

doi: 10.11830/ISSN.1000-5013.201606062



改进的自适应核相关滤波目标跟踪

黄德天^{1,2}, 顾培婷¹, 柳培忠³, 黄炜钦¹

(1. 华侨大学 工学院, 福建 泉州 362021;

2. 华侨大学 机电及自动化学院, 福建 厦门 361021;

3. 厦门大学 信息与通信工程博士后流动站, 福建 厦门 361005)

摘要: 利用核相关滤波器跟踪框架,提出一种改进的自适应颜色属性的目标跟踪方法.首先,构建循环样本矩阵,引进颜色属性作为特征描述目标;然后,采用流行学习局部线性嵌入(LLE)算法自适应地对特征向量进行降维,得到低维特征空间;最后,根据正则化最小二乘分类器获得目标位置.实验结果表明:文中算法的平均中心位置误差减少了 21.29 px;在阈值为 20 px 时,平均距离精度提高了 27.9%,平均跟踪速度为 38 帧·s⁻¹;与传统核相关滤波(KCF)算法相比,文中算法具有良好的光照不敏感性及更高的跟踪精度和鲁棒性.

关键词: 目标跟踪;核相关滤波器;颜色属性;局部线性嵌入

中图分类号: TP 391

文献标志码: A

文章编号: 1000-5013(2017)05-0693-06

Improved Adaptive Target Tracking Based on Kernelized Correlation Filters

HUANG Detian^{1,2}, GU Peiting¹,
LIU Peizhong³, HUANG Weiqin¹

(1. College of Engineering, Huaqiao University, Quanzhou 362021, China;

2. College of Mechanical Engineering and Automation, Huaqiao University, Xiamen 361021, China;

3. Postdoctoral Research Station of Information and Communication Engineering, Xiamen University, Xiamen 361005, China)

Abstract: An improved adaptive color attribute tracking algorithm is proposed based on the kernel correlation filter. Firstly, the cycle matrix is established, and color attribute is used to describe the target. Secondly, the local linear embedding (LLE) algorithm was applied to reduce the dimension of extracted feature to achieve a low-dimensional feature space. Finally, the position is obtained by learning the regularized least-squares classifiers. Experimental results demonstrate that the proposed algorithm reduces the median center location error by 21.29 px, the average distance precision is increased by 27.9% when the threshold is set 20 px, and the average tracking speed is 38 frames·s⁻¹. Compared with the original kernelized correlation filters (KCF) algorithm, the proposed algorithm not only has well illumination insensitivity, but also has higher tracking accuracy and robustness.

Keywords: target tracking; kernelized correlation filters; color attribute; local linear embedding

视频目标跟踪是机器视觉领域的一个重要研究方向,广泛应用于人机交互、军事监控等领域.目前,较为普遍的目标跟踪算法是生成模型方法^[1-2]和判别模型方法^[3-9].在复杂背景下,目标跟踪易受到各种因素影响,如光照变化、目标遮挡、相似目标等,如何更精确地跟踪目标成为一个热门研究课题. Kalal

收稿日期: 2016-06-21

通信作者: 黄德天(1985-),男,讲师,博士,主要从事机器学习、图像处理的研究. E-mail: huangdetian@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(61203242);福建省泉州市科技计划项目(2014Z113)

等^[3]提出 TLD 算法,将跟踪器、检测器和学习相结合. Zhang 等^[4]提出的压缩感知跟踪算法,首次将压缩感知引入视频跟踪中,构造一个投影矩阵,从而获得目标的位置. 近年来,出现了系列相关滤波跟踪算法,其跟踪速度快、跟踪精度高. Henriques 等^[6]提出了循环矩阵(CSK)算法,采用循环移位的方式采集样本,并利用快速傅里叶变换,快速学习分类器,提高了跟踪速度. 在此基础上, Henriques 等^[7]又提出了核相关滤波(kernelized correlation filters, KCF)算法,该算法首先采用循环位移构建样本,训练分类器时通过离散傅里叶变换(DFT)计算,极大地提高了目标跟踪速度,但目标受光照变化影响,使跟踪不精确,甚至跟丢. 因此,本文提出了改进的自适应核相关滤波算法.

