

音频处理方法、装置、可读介质及电子设备

本申请要求于 2020 年 03 月 31 日提交中国专利局、申请号为 202010245598.6、申请名称为“音频处理方法、装置、可读介质及电子设备”的中国专利申请的优先权，其全部
5 内容通过引用结合在本申请中。

技术领域

本公开涉及计算机技术领域，具体地，涉及一种音频处理方法、装置、可读介质及电子设备。

10

背景技术

AR（Augmented Reality，增强现实）强调虚拟与现实的结合，它作为短视频应用玩法的新潮流，能够令用户与真实世界进行交互，例如，在拍摄到特定物体时显示与该物体对应的特效。目前的 AR 产品能够在视觉上为用户实现虚拟与现实的结合，但是，对于带有声音的特效，用户在感知环境和声音时容易产生“错位感”。举例来说，若预先设置识别到某物体时播放特定音频，那么如果用户通过手机拍摄到这一物体时手机会播放特定音频，而当用户逐渐远离或靠近该物体时，用户听到的声音不会变化，如用户逐渐远离该物体，在用户的常识认知中听到的声音应当变小，而实际上用户始终能够听到同样大的声音，这与用户的认知不符，导致用户产生“错位感”，影响用户的体验，降低用户在 AR 场景
15 中的沉浸感。
20

发明内容

提供该发明内容部分以便以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。该发明内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。
25

第一方面，本公开提供一种音频处理方法，所述方法包括：

获取终端抓取到的原始图像；

根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；

根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。

5 第二方面，本公开提供一种音频处理装置，所述装置包括：

第一获取模块，用于获取终端抓取到的原始图像；

第一确定模块，用于根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；

10 音频处理模块，用于根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。

第三方面，本公开提供一种计算机可读介质，其上存储有计算机程序，该程序被处理装置执行时实现本公开第一方面所述方法的步骤。

15 第四方面，本公开提供一种电子设备，包括：

存储装置，其上存储有计算机程序；

处理装置，用于执行所述存储装置中的所述计算机程序，以实现本公开第一方面所述方法的步骤。

20 第五方面，本公开提供一种计算机程序产品，所述程序产品包括：计算机程序，所述计算机程序被处理装置执行时实现本公开第一方面所述方法的步骤。

第六方面，本公开提供一种计算机程序，所述计算机程序被处理装置执行时实现本公开第一方面所述方法的步骤。

25 通过上述技术方案，获取终端抓取到的原始图像，根据该原始图像，确定目标对象相对于终端的三维相对位置，作为第一三维相对位置，并根据第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。其中，目标声音为音效对象对应的音效声音。由此，能够基于目标对象相对于终端的三维相对位置对目标声音进行三维效果处理，得到的音频在用户听来更符合终端所处的实时位置。并且，基于这一方式，在终端持续抓取图像的过程中，能够持续生成符合终端实时位置的带三维效果

的音频，且随终端位置的变化而适应性变化，手持终端的用户能够听到符合当前移动趋势的具有空间感的音频，用户在虚拟场景下的沉浸体验将会有所提升。

本公开的其他特征和优点将在随后的具体实施方式部分予以详细说明。

5 附图说明

结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。在附图中：

图 1 是根据本公开的一种实施方式提供的音频处理方法的流程图；

10 图 2 是根据本公开提供的音频处理方法中，根据原始图像确定目标对象相对于终端的三维相对位置的步骤的一种示例性的流程图；

图 3 是根据本公开的一种实施方式提供的音频处理装置的框图；

图 4 示出了适于实现本公开实施例的电子设备的结构示意图。

15 具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

25 本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

30 需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

5 图 1 是根据本公开的一种实施方式提供的音频处理方法的流程图。如图 1 所示，该方法可以包括以下步骤。

在步骤 11 中，获取终端抓取到的原始图像。

10 终端可以通过具有拍摄功能的装置（例如，设置于终端的摄像头）实现图像抓取。示例地，当用户打开终端的拍摄类程序，屏幕上显示的取景框（例如，AR 取景框）会持续对取景框内的内容进行抓取，也就是持续进行图像抓取，具体的抓取方式可以例如为在取景框被开启后进行周期性抓取（周期可根据经验值预设），在到达每一抓取时刻自动抓取当下取景框内的内容。在实际应用场景中，若开启了终端的取景框，自取景框被开启开始，到该取景框此次被关闭，过程中可能进行了多次抓取，每一次都能抓取到相应的图像，抓取到的每一图像均可作为原始图像。

