

基于相位场建模的鲁棒框架 在高分辨率卫星图像道路网络提取中的应用¹

彭婷^{1),2)} JERMYN Ian H.²⁾ PRINET Véronique¹⁾ ZERUBIA Josiane²⁾

¹⁾ (中国科学院自动化研究所LIAMA&NLPR 中国北京 100190)

²⁾ (法国国立信息与自动化研究院Ariana项目组 06902 法国Sophia Antipolis)

摘要 最近提出的相位场高阶主动轮廓是基于区域和形状的图像分割的强大的模型。高阶的轮廓表达方式,使其可以对具有复杂形状的区域或物体建模。而且与传统的水平集方法相比,相位场建模避免了轮廓的初始化和距离函数的重新初始化等问题。本文研究的是,高噪声图像中的带状或者线状结构物体(道路、血管等)的提取问题。我们提出一个基于相位场高阶主动轮廓模型的鲁棒框架。首先分割一幅低分辨率图像,即 n 阶Haar小波分解的尺度系数。降低分辨率能减少高频噪声的影响,因此,得到的预分割是目标区域的粗略的完整提取。然后,把这个粗略结果加入模型中,驱使最终的分割结果与预分割结果比较接近。我们将所提出的框架应用于密集城区的高分辨率卫星图像的道路网络提取中。量化评价和与其它方法的比较说明了该方法的有效性。

关键词 相位场建模; 先验知识; 主动轮廓模型; 高分辨率遥感图像; 道路网络分割

¹ 本课题得到LIAMA项目和European Union Network of Excellence MUSCLE (FP6-507752)资助。彭婷,女,1982年生,博士,研究方向为计算机视觉和图像处理。JERMYN Ian H.,男,1964年生,博士,研究员,主要研究兴趣涉及形状建模、纹理建模以及用于形状分析的信息几何学。PRINET Véronique,女,1969年生,博士,副教授,主要研究领域为图像处理、计算机视觉、机器学习、场景理解和遥感图像中的应用。ZERUBIA Josiane,女,1957年生,博士,主任研究员,博士生导师,主要研究兴趣是采用概率模型和变分方法的图像处理,她的工作也与参数估计和优化方法有关。

1 引言

本文主要研究的是,基于水平集主动轮廓模型的、高噪声图像中的带状或者线状结构物体的分割问题。自Kass等人^[1]和Osher等人^[2]先驱性的工作之后,主动轮廓模型在分割问题中,得到了非常广泛的研究。主动轮廓模型定义了轮廓的动态变化过程,即随时间,轮廓由初始位置循环迭代到目标的边缘。于是,需要直接或者间接地定义一个演化方程,来表示每个轮廓点的运动速度。首先,Kass提出的是模型的参数表达,即通过沿轮廓采样的点来表示轮廓。但是这种表达方式存在很多问题:对噪声敏感、不易处理拓扑结构变化等。为了解决上述问题,之后,Osher提出了模型的水平集表达。水平集的主要思想是,将闭合轮廓作为零水平集嵌入到高一维的水平集函数中,这样问题就转化为考虑水平集函数的演化。Chan等人^[3]基于Mumford-Shah模型,提出了基于区域的能量函数,模型驱使每个分割区域具有均一的灰度值。Paragios等人^[4]提出同时基于边界和区域的模型,边界项的作用是吸引轮廓到最大梯度处,并约束边界的平滑度;而区域项是每个区域后验概率的最大似然估计。以上这两种方法的局限在于,只能处理前景和背景灰度值比较均一的图像。

近些年,为了提高模型的抗干扰能力,人们将先验形状信息引入主动轮廓模型。按照先验形状的通用程度,文献中已有的方法可以分为两类:特定形状先验和通用形状先验。大部分方法都是采用特定形状先验^[5,6],即事先给定某个固定的参考形状或者服从某个概率函数的参考形状分布,于是主动轮廓的演化过程,不仅受到图像的影响,而且受到参考形状的约束。相反地,通用形状先验不是仅仅适用于某个特定形状,而是适用于与实际应用背景相关的某类形状。Péteri等人^[7]将边界互相平行的约束加入主动轮廓。Rochery等人^[8]利用二阶主动轮廓描述任意长距离的点对之间的相互作用,从而描述了一类“臂状”网络形状。

本文提出了一个新颖的、基于多分辨率

分析和相位场高阶主动轮廓模型的鲁棒框架,并以北京城区QuickBird高分辨率遥感全色图像(见图1)为例,实现了道路网络的提取。与传统的遥感图像的人工判读和识别地物的方法相比,采用计算机对遥感图像目标进行自动识别和理解,具有花费相对较低、效率高、精度高等诸多优点。这项工作无疑在智能交通导航、城市发展规划、环境评估与监测、突发事件处理等许多领域,都有着广泛的应用前景。

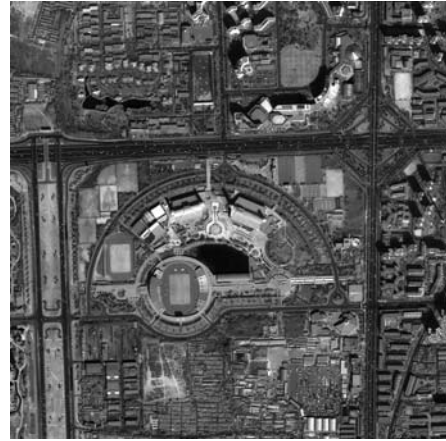


图1 北京城区QuickBird高分辨率遥感全色图像(~0.6m/像素)

本文第2节将详细介绍问题的数学表达和模型各个部分的数学定义。第3节提出了完整的计算框架和优化方法。第4节是模型框架应用于遥感图像的实验结果,以及与其它方法的比较和评价。最后是本文的结论。

2 数学模型

2.1 概述

本小节首先简述图像分割问题的一般Bayes理论框架,然后引入整个模型中的两个关键概念:相位场模型和多分辨率分析。

假设给定一幅图像 $I: \Omega \rightarrow \mathfrak{R}$,其中 Ω 是图像域,希望寻找图像 I 中目标对应的区域 R 。在我们的应用中, R 即为道路网络区域。关于区域 R 的先验知识 K 的提出,例如,区域轮廓几何形状和内部像素的光谱

性质，对问题的解决往往具有非常重要的作用。利用Bayes公式，图像分割可转化为最大化后验概率问题：

$$P(R|I, K) = \frac{P(I|R, K)P(R|K)}{P(I|K)}. \quad (1)$$

或者，从另一个角度说，最小化概率分布的负对数，即能量函数的问题。式(1)中的分母与待估计的数值 R 无关，故在优化过程中，可以忽略。总能量函数表达为：

$$E(R; I) = \omega E_p(R) + E_D(I, R), \quad (2)$$

其中先验能量 E_p 描述了目标区域几何形状

的先验知识，数据能量 E_D 描述了图像中目标和背景的灰度值统计信息， ω 是平衡这两部分能量贡献的权值。

2.1.1 相位场模型

“相位场”是物理学中的一个概念^[9]。假设沉淀物在另一个与之隔绝的基体里扩散，有三种需要考虑的量：沉淀物（Precipitate）、基体（Matrix）和接触面（Interface）。接触面是一个具有一定宽度的活动轮廓，它的运动由相应的力学原理和边界条件决定。在相位场建模中，整个微观结构的状态被表达为一个连续变量，即相位场函数 ϕ ，比如， $\phi = 1$ ， $\phi = -1$ 和 $-1 < \phi < 1$ 分别表示沉淀物、基体和接触面。

