

一种实时的主动跟踪方法^{*})

胡伏原¹ 张艳宁¹ 潘进² 袁金刚¹

(西北工业大学计算机学院 西安 710072)¹ (西安通信学院网络安全与对抗研究室 西安 710106)²

摘要 针对单视角像机的视野局限性以及连续跟踪特定目标时运动目标实时检测的困难,提出了一种离散化活动像机运动的方法。先离散化像机运动,建立像机预置位表和背景索引表;然后,利用目标位置信息和运动信息判断像机转动情况,并结合对应的控制原理,实现了像机的调度;最后,利用通过计算机视觉方法确定了特定目标在不同离散空间的对应关系,实现了对特定目标大范围的主动跟踪和运动目标的实时检测。实验结果表明,对于复杂场景中特定运动目标大范围的主动跟踪具有较好的鲁棒性和实时性,并且可以实时提取场景中的运动区域。

关键词 主动跟踪,活动像机,运动目标检测

Adaptive Visual Object Tracking in Video

HU Fu-Yuan¹ ZHANG Yan-Ning¹ PAN Jin² YUAN Jin-Gang¹

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072)¹

(Lab of Network Security and Countermeasure, Xi'an Communications Institute, Xi'an 710106)²

Abstract This paper presents a novel way for object manipulation using an active camera with pan and tilt control to overcome disadvantages of single camera view and the difficulties of real-time detect moving objects in active camera. There are three stages. The first stage is preprocessing to discretize the actively moving camera according to the surveillant scene and create an index table of camera pose, camera parameters and background, which resolve the difficulty of moving objects real-time detection because of moving camera. Then, the desired viewing angle from objects of interest is allotted by using object position and moving information, and the active camera is rotated by corresponding control theory. Finally, location of moving object in different camera pose is based on the computer vision with real-time and accurate capabilities, which adaptively tracks objects of interest in different camera pose. Results on real image sequences illustrate the real-time and robust of the proposed method in large scene.

Keywords Adaptive visual object tracking, Active camera, Moving object detection, Active camera discretization

1 引言

通过计算机视觉实现对运动目标的检测和跟踪在智能监控、虚拟现实、运动分析、高级用户接口和基于模型的编码等方面有着广泛的应用^[1]。传统的视觉分析处理都采用单目视觉,而单个摄像机在固定视角的观察范围和目标清晰度有限,运动目标的跟踪严格受制于像机的安装角度和可视范围。为了确保对可疑目标或者感兴趣目标的大视野、大场景的跟踪,在不增加摄像头的基础上,利用活动像机实现对特定目标的主动跟踪是很重要的解决手段。它是计算机视觉中的一个重要研究领域,同时是智能监控系统的一个非常重要的组成部分。

采用活动像机实时地跟随目标运动,让目标尽可能处于像机视野中。但是,当像机运动的时候,实现目标检测一般是通过图像稳定,然后利用静止背景下运动目标检测方法或者通过光流方法进行运动目标检测,通过软件基本上不可能实时实现。因此,本文提出一种可以实现实时检测目标的主动跟踪方法。该方法主要是通过离散化像机,当目标处于当前像机视野的边缘位置时,摄像机快速转动到对应位置,让目标重新出现在像机的视野中。然而,由于像机在快速运动的时候图像会产生运动模糊,给目标的跟踪带来极大的困难。本

文提出借助计算机视觉的办法,找到目标在像机转动前后的图像坐标系中位置对应关系,这样就可以很好地实现像机转动前后目标的快速匹配和再跟踪,并且精度高、速度快。对于像机的转动,本文提出利用目标的运动速度和位置信息,判别像机的转动信息,快速稳健地实现了像机的调度。主要有3个关键问题:

(1)预处理。主要是为主动跟踪做准备,从而可以实现对运动目标的实时检测。

(2)运动目标检测和主动跟踪。主要是实现目标的检测和对特定目标的跟踪。这里主要难点是像机运动前后的特定目标跟踪问题。

(3)摄像机调度。主要确定什么时候转动像机和怎样转动像机。

本文将依次对这3部分内容进行阐述,然后给出本文设计的基于计算机视觉的主动跟踪方法,最后给出实验结果。

2 预处理

为了能在主动跟踪的同时实现对目标的实时检测,本文提出离散化像机转动,实现对特定目标的跟踪。因此,在主动跟踪前,必须先离散化像机。并且,为了后续、实时准确地检测和跟踪,本文还需要实现背景图像的获取和像机在各个状

^{*})博士点基金和西北工业大学博士创新基金(CX200418)资助。胡伏原 博士,主要研究方向为视频处理、计算机视觉等。

态时候的内外参数求取。

本文预处理主要是离散化像机运动,并对整个场景建立像机转动索引、像机参数索引和背景索引。通过离散化像机运动,像机大部分时间处于静止状态,便于实现运动目标的实时检测。主要实现过程如下:

先向摄像机发复位命令,然后控制摄像机从初始角度开始,从左到右、从上到下对像机进行水平旋转(pan)和俯仰(tilt)运动,每隔一定角度设置预置位,直到像机的最大视野或需要监控的场所被记录完毕,最终形成一个预置位表。

接下来,对每个预置位,采集一幅背景图像,并标记对应的索引号。也可以对每个预置位所在的场景进行统计分析,比如混合高斯模型^[4],建立背景。然而,对每个预置位进行统计建模,需要较长的时间,对于开始建立的模型可能由于光照等变化而产生不准确。因此,本文选择了差分 and 简单的背景模型相结合的方法在像机静止的情况对运动目标进行提取。

最后,对每个预置位,本文利用简易定标板实现摄像机内参数定标^[2],利用角锥体后方交会法^[3]求解像机外参数。像机标定后,形成对应的像机参数索引表。

3 运动目标主动跟踪

在复杂背景下,对于运动目标的检测和跟踪是一个很难的任务。本节的目标主要是探讨复杂环境在像机静止的情况下运动目标的检测和对于特定目标的跟踪问题。

静止背景下运动目标的检测主要有瞬时差分、背景提取和光流法。本文主要采用三帧差分 and 自适应背景提取相结合^[5]的方法精确检测运动目标。跟踪方面,本系统主要有两个问题:静止背景下特定目标跟踪和像机运动前后特定目标的跟踪。静止背景下目标的跟踪常用的方法有 Kalman、MeanShift 和 Condensation,本文采用 Kalman 和 MeanShift 相结合的方法,并通过 Kalman 预测解决遮挡问题;对于像机运动前后特定目标的跟踪,本文提出采用计算机视觉的方法对不同像机姿态下的目标进行对应图像坐标系的匹配方法,即目标位置映射,解决像机转动前后目标跟踪问题。本节主

$$\begin{cases} \mu_{t,m} = R_m^{-1} * s_m^{-1} * (s_n * R_n * (\mu_{t,n} - \mu_{0,n}, v_{t,n} - v_{0,n}, -f_n)^T + T_n - T_m) + \mu_{0,m} \\ v_{t,m} = R_m^{-1} * s_m^{-1} * (s_n * R_n * (\mu_{t,n} - \mu_{0,n}, v_{t,n} - v_{0,n}, -f_n)^T + T_n - T_m) + v_{0,m} \end{cases} \quad (5)$$

对于相同像机在不同姿态,可以近似认为内参、尺度因子不发生变化;对于光心位置在云台的旋转轴上,可以近似认为平移矩阵不变。式(5)可以简写为

$$\begin{cases} \mu_{t,m} = R_m^{-1} * R_n * (\mu_{t,n} - \mu_{0,n}, v_{t,n} - v_{0,n}, -f_n)^T + \mu_{0,m} \\ v_{t,m} = R_m^{-1} * R_n * (\mu_{t,n} - \mu_{0,n}, v_{t,n} - v_{0,n}, -f_n)^T + v_{0,m} \end{cases} \quad (6)$$

