



# (12)发明专利申请

(10)申请公布号 CN 108769535 A

(43)申请公布日 2018. 11. 06

(21)申请号 201810723144.8

(22)申请日 2018.07.04

(71)申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72)发明人 周景锦 程培 傅斌 高雨

(74)专利代理机构 广州华进联合专利商标代理  
有限公司 44224

代理人 李文渊 何平

(51) Int. Cl.

H04N 5/232(2006.01)

G06T 3/00(2006.01)

G06K 9/00(2006.01)

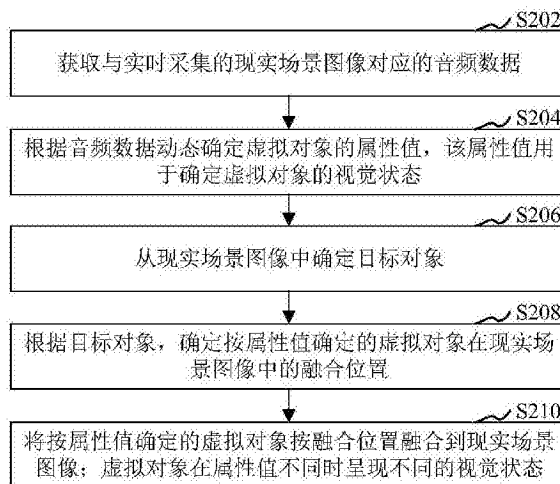
权利要求书3页 说明书17页 附图9页

## (54)发明名称

图像处理方法和装置、存储介质和计算机设备

## (57)摘要

本申请涉及一种图像处理方法和装置、存储介质和计算机设备,所述方法包括:获取与实时采集的实时场景图像对应的音频数据;根据所述音频数据动态确定虚拟对象的属性值,所述属性值用于确定所述虚拟对象的视觉状态;从所述现实场景图像中确定目标对象;根据所述目标对象,确定按所述属性值确定的虚拟对象在所述现实场景图像中的融合位置;将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像;所述虚拟对象在属性值不同时呈现不同的视觉状态。本申请提供的方案可以提升用户与虚拟对象之间的交互性。



1. 一种图像处理方法,包括:

获取与实时采集的现实场景图像对应的音频数据;

根据所述音频数据动态确定虚拟对象的属性值,所述属性值用于确定所述虚拟对象的视觉状态;

从所述现实场景图像中确定目标对象;

根据所述目标对象,确定按所述属性值确定的虚拟对象在所述现实场景图像中的融合位置;

将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像;所述虚拟对象在属性值不同时呈现不同的视觉状态。

2. 根据权利要求1所述的方法,其特征在于,所述获取与实时采集的现实场景图像对应的音频数据包括:

在实时采集现实场景图像时,从当前环境中实时采集与所述现实场景图像对应的音频数据;或者,

从实时采集的现实场景图像的背景音频中,读取与所述现实场景图像所对应时间戳对应的音频数据。

3. 根据权利要求1所述的方法,其特征在于,所述根据所述音频数据动态确定虚拟对象的属性值包括:

获取所述音频数据的参数值;

确定参数值与虚拟对象的属性值之间的预设映射关系;

根据所述预设映射关系,将所述参数值映射为虚拟对象的属性值。

4. 根据权利要求3所述的方法,其特征在于,所述获取所述音频数据的参数值包括:

对所述音频数据进行抽样;

将抽样所得的结果量化并编码,获得编码音频数据;

根据所获得的编码音频数据确定所述音频数据的参数值。

5. 根据权利要求3所述的方法,其特征在于,所述参数值包括频率值;所述根据所获得的编码音频数据确定所述音频数据的参数值包括:

将时域的所述编码音频数据转换为频域音频数据;

将所述频域音频数据分段,获得多个子频域音频数据;

确定各子频域音频数据的振幅;

从各子频域音频数据中选取振幅最大的子频域音频数据;

按照所选取的子频域音频数据确定所述音频数据对应的频率值。

6. 根据权利要求3所述的方法,其特征在于,所述参数值包括音量值;所述根据所获得的编码音频数据确定所述音频数据的参数值包括:

根据所获得的编码音频数据确定音量值;或者,

将时域的所述编码音频数据转换为频域音频数据;根据所述频域音频数据确定音量值。

7. 根据权利要求1所述的方法,其特征在于,所述属性值包括属性调整量;所述按所述属性值确定的虚拟对象,是按所述属性调整量调整相应属性后的虚拟对象;

所述将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像包括:

确定所述虚拟对象所具有的、且与所述属性调整量对应的属性；  
按照所述属性调整量调整所述虚拟对象的所述属性，得到调整属性后的虚拟对象；  
将调整属性后的虚拟对象按所述融合位置融合到所述现实场景图像。

8. 根据权利要求1所述的方法，其特征在于，所述属性值包括属性变化目标值；所述按所述属性值确定的虚拟对象，是相应属性变化至所述属性变化目标值后的虚拟对象；

所述将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像包括：  
确定所述虚拟对象所具有的、且与所述属性变化目标值对应的属性；

将所述虚拟对象的所述属性变化至所述属性变化目标值，得到属性变化后的虚拟对象；

将属性变化后的虚拟对象按所述融合位置融合到所述现实场景图像。

9. 根据权利要求1至8任一项所述的方法，其特征在于，所述从所述现实场景图像中确定目标对象包括：

从所述现实场景图像中识别生物特征；

当所述生物特征满足预设条件时，则

将所述现实场景图像中与所述生物特征对应的生物对象确定为目标对象。

10. 根据权利要求1至8任一项所述的方法，其特征在于，所述根据所述目标对象，确定具有所述属性的虚拟对象在所述现实场景图像中的融合位置包括：

检测所述目标对象的特征；

在所检测到的特征中查找与具有所述属性的虚拟对象匹配的特征；

根据所述匹配的特征，确定具有所述属性的虚拟对象在所述现实场景图像中的融合位置。

11. 根据权利要求1至8任一项所述的方法，其特征在于，还包括：

提取所述音频数据的音频特征；

当所述音频特征符合第一触发条件时，则执行以下至少一种：

新增虚拟对象；

切换虚拟对象；

切换所述视觉状态的类型。

12. 根据权利要求1至8任一项所述的方法，其特征在于，还包括：

根据所述音频数据进行识别，获得识别结果；

确定与所述识别结果匹配的动态效果类型；

按照所述动态效果类型和所述属性值确定所述虚拟对象所呈现的视觉状态；所述视觉状态与所述动态效果类型相匹配。

13. 一种图像处理装置，包括：

音频数据获取模块，用于获取与实时采集的现实场景图像对应的音频数据；

属性值确定模块，用于根据所述音频数据动态确定虚拟对象的属性值，所述属性值用于确定所述虚拟对象的视觉状态；

目标对象确定模块，用于从所述现实场景图像中确定目标对象；

融合位置确定模块，用于根据所述目标对象，确定按所述属性值确定的虚拟对象在所述现实场景图像中的融合位置；

融合模块,用于将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像;所述虚拟对象在属性值不同时呈现不同的视觉状态。

14. 根据权利要求13所述的装置,其特征在于,所述音频数据获取模块还用于在实时采集现实场景图像时,从当前环境中实时采集与所述现实场景图像对应的音频数据;或者,从实时采集的现实场景图像的背景音频中,读取与所述现实场景图像所对应时间戳对应的音频数据。

15. 根据权利要求14所述的装置,其特征在于,属性值确定模块还用于获取所述音频数据的参数值;确定参数值与虚拟对象的属性值之间的预设映射关系;根据所述预设映射关系,将所述参数值映射为虚拟对象的属性值。

16. 一种存储介质,存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行如权利要求1至12中任一项所述方法的步骤。

17. 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,所述计算机程序被所述处理器执行时,使得所述处理器执行如权利要求1至12中任一项所述方法的步骤。

## 图像处理方法、装置、存储介质和计算机设备

### 技术领域

[0001] 本申请涉及图像处理技术领域,特别是涉及一种图像处理方法、装置、存储介质和计算机设备。

### 背景技术

[0002] 随着图像处理技术和计算机技术的快速发展,出现了各式各样的用于录制视频的应用,比如用户可通过客户端录制虚拟和现实相结合的视频。

[0003] 目前在录制视频时,用户可在客户端的录制界面中自由选取虚拟对象,将这些虚拟对象添加到视频对应图像帧的相应位置,使得虚拟对象可跟随视频中运动目标的移动而移动。

[0004] 然而,通过上述方式所录制的视频,虚拟对象仅能够跟随运动目标的移动而移动,交互性差。

### 发明内容

[0005] 基于此,有必要针对虚拟对象仅能够跟随运动目标的移动而移动,使交互性差的技术问题,提供一种图像处理方法、装置、存储介质和计算机设备。

[0006] 一种图像处理方法,包括:

[0007] 获取与实时采集的现实场景图像对应的音频数据;

[0008] 根据所述音频数据动态确定虚拟对象的属性值,所述属性值用于确定所述虚拟对象的视觉状态;

[0009] 从所述现实场景图像中确定目标对象;

[0010] 根据所述目标对象,确定按所述属性值确定的虚拟对象在所述现实场景图像中的融合位置;

[0011] 将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像;所述虚拟对象在属性值不同时呈现不同的视觉状态。

[0012] 一种图像处理装置,包括:

[0013] 音频数据获取模块,用于获取与实时采集的现实场景图像对应的音频数据;

[0014] 属性值确定模块,用于根据所述音频数据动态确定虚拟对象的属性值,所述属性值用于确定所述虚拟对象的视觉状态;

[0015] 目标对象确定模块,用于从所述现实场景图像中确定目标对象;

[0016] 融合位置确定模块,用于根据所述目标对象,确定按所述属性值确定的虚拟对象在所述现实场景图像中的融合位置;

[0017] 融合模块,用于将按所述属性值确定的虚拟对象按所述融合位置融合到所述现实场景图像;所述虚拟对象在属性值不同时呈现不同的视觉状态。

[0018] 一种存储介质,存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行上述的图像处理方法的步骤。

[0019] 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,所述计算机程序被所述处理器执行时,使得所述处理器执行上述的图像处理方法的步骤。

[0020] 上述的图像处理方法、装置、存储介质和计算机设备,获取与实时采集的实时场景图像对应的音频数据,通过音频数据动态确定虚拟对象的属性,从而实现了根据音频数据对虚拟对象属性的控制。通过实时场景图像中的目标对象,确定虚拟对象在实时场景图像中的融合位置,按照融合位置将根据属性值确定的虚拟对象融合到实时场景图像中,由于虚拟对象的属性值由音频数据控制,当音频数据发生变化时,融合到实时场景图像中的虚拟对象的属性值也随之产生变化,提高了交互性。

## 附图说明

[0021] 图1为一个实施例中应用图像处理方法的系统结构图;

[0022] 图2为一个实施例中图像处理方法的流程示意图;

[0023] 图3为一个实施例中将虚拟对象融合到实时场景图像的示意图;

[0024] 图4为一个实施例中将虚拟对象融合到实时场景图像的示意图;

[0025] 图5为一个实施例中根据音频数据确定属性值,将具有该属性值的虚拟对象融合到实时场景图像的示意图;

[0026] 图6为一个实施例中根据音频数据的参数值确定虚拟对象的属性值的步骤的流程示意图;

[0027] 图7为一个实施例中对音频数据进行抽样、量化和编码的示意图;

[0028] 图8为一个实施例中根据编码音频数据确定频率值的步骤的流程示意图;

