



(12) 发明专利申请

(10) 申请公布号 CN 102982341 A

(43) 申请公布日 2013. 03. 20

(21) 申请号 201210434490. 7

(22) 申请日 2012. 11. 01

(71) 申请人 南京师范大学

地址 210097 江苏省南京市鼓楼区宁海路  
122 号

(72) 发明人 宋宏权 刘学军 阎国年 张兴国

(74) 专利代理机构 南京知识律师事务所 32207  
代理人 李媛媛

(51) Int. Cl.

G06K 9/62(2006. 01)

G06K 9/00(2006. 01)

G06K 9/54(2006. 01)

H04N 7/18(2006. 01)

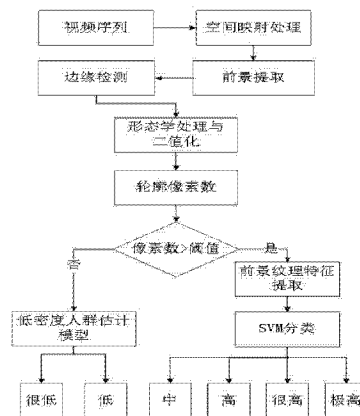
权利要求书 1 页 说明书 5 页 附图 2 页

(54) 发明名称

一种可跨摄像机的自适应人群密度估计方法

(57) 摘要

本发明公开了一种可跨摄像机的自适应人群密度估计方法,其步骤为:捕获视频监控信号,获取视频监控人群图像并对其进行空间映射处理;在地理参考下提取人群活动的前景图像;对前景图像进行边缘检测、形态学处理等操作;若前景边缘像素数目小于设定的阈值,则采用低密度人群估计模型计算人群密度,并根据人群密度等级标准进行分类;若前景边缘像素数目大于设定的阈值,则提取前景图像的纹理特征,利用跨摄像机支持向量机(SVM)人群密度分类器估计人群密度等级。本发明利用视频数据空间映射方法,将视频数据统一至地理参考,解决了不同监控设备的人群图像尺度多样化问题,克服了模型的场景依赖性,大大提高了人群密度估计模型的构建效率。



1. 一种可跨摄像机的自适应人群密度估计方法,其步骤为:

第一步、捕获视频监控信号,获取视频监控人群图像并对其进行空间映射处理;

第二步、自适应人群密度估计:基于空间映射处理的人群图像,在地理参考下提取人群活动的前景图像;对前景图像进行边缘检测、形态学处理和图像二值化操作,获取到前景边缘像素的数量;设定人群密度估计模型自适应选取阈值;若前景边缘像素数目小于人群密度估计模型自适应选取阈值,则采用低密度人群密度估计模型计算人群密度,并根据人群密度等级标准进行分类;若前景边缘像素数目大于该阈值,则提取前景图像的纹理特征,并利用高密度人群密度估计模型的跨摄像机 SVM 人群密度分类器估计人群密度等级。

2. 根据权利要求 1 所述的一种可跨摄像机的自适应人群密度估计方法,其特征在于,所述第一步进行空间映射处理的具体步骤为:

(1) 选取图像中的人群活动区域;

(2) 利用双灭点透视模型对选取的图像区域进行透视校正;

(3) 选取三组以上透视校正后的图像坐标及其对应的地理坐标,根据坐标的对应关系求算图像空间映射变换矩阵,完成监控人群图像的空间映射。

3. 根据权利要求 1 所述的一种可跨摄像机的自适应人群密度估计方法,其特征在于,所述第二步中采用的低密度和高密度人群密度估计模型的具体步骤分别为:

(a) 低密度人群密度估计模型:选取低密度人群图像样本,对其进行空间映射处理;在地理空间提取人群活动范围前景图像;对前景图像进行边缘检测、形态学处理和图像二值化操作;标定人群数量,根据像素数目与人群数量之间的关系,利用最小二乘法拟合跨摄像机低密度人群密度估计模型;根据人群数量与监控图像中的人群活动面积,求算人群密度值;

