

引用格式:李海彪,黄山.基于自适应的核相关滤波的目标跟踪算法[J].电光与控制,2019,26(4):49-53. LI H B, HUANG S. A target tracking algorithm based on adaptive kernelized correlation filtering[J]. Electronics Optics & Control, 2019, 26(4):49-53.

## 基于自适应的核相关滤波的目标跟踪算法

李海彪<sup>a</sup>, 黄 山<sup>b</sup>

(四川大学, a. 电气信息学院; b. 计算机学院, 成都 610065)

**摘 要:** 针对核相关滤波器(KCF)跟踪算法在目标发生尺度变化和严重遮挡的情况下跟踪失败的问题,提出了一种基于自适应的核相关滤波的目标跟踪算法。该算法运用了尺度估计策略,使跟踪框自适应,用多项式核函数来减少计算量,采用了 FHog 目标特征代替原来的 Hog 特征,获取更多的目标特征信息。实验采用 OTB-2013 评估基准的 50 组视频序列进行测试,并与其他 31 种跟踪算法进行对比,测试所提算法的有效性。实验结果表明:所提算法成功率为 0.549,精确度为 0.736,排名第一,与 KCF 算法相比,分别提高了 3.8% 和 1.0%。该算法在目标发生尺度变化、严重遮挡等复杂情况下,均具有较强的稳健性和鲁棒性。

**关键词:** 目标跟踪; 核相关滤波; FHog 特征; 尺度估计; 多项式核

中图分类号: TP391.41 文献标志码: A doi:10.3969/j.issn.1671-637X.2019.04.010

## A Target Tracking Algorithm Based on Adaptive Kernelized Correlation Filtering

LI Hai-biao<sup>a</sup>, HUANG Shan<sup>b</sup>

(Sichuan University, a. College of Electrical Engineering and Information;

b. College of Computer Science, Chengdu 610065, China)

**Abstract:** In order to solve the problem of tracking failure of the Kernelized Correlation Filtering (KCF) tracking algorithm in the case of target scale changes and severe occlusion, an adaptive tracking algorithm is proposed based on kernelized correlation filtering. The algorithm uses a scale estimation strategy to adapt the tracking frame to target scale changes, and uses polynomial kernel functions to reduce the computational complexity. The FHog target feature is used to replace the original Hog feature to obtain more target feature information. In the experiment, 50 sets of video sequences based on the OTB-2013 evaluation benchmark were tested and compared with other 31 tracking algorithms to verify the effectiveness of the proposed algorithm. The experimental results show that: the success rate of this algorithm is 0.549 and the accuracy is 0.736, ranking first, which is improved by 3.8% and 1.0% respectively compared with the KCF algorithm. The algorithm has strong steadiness and robustness under complex conditions such as target scale changes and severe occlusion.

**Key words:** target tracking; kernelized correlation filtering; FHog feature; scale estimation; polynomial kernel

### 0 引言

目前主流的目标跟踪算法主要有两种:一种是基于深度学习的跟踪算法<sup>[1-4]</sup>;另一种是基于相关滤波器的目标跟踪算法<sup>[5-7]</sup>。基于深度学习的跟踪算法需要处理大量的数据,要求运算能力强,对硬件设备要求

高,同时前期要处理大量的数据,移植较为麻烦,目前仍处于理论研究阶段,实际应用还有待发展。基于相关滤波器的跟踪算法,以最小输出平方误差和相关滤波器(MOSSE)跟踪算法、核相关滤波器(KCF)跟踪算法为代表,此类算法运行速度较快,对硬件要求较低,发展比较成熟。基于相关滤波器的目标跟踪虽然取得了巨大的进步,但由于受到目标的尺度变化、形变、严重遮挡、旋转等因素的影响,目标跟踪依然是一个难题。例如:文献[8]提出了一种 MOSSE 的学习跟踪算

