

基于预测的多特征融合 Mean-Shift 跟踪算法

郭 宇 郝晓燕 张兴忠

(太原理工大学 太原 030024)

摘 要 视频监控在生活中的应用已经相当广泛,其中视频目标精确跟踪是计算机视觉中应用较广、难度较大的一部分。在实际视频场景中目标存在复杂的变化,如外形变化、部分遮挡、光照变化等,这对 Mean-Shift 跟踪算法产生了较大的影响。为了解决上述变化导致的跟踪不准确的问题,融合颜色和 Gabor-LBP 纹理特征进行 Mean-Shift 跟踪,并利用二次多项式预测运动目标的位置,以提高跟踪的准确度。

关键词 Mean-Shift, 特征融合, 纹理特征, 目标跟踪, 二次多项式

中图分类号 TP391 **文献标识码** A

Multi-feature Fusion Mean-Shift Tracking Algorithm Based on Prediction

GUO Yu HAO Xiao-yan ZHANG Xing-zhong

(Taiyuan University of Technology, Taiyuan 030024, China)

Abstract The application of video surveillance in life has been quite extensive, especially the main target tracking is widely used in daily life, and it is a difficult part in the computer vision. In the real video scene, there are many complex target appearance changes, such as partial occlusion, light changes, etc. These have a greater impact on the Mean-Shift tracking algorithm. In order to solve the problem of inaccurate tracking caused by the above complex environment, this paper fused the color and Gabor-LBP edge features in Mean-Shift tracking algorithm, and introduced the quadratic polynomial to predict the position of the video target, to improve the tracking accuracy.

Keywords Mean-Shift, Feature fusion, Texture feature, Object tracking, Quadratic polynomial

Mean-Shift 目标跟踪算法运行速度快,能够在短时间内找到匹配度最高的目标,适用于对响应时间要求高的跟踪系统。Mean-Shift 利用迭代的思想,通过计算目标点的偏移均值进行移动,然后重新选择目标点,继续计算偏移均值,重复迭代直到满足特定的条件结束。偏移均值的计算采用颜色直方图之间的相似度进行调节。该算法速度快,计算效率高,但也存在一定的缺陷,在计算偏移值时,如果目标移动速度过快,两帧画面中没有目标的重叠区域,则会直接导致目标跟踪失败。

近年来,在视觉目标跟踪技术方面已有许多研究成果,研究兴趣来源于该技术的实际应用,包括视频监控、交通流量控制、人机交互和视频压缩等。王永忠等^[1]基于 Fisher 构造了可分性度量的权值自适应更新机制,结合颜色及 LBP 特征提出了一种基于多特征自适应融合的核跟踪框架,实现了自适应融合的跟踪算法;黄伟国等^[2]根据角点特征将物体轮廓分解为小片段,并综合长度尺度进行目标识别;刘明华^[3]提出一种多特征自适应融合的 Mean-Shift 目标跟踪算法。在跟踪场景的动态变化过程中,通过选择对目标和背景区分能力强的特征描述目标,建立多特征融合目标模型,并设置特征重要性权值。当目标速度过快或光线过明过暗时,传统的基于颜色特征的 Mean-Shift 算法的跟踪无法满足跟踪要求,性能下

降,甚至丢失目标。上述几种算法通过模型的改进,增加了算法的时间复杂度。本文在原有算法的基础上,结合二次多项式,融合不同权值颜色特征、Gabor-LBP 纹理特征,提出了一种基于预测多特征融合的 Mean-Shift 算法,并通过实验验证了算法的鲁棒性。

1 传统方法描述

基于颜色特征的 Mean-Shift 跟踪算法,通过计算候选区域与目标区域内像素的加权归一化的颜色直方图值相似度,选择使得相似函数最大的候选模型,并不断迭代更新得到目标的最终位置。该算法虽然使用了基础的颜色特征,但是当光照变化时,单独的颜色特征无法达到目标跟踪的要求,因此引入改进后的 LBP 边缘特征来弥补环境条件变化所导致的跟踪失败。