(b) 高密度人群密度估计模型:选取不同密度等级的高密度人群样本图像,对其进行空间映射处理;在地理空间提取人群活动范围前景图像;计算前景图像的灰度共生矩阵,提取纹理特征值;采用一对一判别策略,根据人群密度等级与纹理特征值,训练跨摄像机 SVM 分类器;在地理空间提取人群密度的图像纹理特征,利用 SVM 分类器确定人群密度等级。

4. 根据权利要求 1、2 或 3 所述一种可跨摄像机的自适应人群密度估计方法,其特征在于,所述设定人群密度估计模型自适应选取阈值的具体步骤为:

(1) 训练得到低密度人群密度估计模型;

(2) 利用空间映射处理后的人群图像,得到该摄像机监控场景对应的真实人群活动区域面积;

(3) 根据人群密度等级标准选取人群密度值,当人群密度高于此值时采用高密度人群密度估计模型;

(4) 利用摄像机监控场景的人群活动区域面积和人群密度阈值计算人群数量阈值;

(5) 利用计算得到的各摄像机监控场景中的人群数量阈值,根据低密度人群密度估计模型计算对应的像素数目,即可得到自适应选取模型的阈值。

## 一种可跨摄像机的自适应人群密度估计方法

### 技术领域

[0001] 本发明涉及一种可跨摄像机的人群密度估计模型构建方法,具体说,是一种克服了视频监控场景依赖性的人群密度估计模型构建方法。

### 背景技术

[0002] 随着社会经济的快速发展,娱乐活动、展览活动、体育赛事等大规模人群聚集活动频繁出现,因人群密度过大而造成的人员伤亡事故屡见不鲜。因此人群密度是对公众聚集场所进行风险评估与安全管理的重要指标,实时监控公众聚集场所的人群密度具有重要现实意义。视频本身兼具时空、表达直观、信息丰富、动态实时等特点,近年来,基于视频的智能化人群密度监控技术已成为研究热点。人群密度是人群分析中的重要因素。人群密度用来衡量公共场所的拥挤状况,同时也可用来检测公众聚集场所的危险程度。

[0003] 目前对人群密度的研究可分为三类:基于像素的方法、基于纹理分析的方法和基于人体分割的方法。基于像素的方法最早由 Davies 在文章“基于图像处理的人群监控”(Crowd monitoring using image processing, *Electronics & Communication Engineering Journal*, 1995, 7(1): 37-47; 《电子与通信工程杂志》,1995年,第7卷第1期,37-47)中提出,通过背景减提取人群前景,运用边缘检测法提取前景边缘像素数目,根据标定的人数拟合人群数量估计线性模型,将提取的前景边缘像素数输入估计模型可获得对应的人群数量。由于透视畸变效应的影响,人群前景像素与边缘像素数目随着其真实点距摄像机的远近产生近大远小现象。为了消除透视效应, Ma 等提出一种对人群前景图像进行透视校正处理进而估计人群数量的方法。该方法建立了透视校正后前景像素数和人群数量间的线性关系,并通过分析监控场景中的人群密度,实现对相关异常事件的检测(On pixel count based crowd density estimation for visual surveillance, *Cybernetics and Intelligent Systems Proceedings*, Singapore, 1-3 December 2004; 基于像素数的视频监控人群密度估计,2004年(国际控制与智能系统)学术会议论文集,新加坡,2004年12月1-3日)。

[0004] 基于像素的方法在人群密度较小时具有良好效果,随着人群密度增大,因行人间相互遮挡使得此类方法的线性关系不再成立。Marana 等根据人群图像的纹理模式,提出了基于纹理分析的人群密度估计方法(Automatic estimation of crowd density using texture, *Safety Science*, 1998, 28(3): 165-175; 基于纹理的人群密度自动估计,安全科学杂志,1998年,第28卷第3期,165-175)。此类方法认为:高密度人群在纹理上表现为细模式,低密度人群表现为粗模式。提取人群图像感兴趣区域的纹理特征,基于图像纹理分析,训练人群密度等级分类器,并使用密度等级分类器实现对人群密度的等级分类。

[0005] 基于人体分割的方法是在图像或视频序列中识别人群个体,进而统计人群数量。Lin 等提出一种利用 Haar 小波变换和支持向量机相结合的人群密度估计方法(Estimation of number of people in crowded scenes using perspective transformation, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 2001,

