

(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
2019 年 10 月 17 日 (17.10.2019)



(10) 国际公布号
WO 2019/196312 A1

(51) 国际专利分类号:
G06F 3/16 (2006.01) **G06T 7/73** (2017.01)
G06T 7/60 (2017.01) **G06K 9/00** (2006.01)

(21) 国际申请号: PCT/CN2018/102853

(22) 国际申请日: 2018 年 8 月 29 日 (29.08.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810314093.3 2018年4月10日 (10.04.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 周宸(ZHOU, Chen); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 周宝(ZHOU, Bao); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 王健宗(WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市立智方成知识产权代理事务所(普通合伙)(LIFANG & PARTNERS LTD.); 中国广东省深圳市福田区金田路 4018 号安联大厦 22D03, Guangdong 518000 (CN)。

(54) Title: METHOD AND APPARATUS FOR ADJUSTING SOUND VOLUME BY ROBOT, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 机器人调节音量的方法、装置、计算机设备和存储介质

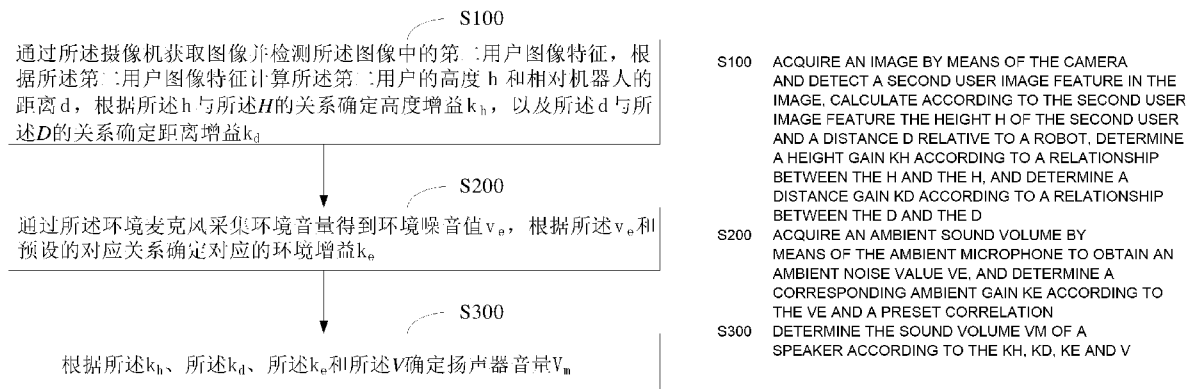


图 2

$$V_m = \frac{k_e(k_d + k_h)}{2} V \quad (1)$$

(57) Abstract: The present application provides a method and an apparatus for automatically adjusting sound volume by a robot. The robot has a camera, a speaker and an ambient microphone for acquiring an ambient sound. It is predefined that the corresponding sound volume of the speaker is V when a first user with a height of H has a distance D from the robot. The method comprises the following steps: acquiring an image by means of a camera and detecting a second user image feature in the image; calculating according to the second user image feature the height h of the second user and a distance d relative to a robot; determining a height gain k_h according to a relationship between the h and the H , and determining a distance gain k_d according to a relationship between the d

(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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(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)。

根据细则4.17的声明:

— 发明人资格(细则4.17(iv))

本国际公布:

— 包括国际检索报告(条约第21条(3))。

and the D; acquiring an ambient sound volume by means of an ambient microphone to obtain an ambient noise value v_e , and determining a corresponding ambient gain k_e according to the v_e and a preset correlation; determining the sound volume of a speaker (I) according to the k_h , k_d , k_e and V, so that the robot can intelligently adjust the sound volume of the speaker according to the actual situation, improving interaction efficiency and user experience. Further provided are a computer device and a storage medium.

(57) 摘要: 本申请提供一种机器人自动调节音量的方法和装置, 机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风, 预定义高度为H的第一用户在离机器人距离为D时对应的扬声器音量为V, 方法包括如下步骤: 通过摄像机获取图像并检测图像中的第二用户图像特征, 根据第二用户图像特征计算第二用户的高度h和相对机器人的距离d, 根据h与H的关系确定高度增益 k_h , 以及d与D的关系确定距离增益 k_d ; 通过环境麦克风采集环境音量得到环境噪音值 v_e , 根据 v_e 和预设的对应关系确定对应的环境增益 k_e ; 根据 k_h 、 k_d 、 k_e 和V确定扬声器音量(I)使得机器人可以根据实际情况智能调节扬声器音量, 提高了交互效率和用户体验。还提供了一种计算机设备和存储介质。

机器人调节音量的方法、装置、计算机设备和存储介质

5 本申请要求于 2018 年 4 月 10 日提交中国专利局、申请号为 201810314093.3，
发明名称为“机器人调节音量的方法、装置、计算机设备和存储介质”的中国专利
申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

10 本申请涉及机器人技术领域，具体而言，本申请涉及一种机器人自动调节音量
的方法和装置，以及一种计算机设备和存储有计算机可读指令的存储介质。

背景技术

 发明人意识到目前的服务机器人一般采用固定的音量进行语音对话和视频播
放等功能，可能会由于各种因素导致环境噪音的分贝值较大，比如人流/其他音响设
备的声音，从而导致用户难以听到机器人的声音，交互效率差，用户体验差。

15 发明内容

 本申请的目的旨在至少能解决上述的技术缺陷之一，特别是交互效率差的技术
缺陷。

 本申请提供一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器
和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为
20 D 时对应的扬声器音量为 V ，所述方法包括如下步骤：通过所述摄像机获取图像并
检测所述图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二
用户的高度 h 和相对所述机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益
 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；通过所述环境麦克风采集环境音
量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；根据
25 所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

 本申请还提供一种机器人自动调节音量的装置，所述机器人具有摄像机、扬声
器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离
为 D 时对应的扬声器音量为 V ，所述装置包括：第一计算模块，用于通过所述摄像
机获取图像并检测所述图像中的第二用户图像特征，根据第二用户图像特征计算所
30 述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增
益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；第二计算模块，用于通过所述

环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定与对应的环境增益 k_e ；音量计算模块，用于根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d+k_h)}{2} V$ 。

本申请还提供一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述机器人自动调节音量的方法包括以下方法的步骤：通过所述摄像机获取图像并检测所述图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d+k_h)}{2} V$ 。

本申请还提供一种存储有计算机可读指令的非易失性存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述机器人自动调节音量的方法包括以下步骤：通过所述摄像机获取图像并检测所述图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d+k_h)}{2} V$ 。

上述的机器人自动调节音量的方法、装置、计算机设备和存储介质，通过判断用户的 h 和用户相对机器人的距离 d 并结合环境麦克风测量的环境噪音值 v_e 来确定扬声器音量 V_m ，使得机器人可以根据实际情况智能调节扬声器音量，从而无论在什么环境都可以给用户最适宜的音量大小，提高了交互效率和用户体验。

