

自适应多阈值图像分割算法

龙建武¹, 李继豪¹, 曾谁飞^{2,3,4}

(1. 重庆理工大学计算机科学与工程学院, 重庆 400054; 2. 青岛海尔电冰箱有限公司, 山东 青岛 266700;
3. 海尔优家智能科技(北京)有限公司, 北京 100006; 4. 北京邮电大学计算机学院(国家示范性软件学院), 北京 100876)

摘要: 针对当前大部分多阈值分割方法存在最优阈值组合定位难、阈值增多导致计算复杂度指数增长的问题, 提出了一种自适应多阈值图像分割算法。首先, 通过双边滤波对直方图进行平滑处理, 采用谷底筛选策略有效压缩阈值搜索空间; 接着, 基于动态规划算法, 将多阈值搜索问题转化为矩阵极值搜索问题, 并结合四边形不等式特性, 使用分治策略搜索代价矩阵最大值, 进一步提高搜索效率; 此外, 构建基于直方图谷底特征的目标函数, 自动确定最佳分割类数, 同时将 RGB 这 3 个通道直方图各自得到的最佳分割类数进行合并, 以获得最佳阈值进而完成彩色图像分割问题; 最后, 在 BSDS500 与 MSRC 数据集上进行系统性实验, 验证其在处理不同场景时的有效性与适用性。

关键词: 多阈值分割; 矩阵搜索; 动态规划; 分治策略

中图分类号: TP391

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025148

Adaptive multi-threshold image segmentation algorithm

LONG Jianwu¹, LI Jihao¹, ZENG Shuifei^{2,3,4}

1. College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China
2. Qingdao Haier Refrigerator Co., Ltd., Qingdao 266700, China
3. Haier Uplus Intelligent Technology (Beijing) Co., Ltd., Beijing 100006, China
4. School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: Aiming at the issues that most existing multi-threshold segmentation methods struggle to accurately locate the optimal threshold combination and face soaring computational complexity with more thresholds, an adaptive multi-threshold image segmentation algorithm (AMTISA) was proposed. First, histograms were smoothed with bilateral filtering and a valley-bottom screening strategy was used to shrink the threshold search space. Then, the multi-threshold search was turned into a matrix extreme value search using dynamic programming. By using quadrilateral inequality properties, a divide-and-conquer approach found the cost matrix's maximum value, boosting search efficiency. A target function based on histogram valley features was constructed to automatically determine the optimal number of segmentation classes. The optimal numbers from RGB channels merged for best color image segmentation threshold. Finally, experiments on BSDS500 and MSRC datasets confirm the algorithm's effectiveness and adaptability.

Keywords: multi-threshold segmentation, matrix search, dynamic programming, divide-and-conquer strategy

收稿日期: 2025-07-17; 修回日期: 2025-08-15

通信作者: 曾谁飞, zengshuifei@139.com

基金项目: 重庆市教育委员会科学技术研究计划青年项目(No.KJQN202201148); 重庆市教育委员会人文社会科学研究青年项目(No.23SKGH263); 重庆理工大学研究生教育高质量发展行动计划基金资助项目(No.gzlxc20253270)

Foundation Items: The Young Scientists Project of Science and Technology Research Program of Chongqing Education Commission (No.KJQN202201148), Humanities and Social Sciences Research Program of Chongqing Municipal Education Commission (No.23SKGH263), The Funding Achievements of the Action Plan for High Quality Development of Graduate Education at Chongqing University of Technology (No.gzlxc20253270)

0 引言

图像分割作为计算机视觉领域的一项核心技术,其目的在于将图像细分为多个具备特定属性且彼此无交集的区域,进而精准提取图像中用户关注的目标对象,使其与背景或其他非关键部分相区分。该技术流程模拟了人类视觉系统对环境的感知及解析机制,通过捕捉图像中的边缘轮廓、几何形态或色彩分布等关键特征,将原本复杂的图像信息转化为结构清晰、便于处理的格式。

图像分割不仅是目标识别、动态追踪、类别区分等任务的重要基石,还在医学影像的精准分析、自动驾驶系统的环境感知、遥感图像的智能解译以及工业产品的质量检测等多个领域展现出不可或缺的应用价值,是推动图像处理向智能化、自动化方向发展的关键支撑。

当前,图像分割领域的算法呈现多样化态势,依据不同原理和特性划分为多种类别。诸如阈值分割法、基于边缘检测的分割技术、基于聚类分析的分割方法、区域生长算法和深度学习策略等。文献[1]提出基于优化模型的快速边缘检测新技术,先按像素细分图像,再依规则合并区域以缩短检测时间。文献[2]创立出一种针对计算机断层扫描(CT, computed tomography)图像中肝脏区域的自动标注方法,解决在医学图像上手动标注耗费大量时长问题。Zhong等^[3]引入一种名为四维簇语义(4D-CS, 4D cluster-semantic)的4D时空激光雷达语义分割方法,利用聚类先验信息提升4D时空激光雷达数据的语义分割性能。Xiao等^[4]提出基于目标簇分割的多视点视域分析并行算法,分解任务为子任务,利用GPU并行计算高效处理,提升计算效率。Zhang等^[5]提出基于Halcon的区域生长算法,应用于舌下静脉分割,从初始种子点依规则拓展区域,实现精确分割。文献[6]提出基于自适应区域生长与迭代阈值分析的红外小目标检测法,能自动确定种子像素、自适应选最优阈值,增强适应性。文献[7]提出基于深度学习的运动人体图像分割识别算法,可自动识别分割目标区域并精准分类识别。文献[8]创建了一种基于Transformer的统一框架DocSAM,能异构数据集联合训练,增强鲁棒性与泛化力,还降低计算存储资源需求。

多阈值分割算法相较于其他算法,具备独特优势。它不需要复杂边缘检测提取轮廓,规避了聚类

分析对初始参数敏感导致的分割误差;与区域生长算法相比,不需要确定种子点,减少人为干扰,且能同时处理多阈值,分割效率高;相较于深度学习,不依赖大量训练数据,更适配数据有限场景,算法实现简单、计算成本低,能快速应用于实际任务,实时性好。

基于多阈值分割算法的上述优点,文献[9]提出一种基于多阈值分割的节能路由协议(EE-RPMS, energy-efficient routing protocol based on multi-threshold segmentation),为无线传感器网络(WSNS, wireless sensor networks)应用提供有效方案,优化簇形成与簇头选择,提升网络能效和寿命。Jiang等^[10]提出基于贝叶斯预测进化算法的快速多阈值图像分割法,借多级阈值技术提升分割效率与质量。Deng等^[11]提出一种结合最大类间方差法(OTSU)阈值分割和自适应角点检测的算法,显著提升角点检测的准确性和鲁棒性。

当前图像处理与分析领域中,尽管多阈值分割方法凭借其能够依据图像不同特征划分多区域的优势被广泛应用,但大部分现有的多阈值分割方法仍面临着一系列棘手难题。一方面,确定最优阈值组合时,由于图像数据本身的复杂性和多样性,阈值之间相互关联且对分割结果影响各异,难以精准定位到真正能使分割效果达到最佳的那组阈值;另一方面,随着阈值数量的增加,分割计算复杂度呈指数级增长,限制了其在实时图像处理场景的应用。针对上述问题,本文在合理选取阈值组合以及有效降低计算复杂度方面展开研究,提出一种自适应多阈值图像分割算法(AMTISA, adaptive multi-threshold image segmentation algorithm)。

本文主要贡献包括以下几点。

1) 基于阈值搜索时有效阈值多集中于谷底位置,本文设计谷底筛选策略,将候选阈值由全灰度级压缩至显著谷底集合,有效缩小搜索空间并抑制噪声干扰。

2) 现有多阈值图像分割方法存在搜索范围广、计算复杂、耗时久的问题。本文将动态规划与四边不等式驱动的分治剪枝相结合,把指数复杂度 $O(L^K)$ 降至 $O(KL\lg(L))$,在保证全局最优的同时实现实时处理。

