

文章编号:1004-4213(2010)06-1047-6

# 一种结合粒子滤波和张量子空间的目标跟踪算法\*

温静,李洁,高新波

(西安电子科技大学 电子工程学院,西安 710071)

**摘要:**由于传统的子空间方法易于丢失图像目标的二维特性,为此本文提出了一种新颖的自适应目标跟踪算法,通过张量的方式建立目标的外观模型——张量子空间,利用在线学习的方法更新其外观模型,同时,利用目标仿射运动的先验信息,通过粒子滤波自适应地跟踪运动目标,并将获得的最优目标观测作为新数据反馈回子空间更新.此外,为了保证子空间更新能获得精确且紧致目标子空间表达,引入动态部分函数滤除样本野点.实验结果表明,本文提出的自适应目标跟踪方法具有较强的鲁棒性,对于存在姿态变化、短时遮挡和光照变化等情况下均可有效地跟踪目标.

**关键词:**模式识别;目标跟踪;张量子空间;在线学习;粒子滤波;仿射运动;DPF

**中图分类号:**TP391.4

**文献标识码:**A

**doi:**10.3788/gzxb20103906.1047

## 0 引言

目标跟踪易受到采集设备分辨率的限制、环境光线的变化以及非刚体运动等因素的影响,从而造成目标的丢失.现有的方法往往提取目标的不变性特征,并结合运动预测进行跟踪. Condensation 算法<sup>[2]</sup>成功将粒子滤波引入视觉跟踪.文献[8]通过建立目标的颜色直方图,结合粒子滤波和均值漂移设计了一种基于颜色分布的混合算法.文献[11]则对图像进行两级塔式分解构建边缘方向直方图,利用目标和背景的纹理差异,通过粒子滤波进行目标跟踪.为了防止出现目标丢失的问题,文献[9]则利用检测和跟踪相结合的方式,将 AdaBoost 和粒子滤波有机的结合在一起.然而,颜色、边缘、纹理等这些特征往往对光照变化以及背景复杂比较敏感,在实际应用上受到了限制.

近年来, Ross 等人<sup>[3-4,12]</sup>借鉴文献[1],通过增量学习更新目标外观子空间,该方法不仅极大地提高了特征跟踪的效率,获得了较 Eigentracking 更为鲁棒的效果.这种基于子空间的方法获得较大的成功,但是这些方法都将图像拉伸为向量,而丢失了目标的部分结构信息和邻域关系,因此,必须采用能够保持图像邻域结构的特征来描述目标.文献[10]提

出一种基于三阶张量的增量子空间学习方法跟踪目标,利用三阶张量表示视频序列段(图像帧数大于 1),并基于三阶张量构建和实时更新张量子空间,该方法将视频空时的变化也引入子空间中,对于存在噪音的视频获得较好的跟踪效果,然而视频中目标外观随时间的变化并不服从正态分布,增加时域的特征基易造成目标二维外观子空间的不精确.

本文采用粒子滤波作为跟踪的框架,建立二阶张量子空间来表示视频中的序贯目标外观数据流.然后,对仿射运动参量进行采样,获得目标位置的最大后验概率估计值;同时,提出了一种张量子空间在线学习(Online Tensor Subspace Learning, OTSL)方法来更新二阶张量子空间.最后,在更新子空间的在线学习过程中,通过<sup>[6]</sup>动态部分函数(Dynamic Partial Function, DPF)滤除不适于用作子空间构造的观测值.

## 1 目标的外观模型——张量子空间

与动态张量分析(Dynamic Tensor Analysis, DTA)相类似<sup>[7]</sup>.在更新子空间时,未考虑到均值和方差的更新.本文提出的在线学习将对这两个参量作出修正. $d$  模的方差矩阵  $C_d$  定义为

$$C_d = \sum_j ((\mathbf{X}_j - \mathbf{M}^{(X)})_{(d)}^T (\mathbf{X}_j - \mathbf{M}^{(X)})_{(d)}) \quad (1)$$

式中

$$\mathbf{M}^{(X)} = \frac{\lambda n \mathbf{M}_{old}^{(X)} + m \mathbf{M}_{new}^{(X)}}{\lambda n + m} \quad (2)$$

$$\mathbf{M}_{old}^{(X)} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j, \mathbf{M}_{new}^{(X)} = \frac{1}{m} \sum_{j=1}^m \mathbf{X}_j \quad (3)$$

