



(12) 发明专利申请

(10) 申请公布号 CN 101770507 A

(43) 申请公布日 2010. 07. 07

(21) 申请号 200910261124. 4

(22) 申请日 2009. 12. 28

(30) 优先权数据

2008-332133 2008. 12. 26 JP

(71) 申请人 索尼株式会社

地址 日本东京都

(72) 发明人 浅野康治

(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 南霆

(51) Int. Cl.

G06F 17/30 (2006. 01)

G10L 15/02 (2006. 01)

G10L 15/00 (2006. 01)

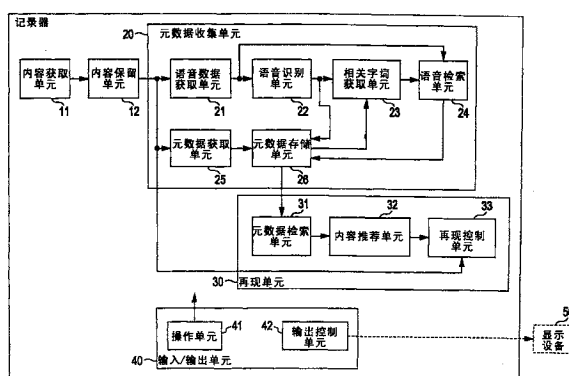
权利要求书 2 页 说明书 16 页 附图 8 页

(54) 发明名称

数据处理设备、数据处理方法和程序

(57) 摘要

本发明提供了数据处理设备、数据处理方法和程序。该数据处理设备包括：语音识别单元，被配置来对语音数据执行连续语音识别；相关字词获取单元，被配置来获取与通过连续语音识别而获得的至少一个字词相关的字词，作为与对应于包括语音数据的内容数据的内容相关的相关字词；以及语音检索单元，被配置来从语音数据检索相关字词的话语来获取已检索出话语的相关字词作为内容的元数据。



1. 一种数据处理设备,包括:
语音识别装置,用于对语音数据执行连续语音识别;
相关字词获取装置,用于获取与通过所述连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括所述语音数据的内容数据的内容相关的相关字词;以及
语音检索装置,用于从所述语音数据检索所述相关字词的话语来获取已检索出话语的相关字词作为所述内容的元数据。
2. 如权利要求 1 所述的数据处理设备,还包括话题估计装置,用于基于所述连续语音识别的结果来估计与所述语音数据相对应的语音的主旨的话题,其中
所述相关字词获取装置获取与所述话题相关的字词作为所述相关字词。
3. 如权利要求 2 所述的数据处理设备,其中所述相关字词获取装置从与通过所述连续语音识别而获得的所述至少一个字词相关的字词中获取除所述连续语音识别应识别的字词之外的字词作为所述相关字词。
4. 如权利要求 2 所述的数据处理设备,其中所述相关字词获取装置获取所述话题的句子中出现的新字词作为所述相关字词。
5. 如权利要求 2 所述的数据处理设备,其中
所述内容数据被指派有所述内容的元数据,并且
所述话题估计装置还基于指派给所述内容数据的元数据来估计所述话题。
6. 如权利要求 5 所述的数据处理设备,其中
所述内容数据是电视广播的广播数据中包括的节目的数据,
所述广播数据除所述节目的数据以外还包括作为所述节目的元数据的电子节目指南 EPG 数据,并且
所述话题估计装置还基于所述广播数据中包括的所述 EPG 数据来估计所述话题。
7. 如权利要求 5 所述的数据处理设备,其中在指派给所述内容数据的元数据包括固有名称的情形中,所述相关字词获取装置还获取与指派给所述内容数据的元数据中包括的所述固有名称相关的固有名称作为所述相关字词。
8. 如权利要求 2 所述的数据处理设备,其中所述相关字词获取装置从网络上的服务器获取所述相关字词。
9. 如权利要求 2 所述的数据处理设备,还包括:
元数据存储装置,用于与所述内容数据相关联地存储所述内容数据的元数据;
元数据检索装置,用于在输入了关键字时从存储在所述元数据存储装置中的所述元数据中检索与所述关键字匹配或类似的元数据;以及
内容推荐装置,用于推荐对应于与由所述元数据检索装置检索出的所述元数据相关联的内容数据的内容。
10. 如权利要求 9 所述的数据处理设备,还包括再现控制装置,用于在从由所述内容推荐装置推荐的内容中指定了要再现的内容时再现所指定的内容。
11. 一种数据处理方法,包括以下步骤:
对语音数据执行连续语音识别;
获取与通过所述连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括所述语音数据的内容数据的内容相关的相关字词;以及

从所述语音数据检索所述相关字词的话语来获取已检索出话语的相关字词作为所述内容的元数据；

所述步骤由数据处理设备执行。

12. 一种程序,用于使计算机充当下述装置：

语音识别装置,用于对语音数据执行连续语音识别；

相关字词获取装置,用于获取与通过所述连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括所述语音数据的内容数据的内容相关的相关字词；以及

语音检索装置,用于从所述语音数据检索所述相关字词的话语来获取已检索出话语的相关字词作为所述内容的元数据。

13. 一种数据处理设备,包括：

语音识别单元,被配置来对语音数据执行连续语音识别；

相关字词获取单元,被配置来获取与通过所述连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括所述语音数据的内容数据的内容相关的相关字词；以及

语音检索单元,被配置来从所述语音数据检索所述相关字词的话语来获取已检索出话语的相关字词作为所述内容的元数据。

数据处理设备、数据处理方法和程序

技术领域

[0001] 本发明涉及数据处理设备、数据处理方法和程序。更具体而言,本发明涉及被配置来辅助获取例如语音或图像内容的元数据的数据处理设备、数据处理方法和程序。

背景技术

[0002] 为了从繁多的电视广播节目的内容中推荐期望内容,例如用户感兴趣的内容,必须检索该期望内容。为了实现内容检索,必须预先向内容指派元数据。

[0003] 语音识别技术已被考虑来作为向内容指派元数据的方法。

[0004] 具体而言,在内容包括语音(例如电视广播节目)并且内容的内容数据包括语音数据的情形中,可对语音数据进行语音识别,通过语音识别获得的字词(word)可用作该内容的元数据。

[0005] 然而,即使在利用能够识别大量字词的大词汇连续语音识别系统执行语音识别的情形中,通过语音识别可获得的字词也限于预先登记在该系统用于语音识别的字典中的那些。

[0006] 因此,难以获得尚未被登记在字典中的字词(下文称作“未登记字词”)用作元数据。

[0007] 未登记字词包括最近变得被频繁使用的新近出现的字词(下文称作“新字词”)和诸如不太出名的地点的名称之类的固有名称(propersname)。

[0008] 为了获得这种新字词和固有名称作为元数据,必须将这些未登记字词作为识别目标字词登记到字典中。

[0009] 然而,将包括新字词和固有名称的未登记字词登记到字典中来增大识别目标字词的数目的作法也导致了执行语音识别过程必需的时间的延长和语音识别的精度降低。

[0010] 为了提高短话语(utterance)中的字词的识别率,提出了一种执行连续语音识别的方法,在该方法中从要识别的语集(corpus)生成连续语音识别字典,还在考虑到该连续语音识别字典的情况下生成用于提高对未登记字词的识别的补充识别字典,然后使用该连续语音识别字典和补充识别字典二者来执行连续语音识别(例如参见日本未实审专利申请 No. 2008-242059)。

发明内容

[0011] 可以通过使用语音检索技术来获得元数据,在该技术中在语音数据中搜索特定字词的话语,并且检测出该特定字词的话语发生的定时(时间)。

[0012] 具体而言,在语音检索中,可以在语音数据中搜索作为内容的元数据的字词的话语,并且可以获得其话语被包括在该语音数据中的字词作为该内容的元数据。

[0013] 然而,存在大量的需要作为内容的元数据而获得的字词。如果要检索这样大量的字词,语音检索将花费可观的时间量,导致难以获取元数据。

[0014] 考虑到前述问题,期望提供一种能辅助获取元数据的技术。

[0015] 根据本发明的实施例,提供了一种数据处理设备,该数据处理设备包括:语音识别单元,被配置来对语音数据执行连续语音识别;相关字词获取单元,被配置来获取与通过连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括语音数据的内容数据的内容相关的相关字词;以及语音检索单元,被配置来从语音数据检索相关字词的话语来获取已检索出话语的相关字词作为内容的元数据。根据本发明另一个实施例,提供了一种用于使计算机充当该数据处理设备的程序。

[0016] 根据本发明又一个实施例,提供了一种数据处理方法,该方法包括以下步骤:对语音数据执行连续语音识别;获取与通过连续语音识别而获得的至少一个字词相关的字词,作为与对应于包括语音数据的内容数据的内容相关的相关字词;以及从语音数据检索相关字词的话语来获取已检索出话语的相关字词作为内容的元数据;所述步骤由数据处理设备执行。

[0017] 在本发明的上述实施例中,对语音数据执行连续语音识别,并且获取与通过连续语音识别而获得的至少一个字词相关的任意字词,作为与对应于包括该语音数据的内容数据的内容相关的相关字词。然后,在该语音数据中搜索该相关字词的话语,并且获得已检索到其话语的相关字词作为该内容的元数据。

[0018] 注意,该数据处理设备可以是独立的设备,或者可以是一个设备中包括的内部模块。

[0019] 此外,程序可以提供为通过发送介质发送的程序或者记录在记录介质上的程序。

[0020] 根据本发明的上述实施例,可以容易地获取元数据。

附图说明

[0021] 图 1 是示出了本发明被应用到的记录器的第一实施例的配置示例的框图;

[0022] 图 2 是示出了元数据收集过程的流程图;

[0023] 图 3 是示出了再现过程的流程图;

[0024] 图 4 是示出了本发明被应用到的记录器的第二实施例的配置示例的框图;

[0025] 图 5 示出了利用矢量空间方法的话题估计方法;

[0026] 图 6A 和 6B 示出了“tf”和“idf”;

[0027] 图 7 是示出了元数据收集方法的另一个流程图;以及

[0028] 图 8 是示出了本发明被应用到的计算机的实施例的配置示例的框图。

具体实施方式

[0029] < 第一实施例 >

[0030] [本发明被应用到的记录器的第一实施例的配置示例]

