[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/30 (2006.01)
G06F 3/16 (2006.01)
G10L 15/00 (2006.01)
G10L 17/00 (2006.01)

[12] 发明专利说明书

专利号 ZL 200510060100.4

[45] 授权公告日 2007 年 8 月 22 日

[11] 授权公告号 CN 1333363C

[22] 申请日 2005.1.14
[21] 申请号 200510060100.4
[30] 优先权
[32] 2004. 1.14 [33] JP [31] 006456/04
[73] 专利权人 索尼株式会社
地址 日本东京都
[72] 发明人 田中出 饭田健一 三原悟史
山田荣一
[56] 参考文献
JP2002 - 333897A 2002.11.22
CN1422494A 2003.6.4
JP2000 - 322077A 2000.11.24
JP2002 - 14691A 2002.1.18
审查员 吴 芸

[74] 专利代理机构 北京市柳沈律师事务所
代理人 黄小临 王志森

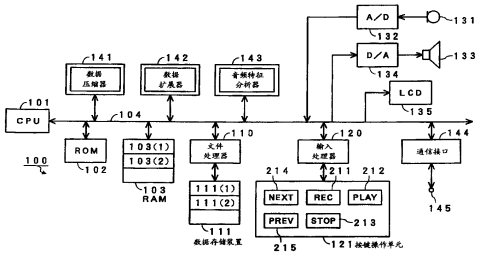
权利要求书 4 页 说明书 27 页 附图 18 页

[54] 发明名称

音频信号处理设备和音频信号处理方法

[57] 摘要

一种自动检测要处理的音频信号中的改变点的音频特征分析器。 中央处理单元(CPU)获得指示音频信号内的改变点位置的改变点信息，该改变点信息被记录在数据存储装置上。 CPU 根据用户通过按键操作单元输入的指令识别改变点信息，定位对应于该识别的改变点信息的音频数据，从而可以从那里开始诸如播放要处理的音频数据等的处理。



1、一种音频信号处理设备，包括:

第一检测装置，用于基于具有预定尺寸的各个处理单元，根据音频信号来检测在要处理的音频信号中的发言人改变;

获取装置，用于获取指示其中所述第一检测单元已检测到发言人改变的音频信号的位置的改变点信息；以及

保存装置，用于保存由获取装置获取的改变点信息。

2、根据权利要求 1 所述的音频信号处理设备，其中第一检测装置能基于各个处理单元来提取该音频信号的特征，并根据所提取的特征来检测从非话音段到话音段的改变点以及话音段内的发言人改变点。

3、根据权利要求 2 所述的音频信号处理设备，进一步包括:

存储装置，用于存储代表一个或多个发言人的语音特征的一个或多个特征信息，以及所述一个或多个发言人的一个或多个识别信息，所述多个特征信息和多个识别信息是分别相互关联的；以及

识别装置，用于通过比较第一检测装置提取的特征和存储在存储装置内的所述多个特征信息来识别发言人;

其中该保存装置保存改变点信息以及由识别装置识别的发言人的识别信息，所述改变点信息和所述识别信息相互联系。

4、根据权利要求 2 所述的音频信号处理设备，进一步包括: 第二检测装置，用于通过分析分别与多个麦克风相关联的多个音频信道的多个音频信号来检测发言人的位置，其中所述获取装置根据第二检测装置检测到的发言人位置中的改变来识别改变点，并获取对应于所识别的改变点的改变点信息。

5、根据权利要求 3 所述的音频信号处理设备，进一步包括:

发言人信息存储装置，用于存储根据分别与多个麦克风相关联的多个音频信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置的多个发言人的多个识别信息，所述多个发言人位置分别与所述多个识别信息相关联；以及

发言人信息获取装置，用于从发言人信息存储装置获取与发言人位置相关联的识别信息，该发言人位置是通过分析多个音频信道的多个音频信号来确定的;

其中该识别装置根据由发言人信息获取装置获取的识别信息来识别该发言人。

6、根据权利要求3所述的音频信号处理设备，进一步包括显示信息处理装置，其中存储装置存储分别与对应于各个识别信息的发言人相关联的多个信息，所述多个信息分别与各个识别信息相关联，以及该显示信息处理装置显示音频信号中的改变点的位置以及与由识别装置识别的发言人相关的信息。

7、根据权利要求1所述的音频信号处理设备，其中第一检测装置根据发言人位置来检测发言人改变，其中所述发言人位置是通过分析各个音频信道的多个音频信号而确定的，所述音频信号是由不同麦克风收集的。

8、根据权利要求7所述的音频信号处理设备，其中保存装置保存改变点信息以及指示由第一检测装置检测到的发言人位置的信息，该改变点信息与指示发言人位置的信息相互关联。

9、根据权利要求7所述的音频信号处理设备，进一步包括:

发言人信息存储装置，用于存储根据分别与多个麦克风相关联的多个信道的多个音频信号而确定的发言人位置，以及在各个发言人位置的发言人的多个识别信息，所述多个发言人位置分别与所述多个识别信息相互关联；以及

发言人信息获取装置，用于从发言人信息存储装置获取与发言人位置相关联的发言人的识别信息，该发言人位置是通过分析多个音频信道的多个音频信号来确定的;

其中保存装置保存改变点信息以及由发言人信息获取装置获取的识别信息，所述改变点信息与所述识别信息相互关联。

10、根据权利要求9所述的音频信号处理设备，进一步包括显示信息处理装置，其中发言人信息存储装置存储分别与对应于各个识别信息的发言人相关的多个信息，所述多个信息分别与各个识别信息相关联，以及该显示信息处理装置显示音频信号中的改变点的位置，以及与所确定的发言人位置相联系的、与发言人相关的信息。

11、一种音频信号处理方法，包括:

第一检测步骤，根据音频信号，基于具有预定尺寸的各个处理单元，检测要处理的音频信号中的发言人改变;

获取步骤，获取指示已在第一检测步骤中检测到发言人改变的音频信号的位置的改变点信息；以及

存储步骤，在记录介质上存储在获取步骤获取的改变点信息。

12、根据权利要求 11 所述的音频信号处理方法，其中在第一检测步骤中，基于各个处理单元提取该音频信号的特征，并根据提取的特征来检测从非话音段到话音段的改变点以及话音段内的发言人的改变点。

13、根据权利要求 12 所述的音频信号处理方法，进一步包括识别步骤，通过比较在第一检测步骤中提取的特征和表示一个或多个发言人的语音特征的一个或多个特征信息来识别发言人，所述多个特征信息和所述一个或多个发言人的一个或多个识别信息相互关联地存储在记录介质上，其中改变点信息和在识别步骤中识别的发言人的识别信息在存储步骤中被相互关联地存储在该记录介质上。

14、根据权利要求 12 所述的音频信号处理方法，进一步包括：第二检测步骤，通过分析分别与多个麦克风相关联的多个音频信道的多个音频信号来检测发言人位置，其中在获取步骤中，根据在第二检测步骤中检测到的发言人位置中的改变来识别改变点，并获取对应于所识别的改变点的改变点信息。

15、根据权利要求 13 所述的音频信号处理方法，进一步包括：

发言人信息存储步骤，在发言人信息存储装置上预先存储根据分别与多个麦克风相关联的多个音频信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置上的多个发言人的多个识别信息，所述多个发言人位置分别与所述多个识别信息相关联；以及

发言人信息获取步骤，从发言人信息存储装置获取与发言人位置相关联的发言人识别信息，该发言人位置是通过分析多个音频信道的多个音频信号来确定的；

其中在该识别步骤中，根据在发言人信息获取步骤获取的识别信息来识别该发言人。

16、根据权利要求 13 所述的音频信号处理方法，进一步包括显示信息处理步骤，其中在记录介质上分别与各个识别信息相关联地存储分别与对应于各个识别信息的多个发言人相关的多个信息，以及在显示信息处理步骤中显示音频信号中的改变点的位置和与在识别步骤中所识别的发言人相关的信息。

17、根据权利要求 11 所述的音频信号处理方法，其中在第一检测步骤中，根据通过分析各个音频信道的多个音频信号而确定的发言人位置来检测改变点，该音频信号是由不同麦克风收集的。

18、根据权利要求 17 所述的音频信号处理方法，其中在存储步骤中相互关联地存储改变点信息和在第一检测步骤中检测到的指示发言人位置的信息。

19、根据权利要求 17 所述的音频信号处理方法，进一步包括:

发言人信息存储步骤，在发言人信息存储装置上预先存储根据分别与多个麦克风相关联的多个信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置的多个发言人的多个识别信息，所述多个发言人位置分别与多个识别信息相关联；以及

发言人信息获取步骤，从发言人信息存储装置获取与发言人位置相关联的发言人的识别信息，该发言人位置是通过分析多个音频信道的多个音频信号来确定的;

其中在存储步骤中相互关联地存储改变点信息和在发言人信息获取步骤中获取的识别信息。

20、根据权利要求 19 所述的音频信号处理方法，进一步包括显示信息处理步骤，其中存储装置存储分别与对应于各个识别信息的多个发言人相关的多个信息，所述多个信息分别与各个识别信息相关联，以及在显示信息处理步骤中显示音频信号中的改变点的位置，以及与所确定的发言人位置相关联的发言人有关的信息。

音频信号处理设备和音频信号处理方法

技术领域

本发明涉及处理音频信号的多种设备，例如IC(集成电路)录音机、MD(迷你盘)录音机或个人计算机，以及涉及在这些设备中使用的方法。

背景技术

如在日本未审查的专利申请第2-206825号中所公开的，已经提出用于对记录的音频数据执行语音识别以将音频数据转换为文本数据的，从而自动创建记录的记录准备设备。这种技术允许快速地自动准备会议记录。然而，在某些情况中，期望准备仅仅重要部分的记录而非基于所有记录的音频数据来准备记录。在此种情况中，需要从记录的音频数据中寻找到感兴趣的部分。

例如，当使用IC录音机、MD录音机等记录下较长会议的进行时，为了从记录的音频数据中找到感兴趣的部分，需要播放音频数据并听取播放的声音。虽然能使用快进或快退找到感兴趣的部分，但这通常需要付出劳动和时间。因此，已提出能嵌入(分配)促进在记录的数据中搜索的标记的记录设备。例如，在MD录音机中，此种功能被实现为附加磁道标记功能。

然而，附加促进对音频数据的搜索的标记的功能如上所述是由用户手动操作的，因此没有用户操作就不能分配标记。因此，甚至记录期间当用户试图执行附加标记到用户认为是重要的部分中的操作时，如当用户专注于会议进行时也会忘记执行附加标记的操作。

此外，甚至当用户对感兴趣的讲话分配了标记，由于是在听取感兴趣的讲话期间执行嵌入标记的操作，也可能在感兴趣的讲话之后才记录该标记。因此，为了听取感兴趣的讲话，用户必须执行将播放位置移到该标记并稍后再向后移动的操作。如果用户操作到感兴趣的部分之前或之后并不得不重复这一操作，这对用户来说是麻烦和紧迫的。

此外，具有标记的这部分的内容是未知的，直至听取之后。如果听取后发现这部分不是感兴趣的那一部分，必须重复移到下一标记的操作，直至发现感兴趣的部分，而这样是很费力的。如上所述，虽然分配帮助搜索音频数

据的标记的功能十分便利，但是当例如用户不习惯该操作时，就不能很好的利用分配标记给感兴趣的音频数据这一功能。

发明内容

因此，本发明的一个目的是提供一种设备和方法，易于允许用户快速地在要处理的音频信号中找到和使用感兴趣的部分。

为实现这一目的，根据本发明的一个方面，提供一种音频信号处理设备。该音频信号处理设备包括：第一检测单元，用于在要处理的音频信号中根据该音频信号基于具有预定尺寸的各个处理单元来检测发言人的改变；获取单元，用于获得其中第一检测单元已检测到发言人改变的音频信号的位置的改变点信息；以及保存单元，用于保存由获取单元获取的改变点信息。

在该音频信号处理设备中，检测单元自动检测在要被处理的音频信号中的改变点，获取单元获取指示在音频信号中的改变点的位置的改变点信息，保存单元保存该改变点信息。保存指示改变点的位置的改变点信息等同于在要被处理的音频信号中给改变点分配标记。

如上所述检测和保存的改变点信息使得能够定位对应于该改变点信息的音频信号，从而可以从该位置开始诸如播放要被处理的音频信号的处理。因此，允许用户参考自动分配给该音频信号中的改变点的标记，从音频信号中快速地找出感兴趣的部分，而无需执行麻烦的操作。

优选地，第一检测单元能基于各个处理单元提取音频信号的特征，并根据提取的特征检测从非话音段到话音段的改变点以及话音段中的发言人改变的点。

因此，检测单元基于具有预定尺寸的各个处理单元检测要被处理的音频信号的特征，并执行诸如将特征和先前所检测到的特征进行比较的处理。因此，检测单元能检测从无声段或噪声段到话音段的改变点以及话音段内的发言人改变点。

因此，标记被至少分配给发言人改变的点，以便能参考发言人改变点快速地从音频数据中找到感兴趣的部分。

该音频信号处理设备可以进一步包括：存储单元，用于存储表示一个或多个发言人的语音特征的一个或多个特征信息，以及所述一个或多个发言人的一个或多个识别信息，特征信息以及识别信息相互关联；还包括识别单元，

用于通过比较由第一检测单元提取的特征和存储在存储单元内的特征信息来识别发言人。在此情况中。保存单元保存改变点信息以及由识别单元识别的发言人的识别信息，该改变点信息和识别信息相互关联。

在该音频信号处理设备中，代表发言人的语音特征的特征信息以及发言人的识别信息被相互关联地存储在存储单元内。识别单元通过比较由第一检测单元提取的特征和存储在存储单元内的特征信息在改变点识别发言人。保存单元保存改变点信息以及识别的发言人的识别信息。

因此，能播放或提取对应于特定发言人的语音的部分，以及根据在各个改变点的发言人的识别从音频数据中快速地找到感兴趣的部分。

该音频信号处理设备可以进一步包括：第二检测单元，用于通过分析分别与多个麦克风相关联的多个音频信道的多个音频信号来检测发言人位置。在此情况中，获取单元根据由第二检测单元检测到的发言人位置中的改变来识别改变点，并获取对应于所识别的改变点的改变点信息。

在该音频信号处理设备中，第二检测单元通过分析各个音频信道的多个音频信号、检测要被处理的音频信号中的改变点来检测发言人的位置。获取单元根据由第一检测单元检测到的改变点和由第二检测单元检测到的改变点来识别实际使用的改变点，并获取指示所识别的改变点的位置的改变点信息。

因此，根据由第二检测单元检测到的改变点，能更精确和更可靠地检测出音频信号中的改变点，允许从音频数据中搜索出感兴趣的部分。

该音频信号处理设备可进一步包括：发言人信息存储单元，用于存储基于分别与多个麦克风相关联的多个音频信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置的发言人的识别信息，该发言人位置和识别信息相互关联；以及发言人信息获取单元，用于从发言人信息存储单元获取与通过分析多个音频信道的多个音频信号而确定的发言人位置相关联的发言人的识别信息。在此情况中，识别单元根据由发言人信息获取单元所获取的识别信息来识别发言人。

在该音频信号处理设备中，发言人信息存储单元存储基于分别与多个麦克风相关联的多个音频信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置的发言人的识别信息。即，基于提供各个麦克风的位置来确定发言人的位置。例如，距离第一麦克风位置最近的发言人是A，距离第二麦克风位置最近的发言人是B。因此，例如基于哪个麦克风与具有最高电

平的音频数据的音频信道相关联，能够确定当前发言人与哪个麦克风相关联。

发言人信息获取单元分析各个音频信道的音频数据，根据与主要用于收集声音的麦克风相关联的那个音频信道来识别发言人的位置。识别单元根据以上述方式获得的标识来识别在改变点的发言人。因此，可以使用准确的信息从要被处理的音频数据中搜索感兴趣的部分，从而提高发言人识别的精确性。

音频信号处理设备可进一步包括：显示信息处理单元。在该种情况中，存储单元存储分别与对应各个识别信息的发言人相关的多个信息，该信息分别和各个识别信息相关，显示信息处理单元显示音频信号中的改变点的位置以及与由识别单元识别的发言人相关的信息。

在音频信号处理设备中，存储单元存储分别与对应于各个识别信息的发言人相关的信息，例如，不同的图像数据或图形数据，如与各个识别信息相关的面部图片数据、图标数据、标记图像数据或动画图像数据。显示信息处理单元显示改变点的位置以及与由识别单元所识别的发言人相关的信息。