收稿日期:2018-05-07

修回日期:2018-06-07

作者简介:李海彪(1989—),男,河南新乡人,硕士生,研究方向为目标跟踪。

法,该算法只需要训练一个目标外观模型,通过离散傅里叶变换将目标与所有候选区域之间的相似度计算转换到频域,提升了跟踪算法的运行速度,但在目标形状、光照发生变化时,跟踪效果较差。文献[9]提出了循环结构的检测跟踪(CSK)算法,该算法利用样本的循环移位来训练分类器,利用离散傅里叶变换在频域实现快速计算,其采用了目标的灰度特征,当在光照变化、目标遮挡时,跟踪很容易失败。文献[10]在CSK算法的基础上提出了KCF跟踪算法,使用方向梯度直方图(Hog)<sup>[11]</sup>特征代替原来的灰度特征,将相关滤波器由单通道扩展到多通道,提高了目标跟踪成功率,但在目标的尺度发生变化、严重遮挡时,其跟踪效果较差。

本文针对KCF算法在目标发生尺度变化、严重遮挡时跟踪失败的问题,提出了一种基于自适应的核相关滤波的目标跟踪算法,该算法采用了尺度估计策略,使跟踪框自适应,引入了多项式核函数,以减少计算量。同时用FHog特征<sup>[12]</sup>来代替原来的Hog特征,获取了目标更多的特征信息,以提高目标跟踪的准确率。

## 1 KCF跟踪算法

KCF跟踪算法是一种基于核相关滤波器的跟踪算法,利用循环矩阵理论在目标区域进行移位,为分类器构造了大量样本,解决了稀疏采样的样本冗余性问题,通过傅里叶变换快速训练分类器和检测目标,最终达到快速跟踪目标的目的。具体步骤如下所述。

### 1) 建立回归函数。

利用样本,通过岭回归(脊回归)问题建立一个回归函数(分类器),即

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

式中: $\mathbf{w}$ 为线性回归的权值; $\lambda$ 为控制过拟合的正则化参数,式(1)为闭式解,即

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

目标图像 $\mathbf{x} \in \mathbf{R}^{M \times N}$ 为基样本, $\mathbf{X}$ 为基样本 $\mathbf{x}$ 构造得到的循环矩阵, $\mathbf{y}$ 为标记样本的实际响应矩阵,两者都为已知量, $\mathbf{I}$ 为单位矩阵。

### 2) 利用移位的样本来训练最小二乘分类器。

在对分类器进行训练时,利用目标的基本样本为正样本,对基本样本进行循环移位得到的样本为负样本,可得

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

由于对角矩阵是对称的,对其进行埃尔米特转换就会剩下复共轭 $\hat{\mathbf{x}}^*$ 。另外,由 $\mathbf{F}$ 的特性可以得到 $\mathbf{F}^H \mathbf{F} = \mathbf{I}$ ,由于操作在对角矩阵上,式(3)可以写成

$$\mathbf{X}^H \mathbf{X} = \mathbf{F} \text{diag}(\hat{\mathbf{x}}^* \cdot \hat{\mathbf{x}}) \mathbf{F}^H \quad (4)$$

将式(4)代入式(2)可得线性回归的权值 $\mathbf{w}$ 的变换形式

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{x}}^* \cdot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \cdot \hat{\mathbf{x}} + \lambda} \quad (5)$$

这样可以将最小二乘分类器的训练时间复杂度从原本的矩阵求逆运算转换为矩阵的相对元素相乘与离散傅里叶变换。

### 3) 非线性回归的核相关滤波器。

可以采用核函数将输入的向量 $\mathbf{x}$ 映射到特征空间 $\phi(\mathbf{x})$ 中,把式(1)的解表示为输入的线性组合,系数为

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

式中: $\mathbf{K}$ 为核矩阵,元素 $K_{ij} = k(x_i, x_j)$ 。使用核函数的分类器将样本特征映射到高维特征空间,使得分类更加精确。但是这样也会使大量样本训练的时间增加,所以通过循环矩阵构建训练样本可以减少计算。根据循环矩阵的性质经推导可得

$$\hat{\mathbf{a}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}_{\hat{\mathbf{y}}} + \lambda} \quad (7)$$

式中, $K_{\hat{\mathbf{y}}}$ 是核矩阵 $\mathbf{K}$ 的第一行, $\hat{\mathbf{k}}_{\hat{\mathbf{y}}}$ 表示 $K_{\hat{\mathbf{y}}}$ 的离散傅里叶变换。

