

文章编号:1002-2082(2009)03-0448-06

基于自适应非参数统计模型的 彩色目标跟踪算法研究

惠宝聚¹, 高雅², 李良福^{2,3}

(1. 海装舰船技术保障部, 北京 100841; 2. 西安应用光学研究所, 陕西 西安 710065;
3. 西安交通大学 系统工程研究所, 陕西 西安 710049)

摘要: 针对复杂环境条件下的视觉跟踪问题, 提出一种基于自适应非参数统计模型的彩色目标跟踪算法。利用目标和背景之间的强度差别, 基于自适应核密度估计模型对运动目标进行了非参数统计建模。为了实现具有较大范围运动目标的跟踪, 在充分考虑目标和背景之间的相关性前提下, 采用目标特征统计的背景加权直方图对搜索区域进行了扩大。为了提高对环境变化的适应能力, 根据目标和环境的变化自适应更新目标特征分布模型。通过对实际图像序列的实验, 结果表明该算法能够有效跟踪运动目标, 并且平均迭代次数比传统方法减少了37.28%。

关键词: 视觉跟踪; 非参数统计模型; 彩色目标

中图分类号: TP391

文献标志码: A

Visual tracking algorithm for colour objects based on adaptive nonparametric statistical model

HUI Bao-ju¹, GAO Ya², LI Liang-fu^{2,3}

(1. Ship Technical Support Division of Naval Equipment Department, Beijing 100841, China;
2. Xi'an Institute of Applied Optics, Xi'an 710065, China;
3. System Engineering Institute, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To realize the visual tracking under the complicated condition, an efficient color object tracking algorithm based on the adaptive is presented in this paper. Based on an adaptive kernel model, the nonparametric statistical modeling of a moving target was carried out with the intensity difference between the target and the background. The search region is extended for searching objects with the background-weighted histogram for statistics of the target feature on the premise of taking the relevance between the target and background into account in order to realize the tracking of the moving gargets in a large area. According to the change of the object and environment, the target model is updated to improve the adaptive ability for environment variation of object tracking. Experimental results on real image sequences show that the algorithm can efficiently track the moving gargets, and the average iteration number reduces 37.28% in comparison with other method.

Key words: visual tracking; nonparametric statistical model; colour object

收稿日期:2009-01-21; 修回日期:2009-02-28

作者简介:惠宝聚(1965—),男,山东沂水县人,工程师,主要从事舰艇武器装备的技术保障研究工作。

引言

现实生活之中,大量有意义的视觉信息包含在运动图像序列之中,甚至有些动物的眼睛经过进化,只能看见运动的物体。视觉跟踪是计算机视觉的重要研究方向,在图像制导、交通监控、机器人视觉导航等领域中都得到了广泛的应用^[1-2]。视觉跟踪是指对包含有目标及背景的序列图像使用跟踪算法提取运动目标并预测它的运动,将图像序列之间同一运动目标关联起来,从而得到运动目标完整的运动轨迹。视觉跟踪技术是以计算机视觉技术为核心,结合了图像处理、计算机、模式识别、人工智能等众多领域的前沿技术,其难点在于快速、可靠和准确地实现目标的跟踪。

最近几年很多学者进行了视觉跟踪技术研究,并提出了许多算法,可以把它们大致地分成由下至上(bottom-up)和由上至下(top-down)的研究方法^[3]。由上至下的研究方法需要基于目标模型通过图像量测数据对状态进行假设检验。这种方法对图像分析的依赖性较少,因为目标的假设能够为图像分析提供强有力的约束,但是它的性能是由产生和检验这种假设的方法所决定。为了达到可靠而有效的跟踪,必须进行一系列的假设,这样就要涉及到更多的计算。而由下至上的方法是通过分析图像的内容对目标进行建模来计算目标的状态,它又称为基于数据驱动(data-driven)的方法。例如,通过曲线拟合来重建参数形状。这种方法不依赖于先验知识,处理效率较高,然而这种算法很大程度上取决于图像分析的能力,因为图像像素的拟合、聚类和轮廓描述等处理可能被杂波和噪声所干扰。为了描述目标的形状与状态,以区别于其他目标,目标模型描述是一个非常重要的问题。许多参数统计技术已经被用于目标模型描述。基于参数估计的方法是假设目标模型满足某种特定分布形式,因此需要对参数进行统计分析,以实现实时跟踪。而非参数统计方法不取决于目标模型的特定分布形式,目标分布不是通常遇到的典型分布,因此更为实用。这种方法利用图像数据进行引导使得匹配函数达到最优解,不受参数模型的限制^[4]。

