

运动图像序列中多目标跟踪的研究与实现

张 嫣, 姚耀文, 唐华松

(华南理工大学 计算机系, 广东 广州 510640)

摘 要: 针对运动图像序列中多个目标如何建立关联, 对进入和离开跟踪窗口的目标如何计数进行研究。提出了一种代价函数, 设置目标链记录多个目标的特征和运动状态, 搜索最佳匹配目标, 使正确轨迹得以延续, 实现了对被跟踪目标正确的统计和处理。

关键词: 运动连续性; 卡尔曼滤波; 目标链; 图像序列

中图法分类号: TP391.41

文献标识码: A

文章编号: 1001-3695(2002)01-0074-03

The Study and Implementation on Tracing Multi-object in Image Sequences

ZHANG Yan, YAO Yao-wen, TANG Hua-song

(Dept. of Computer Science, South China of University Technology, Guangzhou Guangdong 510640, China)

Abstract: This paper aims at how to establish relevant relationship for multi-object. It studies how to take count of targets entering or leaving the tracking-window. Firstly, this paper provides a cost function and establishes object-chain to note tracked targets' character and state. Secondly, It searches the furthest matching targets and continues the right tracks. Finally, some exceptive problems are discussed in order to dispose the tracked targets rightly.

Key words: Motion Continuity; Kalman Filter; Object-chain; Image Sequences

1 运动目标跟踪的前景与意义

运动目标跟踪是计算机视觉研究的一个核心课题, 它在医学研究、交通监视、客流量统计、天文观测等领域有着非常重要的实用价值, 有着广阔的发展前景。它的目的就是通过对传感器拍摄到的图像序列进行分析, 计算出目标在每帧图像上的二维位置坐标, 并根据不同的特征值, 将图像序列中不同帧中同一运动目标关联起来, 得到各个运动目标完整的运动轨迹。

运动目标跟踪是个复杂的问题, 而其中的关键是目标匹配。现有的运动目标跟踪方法各有千秋, 不同的跟踪算法适用于目标不同的运动状态。本文提出了一种代价函数, 采用窗口质心跟踪方法, 对多个目标进行跟踪, 建立每个被跟踪目标的“目标链”, 保证了跟踪的稳定性和精度。

2 运动目标跟踪的框架

本文重点研究二维图像序列多运动目标的跟踪。首先利用卡尔曼滤波对当前已被跟踪目标的下一步运动进行预测, 以缩小目标搜索匹配的范围, 加快目标匹配的速度。在下一帧指定范围内进行目标匹配, 以建立目标的关联关系。可将跟踪算法划分为两大模块: 预测模块和匹配更新模块。

预测模块主要利用卡尔曼滤波估计运动目标的特征值(运动参数, 如: 包围窗口的质心位置、运动速度、包围窗口的长宽等), 预测下一帧中运动目标可能出现的区域; 匹配更新模块则针对相邻帧间运动目标的变化, 利用特征值计算代价函数值, 求出当前帧中运动目标(已被跟踪目标)在下一帧中的对应后继目标, 建立关联关系, 并更新已被跟踪运动目标的“目标链”。跟踪算法的模块结构关系如图 1 所示。

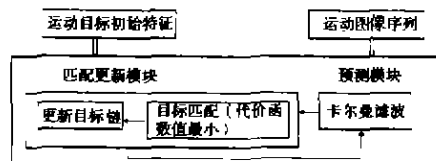


图 1 跟踪算法模块结构关系图

3 代价函数的提出及其适用情况

3.1 图像序列中目标运动的特点

由于图像序列中每两帧(相邻帧)时间间隔 Δt 比较小, 所以可通过人为地控制采样频率得到, 物体的运动不可能突变, 在 Δt 较小的情况下, 可认为同一运动目标在相邻两帧图像的质心位置和包围窗口变化不大, 即运动具有连续性。由于人的运动接近匀速运动, 有一个渐变过程, 所以提出代价函数的依据就是: 对应运动目标在相邻帧上质心间距离和包围窗口目标连通域面积变化不大。

例如:设采样得到一组图像运动序列, $F = \{f_1, f_2, \dots, f_n\}$, 则可对每帧图像建立相同的二维坐标系。设第 n 帧图像 f_n 中 m 个目标的连通域 $S_n^1, S_n^2, \dots, S_n^m$, 用链表存放每帧的所有目标连通域的特征值, 其中连通域 S_n^j 的特征值有下列参数: $X_n^j, Y_n^j, L_n^j, W_n^j$, 而 X_n^j, Y_n^j 分别是第 n 帧图像上第 j 个目标的质心横坐标、纵坐标, L_n^j, W_n^j 分别是其包围窗口的长度和宽度, 另外在图像序列任意一帧上都有可能新目标出现, 旧目标消失的情况。

