

多尺度分解双寻优策略 SPCNN 的果园苹果异源图像融合模型

刘立群¹, 顾任远¹, 周煜博¹, 火久元²

(1. 甘肃农业大学信息科学技术学院, 兰州 730070; 2. 兰州交通大学电子与信息工程学院, 兰州 730070)

摘要: 针对单一的自然场景图像信息不能满足准确识别果实和精准定位目标的要求, 提出一种多尺度分解双寻优策略简化脉冲耦合神经网络 (Simplified Pulse Coupled Neural Network, SPCNN) 的飞行时间 (Time of Flight, ToF) 与可见光果园苹果图像融合模型。对 SPCNN 模型引入带参数优化的双寻优策略, 对非下采样轮廓波变换 (Nonsubsampled Contourlet Transform, NSCT) 融合规则进行改进。模型包括配准模块、编码区、多尺度分解模块、单目标 SPCNN 融合模型、多目标 SPCNN 融合模型、解码区。模型改进了 SPCNN 模型的参数优化方式以及迭代次数, 模型自适应点火次数较低, 在 3~7 次左右, 具有点火次数低、自适应分割、效率高的优点。中光 15:00 时段点火识别成功率达到了 100.00%, 点火分割时间达到最低 91.91s。与其他融合模型比较, 模型在强光 12:00、中光 15:00、弱光 18:20、19:00 时段融合图像识别成功率达到 100.00%; 融合时间低于 SPCNN 模型, 达到最低 92.68 s。模型识别精度最优达到了 100.00%, 融合耗时最低达到了 92.68 s, 模型大小较 SPCNN 低一个数量级, 可补充和完善图像层次融合理论和方法。

关键词: 图像识别; 模型; 非下采样轮廓波变换; 简化脉冲耦合神经网络; 异源图像融合模型; 单目标策略; 多目标策略
doi: 10.11975/j.issn.1002-6819.2022.17.017

中图分类号: TP301

文献标志码: A

文章编号: 1002-6819(2022)-17-0158-10

刘立群, 顾任远, 周煜博, 等. 多尺度分解双寻优策略 SPCNN 的果园苹果异源图像融合模型[J]. 农业工程学报, 2022, 38(17): 158-167. doi: 10.11975/j.issn.1002-6819.2022.17.017 http://www.tcsae.org

Liu Liqun, Gu Renyuan, Zhou Yubo, et al. Heterologous image fusion with multi-scale decomposition and dual optimization SPCNN in an apple orchard[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(17): 158-167. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.17.017 http://www.tcsae.org

0 引言

苹果自动采摘中, 准确识别出苹果果实对苹果目标的精准定位尤为重要^[1]。单一的自然场景图像采集信息已经不能满足准确识别果实和精准定位目标的要求, 异源图像融合 (Image Fusion, IF) 技术应运而生^[2]。不同类型传感器采集的图像数据特性不尽相同, 其中深度传感器成像系统采用飞行时间 (Time of Flight, ToF) 技术, 通过 (850±30) nm 近红外光的飞行时间来完成距离计算, 已作为可见光相机的有益补充, 应用在农业、医疗、质检、视觉等领域, 具有广泛的应用前景和研究价值^[3]。杨勇等用卷积自编码融合网络针对红外与可见光图像进行融合^[4], 结果优于其他主流融合方法, 但仅适用于红外热度信息丰富的图像融合领域。在自然场景果园条件下, 果树、果实本身不具备红外热度感应特性, ToF 深度传感器成像系统具备发射近红外光并通过光线发射和反射时间差计算果实距离、产生深度 (即距离) 信息的条件^[5], 因此适用于构建本文异源视觉系统。学者们针对 ToF 深度传感器成像系统采集的深度信息进行了一系列应用研究^[6-7], 利用 ToF 深度传感器与可见光相机构建异源视觉系统是可行且必要的, 深度图像在融合研究应用中提供

了必要的补充信息。

像素级变换域融合模型是目前研究的热点, 其中非下采样轮廓波变换 (Nonsubsampled Contourlet Transform, NSCT) 是一种灵活的多尺度、多方向、平移不变的变换域融合模型分解方法, NSCT 变换由非下采样塔式滤波器组 (Nonsubsampled Pyramid Filter Bank, NSP) 和非下采样方向滤波器组 (Nonsubsampled Directional Filter Bank, NSDFB) 两部分组成^[8]。利用 NSCT 提取源图像的特征, 可以为融合图像提供更丰富的细节信息^[8]。部分学者尝试将小波变换应用于图像融合领域^[9], 但由于离散小波变换 (Discrete Wavelet Transform, DWT) 仅能分解出有限方向的细节信息, 融合结果仍然不够理想^[10]。部分学者将多尺度变换、卷积神经网络引入, 相继提出了轮廓波图像融合模型^[11-12]、深度卷积神经网络多聚焦图像融合模型^[13]、多序列深度特征融合模型^[14]、变分融合模型^[15]。杨风暴等^[16-17]提出 DWT、NSCT 和改进主成分分析 (Principal Components Analysis, PCA) 的多算法协同组合融合模型。与曲线波和剪切波相比, NSCT 能够充分融合源图像信息, 融合图像具有较好的相关系数和信息熵, 曲线波和剪切波在寻找边缘和轮廓方面具有很好的效果^[10,16-17]。但是自然场景下, 过多边缘信息并不能用于提取更有效的果实图像。现有融合模型具有各自应用领域的局限性, 无法自适应扩展到本研究自然场景果园图像中。因此, 曲线波和剪切波更适合医学等图像等背景简单、更需要提取轮廓的图像, NSCT 更适用于本项目自然场景采集的图像融合。

收稿日期: 2022-02-17 修订日期: 2022-07-20

基金项目: 甘肃农业大学青年导师基金资助项目 (GAU-QDFC-2020-08); 甘肃省科技计划项目 (20JR5RA032)

作者简介: 刘立群, 副教授, 研究方向为计算机技术在农业工程中的应用。
Email: llqhjy@126.com。

简化脉冲耦合神经网络 (Simplified Pulse Coupled Neural Network, SPCNN) [18] 模型是模拟大脑皮层视觉神经细胞活动建立的神经网络模型, 因具有聚集性点亮分割机制, 常被用于图像融合领域, 具有参数少、模型简单、参数受限于人工经验设置等特点, 研究中常被视为多目标参数优化问题。在图像融合的多目标优化领域中, 基于优化理论的自适应算法, 站在一个更客观的角度去审视图像融合过程, 做出最为科学严谨的判断, 大大提升了融合图像的可靠性 [19]。目前有学者将量子粒子群优化算法 [20]、灰狼优化算法 [21]、混合蛙跳算法 [22]、蚁群算法 [23] 等应用于 SPCNN 图像融合的多目标参数求解中, 建立了图像处理优化模型, 提出了有效的评价指标, 取得了较好的效果。

本文针对尚未形成公认的、系统的图像融合理论方法, 提出一种多尺度分解双寻优策略 SPCNN 的 ToF 与可见光果园苹果图像融合模型。模型在自然场景果园苹果图像融合方面, 有效提高 ToF 与可见光图像融合质量, 对补充和完善图像层次融合理论和方法具有重要意义。

1 多尺度分解的异源图像融合模型构建

本文在 SPCNN 基础上改进, 提出多尺度分解双寻优策略 SPCNN 的 ToF 与可见光果园苹果图像融合模型对异源视觉系统采集的图像进行融合。

其中, SPCNN 表达式如式 (1)~式 (5) 所示 [18]。但仍存在参数经验设定、不能自适应终止、易出现过分割的缺陷, 尚需进一步改进。

$$F_{ij}(n) = I_{ij} \quad (1)$$

$$L_{ij}(n) = \exp(-\alpha_L) L_{ij}(n-1) + V_L \sum_{kl} W_{ij,kl} Y_{kl}(n-1) \quad (2)$$

$$U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n)) \quad (3)$$

$$\theta_{ij}(n) = \exp(-\alpha_\theta) \theta_{ij}(n-1) + V_\theta Y_{ij}(n) \quad (4)$$

$$Y_{ij}(n) = \text{step}(U_{ij}(n) - \theta_{ij}(n)) \quad (5)$$

式中 I_{ij} 是神经元外部刺激, 由输入图像的灰度值表示; $F_{ij}(n)$ 是反馈输入域, $L_{ij}(n)$ 是链接输入域; $W_{ij,kl}$ 表示链接系数; β 表示链接强度, 决定耦合链接通道权重大小; $U_{ij}(n)$ 为模型内部状态信号; n 表示当前点火次数; θ_{ij} 为神经元的动态阈值, V_θ 、 V_L 为动态阈值放大系数, 控制神经元激活后增加的门槛阈值; α_L 和 α_θ 分别决定了链接通道反馈项和动态阈值的衰减速度; $Y_{ij}(n)$ 为当前神经元的脉冲输出, 是内部活动项与动态阈值在脉冲发生器中进行比较后输出的响应结果。当满足 $U_{ij}(n) > \theta_{ij}(n)$ 时, 就会达成点火条件输出 $Y_{ij}(n)=1$ 。

