(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101606155 B
(45)授权公告日 2013.03.13

(21)申请号 200880004181.8
(22)申请日 2008.08.05
(30)优先权数据
208274/2007 2007.08.09 JP
(85)PCT申请进入国家阶段日
2009.08.05
(86)PCT申请的申请数据
PCT/JP2008/002111 2008.08.05
(87)PCT申请的公布数据
WO2009/019860 JA 2009.02.12
(73)专利权人 松下电器产业株式会社
地址 日本大阪府
(72)发明人 高田和丰 续木贵史 松浦聪
(74)专利代理机构 永新专利商标代理有限公司 72002
代理人 黄剑锋
(51)Int.Cl.
G06F 17/30(2006.01)
(56)对比文件
JP 特开平 10-134075 A,1998.05.22, 全文.
CN 1381799 A,2002.11.27, 全文.
US 2005/0289128 A1,2005.12.29, 全文.
审查员 田晶

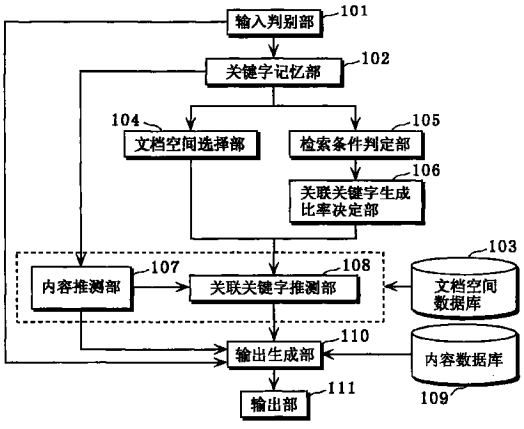
权利要求书 4 页 说明书 17 页 附图 11 页

(54)发明名称

内容检索装置

(57)摘要

内容检索装置(100),在用户的检索对象流动变化的情况下,也能够向该用户出示适当的关联关键字,内容检索装置(100)包括:内容推测部(107),检索按照检索关键字的内容;文档空间数据库(103),存储按照关键字的出现频度的多个文档空间;文档空间选择部(104),按照文档空间的出现频度,从文档空间数据库(103)中选择筛选文档空间以及扩展文档空间,所述文档空间示出检索关键字和对应于该检索关键字的内容的关联度;关联关键字推测部(108),将与筛选文档空间以及扩展文档空间分别相对应的关键字作为筛选关键字以及扩展关键字选择;以及输出部(111),限制所选择的筛选关键字以及扩展关键字。



CN 101606155 B

1. 一种内容检索装置，根据内容数据库，显示与用于检索内容的检索关键字有关的关联关键字，在所述内容数据库中多个内容的每一个与多个所述检索关键字和多个所述关联关键字一起被存储，所述内容检索装置包括：

关键字取得单元，取得检索关键字；

内容检索单元，从所述内容数据库中检索对应于所述检索关键字的内容；

文档空间存储单元，存储文档空间信息，所述文档空间信息是所述内容数据库中包含的各个关键字按照针对所述内容数据库的各个内容出现的出现频度而被分类的文档空间信息，并且，所述文档空间信息示出各个内容和各个关键字之间的关联度；

文档空间选择单元，以示出所述检索关键字和所述检索关键字所对应的内容之间的关联度的第一文档空间信息中的关键字的出现频度为基准，根据所述文档空间存储单元所存储的文档空间信息的出现频度的大小关系，选择第二文档空间信息，所述第二文档空间信息具有用于筛选作为检索对象的内容数量的关键字，所述第二文档空间信息还具有用于扩展作为检索对象的内容数量的关键字；以及

显示单元，显示所述文档空间选择单元所选择的第二文档空间信息中满足规定的基准的关键字，以分别作为关联关键字。

2. 如权利要求 1 所述的内容检索装置，

所述文档空间选择单元，具有第一文档空间选择单元，该第一文档空间选择单元从所述文档空间存储单元中选择与比所述第一文档空间信息的出现频度少的出现频度相对应的所述第二文档空间信息，以作为筛选文档空间信息；

所述显示单元，显示所述筛选文档空间信息中满足规定的基准的关键字，以作为所述关联关键字。

3. 如权利要求 2 所述的内容检索装置，

所述文档空间选择单元，还具有第二文档空间选择单元，该第二文档空间选择单元从所述文档空间存储单元中选择与所述第一文档空间信息的出现频度以上的出现频度相对应的所述第二文档空间信息，以作为扩展文档空间信息；

所述显示单元，显示作为所述筛选文档空间信息中满足规定的基准的关键字的筛选关键字、和作为所述扩展文档空间信息中满足规定的基准的关键字的扩展关键字，以作为所述关联关键字。

4. 如权利要求 3 所述的内容检索装置，

所述内容检索装置还包括关联关键字选择单元，该关联关键字选择单元从所述内容数据库中选择与所述筛选文档空间信息以及所述扩展文档空间信息分别相对应的筛选关键字以及扩展关键字，以便针对所述内容检索单元所检索的内容示出的关联度越大的关键字就越被选择；

所述显示单元，显示所述关联关键字选择单元所选择的筛选关键字以及扩展关键字，以作为所述关联关键字，并且，显示所述内容检索单元所检索的内容的属性。

5. 如权利要求 4 所述的内容检索装置，

所述文档空间存储单元，在将文档频度值作为所述出现频度利用来分类的各个关键字的每个分区，存储所述文档空间信息。

6. 如权利要求 4 所述的内容检索装置，

所述内容检索单元，

在用户选择所述显示单元所显示的筛选关键字，从而所述关键字取得单元取得所述筛选关键字，以作为所述检索关键字的情况下，判断为应该进行筛选检索，并检索与在所述检索关键字之前所述关键字取得单元所取得的其它的关键字以及所述检索关键字相对应的内容。

7. 如权利要求 6 所述的内容检索装置，

所述内容检索单元，

在用户选择所述显示单元所显示的扩展关键字，从而所述关键字取得单元取得所述扩展关键字，以作为所述检索关键字的情况下，判断为应该进行扩展检索，并检索只与所述检索关键字相对应的内容。

8. 如权利要求 7 所述的内容检索装置，

所述内容检索装置还包括检索条件判定单元和比率决定单元，

所述检索条件判定单元，在用户选择所述显示单元所显示的筛选关键字，从而所述关键字取得单元取得所述筛选关键字，以作为所述检索关键字的情况下，判定为由所述内容检索单元进行筛选检索，在用户选择所述显示单元所显示的扩展关键字，从而所述关键字取得单元取得所述扩展关键字，以作为所述检索关键字的情况下，判定为由所述内容检索单元进行扩展检索；以及

所述比率决定单元，按照所述检索条件判定单元所判定的结果，决定所述显示单元所显示的所述筛选关键字的数量和扩展关键字的数量的比率；

所述关联关键字选择单元，选择与所述比率决定单元所决定的比率相对应的数量的筛选关键字以及扩展关键字。

9. 如权利要求 8 所述的内容检索装置，

所述比率决定单元，在所述检索条件判定单元判定为进行筛选检索的情况下，决定比率，以便所述筛选关键字的数量多于所述扩展关键字的数量。

10. 如权利要求 8 所述的内容检索装置，

所述比率决定单元，在所述检索条件判定单元判定为进行扩展检索的情况下，决定比率，以便所述扩展关键字的数量多于所述筛选关键字的数量。

11. 如权利要求 4 所述的内容检索装置，

内容检索单元，判别与所述检索关键字相对应的内容群、和与在所述检索关键字之前所述关键字取得单元所取得的其它的关键字相对应的内容群之间重复的内容的数量是否在阈值以上，

在判别为阈值以上的情况下，检索与所述检索关键字以及所述其它的关键字相对应的内容，

在判别为小于阈值的情况下，检索只与所述检索关键字相对应的内容。

12. 如权利要求 4 所述的内容检索装置，

所述关联关键字选择单元，选择示出阈值以上的关联度的关键字，以作为所述筛选关键字或扩展关键字。

13. 如权利要求 4 所述的内容检索装置，

所述关联关键字选择单元，在选择多个筛选关键字或扩展关键字的情况下，且在与所

述多个筛选关键字或扩展关键字的每一个相对应的内容群中重复的内容的数量在阈值以上的情况下，通过结合所述多个筛选关键字或扩展关键字，从而生成结合关联关键字，

所述显示单元，不显示所述多个筛选关键字或扩展关键字，而显示所述结合关联关键字。

14. 如权利要求 4 所述的内容检索装置，

所述关联关键字选择单元，

从所述内容数据库中选择所述筛选关键字以及扩展关键字，以便在对所述内容检索单元所检索的内容示出的关联度越大、且在联想词典中示出的与所述检索关键字之间的关联性越强的关键字就越被选择，所述联想词典示出各个关键字之间的关联性。

15. 如权利要求 1 所述的内容检索装置，

所述文档空间选择单元，在筛选作为检索对象的内容数量的情况下，选择比所述第一文档空间信息中的关键字的出现频度少的文档空间信息，以作为具有用于筛选作为检索对象的内容数量的关键字的所述第二文档空间信息，在扩展作为检索对象的内容数量的情况下，选择所述第一文档空间信息中的关键字的出现频度以上的文档空间信息，以作为具有用于扩展作为检索对象的内容数量的关键字的所述第二文档空间信息。

16. 一种内容检索方法，根据内容数据库，显示与用于检索内容的检索关键字有关的关联关键字，在所述内容数据库中多个内容的每一个与多个所述检索关键字和多个所述关联关键字一起被存储，所述内容检索方法包括：

关键字取得步骤，取得检索关键字；

内容检索步骤，从所述内容数据库中检索对应于所述检索关键字的内容；

文档空间选择步骤，利用文档空间存储单元，以示出所述检索关键字和所述检索关键字所对应的内容之间的关联度的第一文档空间信息中的关键字的出现频度为基准，根据所述文档空间存储单元所存储的文档空间信息的出现频度的大小关系，选择第二文档空间信息，所述第二文档空间信息具有用于筛选作为检索对象的内容数量的关键字，所述第二文档空间信息还具有用于扩展作为检索对象的内容数量的关键字，所述文档空间存储单元存储文档空间信息，所述文档空间信息是所述内容数据库中包含的各个关键字按照针对所述内容数据库的各个内容出现的出现频度而被分类的文档空间信息，并且，所述文档空间信息示出各个内容和各个关键字之间的关联度；以及

显示步骤，显示在所述文档空间选择步骤所选择的第二文档空间信息中满足规定的基准的关键字，以分别作为关联关键字。

17. 一种内容检索装置，根据内容数据库，显示与用于检索内容的检索关键字有关的关联关键字，在所述内容数据库中多个内容的每一个与多个所述检索关键字和多个所述关联关键字一起被存储，所述内容检索装置包括：

关键字取得单元，取得检索关键字；

内容检索单元，从所述内容数据库中检索对应于所述检索关键字的内容；

文档空间存储单元，存储文档空间信息，所述文档空间信息是所述内容数据库中包含的各个关键字按照针对所述内容数据库的各个内容出现的出现频度而被分类的文档空间信息，并且，所述文档空间信息示出各个内容和各个关键字之间的关联度；以及

文档空间选择单元，以示出所述检索关键字和所述检索关键字所对应的内容之间的关

联度的第一文档空间信息中的关键字的出现频度为基准，根据所述文档空间存储单元所存储的文档空间信息的出现频度的大小关系，选择第二文档空间信息，所述第二文档空间信息具有用于筛选作为检索对象的内容数量的关键字，所述第二文档空间信息还具有用于扩展作为检索对象的内容数量的关键字。

内容检索装置

技术领域

[0001] 本发明涉及用于从多个内容中检索用户想要视听的内容的内容检索装置。

背景技术

[0002] 在以往的检索方法中所报告的较多的方法是，在从服务器等所存储的大量的内容中检索用户想要视听的内容时，生成并出示与用户输入的检索关键字一致的内容的一览。这些方法是，对能够明确估计作为检索对象的内容、且能够想起表示检索对象的关键字的用户有效的方法。

[0003] 但是，不能想起表示检索对象的关键字的用户，由于不能输入检索关键字，因此不能利用该方法来检索内容。这些状况是，例如在检索电视节目、音乐或电影等娱乐内容时常见的现象。

[0004] 对于这些问题，所报告的另一种内容检索技术是，系统方出示多个关键字，用户反复进行从所出示的关键字中选择感兴趣的关键字的操作，并进行内容的筛选，从而以对话方式获得所希望的内容的方法（例如，参照专利文献 1）。

[0005] 所述专利文献 1 的内容检索装置（文档处理装置），在检索开始时，同时取得检索关键字和检索条件，限定检索领域，并出示与该检索关键字有关的关联关键字。在该装置中，通过在检索开始时筛选检索领域，从而能够出示易于筛选内容的关联关键字。因此，用户只要反复选择该内容检索装置出示的关联关键字，就能够获得所希望的内容。该方法，有效于表示检索对象的关键字不明确的用户。

[0006] 专利文献 1 ：日本国特开平 10-134075 号公报

[0007] 然而，在所述专利文献 1 的内容检索装置中存在的问题是，在检索开始时还未明确决定用户的检索对象、且正在检索中想要看的内容流动变化的情况下，不能出示用于检索所希望的内容的关联关键字。

