

时空上下文学习长时目标跟踪

刘 威 赵文杰 李 成

空军航空大学航空航天情报系, 吉林 长春 130022

摘要 针对在线学习跟踪算法中目标模型更新错误而导致跟踪漂移的问题,提出了一种简单但高效的解决方案。在目标区域均匀采样点跟踪器,基于纹理描述对前后两帧点跟踪器进行置信度评估并以此完成目标初步定位,由多维特征时空上下文模型输出目标位置置信图以完成目标精确定位,同时结合置信图决定模型更新速率并给出了一种多尺度更新机制。实验表明,该方法在背景干扰、快速运动、遮挡、光照变化及尺度变化下均能完成稳健跟踪,在320 pixel×240 pixel的视频序列中平均跟踪速度为55.1 frame/s,可以满足实时应用的需求。

关键词 机器视觉; 目标跟踪; 时空上下文; 在线学习

中图分类号 TP391 **文献标识码** A

doi: 10.3788/AOS201636.0115001

Long-Term Visual Tracking Based on Spatio-Temporal Context

Liu Wei Zhao Wenjie Li Cheng

Department of Aerospace Intelligence, Aviation Univ of Air Force, Changchun, Jilin 130022, China

Abstract Aiming at the tracking drift problem due to object model update falsely in the online learning tracking algorithms, a simple but efficient solution is proposed. In the target area point trackers are uniformly sampled, which are assessed based on texture description in two consecutive frames point trackers and then the initial location of target is completed. Multi-dimensional feature spatio-temporal context model is used to output the precise position of object by the confidence map, the model update rate is decided combining with the confidence map and a multi-scale update mechanism is proposed. Experimental results show that the proposed algorithm can complete the robust tracking under the condition of background interference, fast motion, occlusion, illumination changing and scale changing. In the video sequence of 320 pixel×240 pixel, the average tracking speed can keep in 55.1 frame/s, which meets real-time application requirement.

Key words machine vision; object tracking; spatio-temporal context; online learning

OCIS codes 150.0155; 150.1135; 150.1488

1 引 言

目标跟踪在行为识别、视频监控、人机交互中的广泛应用,使其成为计算机视觉领域非常热门的研究课题^[1],而一个稳健的跟踪算法面临的主要挑战有:局部或全局遮挡、光照变化、姿态变化、尺度变化等,为此设计一个能够处理外观变化的目标模型则显得尤为重要。按照是否考虑背景信息,目标跟踪可划分为生成式跟踪算法和判别式跟踪算法。

判别式跟踪算法^[2-5]将跟踪构建为一个二元分类问题,需要采样大量样本完成分类器训练;生成式跟踪算法^[6-8]通过学习目标不同外观变化构建目标模型,抛弃了背景中的有用信息并大大提高了计算量。近年来自适应相关滤波技术^[9-10]在目标跟踪中的应用取得了显著成效,文献[11]提出时空上下文学习跟踪(STC)算法,同时考虑目标及周围环境信息,并利用前一帧目标位置通过滤波器输出当前帧目标位置置信图以完成目标定位并在线更新时空上下文模型,满足算法实时性的同时提高了跟踪精度。但上述算法存在的主要问

收稿日期: 2015-06-30; 收到修改稿日期: 2015-08-11

基金项目: 国家自然科学基金(61301233)

作者简介: 刘 威(1991—),男,硕士研究生,主要从事视觉跟踪和运动分析方面的研究。E-mail: 1224337250@qq.com

导师简介: 赵文杰(1968—),女,博士,教授,主要从事光学成像自动目标识别方面的研究。E-mail: actionlw@163.com

题有:1) 基于帧间目标运动较小的假设,在目标快速运动时导致跟踪失败;2) 在线模型采用固定速率更新,在目标被遮挡时易产生模型更新错误导致目标漂移。

针对上述问题,本文充分利用相关滤波技术的高效性,在目标区域均匀采样点跟踪器,据此完成目标初步定位后,由时空上下文模型输出当前目标位置置信图,完成目标精确定位,此外基于当前目标位置置信图来自适应更新时空上下文模型,避免了目标被遮挡时产生模型更新错误。

