

基于超像素表达的面向目标提取最优分割算法

王金旺

武汉大学电子信息学院, 湖北, 武汉, 430079

摘要: 本文研究了一种基于超像素表达的面向目标提取最优分割算法。首先使用 SLIC 超像素分割算法对图像进行超像素表达, 在超像素表达的基础上, 先后进行基于谱聚类和区域近邻图的超像素合并, 得到最终的最优分割结果。同时, 使用 Berkeley 图像数据库 BSDS500 和无人机航拍图像两类图像进行面向目标提取的最优分割算法实验。并进行了定性分析。实验结果验证了算法的有效性。

关键词: 无人机; 航拍图像; 超像素; 面向目标提取; 最优分割

中图法分类号: P208

文献标识码: A

无人机作为一种新的遥感平台和影像获取方法, 具有高分辨率、高灵活性、高效率和低成本的优势, 其应用范围越来越广。无人机航拍图像的分辨率可达厘米级, 图像中的相邻像素往往具有很强的相关性, 导致“同物异谱”, “异物同谱”现象大量发生, 同时图像数据量大幅度增加, 地物信息呈现高度细节化。因此图像解译的方式必须要从传统的面向逐个像元光谱信号处理的方式转变到以面向对象 - 超像素的空间、结构和几何信息处理为主的方式上来。本文主要围绕无人机航拍图像的超像素表达展开研究, 并在此基础上探索面向目标提取的最优图像分割方法。

超像素是指具有相似感官特征的相邻像素所组成的像素块。与基于单个像素的分割算法相比, 基于超像素的分割算法具有以下优势: 1. 超像素提取了图像的局部特征, 保留了对图像进行后续处理所需的有效信息; 2. 超像素的使用能大大减小图像中的冗余信息, 降低了图像处理算法的运行时间。

经过十余年的发展, 超像素技术已经被广泛地应用于各大计算机视觉和图像处理领域。例如, 目标跟踪、人体姿态估计、显著性分析等。

1 概述

面向目标提取最优分割算法指的是将我们所关注的前景目标从图像中提取出来, 便于后续对前景目标做进一步的处理。例如, 在遥感图像处理领域, 在图像预处理阶段, 使用面向目标提取的最优分割算法将房屋、树木和道路等目标提取出来, 在后续的处理中, 将大大减少对目标信息提取的难度。

本文主要研究基于超像素的面向目标提取最优分割算法。传统的最优分割算法大多是基于单个像素的方法, 这些方法在图像尺寸非常大时, 运算速度很慢, 无法满足目前图像大小越来越大的趋势, 而随着超像素分割算法运行效率的不断提高, 为了提高面向目标提取最优分割算法的效率, 有必要对基于超像素的面向目标提取最优分割算法进行研究。

2 基于像素的最优分割算法

本节主要介绍一种基于像素的面向目标提取最优分割算法 SRM(Statistical Region Merging),

即基于统计学的区域合并算法。

2.1 算法简介

基于统计学的区域合并算法 SRM 是一种像素级的区域合并算法，它的核心思想是将具有相同统计特性的像素进行合并，进而得到面向目标提取的图像分割结果。

2.2 算法原理

SRM 算法的核心可以概括为两条准则：1. 在 RGB 颜色空间的任何一个颜色通道中，同一统计学区域中的像素具有相同的统计期望值。2. 相邻统计学区域的统计期望值在 RGB 颜色空间中至少一个通道不同。

SRM 算法的执行步骤如下：

1. 在图像 I 中，定义集合 S_I 为 I 的四邻域像素集合。在本章节，定义符号 $|\cdot|$ 表示像素数量。

2. 假设像素 p 和像素 p' 为相邻的两个像素，在判断两个像素是否属于同一区域的，需要计算两像素统计学相似度，通过函数 f 来计算， f 的定义如下：

$$P_i = f_a(p, p') = |\overline{N_p(p')} - \overline{N_{p'}(p)}|$$

其中， $N_p(p')$ 为与像素 p' 的距离小于常数 2Δ 的区域。通过 f 计算好相似度之后，将相似度按升序排列。

3. 通过 $P(R, R')$ 判断区域 R 和区域 R' 是否应该合并， $P(R, R')$ 定义如下：

$$P(R, R') = \begin{cases} \text{true} & |\overline{R_a} - \overline{R'_a}| \leq \sqrt{b^2 R + b^2 R'} \\ \text{false} & \text{otherwise} \end{cases}$$

