

# 基于多模态特征图融合的红外热图像目标区域提取算法

朱莉, 张晶, 傅应锴, 沈惠, 张守峰, 洪向共\*

(南昌大学信息工程学院, 江西 南昌 330031)

**摘要:** 红外热图像目标区域(Region of Interest, ROI)提取对故障检测、目标跟踪等有着重要意义。为解决红外热图像干扰多、需人工标记及准确率低等问题,提出一种基于多模态特征图融合的红外热图像 ROI 提取算法。通过对比度、熵及梯度特征构建多模态特征图并进行区域填充,实现 ROI 提取。将新算法应用于实际采集的光伏太阳能板图像中。结果表明,新算法具有平均查准率高(93.0553%)、平均查全率高(90.2841%)、 $F_1$  指数和 J 指数均优于图割法,人工标记少等优点,可有效用于红外热图像 ROI 提取。

**关键词:** 红外热图像; 对比度; 熵; 梯度

**中图分类号:** TP391.41 **文献标识码:** A

## Infrared thermal image ROI extraction algorithm based on fusion of multi-modal feature maps

ZHU Li, ZHANG Jing, FU Ying-Kai, SHEN Hui, ZHANG Shou-Feng, HONG Xiang-Gong\*

(Information Engineering School of Nanchang University, Nanchang 330031, China)

**Abstract:** Infrared thermal image region of interest (ROI) extraction has important significance for fault detection, target tracking and so on. In order to solve the problems of many infrared thermal image disturbances, artificial markers and low accuracy, a ROI of infrared thermal image extraction algorithm based on fusion of multi-modal feature map is proposed. Multi-modal feature maps are constructed by contrast, entropy, and gradient features, and region filling is performed to achieve ROI extraction. New algorithm is applied to actual collected photovoltaic solar panel image. Simulation results show that the proposed algorithm has high average precision (93.0553%), high average recall (90.2841%),  $F_1$  index and J index are better than Grab Cut, less artificial marks, etc.. It can be effectively used for ROI extraction of infrared thermal images.

**Key words:** infrared thermal image, contrast, entropy, gradient

**PACS:** 07.05.Pj

## 引言

红外热成像技术是以接受景物自身各部分辐射的红外线来进行探测,利用景物自身各部分辐射的差异获得图像的细节,实质是一种波长转换技术<sup>[1]</sup>。随着红外热成像技术的不断发展,基于红外

热图像的图像处理技术已经成为一项独立的、具有鲜明特色的研究方向。红外热图像的目标区域(Region of Interest, ROI)提取对故障检测、目标跟踪等有着重要意义<sup>[2-3]</sup>。然而,由于景物热平衡、传输距离、大气衰减及仪器成像系统参数等原因使得红外热图像存在清晰度低、边缘模糊、信噪比低以及干扰

收稿日期: 2018-04-03 修回日期: 2018-05-08

Received date: 2018-04-03 revised date: 2018-05-08

基金项目: 国家自然科学基金资助项目(61463035), 中国博士后科学基金资助项目(2016M592117), 江西省科技厅科学基金资助面上项目(20161BAB202045), 江西省博士后科研择优资助项目(2016KY01), 江西省科技厅杰出青年基金项目(2018ACB21038)

**Foundation items:** Supported by National Natural Science Foundation of China (61463035), China Postdoctoral Science Foundation (2016M592117), Science and Technology Department of Jiangxi Province Science Fund Project (20161BAB202045), Jiangxi Province Postdoctoral Research Excellent Foundation (2016KY01), Science and Technology Department of Jiangxi Province Outstanding Youth Fund Project (2018ACB21038)

**作者简介 (Biography):** 朱莉(1982-), 女, 江西南昌人, 副教授, 博士研究生, 研究领域为图像处理、信号处理和机器学习。E-mail: lizhu@ncu.edu.cn

\* 通讯作者 (Corresponding author): E-mail: lizhu@ncu.edu.cn

复杂等问题<sup>[4]</sup>. 在实际应用中, 干扰主要包括: (1) 导热构件干扰. 不同的构件会出现不同类型的干扰, 由于目标区域与干扰区域存在相似性, 导致难以准确提取目标区域. (2) 天气干扰. 阴天的温度和光线使得目标区域与环境温度差异较小, 造成定位不准. 因此, 提出新的红外热图像的提取算法, 以有效地分割出复杂红外热图像中的目标区域是十分有必要的.

红外热图像的区域提取可以分为基于数字图像处理的方法与基于统计学的方法. 目前, 研究者已经提出多种图像处理方法, Alamri 等提出阈值分割的方法<sup>[5]</sup>; Xia R 等提出纹路分析方法<sup>[6]</sup>; 李仲勋等提出基于微分方程的图像目标提取<sup>[7]</sup>; Zhao 提出红外背景杂波抑制滤波法<sup>[8]</sup>. 当区域部分的信噪比较低, 或目标区域与干扰区域灰度级差异较小时, 以上方法提取出的结果往往存在目标区域与干扰区域杂糅的情况. 近年来, 基于统计学的目标区域提取方法越来越受人们的关注, 如模糊 C 均值法 (Fuzzy C Mean, FCM) 聚类法<sup>[9]</sup>, 脉冲神经网络 (Pulse-Coupled Neural Network, PCNN) 法等<sup>[10]</sup>. 这些方法通常需要对图像数据进行预先的人工标记, 构建模型需要很多人工实验, 不断地对结果进行分析, 费时费力. 图割法 (Grab Cut) 是微软剑桥研究院提出的著名交互式图像语义分割方法, 是一种高效前景背景分割算法, 综合利用了纹理和边界两种信息, 已被广泛应用于红外热图像的分割技术中<sup>[11]</sup>. 但是该方法存在分割速度慢, 迭代效率低, 背景复杂时边缘分割不完整等缺陷; 且 Grab Cut 算法在迭代过程中, 针对不同的图像, 需要手动添加不同的标记.

