

基于多人遮挡的定位跟踪算法

毛 爽, 方 颖, 陈 曙, 王汇源

(山东大学信息科学与工程学院, 济南 250100)

摘 要: 针对传统视频跟踪算法中存在的问题, 提出一种基于多人遮挡问题的定位跟踪算法, 该算法采用改进的投影分析方法对运动目标进行定位, 并结合 Kalman 滤波进行匹配跟踪。仿真实验结果表明, 该算法能够有效区分、定位、跟踪处于 2 种不同遮挡状态下的每个人的位置, 具有一定应用价值。

关键词: 多人跟踪; 遮挡; 投影分析; Kalman 滤波

Location and Tracking Algorithm Based on Multiple Humans Occlusion

MAO Shuang, FANG Ying, CHEN Shu, WANG Hui-yuan

(School of Information Science & Engineering, Shandong University, Jinan 250100)

【Abstract】 Aiming at the problems existed in traditional video tracking algorithm, a novel location and tracking algorithm based on multiple humans occlusion is proposed, which locates the moving object by using an improved projection analysis method, and conducts matching tracking combined with Kalman filter. Simulation experimental results show this algorithm can separate, locate and track each human location in two different occluded situations.

【Key words】 multiple humans tracking; occlusion; projection analysis; Kalman filtering

1 概述

在单目摄像机及背景相对静止的条件下, 用背景动态建模、减除背景的方法将运动像素提取出来, 并对这些运动像素进行联通区域分析和聚类, 从而定位每个运动目标。在已有文献中, 总是假设目标进入场景时是单一的^[1-2], 一个运动块对应单一的运动目标; 在发生部分或完全遮挡情况下, 则将多个运动目标当作整体来对待^[1,3]。 W^4 系统^[4]用投影分析和轮廓分析来区分块中的多个人, 但是当多个人在同一垂直线上时, 该方法会出错。因此, 多人视频跟踪要求算法能够达到以下几个要求: (1) 在发生遮挡时, 算法应能定位或预测遮挡人的位置; (2) 由于运动像素提取有误差, 一个人可能对应几个小块, 需要将小块“合并”; (3) 在做背景减除时, 由于人有可能在几帧中丢失, 因此算法应能预测丢失人的位置。

为解决上述问题, 本文提出一种多人跟踪算法。当多人由于位置接近而导致一个运动块对应多人时, 会认为发生了浅度遮挡; 而当多人因为运动重叠在一起, 以致不能从外形上区分他们时, 则认为发生了深度遮挡。对于这 2 种遮挡, 本文采用不同策略: 预处理提取运动像素; 用一种改进的投影分析法对目标进行定位分析, 找到场景中每一目标的位置身高等信息, 同时分割浅度遮挡情况下的每一个人; 为每个目标建立一个 Kalman 滤波器, 用 Kalman 预测值估计目标在下一帧中的位置, 建立目标帧间的对应关系; 如果目标由于检测的原因或是深度遮挡的原因丢失, 则用 Kalman 滤波预测目标位置。

2 算法介绍

2.1 预处理

采用基于单峰高斯模型背景建模的方法建立动态背景,

然后进行背景减除, 以提取前景像素, 并用中值滤波滤去小块噪声。处理结果如图 1(a)所示, 可以看出在提取的二值化图像中仍有小块噪声, 前景目标人身上会有“小洞”等干扰因素, 一个人也可能对应多个小块。

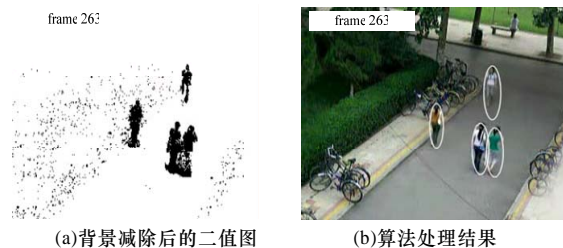


图 1 改进的投影分析法处理过程

2.2 改进的投影分析定位算法

投影分析法将背景减除提取的二值图像分别对水平轴做垂直投影, 即记录每一垂直线上黑色像素点的个数; 同理, 对垂直轴做水平投影。通过选择较大的垂直投影值和水平投影值的范围, 确定目标的位置。当场景中只有单个物体时, 该方法有效; 如有多人则会出错。本文对该方法进行改进, 使其在多人之间发生浅度遮挡时, 也能有效进行分割定位。

