

中图分类号: TP391 文献标志码: A 文章编号: 1006-8961(2010)12-1790-06

论文索引信息: 蒋鹏, 秦小麟. 利用背景聚类的快速前景分割算法[J]. 中国图象图形学报, 2010, 15(12): 1790-1795

利用背景聚类的快速前景分割算法

蒋鹏, 秦小麟

(南京航空航天大学计算机科学与工程系, 南京 210016)

摘要: 提出一种利用背景聚类的快速前景分割算法。该算法首先通过一种专门用于背景聚类的无监督模糊聚类方法将历史像素值进行聚类, 继而用高斯成分来模拟每一个聚类, 构建了基于聚类的时间域的背景模型。前景的分割则采用阈值化方法对像素属于背景的概率进行二分化处理。由于该方法能够根据场景自适应确定背景为单模或多模分布, 避免了耗时的背景模型构建和更新过程, 因而减少了内存使用量并提高了检测速度。对于多种场景下的不同视频进行实验, 结果表明该算法能够在保持检测精度的同时, 大幅提高检测速度。

关键词: 无监督聚类; 高斯模型; 自适应

Foreground detection based on unsupervised background clustering

JIANG Peng, QIN Xiaolin

(Department of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract: A statistical background subtraction technique is proposed based on clustering of temporal color/intensity. An un-supervised clustering method is proposed to model a background with serial of clusters. The unimodal or multimodal distributions of background are detected adaptively. We use a Gaussians model to simulate each cluster which prevents the estimation the parameter of mix of Gaussians model. The foreground will be detected by comparing the background possibility with a threshold. Experimental results show our approach has equal or better segmentation performance and is proved capable of real-time processing.

Keywords: unsupervised clustering; Gaussians model; adaptive

0 引言

背景差法是目前一种常用的运动前景分割方法, 在智能视频监控等领域有广泛应用。背景差法具有速度快、精度高等特点, 因此近年来受到越来越多研究者的关注。如何更有效地对背景建模是目前背景差法研究的重点问题。一般而言, 背景建模可以分为单模和多模两种方式。其中单模方式是指参考背景的每个点是由一个像素值表示。常见的单模方法包括在一段图像序列中, 利用前几帧的平均

值作为背景^[1]; 以像素点处于稳定状态最长的灰度值作为背景像素灰度值^[2], 或者以出现频率最高的像素值作为背景 (MPB)^[3] 等。但背景可能不是绝对静止的, 例如, 背景中包含有风吹动的树叶或者波动的水面等, 因此单模方式无法以一个像素值来表示背景的历史变化情况, 从而引起分割的精度下降。多模方式通过对每个像素进行建模, 保留了背景的多个状态, 背景模型已经不是由一个像素点表示, 因此多模方式可以有效地解决背景细微变化的问题。混合高斯 (MoG) 背景模型^[4] 是一种常见的多模背景模型, MoG 模型中的每个高斯成分对应场景的一

基金项目: 国家自然科学基金项目 (60673127); 国家高技术研究发展计划 (863) 项目 (2007AA01Z404)。

收稿日期: 2009-03-31; **改回日期:** 2009-07-13

第一作者简介: 蒋鹏 (1976—), 男, 南京航空航天大学计算机应用专业博士研究生。主要研究方向为多媒体数据库和计算机视觉。

E-mail: jplus@163.com。

个状态。但现有的 MoG 背景建模^[4-5]大多采用固定高斯成分个数的方法。在实际场景中,不同的区域,背景的状态个数通常是不一样的。而且估计 MoG 的参数较慢,难以满足实时性要求。为了增加对于场景的自适应性,Elgammal 等人^[6]提出了一种非参数背景建模算法——核密度估计(KDE),该方法不用对背景的概率模型做任何假设,直接用最近历史像素值估计当前像素属于背景的概率。但 KDE 的计算是基于历史样本基础之上的,因此需要存储大量的历史数据,内存消耗较大。而且在背景更新时需要历史数据进行 KDE 估计,其速度较慢,实时性较差。

