



(12)发明专利

(10)授权公告号 CN 104978567 B

(45)授权公告日 2018. 11. 20

(21)申请号 201510318709.0

(22)申请日 2015.06.11

(65)同一申请的已公布的文献号
申请公布号 CN 104978567 A

(43)申请公布日 2015.10.14

(73)专利权人 武汉大千信息技术有限公司
地址 430079 湖北省武汉市东湖开发区水
蓝路水蓝郡329号

(72)发明人 严国建

(74)专利代理机构 武汉智盛唯佳知识产权代理
事务所(普通合伙) 42236
代理人 胡红林

(51)Int.Cl.
G06K 9/00(2006.01)
G06K 9/62(2006.01)

(56)对比文件

CN 103632170 A,2014.03.12,
CN 103366572 A,2013.10.23,
CN 103400138 A,2013.11.20,
辛海涛.基于运动目标检测的行人计数方
法.《中国优秀硕士学位论文全文数据库信息科
技辑(月刊)》.2012,(第05期),第1138-1232页.

审查员 张倩

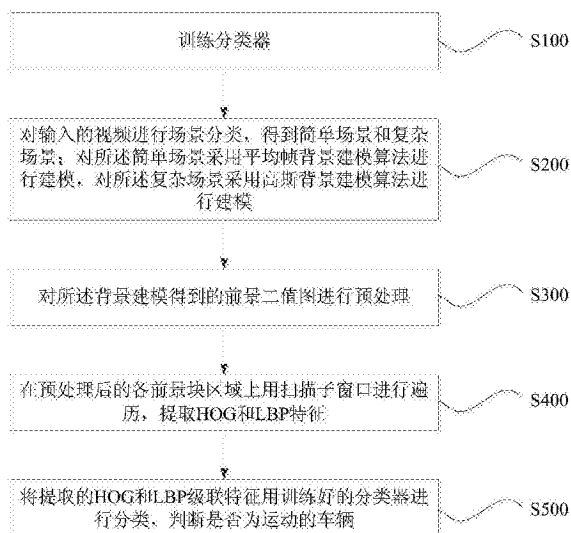
权利要求书2页 说明书5页 附图1页

(54)发明名称

基于场景分类的车辆检测方法

(57)摘要

本发明涉及一种基于场景分类的车辆检测方法,该方法包括:训练分类器;对输入的视频进行场景分类,得到简单场景和复杂场景;对简单场景采用平均帧背景建模算法进行建模,对复杂场景采用高斯背景建模算法进行建模;对背景建模得到的前景二值图进行预处理;在预处理后的各前景块区域上用扫描子窗口进行遍历,提取HOG和LBP特征;将提取的HOG和LBP级联特征用训练好的分类器进行分类,判断是否为运动的车辆。本发明能根据场景复杂度自动选择最匹配的算法,还能同时提取HOG和LBP特征,将两种特征级联作为新的分类器特征,通过这两方面的改进,即保证活动车辆的检测效果,又极大提高系统整体检测速度,从而能更快更准确地定位到要待识别的车辆。



1. 一种基于场景分类的车辆检测方法,其特征在于,包括:

训练分类器;

对输入的视频进行场景分类,得到简单场景和复杂场景;对所述简单场景采用平均帧背景建模算法进行建模,对所述复杂场景采用高斯背景建模算法进行建模;在输入的视频中,对视频中活动目标进行检测,然后统计检测出来的活动目标块数量和面积,当活动目标数量小于 m 个, $10 \leq m \leq 30$,并且活动区域面积少于整张图像的 $n\%$, $40 \leq n \leq 70$,则判定该视频场景为简单场景,采用平均帧背景建模算法;当活动目标数量大于 m 个,则可以判定该视频场景为复杂场景,相应的采用高斯背景建模算法;所述平均帧背景建模算法具体包括:

第一步:从视频中读取连续的 K 帧图像,并将每帧图像转化为灰度矩阵 D_x

$$D_x = \{Y_{i,j}, i \in \{1, \dots, M\}, j \in \{1, \dots, N\}\}$$

式中, M 表示图像帧的行数, N 表示图像帧的列数, $Y_{i,j}$ 是 (i,j) 位置像素转变后的灰度值, $Y_{i,j}$ 通过下式计算:

$$Y_{i,j} = 0.299 \times R_{i,j} + 0.587 \times G_{i,j} + 0.114 \times B_{i,j}$$

式中, $R_{i,j}, G_{i,j}, B_{i,j}$ 分别是图像在第 i 行 j 列上的R、G、B颜色值;