2 时空上下文跟踪器

文献[11]基于贝叶斯框架利用图像灰度特征构建目标及局部背景的相关关系,通过时空上下文模型计算目标位置置信图完成目标定位,给定当前帧目标位置 x^* ,其上下文信息定义为 $X^c = \{c(z) = [I(z), z] | z \in \Omega_c(x^*)\}$,式中 $I(z)$ 代表 z 处的像素值, $\Omega_c(x^*)$ 是当前目标局部区域,为此基于贝叶斯理论目标似然函数可构建为

$$c(x) = P(x|o) = \sum_{c(z) \in X^c} P[x, c(z)|o] = \sum_{c(z) \in X^c} P[x|c(z), o] P[c(z)|o], \quad (1)$$

式中 $c(x)$ 为目标置信图,条件概率 $P[x|c(z), o]$ 构建了目标与其上下文之间的空间关系, $P[c(z)|o]$ 是表征局部上下文的先验概率

$$P[c(z)|o] = I(z)w_\sigma(z - x^*), \quad (2)$$

式中 $I(z)$ 代表图像灰度特征, $w_\sigma(z - x^*)$ 是权重函数,定义为 $w_\sigma(d) = a \cdot \exp\left(-\frac{d^2}{\sigma^2}\right)$, a 是归一化参数, σ 是尺度参数。

条件概率 $P[x|c(z), o]$ 定义为

$$P[x|c(z), o] = h^{sc}(x - z). \quad (3)$$

目标置信图 $c(x)$ 定义为

$$c(x) = b \cdot \exp\left(-\left|\frac{x - x^*}{\alpha}\right|^\beta\right), \quad (4)$$

式中 b 是归一化系数, α 是尺度参数, β 是形状参数,文献[11]实验表明 $\beta = 1$ 时取得较好跟踪效果。结合(2)~(4)式,(1)式可以改写为

$$b \cdot \exp\left(-\left|\frac{x - x^*}{\alpha}\right|^\beta\right) = \sum_{c(z) \in X^c} h^{sc}(x - z) I(z) w_\sigma(z - x^*) = h^{sc}(x) \otimes I(x) w_\sigma(x - x^*), \quad (5)$$

式中 \otimes 代表卷积操作,为此(5)式可以通过快速傅里叶变换转换到频率域

$$\mathcal{F}[c(x)] = \mathcal{F}[h^{sc}(x)] \odot \mathcal{F}[I(x) w_\sigma(x - x^*)], \quad (6)$$

式中 $\mathcal{F}(\cdot)$ 代表快速傅里叶变换,经过傅里叶逆变换 $\mathcal{F}^{-1}(\cdot)$ 可以得到空间上下文模型 $h^{sc}(x)$

$$h^{sc}(x) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}[c(x)]}{\mathcal{F}[I(x) w_\sigma(x - x^*)]}\right\} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\left[b \cdot \exp\left(-\left|\frac{x - x^*}{\alpha}\right|^\beta\right)\right]}{\mathcal{F}[I(x) w_\sigma(x - x^*)]}\right\}. \quad (7)$$

定义 H_t^{sc} 为第 t 帧的时空上下文模型,文献[11]更新第 $t+1$ 帧的时空上下文模型 H_{t+1}^{sc} 的公式为

$$H_{t+1}^{sc} = (1 - \rho) H_t^{sc} + \rho h_t^{sc}. \quad (8)$$

得到第 t 帧时空上下文模型后,在第 t 帧目标位置置信图 $c_{t+1}(x)$ 中找到最大响应值即可完成跟踪任务

$$x_{t+1}^* = \arg \max_{x \in \Omega_c(x^*)} c_{t+1}(x) = \arg \max_{x \in \Omega_c(x^*)} \mathcal{F}^{-1}\left\{\mathcal{F}(H_{t+1}^{sc}) \odot \mathcal{F}[I_{t+1}(x) w_\sigma(x - x_t^*)]\right\}. \quad (9)$$

STC算法利用自适应滤波技术充分提高了算法实时性,但依然存在问题:(2)式表明只利用了图像灰度特征,抛弃了判别性较强的颜色纹理等信息;(9)式表明STC在前一帧目标位置确定的局部区域构建时空上下