### 4) 核相关滤波器的响应。

当训练好分类器之后,将新的一帧里的图像特征输入分类器,来判断目标位置。这样,分类器的响应 $\mathbf{y} = \mathbf{w}^T \mathbf{z}$ 结合之前的公式,可得

$$\mathbf{y} = \mathbf{F}^{-1}(\hat{\mathbf{k}}_{\mathbf{z}} \cdot \hat{\mathbf{a}}) \quad (8)$$

式中, $\hat{\mathbf{k}}_{\mathbf{z}} = k(\mathbf{x}, \mathbf{z}_i)$ , $\mathbf{z}_i$ 为新一帧中获取的训练样本, $\mathbf{x}$ 为上一帧学习获得的目标模型,使 $\mathbf{y}$ 为最大的位置就是新一帧中目标位置。

## 2 本文的改进算法

针对KCF算法在目标发生尺度变化、严重遮挡时跟踪失败的问题,本文对KCF算法的目标特征、目标跟踪框和核函数进行了改进。

### 2.1 FHog特征

KCF算法采用Hog特征,只提取了目标的单通道灰度特征,其跟踪效果较差。本文用FHog特征替换Hog特征,采用有符号梯度和无符号梯度特征信息相结合的方式,以获取更多的目标特征信息,增强了算法的鲁棒性。具体过程如下:该特征取消了原Hog特征中的块(Block),只保留了单元(Cell),但是归一化时,直接将当前单元与其周围的4个单元(Cell)所组成的一个区域归一化,计算梯度方向时可以计算有符号( $0^\circ \sim 360^\circ$ )或无符号( $0^\circ \sim 180^\circ$ )的梯度方向。FHog特征采用了有符号梯度和无符号梯度相结合策略。取FHog特征计算过程,如图1所示。

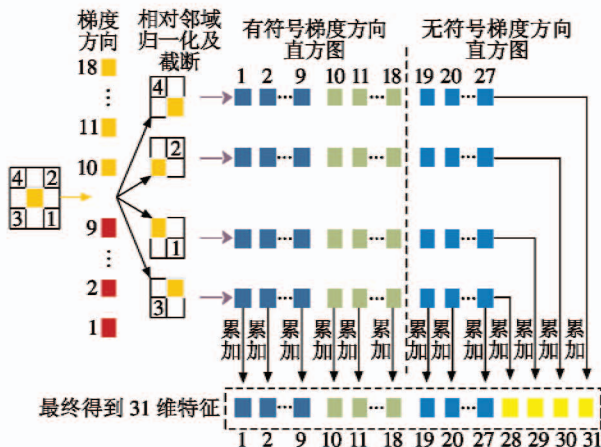


图1 FHog特征图

Fig.1 FHog feature diagram

如果直接将特征向量化,一个 $8 \times 8$ 的单元的特征维度数高达108,由于大量单元的无符号梯度中,每个单元共有36维特征,进行主成分分析,发现前11维特征基本上包含了所有信息,为了计算简便,无符号特征采用了11维,再加上18维有符号特征,共计31维特征向量,有的是32维特征向量,即第32维特征向量为 $\theta$ 。

### 2.2 尺度自适应

KCF算法的跟踪框是固定的,当目标发生尺度变化时,会丢失很多目标特征信息,从而导致跟踪失败。针对这种现象,本文采用了目标尺度估计策略,使目标框能自适应,以获取完整的目标特征信息,提高了跟踪成功率。具体过程如下:本文在程序中设置了不同的跟踪框调整率 $search\_size$ ,在每帧计算目标最佳位置时,由不同的调整率来调整跟踪框大小,利用傅里叶反变换得出不同的跟踪框与当前模板的相关度,再用 $find()$ 函数找到相关度值最大的所在窗口,继而确定出目标最佳的中心位置和长宽。以视频集中的Dog1视频序列为例,如图2所示。

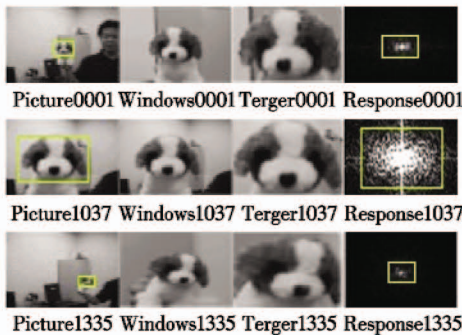


图2 图像的傅里叶频谱反应图

Fig.2 Fourier spectral response of an image

图中:Picture \* 为视频中目标最佳位置图;Windows \* 为目标搜索区域图;Terger \* 为最佳目标区域图;

