

一种用于运动目标检测的快速收敛混合高斯模型

焦波 李国辉 涂丹 汪彦明

(国防科技大学信息系统与管理学院系统工程系,长沙 410073)

摘要 背景模型是交通监控视频中检测运动目标的一种常用方法。混合高斯模型在训练背景模型的过程中效果良好,但其收敛速度较慢。目前各种改进模型,只是提高其初始化的收敛速度;为了加快检测过程中背景改变时的收敛速度,必须实时检测背景是否发生改变,若改变,则需要对模型重新进行初始化。基于以上情况,提出了一种改进的混合高斯模型,该模型不需要重新初始化,避免了实时检测背景是否发生改变的多余步骤,实验结果明显提高了检测过程中的收敛速度。

关键词 快速收敛 混合高斯模型 背景模型 目标检测

中图分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)11-2139-05

A Fast Convergent Gaussian Mixture Model for Moving Object Detection

JIAO Bo, LI Guo-hui, TU Dan, WANG Yan-ming

(Department of System Engineering, School of Info System and Management, National University of Defense Technology, Changsha 410073)

Abstract Background model is a common method for detecting moving object in traffic surveillance video. The effect of Gaussian Mixture Model used in training background model is good, but its convergence velocity is low. At present, many improved models only accelerate the initial convergent velocity. For accelerating the convergent velocity when background changes in the process of surveillance, the models need to detect whether background has changed or not real time and then to be initialized again if background changes. In this paper we put forward a new improved Gaussian Mixture Model, which needn't be initialized again if background information changes and avoids redundant steps of detecting whether background has changed or not real time. Experiment result of the new model shows the convergent velocity in the process of surveillance is improved evidently.

Keywords Fast convergence, gaussian mixture model, background model, object detection

1 引言

交通视频监控系统中通常采用静止摄像机获取交通视频。检测运动目标是分析交通视频过程中一个基础而又关键的任务。在没有运动目标的前提下,视频中像素点可以通过有规律的统计模型来描述,这种统计模型称为像素点的背景模型。运动目标可以通过其所包含像素点是否属于背景模型来检测,这种检测方法通常称为“减背景”

方法。

文献[1]采用单个高斯模型作为背景模型,该模型虽然简单,但其分布比较复杂且需要精细的参数设定。混合高斯模型是由文献[2]提出的,并且由文献[3]、[4]发展为目前常用的表现形式。混合高斯模型由多个高斯分布组成,可以在包含运动目标的视频中自适应的提取背景模型,对背景模型的描述更加准确,但其收敛速度较慢。

目前,国内外提出了各种混合高斯模型参数估

基金项目:国家自然科学基金项目(60273066)

收稿日期:2007-01-10;改回日期:2007-05-31

第一作者简介:焦波(1981~),男。国防科技大学系统工程专业博士研究生。主要研究方向为多媒体信息系统、多媒体数据挖掘和计算机视觉。E-mail: jiaobonudt116@163.com

计的改进方法。文献[5]减少了每帧图片的处理时间,并提高了“减背景”效果,但其对收敛速度没有明显改进;文献[6]在说话人辨认领域,较大幅度提高了收敛速度,但其需要用形态学方法估计混合高斯分量的个数,复杂性较高,不能满足交通视频实时检测的需求。

文献[7]在文献[8]的基础上设计了前 $1/\alpha$ 帧的初始化算法,大大提高了混合高斯模型的初始化速度,但其对检测过程中背景改变的情况效果较差,主要表现为以下两点:(1)该算法在形式上与混合高斯模型不一致,因此,必须给出背景改变的判断规则,并且每帧需要进行判断,当判断出背景改变,再调用该算法进行初始化;(2)当视频中少数几帧出现异常,可能会导致调用该算法进行初始化,从而丢失前几帧属于背景模型的样本点,并且如果初始化次数过多,则会导致算法不稳定。

2 混合高斯模型 (M_1)

Grimson 和 Stauffer 在文献[3]、[4]中给出了混合高斯模型 M_1 的原理,该模型用 $K(3-5)$ 个高斯分布统计每帧中相同像素点在最近 t 帧的值 $\{X_1, X_2, \dots, X_t\}$,从而得到以下该像素点值在当前帧的概率密度函数:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \quad (2)$$

