光电技术及应用

DOI：10.16818/j.issn1001－5868.20250220001

基于动态多层次特征提取的高光谱解混方法

陈丽媛，杨华东*

（沈阳理工大学 信息科学与工程学院，沈阳 110159）

摘　要： 针对基于视觉Transformer进行高光谱解混时，固定大小的Patch划分方式可能导致图像局部结构的坍塌和图像间语义不一致的问题，提出一种数据驱动的方法。该方法旨在根据输入的高光谱图像自适应调整Patch划分的位置和尺度，并通过引入多尺度变形注意力机制有效捕获图像中的局部特征和空间位置信息，同时通过多层次特征提取策略，全面挖掘图像中蕴含的丰富特征。在多个真实数据集上的实验对比结果表明，所提方法能够精确估计丰度图和端元光谱，并在不同场景下，特别是在处理复杂地物光谱特征时，展现出优异的泛化能力和稳定性。

关键词： 高光谱图像；视觉Transformer；自适应图像划分；多尺度变形注意力机制

中图分类号： TP751　**文章编号：** 1001－5868(2025)03－0557－10

Hyperspectral Unmixing Based on Dynamic Multilevel Feature Extraction

CHEN Liyuan, YANG Huadong

(**School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, CHN**)

Abstract: The fixed-size patch division method may cause the local structure of an image to collapse and induce inconsistency in semantics between images during visual Transformer-based hyperspectral unmixing. Therefore, this study proposes a data-driven method that adaptively adjusts the position and scale of patch division according to the input hyperspectral image and effectively captures the local features and spatial positions in the image via a multiscale deformation attention mechanism. In addition, a multilevel feature extraction strategy is used to comprehensively mine the rich features contained in the image. A comparison of experimental results obtained for multiple real datasets shows that the proposed method can accurately estimate the abundance map and endmember spectrum. Moreover, it exhibits excellent generalization ability and stability in different scenarios, especially during the spectral feature processing of complex objects.

Key words: hyperspectral images; visual Transformer; adaptive image segmentation; multi-scale deformation attention mechanism

0　引言

高光谱遥感作为遥感领域的重要分支，其获得的数据称为高光谱遥感数据[1]。相较于传统的遥感数据，高光谱遥感图像具有较高的光谱分辨率和图谱合一的特点[2]，目前已成为遥感工程应用中的支撑数据之一[3]。高光谱遥感数据通常由成像光谱仪获得[4]，但成像光谱仪的空间分辨率较低，混合像元问题普遍存在并严重影响了地物目标识别和分类的精度[5]。解决混合像元问题最有效的方法是高光谱解混，发展有效的高光谱解混方法已经成为研究的重要环节[6]。

收稿日期：2025－02－20.

基金项目：辽宁省属本科高校基本科研业务费专项资金资助项目（LJ212410144018，SYLUGXRC202258）.

* **通信作者：**杨华东

现有的高光谱解混方法大致分为四类，分别是基于几何学的方法、基于统计学的方法、基于稀疏回归的方法及基于神经网络的方法[7]。其中，基于几何学的方法通过几何分析和端元成员选择来提取地物成分[8]；基于统计学的方法利用概率模型和统计推断来估计丰度[9]；基于稀疏回归的方法通过构造稀疏解来解混光谱混合问题[10]。这些方法在一定程度上具有较高的计算效率和较好的可解释性，然而，其往往依赖人工选择端元成员或依赖固定的假设，在复杂场景中的应用受到较大限制[11]。此外，固定的假设使得这些方法在面对高度复杂的地物成分或极端环境下，缺乏足够的灵活性，无法适应数据中的变化和不确定性。

神经网络一向被用于高光谱图像的解混模型中，但由于模型过于复杂、调参困难等问题难以解决，神经网络未能取得很好的成果[12]。自从自编码网络被运用于高光谱图像的解混之中，神经网络便开始发挥其强大的学习能力，取得了优于传统方法的解混效果。在现有的神经网络模型中，包括通过卷积神经网络（Convolutional Neural Network，CNN）进行有监督解混的方法[13]，以及通过循环神经网络（Recurrent Neural Network，RNN）捕获光谱通道间的相关信息的方法[14]。而在无监督解混中，应用最广泛的是自编码器（Autoencoder，AE）网络[15]。与传统方法相比，自编码网络能自动从高光谱数据中提取关键特征，避免了人工选择端元的复杂性。此外，自编码网络能够适应复杂的光谱混合模式，并能在无监督条件下训练，无需大量标注数据，这些优势使自编码器在高光谱解混中表现出更高的灵活性和鲁棒性。

近年来，高光谱解混领域的研究进一步深化。Deng等提出的DiffHU方法[16]，利用扩散模型过程模拟光谱混合，逐步优化丰度估计，在复杂混合场景中表现出色，但计算成本较高。Chen等提出双流协作网络（TCCU-Net）[17]，通过融合Transformer编码器和CNN编码器的输出弥合编码器和解码器之间的语义鸿沟，但其需并行运行双流编码器，导致计算量增多且在资源受限场景中部署困难。Gao等提出的空间-光谱自适应融合网络（SSAF-Net）[18]通过注意力机制自适应融合空间-光谱特征，且在无监督框架中显式建模端元变异性，提升了复杂场景的解混鲁棒性，但其端元扰动模型的假设较强且空间注意力对细小地物的敏感性不足。

Google于2017年提出的Transformer模型在NLP领域取得了巨大成功，这促使研究者尝试将其用于CV领域。视觉Transformer（ViT）模型将Transformer架构应用于计算机视觉任务，并取得了显著的成绩[19]。然而，其将图像划分为固定大小的非重叠图像块，尽管在一定程度上简化了图像处理过程，但在处理高光谱图像时，局部上下文信息丢失、边界效应以及无法适应多样化图像特性等问题极大影响了模型的性能，特别是在需要精确提取空间和光谱特征的高光谱解混任务中。