[0029] 图9为一个实施例中按照属性调整量调整虚拟对象,并将调整后的虚拟对象融合到实时场景图像的步骤的流程示意图;

[0030] 图10为一个实施例中按照属性变化目标值调整虚拟对象,并将调整后的虚拟对象融合到实时场景图像的步骤的流程示意图;

[0031] 图11为一个实施例中根据目标对象的特征确定虚拟对象在实时场景图像的融合位置的步骤的流程示意图;

[0032] 图12为一个实施例中目标对象脸部特征点的示意图;

[0033] 图13为另一个实施例中图像处理方法的流程示意图;

[0034] 图14为一个实施例中图像处理装置的结构框图;

[0035] 图15为一个实施例中计算机设备的结构框图。

## 具体实施方式

[0036] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0037] 图1为一个实施例中图像处理方法的应用环境图。参照图1,该图像处理方法应用于图像处理系统。该图像处理系统可以是一个终端或多个终端的组合,该终端可以是智能手机、电脑或其它可支持AR (Augmented Reality,增强现实) 技术的设备。如图1所示,图像处理系统可以包括:摄像头、场景产生器、图像合成器和显示器;其中:

[0038] 摄像头,用于获取目标对象对应环境的现实场景图像,将获取到的现实场景图像发送至图像合成器,以与增强现实模型的虚拟对象进行合成操作。

[0039] 场景产生器,用于根据图像数据中目标对象的位置信息,如通过分析目标对象的特征确定虚拟对象的融合位置,然后将虚拟对象发送至图像合成器。

[0040] 图像合成器,用于将关于目标对象的现实场景图像和虚拟对象,按照融合位置进行融合,将融合的结果输出至显示屏。

[0041] 显示器,用于将图像合成器输出的融合图像进行显示,形成目标对象和用于增强现实模型的虚拟对象共同显示的效果。

[0042] 如图2所示,在一个实施例中,提供了一种图像处理方法。本实施例主要以该方法应用于上述图1中的终端来举例说明。参照图2,该图像处理方法具体包括如下步骤:

[0043] S202,获取与实时采集的现实场景图像对应的音频数据。

[0044] 其中,现实场景可以指用户通过某种媒介所观看到的现实画面,在该现实画面中至少包括以下之一:人物、自然风景、人文风景和人类的智慧作品等。人类的智慧作品指的是人类通过劳动和智慧所创造的作品。举例来说,现实场景为用户通过肉眼所观看的人与自然的画面,或者用户通过3D眼镜所观看立体电影的画面。现实场景图像可以通过终端采集关于现实场景的图像。例如,现实场景图像为通过图1中摄像头实时采集现实场景的图像。终端采集多个现实场景图像后,将多个现实场景图像按照采集时间进行组合,可得到一个视频。

[0045] 音频数据为时域的音频信号,音频数据中携带有语音、音乐和音效的有规律的声波的频率、幅度变化信息等。根据声波的特征,可把音频数据分类为规则音频和不规则音频。其中规则音频又可以分为语音、音乐和音效。音频数据其来源可以是终端从外界采集所得,也可以是从与现实场景图像的背景音频中读取所得。

[0046] 在一个实施例中,S202具体可以包括:在实时采集现实场景图像时,终端从当前环境中实时采集与现实场景图像对应的音频数据;或者,终端从实时采集的现实场景图像的背景音频中,读取与现实场景图像所对应时间戳对应的音频数据。

[0047] 例如,在实时采集现实场景图像时,终端通过麦克风采集当前环境中的音乐或普通说话的声音。或者,由于在开发过程中,开发人员设置在实时采集的现实场景图像时自动播放背景音乐,终端在自动播放的背景音乐中读取与现实场景图像所对应时间戳对应的音频数据。或者,在实时采集的现实场景图像时,音乐客户端同时在播放音乐,终端则在播放的音乐中读取与现实场景图像所对应时间戳对应的音频数据。

[0048] S204,根据音频数据动态确定虚拟对象的属性值,该属性值用于确定虚拟对象的视觉状态。

[0049] 其中,虚拟对象可以是图像素材,该图像素材如静态贴纸或动态贴纸等。此外,虚拟对象还可以是用于增强目标对象显示效果的虚拟道具。例如,虚拟道具可以是用于装扮目标对象的各种虚拟挂件和虚拟背景。属性值可以包括:属性调整量和属性变化目标值。属性调整量包括虚拟对象的缩放比例、旋转角度和偏移量等。属性变化目标值包括虚拟对象的颜色RGB值。

[0050] 需要说明的是,所确定虚拟对象的属性值为当前的属性值。例如,假设用户拍摄一个具有AR效果的视频,若视频中的第*i*帧现实场景图像为当前的现实场景图像,第*i*帧现实

场景图像对应时刻的虚拟对象的属性值为当前的属性值。视频中的第 $i-1$ 帧现实场景图像为上一时刻的现实场景图像,第 $i-1$ 帧现实场景图像对应时刻的虚拟对象的属性值为上一时刻的属性值。其中, $i$ 为大于等于1的正整数。

[0051] 在一个实施例中,终端根据音频数据动态确定虚拟对象的属性值,根据确定的属性值调整虚拟对象,获得按属性值确定的虚拟对象。

[0052] 在一个实施例中,音频数据具有音量值、频率值和音色等参数值。终端根据音频数据动态确定虚拟对象的属性值。当终端将虚拟对象的原始属性值更新为所确定的属性值时,便可实现对虚拟对象的调整。例如,终端根据音频数据的频率值来调整虚拟对象的缩放比例,或调整虚拟对象的颜色RGB值。

[0053] S206,从现实场景图像中确定目标对象。

[0054] 其中,目标对象可以是自然界中具有生命的物体,如人类。

[0055] 在一个实施例中,S206具体可以包括:从现实场景图像中识别生物特征;当生物特征满足预设条件时,则将现实场景图像中与生物特征对应的生物对象确定为目标对象。

[0056] 其中,生物特征可以是生物的轮廓特征,也可以是生物的细节特征,如人的脸部特征等。

[0057] 在一个实施例中,预设条件可以是预设生物特征或生物特征的预设完整度阈值。终端从现实场景图像中识别生物特征;当生物特征满足预设生物特征,和/或生物特征的完整度达到预设完整度阈值时,则将现实场景图像中与生物特征对应的对象确定为目标对象。

[0058] S208,根据目标对象,确定按属性值确定的虚拟对象在现实场景图像中的融合位置。

[0059] 其中,融合位置指的是:在融合到现实场景图像时,虚拟对象的中心点在现实场景图像中所处的位置。融合位置可以是一个位置点,也可以是多个位置点。不同的虚拟对象可以融合在目标对象的不同部位。

[0060] 在一个实施例中,终端检测目标对象的特征,从目标对象的特征中选取与虚拟对象匹配的特征,根据所选取的特征确定按属性值确定的虚拟对象在现实场景图像中的融合位置。

[0061] 作为一个示例,如图3所示,当虚拟对象为AR潜水眼镜的虚拟道具时,终端检测目标对象的特征,从检测到的特征中确定眼部特征,根据眼部特征确定虚拟对象的融合位置为用户的眼部位置。

[0062] 作为另一个示例,如图4所示,当虚拟对象为AR兔子装扮的虚拟道具时,对于AR兔子装扮中的兔牙虚拟道具,终端检测目标对象的特征,从检测到的特征中确定嘴部特征,根据嘴部特征确定虚拟对象的融合位置为用户的嘴部位置。对于AR兔子装扮中的耳朵虚拟道具,终端检测目标对象的特征,从检测到的特征中确定头部特征(如头发),根据头部特征确定虚拟对象的融合位置为用户的头部位置。

[0063] 在一个实施例中,终端按属性值调整的虚拟对象,从而使虚拟对象进行尺寸的缩放、或更换颜色RGB值、或改变旋转角度等。需要说明的是,终端按属性值调整的虚拟对象,是以上一帧融合后的现实场景图像中的虚拟对象为参考,从而实现当前虚拟对象进行缩放、或更换颜色RGB值、或改变旋转角度等操作。



[0064] S210, 将按属性值确定的虚拟对象按融合位置融合到现实场景图像; 虚拟对象在属性值不同时呈现不同的视觉状态。

[0065] 在一个实施例中, 终端确定虚拟对象的中心点或关键位置点, 将虚拟对象的中心点或关键位置点融合至与融合位置匹配的区域。其中, 关键位置点用于确定虚拟对象所要融合的区域。中心点用于与融合位置匹配, 即在融合时, 将虚拟对象的中心点置于在融合位置。

[0066] 如图3所示,  $p$ 表示虚拟对象的关键位置点, 该关键位置点为AR潜水眼镜的虚拟道具的三个位置。在将AR潜水眼镜的虚拟道具融合到融合在现实场景图像中时, 根据图中3中的 $p$ 点将AR潜水眼镜的虚拟道具对准目标对象的两眼和鼻子, 从而实现AR潜水眼镜的虚拟道具的准确融合, 具体效果如图3 (b) 所示。

[0067] 如图4所示,  $q$ 表示虚拟对象的关键位置点, 该关键位置点在AR兔子装扮的兔牙虚拟道具中的上部, 以及, 在AR兔子装扮的耳朵虚拟道具中下部。在将AR兔子装扮的虚拟道具融合到现实场景图像时, 根据该关键位置点将AR兔子装扮的虚拟道具融合到目标对象的嘴唇部位和头部, 从而实现AR兔子装扮的虚拟道具的准确融合, 具体效果如图4 (b) 所示。

[0068] 虚拟对象在属性值不同时呈现不同的视觉状态, 举例来说, 假设用户拍摄一个具有AR效果的视频, 若视频中的第 $i$ 帧现实场景图像为当前的现实场景图像, 第 $i$ 帧现实场景图像对应时刻的虚拟对象的属性值为当前的属性值。视频中的第 $i-1$ 帧现实场景图像为上一时刻的现实场景图像, 第 $i-1$ 帧现实场景图像对应时刻的虚拟对象的属性值为上一时刻的属性值。其中,  $i$ 为大于等于1的正整数。当各帧现实场景图像和对应属性值的虚拟对象按照时间组合, 那么可获得一个具有AR效果的视频, 在该视频中, 虚拟对象的属性值随着音频数据的变化而发生变化。如音频数据的音量值或频率值发生变化, 虚拟对象的尺寸、或颜色RGB值、或方位随之产生变化。

[0069] 作为一个示例, 如图5所示, 假设图5 (a) 为上一时刻融合后的现实场景图像,  $m$ 为为原始尺寸的虚拟对象。当所获取的音频数据的参数值发生变化, 如音量大小发生变化, 或频率值发生变化, 那么, 对应的虚拟对象的尺寸也发生变化, 变化后的虚拟对象如图5 (b) 中的 $n$ 所示,  $n$ 为尺寸放大的虚拟对象。

[0070] 上述实施例中, 获取与实时采集的现实场景图像对应的音频数据, 通过音频数据动态确定虚拟对象的属性, 从而实现了根据音频数据对虚拟对象属性的控制。通过现实场景图像中的目标对象, 确定虚拟对象在现实场景图像中的融合位置, 按照融合位置将根据属性值确定的虚拟对象融合到现实场景图像中, 由于虚拟对象的属性值由音频数据控制, 当音频数据发生变化时, 融合到现实场景图像中的虚拟对象的属性值也随之产生变化, 提高了交互性。

[0071] 在一个实施例中, 如图6所示, S204具体可以包括:

[0072] S602, 获取音频数据的参数值。

[0073] 其中, 音频数据的参数值包括音频数据的音量值、频率值和音色。音量值可以是以下任一种: 平均频率值、最大音量值或最小音量值。频率值可以是以下任一种: 平均频率值、最大频率值或最小频率值。

[0074] 具体地, 终端通过分析音频数据, 获得音频数据中的音量值、频率值和音色等参数值。

[0075] 在一个实施例中,S602具体可以包括:对音频数据进行抽样;将抽样所得的结果量化并编码,获得编码音频数据;根据所获得的编码音频数据确定音频数据的参数值。

[0076] 具体地,终端采用PCM((Pulse Code Modulation,脉冲编码调制)的方式,对所采集的时间上连续的时域音频数据进行抽样、量化和编码,获得二进制的编码音频数据。终端根据所获得的编码音频数据确定音频数据的参数值,如确定音频数据的音量值。

[0077] 作为一个示例,如图7所示,对音频数据 $U(t)$ 进行抽样,使时间上连续的时域音频数据离散化。将离散化的音频数据进行量化,获得M进制的量化音频数据,其中,M为大于2的正整数。将量化后的音频数据进行编码,得到二进制的编码音频数据。

[0078] S604,确定参数值与虚拟对象的属性值之间的预设映射关系。

[0079] 在一个实施例中,不同类型的参数值对应不同的预设映射关系,终端根据参数值的类型,确定与虚拟对象的属性值之间的预设映射关系。其中,类型包括音量类型、频率类型和音色类型。与类型对应的,参数值包括音量值、频率值和音色。

[0080] 例如,以参数值为音量值,以属性值为缩放比例为例,终端确定音量与虚拟对象的缩放比例之间的预设映射关系,该预设映射关系可以是函数关系式,如下所示:

$$[0081] \quad f(x) = \begin{cases} 1.0, & x < 50 \\ 1.0 + 3.0 \times \frac{x-50}{70}, & x \geq 50 \end{cases}$$

[0082] 其中,x为音频数据的音量值,范围可以是0~120分贝(db), $f(x)$ 为虚拟对象的缩放比例。

[0083] S606,根据预设映射关系,将参数值映射为虚拟对象的属性值。

[0084] 在一个实施例中,终端确定预设映射关系时,将参数值作为变量输入该预设映射关系,得到对于虚拟对象的对应属性值。例如,当音量值为40db时,将音量值40输入函数 $f(x)$ ,得到的属性值为1的缩放比例,即对原始的虚拟对象不做任何放大或缩小处理。又例如,当音量值为120db时,将音量值120输入函数 $f(x)$ ,得到的属性值为4的缩放比例,即将原始的虚拟对象放大到原来的4倍。可以看出,虚拟对象的属性值随着音量数据的参数值变化而变化,从而根据音量数据的参数值实现对虚拟对象的控制,从而使虚拟对象呈现不同的视觉状态。

[0085] 上述实施例中,确定参数值与虚拟对象的属性值之间的预设映射关系,当获得对应的参数值的大小时,便可以通过该预设映射关系获得对应的虚拟对象的属性值,从而实现对虚拟对象的调整,使虚拟对象呈现不同的视觉状态,提升了虚拟对象的多样化变化。

[0086] 在一个实施例中,虚拟对象的属性值可以通过音频数据的参数值确定,该参数值可以是频率值或音量值,当参数值为频率值时;如图8所示,上述根据所获得的编码音频数据确定音频数据的参数值的步骤,具体可以包括:

[0087] S802,将时域的编码音频数据转换为频域音频数据。

[0088] 在一个实施例中,编码音频数据为离散化的音频信号。终端根据离散傅里叶变换,将编码音频数据转换为频域音频数据。其中,频域音频数据包含有音频数据的幅值(即音量值)、频率值和相位。

[0089] S804,将频域音频数据分段,获得多个子频域音频数据。

[0090] 在一个实施例中,终端按照设定的步长,对频域音频数据进行分段,获得多个子频

域音频数据。例如,使用512点傅里叶变换时,最多可以将0到截止频率(若抽样率48kHz时,截止频率为24kHz)的频段等分为256个频段,然后执行S806,即确定各频段内的振幅。

[0091] 在一个实施例中,终端按照不等分的方式,将频域音频数据划分为不等长的多个频段,获得多个子频域音频数据。

[0092] S806,确定各子频域音频数据的振幅。

[0093] 在一个实施例中,各子频域音频数据中包含了幅值、频率值和相位,终端确定各子频域音频数据中的振幅,从而获取到各子频域音频数据中的音量值。

[0094] S808,从各子频域音频数据中选取振幅最大的子频域音频数据。

[0095] 其中,振幅大表示音频数据的功率较大,对于所获取的音频数据而言,当音频数据的功率大对应的有用信号多。例如,终端通过麦克风采集音频数据,当所采集的音频数据的功率小,说明当前所采集的音频数据可能为噪声信号,因此,可以选取振幅最大的子频域音频数据。

[0096] 具体地,当确定各子频域音频数据的振幅时,将各子频域音频数据之间的振幅进行比较,获取振幅最大的子频域音频数据。

[0097] 在一个实施例中,终端对各子频域音频数据按照振幅的大小进行排列,从排列的各子频域音频数据中选取振幅最大的子频域音频数据。

[0098] S810,按照所选取的子频域音频数据确定音频数据对应的频率值。

[0099] 上述实施例中,对频域音频数据进行分段,可以通过分段后所得子频域音频数据确定频率值,一方面,通过该频率值可以实现对虚拟对象的调整,另一方面,分段后的频域音频数据,在计算过程中可降低计算量,加快计算速率。

[0100] 在一个实施例中,虚拟对象的属性值可以通过音频数据的参数值确定,该参数值可以是频率值或音量值,当参数值为音量值时;如图8所示,根据所获得的编码音频数据确定音频数据的参数值的步骤,具体可以包括:根据所获得的编码音频数据确定音量值;或者,将时域的编码音频数据转换为频域音频数据;根据频域音频数据确定音量值。

[0101] 时域的编码音频数据的幅值可以表示音频数据的音量值。在一个实施例中,终端将时域的编码音频数据的幅值,确定为音频数据的音量值。

[0102] 在一个实施例中,编码音频数据为离散化的音频信号。终端根据离散傅里叶变换,将编码音频数据转换为频域音频数据。其中,频域音频数据包含有音频数据对应波形的幅值(即音量值)、频率值和相位。终端将频域音频数据中的幅值确定为音频数据的音量值。

[0103] 在一个实施例中,终端将编码音频数据转换为频域音频数据后,按照设定的步长对频域音频数据进行分段,获得多个子频域音频数据。终端根据各子频域音频数据确定对应的幅值,将最大的幅值确定为音频数据的音量值,或者,将平均幅值确定为音频数据的音量值。

[0104] 上述实施例中,根据所获得的编码音频数据,或者根据由时域的编码音频数据转换的频域音频数据这两种方式确定音量值,得到用于确定虚拟对象视觉状态的属性值,通过该音量值可以实现对虚拟对象的调整。

[0105] 在一个实施例中,虚拟对象的属性值可以是属性调整量和属性变化目标值。属性调整量可以是虚拟对象的缩放比例、旋转角度和偏移量等,属性变化目标值可以是虚拟对象的颜色RGB值。当属性值为属性调整量时,上述的按属性值确定的虚拟对象,是按属性调

整量调整相应属性后的虚拟对象,如图9所示,S210具体可以包括:

[0106] S902,确定虚拟对象所具有的、且与属性调整量对应的属性。

[0107] 其中,属性指的是虚拟对象所拥有的属性,包括缩放、颜色、旋转和偏移等。与属性调整量对应的属性包括缩放、旋转和偏移等。对应的,属性调整量为属性对应的具体值。

[0108] 在一个实施例中,终端根据音频数据的参数值,确定虚拟对象所具有的、且与属性调整量对应的属性。

[0109] S904,按照属性调整量调整虚拟对象的属性,得到调整属性后的虚拟对象。

[0110] 在一个实施例中,终端根据映射关系,确定与音频数据的参数值对应的属性调整量后,按照属性调整量调整虚拟对象,使虚拟对象的属性发生变化,得到调整属性后的虚拟对象。

[0111] 例如,若属性调整量为缩放比例,根据缩放比例调整虚拟对象的尺寸大小,获得调整尺寸后的虚拟对象。

[0112] S906,将调整属性后的虚拟对象按融合位置融合到现实场景图像。

[0113] 上述实施例中,通过属性调整量调整虚拟对象的属性,将调整属性后的虚拟对象按融合位置融合到现实场景图像,获得随音频数据的参数变化的虚拟对象,实现根据音频数据调整虚拟对象,提升了虚拟对象的多样化变化,增强用户的体验。

[0114] 在一个实施例中,虚拟对象的属性值可以是属性调整量和属性变化目标值。属性调整量可以是虚拟对象的缩放比例、旋转角度和偏移量等,属性变化目标值可以是虚拟对象的颜色。当属性值为属性变化目标值时,上述的按属性值确定的虚拟对象,是相应属性变化至属性变化目标值后的虚拟对象,如图10所示,S210具体可以包括:

[0115] S1002,确定虚拟对象所具有的、且与属性变化目标值对应的属性。

[0116] 其中,与属性变化目标值对应的属性包括虚拟对象的颜色。对应的,属性变化目标值为属性对应的具体值,如颜色RGB值。

[0117] 在一个实施例中,终端根据音频数据的参数值,确定虚拟对象所具有的、且与属性变化目标值对应的属性。

[0118] S1004,将虚拟对象的属性变化至属性变化目标值,得到属性变化后的虚拟对象。

[0119] 在一个实施例中,终端根据映射关系,确定与音频数据的参数值对应的属性变化目标值后,按照属性变化目标值调整虚拟对象,使虚拟对象的属性发生变化,得到调整属性后的虚拟对象。

[0120] 例如,若属性变化目标值为目标颜色RGB值,根据目标颜色RGB值调整虚拟对象的显示颜色,使虚拟对象的原始显示颜色调整为目标颜色RGB值所对应的颜色,如原始为红色,通过目标颜色RGB值的调整后获得蓝色的虚拟对象。

[0121] S1006,将属性变化后的虚拟对象按融合位置融合到现实场景图像。

[0122] 上述实施例中,通过属性变化目标值调整虚拟对象的属性,将调整属性后的虚拟对象按融合位置融合到现实场景图像,获得随音频数据的参数变化的虚拟对象,实现根据音频数据调整虚拟对象,提升了虚拟对象的多样化变化,增强用户的体验。

[0123] 在一个实施例中,如图11所示,S208具体可以包括:

[0124] S1102,检测目标对象的特征。

[0125] 具体地,终端通过特征点检测方式,如级联回归CNN、或Dlib、Libfacedetect、或

Seetaface等特征点检测方式检测目标对象的特征。

[0126] 作为一个示例,如图12所示,为目标对象的脸部特征点的检测结果,为了描述方便,采用数字标记识别得到的各个脸部特征点,例如图12中所示的1~17表示脸部边缘特征点,18~22以及23~27对应表示用户的左眉部特征点和右眉部特征点,28~36表示用户的鼻子特征点,37~42表示用户的左眼特征点,43~48表示用户的右眼特征点,49~68表示用户的嘴唇特征点。需要指出的是,以上仅为示例,在可选实施例中可以在以上脸部特征点中仅识别部分或更多的特征点,或采用其他方式标记各个特征点,均属于本发明实施例的范畴。

