



本科毕业设计（论文）

题目：动态场景中目标跟踪方法设计

院（系）计算机科学与工程学院

专 业计算机科学与技术

班 级15060116

姓 名杨拓

学 号15060116201

导 师胡秀华

2019 年 5 月 18 日

动态场景中目标跟踪方法设计

摘 要

运动目标跟踪作为计算机视觉领域的重要研究方向,可以获得目标运动轨迹信息,为行为分析、场景理解等高层次语义处理提供重要数据,得到了各研究学者的广泛关注,相关成果已成功应用在人机交互、安防监控、智能监控等诸多领域。然而,运动场景中目标受障碍物干扰、光照和尺度变化等复杂情况影响,目标表观动态改变,研究能够有效实现不同场景中目标稳定跟踪的方法具有重要的理论研究意义和工程实践价值。文中结合典型核相关滤波方法给出了一种融合多特征的多尺度抗遮挡跟踪方案,相关研究工作描述如下:

针对跟踪过程中出现的目标模糊和形变的情况,提取目标的 HOG 特征和颜色特征,然后进行融合,利用两者对于目标表征颜色变化和形状变化敏感度的互补性,得到一种对形变和模糊都不敏感的融合特征。

针对在跟踪过程中目标尺度会发生变化的情况,在跟踪目标位置发生更新的情况下加入多尺度检测,通过获取目标图像原本尺度,一个较大尺度和一个较小尺度的响应图的最大峰值,比较三个峰值大小,选取峰值最大的尺度作为最适合目标大小的目标跟踪框。

针对目标在跟踪过程中出现的短时遮挡问题,引入目标遮挡情况判断,并给出遮挡情况发生时的目标轨迹预测方法。通过检测最合适尺度的最大响应峰值是否小于设定阈值来判断目标是否发生遮挡,当目标发生遮挡后,通过卡尔曼滤波器来预测目标即将出现的位置。

最后,在基准库上对文中所提方法及相关典型算法进行性能评估,通过定性和定量评价结果可知,在尺度变换、短时部分遮挡、运动模糊、形变等多种复杂情况下,文中所提方法在成功率和准确率两个性能指标上均优于其他滤波类算法,并且具有较好的计算实时性。

关键词: 目标跟踪 核相关滤波 融合特征 多尺度 卡尔曼滤波

Research on method of object tracking in dynamic scenes

Abstract

As an important research direction in the field of computer vision, moving target tracking can obtain target trajectory information, provide important data for high-level semantic processing such as behavior analysis and scene understanding, and has received extensive attention from various research scholars. Relevant results have been successfully applied to people. Machine interaction, security monitoring, intelligent monitoring and many other fields. However, in the motion scene, the target is affected by complex obstacles such as obstacle interference, illumination and scale changes, and the target apparent dynamic changes. It is of great theoretical significance and engineering practice value to study the method of effectively achieving target stable tracking in different scenarios. In this paper, a multi-scale anti-occlusion tracking scheme with multi-features is presented in combination with typical kernel correlation filtering methods. The related research work is as follows:

According to the target blur and deformation in the tracking process, the HOG features and color features of the target are extracted, and then the fusion is performed. The complementarity between the two is used to characterize the color change and shape change sensitivity, and a pair of deformation and blur is obtained. Non-sensitive fusion features.

In the case that the target scale changes during the tracking process, multi-scale detection is added when the tracking target position is updated, and the maximum scale of the response image of a larger scale and a smaller scale is obtained by acquiring the original scale of the target image. Compare the three peak sizes and select the scale with the largest peak as the target tracking frame that best fits the target size.

Aiming at the short-term occlusion problem that occurs during the tracking process, the target occlusion situation is judged, and the target trajectory prediction method when the occlusion situation occurs is given. The target is occluded by detecting whether the maximum response peak of the most suitable scale is less than the set threshold. When the target is occluded, the Kalman filter is used to predict the upcoming position of the target.

Finally, the performance evaluation of the proposed method and related typical algorithms is carried out on the benchmark library. The qualitative and quantitative evaluation results show that in the complex cases of scale transformation, short-term partial occlusion, motion blur and deformation, the paper mentions The method is superior to other filtering algorithms in both performance rate and accuracy, and has good computational real-time performance.

Keywords: object tracking; kernel correlation filtering; fusion feature; multiscale; Kalman filter

目 录

中文摘要	1
英文摘要	1
1 绪论	1
1.1 研究背景及意义.....	1
1.2 目标跟踪方法研究现状.....	1
1.3 目标跟踪所面临的挑战.....	2
1.4 本文主要研究工作.....	3
1.5 本文的内容结构安排.....	4
1.6 本章小结.....	4
2 核相关滤波器理论	5
2.1 核相关滤波器理论基础.....	5
2.1.1 岭回归	5
2.1.2 循环移位和循环矩阵	5
2.1.3 傅里叶对角化	6
2.2 核相关滤波器的算法.....	6
2.2.1 核相关滤波器算法的特点	6
2.2.2 核相关滤波器算法所用到的两个数学模型	7
2.2.3 核相关滤波器算法实现流程	7
2.3 本章小结	8
3 一种基于核相关滤波器的多尺度抗遮挡目标跟踪算法	9
3.1 总体框架	9
3.2 跟踪器目标检测和初始化时涉及内容	10
3.2.1 方向梯度直方图特征 HOG.....	10
3.2.2 Lab 颜色模型	12
3.2.3 汉宁窗以及融合以上两种特征的方法.....	12
3.3 跟踪器模型更新时涉及内容.....	15

3.3.1 卡尔曼滤波跟踪器原理介绍以及在本算法中的应用	15
3.3.2 多尺度介绍以及在本算法中的应用	17
3.4 本章小结	19
4 实验结果及分析	20
4.1 实验所用数据集以及开发环境设置	20
4.2 与其他跟踪算法的定性比较	20
4.3 与其他跟踪算法的定量比较	26
4.3.1 跟踪算法的精确度和成功率指标介绍.....	26
4.3.2 跟踪算法的精确度和成功率评估结果.....	27
4.4 本章小结	32
5 结论与展望	34
致谢	36
参考文献	37
毕业设计（论文）知识产权声明.....	39
毕业设计（论文）独创性声明.....	40
附录	

1 绪论

1.1 研究背景及意义

运动目标跟踪技术已经有近 20 年的研究历史，这项技术的研究内容主要涉及运动目标检测与提取、运动目标跟踪、运动目标识别、运动目标行为分析和理解等诸多方面，是计算机视觉研究的重要分支^[1-3]。在此期间伴随着计算机通信技术、运算技术和信息技术的发展，以及计算机硬件，包括图像处理硬件的不断更新换代，运动目标跟踪技术也得到了快速发展，相关技术研究成果也得到广泛应用。

运动目标跟踪技术通常以图像处理，模式识别和人工智能等基础理论作为支撑，反过来也是这些理论研究的热点，其目的是赋予计算机视觉，模拟人类的双眼对于获得的可视化信息的处理以及产生信息反馈的过程。运动目标跟踪技术在最初应用于视频监控系统，后来在国防军事，智能机器人，智能交通指挥等现实领域中也有广泛的发展和应用。

运动目标跟踪作为数字图像处理和计算机视觉的一个热门研究领域在人类生产生活中具有很大的经济效益和应用价值，一直以来吸引着大量学者对其进行深入研究，但是受环境复杂、目标运动状态多样性等客观因素影响，目前还无法找到一个在所有条件下都适用的方法。近年来，国内外学者提出了许多方法用于改进运动目标跟踪技术，其目的就是为了提高运动目标检测和跟踪技术在不同应用条件下的准确性、实时性、鲁棒性，但是目前对运动目标跟踪的研究依然面临着巨大挑战。所以，进行运动目标跟踪的研究具有一定的研究价值。基于这个理由，本文选题以智能视频监控为应用背景，着重分析在动态场景中跟踪目标发生遮挡和多尺度变化的问题，对现有流行的核相关滤波跟踪算法进行改进，提出了一种基于核相关滤波器的多尺度抗遮挡目标跟踪算法。

1.2 目标跟踪方法研究现状

根据对目标的表现特征建立模型的方式，视觉目标跟踪算法可以分为生成式跟踪和判别式跟踪两种算法。在跟踪目标所在的一系列视频序列中，生成式跟踪算法会在视频序列目前播放的这一帧的目标所在区域提取特征（建立模型），然后在下一帧中根据上一帧中提取目标的特征去寻找与之拥有相同特征的位置，这个位置就会被看成跟踪算法给出的预测目标位置。而判别式跟踪算法的处理方式截然不同，它在视频序列播放的当前帧会采用一种叫正负样本目标判别方式，把跟踪目标所处于的位置区域判定为积极样本，其余的区域判定为消极样本，把这两种样本作为输入去训练一种叫做分类器的分类方法，利用这种分类方法去在下一帧寻找目标最有可能出现的位置，然后不断地重复这个过程；这样在视频每一

帧中通过正负样本的不断训练,分类器对于目标新出现位置和区域的判断也就会更加准确,这类方法通常把图像特征和机器学习结合在一起。目前,以相关滤波和深度学习为代表的判别式跟踪算法逐渐占据了主流地位,而在这之前稀疏编码作为生成式跟踪算法中的代表也拥有广泛的应用。以下分别简要概括上面提到的几种算法的基本思想:

稀疏表示:把输入信号用于一组给定的过完备字典线性表示,将其系数进行稀疏性约束,使得系数向量尽可能的为 0,这个过程被称作稀疏表示^[4,5]。利用稀疏表示的典型算法例如 L1Tracker,把候选样本通过稀疏表示成目标模版和琐碎模版,利用琐碎模版来处理遮挡问题,通过对稀疏系数的非负约束排除目标跟踪过程中的杂斑问题。

相关滤波:从信号处理领域中发展而来,两个信号之间的相似程度用相关性表示,相关操作用卷积^[6-13]表示。所以这种跟踪算法的基本思想就是通过找到一个滤波模版,让视频序列中下一帧图像与滤波模版进行卷积,预测的目标位置就是响应最大的区域。用这种思想实现的算法有很多,早期这种算法有平方误差最小和(MOSSE)算法,其实现的思想为最早出现的相关滤波思想。之后在 MOSSE 算法^[14]的基础上出现了改进的 CSK 算法^[15]和 KCF^[16]算法,这两种算法都使用了核方法,其作用是找出并且学习一组数据之间的相互关系。在这两种方法中,KCF 的算法性能最好,原因是 KCF 把循环矩阵的运算进行了简化,提高了算法的实时性和性能。但是 KCF 算法也具有两大弊端,第一是没有实现多尺度,第二是对于目标遮挡的情况会跟踪失败。所以为了解决这两个问题,两种算法应运而生,DSST 算法^[17]通过实现了尺度变化处理从而解决了第一个问题,基于分块融合特征的算法解决了第二个问题。但是前面提到的这些算法实际上还是拥有一个共同的弊端,在跟踪目标发生快速移动的时候产生的边界效应问题会造成算法跟踪失败。对这个问题随着 SRDCF^[18]算法的出现而得以解决,此算法使得滤波类的算法上了一个新的台阶,通过在时域上对滤波器进行正则化处理来使得算法达到了与深度学习类跟踪算法相媲美的跟踪效果。

深度学习:深度特征^[19,20]对于目标有极强的表征能力,尤其是在目标检测中表现出强大的能力,但是由于目标跟踪具有特殊性,只可以通过初始帧的图片提取目标的特征信息,而深度学习中的神经网络需要大量的数据来学习,所以最开始使用深度学习进行跟踪的效果并不好。后来通过研究发现用于图像分类器训练的 CNN(卷积神经网络)移植到跟踪领域会产生极好的跟踪效果,所以深度学习跟踪算法在最近几年才得到了积极的发展。

根据最近几年的目标跟踪算法赛事 VOT 的排名靠前的结果,使用卷积神经网络和相关滤波的方法占据了主流,未来在目标跟踪领域使用这两种思想的算法将会越来越多。本文选题对于动态场景中跟踪目标发生多尺度变化和遮挡问题的解决正是通过对相关滤波类算法中的核相关滤波算法(KCF)进行改进实现的。

1.3 目标跟踪所面临的挑战

目标跟踪中面临的挑战很多,下面介绍一下在实际应用中比较常见的几种挑战:

遮挡: 目标跟踪中的遮挡情况按照目标被挡住的时间长短来区分可以区分成两种情况:跟踪目标被障碍物长时间或者短时间的挡住。按照目标被障碍物挡住的部分不同也可以区分成两种情况:目标完全或者只有一部分被障碍物挡住^[21,22]。对于要在目标被障碍物长时间完全挡住情况下进行跟踪,至今还没有有效的解决办法。但是对于跟踪目标部分或短时间被障碍物挡住的情况下进行跟踪,有两种有效的解决办法:一是把跟踪目标的图像分成多块提取特征,即使目标被部分遮挡,也能够有效地检测到目标的部分特征;二是通过检测机制检测目标是否被遮挡,如果发现遮挡,不再更新模型。采用另外的方法对目标进行跟踪,本文算法正是基于第二种解决思路来解决目标被障碍物遮挡的问题。

形变: 形变指在跟踪视频序列中,目标表观特征和轮廓不断地在发生着变化,通常会造成跟踪目标的矩形框发生漂移。在这种情况下,解决这种问题的一种思路就是提取目标的表观特征时,采用一种对目标形变不敏感的特征,在本文的算法中使用的就是能够对形变不敏感的 Lab 颜色特征。

背景杂斑: 也称为相似背景干扰,指的是跟踪目标附近有可能存在着与目标具有相似颜色的干扰。对于这个问题解决也有两种思路,一是根据目标的运动信息,预测目标会出现的位置。二是通过让采集目标周围的样本对分类器进行训练,从而加强分类器对目标和背景的辨识能力。

尺度变化: 在跟踪视频序列中,随着目标的运动,目标的大小有可能会发生变化,如何产生准确的适合目标大小的跟踪框是解决尺度变化问题的关键,通常的做法就是在跟踪模型更新时产生大小不一的样本候选框来预测跟踪目标的跟踪结果,通过一些指标(例如峰值)来选取出最优的尺度(跟踪框)作为对于目标位置和大小描述的最佳跟踪结果。

目标跟踪面临的挑战还有很多,例如光照变化、快速运动、超出视野、低分辨率、运动模糊等等。所有的这些因素加起来意味着目标跟踪是一项具有挑战性的任务。

1.4 本文主要研究工作

本课题研究的主要内容为在动态背景下实现对运动目标的跟踪,通过提取目标的表观信息特征中的方向梯度直方图特征和 Lab 颜色空间直方图特征,然后把两种特征进行融合,主要内容涉及目标纹理特征提取,目标跟踪算法实现,利用跟踪结果训练分类器,确定目标位置几个核心模块,设计中要利用跟踪算法得到跟踪结果,利用典型的机器学习方法训练分类器模型,利用分类结果纠正跟踪结果,确定最终目标位置,给出完整的目标移动的状态信息,完成实验分析。

主要完成以下工作:

(1) 分析问题, 整理思路, 确定实现方法

分析算法实现所需要的解决的问题和算法要实现的功能,明确算法的实现思路,

画出程序流程图，确定实现算法的程序语言和开发环境。

(2) 思考算法实现具体细节，搭建所需要的实验环境

学习算法实现中每个功能所需知识和实现技术，确定数据在算法中的处理方式和走向。在个人计算机上搭建实验环境。

(3) 编码和测试

把算法的程序流程图通过代码进行实现，测试已经实现的算法功能是否达到之前确定的要求。

(4) 对比和改进

使用算法性能对比工具，把实现算法和其他类似算法进行综合的性能对比，针对自己算法的不足进行改进。

1.5 本文的内容结构安排

本文主要介绍的内容如下：

绪论主要介绍了本文选题所涉及到目标跟踪目前的发展趋势和背景，罗列并简要地介绍了目前一些已经出现的目标跟踪算法，从中着重介绍了与本文使用算法类型相同的相关滤波算法。然后对于目标跟踪所面临的挑战做了一定的总结。最后对于本文所需要做的研究工作做了简要的说明。

