Generative AI-assisted computational optical imaging

Wenbo Wan（万文博）
School of Information Engineering, Nanchang University

Sept 20-22, 2024 | Xiamen

Menu

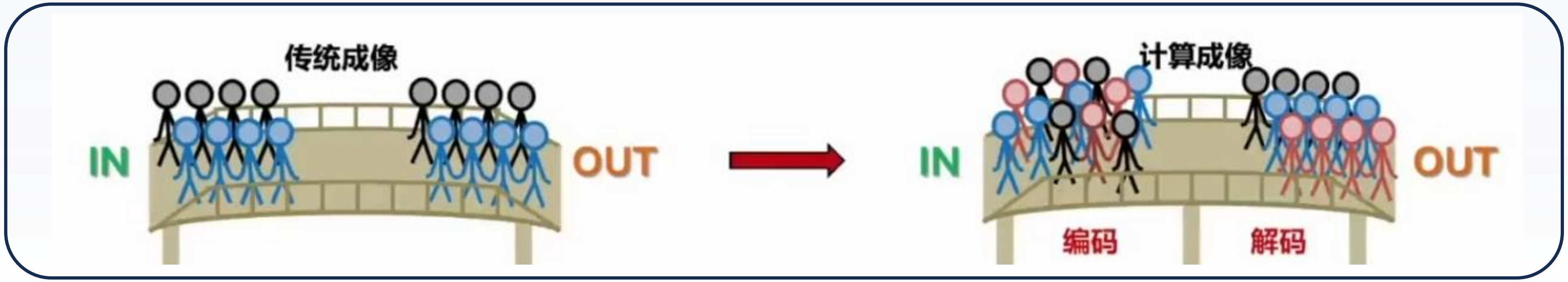
Part I

Computational Optical Imaging and Generative AI

Computational optical imaging

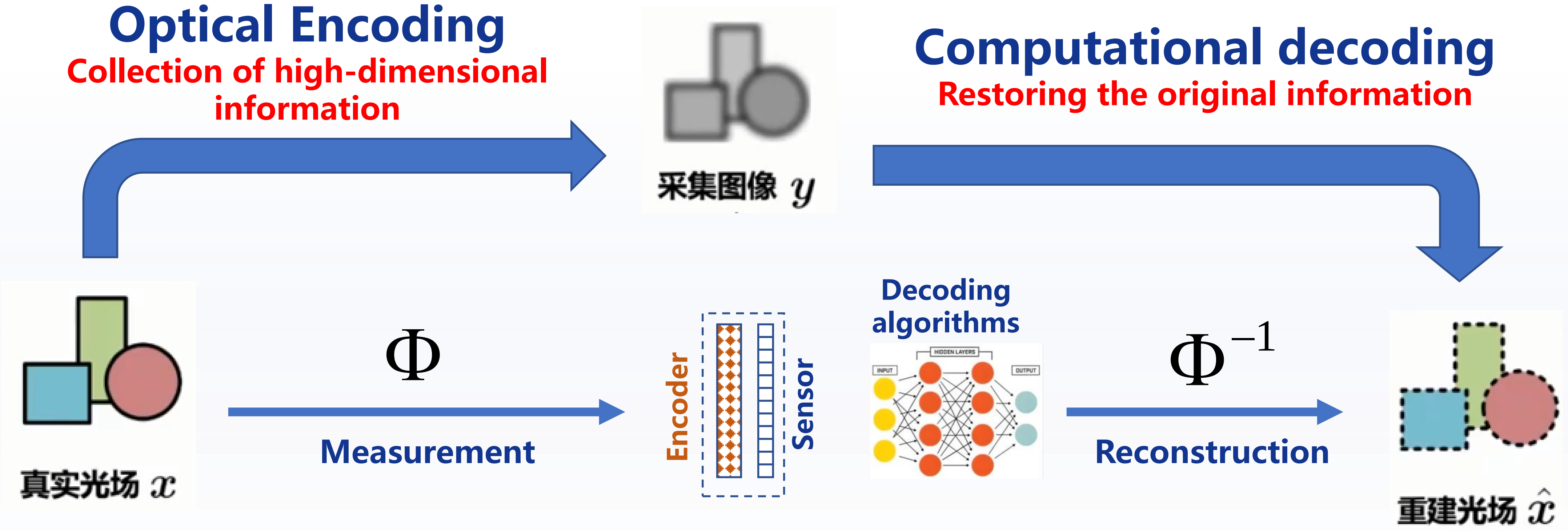
- The optical design based on geometrical optics is simple and easy to use
- The linear approximation method has reached the theoretical limit
- The development of digital signal processing and computer technology provides the possibility of introducing information processing into the imaging process

Information-driven computational imaging techniques



Computational optical imaging
- Optical encoding and computational decoding

- ➢ The photoelectric detection chip constitutes the information bottleneck of computational imaging



Computational optical imaging
- Forward and inverse problems

- The forward model based on physical processes is used to establish the inverse problem of information reconstruction

Forward problem

$$\Phi(x) = y$$

Physical model

Measured data

Inverse problem

$$find\ \ x,\ \ s.t.\ \ \Phi(x) = y$$

- An optimization problem that is easy to solve

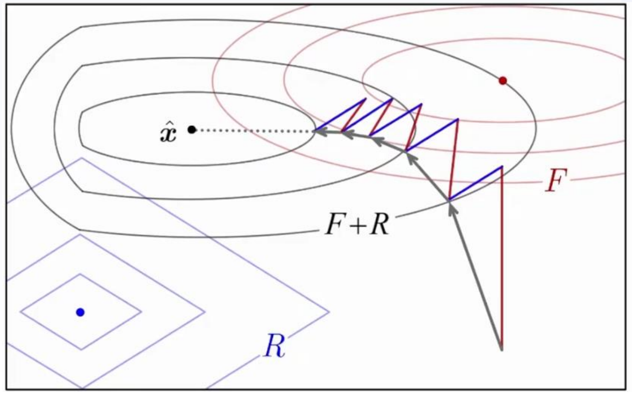
Inverse problem

$$find\ \ x,\ \ s.t.\ \ \Phi(x) = y$$

Optimization problem

$$\arg\min_{x} \frac{1}{2}\left\|\Phi(x) - y\right\|_2^2 + \lambda R(x)$$

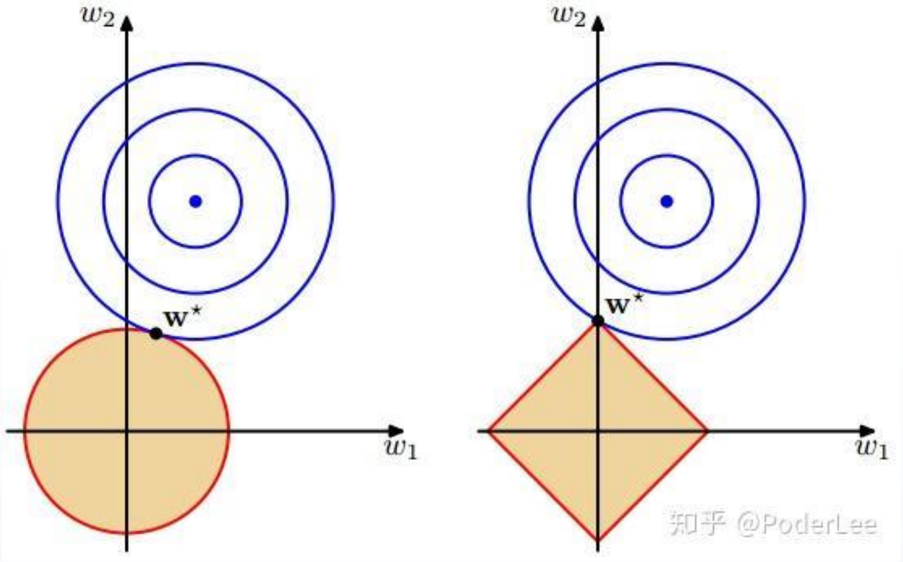
Fidelity **Regularization**



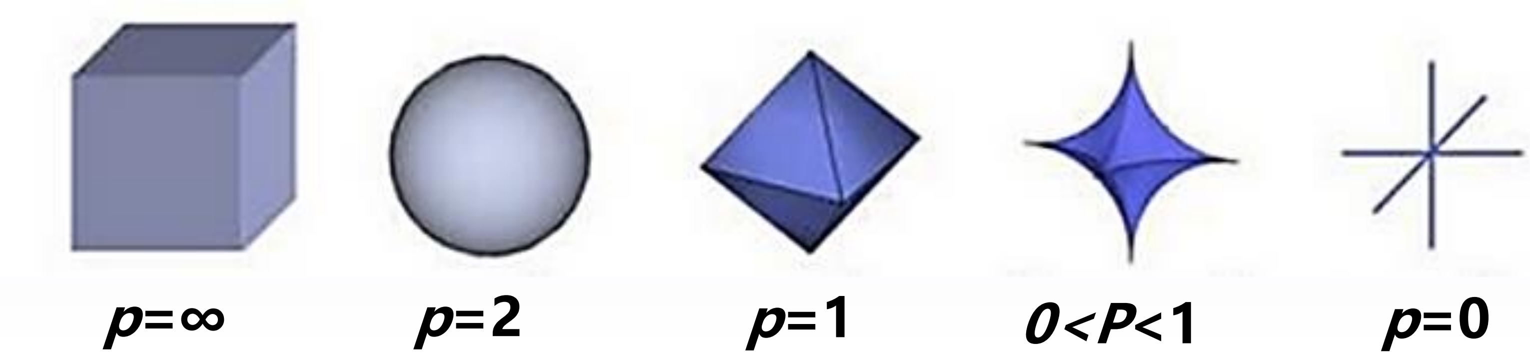
Information representation and regularization constraints

➢ Sparse prior information and regularization constraints

$$\|\mathbf{x}\|_p = \left[\sum_{i=1}^{n} |x_i|^p \right]^{1/p}$$



- L_p norm： The length of a vector
- Prevent overfitting and get sparse solutions



Generative AI and Diffusion Models


年度"十大新词语"发布!

新华社 2023-12-17 14:35 发表于北京

"汉语盘点2023"活动发布年度"十大新词语"

作为年度"汉语盘点"活动重要组成部分，12月16日，国家语言资源监测与研究中心发布2023年度"十大新词语"，它们分别是：**生成式人工智能、全球文明倡议、村超、新质生产力、全国生态日、消费提振年、特种兵式旅游、显眼包、百模大战、墨子巡天。**

要组成部分，12月16日，国家语言资源
们分别是：**生成式人工智能、全球文明**
振年、特种兵式旅游、显眼包、百模大

专家介绍，……技热点，它不仅影响着……的工具和视角；"全球……大国责任和担当；"村……热门话题；"新质生产……动是其中关键；"全国……育活动，共绘美丽中国画卷；2023年确立为"消费提振年"，出台诸多引导消费的利好政策；年轻人

Sora爆火！人工智能将如何影响世界？

央视新闻 2024-02-20 12:06 北京

⇧点蓝色字关注"**央视新闻**"

清澈灵动的眼眸、活泼可爱的萌宠、神秘莫测的海底世界、熙熙攘攘的夏日街区、充满科技感的魔幻都市……

这段场景逼真、色彩丰富、氛围浓厚的短视频，全部由人工智能系统制作生成。

Generative AI and Diffusion Models

图像生成王者不是GAN？扩散模型最近有点火，效果直达SOTA

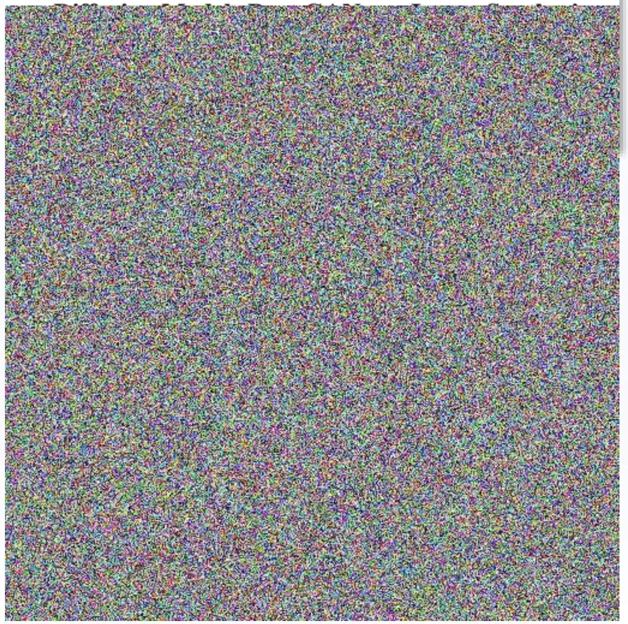
原创 2021-12-29 14:08 · 量子位

博雯 发自 凹非寺
量子位 报道 | 公众号 QbitAI

OpenAI刚刚推出的年末新作GLIDE，又让**扩散模型**小火了一把。

这个基于扩散模型的文本图像生成大模型参数规模更小，但生成的图像质量却更高。

于是，依旧是OpenAI出品，论文标题就直接号称“在图像生成上打败GAN”的**ADM-G模型**也重新进入了大众眼中：



戴上口罩都能还原，用去噪扩散概率模型修复图像，效果「真」极了

2022-01-26 16:04 · 机器之心Pro

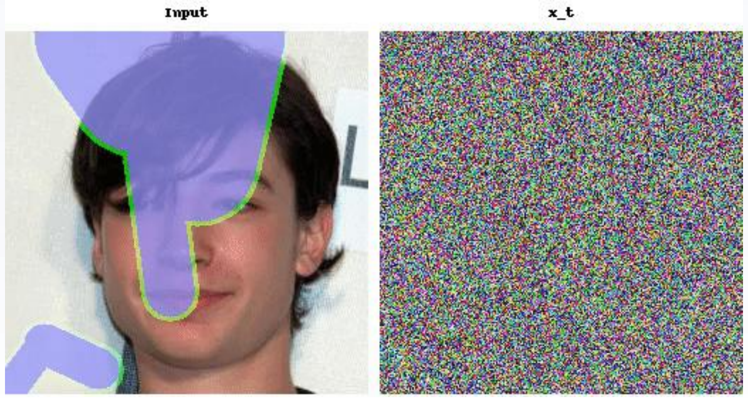
机器之心报道

编辑：杜伟、陈萍

> 无论掩码类型如何多变，苏黎世联邦理工学院计算机视觉实验室（CVL）的图像修复方法都能还原出逼真的图像。

图像修复旨在填充图像中的缺失区域，被修复区域需要与图像的其余部分协调一致，并且在语义上是合理的。为此，图像修复方法需要强大的生成能力，目前的修复方法依赖于 GAN 或自回归建模。

近日，来自苏黎世联邦理工学院计算机视觉实验室（CVL）的研究者提出了 RePaint，这是一种基于 DDPM（Denoising Diffusion Probabilistic Model，去噪扩散概率模型）的修复方法，该方法还可以适用于极端情况下的蒙版。



