



(12) 发明专利申请

(10) 申请公布号 CN 102378050 A

(43) 申请公布日 2012. 03. 14

(21) 申请号 201110204943. 2

H04N 21/472(2011. 01)

(22) 申请日 2011. 07. 13

(30) 优先权数据

1011751. 3 2010. 07. 13 GB

(71) 申请人 索尼欧洲有限公司

地址 英国萨里郡

(72) 发明人 胡·霍普金斯 蒂莫西·埃德蒙兹

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

代理人 宋鹤

(51) Int. Cl.

H04N 21/431(2011. 01)

H04N 21/435(2011. 01)

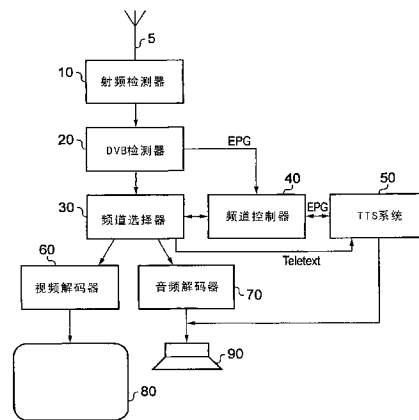
权利要求书 3 页 说明书 9 页 附图 6 页

(54) 发明名称

使用文本转语音转换的广播系统

(57) 摘要

本发明公开使用文本转语音转换的广播系统。广播信号接收机包括:文本数据接收器,接收广播文本数据以用于与用户界面相关地被显示给用户;文本转语音(TTS)转换器,将接收的文本数据转换成音频语音信号,TTS转换器可操作以检测用于转换的单词是否被包括在所存储的用于转换的单词的列表中,并且如果用于转换的单词被包括在所存储的用于转换的单词的列表中,则根据由所存储的列表定义的转换来转换该单词;以及如果用于转换的单词未被包括在所存储的用于转换的单词的列表中,则根据一组预定的转换规则来转换该单词;转换存储器,存储用于TTS转换器的转换的所述单词的列表;以及更新接收器,接收附加单词和相关的转换以用于存储在转换存储器中。



1. 一种广播信号接收机,包括:

文本数据接收器,所述文本数据接收器接收广播文本数据以用于与用户界面相关地被显示给用户;

文本转语音 TTS 转换器,所述 TTS 转换器将接收的文本数据转换成音频语音信号,所述 TTS 转换器可操作以检测用于转换的单词是否被包括在所存储的用于转换的单词的列表中,并且如果用于转换的单词被包括在所存储的用于转换的单词的列表中,则根据由所存储的列表定义的转换来转换该单词,并且如果用于转换的单词未被包括在所存储的用于转换的单词的列表中,则根据一组预定的转换规则来转换该单词;

转换存储器,所述转换存储器存储用于所述 TTS 转换器的转换的所述单词的列表;以及

更新接收器,所述更新接收器接收附加单词和相关联的转换以用于存储在所述转换存储器中。

2. 根据权利要求 1 所述的接收机,其中:

所述 TTS 转换器可操作以通过组装与单词或单词的一部分有关的语音成分来生成所述音频语音信号;并且

所述转换存储器为所述转换存储器中所存储的每个单词定义将用在该单词的转换中所要使用的相应系列的一个或多个语音成分。

3. 根据权利要求 1 或 2 所述的接收机,其中所述更新接收器可操作以通过经由因特网连接访问转换储存库来接收所述附加单词和相关联的转换。

4. 根据权利要求 1 或 2 所述的接收机,其中所述更新接收器可操作以接收所述附加单词和相关联的转换作为广播更新信号。

5. 根据前述任一权利要求所述的接收机,其中所述接收机是电视信号接收机,所述电视信号接收机可操作以接收用于输出给所述用户的包括视频和音频信号的电视信号。

6. 根据权利要求 5 所述的接收机,其中所述广播文本数据包括电子节目指南数据和/或图文数据。

7. 根据权利要求 6 所述的接收机,其中至少所述电子节目指南数据作为服务信息数据被广播。

8. 根据前述任一权利要求所述的接收机,包括远端操控器,所述远端操控器具有用于控制所述接收机的操作的一个或多个用户可操作控件。

9. 根据权利要求 8 所述的接收机,其中所述远端操控器具有音频输出设备,所述音频输出设备用于从由所述 TTS 转换器生成的音频信号生成听得见的输出。

10. 根据权利要求 8 或 9 所述的接收机,其中:

所述远端操控器包括用户控件,所述用户控件以供所述用户操纵来指示所述 TTS 转换器的不正确转换;并且

所述接收机可操作以响应于所述用户控件的操作来发送请求转换信息的提供的消息,所述消息指示转换问题并且指示在所述用户控件被操纵时被转换的文本。

11. 根据权利要求 10 所述的接收机,其中由所述消息指示的、在所述用户控件被操纵时被转换的文本包括以下内容中的一者或两者:在所述用户控件被操纵之前的时段期间被转换的预定数目的单词;以及在所述用户控件被操纵之前的预定时段期间被转换的单词。

12. 一种广播信号接收的方法,所述方法包括以下步骤:

接收广播文本数据以用于与用户界面相关地被显示给用户;

将接收的文本数据转换成音频语音信号,所述转换步骤包括检测用于转换的单词是否被包括在所存储的用于转换的单词的列表中,并且如果用于转换的单词被包括在所存储的用于转换的单词的列表中,则根据由所存储的列表定义的转换来转换该单词,并且如果用于转换的单词未被包括在所存储的用于转换的单词的列表中,则根据一组预定的转换规则来转换该单词;

存储用于转换的所述单词的列表;以及

接收附加单词和相关联的转换以用于存储在所述转换存储器中。

13. 一种广播信号发送系统,包括:

文本数据发送器,所述文本数据发送器发送广播文本数据以用于与接收机处的用户界面相关地被显示给用户;

消息接收器,所述消息接收器接收请求提供文本转语音(TTS)转换信息的信息,所述消息指示用户注意到的转换问题并且指示在所述用户注意到所述转换问题时被转换的文本;以及

更新发送器,所述更新发送器响应于接收的消息来发送单词和相关联的TTS转换以存储在接收机处。

14. 根据权利要求13所述的系统,包括:

检测器,所述检测器检测是否至少阈值数目的消息指示具有共同的一个或多个单词的文本,从而指示关于共同的单词存在潜在的转换问题;

并且其中,所述更新发送器可操作以发送用于所检测到的共同的单词以及相关联的TTS转换。

15. 一种广播信号发送方法,包括以下步骤:

发送广播文本数据以用于与接收机处的用户界面相关地被显示给用户;

接收请求文本转语音(TTS)转换信息的提供的消息,所述消息指示用户注意到的转换问题并且指示在所述用户注意到所述转换问题时被当时转换的文本;以及

响应于接收的消息来发送单词和相关联的TTS转换以用于存储在接收机处。

16. 一种广播信号发送系统,包括:

文本数据发送器,所述文本数据发送器发送广播文本数据以用于与接收机处的用户界面相关地被显示给用户;

文本库,所述文本库维护其文本转语音(TTS)转换信息不需要被发送给文本接收器的单词的列表;

比较器,所述比较器用于将所要发送的文本数据与所述文本库中所存储的单词相比较;以及

更新发送器,所述更新发送器响应于指示在所述文本库中未找到所要发送的单词的比较,来发送单词和相关联的TTS转换以用于存储在接收机处。

17. 一种广播信号发送方法,包括以下步骤:

发送广播文本数据以用于与接收机处的用户界面相关地被显示给用户;