15 在步骤 12 中，根据原始图像，确定目标对象相对于终端的三维相对位置，作为第一三维相对位置。

20 通过原始图像，可以确定目标对象相对于终端的三维相对位置。在这里，目标对象可以是存在于原始图像中的对象，或者，目标对象还可以是未出现在原始图像中的对象（前提是终端在同一拍摄过程中、拍摄到原始图像前的历史拍摄中已拍摄到目标对象，目标对象只是在终端拍摄移动的过程中、在原始图像中暂时“出画”）。

在本公开提供的应用场景中，认为目标对象对应有音效对象，音效对象对应有音效声音（目标对象、音效对象、音效声音之间如何对应均可预先设置），在效果上表现为在确定存在目标对象时播放目标对象所对应的音效对象的音效声音，且在用户听来发声的声源位置（即，音效声音的位置）应当满足用户的认知。

25 其中，音效对象可以与目标对象处于同一位置，也可以处于不同位置（例如，在目标对象周边）。例如，设置在识别到建筑物 A 时播放飞机飞行的声音，且飞机从建筑物 A 左侧飞行至建筑物 A 右侧，则目标对象为建筑物 A，音效对象为飞机，音效声音为飞机飞行的声音，可知，在这一示例中，音效对象与目标对象并未处于相同位置，音效对象是从目标对象的左侧移动至目标对象的右侧。再例如，设置在识别到建筑物
30 B 时播放建筑物 B 的介绍语音，则目标对象为建筑物 B，音效对象为建筑物 B，音效

声音为建筑物 B 的介绍语音，可知，在这一示例中，音效对象与目标对象处于相同位置。

在实际场景中，需要在确定存在目标对象（例如，目标对象在画面中，或者，目标对象暂时“出画”）时播放声音，本方案的目的在于将播放效果一成不变的音频处理成更加具有空间感的音频，营造空间感必不可少的就是空间位置，而营造空间感所需的

5 空间位置自然就是目标对象相对于终端的三维相对位置，即第一三维相对位置。

在一种可能的实施方式中，步骤 12 可以包括以下步骤，如图 2 所示。

在步骤 21 中，对原始图像进行特征提取，获取原始图像中的第一锚点信息。

第一锚点信息可以包括对原始图像进行特征提取后，识别出的原始图像中的至少一个锚点（即，特征点）。其中，特征提取的方法为本领域的常规手段，此处不赘述。

10

在步骤 22 中，根据预先存储的对应于目标对象的参考锚点信息，判断第一锚点信息能否与参考锚点信息相匹配。

在实际应用场景中，若图像中存在目标对象，就可以呈现与目标对象对应的特效、声音等。举例来说，若目标对象为钟楼，并且，设置当图像中存在钟楼时，显示时钟动画，并播放钟声，那么当原始图像内包含钟楼时，应当将时钟附着在钟楼上，同时

15 播放钟声。

目标对象是已建模的对象，在数据处理前期，首先针对包含有目标对象的一系列图像进行特征提取，确定目标对象的参考锚点信息并存储，其中，目标对象的参考锚点信息包括能够辅助识别出目标对象的至少一个锚点。示例地，若目标对象为书桌，

20 则首先收集包含书桌的大量图像，再基于收集到的图像，对每个图像进行特征提取，得到能够表征书桌的特征点，从而确定对应于书桌的参考锚点信息并与书桌这一对象关联存储。实际应用中，可以构建一锚点信息库，其中存储已建模的所有对象对应的参考锚点信息，从而能够更加快速地确定原始图像中是否存在目标对象以及存在何种目标对象。

