

基于协方差矩阵的多特征融合跟踪算法

金泽芬, 侯志强, 余旺盛, 王 鑫

中国人民解放军空军工程大学信息与导航学院, 陕西 西安 710077

摘要 为提高视觉目标跟踪算法的稳健性, 提出一种基于协方差矩阵的多特征融合跟踪算法。在量子遗传算法框架下, 采用区域协方差描述子对颜色、边缘和纹理特征进行融合, 并采用一种快速协方差交叉算法进行模型更新。该算法综合利用了区域协方差描述子维数较低, 量子遗传算法收敛速度快且全局寻优能力强和快速协方差交叉算法快速计算的特点, 能极大地提高了融合、匹配与更新过程的运算效率, 实现了快速有效的多特征融合跟踪。实验结果表明, 该算法能够有效应对遮挡、旋转、形变和运动模糊等多种复杂变化的干扰, 实现对目标的快速稳健跟踪。

关键词 图像处理; 视觉跟踪; 特征融合; 协方差矩阵; 边缘; 局部二值模式

中图分类号 TP391.4 **文献标识码** A

doi: 10.3788/AOS201737.0915005

Multiple Feature Fusion based on Covariance Matrix for Visual Tracking

Jin Zefenfen, Hou Zhiqiang, Yu Wangsheng, Wang Xin

Information and Navigation College, Air Force Engineering University of PLA, Xi'an, Shaanxi 710077, China

Abstract In order to improve the robustness of visual target tracking algorithm, a multiple feature fusion tracking algorithm is proposed based on covariance matrix. Under the framework of quantum genetic algorithm, the region covariance descriptor is used to fuse color, edge, and texture features. A fast covariance intersection algorithm is adopted to update the model. The proposed algorithm makes the most use of low dimension of the region covariance descriptor, fast convergence and strong global search ability of the quantum genetic algorithm, and fast calculation ability of the fast covariance intersection algorithm, which greatly improves the efficiency of fusing, matching and updating process, and effectively realizes fast and efficient multi-feature fusion tracking. Experimental results show that the proposed algorithm can effectively cope with the interference, such as occlusion, rotation, deformation and motion blur, and achieve fast and robust target tracking.

Key words image processing; visual tracking; feature fusion; covariance matrix; edge; local binary pattern

OCIS codes 100.4999; 100.2960; 110.4155

1 引 言

视觉跟踪是计算机视觉领域的一个重要课题, 已广泛应用于视频监控^[1]、人机交互^[2]和增强现实^[3]等领域。近年来, 专家学者们针对该课题进行了大量研究, 提出了许多优秀的跟踪算法。实际跟踪中复杂的场景变化难以预测, 容易出现局部或全局遮挡、光照变化、姿态变化、尺度变化等多种情况^[4], 实现稳健跟踪仍然具有极大的挑战性。

视觉跟踪的主要模块包括目标表示、搜索策略和模型更新, 其中: 目标表示是保证算法实现有效跟踪的关键; 搜索策略主要决定跟踪的实时性; 模型更新则能够提高跟踪的稳健性。因此, 跟踪算法的研究主要从这三个模块出发, 并以有效的方法解决实际跟踪中的问题。

在目标表示上, 单特征描述子因计算简便而被广泛应用, 如灰度、颜色、边缘、Haar-like 以及纹理特征等。然而, 单特征描述子通常不足以应对背景和目标的复杂变化, 多特征描述子则能够提供更充分的目

收稿日期: 2017-04-14; **收到修改稿日期:** 2017-05-15

基金项目: 国家自然科学基金(61473309, 41601436)、陕西省自然科学基金基础研究计划项目(2016JM6050)

作者简介: 金泽芬(1993—), 女, 硕士研究生, 主要从事视觉目标跟踪方面的研究。E-mail: christine123456@163.com

导师简介: 侯志强(1973—), 男, 博士, 教授, 主要从事多传感器信息融合、图像处理、模式识别等方面的研究。

E-mail: hou-zhq@sohu.com

标信息,提高跟踪性能。过去几年,一些多特征融合跟踪的算法相继被提出,通过考虑不同特征的优缺点,融合具有互补性的特征表示目标^[5-7]。常见的融合过程是在跟踪中动态调整特征权值^[8-10],使在某一场景中具有优势的特征发挥更大作用。但实时有效地识别场景并自动筛选出真正有利的特征十分困难。在众多融合方法中,区域协方差描述子^[11]是一种优秀的目标描述模型,其能够在不计算特征权重的情况下自然地融合多个特征,被广泛运用在图像匹配^[12]、图像分割^[13]和目标识别^[14]等领域。本文采用区域协方差描述子对颜色、边缘和纹理特征进行融合,更准确、更丰富地应用目标信息,在跟踪中提高跟踪结果的稳健性。

在搜索策略上,量子遗传算法(QGA)具有收敛速度快、全局寻优能力强的特点,其可以提高跟踪中的搜索和匹配效率。近年来 QGA 在飞机、车辆、机器人和巡航导弹等路径规划^[15],网络调度分配和覆盖^[16],优化神经网络^[17]等领域得以广泛研究。另外,QGA 在图像去噪^[18]、分割^[19]和配准^[20]等方面均具有优势,但目前尚未有应用于视觉跟踪的研究。

在模型更新上,考虑到目标模型为协方差矩阵,应当结合协方差矩阵的特点采用有效方式进行模型更新。考虑到协方差交叉算法^[21]在数据拟合问题中的优化效果,且相较于传统协方差交叉算法,快速协方差交叉算法计算量小,能够极大提高算法效率,因此,本文采用一种快速协方差交叉算法^[22]进行模型更新,最大化保留模型中的有效信息。

综上,本文的主要工作包括:采用区域协方差描述子对颜色、边缘和纹理特征进行融合,以此进行目标描述;采用 QGA 作为搜索策略;采用快速协方差交叉算法进行模型更新,提高算法的跟踪效率;提出一种基于协方差矩阵的多特征融合跟踪算法;在 CVPR 2013 数据平台^[23]上进行 51 组视频的实验,并与近年的优秀算法进行比较,定性和定量地验证本文算法的优越性。

2 相关基础理论

2.1 区域协方差描述子

2.1.1 特征协方差矩阵

区域协方差描述子对图像区域具有很强的描述能力,它包含不同特征的相关性,并在协方差计算的求平均过程中起到去噪作用。其计算步骤如下。

1) 计算图像 I 的特征矩阵

$$\mathbf{F}(x, y) = \boldsymbol{\phi}(I, x, y), \quad (1)$$

式中 $\mathbf{F}(x, y)$ 的维度为 $W \times H \times d$, W 、 H 分别为图像的宽和高, d 为特征的种类数(本文提取的特征为颜色、边缘及纹理信息)。

