(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 110353685 B
(45)授权公告日 2022.03.04

(21)申请号 201910202125.5
(22)申请日 2013.03.28
(65)同一申请的已公布的文献号
申请公布号 CN 110353685 A
(43)申请公布日 2019.10.22
(30)优先权数据
2012901255 2012.03.29 AU
(62)分案原申请数据
201380028268.X 2013.03.28
(73)专利权人 昆士兰大学
地址 澳大利亚昆士兰州
(72)发明人 U·R·阿贝拉特纳 V·斯旺加 Y·A·阿姆鲁罗
(74)专利代理机构 永新专利商标代理有限公司 72002
代理人 刘瑜 王英
(51)Int.Cl.
A61B 5/08(2006.01)
A61B 7/04(2006.01)
(56)对比文件
CN 102325494 A,2012.01.18
US 2011208521 A1,2011.08.25
US 2012071777 A1,2012.03.22
WO 2010066008 A1,2010.06.17
WO 2007001431 A2,2007.01.04
审查员 李易陆

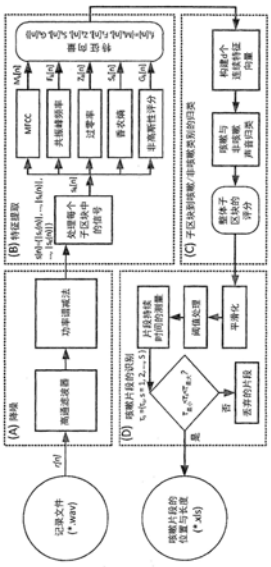
权利要求书2页 说明书49页 附图24页

(54)发明名称

用于处理患者声音的方法与装置

(57)摘要

一种操作计算设备以处理患者声音的方法，所述方法包括以下步骤：从所述患者声音的片段提取特征；以及基于所提取的特征和预定准则，将所述片段归类为咳嗽或非咳嗽声音；以及基于所述患者声音中被归类为咳嗽声音的片段，在所述计算设备的控制下，将对疾病相关状态的诊断呈现在显示器上。



CN 110353685 B

1.一种计算设备,包括与电子存储器通信的至少一个电子处理器,所述电子存储器包含指令,所述指令用于所述处理器以执行一种处理用于诊断呼吸疾病的患者声音的方法,所述方法包括:

将声音接收设备配置为接收来自患者的所述患者声音并且形成对应的数字化声音信号;

计算针对所述数字化声音信号的多个子区块中的每个的特征向量,所述特征向量包括以下中的两个或更多个:梅尔频率倒谱系数(MFCC)、熵特征、过零率、以及非高斯性;

操作训练的咳嗽检测模式归类器以响应所述特征向量,从而产生咳嗽检测模式归类器输出信号;

对所述输出信号进行滤波以产生平滑输出信号;

对所述平滑输出信号进行阈值判断以识别所述数字化声音信号中的候选咳嗽片段;

基于候选咳嗽片段的持续时间长于最小咳嗽长度并且短于最大咳嗽长度而将候选咳嗽片段归类为咳嗽片段;

形成所述咳嗽片段中的每个的多个子片段并且计算针对所述子片段中的每个的特征向量,这些特征向量包括以下中的两个或更多个:梅尔频率倒谱系数(MFCC)、熵特征、过零率、以及非高斯性;

将所述多个子片段的所述特征向量应用到训练的患病咳嗽模式归类器以将所述咳嗽片段中的每个针对特定呼吸疾病归类为患病和非患病中的一种;

计算患病咳嗽指数,所述患病咳嗽指数指示被归类为患病的咳嗽片段相对于所有所述咳嗽片段的比例;

基于所述患病咳嗽指数,将所述患者认为遭受所述特定呼吸疾病;并且

在所述计算设备的控制下,在显示器上呈现对所述特定呼吸疾病的诊断。

2.根据权利要求1所述的计算设备,其中,形成所述咳嗽片段的多个子片段包括形成三个子片段。

3.根据权利要求1所述的计算设备,其中,所述计算设备包括被编程为执行所述方法的移动电话。

4.根据权利要求1所述的计算设备,其中,计算针对所述数字化声音信号的多个子区块中的每个的特征向量包括提取以下特征的组中的任一组:

梅尔频率倒谱系数和共振峰频率和过零率和香农熵和非高斯性特征;或

过零率和香农熵和非高斯性特征;或

梅尔频率倒谱系数和过零率和共振峰频率和非高斯性特征。

5.根据权利要求1所述的计算设备,其中,所述训练的咳嗽检测模式归类器包括人工神经网络。

6.根据权利要求5所述的计算设备,其中,所述人工神经网络包括时延神经网络(TDNN)。

7.根据权利要求6所述的计算设备,其中,所述时延神经网络具有输出层和输入层之间的隐藏层,其中,所述隐藏层中具有10至50个神经元。

8.根据权利要求1所述的计算设备,还包括向所述咳嗽片段应用更高阶的统计分析以对所述咳嗽片段进行进一步归类。

9.根据权利要求8所述的计算设备，其中，所述更高阶的统计分析包括计算双频谱。

10.根据权利要求1所述的计算设备，其中，形成所述咳嗽片段中的每个的多个子片段并且计算针对所述子片段中的每个的特征向量还包括：

确定以下中的一个或多个：针对所述子片段中的每个的双频谱指数(BSG)、共振峰频率(FF)、对数能量(Log E)、以及峰态(Kurt)特征。

11.根据权利要求10所述的计算设备，其中，所述训练的患病咳嗽模式归类器包括以下中的任一个：逻辑回归模型；人工神经网络；贝叶斯归类器；隐马尔科夫模型；以及支持向量机。

12.根据权利要求11所述的计算设备，

其中，所述训练的患病咳嗽模式归类器已经利用训练集被训练，所述训练集包括从遭受以下不适中的一种或多种的患者记录的非肺炎声音：哮喘、支气管炎、鼻咽炎、气喘、扁桃体咽炎、心脏病、喉软骨软化病、疟疾以及异物吸入。

13.根据权利要求1所述的计算设备，其中，所述训练的患病咳嗽模式归类器包括以下中的任一个：逻辑回归模型；人工神经网络；贝叶斯归类器；隐马尔科夫模型；以及支持向量机。

14.根据权利要求13所述的计算设备，

其中，所述训练的患病咳嗽模式归类器已经利用训练集被训练，所述训练集包括从遭受以下不适中的一种或多种的患者记录的非肺炎声音：哮喘、支气管炎、鼻咽炎、气喘、扁桃体咽炎、心脏病、喉软骨软化病、疟疾以及异物吸入。

15.根据权利要求1所述的计算设备，其中，所述特定呼吸疾病包括肺炎。

16.根据权利要求15所述的计算设备，其中，所述患病咳嗽指数是肺炎咳嗽指数。

17.根据权利要求1所述的计算设备，其中，在所述显示器上呈现对患病相关状态的诊断包括呈现对肺炎的诊断。

18.根据权利要求1所述的计算设备，其中，所述患者声音是利用一个或多个麦克风采集的，所述一个或多个麦克风从与所述患者的物理接触移除。

用于处理患者声音的方法与装置

[0001] 本申请是申请日为2013年3月28日、发明名称为“用于处理患者声音的方法与装置”的专利申请201380028268.X的分案申请。

技术领域

[0002] 在第一方面，本发明涉及用于检测患者的咳嗽声音的自动方法与装置。在第二方面，本发明涉及用于根据诸如咳嗽声音的患者声音来诊断疾病状态的方法与装置。

背景技术

[0003] 对现有技术的方法、装置或文件的任意引用均不构成它们形成的或形成公知常识的部分的任意证据或承认。

[0004] 咳嗽是身体的一种防御机制，以清除呼吸道中意外吸入的或因感染而在内部产生的异物[1]。它是一些呼吸疾病，例如哮喘和顿咳(百日咳)以及肺炎的常见症状，肺炎是5岁以下儿童死亡的主要原因。据估计[2]，肺炎每年引起该组中超过160万人死亡。世界卫生组织(WHO)也报道，在那些国家，百日咳已成为主要儿童发病率之一，每年估计有5000万病例和300000人死亡[4]。

[0005] 尽管咳嗽在呼吸疾病中是常见的，并且被认为是重要的临床症状，但并没有评估咳嗽的黄金标准。在典型的会诊中，医师可以听取自然或自发咳嗽的几次发作，以获得诸如咳嗽的“湿润”的定性信息。这种定性信息在对呼吸疾病的诊断以及处置中极其有用。然而，人工分析遭受操作者偏差的问题，并且产生主观结果。

[0006] 在会诊期间，医师也可以寻求关于咳嗽的定量信息，例如咳嗽事件在给定时间间隔上的发生频率。该信息能够被用于确定咳嗽的性质(例如急性、慢性)和严重度，以及用于监测处置效率。然而，为了获得该信息，医师严重依赖患者或他们的照护者的主观报告。存在对于一种能够尤其在儿童疾病中计数咳嗽的次数的自动设备的需要。更重要地，需要能够从长的儿科记录中自动提取咳嗽事件的技术，以便帮助对诸如肺炎、百日咳及哮喘的疾病的诊断。

[0007] 已经采取一些途径来开发自动咳嗽计数系统(例如Hull自动咳嗽计数器(HACC)[5]、莱斯特咳嗽监测器(LCM)[6]、LifeShirt[7]、VitaloJA[8]，以及PulmoTrack[9])。这些设备的性能各异。HACC声称有(80％、96％)的灵敏度和特异性[5]。该数字针对LifeShirt、Pulmotrack、LCM以及Vitalojak分别为(78％、99％)、(94％、96％)、(85.7％、99.9％)，以及(97.5％、97.7％)[6、10-13]。它们依赖声音强度相关的技术，使得它们易受记录条件及所用的特定器械的变化的影响。据发明人所知，这些商业化设备中还没有一个在儿科群体上测试过。

[0008] 在儿童，尤其是较年幼的那些上的咳嗽记录带来一些额外的挑战。较年幼的儿童不能根据要求产生自发咳嗽。靶向于儿科群体的任意方法均应能够使用在感兴趣时期上记录的自然咳嗽。在儿科记录中，发现有大量与咳嗽声音混杂的哭叫、发声和咕哝。因此，针对成年人开发的技术在儿童上的使用不太可能是最优的。来自儿童的咳嗽记录中的另一个问

题是咳嗽声音强度变化。诸如严重儿科肺炎的疾病能急剧降低咳嗽声音的幅度。即使在健康人中，咳嗽声音也能有大的动态范围，覆盖大声咳嗽至几乎听不到。该情况可能使基于强度的技术针对现场使用不可靠。性能也将取决于使用的具体声音捕获器械、校准状态以及测量协议。

[0009] 现有的商品化咳嗽计数设备(例如LifeShirt、Vitalojak和Pulmotrack)采用接触传感器。尽管接触传感器的使用可以有一些优点，但它们也带有一些缺陷。介于中间的肌肉组织严重削减使用接触传感器记录的咳嗽的带宽；自由空气系统对此免疫。接触传感器，与非接触(自由空气)麦克风相比，对抗通过空气传播的背景声音是鲁棒的。然而，它们对通过组织和骨骼传导的声音更为脆弱；因传感器移动造成的假摩擦声音也可能是个问题。在传染性疾病中，需要精细的工作以避免通过接触仪器造成患者的交叉感染。此外，在儿科对象中，接触传感器也可能因患者不舒服而难以附着。

[0010] 咳嗽声音承载有关于气道的状态的非常有用的信息。然而，现有的设备使用仅能够检测事件的存在(“咳嗽检测”)却不能自动提取咳嗽事件(“咳嗽分割”)用于进一步分析的方法。因此，它们被限制在对咳嗽的计数。咳嗽分割，除咳嗽检测能力以外，还要求关于每次咳嗽事件的精确开始以及结束的知识。已知咳嗽间的间隙、咳嗽的持续时间以及咳嗽的幅度可以承载与呼吸疾病相关的信息[14]。

[0011] 其症状包括咳嗽的一种疾病是肺炎。肺炎是世界范围内幼童的主要杀手。其相当于每年多余19％的五岁以下儿童死亡数。它是一种贫穷病，并且与营养不良和差的医疗设施强烈相关。因此，儿童肺炎死亡数在一些国家非常高。肺炎也是全世界老年人中的问题。

[0012] 肺炎被定义为肺部的感染，具有炎性细胞和分泌物在肺泡中的积累。肺炎的常见症状包括：咳嗽、呼吸困难、发热、头痛、食欲不振、流鼻涕、以及气喘。在严重肺炎的情况中，小婴儿费力呼吸并且可能遭受抽搐、皮肤苍白、失去知觉、体温过低以及嗜睡。

[0013] 肺炎是一种难以诊断的疾病。目前的诊断方法包括临床检查(例如：生理指标、胸部听诊)、生化检测(例如：痰液分析、氧饱和度)以及医学成像(例如：胸部X射线以及在一些情况中为X射线CT)。

[0014] 目前的诊断方法的问题是什么：

[0015] 胸部X射线(CXR)被认为是用于诊断肺炎的常用参考准则。然而，其并非黄金标准。在疾病早期，或者当疾病涉及在CXR中不容易看到的肺部分时，单独使用CXR难以诊断肺炎。而且，有时CXR结果能够因肺部疤痕或充血性心力衰竭(它们能够在CXR中模仿肺炎)而是误导性的。尽管X射线CT可以提供更好的结果，但在一些国家它们甚至在三级保健医院中都没有广泛使用。痰液测试要求实验室培养并且最少要花2-3天，使得它们对于初步诊断来说太慢。阳性痰液测试并不一定指示肺炎的存在，因为引起肺炎的病原体中许多都天然地存在于健康人的喉咙中。因此，痰液测试主要是用来检查已在患者上启动的特定抗生素的敏感性。临床检查连同经由听诊器的胸部听诊是在临床环境中被用于对肺炎的初步检查的一线途径；X射线在可用时可以被用于证实诊断。

[0016] 上述方法中没有一个可用于在世界上肺炎肆虐的偏远地区大规模部署。它们昂贵、资源密集，并且要求训练有素的医疗专业人员执行它们。

[0017] 为了克服该问题，世界卫生组织(WHO)已开发出一套高度简化的指南[参考3]，以在世界上资源贫乏且偏远的地区诊断儿童肺炎。根据这些，出现呼吸困难或咳嗽的儿童，如

果他们有呼吸急促(快速呼吸),则被诊断出肺炎。快速呼吸被定义为:在小于2个月的婴儿中每分钟60次呼吸或更多,针对2个月至12个月的婴儿每分钟50次呼吸或更多,以及针对年龄在12个月至5岁的儿童每分钟40次呼吸或更多[参考3、4]。胸部凹入、皮肤苍白以及失去直觉可以指示严重肺炎,并且也属于WHO危险指征。该系统在现场更容易实施,并且被设计为具有高诊断灵敏度(约90%患有疾病的患者被找出)。然而,WHO指南遭受差的诊断特异性的问题;大量没有肺炎的患者也被找出认为患有肺炎。已知WHO算法的特异性为约20%。

[0018] 尽管WHO指南已帮助使死亡率降至每年160万儿童死亡数,但该方法仍有几个问题。由于其低特异性[参考6],大量非肺炎的儿童不必要地接受抗生素。这已导致由社区性抗生素耐药性引起的处置失败。在许多肺炎流行地区,诸如疟疾的疾病也是常见的。肺炎和疟疾两者都有发热、快速呼吸和咳嗽的症状,并且针对肺炎的WHO算法能够导致误诊并延误治疗。不需要抗生素的几种其他疾病/病况(例如COPD、哮喘、肺水肿、肺癌等)可能出现与肺炎类似的临床特征。

[0019] 为了改善WHO标准的特异性,Cardoso等人[参考6]提议包括发热的存在,以诊断肺炎。他们示出,增加发热显著改善了诊断特异性(高达50%)。过去的几个其他研究者已评估了WHO标准在儿童肺炎诊断中的准确度。Harari等人[参考7]在185名儿童中研究了包括呼吸急促的几个变量,以确定哪些临床指征最佳地预测肺炎的影像学证据。他们报道,仅以呼吸急促(针对<12个月的小孩呼吸率(RR)50≥次呼吸/分钟,以及如果年龄为1岁或更大则为RR≥40次呼吸/分钟)作为预测指标,在诊断肺炎时灵敏度为73%和特异性为64%。当他们增加胸部凹入到呼吸急促时,以特异性为代价(下降6%)使灵敏度改善了4%。与其他临床症状(例如鼻翼扇动、发热、睡眠欠佳、咳嗽>2天等)类似地,灵敏度和特异性在20%至90%之间变化[参考6-10]。以特异性为代价获得高灵敏度,反之亦然。

[0020] 本发明的第一方面的目的是提供一种改进的用于识别咳嗽声音的方法。

[0021] 此外,本发明的另外一方面的目的是提供一种用于从咳嗽声音诊断特定疾病状态(例如肺炎、哮喘和鼻咽炎)的方法。

发明内容

[0022] 根据本发明的第一方面,提供一种操作计算设备以处理患者声音的方法,所述方法包括以下步骤:

[0023] 从所述患者声音的片段提取特征;以及

[0024] 基于所提取的特征和预定准则,将所述片段归类为咳嗽或非咳嗽声音。

[0025] 所述患者声音将优选地包括所述患者的声音记录。

[0026] 在本发明的优选实施例中,从所述患者声音记录提取特征的所述步骤包括处理所述片段,以确定预定统计学参数。

[0027] 例如,在优选实施例中,所述预定统计学参数包括对应于所述片段的梅尔频率倒谱系数“M_k”

[0028] 所述预定统计学参数也可以包括非高斯性评分“N_k”,指示所述片段与高斯模型的偏差。

[0029] 从所述患者声音记录提取特征的所述步骤也可以包括估计针对所述片段的共振峰频率“F_k”。

[0030] 优选地，提取特征的所述步骤包括确定针对所述片段的过零率“Z_k”，以由此检测与其相关联的周期性。

[0031] 在本发明的优选实施例中，从所述片段提取特征的所述步骤包括确定小波香农熵值“S_k”、计算峰态(v_k)、音高(p_k)、双频谱评分(B_k)、对数能量(E_k)。

[0032] 优选地，所述方法包括基于所述M_k、N_k、F_k、Z_k、S_k、v_k、p_k、B_k和E_k，计算归一化特征向量f_k。

[0033] 优选地，将所述片段归类为咳嗽或非咳嗽声音的所述步骤包括，通过将所提取的特征或从所提取的特征推导的参数(例如所述归一化特征向量)与预录咳嗽声音的数据集的特征进行比较，确定它们是否满足所述预定标准。

[0034] 归类所述片段的所述步骤可以包括针对对应于测试声音片段的测试特征的后续归类，利用预录咳嗽声音的所述数据集的特征训练人工神经网络，以由此将所述测试声音片段归类为非咳嗽或咳嗽类别。或者，也可以使用其他模式归类器，例如逻辑回归归类方法。

[0035] 优选地，所述方法还包括基于所述患者声音记录中被归类为咳嗽声音的片段，诊断疾病相关状态。

[0036] 所述方法可以还包括监测来自若干传感器的患者数据。

[0037] 例如，所述方法可以包括监测来自以下中的一个或多个的患者数据：呼吸检测传感器、温度传感器、移动传感器、EEG传感器、氧水平传感器。

[0038] 所述方法可以包括将咳嗽声音分类为“湿咳”或“非湿咳”声音。

[0039] 当所述方法包括诊断疾病状态时，其将还包括以下步骤：

[0040] 将所述患者声音的咳嗽事件分类为患病或非患病的；

[0041] 计算患病咳嗽指数，所述患病咳嗽指数指示被分类为患病咳嗽事件的比例；以及

[0042] 如果所述患病咳嗽指数符合预定要求，则认为所述患者遭受所述特定疾病状态。

[0043] 所述预定要求可以是所述患病咳嗽指数超过预定阈值。

[0044] 优选地，将所述咳嗽事件分类为患病或非患病的所述步骤包括，根据归类程序处理从所述咳嗽事件提取的特征。例如，所述归类程序可以包括逻辑回归模型的应用。或者，所述归类器可以包括受训人工神经网络的应用。

[0045] 根据本发明的另一方面，提供一种计算设备，其包括与电子存储器通信的至少一个电子处理器，所述电子存储器包含针对所述处理器的指令，以执行前述用于处理患者的声音记录以将所述记录的片段分类为咳嗽或非咳嗽声音的方法。

[0046] 例如，所述计算设备可以包括被编程为执行前述方法的移动电话或类似设备。

[0047] 根据本发明的另外一方面，提供一种承载有形指令的机器可读介质，所述有形指令用于由电子处理器运行，以执行前述方法。

[0048] 根据本发明的另外一方面，提供一种操作计算设备的方法，以处理患者的患者声音，以由此将所述声音转换为指示对所述患者的特定疾病状态的诊断。例如，所述疾病状态可以包括肺炎。

[0049] 优选地，所述患者声音包括所述患者的声音记录。

[0050] 优选地，所述诊断方法包括以下步骤：

[0051] 将所述声音记录的咳嗽事件分类为患病或非患病的；

[0052] 计算患病咳嗽指数，所述患病咳嗽指数指示被分类为患病咳嗽事件的比例；以及

[0053] 如果所述患病咳嗽指数符合预定要求，则认为所述患者遭受所述特定疾病状态。

[0054] 例如，所述预定要求可以为所述患病咳嗽指数超过预定阈值。

[0055] 所述方法可以包括检测来自以下中的一个或多个的患者数据：移动传感器、EEG传感器、氧水平传感器、呼吸检测传感器和温度传感器。

[0056] 优选地，将所述咳嗽事件分类为患病或非患病的所述步骤包括根据归类程序处理从所述咳嗽事件提取的特征。

[0057] 例如，所述归类程序可以包括逻辑回归模型的应用。或者，所述归类器可以包括受训人工神经网络、贝叶斯归类器、基于隐马尔科夫模型(HMM)的归类器、支持向量机(SVM)等等的应用。在一个实施例中，用于所述归类器的训练集包括从遭受诸如以下的不适的患者记录的非肺炎声音：哮喘、支气管炎、鼻咽炎与气喘、扁桃体咽炎(tonsillopharyngitis)、充血性心脏病、喉软骨软化病(laryngomalacia)及异物吸入。

[0058] 本发明的各个实施例涵盖在诸如疗养院的老年照护设施中对同院病人的长期监测(没有接触传感器)。将认识到，肺炎是老年人中的严重问题，并且最好尽早开始处置。需要医师的干预的任意呼吸病况，应由工作人员紧急标出，并且本文描述的方法的应用有助于满足该要求。这样的方法，根据本发明的一方面的实施例，可以被应用于慢性病患者(患有免疫系统或呼吸系统疾病)，例如艾滋患者。

[0059] 本发明的各方面的实施例涵盖在医院环境内外对等待看特定专科的呼吸患者的分流。

[0060] 根据本发明的另一方面，提供一种计算设备，其包括与电子存储器通信的至少一个电子处理器，所述电子存储器包含针对所述处理器的指令，以执行前述方法，以处理患者的声音记录，以将所述记录的片段分类为咳嗽或非咳嗽声音。

[0061] 根据本发明的另外一方面，提供一种承载有形执行的计算机可读介质，所述有形指令用于由电子处理器运行，以执行前述方法。

[0062] 根据本发明的另一方面，提供一种操作计算设备以处理为数字电子格式的患者声音的方法，所述方法包括以下步骤：操作所述计算设备的至少一个处理器，以从所述患者声音的片段提取特征的步骤，

[0063] 基于所提取的特征和预定准则，将所述片段归类为咳嗽或非咳嗽声音，以及

[0064] 将片段有形地标记为所归类的咳嗽或非咳嗽，以由所述计算设备进一步处理。

附图说明

[0065] 可以从以下具体实施方式领悟本发明的优选特征、实施例和变型，所述具体实施方式为本领域技术人员提供用于执行本发明的足够信息。所述具体实施方式不应被视为以任何方式限制前面的发明内容。所述具体实施方式将参考如下几个附图：

[0066] 图1为根据本发明的实施例的咳嗽处理系统的框图。

[0067] 图2为根据本发明的一方面的优选实施例的方法的流程图。

[0068] 图3，TDNN的结构，其包括输入层、隐藏层和输出层。TDNN的输入为d个连续特征向量f_k。例如，针对d＝5，所述输入为$\{f_{k-4j}, f_{k-3j}, \cdots\cdots, f_{kj}\}$。为了归类所有子区块，将所述输入向右移一子区块。

[0069] 图4,咳嗽片段的持续时间(τ_s)的平滑概率密度函数。其具有为0.373s的平均,以及为0.144s的标准偏差。

[0070] 图5图示降噪过程。(A)记录中的典型信号,(B)为高通滤波器的输出,并且(C)为输出谱减法滤波器。所形成的滤波器显著改善所述信号的SNR(信噪比)。

[0071] 图6,随机选择的梅尔频率倒谱系数的概率密度函数(pdf)(出于显示目的而平滑化)。尽管存在交叠,系数M(4)能够被用于区分咳嗽与器具声音,系数M(9)在咳嗽与哭喊之间区分,并且M(11)在咳嗽与发声之间区分。

[0072] 图7,五个第一共振峰频率的pdf(F(1)-F(5))。尽管所述共振峰频率的分布交叠,但它们具有不同的平均、偏态和峰态,尤其针对F(1)、F(3)和F(5)。

[0073] 图8,能量、过零率(ZCR)、香农熵和非高斯性评分(NGS)的pdf。NGS能够被用于在咳嗽与发生和哭喊之间区分,而ZCR在咳嗽与器具声音之间区分。声音之间的对数能量和香农熵的轮廓相似,但它们具有不同的偏态和峰态。

[0074] 图9图示咳嗽片段识别。(A)来自降噪过程的声音信号,(B)来自TDNN的输出信号,(C)在平滑化过程之后的输出信号,以及(D)阈值处理过程之后的输出信号。来自所述阈值处理的信号可能包括假片段(例如FD片段)。

[0075] 图10,对来自分割算法的输入信号和输出信号的图示。虚线指示分割输出,能够在这里确定每个咳嗽片段的开始和结束。假片段FD(在图9(D)中示出的)被丢弃,因为其持续时间在持续时间的准则以外($\tau_{最小}<\tau_s<\tau_{最大}$)。所述算法成功将所有咳嗽声音分割,包括与说话(SP)交叠的CG3。

[0076] 图11,对信号的分割输出(虚线),其包含咳嗽和哭喊声音。我们的分割算法成功地分割两个咳嗽声音,并丢弃哭喊声音。仅小部分的哭喊声音被错误地检测为咳嗽片段。

[0077] 图12,针对使用咳嗽特征的肺炎诊断所提出的算法的框图。

[0078] 图13,样本肺炎咳嗽声音。在小心聆听咳嗽事件之后,手动标记咳嗽事件的开始和结束。

[0079] 图14,数字高通滤波器的幅值响应。肺炎归类方法。

[0080] 图15,针对湿咳和干咳声音归类所提出的算法的框图。

[0081] 图16,针对200个训练和测试数据集达到的灵敏度和特异性的柱状图。仅选择的特征被用于LR模型设计。

[0082] 图17为根据本发明的另外的实施例,用于对疾病状态(例如肺炎)的诊断的方法的流程图。

[0083] 图18包括呈现与图17的方法的应用相关联的对象的细节的图表。

[0084] 图19示出(a)肺炎咳嗽、(c)哮喘咳嗽、(e)支气管炎咳嗽和(g)毛细支气管炎咳嗽波形以及它们对应的功率谱图的典型范例。

[0085] 图20为根据本发明的一方面另外的且优选的实施例的自动咳嗽归类程序的框图。

[0086] 图21示出来自两位患者的干咳波形和湿咳波形的典型范例。

[0087] 图22A和图22B为针对咳嗽声音的SNR的柱状图。

[0088] 图23示出被用于图20中图示的方法的分析与讨论中的统计学信息。

具体实施方式

[0089] I.用于咳嗽归类的第一方法

[0090] 2.1.咳嗽记录协议

[0091] 咳嗽记录系统包括以下:具有超心型指向性束模式的低噪声麦克风53、55(澳大利亚悉尼Model NT3，RODE®)，之后是前置放大器52和A/D转换器51(美国加利福尼亚Model Mobile Pre-USB，M-Audio®)用于信号采集。Mobile Pre-USB的输出被连接到膝上型计算机52的USB端口50。从所述麦克风到对象42的口部的标称距离为50cm。实际距离能够因对象移动而在40cm至100cm之间变化。然而，所提出的方法不依赖于声音强度，并且结果独立于口部到麦克风的距离。我们保持采样率在44.1k样本/s以及16-bit的分辨率，以获得最佳声音品质。

[0092] 根据本发明的各方面的优选实施例，计算机系统52作为咳嗽/非咳嗽归类系统和/或基于咳嗽的诊断系统而操作，同时运行计算机程序，马上将描述所述计算机程序。个人计算机系统52包括为指向设备60和键盘58的形式的数据输入设备，以及为显示器56的形式的数据输出设备。所述数据输入与输出设备被耦合到包括中央处理单元70的处理箱54。显示器56包括人机接口，用于呈现由本文描述的所述方法实施的各个归类程序的结果。其他人机接口也是可能的，例如智能手机或打印机。

[0093] 中央处理单元(CPU)70与储存设备接口连接，所述储存设备可由机器读取并且有形地体现可由CPU运行的指令的程序。这些储存设备经由主板68包括RAM 62、ROM 64以及二级储存设备，即硬磁盘66和光盘读取器48。所述个人计算机系统也包括USB端口50，用于与外部ADC模块51的通信，外部ADC模块51前置放大、滤波并数字化来自麦克风53和55的信号。所述麦克风从躺在床40上的对象42拾取声音，例如咳嗽声音。

