

文章编号 1004-924X(2008)01-0122-06

# MSMC 跟踪算法在目标跟踪中的应用

孟 勃<sup>1,2</sup>, 朱 明<sup>1</sup>

- (1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;
2. 中国科学院 研究生院, 北京 100039)

**摘要:**针对序贯蒙特卡罗(Sequential Monte Carlo, MC)算法存在的计算量大的缺点,提出了一种新的 Mean Shift Monte Carlo(MSMC)目标跟踪算法。在传统的 MC 算法中采取 Mean Shift 这种梯度最优下降法来寻找局部最大样本值,以目标的颜色特征建立目标状态空间模型,并用 Bhattacharyya 系数作为评价函数给出样本的权值。算法以少于 300 个样本(实验用 200 个样本)来保持对目标运动状态预测的多样性,有效地克服了 MC 算法收敛速度较慢的弱点,将算法的计算时间从 76 ms/frame 降低到了 35 ms/frame(跟踪窗口为 28 pixel×26 pixel)。实验结果表明,提出的算法能够在发生遮挡的情况下实现较稳定的目标跟踪,使算法应用于实际工程成为可能。

**关键词:**目标跟踪;序贯蒙特卡罗算法;Mean Shift MC;局部最优

中图分类号:TP391 文献标识码:A

## Application of MSMC algorithm to visual tracking

MENG Bo<sup>1,2</sup>, ZHU Ming<sup>1</sup>

- (1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*
2. *Graduate School of the Chinese Academy of Sciences, Beijing 100039, China*)

**Abstract:** A new Sequential Monte Carlo(MC) algorithm, Mean Shift Monte Carlo (MSMC) algorithm, is proposed for visual tracking in image sequences. The MSMC takes Mean Shift to converge the samples with smaller weight to look for the local maximum ones. A state-space model of the object is established using the color cue, and the weights of samples are given by adopting the Bhattacharyya coefficients as the evaluation function. The MSMC algorithm can maintain the diversity less than 300 samples (200 samples are used in experiments), so the consumed time can be decreased from 76 ms/frame to 35 ms/frame (the tracking window is 28 pixel×26 pixel). The experimental results in real video data show that the algorithm can track objects stably in the case of obstruction and it is possible to be used in the practical projects.

**Key words:** visual tracking; sequential Monte Carlo (MC) algorithm; Mean Shift Monte Carlo (MSMC); local maximum

## 1 引 言

机动目标的跟踪是计算机视觉中的一个重要研究方向,也是一个很广泛的研究课题。它不仅在军事国防、智能视频监控、运动分析等领域具有广泛的应用前景和潜在的经济价值,在计算机视觉理论研究的层次上也有很高的价值。其中,如何较好地兼顾算法的实时性和稳健性始终是研究的前沿课题。1993年由 Gordon, Salmond 等人<sup>[1]</sup>提出了一种新的基于 Bayes 原理的序贯 Monte-Carlo 方法,1998年,Isard M. 和 Blake A. 等人<sup>[2-5]</sup>首次将算法引入计算机视觉领域进行运动目标的跟踪,该方法的核心是利用一些随机样本本来表示系统随机变量的后验概率密度,能得到基于物理模型的最优数值解,而不是对近似模型进行最优滤波,它作为一种基于贝叶斯估计的非线性滤波算法,在处理非线性时变系统的参数估计和状态滤波问题方面有独到的优势,因此获得了很大的发展。但是,它存在一个致命的弱点,即只有当样本数足够多时,加权样本才能较好地模拟后验分布,因此计算量相当大,这成为它在实时性要求较高的目标跟踪中应用的一大瓶颈。

针对以上问题,已经提出了很多改进算法,如辅助变量粒子滤波算法、局部线性化方法、拒绝采样方法等<sup>[6-10]</sup>,本文提出了一种新的改进算法,MSMC 算法,将 Mean Shift<sup>[11-12]</sup>这种最优梯度下降法结合进 MC 算法中,利用 Mean Shift 快速收敛的特点来寻找局部最优样本值,代替传统 MC 中的重采样过程,克服了重采样过程中对高权值样本重复加权平均所带来的退化问题和样本匮乏问题,从而不仅集成了 MC 的鲁棒性和并行性以及 Mean Shift 迭代算法的快速性的优点,并且能够用较少的样本来保持对目标状态预测的多峰性,有效地克服了 MC 算法收敛速度较慢的弱点,大大减少了算法的计算量,实现了稳定且实时的目标跟踪,并使算法应用于实际工程中成为可能。

