

基于像素可信度和空间位置的运动目标跟踪

施 华¹ 李翠华¹ 韦凤梅² 王华伟¹¹(厦门大学计算机科学系 厦门 361005)²(厦门大学自动化系 厦门 361005)

(shihuang2002@sohu.com)

Moving Object Tracking Based on Location and Confidence of Pixels

Shi Hua¹, Li Cuihua¹, Wei Fengmei², and Wang Huawei¹¹(Department of Computer Science, Xiamen University, Xiamen 361005)²(Department of Automation, Xiamen University, Xiamen 361005)

Abstract Moving object tracking is a critical issue of image sequence processing. In this paper, a moving object tracking algorithm based on location and confidence of pixels is proposed. Firstly, the moving objects are detected by combining the median background model in temporal domain with the minimum cross-entropy in spatial domain. Then the rectangle area of the objects are obtained, and at the same time an HSV color distribution model is used to measure the similarity between target rectangles and hypothetical rectangles. In this process, a weighting function based on location and confidence of pixels is presented to weigh the pixel values in the rectangle area of the tracking. The experimental results show that the algorithm is computationally efficient and robust to scale invariant, partial occlusion and interactions of non-rigid objects, especially similar objects.

Key words tracking; cross-entropy; partial occlusion; color distribution; Bhattacharyya coefficient

摘 要 运动目标跟踪是视频信息处理的重要研究课题之一。首先将时间域上的中值背景建模与空间域上最小交叉熵法相结合,用于检测运动目标所在跟踪区域。在此基础上,提出了跟踪区域内基于像素的可信度与空间位置的权重函数,利用 HSV 色彩分布模型计算出目标模型与预测模型间的相似性,选出最优相似模型作为当前目标模型,从而实现了多目标的跟踪。实验显示,该算法计算简单,对相似目标能实现准确的跟踪,对非刚性目标的尺度变化、多目标的交叉及部分遮挡具有鲁棒性。

关键词 跟踪;交叉熵;部分遮挡;颜色分布;Bhattacharyya 系数

中图分类号 TP391.41

1 引 言

运动目标跟踪是计算机视觉研究的重点与难点,在军事和社会安全领域有着广泛的应用前景。近年来随着集成电路和计算机技术的迅速发展,视

觉监视系统所要求的硬件设备成本大大降低,因此,它获得了日益广泛的研究与应用。许多关于视觉监视的大规模研究项目已经在美国、欧洲等地展开。现在已有许多文章对不同应用条件中的检测与跟踪做了较详细地论述。Huttenlocher 等人提出了 Hausdorff 距离的定义与计算方法^[1],并在文献[2]

收稿日期:2004-05-26;修回日期:2004-12-13

基金项目:国家创新研究群体基金项目(60024301);国家自然科学基金项目(60175008)

实现了物体的跟踪算法. Haritaoglu 等人开发的 W^4 系统^[3]是一个实时的监控系统,它通过建立双峰分布的统计背景模型来实现目标的检测,并为目标建立外观模型来实现多人的跟踪. Stauffer 等人^[4]提出了自适应混合高斯模型,它利用连通区域检测算法对前景目标进行检测,通过匹配目标的位置、大小、形状和颜色等特性,建立运动目标与前景目标间的对应关系来实现对目标的跟踪. Nummiaro 等人^[5]利用支持向量机(SVM)技术对目标进行初始化检测,然后应用粒子滤波方法对目标进行跟踪,有效地解决了目标的部分遮挡问题. 另外,还有一些文章^[8~11]对跟踪算法也做了深入的探讨.

针对静止摄像机的情况,本文通过静止背景下时域上的中值背景建模与空间域的交叉熵相结合的方式在跟踪初始化时检测运动目标,然后利用颜色分布模型实现多目标的跟踪.

2 运动目标检测

运动目标的检测常用传统的背景减除法^[12,13],它只考虑了图像序列的时间相关性,而没有考虑图像序列的空间相关性. 本文拟采用时间域与空间域相结合的方法来设定当前图像与背景图像的差分图像的阈值,在空间域上采用最小交叉熵法计算阈值.

2.1 背景模型构建及更新

选择视频的前 N 帧来构建初始化背景. 设 $I(x)$ 和 $B(x)$ 分别为连续 N 帧图像中像素点 x 上的中值与均方差. 以 $B(x)$ 为背景图像的像素值来构建背景,即在初始化背景中, $B(x) = I(x)$. 然而,随着时间的推移,光照变化、运动场景的混乱、多目标的重叠及各种难以预知的变化均会显现出来. 为了应付这些外来的变化并准确跟踪到运动目标,必须使背景模型具有鲁棒性与自适应性.