[0031] 图 1 是示出了本发明被应用到的记录器的第一实施例的配置示例的框图。

[0032] 参考图 1,该记录器例如是硬盘 (HD) 记录器,并且包括:内容获取单元 11、内容保留单元 12、元数据收集单元 20、再现单元 30 和输入 / 输出单元 40。

[0033] 内容获取单元 11 获取例如组成电视广播节目的图像和语音内容的内容数据,并将该内容数据提供给内容保留单元 12。

[0034] 在内容数据伴有与该内容数据相对应的内容的元数据的情形中,内容获取单元 11

还获取该元数据并将其提供给内容保留单元 12。

[0035] 具体而言,内容获取单元 11 可以是调谐器,该调谐器接收诸如数字电视广播之类的电视广播中的广播数据。内容获取单元 11 接收从广播电台(未示出)发送(广播)来的广播数据,并将所获取的数据提供给内容保留单元 12。

[0036] 在这里,广播数据包括作为内容或节目的数据的内容数据。广播数据还可适当地包括节目的元数据(即,指派给节目(内容)的元数据),例如电子节目指南(EPG)数据。

[0037] 尽管作为节目的数据的内容数据可以包括节目的图像数据和与该图像数据相伴的语音数据,但是要由内容获取单元 11 获取的内容数据仅必须包括至少语音数据,例如音乐数据。

[0038] 注意,内容获取单元 11 可包括通信接口(I/F),该通信接口经由诸如局域网(LAN)或因特网之类的网络来执行通信。在该情形中,内容获取单元 11 接收并获取从网络上的服务器发送来的内容数据和元数据。

[0039] 内容保留单元 12 可配置有大容量记录(存储)介质,例如硬盘(HD)。必要的话内容保留单元 12 记录(或者存储或保留)从内容获取单元 11 提供来的内容数据。

[0040] 在诸如 EPG 数据之类的内容(节目)的元数据被从内容获取单元 11 提供到内容保留单元 12 的情形中,内容保留单元 12 还记录该元数据。

[0041] 注意,将内容数据记录到内容保留单元 12 中对应于“记录”(包括编程记录和所谓的“自动记录”)。

[0042] 元数据收集单元 20 充当数据处理设备,该数据处理设备收集其内容数据已被记录在内容保留单元 12 中的内容的元数据。

[0043] 具体而言,元数据收集单元 20 由语音数据获取单元 21、语音识别单元 22、相关字词获取单元 23、语音检索单元 24、元数据获取单元 25 和元数据存储单元 26 构成。

[0044] 语音数据获取单元 21 通过从内容保留单元 12 读取多个内容项(这些内容项的内容数据已被记录在内容保留单元 12 中)中的正关注的感兴趣的内容的内容数据中所包括的语音数据,来获取该语音数据。语音数据获取单元 21 将所获取的语音数据提供给语音识别单元 22 和语音检索单元 24。

[0045] 语音识别单元 22 可以具有执行可识别大量字词的大词汇连续语音识别的功能。语音识别单元 22 对从语音数据获取单元 21 提供来的语音数据执行(连续)语音识别。

[0046] 此外,语音识别单元 22 将至少一个所获得的字词(字串)作为语音识别的结果提供给相关字词获取单元 23 和元数据存储单元 26。

[0047] 在这里,语音识别单元 22 中结合有字典,并且利用登记在该字典中的字词作为识别目标字词来执行语音识别。因此,语音识别单元 22 通过语音识别获得的字词是登记在该字典中的字词。

[0048] 相关字词获取单元 23 获取与从语音识别单元 22 提供来的通过语音识别获得的字词相关的任何字词,作为与感兴趣的内容相关的相关字词。相关字词获取单元 23 将所获取的相关字词提供给语音检索单元 24。

[0049] 例如,相关字词获取单元 23 可以利用辞典来获取含义与通过语音识别所获得的字词的语义接近的字词作为相关字词。

[0050] 或者,相关字词获取单元 23 可以使用关于字词同时出现的概率的数据来获取可

能与通过语音识别所获得的字词一起出现的字词,即,与通过语音识别所获得的字词同时出现的概率不低于一预定阈值的字词作为相关字词。

[0051] 该辞典或同时出现概率数据可以作为静态数据存储在与相关字词获取单元 23 中。

[0052] 此外,相关字词获取单元 23 还可从网络上的服务器获取相关字词(或者用于获取相关字词的信息)。

[0053] 具体而言,相关字词获取单元 23 可以执行抓取(crawling)来从网络上的服务器收集信息,并且使用所收集的信息来更新该辞典或同时出现概率数据。然后,相关字词获取单元 23 可以使用更新后的辞典或同时出现概率数据来获取相关字词。

[0054] 为了更新辞典,可以向辞典添加字词,或者可以更新该辞典中的字词之间的联接(关系)。为了更新同时出现概率数据,可以向同时出现概率数据添加字词,或者可以更新同时出现概率的值。

[0055] 如上所述,相关字词获取单元 23 能够从网络上的服务器获取相关字词。这允许获取未被登记在语音识别单元 22 中结合的字典中的字词(例如最近变得频繁使用的新字词或固有名称)作为相关字词。

[0056] 语音检索单元 24 在从语音数据获取单元 21 提供来的语音数据中搜索从相关字词获取单元 23 提供来的相关字词的话语。随后语音检索单元 24 获取已找到其话语的相关字词作为感兴趣内容(即,与包括从语音数据获取单元 21 提供来的语音数据的内容数据相对应的内容)的元数据。语音检索单元 24 将所获取的元数据提供给元数据存储单元 26。

[0057] 在感兴趣内容的元数据被记录在内容保留单元 12 中的情形中,元数据获取单元 25 通过从内容保留单元 12 读取感兴趣内容的元数据来获取元数据,并将所获取的元数据提供给元数据存储单元 26。

[0058] 元数据存储单元 26 存储作为语音识别的结果从语音识别单元 22 提供来的字词,作为感兴趣内容的元数据。

[0059] 元数据存储单元 26 还存储从语音检索单元 24 和元数据获取单元 25 提供来的感兴趣内容的元数据。

[0060] 在这里,在元数据存储单元 26 存储的元数据中,作为语音识别的结果从语音识别单元 22 提供来的字词也称作“识别结果元数据”。

[0061] 此外,在元数据存储单元 26 存储的元数据中,从语音检索单元 24 提供来的元数据也称作“检索结果元数据”。

[0062] 另外,在元数据存储单元 26 存储的元数据中,从元数据获取单元 25 提供来的元数据(即,(预先)指派给感兴趣内容的元数据)也称作“预指派元数据”。

[0063] 注意,在元数据收集单元 20 中,元数据存储单元 26 被配置来存储作为语音识别的结果从语音识别单元 22 提供来的所有字词,作为感兴趣内容的元数据。或者,元数据存储单元 26 可被配置来仅存储必要的字词作为感兴趣内容的元数据。

[0064] 具体而言,登记在语音识别单元 22 中结合的字典中的每个字词可以被施加以一个标志,该标志例如指示是否将该字词存储为元数据。在该情形中,在作为语音识别的结果从语音识别单元 22 提供来的字词中,元数据存储单元 26 可以仅将施加有指示该字词应当被存储为元数据的标志的字词存储为感兴趣内容的元数据。

[0065] 此外,在元数据收集单元 20 中,相关字词获取单元 23 可被配置来获取以下字词作

为相关字词：与作为语音识别的结果从语音识别单元 22 提供来的字词相关的字词，以及作为预指派元数据存储单元 26 中的字词相关的字词。

[0066] 具体而言，例如在存储在元数据存储单元 26 中的预指派元数据包括固有名称的情形中，相关字词获取单元 23 可获取与该固有名称相关的固有名称等作为相关字词。

[0067] 更具体而言，例如假设感兴趣内容是 TV 戏剧节目，并且预指派元数据包括作为感兴趣内容的 TV 戏剧节目中出现的演员的姓名。在该情形中，可以获取先前曾与该演员一起表演过的演员的姓名和该演员曾在其中扮演过角色的其他 TV 节目的标题作为相关字词。可以例如从提供 TV 节目的信息的 web 服务器获取这些演员的姓名和 TV 节目的标题作为相关字词。

[0068] 此外，在元数据收集单元 20 中，相关字词获取单元 23 可被配置来从与语音识别单元 22 通过语音识别获得的字词相关的字词中仅获取除应当在语音识别过程中被识别的字词之外的字词来作为相关字词。

[0069] 具体而言，在某个字词 A 是相关字词并且语音检索单元 24 已从语音数据检索到了该相关字词 A 的话语的情形中，该相关字词 A 作为感兴趣内容的元数据被存储到元数据存储单元 26 中。

[0070] 同时，如果该字词 A 是识别目标字词，即，如果字词 A 已登记在语音识别单元 22 中结合的字典中，则假如在语音识别单元 22 中已成功执行了语音识别，那么该字词 A 可以作为识别结果元数据而被存储到元数据存储单元 26 中。

[0071] 因此，在该情形中，语音检索单元 24 不必从语音数据检索字词 A 作为相关字词，因为字词 A（是识别目标字词）将作为识别结果元数据被存储在元数据存储单元 26 中。

[0072] 相关字词获取单元 23 被配置来仅获取除应当被语音识别单元 22 识别的字词之外的字词来作为相关字词。即，相关字词获取单元 23 被配置为不获取语音识别的目标字词作为相关字词。这可以减少成为语音检索单元 24 执行的语音检索的目标字词的相关字词的数目，从而可以确保语音检索单元 24 的语音检索的快速处理。

[0073] 注意，在元数据收集单元 20 中，元数据存储单元 26 被配置来与已记录在内容保留单元 12 中的感兴趣内容的内容数据相关联地存储该感兴趣内容的元数据。例如，元数据存储单元 26 可以一同存储感兴趣内容的元数据和用于标识该感兴趣内容的标识信息。

[0074] 此外，当已在感兴趣内容的语音数据中找到了相关字词的话语时，必要的话，元数据存储单元 26 可以与作为该相关字词的元数据相关联地存储指示该相关字词的话语在语音数据中的定时的定时信息。

[0075] 即，在该情形中，语音检索单元 24 获取在语音数据中已找到了话语的相关字词来作为元数据，并且还检测该相关字词的话语在该语音数据中的定时。语音检索单元 24 随后将作为元数据的该相关字词与指示该相关字词的话语的定时的定时信息一起提供给元数据存储单元 26。