本文提出一种基于多模态特征图融合的红外热图像 ROI 提取的算法. 该算法首先将图像的灰度差异、边缘变化、纹理结构等形成包括对比度、熵和梯度融合成多模态特征图; 再经过滤波和加权, 并进行区域填充; 最终实现目标区域提取. 将本文算法应用到实际采集两个月的光伏太阳能板红外热图像中. 研究表明, 本文算法提取的目标区域的查准率和查全率分别为 93.0553% 和 90.2841%, 比 Grab Cut 算法有显著提高, 且所需人工标记数据及人工干预少, 因此可有效地应用于红外热图像的 ROI 提取.

## 1 多模态特征融合的 ROI 提取算法

### 1.1 算法流程

针对红外热图像, 提出一种融合对比度、熵以及梯度多模态特征的 ROI 提取算法. 算法流程如图 1

所示. 本文提出的红外热图像 ROI 区域提取算法主要包括三个步骤.

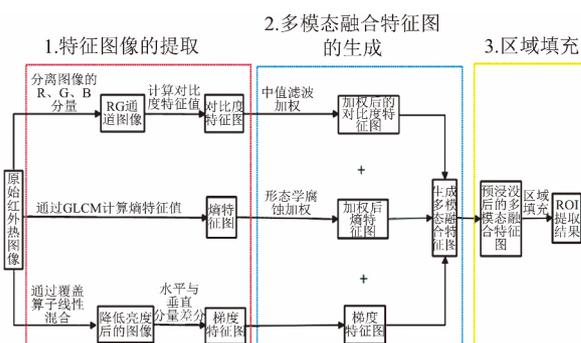


图 1 算法流程图

Fig. 1 The diagram of the proposed algorithm

第一步: 特征图像提取. 包括对比度特征、熵特征以及梯度特征. 通过灰度共生矩阵 (gray level co-occurrence matrix, GLCM) 计算、通道的对比度特征图像和原图像的熵特征图像<sup>[12]</sup>. 通过覆盖算子线性混合, 降低图像亮度, 用来减少数据图中亮度在构建梯度图像中的干扰, 并利用水平与垂直分量差分方法得到梯度特征图像.

第二步: 多模态融合特征图生成. 利用中值滤波获取加权对比度特征, 利用形态学腐蚀获得加权熵特征, 再与梯度特征进行融合, 生成多模态特征图像.

第三步: 对多模态特征图像进行预浸没与区域填充, 进而实现 ROI 提取.

### 1.2 特征图像提取

#### 1.2.1 对比度特征图像提取

在图像中, 对比度反映图像的清晰程度和纹理深浅的程度. 纹理越深, 对比度值越大, 反之越小. 本文中, 采用灰度共生矩阵的方法来求原图像的对比度特征图. 通道图像的红外热图像主要表征轮廓特征, 通道图像则表征高温和低温特征. 合并、分量的图像, 可以同时表征图像的轮廓与区域温度. 本文将和分量进行线性加权, 组合成新的分量, 再转换为灰度图像. 如式 (1) 所示:

$$I_1(x, y) = 0.299 \times R_{f_0} + 0.587 \times G_{f_0} \quad (1)$$

式 (1) 中  $I_1(x, y)$  表示灰度图像,  $R_{f_0}, G_{f_0}$  表示原始输入图像的  $R, G$  分量. 在  $M \times N$  大小的图像中, 设用  $S_{xy}$  表示中心在点  $(x, y)$  处, 上下距离为  $s$  的图像窗口.  $P(i, j)$  表示在滑动像素块区域内, 灰度级为  $i$  的点离开某个固定位置关系  $d$  到目的灰度级为  $j$  的概率. 在固定距离  $d$ , 固定方向  $\theta$ , 便可以得到这个点关

于滑动像素块, 区域内的对比度值  $con(x, y)$ . 如式(2)、(3)所示:

$$P(i, j) = \frac{g_{ij}(i, j \in S_{xy})}{L^2} \quad , \quad (2)$$

$$con(x, y) = \sum_i^L = \sum_j^L (i - j)^2 P(i, j) \quad , \quad (3)$$

式(2)、式(3)中  $i, j$  表示灰度等级,  $g_{ij}$  表示为空间中间隔的像素点个数,  $L$  表示灰度等级 ( $L=32$ ),  $con(x, y)$  表示对比度值. 求四个不同方向 ( $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ ) 的  $con(x, y)$ , 对其求均值  $\overline{con}(x, y)$ , 将均值作为关于点  $(x, y)$  在区域  $S_{xy}$  的灰度共生矩阵. 我们可以得到由一张图像中的像素点的对比度均值  $\overline{con}(x, y)$ , 所组成的矩阵  $c_1(x, y)$ . 选择一个阈值  $K_0$ , 对  $c_1(x, y)$  进行二值化得到对比度特征图为  $c_2(x, y)$ , 此过程提取出的对比度特征是描述红外热图像中目标区域的大致轮廓. 如式(4)、(5)、(6)所示:

$$\overline{con}(x, y) = \frac{1}{4} \sum_{\theta} con(x, y) \quad , \quad (4)$$

$$c_1(x, y) = \overline{con}(x, y) \quad , \quad (5)$$

$$c_2(x, y) = \begin{cases} 0, & \text{otherwise} \\ L_{\max}, & 0 < K_0 < \overline{con}(x, y) \end{cases} \quad , \quad (6)$$

式(4)、(5)、(6)中  $\overline{con}(x, y)$  表示在图像中不同点  $(x, y)$  对比度均值,  $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ ,  $c_1(x, y)$  表示对比度均值矩阵,  $c_2(x, y)$  为进行分层后得到的对比度特征图,  $L_{\max}$  为灰度等级最大值,  $K_0$  为阈值.