利用一阶 Kalman^[14]滤波模型来更新背景,其中, $B_t(x)$ 为 t 时刻 x 点上的背景像素值, $I_t(x)$ 为 t 时刻 x 点上的当前像素值, $B_{t-1}(x)$ 为 $t-1$ 时刻背景图像上 x 点上的均方差值, α 是更新因子.

$$B_{(t+1)}(x) = (1 - \alpha) B_t(x) + \alpha I_t(x), \quad (1)$$

$$\sigma_{t+1}^2(x) = (1 - \alpha) \sigma_t^2(x) + \alpha (I_t(x) - B_t(x))^2. \quad (2)$$

依据上述公式,已知图像上某点的当前背景值,可以计算出差分图像值:

$$D_t(x) = abs(I_t(x) - B_t(x)). \quad (3)$$

差分图像的获取在一定程度上能区分前景与背

景,在第 2.3 节、第 3 节的前景目标检测与跟踪中有很大的用处.

2.2 基于空间最小交叉熵的阈值方法

信息熵是表征信源本身统计特性的一个物理量. 将信息熵引入运动目标的检测就是用熵决定一个合适的阈值,从而将差分图像分割为目标与背景,使二值化后的图像能正确反映原图像的特性.

为选择目标分割的最佳阈值,考虑到差分图像上像素的空间统计相关性,本文引入最小交叉熵函数^[15]来求取最优阈值而使原始差分图像和二值分割图像之间的信息量差异最小. 在理想的差分图像中,可以将直方图看成是目标与背景两个正态的混合分布,在这种假设下,设图像共有 L 级灰度,则像素的后验概率为

$$q(g) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_o(T)} \exp\left\{-\frac{[g - \mu_o(T)]^2}{2\sigma_o^2(T)}\right\}, & 0 \leq g < T, \\ \frac{1}{\sqrt{2\pi}\sigma_b(T)} \exp\left\{-\frac{[g - \mu_b(T)]^2}{2\sigma_b^2(T)}\right\}, & T \leq g < L, \end{cases} \quad (4)$$

式(4)中的类内方差可由原始图的灰度直方图 $h(g)$ 估计出来:

$$\sigma_o^2(T) = \frac{1}{P_o} \sum_{g=0}^{T-1} h(g) [g - \mu_o(T)]^2, \quad (5)$$

$$\sigma_b^2(T) = \frac{1}{P_b} \sum_{g=T}^{L-1} h(g) [g - \mu_b(T)]^2, \quad (6)$$

其中, $h(g)$ 为直方图上灰度值为 g 时的像素点个数,

$$P_o = \sum_{g=0}^{T-1} h(g),$$

$$P_b = \sum_{g=T}^{L-1} h(g),$$

$$\mu_o(T) = \frac{1}{P_o} \sum_{g=0}^{T-1} gh(g),$$

$$\mu_b(T) = \frac{1}{P_b} \sum_{g=T}^{L-1} gh(g).$$

应用最小交叉熵函数,差分图像与二值分割图像的交叉熵为

$$J(T) = \sum_{g=0}^{L-1} [gh(g) \ln \frac{g}{\mu_o(T)} + \mu_o(T) h(g) \ln \frac{\mu_o(T)}{g}] + \sum_{g=T}^{L-1} [gh(g) \ln \frac{g}{\mu_b(T)} + \mu_b(T) h(g) \ln \frac{\mu_b(T)}{g}], \quad (7)$$

使 $J(T)$ 达到最小熵的 T 值即为分割阈值,由此它分割差分图像用到了图像的空间统计相关性.

2.3 运动目标分割

获取了背景图像后,很容易得出当前图像的差分图像,从而区分出背景与运动前景. 设 $M_1(x)$ 是由差分图像与时间域上均方差而得到的二值化掩膜, $S_1(x)$ 是其掩膜图像, 设 $M_2(x)$ 是由差分图像与空间域的最小交叉熵而得到的二值化掩膜, $S_2(x)$ 是其对应的掩膜图像, 则

$$M_1(x) = \begin{cases} 1, & |I_t(x) - B_t(x)| > 2.5 \sigma_t(x), \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

$$S_1(x) = \begin{cases} 1, & M_1(x) = 1, \\ N_1/9, & \text{otherwise,} \end{cases} \quad (9)$$