附图说明

本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图1为一个实施例中计算机设备的内部结构示意图；

图 2 为一个实施例的机器人自动调节音量的方法流程示意图；

图 3 为一个实施例的机器人和用户之间的空间位置俯视图；

图 4 为一个实施例的机器人自动调节音量的装置模块示意图。

具体实施方式

5 图 1 为一个实施例中计算机设备的内部结构示意图。如图 1 所示，该计算机设备包括通过系统总线连接的处理器、非易失性存储介质、存储器和网络接口。其中，该计算机设备的非易失性存储介质存储有操作系统、数据库和计算机可读指令，数据库中可存储有控件信息序列，该计算机可读指令被处理器执行时，可使得处理器实现一种机器人自动调节音量的方法。该计算机设备的处理器用于提供计算和控制能力，支撑整个计算机设备的运行。该计算机设备的存储器中可存储有计算机可读指令，该计算机可读指令被处理器执行时，可使得处理器执行一种机器人自动调节音量的方法。该计算机设备的网络接口用于与终端连接通信。本领域技术人员可以理解，图 1 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 10 15 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。

以下描述的机器人自动调节音量的方法，可以应用于智能机器人，例如客服机器人、儿童教育机器人等等。

图 2 为一个实施例的机器人自动调节音量的方法流程示意图。

本申请提供一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风（还具有采集用户声音的麦克风），预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述方法包括如下步骤：

步骤 S100：通过所述摄像机获取图像并检测图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与 H 的关系确定高度增益 k_h ，以及所述 d 与 D 的关系确定距离增益 k_d 。

可以采用人脸检测方法进行人脸检测，以检测图像中的第二用户。

由于机器人的摄像机可能捕捉到多个人脸，其中有些人只是背景的人物，并没有和机器人进行交互例如对话，因此可以只需要考虑面向摄像机与机器人对话的人。摄像机通常设置在机器人的面对用户的方向，例如如果机器人具有头部，则可以设置在头部的额头或面部的位置；如果机器人具有躯干，则也可以设置在正面躯干的位置。在此不对摄像机的设置位置进行限定，只需要确保第二用户与机器人对话时

能够拍摄到第二用户即可。

对于摄像机来说，拍摄得到的图像（图片或视频帧）是固定大小的，可以在图片中心位置划定预设的矩形位置作为人脸识别区域，只在这个人脸识别区域中进行人脸检测。例如假设使用的是 1920×1080 尺寸的图片，可以在图片中心位置划定

1000×1000 的矩形位置作为人脸识别区域。

人脸检测（Face Detection）技术，即通过图像分析检测出图像中存在的人脸，并把人脸的位置准确地用矩形框框出来的技术，是人脸特征点检测、人脸识别的基础。常用的人脸检测数据集，包括 FDDB（Face Detection Data Set and Benchmark）等。随着近年来随着深度学习的快速发展，涌现出来很多优秀的人脸检测方法。

例如，FDDB 数据库就提交了很多出色的人脸检测方法，例如采用级联 CNN 网络（卷积神经网络，Convolutional Neural Network）的人脸检测方法：A Convolutional Neural Network Cascade，改进的 faster rcnn 做人脸检测：Face Detection using Deep Learning: An Improved Faster RCNN Approach，还有对小脸检测非常成功的 Finding tiny faces 等等。另外，像 opencv、dlib、libfacedetect 等数据库也提供了人脸检测的接口。

常用的人脸检测方法有以下几种：

- 1.单个 CNN 人脸检测方法
- 2.级联 CNN 人脸检测方法
- 3.OpenCV 人脸检测方法
- 4.Dlib 人脸检测方法
- 5.libfacedetect 人脸检测方法
- 6.Seetaface 人脸检测方法

以下简单介绍下单个 CNN 人脸检测方法。

首先训练一个判断人脸和非人脸的二分类器。例如采用卷积神经网络 caffeNet 进行二分类，可以在 imagenet 数据集训练过的模型，利用自己的人脸数据集，进行微调。也可以自定义卷积网络进行训练，为了能检测到更小的人脸目标，我们一般采用小一点的卷积神经网络作为二分类模型，减小图像输入尺寸，加快预测速度。

然后将训练好的人脸判断分类网络的全连接层改为卷积层，这样网络变成了全卷积网络，可以接受任意输入图像大小，图像经过全卷积网络将得到特征图，特征图上每一个“点”对应该位置映射到原图上的感受野区域属于人脸的概率，将属于人脸概率大于设定阈值的视为人脸候选框。

图像上人脸的大小是变化的，为了适应这种变化，最好的办法就是使用图像金字塔的方式，将待检测的图像缩放到不同大小，以进行多尺度人脸检测。对多个尺度下检测出来的所有人脸候选框，做非极大值抑制 NMS，得到最后人脸检测的结果。

5 如果本实施例中的机器人自动调节音量的方法应用于安卓系统，可以应用 FaceDetector 类判断摄像机拍摄的图像是否存在人脸图像。安卓系统内置有人脸识别 API: FaceDetector，该 API 可以通过少量代码完成人脸识别，但是这种识别是最基本的识别，即只能识别出图像中的人脸。

安卓中的人脸识别技术，需要用到的底层库: android/external/neven/，架构层: frameworks/base/media/java/android/media/FaceDetector.java。Java 层的限制: 1，只能接受 Bitmap 格式的数据; 2，只能识别双眼距离大于 20 像素的人脸像（可在 framework 层中修改）; C，只能检测出人脸的位置（双眼的中心点及距离），不能对人脸进行匹配（查找指定的脸谱）。

Neven 库提供的主要方法:

15 A, android.media.FaceDetector .FaceDetector(int width, int height, int maxFaces);
B, int android.media.FaceDetector .findFaces(Bitmap bitmap, Face [] faces)。

在安卓系统中，可以通过 FaceDetector 类获取人脸图像的两眼位置，并根据两眼位置确定人脸图像在桌面的所处位置。具体步骤可以是: 获取人脸图像的两眼位置中心点，获取人脸图像的两眼间距（瞳距），根据两眼位置中心点和两眼间距绘制矩形区域（矩形框），以矩形区域作为人脸图像在桌面的所处位置。可以通过以下代码获取人脸图像的两眼位置中心点:

mFace[i].getMidPoint(eyeMidPoint);

可以通过以下代码获取人脸图像的两眼间距:

eyesDistance = mFace[i].eyesDistance();

25 可以通过以下代码绘制矩形区域:

myEyesDistance = face.eyesDistance(); //得到两眼位置中心点和两眼间距参数，并对每个人脸画框

```

canvas.drawRect(                //矩形框的位置参数
                (int)(myMidPoint.x - myEyesDistance),
                (int)(myMidPoint.y - myEyesDistance),
                (int)(myMidPoint.x + myEyesDistance),

```

30

(int)(myMidPoint.y + myEyesDistance),

myPaint);

