

光学学报

基于时空编码调控的像素超分辨成像:研究现状与发展趋势(特邀)

梁坤瑶^{1,2,3}, 张许^{1,2,3}, 裴子豪^{1,2,3}, 李洪春^{1,2,3}, 刘昕^{1,2,3}, 王博文^{1,2,3**}, 陈钱³, 左超^{1,2,3*}¹南京理工大学电子工程与光电技术学院, 智能计算成像实验室(SCILab), 江苏 南京 210094;²南京理工大学智能计算成像研究院, 江苏 南京 210019;³江苏省视觉传感与智能感知重点实验室, 江苏 南京 210094

摘要 受限于光电效应本征机制与探测器制造工艺,传统基于光强直接探测的光电成像技术在实现“高分辨率、高灵敏度、高信噪比”关键性能指标方面面临诸多瓶颈,特别是所获取图像的空间分辨率以及空间带宽积,长期受限于探测器像元尺寸与器件物理规模。近年来,计算成像作为融合前端时空编码调控与后端数字计算反演的成像新范式,为打破传统成像技术的诸多限制提供创新的解决方案,代表未来先进光学成像的发展方向。系统回顾基于时空编码调控的像素超分辨成像技术的研究进展,重点围绕孔径平面与焦平面调控策略,阐述光场高维调制机制及其后端图像重建算法,二者相互协同以突破探测器奈奎斯特采样极限,实现亚像素空间信息的有效解耦与复原。相关技术在保持成像系统紧凑性与光通量利用率的基础上,显著提升成像系统的空间分辨率与空间带宽积,为复杂环境下远距离、宽视场、高精度光电探测识别等应用提供新颖解决方案与关键技术支撑。

关键词 计算成像; 像素超分辨; 时空编码调控; 光电探测

中图分类号 TP394.1; TH691.9

文献标志码 A

DOI: 10.3788/AOS251177

1 引言

人类对光影最初的探索,并非始于精密的仪器,而是源于对自然现象的细微洞察。正如阳光穿过茂密树叶的间隙,在地面投下斑驳的光影,东方先哲墨子发现“小孔成像”,并在《墨子·经下》中记载“景到,在午有端,与景长”。这标志着人类历史上首次系统地观察、记录并阐释光学成像的基本原理。1969年,贝尔实验室的 Willard Boyle 和 George E. Smith 在研究金属氧化物半导体技术(MOS)时,发明了基于光电效应^[1]的电荷耦合器件(CCD)。6年后,伊士曼柯达公司的工程师 Steven Sasson 首次在半导体的结构中实现了将光信号转换为电信号,并完成图像的数字化采集和记录的过程,成功研制出世界首台 CCD 数码相机,进而拉开了光电成像的时代序幕。在之后的近半个世纪,得益于信号传输与通信、激光与光全息、光电信号数字化及数字信号处理等关键技术的融合与突破,光学成像

技术从早期的光化学成像、胶片成像,发展至当今以光电数字成像为核心的数码时代,如图 1 所示。

然而,传统的光电成像技术仍基于“小孔成像”的思想,即目标场景通过光学系统以“点对点”的方式直接投影到传感器上,将连续的光场信息进行离散采样并量化为“所见即所得”的数字图像。因此,传统光学成像系统仅能记录二维的空间光强度分布,导致光场的振幅^[2]、相位^[3]、偏振^[4]、光谱^[5]等多维物理信息在投影过程中被压缩为单一的强度信息,高维物理场信息获取受到极大限制^[6]。更为关键的是,成像系统的空间分辨率^[7]作为衡量图像质量和信息量的核心指标,长期以来受到光学分辨率(由镜头孔径尺寸所决定的衍射极限)与像素分辨率(由探测器的像元尺寸所决定的采样频率)的双重物理制约^[8],难以实现分辨率的大幅提升。

在过去的十多年里,信息技术的高速发展为光学成像领域注入新的活力,计算成像(computational

收稿日期: 2025-05-30; 修回日期: 2025-06-24; 录用日期: 2025-06-25; 网络首发日期: 2025-06-25

基金项目: 国家自然科学基金(U21B2033, 62227818, 62361136588, 62305162)、国家重点研发计划(2022YFA1205002, 2024YFE0101300, 2024YFF0505603, 2024YFF0505600)、江苏省科技计划生物医药竞争类项目(BE2022847)、江苏省科技计划重点国别产业技术研发合作项目(BZ2022039)、中央高校科研专项资助项目(30923010206, 30925020211)、中央高校基础科研业务费专项资金资助项目(2023102001, 2024202002)、等离子体物理全国重点实验室项目(JCKYS2024212804)、江苏省光谱成像与智能感知重点实验室开放基金(JSGP202105, JSGP202201, JSGPCXZNGZ202401)

通信作者: *zuochao@njust.edu.cn; **wangbowen@njust.edu.cn

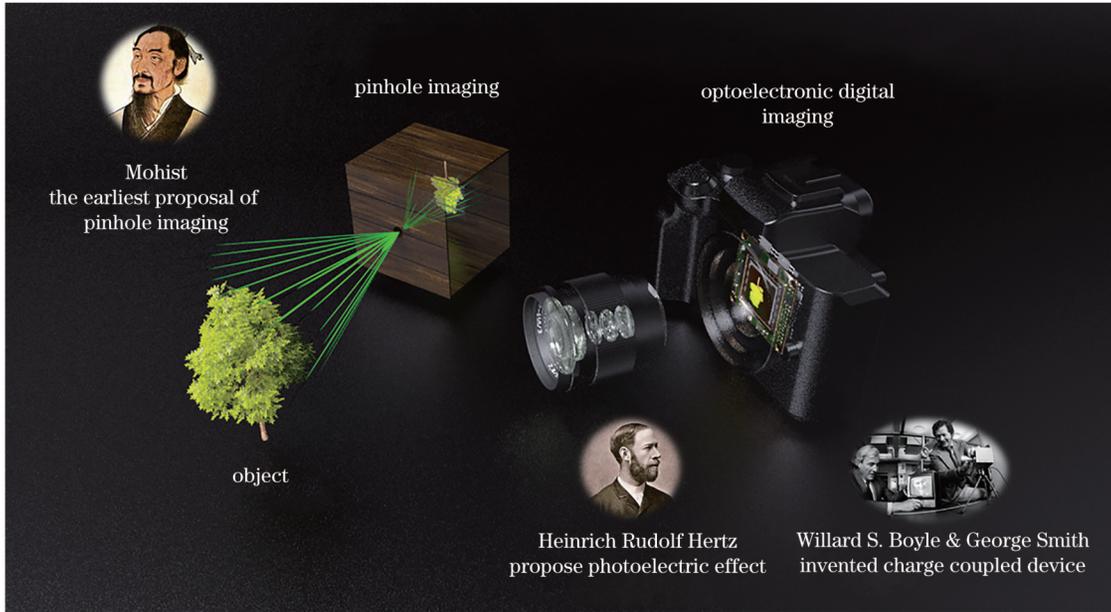


图 1 小孔成像到光电数字成像

Fig. 1 Pinhole imaging to optoelectronic digital imaging

imaging) 技术应运而生^[9],重新定义了感知外界信息的方式。从“所见即所得”的一一映射到对高维光场的耦合编码与计算重构,计算成像将光作为信息载体的一部分,模糊了物理世界与数字世界的边界,从而突破传统光学成像的物理限制。将传统成像中难以获取的高维信息编码至可测量的低维数据,再通过计算重建算法解码恢复,从而突破传统光学成像在衍射极限、采样定理及空间带宽积等方面的物理限制,为光学成像的功能拓展与性能提升提供全新范式。

作为计算成像框架下提升光学系统成像性能的技术策略之一,时空编码调控策略从“时”与“空”两个维度对光场进行编码。本文旨在梳理“时空编码调控”框架下的远场像素超分辨成像技术,探讨其如图 2 所示的在不同成像链路的时空编码调控策略中的研究脉络与应用,并展望其未来发展趋势。第 2 节主要探究传统光电探测器空间分辨率受限因素。第 3 节主要探究在不同成像链路的时空编码调控策略下的像素超分辨成像研究现状。第 4 节对全文进行总结并展望。

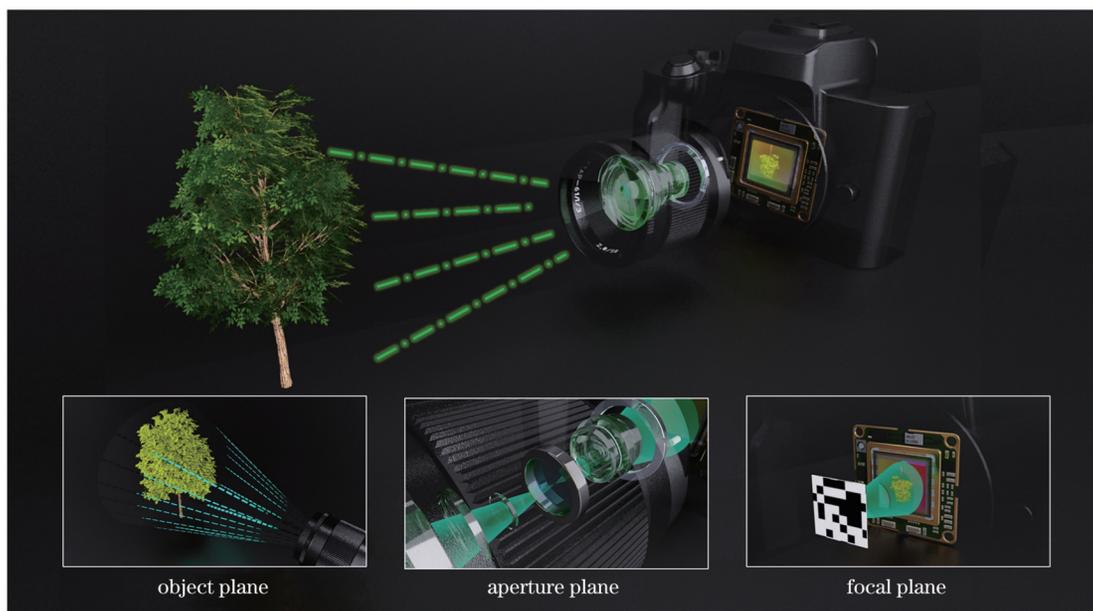


图 2 基于时空编码在成像链路不同位置处调控示意图

Fig. 2 Schematic diagram of spatio-temporal coding-based modulation at different positions in imaging chain

2 传统光电探测器件的空间分辨率

成像系统的空间分辨率^[7]作为衡量图像质量和信息量的核心指标,长期以来受到光学分辨率(由镜头孔径尺寸所决定的衍射极限)与像素分辨率(由探测器的像元尺寸所决定的采样频率)的双重物理制约^[8]。光学分辨率为由瑞利判据或阿贝衍射极限定义的成像系统可分辨两点的最小距离。像素分辨率由探测器最小单元结构感光单元,即像元的尺寸和数量决定。其中,像元尺寸决定探测器的采样间隔,像元数量决定探测器能承载的信息容量。从频域的角度看,光学分辨率决定系统能传递的最高空间频率,而像素分辨率决定最终能转化为数字图像的最高采样频率,如图 3 所示。二者如同“木桶效应”中的两块木板,共同决定成像系统最终的空间分辨率。

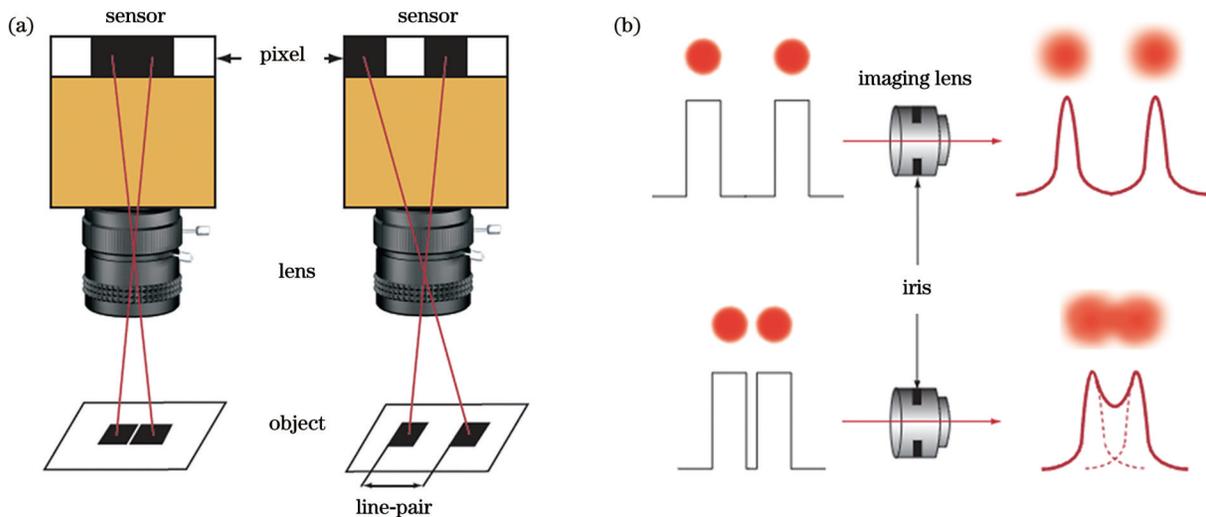


图 3 成像系统的采样极限与光学衍射极限

Fig. 3 Sampling limit and optical diffraction limit of imaging system

从光电转换的物理机制来看,成像系统的探测器端同样面临着灵敏度与分辨率的固有矛盾。首先,像素的物理尺寸直接决定其光子捕获能力:像元尺寸缩小,其接收光子的面积随之减小,导致在同样曝光时间下捕获的光子数量急剧下降。这使得有效信号相对于固有噪声(如暗电流、读出噪声等)而言更加微弱,原有的细节信息可能被噪声淹没,从而限制对细节分辨能力的提升。其次,随着像素尺寸的减小和集成密度的提高,像素之间的光电串扰现象会愈发严重,光或电信号在相邻像素间的“泄漏”“串扰”会造成读出图像模糊,进一步降低对精细结构的辨识能力^[10],如图 4 所示。噪声的存在进一步模糊理论上的可分辨极限,即使一个特征在理论上可以被衍射和采样所分辨,如果其信号强度相对于噪声水平过低,那么在实际图像中也可能无法被有效识别。因此,如何权衡光电成像系统设计,实现“高灵敏”与“高分辨”成像探测,成为亟需

从信息论的角度来看,根据奈奎斯特采样定理,为确保原始光学信号能够无失真地恢复,探测器的采样频率必须至少是光学最高频率的 2 倍。若探测器的像元尺寸过大将导致采样频率不足,图像信息的高频成分被混叠到低频成分当中,并在离散采样后出现“马赛克”效应。在此情况下,即使光学系统能够提供更高的衍射极限分辨率(如具有高数值孔径的显微镜物镜),成像系统的空间分辨率依然受限于探测器的采样极限,难以进一步提升。反之,当探测器像元尺寸提供充足的采样频率时,如果光学系统的艾里斑尺寸大于像元尺寸,那么即使过采样也无法恢复超越衍射极限的频率细节,此时成像系统的空间分辨率主要由光学衍射极限决定。不难理解,当成像系统的分辨率受限于采样极限的“短板”时,其分辨率提升的“天花板”往往由光学衍射极限的“长板”所决定的,反之亦然。

