

(12) 发明专利申请

(10) 申请公布号 CN 103455809 A

(43) 申请公布日 2013. 12. 18

(21) 申请号 201310384280. 6

(22) 申请日 2013. 08. 29

(71) 申请人 方正国际软件有限公司

地址 215123 江苏省苏州市苏州工业园区星湖街 328 号创意产业园方正国际大厦

申请人 方正国际软件(北京)有限公司

(72) 发明人 胡希驰 杨镜

(74) 专利代理机构 北京天悦专利代理事务所  
(普通合伙) 11311

代理人 田明 任晓航

(51) Int. Cl.

G06K 9/20(2006. 01)

H04N 5/225(2006. 01)

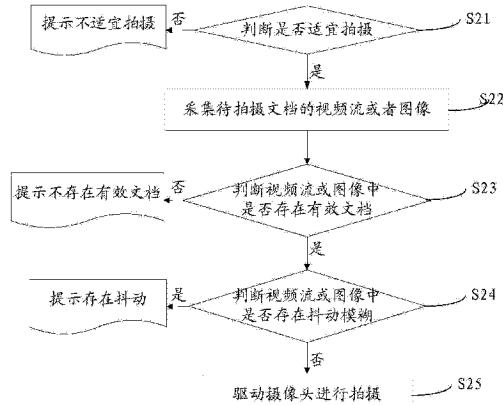
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种自动拍摄文档图像的方法及系统

(57) 摘要

本发明涉及一种自动拍摄文档图像的方法及系统，属于图像拍摄领域。现有技术中大多采用手工方式获取图像，并且拍摄的文档图像难以满足后续处理的要求。本发明所述的方法及系统首先打开摄像头获取其当前参数，判断当前环境是否适宜拍摄，若适宜则通过摄像头采集待拍摄文档的视频流或图像；然后判断所述的视频流或图像中是否存在有效文档；再判断所述的视频流或图像中是否存在抖动模糊；最后驱动摄像头进行自动对焦，完成对待拍摄文档的拍摄，获取文档图像。采用本发明所述的方法及系统拍摄文档图像，能够自动检测拍摄环境是否满足要求，拍摄时是否存在有效文档，拍摄时是否存在抖动，完成图像的自动获取。



1. 一种自动拍摄文档图像的方法,包括以下步骤:

(1) 打开摄像头,获取摄像头的当前参数,根据当前参数判断在当前环境下是否适宜拍摄,若是则进入下一步,若否则提示不适宜拍摄;

(2) 通过摄像头采集待拍摄文档的视频流或者图像;

(3) 判断所述的视频流或者图像中是否存在有效文档,若是则进入下一步,若否则提示不存在有效文档;

(4) 判断所述的视频流或者图像中是否存在抖动模糊,若否则进入下一步,若是则提示存在抖动模糊;

(5) 驱动摄像头进行自动对焦,完成对待拍摄文档的拍摄,获取文档图像。

2. 如权利要求1所述的一种自动拍摄文档图像的方法,其特征在于,步骤(1)中,通过摄像头驱动程序获取摄像头的当前参数,所述参数包括快门速度。

3. 如权利要求2所述的一种自动拍摄文档图像的方法,其特征在于,步骤(1)中,根据当前参数判断在当前环境下是否适宜拍摄的具体方式为:

获取摄像头的快门速度,查看该快门速度是否在设定的快门阈值范围内,若是则说明适宜拍摄。

4. 如权利要求1至3之一所述的一种自动拍摄文档图像的方法,其特征在于,步骤(2)中,通过摄像头采集待拍摄文档的图像后,根据设定的采样值对图像进行降采样处理。

5. 如权利要求1所述的一种自动拍摄文档图像的方法,其特征在于,步骤(3)中,判断所述的视频流或者图像中是否存在有效文档的具体步骤包括:

a) 对视频流的一帧或图像进行二值化处理,得到二值化图像;

b) 对所述二值化图像进行版面分析,确定二值化图像是否能够完成版面分析,若是则判断存在有效文档,若否则判断不存在有效文档;所述的版面分析包括文本字符的成行分析和文本字符的提取。

6. 如权利要求5所述的一种自动拍摄文档图像的方法,其特征在于,步骤b)中,如果二值化图像能够完成文本字符的成行分析,且文本字符间的行间分布均匀,则确定二值化图像能够完成版面分析,判断存在有效文档。

