

引用格式:陈均瑞,李郴荣,戴兴安,等.基于GM(1,1)尺度估计的核相关性滤波跟踪方法[J].电光与控制,2019,26(12):64-68. CHEN J R, LI C R, DAI X A, et al. A kernel correlation filtering tracking method based on GM(1,1) scale estimation[J]. Electronics Optics & Control, 2019, 26(12): 64-68.

基于GM(1,1)尺度估计的核相关性滤波跟踪方法

陈均瑞, 李郴荣, 戴兴安, 盛守照
(南京航空航天大学, 南京 210016)

摘要: 针对在目标视觉跟踪过程中目标尺度变化导致跟踪精度下降的问题,提出了基于GM(1,1)尺度估计的核相关性滤波跟踪方法。在核相关性滤波跟踪算法的基础上引入GM(1,1)灰色预测模型对目标尺度进行预测,并设计尺度估计流程估计当前目标尺度,有效地解决了尺度变化明显导致跟踪性能下降的问题。最后利用OTB-100视频序列数据集测试算法的性能,测试结果表明在目标尺度明显变化时,该算法仍具有较好的跟踪性能,可实现稳定准确的目标跟踪,有效地提高了核相关性滤波跟踪算法的跟踪性能。

关键词: 核相关性滤波; 尺度估计; 目标跟踪; GM(1,1)灰色预测

中图分类号: TP391 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2019.12.013

A Kernel Correlation Filtering Tracking Method Based on GM(1,1) Scale Estimation

CHEN Jun-ruì, LI Chen-rong, DAI Xing-an, SHENG Shou-zhao
(Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Aiming at the problem that the tracking accuracy decreases due to the change of target scale in the process of target visual tracking, a kernel correlation filtering tracking method based on GM(1,1) scale estimation is proposed. The GM(1,1) grey prediction model is introduced to predict the target scale based on the kernel correlation filtering tracking algorithm, and the scale estimation process is designed to estimate the current target scale, which effectively solves the problem of tracking performance declining caused by obvious scale changes. Finally, the OTB-100 video sequence dataset is used to test the performance of the algorithm. The test results show that the algorithm still has good tracking performance when the target scale changes significantly. It can realize stable and accurate target tracking, and effectively improve the tracking performance of the kernel correlation filter tracking algorithm.

Key words: kernel correlation filtering; scale estimation; target tracking; GM(1,1) grey prediction

0 引言

目标视觉跟踪是机器视觉领域一个重要的研究方向,目标视觉跟踪在视频监控、无人车驾驶、视觉导航与指导等多个领域广泛应用^[1]。近几年,各种视觉跟踪算法蓬勃发展,跟踪性能得到了很大的提升^[2]。HARE等提出了Struck算法^[3];KALAL等提出了TLD跟踪算法,将跟踪算法分为跟踪、学习、检测3个模块,较好地解决了目标长时跟踪过程中目标丢失重检测的问题^[4];

BOLME等提出的MOSSE算法首次将相关性滤波应用到视觉目标跟踪领域^[5];HENRIQUES等提出KCF算法,引入循环矩阵,减少了运算复杂度,具有较优的跟踪性能^[6];DANELLJAN等提出了DSST算法^[7],利用尺度滤波器较好地解决了跟踪尺度估计问题;C-COT算法^[8]利用VGG特征,提出了学习连续卷积算子,并解决了特征分辨率不同时的融合问题,跟踪准确率有较大提升,但是速度较慢。核相关性滤波算法在目标快速运动、尺度快速变化等情况下,尺度自适应能力较弱,容易出现跟踪失败或者无法重新跟踪,这是当前亟待解决的一个问题。本文在核相关性滤波的基础上引入GM(1,1)灰色预测算法对跟踪目标进行尺度预测估计,并在OTB-100数据集上实验,验证了本文算法的有效性。

收稿日期:2018-11-05 修回日期:2019-05-08

基金项目:航空科学基金(2016ZC52018)

作者简介:陈均瑞(1994—),男,广东肇庆人,硕士生,研究方向为计算机视觉与目标跟踪。

1 核相关性滤波跟踪算法

核相关性滤波算法的训练本质上是一个岭回归问题。算法的训练目标是训练得到一个函数 $f(z) = \mathbf{w}^T z$, 使得样本 x_i 在此函数下与回归目标 y_i 之间实现最小化均方差, 可以表示为

$$\min_{\mathbf{w}} \sum_i (f(x_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2 \quad (1)$$