[0127] S1104,在所检测到的特征中查找与具有属性的虚拟对象匹配的特征。

[0128] 不同的虚拟对象所对应现实场景图像中的融合位置也不同。如图3所示,AR潜水眼镜的虚拟道具对应现实场景图像中的融合位置,应该是目标对象的眼部位置。如图4所示,AR兔子装扮的虚拟道具的耳朵对应现实场景图像中的融合位置,应该是目标对象的头部位置;而AR兔子装扮的虚拟道具的兔牙对应现实场景图像中的融合位置,应该是目标对象的牙齿部位(或嘴唇部位)。如图5所示,AR小猫装扮对应现实场景图像中的融合位置,应该是目标对象的两边脸部位置。

[0129] 在一个实施例中,终端确定虚拟对象的功能或用途,根据功能或用途确定虚拟对象所要挂载在目标对象的部位,进而确定虚拟对象与目标对象的特征之间的关系。终端在所检测到的特征中,根据所确定的关系获得与具有属性的虚拟对象匹配的特征。

[0130] S1106,根据匹配的特征,确定具有属性的虚拟对象在现实场景图像中的融合位置。

[0131] 上述实施例中,通过目标对象的特征点,确定虚拟对象在现实场景图像中的融合位置,以便于根据该融合位置将虚拟对象融合在现实场景图像中,获得视觉状态发生变化的虚拟对象,提升了虚拟对象的多样性变化。

[0132] 在一个实施例中,该方法还可以包括:提取音频数据的音频特征;当音频特征符合第一触发条件时,则执行以下至少一种:新增虚拟对象;切换虚拟对象;切换视觉状态的类型。

[0133] 其中,音频特征可以是至少以下之一:音频数据的音量值、频率值、音色、对数功率谱和梅尔频率倒谱系数等。对数功率谱和梅尔频率倒谱系数可以反映出音频数据的功率值以及说话人的风格特征和语音表现力等特征。语音表现力可以是语音的声调、轻重和节奏等特征。对应的第一触发条件可以是音量值达到预设音量阈值,或频率值达到预设频率阈值,或音色满足音色条件,或功率值达到功率阈值,或说话人的风格特征满足风格特征条件,或说话人的语音表现力满足表现力条件等。

[0134] 视觉状态的类型可以是虚拟对象的显示尺寸、显示颜色和运动轨迹等。

[0135] 在一个实施例中,终端对时域的音频数据进行分帧和加窗处理,获得各帧的音频数据。终端对各帧的音频数据进行傅里叶变换,获得对应的频谱。终端根据各帧的频谱计算出功率谱,然后对该功率谱进行对数运算,获得对数功率谱。终端可以将该对数功率谱确定为语音特征,或者,将对数功率谱经过离散余弦变换所得的结果确定为语音特征。

[0136] 例如,假设所采集的语音的信号表达式为 $x(n)$ ,分帧和加窗后的语音为 $x'(n) = x(n) \times h(n)$ ,对加窗后的语音 $x'(n) = x(n) \times h(n)$ 进行离散傅里叶变换,得到对应的频谱信

号为:

$$[0137] \quad X(k) = \sum_{n=1}^{N-1} x(n) \times h(n) e^{-j2\pi kn/N}, 1 \leq k \leq N$$

[0138] 其中,N表示离散傅里叶变换的点数。

[0139] 获得各帧语音的频谱时,终端计算出对应的功率谱,并求出功率谱的对数值得到对数功率谱,从而得到对应的语音特征。

[0140] 或者,获得对数功率谱后,终端将对数功率谱输入梅尔尺度的三角滤波器,经离散余弦变换后得到梅尔频率倒谱系数,所得的梅尔频率倒谱系数为:

$$[0141] \quad C(n) = \sum_{m=1}^{N-1} x(n) \cos\left(\frac{\pi n(m-0.5)}{M}\right), n=1,2,\dots,L$$

[0142] 其中,L阶指的是梅尔频率倒谱系数阶数,可以取值取12-16.M指的是三角滤波器个数。

[0143] 上述实施例中,通过提取音频数据的音频特征,在音频特征满足对应的触发条件时,可在原来虚拟对象的基础上新增虚拟对象,或将原来的虚拟对象切换为其它虚拟对象,或将原来所呈现的视觉状态进行切换,使融合在现实场景图像中的虚拟对象多样化,以及所呈现的视觉状态多样化,进而提高了用户与虚拟对象的交互性。

[0144] 在一个实施例中,该方法还可以包括:根据音频数据进行识别,获得识别结果;确定与识别结果匹配的动态效果类型;按照动态效果类型和属性值确定虚拟对象所呈现的视觉状态;视觉状态与动态效果类型相匹配。

[0145] 其中,识别结果可以指音频类型和音频数据中关键字的文本特征。音频类型可以指音乐的类型,如轻音乐、摇滚音乐和爵士音乐等音乐类型。文本特征可以指关键字的轻音和重音等。重音可通过增加音强或音高来表示。

[0146] 动态效果可以是虚拟对象在动态变化过程中所显示出来的效果。具体地,动态效果可以是以下中的任一种或多种的组合:旋转、移动、透明与非透明之间变化和颜色变化等。例如,虚拟对象随音频数据的变化而发生旋转,或边移动边旋转等。对应的,动态效果类型可以包括旋转类型、移动类型、透明与非透明之间变化类型和颜色变化类型。

[0147] 在一个实施例中,终端识别出音频数据所对应的音乐类型时,获取与音乐类型对应的动态效果,按照获取的动态效果和属性值确定虚拟对象所呈现的视觉状态。例如,当所获取的音频数据为摇滚音乐类型时,动态效果可以是比较动感的效果。

[0148] 在一个实施例中,终端识别出音频数据中关键字的文本特征,根据识别的文本特征选择对应的动态效果。例如,当识别出音频中的关键字为重音时,将虚拟对象的动态效果切换为重音所对应的动态效果。

[0149] 上述实施例中,通过音频数据的识别结果确定对应的动态效果,使虚拟对象随音频数据的变化而呈现不同的动态效果,提升了用户与虚拟对象之间的交互性。

[0150] 对于传统的图像处理方案中,大部分相机/短视频类的应用程序都有动效虚拟对象的能力,即虚拟的虚拟对象会跟随人脸的运动;也都有播放背景音乐或者接收麦克风的能力,即录制视频的时候可以带有背景音乐或麦克风声音。但是目前还没有一款应用程序可以实时获取声音,进而分析声音的属性来实时调节虚拟对象的属性值。

[0151] 为了解决上述问题,本发明实施例提供了一种图像处理方法,通过该图像处理方

法,可实现根据音乐变化动态的调整虚拟对象(如动效虚拟对象随音乐变化),使虚拟对象跟随音乐的音量值或频率值变化而发生颜色,或尺寸,或旋转角度的变化,如图13所示,该图像处理办法包括以下步骤:

[0152] S1302,获取音频数据。

[0153] 终端获取音频数据的方式可以是:一是通过终端的麦克风采集的方式,另一是从终端播放的相应背景音乐读取的方式。通过终端的麦克风采集的方式主要是采集外界的音频数据,如利用手机常用的话筒功能采集用户发出的语音。从终端播放的相应背景音乐读取的方式,主要是终端解码所播放背景音乐的音频格式文件,从而获得音频数据。需要说明的是,可以将上述两种方式所获得的音频数据中的一种作为输入源,也可以将上述两种方式所获得的音频数据之间的混合作为输入源。终端通过PCM编码调制方式,将所获取的音频数据编码成二进制的编码音频数据。音频数据也可称为音频信号,在本发明实施例中不做区分。

[0154] 其中,PCM是一种常见的编码方式,主要是将模拟的音频数据按照预设的时间间隔进行抽样,使模拟的音频数据离散化,然后对抽样值进行量化,同时将量化后的抽样值进行编码,获得按二进制码表示抽样脉冲的幅值。

[0155] S1304,解析音频数据,获得对应的参数值,参数值如频率值和音量值。

[0156] 终端获得经过PCM编码后的编码音频数据,从该编码音频数据中解析出与声音相关的属性值,该属性值可以包括:音量值、频率值和音色等。

[0157] 音量值可以通过音频数据的响度或对应波形的幅值表示,表征一段时间内音频数据的音量大小,计算公式如下:

$$[0158] \quad \bar{v} = 20 \log_{10} \left( \frac{1}{N} \sum_{i=1}^N \text{abs}(v_i) \right)$$

[0159] 其中, $v_i$ 表示经过PCM编码之后所得的编码音频数据中一个抽样点的振幅,N表示抽样点的个数,本实施例中N可以取值1024,也可以为其它数值。对于抽样率为48k的音频数据,1秒钟可计算48次音量值。

[0160] 频率值可以是音频数据在单位时间内对应波形上下震动的次数,单位为Hz。音频数据可以被分解为不同频率值、不同幅值的正弦波的叠加,利用FFT算法可以将PCM编码后所得的编码音频数据转化为频域音频数据,通过该频域音频数据可获得频率值和音量值(即幅值)。

[0161] 将PCM编码后所得的编码音频数据转化为频域音频数据时,可使用512点FFT,最多可将0到截止频率(对于48kHz抽样率,截止频率为24kHz)的频段等分为256个频段,并计算出每个频段的幅值,从而得到音频数据的音量值。此外,获取振幅最大的频段,根据振幅最大频段中的子频域音频数据确定频数据对应的频率值。

[0162] 无论是音量值的计算还是FFT的计算,在终端都可以实现实时计算。

[0163] S1306,选取对应的映射式,将获得的参数值输入映射式。

[0164] 终端获得每个时段内的频域音频数据。以48kHz抽样率的音频数据作为输入源为例,1秒钟内会计算48次。终端根据不同的需求,设计不同的映射式。其中,该映射式即为本发明实施例中所述的映射关系。该映射式的输入变量为音量值或频率值,输出为虚拟对象的属性值,属性值如颜色、缩放比例、旋转角度等。以音频数据的音量值为输入,以虚拟对象

的缩放比例为输出为例,可以设计下述分段的映射式:

$$[0165] \quad f(x) = \begin{cases} 1.0, & x < 50 \\ 1.0 + 3.0 \times \frac{x-50}{70}, & x \geq 50 \end{cases}$$

[0166] 其中,x为音频数据的音量值,范围可以是0~120db,f(x)为虚拟对象的缩放比例。

[0167] 根据实际需求,可以配置各种不同的映射式。其中,映射式在3个维度内可配:1)根据映射式的输入类型配置映射式,如输入类型为音量值或频率值;2)根据映射式的输出类型配置映射式,输出类型可以是虚拟对象的缩放比例、颜色、旋转角度、偏移等各个不同维度的属性值;3)函数的输入与输出之间的对应关系配置映射式。

[0168] S1308,输出虚拟对象的属性值。

[0169] 根据上述映射式,当分贝值小于50db时,虚拟对象保持默认大小1.0。当分贝值大于50db时,虚拟对象的缩放比例随分贝值的增大而增大,当分贝值为120db时,缩放比例为4.0。

