

基于先验形状约束水平集模型的建筑物提取方法

田昊¹ 杨剑¹ 汪彦明¹ 李国辉¹

摘要 提出了一种先验形状约束的变分水平集模型,并将其用于单幅遥感图像多建筑物的自动提取中.将多个先验形状竞争模型引入水平集方法中,在标记函数的指导下,利用先验形状能量来约束曲线的演化,在对图像进行分割的同时完成建筑物的检测和提取.标记函数的引入,加强了先验形状与要检测目标之间的匹配关系.同时本文提出的模型具有先验形状的旋转、缩放和平移不变性.最后的实验结果及定量定性的分析说明了本文方法的可行性.

关键词 水平集, 先验形状, 标记函数, 建筑物检测, 分割, 变分方法

DOI 10.3724/SP.J.1004.2010.01502

Towards Automatic Building Extraction: Variational Level Set Model Using Prior Shape Knowledge

TIAN Hao¹ YANG Jian¹ WANG Yan-Ming¹ LI Guo-Hui¹

Abstract A novel variational level set model for multiple-building extraction from a single remote image is proposed in this paper. Multi-competing shapes are considered together with the level set model, the curve evolution is constrained by the prior shape knowledge and the label function which dynamically indicates the region with which the prior shape should be compared. The building extraction is addressed through a level set image segmentation approach that involves the use of the label function as well as the prior shape knowledge. In addition, the proposed model permits translation, scaling, and rotation of the prior shape. Experimental results and the qualitative and quantitative evaluations demonstrate the potential of the approach.

Key words Level sets, prior shape knowledge, label function, building detection, segmentation, variational method

随着高分辨率卫星及航空遥感影像的不断涌现,从高分辨率遥感影像中自动检测或提取人造目标及建筑物的研究引起了人们的广泛关注.在已有的各种提取方法中,曲线演化技术(包括 Snake 方法、活动轮廓线方法、变形模型以及最近的水平集方法)显示出了其特有的特质,并有着相当好的提取结果.其中,水平集方法相比其他自动提取方法最大的优点在于其处理拓扑变化的能力.同时,水平集方法很自然地将边界信息和区域信息整合在了一起^[1-4].传统的水平集方法曾经被用到卫星影像中的道路检测(半自动的检测)^[5-7]上.然而,上述文献中提出的各种方法都是单纯的基于图像底层信息(如灰度、颜色、亮度)的,所以很容易受到诸如阴影和遮挡等噪声的影响和干扰,造成检测失误.而阴影和目标被遮挡的情况在遥感数据中却是一种很常见的噪声数据.

为了克服上述方法的限制,在水平集方法中引入与感兴趣目标的几何属性相关的先验知识,对曲线的演化加以限制,不失为一种可行的思路.文献[8]提出一种使用多区域竞争算法的水平集图像分

割方法,利用不同图像区域的几何属性及各自特点,在多区域竞争算法的支撑下,增强了水平集方法进行分割的能力.文献[9]对整合了颜色知识、纹理知识、形状知识等统计先验知识的水平集方法进行了综述,同时利用贝叶斯方法引入了先验形状来指导目标的检测.文献[10]中引入先验形状,同时使用标记函数来指导先验形状所起的约束作用,但是该方法不具备先验形状的平移、旋转、缩放不变性.综上所述,变分方法中的水平集方法在加入先验知识用于目标识别的时候,对于阴影、噪声、背景杂波以及图像中目标形状的损失有很强的鲁棒性^[11].

本文主要研究一种从遥感影像中进行建筑物检测的方法.具体来说,本文提出方法的主要目的在于克服现有的图像底层数据驱动的分割方法中阴影和遮蔽等底层信息干扰导致的分割失误.本文提出的先验形状约束的变分方法,可以从单幅遥感影像中勾勒出建筑物的边缘,从而完成图像中建筑物的提取.这种无需其他数据或图像的辅助,从单幅图像中进行建筑物提取的方法在国民经济以及军事领域是一种亟需的技术,在地理信息系统中也有着非常好的应用前景^[12-14].

同时,本文提出的基于先验形状的建筑物提取方法在目标形状和位置不定时,依然可以有效地进

收稿日期 2009-11-02 录用日期 2010-03-30
Manuscript received November 2, 2009; accepted March 30, 2010

1. 国防科学技术大学信息系统与管理学院 长沙 410073
1. School of Information System and Management, National University of Defense Technology, Changsha 410073

行多目标检测. 我们构建了一个建筑物先验形状的模板库, 使曲线演化的范围限制在先验形状模板库允许范围之内. 通过构造加入先验形状约束的能量泛函, 使得目标检测、先验形状模型自动选择以及先验形状模型形变参数的确定等一系列的计算, 都在对目标能量泛函最小化这一优化过程中完成. 相应地, 本文还引入了标记函数的概念, 用来控制先验形状, 使其只对特定的图像区域起约束作用. 标记函数在曲线演化过程中也不断发生变化, 根据待检测目标和先验形状模型的数量决定图像中哪些区域受到先验形状能量的约束. 需要特别说明的是, 具体到本文中, 先验形状模型指的就是前面提到的建筑物模板.

1 传统的 CV 水平集分割方法^[15]

近年来, 水平集方法成为图像分割领域中一种常用的技术. 本节主要介绍传统的水平集分割方法. 该方法最早由 Chan 和 Vese 提出, 其思想是通过最大化图像中非连结的同质区域的灰度值来完成分割, 是一种对 Mumford-Shah 方法的简化. 本文提出的方法就是对该方法的一种改进.

Mumford 等在 20 世纪 80 年代提出了一种图像分割的方法^[16], 即 Mumford-Shah 方法. 假设存在一幅图像 $f: \Omega \rightarrow \mathbf{R}^2$, 通过最小化如下泛函:

$$E(u, C) = \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \lambda^2 \frac{1}{2} \int_{\Omega-C} |\nabla u|^2 dx dy + \nu |C| \quad (1)$$

即可得到由曲线 C 划分的曲线 C 内部区域和外部区域, 而 u 是由曲线 C 包围的区域的灰度平均值. 对于 u 的估计和 C 的确定是在对上述能量泛函最小化的过程中同时完成的.