一句话生成3D模型：AI扩散模型的突破，让建模师慌了

机器之心 2022-11-23 12:33 北京

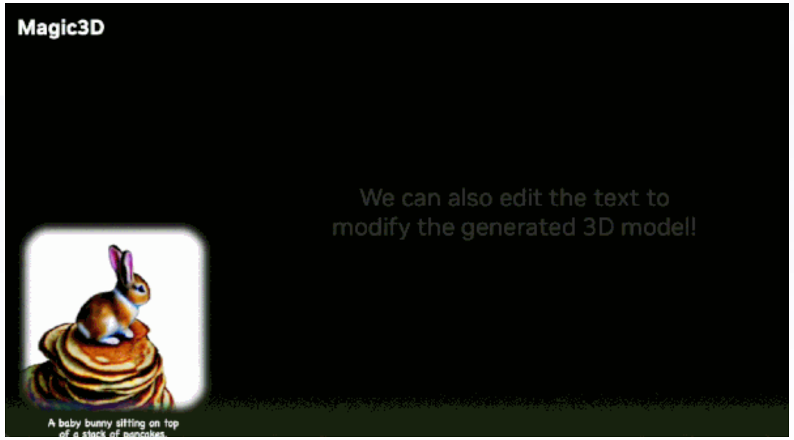
机器之心报道

编辑：泽南、小舟

> 英伟达进入 AI 生成模型领域的研究，直接比别人多一个次元：一句描述生成 3D 模型。

我们生活在三维的世界里，尽管目前大多数应用程序是 2D 的，但人们一直对 3D 数字内容有很高的需求，包括游戏、娱乐、建筑和机器人模拟等应用。

然而，创建专业的 3D 内容需要很高的艺术与审美素养和大量 3D 建模专业知识。人工完成这项工作需要花费大量时间和精力来培养这些技能。



Generative representations from big data

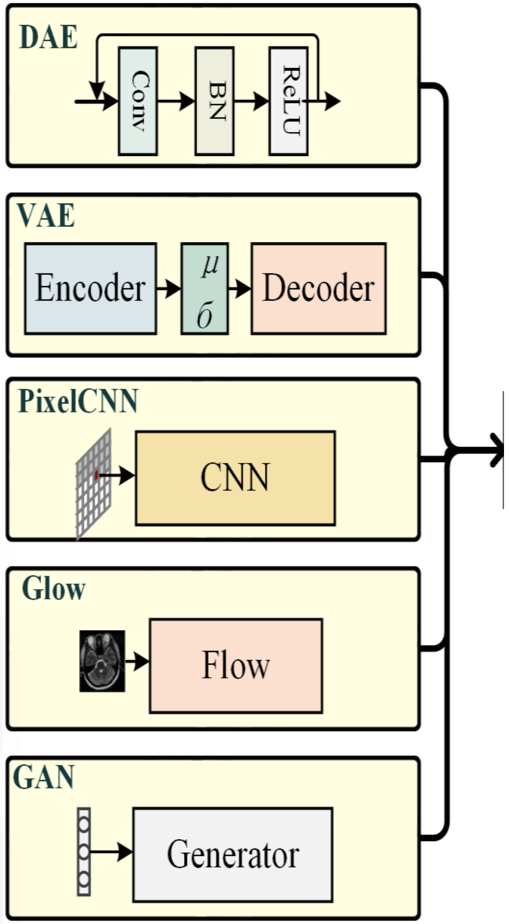
01 Denoising autoencoding(DAE)

02 Variational Autoencoders (VAE)

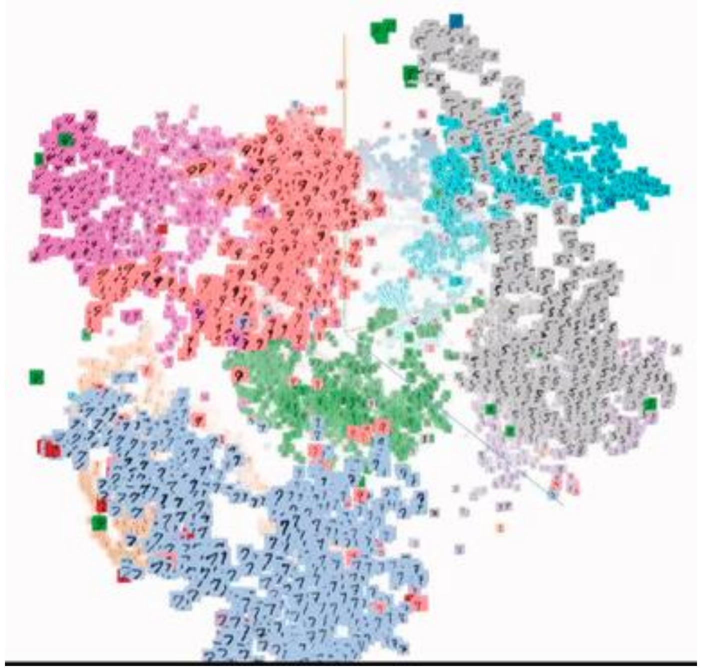
03 Generative Adversarial Network (GAN)

04 PixelCNN

05 Generative Flow (Glow)



Data distribution $\log p(x)$



Z. He, Q. Liu, et al., A comparative study of unsupervised deep learning methods in MRI reconstruction, *iMRI*, 2020.

Generative representations from big data

Generative models -----The training set is constructed with several sample sampled from a given distribution $P_{data}(x)$

The model learns a probability distribution that simulates this distribution $P(x)$

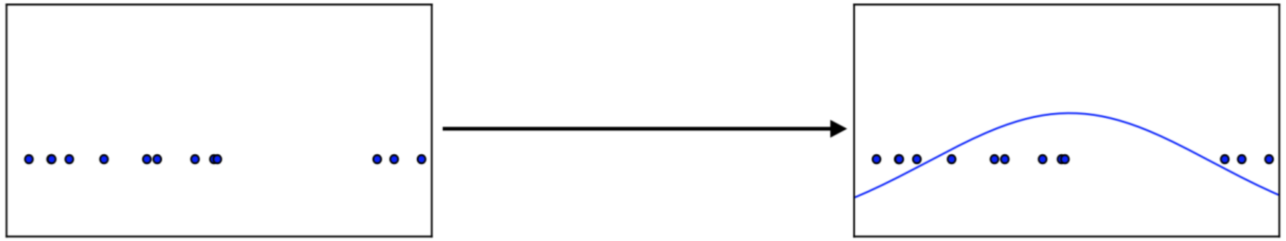
01 **Denoising autoencoding(DAE)**

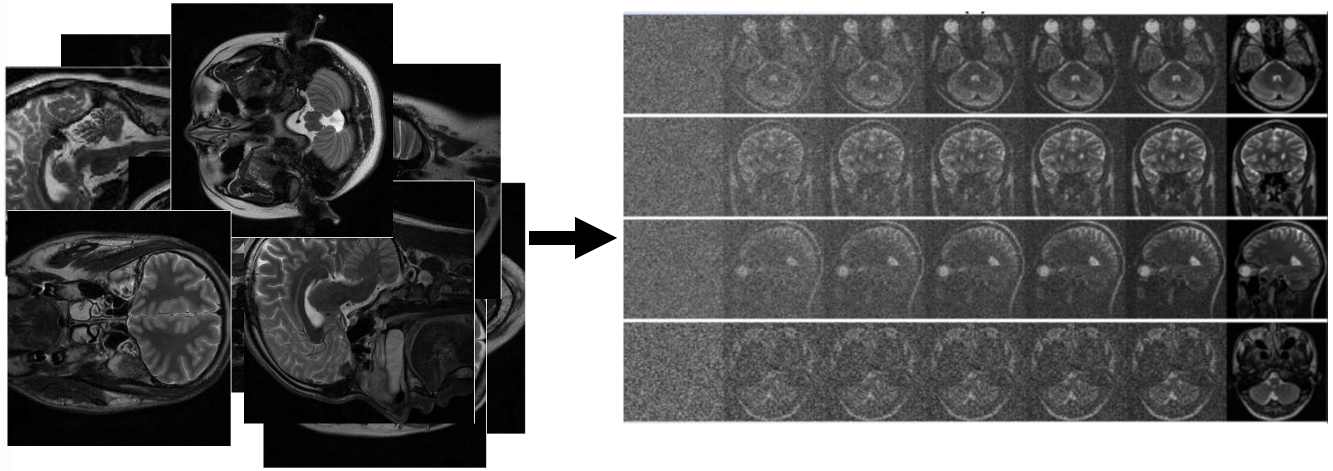
02 **Variational Autoencoders (VAE)**

03 **Generative Adversarial Network (GAN)**

04 **PixelCNN**

05 **Generative Flow (Glow)**





Z. He, Q. Liu, et al., A comparative study of unsupervised deep learning methods in MRI reconstruction, *iMRI*, 2020.

Generative representations from big data

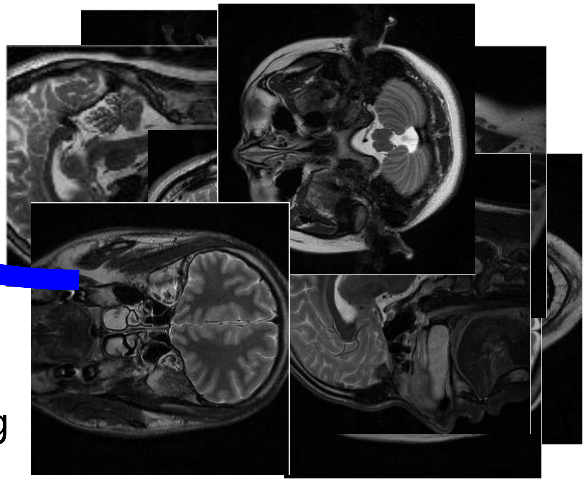
Prior knowledge across modality?

With variable z, not with x itself

Learning prior density in different modality

Prior learned from different modalities for CT recon

CT dataset	$p = 0.8$	47.47/0.9909
	$p = 1$	47.44/0.9908
	$p = 1.5$	47.18/0.9902
	$p = 2$	46.69/0.9888
DIV2K dataset	$p = 0.8$	47.54/0.9908
	$p = 1$	47.52/0.9907
	$p = 1.5$	47.24/0.9901
	$p = 2$	46.71/0.9887
MRI dataset	$p = 0.8$	46.69/0.9890
	$p = 1$	46.65/0.9889
	$p = 1.5$	46.41/0.9882
	$p = 2$	45.82/0.9863



F. Zhang, M. Zhang, B. Qin, Y. Zhang, Z. Xu, D. Liang, Q. Liu, REDAEP: Robust and enhanced denoising autoencoding prior for sparse-view CT reconstruction, *IEEE Trans. Radiat. Plasma Med. Sci.*, 2020.

Unsupervised learning：From DAE to DSM

Q. Liu, Q. Yang, H. Cheng, S. Wang, M. Zhang, D. Liang, Highly undersampled magnetic resonance imaging reconstruction using autoencoding priors, ***Magn. Reson. Med.***, vol. 83, no. 1, pp. 322-336, 2020.

DAE

S. Wang, J. Lv, Z. He, D. Liang, Y. Chen, M. Zhang, Q. Liu, Denoising auto-encoding priors in undecimated wavelet domain for MR image reconstruction, ***Neurocomputing***, vol.437, pp.325-338, 2021.

DAEP

DMSP

DSM

M. Zhang, M. Li, J. Zhou, Y, Zhu, S. Wang, D. Liang, Y. Chen, Q. Liu. High-dimensional embedding network derived prior for compressive sensing MRI reconstruction, ***Med. Image Anal.***, vol. 64, 101717, 2020.

DMSP

C. Quan, J. Zhou, Y. Zhu, Y. Chen, S. Wang, D. Liang, Q. Liu, Homotopic gradients of generative density priors for MR image reconstruction, ***IEEE Trans. Med. Imag.***, 2021.

DSM

Generative model

Part II

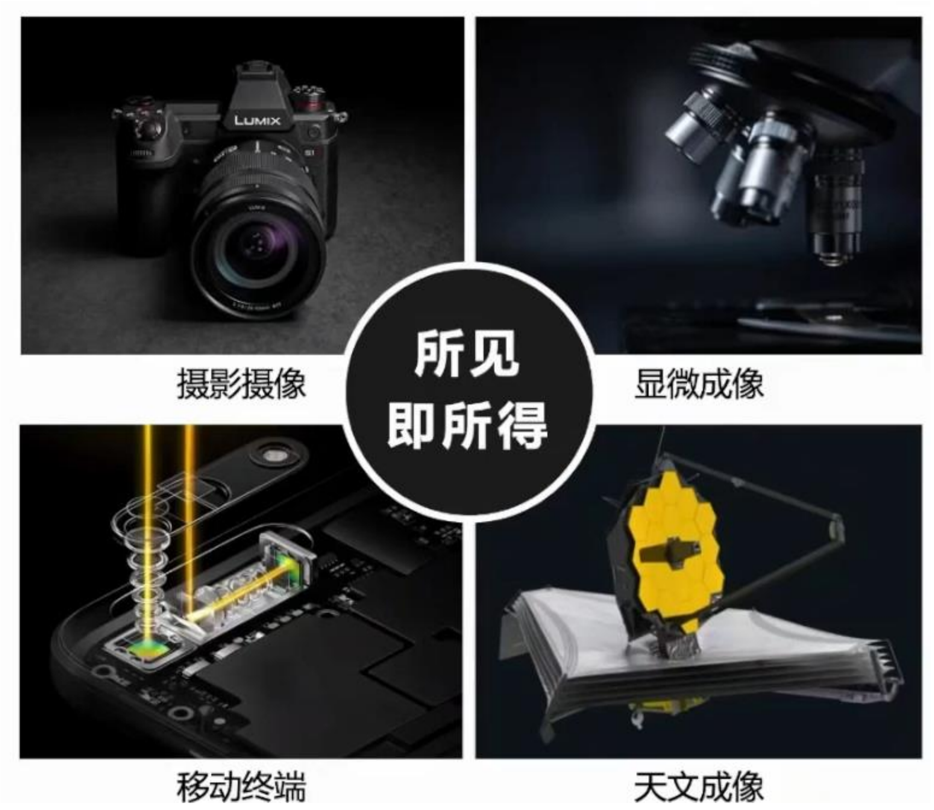
Generative AI assisted multi-phase FZA lensless imaging

The bottleneck of traditional lens-based imaging systems

◆ lens-based imaging

✓ "Point-to-point" conjugate projection

✓ what you see is what you get



◆ Bottleneck

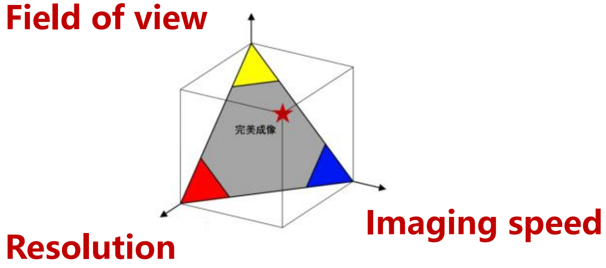
➢ Performance gains come at the cost of increased size, weight, cost, and complexity

