

# 基于工地场景的深度学习目标跟踪算法

马少雄<sup>1,2</sup>, 邱实<sup>3</sup>, 唐颖<sup>4</sup>, 张晓<sup>5</sup>

(1. 西安理工大学, 陕西西安 710048; 2. 陕西铁路工程职业技术学院, 陕西渭南 714000;  
3. 中国科学院西安光学精密机械研究所中科院光谱成像技术重点实验室, 陕西西安 710119;  
4. 成都理工大学, 四川成都 610059; 5. 西北大学, 陕西西安 710127)

**摘要:** 针对施工现场环境复杂, 难以高效管理的问题. 提出了基于工地场景的深度学习目标跟踪算法, 辅助施工顺利进行. 根据工地现场目标的连续性, 构建增强群跟踪器, 提升目标成功跟踪的概率. 然后从滑动窗口、Stacked Denoising Auto Encoder (SDAE) 和 Support Vector Machine (SVM) 三方面组建深度检测器. 在滑动窗口方面: 从梯度角度建立模型实现窗口自适应. 在 SDAE 算法方面: 构建反向算法微调网络参数. 优化 SVM 算法降低跟踪时目标漂移和跟踪失败的概率, 最终实现目标高精度跟踪. 通过实验表明本文提出的算法可有效对目标进行跟踪, 实现动态管理.

**关键词:** 工地场景; 深度学习; 目标跟踪; 增强群滤波器; SDAE; SVM

**中图分类号:** TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2020)09-1665-07

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2020.09.001

## Deep Learning Target Tracking Algorithm Based on Construction Site Scene

MA Shao-xiong<sup>1,2</sup>, QIU Shi<sup>3</sup>, TANG Ying<sup>4</sup>, ZHANG Xiao<sup>5</sup>

(1. Xi'an University of Technology, Xi'an, Shaanxi 710048, China; 2. Shaanxi Railway Institute, Weinan, Shaanxi 714000, China;  
3. Key Laboratory of Spectral Imaging Technology CAS, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China;  
4. Chengdu University of Technology, Chengdu, Sichuan 610059, China; 5. Northwest University, Xi'an, Shaanxi 710127, China)

**Abstract:** Construction site is difficult to be effectively managed owing to its complex environment. A deep learning target tracking algorithm based on construction site scene is proposed to assist the construction progress. Firstly, according to the continuity of the target in the site scene, the enhanced group tracker is constructed to improve the successful probability of target tracking. Then, the depth detector is constructed with sliding window, stacked denoising auto encoder (SDAE) and support vector machine (SVM). Sliding window: a model is built from the gradient angle to realize window adaption. SDAE algorithm: the reverse algorithm is built to fine-tune network parameters. Optimized SVM algorithm reduces the probability of target drift and tracking failure. Finally, high precision tracking is achieved. Experiments show that the proposed algorithm can track the target effectively and realize dynamic management.

**Key words:** site scene; deep learning; target detecting; the enhanced group filter; SDAE; SVM

## 1 引言

随着视频监控技术的飞速发展, 视频监控技术已广泛应用至施工现场. 通过监控视频可动态了解施工进度, 秩序, 安全性等方面. 其中目标跟踪是视频分析的重要组成部分, 当前研究的热点问题.

Li<sup>[1]</sup>从机器学习角度分析视觉跟踪的可行性. Koch<sup>[2]</sup>利用贝叶斯准则实现目标跟踪. Zhang<sup>[3]</sup>使用

network flows 对多目标跟踪. Zhou<sup>[4]</sup>将 SIFT features 和 mean shift 相融合实现目标跟踪. Khan<sup>[5]</sup>将跟踪算法应用于分析拥挤人群. Qiu<sup>[6]</sup>根据红外与自然图像成像特点构建模型实现运动目标提取. Kuo<sup>[7]</sup>根据外观模型实现多目标在线跟踪. Benfold<sup>[8]</sup>利用滑动窗口对目标跟踪. Liu<sup>[9]</sup>建立 local sparse appearance model 和 k-selection 实现跟踪目标. Kalal<sup>[10]</sup>从机器学习角度实现目标跟踪与检测. Meijering<sup>[11]</sup>采用粒子滤波对目标跟踪. Qiu<sup>[12]</sup>