式中  $\cdot_{(d)} \in \mathbf{R}^{(\prod_{i \neq d} N_i) \times N_d}$  为张量模  $d$  矩阵化表示,  $n$

\* 国家自然科学基金(60702061)、教育部长江学者和创新团队支持计划(IRT0645)、深圳大学 ATR 国防科技重点实验室开放基金和总装备部预研基金(9140A06050107DZ0113)资助

Tel: 029-88201838

Email: berbermimi@yahoo.cn

收稿日期: 2008-10-09

修回日期: 2009-03-16

为旧有的张量数,  $m$  为新到达的张量数. 通过对角化  $C_d$ , 我们更新  $C_d = U_d E_d U_d^T$  的投影矩阵, 其中  $U_d$  为正交矩阵,  $E_d$  为对角矩阵. 学习算法如图 1, 保留存储  $U_d$  和  $E_d$ ; 当有新数据到达时, 需要利用  $U_d$  和  $E_d$  重构出旧方差, 联合新数据计算新获得方差

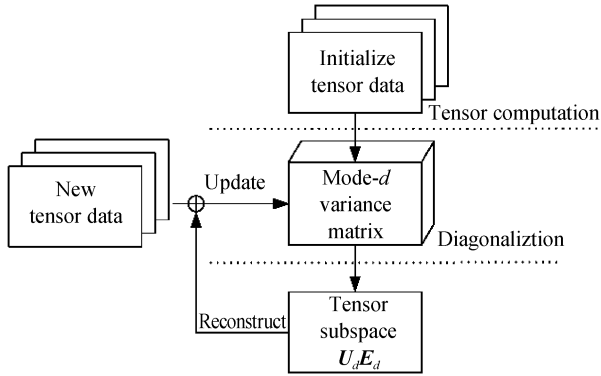


图 1 张量子空间在线学习

Fig. 1 Online learning for tensor subspace

$$C_d^{\text{new}} = \sum_{j=1}^m ((X_j - M_{\text{new}}^{(X)})_{(d)}^T (X_j - M_{\text{new}}^{(X)})_{(d)}) + \frac{nm}{n+m} (M_{\text{old}}^{(X)} - M_{\text{new}}^{(X)})_{(d)}^T (M_{\text{old}}^{(X)} - M_{\text{new}}^{(X)})_{(d)} \quad (4)$$

从而计算出新的  $U_d$  和  $E_d$ , 其中,  $C_d^{\text{new}}$  中后一项为方差矩阵的修正项. 在本文的目标跟踪应用中, 本文采用的二阶张量来表示目标, 张量子空间在线学习算法的详细步骤为:

Step1: 矩阵化初始目标图形成块为  $X_{(d)} \in \mathcal{R}^{\prod_{i \neq d} N_i \times N_d}$ ,  $d=1, 2$ , 保留现有张量数为  $n$ ;

Step2: 用式(2)计算新均值  $M^{(X)}$ , 保留现有均值为旧均值  $M_{\text{old}}^{(X)}$ ;

Step3: 重构每一“模”方差矩阵  $C_d^{\text{odd}} \leftarrow U_d E_d U_d^T$ , 通过式(4)和(1)分别计算  $C_d^{\text{new}}$  以及每一“模”下新方差  $C_d \leftarrow \lambda C_d^{\text{odd}} + C_d^{\text{new}}$ , 并计算  $C_d$  获得  $U_d$  和  $E_d$ , 输出并保留现有  $U_d$  和  $E_d$ ;

Step4: 更新张量数为  $n \leftarrow \lambda n + m$ ;

Step5: 数据结束; 否则跳到 step2.

在张量子空间的在线学习算法中,  $\lambda$  为遗忘因子该在线学习算法的空间存储消耗为  $(m+1) \times \prod_{i=1}^M N_i + \sum_{i=1}^M N_i \times R_i + \sum_{i=1}^M R_i$ .

