



(12) 发明专利申请

(10) 申请公布号 CN 101950116 A

(43) 申请公布日 2011.01.19

(21) 申请号 201010280825.5

(22) 申请日 2010.09.14

(71) 申请人 浙江工业大学

地址 310014 浙江省杭州市下城区朝晖六区

(72) 发明人 郑雅羽 刘连杰 杨雷刚 俞立

(74) 专利代理机构 杭州天正专利事务所有限公司 33201

代理人 王兵 王利强

(51) Int. Cl.

G03B 13/36(2006.01)

G02B 7/36(2006.01)

H04N 5/232(2006.01)

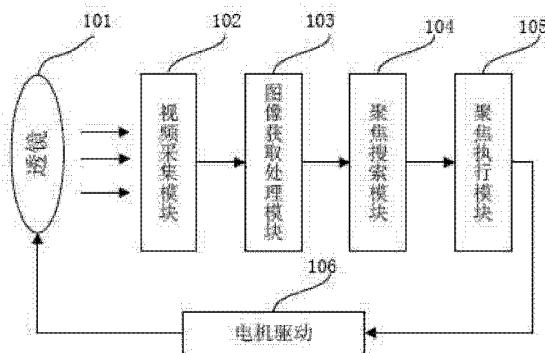
权利要求书 2 页 说明书 6 页 附图 3 页

(54) 发明名称

一种应用于多主体场景的视频自动聚焦方法

(57) 摘要

一种应用于多主体场景的视频自动聚焦方法，包括以下步骤：1) 将镜头所对应的场景分划分为左、中、右三个子窗口；2) 分别对三个子窗口中的图像进行清晰度评价值计算；3) 按照预设的步长将聚焦镜头向近端方向移动；4) 通过连续两帧图像的清晰度值变化情况，选择是否将三个子窗口归入聚焦窗口中；5) 判定聚焦窗口中的子窗口的个数，当子窗口个数不为零时，则说明未搜索到最近主体则重复所述步骤 2) 至 5)，否则进入步骤 6)；6) 控制聚焦镜头向远端方向移动，回溯到清晰度曲线的第一个波峰值位置，将该位置作为镜头的聚焦位置。本发明控制摄像镜头能够在多主体情况下实现准确快速聚焦到最近主体。



1. 一种应用于多主体场景的视频自动聚焦方法,其特征在于:所述的自动聚焦方法包括以下步骤:

- 1) 将镜头所对应的场景分划分为左、中、右三个子窗口;
- 2) 分别对三个子窗口中的图像进行清晰度评价值计算。

3) 按照预设的步长将聚焦镜头向近端方向移动,其中近端方向是指聚焦平面靠近聚焦镜头的方向,其对应的反方向定义为远端方向;

4) 在三个子窗口中分别获得连续两帧图像的清晰度值变化情况,当子窗口中图像的清晰度值增大时,则将该子窗口纳入聚焦窗口中,否则将该子窗口从聚焦窗口中去除;

5) 判定聚焦窗口中的子窗口的个数,当子窗口个数不为零时,则说明未搜索到最近主体则重复所述步骤2)至5),否则进入步骤6);

6) 控制聚焦镜头向远端方向移动,回溯到清晰度曲线的第一个波峰值位置,将该位置作为镜头的聚焦位置。

2. 如权利要求1所述的应用于多主体场景的视频自动聚焦方法,其特征在于:所述步骤2)中采用基于梯度阈值的清晰度计算方法,包括以下步骤:

(201) 对获取到的原始数字图像,通过基于边缘检测的图像 Sobel 算子进行图像的梯度值的计算;

(202) 利用获得图像的梯度值进行直方图计算,得到图像的梯度直方图;

(203) 对图像梯度直方图进行权值修正,得到加权梯度直方图;

(204) 通过修正后的加权梯度直方图,计算获得梯度阈值;

(205) 对图像进行清晰度值计算,当图像梯度大于梯度阈值时才将其纳入梯度值的累加,将最终的梯度累加值作为该图像的清晰度值。

3. 如权利要求1或2所述的应用于多主体场景的视频自动聚焦方法,其特征在于:所述视频自动聚焦方法还包括以下步骤:

7) 聚焦完成后则进入观察模式,判定是否需要进入重聚焦,包括以下步骤:

(701) 初始化各子窗口的清晰度参考值,在聚焦模式结束时,将每个窗口的清晰度值作为各个窗口的清晰度参考值;

(702) 计算每个子窗口的平均清晰度值,平均清晰度值为在观察模式下的清晰度值在时间上的平均;

