

基于动态规划的航空影像中的目标提取

刘少创 林宗坚

(武汉测绘科技大学遥感与信息工程学院 武汉 430070)

摘要 在对 Snakes 能量函数进行了改进之后,将其用于处理航空影像中的目标提取。动态规划方法被用来解决优化求解问题。最后给出了利用这种方法对目标进行提取的例子。

关键词 航空影像,目标提取,能量函数,动态规划,可变模型

1 引言

航空影像中目标的自动提取是计算机视觉界和摄影测量与遥感界共同关注的问题之一。人们经过多年的努力,虽然提出了不少航空影像中的目标提取方法,如文献[1—3]等,但是由于问题本身的难度和技术水平的局限性,要实现影像中目标全自动提取还有很多困难。因此,在航空影像理解系统中要考虑人的因素,这就是航空影像中目标的半自动提取问题。本文就是试图采用半自动的途径实现航空影像目标的提取。在利用本方法进行影像中目标的半自动提取时,先由人工给定一定数量的种子点,这样就得到了影像中目标的位置信息,而由种子点构成的形状也反映了目标的大致形状。对于人来说,给定这些种子点比较容易,但对于全自动化的影像理解系统的高水平处理部分来说却非常困难,这就是采取半自动途径的主要原因。在现有的半自动目标提取方法中,可变模型(Deformable Models)是一种比较有效的方法^[4],而 Snakes (Active Contour Models)^[5]则可以看作是可变模型的一个特例。本文就是在原始的 Snakes 模型的基础上,将能量函数进行修改,并采用动态规划方法进行优化求解,来实现航空影像中目标的提取的。

2 Snakes 的基本概念及其改进

原始的 Snakes 模型是 Kass 于 1987 年提出的^[5]。这一模型已经用在计算机视觉领域的很多方面,如医学图象处理^[6]、目标识别^[7]、三维物体的重建^[8]、光流计算、影像的几何纠正和影像匹配^[9]等,笔者也将这一模型用于航空影像参数的分割^[10],但是,原始的 snakes 的主要缺点之一是在处理过程中要加入许多参数,而这些参数的数值的设计需要凭人的经验进行。为了克服这个困难,本文对 Snakes 的能量函数进行了修改,并将修改后的能量函数用于航空影像中目标的提取,获得了满意的效果。

2.1 Snakes 的基本概念

假设有一条空间可变曲线 $v(s, t)$,它是由两个参数 s (空间位置)与 t (时间)决定的, v 由具有相同参数的两个变量 x, y 构成,即:

$$v(s, t) = [x(s, t), y(s, t)] \quad s \in \Omega, t \in T. \quad (1)$$

Snakes 的势能函数 E_{snake} 可以定义为:

$$E_{snake}(v) = \frac{1}{2} \int_{\Omega} [E_{int}(v) + E_{ext}(v) + E_{Field}(v)] ds \quad (2)$$

其中, $E_{make}(v)$ 表示 Snake 的总能量; $E_{int}(v)$ 代表 Snake 的内部势能, 它是 Snake 的内在的弯曲力与伸展力的函数; $E_{ext}(v)$ 代表由作用在 Snake 上的外部控制力引起的势能; $E_{Field}(v)$ 代表由影象的各种特征引起的势能。

在原始的 Snake 模型中, $E_{int}(v)$ 定义为:

$$E_{int}[v(s, t)] = \omega_1(s) |v_s|^2 + \omega_2(s) |v_{ss}|^2 \quad (3)$$

其中,

$$|v_s|^2 \equiv |\partial v / \partial s|^2 \approx |v_i - v_{i-1}|^2,$$

$$|v_{ss}|^2 \equiv |\partial^2 v / \partial s^2|^2 \approx |v_{i-1} - 2v_i + v_{i+1}|^2$$

一阶项 $\omega_1(s) |v_s|^2$ 的作用是使得 snake 受力时抗伸缩; 二阶项 $\omega_2(s) |v_{ss}|^2$ 的作用是使得 Snake 受力时抗弯曲; 权 $\omega_1(s)$ 所起的作用是调节 Snake 的伸缩力, 它相当于弹簧的弹性系数; $\omega_2(s)$ 是调节 Snake 的刚性。 $\omega_1(s)$ 与 $\omega_2(s)$ 都是位置的函数。 Snake 位置与曲率的非连续性是通过 $\omega_1(s)$ 与 $\omega_2(s)$ 置零得到的。

