

DOI: 10.11830/ISSN.1000-5013.201611030



采用互补特征的核相关滤波 目标跟踪算法

谢维波, 夏远祥, 刘文

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 为了改善跟踪算法的性能, 提出一种自适应加权的融合颜色特征和方向梯度直方图(HOG)特征的多核多通道的相关滤波跟踪算法. 针对核相关滤波算法特征单一的问题, 采用互补特征核空间描述目标, 并根据互补特征响应值的大小, 自适应为互补特征核空间分配权重、更新模型, 提高算法的鲁棒性. 实验结果表明: 所提出的算法不仅能在一定程度上处理目标外观变化问题, 而且完全满足跟踪场景的实时需求.

关键词: 目标跟踪算法; 核相关滤波; 互补特征; 自适应权重; 颜色特征; 方向梯度直方图特征

中图分类号: TP 311 **文献标志码:** A **文章编号:** 1000-5013(2018)03-0429-06

Target Tracking Algorithm Using Complementary Features of Kernelized Correlation Filter

XIE Weibo, XIA Yuanxiang, LIU Wen

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: In order to improve the performance of tracking algorithm, a correlation filter tracking algorithm with multi-kernel and multi-channel using an adaptive weighted fusion method based on color feature and histogram of oriented gradient (HOG) feature is proposed. As kernelized correlation filter can extract few features, this algorithm presents target appearance by using complementary kernel features. According to the magnitude of the response values of the complementary features, the weights of the complementary kernel features and updating model are adaptively assigned, improving the robustness of the algorithm. The results of experiments show that the proposed algorithm not only can handle changes of object's appearance, but also completely meet the tracking demand of real-time scenario.

Keywords: target tracking algorithms; kernel correlation filter; complementary features; adaptive weights; color feature; histogram of oriented gradient feature

目标跟踪是计算机视觉领域中最具挑战的问题之一, 在视频监控、机器人等领域中有着广泛应用. 虽然研究者提出了许多有效的跟踪算法^[1-10], 但实际跟踪过程中的目标外观变化, 包括光照变化、非刚性形变、遮挡等问题, 常常导致跟踪失败. 针对跟踪目标的外观变化问题, 一种方法是将跟踪目标分解成不同的部分, 建立分块模型进行跟踪^[5]. 另外一种有效的做法是, 设计一个互补特征模型, 使用不同的特征描述目标外观. 在互补特征模型中, 不同的特征能获得目标不同通道的信息并取得良好的跟踪效果^[11]. 本文在核相关滤波跟踪算法^[1-4]的基础上, 将两种互补的特征模型进行更合理的组合, 同时, 针对跟踪过程中模型更新参数固定的问题, 提出一种自适应参数更新的方法.

收稿日期: 2016-11-11

通信作者: 谢维波(1964-), 男, 教授, 博士, 主要从事信号处理与视频图像分析的研究. E-mail: xwblxf@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(61271383); 华侨大学研究生科研创新能力培育计划资助项目(1400214007)

1 核相关滤波模型

1.1 训练样本的构建

在判别模型中,需要在目标周围进行采样获得训练样本.因为计算量的限制,只能进行少量随机采样,因此,降低了分类器的训练效果.KCF(kernerlized correlation filter)跟踪算法通过循环矩阵进行浓密采样.

设基本向量为 $\mathbf{x}=(x_0,x_1,\cdots,x_n)$,则其循环矩阵 \mathbf{X} 为

$$\mathbf{X} = C(\mathbf{x}) = \begin{pmatrix} x_0 & x_1 & \cdots & x_{n-1} \\ x_{n-1} & x_0 & \cdots & x_{n-2} \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_0 \end{pmatrix}.$$

(1)

对于循环矩阵 \mathbf{X} ,它的每一行都被看作是训练样本.其中,基本向量 \mathbf{x} 为正样本,其他循环移位向量为负样本,如图 1 所示.