式中, $\omega_{i,t}$ 表示 t 帧的值 $\{X_1, X_2, \dots, X_t\}$ 中落入第 i 个高斯分布的比率, $\mu_{i,t}$ 和 $\Sigma_{i,t}$ 分别表示高斯分布的期望值和方差, $\Sigma_{i,t} = \sigma_{i,t}^2 \cdot \mathbf{I}$ (假设像素点的 RGB 值相互独立,且有相同的方差, \mathbf{I} 为单位矩阵)。为了表达方便,假设本文中的 $\omega_{i,t}/\sigma_{i,t}$ ($i = 1, 2, \dots, K$) 都是按从大到小的顺序排列, $\eta_{i,t}$ 为与 $\omega_{i,t}$ 对应的高斯分布。该模型中有两个重要参数 α 和 H , 其中 $1/\alpha$ 表示训练 $P(X_t)$ 的最大样本空间中的样本点个数,即 $\alpha = 1/t$; H 用来确定作为背景模型的 B 分布 (指 $n_{i,t}$ 中的前 B 个高斯分布 $n_{1,t}, n_{2,t}, \dots, n_{B,t}$ 为服从背景模型的高斯分布):

$$B = \operatorname{argmin}_b \left(\sum_{i=1}^b \omega_{i,t} > H \right) \quad (3)$$

对该像素点的每个新样本点 X_{t+1} , 判断其与已存在的 K 个高斯分布中的哪一个匹配。如果 X_{t+1} 在某高斯分布的 N (N 的大小由视频噪声决定, 试验时选取 $N=3$) 个标准差之内, 则规定 X_{t+1} 与该高斯分布匹配。如果 X_{t+1} 与前 s 个高斯分布都不匹配, 且与第 $s+1$ 个高斯分布匹配, 则令 $M_{s+1,t+1} = 1$ 且 $M_{i,t+1} = 0$ ($1 \leq i \leq K$ 且 $i \neq s+1$), 第 $s+1$ 个高斯分布按下式更新其权重、期望值和方差; 其他高斯分布只更新权重, 期望值和方差不变:

$$\omega_{i,t+1} = (1 - \alpha) \cdot \omega_{i,t} + \alpha \cdot M_{i,t+1} \quad (4)$$

(如果第 1 个匹配 $M_{1,t+1} = 1$, 否则 $M_{i,t+1} = 0$)

$$\mu_{i,t+1} = (1 - \rho) \cdot \mu_{i,t} + \rho \cdot X_{t+1} \quad (5)$$

$$\sigma_{i,t+1}^2 = (1 - \rho) \cdot \sigma_{i,t}^2 + \rho \times (X_{t+1} - \mu_{i,t+1})^T (X_{t+1} - \mu_{i,t+1}) \quad (6)$$

$$\rho = \alpha / \omega_{i,t+1} \quad (7)$$

如果 K 个高斯分布都不匹配, 则 ω/σ 最小高斯分布 $\eta_{K,t+1}$ 的期望值用 X_{t+1} 代替, 方差设为较大的值, 权重设为较小的值。

3 快速收敛的混合高斯模型 (M_2)

在 M_1 模型中, 假设 $\eta_{1,0}$ 不符合背景模型, 则有大量样本点与 $\eta_{1,0}$ 不匹配; 不妨设前 n 个样本点都与 $\eta_{1,0}$ 不匹配, 则 $\omega_{1,n} = (1 - \alpha)^n \omega_{1,0}$; 通常 α 的取值较小, 故 $\omega_{1,n}$ 的减小速度较慢; 为使 $\omega_{1,n}$ 低于符合背景模型的高斯分布的权重, 将需要大量的样本点, 这样导致 M_1 模型收敛速度较慢。

为了加快收敛速度, 需要加快符合背景模型高斯分布权重的增加速度和不符合背景模型高斯分布权重的减少速度。在 M_1 模型的基础上, 设计快速收敛的混合高斯 M_2 模型, 该模型只改变式(3)、式(4), 其他不变。

