연세대학교
YONSEI UNIVERSITY

Spatio-Temporal Coherent Multi-View Video Interpolation

머신러닝과응용
강승윤, 이승호

CVML

Introduction

- **VR contents**
  - **Change in view**

    Requiring high resolution light field for natural interaction.

  - **Property of light field**

    Requiring high dimensionality at least 5D, such as $L(h, w, u, v, c)$.

  - **What if VR content requires 4K resolution?**

Introduction

- **Obstacle of real-time VR streaming**
  - **Limited network environment**

    Assuming maximum 5G network speed as **20Gbps.**

    Minimum throughput for a 4K video is **25Mbps.**

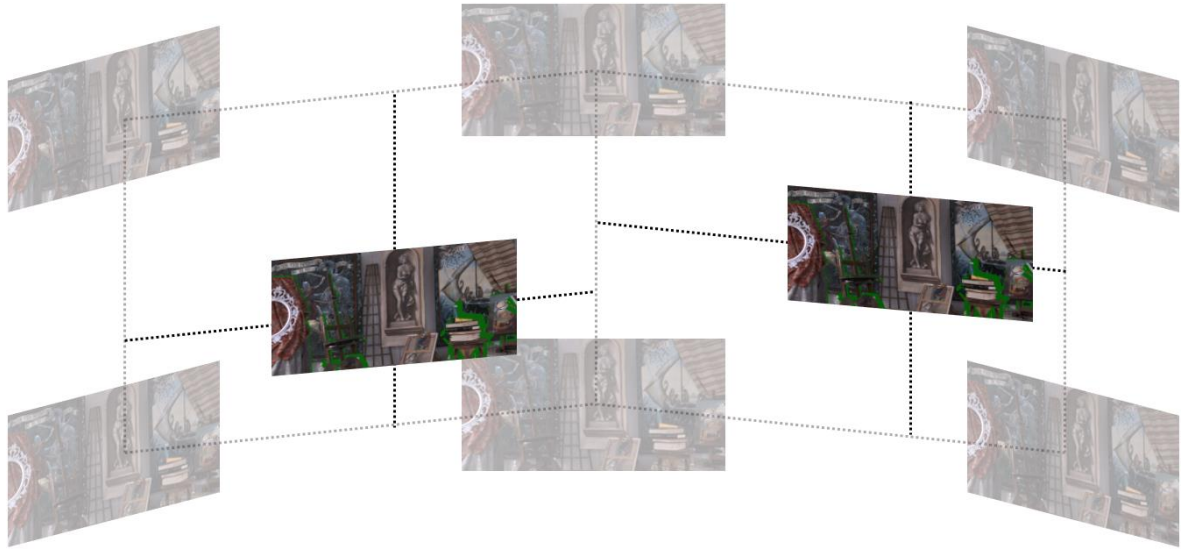
    = 800 4K videos! = 28 x 28 viewpoint.

  - **Demand for reduce light field dimensions.**

Introduction

- **Obstacle of real-time VR streaming**
  - **Possible solution**

    A view synthesis via images from multi-posed camera.

