

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04N 7/18 (2006.01)

G06T 7/20 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200710041860. X

[43] 公开日 2008 年 12 月 17 日

[11] 公开号 CN 101325690A

[22] 申请日 2007. 6. 12

[21] 申请号 200710041860. X

[71] 申请人 上海正电科技发展有限公司

地址 200235 上海市中山西路 1525 号技贸大厦 6 楼

[72] 发明人 张宪民

[74] 专利代理机构 上海专利商标事务所有限公司
代理人 陈亮

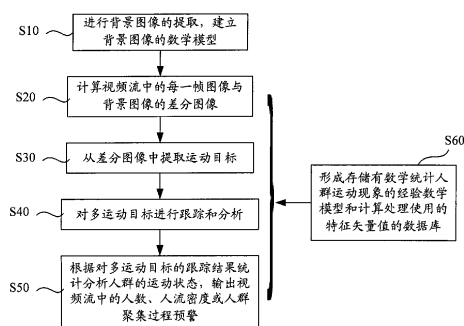
权利要求书 8 页 说明书 19 页 附图 4 页

[54] 发明名称

监控视频流中人流分析与人群聚集过程的检测方法及其系统

[57] 摘要

本发明公开了一种监控视频流中人流分析与人群聚集过程的检测方法和系统，自动监测人群运动状态输出人群运动状态统计数值。其技术方案为：该检测方法包括：提取背景图像，建立背景图像的数学模型；计算视频流中的每一帧图像与背景图像的差分图像；从差分图像中提取运动目标；对多运动目标进行跟踪和分析，建立运动目标的颜色空间分布模型，为每个目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描述搜索运动目标的主要特征，并在跟踪目标过程中对检测到的目标进行特征匹配以判断是否需要跟踪的目标；根据对多运动目标的跟踪结果统计分析人群的运动状态，输出视频流中的人数、人流密度或人群聚集过程预警。本发明应用于视频监控领域。



1 一种监控视频流中人流分析与人群聚集过程的检测方法，包括：

- (1) 进行背景图像的提取，建立背景图像的数学模型；
- (2) 计算视频流中的每一帧图像与该背景图像的差分图像；
- (3) 从差分图像中提取运动目标；

(4) 对多运动目标进行跟踪和分析，建立运动目标的颜色空间分布模型，为每个目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描述搜索运动目标的主要特征，并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否需要跟踪的目标；

(5) 根据对多运动目标的跟踪结果统计分析人群的运动状态，输出视频流中的人数、人流密度或人群聚集过程预警。

2 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，步骤 (1) 进一步包括：

选择视频流的前 N 帧来构建初始化背景，设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续 N 帧图像中像素点上的中值和均方差，以 $\lambda(x)$ 为背景图像的像素值来构建背景，在初始化背景中，初始化背景像素点 $B(x) = \lambda(x)$ ，其中 N 为预设的自然数；

利用一阶 Kalman 滤波模型更新背景图像：

$$B_{t+1}(x) = (1 - \alpha)B_t(x) + \alpha F_t(x)$$

$$\sigma_{t+1}^2(x) = (1 - \alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$$

其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点上的均方差值， α 是更新因子，且 $0 < \alpha < 1$ 。

3 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，步骤 (2) 进一步包括：

根据当前帧图像上某点的像素值，计算出差分图像值，由下式确定差分图像的二值掩膜：

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th=2.5\sigma_t(x)$ ，其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值；

对得到的二值掩膜图像进行滤波处理；

应用形态学算子对二值掩膜图像中的断裂或空洞进行处理，同时使运动目标区域的边界变得更加光滑；

利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点，采用基于前景区域的面积阈值法去除噪声：若相互连接部件的像素总数超过设定阈值，则该部件为独立运动部件，剩下的像素为随机噪声。

4 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，步骤（3）中，根据步骤（2）得到的对象二值模板，清除当前帧内不在模板内的像素点，得到运动目标的分割结果。

5 根据权利要求 4 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，步骤（3）中，对于多运动目标的场景，将多个运动目标分割成多个单运动目标并加以提取。

6 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，步骤（4）进一步包括：

（a）在跟踪前通过初始帧与背景差分获得初始的运动目标，通过区域生长法获得跟踪运动目标的颜色分布模型，如果是多运动目标的跟踪，可以在初始状态为每个跟踪运动目标建立颜色分布模型：

为每个运动目标提取 N 个颜色区域对象，并按区域的相对位置编号： $RO_i=(avc_i, r_i, i)$ ，其中 N 为自然数， i 为 1 至 N 的任一自然数， RO_i 为颜色区域对象， avc_i 为区域对象的代表色， r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率， $r_i = n_i / A_{vop}$ ， n_i 为区域包含的像素数目， A_{vop} 为运动目标的总像素数目， i 为区域编号；

（b）建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop}=(RO_1, RO_2, \dots, RO_N)$ ，计算运动目标的中心坐标 (\bar{x}, \bar{y}) ：
$$\bar{x} = \frac{1}{A_{vop}} \sum_{x \in VOP} x, \quad \bar{y} = \frac{1}{A_{vop}} \sum_{y \in VOP} y;$$

(c) 对当前每一帧图像, 与背景图像做差分后提取到 n 个候选的运动目标, 计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$, 假设前一帧中跟踪到的运动目标为 vop_{i-1}^G , 其颜色区域特征向量 $S_{vop_{i-1}^G} = (RO_1^G, RO_2^G, \dots, RO_N^G)$, 中心坐标 (\bar{x}^G, \bar{y}^G) , 计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离: $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$, $k = 1 \dots n$, 选取距离最小的候选目标 vop_i^D ;

(d) 根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$, 遍历候选目标 vop_i^D 的所有像素, 把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域, 并记录区域的像素数目 n_1 , 计算此区域的平均颜色值 avc_1^D , 编号为 1; 把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域, 并记录区域的像素数目 n_2 , 计算此区域的平均颜色值 avc_2^D , 编号为 2; 依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D = (avc_i^D, r_i, i)$, 并得到特征向量 $S_{vop_i^D} = (RO_1^D, RO_2^D, \dots, RO_N^D)$;

(e) 计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离:

$$d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$$

其中 i 表示跟踪到的目标 vop_{i-1}^G , j 表示候选目标 vop_i^D , m 为颜色区域编号, $s(avc_m(i), avc_m(j))$ 是颜色 $avc_m(i)$ 与颜色 $avc_m(j)$ 的相似度量, Th 为颜色相似度的阈值;

(f) 计算联合距离的公式为: $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j) \times \dots \times d_N(i, j)$, 在当前帧中, 如果 $d_{joint}(vop_{i-1}^G, vop_i^D)$ 小于预设的距离门限值 η_d , 则两者匹配, 此候选目标即为跟踪到的目标; 否则认为不匹配, 选取下一个与目标 vop_{i-1}^G 距离较小的候选目标, 做相同的匹配, 如此重复直到匹配到目标;

(g) 在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程: 对于颜色区域对象 $RO = (color, ratio, number)$ 的主分量 $color$, 使用自适应颜色更新模型: $c_i = (1 - \alpha)\bar{c} + \alpha c_{i-1}$, c_i 为第 i 帧的区域颜色值, c_{i-1} 为第 $i-1$ 帧的区域颜色值, \bar{c} 为前 i 帧区域颜色均值, α 为权重因子, $ratio$ 分量, 用当前帧正确跟

踪到的目标的颜色区域对象的相应的 ratio 分量作为新的值。而对应的编号 number 则保持不变。

7 根据权利要求 6 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，在步骤（4）中，如果当前帧候选运动目标的集合中没有匹配到目标，则继续处理下面的若干帧图像，若连续的若干帧匹配不到目标则认为目标已经消失，若能找到目标则又重新进入目标跟踪阶段，其中帧数是预设的。

8 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，在步骤（5）中，根据人群聚集面积确定是否进行人群聚集预警。

9 根据权利要求 1 所述的监控视频流中人流分析与人群聚集过程的检测方法，其特征在于，在人群运动状态的检测过程中，对视频流中的人群运动现象进行大量的数学统计以形成经验数学模型，建立摄像角度、焦距、空间坐标、人的像素面积和形状的特征矢量值，以数学表格的形式存储该些特征矢量值和该经验数学模型以在检测过程中使用。

10 一种监控视频流中人流分析与人群聚集过程的检测系统，包括：

背景提取装置，进行背景图像的提取，建立背景图像的数学模型；

差分图像计算装置，计算视频流中的每一帧图像与背景图像的差分图像；

运动目标提取装置，从差分图像中提取运动目标；

运动目标跟踪与分析装置，对多运动目标进行跟踪和分析，建立运动目标的颜色空间分布模型，为每个目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描述搜索运动目标的主要特征，并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否需要跟踪的目标；