3) 构造含谷底权重惩罚的自适应目标函数,不需要人工预设即可自动确定最佳分割类别数,并

通过对RGB这3个通道阈值的融合获得高保真彩色分割结果。

1 相关工作

本文相关工作主要涉及基于多阈值分割算法以及OTSU算法,接下来对相关工作进行深入阐述。

1.1 多阈值分割算法

多阈值图像分割^[12-16]的任务就是根据像素的灰度级,将图像划分为多个具有不同灰度特征的区域。

假设一张图像的灰度级为 L ,灰度值 $i = \{0, 1, \dots, L-1\}$;给定阈值 $T = \{T_1, T_2, \dots, T_{K-1}\}$;将像素划分为 K 个类别,即 $\{C_1, C_2, \dots, C_K\}$,其中第 k 个类别阈值区间表示为 $C_k = \{i | T_{k-1} \leq i < T_k\}$,其中 $k \in \{1, 2, \dots, K\}$, $T_0 = 0$, $T_K = L$ 。

归一化直方图中,图像中灰度级 i 出现概率估计近似为

$$p(i) = \frac{n_i}{N} \quad (1)$$

其中, n_i 表示灰度值为 i 的像素点个数, N 表示像素点总数。

对于每个类别,可计算类别概率,定义为

$$\omega[T_{k-1}, T_k] = \sum_{i=T_{k-1}}^{T_k-1} p(i) \quad (2)$$

其中, $\omega[T_{k-1}, T_k]$ 表示阈值区间 $[T_{k-1}, T_k)$ 内类别的权重。

类别均值计算式为

$$\mu[T_{k-1}, T_k] = \sum_{i=T_{k-1}}^{T_k-1} \frac{p(i)i}{\omega[T_{k-1}, T_k]} \quad (3)$$

类别方差计算式为

$$\sigma^2[T_{k-1}, T_k] = \sum_{i=T_{k-1}}^{T_k-1} (i - \mu[T_{k-1}, T_k])^2 \frac{p(i)}{\omega[T_{k-1}, T_k]} \quad (4)$$

最优阈值指的是能够使目标函数取得最小值或最大值的阈值集合。为表述简洁且保证通用性,以下仅呈现目标函数 J 被最大化时的情形,表达式为

$$\{T_1^*, T_2^*, \dots, T_{K-1}^*\} = \arg \max_{0 \leq T_1 < T_2 < \dots < T_{K-1} < L} [J(T_1, T_2, \dots, T_{K-1})] \quad (5)$$

由式(5)得到的一组阈值可以将图像中的像素有效地划分为 K 个不同的类别,以实现最佳的图像分割效果。

1.2 多级最大类间方差算法

在统计学与图像处理领域,方差常用于衡量数据的离散程度。方差数值的大小能够直观体现数据点之间的差异状况,具体表现为:方差值越大,数据点的分布越松散;方差值越小,数据点越紧凑。最大类间方差法^[17-22]正是基于方差的这一特性来实现图像分割的目的。

结合1.1节可知,该图像的总均值 μ 计算式为

$$\mu = \sum_{i=0}^{L-1} ip_i \quad (6)$$

那么,类间方差 σ_B^2 计算式为

$$\sigma_B^2 = \sum_{k=1}^K \omega[T_{k-1}, T_k] (\mu[T_{k-1}, T_k] - \mu)^2 = \sum_{k=1}^K \omega[T_{k-1}, T_k] (\mu[T_{k-1}, T_k])^2 - \mu^2 \quad (7)$$

其中, μ 是常数。因此可将目标函数 J 记为

$$J(T_1, \dots, T_{K-1}) = \sum_{k=1}^K \omega[T_{k-1}, T_k] (\mu[T_{k-1}, T_k])^2。$$

最优阈值就是使式(7)最大化的那些阈值组合,即最大化 $J(T_1, T_2, \dots, T_{K-1})$,依据式(5),便能计算出最佳阈值组合,利用该阈值组合能够达成对图像的精确分割操作。

综上所述,相较传统穷举式多阈值方法,现有工作受人工设置类别数、指数复杂度及多类过分割等限制。本文算法通过谷底权重惩罚项与四边形不等式分治策略,实现类别数自适应且复杂度降至 $O(KL \ln(L))$ 。

2 算法模型

2.1 问题与动机

在阈值分割领域,自适应性差、实时性弱是多阈值分割算法的难题。多数方法采用穷举策略在整个灰度级上寻找最佳阈值,但随着阈值数量与灰度级不断增多时计算复杂度更是随阈值数量呈指数爆炸式增长。

以处理灰度级为 L 的图像并选取 K 个阈值为例,使用传统遍历搜索方法花费的时间复杂度高达 $O(L^K)$ 。如此高昂的时间成本,导致在实时性要求极为严苛的应用场景中受到限制。在实际图像分割任务里,由于难以提前知晓图像的最佳阈值数量,现有的众多方法无法自动选定与图像适配的阈值个数。通常情况下,只能依靠人工预先设定一个固定的阈

值数量，以此为基础开展图像分割工作。然而，这种人工设定固定阈值数量的做法，极大地制约了诸多方法在实际应用中的效果，使它们难以达成理想的图像分割性能。而且，当图像的灰度级增多、所需阈值数量增加时，搜索空间会呈指数级急剧扩大。在这样的情况下，如何实现分类数量的自动确定，同时有效缩小搜索范围、避免大量无效搜索，已成为当前图像分割领域亟待解决的关键问题。

本文通过提升搜索效率与开展快速全局搜索这两大策略来解决上述难题。首先，通过谷底筛选策略，有效压缩阈值搜索范围，进而缩小搜索空间。其次，融合动态规划（DP, dynamic programming）算法和分治策略，将多阈值搜索问题重构为矩阵最优解搜索问题，并结合四边形不等式特性，快速定位代价矩阵最大值，显著降低分割时间复杂度。最后，基于直方图谷底特征，在目标函数中融入谷底权重因子，构建分割类别数度量模型，经全局搜索确定最优阈值，实现图像自适应分割。

2.2 图像直方图预处理

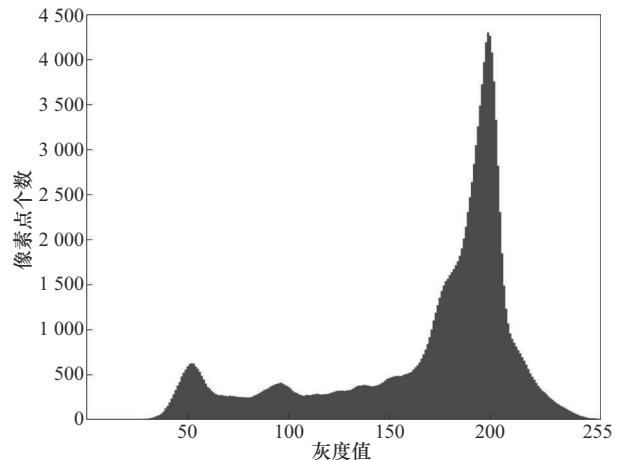
直方图作为一种统计图像灰度分布的常用工具，在实际应用中由于噪声的干扰，其形态往往会出现大量局部波动。这种现象使阈值搜索过程变得更为复杂，极大地增加了阈值搜索的复杂度。为有效确定阈值的分布位置，本文利用双边滤波对直方图进行平滑，减少直方图中噪声的干扰，平滑直方图及其备选阈值如图 1 所示。其中，图 1(b)为原始直方图，图 1(c)为平滑后直方图，其横轴均代表图像的灰度值，用于体现图像像素的亮度层次情况；纵轴则表示每个灰度值在图像中对应的像素数量，反映出对应灰度值在整幅图像里的分布状况。从原始直方图到平滑后直方图的变化，能够直观呈现双边滤波对直方图进行平滑处理以减少噪声干扰的效果，为确定合适的阈值提供更有利的依据。

基于平滑后的直方图，研究发现，阈值通常处于直方图的谷底位置，如图 1(c)中白色竖线所示。原因在于，谷底所在的那个灰度级区间像素数量很稀疏，此类谷底往往与图像中不同类别间的过渡区域相对应，故可利用波谷对图像进行划分。