著名的 CV 算法 (即由 Chan 和 Vese 提出的水平集分割方法) 中, 对 Mumfor-Shah 方法进行了简化和明确, 引入了 Heaviside 函数, 我们考虑一种最简单的情况: 一个只有两个灰度区域 u_+ 和 u_- 的图像 f 可通过最小化下面的能量泛函来完成分割:

$$E_{CV}(u_+, u_-, \phi) = \int_{\Omega} (f - u_+)^2 H(\phi) dx dy + \int_{\Omega} (f - u_-)^2 (1 - H(\phi)) dx dy + \nu \int_{\Omega} |\nabla H(\phi)| dx dy \quad (2)$$

其中, $H(\phi)$ 表示 Heaviside 函数:

$$H(\phi) = \begin{cases} 1, & \phi > \varepsilon \\ 0, & \text{其他} \end{cases} \quad (3)$$

利用变分法, 式 (2) 中的泛函的欧拉-拉格朗日最速梯度可表示为

$$\frac{\partial \phi}{\partial t} = \delta_{\varepsilon}(\phi) \left[\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (f - u_+)^2 + (f - u_-)^2 \right] \quad (4)$$

其中, u_+ 和 u_- 为两个随着水平集演化不断更新的标量, 分别代表了由 ϕ 的零水平集分割出的图像 ϕ 中 $\phi \geq 0$ 和 $\phi < 0$ 的不同区域的灰度均值, Chan 和 Vese 在文献 [17] 中推导了这两个标量的计算方法:

$$u_+ = \frac{\int f H(\phi) dx dy}{\int H(\phi) dx dy} \\ u_- = \frac{\int f (1 - H(\phi)) dx dy}{\int (1 - H(\phi)) dx dy} \quad (5)$$

式中的 $\delta_{\varepsilon}(\phi)$ 为数值化的狄拉克函数, 即 $\delta_{\varepsilon}(\phi) = H'_{\varepsilon}(\phi)$ 是规则化后的 Heaviside 函数, 在计算中一般取如下的形式:

$$H_{\varepsilon}(\phi) = \begin{cases} 1, & \phi > \varepsilon \\ 0, & \phi < -\varepsilon \\ \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{\phi}{\varepsilon} \right) \right), & |\phi| \leq \varepsilon \end{cases} \quad (6)$$

$$\delta_{\varepsilon}(\phi) = \begin{cases} 0, & |\phi| > \varepsilon \\ \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + \phi^2}, & |\phi| \leq \varepsilon \end{cases} \quad (7)$$

从式 (2) 可以看出, 用传统的水平集方法对图像进行分割时, 能量泛函包括了三个主要部分: 1) 一个限制演化曲线长度的平滑项: $\int_{\Omega} |\nabla H(\phi)| dx dy$; 2) 整个图像中的全局目标信息: $\int_{\Omega} r_{\text{obj}}(I) H(\phi) dx dy$; 3) 背景信息: $\int_{\Omega} r_{\text{bg}}(I) (1 - H(\phi)) dx dy$.

上述式子中, r_{obj} 和 r_{bg} 分别用来表示图像中目标区域和背景区域的模型, 可以称之为目标或者背景的描述算子. 针对不同用途或者不同对象的分割, 可使用不同的方法来构建 r_{obj} 和 r_{bg} , 但 r_{obj} 和 r_{bg} 必须能够表征某个区域的特征或者属性, 且为单调递减函数, 这样才能保证能量泛函的最终收敛. 实际应用中, 这样的描述算子主要依据要处理对象的某些自然属性来构建, 如在文献 [17] 中, 作者使用高斯密度函数对目标区域进行建模; 在文献 [18] 中, 作者提出使用更为广义的高斯混合模型对图像中目标和背景的灰度分布进行描述. 同样的, 若在一幅图像中可利用的强度信息不仅只是灰度信息, 而是诸如颜色、亮度甚至雷达或者超光谱等信息都可利用时, 也可构建对应的甚至混合的描述算子来对目标和背景

进行描述. 但是如何对不同类型的图像数据构建和使用不同的描述算子不是本文讨论的重点, 所以本文参考文献 [17] 的方法, 仅使用一种相对简单的情况对背景和目标进行描述, 目的在于更加突出本文的算法.

图 1 所示即为传统的变分水平集方法对遥感影像中的建筑进行检测的结果. 在实验中, 我们与文献 [4, 17] 采取了类似的假设, 使用高斯模型对背景和

$$r_{\text{obj}}(I) = \frac{(\mu_{\text{obj}} - I)^2}{\sigma_{\text{obj}}^2} \quad (8)$$

$$r_{\text{bg}}(I) = \frac{(\mu_{\text{bg}} - I)^2}{\sigma_{\text{bg}}^2} \quad (9)$$

其中, μ_{obj} , μ_{bg} 分别为目标和背景区域的灰度均值, 而 σ_{obj} , σ_{bg} 则分别为相应的协方差.

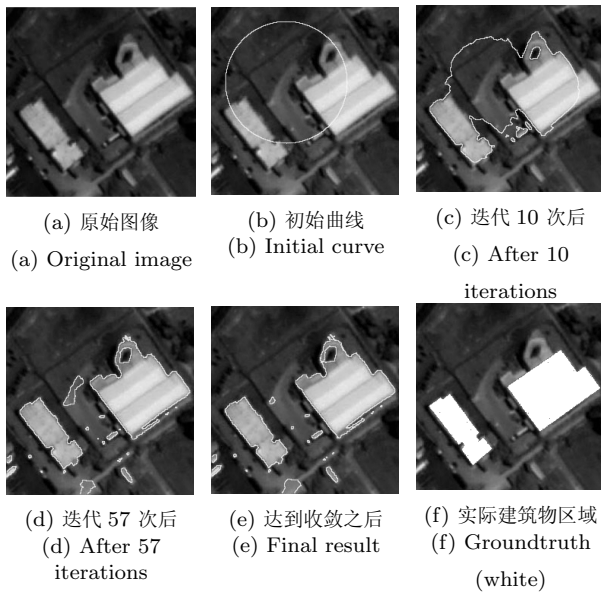


图 1 传统的变分水平集方法的检测结果

Fig. 1 The building detection results of CV method

从图 1 的结果以及式 (8) 和 (9) 中可以看出, 当使用传统的基于图像属性的水平集方法进行建筑物检测时, 在目标或背景区域具有较好的同质性时, 检测结果较好. 但当区域的内同质性被破坏时, 上述模型不能得到令人满意的结果. 而在遥感图像中, 由于光照、阴影、遮挡、杂波等因素的影响, 很难保证区域的一致性. 在图 1 中, 左下角的建筑物和右边建筑物的右下角及左半部分被较好地检测出来, 但是其顶部却出现了检测错误, 这是由于图中的停车场和建筑物屋顶有着相似的反射性质, 同时停车场有较平滑的灰度值, 单纯使用灰度信息不能够准确地描述感兴趣目标.

为了解决上述由于底层信息缺失或描述困难造成的检测问题, 本文引入全局的先验形状模型来指导和限制水平集方法中能量的最小化, 以期达到较为准确的检测结果.

