



单像素高效感知方法概述

边丽衡 詹昕蕊 王华依 刘海燕 索津莉

Overview of efficient single-pixel sensing methods

Bian Liheng, Zhan Xinrui, Wang Huayi, Liu Haiyan, Suo Jinli

在线阅读 View online: <https://doi.org/10.3788/IRLA20220231>

您可能感兴趣的其他文章

Articles you may be interested in

噪声环境下基于单像素成像系统和深度学习的目标识别方法

Target recognition method based on single-pixel imaging system and deep learning in the noisy environment

红外与激光工程. 2020, 49(6): 20200010 <https://doi.org/10.3788/IRLA20200010>

用于光子计数单像素成像的去块状采样网络

Deblocking sampling network for photon counting single-pixel imaging

红外与激光工程. 2021, 50(12): 20210724 <https://doi.org/10.3788/IRLA20210724>

基于光电混合神经网络的单像素快速运动物体分类 (特邀)

Single-pixel fast-moving object classification based on optical-electronical hybrid neural network (Invited)

红外与激光工程. 2021, 50(12): 20210856 <https://doi.org/10.3788/IRLA20210856>

单像素成像理论与方法 (特邀)

Theory and approach of single-pixel imaging (Invited)

红外与激光工程. 2021, 50(12): 20211061 <https://doi.org/10.3788/IRLA20211061>

多阶段深度学习单帧条纹投影三维测量方法

Multi-stage deep learning based single-frame fringe projection 3D measurement method

红外与激光工程. 2020, 49(6): 20200023 <https://doi.org/10.3788/IRLA20200023>

基于深度学习的压缩光子计数激光雷达

Compressed photon-counting laser radar based on deep learning

红外与激光工程. 2020, 49(S2): 20200380 <https://doi.org/10.3788/IRLA20200380>

单像素高效感知方法概述

边丽衡^{1*}, 詹昕蕊¹, 王华依¹, 刘海燕¹, 索津莉^{2*}

(1. 北京理工大学 前沿交叉科学研究院 & 信息与电子学院, 北京 100081;
2. 清华大学 自动化系, 北京 100084)

摘要: 资源受限平台的高效率视觉感知是信息领域的瓶颈难题。不同于传统阵列探测成像, 单像素成像基于压缩感知原理将多维图像编码为一维采集数据, 有效提升了数据压缩率, 且灵敏度高、工作波段宽, 逐渐成为研究热点。然而, 单像素成像重建的图像中仍包含大量对高层语义理解无关的信息, 导致传输、存储、计算的资源浪费。单像素感知是一种直接从一维采集数据解耦高级语义推断结果的新型感知技术, 无需重建多维图像, 相较传统先成像-后感知的技术路径大幅提升了感知效率, 在遥感探测、智慧交通、生物医学、国防军事等众多领域具有广阔的应用前景。文中重点梳理了单像素感知技术的发展历程, 详细介绍了单像素感知技术的方法架构以及在视觉应用中的研究进展, 最后对其未来发展趋势进行了展望。

关键词: 单像素感知; 免成像感知; 深度学习; 调制优化; 联合优化

中图分类号: O438 **文献标志码:** A **DOI:** 10.3788/IRLA20220231

Overview of efficient single-pixel sensing methods

Bian Liheng^{1*}, Zhan Xinrui¹, Wang Huayi¹, Liu Haiyan¹, Suo Jinli^{2*}

(1. Advanced Research Institute of Multidisciplinary Science & School of Information and Electronics,
Beijing Institute of Technology, Beijing 100081, China;
2. Department of Automation, Tsinghua University, Beijing 100084, China)

Abstract: Efficient sensing on resource-limited platforms is a hot research topic in the field of information processing. Different from conventional array image acquisition, single-pixel data recording and compressed sensing-based image reconstruction effectively reduce the bandwidth, but the reconstructed images generally contain many data irrelevant for high-level vision tasks. Single-pixel sensing is an emerging technique that directly infers high-level semantics from one-dimensional encoded measurements without multidimensional image reconstruction. Compared with the conventional first-reconstruction-then-perception scheme, the sensing efficiency is greatly improved. It has broad applications in many fields, such as remote sensing, intelligent transportation, biomedicine, and the national defense military. This overview focuses on the history and development of single-pixel sensing and introduces its technical architecture and research progress in computer vision applications. Finally, we outlook the development trends, hoping to provide some highlights for future studies in this direction.

Key words: single-pixel sensing; image-free sensing; deep learning; modulation optimization; joint optimization

收稿日期: 2022-03-31; 修订日期: 2022-04-27

基金项目: 国家重点研发计划 (2020AA0108202, 2020YFB0505601); 国家自然科学基金 (62131003, 61971045, 61991451)

通讯作者: 边丽衡, 男, 副教授, 博士, 研究方向为计算成像与感知。

索津莉, 女, 副教授, 博士, 研究方向为计算影像学、计算机视觉。

0 引言

传统光学成像技术使用阵列探测器件(如 CCD 或 CMOS 探测器)采集光信号,从而获取目标场景的二维图像。然而,在存储、通信和计算资源紧缺的空基遥感、气象观测等领域^[1-2],传统光学成像技术往往因资源受限而难以应用。以气象观测为例,为及时防御台风等极端气候,亟需实时数据的采集、传输与分析。然而,视觉成像数据量巨大,不仅对观测系统造成较大存储与传输带宽压力,且海量数据处理也对计算平台提出了极高要求。面对上述需求,如何在带宽、算力不足等资源受限条件下实现目标场景信息的快速采集与稳定传输,并高效提取数据中的有效信息,是研究人员一直在探究的重要方向。

自然场景图像中存在大量对目标场景的特征提取与语义感知无用的非目标区域。这些非目标区域信息密度较低,不但占用了大量的成像、通信资源(例如大规模高灵敏度传感器阵列、高带宽通信链路等)和计算资源(例如去噪、去模糊等重建算法),甚至会对后续高层语义信息提取过程产生干扰,降低感知精度。因此,为减轻数据传输压力,研究人员基于自然图像的信息冗余性提出了基于阵列探测的压缩编码成像技术。此类方法在时间^[3-6]、光谱^[7-8]等维度进行数据压缩,虽然在一定程度上提升了信息密度,但针对高层视觉任务信息冗余度仍较高,信息感知效率有待提升。此外,在阵列探测过程中光信号的频率、相位、偏振等高维信息丢失,进一步降低了感知维度和效率。综上,从感知效率的角度,传统的“先成像-后感知”模式在多数情况下并非机器智能的最佳选择。

近年来,随着计算成像技术的蓬勃发展,使用单像素探测器对目标光场进行压缩采集得到了越来越多的关注和应用^[9-14],实现了空间信息的高效压缩。与传统的阵列探测器相比,单像素探测器具有以下优势^[15-22]:(1)灵敏度高,适用于弱光场景及远距离场景的成像^[23-25];(2)光谱响应范围广,在不可见光波段的优势明显,成本低廉;(3)采样速度快,能够对高速变化场景进行信息采集^[26-27]。由于单像素探测器每次采样只输出一个电信号,缺少空间分辨率,因此单像素成像需要采用一系列掩膜图案(即调制掩膜,在主动光照模式下又可称作投影图案)调制目标光场,通

过不断变化的编码调制并累积关联信息,最终使用压缩感知重建算法^[28-32]对目标场景图像进行恢复。

单像素成像需要多次测量,时间分辨率较差。全视场的高精度重构难以同时实现较高的时间分辨率与图像清晰度。根据应用需求,近年来发展出了局部区域的高分辨率单像素成像技术:(1)基于调制掩膜优化,通过设计“类视网膜”的掩膜,实现聚焦重要区域的自适应成像^[33-35];(2)局部区域直接重构,弱化不重要的背景区域信息,实现图像中特定物体的重建^[36],进一步提高了采样效率。然而,这些非目标背景区域的获取仍然造成采集、传输、存储、计算等资源的浪费。

最近,研究人员逐渐聚焦直接从少量的编码测量信息进行任务决策,避免冗余图像的获取与重建,从而进一步提高感知效率。基于与原有单像素成像结构类似的光学系统,场景中的目标特征信息被压缩编码为一维光强度信号,并被作为输入由单像素高效感知技术直接输出相应的语义理解结果,如场景目标的直接识别与追踪等。根据任务需求,对应的高效感知算法将单像素测量值作为特征提取数据进行计算。具体地,该高效感知方法如图 1 所示。与传统成像方法相比,这种高效感知技术无需常规的图像采集、重建过程,直接处理一维测量值,具有更高的数据利用率。

单像素高效感知方法从一维耦合数据中直接提

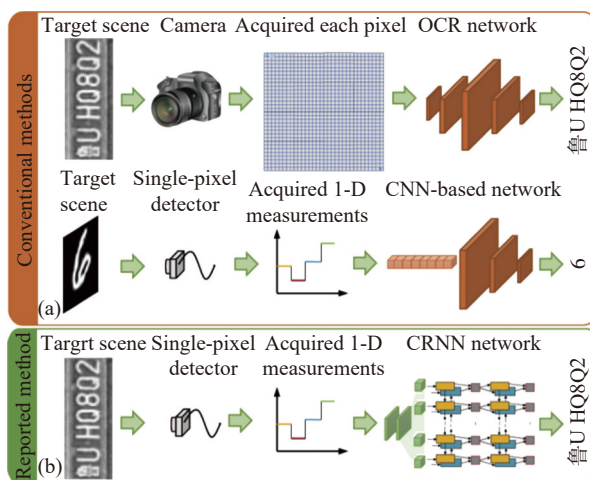


图 1 (a) 传统单像素成像方法(包括阵列传感器扫描成像及单像素成像/分类);(b) 单像素高效感知方法^[37]

Fig.1 (a) Conventional single-pixel imaging method (including array sensor scanning imaging and single-pixel imaging/classification); (b) Single-pixel high-efficiency sensing method^[37]

取感知信息,在低维特征空间中克服了高维数据特征提取时信息冗余、计算复杂的技术瓶颈,实现了高效的高级视觉语义推断。如何精确地压缩并感知场景的语义信息,以及如何提取场景特征信息进行语义计算,是单像素高效感知方法的研究重点与难点。文中将从方法架构及相关算法具体介绍单像素高效感知。

1 单像素高效感知方法架构

单像素高效感知是一种直接从一维采集数据解耦高级语义推断结果的新型感知技术,无需传统的图像获取和重建过程。在实际应用中,单像素感知系统首先利用调制掩膜对目标场景进行光调制,接着使用单像素探测器采集相应的光照强度,并将采集到的一维测量数据输入到感知算法中,即可直接得到高级语义推断结果。

研究人员通过多种方案实现单像素感知算法的设计,文中主要从调制掩膜结构和语义推断算法的设计两个过程分别进行介绍。

1.1 调制掩膜结构设计

单像素感知采用一系列掩膜来调制目标光场。为了达到更好的感知效果,研究人员围绕调制掩膜的设计开展了许多研究,相继提出了随机掩膜、哈达玛掩膜、傅里叶掩膜和网络优化的掩膜^[38]。

具体地,随机掩膜是以随机方式生成服从随机统计分布的矩阵序列^[39-40],具有随机性。参考文献^[41-49]提出利用该随机掩膜进行光场调制。随机掩膜因为不满足正交性,往往需要更多的采样次数。但是,基于随机掩膜调制的单像素成像系统对噪声不敏感,因此对于噪声影响较大的场景较为适用。哈达玛调制掩膜是由哈达玛变换基构成的二值正交矩阵^[50],是广义傅里叶变换的一种形式。参考文献^[51-55]基于哈达玛掩膜进行多次测量。由于存在负值,实际实验通常需要进行差分测量。但是,同一变换域的基底图案是正交的,所以可以减少数据采集中的冗余。傅里叶掩膜调制由 Zhang^[56]等人提出,该掩膜由不同频率不同初始相位的二维条纹图加权叠加表示,形成的多张条纹图案称为傅里叶基掩膜。参考文献^[57-63]利用傅里叶掩膜提取信息。相比于随机掩膜矩阵,哈达玛基和傅里叶基构造的掩膜矩阵效率要高很多。

