

一种视频序列中的运动对象自动分割算法^{*}

张晓燕 赵荣椿

(西北工业大学计算机学院 西安 710072)

摘要 提出了一种自动、准确的运动对象分割算法。首先通过直方图拟合获得准确的背景噪声方差,克服了以往只能依据经验设定背景噪声方差的缺点,并使用显著性测试技术有效地对帧差图进行二值化,确定出运动区域。然后进行形态学和对称差分处理消除噪声及显露背景,获得初始运动对象。但由于分割结果不够精确,再使用梯度向量流场作为外力的改进活动轮廓算法得到运动对象精确轮廓。实验结果表明,该方法能够得到运动对象精确的轮廓,并且具有调整参数少,抗干扰能力强,可并行处理等优点。

关键词 运动对象,背景噪声方差,帧差图,活动轮廓

A New Method for Moving Object Segmentation Automatically in Video Sequence

ZHANG Xiao-Yan ZHAO Rong-Chun

(College of Computer, Northwestern Polytechnical University, Xi'an 710072)

Abstract An automatic and accurate moving object segmentation algorithm is proposed in this paper. Firstly, to overcome the shortcoming of setting the value of background noise variance by experience, we estimate the value by histogram fitting. At the same time, we use the significance test to threshold the difference image and extract the moving areas. Then, the residual noise and the uncovered background due to the motions of objects are fast and effectively eliminated by morphological operations and intersection of two symmetrical moving areas. Thus the initial moving object and its contour can be obtained. However, the segmentation result is not very accurate. To solve this problem, an improved active contour which uses the gradient vector as the external force guides the initial contour moving to the actual video object contour. Experiment results testify that the proposed algorithm not only can obtain the closed and accurate video object contour but also has the properties of few parameters, robust to noise, and parallel processing.

Keywords Moving object, Background noise variance, Difference image, Active contour

1 引言

运动对象分割在视频序列分析中占有重要地位,其结果可以用于二维场景中运动物体特征描述,用于基于特征的视频编码,基于对象的视频控制,以及基于内容的视频提取等。然而,这些应用都需要获得准确的运动对象分割结果,轮廓边缘的错误会使其失去语义含义而导致无效。由此可见,对视频序列中的运动对象进行准确的分割是目前一个极其重要的研究领域。

目前已提出了许多算法用于从视频序列中分割运动对象,根据所采用的物理依据,可简单地分为两类。一类方法以空间属性为分割依据^[1,2],主要思想是首先根据灰度或色彩的相似性将视频的每一帧划分成若干区域,然后估计各区域的运动参数并按运动相似性对区域进行合并,从而实现运动对象分割。这种方法虽然能够得到运动对象准确的边缘轮廓,但运算量巨大,分割速度慢,而且当对象的各部分存在不一致运动时,很难得到完整的运动对象。另一类以时序属性为分割依据^[3,4],它们根据视频的时序属性(如帧间差、光流等)来分割运动对象,其中时序帧差法由于分割效率高而成为最常用的方法,但仅依据时序信息却难以获得准确的对象轮廓。

Snakes,或称为主动轮廓,在图像分割中得到了广泛的应用。它能够获得闭合的对象轮廓。但是传统的主动轮廓存在着过于依赖初始轮廓及不能很好收缩于凹向轮廓的缺点,极大地限制了其在运动对象分割中的应用。

为了能够高效、准确地分割视频序列中的运动对象,本文提出了一种基于显著性测试及改进活动轮廓的运动对象分割算法。该方法对能够获得准确的背景噪声方差,使用显著性测试方法检测出对象的运动区域。通过形态学处理及对称差分法有效地消除噪声及对象运动显露出的背景区域,获得基本准确的初始运动对象。最后采用梯度向量流场作为外力的改进活动轮廓对运动对象轮廓进行修正,有效地克服了传统梯度势能场的弱点,得到了运动对象精准的外轮廓。本文方法具有潜在的并行机制,能够提高分割效率。