[0008] 也就是，在所述专利文献 1 的内容检索装置中，针对一系列的检索，由于只能进行根据某个观点的筛选检索，因此，在作为用户的检索对象的内容流动变化、且用户要看各种内容时，为了重新进行根据其它的观点的检索，而需要每次返回到检索开始画面，从而导致用户的检索负荷增大。

[0009] 例如，在所述专利文献 1 的内容检索装置中，针对检索对象不明确且想要看的内容流动变化的用户，不能选择通过对话来检索不同观点的各种内容的检索模式（扩展检索模式）、和进一步对显示为检索结果的内容进行筛选的检索模式（筛选检索模式）。

发明内容

[0010] 于是，鉴于所述问题，本发明的目的在于提供一种内容检索装置，在作为用户的检索对象的内容流动变化的情况下，也能够向用户出示用于检索所希望的内容的关联关键字。

[0011] 为了实现所述目的，本发明涉及的内容检索装置，根据内容数据库，显示与用于检

索内容的检索关键字有关的关联关键字，在所述内容数据库中多个内容的每一个与多个关键字一起被存储，所述内容检索装置包括：关键字取得单元，取得检索关键字；内容检索单元，从所述内容数据库中检索对应于所述检索关键字的内容；文档空间存储单元，存储文档空间信息，所述文档空间信息是所述内容数据库中包含的各个关键字按照针对所述内容数据库的各个内容出现的出现频度而被分类的文档空间信息，并且，所述文档空间信息示出各个内容和各个关键字之间的关联度；文档空间选择单元，以示出所述检索关键字和所述检索关键字所对应的内容之间的关联度的第一文档空间信息中的关键字的出现频度为基准，根据所述文档空间存储单元所存储的文档空间信息的出现频度的大小关系，选择第二文档空间信息，所述第二文档空间信息具有用于筛选作为检索对象的内容数量的关键字，所述第二文档空间信息还具有用于扩展作为检索对象的内容数量的关键字；以及显示单元，显示所述文档空间选择单元所选择的第二文档空间信息中满足规定的基准的关键字，以分别作为关联关键字。

[0012] 并且，为了实现所述目的，本发明的内容检索装置，根据内容数据库，显示与用于检索内容的检索关键字有关的关联关键字，在所述内容数据库中多个内容的每一个与多个关键字一起被存储，所述内容检索装置包括：关键字取得单元，取得检索关键字；内容检索单元，从所述内容数据库中检索对应于所述检索关键字的内容；文档空间存储单元，存储文档空间信息，该文档空间信息，所述内容数据库中包含的各个关键字按照针对所述内容数据库的各个内容出现的出现频度而被分类的每个分区，示出各个内容和各个关键字之间的关联度；文档空间选择单元，按照第一文档空间信息的出现频度，从所述文档空间存储单元中选择第二文档空间信息，所述第一文档空间信息示出所述检索关键字和所述检索关键字所对应的内容之间的关联度；以及显示单元，显示所述文档空间选择单元所选择的第二文档空间信息中满足规定的基准的关键字，以作为关联关键字。

[0013] 据此，在作为用户的检索对象的内容流动变化的情况下，也能够向用户出示用于检索所希望的内容的关联关键字。

[0014] 例如，所述文档空间选择单元，具有：第一文档空间选择单元，从所述文档空间存储单元中选择与比所述第一文档空间信息的出现频度少的出现频度相对应的第二文档空间信息，以作为筛选文档空间信息；所述显示单元，显示所述筛选空间信息中满足规定的基准的关键字，以作为所述关联关键字。并且，所述文档空间选择单元，具有：第二文档空间选择单元，从所述文档空间存储单元中选择与所述第一文档空间信息的出现频度以上的出现频度相对应的第二文档空间信息，以作为扩展文档空间信息；所述显示单元，显示作为所述筛选空间信息中满足规定的基准的关键字的筛选关键字、和作为所述扩展空间信息中满足规定的基准的关键字的扩展关键字，以作为所述关联关键字。并且，所述内容检索装置，还包括：关联关键字选择单元，从所述内容数据库中选择与所述筛选文档空间信息以及所述扩展文档空间信息分别相对应的筛选关键字以及扩展关键字，以便针对所述内容检索单元所检索的内容示出的关联度越大的关键字就越被选择；所述显示单元，显示所述关联关键字选择单元所选择的筛选关键字以及扩展关键字，以作为所述关联关键字，并且，显示所述内容检索单元所检索的内容的属性。

[0015] 具体而言，文档空间存储单元存储有，出现频度多的关键字所属的文档空间信息、出现频度中等程度的关键字所属的文档空间信息、以及出现频度少的关键字所属的文档空

间信息。

[0016] 据此，选择比检索关键字所对应的文档空间频度信息的出现频度小的出现频度所对应的文档空间信息，以作为筛选文档空间信息。例如，若检索关键字的出现频度为中等程度，则选择出现频度少的关键字所属的文档空间信息，以作为筛选文档空间信息。而且，针对内容检索单元所检索的内容，显示依据该筛选文档空间信息示出大的关联度的关键字，以作为筛选关键字。

[0017] 其结果为，由于比检索关键字的出现频度少的出现频度的关键字被显示为筛选关键字，因此，若用户选择该筛选关键字，以作为新的下个检索关键字，则能够通过内容检索单元，从预先检索并显示了的多个内容中适当地筛选用户所希望的内容。

[0018] 据此，选择检索关键字所对应的文档空间频度信息的出现频度以上的出现频度所对应的文档空间信息，以作为扩展文档空间信息。例如，若检索关键字的出现频度为中等程度，则选择出现频度为中等程度的关键字所属的文档空间信息、和出现频度多的关键字所属的文档空间信息，以作为扩展文档空间信息。而且，针对内容检索单元所检索的内容，显示依据该扩展文档空间信息示出大的关联度的关键字，以作为筛选关键字。

[0019] 其结果为，由于检索关键字的出现频度以上的出现频度的关键字被显示为扩展关键字，因此，若用户选择该扩展关键字，以作为新的下个检索关键字，则能够通过内容检索单元，适当地检索预先检索并显示了的内容的数量以上的、与检索关键字有关的内容。

[0020] 据此，由于同时显示这些筛选关键字以及扩展关键字，以作为关联关键字，因此，在作为用户的检索对象的内容流动变化的情况下，也不像以往的例子那样返回到检索开始画面，而通过筛选关键字以及扩展关键字之中的任何一方，从而用户能够检索所希望的内容。

[0021] 也就是，在以往的技术中，为了筛选预先检索了的多个内容而适合的关联关键字（筛选关键字）、和为了以与检索关键字不同的观点进行再检索而适合的关联关键字（扩展关键字）不被区别，按照从关联度大到小的顺序，依次单纯地显示与检索关键字有关的关联关键字。其结果为，有筛选关键字不被显示的情况，或者，有扩展关键字不被显示的情况。因此，虽不是筛选关键字，也依据该关键字筛选并检索多个内容。

[0022] 于是，在本发明中，如上所述，由于按照出现频度适当地区别筛选关键字以及扩展关键字，并将它们同时显示，因此用户能够容易且适当地检索所希望的内容。

[0023] 并且，也可以是，所述内容检索装置，其中还包括：检索条件判定单元，在用户选择所述显示单元所显示的筛选关键字，从而所述关键字取得单元取得所述筛选关键字，以作为所述检索关键字的情况下，判定为由所述内容检索单元进行筛选检索，在用户选择所述显示单元所显示的扩展关键字，从而所述关键字取得单元取得所述扩展关键字，以作为所述检索关键字的情况下，判定为由所述内容检索单元进行扩展检索；以及比率决定单元，按照所述检索条件判定单元所判定的结果，决定所述显示单元所显示的所述筛选关键字的数量和扩展关键字的数量的比率；所述关联关键字选择单元，选择与所述比率决定单元所决定的比率相对应的数量的筛选关键字以及扩展关键字。例如，所述比率决定单元，在所述检索条件判定单元判定为进行筛选检索的情况下，决定比率，以便所述筛选关键字的数量多于所述扩展关键字的数量。并且，所述比率决定单元，在所述检索条件判定单元判定为进行扩展检索的情况下，决定比率，以便所述扩展关键字的数量多于所述筛选关键字的数量。

[0024] 据此，若用户选择显示单元所显示的筛选关键字，以作为新的下个检索关键字，则显示与该新的检索关键字相对应的内容，并且显示比扩展关键字多的、与该新的检索关键字有关的新的筛选关键字。据此，由于向将执行筛选检索的用户显示多个筛选关键字，因此能够提高针对用户的使用上的方便性。与此相同，若用户选择显示单元所显示的扩展关键字，以作为新的下个检索关键字，则显示与该新的检索关键字相对应的内容，并且显示比筛选关键字多的、与该新的检索关键字有关的新的扩展关键字。据此，由于向将执行扩展检索的用户显示多个扩展关键字，因此能够提高针对用户的使用上的方便性。

[0025] 而且，本发明，除了能够以这些内容检索装置来实现以外，还能够以通过该装置检索内容的方法、用于使该装置检索内容的程序、或者存储该程序的计算机可读的存储介质、集成电路来实现。

[0026] 本发明的内容检索装置具有以下的功能效果，即，在作为用户的检索对象的内容流动变化的情况下，也能够向用户出示用于检索所希望的内容的关联关键字。

附图说明

[0027] 图 1 是示出本发明的实施例中的内容检索装置的结构的方框图。

[0028] 图 2 是示出本发明的实施例中的内容检索装置在检索开始时显示的画面的一个例子的图。

[0029] 图 3 是示出本发明的实施例中的内容检索装置在初始检索画面后显示的画面的一个例子的图。

[0030] 图 4 是示出本发明的实施例中的内容检索装置在检索结果画面后显示的其它的检索结果画面的一个例子的图。

[0031] 图 5 是示出本发明的实施例中的内容数据库所蓄积的多个内容属性信息的一个例子的图。

[0032] 图 6 是示出本发明的实施例中的文档空间数据库所存储的多个文档空间的图。

[0033] 图 7 是示出本发明的实施例中的关键字记忆部所存储的信息的图。

[0034] 图 8 是示出本发明的实施例中的内容推测部的工作的流程图。

[0035] 图 9 是用于说明本发明的实施例中的关联关键字推测部确定关联关键字的候补的工作的说明图。

[0036] 图 10 是用于具体说明本发明的实施例中的关联关键字推测部推测关联关键字的工作的说明图。

[0037] 图 11 是示出本发明的实施例中的关联关键字推测部的工作的流程图。

[0038] 图 12 是示出本发明的实施例中的输出部所显示的检索结果画面的一个例子的图。

[0039] 图 13 是示出本发明的实施例中的内容检索装置整体的工作的流程图。

[0040] 图 14 是示出本发明的实施例中的联想词典的一个例子的图。

[0041] 图 15 是示出本发明的实施例中的内容检索装置的概观的一个例子的图。

[0042] 符号说明

[0043] 100 内容检索装置

[0044] 101 输入判别部

[0045] 102 关键字记忆部

[0046] 103 文档空间数据库

[0047] 104 文档空间选择部

[0048] 105 检索条件判定部

[0049] 106 关联关键字生成比率决定部

[0050] 107 内容推测部

[0051] 108 关联关键字推测部

[0052] 109 内容数据库

[0053] 110 输出生成部

[0054] 111 输出部

具体实施方式

[0055] 以下，参照附图说明本发明的实施例中的内容检索装置。而且，在本实施例中说明一种内容检索装置，从蓄积有多个内容的服务器等中检索用户想要视听的内容，该多个内容是从过去到现在为止的电视节目。

[0056] 图 1 是示出本实施例中的内容检索装置的结构的方框图。

[0057] 本实施例中的内容检索装置 100 是一种装置，在作为用户的检索对象的内容流动变化的情况下，也能够向用户出示用于检索所希望的内容的关联关键字，内容检索装置 100 包括输入判别部 101、关键字记忆部 102、文档空间数据库 103、文档空间选择部 104、检索条件判定部 105、关联关键字生成比率决定部 106、内容推测部 107、关联关键字推测部 108、内容数据库 109、输出生成部 110、以及输出部 111。

[0058] 输入判别部 101，接受输出部 111 所显示的多个选择项中的、通过用户的操作所选择的选择项（以下，称为输入选择项），判别该输入选择项示出关键字、还是示出内容。输入判别部 101，若判别为示出关键字，则将该输入选择项输出到关键字记忆部 102，若判别为示出内容，则将该输入选择项输出到输出生成部 110。在此，在输入判别部 101 最近接受了的输入选择项示出关键字的情况下，以下，将该关键字称为检索关键字。

[0059] 关键字记忆部 102，具有用于记忆从输入判别部 101 输出的输入选择项示出的关键字的区域。

[0060] 检索条件判定部 105，判定由关键字记忆部 102 所存储的检索关键字的检索的模式（检索模式）是筛选模式还是扩展模式。在此，筛选模式是指，组合检索关键字、和从检索开始时到现在为止所输入的过去的关键字，来对内容进行筛选来检索（筛选检索）的模式。另一方面，扩展模式是指，与过去的关键字无关，而进行只符合检索关键字的内容的检索（扩展检索）的模式。