文模型,一旦目标出现快速运动,算法将无法完成跟踪任务;(8)式表明STC采用固定学习率更新时空上下文模型,缺乏对遮挡等因素的判断易导致模型更新错误。

3 本文算法

本文算法采用点跟踪器完成目标初步定位,由多维特征时空上下文模型完成精确定位,并由多尺度滤波器完成尺度更新,算法框架图如图1所示。

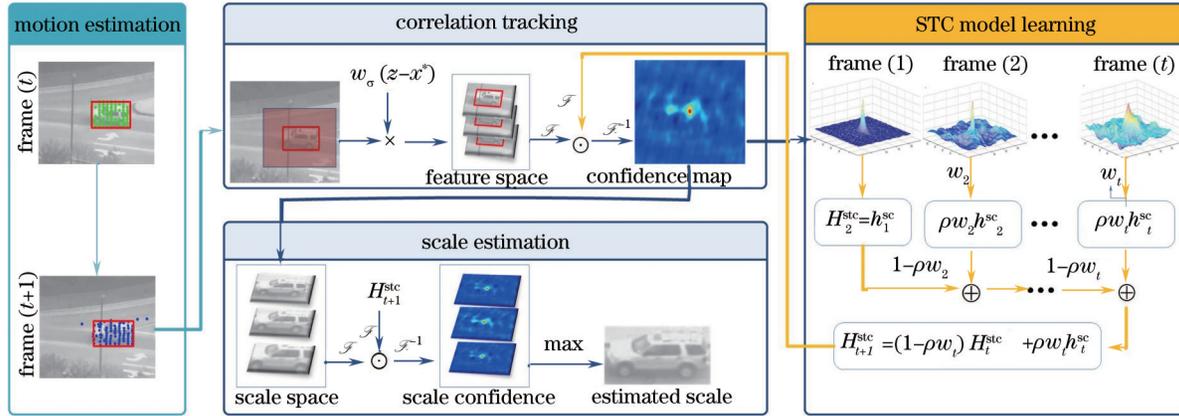


图1 本文算法框架图

Fig.1 Framework of the proposed diagram

3.1 运动估计

运动估计模块用以完成目标初步定位,给定第 t 帧目标框,类似中值流跟踪算法^[12]在框内均匀采样点跟踪器 $X = \{(x_1^t, x_{i+1}^t), (x_2^t, x_{i+1}^t), \dots, (x_n^t, x_{i+1}^t)\}$, 其中 n 为点跟踪器个数,对每个点跟踪器位置 x_i^t 利用 Lucas-Kanada 光流法^[13]预估其在第 $t+1$ 帧位置 x_{i+1}^t , 考虑到每个点跟踪器的结果并非可靠,采用 8 邻域局部相似度纹理描述算子^[14]来评估点跟踪器的状态 $s(x_i^t, x_{i+1}^t)$, 其中 x_i^t 的局部相似度定义为

$$f_{\text{lsn}}(x_i^t) = \sum_{i=0}^8 f(g_i - g_c, d), \quad f(x, d) = \begin{cases} 1, & |x| \leq d \\ 0, & |x| > d \end{cases} \quad (10)$$

式中 g_c 是 x_i^t 处像素值, g_i 是以 x_i^t 为中心的 8 邻域像素值, d 是局部相似度阈值,文献[14]中取为 2。若 $f_{\text{lsn}}(x_i^t)$ 与 $f_{\text{lsn}}(x_{i+1}^t)$ 差异大于阈值 $f_{\text{th_lsn}}$, 表明该点跟踪器的结果不可靠,应予以抛弃,因此定义点跟踪器的状态 $s(x_i^t, x_{i+1}^t)$ 为

$$s(x_i^t, x_{i+1}^t) = \begin{cases} 1, & \text{if } |f_{\text{lsn}}(x_{i+1}^t) - f_{\text{lsn}}(x_i^t)| < f_{\text{th_lsn}} \\ 0, & \text{else} \end{cases} \quad (11)$$

利用剩下可靠的 m 个点跟踪器来预测目标的帧间运动量。以极坐标形式定义点跟踪器的帧间运动量

$$p_i = (r_i, \theta_i) = \left[\sqrt{(d_x)^2 + (d_y)^2}, \quad a \tan 2(d_y/d_x) \right] \quad (12)$$

则目标中心运动量定义为各点跟踪器帧间运动量均值 $p_x = \left(\frac{1}{m} \sum_{i=1}^m r_i, \frac{1}{m} \sum_{i=1}^m \theta_i \right)$, 至此完成了目标在当前帧的初步定位。

3.2 多维特征滤波器跟踪

多维特征表征目标较单一特征具有更好的判别性,因此取得了不错的跟踪效果^[10,15]。假设图像 I 由 d 特征表示,定义特征维数 $l \in \{1, \dots, d\}$, I^l 表示第 l 维特征图像,则第 $t+1$ 帧目标位置可由最优化下式得到

$$x_{t+1}^* = \arg \max_{x \in \Omega_c(x^*)} c_{t+1}(x) = \arg \max_{x \in \Omega_c(x^*)} \mathcal{F}^{-1} \left\{ \mathcal{F} \left(\sum_{l=1}^d H_{t+1,l}^{\text{stc}} \right) \odot \mathcal{F} \left[\sum_{l=1}^d I_{t+1,l}^l(x) w_{\sigma}(x - x_t^*) \right] \right\}, \quad (13)$$