本文模型具体包括单目标 SPCNN 融合模型 (Single Target SPCNN fusion model, ST-SPCNN) 和多目标 SPCNN 融合模型 (Multi-Target SPCNN fusion model, MT-SPCNN) 两个模块。利用 ToF 深度相机与可见光相机构建异源视觉系统, 采用变换域 NSCT 对异源图像进行多尺度多方向分解, 将 SPCNN 模型参数优化问题看作多目标优化问题, 采用改进人工蜂群 (Artificial Bee

Colony, ABC) [24] 算法对 ST-SPCNN 模型和 MT-SPCNN 模型进行双寻优策略参数优化, 包括单目标优化参数策略和多目标参数优化策略, 最后利用优化后的 SPCNN 模型融合异源图像。

1.1 模型结构

多尺度分解的异源图像融合模型结构如图 1 所示, 包括配准模块、编码区、多尺度分解模块、ST-SPCNN、MT-SPCNN、解码区。其中, 输入原图像分别为异源视觉系统 ToF 强度图像、置信图像和对应的可见光图像, 配准模块用于两幅图像的像素级精准配准, 编码区用于对 SPCNN 模型的参数进行编码, 待优化使用; 多尺度分解模块由 NSCT 分解模块组成, 配准后图像分别经过 NSP 多尺度变换以及 NSDFB 多方向分解模块, 将分解出的高频和低频图像, 分别通过 ST-SPCNN 模型、MT-SPCNN 模型不同的融合规则进行融合, 将编码区编码的四个参数向量分别作为单目标人工蜂群算法以及多目标优化算法的初始种群, 采用改进的人工蜂群策略对参数进行自适应的寻优。融合后, 经过 NSCT 逆变换, 融合成一幅图像, 最后通过解码区对每一个二进制向量解码为实数参数。

1.1.1 配准模块

1) 异源视觉系统的标定

本文采用相机标定法 [25], 7×8 棋盘的标定板对异源视觉系统进行标定, 棋盘上的网格为 $22 \text{ mm} \times 22 \text{ mm}$, 使用 MATLAB 摄像机标定工具箱对双目摄像机进行标定。标定时可见光图像是参考图像, ToF 图像是浮动图像, 用彩色相机去标定 ToF 相机。计算得出图像坐标与世界坐标系的对应关系, 其中 ToF 相机相对于彩色相机参考坐标系的 x 、 y 、 z 方向的偏移量是一维向量 $[108.678 \ 15.695 \ 3.800]$, 单位 mm。通过以上操作对异源视觉系统完成相机标定操作。

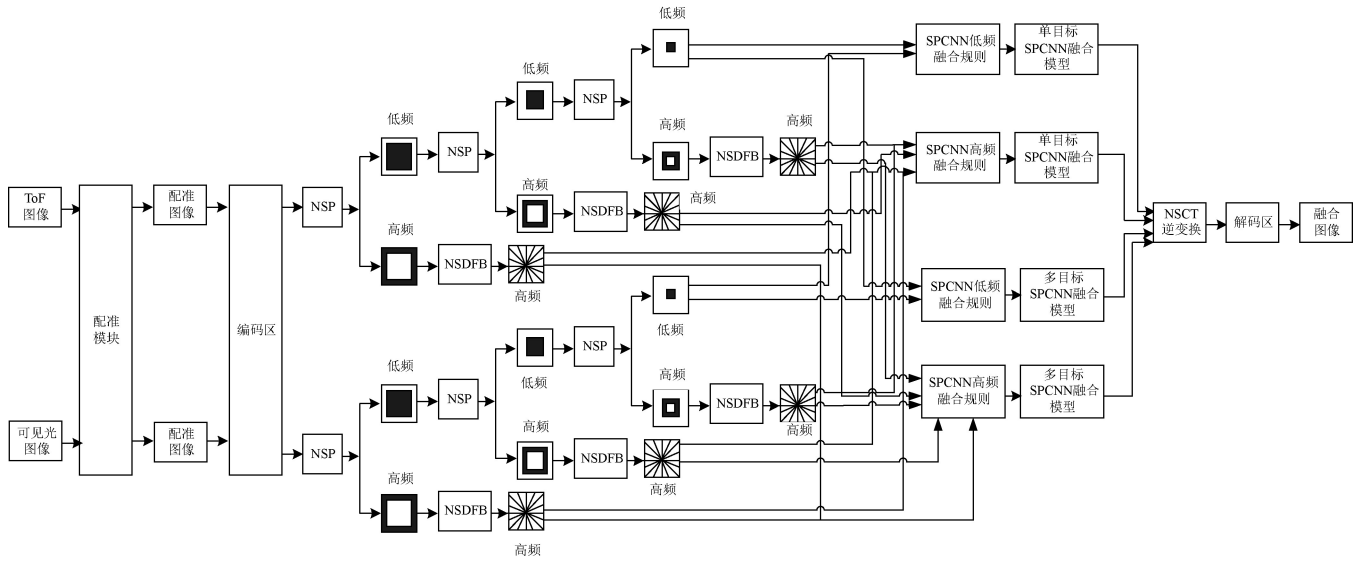
2) 异源图像的配准

本文使用外极线约束用于双目图像的立体校正, 以形成共面配准。然后将尺度不变特征变换 (Scale-Invariant Feature Transform, SIFT) 算法 [26] 应用于校正后的双目图像, 实现立体匹配。最后, 使用随机抽样一致 (Random Sample Consensus, RANSAC) [27] 算法消除失配, 提高精度。

1.1.2 编码区

SPCNN 模型具有参数不能根据图像信息自适应设定、严重依赖人工参数设置的缺陷。编码区针对 SPCNN 模型输入链接通道反馈项 α_L 、链接强度 β 、动态阈值衰减因子 α_θ 以及动态阈值放大系数 V_θ 四个参数进行编码, 每一个参数对应于编码区的一个维度的向量, 构成以下形式矩阵, 如 (6) 式所示。其中维度为 4, 每个参数对应向量采用八位二进制编码。

$$\begin{bmatrix} \alpha_{L1} & \cdots & \alpha_{L8} \\ \vdots & \ddots & \vdots \\ V_{\theta 1} & \cdots & V_{\theta 8} \end{bmatrix} \quad (6)$$



注：NSP 代表 NSCT 变换中非下采样塔式滤波器组；NSDFB 代表 NSCT 变换中非下采样方向滤波器组；NSCT 为非下采样轮廓波变换。

Note: NSP represents nonsubsampling pyramid filter bank in NSCT transform; NSDFB represents nonsubsampling directional filter bank in NSCT transform; NSCT represents nonsubsampling contourlet transform.

图 1 多尺度分解的异源图像融合模型

Fig.1 Multi-scale decomposition of heterogeneous image fusion model

1.1.3 多尺度分解模块

经过配准的两幅图像，经过多尺度 NSCT 分解模块。对图像进行 NSP 变换，通过 nsc 级变换后，会生成一个低频子带图像和 nsc 个高频子带图像，共 $nsc+1$ 个和原图像大小相等的子带图像。对分解出的高频子带进行 q 级 NSDFB 方向分解，可以得到 2^q 个方向子带图像。即对原图像进行 NSCT 变换，其中 NSP 变换是 nsc 级，NSDFB 变换是 q 级，就可以得到 1 个低频子带图像和 nsc 个高频子带图像，每个高频子带又包含 2^q 个方向的子带图像，

即共 $\sum_{q=1}^{nsc} 2^q$ 个不同方向高频子带图像。

本文采用图 1 所示的 NSCT 多尺度分解，其中 nsc 级 q 个方向的参数选择， nsc 和 q 的值均不宜过大也不宜过小， nsc 过大，则导致分解出的高、低频子带数增多，影响后期融合效率， nsc 过小，导致分解层数过少，影响融合效果； q 过大，导致高频细节信息被划分太细腻，丢失重要的纹理信息， q 过小，细节信息未被细分，不利于提供融合质量。因此，本文对两幅图像分别进行 3 级 8 方向分解。图 1 中， $nsc=3$ ， $q=3$ ，对原图像进行 3 级 NSP 多尺度分解，对高频子带进行 3 级 NSDFB 方向分解，最终得到了 1 个低频子带图像和 3 个高频子带图像，每个高频子带又包含 8 个方向的子带图像，共 24 个不同方向高频子带图像。

1.1.4 单目标 SPCNN 融合模型

经过多尺度分解后的图像，变为高频和低频，采用单目标 SPCNN 融合规则分别对高频和低频图像进行融合。由于 SPCNN 模型参数由人工经验手动设定，无法与分割图像自适应匹配，存在将灰度一致性区域点亮误分割的现象。针对 ToF 图像果实目标在局部通常表现为整体灰度显著，结合 ToF 图像自身特性对 SPCNN 模型参数

进行寻优，会得到自适应参数，提高融合质量。这里采用改进的单目标人工蜂群算法优化 SPCNN 参数，将编码区编码的四个参数向量作为单目标人工蜂群算法的初始种群，采用改进的人工蜂群策略对参数进行自适应的寻优。

