



(12) 发明专利申请

(10) 申请公布号 CN 104063885 A

(43) 申请公布日 2014. 09. 24

(21) 申请号 201410353072. 4

(22) 申请日 2014. 07. 23

(71) 申请人 山东建筑大学

地址 250101 山东省济南市历城区临港开发  
区凤鸣路

(72) 发明人 曹建荣 徐阳 李振宇 孙雪梅

(74) 专利代理机构 济南圣达知识产权代理有限  
公司 37221

代理人 张勇

(51) Int. Cl.

G06T 7/20 (2006. 01)

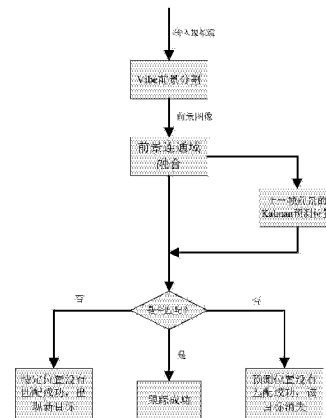
权利要求书2页 说明书5页 附图3页

(54) 发明名称

一种改进的运动目标检测与跟踪方法

(57) 摘要

本发明公开了一种改进的运动目标检测与跟踪方法,包括以下步骤:利用 Vibe 算法对检测的像素点进行分割;根据 Vibe 算法分割得到的连通域的空间分布关系,判断其是否为同一运动目标,如果判断结果为是,则进行连通域合并,直到目标检测完成;对前期获取到的运动对象进行跟踪,利用 Kalman 滤波器结合最小空间距离算法来预测每个运动对象的坐标位置,并进行不断更新、跟踪。本发明采用空间距离来作为判断依据,既能节省存储空间,又同时减少了计算量和计算复杂度。



1. 一种改进的运动目标检测与跟踪方法,其特征是:包括以下步骤:

(1) 利用 Vibe 算法对检测的像素点进行分割:为每个像素点建立一个背景模型,比较待分类像素与背景模型的相似度,若相似度满足设定值,则判断该像素为背景,否则为前景;

(2) 根据 Vibe 算法分割得到的连通域的空间分布关系,判断其是否为同一运动目标,如果判断结果为是,则进行连通域合并,直到目标检测完成;

(3) 对前期获取到的运动对象进行跟踪,利用 Kalman 滤波器结合最小空间距离算法来预测每个运动对象的坐标位置,并进行不断更新、跟踪,清除消失的运动对象并不断记录运动对象的新坐标位置。

2. 如权利要求 1 所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤 (1) 的具体方法为:为每个像素点  $p_t(x)$  建立一个包含  $N$  个样本的背景模型:

$$M(x) = \{p_1, p_2, \dots, p_N\}$$

通过比较待分类像素  $p_t(x)$  与背景模型  $M(x)$  的相似度,即可判断出该像素是前景还是背景,如果相似,即判断为背景,否则为前景。

3. 如权利要求 2 所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤 (1) 中,相似度的判断方法为:定义一个以  $p_t(x)$  为球心, $R$  为半径的球体  $S_R(p_t(x))$ ,用  $A$  来表示球体与背景模型的交集,用来判断相似度:

$$A = \{S_R(p_t(x)) \cap \{p_1, p_2, \dots, p_N\}\}$$

当  $A$  大于一个给定的阈值时,待分类像素即为背景,否则为前景。

4. 如权利要求 1 所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤 (2) 的具体为:通过 Vibe 算法对前景分割存在以下情况:经过 Vibe 算法进行前景分割后,如果同一个目标被分割成了很多个目标块,那这些目标块的最小外接矩形在相对位置上还是存在一定的位置关系的:一是较小的框完全被较大的框所包含;二是两个框之间有交集,但并不完全包含;三是两个矩形框之间并没有交集,但是他们成上下关系,且质心之间的水平与垂直距离都小于一定的阈值。

