

基于自适应背景模型运动目标检测

王陈阳,周明全,耿国华

(西北大学 可视化研究所,陕西 西安 710127)

摘要:随着城市化速度的加快,机动车日益普及,人们在享受机动车所带来的巨大便利的同时,也面临着交通拥挤的困扰。随着计算机硬件技术和计算机视觉技术的发展,基于计算机视觉的交通监控系统成为可能。从一个交通视频序列中识别出运动物体是许多交通监控系统应用系统的重要任务,针对该问题,提出了一种建立在对视频序列中的整个背景情景的统计描述基础上的运动目标的检测的有效方法,该方法能够适应变化的背景,具有较强的鲁棒性和较好的实时性。

关键词:背景差分;背景提取;背景更新;自适应背景模型;运动目标检测

中图分类号:TP391.3

文献标识码:A

文章编号:1673-629X(2007)04-0021-03

Moving Object Detection Based on Adaptive Background Model

WANG Chen-yang, ZHOU Ming-quan, GENG Guo-hua

(Institute of Visualization Technology, Northwest University, Xi'an 710127, China)

Abstract: With the acceleration of urbanization, automobile becomes more and more popular and people are enjoying the convenience that offer; however, people are trapped in the bewilderment of traffic congestion. Along with the computer hardware technology and the computer vision technology development, a computer vision-based traffic monitoring system has become possible. The main task for most a computer vision-based traffic monitoring application system is to identify the moving objects from a video sequence. To solve this problem, authors present a new statistics description of the whole background scene based on the video sequence, establish an effective algorithm for moving object detection. This method can adapt to the changing background with high robustness and excellent real time performance.

Key words: background subtraction; background extract; background update; adaptive background model; moving object detection

0 引言

从视频序列中检测出运动目标在智能视觉监控和人机交互等计算机视觉应用领域具有非常重要的地位,在进行运动检测时,常见的一种情况是摄像机处于静止状态并且焦距是固定的,此时,图像中的背景区域固定不动。在这种情况下,运动目标识别经常使用的方法是背景差分法^[1],在简单的背景差分法使用中,当背景发生一些变化时,如背景中频繁地出现运动物体,或者光照发生变化等等,不能准确地提取背景作为参考图像,从而不能正确地分割出视频序列中的运动物体。为了克服上述问题,文中提出了建立自适应背景模型,并且实现了背景模型的实时更新,能够比较准确地识别出运动目标。

1 建立自适应背景模型

自适应背景模型属于背景差分的范畴,即选取适当的参考帧作为参考图像,用当前帧和参考图像做差分,如果参考图像选取适当,能比较准确地分割出运动物体^[2,3]。所以,建立自适应背景模型的关键在于准确地选取背景图像。

背景是随时间变化而逐渐变化的,而且在一定时间内差别不是很大,人的视觉系统通过使用亮度和色度来区分不同的物体。在通常情况下,输入的视频图像序列包含红、绿、蓝三个颜色分量,混合了亮度和色度信息。这里,用 $X_t = (X_R, X_G, X_B)$ 表示图像中某像素点在时刻 t 的取值。在摄像机镜头固定的情况下,如果没有运动物体,则输入的视频序列被认为是静止的,也就是所说的背景。然而在繁忙的交通路段,直接通过视频图像抓取无运动目标的静态场景是非常困难的,所以需要使用序列图像来合成相对真实的静态交通场景^[4]。

建立背景模型过程如下:

已知图像序列 $(I_1, I_2, I_3, \dots, I_M)$, 从视频序列中

收稿日期:2006-06-25

作者简介:王陈阳(1982-),男,陕西西安人,硕士研究生,研究方向为视频处理、数字图像处理;周明全,教授,博士生导师,研究方向为可视化研究;耿国华,教授,博士生导师,研究方向为可视化研究。

抽取 $N+1$ 帧, 记为 $(X_0, X_1, X_2, \dots, X_N)$ 。在 $N+1$ 帧视频序列中将像素点在 R, G, B 三个颜色分量上的灰度变化同时比较接近的区间称为平稳区间。