由于 $\sum_{i=1}^K \omega_{i,t} \leq 1$, M_2 模型设立一个虚拟的高斯分布

$\eta_{K+1,t}$, 令其权重 $\omega_{K+1,t} = 1 - \sum_{i=1}^K \omega_{i,t}$ 。当 X_{t+1} 与 $\eta_{i,t}$ ($i = 1, 2, \dots, K$) 都不匹配时, 认为 X_{t+1} 与 $\eta_{K+1,t}$ 匹配; 此时 ω/σ 最小的 $\eta_{K,t}$ 被更新, 使得 $\eta_{K,t+1}$ 中只含有一个样本点 X_{t+1} , 因此, X_{t+1} 同时与 $\eta_{K,t}$ 匹配。

这样 M_2 模型由 $K+1$ 个高斯分布组成, $\omega_{i,t}/\alpha$ ($i = 1, 2, \dots, K+1$) 表示属于各高斯分布的样本点个数。

当 X_{t+1} 与 $\eta_{K,t}$ 匹配时, $\eta_{K,t}$ 中只含有 X_{t+1} , 即 $\omega_{K,t+1}/\alpha = 1, \mu_{K,t+1} = X_{t+1}, \sigma_{K,t+1}^2$ 设为较大的值, 故权重变更为 $\omega_{K,t+1} = \alpha, \omega_{i,t+1} = \omega_{i,t} (1 \leq i \leq K-1)$, $\omega_{K+1,t+1} = 1 - \sum_{i=1}^K \omega_{i,t+1}$; 当 X_{t+1} 与 $\eta_{m,t} (1 \leq m \leq K)$ 中的第 1 个分布匹配时, 则在 $\eta_{m,t}$ 中添加 X_{t+1} 得到 $\eta_{m,t+1}$, 即 $\omega_{m,t+1}/\alpha = \omega_{m,t}/\alpha + 1, \mu_{m,t+1}$ 和 $\sigma_{m,t+1}^2$ 由式 (5) ~ 式 (7) 获得, 故权重变更为 $\omega_{m,t+1} = \omega_{m,t} + \alpha$, 为了

保证 $\sum_{i=1}^{K+1} \omega_{i,t+1} = 1$, 需要在 $\eta_{i,t} (1 \leq i \leq K+1, \omega_{i,t}/\alpha > 1)$ 中寻找一个最不可能符合背景模型的高斯分布 $\eta_{s,t}$, 使其样本点个数减 1, 即 $\omega_{s,t+1}/\alpha = \omega_{s,t}/\alpha - 1$ 。

为了寻找 $\eta_{s,t}$, 对 $\eta_{i,t} (1 \leq i \leq K+1)$ 各添加一个变量 F_i , 用于记录 $\eta_{i,t}$ 被最近第几帧的样本点匹配, 即第 F_i 帧的样本点与 $\eta_{i,t}$ 匹配, 且从第 $F_i + 1$ 帧到当前帧 (第 $t+1$ 帧) 没有任何样本点与 $\eta_{i,t}$ 匹配。

如果 $\eta_{p,t}$ 比 $\eta_{q,t}$ 更接近于背景模型, 则适当样本空间中 $F_p > F_q$ 的概率要远大于出现 $F_p \leq F_q$ 的概率。

因此, 在 $\eta_{i,t} (1 \leq i \leq K+1, \omega_{i,t}/\alpha > 1)$ 中寻找 F_i 最小的一个高斯分布 $\eta_{s,t}$ (如果多个 F_i 同时达到最小, 则优先选择 $\eta_{K+1,t}$, 然后选择权重最大的), 使 $\omega_{s,t+1}/\alpha = \omega_{s,t}/\alpha - 1$, 即 $\omega_{s,t+1} = \omega_{s,t} - \alpha, \omega_{i,t+1} = \omega_{i,t} (1 \leq i \leq K+1, i \neq m, i \neq s)$ 。当 $\eta_{m,t}$ 被 X_{t+1} 匹配时, $F_m = t+1$; 因此如果 $s = m$, 则 $\omega_{i,t}/\alpha \leq 1 (1 \leq i \leq K+1, i \neq m)$, 这时 $\omega_{m,t}$ 先加 α 再减 α , 相当于不变。

