

基于融合算法的视频目标跟踪研究

刘姝黎¹, 史健芳¹, 杨 静²

(1 太原理工大学 信息工程学院, 山西 太原 030024;

2 太原理工大学 信息化管理与建设中心, 山西 太原 030024)

摘 要: 提出了一种嵌入型融合算法, 针对 Mean Shift 算法缺乏必要的模板更新、窗口宽度固定的不足, 利用形态学对目标边缘进行检测, 粒子滤波对目标进行预测, 实现 Mean Shift 算法模板实时更新以及核函数宽度自适应。仿真结果表明, 在目标背景复杂多变、遮挡等情况下, 基于本文算法的视频目标跟踪具有较高的鲁棒性和精确度, 且实时性较好。

关键词: 视频目标跟踪; Mean Shift 算法; 形态学; 粒子滤波; 融合算法

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2017)06—0089—05

Research on Video Target Tracking Based on Fusion Algorithm

LIU Shu-li¹, SHI Jian-fang¹, YANG Jing²

(1 College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China;

2 Center of Information Management and Development, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: This paper proposes a fusion algorithm for video target tracking. Mean shift algorithm lacks necessary template update and can't change the width of the window. Particle filter algorithm has a good prediction of the occlusion region. Morphological algorithm has an advantage in edge detection. The fusion of the three algorithms can complement each other. The scenario can update the template in real time, and be updated to the width of the kernel function. Simulation results show that the proposed scheme can improve the robustness, real-time and accuracy of target tracking, and solve the problem of video target tracking, such as complex background, occlusion, tracking delay and so on.

Key words: video target tracking; mean shift; morphology; particle filtering; fusion algorithm

1 引言

随着网络技术的发展, 智能交通也在发生着巨大的变化, 其开始由点控制、线控制发展到面控制, 优化路口通行方案不再以单点车流量为依据, 而是要综合考虑一个区域甚至整个城市各个路口的车流量。视频目标跟踪技术凭借其升级改造方便, 不破坏路面, 处理速度快, 传输便捷, 且可以和视频监控合二为一的优势, 在智能交通中开始发挥越来越大的作用, 和互联网十结合起来更显得意义重大。目前用于视频目标跟踪的算法适应性较差^[1]。因此, 为了满

足视频目标跟踪的准确性和鲁棒性, 将多种算法进行融合是一个可行的方案, 文献[2-3]提出粒子滤波和 Mean Shift 算法相融合的思想, 这种方案对硬件要求高, 耗时过长, 无法满足实时性要求。因此, 本文提出一种混合嵌入型的融合算法, 即应用 Mean Shift 算法实现目标跟踪, 融合粒子滤波更新模板和形态学进行边缘检测以实现核函数的宽度自适应。

2 目标跟踪

用于视频目标跟踪^[4-5]的算法有很多, 综合比较, 本文选择 Mean Shift 算法^[6-7]、粒子滤波及形态

学三种算法以嵌入型融合的方法进行视频目标跟踪。

2.1 Mean Shift 算法

Mean Shift 算法由 Fukunaga 等人^[8]于 1975 年提出,其字面含义就是偏移的均值向量,基本思想是利用概率密度的梯度爬升来达到局部最优。该算法本质上是一种非参数概率密度估计,是指一个迭代的步骤,即首先算出当前点的偏移均值,然后移动该点到其偏移均值,最后以此为新的起始点继续移动,直到满足给定的条件结束。具体如下:

对于 d 维空间 R^d 的 n 个样本点 $x_i (i = 1, 2, \dots, n)$, x 的 Mean Shift 向量定义为

$$M_h(x) = \frac{1}{k} \sum_{x_i \in S_h} (x_i - x) \quad (1)$$

式中, S_h 是半径等于 h 的高维球形区域,并且满足:

$$S_h = \{y: (y - x)^T (y - x) \leq h^2\} \quad (2)$$

式中, k 表示落在 S_h 内采样点个数, $(x_i - x)$ 为 x_i 的偏移向量,从而 Mean Shift 向量表示对 S_h 内 k 个点的偏移向量求平均。一般而言, S_h 内的点都朝着概率密度梯度方向分布(即 Mean Shift 向量 $M_h(x)$ 的方向),Mean Shift 向量示意图如图 1 所示。