(703) 判断每个子窗口的平均清晰度值与其清晰度参考值的差值是否超过预先设定的阈值,若超过则进入步骤2),否则仍然为观察模式,并重复步骤(702)至(703)。

4. 如权利要求1或2所述的应用于多主体场景的视频自动聚焦方法,其特征在于:所述步骤1)中,所述子窗口的大小为 $3/4H \times 1/4W$,其中,H和W分别表示获得的视频场景的高度和宽度。

5. 如权利要求3所述的应用于多主体场景的视频自动聚焦方法,其特征在于:所述步骤1)中,所述子窗口的大小为 $3/4H \times 1/4W$,其中,H和W分别表示获得的视频场景的高度和宽度。

6. 如权利要求1或2所述的应用于多主体场景的视频自动聚焦方法,其特征在于:所述步骤5)中,当聚焦窗口中子窗口个数最多为3个时,判定镜头位置在聚焦位置的远端,控制聚焦镜头向近端移动。

7. 如权利要求 3 所述的应用于多主体场景的视频自动聚焦方法，其特征在于：所述步骤 5) 中，当聚焦窗口中子窗口个数最多为 3 个时，判定镜头位置在聚焦位置的远端，控制聚焦镜头向近端移动。

一种应用于多主体场景的视频自动聚焦方法

技术领域

[0001] 本发明涉及视频自动聚焦技术领域，尤其是应用于多主体场景的视频自动聚焦方法。

背景技术

[0002] 随着数字图像技术的发展，自动聚焦技术的应用范围也不断扩大，现已在照相机、摄像机、显微镜等仪器中得到了广泛应用。总体来讲自动聚焦主要有主动聚焦和被动聚焦两种实现方式：主动聚焦通过激光、红外线或者超声波对目标进行测距，根据所测得距离来调节镜头的聚焦位置。被动聚焦又分为基于装置的聚焦和基于图像的聚焦。其中基于图像的自动聚焦不需要额外的辅助装置所以被广泛的用于数字摄像机中。

[0003] 在目前应用中，基于图像的聚焦方式对环境的要求很高。比如场景中存在多个主体时，这时候得到的清晰度曲线会成双峰性或者是多峰性。例如图 2 中所示，在场景中存在远近不同的两个主体以及更远端的背景。当聚焦窗口包含这三个部分内容时，其清晰度曲线就会成双峰性或者多峰性，图 2 中为了讨论方便忽略背景部分，因为当背景为白墙等低清晰度物体时通常可以被忽略。如图中的清晰度曲线所示，两个峰值点分别代表了近端物体和远端物体的聚焦位置。此时如果用常规的聚焦方法比如公开号为 CN101509998（申请号为 200910097044. X）的中国专利《一种自动聚焦方法及应用该方法的显微镜》和公开号为 CN101494737（申请号位 200910096472. 6）的中国专利《一种一体化摄像机装置及自适应自动聚焦方法》都会因为陷入局部极值，从而不能聚焦到最近的主体。而通常人们使用摄像设备的时候，都会将视频场景中的最近主体作为感兴趣的部分，因此在多主体场景下的自动聚焦目的是为了能够准确、快速的聚焦到最近主体。

发明内容

[0004] 为了克服现有的多主体场景下的自动聚焦方法的不能准确快速的聚焦到最近主体的不足，本发明提供一种摄像镜头能够在多主体情况下实现准确快速聚焦到最近主体的应用于多主体场景的视频自动聚焦方法。

[0005] 本发明解决其技术问题所采用的技术方案是：

一种应用于多主体场景的视频自动聚焦方法，所述的自动聚焦方法包括以下步骤：

- 1) 将镜头所对应的场景分划分为左、中、右三个子窗口；
- 2) 分别对三个子窗口中的图像进行清晰度评价值计算；
- 3) 按照预设的步长将聚焦镜头向近端方向移动，其中近端方向是指聚焦平面靠近聚焦镜头的方向，其对应的反方向定义为远端方向；
- 4) 在三个子窗口中分别获得连续两帧图像的清晰度值变化情况，当子窗口中图像的清晰度值增大时，则将该子窗口纳入聚焦窗口中，否则将该子窗口从聚焦窗口中去除；
- 5) 判定聚焦窗口中的子窗口的个数，当子窗口个数不为零时，则说明未搜索到最近主体则重复所述步骤 2) 至 5)，否则进入步骤 6)；