2 获取运动区域

2.1 背景噪声方差估计

基于显著性测试获取运动区域时,需要使用背景噪声方差 σ^2 。但是背景噪声方差 σ^2 通常只能依据经验设定,为了克服这个缺点,获得更为准确的统计测试量以对帧差图进行有效的二值化,我们用直方图拟合背景上帧差的零均值 Gaussi-

^{*} 本文得到了国家自然科学基金(60141002)和南昌航院测控中心开放实验室基金(KG200104001)的支持。张晓燕 博士研究生,主要研究方向为序列图像处理,模式识别和计算机视觉。

an 分布,求得准确的背景噪声方差值。

令 $h(d), d=0, \dots, 255$ 为帧差图的直方图,每个帧差值 d' ($d' \in [0, 255]$) 的概率

$$p(d') = h(d') / \sum_{d=0}^{255} h(d) \quad (1)$$

假设背景上帧差的最大值为 d_T ($d_T \in [0, 255]$) 且 $h(d_T) \neq 0$ 用直方图 $h(d)$ 拟合背景上方差的分布,得到背景的概率 $p_B(d_T)$, 均值 $\mu_B(d_T)$ 和方差 $\sigma_B^2(d_T)$ 为

$$p_B(d_T) = \sum_{d=0}^{d_T} h(d) / \sum_{d=0}^{255} h(d) \quad (2)$$

$$\mu_B(d_T) = 0 \quad (3)$$

$$\sigma_B^2(d_T) = \sum_{d=0}^{d_T} d^2 \times h(d) / \sum_{d=0}^{d_T} h(d) \quad (4)$$

帧差值 d' 为背景的条件概率为

$$p(d' | B, d_T) = \frac{1}{\sqrt{2\pi\sigma_B^2(d_T)}} e^{-d'^2 / 2\sigma_B^2(d_T)} \quad (5)$$

定义拟合的目标函数为

$$err = \sum_{d=0}^{255} |p(d' | B, d_T) \times p_B(d_T) - p(d')| \quad (6)$$

当 err 的值最小时,直方图与背景帧差的分布间的拟合效果最佳,令 d_B 为使得 err 最小的 d_T 值,则

$$d_B = \arg \min_{d_T} err \quad (7)$$

取与 d_B 对应的 σ_B^2 为最佳的背景噪声方差值,即

$$\sigma^2 = \sigma_B^2 = \sum_{d=0}^{d_B} d^2 \times h(d) / \sum_{d=0}^{d_B} h(d) \quad (8)$$

2.2 基于显著性测试获取运动区域

假设相邻两帧在位置 (x_i, y_i) 上没发生变化(即零假设 H_0),则 (x_i, y_i) 上的帧差 $d_k(x_i, y_i)$ 服从零均值 Gaussian 分布^[5]

$$p(d_k(x_i, y_i) | H_0) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-d_k^2(x_i, y_i) / 2\sigma^2} \quad (9)$$

在以 (x_i, y_i) 为中心,大小为 $n \times n$ 的窗口 $W(x_i, y_i)$ 内计算标准化帧差 d_k / σ 的平方和

$$\Delta_k(x_i, y_i) = \frac{1}{\sigma^2} \sum_{(x,y) \in W(x_i, y_i)} d_k^2(x, y) \quad (10)$$

并以此作为测试统计量,假设两帧图像在窗口 $W(x_i, y_i)$ 内未发生变化,则标准化后的帧差 d_k / σ 服从 $N(0, 1)$ Gaussian 分布且空间无关,所以 d_k / σ 的平方和 $\Delta_k(x_i, y_i)$ 服从 χ^2 分

布^[5],其自由度等于窗口 $W(x_i, y_i)$ 内像素的数目 n^2 。

根据显著性测试理论,阈值 T_α 由给定的显著性水平 $\alpha = p(\Delta_k(x_i, y_i) > T_\alpha | H_0)$ 和 $\Delta_k(x_i, y_i)$ 所服从的分布决定。若 $\Delta_k(x_i, y_i)$ 小于阈值 T_α ,则认为位置 (x_i, y_i) 属于背景区域;反之认为它属于运动区域。 α 的取值与图像噪声相关,图像噪声小, α 的取值应较高,反之,应取较低的 α 值。在实验中,取 $\alpha = 10^{-2}$,并将窗口 $W(x_i, y_i)$ 设为 5×5 ,得到较理想效果。