2) 计算特征的协方差矩阵

$$\mathbf{C}_1 = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{z}_k - \boldsymbol{\mu})(\mathbf{z}_k - \boldsymbol{\mu})^T, \quad (2)$$

式中 n 为图像 I 内像素点个数, \mathbf{z}_k 为第 k 个像素点对应的特征向量(d 维), $\boldsymbol{\mu}$ 为特征的均值(d 维)。

2.1.2 协方差矩阵的距离

在跟踪过程中,通过匹配目标模板和搜索区域的相似度,将相似度最高的区域视为目标区域。计算两个协方差矩阵差异度的距离公式为^[11]

$$\rho(\mathbf{C}_1, \mathbf{C}_2) = \sqrt{\sum_{i=1}^n \ln^2[\lambda_i(\mathbf{C}_1, \mathbf{C}_2)]}, \quad (3)$$

式中 $\lambda_i(\mathbf{C}_1, \mathbf{C}_2)$ 为 \mathbf{C}_1 、 \mathbf{C}_2 的广义特征值, $i=1, \dots, n$ 。

2.1.3 快速协方差

为提高算法速度,采用积分图方法计算特征矩阵的协方差^[11]。在特征协方差矩阵中,每个点 (i, j) 对应的值为

$$\mathbf{C}_1(i, j) = \frac{1}{n-1} \sum_{k=1}^n [\mathbf{z}_k(i) - \boldsymbol{\mu}(i)][\mathbf{z}_k(j) - \boldsymbol{\mu}(j)], \quad i, j = 1, \dots, d. \quad (4)$$

根据均值的一般定义,将(4)式变形,得到

$$C_1(i, j) = \frac{1}{n-1} \left[\sum_{k=1}^n z_k(i) z_k(j) - \frac{1}{n} \sum_{k=1}^n z_k(i) \sum_{k=1}^n z_k(j) \right]. \quad (5)$$

将整个图像的特征矩阵和特征矩阵的平方矩阵分别表示为

$$P(x', y', i) = \sum_{x < x', y < y'} F(x, y, i), i = 1, \dots, d, \quad (6)$$

$$Q(x', y', i, j) = \sum_{x < x', y < y'} F(x, y, i) \cdot F(x, y, j), i = 1, \dots, d. \quad (7)$$

对应于每一个像素点,(6)式和(7)式可分别表示为

$$P_{x,y} = [P(x, y, 1) \quad \dots \quad P(x, y, d)]^T, \quad (8)$$

$$Q_{x,y} = \begin{bmatrix} Q(x, y, 1, 1) & \dots & Q(x, y, 1, d) \\ \vdots & & \vdots \\ Q(x, y, d, 1) & \dots & Q(x, y, d, d) \end{bmatrix}. \quad (9)$$

令 $I(x', y'; x'', y'')$ 为一矩形区域, (x', y') 为左上角坐标, (x'', y'') 为右下角坐标, 区域内每一点为一个维度为 d 的向量, 如图 1 所示^[24]。

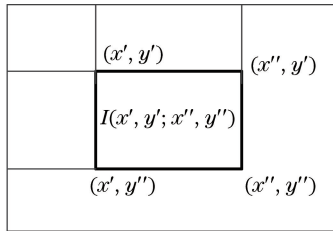


图 1 积分图像

Fig. 1 Integral image

利用积分图 $P_{x,y}$ 和 $Q_{x,y}$, 将图像 I 的协方差矩阵简化为

$$C_1(x', y'; x'', y'') = \frac{1}{n-1} \left[Q_{x'',y''} + Q_{x',y'} - Q_{x',y''} - Q_{x'',y'} - \frac{1}{n} (P_{x'',y''} + P_{x',y'} - P_{x',y''} - P_{x'',y'}) (P_{x'',y''} + P_{x',y'} - P_{x',y''} - P_{x'',y'})^T \right], \quad (10)$$

由此, 协方差矩阵的计算量大幅降低。

2.2 QGA

QGA 的主要搜索过程为: 在搜索范围内随机产生一个初始种群, 用目标函数计算种群内每一个体的适应度, 保留适应度最大的个体, 并用该个体产生量子旋转门对种群进行更新, 进入下一次迭代。迭代过程中, 不断用适应度更高的新个体代替旧个体以更新量子旋转门, 进而更新种群。当迭代次数达到设定值时, 输出最终的适应度最高的个体作为搜索结果。

QGA 包含一个由量子染色体组成的量子种群 $Q(t) = \{q_1^t, q_2^t, \dots, q_N^t\}$, N 为种群的大小, t 为进化代数, q_j^t 为一条量子染色体, 可定义为

$$q_j^t = \begin{bmatrix} \alpha_1^t & \alpha_2^t & \dots & \alpha_m^t \\ \beta_1^t & \beta_2^t & \dots & \beta_m^t \end{bmatrix}, j = 1, 2, \dots, N, \quad (11)$$

式中 m 为量子位数, 表示量子染色体的长度。

在种群更新的过程中, 量子旋转门为

$$U(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad (12)$$

式中 θ 为旋转角, 其大小将影响算法的收敛过程, 其产生规则如表 1 所示^[25]。

表 1 中 $f(x)$ 为计算适应度的函数, θ_i 为染色体第 i 个量子位的旋转角, $s(\alpha_i, \beta_i)$ 表示 θ_i 的符号 (α_i, β_i 为第 i 个量子位的概率幅), b_i 和 x_i 分别为最优解和当前解中的第 i 位基因。以表 1 中的最后一行为例, 当 x_i

和 b_i 均为 1, 且当前 x 的适应度大于最优解 b 的适应度时, 若 $\alpha_i\beta_i > 0$, 则 $\theta_i = 0.025\pi$ 。 θ_i 取值过大时容易导致算法陷入局部收敛, 取值过小时容易导致收敛速度过慢。全局最优解的收敛方向由 $s(\alpha_i\beta_i)$ 决定。

表 1 量子旋转门旋转角生成规则

Table 1 Generative rules of rotation angle of quantum rotation gate

x_i	b_i	$f(x) \geq f(b)$	$\Delta\theta_i$	$s(\alpha_i\beta_i)$			
				$\alpha_i\beta_i > 0$	$\alpha_i\beta_i < 0$	$\alpha_i = 0$	$\beta_i = 0$
0	0	False	0	0	0	0	0
0	0	True	0	0	0	0	0
0	1	False	0	0	0	0	0
0	1	True	0.05π	-1	+1	± 1	0
1	0	False	0.01π	-1	+1	± 1	0
1	0	True	0.025π	+1	-1	0	± 1
1	1	False	0.005π	+1	-1	0	± 1
1	1	True	0.025π	+1	-1	0	± 1