[0061] 关联关键字生成比率决定部 106，按照检索条件判定部 105 所判定的结果，决定应该显示的筛选关键字的数量和扩展关键字的数量的比率（显示比率）。筛选关键字是，能够成为用于执行筛选检索的下一个检索关键字的关键字，也是用于在由过去所输入的关键字所检索的内容内进一步进行细分化来筛选的关键字。另一方面，扩展关键字是，能够成为用于执行扩展检索的下一个检索关键字的关键字，也是用于由关联度与过去所输入的关键字相近的新的关键字重新检索内容的关键字。并且，将这些筛选关键字以及扩展关键字总称

为关联关键字。

[0062] 内容数据库 109，依次存储正在广播中的电视节目等的多个内容。在此，内容包含示出电视节目的运动图像或声音的内容数据、和如电视节目的 EPG 信息那样示出该电视节目的属性的内容属性信息。

[0063] 文档空间数据库 103 存储示出内容数据库 109 所示的各个内容、和各个关键字之间的关系的多个文档空间。而且，依据空间 ID 来识别文档空间数据库 103 所存储的多个文档空间的每一个。

[0064] 文档空间选择部 104，利用检索关键字从文档空间数据库 103 中选择为了推测筛选关键字而应该使用的文档空间（以下，称为筛选文档空间）、和为了推测扩展关键字而应该使用的文档空间（以下，称为扩展文档空间）。而且，文档空间选择部 104，将所选择的筛选文档空间的空间 ID、和所选择的扩展文档空间的空间 ID 输出到关联关键字推测部 108。

[0065] 内容推测部 107，利用关键字记忆部 102 所记忆的关键字、和文档空间数据库 103 所存储的多个文档空间，推测与检索关键字有关的内容。而且，内容推测部 107，将用于识别所推测的内容的内容 ID 输出到关联关键字推测部 108 以及输出生成部 110。

[0066] 关联关键字推测部 108，取得从内容推测部 107 输出的内容 ID、关联关键字生成比率决定部 106 所决定的比率、和从文档空间选择部 104 输出的空间 ID，并且，利用它们和文档空间数据库 103 推测筛选关键字以及扩展关键字，以作为关联关键字。关联关键字推测部 108，将用于识别所推测的关联关键字的关键字 ID 输出到输出生成部 110。

[0067] 输出生成部 110，若从输入判别部 101 取得示出内容的输入选择项，则从内容数据库 109 中抽出对应于该输入选择项的内容数据，并输出到输出部 111。

[0068] 并且，输出生成部 110，若从内容推测部 107 取得内容 ID，并且从关联关键字推测部 108 取得关键字 ID，则从内容数据库 109 中抽出对应于该内容 ID 的内容属性信息、和对应于该关键字 ID 的关联关键字。而且，输出生成部 110，将包含该内容属性信息和关联关键字的检索结果信息输出到输出部 111。

[0069] 输出部 111，具备例如液晶显示器等，若从输出生成部 110 取得内容数据，则再生该内容数据。另一方面，输出部 111，若从输出生成部 110 取得检索结果信息，则显示该检索结果信息的细节。

[0070] 而且，在本实施例中：输入判别部 101 被构成为取得单元，取得检索关键字；内容推测部 107 被构成为内容检索单元，从内容数据库 109 中检索对应于检索关键字的内容；输出部 111 被构成为显示单元，用于显示关联关键字推测部 108 所推测的筛选关键字以及扩展关键字，以作为关联关键字，并且显示内容推测部 107 所推测的内容的属性。

[0071] 在此，对于这些内容检索装置 100 整体的工作，利用该内容检索装置 100 所显示的画面进行说明。

[0072] 图 2 是示出内容检索装置 100 在检索开始时显示的画面的一个例子的图。

[0073] 例如，内容检索装置 100，在检索开始时，在输出部 111 显示初始检索画面。在该初始检索画面显示，作为选择项的“松下花子”、“日本”或“篮球”等，以作为初始检索关键字。在此情况下，输出生成部 110，使输出部 111 显示例如预先设定的“松下花子”或“日本”等的单词，以作为初始检索关键字。而且，输出生成部 110 也可以，使输出部 111 显示“体育”或“实录”等的示出一般类型的单词，以作为初始检索关键字。并且，输出生成部 110 也可

以，利用输出到输出部 111 的内容数据的履历，即，利用用户的视听履历，使输出部 111 显示用户的视听频度多的内容的内容属性信息中包含的单词，以作为初始检索关键字。

[0074] 在此，用户，从输出部 111 的初始检索画面所显示的作为选择项的初始检索关键字中，选择例如初始检索关键字“松下花子”。因此，内容检索装置 100，显示包含与初始检索关键字“松下花子”有关的内容属性信息、和与初始检索关键字“松下花子”有关的关键字（关联关键字）的检索结果信息的细节。

[0075] 图 3 是示出内容检索装置 100 在初始检索画面后显示的画面的一个例子的图。

[0076] 内容检索装置 100，如上所述，若初始检索画面中的任一个初始检索关键字被选择为输入选择项，则检索与该初始检索关键字有关的内容或关联关键字，在输出部 111 显示对应于该初始检索关键字的检索结果画面。在该检索结果画面显示 ：示出与所选择的初始检索关键字有关的内容属性信息（内容的名称或概要）的一览的内容列表 L1 ；示出与所选择的初始检索关键字有关的筛选关键字的一览的筛选列表 L2 ；示出与所选择的初始检索关键字有关的扩展关键字的一览的扩展列表 L3。而且，这些内容属性信息、筛选关键字以及扩展关键字，分别被显示为新的选择项。

[0077] 例如，如上所述，若在初始检索画面选择初始检索关键字“松下花子”，则在内容列表 L1 显示与“松下花子”有关的多个内容的名称以及其概要，在筛选列表 L2 显示与“松下花子”有关的多个筛选关键字，在扩展列表 L3 显示与“松下花子”有关的多个扩展关键字。

[0078] 在此，用户，若在输出部 111 的检索结果画面的内容列表 L1 中有想要视听的内容的名称等，则选择该内容的名称等。因此，内容检索装置 100，从内容数据库 109 中抽出与所选择的内容的名称（示出内容的输入选择项）相对应的内容数据，并进行再生，从而使输出部 111 进行显示。据此，内容的检索结束。

[0079] 另一方面，用户，若在输出部 111 的检索结果画面的内容列表 L1 中没有想要视听的内容的名称等，则从筛选列表 L2 或扩展列表 L3 中选择有可能与想要视听的内容有关的关联关键字。因此，内容检索装置 100，将所选择的关联关键字作为下个检索关键字（输入选择项），检索与该检索关键字有关的内容或关联关键字，从而在输出部 111 显示与该检索关键字有关的检索结果画面。

[0080] 图 4 是示出内容检索装置 100 在检索结果画面后显示的其它的检索结果画面的一个例子的图。

[0081] 例如，若从图 3 示出的检索结果画面的筛选列表 L2 中选择筛选关键字“京都”，则在内容列表 L1 显示与“京都”有关的多个内容的名称以及其概要，在筛选列表 L2 显示与“京都”有关的多个筛选关键字，在扩展列表 L3 显示与“京都”有关的多个扩展关键字。

[0082] 如此，内容检索装置 100，在每次由用户选择关联关键字（筛选关键字或扩展关键字）时，将该关联关键字作为新的检索关键字，反复检索并显示与该检索关键字有关的新的内容以及关联关键字。

[0083] 在此，对于内容检索装置 100 的各个构成要素，进行详细说明。

[0084] 图 5 是示出内容数据库 109 所蓄积的多个内容属性信息的一个例子的图。

[0085] 内容属性信息 CA 中包含例如用于识别内容的内容 ID、作为内容的名称的内容名、内容的视听开始时刻、对应于内容的关键字、以及内容的概要。

[0086] 而且，依据像 EPG 那样的、被记载为内容数据的属性的描述或标题等，进行形态分

析等，从而抽出内容属性信息 CA 中包含的关键字。并且，内容属性信息 CA 中包含的各个关键字，被附上用于识别这些关键字的关键字 ID。

[0087] 图 6 是示出文档空间数据库 103 所存储的多个文档空间的图。

[0088] 例如，文档空间数据库 103，存储以空间 ID“1、2、3”所示的三个文档空间。

[0089] 各个文档空间，以矩阵来示出所述内容属性信息 CA 中包含的多个关键字（关键字 ID）的每一个、和内容数据库 109 所示的所有的内容（内容 ID）的每一个之间的关联度。例如，将 TF(Term Frequency)-IDF(Inverse Document Frequency) 值等用于关联度。

[0090] TF-IDF 值是指，组合了针对文档（内容）的关键字的网罗性或特定性的关键字的加权的指标，例如，在“情报检索と言语处理（信息检索和语言处理）”(pp. 32 ～ 33，东京大学出版会，1999)（非专利文献 1）记载有具体算出方法。

[0091] 例如，也可以以隐性语义索引 (Latent Semantic Indexing；：以下称为 LSI) 来将 TF-IDF 值等的值抽象化。在利用 LSI 的情况下，在高维空间分别处理的关键字，在低维空间有可能相关联，因此，能够进行基于关键字具有的语义或概念的检索。例如，棒球和 Baseball 是不同的关键字，依据一方的关键字的询问不能检索包含另一方的关键字的文档。但是，在低维空间，由于可以期待在语义上相关联的这些关键字被简并为一个维，因此通过输入称为棒球的关键字，从而也能够检索包含 Baseball 的节目。在“情报检索アルゴリズム（信息检索算法）”(pp. 65 ～ 77，共立出版，2002)（非专利文献 2）记载有利用 LSI 的文档检索方法。

[0092] 而且，在图 6 示出的例子中，针对内容（内容属性信息 CA）中出现的关键字示出关联度“1”，针对内容中不出现的关键字示出关联度“0”。例如，空间 ID“1”的文档空间，将关键字 ID“KW1a”的关键字针对内容 ID“C1、C2、…Ck”的关联度示出为“1、1、…0”。

[0093] 并且，三个文档空间，依据关键字的频度信息被区分。关键字的频度信息是指，在内容数据库 109 的各个内容中出现该关键字的出现频度。例如，利用作为与针对关键字的内容数量关联起来的指标的 DF 值 (Document Frequency：文档频度（关键字出现的文档数））等，以作为频度信息。因此，较多的内容（内容属性信息 CA）中出现的关键字的频度信息示出的值大，较少的内容（内容属性信息 CA）中出现的关键字的频度信息示出的值小。

[0094] 例如，空间 ID“1”的文档空间示出频度信息示出的值大的各个关键字和各个内容之间的关联度；空间 ID“2”的文档空间示出频度信息示出的值为中等程度的各个关键字和各个内容之间的关联度；空间 ID“3”的文档空间示出频度信息示出的值小的各个关键字和各个内容之间的关联度。

[0095] 也就是，关键字 ID“KW1a、KW2a、…KWna”的各个关键字，由于出现在较多的内容（内容属性信息 CA）中，因此被分类为空间 ID“1”的文档空间；关键字 ID“KW1b、KW2b、…KWmb”的各个关键字，由于出现在中等程度的数量的内容（内容属性信息 CA）中，因此被分类为空间 ID“2”的文档空间；关键字 ID“KW1c、KW2c、…KW1c”的各个关键字，由于出现在较少的内容（内容属性信息 CA）中，因此被分类为空间 ID“3”的文档空间。

[0096] 而且，按照内容数据库的更新进行文档空间的更新。因此，因内容数据库的更新而发生空间 ID 和关键字（关键字 ID）之间的对应关系的变化。例如，称为“滑雪”的关键字，在冬天出现的情况多，因此，在冬天，该关键字（关键字 ID）与高位的空间 ID 相对应，即，与频度信息示出的值大的文档空间的空间 ID 相对应。反而，在夏天，称为“滑雪”的关键字的

出现频度变少，因此，该关键字（关键字 ID）与低位的空间 ID 相对应，即，与频度信息示出的值小的文档空间的空间 ID 相对应。

[0097] 据此，按照时期的变化，能够准确地使关键字和用户想起的节目数对应起来，也能够出示没有不协调感的筛选关键字以及扩展关键字。

[0098] 本实施例的特点是，如上所述，按照关键字的出现频度，将示出内容数据库 109 中包含的各个内容和各个关键字之间的关系的文档空间分割为多个。

[0099] 输入判别部 101，若接受如图 3 示出的检索结果画面的内容列表 L1 所显示的输入选择项，则判别为该输入选择项示出内容，并将作为该输入选择项的内容 ID 输出到输出生成部 110。

[0100] 另一方面，输入判别部 101，若接受如图 3 示出的检索结果画面的筛选列表 L2 或扩展列表 L3 所显示的输入选择项，则判别为该输入选择项示出关键字，并将作为该输入选择项的关键字 ID、和示出该关键字 ID 的关键字存储到关键字记忆部 102。此时，输入判别部 101，将检索模式标志和检索步骤，与关键字 ID 以及关键字一起存储到关键字记忆部 102，所述检索模式标志示出如图 3 示出的筛选列表 L2 以及扩展列表 L3 中的哪一方中选择了输入选择项，所述检索步骤示出该选择是从检索开始时第几次的选择。