## 2 结合粒子滤波和张量子空间的目标自适应跟踪算法

### 2.1 粒子滤波

由于跟踪目标的状态只与前一时刻的状态有关, 跟踪算法可以看作是概率 Markov 的推理问题, 因此可以通过递归的贝叶斯表述——粒子滤波来整合时域信息. 在时刻  $k$  观测到序列的图像块  $W_k$ , 隐状态变量  $H_k$  为目标运动的仿射参量. 假设给定目标的状态参量  $H_k$  时, 观测到的图像区域  $W_k$  是目标

外观从目标张量子空间中采样生成的. 图像帧间的目标运动可以建模为一阶马尔科夫模型, 服从概率分布  $p(H_k | H_{k-1})$ , 表示在给定目标在时刻  $k-1$  的状态参量  $H_{k-1}$ , 目标出现在  $H_k$  的概率. 给定观测值  $W_k$ , 对于状态参量为  $H_k$  分布为  $p(W_k | H_k)$  的目标建立相似性. 利用状态参量的先验知识, 根据贝叶斯准则合并观测值  $p(H_k | W_k, H_{k-1}) \propto p(W_k | H_k) p(H_k | H_{k-1})$ , 可以得到后验目标参量的最大值  $H_k^*$ .

初始先验状态参量时假设每个仿射参量, 如平移、尺度、旋转和裁剪等, 为相互独立分布的, 即分布在预定义参量  $H_0$  的周围, 分别服从一个正态分布, 本文涉及到仿射参量为:  $H_0 = (x_0, y_0, r_0, s_0, k_0)$ , 其中  $x_0, y_0, r_0, s_0, k_0$  分别为 0 时刻的  $x, y$  方向平移、旋转角度、尺度、高宽比以及裁剪方向. 进而, 任意时刻  $k$  参量  $H_k$  中每个参量可以独立建模为高斯分布, 以前时刻参量  $H_{k-1}$  中对应部分为中心.

$$p(H_k | H_{k-1}) = N(H_k; H_{k-1}, \Psi) \quad (5)$$

式中  $\Psi$  为对角化协方差矩阵, 其每个元素为对应仿射参量的方差.

由于状态参量  $H_k$  不能直接观察到, 而全贝叶斯推理要求计算每个时刻  $P(H_k | W_k, W_{k-1}, \dots, W_1, H_0)$ , 因此无法获得该分布的闭式解. 但是, 我们可以通过一个正态分布来逼近该闭式解, 这个正态分布以  $p(H_k | W_k, H_{k-1}^*)$  的最大值  $H_k^*$  为中心.

近似值  $H_k^*$  可以用简单的采样方法获得, 根据先验  $p(H_k | H_{k-1}^*)$  抽取出一些样本. 对于每个样本  $H_s$ , 其后验概率由  $p_s = p(H_s | W_k, H_{k-1}^*)$  计算获得. 后验概率  $p_s$  为  $H_s$  采样概率与  $H_s$  在概率主成分分析分布下概率值的乘积. 这里不考虑归一化因子, 因为对于所有样本该值为一常量. 最后, 具有最大后验值的样本被选为闭式解的近似值  $H_k^*$ , 即

$$H_k^* = \arg \max_{H_s} p(H_s | W_k, H_{k-1}^*) \quad (6)$$

本文采用第一节的方法对目标外观进行建模, 则  $p(H_k | W_k, H_{k-1})$  转化为采样的图像到目标张量子空间的距离, 而式(6)转化为求解满足最小化该距离的状态参量. 同时, 由于外观随时间变化, 固定的观测模型无法适应目标外观的变化而不能捕捉到目标, 为了适应目标外观的变化, 我们根据在线学习算法, 利用最优的观测图像块更新目标的张量子空间子空间.

### 2.2 野点滤除: DPF 滤波

考虑到跟踪过程中, 由于遮挡或其它外界条件的变化会出现一些干扰的观测值, 而这些观测值由于远离目标外观子空间, 使得这些观测值尽管在不同程度上是目标, 但仍然属于样本野点, 我们希望在

更新张量子空间时,能够去除非目标观测对子空间的误差积累影响.因此,引入距离函数测量两幅图像的相似性<sup>[6]</sup>,以预防干扰值参与构造子空间.本文采用 DPF

$$D(W_k, W_{k-1}^*) = \left( \sum_{\Delta d_i \in \Delta_l} \Delta d_i^r \right)^{\frac{1}{r}} < T \quad (7)$$

式中  $l$  和  $r$  为可调节参量.计算当前时刻的最大观测值  $W_k$  与前一时刻估计值  $W_{k-1}^*$  的距离,允许低于阈值  $T$  的观测值参与构造张量子空间.