建立模型判断局部敏感区域予以跟踪. Zhang<sup>[13]</sup> 提取目标固有特征用于目标跟踪. Danelljan<sup>[14]</sup> 通过估计模型尺寸目标跟踪. Zhang<sup>[15]</sup> 结合上下文信息建立模型予以跟踪. Henriques<sup>[16]</sup> 改进相关滤波器实现高速跟踪. Ma<sup>[17]</sup> 提取多等级卷积特征实现目标跟踪. Ristani<sup>[18]</sup> 将多相机协同来实现目标跟踪. Zhang<sup>[19]</sup> 提取向量表示目标特征. He<sup>[20]</sup> 建立稀疏表达矩阵用于目标跟踪. Lukežič<sup>[21]</sup> 构建 Deformable parts correlation filters 进行目标跟踪. Gundogdu<sup>[22]</sup> 构建基于深度的相关滤波器实现目标跟踪. 以上算法从目标相似性角度出发构建模型实现目标跟踪, 均取得了一定的效果. 但是对于复杂环境下目标跟踪仍是存在较大难度的, 根据文献[1]指出当前目标跟踪所面临的主要挑战有: (1) 目标所处环境复杂, 采用单个跟踪器难以有效跟踪遮挡目标和非刚性形变目标. (2) 目标运动导致跟踪窗口大小和特征难以提取.

基于以上问题, 我们提出完整的目标提取流程: (1) 增强群跟踪器实现目标有效跟踪. (2) 优化滑动滤波窗口, 建立深度学习体系实现目标特征提取.

## 2 算法

深度学习可模拟人工思维方式建模, 对其研究是当前研究的热点. 深度学习算法已在诸多领域展现其出众的优越性. 为此我们从深度学习角度建立模型实现目标跟踪, 流程如图 1 所示. 为提高跟踪鲁棒性, 构建增强群跟踪器. 建立深度探测网络离线学习目标特征, 在线跟踪目标. 为降低训练过程中错误标记带来的误差提出加权 P-N 学习. 最终实现目标跟踪.

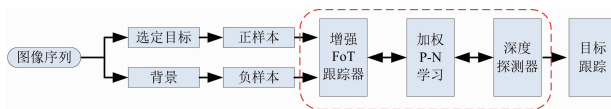


图1 算法流程图

### 2.1 增强群跟踪器

群跟踪器 (Flock of Tracker, FoT) 将局部跟踪器与全局运动模型结合, 可较好的应对目标遮挡、目标非刚性变化的问题<sup>[23,24]</sup>. 为进一步提高预测局部跟踪器的成功率, 本文构建增强 FOT 框图如图 2 所示, 将时空信息融入其中, 增加时空上下文预测器, 并与 NCC (Normalized Cross Correlation) 预测器构建级联预测器, 实现对局部跟踪器的预测. 为进一步降低模型评估失败率, 减少算法迭代次数, 引入快速随机采用一致性 (RANdom SAmple Consensus, RANSAC) 评估跟踪器的全局运动模型, 预测运动目标位置信息.

将 NCC 预测器、时间、空间上下文预测器以级联的方式组合. 当所有预测器均满足条件时, 局部跟踪器跟

踪成功. 从时间维度, 根据当前时刻信息构建马尔科夫模型, 利用时间上下文信息预测下一刻局部跟踪器正确跟踪的概率. 从空间维度, 根据短时间内相邻局部跟踪器之间的运动一致性, 判断局部跟踪器是否成功.

RANSAC 算法<sup>[25]</sup> 可在大数据中快速准确估计最优参数模型, 利用贝叶斯算法更新内点集合概率, 减少迭代次数, 降低模型失败概率. 为此, 根据级联预测器得到局部跟踪器集合, 采用 RANSAC 评估目标最优局部运动模型, 提高跟踪器对目标位置的预测精度, 缓解目标漂移问题.

