

# 一种基于核相关滤波的视觉跟踪算法

黄 健 郭志波 林科军

(扬州大学信息工程学院 江苏 扬州 225127)

**摘 要** 视觉跟踪是计算机视觉的一个重要方向,而核相关滤波(KCF)跟踪是视觉跟踪领域中的一种比较新颖的方法,它不同于传统基于目标特征的方法,不仅具有较高的跟踪精度,而且具有较快的跟踪速度,在实际应用中效果显著。但当物体快速运动或存在较大尺度变化等时,该方法无法准确地跟踪目标。文中提出的基于核相关滤波器的改进算法有效地解决了上述问题,其通过随机更新多模板匹配,确定了核相关滤波的学习因子,从而实现了学习因子自适应更新模型。实验结果表明,该算法根据不同的场景能快速地调整学习因子,从而提高跟踪的成功度。通过自适应学习因子和多模板匹配,该算法对部分遮挡、光照和目标尺度变化具有较强的适应性。

**关键词** 目标跟踪,核相关滤波,多模板匹配,随机更新

**中图分类号** TP391 **文献标识码** A

## Visual Tracking Algorithm Based on Kernelized Correlation Filter

HUANG Jian GUO Zhi-bo LIN Ke-jun

(School of Information Engineering, Yangzhou University, Yangzhou, Jiangsu 225127, China)

**Abstract** Visual tracking is an important part of the computer vision, and kernelized correlation filter tracking is a relatively novel method in visual tracking field. It is different from traditional method based on target feature, which has high accuracy and fast tracking speed. However, when the object moves rapidly or has the larger scale changes, the method cannot track the target accurately. This paper proposed an improved algorithm based on the correlative filter which can effectively overcome the above problems. The learning factors of kernelized correlation filtering and the adaptive updating model of learning factors are determined by using random update multi-template matching. Experimental results show that the algorithm can adjust the learning factors quickly according to different scenarios, thus the success rate of tracking will be improved. Through adaptive learning factor and multi-template matching, this algorithm has robust adaptability to partial occlusion, illumination and target scale.

**Keywords** Target tracking, Kernelized correlation filtering, Multi-template matching, Random update

目标跟踪是计算机视觉的重要研究课题之一,它能提供相应的行为识别、运动分析等方面的数据,并且广泛应用于智能控制、人机交互、行为分析等领域。近年来,很多目标跟踪算法已经被提出,如卡尔曼滤波<sup>[1]</sup>、粒子滤波<sup>[2]</sup>、均值漂移算法<sup>[3]</sup>。虽然视频跟踪技术已经取得了非常大的进步,但是仍然面临着很大的挑战,如遮挡、目标形变。最近,机器学习的理念被引用到视觉跟踪中,当前方法是将目标跟踪问题转化为一种分类任务并且使用在线学习技术更新目标模型,如TLD<sup>[4]</sup>、CSK<sup>[5]</sup>、MIL<sup>[6]</sup>。这些算法能够获得良好的性能,但不能适应当前的挑战。最近,Henriques等<sup>[7]</sup>对CSK方法进行了改进,使用了方向梯度直方图(HOG)特征,从而提出了核相关滤波器和双相关滤波跟踪器。张雷等<sup>[8]</sup>在核相关滤波的基础上提出了一种尺度估计策略。核相关滤波是一种基于相关滤波器的算法。相关性是两种信号之间相似性的度量,两个信号越相似,相关值就越高。为了提高算法的速度,核相关滤波巧妙地应用了循环矩阵来减少训练滤波器的计算复杂度,并利用HOG特征来提高跟踪的准确度。在非平面旋转

和尺度变化的情况下,其表现优异,但在遮挡或快速运动的情况下表现不佳。一旦在框架中出现跟踪错误,在没有错误纠正机制的情况下,跟踪任务将会失败。

采用多模板匹配方法解决了核相关滤波跟踪模型中学习因子和特征不足的问题。在目标跟踪过程中,核相关滤波存在模板和特征单一不变的缺点,引入随机更新策略能弥补该不足。因此,在此基础上,本文利用多模板匹配和随机更新策略,提出了一种基于核相关滤波的多模板匹配随机更新模型。

## 1 核相关滤波目标跟踪

在核相关滤波目标跟踪过程中,通过最小二乘法训练非线性回归函数,公式如下:

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

$$f(x) = \omega^T \varphi(x) \quad (2)$$

式(1)为最小二乘法,式(2)为非线性回归函数,其中, $x_i$ 是样本, $y_i$ 是与样本对应的期望输出, $\lambda$ 是正则化参数。

$$k(\beta, \lambda) = \langle \varphi(\beta), \varphi(\gamma) \rangle \quad (3)$$