[0076] 作为响应，元数据存储单元 26 彼此相关联地存储从语音检索单元 24 提供来的作为元数据的该相关字词及其定时信息。

[0077] 在这里，对于指示相关字词的话语在语音数据中的定时的定时信息，可以采用相对于语音数据的开头（即，与包括该语音数据的内容数据相对应的内容的开头）的时间（例如时间码）。

[0078] 再现单元 30 充当再现内容保留单元 12 中记录的内容数据的数据处理设备。

[0079] 具体而言,再现单元 30 包括元数据检索单元 31、内容推荐单元 32 和再现控制单元 33。

[0080] 在用户操作操作单元 41 (稍后将描述) 来输入用于检索内容的关键字的情形中,元数据检索单元 31 搜索与该输入关键字匹配或类似的元数据。关键字可以例如是用户感兴趣的演员的姓名。

[0081] 具体而言,元数据检索单元 31 从存储在元数据存储单元 26 中的元数据中检索与已通过操作单元 41 的用户操作输入的关键字匹配或类似的元数据。

[0082] 此外,元数据检索单元 31 还向内容推荐单元 32 提供标识信息,该标识信息用于标识对应于与下述元数据相关联的内容数据的内容:所述元数据是元数据存储单元 26 中的与关键字匹配或类似的元数据(下文也称作“匹配元数据”)。

[0083] 内容推荐单元 32 将由从元数据检索单元 31 接收到的标识信息所标识出的内容作为要推荐给观众/听众的推荐内容,并且生成推荐内容的标题的列表。内容推荐单元 32 随后通过输出控制单元 42 使该推荐内容的标题的列表显示在显示设备 50 (例如,电视接收机(TV 机)) 上,以便推荐对推荐内容的观看/收听。

[0084] 此外,在用户操作操作单元 41 来从显示在显示设备 50 上的标题列表中选择要再现的推荐内容的标题时,内容推荐单元 32 向再现控制单元 33 发送对作为要再现的内容的该标题的推荐内容的指定。

[0085] 在接收到来自内容推荐单元 32 的对要再现的内容的指定后,再现控制单元 33 从内容保留单元 12 读取该要再现的内容数据以进行再现。

[0086] 具体而言,再现控制单元 33 对要再现的内容的内容数据执行解码和其他必要的处理,并经由输出控制单元 42 将结果数据提供给显示设备 50。

[0087] 结果,在显示设备 50 中,与要再现的内容的内容数据中包括的图像数据相对应的图像被显示在显示屏幕上,并且与该内容数据中包括的语音数据相对应的声音被从内置扬声器等输出。

[0088] 输入/输出单元 40 充当用于执行该记录器的必要的输入/输出操作的接口。

[0089] 具体而言,输入/输出单元 40 由操作单元 41 和输出控制单元 42 构成。

[0090] 操作单元 41 可以是由用户操作的键盘(具有键和按钮)或遥控器。操作单元 41 将与用户操作相对应的信号适当地提供(输入)到各个模块。

[0091] 输出控制单元 42 控制到诸如显示设备 50 之类的外部设备的数据输出。具体而言,输出控制单元 42 可以向显示设备 50 输出由内容推荐单元 32 生成的推荐内容的标题的列表和要由再现控制单元 33 再现的内容的内容数据。

[0092] [对元数据收集过程的描述]

[0093] 图 1 中示出的记录器执行用于收集内容的元数据的元数据收集过程。

[0094] 下面将参考图 2 描述元数据收集过程。

[0095] 这里假设至少一个内容项的内容数据已被记录在内容保留单元 12 中。

[0096] 该元数据收集过程开始于任意时刻。在步骤 S11 中,元数据收集单元 20 从内容数据已被记录在内容保留单元 12 中的内容数据项中选择要收集其元数据的内容(和其元数据还未被收集的内容)作为关注的感兴趣内容。

[0097] 过程随后从步骤 S11 前进到步骤 S12,在该步骤中元数据获取单元 25 判断感兴趣内容的元数据是否已被记录在内容保留单元 12 中。

[0098] 如果在步骤 S12 中判断感兴趣内容的元数据被记录在内容保留单元 12 中,则过程前进到步骤 S13,在该步骤中元数据获取单元 25 从内容保留单元 12 获取感兴趣内容的元数据。此外,元数据获取单元 25 将感兴趣内容的元数据作为预指派元数据提供给元数据存储单元 26,使元数据存储单元 26 将该元数据与感兴趣内容的内容数据相关联地存储。过程随后从步骤 S13 前进到步骤 S14。

[0099] 如果在步骤 S12 中判断感兴趣内容的元数据未被记录在内容保留单元 12 中,则过程跳过步骤 S13,前进到步骤 S14。

[0100] 在步骤 S14 中,语音数据获取单元 21 从内容保留单元 12 获取感兴趣内容的内容数据中包括的语音数据(语音波形的数据),并将所获取的数据提供给语音识别单元 22 和语音检索单元 24。过程随后前进到步骤 S15。

[0101] 在步骤 S15 中,语音识别单元 22 对从语音数据获取单元 21 接收到的语音数据执行语音识别,并将作为语音识别的结果获得的至少一个字词(字串)提供给相关字词获取单元 23 和元数据存储单元 26。过程随后前进到步骤 S16。

[0102] 在这里,在作为语音识别的结果接收到来自语音识别单元 22 的字词时,必要的话元数据存储单元 26 将接收到的字词作为识别结果元数据与感兴趣内容的内容数据相关联地存储。

[0103] 此外,为了执行语音识别,语音识别单元 22 使用例如隐式 Markov 模型(HMM)作为声学模型,并且使用 N-gram 或其他统计语言模型作为语言模型。

[0104] 在步骤 S16 中,相关字词获取单元 23 获取与作为语音识别的结果从语音识别单元 22 提供来的字词相关的任何字词作为相关字词。

[0105] 相关字词不仅可包括与通过语音识别获得的字词相关的字词,还可包括与在步骤 S13 中存储在元数据存储单元 26 中的感兴趣内容的预指派元数据中包括的字词相关的字词。

[0106] 此外,例如在用户简档已被登记在图 1 中示出的记录器等中的情形中,相关字词获取单元 23 可以根据该简档来估计用户可能感兴趣的对象,并获取代表该对象或者与该对象相关的字词。在该情形中,相关字词获取单元 23 可以将与用户感兴趣的对象相关的字词作为相关字词。

[0107] 一旦获取了相关字词,相关字词获取单元 23 就可以生成在其中登记了相关字词的单词列表,并将该单词列表提供给语音检索单元 24。过程随后从步骤 S16 前进到步骤 S17。

[0108] 在步骤 S17 中,语音检索单元 24 判断从相关字词获取单元 23 提供来的单词列表中是否登记有任何相关字词。

[0109] 如果在步骤 S17 中判断至少一个相关字词被登记在该单词列表中,则过程前进到步骤 S18,在该步骤中语音检索单元 24 选择登记在单词列表中的相关字词之一作为所关注的感兴趣字词。过程随后前进到步骤 S19。

[0110] 在步骤 S19 中,语音检索单元 24 执行语音检索来从自语音数据获取单元 21 提供来的感兴趣内容的语音数据中检索出该感兴趣字词的话语,随后过程前进到步骤 S20。

[0111] 在之类,从语音数据检索感兴趣字词的话语的语音检索可利用所谓的“关键字测定位点”来执行,或者可以下面的方式执行。对于从语音数据获取单元 21 提供到语音检索单元 24 的语音数据,可以生成代表音素的指标(index)和音素在语音数据中的位置,并且可从所述指标检索构成该感兴趣字词的音素的序列。

[0112] 在步骤 S20 中,语音检索单元 24 基于步骤 S19 中执行的语音检索的结果,来判断感兴趣字词的话语(即,感兴趣字词的话语的语音数据)是否被包括在感兴趣内容的语音数据中。

[0113] 如果在步骤 S20 中判断感兴趣内容的语音数据包括感兴趣字词的话语,则过程前进到步骤 S21。

[0114] 在步骤 S21 中,语音检索单元 24 将感兴趣字词作为检索结果元数据提供给元数据存储单元 26,使元数据存储单元 26 将该元数据与感兴趣内容的内容数据相关联地存储。过程随后前进到步骤 S22。

[0115] 在这里,在语音检索单元 24 中,在对感兴趣字词的语音检索时可检测出感兴趣字词的话语在语音数据中的定时,并且指示该定时的定时信息可与作为感兴趣字词的检索结果元数据一起被提供给元数据存储单元 26。

[0116] 在该情形中,元数据存储单元 26 与感兴趣内容的内容数据相关联地存储从语音检索单元 24 提供来的检索结果元数据和定时信息。

[0117] 另一方面,如果在步骤 S20 中判断感兴趣内容的语音数据不包括感兴趣字词的话语,则过程跳过步骤 S21 前进到步骤 S22。

[0118] 在步骤 S22 中,语音检索单元 24 从字词列表中删除该感兴趣字词,并且过程返回到步骤 S17 来重复类似的过程。

[0119] 随后,如果在步骤 S17 中判断没有相关字词被登记在字词列表中,则该元数据收集过程结束。

[0120] 如上所述,在该元数据收集过程中,在语音识别单元 22 中对感兴趣内容的语音数据执行语音识别(连续语音识别),并且在相关字词获取单元 23 中,获取与通过语音识别获得的至少一个字词相关的任意字词作为相关字词。随后,在语音检索单元 24 中,在感兴趣内容的语音数据中搜索相关字词的话语,并且获取找到了其话语的相关字词作为感兴趣内容的元数据。

[0121] 因此,在语音检索单元 24 中,与通过语音识别获得的至少一个字词相关的字词被作为相关字词,并被用作检索(语音检索)的目标字词。由于语音检索的目标字词如上所述被限制到相关字词,因此与对希望获取来作为内容的元数据的所有字词执行语音检索的情形相比,可以在较短的时间段中执行语音检索过程。

[0122] 结果,可以高效容易地获取内容的元数据。此外,即使不是语音识别的目标字词的

[0123] 此外,例如在相关字词获取单元 23 被配置来从诸如因特网之类的网络上的服务器获取相关字词的情形中,可以从逐日更新所存储的信息的服务器上的 web 页面获取新出现的字词(新字词)和固有名称来作为相关字词。因此,可以容易地获取这种新字词和固有名称作为元数据。

