

基于增量子空间自适应决策的目标跟踪

仝小敏¹ 张艳宁¹ 杨涛¹

摘要 基于增量子空间的目标跟踪算法多数不加选择地将检测到的目标作为模板训练的样本,并以固定频率更新模板,这种无反馈闭环机制使得算法在目标外观模型发生变化、光照变化等复杂条件下难以鲁棒跟踪目标,一旦跟踪失败很难从错误中恢复.为此,我们提出一种反馈闭环跟踪算法,在增量子空间粒子滤波跟踪框架下,引入跟踪状态判决作为后续模板更新依据.通过判决反馈信息选择合适的样本适时更新模板,有效克服目标外观模型的变化,持续跟踪目标.实验结果表明,由于引入跟踪状态判决,在目标外观变化、光照变化等情况下,本算法能够以与环境相适应的频率及时更新模板,提高跟踪精度,实验结果验证了本文算法的鲁棒性和有效性.

关键词 自适应更新,跟踪状态判决,子空间增量学习,目标跟踪

DOI 10.3724/SP.J.1004.2011.01483

Robust Object Tracking Based on Adaptive and Incremental Subspace Learning

TONG Xiao-Min¹ ZHANG Yan-Ning¹ YANG Tao¹

Abstract The traditional target tracking algorithm usually trains the template with detected samples and updates the template at a fixed frequency. This close-loop mechanism lacks feedback and often makes it impossible to track targets robustly when target appearance or illumination changes. Besides, it can not recover from tracking failure easily. Therefore, we propose a feedback-loop tracking framework by bringing in the tracking state judgement. In this framework, the tracking state judgement works as the basis of the following template updating. According to the tracking state judgement, we can choose suitable samples to update the template at appropriate time so as to track targets continuously. Experimental results show that our method can get the current template immediately and correctly due to the tracking state judgement and decision mechanism. We can update the template at an adaptive frequency and meanwhile track targets correctly even in the case of target appearance or illumination changing.

Key words Adaptive updating, tracking state judgement, subspace incremental learning, object tracking

目标跟踪技术作为目标识别、运动捕获、视频监控、人机交互、视频会议等高层视频语意理解的基础,一直是计算机视觉领域的研究热点^[1].经典的算法有 Meanshift、卡尔曼滤波、粒子滤波等^[2-5],按照提取特征尺度不同可分为基于特征点的跟踪、基于轮廓的跟踪、基于区域特征的跟踪.多数方法中特征提取基于图像色彩、纹理、灰度等外观信息,一旦目标外观发生变化,容易跟踪失败^[6-15].近年来,子空间特征因其良好的尺度不变性、旋转不变性、平移不变性等优良特性,成为目标跟踪中特征提取的主流方法^[16-24].Ross 等^[16]将增量学习方法引

入,提出了基于增量子空间的粒子滤波跟踪算法,有效更新跟踪模板,一定程度上解决了目标外观、光照等缓慢变化情况下的跟踪问题.Yang 等^[25]利用二维子空间特征来保持模板结构特征,提高跟踪精度的同时降低计算复杂度,文献 [26-27] 也提出利用二阶张量保持图像数据结构信息的完整性.尽管上述方法在外观、环境条件缓慢变化情况下能够正确跟踪目标,但是由于跟踪算法本身是基于实际目标符合假设模型的基本假设,一旦实际目标发生突变,不符合假设模型,容易跟踪失败.究其本质,主要由跟踪算法缺乏反馈机制造成,如图 1(a) 所示,该类算法不加选择地把跟踪结果作为模板更新的正样本,按照固定的频率进行模板更新,在目标外观和环境条件缓慢变化等情况下,可正确跟踪目标.由于环境条件和目标外观变化的复杂性和不规律性,固定的模板更新频率显然不能满足鲁棒跟踪需求,更新频率过高会影响算法效率;反之,更新频率过低则不能适应环境变化.一旦目标或者环境发生突变,也即给定的模板更新频率不再适于目标变化的实际频率,跟踪就有可能发生错误.由于缺乏跟踪状态判定与反馈,将很难从错误中恢复.

一个鲁棒的跟踪算法不仅能在实际情况与假定

收稿日期 2011-01-17 录用日期 2011-07-07
Manuscript received January 17, 2011; accepted July 7, 2011
国家自然科学基金 (60903126, 60872145), 中国博士后特别基金 (201003685), 中国博士后基金 (20090451397), 西北工业大学基础研究基金 (JC201120) 资助
Supported by National Natural Science Foundation of China (60903126, 60872145), China Postdoctoral Special Science Foundation (201003685), China Postdoctoral Science Foundation (20090451397), and the Northwestern Polytechnical University foundation for Fundamental Research (JC201120)
1. 西北工业大学计算机学院陕西省语音与图像信息处理重点实验室 西安 710129
1. Shaanxi Key Laboratory of Speech and Image Information Processing, School of Computer Science, Northwestern Polytechnical University, Xi'an 710129

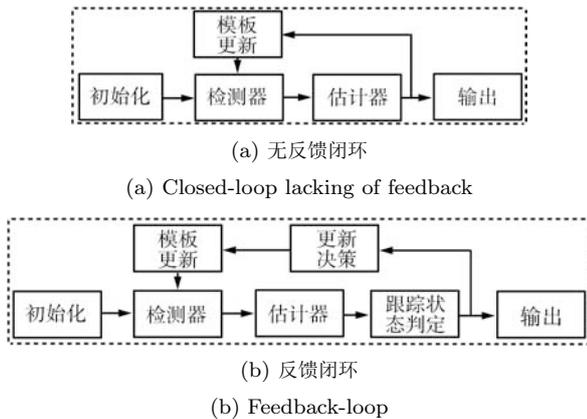


图 1 无反馈闭环和反馈闭环跟踪算法比较

Fig. 1 Comparison of closed-loop tracking algorithm without feedback and feedback-loop tracking algorithm

模型之间一致时达到最优跟踪性能, 发生较小偏差时性能降低的范围不大, 甚至在实际情况与假定模型之间发生较大偏差时, 跟踪性能仍能满足系统指标, 且不出现发散现象, 算法仍然有效. 这就要求算法设计必须能判断实际情况与假定模型之间的符合程度, 并以此为依据构成闭环通路 (如图 1 (b)), 及时反馈修正模型, 从而达到鲁棒跟踪. 作为高层反馈信息, 目标跟踪状态判定应该能够有效指导底层信息处理, 这也符合生物视觉信息处理的逻辑.

基于上述分析, 本文选取增量子空间特征作为模板特征^[16], 在粒子滤波跟踪框架下引入跟踪状态判决与决策机制, 提出了一种基于增量子空间自适应决策的目标跟踪算法, 一方面利用子空间增量学习获得正确、完整、本质的模板描述, 克服目标旋转、缩放、平移等影响; 另一方面引入跟踪状态判定及反馈更新机制, 根据跟踪状态选择合适的样本适时更新模板. 本方法采用粒子滤波进行预测和估计, 提取子空间特征作为目标模板, 无需离线训练, 在跟踪过程中通过自适应决策的增量学习更新模板, 更新频率随目标外观和环境的变化而动态改变, 在目标外观和环境变化等复杂条件下均可鲁棒跟踪目标.

1 自适应决策更新

自动跟踪状态判决和自适应决策更新是鲁棒跟踪算法的必要环节. 在基于子空间特征的粒子滤波跟踪框架下, 本文重点研究如何利用当前跟踪结果的各种冗余信息进行跟踪状态判决和子空间特征自适应更新决策.