### 1.2.2 熵特征图像提取

采用灰度共生矩阵的方法求原图像的熵特征图. 像素点在  $S_{xy}$  区域内, 分别计算四个方向 ( $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ ) 上的熵特征  $ent(x, y)$ , 然后对其求平均值  $\overline{ent}(x, y)$ . 每个小块图像遍历完成后, 得到整张红外热图像的熵特征图像  $e_1(x, y)$ . 如式(7)、(8)、(9)所示:

$$ent(x, y) = - \sum_i^L \sum_j^L P(i, j) \log_c P(i, j) \quad , \quad (7)$$

$$\overline{ent}(x, y) = - \frac{1}{4} \sum_{\theta} ent(x, y) \quad , \quad (8)$$

$$e_1(x, y) = \overline{ent}(x, y) \quad , \quad (9)$$

式(7)、式(8)、式(9)中  $ent(x, y)$  为熵特征,  $\theta$  代表4个方向  $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ ,  $\overline{ent}(x, y)$  为熵特征均值. 选取一个阈值对熵特征图像二值化, 该阈值设为  $L_0$ . 在本文中取  $L_0 = -16$ .  $ent(x, y)$  分层后像素点

的灰度级别对应于0和  $L_{\max}$ . 为了便于后期的叠加并将纹理较少的板区提取出来, 将熵特征值进行取反操作得到熵特征图像. 如式(10)所示:

$$e_2(x, y) = \begin{cases} 0 & \text{otherwise} \\ L_{\max} & L_0 < \overline{ent}(x, y) < 0 \end{cases} \quad (10)$$

式(10)中  $e_2(x, y)$  为熵特征值进行取反操作后的熵特征图像,  $L_{\max}$  为灰度等级最大值. 熵反映了图像中纹理的非均匀程度或复杂程度, 是图像混乱程度的度量. 若图像没有任何纹理, 灰度共生矩阵几乎为0, 熵接近为0; 若图像有较多细小的且不均匀的纹理, 则熵的绝对值较大; 反之图像纹理较少或者纹理具有均一方向, 则图像熵的绝对值较小.

### 1.2.3 梯度特征图像提取

由于商用手持的红外热成像仪成像范围小, 存在不同数据图的灰度级分布不均匀, 容易出现灰度级分布处在两个极端分布的情况, 影响图像梯度特征的提取. 本文采用了通过覆盖算子减弱图像亮度的影响的方法. 覆盖算子如式(11):

$$f_2(x, y) = (1 - \mu)f_0(x, y) + \mu f_1(x, y) \quad , \quad (11)$$

式(11)中  $f_2(x, y)$  表示亮度降低的图像,  $(1 - \mu)$  为衰减因子,  $f_0(x, y)$  为原图像,  $f_1(x, y)$  表示输入图像. 当  $\mu$  从0变化到1时, 衰减因子减弱了原图像  $f_0(x, y)$  的影响,  $f_1(x, y)$  为进行交叉融合的输入图像,  $f_1(x, y)$  为全黑图像. 用画面交叉融合 (cross-dissolve) 的效果, 减小亮度因子的影响, 将图像的亮度降低, 降低亮度后的图像为  $f_2(x, y)$ . 将图像由 RGB 模式切换到 HSI 模式, 并求  $f_2(x, y)$  得 I 分量图像  $I_2(x, y)$ . 如式(12)、(13)所示:

$$I_2(x, y) = 0.299 \times R_{f_2} + 0.587 \times G_{f_2} + 0.114 \times B_{f_2} \quad , \quad (12)$$

$$g(x, y) = |I_2(x, y) - I_2(x + 1, y)| + |I_2(x, y) - I_2(x, y + 1)| \quad , \quad (13)$$

式(12)、(13)中  $I_2(x, y)$  表示  $f_2(x, y)$  中的一个像素点,  $R_{f_2}, G_{f_2}, B_{f_2}$  为 RGB 分量,  $g(x, y)$  表示梯度图像. 求取梯度的方法使用的是水平与垂直分量差分方法,  $I_2(x, y)$  通过水平垂直差分法后得到梯度图像.

### 1.3 多模态特征图像生成

对比度特征图像采用中值滤波, 选择核函数为矩阵进行中值滤波, 得到经过中值滤波的对比度特征图像. 如式(14)所示:

$$P_{idm}(x, y) = Med[c_2(x, y)] \quad , \quad (14)$$

式(14)中  $e_2(x, y)$  为对比度特征图像. 熵特征值特征图像则采用形态学腐蚀的方法, 使用椭圆形算子, 核函数的大小为  $3 \times 3$  的矩阵  $B$ . 腐蚀方法可以使区域连通, 边缘连接. 经过形态学方法得到的熵特征图像. 如式(15)所示:

$$e_3(x, y) = e_2(x, y) \ominus B \\ = \{z | (B)_z \subseteq e_2(x, y)\}, \quad (15)$$

式(15)中  $e_3(x, y)$  为经过形态学方法得到的熵特征图像. 对两幅特征图像进行预处理后, 进而对两幅图像标记权重进行分层. 如图2(a)所示. 分层的目的是需要模拟出与目标图像区域划分的形状的梯度图. 将整个多模态特征图像分出三块区域, 即: 纹理均一的目标区域, 环境区域, 环境与目标区域的边界.