本文简要地介绍了 MSMC 算法的基础理论和基本思想;叙述了算法在目标跟踪中的具体应用,并给出仿真试验结果;最后给出了结论。

## 2 算法的理论基础和基本思想

### 2.1 Mean Shift 算法

Mean Shift 算法是一种基于最优梯度下降的方法,它通过迭代的方法来寻找目标。

给定  $n$  个点  $x_i \in \mathbf{R}^2, i=1, \dots, n$ , 核函数  $K(x)$  的窗口半径  $h$ , 在  $x$  处使用核  $K(x)$  的多变量核密度估计为:

$$\hat{f}_{H,K}(x) = \frac{1}{n |H|^{1/2}} \sum_{i=1}^n K\left(\frac{x-x_i}{H^{1/2}}\right). \quad (1)$$

核函数  $K(x)$  有多种不同的形式,常用的核有 Epanechnikov<sup>[9]</sup>核以及正态核,正态核是最优的核,但是它的计算量较大。Epanechnikov 核是估计密度与真实密度间平均全局误差最小化的核,表达式为:

$$K_E(x) = \begin{cases} 2c(1 - \|x\|^2) & \|x\| \leq 1 \\ 0 & \text{其他} \end{cases}, \quad (2)$$

$c$  为单位球体积,对应的轮廓函数为

$$K_E(x) = \begin{cases} 1-x & 0 \leq x \leq 1 \\ 0 & \text{其他} \end{cases}, \quad (3)$$

$$g_E(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{其他} \end{cases}.$$

使用核  $K(x)$  的核密度梯度估计为:

$$\hat{\nabla} f_{H,K}(x) \equiv \nabla \hat{f}_{H,K}(x) = \frac{c_k}{n |H|^{1/2}} \sum_{i=1}^n \nabla k\left(\left\|\frac{x-x_i}{H^{1/2}}\right\|^2\right) = \frac{2c_k}{n |H|^{1/2}} H^{-1} \sum_{i=1}^n (x-x_i) k'\left(\left\|\frac{x-x_i}{H^{1/2}}\right\|^2\right). \quad (4)$$

注意到  $g(x) = -k'(x)$ , 进行代换有:

$$\hat{\nabla} f_{H,G}(x) = \frac{2c_k}{c_g} H^{-1} \hat{f}_{H,g}(x) M_{H,G}(x), \quad (5)$$

即:

$$M_{H,G}(x) = \frac{c_g H \hat{\nabla} f_{H,K}(x)}{2c_k \hat{f}_{H,G}(x)}. \quad (6)$$

结果发现一般意义下的 Mean Shift 向量不一定指向核密度梯度估计的方向,但是不影响算法的收敛性。

### 2.2 序贯 Monte Carlo 算法

算法的基本思想为:用一组带有相关权值的随机本来表示状态矢量,并且基于这些样本和权值可以计算出状态估算值。当样本数非常大

时,这种对于状态的估算值将等同于后验概率密度函数 PDF。

在目标跟踪系统中,常用状态转移方程和观测方程来表示系统的动态方程,令:

$$\begin{cases} x_{k+1} = \Phi x_k + \omega_k \\ y_k = H x_k + v_k \end{cases}, \quad (7)$$

其中,  $\Phi$  为系统的状态转移矩阵,  $H$  为系统的观测矩阵,  $\omega_k, v_k$  分别为系统的过程噪声和观测噪声。在算法中,概率分布被近似为用样本及其权值定义的离散随机量测值。令  $\{x_k^i, \omega_k^i\}_{i=1}^{N_s}$  表示代表后验 PDF  $p(x_k | y_{1:k})$  的随机样本,权值  $\omega_k^i$  经过归一化处理,在时刻  $k$  的后验概率密度可以近似为

$$p(x_k | y_{1:k}) \approx \sum_{i=1}^{N_s} \omega_k^i \delta(x_k - x_k^i). \quad (8)$$