关于人脸检测方法在此不再赘述，在本实施例中，并不对人脸检测方法进行限定。检测出人脸后，计算所检测到的第二用户的高度 h 和相对机器人的距离 d ，根据
5 h 与 H 的关系确定高度增益 k_h ，以及 d 与 D 的关系确定距离增益 k_d 。

第二用户相对于机器人的距离 d 可以通过相关的测距传感器的感测值来计算确定，例如通过红外传感器进行红外测距、使用激光传感器进行激光测距等等。在本实施例中，以图像分析方法来计算确定 d 。

在计算确定 d 时，基于两个假设。第一，对于绝大部分人群来说，人的瞳距的
10 差距相差不大（ $\pm 2\text{cm}$ 左右）。第二，与机器人对话的用户，用户与机器人的距离变化限定在比较小的范围。原理是，通过对比拍摄图片中的人脸瞳距与标定图片中的人脸瞳距的比值来估算距离。用户人脸离摄像机越近，人脸的尺寸越大，这个关系近似成线性关系。

在本实施例中，第二用户图像特征包括人像瞳距，预定义第一用户在离机器人
15 距离为 $D1$ 时图像中的人像瞳距为 $A1$ （实际瞳距），第一用户在离机器人距离为 $D2$ 时图像中的人像瞳距为 $A2$ （实际瞳距），则通过以下公式计算第二用户相对机器人的距离 d ：

$$d = k(a - A1) + D1$$

其中 $k = \frac{D2-D1}{A2-A1}$ ， a 为应用人脸检测方法进行人脸检测时检测到的第二用户人像
瞳距，即图像中的人像瞳距。

当然，还可以采用其他的图像分析方法来计算确定 d ，例如采用其他的计算公
20 式，只需要遵循上述两个假设以及原理即可，在此不赘述。

确定了 d 之后，根据 d 与 D 的关系确定距离增益 k_d 。距离增益 k_d 与 d 是具有正向
关系的，例如正比例关系。在本实施例中，距离增益 $k_d = d/D$ 。当然在其他的实施例中，还可以是其他的计算方式， $k_d = d/D+m$ （ m 为预设系数）等等，在此不赘
25 述。

在本实施例中，预定义第一用户瞳距为 C （实际瞳距）时对应图像中的人像瞳
距为 c （图像中的人像瞳距），则通过以下公式计算第二用户高度 h ：

$$h = H1 + \frac{C}{c}\Delta h$$

其中， $H1$ 为摄像机高度， Δh 为应用人脸检测方法进行人脸检测时检测到的人
脸矩形框中心与图像中心的像素差值。

同样的，还可以采用其他的图像分析方法来计算确定 h ，例如采用其他的计算公式，只需要遵循上述两个假设以及原理即可，在此不赘述。

确定了 h 之后，根据 h 与 H 的关系确定高度增益 k_h 。同样的， h 与 H 是具有正向关系的，例如正比例关系。在本实施例中，高度增益 $k_h = (h - \Delta)/(H - \Delta)$ ，其中
5 Δ 为扬声器高度。当然，在一些实施例中，可以忽略扬声器高度，即 $\Delta=0$ 。

步骤 S200：通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据 v_e 和预设的对应关系确定对应的环境增益 k_e 。

具体的，通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据 v_e 所处的区间范围确定与该区间范围对应的环境增益 k_e 。

10 可以预设多个区间范围，每个区间范围都有对应的预设环境增益。即 (v_1, v_2) 区间范围对应环境增益 k_1 、 (v_2, v_3) 区间范围对应环境增益 k_2 、……、 (v_{n-1}, v_n) 区间范围对应环境增益 k_{n-1} 。

例如，设定噪音标准为 70dB。

对于安静环境下 $(v_e < 40\text{dB})$ ， $k_e = 0.8$ ；

15 对于普通环境下 $(40\text{ dB} < v_e < 70\text{dB})$ ， $k_e = 1$ ；

对于嘈杂环境下 $(70\text{dB} < v_e < 90\text{dB})$ ， $k_e = 1 + (v_e - 70)/100$ ；

对于极其嘈杂情况下 $(v_e > 90\text{dB})$ ， $k_e = \infty$ ；

当然，在一些实施例中，可以通过 v_e 和预设的计算公式确定环境增益 k_e 。

在其中一个实施例中，环境麦克风至少包括第一麦克风和第二麦克风，第一麦克风和第二麦克风位于机器人两侧（以机器人正面方向为基线），例如机器人头部
20 两侧，或者机器人躯干两侧，见图 3；通过所述环境麦克风采集环境音量得到环境噪音值 v_e 的过程包括：

通过第一麦克风采集环境音量得到第一环境噪音值 v_1 ，通过第二麦克风采集环境音量得到第二环境噪音值 v_2 ，将 v_1 和 v_2 之中的最大者确定为环境噪音值 v_e ，即 $v_e =$
25 $\max(v_1, v_2)$ 。

确定 v_e 后，即可通过在数据表中查询 v_e 所述区间范围，然后得到 v_e 所述区间范围对应的环境增益 k_e 。

步骤 S300：根据 k_h 、 k_d 、 k_e 和 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

k_h 、 k_d 、 k_e 都是与扬声器音量 V_m 具有正向关系（例如正比例关系）的， $\frac{(k_d + k_h)}{2}$ 相
30 当于音源增益， $\frac{k_e(k_d + k_h)}{2}$ 相当于总增益，因此任何基于该正向关系而对上述扬声器音量 V_m 的计算公式的适当变形，都可以认为是合理的，在此不赘述。

当然,还可以预设最大音量 V_{max} 和最小音量 V_{min} ,如果 $V_m < V_{min}$,则 $V_m = V_{min}$;如果 $V_m > V_{max}$,则 $V_m = V_{max}$ 。

上述机器人自动调节音量的方法,通过判断第二用户的 h 和第二用户相对机器人的距离 d 并结合环境麦克风测量的环境噪音值 v_e 来确定扬声器音量 V_m ,使得机器人可以根据实际情况智能调节扬声器音量,从而无论在什么环境都可以给用户最适宜的音量大小,提高了交互效率和用户体验。

图4为一个实施例的机器人自动调节音量的装置模块示意图。对应机器人自动调节音量的方法,本申请还提供一种机器人自动调节音量的装置,所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风(还具有采集用户声音的麦克风),预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ,所述装置包括第一计算模块100、第二计算模块200和音量计算模块300。

第一计算模块100用于通过所述摄像机获取图像并检测图像中的第二用户的图像特征,根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ,根据 h 与 H 的关系确定高度增益 k_h ,以及 d 与 D 的关系确定距离增益 k_d ;第二计算模块200用于通过所述环境麦克风采集环境音量得到环境噪音值 v_e ,根据 v_e 和预设的对应关系确定对应的环境增益 k_e ;音量计算模块300用于根据 k_h 、 k_d 、 k_e 和 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