[0124] [对再现过程的描述]

[0125] 除了元数据收集过程之外,图 1 中示出的记录器还执行再现过程,在再现过程中,利用在元数据收集过程中收集的元数据对内容进行推荐和再现。

[0126] 现在将结合图 3 来描述该再现过程。

[0127] 在这里假设元数据收集过程已执行完毕,并且元数据存储单元 26 存储了其内容数据被记录在内容保留单元 12 中的至少一个内容项的元数据。

[0128] 在再现过程中,首先在步骤 S41 中,元数据检索单元 31 判断是否已输入了关键字。

[0129] 如果在步骤 S41 中判断尚未输入关键字,则过程返回到步骤 S41。

[0130] 如果在步骤 S41 中判断已输入了关键字,即,在用户已通过操作操作单元 41 输入了关键字时,过程前进到步骤 S42。

[0131] 在该示例中,关键字是通过操作单元 41 的用户操作输入的。或者,例如在用户简档已被登记在图 1 中示出的记录器等中的情形中,可以使用该简档来输入关键字。即,可以根据用户简档来估计用户感兴趣的对象,从而可以输入代表该对象的字词等作为关键字。

[0132] 在步骤 S42 中,元数据检索单元 31 在存储在元数据存储单元 26 中的元数据中搜索与通过操作单元 41 的用户操作输入的关键字匹配或类似的元数据(匹配元数据)。过程随后前进到步骤 S43。

[0133] 在步骤 S43 中,元数据检索单元 31 检测与匹配元数据相关联的内容数据(其中匹配元数据与通过步骤 S42 中的检索获得的关键字匹配或类似),并将用于标识与检测出的内容数据相对应的内容的标识信息提供给内容推荐单元 32。

[0134] 过程随后从步骤 S43 前进到步骤 S44,在步骤 S44 中内容推荐单元 32 将从元数据检索单元 31 接收到的标识信息所标识出的内容作为推荐内容推荐,然后过程前进到步骤 S45。

[0135] 具体而言,内容推荐单元 32 生成推荐内容的标题的列表,并将该列表提供给输出控制单元 42。

[0136] 作为响应,输出控制单元 42 将从内容推荐单元 32 接收到的标题列表提供给显示设备 50 进行显示。

[0137] 在步骤 S45 中,再现控制单元 33 判断是否指定了要被再现的内容。

[0138] 如果在步骤 S45 中判断已指定了要被再现的内容,即,在用户已操作了操作单元 41 来从在显示设备 50 上显示的标题的列表中选择要再现的推荐内容的标题,并且内容推荐单元 32 已响应于操作单元 41 的用户操作而指示再现控制单元 33 再现由用户选择的标题的推荐内容的情形中,过程前进到步骤 S46。在步骤 S46 中,再现控制单元 33 通过从内容保留单元 12 读取要再现内容的内容数据来再现该内容。

[0139] 具体而言,再现控制单元 33 对要再现的内容的内容数据执行解码和其他必要的处理,并将结果数据提供给输出控制单元 42。输出控制单元 42 从再现控制单元 33 接收内容数据,并将该数据提供给显示设备 50。因此,在显示设备 50 中,显示与要再现的内容的内容数据中包括的图像数据相对应的图像,并且同时输出与该内容数据中包括的语音数据相对应的声音。

[0140] 此后,例如在完成了对要再现的内容的所有内容数据的再现时,再现过程结束。

[0141] 另一方面,如果在步骤 S45 中判断尚未指定要再现的内容,则过程前进到步骤 S47,在该步骤中元数据检索单元 31 判断操作单元 41 是否已被操作来请求重新输入关键

字。

[0142] 如果在步骤 S47 中判断操作单元 41 已被操作来请求重新输入关键字,则过程返回到步骤 S41,重复类似的过程。

[0143] 如果在步骤 S47 中判断没有操作操作单元 41 来请求重新输入关键字,则过程前进到步骤 S48,在该步骤中元数据检索单元 31 判断操作单元 41 是否被操作来终止再现过程。

[0144] 如果在步骤 S48 中判断操作单元 41 尚未被操作来终止再现过程,则过程返回到步骤 S45,重复类似的过程。

[0145] 如果在步骤 S48 中判断操作单元 41 已被操作来终止再现过程,则终止再现过程。

[0146] 如上所述,根据元数据收集过程,可以获取诸如新字词和固有名称之类的不是语音识别的目标字词的单词作为元数据。此外,根据利用这种元数据执行的再现过程,可以正确地(精确地)检索、推荐和再现用户感兴趣的内容。

[0147] < 第二实施例 >

[0148] [本发明被应用到的记录器的第二实施例的配置示例]

[0149] 图 4 是示出了本发明被应用到的记录器的第二实施例的配置示例。

[0150] 在图 4 中,与图 1 中的部分相对应的部分由类似的标号标注,并且适当的话将不重复对它们的描述。

[0151] 图 4 中示出的记录器在配置方面与图 1 中示出的记录器相同,只是向元数据收集单元 20 添加了话题估计单元 61。

[0152] 话题估计单元 61 从语音识别单元 22 接收作为语音识别的结果而获得的至少一个单词。

[0153] 话题估计单元 61 基于作为语音识别的结果从语音识别单元 22 提供来的至少一个单词,来估计与感兴趣内容的语音数据相对应的语音的主旨的话题。话题估计单元 61 将估计出的话题提供给相关字词获取单元 23,作为感兴趣内容的话题。

[0154] 具体而言,话题估计单元 61 估计与通过语音识别而获得的至少一个单词(字串)类似的句子(文本)的话题,作为感兴趣内容的话题。

[0155] 在该情形中,相关字词获取单元 23 获取与从话题估计单元 61 提供来的感兴趣内容的话题相关的任意单词作为相关字词。

[0156] 话题估计单元 61 不仅可以基于作为语音识别的结果从语音识别单元 22 提供来的单词,而且可以基于存储在元数据存储单元 26 中的预指派元数据中包括的单词(其例如包括诸如 EPG 数据中包括的演员的名称和节目标题之类的固有名称和构成介绍节目大意的文本的单词)来估计感兴趣内容的单词。

[0157] 此外,在图 4 中,相关字词获取单元 23 所获取的相关字词不限于与感兴趣内容的话题相关的单词。如在图 1 的情形中一样,相关字词获取单元 23 还可以获取与存储在元数据存储单元 26 中的预指派元数据中包括的单词相关的单词作为相关字词。

[0158] 注意,相关字词获取单元 23 可以预先生成与各种话题相关的单词列表来作为话题相关单词列表。在该情形中,相关字词获取单元 23 可以获取登记在话题相关单词列表中的与感兴趣内容的话题相对应的单词作为相关字词。

[0159] 话题相关单词列表可以作为静态数据被存储在相关字词获取单元 23 中。

[0160] 此外,相关字词获取单元 23 还可从网络上的服务器获取相关字词(和用于获得相

关字词的信息)。

[0161] 具体而言,相关字词获取单元 23 可以执行抓取来从网络上收集诸如构成网页的文本(句子)之类的信息,并且使用该信息来更新话题相关字词列表。随后,相关字词获取单元 23 可以使用更新后的话题相关字词列表来获得相关字词。

[0162] 在这里,在更新话题相关字词列表时,登记在该话题相关字词列表中的字词可以被更新(修改)成以下字词:在通过抓取从网络收集的句子中的、与该话题相关字词列表相对应的话题的句子中,出现的次数不小于一预定阈值的字词,或者在出现次数方面排名较高的字词。

[0163] 如上所述,在相关字词获取单元 23 中,从网络上的服务器获取相关字词(登记在话题相关字词列表中)。这使得可以获得未登记在语音识别单元 22 中结合的字典中的字词(包括最近变得频繁使用的新字词和固有名称)作为相关字词。

[0164] [对话题估计方法的描述]

[0165] 下文将描述由图 4 中示出的话题估计单元 61 执行的估计感兴趣内容的话题的方法。

[0166] 可通过利用了所谓的话题模型的方法来估计话题,话题模型例如是概率性潜在语义分析(PLSA)或者潜在狄利克雷(Dirichlet)分配(LDA)。

[0167] 或者,可通过利用矢量空间方法的方法来估计话题,在矢量空间方法中,基于构成句子的字词来用矢量表达每个句子(字词串),并且使用矢量来获得要估计其话题的句子(“下文称作“输入句子””)和已知其话题的句子(下文也称作“示例句子”)之间的余弦距离。

[0168] 现在将参考图 5 来描述利用矢量空间方法的话题估计方法。

[0169] 根据矢量空间方法,每个句子(字词串)用一个矢量表达,并且获得由句子的矢量形成的夹角(余弦距离)作为句子之间的相似度或者句子之间的距离。

[0170] 更具体而言,在矢量空间方法中,准备了话题已知的句子(示例句子)的数据库(下文也称作“示例句子数据库”)。

[0171] 在图 5 中示例句子数据库存储从 #1 到 #K 的 K 个示例句子,并且在从 #1 到 #K 的 K 个示例句子中出现的字词中,采用被彼此不同表达的 M 个字词作为矢量的元素。

[0172] 在该情形中,如图 5 所示,该示例句子数据库中存储的每个示例句子可由 M 维矢量表达,这 M 维矢量具有 M 个字词 #1、#2、...、#M 作为其元素。

[0173] 对于与代表示例句子的矢量中的字词 #m(m = 1、2、...、M) 相对应的元素值,可以例如采用字词 #m 在该示例句子中出现的次数。

[0174] 与在示例句子的情形中一样,输入句子也可以用 M 维矢量表达。

[0175] 参考图 5,当代表某一示例句子 #k(k = 1、2、...、K) 的矢量被表达为 x_k ,代表输入句子的矢量为 y ,并且矢量 x_k 和 y 形成的夹角为 θ_k 时,可根据下式 (1) 获得余弦 $\cos \theta_k$ 。