因此，用户可以直观地在要被处理的音频数据中找到对应于各个发言人的讲话的部分。因此，用户可以快速地在要被处理的音频数据中找到感兴趣的部分。

在音频信号处理设备中，第一检测单元可以根据通过分析各个音频信道的多个音频信号而确定的发言人位置来检测发言人的改变，其中该音频信号由不同麦克风收集。

在音频信号处理设备中，通过分析各个音频信道的多个音频信号来识别发言人位置，发言人位置中的改变点被检测为改变点。

因此，通过分析各个音频信道的音频信号，可以容易并准确地检测出要被处理的音频信号中的改变点，以及可将标记分配给发言人改变的点。此外，还能参考发言人改变的点来快速地从音频数据中找到感兴趣的部分。

优选地，在音频信号处理设备中，保存单元保存改变点信息以及指示由第一检测单元所检测到的发言人位置的信息，改变点信息和指示发言人位置的信息相互关联。

在音频信号处理设备中，可将保存在保存单元内的信息提供给用户。因此，允许用户寻找在每个改变点发言的发言人的发言人位置，并从要被处理的音频数据中寻找感兴趣的部分。

音频信号处理设备可进一步包括：发言人信息存储单元，用于存储根据分别与多个麦克风相关的多个音频信道的多个音频信号而确定的多个发言人位置，以及在各个发言人位置的发言人的识别信息，多个发言人位置分别与多个识别信息分别相关联；以及发言人信息获取单元，从发言人信息存储单元获得与发言人位置相关联的发言人的识别信息，所述发言人位置是通过分析多个音频信道的多个音频信号来确定的。在此种情况中，保存单元保存改变点信息以及由发言人信息获取单元所获取的识别信息，所述改变点信息和所述识别信息相互关联。

在音频信号处理设备中，发言人信息存储单元存储根据麦克风位置确定的发言人位置以及在各个发言人位置的发言人的识别信息，所述发言人位置与所述识别信息相互关联。发言人信息获取装置通过分析各个音频信道的多个音频信号来识别发言人位置。保存单元保存改变点信息以及发言人信息获取单元获取的识别信息，所述改变点信息和所述识别信息相互关联。

因此，能够在各个改变点识别发言人，以及将该信息提供给用户。因此，能容易并准确地从要被处理的音频数据中找到感兴趣的部分。

音频信号处理设备还可以包括显示信息处理单元。在此种情况中，发言人信息存储单元存储分别与对应于各个识别信息的发言人相关的信息，所述信息分别与各个识别信息相关联，显示信息处理单元显示音频信号内的改变点的位置以及与关联于所确定的发言人位置的发言人相关的信息。

在音频信号处理设备中，发言人信息存储装置存储分别与对应于各个识别信息的发言人相关的多个信息，例如，与各个识别信息相关的不同的图像数据或图形数据，如面部图片数据、图标数据、标记图像数据、或动画图像数据。显示信息处理单元显示改变点的位置以及关于由识别单元识别的发言人的信息。

因此，用户可以直观地在要被处理的音频数据中找到对应于各个发言人的讲话的部分。因此，用户能够快速地在要被处理的音频数据中找到感兴趣的部分。

根据本发明的另一方面，提供一种音频信号处理方法。该音频信号处理方法包括：第一检测步骤，基于具有不同尺寸的各个处理单元，根据音频信号来检测要被处理的音频信号中的发言人改变；获取步骤，获取指示在第一检测步骤中检测到发言人改变的音频信号的位置的改变点信息；以及存储步

骤，存储在获取步骤获取的改变点信息到记录介质上。

根据本发明，甚至当记录较长会议时，每当发言人改变发生时就自动分配发言人改变标记。这样更易于在准备会议记录过程中搜索讲话，允许容易和快速地重复播放对应于感兴趣的发言人的讲话的部分。

此外，能在音频数据的改变点识别发言人以及管理指示与改变点相关联的发言人的信息。因此，能够容易和快速地找到对应于特定发言人的讲话的部分而无需播放该音频数据。

此外，对创建会议记录的人的记忆力的依赖降低了。这有助于提高准备会议记录的工作的效率，其中准备会议记录曾经是费力的和耗时的。此外，还能将记录的数据用作音频数据形式的会议记录，而无需创建会议记录。这使搜索更容易。

附图说明

图 1 是根据本发明实施例的记录/播放设备的框图;

图 2 是用于解释给由记录/播放设备记录的所收集的音频信号中的改变点分配标记的处理方案的图;

图 3 是示出在播放记录的音频信号期间将播放位置设定到标记时，LCD 上所显示的信息是如何根据这些操作改变的图;

图 4 是图 1 所示的记录/播放设备所执行的记录处理的流程图;

图 5 是图 1 所示的记录/播放设备所执行的播放处理的流程图;

图 6 是示出在图 1 所示的记录/播放设备的外部存储装置的存储区域内创建的音频特征数据库的例子的图;

图 7 是用于解释在图 1 所示达到记录/播放设备内的用于分配标记给收集到的音频信号的处理方案的图;

图 8 是示出在播放记录的音频信号期间将播放位置设定到标记时，LCD 上所显示的信息是如何根据这些操作改变的图;

图 9 是在记录处理之后给记录的音频信号中的改变点分配标记的处理流程图;

图 10 是示出根据从图 1 所示的记录/播放设备传输到个人计算机的数据，在显示器的屏幕上显示的改变点信息的例子的图;

图 11 是示出根据从图 1 所示的记录/播放设备传输到个人计算机的数据，

在显示器的屏幕上显示的改变点信息的例子的图；

图 12 是根据本发明的另一实施例的记录/播放设备的框图；

图 13 是示出麦克风和音频信号处理器的例子的图；

图 14 是示出麦克风和音频信号处理器的另一例子的图；

图 15A 和 15B 是用于解释在记录处理之后给记录的音频信号内的改变点分配标记的处理的图；

图 16 是示出发言人位置数据库的例子的图；

图 17A 和 17B 是用于解释根据从麦克风输出的信号通过识别发言人位置来识别发言人的其他示例方案的图；和

图 18 是根据本发明的另一实施例的记录/播放设备的框图。

具体实施方式

现在，将参考附图说明根据本发明实施例的设备、方法以及程序。将以实例来说明实施例，在实例中，将本发明应用到 IC 录音机中，所述 IC 录音机是用于记录并播放音频信号的设备。

第一实施例

IC 录音机结构和操作概览

图 1 是 IC 录音机的框图，该录音机是根据本发明第一实施例的记录/播放设备。参考图 1，根据第一实施例的 IC 录音机包括用微计算机实现的控制器 100。控制器 100 包括：中央处理单元(CPU)101，存储程序和各种数据的只读存储器(ROM)102，主要用作工作区域的随机访问存储器(RAM)103，这些部件经 CPU 总线 104 相互连接。如下面将要说明的，RAM 103 包括压缩数据区域 103(1)和 PCM(脉码调制)数据区域 103(2)。

控制器 100 经文件处理器 110 被连接到数据存储装置 111，并经输入处理器 120 被连接到按键操作单元 121。此外，控制器 100 经模拟/数字转换器(下文中简写为 A/D 转换器)132 被连接到麦克风 131，并经数字/模拟转换器(下文中简写为 D/A 转换器)134 连接到扬声器 133。此外，控制器 100 还连接到液晶显示器(LCD)135。在此实施例中，LCD 135 包括 LCD 控制器的功能。

此外，控制器 100 连接到数据压缩器 141、数据扩展器 142、音频特征分析器 143 以及通信接口(下文中简写为通信 I/F)144。数据压缩器 141、数据扩

展器 142 以及音频特征分析器 143 的功能由图 1 中的双线指示，也可在控制器 100 的 CPU 101 执行的软件(例如程序)内实现。

在第一实施例中，通信 I/F144 是数字接口，诸如 USB(通用串行总线)或 IEEE(电子和电气工程师协会)-1394 接口。通信 I/F144 允许与连接到连接终端 145，如个人计算机或数字照相机的各种电子装置交换数据。

在根据第一实施例的 IC 录音机中，当按下按键操作单元 121 的 REC 按键(录音按键)211 时，CPU 101 控制相关部件执行记录处理。在记录处理中，麦克风 131 收集声音，收集的声音由 A/D 转换器 132 进行 A/D 转换，所得到的数字数据由数据压缩器 141 压缩，经文件处理器 110 将所获得的音频信号记录在数据存储装置 111 的预定存储区域。

第一实施例中的数据存储装置 111 是闪存或包括闪存的存储器卡。如下面要说明的，数据存储装置 111 包括数据库区域 111(1)和音频文件 111(2)。

在记录处理中，根据第一实施例的 IC 录音机，由音频特征分析器 143 的功能分别分析具有预定尺寸的每一处理单元的所记录的收集到的音频信号的特征。当检测到特征改变时，IC 录音机分配标记到该改变点。这些标记允许从记录的音频信号中快速搜索到期望的音频信号段。

图 2 是用于解释在记录的收集音频信号内的改变点处分配标记的处理方案的图。如上所述，在根据第一实施例的 IC 录音机中，分别分析由麦克风 131 收集到的具有预定尺寸的每一处理单元的音频信号的特征。

通过比较当前处理单元的特征分析结果和前一处理单元的特征分析结果，检测到从静默段或噪声段到话音段的改变点或在某话音段内发言人改变的点，标识音频信号内改变的时间位置。然后，在数据存储装置 111 内将识别的位置存储为改变点信息(标记信息)。这样，通过存储指示音频信号内的改变点位置的改变点信息来实现标记所记录的收集到的音频信号。

作为例子，可以考虑记录会议进行的例子。假设在记录开始之后 10 秒钟 A 开始讲话，如图 2 所示。在此情况中，在 A 开始讲话之前，收集的是静默，或纯语音之外的无意义的声音，即诸如干扰的噪声、拉椅子的声音、或某个物品撞击桌面的声音。当 A 开始讲话并收集 A 的讲话时，收集到的音频信号特征分析结果与 A 开始讲话之前的结果明显不相同。

通过音频特征分析器 143 检测到记录的收集音频信号中的改变点，识别(获得)音频信号中改变点的位置，在数据存储装置 111 内将指示音频信号内的

识别位置的改变点信息存储为图 2 中的标记 MK1。图 2 显示将记录开始之后过去的时间存储为改变点信息的例子。

进一步假定 B 在 A 结束讲话之后不久开始讲话。在 B 开始讲话之前的时间段是静默段或噪声段。在此例子中一样，当 B 开始讲话并且收集 B 的讲话时，对收集的音频信号的特征分析结果明显和 B 开始讲话之前的不同。因此，如图 2 标记 MK2 所示，在数据存储装置 111 内存储改变点信息(标记 MK2)，从而将标记分配给 B 讲话的开始点。

此外，当 B 讲话时可能发生 C 插话的情况。在此情况中，由于 B 的声音和 C 的声音不同，对收集的音频信号的分析结果在 B、C 之间不同。因此，如图 2 标记 MK3 所示，在数据存储装置 111 内存储改变点信息(标记 MK3)，从而将标记分配给 C 讲话的起始点。

如上所述，在根据第一实施例的 IC 录音机的记录处理中，分析收集的音频信号的特征，并存储音频信号特征中的改变点。因此，可将标记分配给音频信号特征内的改变点。

参考图 2，标记 MK1、MK2、MK3 的“其他”部分允许与该标记相联系地存储相关信息。例如，如果通过语音识别将语音转换为文本数据，则该文本数据和相关标记一起存储。

在根据第一实施例的 IC 录音机中，当按下按键操作单元 121 的 PLAY 按键(播放按键)212 时，CPU 101 控制相关部件执行播放处理。更具体地，经文件处理器 110 读取记录在数据存储装置 111 预定存储区域的压缩数字音频信号，由数据扩展器 142 扩展该数字音频信号，借此恢复压缩之前的原始数字音频信号。恢复的数字音频信号被 D/A 转换器 134 转换为模拟音频信号，将该模拟信号提供到扬声器 133。因此，产生要播放的对应于记录的音频信号的声音。

在根据第一实施例的 IC 录音机的播放处理中，当操作按键操作单元 121 的 NEXT 按键(用于定位下一标记的按键)214 或 PREV 按键(用于定位前一标记的按键)215 时，播放位置被快速地设定到相关标记的位置，从而播放从那里开始。

图 3 是示出根据操作在 LCD 135 上所显示的信息的改变的图，该图用于解释当播放记录的音频信号时，用于定位记录的音频信号上的由标记所指示的位置的操作。参考图 3，当按下 PLAY 按键 211 时，如前所述，CPU 101

控制相关部件从指定的记录音频信号的开头开始播放。

在对应于 A 的讲话的部分，根据参考图 2 所述在记录处理中分配的标记 MK1，显示 A 讲话开始的时刻，以及指示该标记是在开始记录之后所分配的第一个标记的“序列号 1”，如图 3 A 部分所示。

当播放继续并且对应 B 讲话的部分开始播放时，显示 B 讲话的开始时刻，以及指示该标记是在开始记录之后所分配的第二个标记的“序列号 2”，如图 3 B 部分所示。然后，当按下 PREV 按键 215 时，CPU 101 将播放位置设定到 A 讲话的起始点，即标记 MK1 指示的开始点之后的 10 秒(0 分钟，10 秒)，从而播放从那里重新开始，如图 3 C 部分所示。

然后，当按下 NEXT 按键 214 时，CPU 101 将播放位置设定到 B 的讲话的开始点，即在由标记 MK2 指示的开始点之后的 1 分钟 25 秒，从而播放从那里恢复，如图 3 D 部分所示。当再次按下 NEXT 按键 214 时，CPU 101 将播放位置设定到 C 的讲话的起始点，即由标记 MK3 指示的开始点后的 2 分钟 30 秒，播放从那里恢复，如图 3 E 部分所示。

如上所述，在根据第一实施例的 IC 录音机中，在记录处理中，自动分析收集的音频信号的特征，并分配标记到特征内的改变点。此外，在播放处理中，通过操作 NEXT 按键 214 或 PREV 按键 215，可将播放位置快速地设定到由分配的标记指示的记录音频信号的点， 从而播放可以从那里开始。

这允许用户将播放位置快速地设定到感兴趣的发言人的讲话，并播放和收听记录的音频信号的这一部分。因此，用户可以快速地准备关于感兴趣的讲话的会议记录。

虽然为简化起见，在第一实施例中将指示从记录开始过去的时间的信息用作改变点信息，但不限于此，例如，记录在数据存储装置 111 的记录介质上的音频信号的地址也可用作改变点信息。

IC 录音机操作详细说明

下面，参考图 4 和 5 所示的流程图详细说明根据第一实施例的 IC 录音机所执行的记录处理和播放处理。

记录处理

首先，将说明记录处理。图 4 是显示由根据第一实施例的 IC 录音机所执

行的记录处理的流程图。图 4 所示处理由控制相关部件的 CPU 101 执行。

当被供电但未操作时，根据第一实施例的 IC 录音机等待用户的输入操作(步骤 S101)。当用户按下操作单元 121 的操作按键时，输入处理器 120 检测到该操作，并通知 CPU 101 该操作。CPU 101 确定所接受的操作是否是按下 REC 按键 211(步骤 S102)。

如果在步骤 S102 中确定接受的操作不是 REC 按键 211 的按下，CPU 101 执行对应于用户操作的按键的处理，例如，对应 PLAY 按键 212 的播放处理，对应 NEXT 按键 124 的定位下一标记的处理，或对应 PREV 按键 215 的定位前一标记的处理(步骤 S103)。显然，也允许快速前进或快速后退。

如果在步骤 S102 确定按下 REC 按键，CPU 101 命令文件处理器 110 执行文件记录处理。响应于该命令，文件处理器 110 在数据存储装置 111 内创建音频文件 111(2)(步骤 S104)。

然后，CPU 101 确定按键操作单元 121 的 STOP 按键 213 是否被按下(步骤 S105)。如果在步骤 S105 确定按下了 STOP 按键 213，则执行如后面所述的预定终止处理(步骤 S114)，并退出图 4 所示的处理。

如果在步骤 S105 确定没有按下 STOP 按键 213，CPU 101 命令 A/D 转换器 132 将经麦克风 131 输入的模拟音频信号转换为数字音频信号，从而收集的声音是数字化的(步骤 S106)。

响应于该命令，A/D 转换器 132 将经麦克风 131 输入的模拟音频信号转换为规则周期(即对于具有预定尺寸的每一处理单元)的数字音频信号，将数字音频信号写入到 RAM 103 的 PCM-数据区域 103(2)，并通知 CPU 101 该写入(步骤 S107)。