对于 Mean-Shift 对应的输入点的位置要求不是特别的准确,因此可以用式(6)计算目标在同一时刻不同像机姿态下图像坐标系中的对应关系。

3.2 目标参数更新

在像机运动后,由于可以通过计算机视觉的方法计算目标在不同姿态下对应的图像坐标系中目标位置关系,因此更新 Kalman 和 MeanShift 参数就可以实现目标的继续跟踪。

为了减少计算量,以精度换取时间,不用重新映射速度分量,仅仅更新目标的位置信息。对于 kalman 滤波器,重新初始化;对于 Mean-Shift,用 3.1 节中计算的新目标位置点代入输入候选点,继续实现目标跟踪。

3.3 遮挡处理

若目标在 $t_0 \sim t_1$ 时间范围内被背景(如树木、建筑物等)

要就目标位置映射和遮掩问题的解决进行研究。

3.1 目标位置映射

转动前后在各自图像坐标系中目标位置的关系可以通过计算机视觉来解决。假设像机转动前处于预置位 n ,转动后处于预置位 m , t 时刻目标底部中心在预置位 k 中的图像点坐标 $(\mu_{t,k}, v_{t,k})$,在预置位 n 转动到预置位 m 时目标可能的运动位移为 $(\delta_{t,n,x}, \delta_{t,n,y})$,则 $(\mu_{t,n} + \delta_{t,n,x}, v_{t,n} + \delta_{t,n,y})$ 和 $(\mu_{t,m}, v_{t,m})$ 为同一目标在不同图像坐标系下的坐标。为了讨论方便,仅讨论目标在同一时刻不同像机姿态下的对应关系。

对于给定多个预置位均已标定好的摄像机。设预置位 k 摄像机的光心位置在世界坐标系中为 $(X_{0,k}, Y_{0,k}, Z_{0,k})$;目标底部在图像平面中的坐标为 $(\mu_{t,k}, v_{t,k})$,目标底部在世界坐标系中为 $(X_{t,k}, Y_{t,k}, Z_{t,k})$ 。根据摄像机坐标系和图像坐标系之间的转换关系,可以计算出目标像点在摄像机坐标系中的位置坐标为 $(x_{t,c,k}, y_{t,c,k}, z_{t,c,k})$,见式(1)。

$$\begin{cases} x_{t,c,k} = \mu_{t,k} - \mu_{0,k} \\ y_{t,c,k} = v_{t,k} - v_{0,k} \\ z_{t,c,k} = -f \end{cases} \quad (1)$$

其中 $(\mu_{0,k}, v_{0,k})$ 为主点在图像坐标系中的坐标; f 是像机焦距,通过内参标定计算可得。

对于摄像机坐标系中的目标像点映射到世界坐标系中,可以通过旋转矩阵和平移矩阵变换进行,即

$$(X_{t,k}, Y_{t,k}, Z_{t,k})^T = s * R * (x_{t,c,k}, y_{t,c,k}, z_{t,c,k})^T + T \quad (2)$$

其中 s 为尺度因子, R 为旋转矩阵, T 为平移矩阵,均可通过外参标定求得。因此,可以得到像机在姿态 m 和 n 下的目标在各自图像坐标系中的关系:

$$s_m * R_m * (x_{t,c,m}, y_{t,c,m}, z_{t,c,m})^T + T_m = s_n * R_n * (x_{t,c,n}, y_{t,c,n}, z_{t,c,n})^T + T_n \quad (3)$$

即

$$(x_{t,c,m}, y_{t,c,m}, z_{t,c,m})^T = R_m^{-1} * s_m^{-1} * (s_n * R_n * (x_{t,c,n}, y_{t,c,n}, z_{t,c,n})^T + T_n - T_m) \quad (4)$$

由式(1)和(4)可得:

所遮掩,在 $t_0 \sim t_1$ 的图像帧中便跟踪不到目标。则在 t_0 时刻,对目标的运动参数根据前帧状态以及误差参量进行 Kalman 预测,将预测值作为遮掩时的真实值保留下来,成为 $t_0 + 1$ 帧时刻目标的状态参数值。如此继续, $t_1 + 1$ 时刻目标进行 Kalman 预测前的运动参数便是 t_1 时刻的预测值。目标在 $t_1 + 1$ 时刻 MeanShift 跟踪成功,进行状态更新。本文通过对室外场景分别在雨天和雾天进行了目标的检测和跟踪,结果见图 2 所示。