[0094] 二级储存设备66为带有用于由中央处理器70运行的有形指令的磁性数据储存介质。典型地将已从安装盘(例如光盘46)安装了这些指令，尽管也可以在存储器集成电路中，或经由计算机网络从远程服务器安装设施来提供它们。所述指令构成软件产品72，其被载入到RAM 62的电子存储器中。在被运行时，所述指令使计算机系统52作为基于咳嗽的诊断系统，和/或咳嗽或非咳嗽患者声音归类器而操作，并且尤其实施马上将描述的多种方法之一。

[0095] 本领域技术人员将认识到，软件产品72的编程直接鉴于本发明的所述方法，现在将描述所述方法的实施例。在以下方法中，各个变量均被操纵。将认识到，在计算机系统52用于实施所述方法的操作期间，CPU 70的对应的寄存器将被增量，并且借助电子信号沿被蚀刻在主板68上的传导总线行进，而将数据写到二级储存设备66和RAM 62并从二级储存设备66和RAM 62检索所述数据。因此，随着计算机系统72运行软件以实施现在将描述的方法，在计算机系统52内发生物理效应及变换。

[0096] 尽管所述计算设备已被视为包括个人计算机(例如台式或膝上型计算机)，但本发明的实施包含其他计算设备。例如，本发明的实施例包含专用医学设备以及还有装载有应用(即用于实施本文描述的方法中的一个或多个的“app”)的智能手机。在智能手机的情况中，所述手机的内置麦克风可以被用于监测所述患者声音，或者备选地一个或多个高品质外部麦克风可以被连接到智能手机用于这样的监测。根据本文描述的方法，智能手机可以

经由移动电话网络和/或因特网，将患者声音的数字记录传输到远处的计算机，所述计算机然后处理所述数字记录。可以将所述处理的结果传输回智能手机，用于在其上的显示，或者备选地在远处的计算机的控制下被显示在电子显示设备上。

[0097] 所述计算设备也可以被接口连接到多个患者传感器，例如移动传感器；EEG传感器；氧水平传感器、呼吸检测传感器、温度传感器。也可以在本文描述的各个方法的执行期间，使用来自这些传感器的数据。

[0098] 针对该工作的数据是在印度尼西亚日惹的Sardjito医院，从针对呼吸不适而被收治的儿科患者记录的。我们在自然医院环境中采集数据，除了将我们的声音记录系统放置在床边以外，没有以任何方式修改它(参见图1)。在白天期间，在两种类型的病房(单人和双人房间)中进行记录。研究协议已接收来自Sardjito医院和澳大利亚昆士兰大学的伦理审查。

[0099] 我们的数据库包括来自跨度5-64个月年龄范围的儿童的咳嗽声音。纳入与排除准则的细节参见表1。

[0100] 表1

[0101] 研究中使用的纳入与排除准则

[0102]

纳入准则	排除准则
- 呼吸道感染的症状，以下中的至少 2 个：咳嗽、有痰、呼吸困难以及体温＞37.5°	- 不预期恢复的重病 - 飞沫预防措施 - 具有无创通气处置
- 知情同意完成	- 无知情同意

[0103] 在医师已检查对象并且初始处置已开始之后，启动所述记录。针对每个对象的记录的持续时间为4-6小时。在表2中，我们说明所了对象的人口统计学细节。

[0104] 表2

[0105] 该研究中涉及的对象的人口统计学信息。

[0106]

对象	年龄（月）	性别	诊断
1	6	M	心力衰竭，肺炎
2	8	M	肺炎
3	2	M	鼻咽炎
4	8	M	肺炎
5	5	M	肺炎
6	12	F	肺炎
7	64	F	支气管肺炎
8	16	F	肺炎
9	15	F	扁桃体咽炎
10	16	F	鼻咽炎
11	9	M	肺炎
12	11	M	肺炎
13	30	M	支气管炎
14	11	M	肺炎

[0107] 2.2.咳嗽分割方法

[0108] 能够如下将离散化的声音记录，r[n]，建模为咳嗽声音$s_c[n]$、背景噪声b[n]与非咳嗽声音$s_{nc}[n]$(例如说话、哭喊、发声、器具移动等)的加和：

[0109] $r[n]=s_c[n]+b[n]+s_{nc}[n]$ (1)

[0110] 本发明的实施例提供一种用于从记录r[n]提取$s_c[n]$的自动方法。在图2中示出总体方法的框图。其包括四个主要过程：(a)降噪，(b)从数据子区块的特征提取，(c)数据子区块到咳嗽组(CG)与非咳嗽组(NG)类别的归类，以及(d)通过合适地分组CG/NG归类的连续数据子区块，识别/形成咳嗽片段。

[0111] 在2.2.1-2.2.4节，我们提供对所述方法的详细描述。

[0112] 2.2.1降噪

[0113] 为了降低背景噪声b[n]，我们通过两个不同的滤波器处理r[n]：(a)高通滤波器(HPF)，以及(b)功率谱减法(PSS)滤波器。HPF被设计为四阶巴特沃斯滤波器。其被用于降低可能来自麦克风架或床的移动的低频干扰。PSS滤波器被用于降低局部静止的背景噪声。其通过以原始声音信号的功率谱减去估计的背景噪声，估计干净声音信号[15]。能够如下改写滤波之后的输出信号$\hat{s}[n]$：

[0114] $\hat{s}[n]=\hat{s}_c[n]+\hat{s}_{nc}[n]$ （2）

[0115] 我们通过在降噪过程之前和之后计算信噪比[SNR]，测量了所述滤波器的性能，所述信噪比[SNR]被定义为信号幅值相对于背景噪声的量度。所述SNR被计算为：

[0116] $$SNR=20\log\left(\frac{R_s}{R_b}\right)$$ （3）

[0117] 其中R_s和R_b分别为信号$\hat{s}[n]$和背景噪声b[n]的均方根(rms)值。

[0118] 在所述降噪过程之后，$\hat{s}[n]$被用作对特征提取单元的输入。

[0119] 2.2.2对声音信号的特征提取

[0120] 在该节，我们描述咳嗽声音产生的模型以及特征提取单元的细节。

[0121] (a)咳嗽声音建模

[0122] 基于生理学考虑，咳嗽声音通常被视为四个不同相位的组合[1]：吸气、收缩、压缩和排出。所述吸气相位由吸入开始并且由声门、上喉部括约肌或两者的关闭而结束。在所述收缩相位，呼吸肌肉群收缩抵住声门结构，产生以齿槽、胸膜和声门下气道压明显升高为特征的压缩相位。在所述排出相位中，声门快速打开随后是空气在大压力梯度下的快速呼出。从肺部排出的空气的快速移动生成咳嗽声音，其具有来自呼吸系统的不同区域的贡献。咳嗽声音产生的机制与语言产生共有一些相似点。

[0123] (b)特征设计与提取

[0124] 为了获得声音信号的特征，我们对$\hat{s}[n]$应用长度为N的矩形滑动窗口$w_r[n]$，生成数据子区块。使第k个数据子区块由$s_k[n]$指代，因此$\hat{s}[n]$能够被表达为子区块$s_k[n]$的串联，即$\hat{s}[n]=\{|s_1[n]|,|s_2[n]|,\ldots,|s_k[n]|\}$，其中K为$\hat{s}[n]$中子区块的总数目。我们针对每个子区块$s_k[n]$计算以下特征。

[0125] i)梅尔频率倒谱系数(MFCC)：MFCC被广泛用于语言处理[16、17]，并且被发现针对打鼾分析也高度有用[18-21]。在该工作中，受咳嗽/呼吸声音与打鼾和语言的相似性的启示，我们探索了MFCC在咳嗽分割中的应用。我们计算每个子区块$s_k[n]$的第一L梅尔频率倒谱系数$M(l)_k=\{M(l)_k,1=1,2,\cdots,L\}$。通过利用信号$s_k[n]$乘以汉明窗口$w_h$，随后应用快速傅立叶变换(FFT)、梅尔频率滤波器组f_m滤波和离散余弦变换的连续过程，计算所述系数。在(4)中限定所述梅尔频率滤波器f_m，并且在(5)中给出$M(l)_k$。

[0126] $$f_m=\left[2595\log\frac{(1+f(\mathrm{Hz}))}{700}\right] \quad (4)$$

[0127] $$M(l)_k=\sqrt{\frac{2}{c}}\sum_{c=l}^{c}(\log D_c)\cos[n(c-0.5)\pi/c],n=1,2,...,N \quad (5)$$

[0128] 其中D_c(c=1,2,…,C)为梅尔滤波器组的输出。

[0129] ii)共振峰频率：在语言中，共振峰频率示出声道共振的特性；在打鼾声音分析中，它们指示上气道的共振。我们假设在咳嗽/呼吸声音中，共振峰可以携带整个呼吸道的共振。例如，源于肺部的支气管的振动的气喘声音可以贡献于咳嗽声音中较高频率的共振峰(共振频率)。在该工作中，我们估计第一P共振峰频率$F(p)_k=\{F(p)_k,p=1,2,...,P\}$。通过经由莱文森德宾递归求解Yule-Walker方程，确定LPC谱及其参数[22]。

[0130] iii)过零率(ZCR)：被定义为信号跨过零轴的总时间的ZCR是用于检测信号的周期性质而无论其幅值的简单但有用的方法。如下计算ZCR特征Z_k。

[0131] $$Z_k=\frac{1}{N-1}\sum_{m=l}^{N-1}\Pi\{S_k(m)S_k(m-1)<0\} \quad (6)$$

[0132] 其中，如果自变数A为真，则指示函数Π{A}为1，否则为0。

[0133] iv)非高斯性评分(NGS):非高斯性评分(NGS)提供一种容易的方法,以量化给定信号与高斯模型的偏差。在我们关于打鼾声音分析的先前工作中[23],该特征显示筛查阻塞性睡眠呼吸暂停的能力。为了获得NGS,在每个且每一个$s_k[n]$中,我们计算正太累积分布函数(γ)的倒数(F^{-1}),如在(7)中给出的。

[0134] $$\gamma = F^{-1}(p|\mu,\sigma) = \{\gamma : F(\gamma|\mu,\sigma) = p\} \quad (7)$$

[0135] 其中,μ,σ分别为$s_k[n]$的平均和标准偏差,并且在(8)中定义p。

[0136] $$p = F(\gamma|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\gamma} e^{\frac{-(1-\mu)^2}{2\sigma^2}dt} \quad (8)$$

[0137] NGS(N_k)为$s_k[n]$(γ)的概率图与在(9)中给出的其参考高斯概率图(g)的偏差,其中g[n]和γ[n]分别表示参考正态数据与被分析数据的概率。

[0138] $$N_k = 1 - \left(\frac{\sum_{n=1}^{M} g[n] - \bar{g}}{\sum_{n=1}^{M} \gamma[n] - \bar{\gamma}}\right) \quad (9)$$

[0139] v)香农熵:咳嗽声音是表示来自呼吸道的各个子结构的贡献的复杂信号。这些分量中的一些显示伪周期性结构,而其他的具有随机性的随机特性。在一些情况中,咳嗽声音具有从吸气相位到呼气相位的突然转变。该时期可以包含立即中段以及高频分量。在该工作中,我们计算了香农小波熵,以捕获这些特征。在小波分析中,能够使用互补滤波器,将在刻度i的信号$s_k[n]$分解成其高频分量和低频分量,所述互补滤波器包括低通和高通滤波器。将来自低和高滤波器的输出定义为近似与细节(分别为A_i和D_i)。使用在(9)中的定义,通过计算如在(10)中的小波能量谱E_i,获得小波香农熵S_k[24]。

[0140] $$S_k = -\sum_{i=1}^{l} E_i \log E_i \quad (9)$$

[0141] $$E_i = |D_i[n]^2| \quad (10)$$

[0142] 针对每个子区块$s_k[n]$计算在B(i)-B(v)中描述的特征。在每个特征中,我们计算M_k、F_k、Z_k、S_k和N_k中每个特征的最小值和最大值。我们使用这些值,以将对应的特征的分量归一化到-1至1的范围中。

[0143] 然后将$s_k[n]$的归一化参数向量f_k定义为:$f_k = \{[M_k\ F_k\ Z_k\ S_k\ N_k]^T\}$。针对第j个患者的总体特征矩阵$G_j$则由$G_j = \{f_{1,1}, f_{1,2}, \cdots, f_{k,j}, \cdots, f_{K,J}\}$给出。

[0144] 如图2中图示的,根据本发明的实施例的咳嗽分割技术通过以下而开始,即基于特征向量$f_{k,j}$,将每个且每一个数据子区块$s_k[n]$归类为咳嗽组(CG)和非咳嗽组(NG)的两个非重叠类别。

[0145] 在2.2.3节,我们描述我们在子区块水平针对该CG NG归类开发的模式归类方案。

[0146] 2.2.3子区块到咳嗽和非咳嗽类别的归类

[0147] 本发明的一个实施例利用人工神经网络(ANN)作为在子区块水品的CG/NG模式归类器。我们使用了通过人脑的功能启示的ANN,以识别不同类型的咳嗽声音,而无论它们的强度、持续时间或湿度。而且,ANN具有几个优点,在于其能够基于具有一套给定范例的监督式学习的过程,使用非线性决策边界归类数据。相比较常规方法,已证明其具有利用有限的训练数据集工作的能力[25]。在该工作中,我们使用了被称为时延神经网络(TDNN)的特定形式的ANN[26],其已在语言识别应用中获得成功。TDNN能够归类减损输入特征集的时间平

移[26]的数据子区块$s_k[n]$。

[0148] 我们在图3中示出TDDN的结构。其包括输入层、隐藏层和输出层。在TDNN中，我们使用矩阵G_j的d个连续特征向量，以将子区块归类为CG/NG类别。为了归类下一个子区块，我们将输入右移一子区块，并且使用下一组的特征向量。我们从矩阵G_j的开始到结束执行该过程，因此能够将$\hat{s}[n]$中的全部子区块都归类成CG/NG类别。

[0149] 贯穿训练、验证和测试的整个过程，执行使用TDNN对子区块的归类。为了完成这些过程，我们准备数据集，选择训练参数，训练TDNN，并且验证归类结果。在以下子节中描述这些过程的细节。

[0150] (a)TDNN数据集准备

[0151] 我们形成了包括来自记录的咳嗽和其他声音的数据集。在我们的数据集中有665个咳嗽声音。我们将我们的数据集划分成三个非重叠的随机分区，并将它们分别命名为训练(TDS)、验证(VDS)和测试(TeDS)。TDS、VDS和TeDS占数据库的比例(分别为40％、30％和30％)。我们用符号C_m指代第m个这样的随机分区，m＝1,2,…,M，其中M为该工作中使用的随机分区的总数目。独立地分析全部归类数据集。我们通过应用使用随机子采样(sub-sampling)技术的K-折交叉验证方法，有效地使用我们的数据集。

[0152] 通过使用归类数据集的信息咳嗽片段位置，我们定义矩阵$W_k＝[w_1,w_2,...,w_k,\cdots,w_K,w_k＝\Pi\{\text{CG类别}\}]$，其中$w_k$为当$\Pi\{\text{CG类别}\}$的自变数为正确时由“1”填充否则由“0”填充的每个子区块的类别。矩阵W_k被用作训练过程中的目标。注意，W_k的维数类似于G_j。

[0153] (b)TDNN训练过程

[0154] 用于该研究的TDNN具有如在表3中规定的参数。

[0155] 表3

[0156] TDNN训练参数

[0157]

参数	细节
层数	3
每层的神经元数目	输入层（L_i）=110， 隐藏层（L_h）=20， 输出层（L_o）=1
激活函数权重与初始化	正切sigmoid与线性 Nguyen-widrow
训练方法	弹性反传（RPROP）
停止准则	均方差=10^{-5}， 最小梯度=10^{-10}， 最大验证失败=100
连续输入（d）	5

[0158] 我们在3.4节解释了被用于确定对神经元的数目的选择的程序。

[0159] 我们构建了TDNN的网络，用net_c指代，其包括输入层L_i、隐藏层L_h和输出层L_o。线性与sigmoid函数被选择为层之间的激活函数。为了确定初始权重和偏差，我们使用了

Nguyen-Widrow初始化方法。

[0160] 为了在训练过程期间更新神经元的偏项和权重，我们采用弹性反传(RPROP)算法[27]。发现RPROP成功地克服了确定学习率用于更新神经元的权重的问题。在梯度下降算法中，太小的学习率增加了计算时间，而太大的学习率将在以合理误差到达最佳点之前产生振荡。相反，RPROP使用自适应的值以更新权重，因此其能够快速达到最大优化。在[27]中描述了该算法的细节。

[0161] 在所述训练过程中，我们定义了停止准则：(a)训练数据的均方差(mse)小于10^{-5}，(b)验证误差开始增大，以及(c)达到训练性能中为10^{-10}的最小梯度。除非训练被更早停止，否则允许神经网络训练直到100次。

[0162] 一旦已设定了这些参数，我们开始训练TDNN。在训练过程中，我们让TDNN的网络net_c学习将子区块归类成CG/NG类别。通过提供d个从G_j到net_c的连续特征向量以逼近W_k中的目标，进行所述过程。我们调查了f_k中的特征的几个组合，以寻找在归类CG和NG类别中提供最佳结果的输入的组合。我们在训练过程期间将net_c的输入指代为$\hat{W}_k$，W_k的近似值。为了识别咳嗽片段，如在以下节中处理$\hat{W}_k$。

[0163] 2.2.4咳嗽片段的识别

[0164] 在该节中，我们描述在片段识别过程中用于通过处理TDBB的输出($\hat{W}_k$)，获得咳嗽片段的技术。所述过程包括以下：平滑化TDNN的输出，应用阈值到平滑化的信号以获得片段，对片段的持续时间的测量，以及基于准则对片段的选择。在以下步骤中描述总体过程：

[0165] (S1)定义参数β，确定移动平均滤波器(H)的跨度的小的正整数。对滤波器H应用$\hat{W}_k$以产生平滑信号$\hat{w}_k$。

[0166] (S2)应用阈值(λ)到平滑信号$\hat{w}_k$。来自该过程的输出信号被指代为在(11)中给出的$\ddot{w}_k$。$\ddot{w}_k$中连续元素的组被定义为咳嗽片段候选。

[0167] $$\ddot{w}_k \begin{cases} 1, \text{针对} \hat{w}_k \geq \lambda \\ 0, \text{针对} \hat{w}_k < \lambda \end{cases} \quad (11)$$

[0168] (S3)计算$\ddot{w}_k$中包含‘1’的对应片段的持续时间(τ_s)。

[0169] 通过确定每个片段的开始和结束执行该过程。

[0170] a.开始k＝1，寻找k的最小值(称作k_a)，其中全部子区块$\ddot{w}_k$对应于NG类别的类。取子区块$\ddot{w}_k$，k＝k_a的开始为NG片段的开始。

[0171] b.为了识别咳嗽片段的开始，在k_a<k<K的域中开始搜索。寻找该域中k的最小值(称作k_b)，使得$\ddot{w}_k$中没有一个子区块对应于NG。取子区块$\ddot{w}_k$，k＝k_b的开始为CG片段的开始。

[0172] c.在类似于步骤(S3.a)的过程中确定NG的结束，其中搜索针对k被设定为k_b<k<K的域。

[0173] (S4)定义$\tau_{最小}$和$\tau_{最大}$的参数，小的正分数，其确定来自儿童的典型咳嗽声音的最小和最大片段持续时间。持续时间在该约束外的片段将被丢弃。我们将该过程的输出信号指

定为$\dot{w}_k$。

[0174] 步骤(S1)-(S4)描述了用于确定$s_c[n]$每个片段的开始与结束，以及用于削减记录中$s_{nc}[n]$的片段的方法。为了验证结果，将在$\dot{w}_k$中全部获得的咳嗽片段与W_k中的咳嗽片段进行比较。

[0175] II.结果与讨论

[0176] 3.1咳嗽片段持续时间

[0177] 在我们的临床数据中，我们通过视觉观察与倾听的组合过程，识别了每个咳嗽片段的持续时间(τ_s)。在图4中图示τ_s的平滑化概率密度函数。咳嗽片段具有为($\tau_{最小}$=0.18s和$\tau_{最大}$=11.6s)的最小和最大持续时间。咳嗽片段的平均和标准偏差为(μ_s=0.373，σ_s=0.144)。在分割过程中，我们使用$\tau_{最小}$和$\tau_{最大}$作为准则之一，以将预测片段确定为咳嗽或非咳嗽声音。我们在3.5节中描述该过程。

[0178] 3.2降噪

[0179] 非受控环境中的记录使得数据更容易受背景噪声影响。在我们的记录中，我们识别了噪声的两个独立分量：(a)10Hz以下的低频噪声和(b)白高斯噪声(参见图5(A))。从所述图可见，记录在1.54-5.92dB的范围上具有差的信噪比(SNR)。为了减少这些噪声，我们使用了具有10Hz截止频率的四阶巴特沃斯高通滤波器(HPF)以及功率谱减法(PSS)滤波器。

[0180] 这些滤波器显著改善了SNR。HPF使其增大到11.98-17.19dB(图5(B))并且PSS进一步使其增大到12.62-17.44dB(图5(C))。

[0181] 这些结果显示，所开发的滤波器能够降低背景噪声，并改善记录信号的SNR。

[0182] 3.3声音特征的特性

[0183] 如在2.2.2节中所述，我们针对$s_c[n]$的每个子区块，计算了特征向量f_k={M_k F_k Z_k S_k N_k}。我们使用子区块大小N=882(20ms)。所选择的子区块大小为用于对MFCC的特征提取的基本单元。更小的大小(例如10ms)由于在梅尔滤波器积分中涉及很少的样本而一般导致失真的结果；更大的大小倾向于增大误字率[28]。

[0184] 特征向量f_k包含22个元素：14个M_k的系数(包括对数能量以及第0-第12倒谱系数)，5个F_k的系数，以及每个Z_k、S_k和N_k的1个系数。为了揭示这些特征的特性，我们计算了在CG和NG类别中的特定声音的概率密度函数(pdf)。NG表示宽范围的非咳嗽声音，因此在该工作中，我们选择最主要的声音，例如哭喊(CY)、发声(简写为VC，例如讲话、典型的宝宝声)以及器具声音(简写为AS，例如来自门框、推车、床的声音)。

[0185] 我们在图6-8中图示了f_k的每个特征的平滑pdf。如从这些图可见，特征(M_k、F_k、Z_k、S_k和N_k)在咳嗽与其他声音之间的分布是重叠的。然而，特征的每个分量具有独特的分布。

[0186] 图6示出来自MFCC的随机选择的元素(M(1)、M(4)、M(9)和M(11))的pdf。从图5(A)和图5(B)可见，M(1)的pdf在CG和AS类别中具有不同的平均值(μ=0.31对0.18)。M(9)的pdf在图5(C)中显示，CG具有比CY低的平均值(μ分别=-0.01对=-0.13)。

[0187] 图7展现共振峰频率的pdf。F(1)、F(3)和F(5)的共振峰频率的统计分布(平均，标准偏差，偏态，以及峰态)在CG与AS之间是可区别的。F(2)的分布在CG、CY、VC和AS中的分布看起来相似，然而CG具有最低的平均值(-0.06)。而且，F(4)在TS中的分布在类别中具有最低的平均值(-0.003)。

[0188] 在图8中，我们图示CG、CY、VC和AS的NGS指数。NGS具有将CG与VC和CY区分的潜力

(图7(D))。类似地,从图6(B)可见,ZCR能够被用于区分CG与AS。

[0189] 特征的pdf显示,没有显著特征能够被单独用作针对TDNN的输入以归类CG/NG类别。因此,为了获得所述特征的每个分量的最大利益,我们将它们组合并将d个连续特征向量用作TDNN的输入,以归类CG/NG类别。

[0190] 3.4在CG/NG归类中的基于TDNN的分析

[0191] 在该工作中,我们使用了五个连续特征向量(d=5);因此输入层中神经元的数目为(L_i=110)。在所述输出层中,仅有一个神经元(L_o=1)以代表CG或NG类别。为了确定L_h中神经元的数目,我们在训练过程期间测试了几个数目的神经元。

[0192] 为了评价网络(net_c)的性能,我们计算了有以下等式定义的灵敏度和特异性。

[0193] 灵敏度=TP/(TP+FN)×100 (12)

[0194] 特异度=TN/(FP+TN)×100 (13)

[0195] 其中TP=真阳性,TN=真阴性,FP=假阳性,并且FN=假阴性。

[0196] 在表4中,我们示出针对不同值的L_h的net_c的性能。

[0197] 表4

[0198] TDNN针对隐藏层(L_H)不同组合的性能。具有20个神经元的隐藏层显示最好的归类性能

[0199]

隐藏层（L_h）中神经元的数目	数据集	性能	
		灵敏度	特异性
		y	y
10	TDS	98.34	98.35
	VDS	85.24	85.24
	TeDS	90.32	90.35
20	TDS	98.12	98.13
	VDS	86.80	86.61
	TeDS	91.51	91.51
30	TDS	98.31	98.32
	VDS	85.03	85.05
	TeDS	90.24	90.25
40	TDS	98.41	98.41
	VDS	85.57	85.57
	TeDS	91.45	91.45
50	TDS	98.43	98.43
	VDS	85.69	85.69
	TeDS	89.89	89.92

[0200]

[0201] 从表4可见，在测试集(TeDS)中，在隐藏层中具有20个神经元的net_c具有最高的灵敏度和特异性(两者均为91.51％)。因此，针对该讨论的剩余部分，我们固定网络net_c(L_i＝110，L_h＝20，L_o＝1)。net_c中神经元的数目足够紧凑以被用于具有有限的计算资源的系统中。小数目的层和神经元也改进了网络用于开发针对CG/NG归类的最佳模型以及用于避免存储训练中给出的样本的能力。

[0202] 我们在表5中示出使用特征的不同组合，子区块$s_k[n]$到咳嗽(CG)和非咳嗽(NG)类别的归类的性能。

[0203] 表5

[0204] TDNN在特征的不同组合中的性能。{M_k F_k Z_k S_k N_k}的完整组合具有最优的灵敏度和特异性。

[0205]

数据集	M_k+F_k		$Z_k+S_k+N_k$		$M_k+F_k+Z_k+S_k+N_k$	
	灵敏度	特异性	灵敏度	特异性	灵敏度	特异性
TDS	99.05	99.05	87.00	87.00	98.12	98.13
VDS	85.69	85.73	81.44	81.48	86.80	86.81
TeDS	89.37	89.37	79.95	79.98	91.51	91.51

[0206] 在所述测试集(TeDS)中，使用{Z_k S_{hk} N_k}的归类结果显示最低的灵敏度和特异性(分别为79.95和99.98)。{M_k F_k}的组合具有较高的结果，89.37％的灵敏度和特异性。与之对比，{M_kF_k Z_k S_k N_k}的组合达到为91.51％的最优灵敏度和特异性。接下来，我们探讨了组合{M_k F_k Z_k S_k N_k}作为针对TDNN的输入。

[0207] 在表6中示出TDNN针对M＝14，(C_m，m＝1，2，…，14)的性能。

[0208] 表6

[0209] 针对M＝14，(C_m，m＝1，2，…，14)，来自TDNN的NG/CG归类结果的汇总。

[0210]

数据集	灵敏度		特异性	
	平均	标准偏差	平均	标准偏差
TDS	96.77	0.0121	96.78	0.0121
VDS	89.40	0.0170	89.41	0.0170
TeDS	91.13	0.0073	91.14	0.0073

[0211]

[0212] 计算针对每次实现的归类结果，C_m，m＝1，2，…，14，并且估计灵敏度和特异性的平均和标准偏差(分别为91.13±0.0073％和91.14±0.0073％)。2.3.咳嗽片段从子区块的形成

[0213] 我们在图9中图示了从在3.4节中归类的子区块开始，咳嗽片段的形成。图9(A)示出来自降噪过程的信号，而图9(B)图示来自基于TDNN的归类($\hat{w}_k$)的输出信号。在图9(B)中，我们示出网络net_c的输出，表示对CG和NG类别的状态的近似值(分别为1和0)。如可见，与咳嗽声音相关联的子区块的组形成具有接近于1的值的片段，并且否则为来自非咳嗽声音的子区块的组。从图9(B)也可见，网络net_c有效地将子区块归类成CG和NG类别，即使它们

与讲话重叠(在CG3信号中图示)。然而,CG2和CG5片段的中间部分中的子区块中的一些被错误归类为NG类别,由从高到低的值的尖锐转变指示。

[0214] 为了细化片段并减少片段中被错误归类的子区块,我们基于移动平均滤波器,将信号处理成平滑过程。在该阶段中,我们小心地选择移动平均滤波器的跨度为(β=9)。我们在图9(C)中示出从平滑过程获得的平滑信号($\ddot{w}_k$)。如从图9(C)可见,对于的片段中的尖锐转变部分得以减少。

[0215] 为了确定咳嗽片段的开始和结束,将$\ddot{w}_k$中的全部元素与阈值(λ)进行比较。如果$\ddot{w}_k > \lambda$,则对应的元素将被设置为1,否则对于($\ddot{w}_k < \lambda$)为0,其中λ被设置为0.05。选择该λ的值以使归类器性能最大化。我们在图9(D)中示出来自阈值处理过程的输出信号($\dot{w}_k$)。其现在包含两组从1与0元素形成的片段。我们将具有为1的元素的片段表示为咳嗽片段候选。

[0216] 接下来,我们将咳嗽片段候选中的第一元素定义为片段的开始,并且将最后的元素定义为片段的结束。从3.3和3.4节我们知道,片段中的每个元素都代表20ms子区块的类别。因此,我们能够通过将片段的开始与结束之间的元素的数目乘以20ms,计算片段的持续时间。我们计算了每个片段的持续时间(τ_s),并将τ_s相对于在3.1节中获得的最小和最大咳嗽持续时间进行测试。仅将持续时间在边界以内($\tau_{最小} < \tau_s < \tau_{最大}$)的咳嗽片段选择为咳嗽片段。在图10中示出对该过程的图示。从该图可见,所提出的方法有效地分割咳嗽(CG1-CG6),即使它们具有不同的强度和持续时间。而且,尽管咳嗽声音CG3与讲话重叠,但所述方法能够提取对应的咳嗽声音。图9(D)中示出的假片段(FD)也已被消除。

