Sequential Deep Unrolling with Flow Priors for Robust Video Deraining

Xinwei Xue [1,2,*], Ying Ding [1,2], Pan Mu [1,2], Long Ma [1,2], Risheng Liu [1,2], Xin Fan [1,2]

1 DUT-RU International School of Information Science & Engineering, Dalian University of Technology
2 Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province
ICASSP 2020

目录

Contents

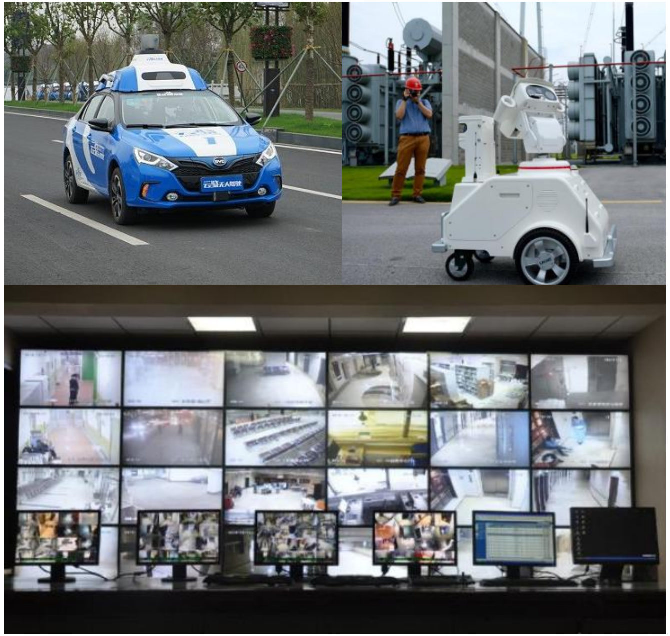
PART
01

Introduction

• **01** Introduction

Background:

Outdoor vision systems are employed for a wide range of applications. Rain is considered as a common and challenging problem in outdoor vision systems, which seriously deteriorates the quality of image and video.



• 01 Introduction

Related work:

- **Video deraining methods based on traditional optimization**

FastDerain (tensor and intrinsic priors), MS-CSC (multi-scale convolutional sparse coding) ;

Comlex prior limitation & Detail information loss in real-world scenarios

- **Video deraining methods based on deep learning**

SpacCNN (super-pixel alignment and compensation CNN framework); J4RNet (joint recurrent rain removal and reconstruction network);

Background smoothed or blurred & Details and textures loss

• **01** Introduction

Contributions:

· **A new video deraining model with flow priors**

The model is established to simultaneously introduce spatial and temporal information for accurately depicting the enhancement model of the current frame.

· **A sequential deep unrolling network**

The network is constructed based on the solving process derived from the single-frame deraining model with flow priors, to obtain data support for guaranteeing performance.

· **Superior performance**

Ablation study demonstrates the effectiveness and necessity of our network. Extensive experiments on synthetic and real rainy video sequences fully verify our superiority.

PART 02

Proposed Method

• 02 Proposed method

Video Deraining Model with Temporal and Spatial Priors

$$\mathbf{O} = \mathbf{B} + \mathbf{R},$$ (single image deraining model)

$$\mathbf{O}_t = \mathbf{B}_t + \mathbf{R}_t,\ t = 1, 2, ..., N,$$

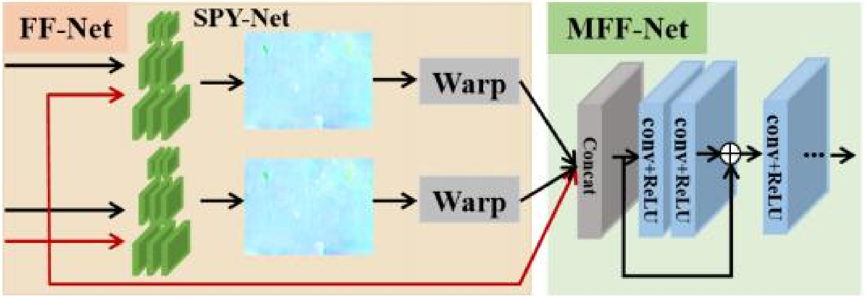
$$\min_{\mathbf{B}_t, \mathbf{R}_t} \|\mathbf{O}_t - \mathbf{B}_t - \mathbf{R}_t\|_F^2 + \Psi(\mathbf{B}_t) + \Phi(\mathbf{R}_t),$$

$$\Psi(\mathbf{B}_t) = \Psi_{tem}(\mathbf{B}_t, \{\mathbf{B}_{t-i}, \mathbf{u}_i\}_{i=1}^{N}) + \Psi_{spa}(\mathbf{B}_t),$$

