



(12) 发明专利申请

(10) 申请公布号 CN 118714981 A

(43) 申请公布日 2024. 09. 27

(21) 申请号 202380021875.7

(22) 申请日 2023.01.30

(30) 优先权数据

22156800.9 2022.02.15 EP

(85) PCT国际申请进入国家阶段日

2024.08.14

(86) PCT国际申请的申请数据

PCT/EP2023/052131 2023.01.30

(87) PCT国际申请的公布数据

W02023/156174 EN 2023.08.24

(71) 申请人 皇家飞利浦有限公司

地址 荷兰艾恩德霍芬市

(72) 发明人 R·K·古杰拉尔 R·B·帕蒂尔

S·古塔

(74) 专利代理机构 北京市金杜律师事务所

11256

专利代理师 陈雪梅

(51) Int.Cl.

A61B 18/20 (2006.01)

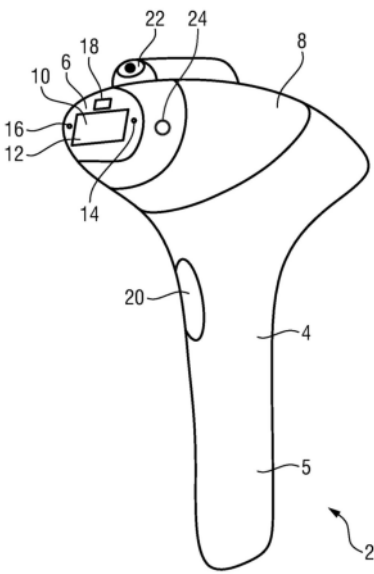
权利要求书4页 说明书17页 附图9页

(54) 发明名称

与处理设备一起使用的装置和方法

(57) 摘要

本发明涉及用于与处理设备一起使用的装置和方法,该处理设备被配置为对受试者的身体部位执行处理操作,其中处理设备被配置为将光脉冲施加到身体部位的皮肤以执行处理操作并针对每个光脉冲输出闪光声音。在实施例中使用经训练的算法或计算系统来可靠地检测闪光声音并通知用户避免过度处理。



1. 一种被配置为与处理设备 (2) 一起使用的装置 (42), 所述处理设备被配置为对受试者的身体部位执行处理操作, 其中所述处理设备被配置为将光脉冲施加到所述身体部位的皮肤以执行所述处理操作, 并且针对由所述处理设备所生成的每个光脉冲而生成闪光声音, 所述装置 (42) 包括:

- 输入 (48), 所述输入被配置为获得在所述处理操作期间感测到的音频信号;
- 处理单元 (50), 所述处理单元被配置为:
 - 将所述音频信号分割成音频片段, 特别是分割成具有设定或预定持续时间的音频片段,
 - 针对每个音频片段从所述音频片段中提取一个或多个特性特征,
 - 通过应用经训练的算法或计算系统来针对每个音频片段检测所述音频片段是否包括针对每个光脉冲而由所述处理设备所生成的闪光声音, 所述经训练的算法或计算系统已在多个音频片段上训练过, 以检测在所述音频片段中针对每个光脉冲而由所述处理设备所生成的闪光声音, 所述多个音频片段各自包括所述特性特征中的一个或多个特性特征, 以及
 - 改变闪光声音计数器, 所述闪光声音计数器计数自所述处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量; 以及
 - 输出 (52), 所述输出被配置为输出以下中的一项或多项:
 - 所述闪光声音计数器,
 - 关于所述闪光声音计数器是否超过闪光声音阈值的指示, 以及
 - 关于所述处理操作已完成的指示。

2. 根据权利要求1所述的装置,

其中所述处理单元 (50) 被配置为使用学习系统、机器学习、分类模型、二元分类模型、神经网络或卷积神经网络作为经训练的算法或计算系统。

3. 根据前述权利要求中任一项所述的装置,

其中所述处理单元 (50) 被配置为提取与所述音频片段的时间、幅度和频率相关的一个或多个特征作为特性特征。

4. 根据权利要求3所述的装置, 其中所述处理单元 (50) 被配置为提取以下中的一项或多项作为音频片段的特性特征: 开始时间、结束时间、幅度、梯度、乐器、主和弦、旋律、节奏、音高、能量、音符偏移、一个或多个梅尔频率倒谱系数 (MFCC)、梅尔频谱图、频谱质心、频谱通量、零交叉率、节拍、音色、音高、和声、均方根能量、幅度包络、频带能量比、频谱图、常数Q变换、频谱滚降和频谱扩展。

5. 根据权利要求3或4所述的装置, 其中所述处理单元 (50) 被配置为将与所述音频片段的时间、幅度和频率相关的所提取的特征排列成阵列, 并使所述阵列接受所述经训练的算法或计算系统以针对每个音频片段检测所述音频片段是否包括闪光声音。

6. 根据前述权利要求中任一项所述的装置,

其中由所述处理单元 (50) 将所述经训练的算法或计算系统应用到音频片段上包括基于所述一个或多个特性特征来构建分类边界。

7. 根据前述权利要求中任一项所述的装置,

其中所述处理单元 (50) 被配置为使用所述检测的结果来进一步训练所述算法或计算系统。

8. 一种被配置为与处理设备 (2) 一起使用的装置, 所述装置特别是如前述权利要求中任一项所述的装置, 所述处理设备被配置为对受试者的身体部位执行处理操作, 其中所述处理设备被配置为将光脉冲施加到所述身体部位的皮肤以执行所述处理操作, 并且针对由所述处理设备所生成的每个光脉冲而生成闪光声音, 所述装置 (42) 包括:

- 输入 (48), 所述输入被配置为获得人工创建的或在所述处理操作期间感测到的音频信号;

- 处理单元 (50), 所述处理单元被配置为:

- 将所述音频信号分割成音频片段, 特别是分割成具有设定或预定持续时间的音频片段,

- 针对每个音频片段从所述音频片段中提取一个或多个特性特征,

- 针对每个音频片段从所述音频信号的注释中导出闪光声音信息, 所述闪光声音信息指示音频片段是否包括针对每个光脉冲而由所述处理设备所生成的闪光声音, 所述音频信号的所述注释指示在所述音频信号内针对每个光脉冲而由所述处理设备所生成的闪光声音, 特别是针对每个光脉冲而由所述处理设备所生成的闪光声音的时间信息, 以及

- 针对每个音频片段将所提取的所述一个或多个特性特征和所导出的所述闪光声音信息提供给算法或计算系统, 以训练所述算法或计算系统用于基于一个或多个特性特征来检测在音频片段内针对每个光脉冲而由所述处理设备所生成的闪光声音。

9. 根据权利要求8所述的装置,

其中所述输入 (48) 被配置为获得所述音频信号的所述注释, 和/或

其中所述处理单元 (50) 被配置为生成所述音频信号的所述注释, 特别是自动地或基于用户输入来生成所述音频信号的所述注释。

10. 根据权利要求1-7中任一项所述的装置, 包括输出 (52), 所述输出被配置为控制用户界面以:

- 显示第一屏幕, 所述第一屏幕请求关于所述处理操作的用户输入;

- 经由所述第一屏幕获得所述用户输入, 所述用户输入包括待处理的所述受试者的所述身体部位的图像和/或待处理的所述受试者的所述身体部位的尺寸的尺寸信息;

- 显示第二屏幕, 所述第二屏幕指示所述装置或包括所述装置的电子设备与所述处理设备的连接性;

- 显示第三屏幕, 所述第三屏幕指示以下中的至少一项:

- 所述皮肤处理设备的操作状态,

- 所述闪光声音计数器, 所述闪光声音计数器计数自所述处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量,

- 关于所述闪光声音计数器是否超过闪光声音阈值的指示, 以及

- 关于所述处理操作已完成的指示。

11. 一种系统, 包括:

处理设备 (2), 所述处理设备包括一个或多个光源 (14、16), 所述一个或多个光源被配置为生成光脉冲以施加到受试者身体的皮肤以执行处理操作, 其中所述处理设备 (2) 被配置为针对由所述处理设备所生成的每个光脉冲而生成闪光声音;

音频传感器 (44), 所述音频传感器被配置为在所述处理操作期间感测音频信号; 以及

如前述权利要求中任一项所述的装置(42)。

12. 一种被配置为与处理设备(2)一起使用的方法,所述处理设备被配置为对受试者的身体部位执行处理操作,其中所述处理设备被配置为将光脉冲施加到所述身体部位的皮肤以执行所述处理操作,并且针对由所述处理设备所生成的每个光脉冲而生成闪光声音,所述方法包括:

- 获得在所述处理操作期间感测到的音频信号,
- 将所述音频信号分割成音频片段,特别是分割成具有设定或预定持续时间的音频片段,
- 针对每个音频片段从所述音频片段中提取一个或多个特性特征,
- 通过应用经训练的算法或计算系统来针对每个音频片段检测所述音频片段是否包括针对每个光脉冲而由所述处理设备所生成的闪光声音,所述经训练的算法或计算系统已经在多个音频片段上训练过,以检测在所述音频片段中针对每个光脉冲由所述处理设备所生成的闪光声音,所述多个音频片段各自包括所述特性特征中的一个或多个特性特征,以及
- 改变闪光声音计数器,所述闪光声音计数器计数自处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量;以及
- 输出以下中的一项或多项:
 - 所述闪光声音计数器,
 - 关于所述闪光声音计数器是否超过闪光声音阈值的指示,以及
 - 关于所述处理操作已完成的指示。

13. 一种配置为与处理设备(2)一起使用的方法,特别是如权利要求12中所述的方法,所述处理设备被配置为对受试者的身体部位执行处理操作,其中所述处理设备被配置为将光脉冲施加到所述身体部位的皮肤以执行所述处理操作,并且针对由所述处理设备所生成的每个光脉冲而生成闪光声音,所述方法包括:

- 获得人工创建的或在处理操作期间感测到的音频信号,
- 将所述音频信号分割成音频片段,特别是分割成具有设定或预定持续时间的音频片段,
- 针对每个音频片段从所述音频片段中提取一个或多个特性特征,
- 针对每个音频片段从所述音频信号的注释中导出闪光声音信息,所述闪光声音信息指示音频片段是否包括针对每个光脉冲而由所述处理设备所生成的闪光声音,所述音频信号的所述注释指示在所述音频信号内针对每个光脉冲而由所述处理设备所生成的闪光声音,特别是针对每个光脉冲而由所述处理设备所生成的闪光声音的时间信息,以及
- 针对每个音频片段将所提取的所述一个或多个特性特征和所导出的所述闪光声音信息提供给算法或计算系统,以训练所述算法或计算系统用于基于一个或多个特性特征来检测在音频片段内针对每个光脉冲而由所述处理设备所生成的闪光声音。

14. 根据权利要求12所述的方法,所述方法被配置为控制用户界面以:

- 显示第一屏幕,所述第一屏幕请求关于所述处理操作的用户输入;
- 经由所述第一屏幕获得所述用户输入,所述用户输入包括待处理的所述受试者的所述身体部位的图像和/或待处理的所述受试者的所述身体部位的尺寸的尺寸信息;
- 显示第二屏幕,所述第二屏幕指示包括所述用户界面的装置或电子设备与所述处理