1.1 跟踪状态判定

为鲁棒跟踪目标, 需判定是否跟踪正确目标及跟踪失败的原因, 以便从错误中恢复. 这里将目标跟踪状态分为稳态跟踪、突变态跟踪和渐变态跟踪三

种. 与这三个状态判定密切相关的度量包括目标模板相似性度量模型、目标运动模型、重构误差模型、粒子群分布模型等. 相似性度量用来度量目标与给定模板之间的相似性, 其相似程度可作为目标跟踪状态判定的依据之一; 目标运动模型是根据目标运动速度建立的模型, 假定目标运动速度具有连续性, 运动突变的点可作为跟踪序列中的野点, 从而为失跟判定提供依据; 重构误差模型是为目标重构误差建立的模型, 假定目标与模板相似时重构误差最小, 时间序列上重构误差的分布也呈现一定的规律性, 当不满足这种规律性时, 可能发生了失跟; 粒子群分布在正确跟踪时呈现一定的规律性, 一旦跟踪失败粒子分布就会杂乱无章, 然而粒子群表现出来的不规律性较之于跟踪失败存在一定的滞后性, 可作为失跟判定的辅助依据.

1) 相似性度量

用以度量目标与给定模板之间的相似性, 并根据相似程度计算目标跟踪状态概率分布. 历史样本训练的模板是目标在时间轴上的特征集合, 可以很好地描述目标外观缓慢变化、环境条件缓慢变化等稳态情况下目标的特征, 目标与其越相似, 说明稳态跟踪的可能性也越大; 上一帧目标训练的瞬时模板反映的是目标瞬时特征, 目标与其差异越大, 则发生突变 (包括环境突变或者遮挡等引起的目标外观突变) 的可能性也越大. 度量目标与这两者之间的相似性, 可为目标跟踪状态提供初步判据.

假设目标与历史模板的相似度为 p_1 , 与瞬时模板的相似度为 p_2 , 我们采用如下三个公式计算稳态概率 p_s 、渐变态概率 p_w 和突变态概率 p_c .

$$p_s = p_1 \times p_2 \quad (1)$$

$$p_w = (1 - p_1) \times p_2 \quad (2)$$

$$p_c = (1 - p_1) \times (1 - p_2) \quad (3)$$

目标与历史模板和瞬时模板相似度都很高, 说明跟踪状态为稳态的概率较大, 如式 (1) 所示; 当目标与历史模板相似度低, 而与瞬时模板相似度很高时, 如式 (2), 说明历史模板已经不能够很好地描述模板特征, 主要由目标在运动过程中外观的缓慢变化造成, 此时往往需更新模板; 当目标与瞬时模板和历史模板的相似度都很低时, 如式 (3), 说明目标或者环境条件很可能发生了突变.

简单选取有效历史训练数据的均值 $Mean$ 作为历史模板, 瞬时模板选取上一帧目标 $Object_{old}$, 假定目标灰度值分布服从正态分布, 那么目标 $Object_{new}$ 与历史模板相似度 p_1 、与瞬时模板相似度 p_2 计算如式 (4) 和式 (5), 其中 σ_1 和 σ_2 分别表示目标历史模板、瞬时模板高斯分布的标准差, 在实际算法实现中

可取相同的值, σ_1 和 σ_2 取值过大会削弱目标外观较大变化情况下的状态判决性能; 反之取值过小则会使得目标外观发生缓慢变化时跟踪性能下降, 需要根据具体的数据选择合适的取值.

$$p_1 = e^{-\frac{\|Object_{new} - Mean\|_2^2}{2\sigma_1^2}} \quad (4)$$

$$p_2 = e^{-\frac{\|Object_{new} - Object_{old}\|_2^2}{2\sigma_2^2}} \quad (5)$$

通过计算并比较稳态、突变态和渐变态概率, 可初步判定跟踪器处于哪个状态的概率最大.

2) 目标运动模型

假定相邻两帧间目标的运动具有连续性, 运动发生突变是小概率事件. 在上述相似性度量的基础上, 进行运动分析, 可进一步判定目标跟踪正确与否. 若初步判定为跟踪目标处于突变态, 那么可依据跟踪结果中运动状态是否发生突变来判定是由于跟踪失败引起的突变, 还是由于环境条件及目标本身外观变化引起的突变; 若初步判定为稳态, 那么可依据运动突变与否判定是否正确跟踪; 若初步判定为渐变态, 且运动未发生突变, 则认为正确跟踪, 否则为跟踪失败.

设第 t 帧和 $t-1$ 帧的目标在图像中的位置分别为 (x_t, y_t) 和 (x_{t-1}, y_{t-1}) , 那么目标在 t 时刻的运动速度为 $\mathbf{v}_t = (x_t - x_{t-1}, y_t - y_{t-1})$, 速度大小 $|\mathbf{v}_t| = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}$, 方向为 $\theta_t = \arctan \frac{y_t - y_{t-1}}{x_t - x_{t-1}}$. 由运动连续性假设可知, t 时刻目标运动速度近似服从以 \mathbf{v}_{t-1} 为中心的正态分布, 即 $\mathbf{v}_t \sim N(\mathbf{v}_{t-1}, \sigma_{t-1}^2)$, 按照经验值取 $\sigma_{t-1}^v = \text{mean}_{t_0=1}^{t-1} |\mathbf{v}_{t_0} - \mathbf{v}_{t_0-1}|$, $\sigma_{t-1}^\theta = \text{mean}_{t_0=1}^{t-1} |\theta_{t_0} - \theta_{t_0-1}|$. 若 \mathbf{v}_t 落在 \mathbf{v}_{t-1} 的 $3\sigma_{t-1}$ 置信区间内, 则认为运动连续, 否则认为发生了运动突变.

3) 重构误差模型

目标与训练的模板之间存在着差异, 随着时间的变化, 这个差异呈现出规律性, 即在模板建立之初, 训练模板所使用的样本和目标在时间上间隔较小, 真实目标与模板最为相近, 重构误差最小; 随着时间的推移, 目标与模板间的差异有明显增大趋势, 如图 2 所示, 重构误差在 10~40 帧之间呈上升趋势. 若跟踪过程中一旦发生错误, 这个规律性就会被破坏, 如图 2 中 200 帧以后, 跟踪目标丢失以后, 重构误差没有增大趋势. 误差模型构建如式 (6), 可为跟踪错误与否提供进一步的判据.

$$ErrorModel_t = \begin{cases} 1, & Error_t > Error_{t-1} \\ 0, & \text{其他} \end{cases} \quad (6)$$

然而, 误差降低并不一定由跟踪错误造成, 例如跟踪过程中目标外观变化不大时, 跟踪误差可能会上下波动, 而不一定上升. 因此, 需要在初步判

定跟踪状态以及运动分析的基础上利用误差模型 $ErrorModel_t$ 确定跟踪状态.

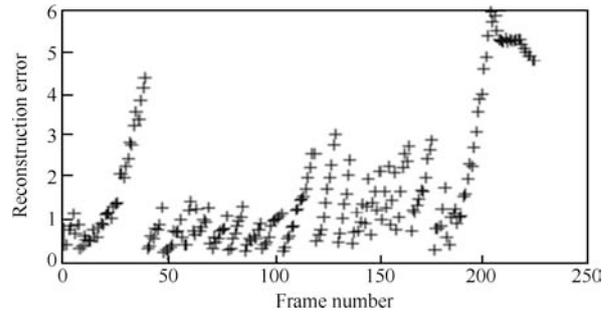


图 2 重构误差随时间的变化

Fig. 2 Reconstruction errors at different time

4) 粒子群分布

粒子滤波跟踪框架中, 利用一组带权重的粒子来模拟目标后验分布, 当模板正确且上一帧跟踪正确时, 权重会集中于少数粒子; 否则, 粒子分布会出现发散, 权重分布比较均匀. 利用粒子群分布的这一特性, 可以判定模板是否需要更新或者上一帧是否跟踪错误. 首先将 N 个粒子权重按照从大到小排序, 得到 N 个有序粒子权重 $w_t^{(i)}$, $i = 1, 2, \dots, N$, 计算粒子分布函数 $F_t(n)$, $n = 1, 2, \dots, N$, 如式 (7), $F_t(n)$ 表示前 n 个最大粒子权重的和, $F_t(N) = 1$, 正确跟踪与错误跟踪情况下粒子群分布比较如图 3 所示. 可见, 粒子群分布集中程度能一定程度上反映跟踪状态.

$$F_t(n) = \sum_{i=1}^n w_t^{(i)}, \quad n = 1, \dots, N \quad (7)$$

由上述分析可知, 相似性度量、目标运动模型、重构误差模型、粒子群分布可作为跟踪状态判定的有效依据, 通过原因分析和状态判定可为反馈决策提供依据. 跟踪状态判定具体方法如表 1 所示.