3.2 代价函数的表达

代价函数形式如下:

$$V(t, j) = \alpha \cdot D(t, j) + \beta \cdot A(t, j) \quad (1)$$

$$D(t, j) = \frac{|C_t^j C_{t+1}^j|}{\max_k |C_t^k C_{t+1}^k|} \quad A(t, j) = \frac{|S_t^j - S_{t+1}^j|}{\max_k |S_t^k - S_{t+1}^k|} \quad (2)$$

其中: $|C_t^j C_{t+1}^j| = \sqrt{(x_t^j - x_{t+1}^j)^2 + (y_t^j - y_{t+1}^j)^2}$, $|S_t^j - S_{t+1}^j| = |L_t^j \cdot W_t^j - L_{t+1}^j \cdot W_{t+1}^j|$, $x_t^j, y_t^j, L_t^j, W_t^j$ 分别是图像序列第 t 帧第 j 个运动目标的包围窗口的质心横坐标、纵坐标、长度和宽度, $1 \leq n \leq \text{num}(t+1)$, $\text{num}(t+1)$ 是第 $t+1$ 帧上运动目标的个数。这里的 $D(t, j)$ 反映了第 t 帧上第 j 个目标与第 $t+1$ 帧上第 j 个目标包围窗口质心距离的大小, 其值越小说明两目标质心越靠近, 而 $A(t, j)$ 反映了第 t 帧上第 j 个目标与第 $t+1$ 帧上第 j 个目标的包围窗口面积的形变程度, 其值越小, 说明这两目标形状越接近, 代价函数值越小, 这两个目标有对应关系的可能性越大。可取 $\alpha = 0.8, \beta = 0.2$ 。

4 运动目标跟踪算法

4.1 预测过程确定搜索区域

跟踪算法的具体实现过程是: 首先在查找匹配过程中, 利用卡尔曼滤波来缩小搜索范围, 例如: 根据第 t 帧第 i 个目标的特征值作为卡尔曼滤波的输入参数, 得到它的预测值, 算出查找区域。然后在第 $t+1$ 帧指定区域中搜索, 依次计算目标 i 与 $t+1$ 帧中各个目标的代价函数值, 找出最小值 (假设是目标 j 与 i 的代价函数值最小), 则说明目标 j 是 i 的后续, 更新该目标的特征值作为下次卡尔曼滤波器的输入。

我们设计了一个线性状态模型:

$$S(t) = 9 \cdot S(t-1) + w(t-1) \quad (3)$$

其中 $S(t)$ 表示了在第 t 帧图像的系统状态, θ 则是状态转换矩阵, $W(t-1)$ 是估计误差, $S(t)$ 是一个包含运动目标位置和大小八维向量, 如式 (4) 所示:

$$S(t) = \begin{bmatrix} \Delta x(t) \\ \Delta y(t) \\ x_s(t) \\ y_s(t) \\ \Delta x'(t) \\ \Delta y'(t) \\ x_s'(t) \\ y_s'(t) \end{bmatrix} \quad W(t) = \begin{bmatrix} W_{\Delta x(t)} \\ W_{\Delta y(t)} \\ W_{x_s(t)} \\ W_{y_s(t)} \\ W_{\Delta x'(t)} \\ W_{\Delta y'(t)} \\ W_{x_s'(t)} \\ W_{y_s'(t)} \end{bmatrix} \quad (4)$$

式 (4) 中 $\Delta x, \Delta y$ 是目标包围窗口质心在单位时间间隔 Δt 内的位移量, $\Delta x', \Delta y'$ 是单位时间间隔 Δt 内的位移变化速率, x_s, y_s 是目标包围窗口的长度和宽度, $x_s'(t), y_s'(t)$ 是时间间隔 Δt 内包围窗口长宽的变化速率, 设目标

以恒定的加速度运动, 并且目标大小的变化是线性的, 则可得到状态转换矩阵为式 (5):

$$\hat{\theta} = \begin{bmatrix} 1 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

卡尔曼滤波算法是基于一系列假设来预测系统状态的, 假设系统状态与观测值之间是线性关系, 则:

$$m(t) = H \cdot s(t) + v(t) \quad (6)$$

这里 $m(t)$ 表示测量向量, H 则是观测矩阵, $v(t)$ 是观测误差, 其中:

$$m(t) = \begin{bmatrix} \Delta x \\ \Delta y \\ x_s \\ y_s \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