第二步:将前 K 帧灰度矩阵叠加,然后再对叠加结果求平均值得到背景模型 I_{bgm} ;

第三步:当输入一帧图像 $I_{present}$,将其与背景模型 I_{bgm} 求差值,得到差值图像 I_{abs} :

$$I_{abs} = |I_{present} - I_{bgm}|$$

第四步:将差值图像 I_{abs} 二值化,获得前景二值图,即活动目标信息 $I_{foreground}$;

对所述背景建模得到的前景二值图进行预处理;

在预处理后的各前景块区域上用扫描子窗口进行遍历,提取HOG和LBP特征;

将提取的HOG和LBP级联特征用训练好的分类器进行分类,判断是否为运动的车辆;

其中,所述训练分类器,包括:

采集车辆正、负样本;采集车辆的正样本图片包括:在实际监控视频中对于车辆在多段不同场景的道路监控视频中,人工截取多张长宽为 $b*b$, $50 \leq b \leq 200$,像素为 $352*288$ 的车辆图片,所述车辆图片包含完整的车辆;采集车辆的负样本图片包括:在实际监控视频中对于车辆在不同场景的道路监控视频中,采用软件对监控视频的每帧自动切割为长宽为 $b*b$ 的图片并保存,其中 $50 \leq b \leq 20$,在这些图片中选至少20000张不含车辆的图片作为负样本;

对每一张正、负样本的图片进行特征选取和提取;

根据提取的特征对SVM线性分类器进行训练,得到训练好的分类器。

2. 根据权利要求1所述基于场景分类的车辆检测方法,其特征在于:高斯背景建模算法具体包括

在视频序列中,对于任意时刻 t 在位置 $\{x_0, y_0\}$ 上,它的历史像素表示为: $\{X_1, \dots, X_t\} = \{I(x_0, y_0, s) : 1 \leq s \leq t\}$,其中 I 表示图像序列;对背景建立 K -高斯模型,则在 X_t 属于背景的概率为:

$$P(X_t) = \sum_{s=1}^K \omega_{s,t} * \eta(X_t, \mu_{s,t}, \Sigma_{s,t})$$

式中, K 为模型数量, $\omega_{s,t}$ 为第 s 个高斯模型在 t 时刻属于背景的权重, $\mu_{s,t}$ 为第 s 个高斯模型在 t 时刻的均值, $\Sigma_{s,t}$ 为第 s 个高斯模型在 t 时刻的方差, η 为高斯密度函数;其中 η 为:

$$\eta(X, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma^{-1} (X_t - \mu_t)}$$

式中, $P(X_t)$ 值越大, 则说明当前像素越符合背景模型, 当 $P(X_t)$ 大于设定的阈值, 则将该像素判断为背景, 否则将其判断为前景。

3. 根据权利要求2所述基于场景分类的车辆检测方法, 其特征在于所述预处理包括:

运用膨胀腐蚀、形态滤波、中值滤波和前景块的面积阈值, 对背景建模得到的前景二值图进行预处理。

4. 根据权利要求3所述基于场景分类的车辆检测方法, 其特征在于所述在预处理后的各前景块区域上用扫描子窗口进行遍历包括:

首先将车辆检测子窗口Block的大小设为 $2a \times 2a$, 每个Block划分为4个Cell, 每个Cell的大小设为 $a \times a$; 用车辆检测子窗口Block对视频帧从左到右从上至下进行扫描, 每次在X轴方向移动的步长设为 a 个像素, 在Y轴方向移动的步长设为 a 个像素; 然后将每个 $2a \times 2a$ Block的大小的图像块, 归一化为 $b \times b$ 的大小图像块。

5. 根据权利要求4所述基于场景分类的车辆检测方法, 其特征在于判断是否为运动的车辆包括: 将每帧提取的HOG特征和LBP特征级联, 得到一个行向量, 用训练好的SVM分类器进行分类, 判断是否为运动的车辆。

基于场景分类的车辆检测方法

技术领域

[0001] 本发明涉及监控视频的检索,具体地指一种基于场景分类的车辆检测方法。

背景技术

[0002] 在刑侦业务中,车辆检测就是通过输入的视频判断是否包含待检测的运动车辆,如果包含则确定目标所在位置。在实际视频侦查中,侦查员常常根据车辆的活动画面和轨迹来快速锁定、排查和追踪嫌疑目标。传统的人工浏览视频侦查模式需要耗费大量的人力和时间,容易贻误破案时机。因此,精准快速的车辆检测技术便于视频侦查员快速、准确地发现嫌疑目标活动画面和轨迹,对公安部门提高破案率、维护人民群众生命财产安全具有重要意义。

