

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2020年12月17日 (17.12.2020)



(10) 国际公布号  
**WO 2020/248900 A1**

- (51) 国际专利分类号:  
*H04N 21/234* (2011.01)
- (21) 国际申请号: PCT/CN2020/094507
- (22) 国际申请日: 2020年6月5日 (05.06.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201910497871.1 2019年6月10日 (10.06.2019) CN
- (71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街30号院3号楼2层B-0035房间, Beijing 100041 (CN)。
- (72) 发明人: 白鹤群 (BAI, Hequn); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。
- (74) 代理人: 泰和泰律师事务所 (TAHOTA LAW FIRM); 中国北京市朝阳区东四环中路56号远洋国际中心A座12层, Beijing 100025 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(54) Title: PANORAMIC VIDEO PROCESSING METHOD AND APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 全景视频的处理方法、装置及存储介质

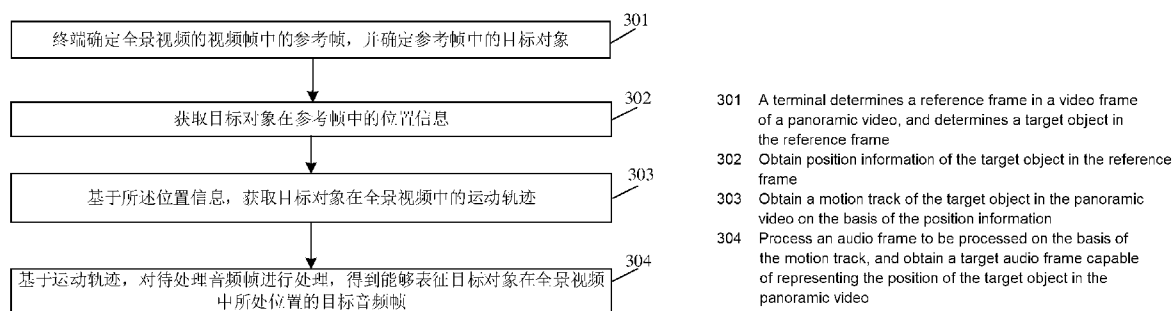


图 3

(57) Abstract: Embodiments of the present invention provide a panoramic video processing method and apparatus, and a storage medium. The method comprises: determining a reference frame in a video frame of a panoramic video, and determining a target object in the reference frame; obtaining position information of the target object in the reference frame; obtaining a motion track of the target object in the panoramic video on the basis of the position information; and processing an audio frame to be processed on the basis of the motion track, and obtaining a target audio frame capable of representing the position of the target object in the panoramic video, the target audio frame being adapted to be combined with the video frame of the panoramic video to obtain a panoramic video file.

(57) 摘要: 本公开实施例提供一种全景视频的处理方法、装置及存储介质; 其中, 方法包括: 确定全景视频的视频帧中的参考帧, 并确定所述参考帧中的目标对象; 获取所述目标对象在所述参考帧中的位置信息; 基于所述位置信息, 获取所述目标对象在所述全景视频中的运动轨迹; 基于所述运动轨迹, 对待处理音频帧进行处理, 得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧; 所述目标音频帧, 用于与所述全景视频的视频帧合成得到全景视频文件。

WO 2020/248900 A1

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

## 全景视频的处理方法、装置及存储介质

### 相关申请的交叉引用

本申请要求于 2019 年 06 月 10 日提交的，申请号为 201910497871.1、发明  
5 名称为“全景视频的处理方法、装置及存储介质”的中国专利申请的优先权，  
该申请的全文通过引用结合在本申请中。

### 技术领域

本公开实施例涉及多媒体技术，尤其涉及一种全景视频的处理方法、装置  
及存储介质。

### 10 背景技术

通信基础设施的不断发展，使得终端网络通信的速率不断提升，从而各种  
媒体形式特别是视频在网络中得到前所未有地广泛传播。全景视频是对用户感  
知视频方式的拓展，视频中的内容的展现能够呈现丰富多样的方式。

为了更好的展现全景视频的三维播放效果，全景视频中的音频数据需要体  
15 现相应声源在全景视频中的位置，相关技术中，对全景视频的音频数据的处理  
依赖人工手动处理，处理效率低。

### 发明内容

有鉴于此，本公开实施例提供一种全景视频的处理方法、装置及存储介质。

本公开实施例提供了一种全景视频的处理方法，所述方法包括：

20 确定全景视频的视频帧中的参考帧，并确定所述参考帧中的目标对象；

获取所述目标对象在所述参考帧中的位置信息；

基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹；

基于所述运动轨迹，对待处理音频帧进行处理，得到能够表征所述目标对

象在所述全景视频中所处位置的目标音频帧；

所述目标音频帧，用于与所述全景视频的视频帧合成得到全景视频文件。

上述方案中，所述获取所述目标对象在所述参考帧中的位置信息，包括：

识别所述参考帧中所述目标对象的多个特征点，得到对应所述目标对象的

5 特征点集合；

基于所述特征点集中的特征点，获取所述目标对象在所述参考帧中的位置信息，所述位置信息包括：

所述目标对象在所述全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考  
10 帧对应的三维模型的中心的距离。

上述方案中，所述基于所述特征点集中的特征点，获取所述目标对象在所述参考帧中的位置信息，包括：

确定所述特征点集中多个特征点的重心位置；

基于所述重心位置，确定所述目标对象在所述全景视频中所处位置对应的  
15 角度信息；

获取所述多个特征点在所述参考帧中占据的区域的像素面积；

基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心  
距离。

上述方案中，所述基于所述重心位置，确定所述目标对象在所述全景视频  
20 中所处位置对应的角度信息，包括：

确定所述重心位置对应的像素点的像素点坐标；

将所述像素点坐标转换成对应所述三维模型的球面坐标，所述球面坐标指示所述目标对象在所述全景视频中所处位置对应的角度。

上述方案中，所述基于所述像素面积，确定所述目标对象在所述全景视频  
25 中所处位置的中心距离，包括：

根据所述像素面积与所述中心距离的平方所呈的反比例关系，确定所述目标对象在所述全景视频中所处位置的中心距离。

上述方案中，所述方法还包括：

将所述特征点集合中的多个特征点，与所述全景视频的非参考帧中的特征点进行特征点匹配，得到特征点匹配结果；

基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置。

5 上述方案中，所述基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置，包括：

基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度大于第一相似度阈值的特征点；

10 基于选取的所述大于第一相似度阈值的特征点，确定所述非参考帧中所述目标对象的位置。

上述方案中，所述基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹，包括：

获取所述全景视频的每个非参考帧中所述目标对象的位置信息；

15 基于所述目标对象在所述参考帧中的位置信息，以及每个非参考帧中所述目标对象的位置信息，确定所述目标对象在所述全景视频中的运动轨迹。

上述方案中，所述方法还包括：

基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度小于第二相似度阈值的特征点；

20 基于选取的所述小于第二相似度阈值的特征点，更新所述特征点集合中的特征点。

上述方案中，所述方法还包括：

获取预设数量的非参考帧，与所述特征点集合中特征点的匹配结果；所述预设数量的非参考帧为所述全景视频的连续视频帧；

25 基于获取的特征点匹配结果，确定所述特征点集合中的待删除特征点，所述待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；

将所述待删除特征点从所述特征点集合中删除。

上述方案中，所述确定所述参考帧中的目标对象，包括：

接收到针对所述参考帧的对象选取操作；

基于所述对象选取操作在所述参考帧中所选定的区域，确定所述参考帧中的目标对象。

上述方案中，所述基于所述运动轨迹，对待处理音频帧进行处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧，包括：

调用三维音频引擎，对待处理音频帧根据所述运动轨迹进行混音处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧。

本公开实施例还提供了一种全景视频的处理装置，所述装置包括：

确定单元，用于确定全景视频的视频帧中的参考帧，并确定所述参考帧中的目标对象；

获取单元，用于获取所述目标对象在所述参考帧中的位置信息；

以及，用于基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹；

处理单元，用于基于所述运动轨迹，对待处理音频帧进行处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧；

所述目标音频帧，用于与所述全景视频的视频帧合成得到全景视频文件。

上述方案中，所述获取单元，还用于识别所述参考帧中所述目标对象的多个特征点，得到对应所述目标对象的特征点集合；

基于所述特征点集合中的特征点，获取所述目标对象在所述参考帧中的位置信息，所述位置信息包括：

所述目标对象在所述全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考帧对应的三维模型的中心的距离。

上述方案中，所述获取单元，还用于确定所述特征点集合中多个特征点的重心位置；

基于所述重心位置，确定所述目标对象在所述全景视频中所处位置对应的角度信息；

获取所述多个特征点在所述参考帧中占据的区域的像素面积；

基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心距离。

上述方案中，所述获取单元，还用于确定所述重心位置对应的像素点的像素点坐标；

将所述像素点坐标转换成对应所述三维模型的球面坐标，所述球面坐标指示所述目标对象在所述全景视频中所处位置对应的角度。

上述方案中，所述获取单元，还用于根据所述像素面积与所述中心距离的平方所呈的反比例关系，确定所述目标对象在所述全景视频中所处位置的中心距离。

上述方案中，所述装置还包括匹配单元；

所述匹配单元，用于将所述特征点集合中的多个特征点，与所述全景视频的非参考帧中的特征点进行特征点匹配，得到特征点匹配结果；

所述获取单元，还用于基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置。

上述方案中，所述获取单元，还用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度大于第一相似度阈值的特征点；