1.2 更新策略选择

根据跟踪状态判定结果及原因分析, 可以作出相应的更新策略选择. 若判定为正确跟踪, 多数情况下目标与模板匹配度较高, 无需更新模板, 只将跟踪结果存入模板训练集. 而对于判定为跟踪错误, 由于造成错跟的原因不同, 应采取不同的更新策略.

稳态下, 若判定为正确跟踪, 说明当前目标与模板匹配度较高, 无需进行更新; 若判定为错误跟踪, 其原因可能是场景中出现了相似目标, 这时, 无需做特殊处理, 只需将跟踪结果对应的粒子抛弃, 从其余粒子中产生跟踪结果, 重新进行跟踪状态判定.

渐变态下, 若判定为正确跟踪, 表明目标与历史模板的相似度虽然低于其与瞬时模板的相似度, 但是

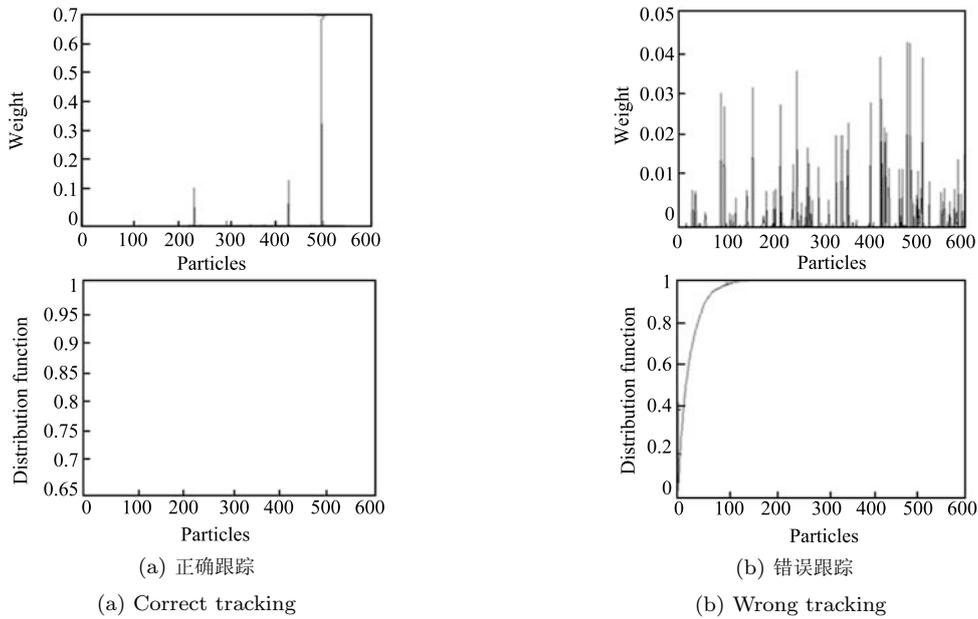


图 3 粒子群分布

Fig. 3 Particle distribution

表 1 跟踪状态判定

Table 1 Tracking state judgement

相似性度量	运动是否突变	误差上升还是下降	粒子权重分布是否集中	跟踪状态判定	原因分析
稳态	否	—	—	正确跟踪	—
稳态	是	上升	—	正确跟踪	—
稳态	是	下降	—	错误跟踪	相似目标出现
渐变态	是	下降	—	错误跟踪	需要模板更新
渐变态	是	上升	是	正确跟踪	—
渐变态	是	上升	否	错误跟踪	上一帧跟错
渐变态	否	上升	—	正确跟踪	—
渐变态	否	下降	是	正确跟踪	—
渐变态	否	下降	否	错误跟踪	上一帧跟错
突变态	是	—	—	错误跟踪	环境条件突变
突变态	否	上升	—	正确跟踪	—
突变态	否	下降	—	错误跟踪	相似目标出现

历史模板还足以描述当前目标的特征, 因此暂不需要更新, 而是将跟踪结果存入模板训练集; 若判定为上一帧跟踪发生错误而引起跟踪错误, 需将两帧前的跟踪结果作为当前粒子的初始位置, 重新进行跟踪; 若判定为由于模板需要更新造成的跟踪错误, 则采用增量空间学习方法对有效模板训练集进行训练, 获得新的模板。

突变态下, 若判定为正确跟踪, 说明目标外观发

生了较大变化, 此时需利用当前跟踪结果对模板进行更新, 并重新开始跟踪; 若判定为由相似目标出现引起的错误跟踪, 则需将跟踪结果对应的粒子抛弃, 从其余粒子中产生跟踪结果, 并进行跟踪状态判定; 若判定为由环境条件突变引起的错误跟踪, 则将上一帧的跟踪粒子作为本帧估计输出, 并取出相应的图像作为跟踪结果, 重新初始化粒子, 开始新的跟踪。

图 4 中, 将跟踪器估计结果输入判决器, 可输出相应的判决结果. 自适应决策器根据判决结果及原因分析作出相应的决策. 由图 4 可知, 判决结果的 8 种情况中, 只有当渐变态下由于模板不符合当前目标而造成错跟时需进行模板更新; 而当突变态正确跟踪时, 模型发生了切换, 需利用突变后的目标作为训练样本重新获取模板. 当相似目标出现在场景中引起错跟时, 需将该目标排除并重新估计结果, 输入决策器, 这是一个反馈迭代的过程, 有利于正确跟踪并且不增加额外的计算量. 渐变态上一帧目标跟踪错误时, 需结合该假设先验重新对本帧图像进行目标跟踪, 这也是一个反馈修正的过程, 然而这种反馈修正需要以时间消耗为代价, 在实际应用中, 可选择丢弃当前帧, 从下一帧开始跟踪.

2 基于增量子空间自适应决策的目标跟踪

基于上述跟踪状态判定和自适应更新决策, 我们提出了基于增量子空间自适应决策的粒子滤波目标跟踪算法, 不仅能够及时更新模板, 更能作出错误判定和处理, 使得算法具有一定的自恢复能力. 在 Bayes 跟踪框架下, 提取样本子空间特征作为跟踪模板, 利用粒子滤波进行预测和估计, 得到估计结果

后将其输入跟踪状态判决器进行状态判决并作出相应的更新决策, 或者通过增量学习的方法更新模板, 或者重新训练模板, 或者重新初始化粒子开始新的跟踪, 算法流程如图 5, 详述如下:

1) 参数初始化

首先进行模板初始化. 子空间特征即是通过 PCA 从训练样本集获得跟踪模板, 主要包含样本均值 $Mean$ 、特征向量集 $Basis$ 、特征值向量 Σ 、有效样本数量 $NumSample$ 四个参数. 初始确定的目标图像作为跟踪模板参数 $Mean$ 的初始值, $Basis$ 、 Σ 初始化为空, $NumSample$ 初始化为 0, 历史模板和瞬时模板初始化为 $Mean$.