人工蜂群算法是模拟群体蜜蜂特性提出的一种群体智能优化算法，是一种典型的全局寻优算法。蚁群算法具有收敛速度慢和易陷入局部最优的缺点。模拟退火算法存在收敛速度慢、性能与初始参数敏感等缺点。遗传算法的局部搜索能力较差，导致单纯的遗传算法比较费时，在进化后期搜索效率较低。禁忌搜索算法具有局部开发能力强、全局开发能力弱的缺点。粒子群算法收敛速度快但容易陷入局部最优解；而人工蜂群算法具有全局寻优能力强、参数少、精度高、鲁棒性强的优点，适于本研究中的参数优化策略。ABC 算法中，雇佣蜂、观察蜂更新食物源信息策略具有单一性及随机性的缺陷，使得算法存在早熟、收敛停滞等问题，为了加速人工蜂群算法收敛速度，将 ABC 算法寻优策略进行改进。

1) 雇佣蜂更新策略

随机生成 NP 个食物源信息，一次食物更新进化中，在蜂群中随机选择一个雇佣蜂所依附的食物源 $X_k=(x_{k1},x_{k2},\dots,x_{kd})$ ，找出此次更新中的全局最优食物源记为 $X_g=(x_{g1},x_{g2},\dots,x_{gd})$ ，在 d 维空间中，对食物源信息空间中每个食物源 $X_k=(x_{k1},x_{k2},\dots,x_{kd})$ 随机选择第 j 维分量 x_{ij} 通过以下雇佣蜂更新策略进行进化，如式 (7) 所示，得到一个新的食物源 $X_i^{new}=(x_{i1},x_{i2},\dots,x_{ij}^{new},\dots,x_{id})$ 。其中， $i,k\in[1,2,\dots,t],i\neq k,j\in[1,2,\dots,d],r\in[-1,1]$ 。

$$x_{ij}^{new} = x_{gd} + r(x_{ij} - x_{kj}) \quad (7)$$

2) 观察蜂更新策略

一次食物更新进化中，找出此次更新中的全局最优食物源作为观察蜂所依附的食物源，记为

$X_g=(x_{g1},x_{g2},\cdots,x_{gd})$, 对雇佣蜂更新策略进化得到的食物源 $X_i^{new}=(x_{i1},x_{i2},\cdots,x_{ij}^{new},\cdots,x_{id})$ 在 d 维空间中随机选择第 j 维分量 x_{ij}^{new} , 通过以下观察蜂更新策略进行进化, 如式 (8) 所示, 得到另一个新的食物源 $X_i^{new2}=(x_{i1},x_{i2},\cdots,x_{ij}^{new2},\cdots,x_{id})$ 。

$$x_{ij}^{new2}=x_{ij}^{new}+r(x_{gd}-x_{ij}) \quad (8)$$

3) 单目标函数计算方法

图像融合质量采用相关系数 (Correlation Coefficient, CC)、交叉熵 (Cross Entropy, CE)、熵 (Entropy, E)、互信息 (Mutual Information, MI) 共 4 个评价指标进行评价。由于图像融合后的质量需要 4 个评价指标共同衡量, 四个评价指标在图像融合质量评价中起的作用相当, 因此, 对于单目标寻优问题, 采取权重相加、权重相等的方式计算单目标函数, 公式如 (9) 所示。其中, 各指标对应的权重 $\omega_i=0.25, i=1,\cdots,4$ 。

$$\text{fitness_single}=\omega_1\text{CC}+\omega_2\text{CE}+\omega_3\text{E}+\omega_4\text{MI} \quad (9)$$

4) SPCNN 融合规则

经过分解的高频和低频图像, 需要通过 SPCNN 融合规则进行融合。其中, SPCNN 的参数链接域的链接系数 $W_{ij,k}$ 采用 $n \times n, n=5$ 的矩阵, 其中 n 代表 $W_{ij,k}$ 与周围相邻神经元 (k,l) 的神经元个数, 这里 $n=5$ 代表当前神经元与周围 5×5 神经元相邻进行链接, 如式 (10) 所示。

$$W_{ij}=\begin{cases} 0, i=\frac{n}{2}, j=\frac{n}{2} \\ \frac{1}{\sqrt{\left(i-\frac{n}{2}\right)^2+\left(j-\frac{n}{2}\right)^2}}, \text{else} \end{cases} \quad (10)$$

低频图像通过选取 3×3 区域, 求解对应区域的方差 σ , 融合规则如式 (11) 所示。其中, $C_{j0}^L(\cdot)$ 代表融合后的低频分量, $C_{j0}^{\text{ToF}}(\cdot)$ 代表 ToF 图像的低频分量, $C_{j0}^{\text{RGB}}(\cdot)$ 代表可见光 (Red Green and Blue, RGB) 图像的低频分量。

$$C_{j0}^L(\cdot)=\begin{cases} C_{j0}^{\text{ToF}}(\cdot), \sigma(C_{j0}^{\text{ToF}}(\cdot))>\sigma(C_{j0}^{\text{RGB}}(\cdot)) \\ C_{j0}^{\text{RGB}}(\cdot), \text{else} \end{cases} \quad (11)$$

高频图像通过选取 3×3 区域, 求解对应区域的区域能量 (Regional Energy, RE), 融合规则如式 (12) 所示。 $C_{jk}^L(\cdot)$ 代表融合后的高频分量, $C_{jk}^{\text{ToF}}(\cdot)$ 代表 ToF 图像的高频分量, $C_{jk}^{\text{RGB}}(\cdot)$ 代表 RGB 图像的高频分量。

$$C_{jk}^L(\cdot)=\begin{cases} C_{jk}^{\text{ToF}}(\cdot), \text{RE}(C_{jk}^{\text{ToF}}(\cdot))>\text{RE}(C_{jk}^{\text{RGB}}(\cdot)) \\ C_{jk}^{\text{RGB}}(\cdot), \text{else} \end{cases} \quad (12)$$

1.1.5 多目标 SPCNN 融合模型

从控制理论的角度讲, 异源图像融合本质上可以看作是将多幅图像组合为一幅图像的信息寻优过程。由于图像融合效果的评价指标具有多样性, 各种指标之间关系复杂甚至相互矛盾, 因此这种寻优过程可以采用多目标优化问题 (Multi-Objective Optimization Problems, MOOP) [29] 的理论来解决。

在 ST-SPCNN 模型中, 将 4 个图像融合质量评价函数进行了加权求和的方式, 忽略了 4 个评价指标的多样

性。可以把这种寻优过程理解为多目标优化问题, 含有多个目标函数, 求解时存在多个相互冲突的目标, 某个目标的改善可能引起其他目标性能的降低, 需要在各个目标之间权衡协调, 使所有的目标尽可能达到最优 [28], 最终得到一个帕雷托最优解集, 又称 Pareto 最优解集。

1) 多目标密集度构造方法

在多目标优化问题中, 个体的优劣是通过支配关系和密集信息来判断的。为保证 Pareto 最优解之间的分布不至于密集, 本文采用一种新的网格构造方法对 Pareto 最优解集密集度进行衡量, 如式 (13) ~ (14) 所示。式中, low_i 和 uper_i 分别表示网格的最小和最大边界值, α 表示网格扩张系数, $\alpha=0.1$ 表示网络的边界值可以在一定范围内动态调整, Target 表示目标函数的数量, Target=4 代表多目标优化问题有四个目标函数 CC, CE, E, MI。因此, 对于最优解集中的每一个维度, 在非支配解集中找到它的最大以及最小值。适当放大最大值与最小值之间的区间, 以防止非支配解落在网格的边界上。

$$\text{low}_i=\text{low}_i-\alpha(\text{uper}_i-\text{low}_i) \quad (13)$$

$$\text{uper}_i=\text{uper}_i+\alpha(\text{uper}_i-\text{low}_i) \quad (14)$$

$$i=1, \cdots, \text{Target}, \text{Target}=4, \alpha=0.1$$

2) 多目标删除选择概率构造方法

构造非支配解集时, 需要一定的概率随机删除多余的非支配解, 以便满足最优解集密集度要求, 本文采用一种新的删除选择概率方法, 使用非支配解同维度网格索引 grid_w 之间距离的 L^1 误差进行度量, 公式如 (15) 所示。

$$L_i^1=\frac{1}{k} \sum_{l=1}^{\text{Rep}} \sum_{j=1}^{\text{Rep}} \sum_{w=1}^{\text{Target}} |\text{grid}_{iw}-\text{grid}_{jw}| \quad (15)$$

3) 多目标函数计算方法

为解决图像融合质量评价函数多样性的问题, 采用单目标中雇佣蜂、观察蜂更新策略, 使用新的网格构造方法解集 Pareto 最优解集密集度问题, 采用网格索引距离的 L^1 误差作为非支配解选择删除概率, 同时选择上述四个图像融合质量评价函数, 形成四个目标的多目标优化问题, 公式如 (16) 所示。

$$\text{fitness_pareto}=\max\{\text{CC}, \text{CE}, \text{E}, \text{MI}\} \quad (16)$$