在视频监控或智能交通等应用中,往往对算法的实时性要求较高。而过去的多模方法^[4-7]在检测速度方面常常难以满足要求。

为了提高检测速度,文献[8]提出一种在时域上对像素值进行聚类的背景模型。在该方法利用 K 均值算法对像素的灰度值进行聚类,其背景模型由 K 个聚类表示。再根据待检测像素灰度值与这 K 个聚类的相似程度判断是否为前景点。对于背景的更新是动态的创建新类,不需要对历史数据进行重新估计,因此该方法速度较快。但该方法采用固定的聚类数($K=5$),因此每个背景点都需要由 K 个聚类表示,需要消耗大量的内存。文献[7,9]等研究指出,不同点的背景的变化实际上是不相同的,因此采用固定的高斯成分或者固定的聚类数会造成计算资源的浪费。一种解决方法是根据背景的实际变化,自适应用单模或多模表示背景。对于背景静止不变的区域,可以由一个值来表示背景,而对于背景变化较大的区域,则可以保留背景的多个状态。

研究表明,多模方式的耗时主要是在背景模型的建立以及更新上,因此如何在保证检测精度的同时提高检测速度是一个重要研究内容。为了提高检测速度,从背景建模以及更新方面对传统多模方法进行改进。首先在时间域上对历史样本进行无监督聚类,然后采用一种近似的方法,对每一类用一个高斯成分来表示。这样的好处是可以根据背景的变化自适应确定高斯成分数,而且高斯分布是通过聚类近似获得的,因此避免了估计高斯模型参数等计算量较大的操作。背景的更新过程则是通过动态创建新的类或者更新现有类的聚类中心完成的。这样背景更新过程中不需要对整个背景模型进行重新更新,提高了背景更新的效率。前景的检测是根据像

素点与背景高斯模型的相似程度获得背景相似概率,如果概率小于一个阈值的话,则判断成前景,否则为背景。

1 基于时间域无监督聚类的背景模型

1.1 时间域无监督聚类

过去的方法常常采用固定一个或多个的背景状态数。这种方法或者无法保留背景多个状态,或者对所有背景点一概用多个状态表示,导致系统计算资源的浪费。因此有必要根据背景的变化规律自适应确定背景状态数。为此提出一种自适应确定背景状态数的聚类方法。与传统的 K 均值和模糊 C 均值等聚类方法不同的是本文方法可以根据背景的变化情况自动确定背景状态数。首先以像素颜色为特征,提出一种专门用于背景聚类的聚类算法,该算法通过改进 FCM 算法进行背景聚类,其主要思想是首先设置一个较大的聚类数 C,聚类过程中动态合并相似的类别,最后获得最佳聚类数同时完成聚类。在进行无监督聚类前,首先定义两个聚类中心的分散度为

$$S(i, j) = \frac{C(i) + C(j)}{D(i, j)} \quad (1)$$

式中, $C(i)$ 为一个类的聚集度,其中聚集度定义为

$$C(i) = \frac{1}{N_i} \sum_{x_k \in V_i} |x_k - V_i| \quad (2)$$

N_i 为聚类于 V_i 的像素个数。 D 为两个类之间的距离,定义为

$$D(i, j) = |V_i - V_j| \quad (3)$$

设置的类合并条件是:类的分散度较小或者两个类之间的距离过小。即当 $S(i, j) \geq 1$ 或者 $D(i, j) < T_c$,则认为两个聚类过于相似,其中 T_c 为判断类相似的阈值,该阈值可以根据场景设置成 10~20。当两个类相似时,将类 i 和类 j 合并成新类 k 。合并方法如下:

$$V_k = (V_i + V_j) / 2$$

$$u_k = (u_i + u_j)$$

$$N_c = N_c - 1$$

V_k 为合并后类的聚类中心, u_k 为新类的隶属度, C 为合并后类总数。

因此本文的背景聚类算法如下:

1) 设置一个较大的聚类数初始值 N_c 并用随机数初始化隶属矩阵 U 以及聚类中心。

2) 利用 FCM 更新规则更新 N_c 个聚类中心 V 和隶属矩阵 U 。

3) 计算 FCM 价值函数 J 。如果 J 大于某个固定的阈值,或它相对上次价值函数值的改变量大于某个阈值,则转至步骤 2)。

4) 计算各个类间分散度和相似度,并合并相似的类。如果 C 值减小,转到步骤 2),如果 N_c 值稳定则退出算法。

FCM 更新规则和价值函数 J 参见文献[10]。通过比较实验,聚类初始值设置成 $N_c = 5$ 可以获得良好的效果,即认为背景的状态最多为 5 个。

当背景点变化较小时,利用本文的类合并方法可以将相似的类进行合并,聚类数较少。而当背景变化较大时,则类之间差距较大,聚类数也相应较多,因此可以有效保留背景的多个状态。如果将每个点处的聚类数用颜色的深浅来显示的话,可以形成一个背景状态分布图。颜色越深,表示其聚类数越多,反之亦然。

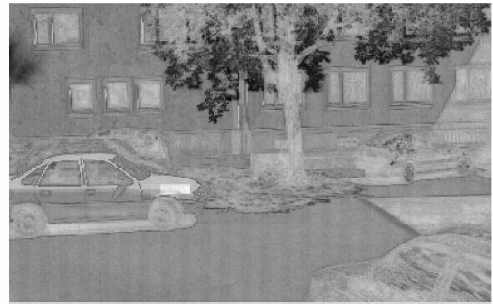
图 1 为以 Vssn06(video 7)为例的背景状态分布图。该场景为一个典型的户外场景,图像上方的树叶被风吹过,该处的背景有细微的运动。从图 1(b)可以看出,在树叶处,其背景变化较大,因此聚类数也相应较多(文本聚类结果为 3),在图 1(b)中显示颜色较深。其他区域例如马路,背景几乎不变,本文获得的聚类数为 1,在图 1(b)中显示的颜色较浅。从图 1 中可以看出,无监督聚类算法可以有效地根据背景变化自适应确定背景状态数。对于该场景而言,约 91% 的面积都可以用一个聚类来表示。这样动态决定聚类数的好处是无需对场景中的路面或车辆等静态区域保留多个状态,这样可以有效地减少计算量和内存使用量。

1.2 基于聚类的快速背景构造

在获得背景聚类后,采用一种基于聚类的快速背景构造方法。首先回顾文献[4]的 MoG 方法,MoG 方法的背景是由多个高斯成分 $M_g(\mu, \sigma^2)$ 组成,其中 μ 为高斯成分的均值,而 σ^2 为方差。由于聚类算法是将具有相似颜色的像素点进行聚类,聚类中心可以看成是该聚类元素的平均值,那么可以用近似的方法将每个聚类看成是一个高斯模型 $M_g(\mu, \sigma^2)$ 。高斯模型的均值可以由聚类的聚类中心表示,即 $\mu = V_i$ 。方差值 σ^2 是用于描述样本分散程度,方差较大时,离聚类中心较远的样本也可以得到



(a) 原图



(b) 背景状态分布图

图 1 背景状态分布图示例

Fig. 1 Example of background modal number distribution map

较大的概率值,这样有利于抑制背景的噪声。因此本文对所有的聚类均采用相同的方差值以减少计算量,方差值设置成任意一个聚类数为一的聚类的方差值。

设一个点处背景样本集共有 N_c 个聚类,如果将每个聚类表示成高斯分布,则待检测像素 x 属于背景的概率为

$$p(x) = \sum_{i=1}^{N_c} \frac{w_i}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - V_i)^2}{2\sigma^2}\right) \quad (4)$$

式中, w_i 为每个高斯分布所代表聚类的权重,定义为

$$w_i = \frac{N(i)}{\sum_{i=1}^{N_c} N(i)} \quad (5)$$

即包含元素越多的聚类,其权重也就越大。注意式(4)是采用概率和的形式,即每个像素是与所有的高斯分布的相似性之和。这样能更准确地反映像素颜色与多模背景模型的相似度。