基于选取的所述大于第一相似度阈值的特征点，确定所述非参考帧中所述目标对象的位置。

上述方案中，所述获取单元，还用于获取所述全景视频的每个非参考帧中所述目标对象的位置信息；

基于所述目标对象在所述参考帧中的位置信息，以及每个非参考帧中所述目标对象的位置信息，确定所述目标对象在所述全景视频中的运动轨迹。

上述方案中，所述装置还包括：

更新单元，用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度小于第二相似度阈值的特征点；

基于选取的所述小于第二相似度阈值的特征点，更新所述特征点集合中的特征点。

上述方案中，所述装置还包括：

删除单元，用于获取预设数量的非参考帧，与所述特征点集合中特征点的  
5 匹配结果；所述预设数量的非参考帧为所述全景视频的连续视频帧；

基于获取的特征点匹配结果，确定所述特征点集合中的待删除特征点，所述待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；

将所述待删除特征点从所述特征点集合中删除。

上述方案中，所述确定单元，还用于接收到针对所述参考帧的对象选取操  
10 作；

基于所述对象选取操作在所述参考帧中所选定的区域，确定所述参考帧中的目标对象。

上述方案中，所述处理单元，还用于调用三维音频引擎，对待处理音频帧根据所述运动轨迹进行混音处理，得到能够表征所述目标对象在所述全景视频  
15 中所处位置的目标音频帧。

本公开实施例还提供了一种全景视频的处理装置，包括：

存储器，用于存储可执行指令；

处理器，用于执行所述可执行指令时，实现本公开实施例提供的全景视频  
的处理方法。

20 本公开实施例还提供了一种非暂态存储介质，存储有可执行指令，所述可执行指令被执行时，用于实现本公开实施例提供的全景视频的处理方法。

本公开实施例具有如下有益效果：

应用本公开上述实施例，自动获取目标对象在全景视频中的运动轨迹，并根据运动轨迹，自动实现对待处理音频帧的处理，进而得到能够表征目标对象  
25 在全景视频中所处位置的目标音频帧，提高了对全景视频的音频数据的处理效率。



## 附图说明

图 1 是本公开实施例提供的全景视频的处理系统 100 的架构示意图；

图 2 是本公开实施例提供的全景视频的处理装置的组成结构示意图；

图 3 是本公开实施例提供的全景视频的处理方法的流程示意图；

5 图 4 是本公开实施例提供的确定参考帧中的目标对象的示意图；

图 5 是本公开实施例提供的全景视频的三维模型到二维图像的映射关系示意图；

图 6 是本公开实施例提供的全景视频对应的球坐标系的示意图；

图 7 是本公开实施例提供的全景视频的三维模型投影示意图；

10 图 8 是本公开实施例提供的全景视频中对象的中心距离的示意图；

图 9 是本公开实施例提供的全景视频的处理方法的流程示意图；

图 10 是本公开实施例提供的全景视频的处理装置的组成结构示意图。

## 具体实施方式

为了使本公开的目的、技术方案和优点更加清楚，下面将结合附图对本公开  
15 作进一步地详细描述，所描述的实施例不应视为对本公开的限制，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本公开保护的范  
围。

在以下的描述中，涉及到“一些实施例”，其描述了所有可能实施例的子集，  
但是可以理解，“一些实施例”可以是所有可能实施例的相同子集或不同子集，  
20 并且可以在不冲突的情况下相互结合。

在以下的描述中，所涉及的术语“第一\第二\第三”仅仅是是区别类似的对象，不代表针对对象的特定排序，可以理解地，“第一\第二\第三”在允许的情况下可以互换特定的顺序或先后次序，以使这里描述的本公开实施例能够以除了在这里图示或描述的以外的顺序实施。

25 对本公开实施例进行进一步详细说明之前，对本公开实施例中涉及的名词和术语进行说明，本公开实施例中涉及的名词和术语适用于如下的解释。

1) 全景视频, 是指在各种电子设备中播放能够供用户裸眼观看的视频, 且播放窗口的视场能够进行方向和放大倍数的调整。

2) 三维模型, 是模拟全景视频所表达的空间的模型, 用于将全景视频的视频帧映射到三维模型的表面以形成表面纹理, 三维模型的表面一般采用球面或圆柱面。

3) 视场, 在三维模型的中心的虚拟镜头感知三维模型的表面纹理的视线的集合, 更一般地, 是指通过一个窗口在全景视频的视频帧中能够观看的区域。

4) 播放窗口, 全景视频的客户端中默认用于播放全景视频的全屏或非全屏的窗口, 至少包括全景视频的视频播放区域, 还可以包括设置有相关操作的入口的操作区域。

5) 播放窗口视场, 即与播放窗口对应的视场, 控制在播放窗口中能够感知到视频帧中处于相应视场的部分内容。

6) 特征点, 在图像中能够反映对象的局部特征(如颜色特征、形状特征和纹理特征)的点, 一般为多个像素点的集合, 以人脸图像为例, 特征点可以是眼睛特征点、嘴巴特征点或鼻子特征点等。

基于上述对本公开实施例中涉及的名词和术语的解释, 接下来对本公开实施例提供的全景视频的处理装置进行说明。本公开实施例的全景视频的处理装置可以以各种形式来实施, 如: 由智能手机、平板电脑和台式机等终端单独实施, 或者由服务器单独实施, 或者由终端、服务器协同实施。

作为由终端、服务器协同实施本公开实施例的全景视频的处理装置的示例, 参见图 1, 图 1 是本公开实施例提供的全景视频的处理系统 100 的一个可选的架构示意图, 为实现支撑一个示例性应用, 终端 400 (包括终端 400-1 (包括图形界面 410-1) 和终端 400-2 (包括图形界面 410-2)), 终端 400 通过网络 300 连接服务器 200, 终端上设置有全景视频客户端, 用户可通过该全景视频客户端实现对全景视频的显示, 网络 300 可以是广域网或者局域网, 又或者是二者的组合, 使用无线链路实现数据传输。

终端 400 (如终端 400-1), 用于加载全景视频数据, 进行全景视频的视频帧显示, 基于显示的视频帧, 确定全景视频的视频帧中的参考帧, 并确定参考帧中的目标对象, 发送参考帧标识及目标对象的标识给服务器 200;

服务器 200, 存储有全景视频数据, 用于基于参考帧标识及目标对象的标识, 获取目标对象在参考帧中的位置信息; 基于位置信息, 获取目标对象在全景视频中的运动轨迹; 基于运动轨迹, 对待处理音频帧进行处理, 得到能够表征目标对象在全景视频中所处位置的目标音频帧;

服务器 200, 还用于将得到的目标音频帧与全景视频的视频帧合成, 得到全景视频文件, 以及发送合成得到的全景视频文件给终端 400;

终端 400, 还用于基于接收到的全景视频文件, 进行全景视频播放。

本公开实施例提供的全景视频的处理装置可以实施为硬件或者软硬件结合的方式, 下面说明本公开实施例提供的全景视频的处理装置的各种示例性实施。

对本公开实施例的全景视频的处理装置的硬件结构做详细说明, 图 2 为本公开实施例提供的全景视频的处理装置的组成结构示意图, 图 2 示出的装置组成仅仅是一个示例, 不应对本公开实施例的功能和使用范围带来任何限制。

如图 2 所示, 全景视频的处理装置可以包括处理装置 (例如中央处理器、图形处理器等) 210, 其可以根据存储在只读存储器 (ROM, Read-Only Memory) 220 中的程序或者从存储装置 280 加载到随机访问存储器 (RAM, Random Access Memory) 230 中的程序而执行各种适当的动作和处理。在 RAM 230 中, 还存储有全景视频的处理装置操作所需的各种程序和数据。处理装置 210、ROM 220 以及 RAM 230 通过总线 240 彼此相连。输入/输出 (I/O, Input/Output) 接口 250 也连接至总线 240。

通常, 以下装置可以连接至 I/O 接口 250: 包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 260; 包括例如液晶显示器 (LCD)、扬声器、振动器等的输出装置 270; 包括例如磁带、硬盘等的存储装置 280; 以及通信装置 290。通信装置 290 可以允许全景视频的处理装置

与其他设备进行无线或有线通信以交换数据。虽然图 2 示出了具有各种装置的全景视频的处理装置,但是应理解的是,并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地,根据本公开的实施例,所提供的流程图描述的过程可以被实现为  
5 计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,计算机程序可以通过通信装置 290 从网络上被下载和安装,或者从存储装置 280 被安装,或者从 ROM220 被安装。在计算机程序被处理装置 210 执行时,执行本公开实施例的方法中限定的上述功  
10 能。

需要说明的是,本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以包括但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不  
15 限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (ROM)、可擦式可编程只读存储器 (EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