5. 如权利要求 4 所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤 (2) 的具体方法为:假设有两个矩形窗分别为  $A$  和  $B$ ,宽度分别记为  $W(A)$  和  $W(B)$ ,高度记为  $H(A)$  和  $H(B)$ ,中心点坐标记为  $(x(A), y(A))$ 、 $(x(B), y(B))$ ,则两个框中心点的水平距离为:

$$X_d = |x(A) - x(B)|$$

垂直距离为:

$$Y_d = |y(A) - y(B)|$$

所述步骤 (2) 中,如果是较小的框完全被较大的框所包含,则符合判决条件 1,即:

$$X_d \leq |W(A) - W(B)| / 2$$

和

$$Y_d \leq |H(A) - H(B)| / 2$$

直接将包含在内部的矩形框去掉即可。

6. 如权利要求 5 所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤 (2) 中,如果两个框之间有交集,但并不完全包含,则符合判决条件 2,即:

$$X_d \leq \alpha_1 \times |W(A) + W(B)| / 2$$

和

$$Y_d \leq \alpha_2 \times |H(A)+H(B)|/2。$$

7. 如权利要求5所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤(2)中,如果两个矩形框之间并没有交集,但是他们成上下关系,且质心间的水平与垂直距离都小于一定的阈值,则符合以下条件,即:

$$X_d \leq \beta_1 \times |W(A)+W(B)|/2$$

和

$$Y_d \leq \beta_2 \times |H(A)+H(B)|/2$$

则把这几个矩形框的最大外接矩形画出来,把内部的矩形框全部删除即可,其中, $\alpha_1$ 、 $\alpha_2$ 、 $\beta_1$ 和 $\beta_2$ 是系数。

8. 如权利要求1所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤(3)的具体方法包括:

第一步,针对已经检测到的所有的运动目标,分别对每一个目标*i*,首先利用Kalman滤波器来预测该目标在下一帧中的坐标位置,此位置记为预测位置;

第二步,针对下一帧中检测到的所有目标,对每一个目标*j*的坐标位置记为待定位置;

第三步,对某个*i*的预测位置,如果与其距离最近的待定位置是目标*j*的位置,同时针对该目标*j*的待定位置,与其距离最近的预测位置是*i*的位置,则说明目标*i*与目标*j*匹配成功,即跟踪成功,然后再用*j*的信息来更新旧目标*i*的信息,并同时根据目标*j*的位置与运动信息来更新Kalman滤波器,以调整其参数,进行更加鲁棒性的预测;

第四步,对于没有匹配成功的某个目标*i*的预测位置,说明该目标已经消失,因而将该目标从跟踪列表中清除即可;对于没有匹配成功的目标*j*的待定位置,说明目标*j*为新出现的目标,则为其开辟新的存储空间,并初始化其运动轨迹。

9. 如权利要求1所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤(3)中,将已经成功跟踪的所有目标的最小外接矩形彩色前景保存到前景容器中,将每个目标在每一帧的位置及大小信息保存到目标列表中,将经过Vibe背景建模得到的所有背景保存在背景容器中。

10. 如权利要求1所述的一种改进的运动目标检测与跟踪方法,其特征是:所述步骤(3)中,已知在视频的第*n*帧中的某个目标,去搜寻第*n+1*帧及后续帧中该目标的位置等信息,如果找到了,就说明跟踪成功;否则,可能就跟踪失败或者是目标消失。

## 一种改进的运动目标检测与跟踪方法

### 技术领域

[0001] 本发明涉及一种改进的运动目标检测与跟踪方法。

### 背景技术

[0002] 作为智能视频监控与分析系统的基础,运动目标检测与跟踪模块在其中起着非常重要的作用,它为后期的物体识别、行为与轨迹分析等后处理算法奠定了基础。人们已经对前景检测与跟踪算法进行了大量的研究。

[0003] 在前景检测算法方面,主要分为光流法、帧差法以及背景建模法。其中光流法对硬件要求较高,因而一般硬件很难满足算法的实时性;帧差法由于是相邻的两帧图像相减,而相邻的两帧图像中的运动目标在空间位置上相距很近,因而会在目标区域中产生很大的空洞,使得检测精度降低;背景建模法则是背景图像建立模型,使得在前景检测的过程中背景可以不断的更新,能够适应存在细微变化的背景环境,但是当前景的颜色与背景颜色相近,即前景与背景对比度较低时,背景建模算法也很难完整的检测出前景目标。

