

基于 MR-MS 的 snake 模型及其应用

丁 炜, 成思源, 张湘伟

(广东工业大学 机电工程学院, 广东 广州 510090)

摘要: 由 Kass 等人 1987 年提出的 snake 模型在计算机视觉领域引起了研究人员的广泛关注, 并且被越来越多的应用于计算机视觉的许多实际问题. 本文针对其缺陷提出了基于多分辨率和多尺度(简称 MR-MS) 的 snake 型, 并将其应用于医学的图像处理, 实验结果证明了新模型的有效性.

关键词: snake 模型; 多分辨率; 多尺度

中图分类号: TP391

文献标识码: A

文章编号: 1007-7162(2005)01-0037-05

Kass 等人在 1987 年第一届计算机视觉国际会议上提出活动轮廓模型(又称为(snake)), 用以解决计算机视觉领域的经典难题——如何锁定图像中感兴趣的部分^[1]. 这一方法自提出以来受到研究人员的广泛关注, 并被广泛应用于边缘检测、运动跟踪、3-D 重建、图像匹配等方面. 本文首先对经典 snake 模型做了简要回顾, 并指出其存在的不足; 在此基础上引入多分辨率(multi-resolution)和多尺度(Multi-scale)的概念, 并建立了基于多分辨率和多尺度(简称 MR-MS)的新型 snake 模型; 最后通过实验结果证明了本文所提出方法的有效性.

1 经典 snake 模型的原理和实现

Snake 是一条受控连续样条可变形模板, 采用低层与高层信息相结合的原理来锁定感兴趣的图像特征. 当 snake 被放置在感兴趣的图像特征附近时, 作用在 snake 的力将使 snake 滑向图像特征. 作用在 snake 上的力取决于 snake 被放置的位置和它的形状在空间如何改变, 通常用内部力和外部力的和来表示. 内部力保证 snake 的光滑性, 外部力使 snake 滑动. 当由内部力和外部力的积分得到的全部能量达到极小时, snake 便完成滑动而锁定感兴趣的图像特征.

Snake 的参数描述为 $\mathcal{U}(s) = (x(s), y(s))$, 这里参数 $s \in [0, 1]$. Snake 的总能量可表示为

$$E_{\text{snake}} = \int_0^1 (E_{\text{int}}(\mathcal{U}(s)) + E_{\text{image}}(\mathcal{U}(s)) + E_{\text{con}}(\mathcal{U}(s))) ds, \quad (1)$$

其中, $E_{\text{int}}(\mathcal{U}(s)) = [\alpha(s)|\mathcal{U}_s(s)|^2 + \beta(s)|\mathcal{U}_{ss}(s)|^2]/2$,

$$E_{\text{image}}(\mathcal{U}(s)) = \omega_{\text{line}} E_{\text{line}}(\mathcal{U}(s)) + \omega_{\text{edge}} E_{\text{edge}}(\mathcal{U}(s)) + \omega_{\text{tem}} E_{\text{tem}}(\mathcal{U}(s)),$$

$$E_{\text{con}}(\mathcal{U}(s)) = -k(x_1 - x_2)^2.$$

根据变分原理, 使式(1)满足极小化条件, 可得到欧拉—拉格朗日方程^[1]

$$\begin{cases} \mathbf{A} \cdot \mathbf{X} + \mathbf{F}_x(x, y) = 0 \\ \mathbf{A} \cdot \mathbf{y} + \mathbf{F}_y(x, y) = 0 \end{cases} \quad (2)$$

收稿日期: 2004-02-19

基金项目: 国家自然科学基金(50105013)

作者简介: 丁 炜(1976), 男, 硕士研究生, 主要研究方向为图像处理与模式识别.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.

在方程(2)中, $\mathbf{A}$ 是一个由系数 $a(s)$ 和 $\beta(s)$ 决定的五对角矩阵, $\mathbf{F}$ 由图特征决定. 令 $\mathbf{v}$ 为欧拉步长, snake 在各个时刻的位置表达式为

$$\begin{cases} x_{t+1} = (\mathbf{A} + \mathbf{v}\mathbf{I})^{-1}(\mathbf{v}\mathbf{x}_t - \mathbf{F}_x(x_t, y_t)) \\ y_{t+1} = (\mathbf{A} + \mathbf{v}\mathbf{I})^{-1}(\mathbf{v}\mathbf{y}_t - \mathbf{F}_y(x_t, y_t)). \end{cases} \quad (3)$$

从上面的求解过程可以发现, 复杂的图像处理难题被转化成了简洁的数学表达式求解. 与传统的方法相比, 经典 snake 模型主要具有以下优点^[2]:

1) 图像特征提取过程中, 采用一个能量方程来表达复杂的图像提取过程, 将图像处理问题转换成数值计算问题;

2) 经过适当的初始化, snake 能够自主的收敛于能量极小值状态;

3) 可以采用高斯平滑技术对图像进行预处理, 不仅提高算法的鲁棒性(robustness), 而且扩大了 snake 模型的捕获范围和降低计算复杂度.