在本公开实施例中,计算机可读存储介质可以是任何包含或存储程序的有  
20 形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开实施例中,计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读  
25 介质,计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、射频 (RF) 等等,

或者上述的任意合适的组合。

上述计算机可读介质可以是上述全景视频的处理装置中所包含的；也可以是单独存在，而未装配入全景视频的处理装置中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序  
5 被该全景视频的处理装置执行时，使得全景视频的处理装置执行本公开实施例提供的上述全景视频的处理方法。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的  
计算机程序代码，上述程序设计语言包括面向对象的程序设计语言—诸如 Java、  
Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的  
10 程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络，包括局域网(LAN)和广域网(WAN)，以连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。  
15

描述于本公开实施例中涉及到的单元和/或模块可以通过软件的方式实现，也可以通过硬件的方式来实现。

对于硬件的方式来说，实现本公开实施例的全景视频的处理装置的单元和/或模块可以被一个或多个应用专用集成电路（ASIC，Application Specific  
20 Integrated Circuit）、DSP、可编程逻辑器件（PLD，Programmable Logic Device）、复杂可编程逻辑器件（CPLD，Complex Programmable Logic Device）、现场可编程门阵列（FPGA，Field-Programmable Gate Array）或其他电子元件实现，用于执行实现本公开实施例提供的方法。

25 接下来对本公开实施例提供的全景视频的处理方法进行说明，在实际实施时，本公开实施例的全景视频的处理方法可由终端实施，或由终端及服务器协同实施，接下来以终端实施为例进行说明。图3为本公开实施例提供的全景视

频的处理方法的流程示意图，参见图 3，本公开实施例的全景视频的处理方法包括：

步骤 301：终端确定全景视频的视频帧中的参考帧，并确定参考帧中的目标对象。

5 在实际实施时，终端解码全景视频文件得到多个连续的视频帧，显示解码得到的视频帧，用户基于终端显示的视频帧进行参考帧的选取，在实际应用中，通常选取多个连续的视频帧中出现目标对象的第一个视频帧作为参考帧，相应的，终端接收到用户触发的对象选取操作，确定对象选取操作所针对的视频帧作为参考帧，相应的，全景视频中除参考帧以外的视频帧均为非参考帧。

10 在一些实施例中，终端可通过如下方式确定参考帧中的目标对象：接收到针对参考帧的对象选取操作，基于对象选取操作在参考帧中所选定的区域，确定参考帧中的目标对象。在实际应用中，对象选取操作既可以为用户的触控操作，也可以为用户通过鼠标触发的指针选取操作；图 4 为本公开实施例提供的确定参考帧中的目标对象的示意图，如图 4 所示，目标对象为全景视频中某  
15 个特定的人，用户可通过鼠标在显示该人的第一个视频帧中进行勾画，使得鼠标指针的滑动轨迹包括该人或对应该人的轮廓。如此，通过人机交互实现参考帧及目标对象的选取，使得用户能够自主决定待处理音频数据对应的目标对象，提高了用户体验。

步骤 302：获取目标对象在参考帧中的位置信息。

20 在一些实施例中，终端可通过如下方式获取目标对象在参考帧中的位置信息：识别参考帧中目标对象的多个特征点，得到对应目标对象的特征点集合；基于特征点集合中的特征点，获取目标对象在参考帧中的位置信息；

其中，目标对象在参考帧中的位置信息包括：目标对象在参考帧中所处位置对应的角度信息；目标对象在全景视频中所处位置的中心距离，该中心距离  
25 为所述位置与参考帧对应的三维模型的中心的距离。

这里，对于参考帧中目标对象的特征点的识别可采用尺度不变特征变换 (SIFT, Scale-Invariant Feature Transform) 技术，或者方向梯度直方图 (HOG,

Histogram of Oriented Gradient) 技术。

5 在一些实施例中，终端可通过如下方式获取目标对象在参考帧中的位置信息：终端确定特征点集合中多个特征点的重心位置，基于重心位置，确定目标对象在全景视频中所处位置对应的角度信息；终端获取所述多个特征点在参考帧中占据的区域的像素面积，基于获取的像素面积，确定目标对象在全景视频中所处位置的中心距离。

10 这里，首先对全景视频中某特定位置对应的角度信息进行说明。在实际应用中，全景视频的视频帧中的像素点坐标和该像素点对应的三维模型的球面坐标存在固定的映射关系，图 5 为本公开实施例提供的全景视频的三维模型到二维图像的映射关系示意图，参见图 5，图 5 为根据 Tissot Indicatrix 投影的等距柱状图，该图示出了  $360^\circ$  全景空间与该  $360^\circ$  全景视频画幅的映射关系。图 6 为本公开实施例提供的全景视频对应的球坐标系的示意图，参见图 6，对于全景视频中的任一位置 A（即球面上任一点）来说，位置 A 的角度信息包括水平转角  $\psi$  及垂直转角  $\beta$  表示。

15 基于上述对角度信息的说明，在一些实施例中，可通过如下方式确定参考帧中目标对象在全景视频中所处位置对应的角度信息：

确定多个特征点的重心位置对应的像素点的像素点坐标；将像素点坐标转换成对应三维模型的球面坐标，球面坐标指示目标对象在全景视频中所处位置对应的角度。

20 接下来对目标对象在全景视频中所处位置的中心距离进行说明。以全景视频的三维模型为球面模型为例，参见图 7，图 7 为本公开实施例提供的全景视频的三维模型投影示意图，在球面模型 71 的中心模拟一个对应人眼的虚拟镜头 72，虚拟镜头与播放窗口具有相同的视场，从而播放窗口视场在球面模型的表面纹理中的投影区域，即是通过播放窗口能够在视频帧中观看到的内容，视频帧中目标对象与虚拟镜头的距离，即为目标对象在全景视频中所处位置的中心距离。当目标对象在全景视频中所处位置的中心距离变大时，相应的目标对象在全景视频中的像素面积会变小；当目标对象在全景视频中所处位置的中心距

离变小时，相应的目标对象在全景视频中的像素面积会变大。图 8 为本公开实施例提供的全景视频中对象的中心距离的示意图，参见图 8，其中，编号 1、2 及 3 所指示的均为全景视频帧中的对象，编号 0 指示的是全景视频的三维模型中的虚拟镜头，以 1 所指示的对象为目标对象为例进行说明，目标对象所处的  
5 位置的 5 中心距离为 4m。

基于上述对中心距离的说明，在一些实施例中，可通过如下方式确定目标对象在全景视频中所处位置的 10 中心距离：根据多个特征点在参考帧中占据的区域的像素面积与中心距离的平方所呈的反比例关系，确定目标对象在全景视频中所处位置的 10 中心距离。

10 这里，多个特征点在参考帧中占据的区域的像素面积，可以理解为目标对象在参考帧中所占据的像素面积。

在一些实施例中，目标对象在全景视频中所处位置的 15 中心距离  $d$ ，与目标对象在当前视频帧中对应的像素面积  $Area$  的关系如下所示：

$$d = k \sqrt{\frac{1}{Area}};$$

15 其中， $k$  为调整系数，具体数值的大小可依据实际情况进行设定。

步骤 303：基于所述位置信息，获取目标对象在全景视频中的运动轨迹。

在一些实施例中，可通过如下方式获取目标对象在全景视频中的运动轨迹：

获取全景视频的每个非参考帧中目标对象的位置信息；基于目标对象在参考帧中的位置信息，以及每个非参考帧中目标对象的位置信息，确定目标对象  
20 在全景视频中的运动轨迹。

接下来对非参考帧中目标对象的位置信息的获取进行说明。在一些实施例中，可通过如下方式确定非参考帧中目标对象的位置：将特征点集合中目标对象的多个特征点，与全景视频的非参考帧中目标对象的特征点进行特征点匹配，得到特征点匹配结果；基于特征点匹配结果，确定非参考帧中目标对象的位置。

25 这里，需要说明的是，对于非参考帧中目标对象位置信息的获取，为逐帧顺序获取，也即对于连续的多个非参考帧，依据非参考帧对应的播放时间的先



后，顺序的获取目标对象在每个非参考帧中的位置信息，也即，顺序的执行特征点集合中特征点与非参考帧中特征点的匹配。在执行特征点集合中特征点与非参考帧中特征点的匹配时，首先通过特征点集合中特征点可定位非参考帧中目标对象，然后基于识别得到的非参考帧中的目标对象，进行特征点提取，进而将提取得到的特征点与特征点集合中的特征点进行匹配。

在一些实施例中，基于特征点匹配结果，可通过如下方式确定非参考帧中目标对象的位置：基于特征点匹配结果，在非参考帧中，选取与参考帧中的特征点的相似度大于第一相似度阈值的特征点；基于选取的大于第一相似度阈值的特征点，确定非参考帧中目标对象的位置。

这里，在实际应用中，第一相似度阈值的取值可基于实际需要进行设定，具体地，第一相似度阈值的设定可保证如下：当参考帧中目标对象的第一特征点与特征点集合中第二特征点的相似度大于第一相似度阈值时，第一特征点与第二特征点为目标对象的同一特征点。

