

# 基于动态模板匹配的自适应尺度目标跟踪算法

陈方芳, 宋代平\*

重庆大学机械与运载工程学院, 重庆 400030

**摘要** 为保证大型光学模块装校装置在装校过程的稳定性和准确性, 需要对导向模块进行监测跟踪。针对传统模板匹配跟踪算法在工业应用中存在无法应对目标尺度变化、缺少有效的模板更新策略等问题, 提出一种基于动态模板匹配的自适应尺度目标跟踪算法。首先对第一帧图像检测运动目标区域, 提取目标模板中心点, 生成模板图像金字塔; 在之后的每一帧中, 先计算各尺度下模板图像和目标图像的相似性, 取相似度最大的尺度因子作为尺度变化量, 借助动态模板更新策略更新模板; 最后融合卡尔曼滤波算法, 预测出候选目标范围。在 OTB 数据集上进行实验, 结果表明, 所提算法在满足实时性要求的同时, 相较于传统算法, 重合率提高约 21 个百分点。

**关键词** 目标跟踪; 动态模板; 自适应尺度; 卡尔曼滤波

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP213163

## Adaptive Scale Target Tracking Algorithm Based on Dynamic Template Matching

Chen Fangfang, Song Daiping\*

College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400030

**Abstract** To ensure the stability and accuracy of a large-scale optical module in the installation process, it is necessary to monitor and track the guide module. An adaptive scale target tracking algorithm based on dynamic template matching is proposed to address the problems of traditional template matching tracking algorithms in industrial applications, such as their inability to cope with target scale change and lack of a template adaptive update mechanism. First, the moving target area is identified from the first frame image, the center point of the target template is extracted, and the template image pyramid is generated. The template is updated using a dynamic template update strategy, and then the Kalman filtering algorithm is used to predict the potential target range in each subsequent frame. The similarity between the template image and the target image at each scale is calculated, and the scale factor with the highest similarity is taken as the scale change. In comparison to conventional algorithms, experiments on the OTB dataset demonstrate that the proposed algorithm can meet real-time requirements while increasing the coincidence rate by about 21 percentage points.

**Key words** target tracking; dynamic template; adaptive scale; Kalman filtering

## 1 引言

目标跟踪作为计算机视觉的一个基本研究领域, 主要方法分为生成式跟踪和判别式跟踪<sup>[1]</sup>。前者基于目标特征的有效表达如灰度特征、几何特征等对物体进行跟踪<sup>[2]</sup>, 代表性方法包括均值漂移(MeanShift)算法<sup>[3]</sup>、卡尔曼滤波算法、模板匹配算法等。均值漂移方法仅依靠色彩直方图作为匹配特征, 在目标与背景颜色相近时会出现误匹配。卡尔曼滤波算法无法应对物

体消失情况。而传统模板匹配算法虽具有定位精准、方便实现等特点, 也存在着自适应更新机制缺乏、全局搜索范围大等问题。不同于生成式跟踪方法, 判别式跟踪方法将目标跟踪看作划分目标与背景的二分类问题, 除学习目标自身特征外, 也充分学习目标背景信息, 包括跟踪学习检测(TLD)算法等<sup>[4-5]</sup>。判别式跟踪模型充分利用了背景信息, 具有较高的准确性, 但计算量大, 容易学习到错误样本, 同时在线分类器的训练复杂度也直接影响到跟踪器的实时性。近年来, 融合深

收稿日期: 2021-12-07; 修回日期: 2021-12-26; 录用日期: 2022-01-05; 网络首发日期: 2022-01-13

基金项目: 国家自然科学基金(U1530138)

通信作者: \*songdp@cqu.edu.cn

度学习的目标跟踪算法<sup>[6-10]</sup>成为热门方向,无需手工设计特征即可实现高准确率分类,但同样有局限,一是需要大量的训练数据,二是对硬件性能要求高。

相比于使用单一算法,研究人员往往将不同的算法融合在一起以进行优势互补。文献[11]将相关滤波算法融合,实现了稳定跟踪效果。文献[12]分别采用MeanShift算法和粒子滤波算法,综合得出正确的目标位置。文献[13]将深度特征与相关滤波算法相结合,以设计出包含更多目标信息的精确特征。针对复杂背景和相似干扰问题,文献[14]将自适应模型融合到判别式尺度空间跟踪(DSST)模型中以提高算法的鲁棒性,文献[15]融合深度特征、梯度直方图特征和颜色直方图特征以提高跟踪性能。针对短时遮挡问题,文献[16]运用卡尔曼滤波算法的预测机制来优化连续自适应均值漂移(Camshift)算法,取得了不错的跟踪效果。

模板匹配是目标跟踪领域常用的算法之一,指将已知模板作为窗口,在待匹配图中进行滑动搜索,以寻找到最相似匹配位置的跟踪方法。模板匹配跟踪算法是通过衡量已知模板和待匹配图间的相似性来对实时帧进行目标跟踪的方法,具有匹配精度高、实现简单等特点,但难以应对目标形态发生变化及运动的场景。不少学者采用动态模板匹配来解决运动物体的识别和跟踪问题<sup>[17-18]</sup>,取得了较好的跟踪效果。不同于仅采用固定目标物体模板的传统模板匹配,动态模板匹配根据应用场景的不同,采用不同的模板更新策略实时更新模板,显著提高了目标跟踪的准确率。