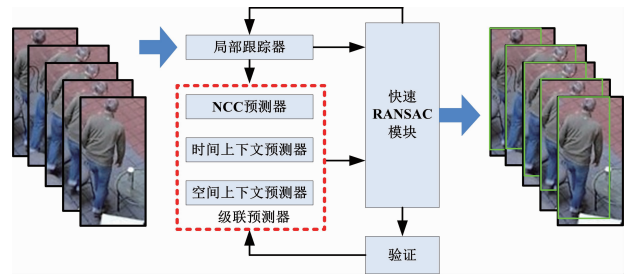


图2 增强FOT框图

### 2.2 深度检测器

深度检测器包含: (1) 滑动窗口, 利用多尺度滑动窗口全局扫描输入图像; (2) 深度去噪自编码器 SDAE, 提取多尺度扫描窗口内图像特征, 对图像信息表达; (3) 线性 SVM 分类器提取特征向量, 预测图像属性.

传统的滑动滤波<sup>[8]</sup> 过程为:

$$\begin{cases} \hat{\mathbf{X}}_{k|k-1} = \Phi_{k,k-1} \hat{\mathbf{X}}_{k-1} \\ \mathbf{P}_{k|k-1} = \Phi_{k,k-1} \mathbf{P}_{k-1} \Phi_{k,k-1}^T + \Gamma_{k-1} \mathbf{Q}_{k-1} \Gamma_{k-1}^T \\ \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{D}_{v_k}^{-1} \\ \hat{\mathbf{X}}_k = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_k (\mathbf{Z}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1}) \\ \mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \end{cases} \quad (1)$$

假定噪声稳定, 对应的滤波器增益函数  $\mathbf{K}_k^*$  为:

$$\begin{cases} \mathbf{K}_k^* = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \hat{\mathbf{D}}_{v_k}^{-1} \\ \hat{\mathbf{D}}_{v_k} = \frac{1}{N} \sum_{j=1}^k \bar{\mathbf{Z}}_{j|j-1} \bar{\mathbf{Z}}_{j|j-1}^T \\ \bar{\mathbf{Z}}_{j|j-1} = \mathbf{Z}_j - \mathbf{H}_j \hat{\mathbf{X}}_{j|j-1} \end{cases} \quad (2)$$

传统自适应滤波采用固定窗口宽度  $N$ ,  $N$  与自适应滤波结果的无偏性成正比, 而与自适应滤波跟踪灵敏度成反比. 为综合考虑无偏性与灵敏度, 选择合适的  $N$  尤为重要.

为此引入梯度检测函数对滑动窗口增益函数进行优化, 根据式(1)可知:

$$\begin{cases} \hat{\mathbf{R}}_k = \hat{\mathbf{D}}_k - \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T \\ \mathbf{R}_0 = \beta_k^{-1} \hat{\mathbf{R}}_k \end{cases} \quad (3)$$

其中  $\beta_k$  为预设量测噪声方差修正函数, 则

$$\beta_k^{-1} = \mathbf{R}_0 (\hat{\mathbf{D}}_{V_i} - \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T)^{-1} \quad (4)$$

通过实时更新  $\beta_k$ , 得到新的自适应滤波增益函数:

$$\mathbf{K}_k^* = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \beta_k \mathbf{R}_0)^{-1} \quad (5)$$

由于量测噪声方差矩阵随时间分段变化, 则必然存在明显梯度值, 可有效识别量测噪声方差阵变化, 定义梯度检测函数  $E$  为  $\beta_i$  的相关函数:

$$E = S \left| \left( \frac{1}{k - k_2 + 1} \sum_{i=k_2}^k \beta_i - \frac{1}{k_1 - k + L} \sum_{i=k-L+1}^{k_1} \beta_i \right) \times \left( \frac{1}{d(k_1 - k + L)} \sum_{i=k-L+1}^{k_1} \beta_i \right)^{-1} \right| \quad (6)$$

其中  $k_1, k_2 \in (k - L, k)$ , 根据式(3)检测梯度值变化, 定义滑动串口的宽度区间为  $N_k \in [N_{\max}, N_{\min}]$ , 则

$$N_k = \begin{cases} N_{\max}, & e < T \\ N_{\min} + (k - k_b)(k - k_b) \in [0, N_{\max} - N_{\min}], & e \geq T \\ 2N_{\max}(k - k_b) \notin [0, N_{\max} - N_{\min}], & e \geq T \end{cases} \quad (7)$$

当噪声统计特性发生变化时调整  $N_k$ . 当  $N_k$  减小, 使得噪声方差估计值跟踪实际噪声统计特性, 有效防止实际噪声统计特性发生变化的过程中, 噪声方差估计值跟踪不及时导致估计精度下降的问题. 在实际噪声统计特征稳定, 不调整  $N_k$ , 噪声方差估计值跟踪精度高, 滤波结果无偏性好, 自适应滤波精度可保证.