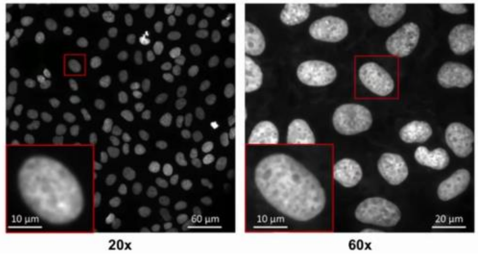
Leica Apo-Telyt-R

1:5.6 1600mm

60Kg、$ 2.06 million

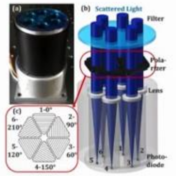
➢ FoV, resolution, and speed are mutually restrictive





➢ High-dimensional information puts forward high requirements for sampling bandwidth and time

多光谱阵列成像



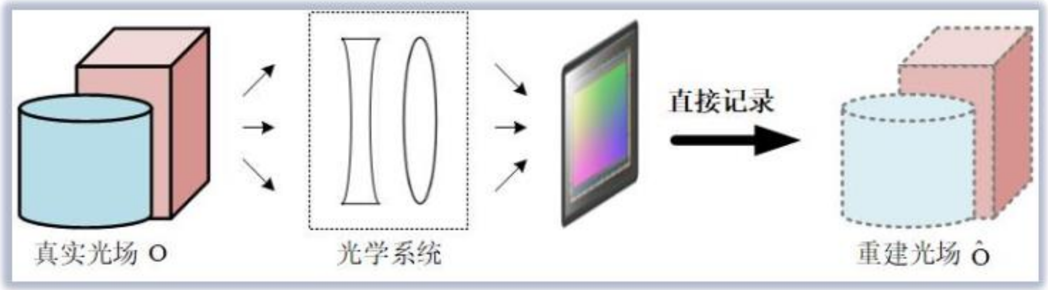
偏振成像

结构光三维成像

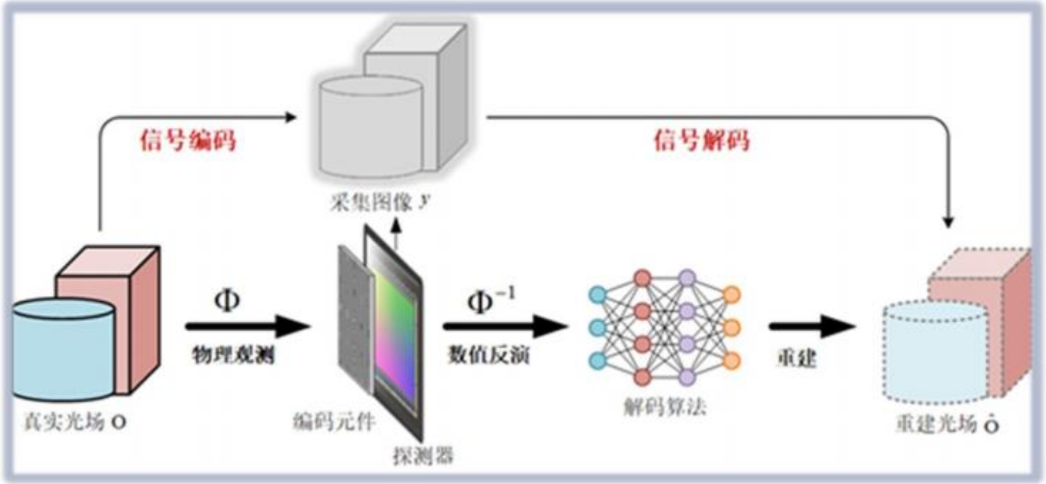
Lensless imaging technology

- No longer only relies on physical optical components
- Combines **front-end planar optical coding devices** and **back-end computational reconstruction methods** to achieve light field reconstruction

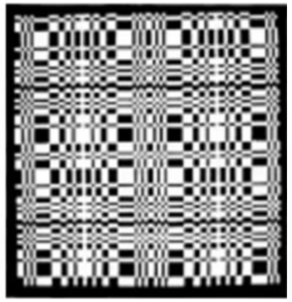
Lens-based imaging



Lensless imaging

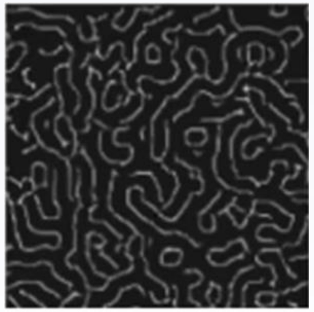


- **Encoding masks**


URA


Spiral Gratings


FZACam

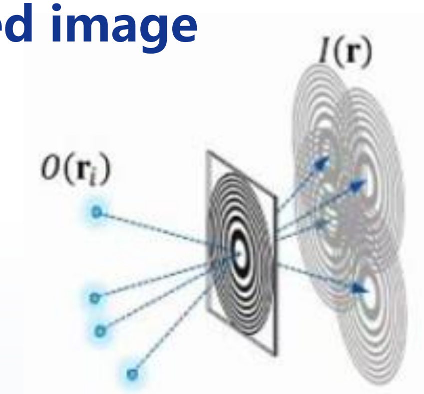

PhlatCam

Fresnel Zone Aperture Coded Lensless Imaging

- **Incoherent encoding**

Each point source projects the FZA onto the sensor to form an encoded image

$$I(r)=\frac{1}{2}\sum_{k}^{N}I_k\left[1+\cos\left(\frac{\pi}{r_1^2}|r-r_k|^2\right)\right]$$



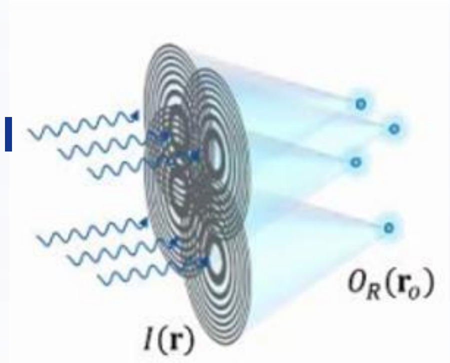
- **Coherent decoding**

$$O_R(r_o)=C\iint I(r)\exp\left\{\frac{i\pi}{\lambda d}|r-r_o|^2\right\}dS$$

$$=\frac{ir_1^2}{2}\sum_{k}^{N}I_k+\boxed{\frac{r_1^4}{4}\sum_{k}^{N}I_k\delta(r_o-r_k)}+\boxed{\frac{r_1^2}{8}\sum_{k}^{N}I_k\exp\left(\frac{i\pi}{2r_1^2}|r-r_k|^2\right)}$$

Original image

Twin image



The reconstructed image is disturbed by the twin image

- **How to eliminate the twin image ?**

Inverse Problem:

$$\text{find } O, \text{ s.t. } HO=I$$

Least Squares Method

$$\min_{O}\ \frac{1}{2}\|I-HO\|_2^2$$

Optimization

$$\boxed{\min_{O}}\ \frac{1}{2}\|I-\boxed{H}O\|_2^2+\boxed{\lambda\Psi}$$

Optimization algorithm

Physical modeling

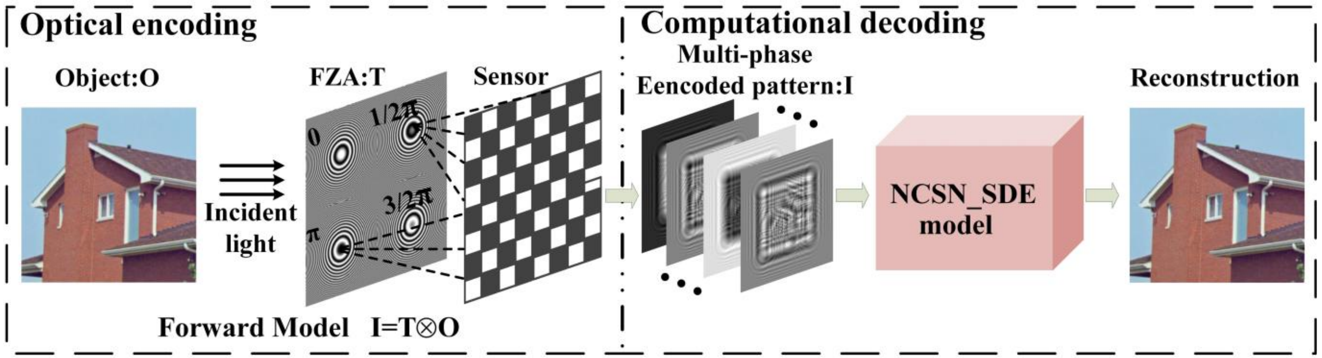
Regularization

Multiphase FZA lensless imaging

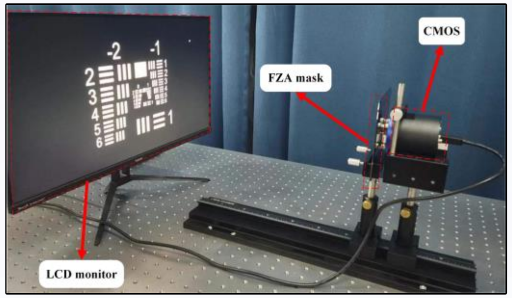
Multiphase FZA encoder

FZA encoder Phase

$$t(x_p, y_p; \phi) = \frac{1}{2}[1 + \cos\{\boxed{\frac{\pi}{r_1^2}(x_p{}^2 + y_p{}^2)} + \boxed{\phi}\}]$$

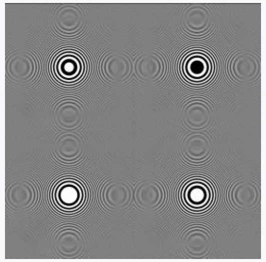


- Obtain multi-encoded information with single-frame snapshot
- Acquisition and utilization of high-dimensional prior information
- Eliminate twin image artifacts and chromatic aberrations



Lensless imaging system

=



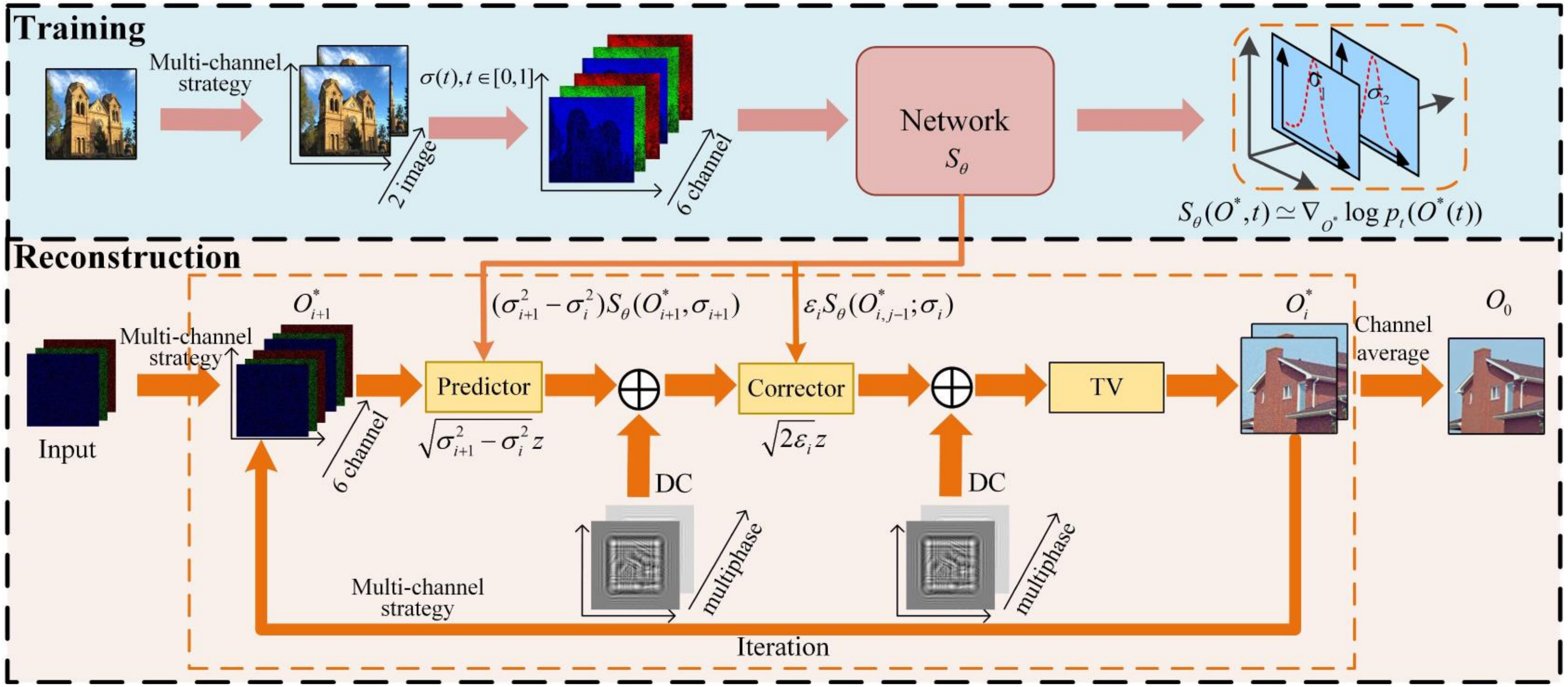
Multiphase FZA encoder

+

CMOS sensor

Wan W. et al., Multi-phase FZA Lensless Imaging via Diffusion Model, Optics Express, 2023, 31(12): 20595-20615

Generative AI-assisted lensless imaging



Prior Learning:

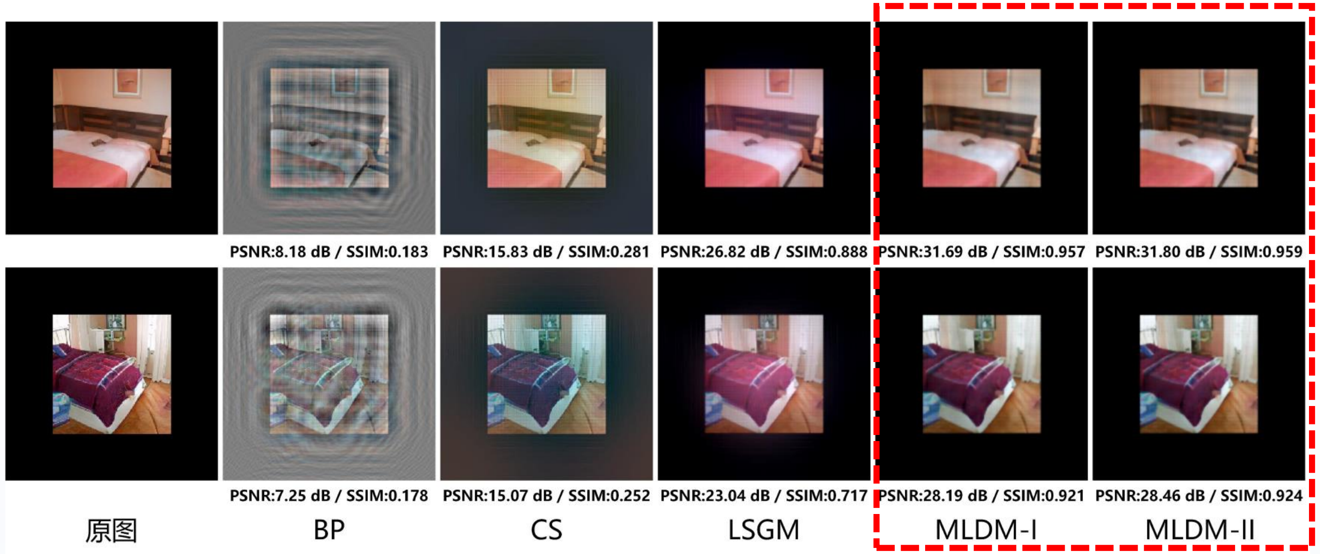
- Learn the representation of prior information in high-dimensional images with multiple color channels

Image Reconstruction:

- Establish the information correlation between the color image channel and the encoded phase channel
- The prior information of data distribution is used to constrain the prediction-correction iterative reconstruction process
- Improve the quality of image reconstruction

Wan W. et al., Multi-phase FZA Lensless Imaging via Diffusion Model, Optics Express, 2023, 31(12): 20595-20615

Simulative validation



Method	BP	CS	LSGM	MLDM-I	MLDM-II
PSNR(dB)	7.40	17.68	24.46	29.42	29.91
SSIM	0.182	0.268	0.711	0.933	0.941

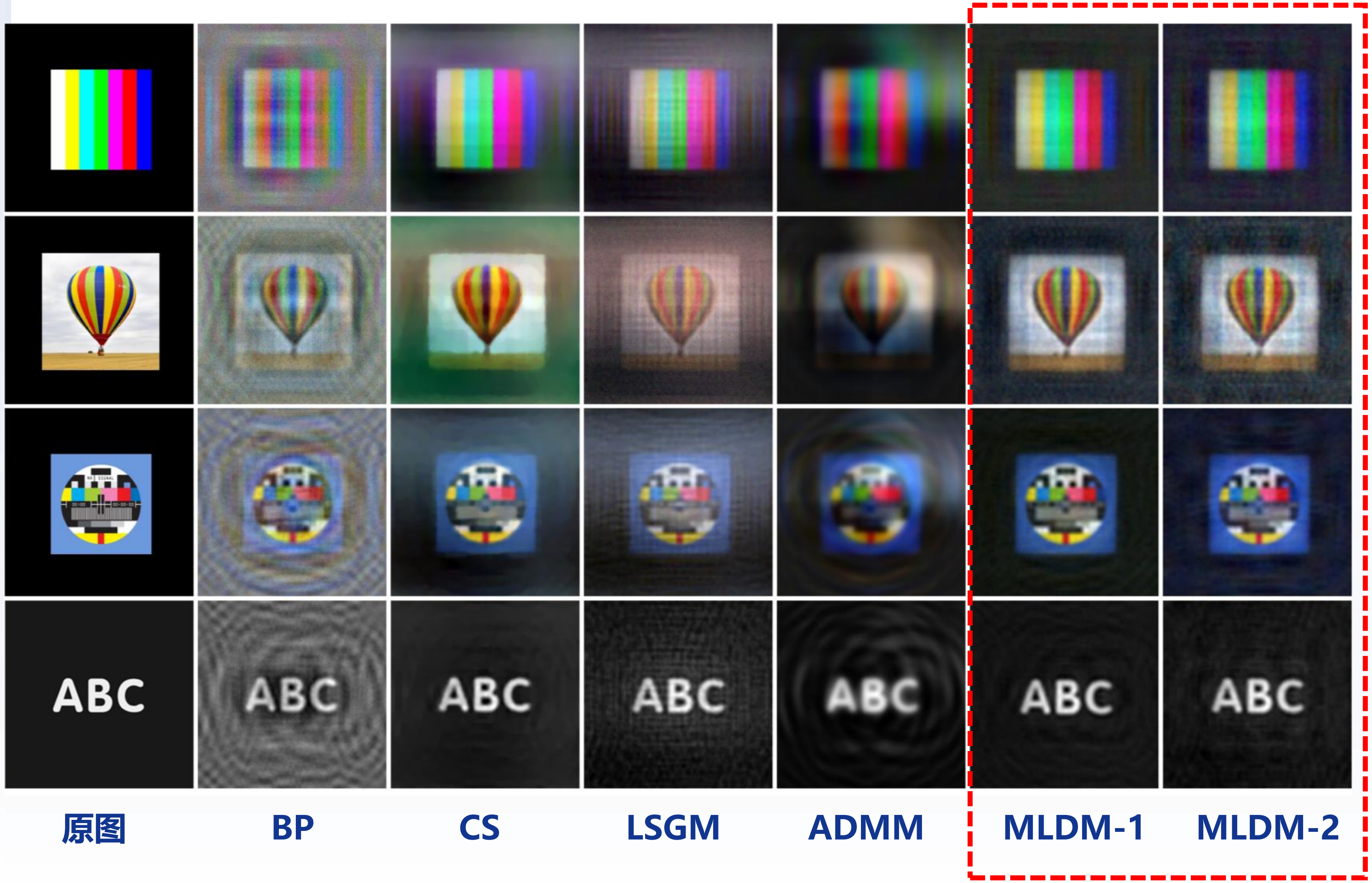
- **BP:** The quality of the reconstruction is the worst, and a significant twin images can be observed
- **CS:** Intermediate color diffusion, discoloration at the edges of the image, and noticeable color distortion
- **LSGM:** The sharpness and color accuracy of the reconstructed image have improved, but there is still a noticeable meshing effect
- **MLDM I/II:** Eliminates the interference of twin image artifacts for better color fidelity

Generalization validation

- Good reconstruction results for cross-dataset targets
- Strong generalization can been demonstrated

Method	BP	CS	LSGM	MLDM-I	MLDM-II
PSNR(dB)	7.43	15.35	24.36	**28.32**	**28.87**
SSIM	0.175	0.251	0.784	**0.925**	**0.933**

Experimental validation



Conclusion:

- Excellent performance in restoring the target image, eliminating the twin image effect, and improving the image clarity
- Images reconstructed by MLDM have high color fidelity
- The MLDM method has higher PSNR and SSIM values

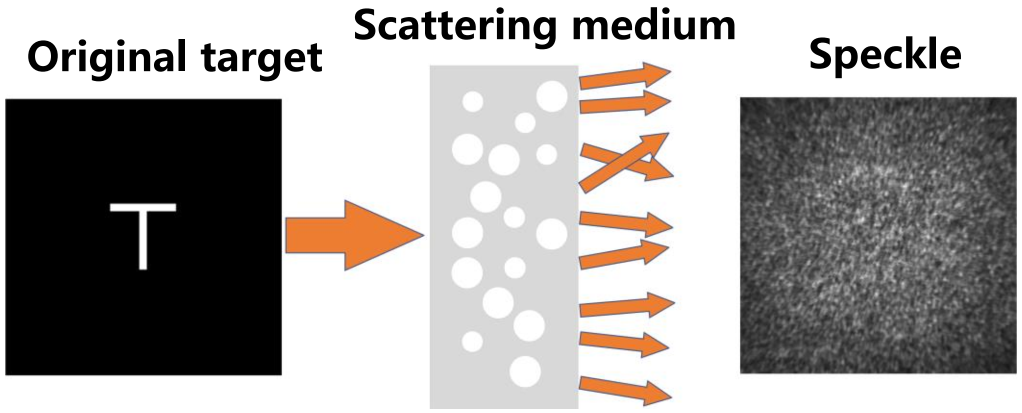
Part III

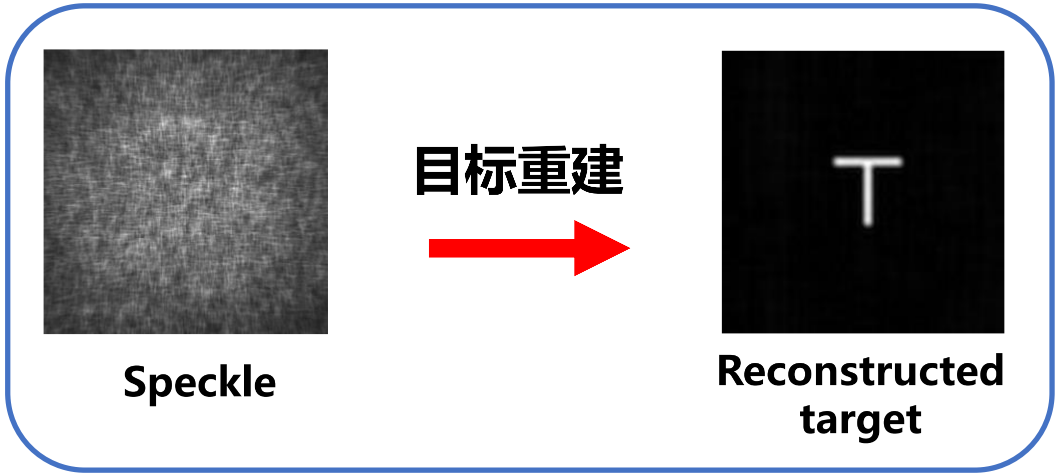
Generative AI assisted imaging through scattering medium

Imaging through scattering medium

Scattered Imaging:
Imaging using scattered light that carries information about the target

- As the degree of **turbidity of the medium increases**
- The image target gradually **changes from clear to unrecognizable**
- The target information is scrambled, and **only optical speckle can be obtained**

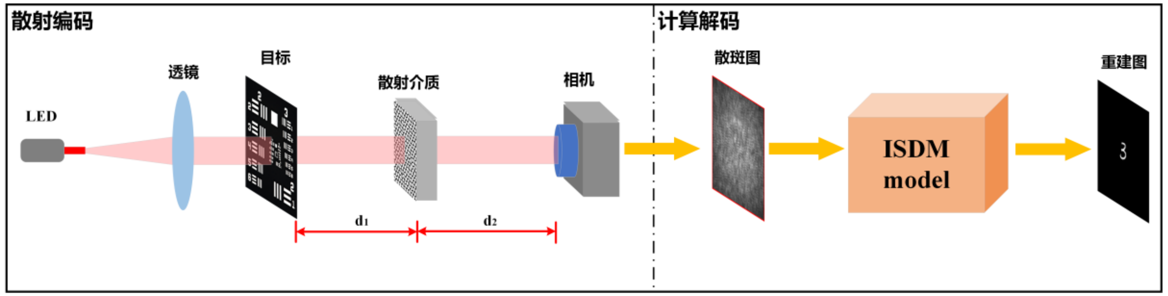




Computational optical imaging offers the possibility of recovering targets hidden behind scattering media

Imaging through scattering medium

◆ Scattering Imaging: Scattering Coding and Computational Decoding



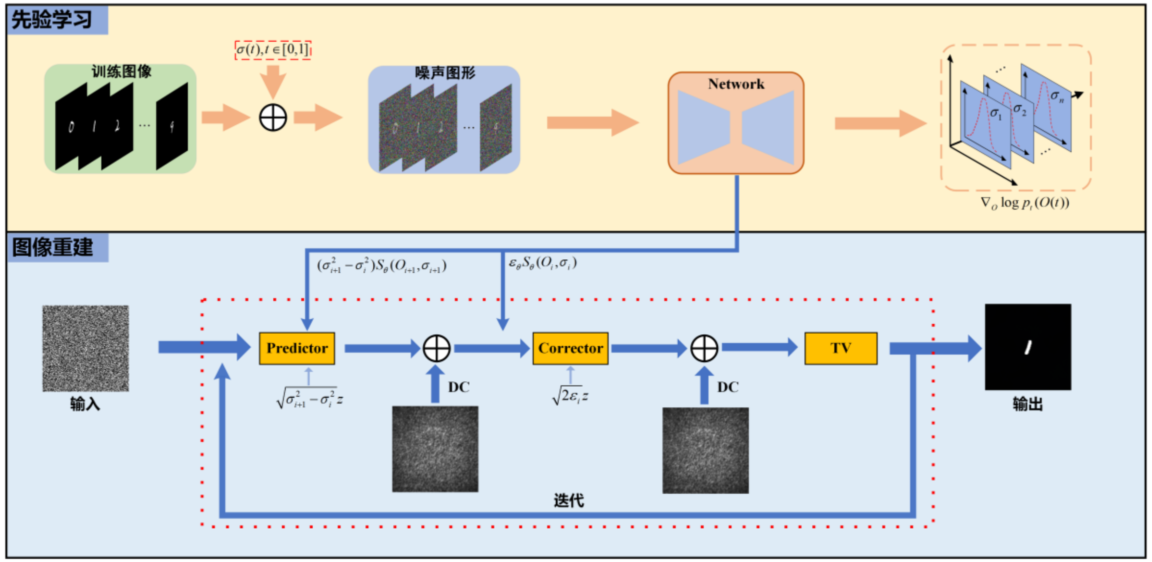
➢ **Optical encoding**

$$I(x,y) = O \otimes P = \iint O(x,y)\cdot P(x,y)dxdy$$

➢ **Computational decoding**

$$O_i = \mathbb{F}^{-1}\left(\frac{H^*\mathbb{F}(I)}{\| H \|^2 + q}\right) + \lambda W_{i,j} \qquad W_{i,j} = O_{i,j-1} + \varepsilon_i S_\theta(O_{i,j-1},\sigma_i) + \sqrt{2\varepsilon_i}\, z$$

◆ ISDM算法流程图



Prior Learning:

- Gradient distribution prior information is learned using a score-based diffusion model

