



## (12) 发明专利申请

(10) 申请公布号 CN 104574439 A

(43) 申请公布日 2015. 04. 29

(21) 申请号 201410819971. 9

(22) 申请日 2014. 12. 25

(71) 申请人 南京邮电大学

地址 210023 江苏省南京市亚东新城区文苑  
路 9 号

(72) 发明人 朱松豪 刘佳伟 胡荣林

(74) 专利代理机构 南京知识律师事务所 32207

代理人 汪旭东

(51) Int. Cl.

G06T 7/20(2006. 01)

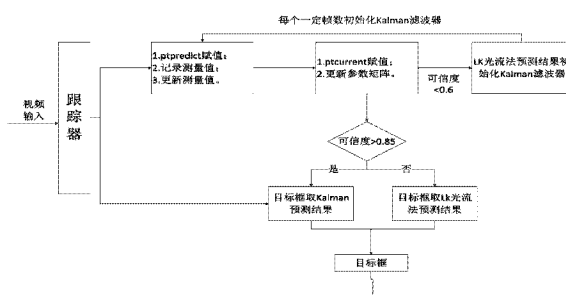
权利要求书3页 说明书8页 附图2页

### (54) 发明名称

一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法

### (57) 摘要

本发明公开了一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,该方法包括以下步骤:首先分析卡尔曼滤波器的原理,并对其加以改进;然后,利用改进的卡尔曼滤波器增强 TLD 算法,提高系统的可靠性;最后,利用改进的卡尔曼滤波器所得到的预测结果,对检测模块进行改进,缩小 TLD 算法的检测区域,进一步提高跟踪的实时性。



1. 一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,所述方法包括如下步骤:

步骤 1:分析和利用卡尔曼滤波器的原理,并对其加以改进;

卡尔曼滤波器用于建模的系统是线性系统;预测和测量的噪声都是白噪声,噪声服从高斯分布;

步骤 2:利用改进的卡尔曼滤波器增强 TLD 算法;

通过改进的卡尔曼滤波来增强 TLD 算法,即根据目标位置的时间步长,获得同步更新的参数传递矩阵;

步骤 3:利用改进的卡尔曼滤波器所得到的预测结果,对检测模块进行改进,缩小 TLD 算法的检测区域;

根据卡尔曼滤波的功能,预计当前帧中目标存在的区域,将预计区域作为 TLD 算法的目标检测区域;

通过改进卡尔曼滤波器的预估结果来设置子窗口的位置范围。

2. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,所述方法步骤 1 包括:第一条条件是通过  $k-1$  时刻的系统状态与一个参数矩阵的乘积得到  $k$  时刻的系统状态;其余两个条件表示噪声是不随时间变化而变化,用协方差和均值就能准确地为幅值建模;

在实时跟踪过程中,两帧间的时间间隔非常小,相邻帧的目标运动可视为成线性;这样就满足卡尔曼滤波三大条件中的第一个条件,利用卡尔曼滤波估计视频序列中目标的运动状态,运动估计过程包括如下步骤:

(1) 运动预测阶段;

假设运动目标的坐标为  $(x, y)$ ,运动速度为  $(v_x, v_y)$ ,则  $k$  时刻的状态  $s_k$  为  $(x, y, v_x, v_y)^T$ ,测量状态为  $z_k$  为  $(x_k, y_k)$ ,  $A_k$  为系统状态预测方程中的传递矩阵为:

$$A_k = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

由于该系统中没有输入控制变量,因此  $B_k$  为 0,测量方程  $H_k$  中的参数矩阵为:

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (2)$$

卡尔曼滤波第二、三两个条件要求噪声必须是白噪声,且服从高斯分布,所以测量方程和预测方程中的噪声均为服从高斯分布的白噪声,它们的协方差矩阵  $P_k$ 、 $Q_k$  分布分别为:

$$P_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

系统的测量方程为:

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4)$$

系统的预测方程为：

$$\begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ v_{xk-1} \\ v_{yk-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

(2) 运动更新阶段；

利用预测阶段得到的系统估计  $P_k$  和系统状态  $x_k$ ，不断重复预测步骤，直到系统结束；

卡尔曼滤波在相对稳定的系统中，能取得预测效果；但当目标运动带有随机性时使用卡尔曼滤波器，得到的预测效果往往会出现较大的误差；需对卡尔曼滤波进行适当的改进；

卡尔曼滤波在得到上一时刻的状态后，通过状态转移矩阵决定当前时刻状态的预测值，同时状态转移矩阵与  $k$  时刻的状态估计也有关，相对准确的测量值能保证预测值的准确度；

若从  $k$  到  $k+1$  时刻中的某一时间  $k_t$  发生状态突变，则从  $k$  到  $k_t$  时间段的状态转移矩阵为  $k$  时刻的滤波值，但  $k_t$  到  $k+1$  时刻内的转移矩阵为  $k+1$  时间观测的状态；假设  $k$  到  $k+1$  时间段内  $k+1/2$  时刻为中间时刻，则有以下式：

$$X(k|k-1) = A\left(k, k-\frac{1}{2}\right) * \left[ A\left(k-\frac{1}{2}\right) * X(k-1) \right] \quad (6)。$$

3. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法，其特征在于，所述方法步骤 2 包括：(1) 初始化系统，将初始边框的中心位置赋予  $ptcurrent$  和  $ptpredict$ ，同时修改卡尔曼滤波器系统状态的初始数值和初始预测，并利用这两个数值初始化卡尔曼滤波器；(2) 通过卡尔曼滤波器对后续帧进行预测，且将该时刻预测的系统状态赋给  $ptpredict$ ；同时，将 LK 光流法跟踪得到的目标边界框的中心位置点赋给  $ptcurrent$ ；(3) 更新  $ptcurrent$  之前，用  $ptlast$  保存原先的系统状态测量；(4) 用更新系统的状态预测量和状态测量对卡尔曼滤波器进行纠正，并预测目标边框的中心位置点，采用原来有的目标边框的宽高比来还原出边框；(5) 利用卡尔曼滤波器预测每一帧目标边框的可信度，即计算目标边框与最近邻分类器中目标模板的相似度；若该值大于 0.85，则用该目标边框取代 LK 光流法的检测边框结果，并将该边框检测结果传递给后续的学习模块和检测模块；相反，则保留 LK 光流法所得的目标边框；若可信度小于 0.6 或满足一定帧数后，则用 LK 光流法的边框检测结果重新初始化卡尔曼滤波器。

4. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法，其特征在于，所述方法步骤 2 的 TLD 算法模块包括：跟踪模块、学习模块、检测模块，其中跟踪模块的功能是通过一个短周期的自适应跟踪器，在帧与帧之间运动的有限性、可见性的情况下，预测目标的运动趋势；学习模块的功能是评价跟踪模块和检测模块的性能，并利用算法生成的有效训练集来更新检测模块，消除检测误差；检测模块的功能是对目标进行实时检测，同时在必要时纠正跟踪器。

5. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,所述方法步骤 2 的 TLD 算法包括:首先,设定第一帧中目标的位置和大小,完成 TLD 算法的初始化;然后,在跟踪过程中,检测模块和跟踪模块共同对每一帧图像进行处理,其中跟踪模块依据前一帧中目标的位置信息来预测当前帧中同一目标的位置,检测模块则对当前帧进行全局扫描以检测目标,且融合跟踪结果和检测结果,给出当前帧是否存在运动目标、目标所在位置、以及跟踪轨迹是否有效等信息;最后,将检测结果和跟踪结果输入至学习模块,由学习模块决定是否对跟踪模块和检测模块进行更新。

6. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,所述方法步骤 3 包括:(1) 通过改进的卡尔曼滤波器估计当前帧中运动目标的中心位置;(2) 若某一矩形区域的长宽比与上一帧中目标边界框的长宽比一致,大小是上一帧边界框的 4 倍,即认为该矩形区域包含检测目标;(3) 求与所有与该划定矩形区域有交集的子窗口,然后将这一系列子窗口送入分类器,由分类器检测子窗口中是否包含目标。

7. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,所述方法步骤 3 所述的检测模块包括:方程分类器;随机森林分类器;最近邻分类器;每个阶段都将不满足条件的图像块去除,同时将满足条件的图像块送入下一个分类器。

8. 根据权利要求 1 所述的一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于:所述方法应用于卡尔曼滤波。

## 一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法

### 技术领域

[0001] 本发明涉及一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,属于图像处理技术领域。

### 背景技术

[0002] 运动目标的检测与跟踪一直都是计算机视觉研究领域中热点问题,在国防和民用的各个领域都能广泛地应用,如智能监控、航天航空、人工智能、导航制导等方面。在目标检测与跟踪过程中,会面临多个方面的挑战,如目标的遮挡、光线的变化、目标轨迹的改变、目标的运动速度过快、相似目标的出现、复杂的背景等。

[0003] 随着计算机性能地不断提升,处理视频图像等大规模数据的能力也随之不断提升,从而使目标检测与跟踪技术得到了极大发展。1950 年,Gibson 首先提出采用光流方法,实现由静态向动态的图像处理。随着研究的深入,光流法已很难满足系统对实时性提出的要求,由此,许多其它方法应运而生。如:Meanshift 算法以其无参数、计算量小的优势得到重视;Condesation 算法利用了粒子滤波的思想;VSAM 算法实现了对复杂城市进行全方位视频监控。