3 获取初始运动对象及轮廓

使用上述方法虽然可分离出运动的区域与背景,但提取的运动区域模糊、不完整,而且还存在一些孤立噪声点。首先使用连通组建分析去除噪声小区域,然后进行形态运算,去除一些细小的突出,填补一些小洞和缝隙,融合边界上一些较窄的缺口,得到完整的运动区域。由于运动区域包含了运动对象区域以及遮挡区域,因此我们利用对称差分法来提取中间帧运动对象。图 1 显示了对称差分法获取运动对象原理图。

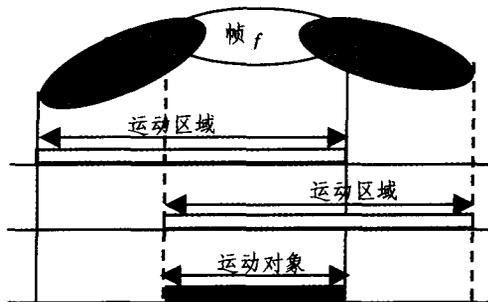


图 1 对称差分法获取运动对象原理图

从图中可看出,利用当前帧与前后帧分别求出帧差图并分别进行显著性测试提取出前后两个运动区域,由于它们都包含有当前帧运动对象以及不同的显露背景,因此对运动区域取交集即可获得当前帧运动对象。

将运动对象的初始分割模板与原图进行“与”操作获得初始运动对象,用形态学方法去除内部像素得到运动对象轮廓。图 2 显示了 Trevor 序列用上述方法获取的初始运动对象及轮廓。我们发现提取出的轮廓准确性不高,将在后续步骤中进一步提高分割准确性。



图 2 Trevor 序列获取的初始运动对象及轮廓

4 基于改进活动轮廓的精确分割

4.1 改进的活动轮廓

活动轮廓是定义在图像范围内的一条曲线或一个表面,它能够曲线或表面本身的内力影响和由图像产生的外力影响共同作用下发生变形。形变的最终目的是使曲线或平面本身与目标物体的边缘或者是图像中期望检测到的特征形

状相一致。

第一个活动轮廓的模型由 Kass^[6] 在 1987 年提出,其能量表达式:

$$E = \int_0^L \{E_{int}(X(s)) + E_{ext}(X(s))\} ds \quad (11)$$

其中, S 表示弧长的参数,活动轮廓在图像中的位置通过参数向量 $X(s) = [x(s), y(s)]$ 表示, E_{int} 表示曲线本身内部的力量,它使得曲线伸缩、弯曲。其定义为:

$$E_{int} = \frac{\alpha |X'(s)|^2 + \beta |X''(s)|^2}{2}, s \in [0, 1] \quad (12)$$

式中, $X'(s)$ 表示参数向量 $X(s)$ 关于 s 的一阶微分, $X''(s)$ 表示参数向量 $X(s)$ 关于 s 的二阶微分。 α 为弹力系数, β 为强度系数。

E_{ext} 表示由于图像的性质而产生的力量, 它使得曲线向着目标移动。

当曲线到达目标物体(通常是物体的边缘线)的时候, 期望能量函数也应达到较小的值。根据(11)式和(12)式, 我们可以写出完整的能量表达式:

$$E = \int_0^1 \frac{1}{2} (\alpha |X'(s)|^2 + \beta |X''(s)|^2) + E_{ext}(X(s)) ds \quad (13)$$