第二章主要介绍了本文所用的核相关滤波算法涉及到的一些数学基础理论知识，包括岭回归，循环移位和循环矩阵，傅里叶对角化。之后对于核相关滤波类算法的特点和涉及到使用的两个数学模型做了简要的介绍，最后详细地介绍了核相关滤波算法的实现流程。

第三章本文针对核相关滤波算法的缺陷，提出了一种基于核相关滤波器的多尺度融合特征的抗遮挡改进算法，首先大体对改进算法的总体框架和运行的详细流程做了介绍。之后对于算法中涉及到的一些关键内容，如 HOG 特征，Lab 颜色特征，Kalman 滤波器等内容做了详细地介绍。通过这些内容来说明本文算法是如何实现跟踪目标的多尺度和抗遮挡功能的。

第四章主要将本文实现的算法与其他滤波类算法进行目标跟踪实验的结果进行对比，首先对于实验环境进行了详细地介绍，然后分别对本文算法和其他滤波类算法的跟踪效果和性能进行定性和定量分析。

第五章对于本文算法已完成的研究工作进行总结，并且对本文算法仍然存在的不足进行了分析，并且提出了改进方法。

1.6 本章小结

本章首先简要介绍了选题所涉及的目标跟踪领域的发展背景，分别阐述了目标跟踪的理论研究和实际应用的意义，指出了本课题研究的重要价值。然后对于目前在目标跟踪领域所常见的主要三类跟踪算法和常见的跟踪挑战分别做了分析和介绍。最后对研究课题要进行的主要工作和论文的结构安排都做了详细地说明，为下面对课题研究内容的详细介绍打下基础。

2 核相关滤波器理论

2.1 核相关滤波器理论基础

2.1.1 岭回归

对于一个数据集 X ，拥有 m 个观测训练数据集，对于新输入的 n 维 X ，回归就是让其预测一个或者多个的目标输出。线性回归的目的是为了构造如(2.1)公式所示的回归模型：

$$f(\mathbf{z}) = \mathbf{w}^T \mathbf{z} \quad (2.1)$$

其中， \mathbf{z} 是输入的数据， \mathbf{w} 是回归模型参数矩阵， $f(\mathbf{z})$ 是回归结果，即响应值。在通过最小二乘法得到的回归再加上列向量权重矩阵 \mathbf{w} 的 F 范数作为一个新的数据项，如公式(2.2)所示：

$$\min_{\mathbf{w}} \sum (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2 \quad (2.2)$$

在公式(2.2)中，第二项的系数是 λ ($\lambda > 0$)，系数增大第二项的作用也会越来越明显，反之则不然。当 λ 的值取 0 时，岭回归公式通过化简变成最小二乘回归公式。对于公式(2.2)有关于 \mathbf{w} 的闭式解，对矩阵 \mathbf{w} 求导，令导数为 0，可以求得 \mathbf{w} 的解，如公式(2.3)所示：

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.3)$$

将公式(2.3)扩展到傅里叶域，得到公式(2.4)，为最终得到的岭回归公式。

$$\mathbf{w} = (\mathbf{X}^H \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^H \mathbf{y} \quad (2.4)$$

2.1.2 循环移位和循环矩阵

对于一个 $n \times 1$ 的向量 \mathbf{x} ，用循环矩阵 P (公式 2.7)把 \mathbf{x} (公式 2.5)进行循环移位，循环移位的计算公式如公式(2.6)所示：

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T \quad (2.5)$$

$$P\mathbf{x} = [x_n, x_1, x_2, \dots, x_{n-1}]^T \quad (2.6)$$

$$P = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \quad (2.7)$$

按照行将 \mathbf{x} 各个循环移位得到的结果按照行拼接得到的矩阵叫做 \mathbf{x} 构成的循环矩阵。如公式(2.8)所示：

$$X = C(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \quad (2.8)$$

2.1.3 傅里叶对角化

对于任意一个循环矩阵来说,在傅里叶空间中都可以进行对角化计算,根据公式(2.9)所示:

$$X = F \text{diag}(\mathbf{x}) F^H \quad (2.9)$$

其中向量 \mathbf{x} 构成了式子中的 X , 为一个循环矩阵, 向量 \mathbf{x} 经过傅里叶变换, 组成了等式右边的对角阵的主对角元素。傅里叶对角化把循环矩阵的计算变成了向量运算。

用四阶矩阵作为例子, F 是傅里叶变换矩阵, 形式如公式(2.10)所示:

$$F = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega^9 \end{bmatrix} \quad (2.10)$$

$$\omega = e^{-2\pi i/n} \quad (2.11)$$

在公式(2.11)中, n 代表矩阵的阶数, 上面提到的矩阵中 n 就为 4, 然后, 通过使用 DFT 矩阵, 可以计算出信号, 也就是矩阵和向量的傅里叶变换, 如公式(2.12)所示:

$$\mathcal{F}(\mathbf{z}) = \sqrt{n} F \mathbf{z} \quad (2.12)$$

2.2 核相关滤波器的算法

之前提到过目标跟踪算法有两种: 生成式算法和判别式算法。而核相关滤波器算法属于后者, 与判别式算法的原理类似, 在其实现原理中的分类器被替换成了滤波器来对跟踪目标的位置进行判断。在对分类器进行训练的时候使用的仍然是 Positive-Negative Sample 方法, 即正负样本训练方法。目标区域一般被看作是正样本, 目标周围区域被看作是负样本。在图像中寻找的位置越靠近目标, 其周围的区域越有可能是正样本。

2.2.1 核相关滤波器算法的特点

(1) 把循环矩阵在跟踪目标所在的周围附近区域进行循环移位, 采集正负样本, 采用岭回归方法去训练滤波器。由于循环矩阵能通过傅里叶变换矩阵进行对角化运算, 所以得到循环矩阵的方法得到了简化: 通过向量的哈达玛积(向量中元素的点乘计算), 与此同时算法的运行速度和实时性也通过这种运算的简化

而得到了提升。

(2) 使用核函数能够把在向量空间中的岭回归映射到非向量空间, 如果要在非向量空间内进行对偶问题或者是常规约束的求解, 也可以通过利用循环矩阵能够通过傅里叶变换矩阵进行对角化运算的性质来简化运算过程。

(3) 相比起单通道特征(如灰度特征), 核相关滤波器可以采用例如 HOG 这种的拥有多个梯度和直方图通道的特征, 也可采用像 RGB 和 HSV 这种拥有三个通道的颜色特征, 它为本节算法集成多通道特征提供了一种方法。

2.2.2 核相关滤波器算法所用到的两个数学模型

(1) 线性回归

$$\mathbf{w} = \frac{\mathbf{x}^* \odot \mathbf{y}}{\mathbf{x} \odot \mathbf{x} + \lambda} \quad (2.13)$$

如公式(2.2)和公式(2.13)所示, x_i 代表样本, x_i 样本总体就组成一个样本总体用 \mathbf{x} 表示, y_i 表示期望相应, y_i 组成的总体用 \mathbf{y} 表示, 这里的 λ 表示正则系数, 防止过拟合。(2.13) 公式的作用在于将样本 \mathbf{x} 利用快速傅氏变换到傅里叶域, 并且将 \mathbf{y} 进行傅里叶变换得到回归器的傅里叶域形式。这是核相关滤波器算法中使用的第一个回归器。

(2) 非线性回归

$$\mathbf{k}^{xx'} = \exp\left(-\frac{1}{\sigma^2}\left(\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2\mathcal{F}^{-1}\left(\mathbf{x}^* \odot \mathbf{x}'\right)\right)\right) \quad (2.14)$$

$$\mathbf{k}^{xx'} = \exp\left(-\frac{1}{\sigma^2}\left(\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2\mathcal{F}^{-1}\left(\sum_c \mathbf{x}_c^* \odot \mathbf{x}_c'\right)\right)\right) \quad (2.15)$$

公式(2.14)表示通过高斯核函数(RBF)可以对两个单通道的向量求相关内积, 公式(2.15)同样是利用高斯核函数公式对输入的参数进行核相关运算, 这里输入的参数可以是拥有多个通道的特征图或者是彩色图。

2.2.3 核相关滤波器算法实现流程

(1) 在视频序列的第一帧中框选出所要跟踪的目标的位置, 选择的矩形框区域大小将会被扩大 2.5 倍, 区域的大小被记为 $\mathbf{m} \times \mathbf{n}$ 。

(2) 将框定的目标位置的矩形框进行余弦加权操作, 然后计算 HOG 特征算子, 得到 N 维度的 HOG 特征图。组成特征的每个维度被看作样本以大小为 $\mathbf{m} \times \mathbf{n}$ 作为输入, 样本被记作 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N$ 。

(3) 使用二维高斯函数, 生成标签矩阵 \mathbf{y} , 其大小和样本 $\mathbf{m} \times \mathbf{n}$ 一致。

(4) 使用上文中提到的非线性回归模型(2.14)或者是(2.15), 计算出 \mathbf{k}^{xx} , HOG 特征作为输入, 训练形成一个非线性回归器如(2.16)所示:

$$\mathbf{a} = \frac{\mathbf{y}}{\mathbf{k}^{xx} + \lambda} \quad (2.16)$$

(5) 在视频序列播放的下一帧中, 在之前帧目标出现的位置中提取 $m \times n$ 大小的区域作为样本进行余弦加权, 然后得到这个区域的 HOG 特征图并且把每个维度记作 $z_1, z_2, z_3, z_4, \dots, z_N$ 作为样本输入, 利用公式(2.17)求出在核空间中的测试样本和训练样本的核矩阵元素 K_{ij}^z , 然后元素共同组成核矩阵 K^z (2.18), 再求出 K^{xz} , 其为矩阵 K^z 转置矩阵的第一行。

$$K_{ij}^z = \phi(z_i)^T \phi(x_j) \quad (2.17)$$

$$K^z = \phi(X)\phi(Z)^T \quad (2.18)$$

(6) 使用以下公式(2.19), 求得在傅里叶域下面的响应矩阵, 之后通过傅里叶逆变换得到响应矩阵 $\hat{\mathbf{f}}(\mathbf{z})$, 如式子(2.19)所示:

$$\hat{\mathbf{f}}(\mathbf{z}) = \mathbf{k}^{xz} \odot \mathbf{a} \quad (2.19)$$

(7) 最大的响应值所处的位置在响应矩阵 $\hat{\mathbf{f}}(\mathbf{z})$ 中寻找, 如果在矩阵中一个位置的响应值大于响应矩阵中所有其它位置的响应值, 那么此位置为当前帧中目标所处的位置。但是如果情况恰好相反, 可以采取另外的措施(如重新进行全局搜索匹配)来重新获得一个目标所在的区域, 之后重新从步骤(1)中开始。

(8) 之后的步骤主要就是更新模型, 从最新找到的目标位置处提取样本, 然后紧接着重复步骤 2 到步骤 5 这四个操作。对于当前帧, 记模型为 α' , 则当前帧的模型和最开始计算出来的模型进行插值运算可以得到下一帧使用的模型, 如公式(2.20)所示:

$$\alpha_{new} = m\alpha_{old} + (1-m)\alpha' \quad (2.20)$$

这里的 m 代表着学习率, 值介于 0-1 之间, 表示的是之前的模型对当前模型的影响程度, 其值过大或者是过小都会造成跟踪失败。

(9) 一直重复从步骤(6)开始的检测过程。

2.3 本章小结

本章主要对涉及到核相关滤波器的数学基础理论, 包括岭回归, 循环矩阵, 循环移位, 傅里叶对角化和核相关滤波器要用到的两个数学模型进行了简要的介绍。之后对核相关滤波器的特点进行分析, 其在目标跟踪过程中的最大的优点就在于性能优势, 但是缺点也很明显, 不能应对跟踪过程中目标发生的多尺度变化以及遮挡问题。这两个缺点的提出为后面提出通过提取融合特征, 并且实现多尺度和抗遮挡的解决方案奠定了基础。最后本章对核相关滤波算法的实现流程进行了介绍。

3 一种基于核相关滤波器的多尺度抗遮挡目标跟踪算法

核相关滤波器算法（KCF）虽然有明显的速度优势，通过利用在傅里叶空间中循环矩阵可对角化这个性质来简化了正负样本训练中矩阵计算过程，大大降低了运算量，提高了算法的性能和实时性，但是它仍然具有一定的缺陷：

第一、在目标跟踪过程中的框定目标的跟踪框大小已经事先由程序设定，自始至终跟踪框的大小没有发生变化。但是实际上在跟踪视频序列中跟踪目标的大小会经常发生变化，因为 KCF 对循环矩阵的依赖，对于多尺度的目标跟踪效果并不是特别好，在跟踪过程中跟踪框有可能会漂移的情况，最后导致的结果就是跟踪失败。

第二、KCF 在跟踪过程中当目标被遮挡时跟踪也会失败。之前在 KCF 算法的实现流程中本文提到算法中会使用公式(2.17)对模型系数 α 进行线性插值，但是如果跟踪的目标已经被遮挡了若干帧，那么新插入的模型系数 α' 将会造成错误的线性插值，使用模型去寻找的目标位置将会是错误的，也会造成跟踪框结果发生漂移的情况，此时遮挡物已经污染了模型。

3.1 总体框架

本文所设计的算法主要分成了两个模块，跟踪器目标检测和初始化模块和跟踪器模型更新模块，算法详细的实现原理和实现过程中的中间结果如图 3.1 所示。

第一个模块主要流程是通过视频每一帧和在每一帧中标定的所要跟踪目标的跟踪框信息来提取目标图像中的融合特征信息（HOG+Lab 颜色特征）并且在框选目标的中心位置附近进行采样，通过训练得到一个模型，这个模型能够计算图像中各个位置像素的响应值。同时在视频第一帧播放时初始化卡尔曼滤波器，开始从视频的第一帧统计所要跟踪目标的运动信息，以便于跟踪目标发生遮挡时使用预测目标位置信息确定被遮挡目标的位置。

第二个模块主要流程是当视频序列即将播放下一帧图像时，判断上一帧的跟踪框的位置是否发生了不正常的情况，例如出界，如果有则对跟踪矩形框的位置进行相应的调整。同时检测原本检测目标跟踪框大小（尺度）是否发生了变化，如果没有发生变化，在前一帧目标附近进行采样，用训练出来的模型与图像进行相关操作，保存每个采样点像素的响应值，响应值最高且满足给定的阈值条件的位置则为目标此时的中心位置。然后通过特定的函数计算目标所在位置跟踪框的总峰值。然后通过原本设定的尺度步长，在这里起到倍数的作用，对矩形框和尺度值大小进行放大和缩小，得到一大一小两个尺度，分别计算两个尺度区域的总峰值，然后与之前最开始得到的那个峰值进行大小比较，选取出拥有最大峰值的尺度，最后通过尺度和之前得到的目标中心位置调整本帧目标跟踪框的位置，然后训练和更新模型，用于在下一帧图像的采样位置与图像进行相关操作。

在第二个模块中要注意的是，目标发生遮挡的时候，此时峰值最大的尺度会

小于一个阈值 λ ，这个值可以作为判定目标是否被遮挡的依据，如果目标发生遮挡，这种情况下可以使用卡尔曼滤波器接替跟踪并且不再更新模型。因为从第一帧卡尔曼滤波器初始化开始就一直在搜集有关跟踪目标运动的信息并且一直对下一帧目标将要出现的位置进行预测；使用卡尔曼滤波器能够进行短时目标位置的检测和跟踪，当目标出遮挡时，峰值恢复正常，这时可以继续更新模型，不影响正常的跟踪。

