

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日
2015年12月23日 (23.12.2015) WIPO | PCT

(10) 国际公布号
WO 2015/192395 A1

- (51) 国际专利分类号:
H04B 17/00 (2006.01)
- (21) 国际申请号: PCT/CN2014/081156
- (22) 国际申请日: 2014年6月30日 (30.06.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410269839.5 2014年6月17日 (17.06.2014) CN
- (71) 申请人: 公安部第三研究所 (THE THIRD RESEARCH INSTITUTE OF MINISTRY OF PUBLIC SECURITY) [CN/CN]; 中国上海市徐汇区岳阳路76号, Shanghai 200031 (CN)。上海精汇电子设备闵行有限公司 (SHANGHAI JINGHUI ELECTRONICS EQUIPMENT MINHANG CO., LTD.) [CN/CN]; 中国上海市闵行区黎安西路1398号, Shanghai 201100 (CN)。
- (72) 发明人: 戎玲 (RONG, Ling); 中国上海市徐汇区岳阳路76号, Shanghai 200031 (CN)。赵澎 (ZHAO, Peng); 中国上海市闵行区黎安西路1398号, Shanghai 201100 (CN)。鲍逸明 (BAO, Yiming); 中国上海市徐汇区岳阳路76号, Shanghai 200031 (CN)。唐德成 (TANG, Decheng); 中国上海市闵行区黎安西路1398号, Shanghai 201100 (CN)。胡融 (HU,

Rong); 中国上海市闵行区黎安西路1398号, Shanghai 201100 (CN)。史源 (SHI, Yuan); 中国上海市徐汇区岳阳路76号, Shanghai 200031 (CN)。

(74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区西坝河西里28号英特公寓C座104, Beijing 100028 (CN)。

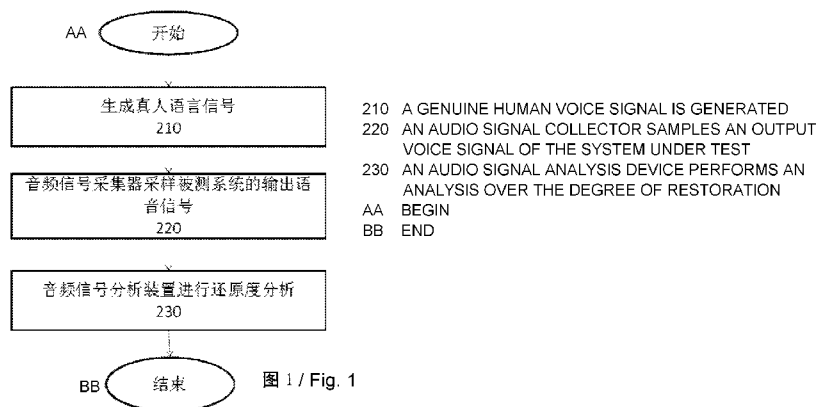
(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

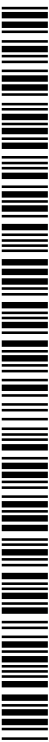
(54) Title: METHOD AND SYSTEM FOR SCORING HUMAN SOUND VOICE QUALITY

(54) 发明名称: 人声语音质量评分方法及系统



(57) Abstract: A method and system for scoring human sound voice quality, the method comprising: a sound source part (110) generating a standard genuine human voice signal as an input voice signal, and inputting the input voice signal to the system under test (120) from the sound source part (110); transmitting the input voice signal in the system under test (120) and outputting the input voice signal from an output end of the system under test (120) as an output voice signal; collecting the continuous output voice signal; and segmenting and analyzing the signal flow of the collected output voice signal to determine the degree of restoration.

(57) 摘要: 一种人声语音质量评分方法及系统, 该方法包括: 声源部分 (110) 产生标准的真人语音信号作为输入语音信号, 从声源部分 (110) 输入被测系统 (120); 输入语音信号在被测系统 (120) 中传递, 并作为输出语音信号从被测系统 (120) 的输出端输出; 采集输出的连续的语音信号; 将采集的输出语音信号的信号流切片和分析处理而确定还原度。