• **02** Proposed method

Sequential Deep Unrolling Framework

$$
\begin{cases}
\mathbf{R}_t^{k+1} = \dfrac{\mathbf{O}_t - \mathbf{B}_t^k}{1-\rho_0} + \dfrac{\rho_0}{\rho_0 - 1}\mathbf{P}_t^k, \\
\mathbf{B}_t^{k+1} = \dfrac{\mathbf{O}_t - \mathbf{R}_t^{k+1}}{1-\rho_1-\rho_2} + \dfrac{\rho_1}{\rho_1+\rho_2-1}\mathbf{Q}_s^k + \dfrac{\rho_2}{\rho_1+\rho_2-1}\mathbf{Q}_t^k,
\end{cases}
$$



$$
\begin{cases}
\mathbf{P}_t^{k+1} = \arg\min_{\mathbf{P}_t} \Phi(\mathbf{P}_t) + \dfrac{\rho_0}{2}\|\mathbf{R}_t^{k+1} - \mathbf{P}_t\|^2, \\
\mathbf{Q}_s^{k+1} = \arg\min_{\mathbf{Q}_s} \Psi_{spa}(\mathbf{Q}_s) + \dfrac{\rho_1}{2}\|\mathbf{B}_t^{k+1} - \mathbf{Q}_s\|^2, \\
\mathbf{f}_i^{k+1} = \arg\min_{\mathbf{f}_i} \Psi_{tem}(\mathbf{B}_t^{k+1}, \{\mathbf{B}_{t-i}, \mathbf{f}_i\}_{i=1}^N) + \dfrac{\rho_3}{2}\|\mathbf{u}_i^k - \mathbf{f}_i\|^2, \\
\mathbf{Q}_t^{k+1} = \arg\min_{\mathbf{Q}_t} \Psi_{tem}(\mathbf{Q}_t, \{\mathbf{f}_i^{k+1}, \mathbf{u}_i^{k+1}\}_{i=1}^N) + \dfrac{\rho_2}{2}\|\mathbf{B}_t^{k+1} - \mathbf{Q}_t\|^2,
\end{cases}
$$

Experimental Results

PART
03

• 03 Experimental Results

Data Set and Related Works

Synthesized data:

- RainSynLight25 from J4RNet[9] -> light rain streaks
- RainSynComplex25 from J4RNet[9] -> heavy rain streaks and sparkle noises

Real data:

- The data from SpacCNN[8]

RainSynLight25

RainSynComplex25

Real

[9] Jiaying Liu, Wenhan Yang, Shuai Yang, and Zongming. Guo, "Erase or fill? deep joint recurrent rain re_x0002_moval and reconstruction in videos," in CVPR, 2018, pp. 3233–3242.
[8] Jie Chen, Cheen-Hau Tan, Junhui Hou, Lap-Pui Chau, and He. Li, "Robust video content alignment and compensation for rain removal in a cnn framework," in CVPR, 2018, pp. 6286–6295.

