

Deep-learning based on-chip rapid spectral imaging with high spatial resolution

摘要

光谱成像扩展了传统彩色摄像机的概念，能够在多个光谱通道上捕捉图像，具有广泛的应用前景。基于扫描方法的传统光谱相机存在采集速度慢和体积大的缺点。基于芯片上计算的元表面滤波器的光谱成像为便携应用提供了一种有前途的方案，但由于逐点迭代光谱重建和重建后的光谱图像中的马赛克效应，它耐受长时间的计算。在这项研究中，演示了芯片上的快速光谱成像，通过基于深度学习的光谱数据立方体重建消除了光谱图像中的马赛克效应。实验结果表明，比迭代光谱重建快了 4 个数量级，对标准彩色板的光谱重建的保真度超过了 99%。特别是，在移动物体和户外驾驶场景下演示了视频速率的光谱成像，对于识别异色现象表现良好，可以通过它们的光谱区分颜色相似的天空和白色汽车，展现了在智能感知领域，尤其是在自动驾驶和其他实际应用中的巨大潜力。

关键词：光谱成像；深度学习；超表面

1 引言

光谱一般是指波长覆盖紫外到红外波段的电磁波频谱，反映了光波与物质相互作用的丰富信息。不同波段的光谱记录着不同类型能级跃迁的辐射信息。其中，电子跃迁对应紫外和可见光波段，分子的振动和转动能级跃迁则对应于红外乃至微波波段。每种物质都有特征吸收、发射或散射光谱，因此光谱也被称为物质的“指纹”，通过光谱分析可以鉴别物质并确定其化学组成。自 20 世纪 60 年代以来，随着遥感技术的兴起，人们期望同时探测物体的图像和光谱信息，这催生了光谱成像技术，它能够获得包含二维空间 (x, y) 和光谱 λ 信息的三维数据立方体，超越人眼感知能力，在疾病诊疗 [1,2]、精准农业 [3]、食品安全 [4]、天文探测 [5]、环境监测 [6] 等诸多领域有着重要的应用前景。

根据光谱分辨能力的不同，可以将光谱成像划分为多光谱成像、高光谱成像和超光谱成像。具体来说，通常采用工作中心波长与分辨率的比值 $\lambda/\Delta\lambda$ 来进行区分。

- 1) 多光谱成像 (multispectral imaging): 对应的 $\lambda/\Delta\lambda \sim 10$ ，在可见光波段对应几个波长通道，一般用于地带分类领域；
- 2) 高光谱成像 (hyperspectral imaging): 对应的 $\lambda/\Delta\lambda \sim 100$ ，在可见光波段对应数十个波长通道，被广泛应用于遥感、农业等领域；
- 3) 超光谱成像 (ultraspectral imaging): 对应的 $\lambda/\Delta\lambda \sim 1000$ ，在可见光波段对应数百个波长通道，通常用于大气微粒探测、天文观测等领域。

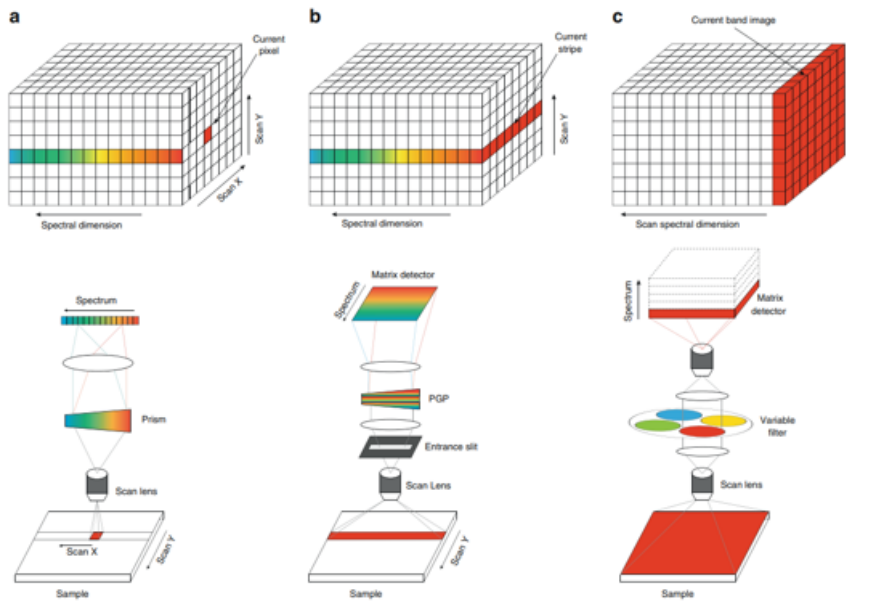


图 1. 传统光谱成像示意图。a. 点扫描。b. 线扫描。c. 波长扫描

2 相关工作

2.1 传统光谱成像技术

传统的光谱成像仪器依赖于空间扫描，如扫帚扫描 [7] 和推扫扫描 [8]，或者依赖于光谱扫描，如滤光轮 [9] 和可调滤光器 [10, 11]。根据其扫描方式的不同，可将光谱成像分为点扫描、线扫描、波长扫描三类。如图1所示。点扫描采用线阵列探测器，每次只获取目标物体上一个点的光谱信息，通过二维空间扫描获取整个数据立方体，采集速度慢。线扫描采用面阵列探测器，每次可以获取目标物体一行的光谱信息，通过一维空间扫描实现完整的光谱成像。波长扫描采用面阵列探测器。并通过一组前置滤光片或可调谐滤光片实现波长通道的切换或扫描依次采集各个波长下的图像信息。然而，这些扫描方法存在采集速度慢的问题，不适用于对移动目标进行动态记录。

