

文章编号 1004-924X(2015)07-2093-07

视频监控系统中的概率模型单目标跟踪框架

李静宇*, 刘艳滢, 田睿, 王延杰, 姜瑞凯

(中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033)

摘要:针对视频监控的特点与跟踪目标的强机动性,提出了一种新的基于概率模型的目标跟踪框架,从目标表观模型、系统动态模型以及系统观测模型 3 个方面对当前标准的粒子滤波目标跟踪方法进行了改进。首先,考虑人眼细胞的分布特点,基于人眼分布结构建立目标表观模型来提高跟踪系统抵抗局部遮挡的能力;然后,建立基于自适应目标运动的系统动态模型,提高跟踪算法对快速机动目标的鲁棒性;最后,采用实时更新的系统观测模型,有效避免目标在遇到遮挡、光照变化、剧烈变形等情况下发生的跟踪漂移现象。实验结果表明,本文算法的正确跟踪率可达 98%;平均跟踪误差小于 6 个像元。实验证明本文算法在保证系统跟踪精度要求的同时,具有计算量小、抗干扰能力强等特点。

关键词:视频监控;概率模型;目标跟踪;表观模型;实时更新

中图分类号:TN941.1 **文献标识码:**A **doi:**10.3788/OPE.20152307.2093

Probabilistic model single target tracking framework for video surveillance system

LI Jing-yu*, LIU Yan-ying, TIAN Rui, WANG Yan-jie, JIANG Rui-kai

(Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China)

* Corresponding author, E-mail:18686653384@qq.com

Abstract: According to characteristics of video surveillance and strong mobility of tracking targets, a novel target tracking framework based on a probability model was proposed. The current standard particle filtering target algorithm was improved based on a target appearance model, a systemic dynamic model, and a systemic observation model. Firstly, the target appearance model was established by taking the distribution of human eye cell into account to improve its resistance capability for partial occlusion of local occlusion. Then, the systemic dynamic model based on the adaptive target movement was built to improve the robustness of tracking framework for the fast moving target. Finally, the systemic observation model with online update was established to prevent the tracking shift when the target faced the occlusion, illumination changes, severe deformation, etc., effectively. Experimental results show that the proposed algorithm achieves 98% of correct tracking rate, and the average tracking error is less than 6 pixels. The proposed method satisfies the video surveillance system

收稿日期:2014-12-15;修订日期:2015-02-10.

基金项目:国家 863 高技术研究发展计划资助项目(No. 2014AA7031010B)

requirements for stabilization, reliability, higher precision, less computing cost, as well as strong anti-jamming.

Key words: video surveillance; probabilistic model; target tracking; appearance model; online update

1 引言

随着视频监控规模日益扩大,通过安保人员进行人工实时监控的局限性也越发突出。无论是观看实时视频流,还是观看回放录像,由于人类自身生理上的弱点,经常会因为过度疲劳导致漏报现象。另外,由于传统视频监控系统无法有效地对存储录像数据进行分类,也给数据分析工作带来了很大的难度。因此,监控系统的智能化成为当前的研究热点。智能监控系统与传统监控系统的研究目标有所不同,快速的运动目标识别、分类与跟踪算法是这一研究课题的关键。

按照模型类别,运动目标跟踪方法可划分为基于确定性模型和基于概率模型的方法。与确定性算法相比,以粒子滤波技术为代表的概率模型算法在对非线性问题的处理能力、对背景干扰的鲁棒性以及算法的灵活性等方面均有着明显的优势,因此得到国内外学者的广泛研究^[1-8]。