同时, snake 模型也存在如下缺点:

1) 对初始位置敏感, 需要依赖其他机制将 snake 放置在感兴趣的图像特征附近;

2) 虽然使用变分法求解偏微分方程可以将图像处理的问题转换为适于计算机处理的数值计算, 但它缺乏数值稳定性, 且可能收敛于局部而非全局极值点;

3) 由于计算过程中需要对离散点求四阶导数, 所以不仅对噪声敏感, 而且轮廓很难收敛到曲率高的边缘, 从而对角点的表达束手无策.

2 MR-MS 模型的建立

2.1 MR-MS 概念的提出

分辨率用以描述图像精细的程度, 即单位面积内的像素(pixel)的数目. 从第一节的论述中可以发现: snake 模型本质是求解一个包含图像特征、轮廓特性和人为定义的约束力的能量方程, 能量方程的解即为目标曲线. 不幸的是: 由于可能轮廓 Ω 的空间通常非常巨大和由于强噪声的存在引起的能量函数收敛于局部极值, 致使求解这样一个能量函数是一件非常棘手的事. 更进一步的, 由于这种方法的极小化是建立在图像梯度的基础上的, 所以要求 snake 的初始位置应该恰当选定且不能离目标轮廓太远, 以防止 snake 收敛于错误的能量极值. 为了克服经典 snake 的这一缺陷, 科研工作者提出了许多方法, 建立在模拟退火算法(Simulated Annealing)基础上的随机方法就是其中的一种, 它可以避免 snake 模型收敛于极值和不需要人工对 snake 进行初始化^[3]. 但这些方法的一个致命的弱点在于它的计算量大到让人无法忍受. 为了在实际应用中缩短运算时间, 学者们提出将多分辨率引入活动轮廓模型^[4]. 多分辨率方法(又称多格法)应用于计算机视觉的很多领域. 当研究者希望进行一种迭代性的和需要多次连续更新的图像处理计算时, 多分辨率通常是首选, 如光流分析和图像分割等许多领域被作为一种有效的工具. 采用它不仅具有潜力找到问题的较优解, 而且一般可以减少算法的运行时间.

多分辨率的基本原则如下: 构造一个图像金字塔(image pyramid)并且开始在一个非常低分辨率的图像上进行某种过程(或称为处理), 结合本文是求能量函数的解, 把这一层次上的处理结果(轮廓收敛的位置)作为位于金字塔上面的下一个较高层次的初始位置, 重复这一过程直到到达金字塔的最顶层(对应于原始图像本身)^[5].

多分辨率方法实质上是一种分层优化、由粗糙到精细的算法, 在给出其一般算法之前先对下面各量的含义进行说明. 把原始图像的分辨率定义为 $2^N \times 2^N$, 2^M 表示最精细的图像上面的

用以表示 snake 轮廓的离散点数, S 表示分辨层数, $2^{N-S} \times 2^{N-S}$ 表示初始图像所投影到的粗糙平面的分辨率, 2^{M-S} 表示在该分辨率下表示活动轮廓所需的离散点数. 当要求解能量方程时, 首先在分辨率为 $2^{N-S} \times 2^{N-S}$ 的 S 层上进行, 当 snake 在 S 层上收敛于极小值后, 把这个极小值作为 $(S-1)$ 层的初始位置, 然后在图像分辨率为 $2^{N-S+1} \times 2^{N-S+1}$ 的 $(S-1)$ 层上进行求解能量方程, 同样的这一层的解将作为下一层的初始位置, 重复这一过程直到第 1 层, 即分辨率为 $2^N \times 2^N$ 的原始图像. 用公式表达上述过程如下:

$$\mathbf{V}_p^0 = \prod_p (\mathbf{V}_{p+1}^t), \quad (4)$$

$$\mathbf{V}_p^t = (\mathbf{Id} + \mathbf{A}_p)^{-1} (\mathbf{V}_{p+1}^{t-1} + \mathbf{F}_p(\mathbf{V}_{p+1}^{t-1})). \quad (5)$$