为了确保稳定性, 训练 $P(X_t)$ 的样本空间比较大 (通常 $1/\alpha = 500$)。假设 $\eta_{s,t} (s \neq K+1)$ 不符合背景模型且 $\omega_{s,t}$ 接近于 1, 则 $\omega_{s,t}/\alpha$ 将是比较大的数; 即使 $\omega_{s,t}$ 每帧减 α , 收敛速度也会较慢。因此, 可以设定阈值 D , 当 $t+1 - F_s > D, \omega_{s,t} > A \cdot \alpha$ 且 $s \neq K+1$ 时, 令 $\omega_{s,t+1} = \omega_{s,t} - A \cdot \alpha (A > 1)$, 同时更新 $\omega_{K+1,t+1} = 1 - \sum_{i=1}^K \omega_{i,t+1}$ 。 D 与运动目标经过固定像素点所需帧数的平均值有关, 且 D 远小于 $1/\alpha$ 。

收敛过程中通常会出现小样本情况, 即 $\sum_{i=1}^K \omega_{i,t} \ll 1$, 这时式 (3) 需要改变为

$$B = \operatorname{argmin}_b \left(\sum_{i=1}^b \omega_{i,t} / \sum_{i=1}^K \omega_{i,t} > H \right) \quad (8)$$

4 M_2 模型适应性分析

已知模型 M_1 不仅适应于背景自身缓慢变化的

情况, 而且适应于复杂背景的目标检测。这主要是因为:

(1) M_1 模型采用式 (5) ~ 式 (7) 不断更新高斯分布的期望值和方差。当背景自身缓慢变化时, 新背景模型高斯分布的期望值和方差会有少量偏移, 因此, 大量样本点还会落入原背景模型的高斯分布, 通过式 (5) ~ 式 (7), 原背景模型高斯分布会得到不断修正而趋向于新背景模型高斯分布。这样 M_1 模型就能适应于背景自身缓慢变化的情况。

(2) 在复杂背景的目标检测过程中, 背景模型通常需由多个高斯分布来描述。 M_1 模型采用 K 个高斯分布来统计像素点在每帧的样本点, 并按式 (3) 将样本点个数最多的几个高斯分布作为背景模型。这样 M_1 模型就能合理找到符合背景模型的多个高斯分布。

M_2 模型与 M_1 模型相比, 只是改变了式 (3)、式 (4)。将式 (3) 改为式 (8) 是为了避免出现小样本时权重和不等于 1 的情况, 更改式 (4) 后匹配率高的高斯分布始终会优先被判定属于背景模型。 M_2 模型在上述两个原因上与 M_1 模型并不冲突, 因此, M_2 模型同样适应于背景自身缓慢变化和复杂背景目标检测的情况。

5 试验结果及其讨论

本文提出 M_2 模型的收敛性用图 1 中的户外高速公路视频进行检测, 从第 61 帧开始该视频的亮度发生改变。为了与文献 [7] 中 M_3 模型对比, M_2 模型在第 1 帧将 K 个高斯分布设置为

$$\begin{cases} \omega_{i,1} = \alpha & i = 1, 2, \dots, K \\ \omega_{K+1,1} = 1 - \sum_{j=1}^K \omega_{j,1} \\ \mu_{i,1} = X_1 & i = 1, 2, \dots, K \\ \sigma_{i,1}^2 = 20 & i = 1, 2, \dots, K \end{cases} \quad (9)$$

并设置参数 $K=3, \alpha=0.002, H=0.6, D=20, A=20$ 。

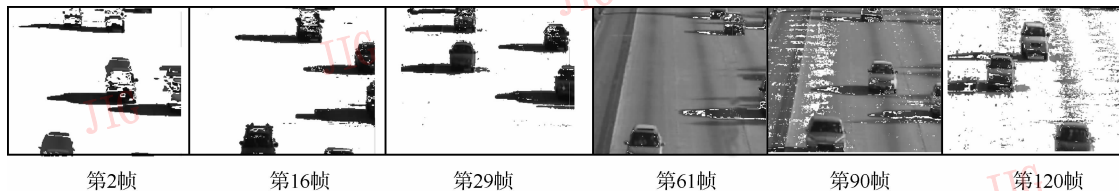
图 2、图 3 给出了 M_3 模型和 M_2 模型收敛性的对比效果。从第 2 帧到第 29 帧, 两模型的收敛效果基本相同; 但从第 61 帧到第 120 帧, M_2 模型的收敛效果明显优于 M_3 模型。