与灰度特征相比,颜色特征具有更为丰富的感知信息,是计算机视觉中的重要特征。针对颜色特征的直方图统计方法能够给出目标特征的概貌性描述,因此在复杂的背景中具有良好的鲁棒性,并且具有统计特征旋转不变性和尺度不变性等优点^[5-6]。核密度估计是另一种非参数统计方法,已经

被用作重要的数据分析工具^[7]。利用颜色直方图与空间核函数的优点,Comaniciu^[8-9]提出了一种基于Mean shift的非参数跟踪方法(MST)。它利用Bhattacharyya系数作为对目标模板和候选模板的相似性度量,利用梯度优化方法来减少特征匹配搜索的计算时间,以实现快速的目标定位。然而,这种方法采用传统的核函数不能够适应颜色分布直方图的变化^[10],而且不能够跟踪两帧之间具有很大运动范围的目标。

在具有杂波、粘连、遮挡、光照变化等复杂环境下实现准确的目标跟踪是计算机视觉的一个难题^[11],在复杂环境中跟踪大范围运动的非刚性目标具有重要的研究意义。在大部分情况下目标特征可能混杂着部分背景,背景信息的利用也扩大了搜索范围,因此背景信息对于特征分析来说非常重要。本文提出了一种有效的视觉跟踪算法(EVT),它结合了背景加权直方图和自适应核密度估计的优点,根据局部直方图建立了自适应核估计模型。实验结果验证了本文算法能够对具有遮挡的大范围运动目标进行快速而且准确的视觉跟踪与定位。

1 自适应核密度估计模型

核密度估计和Parzen窗估计是非参数估计的常用方法。给定一个具有密度函数 $f(x)$ 的 d 维样本 X_1, X_2, \dots, X_n ,则其在 x 处的Parzen窗密度 $\hat{f}(x)$ 的表达式为

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\frac{x - X_i}{h}\right) \quad (1)$$

式中: $k(\cdot)$ 是核函数; h 是核窗宽。通常满足条件 $\int k(x)dx = 1, k(\cdot)$ 为对称函数,即 $k(x) = k(-x)$ 。

假设 V 是图像中需要处理的像素集合, x 为 V 中正在处理的像素, $I(x)$ 是该像素的颜色值。通常,一个平核函数可以定义为:

$$k(x) = \begin{cases} 1, & \text{if } \|x\| \leq \eta \\ 0, & \text{其他} \end{cases} \quad (2)$$

式中 η 是图像空间中该邻域的半径。同样,高斯核函数可以定义为

$$k(x) = e^{-\|x\|^2} \quad (3)$$

由于每一个像素把整个图像看作它的邻域,高斯核函数难以进行实际应用。因此可以把平核函数和高斯核函数结合起来,组成一个截断高斯核函数:

$$k(x) = \begin{cases} e^{-\varphi \|x\|^2}, & \text{if } \|x\| \leq \eta \\ 0, & \text{其他} \end{cases} \quad (4)$$

这个核函数对图像空间中所有点的处理方式完全一样。在其邻域内具有足够多相似强度像素的条件下,它能够保留所有的目标,反之则不行。通过修改以上核函数以适应不同的局部直方图,我们可以得到:

$$k(x) = \begin{cases} e^{-\varphi(S) \|x\|^2}, & \text{if } \|x\| \leq \eta \\ 0, & \text{其他} \end{cases} \quad (5)$$