调制掩膜也可作为网络的一部分进行训练,与后

续网络一起学习达到共同最优的结果,参考文献^[37, 64-70]将调制掩膜设计为全连接层权重或者卷积核与后续网络一起训练进行共同优化。相比于数值固定的掩膜设计,网络优化的掩膜可以在网络多次训练中优化参数,提高效率,但也将导致网络参数激增,容易造成过拟合。

1.2 语义推断算法设计

语义推断算法将采集到的一维测量数据作为输入,直接输出相应的高级语义推断结果。研究人员主要通过模型驱动的机器学习和数据驱动的深度学习这两类方法进行语义推断算法的设计。其中,模型驱动的机器学习算法分别基于分类器(如最大似然分类器^[41]、支持向量机^[42-43]、高斯分类器^[46]等)以及基于不同测量值距离(如哈希算子^[61]、欧氏距离^[45]、投影距离^[54-55,60,62]等)进行高效感知。随着数据驱动深度学习的不断发展,研究人员相继提出了新型网络框架来实现语义推断算法的设计,如在分类任务中,运用传统的多步训练网络^[47-48,71]与联合调制掩膜优化的端到端的网络(如 EfficientNet 网络^[38]、全连接网络^[65,69]、Transformer^[68]等),实现高效感知。进一步地,针对不同采样率任务,运用自适应的动态网络^[66-67]可以有效减少计算与存储资源消耗。在识别任务中,研究人员运用 LSTM^[37]网络能够理解多个对象特征序列之间的内部联系。在分割任务中,运用 U-Net++^[70]网络能够整合不同尺度特征图提升感知精度。

下文详细介绍上述调制掩膜结构设计和对应语义推断算法及相关工作。

2 单像素高效感知方法

在高效率感知的策略中,由于许多应用聚焦于对图像做出决策,而并非计算重建,因此仅需要分析测量值即可。对于一维测量值的直接分析,有较为成熟的算法,例如雷达检测技术、多普勒测距技术等。该类方法通过主动发射电磁波、声波等探测信号,接收经由目标物体调制后的信号,再对信号进行分析处理,最终求得场景的有效信息,如场景中目标的识别、分类、定位、检测等,无需完全成像。

单像素高效感知算法近年来也分别在模型驱动机器学习算法及数据驱动深度学习算法中逐步发展。其中,模型驱动机器学习算法多以调制掩膜与一

维测量值的相互匹配,或是一维测量值之间的直接关系来提取有效信息;数据驱动深度学习算法则侧重于利用感测能力较好的网络,或是优化调制掩膜联合相应网络的策略,进行高级语义感知。

文中将以任务为导向,分别基于数据驱动机器学习算法与数据驱动深度学习算法,依次介绍单像素高效感知方法应用于场景中目标的分类、识别、分割、追踪任务。

2.1 分类

分类问题一直是场景特征研究中的重点,目标分类在遥感、军事防御等多领域中都得到了广泛应用,如何实现快速、高精度的目标分类是研究的重点问题。

2.1.1 模型驱动的机器学习方法

基于模型驱动的机器学习算法主要通过研究不同采集数据之间的数值关系实现“免成像”感知。评判一维耦合测量序列相似度的指标主要包括距离、关联性、相似性等。下文中将以典型的应用为例,说明如何利用单像素测量值之间的差值或距离关系,进一步结合机器学习的模型及相关算法实现高效感知。

2.1.1.1 基于分类器模型估计的单像素感知

最大似然分类 (maximum likelihood classification) 方法针对两类或多类判决问题,用统计方法根据最大似然比的贝叶斯判决准则法建立非线性判别函数集,并假定各类分布函数为正态分布,在此基础上选择训练区,计算各个待分类样区的归属概率,最终进行分类。在分类时,该分类器不仅考虑了待分类样本到已知类别中心的距离,而且还考虑了已知类别的分布特征,所以其分类精度高于最小距离分类法。

在 2007 年初期,莱斯大学^[41]提出了一种算法来支持一种新的压缩分类理论,该理论与压缩感知具有相同的优势。该方法基于广义最大似然分类器 (GMLC),适用于各种信号的分类问题。针对图像分类问题,该文利用具有不同成像参数 (平移、尺度、视角、光照等) 的固定场景的图像集在高维环境图像空间中形成低维、非线性流形^[72]。同时,利用随机掩膜还设计了一种伪随机测量方案和一种新的分类算法,将生成的降维匹配滤波器称为粉碎过滤器,该过滤器为经典匹配过滤器的推广。

与原始图像的维度相比,粉碎过滤器仅使用一小部分测量值即可实现高精度的分类。这项工作建立

在压缩分类的初步调查基础之上^[73-77],基于两个关键事实:1) 简单的参数模型在高维图像空间内的信号类上强加了低维流形结构;2) 流形的几何结构在它们的投影下被保留到一个随机的低维子空间。给定分类性能级别所需的测量数量不取决于图像的稀疏性或可压缩性,而仅取决于噪声水平和流形的结构,在流形的维度上线性增长,但在采样率及图像类别的数量上呈对数增长。文中基于 GMLC 的降维匹配滤波器很容易用单像素压缩成像相机实现。

(2) 支持向量机分类器

2015 年, Kulkarni^[43] 使用支持向量机分类器,在不需要图像重建的情况下,从压缩摄像机中执行动作识别推理。同样,该方法不实行图像重建过程,验证了在没有复杂的图像重建过程下视觉推断是否可行的问题 (重建通常需要高采样率和高信噪比),证明了在导致图像重建失败的压缩率和噪声水平的情况下,可直接进行动作识别,无需完整成像。

在此基础上,在 2018 年时, Ota 等人^[78] 在生物领域创新性地提出基于单像素无成像感知的超快速细胞分类技术,发表在《Science》期刊中。其完整流程及示意图如图 2 所示。

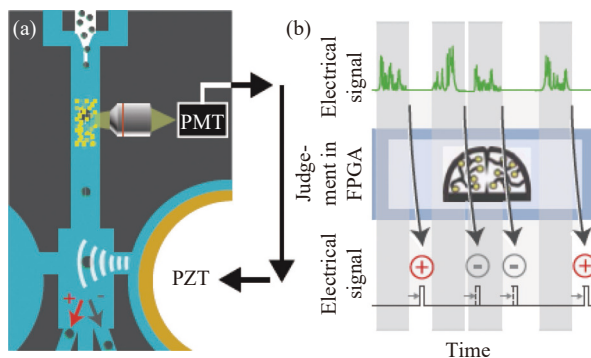


图 2 (a) 用于细胞分选的微流体系统; (b) 所探测电信号的处理^[78]
Fig.2 (a) A microfluidic system for cell sorting; (b) Processing of detected electrical signals^[78]

该研究操控细胞使其聚焦,并从固定的掩膜板前流过以完成随机掩膜的调制,经由单像素探测器记录每个细胞发出的荧光光波,并得到一个一维时序信号,根据此信号在 FPGA 上进行实时分析,其中 FPGA 上装载了经过训练的基于支持向量机的分类器。实验以最多 3000 个/s 细胞的吞吐率进行,最终依据类别信息进行实时的细胞分流控制。

(3) 高斯分类器

在探测过程中,为了提升感知速率,如何在最小的采样率下实现高效率的感知,是一个需要解决的实际问题。2019年,Latorre-Carmona等人^[52]基于单像素探测系统提出了一种实时的分类方法。在同一场景的探测过程中,通过多次采集可累积一系列的测量值,在每次测量之后,系统都会通过高斯分类器做出如下判断:继续获取目标场景的调制信息或直接输出预测结果,以此来提高感知效率。

具体地,对于 $M \times N$ 大小的场景 O ,其调制掩膜序列为 $H_k(k=1,2,\dots)$,第 k 个掩膜调制时获取的测量值 r_k 可表示为:

$$r_k = \sum_{m=1}^M \sum_{n=1}^N H_k(m,n) \cdot O(m,n) \quad (1)$$

因此,基于前 k 个掩膜调制的测量值与掩膜信息,通过简单的计算即可重构场景 $O(m,n)$:

$$O(m,n) = \sum_{k=1}^K r_k \cdot H_k(m,n) \quad (2)$$

在每次测量后,按照上述步骤重构场景 $O(m,n)$,即可得到目标概率密度函数:

$$P(O|C_i) = \frac{\exp\left(-\frac{(O-\mu_i)^T \sum_i^{-1} (O-\mu_i)}{2}\right)}{(2\pi)^{\frac{N}{2}} \left|\sum_i\right|^{\frac{1}{2}}} \quad (3)$$

式中: C_i 表示预测分类结果; μ_i 与 \sum_i 分别是不同类别 i 的均值与协方差矩阵,分别通过训练数据集 $\{S_i\}_1^Q$ 计算得到。因此,后验概率分布可表示为:

$$P(r_k|C_i) = \frac{\exp\left(-\frac{(r_k-\mu_i^k)^T \left[\sum_i^k\right]^{-1} (r_k-\mu_i^k)}{2}\right)}{(2\pi)^{\frac{Mk}{2}} \left|\sum_i^k\right|^{\frac{1}{2}}} \quad (4)$$

在实验中,研究人员为后验概率函数设定阈值,当其达到阈值,即判断已有较高概率输出正确预测结果,停止采集,输出预测类别,反之则继续测量。这种实时决策机制极大地提高了分类效率,且更加适用于实际应用。在此基础上,引申出了更多的先预判耦合数据,再输入至分类器的研究及应用。

2.1.1.2 基于距离模型计算的单像素感知

(1) 哈希距离

哈希算法(PHA)^[46,79]是一种应用于相似图像的

搜索算法。感知哈希技术作为PHA的一种,将图像数据转换为数千个二进制序列^[80],可以有效解决图像信息验证问题。PHA^[81]通过计算离散余弦变换系数矩阵的图像来计算哈希值。在傅里叶鬼成像^[59]方案中,可以通过使用正弦结构照明图案的差分测量直接获得物体图像的傅里叶系数矩阵。PHA结合傅里叶鬼成像在误码率(BER)或归一化汉明距离^[59]中都有实际应用。

2018年,基于哈希距离,上海交通大学的课题组^[63]提出了一种通过傅里叶光谱实现单像素高效感知的方法。该文结合离散傅里叶变换来计算散列值,并计算目标物体测量值和对比图像库中物体测量值之间的哈希差值以得到哈希距离。通过设置合适的阈值,可以快速准确地识别物体图像的类别。所提出的方案通过使用较少的测量来实现单像素感知哈希对象识别。

参考文献[63]利用了傅里叶域低频区域中自然对象的稀疏性和集中性特征以及计算鬼成像的框架。具体地,PHA为每张图像生成一个“指纹”,通过将对象的哈希“指纹”与图像库中的“指纹”进行比较,将其用于识别。为了获得哈希“指纹”,首先通过使用灰度、解析、谐波四步相移正弦模式进行照明^[61,82-84],也就是上文所使用的傅里叶掩膜,同时利用一个没有空间分辨率的单像素探测器收集反射光,并最终直接在傅里叶域中获取哈希值并计算哈希距离。

在得到场景目标的傅里叶系数后,需要计算场景目标的哈希值。对于傅里叶单像素成像来说,重建一个 $M \times N$ 像素图像,需要 $4 \times M \times N$ 次测量才可获得目标图像的完整傅里叶频谱。然而,哈希值计算不需要场景的完整傅里叶频谱,只需要收集傅里叶系数的实部,即 $D0 \sim D\pi$,或者傅里叶系数的虚部,即 $D\pi/2 \sim D3\pi/2$ 。只需要 $2 \times M \times N$ 次测量即可实现非成像对象的分类。为了制定统一的标准,参考文献[63]定义了哈希距离(HD)来表示图像库中对象图像和对比度图像之间的哈希值差异。

$$D_{objcon} = \frac{1}{MN} \sum_{i=1}^{MN} |obj_i - con_i| \quad (5)$$

式中: D_{objcon} 是物体图像和对比度图像之间的哈希距离; obj_i 表示对象图像的哈希值; con_i 表示对比图像的

哈希值。上式中归一化处理 M 、 N 、 D_{objcon} , 将其映射至 0~1 区间, 并通过多次实验找到了一个合适的阈值。如果目标图像与对比图像之间的哈希距离小于或等于该阈值, 则识别成功。否则, 将继续比较对象图像的“指纹”与“指纹”库, 直到分类成功。上述方法省去了成像过程, 实现了高效率的分类与识别。该方法基于测量值之间的距离进行搜索, 从而获得最佳阈值, 也为后续的研究提供了新的思路。