其次对粒子初始化. 每个粒子代表一个可能的目标窗口, 采用仿射参数模型来表示 (如式 (8)), 保证参数之间的独立性, 从而确保粒子漂移的随机性.

$$AffineParam = \begin{bmatrix} xCenter & \lambda_1 & \frac{\lambda_1}{\lambda_2} \\ yCenter & \theta & \phi \end{bmatrix} \quad (8)$$

其中, $(xCenter, yCenter)$ 表示粒子中心坐标, λ_1, λ_2 分别是横轴和纵轴方向的尺度因子, θ, ϕ 分别是旋转角和剪切角. 由初始目标区域的四个顶点获得仿射参数模型, 并作为所有粒子的初始参数.

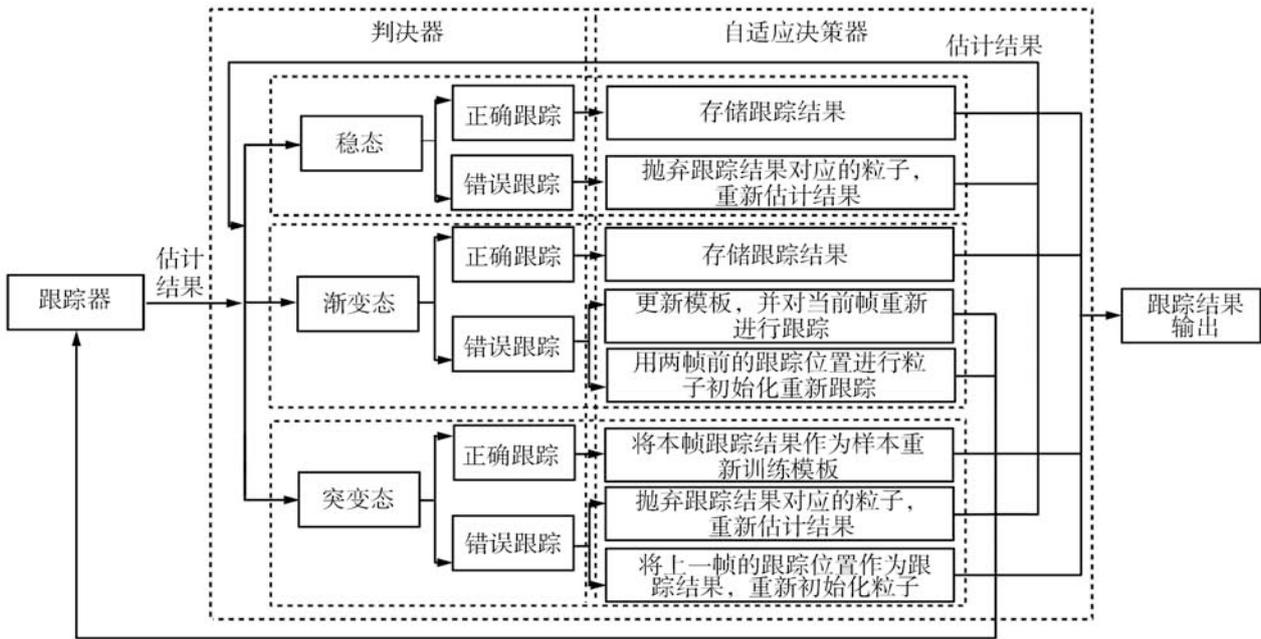


图 4 跟踪状态判定及决策图

Fig. 4 Tracking state judgement and updating decision

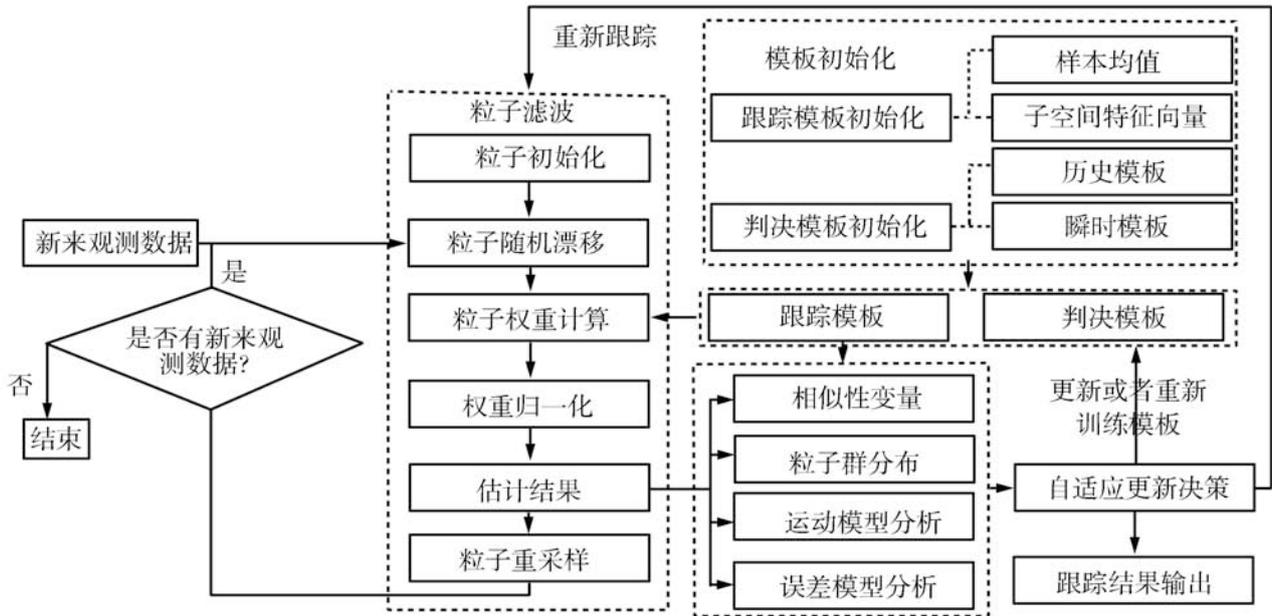


图 5 基于增量子空间自适应决策的跟踪流程图

Fig. 5 Incremental learning and self adaptive decision based tracking algorithm

2) 粒子滤波

在 Bayes 跟踪框架下, 粒子滤波是利用一组带有权值的粒子 $\{Particle_i, w_i | i = 1, \dots, M\}$ 来模拟跟踪状态分布函数, 用 $t-1$ 时刻粒子重要性采样结果来模拟 t 时刻先验分布, 通过计算条件概率估计后验分布, 选择后验分布期望最优跟踪状态. 由于重采样后所有粒子的概率相等, 后验概率分布与条件概率分布相一致.

利用子空间特征计算重构误差, 重构误差越小, 说明与模板一致性越高, 是跟踪目标的概率越大, 否则与模板差异越大, 是跟踪目标的可能性越小, 因此重构误差能够反映目标概率分布. 重构系数 $Coef$ 、重构误差 $reconError$ 计算如下:

$$Coef = \mathbf{Basis}^T \times (A - Mean) \quad (9)$$

$$reconError = A - Mean - \mathbf{Basis} \times Coef \quad (10)$$

其中, \mathbf{Basis} 为特征向量, \mathbf{Basis}^T 为特征向量的转置, A 为新样本, $Mean$ 为历史样本均值. 根据重构误差大小可以为粒子赋予相应的权重, 重构误差越大, 权重越小; 反之, 权重越大. 权重分布反映了后验概率分布, 设 w_i^t 表示 t 时刻第 i 个粒子的权重, h_i^t 为 t 时刻第 i 个粒子对应的状态仿射参数, $p(h_i^t | h_i^{t-1})$ 服从高斯分布, 并且 h_i^t 由 h_i^{t-1} 按高斯概率随机漂移得到, 如式 (11), ψ 表示高斯分布产生的随机数, 那么有 $p(h_i^t | w_i^t, h_i^{t-1}) \propto w_i^t$, 权重按照式 (12) 计算, 其中 σ_w 为重构误差高斯分布的标准差, 由于一般情况下粒子群中总存在与模板相似度很高

的粒子使得重构误差很小, 取值过大使得粒子权重之间的差异性减小, 不利于持续跟踪, 因此重构误差高斯分布的标准差通常很小.