6) 控制聚焦镜头向远端方向移动,回溯到清晰度曲线的第一个波峰值位置,将该位置作为镜头的聚焦位置。

[0006] 作为优选的一种方案:所述步骤 2) 中采用基于梯度阈值的清晰度计算方法,包括以下步骤:

(201) 对获取到的原始数字图像,通过基于边缘检测的图像 Sobel 算子进行图像的梯度值的计算;

(202) 利用获得图像的梯度值进行直方图计算,得到图像的梯度直方图;

(203) 对图像梯度直方图进行权值修正,得到加权梯度直方图;

(204) 通过修正后的加权梯度直方图,计算获得梯度阈值;

(205) 对图像进行清晰度值计算,当图像梯度大于梯度阈值时才将其纳入梯度值的累加,将最终的梯度累加值作为该图像的清晰度值。

[0007] 进一步,所述视频自动聚焦方法还包括以下步骤:

7) 聚焦完成后则进入观察模式,判定是否需要进入重聚焦,包括以下步骤:

(701) 初始化各子窗口的清晰度参考值,在聚焦模式结束时,将每个窗口的清晰度值作为各个窗口的清晰度参考值;

(702) 计算每个子窗口的平均清晰度值,平均清晰度值为在观察模式下的清晰度值在时间上的平均;

(703) 判断每个子窗口的平均清晰度值与其清晰度参考值的差值是否超过预先设定的阈值,若超过则进入步骤 2),否则仍然为观察模式,并重复步骤(702)至(703)。

[0008] 再进一步,所述步骤 1) 中,所述子窗口的大小为 $3/4H \times 1/4W$,其中, H 和 W 分别表示获得的视频场景的高度和宽度。

[0009] 更进一步,所述步骤 5)中,当聚焦窗口中子窗口个数最多为 3 个时,判定镜头位置在聚焦位置的远端,控制聚焦镜头向近端移动。

[0010] 本发明的技术构思为:本发明第一个要解决的技术问题是设计出一种精度高、速度快的清晰度计算方法以表征图像的清晰程度。本发明的第二个要解决的技术问题是设计一种多主体聚焦策略,该聚焦方法能够克服传统聚焦方法在多主体场景下陷入局部极值的问题。本发明第三个要解决的技术问题是要设计一种重聚焦机制,该机制能够决定何时进行重聚焦。

[0011] 本发明针对三可变镜头的摄像设备。所谓三可变镜头是指焦距、聚焦、光圈三个参数均可以通过电动马达驱动变化的镜头。首先通过这些摄像设备实时采集原始视频信号,然后从这些原始视频信号中经过图像获取以及预处理得到符合标准格式的数字图像。接着由视频自动聚焦方法改变聚焦镜头的位置,实现对观察场景中的最近主体得自动聚焦。其中视频自动聚焦方法分为聚焦模式和观察模式。

[0012] 在上述步骤 (2) 中的清晰度计算中采用了一种基于梯度阈值的清晰度计算方法。

[0013] 在上述步骤 (4) 中聚焦区域选择机制为:在整个聚焦过程中,聚焦窗口依据各子窗口中的清晰度值变化而动态地进行变化;当子窗口中图像的清晰度值增大时,则将该子窗口纳入聚焦窗口中,否则将该子窗口从聚焦窗口中去除。聚焦窗口中子窗口个数最多为 3 个,说明此时镜头位置在聚焦位置的远端;当聚焦窗口中的子窗口个数最少为 0 个,说明此时镜头位置在聚焦位置的近端。

[0014] 本发明有益的效果是：能够在多主体场景下，根据图像的方法，快速、准确地搜索到场景中的最近主体，克服了传统聚焦方法在多主体场景下陷入局部极值的问题。

附图说明

- [0015] 图 1 是多主体自动聚焦方法的系统模块框图。
- [0016] 图 2 是多主体场景下的清晰度曲线示意图。
- [0017] 图 3 是多主体自动聚焦方法的聚焦主程序流程图。
- [0018] 图 4 是多主体自动聚焦方法的聚焦模式流程图。
- [0019] 图 5 是多主体自动聚焦方法的聚焦窗口的划分示意图。
- [0020] 图 6 是多主体自动聚焦方法的清晰度计算流程图。
- [0021] 图 7 是多主体自动聚焦方法的观察模式流程图。