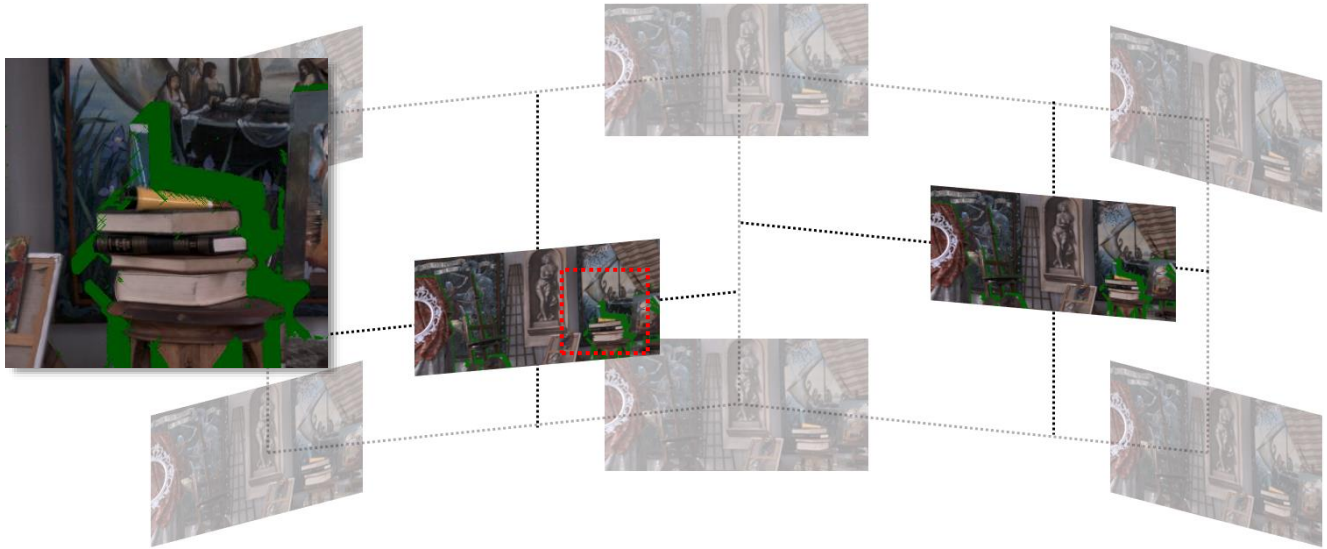


Reference View Synthesizer, MPEG. https://mpeg.chiariglione.org/meetings/124

Introduction

- **Obstacle of real-time VR streaming**
  - **Problem of view synthesis with spatially quantized video**

    Blind spot occurs. No view captures the point!



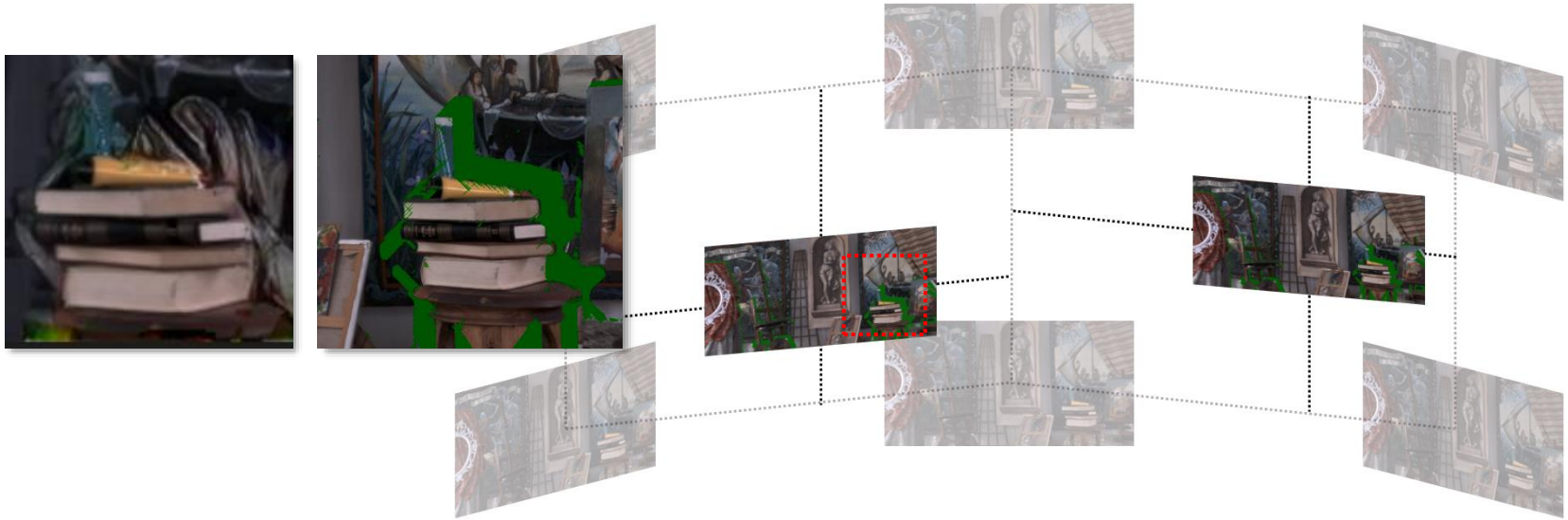
Blind spot of synthesized view.

Reference View Synthesizer, MPEG. https://mpeg.chiariglione.org/meetings/124

Introduction

- **Obstacle of real-time VR streaming**
  - **Solution for blind spot?**

    Exploiting image inpainting technology.



Blind spot of synthesized view.