解决的目标。

3 基于时空编码调控的像素超分辨成像研究现状

像素超分辨的核心在于精确估计出目标高分辨率(HR)像素网格与原始低分辨(LR)像素采样之间的响应关系或映射模型^[7],即从观测到的低分辨率图像中恢复出丢失的高频细节。从数学的角度来看,这一过程本质上是信息欠定的观测数据中恢复原始高频信号的病态逆问题^[11-12],且当成像过程伴随不确定性噪声干扰时,该逆问题的求解矩阵条件数会恶化,使得重建结果对噪声高度敏感,并产生严重伪影。为此,必须引入有效的正则化策略,通过融入先验知识——如精确的成像模型、统计噪声模型,或从大量外部样本中学习到的图像先验(如稀疏性、自相似性、深度学习特征等)——来构建合理的约束条件,从而将病态问题转化

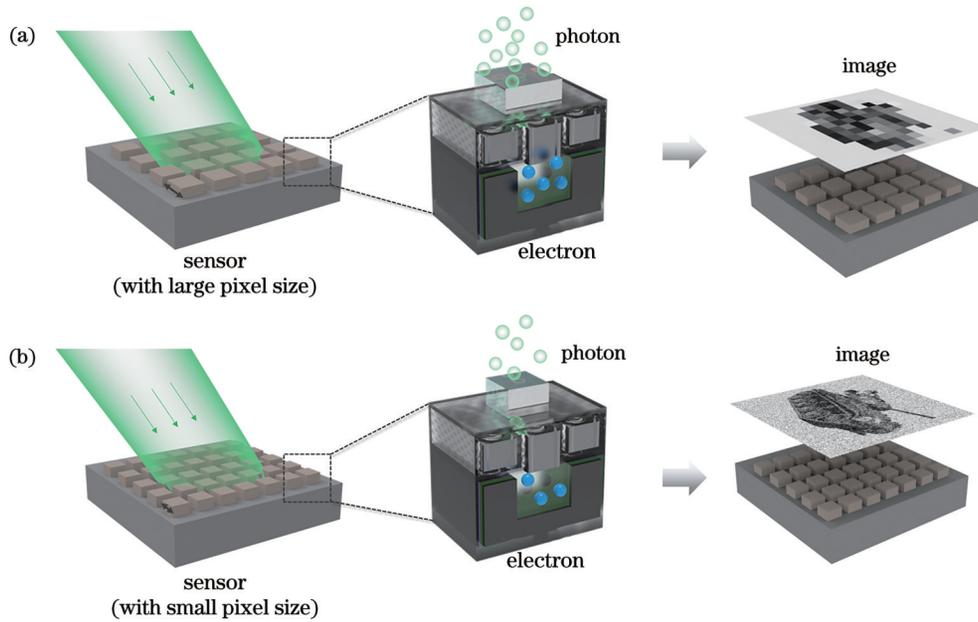


图 4 基于光电转换的物理机理探究分辨率与灵敏度之间的关系。(a)大像素尺寸的探测器在相同积分时间内能够捕获的光子多,生成的图像信噪比高,但图像分辨率低;(b)小像素尺寸的探测器在相同积分时间内能够捕获的光子少,生成的图像信噪比低,但图像分辨率高

Fig. 4 Relationship between resolution and sensitivity based on physical mechanism of photoelectric conversion. (a) Detectors with large pixel sizes can capture more photons within the same integration time, generating images with high signal-to-noise ratios but low resolution; (b) detectors with small pixel sizes can capture fewer photons within the same integration time, generating images with low signal-to-noise ratios but high resolution

为适定或近似适定的问题,进而稳定地重建出高质量的高分辨率图像。图 5 是基于像素超分辨成像技术的分类。计算成像对多维度光场信息进行灵活调控与

高效利用,可调控维度包括时间、空间、振幅、相位、偏振及光谱^[13-15]等。通过对这些维度进行巧妙编码与信息置换——例如牺牲时间分辨率换取空间分辨率

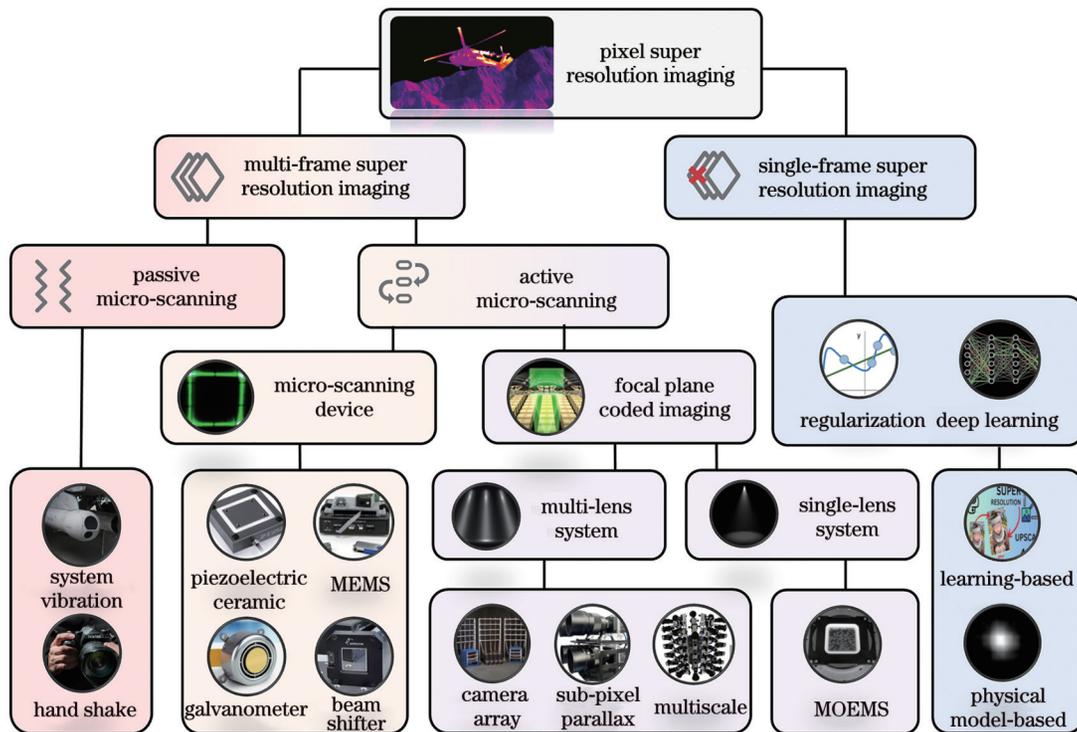


图 5 基于像素超分辨成像技术的分类

Fig. 5 Pixel super-resolution imaging technology-based classification

提升,或牺牲部分视场以增强特定区域细节分辨能力——计算成像技术能在系统有限信息带宽内,最大限度提升空间分辨率。

作为计算成像框架下提升光学系统成像性能的技术策略之一,时空编码调控策略从“时”与“空”两个维度对光场进行编码。其中,“时”(temporal)维度指对时间资源的有效利用与调控,既可以表现为单次曝光时间内对光信号进行高速时序调制,也涵盖在通过多次曝光捕获目标的场景中不同的瞬时信息。“空”(spatial)维度则具有两重含义:从调控机制层面,其旨在对成像链路中的特定位置(如物面、孔径面、中间像面或焦平面)引入特定的照明策略或光学编码器件^[16];从调控目标层面,其旨在通过编码调控,使得最终投射至探测器处的光场呈现特殊的空间强度结构特征。在

实际应用中,时域编码与空域编码两者往往并非孤立存在,而是紧密结合、相辅相成,协同突破传统成像模式在信息获取上的固有局限,进而显著提升光学系统的综合成像性能,例如在时空分辨率、视场范围、成像景深、信噪比乃至多维信息(如光谱、偏振)获取能力等方面实现增强^[17]。

在“计算光学成像”与“时空编码调控”的框架下,像素超分辨技术通过在成像链路的孔径平面、焦平面引入如图 6 所示的时空编码器件进行高维光场调控,牺牲时间维度(时间分辨)换取空间维度(空间分辨)提升,结合后端重建算法与深度学习技术,解耦亚像素信息并突破探测器空间采样的物理约束,实现混叠信号的无失真或近似的恢复。

本节旨在梳理“时空编码调控”框架下的远场像素

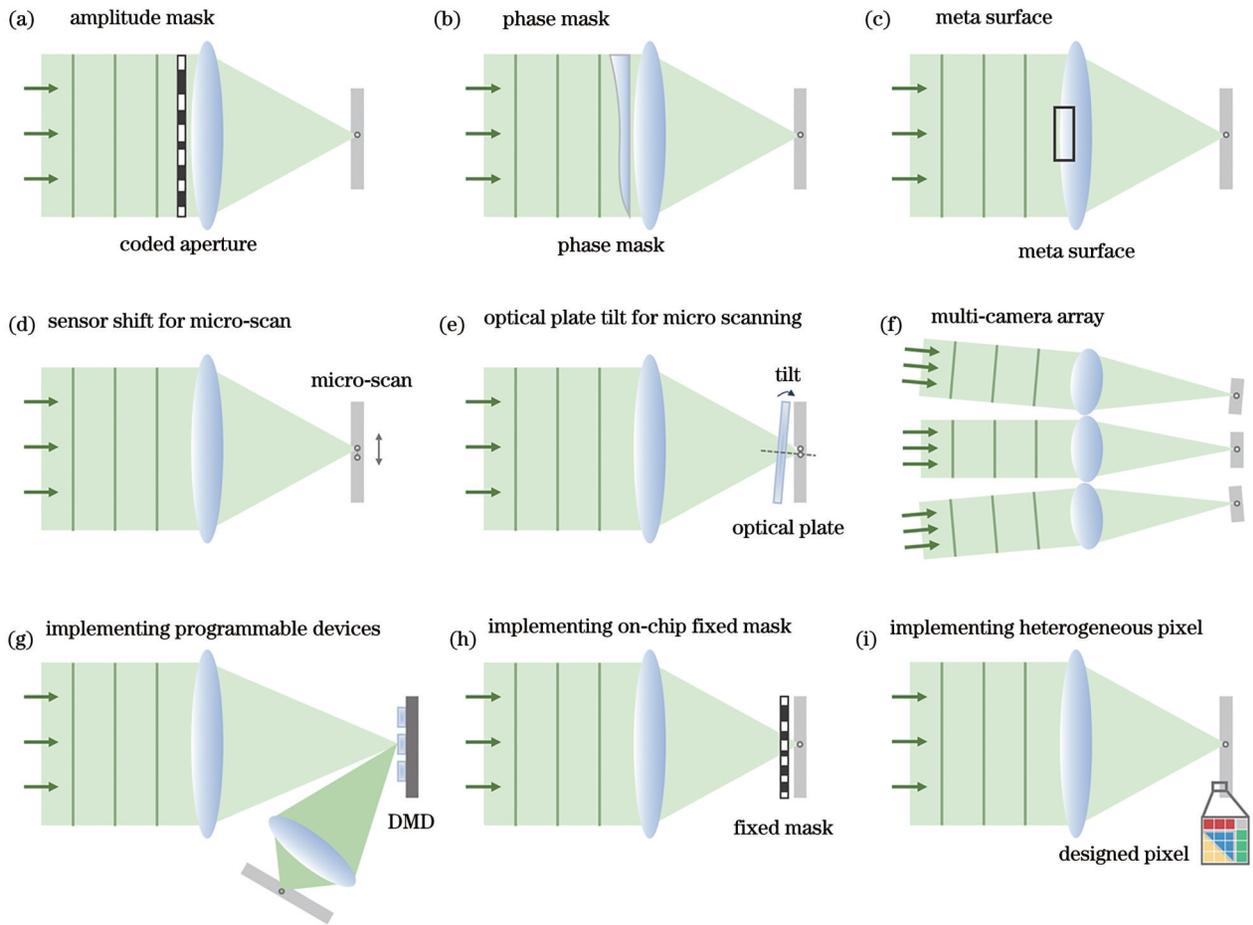


图 6 基于不同成像链路中的时空编码调控原理图。(a)基于孔径平面的振幅编码调控;(b)基于孔径平面的相位编码调控;(c)基于孔径平面的超表面编码调控;(d)基于焦平面的微扫描编码调控;(e)基于焦平面的光学元件编码调控;(f)基于焦平面的相机阵列编码调控;(g)基于焦平面的可编程器件编码调控;(h)基于焦平面的固定掩模编码调控;(i)基于焦平面的异构像元探测器编码调控

Fig. 6 Principle diagram of spatio-temporal encoding modulation in different imaging links. (a) Aperture plane-based amplitude encoding modulation; (b) aperture plane-based phase encoding modulation; (c) aperture plane-based metasurface encoding modulation; (d) focal plane-based micro-scanning encoding modulation; (e) focal plane-based optical element encoding modulation; (f) focal plane-based camera array encoding modulation; (g) focal plane-based programmable device encoding modulation; (h) focal plane-based fixed mask encoding modulation; (i) focal plane-based heterogeneous pixel detector encoding modulation

超分辨成像技术,探讨其在不同成像链路的时空编码调控策略中的研究脉络与应用。第 3.1 节主要聚焦于孔径平面的时空编码调控策略,以振幅编码掩模、相位编码掩模以及超表面等调控方式为核心展开讨论,如图 6(a)~(c)所示。第 3.2 节介绍基于焦平面端的时空编码调控策略的相关研究,并根据调控的方式不同,从微扫描器件调控、可编程器件调控、固定编码掩模调控以及基于异构像元探测器编码调控这四个角度对焦平面调控的实现方法进行总结,如图 6(d)~(i)所示。第 3.3 节主要介绍基于深度学习的像素超分辨成像技术研究,并将其分为基于数据驱动和物理模型驱动的像素超分辨成像分别展开。

3.1 基于孔径平面的时空编码调控像素超分辨成像技术

在计算光学成像链路中,孔径平面是指决定成像系统所能探测的最高空间频率的临界光瞳位置。在该平面引入编码调控,直接对系统点扩散函数(PSF)与光学传递函数(OTF)进行全局调控。本节聚焦孔径平面时空编码像素超分辨的机制与进展:首先论述基于经典振幅/相位调制元件的编码策略及其在分辨率增益中的作用机理,如图 7(a)和(b)所示;随后探讨超表面等新兴微纳器件因扩展光场自由度在构建紧凑高效孔径编码系统中具有的潜在优势及其相关前沿探索,如图 7(c)和(d)所示。

