NEURAL INFORMATION PROCESSING SYSTEMS

# Toward Dynamic Non-Line-of-Sight Imaging with Mamba Enforced Temporal Consistency

*Yue Li, Yi Sun, Shida Sun, Juntian Ye, Yueyi Zhang, Feihu Xu, Zhiwei Xiong*

University of Science and Technology of China
{yueli65, sunyi2017, jt141884, sdsun}@mail.ustc.edu.cn
{zhyuey, feihuxu, zwxiong}@ustc.edu.cn

**Poster**: Sat 14 Dec 8:30 a.m. CST — 11:30 a.m. CST

**Paper**: https://openreview.net/pdf?id=QiCJomIW3l

**Project**: https://github.com/Depth2World/Dynamic_NLOS

中国科学技术大学 University of Science and Technology of China

NEL-BITA National Engineering Laboratory for Brain-Inspired Intelligence Technology and Application

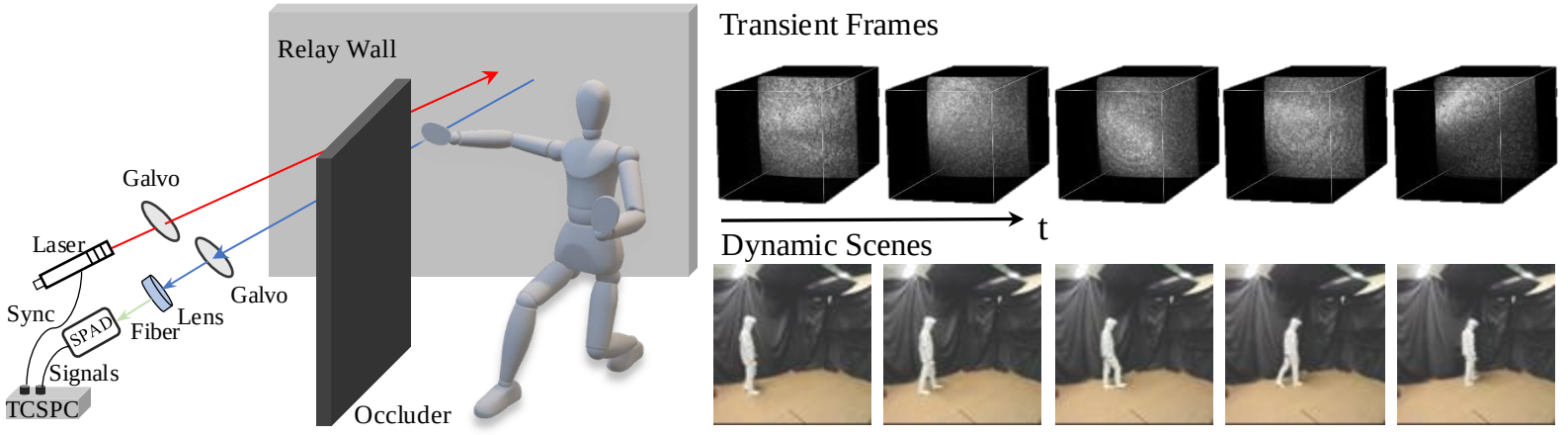
VIDAR Visual Information Discovery And Recovery

NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Dynamic Non-Line-of-Sight Imaging




1) Insufficient information fusion across adjacent transient frames: Existing methods, whether traditional or deep-based, typically concentrate on individual transient frames, overlooking the temporal consistency between them.