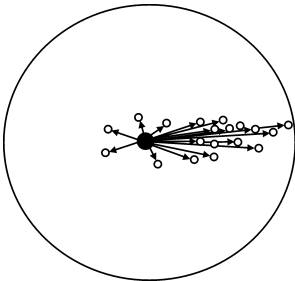


图 1 Mean Shift 向量图

图 1 中,大圆圈表示 S_h 区域,小圆圈表示样本点分布,中心的黑点是 Mean Shift 向量的基准点,箭头表示该样本点对基准点的偏移向量,偏移向量的平均值 Mean Shift 向量 $M_h(x)$ 自然会指向样本点分布最多的地方,即概率密度函数的梯度方向。该算法在目标跟踪中具体应用如下所示。

(1) 目标模型叙述

假设目标区域的中心为 x_0 ,其中有 n 个像素用 $\{x_i\}_{i=1,2,3,\dots,n}$ 来表示, m 为特征值的个数,则特征值 $u = 1, 2, 3, \dots, m$ 的概率密度的估计为

$$\hat{q}_u(x_0) = c \sum_{i=1}^n k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (3)$$

$k(x)$ 是核函数的轮廓函数,由于目标中心附近的像素比其他位置的像素对描述目标来说更可靠,

因此核函数对中心附近的像素赋予大的权值,其他位置的像素赋予小的权值; $b(x_i)$ 是 x_i 像素的量化值,即图像灰度值索引函数; $[b(x_i) - u]$ 的总体作用是判断目标区域中像素值是否属于第 u 个特征值,若属于,则值为 1,否则为 0; h 是核函数宽度,用来决定候选目标区域大小; c 是一个标准化常量系数,其表达式为

$$c = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right)} \quad (4)$$

(2) 候选目标叙述

同理,假设候选目标区域 $\{y_i\}_{i=1,2,3,\dots,n_h}$ 的中心坐标区域为 y ,则候选模型特征值 $u = 1, 2, 3, \dots, m$ 的概率密度的估计为

$$\hat{p}_u(y) = c_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - y_i}{h}\right\|^2\right) \delta[b(y_i) - u] \quad (5)$$

$$\text{式中, } c_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{y - y_i}{h}\right\|^2\right)}.$$

(3) 目标定位

一般使用 Bhattacharyya 系数作为 $\hat{p}$ 和 $\hat{q}$ 之间的相似性度量函数,具体表达式为

$$\hat{\rho}(\hat{p}(y), \hat{q}(x_0)) = \rho[\hat{p}(y), \hat{q}(x_0)] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u(x_0)} \quad (6)$$

$$\text{式中, } p(y) = \{p_u(y)\}_{u=1,\dots,m}, \sum_{u=1}^m p_u = 1;$$

$$q(x_0) = \{q_u(x_0)\}_{u=1,\dots,m}, \sum_{u=1}^m q_u = 1.$$

由此目标位置即为式中 $\hat{p}(y)$ (其值在 0 ~ 1 之间) 局部极大值对应的位置。为使 $\hat{p}(y)$ 最大,将前一帧中目标中心位置 y_0 作为当前帧的目标中心,即从这一点开始寻找最优匹配目标,其中心设为 y 。将上式在 $\hat{p}_u(y_0)$ 处进行泰勒级数展开,舍去高阶项综合可得:

$$\rho[\hat{p}(y), \hat{q}(x_0)] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u(x_0)} + \frac{c_h}{2} \sum_{i=1}^{n_h} \quad (7)$$

式中, $w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u(x_0)}{\hat{p}_u(y_0)}} \delta[b(x_i) - u]$ 。分析可知,要使式(7) 左边值最大,只需最大化其右边第二项即可,而第二项表示的是当前帧位置 y 利用 w_i 加权核函数 k 估算的概率密度。此密度的极值可以通过计

算 Mean Shift 向量得到目标的新位置,从而得到候选区域中 y_0 移向真实目标区域 y 的向量,具体表达式为

$$m(y) = y_1 - y_0 = \frac{\sum_{i=1}^{n_h} x_i w_i g(\| \frac{y_0 - x_i}{h} \|^2)}{\sum_{i=1}^{n_h} w_i g(\| \frac{y_0 - x_i}{h} \|^2)} - y_0 \quad (8)$$

式中, $g(x) = -k'(x)$.

(4) 整个算法流程

Step1 在当前帧以 y_0 为起点,计算候选目标的特征 $\{p_u(y)\}_{u=1 \dots m}$,并估算巴氏系数:

$$\rho[\hat{p}(y_0), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} \quad (9)$$

step2 计算权值 $\{w_i\}_{i=1, \dots, m}$;

step3 利用 Mean Shift 算法计算目标的新位置:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g(\| \frac{y_0 - x_i}{h} \|^2)}{\sum_{i=1}^{n_h} w_i g(\| \frac{y_0 - x_i}{h} \|^2)} \quad (10)$$

step4 更新候选目标核函数的直方图为 $\{p_u(y_1)\}_{u=1, 2, 3, \dots, m}$,并再次估算巴氏系数:

$$\rho[\hat{p}(y_1), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y_1) \hat{q}_u} \quad (11)$$

step5 如果 $|y_1 - y_0| < \epsilon$,则跟踪结束,否则 $y_0 = y_1$,转到 step2.

