

# Snake 模型综述

李天庆, 张毅, 刘志, 胡东成

(清华大学自动化系, 北京 100084)

**摘要:** 基于轮廓线模型的目标分割、物体跟踪方法是近十几年来图像和视频领域研究的热点。因为轮廓线模型可以将待处理问题的先验知识与各种图像处理算法有效地融合在一起, 所以比以往的计算视觉理论具有更强的实用性。虽然很多模型算法已经和最初的 Snakes 大相径庭, 但它们的指导思想却是一脉相承的。该文以 Snakes 模型为基础, 系统地介绍了轮廓线模型的发展脉络和各种最新成果。

**关键词:** 轮廓线模型; Snakes; 图像分割; 跟踪

## An Overview on Snakes Models

LI Tianqing, ZHANG Yi, LIU Zhi, HU Dongcheng

(Department of Automation, Tsinghua University, Beijing 100084)

**【Abstract】** Active contour models are the hotspot in the field of computer vision. They are more practically and powerful than other computer theories because the contours could merge prior knowledge and image processing algorithms. Although the forms of these models have many differences, the essence is the one. The paper tries to introduce the whole active contour models family.

**【Key words】** Active contour model; Snakes; Segmentation; Tracking

自 1987 年 Kass<sup>[1]</sup>提出 Snakes 模型以来, 各种基于主动轮廓线的图像分割、理解和识别方法如雨后春笋般蓬勃发展起来。Snakes 模型的基本思想很简单, 它以构成一定形状的一些控制点为模板 (轮廓线), 通过模板自身的弹性形变, 与图像局部特征相匹配达到调和, 即某种能量函数极小化, 完成对图像的分割。再通过对模板的进一步分析而实现图像的理解和识别。

在 Snakes 这样一个简单的模型下, 蕴含了深刻的计算机视觉理论的发展成果。以 Marr 为代表的计算机视觉理论, 采用自底而上的处理方法。它将各种检测算子作用于图像得到初始简图, 再通过一定规则对这些特征进行聚集而得到 2.5 维简图, 最终求解三维模型; 与此相对应的是从心理学发展起来的基于推理的视觉理解理论, 采用从顶而下的处理方法。进入 20 世纪 80 年代以后, 人们逐渐认识到放之四海皆准的视觉理论遥不可及, 而针对具体问题, 能够有效融合上层知识和底层图像特征的方法才是最有效、最切实可行的。Snakes 模型正是在这样一种背景下提出的。

构造 Snakes 模型的目的是为了调和上层知识和底层图像特征这一对矛盾。无论是亮度、梯度、角点、纹理还是光流, 所有的图像特征都是局部的。所谓局部性就是指图像上某一点的特征只取决于这一点所在的邻域, 而与物体的形状无关; 但是人们对物体的认识主要是来自于其外形轮廓。如何将两者有效地融合在一起, 正是 Snakes 模型的长处。Snakes 模型的轮廓线承载了上层知识, 而轮廓线与图像的匹配又融合了底层特征。这两项分别表示为 Snakes 模型中能量函数的内部力和图像力。

正是因为 Snakes 模型有这样深刻的理论背景, 它一经提出就成为计算机视觉领域研究的热点。10 多年来, 不仅针对 Snakes 模型本身的能量函数构造和求解算法作了很大改进,

更在其基础上衍生出了许多新模型。它们有些在形式上已经与 Snakes 相去甚远, 而且也要复杂得多, 但其指导思想却是一脉相承的。本文将对 Snakes 这个大家族做一梳理, 着重对新成果进行介绍。

### 1 基本 Snakes 模型家族

基于 Snakes 的思想, 在不同的应用背景下, 衍生出多种轮廓线模型 (表 1)。

表 1 基本 Snakes 模型家族一览

名称	提出者	年代	主要参考文献
Snakes	Kass M, Witkin A, Terzopoulos D	1987	[1]
Snakes (DP)	Amini AA	1990	[2]
Snakes (balloons)	Cohen L D	1991	[3]
GVF Snakes	Xu C, Prince J L	1998	[4]
Deformable template	Terzopoulos D	1987	[5]
Level set	Osher S	1990	[6]