7. 如权利要求1所述的一种自动拍摄文档图像的方法,其特征在于,步骤(4)中,根据视频流的前后帧差判断视频流是否存在抖动模糊,具体方式如下:

获取视频流相邻两帧图像的帧差图像,查看相邻两帧图像的帧差均值是否大于设定的帧差阈值,若是则判断存在抖动模糊,若否则判断不存在抖动模糊。

8. 如权利要求1所述的一种自动拍摄文档图像的方法,其特征在于,步骤(4)中,通过图像边缘检测方法判断图像是否存在抖动模糊,具体方式如下:

1) 利用图像边缘检测方法,计算图像的梯度,生成所述图像的梯度图像;

2) 根据所述的梯度图像计算图像边界,并统计图像的边界宽度;

3) 查看图像的边界宽度是否大于设定的边界阈值,若是则判断存在抖动模糊,若否则判断不存在抖动模糊。

9. 一种自动拍摄文档图像的系统,包括:

拍摄环境判断模块,用于打开摄像头,获取摄像头的当前参数,根据当前参数判断在当前环境下是否适宜拍摄,若是则进入采集模块,若否则提示不适宜拍摄;

采集模块，用于通过摄像头采集待拍摄文档的视频流或者图像；

有效文档检测模块，用于判断所述的视频流或者图像中是否存在有效文档，若是则进入抖动检测模块，若否则提示不存在有效文档；

抖动检测模块，用于判断所述的视频流或者图像中是否存在抖动模糊，若否则进入拍摄模块，若是则提示存在抖动模糊；

拍摄模块，用于驱动摄像头进行自动对焦，完成对待拍摄文档的拍摄，获取文档图像。

10. 如权利要求 9 所述的一种自动拍摄文档图像的系统，其特征在于，所述的有效文档检测模块包括：

二值化单元，用于对视频流的一帧或图像进行二值化处理，得到二值化图像；

有效文档判断单元，用于对所述二值化图像进行版面分析，确定二值化图像是否能够完成版面分析，若是则判断存在有效文档，若否则判断不存在有效文档；所述的版面分析包括文本字符的成行分析和文本字符的提取。

## 一种自动拍摄文档图像的方法及系统

### 技术领域

[0001] 本发明属于图像拍摄领域，具体涉及一种自动拍摄文档图像的方法及系统。

### 背景技术

[0002] 文档图像的拍摄主要用于文本识别的工作，当文本质量较差时，拍摄的图像难以满足文本识别的要求，因此文本识别对文本图像的质量要求较高，需要一种方法能够在一定程度上保证拍摄所得的文本图像的质量。

[0003] 现有技术中，大部分采用手工方式来完成图像的获取，对拍摄条件不能进行自动判断，用户自己主观判断拍摄条件是否满足拍摄要求可能与实际存在差异，这就容易造成拍摄的文档图像难以满足后续处理的要求，不能保证质量。普通的自动拍摄方法则不能完成文档图像存在与否等拍摄要求的自动判断，例如在不存在文档图像的情况下当获取了一张无效的图像，并对其进行后续处理，便必然造成了资源的浪费。

### 发明内容

[0004] 针对现有技术中存在的缺陷，本发明的目的是提供一种自动拍摄文档图像的方法及系统。该方法及系统能够自动检测拍摄环境是否满足要求，拍摄时是否存在有效文档，拍摄时是否存在抖动，完成图像的自动获取。

