

融合颜色特征的核相关滤波器目标跟踪

忽晓伟, 陈娟

(郑州大学西亚斯国际学院电子信息工程学院, 郑州 451150)

摘要: 目标特征表示是视觉目标跟踪领域中的一个热点话题。近年来,核相关滤波器因其时效性在目标跟踪中得到广泛应用,但简单的灰度特征表示难以应对复杂环境中的目标跟踪问题。基于此,融合颜色特征提出了一种实时的目标跟踪方法。实验结果表明,颜色信息可以有效提高目标跟踪的整体性能,该方法能够适应姿态变化、光照变化等多种目标外观变化,平均跟踪速度为62.1帧/s,可以满足实时应用的需求。

关键词: 目标跟踪; 特征表示; 相关滤波器; 颜色特征

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-637X(2017)06-0043-04

Kernel Correlation Filter Object Tracking Integrated with Color Features

HU Xiao-wei, CHEN Juan

(College of Electronic Information Engineering, Sias International University, Zhengzhou 451150, China)

Abstract: Object feature representation is a hot topic in visual object tracking. Recently, kernel correlation filter has found wide application in object tracking with its high efficiency, while the simple pixel representation is not adaptable to tracking in complex scenes. A real-time tracking method integrated with color features is proposed here. Experimental results show that color information can improve the overall performance effectively, and the proposed algorithm can perform superiorly under the condition of different appearance variations, such as pose variation, illumination variation etc. The average tracking speed is 62.1 frames per second, which meets the real-time application requirement.

Key words: object tracking; feature representation; correlation filter; color feature

0 引言

视觉目标跟踪是当前监控场景中一个非常具有挑战性的课题^[1],其目的是在一段序列图像中连续估计出目标的位置。在当前的人机交互、视频监控、自动驾驶等诸多领域有着广泛的应用前景。

近年来,大部分跟踪算法都采用基于检测的跟踪框架^[2-4],其主要思想是训练一个在线更新的分类器,由该分类器在每一帧中给出目标位置。在这一类方法中,相关滤波器技术将复杂的卷积操作转换到频域,提高了目标跟踪的时效性,因而受到学者的广泛关注^[5-6]。

文献[7]通过探索目标样本之间的循环移位对相关滤波器进行了拓展,提出了基于核相关滤波器(Ker-

nelized Correlation Filters, KCF)的目标跟踪方法,KCF跟踪器采用单一的灰度信息来表征目标,在目标与背景灰度信息较为相似时会导致跟踪失败。为此,本文通过探索更具有判别力的颜色特征来表征目标,并基于核相关滤波器的跟踪框架提出了一种新的跟踪算法,提高了KCF的跟踪性能。

1 核相关滤波跟踪器(KCF)

KCF的基本思路是:第一帧给定目标的初始位置和大小,现实场景中视频帧率一般为25帧/s,因此目标帧间运动量不会太大,以当前目标位置为中心的局部上下文区域为下一帧的跟踪区域,通过在线更新的目标模型判别出当前帧的目标位置,进而完成跟踪任务。假设跟踪区域的大小为 $M \times N$,KCF利用该区域的所有循环移位 $x_{m,n}$ 作为分类器的训练样本,其中, $(m,n) \in \{0,1,\dots,M-1\} \times \{0,1,\dots,N-1\}$,样本标签是一个高斯岭回归函数,即

收稿日期:2016-10-17

修回日期:2016-12-21

基金项目:2016年度河南省科技计划项目(162102210316)

作者简介:忽晓伟(1980—),男,河南南阳人,硕士,讲师,研究方向为机器视觉、图像处理。

$$y(m, n) = \exp(-((m - m_0)^2 + (n - n_0)^2)/\sigma^2) \quad (1)$$

式中: m_0 和 n_0 为区域中心位置; σ 为尺度参数。则样本训练的目的就是学习到一个滤波器 w 来表征目标模型

$$w = \arg \min_w \sum_{m,n} |\langle \phi(x_{m,n}), w \rangle - y(m, n)|^2 + \lambda \|w\|^2 \quad (2)$$