在文本库中维护其文本转语音(TTS)转换信息不需要被发送给文本接收器的单词的

列表；

将所要发送的文本数据与被存储在所述文本库中的单词相比较；以及
响应于指示在所述文本库中未找到所要发送的单词的比较，发送单词和相关联的 TTS
转换以用于存储在接收机处。

18. 一种计算机软件，用于实现根据权利要求 12、15 和 17 中任一项所述的方法。

使用文本转语音转换的广播系统

技术领域

[0001] 本发明涉及一种使用文本转语音 (text-to-speech, TTS) 转换的广播系统。

背景技术

[0002] 本发明可适用于广播播送以及各种广播信号接收机, 诸如电视接收机或移动手持电话。为了说明本发明的背景技术, 以下仅在电视接收机的上下文中描述问题。

[0003] 已经提出了利用 TTS 转换来辅助目盲或视力部分残疾的用户的电视接收机。在 GB-A-2 405 018 和 GB-A-2 395 388 中公开了两个示例。在这些示例中, TTS 技术被用来以听得见的形式再现诸如电子节目指南 (EPG) 数据和图文 (teletext) 数据之类的的数据。

[0004] 此上下文中的 EPG 数据是指由广播商预先提供的以允许用户选择要观看和 / 或记录的节目的节目列表, 以及定义在特定频道上广播的当前的和接下来的节目的数据。图文数据是指由广播商提供的作为信息服务一部分的文本数据 (textual data)。图文数据的示例可以包括新闻文本、天气信息、电影列表等的页面。所有这些数据都具有如下共同的特征: 他们通常通过在电视屏上显示文本而可被用户获得, 并且实际而言, 他们具有无限的词典 (lexicon) (词汇表; 可用单词 (word) 集)。正是此无限词典的特征给 TTS 系统造成了困难。

[0005] TTS 技术或者依赖于回放预先记录的与将被 TTS 设备转换成语音的单词有关的语音, 或者是通过从称为音素的发音子元素建立完整单词的。音素是语声的基本单位, 并且基本上表示语言中能够表示含义的差别的最小语音单位。TTS 系统使用规则集来从将被转换成语音的单词的拼写生成音素的连续。在包含许多不规则发音的语言 (诸如英语) 中, 这些规则可能是复杂的, 特别是当类似的拼写具有不同的发音时 (例如, 英语单词 “through”、“though”、“cough”、“rough”、“plough”、“ought”、“borough”、“lough” 等中的字符集 “ough”, 他们中的这 4 个字符具有不同的发音)。但是, 尽管有这些复杂因素, 基于音素或基于预先记录的语音的 TTS 系统一般被布置来应对系统设计者预先知道的单词复杂性。然而, 实际上不可能预先预测在 EPG 数据、图文数据等中将出现什么单词。例如, 广播商可能引入缩写词 (例如, 针对 “sports” (体育) 频道的 “Spts”)。在另一示例中, 节目主持人或新闻人物的名称可能已经被普遍使用但是通常可能还未被包括在 TTS 系统的词典中——例如 “George Papandreou”、“Lembit Opik”、“Albus Dumbledore”。

[0006] Adobe® Captivate 4 TTS 系统提供了通过用户以 TTS 系统能够识别和发音的更语音的形式重写难以发音的单词来定制 TTS 发音的工具。但是在 EPG 或图文数据的 TTS 转换的上下文中, 该布置对于基于音素的 TTS 系统将几乎没什么用。首先, EPG 或图文数据是瞬变的; 用户可能只访问其一次, 因此用户将不会选择花时间设计和输入替换语音拼写来协助 TTS 系统。其次, 用户可能甚至不知道特定的单词——例如诸如 “Spts” 之类的缩写应当如何发音。第三, 在针对视力部分残疾或目盲的用户的系统中, 期望用户来重新键入替换语音拼写是过分的负担。

[0007] Adobe Captivate 4 的布置与基于预先记录的发音的 TTS 系统没有关系。

发明内容

[0008] 本发明提供一种广播信号接收机,包括:文本数据接收器,接收广播文本数据以用于与用户界面相关地被显示给用户;文本转语音(TTS)转换器,将接收的文本数据转换成音频语音信号,TTS转换器可操作用于检测用于转换的单词是否被包括在所存储的用于转换的单词的列表中,并且如果用于转换的单词被包括在所存储的用于转换的单词的列表中,则根据由所存储的列表定义的转换来转换该单词;以及如果用于转换的单词未被包括在所存储的用于转换的单词的列表中,则根据一组预定的转换规则来转换该单词;转换存储器,存储用于TTS转换器的转换的所述单词的列表;以及更新接收器,接收附加单词以及相关转换以用于存储在转换存储器中。

[0009] 本发明的各种进一步的各个方面和特征在所附权利要求中被限定。

[0010] 本发明有利地提供对由例如电视接收机中的TTS系统使用的词典数据的广播更新。

附图说明

[0011] 现在将仅通过举例的方式参考附图来描述本发明的实施例,在附图中:

[0012] 图1示意性地图示出电视接收机;

[0013] 图2示意性地图示出TTS系统;

[0014] 图3示意性地图示出TTS转换器;

[0015] 图4示意性地图示出转换词典或规则数据库;

[0016] 图5示意性地图示出具有网络连接的接收机;

[0017] 图6示意性地图示出具有远端操控器(remote commander)的接收机;

[0018] 图7示意性地图示出问题消息的生成;

[0019] 图8示意性地图示出广播商对问题消息的响应;

[0020] 图9示意性地图示出用于生成更新数据的另一技术。

具体实施方式

[0021] 图1示意性地图示出作为广播信号接收机的一个示例的电视接收机。电视接收机的许多操作是传统的,因此那些方面将仅以概要的形式来描述。图1中示出的示例是根据一个或多个数字视频广播标准(诸如DVB-T标准)进行操作的接收机。

[0022] 天线5可以是地面或卫星天线,接收广播数字电视信号。这些被传递给射频(RF)检测器10,RF检测器10将接收的RF信号向下解调到基带。注意,尽管该示例使用基于天线的接收,但是这里所描述的技术可等同地应用于其它广播递送系统,诸如有线或IPTV(因特网协议电视)系统。

[0023] 基带信号随后被传递给DVB检测器20。DVB检测器20是已知的DVB接收机中从基带广播信号导出所谓的数字视频传输流(TS)的那些部分和充当从基带广播信号导出图文数据和诸如电子节目指南(EPG)数据之类的服务信息(DVB-SI)的文本数据接收器的那些部分的示意性表示。传输流被传递给频道选择器30,频道选择器30在频道控制器40的控制下允许用户选择用于观看的特定频道。与所选频道相对应的音频和视频数据流被分别传递给音频解码器70(进而从音频解码器70传递给放大器和扬声器布置90)和视频解码

器 60 (进而从视频解码器 60 传递给显示屏 80) 。

[0024] 显示屏 80 以及放大器和扬声器 90 可以被设置为接收机的一部分, 如集成数字电视接收机的情形中那样, 或者可以是在分离的单元中, 如机顶盒 (STB) 包含的数字接收器耦接到用于显示接收的信号的情况中那样。

[0025] 由 DVB 检测器 20 导出的 EPG 数据被 DVB 检测器缓冲, 并且在需要时, 被传递给频道控制器 40。响应于适当的用户命令 (例如使用远端操控器, 图 1 中未示出), EPG 数据被显示在显示屏 80 上, 使得用户能够操作进一步的控件来选择可用于观看的频道中的某一频道。

[0026] 另一类型的 EPG 数据是所谓的“当前和接下来”(now and next) 数据, 其提供对频道上可观看的当前节目的名称 (和简要细节) 以及此频道上接下来的节目的名称 (和简要细节) 的频繁更新的指示。