设  $x_{0:k} = \{x_j, j=0, \dots, k\}$  和  $y_{1:k} = \{y_j, j=1, \dots, k\}$  分别表示各个时刻的系统状态和观测状态,  $\{\omega_j^i, i=1, \dots, N\}$  表示  $j$  时刻所对应样本  $\{x_j^i, i=1, \dots, N\}$  的归一化权值,即  $\sum_{i=1}^N \omega_j^i = 1$ 。算法就是用  $\{x_{0:k}^i, \omega_j^i\}_{i=1}^N$  来完全描述后验概率分布  $p(x_{0:k} | y_{1:k})$ 。

### 3 MSMC 算法在目标跟踪中的应用

#### 3.1 目标状态模型描述

本文拟采用目标图像部分区域的颜色特征作为目标的模板,建立出目标模板的颜色分布<sup>[7-8]</sup>,以表示目标的状态空间,建立状态模型<sup>[13-14]</sup>。

$$q_u = C \sum_{x_i \in \mathbf{R}} K(\|x_i\|^2) \delta[b(x_i) - u], \quad (9)$$

假设模板区域的中心位置为  $y_0$ ,  $b(x_i)$  表示在位置  $x$  处的灰度级,固定区域中所有像素的位置可以表示为  $\{x_i\}_{i=1}^M$ ,  $M$  为区域内所有像素的总数。用  $q = \{q_u\}_{u=1, \dots, m}$  来表示目标区域的颜色分布,  $C$  为归一化常数,  $K(x)$  为 Epanechnikov 核密度函数<sup>[15]</sup>。

从状态模型中随机抽取样本,作为样本集,并赋予一定的权值,用  $\{S_i^{(n)}, \omega_i^{(n)}\}_{n=1, \dots, N}$  来表示,所得到的候选目标的颜色密度分布函数即为:

$$p(s) = \{p_u(s)\}_{u=1, \dots, m} = C \sum_{i=1}^n K\left(\left\|\frac{s - x_i}{h}\right\|^2\right) \delta(b(x_i) - u). \quad (10)$$

#### 3.2 目标的动态模型

其次确立目标的运动模型,建立状态转移矩阵,状态转移矩阵的选取十分重要,它决定了各样本在时间序列上的传播方向,直接决定了整个跟踪系统的性能,这里假设样本均满足匀速直线运动。

$$x_i = A x_{i-1} + B w_i, \quad (11)$$

这里,  $x_i = (x, y, \dot{x}, \dot{y})_i^T$ , 为目标的位置及速度参数,  $w_i$  可认为是高斯白噪声,作为模型的过程噪声,而  $A$  为动态模型的系数,决定了动态模型的性能。

#### 3.3 权值计算

根据对目标的观测值来对各样本赋权值,依据算法所选取的似然函数,这里采用 Bhattacharyya 系数<sup>[16]</sup>

$$\rho(y) \equiv \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u}. \quad (12)$$

如果候选目标的中心位置  $y$  距离它的初始值  $\{p_u(y_0)\}_{u=1, \dots, m}$  并不是很大,可以对式(12)进行 Taylor 展开,得到 Bhattacharyya 系数的近似值,以衡量候选目标与目标模板之间的颜色相似度,得到相似系数:

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \frac{C}{2} \sum_{x_i \in \mathbf{R}} \omega_i K(\|d(x_i)\|^2), \quad (13)$$

这里,

$$\omega_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta(b(x_i) - u). \quad (14)$$

由于式(14)的等式右边的值与位置  $y$  无关,所以,式(13)的第二项的最大值决定了 Bhattacharyya 系数的最大值。因此,通过下述的 Mean Shift 过程得到对目标估计的最优值,充分利用 Mean Shift 算法快速收敛于局部最优值的优点,在算法中,对每一个样本都应用 Mean Shift 迭代,使权值小的样本收敛于临近的权值较大的样本处,通过 Mean Shift 的迭代过程得到局部最优的样本位置:

$$y_i = \frac{\sum x_i \omega_i g(d(x_i))}{\sum \omega_i g(d(x_i))}. \quad (15)$$