响应于该通知，CPU 101 命令数据压缩器 141 压缩存储在 RAM 103 的 PCM 数据区域 103(2)内的数字音频信号(PCM 数据)(步骤 S108)。响应于该命令，数据压缩器 141 压缩 RAM 103 的 PCM 数据区域 103(2)内的数字音频信号，并将压缩的数字音频信号写入到 RAM 103 的压缩数据区域 103(1)(步骤 S109)。

然后，CPU 101 命令文件处理器 110 将 RAM 103 的压缩数据区域 103(1)内的压缩数字音频信号写入到创建在数据存储装置 111 中的音频文件 111(2)内。因此，文件处理器 110 将 RAM 103 的压缩数据区域 103(1)内的压缩数字音频信号写入到数据存储装置 111 中的音频文件 111(2)内(步骤 S110)。

当完成将压缩数字音频信号写入到音频文件 111(2)后，文件处理器 110 通知 CPU 101 该完成。然后，CPU 101 命令音频特征分析器 143 分析先前记录在 RAM 103 的 PCM 数据区域 103(2)的数字音频信号的特征，从而音频特征分析器 143 提取 RAM 103 的 PCM 数据区域 103(2)内的数字音频信号的特征(步骤 S111)。

音频特征分析器 143 的数字音频信号的特征分析(特征提取)可以基于不同方法，例如，声波纹分析、语音速率分析、停顿分析、或重音分析。为简化说明起见，在此假定根据第一实施例的 IC 录音机的音频特征分析器 143 使用声波纹分析来提取要分析的数字音频信号的特征。

音频特征分析器 143 比较当前提取的音频特征(声波纹数据)和先前提取的声波纹数据，以确定从输入音频信号提取的特征是否已从先前的特征发生改变，并通知 CPU 101 该结果。基于该结果，CPU 101 确定收集的声音的特征是否改变(步骤 S112)。

如果在步骤 S112 确定特征未改变，则 CPU 101 在下一时段(下一处理单元)重复从步骤 S105 到步骤 S112 对音频信号的处理。

如果在步骤 S112 确定特征已经改变，则 CPU 101 确定发言人已改变，并命令文件处理器 110 分配标记到要处理的音频信号的特征中的改变点(步骤 S113)。响应于该命令，文件处理器 110 将指示关于音频文件 111(2)的音频特征中的改变点的信息，如指示从音频文件 111(2)开始的时间的信息或指示记录地址的信息写入到数据存储装置 111 的数据库区域 111(1)。此时，音频文件 111(2)和指示音频特征改变点的信息相互关联的存储。

在步骤 S113 之后，CPU 101 在下一时段(下一处理单元)重复从步骤 S105 到步骤 S112 的对音频信号的处理。

如果在步骤 S105 确定用户按下了 STOP 按键 213，则 CPU 101 执行预定的终止处理，所述处理包括：命令文件处理器 110 停止将数据写入到数据存储装置 111 的音频文件 111(2)内；命令数据压缩器 141 停止压缩以及命令 A/D 转换器 132 停止到数字信号的转换(步骤 S114)。然后退出图 4 所示处理。

音频特征分析器 143 通过保持先前提取的音频特征数据(声波纹数据)，并将先前的音频特征数据和最近提取的音频特征数据(声波纹数据)比较，来确定音频特征是否改变。如果它满足将最近提取的特征数据仅和前一组特征数据比较，则音频特征分析器就能恒定地仅保持前一组特征数据。如果要将最

近提取的特征数据和两组或三组先前的特征数据比较以提高精确性，即当观察到来自先前两组或三组特征数据之一不同时确定特征已改变，则必须保持两组或三组先前的特征数据。

如上所述，在根据第一实施例的 IC 录音机中，能分析记录的收集音频信号的特征，检测收集的音频信号的特征内的改变点，并分配标记到收集的音频信号中的改变点的位置。

播放处理

下面，将说明播放处理。图 5 是显示由根据第一实施例的 IC 录音机执行的播放处理的流程图。图 5 所示处理由控制相关部件的 CPU 101 执行。

在根据第一实施例的 IC 录音机的播放处理中，能够使用如参考图 4 所述的在记录处理中分配到所收集和记录的音频信号的特征中的改变点的标记，从记录的音频信号中快速地找到期望的音频信号段。

当被供电但不操作时，根据第一实施例的 IC 录音机等待用户的输入操作(步骤 S201)。当用户按下按键操作单元 121 的操作按键时，输入处理器 120 检测到该操作，并通知 CPU 101 该操作。CPU 101 确定接受的操作是否是按下 PLAY 按键 212(步骤 S202)。

如果在步骤 S202 确定接受的操作不是 PLAY 按键 212 的按下，CPU 101 执行对应于用户操作的按键的处理，例如，对应 REC 按键 212 的记录处理、对应 NEXT 按键 124 的定位下一标记的处理或对应 PREV 按键 215 的定位前一标记的处理(步骤 S203)。显然，也允许快速前进或快速向后。

如果在步骤 S202 确定接受的操作是播放按键 212 的按下，CPU 101 命令文件处理器 110 读取数据存储装置 111 上的音频文件 111(2)(步骤 S204)。然后，CPU 101 确定是否按下了按键操作单元 121 的 STOP 按键 213(步骤 S205)。

如果在步骤 S205 确定操作了 STOP 按键 213，则将执行如后所述的终止处理(步骤 S219)。然后退出图 5 所示处理。

如果在步骤 S205 确定没有操作 STOP 按键 213，则 CPU 101 命令文件处理器 110 读取存储在数据存储装置 111 的音频文件 111(2)内的压缩数字音频信号的数量，所述数量对应于系统所规定的处理单元的尺寸，并将该数字音频信号写入到 RAM 103 的压缩数据区域 103(1)(步骤 S206)。

当完成写入时，通知 CPU 101 该完成。然后，CPU 101 命令数据扩展器

142 扩展 RAM 103 的压缩数据区域 103(1)内的压缩数字音频信号。然后，数据扩展器 142 扩展压缩数字音频信号，并将扩展后的数字音频信号写入到 RAM 103 的 PCM 数据区域 103(2)(步骤 S207)。

当完成写入时，通知 CPU 101 该完成。然后，CPU 101 命令 D/A 转换器 134 将 RAM 103 的 PCM 数据区域 103(2)内的扩展数字音频信号转换为模拟信号，并提供该模拟信号到扬声器 133。

因此，对应于存储在数据存储装置 111 的音频文件 111(2)内的数字音频信号的声音从扬声器 133 输出。然后，D/A 转换器 134 通知 CPU 101 已经输出通过 D/A 转换获得的模拟音频信号。然后，CPU 101 确定是否操作了按键操作单元 121 的操作按键(步骤 S209)。

如果在步骤 S209 确定没有操作任何操作按键，则重复从步骤 S205 的处理从而继续播放数据存储装置 111 的音频文件 111(2)内的数字音频信号。

如果在步骤 S209 确定操作了操作按键，CPU 101 确定被操作的按键是否是 PREV 按键 215(步骤 S210)。如果在步骤 S210 确定操作了 PREV 按键 215，则 CPU 101 命令文件处理器 110 停止从音频文件 111(2)内读取数字音频信号，命令数据扩展器 142 停止扩展，并命令 D/A 转换器 134 停止到模拟信号的转换(步骤 S211)。

然后，CPU 101 命令文件处理器 110 从数据存储装置 111 的数据库区域 111(1)读取当前播放位置之前一个的标记信息(改变点信息)，从而将播放位置设定到由该标记信息指示的音频信号位置，并从这里开始播放(步骤 S212)。此时，如参考图 3 所述的，显示对应于该用于设定播放位置的标记信息的播放位置信息(步骤 S213)。然后，处理从步骤 S205 重复。

如果在步骤 S210 确定被操作的按键不是 PREV 按键 215，则 CPU 101 确定操作的按键是否是 NEXT 按键 214(步骤 S214)。如果在步骤 S214 确定操作了 NEXT 按键 214，则 CPU 101 命令文件处理器 110 停止从音频文件 111(2)读取数字音频信号，命令数据扩展器 142 停止扩展，并命令 D/A 转换器 134 停止到模拟信号的转换(步骤 S215)。

然后，CPU 101 命令文件处理器 110 从数据存储装置 111 的数据库区域 111(1)读取当前播放位置之后一个的标记信息(改变点信息)，从而将播放位置设定到由该标记信息指示的音频信号位置，并从这里开始播放(步骤 S216)。此时，如参考图 3 所述的，显示对应于该用于设定播放位置的标记信息的播

放位置信息(步骤 S217)。然后，处理从步骤 S205 重复。

如果在步骤 S214 确定操作的按键不是 NEXT 按键 214，则 CPU 101 执行对应于操作的按键的处理，例如快进或快退。然后，处理从步骤 S205 重复。

如上所述，在记录处理中，当检测到音频特征改变时，IC 录音机假定发言人改变，并自动分配标记到该改变点。因此，在播放处理中，允许用户仅仅通过按下 PREV 按键 215 或 NEXT 按键 214 到达各个讲话的开始。这极大的促进了会议记录的准备，例如，当重复播放特定讲话或当搜索重要讲话时。也就是说，能够快速地从记录的音频信号中找到期望的段。

此外，自动检测收集的音频信号的特征中的改变点，并自动分配标记到改变点。因此，无需用户的任何操作就将标记分配给改变点。

第一实施例的修改

当记录会议进行并根据该记录准备会议记录时，如果能够不用播放记录的声音就找到谁在讲话会更加方便。因此，在根据第一实施例的修改的 IC 录音机中，通过分析会议参加者的声音特征而获得的声波纹数据与符号相联系地存储，其中所述符号用于识别各个参加者，由此分配允许识别发言人的标记。

根据该修改的 IC 录音机和图 1 所示根据第一实施例的 IC 录音机结构相类似。然而，在根据修改的 IC 录音机中，例如在数据存储装置 111 或 RAM103 的存储区域中创建关于会议参加者的音频特征数据库。在下面的说明中，假定在数据存储装置 111 的存储区域内创建音频特征数据库。

图 6 是显示在根据修改的 IC 录音机的数据存储装置 111 的存储区域内创建的音频特征数据库的例子的图。如图 6 所示，该例中的音频特征数据库包括: 用于识别会议参加者的识别符(例如根据注册顺序的序列号)、会议参加者的姓名、通过分析会议参加者的声音特征获取的声波纹数据、诸如会议参加者面部图像的图像数据、分配给各个会议参加者的图标数据、以及诸如文本数据的其他数据。

声波纹数据、图像数据、图标数据以及其他数据的每一个都以文件形式存储在数据存储装置 111 内，将各个会议参加者的识别符作为关键信息(联合信息)。通过特征分析获得的声波纹数据是在会议之前通过收集会议参加者的声音并分析这些声音的特征预先获得的。

即，根据修改的IC录音机具有音频特征数据库创建模式。当选择音频特征数据库创建模式时，收集会议参加者的声音，分析收集的声音的特征以获得声波纹数据。在数据存储装置111的存储区域内与诸如序列号码的识别符相关联地存储该声波纹数据。

除了识别符以及声波纹数据的其他信息，例如姓名、图像数据、以及图标数据经连接到连接终端145的个人计算机等提供给根据修改的IC录音机，并和识别符以及声波纹数据相关联地存储，如图6所示。显然，例如，可通过操作提供在IC录音机按键操作单元121上的操作按键来输入姓名，并可用连接到连接终端145的数字照相机捕获图像数据。

如参考图1，2和4所述，在根据修改的IC录音机中也分析收集的声音的特征来检测声波纹数据内的改变点，并自动分配标记到对应于改变点的音频信号位置。当检测到改变点时，检查最近收集的声音的声波纹数据以及音频特征数据库内的声波纹数据的匹配，在分配的标记中包括具有匹配声波纹数据的参加者的识别符。

图7是用于解释分配标记给根据修改的IC录音机记录并收集的音频信号的处理方案的图。该分配标记的处理基本上和参考图2所述的相同。然而，将发言人的识别符附加到标记上。

作为例子，将考虑记录会议进行的情况。假定A在开始记录之后10秒开始讲话，如图2所示。在此情况中，在A开始讲话之前，收集的是静默，或纯语音之外的无意义的声音，即诸如干扰的噪声、拉椅子的声音、或某个物品撞击桌面的声音。因此，收集到的音频信号特征分析结果与A开始讲话之前的结果明显不相同。标识(获取)在音频信号中的改变点的位置，将识别的改变点信息存储为图7中的标记MK1。

在此情况中，检查最近的声波纹数据和音频特征数据库的声波纹数据之间的匹配，并在标记MK1中包括具有匹配声波纹数据的发言人的识别符(会议参加者)。图7还显示将从记录开始之后过去的时间存储为改变点信息的例子。

进一步假定在A停止讲话后不久B开始讲话，B开始讲话之前的一小段时间段是静默或噪声段。在此情况中，同样，当B开始讲话并收集B的声音时，收集的音频信号的特征分析结果和B开始讲话之前明显不相同。因此，如图7中标记MK2所示，存储改变点信息(标记MK2)，以便将标记分配给B

讲话的起始点。

在此情况中，同样的，检查最近的声波纹数据和音频特征数据库的声波纹数据之间的匹配，并在标记 MK2 中包括具有匹配声波纹数据的发言人的识别符(会议参加者)。

此外，还可能出现当 B 讲话时 C 插话的情况。在此情况下，由于 B 的声音不同于 C 的声音，收集的音频信号的分析结果在 B、C 之间不同。因此，如图 7 中标记 MK3 所示，在数据存储装置 111 内存储改变点信息(标记 MK3)，从而将标记分配给 C 讲话的起始点。

在此情况中，同样的，检查最近的声波纹数据和音频特征数据库的声波纹数据之间的匹配，并在标记 MK3 中包括具有匹配声波纹数据的发言人的识别符(会议参加者)。

这样，能够识别记录的音频信号中的哪一部分是谁的讲话。例如，能够轻易地仅播放 A 的讲话并概括 A 的讲话。

对于该修改中的标记的其他信息，例如收集的声音通过语音识别被转换为文本数据，将该文本数据以文本数据文件的形式存储为其他信息。通过使用该文本数据文件，能快速地准备会议记录或讲话概要。

在根据修改的 IC 录音机中，能以类似参考图 1、3 和 5 所述的方式播放记录的声音。此外，在根据修改的 IC 录音机的情况中，能无需播放记录的声音就能识别记录的声音中的每一发言人的讲话。

图 8 是显示 LCD 135 上显示的信息是怎样随着操作而改变的图，该图用于说明当播放记录的音频信号时，将播放位置设定到标记位置的操作。如图 8 所示，当按下 PLAY 按键 211 时，如前所述，CPU 101 控制相关部件，从而从指定的所记录的音频信号的起始处开始播放。

在对应于 A 的讲话的部分，根据如参考图 7 所述的记录处理中分配的标记 MK1，显示关于 A 的讲话的开始时刻 D(1)、对应于发言人图像数据的面部图片 D(2)、发言人姓名 D(3)、以及该讲话起始部分的文本数据 D(4)，并显示播放标记 D(5)，如图 8 的 A 部分所示。

然后，继续播放，当开始播放对应于 B 的讲话的部分时，根据记录处理中分配的播放标记 MK2，显示关于 B 的讲话的开始时刻 D(1)、对应于发言人的图像数据的面部图片 D(2)、发言人姓名 D(3)、以及该讲话起始部分的文本数据 D(4)，和显示播放标记 D(5)，如图 8 的 B 部分所示。

然后，当按下 PREV 按键 215 时，CPU 101 将播放位置设定到由标记 MK1 指示的 A 讲话的起始点，即开始之后的 10 秒(0 分钟，10 秒)，因此播放从那里开始，如图 8 的 C 部分所示。在此情况中，与图 8A 部分所示情况类似地显示关于 A 的讲话的开始时刻 D(1)、对应于发言人的图像数据的面部图片 D(2)、发言人姓名 D(3)、以及该讲话起始部分的文本数据 D(4)，并显示播放标记 D(5)。

然后，当按下 NEXT 按键 214 时，CPU 101 将播放位置设定到 B 的讲话的起始点，即由标记 MK2 指示的开始之后的 1 分钟 25 秒，从而播放从那里开始，如图 8 的 D 部分所示。在此情况中，与图 8B 部分所示情况类似地显示关于 B 的讲话的开始时刻 D(1)、对应于发言人的图像数据的面部图片 D(2)、发言人姓名 D(3)、以及该讲话起始部分的文本数据 D(4)，并显示播放标记 D(5)。