在获得检测像素 x 属于背景的概率 $p(x)$ 值之后,一个简单的前景分割方法是比较 $p(x)$ 和一个阈值 T_g ,当 $p(x)$ 满足式(6)时

$$p(x) < T_g \quad (6)$$

则认为检测到前景,否则为背景。采用文献[11]中

设置优化阈值的方法设置阈值 T_g ,即用直方图统计 $p(x)$ 的分布,查找直方图的第一个波谷作为优化阈值。具体实现见文献[11]。

2 背景模型的自适应更新

在背景初始化完毕以后,算法进入背景模型更新阶段。过去的多模方法在更新背景时,往往需要重新更新整个背景模型,因此速度较慢。为了解决背景更新的速度问题,用一种动态更新和创建聚类的方法自适应更新背景。自适应背景更新包括两大部分:1)当光照缓慢变化时,背景聚类的聚类中心可能会随光照的变化而发生漂移,因此需要根据聚类内元素的分布自适应调整聚类中心;2)当前景停止成为背景的一部分时,背景需要将静止的前景更新到背景中。如果 x 与某个聚类中心的距离小于 T_c ,则将 x 分配到该聚类中,并更新该聚类的聚类中心,更新后的聚类中心为

$$V'_i = \frac{V_i \cdot N_i + x}{N_i + 1} \quad (7)$$

式(7)实际上是利用 x 与聚类中所有元素的均值对聚类中心进行调整。这样当有缓慢的光照变化时,聚类中心可以随着光照的变化不断地向匹配次数较多也就是真实背景颜色处移动。

当 x 与所有的聚类中心的距离都大于 T_c 时,则认为 x 可能是新出现的前景点。本文的做法是为 x 创建一个新的类,并用 x 的颜色值作为聚类中心。这种方法与文献[8]的方法相似。这样的设计也可以解决一些高斯模型无法容忍的噪声问题。当有前景出现时,由于新创建的聚类元素数目相对较少,因此根据式(5),其权重也较少,并不会影响前景分割。当前景停止运动时,前景类的元素逐渐增多,前景类的权重也随之加大,最后停止运动的前景会以一个新类的形式更新到背景中。

可以看出,背景更新方法是采用对已有类进行更新或增加新类的方法实现的,因此无需对背景历史数据重新更新。由于避免了耗时的背景更新操作,其背景更新速度较快。

3 实验结果及分析

为了测试本文整体算法的有效性和鲁棒性,对大量标准视频和实际视频进行了测试。测试环境为

一台 P-IV 1.5G, 512M 内存的 PC,并用 VC++ 实现算法。受文献[7]以查表的形式计算 KDE 的启发,在实现中本文用一种查表的形式计算式(4),以优化算法速度,并采用基于连通成分^[7]后处理方法,去除面积过小的孤立区域。测试视频包括各种条件下的室内和室外场景,并与一种改进的自适应高斯混合模型方法(MoG)^[6]和单模方法——出现概率最大的点为背景点(MPB)^[3]的方法进行了对比。图2为实验对比结果图。第1行为对 Vssn06 的测试结果。由于该视频为户外场景,在该场景中,树叶随风摆动,光照也在缓慢变化。在该场景中约 91% 的面积的背景可以用一个聚类表示。在树叶区域,其背景的聚类数目较多,因此其背景也相应的由多个高斯模型组成。在树叶区域,算法就近似成 MoG 方法。从图2可以看出,由于采用多个高斯模型对树叶区域进行建模,本文算法和 MOG 在树叶区域只产生了少量的虚假目标信息。本文方法与 MOG 相比,摆动的树叶的对前景造成的干扰更少。而单模的 MPB 算法采用单一背景值作为背景,所以当树叶晃动时,该方法无法保持背景的多个状态,因此产生了较多的虚假前景信息。

图2第2行为对 tower1_set2 的测试结果。该场景为一个真实的户外监控场景。该场景中,利用本文算法,约 98% 的面积都可以用一个聚类来表示。从前景分割结果来看,本文算法提取的前景外形更加清晰,背景噪声也相应较少。这说明本方法的背景聚类以及高斯模型近似方法能够较好地表现背景的状态。而自适应的聚类数又可以更准确的描述背景的变化程度。