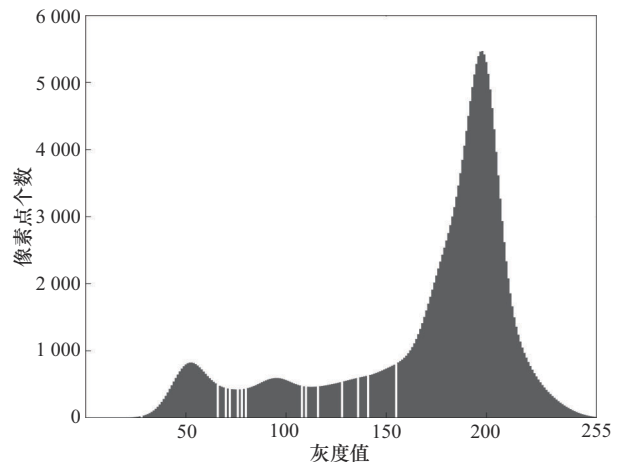
本文通过分析像素点的邻接值与它自身的关联，将直方图灰度变化分为上升状态、下降状态、极大值状态以及谷底状态这 4 种情形。为筛选出备选阈值，定义波谷度量函数 $V(i)$ 为



(a) 原始图像



(b) 原始直方图



(c) 平滑后直方图

图 1 平滑直方图及其备选阈值

$$V(i) = (p[i - \Delta] - p[i]) + (p[i + \Delta] - p[i]) \quad (8)$$

其中， Δ 为计算差值的窗口大小，本文设 Δ 为 1。在谷底筛选策略中，从信息粒度与噪声抑制权权衡看， $\Delta = 1$ 以最小可分辨尺度滑动检测，能确保对真实谷底响应无时延且不丢失细节；与 Nyquist-Shannon 采样定理^[23-24]离散事件类比，其满足“采样频率”要求，可最大限度保留极值信息且不引入

假阳性。当 $V(i) > 0$ 时, 其值越大, 对应的谷就越显著; 当 $V(i) < 0$ 时, 则该波谷点肯定不是谷底。

为了更精确地确定备选阈值范围, 那些 $V(i)$ 值较小对应的亮度值将被忽略, 因为它们成为谷底的概率极低。于是本文先根据波谷度量函数计算 $V(i)$ 的值, 对 $V(i) > 0$ 的波谷点按从大到小的顺序进行排序, 并将 $V(i) > 0$ 的部分记作集合 S 。在集合 S 中从大到小选取前 M 个亮度值作为备选阈值空间 Ω , 计算式为

$$\Omega = [\text{sort}(S)]_M \quad (9)$$

其中, $\text{sort}(\cdot)$ 表示对集合 S 中的波谷点进行降序排列, $[\cdot]_M$ 表示从排序后的类别中选择概率最高的前 M 个。

通过上述处理, 能够筛选出有效的波谷作为候选阈值。这一方法相比直接在原始灰度级上搜索, 显著缩小了阈值的搜索范围, 因此, 后续操作只需在这个精简后的灰度级集合上进行阈值选取, 从而大幅提升了处理效率。

2.3 基于 DP 算法的多阈值搜索

本文经过分析可知, 多阈值分割方法的核心在于通过设定一系列有序阈值, 将图像的灰度空间划分成多个不同的区域子区间。

具体而言, 设定灰度范围 $[0, L)$ 为主区间。当按照一定顺序插入若干个阈值, 主区间就会被分割成 k 个连续的子区间, 用数学表达式表示就是 $[0, T_1), [T_1, T_2), \dots, [T_{K-1}, L)$ 。本文将求解整个区间 $[0, L)$ 的最大总代价问题, 拆解为在各个子区间内分别求解最值问题。而且, 每个子区间的最值求解过程, 都是在前一个阶段最值求解结果的基础上进行的。结合 DP 算法的思想, 可以将多阈值搜索问题转化为在每个子区间内求解最大类代价问题。

子问题可被定义为: 在给定的灰度值范围 $[0, L)$ 内, 探寻能将其划分为 K 个类的最优阈值组合。其最优解可通过式(10)来明确给出, 即

$$J^*(T_1^*, T_2^*, \dots, T_{K-1}^*) = \max_{0 \leq T_1 < \dots < T_{K-1} < L} \sum_{k=1}^K l[T_{k-1}, T_k] \quad (10)$$

其中, J 表示在区间 $[0, L)$ 内的最大代价值, $l[T_{k-1}, T_k]$ 表示在 2 个阈值 T_{k-1} 和 T_k 之间的代价成本, 其计算式为

$$l[T_{k-1}, T_k] = \sum_{i=T_{k-1}}^{T_k-1} \omega[T_{k-1}, T_k] (\mu[T_{k-1}, T_k])^2 \quad (11)$$

通过对子问题最优解进行重新表述与推导, 式(5)可等价于

$$T^* = \arg \max_T \sum_{k=1}^K l[T_{k-1}, T_k] = \arg \max_T (\sum_{k=1}^{K-1} l[T_{k-1}, T_k] + l[T_{K-1}, T_K]) \quad (12)$$

其中, T^* 表示最优解, 得到一组阈值 $T = \{T_1, T_2, \dots, T_{K-1}\}$, 使总代价最大化。由式(13)可以得到每段区间最佳阈值, 其递推式为

$$\begin{aligned} T_{K-1}^* &= \arg \max_{T_{K-1}} \{ J^*(T_1^*, T_2^*, \dots, T_{K-2}^*) + l[T_{K-1}, T_K] \} \\ T_{K-2}^* &= \arg \max_{T_{K-2}} \{ J^*(T_1^*, T_2^*, \dots, T_{K-3}^*) + l[T_{K-2}, T_{K-1}] \} \\ &\vdots \\ T_k^* &= \arg \max_{T_k} \{ J^*(T_1^*, T_2^*, \dots, T_{k-1}^*) + l[T_k, T_{k+1}] \} \\ &\vdots \\ T_1^* &= \arg \max_{T_1} \{ l[T_1, T_2] \} \end{aligned} \quad (13)$$

其中, $K-1$ 表示最终期望的阈值个数; $T_1^*, T_2^*, \dots, T_{K-1}^*$ 表示当前阶段正在求解的最佳阈值; J^* 表示此前累积的最大代价; l 表示当前阶段新增代价。

通过递推式(13), 把一次性求 $K-1$ 个最优阈值的全局问题, 拆解成 $K-1$ 个局部决策, 并用动态规划从后向前回溯出唯一的最优阈值序列。本文以灰度级 $L=5$ 和分割类数 $K=4$ 为例构建 DP 状态矩阵 A 。具体步骤如下。

1) 确定阈值搜索范围。根据上述分析可知, 第 k 个阈值的搜索范围是 $[k, L - K + k]$ 。表 1 列出了阶段数与每阶段阈值的候选范围。

阶段	候选灰度级	含义
1	1, 2	第 1 个阈值候选位置
2	2, 3	第 2 个阈值候选位置
3	3, 4	第 3 个阈值候选位置

2) 构建 DP 状态矩阵。在 DP 中, 阶段路径的构建是一个核心步骤, 它涉及如何通过一系列的决策步骤达到最优解。具体而言, 阶段路径描述了在不同的灰

度级上挑选最佳阈值组的过程，每个阶段都与一个阈值的选择紧密相连，而路径则完整地呈现了从初始状态逐步过渡到结束状态过程中所做出的一系列最优决策。在实际操作中，需要逐一权衡各个阶段的代价值。具体做法是，通过对比各阶段二维搜索矩阵中的元素，从中找出最大值，以此作为确定最优路径的重要依据。上述例子中，可能的阈值组合有 $\{T_1, T_2, T_3\}$ 、 $\{T_1, T_2, T_4\}$ 、 $\{T_1, T_3, T_4\}$ 、 $\{T_2, T_3, T_4\}$ 。

DP 阶段路径如图 2 所示，横轴代表灰度级，纵轴代表阶段。每个阶段的搜索范围由前一阶段的阈值决定，形成了一个逐步缩小的搜索空间。可以发现，上述穷举出的阈值组合刚好对应图 2 中的各条路径，因此原最优化问题式(12)将被转化为求解最优路径问题。

