

DOI: 10.13973/j.cnki.robot.2016.0420

# 基于改进核相关滤波器的 PTZ 摄像机控制方法

潘振福, 朱永利, 周国亮

(华北电力大学控制与计算机工程学院, 河北 保定 071003)

**摘要:** 针对传统的 PTZ (pan-tilt-zoom) 摄像机控制方法是依靠人工操作, 无法连续、实时跟踪动态目标, 且跟踪目标准确度低等缺陷, 提出了一种基于改进的核相关滤波器 (KCF) 目标跟踪算法的 PTZ 摄像机控制方法。首先, 对传统的 KCF 目标跟踪算法做了运动状态估计和尺度估计方面的改进。在目标运动状态估计中, 将粒子滤波框架与传统 KCF 算法相结合, 估算出运动目标的位置。这种基于概率的运动状态估计方法可以获得更加稳定的目标信号并同时减少背景干扰信息的引入, 从而可以在复杂场景下具有更强的抗干扰性。目标尺度估计中采用相关滤波器在尺度金字塔中估算目标的尺度, 使算法对尺度变化的运动目标具有更强的适应能力。其次, 根据跟踪结果信息, 通过 PELCO.D 协议控制 PTZ 摄像机, 始终保持目标在视频画面内。最后, 将改进 KCF 算法与其他跟踪算法在 Benchmark 数据集中做对比实验, 验证改进算法的鲁棒性与有效性。将算法应用于 PTZ 摄像机的控制中, 并用 C++ 语言实现了改进 KCF 算法控制 PTZ 摄像机上位机, 实验结果表明该 PTZ 摄像机控制方法能准确跟踪被遮挡目标, 使其稳定地呈现在取景框中。

**关键词:** 目标跟踪; 云台摄像机; 相关滤波器; 粒子滤波; 运动状态估计; 尺度估计

中图分类号: TP391

文献标识码: A

文章编号: 1002-0446(2016)-04-0420-08

## The PTZ Camera Control Method Based on the Improved Kernelized Correlation Filter

PAN Zhenfu, ZHU Yongli, ZHOU Guoliang

(School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China)

**Abstract:** Traditional PTZ (pan-tilt-zoom) camera control methods rely on manual operation. By those methods, dynamic objects can't be tracked continuously in real time, and the accuracy of object tracking is low. To solve those problems, a PTZ camera control method based on an improved KCF (kernelized correlation filter) object tracking algorithm is proposed. Firstly, the traditional KCF object tracking algorithm is improved in terms of motion state estimation and scale estimation. In target motion state estimation, the particle filter framework is combined with the traditional KCF algorithm to estimate the position of the moving target. With the motion state estimation method based on probability, stabler target signals can be obtained, and background interference information is reduced, thereby stronger anti-jamming ability is accomplished in complex scenarios. In target scale estimation, correlation filter is applied to estimating the scale of the target in scale pyramid, which improves the algorithm adaptability to the scale changes of moving targets. Secondly, the PTZ camera is controlled with the PELCO.D protocol according to the information of tracking results in order to keep the target within the viewfinder. Finally, the comparison experiments are carried out between the improved KCF algorithm and the other tracking algorithms using the Benchmark data sets in order to verify the effectiveness and robustness of the improved algorithm. The algorithm is applied to the PTZ camera control, and the PC system of PTZ camera is controlled by the improved KCF algorithm with C++ language. The experiment results show that the PTZ camera control method can track the obscured target accurately and keep it in the viewfinder stably.

**Keywords:** object tracking; PTZ (pan-tilt-zoom) camera; correlation filter; particle filtering; motion state estimation; scale estimation

## 1 引言 (Introduction)

PTZ (pan-tilt-zoom) 摄像机是将 CCD (charge-coupled device) 摄像头、变焦变倍镜头和全景云台等部件集成在一个单元中所构成的摄像机系统。传

统的 PTZ 摄像机监控系统通过人工操控键盘或者摇杆对监控目标进行跟踪, 如果目标较远, 通过手动控制镜头变焦。这种人工操控的方式造成跟踪误差太大, 系统摄像不稳定、不连续。近年来, 基于

主动视觉的 PTZ 摄像机系统逐渐发展起来, 陈双叶等<sup>[1]</sup>提出了基于 MeanShift 算法的 PTZ 摄像机控制方法, 简单地采用颜色直方图作为跟踪匹配特征, 容易受到光照变化、运动模糊遮挡等因素影响而造成跟踪失败. Wang 等<sup>[2]</sup>提出的高斯混合模型目标检测定位的 PTZ 控制方法仅仅解决了镜头对焦问题. 而 Dong 等<sup>[3]</sup>除了应用高斯混合模型目标检测之外, 还融合了卡尔曼滤波器与 CamShift 算法对目标进行跟踪, 从而控制 PTZ 摄像机. 但相比近年来先进的目标跟踪技术, 这些跟踪算法相对落后, 其性能容易受众多因素影响, 包括光照变化、遮挡和背景混杂等. 视觉目标跟踪 (visual object tracking, VOT) 2014 年竞赛的结果显示<sup>[4]</sup>, 基于分类相关滤波器的跟踪器是目前最优秀的目标跟踪算法之一, 提供了先进的跟踪性能. Danelljan 等<sup>[5]</sup> 2014 年提出一种自适应颜色属性 CN (color name) 跟踪方法, 将颜色特征属性集成到 CSK (circulant structure of tracking-by-detection with kernels) 跟踪器<sup>[6]</sup>中, 并通过把多维特征映射到高斯核空间中进行学习, 得到一个自适应相关滤波器; Henriques 等<sup>[7]</sup>对 CSK 方法进行了扩展, 使用了梯度方向直方图 (HOG, histogram of oriented gradient) 特征, 提出了核相关滤波器 (KCF, kernelized correlation filter) 及双相关滤波器 (DCF, dual correlation filter) 跟踪器, 运算速度能达到 300 bit/s 以上. 由 Danelljan 等在 CSK 以及最小输出均方误差和 (MOSSE, minimum output sum of squared error) 滤波算法<sup>[8]</sup>的基础上提出的基于相关滤波器的分类尺度空间跟踪器 (DSST, discriminative scale space tracker)<sup>[9]</sup>的性能是视觉目标跟踪 VOT2014 竞赛中最佳的.