• 03 Experimental Results

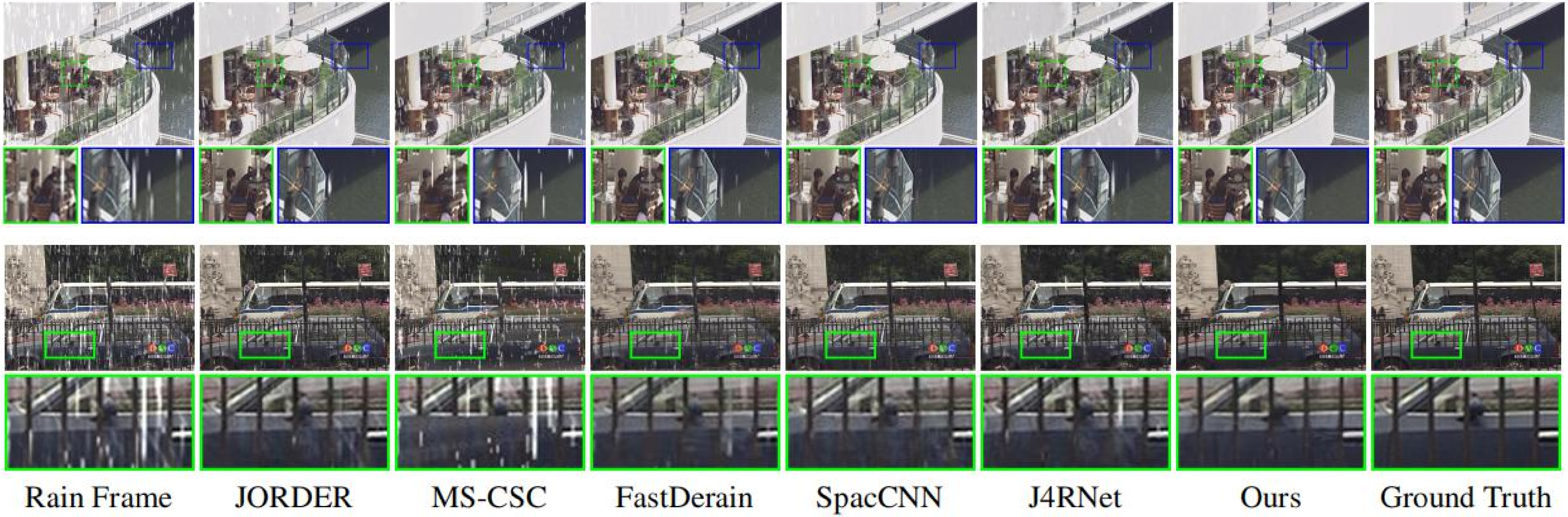
Results for synthesized data

Dataset	Metric	JORDER[15]	DID-MDN [16]	MS-CSC[6]	FastDerain[5]	SpacCNN[8]	J4RNet[9]	Ours
		Single image deraining method		Video deraining method				
RainSynLight25	PSNR	31.03	23.78	24.43	31.57	31.52	31.71	**33.04**
	SSIM	0.9134	0.8140	0.7312	0.9058	0.8980	0.8971	**0.9643**
RainSynComplex25	PSNR	19.99	17.51	16.57	26.91	21.46	22.46	**28.03**
	SSIM	0.6085	0.5888	0.5833	0.8011	0.5925	0.7336	**0.9046**
	Time (s)	0.29	0.14	3.64	0.42	4.49	6.42	1.01
		Caffe (matlab)	Python	Matlab	Matlab	MatConvNet	Caffe (matlab)	Pytorch (python)

[15] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng. Yan, “Deep joint rain detection and removal from a single image,” in CVPR, 2017, pp. 1357–1366.
[16] He Zhang and Vishal M. Patel, “Density-aware single image de-raining using a multi-stream dense network,” in CVPR, 2018, pp. 695–704.
[6]Minghan Li, Qi Xie, Qian Zhao, Wei Wei, Shuhang Gu, Jing Tao, and Deyu. Meng, “Video rain streak removal by multiscale convolutional sparse coding,” in CVPR, 2018, pp. 6644–6653.
[5] Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Liang Jian Deng, and Yao. Wang, “A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors,” in CVPR, 2017, pp. 4057–4066.