[0217] 图11图示所述分割方法在典型的信号中的输出,所述信号包含与哭喊生混杂的咳嗽声音。所述方法成功地从所述信号提取到两个咳嗽声音,并且仅从长持续时间的哭喊中检测到一个假片段。

[0218] 我们在表7中示出我们的分割方法针对M=14,(C_m,m=1,2,…,14)的性能。

[0219] 表7

[0220] 咳嗽分割方法针对M=14,(C_m,m=1,2,…,14)的性能。

[0221]

数据集	灵敏度		特异性	
	平均	标准偏差	平均	标准偏差
TDS	99.20	0.77	98.80	0.68
VDS	94.85	2.86	95.00	2.18
TeDS	95.07	3.43	94.76	4.02

[0222] 计算针对每次实现,C_m,m=1,2,…,14,的分割结果并估计灵敏度和特异性的平均和标准偏差(分别为95.07±3.43%和94.76±4.02%)。结果也示出,相比较子区块归类的结构,咳嗽片段从子区块的形成过程使灵敏度和特异性改善了约4%。所提出的方法成功地削减了诸如哭喊、发声的声音,以及其他环境声音。

[0223] 在该工作中,我们使用非接触式传感器记录了来自印度尼西亚儿科病房中的儿童的咳嗽。我们发现,非接触式传感器因容易部署和感染控制,而针对该目的是优选的。

[0224] 我们的工作集中在小于5.5岁的儿科群体,这是先前的研究所未触及的区域。该年

龄范围的儿童更易受呼吸系统疾病侵害，显示有高发病率和死亡率。因此，用于开发分析来自该群体的信号的方法的研究极为重要。

[0225] 肺炎/非肺炎以及湿咳/非湿咳声音归类方法

[0226] 现在参考图12，示出图示根据本发明另外一方面的优选实施例的方法的框图。开发图12中图示的方法用于对与患者相关联的特定疾病状态的诊断，例如肺炎/非肺炎归类。在该说明书的最后讨论根据本发明的实施例的归类方法另外的实施例。

[0227] 临床评估及纳入-排除准则

[0228] 该工作的数据采集环境为印度尼西亚Gadjah Mada大学Sardjito医院的的呼吸医学单元。我们的对象群体包括具有呼吸系统疾病的症状的个体。由临床医师针对存在的症状评估医院的普通病房收治的患者。在标准纸上记述全部观察结果。观察结果包括咳嗽、流涕、呼吸困难、腹泻等等的存在。临床医师也记录温度、呼吸速率、SpO2和心率。

[0229] 表8列出纳入与排除准则。接洽满足纳入标准的全部患者。完成知情同意。在患者收治的前12小时内招募他们。在知情同意之后，针对接下来的4-6小时进行连续声音记录。

[0230] 表8.研究中使用的纳入与排除准则

[0231]

纳入准则	**排除准则**
-患者有胸部感染的症状：以下中的至少 2 个 -咳嗽 -有痰 -呼吸困难 -体温＞37.5° -知情同意	-不预期恢复的重病，例如晚期肺癌 -飞沫预防措施 -需要 NIV -无知情同意

[0232] A.数据采集与记录系统。

[0233] 声音记录是在医院的普通成人病房中进行的。患者与通过窗帘隔开的5名其他患者共用房间。患者由它们的家人陪伴。主治医师定期访问患者，然而不记录与患者有关的机密信息。记录中存在的常见噪声来自吊扇、脚步、讲话、敲门声、推车移动以及来自房间外部的其他模糊的噪声。

[0234] 使用两个系统进行咳嗽声音记录，

[0235] (i)计算机数据采集系统——高保真度专业品质的前置放大器和A/D转换器单元，(美国加利福尼亚M-Audio的Model Mobile-Pre USB)，带有相匹配的一对具有超心型指向性束模式的低噪声麦克风(澳大利亚悉尼RODE的Model NT3)。

[0236] (ii)便携式记录系统——高端轻质便携式2-AA电池供电的录音机(Olympus LS-11)，具有两个精密电容式麦克风。

[0237] 在两种记录设置中，我们使用44.1kHz的采样速率，具有16bit的分辨率。从麦克风到患者的口部的标称距离为50cm，但由于患者移动而能够从40cm到70cm变化。

[0238] 我们接收到数字声音数据，辅以来自医院的临床诊断信息。在严格的口令控制方案下，将数据完全去除识别信息并且储存在安全防火墙之后。对数据的访问被限制到由主要研究者个体授权的参与研究者。

[0239] B.根据咳嗽声音数据的特征计算

[0240] 图12示出针对肺炎/非肺炎归类形成的总体算法的框图，而图15示出针对湿咳/非湿咳声音归类形成的算法的框图。

[0241] 图13示出具有两个咳嗽事件的咳嗽声音的样本。在仔细聆听之后手动标记咳嗽事件的开始和结束。在对咳嗽事件的开始和结束的手动评分之后，我们遵循之后的步骤，以根据咳嗽事件数据计算数学特征。

[0242] [C1].使x[k]指代离散时间咳嗽声音的第k个样本。滤波器x[k]使用数字高通滤波器以得到y[k]。由(1)给出所述滤波器在z-变换中的传递函数。图14示出具有B＝[1-0.85]和A＝[1]的滤波器响应。

[0243] $$Y(z)=\frac{B(1)-B(2)z^{-1}}{1+Az^{-1}}X(z) \quad (1)$$

[0244] 针对湿咳/非湿咳，声音归类将y[k]划分成‘n’个相等大小的片段。使$y^n[k]$表示y[k]的第n个片段。

[0245] 针对肺炎/非肺炎归类，不考虑该子分割。

[0246] [C2].在肺炎归类的情况中从经滤波的咳嗽片段y[k]，并且在湿咳/非湿咳归类的情况中从每个经滤波的子片段$y^n[k]$，计算以下特征。

[0247] 1.非高斯性评分(NGS)——NGS给出数据的给定片段的非高斯性的量度。正常概率图能够被用于获得对一组数据的高斯性的视觉量度。能够使用(2)计算数据片段y[k]的NGS。注意在(2)中，p(k)和q(k)表示概率，并且p和q分别代表参考正常数据与被分析数据的平均值。能够在[29]中找到计算NGS的详细方法。

[0248] $$NGS=1-\left(\frac{\sum_{k=1}^{N}p[k]-\bar{p}}{\sum_{k=1}^{N}q[k]-\bar{q}}\right) \quad (2)$$

[0249] 2.共振峰频率——在人类语音分析中，共振峰被称作人类声道的共振。它们表现为在声学信号的LPC谱中的最大幅度峰。我们在我们的特征集中包含了第1四个共振峰频率(F1、F2、F3、F4)。以往在语言和声学分析中的研究已显示，F1-F4对应于上呼吸道的各个声学特征[30]。我们通过峰值拾取所述LPC谱，计算F1-F4。针对该工作，我们使用了第14阶LPC谱，并且连同Levinson-Durbin递归程序一起经由Yule-Walker自回归方法来确定其参数[31]。

[0250] 3.对数能量(LogE)——使用等式3计算针对片段y[k]的对数能量。

[0251] $$\log E=10\log_{10}\left(\varepsilon+\frac{1}{N}\sum_{k=1}^{N}\left(y(k)^2\right)\right) \quad (3)$$

[0252] 其中ε为任意小的正常数，添加其以防止任意偶然的为0的对数计算。

[0253] 4.过零率(Zcr)——针对每个咳嗽事件计数的过零的数目。

[0254] 5.峰态(Kurt)——峰态为使用(4)计算的与咳嗽事件数据y[k]的概率分布相关联的峰度的量度，其中(4)中的μ和σ为y[k]的平均和标准偏差。

[0255] $$kurt=\frac{E(y[k]-\mu)^4}{\sigma^4} \quad (4)$$

[0256] 6.梅尔频率倒谱系数(MFCC)——MFCC通常被用于语音分析系统[32]。它们基于非

线性梅尔刻度频率上的对数功率谱的余弦变换，表示声学信号的短期功率谱。我们在我们的特征集中包含12个MFCC系数。

[0257] 7.双频谱分(BS)——信号的三阶谱被称作双频谱。能够经由估计三阶累积量并然后进行2D傅立叶变换，估计所述双频谱，遵循该方法，其被称为估计所述双频谱的间接方法。使用(5)估计所述三阶累积量$C(\tau_1,\tau_2)$，如在[33]中定义的。

[0258] $$C(\tau_1,\tau_2)=\frac{1}{L}\sum_{k=0}^{L-1}y(k)y(k+\tau_1)y(k+\tau_2),\quad |\tau_1|\leq Q,\quad |\tau_2|\leq Q \qquad (5)$$

[0259] 通过向累积量估计应用双频谱窗口函数，获得窗口累积量函数$C_w(\tau_1,\tau_2)$。为此目的，我们使用在[34]中描述的最低双频谱-偏置上确界窗口(minimum bispectrum-bias supremum window)。

[0260] 使用(6)将片段y[k]的双频谱$B(\omega_1,\omega_2)$估计为累积量估计$C_w(\tau_1,\tau_2)$的2D傅立叶变换。我们使用512个点的FFT长度。

[0261] $$B(\omega_1,\omega_2)=\sum_{\tau_1=-\infty}^{\tau_1=+\infty}\sum_{\tau_2=-\infty}^{\tau_2=+\infty}C^{yi}(\tau_1,\tau_2)e^{-i(\tau_1\omega_1+\tau_2\omega_2)} \qquad (6)$$

[0262] 在频域中，可以针对数据片段y[k]定义量$P(\omega;\varphi,\rho)$，使得

[0263] $$P(\omega;\varphi,\rho)=B(\omega,\varphi\omega+\rho) \qquad (7)$$

[0264] 描述了以角度$\tan^{-1}\Phi$偏向于ω_1轴并且沿ω_2轴从原点移动量ρ，$(-\pi<\rho<\pi)$的一维切片[5]。针对该工作，我们设$\Phi=1$并且$\rho=0$，使得所考虑的双频谱的切片偏向于ω_1轴45度，并且穿过原点(即，在(ω_1,ω_2)平面中由$\omega_1=\omega_2$描述的线)。

[0265] 然后使用(8)计算双频谱分(BS)。在(8)中我们使用$\omega_1=90$hz，$\omega_2=5$khz，$\omega_3=6$khz并且$\omega_4=10.5$khz。

[0266] $$BI=\frac{\int_{\omega1}^{\omega2}P(\omega)}{\int_{\omega3}^{\omega4}P(\omega)} \qquad (8)$$

[0267] 在特征计算之后，针对每个咳嗽片段，在肺炎归类的情况中我们形成包含21个特征的特征向量(12个来自MFCC;4-共振峰频率;来自NGS、LogE、Zcr、Kurt和双频谱中的每个各1个)，并且在湿咳/非湿咳归类的情况中我们得到包含63个特征的向量(36个来自MFCC;12-共振峰频率;来自NGS、LogE、Zcr、Kurt和双频谱中的每个各3个)。

[0268] C.肺炎归类模型与特征优化

[0269] 我们使用基于咳嗽的特征，遵循针对肺炎归类的三步过程，所述三步过程如下：

[0270] 步骤1.在第一步骤中，我们将每个咳嗽事件归类为两个类别之一：‘肺炎咳嗽’或‘非肺炎咳嗽’。为了解决该问题，我们针对每个咳嗽事件形成特征向量，并且推导模型，所述模型能将咳嗽事件分成两个类别。

[0271] 利用如在表9中所示的不同的特征集推导不同的模型。

[0272]

特征模型	年龄（月）	发热的存在	呼吸指数	咳嗽特征
LR-模型 1				√
LR-模型 2			√	√
LR-模型 3		√		√

[0273]

LR-模型 4		√	√	√
LR-模型 5	√	√	√	√
年龄		针对快速呼吸的呼吸阈值		
1 个月-60 个月		每分钟 40 次呼吸		
>60 个月		每分钟 20 次呼吸		

[0274] 表9.用于咳嗽事件到类别‘肺炎咳嗽’和‘非肺炎咳嗽’的归类的推导的逻辑回归模型。

[0275] ‘√’指示该特征在模型设计中的纳入。呼吸指数是通过从所记录的呼吸速率减去呼吸阈值而计算的。

[0276] 针对所述咳嗽归类，我们使用逻辑回归统计模型。其为广义线性模型，使用几个预测器(自变量)以估计分类事件的概率(因变量)。

[0277] 在该工作中(肺炎归类建模)，假设因变量Y针对肺炎咳嗽等于“一”(Y＝1)并且针对非肺炎咳嗽等于“零”(Y＝0)。

[0278] 使用逻辑回归函数推导模型，以估计概率Y＝1(即属于‘肺炎咳嗽’的种类的咳嗽事件)，假设自变量(即特征集)如下：

[0279] $$P(Y=1|_{x1,x2,x3,\ldots,xn})=\frac{e^{z}}{e^{z}+1} \quad (9)$$

[0280] $z=\beta_0+\beta_1x_1+\beta_2x_2+\ldots+\beta_nx_n$ (10)

[0281] 在(10)中β_0被称作截距，并且β_1，β_2等等分别被称作自变量(特征)x_1，x_2的回归系数。

[0282] 使用MATLAB统计工具箱7.5版设计逻辑回归模型。选择特征以在最终模型中仅包含有助于所述归类的最优自变量(具有低‘p’值的变量)。然后将所述最终模型用于估计概率P，并使用概率阈值P_{th}将每个咳嗽事件归类为属于两个类别中的任一个。

[0283] 步骤2.在第二步骤中，针对每位患者，我们加和‘肺炎咳嗽’的数目并使用(9)计算被称作‘肺炎咳嗽指数(PCI)’的指数。

[0284] $$PCI=\frac{\text{肺炎咳嗽的总数目}}{\text{咳嗽的总数目}} \quad (11)$$

[0285] 步骤3.在第三步骤中，我们应用(在保持特异性>75％的同时针对高灵敏度而优化的)预定阈值PCI_{th}到所述PCI，以将患者归类为两类，‘I类-肺炎’和‘II类-非肺炎’。

[0286] 湿咳/非湿咳归类模型

[0287] 针对湿咳和非湿咳归类，我们再次使用逻辑回归统计模型。在该情况中，假设因变

量Y针对湿咳等于“一”(Y＝1)并且针对非湿咳等于“零”(Y＝0)。

[0288] 使用逻辑回归函数推导模型,以估计概率Y＝1(即属于‘湿咳’的种类的咳嗽事件),假设自变量(即特征集)使用(9)和(10)。然后选择特征以仅包含有助于归类的最优自变量(具有低‘p’值的变量)。然后将最终模型用于估计概率P,并使用概率阈值将每个咳嗽事件归类为属于两个类别中的任一个。

[0289] III.3.肺炎归类结果

[0290] A.数据库和临床诊断

[0291] 从81个对象声音记录分析总计541个咳嗽事件(最少2个、最多12个,并且平均6.6+2.14个咳嗽事件)。81个对象中,50个为肺炎患者,并且31个为非肺炎患者。非肺炎患者患有以下疾病,例如,哮喘-11、支气管炎-8、鼻咽炎-6,以及其他(哮喘、扁桃体咽炎、心脏病、喉软骨软化病、异物吸入)-6。在全部疑似肺炎患者上执行胸部X光。未在其上执行胸部X光的十九名患者中,8名为哮喘、5名鼻咽炎、2名支气管炎并且4名其他疾病患者。表10给出患者数据库的年龄与性别统计资料。

[0292]

平均年龄	最小年龄	最大年龄	＜2个月	2-11个月	12-60个月	＞60个月	男性对象	女性对象
41.75±1.45	1	191	3	28	30	20	43	38

[0293] 表10.该研究中使用的患者的年龄和性别统计资料

[0294] 针对在2.D节步骤1中使用逻辑回归推导咳嗽归类模型,我们将我们的数据库随机划分成两组,训练集和测试集。为了验证所述模型,我们从81个对象的数据库生成了200个不同的训练和测试数据集。每个训练集和测试互相排斥并且分别有56和25个对象。

[0295] B.针对肺炎诊断的WHO准则vs临床诊断

[0296] 表11示出针对我们81个对象的数据库,使用WHO准则的肺炎诊断与临床诊断的肺炎病例的列联表。

[0297]

WHO准则	灵敏度	特异性	准确度
[咳嗽或呼吸困难&快速呼吸]	92	26	66.67
针对快速呼吸使用的阈值为		临床上证实的诊断	
<2个月——60 BPM	WHO准则	46	23
2-11个月——50 BPM		4	8
12-60个月——40			
>60个月——20 BPM			

[0298] 表11.使用WHO准则的肺炎诊断vs临床诊断的肺炎的列联表。BPM—每分钟呼吸次数。

[0299] C.基于肺炎咳嗽指数的肺炎诊断

[0300] 表12示出在使用表9中给出的设计的4个逻辑回归模型将咳嗽事件分成两类时的训练与测试归类结果。

[0301]

特征模型	训练数据集（来自 56 个对象的 374±9 咳嗽事件）			测试数据集（来自 25 个对象的 166±9 咳嗽事件）		
	灵敏度	特异性	准确度	灵敏度	特异性	准确度
LR-模型 1	68.84±2.4	69.12±2.4	68.94±2.4	65.2±8	64.5±9	65±5
LR-模型 2	78.7±2.4	79±2.4	79±2.4	73.8±10	75.5±10.4	74.3±5.5
LR-模型 3	77.9±2.9	78±2.9	78±2.9	75.4±9	72.6±12	74.4±5.7
LR-模型 4	87.1±2	87.4±2	87.2±2	83.5±8.5	80.5±10	82.3±5
LR-模型 5	90±2.6	90.1±2.6	90±2.6	85.6±8	80.1±12	83.5±5

[0302] 表12.在2.D节步骤1中给出的方法之后，使用表3中示出的不同逻辑回归模型将咳嗽事件归类成两个类别（‘肺炎咳嗽’和‘非肺炎咳嗽’）的归类结果。被用于形成所述模型的选择的咳嗽特征为BSG、FF1、FF2、FF4、ZCR、MFCC3和MFCC6。

[0303] 表13给出使用肺炎咳嗽指数的肺炎归类结果。根据表12，LR-模型5提供针对咳嗽事件归类的最优归类结果(平均灵敏度和特异性分别为85.6±8%和80.1±12%)。然而，在肺炎归类任务中，LR-模型4比LR-模型5表现更好。其达到分别为86.2±9.6与84.3±15的平均灵敏度与特异性。

[0304]

特征模型		训练数据集（56 个对象）			测试数据集（25 个对象）		
	PCI_{th}	灵敏度	特异性	准确度	灵敏度	特异性	准确度
LR-模型 1	0.4	83.3±4	65.8±6	76.7±4	76.9±12	60.5±16	70.6±8.3
LR-模型 2	0.38	84±3.2	81±4	83±3	80.8±11.5	78.5±12	79.7±7.6
LR-模型 3	0.4	87±3.7	77±4.8	83.4±3	84.5±9	70.8±16	79±7.5
LR-模型 4	0.4	90±3	91.5±4	90.7±2.9	86.2±9.6	84.3±15	85.2±7
LR-模型 5	0.4	92.3±2.8	88±4	90.6±2.9	87.4±8.8	77.9±14	83.6±6.3

[0305] 表13.在仅选择的咳嗽特征被用于所述模型训练之后，针对基于肺炎咳嗽指数的肺炎诊断的结果。选择的咳嗽特征为Bl，FF1，FF2，Zcr，MFCC系数4、6和9。

[0306] IV.湿咳/非湿咳归类结果

[0307] A.训练与测试数据集

[0308] 分析了来自46个对象的总计178个咳嗽事件。数据库中的对象的男女比率为1:1。对象的平均年林为3岁1个月。在儿童疾病领域专攻慢性咳嗽哮喘以及其他呼吸系统疾病的具有大于20年临床与研究经验的儿科医师，在小心聆听之后，手动将178个咳嗽事件归类为

湿的和非湿的。我们将该手动归类视为‘参考标准’，将通过设计的LR模型的自动归类的结果与之进行比较。

[0309] 178个咳嗽事件中，82个为湿的并且96个为非湿，如通过参考标准归类的。我们使用咳嗽事件的70％(124个咳嗽事件)用于训练，并且30％(54个咳嗽事件)用于测试所述模式。使用随机数生成器随机选择训练和测试咳嗽事件。为了验证所述模型，我们从所述178个咳嗽事件生成了200个不同的训练与测试数据集。

[0310] B.归类结果

[0311] 当全部咳嗽特征均被用于训练所述模型时，针对测试数据集使用LR-模型的湿/非湿归类的平均灵敏度和特异性分别为74.8±9％和69.9±9.4％。当仅使用选择的咳嗽特征时，平均灵敏度和特异性值跳至79±9％和72.7±8.7％。在特征优化之后，从63个中选择共22个特征。所选择的特征为来自BSG、LogE和Kurt中每个各1个；2个来自NGS；3个来自ZCR；5个来自共振峰频率；以及9个来自MFCC。

[0312] 表14示出针对训练和测试数据集的平均灵敏度、特异性、准确度和kappa结果。

[0313] LR-模型与参考方法之间的kappa一致性在使用全部咳嗽特征时为0.45±0.12，并且在仅使用选择的咳嗽特征时为0.52±0.1。

[0314]

200 个数据集	使用全部咳嗽特征		使用选择的咳嗽特征	
	训练	测试	训练	测试
灵敏度	100±0.1	74.8±9	95.93±2	79±9
特异性	100±0	69.9±9.4	83.4±4	72.7±8.7
准确度	100±0.1	72.6±6.1	88.7±2.7	76.1±5.5
Kappa	1±0	0.45±0.12	0.77±0.05	0.52±0.1

[0315] 表14

[0316] 针对200个设计的LR模型，灵敏度、特异性、准确度和Kappa的平均±STD值

[0317] 图16示出使用200个训练和测试数据集的灵敏度和特异性的柱状图。

[0318] 表15示出针对200个中的最优LR-模型的列联表。其具有90％的灵敏度、80％的特异性以及0.71的高kappa一致性。

[0319]

		测试方法		
		湿	非湿	
参考标准	湿	27	3	30
	非湿	5	20	25
		32	23	55

[0320] 表15

[0321] 针对最优LR模型的列联表(LR＝149)

[0322] 现在将描述根据本发明一方面的另外的且优选的实施例的第二肺炎诊断方法。

[0323] V.2.第二肺炎诊断方法的材料与方法

[0324] 本发明的优选实施例的整体途径汇总于图17中。所述方法包括三个主要阶段:从对象对数据的采集,特征的形成以及模式归类器的训练,以及对技术的预期验证及性能评价。在2A-2C节,我们提供这些步骤的细节。

[0325] D.咳嗽声音数据库的形成

[0326] 针对该工作的临床数据采集环境为印度尼西亚Gadjah Mada大学Sardjito医院的呼吸医学单元。表16列出对象的纳入与排除准则。

[0327]

纳入准则	**排除准则**
-患者有胸部感染的症状:以下中的至少 2 个 -咳嗽 -有痰 -呼吸困难 -体温＞37.5° -知情同意	-不预期恢复的重病,例如晚期肺癌 -飞沫预防措施 -需要 NIV -无知情同意

[0328] 表16

[0329] 针对该研究招募疑似急性呼吸系统疾病(例如肺炎、细支气管炎、支气管炎、哮喘、鼻咽炎等等)的患者。使用由Gadjah Mada大学和昆士兰大学的人类伦理委员会认可的表单完成知情同意。在患者收治的前12小时内招募他们。

[0330] 儿科临床医师针对存在的症状评估患者。在标准纸上记述全部观察结果(例如咳嗽、流涕、发热、呼吸困难、腹泻等等)。数据库也包含常规人口统计信息和临床结果(例如胸部听诊、呼吸速率、血样定量、温度)以及实验室(例如血液、痰液分析)调查。在该讨论中针对肺炎诊断使用的参考标准是由医师在临床表现、实验室测试、胸部X光以及疾病的临床病程的基础上提供的整体诊断。为了使对儿童的辐射暴露最小化,仅在临床上疑似肺炎的对象上,或者存在对X射线的明确临床需要时,执行X射线。因此,在我们的数据库中并非所有对象均经历X射线成像。

[0331] 在初步医学评估之后,在呼吸病房的自然环境中进行声音记录。我们使用床边麦克风(Rode®NT7或Olympus®LSI 1,44.1kHz采样率),收集来自每位患者的6-8小时持续时间的声音数据记录。取决于所述患者的头的位置,从口部到所述麦克风的距离能够在40cm至70cm之间变化。发明人的目的是开发对抗强度变化的鲁棒技术,使得从口部到记录设备的距离在诊断中不具有显著作用。

[0332] 我们使用总计91名患者(63个肺炎和28个非肺炎对象),以开发并验证我们的技术。诸如细支气管炎、哮喘、支气管炎、咽炎、喉骨软化病的疾病被集中在非肺炎组内。在图18中给出对象的细节。我们支配的整体数据集被分成两个非重叠的组:模型开发数据集(D_{md})和预期验证数据集(D_{pv})。这两个数据集完全独立于彼此。集D_{md}和D_{pv}不共享任意咳嗽或任意对象。基于到医院的呼吸诊室的出现次序,将患者分派到每个组。D_{md}和D_{pv}分别包括N_{md}＝66和N_{pv}＝25个对象。

[0333] 来自每个对象的声音数据包括约6-8小时或连续的记录。在仔细聆听过程之后,手

动分割咳嗽声音。没有用于对咳嗽的自动识别的已被接受的方法，并且手动分析仍在临床工作以及研究文献中被用作黄金标准。图19示出(a)肺炎咳嗽、(c)哮喘咳嗽、(e)支气管炎咳嗽和(g)细支气管炎咳嗽波形的典型范例，以及它们对应的功率谱图。可见肺炎咳嗽与其他咳嗽之间在持续时间和功率谱图幅值中明显的特性差异。

[0334] E.特征提取与模型归类器设计

[0335] 我们针对该节描述的工作使用模型开发数据集D_{md}。设C_{md}为D_{md}中来自对象的咳嗽事件的总数目。这里采取的途径使用留一法模型构建与验证过程以形成特征，并优化模型参数。在该框架内，我们的方法能够被描述为四个主要过程步骤(参见图17)，如下所述的步骤1至步骤4。

[0336] 在第一步骤中，我们计算特征矩阵。在第二步骤中，我们使用来自步骤1的特征矩阵设计自动归类器，以将咳嗽声音归类成‘肺炎咳嗽’和‘非肺炎咳嗽’类别。在第三步骤中，我们选择最佳归类器，并且在第四步骤中，我们定义新的指数(被称作肺炎咳嗽指数(PCI))，以识别患有肺炎的患者。

[0337] 步骤1:咳嗽特征的提取与增大

[0338] 在该步骤中，我们的第一个目标是从咳嗽声音提取要被用于肺炎诊断算法的特征。来自D_{md}中的每个咳嗽事件的数学特征如下被计算：

[0339] [1].设x指代来自任意咳嗽事件的离散时间声音信号。

[0340] [2].将x分割成’n＝3’相等大小的非重叠子片段。在文献[19]中，临床医师和科学家均已描述咳嗽声音包括3个相位，(i)初始开放爆发，(ii)继之以嘈杂气流，以及最后的(iii)声门闭合。已显示，这些相位携带有对咳嗽的质量而言特异的不同显著信息。在此基础上，我们将每个咳嗽片段划分成n＝3个子片段。设x_i代表x的第i个子片段，其中i＝1,2,3,...,n。

[0341] [3].针对子片段x_i中的每个，计算如下特征：双频谱分(BGS)、非高斯性评分(NGS)、前四个共振峰频率(FF)、对数能量(LogE)、过零率(ZCR)、峰态(Kurt)以及十二个梅尔频率倒谱系数(MFCC)。注意，我们不利用MFCC的第0个系数，其代表信号x_i中的能量。

[0342] [4].针对D_{md}中的全部C_{md}咳嗽事件重复步骤(i)-(iii)。

[0343] 该过程得到针对每个子片段x_i大小为$C_{md}\times C_f$的候选咳嗽特征矩阵M_c。其中，C_f＝63代表基于咳嗽的特征，并且C_{md}为数据库D_{md}中的总咳嗽事件。

[0344] 在最简单形式的诊断算法中，我们将仅使用基于咳嗽的特征以诊断肺炎。然而，我们认识到一些简单临床测量结果的存在，其能够被用于以在复杂度方面最低的成本改进我们的算法。针对资源贫乏区域的WHO算法，表17，使用年龄和呼吸速率，并且其他研究者已经使用发热的存在。

[0345]

准则的筛查	咳嗽或呼吸困难
针对肺炎的准则	咳嗽和快速呼吸。针对快速呼吸的阈值取决于儿童的年龄 2-11 个月——每分钟 50 次呼吸 12-60 个月——每分钟 40 次呼吸
针对重度肺炎或非常严重的肺炎的准则	咳嗽和快速呼吸，具有一般危险体征，例如 • 较低的胸壁内收 • 平静儿童的喘鸣

[0346] 表17

[0347] 尽管这些中没有一个单独或组合地已在偏远地区获得期望的诊断性能，但这些测量结果具有增大咳嗽衍生特征的潜力。受使用年龄作为参数之一的WHO算法启示，我们使用以月为单位的年龄，作为我们的模型中的候选参数。我们还使用发热的存在或不存在作为二元变量。在WHO算法中，呼吸速率在诊断肺炎中被用作主要参数。在我们的工作中，我们提出一种新量度(参见(1))，我们将其称作呼吸指数(BrI)，以捕获肺炎中的呼吸速率升高。

[0348] $$BrI=\begin{Bmatrix}\text{如果年龄}\geq 60\text{个月，}\ BR-20\\ \text{否则，}\ BR-40\end{Bmatrix}\quad(1)$$