在大型光学模块装校装置工作过程中,观测相机固定,导向模块进行规律的匀速运动。整个运行过程中,存在着跟踪目标尺度变化、背景复杂、光照变化等干扰因素。鉴于此,本文提出一种基于动态模板匹配的自适应尺度目标跟踪算法,融合图像金字塔和卡尔曼滤波算法进行优化,同时提出有效的模板更新机制,并在多个实验场景下对所提算法进行验证。

## 2 基于动态模板匹配的自适应尺度目标跟踪算法原理

所提算法的流程如图1所示。首先对第一帧图像检测运动目标区域,并将其作为初始模板,借助尺度因子集生成模板图像金字塔,进行模板参数初始化;在之后的每一帧中,先计算各尺度下模板图像和目标图像的相似性,取相似性最大的尺度因子作为这一帧图像的尺度变化量;同时提出一种加权模板更新策略,对当前模板和目标跟踪结果进行加权,产生一个新模板,增加在运动过程中目标跟踪的鲁棒性和稳定性。针对模板跟踪全局搜索范围大的问题,借助卡尔曼滤波算法,预测出下一时刻的运动目标中心点,减小搜索范围。后文将所提算法简称为SSTM,以便于描述。

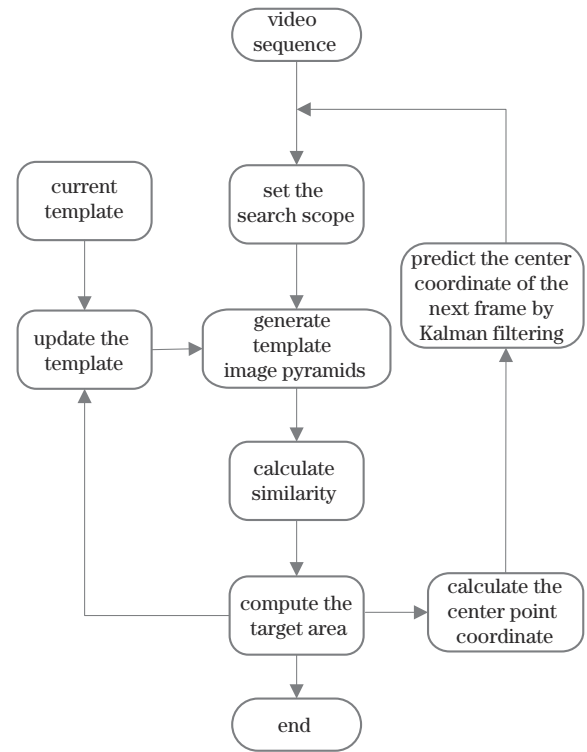


图1 所提算法的流程

Fig. 1 Flowchart of the proposed algorithm

### 2.1 生成模板图像金字塔

应用场景为装校过程中导向模块的监测跟踪,在装校装置的对接过程中,导向模块与观测相机之间的距离由远及近,呈现出一定的尺度变化,此时若采用原始的模板匹配跟踪方法,则会出现目标漂移、定位不准等问题。图像金字塔指来源于同一张原始图,但分辨率不同的相似图像集合,分为高斯金字塔和拉普拉斯金字塔。前者金字塔由下至上分辨率逐级降低,拉普拉斯金字塔则相反。金字塔作为一种直观的图像多尺度表达结构,在目标运动造成尺寸变化的情况下有相当出众的稳定性。

尺度因子集可以看作是不同尺度因子的集合, $s=1$ 时代表原始尺寸大小, $s<1$ 时尺度变小,当 $s>1$ 时尺度变大。由于目标物体的运动是连续的,尺度变化也是连续的,相邻图像帧之间的尺度变化的大部分在0.85~1.15变化范围内,因此设立尺度因子集为

$$S = \{S_1, S_2, \dots, S_k, \dots, S_{2k-1}\} \quad (1)$$

得到的集合为一个关于 $S_k=1$ 对称且单调递增的等差数列,按照尺度因子池中的各个尺度因子对模板图像进行尺寸变化,则可得到对应的模板图像金字塔。

### 2.2 模板更新策略

传统的模板匹配算法使用第一帧的固定目标物体模板进行全程跟踪,缺乏一种有效的模板更新策略,无法处理尺度变化、光照变化、背景复杂等干扰因素的影

响。实时更新策略对当前帧的跟踪结果进行处理,并将结果作为下一帧的更新模板,优点是能够实时反映目标状态和环境变化,但同时存在误差逐渐累积、模板漂移等问题<sup>[19]</sup>。如图 2 所示,在目标跟踪过程中,目标状态发生变化,图 2(b)为人工标注的真实结果。若使用传统模板匹配算法的固定模板,如图 2(c)所示,则无法学习到有效的特征变化。所提加权模板更新策略对当前模板和目标跟踪结果进行加权,产生一个新模

板,各帧目标跟踪结果的权重系数指数式衰减,越靠近当前时刻的数据,权重越大,使模板图像自适应更新的同时避免学习到错误模板。相比固定模板策略,加权模板更新策略得到的模板能够有效地应对目标尺度变化等一系列问题。由于更新得到的模板是过去所有模板的滑动平均,因而具有一定的鲁棒性,不因某次的异常取值产生波动。图 2(d)为使用改进加权模板更新策略之后生成的模板,与真实结果更加接近。