Isard 和 Blake^[9]首次将粒子滤波算法应用到视频跟踪,以解决非线性、非高斯分布条件下的递归状态估计问题。在此基础上,王书朋等^[10]提出了一种根据目标运动特征的自适应目标跟踪算法。该算法利用粒子数与运动估计准确度的关系,自动确定粒子滤波所需粒子数及其分布,提高了粒子的使用效率,避免了由于使用过多粒子而增加跟踪器计算量的问题。但由于其采用的是固定目标模板,当目标发生旋转、变形等变化时,目标跟踪易发生漂移现象。Ross 等^[11]提出了一种增强的实时更新目标子空间的跟踪方法(IVT),大量实验验证显示,当目标发生大尺度缩放、旋转以及光照变化时,其仍能实现稳定跟踪。但由于该算法采用简单的随机游走方式的系统动态模型,仅适用于目标匀速运动,当目标加速或减速前进以及发生严重遮挡时,跟踪效果不理想。

在上述研究基础上,本文提出一种基于概率模型的目标跟踪算法,主要从目标外观模型、系统动态模型以及系统观测模型 3 个方面对当前标

准的粒子滤波的目标跟踪方法进行改进。首先,利用人眼细胞的分布特点,采用基于人眼分布结构建立目标外观模型描述方法,提高了跟踪系统抵抗局部遮挡的能力;然后,建立基于自适应目标运动的动态系统模型,以提高算法对快速机动目标的跟踪鲁棒性;最后,采用实时更新的系统观测模型,当目标遇到遮挡、光照变化、剧烈变形等情况时,可以有效避免跟踪漂移现象的发生。实验证明本文算法在保证系统跟踪精度要求的同时,具有快速高效、自适应性强等特点。

2 视频监控中的概率模型目标跟踪框架设计

2.1 基于人眼分布结构的目标外观模型描述

锥体细胞是人眼中主要的光感受器,其功能是接收外部的光信号。与眼球中心距离越近的区域,锥体细胞分布越密集,与眼球中心距离越远,锥体细胞分布的密度指数级越低。这一生理结构使得人类可以将注意力集中到某一物体,并能够用余光观测其他物体。

本文利用人眼分布结构建立了目标外观模型描述方法,即与目标中心距离越近的区域,采样频率越密集,与中心越远的区域,样本采样越稀疏。目标外观模型描述如图 1 所示。

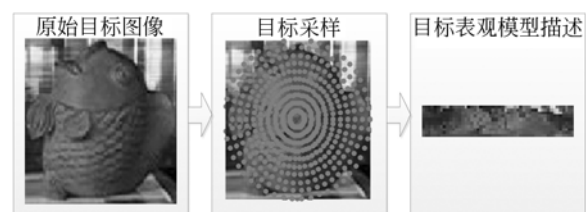


图 1 基于人眼分布结构的目标外观模型描述图

Fig. 1 Schematic diagram of object appearance model based on human eye distribution structure

基于人眼分布结构的目标外观模型描述算法实现的伪代码为:

```

输入:目标图像:img_i
    生成样本数:simple_num
    样本离散程度:k
输出:样本集合:img_o
[height width]=size(img);
for I=1:k
    W(I)=(width*(k+1-I)/k)/2;
    H(I)=(height*(k+1-I)/k)/2;
end
for I=1:k
    n=simple_num/k;
    for J=1:n
        x=W(I)*cos(2*pi*J/n)+width/2;
        y=H(I)*sin(2*pi*J/n)+height/2;
        img_o(I,J)=img_i(y,x);
    end
end
    
```

2.2 基于自适应目标运动的系统动态模型的建立

目标在视频图像中的变化一般可用仿射变换来描述,即:

$$\begin{bmatrix} u_{t+1} \\ v_{t+1} \end{bmatrix} = \begin{bmatrix} a_{1t} & a_{2t} \\ a_{3t} & a_{4t} \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} + \begin{bmatrix} T_{xt} \\ T_{yt} \end{bmatrix}, \quad (1)$$