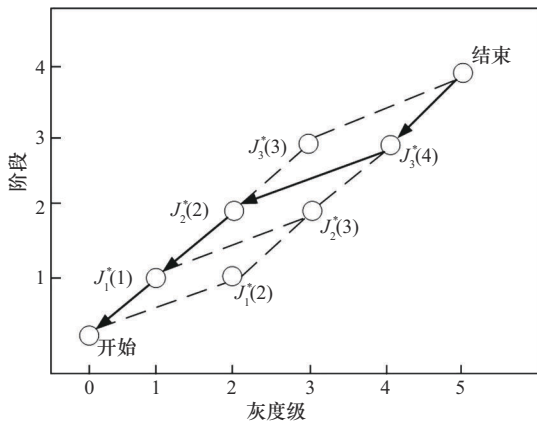


图 2 DP 阶段路径

各个阶段的搜索矩阵构建过程是一个关键步骤，它涉及为每个阶段确定可能的阈值选择及其相应的代价。这一过程开始于定义搜索矩阵的维度，通常基于当前阶段的终点灰度级和起点灰度级。对于每个阶段，搜索矩阵的列 (r) 代表起点灰度级，行 (j) 代表终点灰度级。由于计算区间代价时，后灰度级必须大于前灰度级，因此矩阵的上半部分都是 $-\infty$ 。表 2~表 5 表示 4 个阶段的搜索矩阵。基于 DP 搜索最优路径算法 (DPSEARCH) 如算法 1 所示。

表 2 第一阶段搜索矩阵

j	r
	0
1	$J_1^*(1)$
2	$J_1^*(2)$

表 3 第二阶段搜索矩阵

j	r	
	1	2
2	$J_1^*(1) + l_{1,2}$	$-\infty$
3	$J_1^*(1) + l_{1,3}$	$J_1^*(2) + l_{2,3}$

表 4 第三阶段搜索矩阵

j	r	
	2	3
3	$J_2^*(2) + l_{2,3}$	$-\infty$
4	$J_2^*(2) + l_{2,4}$	$J_2^*(3) + l_{3,4}$

表 5 第四阶段搜索矩阵

j	r	
	3	4
2	$J_3^*(3) + l_{3,5}$	$J_3^*(4) + l_{4,5}$

3) 回溯找到最佳阈值。动态规划求解过程中，本文记录每个阶段产生的最大代价所对应的灰度值。待所有阶段的计算与记录完成后，通过回溯方法，沿着逻辑路径 (如图 2 中黑色实线箭头所指方向) 逐步追溯，从而精准定位能够达成最优效果的最佳阈值。本例最佳阈值组合选取 $T = \{T_1, T_2, T_4\}$ 。

从上述过程可知第一阶段和最后一阶段的搜索空间为一维向量，在线性时间内即可完成最优值的搜索。因此，依据上述步骤规律，对剩余 $K - 2$ 个阶段均可构造出状态矩阵 A ，其具体计算式为

$$A(j,r) = \begin{cases} -\infty, & r > j \\ J_{k-1}^*(r+k-2) + l_{r+k-2j+k-1}, & r \leq j \end{cases} \quad (14)$$

算法 1 DPSEARCH

输入 直方图，类别数 K ，灰度级个数 L

输出 最佳阈值组合 $\{T_1^*, T_2^*, \dots, T_{K-1}^*\}$

- 1) 初始化：起点代价 $J = 0$ ；前一阶段的最优阈值位置 $\text{pred.node} = \text{Null}$
- 2) for $s = 1$ to K do
- 3) for $i = s$ to $L - K - s$ do
- 4) $J_{\max} = -\infty$
- 5) for each i in $s - 1$ do
- 6) $J_{\text{temp}} = J_{s-1}^* + l_{rj}$
- 7) if $J_{\text{temp}} > J_{\max}$ then
- 8) $J_{\max} = J_{\text{temp}}$

```

9)      pred.node = i
10)     end if
11)     end for
12)  end for
13)end for
// 反向跟踪最佳阈值过程
14)初始化: s = K,i = L - 1
15)while s ≥ 2 do
16)    Ts-1* = pred.node
17)    i = Ts-1*
18)end while
return {T1*,T2*,⋯,TK-1*}

```

综上所述, DP 算法遵循一种循序渐进的逻辑推理与计算机制。在算法的推进过程中, 首先要明确前一阶段的最大代价, 然后再计算后一阶段的最大代价值。这些可能的代价值写成如式(15)所示的一个维度为 $(L - K + 1) \times (L - K + 1)$ 的矩阵。其中, $K - 1$ 个阶段对应着 $K - 1$ 个相同大小的矩阵, 在采用动态规划算法进行求解的过程中, 经过分析可知, 时间复杂度为 $O(KL^2)$ 。

2.4 融合分治策略的多阈值搜索

在 2.3 节基于 DP 算法搜索代价矩阵最大值的过程中, 常规做法是借助两层嵌套循环遍历矩阵内的所有元素, 此时算法的时间复杂度为 $O(L^2)$ 。但是随着灰度级 L 的增大, 这种计算方式会导致计算效率急剧降低。通过分析可知, 若目标函数具备特定性质, 便能够运用分治策略对矩阵展开搜索操作, 有效缩小搜索范围, 进而提升搜索效率。

假定代价函数 $l[T_{k-1}, T_k]$ 有如下关系, 即

$$\begin{aligned}
 & l[a,u] + l[b,v] \geq l[a,v] + l[b,u] \\
 & \text{s.t. } 0 < a < b < u < v < L
 \end{aligned} \tag{15}$$

该性质被定义为凸四边形不等式。若类成本具备上述四边形不等式性质, 可从矩阵 A 的下三角区域中选取 4 个元素, 且这 4 个元素需满足条件: $0 < r_1 < r_2 < j_1 < j_2 \leq L - K + 1$, 根据式(15)可以得出

$$A(j_1, r_1) + A(j_2, r_2) \geq A(j_1, r_2) + A(j_2, r_1) \tag{16}$$

式(16)之所以满足凸四边形不等式, 是因为类间方差函数本身对灰度区间的凸性: σ_B^2 在灰度级上是凸的, 因为其由概率加权的一阶和二阶统计量构成, 而平方差运算保持凸性; 当把区间代价 $A(r_1, j_2)$ 定义为该凸函数在区间 $[r_1, j_2]$ 上的增量时, 根据凸函数的二

阶差分非负特性, 对任意 $0 < r_1 < r_2 < j_1 < j_2 \leq L - K + 1$, 式(16)必将满足, 从而直接导出四边形不等式, 使分治策略在动态规划搜索中有效剪枝。

基于上述性质, 结合分治策略和四边形不等式特性, 可简化代价矩阵的最大值搜索。

使用分治策略来寻找 $m \times n$ 矩阵 A 的行最大值时, 首先在矩阵的 $\left\lfloor \frac{m}{2} \right\rfloor$ 中间行中搜索最大值。然后对无效区域进行剪枝, 当 $A(j_1, r_1) < A(j_1, r_2)$ 时, 根据式(16)易知 $A(j_2, r_1) < A(j_2, r_2)$ 。据此, 矩阵左下角区域的元素不可能是所在行的最大值, 故可忽略该区域元素的代价计算。最后递归分解, 将矩阵分为右下和左上 2 个子矩阵, 再对剩余子矩阵进行上述递归操作, 直至定位所有行的最大值。

以灰度级 $L = 8$ 和分割类数 $K = 4$ 为例构建 DP 状态矩阵 A , 其第三阶段的搜索矩阵如表 6 所示。首先, 需要搜索矩阵的中间行 (第 3 行) 中的最大代价值, 假设 $J_4^*(4) + l_{4,5}$ 是中间行最大值。基于分治策略和四边形不等式性质可知, 在执行其余行搜索操作时, 算法仅需聚焦于表中左上和右下 2 个方框所界定的区域来计算代价值。与算法 1 采用逐行遍历矩阵以搜索最大代价值的方式相比, 这种有针对性的搜索策略能够显著规避大量不必要的计算步骤, 进而在极大程度上降低算法的时间复杂度, 有效提升算法的运行效率。

表 6 第三阶段搜索矩阵

j	r				
	2	3	4	5	6
3	$J_2^*(2) + l_{2,3}$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
4	$J_2^*(2) + l_{2,4}$	$J_3^*(3) + l_{3,4}$	$-\infty$	$-\infty$	$-\infty$
5	$J_2^*(2) + l_{2,5}$	$J_3^*(3) + l_{3,5}$	$J_4^*(4) + l_{4,5}$	$-\infty$	$-\infty$
6	$J_2^*(2) + l_{2,6}$	$J_3^*(3) + l_{3,6}$	$J_4^*(4) + l_{4,6}$	$J_5^*(5) + l_{5,6}$	$-\infty$
7	$J_2^*(2) + l_{2,7}$	$J_3^*(3) + l_{3,7}$	$J_4^*(4) + l_{4,7}$	$J_5^*(5) + l_{5,7}$	$J_6^*(6) + l_{6,7}$