设备的连接性；

- 显示第三屏幕,所述第三屏幕指示以下中的至少一项:
- 所述皮肤处理设备的操作状态,
- 所述闪光声音计数器,所述闪光声音计数器计数自所述处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量,
- 关于所述闪光声音计数器是否超过闪光声音阈值的指示,以及
- 关于所述处理操作已完成的指示。

15.一种计算机程序,包括程序代码部件,用于当所述计算机程序在计算机上执行时使用所述计算机执行根据权利要求12、13或14所述的方法的步骤。

与处理设备一起使用的装置和方法

技术领域

[0001] 本发明涉及与处理设备一起使用的装置和方法,该处理设备被配置为对受试者的身体部位执行处理操作,其中该处理设备被配置为将光脉冲施加到身体部位的皮肤以执行处理操作,并且针对由处理设备生成的每个光脉冲生成闪光声音。本发明还涉及此类装置的计算机实现的方法、包括处理设备的系统和计算机程序。

背景技术

[0002] 去除多余毛发的技术包括剃须、电解、拔毛、激光和光疗法(称为光脱毛)以及治疗性抗雄激素的注射。基于光的技术还被用于其他类型的皮肤病处理,包括减少毛发生长和处理痤疮。

[0003] 通过使用适当的光能配置,即波长、强度和/或脉冲持续时间(如果光是脉冲的)方面,可以实现对毛根的选择性加热,随后对毛囊造成暂时或永久性损伤。家用光脱毛设备(例如飞利浦Lumea设备)使用来自高强度光源(例如氙气闪光灯)的强脉冲光(IPL),其可以产生高输出的广谱光脉冲。

[0004] 光脱毛处理的特点是光脱毛设备的用户处理相对较小的皮肤区域以达到脱毛的目的。光脱毛处理使用强光加热头发和毛根中的黑色素,使毛囊进入休止期,防止毛发再生。为了有效利用这项技术脱毛,用户必须完全处理皮肤,而不留任何缝隙。由于这种效果仅保持有限的持续时间,因此必须定期重复处理:通常,在最初约两个月的处理期(每两周进行一次处理)之后,在维持阶段每4至8周进行一次。

[0005] Lumea设备由Lumea app(即,可以在计算机、笔记本电脑、平板电脑和/或智能手机上运行的应用程序或软件应用,其可以无线耦合到处理设备)支持,以帮助用户进行脱毛处理。这增强了用户体验和用户对Lumea设备的参与度。

[0006] 在许多光脱毛设备中,每次发出用于处理的闪光/爆发光后都会立即发出闪光声音,这是一种在时间(例如大约一秒)和频率(例如大约9kHz)上都特定的短声脉冲。处理所需的闪光次数取决于要被应用脉冲的身体部位的尺寸。高强度激光脉冲曝光持续时间超过推荐量可能会导致过度处理,并且严重损伤皮肤。

[0007] WO 2019/079972 A1公开了一种特定语音识别方法,借助于该方法能够以简单的算法、较小的计算量和较低的硬件设备要求识别特定语音。该方法包括对语音信号进行采样,获得语音信号的梅尔频率倒谱系数特性参数矩阵,从特性参数矩阵中提取特性参数。将特性参数供应给基于深度神经网络的特定语音特征模型进行识别,以便确定语音信号是否包括特定语音特征。特定语音识别装置被配置为接收用户的语音并确定语音是否包含特定语音特征。WO 2019/079972 A1中提到的特定语音特征的一个例子是咳嗽声。

发明内容

[0008] 本发明的目的是提供避免或至少降低过度处理风险的措施。

[0009] 在本发明的第一方面,提出了一种被配置为与处理设备一起使用装置,该处理设

备被配置为对受试者的身体部位执行处理操作,其中处理设备被配置为将光脉冲施加到身体部位的皮肤以执行处理操作,并且针对由处理设备所生成的光脉冲而生成闪光声音,该装置包括:

- [0010] -输入,其被配置为获得在处理操作期间感测到的音频信号;
- [0011] -处理单元,其被配置为:
- [0012] -将音频信号分割成音频片段,特别是分割成具有设定或预定持续时间的音频片段,
- [0013] -针对每个音频片段从音频片段中提取一个或多个特性特征,
- [0014] -通过应用经训练的算法或计算系统来针对每个音频片段检测音频片段是否包括针对每个光脉冲而由处理设备所生成的闪光声音,该经训练的算法或计算系统已在多个音频片段上训练过,以检测在音频片段中针对每个光脉冲而由处理设备所生成的闪光声音,所述多个音频片段各自包括这些特征中的一个或多个特性特征,以及
- [0015] -改变闪光声音计数器,该闪光声音计数器计数自处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量;以及-输出,其被配置为输出以下中的一项或多项:
- [0016] -闪光声音计数器,
- [0017] -关于闪光声音计数器是否超过闪光声音阈值的指示,以及
- [0018] -关于处理操作已完成的指示。

[0019] 在本发明的第二方面中,提出了一种被配置为与处理设备一起使用的装置,该处理设备被配置为对受试者的身体部位执行处理操作,其中处理设备被配置为将光脉冲施加到身体部位的皮肤以执行处理操作,并且针对由处理设备所生成的每个光脉冲而生成闪光声音,该装置包括:

- [0020] -输入,其被配置为获得人工创建的或在处理操作期间感测到的音频信号;
- [0021] -处理单元,其被配置为:
- [0022] -将音频信号分割成音频片段,特别是分割成具有设定或预定持续时间的音频片段,
- [0023] -针对每个音频片段从音频片段中提取一个或多个特性特征,
- [0024] -针对每个音频片段从音频信号的注释中导出闪光声音信息,该闪光声音信息指示音频片段是否包括针对每个光脉冲而由处理设备所生成的闪光声音,音频信号的注释指示在音频信号内针对每个光脉冲而由处理设备所生成的闪光声音,特别是针对每个光脉冲而由处理设备所生成的闪光声音的时间信息,以及
- [0025] -针对每个音频片段将所提取的一个或多个特性特征和所导出的闪光声音信息提供给算法或计算系统,以训练算法或计算系统用于基于一个或多个特性特征来检测在音频片段内针对每个光脉冲而由处理设备所生成的闪光声音。
- [0026] 在本发明的一个实施例中,该装置包括输出,其被配置为控制用户界面以:
- [0027] -显示第一屏幕,其请求关于所述处理操作的用户输入;
- [0028] -经由第一屏幕获得用户输入,该用户输入包括待处理的受试者的身体部位的图像和/或待处理的受试者的身体部位的尺寸的尺寸信息;
- [0029] -显示第二屏幕,其指示装置或包括该装置的电子设备与处理设备的连接性;

- [0030] -显示第三屏幕,其指示以下中的至少一项:
- [0031] -皮肤处理设备的操作状态,
- [0032] -闪光声音计数器,其计数自处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量,
- [0033] -关于闪光声音计数器是否超过闪光声音阈值的指示,以及
- [0034] -关于处理操作已完成的指示。
- [0035] 在本发明的第三方面中,提出了一种系统,包括:
- [0036] 处理设备,其包括一个或多个光源,被配置为生成光脉冲以施加到受试者身体的皮肤以执行处理操作,其中处理设备被配置为针对由处理设备所生成的每个光脉冲而生成闪光声音;
- [0037] 音频传感器,其被配置为在处理操作期间感测音频信号;以及
- [0038] 根据本文所公开的不同方面中的任一个装置。
- [0039] 在本发明的其他方面,提供了对应的(计算机实现的)方法、计算机程序,其包括程序代码部件,用于当计算机程序在计算机上执行时使计算机执行本文所公开的对应方法的步骤,以及非暂时性计算机可读记录介质,其中存储计算机程序产品,该计算机程序产品当由处理器执行时导致执行本文所公开的对应方法。
- [0040] 本发明的优选实施例在从属权利要求中定义。应当理解,所要求保护的方法、系统、计算机程序和介质具有与所要求保护的装置相似和/或相同的优选实施例,特别是如从属权利要求中所定义和本文所公开的。
- [0041] 本发明的所有方面都基于这样的共同思想:考虑到激光脱毛处理的有害方面,用户应该精确地知道处理所需的闪光次数和/或应该知道期望区域的处理何时已完成,并且已用足够数量的闪光进行处理。单次处理通常需要例如大约30-40次闪光(取决于期望处理区域的长度)。用户可以手动计算在一次处理期间所应用的闪光次数,并且没有推荐任何闪光次数,这导致用户体验不佳。然而,用户可能忙于其他任务或受到背景噪声的干扰,因此他们可能难以准确地跟踪闪光的总数。因此,本发明(第一方面)使用经训练的算法或计算系统提供在处理期间自动计数闪光声音(以及因此的发射的闪光次数)。这种经训练的算法或计算系统可以被训练以精确地计数每次处理期间的总闪光次数并应对来自处理设备和环境的背景噪声。此外,可以考虑信号的非周期性。因此,可以以更高的准确度获得皮肤处理设备在处理期间发出的闪光次数。
- [0042] 在本发明的上下文中,改变用于对处理期间的闪光声音数量进行计数的计数器应被理解为,使得每次检测到闪光声音时,计数器值递增(例如从零开始)或递减(例如从用于处理某个处理区域的估计最大计数器值开始)。
- [0043] 根据本发明(第二方面),通过使用多个音频片段进行训练,每个音频片段包括从音频片段(例如持续时间为一秒的片段)中导出的一个或多个特性特征,以检测音频片段中的闪光声音。
- [0044] 此外,根据本发明(第三方面),提供了一种用户界面,该用户界面有效地获得并向用户呈现各种信息,包括处理期间发出的闪光次数。它有助于用户轻松理解处理的当前进展并避免过度处理。
- [0045] 优选地,本发明使用学习系统、机器学习、分类模型、二元分类模型、神经网络或卷

神经网络作为经训练的算法或计算系统。例如,在实现中,可以使用传统上已知的XGBoost分类器。XGBoost是一种基于决策树的集成机器学习算法,它使用梯度提升框架。它可以被用于广泛的应用,例如解决回归、分类、排名和用户定义的预测问题。它可以在训练阶段(当系统/算法被训练时)以及在操作阶段(当系统/算法被用于闪光声音检测时)以最短的处理时间提供高预测性能。

[0046] 在另一个实施例中,可以使用与音频片段的时间、幅度和频率相关的一个或多个特征作为(多个)特性特征,其可以从相应的音频片段中被提取。

[0047] 通常,有多种特征可以用作音频片段的(多个)特性特征,包括以下中的一项或多项:开始时间、结束时间、幅度、梯度、乐器、主和弦、旋律、节奏、音高、能量、音符偏移、一个或多个梅尔频率倒谱系数(MFCC)、梅尔频谱图、频谱质心、频谱通量、零交叉率、节拍、音色、音高、和声、均方根能量、幅度包络、频带能量比、频谱图、常数Q变换、频谱滚降和频谱扩展中的一个或多个。在一个实现中,可以使用具有8-13个倒谱系数的选定MFCC特征,而不是通常可用于音频片段的大量、例如39个MFCC特征。最好不使用表示仅携带少量说话者特定信息的输入信号的平均对数能量的第零系数,因为它可能导致检测到背景噪声(闪光声音通常是瞬态的~0.1秒)。因此,根据本发明,可以有效地检测光脱毛设备在存在各种背景噪声的情况下产生的具有低幅度和短持续时间的闪光声音。此外,可以提供所需的低延迟和记忆。