1 KCF 跟踪算法

KCF 跟踪器^[7]所用的分类器为正则化最小二乘(RLS)分类器. 假设 \mathbf{x} 是大小为 $m \times n$ 的图像块,并设定一组训练样本及其回归值 $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots\}$, 其训练的最终目标是找到一个函数 $f(z) = \mathbf{v}^T \mathbf{z}$, 使得如下残差函数最小, 即

$$\min_{\mathbf{v}} \sum_i (f(x_i) - y_i)^2 + \lambda \|\mathbf{v}\|^2. \tag{1}$$

式(1)中: λ 是正则化参数.

文献[7]将核函数引入到非线性回归中. 假设 $\varphi(\mathbf{x})$ 是一个到高维核希尔伯特空间的映射, 则核函数 $k(\mathbf{x}, \mathbf{x}') = \langle \varphi(\mathbf{x}), \varphi(\mathbf{x}') \rangle$, 权重可表示为 $\mathbf{v} = \sum_i \alpha_i \varphi(x_i)$, 系数向量为 $\boldsymbol{\alpha}$, 其元素为 α_i , 最终要求解的参数就由 \mathbf{v} 变为 $\boldsymbol{\alpha}$. 因此, 核正则化最小二乘分类器(KRLS)具有一个简单的闭式解, 即

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}. \tag{2}$$

式(2)中: \mathbf{K} 为核函数矩阵, 矩阵元素为 $K_{i,j} = k(x_i, x_j)$; \mathbf{I} 为单位矩阵; 向量 \mathbf{y} 的元素为 y_i . 由于 \mathbf{K} 是循环矩阵, 则式(2)可转换到 DFT 域, 即

$$\hat{\boldsymbol{\alpha}}^* = \hat{\mathbf{y}} / (\hat{\mathbf{k}}^{xx} + \lambda). \tag{3}$$

式(3)中: $\hat{\mathbf{k}}^{xx}$ 是核函数矩阵 \mathbf{K} 的第一行元素组成的向量, 符号 $\hat{\cdot}$ 表示傅里叶变换.

KCF 算法利用核函数法及傅里叶变换极大地提高了算法实时性. 但是, KCF 算法采用 HOG 特征, 无法正确提取复杂多变的目标特征^[7], 跟踪性能易受光照和目标遮挡等因素影响.

2 自适应 KCF 跟踪算法

在 KCF 框架上实现自适应目标跟踪, 利用颜色属性具有对光照变化及部分遮挡不敏感等优点作为目标特征^[8], 使算法在复杂背景下能准确跟踪目标, 从而提高算法的跟踪精确度和鲁棒性. 通过局部线性嵌入(locally linear embedding, LLE)降维法自适应地对高维颜色特征降维, 从而在保持特征原始拓扑结构的前提下, 提高算法运行效率.

2.1 颜色属性

引入颜色属性表示目标特征解决视觉跟踪的光照问题. Berlin 等^[9]对其进行研究, 把颜色分为 11 类: 黑色、蓝色、棕色、灰色、绿色、橙色、粉色、紫色、红色、白色和黄色. 由于文中使用 11 维的颜色特征表示目标, 高维的特征表示会影响跟踪速度, 所以采用低维自适应方法简化特征表示, 从而提高跟踪速度.

2.2 低维颜色自适应属性

LLE 算法^[10]利用线性重建反映高维数据空间中的非线性结构, 能够使降维的数据保持原有的拓扑结构, 可以有效地降低由于降维引起的大量目标信息损失. 因此, 采用 LLE 算法对高维颜色特征降维, 其主要流程如下.