为提高得到的丰度图和端元光谱的质量，针对以上问题，本文提出动态的图像划分策略，通过自适应的方式调整patch的大小和位置，避免了固定块划分带来的信息丢失和边界效应。具体而言，通过卷积层对输入的高光谱图像进行初步划分，并根据图像尺寸和预设参数动态调整Patch的大小和数量。每个patch的参考位置由锚点网络提供，同时结合相对位置编码预测目标框与锚点之间的偏移量。为了适应变形或位移问题，引入偏移量预测机制对每个Patch进行微调。多尺度变形注意力机制进一步根据预测的偏移量生成采样位置，并加权融合图像特征，提升模型对空间变形和复杂物体的处理能力。

所提出的动态多层次特征提取方法具有以下优势：（1）动态适应性：通过锚点网络和偏移量预测机制动态调整图像划分策略，避免了固定块划分导致的信息丢失问题；（2）多尺度特征融合：结合卷积层与多尺度变形注意力机制，在保持计算效率的同时显著提升了对复杂地物边界的解混精度；（3）层级递进的单流动态融合：通过采用层级递进的单流动态融合网络，在保持光谱特征完整性的同时降低了计算效率，有效解决了传统方法中特征拼接导致的信息损失问题。

1 动态多层次高光谱解混网络

所提出的高光谱解混网络结构如图1所示，整体可分为图像预处理、多层次特征提取以及解码三个部分。在多层次特征提取模块，采用动态图像块划分（DePatch）策略，帮助模型更好地学习局部的光谱和空间信息，增强模型对细节的处理能力。最后，通过解码器对特征进行解码，得到最终的丰度图和端元光谱。

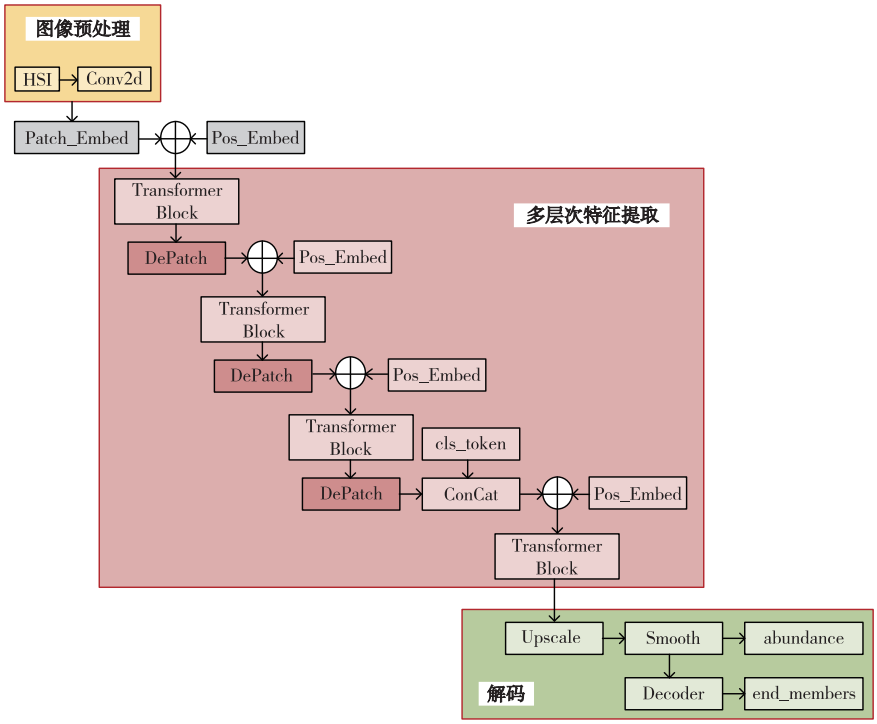


图 1　高光谱解混网络

1.1　图像预处理

由于高光谱图像的高维性，直接对其进行处理会导致巨大的计算开销和内存负担。为减少后续网络层计算负担的同时保留与解混任务相关的关键信息，在网络前端使用卷积层逐步减少输入高光谱数据的维度以实现特征压缩和降维处理。

1.2　动态图像块划分（DePatch）

为了有效处理不同尺度和形态的目标，并增强对高光谱图像中多个地物目标的解混能力，采用动态的块划分方式，根据输入图像的内容对图像划分区域进行自适应调整[20]，从而提升模型对复杂结构和多尺度信息的处理能力。所提出的动态图像块划分机制如图 2 所示。

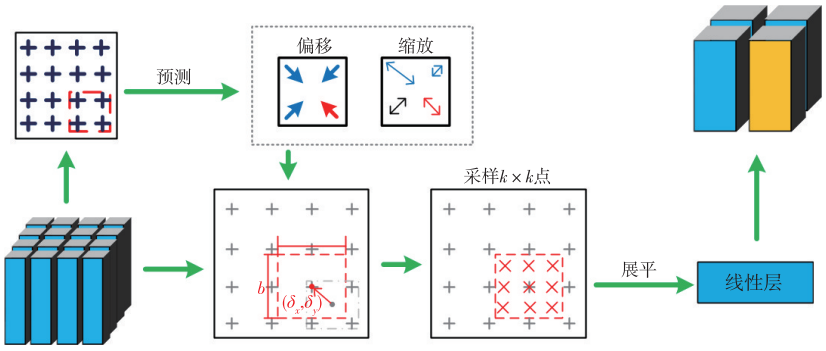


图 2　DePatch 模块[20]

1.2.1　自适应图像块划分

根据输入的高光谱图像 $\boldsymbol{X}\in R^{H\times W\times C}$ 生成一组均匀分布的锚点 $A=\left\{\left(x_{ct}^{(i)},y_{ct}^{(i)}\right),1\leqslant i\leqslant N\right\}$（$N$ 为预设锚点数量），其作用是为每个目标框提供一个起始位置，模型根据这些锚点的分布对图像中的目标进行预测。这种均匀分布能够确保目标框覆盖到图像的各个区域，从而提高目标检测的准确性。