1.1 Mean-Shift 传统算法

Mean-Shift 跟踪算法以数据点的形式可形象地解释为:算法从数据点 x_i^0 开始跟踪目标(数据点 x_i^0 的上标表示迭代次数,下标表示位置点序列)。其原理如图 1 所示,空心圆圈表示当前中心点,黑色实心点表示移动的窗口样本点,虚线较大圆圈代表密度估计窗口的大小,箭头指向为样本点相对于中心点的漂移矢量。在当前帧中反复迭代计算相似度,搜索

本文受国网山西省电力公司科技项目(520530150015,5205301500W)资助。

郭 宇(1994—),女,硕士生,CCF 学生会员,主要研究方向为图形图像识别及处理,E-mail:646764407@qq.com;郝晓燕(1970—),女,博士,副教授,硕士生导师,主要研究方向为计算机语言学、图像处理;张兴忠(1964—),男,硕士,副教授,硕士生导师,主要研究方向为物联网与嵌入式系统应用、视频跟踪。

特征空间中样本点最密集的区域,数据点沿着样本点密度增加的方向移动更新到局部密度极大点 x_i^N ,即为检测到的目标移动后的位置,从而达到跟踪的目的。

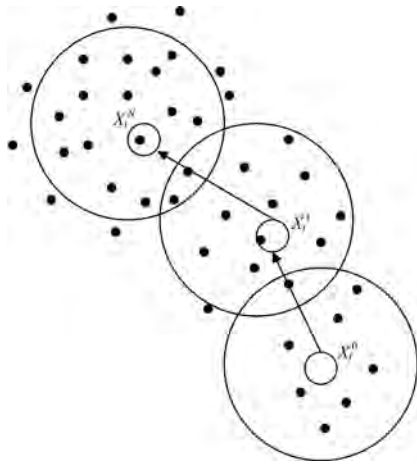


图1 Mean-Shift 算法原理示意图

1) 目标模型的构建

对目标区域的灰度颜色空间进行 m 等分均匀划分,目标模型 $q_u (u=1, 2, \dots, m)$ 的概率密度为:

$$q_u = C \sum_{i=1}^n K(\|z_i^*\|^2) \delta[b(z_i) - u]$$

$$C = 1 / \sum_{i=1}^n \|z_i^*\|^2$$

$$z_i^* = \left\{ \frac{(x_i - x_u)^2 + (y_i - y_u)^2}{x_u^2 + y_u^2} \right\}$$
(1)

其中, z_i 表示归一化后,以目标区域中心为原点的像素点位置; (x_u, y_u) 为目标中心位置; C 为归一化常数; K 为核函数,本文使用 Epanechnikov 核函数; $b(z_i) - u$ 用于判断 z_i 是否在相对应的灰度直方图区间中,若在则置 1,否则置 0。

2) 候选模型的构建

计算当前帧候选区域的直方图,用 $\{z_i (i=1, 2, \dots, n)\}$ 表示像素,则候选模型的概率密度为:

$$p_u(f) = C \sum_{i=1}^n K\left(\left\| \frac{f - z_i}{h} \right\|^2\right) \delta[b(z_i) - u]$$
(2)

其中, h 为核函数窗口大小,其他参数同目标模型。

3) 相似性度量

本文采用 Bhattacharyya 系数作为描述目标模型与候选模型之间相似度的相似函数,数值越大,二者的相似度越高。该系数表示为:

$$\rho(p, q) = \sum_{u=1}^m \sqrt{p_u(f) q_u}$$
(3)

4) 迭代

为使相似函数最大,对式(3)进行泰勒展开:

$$\rho(p, q) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(f_0) q_u} + \frac{C}{2} \sum_{i=1}^n w_i K\left(\left\| \frac{f - z_i}{h} \right\|^2\right)$$

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(f)}} \delta[b(z_i) - u]$$
(4)