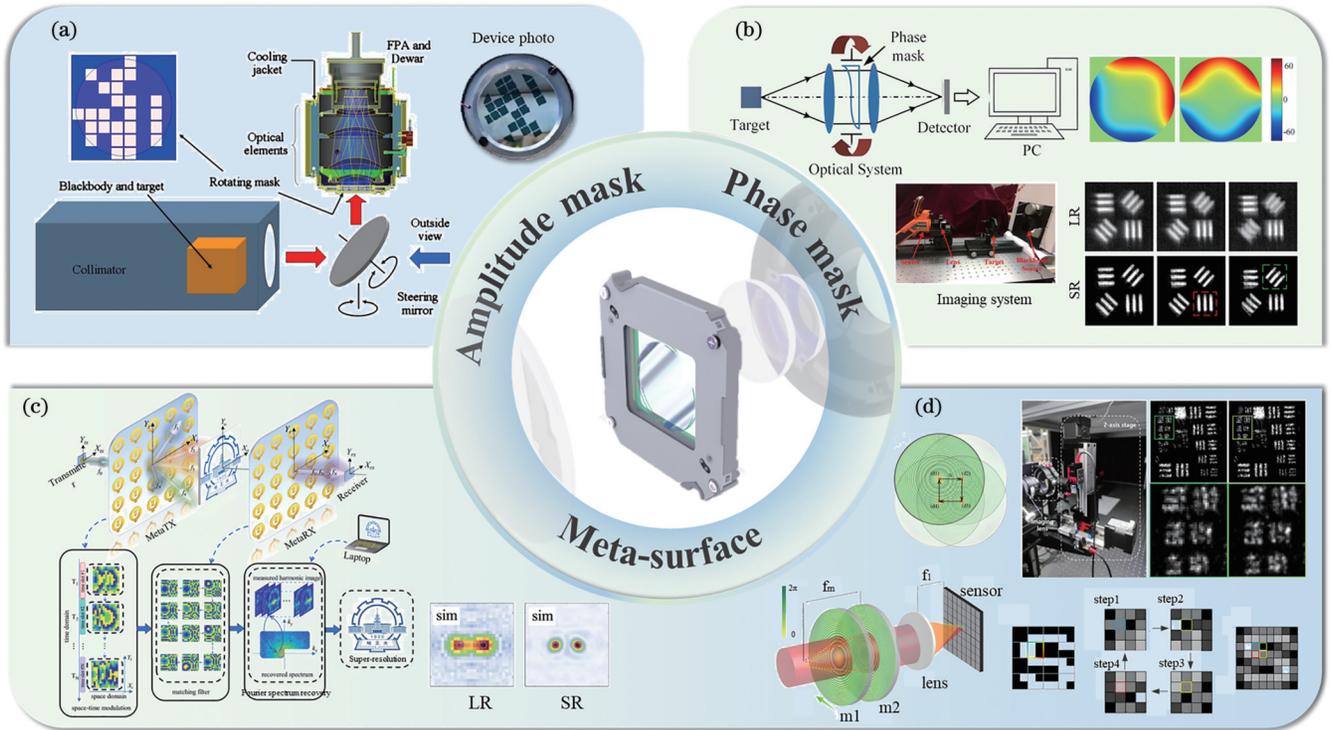


图 7 基于孔径平面时空编码的方法分类及其代表性工作。(a) Slinger 等^[18-19]提出的基于振幅编码调控实现超分辨成像;(b) Yi 等^[20]提出的基于相位编码调控实现超分辨成像;(c) Li 等^[21]提出的基于超透镜调控实现超分辨成像;(d) Zhang 等^[22]提出的基于超透镜调控实现超分辨成像

Fig. 7 Classification of aperture plane-based spatio-temporal encoding methods and their representative works. (a) Super-resolution imaging via amplitude encoding modulation, proposed by Slinger et al.^[18-19]; (b) super-resolution imaging via phase encoding modulation, proposed by Yi et al.^[20]; (c) super-resolution imaging via metalens modulation, proposed by Li et al.^[21]; (d) super-resolution imaging via metalens modulation, proposed by Zhang et al.^[22]

3.1.1 基于振幅/相位调控的时空编码技术

基于孔径平面的振幅或相位调控是实现像素超分辨成像的关键技术路径之一。在孔径面处加入可编程掩模或固定掩模图案,调制孔径平面的波前信息,从而改变系统的点扩散函数,使得高频信息能够以特定方式编码复用于多帧低分辨率图像的采样过程中。结合重建算法从这些编码的低分辨率数据中恢复超越奈奎斯特采样极限的高分辨率图像,从而有效提升成像系统的空间分辨率。本质上,孔径编码是对光瞳函数进行精确的强度或相位设计,以生成具有

特定空间分布的光学传递函数。这一调制过程可数学地描述为

$$P_c(\xi, \eta) = P_{\text{mask}}(\xi, \eta) \times \exp[j\phi_{\text{mask}}(\xi, \eta)] \quad (1)$$

$$F_c(f_x, f_y) = \frac{\iint P_c(\xi + f_x, \eta + f_y) P_c^*(\xi, \eta) d\xi d\eta}{\iint |P_c(\xi, \eta)|^2 d\xi d\eta} \quad (2)$$

式中: $P_c(\xi, \eta)$ 为 ξ - η 空间坐标系下的调制广义光瞳函数; $P_{\text{mask}}(\xi, \eta)$ 为编码掩模的强度函数; $\phi_{\text{mask}}(\xi, \eta)$ 为编码掩模的相位函数; $F_c(f_x, f_y)$ 为 f_x - f_y 频域坐标系下的

光学传递函数;下标 c 表示编码(coded);下标 mask 表示掩模图案(mask)。

2010年,Gorden等^[23]提出一种自适应编码孔径成像(ACAI)技术,通过在小型实验系统中采用微光机电系统(MOEMS)的方式以实现高速切换可编程孔径掩模图案,如图8(a)所示,从而实现像素分辨率的提升,如图8(b)所示。2012年,Slinger等^[19]在前述小型

化原理样机试验成功的基础上将编码孔径技术应用至更大型的工程样机,如图8(c)所示,但在实际搭建中该工作通过在红外成像系统孔径面中引入固定编码孔径掩模,记录8或16个不同旋转角度下编码后的图像。实验结果表明,通过采集点扩散函数调制下的图像,实现2倍的像素超分辨率成像以重建目标场景的高频细节,如图8(d)所示。

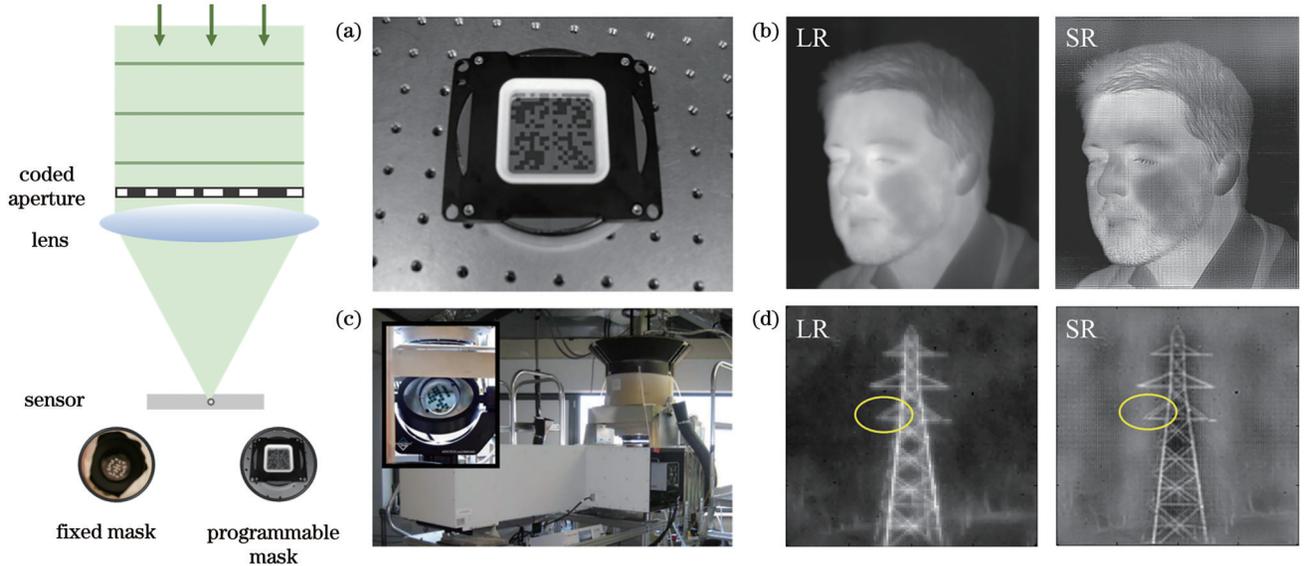


图8 利用振幅掩模实现孔径平面编码调控的像素超分辨成像技术相关工作。(a)~(b) Gorden等^[23]提出的基于微光电机系统的中波红外超分辨成像系统及其超分辨前后对比结果;(c)~(d) Slinger等^[18-19]提出的基于固定编码孔径的中波红外超分辨成像系统及其超分辨前后对比结果

Fig. 8 Related work on pixel super-resolution imaging technology achieving aperture plane coding control using amplitude masks. (a)~(b) Mid-wave infrared super-resolution imaging system based on micro-opto-electro-mechanical systems proposed by Gorden et al.^[23] and its comparison of results before and after super-resolution; (c)~(d) mid-wave infrared super-resolution imaging system based on fixed coded aperture imaging proposed by Slinger et al.^[18-19] and its comparison of results before and after super-resolution

与振幅编码调控相比,相位编码是一种信噪比更高的编码调制方案。由于相位掩模通过引入空间变化的光程差来调制波前相位而不衰减入射光能,其在同等条件下能够获得比振幅编码更高的图像信噪比。2016年,Zhao等^[24]提出在红外成像系统中加入双三次相位板编码调控,同时实现系统景深扩展以及超分辨成像,系统如图9(a)所示。通过光学仿真设计软件精确获取系统相位编码调控后的点扩散函数,对理想的点扩散函数进行重采样后与采集到的编码图像进行反卷积以获取高分辨率图像,结果如图9(b)所示。尽管如此,该重建方法面临两大挑战:一是算法对噪声的鲁棒性较差,在反卷积过程中容易放大噪声并引入计算伪影。二是该方法对系统模型的精确度要求高,任何理论点扩散函数与实际点扩散函数的标定偏差都可能导致重建质量的恶化。2018年,Yi等^[20]发现掩模旋转策略虽不改变点扩散函数的幅值分布,却可引入可控的空间方向性差异。搭建基于旋转相位编码成像系统,如图9(c)所示,由此获取的多帧编码图像

在拉普拉斯金字塔与频谱域进行融合后,超分辨重建图像的细节恢复和伪影抑制效果均得到显著增强,如图9(d)所示。为进一步增强系统对多方向高频信息的响应,Du等^[25]在成像系统对同一块非对称相位掩模进行旋转,使得每帧图像对应的光学传递函数在频域具有方向性差异,进而在频率融合阶段获得合成光学传递函数以增强系统光通量,从而提升超分辨重建精度。

3.1.2 基于超表面的时空编码技术

超表面凭借其亚波长单元结构,通过特定的物理机制对光场的相位、振幅与偏振进行精确的离散操控,可在超薄尺度内实现传统光学元件难以完成的波前调制^[26-28]。其本质是通过在孔径平面直接扩展系统可用自由度,在“空间带宽积调控”框架下重分配传输信道容量,实现波前的编码-解码复用。由于单元结构尺寸与探测器像元尺寸相当甚至更小,这类器件天然契合如像素级编码与亚像元位移的提高空间分辨的编码策略,为突破奈奎斯特采样极限提供新型硬件平台。

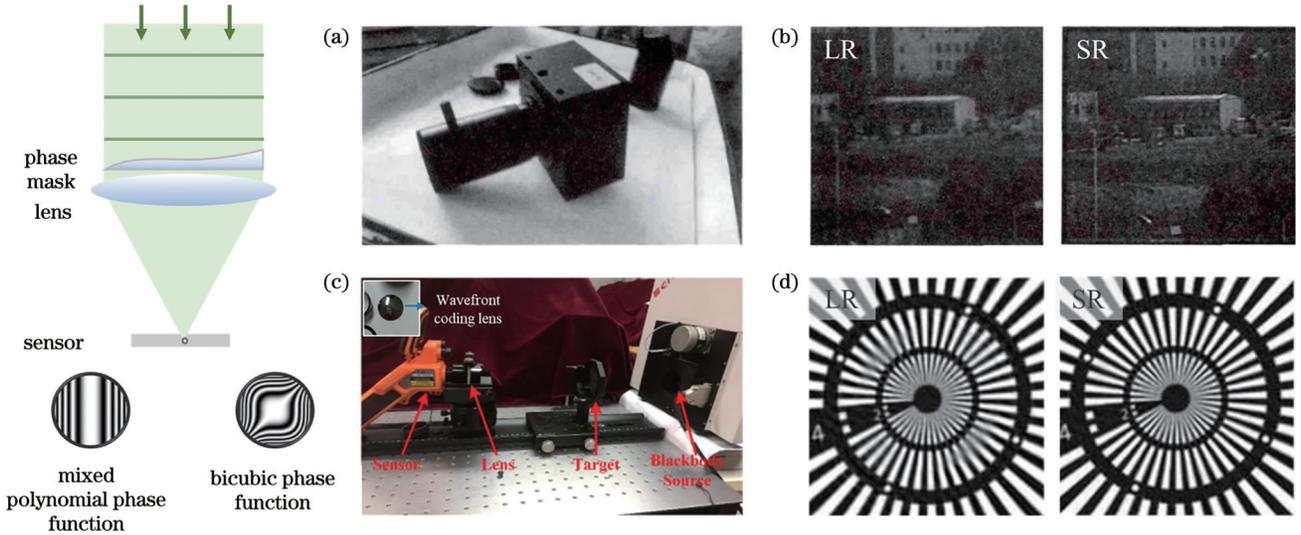


图9 利用相位掩模实现孔径平面编码调控的像素超分辨率成像技术相关工作。(a)~(b) Zhao等^[24]提出的基于波前编码超分辨率成像系统及其超分辨前后对比结果;(c)~(d) Yi等^[20]提出的基于相位编码的旋转多帧像素超分辨率成像实验系统及其超分辨前后对比结果

Fig. 9 Related work on pixel super-resolution imaging technology achieving aperture plane coding control using phase masks. (a)–(b) Wavefront coding-based super-resolution imaging system proposed by Zhao et al.^[24] and its comparison of results before and after super-resolution; (c)–(d) experimental system for rotating multi-frame pixel super-resolution imaging based on phase coding proposed by Yi et al.^[20] and its comparison of results before and after super-resolution

位于孔径平面的超表面,其整体作用相当于一个复杂空间滤波器,即广义光瞳函数 $P_{meta}(\xi, \eta)$ 。超表面上成千上万的单元结构贡献的局域相位共同塑造宏观的相位分布,即光瞳函数的相位项 $\phi_{meta}(\xi, \eta)$,改变任意一个单元的结构参数(如一个纳米柱的直径或旋转角度)则会改变该点的局部相位 $\phi_{local}(\xi, \eta)$,从而改变整体的光瞳函数。该调制关系可用数学公式表达为

$$P_{meta}(\xi, \eta) = P_{pupil}(\xi, \eta) \times \exp[j\phi_{meta}(\xi, \eta)] \quad (3)$$

式中: $P_{meta}(\xi, \eta)$ 为 ξ - η 空间坐标系下的基于超表面调制的广义光瞳函数; $P_{pupil}(\xi, \eta)$ 为广义光瞳函数的强度分布; $\phi_{meta}(\xi, \eta)$ 为超表面全局相位函数; 下标 meta 表示超表面 (metasurface); 下标 pupil 表示光瞳形状。随着可编程、主动可调或级联式超表面技术的成熟,研究者开始将该类超薄器件嵌入成像链路并与计算反演算法协同,形成面向像素超分辨的混合硬件-软件体系。2024年, Li等^[21]提出一种基于时空联合调制的计算超表面成像架构,如图10(a)所示。该架构在收发端各部署一个可重构超表面,并结合结构化照明显微术与傅里叶叠层成像的核心思想,实现对目标场景的超分辨探测。具体而言,发射端超表面(Tx-RIS)采用粒子群优化算法生成最优的照明时序,以合成特定光场;同时,接收端超表面(Rx-RIS)被设计为一个匹配滤波器(matched filter),能够将目标上特定位置的散射场会聚至单像素探测器上,将无法被直接探测到的高频空间信息通过编码照明的方式,搬移到接收系统