通过对所有粒子计算权重并归一化可得后验概率分布, 为了保证跟踪的准确性, 选取后验概率极大值点作为跟踪估计结果.

3) 跟踪状态判决与自适应更新决策

在 t_1 时刻前不进行判决, 默认为正确跟踪并存储跟踪结果. $t = t_1$ 时进行第一次模板训练, 之后按照第 1.1 节和第 1.2 节提出的跟踪状态判决和自适应决策方法进行自适应模板更新并输出跟踪结果.

$$h_i^t = h_i^{t-1} + \psi \quad (11)$$

$$w_i^t = \exp\left(\frac{-\|reconError_i^t\|_2^2}{2\sigma_w^2}\right), i = 1, \dots, M \quad (12)$$

模板更新过程是利用有效样本计算新的子空间特征, 主要包括去均值、计算协方差阵、奇异值分解计算特征值和特征向量三个关键步骤. 然而跟踪过程中随着有效样本数目的增加, 计算子空间特征的耗时和存储消耗也会越来越大, 这使得算法无法实时跟踪目标. 这里采用 SKL (Sequential Karhunen-Loeve transform) 方法^[16, 28] 对模板进行在线增量学习, 无需存储所有样本, 仅利用已有样本的模板训练结果和新来样本计算新模板, 大大降低了计算复杂度. 假设当前模板子空间特征的特征值为 Σ , 特征向量为 \mathbf{Basis} , 新样本为 B , 采用 SKL 方法增量学

习新的子空间特征的步骤如下:

步骤 1. 对 $B - \mathbf{Basis Basis}^T B$ 进行 QR 分解, 求得 $\tilde{B} = orth(B - \mathbf{Basis Basis}^T B)$, 并代入 $R = \begin{bmatrix} \Sigma & \mathbf{Basis}^T B \\ 0 & \tilde{B}^T B \end{bmatrix}$, 计算 R ;

步骤 2. 计算 R 的奇异值分解, 求得对应的特征值 $\tilde{\Sigma}$ 和特征向量 $\tilde{\mathbf{Basis}}$;

步骤 3. 计算新的特征向量 $\mathbf{Basis}' = [\mathbf{Basis} \tilde{B}] \tilde{\mathbf{Basis}}$ 和特征值 $\Sigma' = \tilde{\Sigma}$.

步骤 4. 重复步骤 2 和步骤 3, 直到视频序列结束.

从图 5 可以看出, 本文所提出的算法是在子空间增量学习基础上, 进行了自适应更新决策, 使得增量学习更加自适应于环境、目标外观等不规则变化, 相比于固定频率的模板更新, 自适应更新机制更加符合实际环境条件变化规律, 能够提高跟踪鲁棒性.

3 实验结果与分析

以本文算法为核心, 用 VC++ 开发了一个实时目标跟踪系统. 实验中硬件系统采用 Intel Core2 CPU 2.66 GHz 处理器, 内存 2 G. 为测试和改进算法, 采用美国 DARPA 和文献 [16] 发布的数据库作为测试数据库. 其中 DAPAR VIVID 数据是一段长度为 1821 帧的无人机航拍视频序列, 这段数据包含有目标角度、旋转、尺度、光照等缓慢变化与多次突变, 能够很好地验证本文算法有效性. 文献 [16] 发布的跟踪测试数据库包含有人脸、车辆、玩具等多种类型的目标, 并存在光照、姿态等引起的目标外观变化等复杂情况, 可作为验证数据库, 对本文算法在跟踪精度和效率方面的提高进行验证. 由于文献 [16] 所提出的 IVT 算法能够在线更新目标模板, 一定程度上能较好地解决目标外观、光照变化等复杂情况下的目标跟踪问题, 而本文算法在跟踪中引入跟踪状态判决和自适应决策, 实验验证主要与其进行对比, 可以验证本文所提出判决反馈机制的有效性. 实验中, 系统对 320 像素 \times 240 像素的图像序列跟踪速度平均可达为 10 帧/秒.

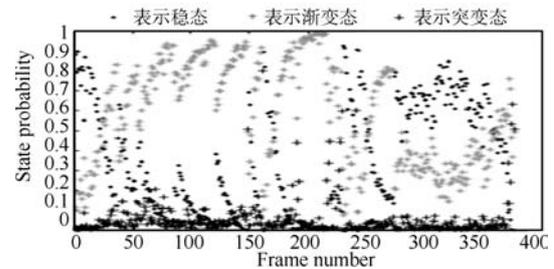
3.1 本文算法性能分析

本文算法主要通过相似性度量、运动模型分析、误差模型分析、粒子分布这四个模型引入跟踪状态判定和原因分析, 并由此做出自适应的更新决策. 为验证跟踪状态判决和自适应决策机制的有效性, 下面将从相似性度量、运动模型分析、误差模型分析、粒子分布四个方面进行实验.

1) 相似性度量

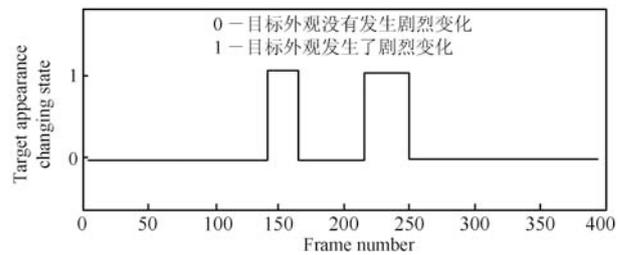
如图 6 (a) 所示为跟踪过程中利用相似性度量进行的跟踪状态预判定, 横轴代表帧数, 纵轴代表不

同状态的概率, “.” 表示稳态、“*” 表示渐变态、“+” 表示突变态. 判定结果表明, 在 150 帧、162 帧、230 帧、242 帧跟踪状态发生了突变. 图 6 (b) 所示为跟踪状态真值, 1 表示目标外观或者环境发生突变的时段, 0 表示未发生突变的时段. 对比图 6 (a) 和图 6 (b) 可知, 跟踪状态预判定结果与真值相符. 从图 7 (a) 为 150 帧前后图像可看出目标灰度相比于模板确实发生了较为剧烈的变化, 图 7 (b)~7 (d) 分别为 162, 231, 242 帧前后目标图像, 目标分别在灰度、姿态等外观上发生了突变, 从而验证了相似性度量模型的有效性.



(a) 跟踪状态预判定

(a) Tracking state pre-judgement



(b) 跟踪状态真值

(b) Tracing state ground truth

图 6 跟踪状态判决

Fig. 6 Tracking state judgement by similarity evaluation

2) 运动模型分析

主要验证运动模型能否及时反映目标运动不规则, 并为跟错判定提供依据. 如图 8 所示为运动速度随时间的变化, 可以看到, 在 82 帧时运动发生了突变, 运动模型判定为突变态, 由于运动目标相似性度量判定为突变态, 跟踪状态判定为跟踪错误, 并将上一帧跟踪位置作为本帧跟踪的结果. 自适应更新决策前后的跟踪结果如图 9 所示. 可见, 自适应决策提高了跟踪算法精度, 并能够从错误中恢复.

