

# 面向小目标图像的快速核密度估计图像阈值分割算法

王 骏<sup>1</sup> 王士同<sup>1</sup> 邓赵红<sup>1</sup> 应文豪<sup>1</sup>

**摘 要** 针对当前小目标图像阈值分割研究工作面临的难题, 提出了快速核密度估计图像阈值分割新方法. 首先给出了基于加权核密度估计器的概率计算模型, 通过引入二阶 Renyi 熵作为阈值选取准则, 提出了基于核密度估计的图像阈值分割算法 (Kernel density estimator based image thresholding algorithm, KDET), 然后通过引入快速压缩集密度估计 (Fast reduced set density estimator, FRSDE) 技术, 得到核密度估计的稀疏权系数表示形式, 提出快速核密度估计图像阈值分割算法 fastKDET, 并从理论上对相关性质进行了深入探讨. 实验表明, 本文算法对小目标图像阈值分割问题具有更广泛的适应性, 并且对参数变化不敏感.

**关键词** 核密度估计, 快速压缩集密度估计, 阈值分割, 小目标图像

**引用格式** 王骏, 王士同, 邓赵红, 应文豪. 面向小目标图像的快速核密度估计图像阈值分割算法. 自动化学报, 2012, 38(10): 1679–1689

**DOI** 10.3724/SP.J.1004.2012.01679

## Fast Kernel Density Estimator Based Image Thresholding Algorithm for Small Target Images

WANG Jun<sup>1</sup> WANG Shi-Tong<sup>1</sup> DENG Zhao-Hong<sup>1</sup> YING Wen-Hao<sup>1</sup>

**Abstract** In order to threshold the image containing small targets well, a novel fast kernel density estimator based image thresholding algorithm is proposed. Firstly, a novel computation model for probability density estimation based on the kernel density estimator with weighting coefficients is presented. By introducing the 2nd-order Renyi entropy as the threshold selection criterion, a novel kernel density estimator based image thresholding algorithm (KDET) is proposed. Then a fast version for KDET, named fastKDET, is proposed by integrating fast reduced set density estimator (FRSDE) and RSDE into the data condensation procedure. Moreover, some fundamental theoretical properties are fully studied. At last, several experiments are conducted and show that our fastKDET is more general than some of the existing algorithms and is insensitive to the parameters.

**Key words** Kernel density estimation, fast reduced set density estimator (FRSDE), image thresholding, small target image

**Citation** Wang Jun, Wang Shi-Tong, Deng Zhao-Hong, Ying Wen-Hao. Fast kernel density estimator based image thresholding algorithm for small targets images. *Acta Automatica Sinica*, 2012, 38(10): 1679–1689

在很多图像处理应用中, 图像中目标和背景区域的像素点可以通过使用某一合理的灰度值 (阈值) 有效地分开, 图像的阈值分割技术对这类图像尤为适用. 近年来, 新思想和新理论不断出现, 日益丰富着这项研究的内容<sup>[1–2]</sup>.

在小目标图像中, 背景区域包含了图像中的绝大多数像素点, 而目标区域则仅包含了很少数量的像素点. 诸如大津法<sup>[3]</sup>、最小方差法<sup>[4]</sup> 等基于类内方差的阈值分割算法对这类图像通常不能取得理想的分割效果. 为了对小目标图像进行有效的阈值分割, 研究人员从不同角度提出了解决问题的新方法. 如 Ng 改造大津法的目标函数, 通过向目标函数中加入新的因子来突出小目标区域所处灰度范围在灰度直方图上的权重, 提出了谷底强调法<sup>[5]</sup>; Li 等根据红外图像目标和背景区域具有相似的统计分布这一特点, 基于标准差设计了新的阈值选取准则, 提出了面向红外图像的统计阈值分割算法<sup>[6]</sup>; 陶文兵等通过在阈值分割过程中融入像素点的空间分布信息, 并以标准割作为阈值选取准则, 提出了一种基于图谱理论的阈值分割新方法<sup>[7]</sup>. 实验表明, 这些方法都可以有效地应用于小目标图像的阈值分割.

尽管如此, 这些算法仍然存在着不足. 首先, 这

收稿日期 2011-05-05 录用日期 2011-07-16

Manuscript received May 5, 2011; accepted July 16, 2011

国家自然科学基金 (60903100), 江苏省自然科学基金 (BK2009067, BK2011417), 中央高校基本科研业务费专项资金 (JUSRP111A38), 江苏省信息融合软件工程技术研究开发中心开放基金项目 (SR-2012-01) 资助

Supported by National Natural Science Foundation of China (60903100), Natural Science Foundation of Jiangsu Province (BK2009067, BK2011417), Fundamental Research Funds for the Central Universities (JUSRP111A38), and Opening Project of Jiangsu Engineering Research and Development Center for Information Fusion Software (SR-2012-01)

本文责任编辑 刘一军

Recommended by Associate Editor LIU Yi-Jun

1. 江南大学数字媒体学院 无锡 214122

1. School of Digital Media, Jiangnan University, Wuxi 214122

些算法通常建立在某些假设的基础上. 例如, 谷底强调法假设图像的灰度直方图呈单峰分布, 当图像中背景区域对应的直方图呈现出双峰或多峰特征时, 它常常会把部分背景区域的像素点误分到目标区域中; 统计阈值分割算法<sup>[6]</sup> 假设图像的目标和背景区域具有相似的统计分布, 这一假设具有一定的局限, 从而影响算法的通用性. 其次, 这些算法的性能依赖于相关参数的合理设置. 由于难以获取相关的先验信息, 无监督学习中的参数设置问题一直是困扰研究人员的难题. 如陶文兵等基于图谱理论的阈值分割算法需要根据不同的图像合理设置参数  $d_1$  和  $d_x$  的值, 当用户没有找到合适的参数值时, 分割算法可能会得到失败的分割结果.

本文中, 我们将从全新的角度来探讨面向小目标图像的阈值分割问题. 对于大多数图像来说, 相同灰度值的像素点在图像中的分布相对集中, 空间上相互靠近的像素点具有相同灰度值的概率也比较大. 小目标图像也同样遵循这一规律. 基于这一事实, 我们使用核密度估计器来反映像素点在空间中出现的概率密度. 但是, 核密度估计通常会伴随着相当大的计算量. 幸运的是, 核密度估计器的稀疏权系数表示形式为我们提供了解决问题的新途径. 我们在提出核密度估计图像阈值分割算法的基础上, 将进一步研究核密度估计器的稀疏表示形式和压缩集密度估计技术在图像阈值分割技术中的应用, 以及由此为图像阈值分割算法带来的新特性. 与已有方法相比, 本文方法的优点在于, 它对图像的灰度直方图形态没有过多的假设, 因此适用范围更为广泛; 另一方面, 分割结果对参数变化不敏感, 这也有利于用户的实际使用.

本文如下组织: 第 1 节提出了一种新的基于加权核密度估计器的概率计算模型, 并引入二阶 Renyi 熵作为阈值选取准则, 提出核密度估计图像阈值分割算法 (Kernel density estimation thresholding, KDET); 第 2 节通过引入压缩集密度估计器 (Reduced set density estimator, RSDE) 和快速压缩集密度估计器 (Fast reduced set density estimator, FRSD) 对像素点进行压缩, 提出快速核密度估计图像阈值分割算法 fastKDET, 并从理论上深入探讨相关性质; 第 3 节通过实验验证本文算法宽广的适应性和对参数变化的鲁棒性; 最后给出结论.