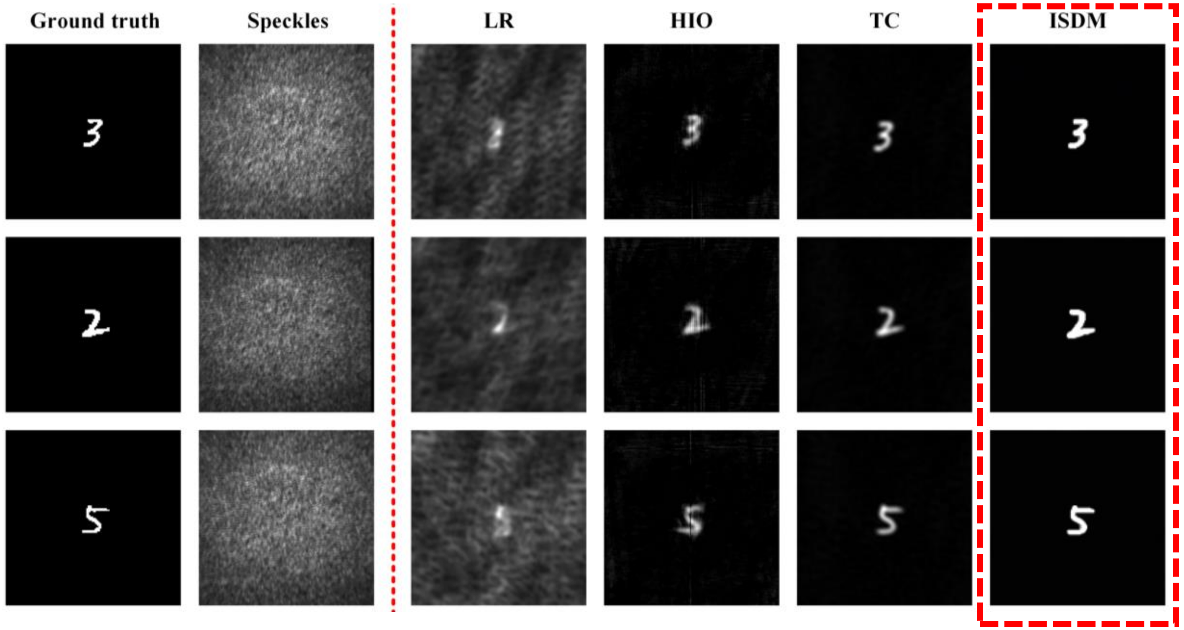
Image Reconstruction:

- Establish the information association between the image target and the encoded image
- The prior information of gradient data is used to constrain the reconstruction process
- Achieve high-quality reconstruction under the premise of satisfying explainability

Chen Z. et al., Imaging through scattering media via generative diffusion model， Applied Physics Letters, 2024, 124: 051101

Simulative validation and generalization verification

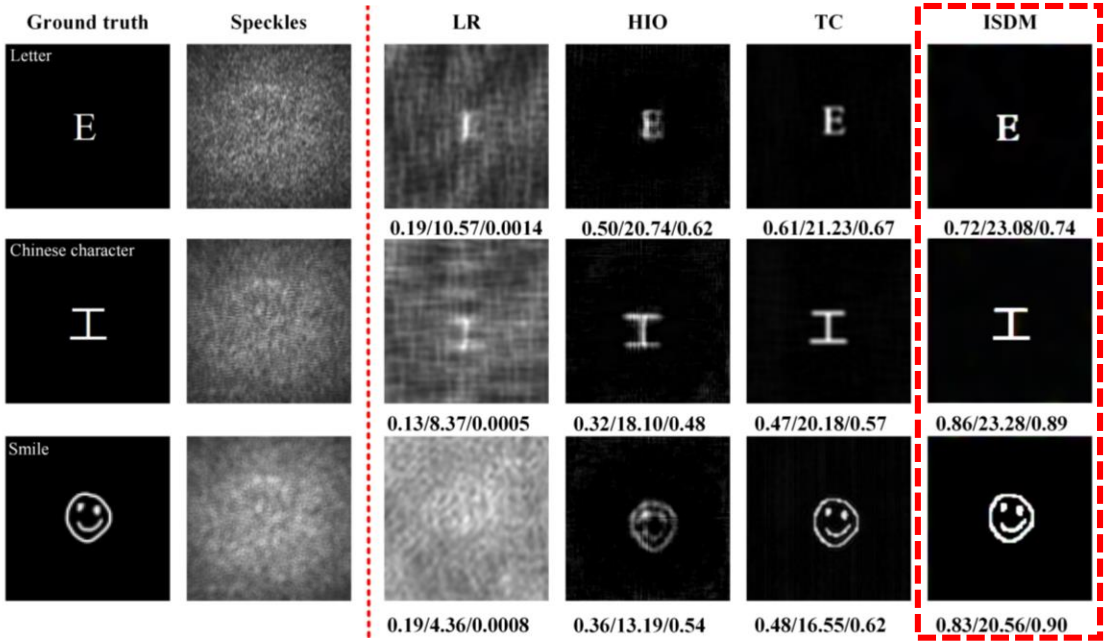
◆ Comparison of reconstruction results of MNIST



Method	PCC	PSNR	SSIM
LR	0.45	14.43	0.0058
HIO	0.66	23.89	0.66
TC	0.73	24.07	0.67
ISDM	**0.87**	**26.11**	**0.74**

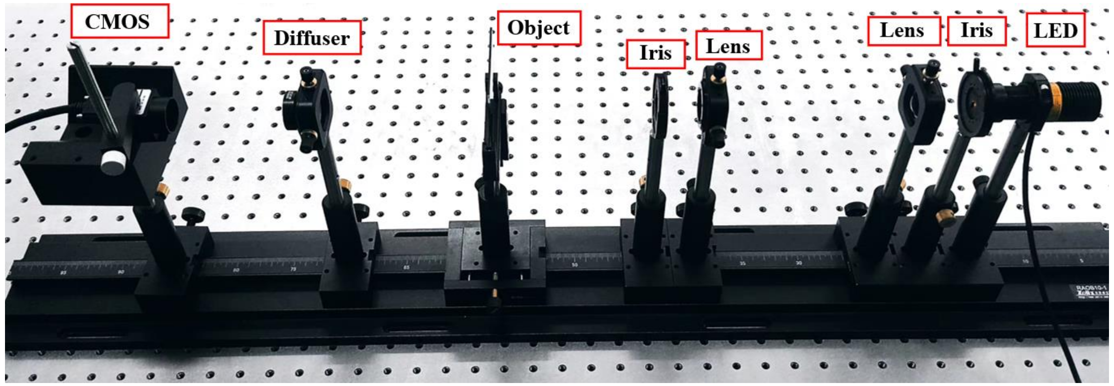
- ISDM removes background noise and improves image clarity
- Fewer artifacts than traditional methods, resulting in higher image quality

◆ Comparison of cross-dataset reconstruction



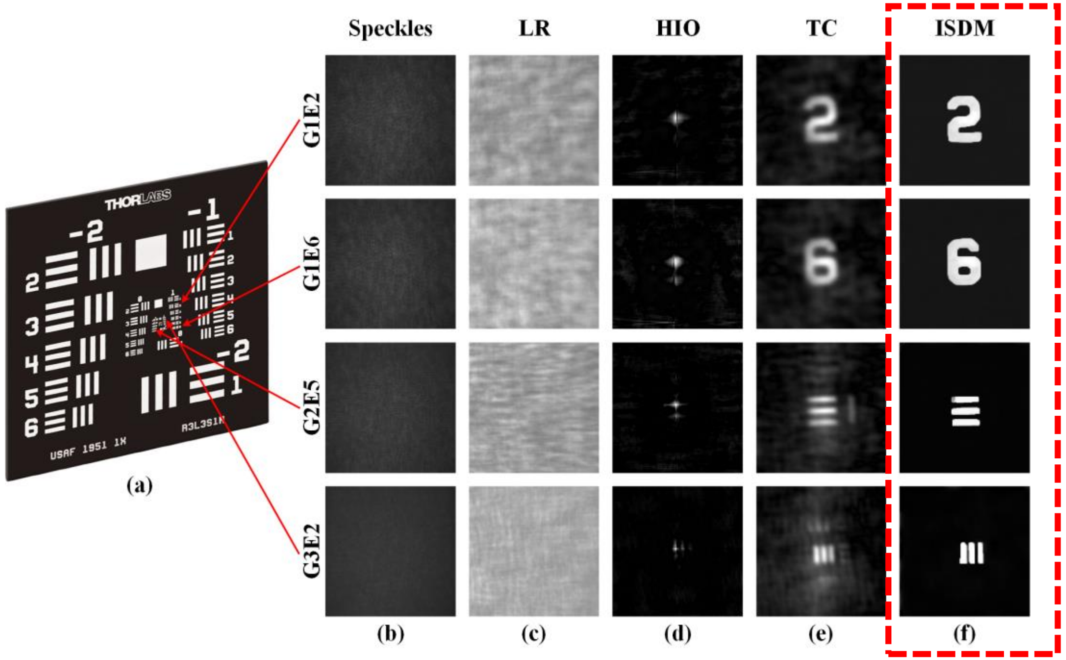
- ISDM can effectively reconstruct different types of target images
- It shows that the ISDM method has strong generalization
- The PCC, PSNR, and SSIM values of the ISDM reconstructed images were higher

Experimental validation



- The beam emitted by the LED (625 nm) is collimated to illuminate the USAF resolution test targets
- The light field is modulated by ground glass
- Captured by sCMOS with a resolution of 1920×1200

◆ Comparison of experimental results



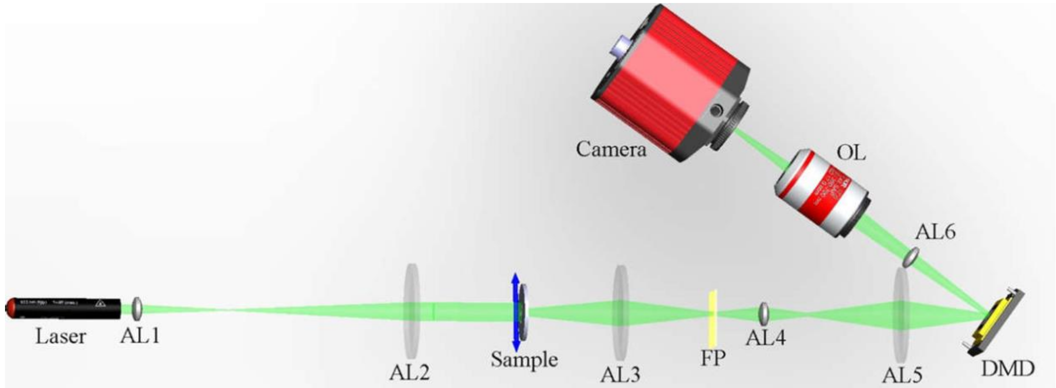
- ISDM can effective in removing artifacts
- Dramatically improve image clarity
- The target image reconstructed by ISDM has a higher fidelity
- High-quality results can be achieved in real-world scenarios
- The spatial resolution reaches 8.98 lp/mm

Part IV

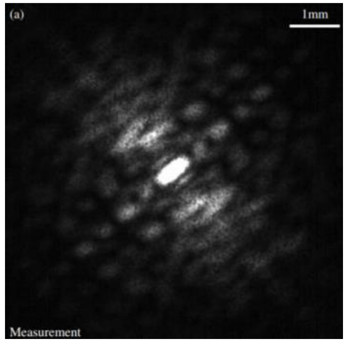
Generative AI assisted temporal compressive coherent diffraction imaging

Snapshot time compressive coherent diffraction imaging

- Snapshot time compressive coherent diffraction imaging (TC-CDI) was proposed by Professor Yuan Xin of Westlake University [1]
- Data is mesured in the frequency domain
- DMD is used to compress multi-frame spectral data into a single image
- Achieve coherent diffraction reconstruction that exceeds the frame rate of the camera

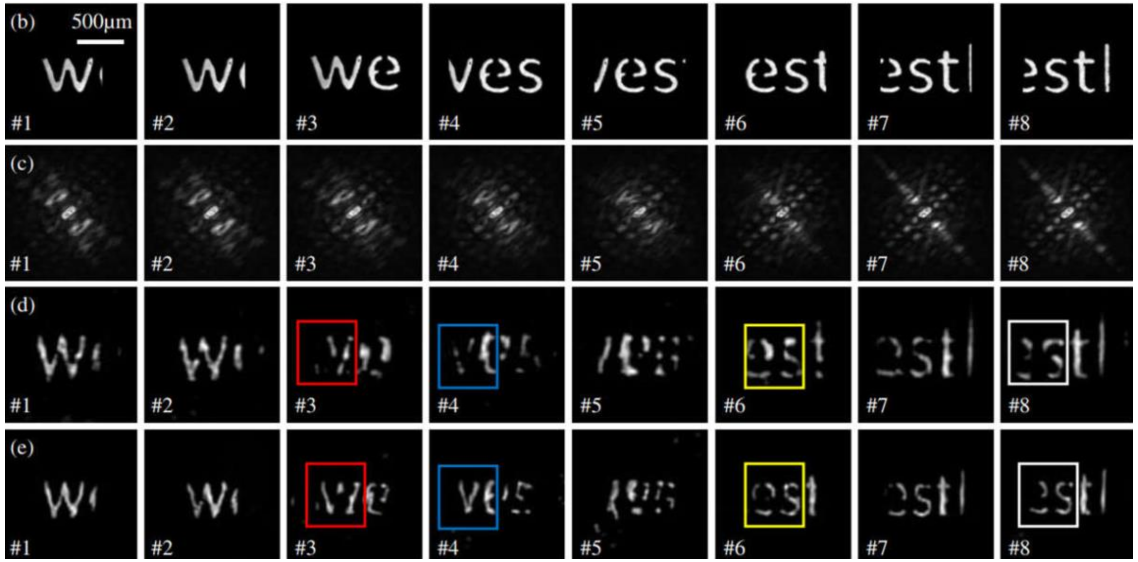


[1] Z. Chen, S. Zheng, Z. Tong, et al., “Physics-driven deep learning enables temporal compressive coherent diffraction imaging,” Optica 9(6), 677–680 (2022).



