

基于感兴趣区域的图像目标提取方法

高 晶¹, 吴育峰², 孙 伟¹, 陈仲华²

(1 第二炮兵工程大学, 陕西 西安 710025; 2 96630 部队, 北京 102206)

摘 要: 针对红外图像目标识别技术难的问题, 提出了一种基于感兴趣区域的目标提取方法. 首先利用二维最大熵方法检测红外图像中的感兴趣区域, 在此基础上将局部分类方法应用于感兴趣区域中每一个像素点的局部邻域, 以分析该像素的灰度特性; 然后运用像素点的斜面距离构建该像素点的空间位置特性, 并与灰度特性相结合; 最后定义新的阈值判断方法进行像素归属性判断, 以达到分离背景提取目标的效果. 经实验验证, 在红外目标提取技术上本文方法较经典的 otsu 方法和最大熵方法相比具有较高的准确性及稳定性, 对于红外图像对比度低、识别难的问题具有一定的应用价值.

关键词: 目标提取; 感兴趣区域; 二维最大熵; 局部分类; 斜面距离

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2013)12-0035-06

Image Target Extraction Method Based on ROI

GAO Jing¹, WU Yu-feng², SUN Wei¹, CHEN Zhong-hua²

(1 Second Artillery Engineering University, Xi'an 710025, China; 2 96630 Army, Beijing 102206, China)

Abstract: A target extraction method based on region of interest is presented to solve the difficult problem of infrared image target identification technology. Firstly, the regions of interest are detected by using two-dimensional maximum entropy method, on the basis of this, the local classification method is applied to the local neighborhood of each pixel in the region of interest, in order to analyse the pixel gray characteristics; Secondly, the spatial position characteristics of every pixel point is builded by calculation of chamfer distance, and is combined with the gray characteristic; Finally, the new determination method of threshold value is defined to determine the ownership of every pixel point, to achieve the separation of background and extraction of target. The results of experiments indicate that the proposed method has a higher accuracy and stability comparing with the classic otsu method and maximum entropy method in infrared target extraction, it has a certain value for problems of low contrast and difficult recognition of infrared image.

Key words: object extraction; region of interest; two-dimensional maximum entropy; local classification; chamfer distance

1 引言

红外图像目标识别技术是一个经典的难题, 目标和背景的界限不明显甚至模糊, 且存在很多噪声点, 这样使得完整提取目标非常困难. 图像的主要信息一般集中在几个少数的关键区域内, 这些区域就称之为感兴趣区域(region of interest, ROI). Prewitt 众数法和 OTSU 方法等传统阈值方

法都要求目标大小占图像面积大小的 30% 以上^[1], 对于面积相对较小的目标, 算法的性能将迅速下降, 利用熵来描述图像的特征, 使得图像特征的维数大大降低^[2].

但由于图像熵的计算仅考虑了图像中像素点的数目, 没有考虑各个像素点在灰度空间上的分布情况, 且计算量较大. 因此, 本文提出一种基于感兴趣区域的目标提取方法.

2 红外图像感兴趣区域提取

2.1 信息熵

信息熵是描述图像中所有信息量的有效途径,图像熵则反映了该图像中像素的分布情况,可用来度量图像的不确定性.假设图像 A 的灰度值为 $\{a_1, a_2, \dots, a_k\}$, 灰度值出现的概率分布为 $\{p(a_1), p(a_2), \dots, p(a_k)\}$, 则可得图像 A 的熵为

$$H(A) = - \sum_{k=1}^k p(a_k) \log p(a_k) \quad (1)$$

针对红外图像,不同亮度的像素点在空间中占据不同的灰度区间,使得图像表现出不同的形状,而不同形状的图像所包含的信息量则不同. Abutela 将 Kapur 等人提出的一维最大熵法推广至二维,基于二维最大熵原理的图像分割首先根据图像中的各个像素的点灰度值及其相邻区域灰度均值生成二维直方图,并以此为依据选取最佳阈值,二维灰度直方图表示如图 1 所示.