可变轮廓线若使得能量 E 最小, 应当满足欧拉方程, 于是有:

$$\alpha X''(s) - \beta X''''(s) - \nabla E_{ext} = 0 \quad (14)$$

为了对方程(14)进行动态的求解, 可以把 X 看作是时间 t 以及弧长 s 的函数。于是, X 关于 t 的偏微分等于(14)式的左端, 即:

$$X_t(s, t) = \alpha X''(s, t) + \beta X''''(s, t) - \nabla E_{ext} \quad (15)$$

当方程的解 $X(s, t)$ 趋于平稳时, $X_t(s, t)$ 趋于零, 此时就可以得到方程(14)的解。

在传统的活动轮廓中, 外力通常由边缘点的梯度势能场给出。导致力量场的有效捕捉范围很小, 而且使得缓变区域不能够获得外力。采用梯度向量流场 $V(X)$, 来替代式(15)中的势能函数 $-\nabla E_{ext}$, 可得:

$$X_t(s, t) = \alpha X''(s, t) + \beta X''''(s, t) + V(X) \quad (16)$$

称通过解上述动态方程而得到的参数曲线为梯度向量流(Gradient Vector Flow, GVF)可变轮廓^[7]。

梯度向量流场的向量 $V(x, y) = (u(x, y), v(x, y))$ 应当使得下式的能量最小:

$$\epsilon \int \int (\mu (u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2) dx dy \quad (17)$$

式(17)右端第一项为平滑项, 它可以使得向量场缓慢变化; 第二项为数据项, 它使得向量场 V 尽量地接近从图像数据中计算得到的 ∇f 。要使得整体能量最小, 应取 $V = \nabla f$ 。

这就满足了设计目标: 当向量场的值较大时, 向量 V 尽量保持与边界图的梯度相一致, 而在图像的缓变区, 向量 V 的变换较平缓。

利用变分, 解以下的 Euler 方程可以得到梯度向量流场,

$$\begin{aligned} \mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) &= 0 \\ \mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) &= 0 \end{aligned} \quad (18)$$

其中, ∇^2 为 Laplace 算子。式(18)对于梯度向量流场提供了更加直观的理解, 注意到在均匀区域中, 因为 $f(x, y)$ 的梯度为零, 所以每个方程的第二项都是零。因此在这样的区域内, u, v 由 Laplace 方程所决定, 从而使梯度向量流场从该区域的边界插值获得, 并且在一定程度上反映了边界向量的竞争, 因此梯度向量流场可以生成指向凹边界的向量。

4.2 轮廓精确化

首先对需分割的图像使用文[8]的方法进行边缘检测, 该边缘检测算法适合于检测强边缘点并去除了—些边缘短线, 正好符合视频对象边缘轮廓线的特点。计算出该边缘图像的 GVF 场。然后以前文中获取的初始运动对象轮廓曲线为初始轮廓, 进行迭代收缩, 获得运动视频对象精确的轮廓曲线。图3显示了 Trevor 第6帧的边缘图像, 轮廓修正过程以及轮廓精确化后分割结果。从图中可以看出, 除了左手肘部受桌子边缘影响外, 其他部分的轮廓都更加准确。

5 实验结果

用本文方法还对标准序列 Alex 和 Table tennis 进行了检测, 试验结果如图4和图5所示。Alex 序列是背景静止且较复杂的头肩序列, 从图4可以看出, 初始分割结果较准确, 所以只需要很少的迭代收缩就可得到准确轮廓。但由于手指轮廓受到桌子上物体边缘轮廓的影响, 因此稍有偏差。图5的 Table tennis 序列获得的初始视频对象轮廓较大, 不准确。但是经过改进活动轮廓精确化后仍然得到了准确的结果。从图中可看出, 本文算法对单个运动对象和多个运动对象都能得到准确的分割结果。