为进一步提高目标框的精确度，通过相对位置编码预测锚点与真实目标框之间的偏移量，确保锚点能根据目标的位置和尺度进行调整。这种相对偏移量的预测使得模型能更精确地拟合目标的实际位置和大小，尤其在面对形态和尺度差异较大的目标时，对目标框的尺寸和位置进行动态调整。

为生成更加精细的目标框，在基于相对位置编码解码目标框坐标的过程中额外考虑目标框的尺

寸信息，即高度和宽度。同时，引入可学习的框偏置参数，通过该参数调整目标框的宽度和高度，使得模型在处理不同形态目标时具备更好的适应性，从而提升模型的表达能力。对解码后的框坐标以插值的方式生成细粒度的网格点，从而提高对目标区域的细节描述能力。

假设$\left(\boldsymbol{x}_{ct}^{(i)},\boldsymbol{y}_{ct}^{(i)}\right)$为第$i$个锚点的中心坐标，$\left(\boldsymbol{x}_1^{(i)},\boldsymbol{y}_1^{(i)}\right)$为动态调整后第$i$个图像块的左上角坐标，$\left(\boldsymbol{x}_2^{(i)},\boldsymbol{y}_2^{(i)}\right)$为动态调整后第$i$个图像块的右下角坐标，$\left(\boldsymbol{\delta}_x^{(i)},\boldsymbol{\delta}_y^{(i)}\right)$为预测的$x/y$方向偏移量，$\left(\boldsymbol{s}_w^{(i)},\boldsymbol{s}_h^{(i)}\right)$为预测的图像块宽度，则锚点的位置和尺度通过以下公式动态调整：

$$\left(\boldsymbol{x}_1^{(i)},\boldsymbol{y}_1^{(i)}\right)=\left(\boldsymbol{x}_{ct}^{(i)}+\boldsymbol{\delta}_x^{(i)}-\boldsymbol{s}_w^{(i)/2},\boldsymbol{y}_{ct}^{(i)}+\boldsymbol{\delta}_y^{(i)}-\boldsymbol{s}_h^{(i)/2}\right) \quad (1)$$

$$\left(\boldsymbol{x}_2^{(i)},\boldsymbol{y}_2^{(i)}\right)=\left(\boldsymbol{x}_{ct}^{(i)}+\boldsymbol{\delta}_x^{(i)}+\boldsymbol{s}_w^{(i)/2},\boldsymbol{y}_{ct}^{(i)}+\boldsymbol{\delta}_y^{(i)}-\boldsymbol{s}_h^{(i)/2}\right) \quad (2)$$

其中偏移量$\left(\boldsymbol{\delta}_x^{(i)},\boldsymbol{\delta}_y^{(i)}\right)$和尺度$\left(\boldsymbol{s}_w^{(i)},\boldsymbol{s}_h^{(i)}\right)$通过轻量级预测网络获得：

$$\left(\boldsymbol{\delta}_x^{(i)},\boldsymbol{\delta}_y^{(i)}\right)=\tanh\left(\boldsymbol{W}_{\text{offset}}\cdot f_p^{A^{(i)}}\right) \quad (3)$$

$$\left(\boldsymbol{s}_w^{(i)},\boldsymbol{s}_h^{(i)}\right)=ReLU\left(\tanh\left(\boldsymbol{W}_{\text{scale}}\cdot f_p^{A^{(i)}}+\boldsymbol{b}_{\text{scale}}\right)\right) \quad (4)$$

式中，$f_p(\cdot)$为特征提取函数，采用单层线性变换；$\boldsymbol{W}_{\text{offset}}\in R^{2\times d}$，$\boldsymbol{W}_{\text{scale}}\in R^{2\times d}$为可学习参数；$\boldsymbol{b}_{\text{scale}}$确保训练初始阶段保持原始划分。

1.2.2 偏移量预测与图像特征提取

传统基于Transformer的方法通常采用均匀网格划分策略，然而这种刚性划分难以适应实际场景中复杂的地物空间分布特性，往往导致部分图像块跨越多个端元区域，从而引入显著的混合像元效应。此外，关键光谱特征在空间上可能呈现局部聚集分布，均匀采样会稀释这些判别性特征的信息密度。针对上述问题，本文提出的动态图像块划分机制通过引入可学习的偏移量预测模块，并结合端到端训练，使图像块中心向信息密度较高的区域（如端元边界、高对比度区域）自适应偏移。这种动态调整显著提升了局部特征的纯度，更符合线性混合模型的局部线性假设。

为了将输入的高光谱图像划分为具有空间局部特征的图像块，使用卷积层对图像进行划分并根据输入图像的尺寸和预设参数调整图像块的大小和数量，将划分后的图像块映射到一个高维的嵌入空间。针对不同区域间存在变形或位移的问题，引入偏移量预测机制，通过一个额外的线性层预测每个图像块的偏移量，从而能够在空间上对图像的局部区域进行微调，为图像特征提取提供了空间上的调整信息。这一过程增强了模型的空间定位能力，使得模型能够更加精细地处理图像中的细节。

对于每个动态划分的图像块，采用双线性插值采样$k\times k$个特征点：

$$\tilde{\boldsymbol{\alpha}}^{(j)}=\sum(q_x,q_y)G(p_x^{(j)},p_y^{(j)};\boldsymbol{q}_x,\boldsymbol{q}_y)\cdot\boldsymbol{X}(\boldsymbol{q}_x,\boldsymbol{q}_y) \quad (5)$$

式中，$\tilde{\alpha}^{(j)}\in R^C$为第$j$个采样点的特征向量，$(p_x^{(j)},p_y^{(j)})$为第$j$个采样点的分数坐标，$(q_x,q_y)$为整数空间位置坐标，$X(q_x,q_y)\in R^C$为输入图像在$(q_x,q_y)$位置的特征，$G(\cdot)$为双线性插值核函数，表示为：

$$G(p_x,p_y;q_x,q_y)=\max\left(0,1-\left|p_x-q_x\right|\right)\cdot\max\left(0,1-\left|p_y-q_y\right|\right) \quad (6)$$