如上所述，目标对象的参考锚点信息能够很好地辅助识别出目标对象。因此，通过比较第一锚点信息与参考锚点信息，可以判断原始图像中是否存在与目标对象相似的对象，从而确定原始图像中是否存在目标对象。具体地，可以判断第一锚点信息能否与参考锚点信息相匹配。示例地，可以计算第一锚点信息和参考锚点信息的相似程度，若二者相似程度高于相似度阈值（可根据经验值设定），就判定二者相匹配，而

25

若二者相似度未达到相似度阈值，就判定二者不匹配。计算第一锚点信息和参考锚点信息的相似度可以采用余弦相似度等常规计算方式，此处不赘述。

在步骤 23 中，若确定第一锚点信息与参考锚点信息相匹配，根据第一锚点信息，确定目标对象相对于终端的三维相对位置，作为第一三维相对位置。

- 5 若经步骤 22 确定第一锚点信息与参考锚点信息相匹配，则说明原始图像中存在目标对象，且该目标对象是预先已存储的对象之一。从而，可以直接根据原始图像的第一锚点信息，确定目标对象相对于终端的三维相对位置，即第一三维相对位置。其中，根据图像以及图像中某个点（或者，多个点）确定该点相对于拍摄图像终端的三维相对位置属于本领域的常用手段，此处不过多描述，例如，对于 iOS 系统来说，可以使用 AR 工具箱（ARKit）、AR 核心（ARCore）这一类的接口完成上述数据处理。
- 10

在另一种可能的实施方式中，步骤 12 可以包括以下步骤：

若获取到用户对原始图像进行锚点标记操作而产生的第二锚点信息，根据第二锚点信息确定目标对象相对于终端的三维相对位置，作为第一三维相对位置。

- 其中，第二锚点信息可以包括用户对原始图像进行锚点标记操作而产生的至少一个锚点。在这一实施方式中，用户可以手动对取景框内进行锚点标记操作，也就是说，不再需要对原始图像进行复杂的图片分析，而是基于用户锚点标记操作而产生的第二锚点信息，确定原始图像中与第二锚点信息所对应的目标对象。示例地，根据第二锚点信息可以在原始图像中圈定出一个范围，并认为该范围所包含的内容就是目标对象。在后续的图像抓取中，对目标对象的定位均可以以此为基础，举例来说，若用户
- 15 标记锚点在原始图像的正中心，则当前确定的目标对象就是原始图像中心所对应的内容，在后续的抓取时刻中，若终端移动，例如，终端向左上移动，则目标对象的定位向右下移动，而并非仍处于中心。
- 20

- 根据第二锚点信息确定第一三维相对位置与步骤 23 相似，即，直接根据原始图像的第二锚点信息确定目标对象相对于终端的三维相对位置，作为第一三维相对位置。其中，如上文所述，根据图像以及图像中某个点（或者，多个点）确定该点相对于拍摄图像终端的三维相对位置属于本领域的常用手段，例如，对于 iOS 系统来说，可以使用 ARKit、ARCore 这一类的接口完成上述数据处理。
- 25

采用上述方式，用户可以直接定位目标对象，无需对图像进行复杂的分析计算，能有效减少数据处理量，且提供给用户的灵活性更高，用户体验提高。

- 30 在另一种可能的实施方式中，步骤 12 可以包括以下步骤：

若原始图像中不存在目标对象，获取目标对象对应的历史运动轨迹信息；

根据历史运动轨迹信息，确定目标对象相对于终端的三维相对位置，作为第一三维相对位置。

如上文所述，在终端移动拍摄过程中，可能会因终端的移动而导致目标对象暂时“出画”，在这一情形下，虽然目标对象并未出现在原始图像中，但是目标对象相对于终端的三维相对位置仍是可以获得。

若原始图像中不存在目标对象，说明此时目标对象可能暂时移动到终端拍摄取景框之外，因此，可以获取目标对象对应的历史运动轨迹信息。其中，历史运动轨迹信息是通过对终端在原始图像前抓取的图像进行运动追踪得到的。