WO 2015/192395 A1

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

人声语音质量评分方法及系统

技术领域

- 5 本申请涉及语音传输检测领域，更具体地涉及人声语音质量评分方法及系统。

背景技术

- 在现有技术中，需要对语音或视频传输使用的设备进行语音传输、视频传输等性能是否符合规定的各项指标做检测（检验）或测试，只有当这些检测或测试所获得结果符合规定的各项标准要求，才能保证这些语音或视频设备是可用且好用、质量是有保障的，进而，才能将这些设备进行生产和实际应用。随着语音、视频传输设备的各项技术的不断进步，对设备以及设备配合使用时的各方面的性能要求也在不断更新进步中。因此，对于设备检测（检验）的技术，也随之不断改进以便提高检测检验结果的准确性和精确度。以应用于住宅楼宇出入口控制的楼宇对讲系统为例，对其使用的设备以及系统整体上的音频传输特性的检测（检验）或测试就是整个对讲系统检测最为重要的部分。已知的楼宇对讲系统的音频传输质量测试的试验方法是评价楼宇对讲系统产品语音传输质量的一整套全程声测试的试验方法，其包括了5项技术参数（如：响度评定值、频率响应/频响、失真、信噪比、侧音掩蔽评定值）的测试和计算。该已有的测试方法中，对输出端的音频信号检测则专注于对其频响、失真及振幅特性的技术特性的评测。例如，如图3所示，声源/音源即信号源301产生音频信号后，通过功放装置302、被测系统的音频信号通路等，在被测系统的音频输出端（如：被测的喇叭、听筒等）303输出；输出端303的输出口固定在隔声挡板上，隔音挡板作为模拟的设备安装用的墙体以保证测试结果更准确性，而正对输出口方向，设有传声器（如：麦克风等）304，传声器304内设的中央处理器CPU采集输出信号，可以将音频信号转换为电信号；传声器304将电信号经过一放大装置305（如信号放大器）传给测量装置测量后由频谱仪306输出结果，其中，测量装置对采样
- 10
- 15
- 20
- 25

的输出信号进行相应技术参数评测包括:

5 频响测试: 调节音频信号源的频率在 200-4000Hz 范围, 记录在 MIC 端的 CPU 采样值 (输出端的音频响应信号), 由此得到音频信号源调节到的每个频点的声压 (如: 声音强弱、声音能量) 有效值, 从而计算 (如: 傅里叶变换、电频表测算等) 得到输出端 (如: 喇叭) 的频响;

失真测试: 调节音频信号源的频率至要测试的频率, 记录 MIC 端 CPU 的采样值, 通过频谱转换 (如: 傅里叶转换) 得到该采样值信号的频谱图, 再通过该频谱图计算出喇叭在该信号所对应的信号源输入信号频率下的失真值;

10 振幅特性测试: 在同一频率下改变音频信号源的输出幅度, 记录 MIC 端 CPU 的采样值, 通过频谱转换 (如: 傅里叶变换) 得到该采样值信号的频谱, 观察在不同幅度信号输入下的输出值是否与输入值相对应, 比如, 是否符合线性关系, 例: 输入提高 10db, 输出是否也提高 10db。

15 进而, 仅基于上述参数间接判别其输出信号的还原度, 比如: 失真越小, 频响越平 (即: 尽可能输入输出之间的变化是趋近于稳定、线性的), 则判断输出的声音越好, 等等。其中, 还原度, 是指信号源输入的原帧频谱的形状与输出端的采样帧频谱形状的一致性, 其是音频测量中评测设备和系统整体性能尤其是输出性能的一个重要技术参数。而由于现有技术缺少对输出端的输出信号还原度进行直接测试和判断, 导致对设备和系统性能尤其输出性
20 能的测评并不精确。并且, 现有的测试方式中还存在以下缺陷: 不能对所有关心的连续频率都测试从而导致的测试结果不精确; 实际设备和系统应用时, 实际存在的语音信号都是多频点信号 (如: 人发声为 N 个频点/频率的各种叠加), 而现有的失真测试是使用单频点 (即单纯的 200Hz、400Hz 等) 进行, 与实际的多频点的失真情况并不一致, 导致利用该测试的失真来判别还原度
25 也不准确。因而, 最终导致现有的音频信号的检测结果并不准确或者说精确度不高, 与实际的主观评测差距较大。。

发明内容

基于现有技术中存在的上述缺陷, 本申请的主要目的是提供一种人声语

音质量评分方法及系统，以提高音频信号还原度测试的准确度、精确度。进一步，提高了对设备、系统等音频信号性能测试整体的精确度。为了解决上述现有技术中的技术缺陷，本申请的目的是通过以下技术方案来实现的。

本申请提供了一种音频信号还原度测试方法，包括：声源部分、被测系统、音频信号采集装置、音频信号分析装置；声源部分产生标准的真人语音信号，将所述真人语音信号作为输入语音信号，从声源部分输入到所述被测系统；输入语音信号在被测系统中传递，并作为输出语音信号从被测系统的输出端输出；音频信号采集装置采集输出端输出的连续的语音信号，并且，将采集的输出语音信号传送到音频信号分析装置；音频信号分析装置对输出语音信号的信号流进行切片和分析处理，确定还原度。

其中，所述音频信号采集装置，包括：将采集的连续语音信号转换成相应的数字信号流，传送给音频信号分析装置进行信号流的切片和分析处理。

其中，音频信号分析装置，包括：获取声源部分产生的作为输入语音信号的所述真人语音信号；将所述输入语音信号的信号流，以时间间隔进行切片划分，以将信号流划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