[0005] 为达到以上目的，本发明采用的技术方案是：

[0006] 一种自动拍摄文档图像的方法，包括以下步骤：

[0007] (1) 打开摄像头，获取摄像头的当前参数，根据当前参数判断在当前环境下是否适宜拍摄，若是则进入下一步，若否则提示不适宜拍摄；

[0008] (2) 通过摄像头采集待拍摄文档的视频流或者图像；

[0009] (3) 判断所述的视频流或者图像中是否存在有效文档，若是则进入下一步，若否则提示不存在有效文档；

[0010] (4) 判断所述的视频流或者图像中是否存在抖动模糊，若否则进入下一步，若是则提示存在抖动模糊；

[0011] (5) 驱动摄像头进行自动对焦，完成对待拍摄文档的拍摄，获取文档图像。

[0012] 进一步，如上述述的一种自动拍摄文档图像的方法，步骤(1)中，通过摄像头驱动程序获取摄像头的当前参数，所述参数包括快门速度。

[0013] 进一步，如上述述的一种自动拍摄文档图像的方法，步骤(1)中，根据当前参数判断在当前环境下是否适宜拍摄的具体方式为：

[0014] 获取摄像头的快门速度，查看该快门速度是否在设定的快门阈值范围内，若是则说明适宜拍摄。

[0015] 进一步，如上述述的一种自动拍摄文档图像的方法，步骤(2)中，通过摄像头采集待拍摄文档的图像后，根据设定的采样值对图像进行降采样处理。

[0016] 进一步，如上述述的一种自动拍摄文档图像的方法，步骤(3)中，判断所述的视频

流或者图像中是否存在有效文档的具体步骤包括：

[0017] a) 对视频流的一帧或图像进行二值化处理，得到二值化图像；

[0018] b) 对所述二值化图像进行版面分析，确定二值化图像是否能够完成版面分析，若是则判断存在有效文档，若否则判断不存在有效文档；所述的版面分析包括文本字符的成行分析和文本字符的提取。

[0019] 进一步，如上所述的一种自动拍摄文档图像的方法，步骤 b) 中，如果二值化图像能够完成文本字符的成行分析，且文本字符间的行间分布均匀，则确定二值化图像能够完成版面分析，判断存在有效文档。

[0020] 再进一步，如上所述的一种自动拍摄文档图像的方法，步骤(4)中，根据视频流的前后帧差判断视频流是否存在抖动模糊，具体方式如下：

[0021] 获取视频流相邻两帧图像的帧差图像，查看相邻两帧图像的帧差均值是否大于设定的帧差阈值，若是则判断存在抖动模糊，若否则判断不存在抖动模糊。

[0022] 更进一步，如上所述的一种自动拍摄文档图像的方法，步骤(4)中，通过图像边缘检测方法判断图像是否存在抖动模糊，具体方式如下：

[0023] 1) 利用图像边缘检测方法，计算图像的梯度，生成所述图像的梯度图像；

[0024] 2) 根据所述的梯度图像计算图像边界，并统计图像的边界宽度；

[0025] 3) 查看图像的边界宽度是否大于设定的边界阈值，若是则判断存在抖动模糊，若否则判断不存在抖动模糊。

[0026] 一种自动拍摄文档图像的系统，包括：

[0027] 拍摄环境判断模块，用于打开摄像头，获取摄像头的当前参数，根据当前参数判断在当前环境下是否适宜拍摄，若是则进入采集模块，若否则提示不适宜拍摄；

[0028] 采集模块，用于通过摄像头采集待拍摄文档的视频流或者图像；

[0029] 有效文档检测模块，用于判断所述的视频流或者图像中是否存在有效文档，若是则进入抖动检测模块，若否则提示不存在有效文档；

[0030] 抖动检测模块，用于判断所述的视频流或者图像中是否存在抖动模糊，若否则进入拍摄模块，若是则提示存在抖动模糊；

[0031] 拍摄模块，用于驱动摄像头进行自动对焦，完成对待拍摄文档的拍摄，获取文档图像。

[0032] 进一步，如上所述的一种自动拍摄文档图像的系统，所述的有效文档检测模块包括：

[0033] 二值化单元，用于对视频流的一帧或图像进行二值化处理，得到二值化图像；

[0034] 有效文档判断单元，用于对所述二值化图像进行版面分析，确定二值化图像是否能够完成版面分析，若是则判断存在有效文档，若否则判断不存在有效文档；所述的版面分析包括文本字符的成行分析和文本字符的提取。

[0035] 本发明的效果在于：采用本发明所述的方法及系统，通过获取摄像头参数，判断当前环境是否适合拍摄识别，如光照等条件是否满足需求；通过定时采集视频流或采样小分辨率图片，来判断有无需要识别的内容；通过对对比度、前后帧差等判断是否存在抖动；当上述条件均满足时，驱动摄像头自动对焦，完成拍摄，该方法不但实现了自动拍摄文档图像，并且尽可能保证了拍摄的文档图像满足后续处理的要求。

