

(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
2011 年 5 月 19 日 (19.05.2011)

PCT

(10) 国际公布号
WO 2011/057511 A1

- (51) 国际专利分类号:
H04N 7/15 (2006.01) G10K 15/08 (2006.01)
H04M 3/56 (2006.01)
- (21) 国际申请号: PCT/CN2010/075891
- (22) 国际申请日: 2010 年 8 月 11 日 (11.08.2010)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
200910207184.8 2009 年 11 月 13 日 (13.11.2009) CN
- (71) 申请人 (对除美国外的所有指定国): 华为终端有限公司 (HUAWEI DEVICE CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地 B 区 2 号楼, Guangdong 518129 (CN)。
- (72) 发明人: 及
- (75) 发明人/申请人 (仅对美国): 詹五洲 (ZHAN, Wuzhou) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。王东琦 (WANG, Dongqi) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, 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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

根据细则 4.17 的声明:

- 关于申请人有权要求在先申请的优先权(细则 4.17(iii))

[见续页]

(54) Title: METHOD, APPARATUS AND SYSTEM FOR IMPLEMENTING AUDIO MIXING

(54) 发明名称: 实现混音的方法、装置和系统

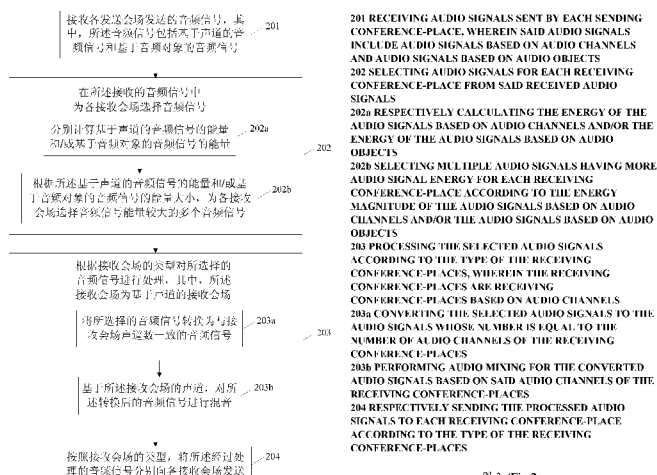


图 2 / Fig.2

(57) Abstract: A method, apparatus and system for implementing audio mixing are provided, and said method for implementing audio mixing includes: receiving audio signals sent by each sending conference-place, wherein said audio signals include audio signals based on audio channels and audio signals based on audio objects; selecting audio signals for each receiving conference-place from said received audio signals; processing the selected audio signals according to the type of the receiving conference-places, wherein the receiving conference-places include receiving conference-places based on audio channels and receiving conference-places based on audio objects; and respectively sending the processed audio signals to each receiving conference-place according to the type of the receiving conference-places. The present invention is applicable to audio mixing for the audio signals based on audio channels and the audio signals based on audio objects in the multipoint conference.

[见续页]

**本国际公布:**

— 包括国际检索报告(条约第 21 条(3))。

(57) 摘要:

一种实现混音的方法、装置和系统,所述实现混音的方法包括:接收各发送会场发送的音频信号,其中,所述音频信号包括基于声道的音频信号和基于音频对象的音频信号;在所述接收的音频信号中为各接收会场选择音频信号;根据接收会场的类型对所选择的音频信号进行处理,其中,所述接收会场包括基于声道的接收会场和基于音频对象的接收会场;按照接收会场的类型,将所述经过处理的音频信号分别向各接收会场发送。本发明适用于在多点会议中对基于声道的音频信号和基于音频对象的音频信号进行混音处理。

实现混音的方法、装置和系统

本申请要求于 2009 年 11 月 13 日提交中国专利局、申请号为 200910207184.8、发明名称为“实现混音的方法、装置和系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本发明涉及通信技术领域，特别涉及一种实现混音的方法、装置和系统。

背景技术

目前，随着通讯技术的发展，视讯会议得到了广泛的开展和应用。视讯会议可以理解为通常意义上的电视会议业务。通过多媒体通信手段，利用电视设
10 备和通信网络召开会议，可以同时实现两地或多个地点之间的图像、语音、数据的交互功能。通常，视讯会议系统包括视讯终端设备、通信网络、多点控制单元（Multipoint Control Unit, MCU）等几部分。

传统的会议终端通常只有单声道或双声道，一般没有空间方位感，或者只能区分左右方位。下一代会议终端一般采用多屏方案，图像与真人大小相同，
15 为了具有很强的临场感和沉浸感，一般要求有很强的声音方位感和空间感，传统的双声道已不能满足要求。为了体现较强的方位感和空间感，现有技术采用能够两种方案：一种是基于多声道的方式进行编码和传输；另外一种是基于音频对象的音频协议进行编码和传输，从而实现在增加较少码率的情况下，可以携带声音的方位和空间感。现有的MCU混音方法通常为基于声道的混音方案。

20 在实现本发明的过程中，发明人发现现有技术中至少存在如下问题：

现有的基于声道的混音方法，在同一会议中只能兼容传统的会议终端，而对于基于多声道的下一代终端以及基于音频对象的下一代终端，不能实现兼容。

发明内容

本发明的实施例提供一种实现混音的方法、装置和系统，能够提高对不同会议终端的兼容性。

5 本发明实施例采用的技术方案为：

一种实现混音的方法，包括：

接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

在所述接收的音频信号中为各接收会场选择音频信号；

10 根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

一种实现混音的装置，具体为多点控制单元，包括：

15 接收模块，用于接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

选择模块，用于在所述接收的音频信号中为各接收会场选择音频信号；

处理模块，用于根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

20 发送模块，用于按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

一种实现混音的系统，包括：多个发送会场和接收会场、多点控制单元，其中，

所述发送会场，用于向多点控制单元发送音频信号，其中，所述音频信号

包括基于声道的音频信号和基于音频对象的音频信号；

所述多点控制单元，用于接收所述发送会场发送的音频信号，在所述接收的音频信号中为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

所述接收会场，用于由所述多点控制单元接收所述经过处理的音频信号。

本发明实施例提供的实现混音的方法、装置和系统，多点控制单元接收到各发送会场发送的基于声道的音频信号和基于音频对象的音频信号后，为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，并将所述经过处理的音频信号分别向各接收会场发送。与现有技术相比，多点控制单元能够对基于声道的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

15

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

20

图1为本发明实施例一提供的实现混音的方法流程图；

图2、图3为本发明实施例二提供的实现混音的方法流程图；

图4为本发明实施例二提供的将所选择的音频信号转换为与基于声道的接

收会场声道数一致的音频信号的流程图；

图5为本发明实施例二提供的网真终端的音频呈现方式示意图；

图6为本发明实施例三提供的实现混音的装置结构示意图；

图7、图8为本发明实施例四提供的实现混音的装置结构示意图；

5 图9为本发明实施例五提供的实现混音的系统结构示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，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它实施例，都属于本发明保护的范围。

为使本发明技术方案的优点更加清楚，下面结合附图和实施例对本发明作详细说明。

实施例一

本实施例提供一种实现混音的方法，如图1所示，所述实现混音的方法包括：

15 101、接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

102、在所述接收的音频信号中为各接收会场选择音频信号；

103、根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

20 104、按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

其中，同一个会场既可以发送音频信号，也可以接收音频信号，也就是说，发送会场和接收会场可以为同一个会场。

本发明实施例实现混音的方法，多点控制单元接收到各发送会场发送的基于声道的音频信号和基于音频对象的音频信号后，为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，并将所述经过处理的音频信号分别向各接收会场发送。与现有技术相比，多点控制单元能够对基于声道
5 的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

