



(12)发明专利申请

(10)申请公布号 CN 109743566 A

(43)申请公布日 2019.05.10

(21)申请号 201811572568.5

(22)申请日 2018.12.21

(71)申请人 重庆爱奇艺智能科技有限公司  
地址 401133 重庆市江北区鱼嘴镇永和路  
39号2层208室

(72)发明人 史明 王西颖

(74)专利代理机构 北京汉昊知识产权代理事务  
所(普通合伙) 11370

代理人 罗朋

(51) Int. Cl.

H04N 17/00(2006.01)

H04N 13/106(2018.01)

H04N 21/234(2011.01)

H04N 21/44(2011.01)

H04N 21/81(2011.01)

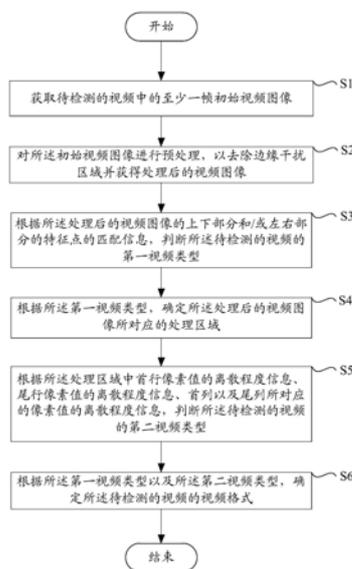
权利要求书3页 说明书16页 附图5页

(54)发明名称

一种用于识别VR视频格式的方法与设备

(57)摘要

本申请的目的是提供一种用于识别VR视频格式的方法与设备。本申请实现了VR视频格式的自动识别:首先通过预处理操作,去除了视频图像的边缘干扰区域,从而提高了识别准确性;接下来通过两次识别过程,分别识别了待检测的视频的第一视频类型以及第二视频类型,最终能够识别至少9种格式的VR视频,使得VR视频的格式识别快速、高效且全面。此外,本申请的整个识别过程对用户透明,进而播放器能够以正确的播放方式实现VR视频的播放,提高了应用的友好性,改善用户体验。



1. 一种用于识别VR视频格式的方法,其中,该方法包括以下步骤:
  - a获取待检测的视频中的至少一帧初始视频图像;
  - b对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像;
  - c根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型;
  - d根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域;
  - e根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频;
  - f根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。
2. 根据权利要求1所述的方法,其中,所述步骤b包括:
  - b1将所述初始视频图像转换为灰度图;对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;  
根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;  
去除所述边缘干扰区域并获得处理后的视频图像。
3. 根据权利要求2所述的方法,其中,该方法还包括:  
将所述初始视频图像缩放至预定大小;  
其中,所述步骤b1包括:  
将所述缩放后的初始视频图像转换为灰度图。
4. 根据权利要求1至3中任一项所述的方法,其中,所述步骤c包括:  
确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;  
c1若有任意一个所述匹配信息大于预定阈值,则判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。
5. 根据权利要求4所述的方法,其中,所述步骤c1包括:  
若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断所述待检测的视频的第一视频类型为上下3D类型;和/或  
若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;  
若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则判断所述待检测的视频的第一视频类型为非3D类型。
6. 根据权利要求1至5中任一项所述的方法,其中,所述步骤e包括:  
确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;  
若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;  
若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离

散阈值,则所述第二视频类型为180度内容视频;

若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

7. 根据权利要求1至6中任一项所述的方法,其中,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和。

8. 一种用于识别VR视频格式 的识别设备,其中,所述识别设备包括:

第一装置,用于获取待检测的视频中的至少一帧初始视频图像;

第二装置,用于对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像;

第三装置,用于根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型;

第四装置,用于根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域;

第五装置,用于根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频;

第六装置,用于根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。

9. 根据权利要求8所述的识别设备,其中,所述第二装置用于:

将所述初始视频图像转换为灰度图;

对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;

根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;

去除所述边缘干扰区域并获得处理后的视频图像。

10. 根据权利要求9所述的识别设备,其中,所述识别设备还包括:

第七装置,用于将所述初始视频图像缩放至预定大小;

其中,所述第二装置用于:

将所述缩放后的初始视频图像转换为灰度图;

对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;

根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;

去除所述边缘干扰区域并获得处理后的视频图像。

11. 根据权利要求8至10中任一项所述的识别设备,其中,所述第三装置包括:

三一单元,用于确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;

三二单元,用于若有任意一个所述匹配信息大于预定阈值,则判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

12. 根据权利要求11所述的识别设备,其中,所述三二单元用于:

若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断

所述待检测的视频的第一视频类型为上下3D类型;和/或

若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;

若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则判断所述待检测的视频的第一视频类型为非3D类型。

13. 根据权利要求8至12中任一项所述的识别设备,其中,所述第五装置用于:

确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;

若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;

若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为180度内容视频;

若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

14. 根据权利要求8至13中任一项所述的识别设备,其中,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和。

15. 一种计算机设备,所述计算机设备包括:

一个或多个处理器;

存储器,用于存储一个或多个计算机程序;

当所述一个或多个计算机程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1至7中任一项所述的方法。

16. 一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序可被处理器执行如权利要求1至7中任一项所述的方法。

## 一种用于识别VR视频格式的方法与设备

### 技术领域

[0001] 本申请涉及虚拟现实技术领域,尤其涉及一种用于识别VR视频格式的技术。

### 背景技术

[0002] 虚拟现实技术(VR,Virtual Reality)的发展激发了人们的观影兴趣,同时对VR影片的质量要求也越来越高。为满足不同用户的观影需求,影片的生产者不断地创造新的VR视频格式;而只有预先知道VR视频的格式,并采用正确的播放方式,才能给用户带来最佳的观影体验。因此,预先识别VR视频格式是非常关键的。

[0003] VR视频可以分为2D视频和3D视频。2D视频又可以分为普通视频、180度视频以及全景视频。3D视频同样也可以分为普通3D视频、180度3D视频以及360度3D视频。进一步地,3D视频内容的排列又可以分为上下和左右两种方式。综上所述,2D视频共包括三种情况,即普通视频、180度视频以及全景视频,而3D视频共包括六种情况,即普通上下排列视频、普通左右排列视频、180度上下排列视频、180度左右排列视频、全景上下排列视频以及全景左右排列视频。

[0004] 对于某些VR视频,应用程序可以根据服务器接口获取VR视频格式,但是一些用户提交的个人视频却无法确定格式。因此,用户只能先将VR视频播放出来,然后根据个人经验选取一种播放方式进行播放。整个播放过程大大降低了应用的友好性,也产生了负面的用户体验。

[0005] 此外,现有的VR视频格式识别技术所识别的类型覆盖较少,不能满足上述各类VR视频格式的设备情况。此外,还有一些设备方法对于识别全景视频存在缺陷,例如,当视频宽高比为2:1的时候就认定该VR视频为全景视频,实际上宽高比为2:1的视频既可以是普通视频,也可以是3D视频甚至其他视频。还有一些办法认为如果VR视频左右两侧可以拼接在一起就认定为全景视频,这种方法也是有缺陷的,例如有些视频源的左右或者上下是带有黑边的,这样势必会导致拼接成功而被错误的认定为全景视频。

### 发明内容

[0006] 本申请的目的是提供一种用于识别VR视频格式的方法与设备。

[0007] 根据本申请的一个实施例,提供了一种用于识别VR视频格式的方法,其中,该方法包括以下步骤:

[0008] a获取待检测的视频中的至少一帧初始视频图像;

[0009] b对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像;

[0010] c根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型;

[0011] d根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域;

[0012] e根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、

首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频;

[0013] f根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。

[0014] 可选地,所述步骤b包括:

[0015] b1将所述初始视频图像转换为灰度图;

[0016] 对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;

[0017] 根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;