式(4)后半部分受 f 值变化影响,因此迭代过程可整理为:

$$f_{k+1} - f_k = \frac{\sum_{i=1}^n w_i (f_k - z_i) g\left(\left\| \frac{f_k - z_i}{h} \right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\| \frac{f_k - z_i}{h} \right\|^2\right)}$$
(5)

当 $f_{k+1} - f_k$ 小于阈值时,所得到的点即为当前帧目标的位置,以该位置作为下一帧搜索中心继续进行相似性计算。

1.2 纹理特征

纹理特征利用图像的灰度值在边缘部分产生的变化,对灰度单调变化不敏感,能够很好地弥补单独使用颜色特征所导致的缺陷。因此引入方向直方图来描述运动目标,使用速度较快的局部二值模式(LBP),但原始 LBP 特征本身不包含方向尺度特征,故将 Gabor 小波变换和 LBP 算子相结合作为检测目标特征的纹理特征模型,并计算其相似度。

二维 Gabor 滤波器能够精确定向地提取图像局部特征,其定义如下:

$$\psi_{\mu, \nu}(x) = \frac{\|k_{\mu, \nu}\|}{\sigma^2} \exp\left(-\frac{\|k_{\mu, \nu}\|^2 \|x\|^2}{2\sigma^2}\right) * [\exp(ik_{\mu, \nu}, x) - \exp(-\frac{\sigma^2}{2})]$$
(6)

其中, x 为像素点位置, μ 和 ν 表示 Gabor 滤波器的方向和尺度, σ 为窗口宽度和波长的比例。式(6)后半部分用交流分量与补偿直流分量相减来抵消对图像亮度的依赖性。 $k_{\mu, \nu} = k_v [\cos \phi_\mu \sin \phi_\mu]^T$ 为小波矢量, $k_v = k_{\max} / f^\nu$, $\phi_\mu = \pi \mu / U$, k_{\max} 为最大频率, f 为尺度因子, U 为滤波器方向的个数。将不同尺度、方向的二维 Gabor 滤波器与目标图像进行卷积运算,取每个卷积结果的幅值部分,得到 Gabor 幅值图。

$$LGBP_{P,R} = \begin{cases} \sum_{p=0}^{P-1} S(G_p(x, u, v) - G_R(x, u, v)) U(LGBP_{P,R}) \leq 2 \\ P+1 \end{cases}$$
(7)

其中, $U(LGBP_{P,R}) \leq 2$ 是使用 LBP 纹理特征对幅值图进行提取的结果,从而可得到纹理模式映射图。对应同一方向尺度建立纹理联合直方图。

$$p_{u1, u2, v}(m, n) = \sum_{1 \leq i \leq w * h} [\delta(m - LGBP_{P,R}(x, u_1, v)) * \delta(n - LGBP_{P,R}(x, u_2, v))]$$
(8)

其中, w 和 h 为图像高度和宽度, $v \in \{0, 1, \dots, V\}$ 为尺度, $u = \{0, 1, \dots, U\}$ 为方向,且 u_1 和 u_2 正交。通过整合多项数据,得出更加合理地描述目标的纹理模式联合直方图。

1.3 二次多项式拟合

Mean-Shift 算法形式简洁,具有较好的效率,但是当目标运动过快时,由于两帧之间的目标区域不重叠,超出滑动窗口步长,不能在局部区域内找到相似性函数得到极值的最优值,直接影响了跟踪目标的准确性。本文引入二次多项式作为拟合函数,来简单表示运动目标的位置,实现简单且速度较快。利用运动目标前十几帧作为前期拟合值,预测物体运动轨迹,当出现遮挡或速度过快时也不影响目标的检测,且可以按照轨迹加速检测。

假设二次多项式形式为 $f(X) = P_0 + P_1 X + P_2 X^2$,拟合 M 个数据点 $(X_i, Y_i) (i=1, 2, \dots, m)$ 。拟合出的结果和实际的结果可能存在偏差,表示为以下方程组:

$$\begin{cases} P_0 + P_1 X_1 + P_2 X_1^2 - Y_1 = R_1 \\ P_0 + P_1 X_2 + P_2 X_2^2 - Y_2 = R_2 \\ \dots \\ P_0 + P_1 X_m + P_2 X_m^2 - Y_m = R_m \end{cases}$$
(9)