其中, N_1 为 8 邻域中由式 (8) 判断为前景像素的个数.

$$M_2(x) = \begin{cases} 1, & |I_t(x) - B_t(x)| > T, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

$$S_2(x) = \begin{cases} 1, & M_2(x) = 1, \\ N_2/9, & \text{otherwise,} \end{cases} \quad (11)$$

其中, T 是由式 (7) 得到的阈值, N_2 为 8 邻域中由式 (10) 判断为前景像素的个数.

$$M(x) = \begin{cases} 1, & S_1(x) \times S_2(x) > 1/3, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

这种阈值的设定结合了时间相关性与空间相关性, 具有较好的抗噪性, 对环境也有很好的鲁棒性. 但是单独的阈值分割是不够的, 当目标区域与背景区域的像素值相近时, 分割出的前景区域可能发生断裂或出现空洞. 所以在视频图像的后处理中, 应用膨胀算子, 对断裂或者空洞部分进行处理. 然后, 采用基于前景区域面积阈值法去除噪声, 检测人时, 将面积小于 20 像素的前景区域去除. 最终的分割与检测结果可以参看图 1(b) 与 1(c).

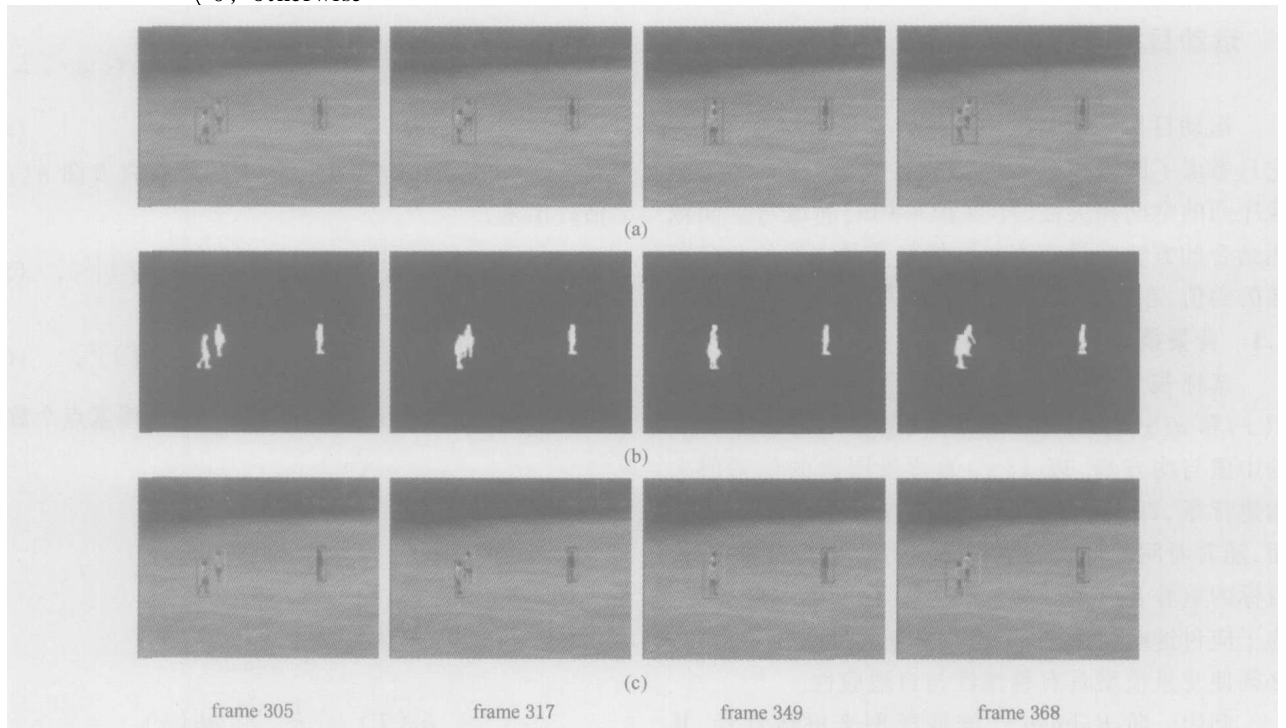


Fig. 1 Football sequence 1. (a) The targets was detected; (b) Binary image; and (c) The color distributions are similar when they interact with each other.