- 1) 寻找相对于每个给定的样本的 β 个近邻样本.
- 2) 定义一个误差函数, 即

$$\min \epsilon \mathbf{W} = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^{\beta} w_{i,j} x_{i,j} \right\|^2. \tag{4}$$

式(4)中: $x_{i,j}$ ($j=1, 2, \dots, \beta$) 为 x_i 的第 j 个近邻点; $w_{i,j}$ 是 x_i 与 $x_{i,j}$ 之间的权值, 并且满足两个约束条件,

即每个数据点 x_i 必须由它的邻域点重构, 否则, $w_{i,j}=0$ 和 $\sum_{j=1}^{\beta} w_{i,j} = 1$.

为了求解矩阵 \mathbf{W} , 应使式(4)最小, 从而构建局部最优化重建权值矩阵, 即

$$w_{i,j} = \sum_{m=1}^k (\mathbf{Q}^i)^{-1}_{j,m} / \sum_{p=1}^k \sum_{q=1}^k (\mathbf{Q}^i)^{-1}_{p,q}. \tag{5}$$

一般情况下, 式(5)中的 \mathbf{Q}^i 是一个奇异值矩阵.

3) 由式(5)的 $w_{i,j}$ 及其近邻点 x_i 计算出样本点的输出向量. 为了将所有样本数据映射到低维空间上, 需要构造一个代价函数, 且映射过程中代价函数达到最小值, 那么有

$$\min \epsilon(\mathbf{U}) = \sum_{i=1}^N \left\| \mathbf{u}_i - \sum_{j=1}^{\beta} w_{i,j} \mathbf{u}_{i,j} \right\|^2. \tag{6}$$

式(6)中: $\epsilon(\mathbf{U})$ 为损失函数值; \mathbf{u}_i 是 x_i 的输出向量; $\mathbf{u}_{i,j}$ ($j=1, 2, \dots, \beta$) 是 \mathbf{u}_i 的 j 个近邻点, 且要满足

$\sum_{i=1}^N \mathbf{u}_i = 0, \sum_{i=1}^N \mathbf{u}_i \mathbf{u}_i^T = \mathbf{I}$. 其中, \mathbf{I} 为单位矩阵. 然后, 求得最优解 \mathbf{u}_i , 使式(6)为最小, 则损失函数可重写为

$$\min \epsilon(\mathbf{U}) = \sum_{i=1}^N \sum_{j=1}^N \mathbf{M}_{i,j} \mathbf{u}_i^T \mathbf{u}_j. \tag{7}$$

式(7)中: \mathbf{M} 是 $N \times N$ 的对称矩阵, $\mathbf{M} = (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W})$. 由式(7)可知, 要使损失函数值达到最小, 则取 \mathbf{U} 为 \mathbf{M} 的最小 d 个非零特征值所对应的特征向量.

LLE 算法中近邻点个数 β 和输出维数 d 决定了算法的性能^[11]. 文中所选的 β 值应尽量使式(4)的误差函数值达到最小, 所以取 β 为 7. 若 d 取值高, 则选取样本多会受到噪声影响; 反之, 样本数据的特征不能得到体现. 通过大量实验, 取 d 为 2. 在跟踪过程中, 通过 LLE 算法降维, 将 11 维的特征向量降为 2 维, 提高了算法的实效性, 并且保持原目标信息不被大量破坏, 从而保证跟踪算法的鲁棒性.

2.3 目标检测与模型更新

利用滑动窗口形式计算图像块 \mathbf{z} 的所有子窗口的响应, 响应值最大的子窗口即为目标的真实位置. 利用核矩阵是循环矩阵这一性质, 分类器的响应为

$$\hat{f}(\mathbf{z}) = (\hat{\mathbf{K}}\mathbf{x}, \mathbf{z})^* \odot \hat{\alpha}. \tag{8}$$

式(8)中: $\hat{\mathbf{K}}\mathbf{x}, \mathbf{z}$ 是这个循环矩阵的第一行组成的向量.

采用文献[7]提供的高斯核函数, 通过式(8)获得目标的响应值, 模型更新 α 和 x_t 为

$$\alpha_t = (1 - \gamma)\alpha_{t-1} + \gamma\hat{\alpha}_t, \tag{9}$$

$$x_t = (1 - \gamma)x_{t-1} + \gamma\mathbf{x}. \tag{10}$$

文中算法的步骤流程如下.