1.1.6 解码区

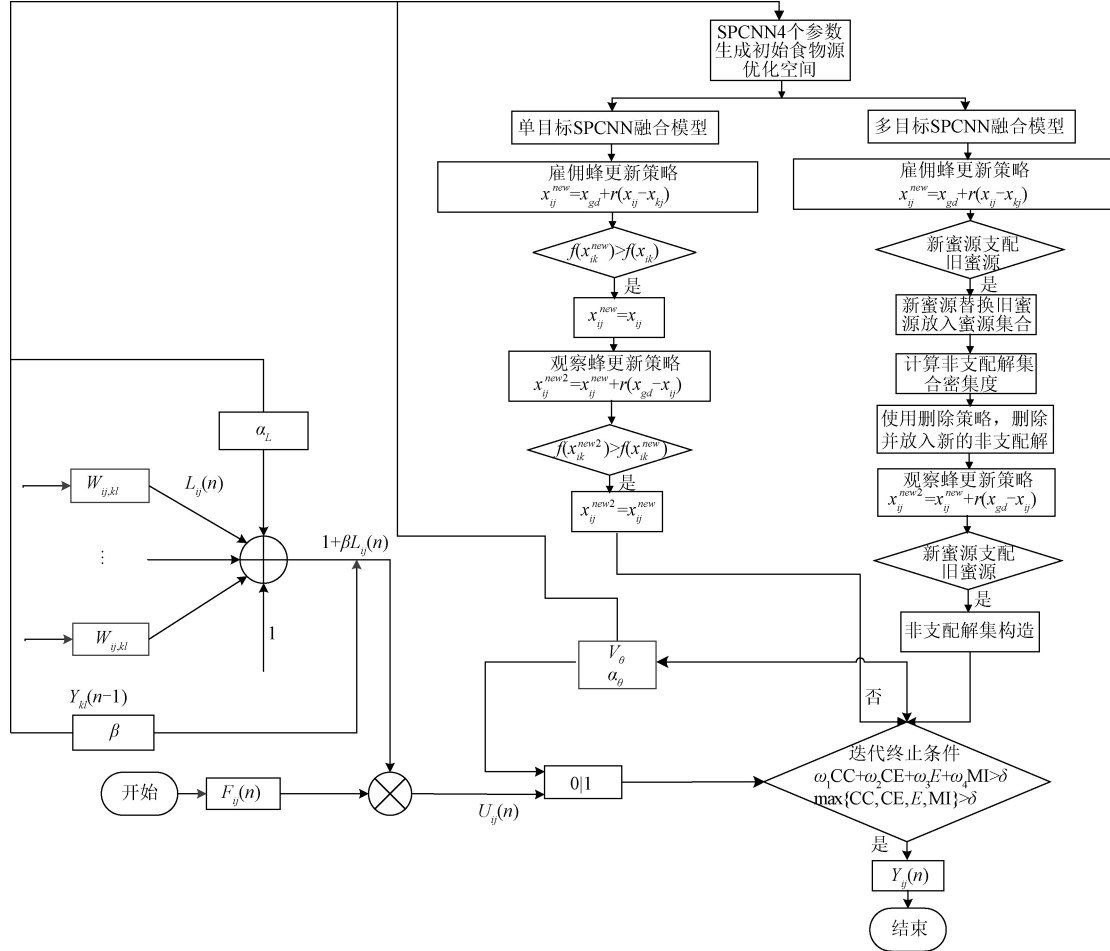
利用多尺度 SPCNN 融合规则对两幅图像融合后, 经过 NSCT 逆变换, 融合成一幅图像, 经过目标函数的对比, 迭代更新后得到融合图像, 此时, 对改进 SPCNN 模型输入链接通道反馈项 α_L 、链接强度 β 、动态阈值衰减因子 α_θ 以及动态阈值放大系数 V_θ 四个参数进行二进制解码, 将每一个二进制向量对应转换成实数参数。

1.2 模型运行流程

模型运行流程是: 将编码区编码的四个参数向量作为单目标人工蜂群算法的初始种群, 采用改进的人工蜂群算法单目标优化参数策略和多目标参数优化策略对参数进行自适应的寻优。SPCNN 模型因迭代终止条件通常人工手动设置为较大次数, 具有不能自适应终止的缺陷。将双寻优策略目标函数作为 SPCNN 模型的迭代终止条件, 加快了 SPCNN 模型点火分割的速度。SPCNN 改进模型融合

流程见图 2 所示。具体流程是：配准后图像经过图 1 的 NSP、NSDFB 多尺度分解模块，分解出高频和低频图像作为 ST-SPCNN 模型、MT-SPCNN 模型的反馈输入域 $F_{ij}(n)$ 的输入，图像中每一个像素点与 SPCNN 中每一个神经元对应。神经元通过链接系数 $W_{ij,kl}$ 与周围相邻神经元 (k,l) 的上一次脉冲输出 $Y_{kl}(n-1)$ 耦合，将编码区中链接通道反馈

项 α_L 、链接强度 β 、动态阈值衰减因子 α_θ 以及动态阈值放大系数 V_θ 四个参数矩阵作为人工蜂群算法的初始种群，分别采用改进的人工蜂群算法单目标优化参数策略和多目标参数优化策略对参数进行自适应的寻优，将双寻优策略目标函数作为 SPCNN 模型的迭代终止条件，达成点火条件时输出 $Y_{ij}(n)=1$ ，输出点火分割结果，完成一次点火过程。



注： $F_{ij}(n)$ 代表 ToF 或可见光图像经 NSCT 分解后的高频或低频分量灰度值； $L_{ij}(n)$ 代表链接输入域； α_L 代表链接通道反馈项； β 代表链接强度； α_θ 代表动态阈值衰减因子； V_θ 代表动态阈值放大系数； $Y_{kl}(n-1)$ 代表上一次点 (k,l) 处的脉冲输出； $W_{ij,kl}$ 表示与周围相邻神经元的链接系数； $(1+\beta L_{ij})$ 表示通过链接强度 β 将输入域 $L_{ij}(n)$ 结合； $U_{ij}(n)$ 为模型内部状态信号； $Y_{ij}(n)$ 为当前神经元的脉冲输出； δ 代表 SPCNN 模型的迭代终止条件。

Note: $F_{ij}(n)$ represents the gray value of high-frequency or low-frequency components of ToF or visible image after NSCT decomposition; $L_{ij}(n)$ represents the link input field; α_L represents link channel feedback item; β represents link strength; α_θ represents dynamic threshold attenuation factor; V_θ represents dynamic threshold amplification factor; $Y_{kl}(n-1)$ represents the pulse output at the last point (k,l) ; $W_{ij,kl}$ represent the link coefficient with surrounding adjacent neurons; $(1+\beta L_{ij})$ indicates the strength through the link β Combine the input domain $L_{ij}(n)$; $U_{ij}(n)$ is the internal state signal of the model; $Y_{ij}(n)$ is the pulse output of the current neuron; δ represents the iteration termination condition of SPCNN model.

图 2 多尺度分解的异源图像融合模型运行流程

Fig.2 Multi-scale decomposition of heterogeneous image fusion model operation process

1.3 模型评价指标

1.3.1 双寻优策略 SPCNN 点火分割评价指标

试验采用原始 SPCNN 模型^[18]，本文 ST-SPCNN 模型、MT-SPCNN 模型测试模型的点火分割性能。分别采用点火次数、分割效果、以及参数寻优结果进行评价。定义识别成功率 R_{Success} 和漏识率 R_{Miss} 两个指标对分割效果进行评价，两个指标定义表达式为 (17) ~ (18) 式。其中 N_m 为模型分割识别出的总目标数， N_f 为人工目视识别的果实数量。为保证分割完整性，视分割出的一个完整无空洞的为一个目标果实，人工目视中忽略遮挡 2/3 果

实、半个果实。

$$R_{\text{Success}} = \frac{N_m}{N_f} \times 100\% \quad (17)$$

$$R_{\text{Miss}} = 100\% - R_{\text{Success}} \quad (18)$$

1.3.2 图像融合评价指标

试验选择六种模型，包括 NSCT 模型^[8]、基于 NSCT 的红外和可见光图像融合模型 (Fusion method for infrared and visible light images based on NSCT, ImNSCT)^[29]、DWT 模型^[10]、原始 SPCNN 模型^[18]和本文 ST-SPCNN 模型、MT-SPCNN 模型对异源视觉系统图像融合性能进行

对比测试。图像融合质量采用 (Average Gradient, AG)、相关系数 CC、交叉熵 CE、熵 E 、互信息 MI、空间频率 (Spatial Frequency, SF) 等评价指标^[30]进行评价。其中 CE 值越小代表融合图像质量越好,其他评价指标值越大代表融合图像质量越好。融合图像的识别成功率 R_{Success} 和漏识率 R_{Miss} 两个评价指标沿用式 (17) ~ (18) 表达式。模型复杂度用时间复杂度和空间复杂度两个指标进行评价,分别衡量模型运行时间效率以及模型所占内存大小。

