

基于视频序列的目标检测与跟踪技术研究

刘翔, 吴谨, 祝愿博, 康晓晶

(武汉科技大学信息科学与工程学院, 湖北武汉, 430081)

摘要:针对计算机智能监控环境,文中提出一种改进的基于像素灰度出现概率最大值的背景建立方法,该方法克服了光照变化对背景重建的影响,使得背景建立的时间大大缩短。并采用一种新的自适应背景更新算法获得背景图像以进行目标检测,这种方法较好地克服了IIR法更新速度难以取值的缺点,使得更新速率可以达到自适应的效果;在目标跟踪阶段,使用基于卡尔曼滤波的方法对检测出的运动目标进行跟踪,由于卡尔曼预测可以大大减小特征匹配的搜索范围,因此提高了跟踪的实时性。实验结果表明,该文的算法能够快速有效地获得、更新背景,并且能够实时地对运动目标进行跟踪。

关键词:背景模型;目标检测;目标跟踪

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2009)11-0179-04

A Study of Object Detecting and Tracking Based on Video Sequences

LIU Xiang, WU Jin, ZHU Yuan-bo, KANG Xiao-jing

(School of Info. Science and Engineering, Wuhan Univ. of Sci. and Techn., Wuhan 430081, China)

Abstract:This thesis was focused on intelligent surveillance. It proposed an improved method of building background, which is based on the maximum probability value of the pixel. This method overcame the light change on the background of the impact of the reconstruction, making the background of the establishment of the time significantly shortened. Then, adopted a new self-adaptive background updating algorithm to gain a background image for the purpose of target detecting, the new method was proved better than IIR method, for it could obtain an adaptive effect on updating rates; In order to track the target detected, the Kalman filtering algorithm was used. The results showed that this algorithm could obtain and update background image in a short time. It also could track the detected target in real-time.

Key words:background model; target detecting; target tracking

计算机智能监控目前已经成为计算机视觉领域的一个倍受关注的前沿研究方向。运动目标的检测与跟踪技术是智能监控的核心研究课题,同时也是各种后续高级处理的基础。因此,做好这方面的研究工作有十分重大的意义。

1 运动目标的检测

1.1 运动检测常用方法概述

常用的运动检测的经典方法有:帧间相差法^[1]、光流法^[2]和背景相减法^[3]。

1)光流法。

光流法由 B. K. P. Horn 和 B. G. Schunck 于 1980 年首先提出。光流计算基于如下两个假设:

(1)任何物体点所观察到的亮度随时间是恒定不变的。

(2)图像平面内邻近点以类似方式进行移动。

根据这两个假设,可以得到光流法的基本方程。设 t 时刻图像上点 (x, y) 处的灰度记为: $I(x, y, t)$, 在 $t + \Delta t$ 时刻,该点的灰度级记为: $I(x + \Delta x, y + \Delta y, t + \Delta t)$ 。由假设可知: $I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$, 将右边的式子在 (x, y, t) 处用泰勒公式展开,化简并且略去二次项得:

$$\frac{\partial I}{\partial x} \cdot \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \cdot \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0 \quad (1)$$

令: $u(x, y, t) = \frac{\Delta x}{\Delta t} = \frac{dx}{dt}$, $v(x, y, t) = \frac{\Delta y}{\Delta t} = \frac{dy}{dt}$ 。

则有:

$$\frac{\partial I}{\partial x} \cdot u + \frac{\partial I}{\partial y} \cdot v + \frac{\partial I}{\partial t} = 0 \quad (2)$$

称 (u, v) 为光流场,式(2)就是光流场的基本方程。

由于光流中携带了物体的运动信息和景物的三维信息,因此它能够检测独立运动的对象,而且不需要预

收稿日期:2009-03-14;修回日期:2009-06-15

基金项目:湖北省自然科学基金(2008CDB311)