[0048] 闪光声音可以通过触发光源、诸如闪光灯来生成,或者在使用光源、诸如LED的情况下可以人工合成。例如,闪光灯或氙弧灯在操作期间由于脉冲放电而创建声音(表示闪光声音)。激光辐射会引起空气中的电击穿,从而产生声音。类似地,当强脉冲光与皮肤接触时会生成声音。然而,处理设备的其他实施例可以使用半导体光源,诸如LED,由于它们不同地起作用,因此不会产生放电声。然后,每次发射用于处理的光脉冲时,可以使用附加的声音发生器来合成闪光声音。

[0049] 根据第一方面的另一个实施例,处理单元可以被配置为将与音频片段的时间、幅度和频率相关的所提取的特征排列成阵列,并使该阵列接受经训练的算法或计算系统以针对每个音频片段检测音频片段是否包括闪光声音。阵列格式可以例如被配置为使得所提取的特征根据与时间相关、与幅度相关和与频率相关的特征来排列或排序,这使得可以更容易甚至可能用经训练的算法或计算系统、诸如XGBoost分类模型处理它们。

[0050] 由处理单元将经训练的算法或计算系统应用到音频片段上可以包括基于一个或多个特性特征来构建分类边界。例如,XGBoost使用基于集成树的技术基于所传递的输入特征来构建分类边界。它使用并行处理、树修剪和正则化来避免过度拟合并提供平衡模型。基于分类边界,每个音频片段可以被分类为闪光和非闪光,以使得最终可以预测处理期间的闪光总数。

[0051] 在第一方面的另一个实施例中,处理单元被配置为使用检测的结果来进一步训练算法或计算系统。因此,在操作期间,检测结果不仅被用于闪光声音检测,而且同时被用于进一步训练以进一步改进经训练的算法或计算系统。

[0052] 根据本发明的第二方面,人工创建的或在处理操作期间感测到的音频信号被用于训练算法或计算系统。以与实际处理操作期间相同的方式,将音频信号分割成音频片段,并从音频片段中提取一个或多个特性特征。从音频信号闪光声音的注释中,根据一个或多个

特性特征导出指示音频片段是否包括闪光声音的信息。在该上下文中,音频信号的注释应被广泛理解为针对音频信号内的闪光声音的指示,特别是闪光声音的时间信息(例如,时间戳或其他指示,其指示音频信号在哪个时间点包括闪光声音,诸如开始时间、结束时间、峰值时间、持续时间等)。然后使用所提取的特性特征和所导出的闪光声音信息来基于一个或多个特性特征来训练算法或计算系统,以检测在音频片段内的闪光声音。

[0053] 音频信号的注释可以由用户预先进行,以使得输入可以获得音频信号的注释,例如与音频信号一起存储或与音频信号分开存储。替代地或另外地,可以自动地进行注释,例如由算法或计算系统或另一个处理实体进行。例如,处理单元可以被配置为生成音频信号的注释,例如自动地或基于用户输入来生成音频信号的注释。

[0054] 根据本发明的第三方面,可以获得不同的发明并将其呈现给用户。该信息可以被显示在各种屏幕上,其中第一屏幕请求关于处理操作的用户输入,第二屏幕指示装置或包括该装置的电子设备与处理设备的连接性,第三屏幕指示皮肤处理设备的操作状态、用于计数自处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量的闪光声音计数器、关于闪光声音计数器是否超过闪光声音阈值的指示、以及关于处理操作已完成的指示。这使得用户能够知道处理操作的当前进度以及何时结束处理,从而可以避免或至少减少过度处理。

[0055] 根据本发明的系统包括处理设备、音频传感器和如上所述的装置中的一个或多个。根据系统的实施例,音频传感器、例如麦克风是装置或包括该装置的电子设备的一部分,特别是计算机、笔记本电脑、平板电脑或智能手机(或仅表示装置的单独实体)的一部分。在另一个实施例中,该装置是处理设备或电子设备的一部分,特别是计算机、笔记本电脑、平板电脑或智能手机。通常,上述方面的装置可以在同一设备上被实现,也可以在不同设备上被实现。在实际实现中,所有方面的装置和方法都在同一设备上被实现,诸如平板电脑或智能手机。

附图说明

[0056] 本发明的这些和其他方面将从下文描述的(多个)实施例中显而易见并参考这些实施例进行阐述。在以下附图中

[0057] 图1示出了可以使用本发明的处理设备的示例性实施例的图示;

[0058] 图2示出了根据本发明的系统的示例性实施例的示意图;

[0059] 图3示出了根据本发明的第一方面的方法的实施例的流程图;

[0060] 图4示出了根据本发明的第二方面的方法的实施例的流程图;

[0061] 图5示出了具有对闪光声音和背景噪声的指示的音频信号的可视化;

[0062] 图6示出了两个闪光声音的放大可视化;

[0063] 图7示出了根据本发明的预处理的示意图;

[0064] 图8示出了根据本发明的特征提取的示意图;

[0065] 图9示出了根据本发明的从模型部署到生产的过程的流程图;

[0066] 图10示出了根据本发明的第三方面的方法的实施例的流程图;以及

[0067] 图11示出了图示出根据本发明的用户界面的不同屏幕。

具体实施方式

[0068] 图1是可以被用于将光脉冲(也称为闪光)施加到皮肤区域的示例性处理设备2的图示。应当了解,图1中的处理设备2仅作为本发明可以使用的手持式处理设备2的示例呈现,并且处理设备2不限于图1中所示的形式或手持式处理设备。处理设备2用于使用在受试者(例如人或动物)的身体上,并且在使用过程中由用户的一只手或两只手握住。处理设备2用于在其接触或靠近受试者的身体部位时使用一个或多个光脉冲对受试者的皮肤或身体执行某些处理操作。处理操作可以通过激光和/或光疗法去除不需要的毛发(称为光脱毛处理或强脉冲光IPL处理)。替代地,处理操作可以是用于去除或防止毛发生长之外的其他目的的激光和/或光疗法,诸如皮肤病学(皮肤)处理、处理痤疮、用于其他类型皮肤状况的光疗法处理、皮肤再生、皮肤紧致或葡萄酒色斑处理。

[0069] 如本文所述,处理设备2由“用户”操作或使用,并且处理设备2被使用在“受试者”的身体上。在一些情况下,用户和受试者是同一个人,即,处理设备2被用户握在手中并由用户使用于自己(例如,使用在腿部皮肤上)。在其他情况下,用户和受试者是不同的人,例如处理设备2被用户握在手中并且由用户使用于其他人。在任一情况下,用户都难以完全覆盖身体部位和/或避免过度处理身体部位的某些区域,因为在应用光脉冲时或应用光脉冲后不久,皮肤几乎没有或完全没有用户可察觉的变化。

[0070] 示例性处理设备2包括壳体4,该壳体至少包括手柄部分5和头部部分6。手柄部分5被成形为使得用户能够用一只手握住处理设备2。头部部分6位于壳体4的头端8处,并且头部部分6用于被放置成与受试者接触,以便在头部部分6与受试者的身体或皮肤接触的位置处对受试者的身体或皮肤执行处理操作。

[0071] 处理设备2用于使用光脉冲来执行处理操作。因此,在图1中,头部部分6包括孔10,该孔被布置在壳体4中或壳体上,使得孔10可以被放置在受试者的皮肤附近或皮肤上(即接触皮肤)。处理设备2包括一个或多个光源12,其用于生成将经由孔10而被施加到受试者的皮肤上并实施处理操作的光脉冲。一个或多个光源12被布置在壳体4中,使得通过孔10从一个或多个光源12提供光脉冲。孔10可以是壳体4的头端8处的开口形式,或者它可以是对光脉冲是透明或半透明的(即光脉冲可以穿过窗口)窗口(包括波导)的形式。

[0072] 在图1中所示的示例性实施例中,孔10具有大致矩形的形状,这导致皮肤上大致矩形的皮肤处理区域。可以了解,孔10可以具有任何其他期望的形状。例如,孔10可以是正方形、椭圆形、圆形或任何其他多边形。

[0073] 虽然图1中未示出,但是处理设备2可以具有用于处理设备2的头部部分6的一个或多个可拆卸附件,每个附件都包括孔10。附件可以用于使用在不同的身体部位上,并且具有不同尺寸的孔10。例如,一个附件可以用于使用在较大的身体部位(诸如腿)上,并且因此具有较大的孔10,而另一个附件可以用于使用在上唇上方的皮肤/头发上,并且因此具有较小的孔10。

[0074] 一个或多个光源12可以生成任何合适或期望波长(或波长范围)和/或强度的光脉冲。例如,光源12可以生成可见光、红外(IR)光和/或紫外(UV)光。每个光源12可以包括任何合适类型的光源,诸如一个或多个发光二极管(LED)、(氙气)闪光灯、一个或多个激光器等。光源12可以提供光谱内容在450-1200nm范围内的光脉冲,持续时间约为2.5ms,因为这些波长通过吸收而加热头发和发根中的黑色素,这使毛囊处于休止期,防止毛发再生。

[0075] 一个或多个光源12被配置为提供光脉冲。也就是说, (多个) 光源12被配置为在短持续时间内(例如, 少于1秒)生成高强度的光。光脉冲的强度应足够高, 以对邻近孔10的皮肤或身体部位实施处理操作。

[0076] 所图示的处理设备2可以可选地包括两个皮肤接触传感器14、16, 它们被定位在头部部分6上或头部部分中, 其被用于确定头部部分6是否与皮肤接触。皮肤接触传感器14、16测量指示头部部分6是否与皮肤接触的参数, 并生成包括该参数的时间序列测量值的相应测量信号。测量信号可以被处理以确定头部部分6是否与皮肤接触。通常, 皮肤接触传感器被用于处理设备2、特别是光脱毛器中, 以确保在生成光脉冲之前处理设备2与皮肤正确接触, 以避免光脉冲被引导到用户或受试者的眼睛中。

[0077] 在一些实施例中, 该参数可以是电容, 并且因此皮肤接触传感器14、16可以经由头部部分6的表面上的相应一对电触点或电极来测量电容, 其中测量到的电容指示是否存在皮肤接触。在替代实施例中, 该参数可以是光的强度或水平, 并且因此皮肤接触传感器14、16可以是光传感器, 其测量入射到光传感器上的光的强度或水平, 其中测量到的强度或水平指示是否存在皮肤接触(例如, 当皮肤遮挡光传感器14、16时, 较少或没有光可能指示皮肤接触, 反之亦然)。在其他替代实施例中, 该参数可以是接触压力的测量, 并且因此皮肤接触传感器14、16可以经由相应的压力传感器或机械开关来测量接触压力, 其中测量到的接触压力指示是否存在皮肤接触。

