Joint Video Rolling Shutter Correction and Super-Resolution

Akash Gupta[†], Sudhir Kumar Singh[†], Amit K. Roy-Chowdhury[‡]

[†] Vimaan [‡] University of California, Riverside

WACV WAIKOLOA HAWAII JAN 3-7 • 2023

Key Takeaways

Rolling shutter artifacts and low-resolution imaging are often commonplace in mobile applications using CMOS cameras.

To address this problem, we propose Patch Attention Network (**PatchNet**) to *jointly* optimize for rolling shutter correction and super-resolution in the feature space.

PatchNet leverages combination of local information, using deformable convolution, and motion field driven global patch-level information from neighbouring patches to recover a high-resolution global shutter video.

Related Work

Table 1: **Characteristic comparison of prior works in rolling shutter correction (RSC) and super-resolution (SR).** Different from the state-of-the-art approaches, **PatchNet** demonstrates patch-level attention in latent space to exploit internal patch recurrence and global information along with pixel-level attention using deformable convolution.

Methods	Task		Attention	
	RSC?	SR?	Pixel?	Patch?
DUN [4]	✓	✗	✗	✗
VSR-T [1]	✗	✓	✗	✓
JCD [5]	✓	✗	✓	✗
PatchNet	✓	✓	✓	✓

Conclusions

Our work addresses the task of joint rolling shutter correction and super-resolution.

We propose Patch Attention Network that employs patch-level attention in feature space to extract global information from neighboring patches using key-query similarity and local information using deformable convolution.

We show that **PatchNet** is able to recover high-resolution global shutter frames from low-resolution rolling shutter video.

References

[1] J. Cao, Y. Li, K. Zhang, and L. Van Gool, "Video super-resolution transformer," *arXiv preprint arXiv:2106.06847*, 2021.

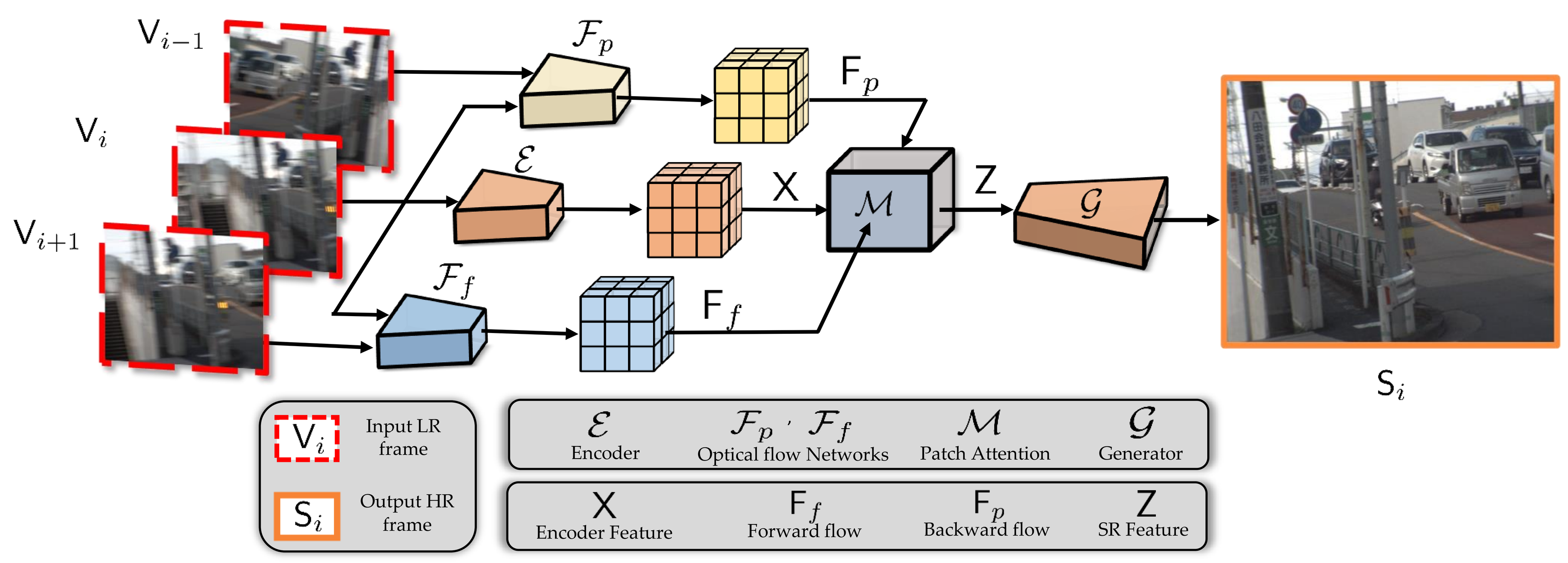
[2] T. Dai, J. Cai, Y. Zhang, S.-T. Xia, and L. Zhang, "Second-order attention network for single image super-resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11065–11074, 2019.

[3] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 136–144, 2017.