2 引入单一先验形状约束的方法

2.1 先验形状数据库的构建

从图 1 及前述的内容可知, 变分水平集方法确实能够很好地处理图像中目标的拓扑变化, 同时能够较好地根据目标的区域属性进行分割, 但是当受到诸如光照条件的变化、背景杂波、阴影、遮挡等因素的影响, 造成底层信息的缺失或不准确时, 传统水平集方法很难达到令人满意的收敛结果. 一个比较自然的想法, 就是对水平集方法中的能量函数进行改造, 可以考虑引入一些先验的形状知识. 而最直接的做法就是在能量函数中加入一个或多个“形状能量项”来约束曲线的演化. 即这些加入的约束项都是一些特定形状的能量表示:

$$E_{\text{total}}(\phi) = E_{\text{CV}}(r_{\text{obj}}, r_{\text{bg}}, \phi) + \alpha E_{\text{prior}}(\phi), \quad \alpha > 0 \quad (10)$$

式中, E_{total} 为加入先验形状约束后的能量函数, E_{CV} 是基于 Chan 和 Vese 提出的水平集方法的能量项, E_{prior} 即为先验形状能量项. E_{prior} 会在整个图像定义域 (整幅图像中) 内对演化曲面 ϕ 进行约束, 从而控制其零水平集, 也就是曲线的演化. 针对最简单的情况, 即先验形状不进行缩放、平移、旋转等变化时, 文献 [10] 提出先验能量项可采取如下的形式:

$$E_{\text{prior}}(\phi) = \int_{\Omega} (\phi - \phi_0)^2 dx dy \quad (11)$$

其中, ϕ_0 是表示已知形状的水平集函数. 在本文中, 为了计算方便并与前述内容保持一致, 对于先验能量项采取如下的描述形式:

$$E_{\text{prior}}(\phi) = \int_{\Omega} (H(\phi) - H(\phi_0))^2 dx dy \quad (12)$$

其中, $H(x)$ 和前文中一样, 是 Heaviside 函数, 它同时作用于 ϕ 和 ϕ_0 , 独立于图像定义域 Ω . 有了对先验形状能量项的定义, 针对本文要讨论的对象, 即建筑物提取, 我们设计了如下的一些“模板”, 如图 2 所示. 因为建筑物均为人造对象, 普遍具有一定的规则的形状, 同时一般具有较为尖锐的棱角, 故而在遥感影像中主要呈现为正方形、长方形、圆形或这些形状组合的匀质区域 (如图 2 (d) 中表示的是由一个矩形和两个半圆组成的形状, 用来检测类似操场的目标).

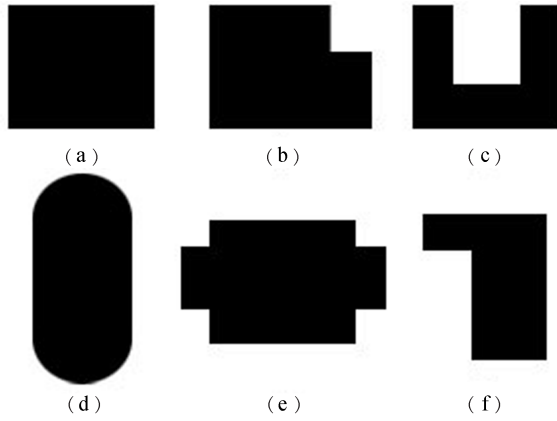


图2 先验形状模板

Fig.2 Prior templates

单纯使用上述的模板在进行建筑物提取时, 由于没有考虑任何目标形变或位置变化等因素, 所以只能对形状和位置与模板严格一致的区域进行检测和提取. 然而在实际的遥感图像建筑物提取问题中, 由于拍摄角度、光照等拍摄因素的不同, 会造成建筑物的朝向和位置是不确定的. 文献 [19] 中作者提出使用先验形状的统计模型来解决建筑物的多样性问题. 但同样的问题在于, 当且仅当对模型的假设与实际数据相符时, 才能得到满意的检测结果, 实际却很难对所有情况都构建对应的统计模型, 这需要耗费大量的人力物力. 甚至可以说, 实际的建筑物提取中, 单纯使用一种简单的、甚至是复杂的统计模型对纷繁复杂的遥感图像中的目标进行建模无疑是杯水车薪, 所以必须同时考虑先验形状的变化.

2.2 先验形状的表达及变换

在文献 [20–21] 中, Paragios 等使用符号距离函数对形状进行表示, 这种符号距离函数实际上是一种特殊的水平集函数: 给定一个封闭且有边界的目标 (或目标区域) $\Omega \rightarrow \mathbf{R}^2$, 其边界为 C , 水平集函数可由下面的关系表示:

$$\phi(x, C) \begin{cases} 0, & x \in C \\ +D(x, C) > 0, & x \in \Omega \setminus C \\ -D(x, C) < 0, & x \in \mathbf{R}^2 \setminus \Omega \end{cases} \quad (13)$$

其中, $D(x, C)$ 是图像中像素 x 到边界 C 的距离. 从而, 图像中的任意目标都可由一个符号距离函数来表示, 反之亦然.

考虑形状发生变化时的几种可能: 缩放、平移、旋转. 可以通过定义一种等价的关系来表示形状发生的变化. 本文中我们考虑使用如下的变化关系: 假设 Ω_1, Ω_2 为图像中两个相关联的形状, ϕ_1, ϕ_2 分别为相应的符号距离函数 (水平集函数), 那么存在一

种四元的对应关系:

$$\phi_2(x, y) = r\phi_1 \left[\frac{(x-a)\cos\theta + (y-b)\sin\theta}{r}, \frac{-(x-a)\sin\theta + (y-b)\cos\theta}{r} \right] \quad (14)$$

式中, (a, b) 表示形状中心, r 表示缩放参数, θ 表示旋转角度. 这样, 给定任意一个形状及其对应的符号距离函数, 可以通过上述的四元关系 (a, b, r, θ) 得到与其相关的另外一个形状的符号距离函数, 即其水平集函数. 由此, 根据式 (12) 和 (14), 我们可以给出如下定义:

令 $f: \Omega \rightarrow \mathbf{R}^2$ 为定义在 Ω 上的图像, ϕ 为目标水平集函数, 也就是对图像进行分割, 或对建筑物进行提取的曲线的水平集表示, 而 ψ_0 为模板数据库中某一已知模板的符号距离函数表示. ψ 为 ψ_0 经式 (14) 变换后, 用于对能量泛函 (10) 进行约束的符号距离函数, 即是与图像中某一目标或目标区域对应的先验形状的水平集表示. 于是, 式 (12) 可以写为如下的形式:

$$E_{\text{prior}}(\phi, \psi) = \int_{\Omega} (H(\phi) - H(\psi))^2 dx \quad (15)$$

在完成先验形状的表达并确定其变化方式后, 我们就可以使用先验形状来约束水平集方法中曲线的演化, 由于先验形状的符号距离函数 (水平集函数) 是已知的, 利用式 (14) 得到相应的能量约束项带入式 (12), 即可确定先验形状能量项的表示. 注意到水平集函数 ϕ 的求解, 四元关系 (a, b, r, θ) 的确定, 以及整个建筑物的检测过程都是在对 $E_{\text{prior}}(\phi)$ 能量最小化的过程中同时完成的, 即为了最小化式 (10) 的能量泛函, 首先由形状模板中的某一先验形状, 通过式 (14) 的变换得到先验形状水平集函数 ϕ_0 (此时的四元关系参数的具体值尚未确定), 随后在对水平集函数 ϕ 不断演化 (即使能量泛函最小化) 的过程中, 完成对区域描述模型 r_{obj} 和 r_{bg} 及四元关系 (a, b, r, θ) 的参数估计.