31(6): 645 - 654; 基于透视变换的人群数量估计, IEEE 系统、人和控制论汇刊, A 辑: 系统与人类, 2001 年第 31 卷第 6 期, 645-654)。首先利用 Haar 小波变换提取图像中类似头部的轮廓特征, 然后训练支持向量机分类器用于判断头部轮廓, 最后对提取的头部轮廓进行几何校正实现对人群数量的估计。此类方法的精度高于基于像素的方法和基于纹理分析的方法, 但只能应用于人群密度很低的场景, 当人群密度较高时, 由于遮挡和聚集导致很难识别分割出人群个体。

[0006] 现有的密度估计方法大都具有场景依赖性, 由于不同的摄像机具有不同的参数和安装设置, 针对不同的摄像机需要分别一一训练估计模型, 此工作不仅繁琐且浪费了大量的人力与时间。近年来, 相关学者在上述三种人群密度估计方法的基础上, 提出了跨场景人群密度估计模型构建方法。Kong 等将提取的图像特征标准化, 以解决透视效应和摄像机的不同方向问题, 可通过对某一个摄像机进行模型构建实现对其他摄像机的人群密度估计 (A viewpoint invariant approach for crowd counting, Pattern Recognition Proceedings, Hong Kong, 20-24 August 2006; 无视点依赖的人群数量估计方法, 2006 年 (国际模式识别) 学术会议论文集, 香港, 2006 年 8 月 20-24 日)。为了实现人群密度估计模型的跨场景使用, 这类方法需要利用摄像机标定、标准化处理、尺度相关性描述等方法计算不同摄像机间的尺度转换因子, 但此类方法同样相对比较繁琐, 如多尺度特征块的生成、特征模板的建立与匹配、摄像机标定等。

[0007] 总的来说, 目前大都针对特定监控场景研究人群密度估计算法, 其场景依赖性强, 且对监控设备的稳定性要求较高, 监控探头姿态有少许变化便会影响模型的精度。对于具有大量监控探头的监控区域, 模型训练工作量大, 浪费了大量的人力与时间。究其原因, 模型的场景依赖性是由监控图像的尺度多样化造成的。由于监控探头的安装高度、姿态以及内参不同, 导致监控图像中相同目标具有不同的表现特征, 使得特定场景的人群分析模型不具有普适性。

## 发明内容

[0008] 本发明所要解决的关键问题是克服人群密度估计模型的监控场景依赖性, 使得利用某监控摄像机样本数据构建的人群密度估计模型同样适用于其他摄像机, 实现人群密度估计模型的跨摄像机应用, 并根据人群密度高低程度自动选取相应的估计模型。因此, 本发明提出了一种可跨摄像机的自适应人群密度估计模型构建方法。

[0009] 本发明的基本思路: 利用视频数据空间映射方法将视频数据统一至地理参考, 以实现人群图像的尺度统一; 利用空间映射处理后的人群样本图像, 在地理空间构建人群密度估计模型, 包括低密度和高密度人群密度估计模型两类, 其中低密度模型采用基于像素的方法, 高密度人群密度估计采用纹理的方法, 利用支持向量机 (SVM) 实现人群密度的等级分类; 设定阈值用于判断低密度或高密度估计模型的选择, 从而实现跨摄像机的自适应人群密度估计。

[0010] 本发明一种可跨摄像机的自适应人群密度估计方法的基本步骤为:

[0011] 第一步、捕获视频监控信号, 获取视频监控人群图像并对其进行空间映射处理;

[0012] 第二步、自适应人群密度估计: 基于空间映射处理的人群图像, 在地理参考下提取人群活动的前景图像; 对前景图像进行边缘检测、形态学处理和图像二值化操作, 获取到前

景边缘像素的数量；设定人群密度估计模型自适应选取阈值；若前景边缘像素数目小于人群密度估计模型自适应选取阈值，则采用低密度人群密度估计模型计算人群密度，并根据人群密度等级标准进行分类；若前景边缘像素数目大于该阈值，则提取前景图像的纹理特征，并利用高密度人群密度估计模型的跨摄像机 SVM 人群密度分类器估计人群密度等级。

