

基于增量张量子空间学习的自适应目标跟踪

温 静,李 洁,高新波

(西安电子科技大学电子工程学院,陕西西安 710071)

摘 要: 传统的基于子空间的跟踪方法易于丢失图像所固有的部分结构和邻域信息,从而降低了目标匹配和跟踪的精度.为此,本文提出了一种增量张量子空间学习算法,用于跟踪目标的建模与模型更新.同时,将该模型与贝叶斯推理相结合,提出一种自适应目标跟踪算法.新方法首先对跟踪目标的外观进行建模,然后利用贝叶斯推理获得目标外观状态参数的最优估计,最后利用最优估计的目标观测更新目标张量子空间.实验结果表明,由于保持了目标外观的结构信息,本文提出的自适应目标跟踪方法具有较强的鲁棒性,在跟踪目标在姿态变化、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: 张量子空间;增量学习;贝叶斯推理;仿射运动

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2009) 07-1618-06

Adaptive Object Tracking with Incremental Tensor Subspace Learning

WEN Jing, LI Jie, GAO Xin-bo

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

Abstract: The conventional subspaces based tracking methods usually have low precision of object matching and tracking, because they lose the inherent partial structure and neighborhood information. In this paper, an incremental tensor subspace learning algorithm is proposed to model and update the object appearance in tensor subspace. Simultaneously, by combining the proposed learning algorithm with Bayesian inference, an adaptive object tracking method is presented. Firstly, we represented the appearance of the object in tensor subspace; secondly, obtained the optimal estimation of the state parameters by Bayesian inference; finally updated the tensor subspace by using the optimal observation. Due to the construction information is maintained, 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: tensor subspace; incremental learning; bayesian inference; affine motion

1 引言

目标跟踪是计算机视觉领域中的经典问题,也是视频序列中人脸识别的前提和关键环节,广泛应用于视频监控、人机交互、视频会议系统等领域.现有的一些方法往往通过提取目标的不变性特征,如颜色、纹理、轮廓等,与运动预测的方法相结合进行跟踪,如,粒子滤波^[3],卡尔曼滤波,EM算法,贝叶斯推理^[11]和均值漂移^[4]等.但是这些特征往往对光照变化以及背景复杂比较敏感,从而限制了其应用的范围.为寻求更加稳定的目标外观描述特征,基于子空间的方法成为当前跟踪算法的主流方向^{[1,2,4][6~8]}之一,其核心^[1]是对训练数据进行主成分分析,从而建立张成目标子空间的一组特征基,对新获得的观测只需计算其与子空间的距离,即可

判断是否为待跟踪的目标.在此基础上,Ross等人^[6,7]借鉴文献^[5]的增量学习,提出了一种基于增量子空间学习(Incremental Subspace Learning, ISL)的目标跟踪,即时跟踪目标的外观变化,这不仅极大地提高了跟踪的效率,同时也使得所建立的目标子空间模型能刻画所有已经获得的目标观测,对于运动缓慢的目标获得较为鲁棒的跟踪效果.但是,这些子空间方法由于将图像拉伸为向量,丢失了目标图像信息的部分结构信息和邻域关系,为此,必须提取能够保持目标图像邻域结构的特征来描述目标,使得高维线性几何(即高阶张量几何)成为了模式识别领域新的研究热点^[9,10].文献^[10]利用三阶张量表示视频序列段(图像帧数大于1),由于引入时域信息,该方法对于运动缓慢且目标外观变化较小的运动获得成功的跟踪结果;但是,对于时间上运动剧烈或外

收稿日期:2008-07-31;修回日期:2008-11-25

基金项目:国家自然科学基金(No. 60702061);教育部长江学者创新团队支持计划(No. IRT0645);深圳大学 ATR 国防科技重点实验室开放基金;总装备部预研基金(No. 9140A06050107DZ0113)

观变化大的情况,算法的泛化能力变差.同时,三阶张量带来了大量的运算代价以及更多的存储空间 $\sum_{i=1}^3 N_i + \sum_{i=1}^3 N_i \times R_i + \sum_{i=1}^3 R_i + \sum_{i=1}^2 N_i$, 其中 R_i 为第 i 维所降的维数.