[0027] 用户可以选择的一个选项是图文信息的显示。图文是提供用于显示的文本和简单图形的低比特率服务 (相比于视频服务的比特率而言)。该术语一般是指与广播音频和 / 或视频系统相关联的广播文本服务, 并且包括按照诸如 DVB 标准之类的模拟或数字广播标准定义的图文, 由包括 Java ® 应用等的多媒体和超媒体信息编码专家组 (MHEG) 或多媒体家用平台 (MHP) 系统定义的文本和交互式服务, 以及用于向广播接收机递送文本和 / 或交互式服务的其它这样的协议。图文服务可以是可选择的, 就像他们是独立的单独频道一样, 而获取由广播商提供的图文服务的另一途径是在观看由该广播商提供的视频频道的同时, 操作特定用户控件。当图文服务被用户选择时, 频道选择器将图文数据路由到视频解码器 60 以被译成可观看的信息页面。

[0028] 因此, 文本数据接收器被布置以便接收广播文本数据以用于与用户界面相关地显示给用户。

[0029] 还提供文本转语音 (TTS) 系统 50。文本转语音 (TTS) 系统 50 对被显示在显示屏 80 上的某些文本类别起作用, 并且将所显示 (或所接收) 的文本数据转换成音频语音信号, 以便由放大器和扬声器 90 输出。在本示例中, TTS 系统对 EPG 数据 (包括当前和接下来数据) 和图文数据进行操作。然而, 在其它实施例中, TTS 系统可使用已知的字符识别并且可对作为接收的视频和 / 或数据服务的一部分而被显示的任意文本进行操作。

[0030] 在这里所讨论的示例中, TTS 操作被应用于显示屏上正显示的文本。然而, TTS 操作可以应用于其他文本, 诸如未被显示的文本。

[0031] 为了将 TTS 技术应用于 EPG 和图文数据, TTS 系统从频道控制器 40 接收当前显示的 EPG 数据以及用户作出的任意选择的文本 (诸如所选频道上特定时间的特定节目的文本描述) 来作为文本数据。TTS 系统从频道选择器 30 接收任意当前显示的图文数据来作为文本数据。TTS 系统进行操作来 (例如, 至少关于英语文本) 从显示的文本的左上开始并且要么按照正常的读取次序 (在图文数据的情况中) 要么按照用户当前选择了文本的哪个部分的次序 (在 EPG 数据的情况中), 将这些类型的显示文本转换成语音信号。在后一情况中, 通常用户操作可移动光标来对 EPG 数据进行导航, 可能将光标从用于一个频道的列表移动到用于另一个频道的列表。TTS 操作可以根据特定电视接收机上使用的用户界面以常规方式设置。例如, 如果用户使用“上 / 下”光标控件来在频道之间移动并且使用“左 / 右”光标控件来改变 EPG 列表的信息被显示的时间段, 则在光标移动中在预定停顿 (例如 0.8 秒)

之后, TTS 系统可以开始对所显示的 EPG 数据中的针对当前选择的频道和当前选择的时间段的时间和节目名称进行转换。

[0032] 现在将描述 TTS 系统 50。图 2 到图 4 是图示出 TTS 系统 50 的操作的示意性示图。TTS 系统 50 包括 TTS 转换器 100、转换字典 110、规则数据库 120 和数字转音频转换器 (DAC) 130。

[0033] TTS 系统将常规语言 (非语音的表示) 转换成语音。语音可以以不同的方式被合成。在具有有限词典或词汇表的系统 (诸如汽车卫星导航系统) 中, 全部单词或甚至短语可以被预先记录, 这针对使用的有限集合的单词和短语提供高质量的输出。在具有更宽泛的词典的系统中, 合成的语音可以通过连接诸如音素之类的语音成分来创建。用于 TTS 系统的另一中替代方式是模仿人类声道的运作和其它语音特征。将参考图 2 至图 4 讨论的示例是基于音素的 TTS 系统。

[0034] 如图 2 至图 3 中所示的基本语音合成处理以一般传统方式操作, 因此这里将仅以概要形式描述。在第一阶段 102 (图 3), TTS 系统尝试将进入的文本转换成可以被稍后的阶段正确处理的单词。该处理有时称为文本标准化、预处理或字元化 (tokenisation)。例如, 在进入文本的流中单独出现的数字“5”将被转换成“five”(五), 而相邻符号的组“523”将被转换成“five hundred and twenty three”(五百二十三)。符号“+”将被转换成单词“plus”(加)。所有这些转换都是基于查找表来执行的, 查找表 (对于图 3 的目的) 被认为是规则数据库 120 的一部分。不能被解析为单词的文本可能被转换成首字母集合: 例如“Spts”将被转换成 4 个连续的首字母“SPTS”。

[0035] 预处理阶段 102 的输出被传递给语言分析器 104, 语言分析器 104 向每个预处理后的单词指派音标 (phonetic transcriptions)。如上所述, 音素单个的语音成分, 其被认为是能够指示含义的差别的最小成分。语言分析器 104 利用相关联的时态 (phasing)、语调和持续时间值来为每个预处理后的单词选择一组或一系列一个或多个音素或其它语音成分。

[0036] 当然, 对于特别常用的单词, 或者对于广告商倡议的单词, 该整个单词的数字化版本可以作为单个成分被存储以供语言分析器选择 (而不必从单个的音素来构建单词)。这里的示例可以是广播商或频道的名称或电视制作者的名称。

[0037] 语言分析器使用两种一般方法的组合来指派音素。第一种是基于所存储的列表或词典的方法, 其中, 大的词典 (被实现为转换字典 110, 并且实际上提供存储的用于转换的单词的列表) 实际上包含将单词映射到音素集合的查找表。语言分析器在词典中查找每个单词并且获取正确的音素集合。该方法是快速和准确的, 如果单词在该词典中被找到的话; 否则, 该方法会失败。另一方法是基于规则的方法, 其中 (存储在规则数据库 120 中的) 一组预定发音规则被应用于单词, 以基于它们的拼写并且某一程度上基于它们的上下文, 即周围的单词, 来确定它们的发音。基于规则的方法至少可以尝试处理任意单词, 但是当该系统尝试处理更多单词时, 规则本身变得越来越复杂。因此, 许多 TTS 系统 (包括如本实施例所示出的 TTS 系统) 使用这些方法的组合。简而言之, 这可能意味着, 如果在转换词典中所存储的用于转换的单词的列表中找到单词, 则使用基于词典的方法, 否则, 使用基于规则的方法, 但是这无法应付同形异音字, 即基于上下文被不同地发音的拼写。英语同形异音字的简单示例包括单词“close”、“rebel”、“moped”和“desert”。因此, 在本实施例中, 向具有该特性的单词被提供基于规则的辅助方法, 以根据单词的上下文, 即, 特定单词周围

的单词,来选择两个以上基于词典的发音中的一个发音。然而,如果语言分析器没有在词典中找到该单词,则仅使用基于规则的方法来尽力尝试发音。

[0038] 所选音素随后被传递给波形生成器 106,波形生成器 106 根据由语言分析器 104 设置的时态、语调和持续时间值,将语音成分或音素连接或组装成与那个单词相关的输出数字化波形。音素一般被布置以便一个到下一个地延续,即在单个的单词中间没有停顿地继续。波形由 DAC 130 转换成模拟形式,以由例如放大器和扬声器 90 输出。

[0039] 简而言之,因此,TTS 转换系统 50 在预处理和语言分析两个阶段期间都利用转换字典 110(充当转换存储器)中所存储的信息以及规则数据库 120 中所存储的信息。

