

单目标跟踪算法的研究与分析^①

王新洁

(大连医科大学 附属第一医院, 大连 116011)

摘要: 运动目标跟踪是计算机视觉的一个中心研究问题, 为视频内容的理解提供重要的信息。首先介绍了目标跟踪的国内外研究现状, 重点归纳分析了运动目标跟踪方法的分类及其发展过程中的提出的各种算法, 对其关键技术进行了剖析和比较。

关键词: 目标跟踪; 点跟踪; 核跟踪; 剪影跟踪

Progress in Object Tracking Algorithms

WANG Xin-Jie

(The First Affiliated Hospital, Dalian Medical University, Dalian 116011, China)

Abstract: Object tracking has long been one of the central research problems in computer vision, which provides important information to understand video content. Firstly, international research status of object tracking is introduced. Then especially, the categories of object tracking methods and various algorithms in the process of development are emphatically analyzed. The key techniques are compared and addressed, and some major problems and possible solutions are also summarized.

Key words: object tracking; point tracking; nuclear tracking; silhouette tracking

1 引言

随着数码电子设备以及数字存储设备的发展, 图像和视频在社会生活中得到越来越多的应用, 随之而来的是计算机视觉技术 (computer vision) 的蓬勃发展。使用计算机技术来帮助我们处理、理解, 分析图像和视频。计算机视觉扩展了人类视觉范围, 并可以长时间工作于恶劣的环境, 在智能监控、人机交互、机器人工程、视频检索等领域具有广阔的应用前景和潜在的商业价值。

基于视频的运动目标跟踪是计算机视觉中一个热门研究问题, 为视频内容的理解提供重要的信息。特别是在光照变化、噪音干扰、几何形变、遮挡等自然场景下的目标跟踪问题, 一直是研究领域的经典难题。

国内外很多研究机构已经对运动目标的跟踪问题开展了大量相关项目的研究。例如, 1997 年美国国防高级研究项目署设立的视觉监控重大项目 VSAM, 主人对雷达和光电仪器的多目标跟踪问题进行了深入的研究; 英国雷丁大学开展了对车辆和行人的跟踪及其

交互作用识别的相关研究; 德国法兰克福理工学院的 Nagel 等从事交通监控系统的研究。几十年来, 目标跟踪中无论是目标搜索算法还是特征选择与更新都不乏研究成果。

2 目标的描述

跟踪的场景中, 目标可以被定义为任何被用于将来分析的感兴趣事物。要跟踪一个目标, 首先要建立对该目标的描述, 以便于算法的识别。目标可以通过形状和表现来描述。

目标的形状定义了所关注的目标在图像平面的空间信息, 包括点 (用质心或者是点集来描述目标), 简单几何形状 (矩形、椭圆), 剪影和轮廓, 铰链模型, 骨架模型等。目标的表现定义了图像中目标的纹理、颜色等局部特征, 如基于统计的表现概率密度 (混合高斯模型、直方图), 模板, 多视角模型 (主成分分析

^① 基金项目: 安徽省教育厅自然科学基金(2005KJ004ZD)

收稿时间: 2011-10-29; 收到修改稿时间: 2011-12-26

(PCI)、独立成分分析^[5])等。处理跟踪问题时,形状描述通常和表观描述联合使用^[6]。

目标的描述方式依据跟踪的应用领域。对于那些在图像上很小的跟踪目标,合适采用点描述。对于形状可以用矩形或是椭圆形近似的物体,简单的几何形状描述更合适。对于跟踪形状复杂的目标,轮廓或是剪影描述更合适。

3 单目标跟踪

目标跟踪的目的是在视频的每一帧图像中确定目标的位置,从而产生目标随时间的轨迹。也可能同时提供目标在每一帧图像中所占的完整区域。通常使用目标检测算法得到图像中可能的目标区域,然后在各帧之间建立目标的对应联系。或者通过迭代更新前一帧图像得到的目标的区域和位置信息,联合估计出目标的区域和对应联系。依据其跟踪目标数目的不同,跟踪算法可分为单目标跟踪与多目标跟踪,本文主要论述单目标跟踪算法。目标模型的选择限制了该目标所能经历的运动和变形。依据所采用目标描述方式的不同,单目标跟踪可分为点跟踪、核跟踪和剪影跟踪,如图 1 所示。