M_2 模型具有以下两个性质:



图 1 原始视频(从第 61 帧开始背景发生改变)

Fig. 1 Original Video (background has changed since the 61 frame)

图 2 文献[7]中的 M_3 模型的收敛效果Fig. 2 The convergence effect of model M_3 proposed by reference [7]图 3 本文提出的 M_2 模型的收敛效果Fig. 3 The convergence effect of model M_2 proposed by this paper

性质 1 当 $\eta_{i,t} (1 \leq i \leq K)$ 中包含背景模型时, $\eta_{K+1,t}$ 被匹配的概率会很低, $\omega_{K+1,t}$ 将快速减小, 使 $\eta_{i,t} (1 \leq i \leq K)$ 中有效样本点(属于背景模型的样本点)个数趋于 $1/\alpha$;

性质 2 当 $\eta_{i,t} (1 \leq i \leq K)$ 中不包含背景模型时, $\eta_{i,t} (1 \leq i \leq K-1)$ 被匹配的概率远低于 $\eta_{K+1,t}$, $\eta_{i,t} (1 \leq i \leq K-1)$ 中无效样本点(不属于背景模型的样本点)将快速减少, 因为 $\eta_{K+1,t}$ 被匹配时 $\eta_{K,t}$ 同时被更新, 所以包含背景模型样本点的 $\eta_{K,t}$ 的权重 $\omega_{K,t}$ 将快速大于 $\omega_{i,t} (1 \leq i \leq K-1)$ 。

根据上述两个性质, 可以分析 M_2 模型在图 3 中的收敛效果。从第 2 帧到第 29 帧: 由式(9)知, M_2 模型在第 1 帧时, $\eta_{i,1} (1 \leq i \leq K)$ 中无效样本点个数不大于 K , 根据性质 2, $\eta_{i,t} (1 \leq i \leq K-1)$ 将快速包含背景模型, 再根据性质 1, $\eta_{i,t} (1 \leq i \leq K)$ 中有效样本点个数将快速增加, 使 M_2 模型在第 29 帧时趋于稳定。从第 61 帧到第 120 帧: M_2 模型在第 60 帧时, $\eta_{i,60} (1 \leq i \leq K)$ 中不包含改变后的背景模型, 根据性质 2, $\eta_{i,t} (1 \leq i \leq K-1)$ 中无效样本点个数快速减少, 因此与 M_3 模型相比, 改变后的背景

模型在 M_2 模型中将更快占据主导地位。

文献[7]的 M_3 模型在检测过程中, 没有对背景的改变进行判断。如果加上背景改变的判断规则, 则从第 61 帧以后, 可以重新对模型进行初始化, 从而加快收敛速度, 但这样就会出现引言中讨论的问题(1)。因为背景模型是统计模型, 所以要用一定数量的样本点来判断背景是否发生改变; 如果样本点个数太少, 则会出现引言中讨论的问题(2)。如果用 30 帧的样本点来判断背景是否改变, 则 M_3 模型将在第 90 帧时对模型进行初始化。图 4 给出了第 90 帧初始化后 M_3 模型的收敛效果, 图 5 给出了 M_2 模型在相应帧的收敛效果。

通过图 4、图 5 的对比, 可以发现, 从第 90 帧到第 120 帧, 加上背景改变判断规则后的 M_3 模型与 M_2 模型的收敛效果相差不大。

因为 M_2 模型用于训练 $P(X_t)$ 的样本空间中样本点个数 $1/\alpha$ 较大, 所以少数帧的异常不可能导致 $\eta_{i,t}$ 快速失去第一的位置; 因此, M_2 模型在确保稳定性的前提下, 加快了收敛速度。