### 2.3 自适应目标跟踪算法

结合粒子滤波和张量子空间的自适应目标跟踪算法,通过在张量子空间中对跟踪目标的外观建立观测模型,利用在线学习的方式更新随运动而变化的目标外观子空间,并在贝叶斯推理的预测下,获得准确的目标观测.在没有训练图像的情况下,给定目标的初始参量,跟踪仍然能逐步地根据输入图像构建并更新张量子空间.

图 2 给出了该自适应目标跟踪算法的基本流程,首先初始化视频数据的状态参量,根据粒子滤波理论进行参量采样,并根据已经建立的张量子空间估计出最优参量,同时将最优参量对应的观测目标进行野值点滤除,对于非野值点的观测值,则利用张量子空间的在线学习算法进行更新.具体步骤为:

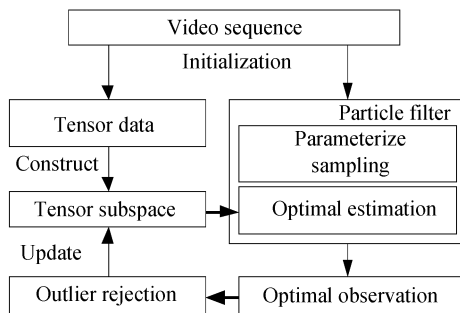


图 2 自适应目标跟踪算法流程

Fig. 2 Flowchart of adaptive object tracking algorithm

Step1: 初始化目标状态的仿射运动参量;

Step2: 利用文献[7]中的 OTA 方法初始化目标外观的张量子空间;

Step3: 利用上一时刻的状态参量  $H_{k-1}$  以及这些参量的先验信息进行采样,获得采样状态参量对应的观测图像块  $W_k$ ;

Step4: 利用式(6)估计该时刻最优的目标状态参量并输出目标观测;

Step5: 用 2.2 节的方法判断 Step4 输出的观测值是否是野点,如果是,则跳到 Step3 处理  $k+1$  时刻的数据,如果不是,则跳到 step6;

Step6: 通过第 1 节的在线学习方法更新张量子空间,直到视频数据结束,算法流程图参见图 2.

## 3 实验结果与分析

为了验证所提出的跟踪算法在精确跟踪目标位置和更新外观模型的性能,设计了五个测试实验.实验中使用了人脸图像序列和两段自然的视频序列,其中,人脸图像序列来自 Oriental Face 数据库,“dudek”序列来自文献[12],其余一段是人工采集的视频.实验一通过对单人图像序列进行重构,验证本文所提出的在线学习方法的有效性;实验二通过“dudek”序列验证本文算法能够获得与文献相同的跟踪效果;实验三通过人工采集的三段视频验证本文提出的方法分别对于目标有遮挡,光照不均匀,姿态存在较大变化的情况都能取得较好的跟踪效果,实验四通过以上四段视频的一组对比实验验证本文提出的算法优于文献[12].

实验中参量设定:张量子空间实行每 5 帧更新,即张量子空间在线学习算法中的  $m$  为 5.

实验一:如图 3,实验选取了 Oriental Face 数据库中单人不同姿态的 18 幅图像,第一行为该人物姿态的原始图像,对该图像序列分别通过 ISL<sup>[12]</sup>以及本文提出的 OTSL 法构建人物姿态的子空间,再用各子空间重构这 18 幅图像,图 3(b)和(c)分别是这两种方法重构出的人物姿态图像.图 3(b)中,通过对 PCA 引入在线学习方法,可以及时地捕捉人脸姿态的变化,但是由于子空间图像被拉成向量,丢失了图像结构信息,使得重构出的图像出现混叠现象,注意,该实验中子空间实行每帧更新,当在线学习的数目增大时,会出现更严重的混叠问题;图 3(c)中,本文提出的方法不仅能及时捕捉姿态的变化,同时保留了较完整的图像信息,获得优于以上三种方法的重构效果.实验表明本文提出的张量子空间在线学习方法能有效的刻画图像结构和邻域信息.

