

复杂环境下运动摄像机稳定跟踪运动目标算法

向桂山

(浙江科技学院 信息与电子工程学院,杭州 310023)

摘要:在复杂环境中,运动摄像机跟踪运动目标是一项相当困难的工作。在基于目标颜色特征的 Mean Shift 跟踪算法中,引入感兴趣区域(ROI, Region of Interest),减少背景干扰及降低计算消耗。提出基于目标强度和目标面积的目标危机判别函数,对强干扰、遮挡情况进行识别;采用直方图维数和量化等级数自适应选取策略解决强干扰,采用子区域搜索选优策略解决目标遮挡和重新捕获的问题。为使被跟踪目标锁定在摄像机视野中央区域,采用基于速度调节的闭环控制模型,驱动 PTZ 摄像机,跟踪运动目标。实验结果表明,算法对背景干扰和遮挡具有较强的适应性,摄像机可以平滑稳健地跟踪快速运动目标,而且系统计算代价小,完全达到了实时的运行速度。

关键词: Mean Shift;感兴趣区域;危机判别函数;随动跟踪

中图分类号: TP391

文献标识码: A

文章编号: 1671-8798(2009)04-0339-05

Stably tracking moving object with active camera in complex environment

XIANG Gui-shan

(School of Information and Electronic Engineering, Zhejiang University of Science and Technology,
Hangzhou 310023, China)

Abstract: Stably tracking moving object in complex environment is troublesome. With the tracking method of Mean Shift based on the object's color feature, region of interest (ROI) is introduced to reduce the disturbance from background and the computational consumption. A discriminant function based on the object's intensity and the object's area is proposed to judge disturbance and occlusion. The adaptive selection strategy of dimensions and bins of histogram is suggested to escape disturbance, and searching strategy in sub-region is suggested to handle occlusion. To center the moving object, a closed loop control model based on speed regulation is applied to drive a PTZ camera to center the target. The results of our experiments show that the active camera can follow-up track moving object stably, even when encountering large area intense disturbance from background. Also the tracker can handle severe or complete occlusion and can recapture the target when it appears again. The system is computationally efficient and can run in real-time speed completely.

Key words: Mean Shift; region of interest; crisis discriminant function; follow-up tracking

收稿日期: 2009-09-20

基金项目: 浙江省教育厅科研项目(Y200804686)

作者简介: 向桂山(1980—),男,土家族,湖北恩施人,讲师,博士,主要从事智能监控、模式识别、机器视觉等研究。

基于视频图像序列的运动目标跟踪具有广泛的应用范围,比如视觉伺服控制、智能监控、人机接口、图像压缩、体育分析以及行为识别等^[1-3]。通常,运动目标跟踪主要可以分为如下 2 种方法:自下而上的跟踪方法和自上而下的跟踪方法。在自下而上的跟踪方法中,先从图像中分割检测出目标,然后进行跟踪,帧差法、背景减法、blob 检测等方法可以用于此模式进行目标提取跟踪。自上而下的跟踪方法是先对目标模式进行预测假设,然后在图像中进行验证,基于模型的方法和模板匹配方法就属于这种跟踪方法。在目标运动—摄像机静止模式下,由于背景变化较小,自下而上的方法可以取得较好的跟踪效果。但由于摄像机视野的限制等因素,目标运动—摄像机运动模式在应用中更为普遍,在这种模式中,因为摄像机的跟随运动,摄像机自身的扰动以及环境的干扰,传统的帧差法、背景减、光流等方法受到了限制。而本研究正是为了寻求一种运算代价小、抗背景干扰和遮挡能力强、能跟踪无规则运动目标的算法,并辅以适当的控制策略,驱动具有水平转动、垂直俯仰 2 个自由度的摄像机始终锁定运动目标并力图使之位于视野的中央。