$g_{\max}(x, y)$  表达了梯度图像整体灰度变换的强度, 即边缘明显程度, 如果图像降低亮度后的  $g_{\max}(x, y)$  是一个较大的值, 可以说明这幅图像的边缘比较明显. 如式(16)所示:

$$g_{\max}(x, y) = \max\{g(x, y)\}, \quad (16)$$

式(16)中  $g_{\max}(x, y)$  梯度图像整体灰度变换的强度. 以含有  $g_{\max}(x, y)$  值来加权两种特征图像:  $c_3(x, y)$ ,  $e_2(x, y)$ . 得到对应的加权系数  $\gamma$ , 参数为  $\alpha, \beta$  有:

$$\gamma = \alpha \times g_{\max}(x, y) + \beta, \quad (17)$$

式(17)中  $\gamma$  为图像对应的加权系数,  $\alpha, \beta$  为参数. 对比度特征图像完成了对梯度图像大轮廓增强, 将环境边界和板区树立屏障. 本实验使用的是参数为  $\alpha = -1$ ,  $\beta = L_{\max}$ , 加权值均衡了图像梯度不明显与梯度明显在求  $g_{\max}(x, y)$  时差异大的情况. 如式(18)、(19)所示:

$$\gamma_{idm} = L_{\max} - g_{\max}(x, y), \quad (18)$$

$$f_{idm}(x, y) = \begin{cases} \gamma_{idm} & e_3(x, y) = L_{\max} \\ 0 & \text{otherwise} \end{cases}. \quad (19)$$

式(18)、(19)中  $f_{idm}(x, y)$  为对比度特征图像加权图像. 取反后的熵特征图像进行标记的目的是将板区区域分层. 实验中选择的参数为  $\alpha = 0.6$ ,  $\beta = 0$ .  $\alpha = 0.6$  形成与以梯度为基础, 熵特征在处于多模态特征图像的第一层结构. 如式(20)、(21)所示:

$$\gamma_{ent} = 0.6 \times g_{\max}(x, y), \quad (20)$$

$$f_{ent}(x, y) = \begin{cases} \gamma_{ent} & e_3(x, y) = L_{\max} \\ 0 & \text{otherwise} \end{cases}. \quad (21)$$

式(20)、(21)中  $f_{ent}(x, y)$  为取反后的熵特征图像加权图像. 之后将经过灰度权重分配的两幅特征图像: 标记  $\gamma_{idm}$  的对比度特征图, 标记  $\gamma_{ent}$  的熵特征图与梯度图像  $g(x, y)$  进行叠加, 如图2(b)所示. 得

到多模态特征图像的表达式如式(22):

$$f(x, y) = g(x, y) + f_{idm}(x, y) + f_{ent}(x, y). \quad (22)$$

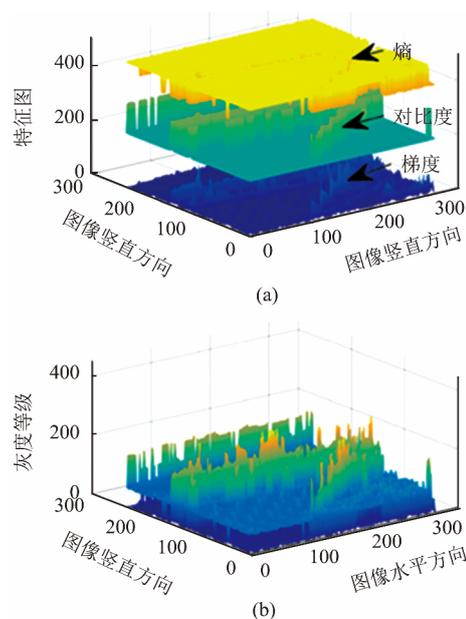


图2 特征图像和加权叠加后的多模态特征图  
Fig. 2 Feature images and multi-modal feature maps weighted by superposition

#### 1.4 预浸没与区域填充

对多模态特征图像进行预浸没和区域填充, 如图3所示. 对多模态图像进行种子点区域填充时, 填充区域受到熵特征图像的孔洞较多和梯度特征图像梯度突变的干扰, 使得区域填充只填充了小块的可连通区域. 进行随机选择种子点时, 这些不连续区域将会使得区域填充未能够在目标范围填充. 预填充的目的是减少那些不连续区域带来的干扰, 减少参数的多样性, 提高区域填充的准确度.

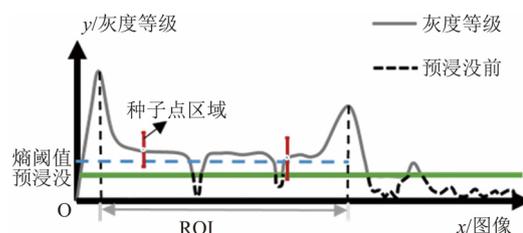


图3 多模态特征图像预浸没和区域填充过程  
Fig. 3 Pre-immersion and region filling process of multi-modal feature maps

对多模态特征图像进行较低范围的浸没, 方法是通过阈值分割实现反向填充, 此过程称为预浸没过程, 目的是在一定程度上减少负差值参数难定的问

题. 设这个取值为  $\tau$ , 定位在区间  $(0 < \tau < \gamma_{ent})$  的值中, 即不大于熵特征图像的填充值, 对多模态特征图像进行浸没的  $\tau$  值. 如式(23)所示:

$$f_{dst}(x, y) = \begin{cases} f(x, y) & \text{otherwise} \\ \tau & f(x, y) \leq \tau \end{cases}, \quad (23)$$

其中  $f_{dst}(x, y)$  为浸没后的图像. 区域填充是根据预先生长的生长准则将像素或子区域组合成为更大的区域的过程. 正负差值用来控制像素种子点填充的灰度级的范围, 用  $[D_{up}, D_{low}]$  来表示. 种子点和其连通区域像素点  $(x, y)$  灰度级的差在  $[D_{up}, D_{low}]$  范围内, 连通区域像素点标记为 1, 差不在此范围内像素点标记为 0. 将标记为 1 的点再添加入种子点的集合  $S$  中, 直至停止标记.