式中, λ 为正则化参数。

通过式(1)回归器预测到的样本标签与真实目标标签差距最小。得到最小化平方误差闭式解为

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (2)$$

式中: \mathbf{X} 为样本矩阵; \mathbf{I} 为单位矩阵; \mathbf{y} 为回归目标向量。

文献[6]中采用循环位移得到样本, 根据循环矩阵的性质, 可避免矩阵求逆。则样本矩阵 \mathbf{X} 为

$$\mathbf{X} = \mathbf{F} \text{diag}(\hat{x}) \mathbf{F}^H \quad (3)$$

核化后, 训练样本可表示为

$$\mathbf{w} = \sum_i \alpha_i \phi(x_i) \quad (4)$$

式中: α 是对偶空间变量; $\phi(\cdot)$ 为核函数。这样滤波器可表示为

$$f(x) = \sum_i \alpha_i \kappa(x, x') \quad (5)$$

式中, $\kappa(x, x') = \langle \phi(x), \phi(x') \rangle$ 。由于核函数具有循环性质, 因此最后得到的核化的岭回归的求解算式为

$$\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (6)$$

式中, \mathbf{K} 为具有循环性质的核矩阵。这样即可得到训练后的滤波器 f 。

训练完成后, 以上一帧检测得到的目标位置为中心, 利用循环矩阵选择图像块 z 使用滤波器 f 进行检测, 最后可以得到滤波响应图为

$$\hat{f}(z) = \hat{\mathbf{K}} \cdot \hat{\alpha} \quad (7)$$

式中, $\hat{\cdot}$ 为离散傅里叶变换符号。则响应图中最大的响应值为

$$f_t = \max(\hat{f}(z)) \quad (8)$$

式中, f_t 为 t 时刻的滤波检测峰值。

2 基于 GM(1,1) 的尺度估计算法

在视觉跟踪的实际应用中目标容易出现纵向云, 进而导致目标尺度的变化, 如图 1 所示, 运动目标尺度逐步变化, 基于原始的 KCF 算法没有自适应目标尺度, 容易导致跟踪框内目标特征减少或者引入噪声, 最终导致跟踪失败。因此, KCF 算法采用固定尺度跟踪目标, 易产生模型漂移, 进而跟踪失败。一般而言, KCF 算法采用金字塔法估计目标尺度, 如果金字塔层数多, 尺度估计精度高, 但是运算速度下降明显; 层数少则尺度估计不准确。



图 1 目标尺度变化示意图

Fig. 1 Schematic diagram of target scale change

2.1 基于 GM(1,1) 的尺度预测

灰色预测模型是通过少量的、不完全的信息, 建立数学模型并做出预测的方法, 因此, GM(1,1) 模型适用于对小样本且单调变化的过程进行预测^[9-10]。考虑到目标尺度在短时间内可认为其变化过程是单调的, 因此可取最新的几个数据作为原始序列, 对目标尺度进行一步预测。

首先建立 GM(1,1) 灰色预测模型。设已知序列 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 对原始序列累加生成序列 $x^{(1)}$ 为

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (9)$$

然后计算 $x^{(1)}$ 的均值生成序列 $z^{(1)}$, 为

$$z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)) \quad (10)$$

式中,

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) \quad k=2, 3, \dots, n \quad (11)$$

进一步地构造列向量 \mathbf{Y}_n 为

$$\mathbf{Y}_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T \quad (12)$$

构造数据矩阵

$$\mathbf{B} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (13)$$

通过累加得到的序列 $x^{(1)}$ 与原始序列 $x^{(0)}$ 相比, 有效地抑制了序列的随机性。其变化趋势可以近似用如下微分方程描述

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (14)$$

式中, a 和 u 可以通过如下最小二乘法拟合得到

$$\begin{bmatrix} a \\ u \end{bmatrix} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}_n \quad (15)$$