(2) 欧氏距离

最常用的距离测量方法包括欧氏距离、切比雪夫距离、切线距离、豪斯多夫距离等^[45]。欧氏距离因其简单高效而被广泛应用于各种图像识别算法中。假设 X 和 Y 是大小为 $M \times N$ 的两幅图像, $X=(X_{11}, X_{12}, \dots, X_{MN}), Y=(Y_{11}, Y_{12}, \dots, Y_{MN})$, 其中 $i=1, 2, \dots, M, j=1, 2, \dots, N$, 传统的欧式距离可表示为:

$$dist(X, Y) = \left[\sum_{i=1, j=1}^{M, N} (X_{ij} - Y_{ij})^2 \right]^{1/2} \quad (6)$$

式中: X_{ij} 和 Y_{ij} 分别表示图像 X 和 Y 。传统的欧氏距

离计算方法只衡量两幅图像对应像素的方差, 但是当图像稍有偏移或扭曲时, 可能会产生较大的欧氏距离, 在测量图像相似度时将会出现误差。为以进一步改进欧氏距离算法, 部分研究者提出了图像欧氏距离 (IMED)^[85]。

2021 年, 同济大学的 Ye^[45] 等人借鉴 IMED 算法, 提出了一种基于投影矩阵和 IMED 的“免成像”的目标分类识别技术, 他们通过主成分分析方法提取空间样本图像的主要特征。在计算重影成像算法中, 将本征空间的投影矩阵作为调制矩阵, 以降低数据的冗余性。在此基础上, 通过计算图像欧氏距离, 在空间位置上充分考虑像素的相关性, 从而提高传统目标分类算法的精度。具体的“免成像”的目标分类算法如图 3 所示。

在计算距离这一过程中, 参考文献 [45] 选取大小为 $m \times n$ 的 M 幅图像 X_1, X_2, \dots, X_M 组成样本图像集, 每幅图像作为列向量 Γ_i 首尾排列 ($1 \leq i \leq M$), 则构成训练样本矩阵 $\Phi = [\Gamma_1, \Gamma_2, \dots, \Gamma_M]$ 的 M 行和 $m \times n$ 列。样本

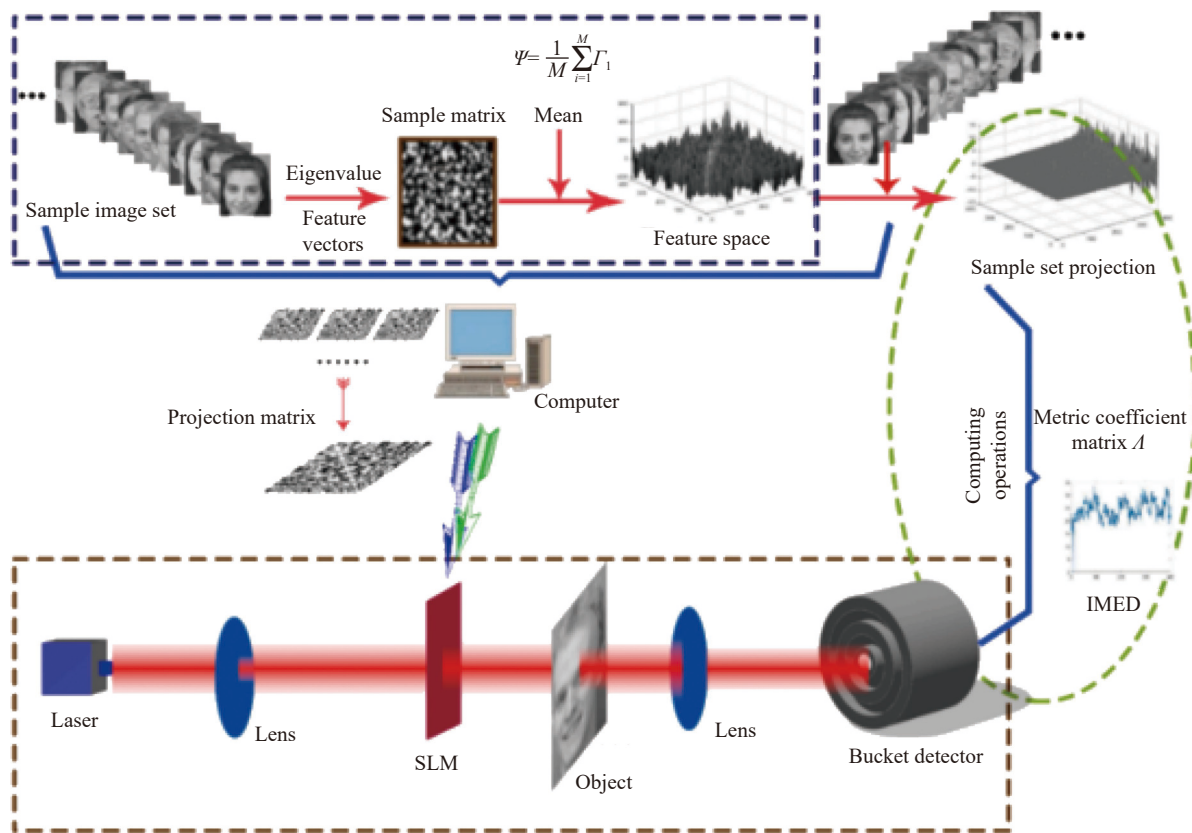


图 3 基于投影矩阵和 IMED 的非成像目标识别算法^[45]

Fig.3 Non-imaging target recognition algorithm based on projection matrix and IMED^[45]

图像集列向量的均值由 PCA 算法计算:

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (7)$$

均值向量包含样本图像集中所有图像的公共信息。为了最大化图像类别之间的差距,扩大 Γ_i 和 Ψ 之间的差异性,计算

$$\hat{\Phi} = [\Gamma_1 - \Psi \Gamma_2 - \Psi \dots \Gamma_M - \Psi] \quad (8)$$

协方差矩阵为:

$$C = \hat{\Phi} \hat{\Phi}^T \quad (9)$$

根据上式即可实现距离计算,通过距离判断即可完成基本分类。

上述研究基于特征空间投影矩阵和图像欧氏距离的计算,直接利用计算鬼成像系统的探测器观测值实现目标分类。该算法在不重建图像的情况下具有更好的识别性能,可结合其他距离模型的计算方法进一步提高分类性能。

上述基于模型驱动的机器学习分类算法各有优劣,适用于不同的应用场景。首先,在基于分类器模型^[41-43, 52]的方法中,最大似然分类器考虑了待分类样本到已知类别中心的距离及分布特征,算法易于实现,但要求统计量服从特定分布;支持向量机分类器可以解决小样本情况下的机器学习问题,有效提高泛化性能,并适用于高维求解问题,然而对缺失数据敏感,对非线性问题缺乏通用解决方案;高斯分类器基于较为完整的样本或特性信息可以实现实时决策,提高了分类效率。其次,在基于距离模型^[45, 53-54, 60-62]的方法中,哈希距离应用于相似图像的搜索算法,可以有效解决图像信息验证问题;欧氏距离易实现且直观,但由于每个坐标对欧氏距离的贡献是平等的,未考虑总体变异对距离的影响,容易产生误差。整体而言,基于模型驱动的机器学习方法计算复杂度低,分类精度易受缺失数据或复杂分布影响,因此在快速浏览分类概况中广泛使用。

2.1.2 数据驱动的深度学习方法

随着深度学习技术的发展与硬件平台计算能力的提高,研究人员提出了一系列数据驱动的深度学习方法用于单像素感知。在基于单像素耦合信息的特征提取等非线性问题的求解上,深度学习的梯度下降与反向传播算法相对于基于模型驱动的传统算法展

现出了很大的优势。

2016年, Lohit 等人^[52]与 Adler 等人^[86]于同年提出使用卷积神经网络从压缩数据中提取非线性特征,并根据特征的差异性直接实现有效的高级语义推理。算法流程如图 4 所示。

具体地,在获取场景光调制的耦合测量信息后,首先将该一维测量值映射为与原图同样维度、同样大小的“伪图像”,并将该“伪图像”作为输入传递至分类网络。网络的结构根据场景类别与任务需求可作调整,大致包括卷积层、池化层与全连接层。分类网络先通过多个卷积层提取图像的特征信息,卷积层包括线性的滤波卷积操作,以及非线性的激活操作,卷积层的输出为特征图。通过在卷积层之间加入最大池化层,可降低数据维度并尽量保持图像的平移不变性。网络末端一般是数个全连接层,以学习特征图中的全局特征,并将数据映射为一维。最后一层为全连接层,其输出长度为预测种类的数目,数值为目标场景在该种类下的匹配得分。最后,通过 softmax 等分类器,网络输出最终的感知结果。

类似地, Jiao 等人^[54]于 2019 年提出了一种光学深度学习模型 (MLSPI)。此方法针对光学衍射神经网络 (DNN) 仅可用于相干光调制解码的局限性,设计线性神经网络,实现了在单像素调制系统等非相干照明条件下的智能感知。该方法具有较低的实验复杂度,易于编程。然而,该方法的感知结果精度较低,仅能实现简单的数字分类,仍需进一步地工作以提高网络的鲁棒性。

在上述工作的基础上,为进一步提高基于耦合数据的分类效率,在系统层面,可通过引入新型先进光电器件或精细的光学工程设计提高感知效率;在算法层面,可改进调制过程与优化感知网络,通过传统的多步优化或端到端的联合优化方式实现;在应用层面,针对具体问题,可以通过设计自适应优化等算法,实现采样率自适应感知等具体目标。