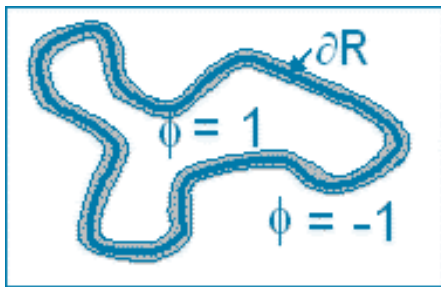


图2 相位场函数 ϕ

相位场模型由Rochery等人^[10]第一次引入图像处理领域，应用于图像分割。相应地，待分割的目标区域、背景区域和边界轮廓附近的窄带，分别对应于上述的物理问题中的沉淀物、基体和接触面。目标区域由相位场

函数 $\phi: \Omega \rightarrow \Re$ 定义为

$R = \{x \in \Omega : \phi(x) > z\}$ ， z 是一个阈值。正如2.2.1节将要说明的，使能量最小化的相位

场函数满足以下条件：当 $x \in R$ 时， $\phi \approx 1$ ；

当 $x \in \bar{R} = \Omega \setminus R$ 时， $\phi \approx -1$ （见图2）。因

此， $\phi_{\pm} = (1 \pm \phi)/2$ 近似是 R 和 $\bar{R}$ 的特征函数。

与传统的水平集方法将水平集函数约束为距离函数相比，相位场函数处于一个线性空间中，模型不依赖初始化，不需要重新初始化，具有更大的拓扑自由度^[10]。如何处理拓扑自由度，无疑是复杂形状建模的难点之一，而相位场模型很自然地解决了这个问题。

2.1.2 多分辨率分析

在整个计算框架中，我们采用了多分辨率分析的方法。多分辨率能量模型的建立，主要考虑到以下三个因素：

- 高分辨率遥感图像包含了不同尺度的地物。为了捕捉这些复杂特征，最直接的想法是在多个分辨率上分析图像。
- 高分辨率图像数据量巨大，对算法速度的要求很高。而多分辨率算法往往可以缩小在高分辨率下算法的搜索空间范围，因而比单个分辨率的算法直接作用于高分辨率数据要快。
- 同一个物体在不同尺度下观测的特性明显不同。比如，在低分辨率图像中，背景被同化为噪声，而道路仍然相对清晰地与背景区分开来，道路分割问题被简化但是提取结果不太准确；而高分辨

率图像提供了更准确的位置和尺寸信息。多分辨率分析的应用可以结合低分辨率数据和高分辨率数据的优势：在低分辨率数据中，大部分可能影响识别过程的细节被去除，可得到粗略的初步结果；在初步结果的约束下，利用高分辨率数据进一步提高分割的精度。

下面将详细介绍相位场能量泛函的各个能量项²。

2.2 先验能量

先验能量 $E_p^{[10]}$ 由两个部分组成：基本的相位场模型 $E_{p,0}$ 和相位场二阶主动轮廓模型 $E_{p,NL}$ 。

传统的主动轮廓^[1,3,4]是图像空间的线性泛函，被表达为轮廓的单重积分，例如，轮廓的周长、轮廓内部的面积。其局限是，由于仅仅描述轮廓点之间非常短距离的局部的相互作用，所以，它只能包含非常有限的关于区域几何的先验知识。高阶主动轮廓^[8]是图像空间的多项式泛函，被表达为轮廓的多重积分。也就是说，它们能描述边缘轮廓点集内任意长距离的点之间的相互作用。通过这种任意长距离的相互作用，高阶主动轮廓能够包含复杂的先验几何约束。因此，与传统的主动轮廓相比，高阶主动轮廓鲁棒性更强，对初始轮廓的位置不敏感。相位场高阶主动轮廓^[10]是用相位场函数表示的高阶主动轮廓。下面将介绍先验能量 E_p 的两个能量项，更多的细节请参考[10]。

2.2.1 $E_{p,0}$

为了利用相位场函数对目标区域建模，基本的相位场模型 $E_{p,0}$ 被定义为^[10]：

$$E_{p,0}(\phi) = \int_{\Omega} dx \left\{ \frac{1}{2} \nabla \phi(x) \cdot \nabla \phi(x) + U(\phi(x)) \right\}. \quad (3)$$

$\nabla \phi(x)$ 是点 x 的梯度矢量，即点 x 的法向量。势函数的定义是：

$$U(y) = \lambda \left(\frac{1}{4} y^4 - \frac{1}{2} y^2 \right) + \alpha \left(y - \frac{1}{3} y^3 \right), \quad (4)$$

其中 λ 和 α 都是常数。当 $\alpha = 0$ 时， $E_{p,0}$ 则为 Ginzburg-Landau 模型。如图3所示，当 $\lambda \geq \alpha > 0$ 时， U 在 $y = \pm 1$ 有两个极小值，在 $y = \alpha / \lambda$ 有一个极大值。

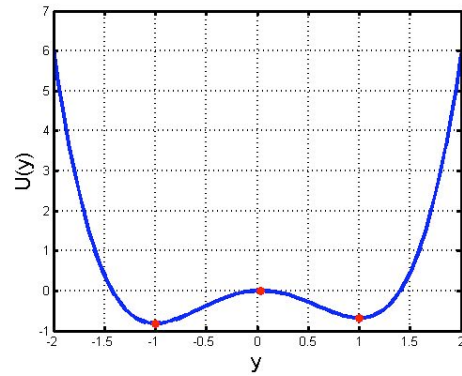


图3 势函数 U ($\lambda = 3, \alpha = 0.1$)

$E_{p,0}$ 的作用等价于传统的线性主动轮廓模型 $\lambda_c L(\partial R) + \alpha_c A(R)$ ，其中 L 是轮廓 ∂R 的周长， A 是轮廓内部面积， λ_c 和 α_c 是常数。如果忽略梯度点乘项，对于某个处处 $\phi = \alpha / \lambda$ 的图像域，最小化能量函数后， ϕ 将在目标 R 内取 1，背景 $\bar{R}$ 内取 -1 。梯度点乘项的作用是惩罚大的梯度变化，使 ϕ 在轮廓附近产生一个从 1 到 -1 的平滑过渡区域。因此， $E_{p,0}$ 能够保证模型的稳定，边界的平滑和前面提到的 R 和 $\bar{R}$ 的特征函数

² 以下的公式中，我们将用 ϕ 代替 R 表示区域。

$\phi_{\pm} = (1 \pm \phi) / 2$ 的性质。

2.2.2 $E_{P,NL}$

为了将复杂的几何形状的先验知识引入模型，需要定义一个包含任意距离 ϕ 值之间相互作用的能量。相位场二阶主动轮廓模型 $E_{P,NL}$ 的定义如下^[10]：

$$E_{P,NL}(\phi) = -\frac{\beta}{2} \iint_{\Omega^2} dx dx' \nabla \phi(x) \cdot \nabla \phi(x') \Psi\left(\frac{|x-x'|}{d}\right), \quad (5)$$

