第 59 卷 第 8 期/2022 年 4 月/激光与光电子学进展 研究论文

激光与光电子学进展

先进成像

基于抗遮挡目标模型的跟踪算法综述

谢郭蓉[1,2]，曲毅[2*]，蒋镕圻[1,2]

[1]武警工程大学研究生大队，陕西 西安 710086；

[2]武警工程大学信息工程学院，陕西 西安 710086

摘要 遮挡问题是导致目标跟踪任务失败的重要因素，如何提升算法的抗遮挡性能是跟踪领域的研究热点。本文首先剖析了遮挡易导致跟踪失败的原因，论述了构建强判别性的鲁棒目标模型对提高跟踪算法抗遮挡性能的重要意义，分析了抗遮挡目标模型的构建方案。其次依据目标模型利用的信息类型，将代表性抗遮挡性能较优的算法分为基于有效特征信息、状态估计信息与稳定时空信息三类。而后详尽分析了基于卡尔曼滤波、粒子滤波、局部空间信息、时间上下文信息、时空上下文信息跟踪算法的抗遮挡思路方案、适用遮挡场景、优缺点及改进方案。最后通过不同类型算法在遮挡场景下的跟踪性能比较，对目标模型构建方案抗遮挡的有效性提出思考与分析，并指出学习语义信息轻量化网络设计、场景上下文预测、仿生视觉机理的应用发展方向。

关键词 机器视觉；目标跟踪；抗遮挡；状态估计信息；时空上下文

中图分类号 TP391.4 **文献标志码** A **doi**：10.3788/LOP202259.0815001

Tracking Algorithms Based on Antiocclusion Object Models

Xie Guorong[1,2], Qu Yi[2*], Jiang Rongqi[1,2]

[1]*Postgraduate Brigade, Engineering University of PAP, Xi'an, Shaanxi* 710086, *China*;

[2]*School of Information Engineering, Engineering University of PAP, Xi'an, Shaanxi* 710086, *China*

Abstract Occlusion is an essential factor that often leads to the failure of object tracking. Improving antiocclusion performance of the algorithm has been a research hotspot in tracking. First, this paper analyzes why occlusion easily leads to tracking failure. Furthermore, the importance of constructing a strong discriminant and robust object model to improve the antiocclusion performance of the tracking algorithm and an effective scheme to improve the antiocclusion performance of the target model are discussed. Then, based on the utilization information type of constructing object model, the representative methods with better antiocclusion performance are divided into three categories on the basis of effective feature, state estimation, and stable spatiotemporal informations. Further, the antiocclusion idea scheme, suitable occlusion scene, pros and cons, and improvement schemes of object tracking algorithm based on Kalman filter, particle filter, local spatial information, time context information, and spatiotemporal context information are analyzed in detail. Finally, through performance comparison with the tracking performance of different types of methods in occlusion scenarios, the antiocclusion effectiveness of the object model construction scheme is analyzed. The application and development direction of learning semantic information lightweight network design, scene context prediction, and bionic vision mechanism are presented.

Key words machine vision; object tracking; anti-occlusion; state estimation information; spatiotemporal context

收稿日期：2021-02-24；**修回日期**：2021-03-29；**录用日期**：2021-04-15

基金项目：国家自然科学基金(No. 61101238)

通信作者：*wjquyi@sina.com

1 引 言

作为计算机视觉领域的重要热点分支，目标跟踪在无人驾驶、智能监控、人机交互、战场态势侦察等领域具有广泛应用[1]，是完成场景感知、行为理解等高层级视觉任务的前提和基础[2]。跟踪任务中的挑战因素主要有目标形变、光源变化、背景杂乱、尺度变化和遮挡等。随着目标跟踪算法不断发展，光源变化、尺度变化等场景下，跟踪算法性能已有显著提高，但遮挡场景下的算法性能仍存在较大提升空间。因此，提升算法抗遮挡性能仍是目标跟踪领域的重点与难点问题。当目标遭遇遮挡时，对跟踪模型有以下不利影响：1)遮挡物在目标候选框内引入了语义或背景干扰，可能导致模型误判，漂移至干扰物上；2)目标因受遮挡而发生信息丢失，可能导致模型漏检，跟丢目标；3)完全遮挡或长时遮挡时，目标信息处于完全消失状态，可能导致模型在目标出现后无法从漂移状态中恢复正确跟踪。

经许多算法论证，构建有强判别力、足够鲁棒的目标模型，是解决误判干扰物、漏判变化目标等问题的根本方案。因此，研究先进算法的抗遮挡目标模型，对提高算法对遮挡等实际复杂场景的适用性有重要意义。为提高跟踪算法的抗遮挡性能，已有国内外学者对提高目标模型的判别鲁棒性进行了研究。文献[3]阐述了遮挡场景下各类跟踪算法的基本思想与目标模型的遮挡处理功能，但研究对象为早期跟踪算法。文献[4-6]针对各类跟踪算法不足，阐述了主流改进方案，但针对抗遮挡方案的论述较少，且缺少抗遮挡的本质功能分析。本文对先进跟踪算法的抗遮挡目标模型进行了系统梳理，旨在为提高复杂场景下的跟踪算法性能提供参考。

先进算法提高目标模型的抗遮挡性能主要有以下方案：1)通过选择提取目标的有效外观运动信息，一方面使目标模型学习到能区分目标对象与干扰物的本质特征，缓解遮挡物对模型的干扰问题，另一方面是选择对遮挡场景不敏感的特征，保持模型对遮挡目标的有效判别；2)利用目标运动状态信息进行状态估计，降低模型对目标外观的依赖性，从而避免漏检外观已变的目标；3)引入跟踪序列时空信息构建目标模型，一是在部分遮挡下仍存在稳定可靠的目标局部空间信息，其区域间的关联有利于灵活重建；二是目标的时序上下文记忆信息有益于构建长期稳健的目标模型，避免过拟合集中时段的遮挡受损样本；三是目标周围区域时空上下文信息有遮挡不变性，可提升部分遮挡时目标跟踪的稳健性。

本文依据提高目标模型抗遮挡性能的构建方案与利用信息类型，将有较优抗遮挡性能的跟踪算法分为基于有效特征信息、基于状态估计信息、基于稳定时空信息的三种跟踪算法。同时分析了各类算法的抗遮挡思路方案、适用遮挡场景、优缺点及改进方案，并分析比较了目标模型抗遮挡的有效性，指出了有益于提升目标模型抗遮挡性能的应用技术研究方向。

2 基于有效特征信息的目标跟踪算法

Wang等[7]提出，与跟踪系统框架的其他模块相比，特征提取模块的质量是影响跟踪算法性能最关键的因素。融合特征可弥补单一特征场景的局限性，选择融合多类场景适应性最优的特征有益于全面刻画目标外观、运动等信息。分析了常见特征的场景适应性特点，介绍了利用融合特征表征目标的代表性跟踪算法的目标模型特点及其遮挡适应性。

2.1 特征的场景适应性分析

常用于跟踪的特征有：基于相邻帧像素位移变化矢量计算的光流运动特征、根据需求设计的手工外观特征及无监督学习得到的深度特征。表1为其代表性特征描述子的场景适应性特点和代表跟踪算法。