鉴于此,本文提出了一种自适应目标跟踪方法,该方法采用二阶张量子空间,比三阶张量空间节省了存储空间和运算时间,同时更加客观地保持了图像的二维结构信息.本文方法采用贝叶斯推理作为跟踪的框架,建立张量子空间来表示视频中的序贯式目标外观数据流.无需预先的训练过程,跟踪算法只需要初始化目标状态参数,然后,对仿射运动参数进行采样,获得目标位置的最大后验概率估计值;同时,提出了一种增量张量子空间学习方法来更新二阶张量子空间.最后,将该增量学习方法纳入贝叶斯推理的跟踪框架中,从而获得一种新的自适应目标跟踪算法.

2 增量张量子空间学习(ITSL)

张量可看作多维矩阵^[12,13],一个 M 阶张量 $\mathbf{X} \in \mathbf{R}^{N_1 \times N_2 \times \dots \times N_M}$ 表示为 X_{n_1, n_2, \dots, n_M} , 其中每个元素为 $X_{n_1, n_2, \dots, n_M}, 1 \leq n_i \leq N_i, 1 \leq i \leq M, N_i$ 为第 i 模的维度.在张量理论中,张量每一维对应一个“模”.

本文采用张量的形式表达数据,以保持图像数据结构信息的完整性,同时,受基于向量子空间的增量学习的启发,提出了一种增量的基于张量子空间学习算法(Incremental Tensor Subspace Learning, ITSL).本节将从张量子空间的初始化和增量学习两个方面进行阐述.

2.1 张量子空间初始化:TensorPCA

本质上, TensorPCA 是基于传统主成分分析理论的,它将数据投影到方差最大的张量子空间中,以保证重构误差最小.其目标函数表述如下:

$$\max_U \sum_j \|Y_j - M^{(Y)}\|^2 \quad (1)$$

其中 $Y = \sum_{i=1}^M X_i U_i$ (2)

$$M^{(Y)} = \frac{1}{n} \sum_{j=1}^n Y_j \quad (3)$$

其中 $U = [U_1, \dots, U_M]$, U_i 为模 i 的子空间, $M^{(Y)}$ 为 X 投影到 U 后 Y 的均值.这里同样也是通过保留能量最大的主成分,以最小化二阶张量的重构误差.

2.2 张量子空间的增量学习算法

本文提出的增量学习与动态张量分析(Dynamic Tensor Analysis, DTA)相类似^[12],然而,动态张量分析在更新子空间时,并未考虑到均值的更新.这里以处理一个“模”为例,定义 C_d 为 d 模下的方差矩阵,

$$C_d = \frac{1}{n} \sum_{j=1}^n \left((X_j - M^{(X)})_{(d)}^T (X_j - M^{(X)})_{(d)} \right) \quad (4)$$

其中 $M^{(X)} = \frac{nM_{old}^{(X)} + mM_{new}^{(X)}}{n+m}$ (5)

$$M_{new}^{(X)} = \frac{1}{m} \sum_{j=1}^m X_j \quad (6)$$

其中 $(\cdot)_{(d)} \in \mathbf{R}^{(i, d, N)} \times N_d$ 为张量模 d 矩阵化表示, n 为旧有张量数, 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 所示,保留存储虚线框中的数据量、均值、 U_d 和 E_d 四个变量;当有新数据到达时,可直接更新数据量和均值,而对于新的 U_d 和 E_d ,则需要利用已有的 U_d 和 E_d 重构出旧方差,联合新数据获得新方差,通过 SVD 分解计算出新的 U_d 和 E_d ,详细步骤如图 1 所示.

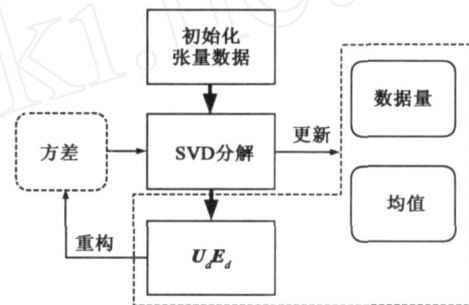


图1 基于张量子空间的增量学习示意图

Step 1: 矩阵化初始目标张量为 $X_{(d)} \in \mathbf{R}^{(i, d, N)} \times N_d$, 存储现有张量数为 n ;

Step 2: 用式(5)计算新均值 $M^{(X)}$, 存储现有均值为旧均值 $M_{old}^{(X)}$;

Step 3: 重构每一“模”方差矩阵 $C_d^{odd} = U_d E_d U_d^T$, 其中 $U_i |_{i=1}^M \in \mathbf{R}^{N_i \times R_i}, E_i |_{i=1}^M \in \mathbf{R}^{R_i \times R_i}$;

Step 4: 通过式(4)计算 C_d^{new} , 计算每一“模”下新方差 $C_d = C_d^{odd} + C_d^{new}$, 并对角化 C_d , 计算获得 U_d 和 E_d , 此时, $U_i |_{i=1}^M \in \mathbf{R}^{N_i \times R_i}, E_i |_{i=1}^M \in \mathbf{R}^{R_i \times R_i}$ 输出并保留现有 U_d 和 E_d ;

Step 5: 更新张量数为 $n = n + m$;

Step 6: 数据结束;否则跳到第二步.