作者简介:刘翔(1983-),男,硕士研究生,研究方向为图像处理与模式识别研究;吴谨,教授,博士生导师,研究方向为图像处理与模式识别研究。

先知道场景的任何信息,并且可用于摄像机运动的情况。但是光流法的计算方法相当复杂,它需要对所有帧中的像素进行计算,计算量大,如果没有专门的硬件支持,很难实时检测。

2) 帧间相差法。

相邻帧差法选取时间间隔相邻的两帧图像相减来检测运动目标,两帧图像之差不为零处表示有运动发生。设 t 时刻图像中 (x, y) 处的灰度为: $I(x, y, t)$, 下一帧图像在 (x, y) 处的灰度为: $I(x, y, t + 1)$, 记这两帧图像之差为: $D(x, y)$ 。则:

$$D(x, y) = I(x, y, t + 1) - I(x, y, t) \quad (3)$$

如果该处没有发生运动,则: $D(x, y) = 0$, 如果该点处有运动发生,则: $D(x, y) \neq 0$ 。取阈值 Th 将差分图像二值化:

$$D(x, y) = \begin{cases} 1, & |D(x, y)| \geq Th \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

该二值图像中包含了目标的运动信息。

与前面介绍的光流法相比较,相邻帧差法的计算量少得多,可以满足实时性的要求。由于连续两帧图像之间的间隔时间很短,这种方法对光照条件也并不敏感。但是,采用这种方法,检测出的只是较小时间间隔内图像变化的部分,而且会使目标内部产生“空洞”。在某些极端情况下,如目标物体的运动速度极慢,采用这种方法就可能根本检测不出这些缓慢运动的物体。

3) 背景相减法^[4]。

背景相减法选取参考图像作为背景图像,并用当前图像与背景图像相减来检测运动目标。设参考图为: $I_1(x, y)$, 一幅包含有运动目标的图像为: $I_2(x, y)$, 记两图像之差为: $\Delta I(x, y)$ 。则:

$$\Delta I(x, y) = I_2(x, y) - I_1(x, y) \quad (5)$$

取阈值 Th 将二值化:

$$\Delta I(x, y) = \begin{cases} 1, & |\Delta I(x, y)| \geq Th \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

这样得到的二值图像比起使用相邻帧差法得到的二值图像来说,能够提供比较完整的特征数据;而比起光流法,这种方法同样计算量较小,可以满足实时性的要求。

1.2 背景建立方法

背景相减法是目前使用最多的目标检测方法。如何获得一副完整、“干净”的背景图像一直以来都是运动检测的研究热点。背景模型的建立是背景相减法目标检测的关键。最简单的方法是直接选取不含运动目标的场景图像作为背景,这是一种理想的情况,但是在多数场合下是不现实的,而且这种方法缺乏灵活性,难以适应场景中光照等情况的变化。一种解决这一问

题的经典方法是对一段时间内的连续多帧图像求平均作为近似的背景图像,但是使用这种方法还是有可能将运动目标混入其中,而且这种方法的实时性比较差,有时为了获得较好的背景图像需要将很长一段时间的多帧图像做平均。目前对于背景模型建立的研究主要可分为基于模型的方法和基于重构的方法^[5]。文中在总结以上两种背景模型建立方法的基础上提出一种改进的基于像素灰度出现最大概率的方法建立背景。

1.3 基于像素灰度出现概率最大值的背景建立方法及其改进方法

选取连续多帧图像中与背景图像相关的像素灰度用于建立背景模型。这种方法的基本思想是:在一段时间内,某一点的灰度值出现的次数最多,则认为该点就是属于背景的。笔者选取 $[0, T]$ 这段时间内的连续多帧图像用于建立背景。设 $f_n(x_0, y_0)$ 为第 n 帧图像中 $(0, 0)$ 点的灰度,其中 $n = 0, 1, 2, \dots, T$ 。统计 T 时间内 $(0, 0)$ 点处某灰度出现的频率,记为: $p_n(x_0, y_0)$, 那么在 $(0, 0)$ 处的原始背景估计为: $B_0(x, y) = \max(p_n(x_0, y_0))$ 。其他各点处的原始背景估计可以用类似的方法求得。该方法原理简单,也比较有效,但是由于需要对每个像素的变化都要做统计因此比较麻烦。尤其是背景中的点往往受光照的变化发生微小的变化,这样会大大影响统计的时间。因此文中对这种方法做了一些改进,具体做法如下:

将 $0 \sim 255$ 的灰度区间划分为 N 个区间, N 取 $10 \sim 25$ 均能够正确地重构图像背景。统计各帧中 $(0, 0)$ 属于的灰度区间出现的概率 $h_n(x_0, y_0)$, 那么在 $(0, 0)$ 处的原始背景估计为该灰度区间内的灰度的均值,即: $B_0(x, y) = \bar{f}_m(x, y)$, 其中, $f_m(x, y)$ 是取 $\max(h_n(x_0, y_0))$ 对应的那个区间中的各灰度。 $f_m(x, y)$ 上面的横线表示取平均,其他各点处的原始背景估计可以用类似的方法求得。

1.4 背景图像的更新

建立好的背景图像由于受到场景中光照等因素的影响不可能是不变的,因此需要在检测过程中对其进行及时的更新。通常使用较多的一种背景更新方法是采用加权平均的方法,对于那些近的图像帧赋予高的权值,这种方法的缺点是不能应付变化较快的场景,可能会导致检测的错误;另外一种常用的方法是采用 IIR 滤波的方法^[6], 即设 $B(x, y)$ 为背景图像, $F(x, y)$ 为当前帧图像, 则 IIR 滤波更新的方法如下式:

$$B_i(x, y) = (1 - \alpha)B_{i-1}(x, y) + \alpha F(x, y) \quad (7)$$

其中 α 是调节更新速度的常数,其取值范围是: $[0, 1]$, α 越大,则更新的速度越快;反之,则越小,通常

$\alpha = 0.05$ 。这种方法的缺点是 α 是一个经验值,且图像中各像素点的更新权重都一样,但在实际中各像素点的更新速率并不一样。

因此,该文提出一种能够自适应调节更新速率的算法来解决这个问题。如果当前输入帧中某位置上的像素灰度和上一帧背景图像对应位置的灰度之间的绝对差较小,则说明该像素可能是背景,因此更新的速度就要快一些;相反的,如果这个差值比较大,就可以适当放慢更新的速率。设 $B(x, y, t - 1)$ 是上一帧得到的背景模型, $F(x, y, t)$ 是当前帧。令 $\beta = |F(x, y) - B(x, y, t - 1)|$, 根据上面的分析,得到自适应的背景更新计算式为:

$$B(x, y, t) = \begin{cases} (1 - 1/\sqrt{\beta})B(x, y, t - 1) + \\ 1/\sqrt{\beta} \cdot F(x, y) & (\beta \neq 0) \\ B(x, y, t - 1) & (\beta = 0) \end{cases} \quad (8)$$

这样通过差值 β 的大小来自动调节更新速率的变化达到背景更新自适应的目的。

2 运动目标的跟踪

2.1 运动目标跟踪算法分类

传统的跟踪方法大致可以分为三种^[7-9]:

第一种是基于相关的跟踪方法,它的主要思想就是在当前帧中寻找与上一帧中目标区域相关性最大的区域,具体做法就是系统的基准图像在实时图像上以不同的偏移值位移,然后根据一定的相似度量准则在每一个偏移量下重叠两幅图像,基准图像和与基准图像同样大小的实时图像进行相关处理,根据判别准则和相关处理结果,判断目标在实时图像中的位置;

第二种是基于特征的跟踪方法,为了实现跟踪并不需要跟踪目标的每一点,只需跟踪目标上的某一特征点就可以实现对整个目标的跟踪。这个特征点可以是目标的质心,也可以是目标上的任意一点,但前提条件是必须保证这一点的高度稳定性,即不易受到外界因素如光照强度变化、噪声等因素的干扰;