实施例二

在本实施例中，多点会议系统中有多个发送会场和接收会场，所述发送会场
10 包括基于声道的发送会场和基于音频对象的发送会场，所述发送会场发出的音频信号由MCU进行混音；其中，同一个会场既可以发送音频信号，也可以接收音频信号，也就是说，发送会场和接收会场可以为同一个会场。

如图2、图3所示，所述实现混音的方法包括：

201、MCU接收各发送会场发送的音频信号，其中，所述音频信号包括基于
15 声道的音频信号和基于音频对象的音频信号；

其中，所述音频信号的类型根据发送会场的类型确定，若发送会场为基于声道的发送会场，则该发送会场发出的音频信号为基于声道的音频信号；若发送会场为基于音频对象的发送会场，则该发送会场发出的音频信号为基于音频对象的音频信号；所述基于声道的发送会场可以为基于单声道或多声道的发送
20 会场，相应地，所述基于声道的音频信号可以为基于单声道或多声道的音频信号。

音频对象指的是将音源作为一个对象来看待，一个音频对象除了包括音频信号之外，还包括辅助信息，所述辅助信息中包括该音频对象的最大绝对能量、

能量比值、空间信息、方位信息、不同播放方式及其对应参数等信息。

202、MCU在所述接收的音频信号中为各接收会场选择音频信号；

其中，所述MCU为各接收会场选择音频信号可以根据各音频信号的能量大小进行选择，具体选择过程可以包括：

- 5 202a、MCU分别计算基于声道的音频信号的能量和/或基于音频对象的音频信号的能量；

 (1) 计算基于声道的音频信号的能量

 当所述声道为单声道时，所述基于声道的音频信号的能量即为该单声道的能量；

- 10 当所述声道为多声道时，分别计算各个声道的能量，然后取最大的声道能量作为该基于声道的音频信号的能量；或者取各个声道的能量的平均值作为该基于声道的音频信号的能量。

 (2) 计算基于音频对象的音频信号的能量

 所述音频对象的辅助信息中携带最大绝对能量和该音频对象的能量比值，

- 15 根据音频对象中最大绝对能量以及该音频对象的能量比值，来计算该音频对象的绝对能量。例如，最大绝对能量为 $E_{\max}$ ，音频对象S1、S2、S3的能量比值分别是 a_1 、 a_2 、 a_3 ，则这三个音频对象的绝对能量分别是 $E_{\max} \times a_1$ 、 $E_{\max} \times a_2$ 、 $E_{\max} \times a_3$ 。

- 202b、MCU根据所述基于声道的音频信号的能量和/或基于音频对象的音频信号的能量大小，为各接收会场选择音频信号能量较大的多个音频信号；其中，MCU为各接收会场选择的音频信号可以相同，也可以不同。

 例如，多点会议系统中有A、B、C、D、E五个会场，所述五个会场既可以接收音频信号，也可以发送音频信号，A、B、C、D、E五个会场发出的音频信号对

应为A1、B1、C1、D1、E1；MCU根据各会场发出的音频信号的能量大小选择出B1、C1、D1三个音频信号，由于各会场通常不接收自己会场发出的音频信号，因此，MCU为各会场选择的音频信号如表1所示：

接收会场	选择的音频信号
A	B1、C1、D1
B	C1、D1
C	B1、D1
D	B1、C1
E	B1、C1、D1

表1

5 进一步地，为了实现多点会议中两个会场的私聊，对于需要进行私聊的两个会场发出的音频信号，只能由私聊的另外一方来选择，而不能被其它会场选择，因此，MCU为各会场选择的音频信号如表2所示：

接收会场	选择的音频信号
A	B1、C1、D1
B	A1、C1、D1
C	D1
D	C1
E	C1、D1

表2

203、根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收
10 会场可以为基于声道的接收会场或基于音频对象的接收会场；

（1）当所述接收会场为基于声道的接收会场时，如图2所示，所述根据接

收会场的类型对所选择的音频信号进行处理包括:

203a、将所选择的音频信号转换为与接收会场声道数一致的音频信号;

如图4所示,所述将所选择的音频信号转换为与接收会场声道数一致的音频信号具体可以包括:

5 L1、判断所选择的音频信号的类型,若所选择的音频信号为基于声道的音频信号,执行步骤L2;若所选择的音频信号为基于音频对象的音频信号,执行步骤L3;

L2、将所述基于声道的音频信号转换成与所述接收会场声道数一致的信号;

所述基于声道的会场其声道数可以是单或多声道,对于多声道,一般是将
10 麦克风放置于不同空间位置,然后对各个麦克风采集的信号各用一个声道进行编码,由于声道之间已经包含了空间信息,因此在接收会场用相同数目的扬声器进行播放时,就可以呈现出原始发送会场的声音空间信息。

例如,所述接收会场为双声道信号,MCU为该接收会场选择的基于声道的音频信号为单声道信号和三声道信号,则MCU将所述单声道信号同时复制到该接收
15 会场的左声道和右声道,即左右声道信号内容相同;而对于所述三声道信号,MCU将第一声道信号复制到该接收会场的左声道,将第三声道信号复制到该接收会场的右声道,将第二声道信号乘上0.707的增益,然后同时加到该接收会场的左声道和右声道上,这样所述单声道信号和三声道信号都转换成了双声道信号。

L3、判断所述音频对象的辅助信息中是否携带与所述接收会场声道数一致
20 的播放方式,若所述音频对象的辅助信息中携带与所述接收会场声道数一致的播放方式,执行步骤L4;若所述音频对象的辅助信息中没有携带与所述接收会场声道数一致的播放方式,执行步骤L5;

其中,所述音频对象的辅助信息中携带多种播放方式及其对应参数;所述

播放方式指的是采用几声道进行播放，例如：双声道、5声道等，每种播放方式的参数表示所述音频对象在各个声道的能量分配情况，该能量分配可以随时间变化。

L4、根据所述辅助信息中相应播放方式的参数，将所述音频对象转换成与
5 所述接收会场声道数一致的音频信号；

例如，所述接收会场为双声道，且所述辅助信息中携带双声道的播放方式，则MCU提取双声道的播放方式对应的参数，根据该参数将所述音频对象信号分配到该接收会场的左右声道上。

L5、根据所述音频对象的方位信息将所述音频对象转换成与接收会场声道
10 数一致的信号，其中，所述方位信息在所述音频对象的辅助信息中携带。

例如，所述辅助信息中只有双声道的播放方式和5声道的播放方式，而所述接收会场为6声道，则MCU根据所述音频对象的方位信息将所述音频对象转换成6声道的音频信号。

其中，所述根据所述音频对象的方位信息将所述音频对象转换成与接收会
15 场声道数一致的信号具体可以为：根据音频对象的方位和所述接收会场各个声道对应的扬声器的位置，确定与所述音频对象距离最近的扬声器；将所述基于音频对象的音频信号复制给与所述音频对象距离最近的扬声器对应的声道，而其它声道不赋予任何信号。

203b、基于所述接收会场的声道，对所述转换后的音频信号进行混音。

20 (2) 当所述接收会场为基于音频对象的接收会场时，如图3所示，所述根据接收会场的类型对所选择的音频信号进行处理包括：

203c、根据接收会场的呈现方式，将所选择的音频信号转换成音频对象；

其中，所述根据接收会场的呈现方式，将所选择的音频信号转换成音频对

象具体可以包括:

S1、根据接收会场的呈现方式, 设定所选择的音频信号的呈现方式;