[0004] 总体而言,目前常见的目标检测与跟踪主要分为如下几类:(1) 基于知识的方法:虽然此类算法解决了统计模式方法的局限性,但同时也引入了一定的问题,如:验证困难,成本较大,在新的场景中需要重新定义与组织知识等。(2) 基于模型的方法:该类算法首先提取目标特征;然后,建立目标的空间模型;接下来,结合这些特征和其它参数对目标的特征进行筛选,建立有效的初始假设;最后,依据特性对目标进行预测,当标记达到特定阈值后,则认为匹配有效。(3) 基于统计模式的方法:此类算法首先对系统进行大规模的训练,统计出目标的特征分布;然后,以模式空间距离为标准,对特征匹配分类计数。该类算法主要应用在目标与周边背景单一的情况下。(4) 基于人工神经网络和专家系统的方法:此类算法借助于神经网络,解决传统算法不能解决的问题,但同时由于神经网络的实时性不佳,所以此类算法的实时性相对较差。而本发明能够很好地解决上面的问题。

### 发明内容

[0005] 本发明目的在于解决了现有技术的不足,提供了一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法,该方法应用于卡尔曼滤波,是利用一种改进的半监督学习机制,不断地对检测和跟踪模块进行更新,从而提高了系统的稳定性、鲁棒性、可靠性。

[0006] 本发明的融合卡尔曼滤波与 TLD 算法的目标跟踪方法,其特征在于,包括以下几个步骤:

[0007] 步骤 1:分析和利用卡尔曼滤波器的原理,并对其加以改进。

[0008] 卡尔曼滤波器用于建模的系统是线性系统;预测和测量的噪声都是白噪声,噪声服从高斯分布,第一条条件是通过  $k-1$  时刻的系统状态与一个参数矩阵的乘积得到  $k$  时刻的系统状态;其余两个条件表示噪声是不随时间变化而变化,用协方差和均值就能准确地

为幅值建模；

[0009] 在实时跟踪过程中，两帧间的时间间隔非常小，相邻帧的目标运动可视为成线性；这样就满足卡尔曼滤波三大条件中的第一个条件，利用卡尔曼滤波估计视频序列中目标的运动状态，运动估计过程包括如下步骤：

[0010] (1) 运动预测阶段；

[0011] 假设运动目标的坐标为  $(x, y)$ ，运动速度为  $(v_x, v_y)$ ，则  $k$  时刻的状态  $s_k$  为  $(x, y, v_x, v_y)^T$ ，测量状态为  $z_k$  为  $(x_k, y_k)$ ， $A_k$  为系统状态预测方程中的传递矩阵为：

$$[0012] \quad A_k = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

[0013] 由于该系统没有输入控制变量，因此  $B_k$  为 0，测量方程  $H_k$  中的参数矩阵为：

$$[0014] \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (2)$$

[0015] 卡尔曼滤波第二、三两个条件要求噪声必须是白噪声，且服从高斯分布，所以测量方程和预测方程中的噪声均为服从高斯分布的白噪声，它们的协方差矩阵  $P_k$ 、 $Q_k$  分布分别为：

$$[0016] \quad P_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

[0017] 综上所述，系统的测量方程为：

$$[0018] \quad \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4)$$

[0019] 系统的预测方程为：

$$[0020] \quad \begin{bmatrix} x_k \\ y_k \\ v_{xk} \\ v_{yk} \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ v_{xk-1} \\ v_{yk-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

[0021] (2) 运动更新阶段；

[0022] 利用预测阶段得到的系统估计  $P_k$  和系统状态  $x_k$ ，不断重复预测步骤，直到系统结束；

[0023] 卡尔曼滤波在相对稳定的系统中，能取得预测效果；但当目标运动带有随机性时使用卡尔曼滤波器，得到的预测效果往往会出现较大的误差；需对卡尔曼滤波进行适当的改进；

[0024] 卡尔曼滤波在得到上一时刻的状态后，通过状态转移矩阵决定当前时刻状态的预测值，同时状态转移矩阵与  $k$  时刻的状态估计也有关，相对准确的测量值能保证预测值的准确度；

[0025] 若从  $k$  到  $k+1$  时刻中的某一时间  $k_t$  发生状态突变, 则从  $k$  到  $k_t$  时间段的状态转移矩阵为  $k$  时刻的滤波值, 但  $k_t$  到  $k+1$  时刻内的转移矩阵为  $k+1$  时间观测的状态; 假设  $k$  到  $k+1$  时间段内  $k+1/2$  时刻为中间时刻, 则有以下式:

$$[0026] \quad X(k|k-1) = A\left(k, k - \frac{1}{2}\right) * \left[ A\left(k - \frac{1}{2}\right) * X(k-1) \right] \quad (6)$$