势能 $E_{Field}(v)$ 模仿经典的重力势能公式表示^[7], 即:

$$E_{Field}[v(s, t)] = \mu \cdot g \cdot I[v(s, t)] \quad (4)$$

其中, μ 为 Snake 的质量密度; g 表示重力加速度; $I[v(s, t)]$ 代表 t 时刻 Snake 上的点 $v(s, t)$ 处势曲面的高度。根据问题的性质, 势曲面可以以不同的形式出现, 可以是影象的强度或影象的梯度, 也可以是影象强度和梯度强度的组合, 但采用组合形式时, 虽然不会改变问题的性质, 却容易引起物理意义的不明确。

外部约束引起的势能 $E_{ext}(v)$ 是由高水平处理机制或人工施加的一些外部约束力引起的势能。这些外部约束力引起的势能能够保证 Snake 在运行过程中不会“陷入”某些局部最小点(如孤立的噪音点)或“逃离”某些可靠的特征点(如航空影象中建筑物的角点)。本文在进行目标提取过程中没有加入外部约束力。

2.2 原始 Snake 的缺陷及改进

从 $|v_s|^2$ 的表达式可以看出: Snake 的内部势能 $E_{int}(v)$ 中的连续性约束, 实际就是 Snake 上相邻两点之间距离的平方和。在优化求解过程中, 需要使 Snake 的总能量不断减小, 因此优化过程总是试图缩短 Snake 上的点间距。在某些情况下, 由于这个约束的作用, 可能会使 Snake 上的所有点都收缩到一点。因此, 采用这项约束作为连续项约束不是很恰当。Williams 等^[11]建议用 Snake 的点间距与其上的

点之间的平均距离之差的平方来代替两点之间的距离的平方作为内部约束的连续项约束。本文采用了 $\|v_i - v_{i-1}\|/d_0$ 作为连续性约束, 在 Snake 上的每一点, 这个能量值均被限制在略大于 1 或略小于 1 的范围内。 d_0 表示 Snake 上相邻两点之间距离的平均值, 也就是根据种子点进行其它点的内插时人工给定的点间距。

$\omega_2(s) |v_{ss}|^2$ 是对表示目标轮廓线的 Snake 的曲率的约束。在利用 Snake 方法提取影象中的目标时, 人工给定的种子点一般都是选择在目标的轮廓线上曲率较大的点附近, 在处理过程开始时, 首先计算各个种子点处的曲率的初始值 C_i^0 , 在处理过程中, 需要计算各个种子点处的曲率值, 将这个曲率值与该点的曲率初始值的差的平方作为该点的平滑性约束, 即: $|C_i - C_i^0|^2$ 。种子点的曲率值可以按下式计算^[11]

$$C_i = |\vec{u}_i / |\vec{u}_i| - \vec{u}_{i+1} / |\vec{u}_{i+1}||^2 \quad (5)$$

其中, C_i 代表第 i 点的曲率的当前值; \vec{u}_i 代表第 i 点与第 $i+1$ 点构成的向量。种子点移动的大致范围和种子点之间的距离是已知的, 由此, 通过计算可以得到种子点曲率变化的最大值, 种子点处的平滑性约束项的权系数 $\omega_2(s)$, 可以设置为该点曲率变化最大值的倒数; Snake 上每个内插点处的曲率值变化范围是大致相同的, 因此可取其最大值的倒数作为该点平滑性约束项的权系数 $\omega_2(s)$ 。

经过上述改进以后得到的 Snake 的内部势能 $E_{int}(v)$ 与原始的 Snake 模型相比更加合理。改进后的连续项可以保证处理过程中不会使 Snake 收缩为一点; 改进后的平滑性约束可以保证 snake 不会蜕化为一条直线。因此, 在进行目标提取过程中, 采用改进以后的内部势能, Snake 可以保持由人工给定的种子点得到的有关目标的大小与形状信息。另外, 在将 Snake 上每一点的连续性和平滑性约束的能量之和作为 Snake 内部约束的能量之前, 应该首先将其乘以 Snake 上总的点数的两倍的倒数, 这样就规定了 $E_{int}(v)$ 的数值的大致范围。