例如, 如图5所示, 为一种网真终端的音频呈现方式, 图像显示屏幕由三个
5 屏幕组合而成, 在每个屏幕下方有两个扬声器, 在屏幕两侧各有一个扬声器,
共有8个扬声器。为了实现图像和声音的匹配, 当屏幕中某个位置有人在发言时,
由该位置下方附近的扬声器发出声音, 从而实现声像匹配。在多点会议时, 正
在发言的会场可能并不是当前屏幕中显示的会场, 此时该会场的声音可以分配
到屏幕两侧的两个扬声器。例如, 有三个音频信号s1、s2、s3被选择, 其中音
10 频信号s1对应的会场的视频信号正在被该接收会场观看, 则可设定音频信号s1
由屏幕下方的6个扬声器来播放, 音频信号s2、s3对应的会场的视频信号当前并
没有被观看, 则可指定音频信号s2、s3分别由屏幕两侧的两个扬声器来播放。

S2、当所选择的音频信号为基于声道的音频信号时, 根据所述设定的呈现
方式, 将所选择的音频信号转换成音频对象;

S3、当所选择的音频信号为基于音频对象的音频信号时, 根据所述设定的
15 呈现方式, 对所述音频对象原来的相关参数进行修改, 以满足所设定的呈现方
式的要求。

203d、将所述转换后的音频对象合并为一个音频对象流。

204、按照接收会场的类型, 将所述经过处理的音频信号分别向各接收会场
发送。

20 本发明实施例实现混音的方法, 由多点控制单元接收到各发送会场发送的
基于声道的音频信号和基于音频对象的音频信号后, 为各接收会场选择音频信
号, 根据接收会场的类型对所选择的音频信号进行处理, 并将所述经过处理的
音频信号分别向各接收会场发送。与现有技术相比, 多点控制单元能够对基于

声道的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

实施例三

5 本实施例提供一种实现混音的装置，如图6所示，所述实现混音的装置，包括：

接收模块61，用于接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

其中，所述音频信号的类型根据发送会场的类型确定，若发送会场为基于
10 声道的发送会场，则该发送会场发出的音频信号为基于声道的音频信号；若发送会场为基于音频对象的发送会场，则该发送会场发出的音频信号为基于音频对象的音频信号；所述基于声道的发送会场可以为基于单声道或多声道的发送会场，相应地，所述基于声道的音频信号可以为基于单声道或多声道的音频信号。音频对象指的是将音源作为一个对象来看待，一个音频对象除了包括音频
15 信号之外，还包括辅助信息，所述辅助信息中包括该音频对象的最大绝对能量、能量比值、空间信息、方位信息、不同播放方式及其对应参数等信息。

选择模块62，用于在所述接收的音频信号中为各接收会场选择音频信号；

处理模块63，用于根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

20 发送模块64，用于按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

其中，同一个会场既可以发送音频信号，也可以接收音频信号，也就是说，发送会场和接收会场可以为同一个会场。

本发明实施例实现混音的装置，多点控制单元接收到各发送会场发送的基于声道的音频信号和基于音频对象的音频信号后，为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，并将所述经过处理的音频信号分别向各接收会场发送。与现有技术相比，多点控制单元能够对基于声道
5 的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

实施例四

如图7、图8所示，所述实现混音的装置，包括：

10 接收模块61，用于接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

其中，所述音频信号的类型根据发送会场的类型确定，若发送会场为基于声道的发送会场，则该发送会场发出的音频信号为基于声道的音频信号；若发送会场为基于音频对象的发送会场，则该发送会场发出的音频信号为基于音频
15 对象的音频信号；所述基于声道的发送会场可以为基于单声道或多声道的发送会场，相应地，所述基于声道的音频信号可以为基于单声道或多声道的音频信号。音频对象指的是将音源作为一个对象来看待，一个音频对象除了包括音频信号之外，还包括辅助信息，所述辅助信息中包括该音频对象的最大绝对能量、能量比值、空间信息、方位信息、不同播放方式及其对应参数等信息。

20 选择模块62，用于在所述接收的音频信号中为各接收会场选择音频信号；

处理模块63，用于根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

发送模块64，用于按照接收会场的类型，将所述经过处理的音频信号分别

向各接收会场发送。

其中，同一个会场既可以发送音频信号，也可以接收音频信号，也就是说，发送会场和接收会场可以为同一个会场。

其中，所述选择模块62包括：

- 5 计算单元621，用于分别计算基于声道的音频信号的能量和/或基于音频对象的音频信号的能量；

选择单元622，用于根据所述基于声道的音频信号的能量和/或基于音频对象的音频信号的能量选择音频信号。所述选择单元622可以根据所述基于声道的音频信号的能量和基于音频对象的音频信号的能量大小，为各接收会场选择音
10 频信号能量较大的多个音频信号；所述选择单元622为各接收会场选择的音频信号可以相同，也可以不同。

当所述接收会场为基于声道的接收会场时，如图7所示，所述处理模块63可以包括：

- 第一转换单元631，用于将所选择的音频信号转换为与接收会场声道数一致
15 的音频信号；

混音单元632，用于基于所述接收会场的声道，对所述转换后的音频信号进行混音。

进一步，所述第一转换单元631可以包括：

第一判断子单元6311，用于判断所选择的音频信号的类型；

- 20 第一转换子单元6312，用于当所选择的音频信号为基于声道的音频信号时，将所述基于声道的音频信号转换成与所述接收会场声道数一致的信号；

第二判断子单元6313，用于当所选择的音频信号为基于音频对象的音频信号时，判断所述音频对象的辅助信息中是否携带与所述接收会场声道数一致的

播放方式;

第二转换子单元6314, 用于当所述音频对象的辅助信息中携带与所述接收会场声道数一致的播放方式时, 根据所述辅助信息中相应播放方式的参数, 将所述音频对象转换成与所述接收会场声道数一致的音频信号;

- 5 第三转换子单元6315, 用于当所述音频对象的辅助信息中没有携带与所述接收会场声道数一致的播放方式时, 根据所述音频对象的方位信息将所述音频对象转换成与接收会场声道数一致的信号, 其中, 所述方位信息在所述音频对象的辅助信息中携带。

当所述接收会场为基于音频对象的接收会场时, 如图8所示, 所述处理模块

- 10 63可以包括:

第二转换单元633, 用于根据接收会场的呈现方式, 将所选择的音频信号转换成音频对象;

合并单元634, 用于将所述转换后的音频对象合并为一个音频对象流。

进一步, 所述第二转换单元633可以包括:

- 15 设定子单元6331, 用于根据接收会场的呈现方式, 设定所选择的音频信号的呈现方式;

第四转换子单元6332, 用于当所选择的音频信号为基于声道的音频信号时, 根据所述设定的呈现方式, 将所选择的音频信号转换成音频对象;

修改子单元6333, 用于当所选择的音频信号为基于音频对象的音频信号时,

- 20 根据所述设定的呈现方式, 对所述音频对象的参数进行修改。

本发明实施例实现混音的装置接收到各发送会场发送的基于声道的音频信号和基于音频对象的音频信号后, 为各接收会场选择音频信号, 根据接收会场的类型对所选择的音频信号进行处理, 并将所述经过处理的音频信号分别向各

接收会场发送。与现有技术相比，该装置能够对基于声道的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

5 实施例五

本实施例提供一种实现混音的系统，如图9所示，所述实现混音的系统，包括：多个发送会场91和接收会场93、多点控制单元92，其中，

所述发送会场91，用于向多点控制单元92发送音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

10 所述多点控制单元92，用于接收所述发送会场91发送的音频信号，在所述接收的音频信号中为各接收会场93选择音频信号，根据接收会场93的类型对所选择的音频信号进行处理，按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场93发送，其中，所述接收会场93包括基于声道的接收会场和基于音频对象的接收会场；