#### 1.1 基本 Snakes 模型

Kass 等提出的原始 Snakes 模型由一组控制点

$$v(s) = [x(s), y(s)] \quad s \in [0, 1] \quad (1)$$

组成, 这些点首尾以直线相连构成轮廓线。其中  $x(s)$  和  $y(s)$  分别表示每个控制点在图像中的坐标位置,  $s$  是以傅立叶变换形式描述边界的自变量。在 Snakes 的控制点上定义能量函数

**基金项目:** 国家社会科学基金资助项目(99BTY004)

**作者简介:** 李天庆(1977—), 男, 博士生, 主研方向: 图像和视频处理; 张毅, 教授、博导; 刘志, 副研究员; 胡东成, 教授、博导  
**定稿日期:** 2004-04-27 **E-mail:** ltiq99@mails.tsinghua.edu.cn

$$E_{total} = \int_s \left( \alpha \left| \frac{\partial v}{\partial s} \right|^2 + \beta \left| \frac{\partial^2 v}{\partial s^2} \right|^2 + E_{ext}(v(s)) \right) ds \quad (2)$$

其中第 1 项称为弹性能量，是  $v$  的一阶导数的模；第 2 项称为弯曲能量，是  $v$  的二阶导数的模；第 3 项是外部能量（外部力），在基本 Snakes 模型中一般只取控制点或连线所在位置的图像局部特征，例如梯度

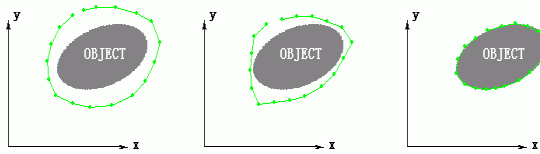
$$E_{ext}(v(s)) = P(v(s)) = -|\nabla I(v)|^2 \quad (3)$$

也称图像力。弹性能量和弯曲能量合称内部能量（内部力），用于控制轮廓线的弹性形变。选取适当的参数  $\alpha$  和  $\beta$ ，将能量函数  $E_{total}$  极小化，所对应的  $v(s)$  就是对物体的分割。在能量函数极小化过程中，弹性能量迅速把轮廓线压缩成一个光滑的圆；弯曲能量驱使轮廓线成为光滑曲线或直线；而图像力则使轮廓线向图像的高梯度位置靠拢。基本 Snakes 模型就是在这 3 个力的联合作用下工作的。

最终对图像的分割转化为求解能量函数  $E_{total}(v)$  极小化。很明显，这是一个典型的变分问题，在离散化条件（数字图像）下，由欧拉方程可知，最终问题的答案等价于求解一组差分方程

$$-\alpha v'' - (\alpha - \beta''')v'' + 2\beta v'' + \beta v'''' = -\nabla P(v) \quad (4)$$

记外部力  $F = -\nabla P$ 。Kass 等将上式离散化后，对  $x(s)$  和  $y(s)$  分别构造两个五对角阵的线性方程组，通过迭代计算进行求解。在实际应用中，一般先在物体周围手动点出控制点，作为 Snakes 模型的起始位置，然后对能量函数迭代求解。具体过程如图 1。



注：图 1 中 3 幅图分别表示 Snakes 模型的起始位置，第 8 步和第 32 步迭代求解的位置。

图 1 Snakes 模型示例

## 1.2 模型力的改进

### (1) Snakes (balloons)

Kass 等提出的基本 Snakes 模型，在没有图像力平衡的条件下，内部力将把所有控制点收缩为一点或一条直线。也就是说，被分割物体必须完全包含在 Snakes 的初始位置之内，否则陷在内部的控制点将无法回到物体的边界（图 2）。造成问题的原因在于离散化的数值求解过程。式（4）中  $F$  的方向就是  $P$  的最速下降方向，而迭代步长正比于  $P$  的梯度绝对值，过大的步长会使轮廓线轻易地越过平衡点，无法收敛到物体边界。另一方面在内部力的平衡下，最终只有少数梯度很大的图像点才能影响 Snakes 的形变，而绝对值不很大的边界点将不起任何作用。为了克服这个缺点，Cohen 等提出归一化的外部力  $F = -k \frac{\nabla P}{\|\nabla P\|}$ ，使得不论图像梯度绝对值的大或小，