4 像机调度

本文通过特定目标的位置和运动方向信息确定像机可能的转动状态。为了简单起见,优先 x 方向的旋转,见式(8)。

$$M_{\text{camera}} = \begin{cases} 1, & x < T \& V_x < 0 \\ 2, & x < T \& V_x > 0 \\ 3, & y < T \& V_y < 0 \\ 4, & y < T \& V_y > 0 \\ 0, & \text{其他} \end{cases} \quad (7)$$

其中 M_{camera} 表示像机的转动状态:1 表示像机往左转,2 表示像机往右转,3 表示像机向上转,4 表示像机向下转,0 表示像

机不发生偏转。 x, y 分别表示目标的在图像中的位置, T 为给定的阈值。

像机的下一个转动预置位可以通过预置位设置表和像机的转动状态 M_{camera} 确定, 本文假设预置位表的初始记号为 0。像机的下一个转动预置位可通过式(9)求得。

$$P_{i+1} = \begin{cases} \max(P_i - 1, W_{PT}(P_i/W_{PT})), & M_{camera} = 1 \\ \min(P_i + 1, W_{PT}(P_i/W_{PT} + 1) - 1), & M_{camera} = 2 \\ \max(P_i - W_{PT}, 0), & M_{camera} = 3 \\ \min(P_i + W_{PT}, W_{PT}H_{PT}), & M_{camera} = 4 \\ P_i, & M_{camera} = 0 \end{cases} \quad (8)$$

其中, P_{i+1} 为下一个预置位, P_i 为当前预置位, W_{PT} 和 H_{PT} 分别表示预置位表的宽度和高度。

在像机的下一个转动状态确定后, 即可利用云台和对应的协议控制像机转动到对应的预置位。

5 主动跟踪系统

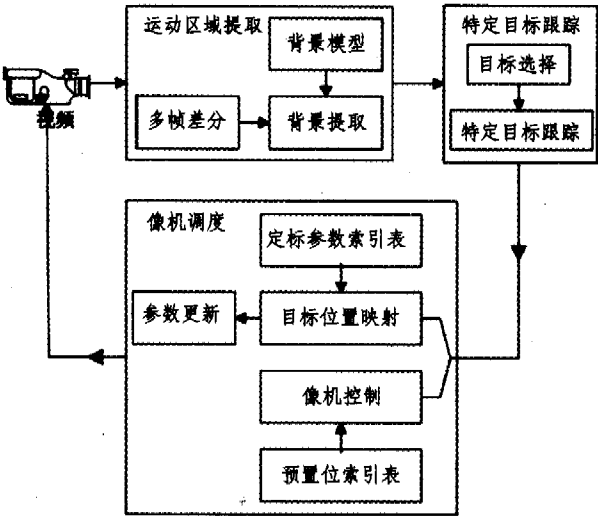


图1 主动系统跟踪流程图

在主动跟踪开始前, 系统先按照从左到右、从上到下设定预置位, 并且在每个预置位拍摄一幅背景图。系统工作时, 首先通过结合 Kalman 滤波器和 Mean-Shift 方法实现对特定目标的跟踪; 然后, 通过特定目标的运动参数和位置信息, 判断是否需要转动像机。如果要转动, 则计算对应的预置位, 并根据对应的控制原理, 实现对像机的旋转; 与此同时, 通过计算机视觉的方法, 把像机转动前的特定目标位置映射到像机转动后的图像坐标系中, 并修改 Kalman 和 Mean-Shift 对应的