可以探测到的低频通带范围内。实验实现了超阿贝分辨率极限 42.3% 的分辨率提升,如图10(b)所示。2025年, Zhang等^[22]通过设计共轭二次相位超表面形成尺度横向平移模块,如图10(c)所示,实现毫米级位移下的亚像元图像平移。具体而言,通过平移其中一个平面光学元件,使其发生毫米级横向位移时,整个模块的光束会发生微小的角度偏转,经过镜头聚焦后最终在焦平面探测器上形成可控的亚像素位移。通过对该元件进行步进式平移,系统便可高效采集一系列包含互补空间信息的低分辨率图像,进而利用多帧超分辨算法重构出高分辨率结果。实验结果表明,该技术成功将远场傅里叶叠层成像的分辨率增强因子由2倍提升至2.8倍,显著提高长距离探测的空间分辨能力,如图10(d)所示。

上述位于孔径平面的编码器件,无论是传统的振幅/相位编码器件还是超表面等新型编码器件,均为提升有效奈奎斯特采样率、压缩系统体积提供可行路径,有望在未来的时空编码计算成像框架内进一步释放像素超分辨潜能。

3.2 基于焦平面的时空编码调控像素超分辨率成像技术

在计算光学成像链路时,焦平面是光电探测器进行光电转换、实现光学图像到数字信号转换的关键位置。通过对探测器的空间采样特性进行精细建模与操控,基于焦平面的时空编码调控技术能够在成像链路的末端以片上像元级调控,编码并重构出亚像素级光场信息。本节将重点探讨在成像系统焦平面引入时空

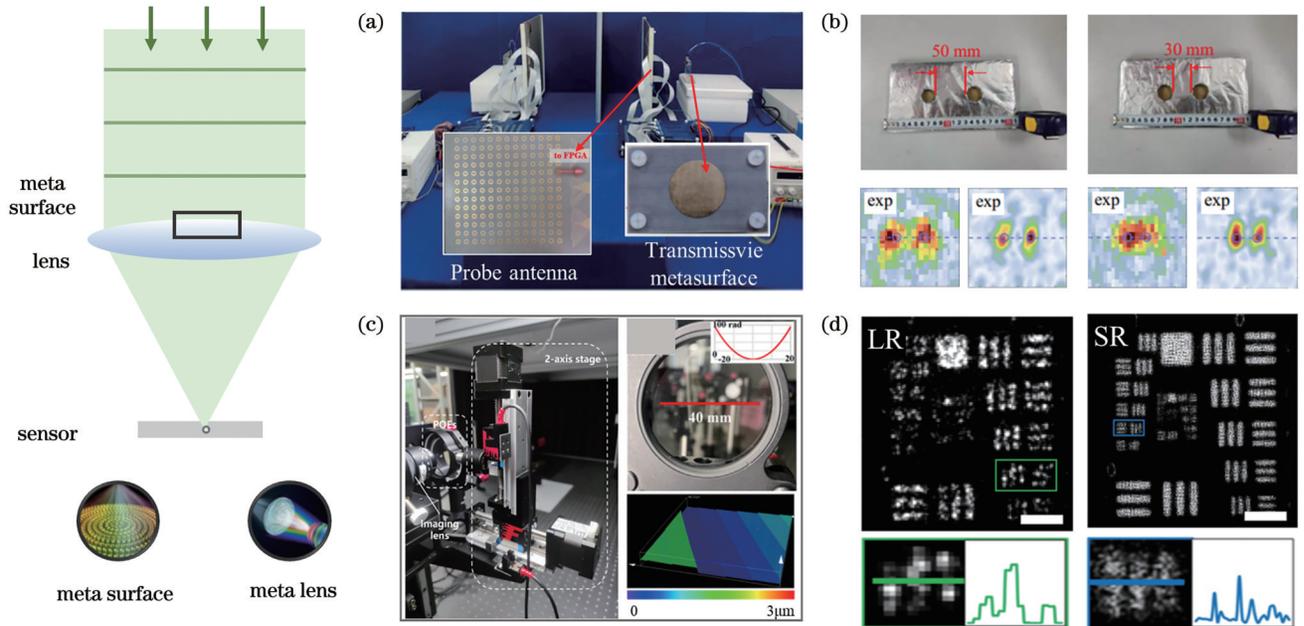


图 10 利用超表面器件实现孔径平面调控的像素超分辨成像技术相关工作。(a)~(b) Li 等^[21]提出的基于可重构超表面的单像素超分辨成像系统及其超分辨前后对比结果;(c)~(d) Zhang 等^[22]提出的基于共轭二次相位超表面对应的远距傅里叶叠层超分辨成像系统及其超分辨前后对比结果

Fig. 10 Related work on pixel super-resolution imaging technology achieving aperture plane control using metasurface devices. (a)–(b) Single-pixel super-resolution imaging system based on a reconfigurable metasurface proposed by Li et al.^[21] and its comparison of results before and after super-resolution; (c)–(d) remote Fourier ptychographic super-resolution imaging system based on a conjugate quadratic phase metasurface pair proposed by Zhang et al.^[22] and its comparison of results before and after super-resolution

编码调控机制以实现像素超分辨的原理、方法及最新进展。如图 11 所示,首先阐述经典的基于焦平面时空编码的像素超分辨成像技术,如微扫描技术(包括主动与被动方式)如何通过引入精确或随机的亚像素位移,以及多孔径相机阵列如何利用空间并行采样获取亚像素级多样性信息。其次,探讨基于可编程器件的焦平面时空编码技术,例如数字微镜器件(DMD)、空间光调制器(SLM)等器件如何通过动态加载和投影特定编码图案,对入射至焦平面附近的光场进行多维度精细时空调制。最后,讨论异构像元结构工程(如非周期性像素排布)等新兴的焦平面编码创新策略,并分析这些方法在构建高性能、高集成度像素超分辨成像系统中所具备的独特优势。通过对上述技术的系统性梳理与深入剖析,揭示焦平面时空编码在拓展成像系统信息获取维度、突破传统采样限制以及提升像素级分辨率方面的核心科学价值与未来发展趋势。

3.2.1 基于微扫描的像素超分辨成像技术

微扫描是一种经典的时空编码调控技术,其通过在系统焦平面上引入图像间的亚像素级相对位移,并利用探测器记录不同时刻下不同亚像素采样位置的对应强度信息,从而实现时空编码的调控。这些经过时空编码的低分辨率图像序列,隐含着超越奈奎斯特采样的高频信息,为后续的超分辨重建提供关键先验信息。根据产生亚像素位移的驱动方式不同,微扫描技术通

常分为主动微扫描技术和被动微扫描技术。

主动微扫描是指利用精密扫描装置或位移平台来控制探测器或光学组件,从而产生可控且已知的亚像素位移。在这种模式下,序列图像之间的相对位移是规则且预先确定的,这种可预测性和精确控制是主动微扫描技术的显著特征。该技术的实现主要依赖高精度微扫描装置,如压电陶瓷(PZT)和微机电系统(MEMS)。这些器件主要通过直接驱动图像传感器产生微小位移,或是调控光学系统中特定元件的位置或姿态以改变光路,从而实现亚像素位移。最早在 1996 年, Sanders 等^[29]提出利用微扫描实现超分辨成像,并开发了一套型号为 Ultra-3000/D 的机载中波红外成像系统,该系统通过压电陶瓷来驱动倾斜光学系统中的一个光学透镜实现图像的微小位移。2004 年, Ben-Ezra 等^[30]开发了一种“抖动相机”系统以实现主动微扫描图像采集,如图 12(a)所示。通过微型致动器控制探测器精确快速地产生亚像素位移,将采集的低分辨率图像进行重建,实现分辨率的提升,如图 12(b)所示。2019 年, Zhang 等^[31]同样通过压电致动器控制光学系统中可移动偏心透镜实现微扫描位移,如图 12(c)所示,其将红外成像系统的分辨率提升 2 倍,如图 12(d)所示。

此外,通过倾斜玻璃平板的光学折射或偏转效应,同样可以实现亚像素位移。该方法通常是在透镜和探

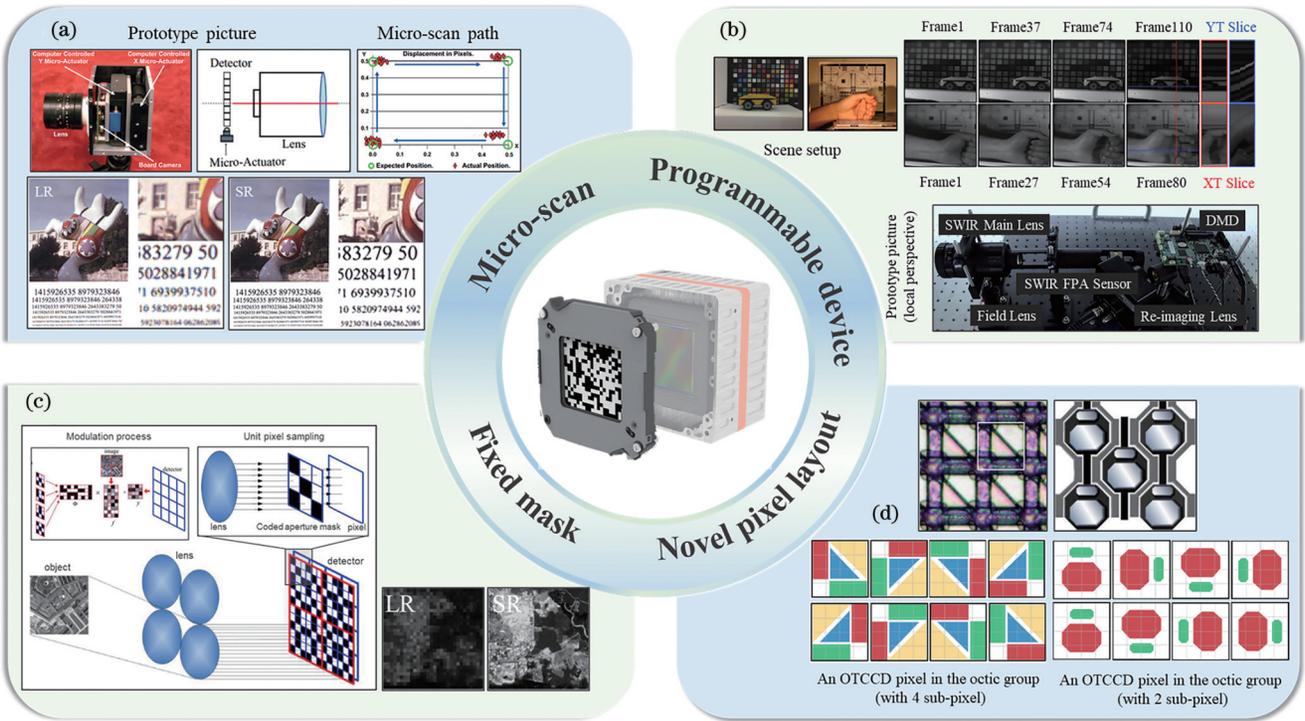


图 11 基于焦平面上空编码的方法分类及其代表性工作。(a) Ben-Ezra 等^[30]提出的基于微扫描实现超分辨成像；(b) Chen 等^[32]提出的基于可编程孔径调控实现超分辨成像；(c) Xiao 等^[33]提出的基于固定掩模调控实现超分辨成像；(d) Shi 等^[34]提出的基于异构像元探测器实现超分辨成像

Fig. 11 Classification of focal plane-based spatio-temporal encoding methods and their representative works. (a) Super-resolution imaging via micro-scanning, proposed by Ben-Ezra et al.^[30]; (b) super-resolution imaging via programmable aperture modulation, proposed by Chen et al.^[32]; (c) super-resolution imaging via fixed mask modulation, proposed by Xiao et al.^[33]; (d) super-resolution imaging via heterogeneous pixel detectors, proposed by Shi et al.^[34]

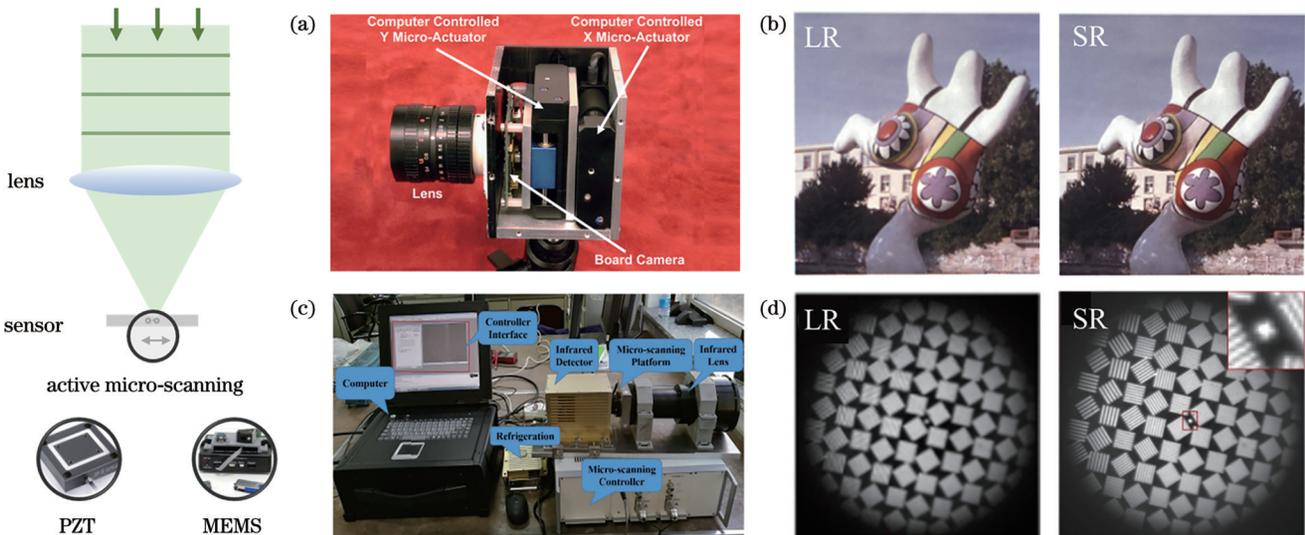


图 12 利用微扫描器件实现主动微扫描焦平面调控的像素超分辨成像技术相关工作。(a)~(b) Ben-Ezra 等^[30]提出的“抖动相机”系统原型及其超分辨前后对比结果；(c)~(d) Zhang 等^[31]提出的红外微扫描光学系统实验图及其超分辨前后对比结果

Fig. 12 Related work on pixel super-resolution imaging technology achieved through active micro-scanning focal plane control using micro-scanning devices. (a)~(b) “Jitter camera” system prototype proposed by Ben-Ezra et al.^[30] and its comparison of results before and after super-resolution; (c)~(d) experimental diagram of infrared micro-scanning optical system proposed by Zhang et al.^[31] and its comparison of results before and after super-resolution

测器之间放置一块平行的光学玻璃板(如光束移位器),通过精确控制该玻璃板的微小倾斜角度,使到达

焦平面的光束发生横向平移,无需移动传感器本身,即可在探测器上产生亚像素位移。2005年,Wiltse 等^[35]

提出一种四点微扫描技术,利用压电陶瓷移动光学元件实现微小的亚像素位移,以改进凝视红外成像的成像质量。2011年,Sui等^[36]建立了一个锗透镜倾角模型以优化平板光学元件的微扫描热成像系统性能。2013年,Sun等^[37]提出绕 z 轴旋转的倾斜光学平板以

实现光轴位移,如图 13(a)所示,从而实现亚像素信息的获取,重建结果如图 13(b)所示。2023年,Gui等^[38]提出利用 Risley 棱镜系统实现光束偏转,如图 13(c)所示。以亚像素步长进行微扫描获取低分辨率图像序列后进行超分辨率重建,结果如图 13(d)所示。