到此为止已经定义了系统的状态方程和观测方程, 通过卡尔曼滤波算法获得运动参数的估计, 利用 $\Delta x, \Delta y$ 可以计算得到搜索区域的中心, x_s, y_s 则可以得到搜索范围:

$$\begin{aligned} w_x &= x_s \cdot 1.5, & w_y &= y_s \cdot 1.5 \\ x &= x_t^{(m)} + \Delta x, & y &= y_t^{(m)} + \Delta y \end{aligned} \quad (8)$$

即搜索范围为:

$$x_t - w_x/2 \leq x \leq x_t + w_x/2, \quad y_t - w_y/2 \leq y \leq y_t + w_y/2.$$

4.2 建立各运动目标对应关系和更新特征值

设实验研究是针对通过某一区域 (例如门) 的人的运动 (进入或走出), 通过实验取得数据 x_m, x_{out} 将整个窗口分为三个区域: 1, 3 为预留区, 2 为进入区 (如图 2 所示)。当目标出现在 1, 3 区域时, 状态可能为刚进入或将要离开; 在区域 2 的目标说明已经完全进入跟踪区域。通过对图像序列中同一目标跟踪计算得到: 在相邻帧时间间隔 Δt 内同一目标运动的质心距离上限为 λ 。

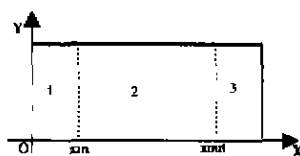


图2 跟踪区域坐标系

该算法的难点也就在于如何将相邻帧上同一运动目标建立对应关系, 不致产生误配。本文的具体处理过程为: 对各个帧中的目标进行识别分割得到各连通区域 (目标包围窗口), 再根据各个被跟踪目标的目标链, 对多个目标进行匹配, 建立目标的对应关系, 然后刷新目标链, 以便后继图像跟踪使用, 并保留运动目标的运动轨迹。这里的目标链记录了目标的特征和运动状态, 便于随时掌握目标的运动情况, 目标链是由目标特征组成, 在此选择为包围窗口的质心位置, 包围窗口的长度和宽度, 窗口质心在相邻帧中的距离。为每个出现在跟踪窗口的目标设置一个目标链, 且保存的是该目标最新图像帧中的特征值, 目标离开跟踪窗口则消除其对应的目标链, 下面是跟踪算法步骤:

(1) 图像序列中第一帧出现的目标全部作为新出现目标, 为每个目标建立目标链、启动进出两个计数器, 并规定沿 x 方向为进, 逆 x 方向为出;

(2) 接收后续图像第 k 帧, 则对当前第 i 个目标链 ($1 \leq i \leq n$, n 为已被跟踪目标数目), 利用其目标链中的特征值启动卡尔曼滤波预测搜索匹配范围, 计算目标 i 与第 k 帧指定范围内所有目标的代价函数值, 并找出其中最小值 (设找到与目标 j 代价函数值最小), 再判断目标 i 与目标 j 的质心距离 d 与 λ 的大小:

1) 如果 $d \leq \lambda$, 则目标 j 为目标 i 的后续, 将目标 j 的特征值替代目标链 i 的值, 并对该目标 j 做标记;

2) 如果 $d > \lambda$, 则说明目标 j 在第 k 帧没有后续目标, 这可能有两种情况: 离开观测窗口或者暂时静止。需判断目标链 i 的质心坐标: ① 当 $x_1 \leq x_m$, 如果该目标运动轨迹的第一个质心横坐标 $x_1 \leq x_m$, 则说明该目标一直在窗口内徘徊, 任何计数器不计数; 如果 $x_1 > x_m$, 则说明目标离开跟踪窗口, 方向为出, 出计数器加 1, 并清除目标链 i ; ② 当 $x_1 > x_m$, 如果该运动目标轨迹第一个质心坐标横坐标 $x_1 \geq x_m$, 任何计数器不计数; 如果 $x_1 < x_m$, 则说明该目标离开跟踪窗口, 方向为进, 进计数器加 1, 并清除目标链 i ; ③ 当 $x_1 < x_m$, 则说明目标 j 暂时静止, 保留其特征值, 等待重新运动;

(3) 当对所有被跟踪目标进行匹配后, 检查该第 k 帧图像上所有目标是否被标记, 如果全部目标都被做了标记, 则说明该帧上所有目标与前帧目标全部建立了后续关系, 如果有目标未被做标记, 则判断其质心横坐标是否满足: $x < x_m$ 或者 $x > x_m$, 满足则是有新目标出现, 为其建立目标链, 设置特征值, 不满足则可能为干扰, 丢弃;