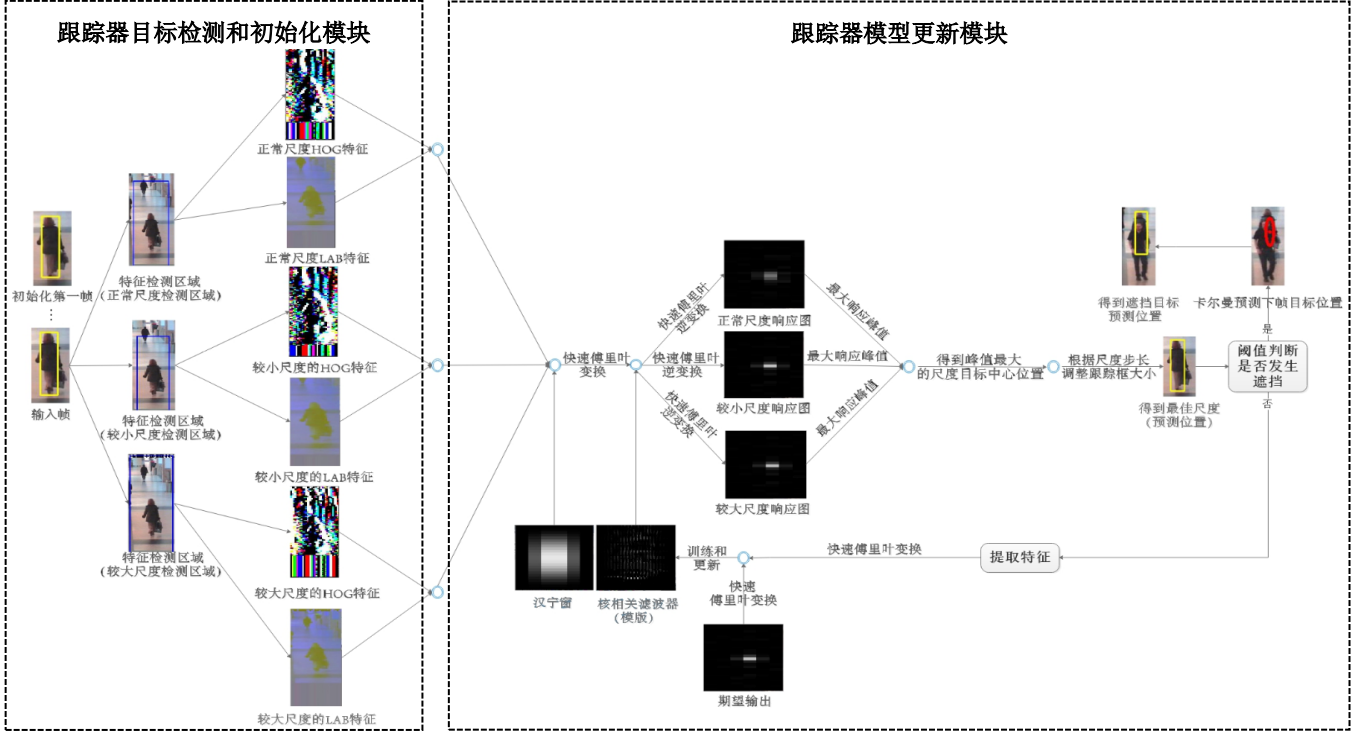


图 3.1 本章算法实现原理图

3.2 跟踪器目标检测和初始化时涉及内容

3.2.1 方向梯度直方图特征 HOG

原理：在一个图像直方图中的每个像素，假设其特征值为 x_i ，那么共同组成的特征直方图如公式(3.1)所示：

$$H(I) = [h(x_1), h(x_2), \dots, h(x_n)] \quad (3.1)$$

其中 $h(x_i)$ 是像素个数的标准化处理值，作为某种特征概率分布出现在特征直方图中，经过图像灰值化的得到的灰度图像的直方图实际上也就是一种经过统计的概率分布。对 $h(x_i)$ 进行累计，对应累计直方图得到公式(3.2)：

$$S(I) = [s(x_1), s(x_2), \dots, s(x_n)] \quad (3.2)$$

方向梯度直方图(Histogram of Oriented Gradient, HOG) 特征最适合于人体对象特征检测。由于人体是一个非刚性的物体，在跟踪过程中存在姿势变化众多，衣着服饰复杂和所处周围环境复杂等干扰因素，这些因素加在一起导致了人体检测的难度变大。但是，HOG 特征能够实现对人体对象轮廓的描述，所以在机器

图像视觉领域中的感兴趣区域检测和分析特征中拥有广泛的应用。一幅图像的 HOG 的特征表示可以通过对这幅图像的部分区域进行统计和计算,从而得到这个区域的梯度方向直方图,假设对于一幅图像 $f(x, y)$,它在点 (x, y) 处的梯度幅度和方向角 θ 使用公式(3.3)和(3.4)可以计算:

$$|\nabla f(x, y)| = \text{mag}(\nabla f(x, y)) = \sqrt{G_x^2 + G_y^2} \quad (3.3)$$

$$\theta = \arctan(G_y / G_x) \quad (3.4)$$

其中 G_x 和 G_y 分别是 x 方向和 y 方向计算出来的梯度, HOG 的特征描述是将梯度方向在区间之中划分成为 N 个均匀空间,像素点 (x, y) 在 k 个梯度方向上的梯度投影如公式(3.5)所示:

$$L_k(x, y) = \begin{cases} |\nabla f(x, y)| & \theta \in S_k \\ 0 & \theta \notin S_k \end{cases} \quad (3.5)$$

提取 HOG 特征的步骤:

HOG 特征的核心是通过计算密集的局部范围上的直方图来对跟踪目标的轮廓信息进行构造。对于图像 $I(x, y)$,本文算法提取跟踪目标的 HOG 特征的实现思路如下:

(1) 为了减少光照突变、环境变化等因素影响,需要规范化图像信息,如规范化 Gamma 空间和颜色空间。对图像进行适当压缩可以有效地应对光照、部分区域阴影等环境的变化。另外,在检测时为了提高本文算法的实时性,一般把彩色图像转化成灰色图像进行处理。

(2) 计算图像的一阶梯度。该一阶求导操作在弱化光线突变干扰的同时,还能对跟踪目标的轮廓信息,图像的纹理信息等进行捕获。与梯度对应的梯度方向计算如公式(3.6)所示:

$$\text{Ang}(x, y) = \arccos \frac{I(x+1, y) - I(x-1, y)}{r} \quad (3.6)$$

梯度计算如公式(3.7)所示:

$$R(x, y) = \sqrt{I((x+1, y) - I(x-1, y))^2 + I((x, y-1) - I(x, y+1))^2} \quad (3.7)$$

(3) 对单元格进行梯度投影。这样做可以降低对跟踪目标的外观变化的敏感度,这一步对局部图像的信息进行了提取。在获得最终的特征之前,需要先把图像分成若干个单元格,然后统计每个单元格中像素的方向梯度直方图并且进行累加,通过映射而获得最终的特征信息。

(4) 归一化块中的所有单元格。通常,图像中的单元格被共享在多个不同的块上,由于归一化操作是针对不同块进行的,因此每次运算结果都将不同。通过对单元格进行归一化操作,可以对干扰因素进行进一步压缩。

(5) 收集 HOG 特征。在待识别图像中对全部重叠的块进行 HOG 特征提取,并将他们组合成最终的特征向量,以用于识别。

本文在提取目标特征时使用的是通过主成分分析方法 (PCA) 进行过降维的 PCA-HOG 特征, 假如目标所在区域图像的分辨率是 100×100 , 那么目标图像的一共拥有 10000 个特征向量, 这对算法的处理是极大的负担, 所以通过 PCA 降维方法, 将表示目标特征的特征向量的个数尽量减少, 提升了算法的性能。

3.2.2 Lab 颜色模型

Lab 颜色模型是由 CIE (国际照明委员会) 在 1976 年公布的一种色彩模式。这种模型的特性之一就在于它同设备图像颜色显示能力无关, Lab 的出现正好弥补了之前出现的 RGB 和 CMYK 颜色模型都必须依赖设备的颜色特性这一不足; 此外, Lab 模型与 RGB 和 CMYK 模型相比具有的一个最大的特点就是它表示的色彩范围是最广的, 所以 RGB 和 CMYK 模型都可以和 Lab 模型进行相互转换。

Lab 颜色空间如图 3.2 所示, 总共拥有 L, a, b 三个通道, 其中 L 通道代表明度 (Luminance), a 通道由负到正方向颜色从绿色变为红色, b 通道由负到正方向颜色从蓝色变为黄色。

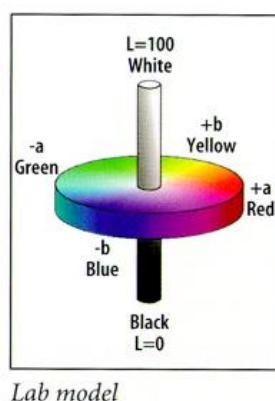


图 3.2 Lab 模型示意图

因为 Lab 模型与设备的颜色显示能力无关, 所以在彩色图像的检索中拥有了大量的应用。此外, 如果希望经过图像处理后的图像能够保留范围较广的色域和充实的色彩, 那么图像可以在使用 Lab 颜色模型的情况下进行处理, 最后再根据需求把 Lab 模型转变成 RGB 模型 (用于显示) 或者是 CMYK 模型 (用于打印) 进行输出。这种处理方式能够让图像在经过处理之后能够尽量地使用更加丰富和优质的色彩进行输出。

综上所述, Lab 颜色模型作为一种特征对于色彩的敏感性是很高的, 但是对于形变和运动模糊敏感度却比较低, 可以适用于在目标发生形变和运动情况下的跟踪。

3.2.3 汉宁窗以及融合以上两种特征的方法

(1) 使用窗函数的原因

在第一章提及到的核相关滤波算法实现过程中用到了傅里叶变换, 其经常用于时域和频域之间的转换。当对拥有无限长度的数字信号 (包括图像) 进行测量

和计算时，不可能分析全部信号，而是分析从信号中截取的有限内容。在这个截取信号的过程中，假设当前处理的信号是一个一维正弦波信号，如果对于这个信号的截取部分的周期不是正弦函数周期的正数倍，那么截取出来的信号就会在信号图上不连续，那么这种截取称为非周期截取。它会导致这个信号的频谱发生泄露，这种情况在频谱上主要表现为“拖尾”现象。而通过加入窗函数则能够减少这种泄露。为了更加形象地说明频谱泄露的现象和窗函数减少频谱泄露的效果，这里选用数字图像标准测试集中的 cameraman 图片的频谱图进行演示说明。

如图 3.3 所示，右边图像为左边图像的频谱图。从图像中可以看出，频谱图在纵轴方向有一条细白线，实际上这种现象由频谱泄露所造成。前面提到，当对数字信号进行有限截取分析时，如果是非周期截取将会造成频谱泄露。把这种情况推广到二维信号，当使用傅里叶变换对目标采样样本进行 x 方向周期延拓和 y 方向周期延拓时，从图 3.4 来看，车辆行驶中车顶上的天空比较亮，但是车轮行驶的马路却比较暗，对应的频谱图在 y 方向上产生了灰度突变，在频谱图上表现出来就是 y 方向上的频谱泄露(y 方向的白线)。观察纵向边缘，灰度变化不是特别大，所以 x 方向上的白线不是很明显，也表明 x 方向上的频谱泄露也不是很明显。但是总归来说，白线的出现代表着图像频谱发生了泄露。



图 3.3 Car2 图像及其它的频谱图

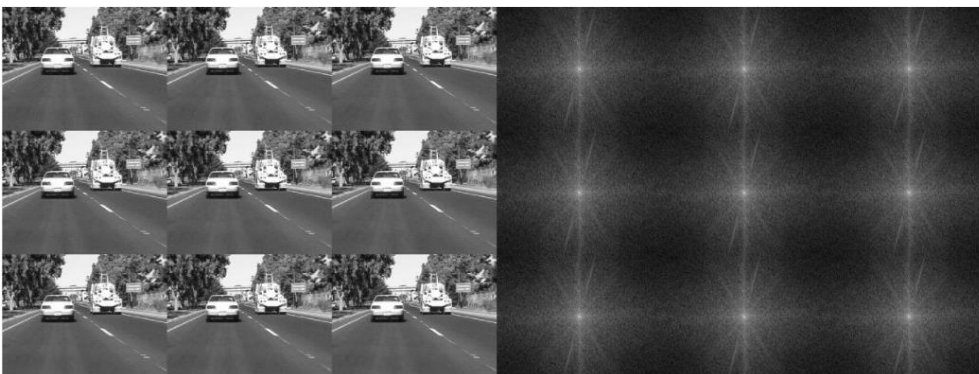


图 3.4 傅里叶变换进行横向和纵向循环赋值的结果和频谱图

为了解决这一问题，这个演示中采用了加余弦窗（窗函数的一种）的方式

解决频谱泄露的问题。图像加上余弦窗的效果如图 3.5 所示,此时频谱的变化如图 3.6 所示,图像增加余弦窗之后,y 轴方向上的频谱泄露的现象不那么明显(白线变淡)了,而且在频谱图中的能量比之前更加集中和清晰。不过在这里要强调一点,窗函数只能减少频谱泄露和不会让频谱泄露消失。



图 3.5 Car2 图像加上余弦窗的显示效果



图 3.6 Car2 图像加窗前后图像频谱变化

(2) 添加汉宁窗(余弦窗)的作用

汉宁窗是窗函数的一种,快速运动作为目标跟踪挑战之一会产生边界效应,分类器(也就是之前本文算法提到的模型)的判断力会因为边界效应产生的错误样本而降低。为了将这种影响减小,所以使用余弦窗在训练模型阶段进行采样,下面对其如何发挥作用进行解释:

训练阶段发挥的作用: 本文使用的 KCF 算法中进行样本采集时会发生循环移位,如果没有余弦窗,生成的都是合成样本(指的是如下图 3.7 所示的移位样本中的样本表示出来的是不完整的跟踪目标信息),采样生成的移位样本如图 3.7 右图所示。

假如循环移位生成 1000×1000 个图像块,那么其中只有 $1/1000000$ 的样本是正确的,这一堆图像块作为训练集不能用于训练模型。加入汉宁窗能够使得本文使用的核相关滤波算法在采样中的循环移位过程中始终保持目标的采样图像用肉眼观察是完整的,这种图片也能作为训练集中的合理的样本。



图 3.7 循环移位生成的样本示意图

在图 3.7 右边图中除了用绿框标注的完整的原始样本,其余的样本都是被“合成”的。

总结起来,余弦窗的作用就是增强 KCF 算法中训练模型时采集样本的判断力,避免因循环移位采集的样本不准确而导致训练了错误的模型的现象。

(3) 融合两种特征的方法

在目标跟踪中通过提取感兴趣区域的 HOG 特征和 Lab 颜色特征,得到描述这两种特征的特征向量,将这两种特征向量拼接起来得到一个新的特征向量矩阵,然后通过余弦窗获得目标的 hog 特征向量矩阵,与之前的那个矩阵进行乘法运算,得到的结果就作为最终描述目标特征的特征向量。

3.3 跟踪器模型更新时涉及内容

3.3.1 卡尔曼滤波跟踪器原理介绍以及在本算法中的应用

原理介绍: 卡尔曼(Kalman)滤波跟踪器包括预测和更新两个组成部分: 预测部分预测下一帧目标的位置信息,利用观测值提供给系统更新参数;系统更新参数后又可修正预测部分的预测结果。因此,整个 Kalman 滤波就是预测和更新两个部分的相互迭代,如图 3.8 所示。