种群更新过程可表示为

$$\begin{bmatrix} \alpha'_i \\ \beta'_i \end{bmatrix} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \quad (13)$$

3 本文跟踪算法

在上述基本理论的基础上, 提出一种基于协方差矩阵的多特征融合跟踪算法, 利用区域协方差描述子进行目标表示, 并借助 QGA 在候选区域中搜索目标。此外, 该算法采用协方差交叉算法进行模型更新。

3.1 目标表示

3.1.1 特征提取与融合

在图像处理领域中, 颜色、纹理和边缘特征均为广泛使用的特征。其中, 颜色特征描述图像区域所包含事物的表观性质, 对图像区域的方向、缩放等变化不敏感, 但容易受光照变化的影响; 纹理特征通过对包含多个像素点的区域进行统计得到, 因而不会因为局部偏差导致匹配失败, 且对光照不敏感, 在模式匹配中具有较大的优越性; 边缘特征提取图像的边缘结构和轮廓信息, 广泛存在于物体与背景、物体与物体之间, 为目标描述提供了有效的信息。

因此, 利用上述三种特征, 对目标区域 I 内的每个像素构造 6 维特征向量

$$\mathbf{F}(x, y) = [R(x, y), G(x, y), B(x, y), G_r(x, y), \theta(x, y), L_{bp}(x, y)], \quad (14)$$

式中 $R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$ 分别为图像中对应像素点的 RGB 颜色分量, $G_r(x, y)$ 和 $\theta(x, y)$ 分别为对图像像素点进行方向提取得到的梯度幅值和方向角, $L_{bp}(x, y)$ 为对应像素点的局部二值模式 (LBP) 特征值。对于边缘特征的提取, 采用梯度算子分别计算每个像素点的梯度幅值 $G_r(i, j)$ 和方向角 $\theta(i, j)$ 。纹理特征的提取采用基础的 3×3 LBP 算子实现。

得到目标区域中每一像素的特征向量 $\mathbf{F}(x, y)$ 后, 代入 (1) 式, 通过 2.1 节的方法计算该区域协方差矩阵 \mathbf{C}_1 , 并将其作为目标表示。

3.1.2 特征对比分析

为比较三种特征的特性并验证融合的有效性, 对比三种特征在任意帧中分别进行匹配与融合再进行匹配跟踪的结果。以 Basketball 序列的第 649 帧和 Liquor 序列的第 873 帧为例, 如图 2 所示, 可以看出, 颜色特征容易被颜色相似的干扰项影响; 边缘特征容易倾向于边界明显、前景与背景对比度较大的区域; LBP 特征则容易倾向于纹理信息丰富的区域。通过协方差区域描述子进行特征融合, 能够综合考虑三种特征携带的目标信息, 得到更准确的目标位置。

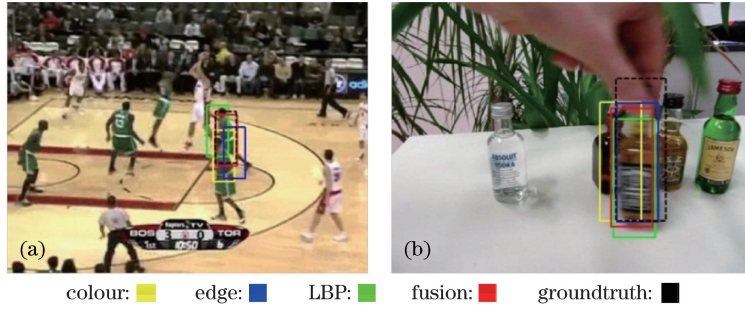


图 2 不同特征跟踪结果比较。(a) Basketball 序列第 649 帧；(b) Liquor 序列第 873 帧

Fig. 2 Comparison of tracking results of different features. (a) The 649th frame of Basketball sequence; (b) the 873rd frame of Liquor sequence

3.2 搜索策略

在跟踪中,搜索的速度,并在一定程度上影响匹配结果的精度。考虑到 QGA 具有收敛速度快、全局寻优能力强等特点,采用该算法作为搜索策略,可以提高整体跟踪效率。基于 QGA 设计跟踪过程如图 3 所示,红色框为上一帧跟踪得到的目标区域。以该区域的中心点为中心,以 $32 \times 32 (2^5 \times 2^5)$ 的矩形框为搜索范围,范围内的像素点位置对应种群中的个体。在搜索范围内随机生成初始种群后,以个体为中心的跟踪框区域的特征协方差与目标模板之间的相似度作为适应度进行计算。通过迭代,输出搜索范围内适应度最大的点,则其对应的跟踪框区域为跟踪结果。

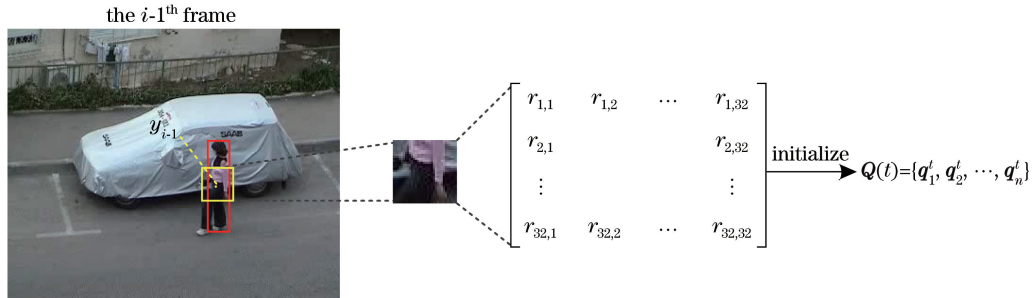


图 3 第 i 帧的种群生成方式

Fig. 3 Population generation method of the i^{th} frame

两帧之间目标的移动范围有限,对数据平台内的视频数据进行统计分析可知,当搜索范围设定为 32×32 的矩形框时,新一帧的中心位置能够将目标包含在内。

此外,在 QGA 中,参数设定至今没有理论判断依据,主要依靠多次运行和调整。本文实验通过反复测试,设定初始种群数为 30,繁衍代数为 10,染色体位数为 5 时,能够获得较好的跟踪效果和较高的跟踪效率。

3.3 模型更新

由于在跟踪的过程中,目标和背景很可能持续发生变化,如果不采用有效的方式对目标模板进行更新,容易在长时间跟踪中丢失目标。针对目标模板为特征协方差矩阵的情况,为最大化地保留协方差矩阵中的有效目标信息,采用协方差交叉算法进行模型更新。

协方差交叉算法被广泛应用于数据拟合问题,在被估计矩阵之间的相关性未知时,能够进行融合估计优化^[21]。其通过对不同估计量进行拟合,得到最优的融合估计量。为提高更新效率,本文采用文献[22]中的快速协方差交叉算法进行模型更新。具体过程如下:

在初始化阶段,以第一帧的目标区域计算得到的特征协方差矩阵 C_{tem} 作为目标模板;在后续跟踪中,当特征协方差矩阵匹配的差异度小于设定的更新阈值 T 时,利用快速协方差交叉算法将该跟踪结果的特征协方差矩阵 C_{best} 与目标模板协方差矩阵 C_{tem} 进行融合,得到新的目标模板 $C_{\text{new}} \rightarrow C_{\text{tem}}$ 。其中

$$\epsilon_1 = \frac{1}{\text{tr}(C_{\text{best}}^{-1})}, \epsilon_2 = \frac{1}{\text{tr}(C_{\text{tem}}^{-1})}, \quad (15)$$

$$\omega_1 = \frac{\epsilon_1}{\epsilon_1 + \epsilon_2}, \omega_2 = \frac{\epsilon_2}{\epsilon_1 + \epsilon_2}, \quad (16)$$

$$\mathbf{C}_{\text{new}}^{-1} = \omega_1 \mathbf{C}_{\text{best}}^{-1} + \omega_2 \mathbf{C}_{\text{tem}}^{-1}. \quad (17)$$

通过实验观察和比较可知,当更新阈值为 1.4~1.5 时能够取得相对较好的效果,该算法的匹配结果较准确,此时进行更新能够有效避免新的误差引入,从而达到提升跟踪性能的目的。因此,本文算法的更新阈值 T 设定为 1.45。

4 算法流程

基于协方差矩阵的多特征融合跟踪算法的流程总结如下:

读入视频序列,选取目标窗口以获取目标初始位置 y_0 及跟踪框大小。

根据 2.1、3.1 两小节中的计算方式,分别提取各特征并计算目标模板 \mathbf{C}_{tem} 。

for $i = 1, \dots, N_{\text{frame}}$ **do**

 初始化种群 \mathbf{Q} ;

for $t = 1, \dots, N_{\text{gen}}$ **do**

 对种群 $\mathbf{Q}(t)$ 内个体进行测量得到相应解集合 $P(t)$;

 计算以每一个体 Ind 为中心的跟踪框区域的特征协方差矩阵 \mathbf{C}_{Ind} ;

 根据(3)式计算集合内个体的适应度,并记录最优适应度 ρ_{max} 及相应解 y ;

 根据 y 按照表 1 中规则生成量子旋转门 $\mathbf{U}(t)$,并用 $\mathbf{U}(t)$ 对种群更新为 $\mathbf{Q}(t+1)$;

end

 输出最终的 y 作为第 i 帧的跟踪结果。

If $\rho_{\text{max}} > T$ **do**

 计算跟踪结果 y 相应的特征协方差矩阵 \mathbf{C}_{best} ;

 用快速协方差交叉算法融合 \mathbf{C}_{best} 和 \mathbf{C}_{tem} 得到新的目标模板。

end

end

其中, N_{frame} 为视频帧数, N_{gen} 为繁衍代数, T 为更新阈值。

5 实验分析

为验证算法的有效性,在 Matlab 2009a 平台上对算法进行编程实现。实验平台配置为 3.70 GHz 的 CPU(Intel(R) Core(TM) i3-4170),内存为 4 GB 的计算机。进行 51 组视频测试(来源于文献[23]提供的数据库),对算法在不同场景环境下的性能进行考量。采用的对比算法为深度学习跟踪(DLT)算法^[26]、跟踪、学习与检测(TLD)算法^[27]、基于核的结构输出跟踪(Struck)算法^[28]、基于自适应结构局部稀疏表达模型的跟踪算法(ASLA)^[29]、精确尺度估计的稳健视觉跟踪(DSST)算法^[30]、结构稀疏跟踪(SST)算法^[31]以及在线多实例学习的视觉跟踪(MIL)算法^[32]。其中,DLT 算法为基于深度学习的跟踪算法;TLD 算法结合了检测与跟踪;Struck 算法采用结构输出 SVM 学习方法;ASLA 采用局部稀疏表示和相似性匹配的方式跟踪;DSST 算法是基于相关滤波的跟踪方法;SST 算法结合了全局稀疏、局部稀疏和联合稀疏三种外观模型;MIL 算法则采用了多示例学习方法。以上算法均为近年来提出的跟踪性能较好的算法。

5.1 定性分析

图 4 给出了跟踪实验的部分具有代表性的结果,体现算法在不同场景下的跟踪效果。从 6 个角度进行具体分析:

1) 相似物干扰:在 Basketball 序列中,目标周围不断出现的身着相似球衣的球员造成了干扰,从图 4(a)中可以看出,本文算法能够始终准确地跟上目标;Liquor 序列中,目标附近的瓶子为干扰项,从图 4(b)中可以看出,本文算法能够较好地应对相似物干扰,准确地跟踪目标。

2) 尺度变化:在 Skiing 序列中,目标具有明显的旋转和尺度变化,且运动速度较快。从图 4(h)中可以看出,在第 25 帧和第 34 帧中,仅本文算法跟上了目标,其余算法均将目标丢失;在 Lemming 序列中,目标同样出现旋转和尺度变化,且从图 4(i)中可以看出,本文算法较好地跟上了目标。

3) 遮挡:在 David3 序列中,目标分别经历了交通指示牌和树干的遮挡。在所有算法中,只有本文算法能够全程准确地跟上目标;在 Jogging1 序列中,目标经历了一次全遮挡,本文算法在目标重新出现的瞬间立刻捕捉到目标,并继续跟踪。

4) 运动模糊:在 Jumping 序列和 Football1 序列中,目标均因剧烈运动出现模糊。从图 4(d)和图 4(e)中可以看出,本文算法能够很好地应对运动模糊的情况,准确跟踪目标。

5) 局部非刚性形变:在 Bolt 序列中,目标在运动的过程中不断发生局部非刚性形变。只有本文算法能够完整地跟踪目标,体现出其优越性。



图 4 8 种不同算法的定性比较。(a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

Fig. 4 Qualitative comparison of eight tracking algorithms. (a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

6) 背景扰乱:在 Matrix 序列中,背景为下雨的夜晚,整体颜色较暗且背景信息丰富。在此背景下,本文跟踪算法体现出比其他算法更优越的性能,较好地完成了跟踪任务;在 Ironman 和 MotorRolling 序列中,背景十分复杂,本文算法均呈现出良好的跟踪性能。