图 1 构造循环结构示意图
Fig. 1 Cycle structure diagram

1.2 核相关滤波的构建

针对线性模型分类能力不足的问题,通过引入核函数,将输入数据映射到非线性特征空间提高分类器的性能,则有 $k(\mathbf{x},\mathbf{x}')=\boldsymbol{\varphi}^T(\mathbf{x})\boldsymbol{\varphi}(\mathbf{x}')$. 其中: k 为核函数; $\boldsymbol{\varphi}$ 为映射函数.此时,分类器^[4]可以表示为

$$f(\mathbf{z}) = \mathbf{w}^T \mathbf{z} = \sum_{i=1}^n \alpha_i K(\mathbf{z}, x_i).$$

(2)

式(2)中: $\mathbf{w} = \sum \alpha_i \boldsymbol{\varphi}(x_i)$; $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_n]$,对偶空间系数 $\boldsymbol{\alpha}^{[4]}$ 可表示 $\hat{\boldsymbol{\alpha}}^* = \hat{\mathbf{y}}/(\hat{\mathbf{k}}^{\mathbf{x},\mathbf{x}} + \lambda)^{[4]}$. 其中, $\mathbf{k}^{\mathbf{x},\mathbf{x}}$ 为目标外观模型的自相关; \mathbf{y} 为以目标为中心的高斯函数.则有

$$\mathbf{k}^{\mathbf{x},\mathbf{x}'} = \exp(-\frac{1}{\sigma^2}(\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2) - 2F^{-1}(\hat{\mathbf{x}} \odot \hat{\mathbf{x}}^*)).$$

(3)

式(3)中: \odot 表示像素点乘操作训练标签.

1.3 快速检测

当新一帧的图像 \mathbf{z} 输入时,分类器的置信输出^[4] $\hat{\mathbf{y}} = (\hat{\mathbf{k}}^{\mathbf{x},\mathbf{z}})^* \odot \hat{\boldsymbol{\alpha}}$. 其中: \mathbf{x} 为上一帧学习到的目标外观模型; \mathbf{y} 值最大的位置即为跟踪目标的位置.更新外观模型和滤波模型^[4] 为

$$x_{1:t} = (1 - \lambda)x_{1:t-1} + \lambda x_t, \quad \alpha_{1:t} = (1 - \lambda)\alpha_{1:t-1} + \lambda \alpha_t.$$

(4)

式(4)中: λ 为学习因子.

2 多通道互补特征核相关滤波模型

2.1 特征描述

在计算机视觉中,不同属性的特征在不同干扰因素下稳定性不同. 1) 方向梯度直方图(histogram of oriented gradient,HOG)特征^[12].这是视觉领域中常用的描述图像局部纹理的特征.它受光照变化较小,但对目标形变比较敏感,同时,忽视了目标的颜色信息. 2) 颜色特征.CN(color names)^[13] 是人为描述颜色的语言颜色标签,CN 特征能充分对目标颜色特征进行描述,但易受光照变化的影响.为了增强

算法的鲁棒性, 利用 HOG 特征和 CN 特征在相同类型干扰因素下稳定性不同, 采用多通道的互补特征对目标进行描述.

2.2 多通道特征融合

由式(3)可知, 用 $\mathbf{x}=[x_1, x_2, \dots, x_c] \in \mathbf{R}^{M \times N \times C}$ 表示目标的多通道特征描述. 其中, x_c 表示第 c 个通道特征, 则对应的多通道高斯核为

$$\mathbf{k}(\mathbf{x}, \mathbf{z}) = \exp\left(-\frac{1}{\sigma^2}(\|\mathbf{x}\|^2 + \|\mathbf{z}\|^2) - 2F^{-1}\left(\sum_c \hat{\mathbf{x}}_c \odot \hat{\mathbf{z}}_c^*\right)\right). \quad (5)$$

对于 HOG 特征, c 表示统计的梯度方向个数; 对于 CN 特征, c 表示颜色的维度. 由于只需计算点积及向量范数, 保证了多通道下模型跟踪的速度. 设 $f_{\text{HOG}}, f_{\text{CN}}$ 分别表示 HOG 特征核和 CN 特征核的置信值, 则多通道多核特征模型下置信值可表示为

$$f(\mathbf{x}) = r_{\text{HOG}} f_{\text{HOG}} + r_{\text{CN}} f_{\text{CN}}(\hat{\mathbf{x}}). \quad (6)$$

式(6)中: $r_{\text{HOG}}, r_{\text{CN}}$ 是与相应特征核跟踪效果有关的权重参数.