其中, R_i 代表实际值和预测值间的差值。使用最小二乘法,取 σ 值最小的 P_i 值,获取最可靠的预测值。

$$\sigma = \sum_{i=1}^m R_i^2$$
(10)

使用上述计算方法,得到下一目标最可靠的位置预测值 Y_{m+1} :

$$Y_{m+1} = \sum_{j=0}^n P_j X_{m+1}^j \quad (11)$$

通过二次多项式的可靠预测,一方面缩减了目标匹配的范围;另一方面减少了算法计算量,提高了目标跟踪算法的实效性。

2 算法步骤

本文算法的具体步骤如下。

步骤 1 首先进行目标的初始化,首帧初始化目标位置,然后以前 2 s 图像为基础,用二次多项式预测目标中心,记为 y_0 。

步骤 2 使用预测目标中心 y_0 和当前帧中的数据,计算出备选模型的 Gabor-LBP 特征模型。

步骤 3 对当前帧数据使用基于颜色特征的 Mean-Shift 算法得到基于颜色的目标中心 y_c ,使用步骤 2 中 Gabor-LBP 特征的 Mean-Shift 算法对当前帧数据进行计算,得到另外一个目标中心 y_g ,融合两个不同的目标中心,得到目标在当前帧中的目标中心 y_1 。

$$y_1 = \lambda y_c + (1 - \lambda) y_g \quad (12)$$

其中,参数 $\lambda = \frac{\rho(y)}{\rho_c(y) + \rho_g(y)}$, ρ 为颜色特征和 Gabor-LBP 特征的 Bhattacharyya 系数。

步骤 4 如果 $\|y_1 - y_0\| < \epsilon$ 成立,则计算图像的下一帧,否则重复步骤 2。

3 实验结果

本文在 MatlabR 2013b, i5 处理器和 Windows 7 环境下进行实验。为了验证算法跟踪的准确性,选用 Car、Caltech 行人数据库以及变电站监控视频片段,分别验证算法在高速移动、遮挡、光照变化时的准确性,并利用二次多项式预测并显示轨迹。图 2(a)为 Car 系列高速移动视频中第 403, 568, 600 帧图像,图 2(b)为 Caltech 行人数据库简单背景下的检测。



(a) Car 视频序列检测结果



(b) Caltech 行人数据库检测结果

图 2 简单场景目标检测

本文实验中采用 Car 视频序列以及 Caltech 行人数据库进行实验测试,算法以第 15 帧作为初始帧开始计算,分别利用 Mean-Shift 算法和本文算法进行检测,并进行对比。由于本文算法引入二次多项式进行预测,在目标速度较快时,也能进行准确的跟踪并显示轨迹,每帧的平均处理速度为 0.108 s,可以满足实时性要求。图 3 中目标较小,环境复杂,且有弯腰等形变较大的动作,同时监控中出现另一目标但原始跟踪并未受到干扰。该算法在两次弯腰时均跟踪准确且跟踪轨迹能够显示目标中心垂直移动的距离。



(a) 第一次弯腰形变



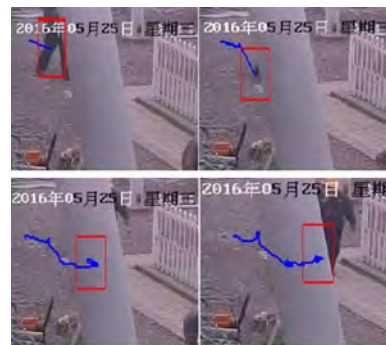
(b) 第二次弯腰形变

图 3 电力系统目标形变检测图

对有遮挡的电力系统实地监控情况进行检测的对比结果如图 4 所示。图 4(a)中,目标由视频左上角出现,手动框选后,红色框为基于颜色的 Mean-Shift 算法的跟踪框,黑色框为颜色轮廓特征融合后的跟踪框。当目标被柱子遮挡时,轨迹及外框开始自由移动,跟踪性能下降,跟踪失败。图 4(b)中使用本文算法,当目标被遮挡时,由于预测作用,轨迹在初始点进行检索并向预测方向前进,当目标再次出现在模型周围时,跟踪继续。