图 1 足球视频序列. (a) 检测图; (b) 分割后的二值图像; (c) 在颜色分布图相似的目标交叉时的跟踪

3 运动目标跟踪

运动目标的跟踪应用了视频图像的颜色分布模型. 通常使用的 RGB 色彩空间并不能很好地反映视觉上的颜色感知性, 两个彩色点的 RGB 空间距离也不能反映其感知度的相似性. 故将 RGB 空间转

化为 HSV 空间来衡量, 每个像素点的值用一个三维模型来表示: $I(i) = [H(i), S(i), V(i)]^T$, 其中, H (hue) 是色度, S (saturation) 是饱和度, V (value) 是亮度值.

3.1 颜色模型的建立

将当前 HSV 彩色图像离散化为 m 个区域, 在实验中为 $8 \times 8 \times 8$ 区域(为了降低光线的敏感度, 也

可以采用 $8 \times 8 \times 4$ 区域). $h(i)$ 是图像在点 i 处的像素值对于上述 m 个离散化区域的映射. 为了降低计算的复杂性, 本文采用了运动检测时的矩形框区域来进行跟踪. 对此矩形框, 我们设定了一个 3 维数据对变量:

$$S = \{ (x, y), ([\dot{x}, \dot{y}]), (H_x, H_y) \}, \quad (13)$$

其中, (x, y) 是当前像素点的坐标, (\dot{x}, \dot{y}) 是当前像素点在 x, y 方向的变化率, (H_x, H_y) 代表矩形框的长与宽的半轴长. 实验所跟踪的目标是人, 故将矩形框的长短半轴比例限定在 $H_x < H_y \leq 4H_x$ 范围内.

检测时, 为了让可靠的像素点有较大的权重, 引入了如下的权重函数:

$$w(i) = (1 - r_i^2) \times D(i) \quad (14)$$

$$r_i = \max\left\{ \frac{x}{1.1 \times H_x}, \frac{y}{1.1 \times H_y} \right\}, \quad (15)$$

$$D(i) = \begin{cases} \log(D_i(i) - THRE), & \text{if } D_i(i) > THRE, \\ 0, & \text{otherwise,} \end{cases} \quad (16)$$

其中, r_i 为方框内像素点 i 远离矩形中心点的归一化距离, x 是 i 的横坐标到当前矩形框中心的水平距离, y 是 i 的纵坐标到跟踪矩形框中心的垂直距离, $D_i(i)$ 由式(3)决定. 因子 $(1 - r_i^2)$ 能够在一定程度减少方框边界像素点的权重, 因为它们可能属于背景或者已被遮挡, 从而增强颜色分布模型的可靠性; 因子 $D(i)$ 与差分图像中的像素点相关, 方框内像素点值的差分值越大, 它属于前景目标的可信度就越高, 其中 $THRE$ 是阈值, 是为了克服噪声而设, 它的值以等于或稍小于由上面第 2.2 节最小交叉熵决定的差分图像的阈值为宜. 实验中依据经验设定为固定阈值 $THRE = 8 \sim 16$. 当然, 为了增强算法的自适应性, 可以周期性地利用第 2.2 节的方法更换设定的阈值. 由上面的权重函数, 采用如下的公式计算区域像素的颜色分布 $P = \{ P^{(u)} \}_{u=1,2,\dots,m}$, 其中 I 是区域中像素点的总个数:

$$P^{(u)} = \frac{1}{f} \sum_{i=1}^I \{ w(i) [h(i) - u] \}, \quad (17)$$

其中, $\frac{1}{f}$ 是归一化因子, $f = \sum_{i=1}^I w(i)$, 是 Kro-necker delta 函数.

考虑到人运动的速度, 认为人的运动在视频中不会超越某个范围. 基于这个假设, 跟踪时, 以矩形框的中心为跟踪点, 预测位置在上一帧的邻域搜索. 在实验中, 采用了 7×7 模板, 并且矩形框以每帧最

大 5% 的放缩面积来进行搜索. 为了度量假设状态与当前状态的相似度, 引入 Bhattacharyya 系数^[16]来衡量两分布 $p(u)$ 与 $q(u)$ 的相似度. 考虑到离散密度, 例如颜色直方图 $p = \{ p^{(u)} \}_{u=1,2,\dots,m}$ 和 $q = \{ q^{(u)} \}_{u=1,2,\dots,m}$, 系数被定义为^[15]

$$[p, q] = \sum_{u=1}^m \sqrt{p^{(u)} q^{(u)}}, \quad (18)$$

越大, 则分布的相似度越高. 对两个同样的标准化直方图而言, $[p, q] = 1$ 表示一个完美的匹配. 因此, 在 7×7 模板中找出此最佳匹配, 即 $[p, q]$ 系数最大的匹配认为是下一状态的目标位置.