第一计算模块100通过所述摄像机获取图像并检测图像中的第二用户图像特征,根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ,根据 h 与 H 的关系确定高度增益 k_h ,以及 d 与 D 的关系确定距离增益 k_d 。

第一计算模块100可以采用人脸检测方法进行人脸检测,以检测图像中的第二用户。

由于机器人的摄像机可能捕捉到多个人脸,其中有些人只是背景的人物,并没有和机器人进行交互例如对话,因此可以只需要考虑面向摄像机与机器人对话的人。摄像机通常设置在机器人的面对用户的方向,例如如果机器人具有头部,则可以设置在头部的额头或面部的位置;如果机器人具有躯干,则也可以设置在正面躯干的位置。在此不对摄像机的设置位置进行限定,只需要确保第二用户与机器人对话时能够拍摄到第二用户即可。

对于摄像机来说,拍摄得到的图像(图片或视频帧)是固定大小的,可以在图片中心位置划定预设的矩形位置作为人脸识别区域,只在这个人脸识别区域中进行人脸检测。例如假设使用的是 1920×1080 尺寸的图片,可以在图片中心位置划定

1000×1000 的矩形位置作为人脸识别区域。

人脸检测（Face Detection）技术，即通过图像分析检测出图像中存在的人脸，并把人脸的位置准确地用矩形框框出来的技术，是人脸特征点检测、人脸识别的基础。常用的人脸检测数据集，包括 FDDB（Face Detection Data Set and Benchmark）

等。随着近年来随着深度学习的快速发展，涌现出来很多优秀的人脸检测方法。

例如，FDDB 数据库就提交了很多出色的人脸检测方法，例如采用级联 CNN 网络（卷积神经网络，Convolutional Neural Network）的人脸检测方法：A Convolutioanal Neural Network Cascade，改进的 faster rcnn 做人脸检测：Face Detection using Deep Learning: An Improved Faster RCNN Approach，还有对小脸检测非常成功的 Finding tiny faces 等等。另外，像 opencv、dlib、libfacedetect 等数据库也提供了人脸检测的接口。

常用的人脸检测方法有以下几种：

1.单个 CNN 人脸检测方法

2.级联 CNN 人脸检测方法

3.OpenCV 人脸检测方法

4.Dlib 人脸检测方法

5.libfacedetect 人脸检测方法

6.Seetaface 人脸检测方法

以下简单介绍下单个 CNN 人脸检测方法。

首先训练一个判断人脸和非人脸的二分类器。例如采用卷积神经网络 caffeNet 进行二分类，可以在 imagenet 数据集训练过的模型，利用自己的人脸数据集，进行微调。也可以自定义卷积网络进行训练，为了能检测到更小的人脸目标，我们一般采用小一点的卷积神经网络作为二分类模型，减小图像输入尺寸，加快预测速度。

然后将训练好的人脸判断分类网络的全连接层改为卷积层，这样网络变成了全卷积网络，可以接受任意输入图像大小，图像经过全卷积网络将得到特征图，特征图上每一个“点”对应该位置映射到原图上的感受野区域属于人脸的概率，将属于人脸概率大于设定阈值的视为人脸候选框。

图像上人脸的大小是变化的，为了适应这种变化，最好的办法就是使用图像金字塔的方式，将待检测的图像缩放到不同大小，以进行多尺度人脸检测。对多个尺度下检测出来的所有人脸候选框，做非极大值抑制 NMS，得到最后人脸检测的结果。

如果本实施例中的机器人自动调节音量的装置应用于安卓系统，第一计算模块 100 可以应用 FaceDetector 类判断摄像机拍摄的图像是否存在人脸图像。安卓系统内置有人脸识别 API: FaceDetector，该 API 可以通过少量代码完成人脸识别，但是这种识别是最基本的识别，即只能识别出图像中的人脸。

5 安卓中的人脸识别技术，需要用到的底层库: android/external/neven/，架构层: frameworks/base/media/java/android/media/FaceDetector.java。Java 层的限制: 1，只能接受 Bitmap 格式的数据; 2，只能识别双眼距离大于 20 像素的人脸像（可在 framework 层中修改）; C，只能检测出人脸的位置（双眼的中心点及距离），不能对人脸进行匹配（查找指定的脸谱）。

10 Neven 库提供的主要方法:

A, android.media.FaceDetector .FaceDetector(int width, int height, int maxFaces);
B, int android.media.FaceDetector .findFaces(Bitmap bitmap, Face [] faces)。

在安卓系统中，第一计算模块 100 可以通过 FaceDetector 类获取人脸图像的两眼位置，并根据两眼位置确定人脸图像在桌面的所处位置。具体步骤可以是: 获取
15 人脸图像的两眼位置中心点，获取人脸图像的两眼间距（瞳距），根据两眼位置中心点和两眼间距绘制矩形区域（矩形框），以矩形区域作为人脸图像在桌面的所处位置。可以通过以下代码获取人脸图像的两眼位置中心点:

```
mFace[i].getMidPoint(eyeMidPoint);
```

可以通过以下代码获取人脸图像的两眼间距:

```
20 eyesDistance = mFace[i].eyesDistance();
```

可以通过以下代码绘制矩形区域:

myEyesDistance = face.eyesDistance(); //得到两眼位置中心点和两眼间距参数，并对每个人脸画框

```
canvas.drawRect( //矩形框的位置参数
```

```
25 (int)(myMidPoint.x - myEyesDistance),
```

```
(int)(myMidPoint.y - myEyesDistance),
```

```
(int)(myMidPoint.x + myEyesDistance),
```

```
(int)(myMidPoint.y + myEyesDistance),
```

```
myPaint);
```

30 关于人脸检测方法在此不再赘述，在本实施例中，并不对人脸检测方法进行限定。第一计算模块 100 检测出人脸后，计算所检测到的第二用户的高度 h 和相对机

器人的距离 d ，根据 h 与 H 的关系确定高度增益 k_h ，以及 d 与 D 的关系确定距离增益 k_d 。

第二用户相对于机器人的距离 d 可以通过相关的测距传感器的感测值来计算确定，例如通过红外传感器进行红外测距、使用激光传感器进行激光测距等等。在本
5 实施例中，第一计算模块 100 以图像分析方法来计算确定 d 。

在计算确定 d 时，基于两个假设。第一，对于绝大部分人群来说，人的瞳距的差距相差不大（ $\pm 2\text{cm}$ 左右）。第二，与机器人对话的用户，用户与机器人的距离变化限定在比较小的范围。原理是，通过对比拍摄图片中的人脸瞳距与标定图片中的人脸瞳距的比值来估算距离。用户人脸离摄像机越近，人脸的尺寸越大，这个关
10 系近似成线性关系。

