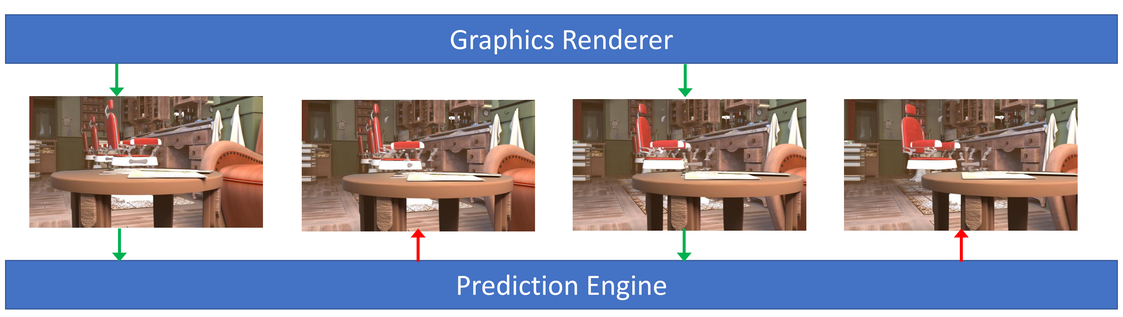
Temporal View Synthesis of Dynamic Scenes through 3D Object Motion Estimation with Multi-Plane Images

Nagabhushan Somraj **Pranali Sancheti** **Rajiv Soundararajan**

Indian Institute of Science (IISc), Bengaluru, India.

Scan for paper, code, dataset and more!

Temporal View Synthesis of Dynamic Scenes



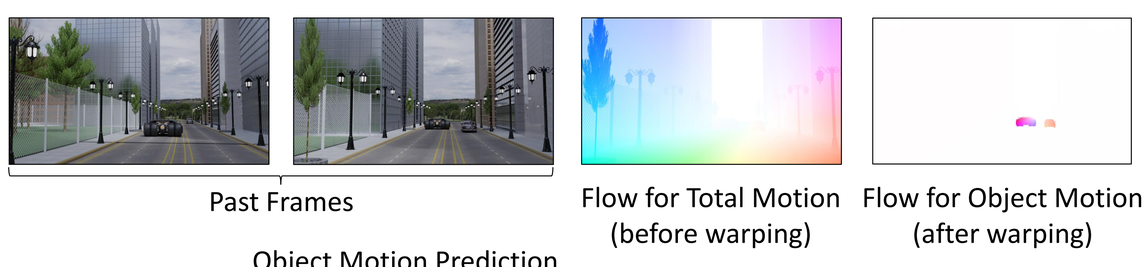
Goal: Predict future frames given past RGB-D frames and relative camera motion.

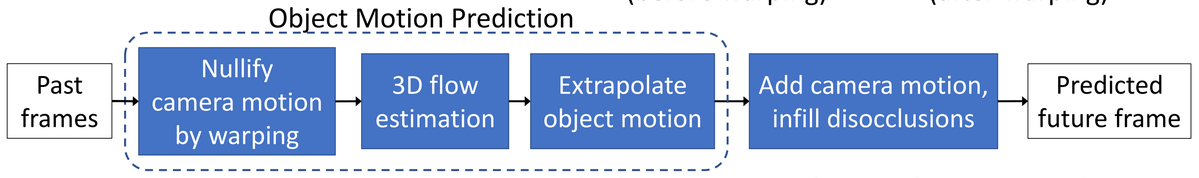
Application: Causal frame-rate upsampling of graphically rendered videos in virtual reality or first-person PC games.

Challenge: Camera/User motion is known;
Object motion is unknown – to be predicted.

Contribution 1 – Object Motion Isolation

Approach: Estimate past object motion as optical flow and extrapolate it.
Challenge: Motion between past frames is mix of camera and object motion.
Solution: Isolate object motion by warping past frames to same view.





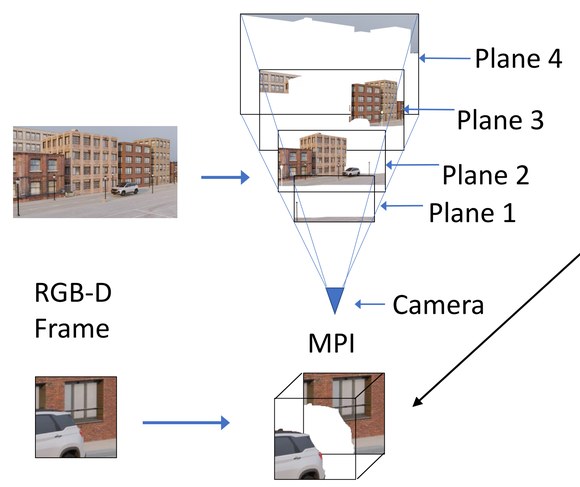
DeCOMPnet: **De**composing **C**amera and **O**bject **M**otion for **P**redicting next frame.