(a) 传统 Mean-Shift 算法的遮挡跟踪



(b) 本文算法的遮挡跟踪

图 4 电力系统监控视频跟踪对比

表 1 不同算法每帧运行的时间比较

视频	颜色特征	颜色轮廓融合跟踪	本文算法
Car	0.44 s	0.28 s	0.108 s
停车场	0.67 s	0.31 s	0.22 s
电力系统	跟踪失败	跟踪失败	0.533 s

本文算法将颜色和 Gabor-LBP 特征进行动态融合,避免了由于环境复杂或者物体的剧烈形变而导致的跟踪失败,同时引入二次多项式对目标轨迹进行预测,避免遮挡时跟踪的漂移现象。

图 5 给出了跟踪算法的精确度,使用帧速为 10 帧/s 的视频进行实验。可以看出,本文算法增加二次多项式后,每帧处理耗时和其他两种算法相比较短;同时随着视频时长的增加,本文算法的目标跟踪准确率逐步提高,且趋于稳定。当目

(下转第 205 页)

- [2] RUDIN L I, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms[J]. *Physica D: Nonlinear Phenomena*, 1992, 60(1): 259-268.
- [3] GOLDSTEIN T, OSHER S. The Split Bregman Method for L1-Regularized Problems[J]. *Siam Journal on Imaging Sciences*, 2009, 2(2): 323-343.
- [4] GILBOA G, OSHER S. Nonlocal Linear Image Regularization and Supervised Segmentation[J]. *Siam Journal on Multiscale Modeling & Simulation*, 2007, 6(2): 595-630.
- [5] GILBOA G, OSHER S. Nonlocal Operators with Applications to Image Processing[J]. *Siam Journal on Multiscale Modeling & Simulation*, 2008, 7(3): 1005-1028.
- [6] VESE L A, CHAN T F. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model[J]. *International Journal of Computer Vision*, 2002, 50(3): 271-293.
- [7] CHAN T, VESE L. An Active Contour Model without Edges [M]. *Scale-Space Theories in Computer Vision*. Springer Berlin Heidelberg, 1999: 141-151.
- [8] 潘振宽, 李华, 魏伟波, 等. 三维图像多相分割的变分水平集方法 [J]. *计算机学报*, 2009, 32(12): 2464-2474.
- [9] DUAN J, PAN Z, LIU W, et al. Color Texture Image Inpainting Using the Non Local CTV Model[J]. *Journal of Signal and Information Processing*, 2013, 4: 43-51.
- [10] HAGEN L, KAHNG A. New spectral methods for ratio cut partitioning and clustering[J]. *IEEE Trans. Comput. -Aided Des. ,* 1992, 11(9): 1074-1085.
- [11] SZLAM A, BRESSON X. Total Variation and Cheeger Cut[C]// *Proceeding of the 27th International Conference on Machine Learning*. 2010: 1039-1046.
- [12] SHI J, MALIK J. Normalized cuts and image segmentation[J]. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, 22(8): 888-905.
- [13] SZLAM A, BRESSON X. A total variation-based graph clustering algorithm for Cheeger ratio cuts[C]// *International Conference on Machine Learning*. 2010: 1039-1046.
- [14] BRESSON X, LAURENT T, UMINSKY D, et al. Convergence and energy landscape for Cheeger cut clustering[J]. *Advances in Neural Information Processing Systems*, 2012: 25: 1394-1402.
- [15] BRESSON X, LAURENT T, UMINSKY D, et al. An adaptive total variation algorithm for computing the balanced cut of a graph[J]. *arXiv preprint arXiv:1302.2717*, 2013.
- [16] BRESSON X, LAURENT T, UMINSKY D, et al. Multiclass total variation clustering[J]. *Advances in Neural Information Processing Systems*, 2013, 26: 1421-1429.
- [17] BRESSON X, TAI X C, CHAN T F, et al. Multi-class transductive learning based on L1 relaxations of Cheeger cut and Mumford-Shah-Potts model[J]. *Journal of Mathematical Imaging and Vision*, 2014, 49(1): 191-201.
- [18] Hu H, LAURENT T, PORTER M A, et al. A method based on total variation for network modularity optimization using the MBO scheme[J]. *SIAM Journal of Applied Mathematics*, 2013, 73(6): 2224-2246.