其中，音频信号分析装置，包括：对来自音频信号采集装置采集的输出语音信号的信号流，与该输出语音信号的信号流相对应的来自声源部分的所述输入语音信号的信号流进行同步；以与切片所述输入语音信号的信号流相同的时间间隔，对采集的连续的语音信号的信号流切片划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

其中，音频信号分析装置，包括：提取属于所述输入语音信号的信号流中的一语音信号组的片段所对应的声音特征值，以及提取与该片段相应的、属于所述输出语音信号的信号流中的一片段所对应的声音特征值，基于相似性原理，计算和分析两个相应片段的对比分数值；将所有相应的、分属于输

入语音信号的信号流和相应的输出语音信号的信号流中的语音信号组的片段的对比分数值，进行统计和/或平均计算，以确定还原度。

5 本申请还提供了一种人声语音质量评分方法，包括：声源部分产生标准的真人语音信号，将所述真人语音信号作为输入语音信号，从声源部分输入到所述被测系统；输入语音信号在被测系统中传递，并作为输出语音信号从被测系统的输出端输出；采集输出端输出的连续的语音信号；将采集的输出语音信号的信号流进行切片和分析处理，确定还原度。

10 其中，还包括：预先获取声源部分产生的作为输入语音信号的所述真人语音信号；将所述输入语音信号的信号流，以时间间隔进行切片划分，以将信号流划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

15 其中，将采集的输出语音信号的信号流进行切片和分析处理，包括：对来自音频信号采集装置采集的输出语音信号的信号流，与该输出语音信号的信号流相对应的来自声源部分的所述输入语音信号的信号流进行同步；以与切片所述输入语音信号的信号流相同的时间间隔，对采集的连续的语音信号的信号流切片划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

20 其中，将采集的输出语音信号的信号流进行切片和分析处理，确定还原度，包括：提取属于所述输入语音信号的信号流中的一语音信号组的片段所对应的声音特征值，以及提取与该片段相应的、属于所述输出语音信号的信号流中的一片段所对应的声音特征值，基于相似性原理，计算和分析两个相应片段的对比分数值；将所有相应的、分属于输入语音信号的信号流和相应
25 的输出语音信号的信号流中的语音信号组的片段的对比分数值，进行统计和/或平均计算，以确定还原度。

其中，还包括：对输入语音信号的信号流和输出语音信号的信号流以相

同的时间间隔进行切片，是以 20ms 时间间隔将信号流切片包含一个或多个语音信号的语音信号组的片段。

5 本申请使用真人说话声音与采集或采样信号进行比较分析，符合设备及其所在系统（例如：楼宇对讲系统）在实际应用中的情况，而由此能得到与实际应用比较一致的测试结果；而对连续发出的音频信号的切片以连续的帧方式进行处理，包括了对所有关心的连续频率（例如：从在 300Hz~3400Hz 范围的真人说话语音中的全部频率）的测试、以及包括了直接基于连续的切片（帧）判断音频输出性能的还原度，其对安全通信设备及系统的音频测试结果、音频输出性能的判别更准确、更精确。

10 附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1 为本申请的人声语音质量评分系统的一实施例的结构框图；

15 图 2 为本申请的人声语音质量评分方法的一实施例的流程图；

图 3 为现有的音频传输质量测试中对输出端的几个测评参数进行测量的一实施例的示意图。

具体实施方式

20 本申请的主要思想在于，检测中，声源采用真人语音作为输入语音信号，使得在检测中传输的语音其音频性能更符合被测设备或系统的实际应用情况，从而对被检测系统的声音输出特征（比如楼宇对讲系统这类对安全性能有要求的通信设备和通信系统）进行检测或测试所获得的更准确、精确的测试结果；进一步，将连续音频信号进行切片连续处理，与真人输入语音信号进行相似性比较，得到还原度的分数，从而更精确、准确地确定被检测系统
25 的声音输出性能。

为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

参见图 1 所示本申请的人声语音质量评分系统的一实施例的结构示意图。该实施例中，以对楼宇对讲系统的音频传输特性进行采样测试为例，基于采样信号的频谱分析以及声源信号进行还原度分析，确定该被测系统的声音输出性能。该实施例的检测系统 100 中，主要可以包括：声源部分 110、被测系统 120、音频信号采集装置（采集器）130、音频信号分析装置 140。

声源部分 110，产生特定语音信号，此特定语音信号可以为标准的人说话的声音，例如：国际电信联盟远程通信标准化组织 P.501 真人语音信号。该真人语音信号作为被检测系统 120 的音频传输特性的测试用传输语音。该语音信号可以经由声源部分 110 作为输入语音信号（例如：信号源 301），输入到被测系统 120 的输入端，在被测系统 120 中传输，最后作为被测输出语音信号由被测系统 120 的输出端（例如：喇叭、听筒等 303）输出。