在这里, 下标 p 表示图像金字塔的层数其取值范围为 $[0, S]$, $\mathbf{V}_p$ 是大小为 2^{M-p} 的矢量, 表示图像金字塔中第 p 层的已经被离散化的活动轮廓曲线; Π_p 是图像金字塔中 $p+1$ 层到 p 层的投影; $\mathbf{F}_p$ 表示在第 p 层 snake 受到的力; $\mathbf{A}_p$ 表示在第 p 层的大小为 $2^{M-p} \times 2^{M-p}$ 的刚度矩阵; τ 是时间步长, t 为时间参数, $\mathbf{V}_p^t$ 表示在第 p 层中时间为 t 时的轮廓位置; $\mathbf{Id}$ 是单位矩阵.

由于 snake 模型是以像素为基础进行搜索, 从初始点开始在它相邻的点进行能量计算和比较, 所以对于一个没有噪声的灰度图像很容易确定其边界, 但是即使是少量噪声的存在, 也会使算法失控而导致 snake 收敛于错误的边界, 因此必须在应用 snake 模型计算前对目标图像进行预处理, 滤掉噪声. 在采用高斯函数进行滤波的过程中, 采用不同的高斯方差得到一组图像, 这组图像就被称为多尺度空间^[1]. 事实上高斯函数不仅可以滤掉噪声, 而且同时提高了 snake 模型的捕捉能力.

2.2 基于多分辨率和多尺度的活动轮廓模型

前面已经分别对多分辨率和多尺度空间的方法进行了阐述, 从前面的分析可发现这两种方法有相似的地方, 主要在于:

- 1) 两种方法都采用了分层优化, 由粗糙到精细的方法;
- 2) 两种方法都增强了活动轮廓模型的捕捉能力, 提高了 snake 的作用范围;
- 3) 两种方法都缩短了活动轮廓模型的运行时间.

当然, 这两种方法也有着本质的差别, 主要体现在:

1) 多尺度方法中, 通过选取不同的平滑函数的参数来改变图像, 平滑函数的作用只是低通滤波, 参数的改变只决定滤除噪声的情况和轮廓边缘的模糊度; 在多分辨率方法中, 图像的分辨率是一个随层而改变大小的量, 能量最小化在不同的分辨率上实现;

2) 多尺度方法中, 图像的特征随着滤波参数的不同而变得模糊; 多分辨率的情况下, 图像特征在低分辨率时消失了.

为了提高 snake 的实用性能, 本文将多尺度方法和多分辨率的方法结合起来, 建立基于多尺度和多分辨率的活动轮廓模型, 并且通过具体的图像处理结果来证明这一方法的有效性. 多尺度和多分辨率的活动轮廓模型的基本思想是: 在多尺度的基础上同时采用多分辨率, 这样可以结合二者的优点, 既能消除噪声的影响, 又能进一步提高 snake 模型的捕捉能力.

基于多尺度和多分辨率的活动轮廓模型的具体算法为

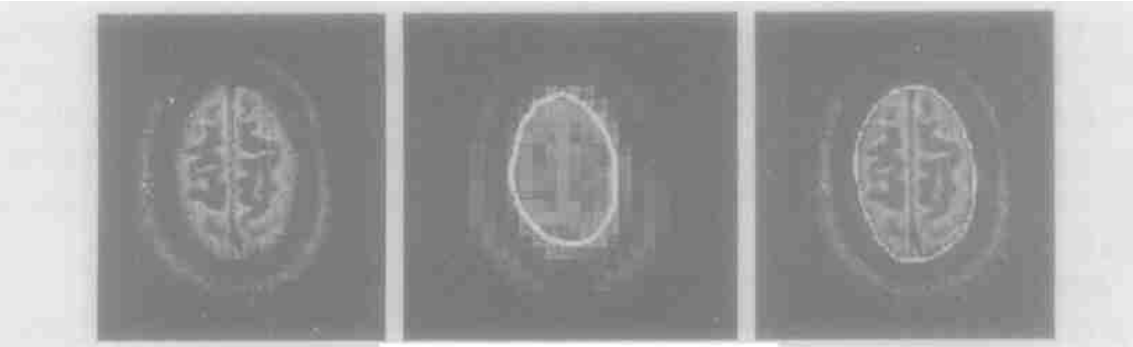
步骤 1, 采用标准差为 σ_0 的大尺度高斯函数对原始图像 $\mathbf{I}(x, y)$ 作滤波处理, 得到图像 $\mathbf{I}(x, y, \sigma_0)$. 建立以 $\mathbf{I}(x, y, \sigma_0)$ 为原始图像的图像金字塔, 然后调用式(4)、式(5)的多分辨率算法对 $\mathbf{I}(x, y, \sigma_0)$ 进行处理, 得到尺度 σ_0 下的最优解.