Off-the-Shelf Optical Flow Estimation

Problem: Off-the-shelf Optical Flow estimation creates distortions in predicted frames.

Cause: Lack of matching points in occluded areas; flow estimated in occluded areas is a mix of neighbourhood foreground and background flow.

Contribution 2 – Multi-Plane Images for 3D Flow Estimation



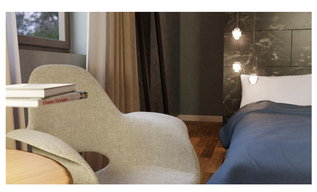
Multi-Plane Image (MPI): Splits a single RGB frame into multiple planes at different depths.

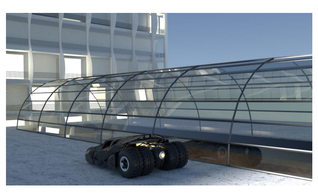
Moving car and static building are separated onto different planes.
Flow in occluded regions depends only on the static background

Challenge: Empty regions in MPI can lead to incorrect matching.
Solution: Introduce masked correlation and convolution layers.

IISc Virtual Environment Exploration Database (VEED) - Dynamic Scenes


Kitchen

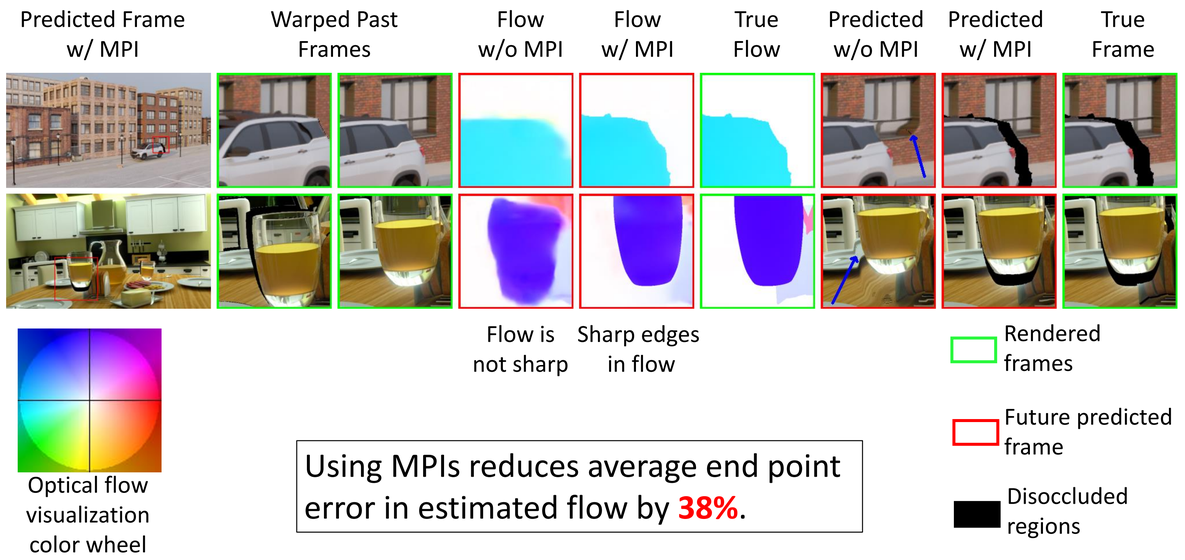

Bedroom


Batmobile on a Building


Millenium Falcon in a Forest

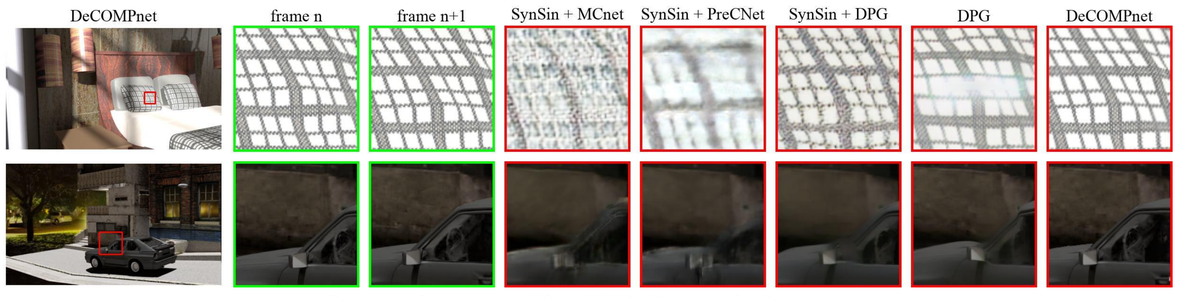
- Developed using Blender
- 200 scenes – 800 videos
- Resolution: 1920 x 1080; 30fps
- Indoor and outdoor scenes
- Data:
  - RGB frames
  - Depth map
  - Camera poses