(1) 系统优化

传统的单像素探测系统光路较为简单,对噪声等干扰的鲁棒性较差。为了提高感知效率,可从系统面对器件架构进行优化,以获取更丰富、更精确的目标特征耦合信息。

2020 年, Bu 等人^[56]基于单像素成像理论,提出

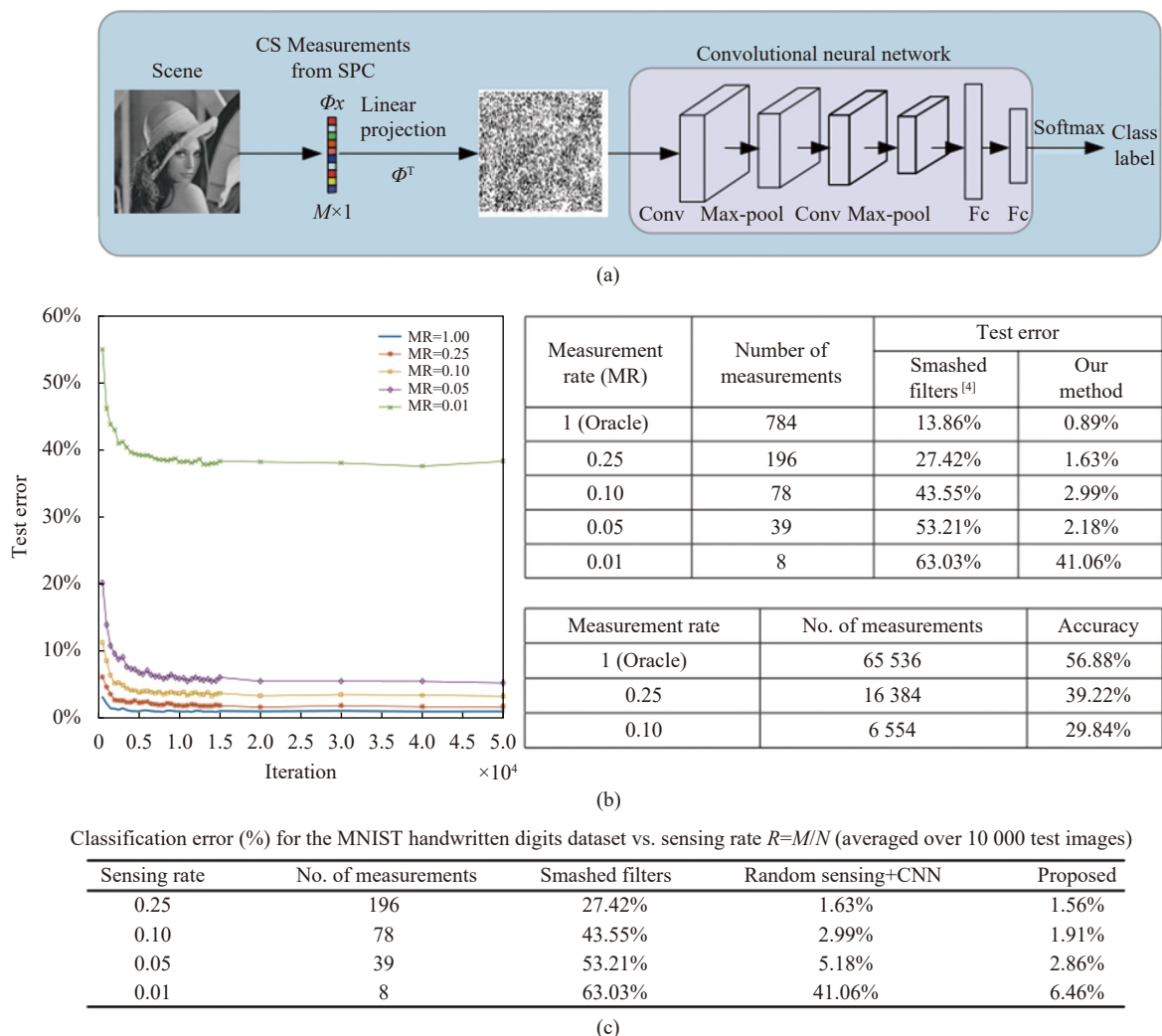


图 4 (a) 基于卷积网络的高效分类系统^[52]; (b) Lohit 等人在 MNIST 与 ILSVRC2012 的实验结果^[52]; (c) Adler 等人在 MNIST 数据集的实验结果^[86]
 Fig.4 (a) High efficient classification system based on the convolutional network^[52]; (b) Lohit's experiment results on MNIST and ILSVRC2012^[52]; (c) Adler's experiment results on MNIST^[86]

并验证了一种基于非线性光学的模式感知架构,如图 5 所示。该方法采用模式选择的图像转换方法 (MSIC), 将原始图像投影为一组相干的空间模式, 并以非线性方式提取其光场特征, 从而实现高精度分类。区别于传统的线性光学系统, 该架构利用了非线性光学过程里蕴含的丰富信息, 且并不直接对信号调制, 从而消除了可以避免的调制损耗与噪声, 实现了高效的智能感知。该方法在大规模图像的在线分类、快速激光雷达数据分析、复杂模式识别等方面均具有较高的潜在应用价值。次年, 该研究团队基于此系统, 使用局部傅里叶变换, 在原始时域与傅里叶频域联合提取图像特征信息, 进一步提高了算法的感知精度以及对噪声等干扰的鲁棒性^[57]。

针对传统单像素探测方法中调制光的分辨率受限于空间光调制器 (SLM) 分辨率, 进而导致连续

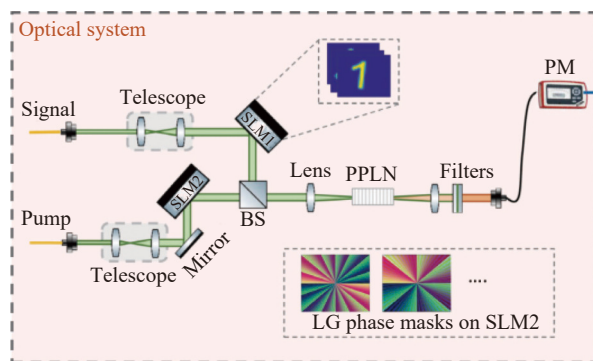


图 5 基于非线性光学的系统架构^[56]
 Fig.5 System setup with nonlinear optics^[56]

调制的频率严重受限的问题, Cao 等人^[71]于 2021 年提出了一种非相干光的调制方法, 他们用激光束照射旋转的毛玻璃, 产生非相干的调制光, 以提升调制效率。在感知阶段, 他们设计了一种由全连接层与卷积层构成的 FCNN 网络, 并在数据处理过程中采用了并行计算方案, 大大缩减了特征信息的提取时间。实验证实, 使用随机二值掩膜对场景光调制时, 该方法在二值图像与灰度图像上, 均可实现极低采样率下的高精度快速分类。

(2) 算法优化

为减小数据量与计算复杂度, 或提高系统的感知精度等, 研究人员在数据采集、特征提取阶段分别提出了多种改进方法, 通过多步优化实现了更高效率的感知。Rizvi 等人^[47]提出了一种基于单像素的快速分类方法。他们在提取耦合数据信息后增加了预处理阶段。首先, 使用经验模态分解 (EMD), 增强信号的特征信息。然后, 计算信号的能量, 将信号的能量代替一维耦合信号作为感知网络的输入, 以进一步降低数据量。在感知部分, 他们采用径向基函数神经网络 (RBF-NN) 代替了传统分类方法中的卷积神经网络, 利用来自目标场景的外部计算特征实现了高速、高精度的场景分类。

上述工作将 DMD 等器件的物理调制与后续的计算感知视作两个不同的阶段, 分别优化。2018 年, Zisselman 等人^[65]提出了一种端到端的高效率的免成像分类方法实现了二者的联合优化。他们用卷积层模拟掩膜调制的物理过程, 将其作为感知神经网络的一部分, 实现了掩膜设计与感知精度的联合优化。具体地, 该方法的网络输入为场景图像, 输出为分类结果。其主要创新点是在网络前端用卷积层代替物理模型中的数值固定的掩膜调制。其中, 卷积层的通道数为采样次数, 卷积核大小与场景保持一致, 数值即为掩膜。每个通道的卷积操作都是对场景光与调制掩膜的哈达玛积求和, 多通道的卷积操作后得到一维向量, 即模拟了多次的掩膜调制后的一维测量值。在网络训练时, 可同时优化卷积核数值与感知网络的参数。最后, 将优化的掩膜对应至物理调制中, 可实现更高精度的免成像分类。

然而, 在上述方法中, 将卷积层的卷积核参数作为优化后的掩膜, 会导致掩膜的数值为小数, 甚至会

出现负数。但是在实际调制中, 为了提高采样速率, 通常选用数值为 0 或 1 的二值掩膜。因此, 如何将优化后的灰度掩膜对应至物理调制的二值掩膜, 是一个需要考虑的实际问题。2020 年, Fu 等人^[38]提出了一种“编码-解码”网络结构, 实现了优化二值掩膜的高效感知, 如图 6 所示。

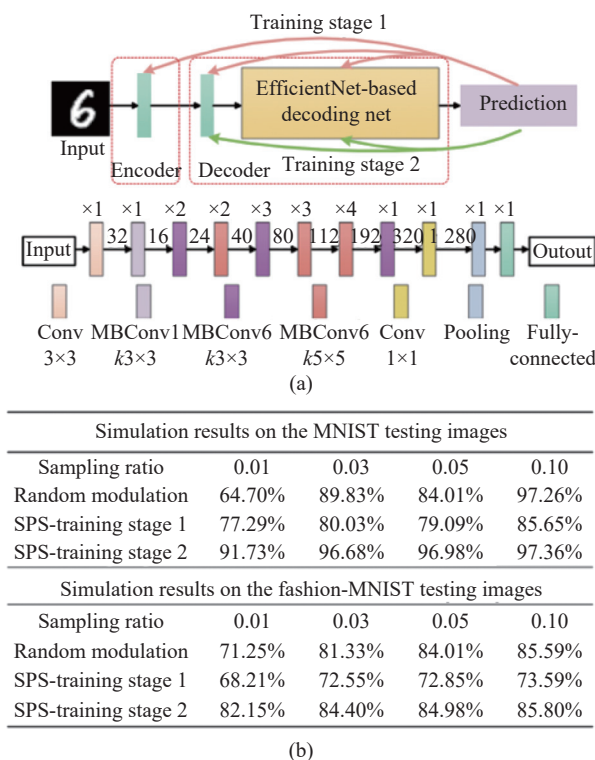


图 6 (a) “编码-解码”网络结构; (b) 分类结果^[38]

Fig. 6 (a) "Encode-decode" network structure; (b) Classification results^[38]

编码网络模拟单像素探测过程, 解码网络为感知网络。网络的训练分为两个阶段。阶段一中, 为了将优化后的灰度掩膜对应至二值, 需要采取如下步骤。从数学上推导, 假设存在一个系数 α , 使得:

$$W \approx \alpha W_b \tag{10}$$

式中: W 是灰度的参数矩阵; W_b 是二值化后的参数矩阵, 并且 $\alpha W_b \in \{\alpha, 0\}$, 可表示为:

$$\alpha W_b = \alpha \text{sign}(W) = \begin{cases} \alpha & W > 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

为了得到最优的二值调制掩膜, 需要解决以下优化函数

$$\alpha^* = \underset{\alpha}{\text{argmin}} W - \alpha W_b^2 \tag{12}$$

进一步可写作:

$$\operatorname{argmin}_{\alpha} F(\alpha) = \alpha^2 W_b^T W_b - 2\alpha W^T W_b + W^T W \quad (13)$$

进一步将 α 置零,便得到其最优解为:

$$\alpha^* = \frac{W^T W_b}{W_b^T W_b} \quad (14)$$

在网络第一阶段训练时,每次前向传播时,首先计算出 α 和 αW_b ,再用二值化的参数替换灰度参数,用于计算网络的损失和相应的梯度。在每一次后向传播时,更新灰度参数。通过这种迭代方式,最终获得最优的编码网络参数。

在网络第二阶段训练时,首先将编码网络的参数固定,只单独训练解码网络。用训练好的最优编码网络对光路进行二值调制,并用单像素探测器采集得到

耦合测量值,然后将该耦合测量值序列输入到解码网络中,完成对目标的类别预测。

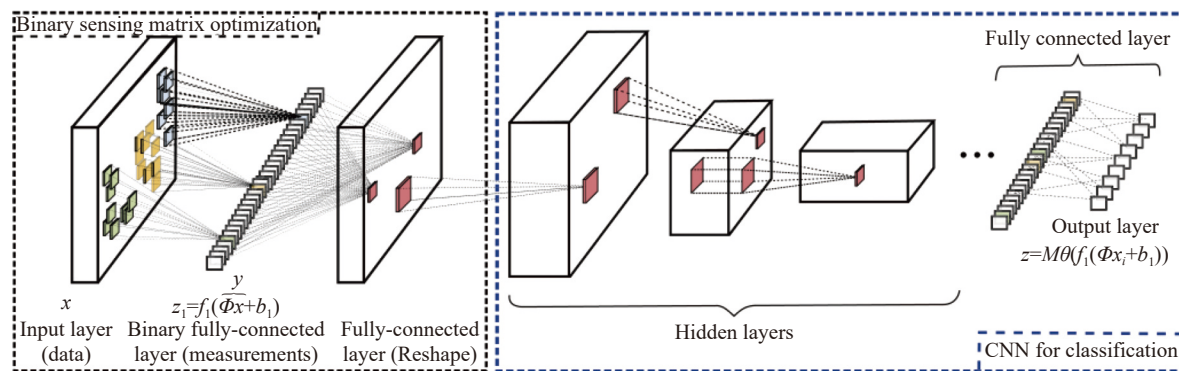
Fu 等人在真实实验中使用上述“编码-解码”网络优化后的二值掩膜,在 MNIST 数据集上以 3% 的低采样率实现了超过 96% 的识别准确率,验证了该方法的高效性。

类似地, Bacca 等人^[87-88]提出了一种端到端的感知方法,方法架构如图 7 所示。

他们将该联合优化问题表述为:

$$\{\Phi, \theta\} = \operatorname{argmin}_{\Phi, \theta} \frac{1}{L} \sum_{l=1}^L \mathcal{L}(M_{\theta}(\Phi_{x_l}), d_l)$$

$$\text{subject to } \Phi \in \{0, 1\}_{k, \pi} \quad k = 1, \dots, K, \quad n = 1, \dots, MN \quad (15)$$



(a)

Table Average classification accuracy for different sensing ratios with the MNIST dataset