在本实施例中，第二用户图像特征包括人像瞳距，预定义第一用户在离机器人距离为 $D1$ 时图像中的人像瞳距为 $A1$ （实际瞳距），第一用户在离机器人距离为 $D2$ 时图像中的人像瞳距为 $A2$ （实际瞳距），则第一计算模块 100 通过以下公式计算第二用户相对机器人的距离 d ：

$$d = k(a - A1) + D1$$

15 其中 $k = \frac{D2-D1}{A2-A1}$ ， a 为应用人脸检测方法进行人脸检测时检测到的第二用户人像瞳距，即图像中的人像瞳距。

当然，第一计算模块 100 还可以采用其他的图像分析方法来计算确定 d ，例如采用其他的计算公式，只需要遵循上述两个假设以及原理即可，在此不赘述。

第一计算模块 100 确定了 d 之后，根据 d 与 D 的关系确定距离增益 k_d 。距离增
20 益 k_d 与 d 是具有正向关系的，例如正比例关系。在本实施例中，距离增益 $k_d = d/D$ 。当然在其他的实施例中，还可以是其他的计算方式， $k_d = d/D+m$ （ m 为预设系数）等等，在此不赘述。

在本实施例中，预定义第一用户瞳距为 C （实际瞳距）时对应图像中的人像瞳距为 c （图像中的人像瞳距），则第一计算模块 100 通过以下公式计算第二用户高
25 度 h ：

$$h = H1 + \frac{C}{c} \Delta h$$

其中， $H1$ 为摄像机高度， Δh 为应用人脸检测方法进行人脸检测时检测到的人脸矩形框中心与图像中心的像素差值。

同样的，第一计算模块 100 还可以采用其他的图像分析方法来计算确定 h ，例如采用其他的计算公式，只需要遵循上述两个假设以及原理即可，在此不赘述。

第一计算模块 100 确定了 h 之后, 根据 h 与 H 的关系确定高度增益 k_h 。同样的, h 与 H 是具有正向关系的, 例如正比例关系。在本实施例中, 高度增益 $k_h = (h - \Delta)/(H - \Delta)$, 其中 Δ 为扬声器高度。当然, 在一些实施例中, 可以忽略扬声器高度, 即 $\Delta=0$ 。

5 第二计算模块 200 通过所述环境麦克风采集环境音量得到环境噪音值 v_e , 根据 v_e 和预设的对应关系确定对应的环境增益 k_e 。具体的, 通过所述环境麦克风采集环境音量得到环境噪音值 v_e , 根据 v_e 所处于的区间范围确定与该区间范围对应的环境增益 k_e 。

10 可以预设多个区间范围, 每个区间范围都有对应的预设环境增益。即 (v_1, v_2) 区间范围对应环境增益 k_1 、 (v_2, v_3) 区间范围对应环境增益 k_2 、 $\dots\dots$ 、 (v_{n-1}, v_n) 区间范围对应环境增益 k_{n-1} 。

例如, 设定噪音标准为 70dB。

对于安静环境下 ($v_e < 40\text{dB}$), $k_e = 0.8$;

对于普通环境下 ($40\text{dB} < v_e < 70\text{dB}$), $k_e = 1$;

15 对于嘈杂环境下 ($70\text{dB} < v_e < 90\text{dB}$), $k_e = 1 + (v_e - 70)/100$;

对于极其嘈杂情况下 ($v_e > 90\text{dB}$), $k_e = \infty$;

当然, 在一些实施例中, 第二计算模块 200 可以通过 v_e 和预设的计算公式确定环境增益 k_e 。

20 在其中一个实施例中, 环境麦克风至少包括第一麦克风和第二麦克风, 第一麦克风和第二麦克风位于机器人两侧 (以机器人正面方向为基线), 例如机器人头部两侧, 或者机器人躯干两侧, 见图 3; 第二计算模块 200 通过所述环境麦克风采集环境音量得到环境噪音值 v_e 的过程包括:

通过第一麦克风采集环境音量得到第一环境噪音值 v_1 , 通过第二麦克风采集环境音量得到第二环境噪音值 v_2 , 将 v_1 和 v_2 之中的最大者确定为环境噪音值 v_e , 即 $v_e = \max(v_1, v_2)$ 。

第二计算模块 200 确定 v_e 后, 即可通过在数据表中查询 v_e 所述区间范围, 然后得到 v_e 所述区间范围对应的环境增益 k_e 。

音量计算模块 300 根据 k_h 、 k_d 、 k_e 和 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

30 k_h 、 k_d 、 k_e 都是与扬声器音量 V_m 具有正向关系 (例如正比例关系) 的, $\frac{(k_d + k_h)}{2}$ 相当于音源增益, $\frac{k_e(k_d + k_h)}{2}$ 相当于总增益, 因此任何基于该正向关系而对上述扬声器音量 V_m 的计算公式的适当变形, 都可以认为是合理的, 在此不赘述。

当然,还可以预设最大音量 V_{max} 和最小音量 V_{min} ,如果 $V_m < V_{min}$,则 $V_m = V_{min}$;如果 $V_m > V_{max}$,则 $V_m = V_{max}$ 。

上述机器人自动调节音量的装置,通过判断第二用户的 h 和第二用户相对机器人的距离 d 并结合环境麦克风测量的环境噪音值 v_e 来确定扬声器音量 V_m ,使得机器人可以根据实际情况智能调节扬声器音量,从而无论在什么环境都可以给用户最适宜的音量大小,提高了交互效率和用户体验。

本申请还提供一种计算机设备,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述计算机可读指令被所述处理器执行时,使得所述处理器执行上述任一项实施例所述机器人自动调节音量的方法的步骤。

本申请还提供一种存储有计算机可读指令的存储介质,所述计算机可读指令被一个或多个处理器执行时,使得一个或多个处理器执行上述任一项实施例所述机器人自动调节音量的方法的步骤。

所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风,预定义高度为 H 的用第一用户在离机器人距离为 D 时对应的扬声器音量为 V ,所述方法包括如下步骤:通过所述摄像机获取图像并检测所述图像中的第二用户图像特征,根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ,根据所述 h 与所述 H 的关系确定高度增益 k_h ,以及所述 d 与所述 D 的关系确定距离增益 k_d ;通过所述环境麦克风采集环境音量得到环境噪音值 v_e ,根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ;根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。通过判断用户的 h 和用户相对机器人的距离 d 并结合环境麦克风测量的环境噪音值 v_e 来确定扬声器音量 V_m ,使得机器人可以根据实际情况智能调节扬声器音量,从而无论在什么环境都可以给用户最适宜的音量大小,提高了交互效率和用户体验。

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,该计算机程序可存储于一计算机可读存储介质中,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,前述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory, ROM)等非易失性存储介质,或随机存储记忆体(Random Access Memory, RAM)等。

权 利 要 求 书

1、一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述方法包括如下步骤：

5 通过所述摄像机获取图像并检测所述图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二用户的高度 h 和相对所述机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；