[0176] $\cos \theta_k = x_k \cdot y / (|x_k| |y|) \dots (1)$

[0177] 在式 (1) 中,“ $\cdot$ ”表示内积,而“ $|z|$ ”表示矢量 z 的模。

[0178] 当矢量 x_k 和 y 在同一方向上时 $\cos \theta_k$ 取最大值“1”,而当矢量 x_k 和 y 在相反方向上时 $\cos \theta_k$ 取最小值“-1”。在这里,输入句子的矢量 y 和矢量句子 #k 的矢量 x_k 的元素取值“0”或更大,因此矢量 x_k 和 y 的 $\cos \theta_k$ 的最小值为“0”。

[0179] 在矢量空间方法中,针对每个示例句子 #k 计算 $\cos \theta_k$ 作为分数,并且例如获得具有最大分数的示例句子 #k 作为与输入句子最相似的示例句子。

[0180] 话题估计单元 61 使用在语音识别单元 22 中通过语音识别获得的至少一个字词串来作为输入句子,并获得与输入句子最相似的示例句子。话题估计单元 61 随后获得与输入句子最相似的示例句子的话题作为估计感兴趣内容的话题的结果。

[0181] 在图 5 中,采用了字词在句子中出现的次数作为代表该输入或示例句子的矢量中的元素的值。字词的这种出现次数称作“词频 (term frequency, tf)”。

[0182] 一半来说,在使用“tf”作为矢量中的元素值时,分数倾向于受出现频率较高的字词的影响。此外,在日语中,助词和助动词倾向于具有较高的出现频率。因此,在使用“tf”作为矢量中的元素的值的情形中,所获得的分数将受到输入或示例句子中包括的助词和助动词的较大影响。

[0183] 作为减少出现频率较高的字词的影响的方法,可以使用“逆文档频率 (idf)”或者作为“tf”和“idf”的組合的“TF-IDF”替换“tf”来作为矢量中的元素的值。

[0184] 在文本的总数 (通过将示例句子和输入句子的数目加总获得的) 表示为“N”,并且这 N 个文本中包括字词 t_i (矢量中的第 i 个元素) 的文本的数目表示为“ df_i ”时,则字词 t_i 的“idf”可例如由下式 (2) 表示。

[0185] $idf = \log_2(N/df_i) \dots (2)$

[0186] 根据式 (2),在某一文本中频繁出现的字词 (即,被认为代表了该文本的主旨 (话题) 的字词) 具有较大的“idf”值,而在多个文本中均匀出现的字词 (一般是助词和助动词) 每个都具有较小的“idf”值。

[0187] 图 6A 和 6B 示出了“tf”和“idf”。

[0188] 注意,图 6A 和 6B 示出了来自 Jin 等的、由 Iwanami Shoten 出版的“GENGO TO SHINRI NO TOUKEI ;KOTOBA TO KOUDOU NOKAKURITSU MODERU NIYORU BUNSEKI”的摘录。

[0189] 图 6A 示出了一个文本集合。

[0190] 在图 6A 中,为了简化说明,该文本集合包括两个文本:文本 #1: “Agrand slam homer smashed in the last inning has reversed the game.” 和文本 #2: “Power relationship between the ruling and opposition parties has been reversed in the Diet.”。

[0191] 图 6B 针对图 6A 中示出的文本集合示出了字词“love”、“reversed”、“Diet”和“homer”中的每个的“tf”和“idf”。

[0192] 在图 6B 中,用逗号隔开了“tf”和“idf”,以示为“tf, idf”的形式。

[0193] 注意,作为“tf”和“idf”的組合的“TF-IDF”例如由下面的式 (3) 表达。

[0194] $W_{i,j} = tf_{i,j} / \max_k \{tf_{k,j}\} \times \log_2(N/df_i) \dots (3)$

[0195] 在式 (3) 中,“ $W_{i,j}$ ”表示文本 #j 中的字词 t_i 的“TF-IDF”,“ $tf_{i,j}$ ”表示文本 #j 中字词 t_i 出现的频率,而“ $\max_k \{tf_{k,j}\}$ ”表示文本 #j 中出现的字词中具有最大出现频率的字词 t_k 的出现频率。此外,“N”表示文本总数 (通过将示例句子和输入句子的数目加总获得的),而“ df_i ”表示这 N 个文本中包括第 i 个字词 t_i 的文本的数目。

[0196] [对元数据收集过程的描述]

[0197] 参考图 7,将描述在图 4 中示出的记录器中执行的元数据收集过程。

[0198] 图 7 中示出的元数据收集过程中的步骤 S61 至 S65 分别与图 2 中示出的步骤 S11 至 S15 相同。

[0199] 在作为在步骤 S65 中语音识别单元 22 对从语音数据获取单元 21 提供来的感兴趣内容的语音数据执行的语音识别的结果而获得了至少一个字词（字串）时，该通过语音识别获得的至少一个字词作为识别结果元数据被提供给元数据存储单元 26 以进行存储，并且还提供给话题估计单元 61。

[0200] 此后，过程从步骤 S65 前进到步骤 S66，在步骤 S66 中话题估计单元 61 估计与作为语音识别的结果从语音识别单元 22 提供来的至少一个字词类似的句子（示例句子）的话题，作为感兴趣内容的话题。话题估计单元 61 随后将得到的话题提供给相关字词获取单元 23，过程然后前进到步骤 S67。

[0201] 在这里，话题估计单元 61 可以估计宽泛类别（较宽概念的类别）的话题，例如政治、经济、体育或杂项，或者可以估计更具体类别的话题。

[0202] 在步骤 S67 中，相关字词获取单元 23 获取与从话题估计单元 61 提供来的感兴趣内容的话题相关的任意字词作为相关字词。

[0203] 具体而言，相关字词获取单元 23 如上所述可存储与各种话题相关的字词列表作为话题相关字词列表，并且获取登记在话题相关字词列表中的与从话题估计单元 61 提供来的感兴趣内容的话题相对应的字词，作为相关字词。

[0204] 在这里，话题是从作为语音识别的结果而获得的至少一个字词估计出的，因此可以说与该话题相关的字词是与通过语音识别而获得的至少一个字词相关的字词。

[0205] 注意，相关字词获取单元 23 还可以获取与存储在元数据存储单元 26 中的预指派元数据中包括的字词相关的任意字词作为相关字词，这与在图 1 中示出的记录器的情形中一样。

[0206] 在相关字词获取单元 23 获取了相关字词时，其生成相关字词被登记在其中的字词列表，并将该列表提供给语音检索单元 24。随后过程从步骤 S67 前进到步骤 S68。此后，执行与图 2 中的各个步骤 S 17 至 S22 相同的步骤 S68 至 S73。

[0207] 注意，图 4 中示出的记录器使用图 7 中示出的元数据收集过程中所收集的元数据来执行推荐和再现内容的再现过程。该再现过程与图 3 中示出的相同，因此这里将不重复描述该过程。

[0208] 根据图 4 中示出的记录器，与图 1 中示出的记录器的情形中一样，可以高效容易地获得内容的元数据。此外，即使不是语音识别的目标字词的（例如新字词和固有名称）也可被获得作为元数据。

[0209] [对本发明被应用到的计算机的描述]

[0210] 上述处理序列可由硬件或软件执行。在要用软件执行这些处理序列的情形中，构成软件的程序被安装到通用计算机等上。

[0211] 图 8 示出了用于执行上述过程的程序被安装到的计算机的实施例的配置示例。

[0212] 程序可预先被记录在硬盘 105 或只读存储器 (ROM) 103 中，硬盘 105 和 ROM 103 是结合在计算机中的存储介质。

[0213] 或者，程序可临时或永久存储（记录）在可移除记录介质 111 中，例如柔性盘、致密盘只读存储器 (CD-ROM)、磁光 (MO) 盘、数字通用盘 (DVD)、磁盘、半导体存储器等。可移

除记录介质 111 可被提供为所谓的封装介质。

[0214] 程序可如上所述从可移除记录介质 111 安装到计算机中,程序也可以经由诸如数字广播卫星之类的人造卫星以无线方式从下载站点传送到计算机,或者经由诸如局域网 (LAN) 或因特网之类的网络以有线方式传送到计算机。以任意上述方式传送的程序可由计算机中的通信单元 108 接收到,并被安装到结合在计算机中的硬盘 105 中。

[0215] 计算机包括中央处理单元 (CPU) 102。CPU 102 经由总线 101 连接到输入 / 输出接口 110。当用户操作输入单元 107 (其包括键盘、鼠标、麦克风等) 来经由输入 / 输出接口 110 输入指令时, CPU 102 根据该指令执行 ROM 103 中存储的程序。或者, CPU 102 可将程序从硬盘 105 加载到随机存取存储器 (RAM) 104 进行执行, 其中程序可以是存储在硬盘 105 中的程序、从卫星或者网络传送来且由通信单元 108 接收到并被安装到硬盘 105 中的程序、或者从安装到驱动器 109 的可移除记录介质 111 读取并被安装到硬盘 105 中的程序。这样, CPU 102 执行上述流程图中示出的过程, 或者执行由上述框图中的配置执行的过程。CPU 102 然后经由输入 / 输出接口 110 从输出单元 106 (其包括液晶显示器 (LCD)、扬声器等) 输出这些过程的结果, 或者从通信单元 108 发送该结果, 或者必要的话将其记录在硬盘 105 等上。

[0216] 注意, 用于编码程序以使计算机执行各个过程的过程步骤不一定必须按照根据流程图所示顺序的时间序列执行。可以并行 (如同在并行处理中) 执行这些过程或者独立 (逐个对象地) 执行这些过程。

[0217] 此外, 程序可由单个计算机处理, 或者可由多个计算机以分布式方式处理。此外, 可将程序传送到远程计算机执行。

[0218] 现在将描述特定示例。从美国开始总统选举的 2008 年起, 美国总统候选人的姓名 “Barack Obama” 和 “John McCain” 突然变得频繁出现在电视广播节目的内容中。

[0219] 然而, 一般来说, 这些姓名不被包括在以前的用于大词汇连续语音识别的字典中, 因此必须更新字典来使得能够语音识别这些姓名。

[0220] 随着字典被重复更新并且字典中所包括的字词的数目增加, 发音上彼此类似的字词也增加了, 这可能导致语音识别的精度降低。

[0221] 同时, 在图 1 或 4 中示出的记录器中, 一般的大词汇连续语音识别被执行一次来分析 (语音识别) 内容的语音数据, 从而获取该语音数据中包括的一般字词。