1.4 模型数据获取

本文建立异源视觉系统,采用德国 Basler 公司生产的 ToF 工业深度相机与佳能公司生产的彩色相机。ToF 相机能输出四种类型的图像,包括 ToF 强度图 (ToF Intensity Image, ToF_intensity), ToF 深度图 (ToF Range Data, ToF_range), ToF 置信图 (ToF Confidence Map, ToF_confidence) 以及 ToF 点云图 (ToF Point Cloud Image, ToF_points)^[31]。自然场景中的采集地点位于甘肃省天水

市秦州区果树研究所试验基地。苹果样品的品种是天旺一号。利用异源视觉系统采集 10:00—19:00 不同光照条件下 ToF 强度图像、深度图像、置信图像和彩色图像各 1 000 余幅。

从果园自然场景中采集的异源图像中选择 6 组数据作为样本,包括强光 12:00、13:00、中光 15:00、16:00 以及弱光 18:20、19:00 六个不同时间段的 ToF 强度图像、置信图像和对应的可见光图像共 18 幅进行试验。分别设计点火分割试验和图像融合试验对提出的融合模型进行性能测试。

2 结果与分析

2.1 双寻优策略 SPCNN 点火分割试验结果分析

2.1.1 改进 SPCNN 点火分割结果分析

以 ToF 置信图像与可见光图像为例,在强光 13:00、中光 15:00 和弱光 18:20 条件下的低频分量点火分割效果见表 1 所示。

表 1 改进 SPCNN 点火分割效果

Table 1 Ignition segmentation effect of improved Simplified Pulse Coupled Neural Network (SPCNN)

| 光照条件 Illumination condition | 模型 Model | 点火次数 Ignition times | 分割目标数 Target number N_m | 目视果实数 Visual fruit number N_f | 识别成功率 Recognition success rate $R_{\text{Success}}/\%$ | 漏识率 Missed recognition rate $R_{\text{Miss}}/\%$ | 分割时间 Segmentation runtime/s |
|-----------------------------------|-------------|------------------------|---------------------------------|---------------------------------------|--|--|--------------------------------|
| 强光 13:00 Strong sunlight 13:00 | SPCNN | 50 | 3 | 5 | 60.00 | 40.00 | 122.01 |
| | ST-SPCNN | 5 | 4 | 5 | 80.00 | 20.00 | 91.91 |
| | MT-SPCNN | 3 | 4 | 5 | 80.00 | 20.00 | 92.01 |
| 中光 15:00 Medium sunlight 15:00 | SPCNN | 50 | 4 | 7 | 57.14 | 42.86 | 121.28 |
| | ST-SPCNN | 6 | 7 | 7 | 100.00 | 0 | 93.32 |
| | MT-SPCNN | 5 | 6 | 7 | 85.71 | 14.29 | 96.01 |
| 弱光 18:20 Weak sunlight 18:20 | SPCNN | 50 | 3 | 5 | 60.00 | 40.00 | 115.18 |
| | ST-SPCNN | 7 | 4 | 5 | 80.00 | 20.00 | 94.12 |
| | MT-SPCNN | 6 | 4 | 5 | 80.00 | 20.00 | 93.64 |

结果显示,SPCNN 模型不能自主停止迭代,完全依靠人工经验设置。依据经验值,迭代次数一般设置在 50~500 次之间,具有点火次数长、点火次数固定、不具有自适应性的缺陷。表 1 表明,随着点火次数增加,SPCNN 模型将属于同灰度的果实区域都分割出来,点火次数增加到 50 次时,果实内部分割出现空洞,出现过分割,不利于提高融合质量。本文提出的 ST-SPCNN 及 MT-SPCNN 模型具有自适应点火机制,在点火分割迭代次数满足双寻优策略目标函数条件时,停止迭代,加快了 SPCNN 模型点火分割的速度。同时,表 1 结果表明,SPCNN 模型点火次数设置为固定次数 50 次,在 13:00 强光、15:00 中光以及 18:20 弱光 3 个时段的分割结果中均不同程度出现过分割现象,这表明固定经验的点火次数对分割结果不利。ST-SPCNN 模型在 3 个时段的自适应点火次数分别为 5、6、7 次,MT-SPCNN 模型自适应点火次数分别为 3、5、6 次,二者自适应点火次数均较低,在 3~7 次左右,远低于 SPCNN 模型固定迭代次数。而且,点火图具有果实分割清晰、无空洞、无过分割的优点。本文模型识别成功率皆高于 SPCNN 模型结果,在 13:00 强光、15:00 中光以及 18:20

弱光 3 个时段 ST-SPCNN 模型和 MT-SPCNN 模型对置信图的识别成功率均达到 80.00% 以上,且 15:00 中光时段达到了 100.00%。分割时间受点火次数影响,提出的模型分割时间皆低于 SPCNN 模型,最低达到 91.91 s。

2.1.2 改进 SPCNN 模型参数分割结果分析

SPCNN 模型 4 个参数人工手动设定值分别为 $\alpha_L = 0.85, \beta = 0.3, \alpha_\theta = 0.8, V_\theta = 20$ 。取多目标策略的非支配解集中的 4 个解作为解集合。其中强光 13:00 条件下的寻优结果见表 2 所示。

SPCNN 模型 4 个参数是通过人工手动设定,具有参数固定、经验取值的缺陷。表 2 结果显示,ST-SPCNN 模型参数寻优简单,运行时间为 91.91 s,低于 MT-SPCNN 模型运行时间 92.01 s。MT-SPCNN 模型参数寻优后具有多个非支配解,表 2 是选取 Pareto 最优解集中 4 个非支配解作为结果,参数彼此之间均符合寻优要求。SPCNN 模型运行时间完全依赖于固定迭代次数的人工设置。双寻优策略寻优依赖于 NSCT 分解、参数寻优时间以及双寻优策略 SPCNN 模型融合时间,但是由于自适应迭代次数,使得改进后 SPCNN 模型点火次数大大减少,提高了效率,在点火几次的情况下就可以达到快速自适应分割,提高了融合效率。

表 2 强光下改进 SPCNN 参数寻优结果

Table 2 Parameter optimization results of improved SPCNN under condition of sunlight with strong

| ST-SPCNN 模型强度+可见光图像 Intensity + visible image of ST-SPCNN | | | | | ST-SPCNN 模型置信+可见光图像 Confidence + visible image of ST-SPCNN | | | | | MT-SPCNN 模型强度+可见光图像 Intensity + visible image of MT-SPCNN | | | | | MT-SPCNN 模型置信+可见光图像 Confidence + visible image of MT-SPCNN | | | | |
|--|---------|-----------------|------------|--------|---|---------|-----------------|------------|--------|--|---------|-----------------|------------|--------|---|---------|-----------------|------------|--------|
| α_L | β | α_θ | V_θ | 分割时间/s | α_L | β | α_θ | V_θ | 分割时间/s | α_L | β | α_θ | V_θ | 分割时间/s | α_L | β | α_θ | V_θ | 分割时间/s |
| 0.09 | 0.57 | 0.68 | 54.65 | 91.11 | 0.39 | 0.80 | 0.75 | 37.73 | 91.91 | 0.87 | 0.51 | 0.94 | 63.77 | | 0.54 | 0.72 | 0.52 | 99.37 | |
| | | | | | | | | | | 0.67 | 0.69 | 0.06 | 25.47 | 91.19 | 0.40 | 0.44 | 0.36 | 76.35 | 92.01 |
| | | | | | | | | | | 0.78 | 0.67 | 0.01 | 60.21 | | 0.19 | 0.13 | 0.69 | 9.38 | |
| | | | | | | | | | | 0.42 | 0.46 | 0.77 | 32.24 | | 0.39 | 0.67 | 0.74 | 52.00 | |

注: α_L 代表链接通道反馈项; β 代表链接强度; α_θ 代表动态阈值衰减因子; V_θ 代表动态阈值放大系数。MT-SPCNN 模型以 4 个非支配解为例。
Note: α_L represents link channel feedback item; β represents link strength; α_θ represents dynamic threshold attenuation factor; V_θ represents dynamic threshold amplification factor. Taking four non dominated solutions as examples for MT-SPCNN model.