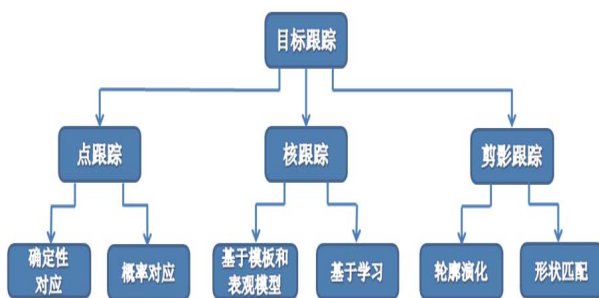


图 1

3.1 点跟踪

图像序列中检测到的目标被表示为其质心所在位置的点,不考虑目标姿态和形状的变化。跟踪过程可以表述为检测到目标在各帧之间的对应。点对应是一个复杂的问题,特别是存在遮挡,漏检,目标的进入和退出的情况。总体来说,点对应方法可以分为两大类,即确定性的和概率性的方法。确定性方法使用运动信息来约束对应问题。概率方法考虑不确定性,使用目标的测量值来建立关联。

(1) 确定性对应方法

点对应的确定性方法使用运动约束集合,定义先前帧中每个目标与当前帧中单个目标关联的相似度。最大化关联的相似度可以表述为一个组合优化问题。通过最优分配算法可以得到所有可行关联中的解^[7]。

(2) 概率对应方法

从视频传感器中得到的测量值包含噪声,目标的运动也会经受随机的扰动。概率对应方法在目标状态的估计时考虑了测量值和模型的不确定性,将目标位置、速度、加速度等性质表示为一个状态空间,认为某一时刻的目标状态为服从一定分布的随机向量。由检测算法得到物体在图像中的位置的概率分布。

状态估计的最优解可以通过迭代的贝叶斯滤波器(Bayesian Filter)^[8]得到,综合观测数据和目标的先验知识,由预测和更新两个步骤组成。建立目标状态的随机动力学模型,目标状态的概率密度在动力学模型的作用下进行扩散和伸展,得到下一时刻的预测值。动力学模型的随机部分导致概率密度伸展,使不确定性增加。确定性的部分造成概率密度的整体平移。观测值通过似然函数对状态的预测分布进行修正。似然函数定义目标状态与观测值的相似程度,描述了目标的观测模型。概率密度在测量值的附近趋向出现一个峰值。

当测量和预测过程中的噪声满足高斯分布,系统方程满足线性假设时,可以通过卡尔曼滤波(Kalman Filter)^[8]递推得到后验概率的均值和方差来描述状态估计的最优解。当线性假设不满足,通过泰勒展式(Taylor)对系统方程进行局部线性化,就得到了扩展卡尔曼滤波器(Extended Kalman Filter)^[9]。周东华等人对扩展卡尔曼滤波器增加了约束条件,要求真实状态值减去估计状态值的残差序列在任意两个时刻都保持正交。又在卡尔曼滤波的增益矩阵中引入了多重渐消因子,进而推导得到了强跟踪滤波器(Strong Tracking Filter)^[10]。对突变状态具有较好的跟踪能力,对模型参数的失配具有较强的鲁棒性。

为了解决扩展卡尔曼滤波器在处理非线性系统的失真问题,Julier 等人又提出了无迹卡尔曼滤波(Unscented Kalman Filter)^[11]。该算法使用确定性抽取的加权采样点(sigma-point)的最小集合,完整地表述了先验随机变量的均值和方差。这一过程被称为无迹变换(Unscented transform)。这些 sigma 点经过非线性系统得到后验的 sigma 点集,仍然很好地描述了