人群运动状态统计分析装置，根据对多运动目标的跟踪结果统计分析人群的运动状态，输出视频流中的人数、人流密度或人群聚集过程预警。

11 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统，其特征在于，该背景提取装置进一步包括：

初始化背景像素点模块，选择视频流的前 N 帧来构建初始化背景，设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续 N 帧图像中像素点上的中值和均方差，以 $\lambda(x)$ 为背景图像的像素值来构建背景，在初始化背景中，初始化背景像素点 $B(x) = \lambda(x)$ ，其中 N 为预设的自然数；

滤波更新模块，利用一阶 Kalman 滤波模型更新背景图像：

$$B_{t+1}(x) = (1-\alpha)B_t(x) + \alpha F_t(x)$$

$$\sigma_{t+1}^2(x) = (1-\alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$$

其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点上的均方差值， α 是更新因子，且 $0 < \alpha < 1$ 。

12 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统，其特征在于，该差分图像计算装置进一步包括：

二值掩膜计算模块，根据当前帧图像上某点的像素值，计算出差分图像值，由下式确定差分图像的二值掩膜：

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th = 2.5\sigma_t(x)$ ，其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值；

滤波处理模块，对得到的二值掩膜图像进行滤波处理；

断裂空洞处理模块，应用形态学算子对二值掩膜图像中的断裂或空洞进行处理，同时使运动目标区域的边界变得更加光滑；

噪声去除模块，利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点，采用基于前景区域的面积阈值法去除噪声：若相互连接部件的像素总数超过设定阈值，则该部件为独立运动部件，剩下的像素为随机噪声。

13 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统,其特征在于,该运动目标提取装置清除当前帧内不在模板内的像素点,得到运动目标的分割结果;对于多运动目标的场景,通过内设的分割模块将多个运动目标分割成多个单运动目标并加以提取。

14 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统,其特征在于,该运动目标跟踪与分析装置进一步包括:

颜色分布模型建立模块,建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop}=(RO_1, RO_2, \dots, RO_N)$, 其中 $RO_i=(avc_i, r_i, i)$, N 为自然数, i 为 1 至 N 的任一自然数, RO_i 为颜色区域对象, avc_i 为区域对象的代表色, r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率, $r_i = n_i / A_{vop}$, n_i 为区域包含的像素数目, A_{vop} 为运动目标的总像素数目, i 为区域编号;

中心坐标计算模块,计算运动目标的中心坐标 (\bar{x}, \bar{y}) :
$$\bar{x} = \frac{1}{A_{vop}} \sum_{x \in VOP} x,$$

$$\bar{y} = \frac{1}{A_{vop}} \sum_{y \in vop} y;$$

候选目标选取模块,对当前每一帧图像,与背景图像做差分后提取到 n 个候选的运动目标,通过该中心坐标计算模块计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$, 假设前一帧中跟踪到的运动目标为 vop_{i-1}^G , 具有通过该颜色分布模型建立模块得到的颜色区域特征向量 $S_{vop_{i-1}^G}=(RO_1^G, RO_2^G, \dots, RO_N^G)$, 通过该中心坐标计算模块得到的中心坐标 (\bar{x}^G, \bar{y}^G) , 计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离: $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$, $k = 1 \dots n$, 选取距离最小的候选目标 vop_i^D ;

特征向量形成模块,根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$, 遍历候选目标 vop_i^D 的所有像素,把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域,并记录区域的像素数目 n_1 , 计算此区域的平均颜色值 avc_1^D , 编号为 1; 把

所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_2 ，计算此区域的平均颜色值 avc_2^D ，编号为 2；依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D = (avc_i^D, r_i, i)$ ，并得到特征向量 $S_{vop}^D = (RO_1^D, RO_2^D, \dots, RO_N^D)$ ；

颜色区域对象距离计算模块，计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离：
$$d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$$
，其中 i 表示跟踪到的目标 vop_{i-1}^G ， j 表示候选目标 vop_i^D ， m 为颜色区域编号， $s(avc_m(i), avc_m(j))$ 是颜色 $avc_m(i)$ 与颜色 $avc_m(j)$ 的相似度量， Th 为颜色相似度的阈值；

匹配判断模块，计算联合距离的公式为： $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j) \times \dots \times d_N(i, j)$ ，在当前帧中，如果 $d_{joint}(vop_{i-1}^G, vop_i^D)$ 小于预设的距离门限值 η_d ，则两者匹配，此候选目标即为跟踪到的目标；否则认为不匹配，选取下一个与目标 vop_{i-1}^G 距离较小的候选目标，做相同的匹配，如此重复直到匹配到目标；

颜色分布模型更新模块，在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程：对于颜色区域对象 $RO = (color, ratio, number)$ 的主分量 $color$ ，使用自适应颜色更新模型： $c_i = (1 - \alpha)\bar{c} + \alpha c_{i-1}$ ， c_i 为第 i 帧的区域颜色值， c_{i-1} 为第 $i-1$ 帧的区域颜色值， \bar{c} 为前 i 帧区域颜色均值， α 为权重因子， $ratio$ 分量，用当前帧正确跟踪到的目标的颜色区域对象的相应的 $ratio$ 分量作为新的值。而对应的编号 $number$ 则保持不变。

15 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统，其特征在于，如果匹配判断模块判断出当前候选运动目标的集合中没有匹配到目标，则继续处理下面的若干帧图像，若连续的若干帧匹配不到目标则认为目标已经消失，若能找到目标则又重新进入目标跟踪阶段，其中帧数是预设的。

16 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统，其特征在于，该人群运动状态统计分析装置是根据人群聚集面积确定是否需要

进行人群聚集预警。

17 根据权利要求 10 所述的监控视频流中人流分析与人群聚集过程的检测系统，其特征在于，系统还包括：

供检测过程使用的人群运动特征数据库，与该差分图像计算装置、该运动目标提取装置、该运动目标跟踪与分析装置以及该人群运动状态统计分析装置连接，以数学表格的形式存储了对视频流中的人群运动现象进行大量的数学统计而形成的经验数学模型以及摄像角度、焦距、空间坐标、人的像素面积和形状的特征矢量值。

监控视频流中人流分析与人群聚集过程的检测方法及系统

技术领域

本发明涉及一种对人流分析和人群聚集过程进行检测的方法与系统，尤其涉及一种应用计算机图像处理技术自动监测人群运动状态和及时准确输出人群运动状态统计数值的方法及系统。

背景技术

目前我国各大中型城市中开始分期分批大量安装社会公共安全视频监控系系统，主要有“街道治安视频监控系统”、“金融机构视频治安监控系统”、“重要公共场所(车站、机场、商业街区、大型运动场、大型超市等)治安视频监控系统”。这些系统的投入使用为保障社会治安的稳定和人民的安定生活服务奠定了坚实的基础，在实际工作中取得了很大的成效。

在一些监控系统中，对人群运动状态：人流方向、人数统计和人群聚集过程等检测十分重要。例如：（1）对政府大院楼门前、重要广场人流、人群聚集过程的检测可以预先告警非法聚众闹事、非法游行上访请愿等事态出现；（2）对奥运场馆、世博会场等大型群体活动广场的人流、人群聚集过程的检测可以预先告警公共安全管理部门对聚众闹事事态的防止、警力调配、人群疏导；（3）对银行金库保管工作人员的人数检测，可以防止违规操作如一人冒领两人打开库门等；（4）对大型超市、车站、码头、机场的人流、人群聚集和人数统计检测可以预先告警公共安全管理部门对拥挤事态的防止、人群疏导以及突发事件的预案启动。

目前的检测方法仅停留在对视频图像的存储记录，需要人工长时间的监控，由人的主观经验来判断人群的运动状态：人多、人少、拥挤还是疏散。这种人工方式存在以下缺点：（1）不能做到数量上的统计；（2）存在难以适应长时间的繁重浏览工作及因眼睛疲劳而遗漏重要信息的问题；（3）不能做到统计意义上的人群数量估计。

发明内容

本发明的目的在于解决上述问题，提供了一种监控视频流中人流分析与人群聚集过程的检测方法，通过计算机图像处理技术实现机器自动监测人群运动状态，及时准确地输出人群运动状态统计数值，为传统的视频监控系统带来新的突破。

本发明的又一目的在于提供了一种监控视频流中人流分析与人群聚集过程的检测系统，通过计算机图像处理技术实现机器自动监测人群运动状态，及时准确地输出人群运动状态统计数值，为传统的视频监控系统带来新的突破。

本发明的技术方案为：本发明提出了一种监控视频流中人流分析与人群聚集过程的检测方法，包括：

- (1) 进行背景图像的提取，建立背景图像的数学模型；
- (2) 计算视频流中的每一帧图像与该背景图像的差分图像；
- (3) 从差分图像中提取运动目标；

(4) 对多运动目标进行跟踪和分析，建立运动目标的颜色空间分布模型，为每个目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描述搜索运动目标的主要特征，并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否需要跟踪的目标；