与外观特征不同，光流特征描述了目标运动信息，短时遮挡时可降低模型对受损外观样本的依赖性。传统光流估计法的约束方程依赖于亮度恒定、时间连续、空间一致的假设限制，因此在求解光流场的实时性、可靠性方面具有提升空间。表2为光流估计法的后续改进方案与代表算法。

相比于手工特征，深度特征具有更强的目标刻画与迁移学习能力[27]。实际上，网络不同层特征表征信息的侧重点不同，表3是以VGG-Net为代表的深层与浅层特征的表征特点。Bhat等[28]认为深度特征偏向于鲁棒性，而浅层特征更关注准确度，因此跟踪任务常融合深层、浅层特征表征目标，以提高复杂场景下的模型判别力与定位精确度。

为提高目标模型抗遮挡性能，应选择提取对遮挡不敏感的特征信息，包括对部分遮挡不敏感的颜

表 1　跟踪特征场景适应性特点及对应代表性算法

Table 1　Scene adaptability of tracking features and representative methods

Feature descriptor		Representation methods	Scene adaptability
Manual visual feature	Intensity	MOSSE[8], CSK[9]	Single stable scene
	Gray		Single stable scene
	Color	CN[10], ASMS[11],	Partial occlusion, fast move, scale invariation
	CN, Lab	DAT[12], CSCT[13]	Inapplicable to light changes
	Gradient	KCF[14], BACF[15],	Translation, rotation, illumination invariation
	HOG, SIFT	LCT[16], DSST[17]	Inapplicable to deformation and motion blur
	Texture	TLD[18]	Illumination, scale, fast motion invariation
	LBP		Inapplicable to deformation, time-consuming
Optical Flow		FlowTrack[19]	Suitable for short-term occlusion and scale change
			Not suitable for severe occlusion, camera shake, and out-of-view
Deep feature		Deep SRDCF[20], CFNet[21]	Strong robustness to complex scenes including occlusion
			Poor real-time performance and interpretability

表 2　光流估计法的发展

Table 2　Development of optical flow estimation methods

Limitations of traditional methods	Related improvements measures	Methods
Time-consuming	Key point optical flow calculation instead of global filed	[22-23]
Limitation of small displacement	Construct an optical flow pyramid for downsampling	[24]
Inaccurate expression	Deep optical flow estimation	[25-26]
Uncertainty under occlusion	Symmetry between optical map and occlusion area	[26]

表 3　VGG-Net不同层提取特征的表征特点

Table 3　Expression traits of features extracted from VGG-Net different layers

Feature type	High-level feature of VGG-Net	Low-level feature of VGG-Net
Extraction layer	Conv4-4, Conv5-4	Conv1-2, Conv2-2
Description	Semantic nature information	Contour and texture
Advantages	Robust to appearance change, occlusion	Essential positioning
Disadvantages	Low spatial resolution	Poor anti-interference ability

色特征、对短期遮挡不敏感的光流特征、适用于复杂场景表征的深层语义特征等。实际场景中遮挡、光照变化、目标形变等挑战因素往往会同时出现，而单一特征有场景局限性，如颜色特征对光照变化敏感，光流特征不适用相机抖动、严重遮挡等场景。因此一般选择融合多类有效特征，以提高目标模型在多类复杂场景下的鲁棒性。

2.2　基于有效特征融合的目标跟踪算法

单一外观特征的表征能力较弱，适用场景有限，因此许多算法融合外观特征，提高多类场景下的模型判别力。文献[29-36]融合颜色与梯度两种互补的手工特征因子，弥补了对光照、形变敏感的缺陷，可适应轻微遮挡等大部分跟踪场景。文献[37-40]融合网络浅层与深层特征，实现了语义特征粗定位到底层特征细定位，提升部分遮挡等复杂场景中的判别力。文献[41-43]构建描述外观、运动信息的协作模型，以更完备的信息表征降低跟踪器对外观模型的依赖。表 4 为基于融合外观特征表示的代表算法目标模型特点及适应场景。

许多算法融合有效特征构建目标模型时，并未提出抵抗遮挡、形变等属性干扰的针对性措施方案，但他们在复杂场景下的性能也远好于简单特征表征的算法，所以特征提取模块质量极大地影响了遮挡等场景下的跟踪性能。应当指出，特征提取模块是跟踪框架中耗时较大的部分，实时性与有效性的提升一直是研究重点；仅依靠对遮挡不敏感的高判别性特征目标模型，依旧无法应对严重遮挡、完全遮挡等更复杂的场景，因此针对遮挡场景引入不变性信息构建目标模型，是进一步提升算法抗遮挡能力的有效途径。

表4　基于融合特征表示的代表算法的目标模型特点及适应场景

Table 4　Object model traits and fitted scenes of representative algorithms based on fusion features

Type	Algorithm	Feature extracted	Object model characteristics	Fitted scenes
Color and gradient feature	SAMF[29]	HOG, CN, Grey	Scale adaptive with multiple features	Partial occlusion, deformation, rotations, interference
	MOCA[30]	MC-HOG	HOG extracted from CN channels	
	Staple[31]	Color, HOG	Global-local features, 80 fps of speed	
	SAT[32]	RGB, SIFT	Encoding structure, keypoints spatial layout	
	DPCF[33]	Color, HOG	Collaborative and deformable model	
	MvCFT[35]	HOG, CN, Grey	Fusion model of multi-perspective feature	
Appearance feature of different layers	HCF[37]	Conv3-4, Conv4-4,Conv5-4	Holistic model with weighted multi-feature filter confidence map	Occlusion, clutters, deformation, semantics-dis tractors
	C-COT[38]	Conv1, Conv5	Continuous spatial interpolation, preciser	
	ECO[39]	Conv1, Conv5	PCA decomposed convolution, fast	
	SANet[44]	CNN and RNN feature	Skip concatenation, self-structure encoding	
Appearance and motion feature	DFCNet[41]	Optical flow, CNN feataure	Adaptive keyframe scheduling mechanism	Clutters, fast movements, deformation, heavy occlusion
	DCTN[43]	Optical flow, CNN feature	Pyramidal feature hierarchy	
	STSGS[42]	Conv3-4, Conv4-4, Conv5-4	Quantum mechanics based saliency detection and motion flow map generation	
	FPRNet[45]	Optical flow, CNN and RNN feature	Multi-scale spatiotemporal representations by flow pyramid recurrent framework	

3　基于状态估计信息的目标跟踪算法

卡尔曼滤波与粒子滤波是跟踪任务中构建状态估计模型的两种经典方法，其通过预测遮挡目标轨迹，避免模型被短期遮挡或瞬间形变的目标污染，是解决遮挡问题的有效方案。下文介绍了卡尔曼滤波与粒子滤波基本理论，并分类阐述了代表性算法的抗遮挡机制。

3.1　基于卡尔曼滤波的目标跟踪算法

作为数据预测任务中公认的最佳方法，卡尔曼滤波方法常用于遮挡场景的跟踪任务，依据状态估计方程预测轨迹信息，所得目标轨迹有助于恢复被严重遮挡目标的状态，所得遮挡物轨迹有助于去除干扰候选框，避免模型漂移的同时提高实了时性。

