

一种结合 Mean-shift 和粒子滤波的视频跟踪算法

李冬, 陈 愚*, 赵学梅, 杨任尔

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 提出一种结合均值偏移算法和粒子滤波理论的视频跟踪算法, 解决了目标旋转、部分遮挡和运动速度过快的问题. 通过均值偏移对粒子滤波中的粒子集进行进一步寻优, 增加了粒子的有效性, 极大减少了粒子采样的数量, 且解决了经过多次重采样后粒子多样性降低的问题. 新的粒子通过与观测值之间的巴氏系数来决定重要性权重. 实验证明: 本算法可以完成实时地对视频目标进行部分遮挡以及目标旋转下的跟踪, 具有较强的鲁棒性.

关键词: 均值偏移; 粒子滤波; 视频跟踪

中图分类号: TN919.8

文献标识码: A

文章编号: 1001-5132 (2011) 01-0024-06

随着计算机视觉应用范围的不断扩展, 近年来视频跟踪技术得到了广泛而深入的研究. 视频跟踪方法主要可 2 种: (1) 有自顶向下的概率目标跟踪方法, 即将目标跟踪转换为在贝叶斯滤波框架下推理目标状态(如位置、速度)后验概率密度的过程; (2) 自底向上的确定目标跟踪方法, 即通过手动或目标检测获得目标模板, 并建立代价函数(Cost Function)来表达目标候选位置和目标模板的相似程度, 然后再利用最优化方法找到代价函数的最大值.

概率目标跟踪方法中的典型代表有基于贝叶斯迭代的卡尔曼(Kalman)滤波和粒子滤波(Particle filter)算法; 确定目标跟踪方法中的主要代表算法有均值偏移(Mean-shift)算法. Kalman 滤波算法近年已取得较大的发展, 扩展 Kalman 滤波(extended Kalman filter)^[1]、无味 Kalman 滤波(unscented Kalman filter)^[2]对于非线性系统都可以进行有效地估计, 但是这些算法都依然限制于噪声高斯性的假设^[3].

Mean-shift 算法是通过多次迭代计算出当前的偏移向量, 并将该点移动到局部最优的一种方法. 1995 年该算法被应用到模式跟踪当中, 并取得了良好的效果^[4], 在此基础上, Bradski 等人^[5]提出了

CamShift 方法, 并在一些简单的应用场景实现了实时跟踪, 效果较好. 但是该方法在目标运动过快或被遮挡时, 跟踪精度会产生较大的误差, 鲁棒性不强. 粒子滤波方法最早于 1998 年由 Isard 等人^[6]应用到目标跟踪中, 该算法对于系统噪声以及观测噪声为非高斯情况下, 也可取得较好的跟踪, 鲁棒性较强. 之后在此基础上, Nummiaro 等人^[7]提出基于颜色的粒子滤波跟踪模型, 并取得非常好的效果, 但其需要采样大量的粒子, 因此算法的复杂度较高. 当下粒子滤波算法的研究已经成为目标跟踪研究的热点, 但是由于传统的粒子滤波中粒子的有效性不高, 通过蒙特卡洛方法必须引入大量的粒子来进行近似计算^[8], 而且通过多次重采样后将会导致粒子的多样性丧失^[9], 这使得跟踪的有效性与实时性就受到了极大的限制.

笔者提出的一种结合均值偏移和粒子滤波的视频跟踪算法, 在粒子滤波框架下对预测的粒子通过 Mean-shift 算法来进行进一步寻优, 并进行重采样后得到最后有效的粒子集, 这样就可以大大地减少所需粒子的数量, 并进一步降低粒子退化现象. 通过应用到真实视频序列的实验结果显示, 对于快速运动和有部分遮挡以及发生旋转的目标, 新算法都可以进行有效跟踪.

收稿日期: 2010-07-03.