在实际应用中，存在如下情况：由于目标对象的移动或转动，某一个或多个在第一视频帧中识别到的目标对象的特征点，在第二视频帧中未出现，即无法检测到，第二视频帧的播放时间晚于第一视频帧，此时，为了保证目标对象在连续旋转、变化时不会被跟踪丢失，当特征点在连续的读多个帧图像中不再出现时，需要将不再出现的特征点从特征点集合中删除；相应的，在一些实施例中，还可通过如下方式更新特征点集合中的特征点：

获取预设数量的非参考帧，与特征点集合中特征点的匹配结果；预设数量的非参考帧为全景视频的连续视频帧；基于获取的特征点匹配结果，确定特征点集合中的待删除特征点，待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；将待删除特征点从特征点集合中删除。

这里，在实际应用中，第三相似度阈值的取值可依据实际情况进行设定，具体地，第三相似度阈值的取值可保证如下：当特征点集合中第三特征点与非参考帧中目标对象的第四特征点的相似度小于第三相似度阈值时，第三特征点与第四特征点为不同特征点，视为在非参考帧中未检测到第三特征点；上述获

取的非参考帧的具体数量可依据实际需要进行设定,如获取连续5个非参考帧,当特征点集合中某特征点在该连续5个非参考帧中均未出现,即特征点匹配的结果表征相似度小于第三相似度阈值时,将该特征点从特征点集合中删除。

在实际应用中,存在如下情况:由于目标对象的移动或转动,在进行特征点匹配时,在非参考帧中检测到的某一个或多个特征点,在特征点集合中不存在,此时,为了保证目标对象在连续旋转、变化时不会被跟踪丢失,当检测到目标对象新增的特征点时,需要将该特征点增加至特征点集合;相应的,在一些实施例中,还可通过如下方式更新特征点集合中的特征点:

基于特征点匹配结果,在非参考帧中,选取与特征点集合中每个特征点的相似度均小于第二相似度阈值的特征点;基于选取的小于第二相似度阈值的特征点,更新特征点集合中的特征点。

这里,在实际应用中,第二相似度阈值的取值可依据实际情况进行设定,具体地,第二相似度阈值的取值可保证如下:当非参考帧中目标对象的第五特征点与特征点集合中每个特征点的相似度均小于第二相似度阈值时,视为特征点集合中不存在第五特征点,即第五特征点为目标对象新增的特征点。

步骤304:基于运动轨迹,对待处理音频帧进行处理,得到能够表征目标对象在全景视频中所处位置的目标音频帧。

在一些实施例中,可通过如下方式对待处理音频帧进行处理,得到能够表征目标对象在全景视频中所处位置的目标音频帧:

调用三维音频引擎,对待处理音频帧根据运动轨迹进行混音处理,得到能够表征目标对象在全景视频中所处位置的目标音频帧。如此,可实现针对目标对象在全景视频中的运动轨迹,自动为目标对象进行配音,准确的还原目标对象的发声方位。

在一些实施例中,在得到能够表征目标对象在全景视频中所处位置的目标音频帧之后,还将目标音频帧与全景视频的视频帧进行合成,得到全景视频文件,该全景视频文件在播放时,目标对象的声音能够准确的反应目标对象在全景视频中所处的方位。

继续对本公开实施例提供的全景视频的处理方法进行说明。图 9 为本公开实施例提供的全景视频的处理方法的流程示意图，参见图 9，本公开实施例的全景视频的处理方法包括：

- 5           步骤 401：确定全景视频的视频帧中的参考帧，并确定参考帧中的目标对象。

在实际实施时，解码后的全景视频为许多帧图像，在目标对象出现的帧图像上，通过软件交互方式，如通过鼠标勾画，划出目标对象所处的区域。例如，如果对象是一个人，则把人的大体轮廓勾画出来。

- 10          步骤 402：识别参考帧中目标对象的多个特征点，得到对应目标对象的特征点集合。

这里，在实际实施时，可采用关键点识别技术，如 SIFT 或 HoG，识别参考帧中目标对象的多个特征点，得到一篮子的特征点，即目标对象的特征点集合。

- 15          步骤 403：将特征点集合中的特征点与非参考帧中目标对象的特征点进行相似度匹配，得到每个参考帧中的有效特征点。

这里，参考帧中的有效特征点为与特征点集合中特征点的相似度大于相似度阈值的特征点。

- 20          在实际实施时，在进行特征点相似度匹配的过程中，需要实时更新特征点集合中的特征点，例如删除特征点集合中连续 N 帧未出现的特征点，增加特征点集合中不存在的目标对象的特征点，具体更新方式可参见前述实施例。

步骤 404：根据参考帧中的特征点获取目标对象在参考帧中的位置信息，根据非参考帧中的有效特征点获取目标对象在每个非参考帧中的位置信息。

- 25          这里，位置信息包括：目标对象在全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考帧对应的三维模型的中心的距离。对于角度信息及中心距离的计算可参见前述实施例，在一些实施例中，可采用目标对象在全景视频中所处位

置对应的球面坐标表示目标对象的位置信息。

步骤 405：根据目标对象在参考帧中的位置信息，以及目标对象在每个非参考帧中的位置信息，获取目标对象在全景视频中的运动轨迹。

这里，以目标对象的位置信息用球面坐标表示为例，根据目标对象对应参考帧的位置的球面坐标，以及对应每个非参考帧的位置的球面坐标，绘制形成对应目标对象的运动轨迹的曲线，经平滑处理，得到表征目标对象在全景视频中运动轨迹的目标曲线。

步骤 406：调用三维音频引擎，对待处理音频帧根据运动轨迹进行混音处理，得到能够表征目标对象在全景视频中所处位置的目标音频帧。

步骤 407：将得到的目标音频帧与全景视频的视频帧合成，得到全景视频文件。

接下来对本公开实施例提供的全景视频的处理装置采用软件单元实施进行说明。图 10 为本公开实施例提供的全景视频的处理装置的组成结构示意图，参见图 10，本公开实施例的全景视频的处理装置包括：

确定单元 11，用于确定全景视频的视频帧中的参考帧，并确定所述参考帧中的目标对象；

获取单元 12，用于获取所述目标对象在所述参考帧中的位置信息；

以及，用于基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹；

处理单元 13，用于基于所述运动轨迹，对待处理音频帧进行处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧；

所述目标音频帧，用于与所述全景视频的视频帧合成得到全景视频文件。

在一些实施例中，所述获取单元，还用于识别所述参考帧中所述目标对象的多个特征点，得到对应所述目标对象的特征点集合；

基于所述特征点集合中的特征点，获取所述目标对象在所述参考帧中的位置信息，所述位置信息包括：

所述目标对象在所述全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考帧对应的三维模型的中心的距离。

5 在一些实施例中，所述获取单元，还用于确定所述特征点集合中多个特征点的重心位置；

基于所述重心位置，确定所述目标对象在所述全景视频中所处位置对应的角度信息；

获取所述多个特征点在所述参考帧中占据的区域的像素面积；

10 基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心距离。

在一些实施例中，所述获取单元，还用于确定所述重心位置对应的像素点的像素点坐标；

将所述像素点坐标转换成对应所述三维模型的球面坐标，所述球面坐标指示所述目标对象在所述全景视频中所处位置对应的角度。

15 在一些实施例中，所述获取单元，还用于根据所述像素面积与所述中心距离的平方所呈的反比例关系，确定所述目标对象在所述全景视频中所处位置的中心距离。

在一些实施例中，所述装置还包括匹配单元；

20 所述匹配单元，用于将所述特征点集合中的多个特征点，与所述全景视频的非参考帧中的特征点进行特征点匹配，得到特征点匹配结果；

所述获取单元，还用于基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置。

25 在一些实施例中，所述获取单元，还用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度大于第一相似度阈值的特征点；

基于选取的所述大于第一相似度阈值的特征点，确定所述非参考帧中所述目标对象的位置。

在一些实施例中，所述获取单元，还用于获取所述全景视频的每个非参考帧中所述目标对象的位置信息；

基于所述目标对象在所述参考帧中的位置信息，以及每个非参考帧中所述目标对象的位置信息，确定所述目标对象在所述全景视频中的运动轨迹。

5 在一些实施例中，所述装置还包括：

更新单元，用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集中的特征点的相似度小于第二相似度阈值的特征点；

基于选取的所述小于第二相似度阈值的特征点，更新所述特征点集中的特征点。

10 在一些实施例中，所述装置还包括：

删除单元，用于获取预设数量的非参考帧，与所述特征点集中特征点的匹配结果；所述预设数量的非参考帧为所述全景视频的连续视频帧；

基于获取的特征点匹配结果，确定所述特征点集中的待删除特征点，所述待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；

15 将所述待删除特征点从所述特征点集中删除。

在一些实施例中，所述确定单元，还用于接收到针对所述参考帧的对象选取操作；

基于所述对象选取操作在所述参考帧中所选定的区域，确定所述参考帧中的目标对象。

20 在一些实施例中，所述处理单元，还用于调用三维音频引擎，对待处理音频帧根据所述运动轨迹进行混音处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧。