2.2 快照式单点成像

为了克服扫描速度慢的限制，人们研究了快照光谱成像方法。早期的快照技术，如积分场光谱学 [12]、多光谱光束分裂 [13] 和图像复制成像光谱仪 [14]，仍然依赖于光的散射，其光学系统也比较庞大。随着压缩感知技术 [15, 16] 的发展，计算快照式光谱成像技术引起了越来越多的研究兴趣 [17]，可分为三类：编码孔径、基于散斑和光谱滤波阵列方法。对于编码孔径方法，经典系统是编码孔径瞬时光谱成像仪（CASSI），它使用固定掩模和色散元件实现波段调制。CASSI 能够利用深度学习技术快速捕捉和重建高光谱图像。然而，复杂的光学元件导致系统体积大，无法满足对便携应用不断增长的需求。基于散斑方法利用散射介质或衍射光学元素的波长依赖性来实现光谱成像。尽管这些系统可以更紧凑，但光谱分辨率受到散斑在波长范围内的相关性的限制。光谱滤波阵列方法可以看作是拜耳滤波器的扩展，采用包含一组光谱滤波器的超像素进行光谱恢复。尽管这一类方法具有紧凑的设备尺寸和高光谱精度的优势，但在重建的光谱图像中存在马赛克效应，即边缘点的恢复光谱不准确。

3 本文方法

3.1 本文方法概述

近年来基于计算光谱重建的新型单点光谱仪在小型化和微型化上取得了突破性进展,但快照式光谱成像技术的研究工作在频谱精度、空间分辨率、成像速度上尚无可同时兼顾的方案。基于超表面的光谱成像技术方案,通过设计不同结构的超表面单元实现对空间各点入射光频谱的宽带调制,采用图像传感器采集调制光信号,通过计算重建得到入射光的光谱信息。超表面调制单元的个数可以显著小于波长通道数,有效减小微型光谱仪的体积:将计算光谱仪进行空间阵列化便可实现光谱成像,其技术方案的优势在于:

- 1) 超表面具有很大的设计自由度,通过改变超表面单元的结构参数,能够实现丰富的透射谱,即获得大带宽、高频谱分辨的光强调制,可突破传统分光型光谱仪由空间光程限制的波长分辨,从而实现高频谱分辨率;
- 2) 超表面单元的数量可以远小于光谱通道数,能够大幅减小每个像素点单个微型光谱仪的体积,实现高空间分辨率,同时也为缩短成像时间提供可能;
- 3) 超表面接收空间光输入时无须耦合结构,适合与现有成像技术融合;
- 4) 超表面采用互补金属氧化物半导体 (CMOS) 兼容的加工工艺,可进一步实现大面积、低成本的晶圆级制备。

该论文采用了为 CASSI 开发的先进深度学习增强算法,演示了芯片上无马赛克的快速光谱成像,应用于我们先前工作中报道的基于超表面的光谱成像器。这个超表面光谱成像器在不同光谱波段产生不同的振幅调制模式,起到了 CASSI 中固定掩模加色散器的作用。具体而言,这个光谱成像器是通过将一个由 360×440 个超表面单元组成的超表面层与自由形状的元子组成的 CMOS 图像传感器 (CIS) 集成而设计的。总共有 400 种超表面单元,每种都具有独特的光谱响应函数。作为原理验证,从 450 到 700nm 中选择了 256×256 个超表面单元和 26 个波长带,步长为 10nm。采用了基于交替方向乘子法 (ADMM) 算法的深度展开网络,称为 ADMM-net,用于快速重建光谱图像。在这里,该网络在包含了 75 万个大小为 $256 \times 256 \times 26$ 的光谱数据立方体的合成数据集上进行了训练,这些数据是从 CAVE [18] 和 KAIST [19] 数据集生成的。此外,为了模拟真实测试情况,对测量值施加了加性白高斯噪声。通过将 ADMM-net 应用于直接重建光谱数据立方体,实现了芯片上消除马赛克效应的快速光谱成像。与逐点迭代光谱重建相比,ADMM-net 实现了四个数量级的速度提升,使光谱数据立方体的重建速率达到每秒 55 个,并且对于 24 块色卡,平均光谱重建保真度超过 99%。在实际应用中,演示了用于移动物体和户外驾驶场景的视频速率光谱图像重建。发现通过它们的光谱,可以有效区分相似颜色的天空和移动的白色汽车,而现有的无人驾驶车辆可能会错误地将白色卡车误认为是天空,从而导致事故。当前工作采用的方法不仅能够解决自动驾驶 [20] 中由于异色现象识别缺陷而引起的巨大安全问题,而且对于智能感知的其他领域也具有潜在的广泛应用前景。