宁波大学学报(理工版)网址: <http://3xb.nbu.edu.cn>

基金项目: 浙江省教育厅科研项目(Y200804700); 宁波市自然科学基金(201001A6001075).

第一作者: 李冬(1985-), 男, 安徽合肥人, 在读硕士研究生, 主要研究方向: 视频目标跟踪质量. E-mail: lidong_cx@sina.com

*通讯作者: 陈愚(1962-), 男, 重庆人, 博士/副教授, 主要研究方向: 图像与视频分析处理. E-mail: chenken@nbu.edu.cn

1 粒子滤波

粒子滤波是运用蒙特卡洛方法的贝叶斯递推算法, 通过动态运动模型来实现递推过程. 在粒子滤波算法中, 假定 t 时刻的状态变量为 X_t , 观测向量为 $Z_t \sim (z_1, z_2, \dots, z_t)$, 即从 $t=1$ 时刻到 t 时刻的历史观测; 又设预测方程和观测方程为 $X_t = f(X_{t-1}, w_t)$, $z_t = h(X_t, v_t)$, 其中 w_t 和 v_t 分别代表系统噪声和观测噪声. 通常粒子滤波中的后验概率 $p(X_t | Z_t)$ 和观测密度函数 $p(z_t | X_t)$ 都无需要求满足高斯分布.

粒子滤波核心思想是通过建立一组带有相应重要性权重的粒子集 $X_t = \{(x_t^{(i)}, w_t^{(i)} | i=1, 2, \dots, N)\}$, 并利用(1)式来近似的计算后验估计概率分布.

$$E_t = \sum_{i=1}^N x_t^{(i)} w_t^{(i)}, \quad (1)$$

其中, $x_t^{(i)}$ 表示 t 时刻的 i 个粒子; $w_t^{(i)}$ 表示 t 时刻 i 个粒子的归一化重要性权重, 满足 $\sum_{i=1}^N w_t^{(i)} = 1$.

对于在序列中进行估计, 序贯重要性采样 (Sequential Importance Sampling, SIS) 是运用粒子滤波的一个重要内容. 在初始时刻, 通过先验信息或者手动选取初始粒子集, 并赋予每个粒子相等的重要性权重 $1/N$; 当到 t 时刻 ($t > 0$) 时, 一般无法从后验概率中直接采样得到所需粒子, 因此, 必须通过 1 个已知的、容易采样的建议分布 $q(x)$ 中采样得到粒子. 通过观测模型 $p(z_t | X_t^{(i)})$ 来计算每个粒子的重要性权重, 算式如下:

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{p(z_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{q(X_t^{(i)} | X_{t-1}^{(i)}, Z_t)}. \quad (2)$$

最后根据(1)式加权求和计算得到所需的估计值, 其中建议分布的选择对于粒子滤波器的设计有着重要影响^[10]. 通常选用先验分布 $p(X_t | X_{t-1})$ 来代替建议分布. SIS 方法有着一个严重的问题就是退化现象, 降低这种现象可以从选取更加合理的重要性函数和进行重采样两方面来考虑. 对于重采样, 现在已经提出了多种方法, 如分层重采样、残差重采样、马尔可夫方法(MCMC)^[11-12]等.

2 Mean-shift 算法

Mean-shift 算法是一种利用数据驱动的无参估计算法, 又称为核密度估计算法, 主要通过均值偏移向量寻找到后验概率局部最优^[13].

首先在初始帧 t 建立目标的特征模型 $F_m(t)$, 并根据先验信息或手动选取搜索窗口尺度 h , 可通过下式提取特征概率:

$$F_m(t) = \sum_{i=1}^N C_n k(\|x_t - s_t^{(i)}\|/h)^2, \quad (3)$$

其中, C_n 为归一化系数, 从而使得 $\sum_{m=1}^n F_m = 1$; n 为总特征数; $k(\|\cdot\|^2)$ 为选用的核函数; x_t 为 t 时刻搜索窗口的中心; $s_t^{(i)}$ 为 t 时刻搜索窗口中 N 个样本中的第 i 个随机样本. 当前时刻 $t+1$ 时提取的特征概率为 $F_m(t+1)$.