[0349] 在(1)中，BR为呼吸速率，并且年龄为以月为单位的患者的年龄。尽管发热为肺炎的常见症状，但其并不特异于肺炎。类似的观察适用于呼吸速率。表18示出候选特征集$F^c=\{C_f,f_1,f_2,\cdots,f_f\}$，其中，$C_f$代表咳嗽衍生的特征，并且剩下的指代在我们的模型中使用的增广特征。

[0350]

特征组合	年龄（月）	发热的存在	呼吸指数	咳嗽特征（#66）
F^1（全部咳嗽特征，C_f）				√
F^2（C_f+BrI）			√	√
F^3（C_f+发热）		√		√
F^4（C_f+年龄）	√			√
F^5（C_f+发热-BrI）		√	√	√
F^6（C_f+年龄-发热+BrI）	√	√	√	√

[0351] 表18-被用于形成2-B节[步骤1]中的特征矩阵的六种不同组合的特征。这里打勾指示该特征在所述模型设计中的纳入。呼吸指数是通过从所记录的呼吸速率减去呼吸阈值而计算的。

[0352] 我们的算法中的最终特征将来自候选特征的组。特征选择、模型开发和验证的细

节将在以下步骤2中描述。

[0353] [步骤2]特征选择与自动归类器设计-本发明人使用逻辑回归模型(LRM)作为优选的模式归类器。然而,本领域技术人员将认识到,其他归类器方法也是适用的。LRM为广义线性模型,其使用几个独立特征以估计分类事件的概率(因变量)。在该工作中,设因变量Y针对肺炎咳嗽等于“一”(Y=1)并且针对非肺炎咳嗽等于“零”($Y^{=}0$)。利用对肺炎的诊断归类从对象抽选的咳嗽事件被标记为肺炎咳嗽,反之亦然。使用回归函数推导模型,以估计概率Y,假设独立咳嗽特征(即$F^c=\{C_f,f_1,f_2,\cdots,f_f\}$)如下:

[0354] $$Prob(Y=1|_{f_1,f_2,f_3,\ldots,f_F})=\frac{e^z}{e^z+1} \quad (2)$$

[0355] $z=\beta_0+\beta_1f_1+\beta_2f_2+\ldots+\beta_nf_F$ (3)

[0356] 在(2)和(3)中,$f_1,f_2,\ldots f_F$为特征向量F^c(自变量)的元素,β_0被称作截距并且β_1、β_2等等被称作自变量的回归系数。为了从Y选择最佳决策阈值λ(如果Y在λ以上则所述咳嗽为肺炎否则为非肺炎),我们使用接收者工作曲线(ROC)分析。

[0357] 我们使用留一法交叉验证(LOV)技术用于LRM设计。如名称所表明的,LOV技术涉及使用来自除一个以外的全部患者的数据用于训练模型,以及来自一个患者的咳嗽事件用于验证模型。系统地重复该过程,使得D_{md}中的每个患者均有一次被用作验证数据。在该过程的最后,我们以N_{md}不同的LRM模型结束。为了评价所设计的N_{md} LRM的性能,计算性能量度,例如灵敏度、特异性、准确度、阳性预测值(PPV)、阴性预测值(NPV)、Cohen's Kappa(K)统计。

[0358] (i)特征选择:特征选择是选择相关特征的子集用于构建鲁棒学习模型的技术。理论上,最佳特征选择要求对特征的全部可能子集的穷举搜索。然而,针对大量特征这么做将是计算密集型且不实际的。因此,我们使用p值搜索了令人满意的特征的集。在LRM设计中,针对每个特征计算p值,并且其指示该特征多么显著地贡献于模型的开发。重要的特征具有低的p值。我们使用LRM的该性质,以在训练阶段期间,在模型中选择有助于归类的特征的合理组合(具有低p值的自变量)。计算针对F^c特征在S_{md} LRM上的平均p值。选择具有小于由p_{ths}给出的阈值的平均p值的特征。使$C_f{}^s$为从C_f选择的咳嗽特征的子集,并且$F^c{}_s$为通过利用选择的咳嗽特征增大的特征形成的候选特征集。

[0359] (ii)鲁棒LRM设计:一旦已知子集$F^c{}_s$,我们在另一次留一法验证过程之后,使用那些特征并再一次构建一套新的LRM。在该过程结束时,我们有N_{md}数量的LRM使用$F^c{}_s$作为输入特征集。

[0360] [步骤3]从N_{md} LRM选择良好模型-从将所选特征$F^c{}_s$用作输入特征的候选LRM,我们基于k均值聚类算法选择一个模型作为最好的。在k均值聚类算法中,目标是将d维空间里的q个数据点划分成k个聚类,使得在所述聚类内,距重心的平方距离得以最小化。我们手里的问题是从对我们可用的N_{md}个模型选择良好模型。为了这么做,我们将d维空间里N_{md}个模型划分成k=2个聚类,即高性能模型聚类和低性能模型聚类。我们设空间维度d等于模型参数加上三个性能量度(灵敏度、特异性和kappa)。然后从所述高性能模型的聚类,我们选择相对于重心具有最低均方差值的模型。设$\mathfrak{R}^{fc}$代表所选的LRM并且$\lambda_{\mathfrak{R}fc}$为针对特征的特定组合的对应的概率决策阈值(使用ROC曲线确定的值,使得归类器性能得以最大化)。

[0361] 一旦选择了$\mathfrak{R}^{fc}$,我们固定模型的全部参数,并完全终止所述训练过程。然后将模

型$\mathfrak{R}^{fc}$用作最优模型，以将每个个体咳嗽事件归类为‘肺炎咳嗽’或‘非肺炎咳嗽’组。

[0362] [步骤4]肺炎咳嗽指数-在该步骤中，针对D_{md}中的每个N_{md}＝66患者，我们使用以下定义计算肺炎咳嗽指数(PCI)。

[0363] 肺炎咳嗽指数(PCI)的定义：使‘P’为从患者记录并分析的咳嗽的总数目。并且使用在步骤3中选择的LRM$\mathfrak{R}^{fc}$将‘P’次咳嗽中的‘Q’次归类为肺炎咳嗽。然后如下计算针对所述患者的PCI指数

[0364] PCI＝Q/P(4)

[0365] 然后使用ROC分析，我们计算阈值PCI_{th}(在保持可接受的特异性的同时针对高灵敏度得到优化)，以将患者归类为两类，‘肺炎’和‘非肺炎’。

[0366] F.在Dpv上对选择的LRM$\mathfrak{R}^{fc}$和PCI的测试

[0367] 遵循在2-B节[步骤1]中描述的程序并使用来自数据集D_{pv}中的N_{pv}＝25名患者的咳嗽事件声音数据，计算大小为$C_{pv}\times C_f$的咳嗽事件特征矩阵M_c^{Dpv}。C_{pv}为D_{pv}中的总咳嗽事件，并且C_f＝63代表基于咳嗽的特征。通过利用选择的咳嗽特征C_f^s扩增临床特征，从M_{fc}^{sDpv}形成M_c^{Dpv}。使用在2-B节[步骤3]中选择的LRM$\mathfrak{R}^{fc}$，以将M_{fc}^{sDpv}中的数据归类为类别‘肺炎咳嗽’和‘非肺炎咳嗽’。然后使用(4)计算针对D_{pv}中的每名患者的PCI。将在2-B节[步骤4]中计算的PCI_{ths}应用于PCI，并在PCI＞PCI_{th}时将患者归类为‘肺炎’，否则为‘非肺炎’。

[0368] 将通过PCI的自动归类的结果与主治临床医师的结果进行比较，并计算性能量度。

[0369] D.数据库与临床诊断

[0370] 使用来自N＝91名患者(48名男性和43名女性)的声音记录。对象的平均年龄为3岁1个月(标准偏差3岁11个月)。对象的年龄范围从1个月至15岁。91个对象中，63个为肺炎患者并且28个为非肺炎患者。非肺炎患者患有诸如哮喘、支气管炎、鼻咽炎及其他(气喘、扁桃体咽炎、心脏病、异物吸入)的疾病。对65名患者执行胸部X射线(CXR)，以证实诊断。未对其进行CXR的26名患者中，八名已被临床诊断为肺炎，并且17名为非肺炎患者。图18给出数据库中的年龄和疾病种类的分布。

[0371] E.咳嗽声音特性

[0372] 图4示出肺炎咳嗽、哮喘咳嗽、支气管炎咳嗽和细支气管炎咳嗽波形，以及它们对应的功率谱图的典型范例。我们能够看到，在肺炎咳嗽与其他咳嗽之间，在持续时间和功率谱图幅值方面，有明显特征差异。肺炎咳嗽持续时间短，具有直到20kHz的广泛分布功率谱。相反，其他咳嗽持续时间长，具有直到15kHz的功率谱。肺炎咳嗽的平均持续时间(使用n＝401次咳嗽为0.26±0.7s)显著小于(2侧t检验，$p<0.005$，t＝-8.6786)其他咳嗽的平均持续时间(使用n＝198次咳嗽为0.32±0.08s)。咳嗽声音波形一般是干净的，具有高信噪比(SNR)。针对D_{md}的平均SNR为15.8±5.6db(最大＝28.05db并且最小＝2.08db)，并且针对D_{pv}的为16.7±5db(最大＝26.7db并且最小＝7.9db)。

[0373] F.基于WHO准则的肺炎诊断

[0374] 表19示出使用WHO准则的肺炎诊断与临床诊断的肺炎的列联表。

[0375]

WHO 准则	灵敏度	特异性	准确度
[咳嗽和/或呼吸困难]	86%	22%	
使用的针对呼吸急促的阈值为：			
<2 个月——60 BPM			

[0376]

2-11 个月——50 BPM			
12-60 个月——40			
>60 个月——20 BPM			

[0377] 针对社区环境中的肺炎诊断的WHO指南是针对2个月至5岁年龄组的儿童设计的[18]。“具有肺炎体征的小于2个月的婴儿被立即送到最近的医疗机构，因为他们有遭受严重疾病或死亡的高风险”。因此表19是使用我们数据库中的#68个对象生成的，其落入2个月至5岁的年龄范围。WHO准则在挑取临床证实的肺炎病例中达到83％的高灵敏度，然而呈现47％的差的特异性。

[0378] G.使用在训练/验证数据集上设计的模型的肺炎诊断

[0379] 从N＝91名患者，分析总计C＝599个咳嗽事件。平均分析每名患者6.6±2个咳嗽事件(每名患者最少＝2并且最多＝12)。在2-A节中，我们将N＝91名患者划分到两个数据集D_{md}(训练/验证数据集)和D_{pv}(预期研究数据集)。D_{md}有来自N_{md}＝66名患者的数据，具有C_{md}＝440个咳嗽事件(平均＝6.7±2，最少＝2，最大＝12)。D_{pv}有来自N_{pv}＝25名患者的数据，具有C_{pv}＝159个咳嗽事件(平均＝6.4±1，最小＝5，最大＝10)。

[0380] 特征矩阵：遵循在2-B节[步骤1]中给出的方法，我们计算特征矩阵M_c。我们使用n＝3以将每个咳嗽片段划分成3个子片段。设置n＝3从每个咳嗽事件给出63个数学特征，包括：(36MFCQ+(12FF)+(BSG、NGS、LogE、ZCR和Kurt每个3个)。利用咳嗽特征6扩增临床特征，创建咳嗽事件特征矩阵M_{fc}。遵循LOV技术，设计N_{md}＝66LRM(由于在D_{md}中N_{md}＝66名患者)。

[0381] 在特征选择之前使用LRM的肺炎咳嗽归类：针对F^1特征组合(仅咳嗽的数学特征)，在将咳嗽事件归类为肺炎和非肺炎咳嗽类别中，在66个LRM上的平均训练灵敏度和特异性为81±1％。验证灵敏度和特异性分别为63％和52％。当包括BrI为特征时，验证集的灵敏度和特异性分别跳至71％和55％。并且当伴随咳嗽的数学特征，全部特征(年龄、发热的存在和BrI)均被包括在训练LRM中时，灵敏度和特异性分别进一步增大到73％和62％。表20给出在使用特征的不同组合将咳嗽事件归类为肺炎与非肺炎咳嗽类别中，所述模型的性能。

[0382]

F^c	训练						验证					
	灵敏度	特异性	准确度	PPV	NPV	K	灵敏度	特异性	准确度	PPV	NPV	K
F^1(全部咳嗽特征，C_f)	81±1%	81±1%	81±1%	90±1%	68±1%	0.59±0.01	63%	52%	59%	73%	40%	0.13
F^2(C_f+BrI)	85±1%	85±1%	85±1%	92±1%	73±2%	0.67±0.02	71%	55%	66%	77%	48%	0.26
F^3(C_f+发热)	83±1%	83±1%	83±1%	91±1%	70±2%	0.62±0.01	68%	54%	63%	75%	45%	0.20
F^4(C_f+年龄)	85±1%	85±1%	85±1%	92±1%	73±1%	0.68±0.02	70%	56%	65%	77%	47%	0.24
F^5(C_f+发热+BrI)	87±1%	87±1%	87±1%	93±1%	76±1%	0.71±0.02	70%	59%	67%	78%	49%	0.28
F^6(C_f+年龄+发热+BrI)	91±1%	92±1%	91±1%	96±1%	83±2%	0.8±0.02	62%	62%	70%	80%	53%	0.34

[0383] 表20-当全部特征均被用于训练所述LRM模型时，所述模型在训练/验证数据集上，将咳

[0384] 嗽事件归类为‘肺炎咳嗽’和‘非肺炎咳嗽’类别中的性能。F^c代表不同的特征组合。

[0385] 在特征选择之后使用LRM的肺炎咳嗽归类：遵循在2-B节[步骤2]中描述的程序并使用p值，我们选择针对每个F^c_s的特征。表6给出所选择的咳嗽特征以及被用于选择的p_{ths}的细节。

[0386]

所用的 P_{ths}	选择的特征						
	BSH	NGS	FF	LogE	Kurt	ZCR	MFCC
0.23	2	1	3	2	0	3	19

[0387] 表21-在所述特征选择步骤之后，从C_f＝63个特征选择30个咳嗽特征。

[0388] 表22给出特征选择之后，在将咳嗽事件归类为肺炎和非肺炎咳嗽类别中所述模型的性能。

[0389]

F^c	训练						验证					
	灵敏度	特异性	准确	PPV	NPV	K	灵敏	特异	准确	PPV	NPV	K

[0390]

			度				度	性	度			
$F^1(C_f)$	79±1%	79±1%	79±1%	89±1%	64±1%	0.54±0.01	69%	64%	68%	80%	50%	0.31
$F^2(C_f+BrI)$	82±1%	83±1%	83±1%	91±1%	70±1%	0.60±0.02	72%	69%	71%	83%	54%	0.39
$F^3(C_f+发热)$	82±1%	82±1%	82±1%	91±1%	69±1%	0.61±0.02	72%	60%	68%	79%	51%	0.31
$F^4(C_f+年龄)$	82±1%	82±1%	82±1%	91±1%	69±2%	0.71±0.02	71%	67%	70%	82%	53%	0.36
$F^5(C_f+BrI+发热)$	86±1%	87±1%	86±1%	93±1%	75±1%	0.70±0.01	77%	74%	76%	86%	60%	0.49
$F^6(C_f+BrI+年龄+发热)$	89±1%	89±1%	89±1%	94±1%	79±1%	0.75±0.01	80%	73%	78%	86%	63%	0.51

[0391] 表22-当选择的特征被用于训练所述LRM模型时，所述模型在训练/验证数据集上，将咳嗽事件归类为‘肺炎咳嗽’和‘非肺炎咳嗽’类别中的性能。

[0392] 根据表22，我们看到针对全部F^c，在所述特征选择之后模型性能的全面改进。当仅来自咳嗽事件的数学特征被用于训练所述模型时，针对F^1_s的验证灵敏度和特异性分别增加到69％和64％，其中K＝0.31。针对F^6_s达到分别为80％和73％的最优灵敏度和特异性。

[0393] 基于PCI的归类：从使用来自D_{md}的数据针对每个F^c_s的N_{md}＝66个设计的LRM，使用如在2-B节[步骤3]讨论的k均值聚类方法选择鲁棒模型$\mathfrak{R}^{fc}$。针对在2-B节的步骤[4]中的使用，固定所选择的模型$\mathfrak{R}^{fc}$以及其全部参数，针对每名患者计算PCI指数。通过ROC分析选择PCI_{th}并将其应用于PCI，以将患者归类为‘肺炎’和‘非肺炎’。表23示出针对6种特征组合F^c_s的基于PCI的肺炎/非肺炎归类结果。

[0394]

选择的 LRM 模型	数据集 D_{md}						
	PCI_{ths}	灵敏度	特异性	准确度	PPV	NPV	K
$\mathfrak{R}^{f1}(C_f^s)$	0.5	93%	90.5%	92%	95%	86%	0.83

[0395]

$\mathfrak{R}^{f2}(C_f^s+BrI)$	0.5	93%	90.5%	92%	95%	86%	0.83
$\mathfrak{R}^{f3}(C_f^s+发热)$	0.5	91%	86%	89%	93%	82%	0.76
$\mathfrak{R}^{f4}(C_f^s+年龄)$	0.5	91%	91%	91%	95%	83%	0.80
$\mathfrak{R}^{f5}(C_f^s+BrI+发热)$	0.5	93%	90.5%	92%	95.5%	86%	0.83
$\mathfrak{R}^{f6}(C_f^s+BrI+年龄+发热)$	0.5	91%	90.5%	91%	95%	83%	0.8

[0396] 表23-在使用肺炎咳嗽指数诊断肺炎中，所选择的LRM$\mathfrak{R}^{fc}$在训练/验证数据集D_{md}上的性能。

[0397] 除F^3_s(其得到为86％的略低的特异性)以外，全部特征组合均达到大于90％的灵敏度和特异性。F^1_s(其仅使用咳嗽特征)具有93％的灵敏度，其中，K＝0.83。

[0398] H.使用所设计的模型在预期研究数据集上的肺炎诊断

[0399] 在包括N_{pv}＝25名患者和C_{pv}＝159个咳嗽事件的完整新数据集D_{pv}上测试在2-B节[步骤3]中选择的模型$\mathfrak{R}^{fc}$。针对在将咳嗽事件归类为肺炎和非肺炎咳嗽中的性能，以及在使用PCI将患有肺炎的患者与非肺炎分离开中的性能两者，测试模型。表24示出所选择的LRM$\mathfrak{R}^{fc}$在将咳嗽事件归类为肺炎和非肺炎咳嗽类别中的性能。

[0400]

选择的 LRM 模型	数据集 D_{md}					
	灵敏度	特异性	准确度	PPV	NPV	K
$\mathfrak{R}^{f1}(C_f^s)$	83%	58%	74%	79%	64%	0.42
$\mathfrak{R}^{f2}(C_f^s$+BrI)	82%	60%	74%	80%	64%	0.42
$\mathfrak{R}^{f3}(C_f^s$+发热)	82%	78%	81%	88%	69%	0.58
$\mathfrak{R}^{f4}(C_f^s$+年龄)	84%	75%	81%	86%	71%	0.57
$\mathfrak{R}^{f5}(C_f^s$+BrI+发热)	82%	69%	77%	83%	67%	0.50
$\mathfrak{R}^{f6}(C_f^s$+BrI+年龄+发热)	88%	85%	87%	92%	78%	0.71

[0401] 表24-在前瞻数据集D_{pv}上，所选择的LRM$\mathfrak{R}^{fc}$在将咳嗽事件归类为‘肺炎咳嗽’和‘非肺炎咳嗽’类别中的性能。

[0402] 表25示出模型在将患者归类为肺炎和非肺炎中的性能。

[0403]

选择的 LRM 模型	预期数据集 D_{pv}					
	灵敏度	特异性	准确度	PPV	NPV	K
$\mathfrak{R}^{f1}(C_f^s)$	94%	75%	88%	89%	86%	0.72
$\mathfrak{R}^{f2}(C_f^s$+BrI)	82%	63%	76%	82%	63%	0.45
$\mathfrak{R}^{f3}(C_f^s$+发热)	94%	87.5%	92%	94%	88%	0.82
$\mathfrak{R}^{f4}(C_f^s$+年龄)	94%	75%	88%	89%	86%	0.72
$\mathfrak{R}^{f5}(C_f^s$+BrI+发热)	88%	75%	84%	88%	75%	0.63
$\mathfrak{R}^{f6}(C_f^s$+BrI+年龄+发热)	82%	100%	88%	100%	73%	0.75

[0404] 表25-在预期数据集D_{pv}上，所选择的LRM$\mathfrak{R}^{fc}$在使用肺炎咳嗽指数诊断肺炎中的性

能。

[0405] 针对咳嗽归类，模型$\mathfrak{R}^{f6}$(所选择的咳嗽特征连同发热的存在、年龄和BrI)达到最优归类，其中，灵敏度＝88％并且特异性＝85％。仅使用咳嗽特征的$\mathfrak{R}^{f1}$具有分别为83％和58％的灵敏度和特异性。在分开肺炎与非肺炎患者中，表现前3的模型为$\mathfrak{R}^{f1}$、$\mathfrak{R}^{f3}$、$\mathfrak{R}^{f4}$和$\mathfrak{R}^{f6}$，全部具有大于0.7的kappa一致性，实质一致。$\mathfrak{R}^{f1}$达到分别为94％和75％的高灵敏度和特异性。

[0406] VI.4.讨论与结论-第二种肺炎诊断方法

[0407] 本发明前面的实施例提供了用于使用咳嗽声音诊断肺炎的自动程序。所述方法基于首先将个体咳嗽事件归类为‘肺炎咳嗽’和‘非肺炎咳嗽’类别，并然后在全部所记录的咳嗽事件上计算肺炎咳嗽指数(PCI)。在来自被诊断患有多种呼吸系统疾病的91名儿科患者的599个咳嗽事件上工作，我们示出能够在保持特异性＞85％的同时以灵敏度＞90％归类肺炎的方法。

[0408] 据本发明人认识，这是世界上首次尝试开发针对集中于咳嗽声音的肺炎诊断的客观模型。结果指示，在资源贫乏地区诊断肺炎时采取以咳嗽为中心的途径的可行性。此外，使用仅从咳嗽衍生的特征，我们以75％的特异性获得灵敏度>90％。能够利用非接触式器械记录咳嗽，并且我们的算法能够在普遍使用的平台(例如智能手机)上，以可由最少受训人员方便使用的方式，得以实施。所述设备将不要求与对象的物理接触，使得容易消毒且要求针对现场使用而言最少的操作者训练。所述技术，以最简单的版本，将要求5-10之间的咳嗽声音，并且将自动地且立即地提供诊断。这样的系统预期将成为在偏远地区肺炎诊断领域中颠覆性的新颖事物。

[0409] 所提出的技术的简单性以及在普遍使用的设备上潜在的低成本实施方式，使得我们的途径在长期监测中有价值。其也将在开发新疫苗以及针对儿童肺炎的管理策略中具有实质性的战略价值。在世界上资源有限地区中实施的新肺炎疫苗的临床试验需要可靠的工具，以测量干预的效力。对于评估靶向于大群体的任何新肺炎管理策略的有效性而言，同样如此。没有用于诊断肺炎的可现场部署的黄金标准，尤其是早期(不严重的)疾病，在其中甚至胸部X光也不起作用[4]。现有的WHO算法因其低的诊断性能而受限。尽管所述算法在挑取肺炎的潜在病例中扮演非常有用的角色，但这么做的代价，低的特异性，使其难以用于评估新干预的效度。

[0410] 使用逻辑回归的自动咳嗽归类器

[0411] 图20示出根据本发明的实施例的自动咳嗽归类过程的框图。其被划分成四个阶段，(A)数据采集过程(B)创建咳嗽声音数据库以及通过专家评分到干/湿类别的归类(C)自动归类器的设计(D)归类器在预期咳嗽声音数据库上的测试。在II-A节至II-D节中，我们描述所述方法的细节。

[0412] A.数据采集

[0413] 针对该工作的临床数据采集环境为印度尼西亚Gadjah Mada大学Sardjito医院的呼吸医学单元。表26列出对象的纳入与排除准则。

[0414]

纳入准则	排除准则
-患者有胸部感染的症状：以下中的至少 2 个 -咳嗽 -有痰 -增加的呼吸困难 -体温＞37.5° -知情同意	-不期待恢复的重病，例如晚期肺癌 -飞沫预防措施 -需要 NIV -无知情同意

[0415] 表26-针对对象的纳入与排除准则

[0416] 接洽满足所述纳入标准的全部患者。使用由Gadjah Mada大学和昆士兰大学的人类伦理委员会认可的表单完成知情同意。在患者收治的前12小时内招募他们。在初步医学评估之后，针对在呼吸病房的自然环境中，针对接下来的4-6小时，进行声音记录。

[0417] 使用两个系统进行声音记录，

[0418] 1.计算机数据采集系统——高保真度系统，具有专业品质的前置放大器和A/D转换器单元，(美国加利福尼亚M-Audio的Model Mobile-Pre USB)以及相匹配的一对具有超心型指向性束模式的低噪声麦克风(澳大利亚悉尼RODE的Model NT3)。使用Adobe audition软件第2版，以将声音数据记录到膝上型计算机上。—参见图1为该类型布置的图示。

[0419] 2.便携式记录系统——高端轻质便携式2-AA电池供电的录音机(Olympus LS-11)，带有两个精密电容式麦克风。

[0420] 在两种声音记录系统中，我们使用44.1khz的采样速率，具有16bit的分辨率(CD品质记录)。从麦克风年到患者的口部的标称距离为50cm，但由于患者移动而能够从40cm到70cm变化。针对每名患者，我们也接收最终诊断以及实验室和临床检查结果。

[0421] B.咳嗽声音数据集以及通过专业人员评分到湿和干的归类

[0422] 设N为使用其声音记录的患者的数目，并且C为来自N名患者的咳嗽事件的总数目。在对每名患者的声音数据的6-8小时的筛查之后，手动分割这C个咳嗽事件。没有被接受的方法用于对咳嗽事件的开始与结束的自动标记。手动标记仍被认为是黄金标准。在仔细聆听之后，手动标记全部咳嗽事件的开始与结束。

[0423] 我们将具有C个咳嗽事件的N名患者划分到两个数据集中，(i)DS1(模型设计数据集)以及(ii)DS2(预期研究数据集)。基于到医院的呼吸诊室的出现次序，将患者划分到DS1和DS2中。数据集DS1与DS2中的患者互不包含。

[0424] (i)DS1——包括来自N1名患者的C1个咳嗽事件。来自该数据集的咳嗽事件被用于设计最佳模型。

[0425] (ii)DS2——包括来自N2名患者的C2个咳嗽事件。来自该数据集的咳嗽事件被用于测试所设计的模型。来自DS2的咳嗽事件对模型设计的过程而言是盲的。

[0426] 在儿科呼吸系统疾病中具有15-20年经验的两名专家评分者然后将来自两个数据集的咳嗽事件评分为两个类别，干的或湿的。评分者对对象的历史和诊断是盲的。将该手动归类视为相对于比较自动归类的结果的参考标准。

[0427] C.咳嗽声音归类器的设计

[0428] 为了设计用于对咳嗽声音的自动归类的系统,我们使用来自DS1的咳嗽事件。设DS11为DS1中包含两名评分者对咳嗽声音的类别意见一致的那些咳嗽事件的子集。我们在DS11中有C11个咳嗽事件。使用DS11中的咳嗽事件以设计自动归类器模型。这是个三步骤过程。

[0429] [步骤1]咳嗽事件特征矩阵计算:在该步骤中,从C11咳嗽事件中的每个,计算包含'F'数学特征的特征向量,并形成大小为C11×F的咳嗽事件特征矩阵'M_{DS11}'。使用以下步骤以从咳嗽事件计算'F'特征。

[0430] (i)设x指代来自咳嗽事件的离散事件声音信号。

[0431] (ii)通过用其除以绝对最大值归一化x。

[0432] (iii)将x分割成'n'个大小相等的非重叠子片段。设x^i代表x的第i个子片段,其中i=1,2,3,...,n。

[0433] (iv)针对每个子片段计算以下特征,并形成包含F特征的特征向量:双频谱分(BGS)、非高斯性评分(NGS)、共振峰频率(FF)、音高(P)、对数能量(LogE)、过零率(ZCR)、峰态(Kurt)以及十二个梅尔频率倒谱系数(MFCC)。

[0434] (v)针对全部C11个咳嗽事件重复步骤(i)-(iii),并形成大小为C11×F的咳嗽事件特征矩阵M_{DS11}。

[0435] [步骤2]自动归类器设计:在本发明目前描述的方面的优选实施例中,我们使用逻辑回归模型(LRM)作为模式归类器。LRM为广义线性模型,其使用几个独立预测器,以估计分类事件的概率(因变量)。在该工作中,假设因变量Y针对湿咳等"一"(Y=1)并且针对干咳等于"零"(Y=0)。使用回归函数推导模型,以估计概率Y=1(即咳嗽事件属于'湿咳'种类),假设自变量(即F特征)如下:

[0436] $$Prob(Y=1|_{x1,x2,x3,\ldots,xn})=\frac{e^z}{e^z+1} \quad (1)$$

[0437] $z=\beta_0+\beta_1f_1+\beta_2f_2+\ldots+\beta_nf_F$ (2)

[0438] 在(1)和(2)中$f_1,f_2,\ldots.f_F$为特征向量(自变量)的元素,β_0被称作截距,并β_1,β_2等等被称作自变量的回归系数。为了从Y选择最佳决策阈值λ(如果Y在λ以上则所述咳嗽为湿的否则为干的),我们使用接收者工作曲线(ROC)分析。

[0439] 使用矩阵M_{DS11}中的数据(来自F自变量的C11观察结果)并采用留一法交叉验证(LOV)技术用于LRM设计。如名称所表明的,LOV技术涉及使用来自除一个以外的全部咳嗽事件的数据用于训练所述模型,并且一个咳嗽事件用于验证所述模型。系统地重复该过程C11次,使得DS11中的每个咳嗽事件均有一次被用作验证数据。这得到L_{C11}数目的LRM。