Sensing ratio	Shots	Method	Test accuracy	Training accuracy
0.25	196	End-to-End	98.48±0.04%	99.99%
		Random-CNN	98.32±0.06%	99.99%
		Proposed method	98.51±0.12%	99.99%
0.10	78	End-to-End	98.29±0.02%	99.99%
		Random-CNN	97.01±0.14%	99.95%
		Proposed method	98.30±0.32%	99.99%
0.05	39	End-to-End	98.09±0.04%	99.99%
		Random-CNN	94.82±0.06%	98.78%
		Proposed method	97.99±0.12%	99.98%
0.01	8	End-to-End	95.18±0.04%	97.19%
		Random-CNN	58.94±0.06%	89.98%
		Proposed method	89.78±0.12%	96.09%

(b)

图 7 (a) 优化二值掩膜的单像素感知方法; (b) 分类结果^[88]

Fig.7 (a) Optimized single-pixel sensing method with optimal binarized patterns; (b) Classification results^[88]

式中: \mathcal{L} 是网络的损失函数; θ 是网络中的参数; d_l 代表不同种类标签。进一步地, 网络的损失函数可表示为:

$$\mathcal{L}(z_l, d_l) = -[d_l \log(z_l) + (1 - d_l) \log(1 - z_l)] \quad (16)$$

式中: z_l 是分类器对第 l 个图像的输出分类。基于上式, 通过同步训练二值掩膜与感知网络的参数, 实现了端到端的高精度感知。

这种端到端的联合优化方法在实际中也得到了一定应用。Zhang 等^[69]将此方法迁移至运动模糊的动态场景分类问题。首先, 他们利用静止的图像数据集训练端到端的免成像感知网络, 得到优化后的掩膜序列。其次, 对于获取的灰度优化掩膜, 他们使用 dithering 算法将其映射为二值。最后, 利用场景在静止与运动阶段的结构相似性等特征, 在物理调制阶段, 使用此掩膜调制高速运动的场景目标, 实现了高

精度的感知。2021 年, Pan 等人^[48]将此方法应用至无透镜的免成像感知装置中。他们采用 mask 对场景光进行调制, 该步骤对应单像素探测器的 DMD 或 SLM 调制过程, 并通过在端到端的感知网络中引入 Transformer 机制, 同步优化掩膜, 实现了高精度的免成像感知。

(3) 自适应优化

在单像素探测的实际实验中, 获取的耦合数据的长度是不固定的, 因此对于不同采样率的任务, 需要分别训练相应的分类网络, 这造成了巨大的计算与存储资源消耗。针对这一问题, Lohit 等人^[67]在 2018 年提出了一种采样率自适应的网络结构, 并基于单像素测量信息, 在分类、追踪等多种感知任务中证实了该方法的可行性。

采样率自适应的感知网络架构如图 8 所示, 网络的输入是长度不固定的单像素测量值, 输出是感知结

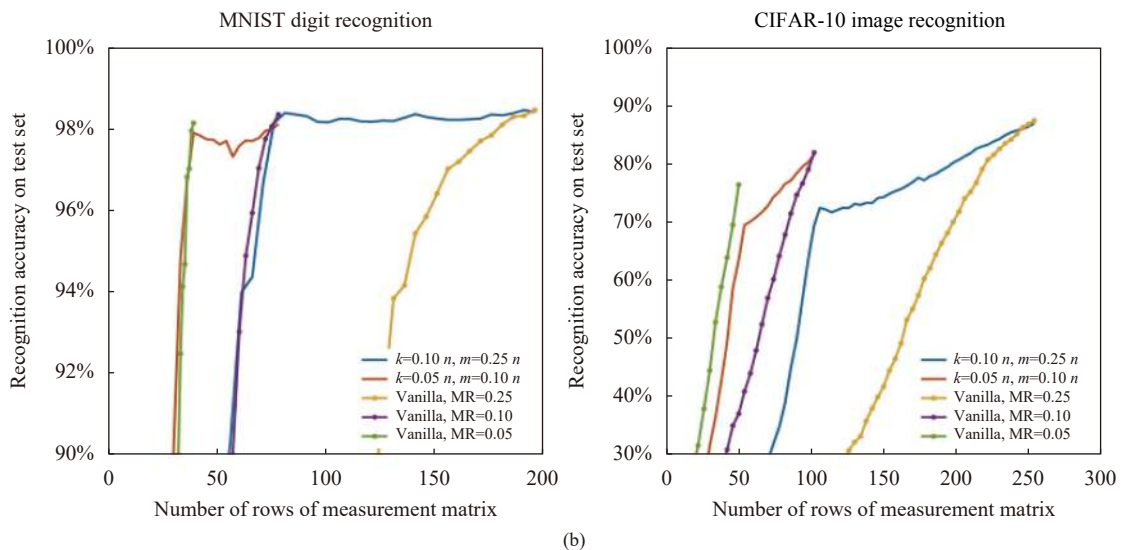
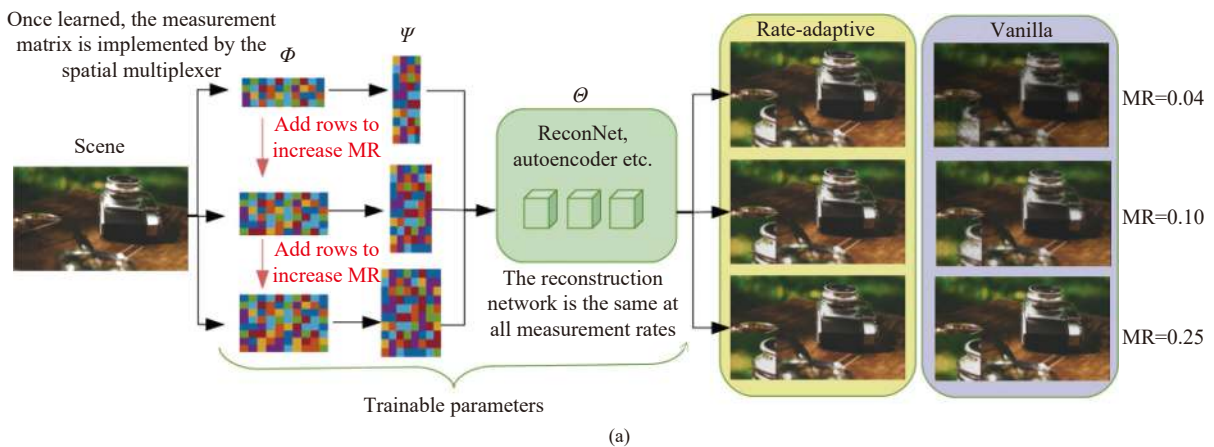


图 8 (a) 采样率自适应的感知网络; (b) 分类结果^[67]

Fig.8 (a) Sampling rate-adaptive sensing network's structure; (b) Classification results^[67]

果。网络采用端到端的训练方法,同步更新调制掩膜与感知网络参数。首先,训练图像通过网络层 ϕ 模拟调制过程,该全连接层的输出即为单像素测量值。然后,通过网络层 ψ ,将压缩信息重新映射至二维空间,再将此特征图输入至感知网络。感知网络可以是 ReconNet, autoencoder, DR2-Net 等重构网络或者 LeNet-5 等分类网络。

类似的,2020年,Xu等人^[66]提出了一种端到端的测量值自适应的网络,免成像网络设计为 DRNN 网络。通过实验,他们证实了该方法在不同数据集(MNIST、CIFAR-10、Fashion_MNIST、COIL-100等)、不同噪声等级下的感知效果,均达到了较高的识别精度,表明了此方法具有较好的泛化性与噪声鲁棒性。

总而言之,相较于传统模型驱动的机器学习方法,基于数据驱动的深度学习方法将物理调制与语义解调进行联合优化,具有更高的感知精度与计算速度。在此基础上,可通过多种优化方式进一步提高感知效率。系统层面的设计优化^[58-59,71]使感知框架具有更高的精度与鲁棒性,但对实验平台要求高,操作

难度大。算法层面的优化主要包括多步优化^[47]与联合优化^[37,65,68-69,87-88],多步优化可结合物理过程实现多阶段优化,可拓展性较高;联合优化通过端到端的学习联合优化编码与解码阶段,可实现实时的高精度分类。但现有的算法层面的优化方法^[66-67]仍局限于较简单的分类任务,方法的泛化性与鲁棒性有待改进。

2.2 识别

利用基于降维匹配滤波器的框架,Lohit团队^[42]利用单像素高效感知提取非线性特征。该文结合支持向量机来对人脸进行识别,具体说明如图9所示。该文收集了使用单像素探测器采集的30个主题的新人脸图像数据集。以人脸识别为例,在100%或更高的压缩比下,仍可执行“免成像”的推理,精度损失极低。

由于降维匹配滤波器可以直接在压缩测量上计算滤波器响应,因此在每个相关平面被划分为不重叠的块,并且对于每个块,计算峰值和峰值旁瓣比(PSR),即为所需的相关特征。PSR使用公式进行计算:

$$PSR = \frac{\text{peak} - \mu}{\sigma} \quad (17)$$

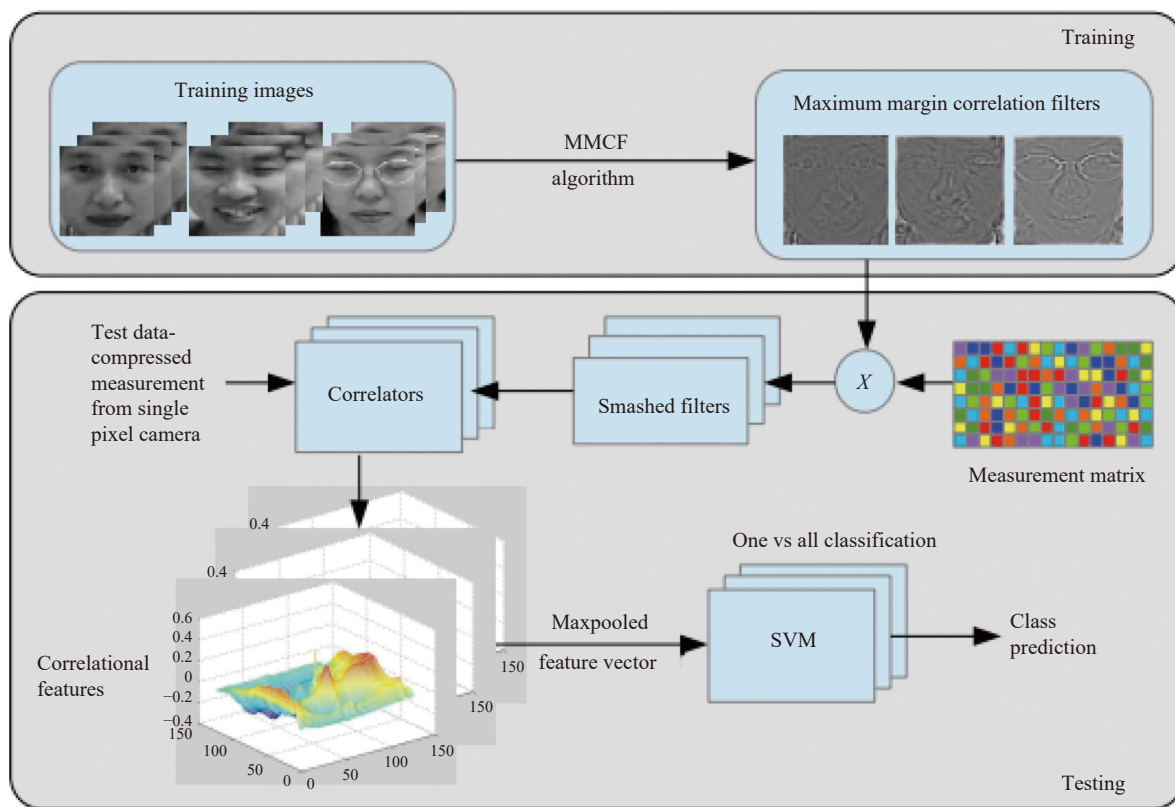


图 9 用于高级视觉推理问题的无重建相关特征估计的流程(以人脸识别为例)^[42]