[0040] 图 4 示意性地图示出转换字典 110 或规则数据库 120,展示与设备的存储数据的更新有关的特征。示意而言,转换词典和规则数据库可被认为具有用于初始数据 150 的存储器装置以及用于接收和存储对初始数据的更新的更新存储器 140。以下将描述更新被接收的方式。但是,基本而言,当转换词典或规则数据库接收到(要转换的单词的形式的)查询时,首先对照初始数据来测试该查询,随后对照更新存储器中所存储的数据来测试该查询。如果初始数据提供了任意响应,则此响应可以用关于更新数据而提供的响应重写。

[0041] 当然,图 2 和图 4 中所示出的布置是示意性的。转换字典 110 和规则数据库 120 不需要是分离的存储器或分离的数据储存库,而是可以实现为返回与被查询的单词有关的规则和转换的单个数据储存库。类似地,初始数据和更新数据不需要被分别地存储;更新数据可以被合并到初始数据中以形成组合的数据结构。在更新数据涉及初始数据中没有包括的单词的情况下,更新数据将简单地作为附加数据。在更新数据涉及初始数据中包括的单词的情况下,更新数据可以被布置来补充或替换对应的初始数据。

[0042] 更新数据可以作为广播数据从转换储存库或通过网络(因特网)连接来接收。在任一情况中,更新数据的发布可以是只通过数据提供者(例如广播商)的决定或响应于来自电视接收机或其用户的自动或手动请求而进行的。例如,更新可以作为广播数据通过使用由 DVB 系统软件更新标准 ETSI TS 102 006 定义的技术(例如参见 <http://broadcasting.ru/pdf-standard-specifications/multiplexing/dvb-ssu/ts102006.v1.3.1.pdf>)来处理。

[0043] 经由网络连接来提供更新数据实际上可以是间接的,例如通过广告商提供作为单独的操作可从其下载更新数据的因特网链接(例如统一资源定位符或 URI)。在例如广播信号接收机没有网络或因特网浏览器能力等的情况中,用户可以使用个人计算机(未示出)将更新数据下载到数据载体(诸如具有 USB 接口(未示出)的存储器)并将数据载体插入广播信号接收机的对应的接口(未示出)。此对应接口可以是广播接收机的 USB 接口或串行端口。

[0044] 图 5 示意性地图示出与结合图 1 描述的接收机类似的电视接收机 200。接收机 200 连接到显示屏 80。除了已经描述的特征以外,电视接收机 200 还包括连接到诸如因特网连接之类的网络连接 230 的检测器 210 和接口 220。

[0045] 检测器 210 具体地利用 TTS 转换器 100、转换字典 110 和规则数据库 120 之间的交互与 TTS 系统相接口。检测器 210 检测用于转换的单词未被包括在转换词典中的情况,并且经由网络连接 230 向广播商发送消息来请求发布与那个单词有关的更新数据,或者访问远程转换储存库(未示出)来搜索与那个单词有关的转换数据,检测器随后可以将此转

换数据下载作为更新数据。因此,在此上下文中,检测器充当更新接收器。

[0046] 远程转换储存库可以是例如由广播商、电视接收机制造商或由视觉残疾慈善团体操纵的网站。

[0047] 图 6 示意性地图示出另一实施例,其中远端操控器 300 与电视接收机 200' 无线地交互。在图 6 中,远端操控器被画得比电视接收机 200' 大,但是将明白这仅仅是示意性的示图,并且事实上,远端操控器很有可能是手持设备。无线交互可以是经由接口 220' (具有图 5 的接口 220 的功能,加上用于与远端操控器 300 交互的无线接口) 和远端操控器中的对应的接口设备(未示出)的。无线交互可以通过已知的红外、无线以太网、蓝牙®或 ZigBee®协议的。

[0048] 远端操控器包括:音频输出设备,诸如扬声器 310(具有对应的放大器,未示出);一个或多个用户可操作控件(用户控制按钮 320),用于操作传统用户遥控功能,诸如频道改变或接收机的其他操作;以及问题按钮 330。

[0049] 扬声器 310 被布置为经由远端操控器 300 和电视接收机 200' 之间的无线连接来接收 TTS 系统 50 的输出。即,所生成的语音由扬声器 310 而不是放大器和扬声器 90 再现。这具有如下优点:在混合观看环境(其中一个用户需要使用 TTS 系统 50 而其他用户可以不用 TTS 系统 50)中,TTS 系统 50 的语音输出不被施加于所有用户,而是仅被导向到需要该语音输出的用户。

[0050] 当用户听到未被 TTS 系统 50 成功地或正确地转换成语音的单词时,用户按压问题按钮 330。这种单词可以是用户能够识别但是被不正确地发音的单词。或者这种单词可以是用户仅仅由于该单词被给出无意义的发音而无法识别的单词。按压问题按钮使得远端操控器指示电视接收机中的消息生成器 240(例如,向广播商)发送用于请求更新数据的消息。消息生成器 240 组织消息(该消息可以指示转换问题,并且可以指示在问题按钮被操纵时所转换的文本),并且经由接口 220' 和网络连接 230 将其发送给广播商。

[0051] 但是这里存在困难,图 7 示出对此困难的解决方案,图 7 是与问题按钮 330 有关的操作的示意性表示。

[0052] 所述困难在于不同的用户具有不同的反应时间,并且所有用户都具有非零的反应时间。这意味着当前即在问题按钮 330 被按压时被转换并被语音化的单词几乎肯定不是触发对问题按钮的按压的单词。

[0053] 参考图 7,在该实施例中,TTS 系统 50 维护最新转换的单词的滚动(rolling)缓冲器 400。该缓冲器可以是覆盖某一预定时间段的缓存器,例如在最后十秒所有单词都被转换,或该缓冲器可以基于预定数目的单词,例如 30 个最新转换的单词,或者甚至基于与最新转换的单词有关的字符或字母的数目,例如最新转换的 200 个字符。当前正被转换的单词由框 410 示出。

[0054] 当问题按钮 330 被用户按压时,远端操控器提供检测按钮操作和向消息生成器 240 发布指令的功能 420。消息生成器随后参考缓冲器 400 来准备消息(430),并且随后经由接口 220' 发送消息(440)(图 6)。

[0055] 消息生成器在问题按钮被按压时参考缓冲器 400。消息生成器从缓冲器 400 中选择文本以用于包括在所述消息中。此文本可以以各种方式来选择:

[0056] (a) 消息生成器可以选择缓冲器 400 中的全部文本;或者

[0057] (b) 消息生成器可以基于用户的反应将不足以快到指示在最新转换的 n 个单词中的问题,来选择缓冲器 400 中除了最新转换的 n 个单词以外的任意单词。值 n 可以例如是 5。在图 7 中示出了值 n 的示意性表示;或者

[0058] (c) 以与 (b) 类似的方式,消息生成器可以使用缓冲器中除了与最近的转换时间段 t 相对应的单词以外的所有单词。 t 的值可以例如是 0.1 秒,并且 t 的值在图 7 中示意性地被示出;或者

[0059] (d) 消息生成器可以(从缓冲器 400 中的单词中)选择利用基于规则数据库的基于规则的转换而不是使用转换词典的基于词典的转换的最新转换的单词。为了实现此,缓冲器 400 可以例如以每个单词单个标记比特的形式来存储与每个单词相关联的元数据,标记比特指示该单词是否是使用转换词典而被转换的。可替换地,接收机可以仅在信息被需要时(即,响应于问题按钮的按压)通过从最新转换的单词开始并且按时间往回进行来检查缓冲器 400 中所存储的每个单词是否在转换词典中被找到来得出这样的信息。在这些任意情形中,在问题按钮被按压之前的阈值时间(例如,0.1 秒)内被转换的单词可以被排除在针对只使用规则数据库的最新转换的单词的搜索以外。如前,这考虑到了用户的反应时间——用户通常不能在问题单词被语音化之后的阈值时间之前按压问题按钮。