步骤 1 划分像素点平稳区间。

$$\begin{cases} f_j^k(x, y) & \textcircled{1} \\ f_{j-1}^k(x, y) & \textcircled{2} \\ \mu_j(x, y) & \textcircled{3} \end{cases} \quad k = R, G, B$$

① 在这里表示在 $N+1$ 帧图像中的第 j 帧 ($j = 0, 1, 2, \dots, N$) 像素点的 (x, y) 处的 R, G, B 灰度值。

② 在这里表示在 $N+1$ 帧图像中的第 $j-1$ 帧 (j 同上) 像素点的 (x, y) 处的 R, G, B 灰度值。

③ 在这里表示划分像素点的判断条件。

$$\mu_j(x, y) = \begin{cases} 0 & |f_j^k(x, y) - f_{j-1}^k(x, y)| \leq \varepsilon \\ 1 & \text{其它} \end{cases} \quad (1)$$

$j = 1, 2, \dots, N; k = R, G, B$ 必须同时成立。

如果 ③ 为 0, 则 ① 与 ② 的 R, G, B 灰度值一致, 表示第 j 帧图像的 (x, y) 处像素与第 $j-1$ 帧图像的 (x, y) 处像素属于同一平稳区间。如果 ③ 为 1, 则 ① 与 ② 的 R, G, B 灰度值至少有一个不一致, 表示第 j 帧图像的 (x, y) 处像素与第 $j-1$ 帧图像的 (x, y) 处像素不属于同一平稳区间。

步骤 2 计算各平稳区间的 R, G, B 三个颜色分量的平均灰度值, 并且进行相似的平均灰度值归类, 称为一致区间 (注: $k = R, G, B$ 同时成立)。

假设一共得到 p 个平稳区间, 各平稳区间所含像素记为 $(m_1, m_2, m_3, \dots, m_p)$ 。则第 s 个平稳区间的 R, G, B 平均灰度值为:

$$\text{ave}_s^k(x, y) = \frac{\sum_{j=q}^{q+m_s} f_j^k(x, y)}{m_s} \quad \begin{cases} s = 1 & q = 0 \\ s = 2, 3, \dots, p & q = m_1 + m_2 + \dots + m_{s-1} \end{cases} \quad (2)$$

k 分别取 R, G, B 来计算

比较所得的 p 个平稳区间的 R, G, B 平均灰度值, 选出 R, G, B 灰度值都一致的全部区间, 并且进行合并。判断第 i 区间和第 j 区间是否为相似平稳区间 (注: $k = R, G, B$ 必须同时成立, 并且 $i \neq j; i, j$ 分别为 $1, 2, 3, \dots, p$):

$$| \text{ave}_i^k(x, y) - \text{ave}_j^k(x, y) | < \varepsilon/2 \quad (3)$$

设第 i 个平稳区间与第 j 个平稳区间合并后组成第 Φ 个一致区间, 则第 Φ 个一致区间的 R, G, B 平均灰度值为: 令 $k = R, G, B$, 则

$$\text{ave}_\Phi^k(x, y) = \frac{\text{ave}_i^k(x, y) \times m_i + \text{ave}_j^k(x, y) \times m_j}{m_i + m_j} \quad (4)$$

利用式 (3)、(4) 将所有的平均灰度值相近的灰度平稳区间归为一类, 统计出该区间像素点出现的频率, 并选出频率最大的, 记为 Max 。

步骤 3 划分像素块, 进行背景模板初始化。

将场景 P_t 按从左到右、从上到下的空间顺序分为一系列 $K \times K (K \geq 8)$ 大小的图像子块, 如图 1 所示。

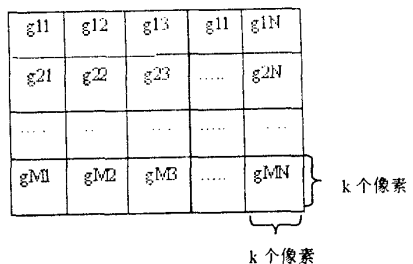


图 1 像素块的划分

设 $\text{Background}_{m,n}^R(x, y), \text{Background}_{m,n}^G(x, y), \text{Background}_{m,n}^B(x, y)$ 分别表示背景模型中第 m 行、第 n 列像素块的第 (x, y) 像素点的 R, G, B 值, 其中第 (x, y) 像素点表示该像素块中第 x 行、第 y 列的像素点。

