

基于视频序列的双背景建模算法^{*}

李文辉^{a,b}, 何林宣^b, 倪洪印^b

(吉林大学 a. 汽车仿真与控制国家重点实验室; b. 计算机科学与技术学院, 长春 130012)

摘要: 在研究传统码本和滑动平均背景建模算法的基础上提出了一种基于视频序列的双背景建模算法, 并利用统计的方法提出了一种基于 HSV 特征的背景更新方法, 有效滤除了噪声点, 解决了光线突变下背景更新的问题。实验表明该算法对光照突变鲁棒性强, 响应速度快, 适合实时图像处理。

关键词: 双背景建模; 目标检测; 码本; 滑动平均

中图分类号: TP391; TP301.6

文献标志码: A

文章编号: 1001-3695(2014)02-0629-04

doi:10.3969/j.issn.1001-3695.2014.02.076

Algorithm on video-based dual background modeling

LI Wen-hui^{a,b}, HE Lin-xuan^b, NI Hong-yin^b

(a. State Key Laboratory of Automotive Simulation & Control, b. School of Computer Science & Technology, Jilin University, Changchun 130012, China)

Abstract: This paper proposed a new dual background model that based on the research of the traditional codebook and moving average background model. These algorithms were all based on the application of the video sequence. This new background modeling algorithm proposed a updating method based on the HSV features and statistical methods, it filtered out the noise points and solved the background updating problems under the illumination mutation effectively. Experiments show that this algorithm has a strong robustness to illumination mutation, a fast response and it is quite fit for real-time image processing.

Key words: dual background modeling; target detection; codebook; moving average

在智能视频监控、运动目标检测等应用中,从视频序列中检测出运动目标是一个基础和关键的环节,它对后续的跟踪、行为识别及分析等处理有着重要的影响。

基于视频序列的目标检测目前主要有光流法^[1,2]、帧差法和减背景^[3,4]三种方法。光流法由于计算复杂度高而很少在实时处理中使用。帧差法由于算法思想较为简单,虽然时间复杂度很低,但存在抗干扰能力较弱、易受环境因素影响等不利于实时处理的明显缺点,在处理复杂场景的应用中显得捉襟见肘,所以越来越多的人把研究的重点放在了基于背景建模算法的目标检测领域中。主流的建模算法有码本建模、高斯建模以及滑动平均建模等。文献[5]对自适应阈值的滑动平均背景建模进行了研究。文献[6]对传统码本背景建模进行了改进,通过缓存码本的方法降低了噪声点。文献[7,8]提出了一种修改后的码本建模方法,改进了背景训练方法,不需要长的训练序列,而用压缩的形式代替传统码本,降低了码本对内存的要求。文献[7]通过基于颜色失真和亮度范围特征的一些码字表示码本背景模型,将像素从 RGB 空间转换为 HSV 空间来计算颜色失真,并将几种计算颜色失真和亮度范围的方法做实验进行了对比。文献[9,10]总结了目前常用的目标检测减背景算法,将这些算法进行了实验和比较,并指出了本领域今后的研究方向和关键问题。文献[11]提出了一种基于码本模型的实时前景背景分割算法,该方法与传统码本算法相比,提高了内存使用效率,并且可以处理光照变化的运动场景。由于码

本背景建模算法的时间复杂度较高,而时间复杂度较低的滑动平均建模方法又存在噪声点较多,不能适应复杂场景变化的缺点,本文主要综合了前人对码本和滑动平均建模两种方法的研究成果,提出了一种新的双背景建模算法,有效地弥补了两种算法中存在的噪声点多、时间复杂度高以及易受光照突变影响的不足。

1 研究背景

基于码本的背景建模算法由码本模型训练、前景提取以及码本模型更新三个阶段构成。采用量化的方法来构建码本模型,针对每一帧构建一个码本模型。码本模型中的码本由循环采样的每一帧中的每个像素点的亮度范围和颜色距离所构成的码字所组成。根据采样值的变化情况,每个码本中的码字数有可能不同。经过对码本背景建模算法研究发现,码本背景建模在室内静态环境下前景提取较为完整,存在较少数量的噪声点,但是码本算法对室外光照突变的处理能力较弱,算法处理的时间复杂度较高,无法满足实时处理的要求。