2.2 图像融合试验结果分析

各时段融合效果比较见表 3 所示, 各时段识别成功率均达到 80.00%以上。其中强光 12:00、中光 15:00 时段强度图、置信图与可见光融合图像识别成功率达到 100%, 弱光 18:20 时段强度图、19:00 时段置信图的融合图像识别成功率达到 100.00%。本文提出的融合模型融合时间低于 SPCNN 模型, 强光 12:00 时段强度图、置信图融合时间分别是 92.68 和 93.69 s, 达到最低。表 4 六种

融合模型的质量评价指标显示, 对于弱光 18:20 时段, 除相关系数 CC (RGB) 低于 SPCNN 模型 0.95、互信息 MI (RGB) 低于 SPCNN 模型 5.73 以外, ST-SPCNN 模型平均梯度 AG 值为 12.03、相关系数 CC (ToF) 值为 0.85、互信息 MI (ToF) 值为 1.00, MT-SPCNN 模型熵 E 值为 7.30、空间频率 SF 值为 23.80, 融合指标值均为最大, MT-SPCNN 模型交叉熵 CE 值为 0.38, 值最小。以上数据说明本文提出的模型图像融合质量高于其他模型。

表 3 6 种融合模型融合效果比较

Table 3 Comparison of fusion effects of six fusion models

| 光照条件 Illumination condition | 时间 Time | 模型 Model | 融合图像总目标数 Total number of fused images N_m | | 目视果实数 Visual fruit number N_f | | 识别成功率 Recognition success rate $R_{Success}/\%$ | | 漏识率 Missed recognition rate $R_{Miss}/\%$ | | 融合时间 Fusion runtime/s | |
|-----------------------------------|------------|-------------|---|------------|---------------------------------------|------------|---|------------|---|------------|--------------------------|------------|
| | | | 强度+可见光 | 置信+ 可见光 | 强度+ 可见光 | 置信+ 可见光 | 强度+ 可见光 | 置信+ 可见光 | 强度+ 可见光 | 置信+ 可见光 | 强度+ 可见光 | 置信+ 可见光 |
| | | | | | | | | | | | | |
| 强光 Sunlight with strong | 12:00 | NSCT | 2 | 2 | 3 | 3 | 66.67 | 66.67 | 33.33 | 33.33 | 53.86 | 52.79 |
| | | ImNSCT | 2 | 2 | 3 | 3 | 66.67 | 66.67 | 33.33 | 33.33 | 54.30 | 53.81 |
| | | DWT | 2 | 2 | 3 | 3 | 66.67 | 66.67 | 33.33 | 33.33 | 1.21 | 1.01 |
| | | SPCNN | 1 | 1 | 3 | 3 | 33.33 | 33.33 | 66.67 | 66.67 | 121.61 | 116.20 |
| | | ST-SPCNN | 3 | 3 | 3 | 3 | 100.00 | 100.00 | 0.00 | 0.00 | 93.78 | 93.83 |
| | | MT-SPCNN | 3 | 3 | 3 | 3 | 100.00 | 100.00 | 0.00 | 0.00 | 92.68 | 93.69 |
| | 13:00 | NSCT | 3 | 3 | 5 | 5 | 60.00 | 60.00 | 40.00 | 40.00 | 52.86 | 52.73 |
| | | ImNSCT | 3 | 3 | 5 | 5 | 60.00 | 60.00 | 40.00 | 40.00 | 53.85 | 53.96 |
| | | DWT | 3 | 3 | 5 | 5 | 60.00 | 60.00 | 40.00 | 40.00 | 1.03 | 1.03 |
| | | SPCNN | 2 | 3 | 5 | 5 | 40.00 | 60.00 | 60.00 | 40.00 | 115.77 | 122.01 |
| | | ST-SPCNN | 4 | 4 | 5 | 5 | 80.00 | 80.00 | 20.00 | 20.00 | 94.00 | 91.91 |
| | | MT-SPCNN | 4 | 4 | 5 | 5 | 80.00 | 80.00 | 20.00 | 20.00 | 94.17 | 92.01 |
| 中光 Sunlight with medium | 15:00 | NSCT | 4 | 4 | 7 | 7 | 57.14 | 57.14 | 42.86 | 42.86 | 52.84 | 52.79 |
| | | ImNSCT | 4 | 5 | 7 | 7 | 57.14 | 71.43 | 42.86 | 28.57 | 54.08 | 53.78 |
| | | DWT | 4 | 4 | 7 | 7 | 57.14 | 57.14 | 42.86 | 42.86 | 1.02 | 1.02 |
| | | SPCNN | 4 | 4 | 7 | 7 | 57.14 | 57.14 | 42.86 | 42.86 | 122.77 | 121.28 |
| | | ST-SPCNN | 7 | 7 | 7 | 7 | 100.00 | 100.00 | 0.00 | 0.00 | 95.28 | 93.32 |
| | | MT-SPCNN | 6 | 7 | 7 | 7 | 85.71 | 100.00 | 14.29 | 0.00 | 95.45 | 96.01 |
| | 16:00 | NSCT | 4 | 4 | 6 | 6 | 66.67 | 66.67 | 33.33 | 33.33 | 52.73 | 52.95 |
| | | ImNSCT | 4 | 4 | 6 | 6 | 66.67 | 66.67 | 33.33 | 33.33 | 53.80 | 53.72 |
| | | DWT | 4 | 4 | 6 | 6 | 66.67 | 66.67 | 33.33 | 33.33 | 1.04 | 1.03 |
| | | SPCNN | 3 | 3 | 6 | 6 | 50.00 | 50.00 | 50.00 | 50.00 | 122.34 | 121.78 |
| | | ST-SPCNN | 5 | 5 | 6 | 6 | 83.33 | 83.33 | 16.67 | 16.67 | 96.03 | 93.50 |
| | | MT-SPCNN | 5 | 5 | 6 | 6 | 83.33 | 83.33 | 16.67 | 16.67 | 95.94 | 95.83 |
| 弱光 Sunlight with weak | 18:20 | NSCT | 4 | 4 | 5 | 5 | 80.00 | 80.00 | 20.00 | 20.00 | 52.97 | 53.35 |
| | | ImNSCT | 4 | 3 | 5 | 5 | 80.00 | 60.00 | 20.00 | 40.00 | 53.76 | 53.82 |
| | | DWT | 4 | 4 | 5 | 5 | 80.00 | 80.00 | 20.00 | 20.00 | 1.02 | 1.02 |
| | | SPCNN | 3 | 3 | 5 | 5 | 60.00 | 60.00 | 40.00 | 40.00 | 121.02 | 115.18 |
| | | ST-SPCNN | 5 | 4 | 5 | 5 | 100.00 | 80.00 | 0.00 | 20.00 | 93.94 | 94.12 |
| | | MT-SPCNN | 5 | 4 | 5 | 5 | 100.00 | 80.00 | 0.00 | 20.00 | 94.01 | 93.64 |
| | 19:00 | NSCT | 5 | 5 | 7 | 7 | 71.43 | 71.43 | 28.57 | 28.57 | 52.77 | 52.68 |
| | | ImNSCT | 5 | 5 | 7 | 7 | 71.43 | 71.43 | 28.57 | 28.57 | 53.63 | 53.66 |
| | | DWT | 5 | 5 | 7 | 7 | 71.43 | 71.43 | 28.57 | 28.57 | 1.04 | 1.01 |
| | | SPCNN | 4 | 5 | 7 | 7 | 57.14 | 71.43 | 42.86 | 28.57 | 116.84 | 122.94 |
| | | ST-SPCNN | 6 | 7 | 7 | 7 | 85.71 | 100.00 | 14.29 | 0.00 | 93.34 | 96.81 |
| | | MT-SPCNN | 6 | 7 | 7 | 7 | 85.71 | 100.00 | 14.29 | 0.00 | 96.74 | 95.14 |

表 4 6 种融合模型融合质量评价指标比较

Table 4 Comparison of fusion quality evaluation indicators of six fusion models

| 指标 Index | NSCT | | ImNSCT | | DWT | | SPCNN | | ST-SPCNN | | MT-SPCNN | |
|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | 强度+可见 光图像 | 置信+可见 光图像 | 强度+可见 光图像 | 置信+可见 光图像 | 强度+可见 光图像 | 置信+可见 光图像 | 强度+可见 光图像 | 置信+可见 光图像 | 强度+可见 光图像 | 置信+可见 光图像 | 强度+可见 光图像 | 置信+可见 光图像 |
| AG | 10.10 | 9.29 | 11.02 | 6.54 | 7.69 | 8.13 | 9.80 | 11.51 | 12.03 | 11.66 | 11.96 | 11.92 |
| CC(ToF) | 0.48 | 0.49 | 0.58 | 0.72 | 0.66 | 0.65 | 0.18 | 0.04 | 0.50 | 0.85 | 0.82 | 0.81 |
| CC(RGB) | 0.84 | 0.75 | 0.75 | 0.50 | 0.75 | 0.65 | 0.84 | 0.95 | 0.72 | 0.20 | 0.42 | 0.24 |
| CE | 1.92 | 2.61 | 0.68 | 3.06 | 0.79 | 1.99 | 0.90 | 2.34 | 0.49 | 2.83 | 0.38 | 2.66 |
| E | 6.72 | 6.50 | 6.87 | 6.22 | 6.83 | 6.75 | 6.76 | 7.10 | 7.10 | 7.02 | 7.30 | 7.06 |
| MI(ToF) | 0.37 | 0.34 | 0.47 | 0.60 | 0.59 | 0.56 | 0.91 | 0.66 | 0.40 | 1.00 | 0.93 | 0.91 |
| MI(RGB) | 1.34 | 0.99 | 0.95 | 0.49 | 1.29 | 1.44 | 3.66 | 5.73 | 1.14 | 0.36 | 0.57 | 0.33 |
| SF | 20.27 | 19.67 | 20.61 | 14.75 | 15.36 | 17.08 | 19.99 | 22.06 | 22.86 | 23.45 | 22.67 | 23.80 |

注：AG 代表融合图像平均梯度；CC(ToF)代表 ToF 图像与融合图像之间的相关系数；CC(RGB)代表 RGB 图像与融合图像之间的相关系数；CE 代表 ToF 图像、RGB 图像与融合图像之间的总交叉熵；E 代表融合图像的熵；MI(ToF)代表 ToF 图像与融合图像之间的互信息；MI(RGB)代表 RGB 图像与融合图像之间的互信息；SF 代表融合图像的空间频率。其中 CE 值越小代表融合图像质量越好，其他评价指标值越大代表融合图像质量越好。

Note: AG represents the average gradient of the fused image; CC(ToF) represents the correlation coefficient between ToF image and fused image; CC(RGB) represents the correlation coefficient between RGB image and fused image; CE represents the total cross entropy between ToF image, RGB image and fused image; E represents the entropy of the fused image; MI(ToF) represents the mutual information between the ToF image and the fused image, MI(RGB) represents the mutual information between the RGB image and the fused image; SF represents the spatial frequency of the fused image. Among them, the smaller CE value means the better quality of the fused image, and the larger the value of other evaluation indicators means the better quality of the fused image.