[0222] 从上述美国总统候选人的姓名出现在其中的内容的语音数据中, 期望语音识别将获取诸如 “美国”、“总统” 和 “选举” 之类的字词作为一般字词。

[0223] 在语音识别之后, 在图 1 或 4 中示出的记录器中, 获取与通过语音识别所获得的至少一个字词相关的任意字词作为相关字词。

[0224] 具体而言, 在图 1 中示出的记录器中, 在相关字词获取单元 23 中, 获取可能与作为语音识别的结果而获得的字词一起出现的字词作为相关字词。

[0225] 可能与作为语音识别的结果而获得的字词一起出现的字词可以通过利用上述字词同时出现概率的数据来获取, 或者可以下面的方式获取。可以使用因特网上的搜索引擎来利用通过语音识别而获得的字词作为关键字来执行搜索。然后, 在通过搜索获得的 web 页面中, 可以选择出现频率较高的字词作为可能与通过语音识别而获得的字词一起出现的字词。

[0226] 在图 4 中示出的记录器的情形中, 在话题估计单元 61 中, 从作为语音识别的结果

而获得的至少一个字词估计内容的话题,并且在相关字词获取单元 23 中,获取出现在该话题的句子中的字词作为相关字词。

[0227] 对于话题估计,可以估计诸如“政治”、“经济”、“体育”等较宽泛的分类的话题,或者可以估计诸如“政治-日本”、“政治-美国”、“政治-中国”等具体分类的话题。

[0228] 一般来说,在估计更具体分类的话题时,可以改善在话题估计单元 61 的后续级中执行对在相关字词获取单元 23 中获取的相关字词的预测。即,由相关字词获取单元 23 获取的相关字词被包括在语音数据的话语中的概率提高了。然而,这导致了用于创建用来估计话题的模型所预先必需的学习数据的量增大。

[0229] 在图 4 中示出的记录器中,作为在相关字词获取单元 23 中获取与话题相关的字词作为相关字词的方法,可以用利用因特网上的新站点的方法来替换上述利用话题相关字词列表的方法。

[0230] 具体而言,例如假设如上所述已获取“美国”、“总统”和“选举”作为通过语音识别而获得的至少一个字词,并且从这些字词估计出的内容的话题为“政治-美国”。

[0231] 在该情形中,相关字词获取单元 23 访问因特网上的新站点来检查与“政治-美国”的话题相关的文章中出现的字词。然后,相关字词获取单元 23 将从当天起预定天数内的文章中出现的任何字词作为新字词(或者最近字词),并且获取这些新字词作为相关字词。

[0232] 例如,预期在美国举行总统选举的 2008 年,美国总统候选人的姓名“Barack Obama”、“John McCain”和“Hillary Clinton”被获得作为话题“政治-美国”的新字词。

[0233] 这确保可以获得诸如“Barack Obama”之类的及时字词作为元数据,而这是仅通过一般的大词汇连续语音识别不容易获取的。

[0234] 在这种情形中,在再现过程中(如图 3 中所示),当用户操作操作单元 41 来输入例如关键字“Barack Obama”时,执行对具有包括“BarackObama”的话语的语音数据的内容的推荐和/或再现。

[0235] 在这里,作为用于获取新字词作为相关字词的信息源,除了因特网上的服务器(站点)中包括的信息之外,也可以使用经由电视广播发送的 EPG 数据、经由数据广播发送的数据和针对听力受损的人的说明字幕。

[0236] 应当注意,图 1 和 4 中示出的记录器在以下几点不同于上述日本未实审专利申请 No. 2008-242059 中公开的技术。在图 1 和 4 中示出的记录器中,可以从诸如因特网之类的网络上的服务器获取相关字词。相反,在上述现有技术中的技术中,从要被识别的语集生成连续语音识别字典,并且在考虑到连续语音识别字典的情况下还生成用于提高对未登记字词的识别的补充识别字典,并且该连续语音识别字典和补充识别字典二者都用于连续语音识别,因此要识别的语集是必需的。

[0237] 此外,图 1 和 4 中示出的记录器与现有技术中的技术的不同在于:图 1 和 4 中示出的记录器通过利用与作为语音识别的结果而获得的字词同时出现的概率、或者通过使用从该字词估计出的话题来获取相关字词,而现有技术的技术在考虑到字词中包括的音节和该字词的语音的该部分的情况下生成补充识别字典。

[0238] 本申请包含与 2008 年 12 月 26 日提交给日本特许厅的日本在先专利申请 JP 2008-332133 中公开的主题相关的主题,该申请的所有内容通过引用结合于此。