图2第3行为对 hall monitor 的实验结果。该视频为室内拍摄的监控录像,光照和背景都相对稳定。因此利用本文方法,超过 99% 的场景都是由一个聚类表示。本文算法就退化成传统的单模方法。从图2可以看出,本文分割结果和 MPB 相似,优于 MoG 方法。这是因为 MoG 方法采用预先设置固定的高斯成分数,当高斯成分设置过多时,会导致漏检。而高斯成分过少时,又会造成误报。

为了定量分析实验结果,对多个已标注的标准测试视频进行了验证,并与相关方法进行了实验比较。实验结果采用文献[12]中定义的查全率(recall),查准率(precision)进行度量。其中查全率是检测正确数与手工标注(ground truth)总数的比值,查准率是检测正确数和总共检测数的比值。

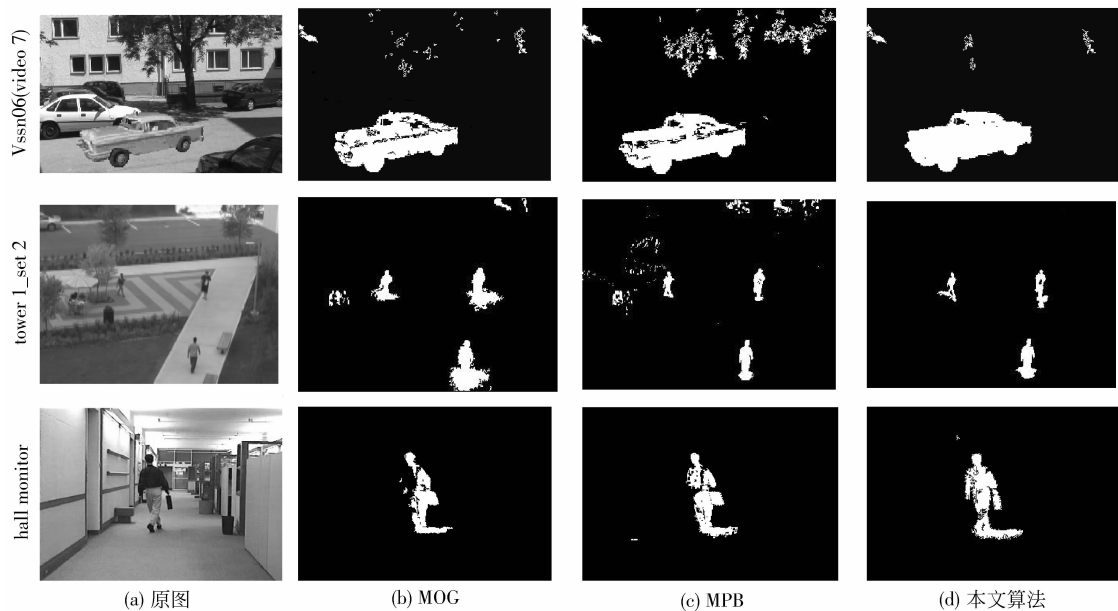


图 2 实验对比结果图

Fig. 2 The comparison results of different methods

表 1 为对 Vssn06-video 7, (300 ~ 400 帧) 和 hall monitor (30 ~ 80 帧) 的平均值的定量实验结果。

从表 1 可以看出, 对 Vssn06, 本文算法查准率优于 MoG, 而查全率略低于 MoG。这是因为本文算法对于背景复杂区域采用多个聚类进行表示, 因此错误的前景数较少。但本文算法是根据聚类的均值来模拟高斯分布的, 因此并未对均值和方差等参数进行精确估计, 造成部分前景区域未检出来, 在前景模板图中表示出来就是前景内小的空洞。这个问题可以采用区域生长算法进行弥补。而在静态场景 (hall monitor), 本文算法查全率和查准率与 MPB 相似, 但高于 MoG 方法。

表 1 实验结果对比表

Tab. 1 Quantitative evaluation on precision, recall and timing

	Vssn06 (video 7) (384 × 240)			hall monitor (352 × 240)		
	查全率	查准率	平均每帧 耗时/ms	查全率	查准率	平均每帧 耗时/ms
MoG	0.93	0.84	85.1	0.78	0.83	82.3
MPB	0.82	0.69	9.5	0.87	0.90	9.2
本文算法	0.89	0.92	65.2	0.89	0.93	60.2