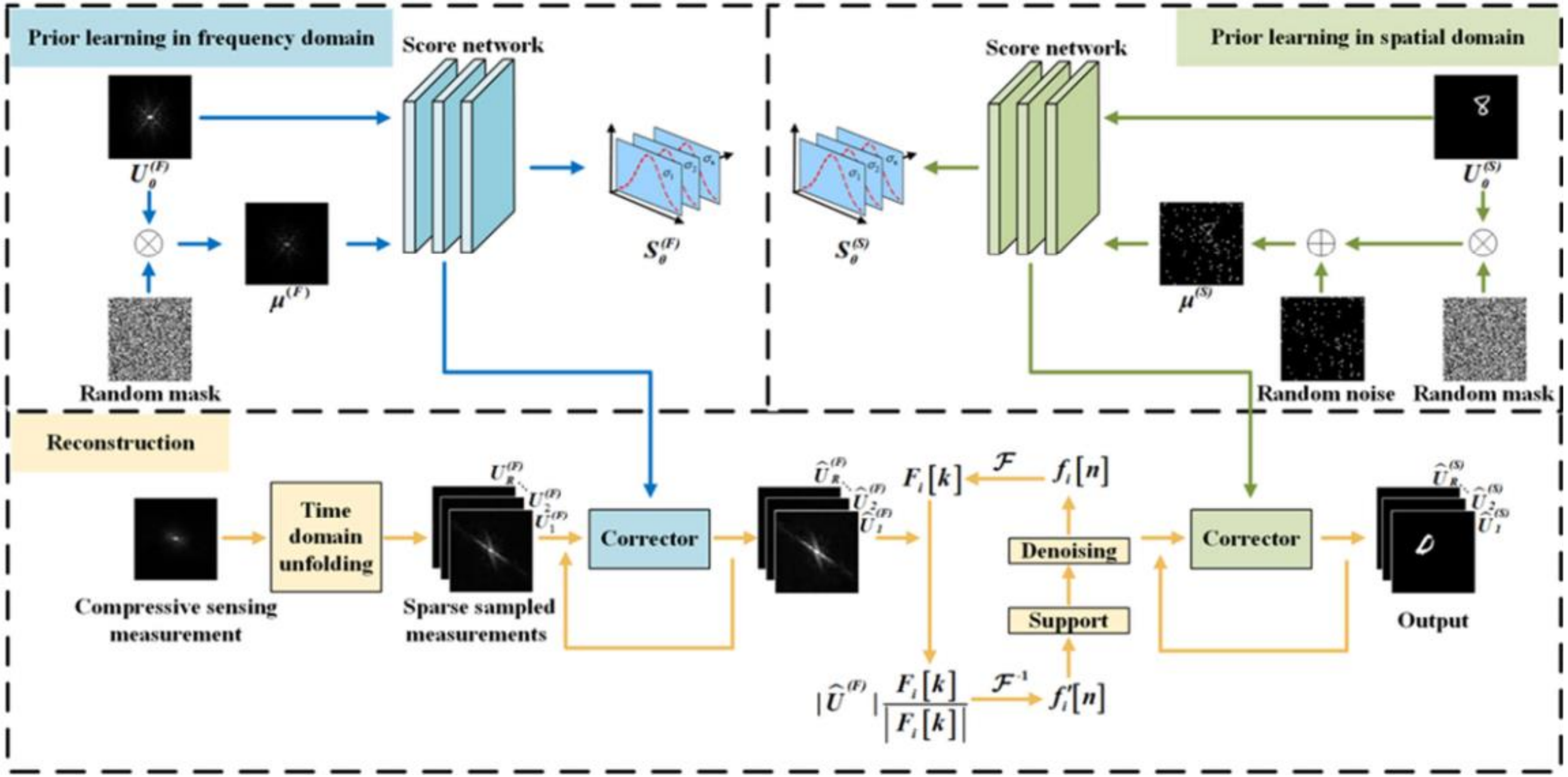
Single-frame spectrum measurement

Restore the multi-frame target



Dual-domain mean-reverting diffusion model-enhanced TCCDI

Priori learning in frequency domain



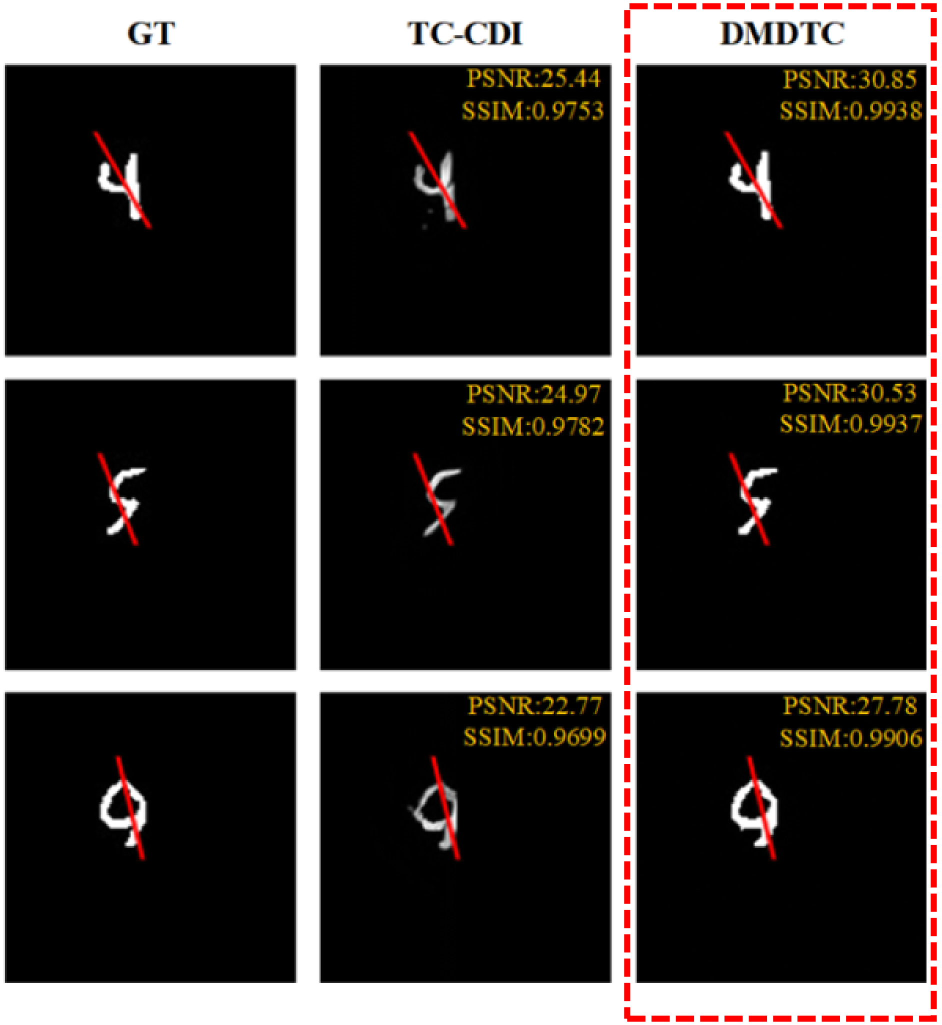
Priori learning in spatial domain

Constraints in both frequency domain and spatial domain

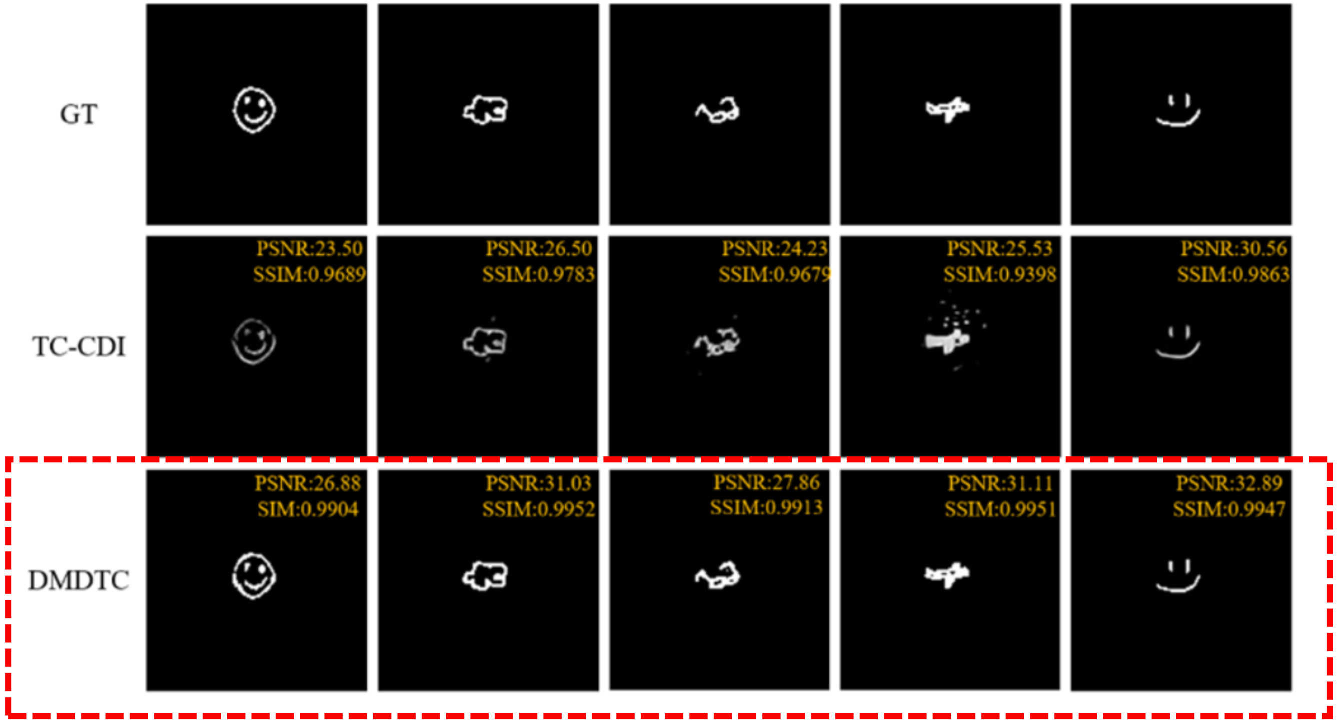
H. Li, et. al., "Dual-domain mean-reverting diffusion model-enhanced temporal compressive coherent diffraction imaging", Optics Express , 2024, 32(9): 15243-15257

Simulative validation and generalization verification

◆ Simulative validation



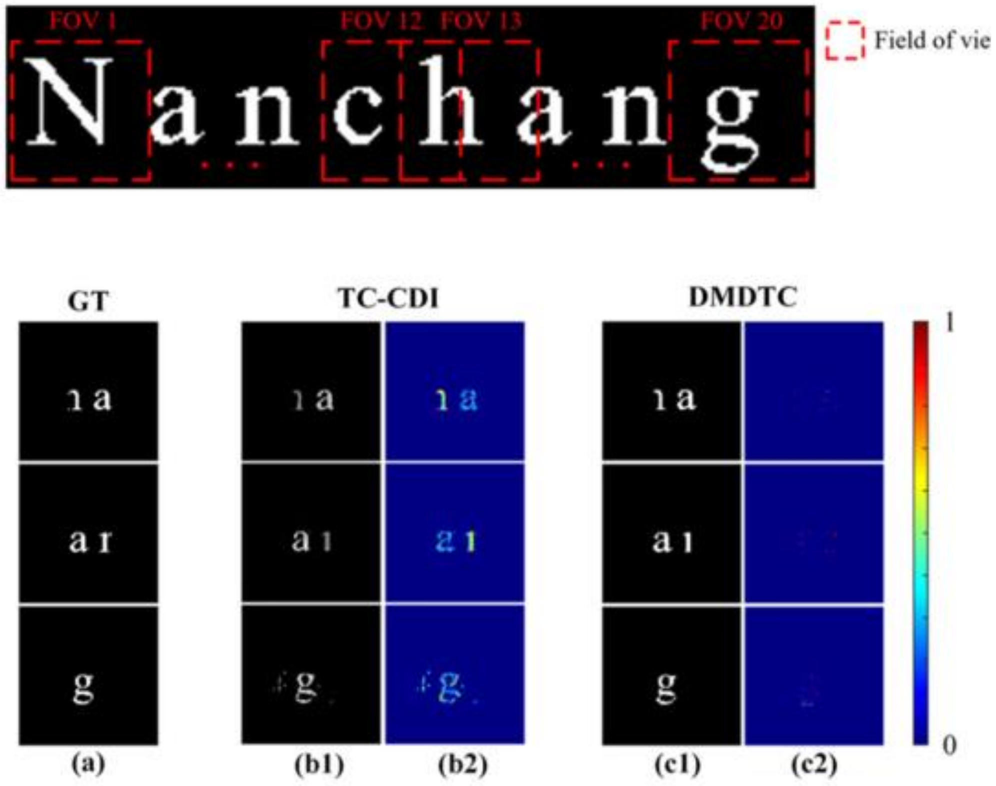
◆ Comparison of cross-dataset reconstruction



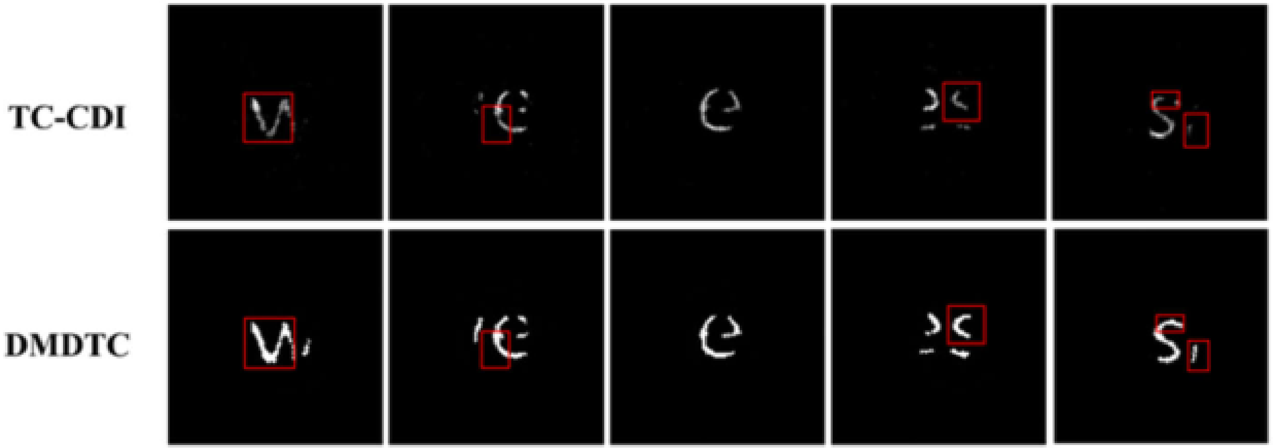
- ➢ The quality of the reconstruction is higher than that of TCCDI
- ➢ High-quality reconstructions are also possible for cross-dataset targets

Dynamic imaging and experimental validation

◆ Simulative validation on dynamic targets



◆ Reconstructions use experimental data



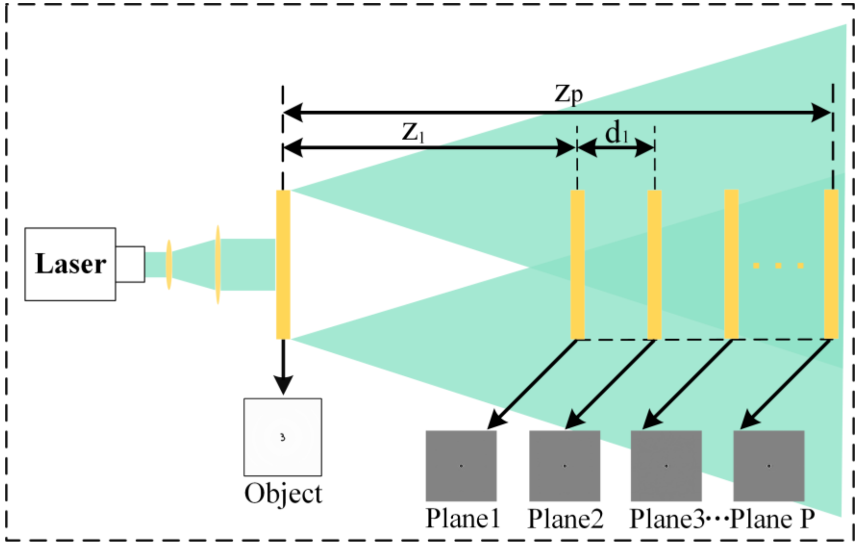
- ➢ High-quality reconstructions are also achieved for dynamic targets
- ➢ The proposed method has smaller residuals
- ➢ Compared with TCCDI, the uniformity of the reconstructed image is better