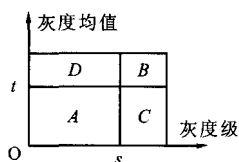


图 1 二维直方图

图 1 二维直方图中任意元素 $r_{i,j}$ 表示灰度级为 i 、邻域平均灰度值为 j 的像素出现的次数. 因为红外图像目标或背景内部像素灰度变化缓慢,灰度和灰度均值差距不大,在二维直方图中主要占据对角线位置,所以对于红外图像而言, A 象限代表了背景, B 象限代表了目标, C 象限和 D 象限主要是目标边缘附近像素点和噪声点^[3].

图像的全局二维熵定义为

$$H(s, t) = \ln[P_A(1 - P_A)] + \frac{H_A}{P_A} + \frac{H_L - H_A}{1 - P_A} \quad (2)$$

式中, $H_A = - \sum_{i=1}^s \sum_{j=1}^t (\frac{P_{ij}}{P_A}) \ln(\frac{P_{ij}}{P_A})$,

$$P_A = \sum_{i=1}^s \sum_{j=1}^t P_{ij}, H_L = - \sum_{i=1}^L \sum_{j=1}^L P_{ij} \ln P_{ij}.$$

在上式中通过求 $H(s, t)$ 最大值,进而找到最佳的 (s, t) , 得到最佳门限值. 尽管二维最大熵法相对于一维最大熵方法抗噪性及鲁棒性好,但从二维直方图中可以看出,求每一个 (s, t) 时,所有计算都需从头开始,难以满足实时性要求,因此需加快二维最大熵图像分割速度.

2.2 二维最大熵阈值递推搜索

对二维最大熵进行优化是可采用递推优化的方法^[3],即求 $H(s, t+1)$ 时,先求 $P_A(s, t+1)$ 和 $H_A(s, t+1)$ 的值,其递推公式为

$$P_A(s+1, t) = P_A(s, t) + \sum_{j=1}^t P_{s+1, j} \quad (3)$$

$$P_A(s, t+1) = \sum_{i=1}^s \sum_{j=1}^{t+1} P_{i, j} = P_A(s, t) + P_A(s-1, t+1) - P_A(s-1, t) + P_{s, t+1} \quad (4)$$

$$H_A(s+1, t) = H_A(s, t) - \sum_{j=1}^t P_{s+1, j} \lg P_{s+1, j} \quad (5)$$

$$H_A(s, t+1) = - \sum_{i=1}^s \sum_{j=1}^{t+1} P_{i, j} \lg P_{i, j} = H_A(s, t) + H_A(s-1, t+1) - H_A(s-1, t) - P_{s-1, t} \lg P_{s-1, t} \quad (6)$$

由以上递推优化,可将原先二维最大熵图像阈值分割算法复杂度从 $O(L^4)$ 降到 $O(L^2)$,减少了对数运算及循环次数,提高了运行效率.

3 目标与背景分离

3.1 感兴趣区域提取

在提取图像的感兴趣区域前,可利用中值滤波及剔除过大过小区域法对图像进行预处理,在不损失目标信息的同时增强目标像素之间的连通性,以减少噪声的影响,由于感兴趣区域的提取是面向于目标的,所以不能直接将已知的目标形状模板大小作为区域的大小,如果分割误差较大的话,目标识别的结果也将受到严重的影响. 因此,本文将目标形状模板大小作为先验知识,运用区域连通原则将图像区域扩展为包含目标的外接矩形,作为感兴趣区域,如图 2 所示.

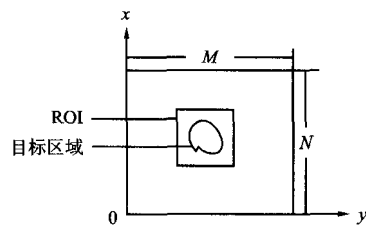


图 2 ROI 示意图

图 3 所示为红外图像的感兴趣区域检测结果图. 由图 3 可看出,在二维熵阈值处理的基础上进行感兴趣区域提取,明显缩小了目标搜索的范围,且有利于后续算法中目标与背景的准确分离.



图3 感兴趣区域

3.2 目标与背景分离

在提取出的感兴趣区域中,有些区域含有目标,有些区域则相反,完全属于背景区域.一般情况下,目标的整体灰度像素高于背景,但是不能排除部分目标区域的像素低于背景,从点像素进行区分会导致大量的误判断,所以在目标与背景的分离中,本文从感兴趣区域内像素点的局部邻域入手分析其属性.