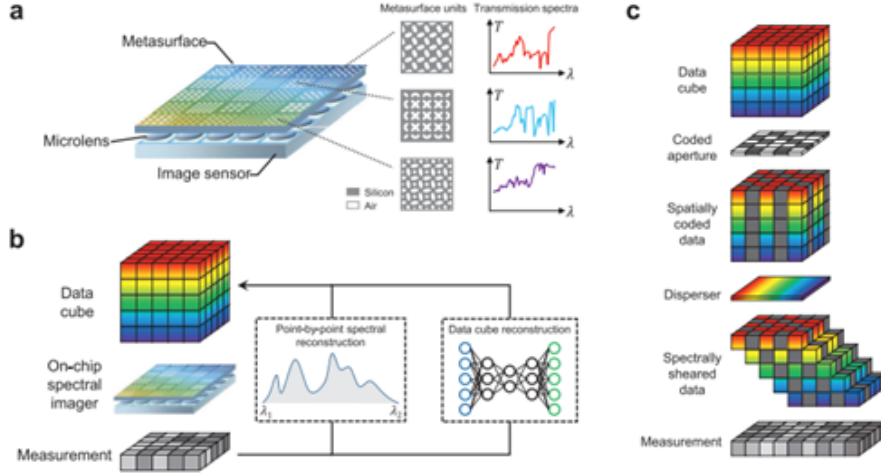


图 2. a, 基于超表面的光谱成像仪的示意图, 包括一个超表面层、一个微透镜层和一个图像传感器层。硅超表面包含 360×440 个超表面单元, 具有自由形状的超原子。有 400 种具有独特透射光谱的超表面单元。通常, 5×5 个超表面单元组合在一起形成一个微光谱仪 (光谱像素)。b, 从 a 中的光谱成像仪测量中进行数据立方体重建的两种方法, 包括使用迭代优化算法进行逐点光谱重建, 以及通过深度学习算法对整个数据立方体进行快速重建。c, 编码孔径快照光谱成像的示意图。光谱数据立方体首先由固定编码孔径 (掩模) 调制, 然后由色散器剪切, 最后由探测器。

3.2 硬件结构

基于超表面的光谱成像器的示意图如图2a所示。在图像传感器上集成了一层硅基超表面与微透镜层。超表面层由 360×440 个超表面单元组成, 这些单元是通过将 20×20 种超表面单元重复排列 18×22 次而获得的。不同的超表面单元具有独特的透射谱。每个超表面单元是一个具有自由形状的元子的周期阵列。相应的 400 种元子的周期和形状被优化以最大程度地增加透射谱之间的差异 [21]。每个超表面单元代表一个空间点, 光经过每个单元调制后将被检测到。对于某个点 (i, j) , 入射光的光谱可以从周围 N 个点 (即 $(i-n, j-n), (i-n, j-n+1), \dots, (i+n, j+n)$, 通常设置 $n = 2$) 的 $N = (2n+1)^2$ 个检测信号中重建。光谱重建是通过解决这样一个线性方程组实现的:

$$\begin{bmatrix} y^{[i-n,j-n]} \\ \vdots \\ y^{[i+n,j+n]} \end{bmatrix} = \begin{bmatrix} M_1^{[i-n,j-n]} & \dots & M_{N_\lambda}^{[i-n,j-n]} \\ \vdots & \ddots & \vdots \\ M_1^{[i+n,j+n]} & \dots & M_{N_\lambda}^{[i+n,j+n]} \end{bmatrix} \begin{bmatrix} x_1^{[i,j]} \\ \vdots \\ x_{N_\lambda}^{[i,j]} \end{bmatrix} \quad (1)$$

在这里, $y^{[i,j]}$ 表示在 (i, j) 点测得的信号, $M_k^{[i,j]}$ 表示在第 k 个光谱通道上 (即超表面单元的透射率) 在 (i, j) 点的调制强度, $x_k^{[i,j]}$ 是在 (i, j) 点目标光谱向量的第 k 个元素, N_λ 是光谱通道的总数。由于总共有 360×440 个点, 需要解决类似于式(1)的 158,400 组方程, 以重建所有点的光谱, 从而恢复整个光谱数据立方体, 这是非常耗时的。对于迭代 CS 算法, 假设 (i, j) 周围 N 个点的入射光谱与式(1)中的光谱相同, 导致在重建的光谱图像中出现马赛克效应。为了解决上述问题, 提出了利用深度学习算法 [22] 直接重建数据立方体的方法, 这受到了 CASSI 的启发, 如图2b所示。

CASSI 的基本原理如图2c所示, 其中光谱数据立方体通过一个固定的物理掩模 (编码孔

径)进行空间编码,然后通过色散元素进行光谱剪切,最后由探测器测量。因此,对于 CASSI,使用了一个固定的掩模和一个色散器,以在不同的光谱通道上实现不同的掩模。固定的掩模可以是传统的阻挡-解除编码孔径,也可以是有色编码孔径。在当前工作中,超表面层被采用以在不同波长下实现不同的掩模,其数学模型可以写为:

$$\begin{bmatrix} y^{[1,1]} \\ \vdots \\ y^{[N_x, N_y]} \end{bmatrix} = \begin{bmatrix} M_1^{[1,1]} & 0 & \dots & 0 & M_{N_\lambda}^{[1,1]} & 0 & \dots & 0 \\ 0 & M_1^{[2,1]} & \dots & 0 & 0 & M_{N_\lambda}^{[2,1]} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_1^{[N_x, N_y]} & 0 & 0 & \dots & M_{N_\lambda}^{[N_x, N_y]} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_\lambda} \end{bmatrix} \quad (2)$$

在这里, $x_k = [x_k^{[1,1]} x_k^{[2,1]} \dots x_k^{[N_x, N_y]}]^T$ 表示光谱数据立方体的第 k 帧, N_x, N_y 分别表示水平和垂直维度上的点数。考虑测量噪声时,式(2)可以以矢量化形式表示如下:

$$y = \Phi x + e \quad (3)$$

在这里, $y \in \mathbb{R}^{(N_x N_y)}$ 是压缩测量, $\Phi \in \mathbb{R}^{(N_x N_y \times N_x N_y N_\lambda)}$ 是感知矩阵, $x \in \mathbb{R}^{(N_x N_y N_\lambda)}$ 是目标信号, $e \in \mathbb{R}^{(N_x N_y)}$ 是测量噪声。作为原理验证,选择了 256×256 个超表面单元进行波长为 $450nm$ 到 $700nm$ 的光谱成像(即 $N_x = N_y = 256$, $N_\lambda = 26$)。