本文利用的是选取种子点进行填充, 种子点可以通过人工手动选择, 或者由对比度特征图通过形态学方法模拟出最大内接多边形的中心为种子点, 实验中选择的是交互式选择种子点的方式. 完成区域填充后, 由于正差值的范围有限, 内部仍将出现小的孔洞, 通过孔洞填充, 使得整个区域提取图像更加完整.

## 2 实验

### 2.1 数据采集

本算法使用的数据集来源为: 中国江苏林洋光伏科技有限公司 2 MW 并网光伏电站, 地图坐标位置为  $(121.639278, 31.817825)$ ; 使用的热成像仪器的型号为: 大立科技 DM63 非制冷焦平面在线式红外热像仪; 图像采集的时间为 2016 年 7 月至 8 月, 拍摄时刻分为早晨、正午及傍晚; 拍摄时天气分为晴朗、多云两种天气; 图片拍摄所用的热成像仪成像像素为像素.

### 2.2 实验过程及参数

提取特征图像, 其中包括梯度图像、对比度特征图以及熵特征图. 通过采用衰减因子为  $(1 - \mu) = 0.4$  求取梯度图像. 由于成像原因存在白色边缘, 故求  $g_{\max}(x, y)$  时将上下边缘减去 3 个像素. 数据集中图像大小均为  $320 \times 240$ , 计算概率时, 为了方便选取特征图像二值化时的阈值, 将  $com(x, y)$ 、 $ent(x, y)$  同时乘上灰度级的平方, 因此实验过程中, 使用  $L^2 \times ent(x, y)$ 、 $L^2 \times con(x, y)$  进行二值化 ( $L$  为用于分层的灰度级). 通过 GLCM 计算对比度特征值来获取原图像对比度特征图时, 设置参数分别为: 上下间距  $s = 5$ , 像素间距  $d = 4$ , 灰度等级  $L = 32$ , 对比度特征图分层阈值  $K_0 = 50$ ,  $L_{\max} = 255$ . 通过 GLCM 计算熵特

征值来获取熵特征图像时, 设置参数为: 下间距  $s = 5$ , 像素间距  $d = 4$ , 灰度等级  $L = 16$ , 熵特征图分层阈值  $L_0 = -16$ , 图像最大灰度级为  $L_{\max} = 255$ . 该过程初步得到关于原始图像的三种特征图像. 对以上得到的三种特征图像进行加权后叠加, 得到多模态特征图像. 该过程已在 1.3 节详细讲述. 对多模态特征图像进行区域填充, 完成对 ROI 区域填充. 该过程中种子点选择为交互式选择种子点, 交互式标记种子点不超过 3 个. 预浸没过程中, 选择  $\tau = 20$ ; 区域填充过程中, 负差值  $D_{low}$  取 20, 正差值  $D_{up}$  取 15.

与本文算法进行对比的算法是“Grab Cut 图像分割法”, 用于对比 Grab Cut 算法使用两种种子点: 一种为前景标记即光伏太阳能板区域, 另一种为背景标记即非光伏太阳能板区域, 进行迭代运算次数为 2 次.

本文所使用的编程以及编译环境为 Microsoft Visual Studio Community 2015. 通过控制台进行交互操作, 所依赖的主要图像算法库及版本为 OpenCV Library 2.4.10. 在部分算法环节, 我们使用了 MATLAB R2016b 对算法进行验证, 其中孔洞填充部分使用 MATLAB 实现.

### 2.3 算法性能分析

该算法和 Grab Cut 算法进行比较, 对这 17 张图进行区域提取测试, 每张图运行 10 次. 选用到评价指标有: 查准率 (Precision  $P$ ), 查全率 (Recall  $R$ ),  $F$  检验值  $F_\alpha$  (置信度  $\alpha = 1$ ) 和  $J$  指数 (Jaccard  $J$ ) 来衡量, 其中  $F$  检验值用于衡量查准率  $P$  与查全率  $R$  之间的关系,  $F$  检验值在置信度  $\alpha = 1$  的条件下, 提取算法结果的  $F_1$  检验值接近 1 为佳.  $J$  指数作为一个统计量, 用于统计算法提取出的区域与标准区域的相似程度.  $J$  值越接近 1, 说明分割结果越接近标准分割.

$$P = \frac{S_1}{S_2} \% \quad , \quad (24)$$

$$R = \frac{S_1}{S_3} \% \quad , \quad (25)$$

$$F_\alpha = \frac{(1 + \alpha^2) PR}{(\alpha^2 P + R)} \quad , \quad (26)$$

$$J(S_1, S_3) = \frac{|S_1 \cap S_3|}{|S_1 \cup S_3|} \quad , \quad (27)$$

其中, 集合  $S_1$  为算法得到的正确像素点 (包括边界点), 集合  $S_2$  算法提取区域的全部像素点, 集合  $S_3$  为人工提取区域的像素集合, 其中运算符表示该区域像素点的统计.

### 3 结果

将该算法与 Grab Cut 算法对图像区域提取的部分结果进行对比,结果如图 4 所示.第一栏图像为原始图像,第二栏为本算法提取到区域,第三栏为对比算法 Grab Cut 对应提取到的图像区域,第四栏为人工手动分割结果.其中黄绿色区域是提取出来的太阳能板区域.其中,图 4(a) 列和图 4(d) 列代表了红外热图像集中 ROI 区域与非 ROI 区域热差异较小,边缘模糊,ROI 区域纹理较少的情形,其成像天气可能为阴天或是傍晚;图 4(b) 列成像特点太阳能板区域与外界环境区域热差异大,纹理丰富;图 4(c) 列图像表现了当目标区域与环境杂糅的较为严重的图像,如图像中板区下部的管道明显,同时又不属于目标区域,这就容易造成后期对图像进行故障诊断的误判(非 ROI 区域故障的误判);图 4(e) 列图像纹理与边缘较为清晰.如图 4 所示,本文算法相比于 Grab Cut 算法,本文算法可以更准确地将红外热图像目标区域提取出来.综上所述,本文算法在区域提取的边缘准确性,查全率均优于 Grab Cut 算法.