15 所述接收会场93，用于由所述多点控制单元92接收所述经过处理的音频信号。

其中，同一个会场既可以发送音频信号，也可以接收音频信号，也就是说，发送会场和接收会场可以为同一个会场。

本发明实施例实现混音的系统中，多点控制单元接收到各发送会场发送的
20 基于声道的音频信号和基于音频对象的音频信号后，为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，并将所述经过处理的音频信号分别向各接收会场发送。与现有技术相比，多点控制单元能够对基于声道的音频信号和基于音频对象的音频信号进行混音处理，从而在同一多点会

议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，提高用户的体验质量。

本发明实施例提供的实现混音的装置和系统可以实现上述提供的方法实施例。本发明实施例提供的实现混音的方法、装置和系统可以适用于在同一多点
5 会议中兼容传统的会议终端、基于多声道的下一代终端以及基于音频对象的下一代终端，但不仅限于此。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。
10 其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）或随机存储记忆体（Random Access Memory, RAM）等。

以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到
的变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围
15 应该以权利要求的保护范围为准。

权 利 要 求

1、一种实现混音的方法，其特征在于，包括：

接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

5 在所述接收的音频信号中为各接收会场选择音频信号；

根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

2、根据权利要求1所述的实现混音的方法，其特征在于，所述在所述接收
10 的音频信号中为各接收会场选择音频信号具体为：

分别计算基于声道的音频信号的能量和/或基于音频对象的音频信号的能量；

根据所述基于声道的音频信号的能量和/或基于音频对象的音频信号的能量选择音频信号。

15 3、根据权利要求2所述的实现混音的方法，其特征在于，所述计算基于声道的音频信号的能量包括：

当所述声道为单声道时，将所述基于声道的音频信号的能量作为该单声道的能量；

当所述声道为多声道时，分别计算各个声道的能量，取最大的声道能量作为该基于声道的音频信号的能量，或者取各个声道的能量的平均值作为该基于
20 声道的音频信号的能量。

4、根据权利要求2所述的实现混音的方法，其特征在于，所述计算基于音频对象的音频信号的能量包括：

根据最大绝对能量和所述音频对象的能量比值，计算所述音频对象的绝对能量，其中，所述最大绝对能量和所述音频对象的能量比值在所述音频对象的辅助信息中携带。

5、根据权利要求1所述的实现混音的方法，其特征在于，当所述接收会场
5 为基于声道的接收会场时，所述根据接收会场的类型对所选择的音频信号进行处理包括：

将所选择的音频信号转换为与接收会场声道数一致的音频信号；

基于所述接收会场的声道，对所述转换后的音频信号进行混音。

6、根据权利要求5所述的实现混音的方法，其特征在于，所述将所选择的
10 音频信号转换为与接收会场声道数一致的音频信号包括：

判断所选择的音频信号的类型；

若所选择的音频信号为基于声道的音频信号，则将所述基于声道的音频信号转换成与所述接收会场声道数一致的信号；

若所选择的音频信号为基于音频对象的音频信号，则判断所述音频对象的
15 辅助信息中是否携带与所述接收会场声道数一致的播放方式；

若所述音频对象的辅助信息中携带与所述接收会场声道数一致的播放方式，根据所述辅助信息中相应播放方式的参数，将所述音频对象转换成与所述接收会场声道数一致的音频信号；

若所述音频对象的辅助信息中没有携带与所述接收会场声道数一致的播放
20 方式，则根据所述音频对象的方位信息将所述音频对象转换成与接收会场声道数一致的信号，其中，所述方位信息在所述音频对象的辅助信息中携带。

7、根据权利要求6所述的实现混音的方法，其特征在于，所述根据所述音频对象的方位信息将所述音频对象转换成与接收会场声道数一致的信号包括：

根据音频对象的方位和所述接收会场各个声道对应的扬声器的位置，确定与所述音频对象距离最近的扬声器；

将所述基于音频对象的音频信号复制给与所述音频对象距离最近的扬声器对应的声道。

- 5 8、根据权利要求1所述的实现混音的方法，其特征在于，当所述接收会场为基于音频对象的接收会场时，所述根据接收会场的类型对所选择的音频信号进行处理包括：

根据接收会场的呈现方式，将所选择的音频信号转换成音频对象；

将所述转换后的音频对象合并为一个音频对象流。

- 10 9、根据权利要求8所述的实现混音的方法，其特征在于，所述根据接收会场的呈现方式，将所选择的音频信号转换成音频对象包括：

根据接收会场的呈现方式，设定所选择的音频信号的呈现方式；

当所选择的音频信号为基于声道的音频信号时，根据所述设定的呈现方式，将所选择的音频信号转换成音频对象；

- 15 当所选择的音频信号为基于音频对象的音频信号时，根据所述设定的呈现方式，对所述音频对象的参数进行修改。

10、一种实现混音的装置，具体为多点控制单元，其特征在于，所述装置包括：

- 20 接收模块，用于接收各发送会场发送的音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

选择模块，用于在所述接收的音频信号中为各接收会场选择音频信号；

处理模块，用于根据接收会场的类型对所选择的音频信号进行处理，其中，所述接收会场包括基于声道的接收会场和基于音频对象的接收会场；

发送模块，用于按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送。

11、根据权利要求10所述的实现混音的装置，其特征在于，所述选择模块包括：

5 计算单元，用于分别计算基于声道的音频信号的能量和基于音频对象的音频信号的能量；

选择单元，用于根据所述基于声道的音频信号的能量和基于音频对象的音频信号的能量选择音频信号。

12、根据权利要求10所述的实现混音的装置，其特征在于，所述处理模块
10 包括：

第一转换单元，用于将所选择的音频信号转换为与接收会场声道数一致的音频信号；

混音单元，用于基于所述接收会场的声道，对所述转换后的音频信号进行混音。

15 13、根据权利要求12所述的实现混音的装置，其特征在于，所述第一转换单元包括：

第一判断子单元，用于判断所选择的音频信号的类型；

第一转换子单元，用于当所选择的音频信号为基于声道的音频信号时，将所述基于声道的音频信号转换成与所述接收会场声道数一致的信号；

20 第二判断子单元，用于当所选择的音频信号为基于音频对象的音频信号时，判断所述音频对象的辅助信息中是否携带与所述接收会场声道数一致的播放方式；

第二转换子单元，用于当所述音频对象的辅助信息中携带与所述接收会场

声道数一致的播放方式时，根据所述辅助信息中相应播放方式的参数，将所述音频对象转换成与所述接收会场声道数一致的音频信号；

第三转换子单元，用于当所述音频对象的辅助信息中没有携带与所述接收会场声道数一致的播放方式时，根据所述音频对象的方位信息将所述音频对象
5 转换成与接收会场声道数一致的信号，其中，所述方位信息在所述音频对象的辅助信息中携带。

14、根据权利要求10所述的实现混音的装置，其特征在于，所述处理模块包括：

第二转换单元，用于根据接收会场的呈现方式，将所选择的音频信号转换
10 成音频对象；

合并单元，用于将所述转换后的音频对象合并为一个音频对象流。

15、根据权利要求14所述的实现混音的装置，其特征在于，所述第二转换单元包括：

设定子单元，用于根据接收会场的呈现方式，设定所选择的音频信号的呈
15 现方式；

第四转换子单元，用于当所选择的音频信号为基于声道的音频信号时，根据所述设定的呈现方式，将所选择的音频信号转换成音频对象；

修改子单元，用于当所选择的音频信号为基于音频对象的音频信号时，根据所述设定的呈现方式，对所述音频对象的参数进行修改。

20 16、一种实现混音的系统，其特征在于，包括：多个发送会场和接收会场、多点控制单元，其中，

所述发送会场，用于向多点控制单元发送音频信号，其中，所述音频信号包括基于声道的音频信号和基于音频对象的音频信号；

所述多点控制单元，用于接收所述发送会场发送的音频信号，在所述接收的音频信号中为各接收会场选择音频信号，根据接收会场的类型对所选择的音频信号进行处理，按照接收会场的类型，将所述经过处理的音频信号分别向各接收会场发送，其中，所述接收会场包括基于声道的接收会场和基于音频对象
5 的接收会场；