Response \* 为最佳目标区域的傅里叶频谱图,图片名字的后4位数字为在该视频序列的第几帧。

### 2.3 多项式核函数

KCF算法采用的是高斯核函数,高斯核函数公式含有除法和指数级运算,计算比较耗时。本文用多项式核函数替换高斯核函数来减少计算量,提高运行速度。多项式核函数的函数形式为

$$K(x, x) = ((x \cdot x) + 1)^q \quad (9)$$

其函数为 $q$ 阶的多项式判别式函数,计算比较简单。经过实验验证,在运行速度上,多项式核函数比高斯核函数在每100帧上提高了约0.5s。

### 2.4 算法框架

本程序的大致框架主要采用了两个for循环函数,遍历所有视频集的图片。程序的具体流程如下。

for 单个视频 in 视频集:

调整率  $search\_size()$ 列表;

if not 第一帧:

for 调整率 in 列表:

$region(img, pos, w\_sz)$  赋给  $p1$ ;

$fft2(features(p1), cell\_size, cos\_windows)$  赋给  $zf$ ;

$p\_corre(zf, m\_xf)$  赋给  $kzx$ ;

$real(iffit2(m\_af \cdot kzx))$  赋给  $response(i)$ ;

$find(max(response(:)))$  赋给  $pos$ ;

$t\_sz \cdot scale$  赋给  $t\_sz$ ;

$region(img, pos, w\_sz)$  赋给  $p2$ ;

$fft2(features(p2), cell\_size, cos\_windows)$  赋给  $xf$ ;

$p\_corre(xf, xf)$  赋给  $kxx$ ;

$yf/(xf + \lambda)$  赋给  $af$ ;

if 第一帧:  $f=1$  else  $f=factor$

$f \cdot af + (1-f) \cdot af0$  赋给  $af0$ ;

$f \cdot xf + (1-f) \cdot xf0$  赋给  $xf0$ ;

其中: $w\_sz$ 为目标搜索区域( $W \times H$ ), $sz$ 为目标区域( $w \times h$ ); $p\_corre()$ 为多项式核相关函数。

## 3 实验结果与分析

为验证本文算法的有效性,实验在CPU为AMD A8-4500M、内存为4 GiB的PC机上使用Matlab2016a开发平台进行编程实现,利用OTB-2013测试平台中50组视频数据集进行评估,并与KCF算法、MOSSE算法以及OTB-2013测试平台<sup>[13]</sup>中包含的29种跟踪算法进行了对比。视频数据集包含以下属性:光照变化(IV)、尺度变化(SV)、遮挡(OCC)、形变(DEF)、运动模糊(MB)、快速运动(FM)、平面内旋转(IPR)、平面外旋转(OPR)、出视野(OV)、复杂背景(BC)、低分辨率(LR)。

### 3.1 实验结果

从32种跟踪算法中挑选出排名在前10的跟踪算



法,并将其在视频数据集上运行。图3所示为从数据集中挑选出目标在发生尺度变化、严重遮挡的情况下,这10种跟踪算法的跟踪结果。



图3 前10种跟踪算法运行结果图

Fig. 3 The results of the top ten tracking algorithms

从图3中可以定性地看出,本文算法在目标发生尺度变化、严重遮挡时,相对于其他几种算法有较好的跟踪效果和较强的鲁棒性。

### 3.2 实验分析

为了评估跟踪算法的性能,实验采用OTB-2013测试平台中两种评估方法作为定量评价指标,即精确度图和成功率图。精确度图能够显示目标中心位置误差小于阈值的帧数占视频总帧数的百分比。实验中阈值选定为20帧。成功率图能够显示重叠率 $S$ 大于给定阈值 $t_0$ 的帧数占视频总帧数的百分比。其中,重叠率 $S$ 为 $S = |R_t \cap R_a| / |R_t \cup R_a|$ ,  $R_t$ 为跟踪的目标框面积,  $R_a$ 为准确的目标框面积,  $\cap$ 和 $\cup$ 分别表示两个区域的交集与并集,  $| \cdot |$ 为区域内的像素点的数量。成功率给出了阈值 $t_0$ 从0到1变化时,成功的帧数所占的比例,利用成功率图曲线下的面积(AUC)作为跟踪算法性能的评价准则。