其中， a 表示 RGB 颜色空间中 R, G, B 三个颜色通道。 $b(R)$ 定义如下：

$$b(R, R') = g \sqrt{\frac{1}{2Q|R|} \ln (R_{|R|}/\delta)}$$

Q 为定义的常量， $\delta = 1/6|I|^2$ ， $g = 256$ 。

对经过排序的相邻像素对 (p, p') 如果 $R(p) = R(p')$ 即像素 p 和像素 p' 不在同一个区域，则计算 $P(R(p), R(p'))$ ，返回值为 *true* 时，合并 $R(p)$ 和 $R(p')$ 。

迭代上述过程，当图像中所有像素都经过上述运算之后，得到最终的合并结果。

3 基于超像素表达的面向目标提取最优分割方法

传统的最优分割方法一般是基于图像像素进行的，而基于超像素表达的最优分割方法是基于超像素的。在对原始图像进行超像素分割之后，在超像素块的基础上，进行超像素块的合并，得到最终最优分割结果。

本部分的超像素表达使用 SLIC 算法进行。

3.1 谱聚类方法简介

聚类就是根据样本之间的相似度，将它们分成不同组的过程。谱聚类的思想是将样本看作图节点 V ，样本间的相似度看作带权的边 E ，这样就能得到一个无向图 $G = (V, E)$ ，从而将聚类问题转化为图分割问题：即找到一种划分方法，使得不同组间边的权重尽可能的低（这意味着组间相似度尽可能低），组内边的权重尽可能高（这意味着组内相似度尽可能高）。

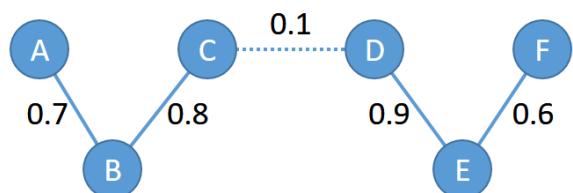


图 1 谱聚类原理图

Fig.1 The Principle Diagram of the Spectral Clustering

在图 1 中，总共有 6 个节点，节点间连线上的数字表示两个节点间的相似度，如果现在要将这图分成两类，根据谱聚类的思想，应该去掉的边是用虚线表示的那一条边，这样就形成了两类节点。

由于本部分只使用到了非标准谱聚类算法，所以这里只介绍非标准谱聚类算法的执行步骤，谱聚类算法的步骤如下：

1. 假设输入数据的聚类数量为 k ，首先，定义数据之间相似度的度量标准，根据数据构造一个图，图的每一个节点对应一个数据点，将图中相邻的节点连接起来，数据之间的相似度用节点间边的权重来表示。把这个图用相似度矩阵的形式表示出来，记为 W ，相似度矩阵一般的计算公式为

$$W_{ij} = \exp\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$$

其中， σ 为需要人工设定的参数，谱聚类算法的分类效果很大程度上取决于参数 σ 选取的合适与否；

2. 把相似度矩阵 W 中每一列的元素加起来得到 N 个数，把他们放到对角线上构成一个 $N \times N$ 的对角矩阵，记为 D ，计算非标准化图的拉普拉斯矩阵 $L = D - W$ ；

3. 求出 L 的前 k 个最小的特征值，以及特征值对应的特征向量 u_1, \dots, u_k ；

4. 将这 k 个特征向量组成一个 $N \times k$ 的矩阵 U ；

5. 将矩阵 U 的第 i 行作为行向量 $(y_i)_{i=1, \dots, n}$ ；
6. 对行向量 $(y_i)_{i=1, \dots, n}$ 使用 k -means 算法进行聚类，得到 k 个聚类 C_1, \dots, C_k [23]。