(上接第 173 页)

标处在较复杂的环境下时,每帧图像的处理速度平均为 0.27s,能够有效保证视频的处理速度,因此所提算法在稳定性和鲁棒性方面都取得了较理想的成果。

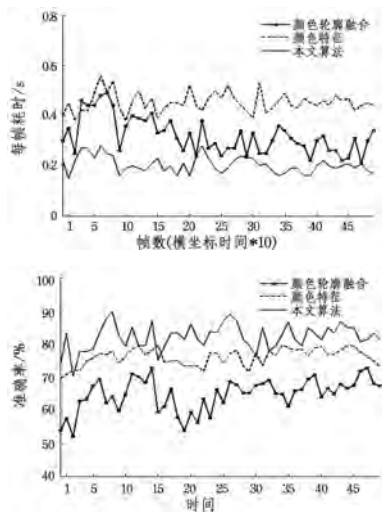


图 5 跟踪算法的精确度

结束语 本文提出了一种基于预测的多特征融合的 Mean-Shift 算法,其将颜色与纹理特征进行动态融合,再用二次多项式对目标位置进行预测。在每一帧图像中,运用基于颜色的 Mean-Shift 算法与基于 Gabor-LBP 的 Mean-Shift 算法对目标模型进行计算,同时结合二次多项式预估的目标位置,通过阈值的规范得到最后的目标位置。本文提出的算法一方面能够有效克服颜色复杂、亮度突变导致的跟踪失败,另

一方面通过预测能够有效解决遮挡问题,与一些经典的 Mean-Shift 算法相比,在速度和准确率上都有一定的提升。

参 考 文 献

- [1] 王永忠,梁彦,赵春晖,等. 基于多特征自适应融合的目标跟踪方法 [J]. *自动化学报*, 2008(4): 393-399.
- [2] 黄伟国,顾超,尚丽,等. 基于轮廓分层描述的目标识别算法研究 [J]. *电子学报*, 2015, 43(5): 854-861.
- [3] 刘明华,汪传生,王宪伦. 基于多特征自适应融合的均值迁移目标跟踪算法[J]. *光电子·激光*, 2015, 26(8): 1583-1592.
- [4] 田莘. 基于 Mean-Shift 算法的目标跟踪问题研究[D]. 西安:西安科技大学, 2010.
- [5] 刘明华. 复杂环境下基于多特征融合的目标跟踪关键技术研究[D]. 青岛:青岛科技大学, 2015.
- [6] 尹宏鹏. 基于计算机视觉的运动目标跟踪算法研究[D]. 重庆:重庆大学, 2009.
- [7] 李鹏飞. 目标跟踪算法的研究与系统构建[D]. 南京:南京师范大学, 2016.
- [8] 袁广林,薛模根,韩裕生,等. 基于自适应多特征融合的 Mean shift 目标跟踪[J]. *计算机研究与发展*, 2010, 47(9): 1663-1671.
- [9] 危自福. 视频目标检测与跟踪关键技术研究[D]. 西安:空军工程大学, 2010.
- [10] CHENG Y. Mean-shift, Mode Seeking, and Clustering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(8): 790-799.
- [11] ZHAO G, PIETIKAINEN M. Local binary pattern descriptors for dynamic texture recognition[J]. *Pattern Recognition*, 2016, 2: 211-214.