采样后的特征通过线性投影得到最终嵌入表示：

$$\boldsymbol{z}^{(i)}=\boldsymbol{W}\cdot\text{Concat}(\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(k\times k)})+\boldsymbol{b} \quad (7)$$

式中，$\boldsymbol{z}^{(i)}\in R^d$为第$i$个图像块的嵌入表示，$d$为第$i$个图像块在隐空间中的表示维度，$\boldsymbol{W}\in R^{(d\times k^2C)}$为投影矩阵，$\boldsymbol{b}\in R^d$为偏置项，$k$为每块采样点数。

1.2.3 多尺度变形注意力机制

为了进一步提升特征提取的灵活性和多样性，结合了多尺度变形注意力机制方法，根据预测的偏移量生成图像采样位置，并计算每个采样点的注意力权重，将采样位置与注意力权重一起输入到变形注意力层，通过加权融合输入图像特征。在此过程中，将变形注意力操作的结果通过全连接层映射到目标嵌入空间，并应用层归一化进行规范化处理。多尺度变形注意力机制有效克服了卷积神经网络在处理不规则空间特征时的局限性，能够捕获图像中的局部特征和空间位置信息，适应图像的空间变形，从而提高高光谱图像中复杂物体的解混性能。

其中，多尺度注意力权重的计算方式为：

$$Attention(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=Softmax((\boldsymbol{Q}\boldsymbol{K}^T)/\sqrt{d}+\boldsymbol{B})\boldsymbol{V} \quad (8)$$

式中，$\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}\in R^{(N\times d)}$为查询、键、值矩阵，$d$为每个头的维度，$\boldsymbol{B}$为相对位置编码矩阵，通过下式计算：

$$\boldsymbol{B}=\boldsymbol{W}_{\text{pos}}\cdot PE(\boldsymbol{\delta P}) \quad (9)$$

式中，$W_{\text{pos}}\in R^{(H\times d)}$为位置编码投影矩阵，$H$为注意力头数，$PE(\cdot)$为正弦位置编码函数，$\boldsymbol{\delta P}$为预测的偏移量矩阵。

1.3 多层次特征提取

对经降维处理的高光谱图像划分为多个尺度的Patch，并通过逐层嵌入和变形操作捕获多尺度的空间信息。在模型中，输入图像经过四个不同尺寸的Patch Embeding层进行处理，每个层对应于不同的空间尺度，逐步提取图像的高维特征。在每个阶段，除基础的Patch嵌入操作以外，还引入了位置编码以帮助模型学习空间位置信息，从而强化高光谱图像中的各个波段之间的空间一致性。为了进一步增强模型的表达能力并提高其泛化能力，模型采用了逐层正则化方法，并通过多尺度变形注意力机制提升了对不同空间尺度和形状变化的适应能力。

通过在各个阶段逐层应用Transformer编码器，利用自注意力机制对局部和全局特征进行建模，从而充分挖掘图像中蕴含的复杂关系。每个Transformer Block都包含了多头注意力（Multi-Hear Attention）、前馈神经网络（MLP）和LayerNorm层。随着网络的深度和计算需求逐渐增大，通过调整每一个多头注意力的头数和多层感知机比率，可以有效控制注意力机制的计算复杂度和特征变换的非线性能力。此外，引入随机跳过层的机制，以提升模型的泛化能力。

1.4 解码

经过所有Transformer编码器层的处理后，其输出通过cls token池化来获取图像的全局表示，并通过上采样操作恢复每个端元的空间分布信息。经过上采样后的特征图通过一个卷积层进行平滑，以提高结果的空间一致性，并通过Softmax操作将每个像素的每个端元的权重映射为一个概率分布。通过解码器将上采样和平滑后的特征图重建为原始高光谱图像，得到端元光谱。

2 实验与分析

2.1 数据集

为了验证所提方法的有效性与适用性，选取了三个具有代表性的高光谱数据集进行实验评估，分别为Samson，Apex和Washington DC Mall数据集。

Samson数据集[21]：包含95×95个像元，每个像元包含156个波段，覆盖波长范围为401～889 nm，包括树木、水、土壤三个端元。

Apex数据集[22]：包含110×110个像元，每个像元包含285个波段，覆盖波长范围为413～2 420 nm，包括树木、水、道路、屋顶四个端元。

Washington DC Mall数据集：包含290×290个像元，每个像元包含191个波段，覆盖波长范围为400～2 400 nm，包括树木、水、道路、屋顶、草、小道六个端元。

2.2 数据集划分策略和参数设置

2.2.1 数据集划分策略

采取空间-光谱双重交叉验证策略以确保模型评估的严谨性。

对于空间维度划分，采用区块划分方法，将原始高光谱图像块划分为非重叠的区块，按照7∶3的比例随机分配训练集与测试集。该策略有效避免了空间自相关对模型性能评估的潜在影响。

针对光谱维度划分，通过分层随机抽样方法，从原始波段中随机选取80%作为训练波段，剩余20%作为验证波段。该划分保持每个端元特征波段的完整性。

此外，所有划分过程均采用固定随机种子确保划分结果的一致性。

2.2.2 参数设置

表1详细列出了本研究的核心参数配置体系，具体包括端元数量P、波段维度L、Patch尺寸$Patch$、重建损失权重β、SAD损失权重γ、训练轮数$Epoch$、学习率LR以及权重衰减率WD。

表1 参数设置

参数	数据集		
	Samson	Apex	DC
P	3	4	6
L	156	285	191
$Patch$	5×5	5×5	10×10
β	5×10^{3}	7×10^{3}	5×10^{3}
γ	3×10^{-2}	7×10^{-2}	1×10^{-4}
$Epoch$	200	200	100
LR	4×10^{-3}	7×10^{-3}	7×10^{-3}
WD	4×10^{-5}	2×10^{-5}	3×10^{-5}