[0101] 图 7 是示出关键字记忆部 102 所存储的信息的图。

[0102] 例如，输入判别部 101，若接受示出如图 2 示出的初始检索画面上的多个选择项中所选择的初始检索关键字“松下花子”的输入选择项，则将其关键字“松下花子”、其关键字 ID“10”以及示出第一次的选择的检索步骤“1”关联起来，存储到关键字记忆部 102。而且，由于此时的输入选择项，不包含在筛选列表 L2 以及扩展列表 L3 的两者中，因此，输入判别部 101 将检索模式标志“初始”与所述关键字关联起来存储。

[0103] 其次，例如，输入判别部 101，若接受示出如图 3 示出的检索结果画面上的多个选择项中所选择的关键字“京都”的输入选择项，则将其关键字“京都”、其关键字 ID “23”、示出第二次的选择的检索步骤“2”、以及示出从筛选列表 L2 中所选择的事宜的检索模式标志“筛选”关联起来，存储到关键字记忆部 102。

[0104] 而且，输入判别部 101，也可以将输入选择项示出的关键字所属的文档空间的空间 ID，与该关键字关联起来，存储到关键字记忆部 102。

[0105] 以下，将由如此在关键字记忆部 102 相关链并存储的检索步骤、关键字、关键字 ID 以及检索模式标志构成的信息，称为关键字信息。

[0106] 内容推测部 107，利用如图 7 示出关键字记忆部 102 所记忆的关键字信息、和如图 6 示出文档空间数据库 103 所存储的多个文档空间，推测与检索关键字有关的内容。

[0107] 在此，内容推测部 107，按照检索关键字的检索模式标志，进行依据筛选模式的内容的推测、和依据扩展模式的内容的推测。

[0108] 图 8 是示出内容推测部 107 的工作的流程图。

[0109] 首先，内容推测部 107，若检索关键字的关键字信息被存储到关键字记忆部 102，则从关键字记忆部 102 取得该关键字信息（步骤 S201）。

[0110] 例如，若关键字记忆部 102 存储作为第三次所选择的关键字“美食”等的检索关键字，内容推测部 107，则从关键字记忆部 102 取得示出关键字“美食”、关键字 ID “54”以及检索模式标志“扩展”的关键字信息。

[0111] 其次，内容推测部 107，根据在步骤 S201 取得的关键字信息中包含的检索模式标志，判别该关键字信息中包含的检索关键字是否为筛选关键字（步骤 S202）。也就是，内容推测部 107，在检索模式标志示出“筛选”时，判别检索关键字为筛选关键字，在检索模式标志示出“扩展”时，判别检索关键字为扩展关键字。

[0112] 在此，内容推测部 107，若判别检索关键字为筛选关键字（步骤 S202 的“是”），进一步，则将直到检索模式标志示出“扩展”或“初始”的最近的关键字被存储为止被存储到关键字记忆部 102 的过去的关键字，从该关键字记忆部 102 中抽出（步骤 S203）。也就是，内容推测部 107，在步骤 S201 以及步骤 S203，取得最近的一个筛选关键字或连续的多个筛选关键字、和一个初始检索关键字或最近的一个扩展关键字。

[0113] 内容推测部 107，在判别检索关键字为扩展关键字时（步骤 S202 的“否”），或者，在步骤 S203 后，从文档空间数据库 103 中选择示出检索关键字或所述过去的关键字的关联度的文档空间（步骤 S205）。

[0114] 例如，内容推测部 107，若在步骤 S201 将关键字 ID“KW2c”的关键字作为检索关键字取得，则选择与该关键字 ID“KW2c”的关键字相对应的空间 ID“3”的文档空间。并且，内容推测部 107，若在步骤 S203 取得关键字 ID“KW2b”的过去的关键字，则选择与关键字 ID“KW2b”的关键字相对应的空间 ID“2”的文档空间。

[0115] 其次，内容推测部 107，在检索关键字为筛选关键字时，生成对应于筛选关键字的输入关键字矢量；在检索关键字为扩展关键字时，生成对应于扩展关键字的输入关键字矢量（步骤 S206）。

[0116] 也就是，在检索关键字为筛选关键字时，内容推测部 107，按照被分类为在步骤 S205 选择的各个文档空间的各个关键字，是否在步骤 S201 以及步骤 S203 已被取得，生成输入关键字矢量。

[0117] 例如，在步骤 S201 取得的检索关键字（筛选关键字）的关键字 ID 为“KW2c”、且在步骤 S203 取得的过去的关键字的关键字 ID 为“KW2b”的情况下，在步骤 S205，与检索关键字相对应的空间 ID“3”的文档空间、和与过去的关键字相对应的空间 ID“2”的文档空间被选择。

[0118] 在此情况下，被分类为空间 ID“2、3”的文档空间的关键字（关键字 ID ＝ KW1b、KW2b、…、KWmb、KW1c、KW2c、…、KW1c）中的关键字 ID“KW2c、KW2b”的关键字被取得。因此，内容推测部 107，生成输入关键字矢量（KW1b，、KW2b、…、KWmb、KW1c、KW2c、…、KW1c）＝（0、1、…、0、0、1、…、0），以便针对所取得的关键字示出“1”，针对未被取得的关键字示出“0”。

[0119] 另一方面，在检索关键字为扩展关键字时，内容推测部 107，按照被分类为在步骤 S205 选择的各个文档空间的各个关键字是否在步骤 S201 已被取得，生成输入关键字矢量。

[0120] 例如，在步骤 S201 取得的检索关键字（扩展关键字）的关键字 ID 为“KW2b”的情况下，在步骤 S205，与检索关键字相对应的空间 ID“2”的文档空间被选择。

[0121] 在此情况下，被分类为空间 ID“2”的文档空间的关键字（关键字 ID ＝ KW1b、KW2b、…、KWmb）中的关键字 ID“KW2b”的关键字被取得。因此，内容推测部 107，生成输入关键字矢量（KW1b、KW2b、…、KWmb）＝（0、1、…、0），以便针对所取得的关键字示出“1”，针对未被取得的关键字示出“0”。

[0122] 其次，内容推测部 107，通过算出在步骤 S206 生成的输入关键字矢量、和在步骤 S205 所选择的文档空间所示的各个内容的列矢量之间的余弦距离，从而算出各个内容对检索关键字的关联度（步骤 S207）。

[0123] 而且，例如，内容推测部 107，将算出了阈值以上的关联度的内容，推测为用户想要视听的内容 ID，并输出该内容的内容 ID（步骤 S208）。

[0124] 而且，在步骤 S206，也可以通过其它的方法生成输入关键字矢量。例如，内容推测部 107，检索与检索关键字有关的内容、和关键字记忆部 102 所记忆的关键字中的、除了检索关键字以外的与过去的关键字有关的内容。其次，内容推测部 107，在如此所检索的内容中的、重复的内容的数量的比率在阈值以上的情况下，利用包含检索关键字和过去的关键字的所有的关键字，制作输入关键字矢量。另一方面，在重复的内容的数量的比率不足阈值的情况下，内容推测部 107 也可以，不利用过去的关键字，而只利用检索关键字，生成输入关键字矢量。据此，能够从关键字记忆部 102 所记忆的过去的关键字中，制作符合想要寻找的内容的输入关键字矢量，也能够适当地推测用户想要寻找的内容。

[0125] 文档空间选择部 104，读出关键字记忆部 102 所存储的、包含检索关键字的关键字信息，从文档空间数据库 103 中找出与该检索关键字相对应的文档空间，即，找出示出检索关键字的关联度的文档空间。

[0126] 在此，文档空间选择部 104，在检索关键字的检索模式标志示出“筛选”的情况下，即，在检索关键字被显示为筛选关键字的情况下，选择比该检索关键字所对应的文档空间的空间 ID 大的空间 ID 所示的文档空间，以作为所述的筛选文档空间。

[0127] 例如，在被显示为筛选关键字的检索关键字的关键字 ID 为“KW2b”的情况下，文档空间选择部 104，从图 6 示出的文档空间数据库 103 中找出与关键字 ID“KW2b”的检索关键字相对应的空间 ID“2”的文档空间。而且，文档空间选择部 104，选择比空间 ID“2”大的空间 ID“3”的文档空间，以作为筛选文档空间。

[0128] 也就是，文档空间选择部 104，从文档空间数据库 103 中选择与比示出检索关键字的关联度的文档空间的出现频度少的出现频度相对应的文档空间，以作为筛选文档空间信息。

[0129] 另一方面，文档空间选择部 104，在检索关键字的检索模式标志示出“扩展”的情况下，即，在检索关键字被显示为扩展关键字的情况下，选择该检索关键字所对应的文档空间的空间 ID 以下的空间 ID 所示的文档空间，以作为所述的扩展文档空间。

[0130] 例如，在被显示为扩展关键字的检索关键字的关键字 ID 为“KW2b”的情况下，文档空间选择部 104，从图 6 示出的文档空间数据库 103 中找出与关键字 ID“KW2b”的检索关键字相对应的空间 ID“2”的文档空间。而且，文档空间选择部 104，选择该空间 ID“2”以下的空间 ID“1、2”的文档空间，以作为扩展文档空间。

[0131] 也就是，文档空间选择部 104，从文档空间数据库 103 中选择与示出检索关键字的关联度的文档空间信息的出现频度以上的出现频度相对应的文档空间信息，以作为扩展文档空间信息。

[0132] 如此，本发明中的文档空间选择部 104 包括：第一文档空间选择单元，从文档空间数据库 103 中选择与示出检索关键字的关联度的出现频度少的出现频度相对应的文档空间，以作为筛选文档空间；以及第二文档空间选择单元，从文档空间数据库 103 中选择与示

出检索关键字的关联度的出现频度以上的出现频度相对应的文档空间，以作为扩展文档空间信息。

[0133] 检索条件判定部 105，读出关键字记忆部 102 所存储的、包含检索关键字的关键字信息，确定该检索关键字的检索模式标志。因此，在检索模式标志示出“筛选”的情况下，检索条件判定部 105，判定针对该检索关键字的检索模式为筛选模式。另一方面，在检索模式标志示出“扩展”的情况下，检索条件判定部 105，判定针对该检索关键字的检索模式为扩展模式。

[0134] 也就是，在本实施例中的检索条件判定部 105，在用户选择输出部 111 所显示的筛选关键字，从而输入判别部 101 将筛选关键字作为检索关键字取得的情况下，判定为由内容推测部 107 进行筛选检索，另一方面，在用户选择输出部 111 所显示的扩展关键字，从而输入判别部 101 将扩展关键字作为检索关键字取得的情况下，判定为由内容推测部 107 进行扩展检索。

[0135] 关联关键字生成比率决定部 106，在检索条件判定部 105 所判定的检索模式为筛选模式的情况下，决定显示比率，以便应该显示的筛选关键字的数量多于扩展关键字的数量；在检索条件判定部 105 所判定的检索模式为扩展模式的情况下，决定显示比率，以便应该显示的扩展关键字的数量多于筛选关键字的数量。

[0136] 例如，关联关键字生成比率决定部 106 保持的阈值为 70％，在检索模式为筛选模式时，决定显示比率为筛选关键字数：扩展关键字数＝ 7 ： 3，在检索模式为扩展模式时，决定显示比率为筛选关键字数：扩展关键字数＝ 3 ： 7。

[0137] 因此，在检索结果画面显示与该显示比率相对应的数量的筛选关键字和扩展关键字。例如，若在检索结果画面所显示的关联关键字的最大数量为 10 个，则在筛选关键字数：扩展关键字数＝ 7 ： 3 时，在筛选列表 L2 显示 7 个筛选关键字，在筛选列表 L3 显示 3 个扩展关键字。另一方面，在筛选关键字数：扩展关键字数＝ 3 ： 7 时，在筛选列表 L2 显示 3 个筛选关键字，在筛选列表 L3 显示 7 个扩展关键字。

[0138] 关联关键字推测部 108，取得文档空间选择部 104 所选择的筛选文档空间以及扩展文档空间的空间 ID，并取得内容推测部 107 所推测的内容的内容 ID。因此，关联关键字推测部 108，在取得的空间 ID 所示的文档空间数据库 103 的筛选文档空间中，确定以阈值以上的关联度来与取得的内容 ID 的内容关联起来的关键字，以作为筛选关键字的候补。进而，关联关键字推测部 108，在取得的空间 ID 所示的文档空间数据库 103 的扩展文档空间中，确定以阈值以上的关联度来与取得的内容 ID 的内容关联起来的关键字，以作为扩展关键字的候补。而且，关联关键字推测部 108，确定除了与检索关键字相同的关键字以外的关键字，以作为筛选关键字或扩展关键字的候补。

[0139] 进而，关联关键字推测部 108，从关联关键字生成比率决定部 106 取得显示比率，根据该显示比率，决定应该显示的筛选关键字以及扩展关键字的各个数量（筛选显示数以及扩展显示数）。因此，关联关键字推测部 108，按从关联度大到小的顺序，从如上述所确定的筛选关键字的候补中选择筛选显示数的候补，以作为应该显示的筛选关键字。进而，关联关键字推测部 108，按从关联度大到小的顺序，从如上述所确定的扩展关键字的候补中选择扩展显示数的候补，以作为应该显示的扩展关键字。