式中 $I_{t+1,l}^l(x)$ 代表第 $t+1$ 帧第 l 维特征图像, $H_{t+1,l}^{\text{stc}}$ 代表第 $t+1$ 帧第 l 维时空上下文模型,由第 t 帧第 l 维空间上下文模型 $h_{t,l}^{\text{sc}}$ 更新得到,与(8)式不同的是本文采用一种自适应更新机制, $h_{t,l}^{\text{sc}}$ 由第 t 帧置信图 $c_t(x)$ 及第 l 维

特征图像 $I'_i(x)$ 学习得到:

$$h_{i,l}^{sc} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}[c_i(x)]}{\mathcal{F}[I'_i(x)w_{\sigma_i}(x-x^*)]} \right\} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F} \left[b \cdot \exp \left(- \left| \frac{x-x^*}{\alpha} \right|^\beta \right) \right]}{\mathcal{F}[I'_i(x)w_{\sigma_i}(x-x^*)]} \right\}. \quad (14)$$

为充分利用图像中的颜色纹理信息并考虑到模型复杂性,采用受光照影响小的色度信息(HSV 颜色空间的 H 分量)和局部二值模式(LBP)纹理信息这二维特征来构建时空上下文模型,当然也可提取其他颜色纹理信息进行扩展。

3.3 尺度更新机制

传统跟踪算法加入尺度更新的同时也大大增加了算法耗时。基于相关滤波器的高效性,提出一种新的尺度更新机制:(2)式中的权重函数 $w_{\sigma}(z-x^*)$ 根据与目标中心的距离决定像素的权重,为此可以通过调整参数 σ 并分别进行滤波操作以获得最佳尺度估计。定义尺度空间 $S = \{a^{-n}, \dots, a^0, \dots, a^n\}$, 对每一个 $s_i \in S$, 有权重

函数 $w_{s_i,\sigma}(d) = a \cdot \exp \left[- \frac{d^2}{(s_i\sigma)^2} \right]$, 该尺度下得到的置信图为

$$c^s(x) = \mathcal{F}^{-1} \left\{ \mathcal{F}(H^{stc}) \odot \mathcal{F} [I(x)w_{s_i,\sigma_i}(x-x_i^*)] \right\}. \quad (15)$$

考虑到算法复杂性(15)式的时空上下文模型只由一维特征(图像灰度)决定,则最优尺度估计为

$$\hat{s} = \arg \max_{s_i \in S} \left\{ \max [c^s(x)] \right\}. \quad (16)$$

3.4 时空上下文模型更新

(8)式表明文献[11]采用固定学习率更新时空上下文模型,一旦目标被遮挡则出现模型更新错误,易导致跟踪漂移或当目标重新出现时无法再定位到目标。而实验表明每一帧得到的目标位置置信图峰值尖锐度^[9](PSR)是不一样的,以Suv序列为例,PSR分布情况如图2所示。