[0239] 应当理解,本发明不限于上述实施方式,而是在不脱离本发明的精神的范围内可

以作出各种修改。

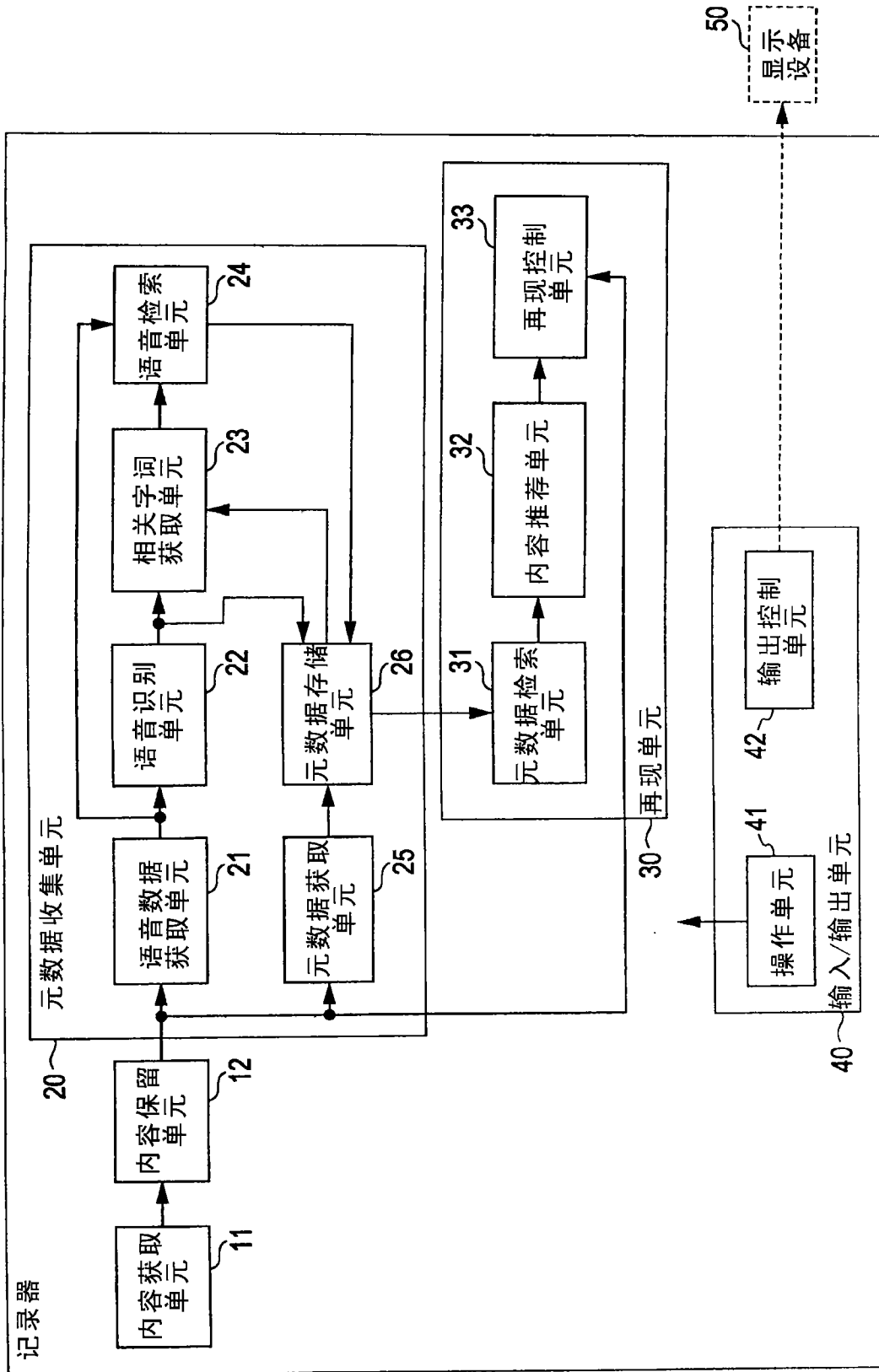


图 1

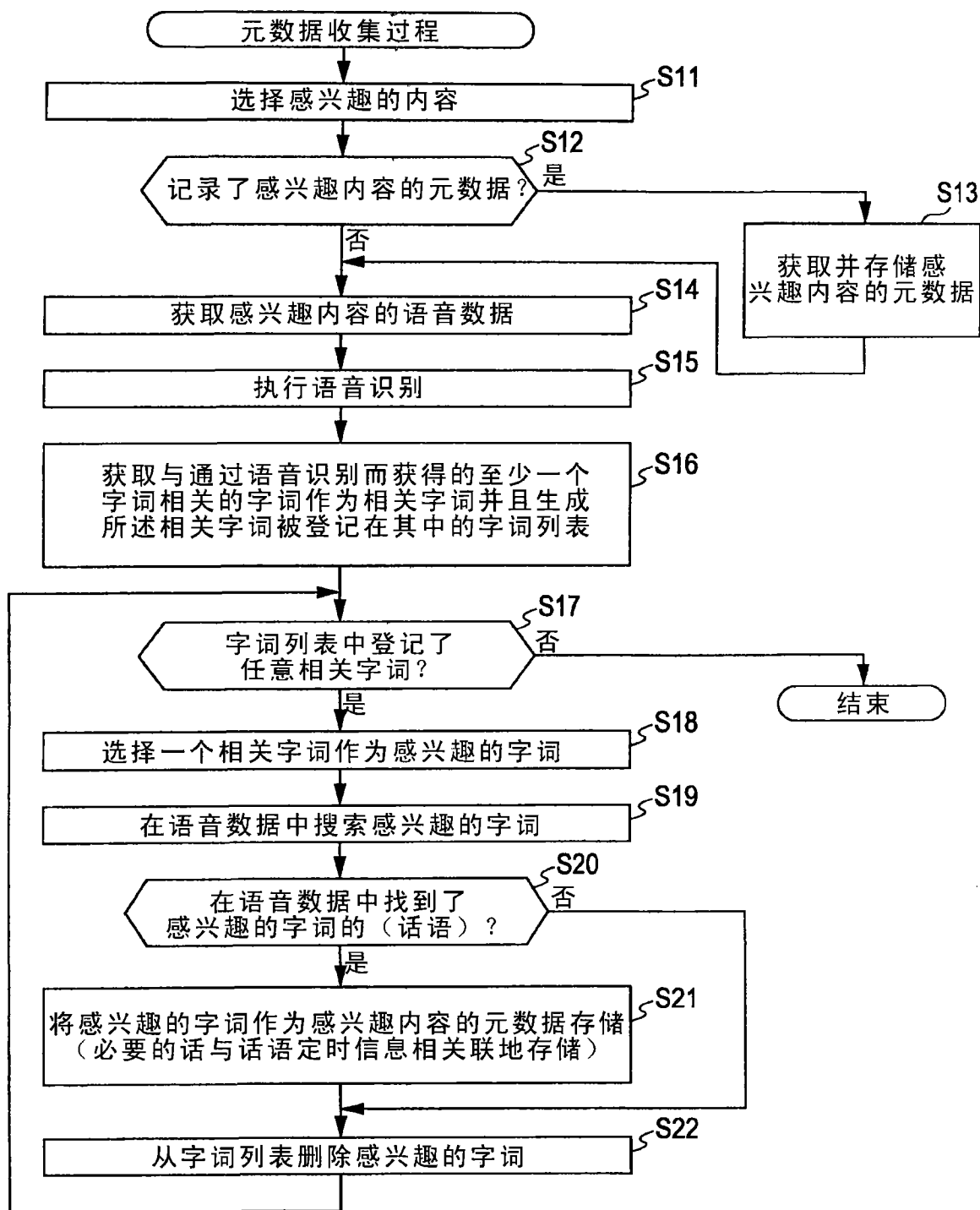


图 2

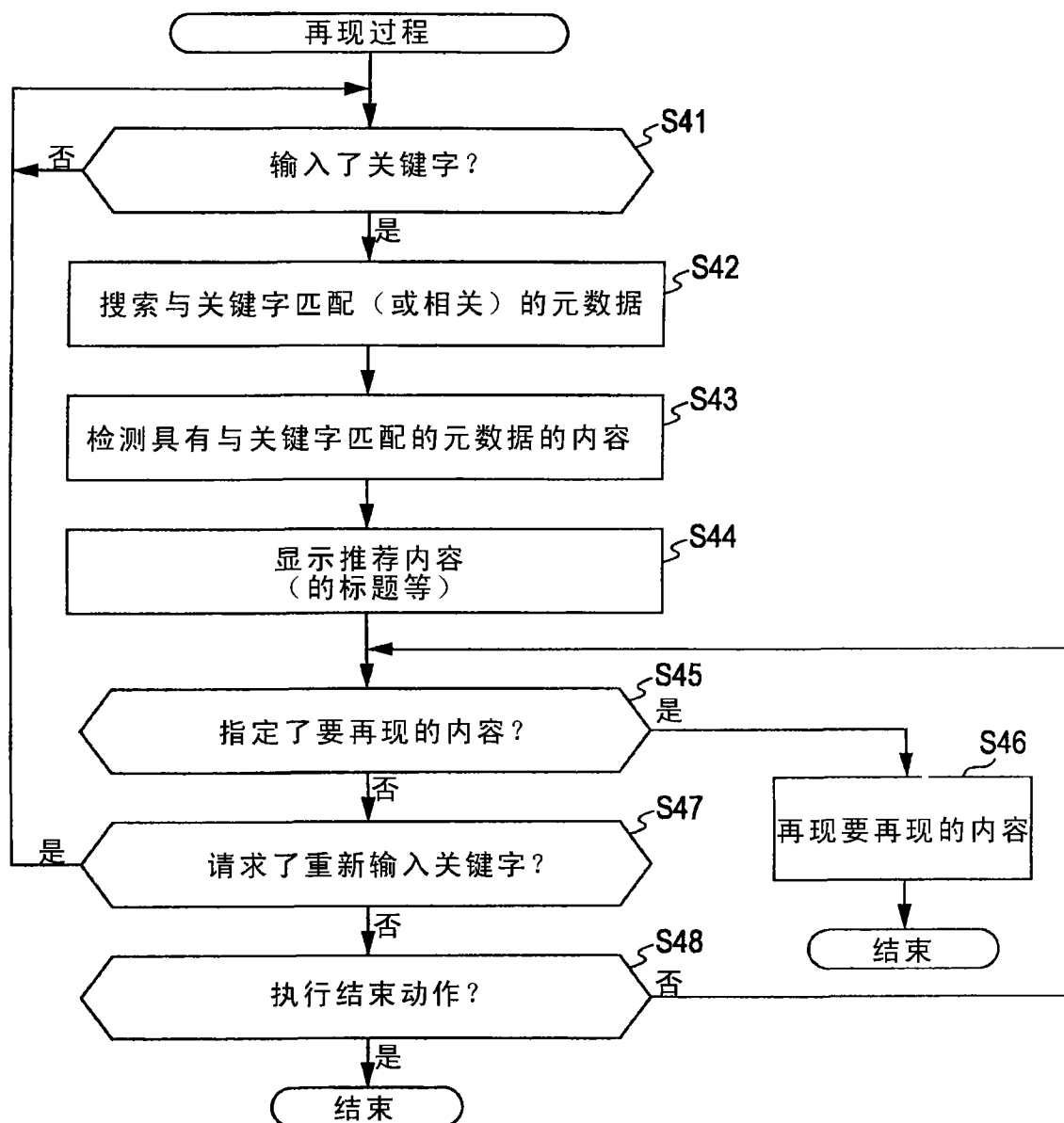


图 3

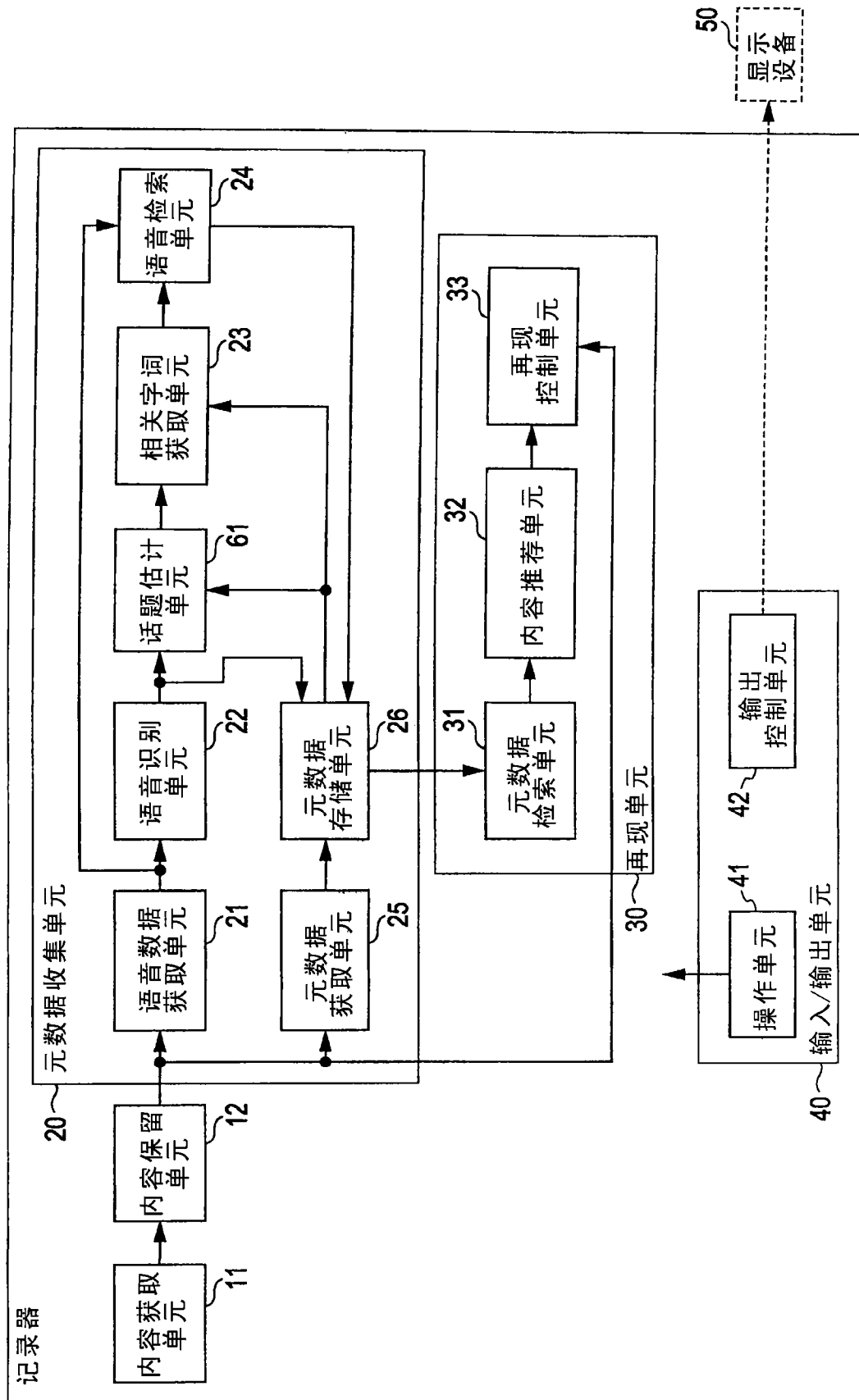


图 4

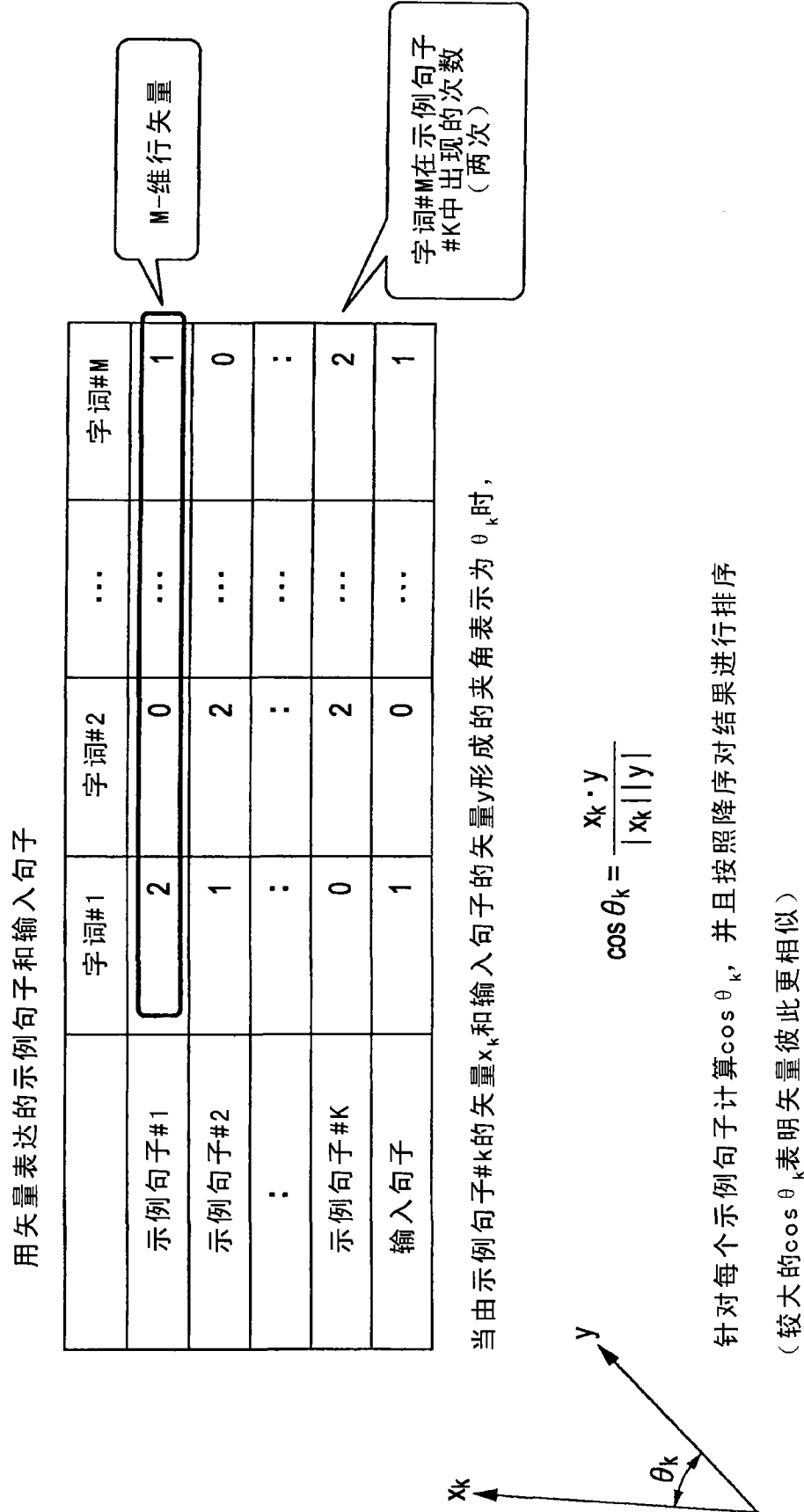


图 5

图 6A