本文通过将分治策略引入 DP 算法, 优化传统 DP 算法在状态转移过程中的矩阵搜索效率, 将阈值分割的时间复杂度进一步降低至 $O(KL \log(L))$, 有效提高阈值分割的效率。DIVCONQ 算法伪代码如算法 2 所示。

算法 2 DIVCONQ

输入 矩阵 $A_{m \times n}$, 初始偏移量 offset

输出 最大代价值位置 max_pos

- 1) 初始化: $j_{\text{mid}} = \left\lfloor \frac{m}{2} \right\rfloor$; 上半子矩阵 $P = A_{(1 \dots j-1) \times (1 \dots r)}$; 下半子矩阵 $Q = A_{(j+1 \dots m) \times (r \dots n)}$; offset = 0; 搜索第 j 行中最大的位置 j_pos
- 2) if $m = 1$ then
- 3) return max_pos = 1
- 4) else
- 5) if $r \neq 1$ then
- 6) $P \leftarrow A_{(1 \dots j_{\text{mid}} - 1) \times (1 \dots j_{\text{mid}})}$
- 7) max_pos = DIVCONQ(P , offset)
- 8) end if
- 9) $Q \leftarrow A_{(j_{\text{mid}} + 1 \dots m) \times (j_{\text{pos}} \dots n)}$
- 10) max_pos = DIVCONQ(Q , offset + r)
- 11) end if

此外, 本文将展开证明过程, 以证实目标函数

$$\begin{cases} l(a,v) = \omega(a,v)\mu^2(a,v) \leq \alpha \frac{\omega(a,v)}{\omega(a,u)} l(a,u) + (1-\alpha)\alpha \frac{\omega(a,v)}{\omega(b,v)} l(b,v) \\ l(b,u) = \omega(b,u)\mu^2(b,u) \leq \beta \frac{\omega(b,u)}{\omega(a,u)} l(a,u) + (1-\beta)\alpha \frac{\omega(b,u)}{\omega(b,v)} l(b,v) \end{cases} \quad (20)$$

将 $l(a,v)$ 、 $l(b,u)$ 相加得

$$l(a,v) + l(b,u) \leq \frac{\alpha\omega(a,v) + \beta\omega(b,u)}{\omega(a,u)} l(a,u) + \frac{(1-\alpha)\omega(a,v) + (1-\beta)\omega(b,u)}{\omega(b,v)} l(b,v) \quad (21)$$

由式(18)易知

$$\begin{cases} \frac{\alpha\omega(a,v) + \beta\omega(b,u)}{\omega(a,u)} = 1 \\ \frac{(1-\alpha)\omega(a,v) + (1-\beta)\omega(b,u)}{\omega(b,v)} = 1 \end{cases} \quad (22)$$

将式(22)代入式(21)中, 得到 $l(a,u) + l(b,v) \geq l(a,v) + l(b,u)$, 故得证。

2.5 自适应多阈值分割算法

传统阈值搜索策略通常直接选取能使类间方差达到最大值时所对应的分类数量。然而, 目标函数的数值会随着分类数量的逐步增多而呈现出单调递增的趋势。这一特性会使算法在运行过程中, 更倾向于选择最大允许的阈值数目。换句话说, 算法最终会选定整个阈值搜索空间中阈值的最大个数, 进而引发过分割的问题。因此, 在以往的诸多阈值分割方法里, 往往需要人工预先设定分割的类数, 以

符合四边形不等式的要求。由式(7)得, $l(j,r) = \omega[T_{k-1}, T_k](\mu[T_{k-1}, T_k])^2$ 。 $0 \leq a < b < u < v < L$ 且平均值是单调非递减函数, 则 $\mu(a,u) \leq \{\mu(a,v), \mu(b,u)\} \leq \mu(a,v)$ 。因此, 存在 $\alpha, \beta \in [0,1]$ 使式(17)成立, 即

$$\begin{cases} \alpha\mu(a,u) + (1-\alpha)\mu(b,v) = \mu(a,v) \\ \beta\mu(a,u) + (1-\beta)\mu(b,v) = \mu(b,u) \end{cases} \quad (17)$$

对式(17)进行整理推导, 可得

$$\begin{cases} \alpha = \frac{\mu(a,v) - \mu(b,v)}{\mu(a,u) - \mu(b,v)} \\ \beta = \frac{\mu(b,u) - \mu(b,v)}{\mu(a,u) - \mu(b,v)} \end{cases} \quad (18)$$

又因为函数 x^2 为凸函数, 根据式(17)可得

$$\begin{cases} \alpha\mu^2(a,u) + (1-\alpha)\mu^2(b,v) \geq \mu^2(a,v) \\ \beta\mu^2(a,u) + (1-\beta)\mu^2(b,v) \geq \mu^2(b,u) \end{cases} \quad (19)$$

对式(19)同时乘以 $\omega(j,r)$, 可得

此来实现图像分割操作。

为攻克前文提及的算法缺陷, 例如, 经典的 OTSU 目标函数随着类别数 K 单调递增, 必然会趋向选择最大 K , 从而导致过分割问题。本文在原有目标函数里融入谷底权重项, 提升算法对直方图谷底特征的敏感性, 该方法可借助最大化目标函数的方式, 自动确定最为适宜的分割类数构建类数度量函数 $F(K)$, 表示为

$$F(K) = (1 - \sum_{i=1}^{K-1} p_{T_i^*}) J_K(T) \quad (23)$$

其中, $1 - \sum_{i=1}^{K-1} p_{T_i^*}$ 为谷底强调项, 这是一个概率的补集, 式(23)利用其概率补集性质对新增阈值施加指数级惩罚, 用于抑制阈值数量过多, 避免过分割; $J_K(T)$ 是当类别数为 K 时的最大代价函数值。随后, 在分割类数的搜索空间 Ω 中逐个计算 $F(K)$, 并取使函数最大的类别个数作为最佳分割类数 K^* , 数学表达式为

$$\begin{aligned} K^* &= \arg \max_K \{ F(K) \} \\ \text{s.t. } K &\in \{ 2, 3, \dots, |\Omega| \} \end{aligned} \quad (24)$$

其中, $|\Omega|$ 表示分割类数的最大搜索范围。

算法3为自适应多阈值分割算法 (AMTISA) 的伪代码。通过该方法能够有效规避人工指定阈值数带来的问题, 并且在保障分割结果准确性的同时, 提升了结果的可靠性。

算法3 AMTISA

输入 图像 I

输出 最佳分割类数 K^*

- 1) 计算灰度直方图 $p[0, \dots, L-1]$
- 2) 利用双边滤波平滑直方图
- 3) 计算总类别均值 μ ; 灰度级 i 出现概率 $p(i)$ 总和
- 4) for $K = 2$ to $|\Omega|$ do
- 5) 根据算法 DPSEARCH 和 DIVCONQ 得到最佳阈值组合 $\{T_1^*, T_2^*, \dots, T_{K-1}^*\}$
- 6) 根据式(23)计算类数度量 $F(K)$
- 7) end for
- 8) 根据式(24)得到最佳类别个数 K^*
- 9) return K^*

2.6 基于彩色图像的多阈值分割算法

在彩色图像的多阈值分割任务中, 首要步骤是对彩色图像的RGB这3个通道直方图执行平滑化操作。此操作旨在降低直方图中的噪声干扰, 让数据分布更为平滑, 进而从经过平滑处理的各通道直方图中筛选出潜在的备选阈值。接下来, 运用算法3分别对RGB这3个通道的直方图进行处理, 从而确定各通道对应的最佳分类数量。随后, 将得到的3个通道的最佳类别数量加以整合。最后, 借助算法从合并后的直方图中筛选出最优阈值组合, 凭借该阈值组合明确目标图像所包含的类别数量, 最终达成对彩色图像的精准分割。

