Exploiting spatial attention mechanism for improved depth completion and feature fusion in novel view synthesis

Anh M. Truong, Wilfried Philips, Peter Veelaert
IPI-TELIN, Ghent University-IMEC

Introduction

Existing novel view synthesis methods:

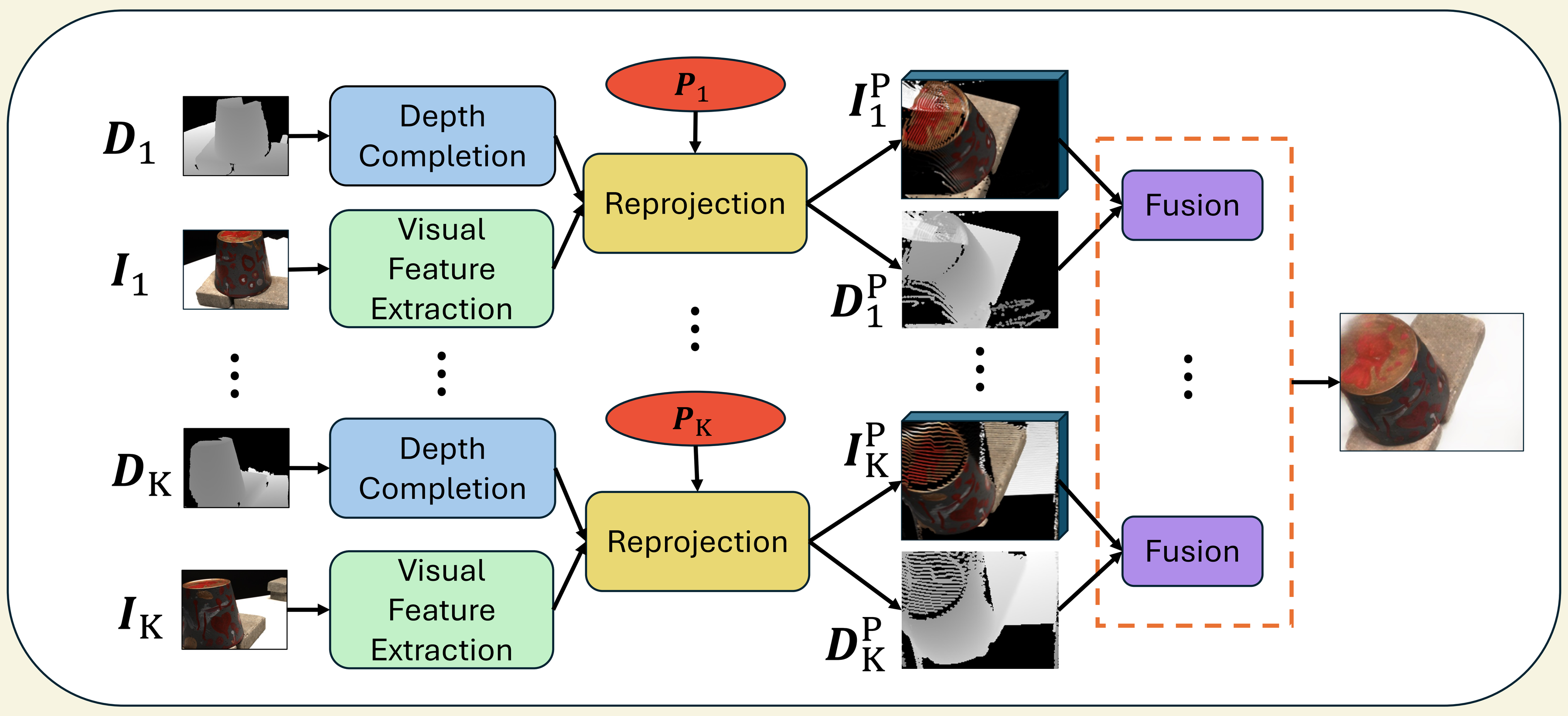
Approaches	Unseen	Sparse	Speed	Flexibility
Image-based rendering	green	green	green	yellow
Neural Scene Representation	red	green	red	green
Generalized Neural Scene Representation	green	yellow	yellow	green
Plane-sweep volume	green	yellow	yellow	green

Goal: Given a set of images and corresponding depth maps from sparse camera setup
➔ render novel view images with deep image-based rendering technique

Limitations of prior methods: artifacts (blurry textures) caused by
- Incorrect geometry estimates.
- Attention mechanism focus solely on the view and channel axes.