3.2 颜色模型的更新

对于非刚性目标, 它不可能长时间保留在同一状态, 甚至短时间内也可能发生很大的变化, 所以在跟踪期间就不可避免地要实行模型的更新. 然而这种更新必须考虑到发生遮挡时的情况, 如果发生遮挡, 必须保持目标的颜色分布不改变. 故颜色模型更新公式为

$$q_t^{(u)} = \begin{cases} (1 - \alpha) q_{t-1}^{(u)} + P_{\max(\cdot)}^{(u)}, & \text{if } (\max(\cdot)) > THRE, \\ q_{t-1}^{(u)}, & \text{otherwise,} \end{cases} \quad (19)$$

其中, q_t 为 t 时刻目标模型, $P_{\max(\cdot)}$ 为 t 时刻的预测的最佳状态, $\max(\cdot)$ 是其对应的匹配度, 即最佳匹配度, $THRE$ 是设定的阈值, 在实验中设定为 0.97, α 是更新速度因子, 实验中设为 0.01 ~ 0.05;

4 运动目标的检测、跟踪结果及分析

为了检验算法的有效性, 对多个图像序列进行了测试. 实验所用的图 1 ~ 3 是实地摄制的 320×240 像素的 AVI 视频, 图 4, 5 是标准的 MPEG 视频. 运动视频的跟踪是假定在初始化检测时, 该人是独立并且完整的, 没有别的人与之混叠或遮挡.

图 1 是在光线很好的情况下录制的足球场视频序列 1, 共 500 帧, 其中, 初始化帧 $N = 200$. 需要指出的是, 图像的初始跟踪必须以准确检测目标为前提, 除去初始化背景帧外, 只在跟踪的第 1 帧需要检测的结果, 其余的跟踪序列帧不需要检测. 一旦检测到目标系统开始跟踪.

图 1(a), (b) 是跟踪序列相应的检测与分割图, 由图 1(a), (b) 序列可以看出, 检测性能良好. 为了简化计算复杂度, 在检测时按灰度图像处理, 依靠图像分割的后处理, 分割的结果比较完整. 图 1(c) 列出了两

相似目标交叉时的跟踪结果图,图中,球员甲(红色方框)和球员乙(绿色方框)在彩色图上的区分度是很小的,但系统在两人交叉时仍实现了成功的跟踪

图 2 是 350 帧的足球场视频序列 2,其中,初始化帧 $N = 200$. 系统在跟踪时,将图像序列的边缘设为检测范围,一旦新目标出现,将跟踪它. 我们记录了整个跟踪过程,先检测然后实现跟踪. 在初始化检测后,跟踪到两个目标甲(红色方框)和乙(蓝色方框). 当第

3 个目标丙(绿色方框)出现,在边缘检测到它,然后实现跟踪. 帧 242 至帧 255 是目标丙逐渐被目标甲遮挡的情况;帧 255 中,目标丙被目标甲部分遮挡,不过头部仍清晰可见,故绿色方框中心定位于头部附近;帧 273 至帧 285 是目标丙逐渐遮挡目标乙的情况;帧 279 中,两目标基本重合,但该算法的定位依然比较准确. 图 2 序列显示该算法对在颜色分布图上有一定区分度的目标有很好的跟踪性能.