[0027] 步骤 2: 利用改进的卡尔曼滤波器增强 TLD 算法, 提高系统的可靠性。

[0028] 通过改进的卡尔曼滤波来增强 TLD 算法, 即根据目标位置的时间步长, 获得同步更新的参数传递矩阵。主要过程描述如下: (1) 初始化系统: 将初始边框的中心位置赋予  $pt_{current}$  和  $pt_{predict}$ , 同时修改卡尔曼滤波器系统状态的初始数值和初始预测, 并利用这两个数值初始化卡尔曼滤波器。(2) 通过卡尔曼滤波器对后续帧进行预测, 且将该时刻预测的系统状态赋给  $pt_{predict}$ ; 同时, 将 LK 光流法跟踪得到的目标边界框的中心位置点赋给  $pt_{current}$ 。(3) 更新  $pt_{current}$  之前, 用  $pt_{last}$  保存原先的系统状态测量。(4) 用更新系统的状态预测测量和状态测量对卡尔曼滤波器进行纠正, 并预测目标边框的中心位置点, 采用原来有的目标边框的宽高比来还原出边框。(5) 利用卡尔曼滤波器预测每一帧目标边框的可信度, 即计算目标边框与最近邻分类器中目标模板的相似度: 若该值大于 0.85, 则用该目标边框取代 LK 光流法的检测边框结果, 并将该边框检测结果传递给后续的学习模块和检测模块; 相反, 则保留 LK 光流法所得的目标边框。若可信度小于 0.6 或满足一定帧数后, 则用 LK 光流法的边框检测结果重新初始化卡尔曼滤波器。

[0029] 步骤 3: 利用改进的卡尔曼滤波器所得到的预测结果, 对检测模块进行改进, 缩小 TLD 算法的检测区域, 进一步提高跟踪的实时性。

[0030] 根据卡尔曼滤波的功能, 预计当前帧中目标存在的区域, 将预计区域作为 TLD 算法的目标检测区域。

[0031] 通过改进卡尔曼滤波器的预估结果来设置子窗口的位置范围, 具体过程如下所述: (1) 通过改进的卡尔曼滤波器估计当前帧中运动目标的中心位置; (2) 若某一矩形区域的长宽比与上一帧中目标边界框的长宽比一致, 大小是上一帧边界框的 4 倍, 即认为该矩形区域包含检测目标; (3) 求与所有与该划定矩形区域有交集的子窗口, 然后将这一系列子窗口送入分类器, 由分类器检测子窗口中是否包含目标。

[0032] 卡尔曼滤波在得到上一时刻的状态后, 通过状态转移矩阵决定当前时刻状态的预测值, 同时状态转移矩阵与  $k$  时刻的状态估计也有关, 所以相对准确的测量值能保证预测值的准确度。

[0033] 有益效果:

[0034] 1、本发明的卡尔曼滤波在系统突变的时间段内增加了系统的采样率, 确保了预测结果的可信度, 且只有在发生突变的情况下采用该方法, 并且对系统的跟踪速度影响不大。

[0035] 2、本发明提高了系统的稳定性、鲁棒性、可靠性。

## 附图说明

[0036] 图 1 为本发明的 TLD 算法框架图。

[0037] 图 2 为本发明光流法的流程图。

[0038] 图 3 为本发明 Kalman 滤波增强 TLD 算法框架图。

[0039] 图 4 为本发明级联分类器示意图。

## 具体实施方式

[0040] 下面结合说明书附图对本发明创造作进一步的详细说明。

[0041] 如图 1- 图 4 所示, 本发明提出了一种融合卡尔曼滤波与 TLD 算法的目标跟踪方法, 该方法首先是分析和利用卡尔曼滤波器的原理, 并对其加以改进; 然后, 利用改进的卡尔曼滤波器增强 TLD 算法, 提高系统的可靠性; 最后, 利用改进的卡尔曼滤波器所得到的预测结果, 对检测模块进行改进, 缩小 TLD 算法的检测区域, 进一步提高跟踪的实时性。

[0042] 本发明的具体实施过程包括以下步骤:

[0043] 步骤 1: 分析和利用卡尔曼滤波器的原理, 并对其加以改进。

[0044] 卡尔曼滤波的实现必须要满足三个重要的条件: (1) 用于建模的系统是线性系统; (2) 预测和测量的噪声都是白噪声; (3) 噪声服从高斯分布。第一条条件是通过  $k-1$  时刻的系统状态与一个参数矩阵的乘积得到  $k$  时刻的系统状态; 其余两个条件表示噪声是不随时间变化而变化, 所以用协方差和均值就能准确地为幅值建模。虽然以上三个条件较为苛刻, 但是卡尔曼滤波在实际应用中非常广泛。

[0045] 由于在实时跟踪过程中, 两帧间的时间间隔非常小, 所以相邻帧的目标运动可视为成线性。这样就满足卡尔曼滤波三大条件中的第一个条件, 因此可利用卡尔曼滤波估计视频序列中目标的运动状态。具体的运动估计过程包括以下两个步骤:

[0046] (1) 运动预测阶段

[0047] 假设运动目标的坐标为  $(x, y)$ , 运动速度为  $(v_x, v_y)$ , 则  $k$  时刻的状态  $s_k$  为  $(x, y, v_x, v_y)^T$ , 测量状态为  $z_k$  为  $(x_k, y_k)$ 。  $A_k$  为系统状态预测方程中的传递矩阵为:

$$[0048] \quad A_k = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

[0049] 由于该系统没有输入控制变量, 因此  $B_k$  为 0。测量方程  $H_k$  中的参数矩阵为:

$$[0050] \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (2)$$

[0051] 卡尔曼滤波第二、三两个条件要求噪声必须是白噪声, 且服从高斯分布, 所以测量方程和预测方程中的噪声均为服从高斯分布的白噪声, 它们的协方差矩阵  $P_k$ 、 $Q_k$  分布分别为:

$$[0052] \quad P_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

[0053] 综上所述, 系统的测量方程为:

$$[0054] \quad \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4)$$



[0055] 系统的预测方程为：

$$[0056] \begin{bmatrix} \hat{x}_k \\ \hat{y}_k \\ \hat{v}_{xk} \\ \hat{v}_{yk} \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1} \\ \hat{y}_{k-1} \\ \hat{v}_{xk-1} \\ \hat{v}_{yk-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (5)$$

[0057] (2) 运动更新阶段

[0058] 利用预测阶段得到的系统估计  $P_k$  和系统状态  $\hat{x}_k$ , 不断重复预测步骤, 直到系统结束。

[0059] 卡尔曼滤波在相对稳定的系统中, 能取得很好的预测效果; 但当目标运动带有随机性时使用卡尔曼滤波器, 得到的预测效果往往会出现较大的误差。因此, 需对卡尔曼滤波进行适当的改进。

[0060] 卡尔曼滤波在得到上一时刻的状态后, 通过状态转移矩阵决定当前时刻状态的预测值, 同时状态转移矩阵与  $k$  时刻的状态估计也有关, 所以相对准确的测量值能保证预测值的准确度。

[0061] 假设从  $k$  到  $k+1$  时刻中的某一时间  $k_t$  发生状态突变, 则从  $k$  到  $k_t$  时间段的状态转移矩阵为  $k$  时刻的滤波值, 但  $k_t$  到  $k+1$  时刻内的转移矩阵为  $k+1$  时间观测的状态。本发明假设  $k$  到  $k+1$  时间段内  $k+1/2$  时刻为中间时刻, 则有以下式：

$$[0062] \quad X(k|k-1) = A\left(k, k-\frac{1}{2}\right) * \left[A\left(k-\frac{1}{2}\right) * X(k-1)\right] \quad (6)$$

[0063] 改进后的卡尔曼滤波在系统突变的时间段内增加了系统的采样率, 确保了预测结果的可信度, 且只有在发生突变的情况下采用该方法。因此, 对系统的跟踪速度影响不大。

[0064] 步骤 2: 利用改进的卡尔曼滤波器增强 TLD 算法, 提高系统的可靠性。

[0065] TLD 算法是一种高效的单目标长时间跟踪算法, 该算法的主要优点在于只需获取很少的先验信息, 就可对目标进行长时间的在线跟踪, 且实时性高、运算速度快、稳定性强, 还能适用于发生目标遮挡、消失等场景。正如其名一样, TLD 算法主要有三个模块组成: 跟踪模块 (Tracking)、学习模块 (Learning)、检测模块 (Detecting), 其中跟踪模块的功能是通过一个短周期的自适应跟踪器, 在帧与帧之间运动的有限性、可见性的情况下, 预测目标的运动趋势; 学习模块的功能是评价跟踪模块和检测模块的性能, 并利用算法生成的有效训练集来更新检测模块, 消除检测误差; 检测模块的功能是对目标进行实时检测, 同时在必要情况下纠正跟踪器。

[0066] TLD 算法的框架如图 1 所示。首先, 设定第一帧中目标的位置和大小, 完成 TLD 算法的初始化; 然后, 在跟踪过程中, 检测模块和跟踪模块共同对每一帧图像进行处理, 其中跟踪模块依据前一帧中目标的位置信息来预测当前帧中同一目标的位置, 检测模块则对当前帧进行全局扫描以检测目标, 且融合跟踪结果和检测结果, 给出当前帧是否存在运动目标、目标所在位置、以及跟踪轨迹是否有效等信息; 最后, 将检测结果和跟踪结果输入至学习模块, 由学习模块决定是否对跟踪模块和检测模块进行更新。

[0067] 本发明在原始 TLD 算法的基础上对其进行改进, 从而获得更为理想的跟踪效果。本发明所涉及的改进有以下两个方面: (1) 对跟踪模块进行改进, 基于改进的卡尔曼滤波器原理对跟踪器进行一定的改进。同时依据目标位置的每个时间步长, 计算并更新参数传

递矩阵 F ; (2) 对检测模块进行改进。根据卡尔曼滤波的功能, 预计当前帧中目标存在的区域, 将预计区域作为 TLD 算法的目标检测区域, 从而减少了 TLD 算法的检测范围, 相对于原 TLD 算法中检测模块要扫描全局图像的图像子窗口而言, 大大提高了检测的效率。

[0068] TLD 算法跟踪模块是基于 Lucas-Kanede (LK) 方法, 即光流法。光流是指图像每个像素运动的速度, 它是景物中可见点的三维速度矢量在成像平面上的投影, 表示景物表面点在图像中位置的瞬间变化。