其中 x, x' 是图像中的两点， β 是常数， d 控制相互作用的范围。相互作用函数 Ψ （见图4）是：

$$\Psi(x) = \begin{cases} \frac{1}{2} \left(2 - |x| + \frac{1}{\pi} \sin(\pi |x|) \right) & \text{if } |x| < 2, \\ 0 & \text{else.} \end{cases} \quad (6)$$

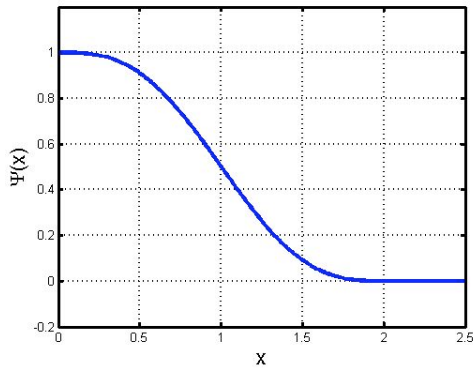


图4 相互作用函数 Ψ

这个相互作用会产生以下两个效果^[10]。一种情况如图5左所示，边界附近的两个点，法向量同方向平行，两点会互相吸引，这适用于道路同一边的点对。另一种情况如图5右所示，边界附近的两个点，法向量反方向平行，两点会互相排斥，并且保持一定的距离 $2d$ ，这适用于道路两边的点对，即 d 与

道路的宽度有关。总之， $E_{P,NL}$ 引入了带状或者线状物体普适的先验几何形状，即具有一定宽度、边缘大致平行、相交于一些连接点的结构。

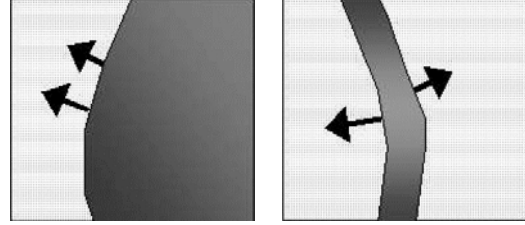


图5 $E_{P,NL}$ 的两种作用效果。黑色区域表示道路，白色区域表示背景，箭头表示道路轮廓点的法向量。

2.3 数据能量

数据能量 E_D 考虑了图像目标和背景的灰度值统计信息。它的定义是式(1)中概率 $P(I | R, K)$ 的负对数。假设目标和背景的概率互相独立，这项则可被分解为 $P(I_R | R, K)P(I_{\bar{R}} | R, K)$ 。数据能量 E_D 由两部分组成，分别对应目标的灰度信息和背景的灰度信息，它的数学表达形式如下：

$$E_D(I, \phi) = - \int_{\Omega} dx \{ \ln P_+(I(x)) \phi_+(x) + \ln P_-(I(x)) \phi_-(x) \}. \quad (7)$$

如前所述， $\phi_{\pm} = (1 \pm \phi) / 2$ 近似是 R 和 $\bar{R}$ 的特征函数，即当 $x \in R$ 时， $\phi_+ = 1, \phi_- = 0$ ；

当 $x \in \bar{R}$ 时， $\phi_+ = 0, \phi_- = 1$ 。 $P_{\pm}$ 是主干道 (+) 和背景 (-) 的图像灰度值 I 的概率密度函数。 $P_{\pm}$ 的定义采用的是监督学习的方法：在分别统计一块目标区域样本和一块背

景区域样本的概率的基础上，假设 $P_{\pm}$ 为混合高斯模型，其中混合高斯模型的系数由EM算法求出。直方图和拟合的概率模型如图6所示。

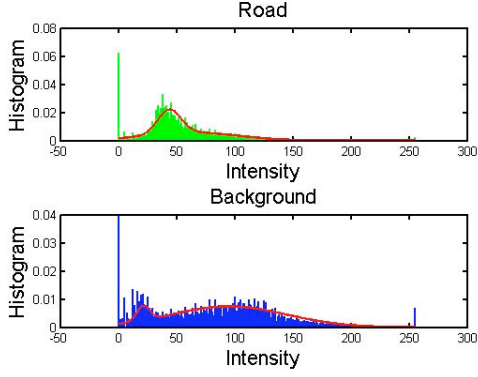


图6 图像灰度值 I 在目标 R (上) 和背景 $\bar{R}$ (下) 的直方图，实线表示的是拟合的混合高斯模型。

3. 实现框架

3.1 流程图

利用Haar小波变换^[11]获得图像的多尺度表达。计算框架中采用的多分辨率分析，主要体现在不同的尺度下，如何定义 $E_{P,NL}$ 的作用区域上。在高尺度下图像的提取过程中，如果用低尺度下的提取结果作为空间约束条件，驱使高尺度下的分割结果处于一个与低尺度下的预分割结果比较接近的空间中，将对整个问题的解决效率和鲁棒性的提高产生非常重要的影响。

算法的具体流程如下：首先，在低尺度下，将能量 $E_{Low} = \omega(E_{P,0} + E_{P,NL}) + E_D$ 作用于低分辨率图像，当算法收敛后，得到低分辨率下粗略的预分割结果。将预分割结果最近邻插值到高分辨率的精度，并用相位场函数表达为 ϕ_{R0} 。然后在高尺度下，将一个

受预分割结果约束的权值

$\phi_{R0+} = (1 + \phi_{R0})/2$ 加入 $E_{P,NL}$ 的积分项，得

到能量 $E_{High} = \omega(E_{P,0} + E_{P,NL,CON}) + E_D$,

其中

$$E_{P,NL,CON}(\phi) = -\frac{\beta}{2} \iint_{\Omega^2} dx dx' \nabla \phi(x) \cdot \nabla \phi(x') \Psi\left(\frac{|x - x'|}{d}\right) \phi_{R0+}(x). \quad (8)$$

这样 E_{High} 就包含了与像素位置有关的空间约束，即二阶主动轮廓引入的带状或者线状的几何先验知识，只在低分辨率提取结果中相应的道路区域起作用。最后将 E_{High} 作用

于高分率图像，当算法收敛后，得到最终的分割结果。算法实现流程图如图7所示。

3.2 能量优化

利用梯度下降法最小化总能量。区别于传统的主动轮廓方法，相位场模型采用“无偏初始化”^[10]，即在整个图像域中，相位场函数的初始值设为常数 α/λ 。因为这个值是势函数 U 的极大值，所以图像中任一点在初始化时，既不倾向目标区域 $\phi \cong 1$ ，也不倾向背景区域 $\phi \cong -1$ ，这就是所谓的“无偏”。整个迭代过程不需要任何重新初始化。

相位场函数 ϕ 的迭代更新步骤如下：

1. 当迭代次数 $n = 0$ 时，

$$\forall x, \phi^0(x) = \alpha/\lambda.$$

2. $\frac{\partial \phi^n(x)}{\partial t} = -\frac{\delta E(\phi^n; I)}{\delta \phi^n}$ ，其中 $\delta E/\delta \phi$