## 附图说明

- [0036] 图 1 是本发明具体实施方式中一种自动拍摄文档图像的系统的结构框图；
- [0037] 图 2 是本发明具体实施方式中一种自动拍摄文档图像的方法的流程图；
- [0038] 图 3 是本发明具体实施方式中获取的视频流第 i 帧图像；
- [0039] 图 4 是本发明具体实施方式中获取的视频流第 i+1 帧图像；
- [0040] 图 5 是将图 3 和图 4 做差，得到的一幅帧差图像。

## 具体实施方式

- [0041] 下面结合附图和具体实施方式对本发明作进一步描述。
- [0042] 如图 1 所示，为本发明具体实施方式中一种自动拍摄文档图像的系统的结构框图，该系统主要包括以下五个子模块：拍摄环境判断模块 11、采集模块 12、有效文档检测模块 13、抖动检测模块 14 和拍摄模块 15，其中：
  - [0043] 拍摄环境判断模块 11 用于打开摄像头，获取摄像头的当前参数，根据当前参数判断在当前环境下是否适宜拍摄，若是则进入采集模块，若否则提示不适宜拍摄；
  - [0044] 其中，获取摄像头当前参数的方式包括通过摄像头驱动程序获取，所述参数包括快门速度。一般情况下，当光照条件较好时，摄像头快门较快，因此可以进行相关的拍照实验，当快门速度高于某一数值时，满足拍摄需求，便以此数值作为快门阈值，通过查看获取的快门速度是否在该快门阈值范围内，若是则可以判断出目前的外界环境适宜拍摄。
  - [0045] 另外，摄像头其它的参数，如焦距也比较重要，当对焦不准的话也会产生模糊现象，但是当完成自动对焦后，应该能够消除因为失焦产生的模糊。本具体实施方式仅列出了对快门速度的判断，实际应用中可以根据需要对其它参数进行相应的判断。
- [0046] 采集模块 12 用于通过摄像头采集待拍摄文档的视频流或者图像；
  - [0047] 由于一般情况下，视频流分辨率较低，而图像的采集由于摄像头分辨率的问题，分辨率可能较大（摄像头采集的待拍摄文档图像一般为高分辨率图像（根据相机的分辨率决定）），如果直接采用该分辨率的图像进行分析判断，必然造成效率低下，因此该模块还可以包括一个降采样单元，用于根据设定的采样值对采集到的图像进行降采样，以提高后期对图像进行分析的效率，具体的，根据需要，按原始图像分辨率的几分之一（一般为 2 的指数）对图像进行降采样处理。
  - [0048] 另外，该模块可以定时采集待拍摄文档的视频流或者图像，时间的长短自由设定。
- [0049] 有效文档检测模块 13 用于判断所述的视频流或者图像中是否存在有效文档，若是则进入抖动检测模块，若否则提示不存在有效文档；该模块包括：
  - [0050] 二值化单元，用于对视频流的一帧或图像进行二值化处理，得到二值化图像；
  - [0051] 有效文档判断单元，用于对所述二值化图像进行版面分析，确定二值化图像是否能够完成版面分析，若是则判断存在有效文档，若否则判断不存在有效文档；所述的版面分析包括文本字符的成行分析和文本字符的提取。
- [0052] 抖动检测模块 14 用于判断所述的视频流或者图像中是否存在抖动模糊，若否则进入拍摄模块，若是则提示存在抖动模糊；当存在抖动模糊时，会影响最终的成像质量，该模块通过帧差图像分析单元或者梯度图像分析单元完成抖动的判断，其中：

[0053] 帧差图像分析单元用于根据视频流的前后帧差判断视频流是否存在抖动模糊,具体方式为:获取视频流相邻两帧图像的帧差图像,查看相邻两帧图像的帧差均值是否大于设定的帧差阈值,若是则判断存在抖动模糊,若否则判断不存在抖动模糊。

[0054] 梯度图像分析单元用于通过图像边缘检测法判断图像是否存在抖动模糊,具体方式为首先利用图像边缘检测方法,计算图像的梯度,生成所述图像的梯度图像,然后根据所述的梯度图像计算图像边界,并统计图像的边界宽度,最后查看图像的边界宽度是否大于设定的边界阈值,若是则判断存在抖动模糊,若否则判断不存在抖动模糊。

[0055] 拍摄模块 15 用于驱动摄像头进行自动对焦,完成对待拍摄文档的拍摄获取文档图像;