3) 重构误差模型分析和粒子群分布

在目标外观随时间变化明显的情况下, 重构误差会随着时间呈明显上升趋势. 而实际情况中由于目标外观变化的不确定性, 重构误差往往会随机波动, 因此重构误差不能够作为跟踪错误的确定性判据. 如图 10 所示, 重构误差虽然下降 (图 10 (a)), 但

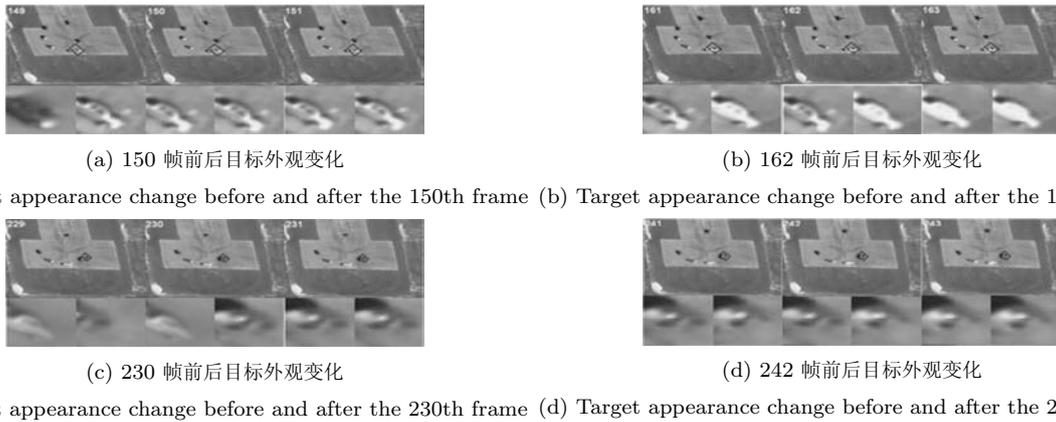


图 7 目标突变前后

Fig.7 Objects before and after mutation

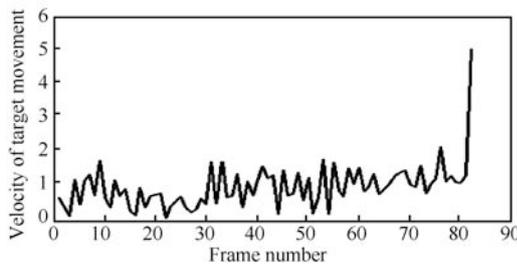


图 8 运动速度随时间的变化
Fig.8 Speed vs. time

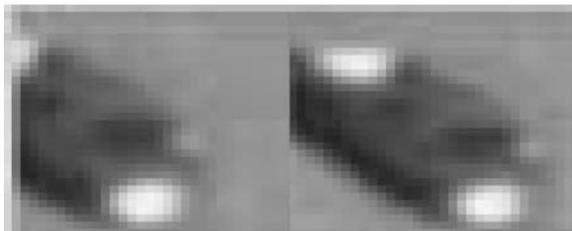


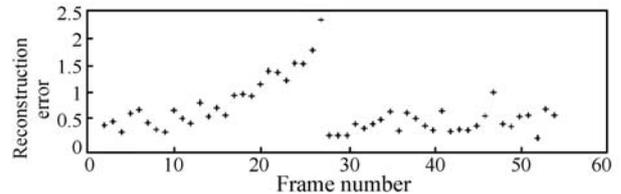
图 9 自适应决策前后跟踪结果比较
Fig.9 Tracking results before and after adaptive updating

由于粒子权重集中分布于少数粒子(图 10(b) 和 10(c)), 所以判定跟踪正确. 图 11 所示为由于光照变化引起目标外观发生突变时模板更新的结果, 测试数据来自文献 [16] 发布的测试数据库. 图 11(a) 所示为 149 帧前后的目标图像, 图 11(b) 为本文算法对应的跟踪结果、重构误差、重构图像和模板, 图 11(c) 为基于增量学习的鲁棒视觉跟踪 (Incremental learning for robust visual tracking, IVT) 算法^[16] 算法对应的处理结果. 可以看出, 149 帧前后目标外观 (包括亮度、姿态等) 发生突变, 164 帧时重构误差很大, 旧模板已不符合当前目标特征, 为持续跟踪目标, 需及时选用当前数据更新模板. 本文利用自适应更新决策机制作出相应判定后将前一帧数

据作为样本重新训练模板, 更新后的特征向量只有一个, 且与当前目标相近程度很高, 重构误差很小, 更新后的模板能更好的描述当前目标的特征, 从而为持续目标跟踪提供了保证. 图 11(c) 中更新后的模板与目标差异程度仍然较大, 重构误差相对明显, 不能及时反映目标外观的变化, 对持续跟踪不利.

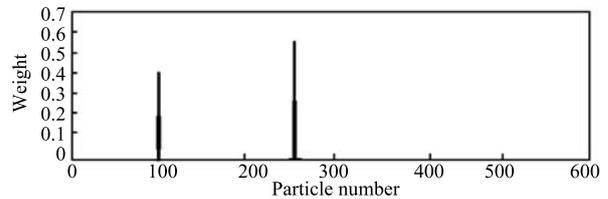
3.2 与 IVT 算法的精度和效率比较

由于引入了自适应更新决策机制, 本文所提出



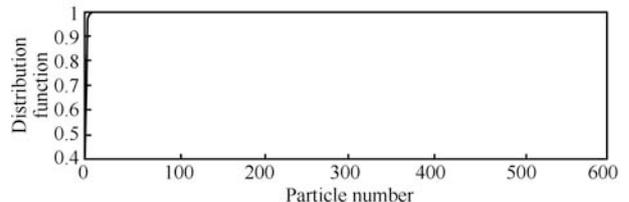
(a) 重构误差

(a) Reconstruction errors



(b) 粒子分权重

(b) Particle weights



(c) 粒子分布

(c) Particle distribution

图 10 误差模型和粒子分布

Fig.10 Error model and particle distribution

的算法更能够鲁棒跟踪目标, 并且具有一定的错误恢复能力. 图 12 所示为本文算法跟踪结果与固定 5 帧更新频率的夜晚车辆跟踪结果对比, 第 1 行为本文算法跟踪结果, 第 2 行为 IVT^[16] 算法的跟踪结果. 可以看到, 本文算法在光照等环境条件发生变化、目标姿态、外观发生剧烈变化等情况下均可持续跟踪目标, 而固定更新频率条件下目标跟踪则出现错误跟踪, 且无法自恢复. 图 13 和图 14 分别是白天车辆和人脸跟踪结果比较. 对于不同类型的目标, 在目标发生缓慢外观变化或者突变等复杂情况下, 本文算法可以对其实现持续跟踪.

表 2 所示为 DARPA 发布的 VIVID 测试数据库下本文算法与 IVT 算法的跟踪精度和更新频率比较, 其中 IVT 算法更新次数是指与本文算法跟

踪相同时间内的更新次数, 跟踪时间之比是指本文算法跟踪时长比 IVT 算法跟踪时长, 更新次数之比是指本文算法更新次数与 IVT 算法更新次数之比. 可以看到, 跟踪时间之比平均约为 2.136, 而更新次数之比平均约为 0.2, 对于短时间内存在的目标, 本文算法能够利用较低的更新频率达到同样的跟踪精度; 而对于存在时间较长的目标, 本文算法更能够鲁棒跟踪. 表 3 所示为文献 [16] 发布测试数据库上本文算法与 IVT 算法跟踪结果的比较, 大量的测试结果表明, 相比于无反馈的 IVT 算法, 本文方法的自适应更新决策机制能够减少不必要的更新, 提高算法效率, 这是由于模板更新的过程涉及奇异值分解、QR 分解等复杂的矩阵运算是本文跟踪算法耗时的主要来源. 未考虑更新决策的跟踪算法需要以某