结束条件 ϵ 应小于 1 个像素,且同时设定最大的迭代次数(一般设为 20).

2.2 形态学

由于噪声的干扰,直接处理图像通常不能得到完整的运动对象边缘,这时需要进一步去除噪声点,而形态学可以用来进行图像的边界检测.在形态学检测中常用的有两种基本运算,如下所示.

(1) 膨胀:假设 f 和 g 分别为目标图像和一个结构元素,则 A 被 B 膨胀的运算定义为

$$f \oplus g = \{z \mid (\hat{g})_z \cap f \neq \emptyset\}.$$

(2) 腐蚀: A 被 B 腐蚀的定义为

$$f \ominus g = \{z \mid (\hat{g})_z \cap f \subseteq f\}.$$

腐蚀是使边界向内部收缩,膨胀是使边界向外部扩张,通过对腐蚀和膨胀的级联使用,又延伸出了以下两种运算:

(1) 开运算: $f \circ g = f \ominus g \oplus g$.

(2) 闭运算: $f \bullet g = f \oplus g \ominus g$.

这两种运算均使图像含有较少细节,轮廓变平滑,边缘更加规则,区别是开运算去掉了细长的突出,而闭运算填充了狭窄的间断.本文实验采用先开后闭运算,去除噪声,消除孤立点.

2.3 粒子滤波

粒子滤波是指通过寻找一组在状态空间中传播的随机样本来近似表示概率密度函数,用样本均值代替积分运算,进而获得系统状态的最小方差估计的过程.其本质是通过非参数化蒙特卡罗模拟方法实现的递推贝叶斯滤波,即采用一组从概率密度随机抽取并附带一定权值的粒子集 $\{x_t^i, w_t^i\}_{i=1}^N$ 逼近后验概率.粒子滤波的具体跟踪流程如下.

Step1 粒子集初始化, $k=0$:随机生成 N 个具有相同概率的粒子,使其位于目标可能的区域,对于 $i=1, 2, \dots, N$, 由先验 $p(x_0)$ 生成采样粒子 $\{x_0^i\}_{i=1}^N$;

Step2 重要性采样:

① 从 $q(x_k^{(i)} \mid x_{k-1}^{(i)}, z_k)$ 采集 N 个新粒子 $\{x_k^i\}_{i=1}^N$,此时粒子集合为 $\{x_{0:k}^i\} = \{x_k^{(i)}, x_{0:k-1}^{(i)}\}_{i=1}^N$;

② 计算粒子权重 $\tilde{w}_k^{(i)}$:

$$\tilde{w}_k^{(i)} = w_{k-1}^{(i)} \frac{p(z_k \mid x_k^{(i)}) p(x_k^{(i)} \mid x_{k-1}^{(i)})}{q(x_k^{(i)} \mid x_{k-1}^{(i)}, z_k)}, i=1 \dots \quad (12)$$

③ 权重归一化: $w_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=1}^N \tilde{w}_k^{(j)}}$;

Step3 重采样:对粒子集 $\{\tilde{x}_k^{(i)}, \tilde{w}_k^{(i)}\}$ 进行重采样,剔除权重较小的粒子,复制权重较大的粒子,得到新的粒子集 $\{x_k^{(j)}, \frac{1}{N}\}_{j=1}^N$

Step4 输出阶:计算当前 k 时刻的状态估计值:

$$\tilde{x}_k = \sum_{i=1}^N w_k^{(i)} x_k^{(i)} \quad (13)$$

Step5 令 $k = k+1$,跟踪继续,转到 Step2.

2.4 融合算法

与粒子滤波相结合,将前一帧的目标位置作为后一帧的迭代起点,先进行预测再进行 Mean Shift 迭代,同时又与形态学算法相结合,适时调整窗口宽度,使多种算法取长补短,达到理想的跟踪效果.算法具体过程说明如下.