由影象特征引起的势能 $E_{Field}(v)$ 也作了改进。将梯度约束的能量表示为梯度沿 Snake 代表的曲线积分的均值的形式

$$E_{Gradient}(v) = 1 - \frac{K_{Gradient}}{LN} \oint Gradient(x, y) ds \quad (6)$$

其中, $E_{Gradient}(v)$ 表示梯度约束的总能量; $Gradient(x, y)$ 表示 Snake 上的点 (x, y) 处的影象梯

度值;LN 表示构成 Snake 的曲线的点数; K_{Gradient} 为一个对梯度值进行标准化的常数,对于有 256 个灰度级的影象来说 $K_{\text{Gradient}}=1/255$ 。

将灰度约束的能量表示为灰度沿 Snake 的曲线积分的均值的形式,即:

$$E_{\text{Grey}}(v) = 1 - \frac{K_{\text{Grey}}}{LN} \oint_S \text{Grey}(x, y) ds \quad (7)$$

其中, $E_{\text{Grey}}(v)$ 代表灰度约束的能量; $\text{Grey}(x, y)$ 表示 Snake 上的点 (x, y) 处影象的灰度值; LN 还是表示构成 Snake 的曲线的点数; K_{Grey} 为对灰度值进行标准化的常数,当影象为 256 个灰度级时, $K_{\text{Grey}}=1/255$ 。这种形式的约束适合于需要提取的轮廓线趋向于影象中目标亮度较大的部分;当需要提取的轮廓线趋向于影象中目标亮度较低的部分时,可采用下式表示的约束,即:

$$E_{\text{Grey}}(v) = \frac{K_{\text{Grey}}}{LN} \oint_S \text{Grey}(x, y) ds \quad (8)$$

各个符号的意义与式(7)相同。

改进后 Snake 的总能量表示为

$$E_{\text{snake}}(v) = E_{\text{int}}(v) + E_{\text{Grey}}(v) + E_{\text{Gradient}}(v) \quad (9)$$

$E_{\text{int}}(v)$ 应该采用以下的形式,即:

$$E_{\text{int}}(v) = \sum_i |(v_i - v_{i-1})/d_o|^2 + \sum_i |C_i - C_i^0|^2 \quad (10)$$

采用改进的影象势能约束后,无论是影象的梯度约束的能量,还是灰度约束的能量,其能量值都在 0 至 1 之间,从而避免了完全由经验来决定权值选取的困难。

3 动态规划的基本原理及 Snake 优化算法

动态规划(Dynamic programming)是解决多阶段决策过程最优化问题的一种有效方法,它是由美国数学家 R. Bellman 等人于 1951 年提出的。动态规划的基本思想是:把含有许多决策变量的问题分解成前后相关的 n 个阶段,由这 n 个阶段构成 n 个子问题,每个子问题仅含有一个或几个需要决策的变量,顺序地利用已解出的前一个子问题的结果,逐个求解后一个子问题,这样,最后一个子问题的解就是原问题的解。动态规划除了能把一个复杂的问题分解成 n 个容易求解的子问题逐个求解外,还能够

通过动态规划求出从初始状态到最终时刻的 n 阶段的最优解,得到从每个阶段的开始状态到最终时刻的各个子问题的最优解。动态规划方法已经用于处理计算机视觉问题^[12]。在利用动态规划处理航空影象中目标提取问题时,需要明确以下几个基本概念。

3.1 阶段变量

在求解过程中,将问题划分成相互联系并编有序号的 n 个阶段,取且仅取这些序号值的变量称为阶段变量。在利用 Snake 进行影象中目标提取时,假设包括种子点和在种子点之间加入的内插点在内共有 n 个点表示 Snake 的轮廓线,则对于湖泊、建筑物等类型的面状目标,求解过程可以分为 n 个阶段,阶段变量 k 的取值为 $1, 2, \dots, n$;而对于较细的道路、河流等线状目标可以分为 $n-1$ 个阶段,阶段变量 k 的取值为 $1, 2, \dots, n-1$ 。