在基于模型的目标跟踪算法中,基于目标颜色特征的 Mean Shift 算法无论从实时性还是鲁棒性,均表现出了较优秀的性能。Mean Shift 算法是一种基于核密度估计的非参数快速模式匹配算法,能沿着梯度方向快速搜索局部密度的峰值。它最早由 Funkunage 和 Honstetler 提出并应用到模式识别中^[4];Comaniciu D^[5-7]等人把基于目标颜色概率分布的 Mean Shift 算法用于跟踪非刚性物体,无论在性能上还是在速度上都达到了较好的效果;G.R. Bradski^[8]把 Mean Shift 算法应用在连续序列图像的脸部跟踪中,并称其为 Camshift 算法,其结果表明此算法对不规则目标运动、图像噪声、虚假目标、部分遮挡和光线变化都有较好的适应性,而且算法的实时性很好。但基于目标颜色特征的 Mean Shift 算法仍然存在无法处理背景中出现大面积强干扰和完全遮挡的情况。为此本研究力图解决这个问题。

1 基于目标颜色特征的跟踪算法

1.1 目标颜色特征模式表达

目标的颜色直方图反映了目标所包含颜色的出现概率,受目标的形状、大小及姿态变化影响较小,所以用颜色直方图来表达目标模式,对刚性和非刚性目标均有较好的稳定性。设目标的中心为 0,

$\{x_i^*\}_{i=1 \dots n}$ 表示目标像素点在图像中的位置, n 为目标的像素总数,函数 $b:R^2 \rightarrow \{1 \dots m\}$ 是位于 x_i^* 的像素向颜色直方图索引的映像, m 为直方图颜色的量化等级。离目标中心越远的像素,越容易受到遮挡或背景干扰的影响,应赋以越小的权重,所以为增加目标模式估计的鲁棒性,对直方图引入一个凸的并单调递减的权重核轮廓函数 $k:[0, \infty) \rightarrow R$ 。假设坐标 x, y 已经分别被目标的 2 个半轴 h_x, h_y 归一化,则目标的任一颜色 u 的直方图可以表示为:

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (1)$$

式(1)中: δ 为 Kronecher delta 函数;归一化常数 C 使得 $\sum_{u=1}^m \hat{q}_u = 1$, 所以有:

$$C = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2)} \quad (2)$$

在当前帧中,如果候选目标中心为 $y, \{x_i\}_{i=1 \dots n_h}$ 表示目标像素点的位置, n_h 为目标像素总数。则候选目标的任一颜色 u 的直方图可以表示为:

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (3)$$

常数 h 决定了候选目标的尺度。同样,归一化常数 C_h 使得 $\sum_{u=1}^m \hat{p}_u = 1$, 所以有

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} \quad (4)$$

1.2 目标颜色特征的相似性度量准则

在当前帧中跟踪目标,就是要找到离散位置 y , 使得候选目标的颜色分布模式 $\hat{p}(y) = \{\hat{p}_u(y)\}_{u=1 \dots m}$ 与被跟踪目标的颜色分布模式 $\hat{q} = \{\hat{q}_u\}_{u=1 \dots m}$ 最相似。为了定义相似性度量, Bhattacharyya 系数被用于评价直方图匹配程度。被跟踪目标与候选目标的相似系数为:

$$\hat{\rho}(y) \equiv \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (5)$$

根据式(5),定义 2 种颜色分布模式之间的距离为:

$$d(y) = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]} \quad (6)$$

两者间的距离越小,则相似度越大。

1.3 基于 Mean Shift 算法的目标跟踪

设上一帧的目标位置为 \hat{y}_0 , 以其为初始位置在当前帧中搜索目标,所以首先计算 \hat{y}_0 处候选目标的颜色分布模式 $\{\hat{p}_u(\hat{y}_0)\}_{u=1 \dots m}$ 。搜索跟踪目标,即求式(6)的最小距离,等价于求式(5)的最大值。在