(a) 第36帧



(b) 第252帧



(c) 第605帧

图3 室内主动跟踪结果

图4为室外情况下主动跟踪结果, 白色为跟踪轨迹, 红色为目标检测结果。当像机处于预置位1时, 实现了对单个行

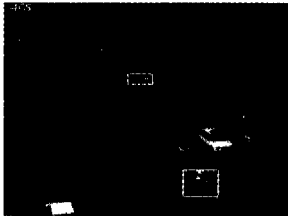
参数, 从而实现对特定目标的继续跟踪。

基于此, 本文提出了一种主动跟踪系统框架。系统主要分为预处理、视频理解和像机调度三个部分。系统简单的工作流程如图1所示。

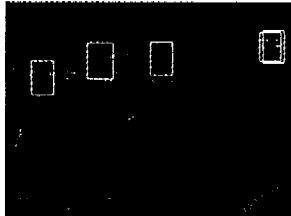
6 实验结果与分析

本文对室外场景进行了多目标的检测和跟踪, 对实验室的运动小火车进行了主动跟踪的模拟试验。图像序列的拍摄帧率约为 25 帧/s, 在 P4 2.66GHz、256MB、摄像机为 Sony FCB-450BP 的平台上, 跟踪和主动跟踪处理后的平均帧率为 20~22 帧/s。

在图2(a), (b)分别对雨天和雾天进行了多目标跟踪。在图2(a)上方汽车被树木遮挡住, 此时遮掩处理对其运动轨迹做了有效的预测分析。图2(b)中汽车被图像中的大树完全遮掩, 目标没有消失, 其运动轨迹仍然延续。无论在何种天气下的复杂自然环境中, 本文的算法都能较好地跟踪到各类真实目标, 能较好处理目标发生部分遮掩、完全遮掩以及形变的问题。



(a) 雨天



(b) 雾天

图2 室外检测和跟踪结果

对于主动跟踪, 绿色的框显示了跟踪的目标的位置及大小, 红色框中白色表示运动检测的目标区域, 目标身后白色的轨迹线表示其在整个图像序列中的运动轨迹。整个随动过程中, 在像机转动过程中不旋转像机, 在像机静止状态对目标检测并且跟踪特定目标。本文一共设置 9 个预置位, 依次从左至右、从上而下对 x 和 y 方向分别设置 3 个预置位。目标出现时的初始预置位为第 7 个预置位。图3(a)为运动小火车在预置位7的位置的检测和跟踪结果, 其中手的点击鼠标操作被检测。由于目标的往上运动, 并且立刻到达图像上边界, 依据第6节可知像机将往上运动, 从而选择预置位4使像机往上转动; 图3(b)为运动小火车在预置位7转动到预置位4的检测和跟踪结果; 同样, 小火车在预置位4中运动到图像右边界时, 像机将像右旋转, 选择预置位3; 图3(c)为预置位4转动到预置位5的情况。

人的跟踪, 目标在经过树的遮挡的时候仍能较好地跟踪; 当日
(下转第 257 页)

其中 h 及 g 分别表示同一小波基的低通和高通滤波器系数。依据式(1)、(2), 可对尺度空间 V_j 作继续分解到任意空间 V_{j+1} 。系数重建快速算法可表示为 $C_{j+1,m} = \sum_m h(m-2k)C_{j,k} + \sum_m g(m-2k)d_{j,k}$ [5]

由一维空间推广到二维空间, 得二维正交小波变换的 Mallat 算法

$$\alpha_{i,l}^j = \sum_{k,m} g(m-2l)g(k-2i)s_{k+2i-2,m+2l-2}^{j-1}$$

$$\beta_{i,l}^j = \sum_{k,m} g(k-2i)h(m-2l)s_{k+2i-2,m+2l-2}^{j-1}$$

$$\gamma_{i,l}^j = \sum_{k,m} h(k-2i)g(m-2l)s_{k+2i-2,m+2l-2}^{j-1}$$

$$s_{i,l}^j = \sum_{k,m} h(k-2i)h(m-2l)s_{k+2i-2,m+2l-2}^{j-1}$$

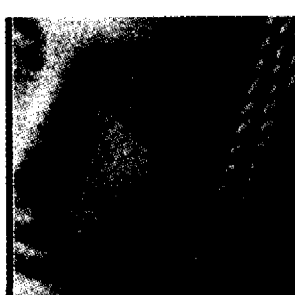
其中上标表示尺度, 下标表示两个方向的位移, $s_{i,l}^0$ ($i, l=1, 2, \dots, N$) 为离散原始图像, 分解图示如图 1。