通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；

10 根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

2、根据权利要求 1 所述的机器人自动调节音量的方法，所述第二用户图像特征包括人像瞳距，预定义所述第一用户在离所述机器人距离为 $D1$ 时所述图像中的人像瞳距为 $A1$ ，所述第一用户在离所述机器人距离为 $D2$ 时所述图像中的人像瞳距为 $A2$ ，则通过以下公式计算所述第二用户相对所述机器人的距离 d ： $d = k(a - A1) + D1$ ，

其中 $k = \frac{D2 - D1}{A2 - A1}$ ， a 为所述图像中第二用户的人像瞳距。

3、根据权利要求 1 所述的机器人自动调节音量的方法，所述第二用户图像特征包括人像瞳距，预定义所述第一用户的真实瞳距为 C 时对应所述图像中的人像瞳距为 c ，则通过以下公式计算第二用户高度 h ：

20
$$h = H1 + \frac{c}{c} \Delta h,$$

其中， $H1$ 为摄像机高度， Δh 为检测到的人脸矩形框中心与图像中心的像素差值。

4、根据权利要求 1 所述的机器人自动调节音量的方法，所述环境麦克风至少包括第一麦克风和第二麦克风，第一麦克风和第二麦克风位于机器人两侧；所述通过所述环境麦克风采集环境音量得到环境噪音值 v_e 包括：

通过第一麦克风采集环境音量得到第一环境噪音值 v_1 ，通过第二麦克风采集环境音量得到第二环境噪音值 v_2 ，将所述 v_1 和所述 v_2 之中的最大者确定为环境噪音值 v_e 。

5、根据权利要求 1 所述的机器人自动调节音量的方法，高度增益 $k_h = (h - \Delta)/(H - \Delta)$ ，其中 Δ 为扬声器高度。

6、根据权利要求 1 所述的机器人自动调节音量的方法，距离增益 $k_d = d/D$ 。

7、一种机器人自动调节音量的装置，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述装置包括：

第一计算模块，用于通过所述摄像机获取图像并检测所述图像中的第二用户图像特征，根据第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；

第二计算模块，用于通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定与对应的环境增益 k_e ；

音量计算模块，用于根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

8、根据权利要求 7 所述的机器人自动调节音量的装置，所述第二用户图像特征包括人像瞳距，预定义所述第一用户在离所述机器人距离为 $D1$ 时所述图像中的人像瞳距为 $A1$ ，所述第一用户在离所述机器人距离为 $D2$ 时所述图像中的人像瞳距为 $A2$ ，则第一计算模块通过以下公式计算所述第二用户相对所述机器人的距离 d ：

$$d = k(a - A1) + D1,$$

其中 $k = \frac{D2 - D1}{A2 - A1}$ ， a 为所述图像中第二用户的人像瞳距。

9、一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行一种机器人自动调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述机器人自动调节音量的方法包括以下步骤：

通过所述摄像机获取图像并检测所述图像中的第二用户的图像特征，根据所述第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；

通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；

根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

10、根据权利要求 9 所述的计算机设备，所述第二用户图像特征包括人像瞳距，预定义所述第一用户在离所述机器人距离为 $D1$ 时所述图像中的人像瞳距为 $A1$ ，所述第一用户在离所述机器人距离为 $D2$ 时所述图像中的人像瞳距为 $A2$ ，则通过以下

$$d = k(a - A1) + D1,$$

其中 $k = \frac{D2-D1}{A2-A1}$ ， a 为所述图像中第二用户的人像瞳距。

11、根据权利要求9所述的计算机设备，所述第二用户图像特征包括人像瞳距，
5 预定义所述第一用户的真实瞳距为 C 时对应所述图像中的人像瞳距为 c ，则通过以下公式计算第二用户高度 h ：

$$h = H1 + \frac{c}{C} \Delta h,$$

其中， $H1$ 为摄像机高度， Δh 为检测到的人脸矩形框中心与图像中心的像素差值。

10 12、根据权利要求9所述的计算机设备，所述环境麦克风至少包括第一麦克风和第二麦克风，第一麦克风和第二麦克风位于机器人两侧；所述通过所述环境麦克风采集环境音量得到环境噪音值 v_e 包括：

通过第一麦克风采集环境音量得到第一环境噪音值 v_1 ，通过第二麦克风采集环境音量得到第二环境噪音值 v_2 ，将所述 v_1 和所述 v_2 之中的最大者确定为环境噪音值
15 v_e 。

13、根据权利要求9所述的计算机设备，高度增益 $k_h = (h - \Delta)/(H - \Delta)$ ，其中 Δ 为扬声器高度。

14、根据权利要求9所述的计算机设备，距离增益 $k_d = d/D$ 。

15、一种非易失性存储介质，所述存储介质内存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行一种机器人自动
20 调节音量的方法，所述机器人具有摄像机、扬声器和用于采集环境声音的环境麦克风，预定义高度为 H 的第一用户在离机器人距离为 D 时对应的扬声器音量为 V ，所述机器人自动调节音量的方法包括以下步骤：

通过所述摄像机获取图像并检测所述图像中的第二用户的图像特征，根据所述
25 第二用户图像特征计算所述第二用户的高度 h 和相对机器人的距离 d ，根据所述 h 与所述 H 的关系确定高度增益 k_h ，以及所述 d 与所述 D 的关系确定距离增益 k_d ；

通过所述环境麦克风采集环境音量得到环境噪音值 v_e ，根据所述 v_e 和预设的对应关系确定对应的环境增益 k_e ；

根据所述 k_h 、所述 k_d 、所述 k_e 和所述 V 确定扬声器音量 $V_m = \frac{k_e(k_d + k_h)}{2} V$ 。

30 16、根据权利要求15所述的非易失性存储介质，所述第二用户图像特征包括人像瞳距，预定义所述第一用户在离所述机器人距离为 $D1$ 时所述图像中的人像瞳距

为 $A1$ ，所述第一用户在离所述机器人距离为 $D2$ 时所述图像中的人像瞳距为 $A2$ ，
则通过以下公式计算所述第二用户相对所述机器人的距离 d ：

$$d = k(a - A1) + D1,$$

其中 $k = \frac{D2-D1}{A2-A1}$ ， a 为所述图像中第二用户的人像瞳距。

17、根据权利要求 15 所述的非易失性存储介质，所述第二用户图像特征包括人像瞳距，预定义所述第一用户的真实瞳距为 C 时对应所述图像中的人像瞳距为 c ，
则通过以下公式计算第二用户高度 h ：

$$h = H1 + \frac{c}{C} \Delta h,$$

其中， $H1$ 为摄像机高度， Δh 为检测到的人脸矩形框中心与图像中心的像素差值。

18、根据权利要求 15 所述的非易失性存储介质，所述环境麦克风至少包括第一麦克风和第二麦克风，第一麦克风和第二麦克风位于机器人两侧；所述通过所述环境麦克风采集环境音量得到环境噪音值 v_e 包括：

通过第一麦克风采集环境音量得到第一环境噪音值 v_1 ，通过第二麦克风采集环境音量得到第二环境噪音值 v_2 ，将所述 v_1 和所述 v_2 之中的最大者确定为环境噪音值 v_e 。