采用相似性函数来表示上一时刻目标特征密度函数与当前时刻目标特征密度函数的匹配度. 根据定义^[14]有:

$$\rho(t+1) \equiv \rho[F_m(t), F_m(t+1)] = \sum_{m=1}^n \sqrt{F_m(t) F_m(t+1)}, \quad (4)$$

其中, $\rho(t+1)$ 为 Bhattacharyya 系数, 即巴氏系数. 当 $\rho(t+1)$ 越大时, 表示匹配程度越高; 当前时刻的状态与上一时刻的越相似, 反之亦然.

通过计算相似性函数的梯度来决定均值偏移向量,

$$m_{h, s_{t+1}^{(i)}} = \frac{\sum_{i=1}^N g(\|x_{t+1} - s_{t+1}^{(i)}\|^2) C_{t+1}^{(i)} s_{t+1}^{(i)}}{\sum_{i=1}^N g(\|x_{t+1} - s_{t+1}^{(i)}\|^2) C_{t+1}^{(i)}} - x_{t+1}, \quad (5)$$

其中, 加权系数为:

$$C_{t+1}^{(i)} = \sum_{m=1}^n \sqrt{\frac{F_m(t)}{F_m(t+1)}}. \quad (6)$$

具体的算法如下:

步骤 1: 建立特征概率函数, 并通过(6)式计算偏移向量;

步骤 2: 将偏移向量赋给 x_{t+1} ;

步骤 3: 若偏移向量小于阈值 ε , 停止迭代, 否则转到步骤 1 继续执行.

3 结合 Mean-shift 的粒子滤波方法

传统粒子滤波主要通过先验概率密度函数来采样得到所需粒子, 因此丢失了当前的观测信息, 当模型不准确或者观测噪声突变情况下, 会产生较大误差. 笔者提出的算法中, 建议分布选用的是通过先验分布预测后的粒子, 接着通过 Mean-shift 算法进一步寻优后得到新的粒子. 由于已经充分

考虑了当前时刻的观测信息,使得建议分布更加合理,鲁棒性也更强.此外,由于通过 Mean-shift 偏移,所有的粒子都重新进行了选择,有效地抑制了经过多次重采样后粒子多样性降低的问题.

3.1 状态转移和观测模型的建立

在粒子滤波框架下进行跟踪,必须建立目标的系统状态转移模型和观测模型.笔者设定的目标状态参数为 $\{x_t, y_t, v_{xt}, v_{yt}, L_x, L_y, a_t\}$, 其中, x_t, y_t 表示目标在 x 方向和 y 方向上的位置; v_{xt}, v_{yt} 分别表示目标在 x 方向和 y 方向上的速度; L_x, L_y 表示粒子浮动窗口的宽与高; a_t 表示窗口变换速度.

在粒子滤波中,粒子随机样本具有多样性,因此对于运动状态的模型依赖性不强.在此,系统模型选用一阶自回归模型作为目标的系统模型:

$$\begin{pmatrix} x_{t+1}^{(i)} \\ y_{t+1}^{(i)} \\ v_{x(t+1)}^{(i)} \\ v_{y(t+1)}^{(i)} \\ L_{x(t+1)}^{(i)} \\ L_{y(t+1)}^{(i)} \\ a_{t+1}^{(i)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 & r_1 & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & r_2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & r_3 \\ 0 & 0 & 0 & 0 & 0 & 1 & r_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_5 \end{pmatrix} \begin{pmatrix} x_t^{(i)} \\ y_t^{(i)} \\ v_{x(t)}^{(i)} \\ v_{y(t)}^{(i)} \\ L_{x(t)}^{(i)} \\ L_{y(t)}^{(i)} \\ a_t^{(i)} \end{pmatrix} + \omega_{t+1}^{(i)} \quad (7)$$