[0078] 所图示的处理设备2可以包括可选的肤色传感器18, 该肤色传感器被定位在头部部分6上或头部部分内, 其可以被用于确定头部部分6所接触的皮肤的肤色。肤色传感器18可以测量指示皮肤的肤色的参数, 并生成包括参数的测量的时间序列的测量信号。测量信号可以被处理以确定头部部分6所接触的皮肤的肤色。通常, 肤色传感器被用于处理设备2、特别是光脱毛器中, 以确保光脉冲具有适合所处理皮肤类型的强度, 或者甚至在皮肤类型不适合光脉冲(例如, 具有高得多的黑色素含量的较暗皮肤)时防止生成光脉冲。

[0079] 在一些实施例中, 可选的肤色传感器18可以是光传感器, 并且由光传感器测量的参数可以是皮肤反射的特定的一个或多个波长的光的强度或水平。在(多个)特定波长的反射光的所测量的强度或水平可以指示肤色。反射光的所测量的强度或水平可以基于皮肤中的黑色素浓度, 因此所测量的强度或水平可以指示黑色素浓度。黑色素浓度可以例如通过测量660nm(红色)和880nm(红外)波长的光反射中被导出。

[0080] 所图示的处理设备2还包括用户控件20, 该用户控件可以被用户操作以激活处理设备2, 使得头部部分6对受试者的身体执行所需的处理操作(例如, 由一个或多个光源12生成一个或多个光脉冲)。用户控件20可以是开关、按钮、触摸板等形式。

[0081] 如上面所指出, 可以可选地使用一个或多个传感器来提供有关处理设备随时间的位置(包括取向)和/或运动的信息, 以使得可以确定合适的反馈, 并且在一些实施例中, 一个或多个传感器可以是处理设备2的一部分或位于该处理设备中。在(多个)传感器是一个或多个运动传感器、诸如加速度计、陀螺仪等的情况下, 一个或多个运动传感器可以是处理设备2的一部分, 例如位于处理设备2内部(并且因此未在图1中示出)。在一个或多个传感器包括成像单元(例如, 相机)的实施例中, 成像单元可以在处理设备2上或在处理设备中。如图1中所示, 成像单元22可以被布置在处理设备2的壳体4上, 以使得它能够获得处理设备2前方的皮肤或身体部位的图像或视频序列。也就是说, 成像单元22被布置在处理设备2的

壳体4上,以使得它从与(多个)光源12通过孔10发射光脉冲的方向大致平行的方向获得图像或视频序列。在一些实施例中,成像单元22被布置在壳体4上,使得当头部部分6与受试者接触时,使得光脉冲可以被施加到(多个)光源12通过孔10可见的皮肤部分,成像单元22无法查看与孔10和头端6接触的皮肤部分。在成像单元22被布置在处理设备2的壳体4上或壳体4中的实施例中,成像单元22相对于处理设备2的其余部分处于固定布置。

[0082] 处理设备2还可以包括音频源24(声音发生器),诸如扬声器,以在每次生成光脉冲(闪光)时发出短音频脉冲(也称为闪光声音)。当光源12是半导体源、诸如LED或VCSEL时尤其如此,这些半导体源不会生成弧光灯特有的脉冲放电闪光声音。在其他实施例中,当生成光脉冲时,光源可以(例如自动地)生成闪光声音;例如,每次发射光脉冲时,可以生成像滴答声一样的声音脉冲。闪光声音例如是短音频脉冲,其持续时间短(例如在0.1至2秒的范围内,诸如大约1秒),具有特定频率(例如9kHz或任何其他选定频率)或频率范围(例如在1至10kHz的范围内),并且具有允许用户听到闪光声音的强度。

[0083] 图2是根据本发明的示例性系统40的示意图,其包括处理设备2和用于与处理设备2一起使用的装置42,例如用于控制处理设备2,例如如图1中所示。系统40还包括音频传感器44(例如麦克风;也可以有多个音频传感器),其被配置为在处理操作期间感测音频信号,其可以优选地是装置42的一部分,但是也可以是外部元件。此外,系统40可以可选地包括反馈单元46(例如,图形用户界面或声音输出),用于向用户和/或受试者提供反馈,特别是关于处理操作的进展或用户要遵循的任何指令。

[0084] 如上面所指出,在一些实施例中,装置42可以是与处理设备2分开的装置,因此装置42可以是电子设备的形式,诸如智能手机、智能手表、平板电脑、个人数字助理(PDA)、笔记本电脑、台式计算机、远程服务器、智能镜子等。在其他实施例中,装置42,特别是由装置42提供的根据本发明的功能性,可以是处理设备2的一部分。

[0085] 装置42包括处理单元50,其通常控制装置42的操作并使得装置42能够执行本文描述的方法和技术。处理单元50还可以被配置为接收来自音频传感器44的测量信号,处理测量信号以确定要被提供的关于处理操作的反馈,并向反馈单元46提供反馈控制信号。因此,处理单元50可以被配置为在传感器44是装置42的一部分的实施例中直接接收来自音频传感器44的测量信号,或者在传感器44与装置42分开的实施例中经由另一组件接收来自音频传感器44的测量信号。在任一情况下,处理单元50可以包括或包含一个或多个输入端口或线路,以用于经由输入单元48接收测量信号。处理单元50还可以包括或包含一个或多个输出端口或线路,以用于经由输出单元52输出反馈控制信号。

[0086] 使用软件和/或硬件,处理单元50可以以多种方式来实现,以执行本文所描述的各种功能。处理单元50可以包括一个或多个微处理器或数字信号处理器(DSP),其可以使用软件或计算机程序代码进行编程以执行所需的功能和/或控制处理单元50的组件以实现所需的功能。处理单元50可以被实现为用于执行一些功能(例如放大器、前置放大器、模数转换器(ADC)和/或数模转换器(DAC))专用硬件和用于执行其他功能的处理器(例如一个或多个编程的微处理器、控制器、DSP和相关电路)的组合。在本公开的各种实施例中可以采用的组件的示例包括但不限于传统微处理器DSP、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、用于实现神经网络的硬件和/或所谓的人工智能(AI)硬件加速器(即,专门为AI应用设计的(多个)处理器或其他硬件,其可以与主处理器一起使用)和/或硬件。

[0087] 处理单元50可以可选地包括或与存储器单元(未示出)相关联。存储器单元可以存储数据、信息和/或信号(包括(多个)图像),以供处理单元50在控制装置42的操作时使用和/或在执行或实行本文所描述的方法时使用。在一些实施例中,存储器单元存储可由处理单元50执行的计算机可读代码,使得处理单元50执行一个或多个功能,包括本文所描述的方法。在特定实施例中,程序代码可以是智能手机、平板电脑、笔记本电脑、计算机或服务器的应用的形式。存储器单元可以包括任何类型的非暂时性机器可读介质,诸如高速缓存或系统存储器,包括易失性和非易失性计算机存储器,诸如随机存取存储器(RAM)、静态RAM(SRAM)、动态RAM(DRAM)、只读存储器(ROM)、可编程ROM(PROM)、可擦除PROM(EPROM)和电可擦除PROM(EEPROM),并且存储器单元可以以存储器芯片、光盘(诸如压缩盘(CD)、数字多功能光盘(DVD)或蓝光盘)、硬盘、磁带存储解决方案或固态设备(包括记忆棒、固态驱动器(SSD)、存储卡等)的形式来实现。

[0088] 在图2中所示的实施例中,由于装置42被示为与(多个)传感器44分离,因此装置42可以包括输入单元48、诸如接口电路系统,以使得装置42能够接收来自(多个)传感器44的(多个)测量信号。装置42中的接口电路系统48使得能够与其他设备(包括(多个)传感器44、处理设备2、服务器和数据库中的一个或多个)建立数据连接和/或进行数据交换。与(多个)传感器44(或任何电子设备,诸如处理设备2)的连接可以是直接的或间接的(例如经由互联网),并且因此接口电路系统48可以经由任何期望的有线或无线通信协议来实现装置42与网络之间的连接,或装置42与另一设备(诸如(多个)传感器44和/或处理设备2)之间的直接连接。例如,接口电路系统48可以使用WiFi、蓝牙、Zigbee或任何蜂窝通信协议(包括但不限于全球移动通信系统(GSM)、通用移动通信系统(UMTS)、长期演进(LTE)、高级LTE等)进行操作。在无线连接的情况下,接口电路系统48(以及装置42)可以包括一个或多个合适的天线,用于通过传输介质(例如空中)进行发射/接收。替代地,在无线连接的情况下,接口电路系统48可以包括装置(例如连接器或插头),以使得接口电路系统48能够连接到装置42外部的一个或多个合适的天线,以用于通过传输介质(例如空中)进行发射/接收。接口电路系统50连接到处理单元50。

[0089] 此外,装置42可以包括输出单元52,诸如输出电路系统,以使得装置能够输出信息,特别是输出到反馈单元46。输出单元通常可以被配置为具有与输入单元48相似或相同的功能。

[0090] 反馈单元46可以包括一个或多个组件,其使得装置42的用户能够(例如经由输入单元48)将信息、数据和/或命令输入到装置42中,和/或使得装置42能够(经由输出单元52)将信息或数据输出到装置42的用户。用户界面可以包括任何合适的(多个)输入组件,包括但不限于键盘、小键盘、一个或多个按钮、开关或拨盘、鼠标、触控板、触摸屏、触控笔、相机、麦克风等,并且用户界面可以包括任何合适的(多个)输出组件,包括但不限于显示单元或显示屏、一个或多个灯或灯元件、一个或多个扬声器、振动元件等。

[0091] 在一些实施例中,反馈单元46附接到处理设备2或是处理设备2的一部分。在其他实施例中,反馈单元46是装置42的一部分。在这些实施例中,反馈单元46可以是装置42的用户界面组件的任何输出组件的一部分,或者利用这些输出组件。在另外的实施例中,反馈单元46可以与处理设备2和装置42分开。

[0092] 应该了解,装置42的实际实现可以包括图2中所示的那些组件之外的附加组件。例

如,装置42还可以包括电源,诸如电池,或用于使得装置42能够连接到主电源的组件。

[0093] 本发明具有不同的方面,特别是在实际处理操作期间实际使用装置42的方面和训练装置42在处理操作期间可靠地识别处理设备2发出的闪光声音的方面。

[0094] 在图3中图示了根据第一方面的方法100的实施例的流程图,其可以在处理设备的实际操作期间被使用。在方法100的第一步骤101中,(由输入单元48)获得在处理操作期间(由音频传感器44)感测到的音频信号,即从音频传感器44接收或检索。在第二步骤102中,将所获得的音频信号分割成音频片段,特别是分割成具有设定或预定持续时间(例如1秒或任何其他合适的持续时间)的音频片段。在另一个(可组合的)实施例中,在处理操作期间,输入单元48直接获得所述预定持续时间的音频片段。在第三步骤103中,针对每个音频片段从音频片段中提取一个或多个特性特征(如下文更详细地解释)。然后,在第四步骤104中,针对每个音频片段检测音频片段是否包括闪光声音。这是通过应用已在多个音频片段上训练过的经训练的算法或计算系统来完成的,每个音频片段包括一个或多个特性特征,以检测在音频片段中的闪光声音。在第五步骤105中,每次检测到闪光声音时,闪光声音计数器都会增加,该计数器计数自处理操作开始或者设定的或用户指示的时刻以来检测到的闪光声音的数量。最后,在步骤106中,经由输出单元52将信息输出到反馈单元46。该信息可以包括以下中的一项或多项:闪光声音计数器(即其当前值)、关于闪光声音计数器是否超过闪光声音阈值的指示以及关于处理操作已完成的指示。