对背景模板进行初始化:

$$\text{Background}_{m,n}^R(x, y) = \text{ave}_{\text{Max}}^R(x, y) \quad (5)$$

$$\text{Background}_{m,n}^G(x, y) = \text{ave}_{\text{Max}}^G(x, y) \quad (6)$$

$$\text{Background}_{m,n}^B(x, y) = \text{ave}_{\text{Max}}^B(x, y) \quad (7)$$

实验结果如图 2 所示。

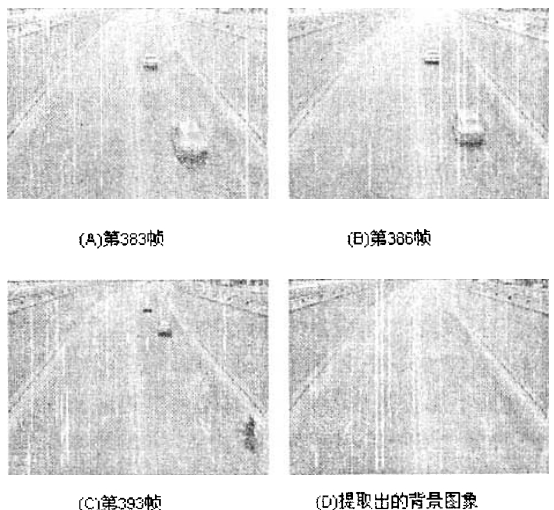


图 2 提取背景

2 更新背景模型

由于交通信息采集系统运行于室外, 而室外的交通静态场景会随着光照强度和方向的变化而变化, 因此系统必须能适应静态场景的缓慢变化, 同时必须具

有一定的抗干扰和抗震能力。

仔细分析发现,运动目标和自然场景变化的速度不同,自然场景变化是一种缓慢变化,而运动目标所引起的图像变化是一种快速变化,因此可以通过低通滤波的方式来阻止快变图像对背景更新的干扰,而光照随时间的缓慢变化则能顺利通过低通滤波器,促使背景的相应更新。在这里采取了一种自适应的背景更新方法,使背景每隔一定时间间隔就更新一次,以达到适应光线变化,理想地分割运动目标。

文中是利用视频序列的当前背景和当前帧的加权平均值来更新背景的^[5,6]。当前帧可能包括前景目标,所以需要先把像素分为前景和背景,然后仅用当前帧的背景像素来更新当前背景,以适应光线等环境因素的变化。如果当前帧中没有运动目标,就用当前帧图像直接对背景模板进行更新,以适应光线的变化;如果当前帧中含有运动目标,就维持目前的背景。设

$$P_{t(m,n)}^r(x,y), P_{t(x,y)}^g(x,y), P_{t(m,n)}^b(x,y)$$

为第 t 帧视频第 m 行、第 n 列像素块的第 (x,y) 像素点的 R,G,B 值,其中第 (x,y) 像素点表示该像素块中第 x 行、第 y 列的像素点。

$$\text{令 } P_{t(m,n)}^k(x,y) \quad k = R, G, B$$

$$\text{Background}_{m,n}^k(x,y) \quad k = R, G, B$$

分别为当前(第 t 帧)某像素的 R,G,B 值和背景模板该像素点的 R,G,B 值。更新背景模板依据下式所示:

$$\text{Background}_{m,n}^k(x,y) = a * P_{t(m,n)}^k(x,y) + (1-a) * \text{Background}_{m,n}^k(x,y) \quad (8)$$

其中, a 的取值直接影响更新的背景质量,若 a 的取值不当甚至会把当前帧图像中的背景误判为前景目标,同时还会使更新速度变慢。在这里经过试验 a 取 0.05 效果比较好。