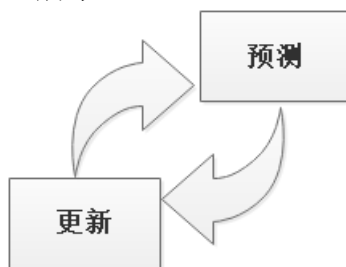


图 3.8 Kalman 滤波跟踪算法

Kalman 滤波跟踪算法是对线性和高斯系统状态序列进行最小方差估计的算法。其基本思想: 如果系统早期创建出可靠的状态观测估计值,那么可以根据先前建立的状态观测来预计当前位置的最优后验概率。Kalman 滤波跟踪算法依赖的系统基于以下的三个假设:

- (1) 线性的系统状态空间。
- (2) 对于目标跟踪的影响噪声服从正态分布。

(3) 噪声干扰信号在功率谱上趋近常数, 即为白噪声。

设系统的状态方程和观测方程分别为式(3.8)和(3.9)

$$X_k = AX_{k-1} + BU_k + W_k \quad (3.8)$$

$$Z_k = HX_k + V_k \quad (3.9)$$

在公式(3.8)和(3.9)中, A 为状态传递矩阵; B 为输入控制矩阵; X_k 为系统在 k 时刻的状态向量矩阵; U_k 为系统在 k 时刻的外部控制向量矩阵; W_k 为影像系统状态的高斯噪声, 其协方差为 Q ; Z_k 为系统在 k 时刻的观测向量矩阵; H 为系统的观测矩阵; V_k 为系统在 k 时刻的测量误差。

Kalman 滤波跟踪通过上一时刻估计结果的状态值和误差, 对下一时刻状态值和误差进行预测。由于通过预测得到的状态值和误差与真实状态相比都有一定的误差, 因此通过新的观测值来更新和修正预测的结果, 并为下一帧提供信息, 利用递归的方法, 就可以不停地预测和修正下一时刻目标的运动状态。

优点和不足: **Kalman 滤波**由于具有准确性高和实时性强的特点, 被应用于解决多目标跟踪的问题。但是它只适用于线性高斯的系统状态空间, 因此当目标出现无规则运动, 交叉遮挡等情况时, **Kalman 滤波跟踪**就会出现发散的现象, 从而导致跟踪失败。因此, 该系统只适用于线性和噪声高斯分布环境下的目标跟踪。

Kalman 滤波跟踪在本文算法中使用的原因: **Kalman 算法**是在当跟踪目标发生短时和部分遮挡的时候用来弥补本文使用的核相关滤波算法的不足。在核相关滤波算法中循环矩阵能通过傅里叶变换矩阵进行对角化运算的性质使得循环矩阵的计算方法得到了简化, 从而让算法的性能和实时性得到了提升。当目标没有被遮挡的时候, 本文算法能够结合 HOG 特征和 Lab 颜色特征融合得到的特征来进行快速并且准确地定位目标, 当目标未被遮挡和出遮挡时对应目标的模型(分类器)也在不断地更新, 但是当目标发生遮挡的时候, 目标基于特征构建的模型发生了污染(模型更新时的线性插值错误), 算法的目标跟踪框很有可能会漂移到障碍物的上面, 宣告跟踪失败。但是卡尔曼滤波器并不是一种基于特征跟踪的跟踪器, 根据公式(3.6)和(3.7)所示, 它会通过充分地利用目标当前的速度和位置信息来预测下一帧目标的位置, 对障碍物遮挡具有很强的抗干扰性。

Kalman 滤波跟踪在本文算法中使用过程: 当本文核心算法 KCF 的模型(分类器)通过第一帧初始化时, 同时也利用第一帧标定的目标位置对 **Kalman 滤波跟踪器**进行初始化, 设置 **Kalman 滤波器**中相应的参数信息, 之后当视频序列每一帧播放时, 都记录下当前跟踪目标的位置和速度信息, 用来对下一帧目标的位置进行预测。当目标发生遮挡时, 也就是说之前提到的用于预测适合目标大小的最佳尺度返回的峰值小于一个指定的阈值时, 停止训练(更新)模型(分类器), 基于之前使用 KCF 算法获得的目标状态信息, 在被判定遮挡的帧中使用 **Kalman 滤波器**对目标位置信息进行预测, 然后重复这个过程, 直到目标出遮挡时, 峰值恢复正常(大于阈值), 之后继续更新模型, 利用模型预测结果对目标跟踪。然而这种处理遮挡的方法也仍有不足, 正如之前提到 **Kalman 算法**的优点和不足中

所说,这种算法只使用于线性下的跟踪问题。也就是说,对于目标下一帧运动状态的预测是通过当前帧的运动状态使用线性方程得到的。

3.3.2 多尺度介绍以及在本算法中的应用

尺度空间理论: 观察日常生活中的任何物体的形态,都有一定的衡量标准,这个衡量标准被称为尺度。对于物体尺度的衡量可以从多方面去理解,衡量物体的重量可以从“公斤”,“吨”,“千克”去描述物体的不同尺度,对于物体大小可以从“米”,“千米”,“毫米”,甚至是“纳米”去描述。对于尺度空间描述比较形象的例子例如 Google Earth,查看地图的时候,随着鼠标滚轮的滑动,地图的比例尺(尺度)也在不断地发生变化,人眼看到的地图上的细节也会发生不断的变化;还有的例子就是相机,通过调焦这种操作相机也可以把让拍摄的景物发生大小变化,由此模拟了景物通过人眼捕捉经过处理在视网膜上进行成像的过程。

通过对未知场景采用机器视觉技术进行分析时,计算机事先并不知道所图像中所要跟踪目标的合适尺度。所以需要考虑使用合适的多尺度描述图像的方法(即在所有尺度下图像的描述),获得感兴趣目标的最佳尺度。获取最佳尺度的同时也要保持尺度的不变性,其指的是在不同的尺度下,跟踪的目标都具有相同的特征关键点,那么对于不同尺度的图像输入就可以检测出相同关键点来进行匹配。

多尺度的两种表达模式: 通常指的是尺度空间与多分辨率金字塔两种方式。

(1) **尺度空间的表达——高斯模糊:** 高斯核函数(RBF)是唯一能够生成多尺度空间的核函数,假设一个图像的尺度空间为 $L(x, y, \sigma)$,它可以被看作由一个原始图像 $I(x, y)$ 和尺度可变的二维高斯函数进行卷积运算来得到,这里再补充一点, x 和 y 代表图像中任意一个像素点的坐标。

尺度空间公式,如公式(3.10)所示:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.10)$$

在公式(3.10)中, $G(x, y, \sigma)$ 代表二维空间高斯(卷积)函数,也称为高斯模版矩阵,它的作用就是对原图像 $I(x, y)$ 进行高斯平滑(高斯模糊),根据高斯模版矩阵的大小,通过合理的设置 σ 来达到调整图像的模糊程度的效果,原始图像 $I(x, y)$ 像素值矩阵通过与高斯模版矩阵相乘,得到一个新的图像 $L(x, y, \sigma)$,这时候图像的模糊程度发生改变,相当于得到了一个新的尺度。

二维空间高斯函数,如公式(3.11)所示:

$$G(x_{is}, y_i, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma^2}\right) \quad (3.11)$$

公式(3.11)实际上是概率论中常见的二维正态分布密度函数, x_i 和 y_i 代表第 i 个尺度高斯模版的中心点, σ^2 代表方差, x 和 y 代表图像中的任意一点, $G(x_{is}, y_i, \sigma)$ 代表编号为 i 的尺度图像中坐标为 (x, y) 的高斯模糊值,通过计算一个尺度中所有像素点的高斯模糊值组成的矩阵称为高斯模版。这两个公式分别代表着图形进行高斯模糊中的两个不断循环重复的过程:高斯平滑和生成高斯模版。

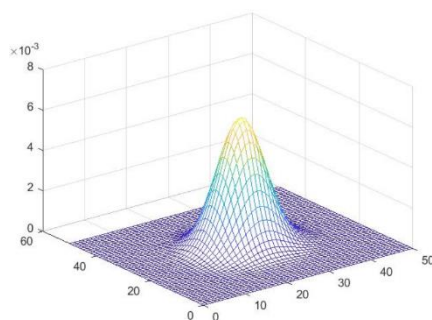


图 3.9 二维空间高斯函数图像

如图 3.9 所示，二维空间高斯函数（二维高斯分布）是从中心开始等高线呈现正态分布的同心圆。

高斯模糊的详细过程：对于原始图像中的任意一个像素点，设定模糊半径，通过加权平均计算得到一个权重矩阵，然后使用原始像素中每个像素的灰度值乘以它对应的权值得到一个新的矩阵，图像每个像素高斯模糊值通过它周围的像素在新矩阵中对应的值相加得到，通过这一系列的操作后，图像中任意像素值和它的邻域内的其他像素的像素值变化程度变小，重复之前提到的这个过程原始图像就可以得到高斯模糊的效果。在图像中，无论是对于整个图像或者是图像的一小部分（例如跟踪目标选取的目标框）来说，它们的边缘始终都是像素值变化程度最大的地方，而其他的区域则相对来说比较平缓。因此，图像高斯模糊最明显地效果就是图像中的物体的边缘轮廓变得不清晰。效果如图 3.10 所示。



图 3.10 对图像进行高斯模糊（图像平滑）的效果图

尺度空间的表达——高斯金字塔：高斯金字塔是对图像不同尺度的模仿，使用高斯核来构建图像的尺度，它的实现方式是对拥有同一尺寸的图像进行不同程度的高斯模糊（平滑），然后进行下采样（指的是对当前图像序列的宽高比例进行缩小，然后再重新得到一个新的序列称为原序列的下采样序列）再平滑，重复这个过程最后得到一个图像的集合，这个集合就构成了图像的高斯金字塔。

(2)多分辨率金字塔表达：它是图像处理中一种早期出现的多尺度表示形式。实现一般通过两个步骤：通过低通滤波器对图像进行平滑，信号处理结果如图 3.11 所示；对平滑过的图像进行降采样，信号处理结果如图 3.12 所示，采样图像长宽大小与原图的比例可以自定。最后得到一个拥有各种缩小尺寸的图像集合。这个集合构成了多分辨率金字塔。

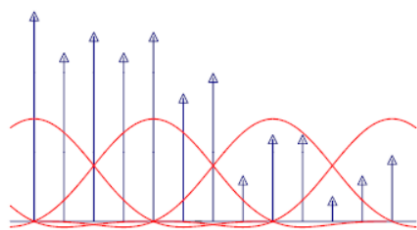


图 3.11 对信号进行低通滤波处理后结果

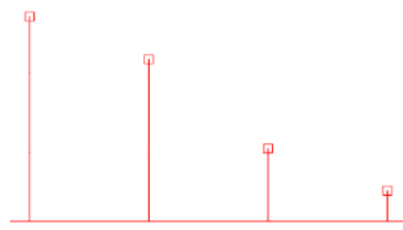


图 3.12 通过降采样处理后的信号

两种表达模式的优缺点:

- 尺度空间的表达是通过不同的高斯核进行平滑卷积来得到的，在所有的尺度上的图像都具有相同的分辨率。
- 多分辨率金字塔在每层的图像都减少固定比率

所以，多分辨率金字塔的优点在于金字塔生成速度很快并且节省存储空间，缺点是无法去分析图像的局部特征；随着尺度参数信息增加，多尺度空间表达的优点也更加明显：图像的局部特征在任何尺度上都可以用简单的信息去描述，缺点是与此同时冗余信息也会增多。

多尺度在本文算法中的应用：本文中实现多尺度是基于第二种方法，在跟踪器初始化时，设定了提取金字塔图像的固定比率（步长）并且初始化高斯核，在跟踪器模型（分类器）发生更新时，在尺度不变的时候检测高斯核函数峰值，然后利用固定比率扩大和缩小目标图像的比例，得到两个新的图像层，同样也可以得到这两个图像对应的核函数峰值，找出最大峰值对应的图像即为最适合目标的尺度。在这个过程中，要进行两个不同的采样操作：升采样和降采样，利用固定比率对原有图像进行放大操作称作升采样，对原有图像进行缩小的操作称作降采样。

3.4 本章小结

本章一开始从算法的总体框架出发，对算法中两个功能模块（跟踪器模初始化模块和目标检测+跟踪器模型更新模块）进行了简要的介绍，然后对本章介绍算法的运行流程包括最开始进行初始化，多尺度融合特征检测，分类器模型更新，遮挡判断和处理等一系列过程进行了详细地阐述，然后对算法中一些使用到的关键知识，例如余弦窗，HOG 特征，Lab 颜色特征，Kalman 滤波跟踪器等内容进行了详细地介绍和剖析，对于它们在算法中使用原因和使用方法进行了详细地解释。

4 实验结果及分析

4.1 实验所用数据集以及开发环境设置

在实验中，使用的是 OTB-100 数据集，是一种专门用于目标跟踪算法的图片序列集，整个数据集中有 100 个文件夹，其中每个文件夹中包含一个视频帧序列集合和一个作为跟踪目标实际位置参考的文本文件 `groundtruth_rect.txt`，文本文件中记录了视频帧序列播放时目标实际跟踪框所在的位置的集合。

本文算法的实验从数据集中抽取了 46 个视频帧序列对本文算法进行测试，其中这些视频帧序列中包含了目标跟踪中出现的各种挑战如尺度变化，目标遮挡，背景干扰等情况。本文算法实现的程序语言是 C++，在集成了 `Opencv3.4.1` 和 `Opencv-contrib` 开发库的 `Visual Studio 2017` 软件上进行编程实现，刚开始算法需要从视频序列目标位置参考文档中读取目标在第一帧中确定的实际跟踪框位置，之后依靠算法自己确定视频帧中的目标的跟踪框位置。遮挡判断的阈值初始设定为 0.25，实验平台是 Intel i5-6200U 2.40GHz 处理器下的 windows10 操作系统。拥有 8GB 的内存和 2GB 的独立显存。通过静态 Release 方式编译生成的 C++ 程序在执行时图像的帧率最高可以达到 94fps。

测试通过使用以上的 46 个视频帧序列来对本文算法以及与本文算法类型相同的跟踪类算法进行测试，把算法产生的跟踪效果和性能曲线进行定性和定量分析，使用 `Visual Tracker Benchmark` 基准测试工具^[23]，其通过 `matlab` 编程语言实现。为了将本文算法集成到基准测试工具中，通过采用 `Visual Studio` 工具下的 Release 静态编译方式对本文算法进行可执行文件生成，然后复制到指定的文件夹。使用 `matlab` 语言编写调用接口文件，采用 `dos` 命令的方式对可执行文件进行调用，其他算法实现程序文件也同时加入到指定的文件夹中。在配置文件中对要跟踪的视频序列和跟踪算法进行相应的设置。之后执行跟踪矩形框画图和性能曲线画图函数，对不同算法的矩形框跟踪效果和不同算法的跟踪性能曲线进行描绘。

4.2 与其他跟踪算法的定性比较

由于用于测试的视频序列太多，总共有 46 个序列，所以选取了其中的 10 个最具有挑战性的跟踪序列，它们是 `singer1`, `singer2`, `blurFace`, `freeman4`, `carDark`, `david2`, `redteam`, `walking2`, `girl2`, `dudek`，其中分别包含了形变，光照变化，模糊，快速移动，背景杂乱，平面外旋转等跟踪挑战情况，对于选择比较的跟踪算法在这 10 个序列上的跟踪效果进行分析和比较，同时选用了六种滤波类跟踪算法与本文算法 `MultipleKCF` 进行对比，它们是 `CXT`^[24], `KCF`^[16], `Struck`^[25], `CSK`^[15], `DSST`^[17]和 `STC`^[26]算法。