3.2 状态变量

求解过程中某一阶段所具有的特征与状况称为状态,描述状态的变量称为状态变量。状态变量与阶段变量有关, $x_k=i$ 就表示求解过程在第 k 个阶段,且处于第 i 个状态下。在利用 Snake 进行影象中的目标提取时,构成 Snake 的每个点可以移动的位置称为所对应的阶段的状态。如:对于每个点都可在一定大小的窗口内移动的情况,假设窗口内的象素的个数为 m ,则状态变量的取值范围为 $1, 2, \dots, m$,即每个阶段都有 m 个状态。

3.3 决策变量

在某一阶段的一个状态给定后求解过程从该状态下一阶段的变化,将随着所作的决定(选择)的不同而达到下一个阶段中不同的状态,这种决定(选择)称为决策。描述决策的变量称为决策变量。用 $u_k(x_k)$ 表示第 k 阶段在状态 x_k 处的决策变量。决策变量的取值范围被称为允许决策集合,可以假设 $D_k(x_k)$ 表示第 k 阶段在状态 x_k 处的决策集合。

3.4 状态转移(变换)方程

动态规划研究的是由一个阶段到下一个阶段的过程。在给出某一阶段的状态 x_k 和该状态下的决策 $u_k(x_k)$ 后,就得到下一阶段状态 x_{k+1} ,有下列的变换关系,即:

$$x_{k+1} = T_k(x_k, u_k) \quad (11)$$

此式称为第 k 阶段的状态转移(或变换)方程。

3.5 策略

由第一阶段的出发点开始到最后阶段的终点为止的过程称为问题的全过程。由各个阶段的决策 $u_k(x_k)$, $k=1, 2, \dots, n$, 所组成的决策序列称为全过程策略, 简称策略, 记为 $P_{1,n}(x_1)$, 即:

$$p_{1,n}(x_1) = \{u_1(x_1), u_2(x_2), \dots, u_n(x_n)\} \quad (12)$$

由第 k 阶段开始的状态到第 n 阶段的终止状态的过程称为 k 子过程, 而相应的决策序列 $\{u_k(x_k), u_{k+1}(x_{k+1}), \dots, u_n(x_n)\}$ 称为 k 子过程策略, 简称子策略, 记为 $P_{k,n}(x_k)$, 则

$$P_{k,n}(x_k) = \{u_k(x_k), u_{k+1}(x_{k+1}), \dots, u_n(x_n)\} \quad (13)$$

K 子过程可供选择的策略称为允许策略集合, 记为 $P_{k,n}(x_k)$ 。

3.6 性能指标(目标)函数和最优指标函数

在多阶段决策过程的最优化问题中, 指标函数是用来衡量所实现过程的一种数量指标, 它是定义在全过程和所有子过程上的数值函数, 这个函数可以用 $V_{k,n}$ 表示, 即:

$$\begin{aligned} V_{k,n} &= V_{k,n}(x_k, u_k, x_{k+1}, u_{k+1}, \dots, x_n, u_n) \\ &= V_{k,n}[x_k, P_{k,n}(x_k)] \\ &k = 1, 2, \dots, n \end{aligned} \quad (14)$$

当初始状态与策略确定时, 指标函数的值也随之确定。性能指标函数的最优值称为最优指标函数, 记为 $f_k(x_k)$, 它与状态 x_k 有关。与最优指标函数值相应的策略称为最优策略。

3.7 阶段指标

当第 k 阶段的状态 x_k 与决策 $u_k(x_k)$ 的值取定后, 第 k 阶段对性能指标的贡献称为阶段指标, 可用 $g_k(x_k, u_k)$ 表示。