Step1 初始化滤波器,输入视频;

Step2 对当前帧选定的区域进行背景差分,平滑滤波,以及二值化(阈值一般 20),然后进行边缘检测,

即在当前帧通过形态学检测出目标的边缘的信息后,统计边缘线中最小的坐标 $x_{\min}$ 、 $y_{\min}$ 和最大的坐标 $x_{\max}$ 、 $y_{\max}$,得出轮廓限定区域宽度和高度分别为

$$w' = x_{\max} - x_{\min}, h' = y_{\max} - y_{\min} \tag{14}$$

从而下一帧跟踪窗口的宽度和高度分别为

$$w = w'(1 + \alpha) \quad h = h'(1 + \beta) \tag{15}$$

式中, $\alpha = k\omega'/\omega, \beta = kh'/h, k$ 为常量,经验值取 5 ~ 10(本实验取 5). 同时得的目标初始中心位置 $x_0(t) = (x_{\min} + w'/2, y_{\min} + h'/2)$,在每个跟踪过程中此步只执行一次.

Step3 根据前一时刻状态值 x_0 ,由粒子滤波获得当前时刻目标的预测位置,即当前 k 时刻的状态估计值:

$$\tilde{x}_k = \sum_{i=1}^N w_k^{(i)} x_k^{(i)} \tag{16}$$

Step4 以预测位置为起点,以 step2 中得到的 w 和 h 为窗口尺寸,由 Mean Shift 算法获得当前时刻的观测位置,从而进一步获得当前时刻的目标估计位置;

Step5 将该估计位置迭代应用到 step2 中,得到下一时刻的预测位置和观测位置以及估计位置;

Step6 反复使用 step2 ~ 5,直至跟踪目标移出追踪范围,跟踪结束.

融合算法流程图如图 2 所示.

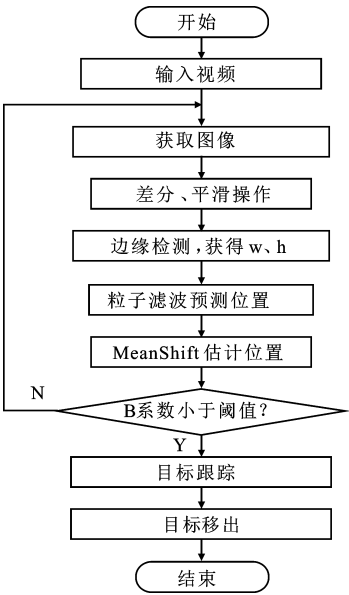


图 2 嵌入型融合算法流程图

3 仿真结果及分析

在 MATLAB2013b 软件环境下,按照图 2 所示流程图进行编程,利用嵌入型融合算法程序对获得

的两段视频素材进行处理,目标跟踪结果如下.

采用嵌入型融合算法对视频 1(选自 MATLAB 自带视频库)中的车辆进行了跟踪和车流量统计,图 3(a)、(b)、(c)分别为视频 1 在不同帧下的跟踪效果. 三幅图中该算法程序分别准确跟踪了一辆、两辆和三辆汽车,并给出了车流量统计结果. 对比实际情况可以看出,该算法对视频目标的跟踪具有较高的准确度.

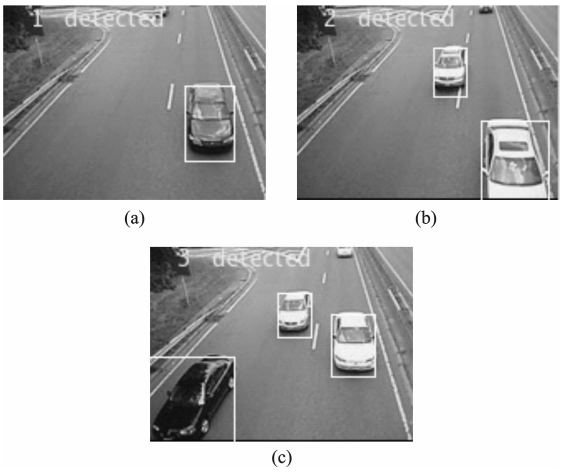


图 3 视频 1 跟踪结果

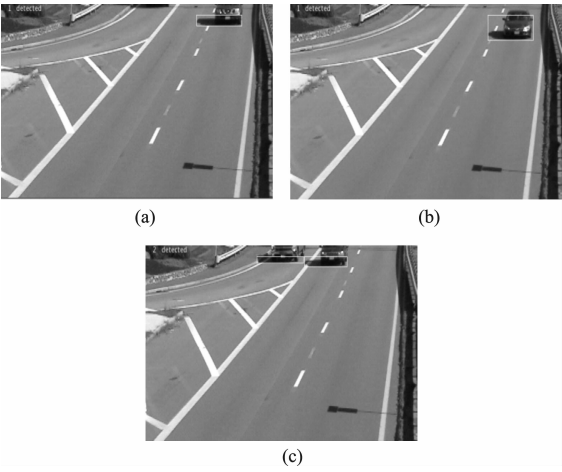


图 4 视频 2 跟踪结果

同理对视频 2(选自 MATLAB 自带视频库)进行跟踪框自适应分析,图 4(a)、(b)、(c)分别为视频 2 在不同帧下的跟踪效果,三幅图中该融合算法程序能够根据迎面驶来车辆目标的大小自适应的调整跟踪框的大小,这样的准确跟踪避免了车辆的重复统计,也对遮挡起到了抑制作用,提高了算法的鲁棒性.