人头顶位置的横坐标对应垂直投影图中局部波峰的横坐标; 而人的高度对应垂直投影图中局部波峰的纵坐标。据

基金项目: 国家自然科学基金资助项目(60675024)

作者简介: 毛 爽(1982—), 男, 硕士研究生, 主研方向: 计算机视觉, 图像处理; 方 颖, 硕士研究生; 陈 曙, 副教授; 王汇源, 教授

收稿日期: 2008-11-06 **E-mail:** maoshuang989@yahoo.com.cn

此, 提出改进的投影分析定位算法。设图像高度为 S_{height} , 图像宽度为 S_{width} , 像素坐标为 $y_i < TL_{\text{height}}$, 像素值 $P_{(a,b)}=0$ (黑色点) 或 $P_{(a,b)}=1$ (白色点)。算法步骤如下:

(1) 将二值化图像对水平轴做垂直投影, 得到投影图。

(2) 对投影图进行分析, 得到局部波峰的坐标 $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n)$ 。

(3) 对所有的波峰坐标, 如果 $y_i < TL_{\text{height}}$, 则删除对应的 (x_i, y_i) ; 如果 $y_i > TL_{\text{height}}$, 令 $y_i = TH_{\text{height}}$ 。

(4) 对所有的波峰坐标, 如果 $|x_i - x_{i-1}| < T_{\text{width}}$, 则删除对应的 (x_i, y_i) 。

(5) 在二值化图像中寻找头顶位置: 对 $1 \leq b_i \leq S_{\text{height}}$, 如果有 $P_{(x_i, y_i)}=0$ 且 $A_{(x_i, b_i, y_i)} > T_{\text{area}}$, 则可确定 (x_i, b_i) 为人的头顶坐标, y_i 为人的身高, 同时令人的宽度为 $y_i/3$, 此时已将人的定位分割出来 (ABC 3 个人)。

(6) 重复步骤(1)~步骤(5), 直到找不到有意义的波峰坐标为止。在步骤(3)中, TL_{height} 为人身高的最小门限, 以消除小波峰噪声的影响。有时会出现图 1(a) 中 2 个人在同一垂直线上的情况, 即在投影时会发生叠加的情况, 所以, 设置 TH_{height} 为人身高的最大门限值, 以防止将多个人画入一个椭圆中。在步骤(4)中, T_{width} 为 2 个人之间宽度的最小门限, 这可以消除如图 1(a) 中目标附近对应噪声的影响。这时 x_i 对应人头顶的水平坐标, y_i 对应人的身高。在步骤(5)中, $A_{(x_i, b_i, y_i)}$ 表示在二值化图像中以 (x_i, b_i) 为顶点, y_i 为长轴, $y_i/3$ 为短轴的椭圆内黑色像素点的个数, 即有意义的面积。设立面积最小门限 T_{area} , 防止头顶上方噪声的影响。在完成步骤(5)后, 可将多数人定位。当找不到有意义的波峰坐标时, 程序结束, 最后得到如图 1(b) 所示的结果。可以看到, 算法准确定位了每个人, 同时区分了浅度遮挡的人。

算法中的门限参数 TL_{height} , TH_{height} , T_{width} 和 T_{area} 可根据视频中人的大小来确定。改进的投影分析法可将场景中多人定位区分出来, 允许多人同时进入场景, 对于浅度遮挡的人群, 则可以将每个人区分定位。即使背景上有小块的噪声, 前景上有“小洞”, 甚至分裂成多块也不会影响结果。然而, 在“不好”的背景减除导致目标丢失和深度遮挡情况下, 采用基于 Kalman 滤波的匹配跟踪算法来预测目标的运动。

2.3 基于 Kalman 滤波的匹配跟踪算法

采用 Kalman 滤波提供最小平方的预测迭代方案能够较精确地对目标运动做出预测。首先进行模式状态预报, 接着引入观测数据, 然后根据观测数据更新模式状态。因为在定位过程中能较准确地统计目标头顶位置的坐标 (P_{xk}, P_{yk}) , 所以可以认为: 当目标在当前时刻某一速度机动时, 下一时刻的速度值是有限的, 且只能在当前速度邻域内。系统的状态方程和观测方程为

$$X_k = AX_{k-1} + BU_k + W_k \quad (1)$$

$$Z_k = HX_k + V_k \quad (2)$$

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (4)$$

其中, $X_k = (P_{xk}, P_{yk}, V_{xk}, V_{yk})^T$, U_k 和 $Z_k = (P_{xk}, P_{yk})^T$ 分别为 k 时刻的状态变量、输入变量和测量变量; A , B 和 H 分别为状