有了上述的基本概念以后, 就可以列出动态规划的函数基本方程

$$f_k(x_k) = \underset{u_k \in D_k(x_k)}{opt} \{g_k(x_k, u_k) + f_{k+1}(x_{k+1})\} \quad (15)$$

$$x_{k+1} = T_k(x_k, u_k) \quad k = n, n-1, \dots, 1 \quad (16)$$

$$f_{n+1}(x_{n+1}) = 0 \quad (17)$$

其中, opt 可以根据问题的性质取 \min 或 \max , 对于本文求 Snake 的最优解的问题来说, 由于能量的最小位置对应着 Snake 的最优解, 所以 opt 取 \min 。

在利用动态规划求 Snake 的最优解时, 如果有

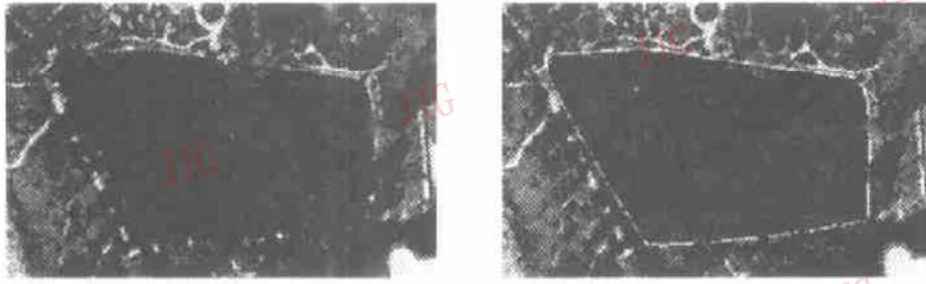
n 个阶段, 而每阶段有 m 个状态, 则整个求解过程的计算复杂性为 $O(nm^3)$ 。为了减小计算量, 需要减少 n 和 m 。由于 n 为构成 Snake 的曲线上的点数, 这个点数是根据目标提取的精度要求提出的, 所以会有一个最小限值; 而 m 为每个阶段状态的个数, 可以采用以每个点为中心的一定大小的窗口内的象元位置为候选点, 这样, 当窗口的大小为 5×5 时, 则 $m = 25$ 。本文设计的优化算法中, 采用了将 Snake 上的点沿与该点的切线垂直的方向移动, 这样就大大减少了计算量。

4 实验与结论

本文以改进后的 Snake 的能量函数为依据, 利用动态规划方法作为优化求解的手段, 进行了航空影像中的目标提取实验。实验是以航空影像中的湖泊与建筑物为例进行的, 由于要提取目标的光谱特性的差异, 对不同的目标灰度约束能量采用了不同的形式, 其中在进行湖泊的提取时, 灰度约束采用的是式(7)表示的能量函数; 在进行建筑物的提取时, 灰度约束采用的是式(8)表示的能量函数。图 1(a) 为航空影像中的湖泊, 影像中的十字表示湖泊提取时人工给定的种子点, 图 1(b) 为最后提取出的湖泊的轮廓线; 图 2(a) 为航空影像中的建筑物, 影像中的十字是在进行建筑物提取时人工给定的种子点, 图 2(b) 为提取出的建筑物的轮廓线。

利用上述方法进行航空影像中目标的提取有下列优点: (1) 由于在目标的提取过程中综合了各种影像信息和连续性、平滑性等约束, 因此, 即使在影像中的目标受到干扰(如某些部分被遮蔽)的情况下, 也能得到理想的处理结果; (2) 在目标提取开始阶段, 由人工给定的种子点可以获得目标的位置、形状等先验信息, 在这些信息的驱动之下进行目标提取时, 可以限制识别过程的搜索空间, 从而大大提高了处理的效率; (3) 改进后的 SNAKE 的能量函数, 解决了原始的 Snake 模型中参数选取的困难; (4) 动态规划方法用于优化求解能够保证得到问题的最优解, 而且其解具有很高的可靠性、稳定性及精度^[11]。尽管如此, 还有下面的问题没有解决: 动态规划方法的计算量大、所占内存空间也大, 尽管这个问题通过限制 Snake 上的点的移动方向得到了一定程度的解决; 在目标提取时没有引入对形状的整体约束, 特别是对房屋这类由直线边构成的规划目标, 整体形状约束的引入将会有利于影像中目标的提取; 种子点