3.3 重建网络

在当前工作中,基于 ADMM 算法的深度展开网络(称为 ADMM-net)被用于数据立方体的重建。ADMM-net 的框架包括 K 个阶段($K = 12$),如图3a所示。与参考文献[18]类似,让 v 表示所需信号的估计值,并通过引入两个辅助变量 x 和 u ,在每个阶段更新变量的三个步骤如下:

$$x^{(k)} = [\Phi^T \Phi + \Upsilon^{(k)} I]^{-1} [\Phi^T y + (v^{(k-1)} + u^{(k-1)})] \quad (4)$$

$$v^{(k)} = D_k (x^{(k)} - u^{(k-1)}) \quad (5)$$

$$u^{(k)} = u^{(k-1)} - (x^{(k)} - v^{(k)}) \quad (6)$$

这里, $\Upsilon^{(k)} > 0$ 是另一个辅助可训练参数, I 是单位矩阵, D_k 表示一个去噪器。方程(4)表示线性投影。由于 Φ 是 N_λ 个对角矩阵的连接,可以 $\Phi \Phi^T$ 看出是一个对角矩阵。因此,方程(4)可以通过逐元运算而不是计算文献[23]中导出的巨大矩阵反演来有效地求解。方程(5)是由 CNN 执行的去噪过程,如图3a所示。采用了一个包含 15 层的 U-net 作为去噪器,其架构描述如图3b所示,其中跳过连接是残差学习的过程,是去噪器中必要的步骤。

从公开可用的高光谱数据集 CAVE 和 KAIST 构建了一个基本数据集,包含 262 个场景,大小为 $512 \times 512 \times 26$ 。基本数据集中选取的 15 个场景的 RGB 图像,通过国际照明委员会的色匹配函数[24]从光谱图像转换而来,如图3c所示。为了模型训练,从基本数据集中随机选择了 252 个场景,实施数据增强以获得大小为 $256 \times 256 \times 26$ 的 750,000 个样本。数据增强操作包括随机裁剪、旋转以及与 LED 光源或阳光光谱相乘。剩余的 10 个场景被降采样到 $256 \times 256 \times 26$ 的大小进行测试。此外,损失函数为地面实况和网络输出结果之间的均方根误

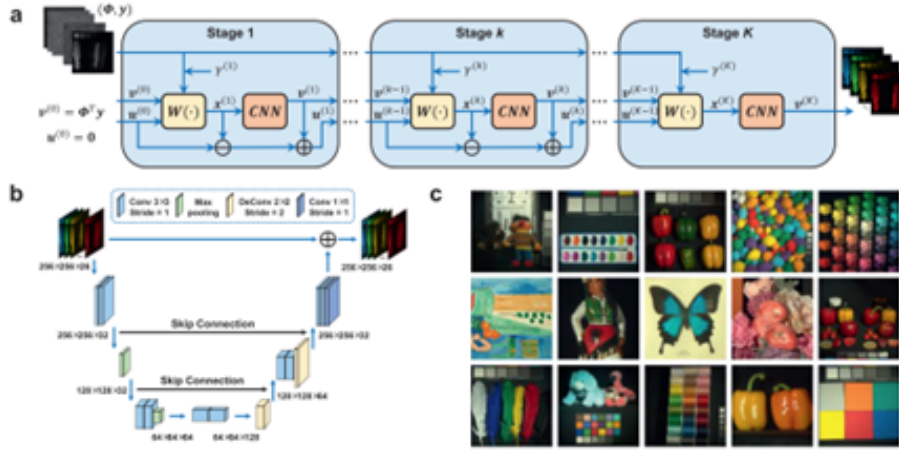


图 3. a, 带有 K 个阶段的 ADMM-net 的数据流图, 其中每个阶段包含一个线性投影 $W(\cdot)$, 表示方程 (4) 中的计算, 以及一个 CNN 去噪器。b, 用于去噪的 a 中使用的 U-net 的架构。c, 基本光谱图像数据集中的 15 个样本的 RGB 图像。

差。网络使用 PyTorch 上的 Adam 优化器进行训练, 使用 NVIDIA GeForce RTX 3080 GPU。总共进行 300 个 epochs, 由于 GPU 内存的限制, 批量大小设置为 4。初始学习率为 0.001, 并在每 20 个 epochs 时缩减为前一个的 90%。

4 复现细节

4.1 原论文的重构结果

该论文无源码, 因此不能实测论文中所述的重建结果。

根据论文描述, 首先在嘈杂条件下进行了网络训练和测试。由于为真实对象获取测量图像以训练网络需要很长时间, 因此采用了模拟测量的方法, 即使用通过方程 (3) 计算得到的带有预校准的测量图像 y , 以及数据集中的光谱图像 x 和适当的噪声分布 e 。 x 和 y 均被归一化到 $[0, 1]$ 。假设方程 (3) 中测量噪声向量 e 的每个元素都遭受独立的零均值高斯分布的影响 [25], 即 $e_i \sim N(0, \sigma_n^2)$ 。这里, 标准差 σ_n 表示噪声水平, 随机选择在 0 和 0.05 之间, 以增加对不同水平噪声的鲁棒性。在图 4 中给出了在 2% 和 4% 噪声水平下的两个测试场景的重建结果。对于 2% 噪声的蜡笔场景, 可以看到所选三个点的恢复光谱的平均保真度超过 99.97%, 如图 4a 的图例所示。当噪声水平增加到 4% 时, 可以看到 ADMM-net 仍然可以可靠地恢复光谱, 平均保真度为 99.87%。这里保真度的定义如下:

$$F(f_1, f_2) = \langle f_1, f_2 \rangle \quad (7)$$

其中, f_1 、 f_2 分别是 l_2 归一化的实际情况和重建结果。在图 4b 中显示了蜡笔高光谱数据的四个重建样本帧, 它们在不同噪声水平下与实际情况高度一致。对于玩具场景, ADMM-net 也能够提供准确的重建结果, 如图 4c 和 4d 所示。当噪声水平从 2% 增加到 4% 时, 与蜡笔场景相比, 光谱重建质量更加严重受损, 因为图 4c 中有更多的细节, 而图 4d 中的四个帧仍然以高质量恢复。