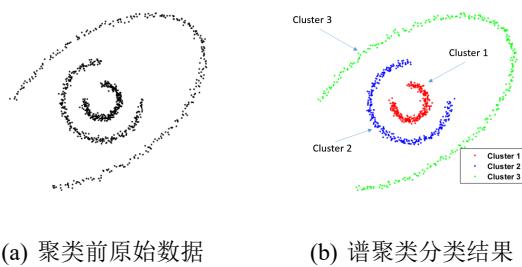


图 2 谱聚类分类效果示意图

Fig.2 Spectral Clustering Classification Effect Diagram

谱聚类算法根据使用的相似度矩阵的不同，可以分成多种算法，其本质上是一类算法。

谱聚类和 k -means 算法比起来有很多优点：

- (1) 谱聚类算法进行聚类的时候，只需要计

算出数据之间的相似度矩阵即可，而 k -means 则要求数据必须是 N 维向量；

- (2) 谱聚类算法对于一般的干扰误差数据不是很敏感；

- (3) 运算复杂度比 k -means 小，在处理图像这种维度比较高的数据时，运行速度比 k -means 快；

- (4) 能够避免得到局部最优解，从而获得全局最优解。

3.2 区域邻接图简介

区域邻接图 (Region Adjacency Graph, RAG) 定义为： $G = (V, E)$ ，其中 V_i 为节点， E_{ij} 为边，相邻两节点间生成边，一般均为无向有权图，边的权值为两相邻节点之间的相似性度量，RAG 表征区域间的空间相邻关系。

在图像分割中，首先使用预处理算法得到过分割的图像，其次，将过分割图像中的区域看作 RAG 中的节点，两相邻节点之间生成边，再根据相似度度量准则，生成过分割图像的 RAG。

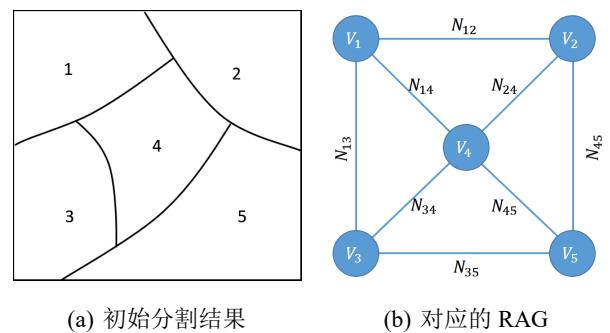


图 3 初始分割结果及其对应的 RAG

Fig.3 The Initial Segmentation Result and Its Corresponding RAG

使用 RAG 时，影响过分割图像合并结果的关键因素在于合并准则，主要有基于局部的合并准则和面向全局的合并准则两类。前者在局部范围内搜索满足合并条件的区域，并进行合并。后者从合并初始阶段开始，逐次合并全局最优的两个相邻区域，直到满足合并终止条件。面向全局的合并准则每次搜索全局最优的相邻区域进行合

并，能及时针对每次合并产生的变化作出调整，精度较高，在合并准则确定的情况下，对于同一幅图像的合并过程相同，分割过程稳定性较好。合并准则中具体涵盖的特征多包括面积、均值、方差等初级视觉特征，而基于人体知觉的语义特征等高级特征较少应用于区域合并。

图4为RAG算法运算结果的示意图，图像中，黑色的边为SLIC算法分割出来的超像素块的边界，黄色的点为RAG中的图节点，绿色的边为RAG中图节点之间的边，不存在RAG边的地方表示图节点相似度较低，RAG算法将图节点之间的边擦除掉了。原始图像中，包括背景的底色共有5种颜色，经过RAG算法的处理，可以明显看出，不同颜色的超像素块之间的RAG边消失，相同颜色的超像素块之间的RAG边依旧存在。