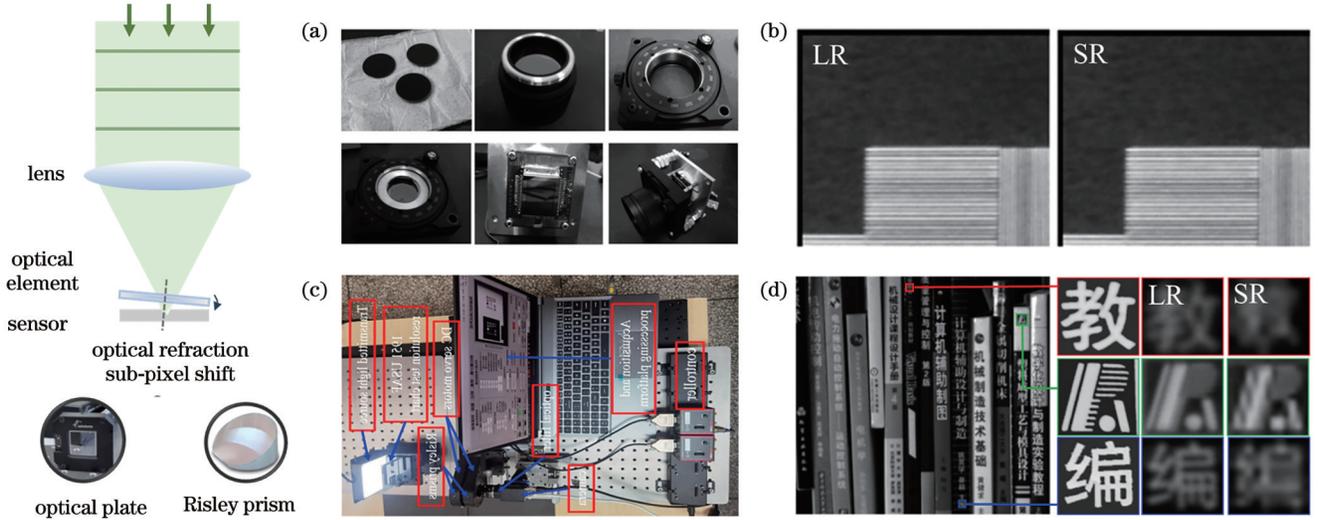


图 13 利用光学元件实现主动微扫描焦平面调控的像素超分辨率成像技术相关工作。(a)~(b) Sun 等^[37]提出的超像素扫描实验系统及其超分辨率前后对比结果;(c)~(d) Gui 等^[38]提出的基于 Risley 棱镜扫描系统及其超分辨率前后对比结果
Fig. 13 Related work on pixel super-resolution imaging technology achieved through active micro-scanning focal plane control using optical elements. (a)~(b) Super-pixel scanning experimental system proposed by Sun et al.^[37] and its comparison of results before and after super-resolution; (c)~(d) Risley prism-based scanning system proposed by Gui et al.^[38] and its comparison of results before and after super-resolution

与主动微扫描技术不同,被动微扫描 (passive micro-scanning) 是一种无需精密位移装置就可以获取亚像素信息的技术,其依赖成像过程中固有的系统抖动或因操作引入的随机扰动,从而在焦平面上产生亚像素位移。1998年,Hardie等^[39]提出一种基于机载平台的红外成像被动超分辨率方法,该方法利用飞行器振动所导致的亚像素位移,通过图像配准算法估算位移信息并将其映射至高分辨率网格,从而显著提升红外图像分辨能力。类似地,2000年 Alam 等^[40]、2014年 Dai 等^[41]同样利用移动平台的非受控振动在红外焦平面阵列上产生随机、亚像素级的位移,从而提升红外视频图像分辨率。通过被动式采集亚像素位移图像的方法随后被成功应用于消费级电子产品。2019年,Google 公司^[42]推出 Super-Res Zoom 功能,其巧妙地利用用户手持拍摄时自然产生的微小抖动,以捕获亚像素位移的图像序列。尽管此类被动微扫描方法无需额外硬件且易于集成,但其核心挑战在于所获取的亚像素位移信息本质上具有随机性和不可控性。这种特性可能导致最终重建图像的质量一致性欠佳且存在较大波动,因此,在对重建结果稳定性、可重复性以及位移精度有较高要求的超分辨率应用场景中,被动微扫描技术的适用性通常会受到限制。

除前述主要依赖单一传感器微动来获取亚像素位

移的技术路线外,另一种并行且高效获得亚像素位移的策略是利用阵列相机系统^[43]。这类系统通过集成多个低分辨率的探测单元,能够在单次曝光下同时捕获同一目标场景的多幅低分辨率图像。由于成像单元之间存在固定的物理空间位移,所采集的多帧图像之间存在亚像素级的相对偏移信息。2005年,Wilburn 等^[44]提出一种由 100 个相机组成的大型复眼阵列成像系统,如图 14(a)所示。通过并行采集目标场景图像,经配准对齐后获得一系列具有亚像素位移的低分辨率图像用于后续超分辨率重建,重建结果如图 14(b)所示。2023年,Yang 等^[45]提出一种四孔径阵列相机,如图 14(c)所示,并通过自适应正则化参数调整方法加快参数更新,用于实现多帧亚像素位移超分辨率重建,实现 2 倍的分辨率提升,如图 14(d)所示。基于多孔径的像素超分辨率技术不仅通过并行采集多视角图像显著提高图像采集效率,能够实现实时超分辨率成像,还克服单相机系统低空间带宽积的问题,未来有望应用于大视场高分辨率成像探测。

综上,基于微扫描的焦平面时空编码调控策略,其核心是在原有的混叠信息基础上,记录焦平面中多个亚像素位移后的“观测样本”。然而单纯依赖空间位移的调制方式从根本上并未改变探测器像素内部固有的混叠模式,亦未能引入额外的编码维度或约束条件以

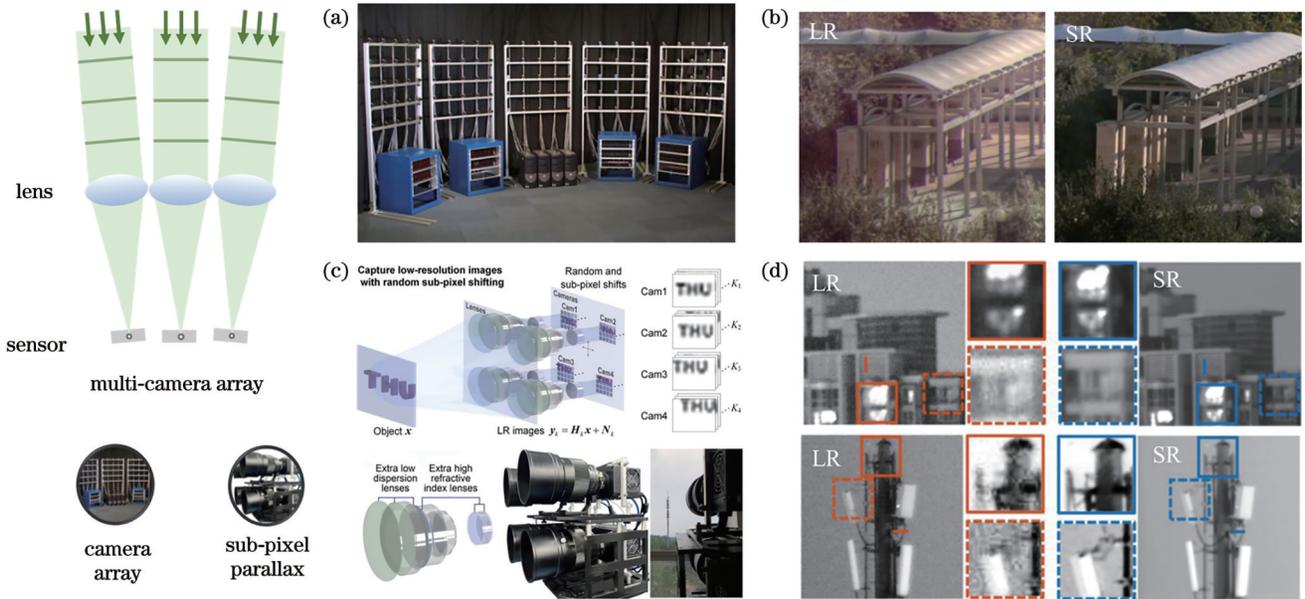


图 14 利用相机阵列实现焦平面调控的像素超分辨成像技术相关工作。(a)~(b) Wilburn 等^[44]提出的大型复眼阵列成像系统及其超分辨前后对比结果；(c)~(d) Yang 等^[45]提出的四相机阵列系统及其超分辨前后对比结果

Fig. 14 Related work on pixel super-resolution imaging technology achieving focal plane control using camera arrays. (a)~(b) Large-scale compound-eye array imaging system proposed by Bennett Wilburn et al.^[44] and its comparison of results before and after super-resolution; (c)~(d) four-camera array system proposed by Yang et al.^[45] and its comparison of results before and after super-resolution

更有效地解译这些已被混叠的高频信息,因此,单纯基于微扫描的编码策略,在像素超分辨率重建中所能达到的分辨率增益幅度通常较为有限,其提升潜力会受到上述内在机制的制约。

3.2.2 基于可编程器件调控的像素超分辨成像技术

除采用微扫描器件实现时空编码调控外,还可采用可编程器件如 DMD 或 SLM 等,将一系列预设或动态生成的空间结构化编码图案投影到系统的中间像面或与物/像共轭的平面上,实现精确的编码调控。由于调控器件的动态可编程性^[46],可以实时、逐点地改变焦平面处光场的振幅、相位等状态,对目标场景进行精细复杂的光学编码,提升成像系统整体编码信息量与编码效率。其中,将可编程器件联合压缩感知算法实现像素超分辨成像是代表性应用之一^[47-49]。该技术通过可编程器件 DMD 加载编码图案,对焦平面光场信息进行空间编码,采集远低于奈奎斯特采样要求的低分辨率图像,再根据信号稀疏先验信息,从压缩的数据中精确重建高分辨率图像。2014 年, Mahalanobis 等^[48]对中波红外实现压缩感知像素超分辨成像,如图 15(a)所示。实验通过在 DMD 中分别采用光栅扫描、随机二值编码图案进行压缩感知以提升像素分辨率,如图 15(b)所示。2015 年, Chen 等^[32]采用压缩感知技术对短波红外探测器实现超分辨成像,所提出的焦平面压缩感知(FPA-CS)系统如图 15(c)所示,采用一个高分辨率的可编程器件 DMD 显示预设的二值编码图案

进行编码调控,并用 64×64 的低分辨率红外焦平面阵列进行压缩采集,其中每个探测器像元积分强度来自 DMD 中 16×16 的特定编码区域。经过压缩感知算法重建后,实现 16 倍的分辨率提升,如图 15(d)所示。

3.2.3 基于固定编码掩模调控的像素超分辨成像技术

除采用可编程器件实现焦平面编码调控外,另一种实现像素超分辨成像的调控策略是在焦平面前方设置固定的编码掩模对入射光场进行编码。与可编程有源器件相比,其优势在于无源低功耗、系统结构简单,以及具备实时成像潜力。2012 年, Xiao 等^[33]提出通过在红外焦平面探测器阵列表面直接集成固定编码掩模,对入射光场进行多路复用压缩编码,如图 16(a)所示。其仅需采集单帧编码低分辨率图像,便可以利用压缩感知算法重建出分辨率提升 2 倍的图像,如图 16(b)所示,尤其适用于对成像速度和系统紧凑性有较高要求的场景。2017 年, Mahalanobis 等^[50]提出一种在中间像面放置固定图案编码掩模的方案,成像系统如图 16(c)所示。其创新之处在于通过精确控制该掩模在 X-Y 平面内的微小机械位移,等效地产生一系列时序变化的编码图案。在此过程中,探测器单元对透过不同掩模有效区域的光信号进行积分,后续则利用对掩模位移所记录的多帧低分辨率图像进行超分辨率重建,实现 8×8 分辨率的提升,结果如图 16(d)所示。

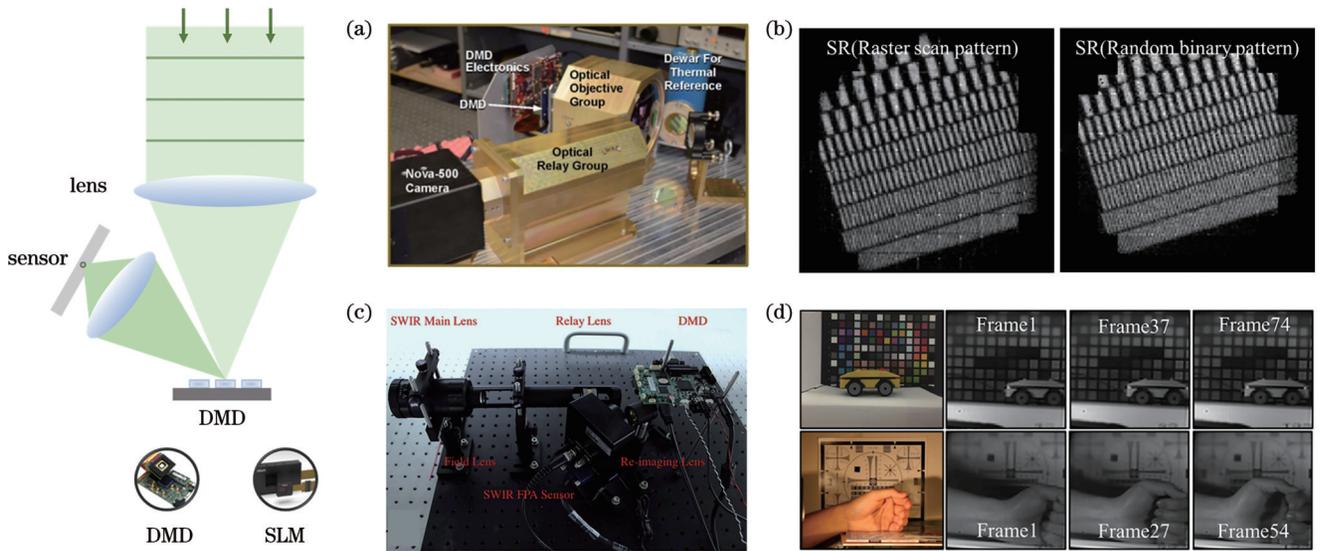


图 15 利用可编程器件实现焦平面调控的像素超分辨成像技术相关工作。(a)~(b) Mahalanobis 等^[48]提出的中波红外压缩感知成像系统的及其超分辨前后对比结果;(c)~(d) Chen 等^[32]提出的基于焦平面压缩感知系统及其超分辨前后对比结果
Fig. 15 Related work on pixel super-resolution imaging technology achieving focal plane control using programmable devices. (a)–(b) Mid-wave infrared compressive sensing imaging system proposed by Mahalanobis et al.^[48] and its comparison of results before and after super-resolution; (c)–(d) system based on focal plane array compressive sensing proposed by Chen et al.^[32] and its comparison of results before and after super-resolution