[0013] 第一步所述视频监控图像的空间映射具体步骤为：

[0014] (1) 选取图像中的人群活动区域；

[0015] (2) 利用双灭点透视模型(罗晓晖, 基于双灭点的图像透视变换方法, 计算机工程, 2009, 35(15):212-214.) 对选取的图像区域进行透视校正；

[0016] (3) 选取三组以上透视校正后的图像坐标及其对应的地理坐标, 根据坐标的对应关系求算图像空间映射变换矩阵, 完成监控人群图像的空间映射。

[0017] 第二步所述低密度人群密度估计模型的具体构建步骤为：

[0018] (1) 选取低密度人群图像样本, 对其进行空间映射处理；

[0019] (2) 在地理空间提取人群活动范围前景图像；

[0020] (3) 基于提取的前景图像, 对其进行边缘检测处理、形态学处理、图像二值化处理操作；

[0021] (4) 标定人群数量, 根据像素数目与人群数量之间的关系, 利用最小二乘法拟合跨摄像机低密度人群密度估计模型；

[0022] (5) 根据人群数量与监控图像中的人群活动面积, 求算人群密度值。

[0023] 所述(2)中的前景图像, 可通过帧差法或已知背景情况下的背景减法得到。

[0024] 第二步所述高密度人群密度估计模型的具体构建步骤为：

[0025] (1) 选取不同密度等级的高密度人群样本图像, 对其进行空间映射处理；

[0026] (2) 在地理空间提取人群活动范围前景图像；

[0027] (3) 计算前景图像的灰度共生矩阵, 提取纹理特征值；

[0028] (4) 采用一对一判别策略, 根据人群密度等级与纹理特征值, 训练跨摄像机 SVM 分类器；

[0029] (5) 在地理空间提取人群密度的图像纹理特征, 利用 SVM 分类器确定人群密度等级。

[0030] 第二步所述模型自适应选取阈值设定的具体步骤为：

[0031] (1) 训练得到低密度人群密度估计模型；

[0032] (2) 利用空间映射处理后的人群图像, 得到该摄像机监控场景对应的真实人群活动区域面积；

[0033] (3) 根据人群密度等级标准选取人群密度值, 当人群密度高于此值时采用高密度人群密度估计模型；

[0034] (4) 利用摄像机监控场景的人群活动区域面积和人群密度阈值计算人群数量阈值；

[0035] (5) 利用计算得到的各摄像机监控场景中的人群数量阈值, 根据低密度人群密度估计模型计算对应的像素数目, 即可得到自适应选取模型的阈值。

[0036] 本发明利用视频数据空间映射方法, 将视频数据统一至地理参考, 解决了不同监控设备的人群图像尺度多样化问题, 克服了模型的场景依赖性, 大大提高了人群密度估计

模型的构建效率。该方法实现简便,具有很强的适应性以及应用前景。

### 附图说明

[0037] 图 1 是视频数据空间映射方法;(a)是各平面空间关系图,(b)是小孔成像的原理图,(c)是两点透视控制点成像,(d)是单灭点透视校正图。

[0038] 图 2 是本发明跨摄像机低密度人群密度估计模型构建方法流程示意图。

[0039] 图 3 是本发明跨摄像机高密度人群密度估计模型构建方法流程示意图。

[0040] 图 4 是本发明跨摄像机自适应人群密度估计方法流程示意图。

### 具体实施方式

[0041] 下面结合附图和实施例,对本发明作进一步详细说明。

[0042] 第一步:相关设备准备。准备一台 ThinkPad X201i 型便携式笔记本电脑,现代 HYC-S200 多功能摄像头三个,Rikaline GPS-6033 型蓝牙 GPS 卫星接收机一个。