Reference View Synthesizer, MPEG. https://mpeg.chiariglione.org/meetings/124

Introduction

- **Obstacle of real-time VR streaming**
  - **Solution for blind spot?**

    Exploiting image inpainting technology.

  - **Problems**

    Inpainting quality

    Temporal consistency

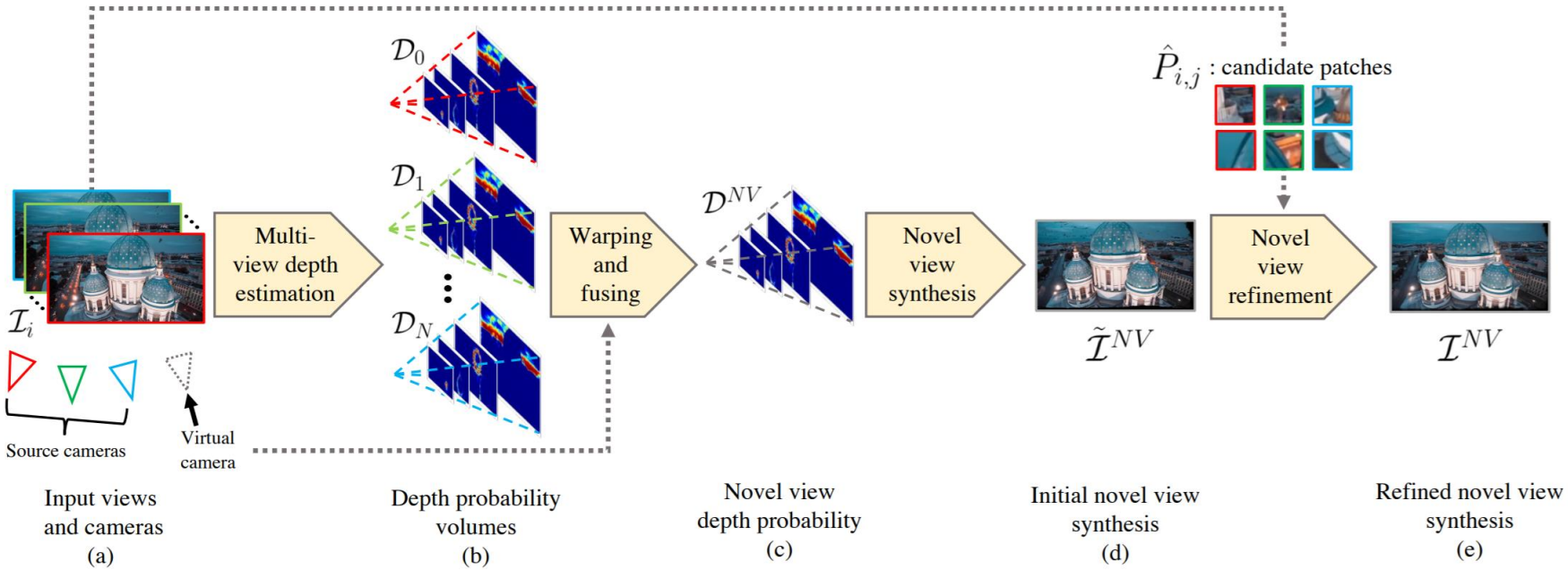
Introduction

- **Problem statement**
  - Building **spatio-temporal coherent** multi-view interpolation framework.

    1. Utilizing view synthesis net and temporal consistent video inpainting net.
    2. Coupling spatial consistency and temporal consistency.