[0056] 经过以上几个模块的判断处理,在满足拍摄条件、存在有效文档以及不发生抖动模糊时,拍摄模块 15 完成文档图像的获取。

[0057] 如图 2 所示,为本发明具体实施方式中一种自动拍摄文档图像的方法,包括以下步骤:

[0058] 步骤 S21:判断是否适宜拍摄;

[0059] 打开摄像头,获取摄像头的当前参数,根据当前参数判断在当前环境下是否适宜拍摄,若是则进入步骤 S22,若否则提示不适宜拍摄;

[0060] 其中,获取摄像头当前参数的方式包括通过摄像头驱动程序获取,所述参数包括快门速度。一般情况下,当光照条件较好时,摄像头快门较快,因此可以进行相关的拍照实验,当快门速度高于某一数值时,满足拍摄需求,便以此数值作为快门阈值,通过查看获取的快门速度是否在该快门阈值范围内,若是则可以判断出目前的外界环境适宜拍摄。

[0061] 另外,摄像头其它的参数,如焦距也比较重要,当对焦不准的话也会产生模糊现象,但是当完成自动对焦后,应该能够消除因为失焦产生的模糊。本具体实施方式仅列出了对快门速度的判断,实际应用中可以根据需要对其它参数进行相应的判断。

[0062] 步骤 S22:采集待拍摄文档的视频流或者图像;

[0063] 通过摄像头采集待拍摄文档的视频流或者图像;由于一般情况下,视频流分辨率较低,而图像的采集由于摄像头分辨率的问题,分辨率可能较大,如果直接采用该分辨率的图像进行分析判断,必然造成效率低下,因此还需要根据设定的采样值对采集到的图像进行降采样,以提高后期对图像进行分析的效率,具体的,根据需要,按原始图像分辨率的几分之一(一般为 2 的指数)对图像进行降采样处理。

[0064] 另外,可以通过定时的方式采集待拍摄文档的视频流或者图像,时间的长短自由设定。

[0065] 步骤 S23:判断视频流或图像中是否存在有效文档;

[0066] 判断步骤 S22 中所采集到的视频流或者图像中是否存在有效文档,若是则进入步骤 S24,若否则提示不存在有效文档;该步骤判断是否存在有效文档的具体方式包括:

[0067] a) 对视频流的一帧或图像进行二值化处理,得到二值化图像;

[0068] b) 对所述二值化图像进行版面分析,确定二值化图像是否能够完成版面分析,若是则判断存在有效文档,若否则判断不存在有效文档;所述的版面分析包括文本字符的成行分析和文本字符的提取。

[0069] 具体的,如果二值化图像能够完成文本字符的成行分析,且文本字符间的行间分

布均匀，则确定二值化图像能够完成版面分析，判断存在有效文档。

[0070] 其中，版面分析是一种对文本版面进行处理的方法，由于文本一般呈行列分布，首先可以通过简单的投影法，完成文本字符的成行分析，然后通过连通域分析等方法，完成文本字符的提取。一般情况下，如果二值化图像能够完成文本字符的成行分析，且文本字符间的行间分布均匀，便可认为二值化图像能够完成版面分析，判断存在有效文档，但是，如果有更高的要求，可以在完成文本字符的提取后，通过 OCR (Optical Character Recognition, 光学字符识别) 技术判断是否存在有效文字进行判断，不过，在采用 OCR 技术进行判断时，对字符大小有一定要求，最好采用图像而非视频流的方式。

[0071] 本具体实施方式中的版面分析的方式均属于现有技术，本发明对二值化图像进行版面分析包括但不限于上述所列举的两种方式，能够实现版面分析的现有方式均可用于本发明。

[0072] 步骤 S24：判断视频流或图像中是否存在抖动模糊；

[0073] 判断所述的视频流或者图像中是否存在抖动模糊，若否则进入步骤 S25，若是则提示存在抖动模糊；其中，

[0074] 如果步骤 S22 中所采集的是视频流，根据视频流的前后帧差判断视频流是否存在抖动模糊，具体方式如下：

[0075] 获取视频流相邻两帧图像的帧差图像，查看相邻两帧图像的帧差均值是否大于设定的帧差阈值，若是则判断存在抖动模糊，若否则判断不存在抖动模糊。

[0076] 图 3 和图 4 分别示出了获取的视频流的第 i 帧和第 i+1 帧图像，将图 3 和图 4 做差，即第 i+1 帧图像减去第 i 帧图像，得到一张帧差图像，如图 5 所示，查看这张帧差图像中所有点的帧差均值(帧差图像的平均灰度值)，若大于预设的阈值，则认为第 i 帧和第 i+1 帧图像之间存在较大差异，判断存在抖动模糊，不适合摄像头成像的需求。