[0140] 图 9 是用于说明关联关键字推测部 108 确定关联关键字的候补的工作的说明图。

[0141] 关联关键字推测部 108，例如，从文档空间选择部 104 取得筛选文档空间的空间 ID“2”，并且，从内容推测部 107 取得内容 ID“C4”。在此情况下，关联关键字推测部 108，在取得的空间 ID“2”所示的文档空间数据库 103 的筛选文档空间中，确定以阈值“0.5”以上的关联度来与取得的内容 ID“C4”的内容关联起来的关键字，以作为筛选关键字的候补。

[0142] 也就是，关联关键字推测部 108，从作为空间 ID“2”的筛选文档空间的矩阵中，参考与内容 ID“C4”相对应的列矢量 (0.8、0.2、0.1、0.4、0.7)。而且，关联关键字推测部 108，确定在该列矢量中的、与阈值为“0.5”以上的关联度“0.8、0.7”相对应的关键字 ID“KW1b、KW5b”的关键字，以作为筛选关键字的候补。

[0143] 图 10 是用于具体说明关联关键字推测部 108 推测关联关键字的工作的说明图。

[0144] 文档空间数据库 103，例如存储示出内容 1、内容 2 以及内容 3 和各个关键字的关系的三个文档空间 (空间 ID ＝ 1、2、3)。

[0145] 空间 ID“1”的文档空间示出，作为频度信息的 DF 值为 100 以上的关键字“体育、报道、电视剧”和内容 1、2、3 的关系 ；空间 ID“2”的文档空间示出，DF 值为大于 20 且 99 以下的关键字“足球、友情、日本”和内容 1、2、3 的关系 ；空间 ID“3”的文档空间示出，作为频度信息的 DF 值为 1 以上且不足 20 的关键字“天气、巴西、学校”和内容 1、2、3 的关系。并且，关键字和内容的关系，以作为它们之间的关联度的 TF-IDF 值来被示出。

[0146] 在这些情况下，关联关键字推测部 108，若从内容推测部 107 取得示出作为关于体育电视剧的节目的内容 3 的内容 ID，并从文档空间选择部 104 取得示出筛选文档空间的空间 ID“2”，则参考如上所述的文档空间数据库 103 中的、与内容 3 以及筛选文档空间 (空间 ID ＝ 2) 相对应的列矢量 (0.7、0.9、0.1)。而且，关联关键字推测部 108，确定在该列矢量中的具有阈值“0.5”以上的 TF-IDF 值“0.7、0.9”的关键字“足球、友情”，以作为筛选关键字的候补。

[0147] 进而，关联关键字推测部 108，从关联关键字生成比率决定部 106 取得显示比率，例如，利用在检索结果画面能够显示的关联关键字的最大数量和其显示比率，决定筛选显示数。关联关键字推测部 108，例如，在该筛选显示数为 1 时，在如上述所确定的筛选关键字的候补“足球、友情”中，只选择 TF-IDF 值最大的筛选关键字“友情”，以作为应该显示在检索结果画面的筛选关键字。

[0148] 关联关键字推测部 108，通过这些选择来推测筛选关键字，将该筛选关键字的关键字 ID 输出到输出生成部 110。

[0149] 另一方面，关联关键字推测部 108，若从内容推测部 107 取得示出内容 3 的内容 ID，并从文档空间选择部 104 取得示出扩展文档空间的空间 ID“1”，则参考如上所述的文档空间数据库 103 中的、与内容 3 以及扩展文档空间 (空间 ID ＝ 1) 相对应的列矢量 (0.8、0、0.7)。而且，关联关键字推测部 108，确定在该列矢量中的具有阈值“0.5”以上的 TF-IDF 值“0.7”的关键字“电视剧”，以作为扩展关键字的候补。

[0150] 而且，关联关键字推测部 108，由于检索关键字为“体育”，因此，若在扩展文档空间有关键字“体育”，则从除了该关键字“体育”以外的其它的关键字中，确定具有阈值以上的 TF-IDF 值的扩展关键字的候补。因此，在所述的例子中，在列矢量中阈值“0.5”以上的 TF-IDF 值有“0.8”和“0.7”，但是，只确定具有 TF-IDF 值“0.7”的关键字“电视剧”，以作为扩展关键字的候补。

[0151] 进而，关联关键字推测部 108，与所述相同，从关联关键字生成比率决定部 106 取得显示比率，例如，利用在检索结果画面能够显示的关联关键字的最大数量和其显示比率，决定扩展显示数。例如，在如上述所确定的扩展关键字的候补有多个时，关联关键字推测部 108，按从 TF-IDF 值大到小的顺序，从所确定的多个扩展关键字的候补中选择该扩展显示数的候补，以作为应该显示在检索结果画面的扩展关键字。

[0152] 关联关键字推测部 108，通过这些选择来推测扩展关键字，将该扩展关键字的关键字 ID 输出到输出生成部 110。

[0153] 图 11 是示出关联关键字推测部 108 的工作的流程图。

[0154] 首先，关联关键字推测部 108，取得内容推测部 107 所推测的内容的内容 ID（步骤 S301），并且，取得文档空间选择部 104 所选择的筛选文档空间的空间 ID（步骤 S302）。

[0155] 其次，关联关键字推测部 108，参考在步骤 S302 取得的空间 ID 示出的筛选文档空间中的、与在步骤 S301 取得的内容 ID 相对应的列矢量，确定具有阈值以上的关联度的关键字，以作为筛选关键字的候补（步骤 S303）。进而，关联关键字推测部 108，根据从关联关键字生成比率决定部 106 输出的显示比率，决定筛选显示数，并且，按从关联度大到小的顺序，选择在步骤 S303 所确定的筛选关键字的候补中的、其筛选显示数的候补，以作为应该显示在检索结果画面的筛选关键字（步骤 S304）。

[0156] 并且，关联关键字推测部 108，取得文档空间选择部 104 所选择的扩展文档空间的空间 ID（步骤 S305）。

[0157] 而且，关联关键字推测部 108，参考在步骤 S305 取得的空间 ID 示出的扩展文档空间中的、与在步骤 S301 取得的内容 ID 相对应的列矢量，确定具有阈值以上的关联度的关键字，以作为扩展关键字的候补（步骤 S306）。而且，关联关键字推测部 108，若在该扩展文档空间中有与检索关键字相同的关键字，则从该关键字以外的关键字中，确定具有阈值以上的关联度的扩展关键字的候补。

[0158] 进而，关联关键字推测部 108，根据从关联关键字生成比率决定部 106 输出的显示比率，决定扩展显示数，并且，按从关联度大到小的顺序，选择在步骤 S306 所确定的扩展关键字的候补中的、其扩展显示数的候补，以作为应该显示在检索结果画面的扩展关键字（步骤 S307）。

[0159] 关联关键字推测部 108，通过执行这些步骤 S301 ～步骤 S307 的处理，从而将筛选关键字以及扩展关键字推测为关联关键字。

[0160] 而且，本实施例中的关联关键字推测部 108 被构成为关联关键字选择单元，从内容数据库 109 中选择与筛选文档空间以及扩展文档空间分别相对应的筛选关键字以及扩展关键字，以便针对内容推测部 107 所检索的内容示出的关联度越大的关键字就越被选择。

[0161] 输出生成部 110，例如，若从内容推测部 107 取得示出内容 3 的内容 ID，则从内容数据库 109 中抽出与该内容 3 相对应的内容属性信息，并输出到输出部 111。进而，输出生成部 110，若从关联关键字推测部 108 取得筛选关键字以及扩展关键字的各个关键字 ID，则从内容数据库 109 中抽出这些关键字 ID 示出的筛选关键字以及扩展关键字，并输出到输出部 111。

[0162] 图 12 是示出输出部 111 所显示的检索结果画面的一个例子的图。

[0163] 输出部 111，针对检索关键字“体育”，若取得从输出生成部 110 如上述所输出的、内容 3 所对应的内容属性信息、和筛选关键字以及扩展关键字，则将它们配置并显示在检索结果画面的规定的地方。

[0164] 因此，在内容列表 L1，显示内容的名称“内容 3”以及概要，以作为针对内容 3 的内容属性信息。并且，在筛选列表 L2，显示用于筛选在内容列表 L1 所示的内容的筛选关键字“友情、足球”；在扩展列表 L3，显示用于以与检索关键字“体育”不同的观点来重新进行检索的扩展关键字“电视剧”。

[0165] 图 13 是示出内容检索装置 100 整体的工作的流程图。

[0166] 首先，内容检索装置 100 的输入判别部 101，接受通过用户的操作所选择的输入选择项（步骤 S101），判别该输入选择项示出关键字还是内容（步骤 S102）。

[0167] 在此，若由输入判别部 101 判别为示出内容（步骤 S102 的“否”），输出生成部 110，则从内容数据库 109 中抽出该内容的内容数据，从而生成向输出部 111 输出的信息（步骤 S109）。而且，输出部 111，若从输出生成部 110 取得该内容数据，则再生并显示该内容数据（步骤 S110）。

[0168] 另一方面，输入判别部 101，若判别为输入选择项示出关键字（步骤 S102 的“是”），则将包含该关键字（检索关键字）以及关键字 ID 等的关键字信息存储到关键字记忆部 102（步骤 S103）。

[0169] 内容推测部 107，若在步骤 S103 包含检索关键字的关键字信息被存储到关键字记忆部 102，如图 8 示出，则利用关键字记忆部 102 以及文档空间数据库 103，检索并推测与检索关键字有关的内容（步骤 S104）。

[0170] 进而，文档空间选择部 104，利用检索关键字，从文档空间数据库 103 中选择筛选文档空间和扩展文档空间（步骤 S105）。

[0171] 并且，检索条件判定部 105，按照关键字记忆部 102 所存储的检索关键字，判定检索模式是筛选模式还是扩展模式（步骤 S106）。而且，关联关键字生成比率决定部 106，按照在步骤 S106 所判定的结果，决定应该显示的筛选关键字的数量和扩展关键字的数量的显示比率（步骤 S107）。

[0172] 若在步骤 S107 决定显示比率，关联关键字推测部 108，则利用该显示比率、在步骤 S104 所推测的内容、在步骤 S105 所选择的筛选文档空间以及扩展文档空间，推测筛选关键字以及扩展关键字（步骤 S108）。

[0173] 若如此推测筛选关键字以及扩展关键字，输出生成部 110，从内容数据库 109 中抽出这些筛选关键字以及扩展关键字，从而生成向输出部 111 输出的检索结果信息（步骤 S109）。而且，输出部 111，若从输出生成部 110 取得该检索结果信息，则显示该检索结果信息中包含的关联关键字（步骤 S110）。

[0174] 并且，若用户选择输出部 111 所显示的关联关键字中的任一个，内容检索装置 100，则接受所选择的关联关键字，以作为新的下个检索关键字。也就是，内容检索装置 100，反复执行从步骤 S101 开始的处理。因此，用户能够从内容数据库 109 中找出所希望的内容。

[0175] 如此，在本实施例中，利用按照关键字的频度信息被分类的多个文档空间，选择比检索关键字所对应的文档空间频度信息小的频度信息的文档空间，以作为筛选文档空间，并显示依据该筛选文档空间示出大的关联度的关键字，以作为筛选关键字，并且，选择检索

关键字所对应的文档空间的频度信息以上的频度信息的文档空间，以作为扩展文档空间，并显示依据该扩展文档空间示出大的关联度的关键字，以作为扩展关键字。

[0176] 以下是以频度信息来将文档空间分类的理由。在关键字的频度信息的值、和与关键字有关的内容数量之间存在正相关关系。因此，为了向用户出示用于筛选内容的关键字（筛选关键字），而需要出示具有比检索关键字的频度信息的值小的频度信息的关键字，以作为筛选关键字。另一方面，为了向用户出示用于检索与向用户所出示的内容不同的其它的内容的关键字（扩展关键字），而需要出示具有正在出示中的内容数量以上的频度信息的值的关键字，以作为扩展关键字。

[0177] 于是，如上所述，预先以频度信息来将文档空间分类，以便汇集与内容数量的关系为相同程度的关键字，来在它们之间能够建立关系，并分别使用用于选择筛选关键字或扩展关键字的文档空间，从而能够同时出示适当的筛选关键字以及扩展关键字。

[0178] 也就是，在作为用户的检索对象的内容流动变化的情况下，也能够向用户出示用于检索所希望的内容的关联关键字，因此用户能够从大量的内容中容易找出所希望的内容。

[0179] 而且，在本实施例中，如图 10 示出，将文档空间分类为三个，将用于分类的阈值(DF 值）为 100 或 20，但也可以，例如，在预先对多名用户进行试验后决定该文档空间的数量或阈值。具体而言，首先，算出文档空间整体中包含的关键字的 DF 值，向用户出示 DF 值不同的多个关键字。而且，将各个关键字分类为，在以所出示的关键字来检索时有可能获得与相同程度的检索结果数的组。据此，能够设定文档空间的数量或阈值。

[0180] 以上，对于本发明涉及的内容检索装置，利用所述实施例进行了说明，但是，本发明不仅限于此。

[0181] 例如，关联关键字推测部 108，在从多个筛选关键字的候补中选择几个筛选关键字时，按从关联度大到小的顺序进行了选择，但也可以，利用联想词典进行选择。

[0182] 一般而言，在用户选择筛选关键字或扩展关键字等的关联关键字的情况下，选择与检索关键字的关联性强的关联关键字。然而，在利用所述的 LSI 制作文档空间数据库 103 的各个文档空间的情况下，有时，在关联关键字推测部 108 所确定的筛选关键字的候补中包含与检索关键字的关联性弱的候补。因此，即使出示与检索关键字的关联性弱的筛选关键字，也该筛选关键字被选择的可能性低。