红外图像在时域上由于场景及传感器自身热分布的不均匀性,导致背景图像是一个缓慢变化且非平稳的二维随机过程.如果目标较小,则可以用二维高斯函数近似^[4],而背景在灰度分布空间上具有较大的相关度,主要占据了图像空间频域的低频空间,同时由于背景图像各个同质区边缘的存在,背景图像中还包括部分高频分量.因此,在空域上可以用一个高斯白噪声来对背景像素进行建模.

Haris^[5]提出的基于统计学的图像局部区域平滑方法可有效提取边缘轮廓,对于提取的每一个感兴趣区域,关键在于将目标检测出来并提取,已达到目标与背景的分离.设感兴趣区域为 R ,大小为 $M \times N$,其中区域内一像素点 P 的邻域大小为 $n \times n$,邻域内所有像素为同质或者不同质组合两种可能性,判断像素点邻域范围的步骤如下:

(1) 假设像素点 P 的邻域 $N_{n \times n}(p)$ 为同质的情况下,高斯随机变量中均值为 μ ,方差为 σ^2 ;不同质的情况认为存在两种高斯随机变量的组合,均值分别为 μ_1 和 μ_2 ,方差 $\sigma_1^2 = \sigma_2^2 = \sigma^2$,则两种高斯随机变量组之和的概率密度函数为

$$p(x) = \frac{p_1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2\sigma^2}} + \frac{p_2}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2\sigma^2}} \quad (7)$$

式中, p_1 和 p_2 为概率.

(2) 设定邻域 $N_{n \times n}(p)$ 内的方差为 s^2 , $s^2 = \frac{1}{N}$

$\sum_{q \in N_{n \times n}(p)} \left[X(q) - \frac{1}{N} \sum_{r \in N_{n \times n}(p)} X(r) \right]^2$,其中 q 为邻域内任意像素, $X(r)$ 为 r 像素点的灰度值, N 为邻域内像素数目,运用极大似然检测法判断邻域的性质可得

当 $s^2 \leq (1+\alpha)\sigma^2$,邻域 $N_{n \times n}(p)$ 为同类;

当 $s^2 > (1+\alpha)\sigma^2$,邻域 $N_{n \times n}(p)$ 为不同类.

其中 α 为概率误差参数.

(3) 当 $N_{n \times n}(p)$ 邻域为不同类时,则需判断像素点 P 的归属情况,根据统计学的计算可得:

$$\begin{cases} p_1 + p_2 = 1 \\ p_1 \mu_1 + p_2 \mu_2 = c_1 = \frac{1}{N} \sum_{r \in N_{n \times n}(p)} X(r) \\ p_1 \mu_1^2 + p_2 \mu_2^2 = c_2 = \frac{1}{N} \sum_{r \in N_{n \times n}(p)} X^2(r) - \sigma^2 \\ p_1 \mu_1^3 + p_2 \mu_2^3 = c_3 = \frac{1}{N} \sum_{r \in N_{n \times n}(p)} X^3(r) - 3\sigma^3 c_1 \end{cases} \quad (8)$$

由上式计算可得

$$\mu_1 + \mu_2 = \frac{c_3 - c_1 c_2}{c_2 - c_1^2}, \mu_1 \mu_2 = \frac{c_1 c_3 - c_2^2}{c_2 - c_1^2}.$$

令: $\mu_1 + \mu_2 = \beta, \mu_1 \mu_2 = \gamma, \mu_1 > \mu_2$.

可计算出

$$\hat{\mu}_1 = \frac{\beta + \sqrt{\beta^2 - 4\gamma}}{2}, \hat{\mu}_2 = \frac{\beta - \sqrt{\beta^2 - 4\gamma}}{2}.$$

进而估算出概率 p_1 和 p_2 的值为

$$\hat{p}_1 = \frac{\hat{\mu}_2 - c_1}{\hat{\mu}_2 - \hat{\mu}_1}, \hat{p}_2 = \frac{c_1 - \hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1}.$$