目标的位置为综合置信值最大的位置, 即为 $\arg \max(f(\mathbf{x}))$.

3 鲁棒的核相关滤波算法

通过定义相似性权重衡量当前跟踪结果, 然后, 根据相似性权重确定不同特征核在确定跟踪结果时所占的比重, 并确定更新模型的学习速率.

3.1 遮挡检测

在相关滤波分类器中, PSR(peak-to-sidelobe ratio)是置信值的峰值旁瓣比, 可以量化相关峰的锐利程度. 在跟踪系统中, 用 PSR 可以衡量当前帧中的候选目标与上一帧目标的相似程度, 即

$$\text{PSR} = (\max(f_t) - \mu) / \sigma. \quad (7)$$

式(7)中: f_t 表示 t 帧时的置信值; μ, σ 分别表示 f_t 的均值和方差.

PSR 值越大, 表示相似程度越高, 跟踪的结果越可靠. 实验发现, 当 PSR 大于 8 时, 跟踪比较稳定, 而小于 5 时, 可认为跟踪目标丢失. 定义衡量相邻两帧目标相似性的权重系数 $w = \text{PSR}$. 当 w 越大时, 表示前、后两帧确定的目标越相似, 发生遮挡的可能性越小. 将相似性权重用于模型更新和确定不同特征核的权值, 可以提高算法的鲁棒性.

3.2 鲁棒的核相关滤波模型

相似性权重可以衡量当前模型的跟踪结果. 当权重较大时, 表明当前的模型比较可靠, 在估计目标位置及更新学习模型时, 应该增大相应的权重, 反之, 则减少. 所以, HOG 特征核和 CN 特征核在确定目标位置时所占的权重分别为

$$r_{\text{HOG}} = w_{\text{HOG}} / (w_{\text{HOG}} + w_{\text{CN}}), \quad r_{\text{CN}} = w_{\text{CN}} / (w_{\text{HOG}} + w_{\text{CN}}). \quad (8)$$

式(8)中: $w_{\text{HOG}}, w_{\text{CN}}$ 分别表示对应特征核的相似性权重.

当更新学习模型时, 采用带权重的自适应模型更新方式. 当相似性权重小于阈值时, 则认为目标被严重遮挡, 此时应停止目标模型的更新.

1) 目标的外观模型更新方式为

$$x_{1:t} = \begin{cases} (1 - \beta w_t) x_t + \beta w_t x_{1:t-1}, & w_t > \text{阈值}, \\ x_{1:t-1}, & \text{其他}. \end{cases} \quad (9)$$

2) 目标的滤波模型更新方式为

$$\alpha_{1:t} = \begin{cases} (1 - \beta w_t) \alpha_t + \beta w_t \alpha_{1:t-1}, & w'_t > \text{阈值}, \\ \alpha_{1:t-1}, & \text{其他}. \end{cases} \quad (10)$$

实验中, 阈值设为 8.

3.3 算法流程

算法完整流程有如下 4 个主要步骤.

1) 在第一帧中, 手动选定跟踪目标, 分别提取 HOG 特征和 CN 特征, 利用式(5)分别获得不同特征核的滤波模型.

- 2) 当新一帧到来时,根据 $\hat{y}=(\hat{k}^{*+}) * \odot \hat{\alpha}$ 确定不同特征核的置信值;然后,根据式(7),(8)计算相似性权重,并由式(6)确定目标位置.
- 3) 根据式(9),(10)对滤波模板和目标外观模型进行更新.
- 4) 输入下一帧,重新获得训练数据;然后,继续下一帧更新.