滑动平均背景建模处理视频序列时,其时间复杂度要远远小于其他背景建模方法的时间复杂度,但是滑动平均依赖其背景更新系数 α 的取值,当 α 值不同算法效果也有所不同。当 α 取值大时,前景点提取不完整;当 α 取值较小时,噪声点较多。并且当出现剧烈的光照变化以及处理复杂环境时,滑动平均的处理效果并不理想。

收稿日期: 2013-04-01; 修回日期: 2013-05-29 基金项目: 吉林省科技发展计划资助项目(20120305)

作者简介: 李文辉(1961-),男,湖南株洲人,教授,博导,博士,主要研究方向为图像处理与虚拟现实(liwh@jlu.edu.cn);何林宣(1988-),男,河南浉池人,硕士研究生,主要研究方向为图像处理与虚拟现实;倪洪印(1984-),男,吉林德惠人,博士研究生,主要研究方向为图像处理与虚拟现实。

为解决码本的算法处理时间较高和滑动平均的前景提取不完整、噪声点较多、背景更新效果不理想、难以适应复杂场景以及易受光照突变影响等问题,本文提出了下面的双背景算法背景建模方法。

2 双背景建模

双背景建模算法在滑动平均和码本两种主要的建模方法的基础上,利用统计的方法提出了一种基于 HSV 特征的背景更新方法。此外,本文的双背景建模还加入了光照突变处理机制,最终形成了一个以两种主要背景建模方法为基础、以 HSV 色彩空间特征为标准更新背景模型的一种双背景建模算法。该算法满足实时要求、有效地抑制了噪声点并受光照突变影响较小。

算法思想流程如图1所示。

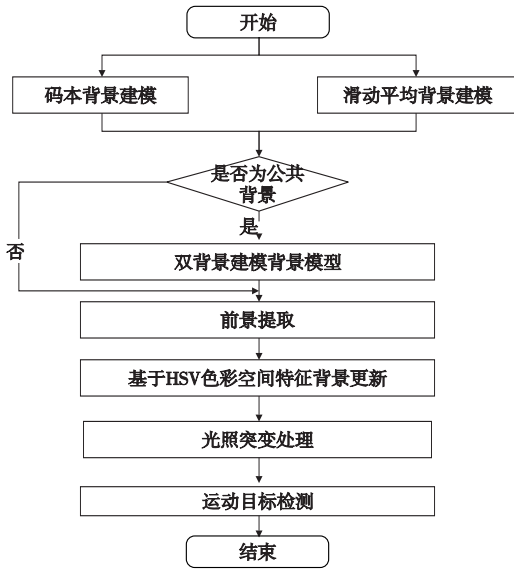


图1 算法思想流程

2.1 双背景建模算法背景模型

本文通过两个步骤得到双背景建模算法的背景模型:a)采用码本训练得到码本背景模型^[5],采用滑动平均背景建模得到滑动平均背景模型;b)根据码本和滑动平均背景得到双背景建模算法的背景模型。

a)采用码本算法训练码本背景模型。首先假设训练序列中单个像素的序列采样值为 $X = \{x_1, x_2, \dots, x_N\}$, 其中的每个元素都是一个 RGB 向量, 设 $C = \{c_1, c_2, \dots, c_L\}$ 为该像素在码本模型中对应的码本, 其中的所有向量都是该像素码本中的 L 个码字。然后对视频每帧采样值和已有像素对应的码本中的码字逐一比较, 如果找到了匹配的码字, 则更新该码字的属性信息, 若未找到匹配的码字, 则新建一个码字并将其链接到相应的码本里。码本的每个码字由两部分信息构成, RGB 向量 $v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$ 和码字属性信息 $\text{aux}_i(\hat{I}, I, f_i, \lambda_i, df_i, dl_i)$, 各属性信息依次代表了最小最大亮度值、码字出现的频率、训练阶段该码字没有出现的最大时间间隔以及首次出现和末次出现的时间。码本建模背景训练的详细过程如下:

(a)初始化。建立与视频帧像素数量大小相同的码本模型, 将每个像素对应的码本初始置空, 每个码本中码字的个数为零。

(b)匹配码字。对于训练视频帧中每个像素的序列采样

值 $X = \{x_1, x_2, \dots, x_N\}$ 中的每一个值 $x_t = (R_t, G_t, B_t)$, $t = (1, \dots, N)$ 定义其亮度值 $I_t = R_t + G_t + B_t$, 并根据同时满足以下颜色距离条件式(1)和亮度范围条件式(2)两个公式的条件找出匹配的码字 c_m 。

$$\text{colordist}(x_t, v_i) \leq \varepsilon_1 \quad (1)$$

$$\text{brightness}(I, \langle \hat{I}, I \rangle) = \text{true} \quad (2)$$

$$\text{其中: } \text{colordist}(x_t, v_i) = \frac{\|x_t\|^2 \|v_i\|^2 - \langle x_t, v_i \rangle^2}{\|v_i\|^2} \quad (3)$$

$$\|x_t\|^2 = R_t^2 + G_t^2 + B_t^2 \quad (4)$$

$$\|v_i\|^2 = R_i^2 + G_i^2 + B_i^2 \quad (5)$$

$$\langle x_t, v_i \rangle^2 = R_t R_i^2 + G_t G_i^2 + B_t B_i^2 \quad (6)$$

$$\text{brightness}(I, \langle \hat{I}, I \rangle) = \begin{cases} \text{true} & \text{if } \hat{I} \leq \|x_t\| \leq I \\ \text{false} & \text{others} \end{cases} \quad (7)$$

其中: x_t 为像素亮度信息; v_i 为码字亮度信息; I 为当前采样值的亮度信息; $\hat{I}$ 为码本模型的相应每个码本中已有码字中的最小和最大亮度值。

(c)有匹配码字则更新码字, 无匹配码字则创建码字。如果步骤(b)条件未满足即码本为空或者未找到匹配的码字, 则创建一个新的码字 c_{new} 。

$$v_L = (R_t, G_t, B_t) \quad (8)$$

$$u_L = \langle I, I, 1, t-1, t, t \rangle \quad (9)$$

否则, 更新匹配的码字 c_m 。具体更新方法如下:

假设码字 c_m 的内容为 $v_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$ 和 $\text{aux}_m(\hat{I}_m, I_m, f_m, \lambda_m, df_m, dl_m)$, 更新码字内容为

$$v_m = \left[\frac{f_m \bar{R}_m + R_t}{f_m + 1}, \frac{f_m \bar{G}_m + G_t}{f_m + 1}, \frac{f_m \bar{B}_m + B_t}{f_m + 1} \right] \quad (10)$$

$$\text{aux}_m = \langle \min\{\hat{I}_m, I_t\}, \max\{\hat{I}_m, I_t\}, f_m + 1, \max\{\lambda_m, t - dl_m\}, df_m, t \rangle \quad (11)$$

(d)消除冗余码字。计算每个码字没有出现的最大时间间隔 λ_i , 对于 $c_i (i = 1, 2, \dots, L)$:

$$\lambda_i = \max\{\lambda_i (N - dl_i + df_i - 1)\} \quad (12)$$

利用 λ_i 来消除冗余的码字, 得到更新后的码本背景模型 $M\{c_k | c_k \in C, \lambda_k \leq T_m\}$, k 为码字的索引, T_m 通常取训练视频帧数的一半, 即 $n/2$, 表示背景的码字至少在最近 $n/2$ 的视频帧中出现过。对视频图像进行滑动平均建模^[5]。以第一帧为背景, 将当前帧与背景帧相减, 建立滑动平均背景模型。如果当前帧与背景帧相减后的差值小于滑动平均的阈值时, 则认为该点是滑动平均模型中的背景点。滑动平均方法利用以下规则对每个像素进行更新:

$$\mu_n(i, j) = \mu_{n-1}(i, j) + \alpha(I_n(i, j) - \mu_{n-1}(i, j)) \quad (13)$$

其中: $\mu_n(i, j)$ 是第 n 帧更新后的背景像素值; $I_n(i, j)$ 为当前帧的像素值; α 为更新系数。

b)由步骤 a)得到码本背景和滑动平均背景, 将其分别存储, 作为双背景建模算法的背景模型的两个组成部分, 共同构成双背景建模算法的背景模型, 在下文前景提取部分中使用。