[0004] 在跟踪算法方面,比较经典的 Camshift 跟踪算法仅仅采用单一的颜色特征,计算量小,运算速度快,但是鲁棒性与适应性较差,在背景复杂时很容易丢失目标;粒子滤波算法的跟踪性能与粒子数成正比,粒子数越多跟踪能力越强,但粒子数的增加极大的增大了算法的计算量,使得实时性难以满足,很难用于对实时性要求较高的监控系统。

### 发明内容

[0005] 本发明为了解决上述问题,提出了一种改进的运动目标检测与跟踪方法,本方法在传统的背景建模方法基础上的改进,即便是在前景与背景颜色相近的情况下也能进行准确的前景检测;在跟踪方法中采用了 Kalman 滤波结合最小空间距离的方法,在保证跟踪鲁棒性的同时也加快了跟踪算法的运算速度。

[0006] 为了实现上述目的,本发明采用如下技术方案:

[0007] 一种改进的运动目标检测与跟踪方法,包括以下步骤:

[0008] (1) 利用 Vibe 算法对检测的像素点进行分割:为每个像素点建立一个背景模型,比较待分类像素与背景模型的相似度,若相似度满足设定值,则判断该像素为背景,否则为前景;

[0009] (2) 根据 Vibe 算法分割得到的连通域的空间分布关系,判断其是否为同一运动目标,如果判断结果为是,则进行连通域合并,直到目标检测完成;

[0010] (3) 对前期获取到的运动对象进行跟踪,利用 Kalman 滤波器结合最小空间距离算法来预测每个运动对象的坐标位置,并进行不断更新、跟踪,清除消失的运动对象并不断记录运动对象的新坐标位置。

[0011] 所述步骤 (1) 的具体方法为:为每个像素点  $p_t(x)$  建立一个包含  $N$  个样本的背景模型:

[0012]  $M(x) = \{p_1, p_2, \dots, p_N\}$

[0013] 通过比较待分类像素  $p_t(x)$  与背景模型  $M(x)$  的相似度,即可判断出该像素是前景还是背景,如果相似,即判断为背景,否则为前景。

[0014] 所述步骤 (1) 中,相似度的判断方法为:定义一个以  $p_t(x)$  为球心,  $R$  为半径的球体  $S_R(p_t(x))$ ,用  $A$  来表示球体与背景模型的交集,用来判断相似度:

$$[0015] \quad A = \{S_R(p_t(x)) \cap \{p_1, p_2, \dots, p_N\}\}$$

[0016] 当  $A$  大于一个给定的阈值时,待分类像素即为背景,否则为前景。

[0017] 所述步骤 (2) 的具体为:通过 Vibe 算法对前景分割存在以下情况:经过 Vibe 算法进行前景分割后,如果同一个目标被分割成了很多个目标块,那这些目标块的最小外接矩形在相对位置上还是存在一定的位置关系的:一是较小的框完全被较大的框所包含;二是两个框之间有交集,但并不完全包含;三是两个矩形框之间并没有交集,但是他们成上下关系,且质心之间的水平与垂直距离都小于一定的阈值。

[0018] 所述步骤 (2) 的具体方法为:假设有两个矩形窗分别为  $A$  和  $B$ ,宽度分别记为  $W(A)$  和  $W(B)$ ,高度记为  $H(A)$  和  $H(B)$ ,中心点坐标记为  $(x(A), y(A))$ 、 $(x(B), y(B))$ ,则两个框中心点的水平距离为:

$$[0019] \quad X_d = |x(A) - x(B)|$$

[0020] 垂直距离为:

$$[0021] \quad Y_d = |y(A) - y(B)|$$

[0022] 所述步骤 (2) 中,如果是较小的框完全被较大的框所包含,则符合判决条件 1,即:

$$[0023] \quad X_d \leq |W(A) - W(B)| / 2$$

[0024] 和

$$[0025] \quad Y_d \leq |H(A) - H(B)| / 2$$

[0026] 直接将被包含在内部的矩形框去掉即可。

[0027] 所述步骤 (2) 中,如果两个框之间有交集,但并不完全包含,则符合判决条件 2,即:

$$[0028] \quad X_d \leq \alpha_1 \times |W(A) + W(B)| / 2$$

[0029] 和

$$[0030] \quad Y_d \leq \alpha_2 \times |H(A) + H(B)| / 2。$$

[0031] 所述步骤 (2) 中,如果两个矩形框之间并没有交集,但是他们成上下关系,且质心之间的水平与垂直距离都小于一定的阈值,则符合以下条件,即:

$$[0032] \quad X_d \leq \beta_1 \times |W(A) + W(B)| / 2$$

[0033] 和

$$[0034] \quad Y_d \leq \beta_2 \times |H(A) + H(B)| / 2$$

[0035] 则把这几个矩形框的最大外接矩形画出来,把内部的矩形框全部删除即可,其中,  $\alpha_1$ 、 $\alpha_2$ 、 $\beta_1$  和  $\beta_2$  是系数。

[0036] 所述步骤 (3) 的具体方法包括:

[0037] 第一步,针对已经检测到的所有的运动目标,分别对每一个目标  $i$ ,首先利用 Kalman 滤波器来预测该目标在下一帧中的坐标位置,此位置记为预测位置;

[0038] 第二步,针对下一帧中检测到的所有目标,对每一个目标  $j$  的坐标位置记为待定位置;

[0039] 第三步,对某个  $i$  的预测位置,如果与其距离最近的待定位置是目标  $j$  的位置,同时针对该目标  $j$  的待定位置,与其距离最近的预测位置是  $i$  的位置,则说明目标  $i$  与目标  $j$  匹配成功,即跟踪成功,然后再用  $j$  的信息来更新旧目标  $i$  的信息,并同时根据目标  $j$  的位置与运动信息来更新 Kalman 滤波器,以调整其参数,进行更加鲁棒性的预测。

[0040] 第四步,对于没有匹配成功的某个目标  $i$  的预测位置,说明该目标已经消失,因而将该目标从跟踪列表中清除即可;对于没有匹配成功的目标  $j$  的待定位置,说明目标  $j$  为新出现的目标,则为其开辟新的存储空间,并初始化其运动轨迹。

[0041] 所述步骤(3)中,已知在视频的第  $n$  帧中的某个目标,去搜寻第  $n+1$  帧及后续帧中该目标的位置等信息,如果找到了,就说明跟踪成功;否则,可能就跟踪失败或者是目标消失。

[0042] 本发明的有益效果为:

[0043] (1) 通过对传统背景建模方法与多连通域融合方法的有效结合,使得在前景与背景对比度较低的环境下也能够进行准确的前景目标检测,为后续运动跟踪与轨迹分析奠定了基础;

[0044] (2) 采用 Kalman 滤波结合最小空间距离法进行运动目标跟踪,其中利用 Kalman 滤波来预测运动目标在下一帧中的大致位置,再结合最小空间距离来精确地进行目标定位,该跟踪方法的适应性更强,同时减少了计算量和计算复杂度。

#### 附图说明

[0045] 图 1 为 Vibe 算法模型示意图;

[0046] 图 2a 为本发明两个连通域的一种位置关系示意图;

[0047] 图 2b 为本发明两个连通域的一种位置关系示意图;

[0048] 图 2c 为本发明两个连通域的一种位置关系示意图;

[0049] 图 3a 为 Vibe 算法检测结果与改进后的算法检测结果对比图;

[0050] 图 3b 为 Vibe 算法检测结果与改进后的算法检测结果对比图;

[0051] 图 4a 为跟踪方法实验结果对比图;

[0052] 图 4b 为跟踪方法实验结果对比图;