在增量张量子空间学习算法中,遗忘因子 λ 用于减小或调节较早数据的对子空间更新的影响.该增量学习算法的空间存储消耗为:

$$(m+1) \times \sum_{i=1}^M N_i + \sum_{i=1}^M N_i \times R_i + \sum_{i=1}^M R_i$$

3 基于 ITSL 的目标跟踪算法

本节将上一节的增量学习算法纳入贝叶斯推理的概率框架下,提出了基于增量张量子空间学习(ITSL)的目标跟踪算法,不仅可以保存图像的结构信息,同时符合视觉跟踪对算法计算效率的要求,此外还节省了处

理数据的存储空间.

3.1 贝叶斯推理

本文在张量子空间中对运动目标建模, 给定目标状态参数 H_k , 假设观测数据 $\{W_k: W_k^i, i=1, 2, \dots, N\}$ 是目标外观从目标张量子空间中采样生成的, 其中, W_k^i 为 k 时刻第 i 个粒子处的观测数据, H_k 为隐状态变量, 一般为目标仿射运动参数. 图像帧间的目标运动可以建立一阶 Markov 模型, 服从概率分布 $p(H_k|H_{k-1})$. 在第 k 帧图像中的第 i 个粒子处的观测数据 W_k^i 的似然函数是通过计算观测数据 W_k^i 在已建立的张量子空间

$$\begin{aligned}
 & \text{在 } \sum_{i=1}^M \mathbf{x}_i \mathbf{U}_i \text{ 中的重构误差来定义, 即} \\
 & p(W_k^i | H_k) \\
 & \exp \left\{ - \left\| (W_k^i - M^{(w)}) - (W_k^i - M^{(w)}) \sum_{i=1}^M \mathbf{x}_i (\mathbf{U}_i \cdot \mathbf{U}_i^T) \right\|^2 \right\} \quad (7)
 \end{aligned}$$

其中, $M^{(w)}$ 为到目前为止所建立子空间的均值; $i=1, 2, \dots, N, N$ 为每帧图像中所取的粒子数目.

通过贝叶斯准则融合状态的先验知识以及观测值 $p(H_k|W_k, H_{k-1}) = p(W_k|H_k)p(H_k|H_{k-1})$, 可以得到目标参数后验的最优值 H_k^* .

由于外观随着时间变化, 固定的观测模型将不能捕捉到目标. 因此, 我们利用到最近的观测值来更新观测概率分布, 即通过第 2 节的增量学习方法更新目标的子空间模型.

3.2 自适应目标跟踪算法

本文提出的基于增量张量子空间学习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, 利用增量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, 具体算法如下:

Step 1: 初始化目标状态参数, 并利用 2.1 节的方法初始化目标外观的张量子空间;

Step 2: 利用 $k-1$ 时刻的目标状态及观测信息, 通过贝叶斯推理预测 k 时刻目标可能的位置;

Step 3: 估计 k 时刻的最优状态参数并输出目标的最优观测值;

Step 4: 通过 2.2 节的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.

目标的状态由其对应的五个仿射参数来描述, 即水平、竖直方向的平移量 x, y 、旋转角度 r 、尺度 s 、裁剪方向 k . 五个参数在第一帧上是通过手工标定的方式来初始化的, 获得初始参数 $H_0 = (x_0, y_0, r_0, s_0, k_0)$. 在下一帧图像中, 目标状态参数通过随机采样获得. 每个状态参数以前一帧的状态参数为均值, 以经验值 $\sigma = (x, y, r, s, k)$ 为方差, 产生高斯分布的随机数而

获得, 即

$$p(H_k|H_{k-1}) = N(H_k; H_{k-1}, \sigma) \quad (8)$$

其中, N 表示高斯函数或正态分布函数.