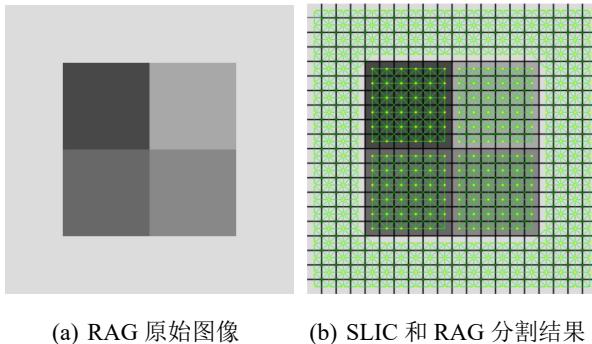


图4 RAG算法运算结果示意图

Fig.4 RAG Algorithm Operation Result Diagram

3.3 基于谱聚类和区域邻接图的最优分割算法

基于谱聚类和区域邻接图的最优分割算法主要包括三个过程：SLIC算法进行图像过分割、谱聚类算法初次合并和区域邻接图进一步合并。下面依次介绍这三个过程。

(1) SLIC算法进行图像过分割

使用SLIC算法对输入原始图像进行过分割，在SLIC算法中，需要设定的参数有两个：超像素数量和紧密度参数。在本算法中，紧密度参数取值为常数20，超像素数量需要根据图像大小、

景物大小、图像内容复杂度和算法应用场景进行设定，经过大量试验，对于Berkeley图像数据库(图像大小为 481×321)中的图像，兼顾算法运行时间，超像素数量设定为500即可。对于更大的图像，适当增加超像素数量能够提高算法分割质量。

(2) 谱聚类算法初步合并

在使用SLIC算法对原始图像进行过分割生成超像素之后，首先进行超像素块的谱聚类过程。

对每一个超像素块，我们提取其5个特征构造相似度矩阵：包括Lab颜色空间超像素块的 l, a, b 颜色分量的均值，以及其等效位置 x 坐标和 y 坐标。

假设SLIC算法生成了 N 个超像素块，则第 i 个超像素块和第 j 个超像素块间的距离将通过以下公式求得：

$$\begin{aligned} D_{i,j} &= \sqrt{A_{i,j} + B_{i,j}} \\ A_{i,j} &= (l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2 \\ B_{i,j} &= \frac{x_i - x_j}{SW}^2 + \frac{y_i - y_j}{SW}^2 \end{aligned}$$

其中， l, a, b 是超像素块在Lab颜色空间的颜色值， x, y 是超像素块的等效位置。

$$SW = \sqrt{(H \times W)/m}$$

H 和 W 为图像的长度和宽度， m 为一个常数，取值为100。

计算出距离矩阵 D 之后，对每一个超像素块，我们选择离它最近的 t 个超像素作为其邻居，则相似度矩阵 S 可以定义如下：

$$S_{i,j} = \begin{cases} \exp\left(\frac{-D_{i,j}^2}{2\sigma_i\sigma_j}\right), & \text{超像素 } i \text{ 和超像素 } j \text{ 相邻} \\ 0, & \text{其他} \end{cases}$$

其中，

$$\sigma_i = \frac{1}{t} \sum_{k \in T} D_{i,k}$$

超像素块 k 属于超像素块 i 的邻居超像素块的集合 T 。为了提高运行速度，将 t 取值为5。

得到相似度矩阵 $S_{i,j}$ 之后，设定谱聚类算法的聚类数量，使用谱聚类算法进行聚类过程，实现超像素块的初步合并。

聚类数量的设定会影响谱聚类算法合并结果，如果聚类数量设定的不合理，会造成原始图像的欠分割，经过反复试验，针对，Berkeley 图像数据库(图像大小为 481×321) 和真实无人机航拍图像(图像大小为 1000×1000) 而言，谱聚类数量设定为 30 时，谱聚类合并结果较为理想，当图像大小更大、内容较为复杂时，增加聚类数量能够得到更好的结果。