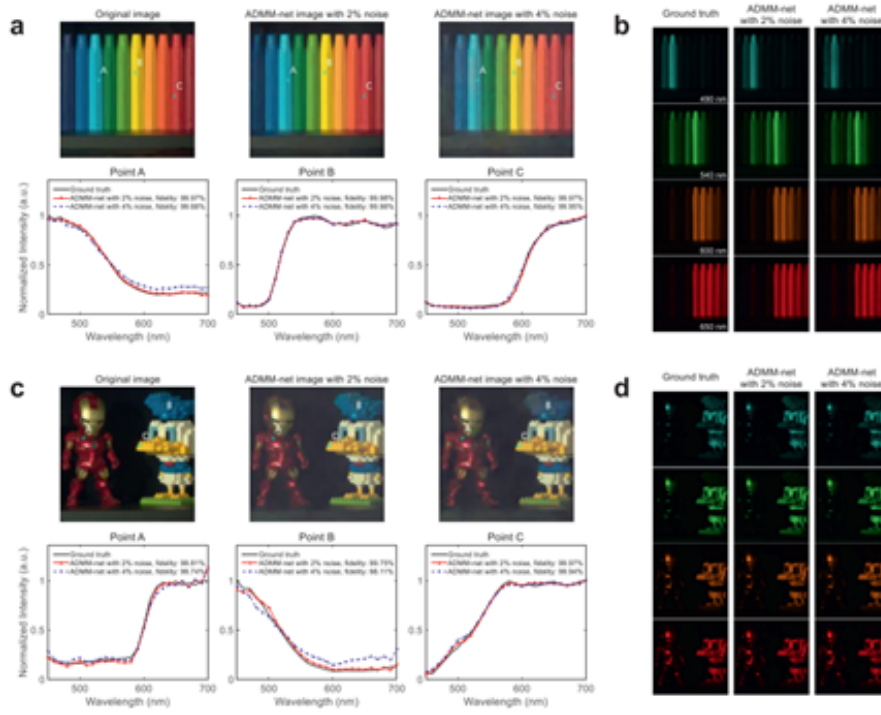


图 4. a-c, 分别在 2% 和 4% 噪声水平下, 对模拟蜡笔和玩具高光数据进行的重建结果。原始和重建的场景 RGB 图像显示在顶行, 大小为 256×256 像素。底行显示了三个选定点的 spectra, 图例中显示了重建 spectra 的保真度。b-d, 分别在 2% 和 4% 噪声水平下, 对模拟蜡笔和玩具高光数据进行的重建帧。

4.2 改进措施

1. 关于设计理念: 超表面滤镜是一种新型的光学器件, 它可以通过调整其表面的形状和大小来控制光的传播方向和相位, 从而实现对光的精确控制。然而, 由于其制造成本较高, 商业化生产面临着很大的挑战。在商业精细化生产及对成本的考虑下, 制作 400 组超表面滤镜是不可行的, 且论文中也没有给出 400 组滤镜的设计参数。因此, 在进行论文复现时, 我们随机生成了 9 条光谱曲线, 将其用作观测矩阵。这些曲线是从高斯分布中随机抽取的, 其随机值遵循均值 $\mu = 0$ 和标准差 $\sigma = 1$ 的高斯分布。如图5所示, 是论文复现时使用的 Φ 值。这些曲线可以在专业软件 FDTD 中转换为超表面的制作参数, 进而制作出符合曲线的硬件设备。复现论文的主要工作就聚焦于找到合适的曲线, 使能重构出的高光数据图片并质量更好。

2. 关于评价指标: 根据随机生成的 Φ 值, 我进行了 16 个点的随机采样, 将其原始值与重构结果进行对比, 并在图6中呈现了对比图。在评估两个像素是否一致时, 我们在论文中采用了保真率作为评价指标。保真率的优点在于其易于理解和解释, 对人们来说相对直观。然而, 需要注意的是, 保真率对于一些人眼感知差异较大的图像变化可能不够敏感。例如, 人眼对某些颜色和纹理变化更为敏感, 而保真率可能无法很好地捕捉这些特征。此外, 保真率对于亮度差异较大的图像更为敏感, 而对于颜色差异较大的图像可能不够敏感, 无法进行精细化的比较。因此, 为了获得更全面的图像质量评价, 我在此基础上引入了均方根误差 (RMSE) 作为额外的评价指标。这样可以更全面地考虑图像的各个方面, 提高评估的综合性和准确性。

3. 关于噪声: 论文中只在公式3中考虑了与信号无关的加性噪声, 而在实际拍摄过程中, 存在与信号相关的各种噪声, 例如散粒噪声等。因此, 在复现过程中, 我对原有的 ADMM-net