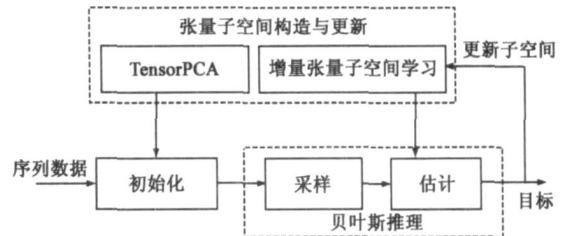


图2 基于增量张量子空间的跟踪算法流程图

4 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, 本节设计了四个测试实验, 实验中使用了一段人脸图像序列和三段自然的视频序列, 其中, 人脸图像序列来自 Oriental Face 数据库, “dudek”序列来自文献[7], 其余两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, 验证本文所提出的增量学习方法的有效性; 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两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, 实验四通过以上三段视频的一组对比实验验证本文提出的算法优于文献[7].

实验中参数设定: 张量子空间实行每 5 帧更新, 即张量子空间增量学习算法中的 m 为 5.

实验一 如图 3 所示, 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, 我们对该图像序列分别通过 ISL^[7]以及本文提出的方法 (ITSL) 构建人物姿态的子空间 (保留 90% 的能量), 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 引入增量学习方法, 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, 当增量学习的数目增大时, 会出现更严重的混叠问题; 图 3 (c) 中,

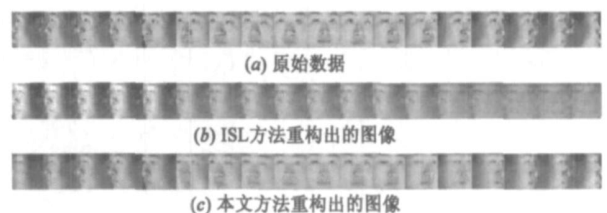


图3 重构姿态图像的对比实验

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增量张量子空间学习方法能有效的刻画图像结构和邻域信息.

实验二 如图 4 所示,显示了普遍使用的视频序列“dudek”序列的跟踪结果.“dudek”视频序列分辨率为 480×720 ,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.序列中跟踪动目标为人脸,图 4 中显示了“dudek”序列中第 104, 107, 183, 185, 461, 492 帧,从图像上可以明显看到三处短时(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,表情变化及快速运动等的情况下,均可有效而鲁棒的跟踪到目标人脸.



图4 视频序列“dudek”跟踪结果

实验三 如图 5 和图 6 所示,人工采集了两段视频序列:“occlusion”和“walker_in_dark”,分辨率均为 384×512 ,采样粒子数为 400.

视频“occlusion”的主要特点是,视频中目标人脸运动速度缓慢,目标任务周围存在其他人脸快速运动以及遮挡干扰,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.

视频“walker_in_dark”的主要特点是,目标人脸存在于光线较暗的环境中,在行进过程中由于光照的问题,会出现面部亮度不均匀的现象,此外由于摄像头的抖动以及目标人脸的非匀速非刚体运动也会为跟踪任务造成困难,而采用本文的方法由于保持了观测图像的内在结构信息,而对光照不均匀这一问题不敏感,从而,稳定地跟踪到目标人脸.