$\hat{p}_u(\hat{y}^0)$ 邻域对 $\hat{\rho}_y$ 进行 Taylor 展开,有:

$$\hat{\rho}_y \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}^0) \hat{q}_u} + \frac{1}{2} \sum_{u=1}^m \hat{p}_u(y) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}^0)}} \quad (7)$$

把式(3)代入式(7),可得:

$$\hat{\rho}_y \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}^0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (8)$$

其中

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}^0)}} \quad (9)$$

式(8)的第一项为与 y 无关,所以求式(5)的最大值即要最大化式(8)的第二项。第二项表示当前帧中核轮廓函数 k 在位置 y 处的密度估计, w_i 为权重。通过 Mean Shift 迭代,此项可以快速达到最大值。

在跟踪算法中,在当前帧的颜色概率分布图像中进行 Mean Shift 迭代。颜色概率分布图像用后向投影(Back Projection)的方法获得^[8],首先把目标的颜色直方图 $\hat{q} = \{\hat{q}_u\}_{u=1 \dots m}$ 按下式从 $[0, \max(\hat{q})]$ 归一化到 $[0, 255]$ 区间中:

$$\hat{q} = \left\{ \min \left(\frac{255}{\max(\hat{q})} \hat{q}_u, 255 \right) \right\}_{u=1 \dots m} \quad (10)$$

然后根据颜色将图像中的像素值用直方图中的对应值代替,这样像素点属于目标像素的概率越大,颜色概率分布图像中相应像素点的值越大。

1.4 感兴趣区域

Mean Shift 是一种搜索局部峰值的匹配算法,不需在整幅图像上作全局运算,由此想到不必计算整幅图像的颜色概率分布,从而引入感兴趣区域(ROI),即只需要计算 ROI 的颜色概率分布,目标跟踪算法在 ROI 内完成。ROI 的引入具有 2 个明显的优点:一是减小计算量,主要是计算颜色概率分布和目标跟踪算法的计算量;二是减小干扰,提高了跟踪的准确性,因为 ROI 之外的区域无论干扰多么强烈,都不会对目标跟踪产生影响。确定 ROI 的位置和大小,是一个相当关键的问题。如果 ROI 太大,则减弱了它的作用;如果 ROI 太小或位置过偏,可能导致部分目标位于 ROI 之外,造成跟踪误差甚至失败。本算法中,ROI 的位置为 Kalman 滤波器预测的目标在当前帧中的位置,大小为上一帧目标大小的 2 倍。若此 ROI 没能包含上一帧目标的中心,则 ROI 向上一帧的目标中心位置伸展直到包

含。实验表明,此包含策略可以减少由于 Kalman 预测误差造成的 ROI 位置误差,Kalman 预测误差主要由目标的不规则运动引起。

2 强干扰、遮挡的判断及处理策略

在跟踪过程中,很可能遭遇背景中的大面积强干扰或者遮挡,因此把对目标的干扰和遮挡定义为目标危机。首先只有正确识别目标危机,才能采取合适的对策。齐苏敏^[9]等人依据 Bhattacharyya 系数的大小来判断目标是否被遮挡,但 Bhattacharyya 系数只是表征了归一化的 2 个目标颜色直方图之间的相似度,目标被部分遮挡时,颜色特征不一定发生较大变化,所以这种判别准则不够准确。朱胜利^[10]等人用 Kalman 滤波残差大小来判断目标是否被遮挡,但由于快速运动目标可能的不规则运动,Kalman 滤波结果本身不可能很精确,所以这种判断准则容易形成误判。而且目前文献中还未对背景中大面积强干扰这种情况进行判断处理,也就不存在相关判断准则。由此笔者提出基于目标强度和目标面积的目标危机判别函数,对强干扰和遮挡情况进行判断。由目标的颜色直方图的性质可知,当直方图的维数和每一维的量化等级数越高时,则对目标的颜色特征描述越精确,越不容易受到干扰,但对光线等的适应性越差;越低时,则对目标的颜色特征描述越粗略,但对光线等的适应性越好。所以根据目标危机函数,自适应选择目标颜色直方图的维数和每一维的量化数,将能在目标的抗干扰性和完整性之间达到平衡。