[0053] 图 5 为改进的算法流程图。

#### 具体实施方式:

[0054] 下面结合附图与实施例对本发明作进一步说明。

[0055] 在运动目标检测方法中,本发明采用改进后的 Vibe 前景检测算法,即基本的 Vibe 背景建模算法和多连通域融合算法相结合的新方法。

[0056] 基本的 Vibe 背景建模算法的思想就是为每个像素点存储一个样本集,样本集中采样值就是该像素点过去的像素值和其邻域的像素值,然后将每一个新的像素值和样本集进行比较来判断是否属于背景点,同时其采用了随机选择机制和邻域传播机制来建立和更新背景模型。它对于各种视频流、不同颜色空间、多种场景内容及变化的场景都适用,从而有效的提高了准确度,增加了抗噪能力和减少了计算负载。

[0057] 如图 1 所示,为每个像素点  $p_t(x)$  建立一个包含  $N$  个样本的背景模型:

[0058]  $M(x) = \{p_1, p_2, \dots, p_N\}$  (1)

[0059] 通过比较待分类像素  $p_t(x)$  与背景模型  $M(x)$  的相似度,即可判断出该像素是前景还是背景。如果相似,即判断为背景,否则为前景。其中相似度按如下方式来定义:如上图 1 所示,定义一个以  $p_t(x)$  为球心,  $R$  为半径的球体  $S_R(p_t(x))$ , 用  $A$  来表示球体与背景模型的交集,即可用来判断相似度:

[0060]  $A = \{S_R(p_t(x)) \cap \{p_1, p_2, \dots, p_N\}\}$  (2)

[0061] 当  $A$  大于一个给定的阈值时,待分类像素即为背景,否则为前景。

[0062] 但由于像建筑物楼道等环境具有光线差,摄像机质量不高,拍出的视频对比度较低等缺点,往往会导致用 Vibe 算法的检测精度变差,因而本发明在此基础上提出了基于空间位置分布特征的多连通域融合算法,通过分析先前由 Vibe 算法分割得到的数个连通域之间的空间分布关系,将属于同一运动目标的不同连通域进行融合。用改进后的算法进行运动检测,就会得到更加精确的检测结果。

[0063] 在经过 Vibe 算法进行前景分割后,如果同一个目标被分割成了很多个目标块,那这些目标块的最小外接矩形在相对位置上还是存在一定的位置关系的:一是较小的框完全被较大的框所包含,如图 2a 所示;二是两个框之间有交集,但并不完全包含,如图 2b 所示;三是两个矩形框之间并没有交集,但是他们主要成上下关系,且距离较近,如图 2c 所示。

[0064] 针对如上所述的三种情况,我们有如下的解决方案:

[0065] 假设有两个矩形窗分别为  $A$  和  $B$ , 宽度分别记为  $W(A)$  和  $W(B)$ , 高度记为  $H(A)$  和  $H(B)$ , 中心点坐标记为  $(x(A), y(A))$ 、 $(x(B), y(B))$ , 则两个框中心点的水平距离为:

[0066]  $X_d = |x(A) - x(B)|$  (3)

[0067] 垂直距离为:

[0068]  $Y_d = |y(A) - y(B)|$  (4)

[0069] 针对上述的第一种关系,即如果符合判决条件 1, 即:

[0070]  $X_d \leq |W(A) - W(B)| / 2$  (5)

[0071] 和

[0072]  $Y_d \leq |H(A) - H(B)| / 2$  (6)

[0073] 则说明它们符合第一种关系,因而直接就把被包含在内部的矩形框去掉即可。

[0074] 而针对后两种关系的处理方式是相似的,即先判断如果符合判决条件 2, 即:

[0075]  $X_d \leq \alpha_1 \times |W(A) + W(B)| / 2$  (7)

[0076] 和

[0077]  $Y_d \leq \alpha_2 \times |H(A) + H(B)| / 2$  (8)

[0078] 或者是满足:

[0079]  $X_d \leq \beta_1 \times |W(A) + W(B)| / 2$  (9)