式中: S 是 x 邻域中所有像素的集合,满足 $S \subseteq V$, 函数 $\varphi(S)$ 可以定义为

$$\varphi(S) = \Gamma \frac{\sum_{u \in U} \Psi(|I(u) - I(x)|)}{M} \quad (6)$$

式中 M 是集合 S 中像素的数目, Γ 是一个常量, u 为特征向量。函数 $\Psi(\cdot)$ 满足:

$$\Psi(x) = \begin{cases} 1, & \text{if } |x| \leq \zeta \\ 0, & \text{其他} \end{cases} \quad (7)$$

式中 ζ 是给定的阈值。当 $\Psi(x)$ 减小时,它像一个平核,平核趋向于平滑掉弱小的目标。但是在 η 内的一些局外像素点将向邻域中高强度的方向运动。在图像中噪声像素之间的强度差别很大,它们所包含的区域也很大,因此具有较低的 $\varphi(S)$ 值。需要跟踪的目标通常是由一个小的区域内具有类似强度的很多像素组成,因此 $\varphi(S)$ 将会较高。高的 $\varphi(S)$ 将使高斯核的标准差较高,由于远离目标中心的像素经过加权后强度变小,不允许背景像素影响目标像素的强度太大。很明显平核函数趋向于把背景混合于目标中,而高斯核趋向于更好地保留目标。因此,借助目标和背景之间的强度差别,利用截断高斯核函数能够对目标进行自适应建模。

2 基于背景加权直方图的跟踪算法

在多数应用情况下,由于目标中可能包含背景,通常很难精确描述目标模型。使用背景信息不准确可能影响跟踪效率,从而无法度量目标和候选区域之间的相似性。因此,我们的方法是使用目标和候选区域之间相对比较重要的背景信息进行相似性度量。

2.1 目标模型的描述

假定在图像中选择目标区域像素位置 $\{x_i\}_{i=1,2,\dots,n}$, $\mu = \{1, 2, \dots, m\}$ 表示目标模型颜色特征的直方图,函数 $b(x_i, t): R^2 \rightarrow \mu$ 表示时间 t 位置 x_i 处的颜色特征直方图。设 $\{g_u\}_{u=1,2,\dots,m}$ 表示满

足归一化约束 $\sum_{u=1}^m g_u = 1$ 的背景颜色特征分布。假定 g_u 的最小非零值是 g_s , 背景的面积等于目标面积的两倍,因此可以得到背景因子:

$$\{d_u = \min(\frac{g_s}{g_u}, 1)\}_{u=1,2,\dots,m} \quad (8)$$

这些背景因子使得具有较低 g_u 特征的重要性降低,尤其在背景区域中。根据(8)式,目标区域的核加权直方图的概率密度分布模型可以表示为

$$q_u = C d_u \sum_{i=1}^n k(x_i - a) \delta[b(x_i, t) - u] \quad (9)$$

式中: a 是核的中心位置; δ 是 Kronecker delta 函数。通过进行归一化约束 $\sum_{u=1}^m q_u = 1$ 可以得到常量 C :

$$C = \frac{1}{\sum_{i=1}^n k(x_i - a) \sum_{i=1}^m d_u \delta(b(x_i, t) - u)} \quad (10)$$