本文融合了 KCF 算法与 DSST 算法的优势, 对运动目标进行跟踪, 并采用粒子滤波框架对目标进行运动状态预测, 提高跟踪准确度 (实现源码: [https://github.com/panzhenfu/KCF\\_DSST\\_Particle\\_tracker.git](https://github.com/panzhenfu/KCF_DSST_Particle_tracker.git)). 然后根据跟踪结果对 PTZ 摄像机进行控制.

## 2 PTZ 摄像机系统架构 (The framework of the PTZ camera system)

本文中 PTZ 摄像机系统由云台摄像机、视频采集卡和运行在 PTZ 控制系统的上位机组成, 如图 1 所示. 将摄像头采集回来的图像传入算法中并进行跟踪运算, 通过 RS232 转 RS485 串口遵循 PELCO\_D 协议控制 PTZ 摄像机转动, 保持跟踪目标在视频图像中, 控制功能如图 2 所示. PELCO\_D 协议定义了水平旋转、上下旋转和变焦变倍等功能

的控制指令格式.

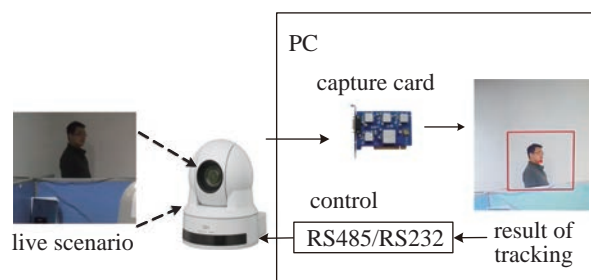


图 1 PTZ 摄像机系统架构

Fig.1 The framework of the PTZ camera system

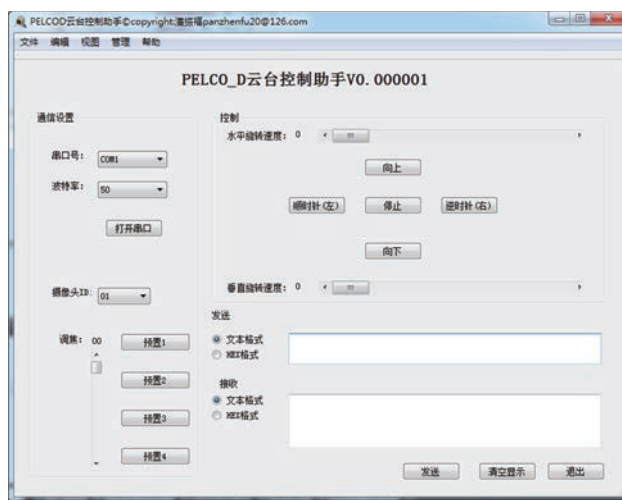


图 2 PELCO\_D 协议控制助手

Fig.2 Control assistant based on PELCO\_D protocol

## 3 目标跟踪方法 (Object tracking method)

本方法把传统的 KCF 目标跟踪算法整合进粒子滤波框架中, 对运动目标进行了运动状态预测, 并融合了 DSST 算法进行尺度估计, 弥补了传统 KCF 算法对固定尺度运动目标跟踪的缺陷. 具体流程如图 3 所示, 其中粒子群  $S$  在 3.2 小节详细介绍,  $p$  为粒子的初始位置,  $v$  为粒子的运动速度,  $r$  为该粒子经过核相关滤波器后得到的最高响应值,  $\pi$  为该粒子的权重值. 下文中以 “.” 运算表示某个粒子  $s$  的某个域, 如  $s.\pi$  表示粒子  $s$  的权重值.  $N$  为粒子数, 该算法先是将粒子滤波框架与位置核相关滤波器结合, 估算出目标位置, 再进行尺度相关滤波来估算目标尺度, 如此循环.

### 3.1 KCF 跟踪算法

在 KCF 跟踪算法中, 主要以核岭回归分类器作为核心. 对样本进行巧妙的变换, 使训练数据矩阵具有循环特性, 构建出的这种循环样本矩阵可以通过离散傅里叶变换实现对角化. 使用这种循环矩阵来训练核岭回归分类器可以减少几个数量级的存储与计算量, 从而达到对目标进行快速有效检测跟

踪的目的, 这就是文 [7] 提出的核岭回归相关滤波器.