## 1 核密度估计图像阈值分割算法

### 1.1 算法推导

令  $f(\mathbf{x})$  为图像中坐标  $\mathbf{x} = (a, b)$  处的像素点灰度值, 大小为  $N = m \times n$  的图像可表示为  $\{f(\mathbf{x})|\mathbf{x} \in [1, 2, \dots, m] \times [1, 2, \dots, n]\}$ . 记  $V_i$  为图像中灰度级为  $i$  的像素点集合,  $N_i$  为图像中灰度为  $i$

的像素点数量, 其中  $i = 1, \dots, L$ ,  $L$  为图像中灰度级总数. 给定灰度值为  $i$  的像素点集合  $V_i$ , 对于图像中任一空间坐标为  $\mathbf{x}$  的像素点, 其灰度级为  $i$  的条件概率密度记为  $p(f(\mathbf{x}) = i|V_i)$ . 使用加权核密度估计器<sup>[8]</sup> 进行概率密度估计, 计算  $p(f(\mathbf{x}) = i|V_i)$  如下:

$$p(f(\mathbf{x}) = i|V_i) = \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \quad (1)$$

其中,  $K(\mathbf{x} - \mathbf{x}_j, \sigma^2) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_j\|^2}{2\sigma^2}\right) / 2\pi\sigma^2$  为核函数,  $\sigma^2 = N_i^{-1/2}$ ,  $\alpha_j \geq 0$  为权重系数, 满足  $\sum_{j=1}^{N_i} \alpha_j = 1$ . 当全部像素点参与概率密度估计时,  $\alpha_j = 1/N_i, j = 1, 2, \dots, N_i$ ; 否则, 可以使用核密度估计器的稀疏权系数表示形式来进行概率密度的估计, 从而达到减少参与计算的像素点数量的目的.

灰度为  $i$  的像素点在图像中出现的概率可计算为  $p(V_i) = N_i/N, i = 1, \dots, L$ . 用  $p(f(\mathbf{x}) = i, V_i)$  表示位置为  $\mathbf{x}$  的像素点具有灰度级  $i$  的概率, 则:

$$p(f(\mathbf{x}) = i, V_i) = p(V_i)p(f(\mathbf{x}) = i|V_i) = \frac{N_i}{N} \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \quad (2)$$

假设图像由阈值  $T$  分为目标和背景两部分, 则位置坐标为  $\mathbf{x}$  的像素点属于目标区域的概率  $p_o^{(T)}(\mathbf{x})$  和  $p_b^{(T)}(\mathbf{x})$  分别计算如下:

$$p_o^{(T)}(\mathbf{x}) = \sum_{i=1}^T p(f(\mathbf{x}) = i, V_i) = \frac{1}{N} \sum_{i=1}^T N_i \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \quad (3)$$

$$p_b^{(T)}(\mathbf{x}) = \sum_{i=T+1}^L p(f(\mathbf{x}) = i, V_i) = \frac{1}{N} \sum_{i=T+1}^L N_i \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \quad (4)$$

信息论中通常使用熵来表示对象的不确定性, 熵的值越大, 对象的不确定性也越大. 用  $p_o(x)$  和  $p_b(x)$  分别表示位置坐标为  $x$  的像素点属于目标和背景区域的概率, 其  $r$  阶 Renyi 熵可以计算如下<sup>[9]</sup>:

$$H(\mathbf{x}) = \frac{1}{1-r} \log(p_o^r(\mathbf{x}) + p_b^r(\mathbf{x})) \quad (5)$$

对于  $r > 1$ ,  $p_o^r(\mathbf{x}) + p_b^r(\mathbf{x})$  越小, 相应的  $H(\mathbf{x})$  越大. 本文取  $r = 2$ .

当图像被阈值  $T$  分为目标和背景两部分时, 位置坐标为  $\mathbf{x}$  的像素点属于目标和背景区域的概率分别记为  $p_o^{(T)}(\mathbf{x})$  和  $p_b^{(T)}(\mathbf{x})$ , 构造目标准则函数如下:

$$f(T) = \int_{\Omega} \left( (p_o^{(T)}(\mathbf{x}))^2 + (p_b^{(T)}(\mathbf{x}))^2 \right) d\mathbf{x} = \int_{\Omega} (p_o^{(T)}(\mathbf{x}))^2 d\mathbf{x} + \int_{\Omega} (p_b^{(T)}(\mathbf{x}))^2 d\mathbf{x} \quad (6)$$

其中,  $\Omega$  为  $\mathbf{x}$  所在特征空间. 最佳阈值  $T_{opt}$  可计算如下:

$$T_{opt} = \arg \min_T f(T) \quad (7)$$

下面讨论式 (6) 的计算. 根据文献 [9] 中式 (6), 有如下两个关系成立:

$$\int_{\Omega} G(\mathbf{x} - \mathbf{x}_1, \sigma_1^2) G(\mathbf{x} - \mathbf{x}_2, \sigma_2^2) d\mathbf{x} = G(\mathbf{x}_1 - \mathbf{x}_2, \sigma_1^2 + \sigma_2^2) \quad (8)$$

所以有:

$$\int_{\Omega} (p_o^{(T)}(\mathbf{x}))^2 d\mathbf{x} = \int_{\Omega} \left( \frac{1}{N} \sum_{i=1}^T N_i \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \right)^2 d\mathbf{x} \quad (9)$$

类似地, 有:

$$\int_{\Omega} (p_b^{(T)}(\mathbf{x}))^2 d\mathbf{x} = \int_{\Omega} \left( \frac{1}{N} \sum_{i=T+1}^L N_i \sum_{j=1}^{N_i} \alpha_j K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \right)^2 d\mathbf{x} = \frac{1}{N^2} \sum_{i=T+1}^L \sum_{l=T+1}^L \sum_{j=1}^{N_i} \sum_{k=1}^{N_l} N_i N_l \alpha_j \alpha_k \times K(\mathbf{x}_j - \mathbf{x}_k, \sigma_j^2 + \sigma_k^2) \quad (10)$$

令  $M$  为  $256 \times 256$  的权重矩阵, 记

$$M(i, l) = \sum_{j=1}^{N_i} \sum_{k=1}^{N_l} N_i N_l \alpha_j \alpha_k K(\mathbf{x}_j - \mathbf{x}_k, \sigma_j^2 + \sigma_k^2) \quad (11)$$

$M$  为其中位置  $(i, l)$  处的元素值, 显然  $M$  为对称矩阵. 在各像素点权重  $\alpha_i$  已知的前提下, 就可以构建权重矩阵  $M$ . 综合式 (9) ~ (11), 有:

$$\int_{\Omega} (p_o^{(T)}(\mathbf{x}))^2 d\mathbf{x} = \frac{1}{N^2} \sum_{i=1}^T \sum_{l=1}^T M(i, l) \quad (12)$$

$$\int_{\Omega} (p_b^{(T)}(\mathbf{x}))^2 d\mathbf{x} = \frac{1}{N^2} \sum_{i=T+1}^L \sum_{l=T+1}^L M(i, l) \quad (13)$$

在图 1 中, 给出矩阵  $M$  的上三角阵, 并通过任意  $T$  值将其划分为三个区域. 不难发现, 对于任意  $T$  值,

式 (12) 和式 (13) 分别对应上三角阵中区域 A 和 B 内的矩阵元素之和. 这样, 积分运算可以通过矩阵元素的求和操作来完成, 即:

$$f(T) = \int_{\Omega} (p_o^{(T)}(\mathbf{x}))^2 d\mathbf{x} + \int_{\Omega} (p_b^{(T)}(\mathbf{x}))^2 d\mathbf{x} = \text{sum}(A) + \text{sum}(B) \quad (14)$$