具体实施方式

[0022] 下面结合附图对本发明作进一步描述。
[0023] 参照图 1~图 7，一种应用于多主体场景的视频自动聚焦方法，所述的自动聚焦方法包括以下步骤：

- 1) 将镜头所对应的场景分划分为左、中、右三个子窗口；
- 2) 分别对三个子窗口中的图像进行清晰度评价值计算；
- 3) 按照预设的步长将聚焦镜头向近端方向移动，其中近端方向是指聚焦平面靠近聚焦镜头的方向，其对应的反方向定义为远端方向；
- 4) 在三个子窗口中分别获得连续两帧图像的清晰度值变化情况，当子窗口中图像的清晰度值增大时，则将该子窗口纳入聚焦窗口中，否则将该子窗口从聚焦窗口中去除；
- 5) 判定聚焦窗口中的子窗口的个数，当子窗口个数不为零时，则说明未搜索到最近主体则重复所述步骤 2) 至 5)，否则进入步骤 6)；
- 6) 控制聚焦镜头向远端方向移动，回溯到清晰度曲线的第一个波峰值位置，将该位置作为镜头的聚焦位置。

[0024] 所述步骤 2) 中采用基于梯度阈值的清晰度计算方法，包括以下步骤：

(201) 对获取到的原始数字图像，通过基于边缘检测的图像 Sobel 算子进行图像的梯度值的计算；

- (202) 利用获得图像的梯度值进行直方图计算，得到图像的梯度直方图；
- (203) 对图像梯度直方图进行权值修正，得到加权梯度直方图；
- (204) 通过修正后的加权梯度直方图，计算获得梯度阈值；

(205) 对图像进行清晰度值计算，当图像梯度大于梯度阈值时才将其纳入梯度值的累加，将最终的梯度累加值作为该图像的清晰度值。

[0025] 所述视频自动聚焦方法还包括以下步骤：

- 7) 聚焦完成后则进入观察模式，判定是否需要进入重聚焦，包括以下步骤：

(701) 初始化各子窗口的清晰度参考值，在聚焦模式结束时，将每个窗口的清晰度值作为各个窗口的清晰度参考值；

(702) 计算每个子窗口的平均清晰度值，平均清晰度值为在观察模式下的清晰度值在

时间上的平均；

(703) 判断每个子窗口的平均清晰度值与其清晰度参考值的差值是否超过预先设定的阈值，若超过则进入步骤 2)，否则仍然为观察模式，并重复步骤(702)至(703)。

[0026] 所述步骤 1)中，所述子窗口的大小为 $3/4H \times 1/4W$ ，其中，H 和 W 分别表示获得的视频场景的高度和宽度。

[0027] 所述步骤 5)中，当聚焦窗口中子窗口个数最多为 3 个时，判定镜头位置在聚焦位置的远端，控制聚焦镜头向近端移动。

[0028] 本实施例的视频自动聚焦方法可以用于显微镜、照相机、摄像机等需要调焦的各种产品中，其系统模块框图，如图 1 所示：

所述的针对多主体自动聚焦的系统模块包括透镜部分 101、视频采集模块 102、图像处理模块 103、聚焦搜索模块 104 和聚焦执行模块 105。首先光线经过透镜 101 上，该透镜为三可变透镜。即该镜头的焦距、聚焦、光圈三个参数均可以通过电动马达驱动变化。然后视频采集模块 102 将采集到的原始视频图像数据进行格式化处理，输出符合 BT. 656 标准的视频信号。然后图像获取处理模块 103 对视频采集模块输出标准的视频信号，进行数字图像的获取，并进行图像的清晰度值评价。聚焦搜索模块 104 根据清晰度评价值，判定聚焦镜头是否已经在聚焦位置，若没有则计算下一步聚焦镜头所需移动的方向和步长，否则保持聚焦镜头位置不变。聚焦执行模块 105 根据聚焦搜索模块 104 计算得到聚焦镜头移动方向和步长，给驱动镜头的步进电机发送控制信号调整镜头位置，从而实现对镜头的调焦移动。

[0029] 参照图 3 所示，为本发明的自动聚焦主程序流程图，本文将按照此流程图对于自动聚焦算法实例予以介绍说明。当启动自动聚焦时，首先判定是否是聚焦状态，是则进入聚焦模式否则进入观察模式。其中聚焦状态的判断采用两次向不同方向移动镜头的方法，首先读取此时的镜头位置和清晰度值 S_t ，然后将聚焦镜头向近端方向移动预设步长得到清晰度值 S_n ，在朝远端移动预设步长得到 S_f 。如果 $S_t > S_n$ 并且 $S_t > S_f$ ，则说明此时的镜头位置为聚焦位置，除此之外的情况说明为离焦位置。