2.3 数值化计算

在使用水平集方法进行图像分割时, 需要对如式 (10) 的能量泛函进行最小化, 一般使用变分法. 我们依然采用 Chan 和 Vese 提出的简化的 Mumford-Shah 方法的思路, 根据其论述, 在计算中可以将长度项省略, 于是上文中的式 (10) 可重写为如下形式:

$$E_{\text{total}}(\phi) = E_{\text{CV}}(r_{\text{obj}}, r_{\text{bg}}, \phi) + \alpha E_{\text{prior}}(\phi) = \int_{\Omega} ((f - u_+)^2 H(\phi) +$$

$$(f - u_-)^2(1 - H(\phi)) \, dx dy + \alpha \int_{\Omega} (H(\phi) - H(\psi))^2 \, dx dy \quad (16)$$

对上述泛函求变分, 可得到 u_+ 和 u_- 依然与式 (5) 中有相同的形式, 而对于先验形状的水平集函数 ϕ_0 , 其四元组的最速梯度下降方向可由下式得到:

$$\frac{\partial a}{\partial t} = \int_{\Omega} (H(\psi) - H(\phi)) [\psi_{0x}(x^*, y^*) \cos \theta - \psi_{0y}(x^*, y^*) \sin \theta] \delta(\psi) \, dx dy \quad (17)$$

$$\frac{\partial b}{\partial t} = \int_{\Omega} (H(\psi) - H(\phi)) [\psi_{0x}(x^*, y^*) \sin \theta - \psi_{0y}(x^*, y^*) \cos \theta] \delta(\psi) \, dx dy \quad (18)$$

$$\frac{\partial r}{\partial t} = \int_{\Omega} (H(\psi) - H(\phi)) [-\psi_{0x}(x^*, y^*) + \psi_{0x}(x^*, y^*)x^* + \psi_{0y}(x^*, y^*)y^*] \delta(\psi) \, dx dy \quad (19)$$

$$\frac{\partial \theta}{\partial t} = \int_{\Omega} (H(\psi) - H(\phi)) [-r\psi_{0x}(x^*, y^*)y^* + r\psi_{0y}(x^*, y^*)x^*] \delta(\psi) \, dx dy \quad (20)$$

式中, x^*, y^* 是式 (14) 中两个元素值的简写, 而 ψ 和 ψ_0 与上一节中提到的是一致的, ψ 是 ψ_0 经式 (14) 变换后将要用于对能量泛函 (10) 进行约束的符号距离函数:

$$x^* = \frac{(x - a) \cos \theta + (y - b) \sin \theta}{r} \quad (21)$$

$$y^* = \frac{-(x - a) \sin \theta + (y - b) \cos \theta}{r} \quad (22)$$

$$\psi_{0x} = \frac{\partial \psi_0}{\partial x}, \quad \psi_{0y} = \frac{\partial \psi_0}{\partial y} \quad (23)$$

通过上述方程, 对于水平集函数 ϕ , 其 E-L 方程为

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_{\text{total}}}{\partial \phi} = -\delta(\phi) [(f - u_+)^2 - (f - u_-)^2 + 2\alpha(H(\phi) - H(\psi))] \quad (24)$$

式中, $H(\phi)$ 及 $(\delta(\phi))$ 计算时使用与式 (6) 和 (7) 相同的形式.

加入先验约束的水平集方法可以很好地解决由于底层信息缺失造成的图像分割错误, 并且可以成功地提取出建筑物的轮廓. 如图 3 所示, 图 3(e) 的结果中, 图 2(b) 所示的模板被恢复出来, 且对应的建筑物边缘被很好地提取出来.

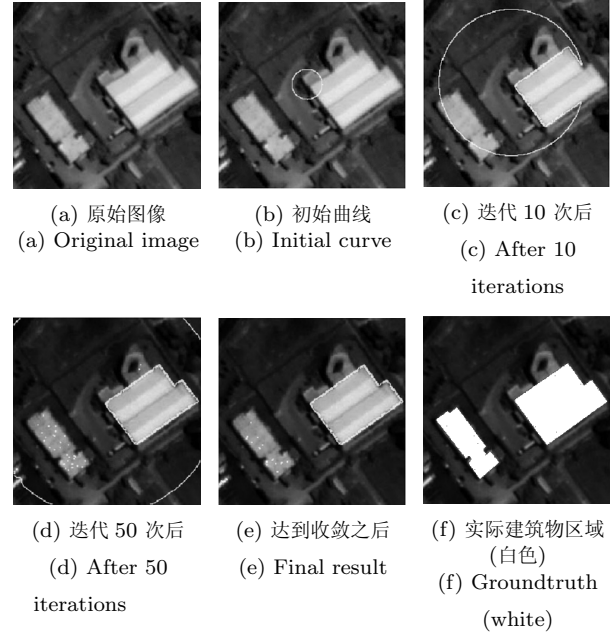


图 3 加入先验形状约束的水平集方法检测结果
Fig. 3 Building detection results with one prior template

但是, 同时可以看出, 如式 (16) 形式的先验水平集方法并不能同时对图像中的多个建筑物进行检测. 依然如图 3 所示, 图像中左下角的建筑物并没有被检测出来, 也就是说, 在这种方法中, 只有一个先验形状模板起作用, 其他形状都被压制了, 这使得水平集方法失去了其能够同时处理多个目标的能力. 所以, 在实际应用中还需要对上述方法进行进一步的改进.

3 多先验形状竞争的建筑物检测

在加入单个先验形状约束后, 传统的水平集方法失去了其能够同时处理多个目标的能力. 因此, 我们引入了一个标记函数来对需要使用的先验形状进行增强或者抑制.

3.1 标记函数的引入

设标记函数 $L: \Omega \rightarrow \mathbf{R}^2$ 是定义在图像定义域上的二值函数, 作用在于增强图像平面中先验形状的作用, 使先验形状不影响非作用区域的曲线演化. 其取值根据图像中的某一位置先验形状的约束作用是否应该被增强而取 +1 和 -1. 引入标记函数后的先验能量项被改写为如下的形式:

$$E_{\text{prior}}(\phi, \psi, L) = \int_{\Omega} (H(\phi) - H(\psi))^2 (L + 1) \, dx + \lambda^2 \int_{\Omega} (L - 1)^2 \, dx + \rho \int_{\Omega} |\nabla H(L)| \, dx \quad (25)$$

式中, λ, ρ 为两个大于零的参数. 对于要检测的目标, 令 ϕ 不变, 最小化式 (25) 的前两项, 将使标记函数 L 有如下的性质:

$$\begin{aligned} L &\rightarrow +1, \quad \text{若 } |H(\phi) - H(\psi)| < \lambda \\ L &\rightarrow -1, \quad \text{若 } |H(\phi) - H(\psi)| > \lambda \end{aligned} \quad (26)$$