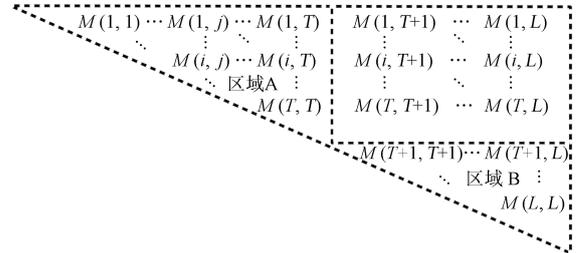


图 1 矩阵  $M$  的上三角阵

Fig. 1 The upper triangular matrix of matrix  $M$

当集合  $V_i$  中的像素点全部用于概率密度估计时, 取  $\alpha_j = 1/N_i, j = 1, 2, \dots, N_i$ . 由式 (11), 矩阵元素  $M(i, l)$  计算如下:

$$M(i, l) = \sum_{j=1}^{N_i} \sum_{k=1}^{N_l} K(\mathbf{x}_j - \mathbf{x}_k, \sigma_j^2 + \sigma_k^2) \quad (15)$$

由此, 得到核密度估计图像阈值分割算法 KDET 如下:

**算法 1. 核密度估计图像阈值分割算法 KDET**

输入. 待分割图像  $I$

输出. 灰度阈值  $T$

读入图像  $I$ ;

设置  $\alpha_i$  为权重向量, 其中  $\alpha_i, i = 1, 2, \dots, N$  表示矩阵中每个像素点的权重;

通过式 (15) 构建矩阵  $M$ ;

for  $t = 1$  to 255

根据式 (14), 计算图 1 中区域 A 和 B 中的元素总和  $f(T)$

end

根据式 (7) 计算求得  $T_{opt}$ .

## 1.2 时间复杂度分析

算法 KDET 的执行时间主要包含两部分: 一部分是计算矩阵  $M$  中每个元素所需的总时间  $t_M$ , 另一部分是对于每个候选阈值  $T$  计算式 (14) 的时间  $t_f$ . 相对于  $t_M, t_f$  基本可以忽略. 因此, 我们主要考察  $t_M$  部分.

根据式 (11), 计算矩阵  $M$  中每个元素所消耗的时间主要集中在乘法运算和幂运算. 这里仅通过计算幂运算  $e^x$  的次数反映其时间复杂度.

**定理 1.** 给定图像  $I$ , 算法 KDET 构建权重矩阵  $M$  所需幂运算总次数的上界为  $L \sum_{i=1}^L N_i^2$ , 其中  $N_i$  为图像中灰度为  $i$  的像素点数量.

**定理 2.** 给定图像  $I$ , 算法 KDET 计算矩阵元素  $M$  所需幂运算总次数的上界为  $\tau_u$ , 即  $\tau_u = L \sum_{i=1}^L N_i^2$ , 则不等式  $N^2 \leq \tau_u \leq LN^2$  成立.

综合定理 1 和定理 2, 算法 KDET 计算矩阵元素  $M$  所需幂运算的总次数  $\tau$  在相当大的范围内变化. 当图像中各灰度级上的像素点数量趋于相等时,  $\tau$  值趋于  $N^2$ ; 相反, 当图像的灰度直方图趋于单峰分布时,  $\tau$  值趋于  $LN^2$ . 因此, 从时间复杂度的角度考虑, KDET 不适用于对包含小目标的高分辨率图像进行分割.

## 2 快速核密度估计图像阈值分割算法 fastKDET

### 2.1 RSDE 和 FRSDE 相结合的像素点压缩技术

精简式 (15) 中参与计算矩阵元素  $M(i, l)$  的像素点数量, 是加快算法 KDET 执行速度的有效方法. 首先给出如下定理:

**定理 3.** 设  $N_i^{(r)}$  为集合  $V_i$  压缩后的大小,  $\sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} = 1$ , 对于任意小正实数  $\delta$  满足

$$\left| N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} K(\mathbf{x} - \mathbf{x}_j, \sigma^2) - \sum_{j=1}^{N_i} K(\mathbf{x} - \mathbf{x}_j, \sigma^2) \right| < \delta \tag{16}$$

则:

$$\left| \sum_{j=1}^{N_i^{(r)}} \sum_{k=1}^{N_l^{(r)}} N_i N_l \alpha_j^{(r)} \alpha_k^{(r)} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) - \sum_{j=1}^{N_i} \sum_{k=1}^{N_l} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) \right| < (N_i + N_l) \delta \tag{17}$$

**证明.** 记

$$M_{il}^{(r)} = \sum_{j=1}^{N_i^{(r)}} \sum_{k=1}^{N_l^{(r)}} N_i N_l \alpha_j^{(r)} \alpha_k^{(r)} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2)$$

根据式 (16), 有:

$$\begin{aligned} \sum_{j=1}^{N_i} K(\mathbf{x} - \mathbf{x}_j, \sigma^2) - \delta &< \\ N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} K(\mathbf{x} - \mathbf{x}_j, \sigma^2) &< \\ \sum_{j=1}^{N_i} K(\mathbf{x} - \mathbf{x}_j, \sigma^2) + \delta & \end{aligned}$$

成立.

记

$$M_{il}^{(r)} = N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} \left( N_l \sum_{k=1}^{N_l^{(r)}} \alpha_k^{(r)} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) \right)$$

则:

$$\begin{aligned} M_{il}^{(r)} &< N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} \left( \sum_{k=1}^{N_l} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) + \delta \right) = \\ \sum_{k=1}^{N_l} \left( N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) \right) &+ N_i \delta < \\ \sum_{k=1}^{N_l} \left( \sum_{j=1}^{N_i} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) + \delta \right) &+ N_i \delta = \\ \sum_{k=1}^{N_l} \sum_{j=1}^{N_i} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) &+ (N_i + N_l) \delta \end{aligned}$$

同理可得:

$$M_{il}^{(r)} > \sum_{k=1}^{N_l} \sum_{j=1}^{N_i} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) - (N_i + N_l) \delta$$

综上所述, 有:

$$\begin{aligned} & \left| \sum_{j=1}^{N_i^{(r)}} \sum_{k=1}^{N_l^{(r)}} N_i N_l \alpha_j^{(r)} \alpha_k^{(r)} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) - \sum_{j=1}^{N_i} \sum_{k=1}^{N_l} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2) \right| < (N_i + N_l) \delta \end{aligned}$$

成立.  $\square$

定理 3 表明, 当  $N_i \sum_{j=1}^{N_i^{(r)}} \alpha_j^{(r)} K(\mathbf{x} - \mathbf{x}_j, \sigma^2)$  的值趋于  $\sum_{j=1}^{N_i} K(\mathbf{x} - \mathbf{x}_j, \sigma^2)$  时, 式 (11) 的值趋于式 (15). 因此如何得到核密度估计器的稀疏权系数表示形式来满足式 (16), 是我们需要解决的问题.

最近, Girolami 等<sup>[8]</sup> 提出了压缩集密度估计器 RSDE, 通过求解二次规划问题, 得到了核密度估计器的稀疏权系数表示形式; Deng 等进一步揭示了 RSDE 与核化中心约束最小包含球 (Center-constrained minimum enclosing ball, CCMEB) 之间的等价关系, 将 RSDE 的求解过程转化为 CCMEB 问题进行求解, 提出了 FRSDE, 成功解决了面向大规模数据集进行概率密度估计时的时间复杂度问题. FRSDE 具有以下优良特性<sup>[10]</sup>: 1) FRSDE 得到的核心集大小和迭代次数的理论上界仅与  $\varepsilon$  值有关, 与原数据集大小无关; 2) 在给定  $\varepsilon$  值的前提下, FRSDE 时间复杂度的理论上界与数据集大小成线性关系; 当在压缩过程中使用抽样策略时, 其时间复杂度与原数据集大小无关, 仅与抽样数据集大小成线性关系.