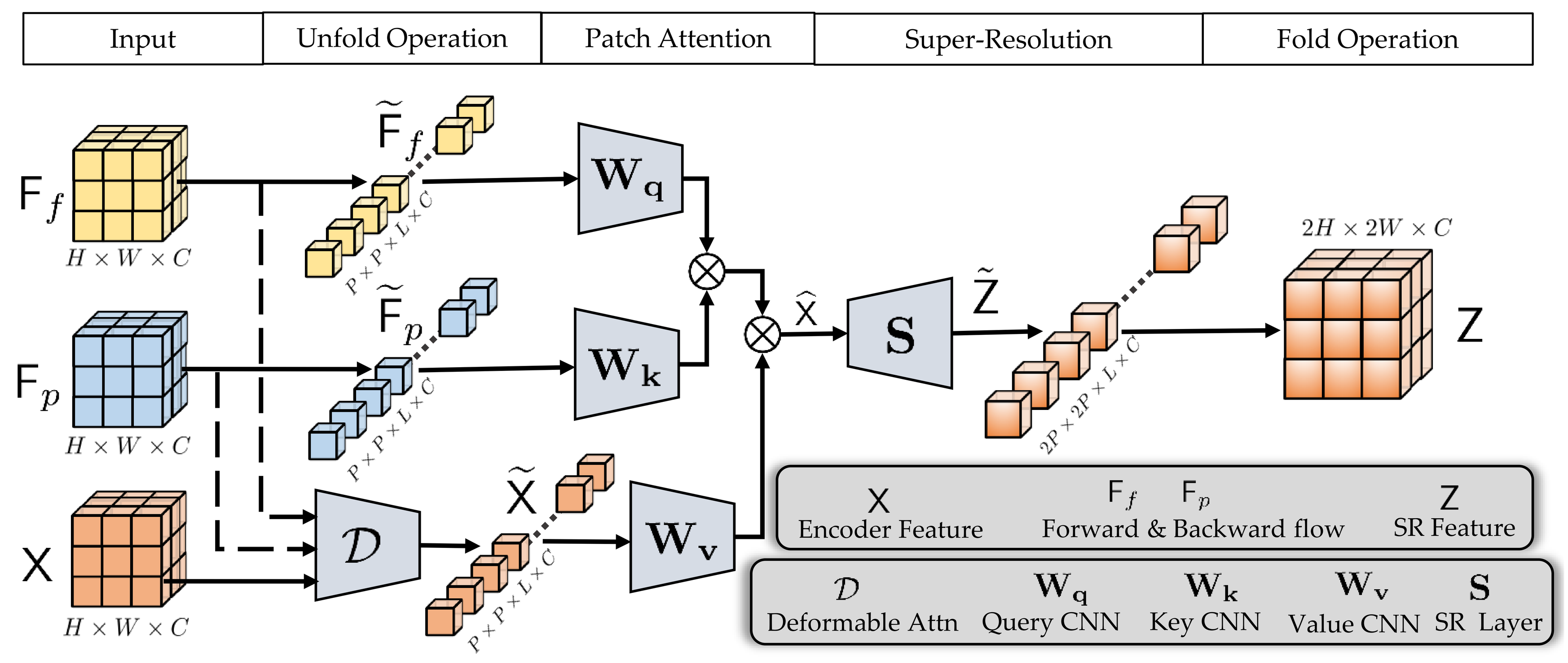
[4] P. Liu, Z. Cui, V. Larsson, and M. Pollefeys, "Deep shutter unrolling network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5941–5949, 2020.

[5] Z. Zhong, Y. Zheng, and I. Sato, "Towards rolling shutter correction and deblurring in dynamic scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9219–9228, 2021.

Approach Overview



Patch Attention Network



Quantitative Results

Table 2: **Quantitative results comparison of PatchNet with the state-of-the-art baselines.** We compare our approach with various combination of RSC model followed by an SR model. We demonstrate that **PatchNet** is able to generate HR global frames, with LR input, compared to approaches which only perform RSC with HR input (highlighted in red).

Methods		BS-RSCD			FastecRS		
RSC	SR	PSNR ↑	SSIM ↑	LPIPS↓	PSNR ↑	SSIM ↑	LPIPS ↓
JCD [5] with LR Input	Bi-linear Interpolation	22.74	0.581	0.463	23.87	0.655	0.339
	Transposed Convolution	24.15	0.628	0.328	24.12	0.632	0.262
	SAN [2]	24.37	0.633	0.305	24.07	0.643	0.281
	EDSR [3]	24.94	0.650	0.263	24.67	0.713	0.187
DUN [4] with LR Input	Bi-linear Interpolation	21.64	0.552	0.489	25.34	0.792	0.185
	Transposed Convolution	24.02	0.602	0.342	25.88	0.801	0.179
	SAN [2]	24.16	0.621	0.322	26.10	0.807	0.165
	EDSR [3]	24.58	0.634	0.286	26.43	0.810	0.147
JCD [5] with HR Input		<u>26.42</u>	<u>0.757</u>	**0.122**	24.84	0.778	<u>0.107</u>
DUN [4] with HR Input		<u>25.14</u>	0.729	0.159	<u>27.00</u>	**0.825**	0.108
PatchNet with LR Input		**27.38**	**0.793**	<u>0.144</u>	**27.12**	<u>0.811</u>	**0.103**

Qualitative Examples