25 这里需要指出的是：以上涉及装置的描述，与上述方法描述是类似的，同方法的有益效果描述，不做赘述，对于本公开实施例所述装置中未披露的技术细节，请参照本公开方法实施例的描述。

本公开实施例还提供一种存储有可执行指令的存储介质，其中存储有可执行指令，当可执行指令被处理器执行时，将引起处理器执行本公开实施例提供

的全景视频的处理方法。

以上描述仅为本公开的实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开实施例中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开实施例构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开实施例中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

## 权 利 要 求 书

1、一种全景视频的处理方法，其特征在于，所述方法包括：

确定全景视频的视频帧中的参考帧，并确定所述参考帧中的目标对象；

获取所述目标对象在所述参考帧中的位置信息；

5 基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹；

基于所述运动轨迹，对待处理音频帧进行处理，得到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧；

所述目标音频帧，用于与所述全景视频的视频帧合成得到全景视频文件。

2、如权利要求 1 所述的方法，其特征在于，所述获取所述目标对象在所述  
10 参考帧中的位置信息，包括：

识别所述参考帧中所述目标对象的多个特征点，得到对应所述目标对象的特征点集合；

基于所述特征点集中的特征点，获取所述目标对象在所述参考帧中的位置信息，所述位置信息包括：

15 所述目标对象在所述全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考帧对应的三维模型的中心的距离。

3、如权利要求 2 所述的方法，其特征在于，所述基于所述特征点集中的特征点，获取所述目标对象在所述参考帧中的位置信息，包括：

20 确定所述特征点集中多个特征点的重心位置；

基于所述重心位置，确定所述目标对象在所述全景视频中所处位置对应的角度信息；

获取所述多个特征点在所述参考帧中占据的区域的像素面积；

25 基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心距离。

4、如权利要求 3 所述的方法，其特征在于，所述基于所述重心位置，确定



所述目标对象在所述全景视频中所处位置对应的角度信息，包括：

确定所述重心位置对应的像素点的像素点坐标；

将所述像素点坐标转换成对应所述三维模型的球面坐标，所述球面坐标指示所述目标对象在所述全景视频中所处位置对应的角度。

5 5、如权利要求3所述的方法，其特征在于，所述基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心距离，包括：

根据所述像素面积与所述中心距离的平方所呈的反比例关系，确定所述目标对象在所述全景视频中所处位置的中心距离。

6、如权利要求2所述的方法，其特征在于，所述方法还包括：

10 将所述特征点集合中的多个特征点，与所述全景视频的非参考帧中的特征点进行特征点匹配，得到特征点匹配结果；

基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置。

7、如权利要求6所述的方法，其特征在于，所述基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置，包括：

15 基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度大于第一相似度阈值的特征点；

基于选取的所述大于第一相似度阈值的特征点，确定所述非参考帧中所述目标对象的位置。

20 8、如权利要求1所述的方法，其特征在于，所述基于所述位置信息，获取所述目标对象在所述全景视频中的运动轨迹，包括：

获取所述全景视频的每个非参考帧中所述目标对象的位置信息；

基于所述目标对象在所述参考帧中的位置信息，以及每个非参考帧中所述目标对象的位置信息，确定所述目标对象在所述全景视频中的运动轨迹。

9、如权利要求6所述的方法，其特征在于，所述方法还包括：

25 基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度小于第二相似度阈值的特征点；

基于选取的所述小于第二相似度阈值的特征点，更新所述特征点集合中的

特征点。

10、如权利要求 6 所述的方法，其特征在于，所述方法还包括：

获取预设数量的非参考帧，与所述特征点集合中特征点的匹配结果；所述  
预设数量的非参考帧为所述全景视频的连续视频帧；

5 基于获取的特征点匹配结果，确定所述特征点集合中的待删除特征点，所  
述待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；

将所述待删除特征点从所述特征点集合中删除。

11、如权利要求 1 所述的方法，其特征在于，所述确定所述参考帧中的目  
标对象，包括：

10 接收到针对所述参考帧的对象选取操作；

基于所述对象选取操作在所述参考帧中所选定的区域，确定所述参考帧中  
的目标对象。

12、如权利要求 1 所述的方法，其特征在于，所述基于所述运动轨迹，对  
待处理音频帧进行处理，得到能够表征所述目标对象在所述全景视频中所处位  
15 置的目标音频帧，包括：

调用三维音频引擎，对待处理音频帧根据所述运动轨迹进行混音处理，得  
到能够表征所述目标对象在所述全景视频中所处位置的目标音频帧。

13、一种全景视频的处理装置，其特征在于，所述装置包括：

20 确定单元，用于确定全景视频的视频帧中的参考帧，并确定所述参考帧中  
的目标对象；

获取单元，用于获取所述目标对象在所述参考帧中的位置信息；

以及，用于基于所述位置信息，获取所述目标对象在所述全景视频中的运  
动轨迹；

25 处理单元，用于基于所述运动轨迹，对待处理音频帧进行处理，得到能够  
表征所述目标对象在所述全景视频中所处位置的目标音频帧；

所述目标音频帧，用于与所述全景视频的视频帧合成得到全景视频文件。

14、如权利要求 13 所述的装置，其特征在于，

所述获取单元，还用于识别所述参考帧中所述目标对象的多个特征点，得到对应所述目标对象的特征点集合；

基于所述特征点集合中的特征点，获取所述目标对象在所述参考帧中的位置信息，所述位置信息包括：

- 5 所述目标对象在所述全景视频中所处位置对应的角度信息；所述目标对象在所述全景视频中所处位置的中心距离，所述中心距离为所述位置与所述参考帧对应的三维模型的中心的距离。