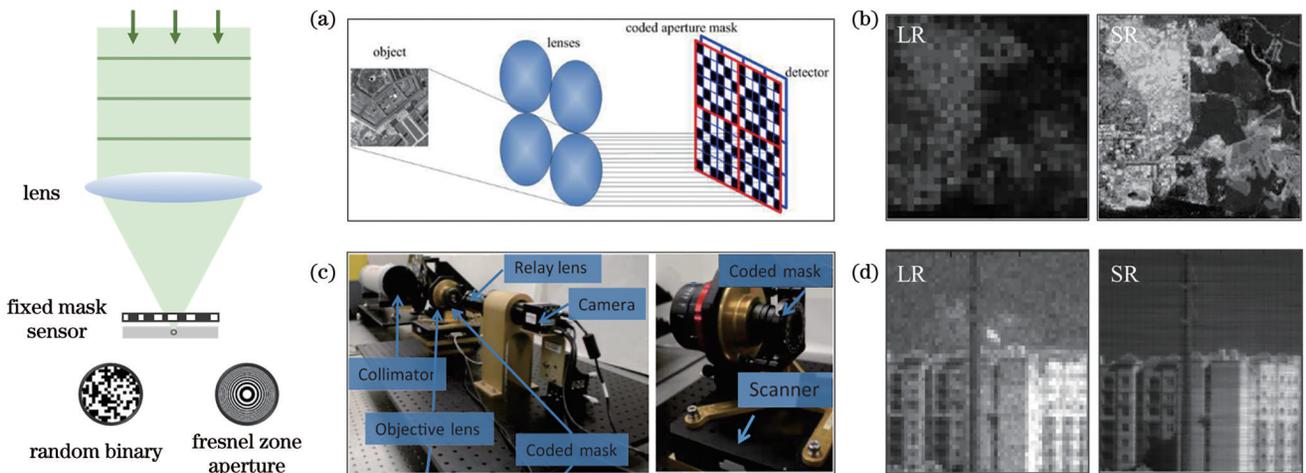


图 16 利用固定掩模实现焦平面调控的像素超分辨成像技术相关工作。(a)~(b) Xiao 等^[33]提出的基于焦平面编码孔径掩模的压缩感知红外成像系统原理及其超分辨前后对比结果;(c)~(d) Mahalanobis 等^[50]提出的基于平移编码掩模焦平面编码系统及其超分辨前后对比结果
Fig. 16 Related work on pixel super-resolution imaging technology achieving focal plane control using fixed masks. (a)–(b) Principle of compressive sensing infrared imaging system based on a focal plane coded aperture mask proposed by Xiao et al.^[33] and its comparison of results before and after super-resolution; (c)–(d) focal plane coding system based on a translating coded mask proposed by Mahalanobis et al.^[50] and its comparison of results before and after super-resolution

3.2.4 基于异构像元探测器的像素超分辨成像技术

基于片上编码的调控虽然有效提升系统的空间分辨能力,但其本质上是降低探测器的光能利用效率以换取空间维度上信息获取能力的提升。为避免引入编码掩模损失的光通量,可通过改变探测器空间采样方式来获取高信噪比的调制图像。相较于传统具有规则周期性排布的传感器,通过构建非周期平移不对称的

探测器采样网格,在图像采集过程中能够采样到更加多样的空间频率成分,为超分辨成像提供一种新的信息调控手段。

20 世纪 70 年代,数学物理学家 Roger Penrose 发现一种著名的非周期性平面镶嵌模式,即彭罗斯铺砌(Penrose tiling)。与传统的周期性铺砌(如正方形或六边形网格会规则重复自身图案)不同,彭罗斯铺砌

在无限延伸的平面上永不精确重复,因而其图案不具备平移对称性。2010年, Ben-Ezra 等^[51]采用基于彭罗斯铺砌的非周期性阵列替代传统规则的探测器像元排列,如图 17(a)所示,并提出一种新型误差反投影超分辨率算法以实现超分辨成像。这种非规则排列的像元传感器旨在克服规则采样栅格的固有局限性,其非周期性使得在亚像素位移采样过程中,每次采样更有可能捕获到该场景中此前未采集到的细节信息,重建结果如图 17(b)所示。2014年, Shi 等^[34]同样借鉴了非对称像素排布的思想,但其方案的独特之处在于:假设探测器像元宏观上仍为正方形,但每个像元内部则由满足特定八元群变换要求的、形状各异

(如近似三角形、L形等组合)的子像素单元构成,如图 17(c)所示。这些子像素的多样化形状及其非对称排布,使得当整个探测器像元进行对称群变换(如旋转或翻转等)操作时,其内部子像素的相对布局便会生成多样化的有效采样模式,从而采集到更丰富的场景信息。对比在周期网格下采集和基于异构探测器采集的低分辨率图像,采用异构像元的低分辨率图像在超分辨重建后能够提升细节分辨能力,如图 17(d)所示。值得注意的是,此类非周期异构像元探测器在实现超分辨成像的过程中,仍需多次采集低分辨率测量值,并在重构时重新映射至高分辨率周期网格中。

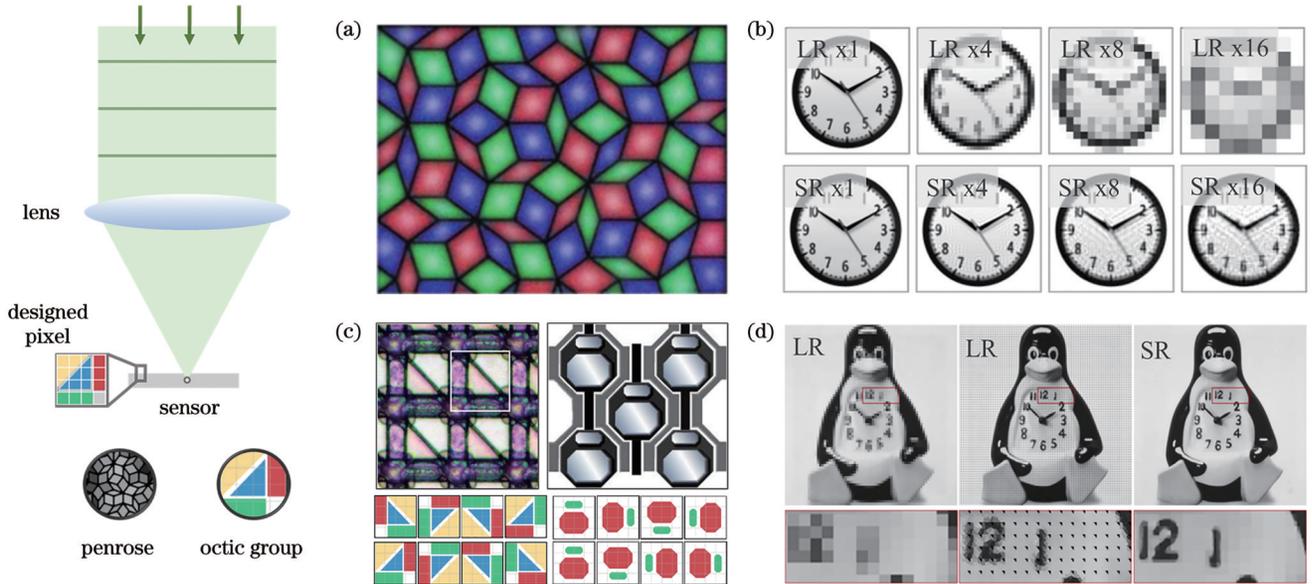


图 17 利用异构像元探测器实现焦平面调控的像素超分辨成像技术相关工作。(a)~(b) Ben-Ezra 等^[51]提出的基于彭罗斯铺砌的异构像元探测器示意图及其超分辨前后对比结果;(c)~(d) Shi 等^[34]提出的基于满足八元群变换的异构像元探测器示意图及其超分辨前后对比结果

Fig. 17 Related work on pixel super-resolution imaging technology achieving focal plane control using heterogeneous pixel detectors. (a)–(b) Schematic diagram of a heterogeneous pixel detector based on Penrose tiling proposed by Ben-Ezra et al.^[51] and its comparison of results before and after super-resolution; (c)–(d) schematic diagram of a heterogeneous pixel detector satisfying octic group transformations proposed by Shi et al.^[34] and its comparison of results before and after super-resolution

综上所述,基于焦平面时空编码的像素超分辨成像技术,无论采用何种编码形式,其核心都在于将场景中的高频信息通过特定编码的方式“嵌入”到低分辨率图像中。其中,基于可编程器件的编码调控方式赋予了成像系统更大的设计自由度与灵活性,能够产生显著且可控的亚像素级光强变化,以提升重建鲁棒性。而基于固定编码掩模的调控方式则提供一个高分辨、高实时、轻量化成像系统的设计思路。未来,基于焦平面编码能够实现像元级的光场精细调控,为构建高度集成化的智能计算超分辨传感器、实现光电成像系统的批量迭代升级奠定基础^[6]。

3.3 基于深度学习的像素超分辨成像技术

上述基于成像链路编码调控的超分辨技术侧重于

如何设计光学成像系统将场景中的高频信息有效地调控到低分辨率图像中,并从采集的编码数据中重建图像;而本节所介绍的基于深度学习的像素超分辨成像技术则侧重于如何在已知的物理模型约束下,从有限的观测数据中恢复出基于统计最优的高分辨率图像。该技术无需对光学系统进行显式的硬件调控或结构优化,主要是在成像链路后端通过大规模数据集的离线训练来实现分辨率的提升,故属于成像链路后端的数字图像阶段实现的超分辨方法。

基于深度学习的像素超分辨技术可分为两类:基于数据驱动的深度学习的像素超分辨成像技术,以及基于物理模型驱动的深度学习的像素超分辨成像技术。这两种路径代表了解决超分辨率病态问题的不同哲学思

想,即数据驱动方法通过从数据分布中学习强先验来对问题进行正则化,而物理模型驱动方法则通过成像的物理机理或施加的退化模型等先验约束来进行正则化。下面对基于两种不同驱动方式的深度学习像素超分辨技术进行介绍。

3.3.1 基于数据驱动的深度学习像素超分辨成像技术

深度学习像素超分辨成像技术发展初期主要依赖于数据驱动,利用神经网络从大量成对的低分辨率-高分辨率图像中学习复杂的映射函数。这一发展经历了从浅层网络到深层网络,从简单卷积到引入残差学习、注意力机制,再到生成对抗网络和 Transformer 架构的不断演进,如图 18 所示。在 2014 年,Dong 等^[52-53]开创性地提出图像超分辨领域的深度学习网络,即 SRCNN (Super-Resolution Convolutional Neural Network),通过浅层的卷积神经网络实现从低分辨率到高分辨率图像的复杂映射。为了解决 SRCNN 计算效率低的问题,Dong 等^[54]随后提出 FSRCNN (Fast Super-Resolution Convolutional Neural Network),通过优化网络结构和在最后进行反卷积上采样,显著提升推理速度。2016 年, Kim 等^[55]受 VGGNet^[56]的启发,所提出的 VDSR (Very Deep Super-Resolution) 通过增加网络深度并采用残差学习等方法,解决了深层网络收敛慢的问题,并将网络深度推向新的高度。直到 2017 年, Lim 等^[57]提出的 EDSR (Enhanced Deep Super-Resolution) 进一步优化残差网络的设计,取得当时最优的性能,在其基础上提出的 MDSR (Multi-Scale Deep Super-Resolution System) 用单个模型实现不同尺度的超分辨率。VDSR 的成功证明了更深的网络结合残差学习能够带来显著的性能提升。2018 年, Zhang 等^[58-59]提出 RCAN (Residual Channel Attention Network),通过引入注意力机制,进一步提升极深网络的特征学习效率。随着网络深度的不断增加,网络所能提取和处理的特征数量也变得异常庞大。此时,如何有效地利用这些特征,筛选出对最终重建结果贡献最大的信息,成为提升性能的关键。EDSR 对网络架构进行针对性优化,从而在性能和效率上都取得了进一步提升^[60-61]。这标志着超分辨领域开始朝着更加任务定制化的架构设计方向发展。

随着深度学习网络在模型深度与参数规模上逐渐显现出性能饱和的趋势,为寻求持续的性能提升,一系列创新思路与研究范式开始不断涌现^[62-64]。2017 年 Ledig 等^[65]提出 SRGAN (Super-Resolution Generative Adversarial Network),其核心在于提出一种新的感知损失函数,并使用 GAN 框架进行超分辨重建,该技术标志着超分辨领域开始关注感知质量而非仅仅是像素级失真。2018 年, Shocher 等^[66]提出首个基于 CNN (Convolutional Neural Network) 的无监督超分辨率网络 ZSSR (Zero-Shot SR),利用图像内部信息统计实现“零样本”超分。尽管 ZSSR 针对每张图像训练小型专

用 CNN 以提取内部信息,但其实际应用存在较大局限性。继 Transformer^[67]在自然语言处理(NLP)领域取得巨大成功后,因其核心的自注意力(self-attention)机制能够有效建模长距离依赖关系,而被引入图像超分辨率任务中以捕捉全局结构信息。2021 年, Liang 等^[68]提出 SwinIR (Swin Transformer for Image Restoration),通过在移位窗口内计算局部自注意力来提高计算效率,并巧妙地将分别用于浅层和深层特征提取的 CNN 网络与 Transformer 模块 RSTB (Residual Swin Transformer Block) 相结合,显著提升高分辨图像重建的质量。

基于数据驱动的超分辨率技术发展路径清晰地展现了网络结构从浅层到深层、从追求像素保真度到兼顾感知质量、从依赖局部信息到探索全局上下文的演进过程。数据驱动超分辨率的发展历程表明,更好的超分辨效果并非一个单目标优化问题,它始终在多个维度上进行权衡与演进:像素准确性(以峰值信噪比为代表)、感知真实感(以 GAN 和感知损失为代表)、计算效率(以 FSRCNN、SwinIR 为代表)以及模型复杂度与深度(以 VDSR、RCAN 为代表)。这种多维度探索的特性,使得基于数据驱动的深度学习超分辨领域保持活力,并不断催生出新的模型和方法。

3.3.2 基于物理模型驱动的深度学习像素超分辨成像技术

尽管数据驱动方法在超分辨率领域取得了巨大成功,但其“黑箱”特性、对大规模高质量训练数据的严重依赖以及在面对测试数据与训练数据分布差异较大时的低泛化性,促使研究者们探索另一条路径:将物理模型与先验知识融入深度学习超分辨率框架。纯数据驱动方法的局限性在于其可解释性较差,性能高度依赖训练数据集的规模和代表性,当测试数据与训练数据存在显著差异时,其性能可能急剧下降。引入物理模型的目的在于弥补纯数据驱动方法的“黑箱”特性和泛化能力不足的问题。数据驱动模型虽然能够学习复杂的函数映射,但其决策过程往往难以理解。当测试数据与训练数据分布一致时,它们表现优异;一旦遇到未知或复杂的降质模型,则可能产生不可预测的错误。通过将物理模型(如系统的正向模型、成像遵循的物理定律等)整合到学习框架中,可以为学习问题引入显式的结构和约束。这不仅有助于重构出与训练真值更加一致的超分辨图像,还可能因为部分知识被“注入”而非完全从数据中学习,从而降低对数据的依赖,并提升模型对未知场景的泛化能力。

2023 年, Bian 等^[69]针对单光子雪崩二极管 (SPAD) 成像提出利用精确的物理噪声模型合成大规模训练数据,并训练门控融合 Transformer 网络,通过亚像素卷积等结构实现 4 倍空间超分辨率。2025 年, Zhang 等^[70]提出的 APD-Net 框架结合了数据训练的扩散模型与单像素成像的物理模型先验,能从仅为 1%