假设输入一张尺寸为 $H \times W$ 的彩色图像, 该图像的像素点数量为 $N = H \times W$, 像素点集合 $x = \{x_1, x_2, \dots, x_N\}$, 类别标签集合 $y = \{y_1, y_2, \dots, y_K\}$ 。对于第 k 个类别标签 y_k , 其对应的R、G、B这3个通道的直方图被划分成不同的区间, 分别用 $C_k^{(R)}$ 、 $C_k^{(G)}$ 、 $C_k^{(B)}$ 表示。其中, $K_R = |C_k^{(R)}|$ 、 $K_G = |C_k^{(G)}|$ 、 $K_B = |C_k^{(B)}|$ 分别表示R、G、B这3个通道的直方图被划分的区间数量。那么, 整张彩色图像所包含的区间数量 K 可表示为

$$K = K_R \times K_G \times K_B \quad (25)$$

RGB直方图分析如图3所示, 其中, 图3(a)是原始图像, 图3(b)、图3(c)、图3(d)分别表示原始彩色图像的R、G、B这3个通道平滑后的直方图, 这些直方图展示了图像中各颜色通道的强度分布情况。

在3张直方图中, 白色竖线表示通过算法筛选出的谷底位置, 这些谷底可以作为图像分割的备用阈值。通过分析这些直方图, 可以了解图像在不同颜色通道中的对比度和亮度分布, 进而进行图像增强、分割或其他处理。

然后, 本文通过整合策略, 先对R、G、B各自分割, 将分割结果合并得到合并后的新直方图。最后, 将得到的直方图再次自适应分割, 即可获得最终的最优阈值。如此一来, 便能借助确定的全局最佳阈值组合, 实现对彩色图像的精确分割。

3 算法实验

为检验本文算法在分割任务中的性能表现, 选取 Mshift(mean shift)^[25]、MCG(multiscale combinatorial grouping)^[26]、CCBCut (compassionately conservative balanced cut)^[27]、ENCut_RWRT(explored normalized cut with random walk refining term)^[28]、NAM_HDBSCAN(nam-based hierarchical density-based spatial clustering)^[29]、HISFBDI(hierarchical image segmentation frame-work based on deep information)^[30]这6种图像分割算法作对比。上述所有比较算法均可在MATLAB2020b编程环境中实现, 并且所有测试均使用上述算法的默认参数进行。

3.1 评价指标与数据集

为了科学、严谨且高效地评估本文算法在分割任务中的性能表现, 经过审慎考量, 精心选取了SC (segmentation covering)、PRI (probabilistic rand index)、VI (variation of information) 以及ACT (average computational time) 作为核心评估指标。这4项指标能够从分割的准确性、一致性、信息变化以及计算效率等多个关键维度, 对本文算法的分割质量展开全面且客观的深度评价, 为算法性能的精准衡量提供坚实保障。

在开展实验验证工作中, 本文选定2个代表性的开源数据集, 即BSDS500^[31]数据集 (Berkeley segmentation data set 500) 与MSRC数据集 (Microsoft research cambridge dataset)^[32]。其中,

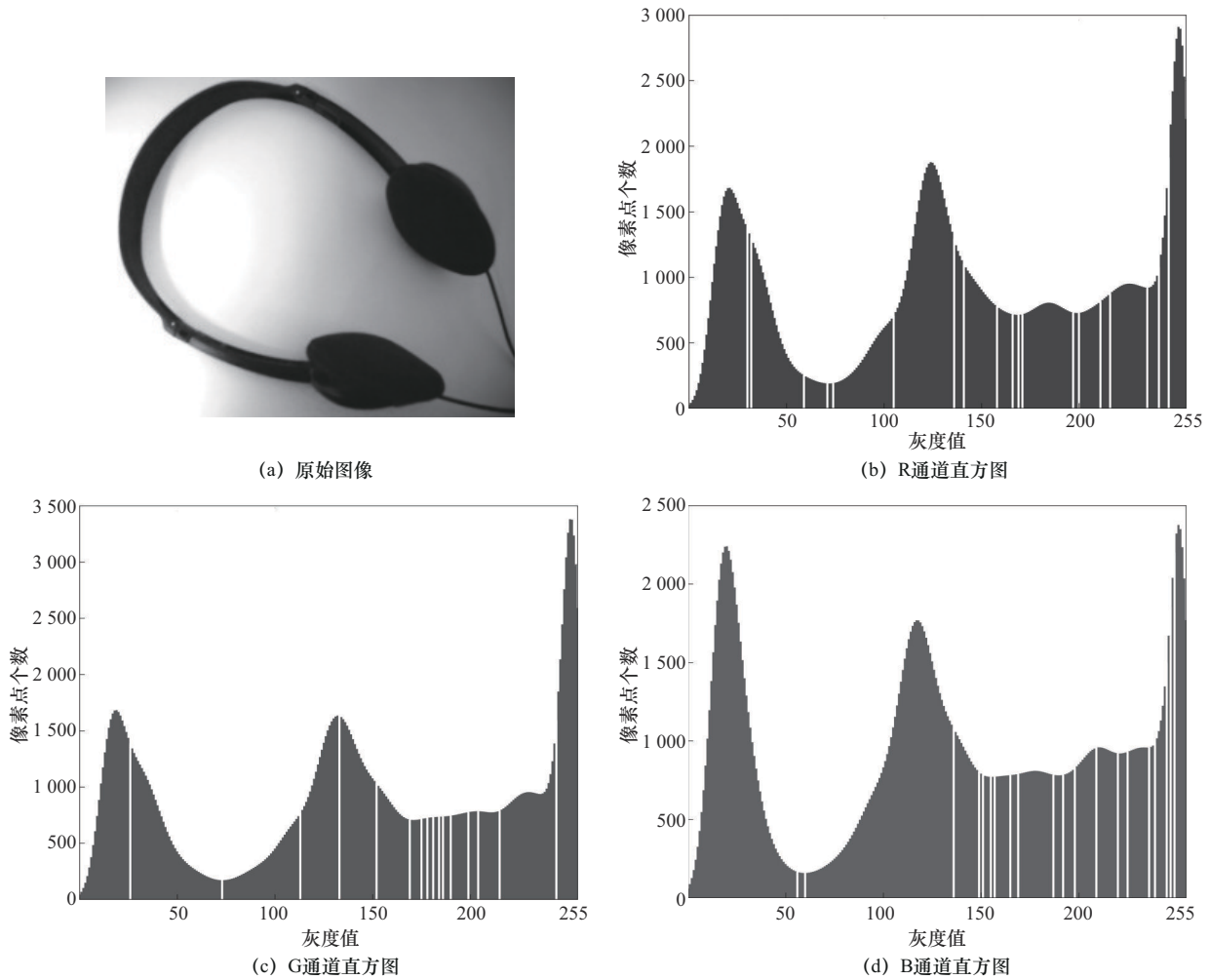


图3 RGB直方图分析

BSDS500 作为图像分割领域极具影响力的数据集，收集了 500 张自然场景图像，并针对每张图像配备了多组由不同标注人员手工绘制的边界标注。该数据集划分为训练集、验证集和测试集，为模型训练与评估提供了完备的数据支撑。其独特之处在于提供了边界定位精度、覆盖曲线下面积等多样化的评估指标，使它成为超像素分割算法、深度学习图像分割模型等众多图像分割技术进行性能评估与对比的重要基准。MSRC 数据集是一个广泛应用的基准数据集，它涵盖了 21 个不同类别，共计 591 张图像，每张图像都经过细致的彩色像素级标注。该数据集采用分层目录设计，搭配精细的标签体系，不仅能支持图像分类、目标检测等基础视觉任务，还能助力场景理解等复杂任务的研究。

3.2 BSD 数据集实验

为系统评估各类算法的性能表现，将文献所涉及的多种分割方法以及本文算法统一应用于 BSD 数据集

开展实验。BSD 数据集上的部分分割结果如图 4 所示，其中，第 1 行为原始图像，第 2 行~第 8 行分别呈现了 Mshift^[25]、MCG^[26]、CCBCut^[27]、ENCut^[28]、NAM^[29]、HISFBDI^[30]以及本文算法的图像分割结果。

从图 4 中呈现的分割可视化结果能够发现，本文算法 AMTISA 具备精准判断并确定目标图像类别数目的能力，且分割效果良好。以第 2 张具有多层次结构的山谷图为例，在多数对比算法中，难以精确选取该图合适的阈值数量，进而导致对山谷区域的分割不够精准。然而，本文算法不仅能较为准确地确定阈值个数，还能实现精确分割，这充分彰显本文算法在处理复杂数据集时具有更为突出的优势。各算法在 BSD 数据集上的实验结果如表 7 所示，表中加粗数字表示该指标的最优值。