[0003] 但是,现有的车辆检测的步骤如图1所示,一般包括以下步骤:当利用车辆正负样本选取和提取特征后,便可利用这些特征训练分类器。得到了分类器后,首先是对输入视频帧进行背景建模,然后提取特征,最后将提取的特征给训练好的分类器,判断是否存在想要检测的目标。

[0004] 其中,在上述现有的车辆检测步骤中,背景建模算法的效果会直接影响到车辆检测的效果,目前,一般的背景建模常常对不同的视频场景,只采用平均帧背景建模算法、VIBE背景建模算法、高斯背景建模算法中的一种算法进行背景建模。具体地,平均帧背景建模算法是通过求取连续视频序列固定位置上像素平均值,用该值表示当该位置像素点的背景模型的算法;高斯背景建模方法,首先对图像的每个像素或其他纹理特征建立一个或多个高斯背景模型,然后通过比较待检图像和背景模型的差异,来检测图像中的前景区域和背景区域;VIBE背景模型针对每个背景点保存了一个样本集,通过比较每一个新的像素点和样本集来判断该新像素点是否属于背景点。

[0005] 然而在实际监控环境下,因为场景光照变化、气候变化或者摄像机本身的抖动等都会导致视频背景发生改变,这些因素给背景建模带来困难。虽然目前对复杂场景下的背景建模已经有很多研究,但目前还没有一种算法能全部解决上述所有问题。例如:平均帧背景建模算法和VIBE背景建模算法速度快,但是只针对简单的视频场景(主要是光照变化小的场景)有效;高斯背景建模算法,对复杂变化的场景具有比较好的鲁棒性,但是速度很慢,背景更新速度缓慢不利于实时检测。

[0006] 此外,在上述现有的车辆检测步骤中,提取特征通常只是提取了图片的HOG特征用于检测,导致图像目标的视觉特征比较单一,辨识度不强,对噪声、光照变化、运动车辆的形变等鲁棒性不强。

发明内容

[0007] 本发明目的在于克服上述现有技术的不足,而提出一种基于场景分类的车辆检测方法,该方法能根据场景复杂度自动选择最匹配的算法;还能在提取梯度特征HOG的基础上,同时提取纹理特征LBP,将两种特征级联作为新的分类器特征,通过这两方面的改进,即

保证活动车辆的检测效果,又极大提高系统整体检测速度,从而能更快更准确地定位到要待识别的车辆。

[0008] 实现本发明目的采用的技术方案是:一种基于场景分类的车辆检测方法,该方法包括:

[0009] 训练分类器;

[0010] 对输入的视频进行场景分类,得到简单场景和复杂场景;对所述简单场景采用平均帧背景建模算法进行建模,对所述复杂场景采用高斯背景建模算法进行建模;

[0011] 对所述背景建模得到的前景二值图进行预处理;

[0012] 在预处理后的各前景块区域上用扫描子窗口进行遍历,提取HOG和LBP特征;

[0013] 将提取的HOG和LBP级联特征用训练好的分类器进行分类,判断是否为运动的车辆。

[0014] 本发明具有以下优点:

[0015] 1、与现有技术相比,本发明方法引入基于场景分类的背景建模,通过将视频按照内容划分为简单场景视频和复杂场景视频两大类,然后计算机根据划分的场景类别自动的选择最优的背景建模算法,这样既能保证检测算法效果,又能实现活动目标快速检测,具有很高的应用价值;

[0016] 2、与现有技术相比,本发明考虑了在检测视频的特征提取,在提取的HOG特征的基础上同时提取了LBP特征,将两种特征级联作为新的分类器特征。相比现有技术提取特征通常只是提取图片的HOG特征用于检测,本方法极大提高运动车辆检测的准确性,也增强了对噪声、光照变化、运动车辆的形变的鲁棒性。

附图说明

[0017] 图1为本发明基于场景分类的车辆检测方法流程图。

具体实施方式

[0018] 下面结合附图和具体实施例对本发明作进一步的详细说明。

[0019] 本发明基于场景分类的车辆检测方法具体包括以下步骤:

[0020] S100、训练分类器。

[0021] 首先,采集车辆的正、负样本。

[0022] 车辆的正样本图片的采集过程为:在实际监控视频中对于车辆在8段不同场景的道路监控视频中,人工截取10000张,长宽为 $b*b$, $50 \leq b \leq 200$ 像素为 $352*288$ 的车辆图片,这些正样本图片应包含完整的车辆而且包含尽可能少的背景,完整的车辆应该包含了车辆的正面、侧面、背面。

[0023] 车辆的负样本图片的采集过程为:在实际监控视频中对于车辆在8段不同场景的道路监控视频中,采用软件对监控视频的每帧自动切割为长宽为 $b*b$ 的图片并保存,其中, $50 \leq b \leq 200$,在这些图片中选至少20000张不含车辆的图片作为负样本。

[0024] 然后,训练正负样本,分别对每一张正、负样本的图片进行特征选取和提取。

[0025] 最后,训练分类器,本实施例采用SVM线性分类器。训练分类器即用分类器训练正、负样本,得到训练好的分类器。

[0026] S200、对输入的视频进行场景分类,得到简单场景和复杂场景;对所述简单场景采用平均帧背景建模算法进行建模,对所述复杂场景采用高斯背景建模算法进行建模。

[0027] 建模算法成立的前提假设是一般监控视频中,单帧图像所包含的活动目标数量不会太多(一般不会超过30个),活动目标面积也较少(不多于整幅图像面积的70%);

[0028] 首先选用平均帧背景建模算法,对视频活动目标进行检测,然后统计检测出来的活动目标块数量和面积。当活动目标数量小于 m 个(m 的取值范围 $10\sim 30$),并且活动区域面积少于整张图像的 $n\%$ (n 的取值范围 $40\sim 70$),则判定该视频场景为简单场景,采用平均帧背景建模算法。当活动目标数量大于 m 个,或者活动区域面积几近覆盖全画面,则可以判定该视频场景为复杂场景,相应的采用高斯背景建模算法。

[0029] 平均帧背景建模算法是通过求取连续视频序列固定位置上像素平均值,用该值表示当该位置像素点的背景模型的算法。本算法成立的依据是:通过大量统计监控视频图像,发现在每一帧视频图像中活动区域只占画面小部分,而大部分区域都是静止的背景。因此对于整个视频序列而言,相同位置上的像素集合中绝大部分都是静止的,只有少数是变动的活动区域。当求取相同位置像素集合的平均值时,少量活动目标像素对该平均值的影响很小,该平均值即能代表图像背景特征。

[0030] 在算法速度测试方面,平均帧算法明显快于高斯背景建模算法和VIBE背景建模算法;VIBE算法速度略高于基于高斯背景建模算法的检测速度。

[0031] 而在算法运行效果方面,清晰场景、模糊场景、夜晚场景下三种算法整体效果都不错,其中背景相对稳定的清晰场景和模糊场景下,平均帧背景建模算法和VIBE背景建模算法稍微好于高斯背景建模算法,而在夜晚和强光变化的场景下,由于平均帧背景建模算法的背景固定,所以效果急剧下降,VIBE算法更新策略上选用随机方式,更新速度相对较慢,所以检测效果也不如高斯背景建模算法。

[0032] 发明在相对简单的场景下采用平均帧背景建模算法,效果最好,速度最快;而在场景相对复杂情况下,采用高斯背景建模算法则是最佳选择。

[0033] 本实施例采用平均帧背景建模算法的具体步骤如下:

[0034] 第一步:从视频中读取连续的 K 帧图像,并将每帧图像转化为灰度矩阵 D_x

[0035] $D_x = \{Y_{i,j}, i \in \{1, \dots, M\}, j \in \{1, \dots, N\}\}$

[0036] 式中, M 表示图像帧的行数, N 表示图像帧的列数, $Y_{i,j}$ 是 (i,j) 位置像素转变后的灰度值, $Y_{i,j}$ 通过下式计算:

[0037] $Y_{i,j} = 0.299 \times R_{i,j} + 0.587 \times G_{i,j} + 0.114 \times B_{i,j}$

[0038] 式中, $R_{i,j}, G_{i,j}, B_{i,j}$ 分别是图像在第 i 行 j 列上的R、G、B颜色值;

[0039] 第二步:将前 K 帧灰度矩阵叠加,然后再对叠加结果求平均值得到背景模型 I_{bgm} ;

[0040] 第三步:当输入一帧图像 $I_{present}$,将其与背景模型 I_{bgm} 求差值,得到差值图像 I_{abs} :

[0041] $I_{abs} = |I_{present} - I_{bgm}|$

[0042] 第四步:将差值图像 I_{abs} 二值化,获得前景二值图,即活动目标信息 $I_{foreground}$ 。

[0043] 高斯背景建模算法具体包括:

[0044] 在视频序列中,对于任意时刻 t 在位置 $\{x_0, y_0\}$ 上,它的历史像素(如灰度值)表示为: $\{X_1, \dots, X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\}$,其中 I 表示图像序列;对背景建立 K -高斯模型,则在 X_t 属于背景的概率为:

$$[0045] \quad P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t})$$

[0046] 式中, K为模型数量, $\omega_{i,t}$ 为第i个高斯模型在t时刻属于背景的权重, $\mu_{i,t}$ 为第i个高斯模型在t时刻的均值, $\Sigma_{i,t}$ 为第i个高斯模型在t时刻的方差, η 为高斯密度函数; 其中 η 为:

$$[0047] \quad \eta(X, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)}$$

[0048] 式中, $P(X_t)$ 值越大, 则说明当前像素越符合背景模型, 当 $P(X_t)$ 大于设定的阈值, 则将该像素判断为背景, 否则将其判断为前景。

[0049] S300、对所述背景建模得到的前景二值图进行预处理。

[0050] 具体地, 本实施例预处理具体为运用膨胀腐蚀、形态滤波、中值滤波和前景块的面积阈值, 对背景建模得到的前景二值图进行预处理。本实施例中面积阈值大小, 车辆设为800~1500。

[0051] S400、在预处理后的各前景块区域上用扫描子窗口进行遍历, 提取HOG和LBP特征。

[0052] 其中, HOG(方向梯度直方图)特征是一种在计算机视觉和图像处理中用来进行物体检测的特征描述子, 它通过计算和统计图像局部区域的梯度方向直方图来构成特征。提取过程包括: 检测窗口; 归一化图像; 计算梯度; 对于每一个cell块对梯度直方图进行规定权重的投影; 对于每一个重叠block块内的cell进行对比度归一化。

[0053] LBP(局部二值模式)是一种用来描述图像局部纹理特征的算子; 它具有旋转不变性和灰度不变性等显著的优点。LBP算子定义为在3*3的窗口内, 以窗口中心像素为阈值, 将相邻的8个像素的灰度值与其进行比较, 若周围像素值大于中心像素值, 则该像素点的位置被标记为1, 否则为0。这样, 3*3邻域内的8个点经比较可产生8位二进制数(通常转换为十进制数即LBP码, 共256种), 即得到该窗口中心像素点的LBP值, 并用这个值来反映该区域的纹理信息。

[0054] 为了解决二进制模式过多的问题, 提高统计性, Ojala提出了采用一种“等价模式”来对LBP算子的模式种类进行降维。Ojala等认为, 在实际图像中, 绝大多数LBP模式最多只包含两次从1到0或从0到1的跳变。因此, Ojala将“等价模式”定义为: 当某个LBP所对应的循环二进制数从0到1或从1到0最多有两次跳变时, 该LBP所对应的二进制就称为一个等价模式类。因此对于3*3邻域内8个采样点来说, LBP特征从原始的256维降到了59维。通过这样的改进, 特征向量的维数减少, 而不会丢失任何信息, 同时减少高频噪声带来的影响。

[0055] 提取HOG和LBP特征的具体操作如下:

[0056] 1) 首先对输入视频进行转码处理, 将其转化为分辨率为352*288, 格式为avi的视频。

[0057] 2) 首先将车辆检测子窗口Block的大小设为 $2a \times 2a$, 每个Block划分为4个Cell, 每个Cell的大小设为 $a \times a$; 用车辆检测子窗口Block对视频帧从左到右从上至下进行扫描, 每次在X轴方向移动的步长设为a个像素, 在Y轴方向移动的步长设为a个像素。

[0058] 3) 然后将每个 $2a \times 2a$ Block的大小的图像块, 归一化为 $b \times b$ 的大小图像块($b \times b$

为训练所用的正负样本的大小)。

[0059] 4) 首先用opencv中自带HOG特征提取函数提取这个图像块的HOG特征,每帧检测视频提取M维的HOG特征向量的维数为M维。

[0060] 5) 然后在自己编写函数,提取LBP特征向量,具体操作如下:

[0061] a. 对于每个cell中的一个像素,将相邻的8个像素的灰度值与其进行比较,若周围像素值大于中心像素值,则该像素点的位置被标记为1,否则为0。这样,3*3邻域内的8个点经比较可产生8位二进制数,即得到该窗口中心像素点的LBP值;