本文受江苏省前瞻性联合研究项目(BY2015061-01)资助。

黄 健(1994-),男,硕士,主要研究方向为目标跟踪;郭志波(1975-),男,博士,副教授,主要研究方向为模式识别、计算机视觉及机器学习, E-mail:zhibo\_guo@163.com;林科军(1992-),男,硕士,主要研究方向为运动目标跟踪。

式(3)的作用是通过核函数将原始空间映射到特性空间。采用非线性回归函数  $f(x)$  通过  $x_i$  的循环移位获得样本  $x_i$ , 而不是通过传统的滑动窗口,从而避免了多次重复的负样本块。通过如下公式计算得到向量系数  $\alpha$ :

$$\alpha = (K + \lambda I)^{-1} y \quad (4)$$

其中,  $I$  是单位矩阵,  $K$  是核矩阵,  $\alpha$  是向量系数,  $y$  为输出响应。所有计算都通过傅里叶域中的计算来提高相关滤波的效率。快速模型训练和检测依赖于核相关,它通过从原始空间到内核空间的映射特征来实现。

$$k(\beta, \gamma) = k(M\beta, M\gamma) \quad (5)$$

如果核函数满足式(5),且  $M$  是任意的排列矩阵,则核矩阵  $K$  是一个循环矩阵。因为循环矩阵的乘积代表了向量的卷积,所以这个计算可以转化为频域中的傅里叶变换。因此在傅里叶域中式(4)可以重新写成式(6):

$$\hat{\alpha} = \hat{y} / (\hat{k} + \lambda) \quad (6)$$

其中,  $k$  是  $K$  的第一行向量,符号  $\wedge$  表示一个向量的离散傅里叶变换(DFT)。快速核函数回归通过核函数自相关性来实现。

$$k_{xx} = K(x, x) \quad (7)$$

$K(\cdot)$  是一个核相关函数。该模型通过式(6)训练得到:

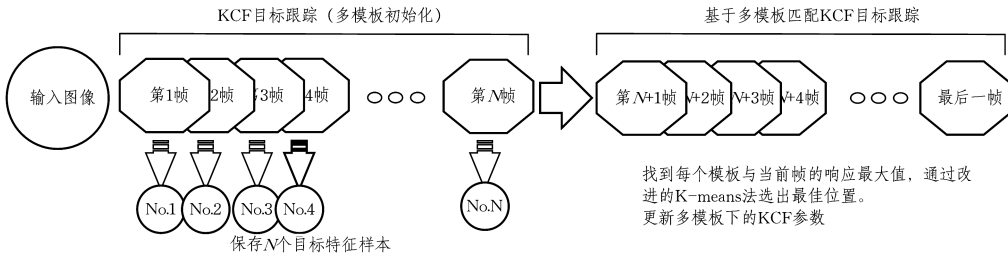


图 1 多模板匹配流程图

多模板匹配下的核相关滤波算法主要包含两个步骤:1)多模板初始化;2)基于多模板匹配的核相关滤波目标跟踪。

在多模板初始化阶段,从第一帧到第  $N$  帧,仍然运用传统核相关滤波方法对目标进行跟踪,但是不同点在于当目标位置每次更新时,通过式(11)计算并且保存目标的每一帧特征  $\hat{x}$ :

$$\arg \max_{\hat{x}} y \quad (11)$$

再通过式(7)和式(8)计算并保存每一帧目标的响应系数  $\hat{\alpha}$ 。

当输入图像运行到第  $N+1$  帧时,将拥有  $N$  个目标特征和  $N$  个响应系数,从而组合成  $N$  个模板。

在基于多模板匹配的核相关滤波目标跟踪阶段,目标的跟踪方式将由多模板匹配的核相关滤波跟踪替代传统的核相关滤波跟踪。从  $N+1$  帧到最后一帧的每一帧中,将多模板初始化阶段的  $N$  个模板替换到传统核相关滤波的单模板。在计算过程中通过传统核相关滤波的块处理的方式来对新一帧进行分割,从而通过式(12)计算输入块与模板的响应值。

$$y = F^{-1}(\hat{\alpha}_{(i)} \odot \hat{k}_{z_x(i)}), i = 1, \dots, N \quad (12)$$