可以看出, λ 是一个类似阈值的参数, 其取值决定了标记函数所起的作用, 实际使用时一般取经验值. 当 λ 取值过大时, 可能会引入一些与要检测的目标不相关的先验形状, 弱化了标记函数的指导作用; λ 取值过小, 又会使 L 稳定在一个比实际需要小得多的状态, 使得先验形状的选择过于严格, 造成了“过约束”. 很明显 L 在能量式中起到的作用是增强图像中与要检测的区域类似的先验形状, 也就是使 L 在 $(H(\phi) - H(\psi))$ 小于阈值 λ 时, 取值为 1. 在式 (25) 中, 后一项的作用主要是保持曲线的平滑及保证曲线在演化时尽量保持最小的长度. ρ 是一个权重参数, 一般根据经验选择, 但在近似的数值计算中该项一般可省略. 将式 (25) 所示的先验能量带入式 (10) 的能量泛函, 并最小化能量泛函, 其结果如图 4 所示. 从图中可见, 不但先验形状起到了约束作用, 将图中右半部分的建筑物精确地通过先验形状模板恢复了出来, 而且与图 3 相比, 图中其他部分的建筑物也被检测了出来 (这时先验模板并未起到约束作用), 在引入标记函数后, 水平集方法重新具有了同时检测多个目标的能力. 也就是说标记函数在能量最小化的过程中选择了先验形状应该起作用的区域, 并加强了该先验形状的约束作用, 同时没有影响到未与先验形状匹配的其他未知的目标的水平集检测.

然而, 引入一个标记函数后, 虽然加强了先验模板所起的作用, 同时也能够使先验模板以外的区域能够不受影响, 但是这种方法依然存在很明显的限制: 不能同时利用多个先验模板对图像中的目标进行检测, 也就是说, 一次只能使用一个先验形状模型. 于是在式 (25) 的基础上, 我们可以继续做如下的改进, 使得能量泛函中可以引入两个不同的先验形状 ϕ_1, ϕ_2 :

$$\begin{aligned} E_{\text{prior}}(\phi, \psi, L) = & \frac{1}{\sigma_1^2} \int_{\Omega} (H(\phi) - H(\psi_1))^2 \times \\ & (L + 1)dx + \frac{1}{\sigma_2^2} \int_{\Omega} (H(\phi) - H(\psi_2))^2 \times \\ & (L - 1)dx + \rho \int_{\Omega} |\nabla H(L)|dx \end{aligned} \quad (27)$$

式中的 σ_1^2 和 σ_2^2 与文献 [22] 中证明的一样, 是两个先验形状模板的方差, 且存在如下的关系: $\sigma_i^2 = \int \phi_i^2 dx - (\int \phi_i dx)^2$.

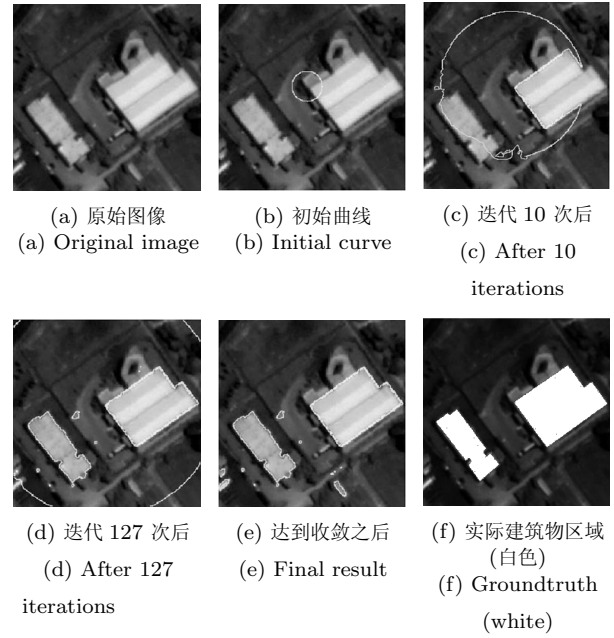


图 4 单标记函数指导的先验形状约束的水平集方法检测结果

Fig. 4 Building detection results with one prior template and label function

3.2 一般情况

这一节中, 我们考虑更为普遍的情况: 同时使用多个先验模板来进行检测. 假设待检测图像中, 部分目标可以从图 2 所示的先验形状模板库中找到相应的模板与之匹配, 同时图像中还存在一些不能与模板库中任一模板匹配的目标, 即我们在使用先验形状模板进行目标提取的同时, 依然需要使用图像中的灰度等强度信息检测出无法与先验形状模板匹配的目标. 为此, 我们引入如下的标记函数 L :

$$L: \Omega \rightarrow \mathbf{R}^k, \quad L(x) = (L_1(x), \dots, L_k(x)) \quad (28)$$

则 $m = 2^k$ 个顶点的多面体 $[-1, +1]^k$, 可以表示 m 个不同的标记函数 $L_j \in (-1, +1)$, 也即 m 个目标区域. 每一个区域由一个指示函数 $x_i = 1, \dots, m$ 表示. 而指示函数 x_i 有如下的形式 [22]:

$$x_i(L) = \frac{1}{4^k} \prod_{j=1}^k (L_j - \omega_j)^2, \quad \omega_j \in (-1, +1) \quad (29)$$

引入上述的 k 维标记函数, 现在可以使用 $m = 2^k$ 个先验形状模板来完成图像中多个目标的检测. 相应的, 如式 (27) 所示的引入标记函数的先验能量项将有如下的形式 [22]:

$$E_{\text{prior}}(\phi, \psi, L) = \sum_{i=1}^{m-1} \int \left(\frac{H(\phi) - H(\psi_i)}{\sigma_i} \right)^2 \times$$

$$x_i(L)dx + \int \lambda^2 x_m(L)dx \times \rho \sum_{i=1}^{m-1} \int |\nabla H(L)|dx \quad (30)$$

将式 (30) 所示的先验能量项带入式 (10), 同样使用梯度下降法对水平集函数 ϕ , 四元组变换参数, 以及区域描述算子在迭代过程中进行更新和估计, 最小化能量泛函, 可以得到水平集函数的演化方程:

$$\frac{\partial \phi}{\partial t} = \frac{\partial E_{\text{total}}}{\partial \phi} = -\frac{\partial E_{CV}}{\partial \phi} - 2\alpha \sum_{i=1}^{m-1} \frac{H(\phi) - H(\psi_i)}{\sigma_i^2} \times x_i(L) \quad (31)$$

令水平集函数 ϕ 和变换参数不变, 标记函数的最速下降梯度 L_j 的演化方程有如下的形式:

$$\frac{L_j}{\partial t} = \frac{\partial E_{\text{total}}}{\partial L_j} = -\alpha \sum_{i=1}^{m-1} \frac{(H(\phi) - H(\psi_i))^2}{\sigma_i^2} \frac{\partial x_i}{\partial L_j} \times \alpha \lambda^2 \frac{\partial x_m}{\partial L_j} - \alpha \rho \nabla \left(\frac{\nabla H(L_j)}{|\nabla H(L_j)|} \right) \quad (32)$$