式中: ϕ 代表可由核空间 κ 推导的映射函数, 例如 x_1 和 x_2 的内积可以表示为 $\langle \phi(x_1), \phi(x_2) \rangle = \kappa(x_1, x_2)$, 常用的核函数有多项式核函数、高斯核函数等; λ 为正则化系数。由于输入被映射到高维特征空间, 式(2)的解可表示为

$$w = \sum_{m,n} \mathbf{a}(m, n) \phi(x_{m,n}) \quad (3)$$

式中, 系数矩阵 $\mathbf{a}(m, n)$ 算式为

$$A = F(\mathbf{a}) = \frac{F(\mathbf{y})}{F(\langle \phi(\mathbf{x}), \phi(\mathbf{x}) \rangle) + \lambda} = \frac{F(\mathbf{y})}{F(\kappa(\mathbf{x}, \mathbf{x})) + \lambda} \quad (4)$$

式中: $F(\cdot)$ 代表傅里叶变换; $\mathbf{y} = \{y(m, n) | (m, n) \in \{0, 1, \dots, M-1\} \times \{0, 1, \dots, N-1\}\}$ 。当下一帧来临时, 在前一帧目标中心处取大小为 $M \times N$ 的图像区域 \mathbf{z} , 目标位置置信图的算式为^[8]

$$\hat{\mathbf{y}} = F^{-1}(A \otimes F(\langle \phi(\mathbf{z}), \phi(\hat{\mathbf{x}}) \rangle)) = F^{-1}(A \otimes F(\kappa(\mathbf{z}, \hat{\mathbf{x}}))) \quad (5)$$

式中: \otimes 代表卷积操作; $F^{-1}(\cdot)$ 代表傅里叶反变换; $\hat{\mathbf{x}}$ 是在线学习的目标表征模型。因此, 在置信图中最大值的位置就是当前帧中的目标位置, 至此即可完成每一帧的目标定位。为了不断学习到目标的外观变化, KCF 采取常见的固定学习率方式更新目标在线模型和分类器参数, 即

$$\hat{\mathbf{x}}_t = (1 - \rho)\hat{\mathbf{x}}_{t-1} + \rho \mathbf{x}_t \quad (6)$$

$$\hat{A}_t = (1 - \rho)\hat{A}_{t-1} + \rho A_t \quad (7)$$

式中: ρ 为固定的学习率; $\hat{\mathbf{x}}_t$ 代表第 t 帧的目标在线模型; \hat{A}_t 为第 t 帧的分类器参数。至此即完成了模型更新。

从以上的分析可知, KCF 通过傅里叶变换操作使得模型更新以及目标定位的计算量大大降低。然而从式(2)可以看出, KCF 所采用的循环移位 $\mathbf{x}_{m,n}$ 只利用了目标的灰度信息。此外, 模型更新中采用固定的学习率, 一旦目标被遮挡或目标发生各种外观变化, 势必会引入模型更新错误, 最终导致跟踪漂移。

2 本文算法

为了提高 KCF 的跟踪性能, 本文提出一种融合颜色信息的核相关滤波器跟踪算法。首先挖掘能够克服场景中光照变化的颜色特征, 同时结合目标灰度特征完成目标表示, 在相关滤波器的框架下完成目标跟踪, 最后结合目标位置的置信度来自适应地更新目标在线模型。算法流程如图 1 所示。

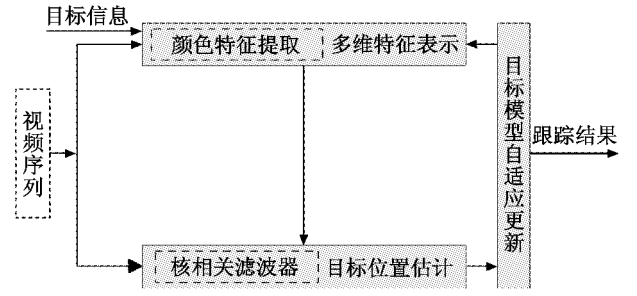


图1 本文算法流程图

Fig. 1 Framework of the proposed algorithm

2.1 抗光照变化的颜色特征表示

传统的颜色空间(如 RGB 颜色空间)各通道之间具有相关性, 极易受到外界光照变化的影响。为此, 通过如下操作去除环境中光照变化的影响。首先将图像从 RGB 空间转换到 $l\alpha\beta$ 空间, 进行归一化处理, 即

$$\begin{cases} l'_s = \frac{\sigma'_l}{\sigma_s} (l_s - m'_l) + m'_l \\ \alpha'_s = \frac{\sigma'_\alpha}{\sigma_s} (\alpha_s - m'_\alpha) + m'_\alpha \\ \beta'_s = \frac{\sigma'_\beta}{\sigma_s} (\beta_s - m'_\beta) + m'_\beta \end{cases} \quad (8)$$