[0170] 如图5所示,图5(a)为默认大小,即缩放比例为1.0;图5(b)为缩放大小约为2.0时的效果。

[0171] S1310,采集现实场景图像。

[0172] 终端可通过摄像头实时采集现实场景图像。其中,现实场景图像可以是摄像头实时所采集视频中的一帧图像。

[0173] S1312,检测现实场景图像中对象的特征。

[0174] 终端对现实场景图像中的目标对象做特征检测,如人脸特征检测。

[0175] 其中,检测的方式可以是:采用开源的opencv或dlib的人脸配准点SDK,或使用优图、商汤等提供的人脸特征点检测SDK进行特征检测。

[0176] S1314,将改变属性值的虚拟对象与现实场景图像融合。

[0177] 终端调整属性值后的虚拟对象融合到现实场景图像中目标对象的固定区域(以目标对象的某个脸部特征点位锚点),即可实现虚拟对象实时跟随人脸,并随音频数据的属性值变化而发生变化。

[0178] 以音频数据的音量值控制虚拟对象的缩放比例为例,效果是虚拟对象可以实时跟随人脸,并且虚拟对象的尺寸会随着麦克风采集到的音量值,或者背景音乐的音量值实时发生变化。

[0179] S1316,输出包括有虚拟对象的现实场景图像。

[0180] 通过上述实施例,可以很大程度上增加自拍/短视频类APP的可玩性,虚拟对象会随着音频数据的参数值变化而发生大小、或颜色、或旋转角度等变化,增加了虚拟对象的多样化变化,提高了用户与虚拟对象之间的交互性。

[0181] 图2为一个实施例中图像处理方法的流程示意图。应该理解的是,虽然图2的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,这些步骤可以以其它的顺序执行。而且,图2中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,这些子步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤的子



步骤或者阶段的至少一部分轮流或者交替地执行。

[0182] 如图14所示,在一个实施例中,提供了一种图像处理装置,该图像处理装置具体包括:音频数据获取模块1402、属性值确定模块1404、目标对象确定模块1406、融合位置确定模块1408和融合模块1410;其中:

[0183] 音频数据获取模块1402,用于获取与实时采集的现实场景图像对应的音频数据;

[0184] 属性值确定模块1404,用于根据音频数据动态确定虚拟对象的属性值,该属性值用于确定虚拟对象的视觉状态;

[0185] 目标对象确定模块1406,用于从现实场景图像中确定目标对象;

[0186] 融合位置确定模块1408,用于根据目标对象,确定按属性值确定的虚拟对象在现实场景图像中的融合位置;

[0187] 融合模块1410,用于将按属性值确定的虚拟对象按融合位置融合到现实场景图像;虚拟对象在属性值不同时呈现不同的视觉状态。

[0188] 在一个实施例中,音频数据获取模块1402还用于在实时采集现实场景图像时,从当前环境中实时采集与现实场景图像对应的音频数据;或者,从实时采集的现实场景图像的背景音频中,读取与现实场景图像所对应时间戳对应的音频数据。

[0189] 在一个实施例中,目标对象确定模块1406还用于从现实场景图像中识别生物特征;当生物特征满足预设条件时,则将现实场景图像中与生物特征对应的生物对象确定为目标对象。

[0190] 上述实施例中,获取与实时采集的现实场景图像对应的音频数据,通过音频数据动态确定虚拟对象的属性,从而实现了根据音频数据对虚拟对象属性的控制。通过现实场景图像中的目标对象,确定虚拟对象在现实场景图像中的融合位置,按照融合位置将根据属性值确定的虚拟对象融合到现实场景图像中,由于虚拟对象的属性值由音频数据控制,当音频数据发生变化时,融合到现实场景图像中的虚拟对象的属性值也随之产生变化,提高了交互性。

[0191] 在一个实施例中,属性值确定模块1404还用于获取音频数据的参数值;确定参数值与虚拟对象的属性值之间的预设映射关系;根据预设映射关系,将参数值映射为虚拟对象的属性值。

[0192] 在一个实施例中,属性值确定模块1404对音频数据进行抽样;将抽样所得的结果量化并编码,获得编码音频数据;根据所获得的编码音频数据确定音频数据的参数值。

[0193] 上述实施例中,确定参数值与虚拟对象的属性值之间的预设映射关系,当获得对应的参数值的大小时,便可以通过该预设映射关系获得对应的虚拟对象的属性值,从而实现虚拟对象的调整,使虚拟对象呈现不同的视觉状态,提升了虚拟对象的多样化变化。

[0194] 在一个实施例中,参数值包括频率值;属性值确定模块1404还用于将时域的编码音频数据转换为频域音频数据;将频域音频数据分段,获得多个子频域音频数据;确定各子频域音频数据的振幅;从各子频域音频数据中选取振幅最大的子频域音频数据;按照所选取的子频域音频数据确定音频数据对应的频率值。

[0195] 上述实施例中,对频域音频数据进行分段,可以通过分段后所得子频域音频数据确定频率值,一方面,通过该频率值可以实现对虚拟对象的调整,另一方面,分段后的频域音频数据,在计算过程中可降低计算量,加快计算速率。

[0196] 在一个实施例中,参数值包括音量值;属性值确定模块1404还用于根据所获得的编码音频数据确定音量值;或者,将时域的编码音频数据转换为频域音频数据;根据频域音频数据确定音量值。

[0197] 上述实施例中,根据所获得的编码音频数据,或者根据由时域的编码音频数据转换的频域音频数据这两种方式确定音量值,得到用于确定虚拟对象视觉状态的属性值,通过该音量值可以实现对虚拟对象的调整。

[0198] 在一个实施例中,属性值包括属性调整量;按属性值确定的虚拟对象,是按属性调整量调整相应属性后的虚拟对象;融合模块1410还用于确定虚拟对象所具有的、且与属性调整量对应的属性;按照属性调整量调整虚拟对象的属性,得到调整属性后的虚拟对象;将调整属性后的虚拟对象按融合位置融合到现实场景图像。

[0199] 上述实施例中,通过属性调整量调整虚拟对象的属性,将调整属性后的虚拟对象按融合位置融合到现实场景图像,获得随音频数据的参数变化的虚拟对象,实现根据音频数据调整虚拟对象,提升了虚拟对象的多样化变化,增强用户的体验。

[0200] 在一个实施例中,属性值包括属性变化目标值;按属性值确定的虚拟对象,是相应属性变化至属性变化目标值后的虚拟对象;融合模块1410还用于确定虚拟对象所具有的、且与属性变化目标值对应的属性;将虚拟对象的属性变化至属性变化目标值,得到属性变化后的虚拟对象;将属性变化后的虚拟对象按融合位置融合到现实场景图像。

[0201] 上述实施例中,通过属性变化目标值调整虚拟对象的属性,将调整属性后的虚拟对象按融合位置融合到现实场景图像,获得随音频数据的参数变化的虚拟对象,实现根据音频数据调整虚拟对象,提升了虚拟对象的多样化变化,增强用户的体验。

[0202] 在一个实施例中,融合位置确定模块1408还用于检测目标对象的特征;在所检测到的特征中查找与具有属性的虚拟对象匹配的特征;根据匹配的特征,确定具有属性的虚拟对象在现实场景图像中的融合位置。

[0203] 上述实施例中,通过目标对象的特征点,确定虚拟对象在现实场景图像中的融合位置,以便于根据该融合位置将虚拟对象融合在现实场景图像中,获得视觉状态发生变化的虚拟对象,提升了虚拟对象的多样性变化。

[0204] 图15示出了一个实施例中计算机设备的内部结构图。该计算机设备具体可以是图1中的终端。如图15所示,该计算机设备包括该计算机设备包括通过系统总线连接的处理器、存储器、网络接口、输入装置和显示屏。其中,存储器包括非易失性存储介质和内存存储。该计算机设备的非易失性存储介质存储有操作系统,还可存储有计算机程序,该计算机程序被处理器执行时,可使得处理器实现图像处理方法。该内存存储中也可储存有计算机程序,该计算机程序被处理器执行时,可使得处理器执行图像处理方法。计算机设备的显示屏可以是液晶显示屏或者电子墨水显示屏,计算机设备的输入装置可以是显示屏上覆盖的触摸层,也可以是计算机设备外壳上设置的按键、轨迹球或触控板,还可以是外接的键盘、触控板或鼠标等。

[0205] 本领域技术人员可以理解,图15中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0206] 在一个实施例中,本申请提供的图像处理装置可以实现为一种计算机程序的形

式, 计算机程序可在如图15所示的计算机设备上运行。计算机设备的存储器中可存储组成该14装置的各个程序模块, 比如, 图14所示的音频数据获取模块1402、属性值确定模块1404、目标对象确定模块1406、融合位置确定模块1408和融合模块1410。各个程序模块构成的计算机程序使得处理器执行本说明书中描述的本申请各个实施例的图像处理方法中的步骤。

[0207] 例如, 图15所示的计算机设备可以通过如图14所示的图像处理装置中的音频数据获取模块1402执行S202。计算机设备可通过属性值确定模块1404执行S204。计算机设备可通过目标对象确定模块1406执行S206。计算机设备可通过融合位置确定模块1408执行S208。计算机设备可通过融合模块1410执行S210。

[0208] 在一个实施例中, 提供了一种计算机设备, 包括存储器和处理器, 存储器存储有计算机程序, 计算机程序被处理器执行时, 使得处理器执行以下步骤: 取与实时采集的现实场景图像对应的音频数据; 根据音频数据动态确定虚拟对象的属性值, 该属性值用于确定虚拟对象的视觉状态; 从现实场景图像中确定目标对象; 根据目标对象, 确定按属性值确定的虚拟对象在现实场景图像中的融合位置; 将按属性值确定的虚拟对象按融合位置融合到现实场景图像; 虚拟对象在属性值不同时呈现不同的视觉状态。

[0209] 在一个实施例中, 计算机程序被处理器执行获取与实时采集的现实场景图像对应的音频数据的步骤时, 使得处理器具体执行以下步骤: 在实时采集现实场景图像时, 从当前环境中实时采集与现实场景图像对应的音频数据; 或者, 从实时采集的现实场景图像的背景音频中, 读取与现实场景图像所对应时间戳对应的音频数据。

[0210] 在一个实施例中, 计算机程序被处理器执行根据音频数据动态确定虚拟对象的属性值的步骤时, 使得处理器具体执行以下步骤: 获取音频数据的参数值; 确定参数值与虚拟对象的属性值之间的预设映射关系; 根据预设映射关系, 将参数值映射为虚拟对象的属性值。

[0211] 在一个实施例中, 计算机程序被处理器执行获取音频数据的参数值的步骤时, 使得处理器具体执行以下步骤: 对音频数据进行抽样; 将抽样所得的结果量化并编码, 获得编码音频数据; 根据所获得的编码音频数据确定音频数据的参数值。

[0212] 在一个实施例中, 参数值包括频率值, 计算机程序被处理器执行根据所获得的编码音频数据确定音频数据的参数值的步骤时, 使得处理器具体执行以下步骤: 将时域的编码音频数据转换为频域音频数据; 将频域音频数据分段, 获得多个子频域音频数据; 确定各子频域音频数据的振幅; 从各子频域音频数据中选取振幅最大的子频域音频数据; 按照所选取的子频域音频数据确定音频数据对应的频率值。

[0213] 在一个实施例中, 参数值包括音量值; 计算机程序被处理器执行根据所获得的编码音频数据确定音频数据的参数值的步骤时, 使得处理器具体执行以下步骤: 根据所获得的编码音频数据确定音量值; 或者, 将时域的编码音频数据转换为频域音频数据; 根据频域音频数据确定音量值。