后验概率密度的一阶和二阶统计特征（均值和方差）。

若目标的状态分布是多模的，如存在遮挡和背景干扰的情况，只能采用蒙特卡罗积分方法来求解贝叶斯滤波器。这种普适性的滤波器即为粒子滤波器（Particle Filter）^[12]，也称作自举滤波器（Bootstrap Filter）^[13]。该算法中，当前状态的概率密度由一群带有权重的随机样本表示，这些样本经过动力学模型演化后，其权重由各自的似然函数进行加权修正，从而得到后验概率的离散化表示。新的概率密度由这些样本和权重得到。这些随机样本从一个易于得到和采样的建议分布（proposal distribution）中采样求得。可以简单地使用该粒子的状态转移条件概率作为建议分布，也可以引入观测值对建议分布的影响，使滤波的结果更为准确，如该粒子经由无迹卡尔曼滤波所得的结果作为建议分布^[14]。为了防止状态估计的方差随着递推的进行不断增大，每次迭代后要对粒子集进行重采样，使得新的采样集中各粒子的权重相等。

3.2 核跟踪

典型的核跟踪通过计算初始区域所表示的目标在相邻帧之间的运动来实现^[15]。通常使用序列帧之间目标的运动参数或者稠密的流场描述目标的运动。各算法的不同主要在于所使用表观模型，跟踪目标数量，以及估计物体运动的方法的不同。

Mean-shift 方法^[15,16]通过比较目标的直方图 Q 与假设目标位置窗口的直方图 P ，迭代地使表观的相似度达到最大值。直方图的相似度是以 Bhattacharya 系数 $\sum_{u=1}^b p(u)Q(u)$ 来衡量的，其中 b 是直方图的格数。每一步迭代过程通过使直方图的相似度增加，计算得到的窗口移动矢量。重复该过程直到收敛，通常为 5 到 6 步。生成直方图时通过空间核函数（kernel）进行加权，接近目标中心的像素具有较高的权重。该方法的显著优点是避免了穷举搜索，减少了目标块平移的计算量。但是需要目标至少有一部分在初始的圆形跟踪区域中。

光流法（optical flow）通过计算每个像素在亮度守恒约束条件 $I(x, y, t) - I(x + dx, y + dy, t + dt) = 0$ 下的流矢量，得到稠密的流场。这种计算通常在该像素周围的区域内使用代数或几何的方式。1994 年，Shi 和 Tomasi 提出了 KLT 跟踪器^[17]。这种跟踪器采用迭

代的方法计算以兴趣点为中心的区域（25*25）的平移 (du, dv) ：

$$\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} = \begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}$$

一旦找到该兴趣点新的位置，KLT 跟踪器计算连续帧相应区域的仿射变换为：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & \\ & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

从而估算跟踪块的质量。如果当前区域和预期的区域之间方差和较小，继续跟踪该特征，否则删除该特征。

最近，用机器学习的方法从分类的角度来使研究跟踪问题，成为了一个新的热点。分类的方法通过在特征集上寻找一个最优的分界面去区分目标与背景从而达到跟踪的目的。它避免了传统方法中构造一个复杂的模型去描述目标。Collins 等人^[18]将跟踪看成是一个二类的分类问题，使用在线的特征选择策略来从不同的颜色空间中选择最有区分性的将目标与背景区分开。

Grabner 和 Bischof 提出了基于在线 Boosting 学习的跟踪方法^[19]。Boosting 算法将一系列仅比随机猜测精度略高的弱分类器组合成一个高精度的强分类器，可以选择出图像区域最有代表性的特征进行分类。由各个弱分类器加权的分类结果投票决定最终的类别归属。在线 Boosting 学习算法采用离线 Booting 算法的训练结果进行初始化。在先前目标位置周围的图像区域估计目标置信度（该置信度一般是由样本与分类面的距离决定）。然后在置信度图上搜索极值点作为估计的目标位置。最后再在这次的跟踪结果上提取目标样本与背景样本用于当前特征模型的更新。由于存在着噪声和遮挡，跟踪过程中的实时更新可能会造成样本标识的错误。Liu 等人^[20]将协同训练（co-training）^[21]引入 adaboost 算法的在线学习过程，使用同一目标相互独立的两组特征训练得到两组分类器，分别为某一图像块中得到的另一组特性进行标识。