SDAE(深度去噪自编码器)从施加噪声的训练集

中学习, 重构原数据. 为降低训练的准确性, 采用无监督特征学习和逐层贪婪算法实现多个自编码器的预训练, 其结构如图 3 所示.

每个编码器隐藏层网络单元都为输入层网络单元的一半, 直至隐藏层网络单元数减至 256 为止, 具体如图 3(a) 所示. 预训练完成后, 将多个自编码器展开 SDAE, 并添加超完备滤波层, SDAE 网络结构如图 3(b) 所示. 为获得更有效的网络参数, 反向传播原始数据与重构数据之间误差微调整个网络, 此时网络结构如图 3(c). 由图 3(b)、图 3(c) 的输出结果可知微调后得到的重构数据更接近原始数据, 实现网络参数优化.

SVM 分类器<sup>[26]</sup>可提高深度检测器对相似目标辨识能力, 为缓解目标漂移或跟踪丢失问题, 将 SVM 优化为:

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{n=1}^N \max(1 - \mathbf{w}^T x_n y_n, 0)^2, y_n \in \{-1, 1\} \quad (8)$$

其中  $(x_n, y_n)$  为训练样本,  $C$  为规则参数, 预测数据  $x$  的类标为  $\arg \max_y (\mathbf{w}^T x) y$ .

为实现网络的准确性, 利用反向传播算法对深度分类神经网络优化, 将 SVM 分类器微调整个深度网络.

$$\frac{\partial l(\mathbf{w})}{\partial h_n} = -2y_n \mathbf{w} (\max(1 - \mathbf{w}^T h_n y_n, 0)) \quad (9)$$

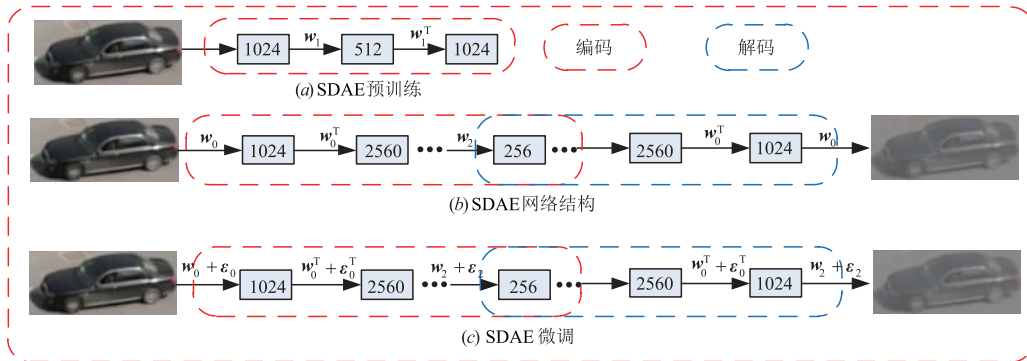


图3 SDAE结构图

### 2.3 加权 P-N 学习

在线 P-N 学习算法<sup>[27]</sup>在标记样本过程中常出现错误标记的问题, 用错误标记的样本训练分类器, 会降低分类器性能, 导致目标偏移甚至跟踪失败.

为解决上述问题, 在分类过程中对训练集中每个样本赋予权值: 正样本权值  $W^+$ 、负样本权值  $W^-$ :

$$\begin{cases} W^+ = W_b^+ + W_c^+ \\ W^- = W_b^- + W_c^- \end{cases} \quad (10)$$

其中  $W_b$  为自扩散分类过程赋予的权值,  $W_c$  为 SVM 分类过程赋予的权值.

$$\begin{cases} W_b^+ = \text{FN}^+ / (\text{FN}^+ + \text{FN}^-) \\ W_b^- = \text{FN}^- / (\text{FN}^+ + \text{FN}^-) \end{cases} \quad (11)$$

$\text{FN}^+, \text{FN}^-$  为样本分为正样本、负样本次数.