上式中的前两项使得标记函数增强了图像中与某目标经过式 (14) 变换后的先验形状最相似的区域. 最后一项则保证了标记函数的空间规则性, 并且防止多个先验形状的互相重叠, 也就是使标记函数在演化时具有光滑的最小长度的曲线形式. 而指示函数 x_i 则由式 (29) 得到. 如在 $k = 2$ 时, 标记函数可以使用四个先验形状, x_i 有如下的形式:

$$\begin{aligned} x_1(L) &= \frac{1}{16}(L_1 - 1)^2(L_2 - 1)^2 \\ x_2(L) &= \frac{1}{16}(L_1 + 1)^2(L_2 - 1)^2 \\ x_3(L) &= \frac{1}{16}(L_1 - 1)^2(L_2 + 1)^2 \\ x_4(L) &= \frac{1}{16}(L_1 + 1)^2(L_2 + 1)^2 \end{aligned} \quad (33)$$

4 实验与讨论

前文中已经讨论过传统的单纯数据驱动水平集方法 (单纯依靠图像中强度信息进行目标检测) 的不足, 同时指出了加入单一先验形状约束的水平集模型用于目标提取时的优点及其局限性, 随后讨论了加入标记函数后的改进算法. 本节主要通过实验来说明本文提出模型的有效性. 如前所述, 本文较传统水平集方法的改进主要有以下三点: 1) 可以不需要

利用多种外界辅助信息 (如阴影、高程信息、地理信息系统信息、光谱信息等), 充分利用构建的建筑物模板作为先验信息, 结合曲线演化方法自身的特点, 从单幅遥感影像中进行建筑物的提取; 2) 对遥感数据中的阴影、杂波、遮蔽等传统方法难以回避的噪声数据不敏感, 较好地克服了噪声的影响; 3) 变分割方法相比其他传统的分割方法具有较好的处理拓扑变换的优势, 加入先验形状约束后, 不但可以很好地解决遥感图像中出现的目标被遮挡、边缘丢失等问题, 而且由于标记函数的引入, 可以根据要检测目标的数量, 结合先验形状模板, 较好地完成多个目标的提取.

针对上述三个改进, 我们模拟了若干种情况进行实验: 1) 在图像中加入遥感影像常见的噪声, 考量本文方法克服噪声的能力; 2) 模拟目标被阴影遮挡的情况, 造成目标部分边缘缺失, 考量本文方法利用先验形状信息处理拓扑变换的能力; 3) 一幅图像中同时出现多个与先验形状模板类似的目标, 考量本文方法在标记函数的指导下利用多个先验形状同时提取多个目标的能力. 实验环境如下: Intel Core Dual CPU 7350, 2.0 GHz, 2 G DDR3 RAM, 操作系统为 Windows 7-32 bit, 实验平台为 Matlab 2009b.

图 5 所示为对某一遥感影像叠加高斯白噪声 (噪声的均值取 0, 方差为 0.02) 后的实验图. 这里需要说明的是, 为了验证本文方法的有效性, 噪声数据的方差值取的相对比较大. 实际操作中, 当出现噪声污染如此严重的情况时会首先对遥感图像进行去噪处理. 图 5(a) 为加入噪声后的图像. 图 5(b) 和 5(c) 分别是迭代 10 次和达到稳定收敛后的结果. 可以看出, 噪声的加入并没有对本文方法进行目标检测的能力造成很大的影响. 本文方法稳定地利用先验模板库中的两个先验形状 (图 2(a) 和 2(b)) 恢复出了建筑物的边缘, 对比图 5(d) 所示的由传统的 CV 算法的检测结果, 本文方法很好地克服了噪声的影响. 传统的 CV 方法依然将建筑物全部检测出来, 但是对比图 1(e) 的结果, 由于噪声的影响出现了不可预料的虚警, 检测结果并不可信.

图 6(a) 所示为人工加入模拟阴影及遮挡了部分建筑物的图像. 在遥感影像中, 由于阴影、云雾及光照和周围目标的影响, 造成目标被遮挡是一种常见的情况. 从图 6(b) 所示的 10 次迭代结果可以看出, 本文的方法由于先验形状的约束和标记函数的指示, 曲线并不会出现杂乱的演化, 而是边演化边恢复变换后的先验形状, 以吻合图中目标. 图 6(c) 为最终达到稳定收敛后的检测结果, 对比图 6(d) 中采用传统的 CV 水平集方法的结果, 可以看出阴影的遮挡并没有对目标的检测产生大的影响, 目标的轮廓

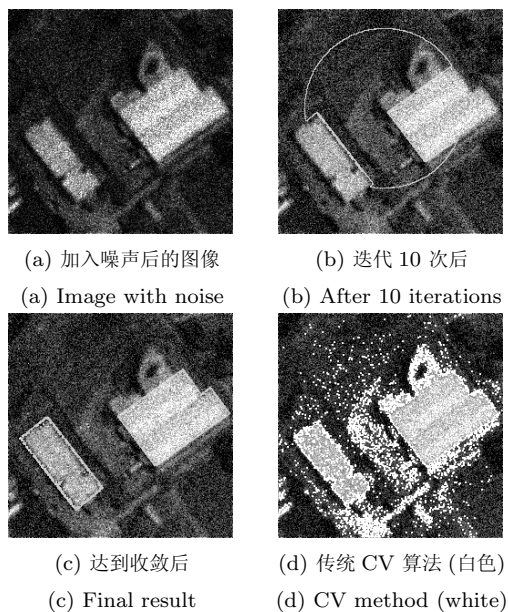


图 5 加入加性噪声后的检测结果
Fig. 5 Building detection results by the proposed model with noise

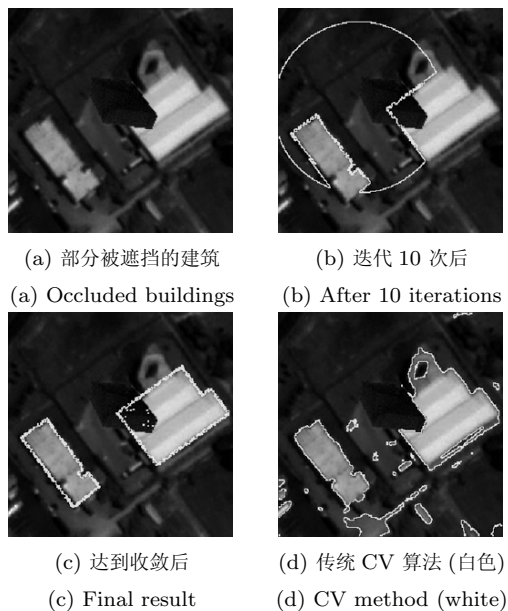


图 6 目标被部分遮挡后的检测结果
Fig. 6 Partial occlusion building detection results by the proposed model

依然被正确地检测出来. 但是从图中可以看出, 阴影部分依然出现了虚警, 这是由于在人工加入模拟的阴影时, 对阴影进行了透明处理, 阴影区域的强度与目标区域发生了交错和重叠, 造成了曲线演化的偏移. 图 6(d) 中传统 CV 水平集方法与图 1(e) 的结果相比, 没有受阴影遮挡的部分目标被检测出来, 但是阴影的遮挡使得图中右半部分的目标的左上角发生了缺损.