被测系统 120，在本实施例中，可以是楼宇对讲系统，接收来自声源部分 110 的输入语音信号，经过功放装置、被测通路、功放装置，传输该输入语音信号直到被测系统的输出端，由该输出端将经过被测系统 120 的该输入的语音信号作为被测输出语音信号而输出。其中，被测通路，可以是被测系统（如被测的楼宇对讲系统）中需要检测的通话通路。

音频信号采集装置 130，采集被测系统 120 输出的语音信号，把采集到的语音信号进行转换后传送到音频信号分析装置 140 进行处理分析。如在输出端设置传声器（例：麦克 MIC304）。

音频信号采集装置 130 可以包括 MIC、功放、音频信号采集仪、等等。

例如：由 MIC 接收被测系统 120 的输出端喇叭播放的经过该被测系统 120 传输过来的语音信号，具体地，这些从输入端进入的连续语音信号经过

被测系统 120 后，通过输出端作为输出的连续语音信号被 MIC 所接收；连续的语音信号，通过功放，传递到音频信号采集仪，由音频信号采集仪将这些连续的语音信号采集到的并传送给音频信号分析装置 140。

5 进一步，例如：由 MIC 接收被测系统 120 的输出端喇叭播放的经过该被测系统 120 传输过来的语音信号，可以由 MIC 将输出的语音信号转换成电信号，再由 MIC 中的 CPU 等处理器进行 A/D 转换等处理，形成数字信号，再将对应语音信号的数字信号传送到音频信号分析装置 140 进行数字信号的处理和分析。由于声源部分 110 的输入语音是连续的，通过被测系统 120 输出的语音信号也是连续的，因而，这些由音频信号采集装置 130 采集到的连续
10 的语音信号所对应的连续的数字信号可以传送到音频信号分析装置 140

音频信号分析装置 140，接收到从音频信号采集器 130 传送来的连续语音信号，或者说，连续语音信号所转换成的相应的连续数字信号，并进行处理和分析，进而，可以确定语音信号的还原度。

15 具体地，音频信号分析装置 140 可以包括内置的处理器（如 CPU 等）或者具有分析处理性能的 PC 机等等。当接收到连续语音信号时，通过 CPU 对连续的语音信号做转换形成连续的数字信号或者说语音的数字信号流；当接收到已经转换成的连续数字信号时，该连续数字信号即为语音的数字信号流。这里，将连续语音信号对应的数字信号流称为语音信号的信号流。

20 进一步，对语音信号的信号流进行切片，如：将信号流划分成 N “片”或者说 N “帧”（N 为大于等于 0 的正整数），再基于这些切片或帧，对采集到的所有测试所关心的所有连续信号进行处理和分析。在一个实施例中，可以基于哈斯效应的原理，考虑人耳在 20 毫秒（ms）的时间段中不会分辨出信号电平频率的先后，选择以时间间隔对信号流切片，每个切片/帧的时间为 20ms。进而，对每个切片/帧进行频谱转换，用转换后的切片/帧的频谱与
25 原始声源部分 110 输入端的语音信号的信号流对应的切片的频谱（即输入信号流的相应帧频谱）进行比较分析，以获得还原度测试/检测结果。

下面将结合图 2 所示本申请的人声语音质量评分方法一实施例的流程

图，对本申请的人声语音质量评分系统及其方法进行更具体的描述。

在步骤 210，声源部分 110 产生特定语音信号，该语音信号可以是标准的人说话的声音，如：国际电信联盟远程通信标准化组织 P.501 真人语音信号。该语音信号作为被测系统 120 的音频传输特性的测试用传输语音信号。

5 该语音信号可以经由声源部分 110 作为输入语音信号，输入到被测系统 120 的输入端，在被测系统 120 中传输，最后作为被测输出语音信号由被测系统 120 的输出端输出。真人语音信号包含所有的交杂失真，采用其作为输入信号，更符合被测系统的使用环境，其测试更准确客观。被测系统 120，例如：楼宇对讲系统。本步骤的具体实施可以参见对系统中关于声源部分 110 及其
10 被测系统 120 的描述。

在步骤 220，真人语音信号通过被测系统 120 输出的连续语音信号被音频信号采集装置 130 进行采集，并发给音频信号分析装置 130 分析。本步骤的具体实施可以参见对系统中关于音频信号采集装置 130 的描述。

15 在步骤 230，音频信号分析装置 140 对采集的音频信号（语音信号）进行切片，然后进行频谱转换，与声源部分 110 产生的特定语音信号进行比较分析，得到还原度测试结果。本步骤的具体实施可以参见对系统中关于音频信号分析装置 140 的描述。

下面将在一个实施方式中，对音频信号的切片处理做进一步描述。

20 首先，可以由该音频信号分析装置 140 的处理器（CPU）预先将标准语音信号的数字信号，即输入端的声源部分 110 产生的输入的连续语音信号对应的数字信号流，进行切片，每段 20ms（即“帧”），信号流切片成 N 端语音信号组（即 N 帧信号）。其中，每段语音信号组（N 帧信号）中包含一个或多个语音信号（或称信号参数），N 帧信号例如：P1，P2，P3...PN，存储这些切片后的帧，而每帧如 P1，由 20ms 时间段内数字信号组成。进而，
25 把每帧信号（即该信号流）转换成对应的频谱，并存储。其中每帧经转换后有相应的声音特征值。