[0440] 为了评价所设计的L_{C11}的性能,计算性能量度,例如灵敏度、特异性、准确度、阳性预测值(PPV)、阴性预测值(NPV)、Cohen's Kappa(K)统计。关于如何解读K值请参见附录A2。

[0441] 设计逻辑回归模型(LRM),用于

[0442] (i)特征选择:特征选择是选择相关特征的子集用于构建鲁棒学习模型的技术。理论上,最佳特征选择要求对特征的全部可能子集的穷举搜索。然而,针对大量特征这么做将是计算密集型且不实际的。因此,我们使用p值搜索了令人满意的特征的集。在LRM设计中,针对每个特征计算p值,并且其指示该特征多么显著地贡献于所述模型的开发。重要的特征

具有低的p值。我们使用LRM的该性质，以在训练阶段期间，在所述模型中选择有助于所述归类的特征的合理组合(具有低p值的自变量)。计算针对‘F’特征在C11 LRM上的平均p值。选择具有小于p_{ths}的平均p值的特征。设F_s为从F选择的特征的子集。

[0443] (ii)鲁棒LRM设计：从M_{DS11}创建大小为C11×F_s的矩阵M'_{DS11}。矩阵M'_{DS11}为仅具有从DS11中的C11咳嗽事件选择的特征F_s的咳嗽事件特征矩阵。使用M'_{DS11}并采用LOV，再训练C11 LRM。

[0444] [步骤3]从L_{C11} LRM选择良好模型：从L_{C11} LRM我们选择一个模型作为最优的，使用k均值聚类算法9，以在预期研究数据集DS2上测试。在k均值聚类算法中，目标是将d维空间中的q个数据点划分成k个聚类，从使得在聚类内，距重心的平方距离得以最小化。

[0445] 我们手里的问题是从对我们可用的L_{C11}个模型选择良好模型。为了这么做，我们我们将d维空间里L_{C11}个模型划分成k＝2个聚类，即高性能模型模型聚类和低性能模型聚类。我们设空间维度d等于模型参数加上三个性能量度(灵敏度、特异性和kappa)。然后从所述高性能模型的聚类，我们选择相对于重心具有最低均方差值的模型。设$\mathfrak{R}$代表所选的LRM并且$\lambda_{\mathfrak{R}}$为对应的概率决策阈值(使用ROC曲线确定的值，使得所述归类器性能得以最大化)。一旦选择了$\mathfrak{R}^{fc}$，我们固定所述模型的全部参数，并使用其用于归类所述预期数据集DS2中的咳嗽声音。

[0446] D.对选择的LRM$\mathfrak{R}$的测试

[0447] 遵循在C节[步骤1]中描述的所述程序，并使用来自数据集DS2的咳嗽事件，计算大小为C2×F的咳嗽事件特征矩阵M_{DS2}。C2为DS2中的总咳嗽事件，并且‘F’为特征向量。通过仅选择鲁棒F_s特征，从M_{DS2}形成M'_{DS2}。使用选择的LRM$\mathfrak{R}$，以将M'_{DS2}中的数据归类为类别湿的或干的。从$\mathfrak{R}$的输出对湿/干类别的决策过程如下：

[0448] 设$\mathfrak{R}$到给定咳嗽输入的输出为$Y_{\mathfrak{R}}$，则如果$Y_{\mathfrak{R}} \geq \lambda_{\mathfrak{R}}$，所述咳嗽被归类为湿的否则为干的。

[0449] 将通过$\mathfrak{R}$的自动归类的结果与专家评分者的结果进行比较，并计算在C节[步骤2]中描述的性能量度。全部算法均使用软件编程语言MATLAB7.14.0.739(R2012a)版被开发。

[0450] 咳嗽声音数据集以及专家评分之间的一致性

[0451] 本发明人使用了来自N＝78名患者(41名男性和37名女性)的声音记录数据。对象的平均年龄为2岁11个月。对象的年龄范围从1个月至15岁，并且患有诸如哮喘、肺炎、支气管炎和鼻咽炎的疾病。表27给出患者的人口统计学与临床细节。

[0452]

性别	男性	41
	女性	37
年龄	新生儿	2
	＜12 个月	31

[0453]

	<60 个月	29
	>=60 个月	16
诊断	肺炎	34
	肺炎+其他	21
	支气管炎	8
	哮喘	3
	鼻咽炎	5
	哮喘+鼻咽炎	1
	其他	6

[0454] 表27-对象的人口统计学与临床细节

[0455] 从N＝78名患者,分析总计C＝536个咳嗽事件。每名患者平均有7个咳嗽事件得到分析(最小＝2并且最大＝13)。数据集DS1具有来自N1＝60名患者的C1＝385个咳嗽事件,并且数据集DS2具有来自N2＝18名患者的C2＝151个咳嗽事件。

[0456] 表28示出在将来自DS1和或DS2的咳嗽事件归类为两个类别湿和干时,两名评分者之间的列联表。

[0457]

数据集 DS11					数据集 DS2				
		评分者 1					评分者 1		
		湿	干				湿	干	
评分者 2	湿	82	55	60%	评分者 2	湿	47	23	67%
	干	20	228	92%		干	11	70	86.4%
		80.4%	80.6%	310			81%	75.3%	117

[0458] 表28-针对将咳嗽归类为湿/Y的人类评分者之间的列联表。针对DS1,K＝0.56并且%一致性＝

[0459] 80.5%,以及针对DS2,K＝0.54并且%一致性＝77.5。

[0460] 在385个咳嗽事件中的DS1中,评分者在被用于形成子集DS11的咳嗽事件的类别上一致C11＝310次(80.5%)。在数据集DS2中,它们在151个中一致117次(77.5%)。评分者1与评分者2之间的kappa一致性针对DS1为0.55并且针对DS2为0.54。在DS11的310个咳嗽事件中,82个属于湿类别并且228个属于干类别。然后将DS11咳嗽事件用于设计LRM模型,如在II-C节中描述的。

[0461] E.我们的数据库中的咳嗽声音特性

[0462] DS11中的干咳的均值持续时间为260+77ms(使用228次干咳计算的),并且湿咳的为238±54ms(使用82次湿咳计算的)。图21示出来自两名患者(id分别为#35和#38)的干咳波形和湿咳波形的典型范例。咳嗽声音波形一般是干净的,具有高信噪比(SNR)。针对DS11的均值信噪比为15.2±5.5db(最大＝28.65db并且最小＝2.9db),并且针对DS2的为18.6±4.5db(最大＝27.8db并且最小＝11.1db)。图22A和图22B为针对DS11和DS2中的咳嗽声音的SNR的柱状图。

[0463] 由人类操作者在聆听咳嗽声音之后仔细标记每次咳嗽的开始与结束，如在图21中所示。遵循在II-C节[步骤1]中给出的所述方法，我们计算特征矩阵M_{DS11}。我们使用n＝3以将每个咳嗽片段划分成3个子片段。在文献中，临床医师和科学家均已描述咳嗽声音包括3个相位，(i)初始开放爆发，(ii)继之以嘈杂气流，以及最后的(3)声门闭合[24,25]。已显示，这些相位携带有对咳嗽的品质而言不同的显著信息。在此基础上，我们将每个咳嗽片段划分成3个子片段。设置n＝3得到长度为66的包括以下特征的特征向量F(n×12MFCC)+(n×4FF)+([n×[BGS,NGS,P,LogE,Zcr,Kurt])。从C11＝310个咳嗽事件和E＝66个特征，创建咳嗽事件特征矩阵M_{DS11}。

[0464] F.使用LRM的自动归类

[0465] 训练阶段期间的特征矩阵及LRM性能:遵循LOV技术，设计L_{C11}＝310个LRM。

[0466] 针对310个LRM的均值训练灵敏度和特异性分别为92±1％和93±0.5％。针对这些模型的验证灵敏度和特异性分别为62％和84％。表4-(A)给出当全部F＝66个特征均被用于训练LRM时的详细归类结果。

[0467]

		灵敏度	特异性	准确度	PPV	NPV	K
（A）当全部特征均被用于形成 LRM 时							
评分者 1	训练	91.76±0.68 [91.69-91.84]	92.65±0.45 [92.6-92.7]	92.45±0.5 [92.36-92.47]	81.80±1 [81.68-81.91]	96.90±0.3 [96.87-96.93]	0.8125±0.1 [0.8112-0.8138]

与评分者2一致	验证	62	84	78	59	86	0.46
评分者1	训练	87.15±0.95 [86.9-87.4]	87.49±0.89 [87.26-87.72]	87.40±0.90 [87.17-87.63]	71.53±1.84 [71-72]	94.97±0.40 [94.87-95.07]	0.6977±0.02 [0.69-0.70]
湿/干类别	验证	53	78	71	47	82	0.3
评分者2	训练	81.96±1.01 [81.7-82.23]	82.24±0.97 [81.98-82.49]	82.14±0.98 [81.89-82.4]	71.83±1.37 [71.48-72.19]	89.18±0.78 [88.98-89.38]	0.6224±0.01 [0.6173-0.6276]
湿/干类别	验证	45	67	59	43	69	0.12

[0468]

（B）当选择的全部特征被用于形成LRM时							
评分者1	训练	87.36±0.61 [87.29-87.43]	87.82±0.43 [87.77-87.87]	87.70±0.46 [87.65-87.75]	72.07±0.87 [71.98-72.17]	95.07±0.25 [95.05-95.10]	0.7041±0.01 [0.7029-0.7053]
与评分者2一致	验证	81	83	82	63	92	0.58
评分者1	训练	82.75±0.57 [82.60-82.89]	83.06±0.52 [82.92-83.19]	82.98±0.52 [82.84-83.11]	63.78±1.18 [63.47-64.08]	93.03±0.27 [92.96-93.10]	0.60±0.01 [0.59-0.60]
湿/干类别	验证	76	79	78	57	90	0.5
评分者2	训练	75.66±0.57 [7.5.51-75.81]	75.92±0.58 [75.77-76.07]	75.83±0.57 [75.68-75.98]	63.44±0.96 [63.19-63.69]	84.95±0.61 [84.79-85.11]	0.49±0.01 [0.4916-0.4975]

[0469]

湿/干类别	验证	72	73	72	59	82	0.43

[0470] 表29:特征选择之前与之后的LRM性能。表中提供的统计数据为均值±标准偏差。

在底部提供训练数据集的均值的95％置信区间。针对评分者1和评分者2，样本大小为来自数据集DS1中的N1＝60名患者的C1＝385个咳嗽事件。在385个咳嗽事件中，评分者在C11-310个咳嗽事件上具有湿/干一致。

[0471] 遵循在II-B节[步骤2]中描述的所述过程，并使用p_{ths}＝0.06，我们选择了F_s＝31个特征。图23示出与在C11＝310个LRM上计算的F＝66个特征相关联的均值‘p值’。选择具有小于p_{ths}＝0.06的均值‘p值’的全部特征。所选择的特征为来自双频谱分、峰态及过零率数的每个1个，来自非高斯性频分和对数能量的每个2个，来自共振峰频率的5个，以及来自梅尔频率倒谱系数的19个。表30给出用于设计最终LRM选择的特征的细节。

[0472]

特征	BSG			NGS			FF1			FF2			FF3			FF4		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
选择的		√			√	√	√	√			√		√		√			
特征	音高			LogE			峰态			ZCR			MRFF0			MFCC1		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
选择的				√	√			√		√				√		√		
特征	MFCC2			MFCC3			MFCC4			MFCC5			MFCC6			MFCC7		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
选择的			√	√		√		√	√	√	√				√		√	√
特征	MFCC8			MFCC8			MFCC10			MFCC11								
	1	2	3	1	2	3	1	2	3	1	2	3						
选择度	√	√	√		√			√		√		√						

[0473] 表30：通过在II-C节[步骤2]使用n＝3，从每个咳嗽片段计算F＝66个特征。‘√’指示在II-C节[步骤2]选择该特征用于设计所述最终模型。

[0474] 根据该表，基于MFCC的特征是最主要的。31个选择的特征中，19个特征是由不同的MFCC分量贡献的。在MFCC共振峰频率之后，利用5个特征进行第二最主要的贡献。而且除了第4个共振峰频率和基于音高的特征(其被完全忽略)以外，全部其他特征贡献于来自至少一个子片段的用于构建最终LRM模型的特征。

[0475] 当仅选择的特征F_s被用于再训练LRM时，均值训练灵敏度和特异性分别被记录为87±1％和88±0.5％，并且验证灵敏度和特异性为81％和83％。LRM与评分者之间的验证kappa一致性在全部所述特征均被用于训练LRM时为0.46，并且当仅使用选择的特征时增加到0.58。表29-(B)给出在特征选择之后的详细训练与验证结果。

[0476] LRM($\mathfrak{R}$)的选择：从使用来自DS11的数据设计的L_{C11}＝310个LRM，使用如在II-C节[步骤3]中讨论的k均值聚类方法，选择最佳模型$\mathfrak{R}$。基于模型参数和性能量度，将模型聚类为两组，高性能模型和低性能模型。310个模型中，202个被聚类到高性能模型组中并且108个被聚类到低性能模型组中。LRM模型#26具有相对于高性能模型的重心的最低均方差值。

选择该模型ℜ并固定其全部参数用于以后使用。在预期数据集DS2上测试ℜ。

[0477] ℜ在前瞻数据集DS2上的性能：表31给出ℜ相对于专家评分者的归类结果。

[0478]

当在来自 DS2 的全部咳嗽事件（151）上测试时相对于个体评分者						
	灵敏度	特异性	准确度	PPV	NPV	K
评分者 1	77.5%	76%	76%	54%	90%	0.47
评分者 2	75%	64%	67%	43%	87%	0.31
仅在评分者 1 和评分者 2 两者关于类别一致的那些事件上测试						
	84%	76%	78%	55%	93%	0.51

[0479] 表31-ℜ在数据集DS2预期研究数据集上的性能

[0480] 当评分者1，湿/干归类被用作参考标准时，ℜ具有77.5％的灵敏度、76％的特异性和0.47的kappa一致性。针对评分者2，结果为灵敏度75％、特异性64％和kappa 0.31。当仅在其中评分者1和评分者2对归类意见一致的那些事件(117个咳嗽事件)上测试模型ℜ时，灵敏度跳至84％并且kappa值跳至0.51。表32示出列联表。

[0481]

		评分者		
		W	nW	
LRM	W	26	21	55%
	nW	5	65	93%
		84%	76%	78%

[0482] 表32：针对在数据集DS2上测试的选择的LRM的列联表K＝0.51。

[0483] 当针对年龄和性别匹配时的LRM结果：表33示出当针对年龄和性别匹配时，LRM在DS11和DS2上的性能。

[0484]

（A）						
针对数据集 DS11 的验证结果。全部所述特征均被用于训练所述 LRM						
	灵敏度	特异性	准确度	PPV	NPV	K
年龄<=60 个月，男性（#121 咳嗽事件）	59%	83%	76%	57%	84%	0.41
年龄<=60 个月，女性（#145 咳嗽事件）	58%	88%	80%	63%	85%	0.47
年龄>60 个月，男性（#20 咳嗽事件）	89%	64%	75%	67%	87.5%	0.51
年龄>60 个月，女性（#24 咳嗽事件）	100%	83%	83%	20%	100%	0.28
针对数据集 DS11 的验证结果。选择的特征被用于训练所述 LRM						
	灵敏度	特异性	准确度	PPV	NPV	K
年龄<=60 个月，男性（#121 咳嗽事件）	73.5%	78%	77%	57%	88%	0.47
年龄<=60 个月，女性（#145 咳嗽事件）	84%	87%	86%	70%	94%	0.67
年龄>60 个月，男性（#20 咳嗽事件）	89%	64%	75%	67%	87.5%	0.51
年龄>60 个月，女性（#24 咳嗽事件）	100%	91%	92%	33%	100%	0.47

[0485]

（B）						
预期研究数据集 DS2						
	灵敏度	特异性	准确度	PPV	NPV	K
年龄<=60 个月，男性（#36 咳嗽事件）	92%	87.5%	89%	78.5%	95%	0.76
年龄<=60 个月，女性（#27 咳嗽事件）	87.5%	95%	92.5%	87.5%	95%	0.82
年龄>60 个月，男性（#30 咳嗽事件）	50%	54%	53%	14%	87.5%	0.02
年龄>60 个月，女性（#24 咳嗽事件）	86%	71%	75%	54.5%	92%	0.48

[0486] 表33-针对年龄和或性别匹配的数据集DS11和前瞻数据集DS2的LRM验证结果

[0487] 由于数据的有限可用性，我们仅考虑4个分类：(i)年龄<＝60个月的男性，(ii)年龄<＝60个月的女性，(iii)年龄>60个月的男性和(iv)年龄>60个月的女性。根据该表，在模型设计阶段期间，与不考虑分类时相比较，在四个分类上的模型验证性能大体上未见显著差异，表29和表33(A)。类似于此，在预期数据集DS2上，除了在第三个分类中(年龄>60的男性)以外(在其中的性能非常差)，选择的模型$\mathfrak{R}$在全部分类上均表现良好(表31和表33(B))。

[0488] 本发明的实施例包含用于将咳嗽声音归类为湿和干种类的自动客观的方法。具发明人所知，这是用于开发儿科咳嗽声音的干/湿归类的客观技术的首次尝试，尤其是在诸如肺炎的疾病中。相比较在它们的描述性分析中使用不多于30次咳嗽的现有工作，本文中呈现的结果基于来自78名对象的536个咳嗽事件。出于这些原因，尚无其他工作可用于与所述结果直接比较。

[0489] 用于对我们的技术的评估的参考方法是由来自不同国家的两名儿科呼吸系统医师对咳嗽声音到湿/干类别的主观归类。这些评分者对对象的实际临床诊断是盲的。在逐个事件的咳嗽归类中，两名专家在中等水平上彼此意见一致(κ＝0.54的kappa值)。在5中，针对湿咳/干咳的临床医师间一致性报道为κ＝0.88。然而应注意，在[5]中，临床医师是在患者水平上而非在个体咳嗽水平上评估咳嗽的湿度。当我们在患者水平上计算评分者之间的一致性时，kappa值增大到κ＝0.66(基本一致)。这些数字进一步说明干/湿归类的主观性。

[0490] 我们的归类器技术是使用仅两名评分者达成共识的事件，在来自训练集(集DS1)的咳嗽上进行训练的。作为训练过程的输出，我们识别良好逻辑回归模型($\mathfrak{R}$)并固定其参数。然后以几种不同的方式在预期集(集DS2)上测试所述模型。当我们相对于在DS2内的共识事件测试$\mathfrak{R}$时，达到最高灵敏度和特异性(84％和76％)的归类。有趣的是我们注意到，这些数字一致地高于我们通过相对于每名评分者的个体归类结果进行测试所得到的。

[0491] 我们的方法的另一突出特征在于，当评分者共识数据被用作基准时，其具有高阴性预测值(NPV＝93％)。这意味着如果所述模型将咳嗽归类为非湿的(干的)，则最有可能的

是两名专家评分者将独立地达到相同的结论。然而，与人类评分者相比，我们的方法的阳性预测值更低(PPV＝55％)。因此，咳嗽中被所述模型归类为湿的相当大部分结果是被人类评分者达成共识地归类为干的。该现象似乎得到Chang等人[5]呈现的结果的解释，他们发现专家人类评分者低评分了湿咳。在[5]中，他们系统地比较了专家临床医师的主观干/湿归类与气道粘液的支气管镜指示。他们报道，临床医师对干咳的归类不必然指示不存在分泌物。气道中的特定情形，例如小量的分密度，可能未在咳嗽声音中以足以被人类观察者检测到的幅值得到反映。在我们的方法中发现的较低PPV值的可能的原因之一可能是黄金标准(人类评分者，用于生成我们的性能统计数据)的该弱点。未来需要相对于支气管镜发现小心验证该假设。

[0492] 正确检测气道粘液的能力能够在化脓性肺部疾病的管理中尤为重要[3,5]。咳嗽是诸如肺炎、支气管炎和细支气管炎的疾病中的早期症状。对该症状的准确评估是在诊断急性疾病或监测慢性症状与处置效力中是关键因素。已知在儿童中，湿咳更可能与下呼吸道感染相关联[3]。作为检测气道粘液的方法，甚至在专家临床医师的方面，对湿咳的主观归类也具有低灵敏度。用于对干/湿咳的归类的准确客观的技术目前在商业或研究水品上均仍是欠缺的。据我们所知，该工作是世界上用于开发这样的技术的首次尝试。

[0493] 我们提供的第一种途经，用于使对咳嗽的干-湿归类自动化。可以通过系统性地优化参数以及精细调谐归类器的训练过程，改善本文中已呈现的结果。已讨论的启发式模型选择过程使得所报道的结果悲观估计。本发明人也相信，能够改进特征设置，并且能够进一步增大所述方法的归类准确度。然而在优化尝试之前，我们需要解决的问题是改进在临床诊断中使用的‘黄金标准’。小心控制的支气管镜检研究将最适合作为黄金标准。

[0494] 对该研究的另一个可能的限制因素是咳嗽声音数据库向干咳的偏差;几乎70％的咳嗽声音为干的，如由专家人类评分者所认识的。然而，利用所有这些因素，本发明的实施例目前能以高灵敏度(84％)和特异性(76％)以及与专家人类评分者的良好一致性(κ＝0.51)归类湿咳与干咳。

[0495] 鉴于以上，将认识到，根据本文描述的本发明的实施例的方法能以高准确度和与儿科医师的良好一致性，将咳嗽声音归类为干的和湿的类别。据发明人所知，这是以在显著大的咳嗽样本上的完整训练与测试结果呈现的，首个已知用于湿/干归类的方法。其也是用于使在患有一系列呼吸系统感染疾病的儿科群体中的湿/干归类自动化的首次尝试。其有潜力成为用于长期咳嗽监测以及在对处置效力的评估中或在对下呼吸道感染的表征中的有用临床工具。其在其中需要咳嗽质量(干/湿)的以小时为基础的时间模式的临床或研究中将是非常有用的。

[0496] 可以与其他可以能够检测肺中实变和粘液的潜在技术一起实施本文描述的方法，所述技术例如为微波成像和超声成像。

[0497] 对在解释前述实施例中使用的术语中的一些的解释如下：

[0498] [A2].Kappa统计被广泛用于在其中应比较两种技术之间的一致性的情形。

[0499] 以下是用于解读Kappa值的指导。

[0500]

Kappa	解读
< 0	小于期望一致性
0.01 - 0.20	极低的一致性
0.21 - 0.40	一般的一致性
0.41 - 0.60	中等的一致性
0.61 - 0.80	高度的一致性
0.81 - 1	几乎完全的一致性

[0501]

[0502] [A3].被用于评价LRM的性能的统计度量的定义。

[0503] 真阳性(TP)-湿咳被LRM正确地识别为‘湿’。

[0504] 假阳性(FP)-干咳被LRM不正确地识别为‘湿’。

[0505] 真阴性(TN)-干咳被LRM正确地识别为‘干’。

[0506] 假阴性(FN)-湿咳被LRM不正确地识别为‘干’。

[0507] $$准确性=\frac{TP+TN}{TP+FP+TN+FN} \quad 10$$

[0508] $$灵敏度=\frac{TP}{TP+FN} \quad 11$$

[0509] $$特异性=\frac{TN}{TN+FP} \quad 12$$

[0510] 本发明一个方面的优选实施例包括一种方法,其能够通过限定咳嗽片段的开始和结束,从记录提取咳嗽声音。从声音信号计算一组不同特征,并将其用作对决策引擎的输入,例如基于自适应神经网络的模式识别算法或逻辑回归模型。在一个实施例中,所提出的方法达到约为95%的分割灵敏度和特异性。该方法能够被用作咳嗽分析系统的前端;因此能够自动分析来自记录中较大量咳嗽声音的定量和定性信息。

[0511] 根据本发明另外一方面的优选实施例,提供一种操作计算设备的方法,以处理患者的声音记录用于对患者的特定疾病状态的诊断。例如,如在本文中详细讨论的,所述疾病状态可以包括肺炎。

[0512] 参考文献

[0513] [1]A.J.Hotaling和G.T.Moynihan,"Cough,"Pediatirc Otolaryngology,第二卷,ed Philadelphia:Saunders,2003,第1395-1404页。

[0514] [2]R.E.Black等,"Global,regional,and national causes of child mortality in2008:a systematic analysis,"The Lancet,第375卷,第1969-1987页,2010。

[0515] [3]Igor Rudan等,"Epidemiology and etiology of childhood pneumonia," Bulletin of the World Health Organization,第86卷,第408-416页,2008。

[0516] [4]WHO,"WHO-recommended standards for surveillance of selected-preventable diseases,"Geneva,WHO/V&B/03.01,2003。

[0517] [5]S.Barry等,"The automatic recognition and counting of cough,"Cough,第2卷,第8页,2006。

[0518] [6]S.Matos等,"An Automated System for 24-h Monitoring of Cough Frequency:The Leicester Cough Monitor,"Biomedical Engineering,IEEE Transactions on,第54卷,第1472-1479页,2007。

[0519] [7]PDACortex,(于2011年10月12日访问),LifeShirt a new era in ambulatory monitoring,可在http://www.pdacortex.corrJVivoMetrics.htm获得。

[0520] [8]J.Thomas,(于2011年10月12日访问),VitaloJAK Cough Monitor.可在http://www.trustech.org.uk/case-study/the-vitaloiak/获得。

[0521] [9]KarmelSonix,(于2011年10月12日访问),PulmoTrack.可在http://www.karmelsonix.com/solution-pulmotrack.html获得

[0522] [10]M.A.Coyle等,"Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease,"Cough,第1卷,第3页,2005。

[0523] [11]E.Vizel等,"Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions,"Cough,第6卷,第3页,2010。

[0524] [12]K.McGuinness等,"Automated cough detection:a novel approach [abstract],"Am J Resp Crit Care Med,第175页:A381,2007。

[0525] [13]J.Smith和A.Woodcock,"New Developments in the Objective Assessment of Cough,"Lung,第186卷,第48-54页,2008。

[0526] [14]S.Ferrari等,"Cough sound analysis to identify respiratory infection in pigs,"Computers and Electronics in Agriculture,第64卷,第318-325页,2008。

[0527] [15]R.Martin,"Noise power spectral density estimation based on optimal smoothing and minimum statistics,"Speech and Audio Processing,IEEE Transactions on,第9卷,第504-512页,2001。

[0528] [16]S..Greenberg等,Speech processing in auditory system.New York:Springer,2004。

[0529] [17]S.Chatterjee和W.B.Kleijn,"Auditory model based modified MFCC features,"in Acoustics Speech and Signal Processing(ICASSP),2010IEEE International Conference on,2010,第4590-4593页。

[0530] [18]W.D.Duckitt等,"Automatic detection,segmentation and assessment of snoring from ambient acoustic data,"Physiological Measurement,第27卷,第1047页,2006。

[0531] [19]E.Goldshtein等,"Automatic Detection of Obstructive Sleep Apnea Using Speech Signals,"Biomedical Engineering,IEEE Transactions on,第58卷,第1373-1382页,2011。

[0532] [20]S.K.Asela等,"Multi-feature snore sound analysis in obstructive sleep apnea-hypopnea syndrome,"Physiological Measurement,第32卷,第83页,2011。

[0533] [21]U.R.Abeyratne等,"Multi-parametric analysis of snore sounds for

the community screening of sleep apnea with non-gaussianity index,"美国专利,2010。

[0534] [22]J.E.Markel和A.H.Gray,Linear Prediction of Speech:Springer-VerlagNew York,Inc.,1982。

[0535] [23]H.Ghaemmaghami等,"Normal probability testing of snore signals for diagnosis of obstructive sleep apnea,"in Engineering in Medicine and Biology Society,2009.EMBC 2009.Annual International Conference of the IEEE,2009,第5551-5554页。

[0536] [24]S.El Safty和A.El-Zonkoly,"Applying wavelet entropy principle in fault classification,"International Journal of Electrical Power&Energy Systems,第31卷,第604-607页,2008。

[0537] [25]G.M.Foody,"Using prior knowledge in artificial neural network classification with a minimal training set,"International Journal of Remote Sensing,第16卷,第301-312页,1995/01/01 1995。

[0538] [26]A.Waibel等,"Phoneme recognition using time-delay neural networks,"Acoustics,Speech and Signal Processing,IEEE Transactions on,第37卷,第328-339页,1989。

[0539] [27]M.Riedmiller和H.Braun,"A direct adaptive method for faster backpropagation learning:the RPROP algorithm,"in Neural Networks,1993.,IEEE International Conference on,1993,第586-591页第1卷。

[0540] [28]V.Tyagi等,"A variable-scale piecewise stationary spectral analysis technique applied to ASR,"in Machine learning for multimodal interaction,Edinburgh,2005,第274-284页。

[0541] [29]1.Ghaemmaghami,H.,U.Abeyratne和C.Hukins,Normal probability testing of snore signals for diagnosis of obstructive sleep apnea.2009:IEEE。

[0542] [30].Ng,A.K.等,Could Formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?Sleep medicine,2008.9(8):第894-898页。

[0543] [31]Oppenheim,A.V.,R.W.Schafer和J.R.Buck,Discrete-time signal processing.第1999卷.1989:Prentice hall Englewood Cliffs,NJ:。

[0544] [32]Zheng,F.,G.Zhang和Z.Song,Comparison of different implementations of MFCC.Journal of Computer Science and Technology,2001.16(6):第582-589页。

[0545] [33]Abeyratne,U,Blind reconstruction of non-minimum-phase systems from 1-D oblique slices ofbispectrum.1999:IET。

[0546] [34]Mendel,J.M.,Tutorial on higher-order statistics(spectra)in signal processing and system theory:Theoretical results and some applications.Proceedings of the IEEE,1991.79(3):第278-305页。

[0547] 3.Tessa Wardlaw,E.W.Johansson和M.Hodge,Pneumonia:The forgotten killer of children.UNICEF/WHO,2006。