是总能量 E 对相位场函数 ϕ 的偏微分。

3. 更新 ϕ 值：

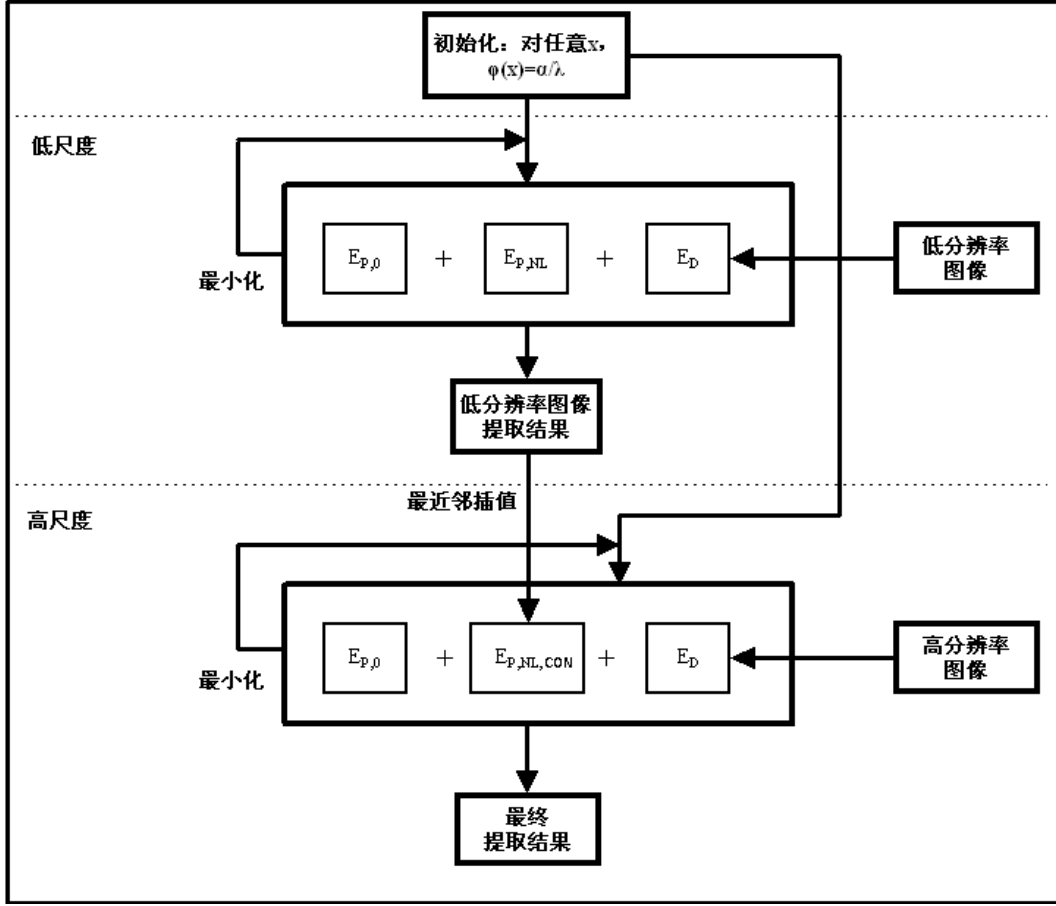


图7 算法实现流程图

$$\phi^{n+1}(x) = \phi^n(x) + \Delta t \frac{\partial \phi^n(x)}{\partial t}, \text{ 其中}$$

Δt 是自适应时间步长；更新迭代次数：
 $n = n + 1$ 。

4. 如果 n 大于最大迭代次数或者 $|\partial \phi^n(x)/\partial t|$ 小于某个给定阈值，迭代结束；否则，回到第2步。

能量 E_{Low} 的演化方程式为：

$$\frac{\partial \phi(x)}{\partial t} = \omega [\nabla^2 \phi - \lambda(\phi^3 - \phi) - \alpha(1 - \phi^2) - \beta \nabla^2 \Psi * \phi] + \frac{1}{2} \ln \left(\frac{P_+(I)}{P_-(I)} \right), \quad (9)$$

其中 $*$ 表示卷积。为了避免计算卷积，我们将演化方程式变换到Fourier域：

$$\frac{\partial \hat{\phi}(k)}{\partial t} = -\omega [k^2 \hat{\phi} + \lambda(\hat{\phi}^3 - \hat{\phi}) + \alpha(1 - \hat{\phi}^2) - \beta k^2 \hat{\Psi} \hat{\phi}] + \frac{1}{2} \ln \left(\frac{P_+(I)}{P_-(I)} \right). \quad (10)$$

实际上，只有 $\omega [\lambda + k^2(\beta \hat{\Psi} - 1)] \hat{\phi}$ 在Fourier域中计算，其余部分仍然在时域中计算^[10]。

类似地，能量 E_{High} 的演化方程式，只需将

式(9)中的 β 代替为 $\beta \phi_{R0+}(x)$ 。

4. 实验结果

输入数据 I 是QuickBird全色图像，如图8左所示，这是典型的北京密集城区场景。图8中、右分别是原分辨率图像（0.6m/像素）和低分辨率图像（经过2次小波变化，即2.4m/像素）的全分辨率局部放大。可见，

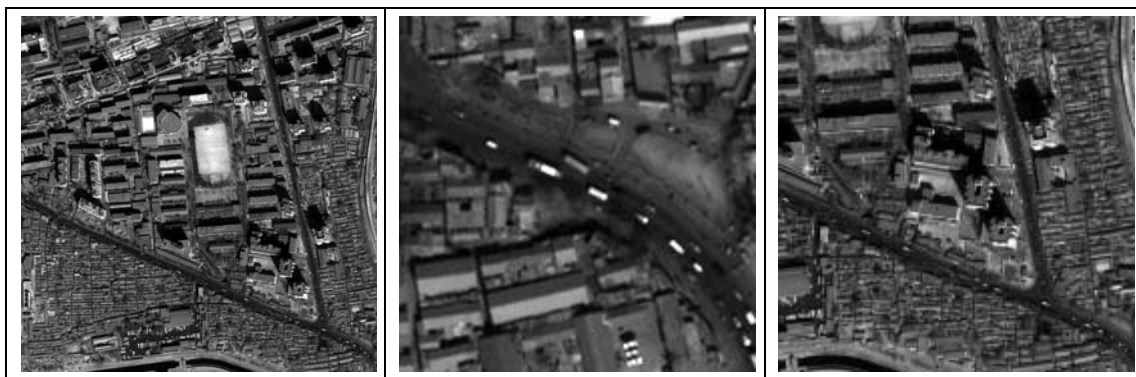


图8 从左至右：QuickBird全色图像（大小：1280×1280 像素）；原分辨率 L_0 下的图像全分辨率放大；低分辨率 L_2 下的图像全分辨率放大。



图9 采用能量 E_{Low} 在低分辨率 L_2 的图像（大小：320×320 像素）上的实验结果。从左到右：分别是二值化的相位场函数，在迭代第1步、第100步和最终收敛第2200步的结果。