通过式(15)计算得到的 a 和 u , 即可得到预测方程

$$x^{(1)}(t+1) = (x^{(0)}(1) - \frac{u}{a})e^{-at} + \frac{u}{a} \quad (16)$$

式中, t 为当前时刻, 则 $x^{(1)}(t+1)$ 为 $t+1$ 时刻的预测值。

将最新的 m 个尺度数据作为原始序列, 则有原始序列 x 为

$$x = [s_t, s_{t-1}, \dots, s_{t-m}] \quad (17)$$

式中, s_t 为 t 时刻的目标检测尺度值。记 GM(1,1) 预测算法为函数 gm , 则有

$$s_{t+1}^p = gm(x) \quad (18)$$

式中: x 为原始序列输入; s_{t+1}^p 为 $t+1$ 时刻的尺度预测输出。

2.2 目标尺度估计流程

在视频流中, 目标在运动过程中目标尺度的变化是连续的且不会发生突变。因此可根据式(8)计算滤波器的响应值, 在下一帧图像中先保持目标尺度不变, 对目标进行滤波检测得到响应值 $f_{s_t}(z)$ 。然后根据 2.1 节, 利用 GM(1,1) 预测算法可以得到 $t+1$ 时刻的目标尺度的预测值 s_{t+1}^p , 将目标尺度调整为预测尺度 s_{t+1}^p , 再进行滤波检测得到响应值 $f_{s_{t+1}^p}(z)$ 。比较 $f_{s_{t+1}^p}(z)$ 与 $f_{s_t}(z)$, 如果 $f_{s_{t+1}^p}(z)$ 比 $f_{s_t}(z)$ 大, 则将尺度选择为 s_{t+1}^p , 否则保持尺度不变。进而得到更新后的尺度 s_{t+1}^w , 然后对更新后的目标尺度 s_{t+1}^w 分别进行尺度缩小与尺度增大。最终对 3 个尺度进行滤波检测, 根据对不同尺度的滤波响应值的大小确定 $t+1$ 时刻的目标尺度, 即

$$s_{t+1} = \begin{cases} \alpha s_{t+1}^w & f_{\max}(z) = f_{\alpha s_{t+1}^w}(z) \\ \beta s_{t+1}^w & f_{\max}(z) = f_{\beta s_{t+1}^w}(z) \end{cases} \quad (19)$$

式中 $f_{\max}(z)$ 为 $f_{\alpha s_{t+1}^w}(z)$, $f_{s_{t+1}^w}(z)$, $f_{\beta s_{t+1}^w}(z)$ 的最大值。目标尺度估计流程如图 2 所示。

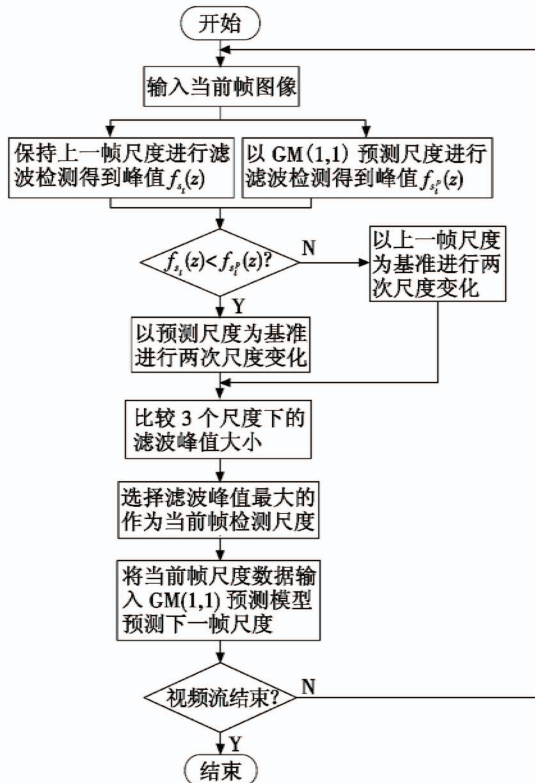


图 2 目标尺度估计流程图

Fig. 2 Flow chart of target scale estimation

2.3 尺度收敛性分析

由 2.2 节可以得知, 本文算法的尺度估计思想与多特征尺度自适应跟踪器 (Scale Adaptive with Multiple Features tracker, SAMF) 相似, 根据估计的尺度与上一时刻尺度进行滤波运算, 选择滤波峰值较大的尺度, 再建立一个 3 级尺度金字塔进行滤波运算。当 GM(1,1) 尺度预测效果不佳时, 算法选择原始尺度建立 3 级金字塔进行目标尺度估计。

设目标真实尺度为 s_{t+1}^t , 那么视频流中两帧图像之间的尺度变化为 $s_{t+1}^t - s_t^t$, 显然两帧之间的尺度变化不会太大, 因此有

$$s_{t+1}^t - s_t^t < \varepsilon \quad (20)$$

估计的尺度越准确, 滤波响应峰值越大。因此根据尺度估计流程可知, 本文算法选择预测尺度与金字塔法生成的尺度中误差最小的。设 t 时刻尺度估计误差为

$$e_t = s_t - s_t^t \quad (21)$$

则 $t+1$ 时刻估计的误差 e_{t+1} 为

$$e_{t+1} < \text{abs}(\alpha - 1)e_t + \varepsilon \quad (22)$$

式中, α 为尺度缩放系数, 一般取 1.05 或 0.95。当 GM(1,1) 尺度预测效果不佳时, 基于尺度金字塔仍然能保证尺度误差不会出现突变, 确保收敛性。同时在 GM(1,1) 尺度预测正常时, 其 $t+1$ 时刻估计的误差 e_{t+1} 必然小于最大估计误差。因此基于这一算法机制确保了尺度估计的收敛性。