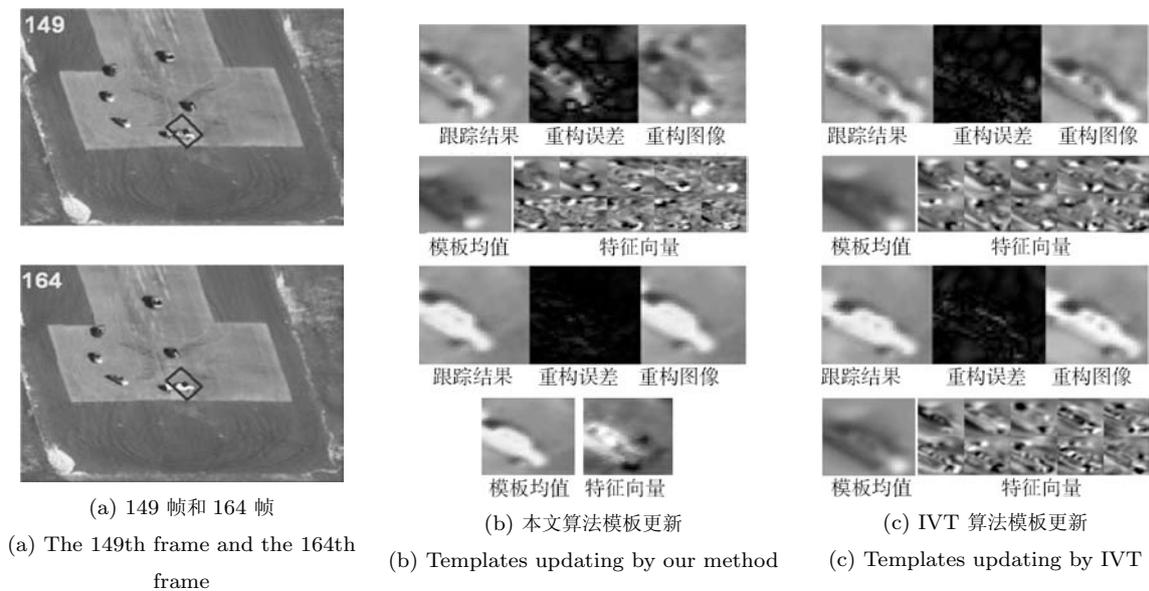


图 11 模板更新结果比较

Fig. 11 Comparison of template updating



图 12 夜晚车辆跟踪结果比较 (数据来自文献 [16] 发布的数据库. 第 1 行为本文算法结果, 第 2 行为 IVT 算法结果)

Fig. 12 Tracking results comparison at night (Data are from [16]. The top row shows the results by our method and the bottom shows the results of IVT.)

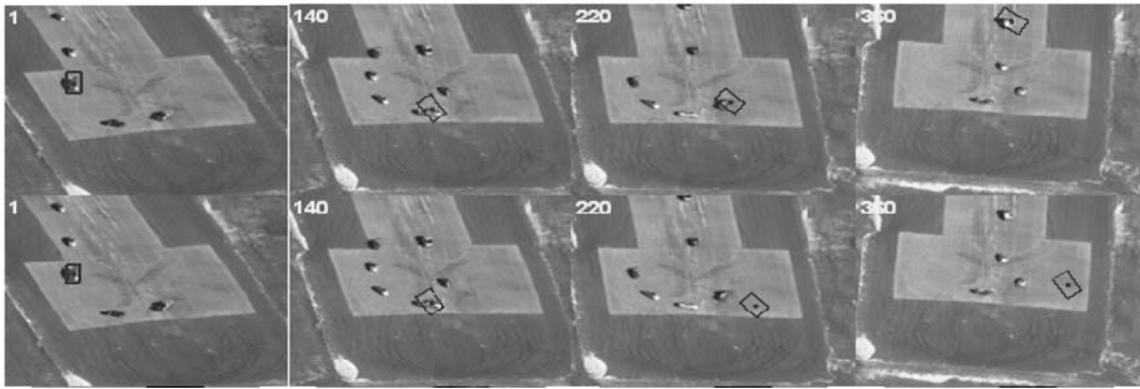


图 13 白天车辆跟踪结果比较 (数据来自 DARPA 发布的 VIVID 数据库. 第 1 行为本文算法结果, 第 2 行为 IVT 算法结果)
 Fig.13 Tracking results comparison in daytime (Data are from DARPA. The top row shows the results by our method and the bottom shows the results of IVT.)

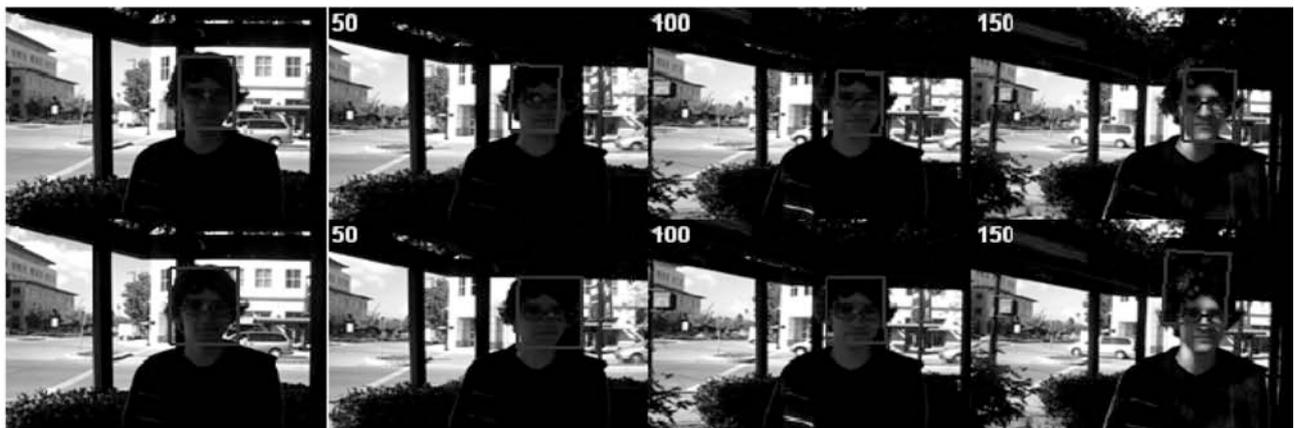


图 14 人脸跟踪结果比较 (数据来自文献 [16] 发布的数据库. 第 1 行为本文算法结果, 第 2 行为 IVT 算法结果)
 Fig.14 Face tracking results comparison (Data are from [16]. The top row shows the result by our method and the bottom shows the result of IVT.)

表 2 DAPAR 数据库下本文算法与 IVT 算法跟踪精度和更新频率比较
 Table 2 Comparison of tracking results and updating frequencies on DARPA dataset

目标编号	存在时间	本文算法跟踪时间 (帧)	IVT 跟踪时间 (帧)	本文算法更新次数	IVT 更新次数	跟踪时间之比	更新次数之比
1	1 ~ 380	1 ~ 380	1 ~ 380	17	76	1.00	0.224
2	1 ~ 280	1 ~ 280	1 ~ 280	14	56	1.00	0.250
3	1 ~ 210	1 ~ 210	1 ~ 210	9	42	1.00	0.214
4	1 ~ 1821	1 ~ 1040	1 ~ 310	31	208	3.35	0.149
5	1 ~ 1821	1 ~ 1810	1 ~ 400	59	362	4.53	0.163

一给定的频率来更新模板, 这个频率通常不会很低, 因为过低会使得算法鲁棒性很差, 在场景发生变化时容易跟踪失败. 考虑更新决策之后, 更新频率与场景发生变化的情况有关, 从表 2 和表 3 统计结果来看, 一般情况下自适应更新决策会使得更新频率会降低很多, 因此计算和时间复杂度也会下降很多.