[0183] 于是，为了避免这些关联性弱的筛选关键字被出示，关联关键字推测部 108，而预先制作示出构成文档空间的关键字之间的关联性的联想词典，并利用它来从筛选关键字的候补中选择应该显示的筛选关键字。

[0184] 联想词典是，例如以“自然言语处理（自然语言处理）”(pp，387 ～ 392，岩波书店，1996)（非专利文献 3）所示的方法来自动制作出的。并且，也可以利用设计者预先制作的联想词典。

[0185] 图 14 是示出联想词典的一个例子的图。

[0186] 如图 14 示出，联想词典被构成为矩阵，以联想值来示出关键字之间的关联强度。

[0187] 关联关键字推测部 108，利用这些联想词典，从所确定的多个筛选关键字的候补中选择应该显示的几个筛选关键字。

[0188] 例如，关键字 ID“KW2”的关键字作为检索关键字被接受到输入判别部 101，并确定

关键字 ID“KW3、KW4、KW5”的关键字，以分别作为筛选关键字的候补。并且，关联关键字推测部 108，根据从关联关键字生成比率决定部 106 输出的显示比率，决定筛选显示数“2”。

[0189] 在这些情况下，关联关键字推测部 108，在图 14 示出的联想词典中，参考针对检索关键字（关键字 ID ＝ KW2）的、筛选关键字的候补（关键字 ID ＝ KW3、KW4、KW5）的联想值(0.4、0.7、0.8)。而且，关联关键字推测部 108，从该三个筛选关键字的候补中，按从联想值大到小的顺序，选择二个候补，以作为应该显示的筛选关键字。也就是，关键字 ID“KW4、KW5”的关键字被选择，以作为筛选关键字。

[0190] 据此，能够防止从筛选关键字的候补中的、与检索关键字的关联性弱的候补被显示为筛选关键字。因此，能够减少因用户不能找出所希望的筛选关键字而重新进行检索来增加的检索步骤数。

[0191] 并且，关联关键字推测部 108，在从扩展关键字的候补中选择应该显示的扩展关键字时，也可以与所述相同利用联想词典来进行选择。据此，能够更减少因用户不能找出所希望的扩展关键字而重新进行检索来增加的检索步骤数。

[0192] 并且，在本实施例中，关联关键字推测部 108，将所选择的多个关联关键字的关键字 ID 输出到输出生成部 110，但也可以，不是将这些关键字 ID 输出，而是根据所选择的多个关联关键字生成新的结合关联关键字，并将该结合关联关键字通过输出生成部 110 显示在输出部 111。

[0193] 而且，关联关键字推测部 108，在选择了二个筛选关键字时，判别在与二个筛选关键字的每一个相对应的内容群中重复的内容的数量是否在阈值以上。因此，若判别为在阈值以上，关联关键字推测部 108，则结合这些二个筛选关键字来生成结合关联关键字，并将该结合关联关键字显示在筛选列表 L2。并且，与所述相同，关联关键字推测部 108，在选择了二个扩展关键字时，判别在与二个扩展关键字的每一个相对应的内容群中重复的内容的数量是否在阈值以上。因此，若判别为在阈值以上，关联关键字推测部 108，则结合这些二个扩展关键字来生成结合关联关键字，并将该结合关联关键字显示在扩展列表 L3。

[0194] 据此，由于将被预测为检索结果都相同的二个关联关键字结合来显示为一个关联关键字，因此，用户能够不迷惑应该选择哪个关联关键字，而容易选择所希望的关联关键字。

[0195] 本发明的内容检索装置可以适应于，从服务器等所存储的大量的内容中检索用户想要视听的内容的检索装置。例如，能够适用于如图 15 示出的、内置有硬盘的 DVD(Digital Versatile Disk：数字多用光盘）刻录器或播放器等的、再生节目的设备。内容是，例如电视节目、电影和音乐等的视听内容、或书籍和论文等的文本内容等，与种类无关，而只要具有作为内容的属性的字符信息，就可以是任何数据。

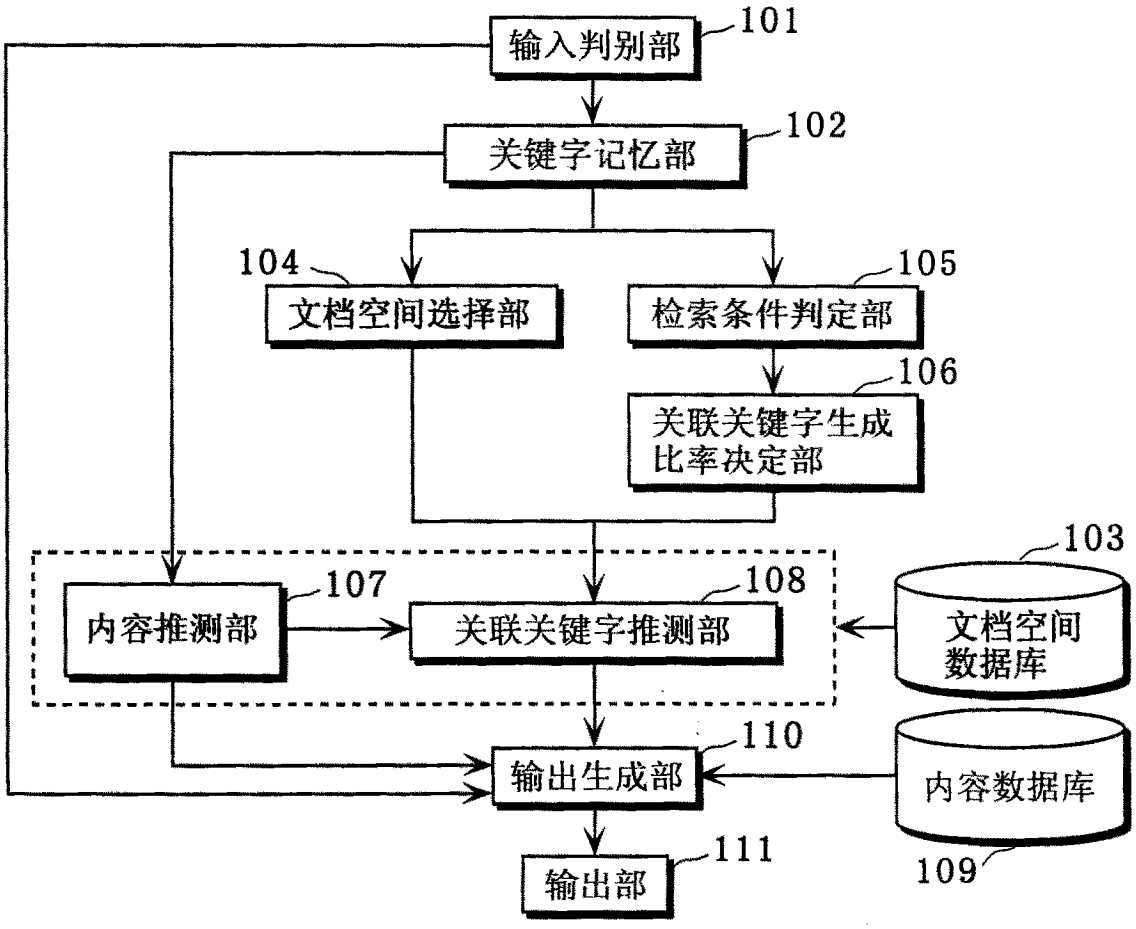

输入判别部
101
关键字记忆部
102
104
文档空间选择部
105
检索条件判定部
106
关联关键字生成
比率决定部
103
107
内容推测部
108
关联关键字推测部
文档空间
数据库
110
输出生成部
内容数据库
111
输出部
109