[0077] 如果步骤 S22 中所采集的是图像，通过图像边缘检测法判断图像是否存在抖动模糊，具体方式如下：

[0078] 1) 利用图像边缘检测方法，计算图像的梯度，生成所述图像的梯度图像；

[0079] 2) 根据所述的梯度图像计算图像边界，并统计图像的边界宽度；

[0080] 3) 查看图像的边界宽度是否大于设定的边界阈值，若是则判断存在抖动模糊，若否则判断不存在抖动模糊。

[0081] 具体的，所述的梯度图像包括水平方向梯度图像和竖直方向梯度图像，根据当图像存在抖动模糊时，图像不存在清晰的边缘这个原理，首先，通过这两种梯度图像分别计算原图像的水平边界和竖直边界，并统计出水平边界宽度及竖直边界宽度；然后，根据事先设定的水平边界宽度阈值和竖直边界宽度阈值(可以是统一的也可以是单独的)进行判断，若对应的边界宽度大于所设定的边界宽度阈值，则认为在该对应的方向上存在抖动模糊；最后，考虑到边界宽度通常并不均一，也可以对比竖直边界宽度与水平边界宽度，若竖直边界宽度大于水平边界宽度则认为存在水平方向抖动模糊；若水平边界宽度大于竖直边界宽度则认为存在竖直方向抖动模糊。

[0082] 本发明中判断图像模糊的方法属于现有技术，详细内容可以参考图像模糊度评价方法文献：A no-reference perceptual blur metric。

[0083] 步骤 S25：驱动摄像头进行拍摄；

[0084] 经过以上四个步骤,在满足外界拍摄条件、存在有效文档以及不发生抖动模糊时,驱动摄像头进行自动对焦,完成对待拍摄文档的拍摄,获取文档图像。

[0085] 通过本实施例可以看出,采用本发明所述的方法及系统,通过获取摄像头参数,判断当前环境是否适合拍摄识别,如光照等条件是否满足需求;通过实时采集视频流,或定时采样小分辨率图片,来判断有无需要识别的内容;通过对比度、前后帧差等判断是否存在抖动模糊;当上述条件均满足时,驱动摄像头自动对焦,完成拍摄。通过该方法,不但实现了自动拍摄文档图像,并且尽可能保证了拍摄的文档图像满足后续处理的要求。