[0018] 去除所述边缘干扰区域并获得处理后的视频图像。

[0019] 可选地,该方法还包括:

[0020] 将所述初始视频图像缩放至预定大小;

[0021] 其中,所述步骤b1包括:

[0022] 将所述缩放后的初始视频图像转换为灰度图。

[0023] 可选地,所述步骤c包括:

[0024] 确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;

[0025] c1若有任意一个所述匹配信息大于预定阈值,则判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

[0026] 可选地,所述步骤c1包括:

[0027] 若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断所述待检测的视频的第一视频类型为上下3D类型;和/或

[0028] 若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;

[0029] 若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则判断所述待检测的视频的第一视频类型为非3D类型。

[0030] 可选地,所述步骤e包括:

[0031] 确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;

[0032] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;

[0033] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为180度内容视频;

[0034] 若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

[0035] 可选地,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和。

- [0036] 根据本申请的另一个实施例,还提供了一种计算机设备,所述计算机设备包括:
- [0037] 一个或多个处理器;
- [0038] 存储器,用于存储一个或多个计算机程序;
- [0039] 当所述一个或多个计算机程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如上述任一项所述的方法。
- [0040] 根据本申请的另一个实施例,还提供了一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序可被处理器执行如上述任一项所述的方法。
- [0041] 根据本申请的另一个实施例,还提供了一种用于识别VR视频格式的识别设备,其中,所述识别设备包括:
- [0042] 第一装置,用于获取待检测的视频中的至少一帧初始视频图像;
- [0043] 第二装置,用于对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像;
- [0044] 第三装置,用于根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型;
- [0045] 第四装置,用于根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域;
- [0046] 第五装置,用于根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频;
- [0047] 第六装置,用于根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。
- [0048] 可选地,所述第二装置用于:
- [0049] 将所述初始视频图像转换为灰度图;
- [0050] 对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;
- [0051] 根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;
- [0052] 去除所述边缘干扰区域并获得处理后的视频图像。
- [0053] 可选地,所述识别设备还包括:
- [0054] 第七装置,用于将所述初始视频图像缩放至预定大小;
- [0055] 其中,所述第二装置用于:
- [0056] 将所述缩放后的初始视频图像转换为灰度图;
- [0057] 对所述灰度图进行边缘检测,并对所述边缘检测的结果进行积分处理;
- [0058] 根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;
- [0059] 去除所述边缘干扰区域并获得处理后的视频图像。
- [0060] 可选地,所述第三装置包括:
- [0061] 三一单元,用于确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;
- [0062] 三二单元,用于若有任意一个所述匹配信息大于预定阈值,则判断所述待检测的

视频的第一视频类型为3D类型,反之,则为非3D类型。

[0063] 可选地,所述三二单元用于:

[0064] 若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断所述待检测的视频的第一视频类型为上下3D类型;和/或

[0065] 若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;

[0066] 若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则判断所述待检测的视频的第一视频类型为非3D类型。

[0067] 可选地,所述第五装置用于:

[0068] 确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;

[0069] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;

[0070] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为180度内容视频;

[0071] 若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

[0072] 可选地,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和。

[0073] 与现有技术相比,本申请实现了VR视频格式的自动识别:首先通过预处理操作,去除了视频图像的边缘干扰区域,从而提高了识别准确性;接下来通过两次识别过程,分别识别了待检测的视频的第一视频类型以及第二视频类型,最终能够识别至少9种格式的VR视频,使得VR视频的格式识别快速、高效且全面。此外,本申请的整个识别过程对用户透明,进而播放器能够以正确的播放方式实现VR视频的播放,提高了应用的友好性,改善用户体验。

## 附图说明

[0074] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本申请的其它特征、目的和优点将会变得更明显:

[0075] 图1示出根据本申请的一个实施例的一种用于识别VR视频格式的识别设备示意图;

[0076] 图2示出根据本申请的一个实施例的一种用于识别VR视频格式的方法流程图;

[0077] 图3示出根据本申请的一个实施例的从待检测的视频中所获取的一帧初始视频图像;

[0078] 图4示出对图3所示初始视频图像进行边缘检测后的灰度图;

[0079] 图5示出对图4所示的灰度图进行积分处理后的积分图;

- [0080] 图6示出对图3所示的初始视频图像进行预处理后的处理后的视频图像；
- [0081] 图7示出一种判断处理后的视频图像的左右部分的特征点的匹配信息的示意图；
- [0082] 图8示出一种全景内容的示意图；
- [0083] 图9示出了可被用于实施本申请中所述的各个实施例的示例性系统。
- [0084] 附图中相同或相似的附图标记代表相同或相似的部件。

## 具体实施方式

[0085] 下面结合附图对本申请作进一步详细描述。

[0086] 本申请所指识别设备包括但不限于用户设备、网络设备、或用户设备与网络设备通过网络相集成所构成的设备。所述用户设备包括但不限于任何一种可与用户进行人机交互的电子产品，例如虚拟现实个人终端、个人电脑、智能手机、平板电脑等，所述电子产品可以采用任意操作系统，如windows操作系统、android操作系统、iOS操作系统等。其中，所述网络设备包括一种能够按照事先设定或存储的指令，自动进行数值计算和信息处理的电子设备，其硬件包括但不限于微处理器、专用集成电路(ASIC)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、数字信号处理器(DSP)、嵌入式设备等。所述网络设备包括但不限于计算机、网络主机、单个网络服务器、多个网络服务器集或多个服务器构成的云；在此，云基于云计算(Cloud Computing)的大量计算机或网络服务器构成，其中，云计算是分布式计算的一种，由一群松散耦合的计算机集组成的一个虚拟超级计算机。所述网络包括但不限于互联网、广域网、城域网、局域网、VPN网络、无线自组织网络(Ad Hoc网络)等。优选地，所述设备还可以是运行于所述用户设备、网络设备、或用户设备与网络设备、网络设备、触摸终端或网络设备与触摸终端通过网络相集成所构成的设备上的程序。

[0087] 当然，本领域技术人员应能理解上述识别设备仅为举例，其他现有的或今后可能出现的设备如可适用于本申请，也应包含在本申请保护范围以内，并在此以引用方式包含于此。

[0088] 在本申请的描述中，“多个”的含义是两个或者更多，除非另有明确具体的限定。

[0089] 图1示出根据本申请的一个实施例的一种用于识别VR视频格式的识别设备示意图；其中，所述识别设备包括第一装置1、第二装置2、第三装置3、第四装置4、第五装置5以及第六装置6。

[0090] 具体地，所述第一装置1获取待检测的视频中的至少一帧初始视频图像；所述第二装置2对所述初始视频图像进行预处理，以去除边缘干扰区域并获得处理后的视频图像；所述第三装置3根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息，判断所述待检测的视频的第一视频类型，其中，所述第一视频类型包括3D类型或非3D类型；所述第四装置4根据所述第一视频类型，确定所述处理后的视频图像所对应的处理区域；所述第五装置5根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息，判断所述待检测的视频的第二视频类型，其中，所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频；所述第六装置6根据所述第一视频类型以及所述第二视频类型，确定所述待检测的视频的视频格式。

[0091] 所述第一装置1获取待检测的视频中的至少一帧初始视频图像。

[0092] 具体地，所述待检测的视频可以是任意需要检测的视频，优选地，所述待检测的视

频是从VR视频播放设备中所获取的视频。所述待检测的视频可以从播放系统中获取的，也可以是用户自行上传的。

[0093] 然后，所述第一装置1从所述待检测的视频中截取至少一帧初始视频图像，例如，所述第一装置1按照预定的提取位置、提取时间等信息，从所述待检测的视频的该提取位置或该提取时间截取至少一帧初始视频图像；或者，所述第一装置1也可以与其他提供初始视频图像的设备相交互，直接获取待检测的视频中的至少一帧初始视频图像。