(5) 根据对多运动目标的跟踪结果统计分析人群的运动状态，输出视频流中的人数、人流密度或人群聚集过程预警。

上述的监控视频流中人流分析与人群聚集过程的检测方法，其中，步骤(1)进一步包括：

选择视频流的前N帧来构建初始化背景，设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续N帧图像中像素点上的中值和均方差，以 $\lambda(x)$ 为背景图像的像素值来构建背景，在初始化背景中，初始化背景像素点 $B(x)=\lambda(x)$ ，其中N为预设的自然数；

利用一阶Kalman滤波模型更新背景图像：

$$B_{t+1}(x) = (1-\alpha)B_t(x) + \alpha F_t(x)$$

$$\sigma_{t+1}^2(x) = (1-\alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$$

其中 $B_t(x)$ 为t时刻x点上的背景像素值， $F_t(x)$ 为t时刻x点上的当前像素值， $\sigma_t(x)$ 为t时刻背景图像上x点上的均方差值， α 是更新因子，且 $0 < \alpha < 1$ 。

上述的监控视频流中人流分析与人群聚集过程的检测方法，其中，步骤(2)

进一步包括:

根据当前帧图像上某点的像素值, 计算出差分图像值, 由下式确定差分图像的二值掩膜:

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th=2.5\sigma_t(x)$, 其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值, $F_t(x)$ 为 t 时刻 x 点上的当前像素值, $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值;

对得到的二值掩膜图像进行滤波处理;

应用形态学算子对二值掩膜图像中的断裂或空洞进行处理, 同时使运动目标区域的边界变得更加光滑;

利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点, 采用基于前景区域的面积阈值法去除噪声: 若相互连接部件的像素总数超过设定阈值, 则该部件为独立运动部件, 剩下的像素为随机噪声。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 步骤 (3) 中, 根据步骤 (2) 得到的对象二值模板, 清除当前帧内不在模板内的像素点, 得到运动目标的分割结果。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 步骤 (3) 中, 对于多运动目标的场景, 将多个运动目标分割成多个单运动目标并加以提取。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 步骤 (4) 进一步包括:

(a) 在跟踪前通过初始帧与背景差分获得初始的运动目标, 通过区域生长法获得跟踪运动目标的颜色分布模型, 如果是多运动目标的跟踪, 可以在初始状态为每个跟踪运动目标建立颜色分布模型:

为每个运动目标提取 N 个颜色区域对象, 并按区域的相对位置编号: $RO_i = (avc_i, r_i, i)$, 其中 N 为自然数, i 为 1 至 N 的任一自然数, RO_i 为颜色区域对象, avc_i 为区域对象的代表色, r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率, $r_i = n_i / A_{vop}$, n_i 为区域包含的像素数目, A_{vop} 为运动目标的总像素数目, i 为区域编号;

(b) 建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop} = (RO_1, RO_2, \dots, RO_N)$,

计算运动目标的中心坐标 (\bar{x}, \bar{y}) : $\bar{x} = \frac{1}{A_{vop}} \sum_{x \in VOP} x$, $\bar{y} = \frac{1}{A_{vop}} \sum_{y \in vop} y$;

(c) 对当前每一帧图像, 与背景图像做差分后提取到 n 个候选的运动目标, 计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$, 假设前一帧中跟踪到的运动目标为 vop_{i-1}^G , 其颜色区域特征向量 $S_{vop_{i-1}^G} = (RO_1^G, RO_2^G, \dots, RO_N^G)$, 中心坐标 (\bar{x}^G, \bar{y}^G) , 计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离: $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$, $k = 1 \dots n$, 选取距离最小的候选目标 vop_i^D ;

(d) 根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$, 遍历候选目标 vop_i^D 的所有像素, 把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域, 并记录区域的像素数目 n_1 , 计算此区域的平均颜色值 avc_1^D , 编号为 1; 把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域, 并记录区域的像素数目 n_2 , 计算此区域的平均颜色值 avc_2^D , 编号为 2; 依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D = (avc_i^D, r_i, i)$, 并得到特征向量 $S_{vop_i^D} = (RO_1^D, RO_2^D, \dots, RO_N^D)$;

(e) 计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离:

$$d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$$

其中 i 表示跟踪到的目标 vop_{i-1}^G , j 表示候选目标 vop_i^D , m 为颜色区域编号, $s(avc_m(i), avc_m(j))$ 是颜色 $avc_m(i)$ 与颜色 $avc_m(j)$ 的相似度量, Th 为颜色相似度的阈值;

(f) 计算联合距离的公式为: $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j) \times \dots \times d_n(i, j)$, 在当前帧中, 如果 $d_{joint}(vop_{i-1}^G, vop_i^D)$ 小于预设的距离门限值 η_d , 则两者匹配, 此候选目标即为跟踪到的目标; 否则认为不匹配, 选取下一个与目标 vop_{i-1}^G 距离较小的候选目标, 做相同的匹配, 如此重复直到匹配到目标;

(g) 在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程: 对于颜色区域对象 $RO = (color, ratio, number)$ 的主分量 $color$, 使用自适应颜色更新模型: $c_i = (1 - \alpha)\bar{c} + \alpha c_{i-1}$, c_i 为第 i 帧的区域颜色值, c_{i-1} 为第 $i-1$ 帧的区域颜

色值, \bar{c} 为前 i 帧区域颜色均值, α 为权重因子, ratio 分量, 用当前帧正确跟踪到的目标的颜色区域对象的相应的 ratio 分量作为新的值。而对应的编号 number 则保持不变。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 在步骤(4)中, 如果当前帧候选运动目标的集合中没有匹配到目标, 则继续处理下面的若干帧图像, 若连续的若干帧匹配不到目标则认为目标已经消失, 若能找到目标则又重新进入目标跟踪阶段, 其中帧数是预设的。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 在步骤(5)中, 根据人群聚集面积确定是否进行人群聚集预警。

上述的监控视频流中人流分析与人群聚集过程的检测方法, 其中, 在人群运动状态的检测过程中, 对视频流中的人群运动现象进行大量的数学统计以形成经验数学模型, 建立摄像角度、焦距、空间坐标、人的像素面积和形状的特征矢量值, 以数学表格的形式存储该些特征矢量值和该经验数学模型以在检测过程中使用。

基于上述方法, 本发明另外提出了一种监控视频流中人流分析与人群聚集过程的检测系统, 包括:

背景提取装置, 进行背景图像的提取, 建立背景图像的数学模型;

差分图像计算装置, 计算视频流中的每一帧图像与背景图像的差分图像;

运动目标提取装置, 从差分图像中提取运动目标;

运动目标跟踪与分析装置, 对多运动目标进行跟踪和分析, 建立运动目标的颜色空间分布模型, 为每个目标建立至少一个颜色区域对象, 以多运动颜色块及其组合结构为描述搜索运动目标的主要特征, 并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否需要跟踪的目标;

人群运动状态统计分析装置, 根据对多运动目标的跟踪结果统计分析人群的运动状态, 输出视频流中的人数、人流密度或人群聚集过程预警。

上述的监控视频流中人流分析与人群聚集过程的检测系统, 其中, 该背景提取装置进一步包括:

初始化背景像素点模块, 选择视频流的前 N 帧来构建初始化背景, 设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续 N 帧图像中像素点上的中值和均方差, 以 $\lambda(x)$ 为背景图像的像

素值来构建背景，在初始化背景中，初始化背景像素点 $B(x) = \lambda(x)$ ，其中 N 为预设的自然数；

滤波更新模块，利用一阶 Kalman 滤波模型更新背景图像：

$$B_{t+1}(x) = (1-\alpha)B_t(x) + \alpha F_t(x)$$

$$\sigma_{t+1}^2(x) = (1-\alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$$

其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点上的均方差值， α 是更新因子，且 $0 < \alpha < 1$ 。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，该差分图像计算装置进一步包括：

二值掩膜计算模块，根据当前帧图像上某点的像素值，计算出差分图像值，由下式确定差分图像的二值掩膜：

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th = 2.5\sigma_t(x)$ ，其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值；

滤波处理模块，对得到的二值掩膜图像进行滤波处理；

断裂空洞处理模块，应用形态学算子对二值掩膜图像中的断裂或空洞进行处理，同时使运动目标区域的边界变得更加光滑；

噪声去除模块，利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点，采用基于前景区域的面积阈值法去除噪声：若相互连接部件的像素总数超过设定阈值，则该部件为独立运动部件，剩下的像素为随机噪声。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，该运动目标提取装置清除当前帧内不在模板内的像素点，得到运动目标的分割结果；对于多运动目标的场景，通过内设的分割模块将多个运动目标分割成多个单运动目标并加以提取。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，该运动目标跟踪与分析装置进一步包括：