图 1

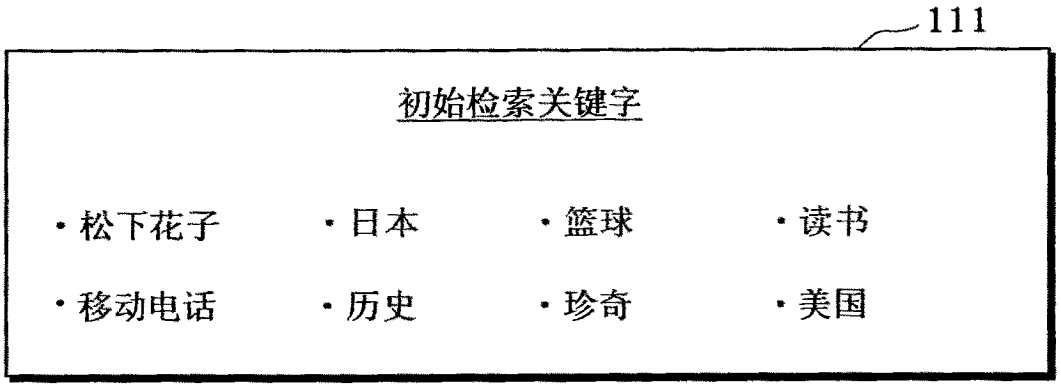

111
初始检索关键字
・松下花子
・日本
・篮球
・读书
・移动电话
・历史
・珍奇
・美国

图 2

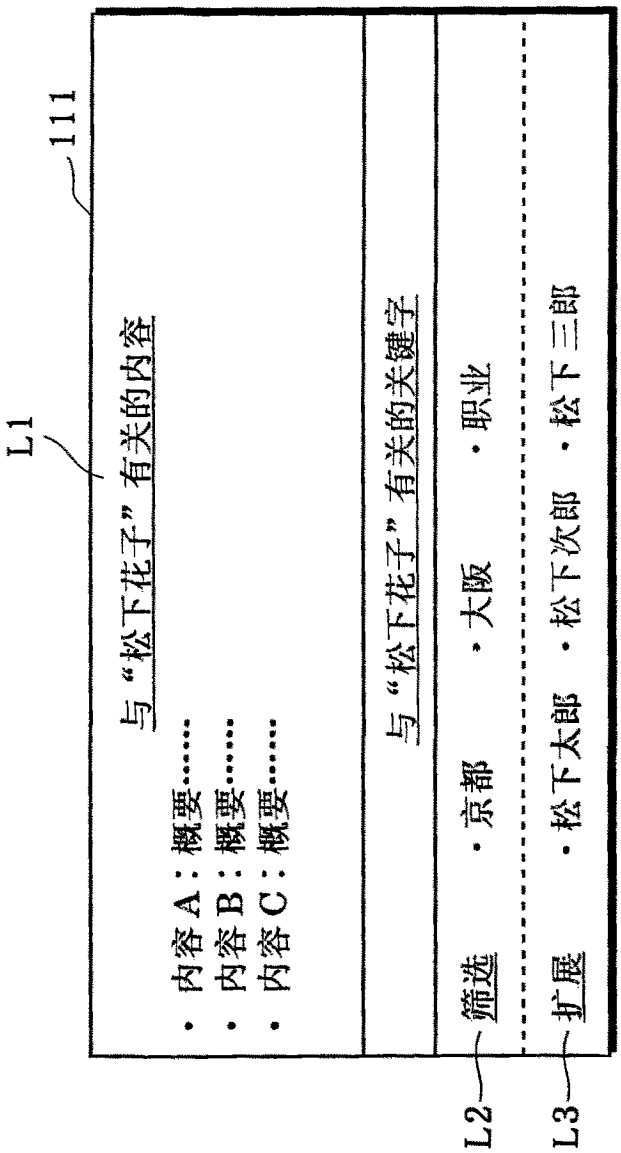

111
L1
与“松下花子”有关的内容
• 内容A：概要……
• 内容B：概要……
• 内容C：概要……
与“松下花子”有关的关键字
筛选
• 京都
• 大阪
• 职业
L2
扩展
• 松下太郎
• 松下次郎
• 松下三郎
L3

图 3

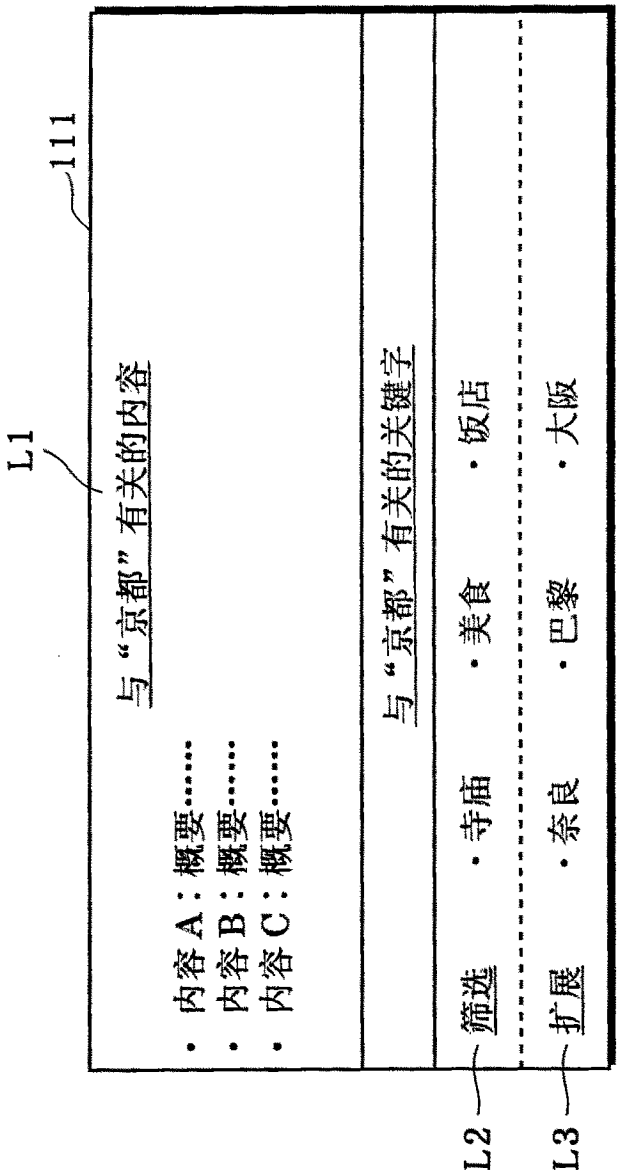
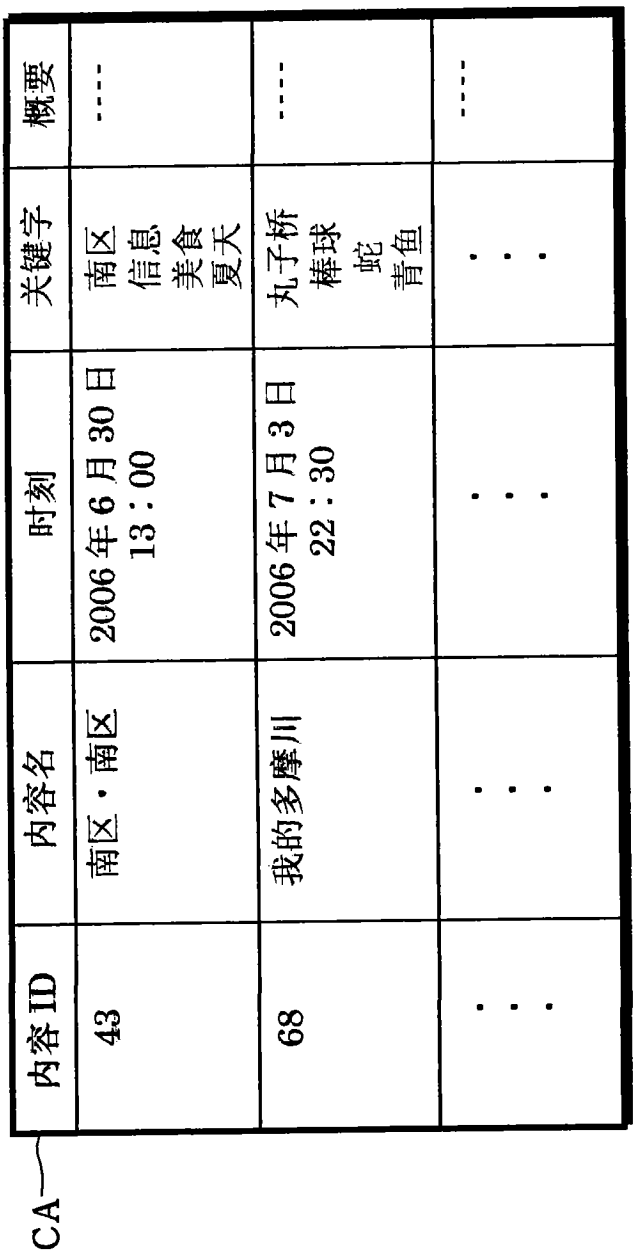

111
L1
与“京都”有关的内容
• 内容A：概要……
• 内容B：概要……
• 内容C：概要……
与“京都”有关的关键字
筛选
• 寺庙
• 美食
• 饭店
扩展
• 奈良
• 巴黎
• 大阪
L2
L3

图 4


CA
内容ID
内容名
时刻
关键字
概要
43
南区・南区
2006年6月30日 13：00
南区 信息 美食 夏天

68
我的多摩川
2006年7月3日 22：30
丸子桥 棒球 蛇 青鱼

图 5

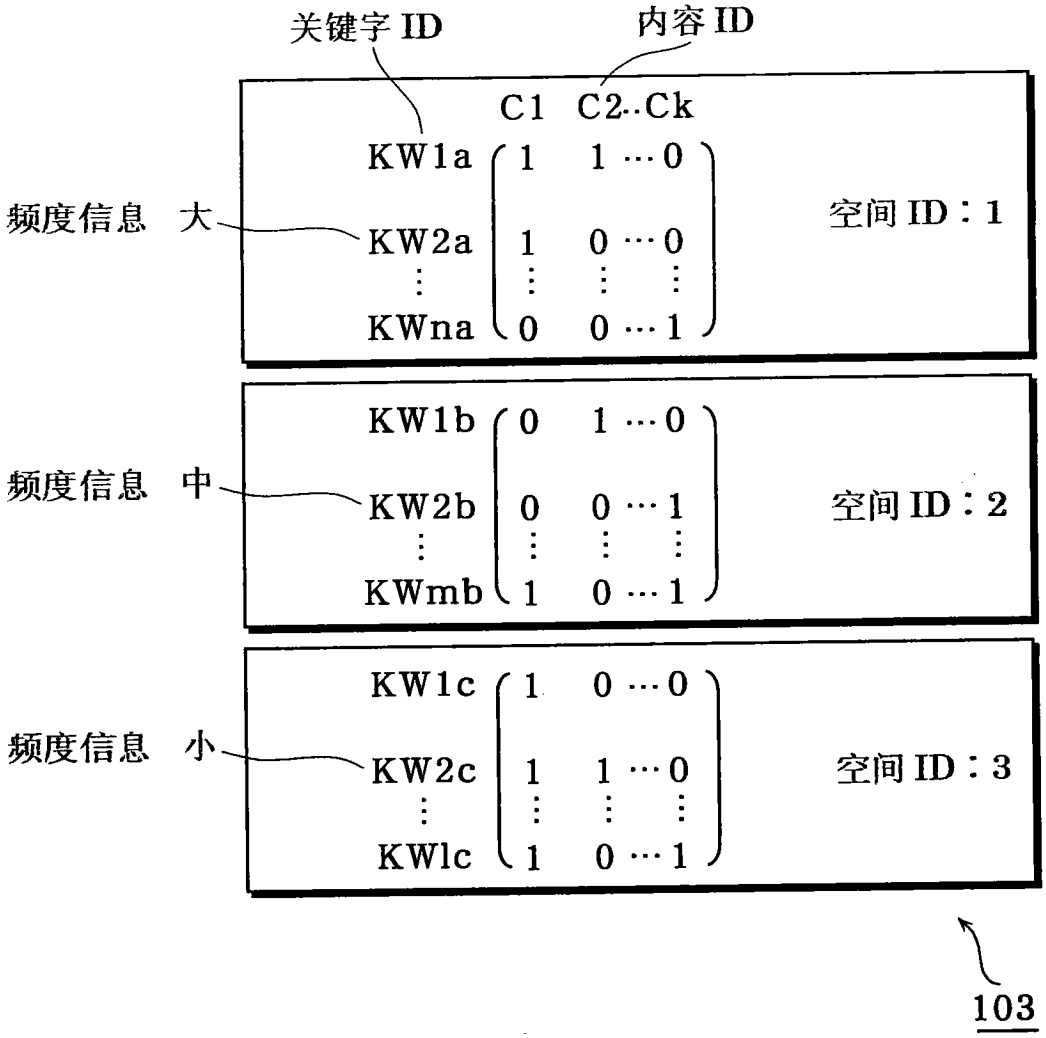


图 6

检索步骤	关键字	关键字 ID	检索模式标志
1	松下花子	10	初始
2	京都	23	筛选
3	美食	54	扩展
…	…	…	…

图 7

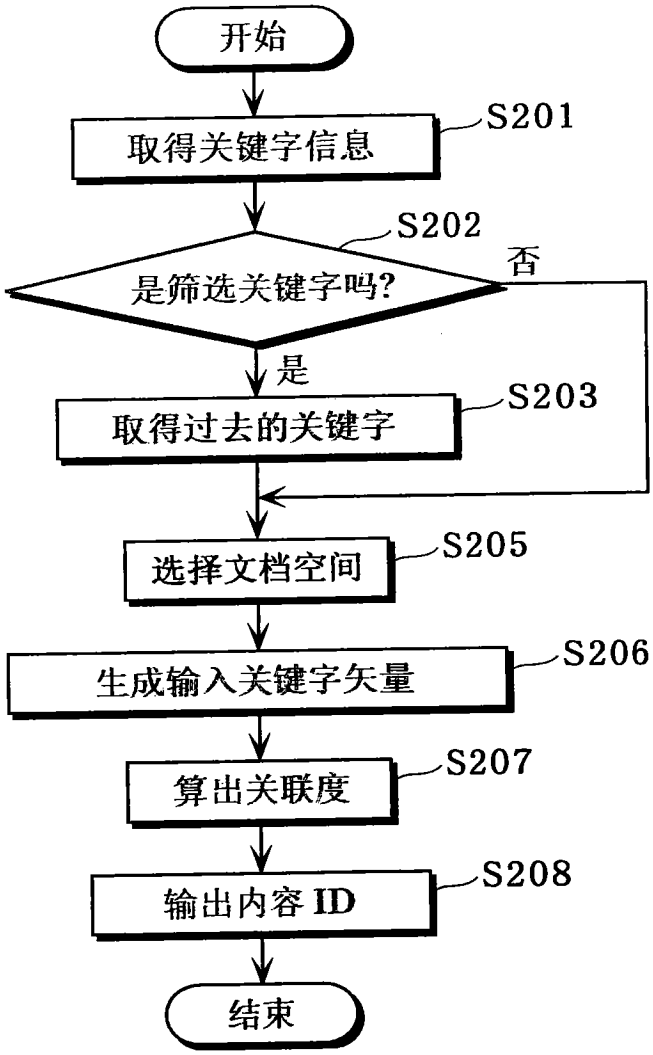

开始
取得关键字信息
S201
S202
是筛选关键字吗?
否
是
取得过去的关键字
S203
选择文档空间
S205
生成输入关键字矢量
S206
算出关联度
S207
输出内容 ID
S208
结束

图 8

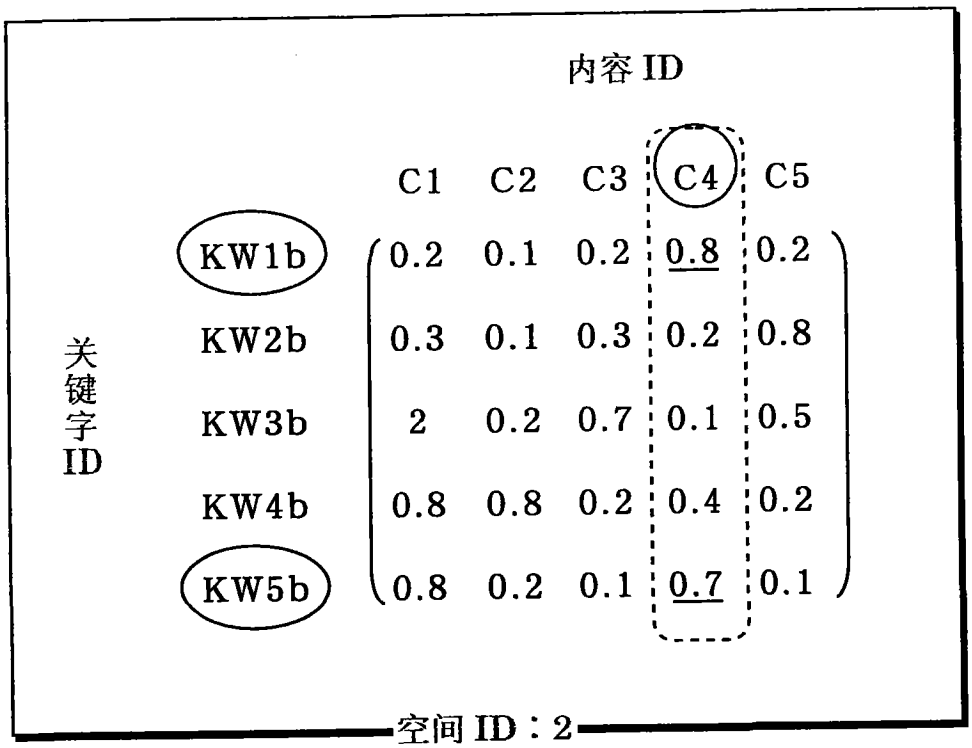

内容 ID
C1 C2 C3 C4 C5
KW1b 0.2 0.1 0.2 0.8 0.2
KW2b 0.3 0.1 0.3 0.2 0.8
KW3b 2 0.2 0.7 0.1 0.5
KW4b 0.8 0.8 0.2 0.4 0.2
KW5b 0.8 0.2 0.1 0.7 0.1
关键字 ID
空间 ID：2