态转换矩阵、最优控制输入矩阵和观测转换矩阵; W_k 和 V_k 为相互独立的系统处理白噪声和观测白噪声, 则预测方程为

$$\hat{X}_k(-) = A\hat{X}_{k-1}(+) \quad (5)$$

$$P_k(-) = AP_{k-1}(+)A^T \quad (6)$$

其中, $(-)$ 表示从 $k-1$ 时刻到 k 时刻的预测值; $(+)$ 表示 $k-1$ 时刻的更新值; P_k 为 k 时刻的预测误差。更新方程为

$$\bar{K}_k = P_k(-)H^T(H P_k(-)H^T)^{-1} \quad (7)$$

$$\hat{X}_k(+) = \hat{X}_k(-) + \bar{K}_k(Z_k - H\hat{X}_k(-)) \quad (8)$$

$$P_k(+) = (I - \bar{K}_k H)P_k(-) \quad (9)$$

其中, \bar{K}_k 为 k 时刻的 Kalman 增益。

跟踪原则为: 为每个进入场景的人建立一个 Kalman 滤波器。如果用预方程式(7)和式(8)预测的 k 帧人头顶的坐标 $(\hat{P}_{xk}, \hat{P}_{yk})$ 与 k 帧的测量值 (P_{xk}, P_{yk}) 距离最近, 且满足如下公式:

$$|\hat{P}_{xk} - P_{xk}| < d_x \quad (10)$$

$$|\hat{P}_{yk} - P_{yk}| < d_y \quad (11)$$

其中, d_x, d_y 分别是 x 和 y 方向允许的最大位置误差, 则预测值和测量值可建立匹配关系, 并用测量值更新测量式(7)~式(9); 如果在 k 帧找不到测量值与之匹配, 且点 $(\hat{P}_{xk}, \hat{P}_{yk})$ 位置在场景中, 则认为目标在检测时丢失或发生深度遮挡, 用预测值估计目标位置, 代替测量值更新测量方程。这就可解决浅度和深度遮挡问题, 且实现对目标的准确定位和跟踪。

3 实验

图 2 是本文算法在一个具体实例中的应用结果, 图 2(a)~图 2(d) 是几个关键帧。其中, 图 2(a) 为 149 帧中 4 号人和 5 号人同时进入场景的情况。可以看到, 3 号~5 号人浅度遮挡, 发生“融合”, 而本文算法正确区分定位了每个人。在做背景减除时, 前景图中目标会存在着小块的噪声, 也会分裂成多个小块, 但本文算法并不受影响。从第 332 帧~第 365 帧, 6 号人和 3 号人发生深度遮挡。在图 2(c) 中, 3 号人大部分被遮住, 2 个人几乎重叠。由于 3 号人头部露出, 本文的投影算法准确定位了 3 号人; 对几乎被“掩埋”的 6 号人, 因为结合了 Kalman 预测, 所以没有将其丢失, 大致准确地跟踪了 6 号人的运动, 如图 2(b)~图 2(d) 所示。6 号人头顶的测量轨迹会有 2 处断开, 一段是由于“不好”的背景减除, 导致 6 号人丢失; 另一段是由于 3 号人和 6 号人发生遮挡, 导致丢失。

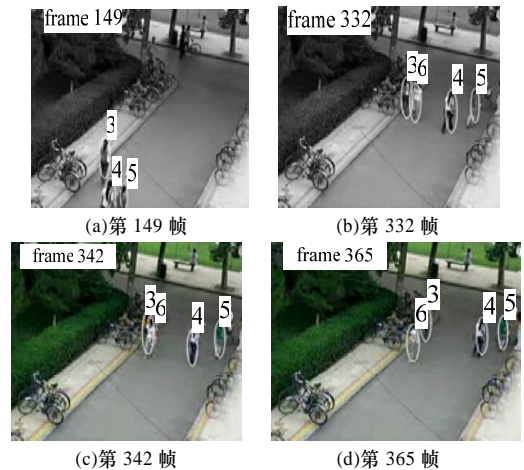


图 2 算法应用结果

(下转第 224 页)