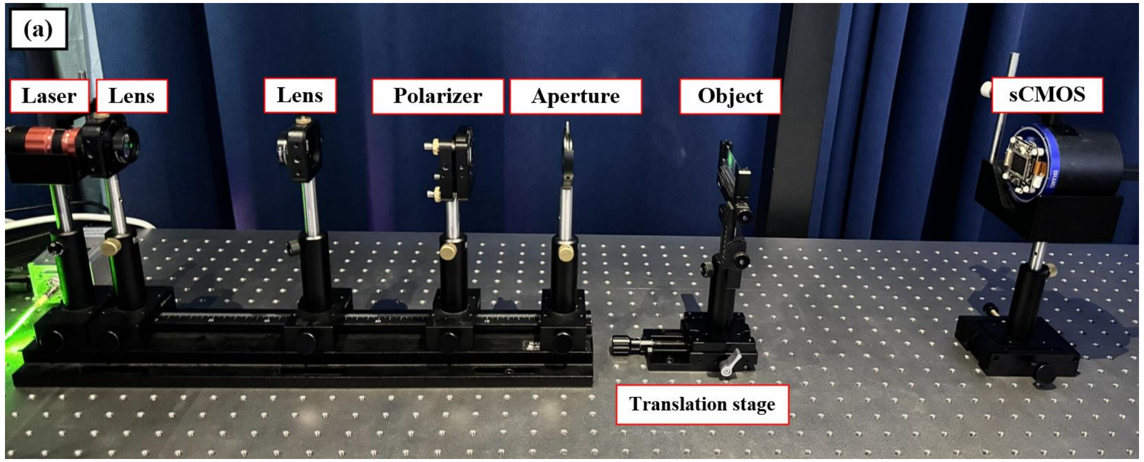
Part V

Generative AI assisted multiplane Digital holographic reconstruction

Multi-plane digital holographic imaging

- Digital holography can reconstruct the amplitude and phase information of the target light field.
- However, the reconstruction quality is largely limited by the size of the hologram.
- Multi-plane holograms can impose constraints for reconstruction, yet the quality of the reconstructed images continues to be restricted owing to the deficiency of effective prior information constraints.

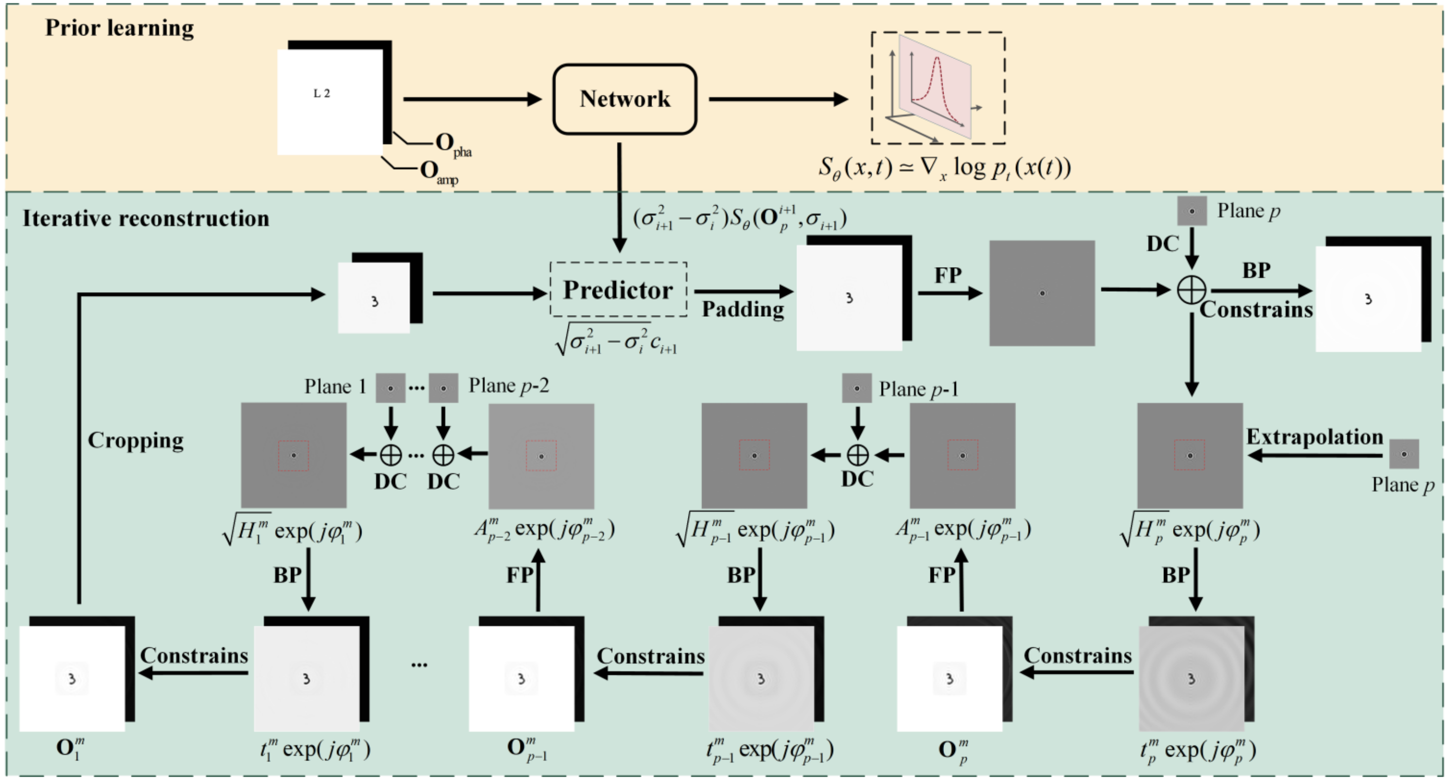




Multi-plane digital holographic imaging system

- For each target, five holograms are measured at different positions.
- The object on the first plane is positioned at a distance from the sCMOS sensor.
- Five holograms are captured at equal intervals.

Diffusion model-boosted multiplane extrapolation for digital holographic reconstruction



Prior learning

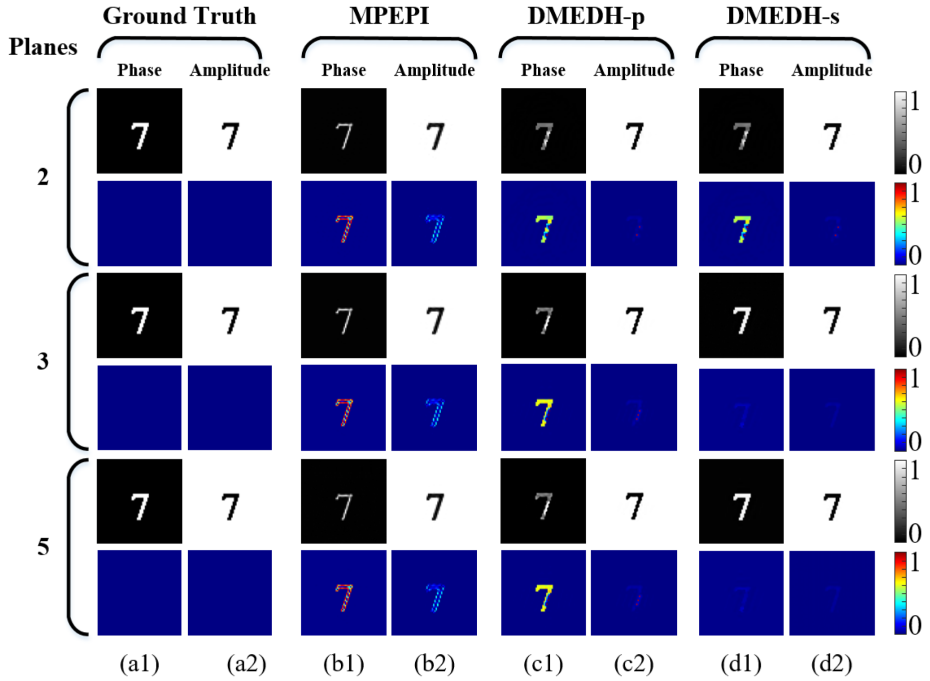
- ✓ The gradient distribution of amplitude and phase is learned by denoising score matching.

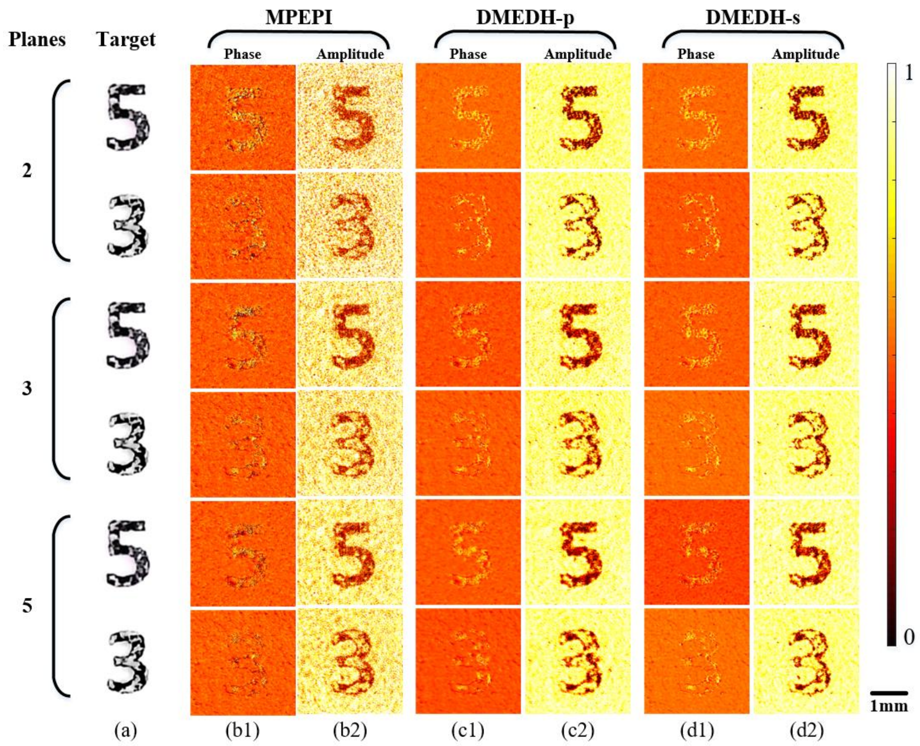
Iterative reconstruction

- ✓ The information correlation between the amplitude and phase image channels and the hologram channel is established.
- ✓ The score function of the training process is used to predict the noisy images.

S. Gao et al., DMEDH: diffusion model-boosted multiplane extrapolation for digital holographic reconstruction, Optics Express, 2024, 32(18):31920-31938.

Results





Planes		MPEPI [dB / NA / NA]	DMEDH-p [dB / NA / NA]	DMEDH-s [dB / NA / NA]
2	Phase	18.53/0.4776/0.0143	20.83/0.7864/0.0091	**20.83/0.7864/0.0091**
	Amplitude	29.77/0.9900/0.0011	30.74/0.9929/0.0010	**30.74/0.9929/0.0010**
3	Phase	18.68/0.4574/0.0139	18.80/**0.7791**/0.0151	**24.11**/0.7072/**0.0068**
	Amplitude	30.69/0.9919/0.0009	30.77/0.9910/0.0009	**34.71/0.9960/0.0005**
5	Phase	18.66/0.4707/0.0139	18.79/**0.7744**/0.0152	**28.38**/0.6441/**0.0027**
	Amplitude	30.73/0.9921/0.0009	30.82/0.9910/0.0009	**38.22/0.9988/0.0002**

The bold numbers represent the best quantitative results.

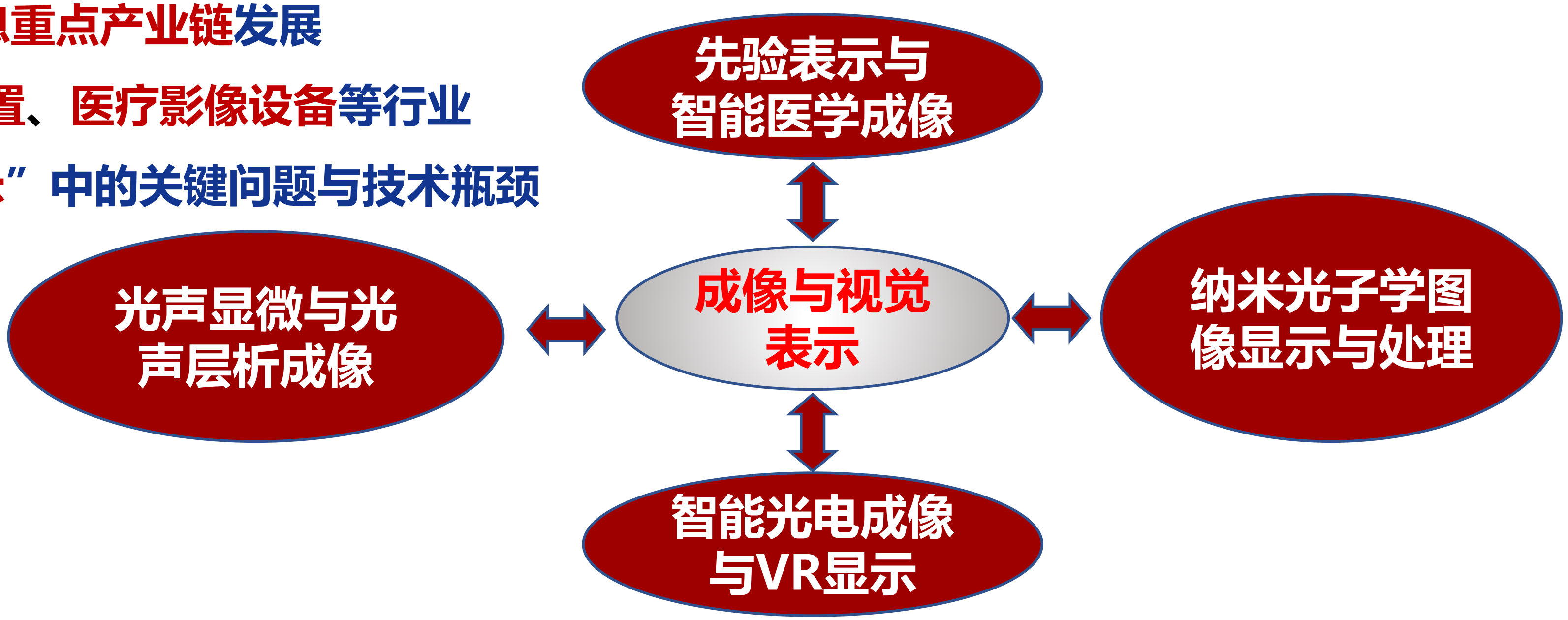
- **Reconstructed image using the DMEDH-s method from 2 planes is better than that of the MPEPI method from 5 planes.**
- **Better imaging results can be obtained with fewer planes using DMEDH**