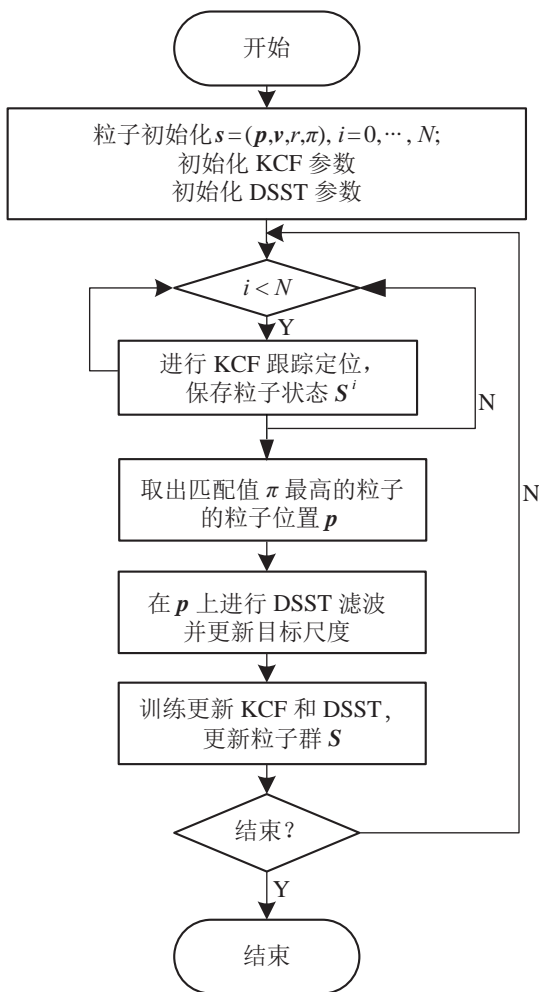


图 3 PTZ 摄像机控制算法流程图

Fig.3 Flow chart of the control algorithm of the PTZ camera

### 3.1.1 核岭回归分类器

KCF 跟踪算法采用文 [10] 的方法在岭回归分类器中引入核函数, 即  $f(z) = \mathbf{w}^T \mathbf{z}$ . 设从样本空间到 Hilbert 特征空间的非线性变换为  $\mathbf{x}_i \rightarrow \boldsymbol{\varphi}(\mathbf{x}_i)$ , 定义核函数为  $\kappa(\mathbf{x}, \mathbf{x}') = \langle \boldsymbol{\varphi}(\mathbf{x}), \boldsymbol{\varphi}(\mathbf{x}') \rangle$ ; 则优化问题的最优解可以表示为

$$\min_{\mathbf{w}} \sum_i ((\mathbf{w}, \boldsymbol{\varphi}(\mathbf{x}_i)) - y_i)^2 + \lambda \|\mathbf{w}\|^2 \quad (1)$$

其中,  $\mathbf{x}_i$  为样本,  $y_i$  为期望输出,  $\lambda$  为一个固定的正则化参数. 根据文 [11] 可知其最优解为样本的线性组合  $\mathbf{w}$ :

$$\mathbf{w} = \sum_i \alpha_i \boldsymbol{\varphi}(\mathbf{x}_i) \quad (2)$$

由式 (2) 可知, 需要优化的变量是  $\alpha_i$ , 而不是  $\mathbf{w}$ .

文 [12] 给出了核岭回归问题的解

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (3)$$

其中,  $\boldsymbol{\alpha}$  为需要优化的参数向量,  $\mathbf{K}$  是核矩阵,  $\mathbf{I}$  是一个单位矩阵,  $\mathbf{y}$  为期望输出.

### 3.1.2 循环矩阵

对样本  $\mathbf{x}$  进行移位变换, 可以获得循环矩阵  $\mathbf{X}$ , 即

$$\mathbf{X} = \mathbf{C}(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \quad (4)$$

注意, 矩阵  $\mathbf{X}$  完全由指定的样本向量  $\mathbf{x}$  (这里的第 1 行) 循环移位生成, 所以  $\mathbf{C}(\mathbf{x})$  表示对样本向量  $\mathbf{x}$  循环移位运算得到的矩阵. 无论生成向量  $\mathbf{x}$  为何值, 所有循环矩阵都能通过离散傅里叶变换 (DFT) 实现矩阵对角化 [13]. 也就是

$$\mathbf{X} = \mathbf{F} \text{diag}(\hat{\mathbf{x}}) \mathbf{F}^H \quad (5)$$

其中,  $\mathbf{F}$  是一个与  $\mathbf{x}$  无关的常量矩阵, 用来计算 DFT, 如  $\mathcal{F}(z) = \sqrt{n} \mathbf{F} z$ ,  $\mathbf{F}^H$  是  $\mathbf{F}$  的共轭转置, 而  $\hat{\mathbf{x}}$  表示生成向量  $\mathbf{x}$  的离散傅里叶变换,  $\hat{\mathbf{x}} = \mathcal{F}(\mathbf{x})$ . 下文都用 “ $\hat{\cdot}$ ” 来表示一个向量的离散傅里叶变换.

### 3.1.3 目标快速检测

根据文 [7] 提出的定理 1, 如果核函数  $\kappa(\mathbf{x}, \mathbf{x}')$  是酉不变的, 则有核矩阵  $\mathbf{K}$  为循环矩阵. 因此, 式 (3) 可根据循环矩阵的特性得到频域下的解:

$$\hat{\boldsymbol{\alpha}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}} + \lambda} \quad (6)$$

其中,  $\mathbf{K} = \mathbf{C}(\mathbf{k})$ ,  $\mathbf{k}$  的元素  $k_i = \kappa(\mathbf{x}, \mathbf{x}_i)$ .