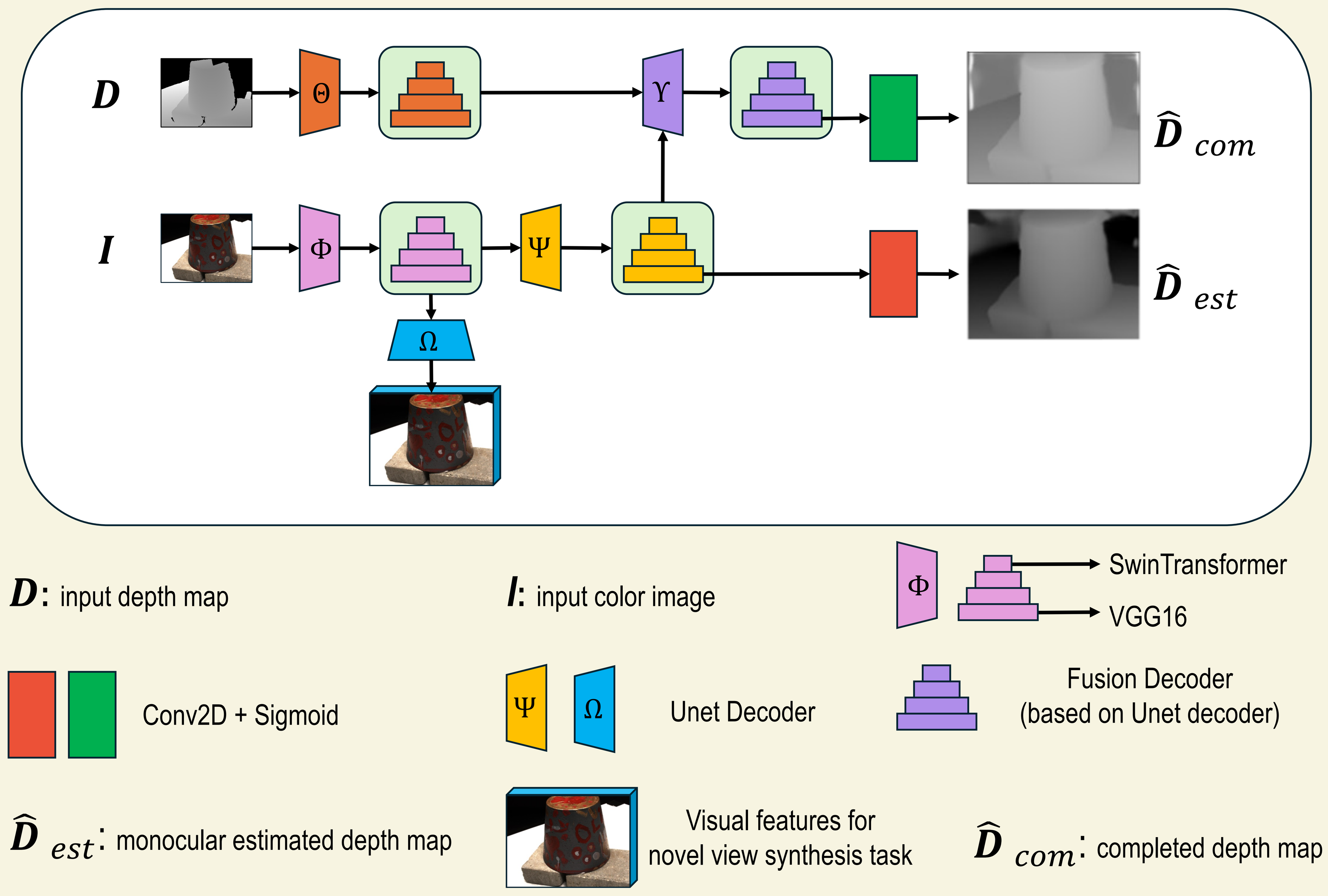
Key Insights:
- Extract **spatial features from color images** by leveraging a monocular depth estimation pipeline.
- Combine spatial features from color images to enhance depth completion.
- Combine an **attention mechanism on spatial, channel, and view axes** to enhance the rendering results (SpatialAttentionGRU).