15、如权利要求 14 所述的装置，其特征在于，

所述获取单元，还用于确定所述特征点集合中多个特征点的重心位置；

- 10 基于所述重心位置，确定所述目标对象在所述全景视频中所处位置对应的角度信息；

获取所述多个特征点在所述参考帧中占据的区域的像素面积；

基于所述像素面积，确定所述目标对象在所述全景视频中所处位置的中心距离。

- 15 16、如权利要求 15 所述的装置，其特征在于，

所述获取单元，还用于确定所述重心位置对应的像素点的像素点坐标；

将所述像素点坐标转换成对应所述三维模型的球面坐标，所述球面坐标指示所述目标对象在所述全景视频中所处位置对应的角度。

17、如权利要求 15 所述的装置，其特征在于，

- 20 所述获取单元，还用于根据所述像素面积与所述中心距离的平方所呈的反比例关系，确定所述目标对象在所述全景视频中所处位置的中心距离。

18、如权利要求 14 所述的装置，其特征在于，所述装置还包括匹配单元；

所述匹配单元，用于将所述特征点集合中的多个特征点，与所述全景视频的非参考帧中的特征点进行特征点匹配，得到特征点匹配结果；

- 25 所述获取单元，还用于基于所述特征点匹配结果，确定所述非参考帧中所述目标对象的位置。

19、如权利要求 18 所述的装置，其特征在于，

所述获取单元，还用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度大于第一相似度阈值的特征点；

基于选取的所述大于第一相似度阈值的特征点，确定所述非参考帧中所述目标对象的位置。

5 20、如权利要求 13 所述的装置，其特征在于，

所述获取单元，还用于获取所述全景视频的每个非参考帧中所述目标对象的位置信息；

基于所述目标对象在所述参考帧中的位置信息，以及每个非参考帧中所述目标对象的位置信息，确定所述目标对象在所述全景视频中的运动轨迹。

10 21、如权利要求 18 所述的装置，其特征在于，所述装置还包括：

更新单元，用于基于所述特征点匹配结果，在所述非参考帧中，选取与所述特征点集合中的特征点的相似度小于第二相似度阈值的特征点；

基于选取的所述小于第二相似度阈值的特征点，更新所述特征点集合中的特征点。

15 22、如权利要求 18 所述的装置，其特征在于，所述装置还包括：

删除单元，用于获取预设数量的非参考帧，与所述特征点集合中特征点的匹配结果；所述预设数量的非参考帧为所述全景视频的连续视频帧；

基于获取的特征点匹配结果，确定所述特征点集合中的待删除特征点，所述待删除特征点与获取的非参考帧中的特征点的相似度小于第三相似度阈值；

20 将所述待删除特征点从所述特征点集合中删除。

23、如权利要求 13 所述的装置，其特征在于，

所述确定单元，还用于接收到针对所述参考帧的对象选取操作；

基于所述对象选取操作在所述参考帧中所选定的区域，确定所述参考帧中的目标对象。

25 24、如权利要求 13 所述的装置，其特征在于，

所述处理单元，还用于调用三维音频引擎，对待处理音频帧根据所述运动轨迹进行混音处理，得到能够表征所述目标对象在所述全景视频中所处位置的

目标音频帧。

25、一种全景视频的处理装置，其特征在于，所述装置包括：

存储器，用于存储可执行指令；

处理器，用于执行所述可执行指令时，实现如权利要求 1 至 12 任一项所述

5 的全景视频的处理方法。

26、一种非暂态存储介质，其特征在于，存储有可执行指令，所述可执行指令被执行时，用于实现如权利要求 1 至 12 任一项所述的全景视频的处理方法。

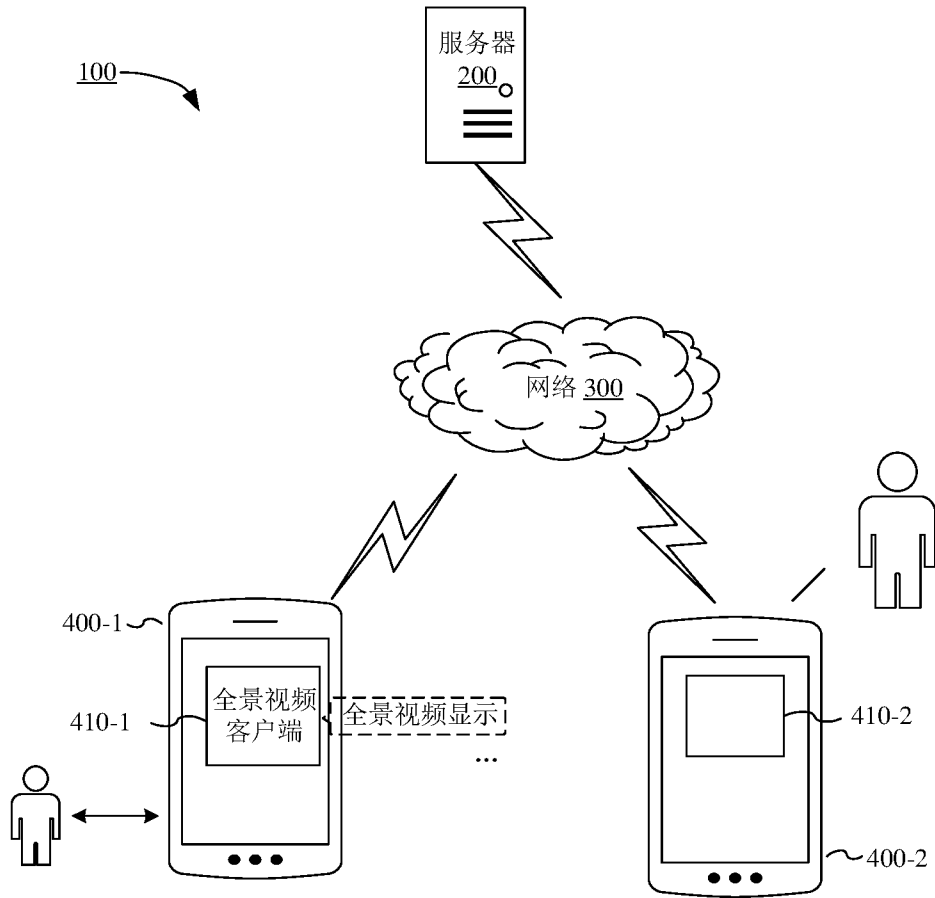


图 1

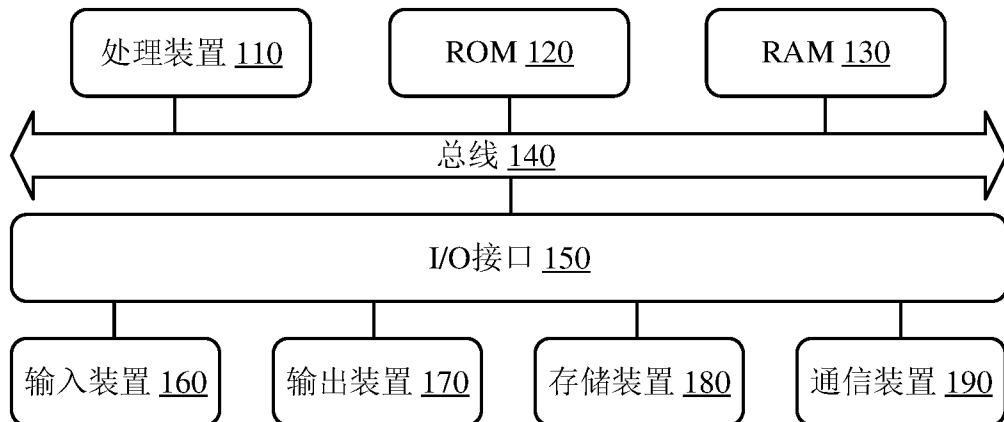


图 2

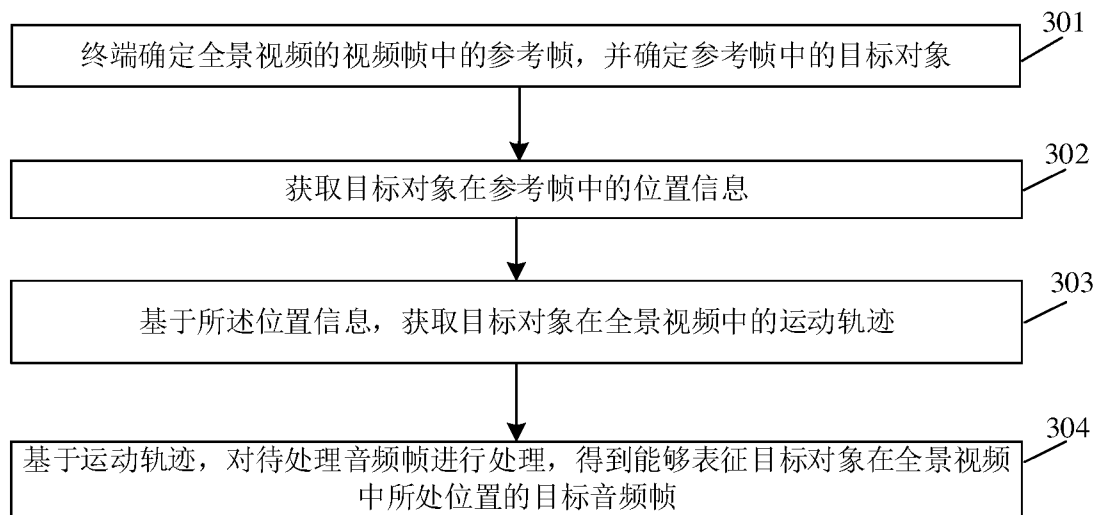


图 3

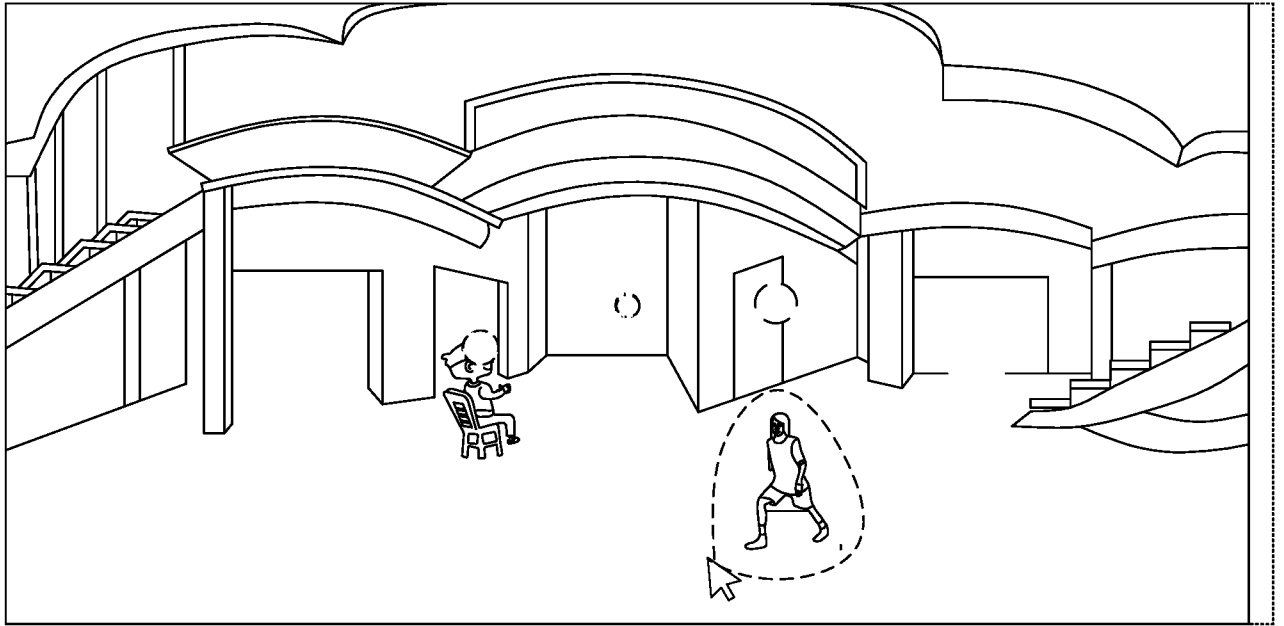


图 4

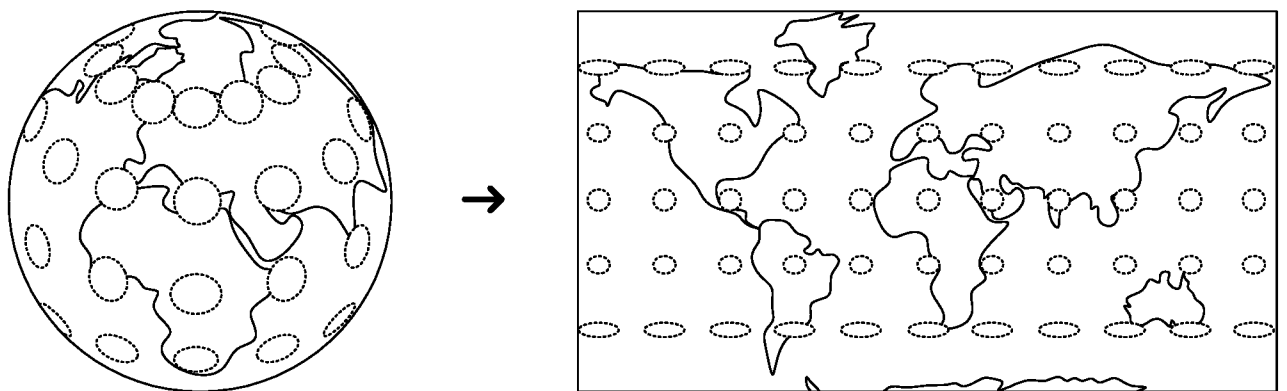


图 5



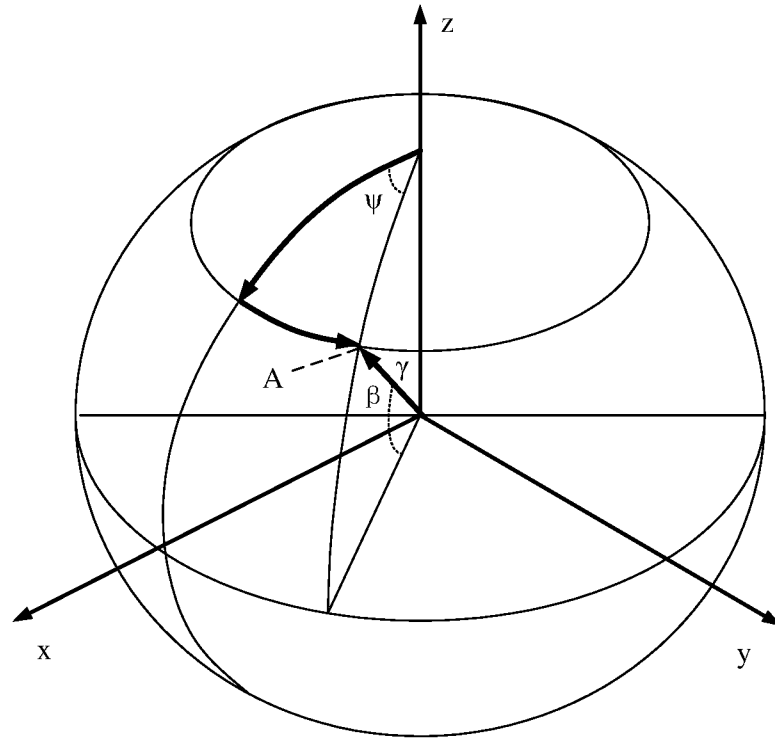


图 6

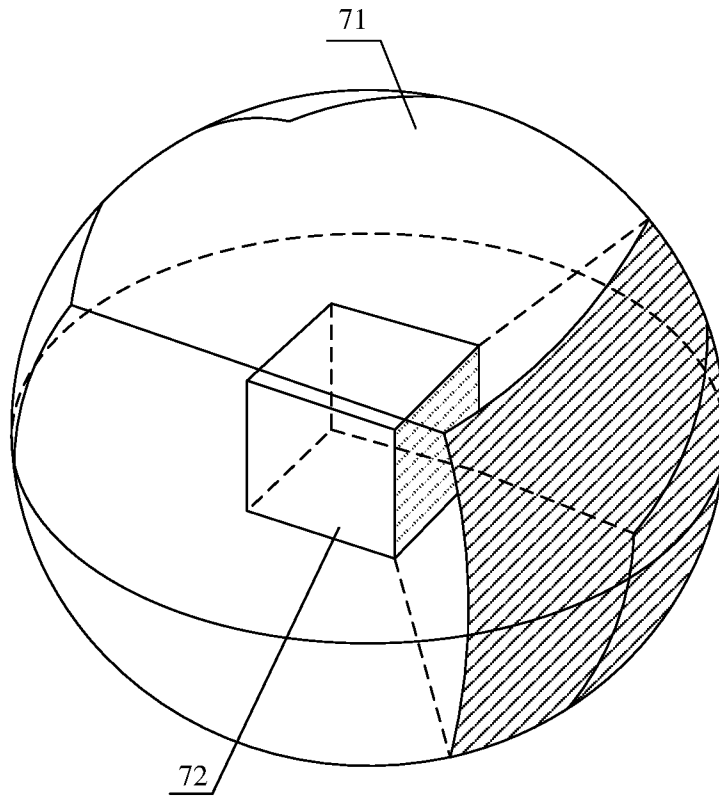


图 7

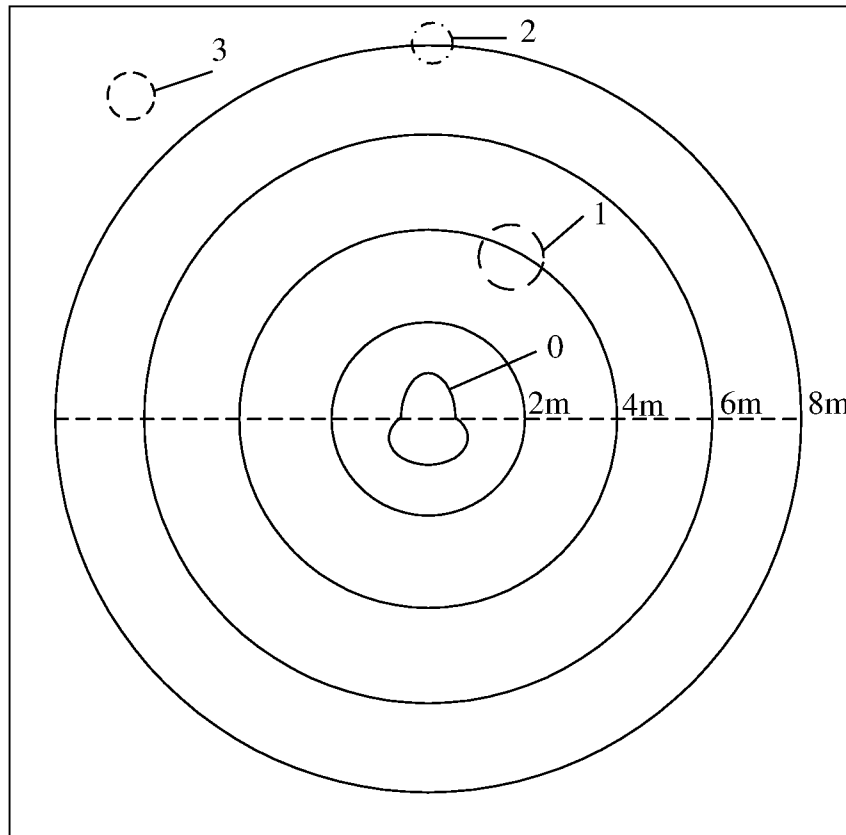


图 8

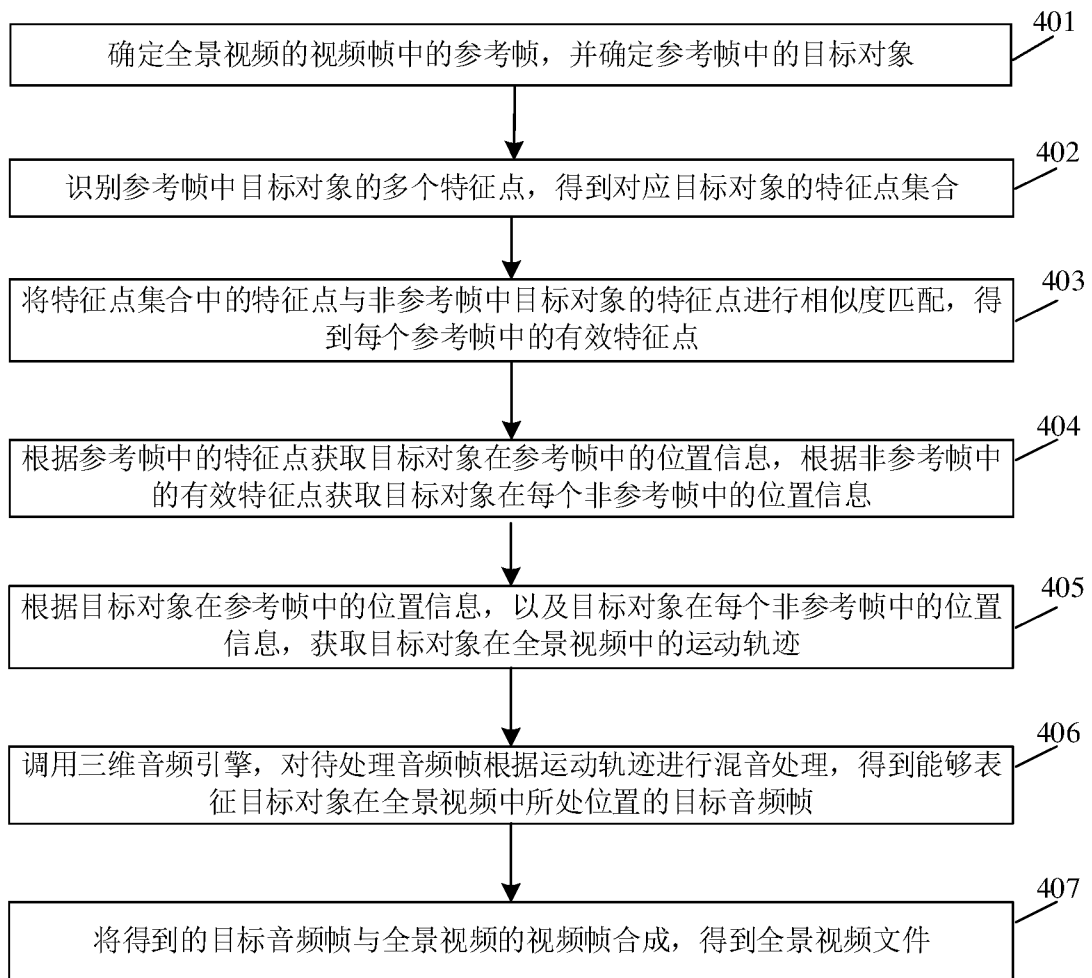


图 9

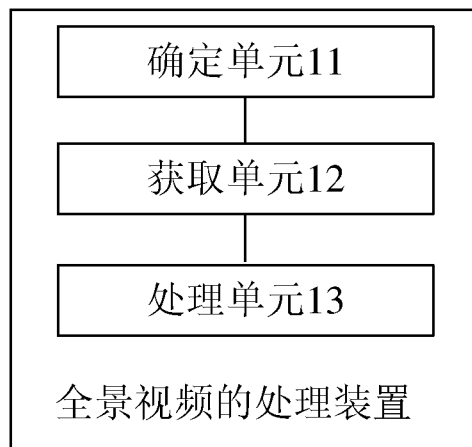


图 10

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/094507

**A. CLASSIFICATION OF SUBJECT MATTER**

H04N 21/234(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, WPI, EPODOC, IEEE: 全景, 视频, 序列, 音频, 声音, 合成, 融合, 目标, 对象, 位置, 轨迹, 定位, 角度, 距离, 重心, 特征, 匹配, panorama, video, audio, sound, combin+, synthesis, merg+, person, people, object, location, position, trace, track, angle, distance, gravity, character+, match

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106331732 A (FENGJING TECHONOLGY CO., LTD.) 11 January 2017 (2017-01-11) description, paragraphs 0067-0077	1-26