从表 7 数据能够看出，本文方法在 SC 指标和 PRI 指标上均有提升，表明本文算法能够更精准地识别图像中的目标区域，有效区分不同类别，极大



(a) 118035分割结果

(b) 55067分割结果

(c) 24063分割结果

(d) 135069分割结果

图4 BSD数据集上的部分分割结果

地提高了分割的准确性；VI 指标相较于大部分对比算法，表现出不同程度的降低，意味着本文算法在处理具有复杂边界的物体时，能够保持分割边界的清晰和准确，减少了过度分割和欠分割的现象；在 ACT 指标方面，AMT 算法以 0.084 s 的用时遥遥领先，仅为 MCG 算法 43.620 s 的约五分之一、CCBCut 算法 34.731 s 的约四百分之一，再次验证了本文算法在保证分割精度的同时将计算效率提升至新的高度。由此可见，在 BSDS500 数据集这一特定环境下，面对图像中复杂的背景以及多样的目标形态，该方法都能维持相对稳定的分割效果，从而为后续的图像分析与处理工作奠定了更为坚实的可靠基础。

表 7 各算法在 BSD 数据集上的实验结果

算法	SC(↑)	PRI(↑)	VI(↓)	ACT/s(↓)
Mshift	0.540	0.790	1.850	0.895
MCG	0.592	0.813	1.548	43.620
CCBCut	0.418	0.734	2.360	34.731
ENCut	0.621	0.784	2.084	22.962
NAM	0.565	0.811	1.805	0.591
HISFBDI	0.669	0.854	1.371	0.186
本文算法	0.764	0.868	1.724	0.084

3.3 MSRC 数据集实验

为全面评估各算法的综合性能，本文在 MSRC 数据集上开展了系统实验，各算法在 MSRC 数据集上的实验结果如表 8 所示，表中加粗数字表示该指标的最优值。

表 8 各算法在 MSRC 数据集上的实验结果

算法	SC(↑)	PRI(↑)	VI(↓)	ACT/s(↓)
Mshift	0.750	0.830	1.800	1.473
MCG	0.668	0.784	1.205	52.432
CCBCut	0.661	0.674	0.989	48.873
ENCut	0.638	0.669	1.382	31.035
NAM	0.750	0.852	0.944	0.353
HISFBDI	0.737	0.826	1.010	0.275
本文算法	0.762	0.896	0.926	0.087

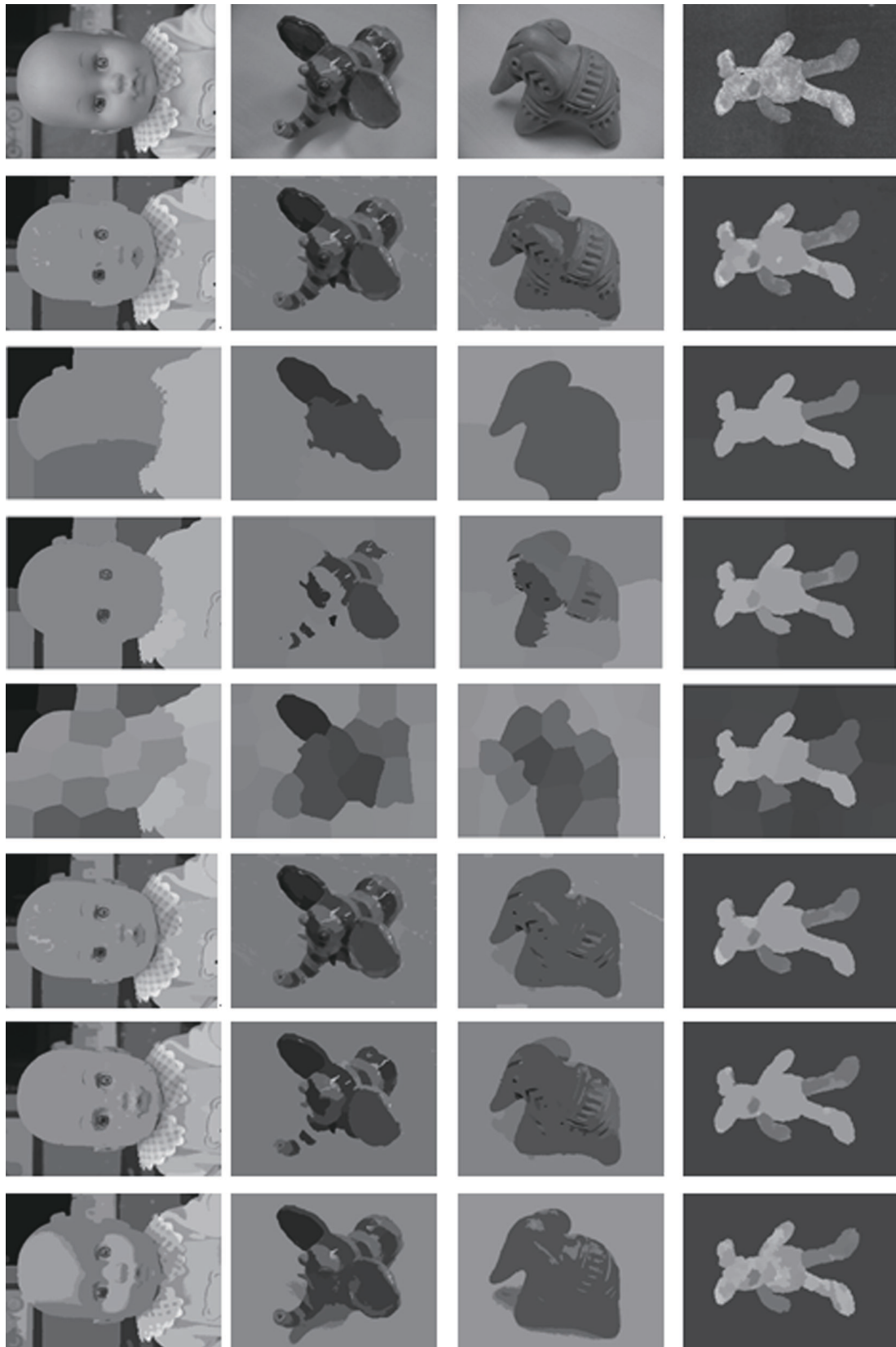
对 MSRC 数据集分割实验的各项性能指标数据进行分析可知，本文算法在 SC 指标和 PRI 指标上均呈现出提升态势，这表明算法在分割的准确性以及区域一致性方面有了显著改进；与此同时，算法还成功使 VI 指标降低，意味着分割结果的不确定性得到有效控制，进一步提升了分割的质量和稳定性；在该数据集上，本文算法的 ACT 指标仅为 0.087 s，依旧保持最低，明显优于 NAM 的 0.353 s、HISFBDI 的 0.275 s 以及 MCG、CCBCut 和 ENCUT 的数十秒级耗时，体现出本文算法在维持分割性能的同时具备更高的计算效率。

MSRC 数据集上的部分分割结果如图 5 所示，其中，第 1 行为原始图像，第 2 行~第 8 行依次为 Mshift、MCG、CCBCut、ENCut、NAM、HISFBDI 以及本文方法的分割结果。以第一张玩偶图像为例，由于该图像中待分割的玩偶细节丰富，其他几种对比方法难以确定阈值个数且无法实现精准分割。本文算法可精准捕捉玩偶的细微特征信息，进而实现玩偶细节的高精度分割，这充分彰显了本文算法在处理复杂纹理图像方面具备优秀的竞争力。

综上所述，在 BSDS500 数据集上，SC 指标升至 0.764，PRI 指标达 0.868；MSRC 数据集上，SC 指标提至 0.762，PRI 指标更是高达 0.896，这主要得益于谷底筛选策略的巧妙运用。该策略将原本 256 级的候选阈值进行大幅压缩，成功过滤掉了噪声产生的峰谷干扰，使后续动态规划在更纯净的搜索空间内锁定最优阈值；同时，MSRC 数据集上 VI 指标降至 0.926，得益于基于四边形不等式的分治剪枝法，此方法确保锁定全局最优解，避免局部极值，防止过分割与欠分割；此外，本文算法 ACT 表现极为出色，在 BSDS500 数据集上仅需 0.084 s。这得益于算法复杂度从 $O(KL^2)$ ，大幅降低至 $O(KL \ln(L))$ 。