颜色分布模型建立模块，建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop}=(RO_1, RO_2, \dots, RO_N)$ ，其中 $RO_i=(avc_i, r_i, i)$ ， N 为自然数， i 为 1 至 N 的任一自然数， RO_i 为颜色区域对象， avc_i 为区域对象的代表色， r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率， $r_i = n_i / A_{vop}$ ， n_i 为区域包含的像素数目， A_{vop} 为运动目标的总像素数目， i 为区域编号；

中心坐标计算模块，计算运动目标的中心坐标 (\bar{x}, \bar{y}) ：
$$\bar{x} = \frac{1}{A_{vop}} \sum_{x \in VOP} x,$$
$$\bar{y} = \frac{1}{A_{vop}} \sum_{y \in vop} y;$$

候选目标选取模块，对当前每一帧图像，与背景图像做差分后提取到 n 个候选的运动目标，通过该中心坐标计算模块计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$ ，假设前一帧中跟踪到的运动目标为 vop_{i-1}^G ，具有通过该颜色分布模型建立模块得到的颜色区域特征向量 $S_{vop_{i-1}^G}=(RO_1^G, RO_2^G, \dots, RO_N^G)$ ，通过该中心坐标计算模块得到的中心坐标 (\bar{x}^G, \bar{y}^G) ，计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离： $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$ ， $k = 1 \dots n$ ，选取距离最小的候选目标 vop_i^D ；

特征向量形成模块，根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$ ，遍历候选目标 vop_i^D 的所有像素，把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_1 ，计算此区域的平均颜色值 avc_1^D ，编号为 1；把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_2 ，计算此区域的平均颜色值 avc_2^D ，编号为 2；依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D=(avc_i^D, r_i, i)$ ，并得到特征向量 $S_{vop_i^D}=(RO_1^D, RO_2^D, \dots, RO_N^D)$ ；

颜色区域对象距离计算模块，计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离：
$$d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$$
，其中 i 表示跟踪到

的目标 vop_{i-1}^G ， j 表示候选目标 vop_i^D ， m 为颜色区域编号， $s(avo_m(i),avo_m(j))$ 是颜色 $avo_m(i)$ 与颜色 $avo_m(j)$ 的相似度量， Th 为颜色相似度的阈值；

匹配判断模块，计算联合距离的公式为： $d_{joint}(i,j)=d_1(i,j) \times d_2(i,j) \times \dots \times d_n(i,j)$ ，在当前帧中，如果 $d_{joint}(vop_{i-1}^G,vop_i^D)$ 小于预设的距离门限值 η_d ，则两者匹配，此候选目标即为跟踪到的目标；否则认为不匹配，选取下一个与目标 vop_{i-1}^G 距离较小的候选目标，做相同的匹配，如此重复直到匹配到目标；

颜色分布模型更新模块，在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程：对于颜色区域对象 $RO=(color, ratio, number)$ 的主分量 $color$ ，使用自适应颜色更新模型： $c_i=(1-\alpha)\bar{c}+\alpha c_{i-1}$ ， c_i 为第 i 帧的区域颜色值， c_{i-1} 为第 $i-1$ 帧的区域颜色值， \bar{c} 为前 i 帧区域颜色均值， α 为权重因子， $ratio$ 分量，用当前帧正确跟踪到的目标的颜色区域对象的相应的 $ratio$ 分量作为新的值。而对应的编号 $number$ 则保持不变。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，如果匹配判断模块判断出当前候选运动目标的集合中没有匹配到目标，则继续处理下面的若干帧图像，若连续的若干帧匹配不到目标则认为目标已经消失，若能找到目标则又重新进入目标跟踪阶段，其中帧数是预设的。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，该人群运动状态统计分析装置是根据人群聚集面积确定是否需要进行人群聚集预警。

上述的监控视频流中人流分析与人群聚集过程的检测系统，其中，系统还包括：

供检测过程使用的人群运动特征数据库，与该差分图像计算装置、该运动目标提取装置、该运动目标跟踪与分析装置以及该人群运动状态统计分析装置连接，以数学表格的形式存储了对视频流中的人群运动现象进行大量的数学统计而形成的经验数学模型以及摄像角度、焦距、空间坐标、人的像素面积和形状的特征矢量值。

本发明对比现有技术有如下的有益效果：本发明通过提取背景图像，从每一帧图像与背景图像的差分图像中提取运动目标，然后为运动目标建立颜色空间分布模型，为每个目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描

述搜索运动目标的主要特征,并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否为需要跟踪的目标,最后根据对多运动目标的跟踪结果统计分析人群的运动状态,输出视频流中的人数、人流密度或人群聚集过程预警。

附图说明

图 1 是本发明的监控视频流中人流分析与人群聚集过程的检测方法的一个较佳实施例的流程图。

图 2 是图 1 实施例中步骤 S10 的子流程图。

图 3 是图 1 实施例中步骤 S20 的子流程图。

图 4 是图 1 实施例中步骤 S40 的子流程图。

图 5 是本发明的监控视频流中人流分析与人群聚集过程的检测系统的一个较佳实施例的框图。

图 6 是图 5 实施例的背景提取装置的子框图。

图 7 是图 5 实施例的差分图像计算装置的子框图。

图 8 是图 5 实施例的运动目标跟踪与分析装置的子框图。

具体实施方式

下面结合附图和实施例对本发明作进一步的描述。

图 1 示出了本发明的监控视频流中人流分析与人群聚集过程的检测方法的一个较佳实施例的流程。请参见图 1,下面是对该流程中各步骤的详细描述。在视频序列中,运动使得目标不同于背景,运动信息成为从背景中提取目标(Video Object Plane, VOP)的重要依据。常见的运动目标检测算法主要有基于相邻帧差分的算法和基于背景图像和当前帧差分的算法。针对获取视频流的摄像机是静止的情况,本发明采用基于背景图像和当前帧差分的算法对运动目标进行检测和提取。

步骤 S10: 进行背景图像的提取,建立背景图像的数学模型。

请同时参见图 2,示出了背景提取的子流程。

步骤 S100: 构建初始化背景。选择视频流的前 N 帧来构建初始化背景,设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续 N 帧图像中像素点上的中值和均方差,以 $\lambda(x)$ 为背景图像的像素值来构建背景,在初始化背景中,初始化背景像素点 $B(x) = \lambda(x)$, 其中 N

为预设的自然数。

步骤 S110: 滤波更新背景图像。考虑到光照变化、摄像头的抖动等不可测的因素的影响, 会使背景发生变化, 导致运动目标的检测出现误差, 因此必须对背景模型进行更新以使之具有鲁棒性与自适应性。

采用一阶 Kalman 滤波模型来更新背景图像, 如下公式所示:

$$B_{t+1}(x) = (1 - \alpha)B_t(x) + \alpha F_t(x)$$

$$\sigma_{t+1}^2(x) = (1 - \alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$$

其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值, $F_t(x)$ 为 t 时刻 x 点上的当前像素值, $\sigma_t(x)$ 为 t 时刻背景图像上 x 点上的均方差值, α 是更新因子, 且 $0 < \alpha < 1$ 。背景模型需要定时更新, 更新频率的策略与具体的应用环境有关。

步骤 S20: 计算视频流中的每一帧图像与背景图像的差分图像。

请参见图 3, 示出了差分图像的计算过程。

步骤 S200: 根据当前帧图像上某点的像素值, 计算出差分图像值, 由下式确定差分图像的二值掩膜:

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th = 2.5\sigma_t(x)$, $B_t(x)$ 为 t 时刻 x 点上的背景像素值, $F_t(x)$ 为 t 时刻 x 点上的当前像素值, $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值。

由于目标与背景区域可能存在颜色相近的像素, 这样提取出的前景区域可能发生断裂或者出现空洞, 同时背景区域也会出现孤立的噪声。因此需要进行以下步骤。

步骤 S210: 对得到的二值掩膜图像进行滤波处理。

步骤 S220: 应用形态学算子对二值掩膜图像中的断裂或空洞进行处理, 同时形态学滤波还可以使得运动目标区域的边界变得更加光滑。

步骤 S230: 利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点, 采用基于前景区域的面积阈值法去除噪声: 若相互连接部件的总数超过设定的阈值, 则该部件为独立运动部件, 剩下的为随机噪声。

步骤 S30: 从差分图像中提取运动目标。根据对象二值模板, 清除当前帧不在模板内的像素点就可以得到运动图像的分割结果。对于多目标的场景, 可以将多个目标分割成一个个的单运动目标区域, 并提取出运动目标。

步骤 S40: 对多运动目标进行跟踪和分析。主要过程是: 建立运动目标的颜色分布模型, 为每个目标建立至少一个颜色区域对象, 以多运动颜色块及其组合结构为描述搜索运动目标的主要特征, 并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否为需要跟踪的目标。

请参见图 4, 示出了对多运动目标的跟踪和分析流程。

步骤 S400: 在跟踪前通过初始帧与背景差分获得初始的运动目标, 通过区域生长法获得跟踪运动目标的颜色分布模型。如果是多运动目标的跟踪, 可以在初始状态为每个跟踪运动目标建立颜色分布模型。