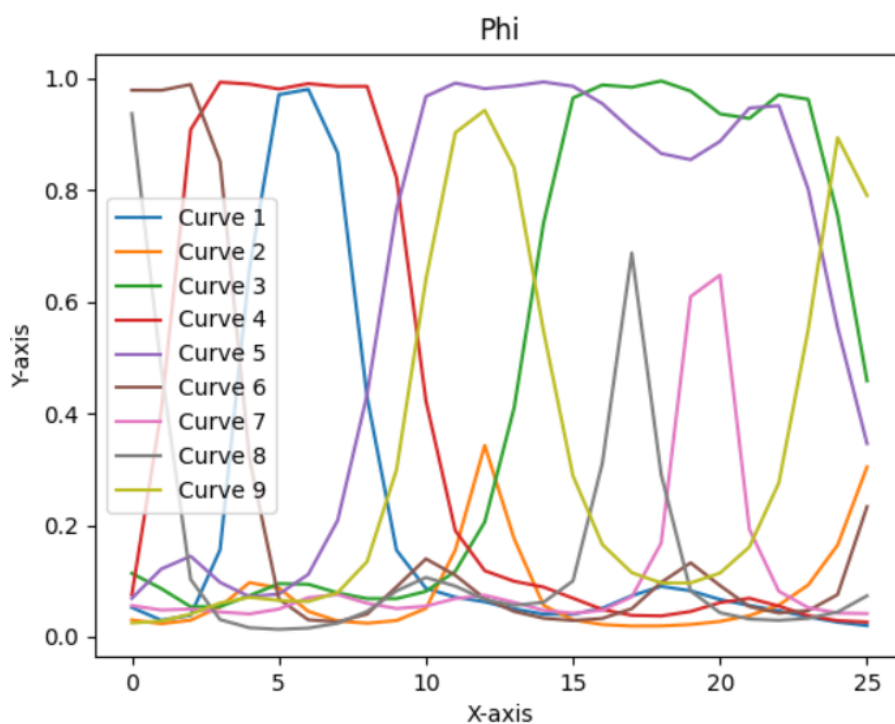


图 5. 随机生成的 Φ 值曲线

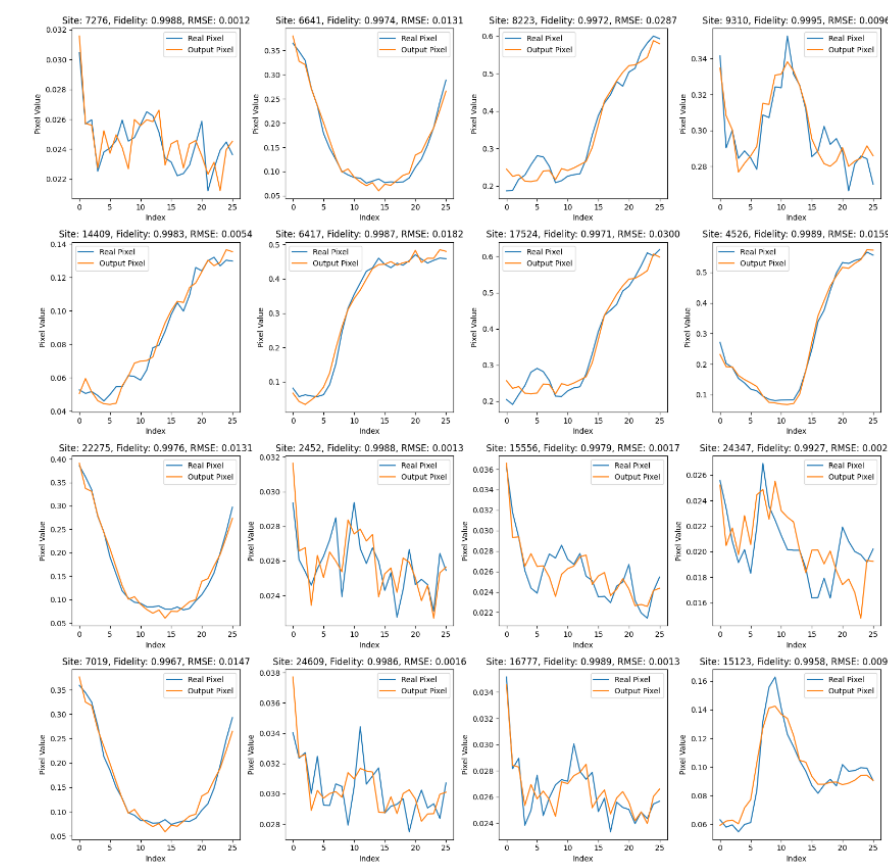


图 6. 在图片中随机采样 16 个像素，将其原始值与重构结果进行对比，使用保真率和均方根误差作为评价指标

网络进行了改进。我添加了噪声层并通过调整参数来控制各种噪声水平，添加原理如下。

设干净信号为 X ，从 X 中计算噪声方差 σ_{u,i_3}^2 和 σ_{t,i_3}^2 ，生成随机过程 u_{i_1,i_2,i_3} 和 t_{i_1,i_2,i_3} 的样本，这些样本对应于在以 db 为单位的指定信噪比下生成噪声张量 $N_{SD}(X)$ 和 N_{SI} [26]。其中， σ_{u,i_3}^2 表示平稳零均值不相关随机过程 u_{i_1,i_2,i_3} 的方差， σ_{t,i_3}^2 是在每个频带中服从零均值白高斯噪声的 t_{i_1,i_2,i_3} 的方差。如果在均匀像素上计算信号的方差，根据定义， $\sigma_{x_{i_1,i_2,i_3}}^2 = 0$ [27]；假设 x 、 u 和 t 是相互独立的，且 u 和 t 均为零均值且稳态的，噪声张量 $N(X)$ 的每个元素 $n(X)$ 的噪声方差可以写成 [26]：

$$\sigma_{n_{i_1,i_2,i_3}}^2(X) = x_{i_1,i_2,i_3} \cdot \sigma_{u,i_3}^2 + \sigma_{t,i_3}^2 \quad (8)$$