[0080] 和

[0081]  $Y_d \leq \beta_2 \times |H(A) + H(B)| / 2$  (10)

[0082] 则把这几个矩形框的最大外接矩形画出来,把内部的矩形框全部删除即可。其中,  $\alpha_1$ 、 $\alpha_2$ 、 $\beta_1$  和  $\beta_2$  是系数,经过大量实验后可得出当这四个系数分别取 1.1、1.1、1.5 和 0.25 时会得到精确的检测结果。

[0083] 部分实验结果如图 3(a)、(b) 所示:

[0084] 其中左侧一纵列是由 Vibe 算法进行前景检测的结果,而右侧一纵列则是通过改进后的方法的检测结果。

[0085] 在运动目标跟踪方法中,实现的主要功能是对前期检测到的前景运动对象进行跟踪,也就是确定同一目标体在图像序列的不同帧中的位置的过程。即已知在视频的第  $n$  帧中的某个目标,去搜寻第  $n+1$  帧及后续帧中该目标的位置等信息,如果找到了,就说明跟踪成功;否则,可能就跟踪失败或者是目标消失。

[0086] 如图 5 所示,本发明在跟踪模块提出了基于空间距离的目标跟踪方法,即是指相邻两帧视频图像中的目标在二维空间中相距有多远。为简单起见,把运动目标的最小外接矩形的中心代表该目标在图像上的位置。采用空间距离来作为判断依据,既能节省了存储空间,又同时减少了计算量和计算复杂度。

[0087] 本发明中运动跟踪方法的具体步骤如下:

[0088] 第一步,针对已经检测到的所有的运动目标,分别对每一个目标  $i$ ,首先利用 Kalman 滤波器来预测该目标在下一帧中的坐标位置,此位置记为预测位置;

[0089] 第二步,针对下一帧中检测到的所有目标,对每一个目标  $j$  的坐标位置记为待定位置;

[0090] 第三步,对某个  $i$  的预测位置,如果与其距离最近的待定位置是目标  $j$  的位置,同时针对该目标  $j$  的待定位置,与其距离最近的预测位置是  $i$  的位置,则说明目标  $i$  与目标  $j$  匹配成功,即跟踪成功,然后再用  $j$  的信息来更新旧目标  $i$  的信息,并同时根据目标  $j$  的位置与运动信息来更新 Kalman 滤波器,以调整其参数,进行更加鲁棒性的预测。

[0091] 第四步,对于没有匹配成功的某个目标  $i$  的预测位置,说明该目标已经消失,因而将该目标从跟踪列表中清除即可;对于没有匹配成功的目标  $j$  的待定位置,说明目标  $j$  为新出现的目标,则为其开辟新的存储空间,并初始化其运动轨迹。跟踪效果图如图 4a、图 4b 所示,本方法能够准确对运动目标进行跟踪。

[0092] 上述虽然结合附图对本发明的具体实施方式进行了描述,但并非对本发明保护范围的限制,所属领域技术人员应该明白,在本发明的技术方案的基础上,本领域技术人员不需要付出创造性劳动即可做出的各种修改或变形仍在本发明的保护范围以内。