式中: l_s, α_s 和 β_s 分别表示处理前图像在 $l\alpha\beta$ 颜色空间各通道的像素值; l'_s, α'_s 和 β'_s 分别表示处理后图像在 $l\alpha\beta$ 颜色空间各通道的像素值; m'_l 和 σ'_l 是处理前图像在 l 通道的均值和方差; m'_l 和 σ'_l 是参考帧图像在 l 通道的均值和方差(实验中选取每个视频序列的第一帧作为参考帧图像); 其他变量同上定义。

经过上述归一化处理之后, 图像再从 $l\alpha\beta$ 空间转换到 RGB 空间。

2.2 多维特征核相关滤波跟踪

多维特征表征目标较单一特征具有更好的判别能力, 经过 2.1 节的处理后得到抗光照变化的 RGB 颜色特征, 再结合目标灰度特征, 本文的特征表示由四维特征组成, 因此原始的 KCF 跟踪器将不再适用。假设图像由 d 维特征表示, 定义特征维数 $i \in \{1, \dots, d\}$ (本文中 $d=4$), 则式(4)应改写为

$$A = F(\mathbf{a}) = \frac{F(\mathbf{y})}{F\left(\sum_{i=1}^d \kappa(\mathbf{x}_i, \mathbf{x}_i)\right) + \lambda} \quad (9)$$

随之最终的目标位置置信图由式(5)改写为

$$\hat{\mathbf{y}} = F^{-1}\left(A \otimes F\left(\sum_{i=1}^d \langle \phi(\mathbf{z}), \phi(\hat{\mathbf{x}}) \rangle\right)\right) = F^{-1}\left(A \otimes F\left(\sum_{i=1}^d \kappa(\mathbf{z}, \hat{\mathbf{x}})\right)\right) \quad (10)$$

式中, \odot 代表点乘操作。

目标位置置信图中, 置信度最大值处即为当前帧

目标位置,至此在 KCF 框架下完成跟踪任务。

2.3 目标模型自适应更新

从式(7)可以发现,原始 KCF 算法中采用固定的学习率进行目标模型更新,一旦目标被遮挡则势必引入模型更新错误,因此,式(7)无法胜任遮挡情况下的目标模型在线更新,受启发于文献[9],本文结合目标位置置信图中的最大值,提出目标模型自适应更新机制为

$$\hat{A}_t = (1 - \rho w_t) \hat{A}_{t-1} + \rho w_t A_t \quad (11)$$

其中:

$$w_t = \begin{cases} 0 & m_t < 0 \\ m_t / \max\{m_i | i = 1, 2, \dots, t\} & \text{其他} \end{cases} \quad (12)$$

式中, m_t 为第 t 帧目标位置置信图中的最大值。式(11)和式(12)表明,当目标位置置信图中的最大值小于 0 时,目标在线模型将不再更新。

3 实验与结果分析

为验证本文算法的有效性,选用 6 段具有不同特性的测试视频进行对比实验,如表 1 所示。

表 1 实验视频的不同特性

Table 1 Different characteristics of video sequences

视频名称	帧数	主要特性
David	730	光照变化,姿态变化
Crossing	120	光照变化,姿态变化
Girl	353	光照变化,局部遮挡
MountainBike	229	光照变化,姿态变化
Skiing	82	尺度变化,姿态变化
Car	659	光照变化,尺度变化

为验证本文提出的跟踪算法(Proposed)的优越性,本文将与原始的 KCF 算法在 3 段实验视频上进行对比实验,同时还将基于相关滤波技术的 CSK^[6]和 STC^[10]跟踪算法以及当前比较热门的 Struck^[11],SCM^[12]和 TLD^[4]算法加入本文的对比实验中。实验平台采用 Intel Core i5 CPU 主频 3.2 GHz 的 PC 机,利用 Matlab2013a 进行调试。