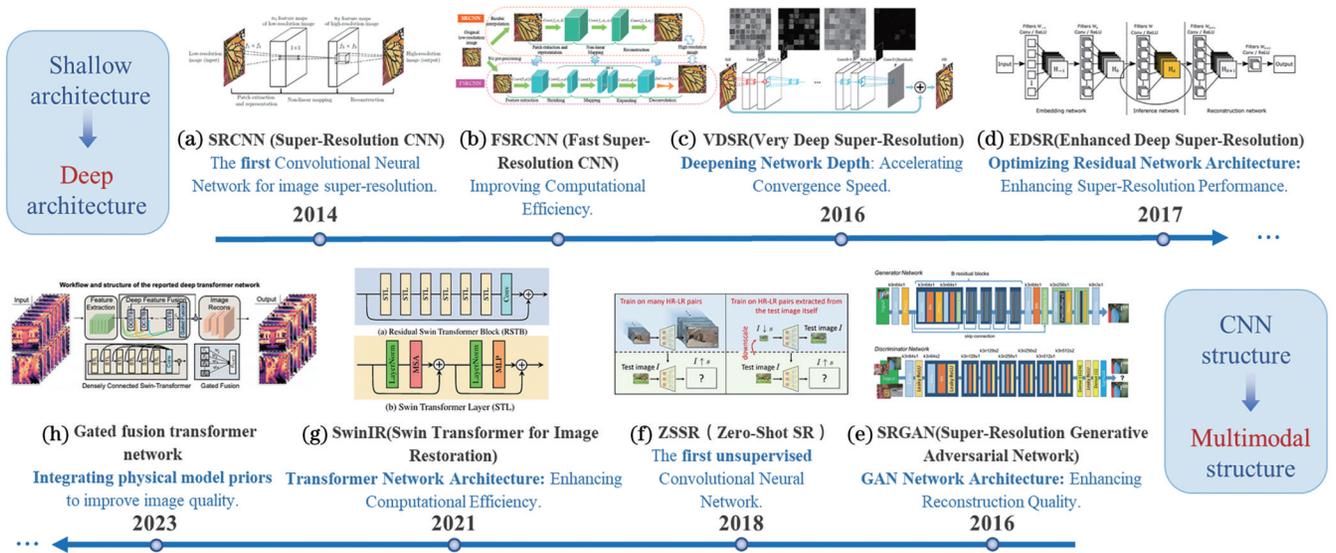


图 18 基于数据驱动深度学习像素超分辨率成像技术的发展历程。(a) Dong 等^[52-53]提出的 SRCNN 结构; (b) Dong 等^[54]提出的 FSRCNN 结构; (c) Kim 等^[55]提出的 VDSR 网络结构; (d) Lim 等^[57]提出的 EDSR 网络结构; (e) Ledig 等^[65]提出的 SRGAN 结构; (f) Shocher 等^[66]提出的 ZSSR 网络结构; (g) Liang 等^[68]提出的 SwinIR 网络结构; (h) Bian 等^[69]提出的门控融合 Transformer 网络结构

Fig. 18 Development history of data-driven deep learning pixel super-resolution imaging technology. (a) SRCNN structure proposed by Dong et al.^[52-53]; (b) FSRCNN structure proposed by Dong et al.^[54]; (c) VDSR network structure proposed by Kim et al.^[55]; (d) EDSR network structure proposed by Lim et al.^[57]; (e) SRGAN structure proposed by Ledig et al.^[65]; (f) ZSSR network structure proposed by Shocher et al.^[66]; (g) SwinIR network structure proposed by Liang et al.^[68]; (h) gated fusion Transformer network structure proposed by Bian et al.^[69]

的极低采样率数据中重建出 256×256 的高分辨率图像。

尽管基于物理模型驱动深度学习像素超分辨率技术已展现出巨大的发展趋势,但仍面临以下挑战:如何在深度学习模型中更有效地平衡物理约束与数据驱动学习,如何选择和表征最恰当的物理先验信息,以及如何提升模型的计算可扩展性、处理物理过程的不确定性与不完整性,并增强模型的可解释性。未来的研究方向可聚焦于将更复杂、多尺度、多物理场耦合的物理模型融入超分辨率网络,为设计更鲁棒和高效的物理驱动超分辨率算法提供坚实基础。

4 总结与展望

基于时空编码调控的像素超分辨率成像作为计算成像的关键分支,其核心在于通过前端物理域的光学编码调控与后端数字域的信息解耦重建,突破传统成像在像素采样与空间带宽积等方面的制约。此框架下的“像素超分辨率”致力于通过计算手段提升有效像素密度、恢复混叠的高频信息,解析乃至重构和感知光场的强度、相位、光谱、偏振、深度/角度等多维与多模态信息。未来,基于时空编码调控的像素超分辨率成像有望从以下三个方面实现突破:

1) 实时鲁棒计算超分辨率成像系统架构:计算超分辨率成像系统的未来演进,旨在构建能在动态复杂环境中实时、鲁棒运行的成像系统,其核心在于对硬件、算

法与系统的深度协同与一体化设计。在硬件层面,需突破现有编码调制器刷新率对数据采集效率的限制,转向探索可重构超表面或集成光子芯片等高速光场编码技术。在算法层面,需应对运动模糊导致物理模型失配的问题,通过系统集成运动估计与补偿模块,或发展嵌入物理先验的深度网络以增强对噪声的鲁棒性。在系统层面,为克服大气湍流与热漂移等环境扰动,必须从静态系统标定升级为动态在线校准机制,通过闭环反馈实现自适应补偿,确保系统保真度。综上,新一代成像架构正向一个感知、决策、反馈一体化的自适应系统演进,通过硬件加速、算法增强与系统自校准的协同,应对未来更严苛的成像挑战。

2) 先进材料与新型时空光场编码器件:像素超分辨率计算成像的性能提升,其核心物理支撑在于能够对光场多自由度(空间、时间、波长、偏振、相位等)实施高速、高精度、高效编码调控的新型器件与材料。一方面,超构表面^[71-73]、衍射光学元件^[74-76]等微纳光学器件,以其精细操控光场的能力,可实现超薄化、集成化的编码模块,调控成像系统的传输矩阵,为紧凑型多模态感知系统提供新路径。另一方面,基于 MEMS 的新型纯相位光调制器^[77-79](PLM)等以其高速度、高光效、宽波段特性,为动态时空编码提供关键技术支撑。这些新型器件的突破将拓展编码维度、提升信息密度与带宽,是实现更高倍率超分辨率、获取丰富多模态信息、驱动系统小型化与集成化的核心动力。

3) 深度学习编码调控联合优化方法:深度学习正深刻变革计算光学成像,不仅提升了信息重建的效率与质量,克服噪声与数据稀疏性挑战,更推动光学编码(物理层)与深度学习解码(算法层)联合优化的端到端(end-to-end)设计理念。被称为深度光学(deep optics)^[80]的技术趋势,正将人工智能从后端处理延伸至前端光学编码设计,有望自主发现并优化超越人类直观的编码策略,以最大化信息获取效率,其本质是在有限“空间带宽积”下实现更优的“信息置换”。同时,面向计算成像的“基础模型”研究,有望通过大规模预训练突破数据瓶颈,实现对新任务、新模态的快速鲁棒适应,引领像素超分辨与多维光场感知向更高阶智能化演进。

总体而言,基于时空编码调控的像素超分辨成像技术,通过物理域的灵活编码与数字域的高效解码,突破传统成像在分辨率、探测维度和信息通量上的限制。这本质上是一种信息置换或空间带宽积调控策略,在遵循物理规律的前提下,通过牺牲某些维度信息换取目标维度(如空间、光谱、时间分辨率)性能的提升或对新维度的感知。未来,“新架构、新器件、新方法”将相互促进,通过多维度信息联合编码与智能解码,进一步拓展感知边界,实现对客观世界更高维度、更精细信息的捕获,推动成像技术从“看清”向“理解”迈进。尽管在前进的道路上仍面临着数据处理、算法泛化性、系统复杂性与成本等挑战,但计算成像技术凭借其独特的“物理-计算”协同优势,必将在科学研究、工业检测、医疗诊断、消费电子等诸多领域展现出更加广阔的应用前景,并持续推动光电成像技术的迭代革新。

参 考 文 献

- [1] Klassen S. The photoelectric effect: reconstructing the story for the physics classroom[J]. Science & Education, 2011, 20(7): 719-731.
- [2] Wang B W, Zou Y, Zhang L F, et al. Multimodal super-resolution reconstruction of infrared and visible images via deep learning[J]. Optics and Lasers in Engineering, 2022, 156: 107078.
- [3] 黄郑重, 曹良才. 数字全息与定量相位成像: 进展与展望(特邀)[J]. 光子学报, 2024, 53(9): 0911001.
Huang Z Z, Cao L C. Digital holography and quantitative phase imaging: advances and prospects (invited) [J]. Acta Photonica Sinica, 2024, 53(9): 0911001.
- [4] 梁健, 巨海娟, 张文飞, 等. 偏振光学成像去雾技术综述[J]. 光学学报, 2017, 37(4): 0400001.
Liang J, Ju H J, Zhang W F, et al. Review of optical polarimetric dehazing technique[J]. Acta Optica Sinica, 2017, 37(4): 0400001.
- [5] Hellerer T, Enejder A M K, Zumbusch A. Spectral focusing: high spectral resolution spectroscopy with broad-bandwidth laser pulses[J]. Applied Physics Letters, 2004, 85(1): 25-27.
- [6] 王博文, 张许, 管海涛, 等. 基于焦平面编码调控的计算成像: 研究现状与发展趋势(特邀)[J]. 光学学报(网络版), 2024, 1(1): 0111001.
Wang B W, Zhang X, Guan H T, et al. Computational imaging based on focal plane-coded modulation: a review (invited) [J]. Acta Optica Sinica (Online), 2024, 1(1): 0111001.
- [7] 左超, 陈钱. 分辨率、超分辨率与空间带宽积拓展: 从计算光学成像角度的一些思考[J]. 中国光学(中英文), 2022, 15(6): 1105-1166.
Zuo C, Chen Q. Resolution, super-resolution and spatial bandwidth product expansion: some thoughts from the perspective of computational optical imaging[J]. Chinese Optics, 2022, 15(6): 1105-1166.
- [8] 李晟, 王博文, 管海涛, 等. 远场合成孔径计算光学成像技术: 文献综述与最新进展[J]. 光电工程, 2023, 50(10): 230090.
Li S, Wang B W, Guan H T, et al. Far-field computational optical imaging techniques based on synthetic aperture: a review [J]. Opto-Electronic Engineering, 2023, 50(10): 230090.
- [9] 左超, 陈钱. 计算光学成像: 何来, 何处, 何去, 何从? [J]. 红外与激光工程, 2022, 51(2): 20220110.
Zuo C, Chen Q. Computational optical imaging: an overview [J]. Infrared and Laser Engineering, 2022, 51(2): 20220110.
- [10] 薛家安, 裘溯, 金伟其, 等. 分焦平面偏振相机串扰的固定线性叠加模型[J]. 光学学报, 2025, 45(8): 0811002.
Xue J A, Qiu S, Jin W Q, et al. Fixed linear superposition crosstalk model of division of focal plane polarization camera [J]. Acta Optica Sinica, 2025, 45(8): 0811002.
- [11] Park S C, Park M K, Kang M G. Super-resolution image reconstruction: a technical overview [J]. IEEE Signal Processing Magazine, 2003, 20(3): 21-36.
- [12] Farsiu S, Robinson D, Elad M, et al. Advances and challenges in super-resolution [J]. International Journal of Imaging Systems and Technology, 2004, 14(2): 47-57.
- [13] Johnson W R, Wilson D W, Fink W, et al. Snapshot hyperspectral imaging in ophthalmology [J]. Journal of Biomedical Optics, 2007, 12(1): 014036.
- [14] Tran M H, Fei B W. Compact and ultracompact spectral imagers: technology and applications in biomedical imaging [J]. Journal of Biomedical Optics, 2023, 28(4): 040901.
- [15] 张润南, 周宁, 周子豪, 等. 光场表征及其分辨率提升技术: 文献综述及最新进展(特邀)[J]. 红外与激光工程, 2024, 53(9): 20240347.
Zhang R N, Zhou N, Zhou Z H, et al. Light field representation and its resolution improvement techniques: an overview (invited) [J]. Infrared and Laser Engineering, 2024, 53(9): 20240347.
- [16] Brady D J. Optical imaging and spectroscopy [M]. Weinheim: John Wiley & Sons, 2009.
- [17] 张润南, 蔡泽伟, 孙佳嵩, 等. 光场相干测量及其在计算成像中的应用 [J]. 激光与光电子学进展, 2021, 58(18): 1811003.
Zhang R N, Cai Z W, Sun J S, et al. Optical-field coherence measurement and its applications in computational imaging [J]. Laser & Optoelectronics Progress, 2021, 58(18): 1811003.
- [18] Slinger C W, Bennett C R, Dyer G, et al. An adaptive coded aperture imager: building, testing and trialing a super-resolving terrestrial demonstrator [J]. Proceedings of SPIE, 2011, 8165: 816511.
- [19] Slinger C, Bennett H, Dyer G, et al. Adaptive coded-aperture imaging with subpixel superresolution [J]. Optics Letters, 2012, 37(5): 854-856.
- [20] Yi R G, Chen X, Dong L Q, et al. An imaging method of wavefront coding system based on phase plate rotation [J]. Proceedings of SPIE, 2018, 10616: 106160W.
- [21] Li W Z, Qi J R, Alu A. Single-pixel super-resolution with a space-time modulated computational metasurface imager [J]. Photonics Research, 2024, 12(10): 2311-2322.
- [22] Zhang Q, Xu X, Guo Y H, et al. Scaled transverse translation by planar optical elements for sub-pixel sampling and remote super-resolution imaging [J]. Nanophotonics, 2025, 14(8): 1203-1211.
- [23] Gordon N T, de Villiers G D, Ridley K D, et al. An experimental infrared sensor using adaptive coded apertures for enhanced resolution [J]. Proceedings of SPIE, 2010, 7818: 781806.