当再次按下 NEXT 按键 214，CPU 101 将播放位置设定到 C 的讲话的起始点，即由标记 MK3 指示的开始之后的 2 分钟 30 秒，从而播放从那里开始，如图 8 的 E 部分所示。在此情况中，显示关于 C 的讲话的开始时刻 D(1)、对应于发言人的图像数据的面部图片 D(2)、发言人姓名 D(3)、以及该讲话起始部分的文本数据 D(4)，并显示播放标记 D(5)。

在该修改中，可以提供一种模式，在该模式中，当快速地按下 NEXT 按键 214 或 PREV 按键 215 两次，例如，当正播放 A 的讲话时，播放位置被设定到对应于 A 的讲话的下一段或前一段，从而播放从那里开始。即，通过重复该操作，能以向前或向后的顺序仅仅播放对应于 A 的讲话的部分。显然，除了 NEXT 按键 214 或 PREV 按键 215，可以提供专用于该模式的操作按键。此时，可以按序自动播放对应于 A 的讲话的部分。

如上所述，在根据修改的 IC 录音机中，在记录处理期间，自动分析收集的音频信号的特征，并对特征中的改变点分配标记。在播放处理中，通过操作 NEXT 按键 214 或 PREV 按键 215，可将播放位置快速设定到分配的标记所指示的记录音频信号的位置，从而从那里开始播放。

此外，在记录音频信号的改变点，能够通过显示发言人的姓名或面部图片来阐明发言人的标识。因此，易于快速地找到感兴趣的发言人，仅播放对应于特定发言人的讲话的部分等。显然，作为识别发言人的信息，可以显示对应特定于每一发言人的图标数据的图标。此外，能够显示讲话开始部分的

文本数据，该文本数据用于区分是否是感兴趣的讲话。

此外，根据该修改的IC录音机的用户被允许通过使用播放时显示的信息快速地将播放位置设定到感兴趣的人的讲话，播放并收听记录的音频信号。因此，用户可以快速地准备关于感兴趣的讲话的会议记录。

即，能够无需播放记录的音频信号就直观地认出是谁在发言，从而易于找到特定发言人的讲话。由于可以使用便于识别发言人的信息，如发言人的面部图片，而不是文本字符串或符号，因此改进了搜索的容易性。

此外，当没有识别出一个发言人时，即当该发言人未注册或当即使该发言人早已注册，但IC录音机未能识别出该发言人时，与该未识别发言人的讲话相关联地分配指示未识别发言人的符号，从而易于找到这一部分。在此情况中，准备会议记录的人播放该未识别发言人的讲话，并识别该发言人。

当将未识别的发言人识别为注册发言人时，可将和该发言人相关的符号分配为标记。当未识别的发言人被识别为未注册发言人时，可执行注册新发言人的操作。从记录的语音中提取该发言人语音的特征，由于该符号与之相关联，可以使用在IC录音机中预先注册的符号或输入到IC录音机的文本字符串、(如果提供的话)由IC录音机的照相机图像形成功能捕获的图像、从外部装置获得的图像数据等。

根据该修改的IC录音机内执行的记录处理与参考图4所述的记录处理类似。然而，当在步骤S113分配指示发言人改变的标记MK1,MK2,MK3,...时，检查和音频特征数据库内d 声波纹数据的匹配以分配相关发言人的识别符。当没有对应的声波纹数据时，分配指示缺少对应声波纹数据的标记。

在根据修改的IC录音机内执行的播放处理与参考图5所述的播放处理类似。然而，当在步骤S217显示指示播放位置的信息时，显示发言人的面部图像、发言人姓名以及代表讲话内容的文本数据等。

虽然从记录起始点过去的时间在根据修改的IC录音机内被用作改变点信息，但不限于此，在数据存储装置111的记录介质上的记录音频信号的地址也可被用作改变点信息。

用于分配标记的执行处理过程定时

在根据第一实施例的IC录音机以及根据第一实施例修改的IC录音机中，检测收集的声音内的改变点，并在记录处理中分配标记给对应于改变点的音

频信号的位置。然而，不限于第一实施例及其修改，可在记录处理完成之后分配标记。即，可在播放处理中分配标记，或可单独地执行标记分配处理。

图 9 是在记录处理完成之后用于分配标记给记录的音频信号内的改变点的处理流程图。即，在播放处理期间给记录的声音内的改变点分配标记时或当单独执行给记录的声音内的改变点分配标记的处理时执行图 9 所示处理。图 9 所示处理也由 IC 录音机的控制相关部件的 CPU 101 执行。

CPU 101 命令文件处理器 110 以预定尺寸的单元读取存储在数据存储装置 111 的音频文件内的压缩记录音频信号(步骤 S301)，并确定是否读取了所有记录的音频信号(步骤 S302)。

如果在步骤 S302 确定没有读取全部记录的音频信号，CPU 101 命令数据扩展器 142 扩展压缩的记录音频信号(步骤 S303)。然后，CPU 101 命令音频特征分析器 143 分析扩展音频信号的特征以获得声波纹数据，并比较该声波纹数据和先前获得的声波纹数据，由此确定记录的音频信号的特征是否改变(步骤 S305)。

如果在步骤 S305 确定记录的音频信号的特征没有改变，从步骤 S301 重复该处理。如果在步骤 S305 确定记录的音频信号的特征改变了，CPU 101 确定发言人已经改变，并命令文件处理器 110 分配标记给音频特征改变的点(步骤 S306)。

因此，文件处理器 110 将指示从文件起始处过去的时间的信息或指示对应记录位置的地址的信息写入到数据存储装置 111 的数据库区域 111(1)，作为指示关于音频文件 111(2)的音频特征中的改变点的信息。在此情况中，音频文件以及指示音频特征中的改变点的信息相互关联的存储。

在步骤 S306 之后，CPU 101 对下一时段(下一处理单元)的音频信号重复从步骤 S301 的步骤。然后，如果在步骤 S302 确定读取了所有的记录音频信号，执行预定的终止处理(步骤 S307)，退出图 9 所示处理。

因此，在记录处理之后，能够在播放处理期间检测到记录的声音内的改变点，并为记录的声音分配标记，或独立执行分配标记给记录的声音的处理。当在播放处理中分配标记时，如 9 所示在步骤 S303 扩展的音频信号被 D/A 转换，得到的模拟音频信号被提供给扬声器 133。

如上所述，通过在记录之后给记录的音频信号的特征中的改变点分配标记，可以降低用于记录的处理负载以及功耗。此外，由于用户可能不希望在

每个记录中自动分配标记，允许设定是否在记录期间自动分配标记。当用户关闭自动标记分配功能来执行记录时，如果稍后希望分配标记，则允许用户甚至如上所述在记录处理之后分配标记给记录的音频信号，这是非常方便的。

此外，由于可如上所述分配标记给记录的音频信号，应用到不具有记录功能但具有信号处理功能的设备是可能的。例如，该实施例可被应用于用在个人计算机的应用软件。在此情况下，由音频记录设备记录的音频信号被传送到个人计算机，从而可由运行在个人计算机上的信号处理应用软件来分配标记。

此外，通过网络等共享由根据本实施例的设备创建的数据，能够将数据自身用作会议记录而无需转录该数据。

因此，本实施例可应用于能进行信号处理的各种电子设备，而不仅仅限于记录设备。因此，使用根据本实施例的电子装置来处理音频信号，对于已经记录的音频信号可以获得类似结果。即，可以更有效地准备会议记录。

此外，如前所述，图 1 所示根据第一实施例的 IC 录音机包括通信 I/F144，从而该 IC 录音机能连接到电子设备，诸如个人计算机。因此，通过传送由该 IC 录音机记录的数字音频信号，包括分配给改变点的标记到个人计算机，能够在具有较大屏幕的个人计算机的显示器上显示更详细的信息。这允许快速搜索感兴趣的发言人的讲话。

图 10 和 11 是显示基于从根据第一实施例的 IC 录音机传送到给个人计算机的记录信号以及分配给记录信号的改变点信息(标记信息)，在连接到个人计算机的显示器 200 的显示屏幕上显示的改变点信息的例子。

在图 10 所示例子中，显示与记录的音频信号相关的时间范围指示 201，以及在时间范围指示 201 的适当位置上显示标记(改变点)MK1、MK2、MK3、MK4…。因此，能一眼就认出多个改变点的位置。此外，例如使用诸如鼠标的点击装置通过放置其上的光标点击某个标记，能够从那里播放记录的声音。

在图 11 所示的例子中，在显示器 200 的显示屏幕上同时显示图 8 所示的多组项。更具体的，显示发言人的面部图片 211(1)、211(2)、211(3)…，以及对应于讲话内容的文本数据 212(1)、212(2)、212(3)…，允许快速搜索感兴趣的发言人的讲话。此外，还能使用个人计算机的功能显示标题指示 210。

在图 11 所示的例子中，左侧的“00”、“01”、“02”、“03”…指示从记录声音的开始点过去的时间。显然，可以实现多种显示模式，例如显示图 8 所示

的多个组的项的模式。

通过将数据传送到诸如个人计算机的具有较大显示器的设备，在该数据中，用识别发言人的信息(符号)来识别记录的讲话，能够无需转录音频数据就准备好会议记录。即，由根据第一实施例的 IC 录音机记录的数据被直接用作会议记录。

此外，使用诸如插入件从而使数据在网页上变得可用并可通过网页浏览器浏览的软件，能通过网络共享会议记录。这有助于极大地降低用于共享信息，例如使信息有效可用的劳动以及时间。

第二实施例

IC 录音机的结构和操作概览

图 12 是根据本发明第二实施例的记录/播放设备 IC 录音机的框图。除了提供两个麦克风 131(1)和 131(2)以及一个音频信号处理器 136 之外，根据第二实施例的 IC 录音机和图 1 所示根据第一实施例的 IC 录音机的构成相同，该处理器用于处理从两个麦克风 131(1)和 131(2)输入的音频信号。因此，对于根据第二实施例的 IC 录音机，用相同标号表示那些对应于根据第一实施例的 IC 录音机的部分，省略对它们的详细说明。

在根据第二实施例的 IC 录音机中，从两个麦克风 131(1)和 131(2)输入的收集的音频信号由音频信号处理器 136 处理，以识别发言人位置(音源位置)，从而可以考虑发言人的位置来识别收集的音频信号中的改变点(发言人改变点)。即，当使用由音频分析而获得的声波纹数据检测到收集的音频信号内的改变点时，根据两个麦克风收集的声音的发言人位置被用作辅助信息，从而可以更准确的识别发言人或改变点。

图 13 是显示麦克风 131(1)和 131(2)以及音频信号处理器 136 的结构例子的图。在图 13 所示的例子中，两个麦克风 131(1)和 131(2)中的每一个都是单向的，如图 13 所示。麦克风 131(1)和 131(2)背靠背相互靠近的分布，从而它们方向性的主方向是相对的。因此，麦克风 131(1)优选地收集发言人 A 的讲话，而麦克风 131(2)优选地收集发言人 B 的讲话。

如图 13 所示，音频信号处理器 136 包括：加法器 1361、比较器 1362、以及 A/D 转换器 1363。由麦克风 131(1)和 131(2)的每一个收集的音频信号被提供给加法器 1361 以及比较器 1362。

加法器1361将麦克风131(1)收集的音频信号和麦克风131(2)收集的音频信号相加，将音频信号之和提供给A/D转换器1363。麦克风131(1)收集的音频信号和麦克风131(2)收集的音频信号之和可用下面的等式(1)表示，等效于由一个无方向性麦克风收集的音频信号。

$$((1+\cos\theta)/2)+((1-\cos\theta)/2)=1 \qquad (1)$$

比较器1362比较麦克风131(1)收集的音频信号和麦克风131(2)收集的音频信号。当麦克风131(1)收集的音频信号电平较高时，比较器1362确定主要是发言人A在讲话，并提供具有值为"1"(高电平)的发言人区别信号给控制器100。另一方面，当麦克风131(2)收集的音频信号电平较高时，比较器1362确定主要是发言人B在讲话，提供具有值为"0"(低电平)的发言人区别信号给控制器100。

因此，根据麦克风131(1)收集的音频信号和麦克风131(2)收集的音频信号确定发言人位置，允许区分发言人A的讲话和发言人B的讲话。

如果第三发言人C从穿过麦克风131(1)和麦克风131(2)的方向性的主方向的方向讲话，即，从斜对发言人A和B的位置(图13的横向方向)，麦克风131(1)和131(2)收集的音频信号的电平实际相等。

为处理在此位置的发言人C的讲话，可为比较器1362定义两个阈值，确定当电平差在±V之内时发言人是位于横向方向的发言人C，当电平差高于+V时发言人是发言人A，而当电平差低于-V时发言人是发言人B。

通过预先认出在麦克风131(1)的方向性方向上的发言人、麦克风131(2)方向性方向上的发言人、横贯麦克风131(1)和麦克风131(2)方向性方向上的发言人，使得可识别发言人。因此，当根据通过分析收集的声音的特征获得的声波纹数据来检测到改变点时，可通过考虑由麦克风收集到的声音电平来更准确地识别发言人。

麦克风和音频信号处理器的另一个例子

可选的，麦克风131(1)和131(2)以及音频信号处理器136可如图14的构成。图14是显示麦克风131(1)和131(2)以及音频信号处理器136的另一种示范结构的图。在图14所示例子中，两个麦克风131(1)和131(2)是无方向性的，如图14所示。麦克风131(1)和131(2)相互临近分布，例如中间的缝隙大约为1厘米。

如图14所示，本例中的音频信号处理器136包括加法器1361、A/D转

换器 1363、减法器 1364、以及相位比较器 1365。由麦克风 131(1)和 131(2)中的每一个收集的音频信号被提供给加法器 1361 和减法器 1364。

从加法器 1361 输出的和信号等效于无方向性麦克风的输出，从减法器 1364 输出的减法信号等效于双向性(**8-figure directivity**)麦克风的输出。双向性麦克风的输出的相位根据声波的入射方向为正或负。因此，加法器 1361 的和输出(无方向性输出)的相位与减法器 1364 的减法输出相位由相位比较器 1365 比较，以确定减法器 1364 的减法输出的极性，从而识别发言人。

即，当减法器 1364 的减法输出的极性是正时，确定收集的是 A 发言人的讲话。另一方面，当减法器 1364 的减法输出的极性是负时，确定收集的是 B 发言人的讲话。

此外，与参考图 13 所述的例子相同，当处理斜对发言人 A 和 B(在图 14 的横向方向上)的发言人 C 的讲话时，对应发言人 C 的讲话的收集的音频信号的减法输出电平较小。因此，通过检查加法器 1361 的和输出以及减法器 1364 的减法输出电平，能够识别发言人 C 的讲话。

虽然图 14 所示音频信号处理器 136 包括加法器 1361，加法器 1361 不是必需的部件。例如，麦克风 131(1)和 131(2)之一的输出信号可以被提供给 A/D 转换器 1363 以及相位比较器 1365。

如上所述，在图 13 和 14 所示的例子中，在记录处理中，能够使用两个麦克风 131(1)和 131(2)收集的声音的电平或极性来识别发言人的位置。此外，通过考虑该识别结果，能检测收集的声音内的改变点，并准确地识别发言人。

当播放处理期间给记录的声音分配标记时或当独立执行给记录的声音分配标记的处理时能够采用图 13 和 14 所示的方案。

例如，当在记录处理后应用参考图 13 说明的方案时，单向麦克风 131(1)和 131(2)收集的音频信号由 2-信道立体声录音来记录，如图 15A 所示。而在播放处理或独立执行分配标记的处理时，扩展从数据存储装置 111 读取的两个信道的压缩音频信号，然后两个信道的扩展音频信号被输入到具有和图 13 所示比较器 1362 相同功能的比较器中。

因此，能够确定是主要使用麦克风 131(1)收集的音频信号还是主要使用麦克风 131(2)收集的音频信号。因此，能够根据该确定结果以及预先知道的相对于每一麦克风的发言人的位置来识别发言人。

类似地，在记录处理之后应用参考图 14 说明的方案时，由 2-信道立体

声录音记录从麦克风131(1)和131(2)输出的信号，并在播放处理期间或独立执行分配标记的处理时，可通过由图14所示的音频信号处理器136执行的相同处理来识别发言人。

当使用从麦克风131(1)和131(2)的输出信号识别发言人时，预先准备的指示发言人相对于每一麦克风131(1)和131(2)的位置的信息存储在IC录音机中，例如以图16所示的发言人位置数据库的形式。

图16是显示发言人位置数据库的例子的图。在此例中，发言人位置数据库包括：对应于来自IC录音机的音频信号处理器136的识别结果的发言人区别信号、与各个发言人区别信号相关的麦克风的识别信息、以及主要使用麦克风的候选发言人的发言人识别符。如图16所示，能够与单个麦克风相联系地登记多个麦克风。