实际上，具有 $\sigma_{x_{i_1,i_2,i_3}}^2 = 0$ 的均匀像素可能非常罕见，理论期望通常用局部平均值来近似 [27]。噪声张量 $N(X)$ 的均方差由 SD 和 SI 噪声方差之和组成：

$$\sigma_{N(X)}^2 = \sigma_{N_{SD}(X)}^2 + \sigma_{N_{SI}}^2 \quad (9)$$

此外，可以用信号 X 的平均功率和信噪比 (db) 来表示：

$$\sigma_{N(X)}^2 = \bar{P}_X \cdot 10^{-(\frac{SNR}{10})} \quad (10)$$

其中， $\bar{P}_X = \frac{\|X\|^2}{I_1 I_2 I_3}$ 。假设有一个参数 α ，该参数控制两个噪声过程对噪声张量 $N(X)$ 的贡献，则：

$$\alpha = \frac{\sigma_{N_{SD}(X)}^2}{\sigma_{N_{SI}}^2} \quad (11)$$

由式(9)和式(11)可知， SI 和 SD 噪声方差均值可以表示为关于参数 α 的形式如下：

$$\begin{aligned} \sigma_{N_{SD}(X)}^2 &= \frac{\sigma_{N(X)}^2 \cdot \alpha}{\alpha + 1}, \\ \sigma_{N_{SI}}^2 &= \frac{\sigma_{N(X)}^2}{\alpha + 1} \end{aligned} \quad (12)$$

5 实验结果分析

在实验中，采用了从高斯分布中随机采样的 9 条曲线生成 Φ 值，并按 3*3 为一个小块进行循环。在评估结果时增加了均方根误差作为评价指标。在模型训练时，对不同类型和强度的噪声进行了系统性的测试，以更全面地评估网络在各种噪声条件下的性能表现。我测试了模型在峰值信噪比在 30、60、90 的情况下模型的重构效果，如图7所示。在信噪比为 30 时，平均保真率为 0.9875，均方根误差为 0.0132；在信噪比为 60 时，平均保真率为 0.9968，均方根误差为 0.0114；在信噪比为 90 时，平均保真率为 0.9979，均方根误差为 0.0078。其结果比论文中给出的值略差，这是由于更全面的考虑了噪声。通过加入信号相关噪声，使得网络更能模拟实际情况下的噪声环境，从而提高了模型的适用性和鲁棒性。这样的改进有助于更真实地反映实际应用场景中的噪声情况，提高了模型在实际环境中的泛化能力。

6 总结与展望

本次复现的论文是 Deep-learning based on-chip rapid spectral imaging with high spatial resolution。主要工作是在论文没有给出源码的情况下复现论文，并从设计理念、评价指标、

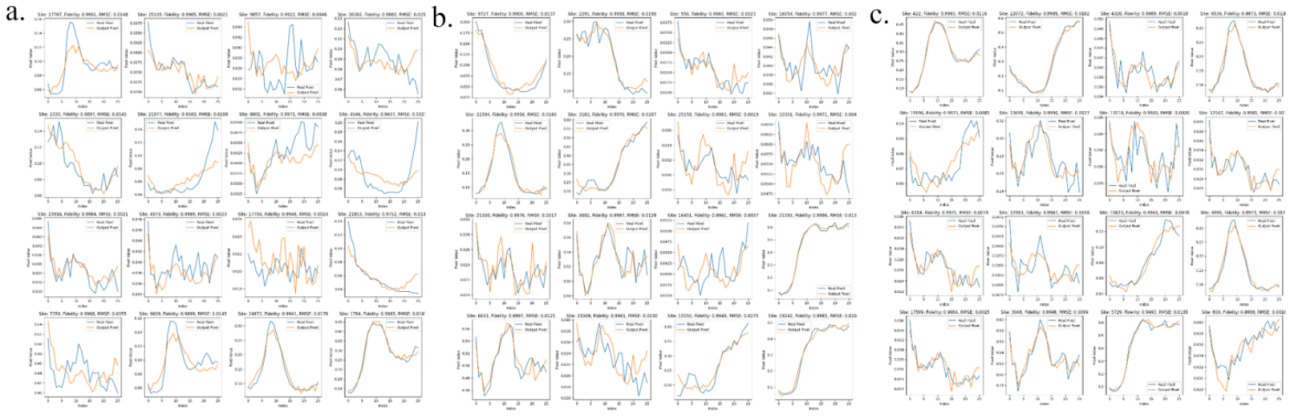


图 7. a. 在信噪比为 30 时，平均保真率为 0.9875，均方根误差为 0.0132；b. 在信噪比为 60 时，平均保真率为 0.9968，均方根误差为 0.0114；c. 在信噪比为 90 时，平均保真率为 0.9979，均方根误差为 0.0078。

噪声方面对论文进行了修改。实验结果比论文中给出的值略差，这是由于更全面的考虑了噪声。总体达到了预期结果。在此基础上，论文还有许多可以改进的方向，在未来可以考虑从重建具有更多波段的光谱图像、加大数据的规模精度和多样性、使用更先进的去噪器、引入 Transformer、3DCNN 等新型网络结构进一步优化光谱图像重建算法等方面继续修改。

参考文献

- [1] G. Lu and B. Fei, “Medical hyperspectral imaging: a review,” *Journal of Biomedical Optics*, vol. 19, no. 1, p. 010901, 2014. [Online]. Available: <https://doi.org/10.1117/1.JBO.19.1.010901>
- [2] X. Gu, Z. Han, L. Yao, Y. Zhong, Q. Shi, Y. Fu, C. Liu, X. Wang, and T. Xie, “Image enhancement based on in vivo hyperspectral gastroscopic images: a case study,” *Journal of biomedical optics*, vol. 21, no. 10, pp. 101412–101412, 2016.
- [3] V. Lebourgeois, A. Bégué, S. Labbé, B. Mallavan, L. Prévot, and B. Roux, “Can commercial digital cameras be used as multispectral sensors? a crop monitoring test,” *Sensors*, vol. 8, no. 11, pp. 7300–7322, 2008. [Online]. Available: <https://www.mdpi.com/1424-8220/8/11/7300>
- [4] Y.-Z. Feng and D.-W. Sun, “Application of hyperspectral imaging in food safety inspection and control: A review,” *Critical Reviews in Food Science and Nutrition*, vol. 52, no. 11, pp. 1039–1058, 2012, pMID: 22823350. [Online]. Available: <https://doi.org/10.1080/10408398.2011.651542>
- [5] B. S. J. Bahaiddin, Shah Mohammad and A. R. Winebarger, “The origin of reconnection-mediated transient brightenings in the solar transition region,” *Nature Astronomy*, vol. 5, no. 3, pp. 237–245, 2021.