为了验证标记函数对同时使用多个先验形状模板的指导作用, 以及本文方法同时精确检测多个目标的能力, 我们截取了一块大小为 $265 \text{ 像素} \times 305 \text{ 像素}$ 的包含多个典型目标的城市遥感影像. 该影像由 IKONOS 卫星 2003 年拍摄于我国唐山地区, 分辨率为 4 m , 如图 7 所示. 其中, 图 7(b) 为人工标识的建筑物区域 (图中带数字部分), 图中共有 11 栋典型的建筑物 (目标), 其中 1 号建筑物为操场, 其余为楼房. 图 7(c) ~ 7(e) 分别为迭代过程到收敛的结果, 可以看出, 对这幅图进行目标检测的过程中, 在 $k = 2, m = 2^k$ 个 (即 4 个) 标记函数的指导下, 图 2 所示的先验形状模板中的四种先验形状 (图 2(a), 2(c), 2(d) 和 2(f)) 被选择出来, 经过式 (14) 表示的关系的变换, 成功地检测出了 11 个目标中的 9 个目标. 而图 7(f) 所示的传统水平集方法, 除了个别目标外, 都没有很好地将目标的轮廓标示出来, 也就是说, 没有很好地完成目标的检测. 而如图 7(b) 中所示的 11 号和 8 号目标, 由于其反射特性与周围环境较为相似, 而与其他目标在图中映应的强度信息有较大的区别, 本文的方法和传统的方法中都被漏检为背景.

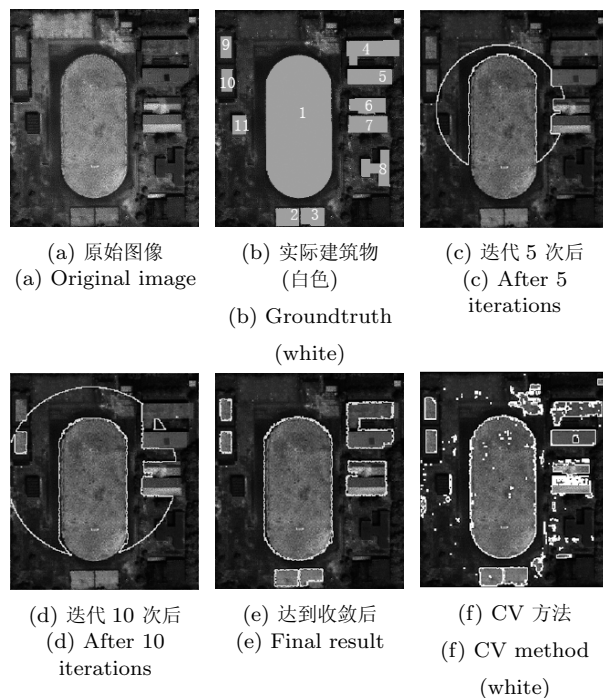


图 7 同时对多个目标进行检测的结果
Fig. 7 Multi-object detection results by the proposed method

图 8 所示的是本文提出的方法和传统的 CV 水平集方法对图 7 检测后的虚警和漏警 (将非目标区域的像素归为目标类别为虚警, 而将目标区域的像素归为背景类别则为漏警) 对比图. 图中白色部分

表示的是两种方法出现虚警的位置, 而灰色部分则是出现漏警的位置, 图 8 (a) 为本文的方法虚警和漏警的比较图, 图 8 (b) 为 CV 方法的对比图. 无需定量分析, 仅从图中对比便可看出, 本文方法的虚警和漏警概率都要低于 CV 方法, 且本文方法的边缘检测的精确性要远高于 CV 方法, 表 1 为两种方法检测效果的量化数据比较.

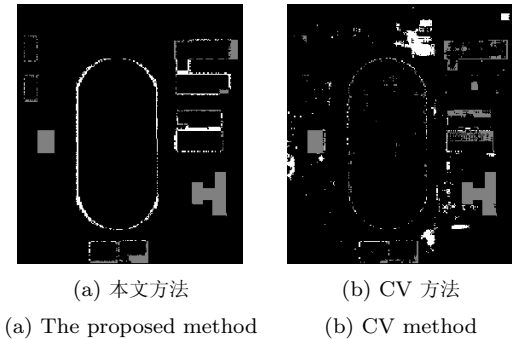


图 8 本文方法和 CV 方法虚警、漏警对比图 (灰色部分为漏警, 白色部分为虚警)

Fig. 8 Comparison of the proposed method and CV method on false alarm and false dismissal where false dismissal is superimposed in gray and false alarm is superimposed in white

表 1 本文方法和 CV 方法检测效果对比 (%)

Table 1 Comparison between the proposed method and CV method (%)

	虚警率	漏警率	误警率 (虚警 + 漏警)
本文方法	6.79	9.16	15.95
CV 方法	7.9	12.64	20.54

从上述的实验结果可以看出, 本文提出的引入先验模型的水平集建筑物提取方法能够较好地完成在一幅图像中对建筑物目标的提取, 且能较好地克服噪声及遮挡的影响, 对建筑物的轮廓还原较为准确, 为后续的各种处理奠定了良好的基础.

需要指出的是, 在本文的实验中, 我们对上述 265 像素 $\times$ 305 像素的图片进行约 300 次迭代计算, 认为达到收敛结果时的耗时约为 11 秒, 而这是在 Matlab 环境中完成的. 在实际应用中, 如果将代码移植为执行效率较高的语言代码, 如 C++ 等, 时间效率将会有非常大的提高. 因此, 对于遥感影像等高分辨率、大范围的建筑物快速检测, 本文的方法依然不失有效性.

为了进一步说明模型的有效性, 最后, 我们选取了一块大尺寸、较为复杂的、建筑密集区域的影像来进行实验, 该图像由 GeoEYE 拍摄于 2009 年 6 月的北京上空, 分辨率为 0.6 m. 我们截取的图像大小为 994 像素 $\times$ 769 像素 (篇幅原因, 文中的图片

进行了缩小), 使用多个先验形状完成提取后的结果如图 9 所示. 可以看出, 影像中大多数的建筑物都被检测了出来, 但是部分边缘存在不准确的现象, 这是因为构建的模板库中的先验形状较少, 不能完全精确地匹配建筑物的轮廓. 标记函数使得建筑物区域得到了加强, 使用了非常类似的先验形状来约束曲线的演化.