Related work

- **Spatial consistent view synthesis**
  - Extreme view synthesis

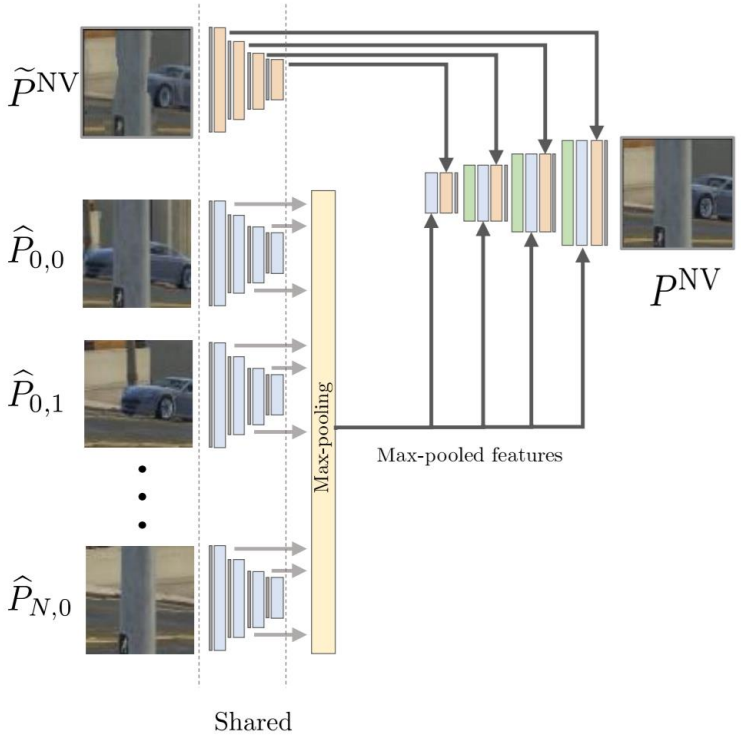


Method overview

Choi, Inchang, et al. "Extreme view synthesis." Proceedings of the IEEE International Conference on Computer Vision. 2019.

Related work

- **Spatial consistent view synthesis**
  - Extreme view synthesis



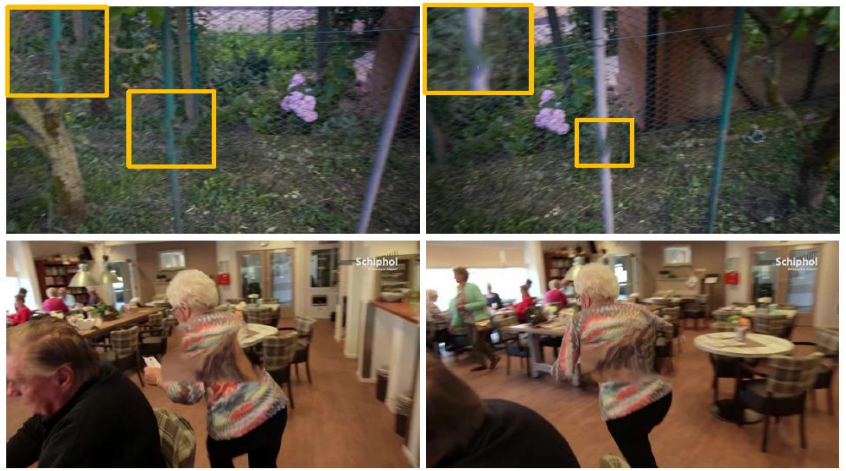
Refinement network

Choi, Inchang, et al. "Extreme view synthesis." Proceedings of the IEEE International Conference on Computer Vision. 2019.