2.3 评价指标

为了全面评估所提方法的性能，选择了均方根误差（Root Mean Square Error，RMSE）和光谱角距离（Spectral Angle Distance，SAD）作为主要的评价指标，其定义如公式（10）和（11）。这些指标能够从不同的角度反映解混算法在光谱和空间域的准确性和鲁棒性。

$$SAD(\hat{\boldsymbol{m}}_i,\boldsymbol{m}_i)=\frac{1}{p}\sum_{i=1}^{p}\arccos\left(\frac{\boldsymbol{m}_i^{T}\boldsymbol{m}_i}{\|\boldsymbol{m}_i\|\bullet\|\boldsymbol{m}_i\|}\right) \tag{10}$$

$$RMSE(\hat{\boldsymbol{y}}_j,\boldsymbol{y}_j)=\sqrt{\frac{1}{N}\sum_{j=1}^{N}\|\hat{\boldsymbol{y}}_j-\boldsymbol{y}_j\|_2^{2}} \tag{11}$$

2.4 实验结果

为了验证所提方法的有效性，将其与五种代表性的传统高光谱解混方法进行比较，分别为四种传统的解混方法 Coolab[23]，FCLSU[24]，NMF[25]，CyCU[26]以及一种基于深度学习的解混方法 DeepTrans[27]。下文详细列出了了不同方法在多个数据集上的性能表现，并综合评估了不同解混方法在不同数据集中的适应性和稳定性，为所提方法的优越性提供了有力的实验支持。对比实验结果参考文献[27]中的结果。

2.4.1 Samson数据集

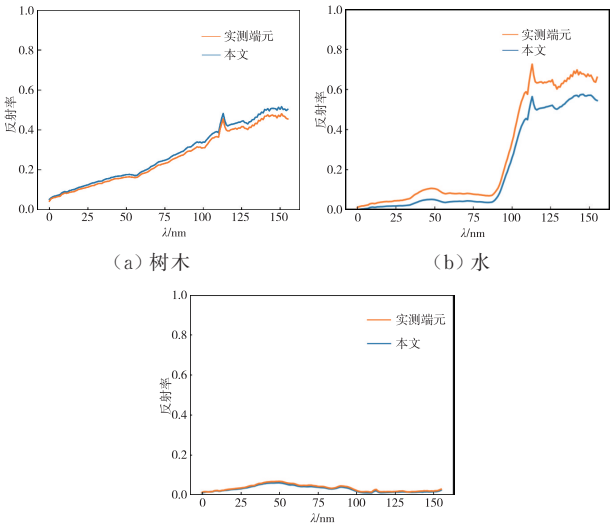
Samson数据集的均方根误差和光谱角距离结果如表2和3，其中红色加粗为最优结果，黑色加粗为次优结果。其可视化结果如图3和表4所示。

表2 Samson数据集RMES结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	**0.060 7**	0.065 3	0.146 6	0.138 6	0.063 9	0.078 1
水	0.118 1	0.149 2	0.206 3	0.265 4	**0.054 2**	0.107 5
土壤	0.150 6	0.176 6	0.201 1	0.241 7	0.093 7	**0.086 3**
整体	0.109 8	0.130 4	0.184 7	0.215 3	0.211 8	**0.090 6**

表3 Samson数据集SAD结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	0.083 2	0.165 5	0.123 9	0.151 7	0.090 9	**0.057 1**
水	0.140 2	0.091 6	0.520 1	0.208 1	0.107 5	**0.057 5**
土壤	0.015 5	0.136 1	0.039 1	0.114 4	0.037 1	**0.009 2**
整体	0.079 6	0.131 1	0.227 7	0.158 1	0.078 5	**0.041 2**



（a）树木 （b）水

（c）土壤

图3 Samson数据集端元光谱可视化结果

Samson数据集的实验结果表明，所提方法在丰度估计方面表现出色，除了对“树木”端元的估计略有不足外，其余结果均位于所有对比方法中的第一或第二，显示出其在处理复杂地物混合时的优势。在端元提取方面，所提方法结果均为最优，且平均SAD为0.041 2，显著优于次优方法，提升幅度达到47.5%，这表明所提方法在恢复真实端元光谱的准确性方面具有明显的优势。

2.4.2 Apex数据集

Apex数据集的均方根误差和光谱角距离结果如表5和6，其中红色加粗为最优结果，黑色加粗为次优结果。其可视化结果如图4和表7所示。

表4 Samson数据集丰度图可视化结果

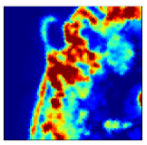
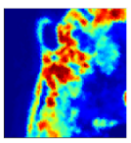
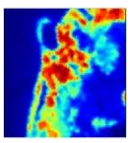
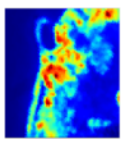
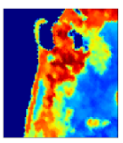
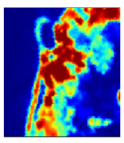
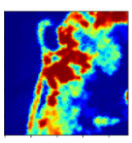
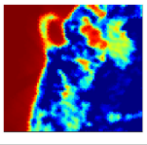
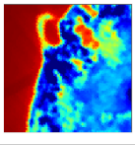
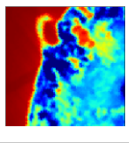
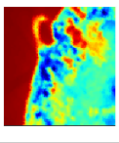
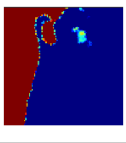
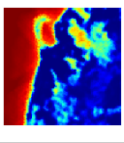
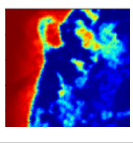
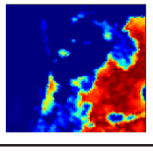
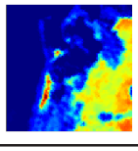
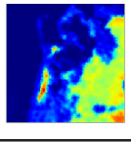
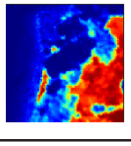
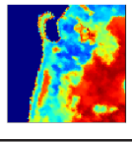
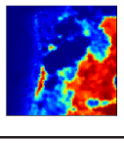
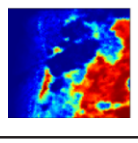
端元	解混方法						
	GT	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木							
水							
土壤							