[0069] LK 光流法有特定的三个条件, 分别为: (1) 亮度恒定性, 即同一像素点在不同帧之间的亮度保持不变; (2) 时间连续性, 即图像内容随时间变化的速度非常缓慢; (3) 空间一致性, 即场景中临近的点, 其在图像上的投影也是邻近点, 且邻近点的速度一致。LK 光流法的流程如图 2 所示, 其工作原理如下所述。首先, 在图像  $I_t$  的目标框  $\beta_t$  中随机选取  $10 \times 10$  点, 用  $x_i^t$  ( $i = 1 \dots 100$ ) 表示; 然后, 通过 LK 光流算法对这 100 个点进行跟踪, 得到下一帧图像  $I_{t+1}$  中这 100 个对应点  $x_i^{t+1}$  的位置, 并根据公式 (7) 得到这 100 个点的前向 - 后向误差, 从中选择一部分点作为可信跟踪点; 最后, 得到目标框  $\beta_{t+1}$  在图像  $I_{t+1}$  中的大小。

[0070]

$$\begin{cases} FB(T_f^k | S) = dis(T_f^k, T_b^k) \\ dis(T_f^k, T_b^k) = \|x_t - \tilde{x}_t\| \end{cases} \quad (7)$$

[0071] 其中  $T_f^k = (x_t, x_{t+1}, \dots, x_{t+k})$  为前向轨迹,  $T_b^k = (x'_{t-k}, x'_{t-k+1}, \dots, x'_t)$  为后向轨迹, 且  $x'_{t-k} = x_{t+k}$ 。

[0072] 如果  $t+1$  时刻的前向 - 后向误差  $FB(T_f^k | S_{t+1})$  小于  $t$  时刻的前向 - 后向误差  $FB(T_f^k | S_t)$ , 则该点为跟踪点可信; 否则, 前向跟踪轨迹是不正确的, 则删除该点。

[0073] 公式 (8) 给出前后两帧的距离变化率:

$$[0074] \quad r = \frac{\|x_i^{t+1} - x_j^{t+1}\|}{\|x_i^t - x_j^t\|}, i, j = 1, 2, \dots, n, i \neq j \quad (8)$$

[0075] 其中,  $n$  为可信跟踪点个数,  $r$  为点  $x_i^t$  和点  $x_i^{t+1}$  在前后两帧中的距离变化率。目标框  $\beta_{t+1}$  大小  $s_{t+1}$  可由目标框  $\beta_t$  大小  $s_t$  获得:

$$[0076] \quad s_{t+1} = s_t * r_m \quad (9)$$

[0077] 其中  $r_m$  为平均的距离变化率。

[0078] 跟踪学习的关键因素, 是跟踪模块是否能够正确跟踪目标。以此同时, 在融合模块中会对检测模块和跟踪模块得到的目标位置进行可信度计算, 从而决定目标的最终位置。因此, 需提高跟踪模块的准确度。

[0079] 本发明通过改进的卡尔曼滤波来增强 TLD 算法, 即根据目标位置的时间步长, 获得同步更新的参数传递矩阵。本发明算法的流程如图 3 所示, 主要包括如下: (1) 初始化系统: 将初始边框的中心位置赋予  $ptcurrent$  和  $ptpredict$ , 同时修改卡尔曼滤波器系统状态的初始数值和初始预测, 并利用这两个数值初始化卡尔曼滤波器。(2) 通过卡尔曼滤波器对后续帧进行预测, 且将该时刻预测的系统状态赋给  $ptpredict$ ; 同时, 将 LK 光流法跟踪得到的目标边框的中心位置点赋给  $ptcurrent$ 。(3) 更新  $ptcurrent$  之前, 用  $ptlast$  保存原先的系统状态测量。(4) 用更新系统的状态预测量和状态测量对卡尔曼滤波器进行纠正, 并预测目标边框的中心位置点, 采用原来有的目标边框的宽高比来还原出边框。(5) 利用卡尔

曼滤波器预测每一帧目标边框的可信度,即计算目标边框与最近邻分类器中目标模板的相似度;若该值大于 0.85,则用该目标边框取代 LK 光流法的检测边框结果,并将该边框检测结果传递给后续的学习模块和检测模块;相反,则保留 LK 光流法所得的目标边框。若可信度小于 0.6 或满足一定帧数后,则用 Lk 光流法的边框检测结果重新初始化卡尔曼滤波器。

[0080] 步骤 3:利用改进的卡尔曼滤波器所得到的预测结果,对检测模块进行改进,缩小 TLD 算法的检测区域,进一步提高跟踪的实时性。

[0081] TLD 算法的检测模块通过扫描窗口检测图像中是否存在目标。其扫描窗口参数设置如下:缩放比例的步长系数为 1.2,水平步长是宽度的 10%,垂直步长是高度的 10%,最小的扫描窗口大小为 20 个像素。