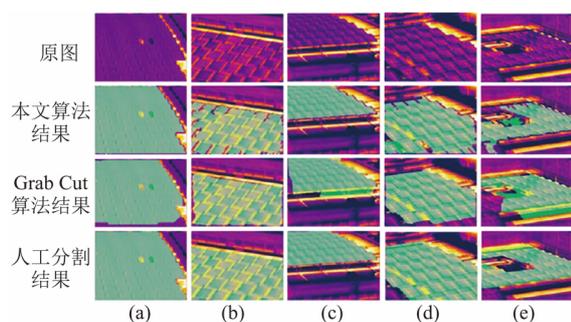


图 4 本文算法与 Grab Cut 算法区域提取结果对比图  
Fig.4 Comparison of extraction results between the algorithm in this paper and the Grab Cut algorithm

该算法区域提取过程图如图 5 所示.其中图 5(a) 到图 5(p) 分别表示为原图 7、通道、图 5(b) 灰度图、对比度特征分层图像、图 5(d) 中值滤波后图像、多模态图像的对比度特征图像、熵特征分层图像、图 5(g) 形态学腐蚀后的图像、为多模态图像的熵特征图像、图 5(a) 降低亮度后的图像、为图 5(j) 梯度特征图像、多模态特征图、图 5(l) 预浸没后的结果、区域填充后的结果、ROI 提取结果、图 5(a) 与图 5(o) 叠加后的结果即本文算法区域提取结果.

特征图像有以下三类情况: I、从图 5(d) 来看,在对比度特征图标记的边缘不连续,如图 5(d) 板区

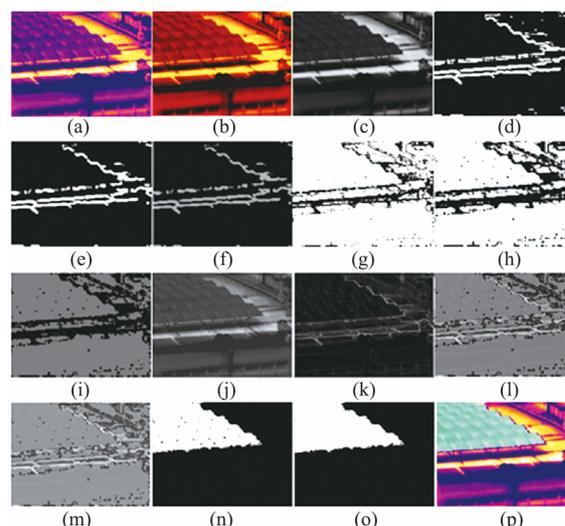


图 5 该算法区域提取过程图  
Fig.5 ROI extraction process of the proposed algorithm

下部分边缘不连通,利用熵特征标记的区域可以弥补填充区域的不连通特点. II、当存在熵特征图标记较为稀疏零散时,如图 5(g) 所示,结合对比度特征图中构建的边缘与下一步预浸没图像,互补及优化提取结果.两种特征值可以互补特征提取上的缺陷,均可以得到较为优秀的提取结果. III、两种特征图均存在缺陷:对比度特征图边缘不连续,熵特征图像存在较多闭合离散区域的情形,这将使得区域填充过程中存在边缘不完整.两种特征图时互补的存在,后期根据图像进行由梯度级的自适应的预填充,在其鲁棒性上有显著提升.进行加权后多模态特征图像如图 5(l) 所示.目标区域图通过对比度特征模拟的区域边界较连续,区域内部较为完整,灰度级方差较小;环境干扰区域图像较不连续,干扰区域的平均灰度级比目标区域低,灰度级方差较大.因此形成了目标区域、边界和干扰区域的三层灰度级的多模态特征图像,用于区域填充.

该算法与 Grab Cut 算法完成分割 17 张图像,每张图像比照人工分割标准结果,得到每张图像的平均查准率,平均查全率,  $F_1$  指数平均值,  $J$  指数平均值,如表 1 所示.

将该算法与 Grab Cut 算法对红外热图像 ROI 提取的比较,本文算法在查全率  $R$  和  $J$  指数上优于 Grab Cut. Grab Cut 算法在本数据集的测试具有有较高的查准确率  $P$ ,但是较低的查全率  $R$ ,说明 Grab Cut 算法存在本数据集普遍存在欠分割的情况.可见仅仅通过自适应阈值对于本数据集区域提取效果不佳,适用于荒漠地区红外热图像 ROI 区域提取.

因此本文主要比较的同样基于种子点的区域提取法 Grab Cut 的算法比较上. 使用 Otsu 算法对该数据进行测试, 得到的 17 张图像的  $\bar{P} = 89.41\%$ ,  $\bar{R} = 93.16\%$ ,  $\bar{J} = 0.6670$ ,  $\bar{F}_1 = 0.7773$ . 分水岭方法得到 17 张图像的  $\bar{P} = 73.23\%$ ,  $\bar{R} = 89.41\%$ ,  $\bar{J} = 0.9564$ ,  $\bar{F}_1 = 0.6906$ . 该算法分别与 Grab Cut、分水岭、Otsu 实验所得  $\bar{P}$ ,  $\bar{R}$ ,  $F_1$  与  $J$  进行单边成对检测, 结果如表 2 所示. 该算法分割结果最好、最稳定, 其性能指标  $R$ ,  $F_1$ ,  $J$  具有统计学意义. 此外, Grab Cut 算法需要标记非区域和目标区域, 且迭代运算两次, 人工标记量与复杂度都要比该算法大.