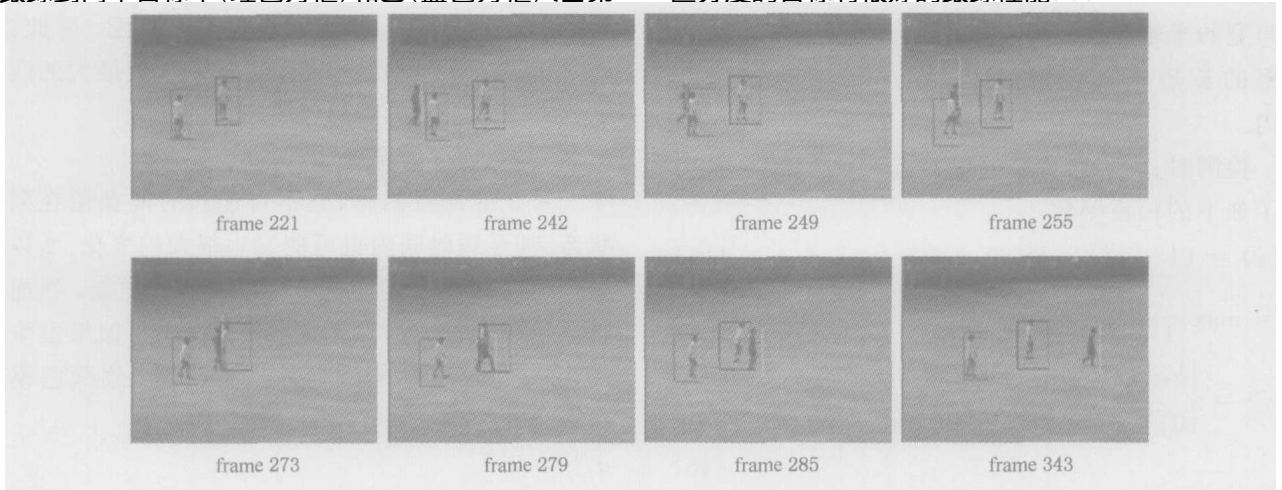


Fig. 2 Football sequence 2: The color distributions are discrepant when they interact with each other.

图 2 在颜色分布图上有较大差异时目标交叉时的跟踪

图 3 是摄制于运动场的视频. 场景左上角的树叶有较大的晃动,但在初始化中我们已去除了虚假目标. 以第 1 帧为初始化背景帧,目标从远到近,该算法仍能准确地跟踪它. 图 4 是标准的 Hall-monitor 视频图像,前 30 帧为初始背景帧. 当目标完全出现后开始跟踪,第 39 帧是跟踪的特定的人;在 99 帧

中,它弯下腰,形体发生一定的变化,然而却并不影响颜色的统计分布,虽然所设的方框并没有完全框住他,但不影响跟踪效果;实验 3,4 显示出该跟踪算法对非刚性物体的尺度变换和目标变形具有较好的鲁棒性.

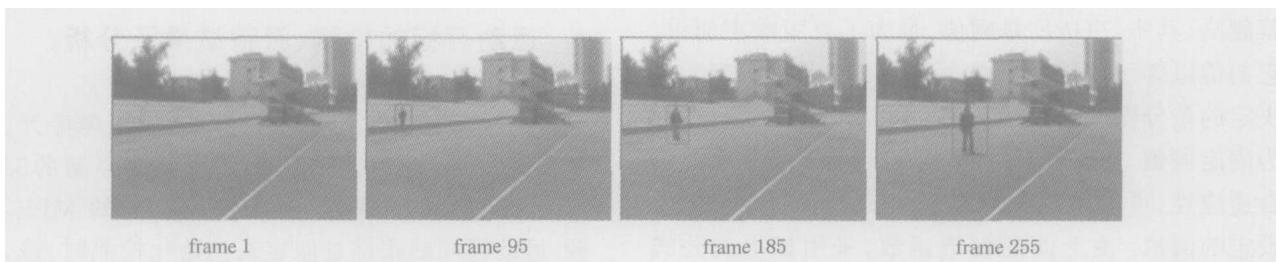


Fig. 3 Tracking for scale invariant, and the background is not completely stationary (motion of tree branches).

图 3 运动目标尺度变换时的跟踪:场景左上角的树叶有较大的扰动.

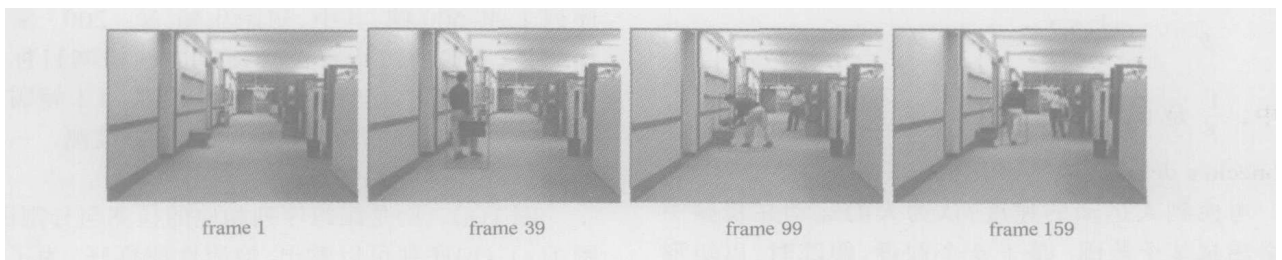


Fig. 4 Hall-monitor sequence: Tracking for shape transformation.