FRSDE 的实际运算时间由二次规划求解算法的时间复杂度和二次规划的求解次数 (即核心集大

小) 共同决定. 假设 FRSDE 所使用的二次规划时间复杂度为  $O(n^2)$ , 最终得到的核心集大小为  $N_Q$ , 则 FRSDE 实际的时间复杂度为  $\sum_{i=1}^{N_Q} O(i^2)$ . 而对于 RSDE, 压缩大小为  $N$  的数据集, 其时间复杂度为  $O(N^2)$ . 因此在数据集规模较小的情况下, FRSDE 与 RSDE 相比, 并不能体现其速度上的优势.

采用 FRSDE 和 RSDE 相结合的方法来压缩图像中各灰度级上的像素点. 引入  $N_{\text{gate}}$  来表示使用 FRSDE 的临界值: 当集合  $V_i$  的大小  $N_i \leq N_{\text{gate}}$  时, 使用 RSDE 进行数据压缩; 当集合  $V_i$  的大小  $N_i > N_{\text{gate}}$  时, 使用 FRSDE 进行数据压缩. 根据相关性质, FRSDE 压缩集合  $V_i$  实际所使用的时间  $t_i$  与  $V_i$  的大小  $N_i$  无关, 仅与  $\varepsilon$  有关. 更重要的是, 数据经压缩后可以很好地满足我们对压缩算法的要求, 即在保证矩阵元素  $M(i, l)$  计算精度的前提下, 尽量减少参与计算的像素点数量, 并使各灰度级上的像素点数量趋于相等.

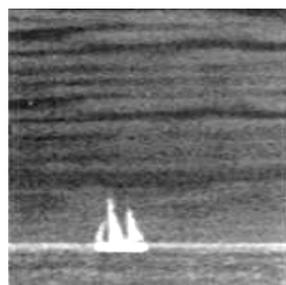
在图 2 中, 我们给出了使用 RSDE 和 FRSDE 对图像 Sailboat 进行压缩, 压缩前后灰度直方图的变化情况. 实验中, 设置  $N_{\text{gate}} = 500$ ,  $\varepsilon = 0.01$ . 在图 2(b) 中, 用灰色表示图像在压缩前的灰度直方图, 用黑色和白色表示压缩后像素点的灰度分布直方图. 从中不难发现, 对于采用 FRSDE 压缩的部分 (即中间的黑色区域), 压缩后各灰度级上像素点的数量几近相等, 而采用 RSDE 压缩的部分 (即黑色部分区域呈小三角形的白色区域), 由于 RSDE 在数据压缩时存在一定的冗余, 所以压缩后各灰度级上像素点的数量与原数据集大小保持着一定的比例关系.

## 2.2 快速核密度估计图像阈值分割算法 fastKDET 及理论分析

本节给出快速核密度估计图像阈值分割算法 fastKDET.

### 算法 2. 快速核密度估计图像阈值分割算法 fastKDET

输入. 待分割图像  $I$ , FRSDE 终止阈值  $\varepsilon$ , 使用



(a) 原图

(a) The original image

FRSDE 的临界值  $N_{\text{gate}}$

输出. 灰度阈值  $T$

读入图像  $I$ ;

for  $i = 0$  to  $L - 1$

if  $N_i > g$

使用 FRSDE 压缩像素点集合  $V_i$ ;

else

使用 RSDE 压缩像素点集合  $V_i$ ;

end

end

通过式 (11) 构建矩阵  $M$ ;

for  $t = 1$  to 255

根据式 (14), 计算图 1 中区域 A 和 B 中的元素总和  $f(T)$

end

根据式 (7) 计算, 求得  $T_{\text{opt}}$ .

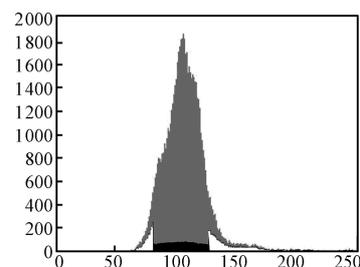
**定理 4.** 给定图像  $I$  和临界值  $N_{\text{gate}}$ , 算法 fastKDET 压缩像素点的时间复杂度上界仅与  $\varepsilon$  有关.

**证明.** 算法 fastKDET 压缩数据集所需运算时间由两部分构成: RSDE 所需时间  $t_{\text{RSDE}}$  和 FRSDE 所需时间  $t_{\text{FRSDE}}$ . 假设图像  $I$  中像素点数目小于  $N_{\text{gate}}$  的灰度级数目为  $C$ , 求解规模为  $n$  的二次规划问题需要时间为  $O(n^2)$ , 则 RSDE 所需要的总时间  $t_{\text{RSDE}}$  满足:  $t_{\text{RSDE}} \leq C \cdot O(N_{\text{gate}}^2)$ .

另一方面, FRSDE 需要的总时间  $t_{\text{reduce}} = t_{\text{RSDE}} + t_{\text{FRSDE}} \leq C \cdot O(N_{\text{gate}}^2) + (L - C) \times O\left(\frac{N}{\varepsilon^2} + \frac{1}{\varepsilon^4}\right)$ . 所以算法 fastKDET 压缩像素点所需的总时间为

$$t_{\text{reduce}} = t_{\text{RSDE}} + t_{\text{FRSDE}} \leq C \cdot O(N_{\text{gate}}^2) + (L - C) \cdot \left(\frac{N}{\varepsilon^2} + \frac{1}{\varepsilon^4}\right) \quad (18)$$

在  $C$  值和  $N_{\text{gate}}$  值确定的前提下, 仅与  $\varepsilon$  有关, 定理成立.  $\square$



(b) 压缩前后的直方图形态

(b) The histogram shapes of Sailboat before and after compression

图 2 Sailboat 经压缩后直方图形态比较

Fig. 2 The comparison of the histogram shapes of Sailboat after compression

**定理 5.** 给定  $\varepsilon$  值和  $N_{\text{gate}}$ , 算法 fastKDET 的时间复杂度上界与图像大小  $N$  呈线性关系.

**证明.** 算法 fastKDET 所需要的时间主要包括两部分: 像素点的压缩时间  $t_{\text{reduce}}$  和矩阵  $M$  的计算时间  $t_M$ . 定理 4 中已经分析了  $t_{\text{reduce}}$ , 下面分析  $t_M$ . 设 RSDE 的数据压缩率为  $\alpha$ , 则灰度级为  $i$  的像素点集合  $V_i$  经 RSDE 压缩后大小  $N_i^{(r)} = \alpha N_i \leq \alpha N_{\text{gate}}$ ; 而给定  $\varepsilon$  值, 像素点集合  $V_i$  经 FRSE 压缩后所得压缩集大小  $N_i^{(r)}$  的上界为  $O(1/\varepsilon)$ , 与像素点集合  $V_i$  的大小  $N_i$  无关. 因此计算矩阵  $M$  所需执行的幂运算总次数  $\sum_{i=1}^L N_i^{(r)}$  的上界与图像大小  $N$  无关. 给定  $\varepsilon$  值和  $N_{\text{gate}}$  值, 算法 fastKDET 所需的总时间计算为  $t_{\text{fastKDET}} = t_{\text{reduce}} + t_M$ . 根据式 (18),  $t_{\text{fastKDET}}$  的上界仅与图像大小  $N$  呈线性关系.  $\square$

### 3 实验

#### 3.1 软硬件环境

本文实验所采用的软硬件环境如表 1 所示.