[0214] 在一个实施例中, 属性值包括属性调整量; 按属性值确定的虚拟对象, 是按属性调整量调整相应属性后的虚拟对象; 计算机程序被处理器执行将按属性值确定的虚拟对象按融合位置融合到现实场景图像的步骤时, 使得处理器具体执行以下步骤: 确定虚拟对象所具有的、且与属性调整量对应的属性; 按照属性调整量调整虚拟对象的属性, 得到调整属性

后的虚拟对象;将调整属性后的虚拟对象按融合位置融合到现实场景图像。

[0215] 在一个实施例中,属性值包括属性变化目标值;按属性值确定的虚拟对象,是相应属性变化至属性变化目标值后的虚拟对象;计算机程序被处理器执行将按属性值确定的虚拟对象按融合位置融合到现实场景图像的步骤时,使得处理器具体执行以下步骤:确定虚拟对象所具有的、且与属性变化目标值对应的属性;将虚拟对象的属性变化至属性变化目标值,得到属性变化后的虚拟对象;将属性变化后的虚拟对象按融合位置融合到现实场景图像。

[0216] 在一个实施例中,计算机程序被处理器执行从现实场景图像中确定目标对象的步骤时,使得处理器具体执行以下步骤:从现实场景图像中识别生物特征;当生物特征满足预设条件时,则将现实场景图像中与生物特征对应的生物对象确定为目标对象。

[0217] 在一个实施例中,计算机程序被处理器执行根据目标对象,确定具有属性的虚拟对象在现实场景图像中的融合位置的步骤时,使得处理器具体执行以下步骤:检测目标对象的特征;在所检测到的特征中查找与具有属性的虚拟对象匹配的特征;根据匹配的特征,确定具有属性的虚拟对象在现实场景图像中的融合位置。

[0218] 在一个实施例中,提供了一种计算机可读存储介质,存储有计算机程序,计算机程序被处理器执行时,使得处理器执行以下步骤:取与实时采集的现实场景图像对应的音频数据;根据音频数据动态确定虚拟对象的属性值,该属性值用于确定虚拟对象的视觉状态;从现实场景图像中确定目标对象;根据目标对象,确定按属性值确定的虚拟对象在现实场景图像中的融合位置;将按属性值确定的虚拟对象按融合位置融合到现实场景图像;虚拟对象在属性值不同时呈现不同的视觉状态。

[0219] 在一个实施例中,计算机程序被处理器执行获取与实时采集的现实场景图像对应的音频数据的步骤时,使得处理器具体执行以下步骤:在实时采集现实场景图像时,从当前环境中实时采集与现实场景图像对应的音频数据;或者,从实时采集的现实场景图像的背景音频中,读取与现实场景图像所对应时间戳对应的音频数据。

[0220] 在一个实施例中,计算机程序被处理器执行根据音频数据动态确定虚拟对象的属性值的步骤时,使得处理器具体执行以下步骤:获取音频数据的参数值;确定参数值与虚拟对象的属性值之间的预设映射关系;根据预设映射关系,将参数值映射为虚拟对象的属性值。

[0221] 在一个实施例中,计算机程序被处理器执行获取音频数据的参数值的步骤时,使得处理器具体执行以下步骤:对音频数据进行抽样;将抽样所得的结果量化并编码,获得编码音频数据;根据所获得的编码音频数据确定音频数据的参数值。

[0222] 在一个实施例中,参数值包括频率值,计算机程序被处理器执行根据所获得的编码音频数据确定音频数据的参数值的步骤时,使得处理器具体执行以下步骤:将时域的编码音频数据转换为频域音频数据;将频域音频数据分段,获得多个子频域音频数据;确定各子频域音频数据的振幅;从各子频域音频数据中选取振幅最大的子频域音频数据;按照所选取的子频域音频数据确定音频数据对应的频率值。

[0223] 在一个实施例中,参数值包括音量值;计算机程序被处理器执行根据所获得的编码音频数据确定音频数据的参数值的步骤时,使得处理器具体执行以下步骤:根据所获得的编码音频数据确定音量值;或者,将时域的编码音频数据转换为频域音频数据;根据频域

音频数据确定音量值。

[0224] 在一个实施例中,属性值包括属性调整量;按属性值确定的虚拟对象,是按属性调整量调整相应属性后的虚拟对象;计算机程序被处理器执行将按属性值确定的虚拟对象按融合位置融合到现实场景图像的步骤时,使得处理器具体执行以下步骤:确定虚拟对象所具有的、且与属性调整量对应的属性;按照属性调整量调整虚拟对象的属性,得到调整属性后的虚拟对象;将调整属性后的虚拟对象按融合位置融合到现实场景图像。

[0225] 在一个实施例中,属性值包括属性变化目标值;按属性值确定的虚拟对象,是相应属性变化至属性变化目标值后的虚拟对象;计算机程序被处理器执行将按属性值确定的虚拟对象按融合位置融合到现实场景图像的步骤时,使得处理器具体执行以下步骤:确定虚拟对象所具有的、且与属性变化目标值对应的属性;将虚拟对象的属性变化至属性变化目标值,得到属性变化后的虚拟对象;将属性变化后的虚拟对象按融合位置融合到现实场景图像。

[0226] 在一个实施例中,计算机程序被处理器执行从现实场景图像中确定目标对象的步骤时,使得处理器具体执行以下步骤:从现实场景图像中识别生物特征;当生物特征满足预设条件时,则将现实场景图像中与生物特征对应的生物对象确定为目标对象。

[0227] 在一个实施例中,计算机程序被处理器执行根据目标对象,确定具有属性的虚拟对象在现实场景图像中的融合位置的步骤时,使得处理器具体执行以下步骤:检测目标对象的特征;在所检测到的特征中查找与具有属性的虚拟对象匹配的特征;根据匹配的特征,确定具有属性的虚拟对象在现实场景图像中的融合位置。

[0228] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一非易失性计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0229] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0230] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对本申请专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