#### 3.4 状态估计

最后,根据这些最优样本及其权值对目标的状态进行加权平均,得到对目标状态较准确的估

计和预测。

$$E[S_t] = \frac{1}{\sum_{i=1}^N \pi_t^{(i)}} \sum_{i=1}^N \pi_t^{(i)} S_t^{(i)}, \quad (16)$$

其中

$$\pi_t^{(i)} = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{d^{(i)2}}{2\sigma^2}\right), \quad (17)$$

$d^{(i)} = \sqrt{1 - \rho(q, p_{(i)})}$  称为 Bhattacharry 距离。

### 3.5 算法流程

采用序贯 MC 方法来预测目标的实际运动状态。在应用 MC 算法的过程中,采用 Mean Shift 快速迭代算法得到局部最优的样本位置,然后根据局部最优的样本及其权值预测目标的实际状态,并根据得到的观测值不断更新运动模型,直至实现对目标准确、稳定的跟踪。流程如下:

(1) 初始目标的选定:本文采取手动选定初始目标的方法选定初始样本集,并根据式(9)计算目

标模板的颜色分布;

(2) 读取下一帧图像,建立动态模型,利用系统转移方程对样本集进行相应的状态传播,并根据式(10)计算候选样本的颜色分布;

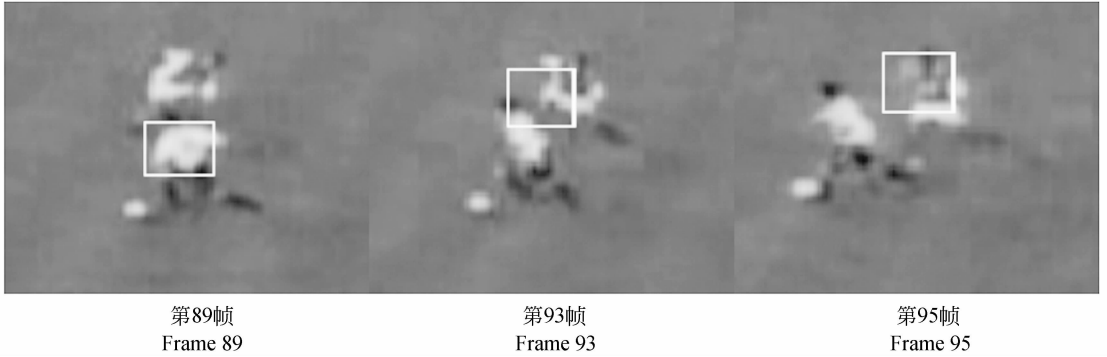
(3) 比较候选样本与模板的颜色分布的相似度,对每个样本赋权值;

(4) 对每个样本进行 Mean Shift 迭代过程,得到局部最优值的位置;