[0094] 优选地，所述初始视频图像为所述待检测的视频中的关键帧。

[0095] 所述第二装置2对所述初始视频图像进行预处理，以去除边缘干扰区域并获得处理后的视频图像。

[0096] 具体地，所述边缘干扰区域包括但不限于任意一种纯色边缘区域，如黑边区域、白边区域、红边区域等，所述边缘干扰区域中不存在图像变换。所述第二装置2通过对所述初始视频图像进行积分处理、扫描像素点等方式，检测所述初始视频图像所对应的黑边区域，并将该边缘干扰区域进行剪裁，以去除所述边缘干扰区域，从而实现所述初始视频图像的预处理。

[0097] 优选地，所述第二装置2将所述初始视频图像转换为灰度图；对所述灰度图进行边缘检测，并对所述边缘检测的结果进行积分处理；根据所述积分处理结果，确定所述初始视频图像所对应的边缘干扰区域；去除所述边缘干扰区域并获得处理后的视频图像。

[0098] 具体地，所述第二装置2根据现有的各类图像转换方式，将所述初始视频图像转换为灰度图；然后，对所述灰度图进行边缘检测，以将边缘响应较强的部分凸显出来，在此，所述边缘检测方法包括但不限于Canny、Sobel等。

[0099] 例如，图3示出根据本申请的一个实施例的从待检测的视频中所获取的一帧初始视频图像，该初始视频图像包括边缘干扰区域，即边缘的黑边部分。通过对所述初始视频图像进行边缘检测，则得到了图4所示的对该初始视频图像进行边缘检测后的灰度图。

[0100] 然后，对该灰度图进行积分处理，以生成积分图。如图5所示，图5即是对图4所示的灰度图进行积分后处理后的积分图。根据所述积分处理结果，能够确定该初始视频图像的图像变化信息，从而确定所述初始视频图像所对应的边缘干扰区域，也即图5所示的黑边区域；最后去除所述边缘干扰区域并获得处理后的视频图像。在此，图6示出对图3所示的初始视频图像进行预处理后的处理后的视频图像。

[0101] 其中，所述积分处理的过程如下所示：

[0102] 以I表示积分图，G表示灰度图，则 $I(x, y) = \text{sum}(G(i, j))$ ，其中， $0 \leq i \leq x, 0 \leq j \leq y$ 。在此，x, y, i, j表示坐标，I(x, y)，以及G(i, j)表示该点像素值，该式的计算意义即为：通过图像累加来显示该图像的变化程度。

[0103] 以所述边缘干扰区域为黑边区域为例，在所述积分图中，黑色部分的数值为0，非黑色部分的数值大于0。通过图5所示的积分图可以看到，当横向扫描到第m列的时候，出现了大量的非零点，即白色像素点，说明从m列开始，图像出现了巨大的变化。这是因为原图存在一定列数的黑边，而当扫描到非黑边时，则原图的数值产生了变化，换言之，是原图的黑边导致了上述变化。因此，可以以m列为分割点，将原图的左侧黑边裁掉。

[0104] 由于黑边的存在往往是对称的，所以也可将右侧的m个像素黑边裁掉；或者，继续对所述图像进行横向扫描，当横向扫描到第m+k列时，出现了大量的零点，即黑色像素点，说

明从 $m+k$ 列开始,图像出现了第二次变化,从而将 $m+k$ 列至右侧顶端的像素黑边裁掉。

[0105] 同理,纵向扫描积分图,在 $n$ 行的时候出现了大量的非零点,即白色像素点,说明从 $n$ 行开始原图出现了巨大变化,这个变化也是由于黑边导致的,因此裁掉原图上侧的 $n$ 行。类似地,也可以裁掉原图对称的下侧 $n$ 行,或者,继续进行扫描,根据扫描结果来裁掉下侧的黑边。

[0106] 更优选地,所述识别设备还包括第七装置(未示出),其中,所述第七装置将所述初始视频图像缩放至预定大小;然后,所述第二装置2对所述缩放后的初始视频图像进行处理。

[0107] 具体地,所述第七装置按照所述初始视频图像的宽高比,将所述初始视频图像进行等比例缩放,以缩放至预定大小;或者,所述第七装置按照预定的比例,将所述初始视频图像进行缩放,以缩放至预定大小;或者,所述第七装置按照预定的图像存储大小,对所述初始视频图像进行缩放,以缩放至预定大小。

[0108] 在此,所述预定大小可以由用户自行设置,也可以根据所述识别设备的处理能力确定。

[0109] 然后,所述第二装置2对所述缩放后的初始视频图像进行处理,以实现快速处理。

[0110] 所述第三装置3根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型。

[0111] 具体地,所述第三装置3将所述处理后的视频图像分为上下两个图像和/或左右两个图像;然后,分别确定所述上下两个图像和/或左右两个图像的上下特征点和/或左右特征点,在此,所述确定方法包括但不限于计算BRIEF特征描述子或ORB特征描述子;接下来,计算所述上下特征点和/或左右特征点的匹配信息,例如,利用Hamming距离来判定所述上下特征点和/或左右特征点是否匹配。最后,基于所计算的匹配信息来判定所述待检测的视频的第一视频类型。

[0112] 在此,所述非3D类型包括2D类型。

[0113] 优选地,所述第三装置3包括三一单元(未示出)以及三二单元(未示出),其中,所述三一单元确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;若有任意一个所述匹配信息大于预定阈值,则所述三二单元判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

[0114] 为便于说明,以下以将该处理后的视频图像分为左右两个图像为例进行说明。

[0115] 具体地,所述三一单元首先将该处理后的视频图像分为左右两个图像,对所述左右两个图像分别检测角点,然后通过例如计算BRIEF特征描述子或ORB特征描述子的方式,计算左右两个图像的特征点。在此,根据3D视频本身的特性可知,左右内容差异是由一定的视差导致的,不存在特征旋转以及尺度变化的情况,因此优选地,可以采用速度较快的BRIEF。

[0116] 然后,所述三一单元使用如hamming距离来计算左右两组特征描述子的距离,若所述hamming距离小于某一阈值,即可表示对应的左右两个特征点匹配。在此,所匹配的特征点的个数即可作为所述处理后的视频图像的左右部分的特征点的匹配信息。

[0117] 图7示出一种判断处理后的视频图像的左右部分的特征点的匹配信息的示意图。

图7示出了左右两部分的特征描述子,以及每对特征描述子的距离信息。

[0118] 若所述上下部分的特征点的匹配信息或所述左右部分的特征点的匹配信息的任意一个大于预定的阈值,则所述三二单元判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

[0119] 例如,继上例,若所述匹配的特征点的个数大于某一数量N时,则判定所述待检测的视频的第一视频类型为3D类型,进一步地,为左右3D类型。

[0120] 优选地,所述三二单元用于:

[0121] 若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断所述待检测的视频的第一视频类型为上下3D类型;和/或

[0122] 若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;

[0123] 若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则判断所述待检测的视频的第一视频类型为非3D类型。

[0124] 本领域技术人员应能理解,所述第一特征阈值可以等同于所述第二特征阈值,所述第一特征阈值也可以不等同于所述第二特征阈值。

[0125] 所述第四装置4根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域。

[0126] 在此,所述处理区域即为ROI (Region of Interest),即对该视频图像的后续处理区域。

[0127] 具体地,若所述第一视频类型是非3D类型,则可以将所述处理后的视频图像的整个图像直接作为处理区域,以进行后续处理;

[0128] 若所述第一视频类型为左右3D类型,则可截取所述处理后的视频图像的左半部分或右半部分作为所述处理区域,以进行后续处理;