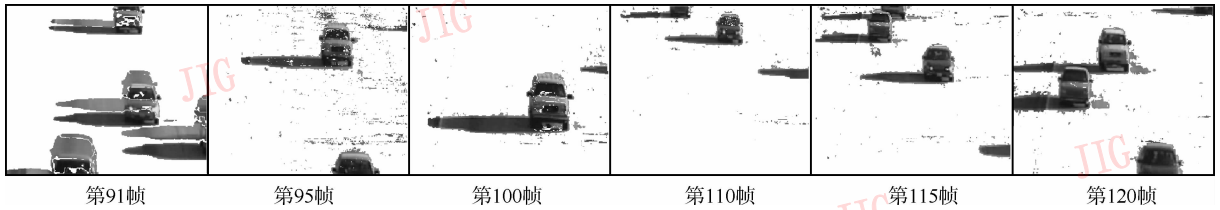


图 4 第 90 帧重新初始化后 M3 模型的收敛效果

Fig. 4 The convergence effect of model M3 initialized in the 90 frame

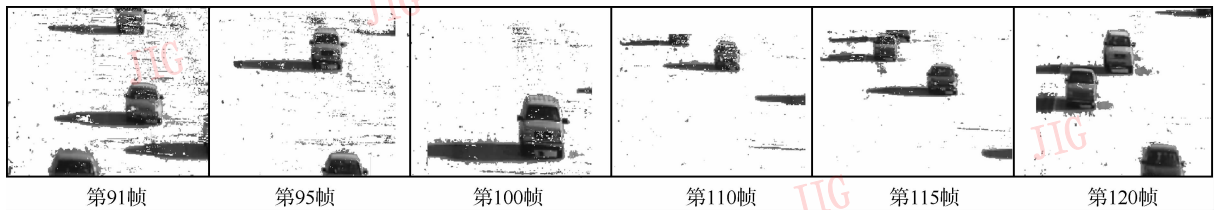


图 5 M2 模型在特定帧的收敛效果

Fig. 5 The convergence effect of model M2 in certain frames

6 结 论

本文提出了快速收敛的混合高斯 M_2 模型,该模型在确保稳定性的前提下,加快了背景模型的收敛速度。 M_2 模型与其他快速收敛算法相比,不需要给出背景改变的判断规则,且不会受少数几帧的影响而完全放弃前面的有效样本点。因此, M_2 模型对混合高斯模型(M_1)收敛性的改进比较明显。

参考文献 (References)

- 1 Wren C R, Azarbayejani A, Darrell T, *et al.* Pfnder: Real-time of the human body[J]. IEEE Transactions on PAMI, 1997, **19**(7): 780 ~ 785.
- 2 Friedman N, Russell S. Image segmentation in video sequences: A probabilistic approach[A]. In: Proceedings Thirteenth Conference on Uncertainty in Artificial Intelligence [C], Providence, Rhode Island, USA, 1997: 175 ~ 181.
- 3 Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], Fort Collins, Colorado, USA, 1999:246 ~ 252.
- 4 Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking[J]. IEEE Transactions Pattern Analysis and Machine Intelligence, 2000, **22**(8):747 ~ 757.
- 5 Zoran Ziykovic. Improved adaptive Gaussian mixture model for background subtraction [A]. In: Proceedings of the International Conference on Pattern Recognition [C], Amsterdam, Netherlands, 2004:23 ~ 26.
- 6 KaewTraKulPong P, Bowden R. An improved adaptive background mixture model for realtime tracking with shadow detection[A]. In : Proceedings of 2nd European Workshop on Advanced Video Based Surveillance System (AVBS01) [C], Kingston, UK: Kluwer Academic Publishers, 2001: 1 ~ 5.
- 7 Sun Yun-da, Li Ming, Wu Wei, *et al.* Background model initialization in moving object detection with shadow elimination[A]. In: Proceedings of the 7th International Conference on Signal Processing [C], Beijing, China, 2004:1288 ~ 1291.
- 8 Wang Chao, Hou Li-min. A new parameter estimation algorithm of Gaussian mixture model[J]. Journal of Shanghai University (Natural Science Edition), 2005, **11**(5): 475 ~ 480. [王超, 侯丽敏. 一种新的高斯混合模型参数估计算法[J]. 上海大学学报(自然科学版), 2005, **11**(5): 475 ~ 480.]