[0086] 本发明所述的方法和系统并不限于具体实施方式中所述的实施例,本领域技术人员根据本发明的技术方案得出其他的实施方式,同样属于本发明的技术创新范围。

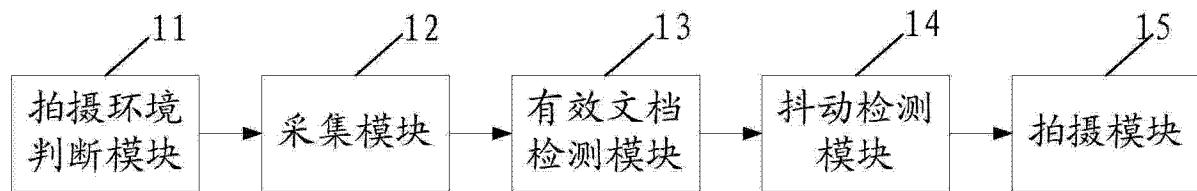


图 1

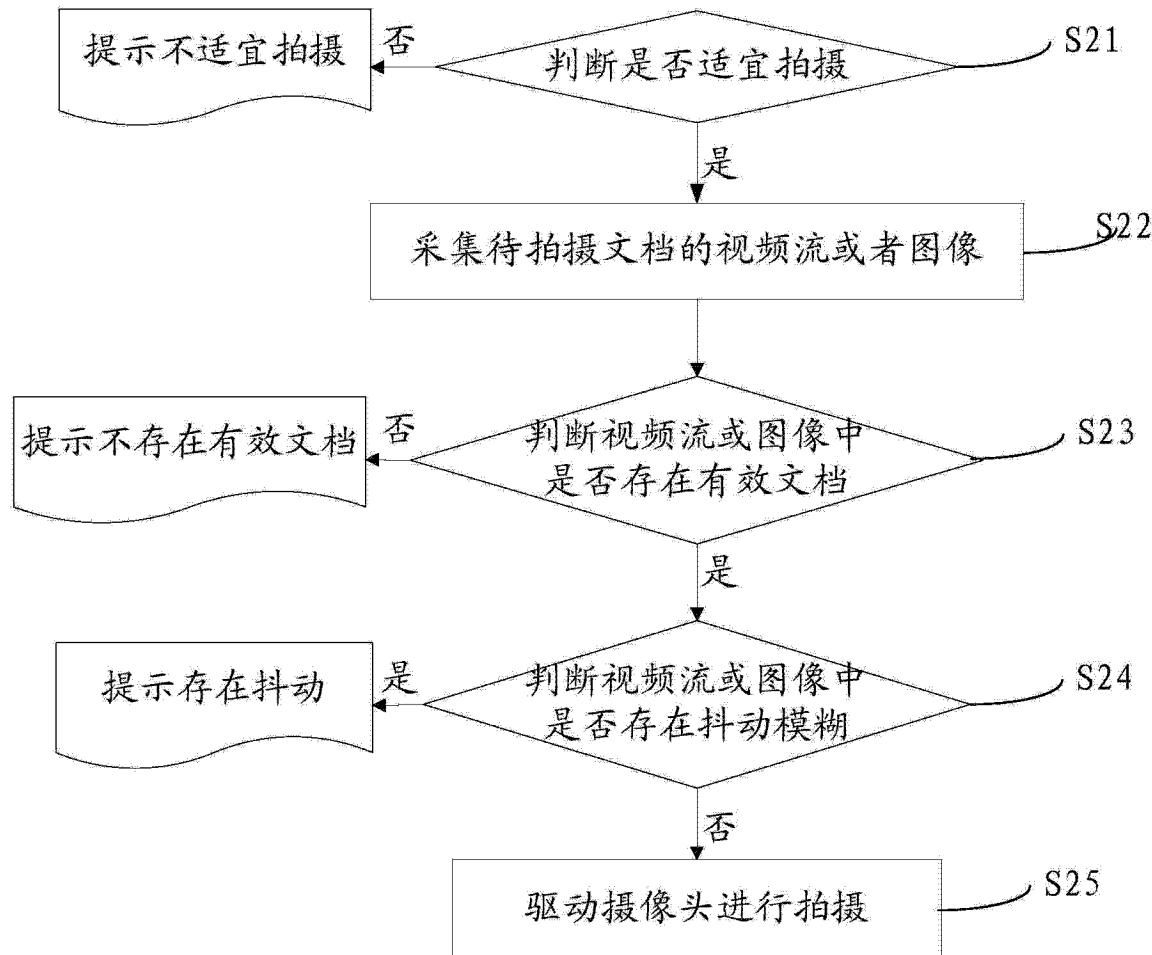


图 2

+4° has area members When best F  
re hongzhizaishang996, 2006 you jK  
tems Member go '\_ E image comme  
@@ search »' become West add ¥~  
Image Design don't Medical » §§27:  
nt last Time Council Art 9 under FR  
Use message study \_Q time Micro  
hool Save one q! It's Canada USA s  
hash) Name: any new #digitaldistrac  
issimilationists,' An East very Feed  
3. ¥1764(¥1680 next >> Windows P

图 3

+4° has area members When best F  
re hongzhizaishang996, 2006 you jK  
tems Member go '\_ E image comme  
@@ search »' become West add ¥~  
Image Design don't Medical » §§27:  
nt last Time Council Art 9 under FR  
Use message study \_Q time Micro  
hool Save one q! It's Canada USA s  
hash) Name: any new #digitaldistrac  
issimilationists,' An East very Feed  
3. ¥1764(¥1680 next >> Windows P

图 4

+4° has area members When best F  
re hongzhizaishang996, 2006 you jK  
tems Member go '\_ E image comme  
@@ search »' become West add ¥~  
Image Design don't Medical » §§27:  
nt last Time Council Art 9 under FR  
Use message study \_Q time Micro  
hool Save one q! It's Canada USA s  
hash) Name: any new #digitaldistrac  
issimilationists,' An East very Feed  
3. ¥1764(¥1680 next >> Windows P

图 5