本文算法得到  $N$  个  $\hat{y}$ , 通过式(13)获得该值所在的位置。

$$\arg \max_{bc} \arg \max_i y_{(i)}, i = 1, \dots, N \quad (13)$$

另外,在基于多模板匹配的核相关滤波目标跟踪阶段,采用了一种改进的 k-means<sup>[9]</sup> 和 C-均值聚类<sup>[10]</sup> 算法,从而获得最终的跟踪目标的位置。该算法使用了最大最小距离积分

$$\hat{\alpha} = \hat{y} / (\hat{k}_{xx} + \lambda) \quad (8)$$

快速位置检测是由核函数互相关实现的,公式如下:

$$k_{xx} = K(x, \hat{z}) \quad (9)$$

其中,  $\hat{z}$  是当前视频帧的候选目标块。

$$y = F^{-1}(\hat{\alpha} \odot \hat{k}) \quad (10)$$

通过式(10)获得分类器的输出响应。式中  $\odot$  表示元素点乘,  $F^{-1}$  表示反离散傅里叶变换。该响应反映了每个位置的检测分数,并且最高分数的位置为目标位置。

## 2 改进的核相关滤波的跟踪流程

在传统的核相关滤波的基础上,提出一种随机模板更新策略的核相关滤波跟踪。首先讨论基于多模板匹配的核相关滤波,然后介绍了关于核相关滤波学习因子更新的细节,最后给出了一种完整的目标跟踪算法。

### 2.1 基于多模板匹配的核相关滤波

传统的核相关滤波跟踪策略以首帧方式选取目标作为模板训练的样本,难以解决遮挡和目标外观形变等问题。且由于样本单一,不具备从复杂模板中恢复的能力,因此将多模板匹配的理论加入到传统的核相关滤波中,所提算法的流程图如图 1 所示。

法<sup>[6]</sup>,与之不同的是,由于多匹配算法中的模板数越大,计算量就越大,跟踪精度随之降低,因此在模板数量不会过多和少量噪点数量的原则下,只需在改进的算法中找到 3 个中心点进行样本聚类。前两个点的选取与最大最小距离积分算法相同,首先随机选取第一个中心点,然后找到与第二个点欧氏距离最大的点作为第二个中心点。第三个中心点的选取则是通过样本点到两个中心点的欧氏距离乘积决定,欧氏距离乘积最大的位置点则为第三个中心点。在选出 3 个中心点位置之后进行聚类,聚类方法采用欧氏距离,将  $N$  个位置点变成 3 个簇,簇中数目最大的则为最终目标点所在的簇。通过平均值法,求出最终目标点所在的簇中的平均位置点。然后通过式(14)计算平均位置点与该簇中其他点的欧氏距离,从而得到相应的权值:

$$\omega_i = \frac{\frac{M}{l_i + 1}}{\sum_{i=1}^N \frac{M}{l_i + 1}}, i = 1, \dots, N \quad (14)$$

其中,  $\omega_i$  为最终目标点所在的簇中第  $i$  个点的权值,  $l_i$  为平均位置点与该簇中第  $i$  个点的欧氏距离,  $l$  为平均位置点与该簇中其他点的欧氏距离之和。通过式(15)得到最终的目标跟踪位置:

$$loc = \sum_{i=1}^N \omega_i loc_i, i = 1, \dots, N \quad (15)$$

在求得新的目标位置的同时,需要通过式(16)和式(17)

更新多模板中的特征参数 $\hat{x}$ 和响应系数 $\hat{\alpha}$ 。

$$\hat{\alpha}_{t(i)} = \beta \hat{\alpha}_{t(i)} + (1-\beta) \hat{\alpha}_{t-1(i)} \quad (16)$$

$$\hat{x}_{t(i)} = \beta \hat{x}_{t(i)} + (1-\beta) \hat{x}_{t-1(i)} \quad (17)$$

## 2.2 多模板匹配下的核相关滤波学习因子更新

多模板匹配的核相关滤波跟踪策略前  $N$  帧选取的目标作为模板训练的样本,无法应对目标后期发生的旋转、尺度变换等问题。并且由于样本固定,无法实现样本集的多样性。因此,将样本随机更新的理论加入到多模板匹配的核相关滤波方法中。从  $N+1$  帧到最后一帧,每间隔  $T$  帧计算当前目标的特征参数 $\hat{z}$ 和响应系数 $\hat{\alpha}$ ,然后通过随机更新策略中选取正态分布的方法选出随机数  $h$ ,接着通过式(18)来替换原始样本中随机数  $h$  对应的样本:

$$\hat{z}_{(h)} = \hat{z}, \hat{\alpha}_{(h)} = \hat{\alpha} \quad (18)$$

改进的核相关滤波视觉跟踪算法如算法 1 所示。

### 算法 1 改进的核相关滤波视觉跟踪算法

输入:训练图像块  $X$ ,高斯核函数下的回归目标  $Y$ ,测试图像块  $Z$

输出:每个位置所对应的响应分数 $\hat{y}$ ,目标跟踪位置  $\text{posfor frame} = N+1, \dots, \text{end}$