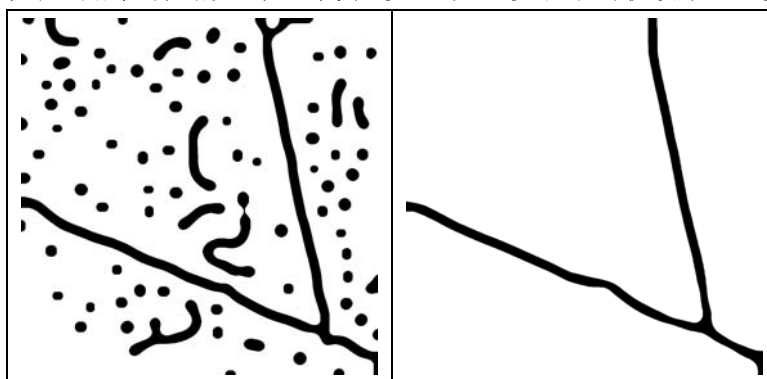


图10 在原分辨率 L_0 的图像（大小：1280×1280 像素）上，从左至右：分别采用能量 E_{Low} 和能量 E_{High} 的分割结果。

在低尺度下，图像中的噪声得到了一定程度的平滑，但是图像仍然比较复杂。以下将原分辨率记为第0层 L_0 ， i 次小波变化后的分辨率记为第 i 层 L_i 。

图9所示是采用能量 E_{Low} 在低分辨率 L_2 图像上迭代第1步、第100步和最终收敛

分辨率	ω	λ	α	β	d
L_2	200	3	0.0905	0.03	20
L_0	2000	3	0.0905	0.03	80

表1 图8中QuickBird图像在两个分辨率下实验的参数设置。注意，除了 d 控制着道路的宽度，以及整个先验能量权值 ω 的变化，其它参数在两个分辨率下设置一致。



图11 在低分辨率 L_2 的图像上，与其它方法的结果对比。从左到右：采用不包含 $E_{P,NL}$ 的 E_{Low} （即 $\beta = 0$ ），Yu^[12]和Wang^[13]的方法分别得到的结果。

指标	E_{Low}	不包含 $E_{P,NL}$ 的 E_{Low}	Yu	Wang
TP/(TP+FN)	0.8700	0.7132	0.9240	0.7701
TP/(TP+FP)	0.6290	0.4741	0.5649	0.4563
TP/(TP+FP+FN)	0.5749	0.3982	0.5398	0.4016

表2 在低分辨率 L_2 的图像上，不同方法的量化评价(T=true, F=false, P=positive, N=negative)。

第2200步的分割结果。虽然无偏初始化加大了问题的难度，但是主干道路仍然被比较完整地提取出来。然而，由于低分辨率的图像信息被一定程度地简化，结果错误地包含了部分停车场，下方道路位置也发生了偏移。

若继续用能量 E_{Low} 作用于原分辨率 L_0 图像，由于图像中噪声和细节的大幅度增加，提取结果并不让人满意（见图10左）。下面将能量 E_{High} 作用于原分辨率 L_0 图像，最终分割结果如图10右所示。多分辨率分析框架的应用，将原分辨率 L_0 中道路提取的搜索空间限制在一个相对较小的子空间中，所以

算法经过较少的迭代就能实现收敛，并且整个道路网络能够完整地、准确地提取出来。在2GHz CPU和1GB内存的电脑上运行MatlabR2007a，得到该结果需要291分钟。虽然这个算法仍然比较耗时，但是如果进行一些算法优化或者用C++语言编写，计算速度将会得到显著的提高。从另外一个角度来说，我们所提出的鲁棒框架基本不需要人工干预，所以以牺牲了一定的计算时间作为代价。在自动化程度和计算效率不能兼顾的情况下，我们选择了提高自动化程度，因为相信随着计算机技术日新月异的发展，在不远的将来，计算时间必将被大大缩短。

表1所示的是，图8中的实验图像在低分

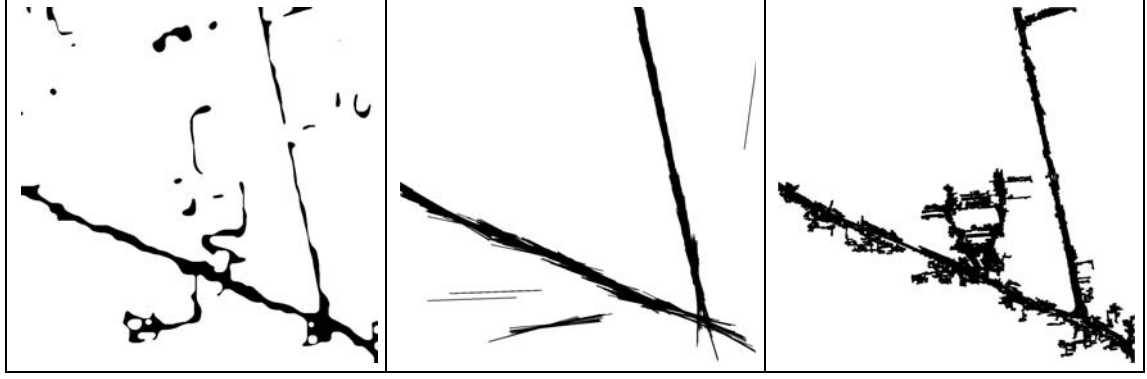


图12 在原分辨率 L_0 的图像上，与其它方法的结果对比。从左到右：采用不包含 $E_{P,NL}$ 的 E_{High} （即 $\beta = 0$ ），Yu^[12]和Wang^[13]的方法分别得到的结果。

指标	E_{High}	E_{Low}	不包含 $E_{P,NL}$ 的 E_{High}	Yu	Wang
TP/(TP+FN)	0.8317	0.6212	0.9040	0.8674	0.8270
TP/(TP+FP)	0.7491	0.4613	0.2792	0.6914	0.3679
TP/(TP+FP+FN)	0.6505	0.3601	0.2711	0.6253	0.3416

表3 在原分辨率 L_0 的图像上，不同方法的量化评价(T=true, F=false, P=positive, N=negative)。

分辨率 L_2 和原分辨率 L_0 下的参数设置。这些参数是在保证模型的稳定^[10]和保证带状结构宽度稳定的约束下手动选取的。除了 d 控制着道路的宽度，以及整个先验能量权值 ω 的变化，其它参数在两个分辨率下设置一致。

在两个分辨率下，我们将本文提出的方法，和没有包含二阶几何先验 $E_{P,NL}$ （即 $\beta = 0$ ）的能量、Yu等人^[12]提出的基于直线密度的方法、Wang等人^[13]提出的采用分类、跟踪和数学形态学的方法进行比较。图11和图12是其它三种方法分别在低分辨率 L_2 和

原分辨率 L_0 下得到的分割结果。两个分辨率下的结果的量化评价^[14]见表2和表3。完整率TP/(TP+FN)是提取的正确道路与正确道路之间的百分比；正确率TP/(TP+FP)是提取的正确道路与提取道路之间的百分比；质量TP/(TP+FP+FN)综合考虑了完整率和正确

率，因此是其中最重要的一个指标。比较表明，一方面，二阶几何先验知识的加入具有非常重要的作用。另一方面，Wang的方法得到的结果道路边缘不准确，容易受到噪声的影响；从量化评价的数值来看，Yu的方法仅次于我们提出的方法，但是这个方法很大的局限性在于只能处理直线道路，对图8右下方稍带弧度的道路无效。

图13所示的是在另一幅QuickBird图像上的实验结果。右上图是能量 E_{Low} 在低分辨率 L_2 图像的提取结果，其中并不包含图像右上方的小路，垂直方向的主干道路的边界位置也不是特别精确。若将能量 E_{Low} 作

用于原分辨率 L_0 图像，如左下图所示，结果仍然不理想。右下图是采用多分辨率模型框架在原分辨率 L_0 图像上的最终结果。值得注意的是，在原分辨率下，图像右上方的小路也被提取出来了，垂直方向的主干道路