[0095] 闪光声音阈值可以例如被预设并反映要被处理的某个身体区域所需的或适当或规定的最大闪光次数。这种设定可以被存储在装置42中或者也可以由这样的用户输入:该用户可以从表格中获得它,例如在用户手册中,或者可以手动地或者在反馈单元或装置42或处理设备2上的用户指导的支持下选择或计算它。例如,闪光声音阈值可能不仅取决于处理区域的尺寸,还取决于皮肤类型、皮肤颜色、先前处理的次数、自上次处理以来的时间等。

[0096] 在图4中图示了根据第二方面的方法200的实施例的流程图,其可以被用于训练。在方法200的第一步骤201中,(例如由装置的声音处理器或外部设备)获得在处理操作期间(由音频传感器44)感测到的或人工创建的音频信号。因此,音频信号是或模拟“真实”音频信号,因为它将在处理期间被记录,并且因此包括多个闪光声音(以及可选的背景噪声)。因此,其可以具有是几秒或几十秒甚至几分钟的长度。此外,音频信号包括注释,该注释可以由用户在收听音频信号时生成,也可以由人工创建音频信号(并同时生成注释)的算法或设备自动生成。音频信号的注释包括指示音频信号内的闪光声音的闪光声音信息,特别是闪光声音的时间信息,诸如闪光声音的开始、结束、最高峰、持续时间等。

[0097] 在第二步骤202中,将(带注释的)音频信号分割成音频片段,特别是分割成具有一组或预定持续时间的音频片段。在第三步骤203中,针对每个音频片段,从音频片段中提取一个或多个特性特征,例如一组MFCC系数(如下文将更详细解释的)。在第四步骤204中,针对每个音频片段从音频信号的注释中导出闪光声音信息,该闪光声音信息指示音频片段是否包括闪光声音,例如,从表示与(带注释的)音频信号相关联或嵌入在其中的注释的文件或元数据中读取闪光声音信息。在第五步骤205中,针对每个音频片段,将所提取的一个或多个特性特征和所导出的闪光声音信息提供给算法或计算系统,以基于一个或多个特性特征来训练算法或计算系统用于检测在音频片段内的闪光声音。

[0098] 通常已知有若干机器学习和深度学习架构用于音频分类。UrbanSound8K就是这样

一种用例,它是一个公共数据集。该数据集包含8732个带标记的城市声音提取,分为十个类别:汽车喇叭、狗吠枪声、儿童玩耍、空调、钻孔、车辆发动机运转、风镐、警报器和街头音乐。使用梅尔频率倒谱系数(MFCC)、特征提取并对其应用零填充,将该输入传递给卷积神经网络(CNN)模型。使用MFCC特征,训练准确度达到98.12%,并且验证准确度达到91.23%。为了提高准确度,应用了特征填充而不是零填充,并使用了其他特征,如零交叉率、音高、RMSE、色度以及MFCC。经过100个时期的训练,训练准确度达到99.21%,并且验证准确度达到97.29%。

[0099] 另一个数据集是ESC-50,它包含50个类别。该数据集包含2000个音频片,每个音频片的持续时间为5秒。ESC-50上的性能为94%的准确度。使用音频频谱图变换器,无卷积架构可用于音频分类。使用音频频谱图变换器可实现94.7%的训练准确度。

[0100] TUT罕见声音事件2017包含每个目标类别的独立声音事件和作为背景的日常声音场景的录音。目标声音事件类别是婴儿哭闹、玻璃破碎、枪声。为解决上述挑战组织了若干事件,在评估数据集上实现了~93%的最佳准确度,在开发数据集上实现了~96%的最佳准确度。解决此问题的方法被划分为4个步骤:特征提取(对数幅度和梅尔频谱图)、用1D convNet转换频谱特征、用RNN-LSTM实现时间依赖性以及用后处理检测音频事件的存在和开始时间。

[0101] Uyen Diep(Sally),《咳嗽事件的分类》,ECE高级顶点项目2017,技术说明披露了作为患者自动咳嗽监测系统的一部分的咳嗽分析,其被划分为两个主要阶段,即事件检测和事件分类。在该方法中,声音事件被分解为32毫秒长度或16毫秒长度的重叠时间帧。计算每个帧的特征并将其馈送到分类器中以确定该帧是否具有咳嗽的特性。然后将结果组合起来对事件做出总体决策。MFCC可以被使用于该方法中。

[0102] 基于信号处理的已知音频分类方法主要是使用深度神经网络来进行多类分类。主要被分类的声音持续几秒钟并且本质上是连续的。相比之下,在当前光脱毛场景中的闪光检测中,闪光通常持续几毫秒并且是非周期性的。在一个实施例,用于构建二元分类模型以检测是否存在闪光的经典机器学习方法被用于闪光声音检测和训练。与深度学习相比,使用这种机器学习(ML)方法的意义在于优化了模型的大小,从而可以轻松在应用程序中提供对模型的部署。

[0103] 在进行脱毛处理时,例如在家中,存在多个挑战,包括高质量的脱毛处理、正确的推荐和无摩擦的体验。特别需要检测处理设备的声音模式。声音模式信号包括在使用处理设备期间发出的闪光声音,但是可能存在潜在的背景噪声,诸如电视声音、收音机、婴儿哭闹、交通噪声、人们的交流、开门或关门等。因此,需要准确地识别音频信号内的预期声音模式。本发明采用信号处理和经训练的算法或计算系统,特别是机器学习(ML),以及一组适当的特性特征(例如12-13个所选MFCC特征),以便可靠且准确地检测闪光声音,而不受任何背景噪声的影响。经训练的算法或计算系统(例如ML模型(诸如XGBoost))在训练过程中学习区分背景噪声(非闪光)和闪光声音。因此,注释有助于模型理解包括闪光和非闪光音频片段的模式。

[0104] 因此,本发明解决了许多挑战,包括以下中的一项或多项:

[0105] • 能够基于从50厘米到1米的距离所捕获的音频准确地检测具有不同幅度的声音

[0106] • 能够抵御各种背景噪声,诸如电视、开门或关门、自来水、收音机等

- [0107] • 自动检测准确的闪光次数(在处理期间被使用)
- [0108] • 减少注释工作量的方法
- [0109] • 隔离与闪光声音重叠的各种背景噪声(诸如电视、滴水等)
- [0110] • 准确检测以处理信号并仅检测闪光声音
- [0111] • 能够具有轻量级且低延迟的架构,用于在传统应用和操作系统中集成人工智能(AI)算法

[0112] 与Uyen Diep在上述文献中公开的咳嗽检测方法等现有技术方法相反,本发明使得能够在存在背景噪声的情况下可靠地检测由处理设备发出的(人工)闪光声音,因为闪光声音特性与人类咳嗽的特性非常不同。由本发明所解决的问题可以被视为检测在处理期间的闪光声音的准确数量。闪光声音本质上不是连续的,并且持续几分之一秒,即闪光声音本质上是非周期性和瞬态的(与非周期性的连续声音和周期性的(简单或复杂)声音相反)。在处理受试者的不同身体部位(如腿部和腋窝)期间所获取的原始音频信号具有不同的持续时间。例如,腋窝音频长度从30秒到105秒不等,而腿部音频长度从30秒到180秒不等。音频信号包括最小间隔为例如0.8秒的闪光声音。音频信号还包括背景噪声。

[0113] 因此,本发明的处理方法不同于现有技术。例如,虽然可以将咳嗽等声音事件分组为重叠帧,但是为了对确切的闪光次数进行计数,本方法在预处理期间将音频信号划分为非重叠音频片段。非重叠音频片段对于避免对闪光的重复计数特别有用。MFCC系统通常仅使用8-13个倒谱系数。例如Uyen Diep在公开的现有技术中通常使用的第零系数表示输入信号的平均对数能量,并且仅携带很少的说话者特定信息。它不被用于本发明的优选实施例中,因为它将导致检测背景噪声。因此,本发明提供了检测由处理设备所产生的低幅度、短持续时间脉冲与变化的背景噪声的能力。MFCC只是一个特征提取步骤,这些特征的正确组合和适当的算法/计算系统(例如适当的ML模型)有助于实现预期效果并解决所解决的问题,特别是提供低延迟、低存储器占用和高准确度,因为它能够由于适当的特征选择而将细微的闪光声音与变化的背景噪声区分开来。

[0114] 在示例性实现中,数据(音频信号)由与处理设备处于近距离(例如在0.5到2.5米的范围内)的智能手机的麦克风收集。音频信号包括闪光声音以及任何现有的背景噪声。

[0115] 在训练期间,可以在出现闪光声音的原始音频信号上进行注释,即,向音频信号添加诸如注释或元数据之类的指示,其指示音频信号中出现闪光声音时的时间时刻。通过将原始音频信号或被注释的信号量化为设定持续时间的片段、例如一秒组块并将片段标记为滴答声(即闪光声音)和非滴答声(非闪光声音),可以执行对音频信号的预处理。此外,可以执行背景噪声过滤,并且可以从音频信号中提取时间和频域特征,诸如MFCC、梅尔频谱图等,从而创建音频特征阵列。在一个实施例中,使用特征阵列来构建ML模型,该模型具有近乎实时地检测闪光声音的能力。ML模型可以例如作为程序文件而被部署在应用程序(应用或软件程序)中,这使得模型能够在期望设备、诸如智能手机上运行。

[0116] 虽然通常可以使用任何算法或计算系统,并且通常可以使用任何类型的学习系统、机器学习、分类模型、二元分类模型、神经网络或卷积神经网络作为经训练的算法或计算系统,但是在实施例中,构建XGBoost二元分类模型来检测音频信号中是否存在闪光声音。如何构建这样的XGBoost分类器在本领域中是众所周知的。可以使用以下超参数构建示例性XGBoost分类器:learning_rate=0.02、n_estimators=600、objective='binary:

logistic’。也可以使用其他参数和其他分类器或ML模型。在这种上下文中,应当注意,通常,经训练的算法或计算系统的特定实现或架构或布局并不是必不可少的,而是例如可以使用具有已知或常规架构的标准学习系统或神经网络来执行检测闪光声音的任务。