卡尔曼滤波基于无偏最小方差估计，对系统状态估计模型进行递归预测修正，其预测模型以当前状态与误差协方差估计得到先验估计，为下步时间状态所用，其修正模型以新观测值与预测模型所得先验估计，对后验估计进行更新校正[46]，最后得到全局最优状态估计值。图1为卡尔曼滤波跟踪器算法流程图。

图中，$\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{H}$分别代表状态转移矩阵、控制输入矩阵和观测矩阵，$\boldsymbol{P}$、$\boldsymbol{Q}$为服从高斯分布的系统预

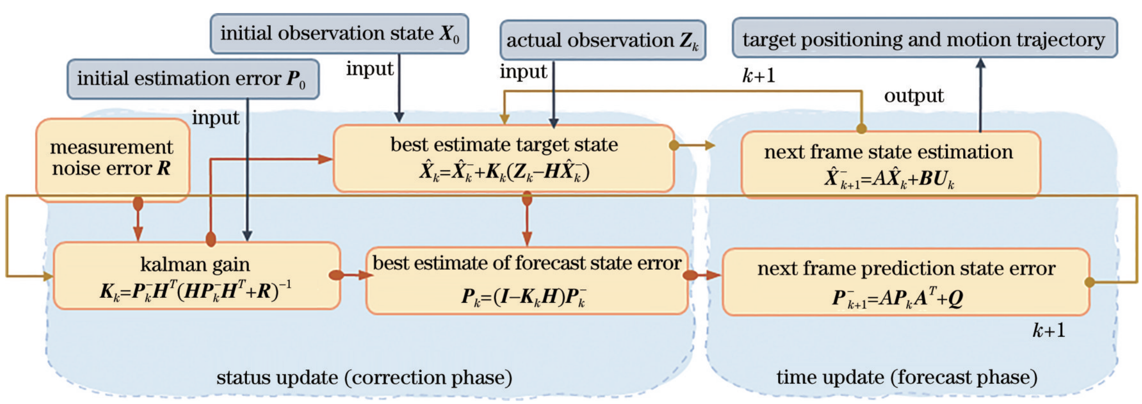

图1　卡尔曼滤波跟踪器流程图

Fig. 1　Kalman filter tracker flow diagram

测误差、观测噪声协方差矩阵，$\boldsymbol{X}_k$、$\boldsymbol{P}_k$、$\boldsymbol{Z}_k$分别为第k帧时的目标状态向量、预测状态协方差矩阵、实际观测向量，$\hat{\boldsymbol{X}}_k^-$、$\boldsymbol{P}_k^-$为未修正估计值，$\hat{\boldsymbol{X}}_k$、$\boldsymbol{P}_k$为最优估计值。

因传统卡尔曼滤波法（KF）构建状态观测方程基于线性动态与高斯噪声，对干扰场景的状态预测不可靠，因此许多学者提出了相应改进方法。主要有使用泰勒展开线性化处理的扩展卡尔曼滤波（EKF），基于无偏变换近似高斯分布的无迹卡尔曼滤波（UKF）。表5分析了三者特点及代表跟踪算法的场景适用性。

表5　三类卡尔曼滤波法特点及代表跟踪算法的场景适用性

Table 5　Characteristics and scenario applicability of three kinds of KF tracking methods

Methods	Type	System requirements		Characteristics and applicability of tracker methods	
		Status	Noise		
[34, 47-50]	KF	Linear	Gaussian	Vulnerable to interference	Applicable to severe occlusion temporarily
[51]	EKF	Non-linear	Gaussian	Easy to accumulate error	
[52-53]	UKF	Non-linear	Non-gaussian	Higher speed and accuracy	

为提高遮挡场景下目标跟踪的定位精确度，文献[34, 49]在检测到遮挡目标时，以卡尔曼滤波预测作为下一帧对象的位置。针对算法KCF（kernelized correlation filters）循环采样易引入遮挡受损样本的问题，文献[34]以峰值旁瓣比作为置信度指标，当低于阈值时认为当前帧目标处于被遮挡状态，此时利用卡尔曼滤波器基于历史帧预测遮挡目标的轨迹。该算法可有效缓解光照变化、相似干扰、遮挡时的跟踪漂移问题。为进一步提高卡尔曼滤波状态模型参数的准确性，文献[47-48]提出了自适应调节卡尔曼滤波状态模型参数的算法，将遮挡对跟踪的影响最小化，文献[48]依据方向梯度直方图（HOG）最大响应与位置误差存在的负相关关系，将HOG检测响应作为位置估计准确性度量，连续调整卡尔曼滤波器协方差参数，从而不断提高遮挡等场景下的跟踪准确性与鲁棒性。

上述算法在检测到遮挡时，利用卡尔曼滤波基于最新历史观测数据对目标轨迹状态进行实时预测，其快速的收敛速度有利于提升算法的实时性。但卡尔曼滤波预测的可靠性对最近观测数据质量的好坏有较大依赖，无法避免非短期遮挡受损样本对模型的污染。因此卡尔曼滤波方法可有效缓解因短暂遮挡的模型漂移问题，但在长期遮挡、出视野等场景下，仍有一定局限性。

3.2　基于粒子滤波的目标跟踪算法

相比卡尔曼滤波，基于蒙特卡洛采样思想的粒子滤波不仅摆脱了线性、高斯分布的约束，且保留了样本多样性的粒子滤波，能避免一段时间内遮挡受损样本对模型的污染。表6为卡尔曼滤波与粒子滤波状态估计的模型特点、抗遮挡机制及场景适用性。

表6　卡尔曼滤波与粒子滤波状态估计模型特点及场景适应性

Table 6　Scenes applicability and traits of status estimation model between Kalman filter and particle filter

Characteristics	Kalman filter	Particle filter
Probability model	State and observation models	Weighted particle swarm estimate
Requirments	Gaussian and linear system	Non-linear and non-gaussian system
Anti-occlusion	Prediction mechanism with fast convergence	Prediction and multi-modal maintenance
Applicable scenes	Short-term occlusion	Partial occlusion, background interference
Disadvantage	Applicable scene limitations	Particles degenerate, time-consuming

与卡尔曼滤波不同，粒子滤波以不同权重离散粒子表示目标状态的后验概率密度，其步骤为：粒子初始化后以建议分布重要性采样，在观测模型中计算粒子权重，依据粒子有效性阈值判断是否进行重采样，最后所得权重最高粒子代表状态最优估计。图2为粒子滤波跟踪流程图，其中$\boldsymbol{x}_k^i$、$\boldsymbol{w}_k^i$分别代表第k时刻、第i个粒子与对应权重，N代表样本集S中的粒子数。

依据提高抗遮挡性能方案的不同，可将结合粒子滤波的跟踪算法分为：1）基于状态估计空间参数进行遮挡检测的算法。因对应遮挡受损样本的粒子状态分布参数异常，该类算法依据参数分布可及时检测遮挡状态并处理；2）利用状态估计模型预测遮挡目标状态轨迹的算法。基于置信度指标检测