5.2 定量分析

为进行定量分析算法,引入中心位置误差曲线、覆盖率曲线、精度曲线和成功率曲线 4 种评价指标进行衡量。其中,中心位置误差为跟踪结果的中心位置与真实中心位置(ground truth,由数据平台给出)之间的欧氏距离;覆盖率为跟踪结果区域 r_t 与真实位置区域 r_a 的交集与并集之比, $S = |r_t \cap r_a| / |r_t \cup r_a|$;精度取决于跟踪结果的平均中心位置误差;成功率取决于跟踪结果的覆盖率,具体计算方式见文献[23]。平均中心位置误差与精度均值越小,表明算法效果越好。覆盖率和成功率则是值越大,表明算法效果越好。

图 5 和图 6 分别给出了图 4 所示视频的中心位置误差曲线和覆盖率曲线,其数值比较结果见表 2 和表 3。

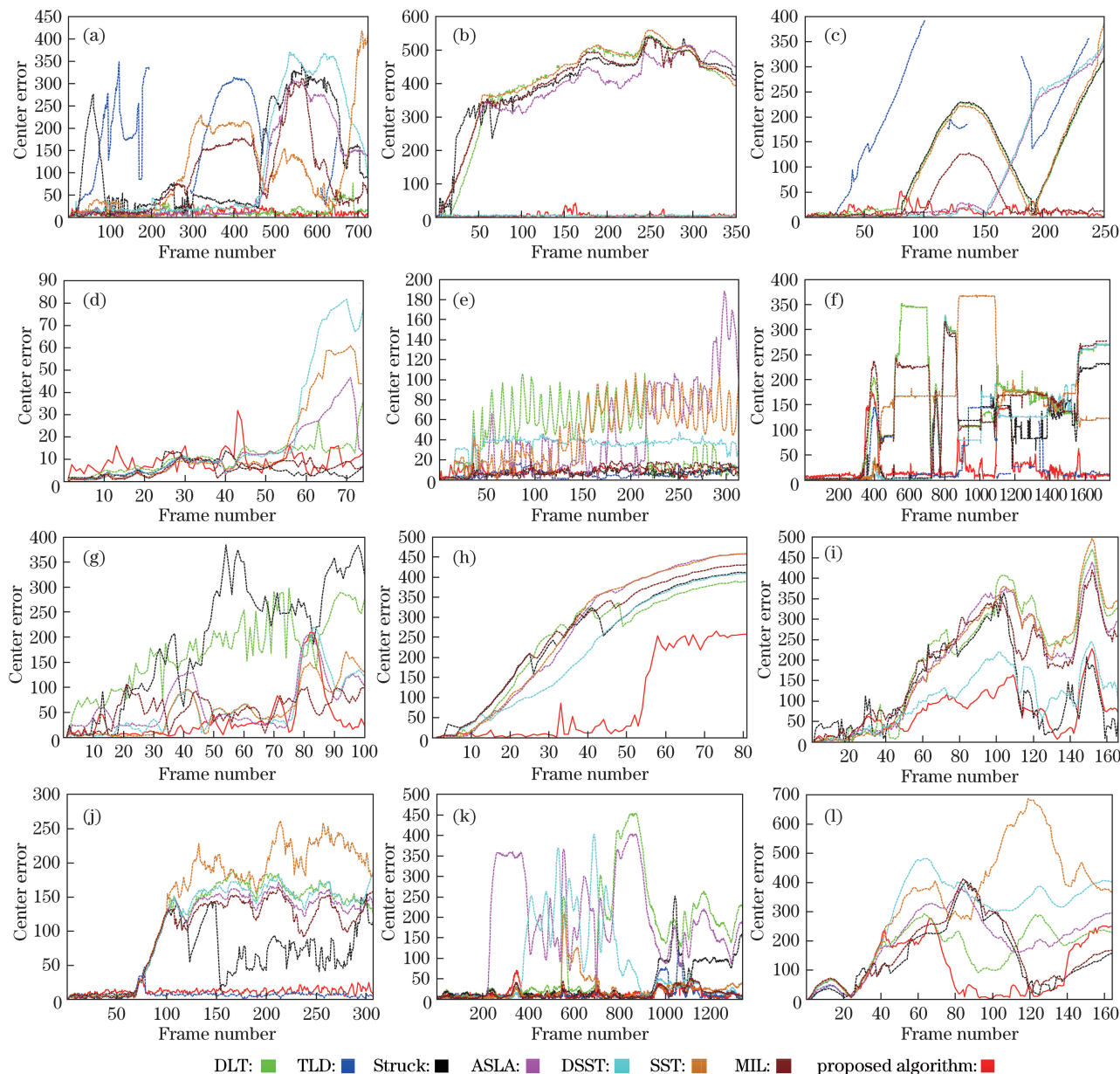


图 5 中心位置误差曲线。(a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

Fig. 5 Center position error curves. (a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