对轮廓线都有相同的影响。Snakes 模型的内部力只能把轮廓线压缩成一个点或者一条直线。Cohen 等引入一个膨胀的外部力，综合在一起成为

$$F = k_1 n(s) - k \frac{\nabla P}{\|\nabla P\|} \quad (5)$$

这样轮廓线可以像“气球”一样膨胀，具有更强的动态行为能力。其中  $n(s)$  是轮廓线上控制点  $v(s)$  的单位法向量， $k_1$  是膨胀力的幅值，取正值或者负值就可以使轮廓线具有膨胀或者收缩的行为能力。一般选取  $k$  略大于  $k_1$ ，使得边界点刚好可以抵消膨胀力。

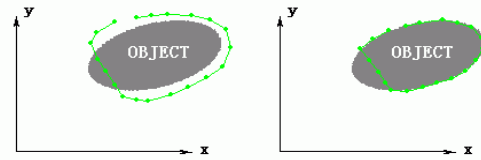


图 2 基本 Snakes 模型对陷入目标内部的控制点失去效力

### (2) GVF Snakes

Kass 等提出的基本 Snakes 模型在工程应用中有两个难点：(1) 轮廓线的初始位置必须离真实边界很近，否则难以得到正确的分割结果。为了解决这个问题，人们提出了 multiresolution 方法、增加外部压力和距离势能等多种改进方法，它们的基本思想都是提高外部力的作用范围，从而引导轮廓线趋向真实边界；(2) 基本 Snakes 模型对无法捕获凹陷边界，尽管也有很多改进方法，但效果都不理想。Xu Chenyang 等提出的 GVF Snake 很好地解决了这些问题。

GVF Snake 的创新在于将传统 Snakes 的图像力（edge map）用扩散方程进行处理，得到整个图像域的梯度向量场（gradient vector flow field）作为外部力。经过扩散方程处理后的 GVF 比 edge map 更加有序、更能体现物体边界的宏观走势（图 3）。

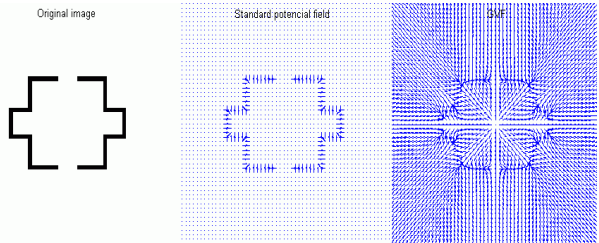


图 3 传统梯度图和经过扩散方程处理后的 GVF 域

由于 GVF 不是一个表达式，无法用能量函数的形式求解，因此 GVF Snake 是利用力的平衡条件进行优化（图 4）。GVF Snakes 具有更大的搜索范围，对轮廓线初始位置不敏感。可以分割凹陷的边界，对梯度绝对值的大小乃至噪声具有更好的鲁棒性。而且它还不必先知道轮廓线是要膨胀（Balloon Snake）还是收缩（Snake）。

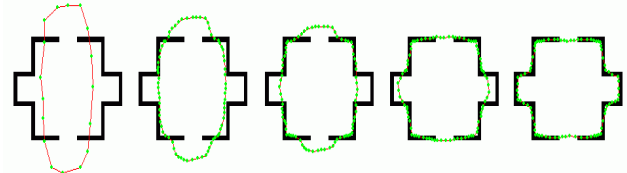


图 4 GVF Snake 模型示例

### 1.3 轮廓线模型的发展

在 Kas 等提出的多边形轮廓线的基础上，为了减少控制点个数、提高模型的灵活性和专用性，又分别衍生出了 B 样条曲线、函数曲线和点分布 3 种新的轮廓线模型。

#### (1) B-spline Snake

从计算机图形学可知，B 样条曲线仅需要少量的控制点就可以柔和剧烈变化的曲线。而且每一个控制点的移动仅引起一段曲线变化，不会影响整体形状。基本 Snakes 模型的多边形轮廓线就可以看作一次 B 样条。实际使用中，常采用 2 次或 3 次 B 样条曲线作为 Snakes 的轮廓线。