模型复杂度比较见表 5 所示。ST-SPCNN 及 MT-SPCNN 模型在时间复杂度上比其他模型分别多出单目标参数寻优时间 $O(\max\text{Cycle} \times \text{NP} \times \text{limit} + \max\text{Cycle} \times \text{NP})$ 和多目标参数寻优时间 $O(\max\text{Cycle} \times (\text{NP} \times \text{limit} \times \text{Rep} \times \text{Rep} + \text{Rep} \times \text{Target} \times \text{nGrid} + \text{Rep} \times \text{Rep} \times \text{Target} + \text{NP} \times \text{Target}))$ 。空间复

杂度比其他模型分别多出参数寻优存储空间 $O(\text{NP} \times d)$ 和 $O(\text{NP} \times d \times \text{Target} + \text{Rep} \times d \times \text{Target} + \text{Target} \times \text{nGrid})$ 。实例分析可知，本文模型时间复杂度比 NSCT 模型多一个数量级，比 SPCNN 模型低一个数量级，在空间复杂度上与之对等不变。

表 5 6 种融合模型复杂度比较

Table 5 Complexity comparison of six fusion models

| 模型 Model | 时间复杂度 Time complexity | 时间复杂度实例 Time complexity instance | 空间复杂度 Spatial complexity | 空间复杂度实例 Spatial complexity instance |
|-------------|---|-------------------------------------|--|--|
| NSCT | $O(M \cdot N + \text{nsc} \cdot 2^q)$ | 307 224 | $O(M \cdot N + \text{nsc} \cdot 2^q)$ | 307 224 |
| ImNSCT | $O(M^2 \cdot N^2 + \text{nsc} \cdot 2^q \cdot M \cdot N)$ | 94 379 212 800 | $O(M \cdot N + \text{nsc} \cdot 2^q)$ | 307 224 |
| DWT | $O(M \cdot N)$ | 307 200 | $O(M \cdot N)$ | 307 200 |
| SPCNN | $O(n \cdot n + \text{iteration} \cdot M \cdot N)$ | 15 360 025 | $O(M \cdot N)$ | 307 200 |
| ST-SPCNN | $O(\max\text{Cycle} \cdot \text{NP} \cdot \text{limit} + \max\text{Cycle} \cdot \text{NP} + n \cdot n + \text{iteration} \cdot M \cdot N)$ | 1 843 285 | $O(\text{NP} \cdot d + M \cdot N)$ | 307 212 |
| MT-SPCNN | $O(\max\text{Cycle} \cdot (\text{NP} \cdot \text{limit} \cdot \text{Rep} \cdot \text{Rep} + \text{Rep} \cdot \text{Target} \cdot \text{nGrid} + \text{Rep} \cdot \text{Rep} \cdot \text{Target} + \text{NP} \cdot \text{Target} + n \cdot n + \text{iteration} \cdot M \cdot N))$ | 1 872 085 | $O(\text{NP} \cdot d \cdot \text{Target} + \text{Rep} \cdot d \cdot \text{Target} + \text{Target} \cdot \text{nGrid} + M \cdot N)$ | 307 596 |

注：nsc 代表 NSCT 分解级数；q 代表 NSCT 分解方向数；M、N 代表输入高低频子带图像的长度和宽度；iteration 代表点火次数；n 代表链接系数 $\mathbf{W}_{ij,kl}$ 与周围相邻神经元 (k,l) 的神经元个数；NP 代表食物源信息；d 代表空间维度；maxCycle 代表 ABC 算法最大迭代次数；limit 代表食物源的停滞最大次数；Rep 表示非支配解的数量；nGrid 表示划分的网格的数量；Target 表示目标函数的数量。试验中，nsc=3，q=3，M=640，N=480，iteration=50(SPCNN)，iteration=6(本文模型)，n=5，NP=3，d=4，maxCycle=5，limit=3，Rep=20，nGrid=7，Target=4。

Note: nsc represents the number of NSCT decomposition levels; q represents the number of NSCT decomposition directions; M and N represents the length and width of the input high and low frequency subband image; iteration represents the number of ignition; n represents the number of neurons between the link coefficient $\mathbf{W}_{ij,kl}$ and the surrounding adjacent neurons (k,l) ; NP stands for food source information; d represents the spatial dimension; maxCycle represents the maximum number of iterations of ABC algorithm; limit represents the maximum number of stagnation of food sources; Rep represents the number of nondominated solutions; nGrid represents the number of divided grids; Target represents the number of objective functions. In the experiment, nsc=3, q=3, M=640, N=480, iteration=50(SPCNN), iteration=6(proposed model in this paper), n=5, NP=3, d=4, maxCycle=5, limit=3, Rep=20, nGrid=7, Target=4.

3 结 论

本文提出一种多尺度分解双寻优策略简化脉冲耦合神经网络的飞行时间与可见光果园苹果图像融合模型。模型通过双分支融合策略，分别利用多尺度分解单目标简化脉冲耦合神经网络融合模型、多尺度分解多目标简化脉冲耦合神经网络融合模型进行融合。

1) 提出的多尺度分解单目标简化脉冲耦合神经网络融合模型以及多尺度分解多目标简化脉冲耦合神经网络融合模型具有自适应点火机制，在点火分割迭代次数满足双寻优策略目标函数条件时，停止迭代，加快了简化

脉冲耦合神经网络模型点火分割的速度。同时，相较于简化脉冲耦合神经网络模型固定较大迭代次数，模型自适应点火次数较低，在 3~7 次左右，点火图果实分割清晰、无空洞、无过分割。

2) 多尺度分解的飞行时间与可见光异源图像融合模型动态优化简化脉冲耦合神经网络模型参数，具有自适应迭代次数，使得改进后简化脉冲耦合神经网络模型点火次数大大减少，提高了效率，在点火几次的情况下就可以达到快速自适应分割，提高了融合效率。中光 15:00 时段点火识别成功率达到了 100.00%，点火分割时间达到最低 91.91 s。

3) 与其他融合模型相比, 本文提出的融合模型各时段识别成功率均达到 80.00% 以上。其中强光 12:00、中光 15:00、弱光 18:20、19:00 时段融合图像识别成功率达到 100.00%; 融合时间低于简化脉冲耦合神经网络模型, 达到最低 92.68 s。提出模型在弱光 18:20 时段平均梯度值为 12.03、相关系数(飞行时间)值为 0.85、互信息(飞行时间)值为 1.00、熵值为 7.30、空间频率值为 23.80, 融合指标值均为最大, 总交叉熵值为 0.38, 值最小。模型时空复杂度与简化脉冲耦合神经网络模型相差不大。

本文提出的融合模型在多尺度分解单目标简化脉冲耦合神经网络融合模型的优化策略、多尺度分解多目标简化脉冲耦合神经网络融合模型的帕雷托最优解集的构造方法、以及融合策略上均可进一步深入研究, 优化策略尚存在寻找全局最优食物源的冗余操作, 增加了一定的时间复杂度, 此外, 本文模型在强光照条件下融合指标表现不佳, 在今后的研究中, 可以融合别的群体智能优化算法思想进行改进和提升, 对强光照条件下的融合质量有待进一步提升。此外, 今后还将利用深度学习方法对果园场景异源图像融合模型进一步研究和探索, 以提高图像融合的实时性和批量性。

[参 考 文 献]