(3) 区域邻接图进一步合并

谱聚类算法对过分割图像的合并过程中，主要的合并依据为相似度矩阵 S 。在构造相似度矩阵 S 的过程中，超像素块之间的相似度主要通过公式进行计算，上式中，主要考虑了某一个超像素块在全局范围内与哪一个超像素块最相似，进而依据全局范围内的相似性进行聚类合并。RAG 则将区域合并的重点放在衡量某一个超像素块与周围超像素块的相似性上。也就是说谱聚类算法的合并过程是一种面向全局的区域合并过程，而 RAG 的合并过程则是一种面向局部的区域合并过程，两种算法配合使用，能够兼顾全局和局部超像素块的相似性度量和合并。

在对谱聚类算法合并结果进行 RAG 区域合并的过程中，主要的参数为区域相似度阈值。当两个区域的相似度小于阈值时，合并这两个区域，反之不进行合并。区域相似度阈值 $threshold$ 通过自适应的方式进行设定，以下对自适应方法进行介绍。

定义阈值 $threshold$ 为变量 T 。同时，定义变量 $MaxEdges$ ，其定义为在 Lab 颜色空间中，超像素块 k 与相邻超像素块集合 L 的最大颜色距离， $l \in L$ 时，颜色距离定义为：

$$E_{k,l} = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}$$

则 $MaxEdges$ 定义为：

$$MaxEdges = \max(E_{k,l}), l \in L$$

$MaxEdges$ 求出来之后，可以求得 T

$$T = mean(MaxEdges) - \frac{std(MaxEdges)}{2}$$

$mean(MaxEdges)$ 为 $MaxEdges$ 的均值， $std(MaxEdges)$ 为 $MaxEdges$ 的标准差。

为了避免阈值 T 过小影响合并过程，重新定义 T 如下：

$$T' = \begin{cases} T, & \text{if } T > 10 \\ 10, & \text{otherwise} \end{cases}$$

4 实验结果

4.1 实验数据介绍

本部分实验数据分为两部分，一部分为 Berkeley 图像数据库 BSDS500，另一部分为真实无人机航拍图像。使用 Berkeley 图像数据库主要是为了对比基于单像素和基于超像素两种最优分割算法的分割结果，并且在 Berkeley 图像数据库中自然景物图像中，目标物体一般只有一个，便于定性评估最优分割算法。使用无人机航拍图像主要是为了评估算法在复杂场景中的表现。



(a) BSDS12003 (b) BSDS118035 (c) BSDS8068

图 5 Berkeley 图像数据库 BSDS500 中的三张图像

Fig.5 Three of the Berkeley Image Database BSDS500 Images



(a) 航拍图像 1 (b) 航拍图像 2 (c) 航拍图像 3

图 6 三张无人机航拍图像

Fig.6 Three Unmanned Aerial Images

4.2 基于单像素的最优分割算法实验结果

本部分使用 Berkeley 图像数据库 BSDS500 进行实验。在 SRM 算法中，需要设定的参数为 Q ， Q 的取值不同，所得分割效果不同。以下实验结果为 9 种 Q 值的 SRM 实验结果，从左到右从上到下依次为 $Q = 256, 128, 64, 32, 16, 8, 4, 2, 1$ 。