#### (2) Deformable Template

Deformable Template 采用参数化的函数曲线描述某一类物体的外形，例如一个表现眼睛的变形模板由两段抛物线和一个圆组成（图 5）。在参数的驱动下，函数曲线产生一定形变来模仿实际物体。当变形模板与理想模板的偏差和变形模板与图像特征的匹配程度，两者在某种规则下达到最小时就完成了图像的分割。

对 Deformable Template 最常用到的是一种新的递归优化算法。它在模板控制点的法线向量上求解图像特征，生成备选的特征模板，再对模板进行一定位移和形变以适应特征模板，由此递归求解。这

种递归算法就是内外力平衡方法的具体实现。

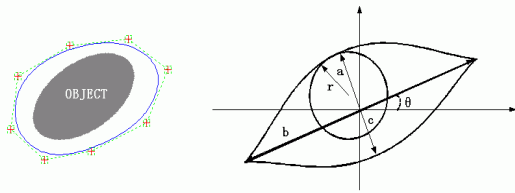


图 5 B-spline Snake 和 Deformable Template 示例

### (3) Point distribution models (PDM)

点分布模型是近几年发展起来的一种强大的图形描述技术。首先要得到一组实际物体轮廓线的样本, 进行归一化处理后, 为每个样本提取一组标志点。对样本集标志点学习得到的概率分布就是点分布模型。PDM 从样本集中学习先验知识, 并把它保存在概率分布中, 比 Deformable Template 更加通用、灵活。使用 PDM 的典型应用是 Active Shape Model。

### 1.4 算法改进

通过求解欧拉方程的数值算法存在一些难以克服的缺点, 比如不易得到全局最优解、需要计算高阶导数、加入硬约束后缺乏稳定性。为了极小化 Snakes 模型的能量函数, Amini 等提出基于动态规划的算法, 从而开启了 Snakes 模型组合优化之先河, 也使得 Snakes 模型在计算机视频领域得以广泛应用。它对 Snakes 模型的  $n$  个控制点的  $m$  邻域, 组合而成的所有备选变形位置, 通过动态规划求得其中某个组合的能量函数相对最小, 由此迭代计算, 最终使轮廓线收敛到物体边界。

另一种更加常用的组合优化方法是贪婪算法。动态规划的每一步要移动  $n$  个节点, 而贪婪算法仅移动一个节点, 通过轮流移动  $n$  个节点最终得到一个解。

以上是 Snakes 模型最常用的 3 个能量函数极小化算法。数值算法形式规范, 每一步迭代步长较大, 计算速度比较快。但对参数太敏感, 且不易增加新约束; 动态规划算法不仅容易增加各种约束, 而且能够得到最优解。但是计算速度慢, 为  $O(nm^3)$ , 且最终轮廓会趋向物体的强特征部分, 使节点间距不均匀; 贪婪算法实现简单, 速度快, 为  $O(nm)$ , 容易加入新约束, 最终节点分布比较均匀, 但无法得到最优解。

前面谈到 Deformable Template 和 GVF Snake 算法, 直接利用内、外部力的平衡关系优化轮廓, 更加灵活、稳定, 是对 Snakes 优化算法的一大改进。

### 1.5 Snakes 测地线模型和 Level set 方法

Caselles 等和 Malladi 等分别独立提出的 Snakes 测地线模型, 给出了一种比导数更加优美而牢固的内部力数学描述。这种模型基于测地流的曲线演化, 演化速度和曲率均值相关。好比水漫过崎岖的地表, 小的沟槽会被迅速填平, 而水表面的曲线长度保持最小。曲线  $C$  的欧氏长度定义为

$$L := \int |C'(q)| dq = \int ds \quad (6)$$

其中  $s$  为欧氏曲线长度。可以证明曲线流为

$$C_t = \kappa \vec{N} \quad (7)$$

其中  $\kappa$  为欧氏曲率,  $\vec{N}$  为法向量, 是对  $L$  的最速最小化途径。该流也称为欧拉曲线收缩流。如果把能量改写为

$$L := \int g(|\nabla I(C(q))|) |C'(q)| dq = \int g(|\nabla I(C(q))|) ds \quad (8)$$