[0129] 若所述第一视频类型为上下3D类型,则可截取所述处理后的视频图像的上半部分或下半部分作为所述处理区域,以进行后续处理。

[0130] 所述第五装置5根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频。

[0131] 在此,所述全景内容表示Equirectangular模式的投影。图8示出一种全景内容的示意图,其表示了从地球仪到世界地图的映射关系。如图8可知,全景图的首行(第一行)是由球面的上极点展开的,尾行(最后一行)是由球面的下极点展开的。因此,全景图的首行像素值应该为同一值,尾行像素值也为同一值;可选地,由于在展开过程中存在插值情况,因此,首行像素值与尾行像素值可以存在一定偏差。此外,由全景图的展开方式可知,全景图左右两侧能够无缝拼接在一起。

[0132] 因此,所述第五装置5分别计算所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;在此,优选地,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和,也即,可以采用

方差来表示所述离散程度信息,也可以采用每个样本值与全体样本值的平均数之差的和来表示所述离散程度信息。

[0133] 优选地,所述第五装置用于:

[0134] 确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;

[0135] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;

[0136] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为180度内容视频;

[0137] 若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

[0138] 例如,假设所述处理区域的宽为 $w$ ,高为 $h$ ,则首行中每个像素的像素值可以用 $P(0, j)$ 表示,其中, $i$ 的取值范围为 $[0, w-1]$ ,尾行中每个像素的像素值可以用 $P(h-1, j)$ 表示,其中 $j$ 的取值范围为 $[0, w-1]$ 。同理,首列中每个像素的像素值可以用 $P(m, 0)$ 表示,其中 $m$ 的取值范围为 $[0, h-1]$ ,尾列中每个像素的像素值可以用 $P(n, w-1)$ ,其中 $n$ 的取值范围为 $[0, h-1]$ 。

[0139] 则以采用每个样本值与全体样本值的平均数之差的和来表示所述离散程度信息为例:

[0140] 首行像素值的离散程度信息 $V_{topPolar}$ 为:
$$V_{topPolar} = \sum_{i=0}^{w-1} (P(0, i) - \sum_{i=0}^{w-1} (P(0, i) / w))$$

[0141] 尾行像素值的离散程度信息 $V_{bottomPolar}$ 为:

$$V_{bottomPolar} = \sum_{j=0}^{w-1} (P(h-1, j) - \sum_{j=0}^{w-1} (P(h-1, j) / w))$$

[0142] 首列以及尾列所对应的像素值的离散程度信息 $V_{diff}$ 为:

$$V_{diff} = \sum_{m=0}^h |P(m, 0) - P(m, w-1)| / h$$

[0143] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,则可认为该图的上下两边是由极点展开的;若 $V_{diff}$ 小于第三离散阈值 $T_3$ ,则可认为该图的左右两边可以无缝拼接在一起的。在此,所述第一离散阈值 $T_1$ 、第二离散阈值 $T_2$ 、第三离散阈值 $T_3$ 可以根据图像由球面展开成柱面时的插值操作来进行取值,例如,若插值较多,则可将所述第一离散阈值 $T_1$ 、第二离散阈值 $T_2$ 、第三离散阈值 $T_3$ 的取值设置的稍大一些。

[0144] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,且 $V_{diff}$ 小于第三离散阈值 $T_3$ ,则所述第二视频类型为全景内容视频;

[0145] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,且 $V_{diff}$ 大于等于第三离散阈值 $T_3$ ,则所述第二视频类型为180度内容视频;

[0146] 若 $V_{\text{topPolar}}$ 大于等于第一离散阈值 $T_1$ 和/或 $V_{\text{diff}}$ 大于等于第三离散阈值 $T_3$ ,且 $V_{\text{diff}}$ 大于等于第三离散阈值 $T_3$ ,则所述第二视频类型为普通内容视频。

[0147] 所述第六装置6根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。

[0148] 具体地,所述第六装置6通过将所述第一视频类型以及所述第二视频类型进行组合,以确定所述待检测的视频的视频格式。

[0149] 由于所述第一视频类型包括3D类型或非3D类型,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频,进一步地,所述3D类型中包括上下3D类型以及左右3D类型,因此,最终确定的视频格式类型包括以下任一种:普通非3D视频、180度非3D视频、全景非3D视频、普通左右排列3D视频、180度左右排列3D视频、全景左右排列3D视频、普通上下排列3D视频、180度上下排列3D视频、全景上下排列3D视频。

[0150] 图2示出根据本申请的一个实施例的一种用于识别VR视频格式的方法流程图。

[0151] 具体地,在步骤S1中,所述识别设备获取待检测的视频中的至少一帧初始视频图像;在步骤S2中,所述识别设备对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像;在步骤S3中,所述识别设备根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型;在步骤S4中,所述识别设备根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域;在步骤S5中,所述识别设备根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频;在步骤S6中,所述识别设备根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。

[0152] 在步骤S1中,所述识别设备获取待检测的视频中的至少一帧初始视频图像。

[0153] 具体地,所述待检测的视频可以是任意需要检测的视频,优选地,所述待检测的视频是从VR视频播放设备中所获取的视频。所述待检测的视频可以从播放系统中获取的,也可以是用户自行上传的。

[0154] 然后,所述识别设备从所述待检测的视频中截取至少一帧初始视频图像,例如,所述识别设备按照预定的提取位置、提取时间等信息,从所述待检测的视频的该提取位置或该提取时间截取至少一帧初始视频图像;或者,所述识别设备也可以与其他提供初始视频图像的设备相交互,直接获取待检测的视频中的至少一帧初始视频图像。

[0155] 优选地,所述初始视频图像为所述待检测的视频中的关键帧。

[0156] 在步骤S2中,所述识别设备对所述初始视频图像进行预处理,以去除边缘干扰区域并获得处理后的视频图像。

[0157] 具体地,所述边缘干扰区域包括但不限于任意一种纯色边缘区域,如黑边区域、白边区域、红边区域等,所述边缘干扰区域中不存在图像变换。在步骤S2中,所述识别设备通过对所述初始视频图像进行积分处理、扫描像素点等方式,检测所述初始视频图像所对应的黑边区域,并将该边缘干扰区域进行剪裁,以去除所述边缘干扰区域,从而实现所述初始视频图像的预处理。

[0158] 优选地,在步骤S2中,所述识别设备将所述初始视频图像转换为灰度图;对所述灰

度图进行边缘检测,并对所述边缘检测的结果进行积分处理;根据所述积分处理结果,确定所述初始视频图像所对应的边缘干扰区域;去除所述边缘干扰区域并获得处理后的视频图像。

[0159] 具体地,在步骤S2中,所述识别设备根据现有的各类图像转换方式,将所述初始视频图像转换为灰度图;然后,对所述灰度图进行边缘检测,以将边缘响应较强的部分凸显出来,在此,所述边缘检测方法包括但不限于Canny、Sobel等。

[0160] 例如,图3示出根据本申请的一个实施例的从待检测的视频中所获取的一帧初始视频图像,该初始视频图像包括边缘干扰区域,即边缘的黑边部分。通过对所述初始视频图像进行边缘检测,则得到了图4所示的对该初始视频图像进行边缘检测后的灰度图。

[0161] 然后,对该灰度图进行积分处理,以生成积分图。如图5所示,图5即是对图4所示的灰度图进行积分后处理后的积分图。根据所述积分处理结果,能够确定该初始视频图像的图像变化信息,从而确定所述初始视频图像所对应的边缘干扰区域,也即图5所示的黑边区域;最后去除所述边缘干扰区域并获得处理后的视频图像。在此,图6示出对图3所示的初始视频图像进行预处理后的处理后的视频图像。

[0162] 其中,所述积分处理的过程如下所示:

[0163] 以I表示积分图,G表示灰度图,则 $I(x,y) = \text{sum}(G(i,j))$ ,其中, $0 < i < x, 0 < j < y$ 。在此,x,y,i,j表示坐标,I(x,y),以及G(i,j)表示该点像素值,该式的计算意义即为:通过图像累加来显示该图像的变化程度。