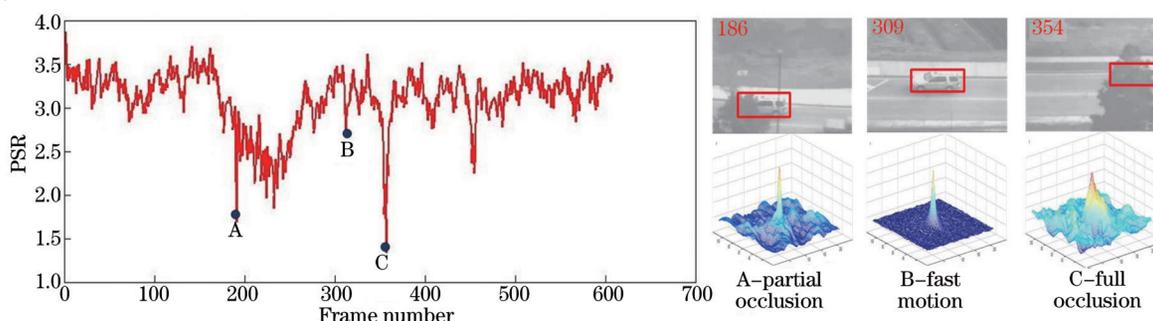


图2 Suv序列PSR分布及分析

Fig.2 Analysis and distribution of PSR in sequence of Suv

PSR大小反映了目标受遮挡、快速运动等因素影响的程度,实验表明目标在不同因素影响下PSR有较大区别,PSR值越大目标位置置信度越高,以图2为例:前180帧目标不受任何因素干扰时PSR值基本稳定在较高水平;第180~200帧当目标被局部遮挡时,PSR值逐步下降直至图2中的A点,当目标逐渐从局部遮挡中走出时PSR值逐步上升;而在第309帧左右目标出现快速运动时,PSR值又下降到谷点(图2中的B点),随着目标在第350~360帧被完全遮挡时,PSR值迅速下降到图中最低点C点。表明当前跟踪结果完全不可靠,时空上下文模型一旦更新势必带来更新错误。图中还发现在第450帧左右光照出现变化时,PSR值又下降到一个谷点,表明此时模型应该以一个较小的学习率进行更新,这样既避免带来更新错误,又能使模型能适应不同的场景和变化,增加了模型对不同因素影响的稳健性和适应能力。

图2的实验表明:当PSR值稳定在2.5~4时跟踪结果可信度较高,而低于2.5时目标可能受到遮挡、快速运动、光照变化等因素的影响,随着影响程度的增加PSR值也逐步降低。因此可以采用每一帧的PSR值来

决定空间上下文模型在更新过程中的权重,则时空上下文模型更新机制为

$$\begin{cases} H_{t+1}^{sc} = (1 - \rho w_t) H_t^{sc} + \rho w_t h_t^{sc} \\ w_t = \begin{cases} 0, & \text{if } R_t < R_{th_psr} \\ \exp[-(R_t - R_{max})^2], & \text{else} \end{cases} \end{cases}, \quad (17)$$

式中 R_t 代表第 t 帧 PSR 值,定义 $R_t = \frac{\max[c_i(x)] - \mu}{\sigma}$, 其中 μ 和 σ 分别是置信图峰值位置周围 11×11 邻域内的均值和标准差, R_{max} 表示当前所有帧中 PSR 的最大值,一旦 R_t 低于阈值 R_{th_psr} , 表明目标被完全遮挡, 应置第 t 帧空间上下文模型 h_t^{sc} 权重为 0, 模型不再更新。

4 实验与结果分析

实验采用具有不同挑战性的 6 段测试视频序列(共 4348 帧), 视频主要特性如表 1 所示。

表 1 实验视频的不同特性

Table 1 Different characteristics for experimental videos

Video	Frame number	Characteristic
Car	1300	background interference, fast movement
Motocross	207	fast movement, motion blur
Suv	609	full occlusion, partial occlusion
Faceocc2	812	partial occlusion, appearance changing
David	761	illumination changing, appearance changing
Car4	659	illumination changing, scale changing

为验证本文算法对不同特性视频的跟踪稳健性,与当前热门的以下跟踪算法进行对比实验: STC^[11]、中值流跟踪(MF)^[12]、压缩感知跟踪(CT)^[5]、检测学习跟踪(TLD)^[4]。实验平台采用 Intel Core(TM)2 Duo CPU 主频 2.5 GHz 的 PC 机, 利用 OpenCV 2.4.9 在 VS2010 上进行调试。所有算法均在第一帧给定框选目标位置后自动启动。

4.1 参数设置

运动估计中 $f_{th_lsn} = 2$, 多维特征滤波器跟踪中参照文献[11] $\alpha = 2.25, \beta = 1$, 尺度更新机制中 $a = 1.02, n = 1$ 。所以尺度空间共 3 个尺度, 参数 σ 初始化为 $\sigma_1 = \frac{s_h + s_w}{2}$, 式中 s_h 和 s_w 分别为初始目标框的高和宽, 尺度每 5 帧更新一次。图 2 的实验表明当低于 2.5 时目标基本处于完全遮挡状态, 故时空上下文模型更新中 $R_{th_psr} = 2.5$, ρ 与文献[11]一样设定为 $\rho = 0.075$ 。

4.2 结果分析

实验采用文献中常见的两个评估标准^[5]对算法性能进行测试。1) 跟踪成功率(SR), 定义 $S_{score} = \frac{R_{ROI_T} \cap R_{ROI_C}}{R_{ROI_T} \cup R_{ROI_C}}$ 代表每一帧跟踪结果得分, 式中 R_{ROI_T} 代表跟踪结果区域, R_{ROI_C} 代表目标真实区域, $S_{score} > 0.5$ 则认为跟踪成功; 2) 中心位置误差(CLE), 为 $\sqrt{(x_T - x_C)^2 + (y_T - y_C)^2}$, 其中 (x_T, y_T) 是跟踪结果的目标位置, (x_C, y_C) 是目标真实位置。