[0548] 4.Berman,S.,E.Simoes和C.Lanata,Respiratory rate and Pneumonia in infancy.Archives of disease in childhood,1991.66(1):第81-84页。

[0549] 5.WHO,Childhood disease,WHO。

[0550] 6.Cardoso,M.R.A.等,Adding fever to WHO criteria for diagnosing Pneumonia enhances the ability to identify Pneumonia cases among wheezing children.Archives of disease in childhood,2011.96(1):第58页。

[0551] 7.Harari,M.等,Clinical signs of pneumonia in children.The Lancet,1991.338(8772):第928-930页。

[0552] 8.Lozano,J.等,Clinical predictors of acute radiological pneumonia and hypoxaemia at high altitude.Archives of disease in childhood,1994.71(4):第323-327页。

[0553] 9.Mulholland,E.等,Standardized diagnosis of pneumonia in developing countries.The Pediatric infectious disease journal,1992.11(2):第77页。

[0554] 10.Palafox,M等,Diagnostic value of tachypnoea in pneumonia defined radiologically.Archives of disease in childhood,2000.82(1):第41-45页。

[0555] 本发明的任意实施例意为仅是说明性的并且不意为限制本发明。因此,应认识到,能够对所描述的任意实施例进行各种其他改变和修改,而不偏离本发明的精神和范围。

[0556] 在本说明书和权利要求书中,要以开放式的意义将词语“包括”及其相关和衍生词解释为包含所记载整体的每个,但不排除包含一个或多个另外的整体。

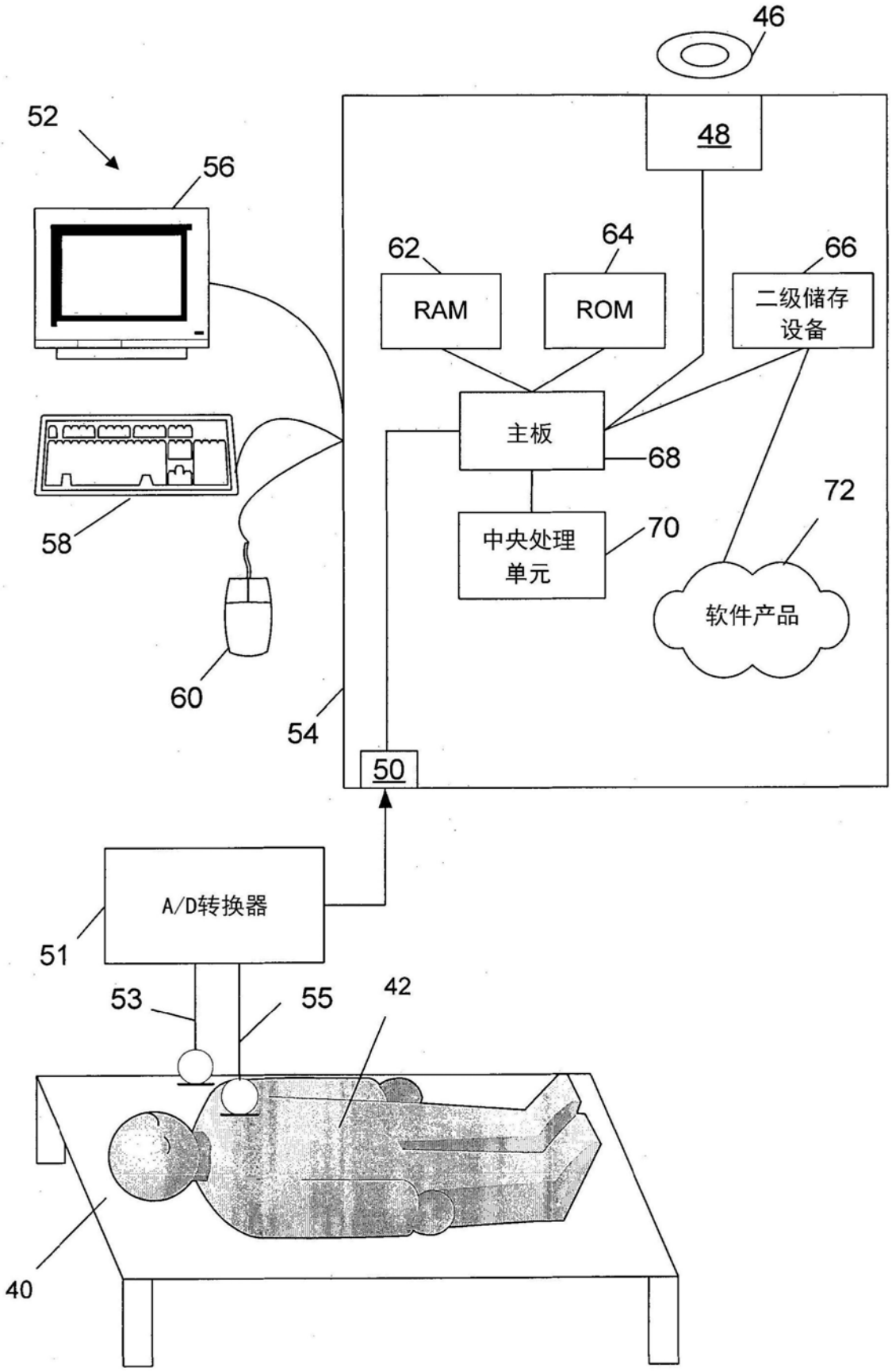

46
48
52
56
62
RAM
64
ROM
66
二级储存
设备
主板
68
72
中央处理
单元
70
软件产品
58
60
54
50
A/D转换器
51
53
55
42
40

图1

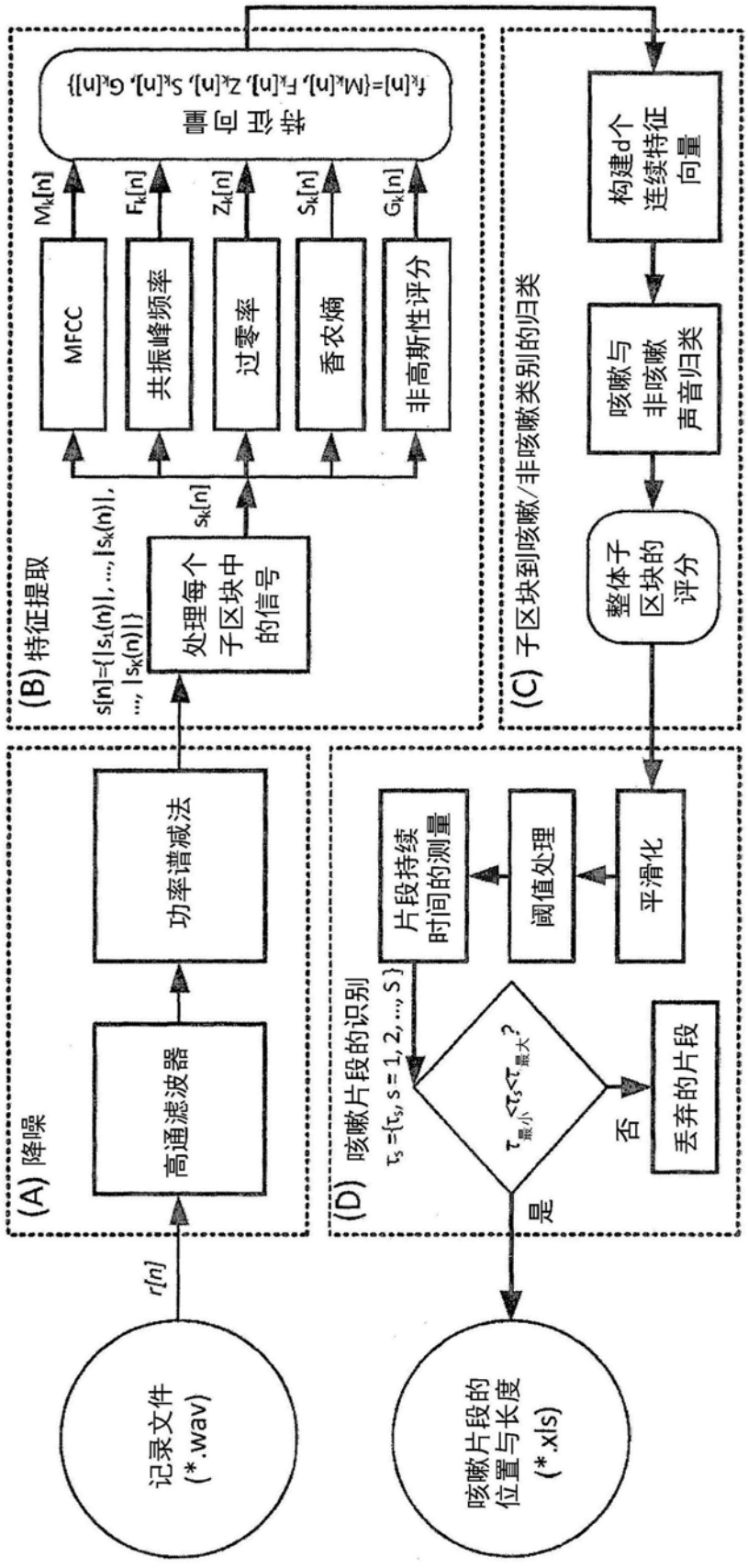

记录文件
(*.wav)
r[n]
(A) 降噪
高通滤波器
功率谱减法
(B) 特征提取
s[n]={|s1(n)|, ..., |sK(n)|, ..., |sK(n)|}
处理每个子区块中的信号
sk[n]
MFCC
共振峰频率
过零率
香农熵
非高斯性评分
Mk[n]
Fk[n]
Zk[n]
Sk[n]
Gk[n]
特征向量
fk[n]={Σk[n], Fk[n], Zk[n], Sk[n], Gk[n]}
(C) 子区块到咳嗽/非咳嗽类别的归类
构建d个连续特征向量
咳嗽与非咳嗽声音归类
整体子区块的评分
(D) 咳嗽片段的识别
τs={τs, s = 1, 2, ..., S}
片段持续时间的测量
阈值处理
平滑化
τ最小<τs<τ最大?
否
丢弃的片段
是
咳嗽片段的位置与长度
(*.xls)

图2

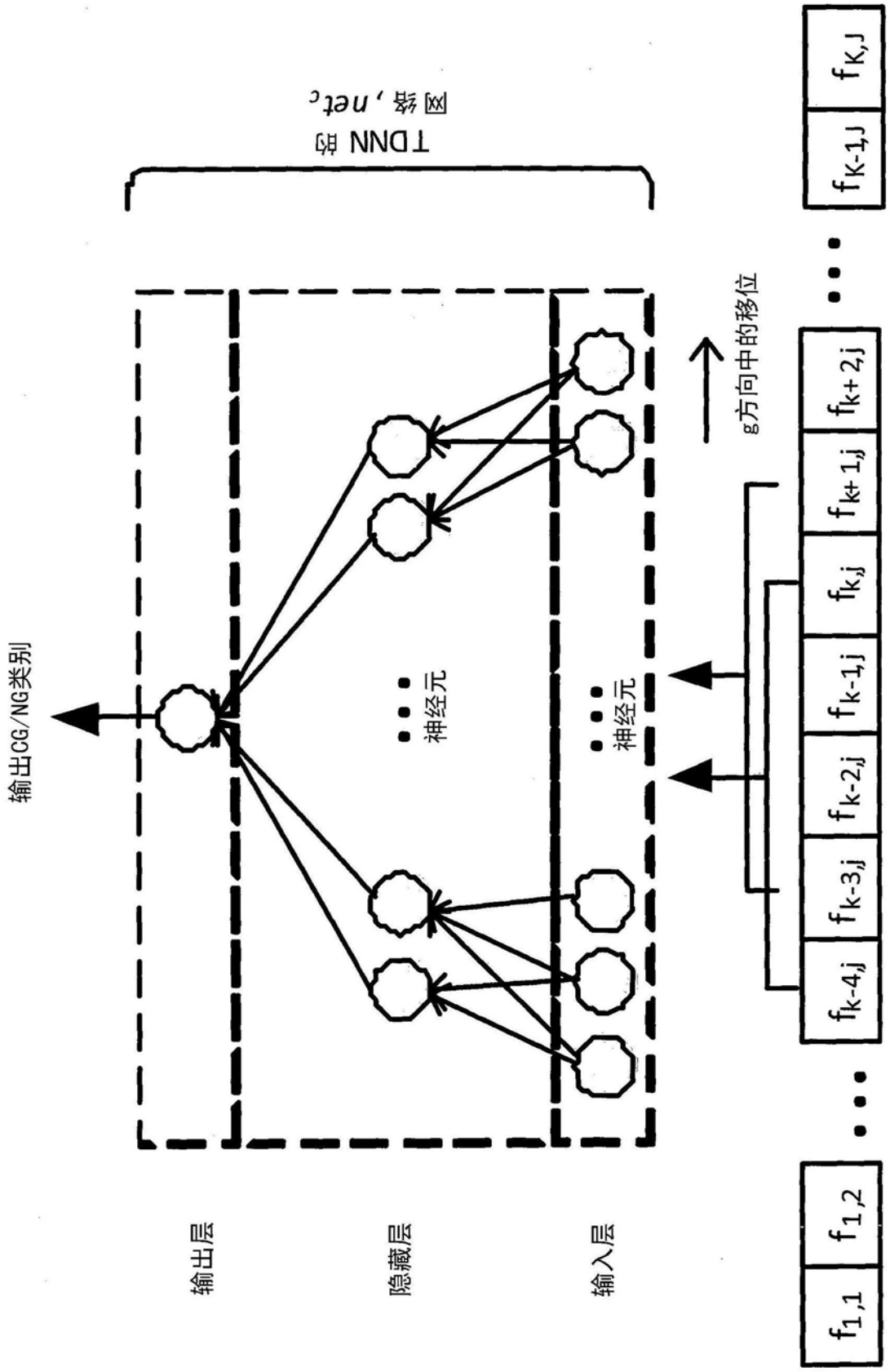

TDNN 的
网络，netc
输出CG/NG类别
神经元
神经元
输出层
隐藏层
输入层
g方向中的移位
f1,1
f1,2
fK-1,J
fK,J
fk-4,j
fk-3,j
fk-2,j
fk-1,j
fk,j
fk+1,j
fk+2,j

图3

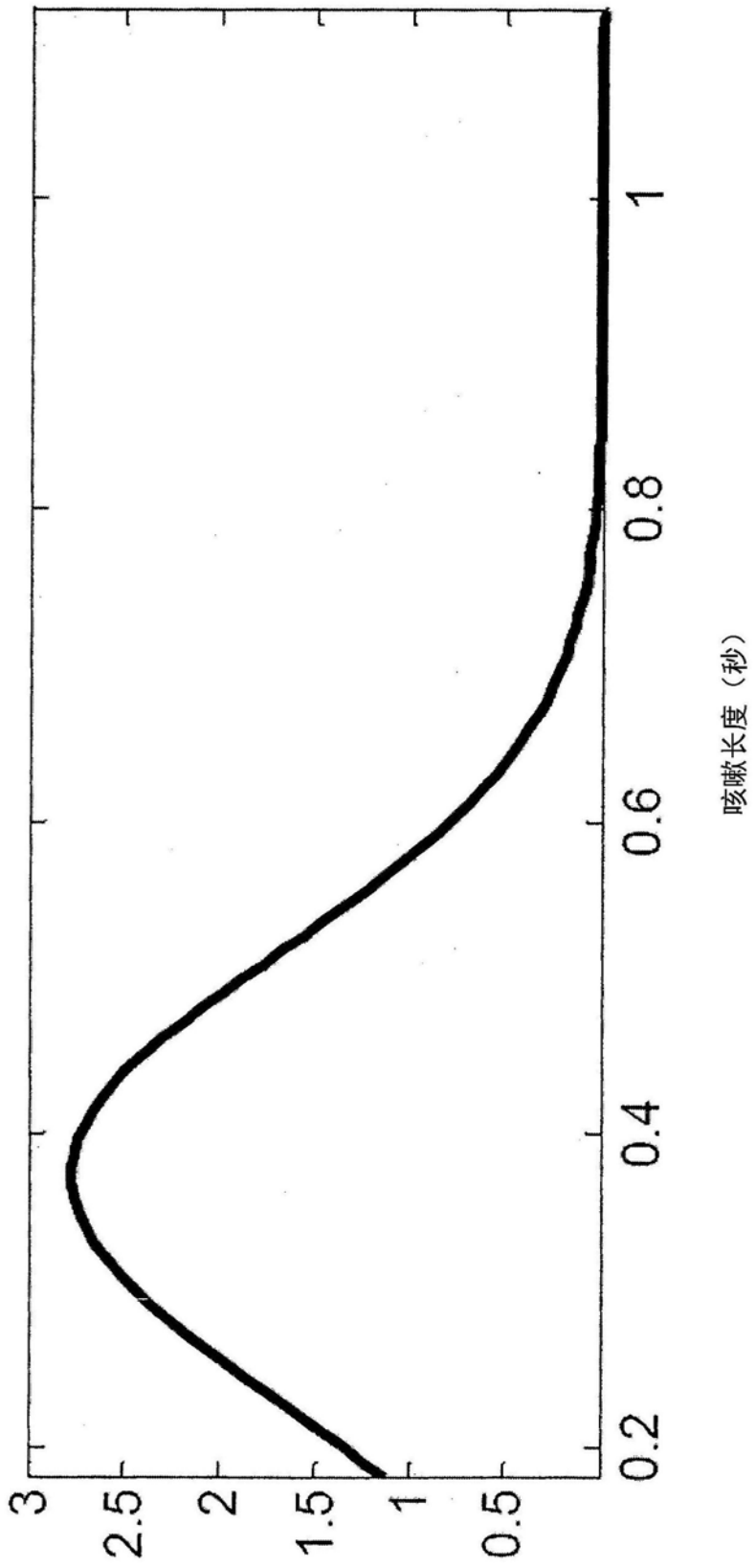


图4

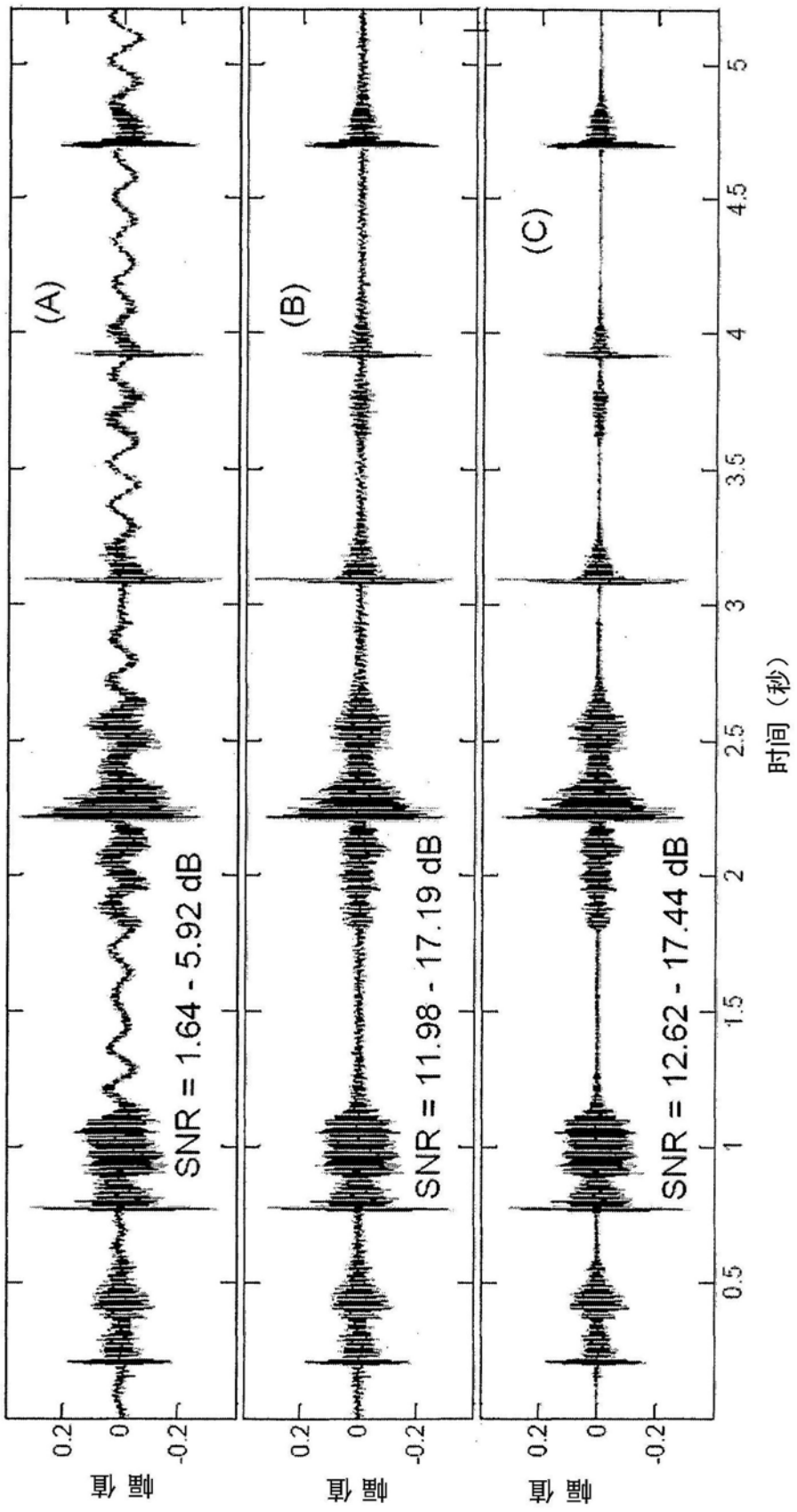

(A)
(B)
(C)
SNR = 1.64 - 5.92 dB
SNR = 11.98 - 17.19 dB
SNR = 12.62 - 17.44 dB
幅值
0.2
0
-0.2
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
时间（秒）