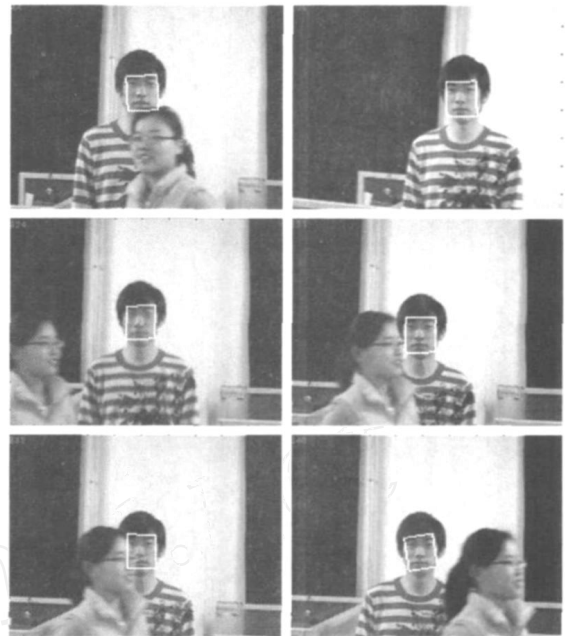


图5 视频序列“occlusion”跟踪结果

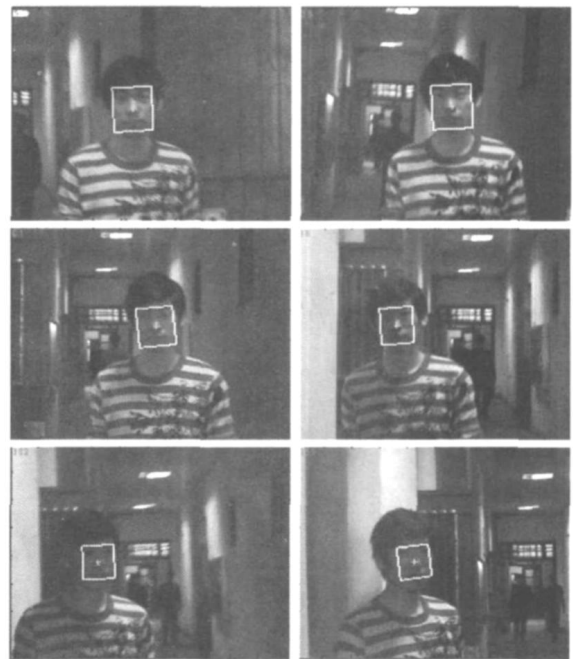
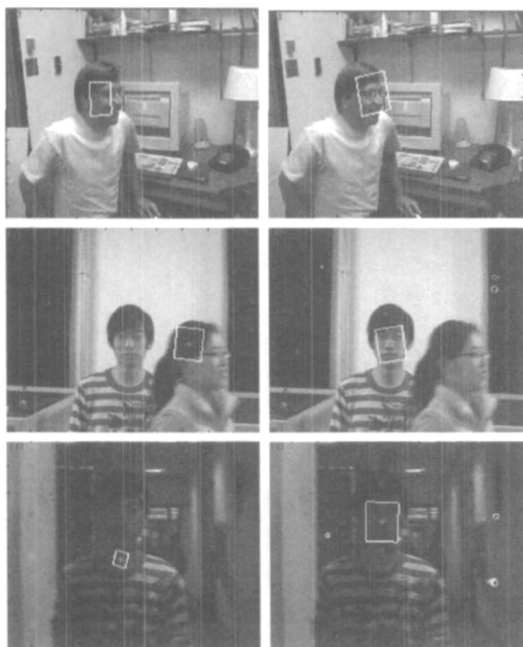


图6 视频序列“walker_in_dark”的跟踪结果

实验四 恢复了文献[7]中的方法.在文献[7]中,跟踪方法采用传统基于向量的方法,即增量空间学习(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[7]和本文方法跟踪效果的视觉对比,如图 7 所示,各列分别为序列“dudek”、“occlusion”以及“walker_in_dark”的对比跟踪结果.

第一行(a)为文献[7]中的方法的跟踪结果.在“dudek”和“occlusion”跟踪过程中,多处出现人脸目标的短视遮挡,非缓变的环境变化以及快速运动;在“walk-



(a) 文献[7]的跟踪结果 (b) 本文的方法
图7 与文献[7]对比视频序列“dudek”, “occlusion”以及“walker_in_dark”的跟踪结果

er. in. dark”的跟踪过程中,所构造的人脸子空间由于失去了目标外观的结构信息,尤其对于光照不均匀的变化会引起相继两帧图像向量较大的差异,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暗淡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.

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.此外,由于在采用贝叶斯推理时,计算参数采样以及相似性的计算量和文献[7]相仿,本文算法和文献[7]的计算量对比主要集中在子空间更新阶段.本文所提出的跟踪算法每更新一次子空间时的计算量可由 $\sum_{i=1}^2 R_i N_i^2 + \sum_{i=1}^2 N_i^2 + \sum_{j=1}^2 N_j^2 + \sum_{i=1}^2 R_i N_i^2$ 表示,且 $N_1 = N_2$, 其中, R_i 和 R_i 分别为张量子空间更新前后投影矩阵 U_i 所保留的列数, $N_i, i = 1, 2$ 为观测数据的行数和列数.考虑实际跟踪算法中 R_i 和 R_i 的取值很小,可以忽略不计,则本文提出的跟踪算法的计算量为 $O(N_i^3)$.