与此同时为了确保公平，采用帧序列文件夹中 `groundtruth_rect.txt` 中第一帧

目标的矩形框位置作为目标的初始位置进行跟踪。

表 4.1 不同视频序列中存在的挑战情况

挑战属性 视频序列	形变	光照变化	运动模糊	快速移动	背景杂乱	平面外旋转	平面内旋转	尺度变化	遮挡	超出视野
freeman4	×	×	×	×	×	√	√	√	√	×
singer1	×	√	×	×	×	√	×	√	√	×
singer2	√	√	×	×	√	√	√	×	×	×
blurFace	×	×	√	√	×	×	√	×	×	×
carDark	×	√	×	×	√	×	×	×	×	×
david2	×	×	×	×	×	√	√	×	×	×
redteam	×	×	×	×	×	√	√	√	√	×
walking2	×	×	×	×	×	×	×	√	√	×
girl2	√	×	×	×	×	√	×	√	√	×
dudek	√	×	×	√	√	√	√	√	√	√

如表格 4.1 所示，呈现出来的是 10 个视频序列中存在的不同的跟踪挑战，如果有对应的挑战属性，标记为√，没有则标记为×，下面分别展示和分析本文算法和其他算法利用 Visual Tracker Benchmark 工具在这 10 个视频序列中绘制出的跟踪效果。

(1) freeman4 序列

freeman4 视频序列描述的是一个人在教室中穿越学生挥舞东西的场景，跟踪目标在移动的过程中身体发生了平面内外旋转的情况，并且目标同时也被挥舞着的报纸遮挡。各种算法跟踪效果如图 4.1 所示，由图可见在整个跟踪过程中只有本文算法 MultipleKCF 在跟踪目标被遮挡或者是部分走出视野的情况下能够对目标进行鲁棒并且精准的跟踪。而其他算法受到遮挡的干扰影响对原始标定目标的跟踪均发生了不同程度的偏移，均未能准确地跟踪目标。

(2) singer1 序列

singer1 序列描述的是演唱会中拍摄歌手的情景，其中对于跟踪的挑战有两个方面，第一个方面是摄像头中跟踪目标发生的尺度变化。第二个方面是舞台灯光的变化对于目标跟踪具有一定的影响。根据跟踪效果图 4.2 可以观察到在整个跟踪过程中只有 DSST 算法和本文算法 MultipleKCF 实现了对目标尺度的变换，能够准确地对视频中跟踪目标的大小进行准确地框定，但是当跟踪目标受到光照变化的影响后，只有本文算法依然能够准确并且精准地跟踪目标，所以由此可以看出本文算法对于光照变化情况下也具有一定的鲁棒性。

(3) singer2 序列

singer2 的场景和 singer1 类似，都是在演唱会中对跟踪目标拍摄的场景，但是 singer2 视频序列中跟踪的挑战与 singer1 有所不同，除了光照变化对目标跟踪带来的干扰之外，还包括目标在平面外旋转带来的干扰。根据图 4.3 的效果所示，除了 CSK 算法，Struck 算法和本文算法 MultipleKCF 能够对目标进行捕捉之外，其他算法的跟踪框均发生了漂移而导致跟踪失败，并且在能够捕捉到跟踪目标的

三个算法之中，根据跟踪效果来看只有本文算法能够准确而精准地跟踪目标。

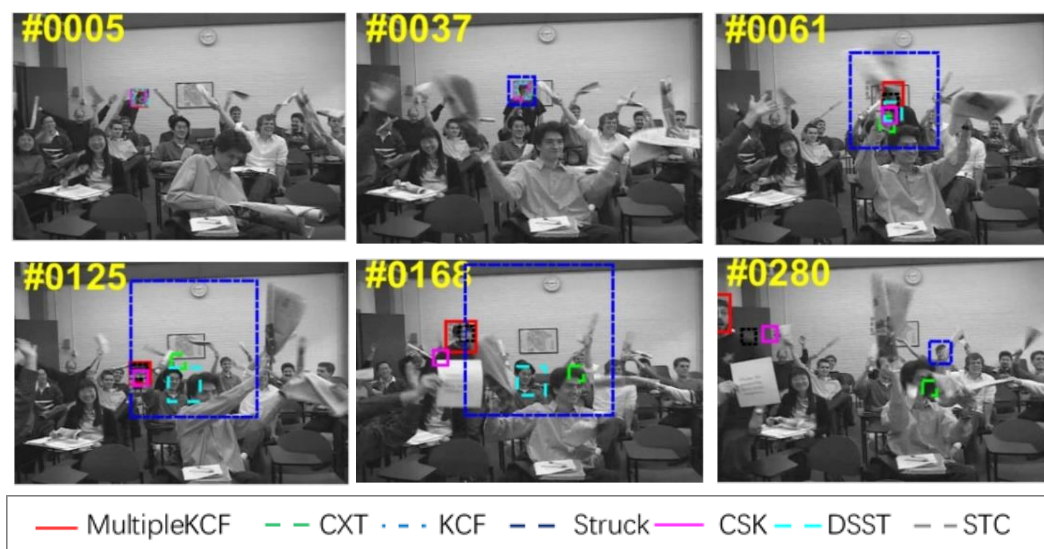


图 4.1 不同跟踪算法对于 freeman4 视频序列中目标的跟踪效果

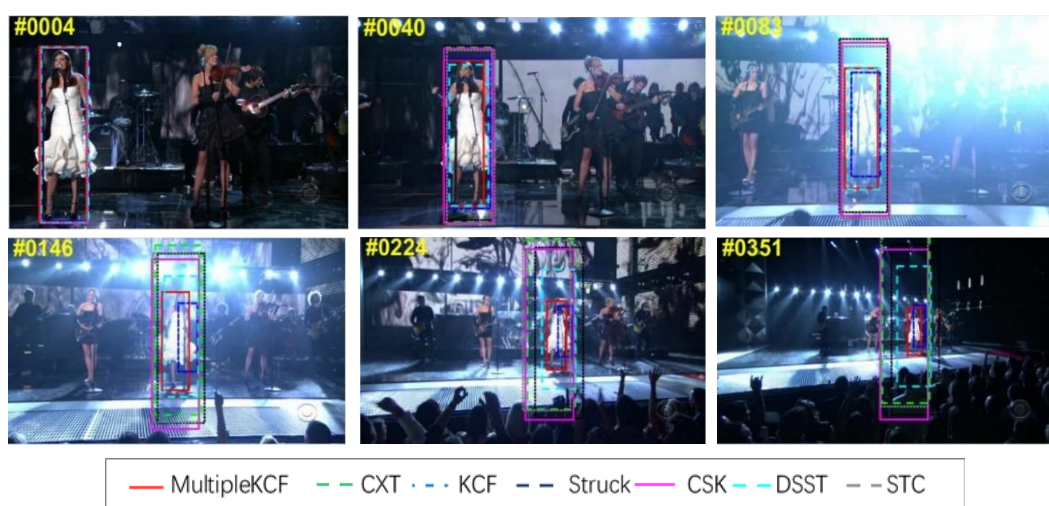


图 4.2 不同跟踪算法对于 singer1 视频序列中目标的跟踪效果

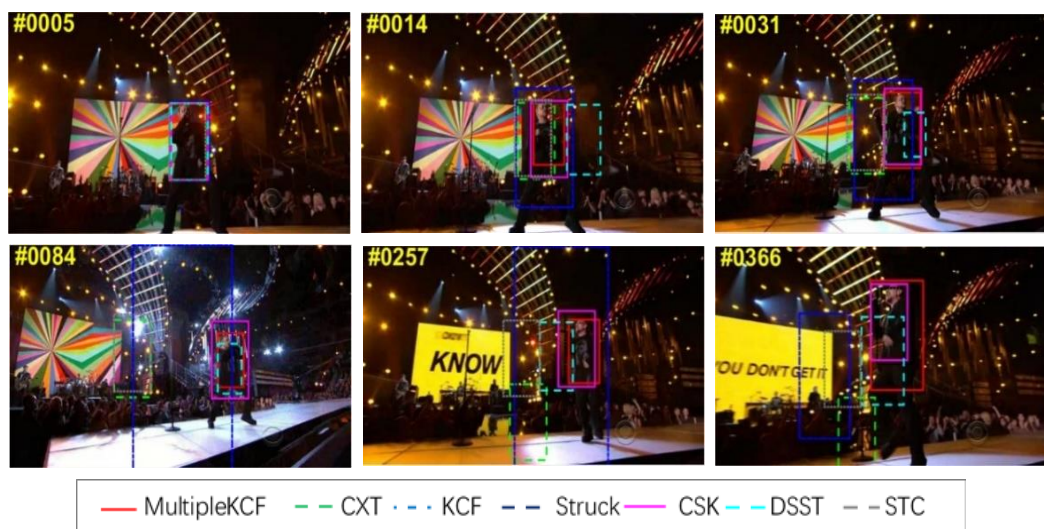


图 4.3 不同跟踪算法对于 singer2 视频序列中目标的跟踪效果

(4) blurFace 序列

blurFace 描述的场景是在摄像头快速移动造成图像模糊的情况下对视频序列中的人脸进行跟踪,其跟踪的主要挑战为快速运动造成的图像模糊,根据图 4.4 跟踪结果来看,除了 STC 算法和 DSST 算法在摄像头快速来回移动的情况下的目标跟踪框发生漂移而导致跟踪失败,本文算法和其他算法对于这种情况下的跟踪都具有一定的鲁棒性。

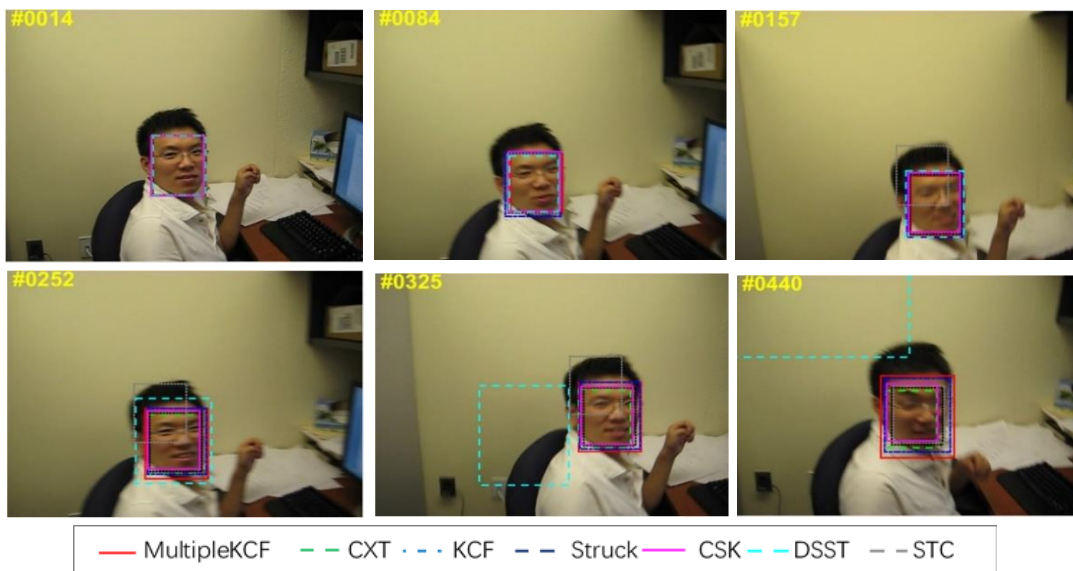


图 4.4 不同跟踪算法对于 blurFace 视频序列中目标的跟踪效果

(5) carDark 序列

carDark 序列的情景描述的是夜晚情况下对拥有杂乱背景干扰的车辆进行跟踪,跟踪的挑战就在于车辆的颜色与背景极为相似,会对车辆的跟踪造成一定的干扰。根据图 4.5 的跟踪效果来看,除了原始版本的 KCF 算法跟踪框漂移而导致跟踪失败以外,其余的跟踪算法均能捕捉到目标,并且在这些算法之中,除了 CSK 算法之外,其他算法都能够对目标的尺度进行合适的描述。

(6) david2 序列

其主要的场景为在跟踪目标为人脸的情况下,人脸发生移动的情况,跟踪挑战主要是跟踪目标发生平面内翻转,根据图 4.6 的跟踪效果,除了 DSST 算法跟踪框发生了偏移之外,其余算法对目标均具有准确并且鲁棒的跟踪效果。

(7) redteam 序列

redteam 视频序列描述的跟踪的主要场景为在摄像头固定,搭载摄像头平台移动的情况下对移动场景中的汽车进行跟踪,跟踪的挑战在于,在汽车移动的过程当中汽车图像会有大小的变化。根据图 4.7 跟踪效果来看,只有本文算法能够对小车移动过程中发生的尺度变化具有准确地描述,本文算法实现的多尺度变换能够对目标进行准确地跟踪。



图 4.5 不同跟踪算法对于 carDark 视频序列中目标的跟踪效果

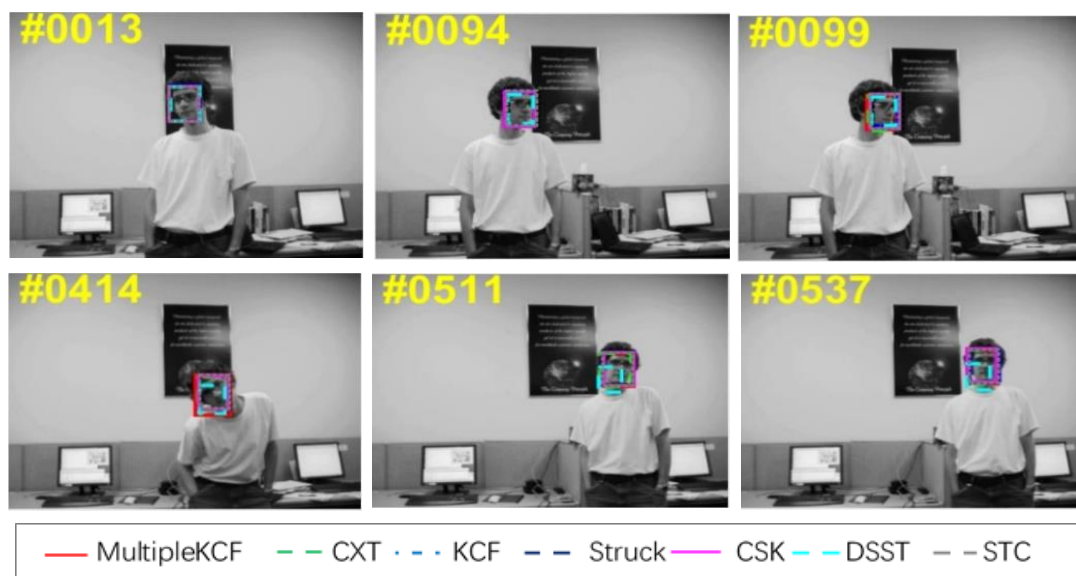


图 4.6 不同跟踪算法对于 david2 视频序列中目标的跟踪效果

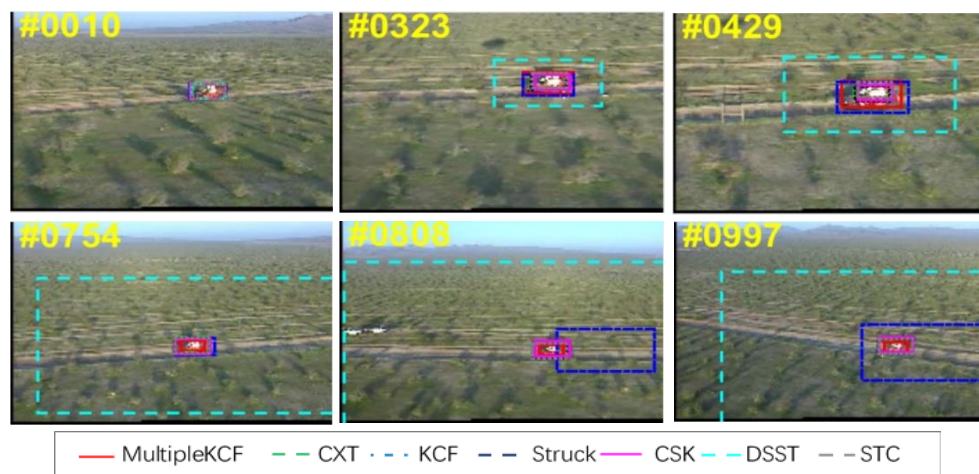


图 4.7 不同跟踪算法对于 redteam 视频序列中目标的跟踪效果

(8) walking2 序列

walking2 视频序列描述的是在视频监控的环境下,跟踪目标发生了遮挡的场景,跟踪的主要挑战为跟踪的目标受到了短时的部分遮挡并且目标发生了尺度的变换,根据图 4.8 所示,当目标发生遮挡的情况下,除了本文算法之外,其他跟踪算法的跟踪框都发生了不同程度上的漂移,只有本文实现的算法 MultipleKCF 从视频序列开始到结束一直都能够对目标的尺度进行很好的描述并且在发生遮挡之后对目标仍然能够准确地跟踪。