[0164] 以所述边缘干扰区域为黑边区域为例,在所述积分图中,黑色部分的数值为0,非黑色部分的数值大于0。通过图5所示的积分图可以看到,当横向扫描到第m列的时候,出现了大量的非零点,即白色像素点,说明从m列开始,图像出现了巨大的变化。这是因为原图存在一定列数的黑边,而当扫描到非黑边时,则原图的数值产生了变化,换言之,是原图的黑边导致了上述变化。因此,可以以m列为分割点,将原图的左侧黑边裁掉。

[0165] 由于黑边的存在往往是对称的,所以也可将右侧的m个像素黑边裁掉;或者,继续对所述图像进行横向扫描,当横向扫描到第m+k列时,出现了大量的零点,即黑色像素点,说明从m+k列开始,图像出现了第二次变化,从而将m+k列至右侧顶端的像素黑边裁掉。

[0166] 同理,纵向扫描积分图,在n行的时候出现了大量的非零点,即白色像素点,说明从n行开始原图出现了巨大变化,这个变化也是由于黑边导致的,因此裁掉原图上侧的n行。类似地,也可以裁掉原图对称的下侧n行,或者,继续进行扫描,根据扫描结果来裁掉下侧的黑边。

[0167] 更优选地,所述方法还包括步骤S7(未示出),其中,在步骤S7中,所述识别设备将所述初始视频图像缩放至预定大小;然后,在步骤S2中,所述识别设备对所述缩放后的初始视频图像进行处理。

[0168] 具体地,在步骤S7中,所述识别设备按照所述初始视频图像的宽高比,将所述初始视频图像进行等比例缩放,以缩放至预定大小;或者,在步骤S7中,所述识别设备按照预定的比例,将所述初始视频图像进行缩放,以缩放至预定大小;或者,在步骤S7中,所述识别设备按照预定的图像存储大小,对所述初始视频图像进行缩放,以缩放至预定大小。

[0169] 在此,所述预定大小可以由用户自行设置,也可以根据所述识别设备的处理能力确定。

[0170] 然后,在步骤S2中,所述识别设备对所述缩放后的初始视频图像进行处理,以实现快速处理。

[0171] 在步骤S3中,所述识别设备根据所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息,判断所述待检测的视频的第一视频类型,其中,所述第一视频类型包括3D类型或非3D类型。

[0172] 具体地,在步骤S3中,所述识别设备将所述处理后的视频图像分为上下两个图像和/或左右两个图像;然后,分别确定所述上下两个图像和/或左右两个图像的上下特征点和/或左右特征点,在此,所述确定方法包括但不限于计算BRIEF特征描述子或ORB特征描述子;接下来,计算所述上下特征点和/或左右特征点的匹配信息,例如,利用Hamming距离来判定所述上下特征点和/或左右特征点是否匹配。最后,基于所计算的匹配信息来判定所述待检测的视频的第一视频类型。

[0173] 在此,所述非3D类型包括2D类型。

[0174] 优选地,所述步骤S3包括步骤S31(未示出)以及步骤S32(未示出),其中,在步骤S31中,所述识别设备确定所述处理后的视频图像的上下部分和/或左右部分的特征点的匹配信息;若有任意一个所述匹配信息大于预定阈值,则在步骤S32中,所述识别设备判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

[0175] 为便于说明,以下以将该处理后的视频图像分为左右两个图像为例进行说明。

[0176] 具体地,在步骤S31中,所述识别设备首先将该处理后的视频图像分为左右两个图像,对所述左右两个图像分别检测角点,然后通过例如计算BRIEF特征描述子或ORB特征描述子的方式,计算左右两个图像的特征点。在此,根据3D视频本身的特性可知,左右内容差异是由一定的视差导致的,不存在特征旋转以及尺度变化的情况,因此优选地,可以采用速度较快的BRIEF。

[0177] 然后,在步骤S31中,所述识别设备使用如hamming距离来计算左右两组特征描述子的距离,若所述hamming距离小于某一阈值,即可表示对应的左右两个特征点匹配。在此,所匹配的特征点的个数即可作为所述处理后的视频图像的左右部分的特征点的匹配信息。

[0178] 图7示出一种判断处理后的视频图像的左右部分的特征点的匹配信息的示意图。图7示出了左右两部分的特征描述子,以及每对特征描述子的距离信息。

[0179] 若所述上下部分的特征点的匹配信息或所述左右部分的特征点的匹配信息的任意一个大于预定的阈值,则在步骤S32中,所述识别设备判断所述待检测的视频的第一视频类型为3D类型,反之,则为非3D类型。

[0180] 例如,继上例,若所述匹配的特征点的个数大于某一数量N时,则判定所述待检测的视频的第一视频类型为3D类型,进一步地,为左右3D类型。

[0181] 优选地,在步骤S32中,所述识别设备用于:

[0182] 若所述处理后的视频图像的上下部分的特征点的匹配信息大于第一特征阈值,则判断所述待检测的视频的第一视频类型为上下3D类型;和/或

[0183] 若所述处理后的视频图像的左右部分的特征点的匹配信息大于第二特征阈值,则判断所述待检测的视频的第一视频类型为左右3D类型;

[0184] 若所述处理后的视频图像的上下部分的特征点的匹配信息并未大于第一特征阈值,且所述处理后的视频图像的左右部分的特征点的匹配信息并未大于第二特征阈值,则

判断所述待检测的视频的第一视频类型为非3D类型。

[0185] 本领域技术人员应能理解,所述第一特征阈值可以等同于所述第二特征阈值,所述第一特征阈值也可以不等同于所述第二特征阈值。

[0186] 在步骤S4中,所述识别设备根据所述第一视频类型,确定所述处理后的视频图像所对应的处理区域。

[0187] 在此,所述处理区域即为ROI (Region of Interest),即对该视频图像的后续处理区域。

[0188] 具体地,若所述第一视频类型是非3D类型,则可以将所述处理后的视频图像的整个图像直接作为处理区域,以进行后续处理;

[0189] 若所述第一视频类型为左右3D类型,则可截取所述处理后的视频图像的左半部分或右半部分作为所述处理区域,以进行后续处理;

[0190] 若所述第一视频类型为上下3D类型,则可截取所述处理后的视频图像的上半部分或下半部分作为所述处理区域,以进行后续处理。

[0191] 在步骤S5中,所述识别设备根据所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息,判断所述待检测的视频的第二视频类型,其中,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频。

[0192] 在此,所述全景内容表示Equirectangular模式的投影。图8示出一种全景内容的示意图,其表示了从地球仪到世界地图的映射关系。如图8可知,全景图的首行(第一行)是由球面的上极点展开的,尾行(最后一行)是由球面的下极点展开的。因此,全景图的首行像素值应该为同一值,尾行像素值也为同一值;可选地,由于在展开过程中存在插值情况,因此,首行像素值与尾行像素值可以存在一定偏差。此外,由全景图的展开方式可知,全景图左右两侧能够无缝拼接在一起。

[0193] 因此,在步骤S5中,所述识别设备分别计算所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;在此,优选地,所述离散程度信息包括方差或每个样本值与全体样本值的平均数之差的和,也即,可以采用方差来表示所述离散程度信息,也可以采用每个样本值与全体样本值的平均数之差的和来表示所述离散程度信息。

[0194] 优选地,在步骤S5中,所述识别设备用于:

[0195] 确定所述处理区域中首行像素值的离散程度信息、尾行像素值的离散程度信息、首列以及尾列所对应的像素值的离散程度信息;

[0196] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息小于第三离散阈值,则所述第二视频类型为全景内容视频;