图 2 模板更新算法比较。(a)初始帧;(b)真实结果;(c)传统算法模板;(d)所提算法模板

Fig. 2 Comparison of template update algorithms. (a) Initial frame; (b) ground truth; (c) traditional algorithm template; (d) proposed algorithm template

所提模板更新策略如下。设初始模板为  $T_{ori}$ , 第  $i$  帧的最优目标匹配区域为  $R_i$ , 第  $i$  帧的模板图像为  $T_i$ , 则有

$$T_i = \begin{cases} T_{ori}, & i = 1 \\ \lambda T_{i-1} + (1 - \lambda) R_i, & i \geq 2 \end{cases}, \quad (2)$$

式中:  $\lambda$  为可调参数,  $0 \leq \lambda \leq 1$ 。  $\lambda$  值越大, 模板更新时的鲁棒性越高, 对原始模板依赖程度越高, 更新幅度也越小;  $\lambda$  值越小, 应对目标物体形态变化的能力越强, 但也越容易出现模板漂移。在目标形态变化较快时  $\lambda$  可取较大值。进行上述过程, 可得到不同尺度下的第  $i$  帧模板图像金字塔。有序列  $T_i^{pyr} = T_i \cdot S$ , 得  $T_i^{pyr} = \{T_i^1, T_i^2, \dots, T_i^k, \dots, T_i^{2k-1}\}$ , 其中  $T_i^k = T_i \cdot S_k$ , 为对应  $S_k$  尺度下的第  $i$  帧模板图像。

### 2.3 确定搜索范围

构建多尺度模板图像序列后, 若通过全局搜索的方式对各个尺度的模板图像进行匹配得出目标图像, 计算量将会成倍提高, 实时性难以得到保证。因此通过引入卡尔曼滤波算法预测出下一时刻目标图像中心位置的坐标值, 将会大大缩小搜索范围, 提高算法效率。如图 3 所示, 方框为待匹配图像, 浅色方框为搜索范围, 深色方框为模板图像。在待匹配图像尺寸为  $720 \times 480$ , 模板图像尺寸为  $101 \times 101$  时, 若使用传统

全局匹配, 需要计算 235600 次, 若采用卡尔曼滤波算法来预测下一帧的目标中心点, 则仅需计算 10404 次。

#### 2.3.1 卡尔曼滤波算法原理

卡尔曼滤波算法在图像跟踪领域有着广泛应用, 其输出只依赖于当前的输入和前一时刻的输出, 滤波过程中使用观测值来修正当前值。卡尔曼滤波用来处理两种噪声的干扰, 从不确定性中以最大概率的方式获得最终状态量。两种噪声分别为运动不确定性和观测噪声。其中运动不确定性表示持续的运动累积造成的运动噪声, 该种噪声符合高斯分布。观测噪声指传感器和观测过程中的外界因素的影响造成的观测不确定性, 与时间完全不相关, 也属于高斯白噪声且无法避免, 在每个时间段为同样噪声概率分布。

$$X_k = F_k X_{k-1} + B_k \mu_k + w_k, \quad (3)$$

$$Z_k = H_k X_k + m_k, \quad (4)$$

式中:  $X_k$  表示  $k$  时刻系统的状态变量;  $Z_k$  表示所获得的测量数据;  $F_k$  表示邻近两时刻状态量之间的转移矩阵;  $B_k$  为作用在控制器变量上的控制变换矩阵;  $w_k$  为运动过程中的白噪声, 服从  $w_k \sim N(0, Q_k)$ ,  $Q_k$  为运动噪声的协方差矩阵;  $H_k$  是观测矩阵, 把真实状态空间映射成观测空间;  $m_k$  是观测噪声, 服从  $m_k \sim N(0, R_k)$ ,  $R_k$  为测量噪声的协方差矩阵。式(3)为系统运动方程, 式(4)为观测方程<sup>[20-21]</sup>。

$$X'_k = F_k X''_{k-1} + B_k \mu_k, \quad (5)$$

$$P'_k = F_k P''_{k-1} F_k^T + Q_k, \quad (6)$$

式中:  $X'_k$  和  $X''_{k-1}$  分别表示  $k$  时刻的先验估计值和最优估计值;  $P'_k$  和  $P''_{k-1}$  分别表示先验和后验误差估计协方差矩阵, 根据  $k-1$  时刻的状态来估计  $k$  时刻的状态量, 以关系矩阵  $F_k$  表示该预测过程;  $B_k$  和  $\mu_k$  主要对预测进

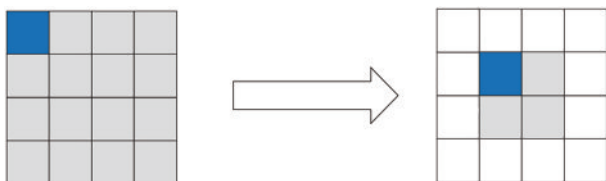


图 3 卡尔曼滤波算法效果示意图

Fig. 3 Effect diagram of Kalman filtering algorithm

行修正,用来表示系统存在的外部控制因素。式(5)和式(6)为预测方程。

$$\mathbf{X}_k'' = \mathbf{X}_k' + \mathbf{K}(\mathbf{Z}_k - \mathbf{B}_k \boldsymbol{\mu}_k), \quad (7)$$

$$\mathbf{P}_k'' = (1 - \mathbf{K}\mathbf{H}_k) \mathbf{P}_k', \quad (8)$$

式中: $\mathbf{K}$ 为卡尔曼增益矩阵。

$$\mathbf{K} = \mathbf{P}_k' \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k' \mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (9)$$