在分类过程中, SVM 分类器二值分类编码器提取的特征向量  $\mathbf{x}$ , 则分类权值为:

$$\begin{cases} W_c^+ = P(y = 1 | \mathbf{x}) \\ W_c^- = 1 - P(y = 1 | \mathbf{x}) \end{cases} \quad (12)$$

其中  $P(y = 1 | \mathbf{x})$  为图像块被分类为正样本的后验概率. 当  $W^+ / W^- \geq 1$  时为正样本, 反之则为负样本.

### 3 算法实现

算法主要分为离线训练和在线跟踪两过程. 离线训练: 将训练图像集归一化, 以 1024 维特征向量表示, 每一维对应一个像素. 将预处理的数据通过无监督特征学习预训练 SDAE. 根据输入数据和重构数据之间误差, 运用反向传播算法微调网络. 通过迁移学习将预训练得到 SDAE 编码器转化到在线跟踪过程中. 为加快预训练过程, 将每帧训练图像分为 5 个图像块, 分别训练这 5 个隐藏单元的自编码器, 根据这 5 个编码器权值初始化第 1 层自编码器.

在线跟踪: 选取待跟踪目标, 并对目标图像建立全视角数据库, 以此作为正样本集, 确保特征鲁棒性. 随机选取目标周围背景区域为负样本集. 将正负样本集通过 SDAE 编码器提取样本特征, 并以此训练 SVM 分类器.

在以上初始化的基础上, 实现跟踪阶段. 增强 FoT 跟踪器和深度检测器逐帧处理图像, 增强 FoT 跟踪器预测当前帧目标所在的位置, 深度检测器检测可能的目标位置, 整合模块整合检测结果和跟踪结果, 得到当前帧是否存在目标、目标位置及当前帧跟踪轨迹是否有效等信息. 加权 P-N 学习更新跟踪器和检测器.

模板更新策略可适应目标外观变化及光照等外界环境影响, 但是使在线模板数量不断增加, 从而增加匹配过程中的计算复杂度. 利用 K-means 算法, 将在线模板分为两个子集合, 构成二值树, 计算复杂度从原来的  $O(n)$  简化为  $O(\lg n)$ , 从而减少了时间复杂度.

### 4 实验结果与分析

实验采用 PETS2009<sup>[28]</sup> 与 BU-Marathon<sup>[29]</sup> 数据库提

供的视频数据, 包括光照变化、背景含噪、室内、室外视角、尺寸变化、快速移动变化的情况. 将目标用矩形窗予以标记作为金标准. 程序编写软件为 VS2015, 程序运行环境为 Win7, 设置滑动窗口大小为  $9 \times 9$ , 步长为 1.

#### 4.1 目标跟踪成功率

为衡量目标跟踪效果, 引入面积交迭度<sup>[30]</sup> (Area Overlap, AO) 作为跟踪效果的评价指标, 其定义为:

$$AO(\alpha, \beta) = \frac{S(\alpha \cap \beta)}{S(\alpha \cup \beta)} \quad (13)$$

AO 为面积交迭度,  $\alpha$  为标注的图像,  $\beta$  算法运行的结果图,  $S(\cdot)$  表示对应区域的像素点数. AO 越高说明效果越好.

设定目标跟踪成功的阈值  $T$ , 来直观展示算法效果. 由图 4 可知随着  $T$  的增加, 所有算法 AO 均呈现下降状态. 当光线变化, 如图 4(a) 所示, 由于光照变化引起区域相似度降低, 导致滑动窗算法<sup>[8]</sup> 效果最低. 如图 4(b) 所示, 由于背景的复杂性, 前景背景相似程度高, 但变化不是很强烈. 所以较光线变化的情况, 效果有所提升. 其中 Zhou<sup>[3]</sup> 建立 SIFT 特征框架提取目标时背景影响严重, 使得效果并不是很好. 如图 4(c) 所示, 由于室外场景目标扰动大, 导致固定模型无法满足需求. 如图 4(d) 所示, 由于室内场景的稳定性较室外强, 所以效果略提升. 如图 4(e) 所示, 前景大幅度变化而背景相对稳定, 算法均取得了较好的效果. 如图 4(f) 所示, 前景目标快速运动. 稀疏表达算法<sup>[20]</sup> 构建鲁棒的信息予以跟踪. 相关滤波算法<sup>[21]</sup> 将目标分块实现跟踪. 我们提出的算法构建增强群跟踪器, 级联判断相似性予以跟踪. 在复杂背景、室内视角条件下, 跟踪效果略低于其他算法. 但从整体上分析总体效果最优, 具有较强的鲁棒性.