然后，当该音频信号分析装置 140 接收到采集的信号流时，即接收到传

送来的从测试的输出端所采集的、对应该输入连续语音信号的信号流时，与该输入信号流进行同步后，同样，以每段 20ms 的方式切片成 N 段语音信号组，即 N 帧。其中，每帧信号即每段语音信号组中包含一个或多个语音信号（或称为信号参数），N 帧信号例如：p1, p2, p3...pN 存储这些切片后的帧，而每帧如 p1，也由 20ms 时间段内数字信号组成。进而，把每帧信号都转换成对应的频谱，并保存。其中每帧经转换后有相应的声音特征值。

其中，语音信号（如：语音信号的数字信号流），可以通过傅里叶变换或者其他声音信号处理实现频谱转换，同时得到各切片的的声音特征值。基于前述对应输入语音信号的经过被测系统 120 传输后的输出语音信号、与作为输入语音信号的标准语音信号一样，在二者同步后，进行同样的切片和频谱转换，得到每段信号组或者说每帧信号的声音特征值。

接着，提取这些声音特征值，即 P1, P2, P3...PN 中每帧对应的声音特征值以及 p1, p2, p3...pN 中每帧对应的声音特征值，基于诸如相似性原理/相似度计算等方式，对每个对应的帧即 P1 对应的声音特征值和 p1 对应的声音特征值进行相似度计算或相似性原理分析，确定还原度。如：P1 特征值与 p1 特征值之间的相似度计算，得到相似度值 0~1（0 相似或 100%相似），为了更清晰的确定分析结果，可以对该值的范围同乘以倍数如 100，采用百分制，即 0~100 分，从而每帧都能得到一个对比分析的分。

例如：P1 的特征值 A、B、C、D 与 p1 的特征值 a、b、c'、d 之间做一一匹配的相似性对比，得出 C 不同于 c'，只有 3 个相似，为 $3/4 * 100 = 75$ 分。

又例如：P1 的特征值为一维数组[A、B、C、D]，p1 的特征值为一维数组[a、b、c、d]，排列都是一维数组，并且，A~D、a~d 编号都为从小到大，则按相似性原理分析该排列趋势、动向相似/相同，因而，P1 与 p1 的频谱对比分数为 $(1/2 + 1/2) * 100 = 100$ 分。

最后，根据每帧的分数，确定所有帧的分数，从而确定还原度，即输出声音对输入声音的还原程度（与输入的相似程度）。例如：统计所有帧的分数并获得该输出语音信号的平均得分，该平均得分就是该被测系统 120 的还

原度得分。进一步，在信号流中还可能不存在间隙帧，而由于间隙帧不包含语音信息，为了减少干扰，可以把间隙帧的分数删除，而只考虑属于有效语音信号的帧的分数，然后统计有效的帧的分数得到该输出信号的平均得分，该平均得分就是该被测系统 120 的还原度得分。

- 5 利用本申请的方案，通过在检测过程中检测用的声源信号采用真人语音，以保证被测系统在检测过程中处于实际工作环境，并且，由于信号是真人语音，其失真包含所有交调失真；进而，通过对信号的切片处理，包括了对连续频率的测试，符合真人语音的信号状况，更能充分显现被测系统的声音输出特性，因此，直接对还原度的连续信号切片检测方式，能获得更准确、更
- 10 精确的被测系统、设备的检测结果。

 需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包

15 括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

 以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围

20 之内。

权 利 要 求 书

1、一种人声语音质量评分系统，至少包括：

声源部分、被测系统、音频信号采集装置、音频信号分析装置；

5 声源部分产生标准的真人语音信号，将所述真人语音信号作为输入语音信号，从声源部分输入到所述被测系统；

输入语音信号在被测系统中传递，并作为输出语音信号从被测系统的输出端输出；

音频信号采集装置采集输出端输出的连续的语音信号，并且，将采集的输出语音信号传送到音频信号分析装置；

10 音频信号分析装置对输出语音信号的信号流进行切片和分析处理，确定还原度。

2、如权利要求 1 所述的系统，其特征在于，所述音频信号采集装置，包括：将采集的连续语音信号转换成相应的数字信号流，传送给音频信号分析装置进行信号流的切片和分析处理。

15 3、如权利要求 1 或 2 所述的系统，其特征在于，音频信号分析装置，包括：

获取声源部分产生的作为输入语音信号的所述真人语音信号；

20 将所述输入语音信号的信号流，以时间间隔进行切片划分，以将信号流划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

4、如权利要求 3 所述的系统，其特征在于，音频信号分析装置，包括：

对来自音频信号采集装置采集的输出语音信号的信号流，与该输出语音信号的信号流相对应的来自声源部分的所述输入语音信号的信号流进行同

步;