表 2 给出了算法在测试视频中的成功率对比(加粗代表最优性能, 斜体加下划线代表次优性能), 可见本文算法较其他算法在各类视频中获得了最优或次优性能, 平均跟踪成功率比 STC 算法提高了 21.3%。此外对于目标跟踪实际应用, 算法实时性也是重要的评估指标, 表 3 给出了算法在不同视频序列上的单帧处理时间对比, 并在最后一列给出了算法的平均跟踪速度, 可见本文方法单帧耗时低于 CT、TLD 算法, 而较 STC 算法有所降低。这是因为算法采用运动估计模块并采用多维特征构建时空上下文模型, 且构建尺度空间进行了尺度估计, 但 55.1 frame/s 的平均跟踪速度完全满足实时性的要求。

表2 跟踪成功率对比

Table 2 Comparison of SR

Sequence	Proposed	STC	MF	CT	TLD
Car/%	90.2	<u>88.9</u>	15.1	49.5	58.2
Motocross/%	98.0	77.2	31.5	<u>85.3</u>	73.2
Suv/%	100	40.3	30.2	39.1	<u>68.2</u>
Faceocc2/%	<u>86.3</u>	41.1	56.9	88.0	69.0
David/%	<u>88.7</u>	93.5	60.8	61.6	77.2
Car4/%	93.2	<u>87.4</u>	22.3	48.1	59.3
Average SR/%	92.7	<u>71.4</u>	37.1	61.9	67.5

表3 各算法单帧耗时和平均跟踪速度

Table 3 Time consuming of each algorithm and average tracking speed

Sequence	Proposed	STC	MF	CT	TLD
Car /ms	19.0	15.9	8.9	29.6	200.5
Motocross /ms	18.4	15.6	8.5	29.1	197.6
Suv /ms	17.7	14.8	7.5	28.4	170.1
Faceocc2 /ms	17.4	14.5	6.9	28.0	171.2
David /ms	17.2	14.3	7.0	27.9	168.4
Car4 /ms	18.1	15.2	7.8	28.5	181.8
Average tracking speed /(frame/s)	55.1	<u>65.7</u>	128.3	35.2	5.5