实验利用OTB-2013测试平台的评估策略,对共计32种跟踪算法进行性能比较。为了显示清楚,本文在精确度图和成功率图上只显示排序靠前的10个跟踪算法的结果,如图4、图5所示。

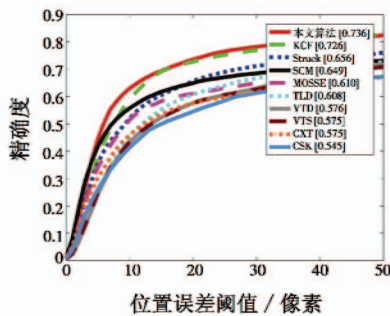


图4 精确度图

Fig. 4 Precision plot

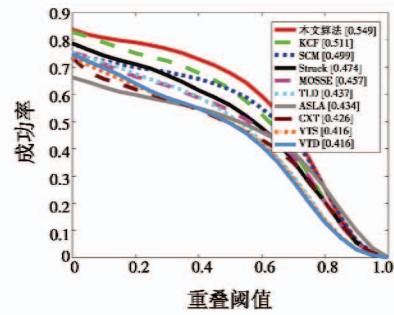


图5 成功率图

Fig. 5 Success rate plot

从图4、图5中可以看出,本文算法均排第一:在精确度图中,本文算法精确度为0.736,相比于KCF算法和MOSSE算法分别提高了1.0%和12.6%;在成功率图中,本文算法成功率为0.549,相比于KCF算法和MOSSE算法分别提高了3.8%和9.2%。这说明本文算法具有较强的鲁棒性,提高了原来算法的整体跟踪性能。

为了充分详细地评估本文算法的跟踪性能,利用OTB-2013测试平台中视频数据集的11个属性进一步评估本文算法,在11个属性的精确度图(如图6所示)中,本文算法在尺度变化(SV)、形变(DEF)、遮挡(OCC)、平面外旋转(OPR)方面有较大的改善,本文算法相对于KCF算法分别提高了5.3%,3.0%,6.8%,0.6%。

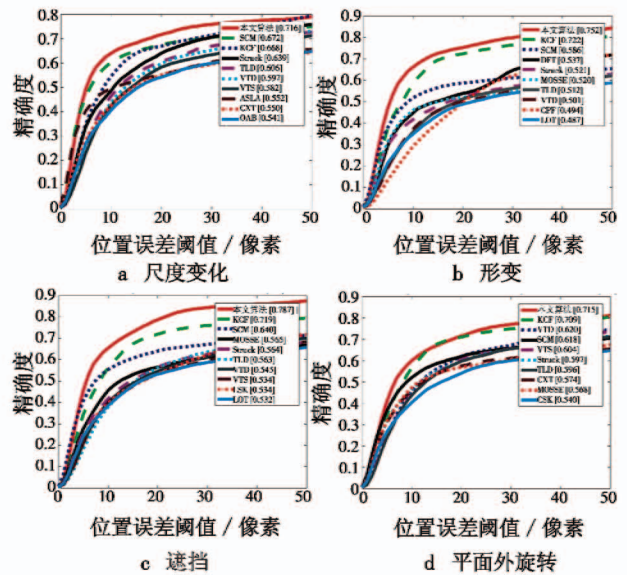


图6 不同属性的精确度图

Fig. 6 Precision plot of attributes

在11个属性的成功率图(如图7所示)中,本文算法在尺度变化(SV)、形变(DEF)、遮挡(OCC)、平面外旋转(OPR)、光照变化(IV)、超出视野(OV)、平面内旋转(IPR)、快速运动(FM)方面有较大的改善,本文算法相对于KCF算法分别提高了14.8%,9.1%,12.3%,