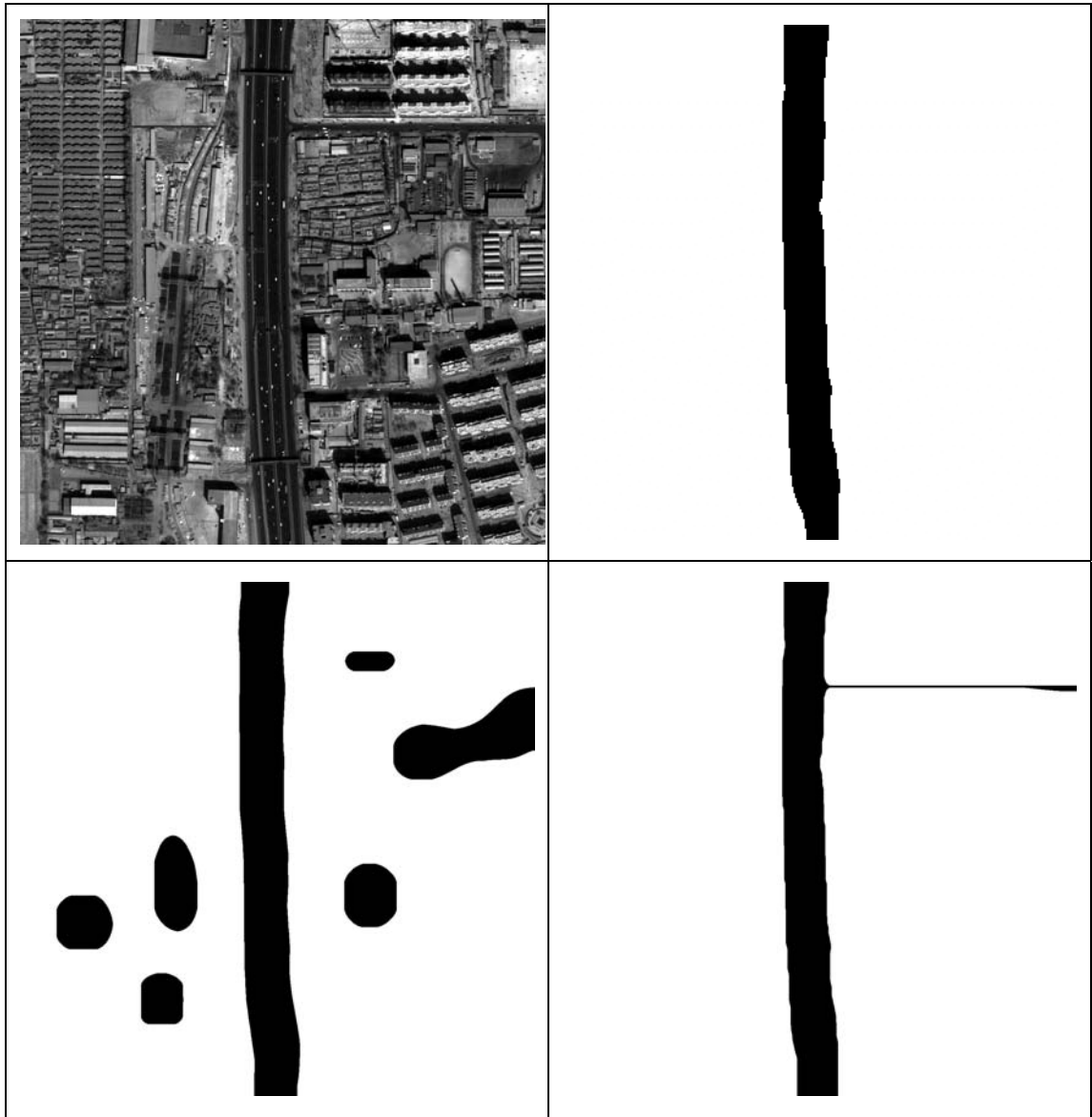


图13 QuickBird图像上的更多实验结果。左上：QuickBird实验图像（大小：1280×1280 像素）；右上：采用能量 E_{Low} 在低分辨率 L_2 的图像（大小：320×320 像素）上的实验结果。

第二行从左至右：分别是采用能量 E_{Low} 和能量 E_{High} ，在原分辨率 L_0 的图像上的分割结果。

的位置和宽度也更加精确。比如，垂直道路最下端，由于白色道路标示线的影响，右下图原分辨率下的提取结果，比右上图低分辨率下的结果更接近图像的实际情况。但是由于先验能量中控制道路宽度的 d 值被设置为主干道路的宽度，所以结果中的小路的宽度有些失真，比实际宽度要小一些。能够精确提取多个不同宽度的模型正是正在进行中的工作。

更多的实验结果见图14和图15。在低分辨率图像上，虽然能量 E_{Low} 可以提取大部分道路，但是由于 L_2 图像中的每一个像素都是 L_0 图像中 4×4 个像素的平均值，即 L_2 图像中的道路边界同时包含了原分辨率 L_0 图像中的道路信息和背景信息，所以结果中道路的



图14 QuickBird图像上的更多实验结果。左上：QuickBird实验图像（大小：1000×1000 像素）；右上：采用能量 E_{Low} 在低分辨率 L_2 的图像（大小：250×250 像素）上的实验结果。

第二行从左至右：分别是采用能量 E_{Low} 和能量 E_{High} ，在原分辨率 L_0 的图像上的分割结果。

边界呈锯齿状，宽度也不是十分准确。正因如此，道路提取需要在原分辨率上进行。但是若将能量 E_{Low} 直接作用于原分辨率图像，与前面的实验相似，结果中存在大量的噪声和错误检测。采用多分辨率模型框架在

原分辨率 L_0 图像上的最终结果的边缘更加平滑，而且在道路交叉口处的提取更加精确。图16所示的是一幅大图像（大小：3300×1100像素）和采用我们的模型在原分辨率上的提取结果。