所述接收会场，用于由所述多点控制单元接收所述经过处理的音频信号。

10

15

20

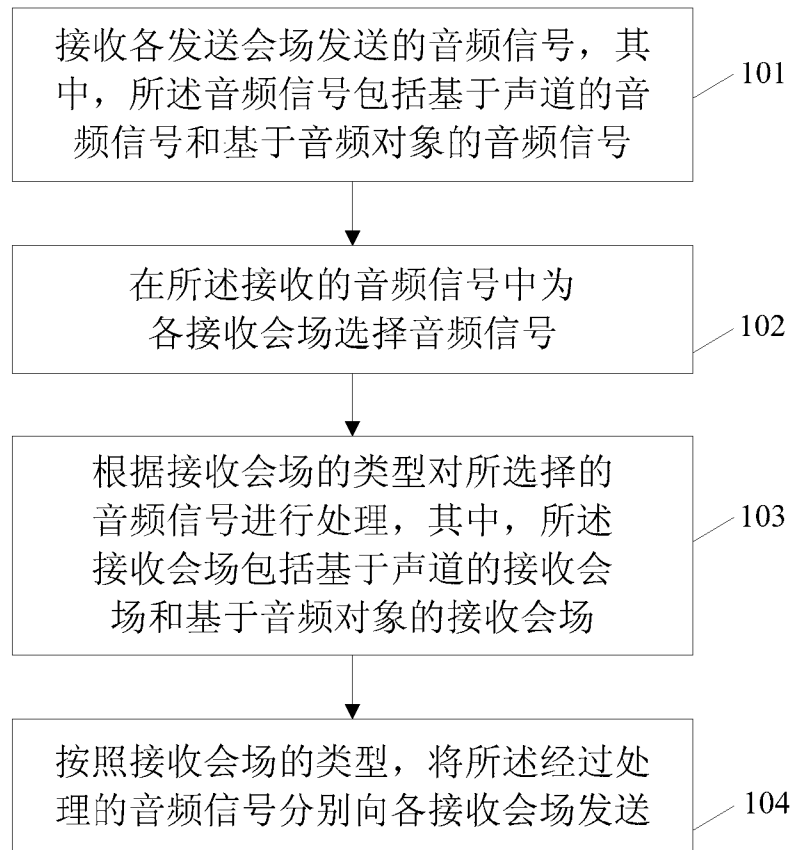


图 1

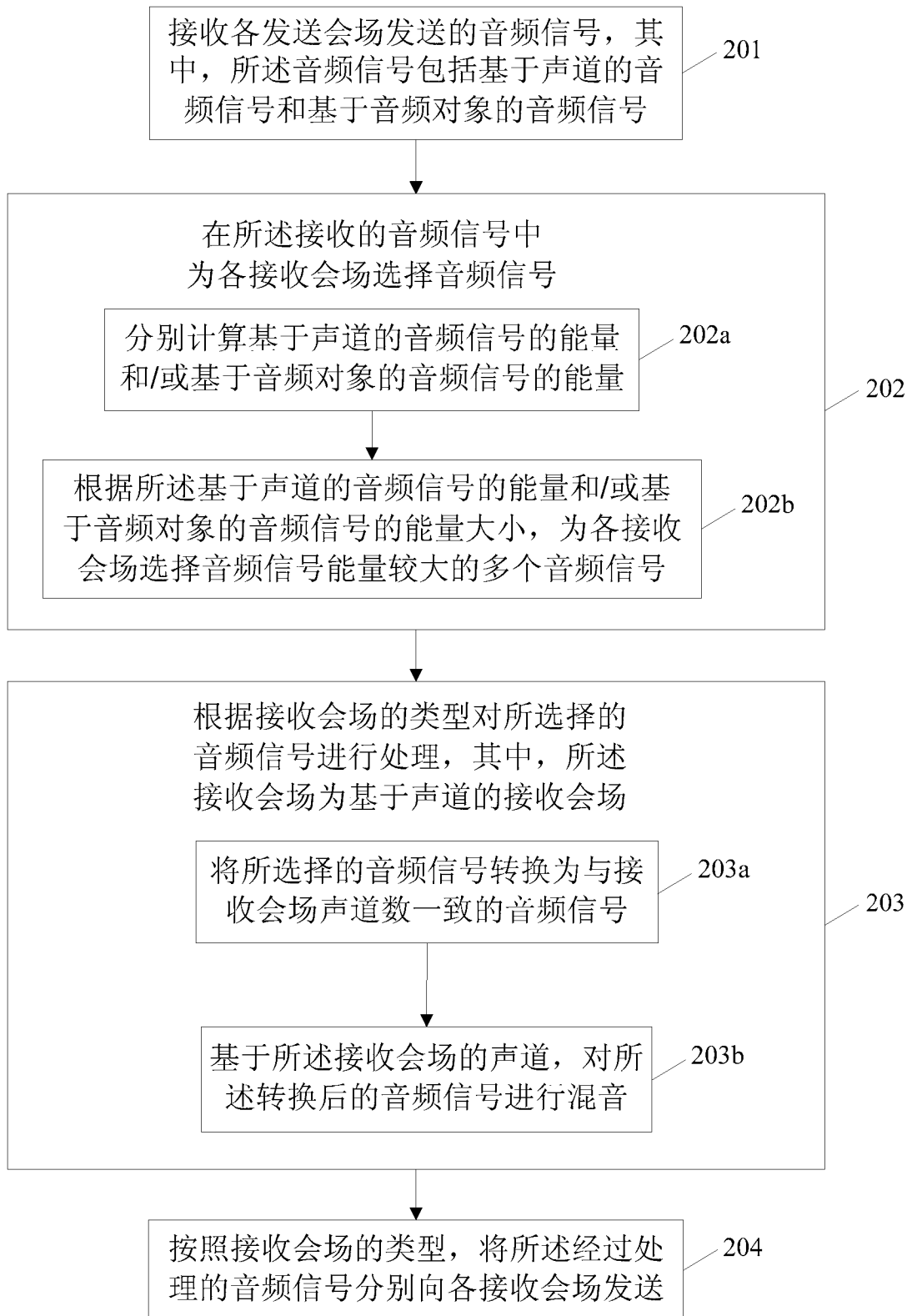


图 2

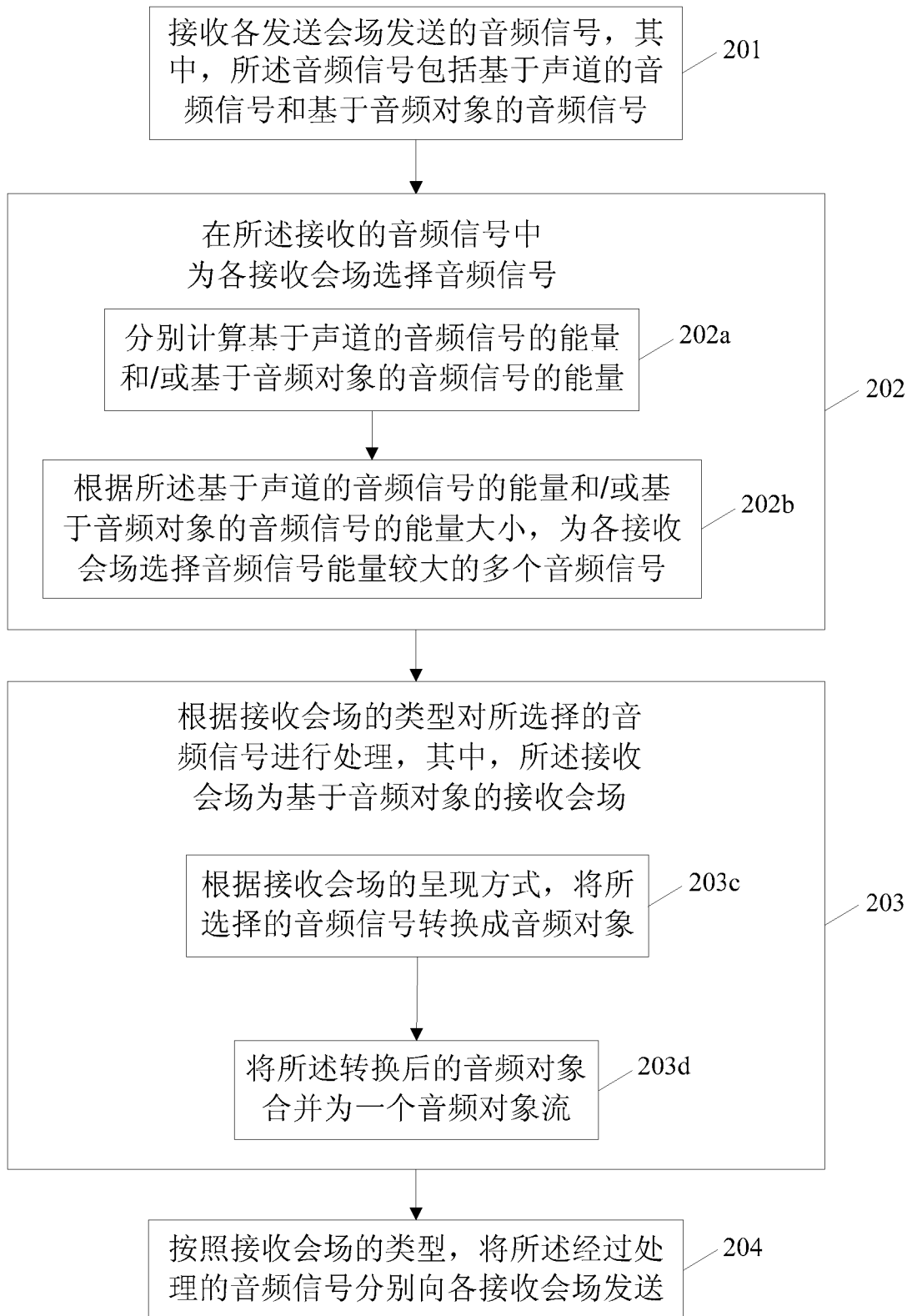


图 3

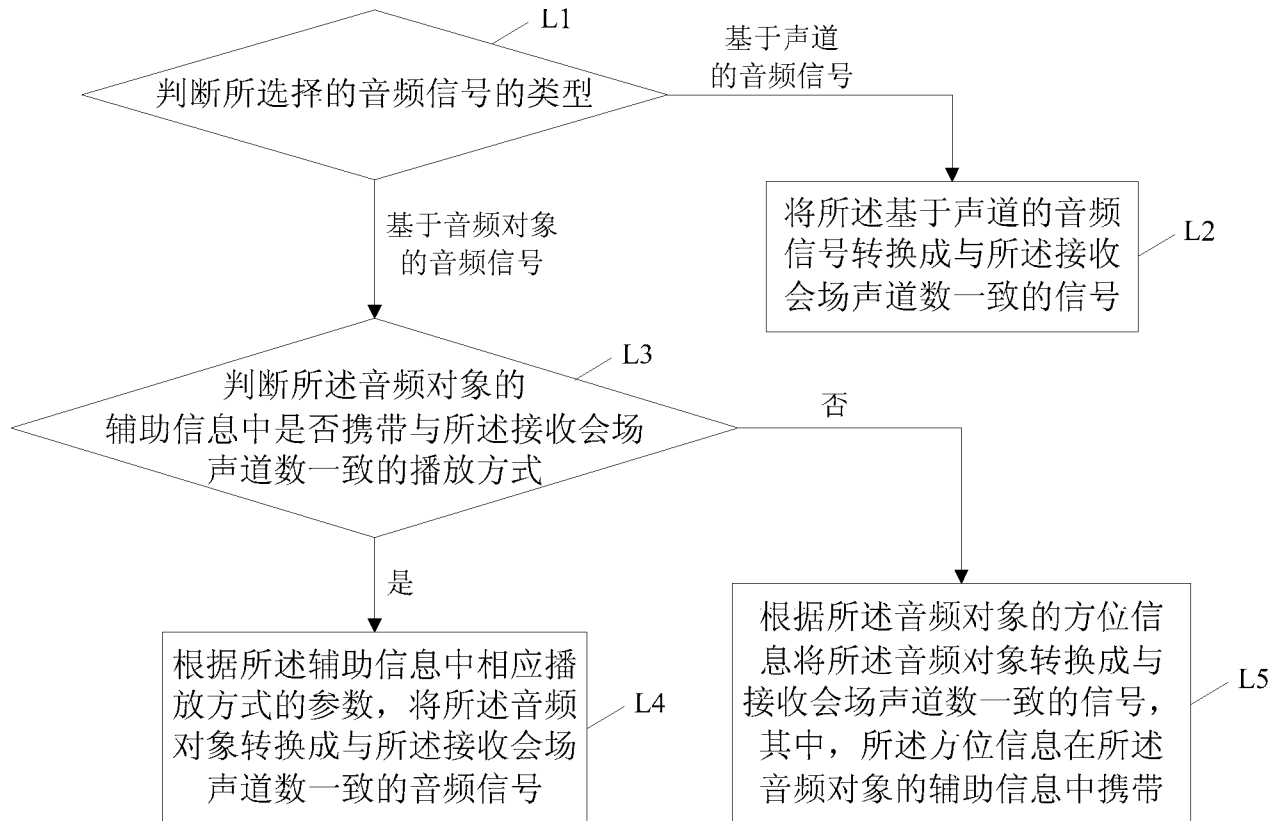


图 4

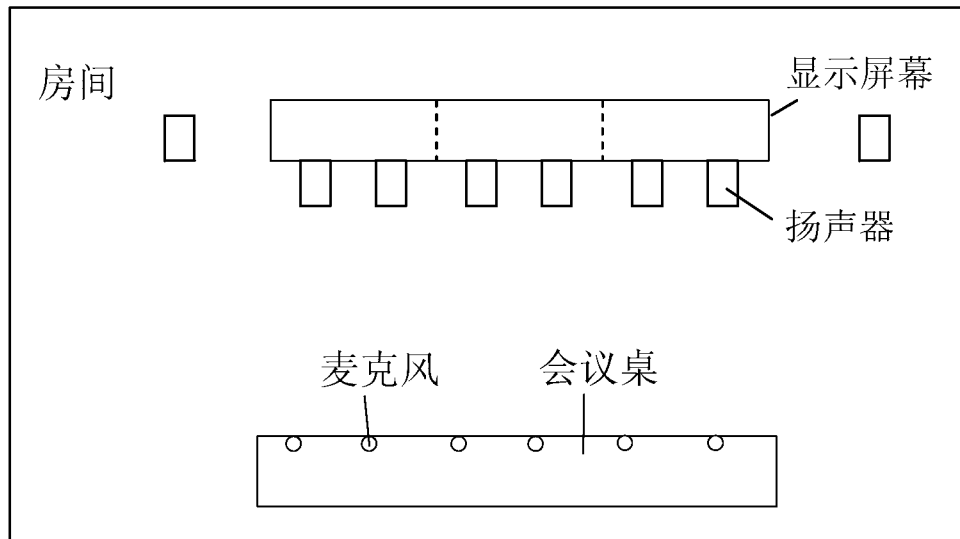


图 5

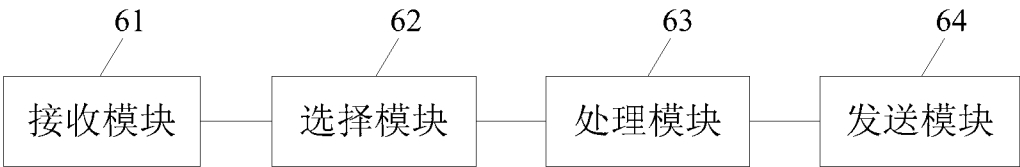


图 6

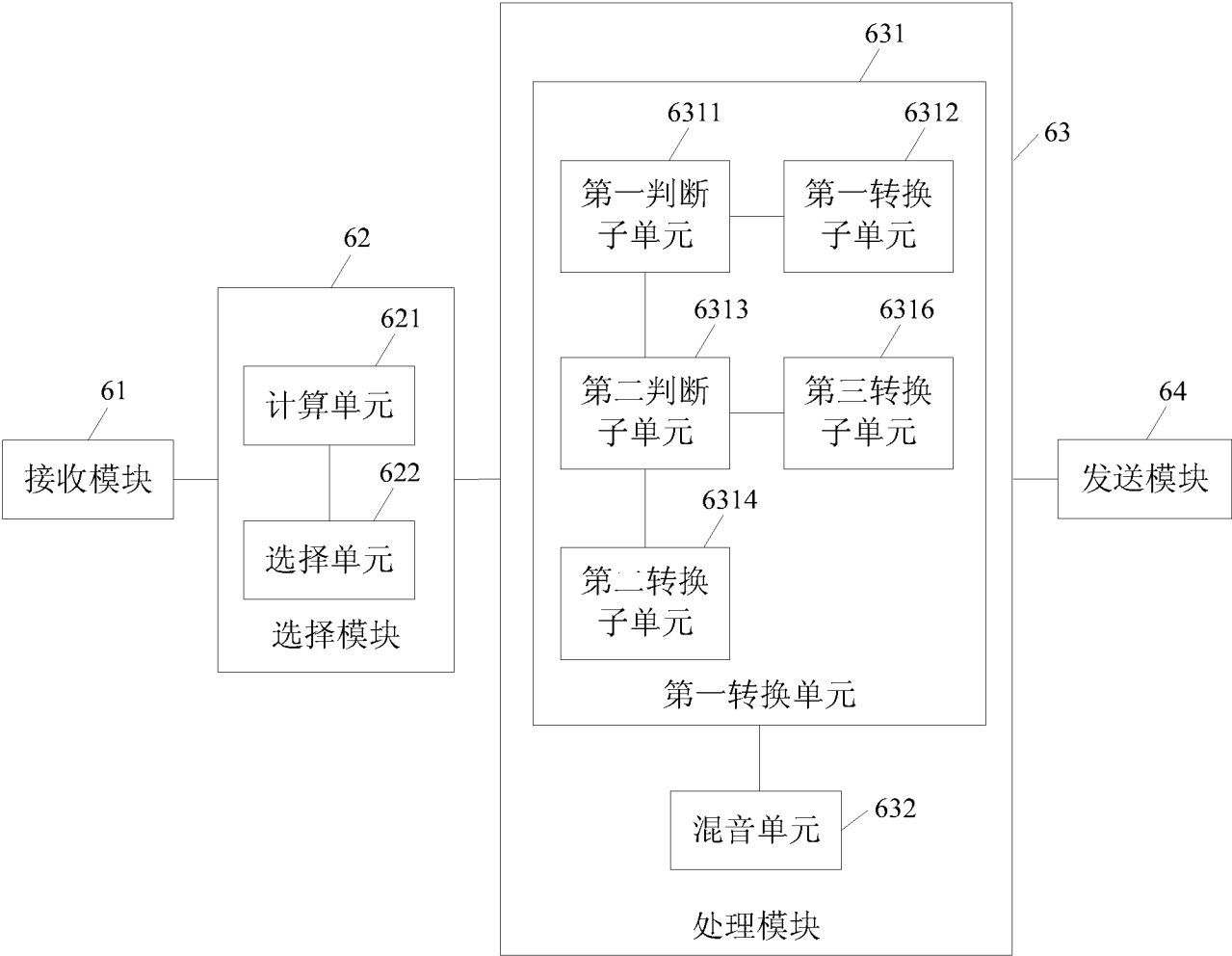


图 7

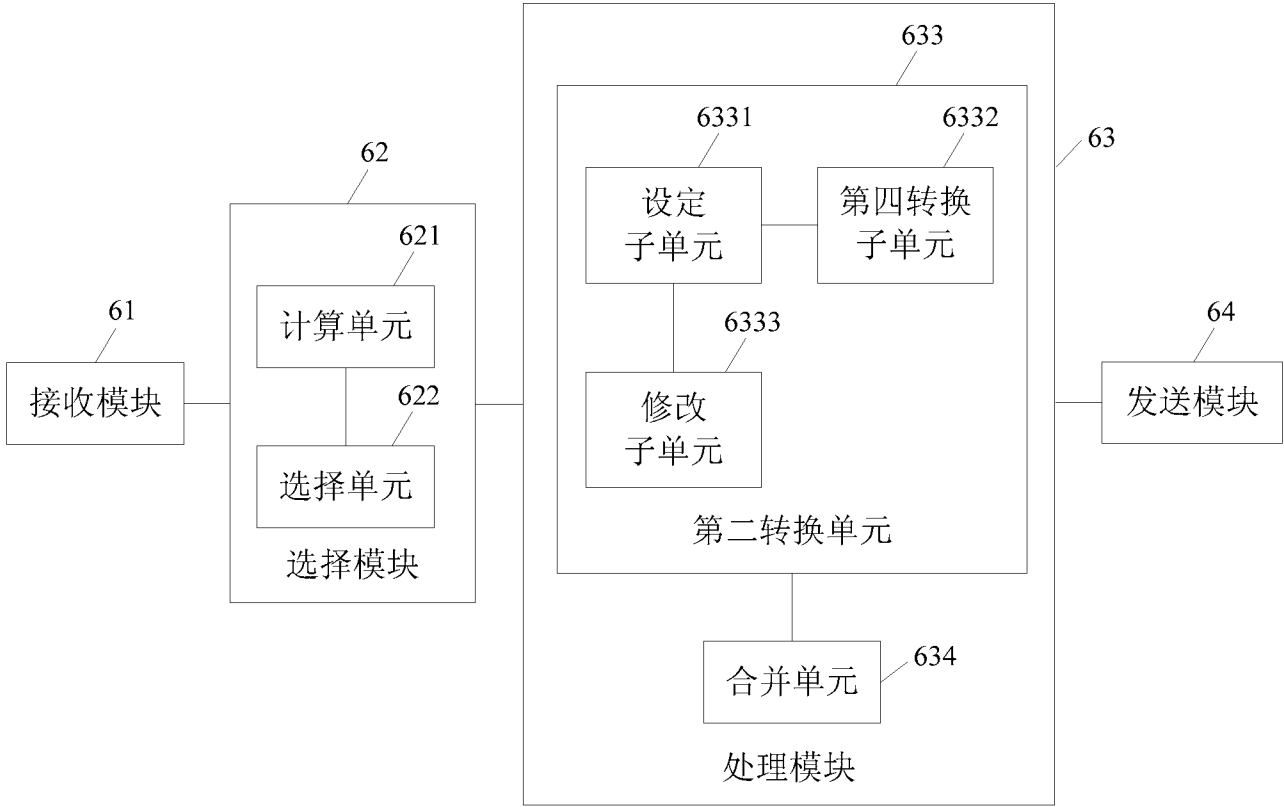


图 8

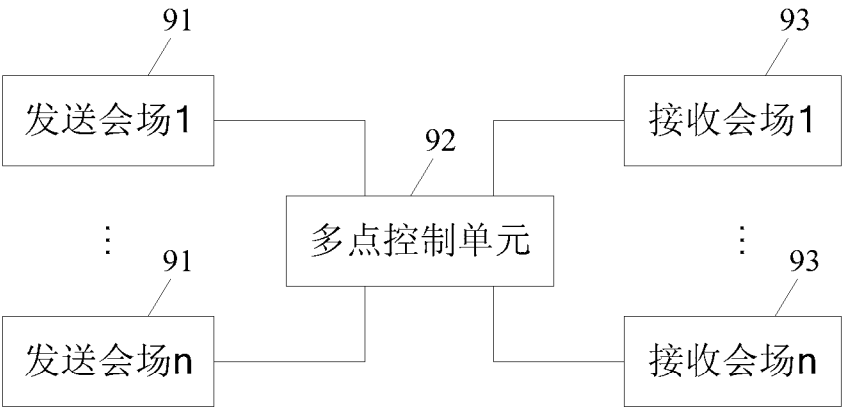


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2010/075891

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

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)

IPC: H04N7/-;H04M3/-;G10K15/-

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)
WPI, EPODOC, IEEE, CNKI, CPRS, SIPOABS, USTXT, CNTXT; mix, place, meeting, conference, channel, object, audio, video, type, energy, select, location, multi-channel, multipoint control, MCU

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN101179693A (SHENZHEN LISHI VIDEO TECHNOLOGY CO.,LTD. et al.) 14 May 2008 (14.05.2008) description, page 3, line 21 to page 5, line 9, claims 1,2, figures 1,2	1-16
A	CN1953537A (BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS) 25 Apr. 2007 (25.04.2007) the whole document	1-16
A	CN101466043A (SHENZHEN HUAWEI COMMUNICATION TECHNOLOGIES CO.,LTD.) 24 Jun.2009 (24.06.2009) the whole document	1-16