图5

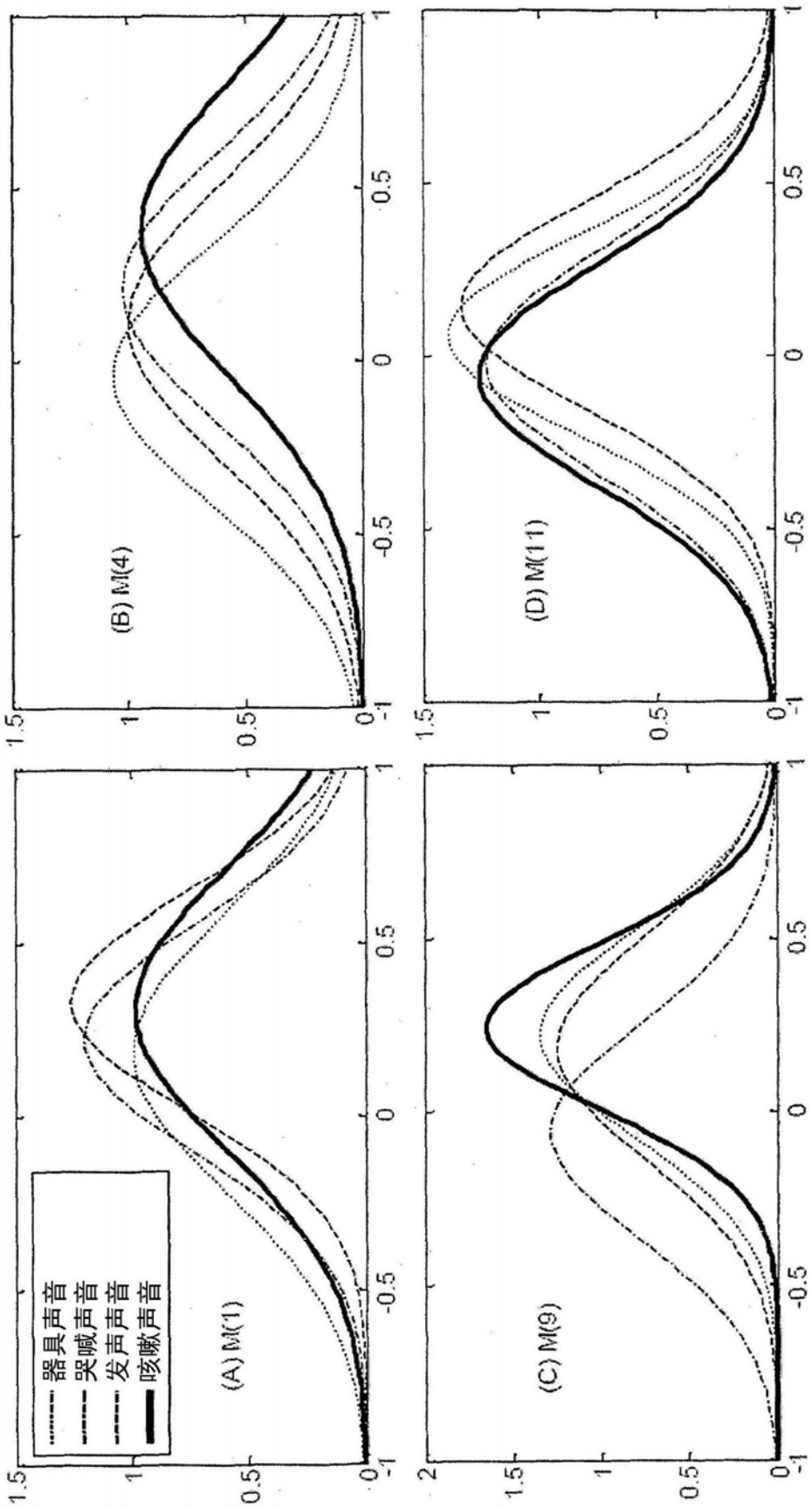


图6

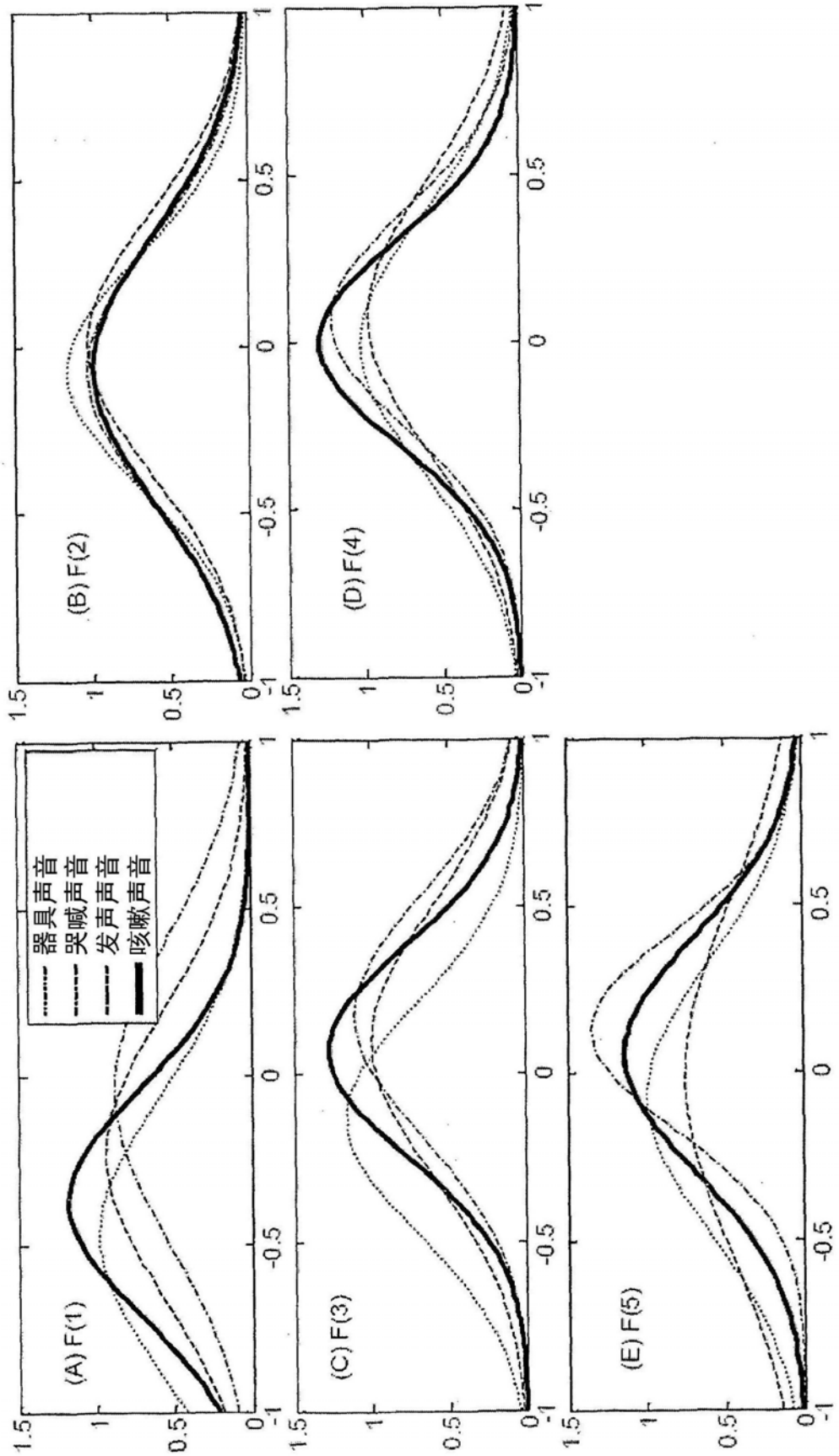


图7

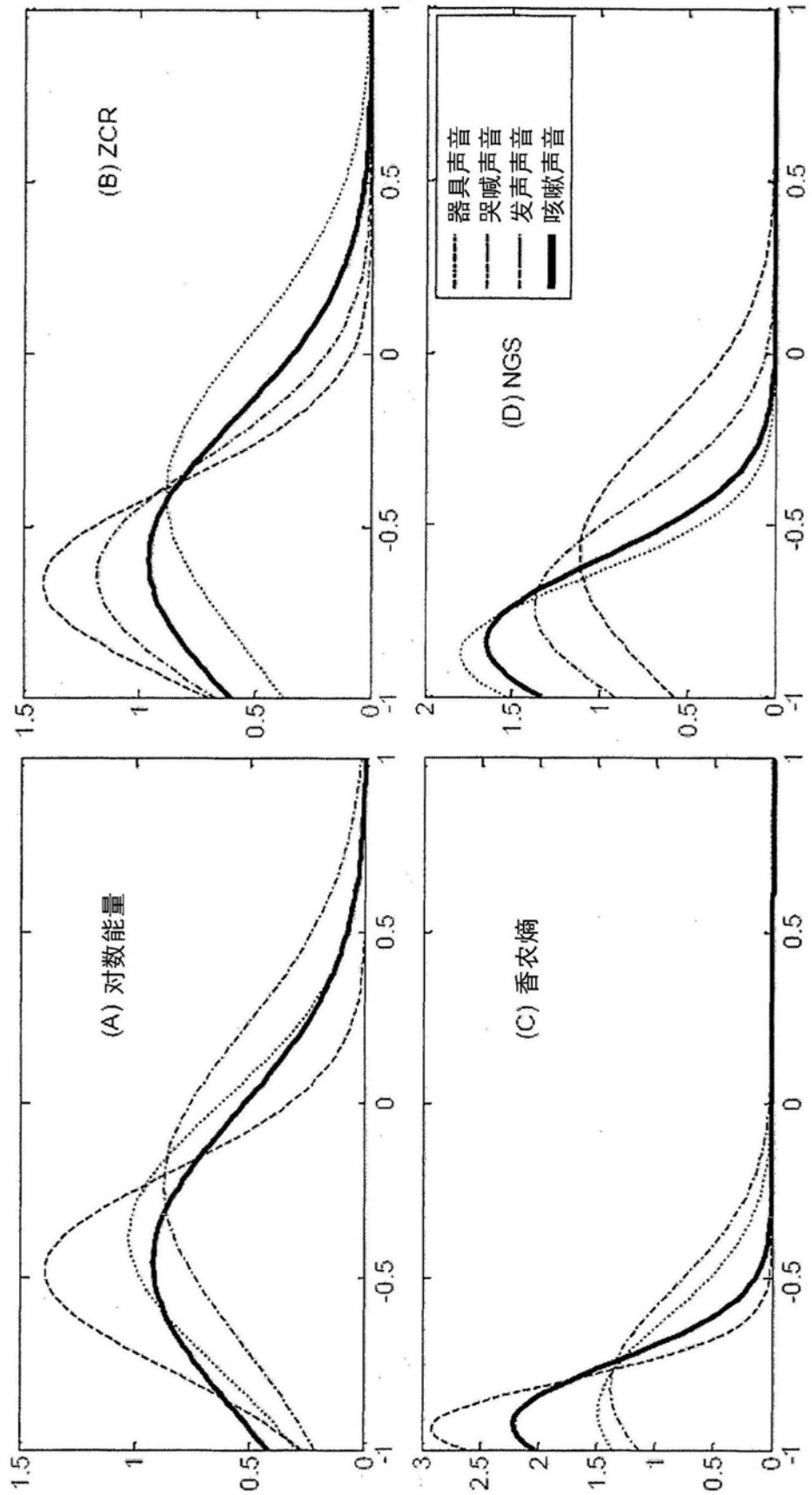


图8

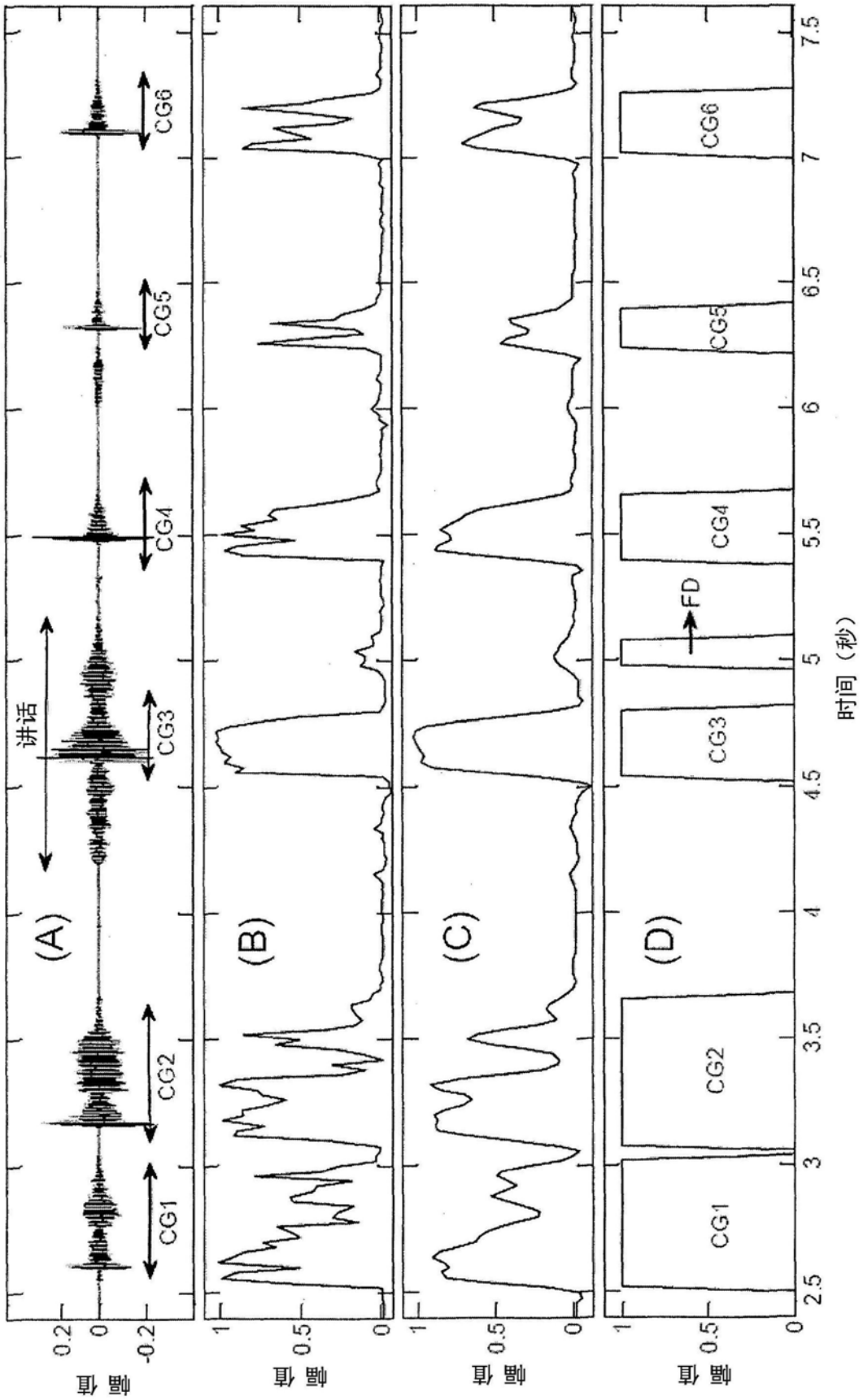

(A)
(B)
(C)
(D)
讲话
CG1
CG2
CG3
CG4
CG5
CG6
FD
幅值
时间（秒）
0.2
0
-0.2
1
0.5
0
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5

图9

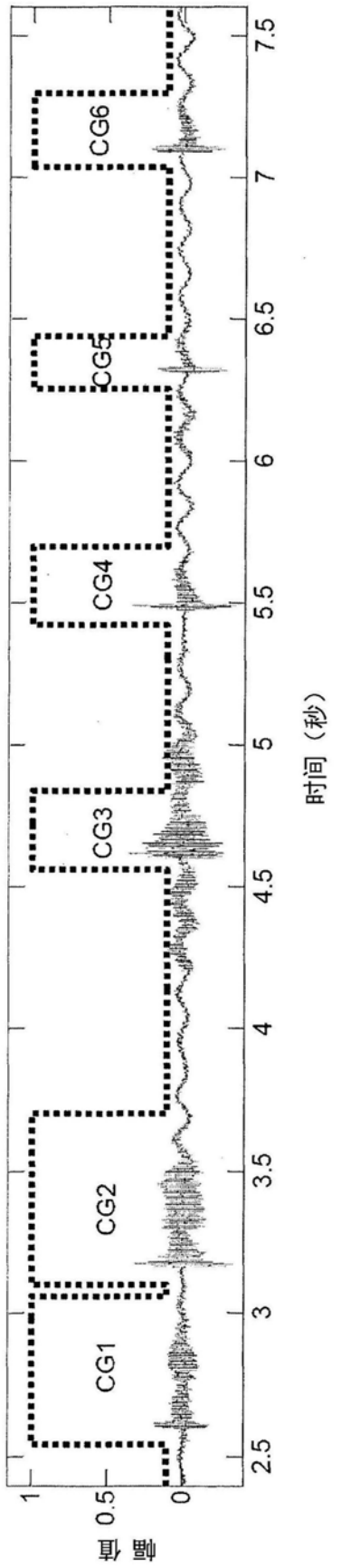


图10

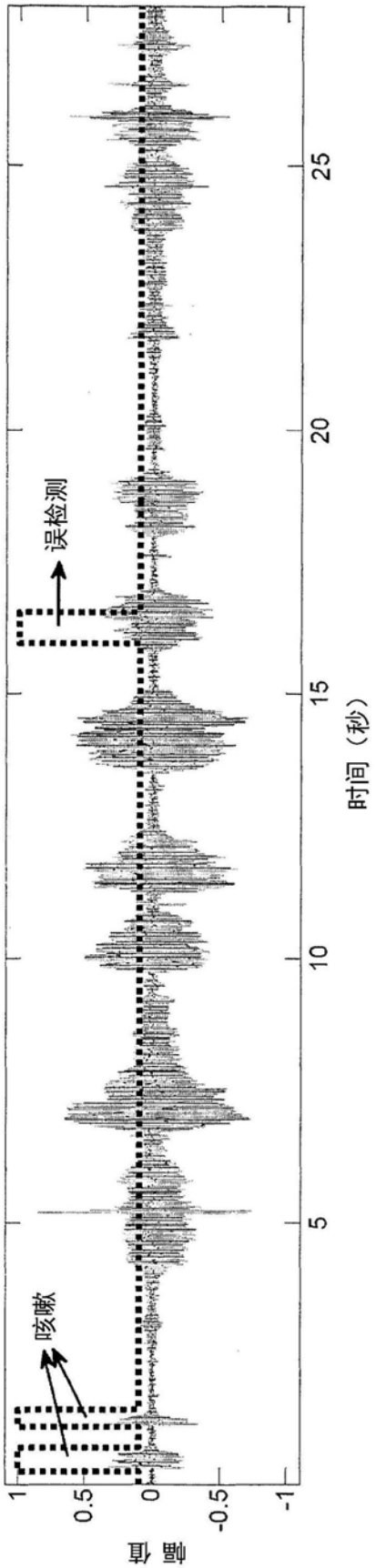


图11

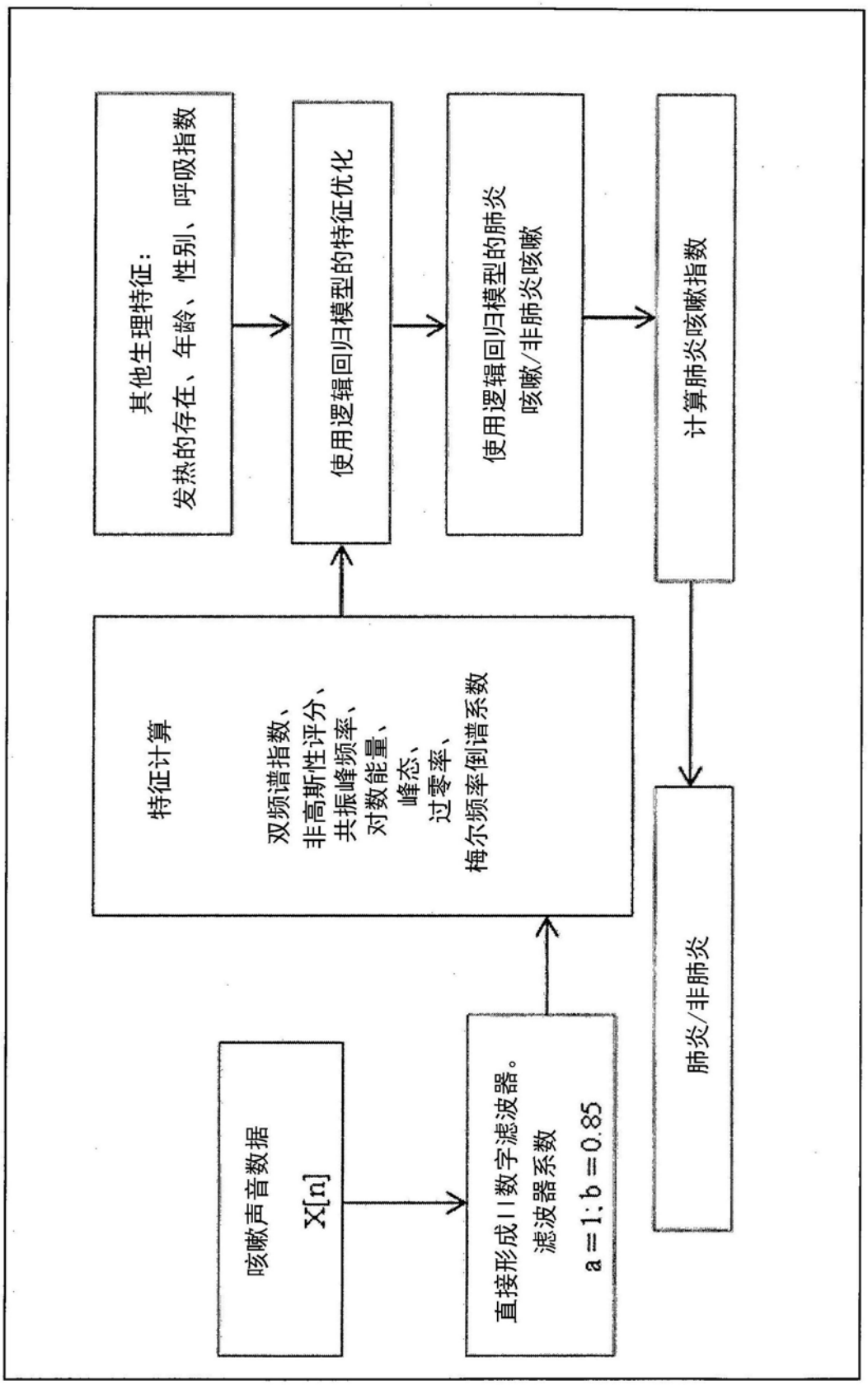

其他生理特征：
发热的存在、年龄、性别、呼吸指数
使用逻辑回归模型的特征优化
使用逻辑回归模型的肺炎
咳嗽/非肺炎咳嗽
计算肺炎咳嗽指数
特征计算
双频谱指数、
非高斯性评分、
共振峰频率、
对数能量、
峰态、
过零率、
梅尔频率倒谱系数
咳嗽声音数据
X[n]
直接形成II数字滤波器。
滤波器系数
a = 1; b = 0.85
肺炎/非肺炎

图12

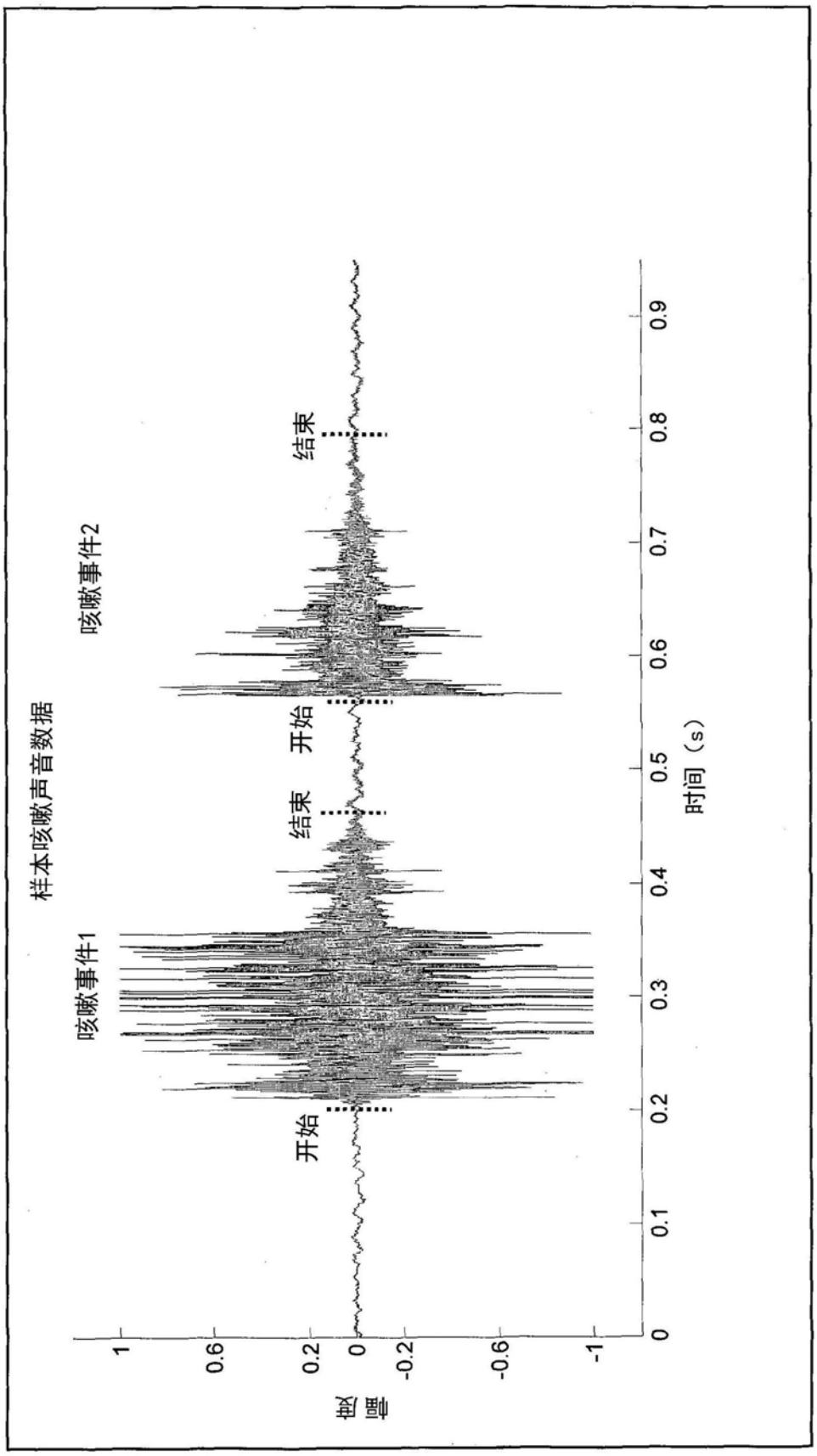

样本咳嗽声音数据
咳嗽事件1
咳嗽事件2
开始
结束
开始
结束
幅值
时间（s）
1
0.6
0.2
0
-0.2
-0.6
-1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

图13

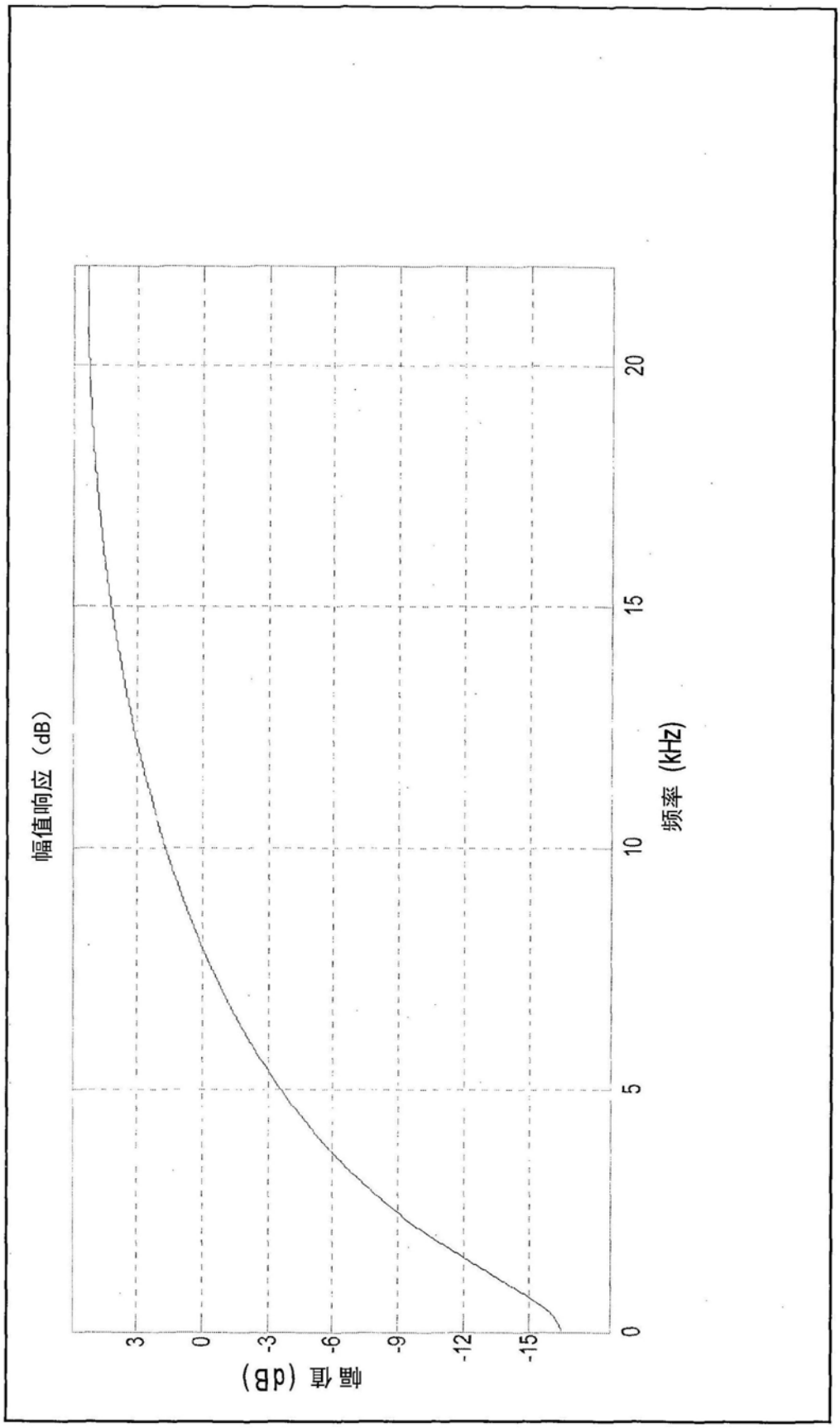


图14

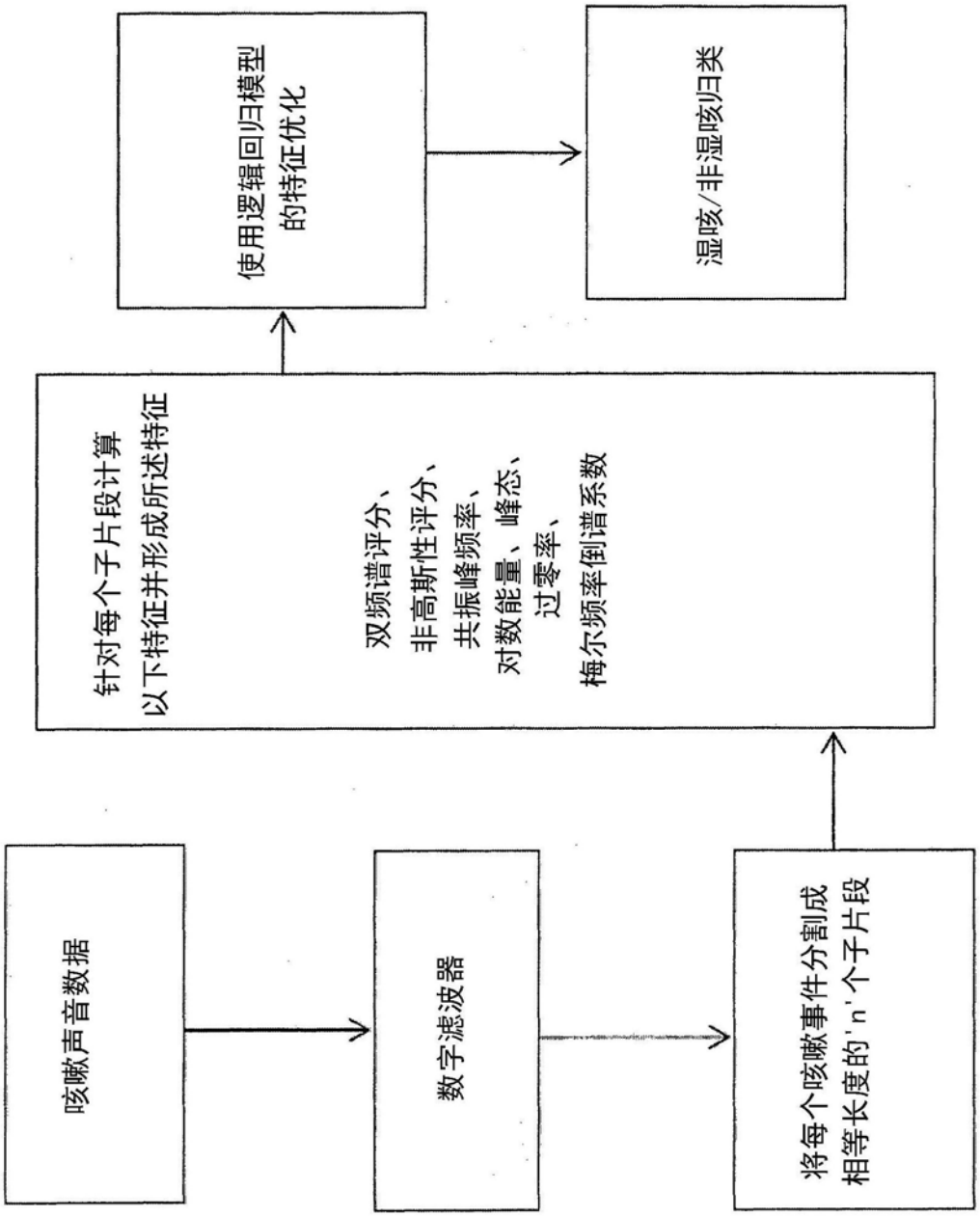

咳嗽声音数据
数字滤波器
将每个咳嗽事件分割成
相等长度的'n'个子片段
针对每个子片段计算
以下特征并形成所述特征
双频谱评分、
非高斯性评分、
共振峰频率、
对数能量、峰态、
过零率、
梅尔频率倒谱系数
使用逻辑回归模型
的特征优化
湿咳/非湿咳归类