[0030] 在自动聚焦算法中，分为聚焦模式和观察模式，且两个模式能够自动切换。在聚焦模式时，算法自动搜索到场景中的最近主体位置，然后进入观察模式；在观察模式时，算法监测每个子窗口清晰度值是否超过阈值，通过重聚焦机制来决定是否进行重聚焦。

[0031] 程序进入聚焦模式如图 4 所示。首先对聚焦区域进行子窗口划分 (S10)，这主要解决多主体聚焦时，搜索算法陷入局部最值以至于不能搜索到最近主体。子窗口的划分如图 5 所示，整个图像区域将被划分成左、中、右三个子窗口。每个子窗口的大小为 $1/4W \times 3/4H$ ，其中 W 和 H 分别代表原始图像的宽带和高度。

[0032] 分别计算每个子窗口的清晰度值 (S20)，清晰度是为了评价图像的清晰程度。因为针对于多主体聚焦，窗口中包含着不同的主体，因此需要对三个子窗口分别进行计算。为了加快计算速度和精度，本发明提出了一种基于梯度阈值的清晰度计算方法。该方法通过加权梯度直方图进行阈值计算，并通过该阈值去除图像中的背景噪声计算的详细步骤如图 6 所示：

首先对原始图像信息用基于边缘的图像算子 Sobel 进行边缘检测得到图像的梯度值 (S201)。边缘检测的算子有很多，因为 Sobel 在精度上优于其它算子，因此在本发明使用 Sobel 算子进行梯度计算。根据式(1) 的图像梯度计算方法：

$$\nabla f(x, y) = \sqrt{G_x^2 + G_y^2} \quad (1)$$

其中, $f(x, y)$ 为数字图像, $\nabla f(x, y)$ 表示对该数字图像在 (x, y) 处的梯度值, G_x 和 G_y 分别为其偏导数。其中偏导数的按照式(2)来计算:

$$G_x = f_{\text{sub}}(x, y) * \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad G_y = f_{\text{sub}}(x, y) * \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (2)$$

这里, $f_{\text{sub}}(x, y)$ 表示图像 f 中以 (x, y) 为中心的 8 领域子图, $*$ 表示卷积运算;为了加快梯度值的计算速度, 提高聚焦速度。这里一副图像的梯度值可以有式(3)进行计算得到近似值。

[0033] $\nabla f(x, y) = |G_x| + |G_y| \quad (3)$

计算得到图像的梯度后, 就可以得到图像梯度直方图(S202)。然后对图像梯度直方图进行修正, 得到加权后的梯度直方图(S203)。修正的过程如下: 对梯度值大于 T_{th} 的值乘 N 修正(其中 $N > 1$); 对梯度值小于 T_{th} 的值乘 M 修正(其中 $0 < M < 1$)。然后从加权梯度直方图中获取梯度阈值 T_N , 获取的方法是取加权梯度直方图中两峰值之间的波谷值作为阈值。阈值选定后, 最后对图像的梯度值进行累加, 作为其清晰度评价值。计算公式如式(4)所示。

[0034] $F(k) = \sum_x \sum_y \nabla f(x, y) \quad \nabla f(x, y) > T_N \quad (4)$

其中, $F(k)$ 表示数字图像 $f(x, y)$ 的清晰度值, T_N 为梯度阈值。

[0035] 此时按照预设的步长将镜头向近端方向移动(S30)。算法的目的是为了在多主体场景中聚焦到最近的主体。因此, 初始方向是向近端移动。

[0036] 计算每个子窗口的清晰度值变化情况(S40)。在进行清晰度计算的时候, 每个窗口都有各自的清晰度值分别记作。将摄像镜头从远端到近端移动, 计算移动后的每个子窗口的清晰度变化值为 DS。如果 $DS > 0$ 则说明此窗口包含着更近端的主体, 否则说明这个窗口中包含着远端主体。聚焦窗口选择的目的是为了让窗口中包含最近主体, 所以将包含着近端主体的子窗口组合起来作为聚焦区域。当三个子窗口的清晰度变化值 DS 都变大 0 的时候, 则聚焦窗口为三个子窗口的和; 当某个子窗口清晰度变化值 DS 小于 0 时, 则将其从聚焦窗口中去除。

[0037] 判断聚焦区域中的子窗口的个数(S50)。当子窗口的个数为零时, 说明已经没有更近的主体了, 此时应该在反向搜索到最近的清晰度曲线的波峰值(S60)。否则重复 S20 至 S50。