输入: 视频序列 I_t .

初始目标位置 p_0 .

初始化: 由初始目标位置 p_0 及第一帧的目标模型 x_1 , 训练分类器得到系数 α .

for $i=2 \cdots \text{frame}$

1) 通过上一帧位置获取图像块 \mathbf{z} , 并得到特征矩阵;

2) 根据式(4)提取图像的低维颜色属性特征矩阵 projection_matrix;

3) 将矩阵 projection_matrix 带入分类器, 通过式(8)计算目标的响应值;

4) 响应最大值 $\arg \max f(\mathbf{z})$, 为目标真实位置;

5) 根据式(9), (10)更新模型 α_t, x_t ;

输出: 检测出每帧的目标位置 p_t .

end for

3 实验结果与分析

为了验证文中算法的有效性, 选用文献[12]提供的视频序列进行实验. 同时, 与 CT 算法^[4]、CSK 算法^[6]、原始 KCF 算法^[7]的跟踪效果进行对比. 在实验中, 跟踪算法的所有参数保持不变, 取 $\beta=7$, 本征

维数 $d=2$, 高斯核的标准差 $\sigma=0.2$, 学习因子 $\gamma=0.075$. 实验采用 MATLAB 编程环境. PC 配置为 Intel(R) Core(TM) i5-4590M CPU@3.30 GHz, RAM 为 8 G.

3.1 实验结果分析

视频序列的部分跟踪结果, 如图 1 所示. 图 1 中: 点框表示文中算法; 线段框表示 CSK 算法; 点线框表示 KCF 算法; 实线框表示 CT 算法.



图 1 部分跟踪结果

Fig. 1 Partial tracking result

根据图 1 所提算法的结果进行分析, 得到以下 3 点结论.

1) 光照变化和尺度变化. 由图 1(a)可知: 吉他手头部晃动剧烈, 且存在光照变化影响, 在目标进入复杂背景区域时受到强光干扰, CT 和 CSK 算法错误的跟踪非目标人物, KCF 算法也出现很大的偏移, 只有文中算法能较好地完成全程跟踪, 能以较高的精度跟踪整段视频. 因此, 与原始的 KCF 算法相比, 文中采用的颜色属性特征对光照变化不敏感, 当受到强光干扰时, 可实现全程准确的稳定跟踪. 图 1(b)

Singer1 视频序列中, 存在光照影响和尺度变化, 4 种算法都有较好的跟踪结果, 但是文中算法的中心位置误差明显优于其他算法, 体现了更好的跟踪稳定性。

2) 复杂背景干扰和旋转. 图 1(c)中, 当车辆行驶时, 目标跟踪易受到周围复杂背景及类似目标的影响, 在第 284 帧时, CT 算法完全跟踪失败, KCF 正在偏移, 累计错误信息, 导致丢失目标; 而文中提出的自适应颜色属性特征具有光照不变性和抗噪性能, 因此, 能有效避免视频中复杂背景等情况, 完成跟踪. 图 1(d)Mountain bike 视频序列存在目标旋转, 除了 CT 算法, 其他 3 种算法均能较好地完成任务, 但是文中算法的精确度更高。

3) 全局、部分遮挡和快速运动. 图 1(e)Soccer1 视频序列后半部分出现目标遮挡和快速晃动情况. CT 和 CSK 算法的中心误差较大, 当出现全部遮挡时, CT 算法无法继续完成任务; 当部分遮挡时, KCF 算法能较好地跟踪; 但是当全局遮挡时, 文中算法体现了其优越性, 取得了稳健的跟踪结果。