A.Saffari 等人提出了基于在线随机森林^[22]（Online Random Forests）跟踪方法。结合了在线打包（online bagging）、极端随机森林（extremely randomized forests）的思想^[23]，提出了一个在线决策树的生长方

法。增加了一个时间权重构建自适应地丢弃某些决策树。该方法得到了比在线 Boosting 更好的跟踪结果。同一个研究组的 C. Leistner 等人提出了多视角在线随机森林的学习方法^[24]，将协同训练(co-training)的思想扩展到任意数目的视角，解决在线更新的样本标识不准确的问题。通过打包(bagging)的方法在随机森林中，为每棵决策树选择一个包含相同数目决策树的父森林(parent forest)。在决策树在线训练之前，使用对应的父森林为更新的样本进行标识。

3.3 剪影跟踪

目标可能会有无法用简单的几何形状很好描述的复杂形状，如手、头、肩等。基于剪影的方法为这些物体提供了准确的描述。基于剪影的目标跟踪器通过先前帧所生成的目标模型找到每一帧中的目标区域。模型可以是颜色直方图，目标的边缘或轮廓。剪影跟踪的方法可以被划分为两大类：形状匹配和轮廓跟踪。形状匹配在当前帧中搜索目标的轮廓^[25]；轮廓跟踪则使用状态空间模型^[26]或是直接最小化某些能量函数^[27]，演化初始轮廓到它在当前帧中新的位置。事实上，这些方法可以认为是时间域内使用先前帧作为先验的目标分割。

4 结论

运动目标跟踪是计算机视觉的一个中心研究问题，有着广阔的工程应用前景。本文主要叙述了目标跟踪整个算法流程，包括在跟踪中用到的目标表示，目标检测，运动模型。总结了各种跟踪算法的设计思想，依据所使用的目标表示方法的不同，大体上将跟踪算法归纳整理成三类：点跟踪，核跟踪，和轮廓跟踪。讨论了当前目标跟踪的研究重点和发展趋势。目标跟踪进一步研究面临着理论和实践两方面的挑战——数据关联和推断的定量理论以及自然场景和认知行为的复杂性，目标跟踪算法的设计还存在许多未知的理论依据。随着研究的不断深入和完善以及人类视觉机理的不断揭示，目标技术的研究将给计算机视觉、智能监控、人机交互、机器人工程、视频检索等研究领域注入新的活力。

参考文献

- 1 Bar-Shalom Y, Daum F, Huang J. The probabilistic data association filter. *IEEE Control Systems Magazine*, 2009,29(6):82-100.
- 2 Hu WM, Tan TN, Wang L, Maybank S. A Survey on Visual Surveillance of Object Motion and Behaviors, *IEEE Trans. on Systems, Man and Cybernetics, Part C: Applications and Reviews*, 2004,34(3):334-352.
- 3 Haag M, Nagel HH. Combination of edge element and optical flow estimates for 3D-model-based vehicle tracking in traffic image sequences. *International Journal of Computer Vision*, 1999,35(3):295-319.
- 4 Paragios N, Deriche R. Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 2002,46(3):223-247.
- 5 Mughadam B, Pentland A. Probabilistic visual learning for object representation. *IEEE Trans. Patt. Analy. Mach. Intell.*, 1997,19(7):696-710.
- 6 Cootes T, Edwards G, Taylor C. Robust real-time periodic motion detection, analysis, and applications. *IEEE Trans. Patt. Analy. Mach. Intell.*, 2001,23(6):681-685.
- 7 Shafique K, Shah M. A non-iterative greedy algorithm for multi-frame point correspondence. *IEEE International Conference on Computer Vision (ICCV)*, 2003: 110-115.
- 8 Isard M, Blake A. CONDENSATION-Conditional Density Propagation for Visual Tracking. *International Journal of Computer Vision*, 1998,29(1).
- 9 Julier S, Uhlmann J. A New Extension of the Kalman Filter to Nonlinear Systems. *Proc. SPIE*, 1997,3068:182-193.
- 10 周东华,席裕庚,张钟俊.一种带多重次优渐消因子的扩展卡尔曼滤波器. *自动化学报*, 1991,17(6):689-695.
- 11 Merwe R, et al. The Unscented Particle Filter, Technical Report CUED/F-INFENG/TR 380, Eng. Dept. Cambridge Univ. 2000.
- 12 Arulampalam MS, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian. *IEEE Trans. Signal Processing*, 2002,50(2):174-188.
- 13 Gordon N. A hybrid bootstrap filter for target tracking in clutter, *IEEE Trans. Aerosp. Electron. Syst.*, 1997,33(1):353-358.
- 14 Doucet A, et al. On sequential Monte Carlo methods for