Related work

- **Temporal consistent video inpainting**
  - Deep Flow-Guided Video Inpainting

missing region



Huang *et al.*

Xu *et al.* *

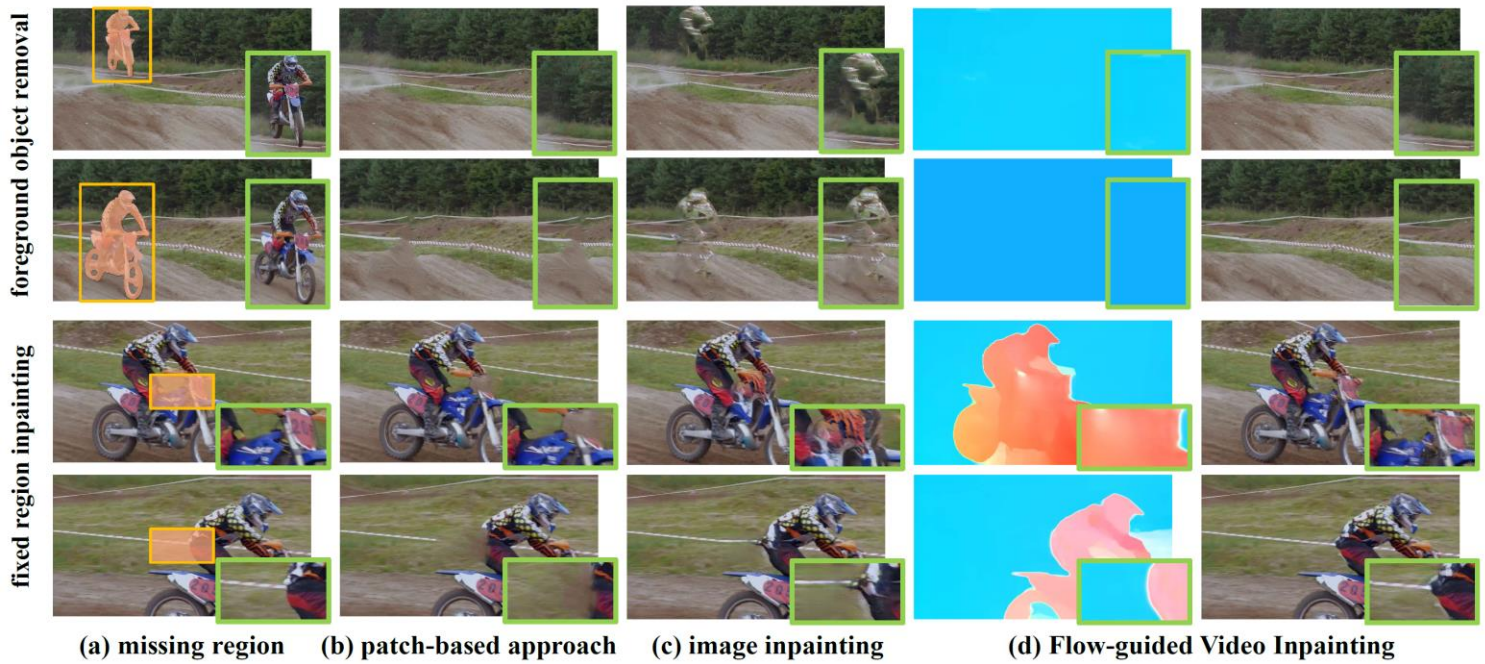
Temporal coherent video inpainting using single video.

Xu, Rui, et al. "Deep Flow-Guided Video Inpainting." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

Huang, Jia-Bin, et al. "Temporally coherent completion of dynamic video." *ACM Transactions on Graphics (TOG)* 35.6 (2016): 196.