图16所示的发言人位置数据库最好在会议之前预先创建。通常，会议参加者以及这些参加者的座位是预先确定的。因此，考虑设定IC录音机的位置，能够在会议之前创建发言人位置数据库。

当会议参加人改变而没有提前通知时，或在会议期间座位改变，例如，没有使用基于麦克风收集的声音的发言人的识别时，仅仅根据音频分析获得的声波纹数据来检测改变点。可选地，在记录处理后可将发言人位置数据库调整为准确的，并重新分配标记给记录的声音。

通过使用图16所示的发言人位置数据库，能够识别发言人位置以及识别在该发言人位置的发言人。

虽然在第二实施例中使用两个麦克风131(1)和131(2)以及涉及两个或三个发言人，但是麦克风的数目不限于2，发言人的数目也不限于3。使用大量麦克风能够识别较大数目的发言人。

此外，通过从麦克风输出的信号来识别发言人的位置进而识别发言人的方案不限于参考图13和14所述的。例如，也可以使用相邻放置的四点麦克风方法或相邻放置的三点麦克风方法。

在相邻放置的四点麦克风方法中，相邻放置四个麦克风M0、M1、M2和M3，从而麦克风之一不在其他三个麦克风所限定的平面上，如图17A所示。考虑到在由四个麦克风M0、M1、M2和M3收集的音频信号的时间结构上的细微差异，由短期相关、声强等计算诸如声源位置或规模的空间信息。这样，通过使用至少四个麦克风，能够准确地识别发言人位置，并根据该发

言人位置(座位位置)识别发言人。

当发言人实际是在一个水平平面的假设可以接受时，则在一个水平平面内提供相互邻近的三个麦克风就足够了，如图 17B 所示。

此外，麦克风的配置无需像图 17A 和 17B 所示的是正交的。例如，在图 17B 所示相邻配置的三点麦克风方法中，麦克风的配置可以是这样的，三个麦克风分布在等边三角形的顶点。

第二实施例的修改

在如上所述根据第二实施例的 IC 录音机中，当使用音频分析获得的声波纹数据来检测收集的音频信号内的改变点时，根据从两个麦克风收集的声音来考虑主要使用的麦克风的区分的结果，从而改进检测音频信号内的改变点的精确性。然而，其他配置也是可能的。

例如，可能提供包括两个麦克风 131(1)和 131(2)以及音频信号处理器 136 但不包括音频特征分析器 143 的 IC 录音机，如图 18 所示。即，除了不提供音频特征分析器 143，如 18 所示的 IC 录音机和图 12 所示的根据第二实施例的 IC 录音机的结构相同。

能够仅根据主要使用的麦克风的区别的结果，两个麦克风 131(1)和 131(2)收集的声音来检测出发言人改变点，根据主要使用的麦克风的区别的结果检测发言人改变，并分配标记到对应于该改变点的音频信号的位置。在该例中，无需分析音频特征的处理，因此降低了 CPU 101 的负载。

虽然在上述实施例中给要处理的音频信号内的改变点分配标记，但是还能仅给发言人的改变点分配标记，从而更有效的搜索变为可能。例如，根据要处理的音频信号的信号电平或声波纹数据，话音段与其他段，例如噪声被非常清楚区分，仅对话音段的起始点分配标记。

此外，根据音频信号频率特征数据或声波纹数据，能够区分发言人是男还是女，报告在改变点的发言人的性别的区别。

此外，根据以如上方式分配的标记信息，能够提供例如，仅用于搜索的搜索模式；用于改变分配的标记的位置、删除标记或增加标记的标记编辑模式；或用于仅播放基于分配的标记所指定的发言人的讲话的特定播放模式，例如仅仅是 A 的讲话。通过添加代码到 CPU 101 执行的程序中可以相对比较容易地实现这些模式。

此外，可提供数据库更新功能，从而例如可以用用于检测改变点的声波纹数据更新图 6 所示的音频特征数据库内的声波纹数据，以此改进音频特征数据库的准确性。例如，甚至在比较声波纹数据的处理中某发言人的声波纹数据未发现匹配时，如果该发言人的声波纹数据实际上存在于该音频特征数据库内，则用最近获得的声波纹数据来替换音频特征数据库内的声波纹数据。

此外，当比较处理中某发言人的声波纹数据和另一个发言人的声波纹数据匹配时，可以做出设置，从而在比较处理中不使用该不同的发言人的声波纹数据。

当声波纹数据和多个发言人的声波纹数据匹配时，定义使用的声波纹数据的优先级，从而该声波纹数据仅和正确的发言人的声波纹数据匹配。

此外，可给讲话的结束点以及起始点都分配标记。此外，分配标记的位置可以改变，例如，考虑各个用户的方便性，在起始点之前或之后几秒。

此外，如前所述，可使用一种或多种不同方法来分析音频信号的特征，而不限制于声波纹分析，以便可以获得精确的分析数据。

虽然以上主要以使用两个麦克风的情况来说明第二实施例，但是麦克风的数目不限于此，可以是不小于 2 的任何数。使用诸如信号电平、极性、或各个麦克风收集声音的延迟时间等不同参数来识别发言人位置，使得能够根据该发言人位置来识别发言人。

此外，虽然第一和第二实施例都是以将本发明应用于记录并播放音频信号的设备——IC 录音机的情况来说明的，但是本发明的应用并不限于 IC 录音机。例如，本发明可应用于记录设备、播放设备以及和不同记录介质一起使用的记录/播放设备，例如诸如硬盘的磁光盘，以及 MD 或诸如 DVD 的光盘。

软件实施

本发明还可以用程序实现，当由 CPU 101 执行时，实现根据上述实施例的音频特征分析器 143、音频信号处理器 136 以及 IC 录音机的其他处理单元的功能，并有效连接这些功能。即，本发明可通过准备用于执行如图 4 和 5 所示的流程的程序以及由 CPU 101 执行该程序来实现。

此外，与上述实施例类似，由录音机记录的音频数据可由个人计算机捕获，该个人计算机内部具有实现音频特征分析器 143 的功能的程序，以便该个人计算机能够检测发言人改变。

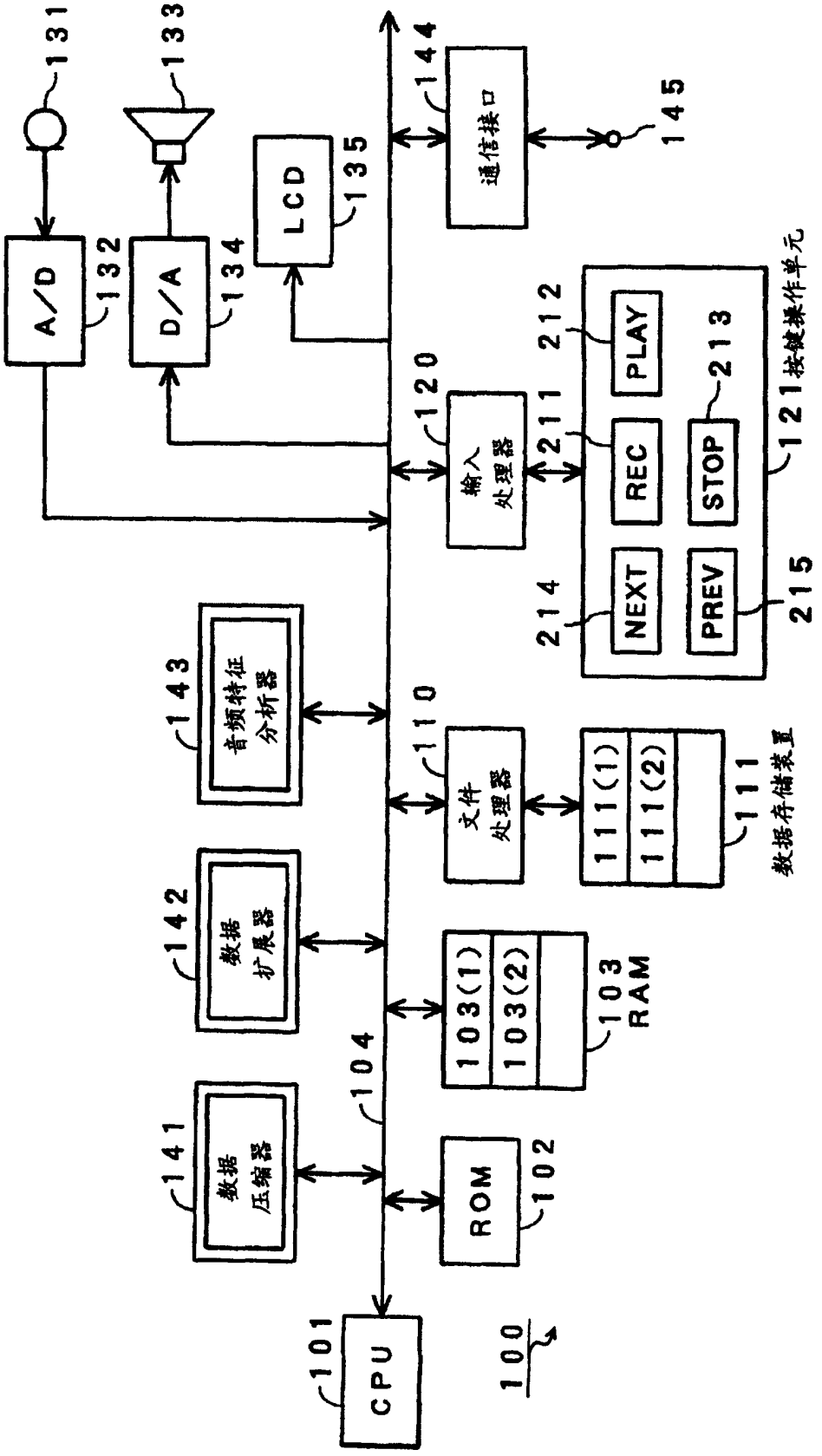

131
133
144
145
A/D
132
D/A
134
LCD
135
通信接口
120
输入处理器
211
212
PLAY
REC
STOP
213
121按键操作单元
214
NEXT
PREV
215
143
音频特征分析器
110
文件处理器
111(1)
111(2)
111
数据存储装置
142
数据扩展器
103(1)
103(2)
103
RAM
104
141
数据压缩器
ROM
102
CPU
101
100

图 1

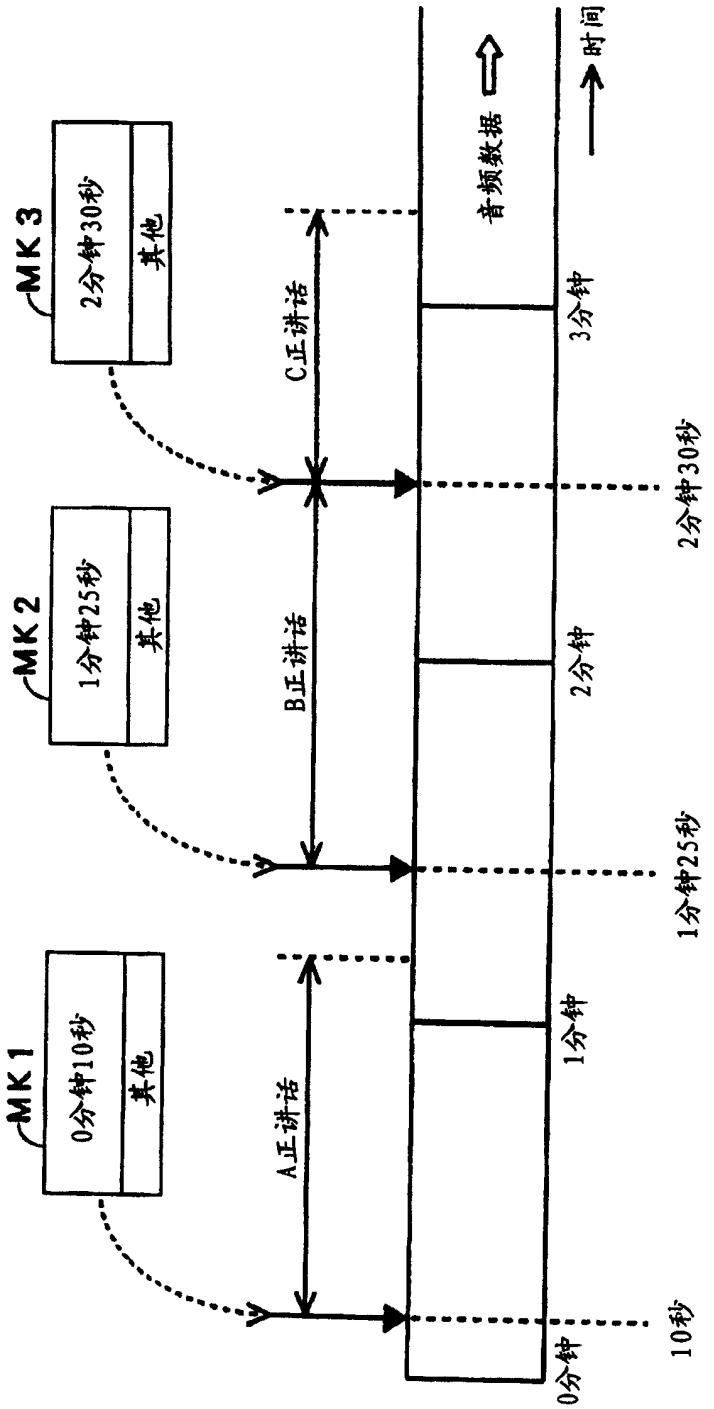

MK 3
2分钟30秒
其他
MK 2
1分钟25秒
其他
MK 1
0分钟10秒
其他
C正讲话
B正讲话
A正讲话
音频数据
时间
3分钟
2分钟
1分钟
2分钟30秒
1分钟25秒
10秒
0分钟

图 2

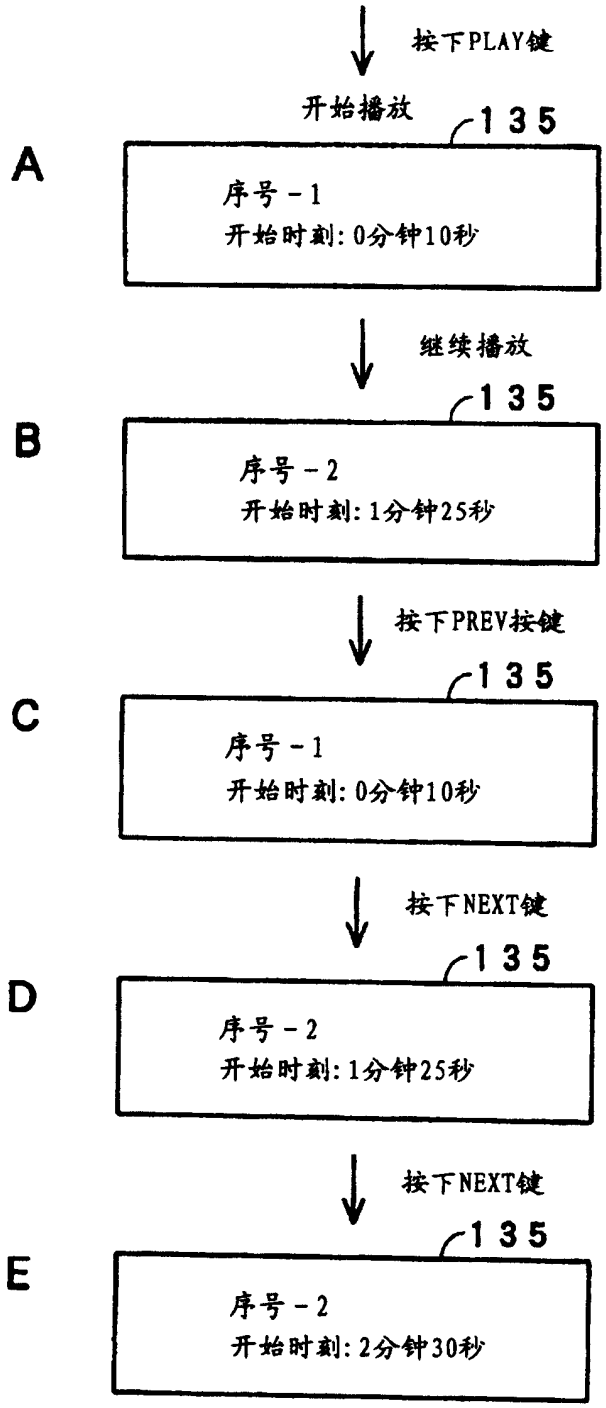

按下PLAY键
开始播放
135
A
序号－1
开始时刻：0分钟10秒
继续播放
135
B
序号－2
开始时刻：1分钟25秒
按下PREV按键
135
C
序号－1
开始时刻：0分钟10秒
按下NEXT键
135
D
序号－2
开始时刻：1分钟25秒
按下NEXT键
135
E
序号－2
开始时刻：2分钟30秒