[0043] 第二步:视频数据空间映射。

[0044] 图 1 (a) 描述了视频数据与空间数据之间的几何关系。图中点 C 为摄像机位置,摄像机拍摄的图像映射至像平面 I,平面 T 为透视校正后的图像,G 为 GIS (Geographic Information System) 空间的地理参考平面。地理空间(G 平面)中的任一点  $P(x_g, y_g)$  在像平面中的位置为  $p(u, v)$ ,图像透视校正后在平面 T 中的位置为  $P_t(x_t, y_t)$ ,视频数据空间映射即为建立点 p 与点 P 的变换关系,实现图像空间 I 到地理空间 G 的映射。图 1 (b) 中平面 I 为二维像平面,0 为光学中心,o 为像主点。点  $p(u, v)$  是三维世界坐标  $P(x, y, z)$  在平面 I 上的投影,f 是焦距。空间中任一点 P 在平面 I 中的位置均可用此针孔模型近似表示,即任一点在平面 I 上的位置为光心 0 与该点连线与像平面的交点,其比例关系式如下:

$$[0045] \quad u = \frac{fx}{z} \quad v = \frac{fy}{z} \quad (1)$$

[0046] 图 1 (c) 中的四边形在监控场景中为矩形时则构成两点透视,其中  $m_1$  和  $m_2$  为灭点。对双灭点中任一灭点进行透视校正得到单灭点透视,再使用相同方法对剩下的灭点进行透视校正。图 1 (d) 为对一个灭点进行校正的示意图,需根据透视原理进行 X 和 Y 方向两次校正。首先旋转图像使边 ab 平行于 x 轴,根据控制点 a, b, c, d 的图像坐标,求取灭点坐标  $(m_x, m_y)$ 。对于 X 方向的校正,可选图像高度内任一水平线宽度作为标准宽度,本实施例选取图像最上端的宽度 W 作为标准宽度。将边 ac 校正为垂直于 x 轴的  $a'c'$ ,根据三角形相似求算 i 高度 ac 边在 x 方向的偏移量  $\Delta x_i$ ,则原图像中点  $(j, i)$  在校正后图像上的坐标为  $(j \pm \Delta x_i, i)$ ,计算公式为:

$$[0047] \quad \begin{cases} i_0 = i \\ j_0 = j + \left( (H - i) \times \frac{m_x - j}{m_y - i} \right) \end{cases} \quad (2)$$

[0048] 其中,  $(j, i)$  是原透视图像的坐标点;  $(j_0, i_0)$  是校正后图像的坐标; H 是图像高度;  $(m_x, m_y)$  是灭点坐标。根据小孔成像原理, Y 方向缩放比例与 X 方向相同,故可根据 X 方向

变换的比例关系对 Y 方向进行同样比例的校正,则 Y 方向校正公式为:

$$[0049] \quad \begin{cases} j_0 = j \\ i_0 = \frac{\frac{mx}{my-1}}{mx - (H-1) \times \frac{mx}{my-1}} \end{cases} \quad (3)$$

[0050] (1) 将三个摄像头以不同的角度放置在人群聚集的广场或路口,用于监控某区域的人群,并将其连接至笔记本电脑;

[0051] (2) 打开基于 Aforge.NET 库(一种开源的计算机视觉库)开发的视频监控程序,用于获取各摄像头的监控图像,根据图 1 所述对监控图像透视校正;

[0052] (3) 在各摄像头的监控图像中选取 4 个控制点,利用 GPS 接收机获取其地理坐标,进而求算各摄像头监控图像中人群活动区域的空间映射矩阵。

[0053] 第三步:跨摄像机自适应人群密度估计。

[0054] (1) 选取其中一个摄像头的低密度人群样本图像,并对其进行空间映射处理,在此基础上,在地理空间进行人群目标的前景提取、边缘检测、形态学处理和图像二值化操作,并根据标定的人数和前景像素数目的关系,进行低密度人群密度估计模型构建,具体参照附图 2;

[0055] (2) 选取其中一个摄像头不同等级的高密度人群样本图像,并对其进行空间映射处理,在地理空间提取人群目标的前景图像,并计算用于描述人群前景图像纹理特征的灰度共生矩阵,根据人群的密度等级和纹理特征值的关系,训练高密度人群的 SVM 分类器,进而对高密度人群进行分类,具体参照附图 3;

[0056] (3) 根据发明内容第二步所述设定模型的自适应选取阈值;