(4) 收第 $k+1$ 帧图像, 处理过程如 (2), (3), 直至图像序列结束。

4.3 算法的改进

对于 4.2 中的算法可能会有下面情况的发生, 假设对于目标链 i 在第 $k+1$ 帧中搜索范围进行匹配, 找到目标 j 与之代价函数值最小, 即 $V(i, j) = \min(V(i, s)), (1 \leq s \leq n, n \text{ 为 } k+1 \text{ 帧中目标个数})$ 。而对于目标链 m 在该帧中搜索匹配也恰好是目标 j 与之代价最小, 即 $V(m, j) = \min(V(m, s)), (1 \leq s \leq n \text{ 且 } s \neq j)$ 。如前所示, 此时对目标 j 重复标记。可作如下处理:

判断目标链 i 和目标链 m 的质心与目标 j 质心间距的大小, 如果 $|C_i C_j| < |C_m C_j|$, 则将目标 j 视为目标 i 的后续, 而目标链 m 重新搜索匹配, 找到 $V(m, k) = \min(V(m, s)), (1 \leq s \leq n \text{ 且 } s \neq j)$, 则目标 k 是目标 m 的后续。在该运动跟踪算法中, 应该考虑运动目标暂时静止不动情况, 保留其目标链等待其重新运动, 以避免产生目标丢失。另外为正确的判断出新目标的出现, 老目标从跟踪区域消失的情况, 以及目标在跟踪区域徘徊的情况, 设置了预留区以保证较高的跟踪正确性。

5 实验结果

实验采用的图像分辨率为 160×120 , 即跟踪窗口的最大横坐标为 160, 摄像头 5 帧/秒, 经过反复实验确定 x_m 为 30, x_{\max} 为 130, λ 为 18

采样的一系列运动图像序列如图 3 所示。

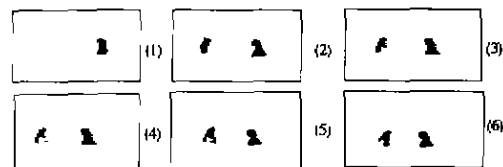


图 3 运动图像序列

表 1 是运动图像序列中运动目标的包围窗口面积和质心坐标。

表 1 目标质心坐标表

序号	目标 1			目标 2		
	Area	centerx	centery	Area	centerx	centery
1	344	101	55			
2	383	87	55	245	14	50
3	379	85	55	249	16	51
4	383	83	55	238	17	51
5	380	80	55	270	20	65
6	380	78	55	255	21	65

根据本文提出的跟踪算法, 得到运动轨迹为:

目标 1: $(101, 55), (87, 55), (85, 55), (83, 55), (80, 55), (78, 55)$;

目标 2: $(14, 50), (16, 51), (17, 51), (20, 65), (21, 65)$;

如图像发生交叉的情况时, 通过连通域划分, 然后利用代价函数搜索匹配, 可保证图像目标的正确跟踪。

6 结束语

本文利用运动目标在相邻帧间运动具有连续性并且包围窗口变化不大的特点, 提出了一种代价函数, 应用于图像跟踪的匹配过程中, 同时启动卡尔曼滤波, 减小了匹配搜索区域, 使用目标链记录了目标最新最近的运动状态和特征值, 保证了运动跟踪的连续性, 在跟踪过程中考虑了目标暂时静止消失的情况, 并对图像序列中间帧出现新目标进行了正确处理, 实现了多目标的正确的跟踪, 进而对运动目标是否在跟踪区域徘徊进行判断, 保证了计数的正确性, 取得了较令人满意的效果。

参考文献:

- [1] 李智勇. 动态图像分析[M]. 北京: 国防工业出版社, 1999.
- [2] 周宏仁, 敬忠良. 机动目标跟踪[M]. 北京: 国防工业出版社, 1991.
- [3] 王宏禹. 随机数字信号处理[M]. 北京: 科学出版社, 1988.
- [4] 姜建国, 曹建中, 高玉明. 信号与系统分析基础[M]. 北京: 清华大学出版社, 1994.
- [5] 汪亚明, 梅正国, 宁昕, 等. 一种非刚体运动图像序列的特征点对应方法[J]. 中国图像图形学报, 2000, (5): 232-236.

作者简介:

张婷(1977-), 女, 计算机专业硕士研究生, 研究方向为计算机应用; 姚耀文(1945-), 男, 副教授, 计算机专业硕士导师, 研究方向为计算机应用; 唐华松(1976-), 女, 计算机专业硕士研究生, 研究方向为计算机应用。