表 5　Apex 数据集 RMES 结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	0.190 7	**0.094 4**	0.246 8	0.202 0	0.093 2	0.119 1
水	0.079 7	0.132 7	0.375 1	0.121 3	0.455 5	**0.097 5**
道路	0.307 8	0.233 1	**0.180 6**	0.292 1	0.440 7	0.176 2
屋顶	0.148 3	0.120 1	0.235 9	0.136 0	0.076 2	**0.109 6**
整体	0.181 9	**0.145 1**	0.259 6	0.187 9	0.266 4	0.125 6

表 6　Apex 数据集 SAD 结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	0.206 3	0.152 0	0.270 1	0.085 0	**0.131 1**	0.136 2
水	**0.513 7**	0.523 7	0.622 3	0.622 3	0.704 6	0.039 8
道路	0.677 2	**0.161 6**	0.400 3	0.454 3	0.943 0	0.090 0
屋顶	**0.100 2**	0.276 3	0.100 2	0.129 8	0.109 5	0.092 4
整体	0.374 4	**0.278 4**	0.348 3	0.322 9	0.472 1	0.089 6

Apex 数据集的实验结果表明，所提方法在丰度估计方面，除了“树木”端元外，其余结果均位于最优或次优位置。虽然“树木”端元的丰度估计结果相较于其他端元略显逊色，但这一差距并未影响整体解混性能，且仍在合理范围内。在端元提取方面，除“树木”端元外，所有端元的提取结果均为最优，且平均 SAD 为 0.089 6，相比次优结果提升了 67.8%，表明所提方法能够准确估计大部分地物的丰度，尤其是在光谱特征差异显著的地物类型中，能够有效分离出各端元。

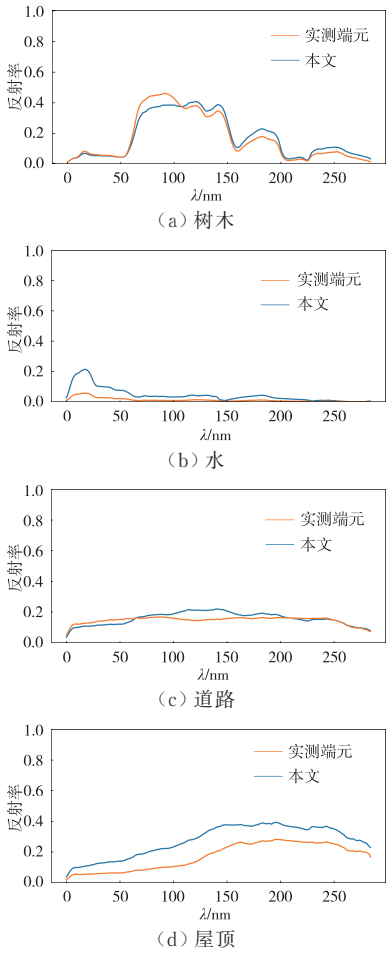


图 4　Apex 数据集端元光谱可视化结果

表 7　Apex 数据集丰度图可视化结果

端元	解混方法						
	GT	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木							
水							
道路							
屋顶							

2.4.3　DC 数据集

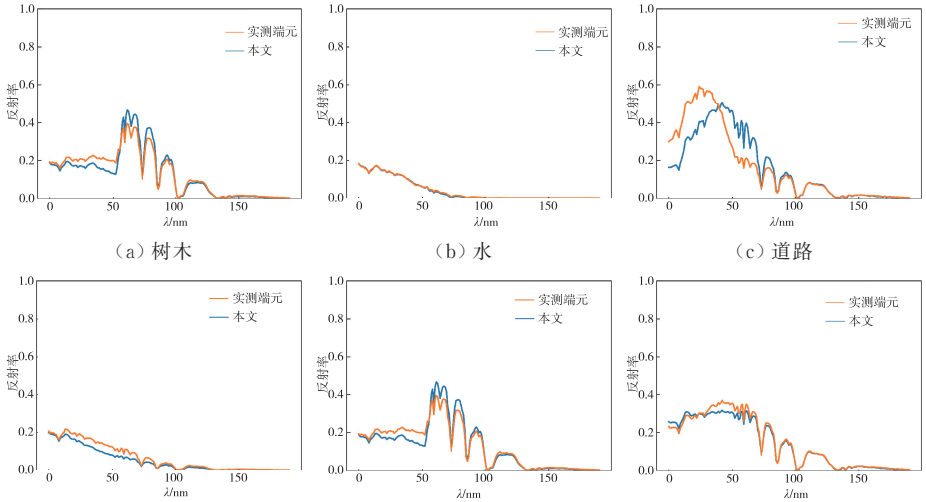
DC 数据集的均方根误差和光谱角距离结果如表 8 和 9，其中红色加粗为最优结果，黑色加粗为次优结果。其可视化结果如图 5 和表 10 所示。

表 8 DC 数据集 RMSE 结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	0.416 7	0.402 5	0.276 1	0.282 4	**0.161 0**	0.151 4
水	0.310 2	0.292 1	0.270 6	0.395 7	0.148 7	**0.163 2**
道路	0.226 3	0.175 7	0.235 1	0.254 5	0.248 0	**0.205 1**
屋顶	**0.043 7**	0.038 0	0.086 2	0.415 7	0.128 3	0.082 7
草	0.290 1	0.309 0	0.362 4	0.410 4	**0.197 2**	0.174 3
小道	0.187 5	0.123 0	0.101 1	0.207 2	0.151 2	**0.119 3**
整体	0.245 8	0.223 4	0.222 0	0.327 7	**0.172 4**	0.154 4