假设核矩阵  $\mathbf{K}^{\mathbf{z}\mathbf{x}}$  由匹配模板和搜索窗口样本核相关所得. 因为匹配模板  $\mathbf{X}$  和检测样本  $\mathbf{Z}$  分别是基于元素向量  $\mathbf{x}$  和  $\mathbf{z}$  移位变换的循环结构矩阵, 相应的,  $\mathbf{K}^{\mathbf{z}\mathbf{x}}$  的每个元素是由  $\kappa(\mathbf{P}^{i-1} \mathbf{z}, \mathbf{P}^{j-1} \mathbf{x})$  组成的一个循环核矩阵, 其中  $\mathbf{P}$  为循环移位的单位阵. 只需第 1 行元素来表示核矩阵

$$\mathbf{K}^{\mathbf{z}\mathbf{x}} = \mathbf{C}(\mathbf{k}^{\mathbf{z}\mathbf{x}}) \quad (7)$$

其中  $\mathbf{k}^{\mathbf{z}\mathbf{x}}$  是  $\mathbf{x}$  与  $\mathbf{z}$  核运算的结果. 可以把回归函数表示成

$$f(\mathbf{z}) = \mathbf{w}^T \mathbf{z} = \sum_i \alpha_i \kappa(\mathbf{x}_i, \mathbf{z}) \quad (8)$$

其中,  $\mathbf{x}$  为上一帧学习的目标匹配模型,  $\mathbf{x}_i$  为  $\mathbf{x}$  第  $i$  次循环移位的结果,  $\mathbf{x}_i = \mathbf{P}^{i-1} \mathbf{x}$ ,  $n$  为  $\mathbf{x}$  元素的个数,

图像块  $\mathbf{z}$  为训练样本. 考虑图像块  $\mathbf{z}$  的循环移位, 可得:

$$\mathbf{F}(\mathbf{z}) = \begin{bmatrix} \sum_i^n \alpha_i \kappa(\mathbf{x}_i, \mathbf{z}_1) \\ \sum_i^n \alpha_i \kappa(\mathbf{x}_i, \mathbf{z}_2) \\ \sum_i^n \alpha_i \kappa(\mathbf{x}_i, \mathbf{z}_3) \\ \vdots \\ \sum_i^n \alpha_i \kappa(\mathbf{x}_i, \mathbf{z}_m) \end{bmatrix} = (\mathbf{K}^{\mathbf{z}\mathbf{z}})^T \boldsymbol{\alpha} \quad (9)$$

其中,  $\mathbf{z}_j$  为图像块  $\mathbf{z}$  的第  $j$  次循环移位结果, 即,  $\mathbf{z}_j = \mathbf{P}^{j-1} \mathbf{z}$ ,  $m$  为  $\mathbf{z}$  的元素个数. 注意到  $\mathbf{F}(\mathbf{z})$  的结果是一个向量, 其元素为向量的所有移位变换形式的输出响应, 为了更有效率地计算式 (9), 根据循环矩阵的特性可得到对角化后的等式:

$$\hat{\mathbf{F}}(\mathbf{z}) = \hat{\mathbf{k}}^{\mathbf{z}\mathbf{z}} \cdot \hat{\boldsymbol{\alpha}} \quad (10)$$

其中, “ $\cdot$ ” 为矩阵点乘运算,  $\mathbf{z}$  是预测包含目标的搜索窗口图像块, 从中检测目标的位置,  $\mathbf{x}$  为从以前帧中学习获得的目标模型. 对式 (10) 进行 DFT 逆运算, 就可得出检测样本  $\mathbf{z}$  的响应矩阵, 响应最大的位置即为目标所在的位置. 更新模型  $\hat{\boldsymbol{\alpha}}$  和  $\hat{\mathbf{x}}$ :

$$\hat{\boldsymbol{\alpha}}_t = (1 - \eta) \hat{\boldsymbol{\alpha}}_{t-1} + \eta \hat{\boldsymbol{\alpha}} \quad (11)$$

$$\hat{\mathbf{x}}_t = (1 - \eta) \hat{\mathbf{x}}_{t-1} + \eta \hat{\mathbf{x}} \quad (12)$$

其中,  $\eta$  为学习因子,  $\hat{\boldsymbol{\alpha}}_t$  和  $\hat{\boldsymbol{\alpha}}_{t-1}$  分别表示当前帧和上一帧更新获得的系数向量傅里叶变换  $\hat{\boldsymbol{\alpha}}$ ,  $\hat{\mathbf{x}}_t$  和  $\hat{\mathbf{x}}_{t-1}$  分别表示当前帧和上一帧更新获得的目标匹配模型的离散傅里叶变换  $\hat{\mathbf{x}}$ .

### 3.2 目标运动状态估计

粒子滤波器 (particle filter) [14] 是一种基于贝叶斯理论和蒙特卡洛方法的求解后验概率的实用性滤波器, 可以根据带噪声的数据估计运动状态. 在目标的状态空间中通过传播带权离散随机变量来近似目标状态的概率分布. 这里的随机变量被形象地称为粒子, 当粒子个数趋于无穷时可以逼近任何形式的概率分布.

粒子滤波器的关键思想是通过带权样本粒子群  $\mathbf{S} = (\mathbf{s}^n | n = 1, \dots, N)$  近似概率分布. 为了算法的代码实现, 本文对粒子群向量化, 采用  $\mathbf{S}$  表示. 每个样本粒子  $\mathbf{s}$  代表核相关滤波器的响应模型.

#### 3.2.1 粒子目标模型与粒子集合

采用第 2 节核相关滤波器的最高响应值分布函数作为目标位置模型. 定义每个核相关滤波

器最高响应值为分布的采样样本, 即粒子模型为

$$\mathbf{s} = (\mathbf{p}, \mathbf{v}, r, \pi) \quad (13)$$

$$\mathbf{p} = (x, y) \quad (14)$$

$$\mathbf{v} = (v_x, v_y) \quad (15)$$

其中,  $\mathbf{p}$  为粒子的初始位置,  $(x, y)$  为粒子对应的坐标.  $\mathbf{v}$  为粒子的运动速度,  $v_x$  为粒子对应  $x$  方向上的速度,  $v_y$  为粒子对应  $y$  方向上的速度,  $r$  为该粒子经过核相关滤波器后得到的最高响应值,  $\pi$  为该粒子的权重值.

样本粒子群的更新通过系统状态变化方程传播:  $\mathbf{S}_t = \mathbf{A} \mathbf{S}_{t-1} + \mathbf{w}_{t-1}$ . 其中,  $\mathbf{A}$  为状态转移矩阵,  $\mathbf{w}_{t-1}$  为高斯噪声.