其中: (u_t, v_t) 为 t 时刻目标在图像中的位置坐标; (T_{xt}, T_{yt}) 为目标平移参数; $(a_{1t}, a_{2t}, a_{3t}, a_{4t})$ 为目标旋转和缩放参数。因此目标在 t 时刻的状态可表示为:

$$s_t = (a_{1t}, a_{2t}, a_{3t}, a_{4t}, T_{xt}, T_{yt}).$$

假设目标运动的动态分布函数为 $p(s_t | s_{t-1})$, 通常情况下它可以用如下的线性随机差分方程来表示:

$$s_t = Cs_{t-1} + D_{t-1} w_{t-1}, \quad (2)$$

其中: $w_{t-1} \sim N(0,1)$ 为高斯分布的系统随机噪声; C 通常是常量; D_{t-1} 为粒子传播距离, 它的取值直接影响算法的跟踪性能。

为了提高算法的鲁棒性, 本文拟采用速度自适应模型更新粒子传播距离 D_{t-1} , 其分量表示为:

$$\bar{s}_t = \frac{1}{k} \sum_{n=t-k}^t |s_n - s_{n-1}|, \quad (3)$$

$$D_t^i = \alpha \bar{s}_t^i, \quad (4)$$

其中: \bar{s}_t^i 表示维度 i 上, 前 k 帧的平均速度, D_t^i 表示维度 i 上粒子的传播距离, α 为常数。

为进一步提高算法的运算效率, 利用粒子数量与粒子传播距离之间的关系自适应确定粒子数目^[10]:

$$N_t = N_{\min} + (N_{\max} - N_{\min}) \left(\frac{2}{1 + \exp(-\beta E)} - 1 \right)$$

$$E = \bar{s}_t^5 \bar{s}_t^6$$

其中: N_t 为 t 时刻的粒子数, N_{\min} 和 N_{\max} 分别为最小和最大粒子数; β 为常数。

2.3 系统观测模型的实时更新

实际被跟踪的目标特征是随时间不断变化的, 那么采用固定观测模型进行长时间跟踪就会导致目标丢失, 因此需要对系统观测模型进行实时更新。

目标在连续的视频序列中一般表现出较强的相关性, 因此目标图像的观测值 I_t 可以通过目标图像均值 μ 和子空间特征向量集 U 表示, 即:

$$p(I_t | s_t) = N(I_t; \mu, UU^T + \epsilon I), \quad (5)$$

其中: I 为单位阵, ϵ 为系统的观测噪声。

子空间特征向量集 U 的迭代求解可通过对数据流向量进行奇异值分解(SVD)获得。假设初始时刻到 t 时刻的视频序列数据流为:

$$A = [I_1 - \mu_t \quad I_2 - \mu_t \quad \cdots \quad I_t - \mu_t]$$

该数组的每一列为序列图像按列重新排序生成的列向量与 t 时刻图像均值 μ_t 的差, 对该数组进行奇异值分解(SVD)结果为 $A = UDV^T$ 。那么, t 时刻到 $t+k$ 得到的新输入的图像流为:

$$B = [I_{t+1} - \mu_t \quad I_{t+2} - \mu_t \quad \cdots \quad I_{t+k} - \mu_t]$$

将新输入的图像流与原数据流合并, 假设新的数据流 SVD 结果为:

$$[A \quad B] = U'D'V'^T, \quad (6)$$

进一步定义 \tilde{B} 为 B 在 U 中正交的部分, 即 $\tilde{B} = B(I - UU^T)$, 且有 $\tilde{B}\tilde{B}^T + UU^T = I$, 则有:

$$[A \quad B] = [U \quad \tilde{B}] \begin{bmatrix} D & U^T B \\ \mathbf{0} & \tilde{B}^T B \end{bmatrix} \begin{bmatrix} V^T & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}. \quad (7)$$

令

$$R = \begin{bmatrix} D & U^T B \\ \mathbf{0} & \tilde{B}^T B \end{bmatrix} = \tilde{\tilde{\tilde{U}}} \tilde{\tilde{\tilde{D}}} \tilde{\tilde{\tilde{V}}}^T$$