3.2 性能分析

为了测试在不同条件下算法的鲁棒性, 采用中心位置误差 (CLE) 和距离精度 (DP) 为评价准则. 其中, 中心位置误差表达式为

$$\epsilon = \frac{1}{n} \sum_{i=1}^n \|C_i - C_i^*\|.$$

式中: n 是图片序列的帧数; C_i 为目标中心位置; C_i^* 为标准中心位置. 距离精度的表达式为

$$DP = m/n.$$

式中: m 为 CLE 小于某个固定阈值 (实验中均选为 20 px) 的图片序列帧数。

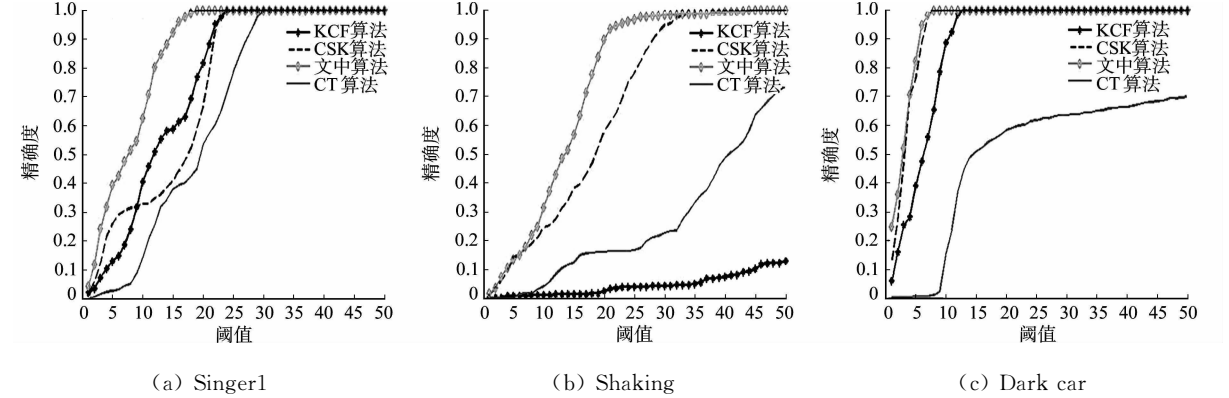
4 种算法的 CLE, DP 的实验数据对比, 如表 1 所示. 由表 1 可知: 文中算法能达到最优或次优性能. 相对于其他 3 种算法, 文中算法具有最优的中心位置误差和距离精度, 其中, 中心位置误差比 KCF 算法减少了 21.29 px, 在阈值为 20 的距离精度比 KCF 算法提高了 27.9%. 在 Shaking 和 Soccer 视频中, 文中算法的精确度也远远高于其他 3 种算法, 证明了文中算法的高精度和不敏感性. 因此, 对于强光、复杂背景等干扰, 文中算法具有较高的精确度和稳定性。

表 1 中心位置误差和距离精度的比较

Tab. 1 Comparison with center location error and distance precision

序列	中心位置误差				距离精度			
	Proposed	CSK	CT	KCF	Proposed	CSK	CT	KCF
Shaking	13.30	17.40	40.10	113.00	88.80	57.90	16.10	2.50
Singer1	7.68	14.00	18.50	12.60	100.00	66.70	53.60	81.50
Dark car	2.93	3.83	38.80	6.05	100.00	100.00	58.30	100.00
Mountain bike	6.85	6.25	211.00	7.66	100.00	100.00	17.30	100.00
Soccer	20.50	71.20	85.50	18.40	79.30	13.30	19.20	44.60
平均值	10.25	22.53	78.78	31.54	93.62	67.58	32.90	65.72

绘制了 4 种算法的跟踪精度曲线, 如图 2 所示. 跟踪精度曲线图用公式表示为 $p = \frac{1}{N} \sum_{i=1}^N \delta[\|O_i - O_i^*\| \leq Th]$. 阈值越低且精度值越高的跟踪器性能更好. 由图 2 可知: 文中算法具有较高的跟踪精度。