图15

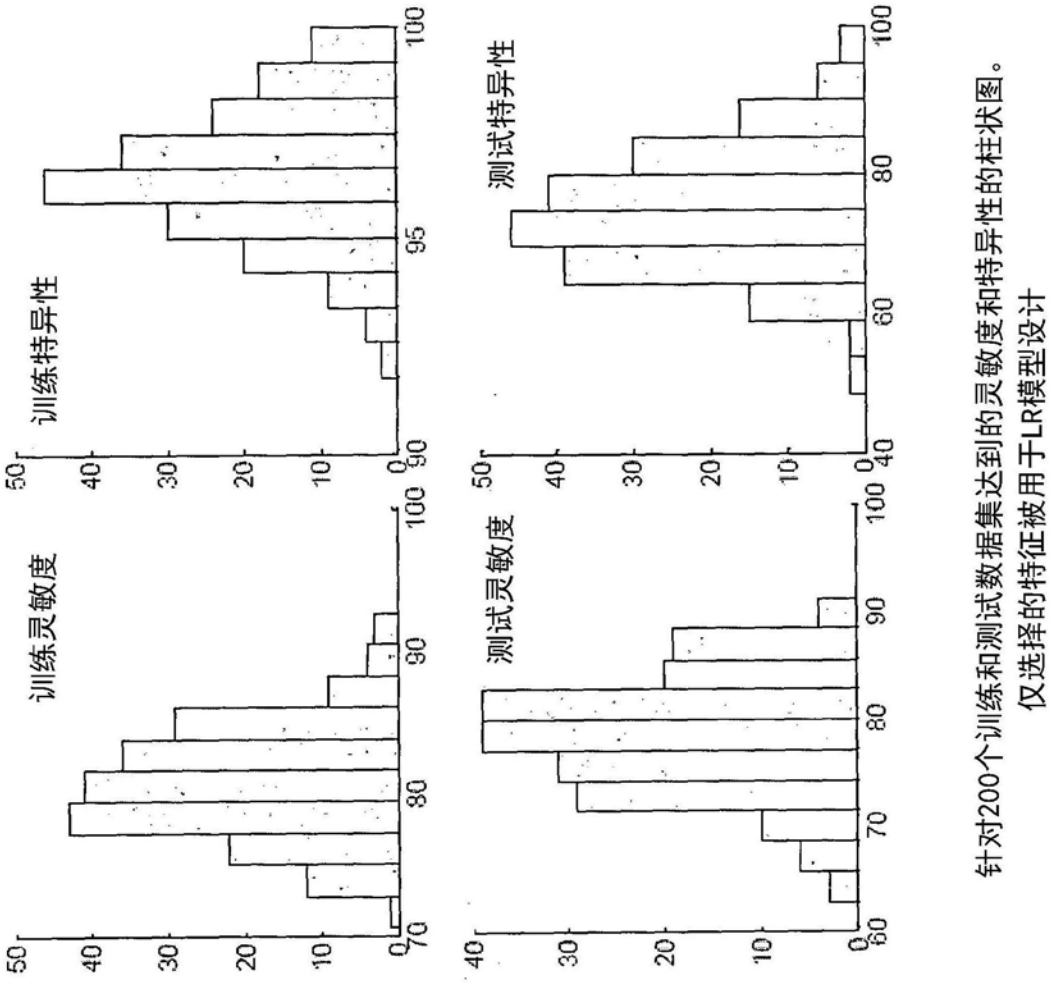

训练特异性
训练灵敏度
测试特异性
测试灵敏度
针对200个训练和测试数据集达到的灵敏度和特异性的柱状图。
仅选择的特征被用于LR模型设计

图16

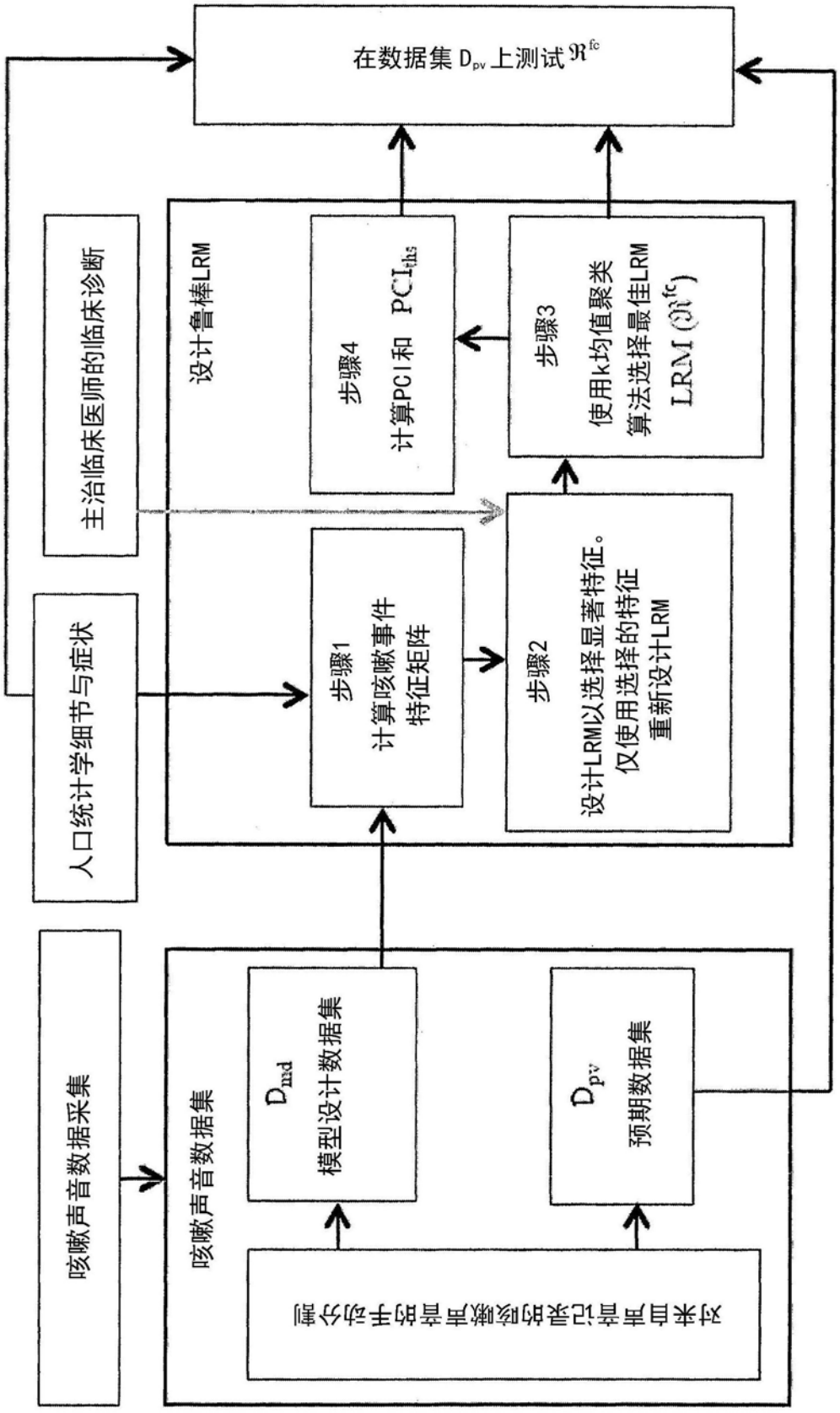

在数据集 D_{pv} 上测试 $\Re^{fc}$
主治临床医师的临床诊断
设计鲁棒LRM
步骤4
计算PCI和 PCI_{ths}
步骤3
使用k均值聚类
算法选择最佳LRM
$LRM(\Re^{fc})$
人口统计学细节与症状
步骤1
计算咳嗽事件
特征矩阵
步骤2
设计LRM以选择显著特征。
仅使用选择的特征
重新设计LRM
咳嗽声音数据采集
咳嗽声音数据集
D_{msd}
模型设计数据集
D_{pv}
预期数据集
对来自声音记录的咳嗽声音的手动分割

图17

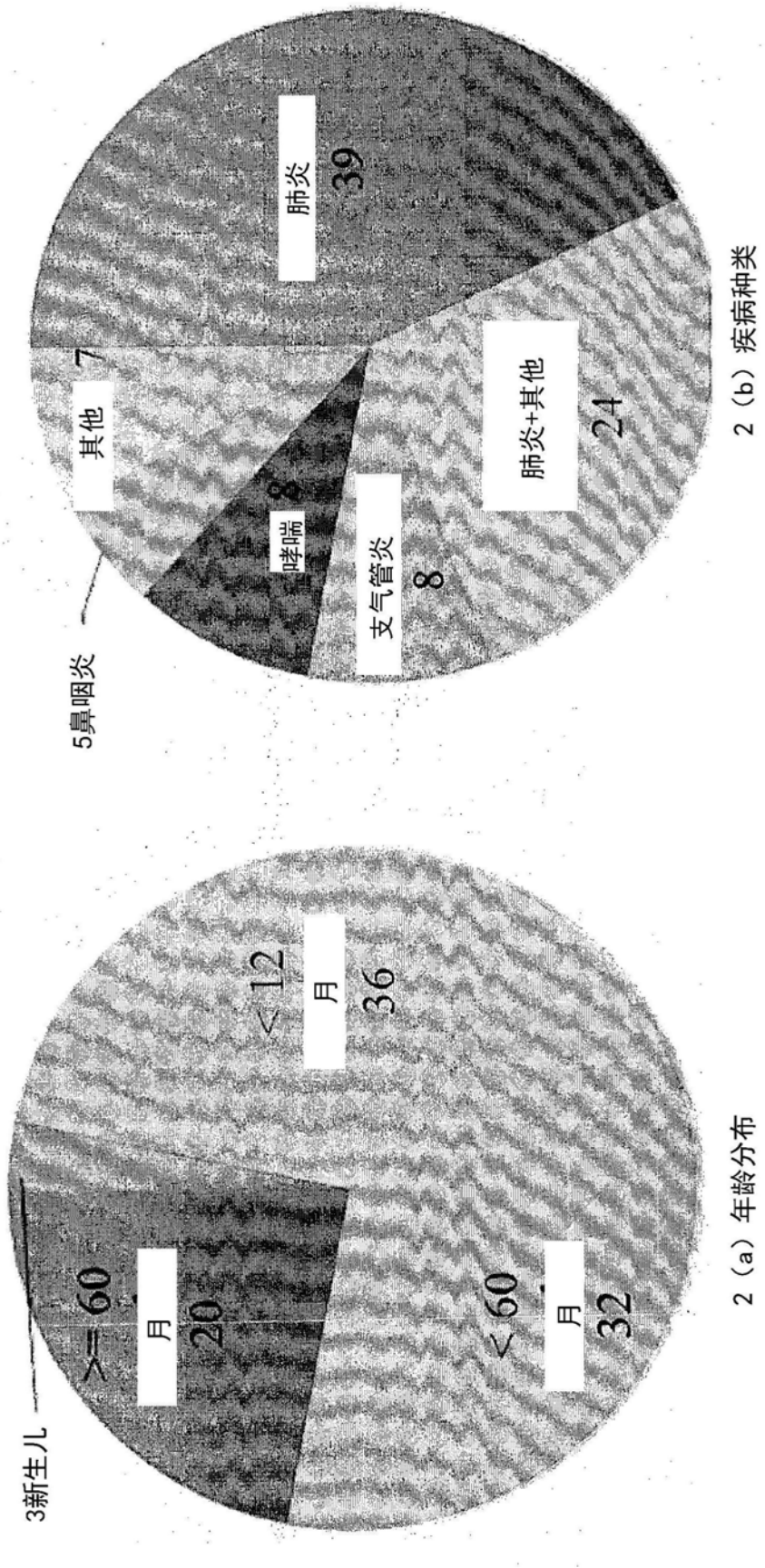

3新生儿
>=60
月
20
<12
月
36
<60
月
32
2（a）年龄分布
5鼻咽炎
其他
7
哮喘
8
支气管炎
8
肺炎+其他
24
肺炎
39
2（b）疾病种类

图18

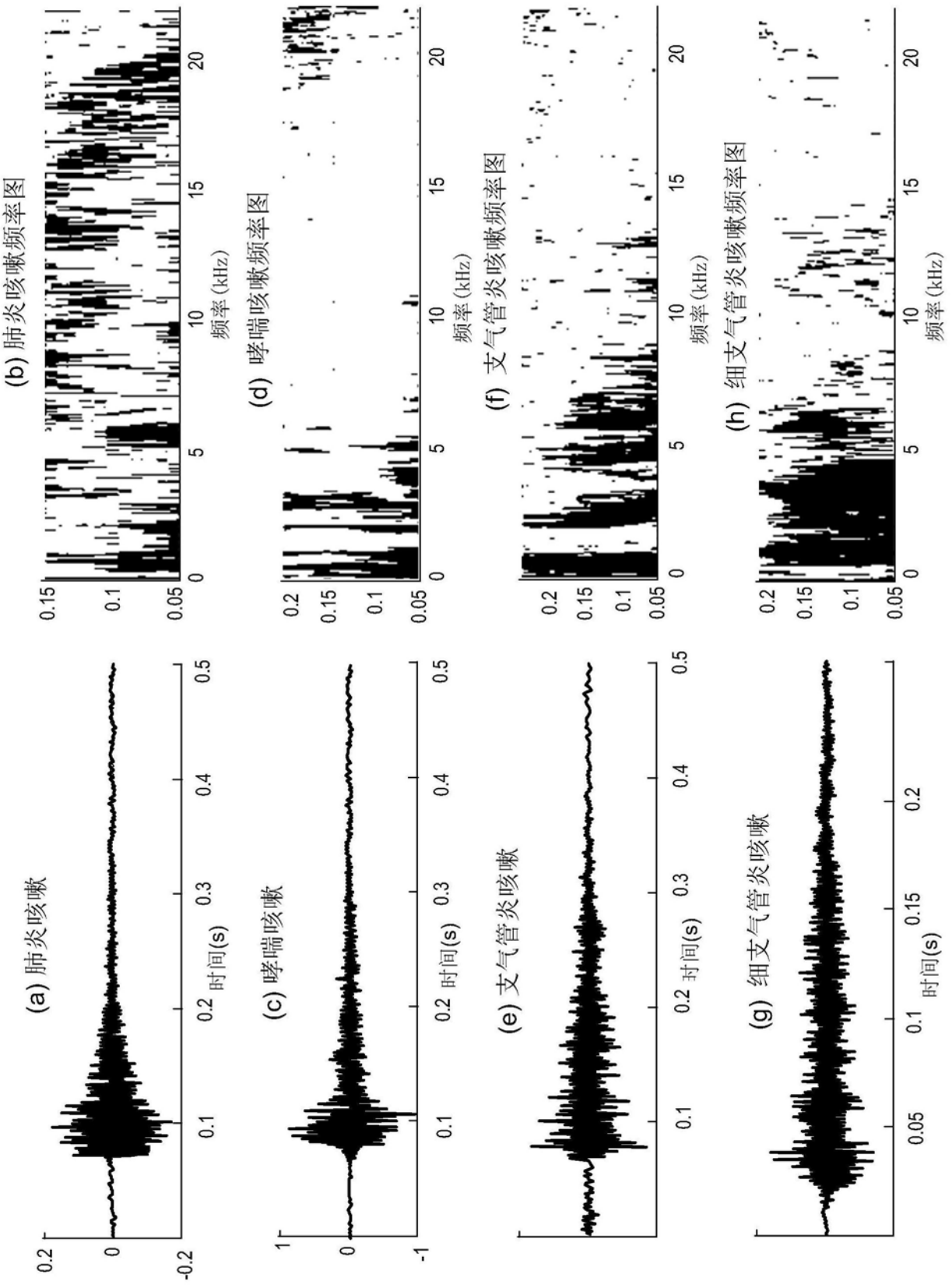

(a) 肺炎咳嗽
(b) 肺炎咳嗽频率图
(c) 哮喘咳嗽
(d) 哮喘咳嗽频率图
(e) 支气管炎咳嗽
(f) 支气管炎咳嗽频率图
(g) 细支气管炎咳嗽
(h) 细支气管炎咳嗽频率图
时间(s)
频率(kHz)

图19

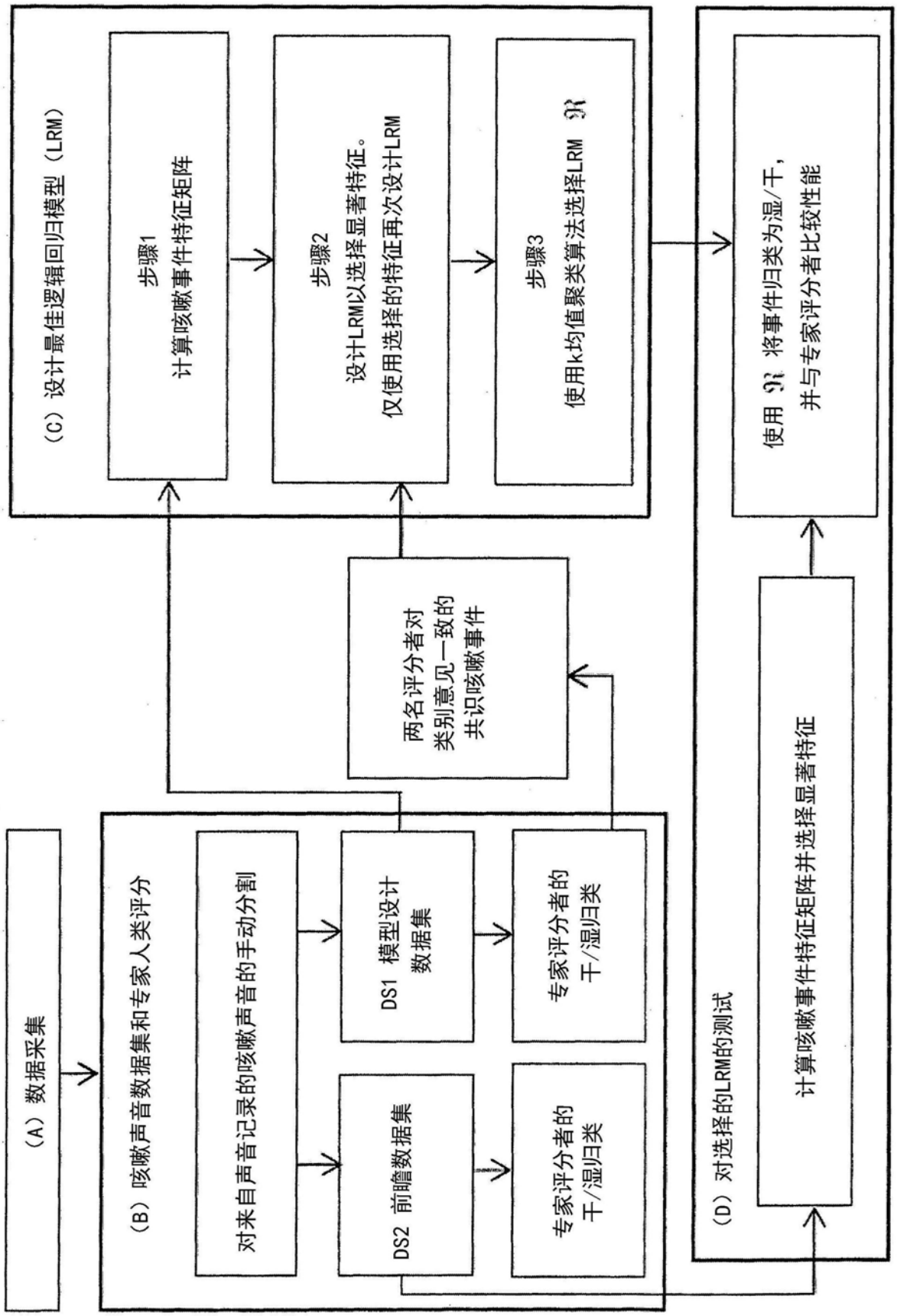

（A）数据采集
（B）咳嗽声音数据集和专家人类评分
对来自声音记录的咳嗽声音的手动分割
DS1 模型设计数据集
专家评分者的干/湿归类
DS2 前瞻数据集
专家评分者的干/湿归类
两名评分者对类别意见一致的共识咳嗽事件
（C）设计最佳逻辑回归模型（LRM）
步骤1
计算咳嗽事件特征矩阵
步骤2
设计LRM以选择显著特征。
仅使用选择的特征再次设计LRM
步骤3
使用k均值聚类算法选择LRM
（D）对选择的LRM的测试
计算咳嗽事件特征矩阵并选择显著特征
使用 将事件归类为湿/干，
并与专家评分者比较性能

图20

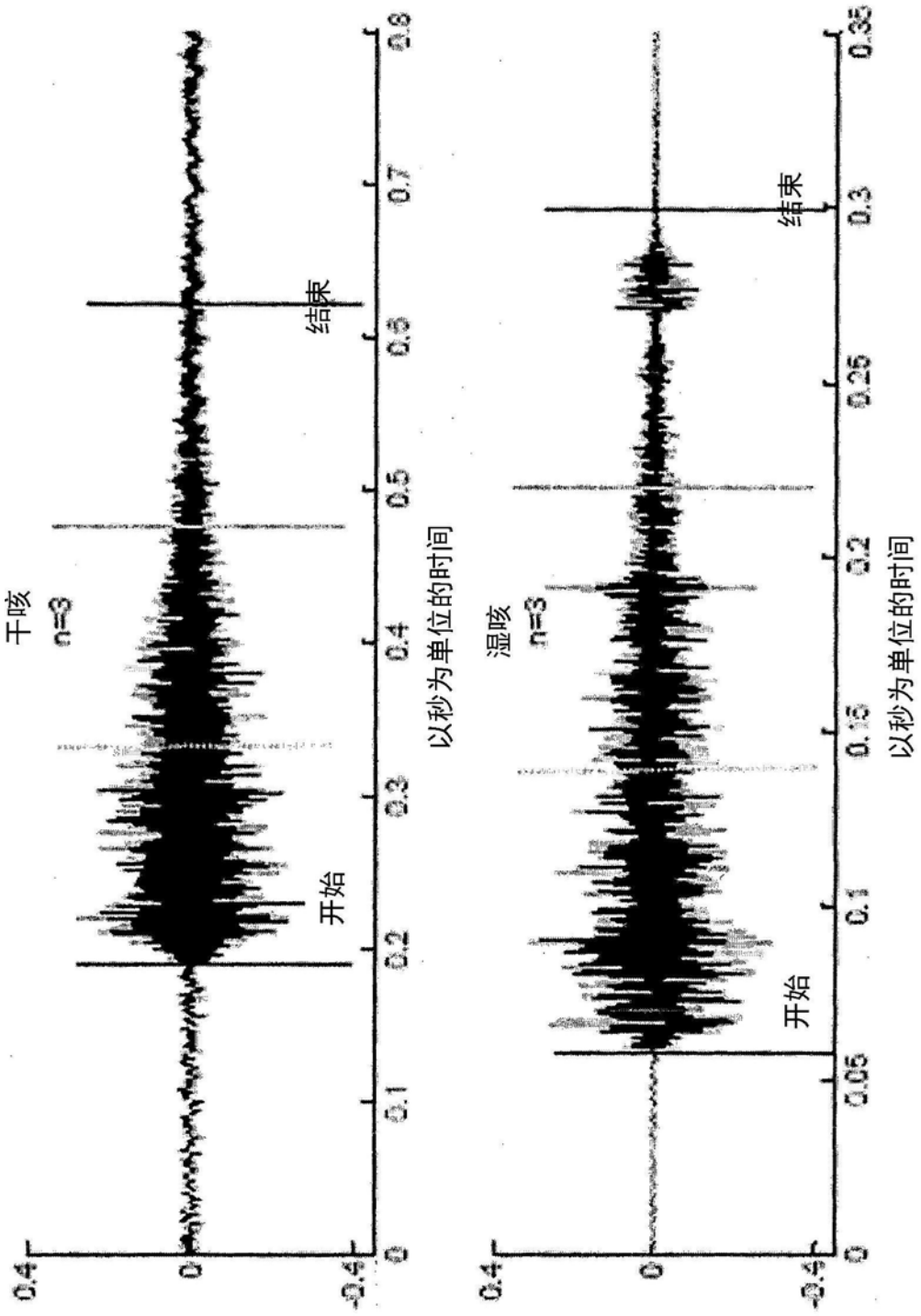


图21

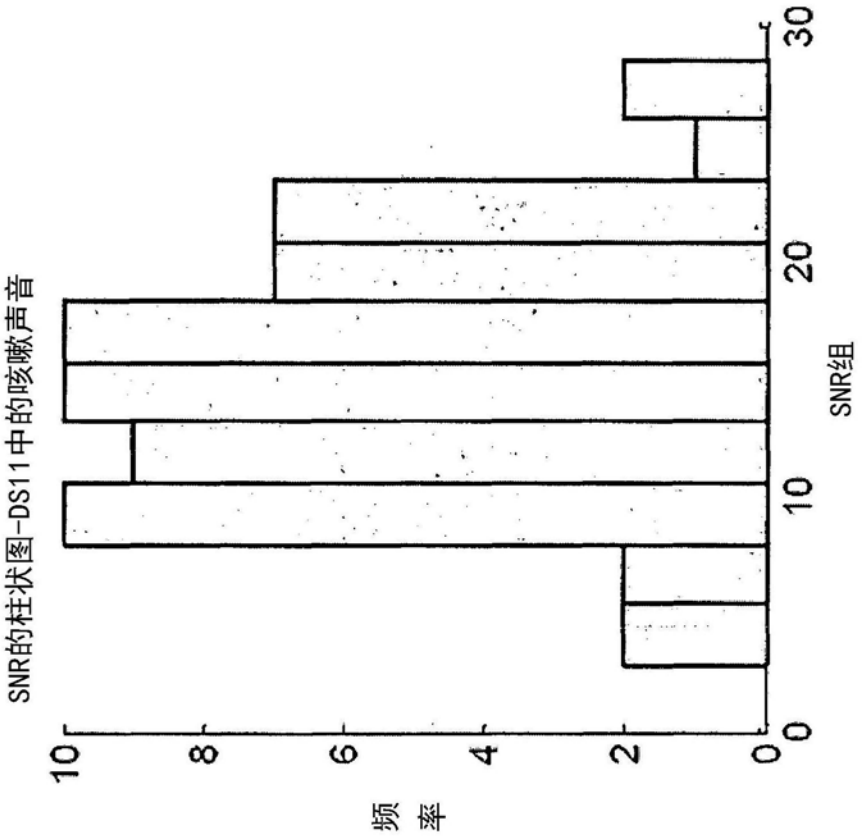


图22A

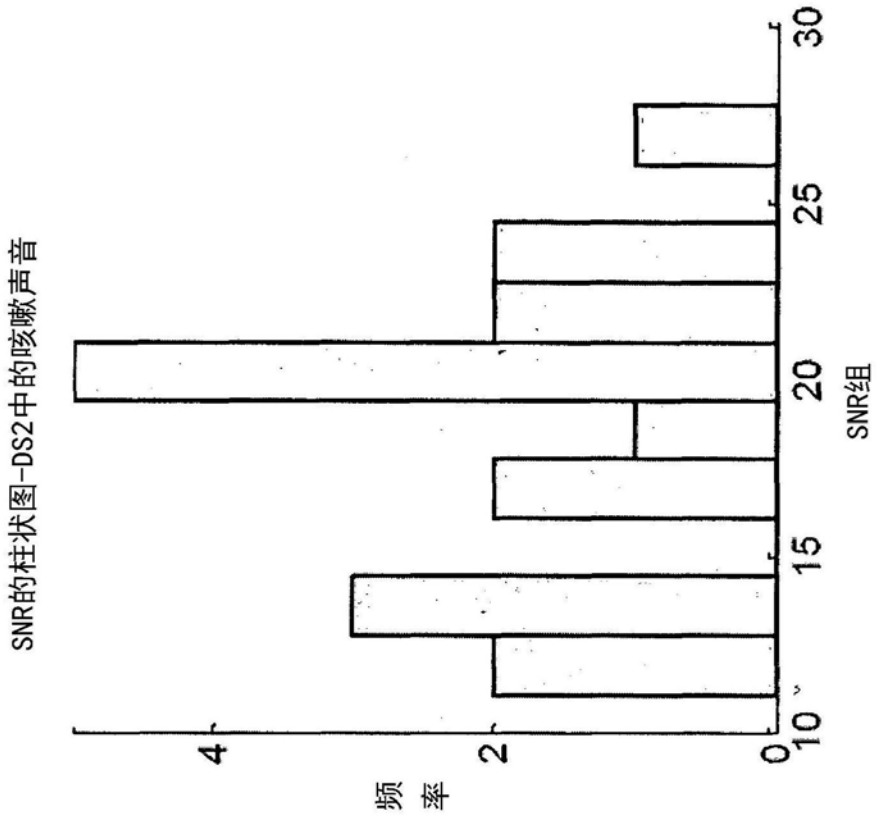


图22B

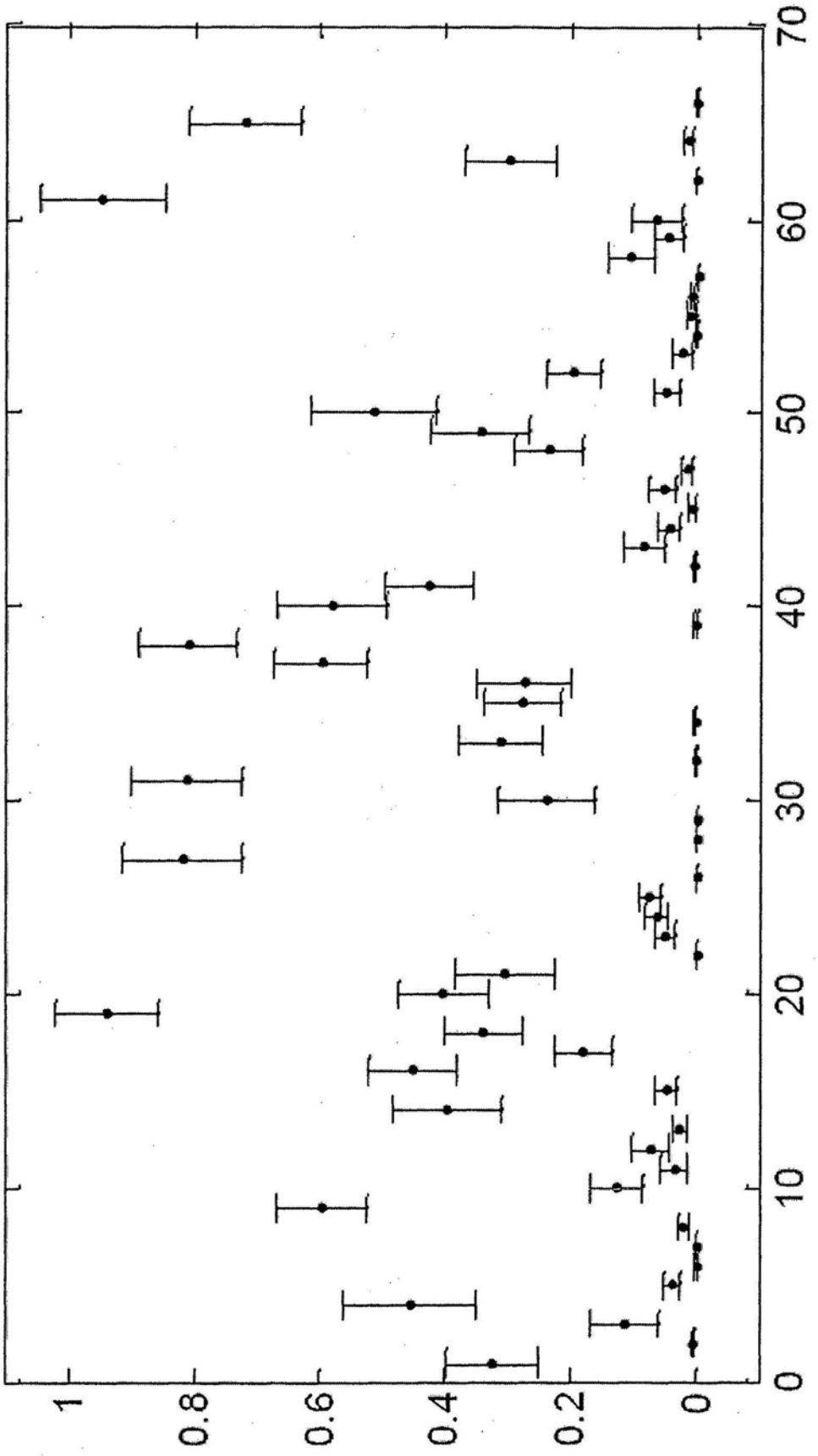


图23