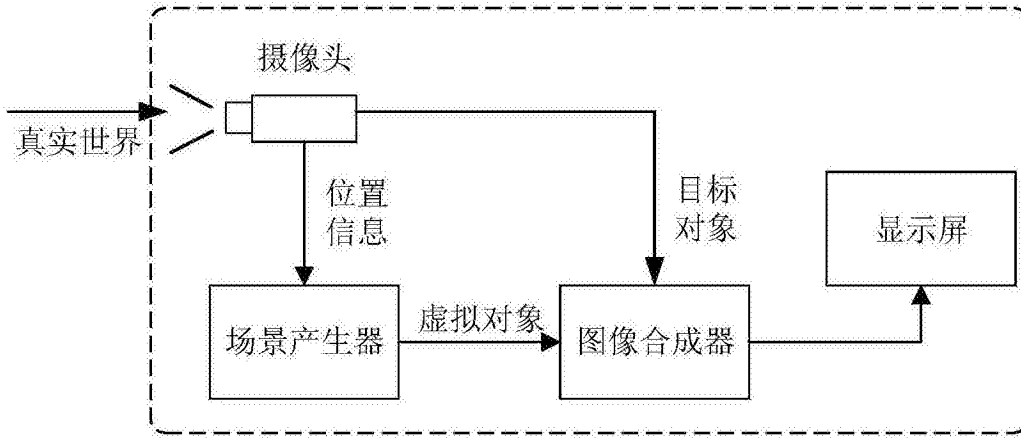


图1

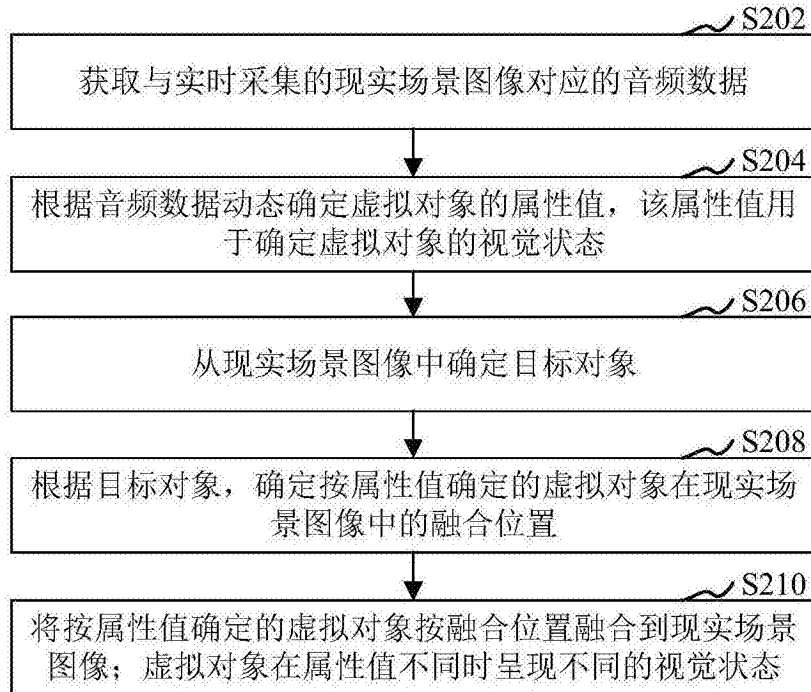


图2

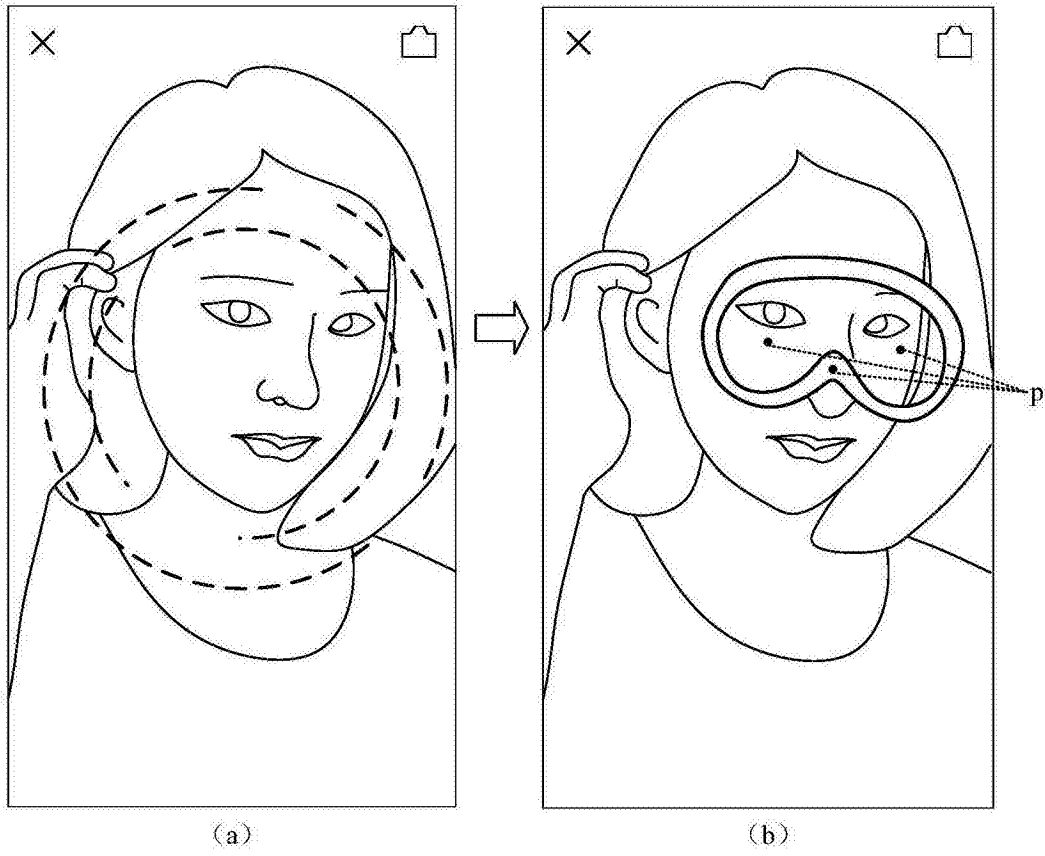


图3

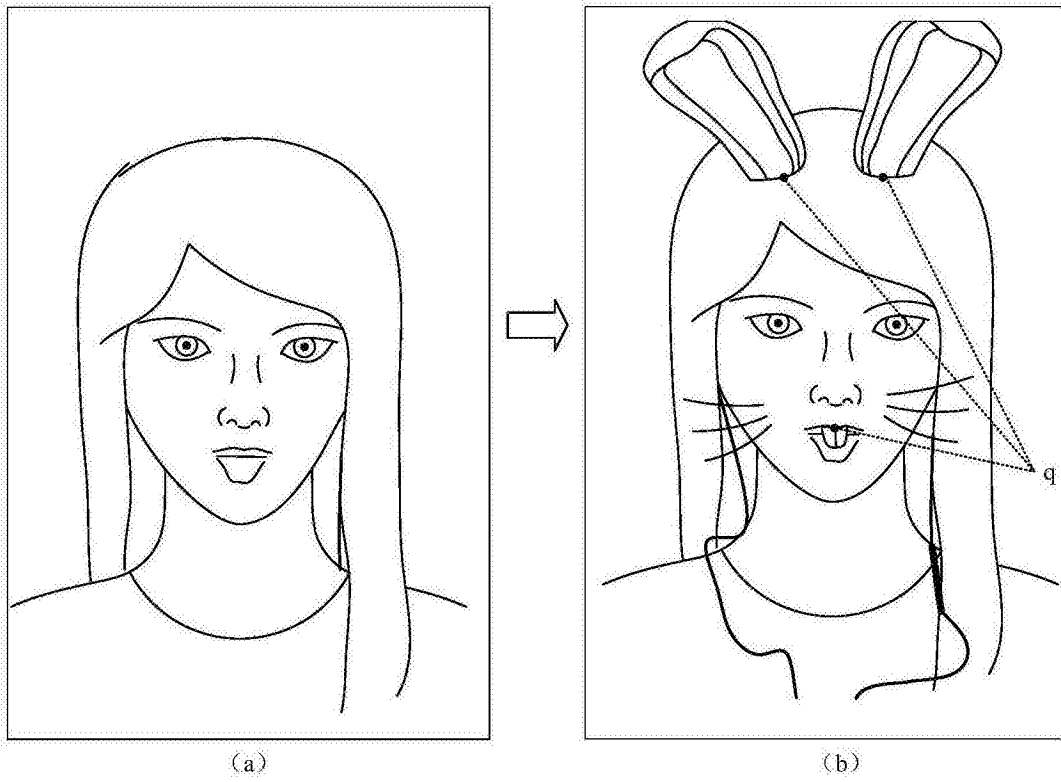


图4

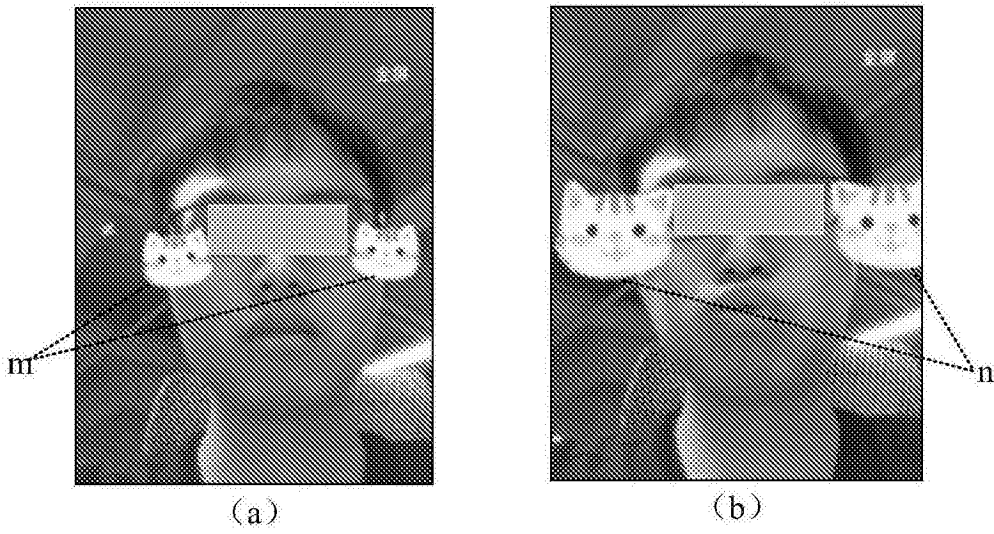


图5



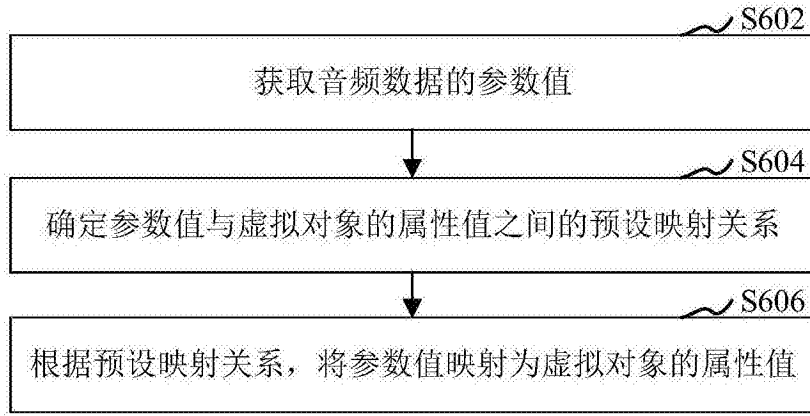


图6

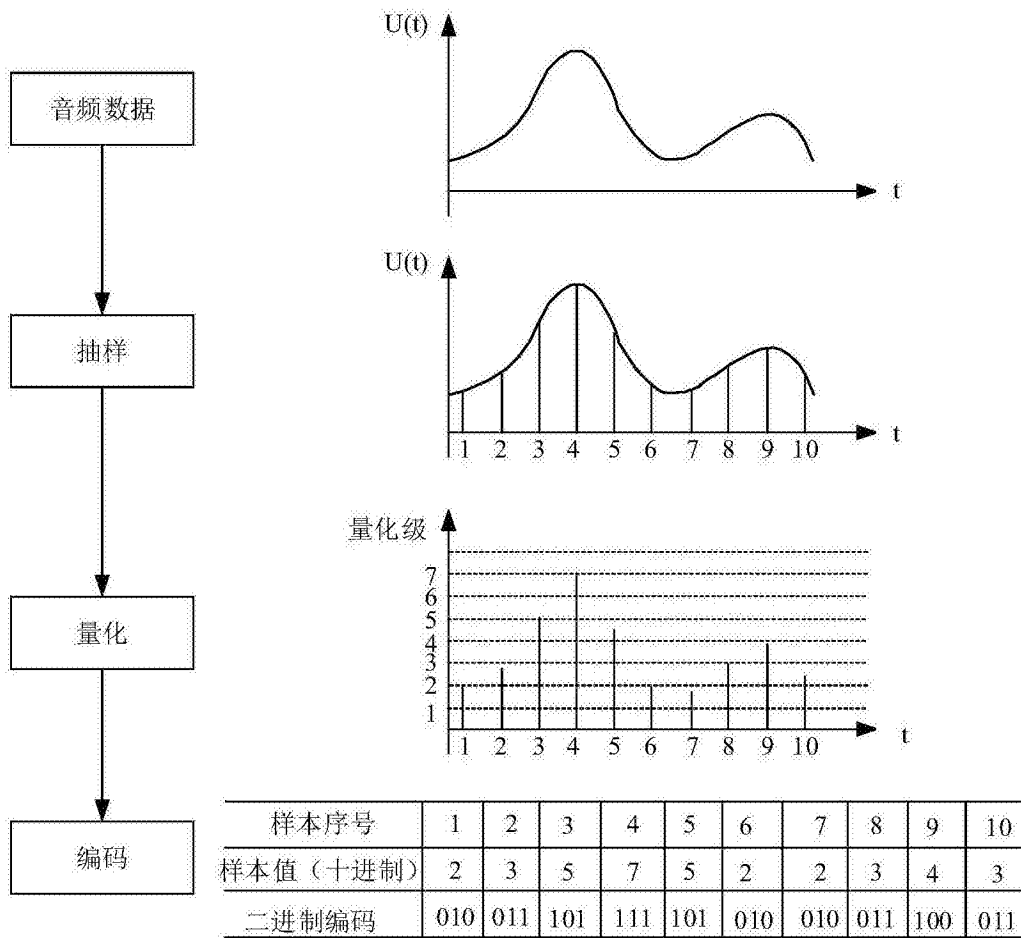


图7

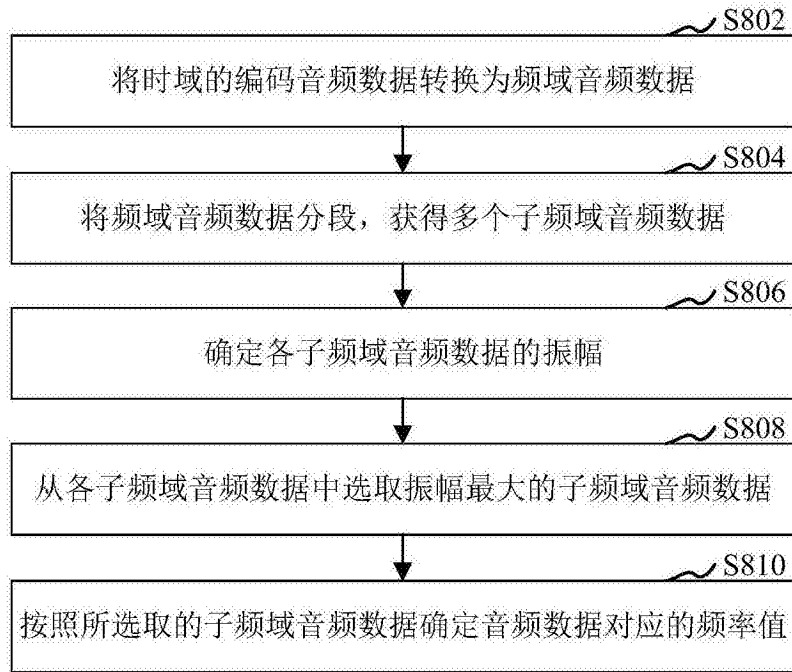