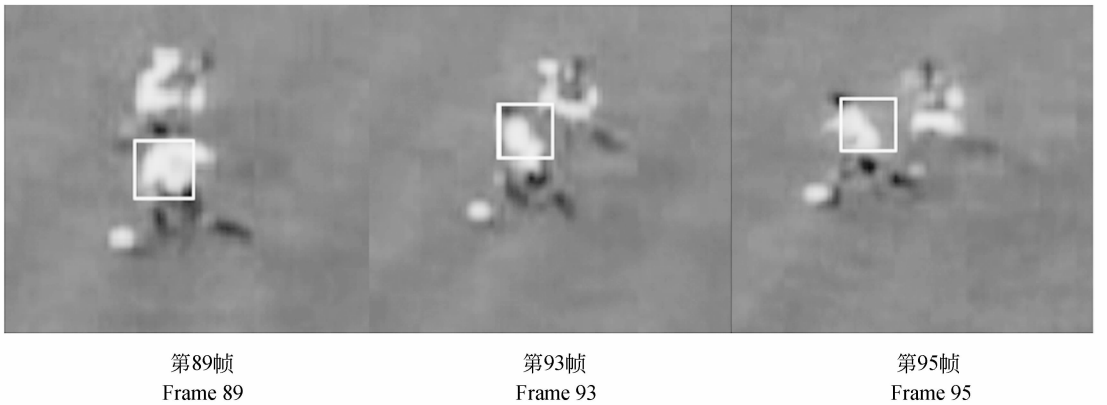
(5) 对局部最优值进行加权平均,得到目标状态的估计值。

## 4 仿真实验

实验是对球场上的运动员进行跟踪,采用模板匹配的方法,当目标发生遮挡时,应用遮挡系数(样本的最大匹配度)来进行判断,并据此得出目



(a) MC 算法跟踪结果图  
(a) Tracking results of MC algorithm



(b) MSMC 算法跟踪结果图  
(b) Tracking results of MSMC algorithm

图 1 目标跟踪结果比较

Fig. 1 Comparison of object tracking

标实际的脱靶量。从下面的实验结果图可以看到：当应用传统 MC 算法，在第 89 帧时，目标运动员受到了其他运动员的干扰，在第 93 帧时跟踪窗口发生抖动现象，到第 95 帧时窗口发生偏离，目标丢失；但是应用 MSMC 算法，在第 93 和 95 帧时，算法仍然能够凭借自身的优越性保证对目标稳定的跟踪。全部实验采用 VC++6.0 语言进行仿真。跟踪结果比较如图 1 所示，窗口大小为  $28 \text{ pixel} \times 26 \text{ pixel}$ ，MC 算法的运算时间为  $76 \text{ ms/frame}$ ，MSMC 算法的运算时间为  $35 \text{ ms/frame}$ 。

图 2 给出了两种算法在每一帧内样本的最大匹配度的比较图，从图中可以看出 MSMC 算法的最大匹配度比 MC 算法的最大匹配度要高很多，这说明，MSMC 算法的跟踪精度要比 MC 算法的

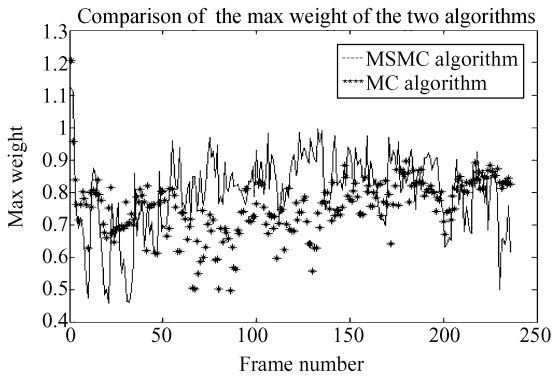


图 2 算法的最大匹配度比较

Fig. 2 Max weight comparison of two algorithms

高。当发生遮挡时(78~110 frame)，MSMC 算法的最大匹配度仍然保持很高，但是 MC 算法的数值则有所下降。

## 5 结 论

本文提出了一种新的改进算法，MSMC 算法，将 Mean Shift 算法这种最优梯度下降法结合到 MC 算法中，利用 Mean Shift 快速收敛的特点来寻找局部最优样本值，代替了传统 MC 中的重采样过程，克服了重采样过程中对高权值样本重复加权平均所带来的退化问题和样本匮乏问题，从而不仅集成了 MC 的鲁棒性和并行性以及 Mean Shift 迭代算法的快速性的优点，且有效地克服了 MC 算法收敛速度较慢的弱点。算法以少于 300 个样本(实验用 200 个样本)来保持对目标运动状态预测的多样性，有效地克服了 MC 算法收敛速度较慢的弱点，将算法的计算时间从  $76 \text{ ms/frame}$  降低到了  $35 \text{ ms/frame}$ (跟踪窗口为  $28 \text{ pixel} \times 26 \text{ pixel}$ )。实验结果表明，本文所提出的算法能够在发生遮挡的情况下实现较稳定的目标跟踪，使算法应用于实际工程成为可能。但是，算法仍然存在许多不足之处，譬如在目标相对速度较大时，算法仍然很难实现实时稳定的跟踪，这将成为以后的工作重点。