[0117] 对于训练(如图4中所图示),使用音频信号的注释。下面将描述如何创建此类注释的实施例。在上述实施例中,通过智能手机的麦克风收集音频数据,例如实时地或以临时存储的方式来收集,优选地与光脱毛设备的距离为0.5m至2.5m。由ML模型提取闪光声音。所提取的数据可以包括音频文件中的每个闪光声音的开始和结束时间。通常,可以使用任何可以将注释添加到音频文件的注释工具来实现此目的。在示例性实现中,使用申请人的注释工具“Barista”。

[0118] 在实施例中,打开原始音频文件,可视化峰值并收听音频(例如仅一次)。在图5中示出了音频文件部分的可视化,其中指示了闪光声音(“Flash Tick”)和背景噪声(“开门/关门或任何其他噪声”)。然后开始播放音频文件,在可以看到或听到闪光声音的地方暂停。然后可以在该区域上放大。图6示出了这种可视化,其中对两个闪光声音进行了一秒放大,其中指示了开始时间和结束时间。然后,可以选择要被注释的区域。可以播放和收听所选区域以进行交叉验证。可以重复此过程直到音频文件结束。

[0119] 在实施例中,如图7中示意性地图示的,可以通过将音频文件量化为具有一定长度(例如,在0.1到5秒的范围内,例如0.5秒或1秒)的音频片段(组块)303来执行原始音频文件301和/或带注释的音频文件302(带有闪光声音的指示)的音频预处理300。然后可以将每个音频片段标记为闪光声音或非闪光声音。此外,可以进行基本的音频操作304,诸如重复预处理多次、改变音频文件的音频信号的音高等。这样做是因为原始音频信号可能并不总是具有足够或最佳的质量,例如可能被背景噪声破坏。

[0120] 在实施例中,如图8中示意性地图示的,执行特征提取400以分析音频数据。音频数据通常可以被视为表示时间、幅度和频率的3D信号。由于3D格式不能(或不容易)由AI模型直接处理,因此可以使用特征提取方法。在将音频信号传递给XGBoost分类模型之前,将信号转换为阵列格式,其然后准备好由AI模型进行进一步处理和分析。通常,在第一步骤401中执行预加重、帧阻塞和窗口化,然后在第二步骤402中应用FFT。随后,应用滤波器组403(例如梅尔标度滤波器组)、对数404以及滤波器组405以获得期望特征,例如MFCC特征406。

[0121] MFCC指的是梅尔标度,它是与音高相关的感知标度。从MFCC特征中获得数个系数值,这些系数值描述音频片段(即一段声音)的特性。在实施例中,在第一步骤401中,给定的音频信号被划分成短帧(音频片段)。执行此步骤是因为信号中的频率随时间变化;因此,计算整个音频信号的傅立叶变换会导致音频信号的频率轮廓随时间丢失。此外,进行窗口化(例如,使用汉明窗口)以交叉检查数据是否无限并减少频谱泄漏。

[0122] 在第二步骤402中,计算离散傅立叶变换以得到傅立叶变换的简化版本并得到实值系数(例如,每帧总共39个系数)。在第三步骤403中,应用梅尔间隔滤波器组,包括:将最低和最高频率转换为梅尔,创建等距点作为应使用的梅尔频带数,将点转换回赫兹,将这些点舍入到最近的频率箱(因为存在离散信号)并创建三角滤波器。由此获得MFCC特征的39个系数。

[0123] 通常,每个音频信号都包括多个特征。然而,需要提取与待解决的问题相关的特性。利用这些音频特征,可以建立和训练智能音频分类模型。

[0124] 在本发明的实施例中,使用梅尔频率倒谱系数(MFCC)作为特性。信号的MFCC是一小组特征(通常约为10-20个),其简明地描述了频谱包络的整体形状。MFCC对人声的特性进行建模。通常每帧有大约39个系数。在实施例中,仅使用前几个系数,例如前12-13个系数(最好是前13个系数),不包括表示输入信号的平均对数能量的第零个分量,因为前几个系数编码了大部分信息(例如频谱包络,共振峰等)。MFCC特征在音乐和语音处理中效果很好,但是缺乏对噪声的鲁棒性。

[0125] 基于下表中所指示的类别,存在多种音频分类:

音频分类	类型	描述
[0126] 抽象级别	高	乐器、主和弦、旋律、节奏
	中等	音高和最佳相关描述符,诸如音符偏移、MFCC
	低	能量、频谱质心、频谱通量、零交叉率
时间范围	瞬时 (~50 ms)	为我们提供有关音频信号的瞬时信息的特征。这考虑非常短的音频信号组块,其通常保持在
		50 到 100 ms 左右
	片段	在 10 或 15 秒的片段范围内计算出的音频特征。
	全局	提供有关整个音频文件的汇总信息的特征。
音乐方面		根据音乐方面进行分类的音频特征,它们可以一窥节拍、音色、音高、和声等。
[0127] 信号域	时域	从波形、x 轴时间、y 轴幅度(零交叉率、RMSE 能量、幅度包络)中提取
	频域	声音以频率为特征。时域没有任何特征。其他特征出自频域(频谱质心、频谱通量、频带能量比)
	时频域	频谱图、梅尔频谱图、常数 Q 变换 频谱图-从时域开始并应用短时傅里叶变换
ML 方法	传统 ML	零交叉率、RMSE 能量、幅度包络、频谱质心、频谱通量、频带能量比、频谱滚降、频谱扩展
	深度学习	频谱图、原始音频

[0128] 音频特征分类可以基于以下域:

[0129]	信号域	时域	零交叉率	$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{\mathbb{R}_{<0}}(s_t s_{t-1})$
			RMSE 能量	$RMS(t) = \sqrt{\frac{1}{T} \sum_{i=1}^T w_i(t) x_i^2(t)}$
			幅度包络	$AE_t = \max_{k=t \cdot K}^{(t+1) \cdot K-1} s(k)$
	频域	频谱质心	$SC_t = \frac{\sum_{n=1}^N m_t(n) \cdot n}{\sum_{n=1}^N m_t(n)}$	
		频带能量比	$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^N m_t(n)^2}$	
	时频	频谱图	频谱图是信号的频率当它随时间不同时的频谱的视觉表示	
[0130]	域	常数 Q 变换	$\delta f_k = 2^{1/n} \cdot \delta f_{k-1} = \left(2^{1/n}\right)^k \cdot \delta f_{\min}$	

[0131] 在预处理步骤中创建音频片段之后,所提取的MFCC特征将被传递给模型。

[0132] 在下文中,将更详细地描述AI模型训练的实施例。模型的输入可以是矩阵(或向量),例如13列作为MFCC系数,并且行数等于音频文件的长度(例如,对于每个长度为1秒的音频片段,有一行)。

[0133] 在实施例中,每个长度为1秒的音频片段使用以下13个MFCC系数(即,这13个MFCC系数表示矩阵的13列):

[0134] mfcc_second_derivative_std_1

[0135] mfcc_second_derivative_std_2

[0136] mfcc_second_derivative_std_3

[0137] mfcc_second_derivative_std_4

[0138] mfcc_second_derivative_std_5

[0139] mfcc_second_derivative_std_6

[0140] mfcc_second_derivative_std_7

[0141] mfcc_second_derivative_std_8

[0142] mfcc_second_derivative_std_9

[0143] mfcc_second_derivative_std_10

[0144] mfcc_second_derivative_std_11

[0145] mfcc_second_derivative_std_12

[0146] mfcc_second_derivative_std_13。

[0147] 应当注意,这些是本领域技术人员通常知道的常用名称;这些特征是基于其识别闪光声音的预测能力来选择和排序的。此处,系数“mfcc_second_derivative_std_1”是MFCC二阶导数标准差(也称为增量-增量(delta-delta)系数)。为了得到二阶导数,提取MFCC特征,然后计算这些MFCC特征的二阶导数。

[0159] -关于闪光声音计数器是否超过闪光声音阈值的指示,以及

[0160] -关于处理操作已完成的指示。

[0161] 在图11中图示了不同阶段的对应用户界面。屏幕601(表示步骤501中的第一屏幕)示出了如何有效进行处理的用户指南,并将询问用户期望处理区域的大小,例如腿长。一旦用户输入腿长(在步骤502中),就会示出屏幕602;将显示一些理想的闪光次数以进行适当的处理。在单击继续之后,用户将进入屏幕603,其中将要求单击开始按钮以开始处理。开始处理将在屏幕604(表示步骤503中的第二屏幕)上询问麦克风访问(或更一般地,指示连接性)。一旦被授权,将开始检测闪光声音,如屏幕605上所显示的,并且整个处理期间使用的总闪光次数将递增,如屏幕5(表示步骤504中的第三屏幕)上所显示的。在另一个实施例中,可以显示用于处理期望处理区域所需的剩余闪光次数,或者可以同时显示所应用的总闪光次数和剩余闪光次数。

[0162] 在图11中所示的屏幕应被理解为示例。每个屏幕上也可以显示其他信息,并且可以有其他设计和数量的屏幕。

[0163] 本发明提供了自动计算每次处理的闪光次数的优势,这将有助于用户计算针对后续处理的次数。处理通常可能需要大约30-40次闪光。目前,用户必须手动跟踪闪光计数,这会导致糟糕的用户体验。因此,对设备的适当推荐将消除设备使用中的摩擦(即,自动计数闪光,而无需用户在处理期间手动计数)并带来更好的用户体验和满意度。所提出的方法已经用不同的背景噪声和能够识别闪光声音的模型进行了测试。每当处理未满足预期水平时,可以检查是否遵循了指导或建议。

[0164] 虽然本发明已在附图和前述说明中进行了详细说明和描述,但是这种说明和描述应被视为是说明性或示例性的,而不是限制性的;本发明不限于所公开的实施例。本领域技术人员在实践所要求保护的发明时,可以通过研究附图、公开内容和附加的权利要求来理解和实现对所公开的实施例的其他变化。

[0165] 在权利要求中,词语“包括”不排除其他元件或步骤,并且不定冠词“一”或“一个”不排除复数。单个元件或其他单元可以实现权利要求中所记述的若干项的功能。某些措施在相互不同的从属权利要求中所记述的事实本身并不指示这些措施的组合不能被充分利用。

[0166] 计算机程序可以被存储/分发给合适的非暂时性介质上,诸如与其他硬件一起提供或作为其他硬件的一部分来提供的光学存储介质或固态介质,但是也可以以其他形式来分发,诸如基于互联网或其他有线或无线电信系统来分发。