图 3

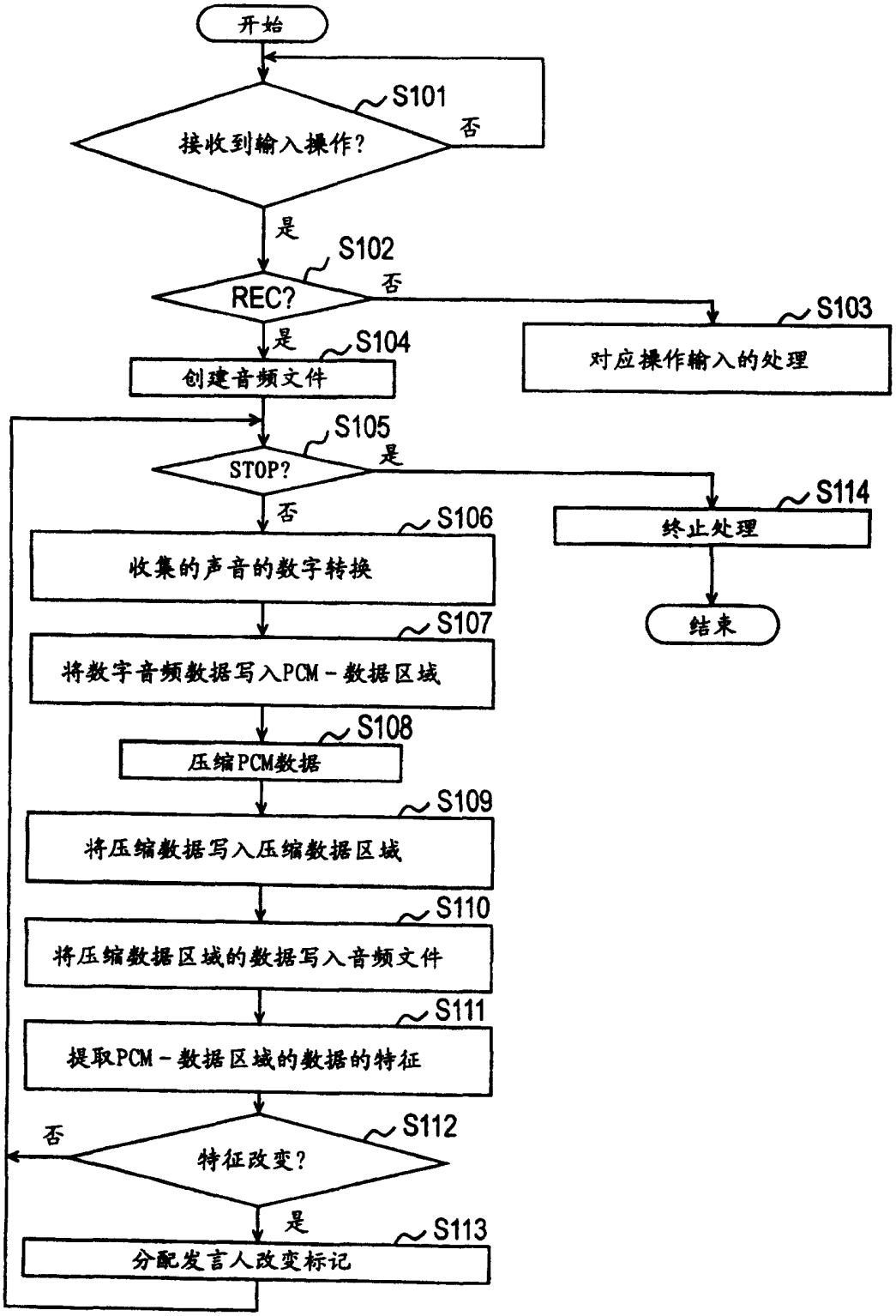

开始
S101
接收到输入操作?
否
是
S102
REC?
否
是
S103
对应操作输入的处理
S104
创建音频文件
S105
STOP?
是
否
S114
终止处理
结束
S106
收集的声音的数字转换
S107
将数字音频数据写入PCM－数据区域
S108
压缩PCM数据
S109
将压缩数据写入压缩数据区域
S110
将压缩数据区域的数据写入音频文件
S111
提取PCM－数据区域的数据的特征
S112
特征改变?
否
是
S113
分配发言人改变标记

图 4

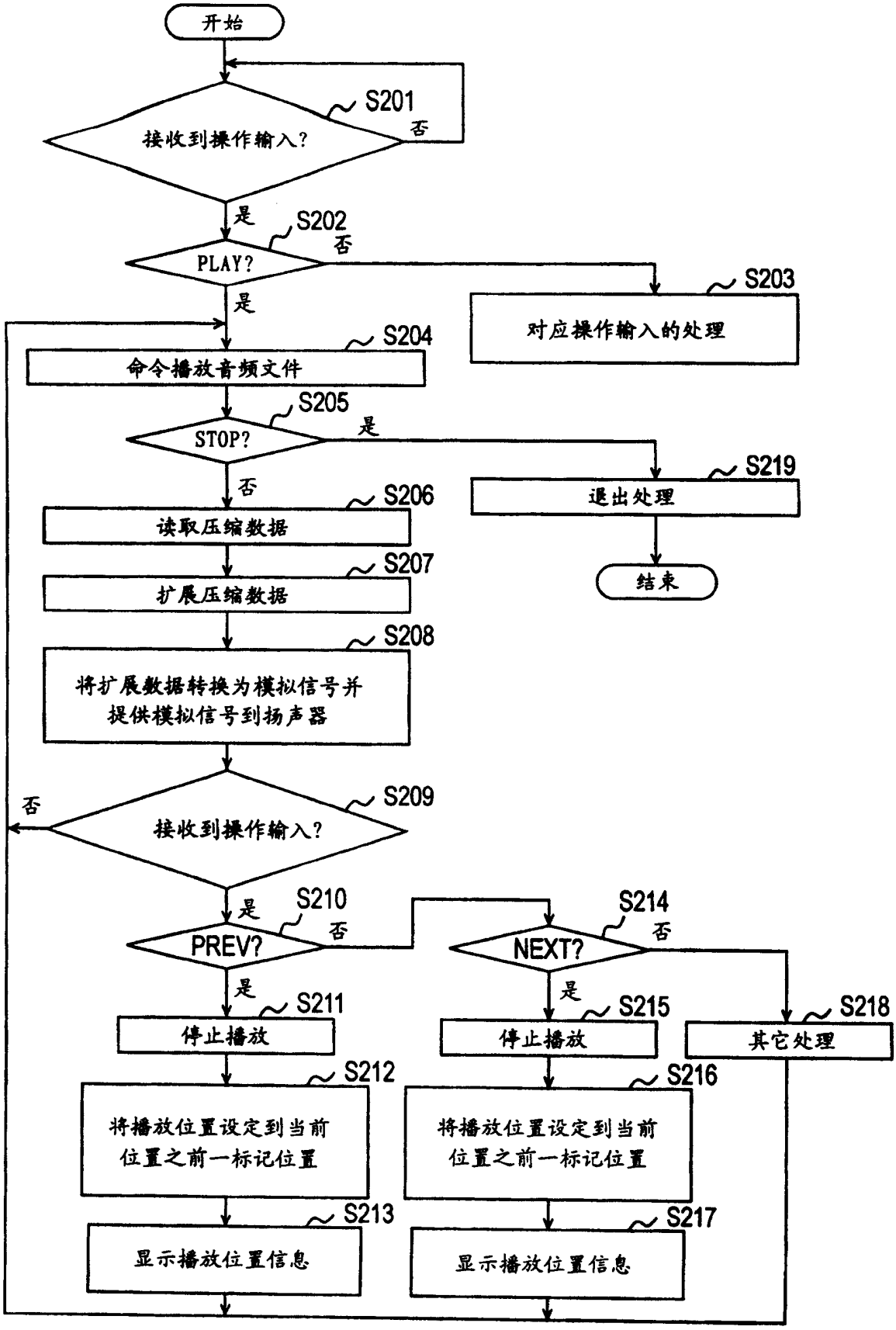
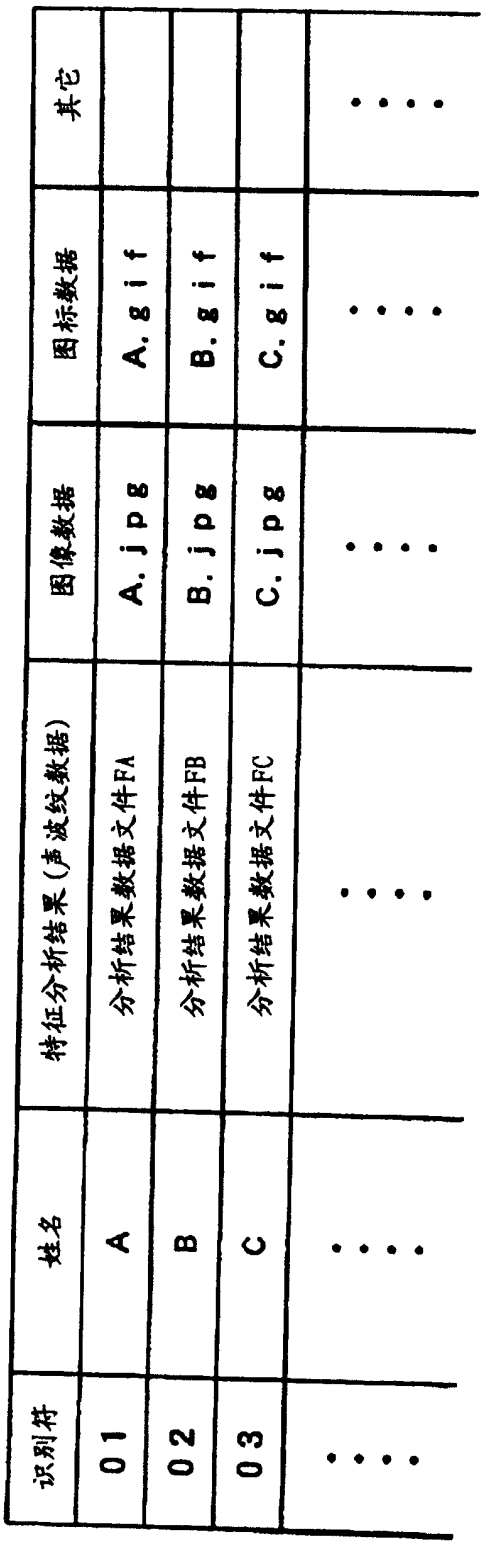

开始
S201
接收到操作输入?
否
是
S202
PLAY?
否
是
S203
对应操作输入的处理
S204
命令播放音频文件
S205
STOP?
是
S219
退出处理
结束
否
S206
读取压缩数据
S207
扩展压缩数据
S208
将扩展数据转换为模拟信号并
提供模拟信号到扬声器
S209
接收到操作输入?
否
是
S210
PREV?
否
是
S211
停止播放
S212
将播放位置设定到当前
位置之前一标记位置
S213
显示播放位置信息
S214
NEXT?
否
是
S215
停止播放
S216
将播放位置设定到当前
位置之前一标记位置
S217
显示播放位置信息
S218
其它处理

图 5



识别符	姓名	特征分析结果(声波纹数据)	图像数据	图标数据	其它
01	A	分析结果数据文件FA	A.jpg	A.gif	
02	B	分析结果数据文件FB	B.jpg	B.gif	
03	C	分析结果数据文件FC	C.jpg	C.gif	
...	...	...	...	...	...

图 6

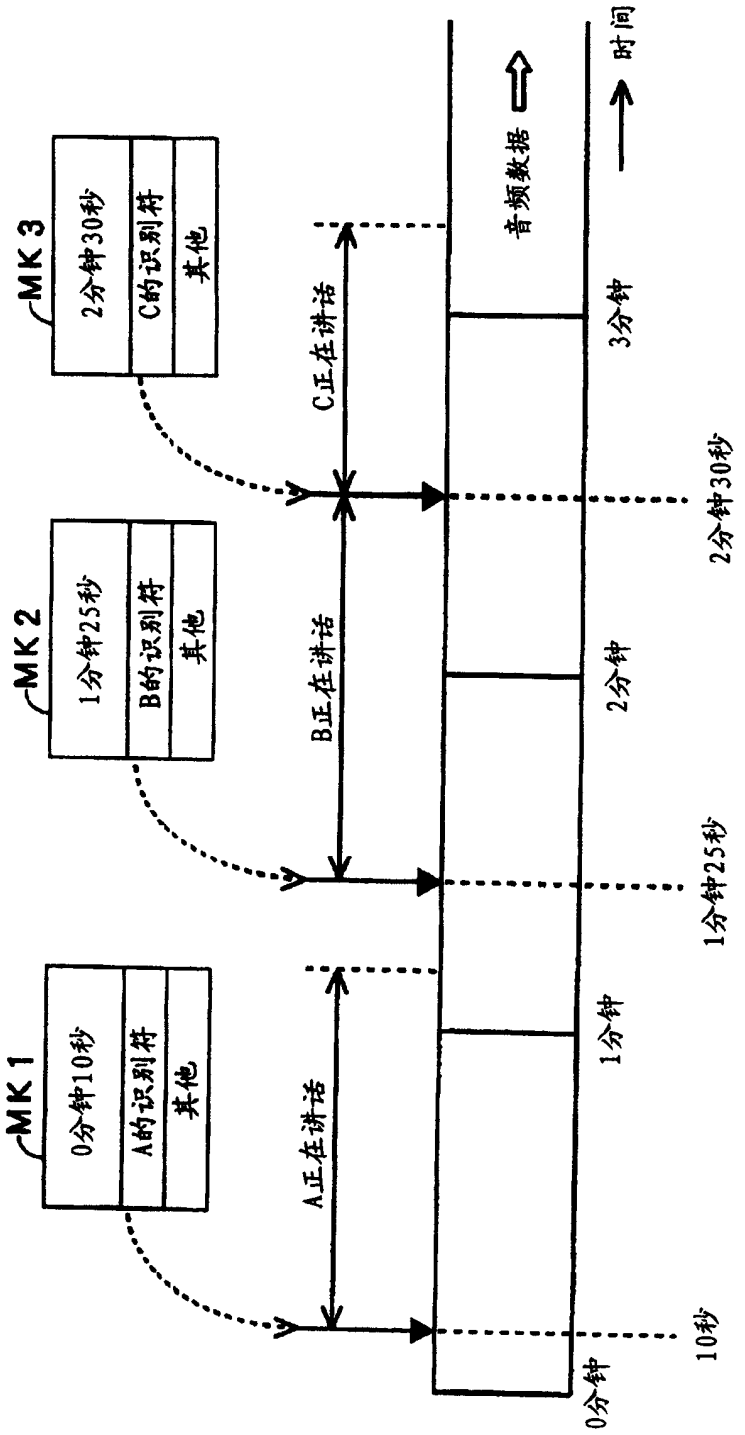

MK 1
0分钟10秒
A的识别符
其他
MK 2
1分钟25秒
B的识别符
其他
MK 3
2分钟30秒
C的识别符
其他
A正在讲话
B正在讲话
C正在讲话
音频数据
时间
0分钟
10秒
1分钟
1分钟25秒
2分钟
2分钟30秒
3分钟