式(7)~(9)为参数更新方程。

### 2.3.2 目标中心点估计

观测相机与目标物体匀速运动状态下,设立 $\Delta_x$ 为图像中目标沿 $x$ 轴方向的移动速度, $\Delta_y$ 为图像中目标沿 $y$ 轴方向的移动速度, $(X_k, Y_k)$ 为当前帧和前一帧目标框图中心像素的 $x$ 轴坐标, $(Y_k, Y_{k-1})$ 为当前帧和前一帧目标框图中心像素的 $y$ 轴坐标,则状态转移矩阵为

$$\begin{pmatrix} X_k \\ Y_k \\ \Delta_x \\ \Delta_y \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{k-1} \\ Y_{k-1} \\ \Delta_x \\ \Delta_y \end{pmatrix}. \quad (10)$$

根据小孔成像模型,具体的观测方程为

$$\begin{pmatrix} X_k \\ Y_k \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \mathbf{m}_k. \quad (11)$$

通过建立目标跟踪的运动方程与观测方程的数学

模型,设定第 $i-1$ 帧的最优目标匹配框中心坐标值为 $(X_{i-1}^c, Y_{i-1}^c)$ ,最优尺度因子为 $S_{i-1}^{\text{next}}$ ,即可运用卡尔曼滤波算法预测出第 $i$ 帧的最优目标匹配框中心坐标值 $(X_i^c, Y_i^c)$ 。同时第 $i$ 帧的搜索候选框即为以 $(X_i^c, Y_i^c)$ 为中心、 $2 \times W_{i-1}^{\text{next}}$ 和 $2 \times H_{i-1}^{\text{next}}$ 作为宽度和高度的矩形框。

### 2.4 目标区域匹配

在确定第 $i$ 帧的模板图像序列和搜索框后,即可在搜索范围内进行模板匹配。如图4所示,模板匹配通过将模板图像作为窗口,在待匹配图像上滑动比对,通过衡量相似度的方法进行匹配。所提算法通过计算各尺度下模板图像和候选框的相似度来确定第 $i$ 帧的最优模板尺度及最优目标区域。以 $T_i^k$ 为例,设定候选框在 $(x, y)$ 处的像素值为 $I(x, y)$ ,当前模板图像 $T_i^k$ 在 $(x', y')$ 处的像素值为 $T_i^k(x', y')$ ,初始模板 $T_{\text{ori}}$ 在 $(x'', y'')$ 处的像素值为 $O(x'', y'')$ ,则本文设定匹配的损失函数为

$$L_i^k = R[I(x, y), T_i^k(x', y')] + R[I(x, y), O(x'', y'')], \quad (12)$$

式中: $R(\cdot)$ 为标准化差值平方和匹配函数,其中两项分别为当前模板图像和待匹配图、初始模板图像和待匹配图的标准化差值平方和匹配函数。损失函数衡量了待匹配图与模板图像和原始图像的相似度,第二项的加入意在防止算法学习到错误的模板。

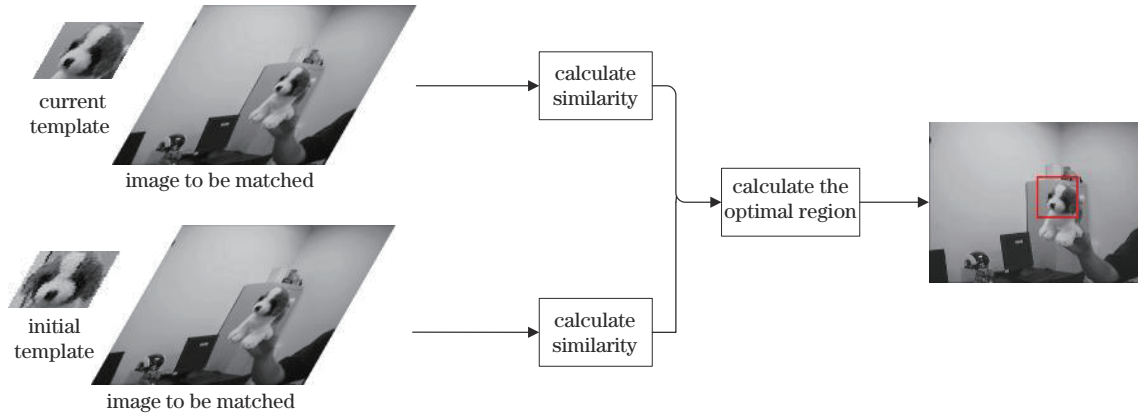


图4 模板匹配示意图

Fig. 4 Diagram of template matching

对模板图像序列集合 $T_i^{\text{pyr}}$ 进行模板匹配操作,则可得到不同尺度下对应的损失函数集 $L_i = \{L_i^1, \dots, L_i^k, \dots, L_i^{2k-1}\}$ ,取 $L_i^{\min} = \min\{L_i^1, \dots, L_i^k, \dots, L_i^{2k-1}\}$ ,其对应的最优目标区域作为第 $i$ 帧的跟踪结果 $R_i$ ,最优目标匹配框中心点设为 $(X_i^c, Y_i^c)$ 。取不同尺度下损失函数最小的尺度作为最优目标尺度,设为 $S_i^{\text{next}}$ ,对应图像的宽度和高度分别为 $W_i^{\text{next}}$ 和 $H_i^{\text{next}}$ ,用来确定下一帧目标搜索范围框的尺寸。