A	US2008/0005246A1 (VANBUSKIRK, Michael R. et al.) 03 Jan.2008 (03.01.2008) the whole document	1-16
A	WO2008/120933A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE et al.) 09 Oct.2008 (09.10.2008) the whole document	1-16
A	WO2006/103584A1 (KONINKLIJKE PHILIPS ELECTRONICS N.V. et al.) 05 Oct. 2006 (05.10.2006) the whole document	1-16

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&”document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 20 Oct.2010(20.10.2010)	Date of mailing of the international search report 18 Nov. 2010 (18.11.2010)
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer DONGGang Telephone No. (86-10)62413649

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2010/075891

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN101179693A	14.05.2008	NONE	
CN1953537A	25.04.2007	NONE	
CN101466043A	24.06.2009	WO2010/075790A1	08.07.2010
US2008/0005246A1	03.01.2008	US7257641B1	14.08.2007
WO2008/120933A1	09.10.2008	EP2143101A1	13.01.2010
		US2010/0121647A1	13.05.2010
		JP2010525378T	22.07.2010
		CN101689368A	31.03.2010
		KR20080089308A	06.10.2008
WO2006/103584A1	05.10.2006	EP1866912A1	19.12.2007
		US2010/0153097A1	17.06.2010
		JP2008535356T	28.08.2008
		AT473502T	15.07.2010
		KR20070118161A	13.12.2007
		MX2007011915A	22.11.2007
		RU2007139918A	10.05.2009