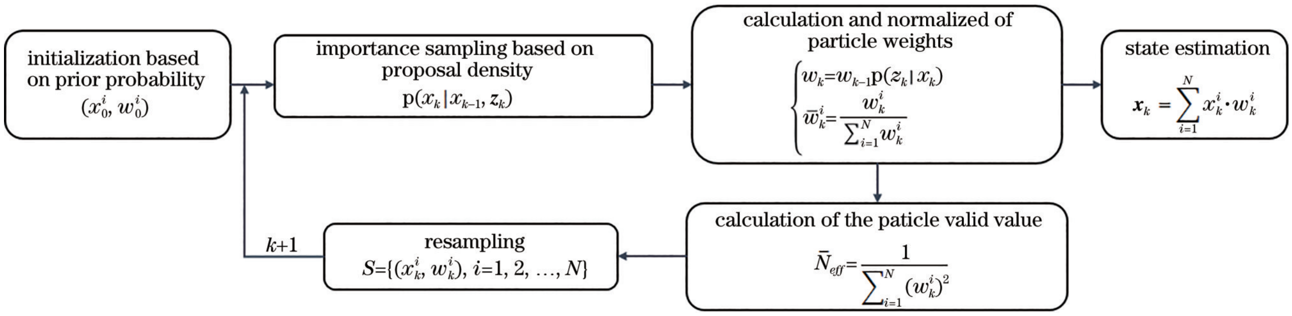

图 2　粒子滤波跟踪流程图

Fig. 2　Particle filter tracker flow diagram

到遮挡后，利用粒子滤波预测目标的状态轨迹，可提升定位准确性；3）基于粒子多样性保留以修正状态模型进行遮挡恢复的算法。为解决粒子退化问题，传统重采样方法丢弃了权重较小粒子，导致粒子样本缺乏多样性，不利于稳定跟踪因遮挡而表观已变的目标。而构建遮挡时会引入噪声的强判别特征观测模型，重建近似遮挡场景真实分布的状态模型，记忆或标记遮挡异常粒子，结合深度学习训练更准确的状态估计方程，都有助于状态估计空间多类分布粒子的保留，从而增强算法对遮挡场景的鲁棒性。表 7 为基于粒子滤波状态估计跟踪算法的抗遮挡方案、创新特点及适应遮挡场景。

表 7　基于粒子滤波状态估计跟踪算法的抗遮挡方案、创新点及适应遮挡场景

Table 7　Anti-occlusion scheme, innovation and occlusion adaptation of algothrims based on PF state estimation

Anti-occlusion scheme		Methods	Contribution and characteristics	Occlusion application
Occlusion detection		[54]	Normalization factor for occlusion detection	Long-term, fully, partial
		[55]	Internality of weight value and distribution region	
		[56]	Third-order cumulants of reconstruction error	
Trajectory prediction		[57]	Bhattacharyya coefficient as judgment criterion	Shot-term
Modified model to maintain particle diversity	Robust observation	[58]	Multiple likelihood models of HSV and HOG	Partial, serious, long-term
		[59]	Color distribution model	
		[60]	Deep feature and color histogram in adaptive mode	
	Redistribution	[61]	Particles generation are independently	Serious
		[62]	Redistribution based on region growth	
	Memory Mechanism	[63]	Memory-based state estimation scheme	Serious, long-term
		[64]	Adaptive update strategy with well model saved	
	Mark flag	[65]	Binary occlusion flag state representation for particles	Partial
	Deep learning	[66]	Observation model built by RBF neural network	Long-term

结合粒子滤波状态模型可靠估计与多样性粒子维护模块，可有效提升遮挡目标跟踪算法的鲁棒性。但目前粒子滤波的计算复杂度较高，且模型单一、粒子退化问题还未能完全解决，因此粒子滤波模型稳定性与实时性的提升仍是难点问题。

4　基于稳定时空信息的目标跟踪算法

部分遮挡时，建模局部空间信息可利用未被覆盖的可靠目标区域，缓解跟踪漂移问题；严重遮挡时，虽然大量目标特征信息因被覆盖而丢失，但跟踪序列中目标状态变化具有时间、空间上的连续性，建模上下文可在信息损失情况下推断目标状态。下文分述了基于局部空间信息、时间上下文信息与时空上下文信息三类代表性跟踪算法的特点及抗遮挡方案。

4.1　基于局部空间信息的目标跟踪算法

基于局部空间信息的目标跟踪算法主要包括基于关键点、分块目标表示的算法，此种算法可降低部分遮挡的不利影响，进行持续跟踪，这是因为：1）未被遮挡物覆盖的目标局部区域依旧保持稳定；2）目标局部区域间的相关运动有空间一致性，利用可靠局部与其他区域间关联，可灵活建模目标几何结构变化，抵抗遮挡物干扰，防止积累误差；3）目标

被遮挡后再次出现时，通过局部与全局空间关联，可快速恢复全局目标模型。

目标关键点间的相关运动具有空间一致性，该信息有利于部分遮挡等非刚性形变下目标模型的灵活重建。SPWT 算法[67]在特征点与光流点表示静态自适应目标对象的基础上，以最大相似度聚类模型跟踪形变对象，以特征点匹配的速度变化近似判定目标被遮挡程度。为避免跟踪漂移至不匹配关键点，利用几何一致性在被遮挡的目标区域添加虚拟特征点，重建当前帧关键点与遮挡目标的匹配，提升了算法抗遮挡性能。SAT 算法[32]设计了基于粒子滤波的结构感知跟踪器，检测与匹配确定性关键点，投票更新校正目标状态，因其外观模型结合了目标关键点的空间布局和内部结构特性，可提高部分遮挡时的跟踪稳定性。完全遮挡时，该算法会因找不到匹配特征而检测到全遮挡，并可在特征再次成功匹配时恢复跟踪。基于关键点的目标空间结构描述有稳定性与灵活性，可适用于部分遮挡、运动模糊等场景，但其信息表征方式具有稀疏性，在严重遮挡、尺度变化等场景下可能会忽略全局信息而丢失跟踪对象。

相比稀疏性的关键点表示，分块表示在遮挡场景下更加稳定，文献[68-70]提出自适应加权分块相关滤波器，增强模型对可靠区域的依赖性，抑制被遮挡部分和背景噪声，结合分块间空间约束关系，提高算法抵抗部分遮挡的能力。RTT 算法[68]以循环神经网络遍历多方向候选区域，获得每分块的置信度，其分块关联性可抵抗单方向遮挡等干扰，但其RNN结构与特征方面还可进一步提高。RPT 算法[71]提出了整合置信度与运动信息的粒子滤波跟踪框架，利用粒子表示不同对象的相对运动信息，以密集无监督方法将同质运动轨迹进行聚类分组，结合有效粒子更新策略，识别利用可靠的图像块进行鲁棒跟踪。针对局部检测框不利于严重遮挡后恢复的问题，文献[33，72]提出“分块-全局”的协同模型，稳定捕捉目标内部结构的同时关联全局搜索。DPCF 算法[33]以不同分块的相对排列构造可变形模型，协同耦合全局相关滤波器进行检测跟踪，可提高严重遮挡、形变、尺度变化等场景下的模型稳定性。

4.2 基于时间上下文信息的目标跟踪算法

与卷积神经网络擅长提取图像空间特征不同，循环神经网络(RNN)因其独有的自连接反馈结构而擅长捕获序列上下文信息，因此RNN被尝试运用于遮挡场景下的跟踪任务。传统RNN因梯度崩溃问题无法处理长序列，其变体长短期记忆网络(LSTM)与门控循环单元(GRU)，分别以遗忘门自连接单元、更新门结构对丢弃或传递的历史信息进行决策，可用于长时序列处理。表8为卷积神经网络、RNN结构及其跟踪应用的特性比较。