## 3 实验

### 3.1 评估参数

对所提算法进行分析可知,尺度因子集和 $\lambda$ 值的设置对算法性能有较大的影响,需根据实际应用场景进行调整。在本文提出的应用场景中,目标物体运动较为规律,不存在或极少有快速运动,因此连续帧间尺度变化较小,基于此,本实验将尺度因子集设为 $\{0.95, 1, 1.05\}$ 。跟踪的目标物体为机械零部件,形变幅度较

小,但跟踪背景复杂,基于此,将 $\lambda$ 值取为 0.7。

目标跟踪标记(OTB)数据集作为目标跟踪算法常用的测试集,涵盖了实际应用场景中的绝大部分干扰因素,如姿态变化、分辨率低、背景复杂等。对所提算法、TLD、核循环矩阵跟踪(CSK)<sup>[22]</sup>、传统 TM 进行对比。TLD 是判别式跟踪方法类的典型算法之一,同时也是长时目标追踪领域的代表算法。CSK 作为经典滤波类跟踪算法之一,使用点乘操作代替滑窗法来完成相关性检测,以提高运算速度,既能保持良好的跟踪速度也能兼顾目标跟踪准确率。选择精确度(Prs)、重合率(Suc)和检测速度 3 个指标对跟踪结果进行分析。精确度指所有距离小于预测位置中心点与真实位置中心点间的特定像素欧氏距离阈值的图像帧数占图像总帧数的百分比。重合率反映预测的位置边框与真实边框的重合程度,为大于对应重合程

度的图像帧数占图像总帧数的百分比。所提算法需要初始帧目标的位置之后才能运行跟踪算法,因此选择 OPE(一次通过性)评估方法。主要有光照变化、尺度变化、背景干扰、平面内外旋转等场景,因此选用 OTB 数据集中具有以上因素的 12 个典型测试集来测试算法的性能。

### 3.2 实验结果

图 5 为 SSTM、TM-V(模板跟踪算法)、CSK、TLD 在 12 个典型场景测试集下的性能对比。图 5(a)的图例为在阈值取 0.5 时的重合率,图 5(b)的图例为阈值取 25 pixel 时的精确度。从图 5 可以看出:所提 SSTM 在 4 个算法中表现最好,TLD 算法和 CSK 算法次之,传统 TM 算法表现稍差;相比于传统 TM, SSTM 的精确度提升 24 个百分点,重合率提升 21.9 个百分点。

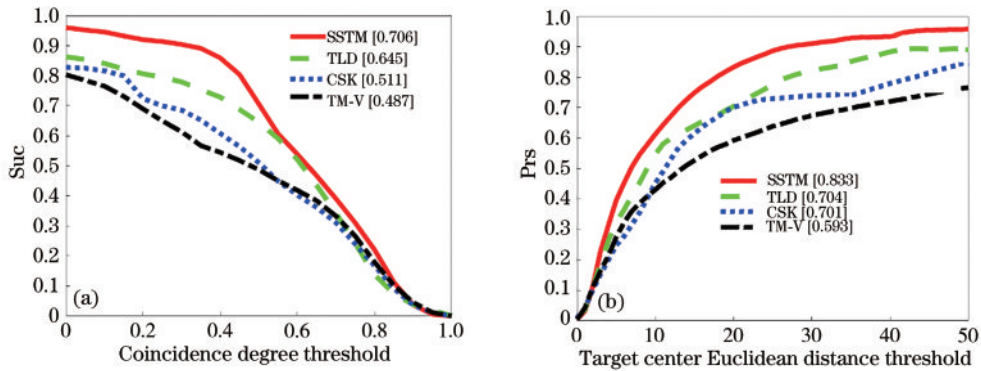


图 5 各跟踪算法的重合率和精确度。(a) 重合率; (b) 精确度

Fig. 5 Coincidence rate and accuracy of each tracking algorithm. (a) Coincidence rate; (b) accuracy

SSTM 算法和 TM 算法的跟踪表现如图 6 所示。从图 6(a)和图 6(b)可以看出:所提算法由于设置了尺度因子集,可以很好地应对目标尺度变化的场景;传统模板匹配跟踪算法由于检测框的尺寸始终不发生变化,在物体尺度变化较大时无法完整识别目标区域,甚至丢失跟踪目标。分析图 6(c)和图 6(d)可知:依靠模

板更新机制,SSTM 算法在背景信息和目标形态发生变化时依然具有一定的稳定性,同时图 6(c)测试集和图 6(d)测试集还受到了遮挡、背景、低分辨率等因素的影响,结果表明所提算法具有较高的普适性。图 6(e)测试集和图 6(f)测试集中的图片序列包含着剧烈的光照变化,SSTM 算法依然得到了不错的跟踪结果。