3 运动目标识别

所谓的运动目标也就是不符合背景判断条件的部分,称之为前景,与背景模型比较起来,前景属于突发事件,其颜色信息和背景的颜色信息都有着巨大的差别。在视频序列中,当有新的帧到来时,系统都会用已建立的背景模型去拟合,逐一比较每个像素点的三个分量 R,G,B 的值,通过每点像素的 R,G,B 值,来判断该点是背景还是运动物体。其中符合背景判断条件的点认为是背景,并且用后续的背景更新算法去更新当前

背景,更新过的背景作为当前背景模型,用在下一帧的处理;而不符合背景判断的点的集合就是前景,也就是上述提到的运动目标。

在实际应用中,还需要对运动目标进行具体的识别,判断它到底是机动车、非机动车,还是行人,也就是所谓的人车识别。

在摄像机镜头固定的情况下,可以通过设置检测区,进行人车识别。方法如下:

步骤1 通过检测到的目标的宽度值,对人车进行粗分类。

$$f_w = \begin{cases} 1 & \text{Width} > T_w \\ 0 & \text{else} \end{cases} \quad (9)$$

T_w 是宽度阈值,如果目标宽度大于 T_w 说明目标是比较大的物体,可能是车辆,也可能是人群,或者是横穿马路的自行车等。如果目标宽度小于 T_w 说明是小目标,可能是人、自行车,或者是摩托车。

步骤2 使用形态学算子紧密度和体态比对人车进行识别。定义紧密度和体态比为:

$$\text{Compactness} = \frac{\text{Area}}{\text{Perimeter}^2} \quad (10)$$

$$\text{Aspect Ratio} = \frac{\text{Height}}{\text{Width}} \quad (11)$$

人群和车辆,两者最大的区别在于紧密度,人群的紧密度要比汽车的小得多,人群的紧密度和汽车的紧密度典型值分别为 0.0144 和 0.0577。可以通过构造一个基于面积—紧密度的线性分割函数 $G(\text{Area}, \text{Compactness})$ 就能将人群和车辆区分开。

人、自行车、摩托车最大区别在于体态比,人的体态比最大,其次是自行车,最小的是摩托车,它们的体态比典型值分别为 11.11, 9.53, 5.76。可以通过构造一个基于面积——体态比平方的线性分割函数 $G(\text{Area}, R^2)$ (R 表示目标的体态比)就能将人群和车辆区分开。

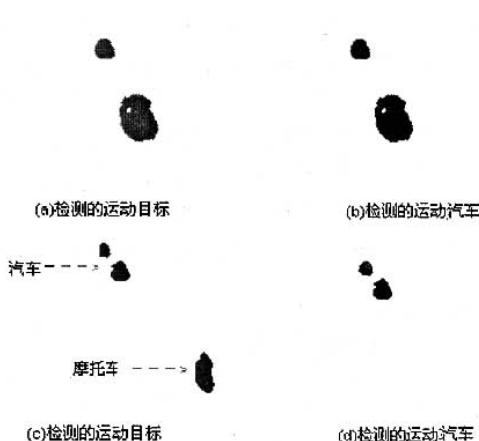


图3 视频序列的目标检测

(下转第26页)

$$K_D = \frac{(\sigma^2 + 4\sigma\xi_p\omega_0 + \omega_0^2)t^2\alpha - 4(t + \alpha)}{4k}$$

3 仿真研究

对于一阶延时系统 $G(s) = \frac{3e^{-s}}{1+s}$, $\xi_p = 0.7$, 通过上述的数值优化方法进行计算可得到控制器的最优参数 $K_P = 0.2820$, $K_I = 0.2336$, $K_D = 0.0834$ 。为了比较该方法的有效性, 另外采用变速积分 PID 控制算法 ($K_P = 4.5$, $K_I = 5.0$, $K_D = 0.055$, $A = 0.4$, $B = 0.6$)^[7] 对上述控制对象仿真, 同样采用阶跃响应来做测试。其仿真结果分别如图 2、图 3 所示。