10 通过历史运动轨迹信息，可以推测出终端抓取到原始图像时，目标对象与终端的相对位置。例如，通过目标对象的历史运动轨迹信息可以确定目标对象的运动方向以及运动速度，从而，可以以原始图像之前的、最后一帧存在目标对象的图像中目标对象所在位置为起点，依据历史运动轨迹中目标对象的运动方向、运动速度，以及最后一帧存在目标对象的图像与原始图像的拍摄时间差确定目标对象相对于终端的三维
15 相对位置，作为第一三维相对位置。

采用上述方式，针对终端未拍摄到目标对象的情况，还能够通过历史对目标对象的运动追踪确定目标对象相对于终端的三维相对位置，避免终端的移动对目标对象的位置确定产生影响。

回到图 1，在步骤 13 中，根据第一三维相对位置，对目标声音进行三维效果处理。
20 根据第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。

其中，目标声音就是音效对象对应的音效声音。

在一种可能的实施方式中，步骤 13 可以包括以下步骤：

25 根据第一三维相对位置、以及目标对象与目标对象对应的音效对象之间的位置关系，确定目标声音的声源位置相对于终端的三维相对位置，作为第二三维相对位置；

根据第二三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。

与目标对象对应的音效对象可以为一个，也可以为多个，若与目标对象对应的音效对象为多个，则每一音效对象各自具有与目标对象之间的位置关系。

根据第一三维相对位置、以及目标对象与目标对象对应的音效对象之间的位置关系，目标对象相对于终端的三维相对位置、目标对象相对于音效对象的声源（即，目标声音的声源位置）的三维相对位置均已知，从而，容易确定出终端相对于音效对象的声源（即，目标声音的声源位置）的三维相对位置，即，第二三维相对位置。之后，根据第二三维相对位置，对目标声音进行三维效果处理，就能够获得经三维效果处理后的音频，并且，在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。

10 其中，对目标声音的三维效果处理可以使用 3D 声音渲染，示例地，3D 声音渲染可以使用环绕声（ambisonic）技术，它是本领域的常规手段，能够基于音频和位置信息模拟音频在空间中的声像位置，进而生成对应于双耳的播放信号。

在另一种可能的实施方式中，本公开提供的方法还可以包括以下步骤：

获取目标声音的附加描述信息。

15 其中，附加描述信息包括但不限于以下中的至少一者：目标声音的声源发声面积、目标声音的声源混响效果、目标声音的声源环境；

目标声音的声源发声面积这一附加描述信息可以影响音频音量的大小。例如，声源发声面积小则音量小，可参考手机等小物体。再例如，声源发声面积大则音量大，可参考建筑物等大物体。目标声音的声源混响效果这一附加描述信息可以影响音频的混响效果，也就是有无混响效果。目标声音的声源环境这一附加描述信息可以影响音频的环境效果，例如，目标声音的声源处于室内，或者，处于室外。

在这一实施方式中，根据第二三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系，可以包括以下步骤：

25 根据第二三维相对位置和附加描述信息，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。

在这一实施方式中，实际上是结合多种声音处理方式，例如，3D 声音渲染、混响效果、音量调节、室内外效果等，以获得处理后的音频。示例地，若目标声音的附

加描述信息为目标声音的声源发声面积小，则对目标声音进行三维处理时，在对目标声源的音频进行 3D 声音渲染的同时，还会调小目标声源的音频的音量。

通过上述技术方案，获取终端抓取到的原始图像，根据该原始图像，确定目标对象相对于终端的三维相对位置，作为第一三维相对位置，并根据第一三维相对位置，
5 对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中目标声音的声源位置与第一三维相对位置符合目标对象与目标对象对应的音效对象之间的位置关系。其中，目标声音为音效对象对应的音效声音。由此，能够基于目标对象相对于终端的三维相对位置对目标声音进行三维效果处理，得到的音频在用户听来更符合终端所处的实时位置。并且，基于这一方式，在终端持续抓取图像的过程中，能够持续生成符合终端实时位置的带三维效果的音频，且随终端位置的变化而适应性变化，手持终端的用户能够听到符合当前移动趋势的具有空间感的音频，用户在虚拟场景下的沉浸体验将会有所提升。
10