Fig.9 Description of non-reconstruction related feature estimation for advanced visual reasoning problems (taking face recognition as an example)^[42]

式中： μ 是平均值； σ 是在以峰值为中心的掩模周围更大区域中相关值的标准偏差^[77]。对于 M 个相关场景中的每一个块，都获得了相似的 PSR 向量。将所有这些向量连接起来，从而形成特定测试图像的单个特征向量。该特征向量被输入到 M 个线性 SVM 中，以进行一对一识别，通过该分类器即可得到人脸识别的结果。

目前大部分的工作只针对单目标识别进行了研究，这些工作只能获得单一的简单语义信息，无法解析出多语义信息。然而在实际应用中，目标场景中往往存在多个目标。Bian^[37] 等人首次提出了一种新颖的单像素感知技术来解决多目标识别的问题。如图 10 所示，常规光学字符识别 (OCR) 需要获取二维图像，然后再提取相应的文本信息；基于 CNN 的网络的方法虽然避免了图像重建过程，但只能提取单一语义信息。而该方法构建的框架可专注于多个对象特征序列之间的内部联系，进而直接感知多个目标，也无需重建图像。

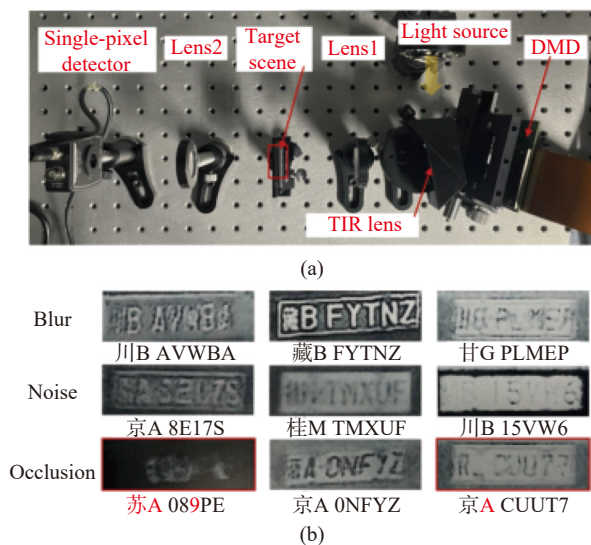


图 10 用于单像素多字符识别的实验验证设置及在 5% 采样率下车牌识别的结果^[37]。(a) 概念验证设置；(b) 各种车牌实验结果

Fig.10 Proof-of-concept setup for single-pixel multi-character recognition and the result of license plate recognition at a 5% sampling rate^[37]. (a) The proof-of-concept setup; (b) The experimental results of various license plates

该方法提出的端到端的递归网络可直接从耦合测量中感知多个目标，与传统成像方法相比，可减少模拟和存储负担；整个网络基于双向 LSTM 的架构层来同时预测多个字符分布，相比独立识别每个符号更

稳定。另外灰度调制掩膜和感知网络一起训练，可以确保以最低的采样率获得最佳的感知效率。

该技术已成功地应用于大规模车牌数据集的识别任务中，实验结果证明，该技术在 100 FPS 刷新率和 5% 的采样率下，可实现 87% 的多目标识别精度。

2.3 分割

场景分割是计算机视觉领域中的一项重要任务，通过将目标场景中的内容进行分类，细分为其组成部分或物体，以便提取感兴趣的目标区域并进行语义分析。现有的方法大多数基于图像进行目标场景分割^[89-92]，单像素感知相关工作只专注于分类和识别等任务，Liu^[70] 等人首次提出了一种单像素感知的场景分割方法。如图 11 所示，该系统架构主要包含了光场调制和单像素探测两部分。首先，使用优化的掩膜对光场进行调制，将目标场景特征信息压缩编码成一维光强度信号，然后通过一个单像素探测器采集耦合的光场总强度。获取的单像素测量数据直接输入设计的深度学习网络中进行场景语义分割并输出分割结果。

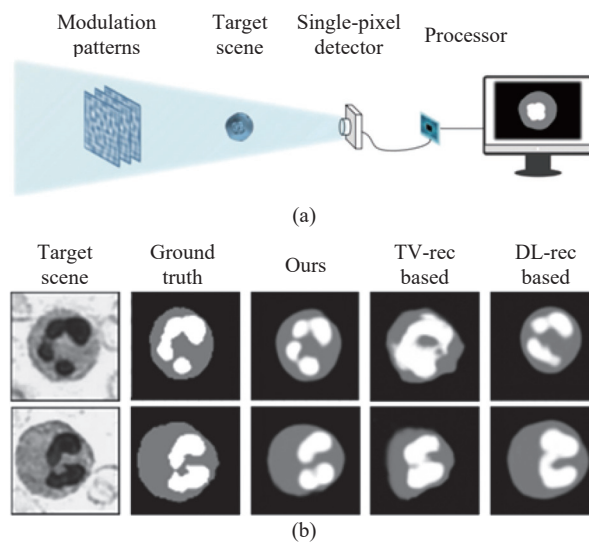


图 11 (a) 免成像的单像素场景分割系统框架；(b) 在 1% 采样率下对比传统先成像后感知方法实验结果^[70]

Fig.11 (a) Imaging-free single-pixel scene segmentation system framework; (b) Comparison of the experimental results of conventional imaging first and then perception methods at 1% sampling rate^[70]

为了更好地学习目标场景的特征实现最优的分割精度和效率，该方法设计了两阶段的网络训练方

法。第一阶段控制仅训练网络的编码子网络和特征提取模块,该过程用于提取目标场景特征,获得最优的感知编码性能。第二阶段控制将梯度回流至整个网络,即编码子网络和解码子网络均同时更新。另外,该方法将调制掩膜与后续网络一起学习,达到共同最优的分割结果。

实验结果证明,在低采样率的情况下 ($\leq 0.1\%$),单像素感知的方法仍然能在白细胞和道路分割数据集上获得较为精确的分割结果。该方法提供了一种在极低采样率下进行场景分割的方案,这使得一些资源受限的平台实现实时感知成为可能。

现有的基于单像素高效感知的分割技术在仿真数据集上性能良好,但在实际应用中受系统误差和环境干扰的影响,精度退化严重,在数据集中增加不同形式的噪声可进一步提升网络的鲁棒性。此外,网络模型可以进一步压缩以实现较快的分割速率,对调制掩膜进行二值化可以大幅提升采集速度等。

2.4 追踪

快速运动目标的实时检测与跟踪在安防、交通、生物医学领域都有着重要的应用,然而,大多数目标追踪的方法^[43-44, 70, 93-94]首先需要重建出每一帧的完整图像,浪费了计算机的存储和处理能力。

在单像素感知目标并追踪的任务中,Zhang^[60]等人直接利用单像素测量值检测和跟踪目标,且在硬件实现上与成像系统相同。该方法使用 6 个傅里叶基掩膜进行结构光调制,与前文分类任务中^[61]的方法类似,仅测量物体图像的完整傅里叶光谱中的 2 个傅里叶系数。由于傅里叶变换是一种全局到点的变换,因此 2 个傅里叶系数能够充分提供关于物体存在和运动的有效信息。此外,空间域中的平移导致傅里叶域中的线性相移,根据这一特性可以估计运动对象的位移。由以下公式可求得物体位移:

$$x_0 = -\frac{1}{2\pi f_x} \cdot \arg\{[\tilde{I}(f_x, 0) - I_{bg} \tilde{I}(f_x, 0)]\} \quad (18)$$

$$y_0 = -\frac{1}{2\pi f_y} \cdot \arg\{[\tilde{I}(0, f_y) - I_{bg} \tilde{I}(0, f_y)]\} \quad (19)$$

式中: $\arg\{\}$ 表示幅角运算; I_{bg} 表示在进行运动目标物体探测前,采集得到的背景傅里叶系数,最后使用均方根误差百分比 (RMSE) 用于量化恢复的投影曲线。实验证明:随着样本数量的增加,恢复曲线更接

近真实的投影曲线。该算法计算效率高,实现了低成本、实时的目标检测和跟踪。

同样利用场景前后帧之间的变换信息,Shi^[53]等人提出以高像素分辨率和超低采样率跟踪运动物体,并且不依赖于捕获连续帧的图像。具体的,文章将高维信息——时间序列上的二维图像转换为低维测量值——一维运动投影曲线。在复杂背景中,运动对象的投影曲线的公式可表示为:

$$f_{i,x}(y) = \sum_n [(I_{nx} - I_{nxb}) \cdot S_{x,n}(y)] / M \quad (20)$$

$$f_{i,y}(x) = \sum_n [(I_{ny} - I_{nyb}) \cdot S_{y,n}(x)] / M \quad (21)$$

式中: $N \times M$ 大小的图像的投影曲线 $f_i(x, y)$ 在 x 轴上表示为 $f_{i,x}(x)$, 在 y 轴上的投影曲线为 $f_{i,y}(y)$ 。

该算法实时获得高分辨率的一维投影曲线,提供场景中运动物体的位置信息,进而实现运动物体的跟踪。实验结果表明,在极低的采样率下 (20%),使用该方法可以达到场景刷新速率 177 FPS,实现实时跟踪 256×256 的移动物体。

将目标场景从二维转换为一维,减少了计算量和存储空间占用,大大提高了系统的计算效率和实时性。受上述方法的启发,Zhou^[54]等人提出了一种对投影曲线进行梯度差分计算的检测与跟踪方法,称为 PCGD,该方法可以在不成像的情况下实现复杂背景场景中快速运动目标的实时检测与跟踪。

该方法进一步对哈达玛掩膜进行分解,更准确地测量哈达玛谱,并在重建投影曲线时获得更多细节。为了提高效率和实时性,他们使用 EAHSI 优化的哈达玛掩膜序列来照亮快速移动的对象。EAHSI 方法根据调制掩膜中能量集中的有效区域的大小来选择投影顺序,得到优化的哈达玛掩膜序列。接着投影优化序列中的少量掩膜用于测量所需的信号,然后由单像素检测器进行收集。该方法具有超低的采样率和较高的计算效率,适用于快速运动目标实时检测和跟踪。

在二维目标追踪的基础上^[60],Deng^[62]等人提出了一种高效感知的三维跟踪方法。该方法借鉴于多目深度成像的光路系统,仅使用两个单像素检测器和一个高速空间光调制器进行数据采集。通过计算物体在两个正交的二维投影平面中的位置,进一步合成物体的立体三维信息。

同样,该方法使用傅里叶掩膜,在实验过程中投影 6 个傅里叶基底图案调制目标运动物体。 \bar{D} 代表 6 个图案的单像素测量值的平均值,如果目标运动物体不在场景中,或目标物体不产生位移,则 \bar{D} 的数值不随时间发生变化;否则,采集到的测量值 \bar{D} 相应发生改变。通过计算 \bar{D} 的前后差值可以判断出目标移动轨迹。同时, \bar{D} 的平稳性可作为判断运动物体存在

的一种方法。

该方法通过相应的单像素测量值直接分析了场景目标的位置,减少了数据存储量,提高了系统检测帧率。实验结果如图 12 所示,所提出的方法可以通过使用 10 000 Hz 数字微镜阵列以 1 666 帧/s 的帧速率检测和跟踪快速移动的物体,所提出的方法适用于隐藏的快速移动对象跟踪。

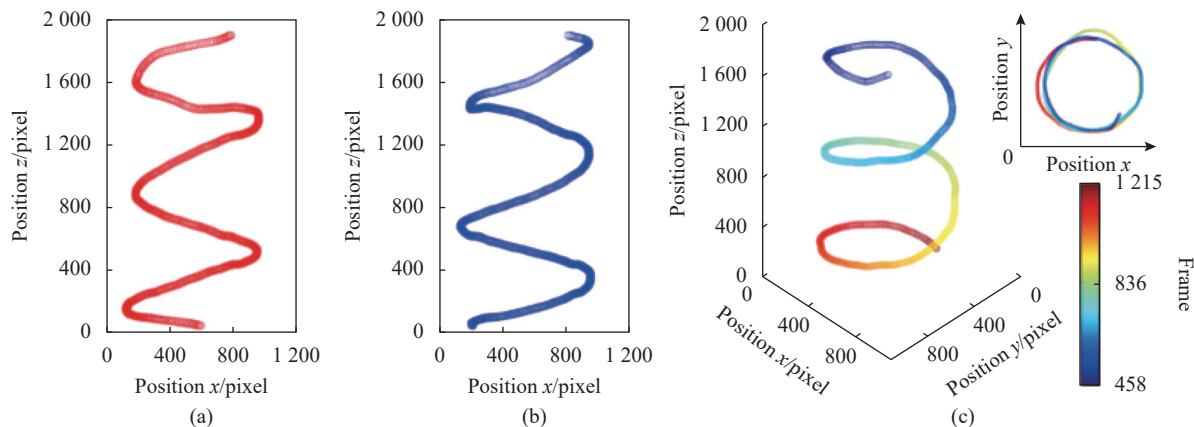


图 12 在 x - O - z 平面 (a)、 y - O - z 平面 (b) 和对应三维的目标 (c) 追踪结果^[62]