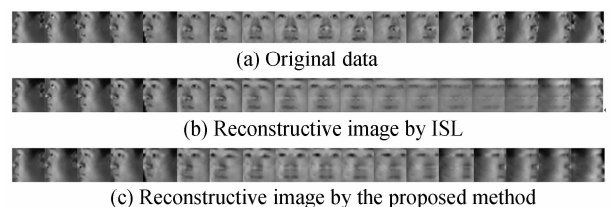


图 3 重构姿态图像的对比实验

Fig. 3 Comparison of reconstructive pose images

实验二:图 4 显示了“dudek”序列的跟踪结果.

“dudek”视频序列分辨率为  $480 \times 720$ ,该视频的主要特点是,人物存在较剧烈的姿态变化,同时伴有不规则的运动,尺度变化,表情变化以及自遮挡等.实验中的采样粒子数为 600 个.序列中跟踪目标为人脸,图 4 中显示了“dudek”序列中第 107,183,461

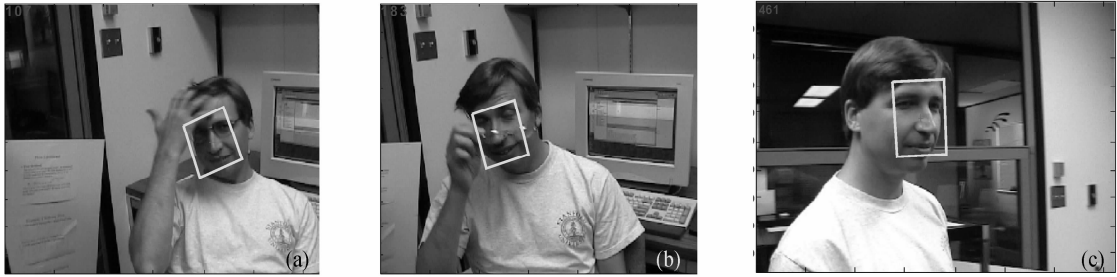


图 4 视频序列“dudek”跟踪结果

Fig. 4 Tracking result by the proposed method on video sequence “dudek”

帧,从图像上可以明显看到两处短时(部分)遮挡以及姿态剧变,本文提出的方法对于存在大幅姿态变化,短时遮挡,表情变化以及快速运动等的情况下,均可鲁棒的跟踪到目标人脸。

实验三:如图 5,视频序列“occlusion2”分辨率



图 5 视频序列“occlusion2”的跟踪结果

Fig. 5 Tracking result by the proposed method on video sequence “occlusion2”

实验四:恢复了文献[12]中的方法.在文献[12]中,跟踪方法采用传统基于向量的方法,即增量子空间学习(ISL),寻找线性子空间表达目标并更新该子空间.该实验主要关注用文献[12]和本文方法跟踪效果的视觉对比,如图 6,各列分别为序列“dudek”、和“occlusion2”的对比跟踪结果。



(a) Tracking result by method in reference[12]



(b) Results by the proposed method

图 6 与文献[12]对比“dudek”和“occlusion2”的跟踪结果  
Fig. 6 Comparison of tracking performance with Ref[12] in video sequence “dudek” and “occlusion2”

第一行(a)为文献[12]中的方法的跟踪结果,在“dudek”和“occlusion2”跟踪过程中,多处出现人脸

为  $384 \times 512$ ,采样粒子数为 400.该视频的主要特点是,视频中目标人脸运动速度缓慢,目标周围存在其他人脸以及自身遮挡干扰,跟踪算法依然可以有效地捕捉到目标人脸,而没有受到遮挡的干扰影响。

目标的短时遮挡,非缓变的环境变化以及快速运动,掺入了非目标噪音或不精确的人脸图像样本,使得丢失了跟踪目标,所构造的人脸子空间由于失去了目标外观的结构信息,而不能客观的描述目标,导致跟踪失败。

第二行(b)为本文的方法的跟踪结果,尽管有目标的遮挡以及光线变化而无法分辨目标和背景,由于保持了图像原始的表达方式,从而较好地保留了图像的结构信息,所建立的张量子空间精确而紧致,进而可以稳定地捕捉到目标。