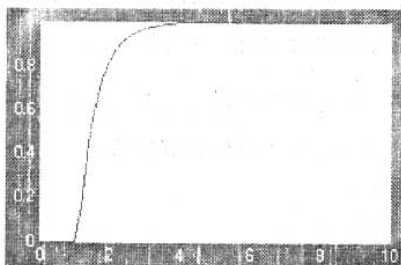


图 2 数值优化 PID 阶跃响应图

由图中可以看出, 数值优化 PID 控制与变速积分 PID 控制一样有效, 其响应性能指标甚至优于变速积分 PID 控制。

4 结论

介绍了一种经过数值优化改进的 PID 控制器, 详细说明了其设计原理与方法, 并做了仿真研究。其与变速积分 PID 控制器的仿真结果的对比表明了该控制器是有效的。文中笔者创新点: 利用数值优化方法

确定最佳的 PID 控制参数, 并给出仿真示例。

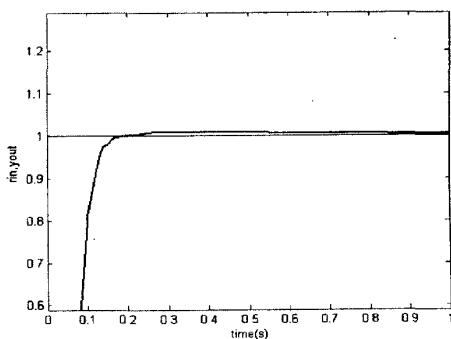


图 3 变速积分 PID 阶跃响应图

参考文献:

- [1] 吴宏鑫, 沈少萍. PID 控制的应用与理论依据[J]. 控制工程, 2003, 10(1): 37-42.
- [2] 刘红军, 韩 璞, 孙海蓉, 等. 基于内模控制的 PID 控制器的设计[J]. 计算机仿真, 2004, 22(3): 207-210.
- [3] 侯勇严, 郭文强. 单神经元自适应 PID 控制器设计方法研究[J]. 微计算机信息, 2005, 21(8): 8-10.
- [4] Muske K R, Badgwell T A. Disturbance modeling for offset-free linear model predictive control[J]. Journal of Process Control, 2002, 12: 617-632.
- [5] Pannocchia G, Laachi N, Rawlings J B. A Candidate to Replace PID Control: SISO - Constrained LQ Control[J]. Journal of Process Control, 2005, 15: 81-88.
- [6] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.
- [7] 刘金琨. 先进 PID 控制 MATLAB 仿真[M]. 北京: 电子工业出版社, 2004.

(上接第 23 页)

经过上面两步, 就可以准确地分类识别出运动目标。实验结果如图 3 所示。

4 结束语

文中提出的基于自适应背景模型运动目标检测方法, 首先初始化背景模型; 然后, 进行模型更新, 运动目标分割; 最后准确识别。得出了比较令人满意的实验结果。但它有自身的缺点, 就是必须满足交道路面被运动目标覆盖的时间占总时间的百分比比较小。

参考文献:

- [1] LONG W, YANG Y. Stationary background generation: An alternative to the difference of two images[J]. Pattern Recognition, 1990, 23(12): 1351-1359.

- [2] Ridder C, Munkelt O, Kirchner H. Adaptive background estimation and foreground detection using Kalman-filter[C]// Proc. of International Conf. on Recent Advances in Mechatronics, ICRAM95, UNESCO chair on Mechatronics. [s. l.]: [s. n.], 1995: 193-199.
- [3] 杨枝灵, 王 开. Visual C++ 数字图像获取、处理及实践应用[M]. 北京: 人民邮电出版社, 2003.
- [4] 王夏黎, 周明全, 耿国华, 等. 交通流视频检测系统的设计与实现[J]. 计算机软件, 2004, 21(9): 68-71.
- [5] 周西汉, 刘 勃, 周荷琴. 一种基于奔腾 SIMD 指令的快速背景提取方法[J]. 计算机工程与应用, 2004, 40(27): 81-84.
- [6] 刘 勃, 周荷琴. 一种基于颜色和运动信息的夜间交通检测方法[J]. 中国图像图形学报, 2005, 12(2): 187-191.