[0197] 若所述首行像素值的离散程度信息小于第一离散阈值、所述尾行像素值离散程度信息小于第二离散阈值且所述首列以及尾列所对应的像素值的离散程度信息大于等于第三离散阈值,则所述第二视频类型为180度内容视频;

[0198] 若所述首行像素值的离散程度信息大于等于第一离散阈值和/或所述尾行像素值离散程度信息大于等于第二离散阈值,且所述首列以及尾列所对应的像素值的离散程度信

息大于等于第三离散阈值,则所述第二视频类型为普通内容视频。

[0199] 例如,假设所述处理区域的宽为 $w$ ,高为 $h$ ,则首行中每个像素的像素值可以用 $P(0, j)$ 表示,其中, $i$ 的取值范围为 $[0, w-1]$ ,尾行中每个像素的像素值可以用 $P(h-1, j)$ 表示,其中 $j$ 的取值范围为 $[0, w-1]$ 。同理,首列中每个像素的像素值可以用 $P(m, 0)$ 表示,其中 $m$ 的取值范围为 $[0, h-1]$ ,尾列中每个像素的像素值可以用 $P(n, w-1)$ ,其中 $n$ 的取值范围为 $[0, h-1]$ 。

[0200] 则以采用每个样本值与全体样本值的平均数之差的和来表示所述离散程度信息为例:

[0201] 首行像素值的离散程度信息 $V_{topPolar}$ 为:
$$V_{topPolar} = \sum_{i=0}^{w-1} (P(0, i) - \sum_{i=0}^{w-1} (P(0, i) / w))$$

[0202] 尾行像素值的离散程度信息 $V_{bottomPolar}$ 为:

$$V_{bottomPolar} = \sum_{j=0}^{w-1} (P(h-1, j) - \sum_{j=0}^{w-1} (P(h-1, j) / w))$$

[0203] 首列以及尾列所对应的像素值的离散程度信息 $V_{diff}$ 为:

$$V_{diff} = \sum_{m=0}^h |(P(m, 0) - P(m, w-1))| / h$$

[0204] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,则可认为该图的上下两边是由极点展开的;若 $V_{diff}$ 小于第三离散阈值 $T_3$ ,则可认为该图的左右两边可以无缝拼接在一起的。在此,所述第一离散阈值 $T_1$ 、第二离散阈值 $T_2$ 、第三离散阈值 $T_3$ 可以根据图像由球面展开成柱面时的插值操作来进行取值,例如,若插值较多,则可将所述第一离散阈值 $T_1$ 、第二离散阈值 $T_2$ 、第三离散阈值 $T_3$ 的取值设置的稍大一些。

[0205] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,且 $V_{diff}$ 小于第三离散阈值 $T_3$ ,则所述第二视频类型为全景内容视频;

[0206] 若 $V_{topPolar}$ 小于第一离散阈值 $T_1$ 、 $V_{bottomPolar}$ 小于第二离散阈值 $T_2$ ,且 $V_{diff}$ 大于等于第三离散阈值 $T_3$ ,则所述第二视频类型为180度内容视频;

[0207] 若 $V_{topPolar}$ 大于等于第一离散阈值 $T_1$ 和/或 $V_{diff}$ 大于等于第三离散阈值 $T_3$ ,且 $V_{diff}$ 大于等于第三离散阈值 $T_3$ ,则所述第二视频类型为普通内容视频。

[0208] 在步骤S6中,所述识别设备根据所述第一视频类型以及所述第二视频类型,确定所述待检测的视频的视频格式。

[0209] 具体地,在步骤S6中,所述识别设备通过将所述第一视频类型以及所述第二视频类型进行组合,以确定所述待检测的视频的视频格式。

[0210] 由于所述第一视频类型包括3D类型或非3D类型,所述第二视频类型包括普通内容视频、180度内容视频或全景内容视频,进一步地,所述3D类型中包括上下3D类型以及左右3D类型,因此,最终确定的视频格式类型包括以下任一种:普通非3D视频、180度非3D视频、全景非3D视频、普通左右排列3D视频、180度左右排列3D视频、全景左右排列3D视频、普通上下排列3D视频、180度上下排列3D视频、全景上下排列3D视频。

[0211] 图9示出了可被用于实施本申请中所述的各个实施例的示例性系统。

[0212] 在一些实施例中,系统900能够作为图1至图8所示的实施例或其他所述实施例中

的任意一个远程计算设备。在一些实施例中，系统900可包括具有指令的一个或多个计算机可读介质（例如，系统存储器或NVM/存储设备920）以及与该一个或多个计算机可读介质耦合并被配置为执行指令以实现模块从而执行本申请中所述的动作的一个或多个处理器（例如，（一个或多个）处理器905）。

[0213] 对于一个实施例，系统控制模块910可包括任意适当的接口控制器，以向（一个或多个）处理器905中的至少一个和/或与系统控制模块910通信的任意适当的设备或组件提供任意适当的接口。

[0214] 系统控制模块910可包括存储器控制器模块930，以向系统存储器915提供接口。存储器控制器模块930可以是硬件模块、软件模块和/或固件模块。

[0215] 系统存储器915可被用于例如为系统900加载和存储数据和/或指令。对于一个实施例，系统存储器915可包括任意适当的易失性存储器，例如，适当的DRAM。在一些实施例中，系统存储器915可包括双倍数据速率类型四同步动态随机存取存储器（DDR4SDRAM）。

[0216] 对于一个实施例，系统控制模块910可包括一个或多个输入/输出（I/O）控制器，以向NVM/存储设备920及（一个或多个）通信接口925提供接口。

[0217] 例如，NVM/存储设备920可被用于存储数据和/或指令。NVM/存储设备920可包括任意适当的非易失性存储器（例如，闪存）和/或可包括任意适当的（一个或多个）非易失性存储设备（例如，一个或多个硬盘驱动器（HDD）、一个或多个光盘（CD）驱动器和/或一个或多个数字通用光盘（DVD）驱动器）。

[0218] NVM/存储设备920可包括在物理上作为系统900被安装在其上的的设备的一部分的存储资源，或者其可被该设备访问而不必作为该设备的一部分。例如，NVM/存储设备920可通过网络经由（一个或多个）通信接口925进行访问。

[0219] （一个或多个）通信接口925可为系统900提供接口以通过一个或多个网络和/或与任意其他适当的设备通信。系统900可根据一个或多个无线网络标准和/或协议中的任意标准和/或协议来与无线网络的一个或多个组件进行无线通信。

[0220] 对于一个实施例，（一个或多个）处理器905中的至少一个可与系统控制模块910的一个或多个控制器（例如，存储器控制器模块930）的逻辑封装在一起。对于一个实施例，（一个或多个）处理器905中的至少一个可与系统控制模块910的一个或多个控制器的逻辑封装在一起以形成系统级封装（SiP）。对于一个实施例，（一个或多个）处理器905中的至少一个可与系统控制模块910的一个或多个控制器的逻辑集成在同一模具上。对于一个实施例，（一个或多个）处理器905中的至少一个可与系统控制模块910的一个或多个控制器的逻辑集成在同一模具上以形成片上系统（SoC）。

[0221] 在各个实施例中，系统900可以但不限于是：服务器、工作站、台式计算设备或移动计算设备（例如，膝上型计算设备、手持计算设备、平板电脑、上网本等）。在各个实施例中，系统900可具有更多或更少的组件和/或不同的架构。例如，在一些实施例中，系统900包括一个或多个摄像机、键盘、液晶显示器（LCD）屏幕（包括触屏显示器）、非易失性存储器端口、多个天线、图形芯片、专用集成电路（ASIC）和扬声器。

[0222] 显然，本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样，倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内，则本申请也意图包含这些改动和变型在内。