图 9

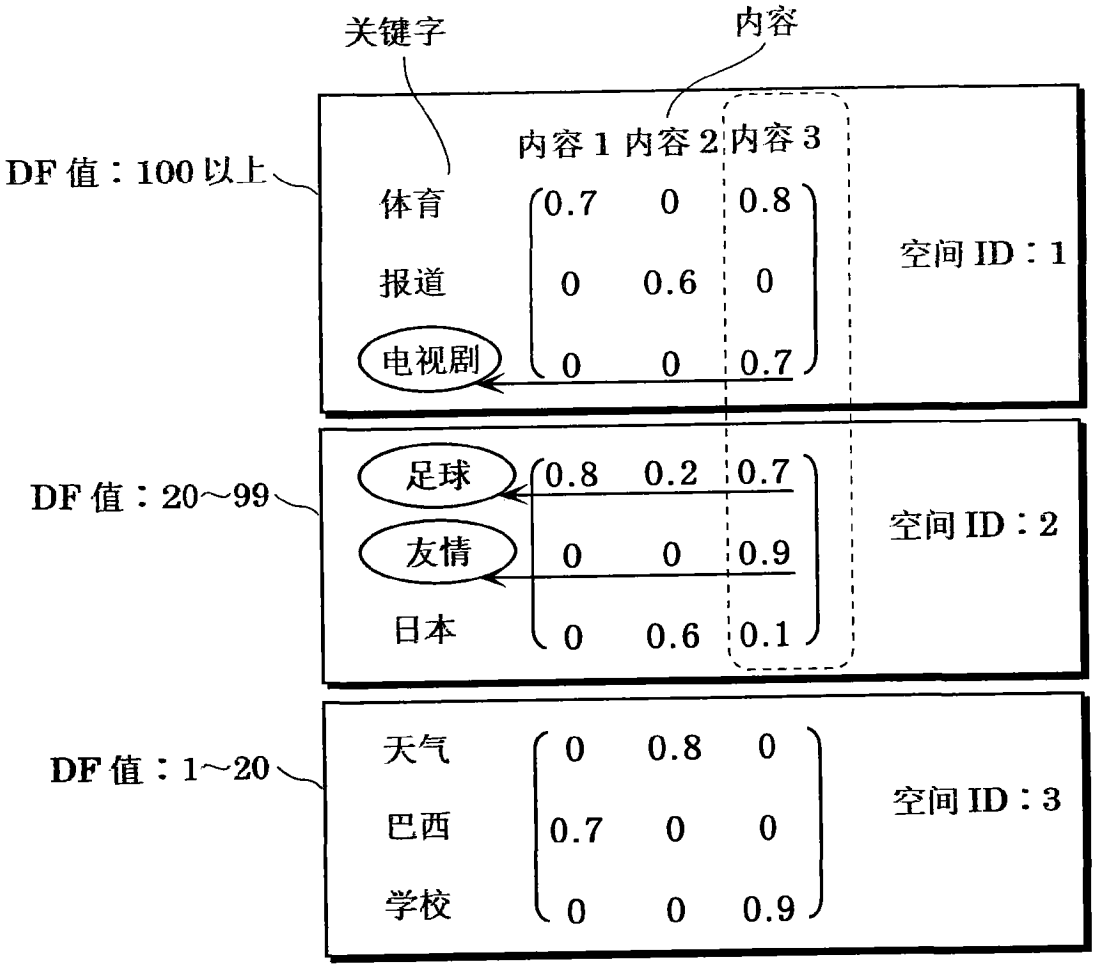

关键字
内容
内容 1 内容 2 内容 3
DF 值：100 以上
体育
报道
电视剧
0.7 0 0.8
0 0.6 0
0 0 0.7
空间 ID：1
DF 值：20～99
足球
友情
日本
0.8 0.2 0.7
0 0 0.9
0 0.6 0.1
空间 ID：2
DF 值：1～20
天气
巴西
学校
0 0.8 0
0.7 0 0
0 0 0.9
空间 ID：3

图 10

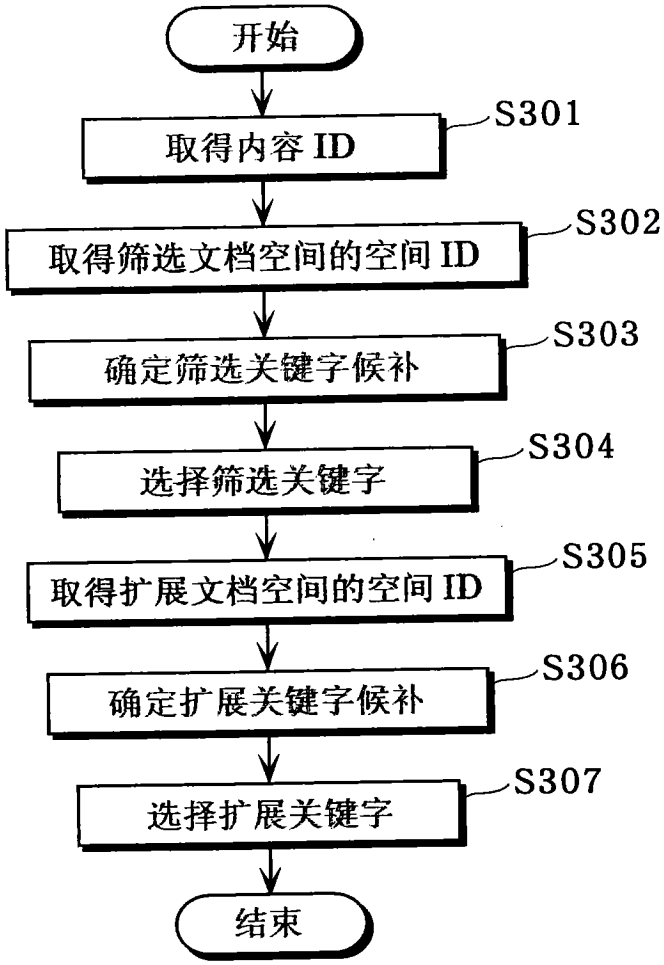

开始
取得内容 ID
S301
取得筛选文档空间的空间 ID
S302
确定筛选关键字候补
S303
选择筛选关键字
S304
取得扩展文档空间的空间 ID
S305
确定扩展关键字候补
S306
选择扩展关键字
S307
结束

图 11

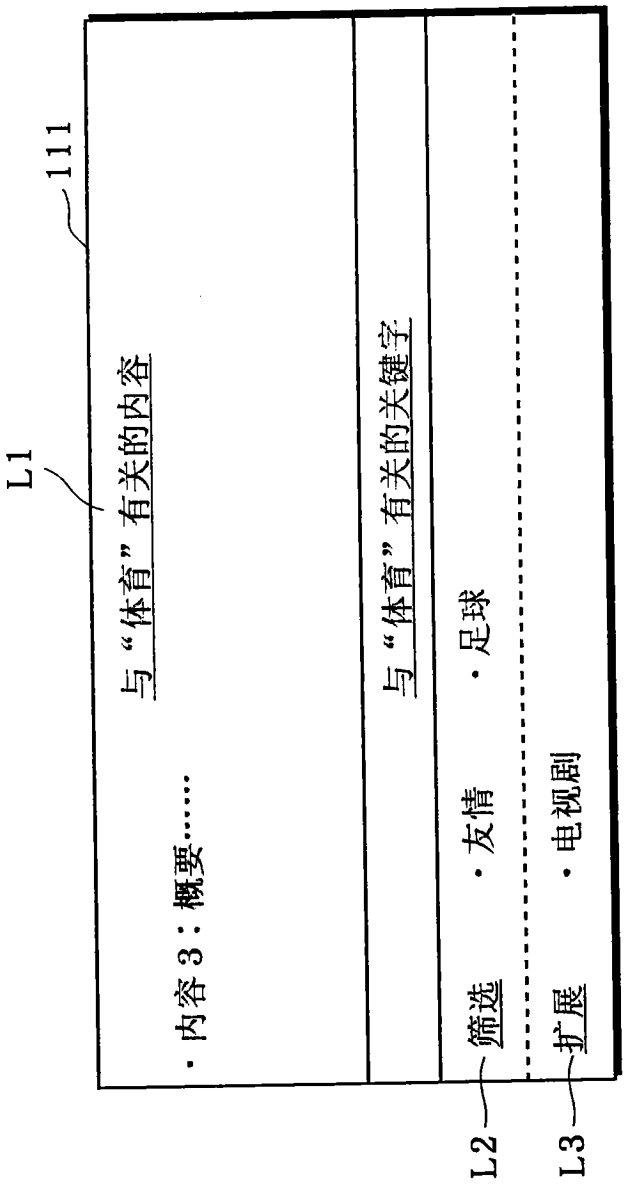

111
L1
与“体育”有关的内容
·内容3：概要……
与“体育”有关的关键字
·足球
·友情
·电视剧
筛选
扩展
L2
L3

图 12

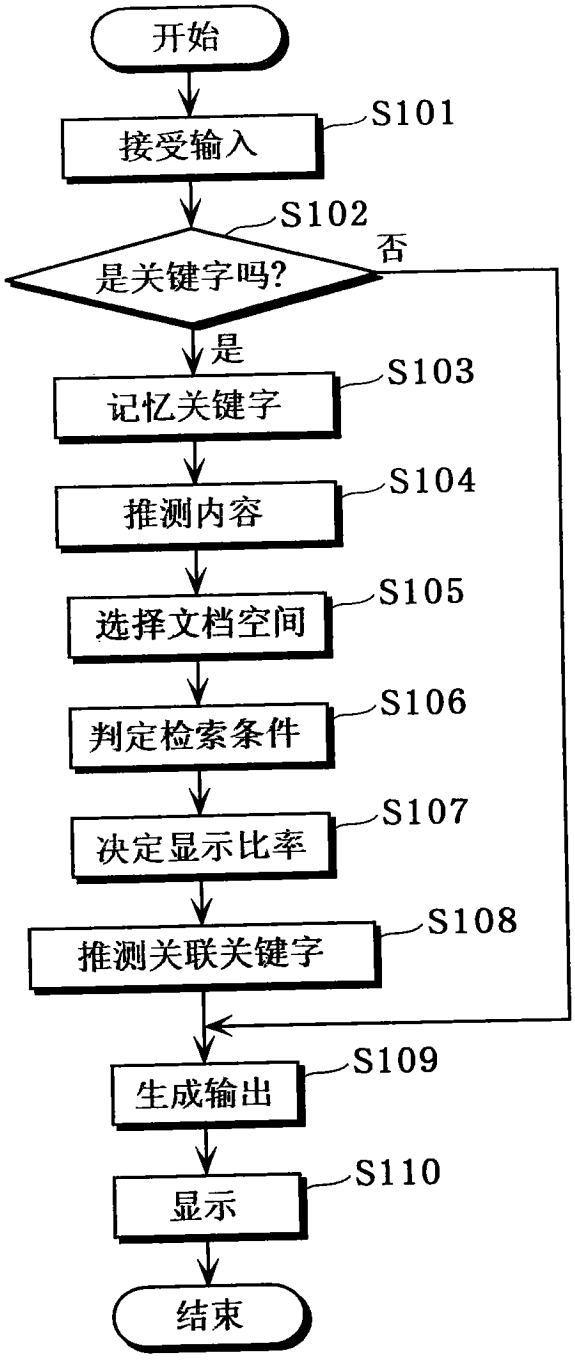

开始
接受输入
S101
S102
是关键字吗?
否
是
记忆关键字
S103
推测内容
S104
选择文档空间
S105
判定检索条件
S106
决定显示比率
S107
推测关联关键字
S108
生成输出
S109
显示
S110
结束

图 13

关键字ID	KW1	KW2	KW3	KW4	KW5
KW1	0.2	0.1	0.2	0.8	0.2
KW2	0.3	0.1	0.4	0.7	0.8
KW3	2	0.2	0.7	0.1	0.5
KW4	0.8	0.8	0.2	0.4	0.2
KW5	0.8	0.2	0.1	0.7	0.1

图 14

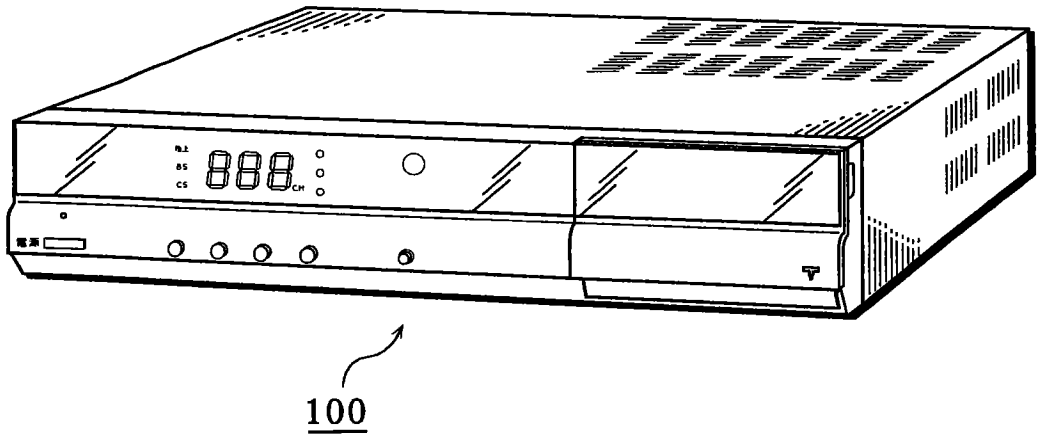


图 15