还是由人工给定,而不是由高水平处理机制自动给出。



(a)原始影象及人工给定的种子点

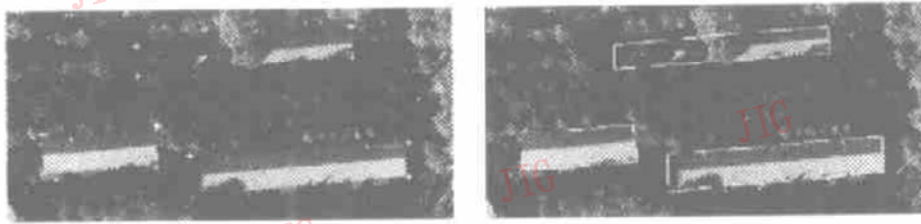
(b)提取结果

图1 由航空影象中提取的湖泊

Fig. 1. Lakes extracted from airphoto

(a)Original image and seeds

(b)Extracted lakes



(a)原始影象及人工给定的种子点

(b)提取结果

图2 由航空影象中提取的建筑物

Fig. 2. Buildings extracted from airphoto

(a)original image and seeds.

(b)Extracted buildings.

参考文献

- [1]Nagao M and Matsuyama T. A Structural Analysis of Complex Aerial Photographs. Plenum, New York, 1980.
- [2]Harlow C A, Traviedi M M, Connors R W and Phillips D. Scene Analysis of High Resolution Aerial Scenes *Optical Engineering*, 1986, Y Vol. 25, No. 3. PP. 345-362.
- [3]Davis L S and Hwang S V. The SIGMA Image Understanding System. IEEE 3rd workshop on Computer Vision, Representation and Control, Bellaire, Michigan, PP. 19-26, October, 12-16, 1995.
- [4]Terzopoulos D and Fleischer K. Deformable Models, Visual Computer, 1988, Vol. 4, pp. 306-331.
- [5]Kass M, Witkin A and Terzopoulos D. Snakes: Active Contour Models. *Inter. J. of Computer Vision*, 1987, Vol. 1, pp. 321-331.
- [6]Cohen L D. On Active Contour Models and Balloons. *CVGIP*, *Image Understanding*, 1991, Vol. 53, NO. 2, pp. 211-218.
- [7]Leymarie F and Levine M D. Tracking Deformable Objects in the plane using Active Contour Model. *IEEE Trans. on PAMI*, 1993, Vol. 15, No. 2, pp. 617-634.
- [8]Wang Y F and Wang J F. Surface Reconstruction Using Deformable Models with Interior and Boudary Constraints. *IEEE Trans. on PAMI*, 1992, Vol. 15, No. 7, pp. 572-579.
- [9]Whitten G Scale Space and Deformable Sheet Models for Computational Vision. *IEEE Trans. on PAME*, 1993, Vol. 15, No. 7, pp. 697-706.
- [10]刘少创,林宗坚. 航空影象分割的 snake 方法,武汉测绘科技大学学报, 1995, 20(1), 7-11.
- [11]Amini A A, Weymouth T E and Jain R C. Using Dynamic Programming for Solving Variational Problems in Vision. *IEEE Trans. on PAMI*, 1990, Vol. 12, No, 9, pp. 855-867.
- [12]Williams D J and Shah M A. A Fast Algorithm for Active Contours and Curvature Estimation. *CVGIP; Image Understanding*, 1992, Vol. 55, No. 1, pp. 14-26.



刘少创,武汉测绘科技大学遥感与信息工程学院博士生,现从事影像理解和遥感应用研究。

Dynamic Programming Based Feature Extraction of Aerial Imagery

Liu Shaocuang ,Lin Zongjian

*(School of Remote Sensing & Information Engineering
Wuhan Technical University of Surveying & Mapping, Wuhan, 430070)*

Abstract In this Paper, an improved Snake (Active Contour Models) is Presented and used for the purpose of feature extraction from aerial imagery. Feature extraction is formulated using dynamic programming as an optimized problem—solving Experimental results are presented to demonstrate the efficiency of the algorithm.

Keywords Aerial imagery, Objects identification, Dynamic Programming, Energy function