Related work

- **Temporal consistent video inpainting**
  - Deep Flow-Guided Video Inpainting

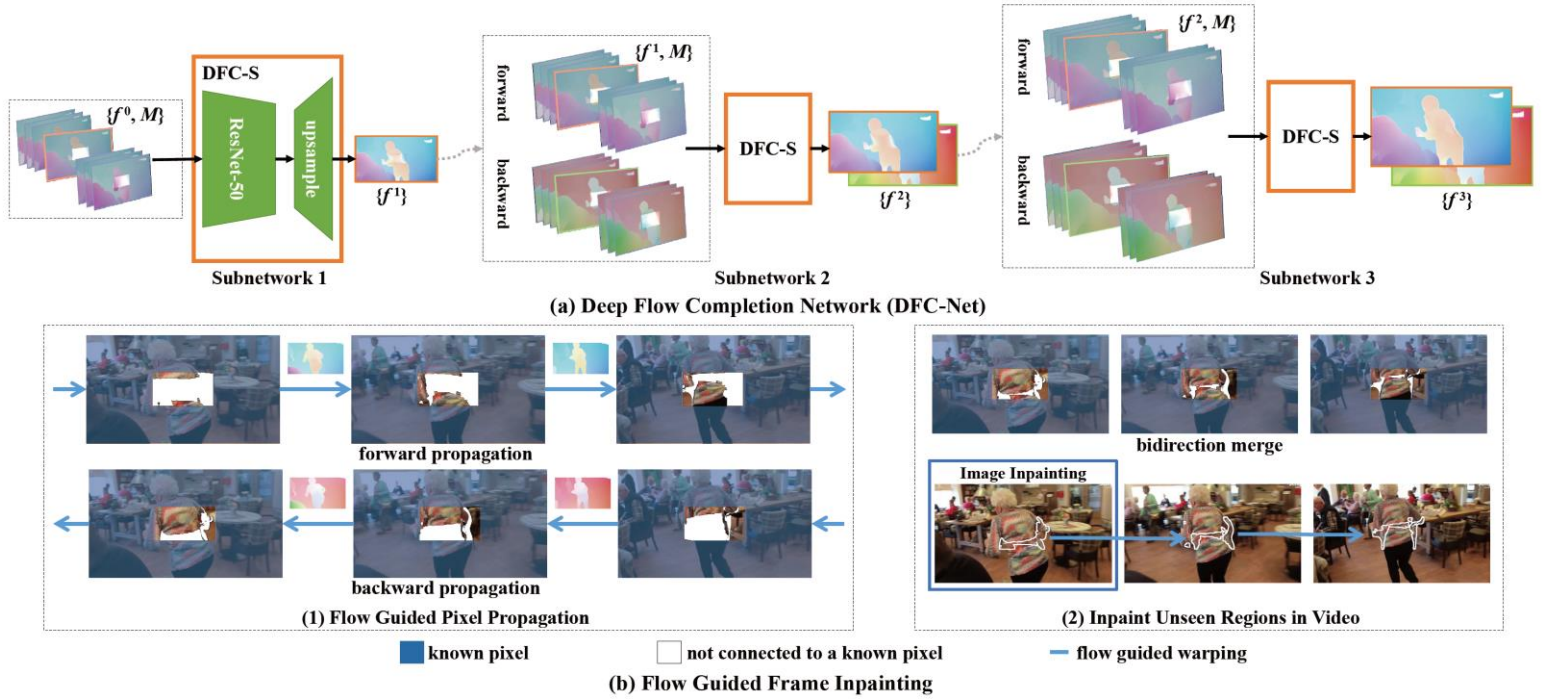


Comparison to several image inpainting works.

Xu, Rui, et al. "Deep Flow-Guided Video Inpainting." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

Related work

- **Temporal consistent video inpainting**
  - Deep Flow-Guided Video Inpainting

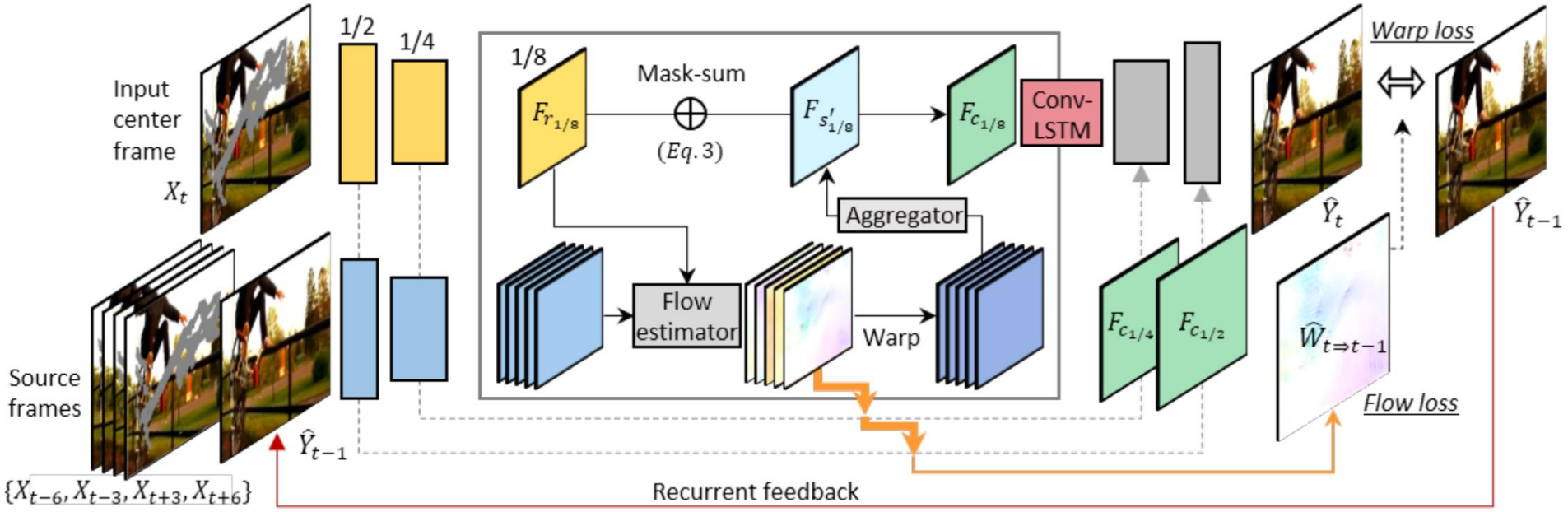


The pipeline of deep flow-guided video inpainting apporach.