3.1 评估标准

实验采用 3 个常见的评估标准:

1) 跟踪成功率(SR),定义 $s = R_T \cap R_C / R_T \cup R_C$ 代表每一帧跟踪结果的得分,其中, R_T 代表算法输出的跟踪结果区域, R_C 代表人工标注的目标真实区域,若 $s > 0.5$ 则认为跟踪结果正确,跟踪成功率定义为跟踪结果正确的帧数与视频总帧数的比值;

2) 中心位置误差(CLE),定义为 $\sqrt{(x_T - x_C)^2 + (y_T - y_C)^2}$,其中, (x_T, y_T) 是算法输出的跟踪结果的目标位置, (x_C, y_C) 是人工标注的目标真实位置,取视频中所有帧的中心位置误差的均值作为评估标准;

3) 平均跟踪速度,定义为所有视频帧数除以算法

处理的总耗时(FPS)。

3.2 结果分析

3.2.1 定量分析

表 2~表 4 分别给出了各种算法的跟踪成功率、中心位置误差以及平均跟踪速度,其中,加粗代表最优表现。从表 2 可以发现,本文算法在两段视频中均具有最高跟踪成功率,只有在 David 序列中稍逊于 Struck 算法,平均跟踪成功率为 94.7%,较原始 KCF 算法提高了 3.2%,居所有对比算法之首。从表 3 可以发现,在中心位置误差这个评估标准下,本文算法也优于原始 KCF 算法,平均中心位置误差为 16.2 像素,比原始 KCF 算法降低了 4.1%,同样在所有对比算法中表现最优。从表 4 可以发现,本文算法的平均跟踪速度低于基于相关滤波技术的 KCF,CSK 和 STC 算法,这是因为本文算法将颜色特征融合进核相关滤波器跟踪框架下增加了算法耗时,但 62.1 帧/s 的平均跟踪速度依然可以满足实时应用的需求。

表 2 算法跟踪成功率

Table 2 Success rate of different methods %

	本文算法	KCF	CSK	STC	SCM	Struck	TLD
David	89.3	85.2	82.1	83.7	89.9	90.1	88.3
Crossing	100	99.1	89.3	90.1	93.2	94.8	95.3
Girl	94.8	90.2	82.1	81.7	85.3	88.6	79.2
MountainBike	99.3	99.7	98.1	93.8	99.3	95.3	97.2
Skiing	88.2	90.3	86.2	85.3	81.2	90.2	91.1
Car	97.3	91.5	90.4	86.4	88.3	84.2	87.3
平均 SR	94.8	92.7	88.0	86.8	89.5	90.5	89.7

表 3 算法中心位置误差

Table 3 Center location error of different methods 像素

	本文算法	KCF	CSK	STC	SCM	Struck	TLD
David	17.7	18.2	19.1	19.3	18.9	17.1	18.3
Crossing	11.6	12.3	15.4	15.9	13.5	12.1	18.4
Girl	19.2	20.3	21.4	24.5	22.1	26.7	22.4
MountainBike	7.2	9.1	10.2	16.7	19.2	10.1	9.7
Skiing	20.3	22.2	19.3	28.3	20.8	12.5	19.4
Car	9.7	11.8	12.2	19.0	17.1	9.9	10.3
平均 CLE	14.3	15.7	16.2	20.6	18.6	14.7	16.4

表 4 算法平均跟踪速度

Table 4 Average tracking speed of different methods 帧/s

	本文算法	KCF	CSK	STC	SCM	Struck	TLD
Average Speed	62.1	132.3	234.4	334.0	1.3	20.6	27.2