Framework

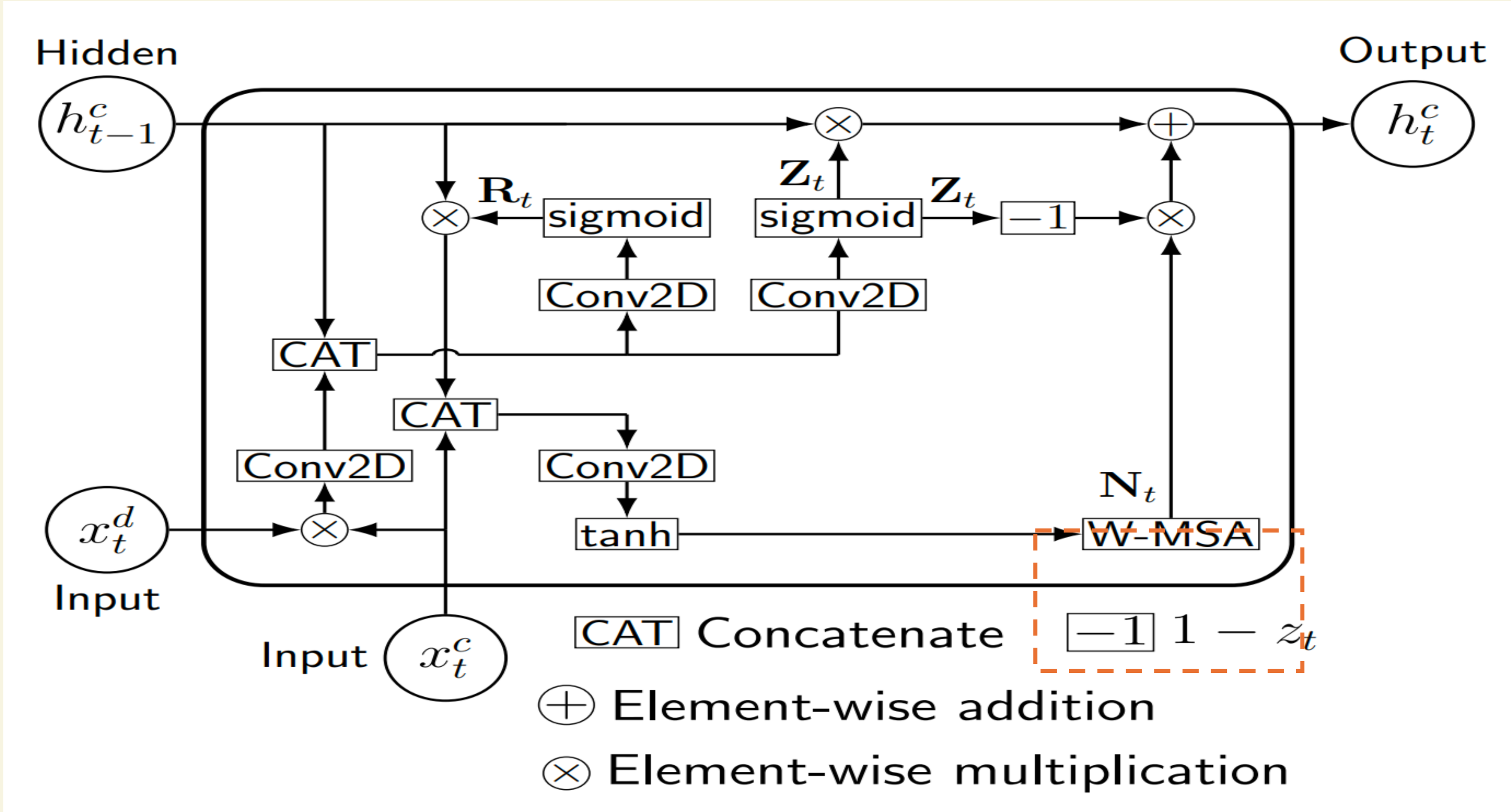


$D_1,\ldots,D_K$: input depth maps
$D_1^P,\ldots,D_K^P$: "reprojected" depth maps
$I_1,\ldots,I_K$: input color maps
$I_1^P,\ldots,I_K^P$: reprojected color maps
$P_1,\ldots,P_K$: projection matrices