		RU2007139922A	10.05.2009
		WO2006/103586A1	05.10.2006
		KR20070116174A	06.12.2007
		MX2007011995A	07.12.2007
		EP1866913A1	19.12.2007
		CN101151660A	26.03.2008
		CN101151658A	26.03.2008
		JP2008535015T	28.08.2008
		AT406651T	15.09.2008
		ES2313646T	01.03.2009
		US2010/0153118A1	17.06.2010

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2010/075891

CLASSIFICATION OF SUBJECT MATTER:

H04N7/15(2006.01)i

H04M3/56(2006.01)i

G10K15/08(2006.01)i

国际检索报告

国际申请号
PCT/CN2010/075891

A. 主题的分类

见附加页

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04N7/-;H04M3/-;G10K15/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, EPODOC, IEEE, CNKI, CPRS, SIPOABS, USTXT, CNTXT; 混音, 会场, 会议, 声道, 对象, 音频, 视频, 类型, 能量, 选择, 多声道, 多点控制, mix, place, meeting, conference, channel, object, audio, video, type, energy, select, location, multi-channel, multipoint control, MCU

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN101179693A (深圳市丽视视讯科技有限公司 等) 14.5 月 2008 (14.05.2008) 说明书第 3 页第 21 行至第 5 页第 9 行, 权利要求 1,2, 图 1,2	1-16
A	CN1953537A (北京航空航天大学) 25.4 月 2007 (25.04.2007) 全文	1-16
A	CN101466043A (深圳华为通信技术有限公司) 24.6 月 2009 (24.06.2009) 全文	1-16
A	US2008/0005246A1 (VANBUSKIRK, Michael R. 等) 03.1 月 2008 (03.01.2008) 全文	1-16
A	WO2008/120933A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE 等) 09.10 月 2008 (09.10.2008) 全文	1-16