4 实验结果分析

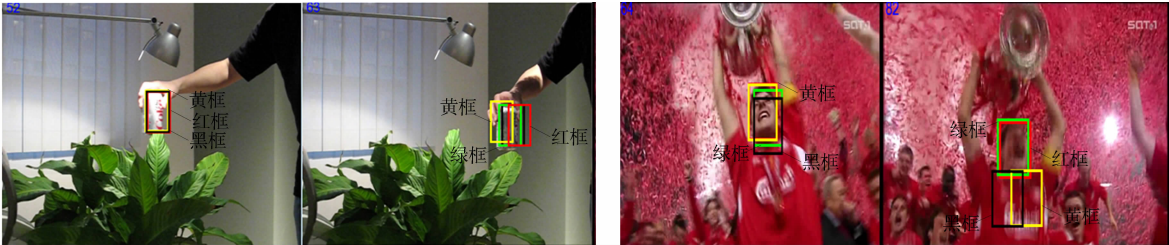
4.1 实验环境与评估标准

为了对所提出算法性能进行更加准确的评估,使用文献[9]中的 10 组目标外观变化较大的标准视频序列对算法进行测试.实验中的对比算法来自相应作者的公开代码,并保留默认的参数设置;正则化参数 λ 设置为 0.01,模型更新因子 β 设为 0.075.系统采用 Windows 7 操作系统,在 Core i5-3470 3.20 GHz CPU 和 8 GB RAM 的台式机配置下进行,算法通过 Matlab 2012 和 Visual Studio 2010 实现.

为了验证算法的有效性,用帧率(FPS)、中心位置误差(CLE)及成功率(SR)等 3 个指标对算法进行评价.其中,帧率用来衡量跟踪速度,平均中心位置误差、成功率用来衡量跟踪结果的准确度.

4.2 KCF 框架下的对比实验

4.2.1 互补特征实验 在 KCF 框架下,不同特征的部分跟踪结果,如图 2 所示.图 2 中:黄色、黑色、红色和绿色矩形实框分别表示灰度特征模型、HOG 特征模型、CN 特征模型和 HOG+CN 特征模型.由图 2 可知:在 Coke 视频中,从第 52 帧至第 63 帧,跟踪目标由为灯下亮度较大的地方运动至亮度较小的地方,易受光照影响的灰度特征模型、CN 特征模型产生了跟踪偏差,HOG 特征模型和互补特征模型受影响较小;在 Soccer 视频中,从第 64 帧至 82 帧,产生了较大的背景干扰,灰度特征模型和 HOG 特征模型产生了较大跟踪偏差,且在之后的跟踪过程中偏差累积,造成跟踪失败.说明充分利用了颜色信息的 CN 特征模型和互补特征模型有着较好的跟踪效果.



(a) Coke 视频 (b) Soccer 视频

图 2 互补特征算法对比实验

Fig. 2 Performance comparison of complementary features algorithms

各种特征模型在 10 组视频序列中 FPS,CLE 和 SR 的平均值,如表 1 所示.由表 1 可知:使用灰度特征的 KCF 算法各方面性能最低,采用 HOG 特征及 CN 特征后的平均跟踪精度均有较大提高.采用多核互补特征的方式(r_{HOG}, r_{CN} 皆设置为 0.5)进一步提高了跟踪精度,但同时损失了速度.

4.2.2 模型更新实验 为了验证模型更新策略的有效性,在多核互补特征模型下进行对比实验,如图 3 所示.图 3 中:红色和绿色矩形实框分别表示固定参数更新和自适应参数更新;所提出的算法采用式(9),(10)中自适应的更新方法.由图 3 可知:在 SUV 视频中,从第 503 帧到第 693 帧,跟踪目标会两次被树木遮挡.采用固定更新方式的模型在遮挡过程中因为学习到了错误模型而逐渐脱离目标,并在之后的过程中始终无法修正直至完全跟踪失败.通过提出的相似性权重更新模型,在两次受到较严重的遮挡后仍能准确跟踪目标.