- [1] 贾伟宽, 赵德安, 刘晓洋, 等. 机器人采摘苹果果实的 K-means 和 GA-RBF-LMS 神经网络识别[J]. 农业工程学报, 2015, 31(18): 175-183.
Jia Weikuan, Zhao De'an, Liu Xiaoyang, et al. Apple recognition based on K-means and GA-RBF-LMS neural network applicated in harvesting robot[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(18): 175-183. (in Chinese with English abstract)
- [2] Cheng B, Jin L, Li G. Infrared and visual image fusion using LNSST and an adaptive dual-channel PCNN with triple-linking strength[J]. Neurocomputing, 2018, 310: 135-147.
- [3] Wang X, Lin T, Jiang X, et al. Reliable fusion of ToF and stereo data based on joint depth filter[J]. Journal of Visual Communication and Image Representation, 2021, 74: 103006.
- [4] 杨勇, 刘家祥, 黄淑英, 等. 卷积自编码融合网络的红外可见光图像融合[J]. 小型微型计算机系统, 2019, 40(12): 2673-2680.
Yang Yong, Liu Jiaxiang, Huang Shuying, et al. Convolutional auto-encoding fusion network for infrared and visible image fusion[J]. Journal of Chinese Computer Systems, 2019, 40(12): 2673-2680. (in Chinese with English abstract)
- [5] 任恒乐, 徐方, 邸霏, 等. 基于深度相机的移动机器人自主跟随技术[J]. 计算机工程与设计, 2020, 41(2): 562-566.
Ren Hengle, Xu Fang, Di Pei, et al. Autonomous following technology of mobile robot based on depth camera[J]. Computer Engineering and Design, 2020, 41(2): 562-566. (in Chinese with English abstract)
- [6] Panigrahy C, Seal A, Mahato N K. MRI and SPECT image fusion using a weighted parameter adaptive dual channel PCNN[J]. IEEE Signal Processing Letters, 2020, 27: 690-694.
- [7] Okorie A, Makrogiannis S. Region-based image registration for remote sensing imagery[J]. Computer Vision and Image Understanding, 2019, 189: 102825.
- [8] Da C A L, Zhou J P, Do M N. The nonsubsampled contourlet transform: theory, design, and applications[J]. IEEE Transactions on Image Processing, 2006, 15(10): 3089-3101.
- [9] Thierry R, Lucien W. The wavelet transform for the analysis of remotely sensed images[J]. International Journal of Remote Sensing, 1993, 14(3): 615-619.
- [10] Chao R, Zhang K, Li Y. An image fusion algorithm using wavelet transform[J]. Acta Electronica Sinica, 2004, 5: 750-753.
- [11] Shah V P, Younan N H, King R L. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets[J]. IEEE Transactions on Geoscience & Remote Sensing, 2008, 46(5): 1323-1335.
- [12] Zhang J, Ma W, Wu Y, et al. Multimodal remote sensing image registration based on image transfer and local features[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(8): 1210-1214.
- [13] Yu L, Xun C, Hu P, et al. Multi-focus image fusion with a deep convolutional neural network[J]. Information Fusion, 2017, 36: 191-207.
- [14] Tang F, Liang S, Zhong T, et al. Postoperative glioma segmentation in CT image using deep feature fusion model guided by multi-sequence MRIs[J]. European Radiology, 2020, 30(7): 823-832.
- [15] Ying Z, Xiaoqun Z. Variational bimodal image fusion with data-driven tight frame[J]. Information Fusion, 2020, 55: 164-172.
- [16] 杨风暴, 董安冉, 张雷, 等. DWT、NSCT 和改进 PCA 协同组合红外偏振图像融合[J]. 红外技术, 2017, 39(3): 201-208.
Yang Fengbao, Dong Anran, Zhang Lei, et al. Infrared polarization image fusion using the synergistic combination of DWT, NSCT and improved PCA[J]. Infrared Technology, 2017, 39(3): 201-208. (in Chinese with English abstract)
- [17] Feng B Y, Anran D, Lei Z, et al. Infrared polarization image fusion based on combination of NSST and improved PCA[J]. Journal of Measurement Science and Instrument, 2016, 7(2): 176-184.
- [18] Yang Z, Ma Y, Lian J, et al. Saliency motivated improved simplified PCNN model for object segmentation[J]. Neurocomputing, 2018, 275: 2179-2190.
- [19] 王俭, 聂仁灿, 周冬明, 等. 多目标粒子群优化 PCNN 参数的图像融合算法[J]. 中国图象图形学报, 2016, 21(10): 1298-1306.
Wang Quan, Nie Rencan, Zhou Dongming, et al. Image fusion algorithm using PCNN model parameters of multi-objective particle swarm optimization[J]. Journal of Image and Graphics, 2016, 21(10): 1295-1306. (in Chinese with English abstract)
- [20] Zhu M, Chen C, Wu X. Infrared and visible image fusion using improved QPSO-PCNN algorithm [C]//2018 8th International Conference on Electronics Information and

- Emergency Communication (ICEIEC). Beijing: IEEE, 2018: 52-58.
- [21] Wang X, Li Z, Kang H, et al. Medical image segmentation using PCNN based on multi-feature grey wolf optimizer bionic algorithm [J]. Journal of Bionic Engineering, 2021, 18(3): 711-720.
- [22] Huang C, Tian G, Lan Y, et al. A new pulse coupled neural network (PCNN) for brain medical image fusion empowered by shuffled frog leaping algorithm[J]. Frontiers in neuroscience, 2019, 13: 210.
- [23] Dharini S, Jain S. A novel metaheuristic optimal feature selection framework for object detection with improved detection accuracy based on pulse-coupled neural network[J]. Soft Computing, 2021: 1-13.
- [24] Karaboga D. An idea based on honey bee swarm for numerical optimization[R]. Kayseri: Erciyes University, 2005.
- [25] Zhang Z. A flexible new technique for camera calibration[R]. Washington State: Microsoft Research, 1998.
- [26] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [27] Leng C, Zhang H, Li B, et al. Local feature descriptor for image matching: A survey[J]. IEEE Access. 2018, 7: 6424-6434.
- [28] Deb K. Multi-objective Optimization Using Evolutionary Algorithms[M]. Chichester: John Wiley & Sons, 2001.
- [29] Zhang J L, Zhao E. Fusion method for infrared and visible light images based on NSCT[J]. Laser & Infrared, 2013, 43(3): 320-323.
- [30] 杨艳春, 李娇, 王阳萍. 图像融合质量评价方法研究综述[J]. 计算机科学与探索, 2018, 12(7): 1021-1035.
- Yang Yanchun, Li Jiao, Wang Yangping. Review of image fusion quality evaluation methods[J]. Journal of Frontiers of Computer Science and Technology, 2018, 12(7): 1021-1035. (in Chinese with English abstract)
- [31] Temer A M. Basler ToF camera user's manual[EB/OL]. (2018-04-09)[2022-02-17]. <https://www.baslerweb.com/cn/sales-support/downloads/document-downloads/basler-tof-camera-users-manual/>.

Heterologous image fusion with multi-scale decomposition and dual optimization SPCNN in an apple orchard

Liu Liqun¹, Gu Renyuan¹, Zhou Yubo¹, Huo Jiuyuan²

(1. College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China;

2. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: Heterologous image fusion has been widely used to integrate multiple images into one. The fused images also present a higher definition, more significant edge intensity, and more information than the source image. There are different characteristics of image data collected by the various types of sensors. Among them, the depth sensor imaging used the Time of Flight (ToF) to realize the distance calculation using the ToF near-infrared light. The beneficial supplement has been commonly used for the visible light camera. The broad application can also be expected in the agriculture, medical treatment, quality inspection, and vision fields. However, the image acquisition of a single natural scene cannot fully meet the requirements for rapid and accurate identification of the fruits and positioning targets. The image fusion can be extended to the heterologous vision system using multi-objective optimization, particularly in the field of natural scenes. In this study, a multi-scale decomposition and dual optimization strategy was proposed to simplify the ToF and visible-light image fusion in an apple orchard using the Simplified Pulse Coupled Neural Network (SPCNN). A double strategy with parameter optimization was introduced into the SPCNN model for the fusion of Nonsubsampled Contourlet Transform (NSCT). The model included the registration module, coding area, multi-scale decomposition module, single target SPCNN fusion model, multi-target SPCNN fusion model, and decoding area. The heterologous vision system was also used to accurately register the ToF and visible light images. Four parameters of SPCNN model were encoded, including the link channel feedback term, link strength, dynamic threshold attenuation factor, and dynamic threshold amplification factor. The NSCT was used to decompose the image at multiple scales. The fusion rules in the SPCNN model were adopted with the improved artificial bee colony algorithm and double optimization, including the single- and multi-objective parameter optimization. Each binary vector was converted into the real parameters using the decoding area. The objective function of the double optimization was used as the iteration termination of the SPCNN model. Finally, the heterogeneous image fusion was implemented after the multi-scale inverse transformation. There was improved parameter optimization and iteration times of SPCNN model. The adaptive ignition times of the model were relatively low (about 3-7 times), indicating low ignition times, adaptive segmentation, and high efficiency. The success rate of ignition recognition reached 100.00%, and the minimum duration of ignition division reached 91.91 s at 15:00. Specifically, the success rate of fusion image recognition also reached 100.00% under different periods, including strong, medium, and weak light at 12:00, 15:00, 18:20, and 19:00, compared with the rest fusion models. The fusion time was much lower than that of SPCNN model, with a minimum of 92.68 s. The four fusion indexes were the largest in the weak light period of 18:20, including the average gradient, correlation coefficient, mutual information, entropy, and spatial frequency. The proposed model presented an excellent performance in accuracy, time-consuming, and model size. The finding can provide a supplement to the image hierarchical fusion.

Keywords: image recognition; models; Nonsubsampled Contourlet Transform; Simplified Pulse Coupled Neural Network; heterologous image fusion model; single objective strategy; multi-objective strategy