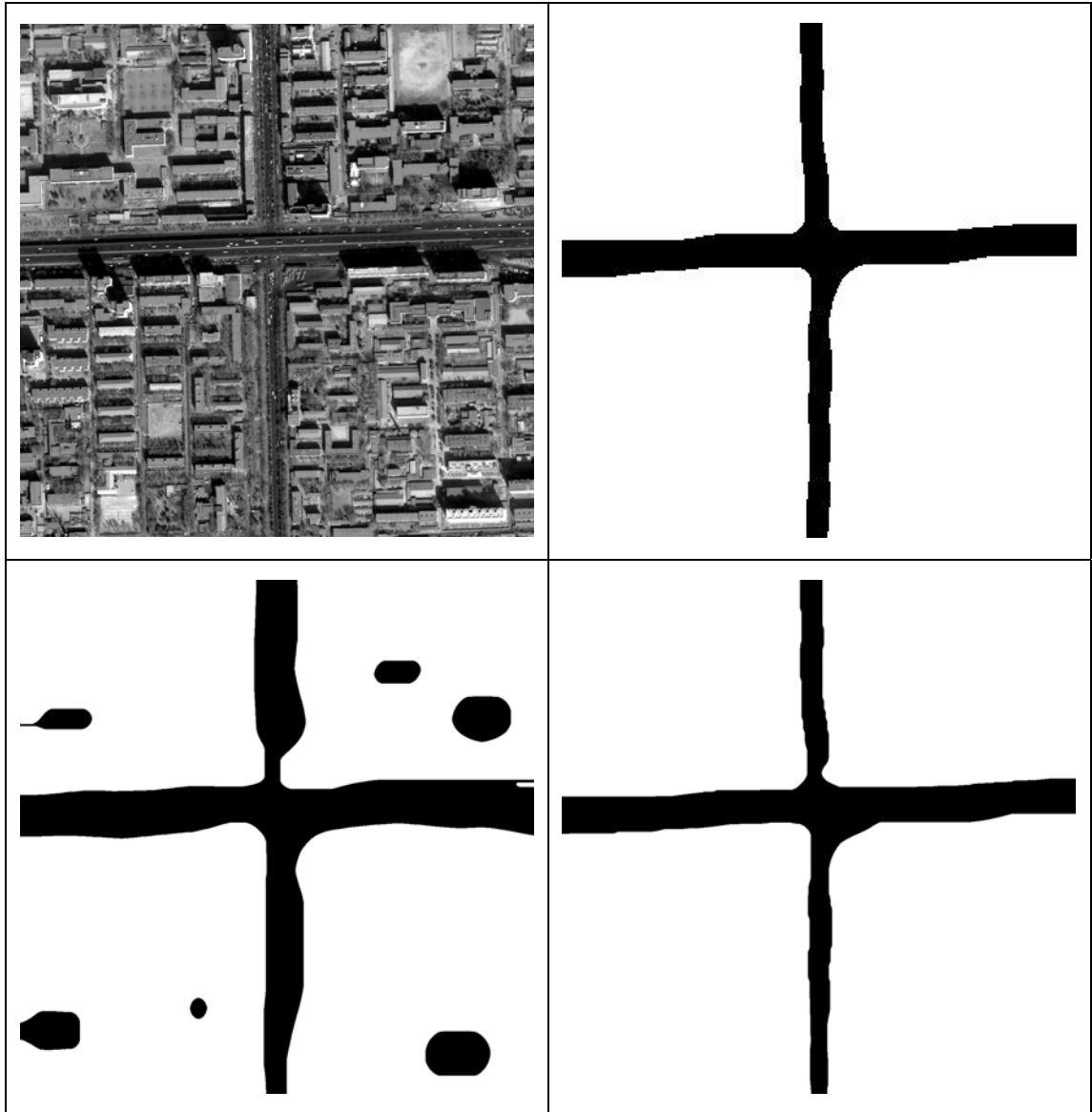


图15 QuickBird图像上的更多实验结果。左上：QuickBird实验图像（大小：1280×1280 像素）；右上：采用能量 E_{Low} 在低分辨率 L_2 的图像（大小：320×320 像素）上的实验结果。

第二行从左至右：分别是采用能量 E_{Low} 和能量 E_{High} ，在原分辨率 L_0 的图像上的分割结果。

5. 结论

本文采用“相位场”对目标区域建模，提出了一种基于多分辨率分析和相位场高阶主动轮廓模型的鲁棒计算框架。一方面，和传统的水平集函数表达区域的方法相比，相位场模型具有很多优势^[10]：相位场函数处于线形空间，因而实现容易；无偏初始化；更大的拓扑自由度。另一方面，高阶主动轮

廓^[8]通过引入轮廓上点对之间长距离的相互作用，描述了带状或者线状物体的几何结构，克服了传统主动轮廓对初始轮廓位置敏感的缺点。

本文的主要贡献是，充分地结合了低分辨率数据和高分辨率数据的优势，提出了一种将多个分辨率数据有效地结合的新思想。在复杂场景建模中，利用低分辨率的初步分割结果，对在原分辨率图像上的分割进行空间约束，驱使最终的分割结果处于一个与预

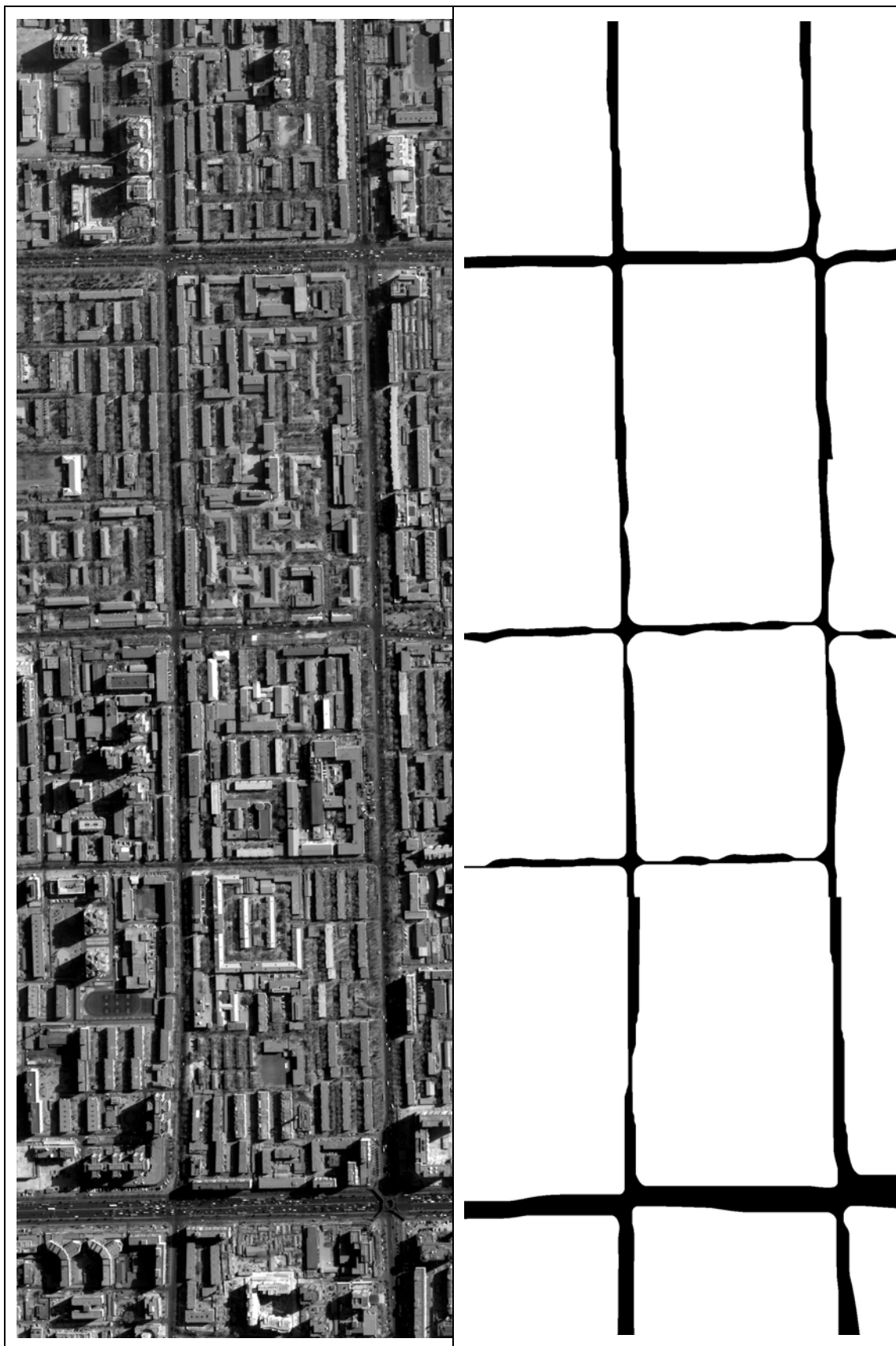


图16 一幅大图像（大小：3300×1100像素）和采用我们的模型在原分辨率上的提取结果。

分割结果比较接近的空间中,从而,有效地抑制了原分辨率图像中噪声和细节大幅度出现对分割过程可能产生的不利影响,极大地提高了模型的鲁棒性。我们将这个计算框架应用于高分辨率密集城区的光学遥感图像的道路网络提取中,分析了模型中不同能量项的作用,并与文献中其它道路提取算法进行了比较。实验结果和与其它方法的比较,充分地证明了算法的有效性:在原分辨率上,道路网络被完全提取出来,并且比其它方法效果更好。