图 4 Hall-monitor 标准视频:目标变形时的跟踪

图 5 是标准的 Laboratory 的 MPEG 视频,首先定位出目标的某一个有特殊性的部位(比如人的头部),然后开始跟踪.在帧 311 时,所跟踪的人打开柜子,此时柜子区域在检测权重上将完全等同于前景,而且其颜色分布与头部相似,虽然帧 311 中人的头部的跟踪定位有一定的偏差,但是从整个跟踪流程看,实验 5 的跟踪结果良好.这种局部跟踪特性对于那些有明显局部颜色特征的目标尤为有效.

与粒子滤波器算法(particle filter)^[5~7]相比,粒子滤波器是基于序贯 Monte Carlo 滤波方法中最具

有代表性的方法之一,它的收敛性在很大程度上依赖于粒子数的多少,因此计算复杂度大,而该算法计算简单,能准确进行多目标跟踪;与均值位移(mean shift)算法^[11]相比,该算法利用目标的检测结果进行加权,有效地限制了跟踪区域,其跟踪效果相对良好.本文所提出的算法实现简单,它在目标邻域搜索,且跟踪区域像素权重基于像素置信度与空间位置,故在一定程度上能跟踪相似目标,如图 1(c)所示.

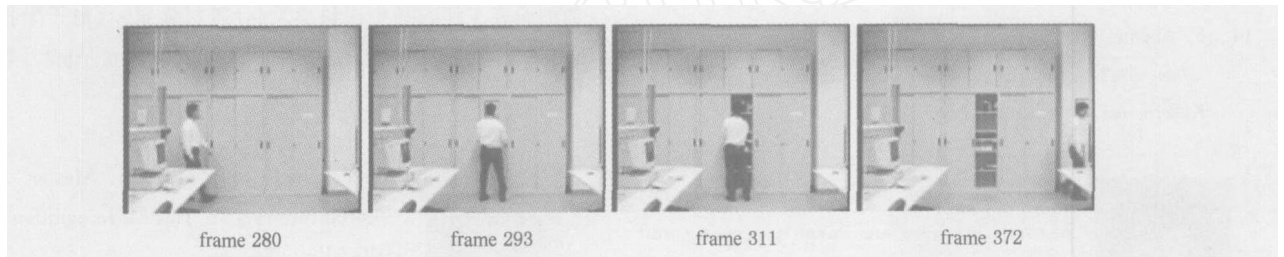


Fig. 5 Laboratory sequence: Tracking for local character (the head of the person).

图 5 Laboratory 标准 mpeg 视频:图像局部(头部)跟踪

5 结束语

针对复杂背景下视频图像运动目标的跟踪问题,采用先检测后跟踪目标的方式,在假设初始化检测时目标不发生混叠、交叉且无遮挡的情况下,目标的特性必须有一个很好的描述,故在初始化检测时采用基于时间域的中值建模与基于空间域的最小交叉熵相结合的方式对跟踪特性进行初始化.在初始化检测的基础上,本文利用 HSV 模型的颜色感知性优势,采用色彩分布模型,在考虑遮挡及背景与前景像素相似的情况下,提出了对跟踪矩形框内像素的加权函数,然后计算预测模型与目标模型之间的最大相似匹配度,将最大匹配概率模型作为下一步的目标模型,从而实现了有效的跟踪.与粒子滤波器跟踪方法相比,它能快速有效地跟踪相似的目标,与均值位移算法相比,它能有效地限制跟踪区域,跟踪效果更为良好.实验显示,该算法计算简单,对相似目标能实现准确的跟踪,能较好地解决多目标的交叉与部分遮挡问题,对于非刚性目标的尺度变化和在一定范围内的变形也有较好的鲁棒性.