19、根据权利要求 15 所述的非易失性存储介质，高度增益 $k_h = (h - \Delta)/(H - \Delta)$ ，
其中 Δ 为扬声器高度。

20、根据权利要求 15 所述的非易失性存储介质，距离增益 $k_d = d/D$ 。

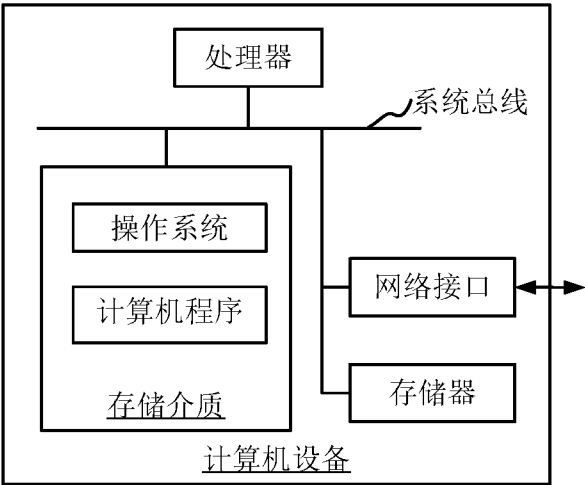


图 1

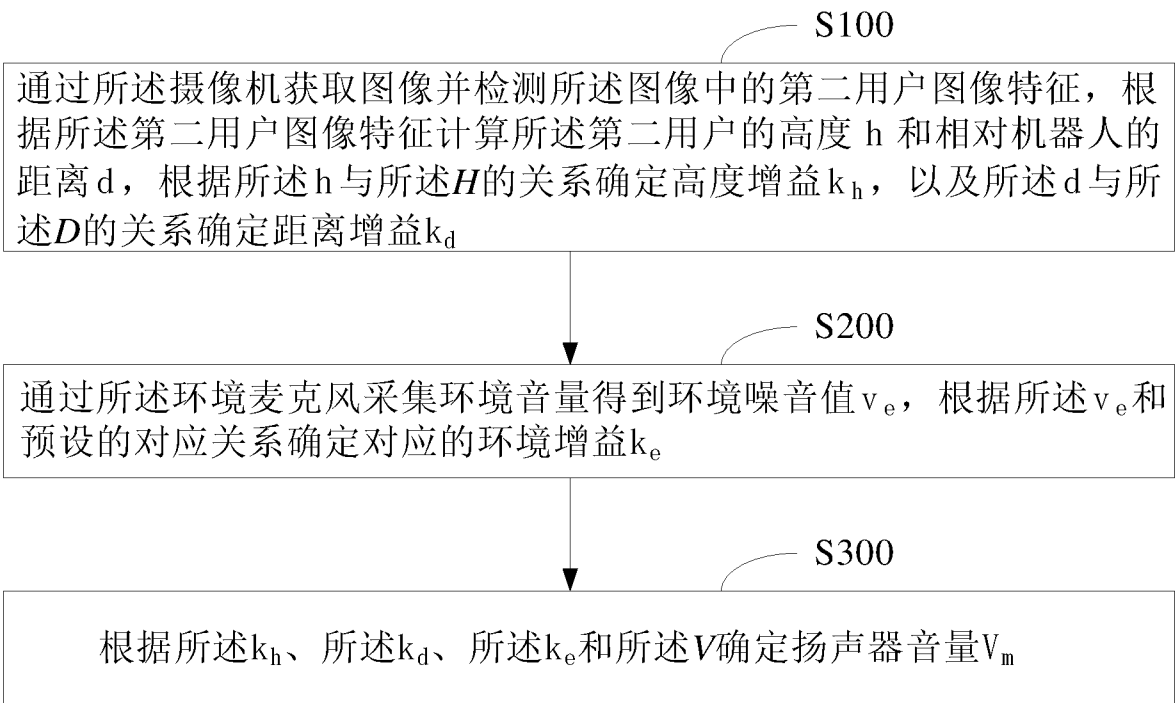


图 2

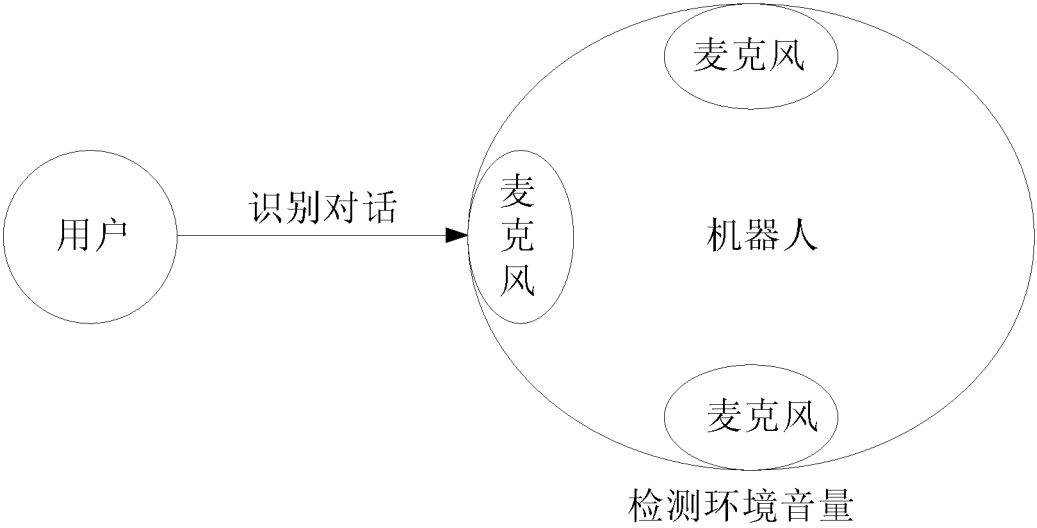


图 3

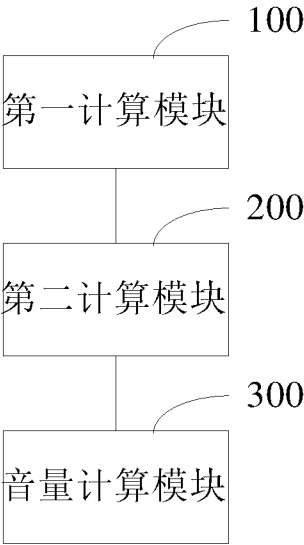


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/102853

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/16(2006.01)i; G06T 7/60(2017.01)i; G06T 7/73(2017.01)i; G06K 9/00(2006.01)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)

G06F; G06T; G06K

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)

CNABS, CNTXT, DWPI, SIPOABS, CNKI: 音量, 声音, 调整, 控制, 扬声器, 高度, 距离, 环境, 噪声, 比例, 瞳距, 图像, 麦克风; volume, sound, voice, adjust, control, height, distance, environment, noise, proportion, pupil, image, microphone