表8 CNN与RNN结构及跟踪应用特性比较

Table 8 Comparison of structure and tracking applications characteristics between CNN and RNN

Characteristic		Convolutional neural networks	Recurrent Neural Network	
			LSTM	GRU
Network	Composition	Convolutional layer, pooling layer, downsampling layer	Forgot gate	Update gate, reset gate
	Structural features	Local receptive field Weight sharing Spatiotemporal downsampling	Better effect	Higher efficiency
			Memory and feedback functions by current and historical states connection	
Tracking	Information	Spatial semantic information	Temporal information	
	Disadvantage	Sensitive to similar interference, lacking relevant information, may lose spatial details	High time consumption	

基于擅于捕获跟踪序列时间上下文信息的特性，RNN可构建状态模型预估，并恢复遮挡目标的轨迹。Deep tracking算法[73]基于传统贝叶斯跟踪框架，构建基于RNN的后验概率模型，引入马尔科夫性质的隐变量用于反映真实环境信息，学习部分观测序列数据到所有对象轨迹的映射关系，实现端到端的完整、无遮挡场景的预测估计，在目标表观的完整性因遮挡物受损时，该算法可恢复遮挡目标状态。但因实际场景较复杂，其适用性还有待证明。

针对目标模型易被遮挡样本污染的问题，文献[74-76]结合RNN隐藏层记忆反馈机制与CNN空间语义信息提取，构建基于上下文感知目标外观模型，可适应长期稳定的外观变化，缓解遮挡时的漂移问题。ROLO算法[74]提出了基于空间监督的循

环卷积神经网络，利用目标上下文位置信息与语义视觉特征，将 YOLO 网络扩展到时空域，实时回归预测卷积层与循环单元处的跟踪位置，同时依据 YOLO 检测框构造空间位置置信度热图，在遮挡时对 LTSM 网络进行分析调整，可解决严重遮挡与运动模糊的问题。Re3 算法[75]结合 CNN 高层语义特征，利用双层 LSTM 学习运动特征，回归目标框位置，可捕获、保留复杂的目标外观，不立即采用观察结果，避免对遮挡物误判。HART 算法[76]结合多层注意力机制降低背景干扰，利用 LSTM 隐藏层存储更新运动信息及关键性视觉特征，通过后续共享预测，粗略估计目标位置，通过数据驱动及存储器权衡目标时空外观信息，提高遮挡、频繁形变时跟踪鲁棒性。

因 LSTM 遗忘门与 GRU 更新门结构可丢弃无效信息，基于 RNN 捕获的上下文可忽略遮挡受损样本，本质上学习到了对遮挡样本的处理，可发挥抵抗遮挡相似干扰、保守适应目标外观变化的作用。但因时间上下文需状态存储模块，且 RNN 在跟踪领域的应用还不够成熟，基于 RNN 捕获时间上下文信息的跟踪算法还有较大提升空间。

4.3 基于时空上下文的目标跟踪算法

在跟踪序列中，部分遮挡可能导致目标外观模型退化，但遮挡目标局部与周围对象的空间上下文不会发生明显变化，利用将空间上下文建立时序关联所得的时空上下文信息，可提升部分遮挡、严重遮挡场景下的跟踪稳定性。

文献[77-80]通过时空上下文模型的空间信息，抑制背景相似物、遮挡物干扰，STT 算法[77]通过增量线性子空间方法，构建存储历史信息的全局上下文模型，以目标分块与周围关键点的外观差异建立局部空间上下文模型，并对关键的弱相关性关系进行多实例增强，提高分类器鲁棒性。STC 算法[78]基于贝叶斯框架，对目标及周围区域底层特征的统计相关性进行建模，在前一帧的目标位置处裁剪局部上下文区域，通过低通滤波构造空间密集上下文特征集，可抵抗部分遮挡、外观变化噪声干扰。为解决相关滤波跟踪器的余弦窗会削减图像块边缘信息，造成因背景信息不足而模型判别力下降的问题，算法 CACF[79]提出了可嵌入任何相关滤波器的 CA 模块，引入降低模板与背景的相关响应值的约束项，学习全局上下文信息。文献[80]将有价值的背景物体与结构信息编码为稠密局部化状态向量，并由卷积门循环单元模块更新，以帧间的密集运动关联映射进行传播，结合外观模型预测目标状态，提高了算法对遮挡的鲁棒性。但是目前时空上下文的应用还不够成熟，构建实际场景下概率模型的可靠性还有待提升。

5 对目标模型构建方案抗遮挡的有效性分析

为进一步分析遮挡等场景下，利用各类信息构建目标模型对提升算法性能的有效性，选择了基于手工外观特征的 CSK、KCF、CN、DSST 算法，基于融合外观特征的 SAMF、MvCFT、MOCA、Staple、C-COT、SANet 算法，基于融合外观运动特征的 FPRNet、SPWT 算法，基于局部空间信息的 RPT、SPWT、RTT、LGCF 算法，基于粒子滤波的 RPT 算法及基于上下文信息的 Re3、CFCA 算法进行比较。表 9 为数据集 OTB50 与 OTB100 上算法的目标模型特点、性能比较。其中 All、Occ 分别代表数据集所有序列及遮挡属性序列，评估指标均为 OPE (one-pass evaluation) 下的成功率、准确率和帧率；AF、G、C、MAF、DAF、DMF 与 RF 分别代表利用手工外观特征、梯度特征、颜色特征、多视角特征、深度外观特征、运动特征及 RNN 自编码特征的目标模型；PF 代表结合粒子滤波的目标模型；Kb、Pb、LGPb 分别代表基于关键点表示、分块表示、分块全局协同的目标模型；T-C、ST-C 分别代表基于时间上下文、时空上下文的目标模型。

由表 9 可知，有效特征信息方面，在保证算法实时性的基础上，选择适应场景的有效特征对提升算法性能十分关键。融合外观特征的算法性能优于单一外观特征，但实时性会下降；使用深度特征表征可大幅增强目标模型的准确性与鲁棒性，但时间成本消耗明显。结合运动特征与外观特征，可进一步提升算法在遮挡场景下的鲁棒性，但其他场景下，性能提升不明显。状态估计信息方面，结合粒子滤波状态估计的目标模型构建，可显著提高算法在遮挡场景下的鲁棒性，但其他场景下，性能提升效果不明显，实时性较差。目标时空信息方面，时间上下文利用对算法性能提升不明显。时空上下文是提升抗遮挡能力的有效方案之一，但其他场景下，算法性能提升不明显。关键点与分块表示利用目标局部空间信息的目标模型构建，有利于提高算法的鲁棒性与准确性，可提高算法抗遮挡性能。

表9 各类目标模型的代表跟踪器在OTB50及OTB100数据集上的性能比较

Table 9 Performance comparison of representative trackers with variety models on OTB50 and OTB100

