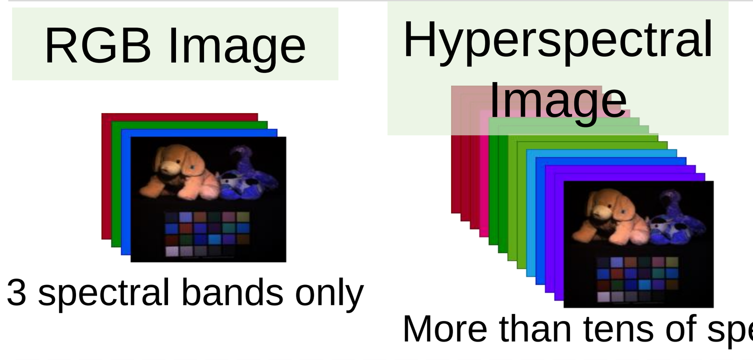
Deep Versatile Hyperspectral Reconstruction Model from a Snapshot Measurement with Arbitrary Masks

Paper ID: #1717

Takumi Takabe (Yamaguchi University), **Xian-Hua Han** (**Rikkyo University**, Japan), Yen-Wei Chen (Ritsumeikan University, Japan)

Research Background and Purpose:

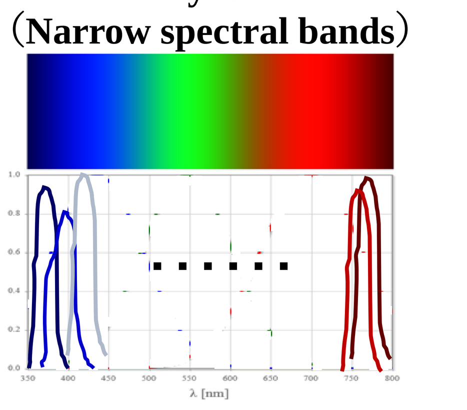


Abundance Spectral information
->Benefiting for many fields
- Remote Sensing
- Agriculture
- Medical diagnosis

Hyperspectral Imaging: Require to capture 3D cube

2D detector: Employ scanning method

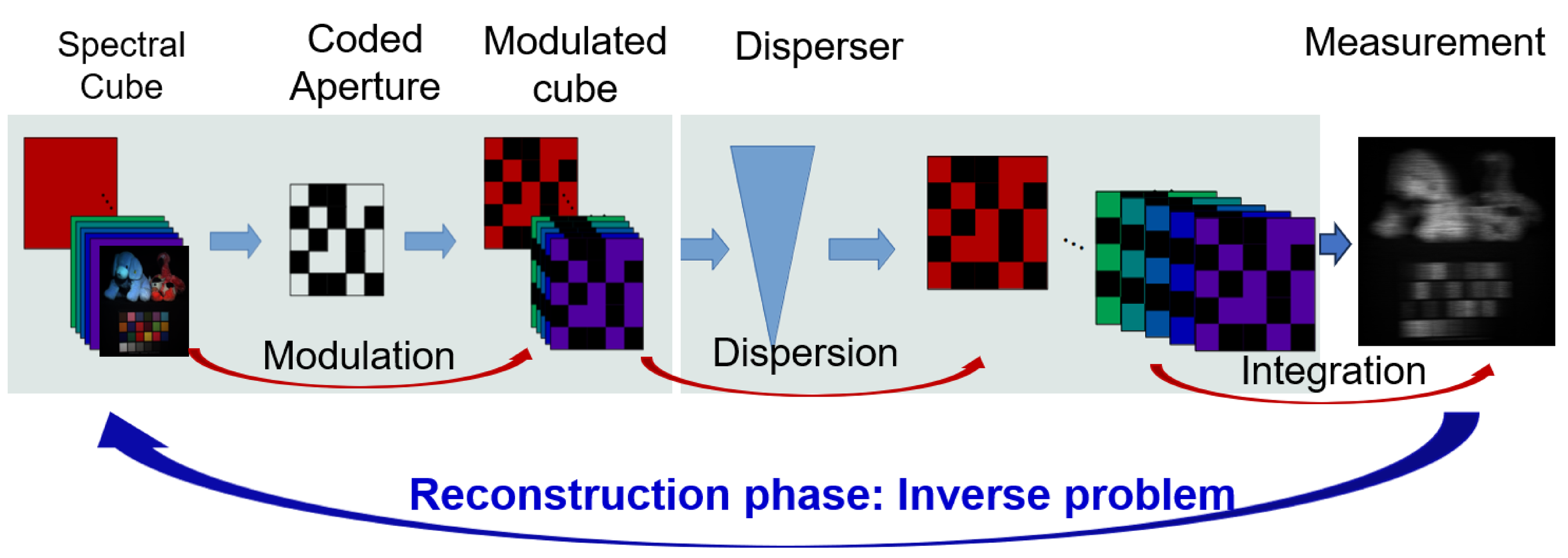
Challenge to capture 3D HSI in dynamic world