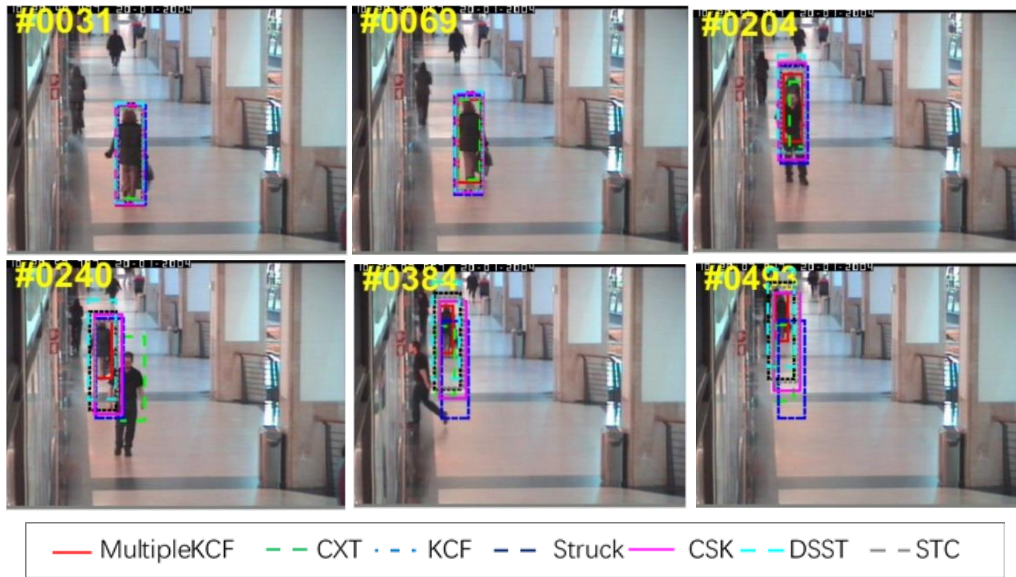


图 4.8 不同跟踪算法对于 walking2 视频序列中目标的跟踪效果

(9) girl2 序列

girl2 序列场景的情况与 walking2 视频序列情况类似,都存在跟踪的目标发生了遮挡的情况,对于此序列的跟踪结果如图 4.9 所示,可以看出,除本文算法之外的其他算法在目标发生遮挡之后跟踪框发生漂移,跟踪失败。结果证明本文算法对于有目标遮挡干扰下情况的目标跟踪具有一定的鲁棒性。

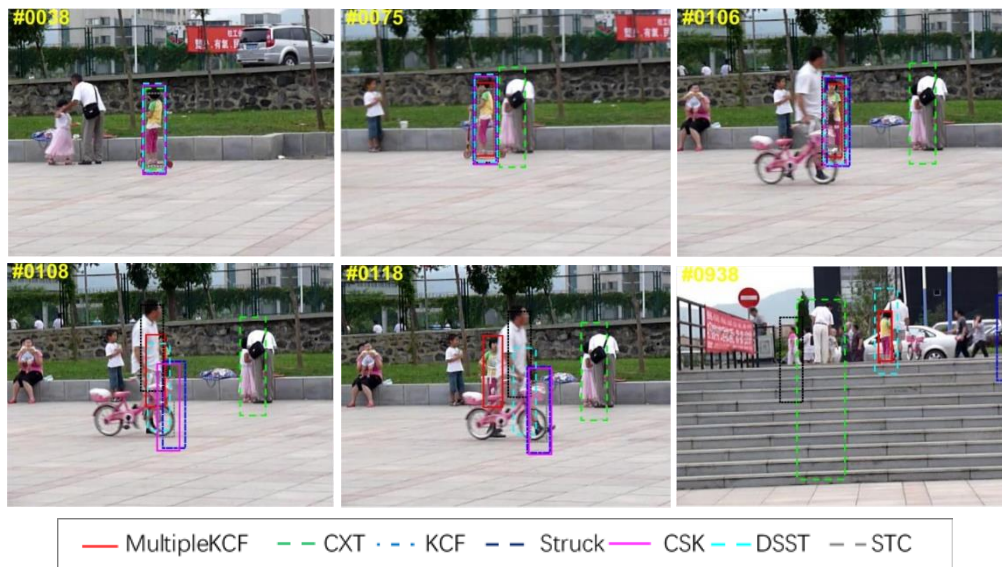


图 4.9 不同跟踪算法对于 girl2 视频序列中目标的跟踪效果

(10) dudek 序列

其场景为跟踪目标发生平面内旋转和部分出视野的情况，这也是此序列跟踪中常见的两大挑战。根据图 4.10 跟踪结果显示，虽然除了本文算法之外的算法均能够准确实现对人脸的跟踪，但是只有本文实现的算法能够对被跟踪的人脸的尺度进行准确地描述。



图 4.10 不同跟踪算法对于 dudek 视频序列中目标的跟踪效果

4.3 与其他跟踪算法的定量比较

对于不同跟踪算法的性能的比较使用的是 Visual Tracker Benchmark 基准测试工具，使用方法在前面已经有介绍，这里不再赘述。接下来本文将对算法鲁棒性评估需要的两个参数进行详细地介绍。

4.3.1 跟踪算法的精确度和成功率指标介绍

精确度图：中心位置误差是衡量跟踪算法的精准度的评估标准之一，指的是通过手动标定的目标真实位置和目标跟踪算法得到的矩形框中心位置之间的欧式距离。对于用于跟踪视频序列中的每一帧，假设手动标定的位置矩形框中心坐标为 (x,y) ，使用跟踪器确定出来目标位置矩形框的中心目标是 (z,w) ，对于这一帧的中心位置误差 ε 使用公式(4.1)计算：

$$\varepsilon = \sqrt{(x-y)^2 + (z-w)^2} \quad (4.1)$$

但是，在实际的跟踪过程中很有可能出现跟踪器跟丢目标的情况，这时候得到的中心误差值是不准确的，所以需要精确度图来评估跟踪器的整体性能，构成图像曲线的每一点是跟踪器预测的目标中心位置 and 实际标定的目标中心位置的中心误差小于规定的阈值(通常为 20 像素)的帧数占据视频序列总帧数的比值。如果跟踪器在当前某一帧跟丢目标，那么当前帧的目标中心位置误差一定会大于阈值，这一帧不会被计入准确率计算。

成功率图：另外一种衡量算法精准度的算法的标准就是边界框重叠率，边界

框重叠率 S 的计算公式如(4.2)所示:

$$S = \frac{|R_G \cap R_T|}{|R_G \cup R_T|} \quad (4.2)$$

其中 R_G 代表的是某一帧对跟踪目标实际标定的区域, R_T 代表这一帧由跟踪器进行对目标位置进行判定的区域。 \cap 与 \cup 分别代表了这个区域的交和并运算,边界框重叠率的算法如图 4.11 所示:

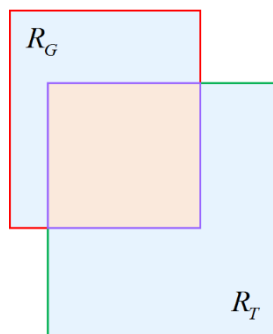


图 4.11 边界框重叠率计算方法

首先求出紫色矩形框中的像素个数,也就是求 $R_G \cap R_T$,然后再求 R_G 和 R_T 的总像素个数,用来减去 $R_G \cap R_T$,得到 $R_G \cup R_T$,最后把两个值进行比值计算得到边界框重叠率。与计算成功率图的原理相同,用满足边界框重叠率阈值的帧数和视频序列总帧数比率画出来的图叫做准确率图,设定的边界框重叠率的阈值一般为 0.6。

鲁棒性评估: 对于跟踪器评估有三种方法, OPE, SRE 和 TRE。OPE 是比较传统的用来评估跟踪器的方法,通过 groundtruth 提供的第一帧中的目标位置进行初始化,然后通过多个测试序列运行算法,最后得到的是关于平均精确度和成功率的结果。所以通常 OPE 也被称作是一次性通过的算法评估方法。另外两种方法分别从不同帧和不同的边界框开始跟踪,这两种评估方法又分别称为时间鲁棒性评估(TRE)和空间鲁棒性评估(SRE)方法。

4.3.2 跟踪算法的精确度和成功率评估结果

本文采用的是传统的跟踪算法评估方法——OPE,针对不同算法之前所有选取的 46 个视频序列跟踪之后保存的矩形框位置结果进行精确度和准确度曲线绘制,得到了不同跟踪算法总体的精确度图和准确度图,如图 4.12 所示,根据目标跟踪中出现如遮挡,尺度变化等等这些挑战对 46 个视频序列分别进行分类,得到了各种跟踪算法在不同的挑战情况下的精确度图和准确度图,如图 4.13~4.23 所示。

如图 4.12 中的性能曲线图像,与之后的图 4.13~4.23 类似,都拥有一张成功率图和一张准确度图用来评估跟踪算法的性能。

对于成功率图来说,拥有两个坐标轴,横坐标参数代表的是之前介绍过的边界框重叠率,横坐标轴上具体的数字代表的是边界框重叠率的阈值,变化的范围

从 0 到 1；纵坐标表示的是随着横坐标边界框重叠率的变化，不同算法在同一视频测试序列集中平均成功率。平均成功率的变化随着横坐标边界框重叠率阈值的变化而发生变化；每当在横坐标上确定一个阈值以后，对于一个跟踪算法，测试工具程序就会去针对测试集中每一个视频序列去统计在运行这个算法的时候序列中边界框重叠率小于阈值的帧数占据当前这个视频序列的总帧数的比率，最后针对测试集中的所有序列的比率取平均值，得到的数值就是这个算法图像上某个阈值对应的成功率。针对不同的算法也可以画出不同的曲线，对于图 4.12 与图 4.13~4.23 的成功率图的差别在于，图 4.13~4.23 用于测试集的视频序列是专门针对某一类目标跟踪挑战的视频序列集，而图 4.12 生成结果的测试集是包括所有选中的 46 个视频序列。

对于精确度图来说，横坐标代表的是代表的是中心位置误差的阈值，变化范围从 0 到 50，纵坐标代表的是跟踪的准确度，变化范围从 0 到 1，其计算思想和方法都与成功率图中的成功率计算方法类似，对于图 4.13~4.23 与图 4.12 的中心误差图的差别也与上面成功率图的情况类似。在此不再赘述。

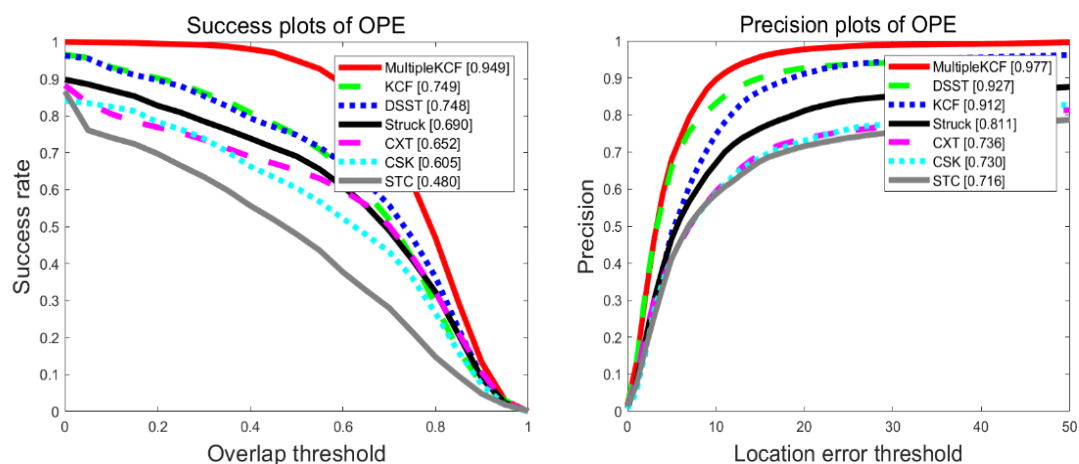


图 4.12 不同算法在所有视频序列中使用 OPE 评估得到的成功率和准确率图

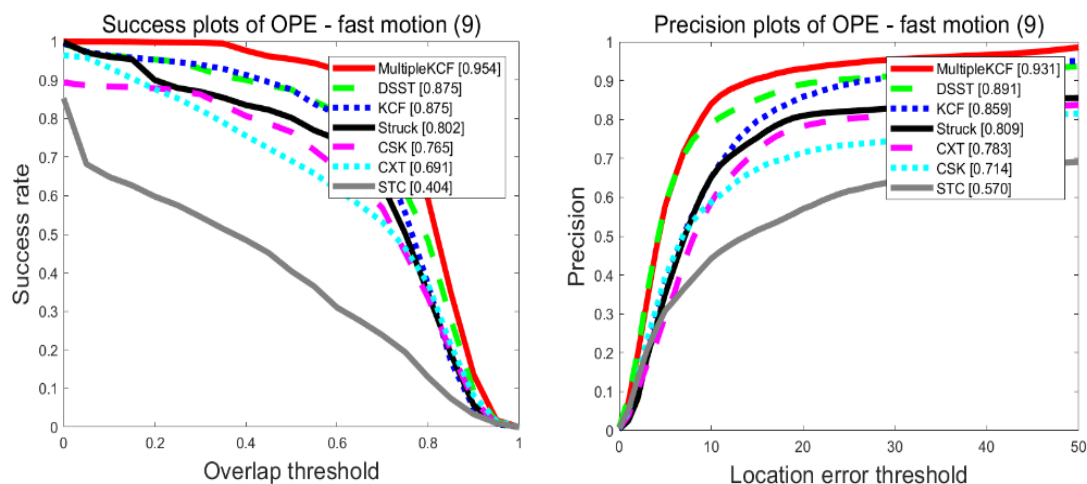


图 4.13 不同算法在快速移动情况下使用 OPE 评估得到的成功率和准确率图，其中右上角的数字代表在测试集中拥有跟踪目标拥有快速移动情况的视频序列有 9 个

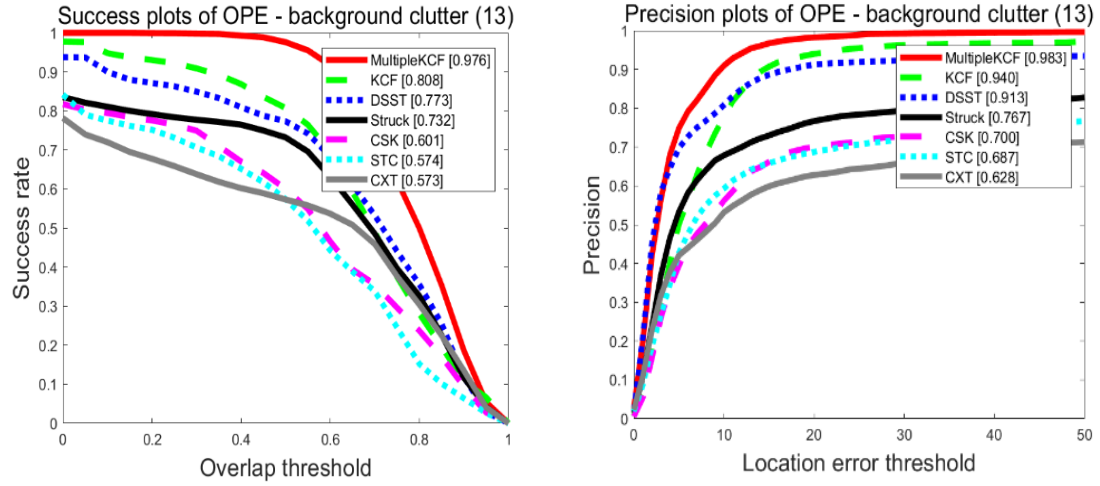


图 4.14 不同算法在背景杂乱情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标拥有背景杂乱干扰的情况的视频序列有 13 个