步骤 2, 减小图像尺度, 即以标准差为 $\sigma_i = \sigma_{i-1} - \delta\sigma$ (这里 δ 为尺度的迭代变化系数) 的高斯函数对原始图像重新滤波. 得到图像 $I(x, y, \sigma_i)$. 建立以 $I(x, y, \sigma_i)$ 为原始图像的图像金字塔, 然后调用式(4)、式(5)中的多分辨率算法对 $I(x, y, \sigma_i)$ 进行处理得到尺度 σ_i 下的最优解.

步骤 3, 重复步骤 2 操作, 直到 $\sigma_i < \sigma_{min}$, σ_{min} 一般可以取为 0.5 或 1, 根据实际情况亦可以采用其它自定义的停止准则.

3 实验结果

根据以上推导, 在 WinXP 系统环境下(计算机 CPU 2.0 GHz, 内存 256 MB), 开发了基于 MR-MS 的活动轮廓模型软件, 并将其应用于医学 MRI 图像脑部轮廓的提取, 结果如图 1 和图 2 所示. 采用本文中提出的方法, 经过 4 次迭代后, snake 模型都能平稳地收敛到目标轮廓, 取得了较好的效果, 证明了本文方法的有效性. 该方法不仅具有经典 snake 模型的优点, 而且可进一步克服其诸如对噪声敏感、捕捉范围小等缺陷.



(a) 原始图像 (b) 低分辨率迭代 (c) 最终迭代结果

图 1 采用 MR-MS 模型处理脑部 MRI 图像的结果 1



(a) 原始图像 (b) 中间迭代结果 (c) 最终迭代结果

图 2 采用 MR-MS 模型处理脑部 MRI 像的结果 2

参考文献:

[1] Kass M, Witkin A, Terzopoulos D. Snake: active contour models[A] . In: Brady I M, Rosenfield A eds. Proceedings of the 1st International Conference on Computer Vision[C] . London: IEEE Computer Society Press, 1987. 259-268.
[2] 李培华, 张田文. 主动轮廓线模型(蛇) 模型综述[J] . 软件学报, 2000, 11(6) : 751-757.
[3] Storvik G. A Bayesian approach to dynamic contours through stochastic sampling and simulated Annealing[J] . IEEE

Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(10) : 970-986.

[4] Wang M, Evans J, Hassebrook L, et al. A multistage, optimal active contour model[J] . IEEE Transactions on Image Processing, 1996, 5(11) : 1586-1591.

[5] 杨杨, 张田文. 基于多分辨率方法的主动轮廓线跟踪算法[J] . 计算机学报, 1998, 21(3) : 210-216.

Snake Model Based on MR-MS and Its Application

DING Wei, CHENG Si-yuan, ZHANG Xiang-wei

(Faculty of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510090, China)

Abstract: The snake model, which was developed by Kass et al. in 1987, has absorbed a lot of attention of researchers in the field of computer vision, and is applied to many cases of computer vision. To solve its limitation, a new snake model based on MR-MS (multi-resolution and multi-scale) is presented and used for medical image segmentation. The result of experiments shows that this new model is feasible.

Key words: snake model; multi-resolution; multi-scale

(上接第 36 页)

The Control Technology Research of Matrix Converter

ZHANG Miao¹, WU Jie², ZHAO Hui¹

(1. Faculty of Automation, Guangdong University of Technology, Guangzhou 510090, China

2. College of Power Supply, South China University of Technology, Guangzhou 510640, China;)

Abstract: This paper studies the control principle of matrix converter, based on the technology of PWM, analyzes the switch on/off mode of matrix converter power modules, and presents a novel control scheme of switching power modules to overcome the surge voltage during the process of the power modules switching on/off. The simulating results based on Simetrix 4. 2 and Matlab are presented at last.

Key words: matrix converter; Bi-direction switch; convert; AG-AC converting control