是一个 $(t+k) \times (t+k)$ 的矩阵, 将其带入式(7), 得到:

$$[\mathbf{A} \ \mathbf{B}] = ([\mathbf{U} \ \tilde{\mathbf{B}}]\tilde{\mathbf{U}})\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T \begin{bmatrix} \mathbf{V}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (8)$$

因此有:

$$\mathbf{U}' = [\mathbf{U} \ \tilde{\mathbf{B}}]\tilde{\mathbf{U}}, \mathbf{D}' = \tilde{\mathbf{D}}, \mathbf{V}'^T = \tilde{\mathbf{V}}^T \begin{bmatrix} \mathbf{V}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (9)$$

图像均值 $\boldsymbol{\mu}$ 的更新表示为:

$$\boldsymbol{\mu}' = \frac{\boldsymbol{\mu}_t + \sum_{i=t}^{t+k} \mathbf{I}_i}{t+k}. \quad (10)$$

观察式(6)~式(9)可知,由于对整个矩阵的SVD可归结为对 \mathbf{R} 矩阵进行SVD,故上述迭代过程的计算量与时间无关。

2.4 极大似然估计

本文采用极大似然方法获取当前时刻的目标状态 s_t 。根据系统动态模型进行随机抽样,对于每一个样本状态 s_i ,计算其似然概率:

$$p(\mathbf{I}_t | s_i) \propto \exp\left[-\frac{1}{2}(\mathbf{I}_t - \boldsymbol{\mu}_i)^T(\mathbf{U}\mathbf{U}^T + \boldsymbol{\epsilon}\mathbf{I})^{-1}(\mathbf{I}_t - \boldsymbol{\mu}_i)\right]$$

似然概率取得最大值时的样本状态将作为新的目标状态,即 $s_t = \arg \max_i p(\mathbf{I}_t | s_i)$ 。

2.5 跟踪框架流程

视频监控中的概率模型的目标跟踪框架流程如图2所示。

a)初始化目标区域,包括目标中心点坐标 (x_0, y_0) ,目标大小 (w_0, h_0) ,并按照2.1节方法进行目标表观建模。目标表观建模结果按列排序并归一化后将作为初始时刻的目标子空间 \mathbf{U} ,初

始化状态参数 $s_0 = (1, 0, 0, 1, 0, 0)$; b)采集图像 \mathbf{I}_t ; c)按照式(2)建立自适应系统动态模型,采用自适应粒子数目进行采样; d)对每一个粒子样本进行目标表观建模; e)按照2.4节方法估计新的状态参数 s_t ; f)利用新的状态参数 s_t 对原目标区域作变换,更新目标区域,根据式(9)~式(10)更新目标子空间 \mathbf{U}' 与图像均值 $\boldsymbol{\mu}'$,回到步骤b),继续进行目标跟踪。

3 实验结果与分析

3.1 实验结果

为了验证本文算法的有效性,采用多伦多大学数据库中用于测试目标跟踪的视频作为实验数据,计算机CPU为Intel Core™ i5处理器,主频为2.5 GHz,内存为4 G,实验平台为MATLAB R2013b。

实验中取 $N_{\min} = 100$, $N_{\max} = 600$, $\alpha = 1.2$, $\beta = 0.02$, $C = 1$ 。

目标中心点初始化坐标和大小分别为(188, 192)和(110, 130),初始状态参数 $s_0 = (1, 0, 0, 1, 0, 0)$,初始粒子数目为100。首先采集图像 \mathbf{I}_0 ,对目标区域进行表观建模,将结果按列排序并归一化,得到初始时刻目标子空间 \mathbf{U} ,然后根据系统动态模型生成采样粒子,对每一个粒子样本进行表观建模并进行似然估计,得到新的状态参数 s_t ,更新目标区域与目标子空间 \mathbf{U}' 。接着采集下一幅图像 \mathbf{I}_t ,如此循环。