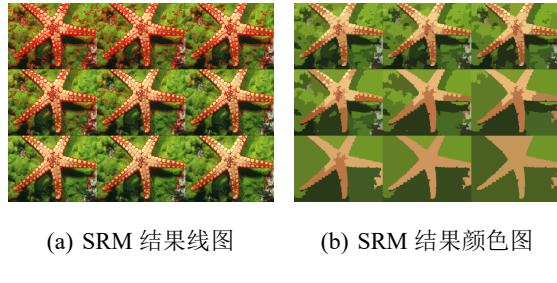


图 7 图像 BSDS12003 的 SRM 结果图
Fig.7 Results Image BSDS12003 SRM are Illustrated

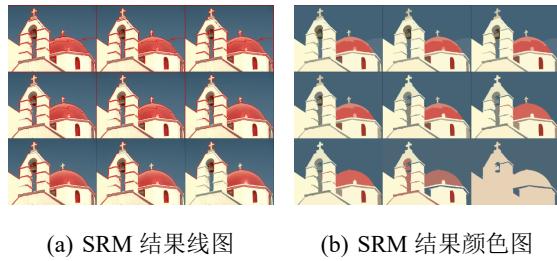


图 8 图像 BSDS118035 的 SRM 结果图
Fig.8 Results Image BSDS118035 SRM are Illustrated

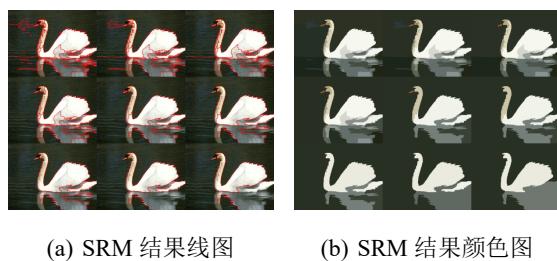


图 9 图像 BSDS8068 的 SRM 结果图
Fig.9 Results Image BSDS8068 SRM are Illustrated

4.3 基于超像素表达的最优分割算法实验结果

本部分实验使用基于超像素分割算法 SLIC 算法、聚类合并算法和区域邻接图 RAG 设计的面向目标提取的最优分割算法。

本部分实验的原始图像分为两部分：Berkeley 图像数据库 BSDS500 和真实无人机航拍图像。针对上述两部分图像，调整必要参数，得到最优分割结果。由于谱聚类算法和 RAG 在本算法中的本质都是进行区域合并，所以本部分实验结果也分成两部分进行展示：包括谱聚类聚类区域合并结果和在谱聚类基础上的 RAG 区域合并结果。

(1) BSDS500 数据库图像实验结果

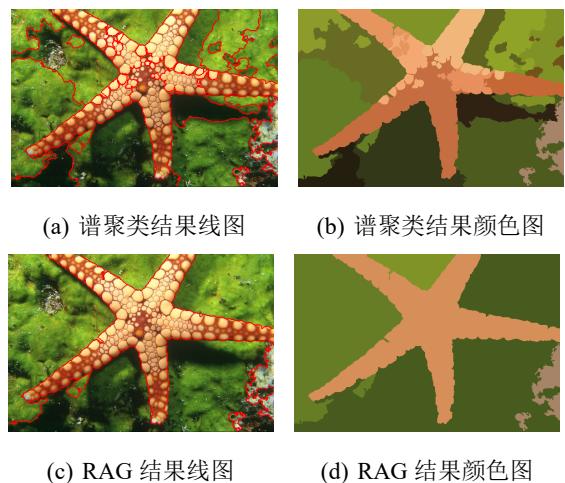


图 10 图像 BSDS12003 的最优分割结果
Fig.10 The image of BSDS12003 optimal segmentation result

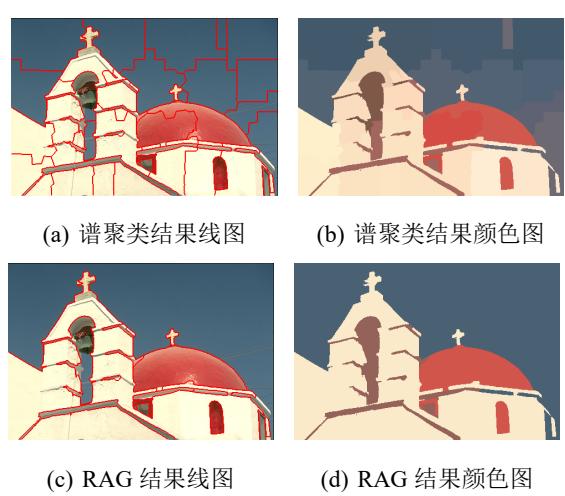


图 11 图像 BSDS118035 的最优分割结果
Fig.11 The image of BSDS118035 optimal segmentation result