把图像信息转化为适当的权重函数  $g(|\nabla I(C(q))|)$ , 以加权欧拉曲线的长度。物体分割的问题以最小化新的能量函数的方式完成。由相应的欧拉方程可以得出能量的最小化梯度流为

$$C_t = g(I) \kappa \vec{N} - (\nabla g \cdot \vec{N}) \vec{N} \quad (9)$$

作为曲线运动, 式(9)第 1 项平滑曲线, 第 2 项使曲线吸引到图像中物体的边界。很显然, Snakes 的测地线模型比 Kass 等的 Snakes 模型在形式上简洁优美, 数学基础更加牢固。

Osher 与 Sethian 提出的 Level set 方法就是此模型的一个实现。设  $\Gamma(t)$  为  $R^n$  中的一个  $n-1$  维分割面, 它包含了一个开区域  $\Omega(t)$ 。它的运动过程表示为一个速度场, 与时间、界面位置、几何特性以及外部物理因素有关。Level set 函数  $\Phi(x, t)$  是一个定义在  $R^n$  的辅助标量函数, 它的零 Level set 总是与  $\Gamma(t)$  相重合, 满足

$$\begin{cases} \Phi(x, t) < 0 & \text{in } \Omega(t) \\ \Phi(x, t) = 0 & \text{on } \Omega(t) \\ \Phi(x, t) > 0 & \text{in } R^n \setminus \overline{\Omega(t)} \end{cases} \quad (10)$$

其中  $x \in R^n$ ,  $t \in R^+$ 。一个典型的 Level set 函数各点的取值为到其零 Level set 曲线的有符号欧拉距离, 可由传播方程

$$\frac{\partial \Phi(x, t)}{\partial t} = \text{sign}(\Phi) (1 - |\nabla \Phi|) \quad (11)$$

得到稳定状态的解。Level set 方法的特点是构造一个 3 维曲面, 和图像平面相交于 Snake 曲线, Level set 通过演变这个 3 维曲面而使 Snake 曲线变形。这种升维处理问题的方法虽然复杂一些, 但好处是 Snake 曲线的拓扑变化不会造成 3 维曲面的拓扑变化。这就使得 Level set 可以动态地同时检测几个物体, 而无须事先知道确切数目。

## 2 总结

轮廓线模型是近十几年来图像视频领域的研究热点。轮廓线模型家族起源于 Snakes, 因融合了先验知识和图像特征而具有很高的实用价值。由其衍生出的各种轮廓线模型分别在形态、结构和算法等不同方向发展, 派生出许多如 Condensation 算法这样重要的工具。本文在 Snakes 模型的基础上, 系统地对各种轮廓线模型加以讨论, 力求把握住这一领域的发展脉络, 从而达到对各种轮廓线模型本质的探讨。

### 参考文献

- 1 Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. International Journal of Computer Vision, 1987: 321-331
- 2 Amini A A, Tehrani S, Weymouth T E. Using Dynamic Programming for Minimizing the Energy of Active Contours in the Presence of Hard Constraints. Proc. Second International Conference on Computer Vision, 1988: 95-99
- 3 Cohen L D. On Active Contour Models and Balloons. CVGIP(53), 1991, (2):211-218
- 4 Xu C, Prince P L. Snakes, Shapes, and Gradient Vector Flow. IEEE Trans. on Image Processing, 1998,7: 359-369
- 5 Terzopoulos D. Matching Deformable Models to Images. Technical Digest Series: Topical Meeting in Machine Vision, 1987,12:160-163
- 6 Osher S, Rudin L. Feature-oriented Image Enhancement Using Shock Filters. SIAM J. Num. Analysis, 1990, 27:919-940