图 7

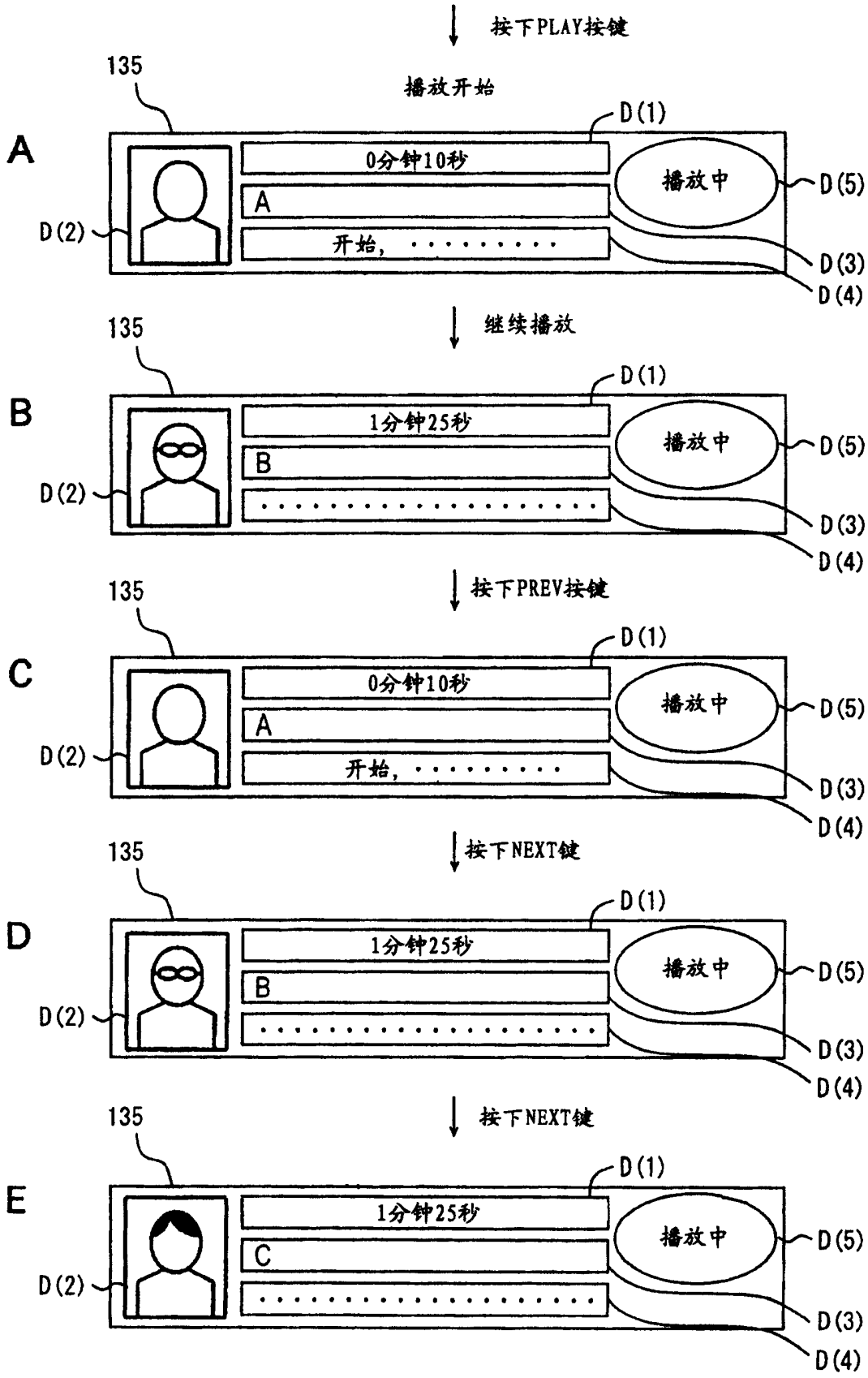

按下PLAY按键
135
播放开始
D(1)
A
0分钟10秒
播放中
D(5)
A
D(2)
开始，・・・・・・・・・
D(3)
D(4)
继续播放
135
D(1)
B
1分钟25秒
播放中
D(5)
B
D(2)
・・・・・・・・・・・・・・・・・・・・
D(3)
D(4)
按下PREV按键
135
D(1)
C
0分钟10秒
播放中
D(5)
A
D(2)
开始，・・・・・・・・・
D(3)
D(4)
按下NEXT键
135
D(1)
D
1分钟25秒
播放中
D(5)
B
D(2)
・・・・・・・・・・・・・・・・・・・・
D(3)
D(4)
按下NEXT键
135
D(1)
E
1分钟25秒
播放中
D(5)
C
D(2)
・・・・・・・・・・・・・・・・・・・・
D(3)
D(4)

图 8

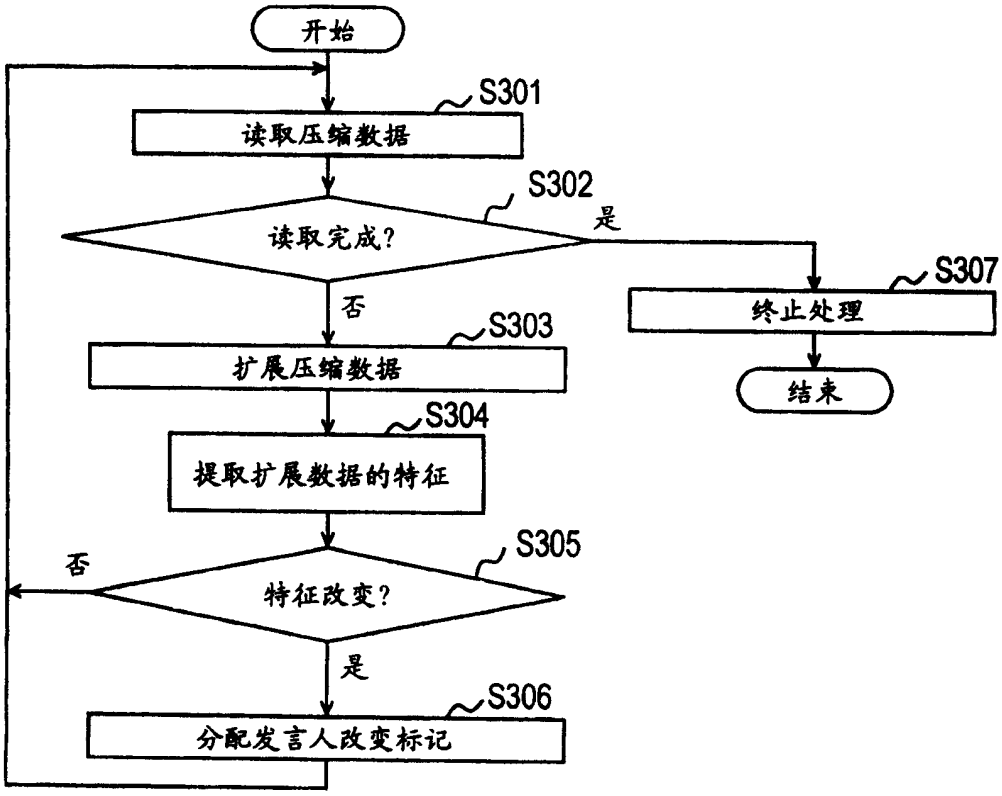

开始
S301
读取压缩数据
S302
读取完成?
是
否
S307
终止处理
结束
S303
扩展压缩数据
S304
提取扩展数据的特征
S305
否
特征改变?
是
S306
分配发言人改变标记

图 9

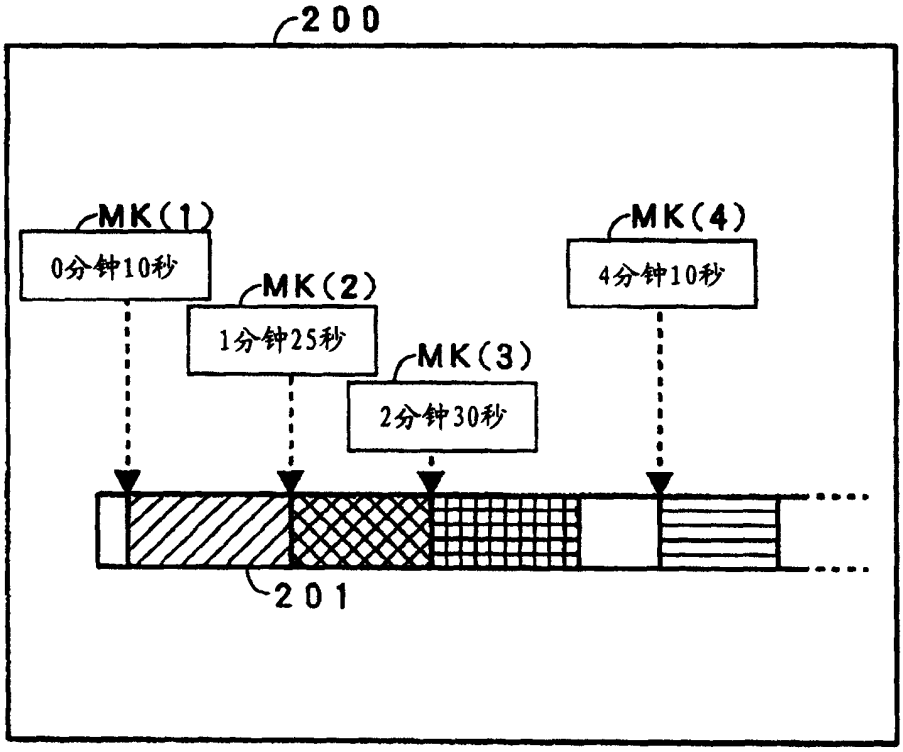


图 10

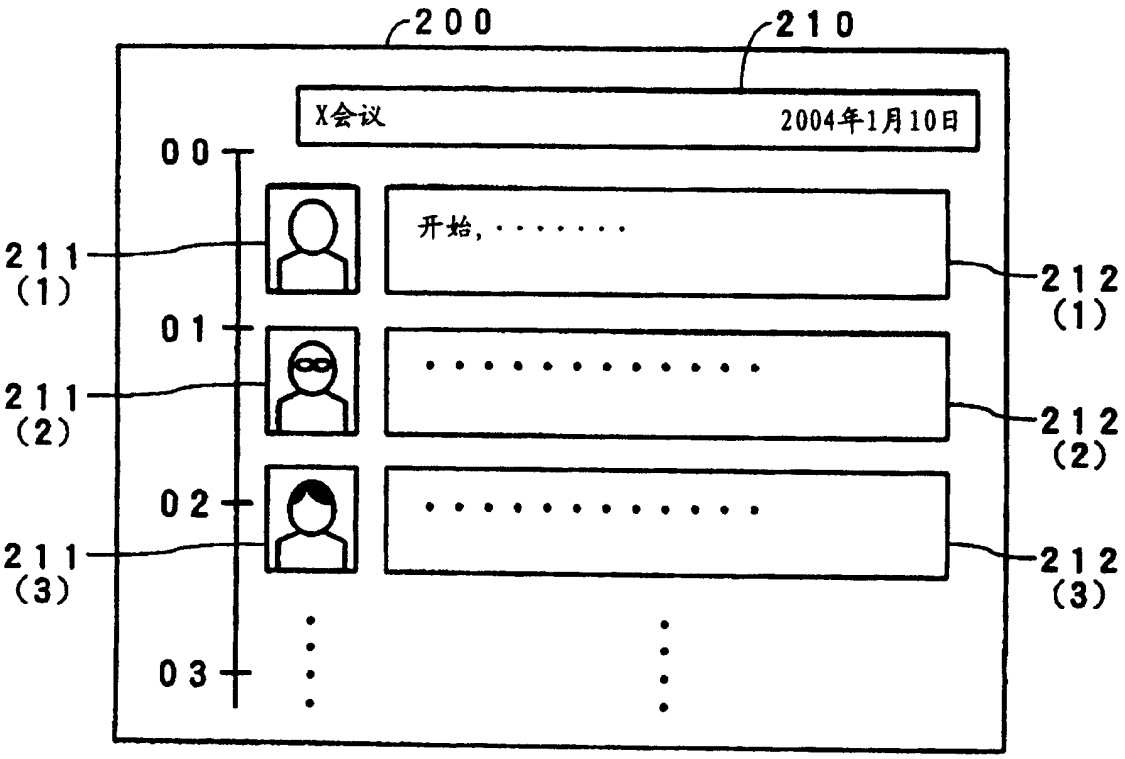

200
210
X会议
2004年1月10日
00
01
02
03
开始,·····
211 (1)
211 (2)
211 (3)
212 (1)
212 (2)
212 (3)

图 11

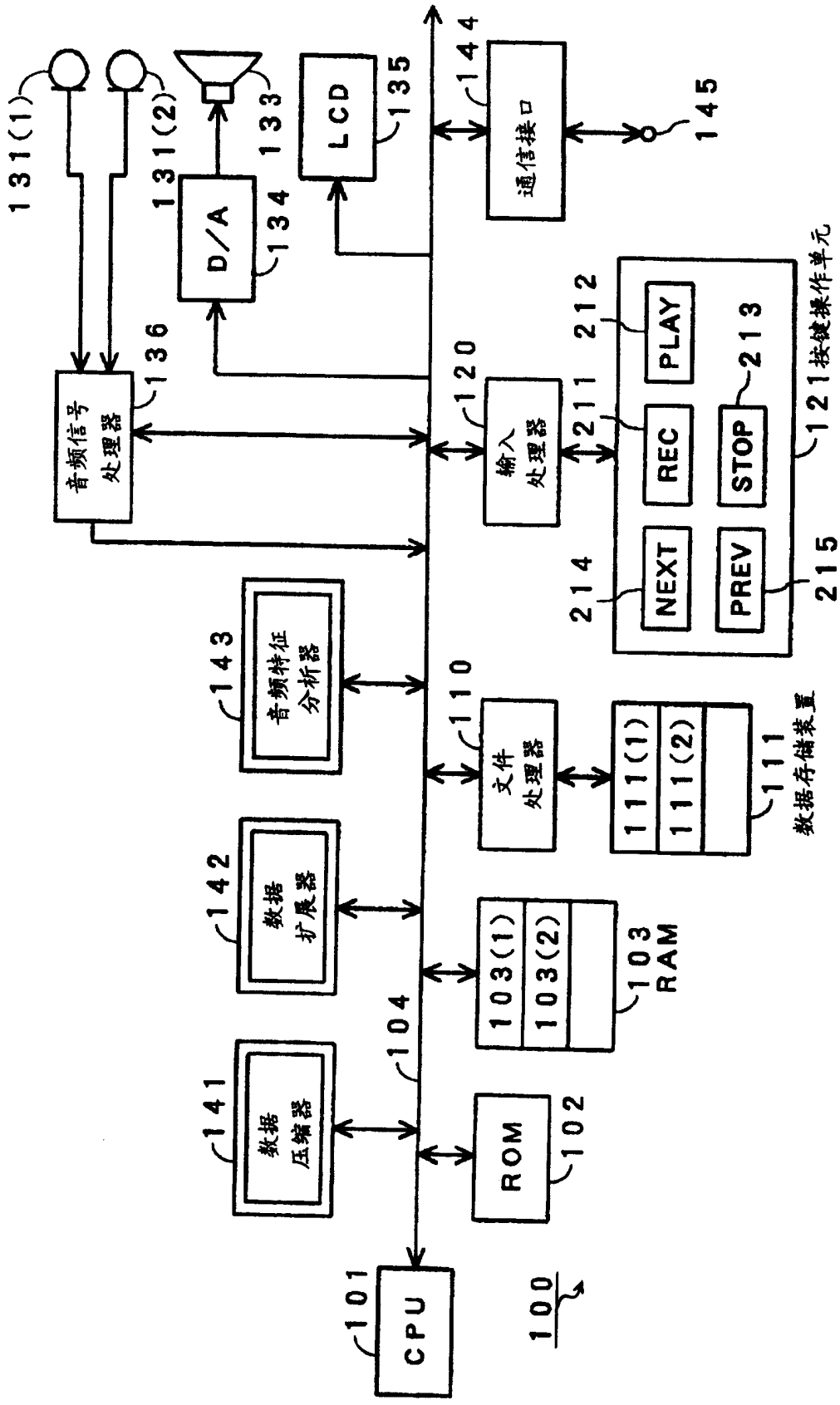

131(1)
131(2)
133
LCD
135
144
145
通信接口
D/A
134
136
音频信号处理器
120
输入处理器
211
212
PLAY
213
REC
STOP
121按键操作单元
214
NEXT
PREV
215
143
音频特征分析器
110
文件处理器
111(1)
111(2)
111
数据存储装置
142
数据扩展器
103(1)
103(2)
103
RAM
104
141
数据压缩器
ROM
102
CPU
101
100