[0057] (4) 启动视频监控程序,对视频数据进行空间映射处理,提取人群目标前景图像,对前景图像进行边缘检测,并对边缘检测后的前景图像进行形态学处理和二值化,得到人群目标的边缘轮廓像素数目,若像素数目小于等于设定阈值,则选择低密度人群密度估计模型进行人群密度估计,否则选择高密度人群密度估计模型进行人群密度估计,具体参照附图 4。

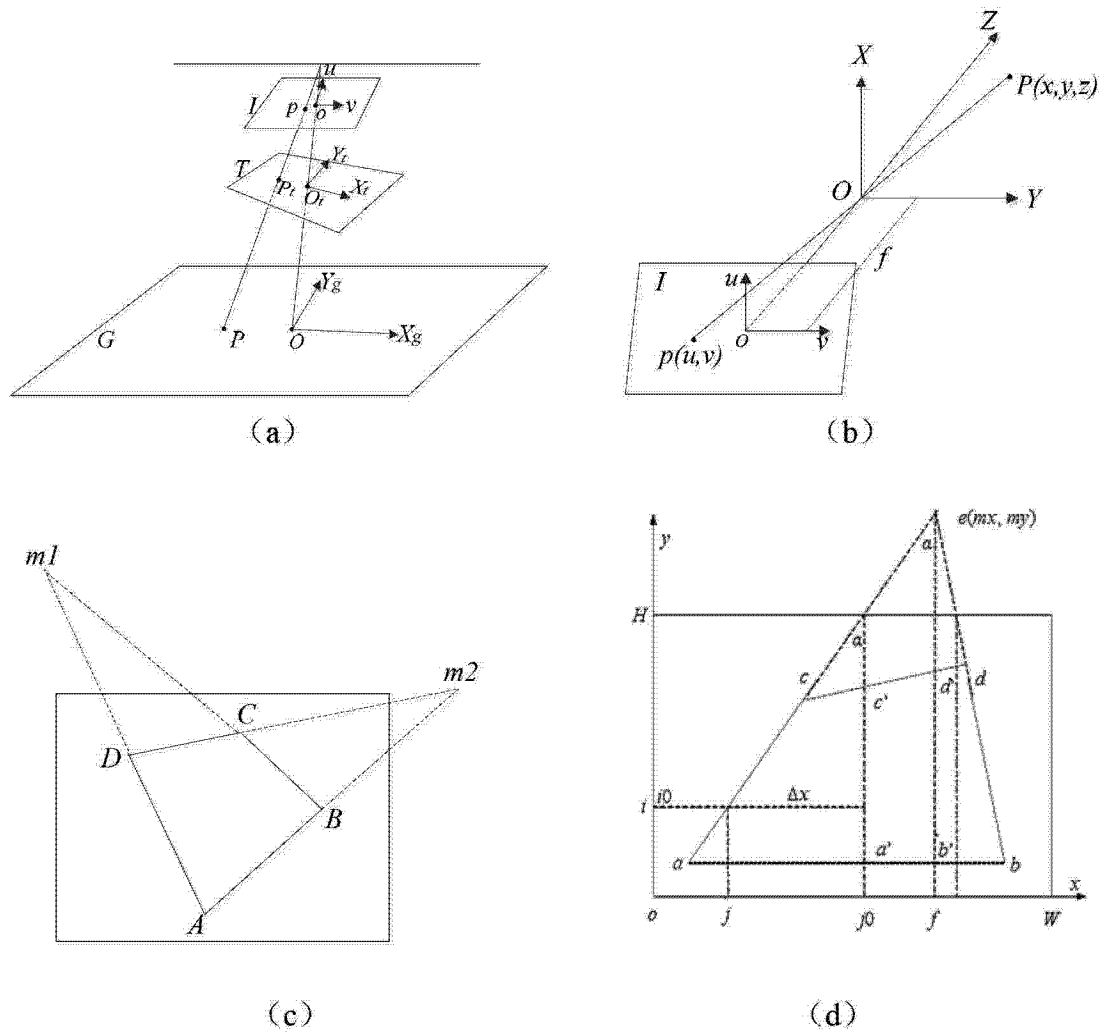


图 1

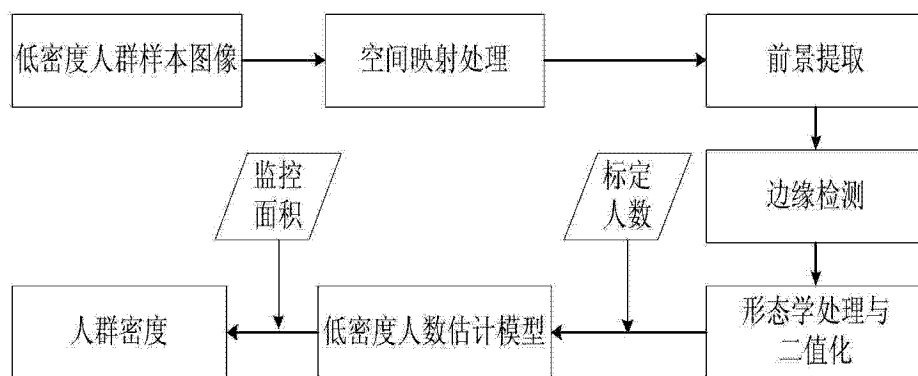


图 2



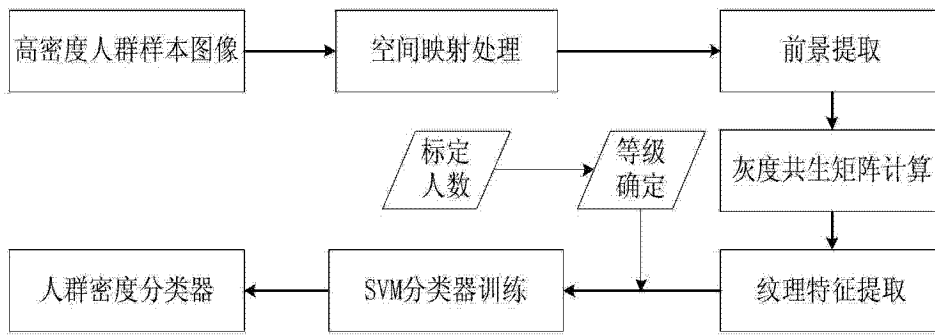


图 3

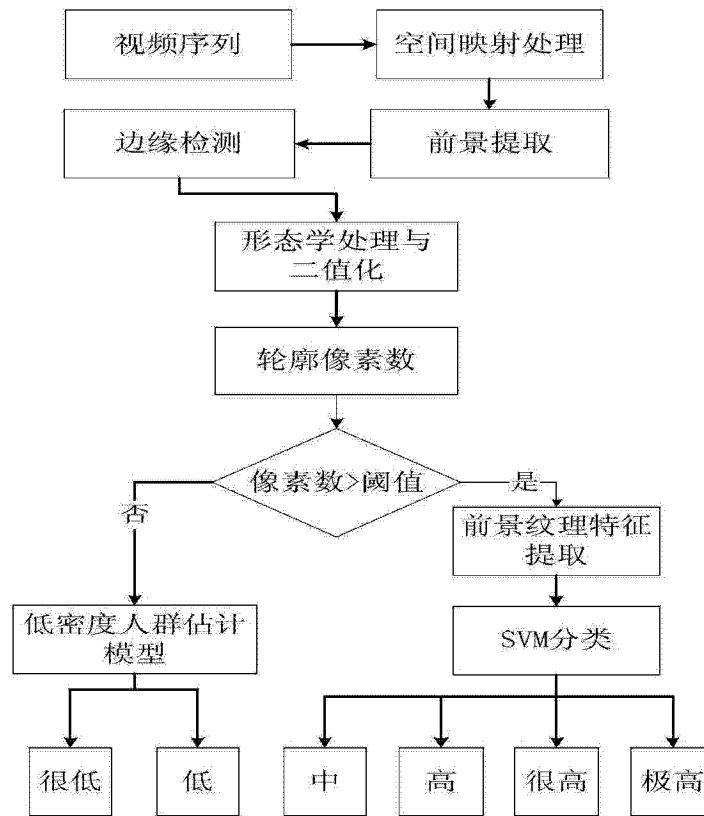


图 4