3 实验结果分析

本文使用 OTB-100^[11-12] 测试序列对本文算法进行实验验证, 同时与 KCF 算法、DSST 算法在跟踪准确率、中心位置误差 (CLE)、平均重叠率和帧率 4 个指标上进行比较, 结果见表 1、图 3。

表 1 3 种算法定量比较结果

Table 1 Quantitative comparison of 3 algorithms

跟踪算法	跟踪准确率/%	CLE	平均重叠率	帧率/(帧·s ⁻¹)
KCF	59.8	0.41	0.38	78
DSST	61.2	0.48	0.42	60
本文	64.3	0.54	0.44	75

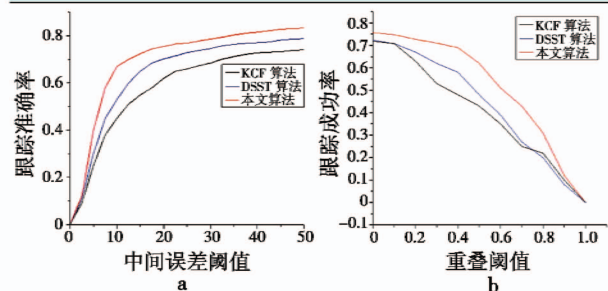


图 3 跟踪准确率和跟踪成功率

Fig. 3 Tracking accuracy and tracking success rate

由比较结果可以看出,本文算法表现最优。由于引入了 GM(1,1) 预测算法对目标尺度进行估计,因此在跟踪精度方面比原始 KCF 算法要高,DSST 算法对尺度进行滤波估计,能较好地估计目标尺度,但是 DSST 算法对尺度估计依赖位置检测的精度,因此 DSST 算法跟踪精度相比本文算法要低些。

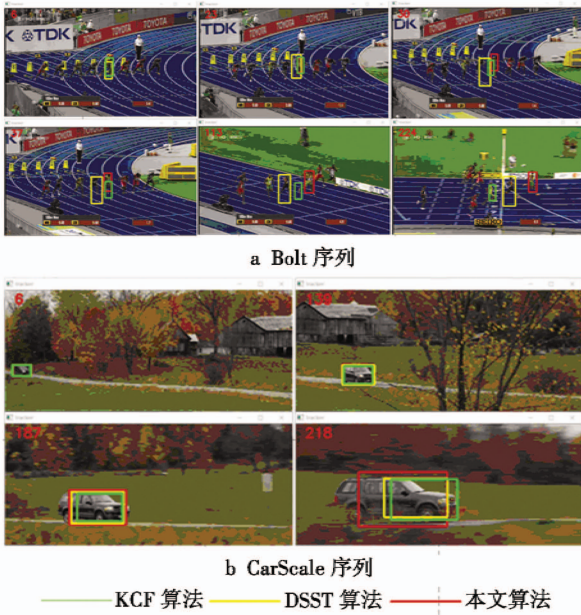


图 4 跟踪效果对比图

Fig. 4 Comparison of tracking effects

本文重点比较了尺度变化较为明显的序列,图 4 为对 KCF 算法、DSST 算法以及本文算法在尺度变化较为明显的序列的测试结果。在 Bolt 序列中,在前 6 帧 3 个算法基本可以对尺度进行较为稳定的估计,但是从第 23 帧开始,KCF 算法的尺度估计明显偏小,DSST 算法则尺度估计明显偏大,尺度不能很好地适应目标,因此 DSST 算法与 KCF 算法均出现了跟踪失败,本文算法对尺度估计较为准确,可以稳定跟踪。在 CarScale 序列前 139 帧的跟踪过程中,汽车的尺度变化并不明显,因此 KCF 算法、DSST 算法以及本文算法对尺度的估计均较为准确,跟踪效果较好,但是在 187 帧以后目标尺度变化明显,DSST 算法与 KCF 算法只能检测到车头部分,尺度估计明显偏小,第 218 帧明显看出 KCF 算法、DSST 算法不能对整个目标进行准确跟踪,本文算法在尺度出现连续较大变化的情况下,效果明显。