图 12

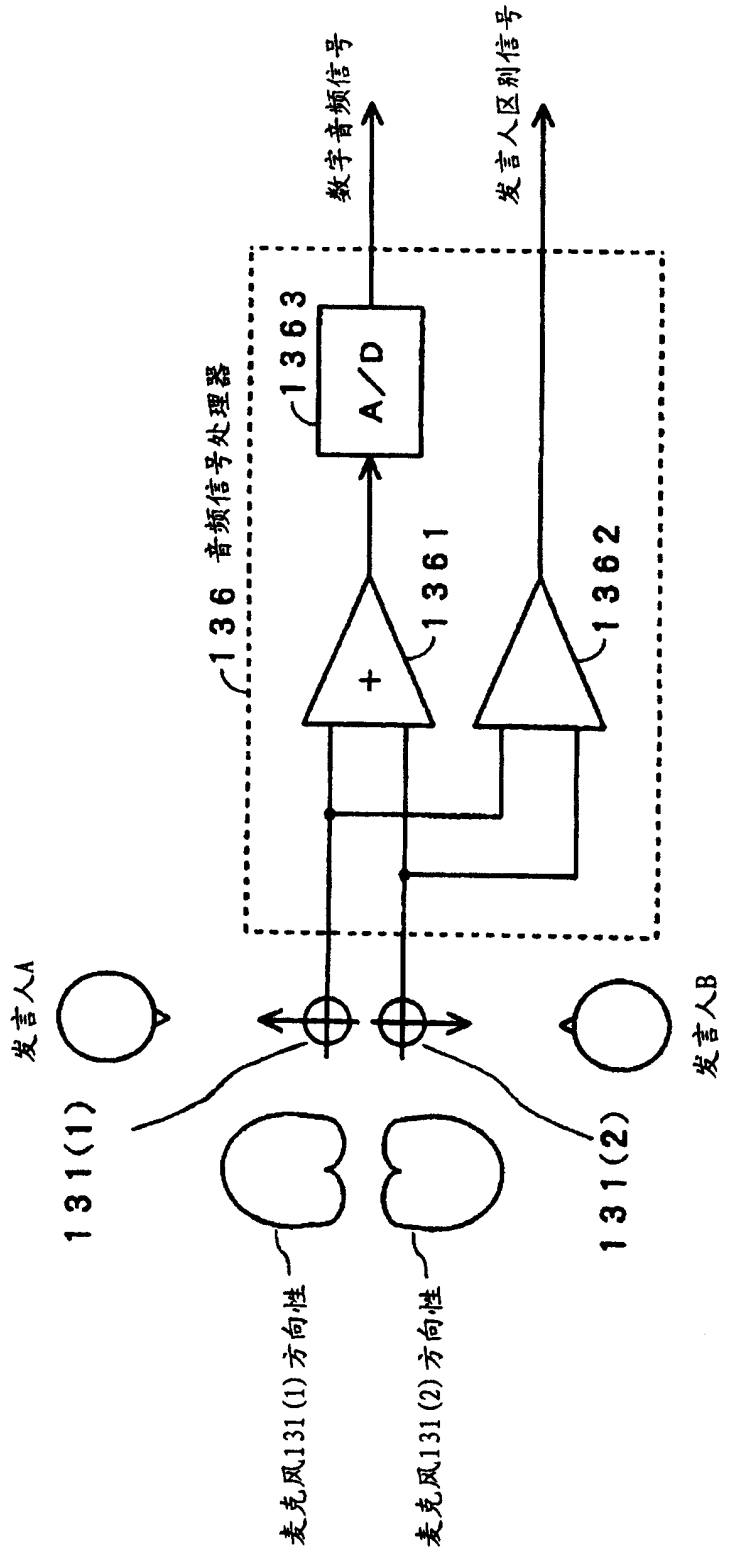

发言人A
131(1)
麦克风131(1)方向性
麦克风131(2)方向性
131(2)
发言人B
136 音频信号处理器
+
1361
1362
A/D
1363
数字音频信号
发言人区别信号

图 13

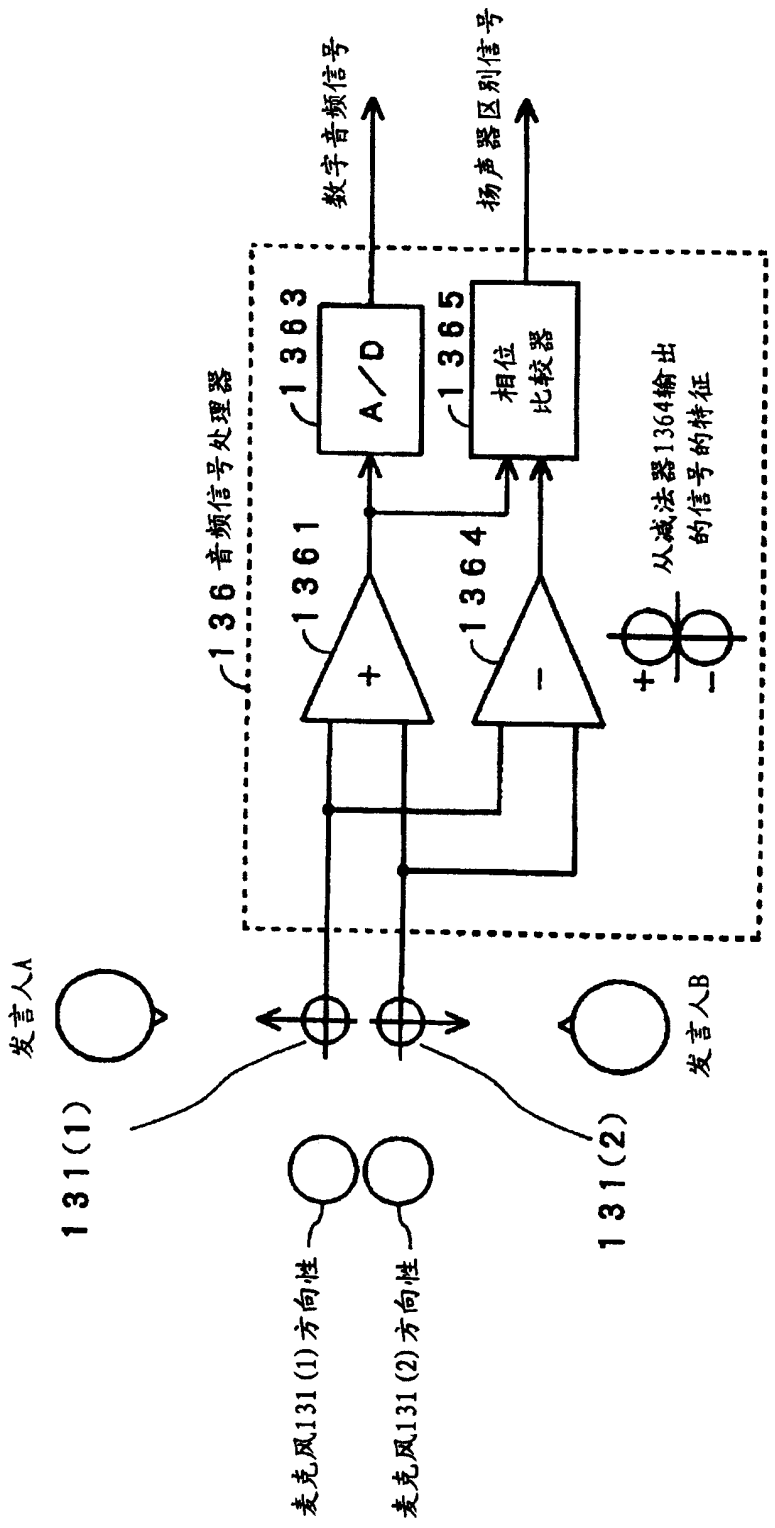

发言人A
131(1)
麦克风131(1)方向性
麦克风131(2)方向性
131(2)
发言人B
136音频信号处理器
1361
1363
A/D
1364
1365
相位
比较器
从减法器1364输出
的信号的特征
数字音频信号
扬声器区别信号

图 14

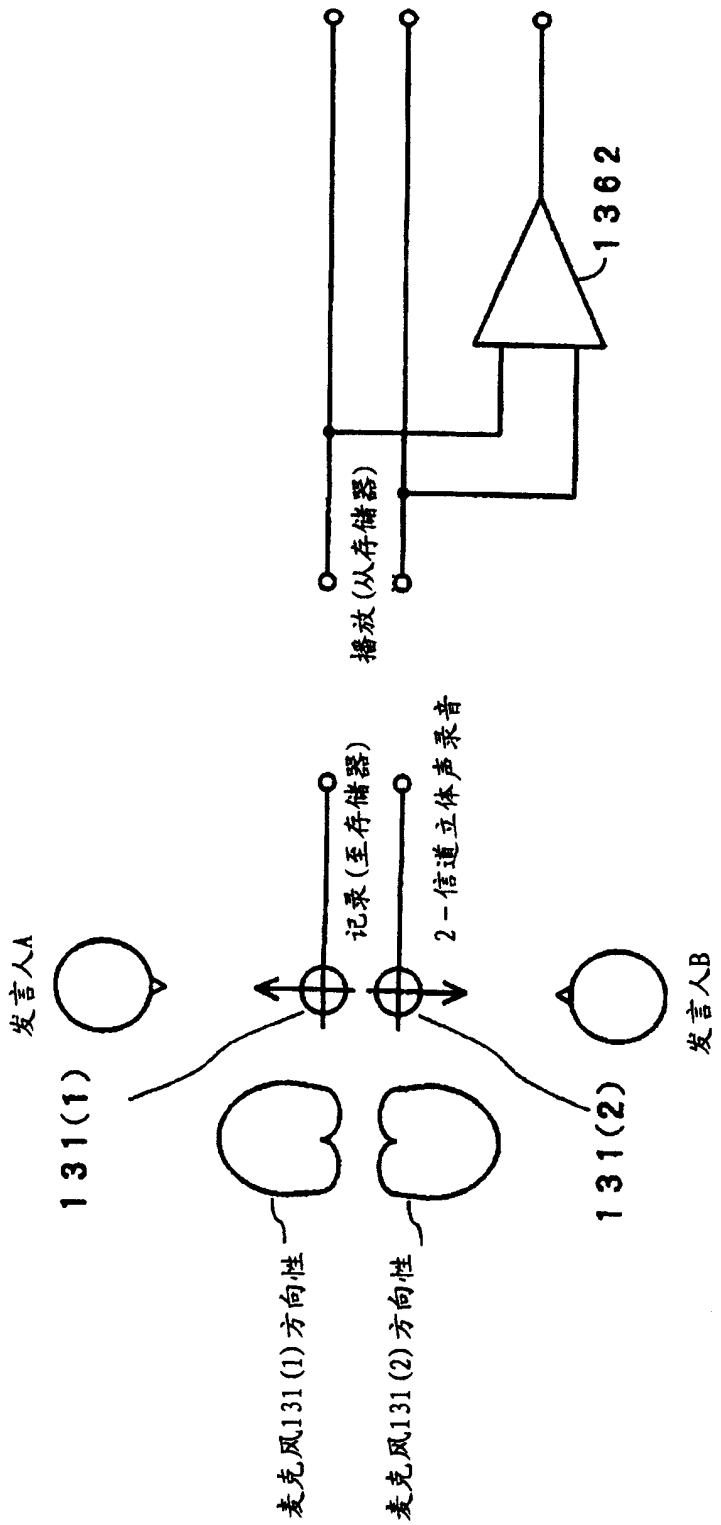

发言人A
131(1)
麦克风131(1)方向性
记录(至存储器)
2-信道立体声录音
麦克风131(2)方向性
131(2)
发言人B
播放(从存储器)
1362

图 15A

图 15B

发言人区别信号	麦克风	发言人识别符1	发言人识别符2	发言人识别符3	.....
1 (+)	麦克风1 (131(1))	发言人A1	发言人A2	发言人A3	.....
2 (-)	麦克风2 (131(2))	发言人B1	发言人B2	发言人B3	.....

图 16

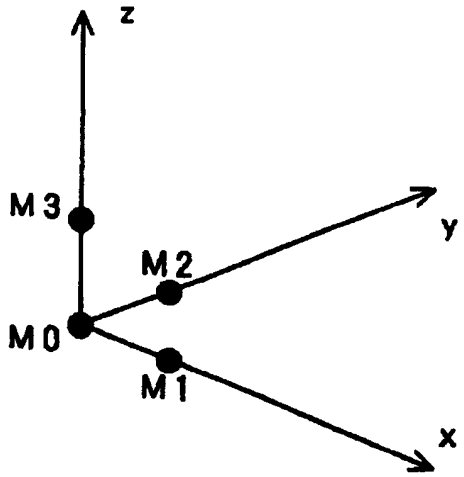


图 17A

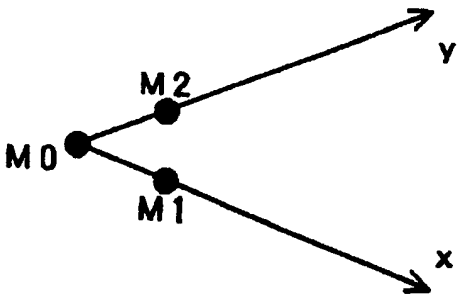


图 17B

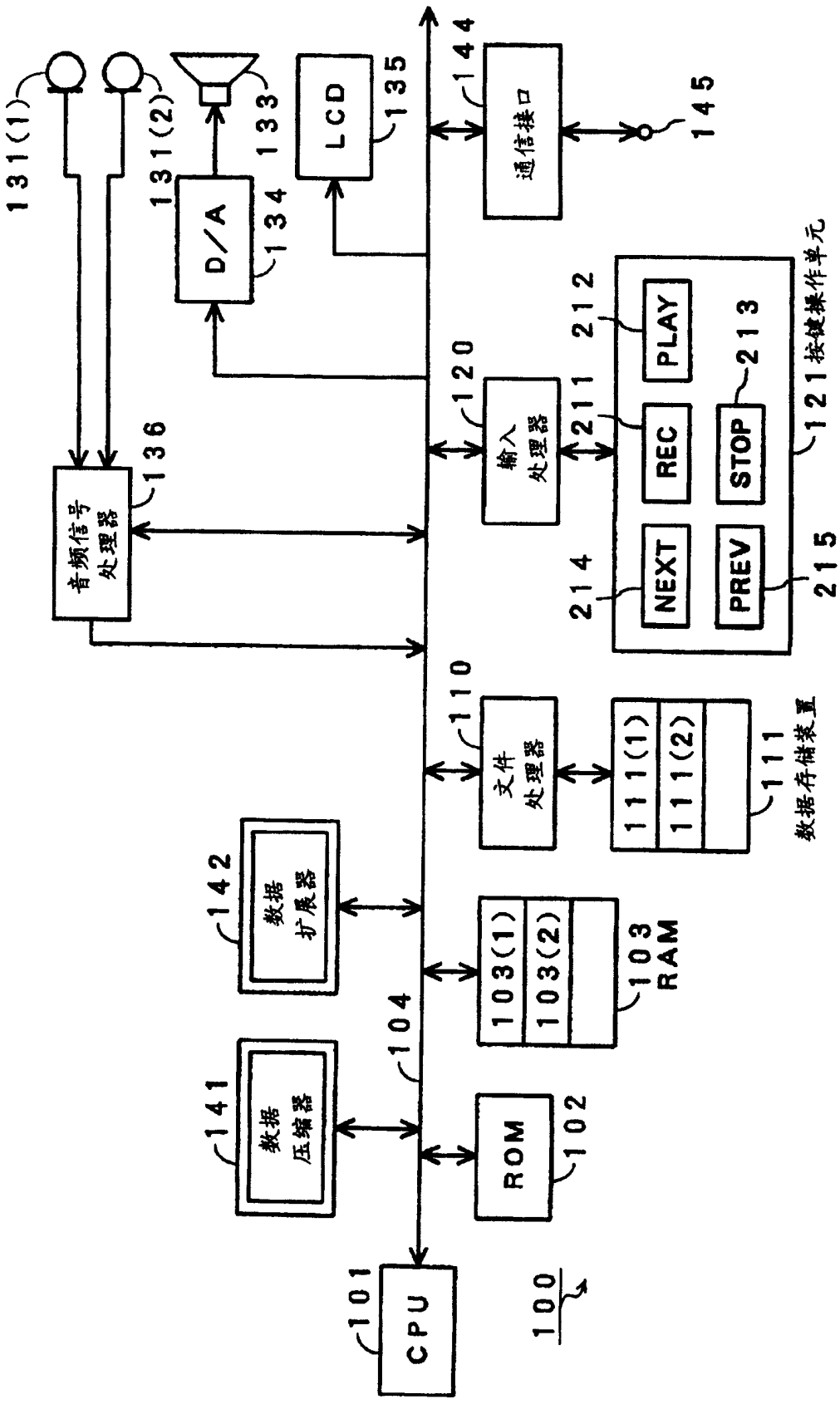

131(1)
131(2)
133
D/A
134
LCD
135
144
通信接口
145
音频信号处理器
136
120
输入处理器
211
212
PLAY
REC
213
STOP
121按键操作单元
214
NEXT
PREV
215
110
文件处理器
111(1)
111(2)
111
数据存储装置
142
数据扩展器
103(1)
103(2)
103
RAM
104
141
数据压缩器
ROM
102
CPU
101
100

图 18