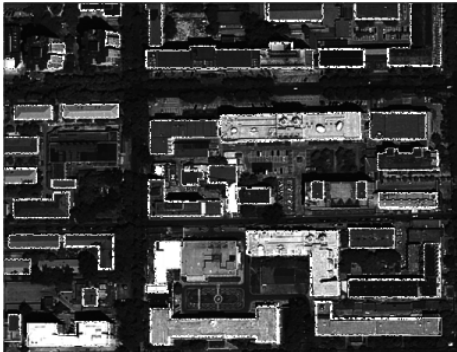


图 9 对北京某城区进行检测的结果图 (白色标注为建筑物边缘)

Fig. 9 Building extraction results of the proposed method from somewhere in Beijing where building edges are superimposed in white

5 小结

本文提出了一种由标记函数指导的、先验形状模型约束的变分水平集遥感图像建筑物提取方法. 本文的方法不但可以在不需要更多辅助信息的情况下从单幅遥感影像中完成多个建筑物的轮廓提取, 而且较好地克服了遥感影像中经常出现的噪声、杂波等噪声污染, 能够在噪声环境下完成建筑物的提取; 同时, 在面对遥感影像中目标被遮挡等信息的缺失情况时, 本文的方法可以很好地处理拓扑变换, 并且利用先验模板的旋转、平移、缩放的变换, 较好地恢复出目标的边缘.

进一步的工作有以下几个方面: 1) 使用执行效率更高的代码实现本文的方法, 获得更为满意的时间效益; 2) 选择更为适合的区域描述算子, 提高曲线演化的效率和检测的准确性. 目前本文主要使用的是图像中的灰度信息, 且假设图像满足高斯模型, 将来可以考虑使用颜色、亮度等强度信息并且使用更精确的概率模型对图像进行描述, 更好地完成检测; 3) 构建更为可信的、普适性更高的先验模板, 可以提高检测的准确性和效率.

References

1 Mayer H. Automatic object extraction from aerial imagery – a survey focusing on buildings. *Computer Vision and Image Understanding*, 1999, **74**(2): 138–149

- 2 Peng J, Zhang D, Liu Y C. An improved snake model for building detection from urban aerial images. *Pattern Recognition Letters*, 2005, **26**(5): 587–595
- 3 Cao G, Yang X, Mao Z Z. A two-stage level set evolution scheme for man-made objects detection in aerial images. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, USA: IEEE, 2005. 474–479
- 4 Karantzalos K, Argialas D. A region-based level set segmentation for automatic detection of man-made objects from aerial and satellite images. *Photogrammetric Engineering and Remote Sensing*, 2009, **75**(6): 667–678
- 5 Samson C, Blanc-Feraud L, Aubert G, Zerubia J. Two variational models for multispectral image classification. In: *Proceedings of the International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Nice, France: Springer, 2001. 344–356
- 6 Ball J E, Bruce L M. Level set segmentation of remotely sensed hyperspectral images. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*. Seoul, Korea: IEEE, 2005. 5638–5642
- 7 Besbes O, Belhadj Z, Boujemaa N. Adaptive Satellite Images Segmentation by Level Set Multiregion Competition, Technical Report 5855, INRIA, Paris, France, 2006
- 8 Mansoni A R, Miticide A, Vazquez C. Multiregion competition: a level set extension of region competition to multiple region image partitioning. *Computer Vision and Image Understanding*, 2006, **101**(3): 137–150
- 9 Cremers D, Rousson M, Deriche R. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 2007, **72**(2): 195–215
- 10 Cremers D, Sochen N, Schnorr C. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In: *Proceedings of the 4th International Conference on Scale Space Methods in Computer Vision*. Isle of Skye, UK: Springer, 2003. 388–400
- 11 Paragios N, Chen Y M, Faugeras O D. *Handbook of Mathematical Models in Computer Vision*. New York: Springer-Verlag, 2005. 145–159
- 12 Gulch E. Digital systems for automated cartographic feature extraction. *International Archives of Photogrammetry and Remote Sensing*, 2000, **33**(B2): 241–256
- 13 Bonnefon R, Dherete P, Desachy J. Geographic information system updating using remote sensing images. *Pattern Recognition Letters*, 2002, **23**(9): 1073–1083
- 14 Doucette P, Agouris P, Stefanidis A. Automation and digital photogrammetric workstations. *Digital Photogrammetric Workstations and Automation: Digital Workstation Architecture and Processing*. Bethesda: ASPRS, 2005. 949–981
- 15 Chan T F, Vese L A. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In: *Proceedings of IEEE Workshop on Variational and Level Set Methods*. Vancouver, Canada: IEEE, 2001. 161–168
- 16 Mumford D, Shah J. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 1989, **42**(5): 577–685
- 17 Chan T F, Vese L A. Active contours without edges. *IEEE Transactions on Image Processing*, 2001, **10**(2): 266–277
- 18 Paragios N, Deriche R. Geodesic active regions: a new framework to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, 2002, **13**(1-2): 249–268
- 19 Riklin-Raviv T, Kiryati N, Sochen N. Prior-based segmentation and shape registration in the presence of perspective distortion. *International Journal of Computer Vision*, 2007, **72**(3): 309–328
- 20 Paragios N, Rousson M, Ramesh V. Matching distance functions: a shape-to-area variational approach for global-to-local registration. In: *Proceedings of the 7th European Conference on Computer Vision*. Copenhagen, Denmark: Springer, 2002. 775–789
- 21 Rousson M, Paragios N. Shape priors for level set representations. In: *Proceedings of the 7th European Conference in Computer Vision*. Copenhagen, Denmark: Springer, 2002. 78–92
- 22 Cremers D, Sochen N, Schnorr C. A multiphase dynamic labeling model for variational recognition-driven image segmentation. *International Journal of Computer Vision*, 2006, **66**(1): 67–81



田昊 国防科学技术大学博士研究生。
2003 年获海军兵种指挥学院学士学位。
主要研究方向为图像处理与模式识别。
本文通信作者。

E-mail: navygz@163.com

(TIAN Hao Ph.D. candidate at National University of Defense Technology. He received his bachelor degree from Naval Academy of Guangzhou in 2003. His research interest covers image processing and pattern recognition. Corresponding author of this paper.)



杨剑 国防科学技术大学硕士研究生。
2007 年获西安交通大学电子与信息工程学院学士学位。主要研究方向为图像处理与目标识别。

E-mail: michael.yangj@sina.com

(YANG Jian Master student at National University of Defense Technology. He received his bachelor degree from Xi'an Jiaotong University in 2007. His research interest covers image processing and auto target recognition (ATR).)



汪彦明 国防科学技术大学博士研究生。
2003 年获国防科学技术大学学士学位。
主要研究方向为计算机视觉。

E-mail: mycroft.won@gmail.com

(WANG Yan-Ming Ph.D. candidate at National University of Defense Technology. He received his bachelor degree from National University of Defense Technology in 2003. His main research interest is computer vision.)



李国辉 国防科学技术大学信息系统与管理学院教授。主要研究方向为多媒体安全。E-mail: guohli@nudt.edu.cn

(LI Guo-Hui Professor at National University of Defense Technology. His main research interest is multimedia security.)