[0060] 在情况 (b) 和 (c) 的任一者中,消息中所包括的单词表示在预定时间段期间转换的单词,或者在按钮被按压之前的预定数目的单词。然而,此单词集合不紧在按钮被按压之前。

[0061] 图 8 和图 9 示意性地示出广播商的以上述形式促成更新数据的准备的操作。

[0062] 图 8 涉及上述情形,其中电视接收机具有如下功能:允许自动和/或手动触发指示转换问题的消息被发送给广播商。图 8 中示出的步骤是例如由在程序控制下进行操作计算机自动执行的。

[0063] 在步骤 500,广播商(经由消息接收机,未示出)接收指示用户注意的转换问题并且请求提供 TTS 转换信息的信息,该消息指示在用户注意到转换问题时被转换的文本。如以上所讨论的,该问题可能与单个单词有关(在自动生成的消息的情况中),或者可替换地,在手动生成消息的情况中,关于一组单词中的哪个单词具有转换问题完全可能具有某些不确定性。

[0064] 在任一情形中,在步骤 510,广播商(使用检测器,未示出)来将当前消息中所包含的文本与被存储在消息库 520 中的、先前接收的消息中所包含的文本相比较。该步骤具有多种益处:

[0065] (a) 如果广播商具有总是在问题单词一但通知之后就提供更新的策略,则在消息库 520 中存在该单词将指示问题已经被处理。进一步的动作不需要并且处理可以跳到步骤 560。如果该单词不在消息库中,则控制进行到步骤 530。

[0066] (b) 广播商可以推迟提供更新,直到已经超过至少阈值数目(例如 20)的问题通知为止。在此情况中,步骤 510 处利用消息库 520 的比较具有检测该单词已被标记为问题的次数的功能。如果此次数少于阈值,则不需要采取动作并且处理跳到步骤 560。如果此次数大于阈值 +1(+1 是确保阈值被超过的可选安全裕度),则广告商可以假定该问题已经被解决,并且也不需要动作。另一方面,如果此次数等于阈值或阈值 +1,则控制可以进行到步骤 530。

[0067] (c) 如果带有多个单词的手动生成的消息被接收, 其中一个单词可以表示问题, 则消息库 520 中所存储的消息的相关性可以指示群组中的问题单词, 特别是如果该问题单词在各种不同的上下文中出现的话。如果在步骤 510 发现单词是当前消息和至少 (比如) 5 个之前的消息所共有的, 则假定关于共有的这 (一个或多个) 单词存在转换问题, 并且控制可以进行到步骤 530。否则, 控制进行到步骤 560。

[0068] 进行到步骤 530 的控制因此假定 (一个或多个) 问题单词已经被识别并且需要被处理。在步骤 530, 广播商从更新提供者 540 预订 (order) 更新。更新的生成仅仅是图 8 中可能需要被手动完成的部分, 但是广播商可以自动访问数字发音信息的储存库来生成更新。更新提供者可以是广播商的雇员、视觉残疾慈善团体等等。

[0069] 在步骤 550, 更新被更新发送器 (未示出) 广播, 该更新发送器响应于接收的消息来发送单词和相关联的 TTS 转换以用于存储在接收机处。以这种方式, 一个用户 (或相对少量数目的用户) 指示出问题的事实致使向所有用户提供更新。这在常常具有一周以上的寿命的 EPG 数据的示例中特别有利, 所以, 如果 TTS 发音问题响应于第一通知或第一少量通知被立即解决了, 则可能大多数用户将从他们第一次访问该 EPG 数据起就将只听到正确的发音。

[0070] 最后, 在步骤 560, 当前消息 (或至少其问题文本部分) 被存储在消息库 520 中, 并且控制返回进行到步骤 500, 以等待下一消息的接收。

[0071] 图 9 示意性地图示出由广播商执行的用于抢先检测潜在的问题单词并向用户发布更新的一组操作。

[0072] 在步骤 600, 广播商准备用于广播的文本 (诸如 EPG 文本或图文信息)。但是在文本被实际广播之前, 步骤 610 至 660 被执行。

[0073] 在步骤 610, 所准备的文本中所使用的单词被与提供所有先前使用的单词的词典或列表 620 的文本库相比较。即, 广播商将词典 620 维护作为在先前广播的 EPG 和图文信息中已经出现的所有单词的有序列表 (例如字母列表)。该词典对于每个单词只需要一个条目——重要的因素是单词之前是否已经被使用过, 而不是其已经被使用多少次。

[0074] 作为维护广播商已经使用过的所有单词的列表的可替换选择, 广播商可以改为维护最新更新的转换词典中出现的在此范围内被提供给用户的所有单词的列表。

[0075] 如果比较器 (未示出) 检测到在词典 620 中没有找到当前准备的文本中的单词, 则在步骤 630, 广播商从与上述更新提供者 540 类似的更新提供者 640 预订更新信息。更新包括用于存储在接收机处的单词和相关联的 TTS 转换。

[0076] 在步骤 650, 广播商使用更新发送器 (未示出) 来广播更新信息并且还将该单词添加到词典 620。

[0077] 最后, 一旦更新信息已经先被广播, 广播商就在步骤 660 使用文本数据发送器 (未示出) 来广播所准备的文本。一般, 文本数据发送器广播文本数据以用于与接收机处的用户界面相关地被显示给用户。

[0078] 广播商可以在发布更新之前应用出现阈值数。这将要求广播商维护用于更新的单词的临时列表 (未示出)。单词不是被存储在词典 620 中, 并且更新信息不在步骤 550 被广播, 直到该单词已经在 EPG 文本或图文中出现了至少阈值数目的次数为止。该阈值例如可以是 3。当临时列表中的单词已经出现了至少阈值数目的次数, 则更新被广播 550, 单词被

存储在词典 620 中并且单词被从临时列表中删除（未示出的步骤）。

[0079] 如之前提及的，更新包括用于转换词典和 / 或规则数据库的条目。更新实际上（作为广播更新信号）在与所使用的特定广播标准相关联的专用字段或用户数据字段中被广播，并且由充当更新接收器的 DVB 检测器接收。例如，更新作为更新信息的回旋馈送（rotating feed）的一部分被广播多次，使得新准备的更新可以在轮播中被添加到所有先前的更新。更新可以被布置使得更新在轮播中再次出现的频率与更新的新度有关，使得较新的更新比较旧的更新被更频繁地重新广播。

[0080] 文本数据发送器是广播发送机系统的传统部分。更新发送机可以是广播发送机系统的传统部分或者可以被实现为如上所述的基于因特网的服务器。结合图 8 和图 9 讨论的其余项目（例如，文本库、比较器等）可以由在软件控制下操作的通用计算机执行。

[0081] 已经结合 DVB 系统讨论了具体实施例，但是这些技术也适用于根据由（例如）使用文本服务信息的 ATSC（高级电视系统委员会）、ARIB（无线行业企业协会）定义的标准操作的广播系统，或适用于针对带有相关联的数字数据（例如图文数据）的模拟广播的 PAL、NTSC 或有关标准。类似地，该技术适用于除了电视广播系统以外的广播系统，例如无线电广播系统，诸如根据 DAB（数字音频广播）标准的数字无线电系统（其中定义当前和将来的节目的辅助文本与音频信号一起被广播），以及诸如 FM 广播之类的模拟无线电系统（相关联的文本经由无线电数据系统（RDS）布置被发送）。这些技术还适用于仅文本的广播系统，例如使用广播文本信息来向用户传递状况或其它广播消息的无线电寻呼机、报警或移动电话系统。