Tracker	Object models	Success rate		Precision		FPS	PC(CPU,RAM,Nvidia GPU)	Test dataset
		All	Occ	All	Occ			
CSK	Grey	0.398		0.545		152	Intel i5-760 2.80 GHz CPU, 16 G RAM	
KCF	G	0.514	0.513	0.740	0.749	172	Intel i5-760 2.80 GHz CPU, 16 G RAM	
CN	C	0.444	0.428	0.635	0.629	105	Intel i5-760 2.80 GHz CPU, 16 G RAM	
SAMF	G+C	0.567	0.613	0.774	0.839	7	Intel i5-760 2.80 GHz CPU, 16 G RAM	
DSST	G	0.555	0.534	0.549		24	Intel-i5-4590 CPU, 8 G RAM	
MvCFT	MAF	0.532	0.555			25.5	Intel-i5-4590 CPU, 8 G RAM	
MOCA	G+C	0.569		0.824	0.877	16.5	Intel Xeon2 2.50 GHz CPU, 256 G RAM	OTB50
Staple	G+C	0.694		0.513		80	Intel Core i7-4790K 4.0 GHz CPU	
FPRNet	DMF, DAF	0.613	0.854			1	Intel Core i7-4790K 4.0 GHz CPU	
RPT	PF, Pb	0.576		0.812		4.15	Intel-i7-3770 3.4 GHz CPU, 16 G RAM	
SPWT	DMF, Kb		0.530		0.792	62.3	Intel i7-6700 CPU	
RTT	Pb, RF	0.588	0.827			3–4	2.80 GHz CPU, 16 G RAM	
Re3	DAF, T-C	0.422	0.390			150	Intel Xeon 2.20 GHz CPU, Nvidia Titan X	
KCF	G	0.476		0.693	0.625	172	Intel i7 3.7 GHz CPU, 12 GB RAM	
C-COT	DAF	0.673		0.902		0.3	Intel i7 3.7 GHz, GTX TITAN Z GPU	
SANet	DAF, RF	0.692		0.928		1	Intel i7 3.7 GHz, GTX TITAN Z GPU	
SAMF	G+C	0.535	0.529	0.743		16.8	Intel Xeon 2.6 GHz CPU, 256 G RAM	
SAMF-CA	ST-C	0.575	0.550	0.793		13	Intel Xeon 2.6 GHz CPU, 256 G RAM	
Staple	G+C	0.579	0.543	0.784		59.8	Intel Xeon 2.6 GHz CPU, 256 G RAM	OTB100
Staple-CA	ST-C	0.579	0.558	0.810		35.2	Intel Xeon 2.6 GHz CPU, 256 G RAM	
DSST	G	0.475		0.695	0.615	24	Intel i7 3.7 GHz CPU, 12 GB RAM	
LGCF	LGPb	0.585		0.782	0.719	8	Intel i7 3.7 GHz CPU, 12 GB RAM	
RPT	Kb	0.715		0.936		20	GeForce GTX 1080Ti GPU	

部分遮挡时，基于局部空间信息、时空上下文信息构建目标模型可有效提升性能，但对于严重、长时遮挡，上述方案不再适用。此时，结合上下文记忆，构建可靠状态估计模型与遮挡检测模块为可行方案。实际上，状态估计信息、局部空间信息、上下文信息是提升抗遮挡性能的辅助手段，这是因为判别对遮挡稳定的局部空间信息、上下文信息依旧依赖于相关区域的特征表示，卡尔曼滤波、粒子滤波法预估轨迹状态的可靠性也与观测样本质量紧密相关。因此提取强判别性、强鲁棒性的特征信息是提升遮挡等复杂场景下跟踪性能的根本方案。

6 总结与展望

遮挡是导致跟踪失败的常见因素之一，同时也是实际长时跟踪应用中不可避免的问题，基于有效信息构建抗遮挡目标模型，可提升遮挡场景下的跟踪算法性能。本文首先概述了目标跟踪任务的难点问题，剖析了遮挡影响跟踪性能的原因；其次基于目标模型构建利用的信息，将遮挡目标跟踪的代表算法分为三类，并分析介绍其抗遮挡思路方案、适用遮挡场景及优缺点；最后通过比较，对三类目标模型提高算法抗遮挡性能的有效性提出了进一步思考。

目前，已有先进算法可有效应对部分遮挡场景，但还无法实现严重遮挡、完全遮挡、长时遮挡等属性场景下的稳健跟踪，算法鲁棒性、准确性与实时性的平衡问题也未能得到解决。提升目标模型抗遮挡性能存在以下难点：1)作为提升抗遮挡性能的根本方案，学习目标本质语义信息极其关键，但目前加深提取特征的网络深度的性能已不能获得明显突破，反而带来了巨大时间消耗；2)人类视觉记忆信息有益于自动恢复受损样本进行场景理解，受其启发，引入上下文信息也是提升算法抗遮挡能力的有效方案之一，但目前构建复杂跟踪场景下的上下文概率模型缺乏可靠性，其理论还有待完善；3)人类视觉依赖于先验知识，在严重、长时遮挡中

可轻松识别定位目标,但基于模板记忆库的跟踪算法无法应对严重遮挡等场景,构建先验知识模型也会消耗较大内存。因此,学习本质信息的轻量化网络设计、场景上下文预测与仿生视觉机理应用是可运用于提升遮挡场景跟踪性能的重要发展方向。

参　考　文　献

[1] Li X, Zha Y F, Zhang T Z, et al. Survey of visual object tracking algorithms based on deep learning[J]. Journal of Image and Graphics, 2019, 24(12): 2057-2080.
李玺, 查宇飞, 张天柱, 等. 深度学习的目标跟踪算法综述[J]. 中国图象图形学报, 2019, 24(12): 2057-2080.

[2] Chen T. Moving objects tracking in aerial videos[D]. Xi'an: Northwestern Polytechnical University, 2018.
陈挺. 无人机对地运动目标跟踪技术研究[D]. 西安: 西北工业大学, 2018.

[3] Xue C, Zhu M, Liu C X. Review of tracking algorithms under occlusions[J]. Chinese Journal of Optics and Applied Optics, 2009, 2(5): 388-394.
薛陈, 朱明, 刘春香. 遮挡情况下目标跟踪算法综述[J]. 中国光学与应用光学, 2009, 2(5): 388-394.

[4] Meng L, Yang X. A survey of object tracking algorithms[J]. Acta Automatica Sinica, 2019, 45(7): 1244-1260.
孟琭, 杨旭. 目标跟踪算法综述[J]. 自动化学报, 2019, 45(7): 1244-1260.

[5] Ge B Y, Zuo X Z, Hu Y J. Review of visual object tracking technology[J]. Journal of Image and Graphics, 2018, 23(8): 1091-1107.
葛宝义, 左宪章, 胡永江. 视觉目标跟踪方法研究综述[J]. 中国图象图形学报, 2018, 23(8): 1091-1107.

[6] Marvasti-Zadeh S M, Cheng L, Ghanei-Yakhdan H, et al. Deep learning for visual tracking: a comprehensive survey[EB/OL]. (2019-12-02)[2021-01-05]. https://arxiv.org/abs/1912.00535.

[7] Wang N Y, Shi J P, Yeung D Y, et al. Understanding and diagnosing visual tracking systems[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3101-3109.

[8] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 2544-2550.

[9] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[M]//Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7575: 702-715.

[10] Danelljan M, Khan F S, Felsberg M, et al. Adaptive color attributes for real-time visual tracking [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 1090-1097.