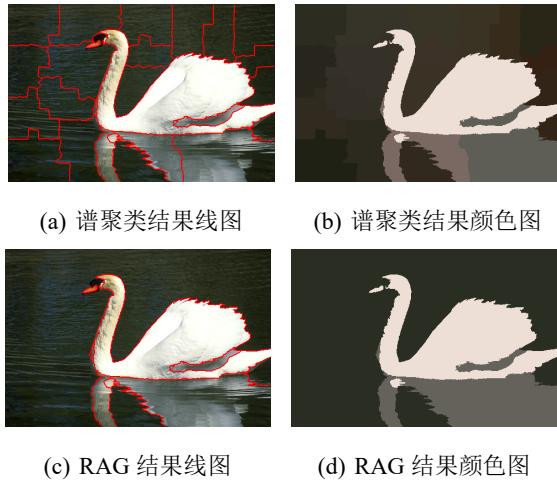


图 12 图像 BSDS8068 的最优分割结果

Fig.12 The image of BSDS8068 optimal segmentation result

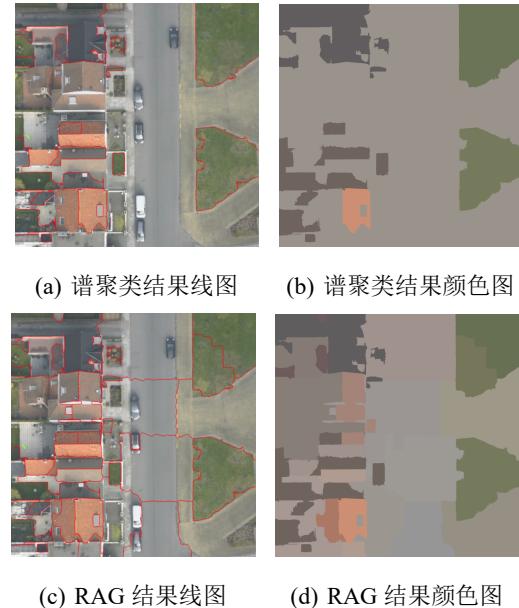


图 14 无人机航拍图像 2 的最优分割结果

Fig.14 2 the optimal segmentation result unmanned aerial images

(2) 无人机航拍图像实验结果

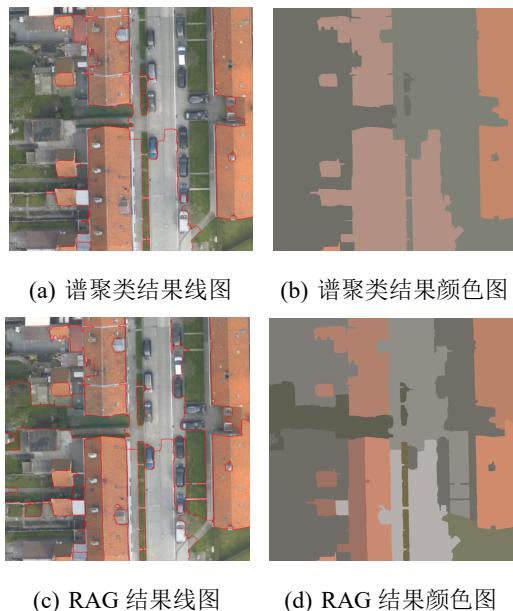


图 13 无人机航拍图像 1 的最优分割结果

Fig.13 1 the optimal segmentation result unmanned aerial images

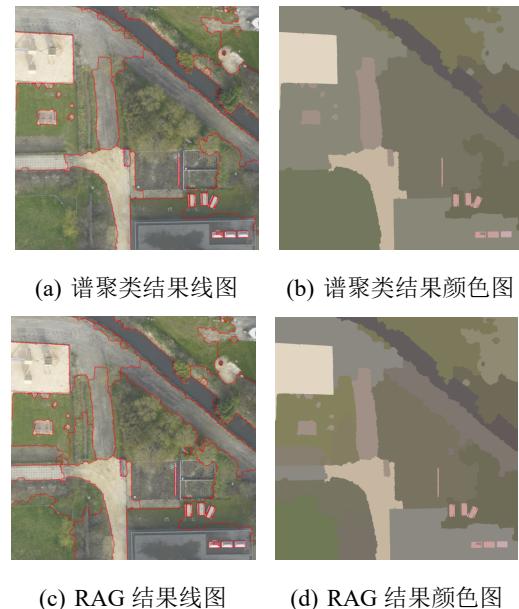


图 15 无人机航拍图像 3 的最优分割结果

Fig.15 3 the optimal segmentation result unmanned aerial images

4.4 实验结果分析

由图 7 和 8 可以清楚的看出， Q 取值不同，分割结果不同， Q 值越小，分割结果中聚类数量越少，当 Q 取值合理的时候，能将前景目标与

背景分离出来。

本部分实验使用两部分数据进行实验，一部分是 Berkeley 图像数据库，另一部分是真实的无人机航拍图像。由图 10、11 和 12 可以清楚的看出，经过谱聚类和 RAG 两次区域合并，能将大部分的超像素块正确进行合并，并将前景目标和背景进行分离。但是，同时由图 13、14 和 15 能够看出，基于超像素的最优分割算法在场景较为复杂的时候分割效果不够理想，许多超像素块被错误合并。

5 结语

在使用 SLIC 算法对图像进行超像素表达的基础上，将谱聚类算法和区域邻接图相结合，完成了基于超像素表达的面向目标提取最优分割算法的设计。并对本部分算法在 Berkeley 图像数据库 BSDS500 和真实无人机航拍图像上的表现进行了定性评估。通过实验，本设计在 BSDS500 这种前景目标较为简单的自然图像上表现较好，但是在真实无人机航拍图像上表现不是很好，有很大的提升空间。

参考文献

- [1] 王春瑶, 陈俊周, 李炜. 超像素分割算法研究综述 [J]. 计算机应用研究, 2014(01 vo v.31; No.267):

6-12.

- [2] 宋熙煜, 周利莉, 李中国等. 图像分割中的超像素方法研究综述 [J]. 中国图象图形学报, 2015(05 vo v.20; No.229): 599-608.
- [3] Nock R, Nielsen F. Statistical region merging[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(11): 1452–1458.
- [4] Ng A Y, Jordan M I, Weiss Y. On Spectral Clustering: Analysis and an algorithm[A]. ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS[C]. MIT Press, 2001: 849-856.