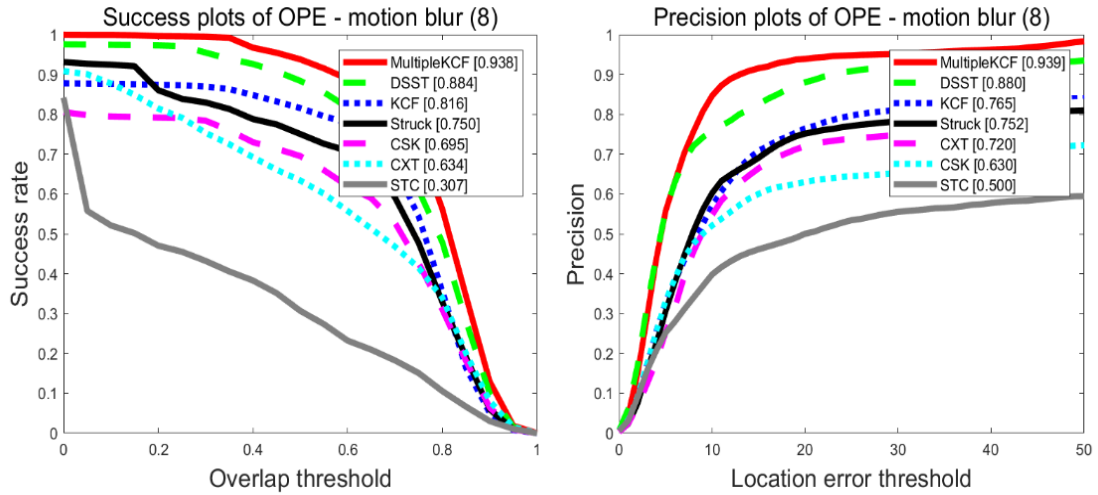


图 4.15 不同算法在运动模糊情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标拥有运动模糊干扰的情况的视频序列有 8 个

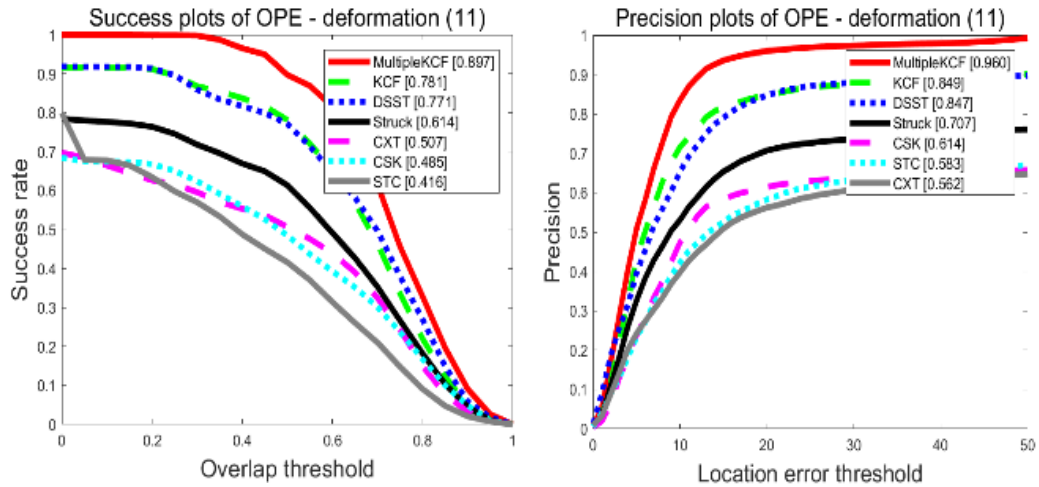


图 4.16 不同算法在目标形变情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生形变的情况的视频序列有 11 个

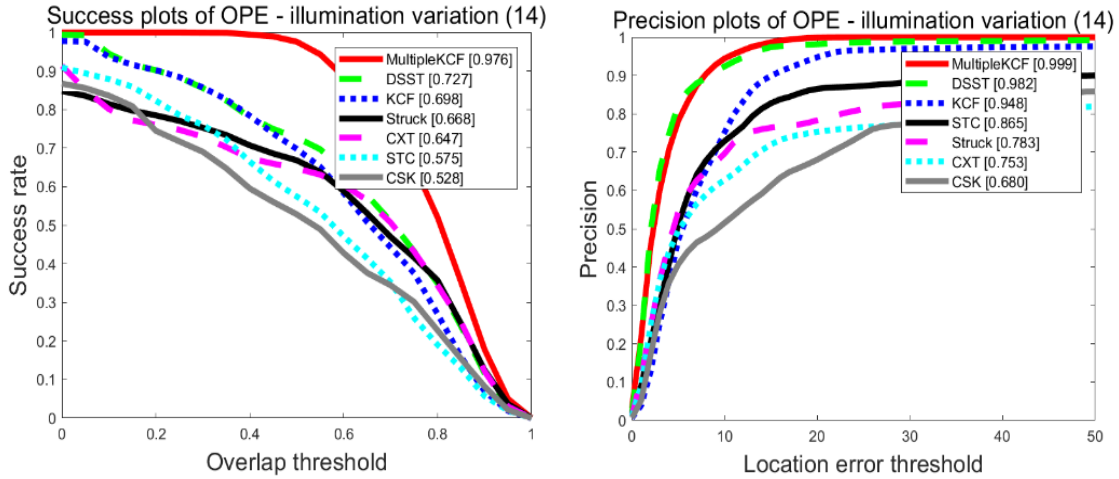


图 4.17 不同算法在光照变化情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生光照变化的情况的视频序列有 14 个

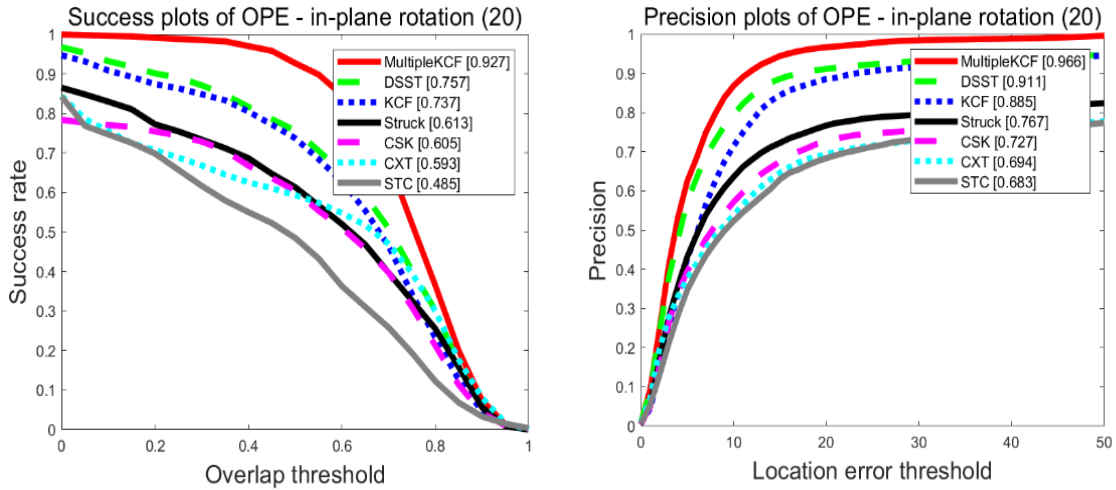


图 4.18 不同算法在平面内旋转情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生平面内旋转的情况的视频序列有 20 个

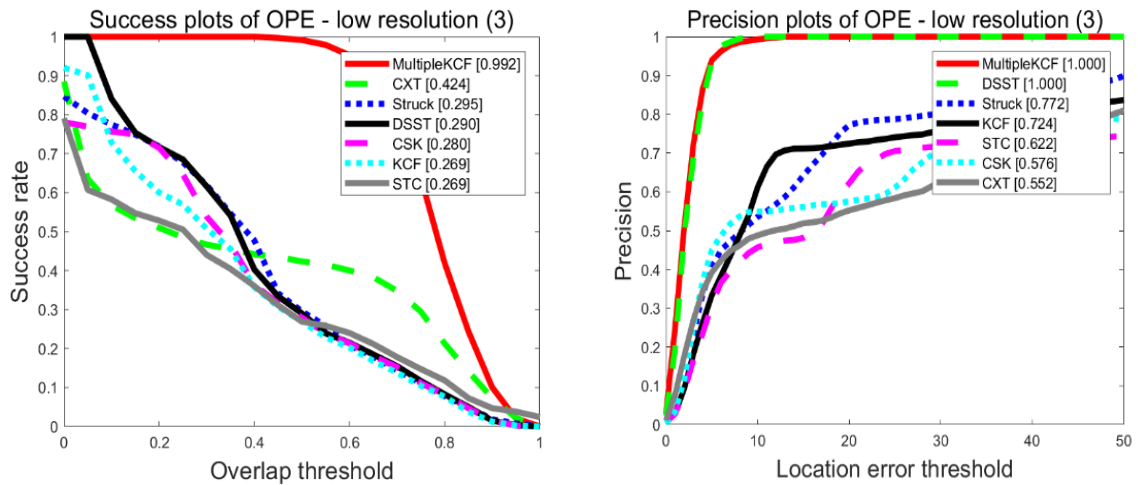


图 4.19 不同算法在低分辨率情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生低分辨率的情况的视频序列有 3 个

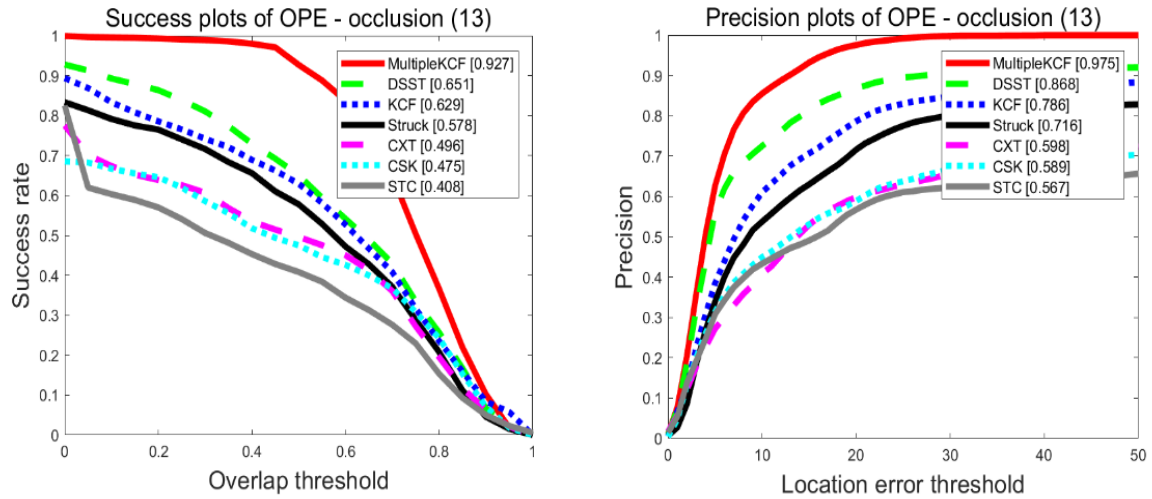


图 4.20 不同算法在遮挡情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生遮挡的情况的视频序列有 13 个

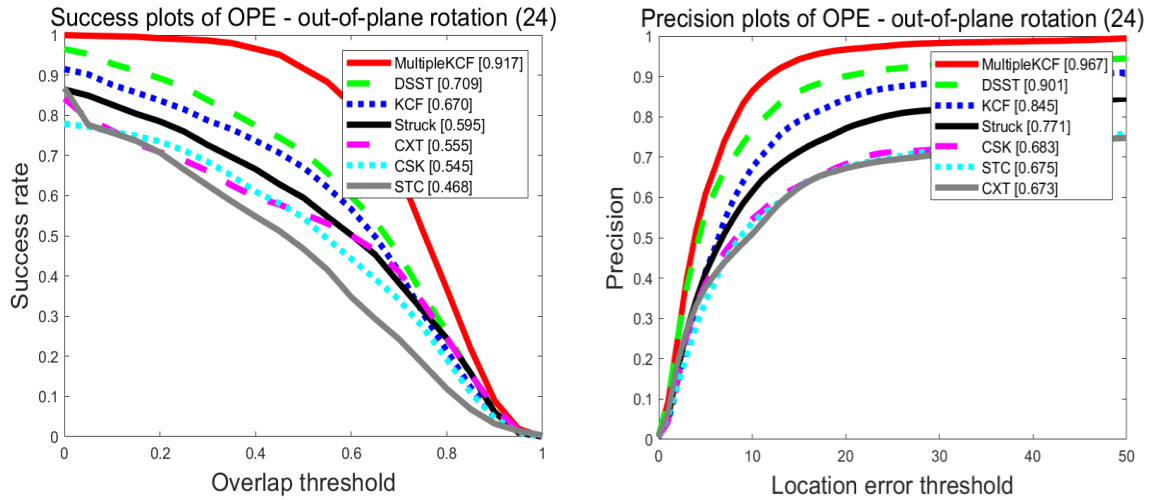


图 4.21 不同算法在平面外旋转情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生平面外旋转的情况的视频序列有 24 个

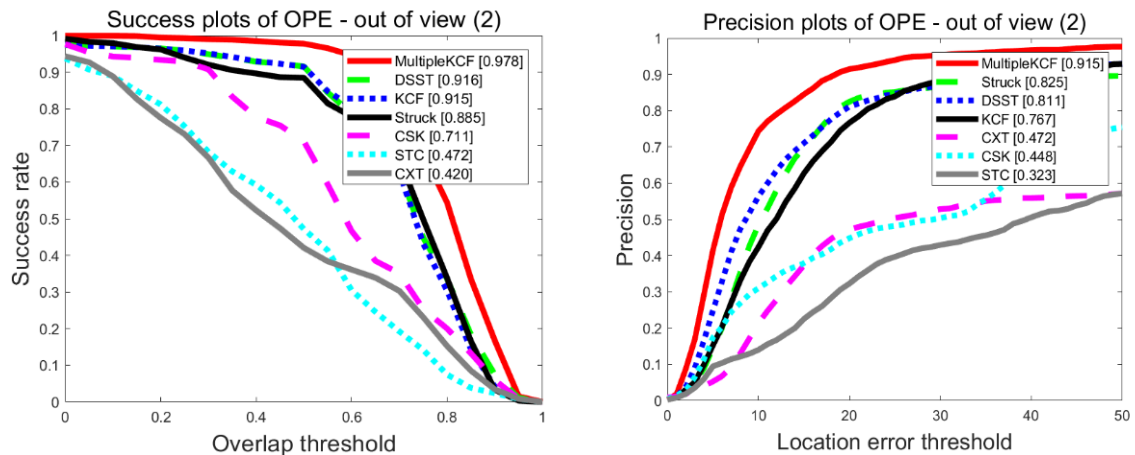


图 4.22 不同算法在超出视野情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生超出视野的情况的视频序列有 2 个