其中, $\{x_t^{(i)}, y_t^{(i)}, v_{x(t)}^{(i)}, v_{y(t)}^{(i)}, L_{x(t)}^{(i)}, L_{y(t)}^{(i)}, a_t^{(i)}\}$ 表示 t 时刻第 i 个粒子的状态; $\omega_t^{(i)}$ 表示 t 时刻第 i 个粒子状态转移时的随机系统噪声向量, $r_1 \sim r_5$ 和均为 $0 \sim 1$ 间的随机高斯噪声.

通过自适应浮动窗口 L 的设计使得每个粒子的预测范围增大,从而使得当目标运动快过的情况下,也可以在目标附近产生粒子.

手动选取感兴趣区域并采样粒子集,样本数为 N ,可通过(8)式对每个粒子进行传播.观测模型的主要目的是为利用观测信息来对粒子的预测

值进行验证.由于每个粒子代表的是真实状态的一种可能性,通过观测值对每个粒子的可能性大小进行判定并分配相应的重要性权重.对于观测值而言,最直接的方法为在当前帧中寻找可靠的特征.本次试验所选用特征主要为颜色特征,通过当前时刻的估计值对下一时刻目标的颜色特征进行实时更新,因此可以适应目标旋转下的跟踪.

3.2 重要性权重的选择

使用 Mean-shift 算法时,选用目标颜色特征在 HSV 空间中的 H 分量建立模型,并通过(6)式进行迭代,最后得到最优位置.偏移后的粒子与观测值通过巴氏系数计算重要性权重.

$$w_t^{(i)} = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1-\rho(x_t^{(i)}, z_t)}{2\sigma^2}}, \quad (8)$$

其中, $\rho(x_t^{(i)}, z_t)$ 表示当前新的粒子所决定窗口中的直方图概率与观测值所在窗口直方图概率的巴氏系数; σ^2 表示巴氏系数的方差.

3.3 重采样

当进行重采样后,虽淘汰了关联性较弱的粒子,但同时也降低了粒子的多样性.因此,选用一个判决门限 N_{th} 来决定是否进行重采样是必要的.一般有效粒子数可定义为:

$$N_{eff} = \frac{1}{\sum_{i=1}^N (w_t^{(i)})^2}. \quad (9)$$

当 $N_{eff} < N_{th}$ 时,表示当前的有效粒子数不够,需要进行重采样.为实现在实时状态下对目标进行跟踪,本次实验采用基于阈值的重采样算法^[15].

3.4 实时跟踪算法

跟踪算法流程如图 1 所示,具体步骤如下:

步骤 1: 手动选取视频中感兴趣区域以初始化粒子集 $\{x_t^{(i)}, y_t^{(i)}, v_{x(t)}^{(i)}, v_{y(t)}^{(i)}, H_{x(t)}^{(i)}, H_{y(t)}^{(i)}, a_{h(t)}^{(i)}\}$, 其中 i

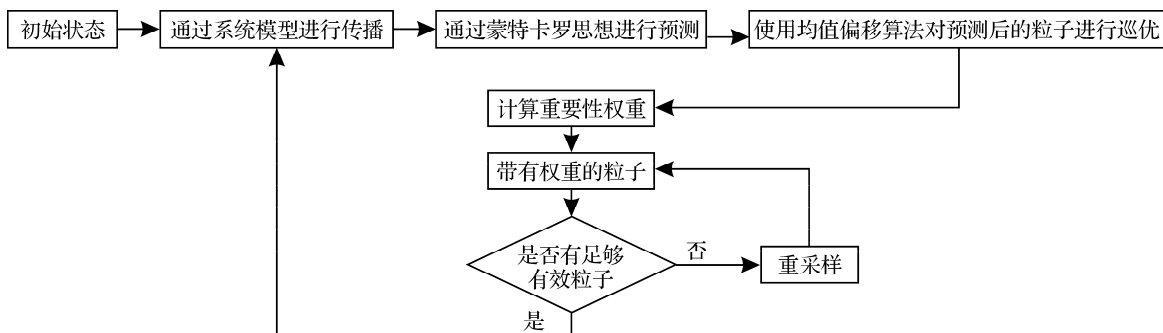


图 1 跟踪算法主要流程图

(1,N), 并在感兴趣区域建立 Mean-shift 算法的目标模型;