- [6] M. B. Stuart, A. J. S. McGonigle, and J. R. Willmott, "Hyperspectral imaging in environmental monitoring: A review of recent developments and technological advances in compact field deployable systems," *Sensors*, vol. 19, no. 14, 2019. [Online]. Available: <https://www.mdpi.com/1424-8220/19/14/3071>
- [7] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, 1998. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0034425798000649>
- [8] P. Mouroulis, R. O. Green, and T. G. Chrien, "Design of pushbroom imaging spectrometers for optimum recovery of spectroscopic and spatial information," *Appl. Opt.*, vol. 39, no. 13, pp. 2210–2220, May 2000. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-39-13-2210>
- [9] C. Zhang, M. Rosenberger, A. Breitbarth, and G. Notni, "A novel 3d multispectral vision system based on filter wheel cameras," in *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2016, pp. 267–272.
- [10] N. Gat, "Imaging spectroscopy using tunable filters: a review," in *Wavelet Applications VII*, H. H. Szu, M. Vetterli, W. J. Campbell, and J. R. Buss, Eds., vol. 4056, International Society for Optics and Photonics. SPIE, 2000, pp. 50 – 64. [Online]. Available: <https://doi.org/10.1117/12.381686>
- [11] J. Antila, U. Kantojärvi, R. Mannila, A. Rissanen, I. Näkki, J. Ollila, and H. Saari, "Spectral imaging device based on a tuneable mems fabry-perot interferometer," in *Proceedings of SPIE 8374*, ser. Proceedings of SPIE, M. Druy, Ed. United States: International Society for Optics and Photonics SPIE, 2012, next-Generation Spectroscopic Technologies V ; Conference date: 23-04-2012 Through 24-04-2012.
- [12] N. Gat, G. Scriven, J. Garman, M. D. Li, and J. Zhang, "Development of four-dimensional imaging spectrometers (4D-IS)," in *Imaging Spectrometry XI*, S. S. Shen and P. E. Lewis, Eds., vol. 6302, International Society for Optics and Photonics. SPIE, 2006, p. 63020M. [Online]. Available: <https://doi.org/10.1117/12.678082>
- [13] B. A. A. J. Stoffels, J. and Jacobus, "Color splitting prism assembly." *United States patent US*, vol. 0844, no. 084, p. 180, 1978.
- [14] A. R. Harvey and D. W. Fletcher-Holmes, "High-throughput snapshot spectral imaging in two dimensions," in *Spectral Imaging: Instrumentation, Applications, and Analysis II*, R. M. Levenson, G. H. Bearman, and A. Mahadevan-Jansen, Eds., vol. 4959, International Society for Optics and Photonics. SPIE, 2003, pp. 46 – 54. [Online]. Available: <https://doi.org/10.1117/12.485557>

- [15] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [16] E. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [17] L. Huang, R. Luo, X. Liu, and X. Hao, “Spectral imaging with deep learning,” *Light: Science & Applications*, vol. 11, no. 1, p. 61, March 16 2022. [Online]. Available: <https://doi.org/10.1038/s41377-022-00743-6>
- [18] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, “Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum,” *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–2253, 2010.
- [19] I. Choi, D. S. Jeon, G. Nam, D. Gutierrez, and M. H. Kim, “High-quality hyperspectral reconstruction using a spectral prior,” vol. 36, no. 6, nov 2017. [Online]. Available: <https://doi.org/10.1145/3130800.3130810>
- [20] J. Liang, J. Zhou, L. Tong, X. Bai, and B. Wang, “Material based salient object detection from hyperspectral images,” *Pattern Recognition*, vol. 76, pp. 476–490, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0031320317304806>
- [21] J. Yang, K. Cui, X. Cai, J. Xiong, H. Zhu, S. Rao, S. Xu, Y. Huang, F. Liu, X. Feng, and W. Zhang, “Ultraspectral imaging based on metasurfaces with freeform shaped meta-atoms,” *Laser & Photonics Reviews*, vol. 16, no. 7, p. 2100663, 2022. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.202100663>
- [22] X. Yuan, D. J. Brady, and A. K. Katsaggelos, “Snapshot compressive imaging: Theory, algorithms, and applications,” *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 65–88, 2021.
- [23] X. Yuan, “Generalized alternating projection based total variation minimization for compressive sensing,” in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 2539–2543.
- [24] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.
- [25] Y. Liu, X. Yuan, J. Suo, D. J. Brady, and Q. Dai, “Rank minimization for snapshot compressive imaging,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2990–3006, 2019.

- [26] X. Liu, S. Bourennane, and C. Fossati, “Reduction of signal-dependent noise from hyperspectral images for target detection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5396–5411, 2014.
- [27] L. Alparone, M. Selva, B. Aiazzi, S. Baronti, F. Butera, and L. Chiarantini, “Signal-dependent noise modelling and estimation of new-generation imaging spectrometers,” in *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009, pp. 1–4.