[0082] 这些技术还适用于字幕系统。首先，TTS 技术（主要针对视觉受损但是听觉完好的用户）不可直接适用于字幕布置（主要针对视觉完好但是听觉受损的用户）。然而，存在这样的情形，其中本技术实际上在字幕系统中非常有用。例如，在双语情形中，可以广播节目，其仅带有单种语言的音频（例如英语语言），并且带有双语字幕（例如针对听觉受损用户的英语字幕以及针对讲威尔士语的用户的威尔士语字幕，而不论他们是否听觉受损）。如上所述的 TTS 系统可以用来输出威尔士语的音频来模拟威尔士语音频流。

[0083] 这样的字幕 / TTS 特征因此不仅可以对视觉受损的用户有用，而且在外语电影被广播时也有用。图文或类似的字幕（他们一般作为编码的文本字符被广播）可以被递送给 TTS 系统。DVB 或类似的字幕一般以位图的形式提供，因此在输入 TTS 系统之前需要进一步的处理（诸如已知的字符识别（OCR）技术）。

[0084] 上述实施例可以用硬件、软件、可编程硬件（诸如 ASIC、FPGA 等）、软件控制的计算机或这些的组合来实现。在涉及软件的实施例的情况中，将明白软件本身以及诸如携带这样的软件的存储介质之类的计算机程序产品被认为是本发明的实施例。

[0085] 上述技术适用于除了电视系统以外的广播系统和接收机，例如：数字无线电广播和接收机（其中 TTS 技术可以用来语音化描述节目的元数据）以及移动电话系统（其中用户菜单甚至文本消息可以由 TTS 系统按照与以上所述的相同方式来处理）。