Y	CN 106992959 A (SHENZHEN PISOFTWARE TECHNOLOGY CO., LTD.) 28 July 2017 (2017-07-28) description paragraphs 0049, 0075	1-26
A	CN 106210703 A (BEIJING MAKEMAGIC TECHNOLOGY DEVELOPMENT CO., LTD.) 07 December 2016 (2016-12-07) entire document	1-26
A	US 2015350521 A1 (CAMERA SLICE, INC.) 03 December 2015 (2015-12-03) entire document	1-26
A	CN 101119482 A (BEIJING ZANB TECHNOLOGY CO., LTD.) 06 February 2008 (2008-02-06) entire document	1-26
A	CN 105895108 A (NANJING QINGJIN INFORMATION TECHNOLOGY CO., LTD.) 24 August 2016 (2016-08-24) entire document	1-26

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

19 August 2020

Date of mailing of the international search report

28 August 2020

Name and mailing address of the ISA/CN

**China National Intellectual Property Administration (ISA/  
CN)  
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing  
100088  
China**

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2020/094507**

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108319362 A (LENOVO (BEIJING) LIMITED) 24 July 2018 (2018-07-24) entire document	1-26
<hr/>		

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2020/094507**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106331732	A	11 January 2017	CN	106331732	B	12 November 2019
CN	106992959	A	28 July 2017	WO	2018082284	A1	11 May 2018
CN	106210703	A	07 December 2016	CN	106210703	B	08 June 2018
US	2015350521	A1	03 December 2015	US	9113068	B1	18 August 2015
CN	101119482	A	06 February 2008	CN	101119482	B	20 July 2011
CN	105895108	A	24 August 2016	CN	105895108	B	24 January 2020
CN	108319362	A	24 July 2018	None			

国际检索报告

国际申请号

PCT/CN2020/094507

<p><b>A. 主题的分类</b></p> <p>H04N 21/234(2011.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																										
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNKI, CNPAT, WPI, EPODOC, IEEE:全景, 视频, 序列, 音频, 声音, 合成, 融合, 目标, 对象, 位置, 轨迹, 定位, 角度, 距离, 重心, 特征, 匹配, panorama, video, audio, sound, combin+, synthesis, merg+, person, people, object, location, position, trace, track, angle, distance, gravity, character+, match</p>																										
<p><b>C. 相关文件</b></p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 106331732 A (北京疯景科技有限公司) 2017年 1月 11日 (2017 - 01 - 11) 说明书第0067-0077段</td> <td>1-26</td> </tr> <tr> <td>Y</td> <td>CN 106992959 A (深圳市圆周率软件科技有限责任公司) 2017年 7月 28日 (2017 - 07 - 28) 说明书第0049、0075段</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 106210703 A (北京美吉克科技发展有限公司) 2016年 12月 7日 (2016 - 12 - 07) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>US 2015350521 A1 (CAMERA SLICE, INC.) 2015年 12月 3日 (2015 - 12 - 03) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 101119482 