目标强度 I 定义为颜色概率分布图像中目标外接矩形内像素值总和,表征目标颜色准确性;目标面积 A 定义为颜色概率分布图像中目标外接矩形的大小,表征目标外形完整性。设背景无干扰时,以 H-S 二维直方图计算的待跟踪目标的强度均值为 I_2 ,目标面积均值为 A_2 ,以 H-S-V 三维直方图计算的待跟踪目标的强度均值为 I_3 ,目标面积均值为 A_3 。目标危机判别函数定义为:

$$\left\{ C_i = \alpha \frac{I_c}{I_i} + (1 - \alpha) \frac{A_c}{A_i} \right\}_{i=2,3} \quad (11)$$

其中 I_c 和 A_c 为由 H-S 二维直方图计算的当前帧中的目标强度和目标面积, I_3 和 A_3 为由 H-S-V 三维直方图计算的当前帧中的目标强度和目标面积; $\alpha < 1$ 为权重系数,若目标是否遭遇危机更加依赖于颜色准确性进行判断,则 α 应取偏大的值,若目标是否遭遇危机更加依赖于外形完整性进行判断,

则 α 应取偏小的值。

在未出现大面积强干扰时,由 H-S 二维直方图来表达目标颜色特征,如果当前帧的目标危机判别函数满足式(12):

$$C_2 > D_2 \quad (12)$$

则表明背景中出现大面积强干扰,此时由 H-S 二维直方图改用更精确的 H-S-V 三维直方图跟踪目标。如果由 H-S-V 三维直方图跟踪目标时,当前帧的目标危机判别函数满足式(13):

$$C_3 > D_3 \quad (13)$$

表明背景中大面积强干扰仍然存在,则此时增加 H-S-V 三维直方图的量化等级数,特别是 V 维的量化等级。其中 D_2 、 D_3 为干扰判别阈值,一般根据目标的最大可能强度和面积决定。

当用 H-S-V 三维直方图跟踪目标时,同时也计算 H-S 二维直方图下的目标强度和目标面积,一旦满足式(14):

$$C_2 < D_2 \quad (14)$$

则由 H-S-V 三维直方图改用 H-S 二维直方图跟踪目标,以保证目标完整性。由于 H-S-V 三维直方图对目标姿态、光线变化敏感,所以为了稳定的跟踪住目标,当用 H-S-V 三维直方图跟踪目标时,必须加快直方图的更新速率。直方图的更新策略为:

$$\{\hat{q}_u \leftarrow (1 - \gamma)\hat{q}_u + \gamma\hat{p}_u\}_{u=1 \dots m} \quad (15)$$

式(15)中: $\gamma < 1$ 为更新速率, \hat{q}_u 为目标颜色直方图, \hat{p}_u 为目标在当前帧的相应颜色直方图。

与干扰判断相似,当用 H-S 二维直方图跟踪目标时,如果满足式(16):

$$C_2 < O_2 \quad (16)$$

则认为目标发生了严重遮挡。当用 H-S-V 三维直方图跟踪目标时,如果满足式(17):

$$C_3 < O_3 \quad (17)$$

则认为目标发生了严重遮挡。 O_2 和 O_3 均为遮挡判别系数,取值根据对遮挡程度的定义而决定。

当目标发生严重或完全遮挡时,设发生遮挡前目标的位置为 P_{pre} ,速度为 V_{pre} ,则据式(18)预测目标将会出现的位置并检测目标是否出现:

$$P_{now} \approx P_{pre} + V_{pre} \times (t_{now} - t_{pre}) \quad (18)$$

为防止捕获虚假目标,根据 H-S-V 三维直方图下的目标强度和目标面积作为判定目标重新出现的准则:

$$\eta\rho_3 < C_3 < \lambda D_3 \quad (19)$$

其中 $\eta > 1$ 、 $\lambda < 1$,引入这 2 个系数,是为了避免捕获虚假目标,所以判断阈值应该比干扰和遮挡的判断

阈值更加严格,即更加接近目标的真实值。

若经过长时间未在预测位置发现目标,则分别以如图 1 所示的 5 个子区域为初始跟踪窗和 ROI,搜索目标。若有多个子区域内的目标均满足式(19),则使式(20)具有最小值的目标作为重新捕获的目标:

$$\alpha \|1 - \frac{I_{3c}}{I_3}\| + (1 - \alpha) \|1 - \frac{A_{3c}}{A_3}\| \quad (20)$$

若以整个图像为初始跟踪窗和 ROI,则由于干扰,即使目标出现,跟踪结果可能不会收敛于目标,而是同时包含了干扰和目标。由于摄像机始终跟随目标运动,所以图像中央区域是目标重新出现概率较大的地方,以图 1 所示的 5 个子区域搜索目标并根据式(20)选优,提高了目标重新被捕获的及时性和准确性。

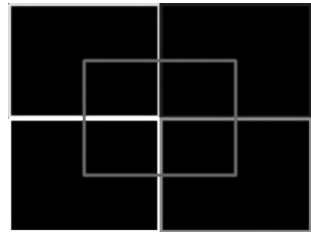


图 1 目标搜索子区域

Fig.1 Sub-areas for searching object

3 基于速度调节的跟踪控制模型

摄像机的视野有限,若摄像机能随着目标运动,并尽力保持目标在视野中央,无疑能大大增加目标的跟踪范围,提高被跟踪目标的信息量。实验采用由步进电机驱动的 PTZ 摄像机,基于摄像机速度调节的伺服控制模型来跟踪运动目标。

设图像中心坐标为 (x_0, y_0) ,作为控制模型的输入,目标在当前帧的跟踪位置为 (\hat{x}, \hat{y}) ,作为控制模型的反馈。根据两者的位置偏差 $e = (e_x, e_y)$ 确定摄像机的控制量:摄像机的水平旋转速度和垂直俯仰速度 (a_{pan}, a_{tilt}) 。由此采用比例控制,比例控制量:

$$\omega = [a_{pan} \quad a_{tilt}]^T = \frac{k_p}{f} \cdot e(t) \quad (21)$$

式(21)中: k_p 为比例增益向量, f 为摄像机焦距。在模型中无需积分和微分控制。 k_p 依据摄像机和目标之间的几何运动模型确定,但精确模型一般比较复杂,较难确定,而且计算量很大,所以本算法中 k_p 由多次实验测定并拟合获得。摄像机运动控制框图如图 2 所示。

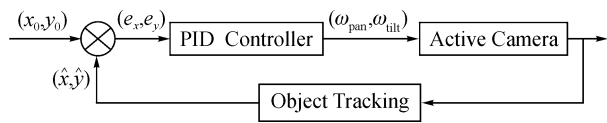


图2 摄像机运动控制框图

Fig.2 Diagram for camera control

4 实验

本算法在奔腾 IV 2.8G CPU、512M DDR 内存配置的 PC 机 VC6.0 环境下开发,摄像机为东芝 VK-S914E,安装于自制云台上,图像采集卡为天敏 SDK-2000,PC 机通过 RS-232 对云台进行运动控制,视频图像大小为 320×240 。

图3为跟踪实验结果,灰色矩形框表示 ROI,白色矩形框表示初始跟踪窗,黑色矩形框为目标外接矩形,椭圆大致拟合物体形状并表示物体方向。在实验中,运动目标的姿态发生极大改变,目标被完全遮挡,背景中出现与目标颜色或形状相似的干扰,目标仍然能被运动摄像机稳定跟踪,当目标完全被遮挡且再次出现时,能被及时捕获和跟踪,其中遮挡等待时间的长短可由外部指定。