从算法的执行速度来看, 本文算法运行速度较 MoG 提高超过 25%。原因是两种方法的计算复杂度很大程度都来源于背景的高斯成分数以及聚类数。本文方法对背景的不同区域进行区别对待。对

于大部分区域, 利用式 (4) 计算前景概率时, 聚类数为 1, 因此仅需要一次查表就可以获得前景概率, 其运算复杂度与 MPB 相仿。而 MoG 算法在设置了高斯成分数后不会再改变, 因此效率较低。另外, 在构建背景模型时, 本文方法的高斯模型是直接根据聚类中心和元素数目模拟获得的, 不需要利用每个元素来计算高斯模型所需的均值和方差, 因此速度很快。

4 结 论

提出一种利用背景聚类的快速运动前景分割算法。该算法通过一种专用的无监督聚类算法实现了时域上像素值的聚类, 继而用一种近似方法将每个聚类模拟成一个高斯分布。由于聚类数是根据背景的复杂度自适应生成的, 因此可以根据场景自适应优化计算资源。而且将聚类近似成高斯分布避免了高斯模型的参数估计, 因此速度较快。但由于本文方法并未对历史样本的均值和方差等参数进行精确估计, 而是通过聚类近似而成的, 因此牺牲了部分的查全率换得了较高的查准率以及运行速度。为了提高查全率, 可以将现有的区域生长算法集成到本文算法中。总体而言, 本文算法在保持分割精度的同时, 提高了分割速度。从定性以及定量的实验都能证明出本文算法的有效性和鲁棒性。

参考文献 (References)

- [1] Friedman N, Russell S. Image segmentation in video sequences: a probabilistic approach [C]//Proceedings of Thirteenth Conference on Uncertainty in Artificial Intelligence. Providence, Rhode Island, USA; Morgan Kaufmann Publishers, 1997:175-181.
- [2] Gutchess D, Trajkonic M, Cohen-Solal E, et al. A background model initialization algorithm for video surveillance [C]//Proceedings of the 8th IEEE Int'l Conf. on Computer Vision. Vancouver, Canada; IEEE Computer. Society, 2001:733-740.
- [3] Hou Z Q, Han C Z. A background reconstruction algorithm based on pixel intensity classification [J]. Journal of Software, 2005, 16(9):1568-1576. [侯志强, 韩崇昭. 基于像素灰度归类的背景重构算法 [J]. 软件学报, 2005, 16 (9): 1568-1576.]
- [4] Stauffer C, Grimson E L. Learning patterns of activity using real-time tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3): 747-757.
- [5] Lee D S. Effective Gaussian mixture learning for video background subtraction [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(5): 827-832.
- [6] Elgammal A, Duraiswami R, Harwood D, et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance [J]. Proceedings of IEEE, 2002, 90(7): 1151-1163.
- [7] Zivkovic Z, van der Heijden F. Efficient adaptive density estimation per image pixel for the task of background subtraction [J]. Pattern Recognition Letters, 2006, 27(7): 773-780.
- [8] Darren E Butler, Michael V. Real-time adaptive foreground background segmentation [J]. Real-Time Imaging, 2005, 6(11): 172-185.
- [9] Stenger B, Ramesh V, Paragios N, et al. Topology free hidden markov models: application to background modeling [C]//Proceedings of IEEE International Conference on Computer Vision. Vancouver, BC, Canada; Publisher: IEEE Computer. Society, 2001: 294-301.
- [10] Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms [M]. New York: Plenum Press, 1981.
- [11] Du-Ming Tsai. A fast thresholding selection procedure for multimodal and unimodal histograms [J]. Pattern Recognition Letter, 1995, 16(6): 653-666.
- [12] Maddalena L. Petrosino. A self-organizing approach to background subtraction for visual surveillance applications [J]. IEEE Transactions on Image Processing, 2008, 17(7): 1168-1177.