目前正在进行的工作是构造可以精确提取多个不同宽度的道路模型,以及改善算法的计算效率。

参 考 文 献

- [1] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision*, 1988, 1 (4): 321-331
- [2] Osher S, Sethian J A. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 1988, 79 (1): 12-49
- [3] Chan T, Vese L. Active contour without edges. *IEEE Transactions on Image Processing*, 2001, 10 (2): 266-277
- [4] 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: 249-268
- [5] Cremers D, Tischhauser F, Weickert J, Schnörr C. Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *International Journal of Computer Vision*, 2002, 50 (3) : 295-313
- [6] Leventon M E, Grimson W E L, Faugeras O. Statistical shape influence in geodesic active contours//*Proceedings of the IEEE Computer Vision and Pattern Recognition*. Hilton Head Island, South Carolina, USA, 2000, 1: 316-322
- [7] Péteri R, Ranchin T. Detection and extraction of road networks from high resolution satellite images//*Proceedings of the IEEE International Conference on Image Processing*. Barcelona, Spain, 2003, 1: 301-304
- [8] Rochery M, Jermyn I H, Zerubia J. Higher order active contours. *International Journal of Computer Vision*, 2006, 69 (1): 27-42
- [9] Chen L. Phase-field models for microstructural evolution. *Annual Review of Materials Research*, 2002, 32: 113-140
- [10] Rochery M, Jermyn I H, Zerubia J. Phase field models and higher order active contours//*Proceedings of the IEEE International Conference on Computer Vision*. Beijing, China, 2005, 2: 970-976
- [11] Mallat S. A wavelet tour of signal processing. Academic Press, 1997
- [12] Yu Z, Prinnet V, Pan C, Chen P. A novel two-steps strategy for automatic GIS-image registration//*Proceedings of the IEEE International Conference on Image Processing*. Singapore, 2004, 3: 1711-1714
- [13] Wang R, Zhang Y. Extraction of urban road network using QuickBird pan-sharpened multispectral and panchromatic imagery by performing edge-aided post-classification//*Proceedings of the International Society for Photogrammetry and Remote Sensing*. Quebec City, Canada, 2003, 1-6
- [14] Heipke C, Mayr H, Wiedemann C, Jamet O. Evaluation of automatic road extraction. *International Archive of Remote Sensing*, 1997, XXXII: 47-56

A Robust Framework Based on Phase Field Modeling for Road Network Extraction from VHR Satellite Images

PENG Ting^{1),2)} JERMYN Ian H.²⁾ PRINET Véronique¹⁾ ZERUBIA Josiane²⁾

¹⁾ (LIAMA&NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China 100190)

²⁾ (Project-Team Ariana, INRIA, 06902 Sophia Antipolis, France)

Abstract Phase field higher-order active contours emerged recently as a powerful model for region-based and shape-based image segmentation. A higher-order representation makes it possible to model regions/objects of sophisticated shapes. Moreover, compared to traditional level set approaches, phase field modeling enables to get rid of the problem of contour initialization and distance function reinitialization. In this paper, we address the detection of band or linear structure objects (roads, vessels, etc.) from very noisy images. We propose a robust framework based on phase field higher-order active contours. We segment at first a low resolution image, which is defined by the scaling coefficients of a Haar wavelet decomposition at order n . Since reducing the resolution eliminates high frequency noise, the resulting pre-segmentation gives a rough but complete detection of the regions of interest. This rough result is then incorporated into the model, such as to force the final segmentation to lie close to the pre-segmentation. We apply the proposed framework to road network extraction from very high resolution (VHR) satellite images in dense urban areas. Quantitative validation and comparison with other methods demonstrate the efficiency of our approach.

Keywords phase field modeling; prior knowledge; active contours; VHR satellite image; road network segmentation



PENG Ting, born in 1982, Ph. D. . Her research interests include computer vision and image processing.

JERMYN Ian H. , born in 1964, Ph. D. , research scientist. His main research interests concern the modeling of shape and texture, and information geometry as applied to inference and shape analysis.

PRINET Véronique, born in 1969, Ph. D. , associate professor. Her current research interests include image processing, computer vision, machine learning, scene understanding, and applications to remote sensing images.

ZERUBIA Josiane, born in 1957, Ph. D. , research director, Ph. D. supervisor. Her current research interests are in image processing using probabilistic models and variational methods. She also works on parameter estimation and optimization techniques.

通讯作者Email pengting@gmail.com。电话：010-82614490（由于第一作者已毕业离所，此电话可以与本文第三作者PRINET Véronique，中文名普林特老师联系）。

Background

This work was a part of a Sino-French cooperation between the LIAMA/NLPR and the INRIA (French National Institute for Research in Computer and Control Sciences). It was partially supported by LIAMA Project “Digital map updating on urban areas using high resolution images”, and by European Union Network of Excellence MUSCLE (FP6-507752). The work of the first author was supported by an MAE/Thales Alenia Space/LIAMA grant. We aim at the study of variational models, with a specific application on remote sensing imagery. In particular, we address the problem of road network extraction from Very High Resolution (VHR) satellite images, in dense urban areas. Detecting roads from remote sensing imagery is critical for many applications, for example cartographic data updating, intelligent navigation, environmental monitoring, disaster management, and so on. However, this is also a very hard problem. The difficulties lie in the following factors. First, much “noise” exists in the road region due to cars, road markings, shadows, etc., while the background is very diverse, containing many features that are locally similar to roads. Second, rather than being simple lines as in images at low resolutions, i.e. more than 5m/pixel, roads appear as elongated, more or less homogeneous surfaces with different widths and curvatures. It is a hard task to retrieve road surfaces. Last but not least, in a dense urban environment, the contrast between roads and background is relatively poor, as opposed to a rural or semi-urban area. All the above factors result in the relative failure of existing road extraction approaches. The development of a reliable automatic algorithm for road extraction in dense urban areas from VHR images is thus a necessity if the increasing demand is to be met. To overcome the difficulties mentioned above, we propose to include a great deal of prior knowledge about the appearance and shape of road networks into our model. The complete work takes advantage of the existence of outdated road maps, the strong geometrical constraints that characterize roads, and the multiscale nature of the phenomena in the images. This paper considers the last two aspects. Our work makes important progress towards an automatic road extraction system for VHR satellite images.