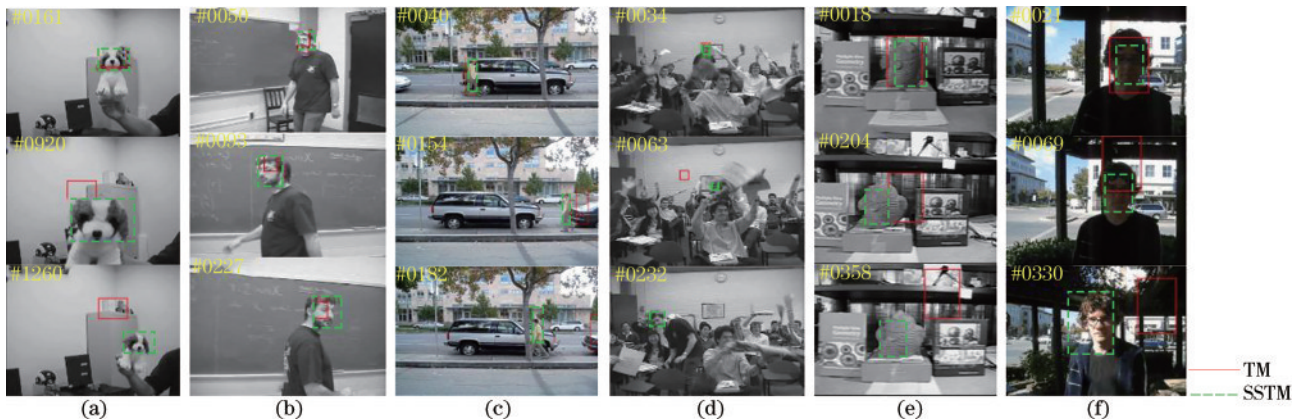


图 6 目标跟踪算法的跟踪结果。(a) dog 1; (b) freemen 1; (c) david 3; (d) freeman 4; (e) fish; (f) trellis

Fig. 6 Tracking results of target tracking algorithms. (a) dog 1; (b) freemen 1; (c) david 3; (d) freeman 4; (e) fish; (f) trellis

为进一步量化所提算法应对不同场景的能力,对 12 个测试集中涉及到的尺度变化(scale variations)、背景复杂(background clutter)、平面内旋转(in-plane rotation)、平面外旋转(out-plane rotation)、光照变化(illumination variations)、遮挡(occlusions)、变形(deformation)属性分别进行重合率(阈值取 0.5)和精确度(阈值取 25 pixel)分析,不同算法的结果如表 1 和表 2 所示。基于稀疏协同模型的鲁棒目标跟踪(SCM)算法利用全局模板和局部表示来建立模型,其模板更新机制同时考虑了原始模板和最新观测结果。在 AlexNet 等融合神经网络的目标追踪算法问世前,带内核的结构化输出跟踪(Struck)算法和 CSK 均为在 OTB 数据集上性能表现前五的算法。分析表 1 和表 2 可得:SCM 算法在各属性下表现均为最优;SSTM 与 Struck 算法表现略逊于前者,Struck 算法在背景复杂、光照变化、变形场景下表现不错,SSTM 算法在尺度变化、平面内外旋转、遮挡场景下性能较好;CSK 算法仅使用灰度空间难以有效利用完整信息,也无法处理尺度变化;TLD 算法作为长时目标追踪领域的代表性算法,在样本量较小时会影响追踪准确率。

表 1 不同算法在各影响因素下的精确度

Table 1 Precision of different algorithms under various influencing factors

Circumstance	TM	SSTM	SCM	CSK	TLD	Struck
Scale variations	0.596	0.888	0.900	0.689	0.680	0.832
Background clutter	0.586	0.765	0.919	0.872	0.581	0.925
In/Out-plane rotation	0.539	0.821	0.890	0.713	0.737	0.850
Illumination variations	0.415	0.742	0.904	0.668	0.863	0.958
Occlusions	0.616	0.783	0.832	0.704	0.610	0.753
Deformation	0.830	0.919	0.941	0.903	0.607	0.948

表 2 不同算法在各影响因素下的重合率

Table 2 Coincidence rate of different algorithms under various influencing factors

Circumstance	TM	SSTM	SCM	CSK	TLD	Struck
Scale variations	0.416	0.746	0.831	0.421	0.524	0.577
Background clutter	0.573	0.622	0.943	0.618	0.610	0.902
In/Out-plane rotation	0.402	0.707	0.823	0.530	0.603	0.623
Illumination variations	0.405	0.570	0.893	0.514	0.764	0.880
Occlusions	0.638	0.791	0.816	0.667	0.679	0.765
Deformation	0.816	0.772	0.988	0.632	0.679	0.961

在算法运行速度方面,通过卡尔曼滤波算法预测下一帧的目标中心将会大大减小搜索范围。表 3 为在尺度因子集设为  $\{0.95, 1, 1.05\}$  时,对 dog 1 测试集使用卡尔曼滤波前后的 SSTM 算法与不同算法的平均耗时对比。改进后的 SSTM 算法由于要对多个尺度

表 3 不同算法对 dog 1 测试集的耗时  
Table 3 Time consuming of different algorithms on dog 1 test set