A (北京智安邦科技有限公司) 2008年 2月 6日 (2008 - 02 - 06) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 105895108 A (南京青衿信息科技有限公司) 2016年 8月 24日 (2016 - 08 - 24) 全文</td> <td>1-26</td> </tr> <tr> <td>A</td> <td>CN 108319362 A (联想北京有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文</td> <td>1-26</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 106331732 A (北京疯景科技有限公司) 2017年 1月 11日 (2017 - 01 - 11) 说明书第0067-0077段	1-26	Y	CN 106992959 A (深圳市圆周率软件科技有限责任公司) 2017年 7月 28日 (2017 - 07 - 28) 说明书第0049、0075段	1-26	A	CN 106210703 A (北京美吉克科技发展有限公司) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-26	A	US 2015350521 A1 (CAMERA SLICE, INC.) 2015年 12月 3日 (2015 - 12 - 03) 全文	1-26	A	CN 101119482 A (北京智安邦科技有限公司) 2008年 2月 6日 (2008 - 02 - 06) 全文	1-26	A	CN 105895108 A (南京青衿信息科技有限公司) 2016年 8月 24日 (2016 - 08 - 24) 全文	1-26	A	CN 108319362 A (联想北京有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文	1-26
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
Y	CN 106331732 A (北京疯景科技有限公司) 2017年 1月 11日 (2017 - 01 - 11) 说明书第0067-0077段	1-26																								
Y	CN 106992959 A (深圳市圆周率软件科技有限责任公司) 2017年 7月 28日 (2017 - 07 - 28) 说明书第0049、0075段	1-26																								
A	CN 106210703 A (北京美吉克科技发展有限公司) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-26																								
A	US 2015350521 A1 (CAMERA SLICE, INC.) 2015年 12月 3日 (2015 - 12 - 03) 全文	1-26																								
A	CN 101119482 A (北京智安邦科技有限公司) 2008年 2月 6日 (2008 - 02 - 06) 全文	1-26																								
A	CN 105895108 A (南京青衿信息科技有限公司) 2016年 8月 24日 (2016 - 08 - 24) 全文	1-26																								
A	CN 108319362 A (联想北京有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文	1-26																								
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																										
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&amp;” 同族专利的文件</p>																										
<p>国际检索实际完成的日期</p> <p>2020年 8月 19日</p>	<p>国际检索报告邮寄日期</p> <p>2020年 8月 28日</p>																									
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN)</p> <p>中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>	<p>授权官员</p> <p>李晶</p> <p>电话号码 86-10-53961705</p>																									

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2020/094507

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	106331732	A	2017年 1月 11日	CN	106331732	B	2019年 11月 12日
CN	106992959	A	2017年 7月 28日	WO	2018082284	A1	2018年 5月 11日
CN	106210703	A	2016年 12月 7日	CN	106210703	B	2018年 6月 8日
US	2015350521	A1	2015年 12月 3日	US	9113068	B1	2015年 8月 18日
CN	101119482	A	2008年 2月 6日	CN	101119482	B	2011年 7月 20日
CN	105895108	A	2016年 8月 24日	CN	105895108	B	2020年 1月 24日
CN	108319362	A	2018年 7月 24日	无			