2.2 双背景建模算法前景提取

双背景建模算法的前景提取分为三个步骤, 具体过程如下:

a)使用码本算法进行前景提取。(a)根据新采样的视频像素点计算其亮度值, 定义与 2.1 节中训练的码本模型中的码

字匹配的布尔值变量 $matched$,且赋初值为 0,并设定恰当的阈值 ε_1 的值,本模块 ε_1 取值 150;(b)根据同时满足颜色距离式(1)和亮度范围式(2)区分前景点,如果找到匹配的码字 c_m ,则 $matched = 1$;(c)根据上述 $matched$ 的值判断每个像素点是前景点还是背景点,如果 $matched = 0$,则该点为前景点,从而得到码本算法的前景点序列 $codebook_foreground$ 。

b)在滑动平均前景阶段,如果当前帧的值与背景帧相减的差值大于阈值,则认为该点为滑动平均模型中的前景点,从而得到滑动平均前景点序列 $average_foreground$ 。

c)将前景点序列 $codebook_foreground$ 、 $average_foreground$ 进行分析判断:如果两个前景序列中都为前景的点,则认定为双背景的前景点序列,否则为背景点。将最终的公共前景点序列以矩阵形式存储作为双背景模型的前景点序列 $foreground$ 。

2.3 HSV 色彩空间特征更新双背景模型

HSV 色彩空间是图像处理中常用的色彩空间之一。其中, H 分量代表色调, S 分量代表饱和度, V 分量代表亮度,三分量相互独立。当 H 分量的值越大时图像帧的色度越大,当 H 分量的值越小时图像的色度越小。当 V 分量较小时,当前图像帧接近于黑色,当 V 分量较大时,图像帧越接近于白色,而当两个像素点的色彩特征差别越大时,相应的两像素的 H 分量差值和 V 分量差值也会有明显变化。利用 HSV 空间分量这一特征,判断出背景噪声点过多的情况,进行背景更新,将其有效地应用到本文算法的背景更新中去。本文以滑动平均更新背景方法进行更新,开始时以视频帧的第一帧和第一帧的 HSV 信息为标准,当满足 HSV 特征变化时,在滑动平均更新方法的基础上对背景进行再更新。具体方法如下:

a)读入视频帧,并记录第一帧的 H 分量和 V 分量信息。

b)用当前帧的 H 分量和 V 分量分别与背景帧的 H 分量与 V 分量做差,结果取绝对值。

c)分别求出每一帧的 H 分量和 V 分量差值绝对值的平均值,对平均值是否满足再更新要求阈值与否进行判断,并决定是否进行再更新。

经统计发现,当前帧中像素点的 H 分量与背景帧的 H 分量的差值的绝对值小于阈值 ε_2 ,并且相应的 V 分量的差值绝对值大于阈值 ε_3 时,原有背景已经不再呈现实际背景的情况,需要对原有背景进行再更新。统计信息如表 1 所示。利用 HSV 空间分量的这一特征,有效保证了背景模型的实时精确度。

表1 H 分量和 V 分量差值平均值统计表部分数据

视频帧截取	H 差值	V 差值	视频帧截取	H 差值	V 差值
	97	19		13	96
	81	32		18	173
	135	35		15	192

表 1 中, H 差值为当前帧中每个像素点的 H 值与背景帧像素点的 H 值做差后取绝对值,然后将所得绝对值求平均值后的值。 V 差值为当前帧中每个像素点的 V 值与背景帧像素点 V 值做差后取绝对值,然后将所得绝对值求平均值后的值。

从统计数据对比可以看出,当背景发生变化时,如光照突变、人群数目突然增多,从而引起原背景中噪声点过多时, H 分

量差值绝对值均值会急剧下降到很小的值。而 V 分量差值绝对值均值会急剧增大到较高的值,因此,设定恰当阈值,并根据阈值变化进行背景更新。

2.4 光照突变的处理

光照突变是复杂场景中必须考虑的环境因素。本文也加入了相应的处理光照突变的方法,在光照突变发生时,将引起场景中几乎所有像素点亮度值的较大变化。针对这个特点,本文采用的方法是:用当前帧与背景帧做差并求绝对值,按照初始阈值得到运动前景图像,统计前景图像中前景点的个数 Ca ,设图像的宽度和高度为 W 和 H ,若

$$\frac{Ca}{W \times H} > T_a \quad (14)$$

则认为在当前帧发生了光线突变,否则没有光线突变发生。 T_a 是检测光线突变的阈值,它的取值可依据场景情况与检测光照突变的速度与误差来进行选取。 $T_a \in (0,1)$ 。