#### 3.2.2 粒子滤波器算法

初始化  $t-1$  时刻样本粒子群  $\mathbf{S}_{t-1}$  和目标模型, 粒子初始权值均为  $\frac{1}{N}$ , 初始状态为  $\mathbf{S}_0$ .

**重采样:** 从粒子群  $\mathbf{S}_{t-1}$  中根据权重  $\mathbf{s}_t^n \cdot \pi$  挑选  $N$  个样本:

(a) 对权重数组计算归一化累计概率  $c_{t-1}^k$ :

$$c_{t-1}^k = \frac{\sum_{i=1}^k \mathbf{s}_{t-1}^i \cdot \pi}{\sum_{i=1}^N \mathbf{s}_{t-1}^i \cdot \pi} \quad (16)$$

(b) 产生  $N$  个  $[0, 1]$  间满足均匀分布的随机变量, 构成数组  $r$ ;

(c) 对  $r$  中每个元素, 分别用二分查找法搜索归一化累计概率数组, 得到含有  $N$  个最小索引的数组  $j$ ,  $j$  满足  $c_{t-1}^{j[i]} \geq r[i]$ ;

(d) 更新样本粒子群,  $\mathbf{s}_{t-1}^i = \mathbf{s}_{t-1}^{j[i]}$ .

**传播:** 根据系统状态方程  $\mathbf{S}_t = \mathbf{A} \mathbf{S}_{t-1} + \mathbf{w}_{t-1}$ , 计算新的样本粒子群  $\mathbf{S}_t$ , 估计新粒子的位置.

**观测:** 根据状态粒子群  $\mathbf{S}_t$  中的每一个样本状态, 进行核相关滤波器位置估算, 然后更新估计, 获得新的权重概率.

(a) 对粒子群  $\mathbf{S}_t$  中的每一个粒子进行位置核相关滤波器滤波, 得到核相关滤波器的最高响应值, 更新粒子的状态:

$$\mathbf{s}_t^i \cdot r = \max(f_{\text{trans}}(\mathbf{s}_t^i) \cdot r) \quad (17)$$

$$\mathbf{s}_t^i \cdot \mathbf{p} = f_{\text{trans}}(\mathbf{s}_t^i) \cdot \mathbf{p} \quad (18)$$

其中,  $f_{\text{trans}}$  为粒子状态传播函数.

(b) 对于样本粒子群  $\mathbf{S}_t$  的每个粒子, 针对前一帧的目标状态  $\mathbf{s}_0$  更新粒子速度:

$$\mathbf{s}_t^i \cdot \mathbf{v} \cdot v_x = \mathbf{s}_t^i \cdot \mathbf{p} \cdot x - \mathbf{s}_0 \cdot \mathbf{p} \cdot x \quad (19)$$

$$\mathbf{s}_t^i \cdot \mathbf{v} \cdot v_y = \mathbf{s}_t^i \cdot \mathbf{p} \cdot y - \mathbf{s}_0 \cdot \mathbf{p} \cdot y \quad (20)$$

(c) 根据样本粒子群  $S_i$  计算各个粒子权重值:

$$s_i^i \cdot \pi = \frac{s_i^i \cdot r}{\sum_{k=0}^N s_i^k \cdot r} \quad (21)$$

**估计:** 根据权重估算样本粒子群  $S_i$  经过核相关滤波器后的最大响应值的最大值, 提取位置坐标作为目标跟踪输出:

$$s_0 = (s_i^k | s_i^k \cdot r = \max(s_i^j \cdot r), i = 1, \dots, N) \quad (22)$$

### 3.3 尺度估计

有尺度估计的 DSST 的跟踪精确度性能在 VOT2014 竞赛中是最佳的. 将尺度估计纳入跟踪器中是需要很高的计算代价的. 而在理想情况下, 精确的尺度估算方法应该鲁棒而且具有较高的计算效率. 为实现这一目标, 本文提出通过学习独立于定位滤波器的滤波器快速地进行尺度估算的方法. 这样有助于把尺度空间的搜索区域限制在更小范围. 此外, 还可以自由地选择每个独立滤波器的特征描述. 因此, 在上述定位框架下, 通过学习另一个 1 维相关滤波器来估算图像中的目标尺度.

#### 3.3.1 1 维相关分类滤波器

采用 1 维的特征映射表示图像信号. 从矩形图像块  $J$  中提取特征映射  $f$  表示目标. 目标是找一个最优的相关滤波器  $h$ , 使得 1 维分类相关滤波器的最小化误差输出为

$$\varepsilon = \|h * f - g\|^2 + \lambda_s \|h\|^2 \quad (23)$$

其中, “\*” 表示卷积运算,  $g$  是希望相关输出值,  $f$  为训练样本. 参数  $\lambda_s > 0$  为正则化参数. 由文 [9] 及式 (23) 可解得

$$\hat{h} = \frac{\hat{g}^* \hat{f}}{\hat{f}^* \hat{f} + \lambda_s} \quad (24)$$

其中, “^” 表示式 (23) 中对应符号的离散傅里叶变换. 上标 “\*” 表示矩阵的共轭矩阵, 例如  $\hat{g}^*$  表示  $g$  的共轭矩阵的离散傅里叶变换. 可以通过最小化输出误差  $\varepsilon$  对所有图像块训练获得最优滤波器  $h$ . 为了获得鲁棒的近似值, 分别更新相关滤波器  $\hat{h}_t$  (式 (24)) 的分子  $A_t$  和分母  $B_t$ :

$$A_t = (1 - \eta) A_{t-1} + \eta \hat{g}_t^* \hat{f}_t \quad (25)$$

$$B_t = (1 - \eta) B_{t-1} + \eta \hat{f}_t^* \hat{f}_t \quad (26)$$

其中, 下标  $t-1$  表示上一帧学习得到的相应的量,  $\eta$  是一个学习速率参数. 使用被检测矩形区域的特

征映射  $Z$  计算相关滤波响应向量  $Y$ . 然后通过寻找最大相关响应值  $y_{\max}$  估算新的目标尺度.

$$Y = \mathcal{F}^{-1} \left( \frac{\bar{A}Z}{B + \lambda} \right) \quad (27)$$

其中,  $\mathcal{F}^{-1}$  是离散傅里叶逆变换.