表 9 DC 数据集 SAD 结果对比

端元	解混方法					
	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木	0.333 5	0.288 3	0.450 7	0.270 4	**0.187 4**	0.054 1
水	**0.030 5**	0.776 6	0.673 6	0.420 5	0.035 1	0.026 7
道路	0.343 9	0.231 6	**0.224 3**	0.464 2	0.268 1	0.169 9
屋顶	0.033 1	**0.034 3**	0.207 8	0.950 0	0.423 7	0.375 6
草	0.317 1	0.317 0	0.195 2	0.089 5	0.283 5	**0.164 2**
小道	0.344 6	0.647 2	0.061 5	0.790 6	0.385 6	**0.086 0**
整体	**0.233 8**	0.383 3	0.302 2	0.497 6	0.263 9	0.145 5

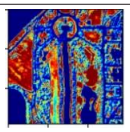
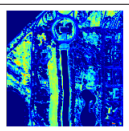
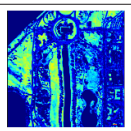
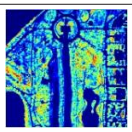
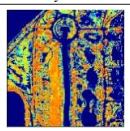
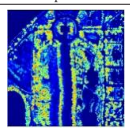
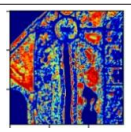
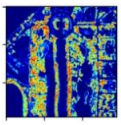
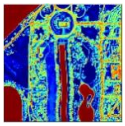
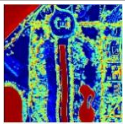
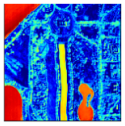
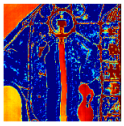
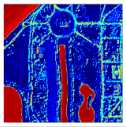
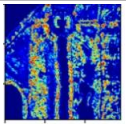
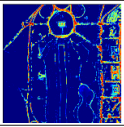
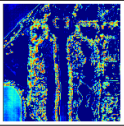
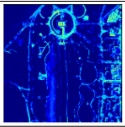
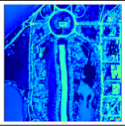
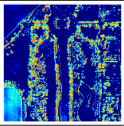
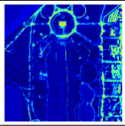
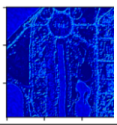
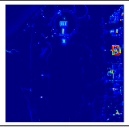
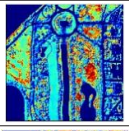
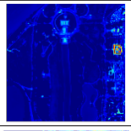
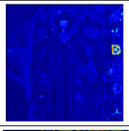
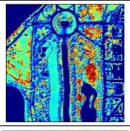
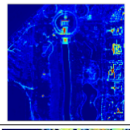
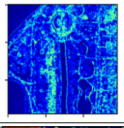
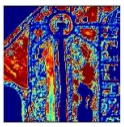
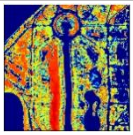
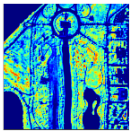
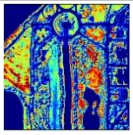
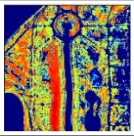
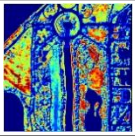
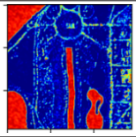
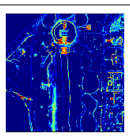
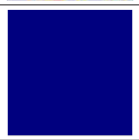
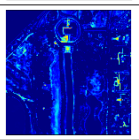
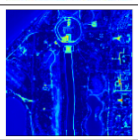


（a）树木 （b）水 （c）道路

（d）屋顶 （e）草 （f）小道

图 5 DC 数据集端元光谱可视化结果

表 10 DC 数据集丰度图可视化结果

端元	解混方法						
	GT	Coolab	FCLSU	NMF	CyCU	DeepTrans	本文
树木							
水							
道路							
屋顶							
草							
小道							

DC数据集的实验结果表明，在丰度估计方面，尽管平均RMSE值未达到最优水平，但其表现位居次优，且仅与最优结果相差较少。除“屋顶”端元外，其他所有端元的丰度估计结果均位于最优或次优位置，显示出所提方法在大多数地物丰度估计上的准确性和可靠性。在端元提取方面，所提方法的平均SAD值为0.145 5，相比次优方法提升了37.8%，显著提高了端元提取的准确性。除了“屋顶”端元外，所有其他端元的提取结果均为于最优或次优，进一步证明了所提方法在端元提取中的卓越表现和稳定性。

2.5　消融实验分析

为了深入评估动态图像块划分和多尺度变形注意力机制对高光谱解混性能的独立贡献及协同效用，本节设计了系统的消融实验。实验在Samson，Apex和DC三个典型的高光谱数据集上进行。各模块在不同数据集上的量化结果表11所示，其中红色加粗表示最优结果，黑色加粗表示次优结果。

表11　模块消融实验的量化实验结果

数据集	参数	DePatch	MSDA	DePatch+MSDA
Samson	*SAD*	**0.039 4**	0.041 4	**0.041 2**
	RMSE	0.115 1	**0.091 7**	**0.090 6**
Apex	*SAD*	**0.098 8**	0.099 7	**0.089 6**
	RMSE	0.143 9	**0.123 4**	**0.125 6**
DC	*SAD*	0.202 1	**0.182 9**	**0.145 5**
	RMSE	**0.132 0**	0.155 6	**0.154 4**

实验结果表明，虽然动态图像块划分模块和多尺度变形注意力机制模块各自都能显著提升解混性能，但二者协同工作时产生了显著的互补增强效应。

3　总结

本文提出一种基于动态图像划分策略的高光谱解混方法，结合了视觉Transformer模型进行端元提取和丰度估计。该方法通过数据驱动的方式自适应地从输入的高光谱图像中提取相关特征，进而增强了解混模型的灵活性与表达能力。在此过程中，所提出的动态图像划分策略不仅能够根据图像的不同特征区域动态调整划分方式，还通过多层次的特征提取机制深入挖掘高光谱图像中潜在的深层次信息，从而提高端元提取和丰度估计的精度。此外，通过在Samson，Apex和Washington DC Mall三个广泛使用的高光谱数据集上进行实验对比，所提出的方法在RMSE，SAD指标上均取得了显著的性能提升，证明了动态图像划分策略和视觉Transformer模型在高光谱解混任务中的优越性。这一方法为高光谱解混领域提供了一种新的思路和解决方案，具有较强的实际应用潜力和理论意义。

参考文献：

[1] 童庆禧，张　兵，张立福. 中国高光谱遥感的前沿进展[J]. 遥感学报，2016，20(5)：689-707.
Tong Qingxi，Zhang Bing，Zhang Lifu. Current progress of hyperspectral remote sensing in China[J]. Journal of Remote Sensing，2016，20(5)：689-707.