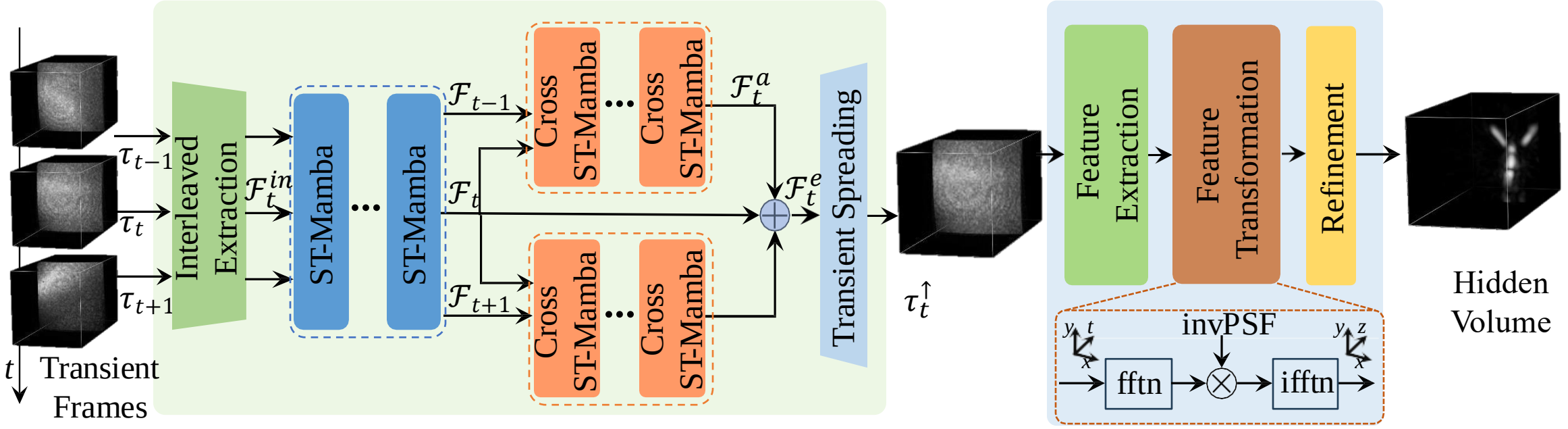
2) Lack of NLOS video datasets, including synthetic data for training and real-world data for evaluation. The rapid exposure time results in a diminished signal-to-noise ratio (SNR) of transient measurement, highlighting the critical need for simulation datasets that accurately emulate real-world conditions.

NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Proposed Method

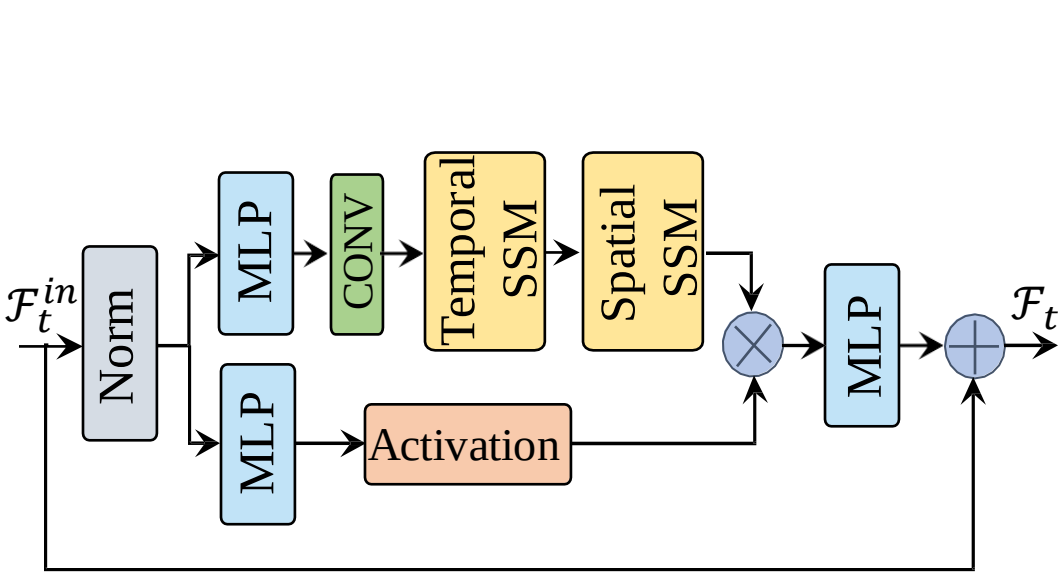


NEURAL INFORMATION PROCESSING SYSTEMS

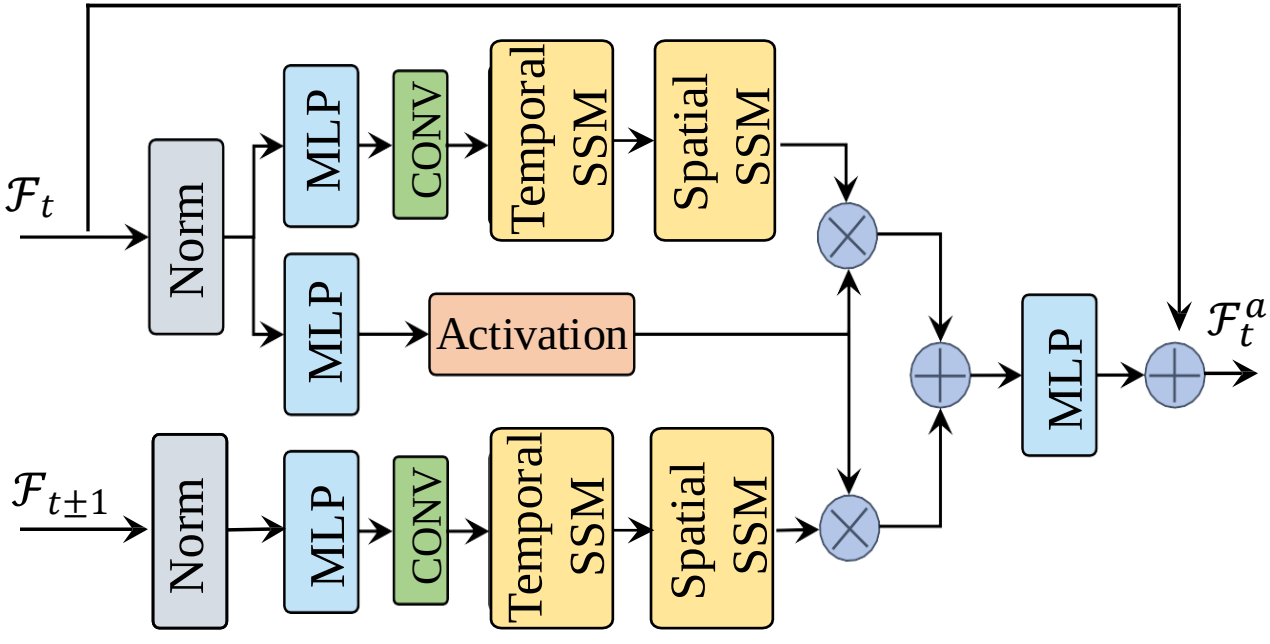
VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Proposed Method




(a) ST-Mamba




(b) Cross ST-Mamba

NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Results about Volume Reconstruction – Synthetic data

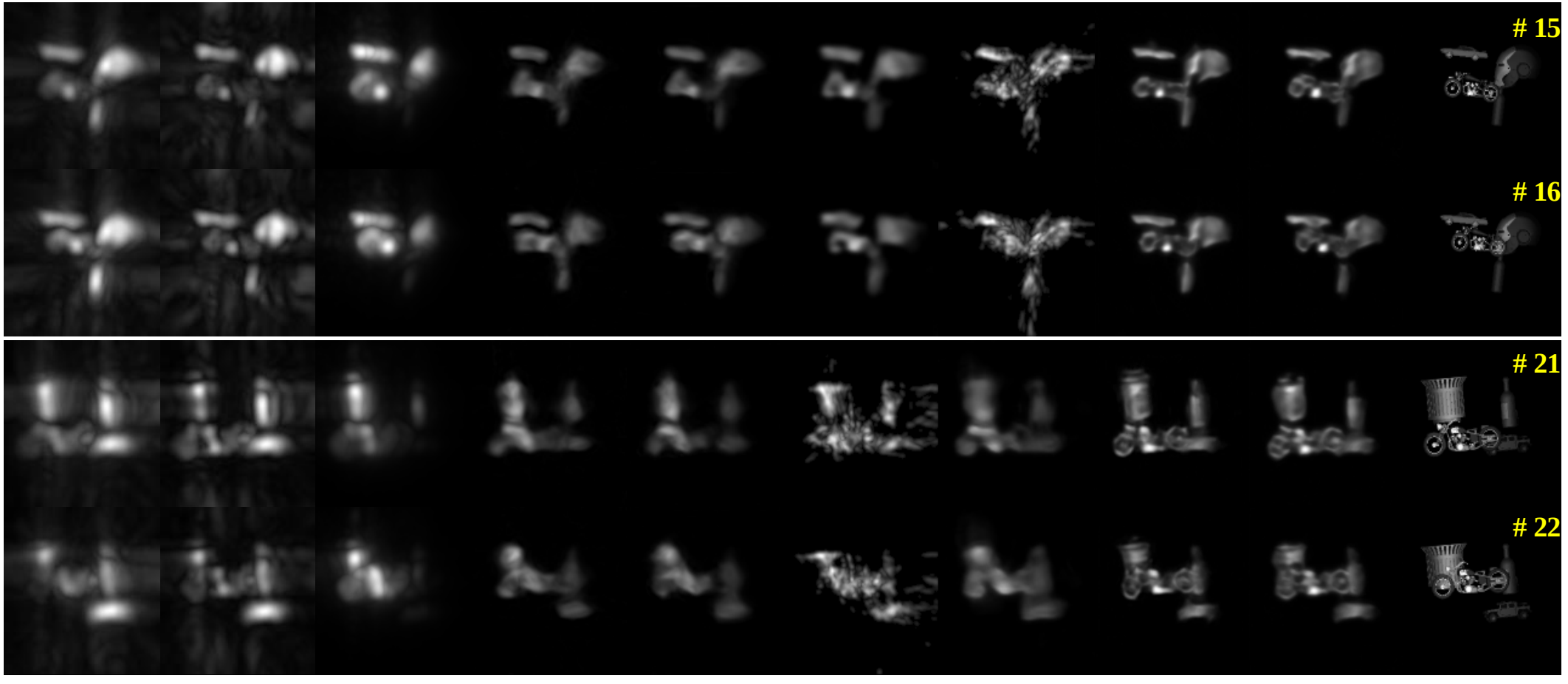
| Methods | Architecture | Intensity | | Depth | |
|---|---|---|---|---|---|
| | | PSNR↑ | SSIM↑ | RMSE↓ | MAD↓ |
| LCT [15] | Linear Optimation | 17.25 | 8.81 | 0.4355 | 0.4103 |
| RSD [42] | Phasor Field Waves | 19.00 | 13.48 | 0.4043 | 0.3844 |
| FK [21] | F-k Migration | 20.90 | 49.84 | 0.3930 | 0.3756 |
| LFE [23] | Physical-based | 23.20 | 78.02 | 0.0993 | 0.0526 |
| I-K [28] | Physical-based | 23.22 | 79.79 | 0.1011 | 0.0468 |
| CSA [1] | Linear Optimation | 20.70 | 71.13 | 0.2647 | 0.1090 |
| USM [4] | Physical-based | 23.80 | 80.85 | 0.0945 | 0.0432 |
| Ours-S | Physical-based | <u>23.97</u> | <u>81.35</u> | <u>0.0939</u> | <u>0.0400</u> |
| Ours | Physical-based | **24.46** | **84.08** | **0.0880** | **0.0397** |

NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Results about Volume Reconstruction – Synthetic data



NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Ablation Studies

- Loss items
- Spatial-temporal Mamba

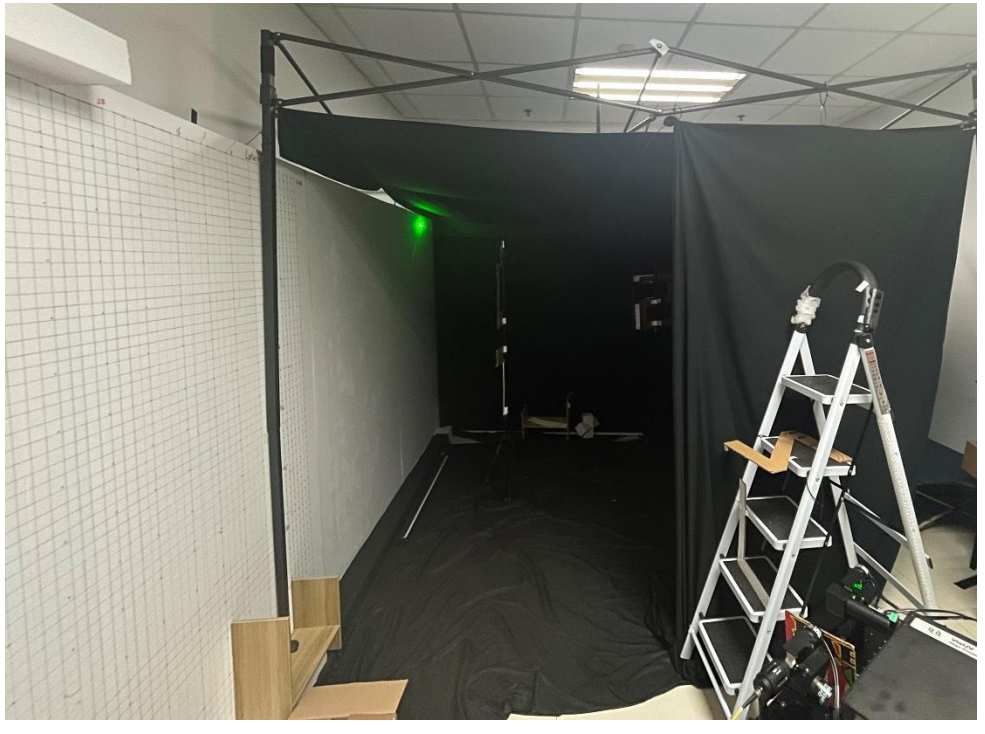
| ST-Mamba | | Loss Items | | | | Intensity | | Depth | |
|---|---|---|---|---|---|---|---|---|---|
| Spatial | Temporal | $\mathcal{L}_{int,dep}$ | $\mathcal{L}_t$ | $\mathcal{L}_{pf}$ | $\mathcal{L}_{ls,tv}$ | PSNR↑ | SSIM↑ | RMSE↓ | MAD↓ |
| S-Mamba | T-Mamba | ✓ | ✓ | × | × | 24.19 | 82.75 | 0.0946 | 0.0409 |
| S-Mamba | T-Mamba | ✓ | ✓ | × | ✓ | 24.18 | 83.10 | 0.0914 | 0.0409 |
| S-Mamba | T-Mamba | ✓ | ✓ | ✓ | × | **24.47** | 83.07 | 0.0905 | 0.0404 |
| S-Mamba | T-Mamba | ✓ | ✓ | ✓ | ✓ | 24.46 | **84.08** | **0.0880** | **0.0397** |
| T-Mamba | T-Mamba | ✓ | ✓ | ✓ | ✓ | 24.32 | 83.68 | 0.0898 | 0.0398 |
| - | T-Mamba | ✓ | ✓ | ✓ | ✓ | 24.38 | 82.49 | 0.0921 | 0.0478 |
| S-Mamba | S-Mamba | ✓ | ✓ | ✓ | ✓ | 24.31 | 83.08 | 0.0911 | 0.0496 |
| S-Mamba | - | ✓ | ✓ | ✓ | ✓ | 24.36 | 82.82 | 0.0938 | 0.0440 |