表 2 本文算法与对比算法性能指标的单边成对  $T$  检验

Table 2 One-sided  $T$  tests among the proposed algorithm

$p$	$P$	$R$	$F_1$	$J$
Grab Cut	0.2911 <sup>#</sup>	0.0283 <sup>*</sup>	0.0229 <sup>*</sup>	0.0238 <sup>*</sup>
分水岭	0.0006 <sup>***</sup>	0.8296 <sup>#</sup>	0.0023 <sup>**</sup>	0.0017 <sup>**</sup>
Otsu	0.0002 <sup>***</sup>	0.0168 <sup>*</sup>	0.0045 <sup>**</sup>	0.0028 <sup>**</sup>

其中, <sup>#</sup>:  $p > 0.5$ , <sup>\*</sup>:  $p > 0.05$ , <sup>\*\*</sup>:  $p > 0.01$ , <sup>\*\*\*</sup>:  $p > 0.001$ .

## 4 结论

针对环境的复杂性所导致对采集红外热图像目标区域误判问题, 提出一种基于多模态特征图融合的红外热图像区域提取算法. 本算法利用红外热图像特征和目标区域的特殊性来构建模型, 并进行目标区域提取. 首先, 该算法将图像的灰度差异、边缘变化、纹理结构等形成包括对比度、熵和梯度融合

成多模态特征图; 随之经过滤波和加权, 并进行预淹没和区域填充; 最终实现目标区域的提取. 研究结果表明本文算法获得的红外热图像目标区域的查准率和查全率对比于 Grab Cut 算法、分水岭和 Otsu 有显著提高, 并且人工标记少, 因此本文算法能应用到实际光伏太阳能板红外热图像的检测中. 下一步的工作需提高在图像信噪比低干扰成分复杂的情况下算法的准确度以及鲁棒性.

## References

- [1] WANG Rui-Feng, YANG Xian-Jiang, WU Wei-Dong. Development of infrared thermal imaging technology [J]. *Infrared and Laser Engineering*. (王瑞凤, 杨宪江, 吴伟东. 发展中的红外热成像技术. 红外与激光工程), 2008, 37 (增刊): 699-702.
- [2] Yasaswi V, Keerthi S, Jainab B S, et al. Infrared thermal image segmentation for fault detection in electrical circuits using watershed algorithm [J]. *International Journal of Engineering Trends and Technology*, 2015, 21(9): 423-429.
- [3] Kim T H, Song T L. Adaptive automatic thresholding in infrared image target tracking [J]. *Journal of Institute of Control*, 2011, 17(6): 579-586.
- [4] Guo F F, Yan G S, Li X D, et al. An improved canny infrared edge detection method based on otsu algorithm [J]. *Infrared*, 2010, 31(7): 24-27.
- [5] Alamri S S, Kalyankar N V, Khamitkar S D. Image segmentation by using threshold techniques [J]. *Computer Science*, 2010, 2(5): 83-86.
- [6] Xia R B, Zhao J B, Hui B, et al. A simple and efficient saliency extraction method based on multi-scale horizon-directional filter for infrared dim small target detection [J]. *Proceedings of SPIE-The International Society for Optical Engineering*, 2011, 8004(3): 510-514.

表 1 本文算法与 Grab Cut 算法性能评价指标比较

Table 1 Performance comparison between the proposed algorithm and Grab Cut

图像	平均查准率 $\bar{P}$	平均查全率 $\bar{R}$	$F_1$ 指数平均值 $F_1$	$J$ 指数平均值 $J$
	本文算法/Grab Cut	本文算法/Grab Cut	本文算法/Grab Cut	本文算法/ $J$ /Grab Cut
1	99.23% / 99.88%	96.99% / 63.20%	0.9809 / 0.7582	0.9626 / 0.6312
2	98.53% / 99.58%	98.30% / 83.67%	0.9841 / 0.9089	0.9688 / 0.8337
3	99.94% / 99.74%	88.17% / 91.38%	0.9365 / 0.9529	0.8812 / 0.9116
4	76.84% / 77.37%	89.89% / 89.72%	0.8276 / 0.8307	0.7068 / 0.7106
5	79.82% / 98.43%	73.21% / 87.92%	0.7601 / 0.9254	0.6189 / 0.8659
6	99.94% / 95.74%	79.66% / 98.46%	0.8859 / 0.9708	0.7962 / 0.9432
7	95.87% / 98.07%	93.16% / 75.46%	0.9449 / 0.8475	0.8956 / 0.7449
8	92.21% / 94.99%	91.82% / 88.70%	0.9198 / 0.9100	0.8515 / 0.8390
9	96.09% / 85.37%	94.93% / 87.67%	0.9539 / 0.8599	0.9130 / 0.7560
10	78.06% / 72.94%	77.55% / 82.83%	0.7754 / 0.7721	0.6343 / 0.6351
11	99.27% / 83.23%	95.40% / 90.54%	0.9729 / 0.8541	0.9473 / 0.7542
12	87.93% / 99.52%	99.65% / 93.04%	0.9342 / 0.9615	0.8766 / 0.9261
13	79.77% / 95.43%	76.79% / 67.13%	0.7816 / 0.7750	0.6431 / 0.6437
14	99.80% / 90.00%	96.24% / 68.06%	0.9799 / 0.7630	0.9606 / 0.6394
15	99.97% / 96.95%	93.16% / 72.03%	0.9645 / 0.8245	0.9313 / 0.7045
16	99.16% / 87.92%	95.10% / 79.84%	0.9709 / 0.8264	0.9434 / 0.7172
17	99.51% / 99.51%	94.81% / 94.81%	0.9711 / 0.9711	0.9438 / 0.9438
均值	93.06% / 91.69%	90.28% / 83.30%	0.9144 / 0.8605	0.8515 / 0.7686