图3 Trevor 序列轮廓精确化过程

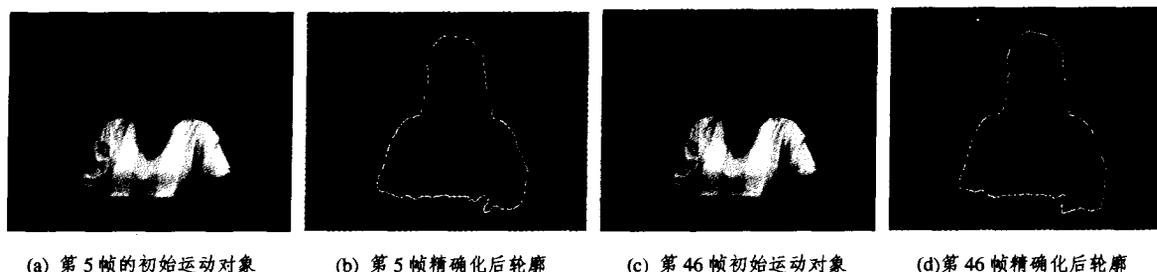


图4 Alex 序列实验结果

其中 $E(x, t)$ 是第 x 个叶子在生长周期 t 时刻产生的潜在的物质, $S(x, t)$ 是第 x 个叶子在生长周期 t 时刻对能源的阻碍, r_1 是单位叶片面积的抵抗系数, r_2 是单位叶柄, 叶脉等抵抗系数。一般的取值为 $r_1 = 800000, r_2 = 0.96$ 。

$$E(x, t) = \begin{cases} E_m \frac{\phi(x, t)}{\phi_m} & \phi(x, t) < \phi_m \\ E_m & \phi(x, t) \geq \phi_m \end{cases}$$

其中 ϕ_m 是阈值, E_m 是生产物质的经验值, $\phi(x, t)$ 是吸收的光能量。

2) 光线模拟。设 E_j 是光源 j 对叶片的光照能量, $I(E_j)$ 是光源 E_j 对于叶片的能量的贡献系数, $b_i = \sum_{E_j \in P_j} I(E_j) E_j$, b_i 是叶面 i 的光接收能量。这样光接收的能量将决定于叶片的面积的大小。叶片的面积大小计算为: 通过由 B 样条控制点构成的四边形网格, 我们取由四边形网格构成的面积来代替叶片的面积, 这样只要求得每一个四边形网格的法向量, 并计算出四边形的面积, 就可以得出垂直光线方向的面积。通过累加可以得到叶片的面积接收到光能量:

$$E_i = \sum_{k=1}^n n_k(u, w) \times n_i s_k$$

其中 $n_k(u, w)$ 是曲面控制点四边形的对称中心法向量, n_i 是第 i 个光源与控制点四边形对称中心的连线方向矢量。 s_k 是控制点四边形构成的面积。

高精度的情况下, 每个叶子的产生的能量不仅取决于叶的面积, 同时涉及到叶片的反射和透射程度^[1]。

4 实验结果

在 java3d 软件包的基础上通过程序实现来验证上述算法的可行性, 我们模拟一种花树的不同成长过程的几幅图像

如下列出。

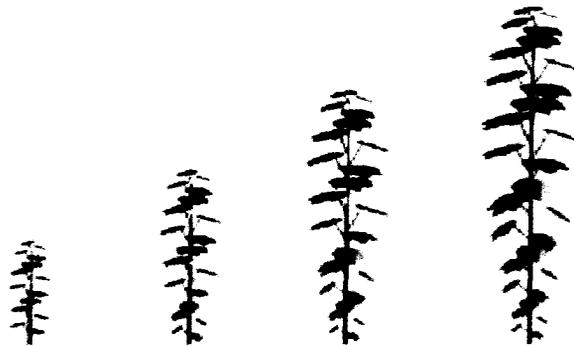
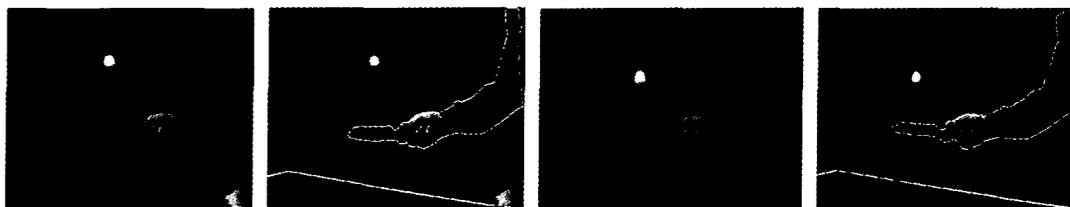