[0082] 为提高检测精度,检测模块中引入了如图 4 所示的一个三级级联分类器,分别是:(1) 方程分类器;(2) 随机森林分类器;(3) 最近邻分类器。每个阶段都将不满足条件的图像块去除,同时将满足条件的图像块送入下一个分类器,这不仅提高了分类器的效率,同时还提高了分类器的稳定性。

[0083] 虽然检测模块有非常有效的分类器,但检测模块需要对目标有可能出现的子窗口进行扫描,以确认目标是否存在图像内;且 TLD 算法只规定了子窗口的大小范围,并未规定子窗口的位置范围。所以,TLD 算法需要检测大量的窗口,其实这些子窗口中绝大多数都不包含目标内容,因此浪费了大量的计算资源。为此,本发明通过改进卡尔曼滤波器的预估结果来设置子窗口的位置范围,具体过程如下所述:(1) 通过改进的卡尔曼滤波器估计当前帧中运动目标的中心位置;(2) 若某一矩形区域的长宽比与上一帧中目标边界框的长宽比一致,大小是上一帧边界框的 4 倍,即认为该矩形区域包含检测目标;(3) 求与所有与该划定矩形区域有交集的子窗口,然后将这一系列子窗口送入分类器,由分类器检测子窗口中是否包含目标。

[0084] 本发明在原始 TLD 算法的基础上对其进行改进,从而获得更为理想的跟踪效果。本发明的改进包括以下两个方面:(1) 对跟踪模块进行改进。基于改进的卡尔曼滤波器原理对跟踪器进行一定的改进。同时依据目标位置的每个时间步长,计算并更新参数传递矩阵  $F$ ;(2) 对检测模块进行改进。根据卡尔曼滤波的功能,预计当前帧中目标存在的区域,将预计区域作为 TLD 算法的目标检测区域,从而减少了 TLD 算法的检测范围,相对于原 TLD 算法中检测模块要扫描全局图像的图像子窗口而言,大大提高了检测的效率。

[0085] 跟踪学习的关键因素,是跟踪模块是否能够正确跟踪目标。以此同时,在融合模块中会对检测模块和跟踪模块得到的目标位置进行可信度计算,从而决定目标的最终位置。因此,需提高跟踪模块的准确度。

[0086] 本发明通过改进的卡尔曼滤波来增强 TLD 算法,即根据目标位置的时间步长,获得同步更新的参数传递矩阵。本发明的算法流程主要包括:(1) 初始化系统:将初始边框的中心位置赋予  $pt_{current}$  和  $pt_{predict}$ ,同时修改卡尔曼滤波器系统状态的初始数值和初始预测,并利用这两个数值初始化卡尔曼滤波器。(2) 通过卡尔曼滤波器对后续帧进行预测,且将该时刻预测的系统状态赋给  $pt_{current}$ 。(3) 更新  $pt_{current}$  之前,用  $pt_{last}$  保存原先的  $pt_{predict}$ ;同时,将 LK 光流法跟踪得到的目标边界框的中心位置点赋给系统状态测量。(4) 用更新系统的状态预测量和状态测量对卡尔曼滤波器进行纠正,并预测目标边框的中心位置点,采用原来有的目标边框的宽高比来还原出边框。(5) 利用卡尔曼滤波器预测

每一帧目标边框的可信度,即计算目标边框与最近邻分类器中目标模板的相似度;若该值大于 0.85,则用该目标边框取代 LK 光流法的检测边框结果,并将该边框检测结果传递给后续的学习模块和检测模块;相反,则保留 LK 光流法所得的目标边框。若可信度小于 0.6 或满足一定帧数后,则用 Lk 光流法的边框检测结果重新初始化卡尔曼滤波器。

[0087] 根据卡尔曼滤波的功能,预计当前帧中目标存在的区域,将预计区域作为 TLD 算法的目标检测区域,从而减少了 TLD 算法的检测范围,相对于原 TLD 算法中检测模块要扫描全局图像的图像子窗口而言,大大提高了检测的效率。

[0088] 虽然检测模块有非常有效的分类器,但检测模块需要对目标有可能出现的子窗口进行扫描,以确认目标是否存在图像内;且 TLD 算法只规定了子窗口的大小范围,并未规定子窗口的位置范围。所以,TLD 算法需要检测大量的窗口,其实这些子窗口中绝大多数都不包含目标内容,因此浪费了大量的计算资源。为此,本发明通过改进卡尔曼滤波器的预估结果来设置子窗口的位置范围,具体过程如下所述:(1) 通过改进的卡尔曼滤波器估计当前帧中运动目标的中心位置;(2) 若某一矩形区域的长宽比与上一帧中目标边界框的长宽比一致,大小是上一帧边界框的 4 倍,即认为该矩形区域包含检测目标;(3) 求与所有与该划定矩形区域有交集的子窗口,然后将这一系列子窗口送入分类器,由分类器检测子窗口中是否包含目标。