步骤 2: 通过(8)式对每个粒子进行预测;

步骤 3: 对预测后的每个粒子进行 Mean-shift 寻优;

步骤 4: 通过观测值与寻优后的粒子间的巴氏系数计算重要性权重 $w_i^{(t)}$, 由(1)式对状态进行估计并建立 Mean-shift 算法目标模型;

步骤 5: 通过(9)式判断粒子有效性, 如果粒子有效性不足则进行重采样, 完毕后执行步骤 4, 否则转到步骤 2 继续执行.

4 实验结果分析与讨论

实验软件环境为 VC 6.0 开发平台, 硬件环境为 Intel Pentium 双核 CPU, 主频 2.8 GHz, 内存 1 GB

的 Dell 个人计算机. 本次实验首先对 Road 视频序列进行了测试研究, 序列大小为 320×240 . 实验结果给出了对于快速运动目标通过一般粒子滤波和结合均值偏移算法粒子滤波跟踪比较的情况, 并进行相关了讨论. 之后, 在实验室运用采样率为 25 f s^{-1} 的摄像头, 实时地对目标进行结果跟踪. 在图 2、图 3 和图 4 中的方框均表示跟踪的窗口. 本实验显示: 运用均值偏移的粒子滤波中采样粒子数大大降低, 算法的复杂度和跟踪效果明显优越于传统的粒子滤波方法; 并在复杂环境下(其中包括目标旋转以及部分遮挡下跟踪的场景)验证了新算法的有效性以及鲁棒性,

图 2 是对目标在快速运动情况下的试验结果. 通过笔者设计的转移模型进行预测, 将粒子转移到目标的真实值附近, 之后再运用均值偏移算法



图 2 结合均值偏移的粒子滤波对快速运动目标的跟踪



(a) 粒子数为50



(b) 粒子数为100

图 3 不同粒子数情况下, 基于颜色的粒子滤波方法对快速运动目标的跟踪情况

表 1 不同跟踪方法的性能比较

跟踪方法	粒子数/个	执行时间/s	跟踪效果
Mean-shift with PF	30	0.032	好
Color based PF	50	0.034	差
Color based PF	100	0.045	相当好



图4 实时的目标的跟踪情况

把粒子转移到局部最优处,提高了粒子的有效性.图3是基于颜色的粒子滤波方法下对目标进行跟踪的结果.由图3可见,粒子数的多少对于目标的跟踪效果有明显的影晌.表1给出了2种方法下的各种参数的比较情况.从图2和图3可见,运用新方法下的跟踪效果明显比传统的粒子滤波要好.在基于颜色的粒子滤波中,当粒子数为50个时,算法所耗费的时间和笔者提出的方法相当,但是不能够进行有效的跟踪;当粒子数为100个时,跟踪效果与笔者的新方法相当,但是进行跟踪所耗费的时间比新方法下的多很多.通过表1显示了新提出的算法的优越性.

图4是在实验室的复杂场景下,通过摄像头对作者本人进行实时跟踪的实验结果.实验中粒子的采样数为30.在粒子滤波框架下,用前一时刻的估计值建立Mean-shift算法的目标模型,因此在目标的颜色发生变换的情况下时,也可以对粒子集进行寻优,从而实现在目标旋转下的有效跟踪.图

4(e)和图4(f)为当目标发生旋转的情况下的跟踪情况,图4(b)和图4(h)分别表示目标被课本和头部分遮挡情况下跟踪的实验结果.