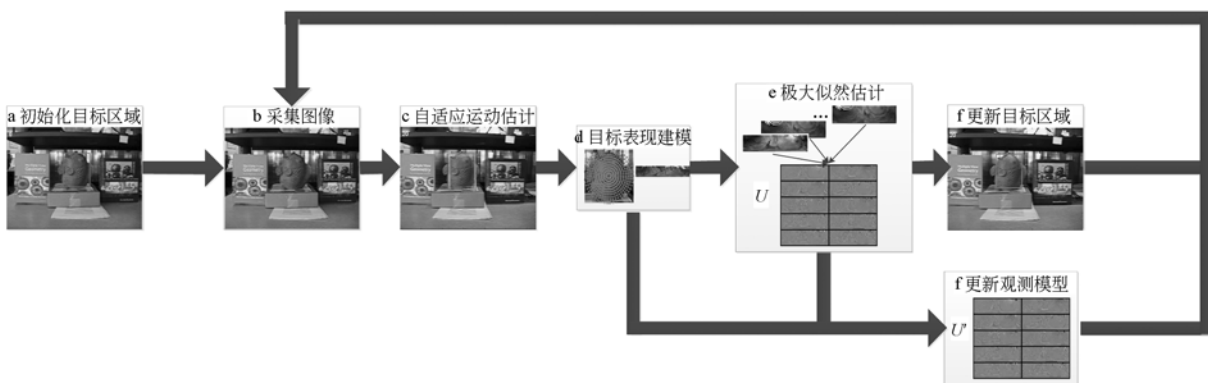


图2 跟踪框架流程

Fig. 2 Flow chart of tracking framework

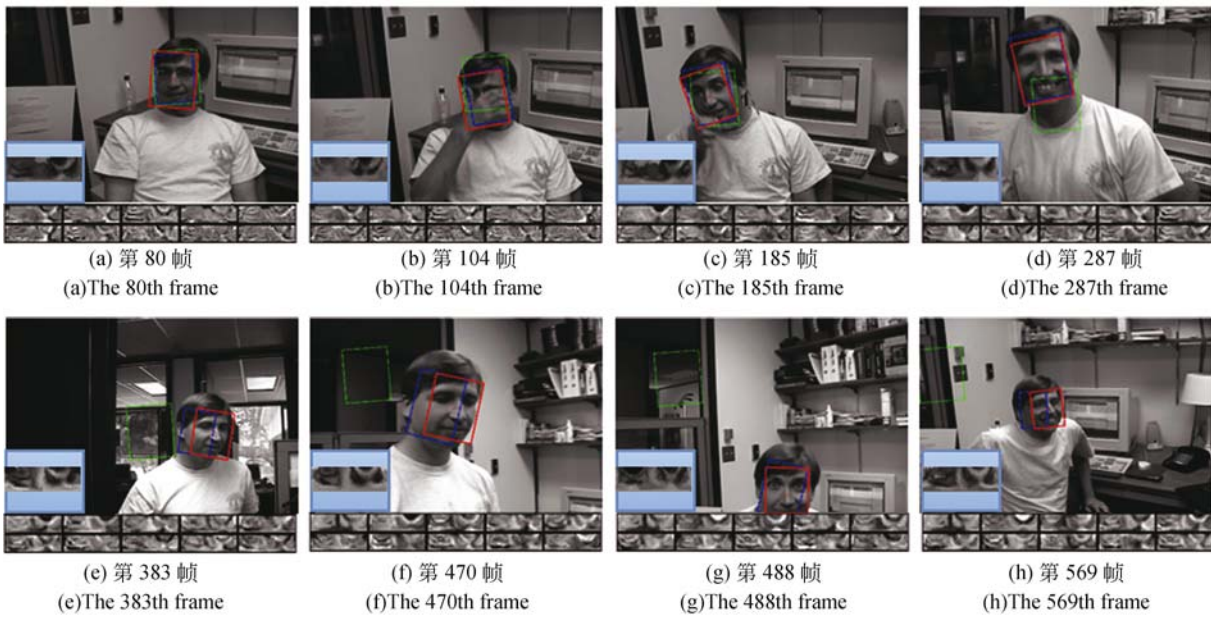


图 3 跟踪实验结果

Fig. 3 Results of tracking experiment

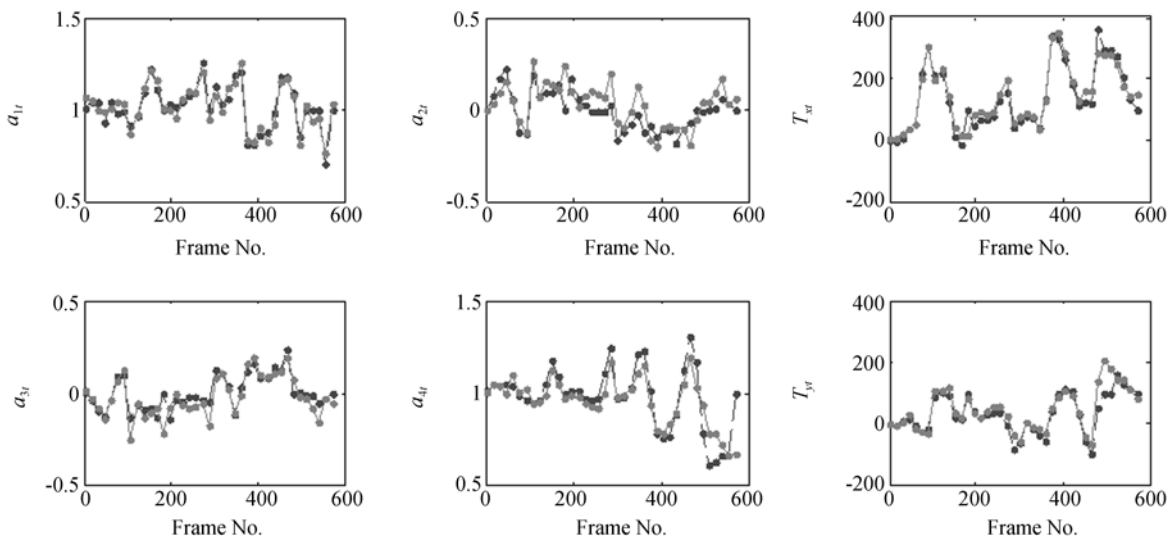


图 4 实际目标状态参数轨迹与跟踪状态参数轨迹

Fig. 4 Parameter trajectories and tracking trajectories for real target state

图 3(彩图见期刊电子版)为对室内人脸图像的跟踪结果。每幅图像包含 3 个部分,上面的主图为跟踪效果图,主图左下角蓝色区域内的辅图为更新后目标区域的表观建模结果。最下面连续的 10 幅辅图为更新的目标子空间,代表目标子空间矩阵的前 10 列。在整个跟踪过程中,目标经历了严重遮挡(第 104、488 帧),表情变化(第 80、287、488、569 帧)以及光线变化(第 185、383、470 帧)等过程,本文算法仍能实现稳定跟踪。

3.2 实验分析

下面对上述图像跟踪结果进行分析。图 3 的主图对比分析了 3 种不同跟踪算法的实验结果。其中,红色实线方框为本文算法结果,绿色虚线方框为文献[10]的跟踪方法结果,蓝色点划线方框为 IVT 算法结果。从图中可以看出,当目标发生表情变化(第 287 帧)后,文献[10]的跟踪方法失效;IVT 算法大部分情况下均能够实现稳定的跟踪,但是当目标变化剧烈时,跟踪误差较大。

而本文算法在上述几种情况下均能实现准确跟踪。

为了进一步量化分析本文算法的跟踪性能,给出了实际目标和由本文跟踪算法估计的状态参数轨迹,6 个状态参数 ($a_{1t}, a_{2t}, a_{3t}, a_{4t}, T_{xt}, T_{yt}$) 的跟踪结果如图 4 所示。从图中可以看出,本文算法的估计结果与实际状态参数轨迹基本吻合。