NEURAL INFORMATION PROCESSING SYSTEMS

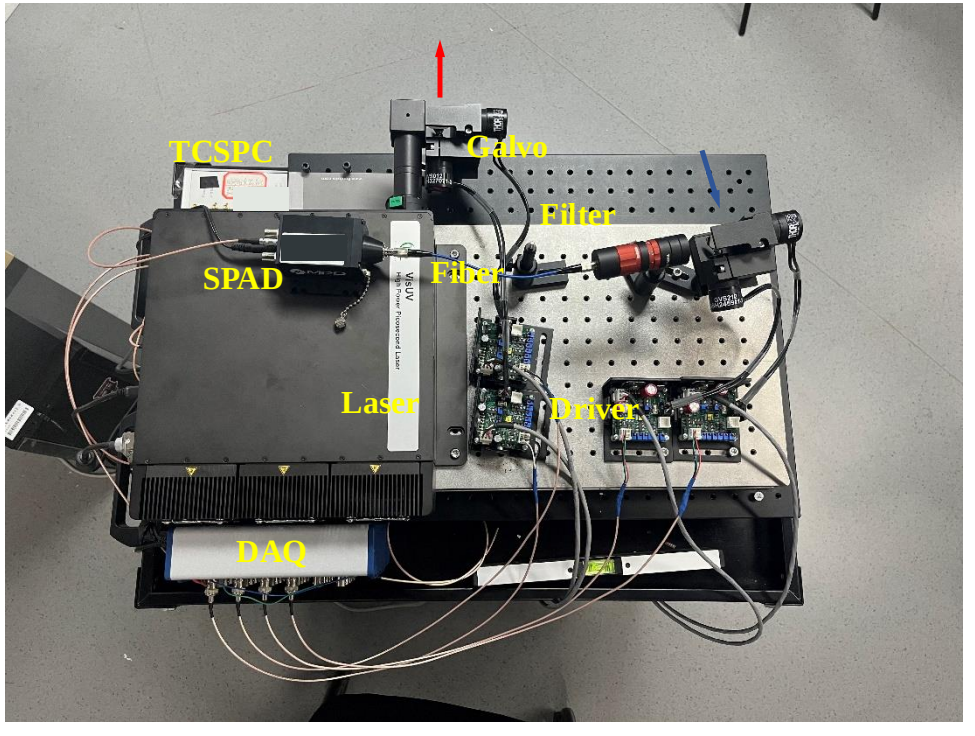
VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Imaging System