南昌大学成像与视觉表示实验室

Laboratory of Imaging and Vision rEpresention, LIVE

成像与视觉表示实验室

- 立足国家战略需求和江西省电子信息重点产业链发展
- 聚焦光电成像系统、VR/AR显示装置、医疗影像设备等行业
- 围绕“传感成像-信号处理-增强显示”中的关键问题与技术瓶颈
- 实现关键技术突破与成果转化

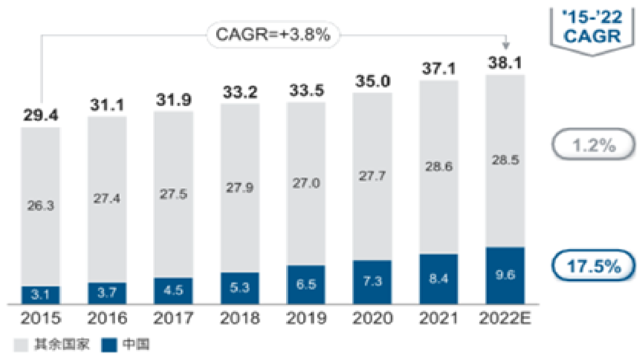


实验室网站：
https://github.com/yqx7150
https://www.labxing.com/lab/1018

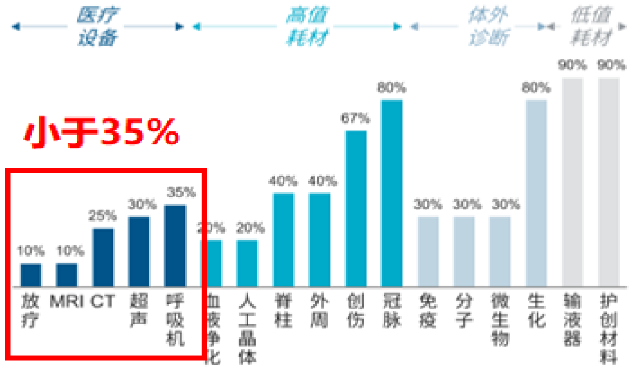
刘且根 教授
国家优青

研究方向1：先验表示与智能医学成像

大型医疗器械国产化替代需求



全球/中国医疗器械市场规模 （千亿元）



医疗器械细分市场国产品牌占比

“企业-高校-医院”协同创新

医院

医疗器械

XGY 鑫高益医疗

RAY SOLUTION 锐世

寰彤核磁

MinFound 明峰

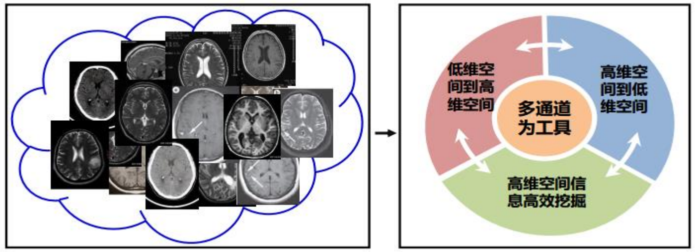
远程医疗

云影医疗

九峰医疗 JF Healthcare

海默股份 Hey Metaverse

- 以高校为主要力量
- 联合企业力量为补充
- 结合医院提出目标任务
- 解决产业链痛点、难点、堵点



高维变量相关和多维空间分析新模式

优秀成果

✓ 转化应用

研究方向2：光声显微与光声层析成像

◆ 乳腺癌检测主要难点

传统单模态成像精度低，仅实现解剖结构检测

- X光钼靶摄影有电离辐射，超声检测特异性低

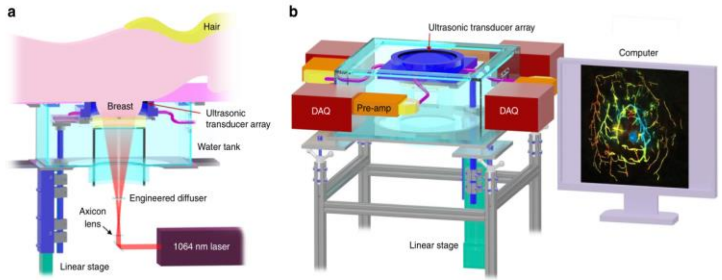
◆ 面向乳腺癌早期检测与诊断

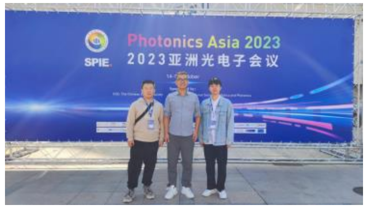
- 实现无损、高对比度、高分辨率深层智能成像
- 融合多模态信息，实现特异性诊断

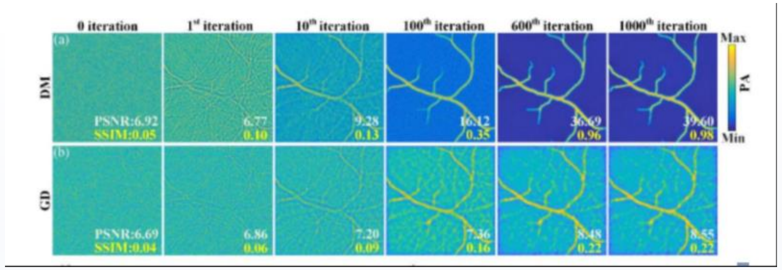
◆ 关键技术难题

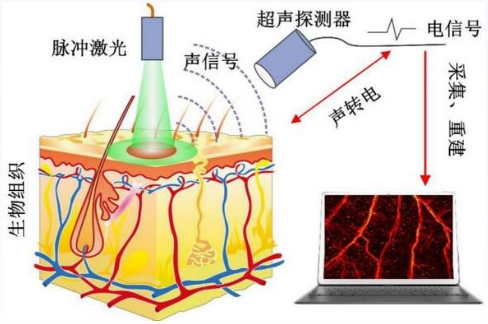
➢ 超声探测器数量有限、受到伪影干扰

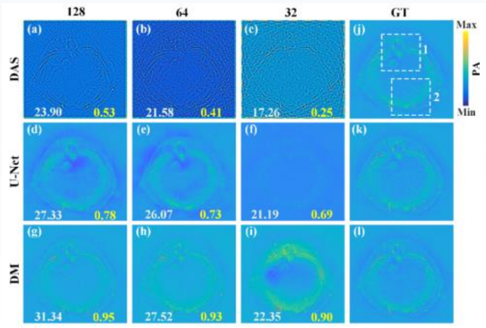
➢ 光学参数背景先验信息缺失

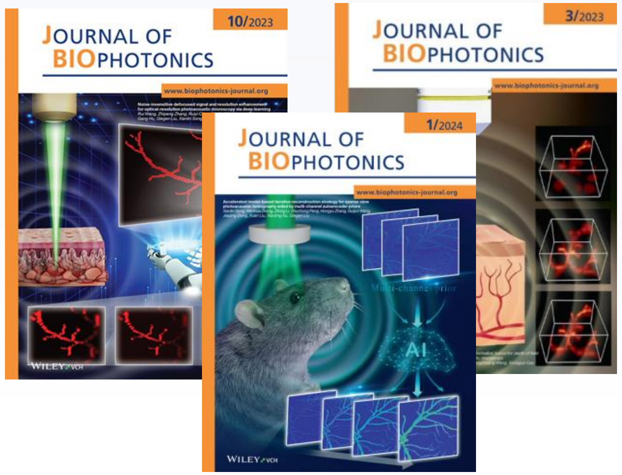












研究方向3：智能光电成像与VR显示

产业需求牵引

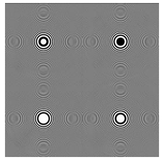
HUAWEI

OFILM 欧菲光

ZETTAVR 境

复杂场景智能光电成像

➢ 多相位菲涅尔孔径编码无透镜成像

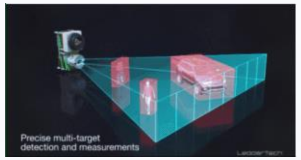


多相位FZA 编码板

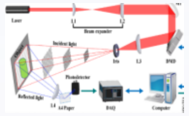
CMOS

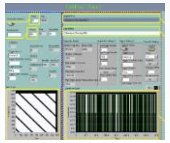
光学计算重建

➢ 直接飞行时间激光雷达模组



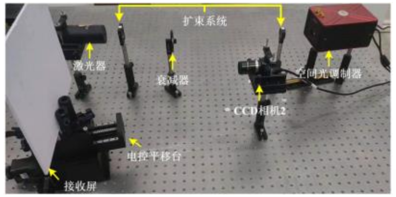
➢ 高速追踪单像素探测系统

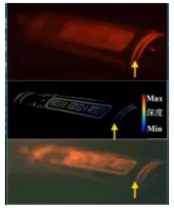




全息VR显示

➢ 三维场景快速采集与全息显示

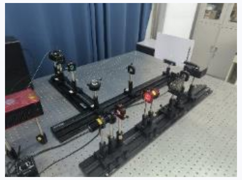


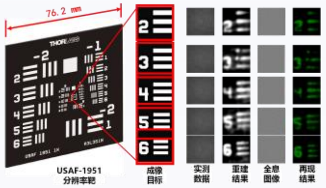


➢ 极简系统结构

➢ 超快三维场景采集速度：～0.6 s

➢ 透过散射介质成像与全息显示





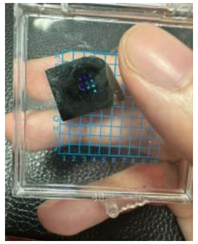
➢ 透过散射介质实现高分辨率成像

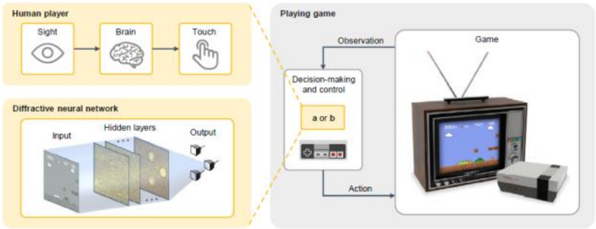
➢ 实现高质量全息显示

研究方向4：纳米光子学图像显示与信息处理

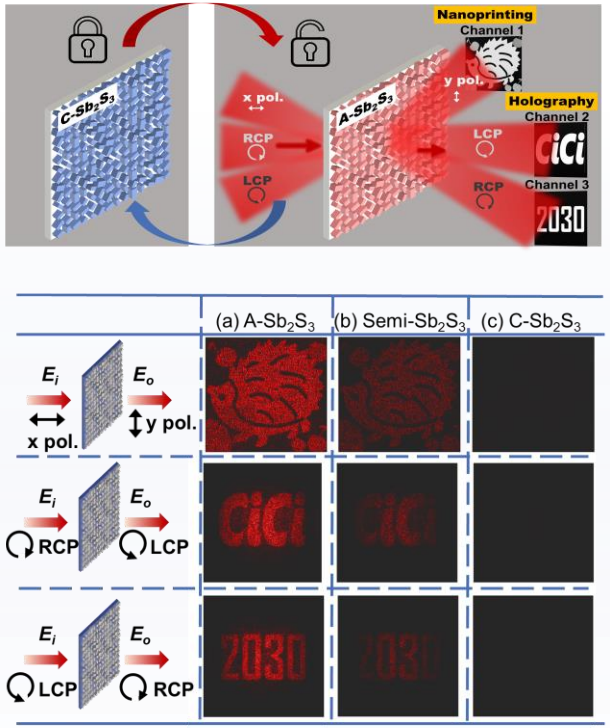
◆ 面向新一代信息技术产业的光学平面集成器件

光子取代电子作为信息载体和能源介质

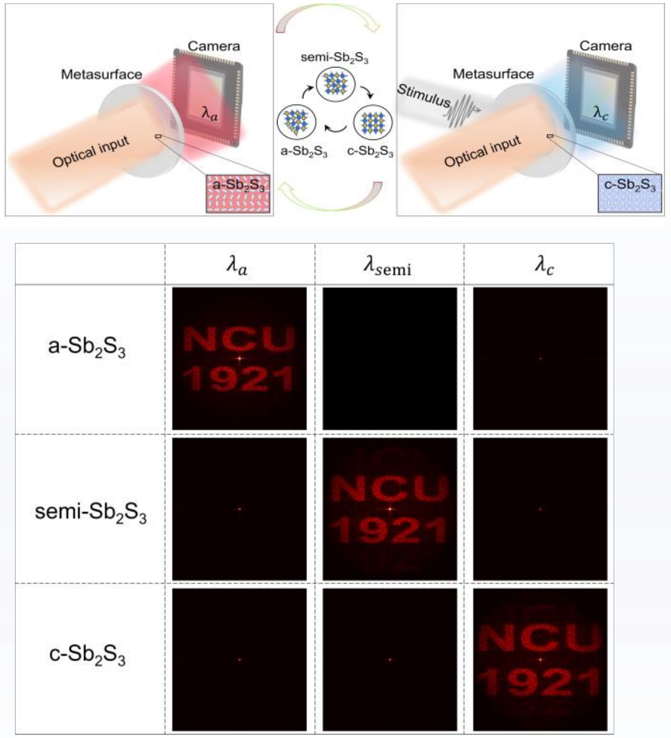




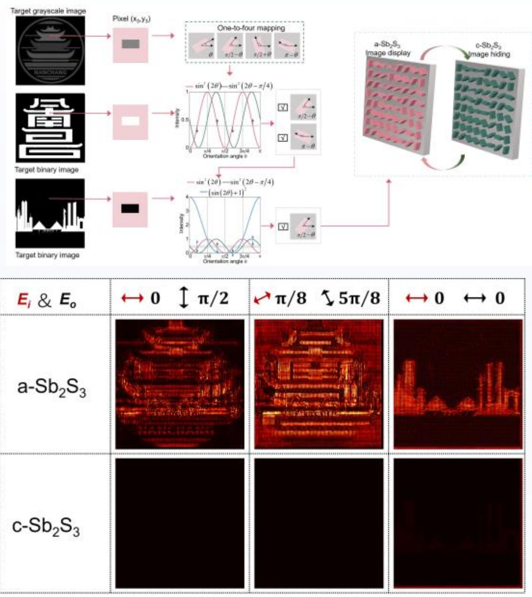
- ✓ 高速度、大带宽、低功耗、高集成
- ✓ 模拟人类的感知能力与控制行为



- ✓ 高容量、高集成度的片上信息存储
- ✓ 高空间分辨率的全息图像显示



- ✓ 光谱和空间调控的高自由度
- ✓ 可调谐波长的全息图像显示



- ✓ 信息复用超表面的极简设计
- ✓ 动态信息显示及加密编码

Thank you!

Wenbo Wan（万文博）

School of Information Engineering, Nanchang University