A	WO2006/103584A1 (KONINKLIJKE PHILIPS ELECTRONICS N.V. 等) 05.10 月 2006 (05.10.2006) 全文	1-16

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期
20.10 月 2010 (20.10.2010)

国际检索报告邮寄日期
18.11 月 2010 (18.11.2010)

ISA/CN 的名称和邮寄地址:
中华人民共和国国家知识产权局
中国北京市海淀区蓟门桥西土城路 6 号 100088
传真号: (86-10)62019451

受权官员

董刚
电话号码: (86-10) **62413649**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2010/075891

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN101179693A	14.05.2008	无	
CN1953537A	25.04.2007	无	
CN101466043A	24.06.2009	WO2010/075790A1	08.07.2010
US2008/0005246A1	03.01.2008	US7257641B1	14.08.2007
WO2008/120933A1	09.10.2008	EP2143101A1	13.01.2010
		US2010/0121647A1	13.05.2010
		JP2010525378T	22.07.2010
		CN101689368A	31.03.2010
		KR20080089308A	06.10.2008
WO2006/103584A1	05.10.2006	EP1866912A1	19.12.2007
		US2010/0153097A1	17.06.2010
		JP2008535356T	28.08.2008
		AT473502T	15.07.2010
		KR20070118161A	13.12.2007
		MX2007011915A	22.11.2007
		RU2007139918A	10.05.2009
		RU2007139922A	10.05.2009
		WO2006/103586A1	05.10.2006
		KR20070116174A	06.12.2007
		MX2007011995A	07.12.2007
		EP1866913A1	19.12.2007
		CN101151660A	26.03.2008
		CN101151658A	26.03.2008
		JP2008535015T	28.08.2008
		AT406651T	15.09.2008
		ES2313646T	01.03.2009
		US2010/0153118A1	17.06.2010

主题的分类:

H04N7/15(2006.01)i

H04M3/56(2006.01)i

G10K15/08(2006.01)i