[0062] b. 然后计算每个cell的直方图,即每个数字(假定是十进制数LBP值)出现的频率;然后对该直方图进行归一化处理;

[0063] c. 最后将得到的每个cell的统计直方图进行连接成为一个特征向量,也就是整幅图的LBP纹理特征向量,每帧检测视频提取的LBP特征向量的维数为N维。

[0064] S500、将提取的HOG和LBP特征级联,得到一个新的M+N维的特征行向量,将得到的新的级联特征向量用训练好的SVM分类器进行分类,判断是否为运动的车辆。

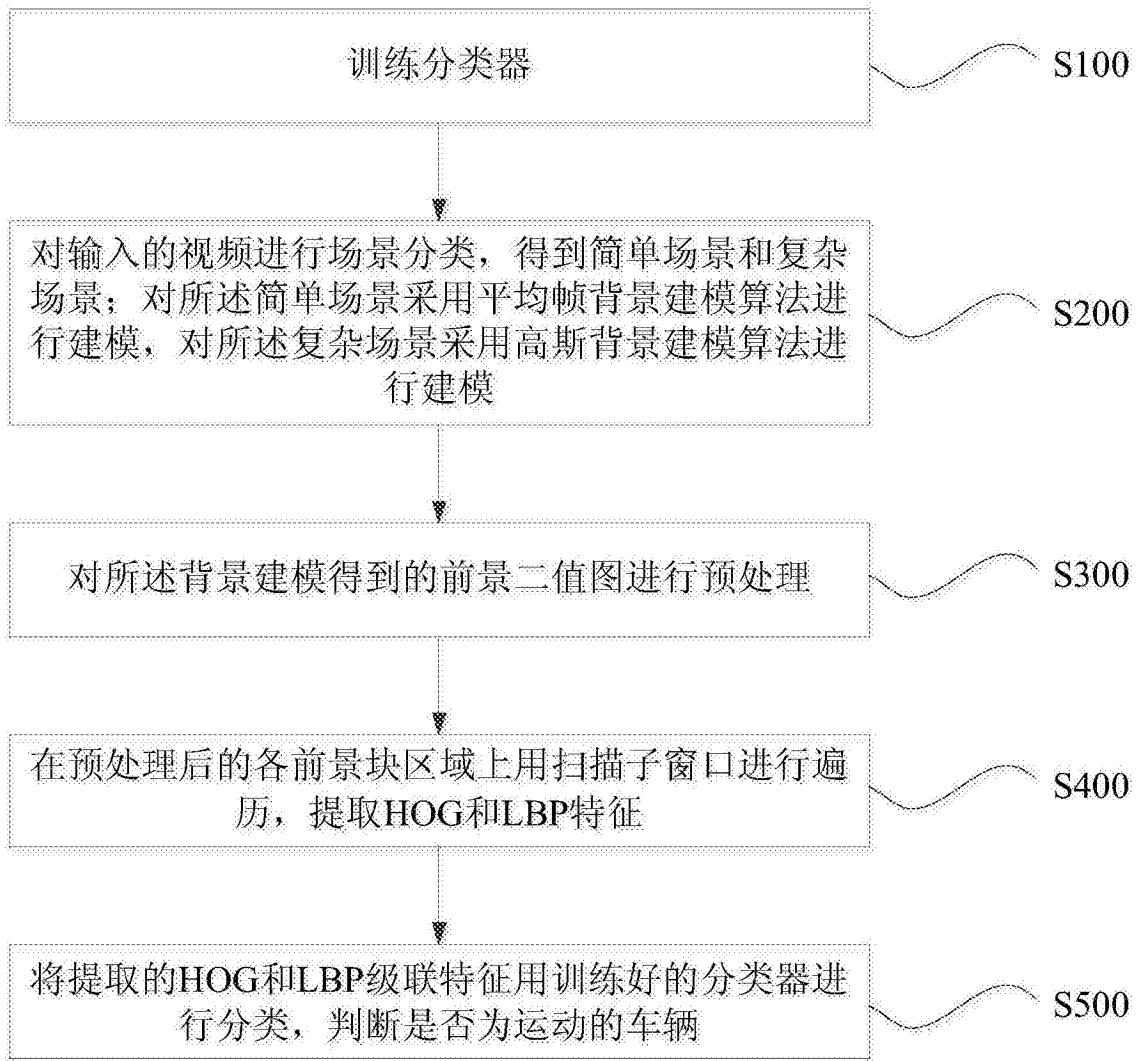


图1