Algorithm	Average time /ms	Speed / (frame·s <sup>-1</sup> )
SSTM without Kalman	67.48	15
SSTM with Kalman	19.91	50
TM	20.50	48
CSK	22.19	45
Struck	45.28	22
SCM	104.60	10
TLD	31.02	32

模板图像进行处理,与传统 TM 算法耗时差距不大;但相比于无卡尔曼滤波的 SSTM 算法,实时性有了成倍的提高;与 SCM 算法相比,在重合率与精确度表现相差不大的情况下,实时性大大得到提高。结合提出的应用场景可知,SSTM 算法在精确度、重合率和检测速度 3 个指标上的综合表现最佳。

### 3.3 真实场景模拟

在实际应用场景中,由于跟踪目标运动较为规律且形变较小,相机呈固定状态,会得到更好的跟踪效果。仿照应用场景,搭建了图 7 所示实验台以获取模拟真实情况的数据集。如图 7 所示,实验相机为分辨率为  $640 \times 480$ 、帧率为 30 frame/s 的固定三角支架相机,导向装置选为外形类似的平移导轮,木纹桌板模拟相似背景,缓慢匀速移动导轮的同时对导轮进行目标跟踪。

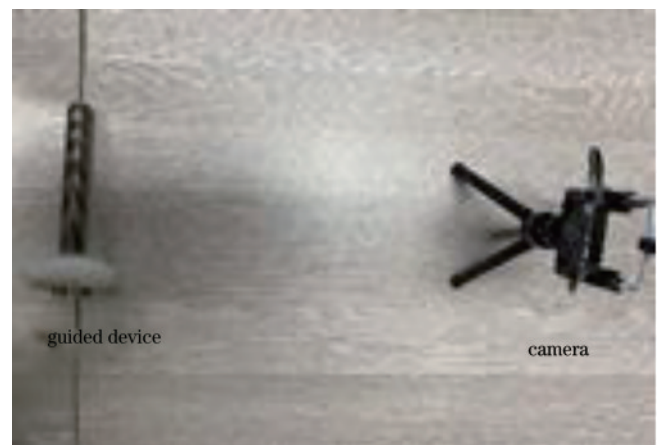


图 7 实验环境示意图

Fig. 7 Schematic of experimental environment

实验效果如图 8 所示,在导轮尺度有较大的变化时,传统 TM 算法几乎丢失目标,而改进后的 SSTM 算法依然能准确地定位跟踪。实验结果证明,相比于传统模板匹配跟踪算法,所提算法在多个场景下均有较大的性能提升。

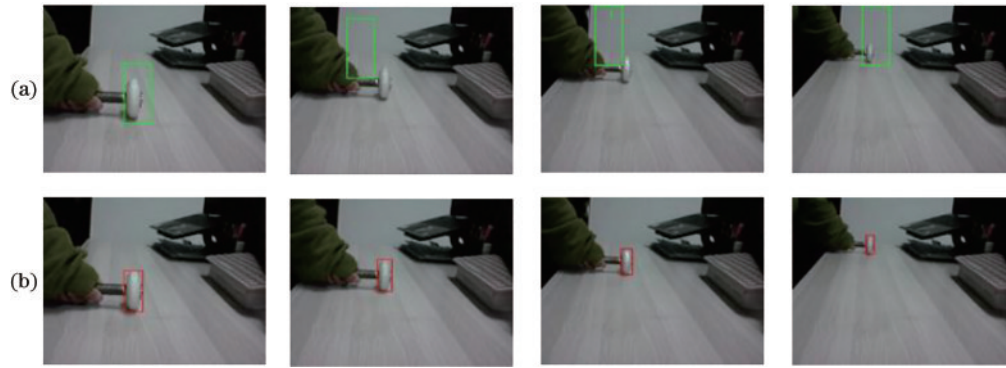


图 8 真实场景下目标跟踪算法的跟踪结果。(a)传统 TM 算法;(b)所提算法

Fig. 8 Tracking results of target tracking algorithms at real scene. (a) Traditional TM algorithm; (b) proposed algorithm

## 4 结 论

沿用传统模板匹配跟踪算法思想,针对其难以应对目标物体形态变化、运动变化等问题,建立模板图像金字塔以解决尺度变化问题。提出一种模板自适应更新机制提高跟踪准确率。利用卡尔曼滤波算法以减小匹配时的搜索区域,并和主流算法 TM、TLD、CSK、SCM、Struck 在多个测试集上进行性能对比。实验结果表明:SSTM 易于实现,精确度高,综合性能领先于其他算法;SSTM 在尺度变化、平面内外旋转、遮挡场景下表现好,应对背景复杂、光照变化场景时略有落后,需要引入更有效的目标表达特征,使相似度衡量更稳健;相比于传统模板匹配跟踪算法,SSTM 在 12 个测试集上的精确度提升 24 个百分点,重合率提升 21.9 个百分点。加权模板更新策略和图像金字塔的引入,使得模板图像具有尺度稳定性的同时更准确地表达出目标物体的实时运动情况,提高了跟踪的精确度和重合率。模板更新策略中的滑动平均思想避免算法学习到错误模板,一定程度上提高了算法的鲁棒性。针对工程领域中观测相机与目标物体间运动相对稳定的情况,引入卡尔曼滤波算法,根据上一次的最优结果来预测当前的目标物体中心坐标值,减小目标搜索区域,提高了算法的实时性。