实验多个实际拍摄视频段,从平均实时性、精确度和鲁棒性三方面对单一算法和其他融合算法的性

能进行了比较,其中实时性的计算方法为车辆进入到离开视频的时间除以算法处理时间得出的百分比,精确度为实际车辆数减去统计偏差数然后再除以实际车辆数得出的百分比,鲁棒性为准确跟踪的次数除以总次数得出的百分比.具体的实验数据统计结果如表 1 所示(表中串行融合指先分类,再 Mean Shift 和粒子滤波二选一,并行融合指两种算法同时进行,然后择优选取).

	表 1 不同算法性能对比					%
	Mean Shift 算法	粒子 滤波	串行 融合	并行 融合	嵌入型 融合算法	
平均实时性	92	88	85	86	89	
平均精确度	82	85	86	85	88	
平均鲁棒性	81	86	87	86	90	

由表 1 可知,在实时性方面,本文嵌入型融合算法不如 Mean Shift 算法效果好,但比粒子滤波、串行和并行融合效果好;在准确度和鲁棒性方面,嵌入型融合算法的处理效果要明显优于其他算法.由综合各评价指标的分析结果可知,本文提出的嵌入型融合方法在实时性、准确度和鲁棒性方面都能够满足视频目标跟踪要求.

4 结束语

本文提出了一种混合嵌入型融合算法用于视频目标跟踪,采用该算法对复杂度不同的两个视频段进行了跟踪分析,研究结果表明:

(1)将 Mean Shift、形态学以及粒子滤波三种算法相互融合用于视频目标跟踪提高了目标跟踪的精确度和鲁棒性以及处理速度,达到了车流量检测的准确性、抗干扰性和实时性要求;

(2)该方案融合了 Mean Shift 算法快速处理和粒子滤波算法可以预测模板的优点,有效解决了视频目标跟踪中遮挡、噪声干扰等问题,避免了串行融合精度不够以及并行融合实时性差的缺点;

(3)该方案结合形态学算法,先对目标进行边缘检测,确定 Mean Shift 跟踪窗口的宽度,使其够自适应跟踪目标,克服了 Mean Shift 算法跟踪窗口宽

度固定的缺点,提高了目标跟踪的自适应性和识别准确度,为智能交通中准确测定车流量奠定了基础;

(4)通过多次实验发现对于车流量大,遮挡严重的视频偶尔会有跟丢现象发生,如何进一步提高精确度,缩短处理时间还有待进一步研究.

参考文献:

[1] 亢朋朋,朱锡芳,南东亮,等.一种改进 Mean shift 算法的视频监控系统[J].微电子学与计算机,2013,30(2):153-156.

[2] 邢浪漫,范蟠果.基于 mean shift 和粒子滤波的混合目标跟踪算法[J].微型机与应用,2011,30(20):47-49.

[3] 徐眯,徐成鸿,舒军.基于 Mean shift 和粒子滤波的目标跟踪方法[J].湖北第二师范学院学报,2016,33(2):22-26.

[4] 张楠,刘承志,范存波,等.基于粒子滤波与局部搜索的视频目标跟踪[J].微电子学与计算机,2013,30(9):1-4.

[5] 张继,王洪元.基于局部粒子密度的粒子滤波跟踪方法[J].微电子学与计算机,2012,29(3):93-97.

[6] Chen X, Zhang M, Kai R, et al. Improved mean shift target tracking based on self-organizing maps[J]. Signal Image & Video Processing, 2014, 8(Supp. 1):103-112.

[7] Deng Y, Liu F, Chen J, et al. Mean shift tracker with chaotic artificial bee colony and space variant resolution [J]. Optik - International Journal for Light and Electron Optics, 2014, 125(16):4572-4577.

[8] 赵晓敏.改进 Mean Shift 跟踪算法的研究与实现[D].沈阳:东北大学,2010.

作者简介:

刘姝黎 男,(1991-),硕士研究生.研究方向为智能仪器、目标跟踪. E-mail:liushuli99@163.com.

史健芳 女,(1966-),博士,教授.研究方向为智能仪器及检测技术、智能信息处理、目标跟踪.

杨 静 男,(1970-),博士,副教授.研究方向为信息融合.