参 考 文 献

1 D. P. Huttenlocher, G. A. Klanderma, W. J. Rucklidge.

- Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Analysis Machine Intelligence*, 1993, 15(9): 850 ~ 863
- 2 D. P. Huttenlocher, J. J. Noh, W. J. Rucklidge. Tracking non-rigid objects in complex scenes. *The 4th Int'l Conf. Computer Vision*, Berlin, Germany, 1993
- 3 I. Haritaoglu, D. Harwood, *et al.* W4: Real-time surveillance of people and their activities. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, 22(8): 809 ~ 830
- 4 C. Stauffer, W. E. L. Grimson. Learning patterns of activity using real-time tracking. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, 22(8): 747 ~ 757
- 5 Katja Nummiaro, Esther Koller-Meier, Luc Van Gool. An adaptive color-based particle filter. *Image and Vision Computing*, 2003, 21(1): 99 ~ 110
- 6 Y. Raja, S. McKenna, S. Gong. Tracking and segmenting people in varying lighting conditions using colour. In: *Int'l Conf. Face and Gesture Recognition*. Nara, Japan: IEEE Computer Society Press, 1998. 228 ~ 233
- 7 S. J. McKenna, Y. Raja, S. G. Gong. Tracking colour objects using adaptive mixture models. *Image and Vision Computing*, 1999, 17(2): 225 ~ 231
- 8 A. Jepson, D. Fleet, T. El-Maraghi. Robust online appearance models for visual tracking. *IEEE Conf. Computer Vision and Pattern Recognition*. Los Alamitos, CA: IEEE Computer Society Press, 2001. 415 ~ 422
- 9 D. Comaniciu, V. Wamash, P. Meer. Real-time tracking of non-rigid objects using mean shift. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition*. South Carolina: IEEE Computer Society Press, 2000. 142 ~ 149

- 10 E. Stringa, C. S. Regazzoni. Real-time video-shot detection for scene surveillance applications. *IEEE Trans. Image Processing*, 2000, 9(1): 69 ~ 79
- 11 K. Toyama, J. Krumm, B. Brumitt, *et al.* Wallflower: Principles and practice of background maintenance. In: *Proc. 7th IEEE Int'l Conf. Computer Vision*. Los Alamitos, CA: IEEE Computer Society Press, 1999, 1: 255 ~ 261
- 12 G. Welch, G. Bishop. An introduction to the Kalman filter. University of North Carolina at Chapel Hill, Technical Report: TR 95-041, 2001
- 13 M. Basseville. Distance measures for signal processing and pattern recognition. *Signal Processing*, 1989, 18(4): 349 ~ 369
- 14 F. Aherne, N. Thacker, P. Rockett. The Bhattacharyya metric as an absolute similarity measure for frequency coded data. *Kybernetika*, 1998, 34(4): 363 ~ 368



Shi Hua, born in 1979. Master. His research interests are image processing and pattern recognition.

施华, 1979年生, 工学硕士, 主要研究方向为图像处理与模式识别。



Li Cuihua, born in 1960. He received his Ph.D. degree in the Institute of Artificial Intelligence and Robotics from Xi'an Jiaotong University. He is currently professor and Ph. D. supervisor of the Computer Science Department at Xiamen University. His research interests include computer visions, image processing and analysis, wavelet transform theory and its applications etc. And he has published more than 60 papers in these fields.

李翠华, 1960年生, 工学博士, 教授, 博士生导师, 主要从事计算机视觉、图像处理与分析、小波变换理论及其应用等方面的研究工作, 在《中国科学》、《计算机学报》、《电子学报》、《生物物理学报》以及《Chinese J. Num. Math. & Appl.》等重要学术刊物上发表论文 60 余篇。



Wei Fengmei, born in 1976. Master. Her research interests are pattern recognition and artificial intelligence.

韦凤梅, 1976年生, 工学硕士, 主要研究方向为模式识别与人工智能。



Wang Hua wei, born in 1981. Master. His research interests are image processing and pattern recognition.

王华伟, 1981年生, 工学硕士, 主要研究方向为图像处理与模式识别。

Research Background

This paper is supported by the National Natural Science Foundation of China under Grant No. 60175008 and the National Innovation Research Group Foundation of China under Grant No. 60024301.

The efficient tracking of visual features in complex environments is a challenging task for the vision community, and it has been applied to many fields such as surveillance and monitoring, perceptual user interfaces etc. Many tracking algorithms work well under normal conditions. However, in some adverse conditions such as scale invariant, partial occlusion and interactions of non-rigid objects, these tracking algorithms don't provide satisfactory results. And further more, they are computationally complicated.

Considering the stationary video camera, we employ the moving object tracking algorithm based on location and confidence of pixels to perform the multi-object tracking. Firstly, the moving objects are detected accurately, and then the color distribution model is established for the detected area. In order to improve the computational efficiency, a weighting function based on the rectangle area is proposed, which takes into account both pixel confidence and pixel position. Our goal is to create a robust, adaptive tracking system to track the non-rigid object, so the color distribution model of foreground object is adapted to this change automatically.

In order to evaluate the performance of the algorithm, some experiments have been carried out. Some of the tracking scenarios are captured video with 320 × 240 resolution which are captured by a digital video camera fixed on a tripod, others are standard MPEG-4 video sequences. Experiments show that the system provides robustness to scale invariant, partial occlusion and interactions of non-rigid objects.