[0223] 需要注意的是,本发明可在软件和/或软件与硬件的组合体中被实施,例如,可采用专用集成电路(ASIC)、通用目的计算机或任何其他类似硬件设备来实现。在一个实施例中,本发明的软件程序可以通过处理器执行以实现上文所述步骤或功能。同样地,本发明的软件程序(包括相关的数据结构)可以被存储到计算机可读记录介质中,例如,RAM存储器,磁或光驱动器或软磁盘及类似设备。另外,本发明的一些步骤或功能可采用硬件来实现,例如,作为与处理器配合从而执行各个步骤或功能的电路。

[0224] 另外,本申请的一部分可被应用为计算机程序产品,例如计算机程序指令,当其被计算机执行时,通过该计算机的操作,可以调用或提供根据本申请的方法和/或技术方案。本领域技术人员应能理解,计算机程序指令在计算机可读介质中的存在形式包括但不限于源文件、可执行文件、安装包文件等,相应地,计算机程序指令被计算机执行的方式包括但不限于:该计算机直接执行该指令,或者该计算机编译该指令后再执行对应的编译后程序,或者该计算机读取并执行该指令,或者该计算机读取并安装该指令后再执行对应的安装后程序。在此,计算机可读介质可以是可供计算机访问的任意可用的计算机可读存储介质或通信介质。

[0225] 通信介质包括藉此包含例如计算机可读指令、数据结构、程序模块或其他数据的通信信号被从一个系统传送到另一系统的介质。通信介质可包括有导的传输介质(诸如电缆和线(例如,光纤、同轴等))和能传播能量波的无线(未有导的传输)介质,诸如声音、电磁、RF、微波和红外。计算机可读指令、数据结构、程序模块或其他数据可被体现为例如无线介质(诸如载波或诸如被体现为扩展频谱技术的一部分的类似机制)中的已调制数据信号。术语“已调制数据信号”指的是其一个或多个特征以在信号中编码信息的方式被更改或设定的信号。调制可以是模拟的、数字的或混合调制技术。

[0226] 作为示例而非限制,计算机可读存储介质可包括以用于存储诸如计算机可读指令、数据结构、程序模块或其它数据的信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动的介质。例如,计算机可读存储介质包括,但不限于,易失性存储器,诸如随机存储器(RAM, DRAM, SRAM);以及非易失性存储器,诸如闪存、各种只读存储器(ROM, PROM, EPROM, EEPROM)、磁性和铁磁/铁电存储器(MRAM, FeRAM);以及磁性和光学存储设备(硬盘、磁带、CD、DVD);或其它现在已知的介质或今后开发的能够存储供计算机系统使用的计算机可读信息/数据。

[0227] 在此,根据本申请的一个实施例包括一个装置,该装置包括用于存储计算机程序指令的存储器和用于执行程序指令的处理器,其中,当该计算机程序指令被该处理器执行时,触发该装置运行基于前述根据本申请的多个实施例的方法和/或技术方案。

[0228] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。装置权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一,第二等词语用来表示名称,而并不表示任何特定的顺序。