Coded aperture snapshot spectral imaging (CASSI)

Leverage Compressive theory to obtain a 2D snapshot measurement for 3D HSI cube

Detection Phase



Reconstruct the underlying 3D HS images from the measure snapshot

- **Challenge task with high compressive rate**
- **HSI reconstruction performance: Bottleneck of the CASSI**
- **Require reconstruction in the CASSI sensor: High-speed**

Related Work

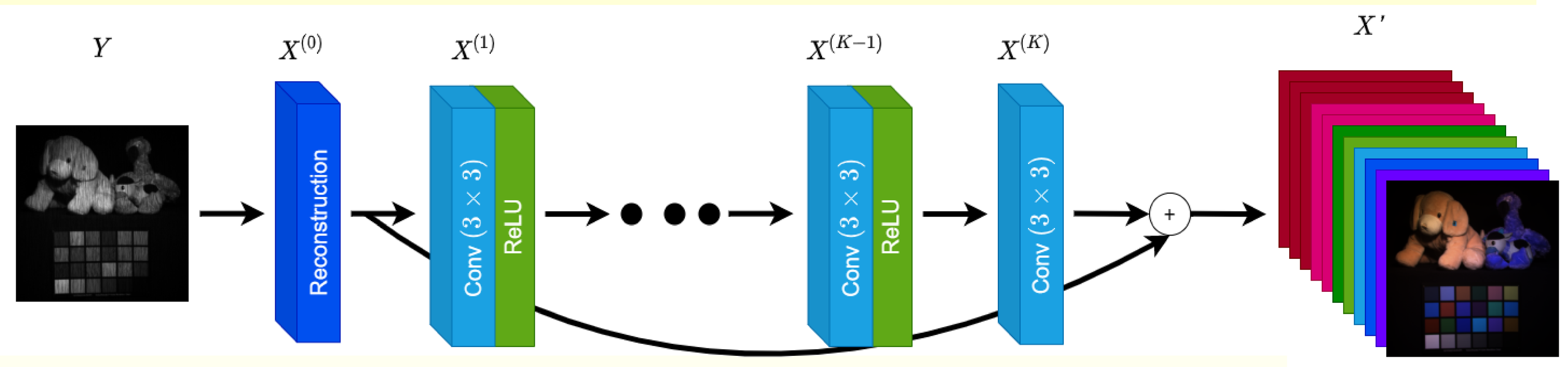
Model-based methods: Formulate the detection process as Mathematical Model

Optimization on the objective function

$$\hat{x} = \underset{x}{\operatorname{argmin}} \frac{1}{2} \| y - \Phi x \|_2^2 + \tau R(x)$$

Regularization term

- Optimization: Time-consuming
- Empirical prior: insufficient to capture diverse structure of HSI



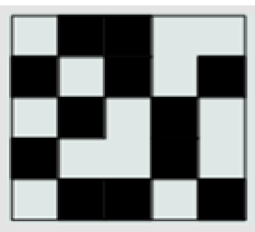
Existing deep learning models

Compared with model-based methods
- Better reconstruction performance
- Faster inference time

Assume

- **Fixed** and **small size** sensing mask in detection phase
  → Low generalization

Sensing mask (coded Aperture) in detection phase



- Different optical designs in the coded aperture
- Different imaging conditions
  → **Diverse sensing masks**