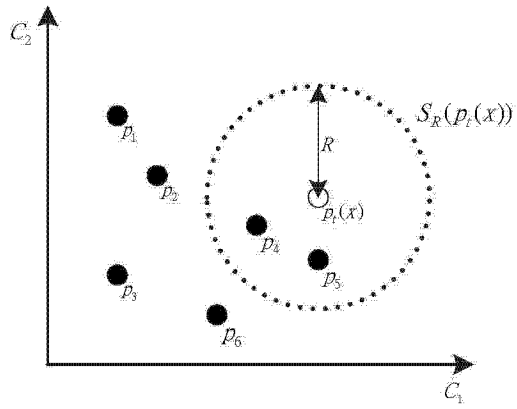


图 1

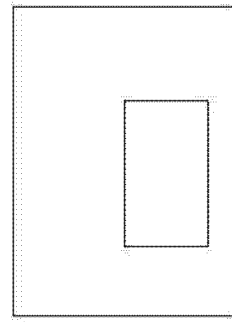


图 2a

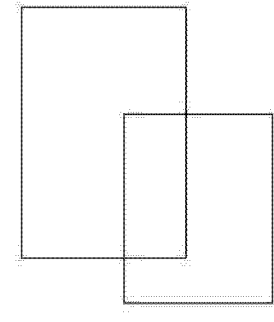


图 2b

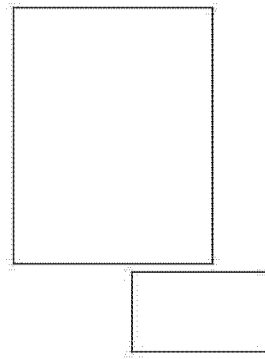


图 2c

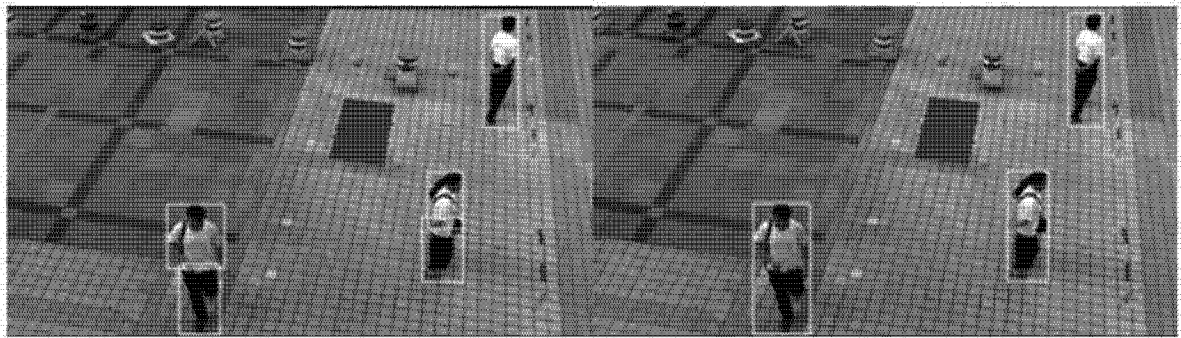


图 3a



图 3b



图 4a



图 4b

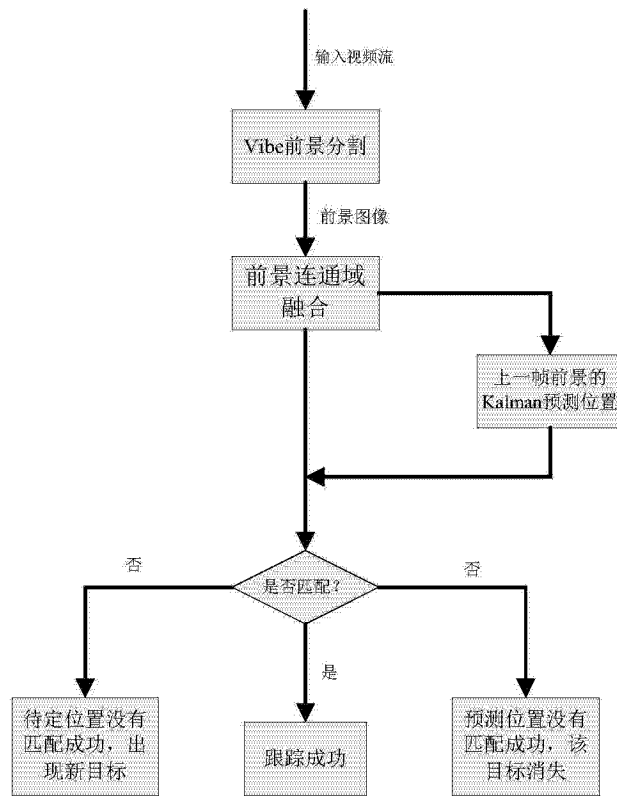


图 5