由以上所有计算可得概率密度函数中各个参数的求解结果,对于像素点灰度特性的计算判断还需首先运用阈值法,根据贝叶斯方法可以用于获取判断阈值计算,如下式所示:

$$T = \frac{\hat{\mu}_1 + \hat{\mu}_2}{2} + \frac{\sigma^2}{\hat{\mu}_2 - \hat{\mu}_1} \ln \frac{p_1}{p_2} \quad (9)$$

(4) 由于(2)部分计算了邻域的方差 s^2 ,结合阈值 T ,本文从整个邻域入手,分析像素点 P 的特性,将其特性值 $A(p)$ 规定在 $[0, 1]$ 区间内,以确保其归一化性质.如下所示:

$A(p) =$

$$\begin{cases} \mu/L & s^2 \leq (1+\alpha)\sigma^2 \\ (1+k) \frac{\hat{\mu}_2}{\hat{\mu}_1 + \hat{\mu}_2} & s^2 > (1+\alpha)\sigma^2, X(p) \geq T \\ (1-k) \frac{\hat{\mu}_1}{\hat{\mu}_1 + \hat{\mu}_2} & s^2 > (1+\alpha)\sigma^2, X(p) < T \end{cases} \quad (10)$$

式中, k 为常数,控制像素点特性值 $A(p)$ 的取值范围,其取值如图4所示.

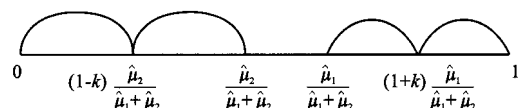


图4 K值分布图

由图 4 可以看出, k 值主要由 $\frac{\hat{\mu}_1}{\mu_1 + \mu_2}$ 和 $\frac{\hat{\mu}_2}{\mu_1 + \mu_2}$ 的值决定, 具体的取值范围主要集中在 $(0, \frac{\hat{\mu}_2}{\mu_1 + \mu_2})$ 和 $(\frac{\hat{\mu}_1}{\mu_1 + \mu_2}, 1)$ 之间. 由此可得出, 当 $\hat{\mu}_1$ 和 $\hat{\mu}_2$ 两值接近时, 不好判断像素点 P 属于目标还是背景, k 值则可以控制 $A(p)$ 取值, 已明确像素点的分类; 同理, 当 $\hat{\mu}_1$ 值很大或者 $\hat{\mu}_2$ 值很小时, k 值都将起着非常好的调节作用.

(5) 本文在分离目标与背景时, 为减少目标提取的误差, 除了应用图像像素点的灰度因素外, 还需考虑像素点空间位置信息, 以往的文献中常常运用欧氏距离进行空间位置的衡量, 但斜面距离相对欧氏距离而言, 更加符合人眼视觉观察. 斜面距离 (Chamfer Distance)^[6] 是 BARROW H. G. 在 1977 年提出来的, 用来描述两个轮廓之间的相似度, 该距离对噪声、旋转、尺度有较好的鲁棒性, 其基本公式如下:

$$d_{\text{cham}}^{(TE)}(x) = \frac{1}{|T|} \sum_{x_t \in T, x_e \in E} \min \| (x_t + x) - x_e \|_2 \quad (11)$$

式中, T 为目标模板, E 为图像检测的边缘, x 为相对位置, $|T|$ 表示模板 T 中边缘的数目, 也可计算为边缘响应点的个数, $\|x\|_2$ 为向量 x 的长度. 由此可知, 斜面距离计算的是模板 T 中的像素点到 E 中的像素点的平均距离. 但如果按照定义图像 E 的距离变化来计算, 时间复杂度将十分高. 因此, 距离变化时, 从边缘点出发, 沿远离边界点方向按所谓的 3-4 法则逐渐增大像素值; 也就是对任意像素点, 其周围八邻域中水平和垂直方向的四个邻域像素的灰度值每次增加 3, 而剩余四个邻域中的像素点则每次增加 4^[7].