Xu, Rui, et al. "Deep Flow-Guided Video Inpainting." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

Related work

- **Temporal consistent video inpainting**
  - Deep Video Inpainting

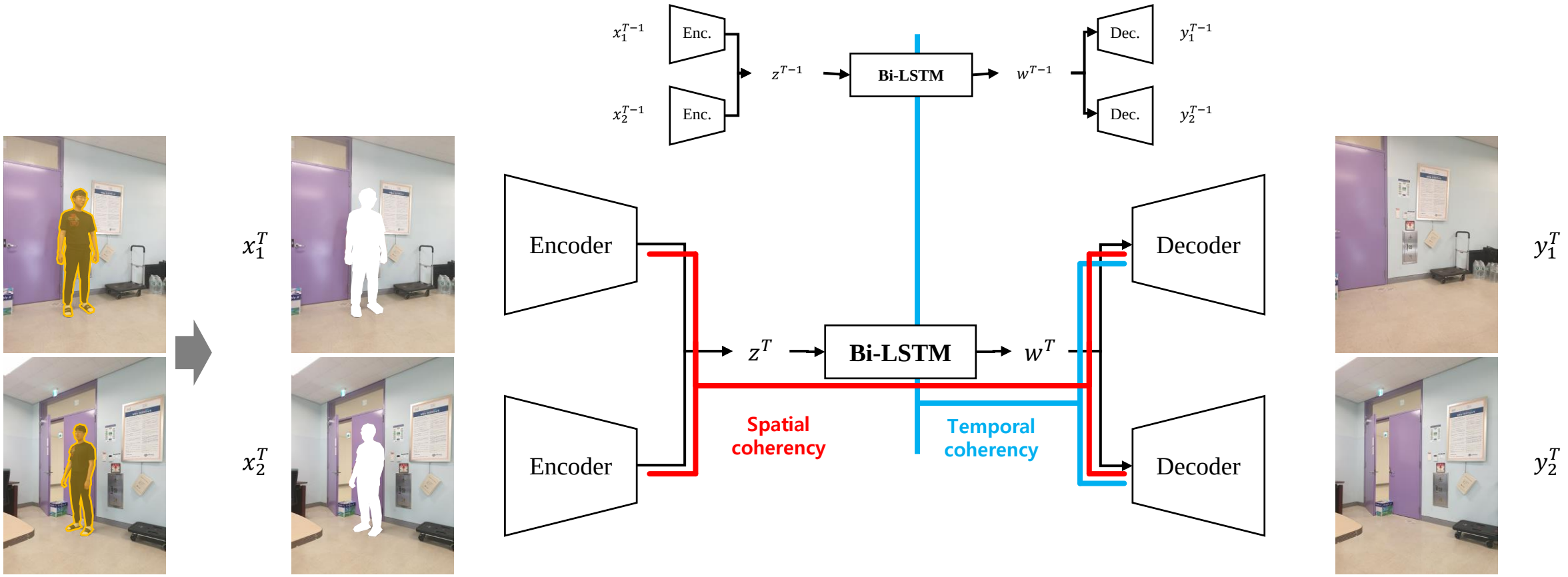


The overview of VINet

Kim, Dahun, et al. "Deep video inpainting." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.