可选地，在上述各实施例的基础上，本公开提供的方法还可以包括以下步骤：

将目标对象对应的图像素材叠加至原始图像的第一三维相对位置，以生成原始图像
15 对应的 AR 图像；

显示 AR 图像。

这一实施例描述的是有关 AR 图像特效的场景。每一目标对象均可对应有图像素材，例如，建筑物对应于能够附着在建筑物外侧的内容，该内容就是建筑物对应的图像素材。因此，在对音频进行处理的基础上，为了保证用户在视觉方面的沉浸感体验，
20 还可以将与目标对象对应的图像素材叠加至原始图像的第一三维相对位置，以生成原始图像对应的 AR 图像，并显示该 AR 图像（同时不再显示原始图像），以提升用户在视觉方面的体验。

进而，就可以通过终端显示生成的 AR 图像，并通过终端播放对目标声音进行三维效果处理得到的音频。示例地，终端可在用户取景过程中实时显示与抓取到的原始
25 图像对应的 AR 图像，并实时播放经三维效果处理后得到的音频。再例如，终端还可以基于已抓取的各个图像，分别执行本公开提供的音频处理方法，得到各个图像对应的 AR 图像以及各个图像对应的经三维效果处理所得到的音频，合成视频并存储，以便用户后续查看（该功能可设置为由用户触发）。

可选地，由于本公开提供的方法是通过声音处理提升用户在虚拟场景中的沉浸感，而若用户所持终端无法将具有空间感的音频提供给用户，则就没有必要进行音频播放，以避免终端不必要的数据交互。因此，本公开提供的方法还可以包括以下步骤：

确定终端是否符合环绕声播放条件；

- 5 若确定终端符合环绕声播放条件，播放在三维效果处理后得到的音频。

示例地，判断终端是否符合环绕声播放条件，可以通过识别终端是否插入双声道耳机确定。若终端插入双声道耳机，可以确定终端符合环绕声播放条件。

采用上述方式，能够保证本公开对于具有空间感的音频的播放，是在终端能够提供环绕声播放条件的前提下进行的，以避免终端不必要的数据交互。

- 10 图 3 是根据本公开的一种实施方式提供的音频处理装置的框图。如图 3 所示，该装置 30 包括：

第一获取模块 31，用于获取终端抓取到的原始图像；

第一确定模块 32，用于根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；

- 15 音频处理模块 33，用于根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。

可选地，所述第一确定模块 32 包括：

- 20 第一获取子模块，用于对所述原始图像进行特征提取，获取所述原始图像中的第一锚点信息；

判断子模块，用于根据预先存储的对应于所述目标对象的参考锚点信息，判断所述第一锚点信息能否与所述参考锚点信息相匹配；

- 25 第一确定子模块，用于若确定所述第一锚点信息与所述参考锚点信息相匹配，根据所述第一锚点信息，确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。

可选地，所述第一确定模块 32 包括：

- 30 第二确定子模块，用于若获取到用户对所述原始图像进行锚点标记操作而产生的第二锚点信息，根据所述第二锚点信息确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。

可选地，所述第一确定模块 32 包括：

第二获取子模块，用于若所述原始图像中不存在所述目标对象，获取所述目标对象对应的历史运动轨迹信息，所述历史运动轨迹信息是通过对所述终端在所述原始图像前抓取的图像进行运动追踪得到的；

- 5 第三确定子模块，用于根据所述历史运动轨迹信息，确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。

可选地，所述音频处理模块 33 包括：

- 第四确定子模块，用于根据所述第一三维相对位置、以及所述目标对象与所述目标对象对应的音效对象之间的位置关系，确定所述目标声音的声源位置相对于所述终端的三维相对位置，作为第二三维相对位置；
- 10

 音频处理子模块，用于根据所述第二三维相对位置，对所述目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系。

可选地，所述装置 30 还包括：

- 15 第二获取模块，用于获取目标声音的附加描述信息，所述附加描述信息包括以下中的至少一者：目标声音的声源发声面积、目标声音的声源混响效果、目标声音的声源环境；