3 实验

3.1 测试环境

在 Intel Core2 Duo CPU E7400,主频 2.80 GHz 2.79 GHz,内存 1 GB 配置的 PC 机上,安装 VS 2005 和 OpenCV 1.0,对自行录制的视频进行检测。六个测试视频皆为录制视频,分辨率为 320×240 。具体各视频参数如表 2 所示。

表2 录制视频

视频序号	时长/s	帧频/fps	视频截图	视频类型描述
1. avi	152	29		室外光照突变
2. avi	486	15		室外树叶晃动
3. avi	152	29		室外多人
4. avi	152	23		室外简单环境
5. avi	34	25		室内单人
6. avi	152	29		室内多人

3.2 测试结果及相关数据分析

经过多次实验测试,确定了阈值参数的取值: $\varepsilon_1 = 30$, $\varepsilon_2 = 20$, $\varepsilon_3 = 80$, $\alpha = 0.8$, $T_a = 0.7$ 。

在阈值参数取值为上述各值时,分别对室内和室外视频进行了实验。室内环境下实验效果和算法处理每帧的时间如图 2、3 和表 3、4 所示。

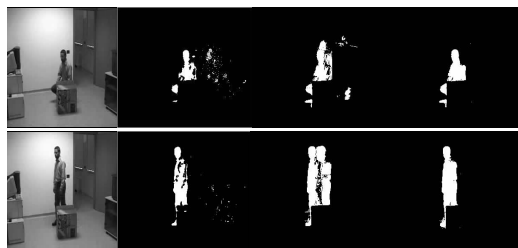


图2 室内环境下测试结果
(从左至右依次为原图、码本、滑动平均、双背景)

表 3 室内环境下(室内视频 5.avi)算法处理每帧时间对比 /ms

算法	最快	最慢	平均
码本	22.341	40.323	25.576
滑动平均	1.443	1.924	1.755
双背景	14.371	16.473	14.447

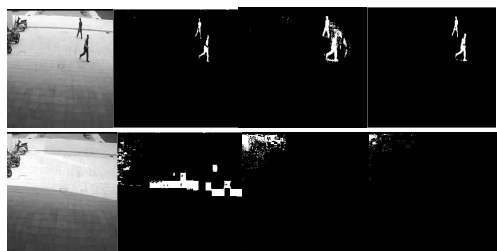
图3 室外环境下测试结果
(从左至右依次为原图、码本、滑动平均、双背景)

表 4 室外环境下(室内视频 3.avi)算法处理每帧时间对比 /ms

算法	最快	最慢	平均
码本	24.671	43.157	25.726
滑动平均	1.454	1.957	1.758
双背景	15.410	31.704	24.437

表 5 中各性能指标为错误前景点率(FP rate)、正确前景点率(TP rate)和精度(precision),计算方法如下^[12]:

$$FP\ rate = \frac{fp}{fp + tn} \quad (15)$$

$$TP\ rate = \frac{tp}{tp + fn} \quad (16)$$

$$precision = \frac{tp}{tp + fp} \quad (17)$$

其中: fp 代表错误的前景点数; tp 代表正确的前景点数; fn 代表错误的背景点数; tn 代表正确的背景点数; $(fp + tn)$ 代表真实前景图像中的背景点总数; $(tp + fn)$ 表示真实前景图像中的前景点总数。

表 5 前景检测结果量化参数对比

算法	技术指标	问题类型			
		室内停留	室内移动	室外多人	光照突变
码本	FP rate	0.052	0.046	0.063	0.097
	TP rate	0.912	0.945	0.859	0.895
	precision	0.468	0.481	0.338	0.275
滑动平均	FP rate	0.047	0.063	0.071	0.017
	TP rate	0.913	0.835	0.787	0.895
	precision	0.475	0.423	0.293	0.563
本文	FP rate	0.113	0.107	0.027	0.031
	TP rate	0.947	0.964	0.816	0.863
	precision	0.791	0.871	0.532	0.861