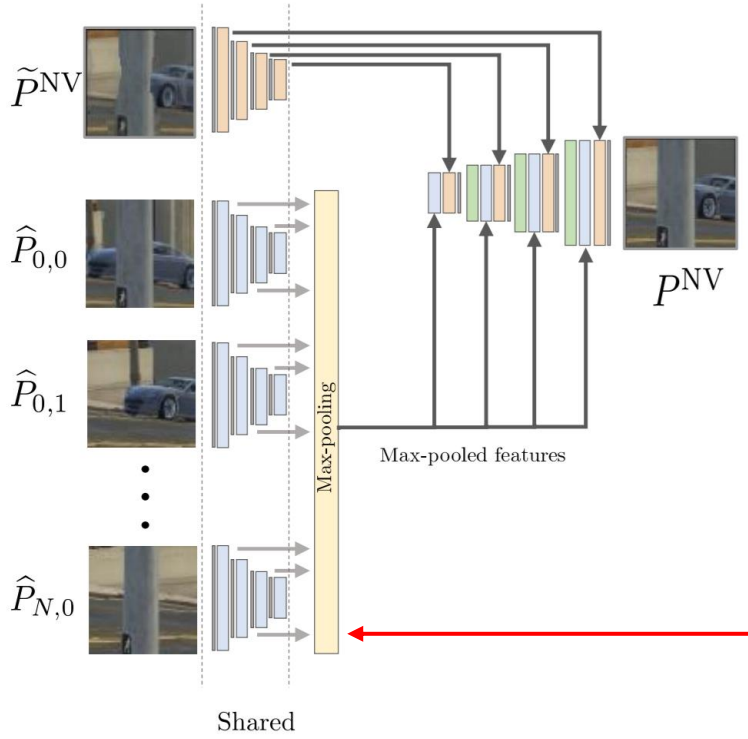
Proposed work

- **Framework overview**

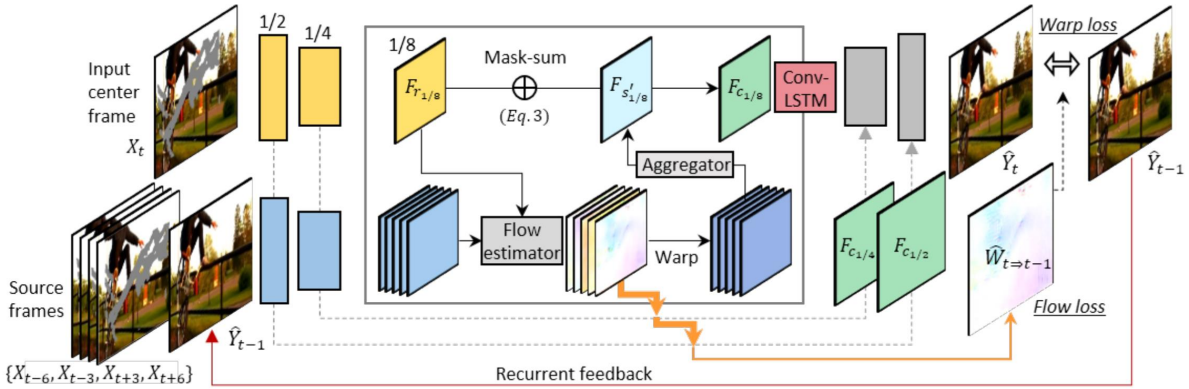


Proposed work

- **Coupling spatial-temporal coherence**



Refinement network



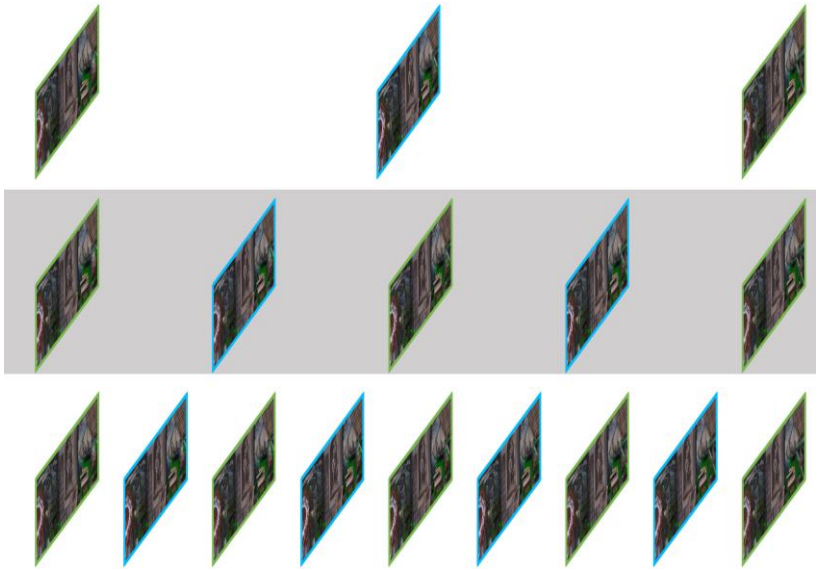
The overview of VINet

Recurrent feedback

Kim, Dahun, et al. "Deep video inpainting." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
Choi, Inchang, et al. "Extreme view synthesis." Proceedings of the IEEE International Conference on Computer Vision. 2019.

Proposed work

- **Reducing computational complexity**
  - Coupling spatial-temporal consistency only on key frames.
  - Then, fill all other blind spots.



Kim, Dahun, et al. "Deep video inpainting." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
Choi, Inchang, et al. "Extreme view synthesis." Proceedings of the IEEE International Conference on Computer Vision. 2019.

Results

- **Qualitative analysis**