- [5] Luxburg U von. A tutorial on spectral clustering[J]. Statistics and Computing, 2007, 17(4): 395-416.
- [6] Zelnik-manor L, Perona P. Self-Tuning Spectral Clustering[A]. see: L.K. Saul, Y. Weiss, L. Bottou. Advances in Neural Information Processing Systems 17[M]. MIT Press, 2005: 1601-1608.
- [7] 林泽桢. 聚类分析中基于密度算法的研究与改进 [D]. 复旦大学, 2013.
- [8] Tremeau A, Colantoni P. Regions adjacency graph applied to color image segmentation[J]. IEEE Transactions on Image Processing, 2000, 9(4): 735-744.
- [9] 肖鹏峰, 冯学智. 高分辨率遥感图像分割与信息提取 [M]. 北京: 科学出版社, 2012.
- [10] 门学敏. 基于图论的超像素分割及其合并算法 [D]. 燕山大学, 2014.

An Optimal Image Segmentation Algorithm for Object Extraction on the Basis of Superpixel Algorithm

WANG Jinwang

School of Electronic and Information, Wuhan University, Wuhan 430079, China

Abstract: In the paper, we have studied an optimal image segmentation algorithm for object extraction on the basis of superpixel algorithm. Using SLIC to generate superpixel on the UAV aerial images, afterwares, on the basis of superpixels, using spectral clustering and RAG(Region Adjacency Graph) methods to merge superpixels. According to the results, we can get the optimal image segmentation. In addition, we use BSDS500, a Berkeley image database and UAV aerial images to conduct experiments over the optimal image segmentation algorithm for object extraction. According to the result, the algorithm turns out to be effective.

Key Words: UAV; aerial image; superpixel; segmentation