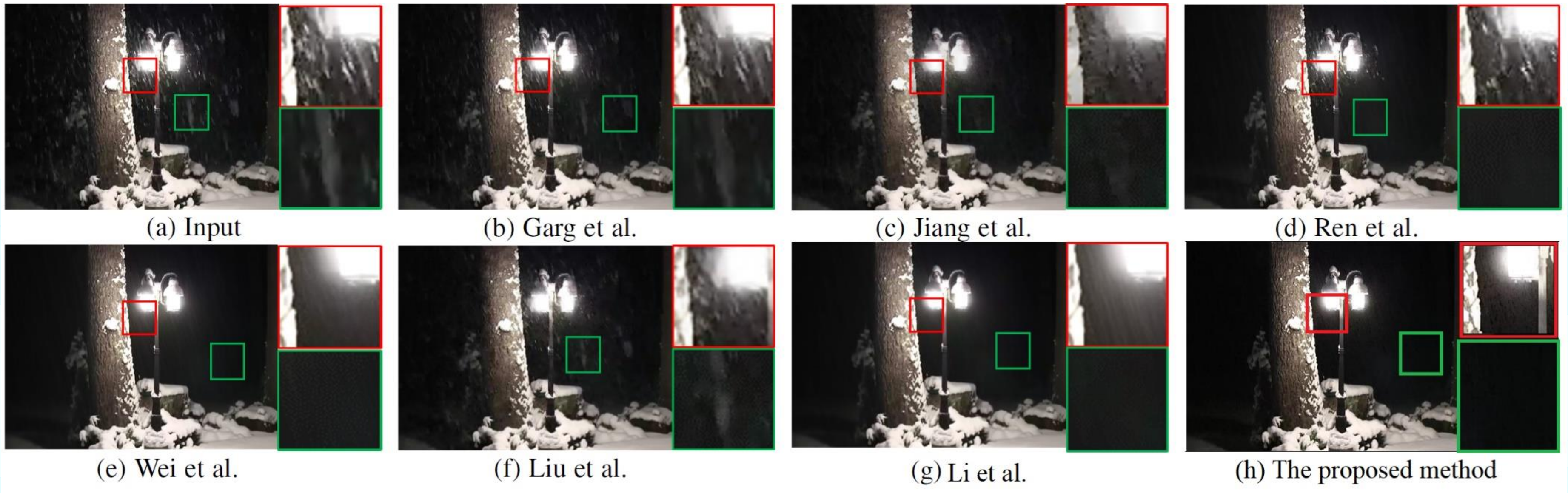
Computational time

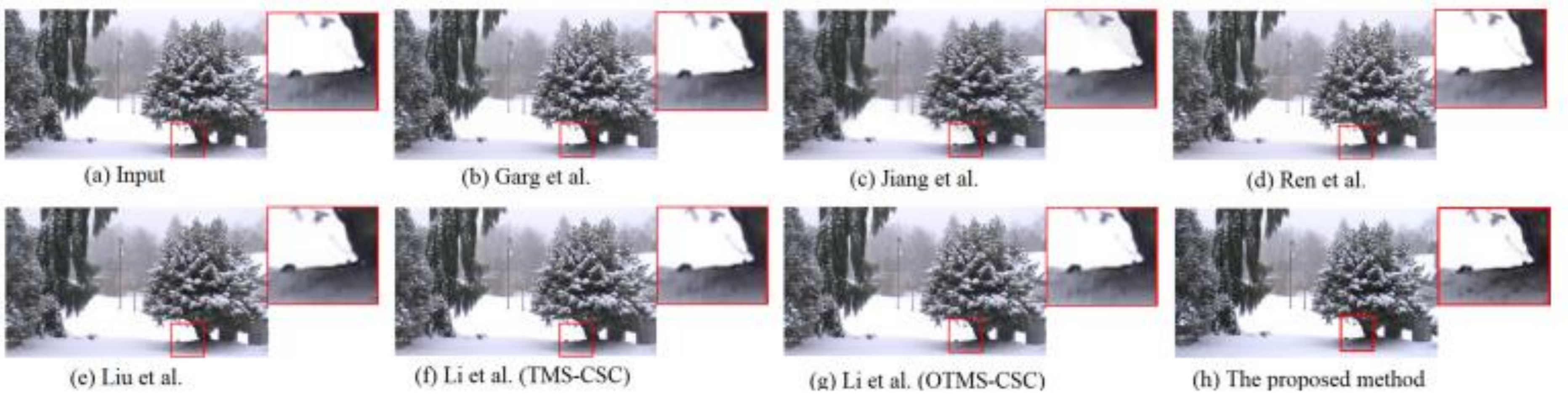


Quantitative Results

Experimental Results

Comparison





Conclusion

Cons/Pros and Future Direction

+ Building a snow and rain dataset for the first time by considering all possible background and camera scenarios

+ Our desnowing/deraining method could efficiently remove the snow/rain without causing any blurriness even for heavy snow scenarios and illumination changes with low computational cost and time.

- Suffering from limitations in dealing with extensive background movements and camera translation, rotation, and zooming in the presence of heavy snow/rain

Improving the method to deal with the dynamic background and moving cameras

Thank you for your attention!