正如在 2.2 节中所讲到的,观测值中存在的干扰会影响到子空间的构造,如果观测值中出现了过多的干扰或样本野点(由于遮挡等情况),会引起子空间构造时误差的积累.为了保证紧致和纯正的表达,有必要通过 DPF 去除干扰。

实验结果表明,本文提出的方法可以有效地解决跟踪过程中姿态较大变化,光照不均匀变化以及遮挡等问题。

实验五:如图 7,分析了上述两段视频序列中目标的子空间学习效果,通过重构误差(Reconstruction error)这一指标定性地对 ISL 和本文提出的 OTSL 子空间学习方法.重构误差显示了子空间学习目标外观的能力情况,重构误差越

低,表明子空间学习时损失的能量越小,从而更好地获得了目标外观信息.由于张量的数据表达形式能够保留图像的二维结构信息,因而所获得的张量子空间可以获得更加紧致的目标表达.图7为子空间均降为两维的情况下,本文的方法获得了略胜于 ISL 方法的重构误差.

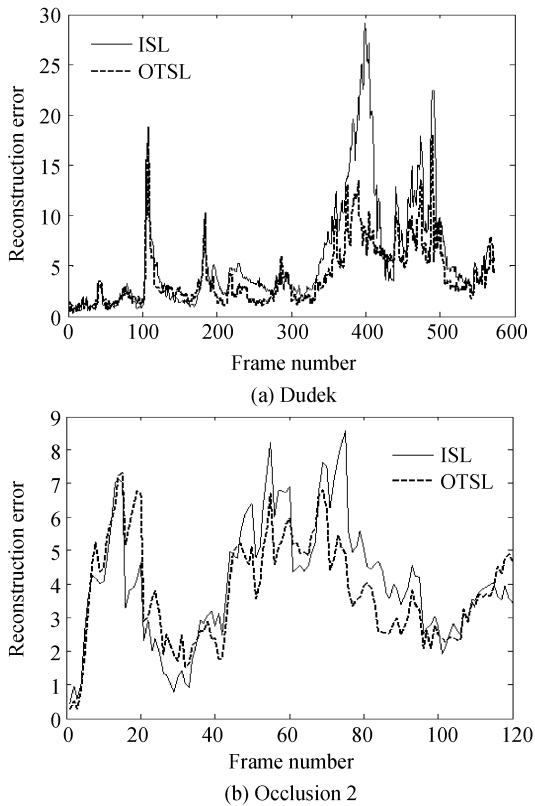


图7 视频“dudek”,以及“occlusion2”中跟踪目标的重构误差对比

Fig. 7 Comparisons of reconstruction error in video sequence “dudek” and “occlusion2”

## 4 结论

本文提出了一种有效的基于外观在线学习的目标跟踪方法.首先,提出了一种基于张量子空间的在线学习算法,其次,构造了一个基于粒子滤波的跟踪框架,通过对目标仿射运动参量进行采样,获得目标下时刻可能的位置及参量,利用后验概率估计的最大值得到预测;然后,提出了利用二阶张量子空间描述运动目标,并在线的更新空间,以保证空间的变化与运动目标的变化相一致;最后,引入了 DPF 滤波,确保了参与构造和更新子空间的观测值来自真正的

目标集合,提高了子空间的精确性.实验结果表明,本文提出的方法利用张量对图像高维线性的表达以及子空间能紧致描述目标的特点,在运动目标的外观变化剧烈,如,光照变化,姿态变化,表情变化,不规则快速运动以及短时遮挡的情况下,均可有效精确的跟踪目标,实验结果验证了本文算法的有效性.