文献[7]中算法每更新一次子空间时的计算量(参见文献[5])为

$$4Dm(K+m) + O(K+m)^3 + 2DK(K+m) \quad (9)$$

其中 $D = N_1 N_2$ 为数据向量的维数, K 为特征向量矩阵

的最大列数, m 为新增加数据的个数.整理式(9),得:

$$D(4m^2 + 6mK + 2K^2) + O(K+m)^3 \quad (10)$$

注意到 $D = N_1 N_2$, 且实际的算法中, $(4m^2 + 6mK + 2K^2)$ 这一项往往比 N_1 或 N_2 大,因而,式(10)中第一项 $D(4m^2 + 6mK + 2K^2) > O(N_i^3)$, 则文献[7]的计算量要大于本文的计算量.综上所述,当保留适当合理的特征向量个数时,本文算法能够获得较文献[7]方法更少的计算量.

5 结论

本文提出了一种有效的基于外观增量学习的目标跟踪方法.首先,提出了一种基于张量子空间的增量学习算法,其次,构造了一个基于贝叶斯推理的跟踪框架,通过对目标仿射运动参数进行采样,获得目标下一时刻可能的位置及参数,利用后验概率估计的最大值得到预测;最后,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.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地描述目标特点,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] M J Black, A D Jepson. Eigentracking: robust matching and tracking of articulated objects using a view based representation [J]. International Journal of Computer Vision (IJCV), 1998, 26(1): 63 - 84.
- [2] M Turk, A Pentland. Face recognition using eigenfaces [A]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Lahaina, Maui, Hawaii, USA, June 3 - 6, 1991. 586 - 591.
- [3] 常发亮, 马丽, 刘增晓, 乔正义. 复杂环境下基于自适应粒子滤波器的目标跟踪[J]. 电子学报, 2006, 34(12): 2150 - 2153.
CHANG Fa-liang, MA Li, LIU Zeng-xiao, QIAO Yi-zheng. Target tracking based on adaptive particle filter under complex background[J]. Acta Electronica Sinica, 2006, 34(12): 2150 - 2153. (in Chinese)
- [4] D Comaniciu, V Ramesh, P Meer. Real-time tracking of non-rigid objects using mean shift [A]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Hilton Head Island, South Carolina, June 13 - 15, 2000, 2: 142 - 149.
- [5] A Levy, M Lindenbaum. Sequential Karhunen-Loeve basis extraction and its application to images [J]. IEEE Transactions on

- Image Processing, 2000, 9(8) :1371 - 1374.
- [6] J Lim, D Ross, R S Lin, M H Yang. Incremental learning for visual tracking[A]. Proceedings of Conference on Advances in Neural Information Processing Systems (NIPS) [C]. Vancouver, Canada :the MIT Press, December 5 - 8, 2004, 793 - 800.
- [7] D Ross, J Lim, R S Lin, M H Yang. Incremental learning for robust visual tracking[J]. International Journal of Computer vision (IJCV), 2008, 77(1 - 3) :125 - 141.
- [8] R S Lin, D Ross, J Lim, M H Yang. Adaptive discriminative generative model and its applications[A]. Proceedings of Conference on Advances in Neural Information Processing Systems (NIPS) [C]. Vancouver, Canada, the MIT Press, December 13 - 18, 2004, 801 - 808.
- [9] M A O Vasilescu, D Terzopoulos. Multilinear subspace analysis for image ensembles[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Madison, Wisconsin, June 16 - 22, 2003, 2:93 - 99.
- [10] Li X, Hu W M, Zhang Z F et al. Robust visual tracking based on incremental tensor subspace learning [A]. Proceedings of International Conference on Computer Vision (ICCV) [C]. Rio de Janeiro, Brazil, October 14 - 20, 2007, 1 - 8.
- [11] M Isard, A Blake. Condensation: Conditional density propagation for visual tracking[J]. International Journal of Computer vision (IJCV), 1998, 29(1) :5 - 28.
- [12] Sun J M, Tao D C, Faloutsos C. Beyond streams and graphs: dynamic tensor analysis[A]. Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) [C]. Philadelphia, USA, August 20 - 23, 2006, 374 - 383.
- [13] Cai D, He X F, Han J W. UIUCDCS-R-2005 - 2572, Subspace Learning Based on Tensor Analysis [S]. Technique Report, May 2005.

作者简介:



温 静 女, 1982 年出生于山西榆次, 博士研究生。2006 在西安电子科技大学获得硕士学位, 目前主要从事模式识别、计算机视觉以及目标跟踪方向的研究。

E-mail: berbemimi@yahoo.cn



李 洁 女, 1972 年出生于陕西西安, 工学博士, 西安电子科技大学副教授。主要从事人工智能、模式识别以及数据挖掘等方面的研究。