5 结语

主要对传统的粒子滤波存在的采样粒子数多、粒子有效性不够以及粒子经过多次重采样后粒子多样性丧失的问题加以了改进.通过Mean-shift算法,对采样后的粒子进行偏移,有效降低了粒子的数量,并克服了粒子经过多次重采样后粒子多样性不足或者丧失的问题.在通过对视频资料序列和实验室中摄取的能够代表复杂场景的序列进行实时跟踪,表明笔者开发的跟踪算法对于目标运动速度过快、旋转以及部分遮挡情况下,都能够对目标进行有效而稳健的跟踪.

参考文献:

- [1] Dong Zhe, You Zheng. A novel extended Kalman filter for a class of nonlinear systems[J]. Progress in Natural

- Science, 2006, 16(9):912-918.
- [2] Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation[J]. Proceedings of the IEEE, 2004, 92(3):401-422.
- [3] Jwo D J, Cho T S. A practical note on evaluating Kalman filter performance optimality and degradation[J]. Applied Mathematics and Computation, 2007, 193(1):490-502.
- [4] Cheng Y. Mean shift, mode seeking, and clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8):790-799.
- [5] Bradski G R, Clara S. Computer vision face tracking for use in a perceptual user interface[M]. Intel Technology Journal, 1998(2):1-15.
- [6] Isard M, Blake A. Condensation-conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1):5-28.
- [7] Nummiaro K, Koller M E, Gool L V. An adaptive color-based particle filter[J]. Image and Vision Computing, 2003, 21(1):99-110.
- [8] 徐钟济. 蒙特卡罗方法[M]. 上海: 上海科学技术出版社, 1985:35-52.
- [9] Arulampalam M S, Maskell S, Gordon N. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50(2):174-188.
- [10] Doucet A, Godsill S, Andrieu C. On sequential Monte Carlo sampling method for Bayesian filtering[J]. Statistics and Computing, 2000, 10(3):197-208.
- [11] Khan Z, Balch T, Dellaert F. An MCMC-based particle filter for tracking multiple interacting targets[C]// Computer Vision, ECCV, 2004:279-290.
- [12] 何萍, 彭宗举, 邵枫. 基于多视点视频编码宏块模式选择的快速算法[J]. 宁波大学学报: 理工版, 2010, 23(1): 29-33.
- [13] Li Xiangru, Wu Fucao, Hu Zanyi. Convergence of a mean shift algorithm[J]. Journal of Software, 2005, 16(3):365-374.
- [14] Dorin C, Visvanathan R, Peter M. Real-time tracking of non-rigid objects using mean shift[J]. Institute of Electrical and Electronics Engineers Computer Society, 2000, 14:142-149.
- [15] 洪少华, 史治国, 陈抗生. 用于纯方位跟踪的简化粒子滤波算法及其硬件实现[J]. 电子与信息学报, 2009, 31(1):96-97.

A Video Tracking Algorithm with Fusion of Mean-shift and Particle Filtering

LI Dong, CHEN Ken^{*}, ZHAO Xue-mei, YANG Ren-er

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: An algorithm for effective video tracking based on fusion of Mean-shift and particle filtering (PF) is proposed. The research efforts are made to tackle the problems with the target rotation, partial blocking and fast-moving, etc. Using Mean-shift in particle filtering is designed to further optimize the particle set, thus increasing the effectiveness of particle identification, reducing the number of particle sampling, and mitigating the drawbacks found in particle diversity reduction arising from multi-sampling process. Bhattacharyya factor is applied to determine the priority weighting between the new particle and measurements. The algorithm is put to real-time target tracking test on partial blocking and target rotation, manifesting the adequate efficiency and robustness.

Key words: mean-shift; particle filtering; video tracking; Bhattacharyya factor

(责任编辑 章践立)