3.2.2 定性分析

图 2 给出了 3 段视频的部分跟踪结果。



a David

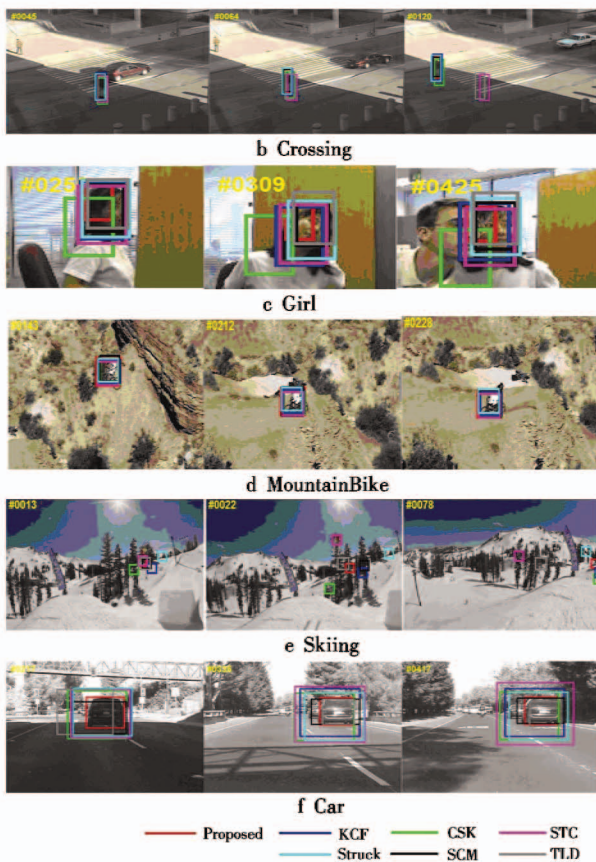


图 2 不同算法部分跟踪结果

Fig. 2 Tracking results of different methods in video sequences

图中不同颜色框代表不同算法的跟踪结果。3 段视频序列中目标都经历了不同程度的光照变化,其中以 David 序列中最为明显,可以看出,本文算法能够较好地锁定目标,充分验证了本文算法提取的颜色特征在抗光照变化上的有效性,在 David 序列第 729 帧时,CSK 和 Struck 算法都出现了严重的跟踪漂移。同时,3 段视频序列中目标也经历了不同程度的姿态变化,其中以 Crossing 序列中行人的姿态变化最为明显,可以发现,在第 120 帧时,STC 和 TLD 算法都无法完成跟踪任务。此外,在 Girl 序列中还出现了目标被局部遮挡的情况,由于本文算法采用自适应的目标模型更新机制,因此在目标被遮挡时能够停止模型更新,从而避免引入更新错误,因而能够较好地完成跟踪任务。

4 结束语

针对视觉目标跟踪中的特征表示这个问题,本文在核相关滤波器的跟踪框架下提出了一种融合颜色特征的目标跟踪算法。为了减少视频场景中经常出现的光照变化的影响,本文提出了一种抗光照变化的颜色特征表示,并结合目标的灰度特征,在核相关滤波器下完成目标跟踪,并构建了一种目标模型自适应更新机

制。实验结果表明,本文算法的跟踪性能较原始 KCF 算法有所提高,并对场景中出现的光照变化、目标姿态变化以及目标遮挡等情况均具有较强的鲁棒性,平均跟踪速度为 62.1 帧/s,完全满足实时应用的需求。

参考文献

- [1] WU Y, LIM J, YANG M H. Online object tracking: a benchmark [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2013:2411-2418.
- [2] BABENKO B, YANG M H, BELONGIE S. Robust object tracking with online multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011(33):1619-1632.
- [3] GRABNER H, LEISTNER C, BISCHOF H. Semi-supervised on-line boosting for robust tracking [C]//European Conference on Computer Vision, 2008:234-247.
- [4] KALAL Z, MATAS J, MIKOLAJCZYK K. Tracking-learning-detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34:1409-1422.
- [5] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2010:2544-2550.
- [6] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C]//European Conference on Computer Vision, 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] MA C, YANG X K, ZHANG C Y. Long-term correlation tracking [C]//IEEE International Conference on Computer Vision, 2015:5388-5396.
- [9] 陈东成,朱明,高文,等.在线加权多示例学习实时目标跟踪[J].光学精密工程,2014,22(6):1661-1667.
- [10] ZHANG K H, ZHANG L, LIU Q S, et al. Fast visual tracking via dense spatio-temporal context learning [C]//European Conference on Computer Vision, 2014:127-141.
- [11] HARE S, SAFFARI A, TORR P H. Struck: structured output tracking with kernels [C]//IEEE International Conference on Computer Vision, 2011:263-270.
- [12] ZHONG W, LU H, YANG M H. Robust object tracking via sparsity-based collaborative model [C]//IEEE International Conference on Computer Vision, 2012:1838-1845.