## 参考文献

- [1] BLACK M J, JEPSON A D. EigenTracking: Robust matching and tracking of articulated objects using a view based representation [J]. *IJCV*, 1998, **26**(1): 63-84.
- [2] ISARD M, BLAKE A. CONDENSATION: Conditional density propagation for visual tracking [J]. *IJCV*, 1998, **29**(1): 5-28.
- [3] LIM J, ROSS D, LIN R S, *et al.* Incremental learning for visual tracking [C]. *Advances in Neural Information Processing Systems (NIPS' 2004)*, Vancouver, Canada, the MIT Press, 2004: 793-800.
- [4] LIN R S, ROSS D, LIM J, *et al.* Adaptive discriminative generative model and its applications [C]. *Advances in Neural Information Processing Systems (NIPS' 2004)* Vancouver, Canada, the MIT Press, 2004: 801-808.
- [5] CAI D, HE X F, HAN J W. Subspace learning based on tensor analysis [P]. *UIUCDCS-R -2005-2572*, 2005.
- [6] QAMRA A, MENG Y, CHANG E Y. Enhanced perceptual distance functions and indexing for image replica recognition [J]. *IEEE Trans. Pattern Recognition and Machine Intelligence*, 2005, **27**(3): 379-391.
- [7] SUN J M, TAO D C, FALOUTSOS Y C. Beyond streams and graphs: dynamic tensor analysis [C]. *The Twelfth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD)*, Philadelphia, USA, 2006, 374-383.
- [8] LU Xiao-peng, YIN Xue-min, ZOU Mou-yan. A hybrid algorithm of object tracking based on color distribution [J]. *Journal of Electronics & Information Technology*, 2006, **34**(12): 2150-2153.  
卢晓鹏, 殷学民, 邹谋炎. 一种基于颜色分布的混合视频跟踪方法 [J]. *电子与信息学报*, 2006, **34**(12): 2150-2153.
- [9] LEI Yun, DING Xiao-qing, WANG Sheng-jin. AdaBoost tracker embedded in particle filtering [J]. *Journal of Tsinghua University (Sci & Tech)*, 2007, **47**(7): 1141-1143.  
雷云, 丁晓青, 王生进. 嵌入粒子滤波中的 AdaBoost 跟踪器 [J]. *清华大学学报(自然科学版)*, 2007, **47**(7): 1141-1143.
- [10] LI X, HU W M, ZHANG Z F, *et al.* Robust visual tracking based on incremental tensor subspace learning [C]. *Proc. Int'l Conf. on Computer Vision*, 2007: 1-8.
- [11] LIU Yang, LI Yu-shan, ZHANG Da-pu, *et al.* Visual Tracking based on dynamic object modeling in particle filter [J]. *Acta Photonica Sinica*, 2008, **37**(2): 375-380.  
刘洋, 李玉山, 张大朴, 等. 基于动态目标建模的粒子滤波视觉跟踪算法 [J]. *光子学报*, 2008, **37**(2): 375-380.
- [12] ROSS D, LIM J, LIN R S, *et al.* Incremental learning for robust visual tracking [J]. *IJCV*, 2008, **77**(1-3): 125-141.

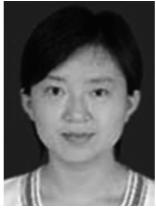
## An Adaptive Object Tracking Algorithm Combined with Particle Filter and Tensor Subspace

WEN Jing, LI Jie, GAO Xin-bo

(School of Electronic Engineering, Xidian University, Xi'an 710071, China)

**Abstract:** In order to overcome the disadvantage that traditional subspace methods usually lose the two-dimensional information of the objects in image, a novel adaptive object tracking method is proposed. The appearance of the object in tensor subspace is modeled and the object model is updated with online learning method. The object is tracked by using particle filter and the prior of affine motion, and the optimal observation is feeded back to the tensor subspace updating. Moreover, DPF is introduced into the subspace updating to reject outliers so as to keep the object subspace precise and compact. The proposed method is able to track targets effectively and robustly under pose variation, short-time occlusion and large lighting and so on in the experiments.

**Key words:** Pattern recognition; Object tracking; Tensor subspace; Online learning; Particle filter; Affine motion; DPF



**WEN Jing** was born in 1982. She is currently a Ph. D. degree candidate, and her research interests focus on computer vision and pattern recognition.



**GAO Xin-bo** was born in 1972. He received the BSc, MSc and Ph. D. degrees in signal and information processing from Xidian University, China, in 1994, 1997 and 1999 respectively. He is a Professor of Pattern Recognition and Intelligent System, and Director of the VIPS Lab, Xidian University. His research interests are computational intelligence, machine learning, computer vision, pattern recognition and artificial intelligence.