#### 3.3.2 特征金字塔

在一个已估算出目标位置的框架下, 在该位置上面提取样本的  $s$  层 PCA-HOG 特征金字塔. 使用 1 维高斯函数作为希望相关输出值  $g$ , 然后使用式 (27) 计算该特征金字塔相关响应向量  $Y$ , 通过最高相关响应值  $y_{\max}$  估算新的目标尺度.

让  $P \times R$  表示当前帧中的目标大小,  $s$  表示尺度大小. 对每个  $n \in \left\{ \left\lfloor -\frac{s-1}{2} \right\rfloor, \dots, \left\lfloor \frac{s-1}{2} \right\rfloor \right\}$ , 提取以目标为中心、大小为  $a^n P \times a^n R$  的图像块  $J_n$ . 其中  $a$  表示各个特征层之间的尺度因子. 提取图像块  $J_n$  的特征映射作为训练样本, 尺度金字塔的第  $n$  层的特征描述值为  $f(n)$ . 最后, 用式 (25) 和式 (26) 更新尺度滤波器  $h_{\text{scale}}$ . 整个跟踪算法流程如图 3 所示.

## 4 实验分析 (Experimental analysis)

### 4.1 跟踪算法比较实验

为了验证算法的有效性, 在 Benchmark 数据集 [15] 中选取 50 组视频图像序列进行测试. 改进的粒子滤波算法对其性能提高的验证如图 4 和图 5 所示, 将加有粒子滤波的算法命名为 OURSP, 只融合 KCF 算法与 DSST 算法而未加有粒子滤波的算法命名为 OURSNoP, 并与目前比较先进的视觉目标跟踪算法, 比如 DSST [9]、KCF [7]、ASLA (adaptive structural local sparse appearance) [16]、CSK [5] 跟踪算法进行对比.

#### 4.1.1 实验环境

实验平台为 Matlab R2010b, 所有的实验均在 Intel Core i3M390 CPU, 主频 2.67 GHz, 4 GB 内存配置的电脑上完成. 详细算法参数见源码.

#### 4.1.2 性能评估

从时间鲁棒性和空间鲁棒性 2 个方面评估本跟踪算法的精度和成功率 [17].

**精度 (precision):** 一个广泛使用的跟踪精度评价指标是中心位置误差, 即平均跟踪目标和手动标注的中心位置之间的欧氏距离. 然后用对所有的帧序列平均中心位置误差来总结算法对该序列的整体性能. 然而, 在追踪过程中失去目标时, 输出位置可以是随机的, 平均误差值可能是对跟踪性能的不正确统计 [18]. 目前的精度评估是采用跟踪估计位置



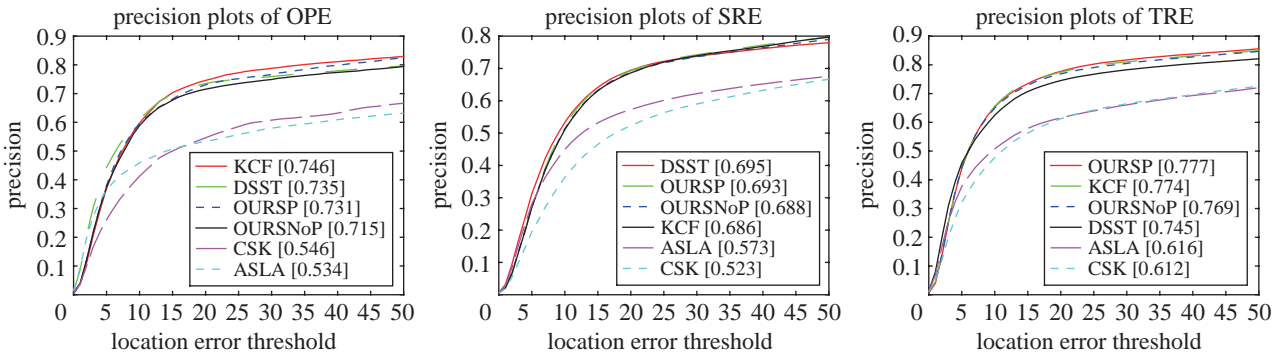


图 4 本文算法与其他算法精度比较结果

Fig.4 The accuracy comparison between the proposed algorithm and the other algorithms