Future Frame Prediction with and without MPI

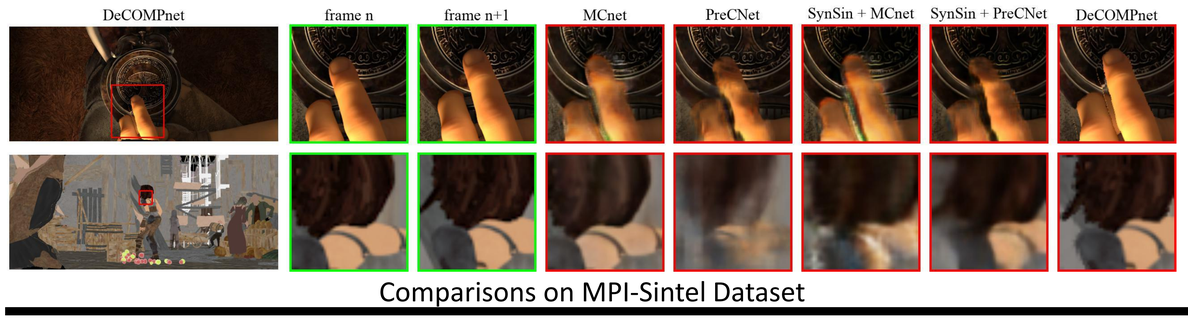


Using MPIs reduces average end point error in estimated flow by **38%**.

Results – Qualitative Comparisons

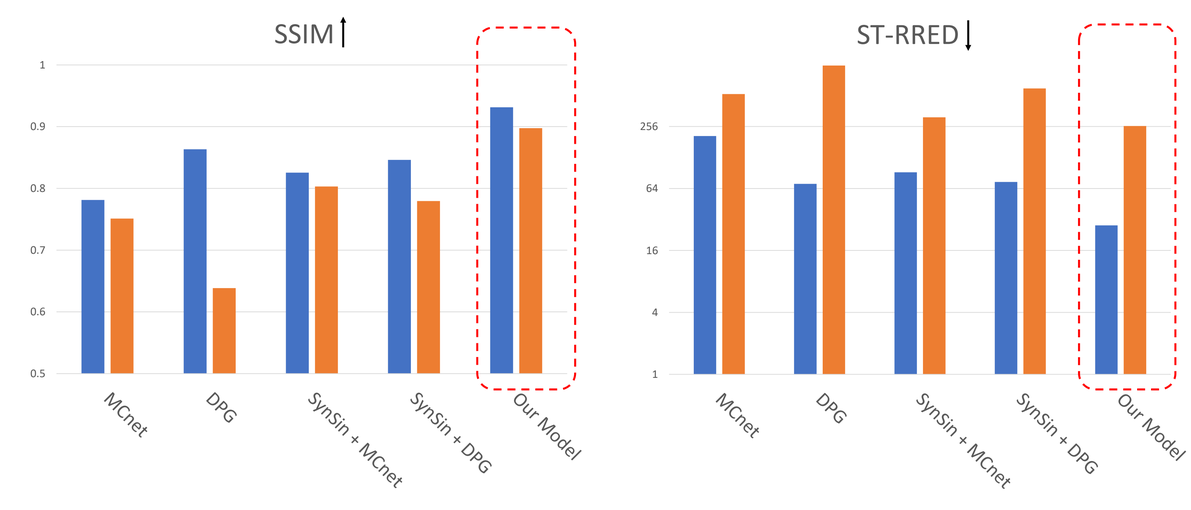


Comparisons on IISc VEED-Dynamic Dataset



Comparisons on MPI-Sintel Dataset

Results – Quantitative Comparisons



Acknowledgements:
This work was supported in part by a grant from Qualcomm. The first author was supported by the Prime Minister's Research Fellowship (PMRF) awarded by the Ministry of Education (MoE), Government of India.