第三种是基于模型的跟踪方法,该方法利用点、线、区域把被跟踪的目标拟合成一个几何模型,运动目标的跟踪变成目标识别问题。这种方法包含高层的语义描述和知识。因此,与其他跟踪方法相比,它的可靠性更强;它的缺点是计算量比较大,因此它需要耗费大量的时间来计算复杂的模型,以处理平移、旋转和变形等,而且需要知道大量关于所要跟踪目标的先验知识。

2.2 基于卡尔曼预测的运动目标跟踪

Kalman 滤波常用于运动估计。将有关位置、速

度、形状、纹理、颜色等特征值看作状态值,先根据前一个时刻的状态值进行预测,然后根据预测值和误差协方差确定搜索范围,在该范围之内查找特征并进行匹配,得到状态的测量值。卡尔曼滤波器是一套线性无偏最小均方误差的递推公式。可以证明,在一定条件下,在最小均方误差准则下得到的最佳线性系统是所有系统中的最佳者^[10]。

形式上,有模型:

$$x_{k+1} = A_k x_k + w_k \quad (9)$$

$$Z_k = H_k x_k + v_k \quad (10)$$

式(9)称为状态方程,式(10)称为量测方程。矩阵 A_k 描述了模型状态的演化, w_k 是均值为零的高斯噪声,并且 w_k 具有协方差 Q_k :

$$Q_k = E[w_k w_k^T] \quad (11)$$

x_k 为特征向量, Z_k 为观测向量。矩阵 H_k 是测量矩阵,它描述了观察是怎样和模型关联; v_k 是另一个均值为零的高斯噪声因子,其协方差 R_k :

$$R_k = E[v_k v_k^T] \quad (12)$$

$$\text{先验估计: } \hat{x}_k^- = A_{k-1} x_{k-1} \quad (13)$$

定义先验估计与后验估计误差如下:

$$e_k^- = x_k - \hat{x}_k^- \quad (14)$$

$$e_k^+ = x_k - \hat{x}_k^+ \quad (15)$$

其协方差分别为 P_k^- 与 P_k^+ 。其中:

$$P_k^- = A_k P_{k-1}^- A_k^T + Q_{k-1} \quad (16)$$

$$P_k^+ = (I - K_k H_k) P_k^- \quad (17)$$

式(17)中的 K_k 称为卡尔曼增益,它以最小化后验协方差 P_k^+ 求得。

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (18)$$

后验估计(即对先验估计用卡尔曼增益进行修正的结果):

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (z_k - H_k \hat{x}_k^-) \quad (19)$$

如果给定初始条件 \hat{x}_0^- 与 P_0^- 。那么,整个预测与修正的过程可以表示如图 1 所示。

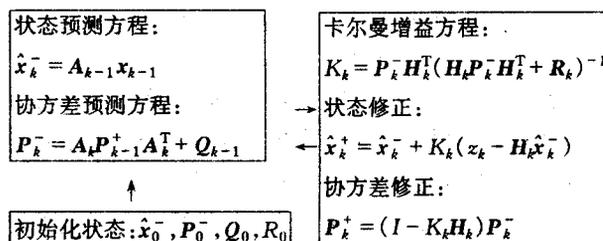


图 1 卡尔曼滤波预测与修正的迭代过程

2.3 运动目标跟踪算法

基于卡尔曼预测的跟踪关键是要找出最佳的相关匹配点。可以在由上一帧图像预测出的位置附近搜索

得到。然后将最佳匹配点作为当前帧的精确位置,修正预测值启动对下一帧的预测,如此反复迭代实现目标的跟踪。整个跟踪过程的算法如下:

第一步:采集视频中初始一段时间内的图像,对图像进行预处理,消除图像中的部分噪声。

第二步:使用文中提出的背景重建方法建立初始背景。

第三步:使用背景相减法检测运动目标,对检测的运动目标进行形态学处理、连通性分析,去掉那些面积较小的非目标噪声干扰。

第四步:设置初始状态,根据 Kalman 滤波器预测出下一帧的状态向量,在预测位置附近找出最佳匹配点,利用状态修正方程修正预测值。

第五步:将修正后的预测值带入状态预测方程中继续跟踪后续帧。

第六步:采集后一帧图像,使用文章的背景更新算法更新背景,重复第三、四、五、六步,直至结束。注意:当从第二次执行到第四步时要使用上次的修正值作为下一次预测器的输入。

3 算法实现与实验结果

该文使用的实验 PC 机的配置为:CPU 主频 2.4GHz,512M 内存,编程环境使用 Visual C++ 6.0。实验结果如图 2~图 5 所示:



图 2 初始序列



图 3 背景图像



图 4 运动目标



图 5 目标跟踪

图 2 是场景的初始状态,从中可以看出场景中一开始就含有运动目标;图 3 是对采集图像进行高斯平滑预处理后采用文中提出的背景重构方法建立起来的背景图像;图 4 是采用背景相减法后经过形态学处理^[11]后检测出来的运动目标;图 5 是采用文中使用的跟踪算法得到的一部分结果,图中检测出的运动目标用矩形框外接显示。

表 1 中给出了跟踪实验中其中连续 5 帧的卡尔曼预测与修正结果:

表 1 实验中卡尔曼预测与修正结果

帧数	卡尔曼预测位置	最佳匹配位置	修正后的预测值
163	(143,68)	(139,68)	(140,68)
164	(142,68)	(140,68)	(141,68)
165	(143,68)	(139,69)	(141,69)
166	(143,69)	(138,69)	(140,69)
167	(142,69)	(138,69)	(139,69)

由表 1 中的数据知,修正后的预测值与实际值的误差减小了。这样不断地修正,保证了跟踪的稳定性。

4 结束语

该文提出了一种改进的基于灰度出现概率最大值的背景重构方法,使用背景相减的方法检测视频序列中的运动目标;对检测出的目标进行基于卡尔曼预测的跟踪,实验证明该方法对于室内环境下的运动目标检测与跟踪有较好的效果。

参考文献:

- [1] Lipton A, Fujiyoshi H, Patil R. Moving target classification and tracking from real-time video[C]//In: Proc IEEE. Workgroup on Applications of Computer Vision. Princeton, NJ:[s. n.],1998:8-14.
- [2] Horn B K P, Schunck B G. Determining optical flow[J]. Artificial Intelligence,1981,17(1-3):185-203.
- [3] Haritaoglu I, David H, Davis L S. Real-time surveillance of people and their activities[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(8):809-830.
- [4] 侯伟,卢炎麟,郑河荣,等.固定背景下的视频分割及在交通视频流的应用[J].计算机技术与发展,2007,17(9):191-193.
- [5] 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.
- [6] 杨俊红,张强,周兵.视频序列中的运动目标检测[J].微计算机信息,2007,23(7-1):226-227.
- [7] Comaniciu D, Ramesh V, Meer P. Kernel-based Object Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2003,25(5):564-577.
- [8] Hager G D, Dewan M, Stewart C V. Multiple Kernel Tracking with SSD[C]//In: Proc IEEE Conference on Computer Vision and Pattern Recognition. [s. l.]: [s. n.],2004:790-797.
- [9] Dai Y P, Yu G H, Hirasawa K. New Development on Tracking Algorithm with Derivation Measurement[C]//In: Proc IEEE International Conference on System, Man and Cybernetics. [s. l.]: [s. n.],2001:3181-3186.
- [10] Haykin, Simon S. Adaptive filter theory[M]. Beijing: Publishing House of Electronics Industry,2002.
- [11] Gonzalez R C, Woods R E. 数字图像处理[M].第2版.阮秋琦,阮宇智译.北京:电子工业出版社,2004:420-453.