表 1 实验的软件硬件环境

Table 1 The hardware and software environment

CPU	内存	操作系统	软件平台
T7100	3G	Vista Home Edition	Matlab R2008b

#### 3.2 与 Parzen 窗图像阈值分割算法<sup>[11]</sup> 进行比较

首先将 fastKDET 与文献 [11] 的 Parzen 窗图像阈值分割算法 (Parzen window estimation, PWT) 进行比较. 实验中, 对图像 airplane (见图 3(a)) 使用不同比例进行缩放, 将其大小调整为 64 像素  $\times$  64 像素、128 像素  $\times$  128 像素、256 像素  $\times$  256 像素, 考察不同大小的图像对算法运行时间的影响. 表 2 给出了算法 PWT 进行阈值分割所用的时间和 fastKDET 压缩像素点后分割图像所用的时间.

根据表 2 所列的结果不难发现, 本文算法和 PWT 在分割效果上相接近. 但是, 在运行时间方面, 随着图像的增大, PWT 所需的时间急剧增加; 而 fastKDET 的运行时间则增长缓慢. 可见, 随着图像中像素点数量的增加, 算法 fastKDET 在运

行速度方面体现出了明显的优势, 这符合前文关于 fastKDET 算法时间复杂度的分析. 在其他图像的实验中也可以得到类似的结论. 由于篇幅有限, 本文不再列出相关结果.

#### 3.3 小目标图像分割效果实验

本实验中, 使用一系列场景中包含小目标的红外图像来测试算法 fastKDET, 并将其与 OTSU<sup>[3]</sup>, MCVT<sup>[4]</sup>, VET<sup>[5]</sup>, MCET<sup>[12]</sup> 这几个算法进行比较. 为了便于分析, 通过手工调整的方式得到各图像的参考阈值. 在确定参考阈值时, 主要从保持目标完整性的角度出发, 兼顾最小化背景区域包含的噪声点数量. 各算法在不同小目标图像上的分割效果如图 3~图 9 所示, 其中, 图 (a) 为待分割的原图, 图 (b) 为对应图像的灰度直方图, 图 (c) 为使用参考阈值所得分割结果, 图 (d)~(h) 分别为 fastKDET ( $\varepsilon = 0.01$ ,  $N_{\text{gate}} = 500$ )、OTSU、MCVT 和 MCET 等算法所得的分割结果.

为了量化算法的分割性能, 采用  $ME$  (Misclassification error)<sup>[13]</sup> 指标对不同算法所得分割结果进行比较.  $ME$  的计算式如下:

$$ME = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o + F_o|} \quad (19)$$

其中,  $B_o$  为标准分割图像的背景区域;  $F_o$  为标准分割图像的目标区域;  $B_T$  为测试图像分割后的背景区域;  $F_T$  为测试图像分割后的目标区域;  $|\cdot|$  表示区域中像素点的数目. 由于标准分割图像难于获取, 本文中把参考阈值下得到的分割图像作为式 (19) 中的标准分割图像. 对于上述所给出的分割结果, 计算  $ME$  指标如表 3 所示.

从图 3 中可以看出, 对于背景单一的小目标图像, 其灰度直方图呈狭窄的单峰分布; 而对于背景相对复杂的小目标图像而言, 其灰度直方图可能呈单峰、双峰或多峰等多种形态, 但是像素的灰度值都集中在某一狭小的区间范围内. OTSU 算法和 MCVT 算法从类内方差的角度考虑问题, 它们趋向于将图像分割为大小相近的两部分, 因此很容易将背景区域中的某些像素点误判为目标. 对于 VET 算法而言, 它对大多数小目标图像有着比较好的分割效果.

表 2 PWT 和 fastKDET 运行时间比较 (秒)

Table 2 The comparison of running time for PWT and fastKDET (s)

	分割所用时间			阈值		
	64 $\times$ 64	128 $\times$ 128	256 $\times$ 256	64 $\times$ 64	128 $\times$ 128	256 $\times$ 256
PWT	4.711	30.982	402.12	192	191	188
fastKDET ( $N_{\text{gate}} = 500$ , $\varepsilon = 0.01$ )	4.188	14.5075	20.225	194	191	189