图8

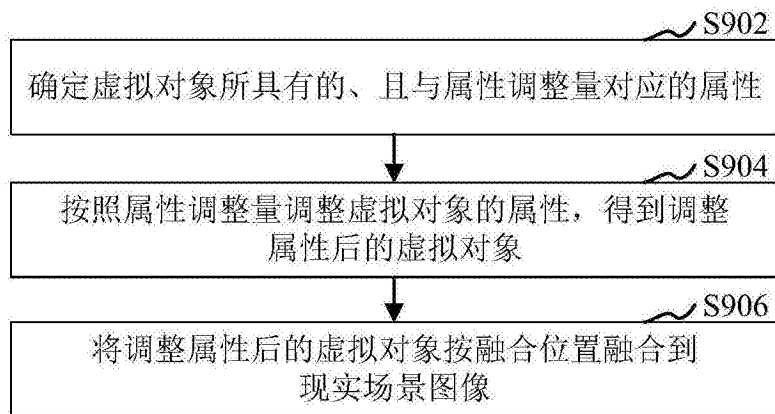


图9

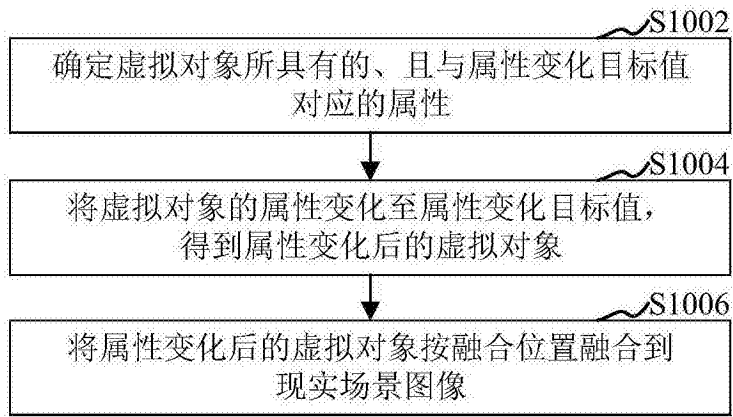


图10

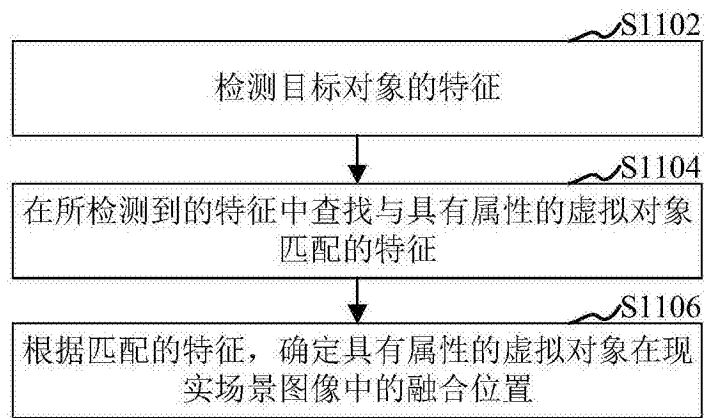


图11

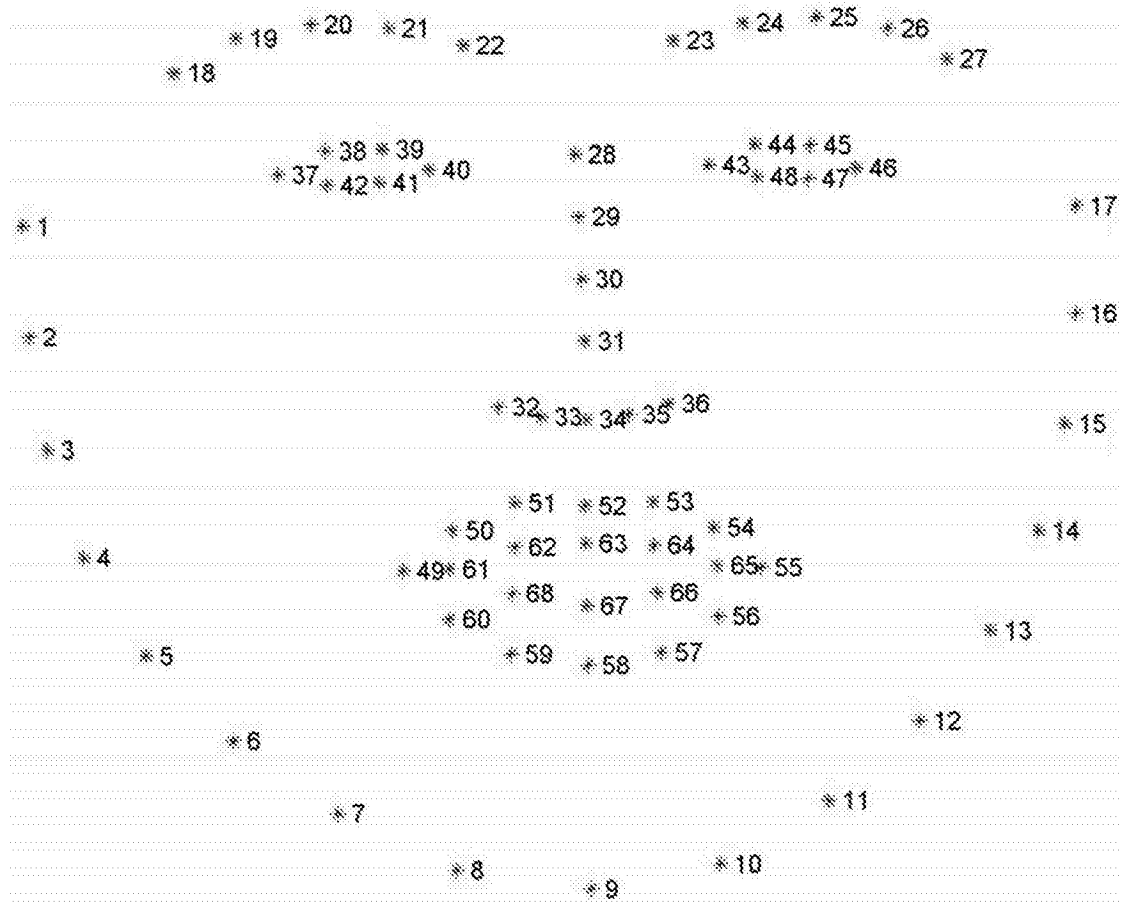


图12

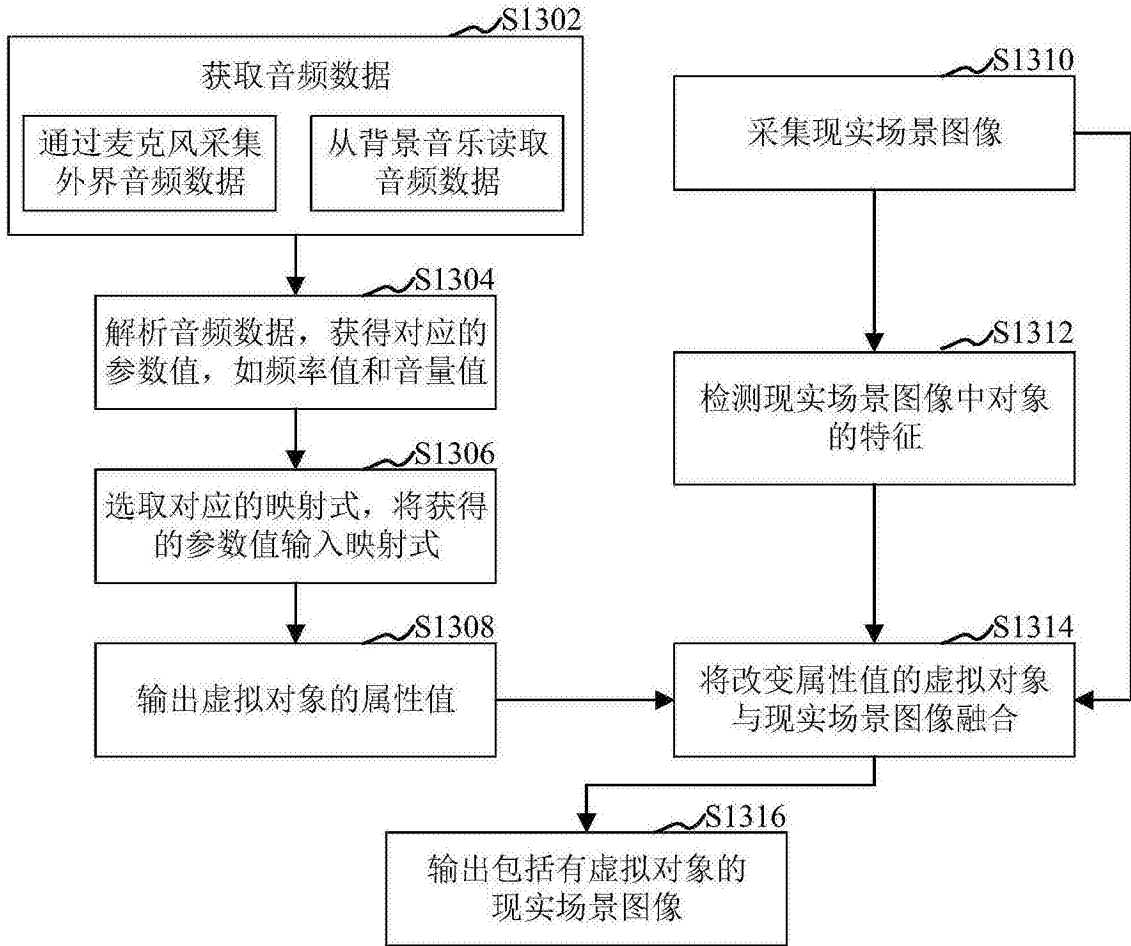


图13

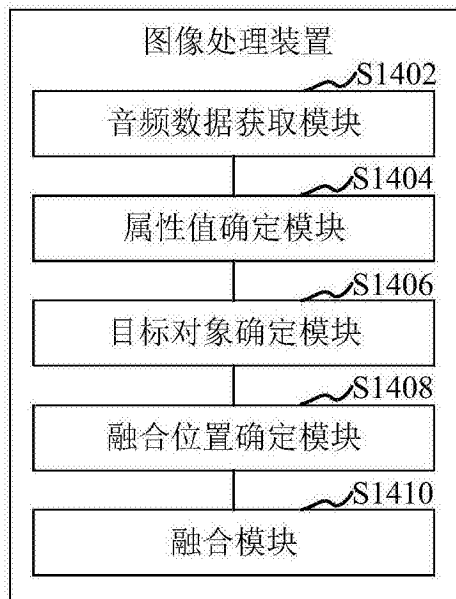


图14

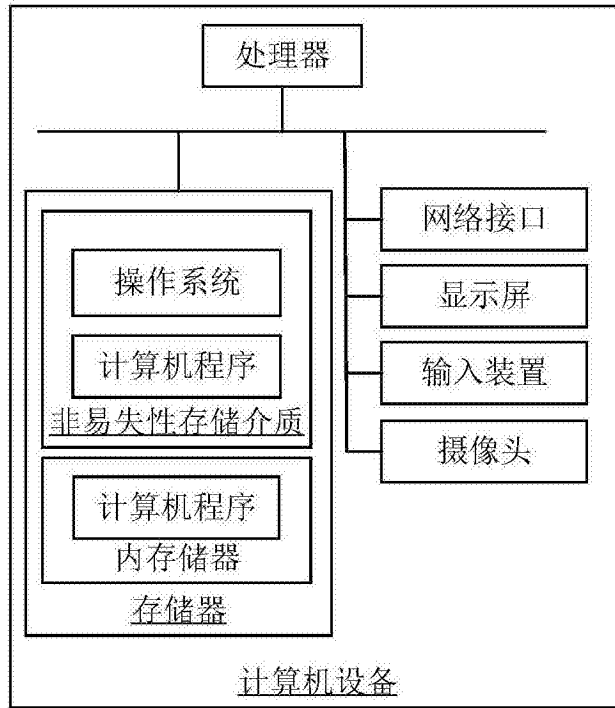


图15