为每个运动目标提取 N 个颜色区域对象, 并按区域的相对位置编号: $RO_i = (avc_i, r_i, i)$, 其中 N 为自然数, i 为 1 至 N 的任一自然数, RO_i 为颜色区域对象, avc_i 为区域对象的代表色, r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率, $r_i = n_i / A_{vop}$, n_i 为区域包含的像素数目, A_{vop} 为运动目标的总像素数目, i 为区域编号。

举例来说, 在视频监控系统中获得图像序列所包含的目标(人物、车辆等)常常具有某种较明显的颜色分布, 人眼很容易给出运动目标的一个或者几个主色调, 分辨出感兴趣目标所在的位置。参考这一现象, 可以将运动目标表示成一个或多个主色区域组成, 这些主色同时构成一种固定的分布关系。如果以一幅视频图像中的人作为运动目标, 主色可分辨为黑色的头发, 上身为黄色上衣, 下身为黑色裤子等。因此将运动目标通过一个或几个颜色区域对象 RO 来表示, 并对区域按相对位置编号。 RO 包括区域的代表色、区域面积比率(区域像素数目占目标总像素数目的比率)以及区域的编号, 即 $RO = (color, ratio, number)$ 。

在跟踪前通过初始帧与背景差分获得初始的运动目标, 通过区域生长法来获得待跟踪目标的颜色分布模型。当然, 对多运动目标跟踪的情形, 可以在初始状态为每个目标建立颜色分布模型。

定义颜色之间的距离阈值 $T1$, 若空间位置相邻的两像素(使用 8 连通方式), 颜色距离低于阈值 $T1$, 则此两个像素可以合并成一个区域, 此区域一直增长到没有相邻像素可以合并为止。为避免同一区域里存在着颜色差异太大的像素, 我们另外设定一个阈值 $T2$, 在区域生长的过程中, 若某像素的颜色与此区域的所有像素的距离大于阈值 $T2$, 则不合并此像素。若目标还有未被合并的像素,

则继续新的区域生长，直到所有的像素属于某个区域为止。对于过小的区域，我们采用的作法是，将之与周围的区域比较，并归入到颜色距离最小的区域中去。计算每个区域的平均颜色，以及区域包含的像素数目占目标像素总数的比率，这样得到最终的几个颜色区域对象 RO 。

步骤 S410：建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop}=(RO_1, RO_2, \dots, RO_N)$ ，计算运动目标的中心坐标 (\bar{x}, \bar{y}) ：

$$\bar{x} = \frac{1}{A_{vop}} \sum_{x \in VOP} x,$$

$$\bar{y} = \frac{1}{A_{vop}} \sum_{y \in vop} y.$$

在本实施例中，以提取 2 个颜色区域对象为例，按照区域的相对位置编号， $RO_1=(avc_1, r_1, 1)$ 和 $RO_2=(avc_2, r_2, 2)$ 。建立运动目标 VOP 的颜色区域特征向量 $s_{vop}=(RO_1, RO_2)$ 。并计算目标的中心坐标 (\bar{x}, \bar{y}) ，其中 A_{vop} 为目标包含的象素数目， $r_i = n_i / A_{vop}$ ， $i=1, 2$ 。

在步骤 S400 和 S410 中，采用区域生长方式建立运动目标的颜色分布模型，由目标的初始像素点开始，将所有像素合并到两个大的区域中，得到两个颜色区域对象 RO ，并建立颜色区域特征向量，进入跟踪状态。

步骤 S420：选取距离最小的候选目标。对当前每一帧图像，与背景图像做差分后提取到 n 个候选的运动目标，计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$ 。假设前一帧中跟踪到的运动目标为 vop_{i-1}^G ，其颜色区域特征向量 $S_{vop_{i-1}^G}=(RO_1^G, RO_2^G, \dots, RO_N^G)$ ，中心坐标 (\bar{x}^G, \bar{y}^G) ，计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离： $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$ ， $k = 1 \dots n$ ，选取距离最小的候选目标 vop_i^D 。

在本实施例中，假定第 $i-1$ 帧跟踪到的目标表示为 vop_{i-1}^G ，并得到目标的颜色区域特征向量 $s_{vop_{i-1}^G}=(RO_1^G, RO_2^G)$ 及中心坐标 (\bar{x}^G, \bar{y}^G) 。对当前第 i 帧图像，与背景做差分，提取 n 个候选目标，并计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$ 。计算待跟踪目标与所有候选目标的中心坐标的距离 $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$ ， $k=1 \dots n$ ，选取距离最小的候选目标 vop_i^D 。

步骤 S430：根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$ ，遍历候选目标 vop_i^D 的所有像素，把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域，并

记录区域的像素数目 n_1 ，计算此区域的平均颜色值 avc_1^D ，编号为 1；把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_2 ，计算此区域的平均颜色值 avc_2^D ，编号为 2；依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D = (avc_i^D, r_i, i)$ ，并得到特征向量 $S_{vop}^D = (RO_1^D, RO_2^D, \dots, RO_N^D)$ 。

在本实施例中，根据 $i-1$ 帧跟踪到的目标 vop_{i-1}^G 颜色区域对象 (RO_{i-1}^G, RO_{i-1}^G) 的平均颜色 (avc_1^G, avc_2^G) ，遍历当前帧与目标 vop_{i-1}^G 距离最小的目标 vop_i^D 的所有像素，把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_1 ，计算此区域的平均颜色值 avc_1^D ，编号为 1；把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_2 ，计算此区域的平均颜色值 avc_2^D ，编号为 2。这样得到候选目标 vop_i^D 的两个颜色区域对象 $RO_1^D = (avc_1^D, r_1, 1)$ ， $RO_2^D = (avc_2^D, r_2, 2)$ ，并得到特征向量 $s_{vop}^D = (RO_1^D, RO_2^D)$ 。

步骤 S440：计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离：

$$d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$$

其中 i 表示跟踪到的目标 vop_{i-1}^G ， j 表示候选目标 vop_i^D ， m 为颜色区域编号， $s(avc_m(i), avc_m(j))$ 是颜色 $avc_m(i)$ 与颜色 $avc_m(j)$ 的相似度量， Th 为颜色相似度的阈值。在本实施例中 m 可取 1 和 2。

步骤 S450：计算联合距离的公式为： $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j) \times \dots \times d_N(i, j)$ ，在当前帧中，如果 $d_{joint}(vop_{i-1}^G, vop_i^D)$ 小于预设的距离门限值 η_d ，则两者匹配，此候选目标即为跟踪到的目标；否则认为不匹配，回到步骤 S420 中选取下一个与目标 vop_{i-1}^G 距离较小的候选目标，做相同的匹配，如此重复直到匹配到目标。

在本实施例中联合距离的计算为： $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j)$ 。如果重复到找完所有的候选目标仍然没有合适的匹配目标，则当前帧跟踪失败。

实际上，相邻两帧的 VOP 跟踪会有如下三种结果：

(a) 对前一帧跟踪得到的 VOP，若当前帧 VOP 候选集合里只有一个 VOP，直接匹配，如果成功即为跟踪到的目标；

(b) 对前一帧跟踪得到的 VOP, 在当前帧 VOP 候选集合里找到最匹配的 VOP, 跟踪成功;

(c) 若当前帧 VOP 候选集合没有匹配到合适的目标, 可能是由于目标被遮挡或者目标运动到镜头之外引起的, 并不能认为目标已经消失, 继续处理下几帧 (是预设的, 通常是 5 帧), 若连续几帧匹配不到合适的目标, 认为目标已经消失, 若能找到目标则又重新进入正常的目标跟踪阶段。

步骤 S460: 在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程: 对于颜色区域对象 $RO=(color, ratio, number)$ 的主分量 $color$, 使用自适应颜色更新模型: $c_i=(1-\alpha)\bar{c}+\alpha c_{i-1}$, c_i 为第 i 帧的区域颜色值, c_{i-1} 为第 $i-1$ 帧的区域颜色值, \bar{c} 为前 i 帧区域颜色均值, α 为权重因子, $ratio$ 分量, 用当前帧正确跟踪到的目标的颜色区域对象的相应的 $ratio$ 分量作为新的值。而对应的编号 $number$ 则保持不变。这样颜色区域对象在每一个新帧中不断得到更新。

对接下来的视频序列中的每一帧都重复上述的步骤 S400~S460, 成功实现了视频流中的运动目标的跟踪分析。

步骤 S50: 根据对多运动目标的跟踪结果统计分析人群的运动状态, 输出视频流中的人数、人流密度或人群聚集过程预警。

在步骤 S40 中得到了运动目标的准确跟踪结果, 根据跟踪结果能够方便得出视频流中的人数统计, 人数统计可以应用在银行金库的监视系统中, 对进出金库大门的人数进行检测, 并接口报警系统。根据跟踪结果也能够获知视频流中的人流密度, 可应用于预先提示公共安全管理部门对拥挤事态的防止、人群疏导以及突发事件的预案启动。同样地, 根据跟踪结果确定人群聚集面积, 再由人群聚集面积的大小判断是否进行人群聚集预警, 可应用于对政府大楼门前, 重要广场的人群聚集过程的预警, 这样可以预先防止非法聚众闹事、非法游行上访请愿等事态的出现。