在以上三项度量指标中,错误前景点率 FP rate 越大,表示前景图像中的噪声越多;正确前景点个数所占比重越多,TP rate 值就越大。

室内和室外两组视频的测试对比中,室内场景下,码本噪声点和滑动平均的运动背景残留现象很明显,经过双背景算法处理后效果有了明显改善;在室外条件下,尤其是在光照突变发生的情况下,双背景建模要优于码本和滑动平均在处理光照突变的效果。在时间复杂度方面,滑动平均的时间复杂度最低,码本的时间复杂度最高,双背景建模的时间复杂度居中但完全符合实时处理的需求;在处理光照突变方面,双背景建模

处理效果最理想,并且处理光照突变的时间复杂度最低。在量化后的前景提取性能数据表中,双背景建模的数据也比其他单独的两种传统建模算法更优。

4 结束语

针对码本的算法处理时间较高和滑动平均的前景提取不完整、难以适应复杂场景以及易受光照突变影响等问题,本文提出了双背景建模算法。经过实验证实,双背景建模能够达到预期的目的,有效地抑制了噪声干扰,对光照突变鲁棒性强,响应速度快,适合实时图像处理。在双背景建模算法中用到了许多阈值,下一步的研究工作就是对该算法作进一步改进并对阈值的自适应设定进行研究。

参考文献:

- [1] YIN Bo, QI Wen-juan, WEI Zhi-qiang, *et al.* Indirect human activity recognition based on optical flow method[C]//Proc of the 5th International Congress on Image and Signal Processing. 2012;99-103.
- [2] SUN Shi-jun, HAYNOR D, KIM Yong-min. Motion estimation based on optical flow with adaptive gradients[C]//Proc of IEEE International Conference on Image Processing. 2000;852-855.
- [3] WANG Jian-lin, YANG Yin-sheng. A moving object detection algorithm based on background subtraction model[C]//Proc of the 3rd IEEE International Conference on Information Management and Engineering. 2011;3273-3276.
- [4] 徐以美, 郭宝龙, 陈龙. 基于 RGB 颜色空间的减背景运动目标检测[J]. 计算机仿真, 2008, 25(3): 214-217.
- [5] 曹昌霞, 沈晓艳, 毕胜, 等. 基于自适应更新率的滑动平均背景更新算法[C]//全国第 18 界计算机技术与应用学术会议论文集. 2007;1001-1005.
- [6] ZHANG Zhao-hui, CHEN Rui-qing, LU Han-qing, *et al.* Moving foreground detection based on modified codebook[C]//Proc of the 2nd International Congress on Image and Signal Processing. 2009;1-5.
- [7] ZHANG Jun, DAI Ke-xue, LI Guo-hui. HSV color-space and codebook model based moving objects detection[J]. Systems Engineering and Electronics, 2008, 30(3): 423-427.
- [8] KIM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Background modeling and subtraction by codebook construction[C]//Proc of International Conference on Image Processing. 2004;3061-3064.
- [9] CHALIDABHONGSE T H, KIM K, HARWOOD D, *et al.* A perturbation method for evaluating background subtraction algorithms[C]//Proc of Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. 2003; 11-12.
- [10] DAI Ke-xue, ZHANG Jun, TU Dan. Prospects and current studies on background subtraction techniques for moving objects detection from surveillance Video[J]. Journal of Image and Graphics, 2006, 11(7): 919-927.
- [11] KIM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Real-time foreground-background segmentation using codebook model[J]. Real Time Imaging, 2005, 11(1): 172-185.
- [12] MADDALENA L, PETROSINO A. A self-organizing approach to background subtraction for visual surveillance applications[J]. IEEE Trans on Image Processing, 2008, 17(7): 1168-117.
- [13] WEN Jun-qin. An adaptive frame difference method for human tracking[J]. Advances in Information Sciences and Service Sciences, 2012, 4(1): 381-387.