本文假设 $C_{i,j}$ 为区域内任一 P 点 (i,j) 的斜面距离, 则 $C_{i,j}$ 的计算步骤如下所示:

1. Initialize: $C_{i,j} = \begin{cases} 0 & (i,j) \in \text{ROI} \\ \infty & \text{otherwise} \end{cases}$
2. Forward;
 - for $i = 1, \dots, M-1$,
 - for $j = 1, \dots, N-1$,
 - $C_{i,j} = \min(C_{i-1,j-1} + 4, C_{i-1,j} + 3, C_{i-1,j+1} + 4, C_{i,j-1} + 3, C_{i,j})$
3. Backward
 - for $i = 1, \dots, M-1$,
 - for $j = 1, \dots, N-1$,

$$C_{i,j} = \min(C_{i,j}, C_{i,j+1} + 3, C_{i+1,j-1} + 4, C_{i+1,j} + 3, C_{i+1,j+1} + 4)$$

由于对像素点灰度特性进行了归一化操作, 则对其斜面距离也进行归一化处理以表示该点的空间位置特性, 公式为:

$$D(p) = \frac{1}{1 + C_{i,j}} \quad (12)$$

(6) 在得知像素点邻域范围内的灰度特性以及空间位置信息的情况下, 可统计像素点的归属性, 用于判断其为背景或者目标, 本文采用将灰度信息与空间信息加权的方法进行判断, 如下式所示:

$$t(p) = \beta \times A(p) + (1 - \beta) \times D(p) \quad (13)$$

β 为权值, 由经验值可得为 0.4, 将 $t(p)$ 值范围限制在 $[0, 1]$ 之间. 步骤如图 5 所示.

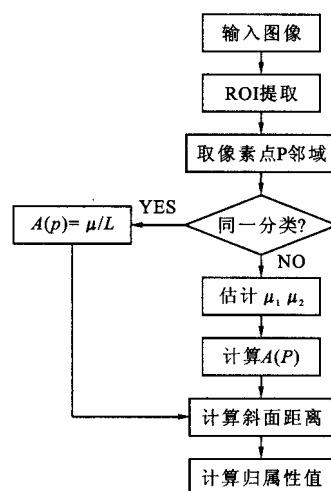


图 5 步骤图

对感兴趣区域中的每一个像素判断其归属性, 也即判断其属于目标或者背景, 以往常用模糊性指数、压缩指数等计算出判断分类的阈值, 但是这些方法只用于灰度判断, 本文综合考虑了像素点的灰度信息即空间位置信息, 本文根据感兴趣区域内所有像素点 $t(p)$ 值的特性, 计算新的阈值步骤如下:

Step 1: 计算出感兴趣区域内所有像素点的 $t(p)$ 值;

Step 2: 将所有的 $t(p)$ 值整数化, 以统计每一个感兴趣区域的直方图, 公式为

$$t'(p) = \text{int}(\frac{t(p)}{\max\{t(p)\}} \times 100),$$

则 $0 \leq t'(p) \leq 100$. (文章中将用 t' 代替 $t(p)$)

Step 3: 以图 2 为例, 将 $h_r(t')$ 定义为感兴趣区域中目标区域内的直方图信息, m_r 为求 $h_r(t')$ 中最大只对应的变量值 t' 的结果值, 公式为: $m_r = \arg \max(h_r(t'))$.

Step 4:与 Step 3 类似,将 $h_b(t')$ 定义为感兴趣区域中目标区域以外的背景信息直方图数据,同理定义 $m_b = \arg \max(h_b(t'))$.

Step 5:在前 4 步的基础上,定义 $h(t')$ 为感兴趣区域内的整体直方图信息,根据 m_r 和 m_b 的取值,定义 $t_f = \arg \min_{m_b \leq \mu < m_r} h(t')$.

根据以上计算步骤得出的阈值,可判断感兴趣区域内任意像素点 P 的归类属性,公式如下:

$$\text{目标区域} = \{p \mid t_f < t'(p)\}$$

3.3 目标提取

对图像感兴趣区域中分离出的目标区域,由目标形状模板等先验知识可进行形状大小筛选,对筛选后的目标区域需进行匹配,以确定最终的目标部分.

本文中运用图像评价指标 SP 值^[8] 对每个提取的目标区域进行判断,由 SP 的计算结果判定目标区域是否提取成功,以确定真正的目标,SP 公式如下式:

$$SP = \frac{\text{num}(S_T \cup S_M) - \text{num}(S_T \cap S_M)}{\text{num}(S_T \cup S_M)} \quad (14)$$

式中, $\text{num}(x)$ 为 x 中元素的个数, S_T 为算法中提取的目标区域内的像素点的集合, S_M 为先验知识中目标形状模板. 当 $SP = 0$ 时,为目标提取效果最佳状态;当 $SP = 1$ 时,完全没有提取到目标.