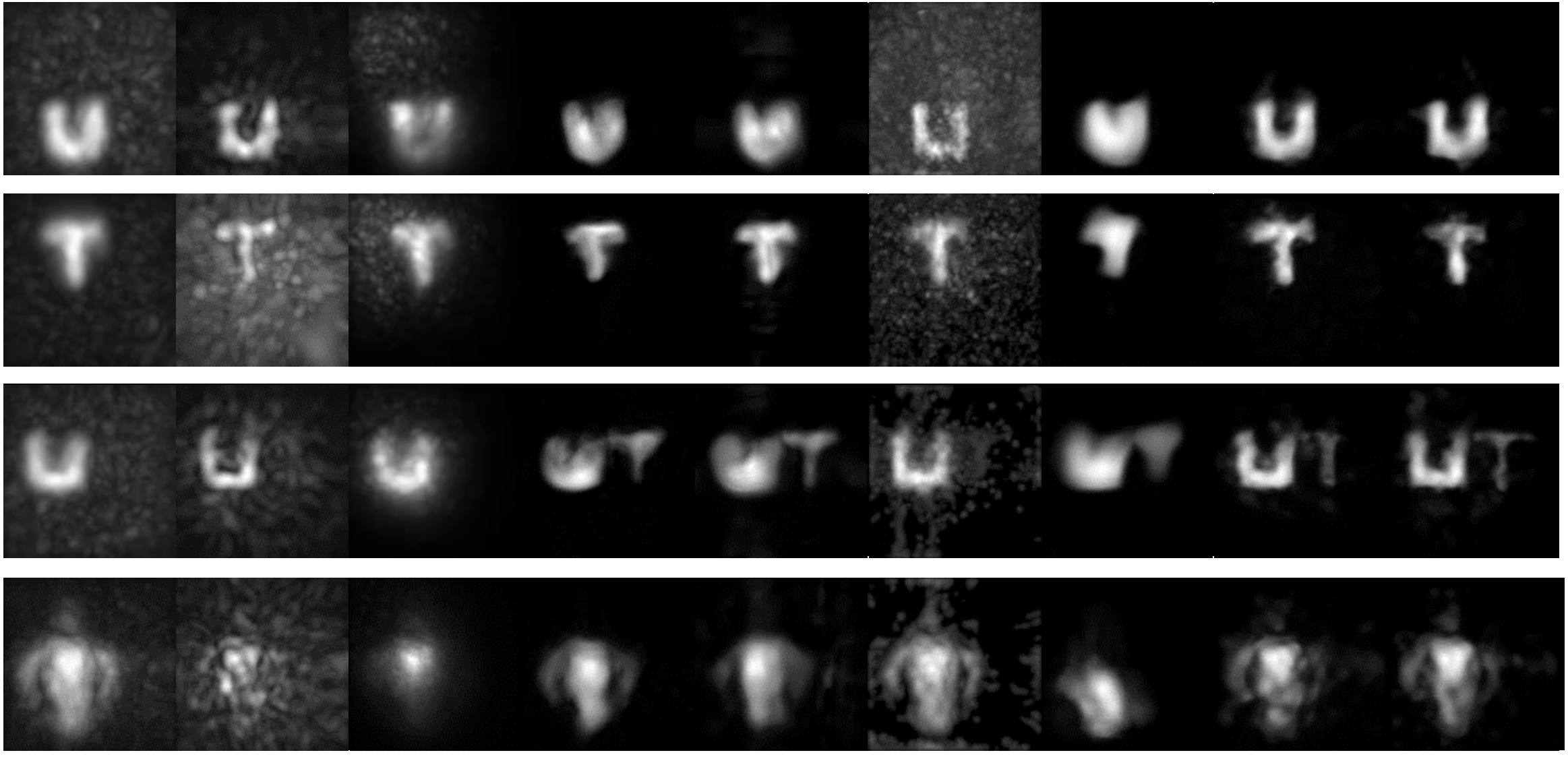


Capture Setup




System Setup

# Real-world Results



NEURAL INFORMATION PROCESSING SYSTEMS

VIDAR
Visual Information Discovery And Recovery

中国科学技术大学
University of Science and Technology of China

# Toward Dynamic Non-Line-of-Sight Imaging with Mamba Enforced Temporal Consistency

Yue Li Yi Sun Shida Sun Juntian Ye Yueyi Zhang Feihu Xu Zhiwei Xiong*

University of Science and Technology of China
{yueli65, sunyi2017, jt141884, sdsun}@mail.ustc.edu.cn
{zhyuey, feihuxu, zwxiong}@ustc.edu.cn