## 参考文献：

- [1] GORDON N J, SALMOND D J, SMITH A F M. Novel approach to nonlinear/non-gaussian Bayesian state estimation[C]. *Proc. of the Radar and Signal Processing*, 1993: 107-113.
- [2] ISARD M, BLAKE A. Condensation-conditional density propagation for visual tracking[J]. *Int. J. Computer Vision*, 1998, 29(1): 5-28.
- [3] 朱明, 鲁剑锋, 胡硕. 采用 DSP 的电视测量跟踪器的研制[J]. *光学 精密工程*, 2005, 13(增): 232-235.  
ZHU M, LU J F, HU SH. Development of TV measuring and tracking system by using DSP[J]. *Opt. Precision Eng.*, 2005, 13(Supp.): 232-235. (in Chinese)
- [4] 王洪建, 李志敏. 基于视频图像的车辆流量实时检测系统[J]. *光学 精密工程*, 2005, 13(增): 222-226.  
WANG H J, LI ZH M. Real-time measuring system based on vehicle flow video image[J]. *Opt. Precision Eng.*, 2005, 13(Supp.): 222-226. (in Chinese)
- [5] ISARD M, BLAKE A. Icondensation: unifying low-level and high-level tracking in a stochastic framework[C]. *Springer-Verlag Berlin Herdelberg 1998, Proc. of the Fifth European Conf. on Computer Vision, Berlin*, 1998: 893-908.
- [6] PITT M K, SHEPHARD N. Filtering via simulation: auxiliary particle filter[J]. *J. of the American Statistical Associatin*, 1999, 94(2): 590-599.

- [7] 林玉池,崔彦平,黄银国. 复杂背景下边缘提取与目标识别方法研究[J]. 光学精密工程, 2006,14(3):509-514.  
LIN Y CH, CUI Y P, HUANG Y G. Study on edge detection and target recognition in complex background[J]. *Opt. Precision Eng.*, 2006,14(3):509-514. (in Chinese)
- [8] 陈浩,谭久彬. 一种用于光电目标跟踪的非线性滤波算法[J]. 光学精密工程, 2006,14(5):917-921.  
CHEN H, TAN J B. Nonlinear filtering algorithm for improving opto-electric target tracking[J]. *Opt. Precision Eng.*, 2006,14(5): 917-921. (in Chinese)
- [9] BERZUINI C, BEST N. Dynamic conditional independence models and Markov chain Monte Carlo methods[J]. *J. of the American Statistical Association*, 1997, 92(5): 1403-1412.
- [10] BELVIKAN E, ACKLAM P J. Monte Carlo filters for non-linear state estimation[J]. *Automatica*, 2001,37(1): 177-183.
- [11] COMANICIN D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using Mean Shift[J]. *Computer Vision and Pattern Recognition*, 2002:142-149.
- [12] DORIN C, PETER M. Mean Shift: a robust approach toward feature space analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002,24(5):603-619.
- [13] ROBERT C T. Mean-shift blob tracking through scale space[C]. *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, USA*, 2003, 2: 234-240.
- [14] NUMMIARO K, MEIER E K, VANGOOL L. A color-based Particle Filter[C]. *1<sup>st</sup> Int. Workshop on Generative-Model-Based Vision in Conjunction with ECCV02*, 2002:53-60.
- [15] NUMMIARO K, MELER E K, VANGOOL L. Object tracking with an adaptive color-based particle filter[C]. *Symposium for Patter Recognition of the DAGM*, 2002:353-360.
- [16] AHERNE F, THACKER N, ROCKETT P. The Bhattacharyya metric as an absolute similarity[J]. *Measure for Frequency Coded Data, Kybernetika*, 1997, 32(4):1-7.

作者简介:孟 勃(1980—),女,吉林人,在读博士研究生,主要研究方向为数字图像的采集与处理。E-mail: mengbonan@gmail.com

朱 明(1964—),男,江西南昌人,研究员,主要从事数字图像处理、电视跟踪和自动目标识别技术的研究。