这里 s^{22} 表示低频分量; $\beta_{i,l}$ 表示水平方向上细节信号; $\gamma_{i,l}$ 表示铅直方向上细节信号; $s_{i,l}$ 表示沿对角线方向的细节信号。

4) 计算逆小波变换 $\hat{x} = W^{-1}x$, 这里 \hat{x} 为还原信号。



(a) "Lena" 原始图



(b) "Lena" 含噪图



(c) "Lena" 去噪图

图 2 "Lena" 图

结论 通过以上仿真可以看出, 对阈值法的常数阈值进行必要改进, 结合小波变换系数分解计算和对阈值处理, 达到对数字图像的去噪和重建, 降低了噪音点, 使得图像变得更清晰。

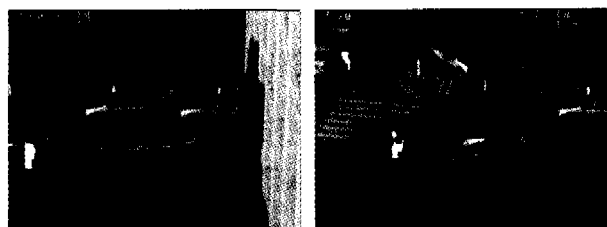
参考文献

1 Daubechies I. The wavelet transform: Time-frequency localization and signal analysis. IEEE Trans Theory, 36: 961~1005

2 Mallat S G. Multifrequency channel decompositions of images and wavelet models. IEEE Trans Acoustics, Speech, and Singnal Processing, 1989, 37(12)
3 谢杰成, 等. 小波图像去噪综述 J. 中国图形图像学报, 2002(4): 209~217
4 文洪雁, 等. 非线性小波变换阈值法去噪改进 J. 测绘通报, 2006(3): 18~21
5 李建平, 唐远炎. 小波分析方法的应用. 重庆: 重庆大学出版社, 1997

(上接第 255 页)

标向左运动时, 像机自动转到预置位 2, 通过计算机视觉计算出行人在下一个预置位中的可能位置, 然后综合检测结果, 实现特定目标的跟踪, 并且能很快速地检测出在该预置位中的其他运动目标, 如图 4(b) 中的左上角行人。



(a) 目标在预置位 1

(b) 目标在预置位 2

图 4 室外主动跟踪结果

小结与展望 本文提出了一个基于计算机视觉的主动跟踪方法。通过建立像机转动索引表、背景索引表和像机参数索引表, 能够实时实现对场景中运动区域的检测; 综合利用多帧差分 and 自适应背景方法提取运动对象, 采用 Kalman 滤波器和 Mean-Shift 实现对特定目标的跟踪, 并结合计算机视觉

的理论, 较好地实现对特定目标的主动跟踪, 从而提高了主动跟踪的速度和精度; 利用目标的运动信息和位置信息, 并结合对应的控制原理, 快速稳健地实现了像机的调度。

然而, 本文在特定的目标的选择方面还是人工进行。可以通过对目标的行为分析, 从语义层面上实现对感兴趣目标的主动跟踪, 这是今后需要努力的方向。

参考文献

1 Gavrilu D M. The Visual Analysis of Human Movement: A Survey. Computer Vision and Image Understanding, 1999, 73(1): 82~98
2 Zhang Z Y. Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. In: Proc. of the Seventh IEEE International Conference on Computer Vision, 1999, 1(1): 666~673
3 冯文源. 近影摄影测量——物体外形与运动状态的摄影法测量. 武汉: 武汉大学出版社, 2002
4 Stauffer C, Grimson W E L. Adaptive Background Mixture Models for Real-time Tracking. In: Proc. Computer Vision and Pattern Recognition, 1999, 2: 23~25
5 Collins R, et al. A system for video surveillance and monitoring: VSAM final report; [Technical Report]. Carnegie Mellon University, CMU-RI-TR-00-12, 2000