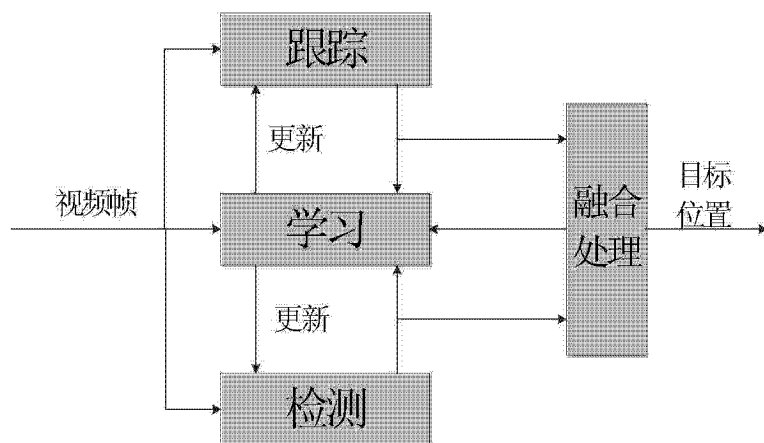


图 1

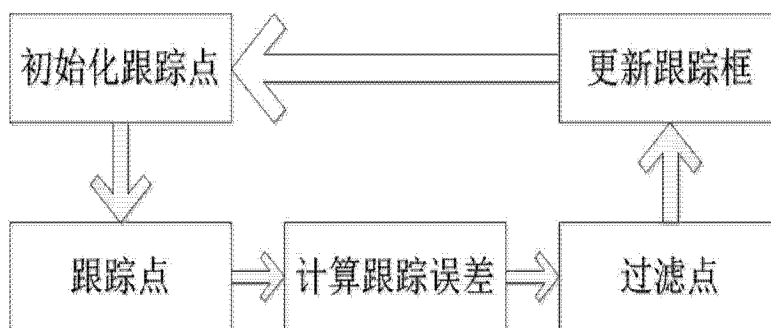


图 2

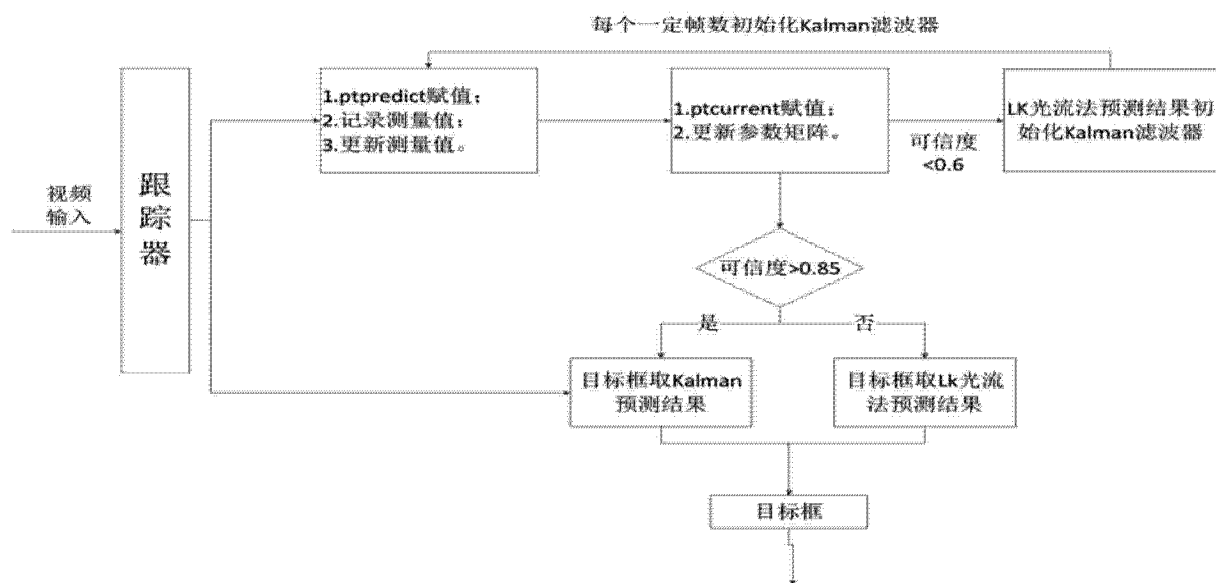


图 3

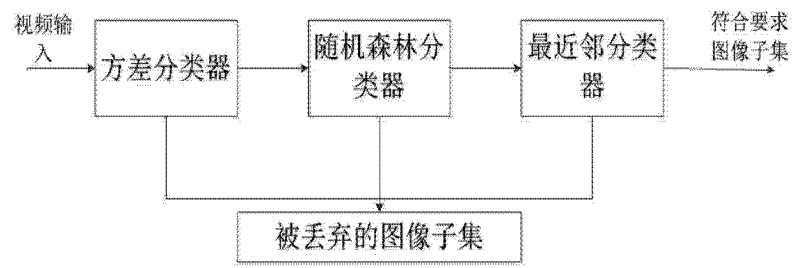


图 4