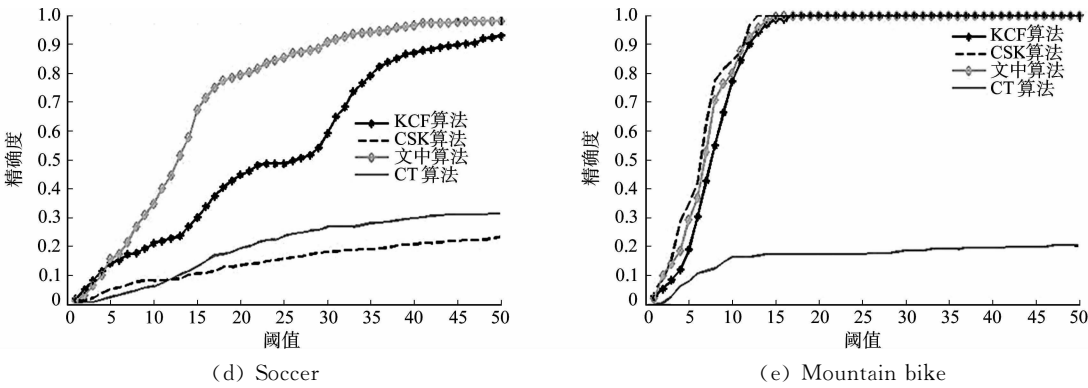


图 2 视频序列跟踪精度曲线图

Fig. 2 Tracking accuracy graph of video sequence

4 结束语

为了解决视频目标跟踪的光照敏感性和提高跟踪精度,基于核相关滤波跟踪算法提出两点改进:首先,针对核相关滤波在光照变化下,会影响跟踪精度,采用颜色属性特征,使算法具有光照不敏感性;其次,由于高维的颜色属性会增加运算复杂度,提出一种自适应降维方法,通过一种非监督学习方法达到低维特征空间,提高目标跟踪的鲁棒性和准确性. 结果表明:文中算法的中心位置误差减少了 21. 29 px,距离精度提高了 27. 9%,在保留传统核相关滤波实时性好、准确度高等特点的同时,能更好地适应光照变化等情况. 此外,在复杂背景下,文中算法能稳定准确地跟踪目标,具有重要的理论研究及应用价值. 在未来工作中,将进一步研究如何提高文中算法的跟踪速度.

参考文献:

[1] 李静宇,刘艳滢,田睿,等. 视频监控系统中的概率模型单目标跟踪框架[J]. 光学精密工程,2015,23(7):2093-2099.

[2] YANG Fan,LU Huchuan,YANG M H. Robust superpixel tracking[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society,2014,23(4):1639-1651.

[3] KALAL Z,MIKOLAJCZYK K,MATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(7):1409-1422.

[4] ZHANG Kaihua,ZHANG Lei,YANG M H. Real-time compressive tracking[C]//Proceedings of the 12th European conference on Computer Vision. Florence:Springer-Verlag,2012:864-877.

[5] CHEN Zhe,HONG Zhibin,TAO Dacheng. An experimental survey on correlation filter-based tracking[J]. Computer Science,2015,53(6025):68-83.

[6] HENRIQUES J F,CASEIRO R,MARTINS P,et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//IEEE Conference on European Conference on Computer Vision. Berlin:Springer-verlag,2012:702-715.

[7] HENRIQUES J F,CASEIRO R,MARTINS P,et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(3):583-596.

[8] 毕笃彦,库涛,查宇飞,等. 基于颜色属性直方图的尺度目标跟踪算法研究[J]. 电子与信息学报,2016,38(5):1099-1106.

[9] BERLIN B,KAY P. Basic color terms: Their universality and evolution[M]. Berkeley:UC Press,1969:11-25.

[10] ROWEIS S T,SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science,2000,2290(5500):2323-2326.

[11] 冯海亮,李见为,黄鸿. 自组织 LLE 算法及其在人脸识别的应用[J]. 光学·精密工程,2008,16(9):1732-1737.

[12] WU Yi,LIM J,YANG M H. Online object tracking: A benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]:IEEE Computer Society,2013:2411-2418.

(责任编辑:黄晓楠 英文审校:吴逢铁)