[0167] 权利要求中的任何参考符号不应被解释为限制范围。

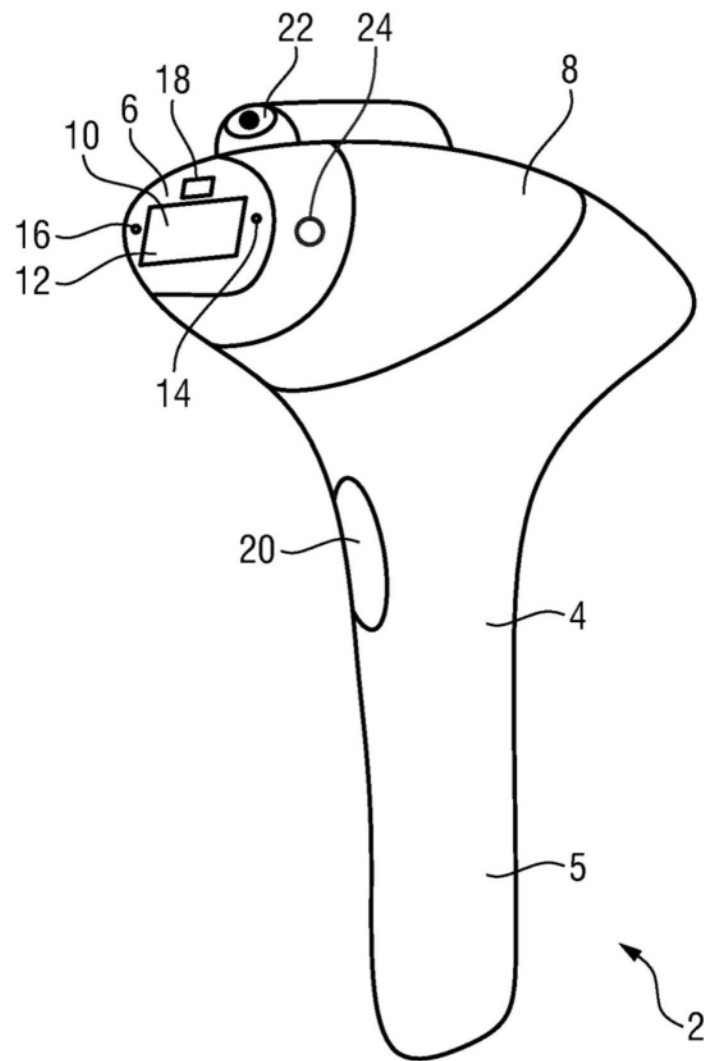


图1

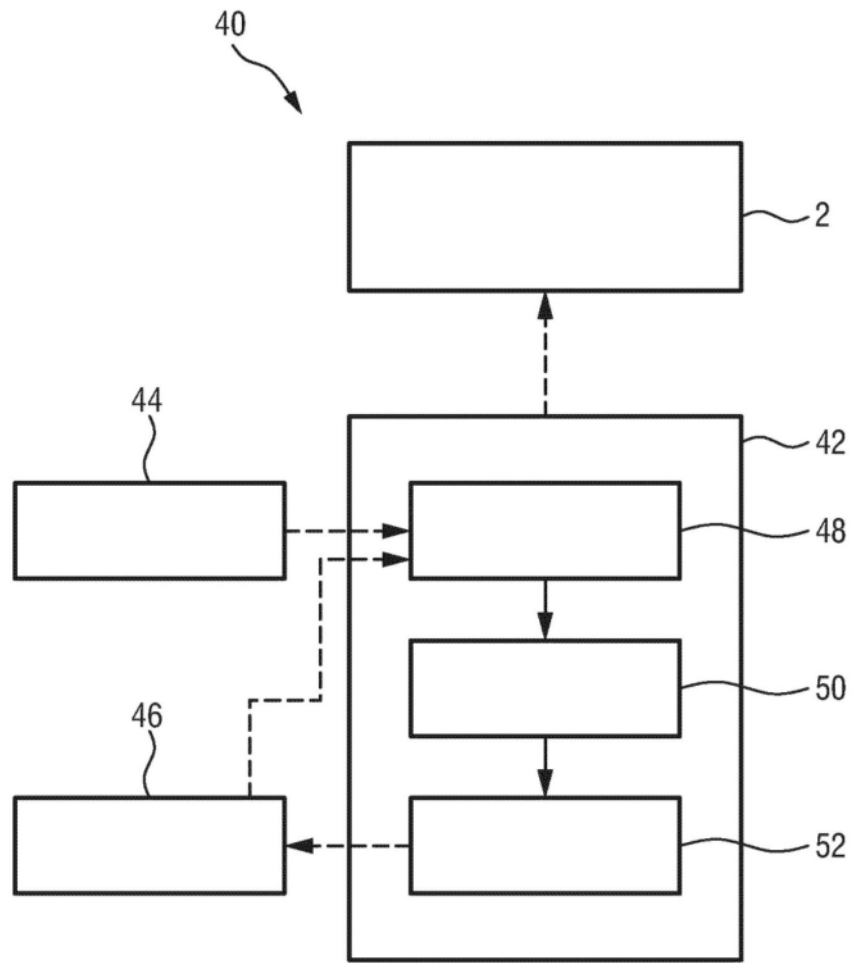


图2

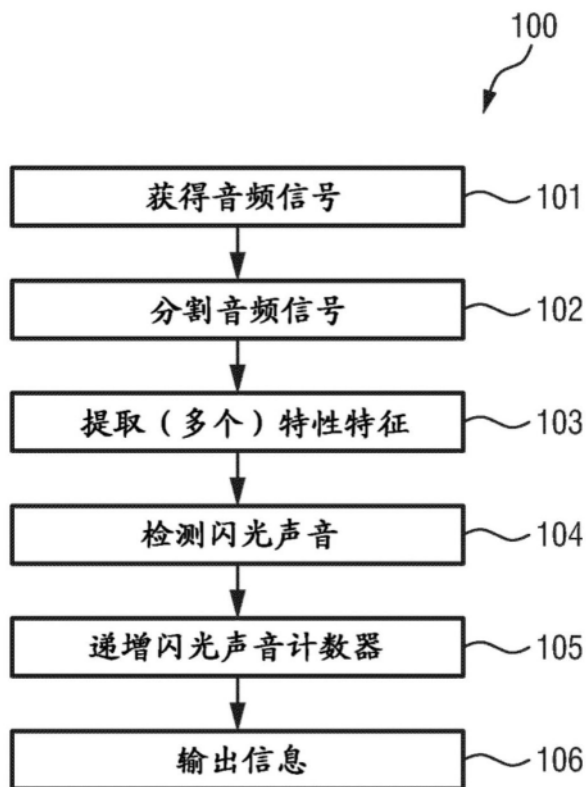


图3

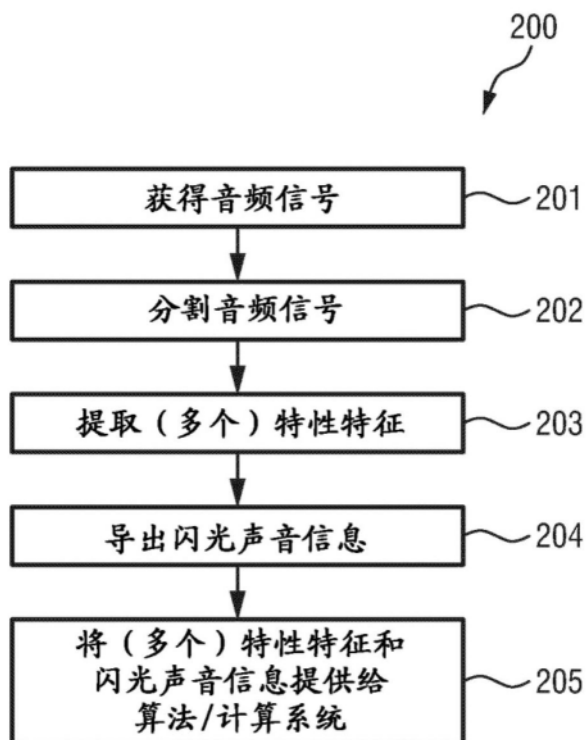


图4

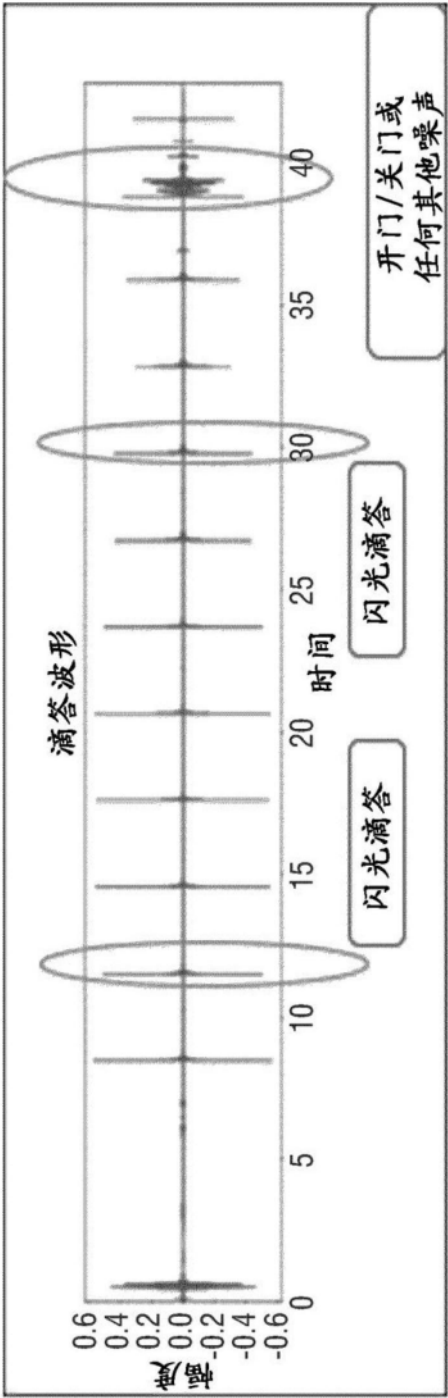


图5