[0038] 回溯到最近峰值后, 则进入观察模式。此时离焦判定机制启动, 用于判定是否需要进入聚焦模式。其具体步骤如图 7 所示。

[0039] 首先计算每个子窗口的清晰度参考值(S701)。这里用 f_x^{focus} 来表示各个子窗口聚焦时的清晰度值, 其中 $x = l, m, r$ 。 l 代表为左窗口, m 代表为中窗口, r 代表为右窗口。

[0040] 然后计算每个子窗口的平均清晰度值(S702), 计算公式如式(5)。

$$[0041] \quad \bar{f}_x^t = \frac{\sum f_x(t)}{t} \quad x = l, m, r \quad (5)$$

这里 $f_x(t)$ 代表 t 时刻的每个子窗口的清晰度水平, \bar{f}_x^t 表示 t 时段内的平均清晰度值。

[0042] 判断所有子窗口的清晰度变化值是否都大于阈值 (S703), 如果是则进入重聚焦, 否则重复步骤 S702 至 S703。判定条件按照公式(6)所示。

$$[0043] \quad |\bar{f}_x^t - f_x^{focus}| > T_h \quad x = l, m, r \quad (6)$$

其中, T_h 为清晰度变化值的阈值。

[0044] 尽管已经示出并描述了本发明的实施例, 但本领域的技术人员应该理解, 在未背离发明精神和原理的情况下, 在所附权利要求及其等效物限定的范围内, 可以对这些实例做各种改变。