10 个数据集中的跟踪结果,如表 2 所示.由表 2 可知:相较于固定参数的更新方法,所提出方法的

表 1 互补特征算法对比实验
Tab.1 Performance comparison of complementary features algorithms

方法	SR/%	CLE 值	FPS 值
原型	56.0	35.10	170.4
HOG 特征模型	70.1	24.90	151.0
CN 特征模型	74.2	20.80	72.1
HOG+CN 特征模型	79.1	12.98	66.2



图 3 模型更新对比实验

Fig. 3 Performance comparison of updating model

平均 SR 值提高 1.7%, 平均 CLE 值降低 1.90%, 同时牺牲了一定的速度。

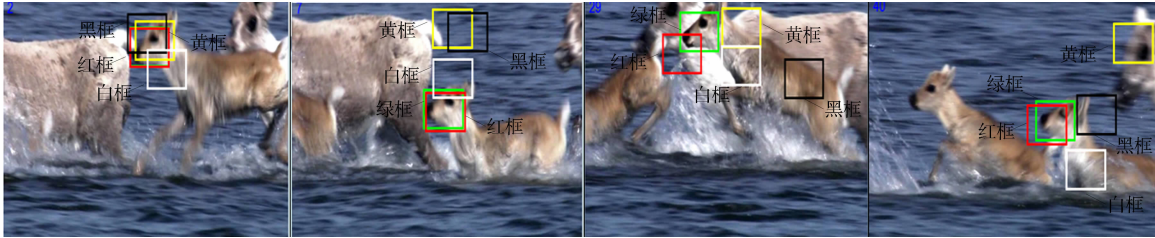
4.3 与其他跟踪算法的对比实验

4.3.1 定性分析 将所提出算法与当前主流的 DFT 算法^[6]、CT 算法^[7]、STC 算法^[3]和基准方法(KCF 算法)在部分视频序列上的部分跟踪结果, 如图 4 所示。图 4 中: 灰色、黑色、黄色、红色和绿色矩形框分别表示 DFT 算法、CT 算法、STC 算法、KCF 和文中算法。

表 2 模型更新对比实验

Tab. 2 Performance comparison of updating model

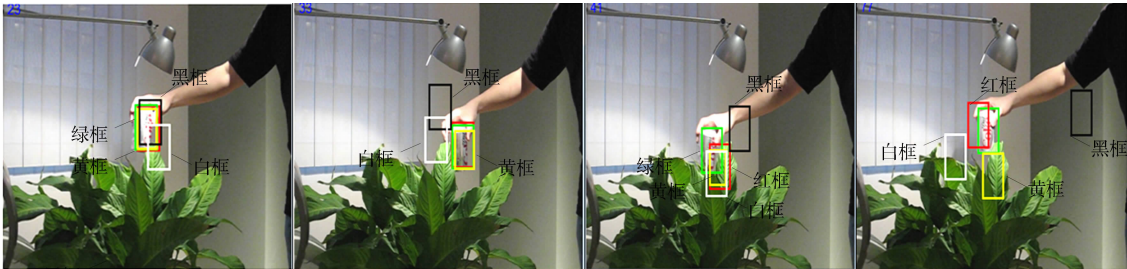
方法	SR/%	CLE 值	FPS 值
HOG+CN 特征模型	79.1	12.98	66.2
文中算法	80.8	11.08	61.8



(a) Deer 视频的部分跟踪结果



(b) Soccer 视频的部分跟踪结果



(c) Coke 视频的部分跟踪结果

图 4 跟踪算法的对比实验

Fig. 4 Performance comparison of tracking algorithms

由图 4(a)可知:Deer 视频中跟踪目标运动较为剧烈,且存在和跟踪目标外观相似的干扰目标。在第 7 帧时,跟踪目标的位置产生了从上到下的较大变化,DFT 算法、STC 算法和 DT 算法脱离了跟踪目标;第 29 帧时,KCF 跟踪算法受到相似目标的干扰,只有文中算法一直有着较好的跟踪结果。由图 4(b)可知:Soccer 视频中存在着较大的光照变化。在第 48,74 和 95 帧中,由于背景的干扰,导致了较大的光线变化,对比算法都产生了一定的跟踪偏差,所提出的算法能一直很好地跟踪目标。由图 4(c)可知:

Coke 视频中对目标存在遮挡情况下进行测试,在第 41,77 帧时,目标被遮挡较多,只有文中算法具有较好的抗遮挡能力.