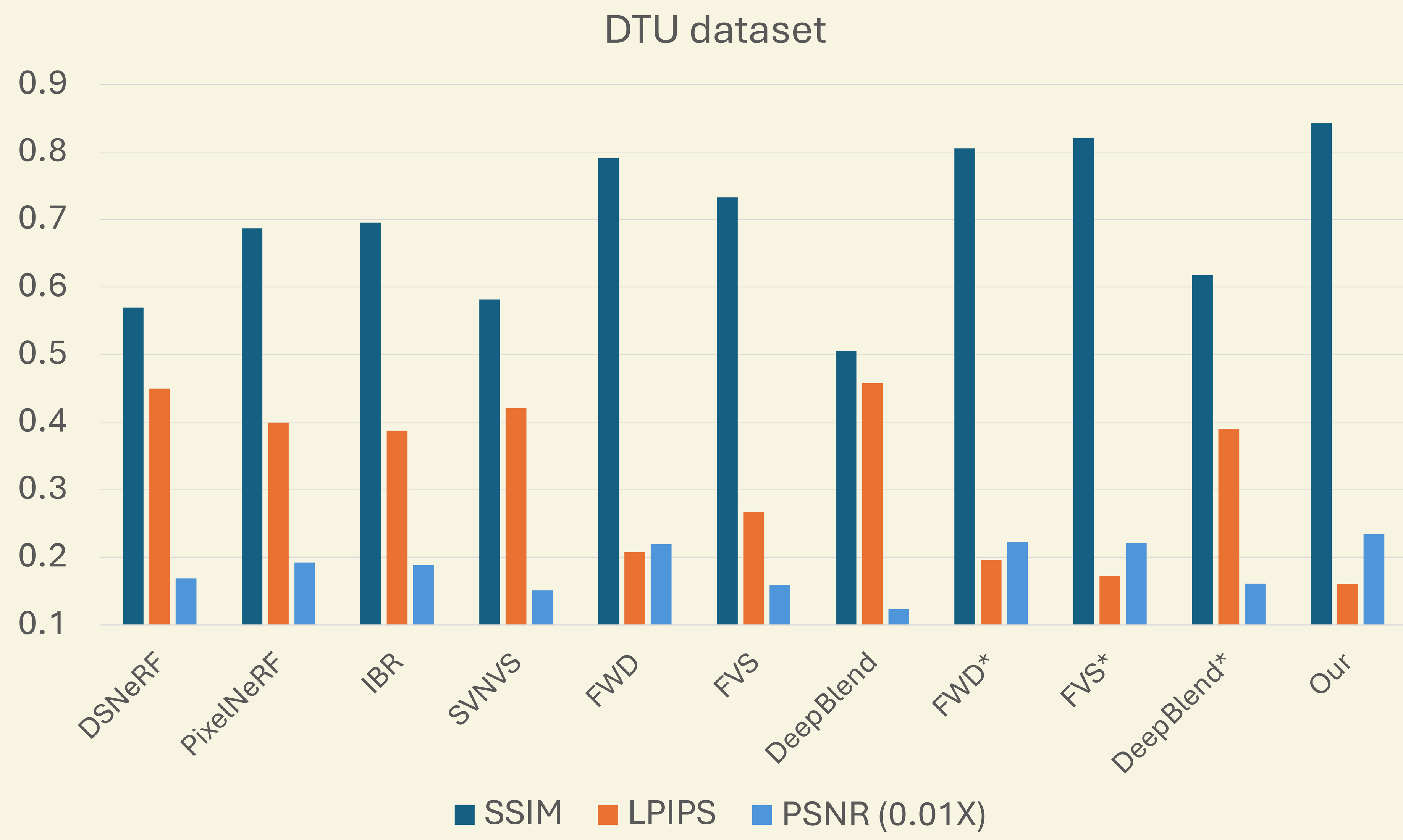
Depth Completion



Novel View Synthesis



Experimental results



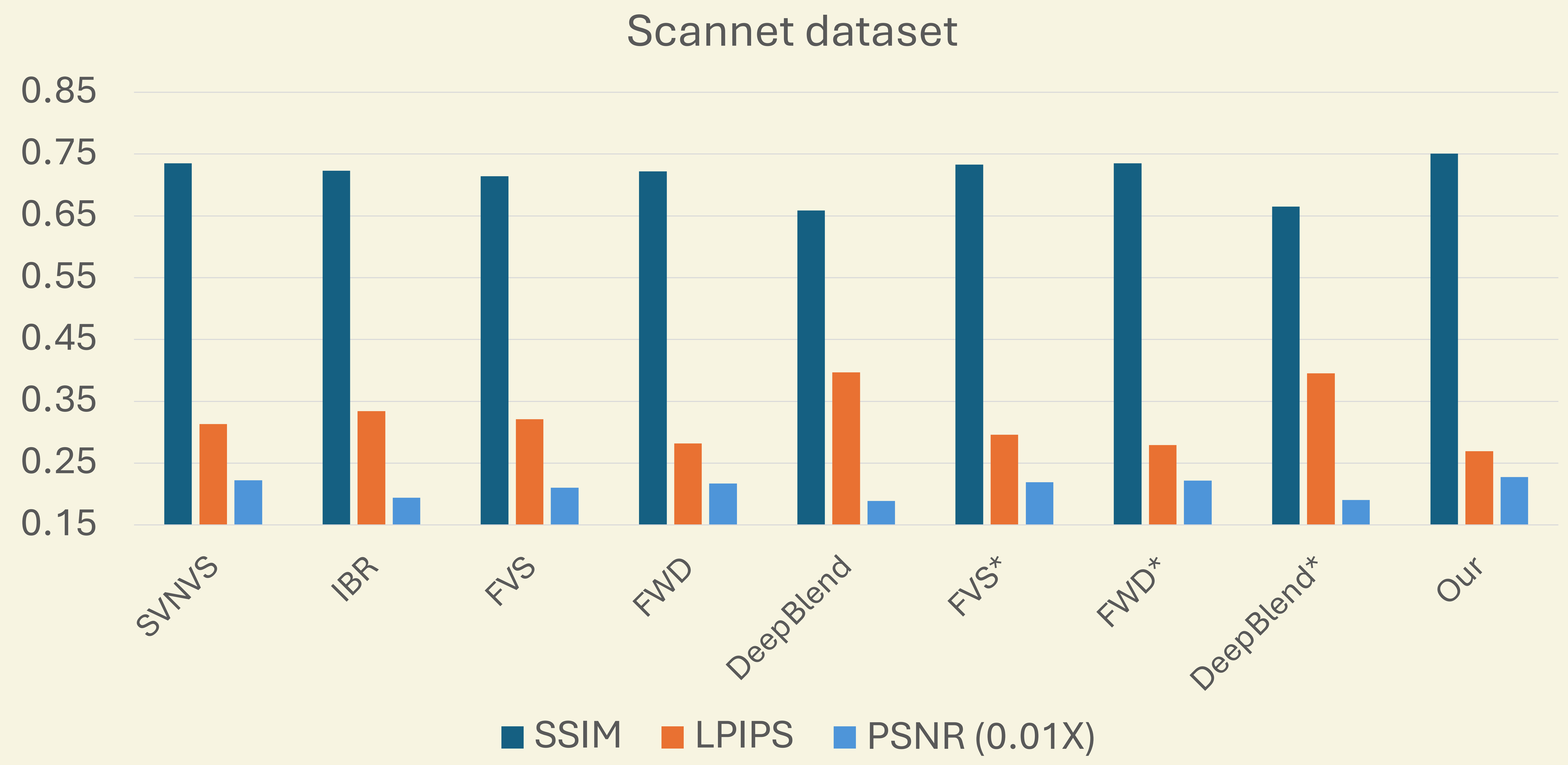


*Table 4.2: **Ablation study on DTU MVS dataset:** We presented the results of the novel view synthesis by replacing the proposed depth completion module and the feature fusion module with different methods. The best results are in bold.*

Feature Extraction	Depth	Fusion Completion	PSNR	SSIM	LPIPS
VGG	-	FVS [125]	15.92	0.73	0.27
VGG	Our	FVS [125]	22.12	0.82	0.17
Swin	-	Our	19.03	0.78	0.22
Swin	FWD [12]	Our	21.36	0.76	0.21
Swin	DDP [128]	Our	21.19	0.81	0.20
Swin	CA [1]	Our	21.71	0.82	0.19
Swin	Our	FVS [125]	23.02	0.84	0.17
Swin	Our	FWD [12]	23.00	0.80	0.19
Swin	Our	Our	**23.44**	**0.84**	**0.16**