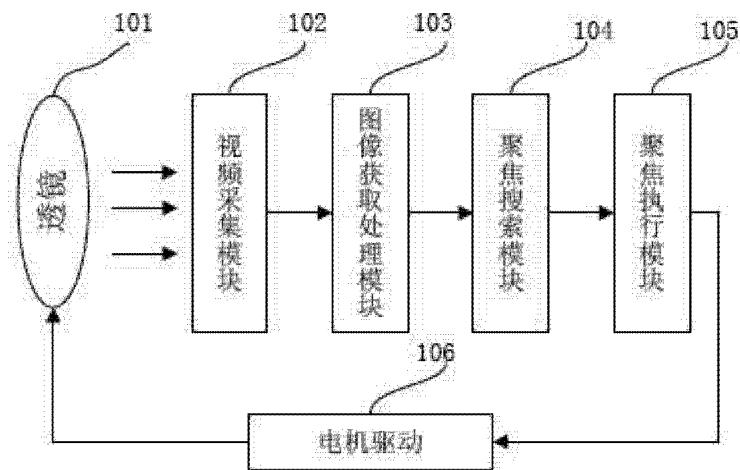


图 1

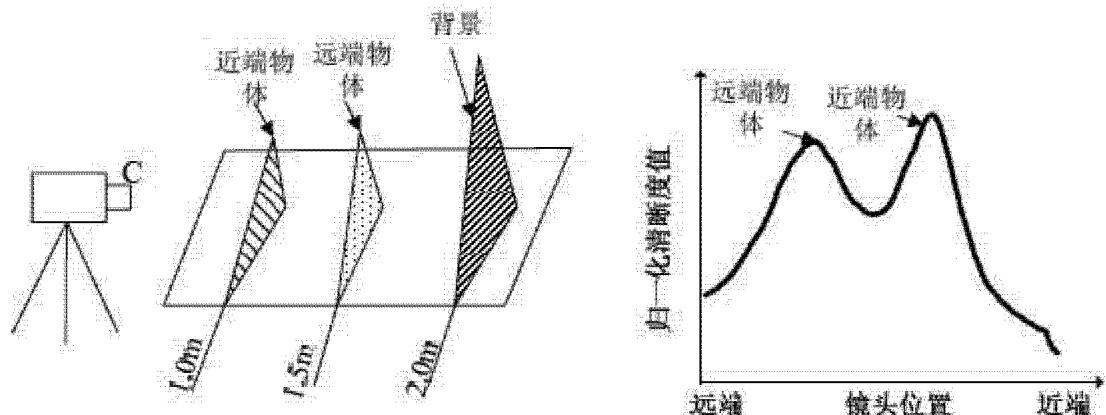


图 2

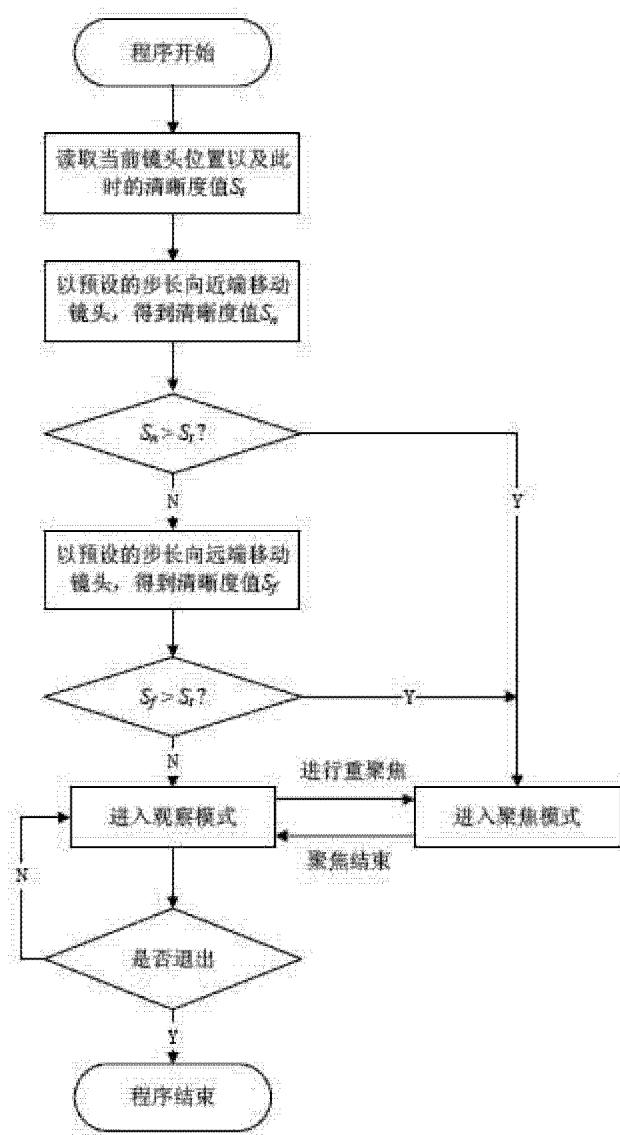


图 3