4.3.2 定量分析 通过 5 种跟踪算法对 10 组标准视频序列进行测试,相应的 FPS,CLE 和 SR 值如表 3 所示.由表 3 可知:文中所提算法在 CLE 和 SR 的平均值取得最优的效果.与 KCF 算法相比,文中算法的 SR 平均值提高 10.7%,CLE 平均值降低 13.01.此外,虽然帧率较 KCF 下降较大,但仍能满足实时要求.

表 3 跟踪算法的对比实验

Tab.3 Performance comparison
of tracking algorithms

算法	SR/%	CLE 值	FPS 值
DFT 算法	37.2	56.33	13.1
CT 算法	21.3	55.17	23.6
STC 算法	52.4	86.30	397.4
KCF 算法	70.1	24.09	151.0
文中算法	80.8	11.08	61.8

5 结束语

在 KCF 算法的基础上,提出一种多核互补特征的核相关滤波跟踪算法.该算法提取图像互补的多通道特征,以多核的形式组合模型,并使用相似性权重更准确地更新和组合模型.通过与原 KCF 算法和其他当前主流跟踪算法对比,验证了文中算法可行性和有效性.当然,文中所提算法仍有不足之处,即当目标外观具有较大的尺度变化、角度旋转等问题时,算法跟踪效果有限.因此,如何处理目标外观变化较大等问题将是下一步的工作.

参考文献:

[1] BOLME D S,BEVERIDGE J R,DRAPER B A,*et al.* Visual object tracking using adaptive correlation filters[C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE Press,2010:2544-2550.

[2] HENRIQUES J F,CASERIO R,MARTINS P,*et al.* Exploiting the circulant structure of tracking-by-detection with kernels[C]//Proceedings of the Europe Conference on Computer Vision. Florence:Springer Press,2012:702-715.

[3] ZHANG Kaihua,ZHANG Lei,LIU Qingshan,*et al.* Fast visual tracking via dense spatio-temporal context learning [C]//Proceedings of the Europe Conference on Computer Vision. Zurich:Springer Press,2014:127-141.

[4] HENRIQUES J F,CASERIO R,MARTINS P,*et al.* High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2014,37(3):583-596.

[5] LIU Ting, WANG Gang, YANG Qingxiong. Real-time part-based visual tracking via adaptive correlation filters [C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. Boston: IEEE Press,2015:4902-4912.

[6] SEVILLA-LARA L,LEARNED-MILLER E. Distribution fields for tracking[C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. Providence:IEEE Press,2012:1910-1917.

[7] ZHANG Kaihua,ZHANG Lei,YANG Mingshuan. Real-time compressive tracking[C]// Proceedings of the Europe Conference on Computer Vision. Florence:Springer Press,2012:864-877.

[8] KWON J,LEE K M. Tracking by sampling trackers[C]// Proceedings of the International Conference on Computer Vision. Barcelona:IEEE Press,2011:1195-1202.

[9] WU Yi,LIM J W,YANG M H. Online object tracking: A benchmark[C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. Portland:IEEE Press,2013:2411-2418.

[10] 钟必能,陈雁,沈映菊,等. 在线机器学习跟踪算法的进展[J]. 华侨大学学报(自然科学版),2014,35(1):41-46. DOI:10.11830/ISSN.1000-5013.2014.01.0041

[11] VARMA M,RAY D. Learning the discriminative power-invariance trade-off[C]// Proceedings of the International Conference on Computer Vision. Rio:IEEE Press,2007:1-8.

[12] FELZENSZWALB P F,GIRSHICK R B. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Software Engineering,2010,32(9):1627-1645.

[13] VAN W J,SCHMID C,VERBEEK J. Learning color names for real-world applications[J]. IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society,2009,18(7):1512-1523.