文本#1: "A GRAND SLAM HOMER SMASHED IN THE LAST INNING HAS REVERSED THE GAME."

文本#2: "POWER RELATIONSHIP BETWEEN THE RULING AND OPPOSITION PARTIES HAS BEEN REVERSED IN THE DIET."

	LOVE	REVERSED	DIET	HOMER
文本#1	0,0	1,0	0,1	1,1
文本#2	0,0	1,0	1,1	0,1

图 6B

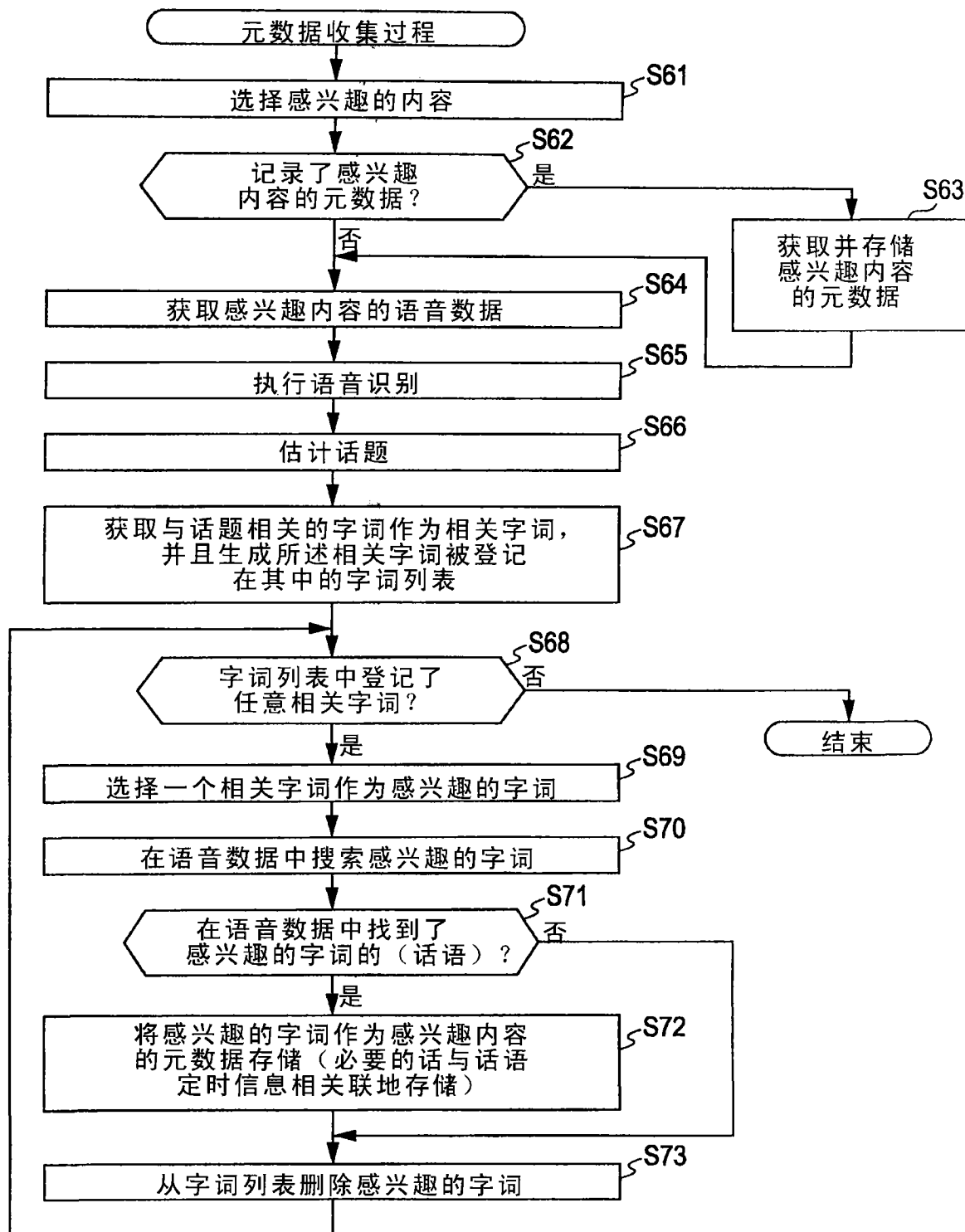


图 7

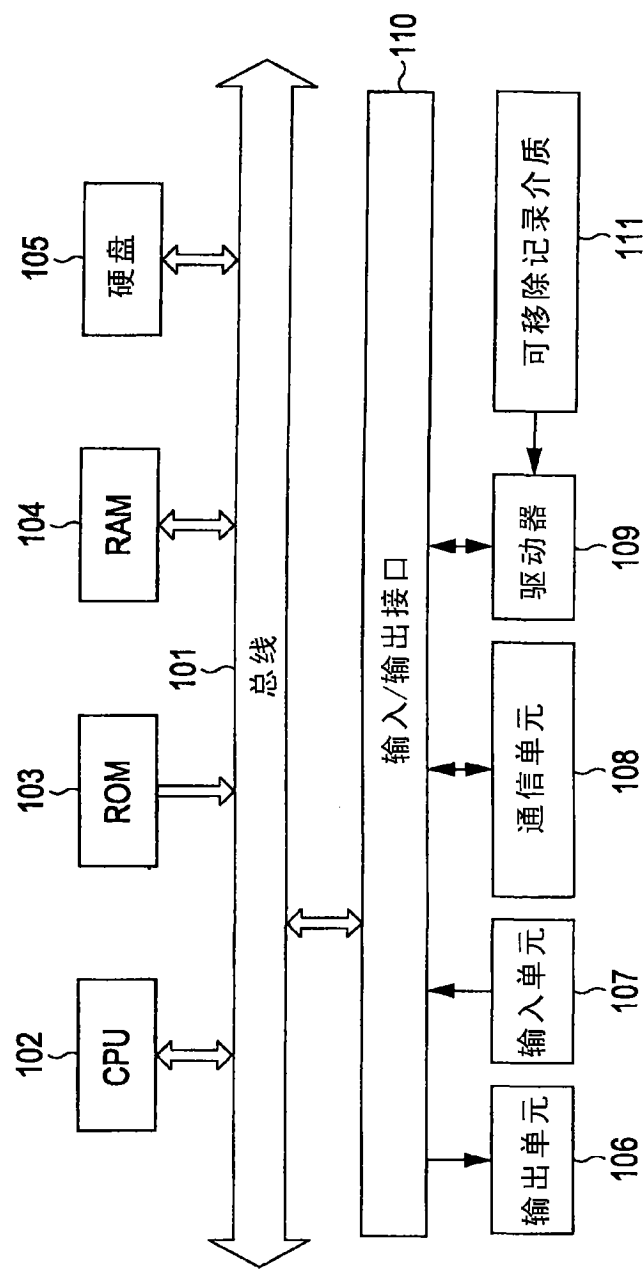


图 8