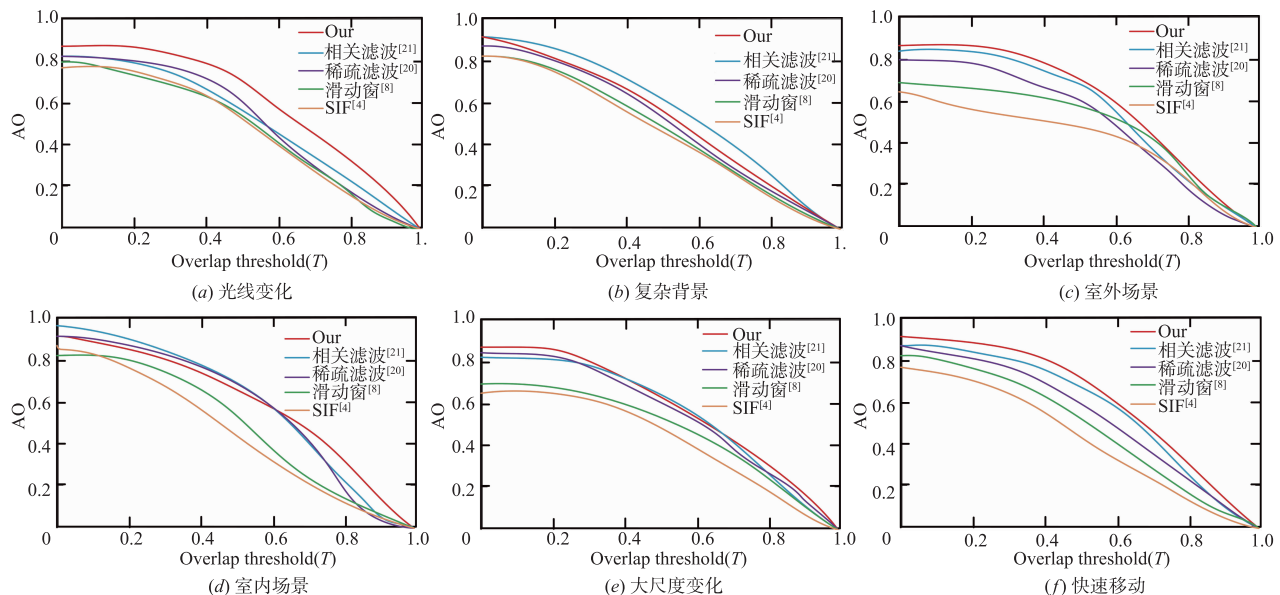


图4 算法响应曲线

### 4.2 目标跟踪精度

在目标跟踪成功的前提下,以 AO, AR (Accuracy Rank), RR (Robustness Rank), EAO (Expected Area Overlap) 衡量目标跟踪精度<sup>[31]</sup>.

跟踪精度如表 1 所示,SIFT<sup>[4]</sup>算法根据局部相似性跟踪目标,跟踪效果不好. 滑动窗<sup>[8]</sup>算法对目标首帧标注,并且动态更新滑动窗,跟踪效果欠佳. 稀疏表达<sup>[20]</sup>从全局角度构建模型,可较好的提取运动目标. 相关滤波<sup>[21]</sup>将目标划分为多部分,利用不同部分的相关性实现跟踪. 我们建立深度检测器,在训练过程中提出加权 P-N 算法提升训练器性能. 在跟踪中改进滑动窗匹配机制,动态更新模板,实现目标跟踪.