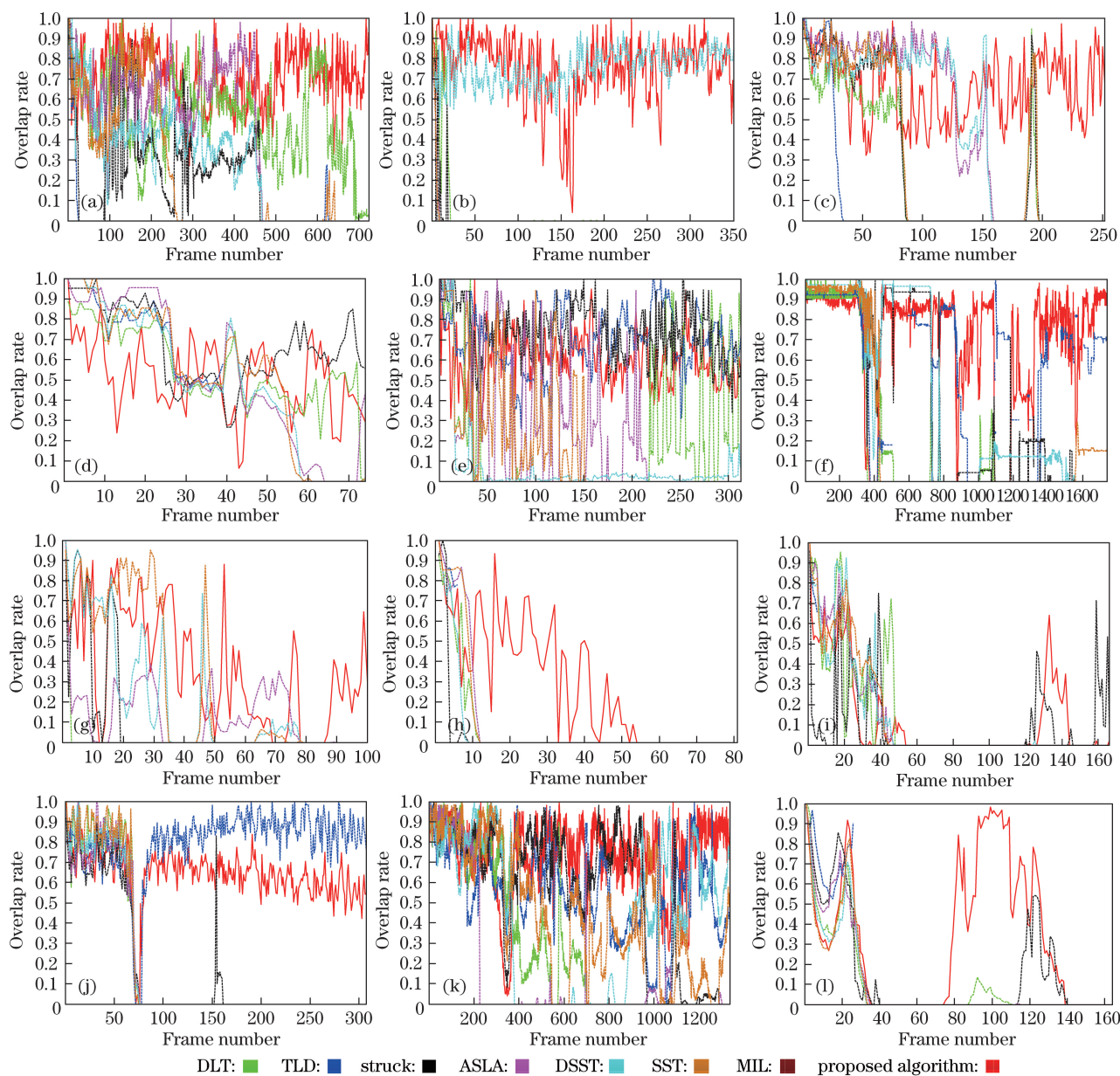


图 6 覆盖率曲线。(a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

Fig. 6 Overlap rate curves. (a) Basketball; (b) Bolt; (c) David3; (d) Football1; (e) Jumping; (f) Liquor; (g) Matrix; (h) Skiing; (i) Ironman; (j) Jogging1; (k) Lemming; (l) MotorRolling

每个图像序列所对应的最优结果标为红色,次优结果标为绿色。“-”为该算法在部分帧中给出异常结果或未给出结果,导致其运算超出平台限制范围。可以看出,本文算法的中心位置误差始终保持在较低位置,覆盖率保持在较高位置,展现出其优越的跟踪性能。Football1 序列中最优算法为 Struck 算法,次优算法为 MIL 算法,主要由于该序列中目标周围的干扰项很多,且干扰项的颜色、纹理和边缘特征都非常接近,影响了跟踪效果;Jumping 序列中本文算法的跟踪结果同样排在第三,主要是在该视频中,背景的纹理和边缘信息十分丰富,导致跟踪结果不够精确;在其他展示的序列中,本文算法均为最优或次优算法。

图 7 给出了 51 组视频的精度曲线和成功率曲线。分别从尺度变化、亮度变化、形变、遮挡以及运动模糊等多个环境对算法进行测试,从图 7 中可以看出,不论是在精度还是成功率方面,本文算法都呈现出良好的跟踪性能;图 8 和图 9 分别给出了不同算法在亮度变化(IV)、尺度变化(SV)、遮挡(OCC)、形变(DET)、运动模糊(MB)、快速移动(FM)、平面内旋转(IPR)、平面外旋转(OPR)、目标超出视野(OV)、背景扰乱(BC)和

表 2 部分跟踪结果的覆盖率比较

Table 2 Comparison of coverage rate of tracking results

%

Name	DLT	TLD	Struck	ASLA	DSST	SST	MIL	Proposed algorithm
Basketball	53.9	2.3	9.93	53.9	14.3	22.2	27.5	90.1
Bolt	4.29	1.4	1.71	1.4	100.0	1.1	1.1	91.6
David3	32.9	10.7	33.73	51.6	54.0	35.7	68.3	78.6
Football1	52.4	46.0	82.4	44.6	41.9	58.1	82.4	56.7
Jumping	16.6	90.1	95.9	17.3	4.8	13.4	62.3	81.3
Liquor	20.5	56.6	41.1	23.6	40.8	23.6	20.2	91.3
Matrix	2.0	1.0	12.0	2.0	21.0	34.0	12.0	36.0
Skiing	7.4	7.4	3.7	11.1	7.4	9.9	7.4	30.4
Ironman	10.8	3.0	4.8	13.3	10.8	14.5	7.2	13.9
Jogging1	22.5	96.4	21.8	22.5	22.5	22.2	22.2	92.8
Lemming	28.0	60.5	66.2	16.9	46.0	43.0	87.2	90.8
MotorRolling	7.3	15.9	16.5	11.0	6.7	7.3	7.3	30.5

表 3 部分跟踪结果的平均中心位置误差比较

Table 3 Comparison of average center position error of tracking results

pixel

Name	DLT	TLD	Struck	ASLA	DSST	SST	MIL	Proposed algorithm
Basketball	12.0	-	118.3	82.6	111.6	105.9	91.9	10.8
Bolt	-	-	398.8	374.7	5.0	409.9	393.5	5.6
David3	107.4	-	106.5	87.8	88.4	104.5	29.7	15.2
Football1	10.4	-	5.5	12.2	20.5	15.7	5.6	8.9
Jumping	41.9	5.9	6.7	46.1	35.2	45.7	10.0	7.6
Liquor	153.3	-	91.0	146.7	99.3	146.7	141.9	23.2
Matrix	171.1	-	194.8	65.2	59.7	54.7	55.0	33.5
Skiing	244.5	-	251.8	266.6	220.1	269.9	267.0	89.4
Ironman	211.4	-	127.6	197.5	105.7	205.4	193.4	74.3
Jogging1	113.0	6.7	62.0	104.6	112.0	114.6	96.3	12.3
Lemming	128.9	-	37.8	178.8	81.5	23.6	12.1	11.8
MotorRolling	170.7	-	145.6	201.4	289.6	377.0	161.0	101.8