为了更好地对尺度估计误差进行分析,计算归一化后的尺度 S_i^* 为

$$S_i^* = S_i / S_1 \quad (23)$$

式中: S_i 为当前帧的尺度; S_1 为第 1 帧的尺度。尺度标准误差 E_s 为

$$E_s = \sqrt{(S_i^* - S_i^{*r})^2} \quad (24)$$

式中, S_i^r 为当前帧的正确尺度值。如图 5 所示,在 150 帧前,3 个算法的尺度误差均较小,但是在 150 帧后,尤其是 200 帧后,DSST 算法与 KCF 算法归一化后的尺度误差明显高于本文算法,可见基于 GM(1,1) 对尺度估计是有效的。图 6 给出了 CarScale 序列中 KCF 算法、DSST 算法与本文算法的跟踪准确率和跟踪成功率,KCF 算法在尺度变化较大的序列中跟踪准确率与成功率均有所改进。

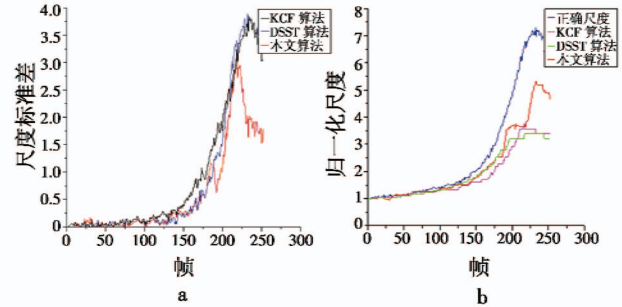


图 5 尺度误差分析图

Fig. 5 Scale error analysis

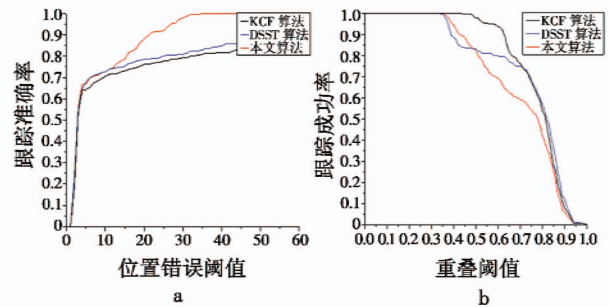


图 6 CarScale 序列中不同算法的跟踪准确率和跟踪成功率

Fig. 6 Tracking accuracy and tracking success rate of different algorithms in CarScale sequence

4 结束语

本文提出了基于 GM(1,1) 尺度估计的核相关性滤波跟踪算法,以 GM(1,1) 预测算法对目标尺度进行预测,实现对目标尺度快速估计,然后通过核相关性滤波算法实现对目标的跟踪。通过实验验证,本文算法在跟踪目标的尺度快速持续变化时,尺度估计准确,在不明显降低跟踪速度的情况下有效提高了跟踪精度。

参考文献

- [1] 黄凯奇,陈晓棠,康运锋,等. 智能视频监控技术综述[J]. 计算机学报,2015, 38(6): 1093-1118.
- [2] 江维创. 基于核相关滤波器的目标跟踪方法研究[D]. 杭州:浙江理工大学,2018.
- [3] HARE S, GOLODETZ S, SAFFARI A, et al. Struck: structured output tracking with kernels[J]. IEEE Transactions

- on Pattern Analysis and Machine Intelligence, 2016, 38 (10):2096-2109.
- [4] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):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. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3):583-596.
- [7] DANELLJAN M, HAGER G, KHAN F S, et al. Accurate scale estimation for robust visual tracking[C]//Proceedings of British Machine Vision Conference, 2014:1-11.
- [8] DANELLJAN M, ROBINSON A, KHAN F S, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking [C]//Proceedings of European Conference on Computer Vision, 2016:472-488.
- [9] 韩中庚. 数学建模方法及其应用[M]. 北京:高等教育出版社,2005.
- [10] 逯程,徐廷学,王虹. 基于改进 GM(1,1)模型的装备故障预测[J]. 火力与指挥控制,2018,43(10):135-138.
- [11] WU Y, LIM J, YANG M H. Online object tracking: a benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2013. doi:10.1109/CVPR.2013.312.
- [12] WU Y, LIM J, YANG M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1834-1848.
-
- (上接第 48 页)
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Neural Information Processing Systems Conference, 2012. doi:10.1145/3065386.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [R/OL]. (2015-04-10) [2019-01-02]. <https://arxiv.org/abs/1409.1556>.
- [13] 塔里克·拉希德. Python 神经网络编程[M]. 林赐,译. 北京:人民邮电出版社,2018.
- [14] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[R/OL]. (2014-09-17) [2019-01-02]. <https://arxiv.org/abs/1409.4842>.