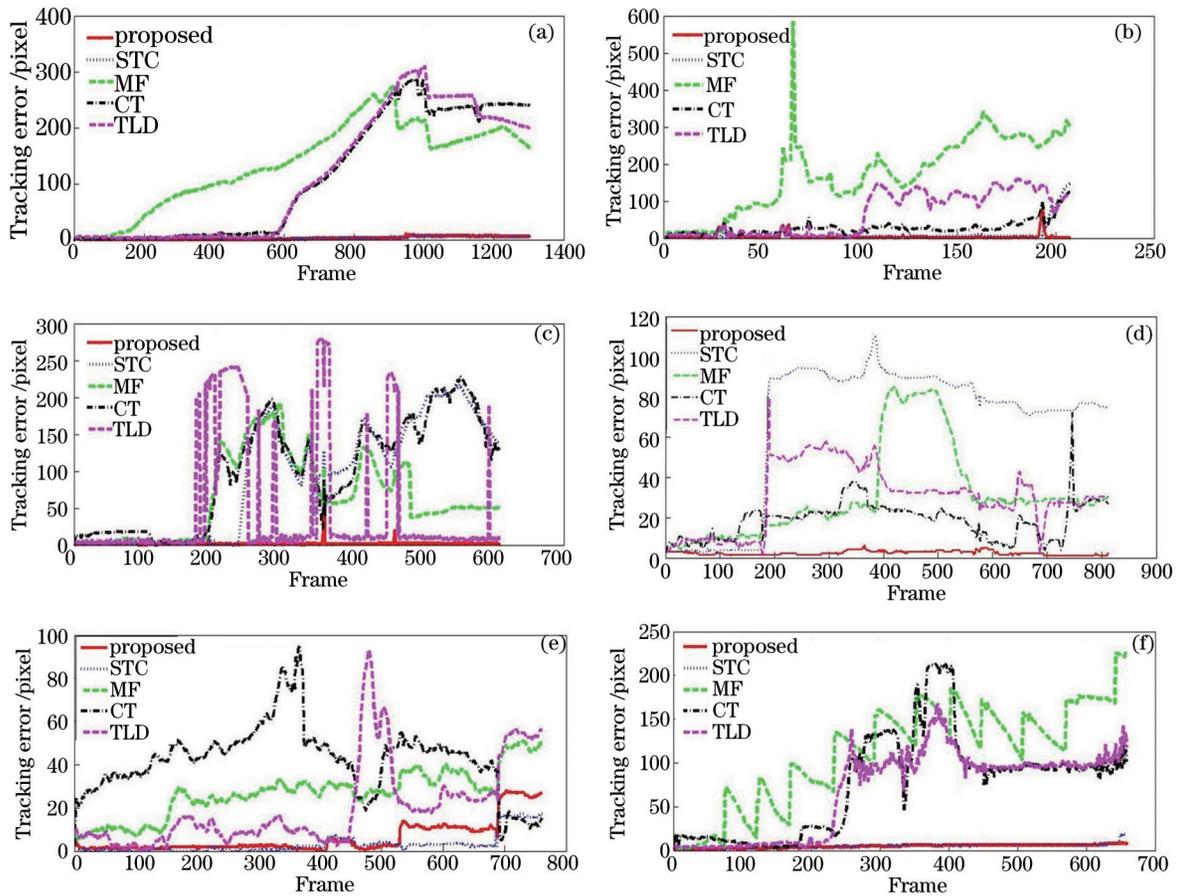


图3 视频序列跟踪误差

Fig.3 Tracking error of video sequences

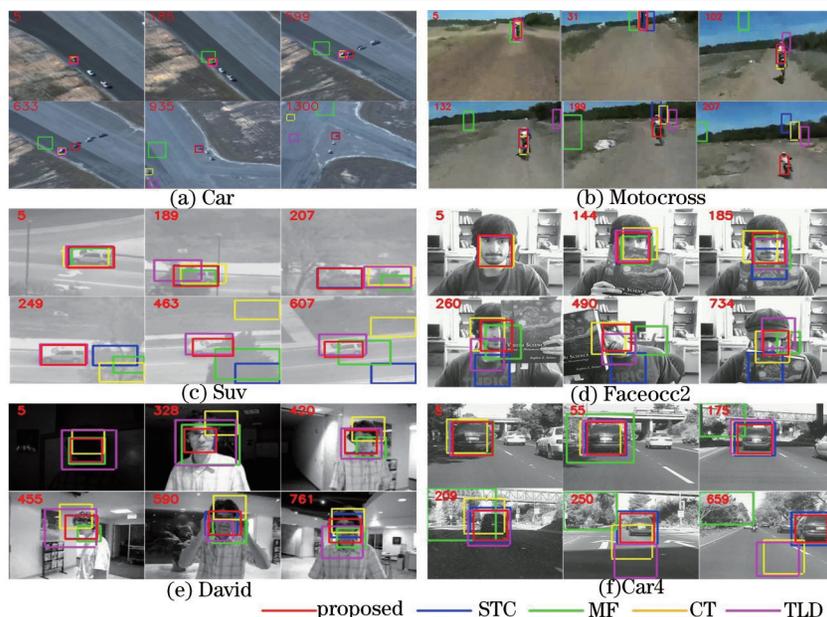


图 4 不同算法跟踪结果

Fig.4 Tracking results of different algorithms

图 3 给出了不同算法的中心位置误差对比,可见本文算法在各类视频中跟踪误差基本保持在 20 pixel 以内,图 4 给出了不同算法的跟踪结果(左上角红色数字代表帧序数),下面结合图 4 针对视频不同特性对本文算法的稳健性进行分析。

1) 背景干扰:图 4(a)Car 序列显示了本文算法的抗背景干扰能力,任务是跟踪机场的红色小车,目标尺寸较小且场景内多辆小车与目标极为相似。同时航拍平台的剧烈运动导致运动模糊的出现, MF 算法在第 185 帧时由于运动模糊导致中值流误差较大出现跟踪失败,目标尺寸过小 CT 算法难以提取足够判别信息而在第 599 帧跟踪到相似目标上(白色小车)。同样 TLD 算法采取全局搜索检测机制也在第 633 帧跟踪到了干扰目标,只有本文算法和 STC 算法全程完成了跟踪任务,体现了时空上下文模型的高效性。