- [7] LI Zhong-Bin, SHI Wen-Zhong. Partial differential equation-based object extraction from remote sensing imagery [J]. *Journal of Infrared & Millimeter Waves*. (李仲玠, 史文中. 基于偏微分方程的遥感图像目标提取. 红外与毫米波学报), 2016, **35**(3): 257-262.
- [8] Zhao Y, Pan H B, Du C P, et al. Bilateral two-dimensional least mean square filter for infrared small target detection [J]. *Infrared Physics & Technology*, 2014, **65**(5): 17-23.
- [9] Bai X Z, Chen Z G, Zhang Y, et al. Spatial information based FCM for infrared ship target segmentation [C]. *IEEE International Conference on Image Processing*, 2015, **46**(12): 5127-5131.
- [10] Qu S R, Yang H H. Automation S O. Infrared image segmentation based on PCNN with genetic algorithm parameter optimization [J]. *High Power Laser & Particle Beams*, 2015, **27**(5): 32-37.
- [11] Rother C, Kolmogorov V, Blake A. "Grab Cut": interactive foreground extraction using iterated graph cuts [C]. *ACM SIGGRAPH*, 2004, **23**(3): 309-314.
- [12] Shijin K P S, Dharun V S. Extraction of texture features using GLCM and shape features using connected regions [J]. *International Journal of Engineering & Technology*, 2016, **8**(6): 2926-2930.

~~~~~

(上接第 124 页)

- [6] Drumetz L, Veganzones M A, Henrot S, et al. Blind hyperspectral unmixing using an Extended Linear Mixing Model to address spectral variability [J]. *IEEE Trans. Image Process.*, 2016, **25**(8): 3890-3905.
- [7] Heylen R, Parente M, Gader P. A review of nonlinear hyperspectral unmixing methods [J]. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2014, **7**(6): 1844-1868.
- [8] YANG Bin, WANG Bin. Review of nonlinear unmixing for hyperspectral remote sensing imagery [J]. *J. Infrared Millim. Waves*. (杨斌, 王斌. 高光谱遥感图像非线性解混综述. 红外与毫米波学报) 2017, **36**(2): 173-185.
- [9] Hapke B W. Bidirectional reflectance spectroscopy. I. Theory [J]. *J. Geophys. Res.*, 1981, **86**: 3039-3054.
- [10] Halimi A, Altmann Y, Dobigeon N, et al. Nonlinear unmixing of hyperspectral images using a generalized bilinear model [J]. *IEEE Trans. Geosci. Remote Sens.*, 2011, **49**(11): 4153-4162.
- [11] Ammanouil R, Ferrari A, Richard C, et al. Nonlinear unmixing of hyperspectral data with vector-valued kernel functions [J]. *IEEE Trans. Image Process.*, 2017, **26**(1): 340-354.
- [12] Chen J, Richard C, Honeine P. Nonlinear unmixing of hyperspectral images based on multi-kernel learning [C]. *Hyperspectral Image and Signal Processing (WHISPERS)*, 2012 4th Workshop on. IEEE, 2012.
- [13] ZHU Fei, Honeine P. Biobjective Nonnegative Matrix Factorization: Linear Versus Kernel-Based Models [J]. *IEEE Trans. Geosci. Remote Sens.*, 2016, **54**(7): 4012-4022.
- [14] Smola A J, Schölkopf B. Learning with kernels [C]. Cambridge: MIT Press, 2008.
- [15] Camps-Valls G, Bruzzone L. Kernel-based methods for hyperspectral image classification [J]. *IEEE Trans. Geosci. Remote Sens.*, 2005, **43**(6): 1351-1362.
- [16] Evgeniou T, Micchelli C A, Pontil M. Learning multiple tasks with kernel methods [M]. *J. Mach. Learn. Res.*, 2005, **6**: 615-637.
- [17] TANG Yi, WAN Jian-Wei, XU Ke, et al. Hyperspectral unmixing based on material spatial distribution characteristics [J]. *J. Infrared Millim. Waves*. (汤毅, 万建伟, 许可等. 地物空间分布特性的高光谱遥感图像解混算法. 红外与毫米波学报) 2014, **33**(5): 560-570.
- [18] Lin C J. Projected gradient methods for nonnegative matrix factorization [J]. *Neural computation*, 2007, **19**(10): 2756-2779.
- [19] Huck A, Guillaume M, Blanc-Talon J. Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data [J]. *IEEE Trans. Geosci. Remote Sens.*, 2010, **48**(6): 2590-2602.
- [20] Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers [J]. *Foundations and Trends® in Machine Learning*, 2011, **3**(1): 1-422.
- [21] Alfonso M V, Dias J M B, Figueiredo M A T. Fast image recovery using variable splitting and constrained optimization [J]. *IEEE Trans. Image Process.*, 2010, **19**(9): 2345-2356.
- [22] Nascimento J M P, Dias J M B. Vertex component analysis: A fast algorithm to unmix hyperspectral data [J]. *IEEE Trans. Geosci. Remote Sens.*, 2005, **43**(4): 898-910.
- [23] Heinz D C, Chang C. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery [J]. *IEEE Trans. Geosci. Remote Sens.*, 2001, **39**(3): 529-545.
- [24] QIAN Yun-Tao, JIA Sen, ZHOU Jun, et al. Hyperspectral unmixing via  $L_{1/2}$  sparsity-constrained nonnegative matrix factorization [J]. *IEEE Trans. Geosci. Remote Sens.*, 2011, **49**(11): 4282-4297.
- [25] LI Xiao-Rui, CUI Jian-Tao, ZHAO Liao-Ying. Blind nonlinear hyperspectral unmixing based on constrained kernel nonnegative matrix factorization [J]. *Signal, Image and Video Processing*, 2014, **8**(8): 1555-1567.
- [26] Swayze G, Clark R, Sutley S, et al. Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada [J]. *Summaries of the Third Annual JPL Airborne Geosciences Workshop*, 1992, AVIRIS Workshop JPL Publication: 47-49.