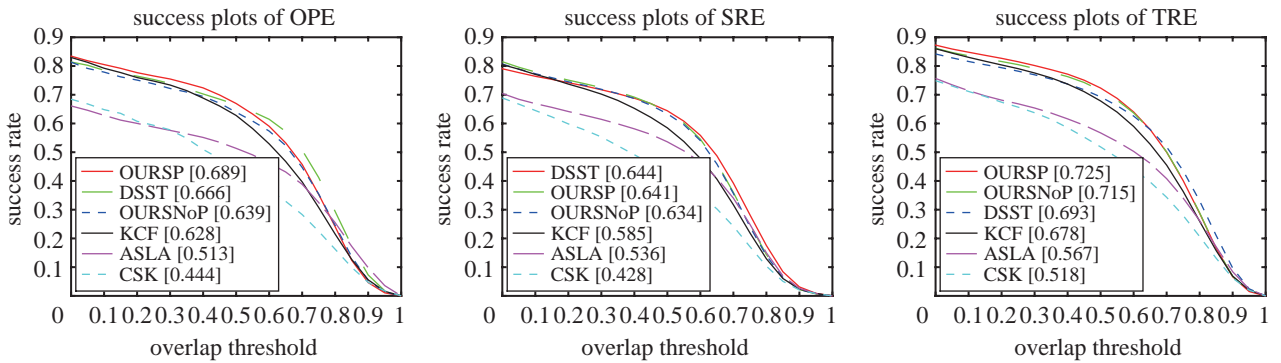


图 5 本文算法与其他算法成功率比较结果

Fig.5 The success rate comparison between the proposed algorithm and the other algorithms

的方法, 利用给定阈值距离范围内帧数的百分比来测量跟踪其整体性能. 一般采用阈值为 20 像素.

成功率 (success rate): 另一个评估指标是边界框重叠率, 即对跟踪边界框  $r_t$  和人工标注真实边界框  $r_a$ , 定义成功率为  $S = \frac{|r_t \cap r_a|}{|r_t \cup r_a|}$ , 其中 “ $\cap$ ” 和 “ $\cup$ ” 分别代表 2 个框区域的交集和并集, “ $|\cdot|$ ” 表示像素的数量. 然后计算跟踪帧数量的重叠率  $S$ , 并将  $S$  大于给定阈值  $t_0$  定义为成功跟踪, 以此测量算法在 1 个帧序列上的性能. 使用曲线下面积 (AUC) 统计估算跟踪算法, 而不是使用一个成功率值在一个特定阈值 ( $t_0 = 0.5$ ) 的评估.

鲁棒性评估 (robustness evaluation): 传统的追踪器评估方法是从人工标注初始化运行的第 1 帧开始测试序列并测量其平均精度或成功率, 称之为一遍过评估 (one-pass evaluation, OPE). 然而一些追踪算法可能对初始化很敏感, 不同的开始帧可能使得算法性能变得更糟或者更好. 因此, 从 2 个方面来分析初始化跟踪算法的鲁棒性, 即扰动初始化时间 (即在不同的帧开始) 和空间 (即从不同的边界框开始). 这些测试方法也称为时间鲁棒性评估 (temporal robustness evaluation, TRE) 和空间鲁棒性评估 (spatial robustness evaluation, SRE).

如图 4 所示, 本文的方法在 OPE 中与 KCF 持平, 比 DSST 高出 1.5%, 比 CSK 高出 36.6%, 比 ASLA 高出 39.7%, 比无粒子滤波的 OURSNoP 高出 2.23%.

在 SRE 精度评估中, 本文算法比 KCF 高出 1%, 比 ALSA 高出 20.9%, 比 CSK 高出 84.1%, 比无粒子滤波的 OURSNoP 高出 0.73%.

在 TRE 精度评估中, 本文算法比 KCF 高出 0.2%, 比 DSST 高出 4.2%, 比 ASLA 高出 26.0%, 比 CSK 高出 26.8%, 比无粒子滤波的 OURSNoP 高出 1.04%.

如图 5 所示, 在 OPE 成功率比较中, 本文算法比 DSST 高出 3.5%, 比 KCF 高出 9.7%, 比 ALSA 高出 34.4%, 比 CSK 高出 55.2%, 比无粒子滤波的 OURSNoP 高出 7.82%.

在 SRE 成功率比较中, 本文算法略逊于 DSST 0.4%, 但依然比 KCF 高出 9.6%, 比 ALSA 高出 19.6%, 比 CSK 高出 49.8%, 比无粒子滤波的 OURSNoP 高出 1.1%.

在 TRE 成功率比较中, 本文算法比 DSST 高出 4.6%, 比 KCF 算法高出 6.9%, 比 ALSA 高出 27.9%, 比 CSK 高出 40.0%, 比无粒子滤波的

OURSNoP 高出 1.39%.

### 4.1.3 质量与效率评估

图 6 所示为 Benchmark 数据集实验中的部分比较实验结果, 其中, 细实线方框为本文算法跟踪结果, 虚线方框为 DSST 跟踪结果, 点线为 ALSA 跟踪结果, 粗实线为 CSK 跟踪结果. 结果显示, 本文算法改善了传统的 KCF 算法固定尺度跟踪的缺陷, 具有很强的稳定性与鲁棒性.

表 1 所示为算法运行速度的比较, 本文算法跟踪效率比 DSST 高出 41.28%, 比 ALSA 高出 1998.13%, 具有很高的时效性, 可以胜任实时任务.

表 1 跟踪算法平均运行速率表 (帧/秒)

Tab.1 The average speeds of the tracking algorithms (frame/s)

算法	OURS	OURSNoP	DSST <sup>[9]</sup>	KCF <sup>[7]</sup>	ASLA <sup>[15]</sup>	CSK <sup>[5]</sup>
	29.185	41.906	20.657	129.867	1.391	209.634



(a) CarScale



(b) Couple



(c) David3



(d) FaceOcc2

— Ours — DSST — KCF — ASLA — CSK

图 6 本文算法与其他算法运行质量比较结果

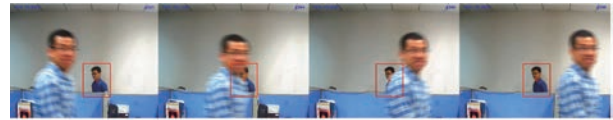
Fig.6 The operation quality comparison between the proposed algorithm and the other algorithms

## 4.2 PTZ 控制方法实验

在 QT5 平台上完成 PTZ 摄像机系统上位机端的 C++ 实现, 运用了 OPENCV 库.