Patch=subwindow(img)

$Z = \text{FFT}(\text{get\_features}(\text{patch}))$

for  $i=1, \dots, N$

$k_i = \text{FFT}(\text{GaussianCorrelation}(z, x_i))$

$\hat{y}_i = F^{-1}(\hat{\alpha}_{(i)} \odot k_z \hat{x}_{(i)})$

$\text{pos} = \text{pos} + \text{improvekmeans}(\hat{y})$

End for

for  $i=1, 2, \dots, N$

$\hat{\alpha}_{t(i)} = \beta \hat{\alpha}_{t(i)} + (1-\beta) \hat{\alpha}_{t-1(i)}$

$\hat{x}_{t(i)} = \beta \hat{x}_{t(i)} + (1-\beta) \hat{x}_{t-1(i)}$

End for

If frame 为  $T$  的倍数

Patch=subwindow(img)

$Z = \text{FFT}(\text{get\_features}(\text{patch}))$

$k_{zz} = \text{FFT}(\text{Gaussian\_Correlation}(z, z))$

$\alpha = y / (k_{zz} + \lambda)$

$\hat{z}_{(h)} = \hat{z}, \hat{\alpha}_{(h)} = \hat{\alpha}$

End if

End for

## 3 实验结果及分析

为了验证本文算法的有效性,从基准中选取了 6 组复杂场景下的视频序列进行测试,并与其他多种流行跟踪算法进行比较。

实验平台为 MatlabR2015b,所有实验均在配置为 Intel Core i5-7300HQ CPU,主频 2.5 GHz,16 GB 内存的电脑上完成。实验中算法的基本参数保持不变,多模板匹配中的  $N$ 、随机模板更新时间  $T$  和变量  $M$  的选取关系到目标跟踪的精准度。经多次实验发现,当  $N$  取 20,  $T$  取 12,  $M$  取 10 时效果最好,数值越大,计算量就越大,数值就越小,精准度不太理想。

在基准下,为了验证新的更新策略的有效性,将本文提出的方法与 4 种流行跟踪方法做比较,4 种流行跟踪方法有 KCF, TLD, CSK, MIL。

为了评估的准确性,本文遵守基准的相关协议,评估结果如图 2 所示。从图 2 中可以很明显地看出,本文提出的方法是 5 种算法中表现最好的,比传统的 KCF 更优秀。

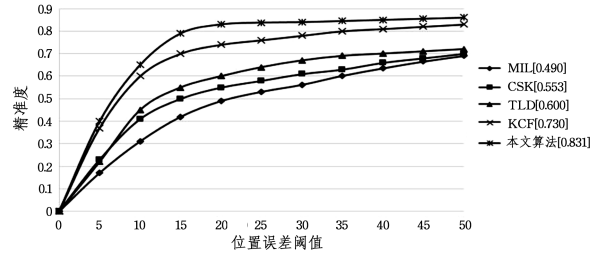


图 2 距离精准度评估结果

通过对图 3 (#84 和 #168 帧)和图 4 (#176 和 #405 帧)的分析可知,目标除了尺度、光照变化的问题外,还存在姿势变化的问题。在 David3 视频序列的整个跟踪过程中, KCF 和本文方法的跟踪效果较好。但由于目标姿势不断变化, KCF 的跟踪窗口引入的干扰信息逐渐增加,最终会出现较大的跟踪偏差。在 David2 视频序列的整个跟踪过程中,目标经历尺度剧烈变化等情况,只有本文方法能够完成整个视频的跟踪,其他方法均不能适应目标的尺度不变。

Jumping 视频和 Tiger2 视频中,有尺度不变、遮挡和快速运动等问题。在 Jumping 视频中(如图 5 中的 Jumping #69 和 #132 帧), TLD 和本文方法的跟踪效果较好。在 Tiger2 视频中(如图 6 中的 Tiger2 #112 和 #257 帧),目标快速运动,并且变化姿态,还有偶尔的局部遮挡,在整个过程中只有本文方法能够适应这些情况,而其他方法会出现快速运动丢失或遮挡时丢失的现象。