为了定量比较 3 种算法的性能差异,采用跟踪误差作为衡量指标。跟踪误差是指每帧图像中目标中心点的真实位置与跟踪算法输出的中心点位置之间的平均距离。跟踪过程中的误差可按下式计算:

$$err(f) = \sqrt{(x - x_T)^2 + (y - y_T)^2},$$

其中: $err(f)$ 代表第 f 帧的误差, x 、 y 为跟踪算法输出的中心点位置, x_T 、 y_T 为中心点的真实位置。

图 5 为 3 种跟踪算法的误差曲线。其中横坐标代表帧号,纵坐标代表跟踪误差,单位为像素。从图中可以看出,本文算法与 IVT 算法远远好于文献[10]的方法。根据算法原理可知,算法计算量与粒子数目成正比,而由于本文采用自适应的粒子数目,故其计算量远远小于 IVT 算法。

图 6 为 3 种算法的粒子数目曲线。其中横坐标代表帧号,纵坐标代表该帧使用的粒子数目。由于粒子滤波算法的计算量与粒子数目成正比,减少粒子数目即减少了算法的计算量。从图中可以看出,本文算法采用的平均粒子数目明显少于另外两种算法。

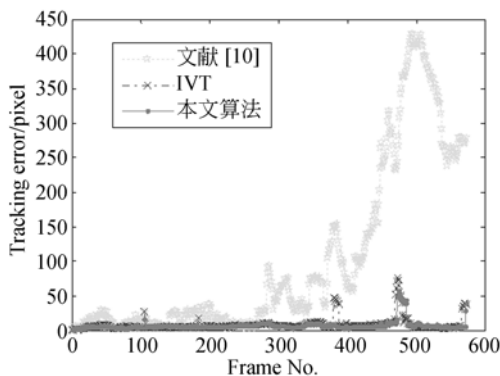


图 5 跟踪误差对比

Fig. 5 Comparison of tracking errors

为进一步验证本文算法的鲁棒性,将实验一用 3 种跟踪算法重复跟踪 50 次,得到表 1 的统计

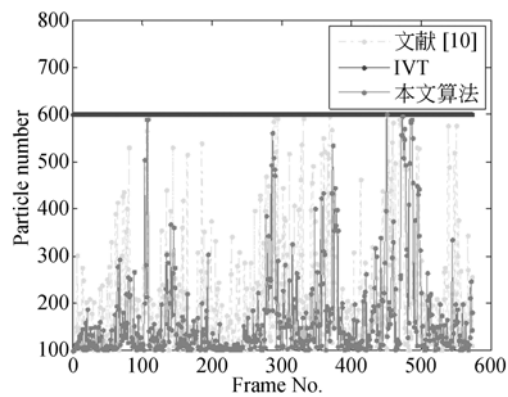


图 6 粒子数比较

Fig. 6 Comparison of particle numbers

结果。表中的跟踪速度为平均每秒实现的跟踪帧数;成功率为跟踪误差小于 15 pixel 的帧数与跟踪视频的总帧数(573 帧)在 50 次实验中的平均比值;跟踪误差为成功跟踪时中心点与实际中心点的偏差。此实验进一步证明,本文算法可以稳定准确地跟踪目标,并兼顾实时性。

表 1 3 种算法跟踪性能对比

Tab. 1 Comparison of tracking performance of three tracking algorithms

算法	跟踪速度/ (frame · s ⁻¹)	平均 粒子数	跟踪误差/ pixel	成功率/ %
文献[10]	32.05	208	11.76	42
IVT	19.12	600	8.26	95
本文算法	29.07	170	5.85	98