为了更为准确的进行检测 (例如在运动目标的提取中, 需要根据目标的形状、面积等因素判断运动目标是否为人), 需要一个数据库供检测过程 (步骤 S20~S50) 使用。这个数据库存储了数学统计人群运动现象的经验数学模型和计算处理使用的特征矢量值, 这些特征矢量值包括摄像角度、焦距、空间坐标和人的像素面积等。如果没有这样的一个数据库, 则通过步骤 S60 来形成数据库: 即通过对视频流中的人群运动现象进行大量的数学统计以形成经验数学模型, 同时建立摄像角度、焦

距、空间坐标、人的像素面积等计算处理用的特征矢量值，将这些特征矢量值和经验数学模型用数学表格的形式存储在系统中以供检测过程使用。

图 5 示出了本发明的监控视频流中人流分析与人群聚集过程的检测系统的一个较佳实施例的框图结构。请参见图 5，检测系统主要包括：背景提取装置 10、差分图像计算装置 20、运动目标提取装置 30、运动目标跟踪与分析装置 40、人群运动状态统计分析装置 50 和人群运动特征数据库 60。

背景提取装置 10 进行背景图像的提取，建立背景图像的数学模型。图 6 示出了背景提取装置 10 的结构，请同时参见图 6。背景提取装置 10 进一步包括：初始化背景像素点模块 100 和滤波更新模块 110。初始化背景像素点模块 100 选择视频流的前 N 帧来构建初始化背景，设 $\lambda(x)$ 和 $\sigma(x)$ 分别为连续 N 帧图像中像素点上的中值和均方差，以 $\lambda(x)$ 为背景图像的像素值来构建背景，在初始化背景中，初始化背景像素点 $B(x) = \lambda(x)$ ，其中 N 为预设的自然数。滤波更新模块 110 利用一阶 Kalman 滤波模型更新背景图像： $B_{t+1}(x) = (1-\alpha)B_t(x) + \alpha F_t(x)$ ， $\sigma_{t+1}^2(x) = (1-\alpha)\sigma_t^2(x) + \alpha(F_t(x) - B_t(x))^2$ ，其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点上的均方差值， α 是更新因子，且 $0 < \alpha < 1$ 。

差分图像计算装置计算视频流中的每一帧图像与背景图像的差分图像。图 7 示出了差分图像计算装置 20 的结构，请同时参见图 7。差分图像计算装置 20 进一步包括：二值掩膜计算模块 200、滤波处理模块 210、断裂空洞处理模块 220 和噪声去除模块 230。二值掩膜计算模块 200 根据当前帧图像上某点的像素值，计算出差分图像值，由下式确定差分图像的二值掩膜：

$$M(x) = \begin{cases} 1 & \text{abs}(F_t(x) - B_t(x)) > Th \\ 0 & \text{其它} \end{cases}$$

其中 $Th = 2.5\sigma_t(x)$ ，其中 $B_t(x)$ 为 t 时刻 x 点上的背景像素值， $F_t(x)$ 为 t 时刻 x 点上的当前像素值， $\sigma_t(x)$ 为 t 时刻背景图像上 x 点的均方差值。滤波处理模块 210 对得到的二值掩膜图像进行滤波处理。断裂空洞处理模块 220 应用形态学算子对二值掩膜图像中的断裂或空洞进行处理，同时使运动目标区域的边界变得更加光滑。噪声去除模块 230 利用构成独立运动部件的像素是相互连接的而独立的像素是相互孤立的小串这一特点，采用基于前景区域的面积阈值法去

除噪声：若相互连接部件的像素总数超过设定阈值，则该部件为独立运动部件，剩下的像素为随机噪声。

运动目标提取装置 30 基于得到的二值模板，清除当前帧内不在模板内的像素点，得到运动目标的分割结果。如果遇到多运动目标的场景，通过内设的分割模块将多个运动目标分割成多个单运动目标并加以提取。

运动目标跟踪与分析装置 40 对多运动目标进行跟踪与分析。总的原理是：该装置建立运动目标的颜色分布模型，为每个运动目标建立至少一个颜色区域对象，以多运动颜色块及其组合结构为描述搜索运动目标的主要特征，并在跟踪目标的过程中对检测到的目标进行特征匹配以判断是否为需要跟踪的目标。

图 8 示出了运动目标跟踪与分析装置 40 的结构，请同时参见图 8。运动目标跟踪与分析装置 40 进一步包括：颜色分布模型建立模块 400、中心坐标计算模块 410、候选目标选取模块 420、特征向量形成模块 430、颜色区域对象距离计算模块 440、匹配判断模块 450 和颜色分布模型更新模块 460。

颜色分布模型建立模块 400 建立待跟踪运动目标 VOP 的颜色区域特征向量 $S_{vop} = (RO_1, RO_2, \dots, RO_N)$ ，其中 $RO_i = (avc_i, r_i, i)$ ， N 为自然数， i 为 1 至 N 的任一自然数， RO_i 为颜色区域对象， avc_i 为区域对象的代表色， r_i 为区域面积比率即区域像素数目占运动目标总像素数目的比率， $r_i = n_i / A_{vop}$ ， n_i 为区域包含的像素数目， A_{vop} 为运动目标的总像素数目， i 为区域编号。

在颜色分布模型建立模块 400 中，在跟踪前通过初始帧与背景差分获得初始的运动目标，通过区域生长法来获得待跟踪目标的颜色分布模型。当然，对多运动目标跟踪的情形，可以在初始状态为每个目标建立颜色分布模型。

定义颜色之间的距离阈值 $T1$ ，若空间位置相邻的两像素（使用 8 连通方式），颜色距离低于阈值 $T1$ ，则此两个像素可以合并成一个区域，此区域一直增长到没有相邻像素可以合并为止。为避免同一区域里存在着颜色差异太大的像素，我们另外设定一个阈值 $T2$ ，在区域生长的过程中，若某像素的颜色与此区域的所有像素的距离大于阈值 $T2$ ，则不合并此像素。若目标还有未被合并的像素，则继续新的区域生长，直到所有的像素属于某个区域为止。对于过小的区域，我们采用的作法是，将之与周围的区域比较，并归入到颜色距离最小的区域中去。计算每个区域的平均颜色，以及区域包含的像素数目占目标像素总数的比

率，这样得到最终的几个颜色区域对象 RO 。

$$\begin{aligned} \bar{x} &= \frac{1}{A_{vop}} \sum_{x \in VOP} x, \\ \bar{y} &= \frac{1}{A_{vop}} \sum_{y \in VOP} y. \end{aligned}$$

候选目标选取模块 420 对当前帧图像，与背景图像做差分后提取到 n 个候选的运动目标，通过中心坐标计算模块 410 计算每个候选目标 k 的中心坐标 $(\bar{x}_i^k, \bar{y}_i^k)$ 。假设前一帧中跟踪到的运动目标为 vop_{i-1}^G ，具有通过该颜色分布模型建立模块得到的颜色区域特征向量 $S_{vop_{i-1}} = (RO_1^G, RO_2^G, \dots, RO_N^G)$ ，已经通过中心坐标计算模块 410 得到的中心坐标 (\bar{x}^G, \bar{y}^G) ，计算跟踪到的运动目标 vop_{i-1}^G 与所有候选的待跟踪运动目标的中心坐标的距离： $d = \sqrt{(\bar{x}_i^k - \bar{x}^G)^2 + (\bar{y}_i^k - \bar{y}^G)^2}$ ， $k = 1 \dots n$ ，选取距离最小的候选目标 vop_i^D 。

特征向量形成模块 430 根据前一帧中跟踪到的目标 vop_{i-1}^G 的颜色区域对象 $(RO_1^G, RO_2^G, \dots, RO_N^G)$ 的平均颜色 $(avc_1^G, avc_2^G, \dots, avc_N^G)$ ，遍历候选目标 vop_i^D 的所有像素，把所有和颜色 avc_1^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_1 ，计算此区域的平均颜色值 avc_1^D ，编号为 1；把所有和颜色 avc_2^G 相似的并在空间上相邻的点合并成一个大的区域，并记录区域的像素数目 n_2 ，计算此区域的平均颜色值 avc_2^D ，编号为 2；依次类推得到候选目标 vop_i^D 的 N 个颜色区域对象 $RO_i^D = (avc_i^D, r_i, i)$ ，并得到特征向量 $S_{vop_i^D} = (RO_1^D, RO_2^D, \dots, RO_N^D)$ 。