4.8%, 2.0%, 0.6%, 0.3% 和 0.9%。

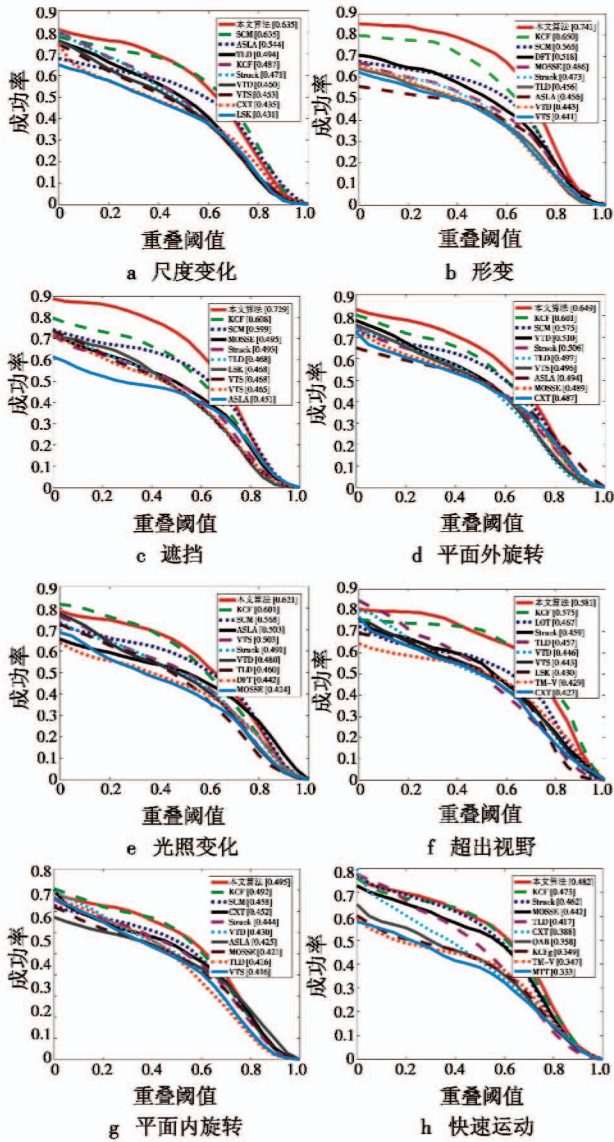


图7 属性的成功率图

Fig.7 Success rate of attributes

通过以上数据的对比分析可知,本文算法在目标发生尺度变化、严重遮挡等复杂情况下,均具有较强的稳健性和鲁棒性。

### 4 结束语

针对 KCF 算法在目标发生尺度变化、严重遮挡等复杂情况下出现跟踪失败的问题,本文提出了基于自适应的核相关滤波的目标跟踪算法。在本文算法中首先采用尺寸估计策略,使跟踪框更好地适应目标的尺寸变化。用 FHog 特征代替原来的 Hog 特征,获取了更多的目标信息,为分类器正确分类提供了更多的依据。对 OTB-2013 的 50 个数据集进行实验,取得较好的效果,对 KCF 算法有了较大的改善,有较强的鲁棒性,整体综合性能较好,但是本文算法只采用 FHog 特征,未

综合采用其他特征,有待改善。

### 参考文献

[1] WANG L, LIU T, WANG G, et al. Video tracking using learned hierarchical features [J]. IEEE Transactions on Image Processing, 2015, 24(4):1424-1435.

[2] WANG L J, OUYANG W L, WANG X G, et al. Visual tracking with fully convolutional networks [C]//Proceedings of the IEEE International Conference on Computer Vision, 2015:3119-3127.

[3] LI H X, LI Y, PORIKLI F. DeepTrack: learning discriminative feature representations online for robust visual tracking [J]. IEEE Transactions on Image Processing, 2016, 25(4):1834-1848.

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

[5] GAO J, LING H B, HU W M, et al. Transfer learning based visual tracking with Gaussian processes regression [C]//European Conference on Computer Vision, 2014:188-203.

[6] 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.

[7] 潘振福,朱永利.多尺度估计的核相关滤波器目标跟踪方法 [J]. 激光与光电子学进展, 2016, 53(10):101501-1-101501-7.

[8] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2010:2544-2550.

[9] 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.

[10] 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.

[11] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:886-893.

[12] FELZENSZWALB P F, GIRSHICK R B, D. MCALLESTER D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9):1627-1645.

[13] WU Y, LIM J, YANG M H. Online object tracking: a benchmark [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2013:2411-2418.