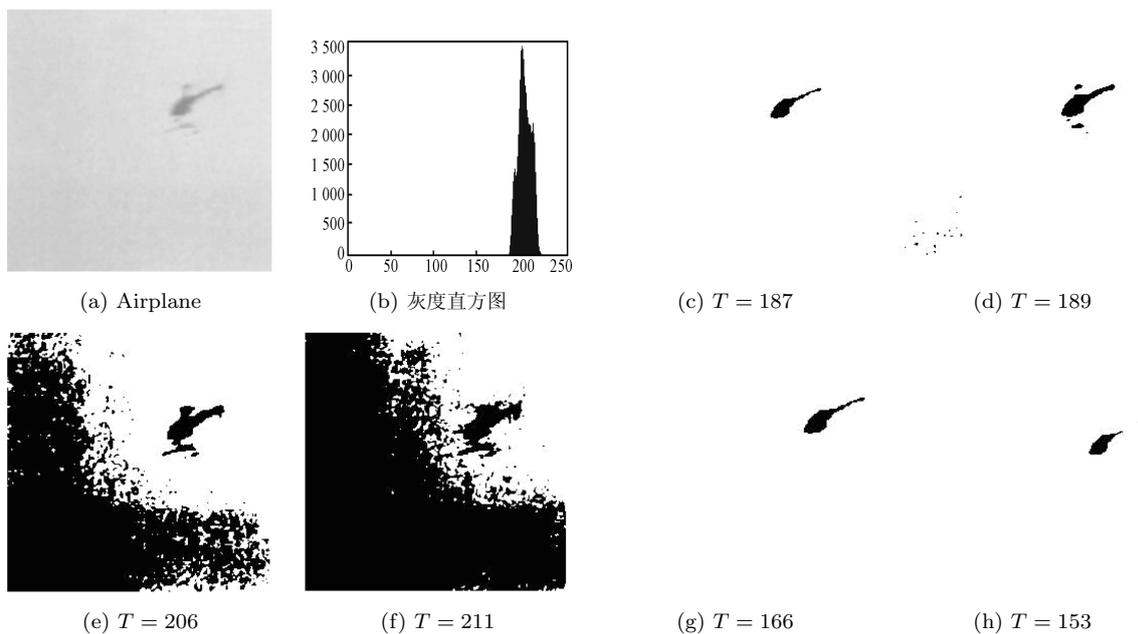


图 3 Airplane 分割结果比较

Fig. 3 The comparison of segmentation results for Airplane

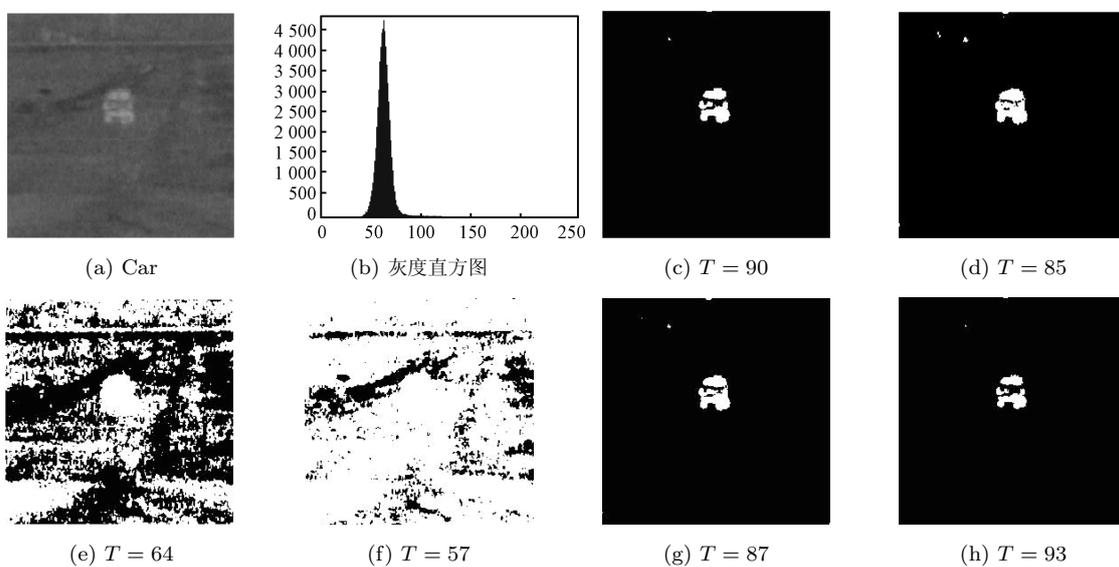
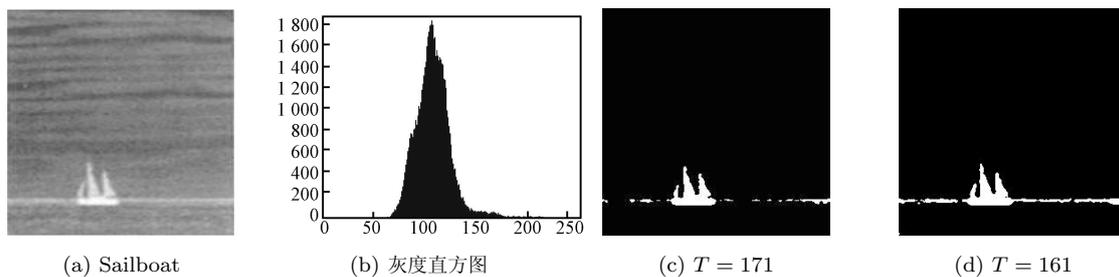