2.2 候选模型的描述

对于目标模型概率密度分布来说,我们可以采用更为紧密的表示方法。对于每一个特征向量 u , 我们可以把这些向量组合成一个矩阵 $U = [u_1, u_2, \dots, u_m]$ 。同样,我们可用 $K(a)$ 来定义核函数 $k(x_i - a)$, 用 D_u 来定义背景因子 $\{d_u\}_{u=1,2,\dots,m}$ 。通过采用这些简化后,我们可以把目标模式表示为更加简单的形式:

$$q = K(a)U(t)D \quad (11)$$

假定我们需要描述一个时间 t' 以 a 为中心的候选区域,则要求核加权直方图的概率密度分布模型为

$$p(a) = K(a)U(t')D \quad (12)$$

2.3 相似性度量

跟踪目标的目的是寻找感兴趣特征的最相似候选区域。处理方法:给定一个目标模型分布 q 和一个候选模型分布 $p(a)$,找到能够最大化目标模型分布 q 与候选模型分布 $p(a)$ 的位置。为了测量它们之间的相似性,Bhattacharyya 系数是一种有效和收敛型的统计度量工具,它可以通过最小化 2 种离散分布之间的距离得到:

$$d(a) = \sqrt{1 - \rho(p(a), q)} \quad (13)$$

式中 $\rho(p(a), q)$ 是相似性度量函数。(13)式等价于最大化 Bhattacharyya 系数的样本估计:

$$\rho(a) = \rho(p(a), q) = \sum_{u=1}^m \sqrt{p_u(a)q_u} \quad (14)$$

在 $p(a_0)$ 处应用泰勒公式展开上式,可以得到 $\rho(a)$ 的一阶线性逼近:

$$\rho(a) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(a_0)q_u} + \frac{1}{2} \sum_{u=1}^m p_u(a) \sqrt{\frac{q_u}{p_u(a_0)}} \quad (15)$$

在连续两帧之间如果候选目标 $\{p_u(a)\}_{u=1,2,\dots,m}$ 没有发生剧烈运动,以上方程始终成立。由于 $p(a)$ 不取决于 a 的值,于是可以得到:

$$\rho(a) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(a_0)q_u} + \frac{C}{2} \sum_{i=1}^n w_i k(x_i - a) \quad (16)$$

式中

$$w_i = \sum_{u=1}^m d_u \sqrt{\frac{q_u}{p_u(a_0)}} \delta[b(x_i, t) - u] \quad (17)$$

通过搜索局部区域核密度估计的最大值,可以递归地计算当前位置 a_0 和下一个位置 a_1 的均值漂移向量,直到到达峰值:

$$a_1 = \frac{\sum_{i=1}^n x_i w_i g(x_i - a_0)}{\sum_{i=1}^n w_i g(x_i - a_0)} \quad (18)$$

式中: $g(x) = -k'(x)$ 。从上式可以看出,新的位置是样本点 $\{x_i\}_{i=1,2,\dots,n}$ 的加权质心。然而,这些样本点的权取决于两部分:第一部分是背景加权的颜色特征权值 w_i ;另一部分是从核函数 $k(x)$ 得到的权,对于远离目标的中心 a_0 的点赋予较低的权值。

2.4 目标模型的更新

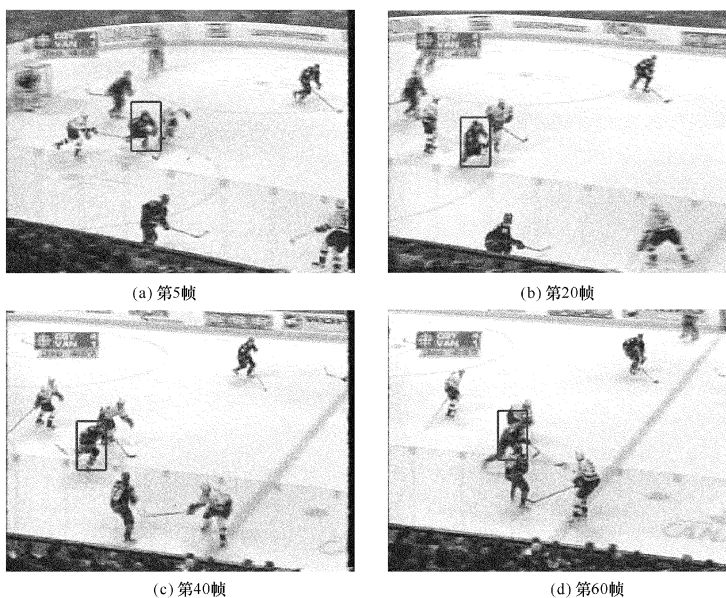
在多数情况下,目标特征的图像受光照变化、复杂背景和几何变形的影响,因此,为了实现有效的视觉跟踪,有必要更新目标模型。设 $\tau \in (0, 1)$ 为更新系数,更新后的目标特征模型 q^* 为