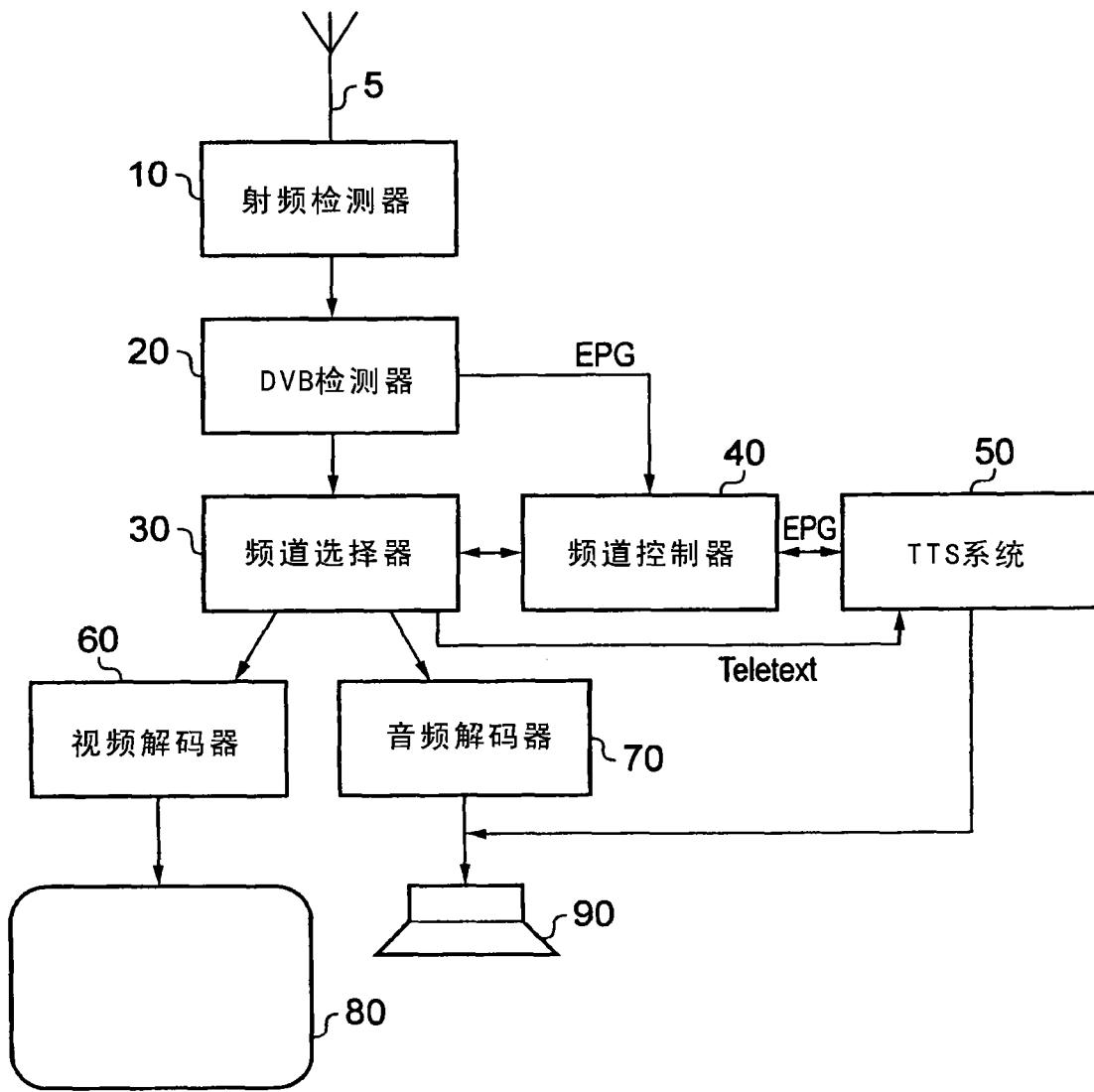


图 1

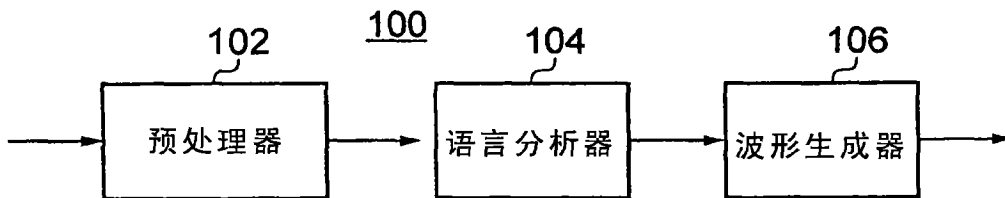


图 3

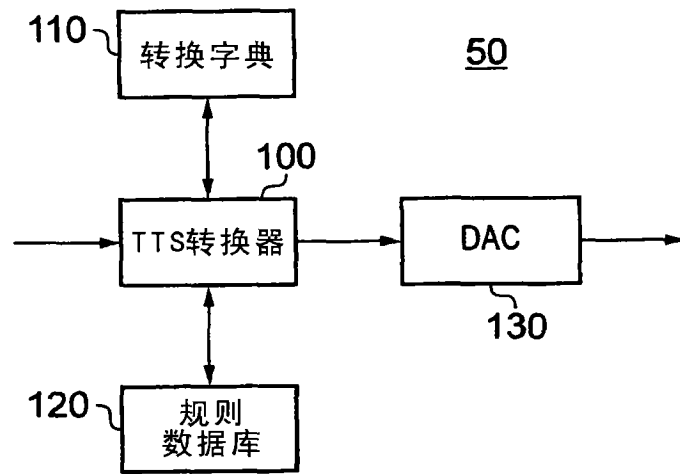


图 2

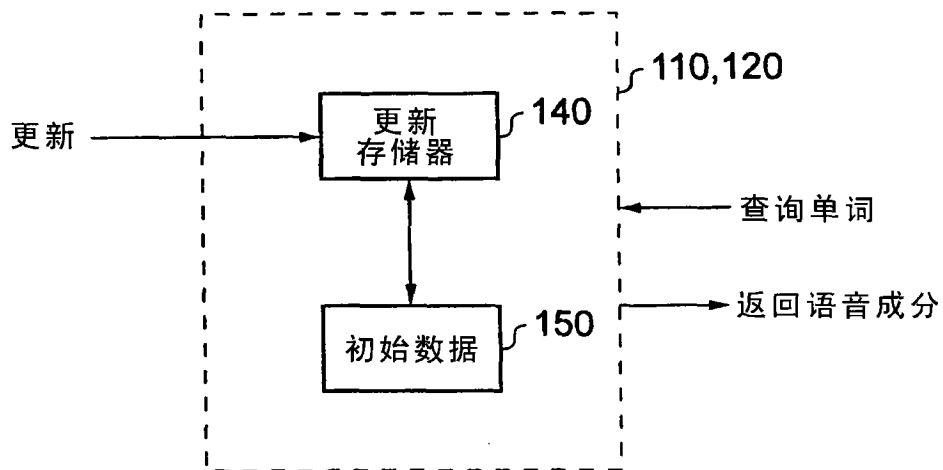


图 4

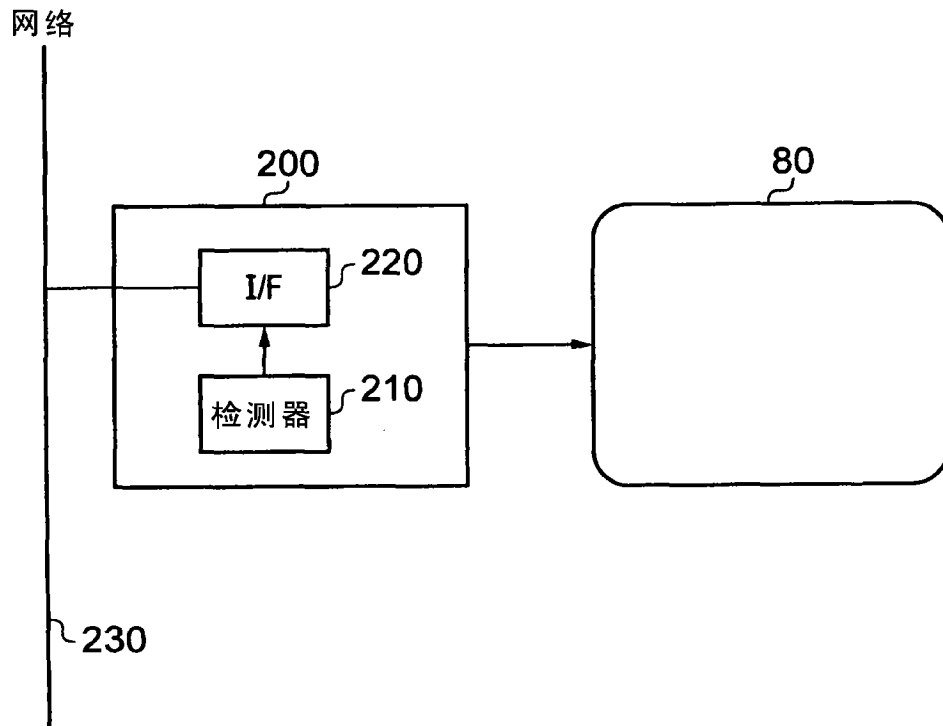