4 结束语

针对阈值图像分割中算法效率与自适应能力难以协同优化，传统算法计算复杂度高且缺乏阈值数量自适应调整机制的问题，本文提出一种融合 DP 与分治策略的自适应多阈值分割算法。算法借助图像直方图的谷底特征构建目标函数，通过对 RGB 这 3 个通道的最佳阈值进行合并，得到全局最佳阈



(a) doll分割结果

(b) ceramic分割结果

(c) elephant分割结果

(d) teddy分割结果

图5 MSRC数据集上的部分分割结果

值, 不需要人工预设分割类别数量, 实现自适应分割。阈值搜索过程建模为多阶段决策问题, 运用 DP 算法将原本的指数复杂度降低至多项式级别。在 DP 算法执行期间, 每个阶段的最值搜索被转化为矩阵最值查找问题。为进一步提高效率, 引入分治策略对满足四边形不等式的最值矩阵进行分层搜索, 将时间复杂度进一步降低至 $O(KL\lg(L))$, 从而显著提升了阈值搜索的整体效率。

参考文献:

- [1] QU H Y, ZHAO W. Image segmentation edge detection technology based on optimization model[C]//Proceedings of the 2023 International Seminar on Computer Science and Engineering Technology (SCSET). Piscataway: IEEE Press, 2023: 131-134.
- [2] HORNG J Y, HO J H, LO C C, et al. Auto labeling for liver CT image based on edge detection and region of interest segmentation [C]//International Conference on Genetic and Evolutionary Computing. Berlin: Springer, 2024: 404-410.
- [3] ZHONG J X, LI Z H, CUI Y B, et al. 4D-CS: exploiting cluster prior for 4D spatio-temporal LiDAR semantic segmentation[J]. IEEE Robotics and Automation Letters, 2025, 10(1): 468-475.
- [4] XIAO T R, DENG J Q, WEN C F, et al. Parallel algorithm for multi-viewpoint viewshed analysis on the GPU grounded in target cluster segmentation[J]. International Journal of Digital Earth, 2024, 17(1): 2308707.
- [5] ZHANG G L, HE Y M. Application of region growing algorithm in segmentation of sublingual veins based on Halcon[C]// Proceedings of the International Conference on Computer Graphics, Artificial Intelligence, and Data Processing. Bellingham: SPIE Press, 2025, 13560: 878-881.
- [6] LI Y S, LI Z Z, GUO Z W, et al. Infrared small target detection based on adaptive region growing algorithm with iterative threshold analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-15.
- [7] HUI D. Research on segmentation and recognition algorithm of moving human image based on deep learning[C]//Proceedings of the 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT). Piscataway: IEEE Press, 2023: 1379-1382.
- [8] LI X H, YIN F, LIU C L. DocSAM: unified document image segmentation via query decomposition and heterogeneous mixed learning[C]// Proceedings of the 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2025: 15021-15032.
- [9] YAO Y D, LI X, CUI Y P, et al. Energy-efficient routing protocol based on multi-threshold segmentation in wireless sensors networks for precision agriculture[J]. IEEE Sensors Journal, 2022, 22(7): 6216-6231.
- [10] JIANG Y Z, YE H W C, LIN Z C, et al. A fast multi-threshold image segmentation method using a Bayesian forecasting evolutionary algorithm[C]//Proceedings of the 2024 International Conference on Sustainable Technology and Engineering (i-COSTE). Piscataway: IEEE Press, 2024: 1-6.
- [11] DENG X L, DENG J, DENG Z Y. Adaptive corner detection algorithm based on OTSU threshold segmentation[C]//Proceedings of the 2024 IEEE 9th International Conference on Computational Intelligence and Applications (ICCIA). Piscataway: IEEE Press, 2024: 120-124.
- [12] PARE S, KUMAR A, SINGH G K, et al. Image segmentation using multilevel thresholding: a research review[J]. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020, 44(1): 1-29.
- [13] JARDIM S, ANTÓNIO J, MORA C. Image thresholding approaches for medical image segmentation - short literature review[J]. Procedia Computer Science, 2023, 219: 1485-1492.
- [14] LUESSI M, EICHMANNI M, SCHUSTER G M, et al. Framework for efficient optimal multilevel image thresholding[J]. Journal of Electronic Imaging, 2009, 18(1): 013004.
- [15] MAHENDRAN N, DURAI RAJ VINCENT P M. Alzheimer's disease classification via pre-trained convolutional neural network variants utilizing multi-level thresholding based on kapur's entropy[C]//Proceedings of the 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON). Piscataway: IEEE Press, 2024: 1-6.
- [16] QIAO L, LIU K, XUE Y F, et al. A multi-level thresholding image segmentation method using hybrid arithmetic optimization and harris hawks optimizer algorithms[J]. Expert Systems with Applications, 2024, 241: 122316.
- [17] TURAJLIC E. Multilevel image thresholding based on OTSU's method and multi-swarm particle swarm optimization algorithm[C]// Proceedings of the 2024 47th MIPRO ICT and Electronics Convention (MIPRO). Piscataway: IEEE Press, 2024: 43-47.
- [18] YU C Z, WU W H, ZHENG J, et al. A multi-stage adaptive otsu thresholding algorithm for pore segmentation in rock thin-section images[J]. Earth Science Informatics, 2025, 18(2): 1-15.
- [19] MA G Y, YUE X F. An improved whale optimization algorithm based on multilevel threshold image segmentation using the OTSU method[J]. Engineering Applications of Artificial Intelligence, 2022, 113: 104960.
- [20] LIU G H, ZHANG Z F, CUI X, et al. Chromosome image segmentation based on OTSU and region growing algorithm[C]//Proceedings of the 2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI). Piscataway: IEEE Press, 2022: 1046-1050.
- [21] VIRTUSIO D T U, TAPAGANAO F J D, MARAMBA R G. Enhanced iris recognition system using daugman algorithm and multi-level OTSU's thresholding[C]//Proceedings of the 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET). Piscataway: IEEE Press, 2022: 1-6.
- [22] MAHENDRAN N, DURAI RAJ VINCENT P M, SAMIAYYA D, et al. Pre-trained convolutional neural networks and OTSU's multi-level thresholding based Alzheimer's classification[C]//Proceedings of the 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). Piscataway: IEEE Press, 2024: 1-7.
- [23] SHANNON C E. Communication in the presence of noise[J]. Proceedings of the IRE, 2006, 37(1): 10-21.
- [24] NYQUIST H. Certain topics in telegraph transmission theory[J]. Proceedings of the IEEE, 2002, 90(2): 280-305.
- [25] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [26] PONT-TUSET J, ARBELÁEZ P, T BARRON J, et al. Multiscale com-

binatorial grouping for image segmentation and object proposal generation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(1): 128-140.

- [27] CAHILL N D, HAYES T L, MEINHOLD R T, et al. Compassionately conservative balanced cuts for image segmentation[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 1683-1691.
- [28] ZHU L, KANG X J, YE L Z, et al. Explored normalized cut with random walk refining term for image segmentation[J]. IEEE Transactions on Image Processing, 2022, 31: 2893-2906.
- [29] ZHENG Y P, WEN D L, SAREM M. A novel NAM-based image segmentation using hierarchical density-based spatial clustering[J]. IET Image Processing, 2024, 18(5): 1245-1257.
- [30] ZHENG Y P, SHU S Q, HUANG J Z, et al. Hierarchical image segmentation based on bi-directional cascade network[C]//Proceedings of the 2025 7th International Conference on Software Engineering and Computer Science (CSECS). Piscataway: IEEE Press, 2025: 1-6.
- [31] ARBELÁEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [32] SHOTTON J, WINN J, ROTHER C, et al. TextonBoost: joint appearance, shape and context modeling for multi-class object recognition and segmentation[C]//Computer Vision - ECCV 2006. Berlin: Springer, 2006: 1-15.

[作者简介]



龙建武 (1984-), 男, 湖北恩施人, 博士, 重庆理工大学副教授、硕士生导师, 主要研究方向为图像处理与机器学习。



李继豪 (2001-), 男, 江西赣州人, 重庆理工大学硕士生, 主要研究方向为图像处理、机器学习、神经网络等。



曾谁飞 (1978-), 男, 博士, 青岛海尔电冰箱有限公司、海尔优家智能科技(北京)有限公司工程师, 北京邮电大学校外导师, 主要研究方向为人工智能、大模型、深度学习、神经网络、机器学习、多模态大模型、语音技术等。