颜色区域对象距离计算模块 440 计算跟踪到的目标 vop_{i-1}^G 和候选目标 vop_i^D 的颜色区域对象距离： $d_m(i, j) = \begin{cases} (r_m(i) - r_m(j))^2 & s(avc_m(i), avc_m(j)) > Th \\ +\infty & \text{其它} \end{cases}$ ，其中 i 表示跟踪到的目标 vop_{i-1}^G ， j 表示候选目标 vop_i^D ， m 为颜色区域编号， $s(avc_m(i), avc_m(j))$ 是颜色 $avc_m(i)$ 与颜色 $avc_m(j)$ 的相似度量， Th 为颜色相似度的阈值。

匹配判断模块 450 计算联合距离的公式为： $d_{joint}(i, j) = d_1(i, j) \times d_2(i, j) \times \dots \times d_N(i, j)$ ，在当前帧中，如果 $d_{joint}(vop_{i-1}^G, vop_i^D)$ 小于预设的距离门限值 η_d ，则两者匹配，此候选目标即为跟踪到的目标；否则认为不匹配，再返回到候选

目标选取模块 420 选取下一个与目标 vop_{i-1}^G 距离较小的候选目标,做相同的匹配,如此重复直到匹配到目标。

实际上,相邻两帧的 VOP 跟踪会有如下三种结果:

(a) 对前一帧跟踪得到的 VOP,若当前帧 VOP 候选集合里只有一个 VOP,直接匹配,如果成功即为跟踪到的目标;

(b) 对前一帧跟踪得到的 VOP,在当前帧 VOP 候选集合里找到最匹配的 VOP,跟踪成功;

(c) 若当前帧 VOP 候选集合没有匹配到合适的目标,可能是由于目标被遮挡或者目标运动到镜头之外引起的,并不能认为目标已经消失,继续处理下几帧(帧数是预设的,通常是 5 帧),若连续几帧匹配不到合适的目标,认为目标已经消失,若能找到目标则又重新进入正常的目标跟踪阶段。

颜色分布模型更新模块 460 在当前帧中找到被跟踪的目标后进入目标颜色分布模型的更新过程:对于颜色区域对象 $RO=(color, ratio, number)$ 的主分量 $color$,使用自适应颜色更新模型: $c_i=(1-\alpha)\bar{c}+\alpha c_{i-1}$, c_i 为第 i 帧的区域颜色值, c_{i-1} 为第 $i-1$ 帧的区域颜色值, \bar{c} 为前 i 帧区域颜色均值, α 为权重因子, $ratio$ 分量,用当前帧正确跟踪到的目标的颜色区域对象的相应的 $ratio$ 分量作为新的值。而对应的编号 $number$ 则保持不变。

人群运动状态统计分析装置 50 根据运动目标跟踪与分析装置 40 所得到的运动目标的准确跟踪结果,基于跟踪结果能够方便得出视频流中的人数统计,人数统计可以应用在银行金库的监视系统中,对进出金库大门的人数进行检测,并接口报警系统。根据跟踪结果也能够获知视频流中的人流密度,可应用于预先提示公共安全管理部门对拥挤事态的防止、人群疏导以及突发事件的预案启动。同样地,根据跟踪结果确定人群聚集面积,再由人群聚集面积的大小判断是否进行人群聚集预警,可应用于对政府大楼门前,重要广场的人群聚集过程的预警,这样可以预先防止非法聚众闹事、非法游行上访请愿等事态的出现。

为了更为准确的进行检测(例如在运动目标的提取中,需要根据目标的形状、面积等因素判断运动目标是否为人),需要一个数据库供检测系统中的其他装置使用。人群运动特征数据库 60 连接差分图像计算装置 20、运动目标提取装置 30、运动目标跟踪与分析装置 40 和人群运动状态统计分析装置 50 以供这些装置检测之

用。该数据库 60 以数学表格的形式存储了数学统计人群运动现象的经验数学模型和计算处理使用的特征矢量值，这些特征矢量值包括摄像角度、焦距、空间坐标和人的像素面积等。

应理解，本发明的发明点在于：（1）采用 HSV 颜色空间算法，应用色彩量化统计分析技术，将颜色距离矢量化；（2）采用背景减除技术得到对象活动区域，再利用数学形态学运算处理，最后利用面积阈值法和区域生长法获得目标所在的最小外接矩形区域，同时采用自适应背景更新策略来逐渐更新用到的背景模型。这种算法可以得到运动目标较为精确的特征，并且算法简捷、快速、有效。（3）在检测与提取出视频对象的基础上，建立目标的颜色空间分布模型，提出颜色区域对象的概念，为每个目标建立一个或几个颜色区域对象，以多运动颜色块及其组合结构为描述搜索目标的主要特征，并在跟踪目标的过程中，对检测到的目标进行特征匹配，以此判断是否需要跟踪的目标。（4）采用基于颜色特征匹配的视频流检索算法来实现目标的更加可靠有效的跟踪。

上述实施例是提供给本领域普通技术人员来实现或使用本发明的，本领域普通技术人员可在不脱离本发明的发明思想的情况下，对上述实施例做出种种修改或变化，因而本发明的保护范围并不被上述实施例所限，而应该是符合权利要求书提到的创新性特征的最大范围。