图 3 David3 (#84 和 #168)

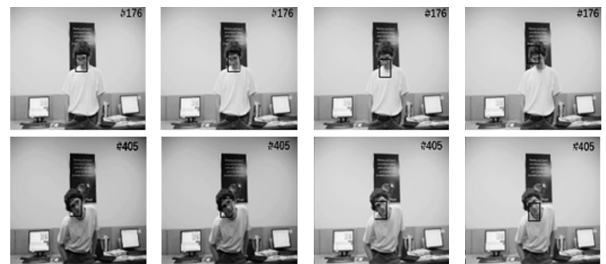


图 4 David2 (#176 和 #405)



图 5 Jumping (#69 和 #132)



(a)本文算法 (b)KCF (c)TLD (d)CSK

图 6 Tiger2(#112 和 #257)

**结束语** 在传统 KCF 的目标跟踪的基础上,提出了一种基于 KCF 的多模板匹配目标跟踪方法。首先通过多模板匹配得到多个样本,然后通过改进的 k-means 方法得到目标位置,最后通过随机模板更新策略更新参数和模板。该方法能在跟踪过程中准确地检测到目标的正确位置,因此可以获得更多的有效目标信息。实验结果表明,在距离精度中,本文方法比传统 KCF 高了 8.41%。本文方法在复杂的场景中获得了更好的鲁棒性和自适应性,提高了跟踪的成功率。因此,改进模板选取、提高特征多样性、采用多模板匹配的 KCF 是我们未来的研究内容。

### 参 考 文 献

- [1] MARCENARO L, FERRARI M, MARCHESOTTI L, et al. Multiple object tracking under heavy occlusions by using Kalman filters based on shape matching[C]// International Conference on Image Processing, 2002:341-344
- [2] ARULAMPALAM M S, MASKELL S, GORDON N, et al. A

tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50(2):174-188.

- [3] COMANICIU D, RAMESH V, MEER P. Real-Time tracking of non-rigid objects using mean shift[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2002: 2142.
- [4] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):1409-1022.
- [5] RUI C, MARTINS P, BATISTA J. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels[C]// Proceedings of European Conference on Computer Vision. ECCV, 2012:702-715.
- [6] 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(8):1619-1632.
- [7] HENRIQUES J F, RUI C, 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] 张雷,王延杰,刘艳涯,等.基于相关滤波器的视觉目标跟踪方法[J].光电子激光,2015,26(7):1349-1357.
- [9] 喻金平,郑杰,梅宏标.基于改进人工蜂群算法的 K 均值聚类算法[J].计算机应用,2014,34(4):1065-1069.
- [10] 刘梦娇,吴成茂.一种改进的局部模糊 C-均值聚类分割算法研究[J].计算机学报,2015,42(6):190-191.

(上接第 229 页)

- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// Proceedings of the Advances in Neural Information Processing Systems, 2012.
- [7] OORD A V D, KALCHBRENNER N, KAVUKCUOGLU K. Pixel recurrent neural networks [J/OL]. <https://cn.arXiv.org/abs/1601.06759>.
- [8] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J/OL]. <https://arxiv.org/abs/1511.06434>.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// Proceedings of the Advances in Neural Information Processing Systems, 2015.
- [10] LIU Z, GAO J, YANG G, et al. Localization and Classification of Paddy Field Pests using a Saliency Map and Deep Convolutional Neural Network [J]. Scientific reports, 2016, 6:20410.
- [11] METZ A R L, RESEARCH I, BOSTON M, et al. Very deep

convolutional networks for large-scale image recognition [J/OL]. [https://www.researchgate.net/publication/294284631\\_Localization\\_and\\_Classification\\_of\\_Paddy\\_Field\\_Pests\\_using\\_a\\_Saliency\\_Map\\_and\\_Deep\\_Convolutional\\_Neural\\_Network](https://www.researchgate.net/publication/294284631_Localization_and_Classification_of_Paddy_Field_Pests_using_a_Saliency_Map_and_Deep_Convolutional_Neural_Network).

- [12] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015.
- [13] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Cognitive modeling, 1988, 5(3):1.
- [14] CARPENTER B. Lazy sparse stochastic gradient descent for regularized multinomial logistic regression [J/OL]. <https://www.mendeley.com/research-papers/lazy-sparse-stochastic-gradient-descent-regularized-multinomial-logistic-regression>.
- [15] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):1.
- [16] KOHAVI R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]// Proceedings of the Ljcai. Stanford, CA, 1995.