[11] Vojir T, Noskova J, Matas J. Robust scale-adaptive mean-shift for tracking[J]. Pattern Recognition Letters, 2014, 49: 250-258.

[12] Possegger H, Mauthner T, Bischof H. In defense of color-based model-free tracking[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 2113-2120.

[13] Fan J Q, Song H H, Zhang K H, et al. Complementary tracking via dual color clustering and spatio-temporal regularized correlation learning[J]. IEEE Access, 2018, 6: 56526-56538.

[14] 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.

[15] Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filters for visual tracking[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1144-1152.

[16] Ma C, Yang X K, Zhang C Y, et al. Long-term correlation tracking[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5388-5396.

[17] Danelljan M, Häger G, Shahbaz Khan F, et al. Accurate scale estimation for robust visual tracking [C]//Proceedings of the British Machine Vision Conference 2014, September 1-5, 2014, Nottingham. London: BMVA Press, 2014.

[18] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409-1422.

[19] Zhu Z, Wu W, Zou W, et al. End-to-end flow correlation tracking with spatial-temporal attention [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 548-557.

[20] Danelljan M, Häger G, Khan F S, et al. Convolutional features for correlation filter based visual tracking[C]//2015 IEEE International Conference on Computer Vision Workshop (ICCVW), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 621-629.

[21] Valmadre J, Bertinetto L, Henriques J, et al. End-to-end representation learning for correlation filter based tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5000-5008.

[22] Liu M, Wu C D, Zhang Y Z. Motion vehicle tracking based on multi-resolution optical flow and multi-scale Harris corner detection[C]//2007 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 15-18, 2007, Sanya, China. New York: IEEE Press, 2007: 2032-2036.

[23] Wu Y, Li L F, Xiao Z S, et al. Optical flow motion tracking algorithm based on SIFT feature[J]. Computer Engineering and Applications, 2013, 49(15): 157-161.
吴垠, 李良福, 肖樟树, 等. 基于尺度不变特征的光流法目标跟踪技术研究[J]. 计算机工程与应用, 2013, 49(15): 157-161.

[24] Bouguet J Y. Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm[J]. Intel Corporation, 2001, 5(1-10): 4.

[25] Liu P P, King I, Lyu M R, et al. DDFlow: learning optical flow with unlabeled data distillation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 8770-8777.

[26] Hur J, Roth S. MirrorFlow: exploiting symmetries in joint optical flow and occlusion estimation[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 312-321.

[27] Li H Y, Bi D Y, Yang Y, et al. Research on visual tracking algorithm based on deep feature expression and learning[C]//The Thirty-Third {AAAI} Conference on Artificial Intelligence, {AAAI} 2019, January 27-February 1, 2019, Honolulu, Hawaii, USA. Menlo Park: AAAI Press, 2015: 8770-8777.
李寰宇, 毕笃彦, 杨源, 等. 基于深度特征表达与学习的视觉跟踪算法研究[J]. 电子与信息学报, 2015, 37(9): 2033-2039.

[28] Bhat G, Johnander J, Danelljan M, et al. Unveiling the power of deep tracking[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11206: 493-509.

[29] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration[M]//Agapito L, Bronstein M M, Rother C. Computer vision-ECCV 2014 workshops. Lecture notes in computer science. Cham: Springer, 2015, 8926: 254-265.

[30] Zhu G B, Wang J Q, Wu Y, et al. MC-HOG correlation tracking with saliency proposal[C]//Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA. Menlo Park: AAAI Press, 2016: 3690-3696.

[31] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: complementary learners for real-time tracking[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1401-1409.

[32] Bouachir W, Bilodeau G A. Structure-aware keypoint tracking for partial occlusion handling[C]//IEEE Winter Conference on Applications of Computer Vision, March 24-26, 2014, Steamboat Springs, CO, USA. New York: IEEE Press, 2014: 877-884.

[33] Akin O, Erdem E, Erdem A, et al. Deformable part-based tracking by coupled global and local correlation filters[J]. Journal of Visual Communication and Image Representation, 2016, 38: 763-774.

[34] Ma C Y, Yu G. An improved kernel correlation filter for occlusion target tracking[C]//2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), July 5-7, 2019, Xiamen, China. New York: IEEE Press, 2019: 674-678.

[35] Li X, Liu Q, He Z Y, et al. A multi-view model for visual tracking via correlation filters[J]. Knowledge-Based Systems, 2016, 113: 88-99.

[36] Liu H F, Sun C, Liang X L. Correlation-filter tracking algorithm with adaptive-feature fusion and anti-occlusion[J]. Laser & Optoelectronics Progress, 2020, 57(14): 141014.
刘海峰, 孙成, 梁星亮. 自适应特征融合与抗遮挡的

相关滤波跟踪算法[J]. 激光与光电子学进展, 2020, 57(14): 141014.

[37] Ma C, Huang J B, Yang X K, et al. Hierarchical convolutional features for visual tracking[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3074-3082.

[38] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9909: 472-488.

[39] Danelljan M, Bhat G, Khan F S, et al. ECO: efficient convolution operators for tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6931-6939.

[40] Li Y, Zhang Y F, Xu Y L, et al. Robust scale adaptive kernel correlation filter tracker with hierarchical convolutional features[J]. IEEE Signal Processing Letters, 2016, 23(8): 1136-1140.

[41] Liu P D, Yan X Y, Jiang Y, et al. Deep flow collaborative network for online visual tracking[C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 4-8, 2020, Barcelona, Spain. New York: IEEE Press, 2020: 2598-2602.

[42] Zhang P, Zhuo T, Huang W, et al. Online object tracking based on CNN with spatial-temporal saliency guided sampling[J]. Neurocomputing, 2017, 257: 115-127.

[43] Jiang X L, Zhen X T, Zhang B C, et al. Deep collaborative tracking networks[C]//British Machine Vision Conference 2018, September 3-6, 2018, Newcastle, UK. London: BMVA Press, 2018: 87.

[44] Fan H, Ling H B. SANet: structure-aware network for visual tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2217-2224.

[45] Ma D, Bu W, Wu X. Multi-scale recurrent tracking via pyramid recurrent network and optical flow[C]//British Machine Vision Conference 2018, September 3-6, 2018, Newcastle, UK. London: BMVC Press, 2018: 242.

[46] Hu P. Research on video object tracking with Kalman filter[D]. Chongqing: Chongqing University, 2010.
胡鹏. Kalman 滤波在视频目标跟踪中的应用研究[D]. 重庆: 重庆大学, 2010.

[47] Heimbach M, Ebadi K, Wood S. Resolving occlusion ambiguity by combining Kalman tracking with feature tracking for image sequences[C]//2017 51st Asilomar Conference on Signals, Systems, and Computers, October 29-November 1, 2017, Pacific Grove, CA, USA. New York: IEEE Press, 2017: 144-147.

[48] Heimbach M, Ebadi K, Wood S. Improving object tracking accuracy in video sequences subject to noise and occlusion impediments by combining feature tracking with Kalman filtering[C]//2018 52nd Asilomar Conference on Signals, Systems, and Computers, October 28-31, 2018, Pacific Grove, CA, USA. New York: IEEE Press, 2018: 1499-1502.