5 以与切片所述输入语音信号的信号流相同的时间间隔,对采集的连续的语音信号的信号流切片划分成相同时间间隔的片段,每段语音信号组中包含有一个或多个语音信号,对每段语音信号组进行频谱转换,获得每段语音信号组转换后对应的声音特征值。

5、如权利要求4所述的系统,其特征在于,音频信号分析装置,包括:

10 提取属于所述输入语音信号的信号流中的一语音信号组的片段所对应的声音特征值,以及提取与该片段相应的、属于所述输出语音信号的信号流中的一片段所对应的声音特征值,基于相似性原理,计算和分析两个相应片段的对比分数值;

将所有相应的、分属于输入语音信号的信号流和相应的输出语音信号的信号流中的语音信号组的片段的对比分数值,进行统计和/或平均计算,以确定还原度。

6、一种人声语音质量评分方法,其特征在于,包括:

15 声源部分产生标准的真人语音信号,将所述真人语音信号作为输入语音信号,从声源部分输入到所述被测系统;

输入语音信号在被测系统中传递,并作为输出语音信号从被测系统的输出端输出;

采集输出端输出的连续的语音信号;

20 将采集的输出语音信号的信号流进行切片和分析处理,确定还原度。

7、如权利要求6所述的方法,其特征在于,还包括:

预先获取声源部分产生的作为输入语音信号的所述真人语音信号;

将所述输入语音信号的信号流,以时间间隔进行切片划分,以将信号流划分成相同时间间隔的片段,每段语音信号组中包含有一个或多个语音信号,

对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

8、如权利要求 7 所述的方法，其特征在于，将采集的输出语音信号的信号流进行切片和分析处理，包括：

5 对来自音频信号采集装置采集的输出语音信号的信号流，与该输出语音信号的信号流相对应的来自声源部分的所述输入语音信号的信号流进行同步；

10 以与切片所述输入语音信号的信号流相同的时间间隔，对采集的连续的语音信号的信号流切片划分成相同时间间隔的片段，每段语音信号组中包含有一个或多个语音信号，对每段语音信号组进行频谱转换，获得每段语音信号组转换后对应的声音特征值。

9、如权利要求 8 所述的方法，其特征在于，将采集的输出语音信号的信号流进行切片和分析处理，确定还原度，包括：

15 提取属于所述输入语音信号的信号流中的一语音信号组的片段所对应的声音特征值，以及提取与该片段相应的、属于所述输出语音信号的信号流中的一片段所对应的声音特征值，基于相似性原理，计算和分析两个相应片段的对比分数值；

20 将所有相应的、分属于输入语音信号的信号流和相应的输出语音信号的信号流中的语音信号组的片段的对比分数值，进行统计和/或平均计算，以确定还原度。

10、如权利要求 8 所述的方法，其特征在于，还包括：对输入语音信号的信号流和输出语音信号的信号流以相同的时间间隔进行切片，是以 20ms 时间间隔将信号流切片包含一个或多个语音信号的语音信号组的片段。