- Bayesian filtering, Dept. Eng., Univ. Cambridge, UK, Tech.Rep, 1998.
- 15 Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking, IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003,25(5):564-577.
 - 16 Hager GD, Dewan M, Stewart CV. Multiple kernel tracking with SSD. Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2004.
 - 17 Shi J, Tomasi C. Good features to track. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1994: 593-600.
 - 18 Collins TR, Liu Y, Leordeanu M. Online selection of discriminative tracking features. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005,27(10):1631-1643.
 - 19 Grabner H, Bischof H. On-line boosting and vision. Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2006. 2006.
 - 20 Liu R, Cheng J, Lu HQ. A Robust Boosting Tracker with Minimum Error Bound in a Co-Training Framework. ICCV 2009. 2009.
 - 21 Levin A, Viola P, Freund Y. Unsupervised improvement of visual detectors using co-training. IEEE International Conference on Computer Vision (ICCV). 2003: 626-633.
 - 22 Breiman L. Random Forests. Machine Learning, 2001, 45(1):5-32.
 - 23 Saffari A, Leistner C, Santner J, et al. On-line Random Forests. 3rd IEEE ICCV Workshop on On-line Computer Vision. 2009.
 - 24 Leistner C, Godec M, Saffari A, et al. Online Multi-View Forests for Tracking. Proc. of Symposium of the German Association for Pattern Recognition (DAGM). 2010.
 - 25 Sato K, Aggarwal J. Temporal spatio-velocity transform and its application to tracking and interaction. Comput. Vision Image Understand. 2004,96(2):100-128.
 - 26 Chen Y, Rui Y, Huang T. Jpdaf based hmm for realtime contour tracking. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2001: 543-550.
 - 27 Bertalmio M., Sapiro G, Randall G. Morphing active contours. IEEE Trans. Patt. Analy. Mach. Intell., 2000,22(7): 33-737.

(上接第 239 页)

恒的主题。通过科技信息资源的聚合,为学科交融、形成优势以及充分利用科技信息资源提供了前提条件。第四、我国科技资源依然存在着配置不均衡的现象,通过科技改革和资源流动等方法,需要长时间逐步改进,而通过搜索引擎技术实现的科技资源深度搜索,则面向广大用户提供各取所需的快捷服务,能够促进科技资源的均衡分配。

参考文献

- 1 孙鸿烈.认清科学数据的战略地位.光明日报,2004.3.12.
- 2 Hanson B. Making Data Maximally Available, Science, Feb. 11,2011
- 3 刘闯.我国科学数据共享机制建设研究.国土资源信息化. 2004,1.
- 4 张景勇.科技部:我国科学数据亟需改变“单兵作战”现状.新华网,2003.2.6.
- 5 刘锋,张玲玲,顾基发.知识管理在互联网中的应用.ISKSS, December 2007,4(4).
- 6 王文钧,李巍.垂直搜索引擎的现状与发展探究.情报科学, 2010,(3):477-480.
- 7 互联网实验室:互联网行业垄断调查及对策研究报告.(2010).2011.
- 8 郑凯明,李义杰.垂直搜索引擎及其应用价值.信息技术, 2008,(4):45-47.
- 9 袁昱明,施建华,陶培根.本体在网络学习主客体构建和交互中的应用.中国远程教育,2009,(12).
- 10 叶德营.默会知识与程序性知识的比较研究[硕士学位论文].杭州:浙江大学,2008.