$$q^* = (1 - \tau)q + \tau p \quad (19)$$

更新系数 τ 是控制目标模型变化到候选模型速度的一个参数。当图像帧之间的光照变化较大时,应该选择较大的更新系数,因而,新的特征模型会更多地受到候选模型的影响。然而,太高的更新系数值也可能导致跟踪停止。当然,当环境变化较小时,应该选择较小的更新系数。

3 实验结果

实验采用分辨率为 320×240 像素的标准 Hockey 视频图像序列进行测试,采用 Matlab 语言在 Pentium IV 2.4 GHz 计算机上编写运行程序。在该图像序列中,大部分图像序列中的球员存在部分遮挡问题,因此应用传统的均值漂移方法得不到实际的跟踪位置,从而导致跟踪失效。在第一帧中手动确定需要跟踪的目标,目标初始尺寸选为 26×46 像素。图1表明球员在复杂环境中存在部分遮挡情况下能够准确跟踪目标,所跟踪的目标采用矩形框标绘出。



(a) 第5帧

(b) 第20帧

(c) 第40帧

(d) 第60帧

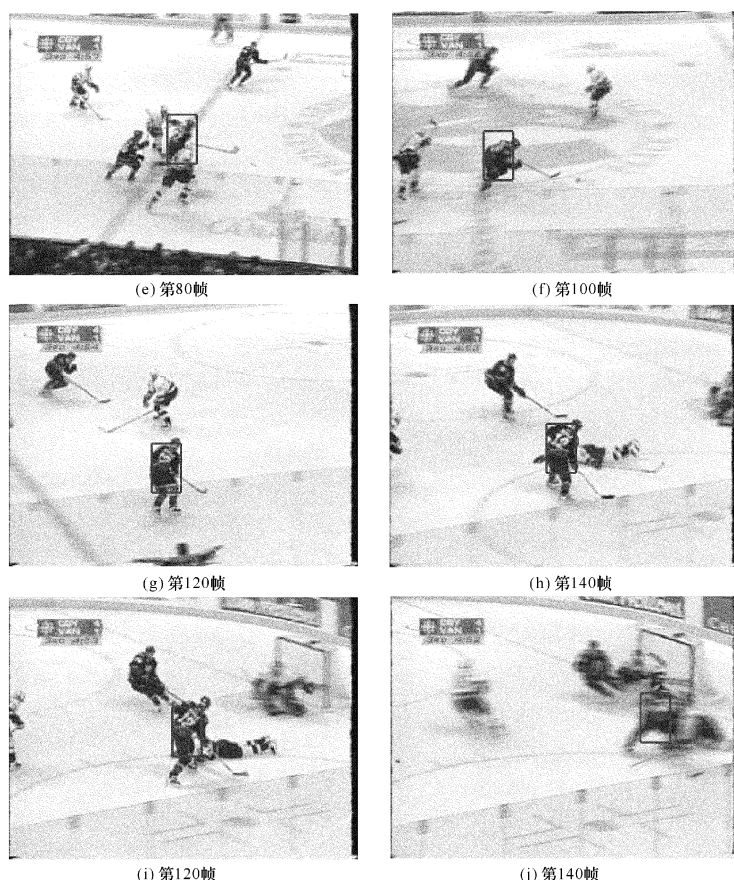


图1 Hockey 图像序列的跟踪结果

Fig. 1 Tracking results of Hockey image sequence

在图 2 中给出了 EVT 算法和传统的 MST 算法在每一帧中需要的迭代次数。传统的 MST 算法平均迭代次数为 6.41 次, 每次迭代的平均计算

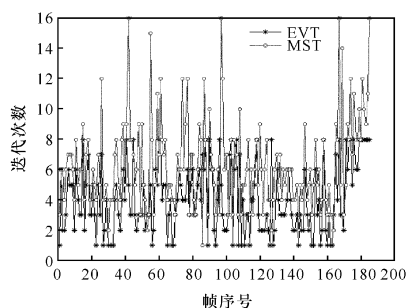


图2 EVT 算法和传统 MST 算法的迭代次数

Fig. 2 Iteration number needed by EVT and MST algorithm in each frame

时间为 0.027 9 s; 而 EVT 算法的平均迭代次数为 4.02 次, 每次迭代的平均计算时间为 0.028 6 s, 因此该算法能够在较短时间内收敛到局部最优解, 平均迭代次数比传统方法减少了 37.28%。对于上述视频图像, 传统的 MST 算法在跟踪时会丢失部分遮挡目标, 而 EVT 算法对于全部图像序列都成功实现了运动目标跟踪。因此, 本文所提出的算法在准确性和实时性方面均优于传统的方法。

4 结论

由于在复杂背景条件下很难精确区分目标与背景, 考虑了目标模型与背景因素的相关性, 提出了一种基于自适应非参数统计模型的视觉跟踪算法。通过建立自适应核密度估计模型, 能够在光照变化时变换模型, 在目标存在遮挡时提高抗干扰能

力。通过采用背景加权直方图目标特征统计方法扩大了运动目标的搜索区域,目标特征分布模型能够根据目标和环境的变化进行更新。在实际图像序列上的实验结果表明本文算法能够有效跟踪运动目标,并且跟踪速度明显优于传统方法。

参考文献:

- [1] 陈东炎, 张王己, 王艳玲, 等. 图像跟踪系统中机动目标预测的实现[J]. 应用光学, 2007, 28(1): 33-37.
CHEN Dong-yan, ZHANG Qi, WANG Yan-ling, et al. Implementation of maneuver target prediction in image tracking system [J]. Journal of Applied Optics, 2007, 28(1): 33-37. (in Chinese with an English abstract)