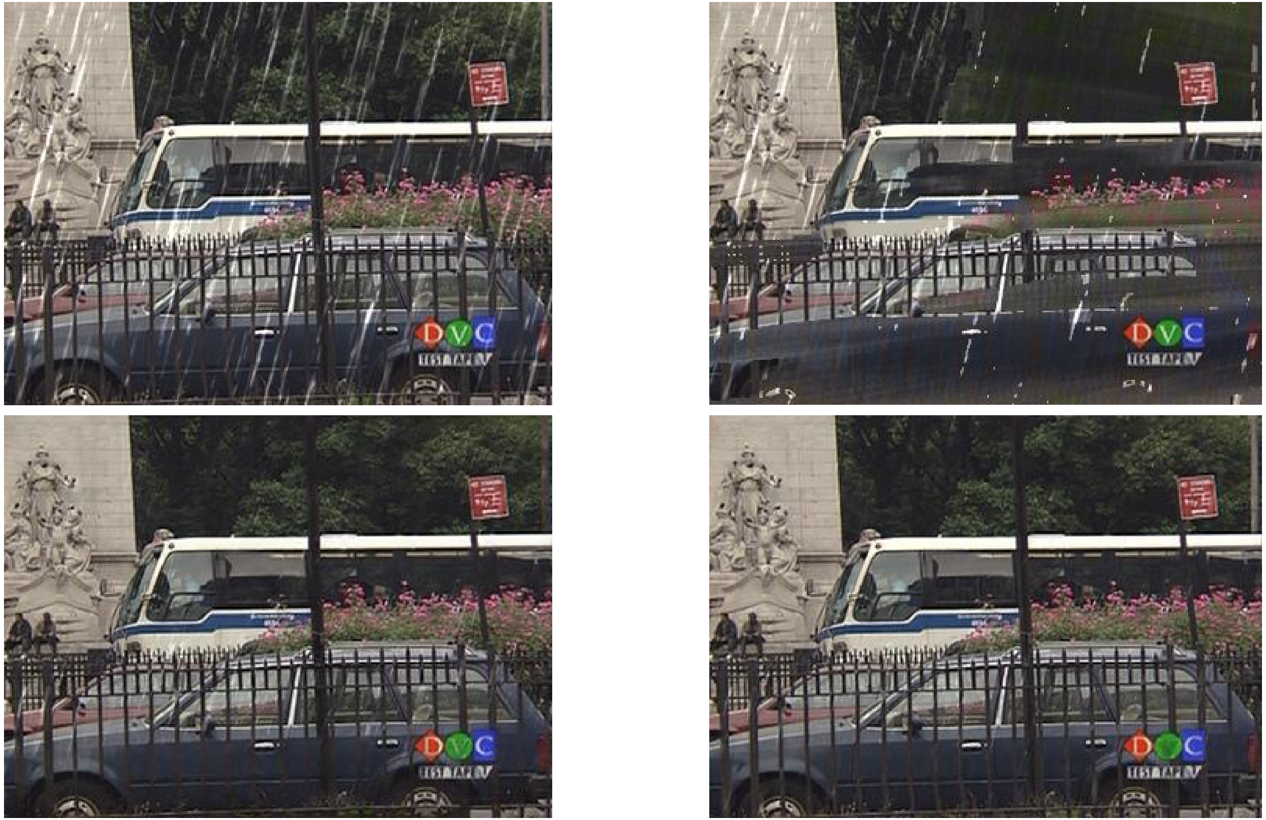
• **03** Experimental Results

Results for synthesized rain image



• **03** Experimental Results

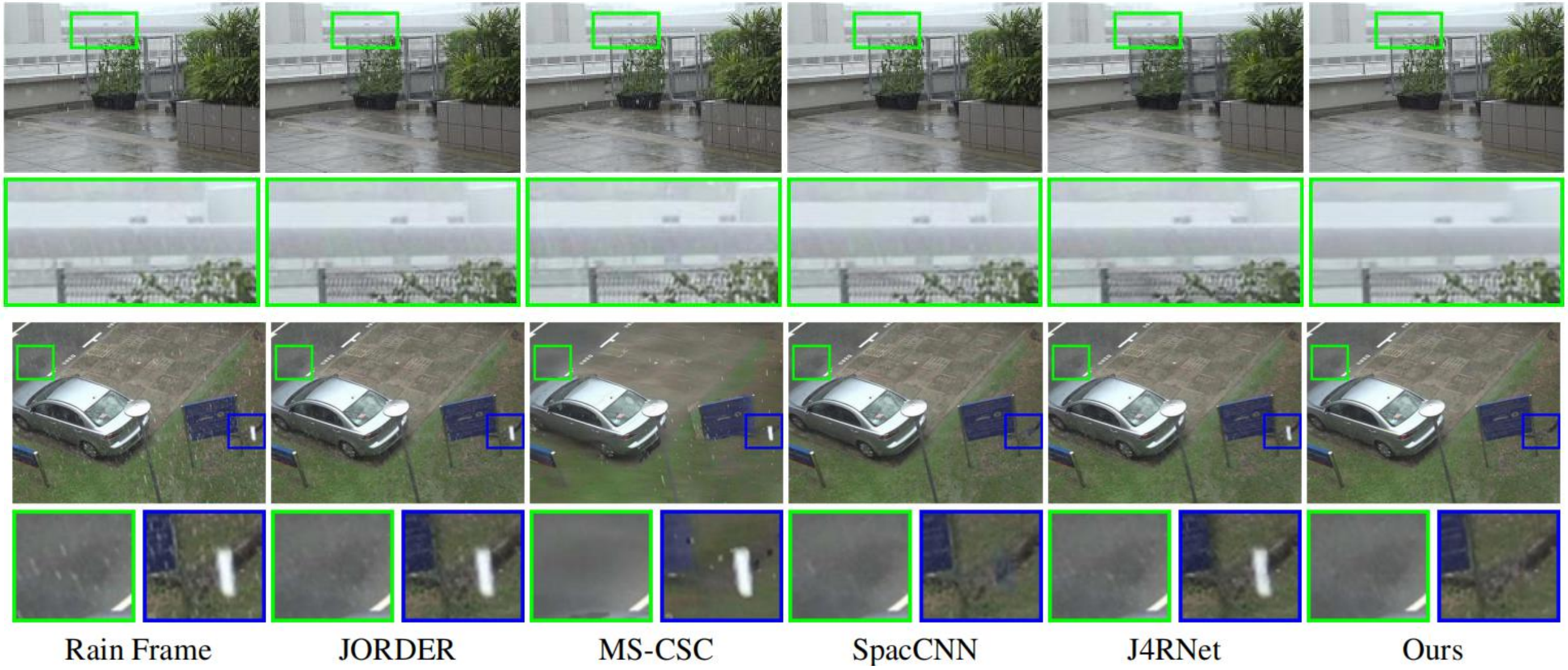
Input MS-CSC



J4RNet Ours

• **03** Experimental Results

Results for real rain data



• **03** Experimental Results

Results for Ablation Study

	M1	M2	M3	M4
SFD-Module	√	×	×	√
FF-Module	×	×	√	√
MFF-Module	×	√	√	√
PSNR	25.24	26.53	27.11	**28.03**
SSIM	0.8083	0.8258	0.8491	**0.9046**

Dataset：RainSynComplex25

M1: SFD-Module (Single-Frame Deraining Module)

M2: MFF-Module (Multi-Frame Fusion Module)

M3: FF-Module (Flow Fusion Module)+ MFF-Module

M4: SFD-Module + FF-Module +MFF-Module

• **03** Experimental Results

Results for Ablation Study

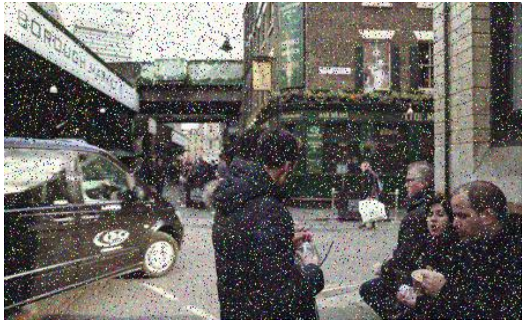
Input	M1	M2	M3	M4
	25.24 / 0.8083	26.53 / 0.8258	27.11 / 0.8491	**28.03 / 0.9046**

Conclusion

PART
04

• **04** Conclusion

- A noval video deraining model with flow priors
- A sequential deep unrolling network
- Superior performance
- Future work -> video desnowing, denoising and dehazing


Thanks for your listening