Proposed Method: Versatile model

Flexible HSI reconstruction for various masks: High generalization

Detection phase

Employ diverse masks

Synthesizing the training samples with different distribution

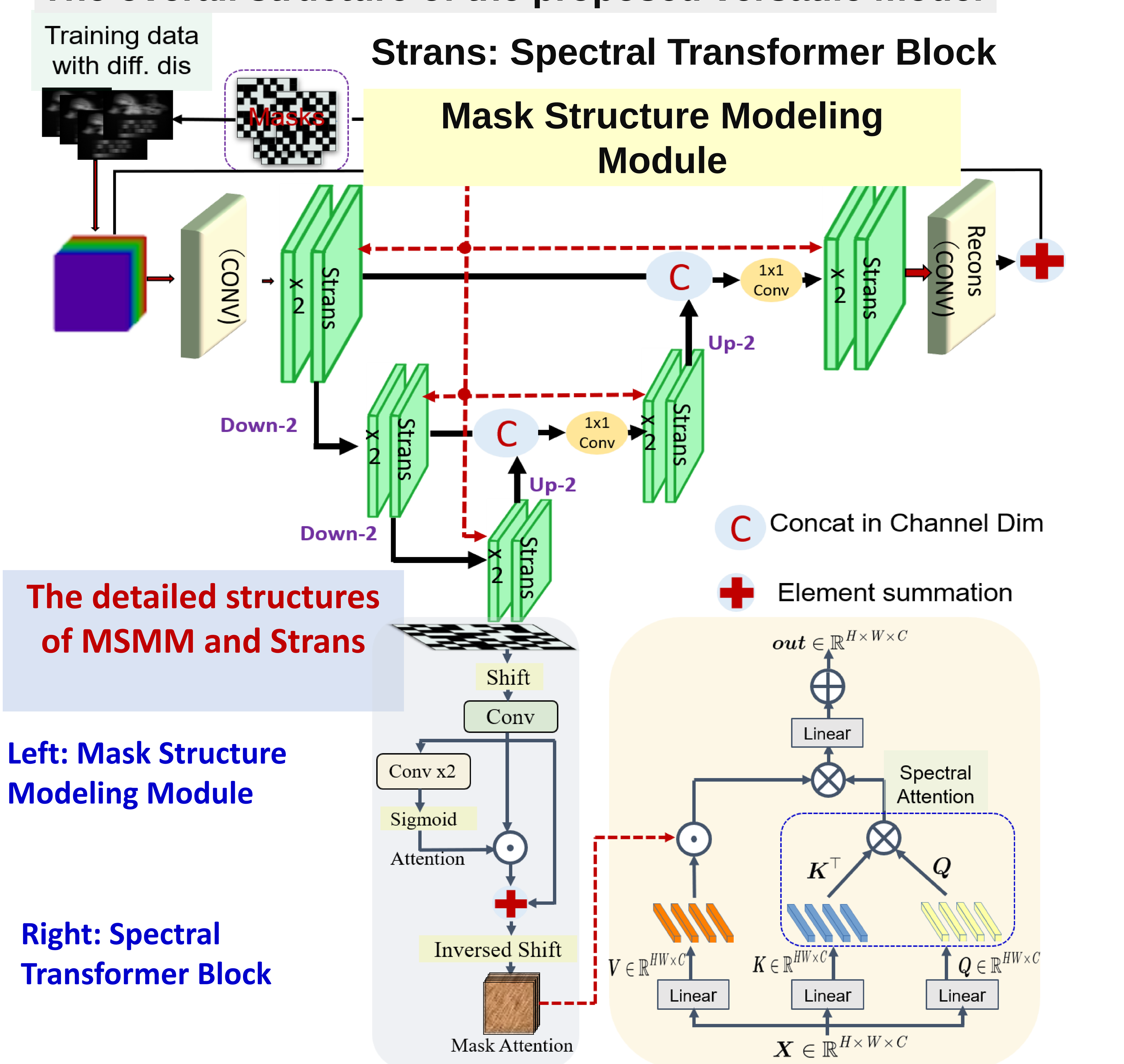
Reconstruction phase

Mask-aware reconstruction model

Backbone architecture: Unet-like Spectral Transformer

→Capture the long-range dependence among spectra

The overall structure of the proposed versatile model



Experimental Results:

Datasets: **CAVE (32 HS Images) 512x512x31**
20 training images; 12 test images

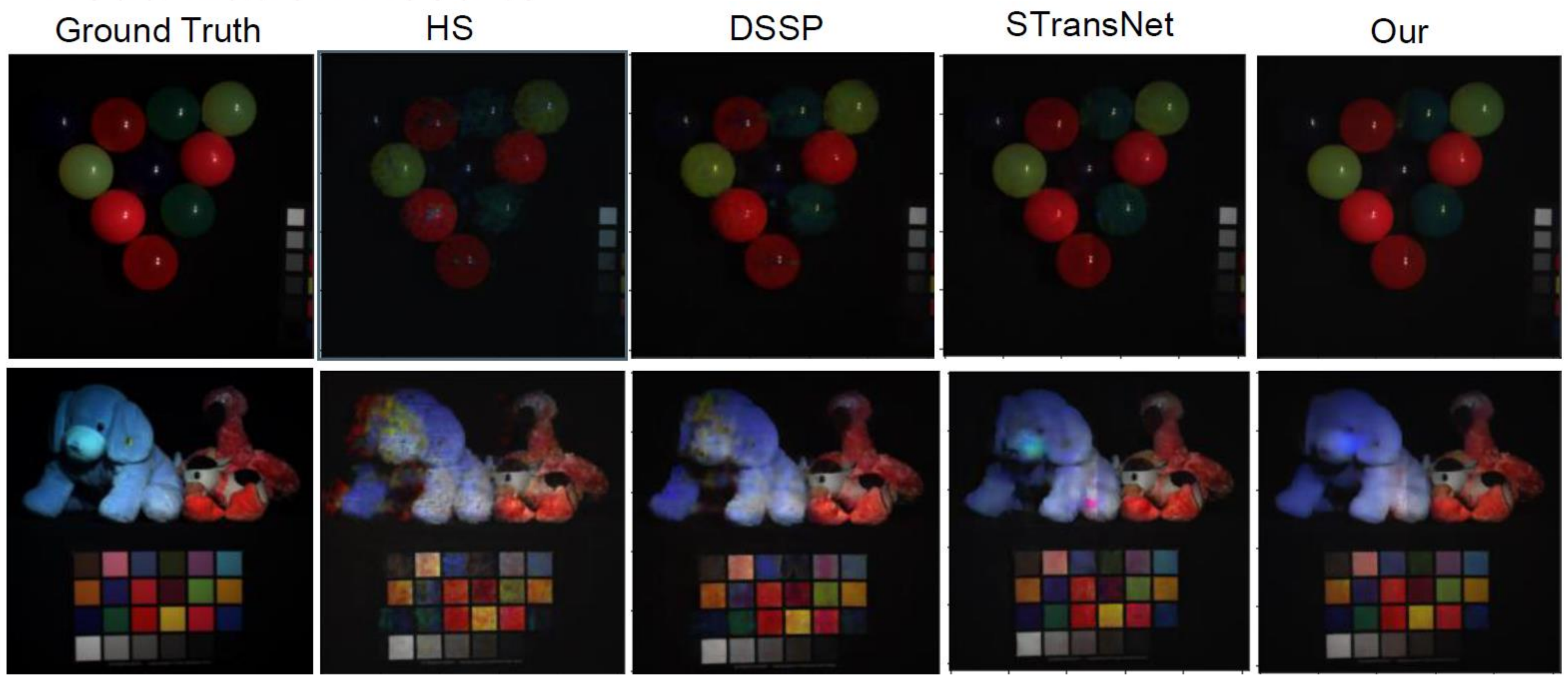
Harvard (50 HS Images) 1042x1392x31
40 training images; 10 test images

Sensing masks: randomly generated binary matrix in a Bernoulli distribution with p = 0.5

Comparison with SoTA methods

Dataset	Metrics	HS [19]	HRNet [20]	DSSP [21]	HMNet [22]	StransNet	Our
CAVE	PSNR↑	25.93	25.82	28.15	27.88	28.97	29.94
	SSIM↑	0.790	0.829	0.851	0.864	0.879	0.903
	SAM↓	0.260	0.305	0.201	0.199	0.188	0.163
Harvard	PSNR↑	34.93	36.04	36.77	36.81	37.05	39.31
	SSIM↑	0.916	0.938	0.934	0.947	0.938	0.955
	SAM↓	0.120	0.166	0.099	0.119	0.100	0.091

Visualization Results



Ablation Study: Verify the effectiveness of different proposed components

Training data generation: using random masks (RM) or a fixed mask (FM)

Test snapshot measurements: generated using random masks (RM) or the fixed mask (FM)

		Test with FM			Test with RM			
Training with FM		✓	✓		✓		✓	
Training with RM				✓		✓		✓
MSMM			✓	✓			✓	✓
CAVE	PSNR↑	28.97	29.94	29.94	25.79	28.54	29.33	29.94
	SSIM↑	0.879	0.896	0.903	0.816	0.879	0.883	0.903
	SAM↓	0.188	0.159	0.163	0.238	0.202	0.170	0.163
Harvard	PSNR↑	37.05	39.27	39.28	30.32	37.77	38.94	39.31
	SSIM↑	0.938	0.955	0.955	0.836	0.947	0.953	0.955
	SAM↓	0.100	0.084	0.091	0.174	0.094	0.087	0.091