- 所述音频处理子模块用于根据所述第二三维相对位置和所述附加描述信息，对所述目标声音进行三维效果处理，以使得在三维效果处理后得到的所述音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系。
- 20

可选地，所述装置 30 还包括：

- 图像处理模块，用于将所述目标对象对应的图像素材叠加至所述原始图像的所述第一三维相对位置，以生成所述原始图像对应的 AR 图像；
- 25 图像显示模块，用于显示所述 AR 图像。

可选地，所述装置 30 还包括：

 第二确定模块，用于确定所述终端是否符合环绕声播放条件；

 音频播放模块，用于若确定所述终端符合所述环绕声播放条件，播放在所述三维效果处理后得到的所述音频。

关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

下面参考图 4，其示出了适于用来实现本公开实施例的电子设备 600 的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理，Personal Digital Assistant）、PAD（平板电脑，Portable Android Device）、PMP（便携式多媒体播放器，Personal Multimedia Player）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字电视（TV）、台式计算机等等的固定终端。图 4 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

10 如图 4 所示，电子设备 600 可以包括处理装置（例如中央处理器、图形处理器等）601，其可以根据存储在只读存储器（Read-Only Memory，ROM）602 中的程序或者从存储装置 608 加载到随机访问存储器（Random Access Memory，RAM）603 中的程序而执行各种适当的动作和处理。在 RAM 603 中，还存储有电子设备 600 操作所需的各种程序和数据。处理装置 601、ROM 602 以及 RAM 603 通过总线 604 彼此相连。
15 输入/输出（Input/Output，I/O）接口 605 也连接至总线 604。

通常，以下装置可以连接至 I/O 接口 605：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 606；包括例如液晶显示器（Liquid Crystal Display，LCD）、扬声器、振动器等的输出装置 607；包括例如磁带、硬盘等的存储装置 608；以及通信装置 609。通信装置 609 可以允许电子设备 600 与其他设备
20 进行无线或有线通信以交换数据。虽然图 4 示出了具有各种装置的电子设备 600，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，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在非暂态计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序
25 代码。在这样的实施例中，该计算机程序可以通过通信装置 609 从网络上被下载和安装，或者从存储装置 608 被安装，或者从 ROM 602 被安装。在该计算机程序被处理装置 601 执行时，执行本公开实施例的方法中限定的上述功能。

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是——
30

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（Electrical Programmable ROM, EPROM 或闪存）、5 光纤、便携式紧凑磁盘只读存储器（Compact Disc ROM, CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，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10 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介质，该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、RF（射频，Radio Frequency）等等，或者上述的任15 意合适的组合。

在一些实施方式中，客户端、服务器可以利用诸如 HTTP（HyperText Transfer Protocol，超文本传输协议）之类的任何当前已知或未来研发的网络协议进行通信，并且可以与任意形式或介质的数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（Local Area Network，“LAN”），广域网（Wide Area Network，“WAN”），20 网际网（例如，互联网）以及端对端网络（例如，自组织（ADaptive Heuristic for Opponent Classification，ad hoc）端对端网络），以及任何当前已知或未来研发的网络。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

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被该电子设备执行时，使得该电子设备：获取终端抓取到的原始图像；根据所述原始图像，25 确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。30 音。

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但不限于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5 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）或广域网（WAN）——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）。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框10 可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图15 和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的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。其中，模块的名称在某种情况下并不构成对该模块本身的限定，例如，20 如，第一获取模块还可以被描述为“获取终端抓取到的原始图像的模块”。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列（Field Programmable Gate Array, FPGA）、专用集成电路（Application Specific Integrated Circuit, ASIC）、专用标准产品（Application Specific Standard Parts, ASSP）、片上系25 统（System on Chip, SOC）、复杂可编程逻辑设备（Complex Programming Logic Device, CPLD）等等。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括30 但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，

或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦除可编程只读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