2) 快速运动和运动模糊:图 4(b)Motocross 序列中目标运动速度较快且摄像平台的剧烈运动也导致运动模糊的出现, MF 算法因误差累积在第 31 帧即跟踪失败。TLD 算法的检测机制在剧烈运动模糊干扰下也基本失效, CT 算法和 STC 算法都是在前一帧目标位置的局部区域内搜索目标,因而在第 199 帧当目标由飞跃状态落地时出现跟踪框漂移,而本文算法的全程稳健跟踪充分验证了运动估计模块的优越性,避免了由目标快速运动带来的跟踪漂移。

3) 全局和局部遮挡:图 4(c)Suv 序列和图 4(d)Faceocc2 序列显示了本文算法的抗遮挡能力。Suv 序列多次出现遮挡:第 185~245 帧长时遮挡、第 345~365 帧和第 450~465 帧短时遮挡。STC 算法和 CT 算法在遮挡出现时依然采用同样速率更新在线模型,图 3(c)的跟踪误差曲线显示 200 帧以后这三类算法误差偏大,当目标第 246 帧重新出现时无法继续完成跟踪任务, TLD 算法能重新检测到目标,而本文算法一旦判断遮挡出现将不再更新时空上下文模型,在 Faceocc2 序列中目标多次出现局部遮挡和姿态变化时也取得了稳健的跟踪效果。

4) 光照和外观变化:图 4(e)David 序列显示了本文算法的抗光照变化和物体形变能力。序列中光照由弱到强,目标在第 420、455、590 帧发生较大姿态变化。本文算法采用目标位置置信图峰值尖锐度来调整模型更新速率,能不断学习到目标外观变化,而且本文算法采用了不受光照变化影响的色度信息和纹理信息来构建时空上下文模型,具有较高的判别性,平均跟踪误差较 STC 算法低 1.23 pixel。

5) 尺度变化:图 4(f)Car4 序列中目标尺度变化较大, MF 算法完全依据光流跟踪的结果进行尺度更新在第 55 帧即出现尺度更新错误, CT 算法无尺度更新在目标第 209 帧从桥下阴影离开时由于在线模型错误累积也导致跟踪失败, STC 算法在跟踪过程中表现稳定,本文算法的基于时空上下文模型的尺度更新机制获得了最佳尺度估计,因此本文算法跟踪成功率较 STC 算法提高了 5.8%。

5 结 论

采用运动估计模块避免了快速运动情况下的目标丢失,基于相关滤波的高效性构建多维特征时空上下文模型并依据目标状态自适应更新模型,避免了模型更新错误而导致跟踪漂移且提高了模型的自适应能力,同时在时空上下文模型的基础上构建尺度空间获得了最佳尺度估计。实验表明:本文方法在背景干扰、快速运动、遮挡、光照变化及尺度变化下均能获得稳健的跟踪结果,55.1 frame/s的平均跟踪速度可以满足实时性的要求。但当目标外观变化较大时,算法跟踪性能有所降低,可以结合现有的重检测机制加以改进,这也是后续的研究重点。

参 考 文 献

- 1 Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411–2418.
- 2 Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking[C]. ECCV, 2008, 5302: 234–247.
- 3 Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619–1632.
- 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 Zhang K H, Zhang L, Yang M H. Fast compressive tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(10): 864–877.
- 6 Ross D A, Lim J, Lin R S, *et al.*. Incremental learning for robust visual tracking[C]. International Journal of Computer Vision, 2008, 77(1): 125–141.
- 7 Mei X, Ling H B. Robust visual tracking and vehicle classification via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(11): 2259–2272.
- 8 Li H X, Shen C H, Shi Q F. Real-time visual tracking using compressive sensing[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011: 1305–1312.
- 9 Bolme D S, Beveridge J R, Draper B A, *et al.*. Visual object tracking using adaptive correlation filters[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 2544–2550.
- 10 Henriques J F, Caseiro R, Martins P, *et al.*. Highspeed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583–596.
- 11 Zhang K H, Zhang L, Liu Q S, *et al.*. Fast visual tracking via dense spatio-temporal context learning[C]. ECCV, 2014, 8693: 127–141.
- 12 Kalal Z, Mikolajczyk K, Matas J. Forward-backward error: Automatic detection of tracking failures[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 23–26.
- 13 Lucey S, Navarathna R, Ashraf A, *et al.*. Fourier lucas-kanade algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1383–1396.
- 14 Tavakoli H R, Moin M S, Heikkila J. Local similarity number and its application to object tracking[J]. International Journal of Advance Robotic Systems, 2013, 10(1): 184–190.
- 15 Danelljan M, Khan F S, Felsberg M. Adaptive color attributes for real-time visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 23–28.

栏目编辑: 苏 岑