Fig. 12 Target tracking results in the x - O - z plane (a), y - O - z plane (b) and the corresponding three-dimensional targets (c)^[62]

目前,追踪算法将运动物体简化为单个空间点,难以描述具有复杂形状的真实物体。另外,追踪算法在前后帧测量值比较上仍旧有较大改进空间,如对于一维测量值的比较,使用余弦相似度、K-L 散度以及使用深度学习中的相似度,进行多方联合比较或将更高效地获取场景目标信息。

3 总结与展望

文中单像素高效感知的理论和方法进行了梳理及总结。在资源受限平台下,单像素感知呈现出独特优势,然而传统“先成像-后感知”的工作模式造成传输与计算资源的浪费并引入重建误差,在诸多领域的应用都面临挑战。一方面,由于目标仅占据部分视场,无需完整成像,因此引申出局部感兴趣区域的单像素成像技术,包括对调制掩膜的优化及直接局部重构,使得单像素成像的复杂性进一步降低,单像素探测器被赋予局部感知能力。另一方面,随着机器学习技术的蓬勃发展,单像素免成像高效感知成为新的研究方向。在模型驱动的机器学习算法中,有基于分类

器的高效感知(如最大似然分类器、支持向量机、高斯分类器等)算法以及基于距离的识别、追踪算法(如哈希算子、欧氏距离、投影距离等)。在数据驱动的深度学习方法中,对于分类任务,一些研究运用掩膜优化的多步训练网络,以及联合优化的端到端的网络(如 EfficientNet 网络、全连接网络等)。进一步地,针对不同采样率任务,提出自适应的动态网络;对于识别任务,如利用 LSTM 网络实现多个目标的识别;对于分割任务,使用 U-Net++ 直接完成血细胞与背景的分等。同样,在时间提升方面,一些研究在上述基础上,进一步提出实时高效感知网络。文中都对其中的算法和工作进行了详细介绍。

在极端探测领域(如非可见光光谱范围、弱光条件等),单像素探测装置展现出了阵列相机不可替代的优势^[95-97]。例如,在特定的波段中^[98-101](如太赫兹波段、X 光波段),阵列相机的制造成本远远高于单像素探测器。在极弱的光照下,经过实验验证^[50],当平均接收光子数小于 1 时,仍能达到 90% 以上的识别准确率。综上,单像素高效感知方法为资源受限的目标

感知任务提供了新思路。

在遥感探测领域,根据国家“十四五”规划关于高质量发展的蓝图布局,针对遥感检测行业的政策相继出台,遥感检测企业的技术研发水平不断提高,促使遥感检测行业市场规模逐步扩大^[102]。单像素遥感成像在地理地形图像^[2, 103]、卫星遥感影像^[104]等领域已有相关应用,而单像素高效感知技术可进一步提升现有的探测效能。例如,结合激光雷达等光学器件可进一步完成三维感知^[105],为远距离的多维度感知任务提供了新方向。

在信息安全领域,加密、信息隐藏等光学图像安全技术研究备受关注^[106]。其中,单像素探测器通过对场景的调制进行信息编码,在保持加密性的同时,也可用来防止恶意攻击,是一种安全性较高的加密手段。单像素高效感知方法针对光学加密系统,可以直接对接收到的密文数据及密钥进行数据分析,实现高精度的密文信息提取^[55, 68],也可在调制过程中伪造攻击,根据截取的密文伪造一组假键,以欺骗或混淆攻击者^[107],为提高未来系统安全性提供了指引。

需要强调的是,目前单像素高效感知技术仍处于初级发展阶段,与传统基于图像的视觉任务相比,基于欠采样低维数据的语义信息提取难度更大。而在计算机视觉任务中,目标分类是相对低级(low level)的语义感知任务,它不仅是物体检测、场景理解等高级(high level)语义感知任务的实现基础,也是验证各项新型感知技术的首选目标。因此,现阶段有关单像素高效感知技术的研究与应用大多集中于简单的分类任务,针对高级语义感知任务有待深入研究。

未来,单像素高效感知技术可结合视觉任务中的图像增强或图像翻译模型,针对所采集的一维数据,在所恢复的灰度图像基础上添加高质量的着色效果;也可以结合医学图像诊断深度模型,使用单像素探测器对不同光谱下的多维度信息联合解耦,实现显微级别的细胞检测及分割,辅助医学诊断,具有重要的研究价值和广阔的应用场景。此外,随着近年来人工智能相关领域的兴起与发展,无人驾驶、智慧医疗、智能航空航天等方向成为研究热点,这些领域对复杂场景下重点目标的识别精度和速度需求将进一步提高,将进一步推动单像素高效感知技术的发展进入新的阶段。

参考文献:

- [1] Ma J. A single-pixel imaging system for remote sensing by two-step iterative curvelet thresholding [J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(4): 676-680.
- [2] Ma J. Single-pixel remote sensing [J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(2): 199-203.
- [3] Llull P, Liao X, Yuan X, et al. Coded aperture compressive temporal imaging [J]. *Optics Express*, 2013, 21(9): 10526-10545.
- [4] Yuan X, Llull P, Liao X, et al. Low-cost compressive sensing for color video and depth[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2014: 3318-3325.
- [5] Yuan X, Liu Y, Suo J, et al. Plug-and-play algorithms for large-scale snapshot compressive imaging[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1447-1457.
- [6] Garipey G, Tonolini F, Henderson R, et al. Detection and tracking of moving objects hidden from view [J]. *Nature Photonics*, 2016, 10(1): 23-26.
- [7] Ashwin A, Wagadarikar, Nikos P, et al. Video rate spectral imaging using a coded aperture snapshot spectral imager [J]. *Optics Express*, 2009, 17: 6368-6388.
- [8] Lin X, Liu Y, Wu J, et al. Spatial-spectral encoded compressive hyperspectral imaging [J]. *ACM Transactions on Graphics (TOG)*, 2014, 33(6): 1-11.
- [9] Shao X, Su Y, Liu J, et al. The connotation and system of computational imaging [J]. *Acta Photonica Sinica*, 2021, 50(5): 0511001. (in Chinese)
- [10] Shao X, Liu F, Li W, et al. Review of computational imaging technology and applications [J]. *Laser & Optoelectronics Progress*, 2020, 57(2): 020001. (in Chinese)
- [11] Huang W, Jiao S, Xiao C. A review of image processing algorithms related to single-pixel imaging [J]. *Laser & Optoelectronics Progress*, 2021, 58(10): 267-284. (in Chinese)
- [12] Baird J L. Apparatus for transmitting views or images to a distance: US Patent, 1,699,270[P]. 1929-01-15. <https://encrypted.google.com/patents/US1699270>.
- [13] Bromberg Y, Katz O, Silberberg Y. Ghost imaging with a single detector [J]. *Physical Review A*, 2009, 79(5): 053840.
- [14] Sen P, Chen B, Garg G, et al. Dual photography [J]. *ACM Transactions on Graphics*, 2005, 24: 745-755.
- [15] Bian L, Suo J, Dai Q, et al. Experimental comparison of single-

- pixel imaging algorithms [J]. *JOSA A*, 2018, 35(1): 78-87.
- [16] Bian L, Suo J, Situ G, et al. Multispectral imaging using a single bucket detector [J]. *Scientific Reports*, 2016, 6: 24752.
- [17] Wang Y, Suo J, Fan J, et al. Hyperspectral computational ghost imaging via temporal multiplexing [J]. *IEEE Photonics Technology Letters*, 2015, 28(3): 288-291.
- [18] Li Z, Suo J, Hu X, et al. Efficient single-pixel multispectral imaging via non-mechanical spatio-spectral modulation [J]. *Scientific Reports*, 2017, 7(1): 1-7.
- [19] Bian L, Suo J, Chung J, et al. Fourier ptychographic reconstruction using Poisson maximum likelihood and truncated Wirtinger gradient [J]. *Scientific Reports*, 2016, 6: 27384.
- [20] Bian L, Suo J, Hu X, et al. Efficient single-pixel imaging in Fourier space [J]. *Journal of Optics*, 2016, 18(8): 085704.
- [21] Han S, Wu L, You L. Single-photon and single-pixel imaging special topic —foreword [J]. *Laser & Optoelectronics Progress*, 2021, 58(10): 1011000. (in Chinese)
- [22] Zheng Y. Research on edge detection based on single-pixel imaging [D]. Xi'an: Xi'an University of Technology, 2021. (in Chinese)
- [23] Graham M G, Steven D J, Miles J P. Single-pixel imaging 12 years on: A review [J]. *Optics Express*, 2020, 28: 28190-28208.
- [24] Pittman T B, Shih Y H, Strekalov D V, et al. Optical imaging by means of two-photon quantum entanglement [J]. *Physical Review A*, 1995, 52(5): R3429-R3432.
- [25] Bennink R S, Bentley S J, Boyd R W. "Two-photon" coincidence imaging with a classical source [J]. *Physical Review Letters*, 2002, 89(11): 113601.
- [26] Sun M J, Edgar M P, Gibson G M, et al. Single-pixel three-dimensional imaging with time-based depth resolution [J]. *Nature Communications*, 2016, 7(1): 1-6.
- [27] Wang H, Bian L, Zhang J. Depth acquisition in single-pixel imaging with multiplexed illumination [J]. *Optics Express*, 2021, 29(4): 4866-4874.
- [28] Cande's E, Romberg J. Sparsity and incoherence in compressive sampling [J]. *Inverse Problems*, 2007, 23: 969-985.
- [29] Duarte M F, Davenport M A, Takbar D, et al. Single-pixel imaging via compressive sampling [J]. *IEEE Signal Processing Magazine*, 2008, 25(2): 83-89.
- [30] Shapiro J H. Computational ghost imaging [J]. *Physical Review A*, 2008, 78: 061802.
- [31] Shi D, Huang J, Wang F, et al. Enhancing resolution of single-pixel imaging system [J]. *Optical Review*, 2015, 22(5): 802-808.
- [32] Pradeep S. On the relationship between dual photography and classical ghost imaging [J]. *arXiv*, 2013: 1309.3007.
- [33] Phillips D B, Sun M J, Taylor J M, et al. Adaptive foveated single-pixel imaging with dynamic supersampling [J]. *Science Advances*, 2017, 3(4): e1601782.
- [34] Zhai X, Cheng Z, Chen Y, et al. Foveated ghost imaging based on deep learning [J]. *Optics Communications*, 2019, 448: 69-75.
- [35] Cao J, Zhou D, Zhang Y, et al. Optimization of retina-like illumination patterns in ghost imaging [J]. *Optics Express*, 2021, 29(22): 36813-36827.
- [36] Mahalanobis A, Muise R. Object specific image reconstruction using a compressive sensing architecture for application in surveillance systems [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2009, 45(3): 1167-1180.
- [37] Bian L, Wang H, Zhu C, et al. Image-free multi-character recognition [J]. *Optics Letters*, 2022, 47(6): 1343-1346.
- [38] Fu H, Bian L, Zhang J. Single-pixel sensing with optimal binarized modulation [J]. *Optics Letters*, 2020, 45(11): 3111-3114.
- [39] Zhai X, Wu X, Sun Y, et al. Theory and approach of single-pixel imaging (Invited) [J]. *Infrared and Laser Engineering*, 2021, 50(12): 2021106. (in Chinese)
- [40] Zhai Y H, Chen X H, Zhang D, et al. Two-photon interference with true thermal light [J]. *Physical Review A*, 2005, 72(4): 043805.
- [41] Davenport M A, Duarte M F, Wakin M B, et al. The smashed filter for compressive classification and target recognition[C]// Proceedings of SPIE, 2007, 6498: 64980H.
- [42] Lohit S, Kulkarni K, Turaga P, et al. Reconstruction-free inference on compressive measurements[C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015: 16-24.
- [43] Kulkarni K, Turaga P. Reconstruction-free action inference from compressive imagers [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(4): 772-784.
- [44] Wei M, Xing F, You Z. A real-time detection and positioning