[2] 苏远超，许若晴，高连如，等. 基于深度学习的高光谱遥感图像混合像元分解研究综述[J]. 遥感学报，2024，28(1)：1-19.
Su Yuanchao，Xu Ruoqing，Gao Lianru，et al. Development of deep learning-based hyperspectral remote sensing image unmixing [J]. National Remote Sensing Bulletin，2024，28(1)：1-19.

[3] 张　达，郑玉权. 高光谱遥感的发展与应用[J]. 光学与光电技术，2013，11(3)：67-73.
Zhang Da，Zheng Yuquan. Hyperspectral remote sensing and its development and application review [J]. Optics & Optoelectronic Technology，2013，11(3)：67-73.

[4] 杜培军，夏俊士，薛朝辉，等. 高光谱遥感影像分类研究进展[J]. 遥感学报，2016，20(2)：236-256.
Du Peijun，Xia Junshi，Xue Zhaohui，et al. Review of hyperspectral remote sensing image classification[J]. Journal of Remote Sensing，2016，20(2)：236-256.

[5] 陈　晋，马　磊，陈学泓，等. 混合像元分解技术及其进展[J]. 遥感学报，2016，20(5)：1102-1109.
Chen Jin，Ma Lei，Chen Xuehong，et al. Research progress of spectral mixture analysis [J]. Journal of Remote Sensing，2016，20(5)：1102-1109.

[6] Nie J，Wu K，Li Y，et al. Advances in hyperspectral remote sensing for precision fertilization decision-making：a comprehensive overview [J]. Turkish Journal of Agriculture and Forestry，2024，48(6)：1084-1104.

[7] 万　升，杨　健，宫　辰. 基于图神经网络的高光谱图像分类研究进展[J]. 电子学报，2023，51(6)：1687-1709.
Wan Sheng，Yang Jian，Gong Chen. Advances of hyperspectral image classification based on graph neural networks [J]. Acta Electronica Sinica，2023，51(6)：1687-1709.

[8] Zhang W，Li X，Dou Y，et al. A geometry-based band selection approach for hyperspectral image analysis [J]. IEEE Transactions on Geoscience and Remote Sensing，2018，56(8)：4318-4333.

[9] Bioucas-Dias J M，Plaza A，Dobigeon N，et al. Hyperspectral unmixing overview：geometrical，statistical，and sparse

regression-based approaches [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2012, 5(2): 354-379.

[10] Huang J, Di W C, Wang J J, et al. Bilateral joint-sparse regression for hyperspectral unmixing [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 10147-10161.

[11] Hong D, He W, Yokoya N, et al. Interpretable Hyperspectral Artificial Intelligence: when nonconvex modeling meets hyperspectral remote sensing [J]. IEEE Geoscience and Remote Sensing Magazine, 2021, 9(2): 52-87.

[12] Bejani M M, Ghatee M. A systematic review on overfitting control in shallow and deep neural networks [J]. Artificial Intelligence Review, 2021, 54(8): 6391-6438.

[13] Wan L, Chen T, Plaza A, et al. Hyperspectral unmixing based on spectral and sparse deep convolutional neural networks [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 11669-11682.

[14] Lei M, Li J, Qi L, et al. Hyperspectral unmixing *via* recurrent neural network with chain classifier [C]// 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020: 2173-2176.

[15] Jin Q, Ma Y, Fan F, et al. Adversarial autoencoder network for hyperspectral unmixing [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(8): 4555-4569.

[16] Deng K, Qian Y, Nie J, et al. Diffusion-model-based hyperspectral unmixing using spectral prior distribution [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5519716.

[17] Chen J, Yang C, Zhang L, et al. TCCU-Net: transformer and CNN collaborative unmixing network for hyperspectral image [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 8073-8089.

[18] Gao W, Yang J, Zhang Y, et al. SSAF-net: a spatial-spectral adaptive fusion network for hyperspectral unmixing with endmember variability [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5506015.

[19] Dosovitskiy A. An image is worth 16×16 words: Transformers for image recognition at scale [J]. arXiv preprint arXiv: 2010.11929, 2020.

[20] Chen Z, Zhu Y, Zhao C, et al. DPT: deformable patch-based transformer for visual recognition [C]// Proceedings of the 29th ACM International Conference on Multimedia, 2021: 2899-2907.

[21] Zhu F, Wang Y, Fan B, et al. Spectral unmixing *via* data-guided sparsity [J]. IEEE Transactions on Image Processing, 2014, 23(12): 5412-5427.

[22] Schaepman M E, Jehle M, Hueni A, et al. Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX) [J]. Remote Sensing of Environment, 2015, 158: 207-219.

[23] Zhang S, Li J, Liu K, et al. Hyperspectral unmixing based on local collaborative sparse regression [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(5): 631-635.

[24] Karami A, Heylen R, Scheunders P. Denoising of hyperspectral images using shearlet transform and fully constrained least squares unmixing [C]//2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2016: 1-5.

[25] Feng X R, Li H C, Wang R, et al. Hyperspectral unmixing based on nonnegative matrix factorization: a comprehensive review [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 4414-4436.

[26] Gao L, Han Z, Hong D, et al. CyCU-Net: cycle-consistency unmixing network by learning cascaded autoencoders [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 5503914.

[27] Ghosh P, Roy S K, Koirala B, et al. Hyperspectral unmixing using transformer network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5535116.

作者简介：

陈丽媛(2000—)，女，硕士研究生，主要从事高光谱图像混合像元分解方面的研究；

杨华东(1986—)，男，博士，副教授，主要从事高光谱图像解译及智能信息处理方面的研究。