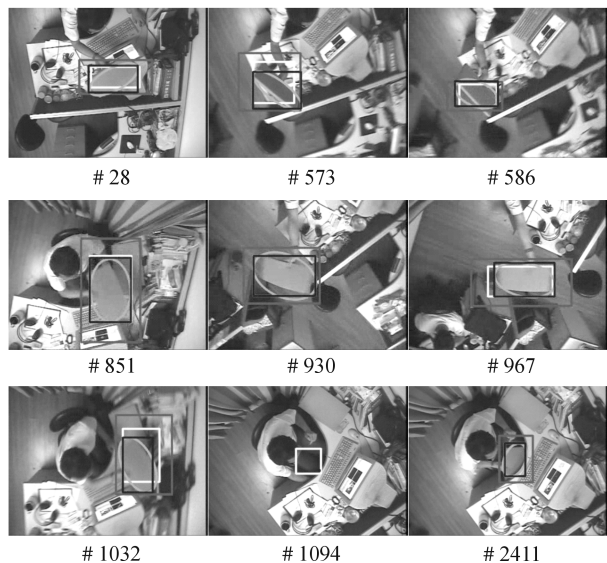


图3 复杂环境下的运动目标跟踪

Fig.3 Experimental results of object tracking in complex environment

在平均占用 37% CPU、28 M 内存的资源消耗下,跟踪速度平均为 6.4 ms/帧,每帧的 Mean Shift 迭代次数平均为 2.9,跟踪成功率达到 93.8%。当遇到背景中非常相似或相同的物体且与目标大部分交叠时,跟踪容易遭遇失败。

5 结 语

在 Mean Shift 跟踪算法中,本研究引入感兴趣区域,从而大大减少了干扰,且减少了计算量和 Mean Shift 迭代次数;基于目标强度和目标面积的目标危机判别函数能够较准确的识别背景强干扰和遮挡情况,而且其对应处理策略解决了背景中的大面积强干扰目标跟踪性能的影响和目标重新捕获的问题;算法的计算代价很小,完全达到了实时的运行速度,辅以基于速度调节的比例控制模型,使摄像机实现了平滑稳定跟踪快速运动目标的目的。

参考文献:

- [1] HARITA OGLU I, HARWOOD D, DAVIS L S. W⁴: Real-time surveillance of people and their activities[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2000, 22(7): 809-830.
- [2] YILMAZ A. Object tracking by asymmetric kernel Mean Shift with automatic scale and orientation selection[C]// IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA: IEEE, 2007: 1-6.
- [3] KHAN S M, SHAH M. A multiview approach to tracking people in crowded scenes using a planar homography constraint[C]// 9th European Conference on Computer Vision. Berlin, Germany, 2006: 133-146.
- [4] FUKUNAGA K, HOSTETLER L D. The estimation of the gradient of a density function, with application in pattern recognition[J]. IEEE Transaction on Information Theory, 1975, 21(1): 32-40.
- [5] COMANICIU D, MEER P. Mean Shift analysis and applications [C]// Proceedings of IEEE International Conference Computer Vision. Kerkyra, Greece, USA: IEEE, 1999: 1197-1203.
- [6] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using Mean Shift [C]// Proceeding of IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, SC, USA: IEEE, 2000, 2: 142-149.
- [7] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [8] BRADSKI G R. Real-time face and object tracking as a component of a perceptual user interface [C]// Proceeding of IEEE Workshop on Application of Computer Vision. Princeton, NJ, USA: IEEE, 1998: 214-219.
- [9] 齐苏敏, 黄贤武, 孟静. 基于自适应活动轮廓模型的实时手势跟踪[J]. 计算机科学, 2006, 33(11): 192-204.
- [10] 朱胜利, 朱善安, 李旭超. 快速运动目标的 Mean Shift 跟踪算法[J]. 光电工程, 2006, 33(5): 66-70.