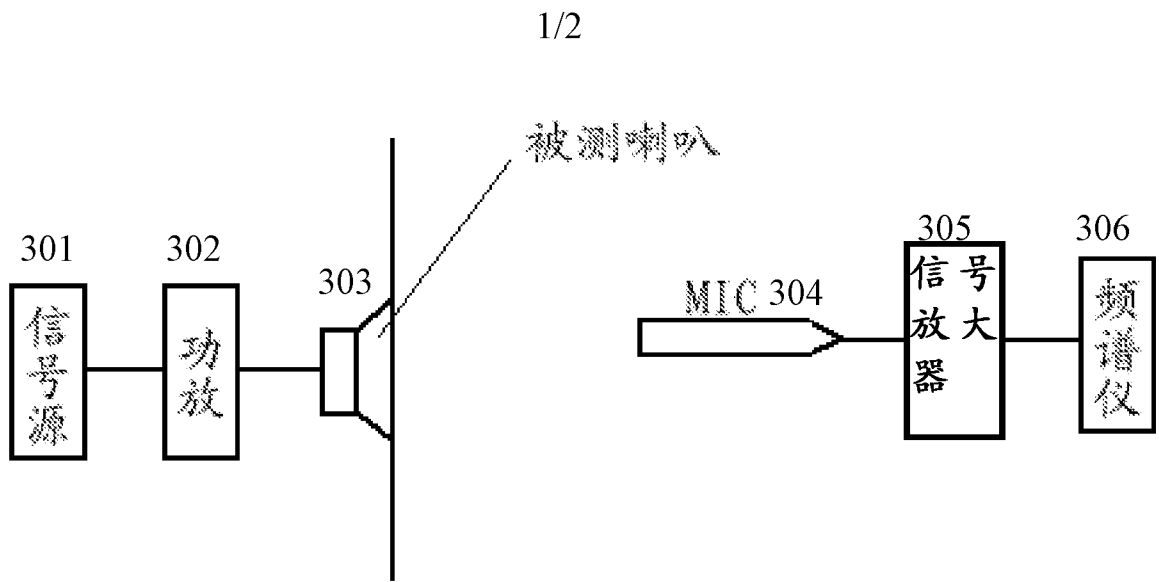


图 3



图 2

2/2

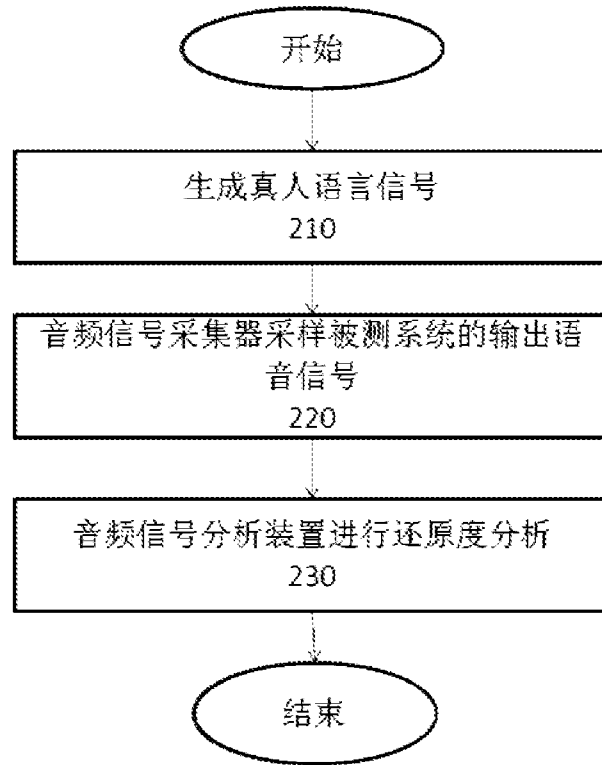


图 1

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/081156

A. CLASSIFICATION OF SUBJECT MATTER

H04B17/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04B17/-; H04Q5/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI: human voice, score, human speak, P501, P.50, subjective, subframe, section, reduction, synchronization, speech, voice, talk, test+, sync, MOS, mean opinion score, evaluat+, quality, frame, human

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 103607669 A (THE THIRD INSTITUTE OF MINISTRY OF PUBLIC SECURITY), 26 February 2014 (26.02.2014), description, paragraphs 0010-0011 and 0056-0066	1-10
Y	CN 101645271 A (INSTITUTE OF ACOUSTICS, CHINESE ACADEMY OF SCIENCES et al.), 10 February 2010 (10.02.2010), description, page 3, line 15 to page 5, line 5	1-10
A	CN 103413558 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS), 27 November 2013 (27.11.2013), the whole document	1-10
A	CN 103730131 A (HUAWEI TECHNOLOGIES CO., LTD.), 16 April 2014 (16.04.2014), the whole document	1-10
A	US 2007027687 A1 (VOXONIC INC.), 01 February 2007 (01.02.2007), the whole document	1-10
A	US 2010318635 A1 (SENDA, Y. et al.), 16 December 2010 (16.12.2010), the whole document	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
09 December 2014 (09.12.2014)

Date of mailing of the international search report
31 December 2014 (31.12.2014)

Name and mailing address of the ISA/CN:
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No.: (86-10) 62019451

Authorized officer
WANG, Yue
Telephone No.: (86-10) **62413592**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2014/081156

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103607669 A	26 February 2014	None	
CN 101645271 A	10 February 2010	CN 101645271 B	07 December 2011
CN 103413558 A	27 November 2013	None	
CN 103730131 A	16 April 2014	WO 2014056326 A1	17 April 2014
US 2007027687 A1	01 February 2007	JP 2008537600 A	18 September 2008
		EP 1859437 A2	28 November 2007
		WO 2006099467 A2	21 September 2006
		CN 101375329 A	25 February 2009
US 2010318635 A1	16 December 2010	CN 101960863 A	26 January 2011
		KR 20100108613 A	07 October 2010
		EP 2252072 A1	17 November 2010
		US 8676935 B2	18 March 2014
		JP 5413364 B2	12 February 2014
		WO 2009110238 A1	11 September 2009