- 5 根据本公开的一个或多个实施例，提供了一种音频处理方法，所述方法包括：
 获取终端抓取到的原始图像；
 根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；
 根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果
- 10 处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。
- 根据本公开的一个或多个实施例，提供了一种音频处理方法，所述根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置，包括：
- 15 对所所述原始图像进行特征提取，获取所述原始图像中的第一锚点信息；
 根据预先存储的对应于所述目标对象的参考锚点信息，判断所述第一锚点信息能否与所述参考锚点信息相匹配；
 若确定所述第一锚点信息与所述参考锚点信息相匹配，根据所述第一锚点信息，确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。
- 20 根据本公开的一个或多个实施例，提供了一种音频处理方法，所述根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置，包括：
 若获取到用户对所述原始图像进行锚点标记操作而产生的第二锚点信息，根据所述第二锚点信息确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。
- 25 根据本公开的一个或多个实施例，提供了一种音频处理方法，所述根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置，包括：
 若所述原始图像中不存在所述目标对象，获取所述目标对象对应的历史运动轨迹信息，所述历史运动轨迹信息是通过对所述终端在所述原始图像前抓取的图像进行运动追踪得到的；

根据所述历史运动轨迹信息，确定所述目标对象相对于所述终端的三维相对位置，作为所述第一三维相对位置。

- 5 根据本公开的一个或多个实施例，提供了一种音频处理方法，所述根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，包括：

根据所述第一三维相对位置、以及所述目标对象与所述目标对象对应的音效对象之间的位置关系，确定所述目标声音的声源位置相对于所述终端的三维相对位置，作为第二三维相对位置；

- 10 根据所述第二三维相对位置，对所述目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系。

根据本公开的一个或多个实施例，提供了一种音频处理方法，所述方法还包括：

- 15 获取目标声音的附加描述信息，所述附加描述信息包括以下中的至少一者：目标声音的声源发声面积、目标声音的声源混响效果、目标声音的声源环境；

所述根据所述第二三维相对位置，对所述目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，包括：

- 20 根据所述第二三维相对位置和所述附加描述信息，对所述目标声音进行三维效果处理，以使得在三维效果处理后得到的所述音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系。

根据本公开的一个或多个实施例，提供了一种音频处理方法，所述方法还包括：

- 25 将所述目标对象对应的图像素材叠加至所述原始图像的所述第一三维相对位置，以生成所述原始图像对应的 AR 图像；
显示所述 AR 图像。

根据本公开的一个或多个实施例，提供了一种音频处理方法，所述方法还包括：
确定所述终端是否符合环绕声播放条件；

若确定所述终端符合所述环绕声播放条件，播放在所述三维效果处理后得到的所述音频。

- 30 根据本公开的一个或多个实施例，提供了一种音频处理装置，所述装置包括：

第一获取模块，用于获取终端抓取到的原始图像；

第一确定模块，用于根据所述原始图像，确定目标对象相对于所述终端的三维相对位置，作为第一三维相对位置；

- 5 音频处理模块，用于根据所述第一三维相对位置，对目标声音进行三维效果处理，以使得在三维效果处理后得到的音频中所述目标声音的声源位置与所述第一三维相对位置符合所述目标对象与所述目标对象对应的音效对象之间的位置关系，所述目标声音为所述音效对象对应的音效声音。

根据本公开的一个或多个实施例，提供了一种计算机可读介质，其上存储有计算机程序，该程序被处理装置执行时实现本公开任意实施例所述方法的步骤。

- 10 根据本公开的一个或多个实施例，提供了一种电子设备，包括：

存储装置，其上存储有计算机程序；

处理装置，用于执行所述存储装置中的所述计算机程序，以实现本公开任意实施例所述方法的步骤。

根据本公开的一个或多个实施例，提供了一种计算机程序产品，程序产品包括：

- 15 计算机程序，该计算机程序被处理装置执行时实现本公开任意实施例所述方法的步骤。

根据本公开的一个或多个实施例，提供了一种计算机程序，该计算机程序被处理装置执行时实现本公开任意实施例所述方法的步骤。

- 20 以上描述仅为本公开的较佳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。

- 25 此外，虽然采用特定次序描绘了各操作，但是这不应理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

5