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106648527 A (LETV HOLDING (BEIJING) CO., LTD. ET AL.) 10 May 2017 (2017-05-10) description, paragraphs [0056]-[0103] and [0107]-[0111], and figures 2 and 4-6	1-20
Y	CN 106534541 A (GUANGDONG XIAOTIANCAI TECHNOLOGY CO., LTD.) 22 March 2017 (2017-03-22) description, paragraphs [0037], [0038] and [0049]	1-20
Y	CN 103793719 A (SHENZHEN UNIVERSITY) 14 May 2014 (2014-05-14) description, paragraph [0035]	2, 3, 8, 10, 11, 16, 17
Y	CN 106377264 A (GUANGZHOU CVTE ELECTRONIC TECHNOLOGY CO., LTD.) 08 February 2017 (2017-02-08) description, paragraphs [0058]-[0061]	3, 11, 17
Y	CN 205754809 U (SHENZHEN DOUBLEWELL TECHNOLOGY LLC) 30 November 2016 (2016-11-30) description, paragraphs [0016]-[0025], and figure 2	1-20
Y	CN 101867739 A (SUN YAT-SEN UNIVERSITY) 20 October 2010 (2010-10-20) description, paragraphs [0030]-[0059], and figures 1 and 2	1-20



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

“&” document member of the same patent family

Date of the actual completion of the international search

25 December 2018

Date of mailing of the international search report

04 January 2019

Name and mailing address of the ISA/CN

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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/102853

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102981422 A (GUANGZHOU HUADUO NETWORK TECHNOLOGY CO., LTD.) 20 March 2013 (2013-03-20) description, paragraphs [0069]-[0108]	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/102853

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106648527	A	10 May 2017	None			
CN	106534541	A	22 March 2017	None			
CN	103793719	A	14 May 2014	None			
CN	106377264	A	08 February 2017	WO	2018072598	A1	26 April 2018
CN	205754809	U	30 November 2016	None			
CN	101867739	A	20 October 2010	None			
CN	102981422	A	20 March 2013	CN	102981422	B	23 December 2015

A. 主题的分类 G06F 3/16(2006.01)i; G06T 7/60(2017.01)i; G06T 7/73(2017.01)i; G06K 9/00(2006.01)i 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																										
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06F; G06T; G06K 包含在检索领域中的除最低限度文献以外的检索文献 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)) CNABS, CNTXT, DWPI, SIPOABS, CNKI: 音量, 声音, 调整, 控制, 扬声器, 高度, 距离, 环境, 噪声, 比例, 瞳距, 图像, 麦克风; volume, sound, voice, adjust, control, height, distance, environment, noise, proportion, pupil, image, microphone																										
C. 相关文件 <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 106648527 A (乐视控股北京有限公司等) 2017年 5月 10日 (2017 - 05 - 10) 说明书第0056-0103、0107-0111段, 图2、4-6</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 106534541 A (广东小天才科技有限公司) 2017年 3月 22日 (2017 - 03 - 22) 说明书第0037-0038、0049段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 103793719 A (深圳大学) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0035段</td> <td>2-3, 8, 10-11, 16-17</td> </tr> <tr> <td>Y</td> <td>CN 106377264 A (广州视源电子科技股份有限公司) 2017年 2月 8日 (2017 - 02 - 08) 说明书第0058-0061段</td> <td>3, 11, 17</td> </tr> <tr> <td>Y</td> <td>CN 205754809 U (深圳市德宝威科技有限公司) 2016年 11月 30日 (2016 - 11 - 30) 说明书第0016-0025段, 图2</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 101867739 A (中山大学) 2010年 10月 20日 (2010 - 10 - 20) 说明书第0030-0059段, 图1-2</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 102981422 A (广州华多网络科技有限公司) 2013年 3月 20日 (2013 - 03 - 20) 说明书第0069-0108段</td> <td>1-20</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 106648527 A (乐视控股北京有限公司等) 2017年 5月 10日 (2017 - 05 - 10) 说明书第0056-0103、0107-0111段, 图2、4-6	1-20	Y	CN 106534541 A (广东小天才科技有限公司) 2017年 3月 22日 (2017 - 03 - 22) 说明书第0037-0038、0049段	1-20	Y	CN 103793719 A (深圳大学) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0035段	2-3, 8, 10-11, 16-17	Y	CN 106377264 A (广州视源电子科技股份有限公司) 2017年 2月 8日 (2017 - 02 - 08) 说明书第0058-0061段	3, 11, 17	Y	CN 205754809 U (深圳市德宝威科技有限公司) 2016年 11月 30日 (2016 - 11 - 30) 说明书第0016-0025段, 图2	1-20	Y	CN 101867739 A (中山大学) 2010年 10月 20日 (2010 - 10 - 20) 说明书第0030-0059段, 图1-2	1-20	Y	CN 102981422 A (广州华多网络科技有限公司) 2013年 3月 20日 (2013 - 03 - 20) 说明书第0069-0108段	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
Y	CN 106648527 A (乐视控股北京有限公司等) 2017年 5月 10日 (2017 - 05 - 10) 说明书第0056-0103、0107-0111段, 图2、4-6	1-20																								
Y	CN 106534541 A (广东小天才科技有限公司) 2017年 3月 22日 (2017 - 03 - 22) 说明书第0037-0038、0049段	1-20																								
Y	CN 103793719 A (深圳大学) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0035段	2-3, 8, 10-11, 16-17																								
Y	CN 106377264 A (广州视源电子科技股份有限公司) 2017年 2月 8日 (2017 - 02 - 08) 说明书第0058-0061段	3, 11, 17																								
Y	CN 205754809 U (深圳市德宝威科技有限公司) 2016年 11月 30日 (2016 - 11 - 30) 说明书第0016-0025段, 图2	1-20																								
Y	CN 101867739 A (中山大学) 2010年 10月 20日 (2010 - 10 - 20) 说明书第0030-0059段, 图1-2	1-20																								
Y	CN 102981422 A (广州华多网络科技有限公司) 2013年 3月 20日 (2013 - 03 - 20) 说明书第0069-0108段	1-20																								
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																										
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “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 “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 “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 “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 “&” 同族专利的文件																										
国际检索实际完成的日期 2018年 12月 25日		国际检索报告邮寄日期 2019年 1月 4日																								
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 白家荣 电话号码 86-(010)-62411469																								

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/102853

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	106648527	A	2017年 5月 10日	无	
CN	106534541	A	2017年 3月 22日	无	
CN	103793719	A	2014年 5月 14日	无	
CN	106377264	A	2017年 2月 8日	WO 2018072598 A1	2018年 4月 26日
CN	205754809	U	2016年 11月 30日	无	
CN	101867739	A	2010年 10月 20日	无	
CN	102981422	A	2013年 3月 20日	CN 102981422 B	2015年 12月 23日