具体跟踪结果如图 5 所示:在红外图像中目标信息较为单一,算法均表现出较高的检测效果. 室内图像中目标存在遮挡,导致相关滤波、稀疏表达、SIFT 算法

跟踪不完全. 室外图像中目标高速运动,受滑动窗算法和 SIFT 算法模板更新速度的限制,导致目标不能完全跟踪. 目标被遮挡,导致提取运动目标不完整. 光学变化图像,SIFT 特征聚焦至阴影区域,导致跟踪失败. 多目标跟踪图像,到目标相邻较近且存在遮挡时,目标会出现误跟踪的情况. 而我们的算法在以上情况均体现出较好的性能.

表 1 不同算法跟踪精度对比

算法	AO	AV	RV	AR	RR	EAO
SIFT 算法 <sup>[3]</sup>	0.32	0.40	1.01	4.12	7.01	0.26
滑动窗算法 <sup>[8]</sup>	0.35	0.43	1.31	6.12	7.46	0.29
稀疏表达算法 <sup>[20]</sup>	0.40	0.48	1.19	6.52	7.86	0.24
相关滤波算法 <sup>[21]</sup>	0.42	0.52	1.20	8.12	8.01	0.31
本文算法	0.48	0.56	0.84	9.53	8.34	0.35



图5 目标跟踪图

## 5 总结

为降低施工现场管理难度,从视频监控角度提出基于工地场景的深度学习目标跟踪算法.从跟踪、检测、训练三个角度构建完整的目标跟踪框架.算法具有较强的鲁棒性,为后续目标分析奠定了基础.

### 参考文献

- [1] Li P, Wang D, Wang L, et al. Deep visual tracking: Review and experimental comparison [J]. *Pattern Recognition*, 2018, 76(4): 323 – 338.
- [2] Koch J W. Bayesian approach to extended object and cluster tracking using random matrices [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2008, 44(3): 1042 – 1059.
- [3] Zhang L, Li Y, Nevatia R. Global data association for multi-object tracking using network flows [A]. *Computer Vision and Pattern Recognition (CVPR)*, 2008 IEEE Conference on [C]. Anchorage, USA; IEEE, 2008. 1 – 8.
- [4] Zhou H, Yuan Y, Shi C. Object tracking using SIFT features and mean shift [J]. *Computer Vision and Image Understanding*, 2009, 113(3): 345 – 352.
- [5] Khan S M, Shah M. Tracking multiple occluding people by localizing on multiple scene planes [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(3): 505 – 519.
- [6] Qiu S, Luo J, Yang S, et al. A moving target extraction algorithm based on the fusion of infrared and visible images [J]. *Infrared Physics & Technology*, 2019, 98(5): 285 – 291.
- [7] Kuo C H, Huang C, Nevatia R. Multi-target tracking by on-line learned discriminative appearance models [A]. *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on [C]. San Francisco, USA; IEEE, 2010. 685 – 692.
- [8] Benfold B, Reid I. Stable multi-target tracking in real-time surveillance video [A]. *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on [C]. Providence, USA; IEEE, 2011. 3457 – 3464.
- [9] Liu B, Huang J, Yang L, et al. Robust tracking using local sparse appearance model and  $k$ -selection [A]. *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on [C]. Providence, USA; IEEE, 2011. 1313 – 1320.
- [10] Kalal Z, Mikolajczyk K, Matas J. Tracking learning detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409 – 1422.
- [11] Meijering E, Dzyubachyk O, Smal I. Methods for cell and particle tracking [J]. *Methods in Enzymology*, 2012, 504(4): 183 – 200.
- [12] Qiu S, Tang Y, Du Y, et al. The infrared moving target extraction and fast video reconstruction algorithm [J]. *Infrared Physics & Technology*, 2019, 97(11): 85 – 92.
- [13] Zhang L, Van Der Maaten L. Preserving structure in model-free tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 36(4): 756 – 769.
- [14] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking [A]. *British Machine Vision Conference* [C]. Nottingham, UK; Springer, 2014. 1 – 11.
- [15] Zhang K, Zhang L, Liu Q, et al. Fast visual tracking via dense spatio-temporal context learning [A]. *European Conference on Computer Vision* [C]. Zurich, Switzerland; Springer, 2014. 127 – 141.
- [16] 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.
- [17] Ma C, Huang J B, Yang X, et al. Hierarchical convolutional features for visual tracking [A]. *The IEEE International Conference on Computer Vision* [C]. San Francisco, USA; IEEE, 2015. 3074 – 3082.
- [18] Ristani E, Solera F, Zou R, et al. Performance measures and a data set for multi-target, multi-camera tracking [A]. *European Conference on Computer Vision* [C]. Amsterdam, The Netherlands; Cham, 2016. 17 – 35.
- [19] Zhang L, Suganthan P N. Visual tracking with convolutional random vector functional link network [J]. *IEEE Transactions on Cybernetics*, 2017, 47(10): 3243 – 3253.
- [20] He Z, Yi S, Cheung Y M, et al. Robust object tracking via key patch sparse representation [J]. *IEEE Transactions on Cybernetics*, 2017, 47(2): 354 – 364.
- [21] Lukežič A, Zajc L Č, Kristan M. Deformable parts correlation filters for robust visual tracking [J]. *IEEE Transactions on Cybernetics*, 2018, 48(6): 1849 – 1861.
- [22] Gundogdu E, Alatan A A. Good features to correlate for visual tracking [J]. *IEEE Transactions on Image Processing*, 2018, 27(5): 2526 – 2540.
- [23] Vojtíř T, Matas J. The enhanced flock of trackers [A]. *Registration and Recognition in Images and Videos* [C]. Berlin, Heidelberg; Springer, 2014. 113 – 136.
- [24] Cheng S, Cao Y G, Sun J X. Target tracking based on enhanced flock of tracker and deep learning [J]. *Journal of Electronics and Information Technology*, 2015, 37(7): 1146 – 1153.
- [25] Li X. Image authenticity decision based on random sample consensus and circular feature selection [A]. *2018 International Conference on Computing and Artificial Intelligence* [C]. New York, USA; ACM, 2018. 91 – 97.
- [26] Singh D, Khan M A, Bansal A, et al. An application of SVM in character recognition with chain code [A]. *Com-*