### 4.2.1 跟踪质量评估

在实验室环境下, 运行结果如图 7 所示, 运动目标在被完全遮挡的情况下, 本系统的跟踪结果仍然很稳定.



(a) 相向运动物遮挡



(b) 同向运动物遮挡

图 7 目标物被遮挡跟踪情况

Fig.7 Tracking when the target is obscured

### 4.2.2 PTZ 控制质量实验

上位机通过 RS232 转 RS485 串口与 PTZ 摄像头相连, 遵循 PELCO\_D 协议控制. 控制结果如图 8 所示. 图中标有“摄像头”字样的设备为 PTZ 摄像机, 贴有白色的标识, 以便明显看到它转动, 图中右下角为摄像头跟踪的图像, 结果显示, 跟踪目标始终保持在图像中.



图 8 PTZ 摄像机跟踪目标旋转结果

Fig.8 Rotation results of the PTZ camera while tracking the target

## 5 结论 (Conclusion)

针对目前现有的基于主动视觉的 PTZ 摄像机控制系统的跟踪性能较差, 导致跟踪丢失等问题, 本文提出了一种基于改进核相关滤波器目标跟踪算法的 PTZ 摄像机控制方法. 在传统的 KCF 算法基础上对目标进行运动状态估计以及目标尺度空间估计, 提高算法鲁棒性. 在 Benchmark 数据集上, 将本文跟踪算法与传统的 KCF 算法及其他跟踪算法进行对比实验, 实验结果表明本文算法在目标尺度变化、光照变化、姿态变化、部分遮挡、旋转及快速运动等复杂情况下均有较强的适应性. 并在 QT5 平台上应用 OPENCV 库实现了 PTZ 摄像机控制系统的 C++ 版. 在实际场景中, 能够准确地跟踪给定目标, 根据目标在屏幕中的位置信息, 控制 PTZ 摄像机转动, 始终保持目标在屏幕中. 验证了本系统在实际场景应用中的适用性与稳定性, 在机器人

视觉导航以及智能监控等领域具有广泛的应用前景.

### 参考文献 (References)

- [1] 陈双叶, 王善喜. PTZ 摄像机跟踪运动目标的智能控制算法的研究[J]. 计算机科学, 2015, 42(11A): 135-139.  
Chen S Y, Wang S X. The study of the intelligent control algorithm for PTZ camera to track moving targets[J]. Computer Science, 2015, 42(11A): 135-139.
- [2] Wang S, Tian Y, Xu Y P. Automatic control of PTZ camera based on object detection and scene partition[C]//IEEE International Conference on Signal Processing, Communications and Computing. Piscataway, USA: IEEE, 2015: 1-6.
- [3] Dong E Z, Yan S X, Tong J G, et al. Design and implementation of a moving object tracking system[C]//IEEE International Conference on Mechatronics and Automation. Piscataway, USA: IEEE, 2015: 274-279.
- [4] Kristan M, Pflugfelder R, Leonardis A, et al. The visual object tracking VOT2014 challenge results[M]//Lecture Notes in Computer Science, vol.8926. Berlin, Germany: Springer-Verlag, 2015: 191-217.
- [5] Danelljan M, Khan F S, Felsberg M, et al. Adaptive color attributes for real-time visual tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2014: 1090-1097.
- [6] Henriques J F, Rui C, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[M]//Lecture Notes in Computer Science, vol.7575. Berlin, Germany: Springer-Verlag, 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, 2014, 37(3): 583-596.
- [8] Bolme D S, Beveridge J R, Draper B, et al. Visual object tracking using adaptive correlation filters[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2010: 2544-2550.
- [9] Danelljan M, Häger G, Khan F S, et al. Accurate scale estimation for robust visual tracking[C]//British Machine Vision Conference. 2014: 1-11.
- [10] Muller K R, Mika S, Ratsch G, et al. An introduction to kernel-based learning algorithms[J]. IEEE Transactions on Neural Networks, 2001, 12(2): 181-201.
- [11] Schölkopf B, Smola A J. Learning with kernels: Support vector machines, regularization, optimization, and beyond[M]. Boston, USA: MIT Press, 2001.
- [12] Rifkin R, Yeo G, Poggio T. Regularized least-squares classification[M]//Suykens J A K, Horvath I, Basu S, et al. Advances in learning theory: Methods, model and applications, vol.190. Amsterdam, Netherland: IOS Press, 2003: 131-154.
- [13] Gray R M. Toeplitz and circulant matrices: A review[J]. Foundations and Trends<sup>®</sup> in Communications and Information Theory, 2006, 2(3): 155-239.
- [14] 王法胜, 赵清杰. 一种用于解决非线性滤波问题的新型粒子滤波算法[J]. 计算机学报, 2008, 31(2): 346-352.  
Wang F S, Zhao Q J. A new particle filter for nonlinear filtering problems[J]. Chinese Journal of Computers, 2008, 31(2): 346-352.
- [15] Wu Y, Lim J, Yang M H. Visual tracker Benchmark[DB/OL]. [2013-10-12]. <http://www.visual-tracking.net>.
- [16] Jia X, Lu H, Yang M H. Visual tracking via adaptive structural local sparse appearance model[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2012: 1822-1829.
- [17] Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2013: 2411-2418.
- [18] 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.

### 作者简介:

潘振福 (1989-), 男, 硕士生. 研究领域: 计算机视觉, 人工智能.

朱永利 (1963-), 男, 博士, 教授, 博士生导师. 研究领域: 人工智能, 机器学习.

周国亮 (1978-), 博士, 副教授. 研究领域: 人工智能及应用.