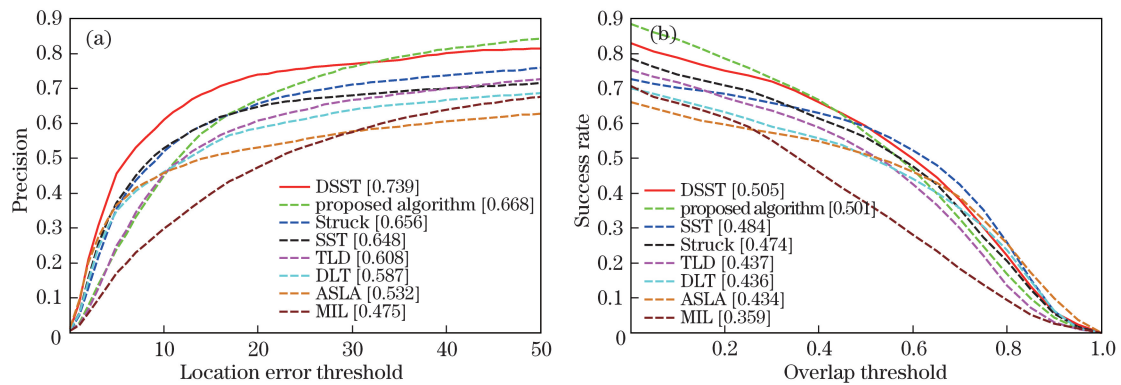


图 7 精度及成功率曲线。(a)精度;(b)成功率

Fig. 7 Curves of precision and success rate. (a) Precision; (b) success rate

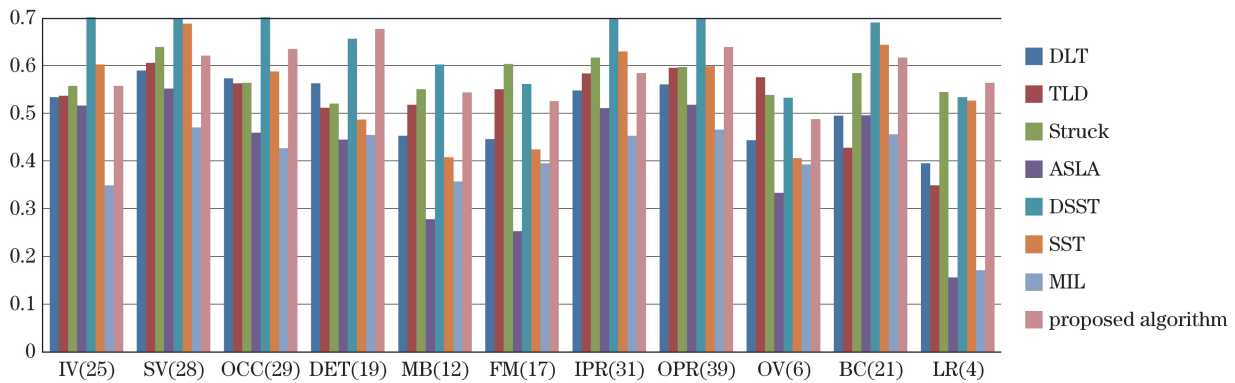


图 8 11 种场景下算法的跟踪精度比较

Fig. 8 Comparison of tracking precision under 11 scenes

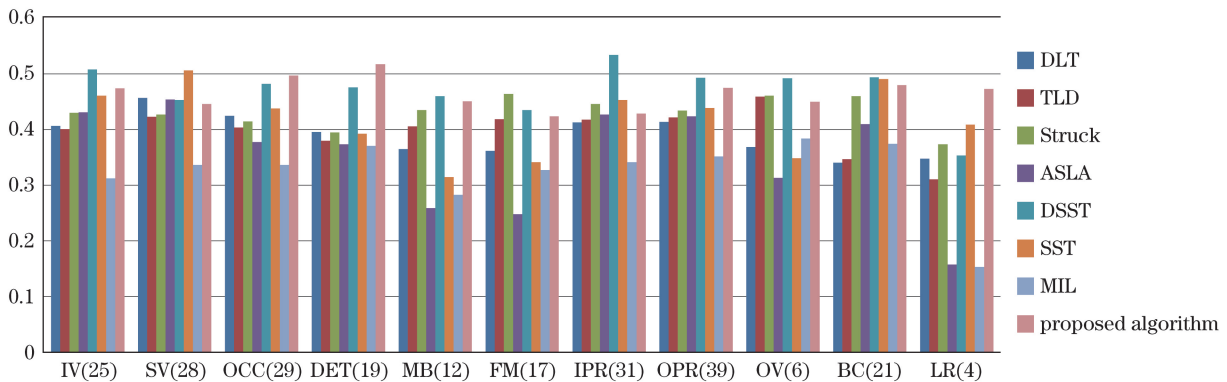


图 9 11 种场景下算法的跟踪成功率比较

Fig. 9 Comparison of tracking success rate under 11 scenes

低分辨率(LR)这 11 种场景下的精度和成功率。分析实验结果可知,本文算法在 OCC、DET、OPR、BC、IV、FM、MB、LR 等多种情况下性能均排在前三位,但整体而言,本文算法的跟踪效果不如 DSST 算法。DSST 算法的突出优势在于加入了尺度变换,利用尺度滤波器计算了目标尺度。而本文算法并未考虑目标的尺度变化问题,跟踪框的大小始终保持不变。未考虑尺度变化这一缺陷在分场景的测试中进一步得到体现。但是,本文算法的贡献在于提供了一种新的多特征融合跟踪思路,所提出的融合框架能够用于任何维度的多特征融合。在今后的研究工作中,将着重考虑尺度的问题,并探索更高效的特征组合,以期提出稳健性更强的跟踪算法。

此外,该算法在 Matlab 2009a 运算平台上进行 51 组视频测试的平均跟踪速度为 3.15 frame/s,算法复杂度在工程上能够达到实时。由于对比算法中,TLD、ASLA 和 DSST 算法采用的是 Matlab 与 C++ 混合编程,Struck 算法和 MIL 算法采用 C++ 编程,DLT 算法应用了图形处理器进行运算,故难以直接比较算法速度。SST 算法采用的运算平台与本文算法最为接近,采用 Matlab 编程并在 2.80 GHz 的 CPU(Intel Core2 Duo)、内存为 8 GB 的计算机上运行,其速度为 2.2 frame/s,略慢于本文算法。

6 结 论

通过结合区域协方差描述子和快速协方差交叉算法,提出了一种快速、有效的多特征融合跟踪算法。算法利用协方差描述子融合了颜色、边缘及纹理特征,利用不同特征所提供的互补信息有效地提升了对目标的描述能力。同时,算法采用快速协方差交叉算法,提取协方差矩阵中的有效信息进行快速模型更新。在整个跟踪过程中,同时进行了像素级与特征级融合,有效地提升了算法性能。

共进行了 51 组视频实验,实验的定量和定性分析结果均能够充分验证,在大多数场景下,本文算法比近年来提出的代表性算法更具有优势。但是,本文算法在应对尺度变化的场景时未能体现出优势,在今后的研究中,将针对尺度变化的问题进行探讨,以期达到更好的跟踪效果。