图 4 花树的生长过程模拟

参考文献

- 1 Soler C, Sillion F, Blaise F, et al. A physiological plant growth simulation engine based on accurate radiant energy transfer: [Rapport de recherche n° 4116]. Fevrier 2001
- 2 范惠琳, 等. 基于分支矩阵的计算机生成树. 华侨大学学报, 1994, 15(2)
- 3 张有会, 等. 自然树木形状的分形模拟. 计算机工程, 2000, 26(9)
- 4 齐东旭. 分形及其计算机生成. 科学出版社, 1996
- 5 Hammel M S, Prusinkiewicz P, Wyvill B. Modelling Compound Leaves Using Implicit Contours
- 6 赵星, 熊范纶, 等. 虚拟植物生长的双尺度自动机模型. 计算机学报, 2001, 24(6)
- 7 de Reffye Ph, Edelin C, Françon J, et al. Plant models faithful to botanical structure and development. Computer Graphics (Proceedings of SIGGRAPH 88), 1988, 22(4): 151~158
- 8 冯金辉, 陈彦云, 等. 树在风中的摇曳—基于动力学的计算机动画. 计算机学报, 1998, 21(9)
- 9 张丽艳, 周来水, 周儒荣. 裁剪 NURBS 曲面面积计算的扫描分片积分法. 工程图学学报, 1997(2-3)

(上接第 241 页)



(a) 第 2 帧的初始运动对象 (b) 第 2 帧精确化后轮廓 (c) 第 13 帧初始运动对象 (d) 第 13 帧精确化后轮廓

图 5 Table tennis 序列实验结果

总结 本文提出了一种基于显著性测试及改进活动轮廓的运动对象分割算法。首先通过直方图拟合获取背景噪声方差, 再基于显著性检测、对称差分以及形态学处理获取初始运动对象及轮廓。但是此时获取的运动对象轮廓与实际轮廓还有一些误差。为了提高分割准确性, 我们对原始图像检测出边缘图像, 并计算出边缘图像的梯度向量流场, 使用梯度向量流场为外力的改进活动轮廓算法对轮廓进行修正。由于运动对象初始轮廓曲线的确定比较准确, 而且改进的活动轮廓算法能够收缩于轮廓的凹边界, 因此得到了更准确的运动对象轮廓曲线。本算法的计算可以采用并行的机制来完成, 算法的运行速度不与帧数成正比, 提高了分割效率。从试验来看, 我们的方法可以实现单个运动对象以及多个运动对象的分割, 结果令人基本满意。本文中的实验都是静态背景的视频序列, 对于有些动态背景的视频序列, 进行全局运动补偿后也可用本算法进行分割。但是对于复杂背景序列, 如何加强边缘图像中的对象轮廓以及去除背景边缘, 从而减少背景边缘

在轮廓精确化过程中的影响, 是我们需要进一步研究的问题。

参考文献

- 1 刘志, 杨杰, 彭宁嵩. 基于假设检验和区域合并的视频对象分割. 数据采集与处理, 2004, 19(2): 124~129
- 2 Mezaris V, Strintzis M G. Video object segmentation using Bayes-based temporal tracking and trajectory-based region merging. IEEE Trans. Circuits and Systems for Video Technology, 2004, 14(6): 782~795
- 3 Kim C, Hwang J N. Fast and Automatic Video Object Segmentation and Tracking for Content-based Application. IEEE Trans on Circuits and Systems for Video Technology, 2002, 12(2): 122~129
- 4 杨莉, 张弘, 李玉山. 视频运动对象的自动分割. 计算机辅助设计与图形学学报, 2004, 16(3): 301~306
- 5 Aach T, Kaup A, Mester R. Statistical model-based change detection in moving video. Signal Processing, 1993, 31(2): 165~180
- 6 Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. International Journal Computer Vision, 1988, 1: 321~331
- 7 Xu C Y, Princ J L. Snake, shapes and gradient vector flow. IEEE Tran. Image Processing, 1998, 7(3): 359~369
- 8 蒋晓悦, 赵荣椿. B-样条子波在图象边缘检测中的应用. 中国体视学与图象分析, 2002, 7(4): 198~201