国际检索报告

国际申请号

PCT/CN2014/081156

<p>A. 主题的分类</p> <p>H04B17/00(2006.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																																	
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04B17/-; H04Q5/-</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNPAT, WPI, EPODOC, CNKI:真人语音, 评分, 真人说话, 人声, 质量, 人发声, 测试, 检测, 评估, 人发音, P. 501, P. 50, 主观, 评测, 测评, 分帧, 切片, 还原, 同步, speech, voice, talk, test+, sync, MOS, mean opinion score, evaluat+, quality, frame, human</p>																																	
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 103607669 A (公安部第三研究所) 2014年 2月 26日 (2014 - 02 - 26) 说明书第0010-0011, 0056-0066段</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 101645271 A (中国科学院声学研究所 等) 2010年 2月 10日 (2010 - 02 - 10) 说明书第3页第15行至第5页第5行</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 103413558 A (南京邮电大学) 2013年 11月 27日 (2013 - 11 - 27) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 103730131 A (华为技术有限公司) 2014年 4月 16日 (2014 - 04 - 16) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2007027687 A1 (VOXONIC INC.) 2007年 2月 01日 (2007 - 02 - 01) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2010318635 A1 (SENDA, YUZO 等) 2010年 12月 16日 (2010 - 12 - 16) 全文</td> <td>1-10</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型:</p> <table border="0"> <tr> <td>“A” 认为不特别相关的表示了现有技术一般状态的文件</td> <td>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</td> </tr> <tr> <td>“E” 在国际申请日的当天或之后公布的在先申请或专利</td> <td>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</td> </tr> <tr> <td>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)</td> <td>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</td> </tr> <tr> <td>“O” 涉及口头公开、使用、展览或其他方式公开的文件</td> <td>“&” 同族专利的文件</td> </tr> <tr> <td>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</td> <td></td> </tr> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 103607669 A (公安部第三研究所) 2014年 2月 26日 (2014 - 02 - 26) 说明书第0010-0011, 0056-0066段	1-10	Y	CN 101645271 A (中国科学院声学研究所 等) 2010年 2月 10日 (2010 - 02 - 10) 说明书第3页第15行至第5页第5行	1-10	A	CN 103413558 A (南京邮电大学) 2013年 11月 27日 (2013 - 11 - 27) 全文	1-10	A	CN 103730131 A (华为技术有限公司) 2014年 4月 16日 (2014 - 04 - 16) 全文	1-10	A	US 2007027687 A1 (VOXONIC INC.) 2007年 2月 01日 (2007 - 02 - 01) 全文	1-10	A	US 2010318635 A1 (SENDA, YUZO 等) 2010年 12月 16日 (2010 - 12 - 16) 全文	1-10	“A” 认为不特别相关的表示了现有技术一般状态的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件	“E” 在国际申请日的当天或之后公布的在先申请或专利	“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性	“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)	“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性	“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件	“P” 公布日先于国际申请日但迟于所要求的优先权日的文件	
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																															
Y	CN 103607669 A (公安部第三研究所) 2014年 2月 26日 (2014 - 02 - 26) 说明书第0010-0011, 0056-0066段	1-10																															
Y	CN 101645271 A (中国科学院声学研究所 等) 2010年 2月 10日 (2010 - 02 - 10) 说明书第3页第15行至第5页第5行	1-10																															
A	CN 103413558 A (南京邮电大学) 2013年 11月 27日 (2013 - 11 - 27) 全文	1-10																															
A	CN 103730131 A (华为技术有限公司) 2014年 4月 16日 (2014 - 04 - 16) 全文	1-10																															
A	US 2007027687 A1 (VOXONIC INC.) 2007年 2月 01日 (2007 - 02 - 01) 全文	1-10																															
A	US 2010318635 A1 (SENDA, YUZO 等) 2010年 12月 16日 (2010 - 12 - 16) 全文	1-10																															
“A” 认为不特别相关的表示了现有技术一般状态的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件																																
“E” 在国际申请日的当天或之后公布的在先申请或专利	“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性																																
“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)	“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性																																
“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件																																
“P” 公布日先于国际申请日但迟于所要求的优先权日的文件																																	
国际检索实际完成的日期	国际检索报告邮寄日期																																
2014年 12月 09日	2014年 12月 31日																																
ISA/CN的名称和邮寄地址	授权官员																																
中华人民共和国国家知识产权局 (ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国	王玥																																
传真号 (86-10)62019451	电话号码 (86-10)62413592																																

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/081156

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	103607669	A	2014年 2月 26日	无			
CN	101645271	A	2010年 2月 10日	CN	101645271	B	2011年 12月 07日
CN	103413558	A	2013年 11月 27日	无			
CN	103730131	A	2014年 4月 16日	WO	2014056326	A1	2014年 4月 17日
US	2007027687	A1	2007年 2月 01日	JP	2008537600	A	2008年 9月 18日
				EP	1859437	A2	2007年 11月 28日
				WO	2006099467	A2	2006年 9月 21日
				CN	101375329	A	2009年 2月 25日
US	2010318635	A1	2010年 12月 16日	CN	101960863	A	2011年 1月 26日
				KR	20100108613	A	2010年 10月 07日
				EP	2252072	A1	2010年 11月 17日
				US	8676935	B2	2014年 3月 18日
				JP	5413364	B2	2014年 2月 12日
				WO	2009110238	A1	2009年 9月 11日

表 PCT/ISA/210 (同族专利附件) (2009年7月)