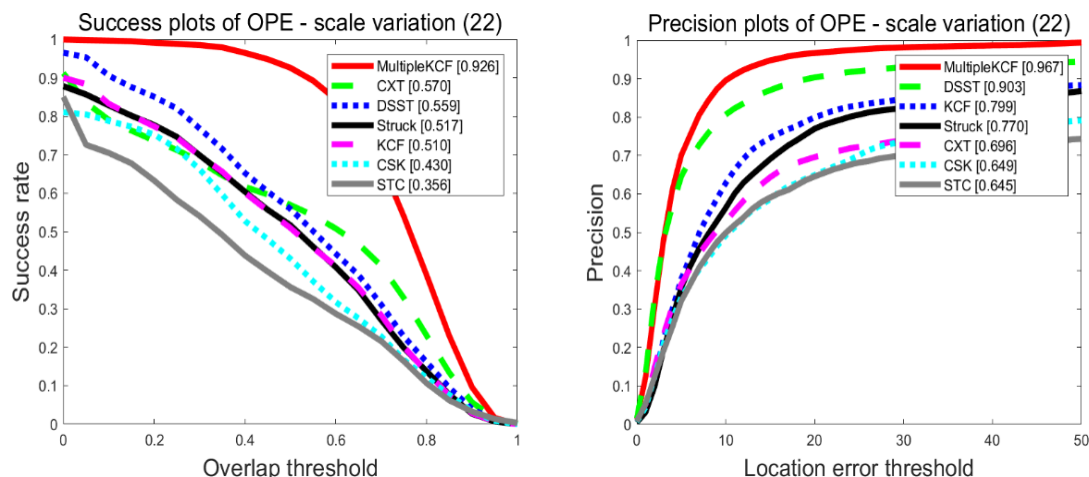


图 4.23 不同算法在尺度变换情况下使用 OPE 评估得到的成功率和准确率图, 其中右上角的数字代表在测试集中拥有跟踪目标发生尺度变换的情况的视频序列有 22 个

通过对以上在不同跟踪挑战下不同算法的性能的曲线的观察和比较。可以从两个角度对性能曲线进行分析:

首先, 无论是对于总体视频序列的还是个别挑战情况下的成功率或精确度曲线图, 根据横纵坐标参数代表的意义和跟踪算法对应的性能曲线图例可以看出: 随着横坐标阈值的变化, 本文算法的曲线位置均在其他滤波类算法之上, 其跟踪的精确度和成功率均优于其他算法。

另外, 对于每一张成功率图或者是曲线图, 右上角除了算法对应的图例表示排名外, 还有它们对应的成功率值或者准确率值, 其计算的方法在本节开头已经提到, 这里不再赘述。统计所有的成功率图, 在不同的跟踪挑战情况下, 排名第一名的跟踪算法 (本文算法) 的成功率一般都比排名第二名的跟踪算法高 1%-30%, 在有的跟踪挑战情况下, 如 in-plane rotation (平面内旋转), occlusion (遮挡), scale variation (尺度变化) 这个比例甚至能分别达到 34.2%, 42% 和 62.4%; 统计所有的准确度图, 在不同的跟踪挑战情况下, 排名第一名的跟踪算法的准确度一般都比排名第二名的跟踪算法高 1%-20%。

从以上的两个角度对各种挑战情况的曲线图进行分析, 可以得出结论, 本文提出的算法在成功率和准确度两个性能指标上均优于与之比较的其他滤波类算法。

4.4 本章小结

本章首先对实验所用数据集以及开发环境设置做了简要介绍, 之后对实验选取的视频序列中出现的目标跟踪挑战情况进行了分析。从用于比较的 46 个视频序列中选取了 10 个最具有代表性的视频序列, 并将本文算法和其他跟踪类算法对这 10 个序列中目标的跟踪效果进行定性分析。然后通过存储跟踪结果的文件画出算法对于选取的视频序列集进行跟踪的精确度图和准确度图, 经过再次细化, 同时得到了在不同跟踪挑战情况下对于不同跟踪算法的精确度图和准确度图, 最后针对这些图对算法的性能进行定量分析。经过对本文算法和其他滤波算法

的定性和定量分析,可以得出结论,本文实现的算法,无论是在各种跟踪挑战情况下,还是在跟踪的精确度和准确度上,效果和鲁棒性都优于与之比较的滤波类算法。

5 结论与展望

在计算机视觉领域,视频序列下的目标跟踪一直都是一个值得研究的重要课题,而通常在目标跟踪的过程中分为两种情况:在静态场景下的跟踪和在动态场景下的跟踪,相比起静态场景下的目标跟踪,在动态场景下的跟踪过程中视频序列中采集的跟踪目标的图像和其周围图像都会发生变化,这会给目标跟踪带来一系列的挑战,如尺度变换,遮挡,光照变化,形变等等。所以在动态背景下的目标跟踪更具有挑战性。本文通过选用和实现了一种适合在动态背景中的判别类跟踪算法——核相关滤波算法(KCF),并且通过分析此算法存在的不足来对算法进行改进。提出了一种基于核相关滤波器的多尺度融合特征抗遮挡算法,并使用基准测试工具 Visual Tracker Benchmark 把本文实现算法与其他目前主流的滤波类算法在标准视频序列测试集 OTB100 中对目标的跟踪效果进行定性分析和比较,同时对本文算法和其他滤波类算法进行跟踪性能评估,通过精确度和准确率两个跟踪性能指标对所有参与评估的算法进行定量分析和比较。本文主要完成的研究工作如下:

(1) 针对动态场景通过实现 KCF 算法来进行目标跟踪,对于 KCF 算法本身具有的缺陷进行改进。针对 KCF 算法未实现多尺度功能的缺陷,通过构建多分辨率金字塔模型的思想来构建对于跟踪目标的多尺度模型,根据不同尺度的响应峰值大小找到最适合跟踪目标大小的尺度。针对 KCF 算法在目标遮挡情况下造成的跟踪框漂移问题,通过最适合目标的尺度的最大响应峰值是否小于设定阈值来检测遮挡,通过卡尔曼滤波器记录目标遮挡前的运动特征信息来对目标在遮挡中会出现的位置进行预测,从而实现在遮挡中对目标的跟踪。针对目标在跟踪过程中出现的模糊和形变问题,通过实现对颜色敏感的 Lab 颜色特征和对形变敏感的 HOG 特征的融合,从而得到一种能够对目标形变和模糊都不敏感的融合特征用于跟踪目标的特征检测。

(2) 对本文算法和其他具有代表性的滤波类算法的跟踪效果进行定性分析。通过从 OTB 跟踪数据集中选取一些拥有不同跟踪挑战的代表性序列对所有算法进行跟踪测试,之后从中选取最具有代表性的几个挑战序列,对本文算法和其他算法在不同挑战情况下的跟踪效果进行分析和比较,得到本文算法在各种挑战环境进行的跟踪效果优于其他滤波类算法的结论。同时对所有算法的跟踪效果进行定量分析比较。在跟踪测试集为事先选定的所有视频序列集的情况下,绘制所有用于测试的跟踪算法的精确度和准确度性能曲线;在跟踪测试集为拥有同一类挑战情况的视频序列集合的情况下,针对不同挑战情况分别绘制出用于测试的跟踪算法的精确度和准确度曲线。由不同算法的性能曲线可以看出本文提出的算法在阈值发生变化时的成功率和精确度均高于选定用于比较的其他滤波类算法。

本文算法针对动态场景中目标跟踪出现的挑战情况对核相关滤波算法进行改进,并且将改进算法与其他目前主流的滤波算法进行了定性和定量分析。但是

鉴于本文作者的水平有限,算法还存在一定的缺陷和不足并且拥有很大的改进空间。具体地介绍以下两点:

(1) 本文的算法针对目标的短时部分遮挡和完全遮挡的跟踪效果较好,但是对于长时跟踪的效果不佳,如果要核相关滤波算法应用于长时跟踪,目前对于这个问题的解决一般有两种思路:一种方法是通过把要跟踪的目标的图像进行分块,然后分别去提取分块中的特征,最后将它们融合用来获得目标的位置^[27];另一种方法是把跟踪器分为两部分,一部分用于对目标的尺度进行预测,另外一部分用于对目标的表观特征变化进行上下文学习。通过 K 近邻分类算法对目标的位置做出一个最好的预测^[28]。

(2) 本文算法目标遮挡检测的方法有待于改进。一般在正常情况下,最适合跟踪目标的尺度的最大峰值(位于跟踪目标的中心点)小于设定的阈值 0.25 即可以判断为发生遮挡,但是在一些少数特殊的视频序列中判断遮挡时这个阈值会发生变动,在这种情况下通过设定的阈值 0.25 来判断目标是否发生遮挡实际上是不可行的,所以要让跟踪器检测到遮挡需要有更好的遮挡检测办法。

致 谢

时光总是匆匆地逝去，转眼间大学四年的生涯即将结束。作者在求学路上遇到了很多无论是从生活方面，还是从学习方面都给予作者很多帮助的人，借此机会要向他们说一声谢谢。

感谢作者的导师胡秀华老师，感谢胡老师在过去几个月里从学生毕业设计开始到结束全程的耐心指导。胡老师渊博的知识，对待学术严谨的作风，认真的做事态度和对人宽容的性格给学生留下了十分深刻的印象。学生平时不仅在研究上得到胡老师的启发，而且在处事方面也在胡老师身上学到了很多。

感谢大学四年教授学生所有课程的老师，没有你们的努力也就没有学生今天取得的成果。你们在学习和做人方面都教会了学生很多，也给学生有了很大的启发。

感谢本人的同学和朋友在学习和生活上给予的帮助，正是因为有了你们的支持，本人的大学四年一直奋发努力，直到现在。

向大学四年所有对作者提供过帮助的人表示衷心的感谢！

参考文献

- [1] 陈纪铭,陈利平.基于智能图像视觉的特定场景监控系统设计.电子制作,2019(01):87-88+86.
- [2] 汪海,王羽中,汪源.基于智能视觉的监控识别系统设计.自动化与仪器仪表,2019(04):49-53.
- [3] Pan Z, Liu S, Sangaiah A K, et al. Visual attention feature (VAF): a novel strategy for visual tracking based on cloud platform in intelligent surveillance systems. *Journal of Parallel and Distributed Computing*, 2018, 120: 182-194.
- [4] 陈小平. 基于稀疏编码视频目标跟踪算法研究.硕士学位论文.江苏: 江南大学,2018
- [5] 储向锋. 稀疏表示在目标跟踪中的应用: 硕士学位论文.江苏: 南京邮电大学,2016
- [6] 李红江.基于高斯核相关滤波的多特征融合的目标跟踪.电子技术与软件工程,2019(03):99-100.
- [7] 吴昭童,左颢睿,徐智勇,张建林.特征融合的核相关滤波跟踪算法.国外电子测量技术,2019,38(02):50-55.
- [8] 李国友,纪执安,张凤煦.一种基于多特征融合的核相关滤波目标跟踪算法.计算机与应用化学,2018(12):1004-1011.
- [9] Wang X, Wang G, Zhao Z, et al. An Improved Kernelized Correlation Filter Algorithm for Underwater Target Tracking. *Applied Sciences*, 2018, 8(11): 2154.
- [10] Hou Y, He Y. Improved Multi-sampling Kernelized Correlation Filter Target Tracking Algorithm. In 3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019). Atlantis Press, 2019.
- [11] Liu M, Diao C, Cheng X, et al. The performance comparison of different feature to kernelized correlation filter tracker. In Tenth International Conference on Graphics and Image Processing (ICGIP 2018). International Society for Optics and Photonics, 2019, 11069: 110694A.
- [12] 李远状,韩彦芳,于书盼.一种核相关滤波器的多尺度目标跟踪方法.电子科技,2018,31(10):1-5+10.
- [13] 江维创,张俊为,桂江生.基于改进核相关滤波器的目标跟踪算法.计算机工程,2018,44(11):222-227.
- [14] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 2544-2550.
- [15] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels. In European conference on computer vision.

- Springer, Berlin, Heidelberg, 2012: 702-715.
- [16]Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters. *IEEE transactions on pattern analysis and machine intelligence*, 2014, 37(3): 583-596.
- [17]Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking. In *British Machine Vision Conference*, Nottingham, September 1-5, 2014. BMVA Press, 2014.
- [18]Danelljan M, Hager G, Shahbaz Khan F, et al. Learning spatially regularized correlation filters for visual tracking. In *Proceedings of the IEEE international conference on computer vision*. 2015: 4310-4318.(SRDCF)
- [19]周启晨,李云栋.基于深度学习的目标跟踪方法及其实现. *工业控制计算机*, 2019, 32(02): 89-90.
- [20]师燕妮.基于深度学习目标检测与跟踪技术的研究. *电子设计工程*, 2019, 27(06): 59-63.
- [21]肖驰. 动态场景中目标跟踪的遮挡处理与研究. 北京: 北京理工大学, 2016
- [22]崔盼果. 复杂场景下遮挡干扰目标鲁棒跟踪技术研究. 安徽: 中国科学院大学 (中国科学院光电技术研究所), 2018
- [23]Wu Y, Lim J, Yang M H. Online object tracking: A benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013: 2411-2418.
- [24]Dinh T B, Vo N, Medioni G. Context tracker: Exploring supporters and distracters in unconstrained environments. In *CVPR 2011*. IEEE, 2011: 1177-1184.
- [25]Hare S, Saffari A, Struck P H S T. Structured output tracking with kernels. In *IEEE International Conference on Computer Vision*. IEEE. 2012: 263-270.
- [26]Zhang K, Zhang L, Yang M H, et al. Fast tracking via spatio-temporal context learning. *arXiv*, 2013. arXiv preprint arXiv:1311.1939.
- [27]Akin O, Mikolajczyk K. Online learning and detection with part-based, circulant structure. In *2014 22nd International Conference on Pattern Recognition*. IEEE, 2014: 4229-4233.
- [28]Ma C, Yang X, Zhang C, et al. Long-term correlation tracking. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015: 5388-5396.

毕业设计（论文）知识产权声明

本人完全了解西安工业大学有关保护知识产权的规定，即：本科学生在校攻读学士学位期间毕业设计（论文）工作的知识产权属于西安工业大学。本人保证毕业离校后，使用毕业设计（论文）工作成果或用毕业设计（论文）工作成果发表论文时署名单位仍然为西安工业大学。学校有权保留送交的毕业设计（论文）的原文或复印件，允许毕业设计（论文）被查阅和借阅；学校可以公布毕业设计（论文）的全部或部分内容，可以采用影印、缩印或其他复制手段保存毕业设计（论文）。

（保密的毕业设计（论文）在解密后应遵守此规定）

毕业设计（论文）作者签名：

指导教师签名：

日期：

毕业设计（论文）独创性声明

秉承学校严谨的学风与优良的科学道德，本人声明所呈交的毕业设计(论文)是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，毕业设计（论文）中不包含其他人已经发表或撰写过的成果，不包含他人已申请学位或其他用途使用过的成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了致谢。

毕业设计（论文）与资料若有不实之处，本人承担一切相关责任。

毕业设计（论文）作者签名：

指导教师签名：

日期：