- [24] 赵惠, 魏静莹, 庞志海, 等. 波前编码超分辨成像技术[J]. 红外与激光工程, 2016, 45(4): 0422003.
Zhao H, Wei J X, Pang Z H, et al. Wave-front coded super-resolution imaging technique[J]. Infrared and Laser Engineering, 2016, 45(4): 0422003.
- [25] Du H Y, Yi R G, Dong L Q, et al. Rotating asymmetrical phase mask method for improving signal-to-noise ratio in wavefront coding systems[J]. Applied Optics, 2018, 57(13): 3365-3371.
- [26] Lee D, Gwak J, Badloe T, et al. Metasurfaces-based imaging and applications: from miniaturized optical components to functional imaging platforms[J]. Nanoscale Advances, 2020, 2(2): 605-625.
- [27] Hu X M, Xu W Z, Fan Q B, et al. Metasurface-based computational imaging: a review[J]. Advanced Photonics, 2024, 6(1): 014002.
- [28] El-Helou A J, Liu Y T, Chen C H, et al. Optical metasurfaces for the next-generation biosensing and bioimaging[J]. Laser & Photonics Reviews, 2025, 19(10): 2401715.
- [29] Sanders J G, Wan W H, Harris V, et al. Compact airborne staring FPA sensor with microscanning[J]. Proceedings of SPIE, 1996, 2743: 158-168.
- [30] Ben-Ezra M, Zomet A, Nayar S K. Jitter camera: high resolution video from a low resolution detector[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004, June 27-July 2, 2004, Washington, DC, USA. New York: IEEE Press, 2004.
- [31] Zhang X F, Huang W, Xu M F, et al. Super-resolution imaging for infrared micro-scanning optical system[J]. Optics Express, 2019, 27(5): 7719-7737.
- [32] Chen H J, Asif M S, Sankaranarayanan A C, et al. FPA-CS: Focal plane array-based compressive imaging in short-wave infrared[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA. New York: IEEE Press, 2015: 2358-2366.
- [33] Xiao L L, Liu K, Han D P, et al. A compressed sensing approach for enhancing infrared imaging resolution[J]. Optics & Laser Technology, 2012, 44(8): 2354-2360.
- [34] Shi B X, Zhao H, Ben-Ezra M, et al. Sub-pixel layout for super-resolution with images in the octic group[M]//Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8689: 250-264.
- [35] Wiltse J M, Miller J L. Imagery improvements in staring infrared imagers by employing subpixel microscan[J]. Optical Engineering, 2005, 44(5): 056401.
- [36] Sui X B, Bai L F, Chen Q, et al. Influencing factors of microscanning performance based on flat optical component[J]. Chinese Optics Letters, 2011, 9(5): 052302.
- [37] Sun M J, Yu K L. A sub-pixel scan method for super-resolution reconstruction[J]. Optik, 2013, 124(24): 6905-6909.
- [38] Gui C, Wang D T, Huang X W, et al. Super-resolution and wide-field-of-view imaging based on large-angle deflection with risley prisms[J]. Sensors, 2023, 23(4): 1793.
- [39] Hardie R C, Barnard K J, Bognar J G, et al. High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system[J]. Optical Engineering, 1998, 37(1): 247-260.
- [40] Alam M S, Bognar J G, Hardie R C, et al. Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames[J]. IEEE Transactions on Instrumentation and Measurement, 2000, 49(5): 915-923.
- [41] Dai S S, Liu J S, Xiang H Y, et al. Super-resolution reconstruction of images based on uncontrollable microscanning and genetic algorithm[J]. Optoelectronics Letters, 2014, 10(4): 313-316.
- [42] Wronski B, Garcia-Dorado I, Ernst M, et al. Handheld multi-frame super-resolution[J]. ACM Transactions on Graphics, 2019, 38(4): 28.
- [43] 曹铭智, 王博文, 齐静雅, 等. 复眼相机阵列动态全景图像合成技术(特邀)[J]. 激光与光电子学进展, 2024, 61(16): 1611016.
Cao M Z, Wang B W, Qi J Y, et al. Dynamic panoramic image synthesis technology using compound eye camera array (invited)[J]. Laser & Optoelectronics Progress, 2024, 61(16): 1611016.
- [44] Wilburn B, Joshi N, Vaish V, et al. High performance imaging using large camera arrays[J]. ACM Transactions on Graphics, 2005, 24(3): 765-776.
- [45] Yang F, Wu J C, Gao Y H, et al. A four-aperture super-resolution camera based on adaptive regularization parameter tuning[J]. Optics and Lasers in Engineering, 2023, 165: 107562.
- [46] Zhang X, Wang B W, Li S, et al. Lensless imaging with a programmable Fresnel zone aperture[J]. Science Advances, 2025, 11(12): eadt3909.
- [47] Zhang L X, Ke J, Chi S, et al. High-resolution fast mid-wave infrared compressive imaging[J]. Optics Letters, 2021, 46(10): 2469-2472.
- [48] Mahalanobis A, Shilling R, Murphy R, et al. Recent results of medium wave infrared compressive sensing[J]. Applied Optics, 2014, 53(34): 8060-8070.
- [49] Wu Z M, Wang X. Focal plane array-based compressive imaging in medium wave infrared: modeling, implementation, and challenges[J]. Applied Optics, 2019, 58(31): 8433-8441.
- [50] Mahalanobis A, Shilling R, Muise R, et al. High-resolution imaging using a translating coded aperture[J]. Optical Engineering, 2017, 56(8): 084106.
- [51] Ben-Ezra M, Lin Z C, Wilburn B, et al. Penrose pixels for super-resolution[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(7): 1370-1383.
- [52] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[M]//Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8692: 184-199.
- [53] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [54] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer International Publishing, 2016, 9906: 391-407.
- [55] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [56] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2025-02-05]. <https://arxiv.org/abs/1409.1556>.
- [57] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [58] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2472-2481.
- [59] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018.

- Lecture notes in computer science. Cham: Springer, 2018, 11211: 294-310.
- [60] Yang W M, Zhang X C, Tian Y P, et al. Deep learning for single image super-resolution: a brief review[J]. IEEE Transactions on Multimedia, 2019, 21(12): 3106-3121.
- [61] Wang Z H, Chen J, Hoi S C H. Deep learning for image super-resolution: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3365-3387.
- [62] Chen W W, Feng S J, Yin W, et al. Deep-learning-enabled temporally super-resolved multiplexed fringe projection profilometry: high-speed kHz 3D imaging with low-speed camera [J]. Photonix, 2024, 5(1): 25.
- [63] Feng S J, Xiao Y L, Yin W, et al. Fringe-pattern analysis with ensemble deep learning[J]. Advanced Photonics Nexus, 2023, 2(3): 036010.
- [64] Wang B W, Chen W W, Qian J M, et al. Single-shot super-resolved fringe projection profilometry (SSSR-FPP): 100, 000 frames-per-second 3D imaging with deep learning[J]. Light: Science & Applications, 2025, 14: 70.
- [65] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21–26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [66] Shocher A, Cohen N, Irani M. Zero-shot super-resolution using deep internal learning[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18–23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3118-3126.
- [67] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[EB/OL]. (2017-06-12)[2024-12-12]. <https://arxiv.org/abs/1706.03762>.
- [68] Liang J Y, Cao J Z, Sun G L, et al. SwinIR: image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11–17, 2021, Montreal, BC, Canada. New York: IEEE Press, 2021: 1833-1844.
- [69] Bian L H, Song H Z, Peng L T, et al. High-resolution single-photon imaging with physics-informed deep learning[J]. Nature Communications, 2023, 14: 5902.
- [70] Zhang Y F, Li Y X, Liu Z H, et al. Physics and data-driven alternative optimization enabled ultra-low-sampling single-pixel imaging[J]. Advanced Photonics Nexus, 2025, 4(3): 036005.
- [71] Hu J, Bandyopadhyay S, Liu Y H, et al. A review on metasurface: from principle to smart metadevices[J]. Frontiers in Physics, 2021, 8: 586087.
- [72] Zaidi A, Rubin N A, Meretska M L, et al. Metasurface-enabled single-shot and complete Mueller matrix imaging[J]. Nature Photonics, 2024, 18(7): 704-712.
- [73] Gopakumar M, Lee G Y, Choi S, et al. Full-colour 3D holographic augmented-reality displays with metasurface waveguides[J]. Nature, 2024, 629(8013): 791-797.
- [74] Orange-Kedem R, Nehme E, Weiss L E, et al. 3D printable diffractive optical elements by liquid immersion[J]. Nature Communications, 2021, 12: 3067.
- [75] Zhang Q, He Z H, Xie Z W, et al. Diffractive optical elements 75 years on: from micro-optics to metasurfaces[J]. Photonics Insights, 2023, 2(4): R09.
- [76] Khonina S N, Kazanskiy N L, Butt M A. Exploring diffractive optical elements and their potential in free space optics and imaging- a comprehensive review[J]. Laser & Photonics Reviews, 2024, 18(12): 2400377.
- [77] Ouyang B, Gong S, Lawrence T, et al. Evaluating texas instruments phase light modulator (PLM) [J]. Proceedings of SPIE, 2022, 12118: 121180I.
- [78] Rocha J C A, Wright T, Bütaitė U G, et al. Fast and light-efficient wavefront shaping with a MEMS phase-only light modulator[J]. Optics Express, 2024, 32(24): 43300-43314.
- [79] Kilcullen P, Liu J D, Wang Y M, et al. High-speed complex field modulation using binary phase-engraved superpixels[J]. Light: Advanced Manufacturing, 2025, 6: 17.
- [80] Wetzstein G, Ozcan A, Gigan S, et al. Inference in artificial intelligence with deep optics and photonics[J]. Nature, 2020, 588(7836): 39-47.

Pixel Super-Resolution Imaging Based on Spatio-Temporal Encoding Modulation: Current Status and Future Trends (Invited)

Liang Kunyao^{1,2,3}, Zhang Xu^{1,2,3}, Pei Zihao^{1,2,3}, Li Hongchun^{1,2,3}, Liu Xin^{1,2,3}, Wang Bowen^{1,2,3**},
Chen Qian³, Zuo Chao^{1,2,3*}

¹Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science & Technology, Nanjing 210094, Jiangsu, China;

²Institute of Intelligent Imaging, Nanjing University of Science & Technology, Nanjing 210019, Jiangsu, China;

³Jiangsu Key Laboratory of Visual Sensing & Intelligent Perception, Nanjing 210094, Jiangsu, China

Abstract

Significance Traditional imaging technology, based on direct intensity detection, is constrained by the intrinsic mechanisms of the photoelectric effect and detector fabrication processes. This presents significant bottlenecks in achieving key performance indicators such as high resolution, high sensitivity, and a high signal-to-noise ratio. In particular, the spatial resolution and space-bandwidth product of the acquired image have long been limited by the detector's pixel size and physical device scale. In recent years, computational imaging has emerged as a new paradigm that integrates front-end spatio-temporal encoding and modulation with back-end digital computational inversion. It offers innovative solutions to

overcome the numerous limitations of traditional imaging techniques and represents the future direction of advanced optical imaging. We systematically review the research progress of pixel super-resolution imaging techniques based on spatio-temporal encoding. Focusing on modulation strategies at the aperture and focal planes, we elucidate the mechanisms of high-dimensional light field modulation and their corresponding back-end image reconstruction algorithms. These two aspects work in synergy to break the detector's Nyquist sampling limit, enabling the effective decoupling and restoration of sub-pixel spatial information. While maintaining system compactness and high light-throughput efficiency, these technologies significantly enhance the spatial resolution and space-bandwidth product of the imaging system, providing novel solutions and key technological support for demanding applications such as long-range, wide-field-of-view, and high-precision optoelectronic detection and recognition in complex environments.

Progress Modulation of amplitude or phase at the aperture plane is one of the key technical paths to achieving pixel super-resolution imaging. By introducing a programmable or fixed mask pattern at the aperture, the wavefront information is modulated, thereby changing the system's point spread function. This allows high-frequency information to be encoded in a specific manner and multiplexed within the sampling process of multiple low-resolution frames. Combined with reconstruction algorithms, a high-resolution image that surpasses the Nyquist sampling limit can be recovered from this encoded low-resolution data, effectively enhancing the spatial resolution of the imaging system. Compared to amplitude encoding, phase encoding is a modulation scheme with a higher signal-to-noise ratio. A phase mask modulates the wavefront phase by introducing spatially varying optical path differences without attenuating the incident light energy. Therefore, it can achieve a higher image signal-to-noise ratio under the same conditions as amplitude encoding. Metasurfaces, with their sub-wavelength unit structures, leverage specific physical mechanisms to precisely and discretely manipulate the phase, amplitude, and polarization of the light field, enabling wavefront modulation at an ultra-thin scale that is difficult for traditional optical elements to achieve. Their essence is to directly expand the available degrees of freedom of the system at the aperture plane, reallocating the transmission channel capacity within the framework of "space-bandwidth product control" to realize wavefront encoding-decoding multiplexing. Since the size of their unit structures is comparable to or even smaller than a detector pixel, these devices are a natural fit for encoding strategies that enhance spatial resolution, such as "pixel-level encoding" and "sub-pixel displacement," providing a novel hardware platform for breaking the Nyquist sampling limit. The aforementioned encoding devices at the aperture plane, whether traditional amplitude/phase encoders or new devices like metasurfaces, all offer feasible paths to increase the effective Nyquist sampling rate and reduce system volume, holding the promise of further unleashing the potential of pixel super-resolution within future spatio-temporal encoding computational imaging frameworks.

In the computational optical imaging chain, the focal plane is the critical location where the photodetector performs photoelectric conversion, transforming an optical image into a digital signal. Through fine-grained modeling and control of the detector's spatial sampling characteristics, spatio-temporal encoding techniques at the focal plane can encode and reconstruct sub-pixel level light field information via on-chip and pixel-level modulation at the end of the imaging chain. Classic focal-plane spatio-temporal encoding techniques for pixel super-resolution include micro-scanning. This technique introduces sub-pixel relative displacements between images on the system's focal plane and uses the detector to record the corresponding intensity information at different sub-pixel sampling positions at different time, thereby achieving spatio-temporal modulation. This sequence of spatio-temporally encoded low-resolution images implicitly contains high-frequency information beyond the Nyquist sampling limit, providing crucial prior information for subsequent super-resolution reconstruction. Based on the driving method used to generate the sub-pixel shifts, micro-scanning is typically categorized into active and passive techniques. Beyond the technical route that relies on the micro-motion of a single sensor to acquire sub-pixel displacements, another parallel and efficient strategy is the use of camera array systems. By integrating multiple low-resolution detector units, these systems can simultaneously capture multiple low-resolution images of the same scene in a single exposure. Due to the fixed physical spatial offset between the imaging units, the captured frames contain inherent sub-pixel relative shift information. In addition to using micro-scanning devices, spatio-temporal encoding can also be achieved using programmable devices such as digital micromirror devices (DMDs) or spatial light modulators (SLMs). These devices project a series of pre-set or dynamically generated spatially structured encoding patterns onto an intermediate image plane or a plane conjugate to the object/image, enabling precise encoding modulation. The dynamic programmability of these devices allows for real-time and point-by-point changes to the amplitude and phase of the light field at the focal plane, enabling sophisticated optical encoding of the target scene and increasing the overall information content and efficiency of the imaging system's encoding.

Alternatively, a fixed encoding mask can be placed in front of the focal plane to encode the incident light field. Compared to programmable active devices, it is passive and low-power, has a simpler system structure, and possesses the potential for real-time imaging. However, while on-chip encoding modulation can effectively enhance the system's spatial resolution, it inherently does so by reducing the detector's light energy efficiency in exchange for improved information

acquisition in the spatial dimension. To avoid the light-throughput loss introduced by encoding masks, modulated images with a high signal-to-noise ratio can be obtained by altering the detector's spatial sampling method. Compared to traditional sensors with a regular and periodic layout, constructing a detector with a non-periodic and translationally asymmetric sampling grid allows for the sampling of more diverse spatial frequency components during image acquisition, providing a new means of information modulation for super-resolution imaging.

Conclusions and Prospects Overall, pixel super-resolution imaging technology based on spatio-temporal encoding and modulation breaks the limitations of traditional imaging in resolution, sensing dimensions, and information throughput by leveraging flexible encoding in the physical domain and efficient decoding in the digital domain. This is essentially a strategy of “information exchange” or “space-bandwidth product control,” where, in adherence to physical laws, information in certain dimensions is sacrificed to enhance performance in target dimensions or to enable the perception of new dimensions. Looking forward, “new architectures, new devices, and new methods” will mutually reinforce one another. Through joint multi-dimensional information encoding and intelligent decoding, they will further expand the boundaries of perception, enabling the capture of higher-dimensional and finer information about the objective world and advancing imaging technology from “seeing clearly” to “understanding”. Although challenges in data processing, algorithm generalizability, system complexity, and cost still lie ahead, computational imaging, by virtue of its unique “physics-computation” synergy, is certain to exhibit ever-broader application prospects in numerous fields, including scientific research, industrial inspection, medical diagnostics, consumer electronics, and even national security, and it will continue to drive the iterative innovation of optoelectronic imaging technology.

Key words computational imaging; pixel super-resolution; spatio-temporal encoding modulation; optoelectronic detection