- [2] 陆培国, 寿少峻. 舰载光电系统高精度跟踪控制技术[J]. 应用光学, 2006, 27(6): 476-484.
LU Pei-guo, SHOU Shao-jun. High accuracy tracking technology and its application in ship-borne electro-optical system [J]. Journal of Applied Optics, 2006, 27(6): 476-484. (in Chinese with an English abstract)
- [3] ISARD M, BLAKE A. Condensation—conditional density propagation for visual tracking [J]. International Journal on Computer Vision, 1998, 29(1): 5-28.
- [4] JIMENEZ J R, MEDINA V, YANEZ O. Nonparametric MRI segmentation using mean shift and edge confidence maps [J]. SPIE 2003, 5032: 1433-1441.
- [5] COMANICIU D, RAMESH V, MEER P. Realtime tracking of non-rigid objects using mean shift [J]. Computer Vision and Pattern Recognition, 2000, 2: 142-149.
- [6] NUMMIARO K, KOLLER-MEIER E, GOOL L V. An adaptive color-based particle filter [J]. Image and Vision Computing, 2003(1): 99-110.
- [7] HAGER G D, DEWAN M, STEWART C V. Multiple kernel tracking with SSD [J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, 1: 790-797.
- [8] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [9] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [10] KING M A, LEE T K, ATKINS M S, MCLEAN D I. Automatic nevi segmentation using adaptive mean shift filters and feature analysis [J]. SPIE, 2004, 5370: 1730-1737.
- [11] ARNAUD E, MEMIN E. Optimal importance sampling for tracking in image sequences: application to point tracking [J]. ECCV, 2004, 3023: 302-314.