4 试验及结果分析

为验证本文所提算法的有效性,在 CPU3.0 GHz、内存 1.0 GB 的计算机上采用 Matlab 环境进行仿真分析,以大小为 256×256 ,复杂背景、目标亮度与背景相比较亮的图像为实验对象,为更加明确的表示算法中目标基准图与红外实时图中相应大小区域的相似性,事先制作基准图^[9] 的大小均选为各自目标大小,作为目标提取的先验信息.

为验证算法的精准度,在检测感兴趣区域的基础上将本文提取目标的算法与经典的 Otsu 算法、最大熵算法进行对比,将每幅图中计算所得最小的 SP 值所对应的区域认定为是目标区域. 图 6 所示为在目标与背景分离的结果中目标提取的结果示意图.

由图 6 可看出,本文方法中目标提取的实验结果较完整准确,且 SP 值最低,这是因为感兴趣区域的检测首先约束了提取目标区域的范围,在此基础上从局部区域的角度出发,对每一个区域内像素点的邻域特性进行判断,以明确该像素点的归属情况,因此完整的目标信息被保留;相比之下,otsu 方法

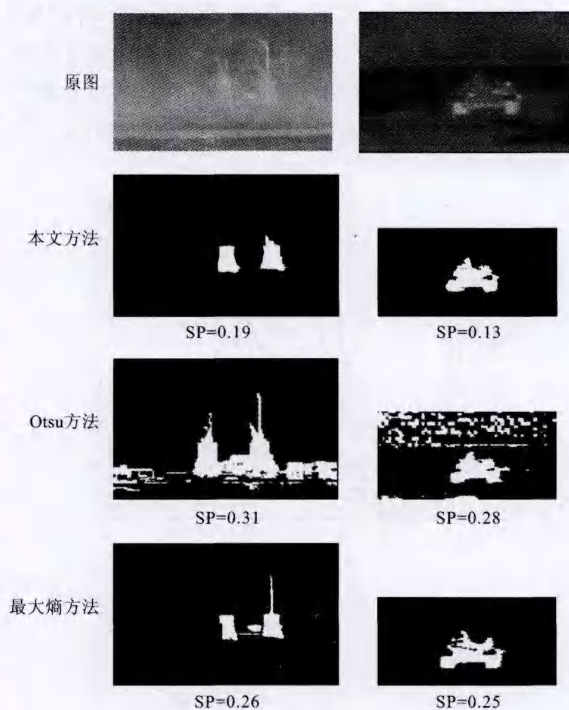


图 6 实验结果图

和最大熵方法虽然也能将目标提取出来,但是由于地面背景与目标存在灰度级交叉,所以实验结果中出现一些将背景错分为目标的情况,且目标提取的不完整,SP 值高.

为测试该算法的整体性能,本文选取两组不同目标对象的红外图像(各 50 幅大小为 256×256 的序列图像),对于每一组图像分别计算其 SP 取值,构成累计概率分布(Cumulative Probability Distribution),如图 7 所示.

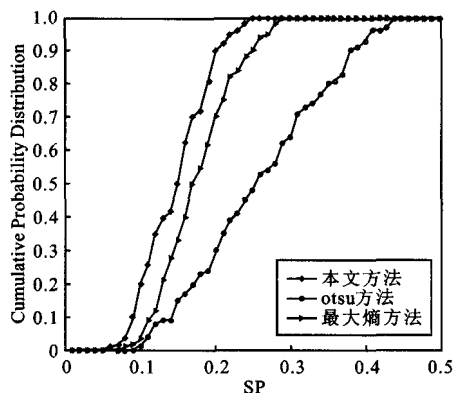
由图 7 可看出对两组序列图而言,本文方法在整体上明显优于 otsu 方法和最大熵方法,具有较高的累计概率.

红外图像目标提取常常用于自动目标识别系统中,因此算法执行时间也尤为重要,然而本文所提算法在目标与背景分离时是从像素点邻域入手,因此,像素点 $n \times n$ 邻域的大小也直接影响着算法运行时间,针对 n 的不同取值进行时间测试,如表 1 所示.