图 4 Car 分割结果比较

Fig. 4 The comparison of segmentation results for Car



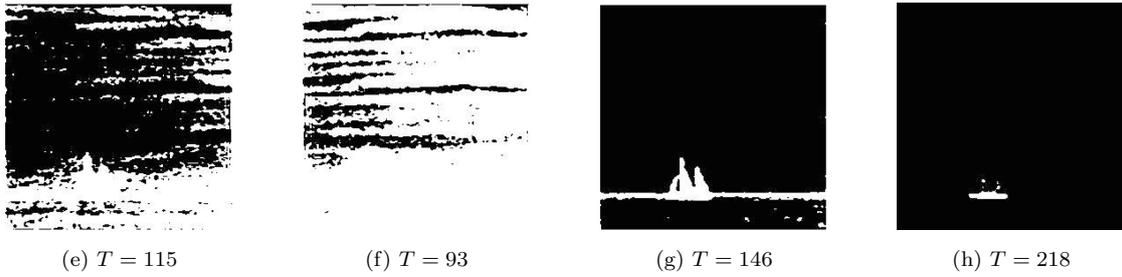


图 5 Sailboat 分割结果比较

Fig. 5 The comparison of segmentation results for Sailboat

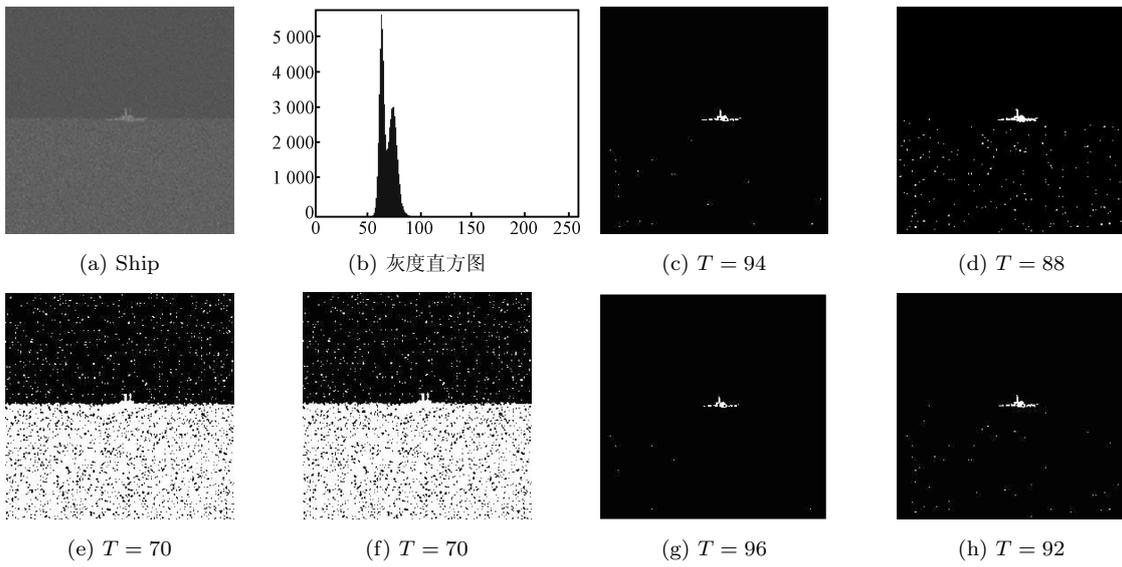


图 6 Ship 分割结果比较

Fig. 6 The comparison of segmentation results for Ship

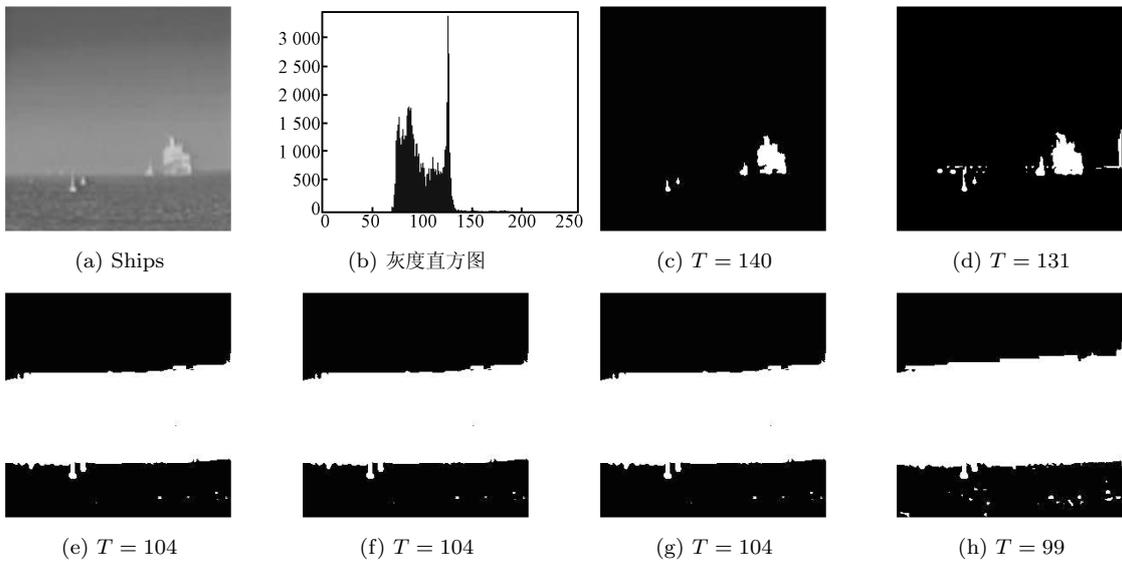


图 7 Ships 分割结果比较

Fig. 7 The comparison of segmentation results for Ships

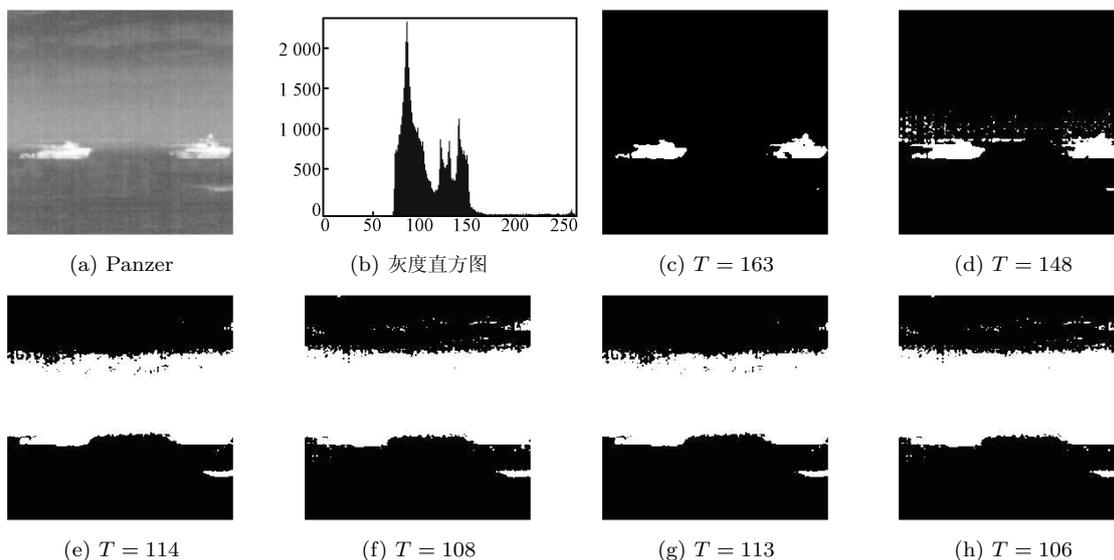


图 8 Panzer 分割结果比较

Fig. 8 The comparison of segmentation results for Panzer

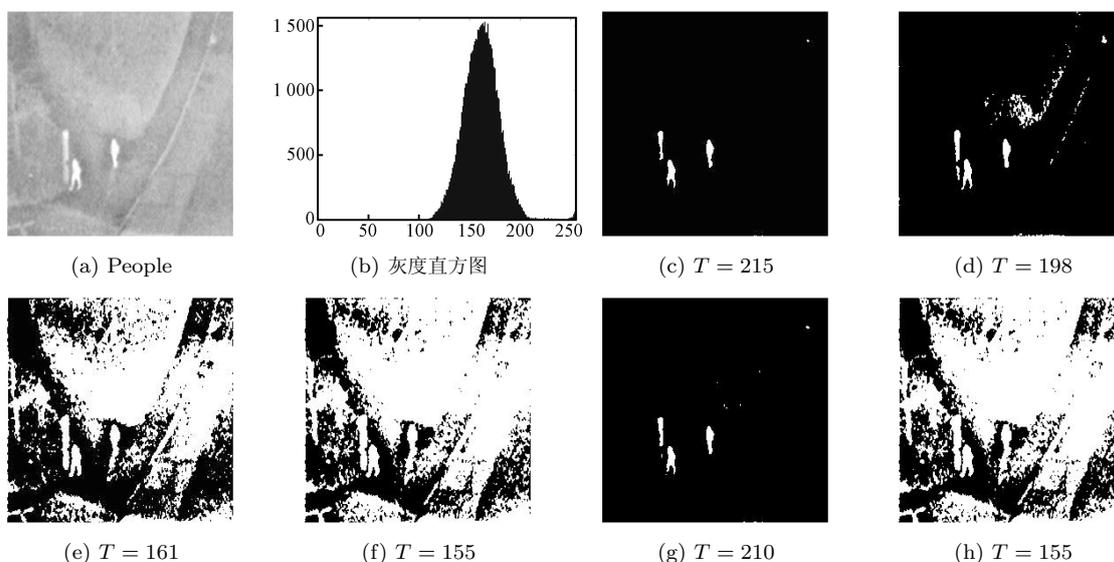


图 9 People 分割结果比较

Fig. 9 The comparison of segmentation results for People

表 3 不同方法 ME 值比较

Table 3 The comparison of ME for different algorithms

图像名	参考阈值	fastKDET	OTSU	MCVT	VET	MCET
Airplane	187	0.0032	0.5247	0.7139	0.0041	0.0058
Car	90	0.0033	0.3621	0.8048	0.0017	0.0014
Sailboat	171	0.0080	0.3127	0.8090	0.0235	0.0132
Ship	94	0.0032	0.4436	0.4436	0.0004	0.0005
Ships	140	0.0067	0.3914	0.3914	0.3914	0.4435
Panzer	163	0.0143	0.3468	0.3761	0.3515	0.3888
People	215	0.0122	0.4917	0.6249	0.0010	0.6249

但是对于图像 Ships 和 Panzer 来说, 其灰度直方图上两个峰值之间出现了较深的低谷, 其强调谷底的加权方式 (参考文献 [5] 式 (8)) 使得谷底区域占了相当大的权重, 因此得到了失败的分割效果. 而 MCET 使用 Poisson 分布来近似估计图像中目标和背景区域中像素点灰度值的分布情况; 当实际图像不符合这一假设时, 分割结果就不令人满意.

而 fastKDET 直接从像素点空间分布出发, 并以核密度估计器的稀疏权系数表示形式进行概率密度的计算, 这一计算模型有着很强的通用性, 所以与其他算法相比, 具有更好的适应性和更广泛的应用范围. 对于个别图像而言, 虽然 fastKDET 不能得到最佳的分割效果, 但是它得到的分割结果与最佳分割结果相差不大, 也是可以接受的.

### 3.4 参数敏感性实验

参数  $\varepsilon$  和  $N_{\text{gate}}$  是算法 fastKDET 中的两个重要参数. 本实验通过调整参数  $\varepsilon$  和  $N_{\text{gate}}$  的值来研究算法 fastKDET 对参数变化的敏感性. 使用 Car、Ship 和 Panzer 这三张分辨率均为 256 像素  $\times$  256 像素的图像, 它们的灰度直方图分别呈单峰、双峰和多峰分布. 使用压缩率来表示图像数据被压缩的程度, 压缩率 CR 可计算为:

$$CR = \frac{\text{The pixels' number obtained by the reduction procedure}}{\text{the pixels' number in the original image}}$$

首先研究参数  $\varepsilon$  对算法性能的影响. 设置  $N_{\text{gate}} = 0$ , 即所有灰度级上的像素点都用 FRSE 进行压缩. 同时, 参数  $\varepsilon$  在 [0.001, 0.064] 的范围内变化. 表 4 给出了在不同的图像上执行算法, 参数  $\varepsilon$  变化时算法 fastKDET 得到的阈值、压缩率和压缩时间.

从表 4 中不难发现, 算法 fastKDET 得到的阈值对参数  $\varepsilon$  的变化不敏感. 根据文献 [10], 参数  $\varepsilon$  的取值直接影响核心集的大小. 由定理 3, 当压缩集里包含的像素点数量  $N_i^{(r)} \rightarrow N_i$  时,  $M(i, l) \rightarrow \sum_{k=1}^{N_l} \sum_{j=1}^{N_i} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2)$ . 在算法的

实际执行过程中, 只要集合  $V_i$  对应的核心集包含的像素点达到一定数量,  $M(i, l)$  就已经接近于  $\sum_{k=1}^{N_l} \sum_{j=1}^{N_i} K(\mathbf{x}_j - \mathbf{x}_k, \sigma^2)$ , 此时由式 (14) 计算得到的目标函数形态几乎不再随着核心集的增大而产生变化. 因此, fastKDET 的分割效果对参数  $\varepsilon$  均不敏感.

另一方面, 从数据压缩率的角度来看, 随着参数  $\varepsilon$  的逐渐变小, 其压缩率的值逐渐变大. 这意味着求解  $M$  矩阵中的每个元素将需要更多的时间. 根据我们的大量实验,  $\varepsilon$  取 0.01 是比较好的选择.