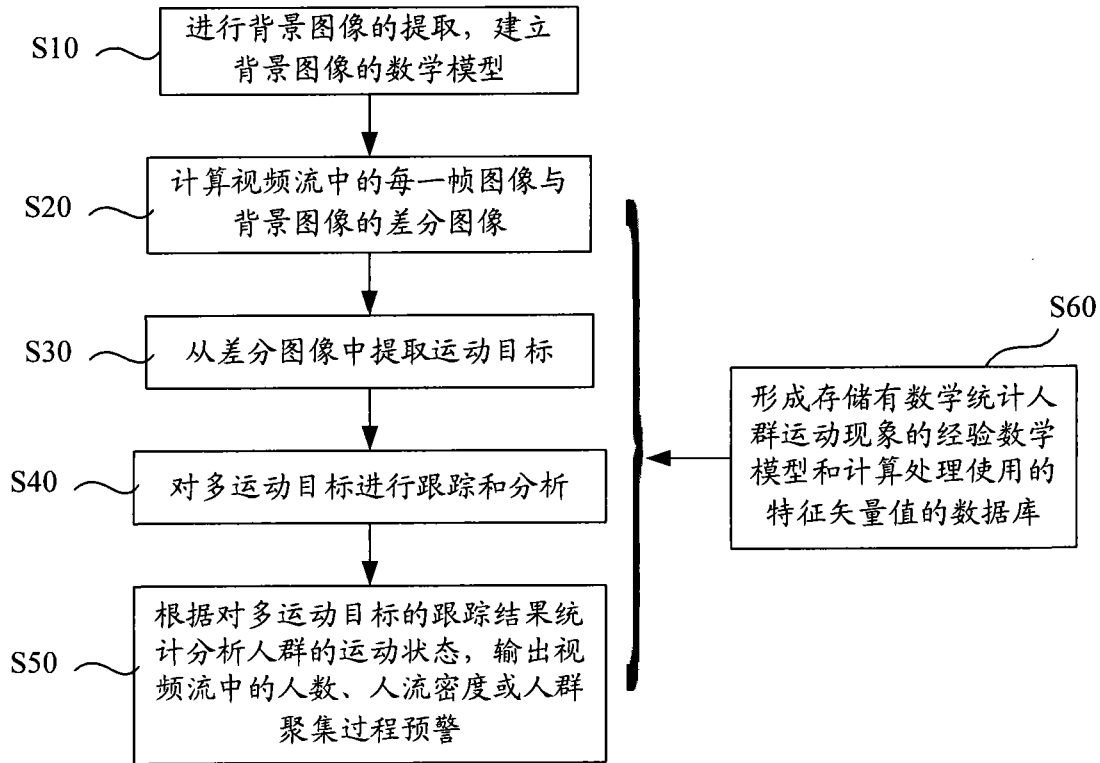


图 1

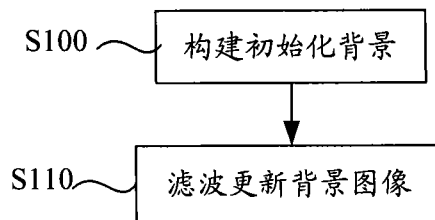


图 2

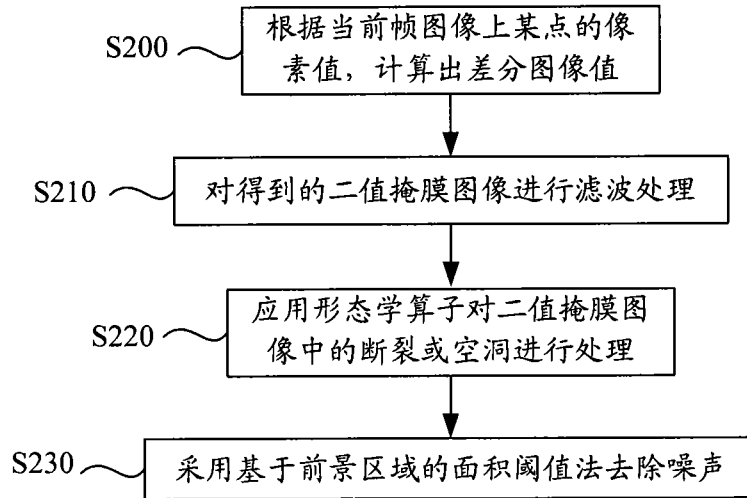


图 3

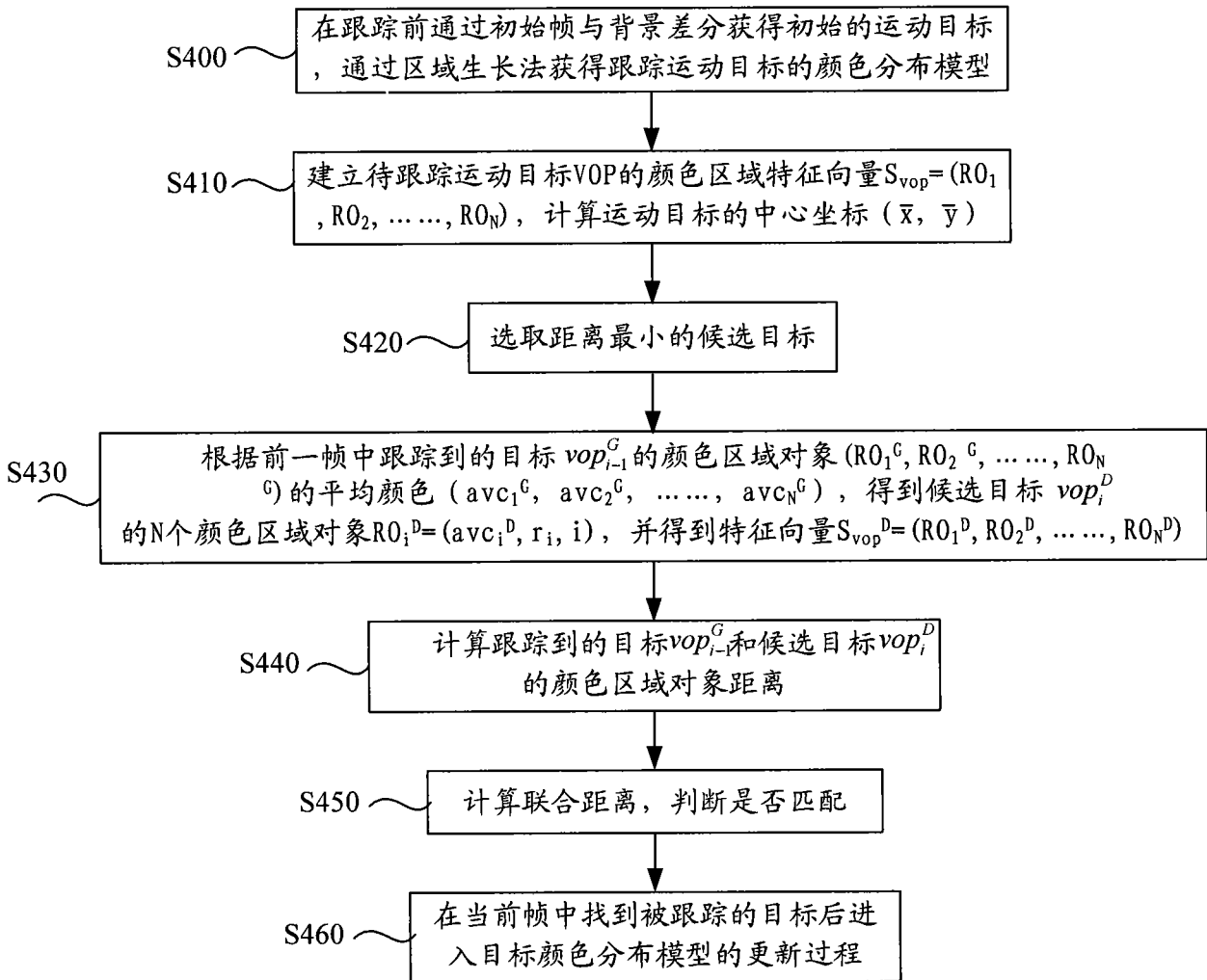


图 4

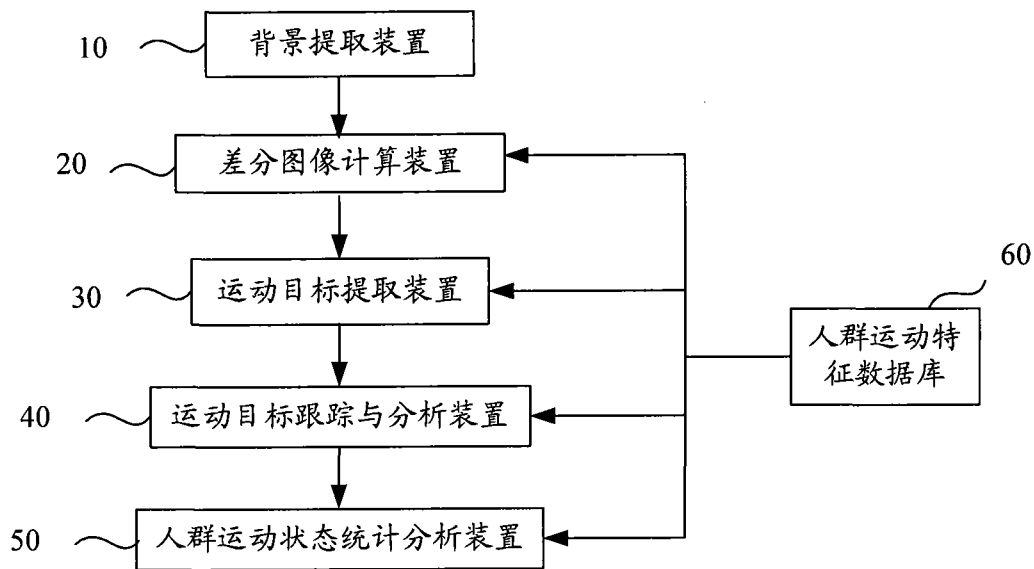


图 5

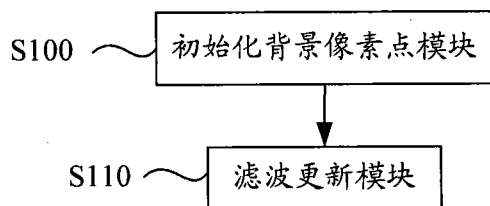


图 6

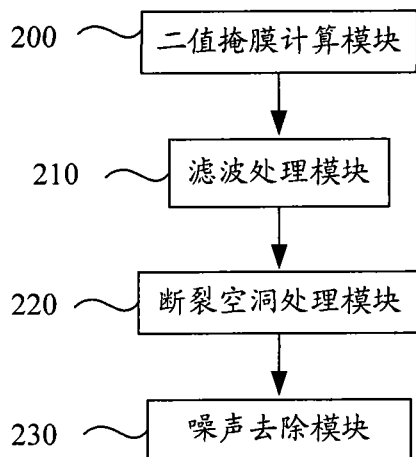


图 7

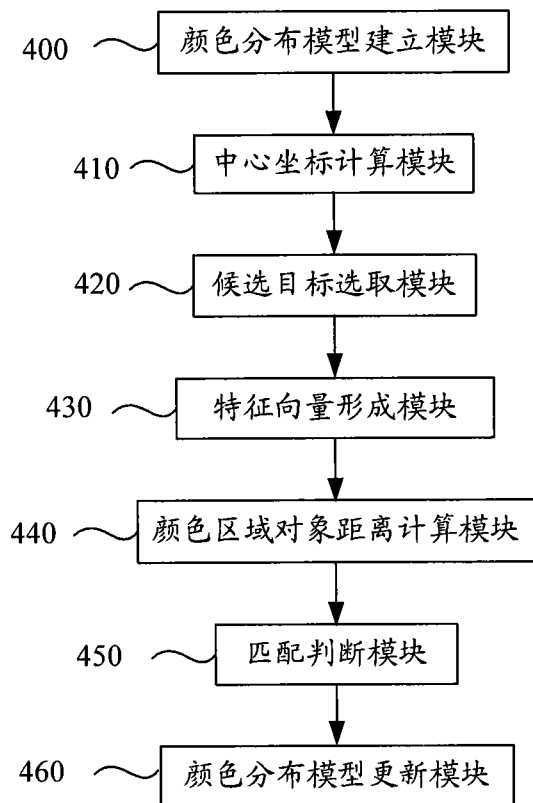


图 8