参 考 文 献

- [1] Qin Z, Shelton C R. Social grouping for multi-target tracking and head pose estimation in video[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2082-2095.
- [2] Chen S, Wang Q Q, Zhang S H, *et al.* Study of target tracking techniques based on non-scanning imaging lidar[C]. SPIE, 2015, 9622: 96220Y.
- [3] Liu K, Li X. Enabling context-aware indoor augmented reality via smartphone sensing and vision tracking[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2015, 12(1s): 15.
- [4] Liu Wei, Zhao Wenjie, Li Cheng. Long-term visual tracking based on spatio-temporal context[J]. Acta Optica Sinica, 2016, 36(1): 0115001.
刘 威, 赵文杰, 李 成. 时空上下文学习长时目标跟踪[J]. 光学学报, 2016, 36(1): 0115001.
- [5] Dou J F, Li J X. Robust visual tracking based on joint multi-feature histogram by integrating particle filter and mean shift[J]. Optik, 2015, 126(15): 1449-1456.
- [6] Guo Pengyu, Su Ang, Zhang Hongliang, *et al.* Online mixture of random naïve Bayes tracker combined texture with shape feature[J]. Acta Optica Sinica, 2015, 35(3): 0315002.
郭鹏宇, 苏 昂, 张红良, 等. 纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器[J]. 光学学报, 2015, 35(3): 0315002.
- [7] Yang K, Xiao Y H, Wang E D, *et al.* Mean shift tracking algorithm based on adaptive fusion of multi-feature[C]. SPIE, 2015, 9675: 96751L.
- [8] Jiang H L, Li J H, Wang D, *et al.* Multi-feature tracking via adaptive weights[J]. Neurocomputing, 2016, 207: 189-201.
- [9] Dou J, Li J. Robust visual tracking based on interactive multiple model particle filter by integrating multiple cues[J]. Neurocomputing, 2014, 135(8): 118-129.
- [10] Ma C, Liu C C, Peng F R, *et al.* Multi-feature hashing tracking[J]. Pattern Recognition Letters, 2016, 69: 62-71.
- [11] Tuzel O, Porikli F, Meer P. Region covariance: a fast descriptor for detection and classification[C]. European Conference on Computer Vision, 2006, 3952: 589-600.
- [12] Faulkner H, Shehu E, Szpak Z L, *et al.* A study of the region covariance descriptor: impact of feature selection and image transformations[C]. IEEE International Conference on Digital Image Computing: Techniques & Applications (DICTA), 2015: 1-8.
- [13] Gu X, Purvis M. Image segmentation with superpixel based covariance descriptor[C]. Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2016: 154-165.
- [14] Zeng M, Wu Z, Tian C, *et al.* Efficient person re-identification by hybrid spatiogram and covariance descriptor[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 48-56.
- [15] Zhang Lei, Fang Yangwang, Chai Dong, *et al.* Cruise missile path planning based on improved quantum evolutionary algorithm[J]. Acta Armamentarii, 2015, 35(11): 1820-1827.
张 磊, 方洋旺, 柴 栋, 等. 基于改进量子进化算法的巡航导弹航路规划方法[J]. 兵工学报, 2015, 35(11): 1820-1827.
- [16] Fan Fuyou, Yang Guowu, Le Qianqi, *et al.* Optimized coverage algorithm of wireless video sensor network based on quantum genetic algorithm[J]. Journal on Communications, 2015, 36(6): 94-104.
樊富有, 杨国武, 乐千杞, 等. 基于量子遗传算法的无线视频传感网络优化覆盖算法[J]. 通信学报, 2015, 36(6): 94-104.
- [17] Du Wenli, Zhou Ren, Zhao Liang, *et al.* Cooperative quantum differential evolution algorithm based method for optimizing neural networks[J]. Journal of Tsinghua University (Science and Technology), 2012, 52(3): 331-335.
杜文莉, 周 仁, 赵 亮, 等. 基于量子差分进化算法的神经网络优化方法[J]. 清华大学学报(自然科学版), 2012, 52(3): 331-335.
- [18] Guo Qiang, Sun Yuxiao. Improved quantum genetic algorithm with double chains in image denoising[J]. Journal of Harbin Institute of Technology, 2016, 48(5): 140-147.
国 强, 孙宇泉. 改进的双链量子遗传算法在图像去噪中的应用[J]. 哈尔滨工业大学学报, 2016, 48(5): 140-147.
- [19] Zhang J, Li H, Tang Z, *et al.* Quantum genetic algorithm for adaptive image multi-thresholding segmentation[J]. International Journal of Computer Applications in Technology, 2015, 51(3): 203-211.
- [20] Zhang Xiujie, Li Shiyong, Shen Yi, *et al.* Application of harmony search quantum genetic algorithm in image

- registration[J]. *Systems Engineering and Electronics*, 2012, 34(10): 2152-2156.
- 张秀杰, 李士勇, 沈毅, 等. 和声量子遗传算法在图像配准中的应用[J]. *系统工程与电子技术*, 2012, 34(10): 2152-2156.
- [21] Yang M, Zhang G. Using covariance intersection for change detection in remote sensing images[J]. *Journal of Electronics*, 2011, 28(1): 87-94.
- [22] Cong J L, Li Y Y, Qi G Q, *et al.* An order insensitive sequential fast covariance intersection fusion algorithm[J]. *Information Sciences*, 2016, 367: 28-40.
- [23] Wu Y, Li J W, Yang M H. Online object tracking: a benchmark [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013: 2411-2418.
- [24] Porikli F. Integral histogram: a fast way to extract histograms in cartesian spaces[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, 1: 829-836.
- [25] Han K H, Kim J H. Genetic quantum algorithm and its application to combinatorial optimization problem[C]. *Proceedings of the 2000 Congress on Evolutionary Computation*, 2000, 2: 1354-1360.
- [26] Wang N, Yeung D Y. Learning a deep compact image representation for visual tracking[C]. *Advances in Neural Information Processing Systems*, 2013: 809-817.
- [27] Kalal Z, Matas J, Mikolajczyk K. P N Learning: bootstrapping binary classifiers by structural constraints[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010: 49-56.
- [28] Hare S, Saffari A, Torr P H S. Struck: structured output tracking with kernels [C]. *International Conference on Computer Vision*, 2011: 263-270.
- [29] 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 (CVPR)*, 2012: 1822-1829.
- [30] Danelljan M, Hager G, Khan F S, *et al.* Accurate scale estimation for robust visual tracking[C]. *British Machine Vision Conference*, 2014, 65: 1-11.
- [31] Zhang T Z, Liu S, Xu C S, *et al.* Structural sparse tracking[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015: 150-158.
- [32] Babenko B, Yang M H, Belongie S. Visual tracking with online multiple instance learning[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009: 983-990.