使用类似的实验方法, 设置  $\varepsilon = 0.01$ , 研究参数  $N_{\text{gate}}$  的变化对算法性能的影响. 结果如表 5 所示.

从表 5 中可以看出, 由于 RSDE 的压缩率高于 FRSE<sup>[10]</sup>, 所以随着参数  $N_{\text{gate}}$  值的逐渐增大, 压缩率也依次增大. 但是, 算法计算所得的阈值几乎不受影响. 可见算法 fastKDET 也同样对参数  $N_{\text{gate}}$  的取值不敏感. 从运行时间的角度考虑,  $N_{\text{gate}}$  的值在 400 ~ 800 之间时, 算法可以在较短的时间内完成数据压缩. 当  $N_{\text{gate}}$  值过小时, FRSE 会使用更多的时间; 反之, 当  $N_{\text{gate}}$  值过大时, RSDE 也将会耗费更多的时间.

## 4 结论

面向小目标图像的阈值分割是图像分割研究领域重点的关注课题, 已有的方法仍然存在着通用性不强、算法性能依赖于参数设置等问题. 本文提出的新算法 fastKDET 则有效地克服了上述缺点. 其原因在于, 一方面, 该算法使用核密度估计器的稀疏权系数表示形式进行概率密度估计, 这一计算模型具有更强的通用性; 另一方面, 算法 fastKDET 使用 RSDE 和 FRSE 相结合的方式对各灰度级的像素点进行压缩, 很好地继承了它们的优良特性.

在参数  $\varepsilon$  固定的前提下, 算法 fastKDET 的时间复杂度上界与图像大小呈线性关系, 这为它高效地应用于高分辨率图像提供了坚实的保证. 本文算

表 4 参数  $\varepsilon$  变化对算法 fastKDET 的影响

Table 4 The effect of the parameter  $\varepsilon$ 's change on fastKDET

$\varepsilon$	Car			Ship			Panzer		
	取值	阈值	压缩率	压缩耗时 (秒)	阈值	压缩率	压缩耗时 (秒)	阈值	压缩率
0.064	83	0.0173	3.28	87	0.0118	2.39	151	0.0356	6.28
0.032	83	0.0262	8.56	87	0.0199	7.34	149	0.0528	15.55
0.016	85	0.0377	20.93	88	0.0319	20.67	149	0.0704	32.81
0.008	85	0.0506	45.39	88	0.046	47.81	148	0.0887	53.36
0.004	85	0.0640	85.62	88	0.0607	91.36	148	0.1048	83.89
0.002	85	0.0758	139.22	88	0.0755	156.25	148	0.1185	119.47
0.001	85	0.0862	209.53	88	0.0896	244.95	148	0.1278	151.20

表 5 参数  $N_{\text{gate}}$  变化对算法 fastKDET 的影响  
Table 5 The effect of the parameter  $N_{\text{gate}}$ 's change on fastKDET

$N_{\text{gate}}$ 取值	Car			Ship			Panzer		
	阈值	压缩率	压缩耗时 (秒)	阈值	压缩率	压缩耗时 (秒)	阈值	压缩率	压缩耗时 (秒)
0	83	0.0461	34.58	88	0.0413	37.62	148	0.0828	44.80
200	83	0.0508	31.90	88	0.0442	33.63	148	0.0863	40.39
400	85	0.0540	30.99	88	0.0505	31.46	148	0.0941	38.76
600	85	0.0569	30.09	88	0.0541	31.72	148	0.1164	37.80
800	85	0.0637	30.37	88	0.0585	31.24	148	0.1505	46.61
1000	85	0.0669	31.44	88	0.0651	32.92	148	0.1842	61.61
1200	85	0.0745	35.15	89	0.0801	38.94	148	0.2121	80.00
1400	85	0.0793	38.95	88	0.0799	39.27	148	0.2244	90.83
1600	85	0.0846	43.69	88	0.0899	46.30	148	0.2328	103.02

法运行时间是在普通 PC 机上的 Matlab 环境下获得的。随着计算机硬件技术的飞速发展、高档计算机的日益普及和高效编程语言的使用, 本文算法的运行速度还会有很大的提升空间。

## References

- Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 2004, **13**(1): 146–165
- Li Zuo-Yong. Research on Image Thresholding Methods Based on Statistic and Spectral Graph [Ph. D. dissertation], Nanjing University of Science and Technology, China, 2010 (李佐勇. 基于统计和谱图的图像阈值分割方法研究 [博士学位论文], 南京理工大学, 中国, 2010)
- Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, **9**(1): 62–66
- Hou Z, Hu Q, Nowinski W L. On minimum variance thresholding. *Pattern Recognition Letters*, 2006, **27**(14): 1732–1743
- Ng H F. Automatic thresholding for defect detection. *Pattern Recognition Letter*, 2006, **27**(14): 1644–1649
- Li Z Y, Liu C C, Liu G H, Yang X B, Cheng Y. Statistical thresholding method for infrared images. *Pattern Analysis and Applications*, **14**(2): 109–126
- Tao Wen-Bing, Jin Hai. A new image thresholding method based on graph spectral theory. *Chinese Journal of Computers*, 2007, **30**(1): 110–119 (陶文兵, 金海. 一种新的基于图谱理论的图像阈值分割方法. 计算机学报, 2007, **30**(1): 110–119)
- Girolami M, Chao H. Probability density estimation from optimally condensed data samples. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(10): 1253–1264
- Torkkola K. Feature extraction by non parametric mutual information maximization. *Journal of Machine Learning Research*, 2003, **3**: 1415–1438
- Deng Z H, Chung F L, Wang S T. FRSDE: Fast reduced set density estimator using minimal enclosing ball approximation. *Pattern Recognition*, 2008, **41**(4): 1363–1372
- Wang S T, Chung F L, Xiong F S. A novel image thresholding method based on Parzen window estimate. *Pattern Recognition*, 2008, **41**(1): 117–129
- Pal N R. On minimum cross-entropy thresholding. *Pattern Recognition*, 1996, **29**(4): 575–580
- Yasnoff W A, Mui J K, Bacus J W. Error measures for scene segmentation. *Pattern Recognition*, 1997, **9**(4): 217–231



王骏 博士, 江南大学数字媒体学院讲师。主要研究方向为模式识别, 数据挖掘, 数字图像处理。本文通信作者。

E-mail: wangjun\_sytu@sina.com

(WANG Jun Ph. D., lecturer at the School of Digital Media, Jiangnan University. His research interest covers pattern recognition, data mining, and digital image processing. Corresponding author of this paper.)



王士同 江南大学数字媒体学院教授。主要研究方向为人工智能, 模式识别和生物信息。

E-mail: wxwangst@yahoo.com.cn

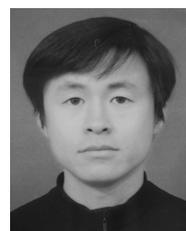
(WANG Shi-Tong Professor at the School of Digital Media, Jiangnan University. His research interest covers artificial intelligence, pattern recognition, and bioinformatics.)



邓赵红 博士, 江南大学数字媒体学院副教授。主要研究方向为模糊建模和智能计算。

E-mail: dzh666828@yahoo.com.cn

(DENG Zhao-Hong Ph. D., associate professor at the School of Digital Media, Jiangnan University. His research interest covers fuzzy modeling and intelligent computation.)



应文豪 江南大学数字媒体学院博士研究生。主要研究方向为模式识别和智能计算。E-mail: cslgywh@163.com (YING Wen-Hao Ph. D. candidate at the School of Digital Media, Jiangnan University. His research interest covers pattern recognition and intelligent computation.)