图 5

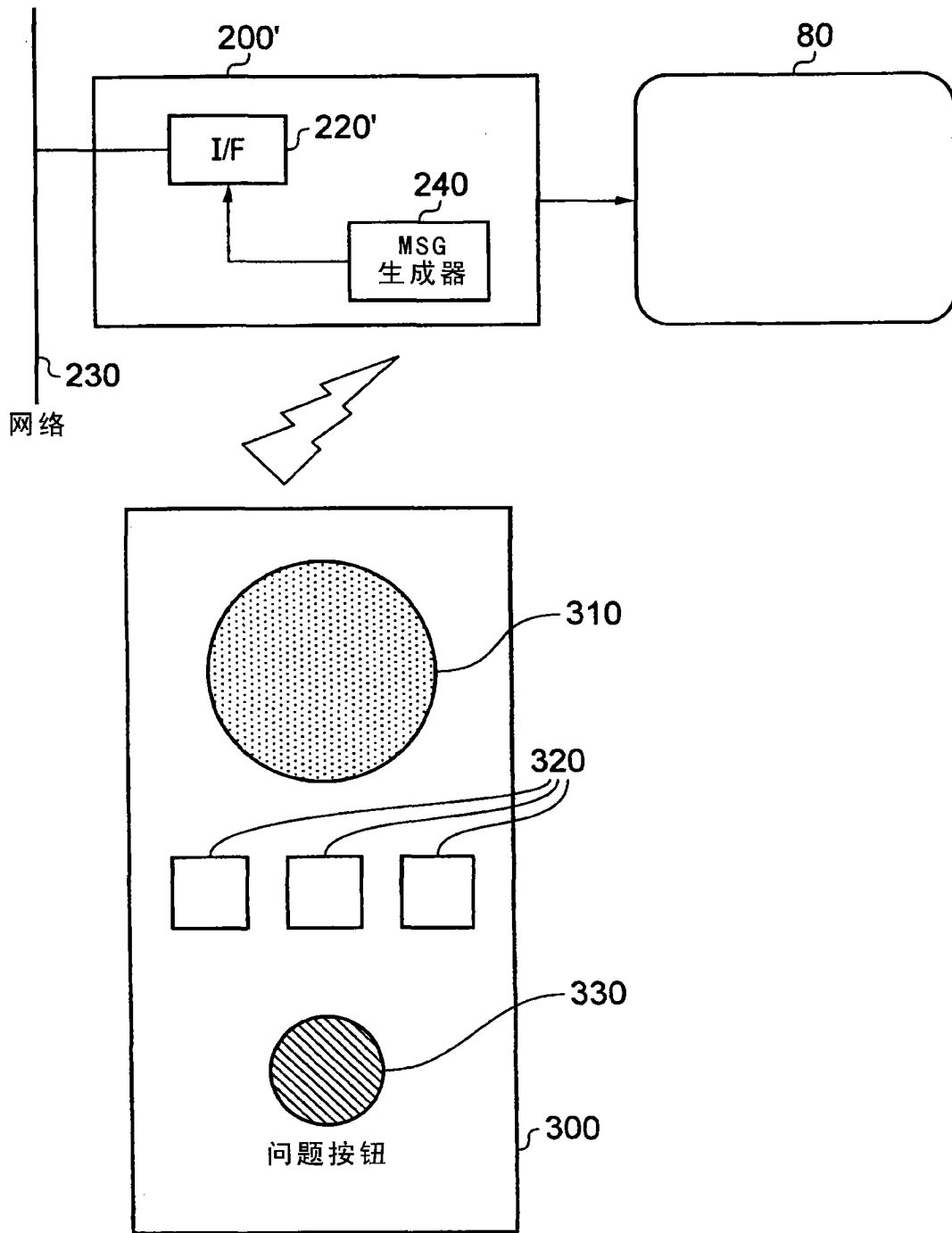


图 6

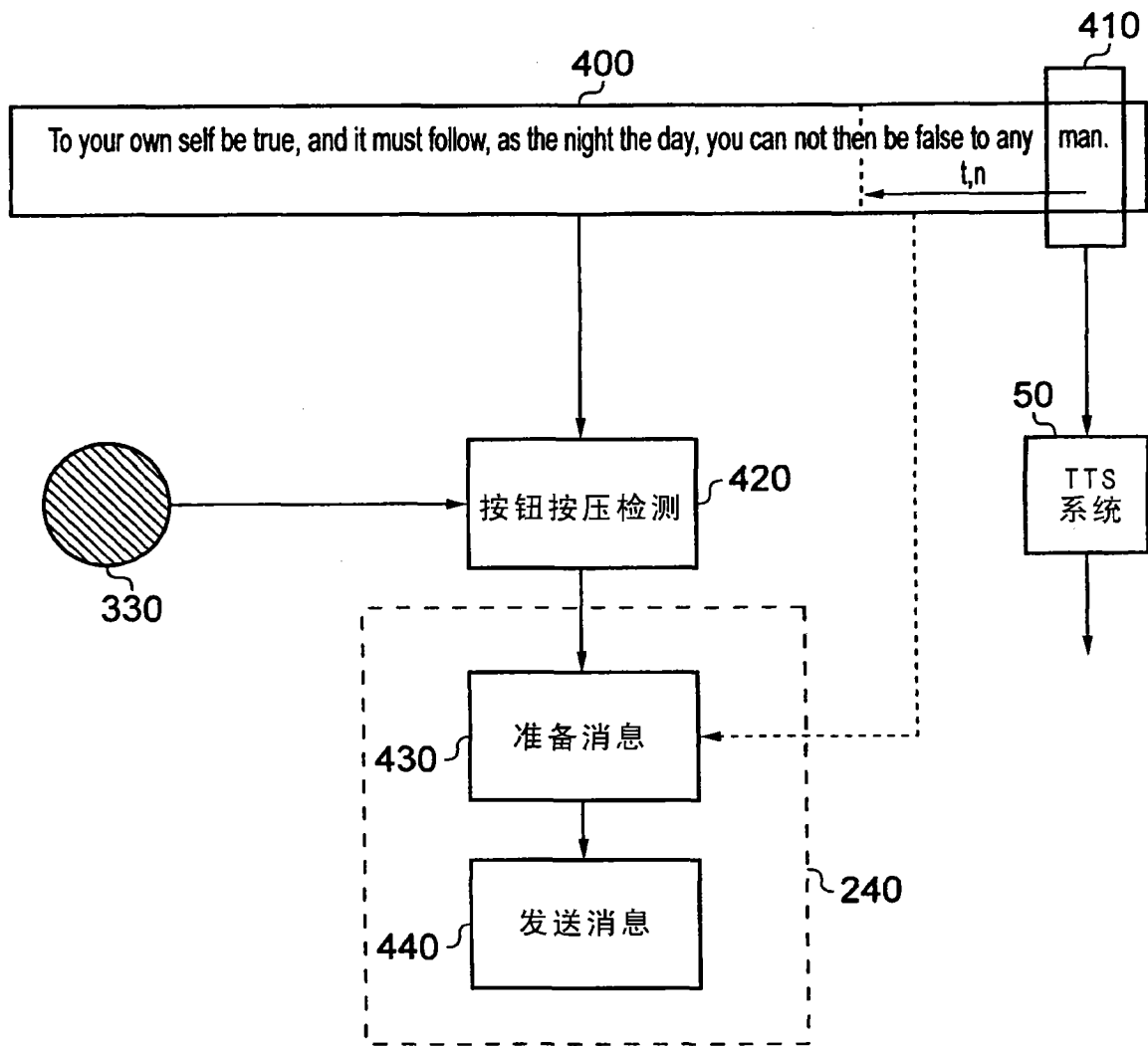


图 7

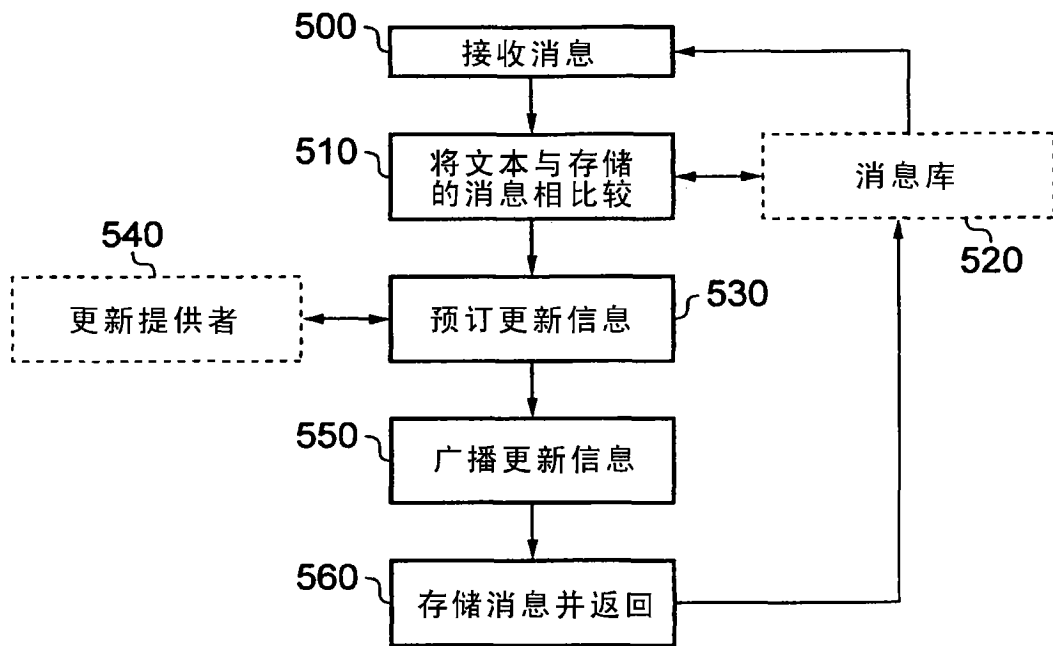


图 8

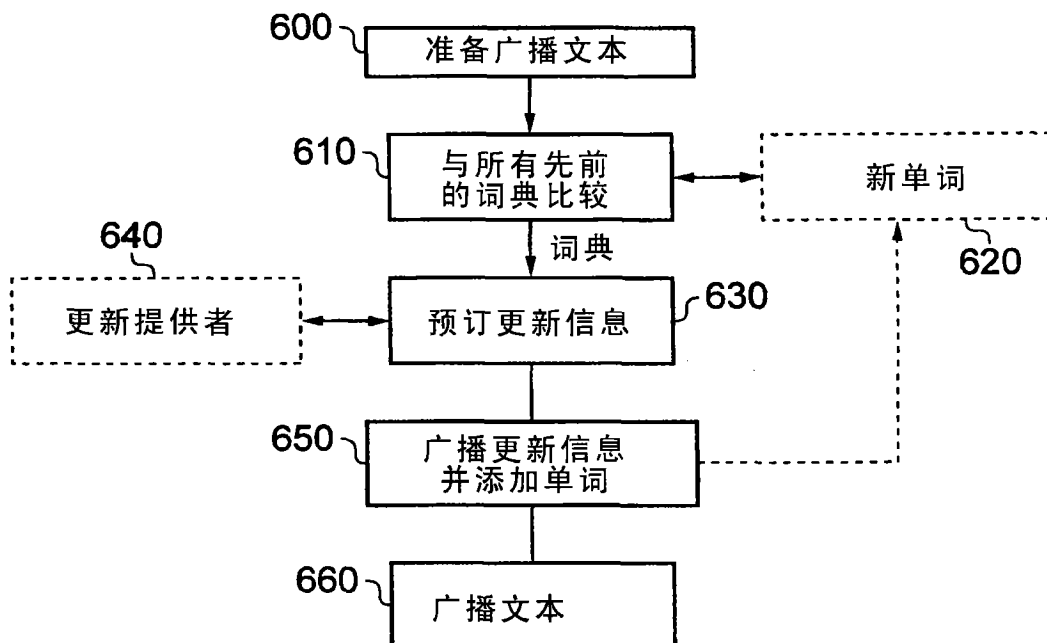


图 9