4 结论

本文提出了一种基于自适应更新决策的目标跟踪算法, 在跟踪中引入跟踪状态判决和自适应更新决策, 根据判决结果等高层知识反馈给底层处理, 并在此基础上, 针对子空间增量学习的粒子滤波跟踪算

表 3 文献 [16] 发布的数据库下本文算法与 IVT 算法跟踪精度和更新频率比较
Table 3 Comparison of tracking results and updating frequencies on dataset in [16]

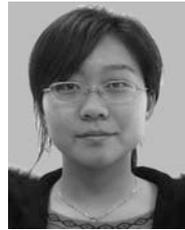
目标	存在时间	本文算法跟踪时间 (帧)	IVT 跟踪时间 (帧)	本文算法更新次数	IVT 更新次数	跟踪时间之比	更新次数之比
鱼	1 ~ 476	1 ~ 476	1 ~ 476	15	95	1.00	0.158
人脸	1 ~ 501	1 ~ 501	1 ~ 132	93	100	3.80	0.930
玩具狗	1 ~ 1344	1 ~ 1280	1 ~ 620	103	256	2.06	0.402
白天车辆	1 ~ 659	1 ~ 659	1 ~ 659	18	131	1.00	0.137
夜晚车辆	1 ~ 393	1 ~ 393	1 ~ 280	12	78	1.40	0.154

法设计了具体的自适应判决和决策准则, 提出了基于增量子空间自适应决策的目标跟踪算法. 实验结果表明, 与 IVT 跟踪算法相比, 本文算法能够大幅度提高跟踪鲁棒性, 在目标外观变化、尺度缩放、旋转、亮度变化等复杂条件下, 仍然能够达到很好的跟踪效果. 此外, 由于自适应更新决策机制更加符合环境条件的变化节奏, 能够省掉不必要的模板更新, 而只在必要的时候对模板进行自适应更新, 从而提高算法的效率. 实验结果验证了本文算法的有效性.

References

- Cannons K. A review of visual tracking [Online], available: <http://www.cse.yorku.ca/techreports/2008/?abstract=CSE-2008-07>, Oct 31, 2010
- Tu J, Tao H, Huang T. Online updating appearance generative mixture model for mean shift tracking. *Lecture Notes in Computer Science*. New York: Springer, 2006. 694–703
- Shan C, Tan T, Wei Y. Real-time hand tracking using a mean shift embedded particle filter. *Pattern Recognition*, 2007, **40**(7): 1958–1970
- Wang X, Wang S, Ma J. An improved particle filter for target tracking in sensor systems. *Sensors*, 2007, **7**(1): 144–156
- Wang T, Gu I, Backhouse A, Shi P. Face tracking using Rao-Blackwellized particle filter and pose-dependent probabilistic PCA. In: Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, USA: IEEE, 2008. 853–856
- Silveira G, Malis E. Real-time visual tracking under arbitrary illumination changes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 1–6
- Buenaposada J, Munoz E, Baumela L. Efficient illumination independent appearance-based face tracking. *Image and Vision Computing*, 2009, **27**(5): 560–578
- Kwon J, Lee K M. Visual tracking decomposition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 1269–1276
- Yu Q, Dinh T B, Medioni G. Online tracking and reacquisition using co-trained generative and discriminative trackers. *Lecture Notes in Computer Science*. Berlin: Springer, 2008. 678–691
- Wang P, Ji Q. Robust face tracking via collaboration of generic and specific models. *IEEE Transactions on Image Processing*, 2008, **17**(7): 1189–1199
- Zhang B, Tian W, Jin Z. Efficient hybrid appearance model for object tracking with occlusion handling. *Optical Engineering*, 2007, **46**(8): 087202-1–087202-11
- Jepson A, Fleet D, El-Maraghi T. Robust online appearance models for visual tracking. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2001. 415–422
- Zhou S, Chellappa R, Moghaddam B. Appearance tracking using adaptive models in a particle filter. In: Proceedings of the 6th Asian Conference on Computer Vision. Jeju, Korea: Asian Federation of Computer Vision Societies, 2004. 1–9
- Babenko B, Yang M, Belongie S. Visual tracking with online multiple instance learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 983–990
- Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking. *Lecture Notes in Computer Science*. Berlin: Springer, 2008. 234–247
- Ross D, Lim J, Lin R, Yang M. Incremental learning for robust visual tracking. *International Journal of Computer Vision*, 2008, **77**(1–3): 125–141
- Ross D, Lim J, Yang M. Adaptive probabilistic visual tracking with incremental subspace update. *Lecture Notes in Computer Science*. Berlin: Springer, 2004. 470–482
- Wu G, Tang Z. Adaptive appearance tracking model using subspace learning method. In: Proceedings of the International Conference on Intelligent Computation Technology and Automation. Changsha, China: IEEE, 2010. 413–416
- Yu G, Hu Z, Lu H. Robust incremental subspace learning for object tracking. *Lecture Notes in Computer Science*. Berlin: Springer, 2009. 819–828

- 20 Li M, Chen W, Huang K, Tan T. Visual tracking via incremental self-tuning particle filtering on the affine group. In: Proceedings of the IEEE Conference on Computer Vision and Patter Recognition. San Francisco, USA: IEEE, 2010. 1315–1322
- 21 Wen Jing, Li Jie, Gao Xin-Bo. Adaptive object tracking with incremental tensor subspace learning. *Acta Electronica Sinica*, 2009, **37**(7): 1618–1623
(温静, 李洁, 高新波. 基于增量张量子空间学习的自适应目标跟踪. 电子学报, 2009, **37**(7): 1618–1623)
- 22 Wen J, Gao X, Li X, Tao D, Li J. Incremental pairwise discriminant analysis based visual tracking. *Neurocomputing*, 2010, **74**(1–3): 428–438
- 23 Wen J, Gao X, Yuan Y, Tao D, Li J. Incremental tensor biased discriminant analysis: a new color-based visual tracking method. *Neurocomputing*, 2010, **73**(4–6): 827–839
- 24 Li X, Hu W, Zhang F, Zhang X, Luo G. Robust visual tracking based on incremental tensor subspace learning. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1–8
- 25 Yang J, Zhang D, Frangi F, Yang J. Two-dimensional PCA: a new approach to appearance based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(1): 131–137
- 26 Li G, Liang D, Huang Q, Jiang S, Gao W. Object tracking using incremental 2d-LDA learning and Bayes inference. In: Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, USA: IEEE, 2008. 1568–1571
- 27 Wang T, Gu I, Shi P. Object tracking using incremental 2D-PCA leaning and ML estimation. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Honolulu, USA: IEEE, 2007. 933–936
- 28 Levy A, Lindenbaum M. Sequential Karhunen-Loeve basis extraction and its application to images. *IEEE Transactions on Image Processing*, 2000, **9**(8): 1371–1374



仝小敏 西北工业大学计算机学院博士研究生. 2010 年获西北工业大学计算机科学与技术专业硕士学位. 主要研究方向为视频目标检测与跟踪, 网络多摄像机协同计算, 视频监控.

E-mail: xmtongnwpu@gmail.com

(TONG Xiao-Min Ph.D. candidate at the School of Computer Science,

Northwestern Polytechnical University. She received her master degree from the School of Computer Science, Northwestern Polytechnical University in 2010. Her research interest covers video target detection and tracking, multi-camera network cooperative processing, and visual surveillance.)



张艳宁 西北工业大学教授. 主要研究方向为计算机视觉, 图像与视频处理, 智能信息处理, 模式识别. 本文通信作者.

E-mail: ynzhang@nwpu.edu.cn

(ZHANG Yan-Ning Professor at Northwestern Polytechnical University.

Her research interest covers computer vision, image and video signal processing, intelligence information processing, and pattern recognition. Corresponding author of this paper.)



杨涛 西北工业大学副教授. 2008 年获西北工业大学控制理论与控制工程专业博士学位. 主要研究方向为视频内容分析, 视频目标检测与跟踪, 图像配准, 相机阵列合成孔径成像.

E-mail: yangtaonwpu@163.com

(YANG Tao Associate professor at the School of Computer Science, North-

western Polytechnical University. He received his Ph.D. degree from the School of Automation, Northwestern Polytechnical University in 2008. His research interest covers video content analysis, visual object detection and tracking, image registration, and camera array synthetic aperture imaging.)