## 参 考 文 献

- [1] 黄凯奇, 陈晓棠, 康运锋, 等. 智能视频监控技术综述[J]. 计算机学报, 2015, 38(6): 1093-1118.  
Huang K Q, Chen X T, Kang Y F, et al. Intelligent visual surveillance: a review[J]. Chinese Journal of Computers, 2015, 38(6): 1093-1118.
- [2] 朱文青, 刘艳, 卞乐, 等. 基于生成式模型的目标跟踪方法综述[J]. 微处理机, 2017, 38(1): 41-47.  
Zhu W Q, Liu Y, Bian L, et al. Survey on object tracking method base on generative model[J]. Microprocessors, 2017, 38(1): 41-47.
- [3] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using Mean shift[C]//Proceedings IEEE Conference on Computer Vision and Pattern Recognition,

June 15, 2000, Hilton Head, SC, USA. New York: IEEE Press, 2000: 142-149.

- [4] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409-1422.
- [5] Kalal Z, Matas J, Mikolajczyk K. Online learning of robust object detectors during unstable tracking[C]//2009 IEEE 12th International Conference on Computer Vision Workshops, September 27-October 4, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 1417-1424.
- [6] Held D, Thrun S, Savarese S. Learning to track at 100 FPS with deep regression networks[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 749-765.
- [7] Danelljan M, Robinson A, Shahbaz Khan F, 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.
- [8] 侯志强, 陈立琳, 余旺盛, 等. 基于双模板 Siamese 网络的鲁棒视觉跟踪算法[J]. 电子与信息学报, 2019, 41(9): 2247-2255.  
Hou Z Q, Chen L L, Yu W S, et al. Robust visual tracking algorithm based on Siamese network with dual templates[J]. Journal of Electronics & Information Technology, 2019, 41(9): 2247-2255.
- [9] Chen K, Tao W B. Once for all: a two-flow convolutional neural network for visual tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(12): 3377-3386.
- [10] Li B, Wu W, Wang Q, et al. SiamRPN: evolution of Siamese visual tracking with very deep networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 4277-4286.
- [11] 余铎, 王耀南, 毛建旭, 等. 基于视觉的移动机器人目标跟踪方法[J]. 仪器仪表学报, 2019, 40(1): 227-235.  
Yu D, Wang Y N, Mao J X, et al. Vision-based object tracking method of mobile robot[J]. Chinese Journal of

- Scientific Instrument, 2019, 40(1): 227-235.
- [12] 陈爱华, 孟勃, 朱明, 等. 多模式融合的目标跟踪算法[J]. 光学精密工程, 2009, 17(1): 185-190.  
Chen A H, Meng B, Zhu M, et al. Multi-pattern fusion algorithm for target tracking[J]. Optics and Precision Engineering, 2009, 17(1): 185-190.
- [13] Bibi A, Mueller M, Ghanem B. Target response adaptation for correlation filter tracking[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9910: 419-433.
- [14] 刘宗达, 董立泉, 赵跃进, 等. 视频中快速运动目标的自适应模型跟踪算法[J]. 光学学报, 2021, 41(18): 1815001.  
Liu Z D, Dong L Q, Zhao Y J, et al. Adaptive model tracking algorithm for fast-moving targets in video[J]. Acta Optica Sinica, 2021, 41(18): 1815001.
- [15] 纪元法, 何传骥, 孙希延, 等. 基于自适应特征融合与上下文感知的目标跟踪[J]. 激光与光电子学进展, 2021, 58(16): 1610011.  
Ji Y F, He C J, Sun X Y, et al. Object tracking based on adaptive feature fusion and context-aware[J]. Laser & Optoelectronics Progress, 2021, 58(16): 1610011.
- [16] Huang S L, Hong J X. Moving object tracking system based on camshift and Kalman filter[C]//2011 International Conference on Consumer Electronics, Communications and Networks (CECNet), April 16-18, 2011, Xianning, China. New York: IEEE Press, 2011: 1423-1426.
- [17] 牛轶峰, 伯晓晨, 沈林成. 基于可变模板的前视目标跟踪算法[J]. 国防科技大学学报, 2003, 25(5): 80-84.  
Niu Y F, Bo X C, Shen L C. A forward looking object tracking algorithm based on deformable template[J]. Journal of National University of Defense Technology, 2003, 25(5): 80-84.
- [18] Latecki L J, Mieziako R. Object tracking with dynamic template update and occlusion detection[C]//Proceedings of the 18th International Conference on Pattern Recognition-Volume 01, August 20-24, 2006, Hong Kong, China. New York: ACM Press, 2006: 556-560.
- [19] Oron S, Bar-Hillel A, Levi D, et al. Locally orderless tracking[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 1940-1947.
- [20] Masazade E, Fardad M, Varshney P K. Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks[J]. IEEE Signal Processing Letters, 2012, 19(12): 845-848.
- [21] Zhu B, Chang L B, Xu J N, et al. Huber-based adaptive unscented Kalman filter with non-Gaussian measurement noise[J]. Circuits, Systems, and Signal Processing, 2018, 37(9): 3842-3861.
- [22] 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.