- munication, Control and Intelligent Systems ( CCIS ), 2015 [ C ]. Mathura, India; IEEE, 2015. 167 – 171.
- [27] Kalal Z, Matas J, Mikolajczyk K. Pn learning; Bootstrapping binary classifiers by structural constraints [ A ]. Computer Vision and Pattern Recognition ( CVPR ), 2010 IEEE Conference on [ C ]. San Francisco, USA; IEEE, 2010. 49 – 56.
- [28] Ferryman J, Shahrokni A. Pets2009; Dataset and challenge [ A ]. 2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance [ C ]. Snowbird, USA; IEEE, 2009. 1 – 6.
- [29] YE R, DENG Y. Multi-target tracking method based on sparse theory and sub-problem coupling [ J ]. Computer Engineering, 2017, 43 ( 6 ): 219 – 224, 229.
- [30] Qiu S, Wen D, Cui Y, et al. Lung nodules detection in CT images using gestalt-based algorithm [ J ]. Chinese Journal of Electronics, 2016, 25 ( 4 ): 711 – 718.
- [31] Qiu S, Luo J, Yang S, et al. A moving target extraction algorithm based on the fusion of infrared and visible images [ J ]. Infrared Physics & Technology, 2019, 98: 285 – 291.

### 作者简介



**马少雄** 男, 1982 年生, 2008 年于西安工业大学获得硕士学位, 2019 年博士毕业于西安理工大学土木工程学院。现为陕西铁路工程职业技术学院铁成(创新)学院副院长, BIM 技术应用研究中心副主任、副教授, 主要研究方向为基本 BIM 技术的信息化管理平台研发及应用、图像处理。

**邱实(通信作者)** 男, 1986 年生, 中国科学院西安光学精密机械研究所中科院光谱成像技术重点实验室副研究员, 博士, 研究方向: 图像处理、模式识别。

E-mail: qiushi215@163.com

**唐颖** 女, 1978 年生, 成都理工大学信息与科学学院电子工程系副教授, 硕士, 研究方向: 嵌入式系统, 图像处理、无线传感器网络。

**张晓** 男, 1992 年生, 西北大学信息科学与技术学院, 博士, 研究方向: 图像处理, 目标识别。