4 结 论

本文根据视频监控系统稳定可靠、精度高、计算量小、抗干扰能力强等要求,提出了一种新的基于概率模型的目标跟踪框架。首先,利用人眼细胞的分布特点,采用基于人眼分布结构建立了目标表观模型的描述方法,提高了跟踪系统抵抗局部遮挡的能力;其次,建立了基于自适应目标运动的动态系统模型,提高了算法对快速机动目标的跟踪鲁棒性;最后,采用实时更新的系统观测模型,使得当目标遇到遮挡、光照变化、剧烈变形等情况,可以有效避免跟踪漂移现象的发生。实验结果证明:本文算法的正确跟踪率可达 98%;平均跟踪误差小于 6 pixel。基本满足了视频监控的要求。

参考文献:

- [1] HARTLEY R, ZISSERMAN A. *Multiple View Geometry in Computer Vision* [M]. Cambridge: Cambridge University Press, 2003, 1:150.
- [2] GORDON N. J., SALMOND D. J., SMITH A. F. M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation [C]. *IEEE Proceedings on Radar and Signal Processing*, 1993, 140(2): 107-113.
- [3] ARULAMPALAM M S, MASKELL S, GORDON N, *et al.*. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking [J]. *IEEE Transactions on Signal Processing*, 2002, 50(2): 174-188.
- [4] JIA X, LU H, YANG M. Visual tracking via adaptive structural local sparse appearance model [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012: 1822-1829.
- [5] 赵鹏,沈庭芝,单宝堂. 一种基于粒子滤波的无人机电视导引系统目标跟踪算法 [J]. *光学精密工程*, 2008, 16(1):134-140.
- ZHAO P, SHEN T ZH, SHAN B T. An object tracking algorithm for TV guiding system of UAV based on particle filter [J]. *Opt. Precision Eng.*, 2008, 16(1):134-140. (in Chinese)
- [6] 陈东成,朱明,高文,等. 在线加权多示例学习实时目标跟踪 [J]. *光学精密工程*, 2014, 22(6): 1661-1667.
- CHEN D CH, ZHU M, GAO W, *et al.*. Real-time object tracking via online weighted multiple instance learning [J]. *Opt. Precision Eng.*, 2014, 22(6): 1661-1667. (in Chinese)
- [7] 杜超,刘伟宁,刘恋. 一种基于卡尔曼滤波及粒子滤波的目标跟踪算法 [J]. *液晶与显示*, 2011, 26(3): 384-389.
- DU C, LIU W N, LIU L. Target tracking algorithm based on Kalman filter and particle filter [J]. *Chinese Journal of Liquid Crystals and Displays*, 2011, 26(3): 384-389. (in Chinese)
- [8] 宋策,张葆,尹传历. 适于机载环境对地目标跟踪的粒子滤波设计 [J]. *光学精密工程*, 2014, 22(4): 1037-1047.
- SONG C, ZHANG B, YIN CH L. Particle filter design for tracking ground targets in airborne environment [J]. *Opt. Precision Eng.*, 2014, 22(4): 1037-1047. (in Chinese)
- [9] ISARD M, BLAKE A. Condensation-Conditional density propagation for visual tracking [J]. *Int. J. Comput. Vis.*, 1998, 29(1): 5-28.
- [10] 王书朋,姬红兵. 用于目标跟踪的自适应粒子滤波算法 [J]. *系统仿真学报*, 2010, 22(3): 630-633.
- WANG SH P, JI H B. Adaptive particle filtering for efficient object tracking [J]. *Journal of System Simulation*, 2010, 22(3): 630-633. (in Chinese)
- [11] ROSS D, LIM J, LIN R, *et al.*. Incremental learning for robust visual tracking [J]. *Internat. J. Comput. Vision*, 2008, 77(3): 125-141.

作者简介:



李静宇(1986—),女,黑龙江齐齐哈尔人,博士,助理研究员,2008年于南开大学获得学士学位,2013年于中国科学院长春光学精密机械与物理研究所获得博士学位,主要从事数字图像处理及机器视觉等方面的研究。E-mail: 18686653384@163.com