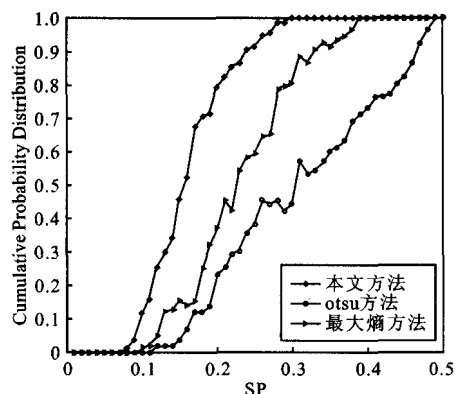
表 1 算法时间

窗口大小	时间/s
3×3	0.235
5×5	0.304
7×7	0.441
9×9	0.622
11×11	0.967
13×13	1.294

由表 1 可看出,随着 n 值的不断变大执行时间也相应变长. 因此本文结合式(13)中的 β , 综合考虑窗口大小和 β 二者对 SP 值的影响. 任意选择十幅红外图像, 分析 n 、 β 和 SP 之间相互影响的情况, 如图 8 所示.



(a) 组一



(b) 组二

图 7 累计概率分布图

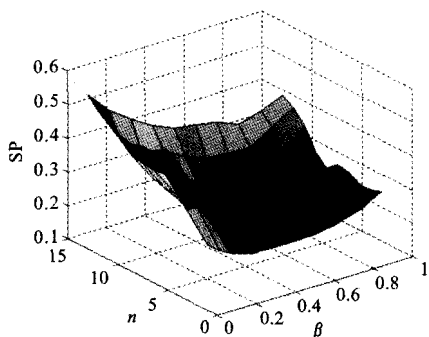


图 8 变量关系图

由图 8 可得出, 当 n 值在 5, 7, 9 之间且 β 在 0.4 到 0.8 之间变化时, SP 的值变化幅度相对较小, 并且值较低, 满足实际工程的需求. 因此, 从整体考虑, 试验中选择 $n=9, \beta=0.6$.

5 结束语

本文针对红外图像目标识别技术难的问题, 提出一种基于感兴趣区域的目标提取方法, 经实验验证, 该方法较经典的 otsu 方法和最大熵方法目标提取较完整且准确, 有效解决了红外图像对比度低、识别难的问题, 充分验证了该算法的有效性和实时性, 对红外图像的目标识别具有很强的实际意义.

参考文献:

- [1] Lee S U, Chuang S Y. A comparative performance study of several global thresholding techniques for segmentation[J]. Computer Vision Graphics and Image Processing, 1990, 52(3): 171-190.
- [2] 王艳娟, 陈晓红, 黄晓欣. 图像感兴趣区域检测技术[J]. 计算机与数字工程, 2007, 35(5): 138-140.
- [3] 郭清风, 王建国. 基于二维最大熵和顺序滤波的红外图像分割[J]. 红外技术, 2008, 30(22): 99-102.
- [4] Eitner P G. Model-based estimation of small targets parameters[J]. Proceedings of SPIE, Signal and Data processing of small targets, 1998(3373): 24-31.
- [5] Haris K, Tziritis G, Orphanoudakis S. Smoothing 2D or 3D images using local classification[C] // Proc. EUSIPCO, Bucharest, Romania, 1994.
- [6] Barrow H G, Tenenbaum J M, Bolles R C, et al. Parametric correspondence and chamfer matching: two new techniques for image matching[C] // Proc. 5th Int. Joint Conf. Artificial Intelligence. Cambridge, 1977: 659-663.
- [7] 李建敏. 基于轮廓片段的图像识别技术研究[D]. 厦门: 厦门大学, 2009.
- [8] Markham K C. Comparison of segmentation processes for object acquisition in infrared images[J]. 1989, 136(1): 13-21.
- [9] 郭文普, 孙继银, 李钊. 基于卫星图像的三维前视基准图制备方法[J]. 无线电工程, 2007, 37(12): 9-11.

作者简介:

高 晶 女, (1983—), 博士研究生. 研究方向为目标识别、图像匹配.

吴育峰 男, (1974—), 硕士, 工程师. 研究方向为人工智能.