- method for small and weak targets using a 1D morphology-based approach in 2D images [J]. *Light: Sci Appl*, 2018, 7(5): 18006.
- [45] Ye Hualong, Zhang Leihong, Zhang Dawei. Non-imaging target recognition algorithm based on projection matrix and image Euclidean distance by computational ghost imaging [J]. *Optics & Laser Technology*, 2021, 137: 106779.
- [46] Zhao Y, Wang S, Zhang X, et al. Robust hashing for image authentication using zernike moments and local features [J]. *IEEE Transactions on Information Forensics & Security*, 2013, 8(1): 55-63.
- [47] Rizvi S, Cao J, Hao Q. High-speed image-free target detection and classification in single-pixel imaging[C]//SPIE Future Sensing Technologies, International Society for Optics and Photonics, 2020, 11525: 115250 X.
- [48] Pan X, Chen X, Nakamura T, et al. Incoherent reconstruction-free object recognition with mask-based lensless optics and the Transformer [J]. *Optics Express*, 2021, 29(23): 37962-37978.
- [49] Zhu Y, Shi J, Wu X, et al. Photon-limited non-imaging object detection and classification based on single-pixel imaging system [J]. *Applied Physics B*, 2020, 126(1): 1-8.
- [50] Pratt W K, Kane J, Andrews H C. Hadamard transform image coding [C]//Proceedings of the IEEE, 1969, 57(1): 58-68.
- [51] Lohit S, Kulkarni K, Turaga P. Direct inference on compressive measurements using convolutional neural networks[C]//2016 IEEE International Conference on Image Processing (ICIP), IEEE, 2016: 1913-1917.
- [52] Latorre-Carmona P, Traver V J, Sánchez J S, et al. Online reconstruction-free single-pixel image classification [J]. *Image and Vision Computing*, 2019, 86: 28-37.
- [53] Shi D, Yin K, Huang J, et al. Fast tracking of moving objects using single-pixel imaging [J]. *Optics Communications*, 2019, 440: 155-162.
- [54] Jiao S, Feng J, Gao Y, et al. Optical machine learning with incoherent light and a single-pixel detector [J]. *Opt Lett*, 2019, 44(21): 5186-5189.
- [55] Kravets V, Javidi B, Stern A. Compressive imaging for defending deep neural networks from adversarial attacks [J]. *Optics Letters*, 2021, 46(8): 1951-1954.
- [56] Bu T, Kumar S, Zhang H, et al. Single-pixel pattern recognition with coherent nonlinear optics [J]. *Opt Lett*, 2020, 45(24): 6771-6774.
- [57] Kumar S, Bu T, Zhang H, et al. Robust and efficient single-pixel image classification with nonlinear optics [J]. *Opt Lett*, 2021, 46(8): 1848-1851.
- [58] Zhou F, Shi X, Chen J, et al. Non-imaging real-time detection and tracking of fast-moving objects using a single-pixel detector [J]. *arXiv preprint*, 2021: 2108.06009.
- [59] Zhang Z B, Ma X, Zhong J G. Single-pixel imaging by means of Fourier spectrum acquisition [J]. *Nature Communications*, 2015, 6(1): 1-6.
- [60] Zhang Z, Ye J, Deng Q, et al. Image-free real-time detection and tracking of fast moving object using a single-pixel detector [J]. *Optics Express*, 2019, 27(24): 35394-35401.
- [61] Gustafsson M G L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy [J]. *Journal of Microscopy*, 2000, 198(Pt 2): 82-87.
- [62] Deng Q, Zhang Z, Zhong J. Image-free real-time 3-D tracking of a fast-moving object using dual-pixel detection [J]. *Optics Letters*, 2020, 45(17): 4734-4737.
- [63] Chen H, Shi J, Liu X, et al. Single-pixel non-imaging object recognition by means of Fourier spectrum acquisition [J]. *Optics Communications*, 2018, 413: 269-275.
- [64] Zhen S, Yao M, Wang S, et al. Single-pixel fast-moving object classification based on optical-electronical hybrid neural network (invited) [J]. *Infrared and Laser Engineering*, 2021, 50(12): 20210856. (in Chinese)
- [65] Zisselman E, Adler A, Elad M. Compressed Learning for Image Classification: A Deep Neural Network Approach[M]//Handbook of Numerical Analysis. Amsterdam: Elsevier, 2018, 19: 3-17.
- [66] Xu Y, Liu W, Kelly K F. Compressed domain image classification using a dynamic-rate neural network [J]. *IEEE Access*, 2020, 8: 217711-217722.
- [67] Lohit S, Singh R, Kulkarni K, et al. Rate-adaptive neural networks for spatial multiplexers [J]. *arXiv preprint*, 2018: 1809.02850.
- [68] Bacca J, Galvis L, Arguello H. Coupled deep learning coded aperture design for compressive image classification [J]. *Optics Express*, 2020, 28(6): 8528-8540.
- [69] Zhang Z, Li X, Zheng S, et al. Image-free classification of fast-moving objects using "learned" structured illumination and single-pixel detection [J]. *Optics Express*, 2020, 28(9): 13269-13278.

- [70] Liu H, Bian L, Zhang J. Image-free single-pixel segmentation [J]. *arXiv preprint*, 2021: 2108.10617.
- [71] Cao J N, Zuo Y H, Wang H H, et al. Single-pixel neural network object classification of sub-Nyquist ghost imaging [J]. *Applied Optics*, 2021, 60(29): 9180-9187.
- [72] Baraniuk R G, Wakin M B. Random projections of smooth manifolds [J]. *Foundations of Computational Mathematics*, 2009, 9(1): 51-77.
- [73] Waagen D, Shah N, Ordaz M, et al. Cassabaum, random subspaces and SAR classification efficacy[C]//Proceedings of SPIE, 2005, 5808: 305-308.
- [74] Duarte M F, Davenport M A, Wakin M B, et al. Sparse signal detection from incoherent projections[C]//IEEE International Conference on Acoustics. IEEE, 2006.
- [75] Davenport M A, Wakin M B, Baraniuk R G. Detection and estimation with compressive measurements[EB/OL]. (2006-11-01)[2022-03-31]. <https://www.mendeley.com/catalogue/3242ecd4-417d-3692-822e-7000c8fd7496/>.
- [76] Haupt J, Castro R, Nowak R, et al. Compressive sampling for signal classification[C]//2006 Fortieth Asilomar Conference on Signals, Systems and Computers. IEEE, 2007.
- [77] Savvides M, Kumar B V, Khosla P. Face verification using correlation filters[C]//3rd IEEE Automatic Identification Advanced Technologies, 2002: 56-61.
- [78] Ota S, Horisaki R, Kawamura Y, et al. Ghost cytometry [J]. *Science*, 2018, 360(6394): 1246-1251.
- [79] Weng L, Preneel B. A secure perceptual hash algorithm for image content authentication [J]. *Lecture Notes in Computer Science*, 2011, 7025: 108-121.
- [80] Niu X M, Jiao Y H. An overview of perceptual hashing [J]. *Acta Electronica Sinica*, 2018, 36(7): 1405-1411.
- [81] Zhang B, Xin Y, Niu X X. Image perceptual hash algorithm based on target character[C]//2012 IEEE 13th International Conference on Communication Technology, 2012: 397-401.
- [82] Chang L, Yan W G, Wang W D. Research on Robust Image Perceptual Hashing Technology Based on Discrete Cosine Transform[M]//Financial Sciences, and Management. Berlin, Heidelberg: Springer, 2012: 799-809.
- [83] Ryu J, Hong S S, Horn B K P, et al. Multibeam interferometric illumination as the primary source of resolution in optical microscopy [J]. *Applied Physics Letters*, 2006, 88(17): 171112.
- [84] Dan D, Lei M, Yao B, et al. DMD-based LED-illumination super-resolution and optical sectioning microscopy [J]. *Scientific Reports*, 2013, 3(1): 1-7.
- [85] Turan A. A weighted Euclidean distance based TOPSIS method for modeling public subjective judgments [J]. *Asia-Pacific Journal of Operational Research*, 2017, 34(3): 1750004.
- [86] Adler A, Elad M, Zibulevsky M. Compressed learning: A deep neural network approach [J]. *arXiv preprint*, 2016: 1610.09615.
- [87] Hinojosa C, Bacca J, Arguello H. Coded aperture design for compressive spectral subspace clustering [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2018, 12(6): 1589-1600.
- [88] Bacca J, Correa C V, Vargas E, et al. Compressive classification from single pixel measurements via deep learning[C]//2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), 2019: 1-6.
- [89] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
- [90] Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation [J]. *IEEE Transactions on Medical Imaging*, 2019, 39(6): 1856-1867.
- [91] Huang H, Lin L, Tong R, et al. Unet 3+: A full-scale connected unet for medical image segmentation[C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020: 1055-1059.
- [92] Cao J, Cholakkal H, Rao M, et al. D2 Det: Towards high quality object detection and instance segmentation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- [93] Stern A, Evladov S, Kashter Y, et al. Anamorphic optics for compressive imaging and compressive motion tracking[C]//2012 11th Euro-American Workshop on Information Optics, 2012: 1-3.
- [94] Lohit S, Kulkarni K, Kerviche R, et al. Convolutional neural networks for noniterative reconstruction of compressively sensed images [J]. *IEEE Transactions on Computational Imaging*, 2018, 4(3): 326-340.
- [95] Zhang Z, Lu T, Peng J, et al. Single-pixel imaging technology and application [J]. *Infrared and Laser Engineering*, 2019,

- 48(6): 0603002. (in Chinese)
- [96] Khakimov R I, Henson B M, Shin D K, et al. Ghost imaging with atoms [J]. *Nature*, 2016, 540: 100-103.
- [97] Ryczkowski P, Barbier M, Friberg A T, et al. Ghost imaging in the time domain [J]. *Nature Photon*, 2016, 10: 167-170.
- [98] Zhao H, Guo Y, Li P, et al. Signal to noise ratio analysis of single pixel imaging and its development in special bands [J]. *Laser & Optoelectronics Progress*, 2021, 58(10): 1011010. (in Chinese)
- [99] Single pixel complex amplitude holography can be applied to microscopic imaging of biological tissue[J]. *Biomedical Engineering and Clinic*, 2021, 25 (6): 714. (in Chinese)
- [100] He Y, Huang Y, Zeng Z, et al. Single-pixel imaging with neutrons [J]. *Science Bulletin*, 2021, 66(2): 133-138.
- [101] Stantchev R, Yu X, Blu T, et al. Real-time terahertz imaging with a single-pixel detector [J]. *Nature Communications*, 2020, 11(1): 1-8.
- [102] Lin R, Zheng Y, Chen S. Research on opportunities and problems of international application and promotion of satellite remote sensing in China [J]. *Satellite Applications*, 2021(10): 45-51. (in Chinese)
- [103] Lu Y, Zhan W, Hu C. Detecting and quantifying oil slick thickness by thermal remote sensing: A ground-based experiment [J]. *Remote Sensing of Environment*, 2016, 181: 207-217.
- [104] Erkmen B I. Computational ghost imaging for remote sensing [J]. *JOSA A*, 2012, 29(5): 782-789.
- [105] Li Z P, Ye J T, Huang X, et al. Single-photon imaging over 200 km [J]. *Optica*, 2021, 8: 344-349.
- [106] Jiao S, Zhou C, Shi Y, et al. Review on optical image hiding and watermarking techniques [J]. *Optics & Laser Technology*, 2019, 109: 370.
- [107] Feng J, Huang W, Jiao S, et al. Generalized forgery attack to optical encryption systems [J]. *Optics Express*, 2021, 29(26): 43580-43597.