[49] Mehmood K, Jalil A, Ali A, et al. Context-aware and occlusion handling mechanism for online visual object tracking[J]. Electronics, 2020, 10(1): 43.

[50] Sun Y J, Zhang L Y, Yun X. Visual tracking algorithm based on region estimation and adaptive classification[J]. Laser & Optoelectronics Progress, 2019, 56(18): 181001.
孙彦景, 张丽颖, 云霄. 基于区域预估与自适应分类的视觉跟踪算法[J]. 激光与光电子学进展, 2019, 56(18): 181001.

[51] Li Y B, Jiu M Y, Sun Q, et al. An improved target tracking algorithm based on extended Kalman filter for UAV[C]//2018 IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP), August 5-8, 2018, Auckland, New Zealand. New York: IEEE Press, 2018: 435-437.

[52] Song J S, Hu G Q. Multi-feature visual tracking using adaptive unscented Kalman filtering[C]//2013 Sixth International Symposium on Computational Intelligence and Design, October 28-29, 2013, Hangzhou, China. New York: IEEE Press, 2013: 197-200.

[53] Ahmed E, Ahmad A, Hadhoud M. A robust framework for object tracking based on corrected background-weighted histogram mean shift and unscented Kalman filter[C]//International Image Processing, Applications and Systems Conference, November 5-7, 2014, Sfax, Tunisia. New York: IEEE Press, 2014: 14947318.

[54] Duan Z H, Cai Z X. Adaptive evolutionary particle filter based object tracking with occlusion handling

[C]//2009 Fifth International Conference on Natural Computation, August 14-16, 2009, Tianjian, China. New York: IEEE Press, 2009: 358-361.

[55] Liu K J, Liu B, Yu N H. A robust occlusion judgment scheme for target tracking under the framework of particle filter[M]//Zhang Y J. Image and graphics. Lecture notes in computer science. Cham: Springer, 2015, 9217: 423-433.

[56] Li L Q, Yan M Y. High-order cumulant-based particle filtering algorithm for pedestrian object tracking[C]//2018 14th IEEE International Conference on Signal Processing (ICSP), August 12-16, 2018, Beijing, China. New York: IEEE Press, 2018: 304-307.

[57] Wen W, Wu L Z. Tracking algorithm of improved spatio-temporal context with particle filter[C]//2017 3rd IEEE International Conference on Computer and Communications (ICCC), December 13-16, 2017, Chengdu, China. New York: IEEE Press, 2017: 1549-1553.

[58] Guo C J, Lu Y, Fang X Z, et al. Multiple likelihoods and state noises based particle filter for long-lived full occlusion handling[C]//2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), September 23-25, 2010, Chengdu, China. New York: IEEE Press, 2010: 11582115.

[59] Zhang T, Fei S M. Improved particle filter for object tracking[C]//2011 Chinese Control and Decision Conference (CCDC), May 23-25, 2011, Mianyang, China. New York: IEEE Press, 2011: 3586-3590.

[60] Qi Y J, Wang Y J, Liu Y C. Object tracking based on deep CNN feature and color feature[C]//2018 14th IEEE International Conference on Signal Processing (ICSP), August 12-16, 2018, Beijing, China. New York: IEEE Press, 2018: 469-473.

[61] Daneshyar M A, Nahvi M. Improvement of moving objects tracking via modified particle distribution in particle filter algorithm[C]//2015 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA), March 11-12, 2015, Rasht, Iran. New York: IEEE Press, 2015: 15323867.

[62] Zhang S P, Wu J P, Tian Y, et al. Robust visual tracking based on occlusion detection and particle redistribution [C]//Proceedings of the Second International Conference on Internet Multimedia Computing and Service-ICIMCS' 10, December 30-31, 2010. Harbin, China. New York: ACM Press, 2010: 159-162.

[63] Qi Y J, Wang Y J. Memory-based state estimation for handling occlusion during object tracking by particle filter[C]//International Conference on Information Science and Technology, March 26-28, 2011, Nanjing, China. New York: IEEE Press, 2011: 953-957.

[64] Zhai Y Y, Song P, Mou Z L, et al. Occlusion-aware correlation particle filter target tracking based on RGBD data[J]. IEEE Access, 2018, 6: 50752-50764.

[65] Meshgi K, Maeda S I, Oba S, et al. An occlusion-aware particle filter tracker to handle complex and persistent occlusions[J]. Computer Vision and Image Understanding, 2016, 150: 81-94.

[66] Han Y X, Ding G Y. Occlusion target tracking based on particle filter and neural network[J]. Computer Integrated Manufacturing Systems, 2020, 26(12): 3229-3235.
韩宇星, 丁刚毅. 基于粒子滤波与神经网络的目标遮挡跟踪[J]. 计算机集成制造系统, 2020, 26(12): 3229-3235.

[67] Cao W M, Li Y H, He Z H, et al. Supplementary virtual keypoints of weight-based correspondences for occluded object tracking[J]. IEEE Access, 2018, 6: 9140-9146.

[68] Cui Z, Xiao S T, Feng J S, et al. Recurrently target-attending tracking[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1449-1458.

[69] Du X F, Dore A, Stoyanov D. Patch-based adaptive weighting with segmentation and scale (PAWSS) for visual tracking[EB/OL]. (2017-08-03) [2021-02-05]. https://arxiv.org/abs/1708.01179.

[70] Yang J F, Zhang J P. Long time target tracking based on kernel correlation filtering[J]. Laser & Optoelectronics Progress, 2019, 56(2): 021502.
杨剑锋, 张建鹏. 基于核相关滤波的长时间目标跟踪[J]. 激光与光电子学进展, 2019, 56(2): 021502.

[71] Li Y, Zhu J K, Hoi S C H. Reliable Patch Trackers: Robust visual tracking by exploiting reliable patches [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 353-361.

[72] Fan H, Xiang J. Robust visual tracking via local-global correlation filter[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California,

USA. Menlo Park: AAAI Press, 2017: 4025-4031.

[73] Ondruska P, Posner I. Deep tracking: seeing beyond seeing using recurrent neural networks[C]//Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA. Menlo Park: AAAI Press, 2016: 3361-3368.

[74] Ning G H, Zhang Z, Huang C, et al. Spatially supervised recurrent convolutional neural networks for visual object tracking[C]//2017 IEEE International Symposium on Circuits and Systems (ISCAS), May 28-31, 2017, Baltimore, MD, USA. New York: IEEE Press, 2017: 17208133.

[75] Gordon D, Farhadi A, Fox D. Re^3: real-time recurrent regression networks for visual tracking of generic objects[J]. IEEE Robotics and Automation Letters, 2018, 3(2): 788-795.

[76] Kosiorek A R, Bewley A, Posner I. Hierarchical attentive recurrent tracking[EB/OL]. (2017-06-28)[2021-01-05]. https://arxiv.org/abs/1706.09262.

[77] Wen L Y, Cai Z W, Lei Z, et al. Online spatio-temporal structural context learning for visual tracking[M]//Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7575: 716-729.

[78] Zhang K H, Zhang L, Liu Q S, et al. Fast visual tracking via dense spatio-temporal context learning [M]//Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8693: 127-141.

[79] Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1387-1395.

[80] Bhat G, Danelljan M, van Gool L, et al. Know your surroundings: exploiting scene information for object tracking[M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12368: 205-221.