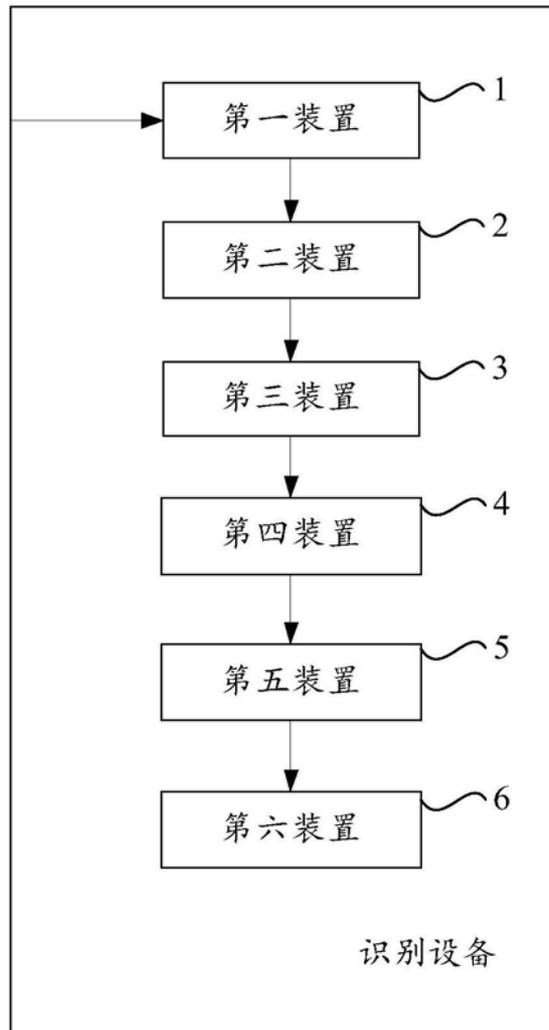


图1

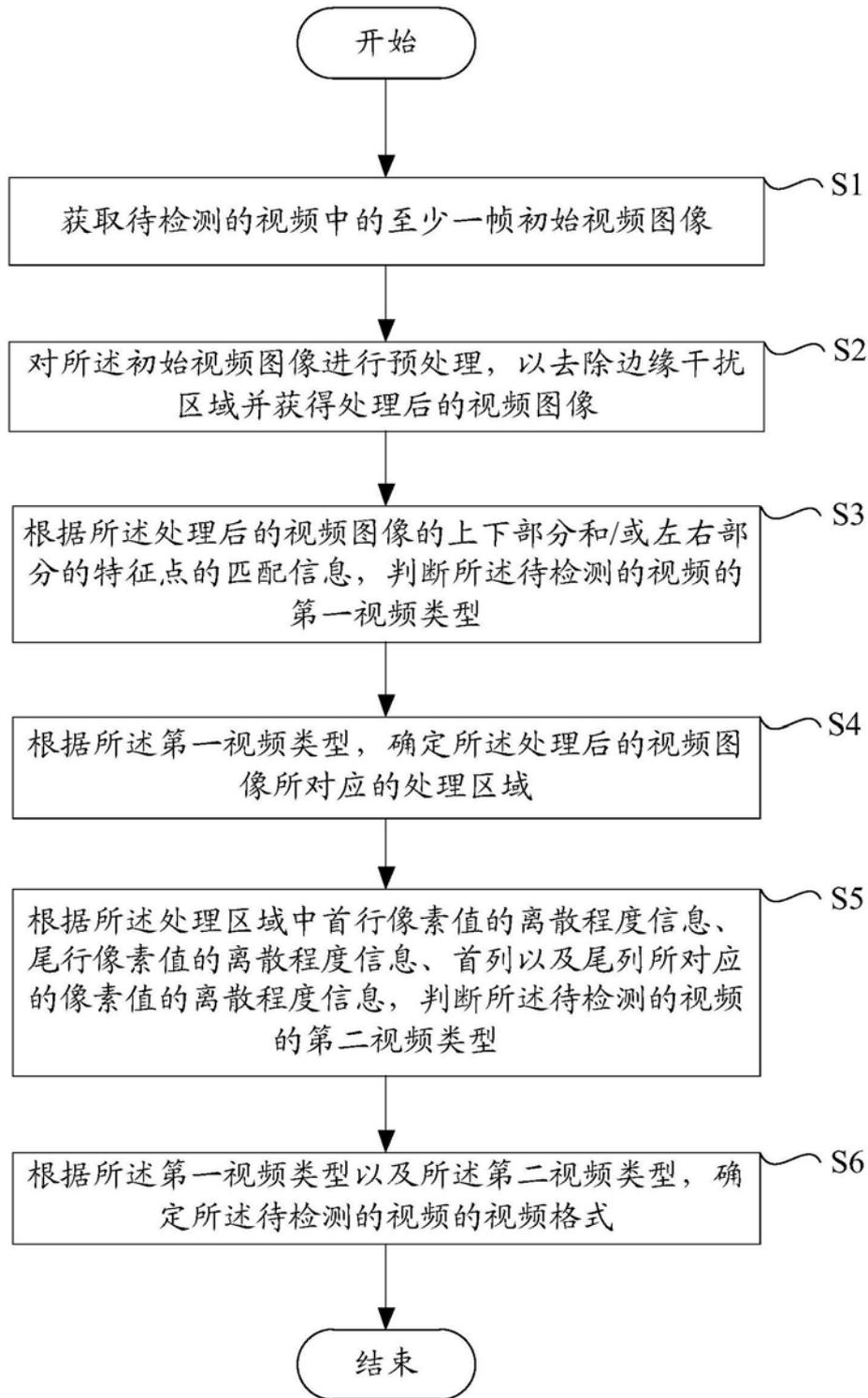


图2



图3

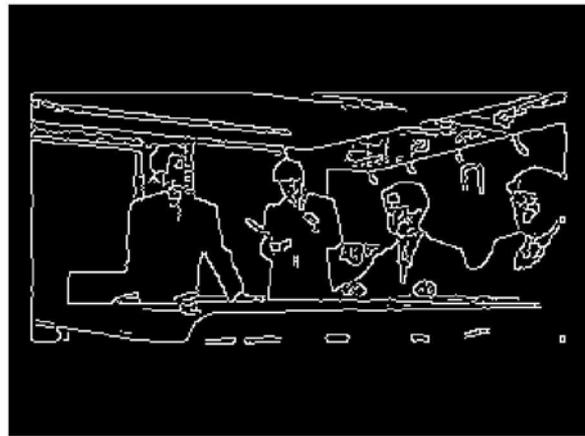


图4



图5



图6

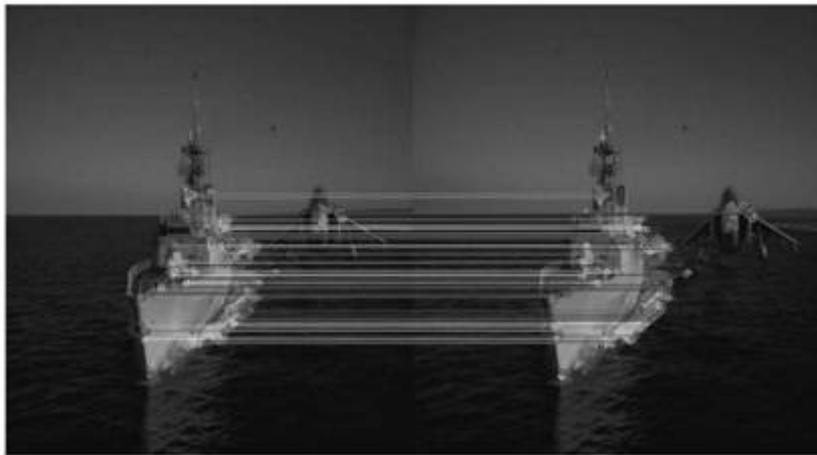


图7

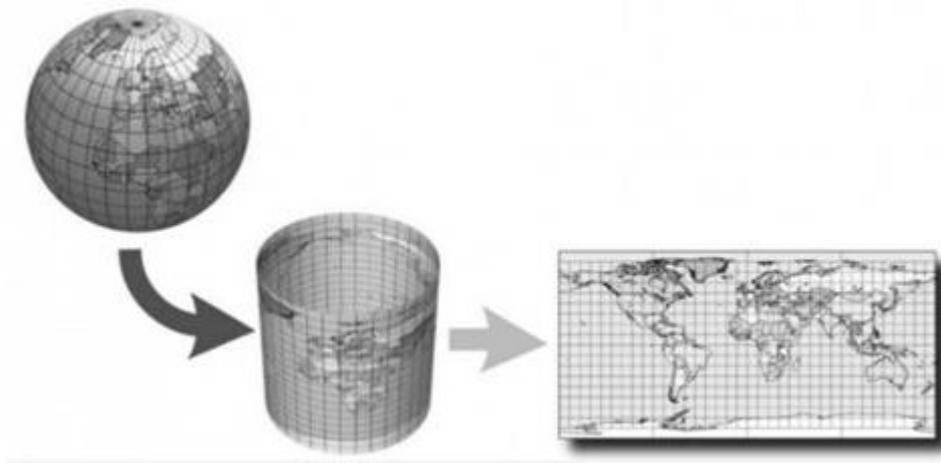


图8

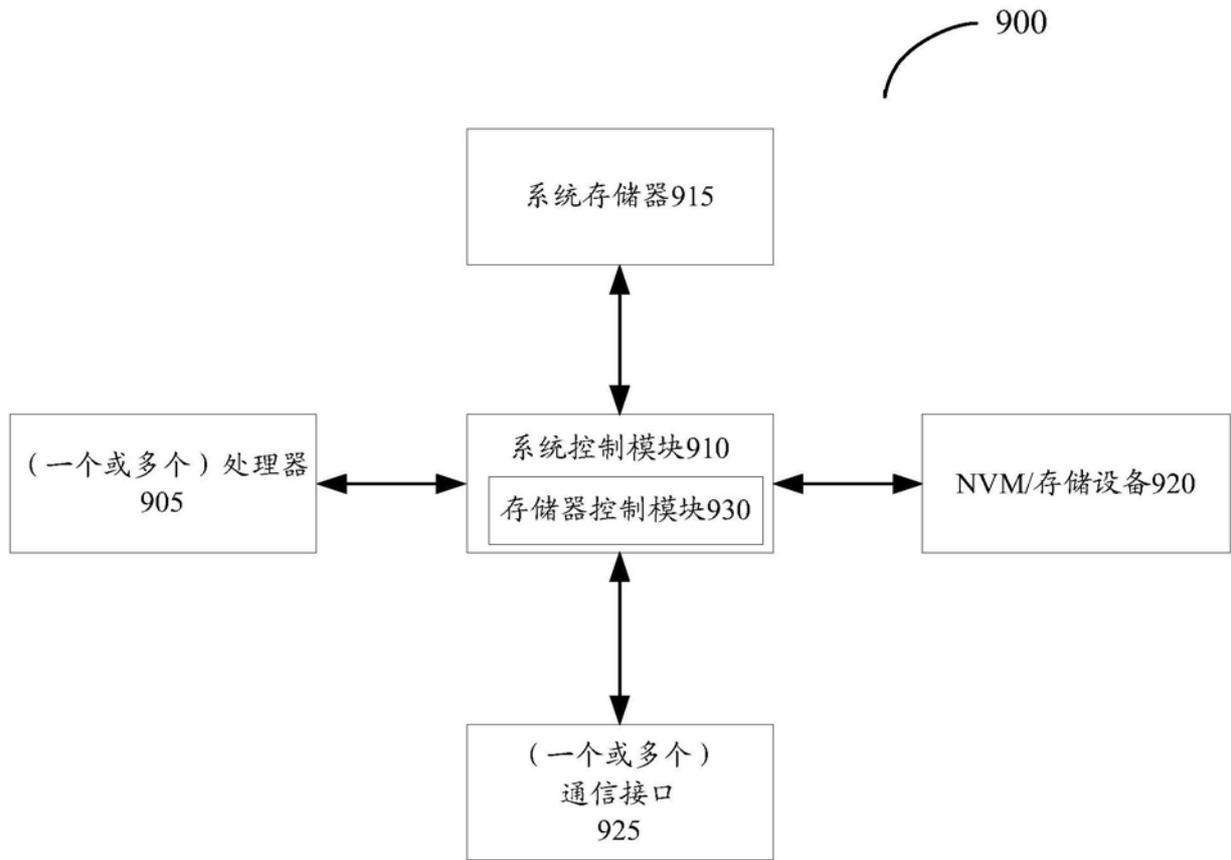


图9