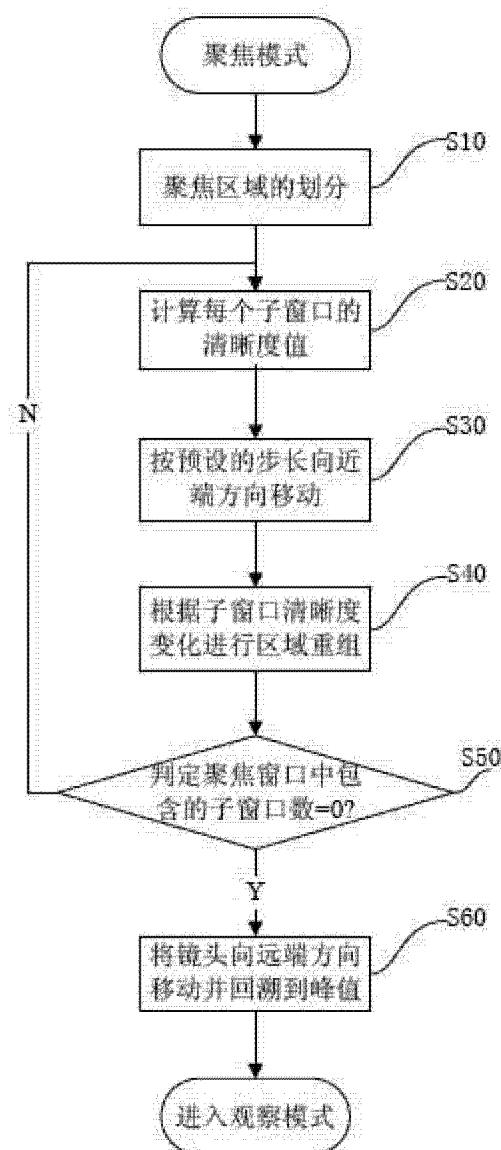


图 4

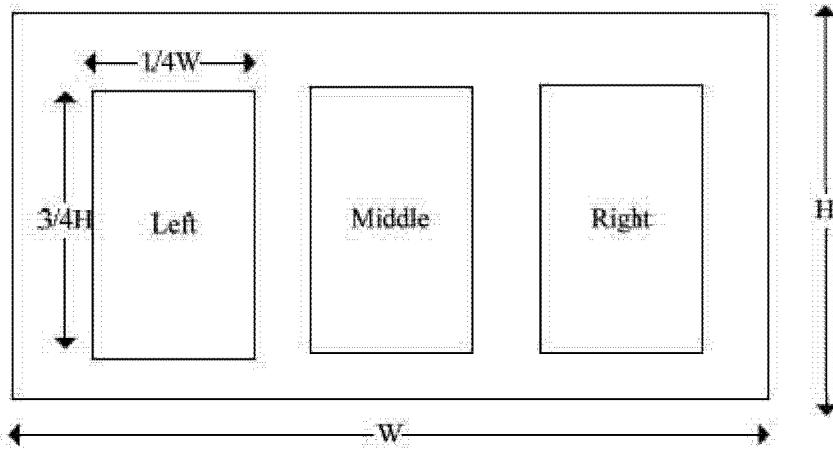


图 5

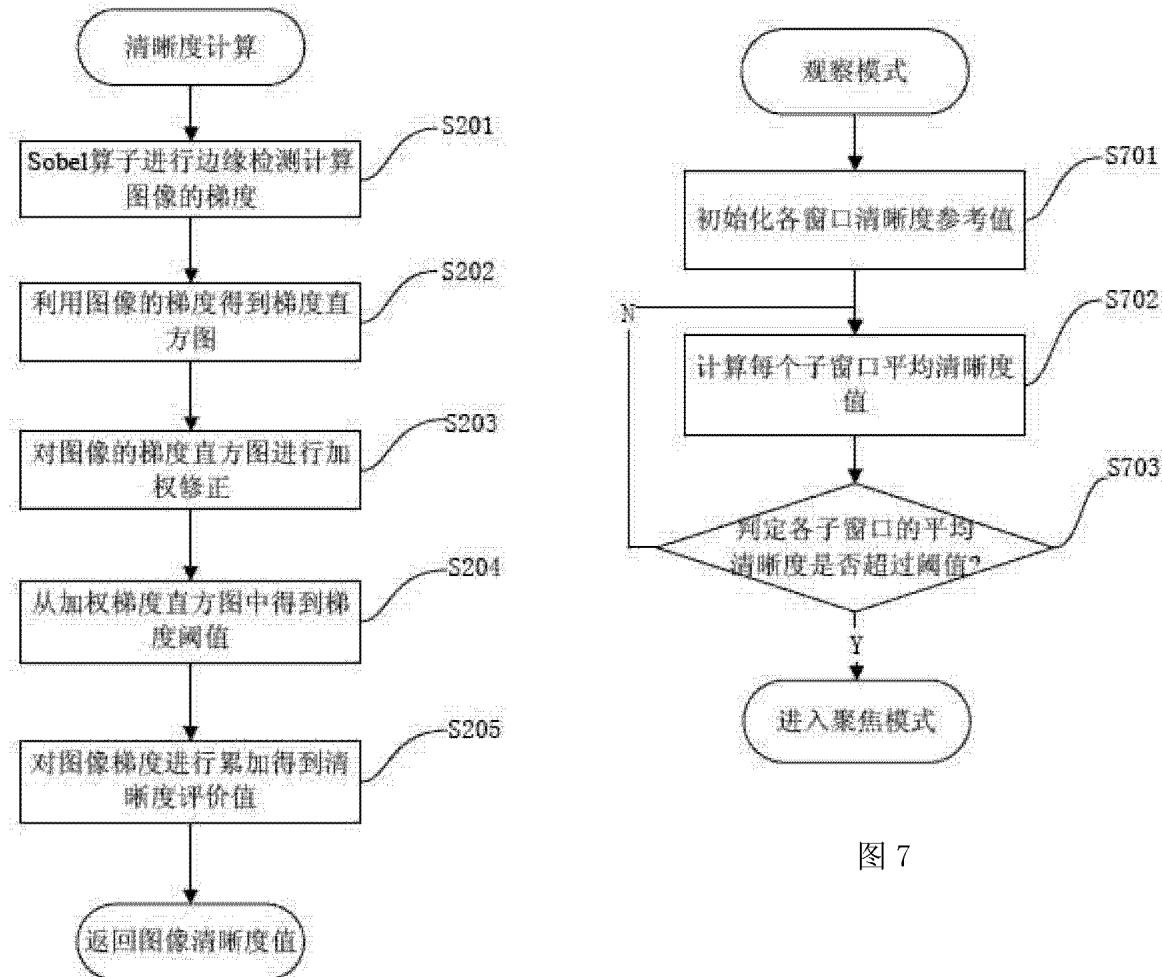


图 6