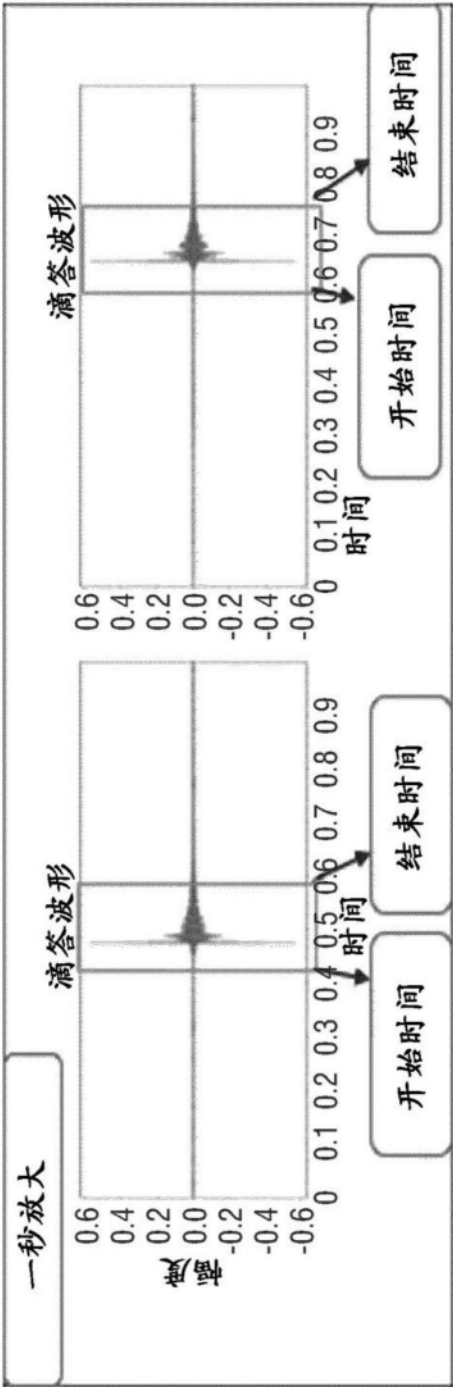


图6

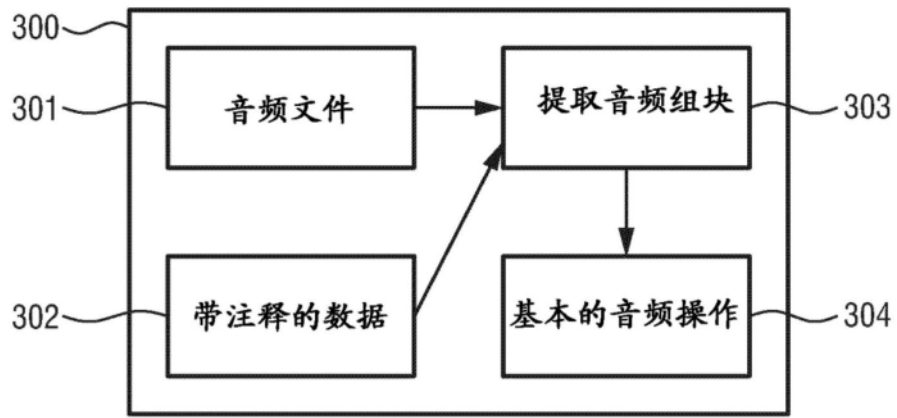


图7

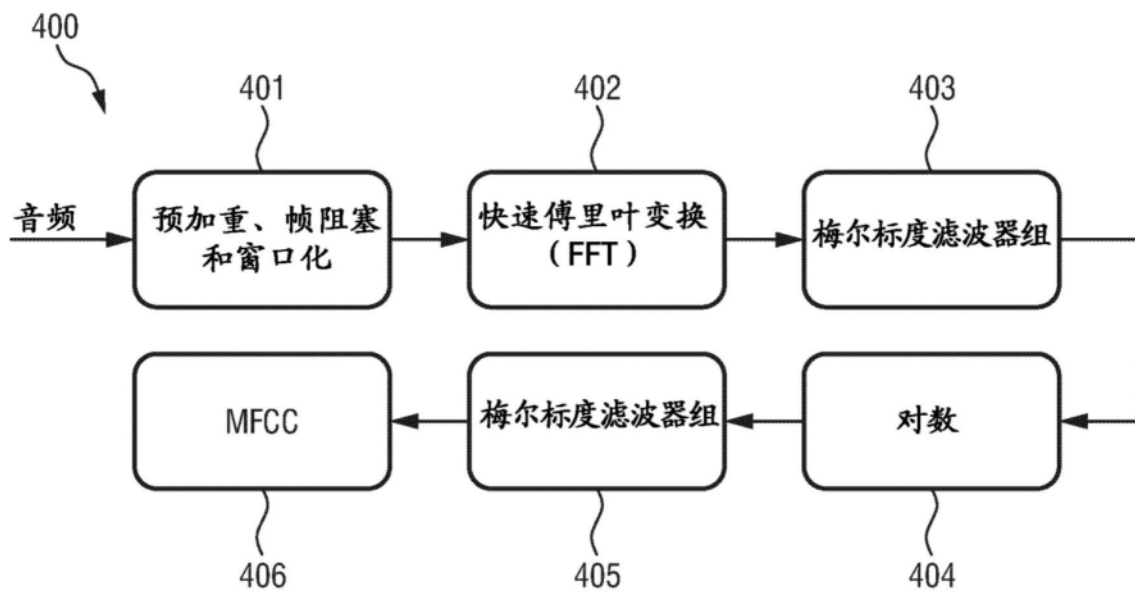


图8

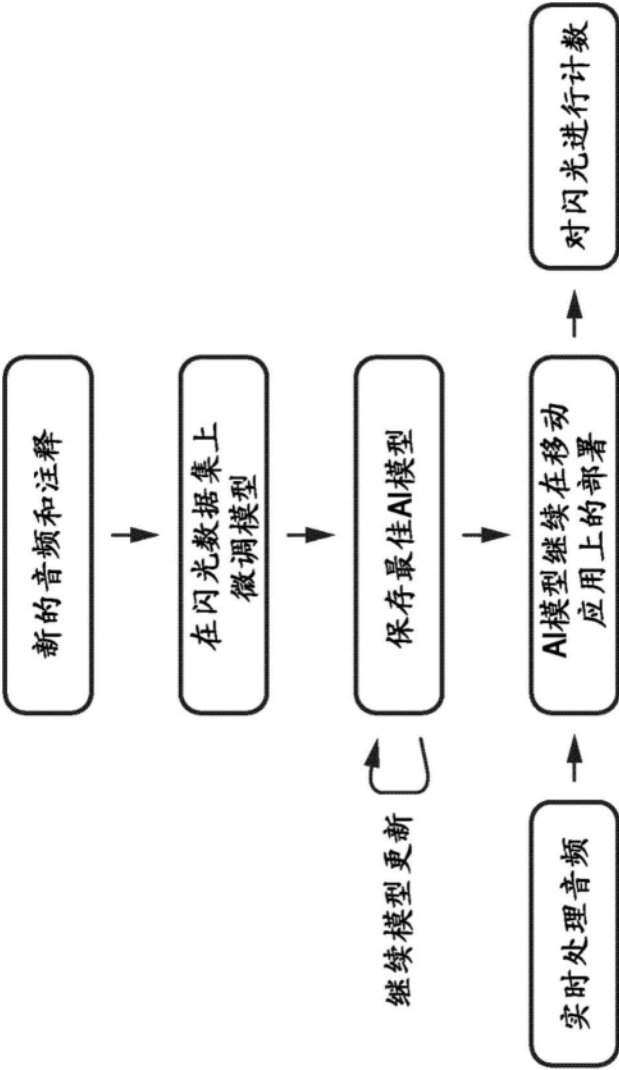


图9

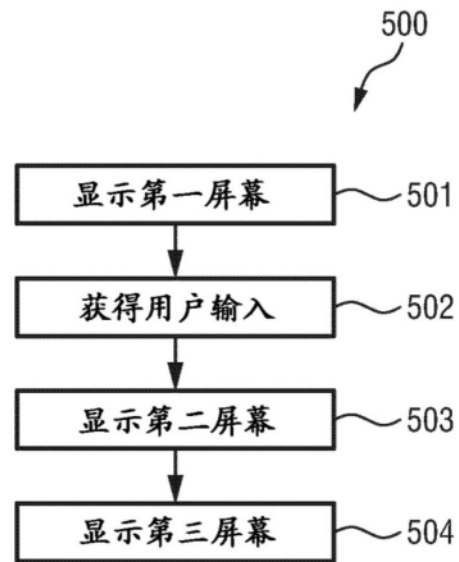


图10

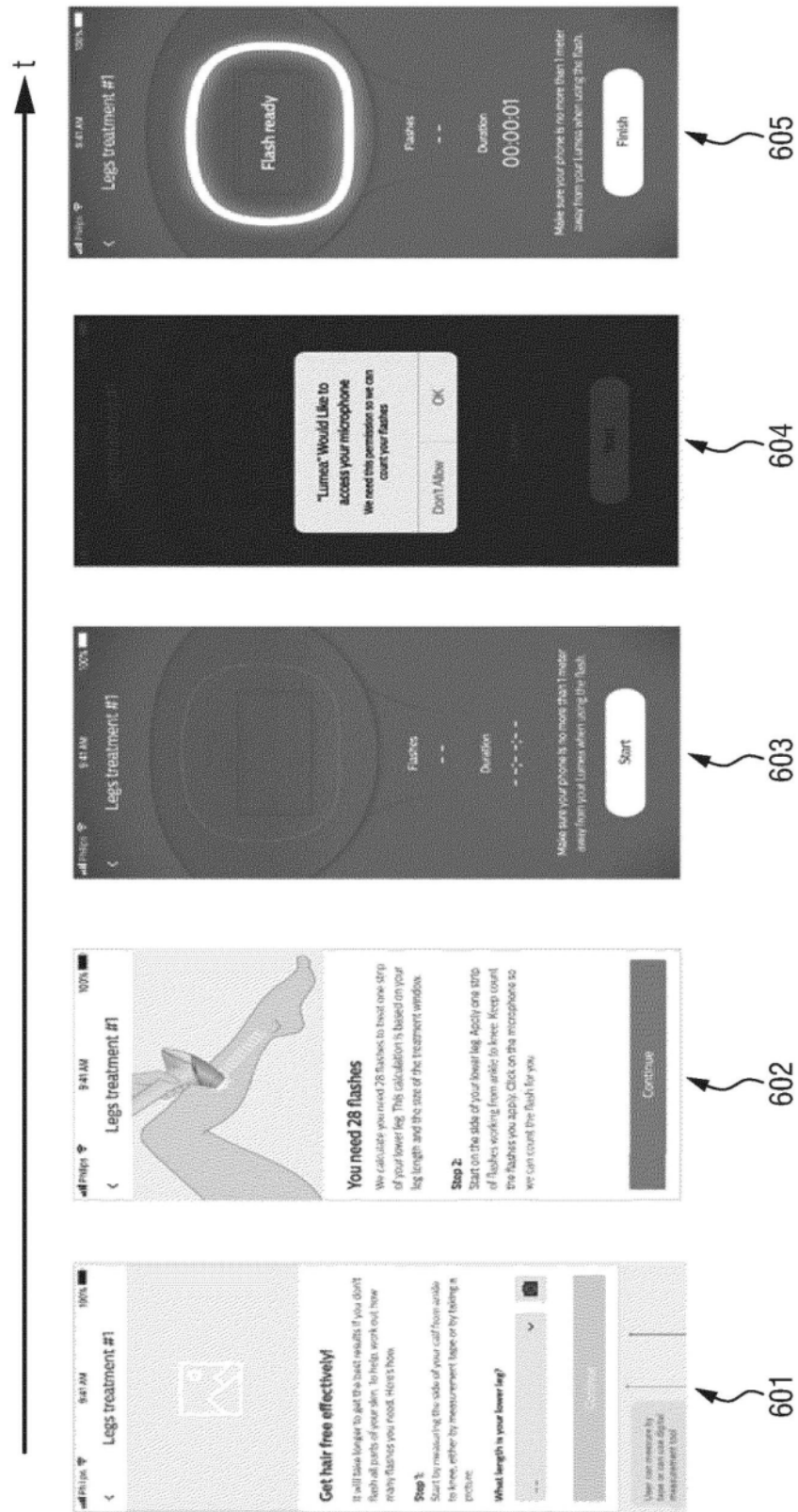


图11