



(12) 发明专利申请

(10) 申请公布号 CN 102622399 A

(43) 申请公布日 2012. 08. 01

(21) 申请号 201210004578. 5

(22) 申请日 2012. 01. 04

(30) 优先权数据

2011-003005 2011. 01. 11 JP

(71) 申请人 索尼公司

地址 日本东京都

(72) 发明人 花井裕也

(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 李晓冬

(51) Int. Cl.

G06F 17/30 (2006. 01)

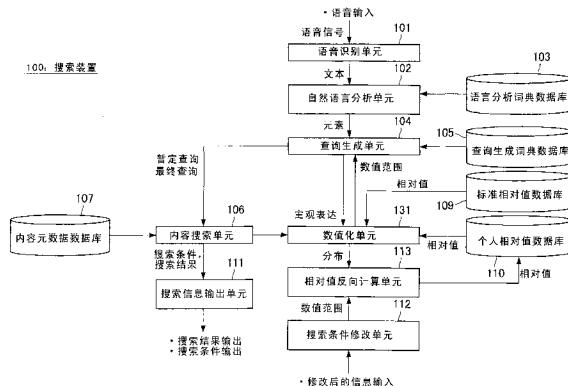
权利要求书 3 页 说明书 22 页 附图 21 页

(54) 发明名称

搜索装置、搜索方法和程序

(57) 摘要

一种搜索装置、搜索方法和程序。公开了用于基于包括至少一个或多个主观元素和一个或多个对象元素的搜索查询来执行搜索的技术。可基于搜索查询的一个或多个对象元素来执行对内容的第一搜索。可确定从第一搜索返回的结果的数值特征的分布，其中数值特征的分布对应于搜索查询的主观元素中的一个。可确定对于主观元素的一个或多个比例阈值，并且可通过应用一个或多个比例阈值来选择数值特征的范围。可基于所选择的数值特征的范围来执行对内容的第二搜索。



1. 一种装置，包括：

至少一个处理器；以及

至少一个存储器，其存储处理器可执行指令，所述处理器可执行指令被配置为使得所述至少一个处理器执行以下处理：

接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；

基于所述一个或多个对象元素执行对内容的第一搜索；

确定从所述第一搜索返回的结果的数值特征的分布，所述数值特征对应于所述一个或多个主观元素中的第一主观元素；

确定对于所述第一主观元素的一个或多个比例阈值；

通过对所述分布应用所述一个或多个比例阈值来选择数值特征的范围；以及

基于所选择的数值特征的范围来执行对内容的第二搜索。

2. 根据权利要求 1 所述的装置，其中，确定所述一个或多个比例阈值包括访问将所述一个或多个比例阈值与所述第一主观元素关联起来的数据库。

3. 根据权利要求 1 所述的装置，其中，确定所述一个或多个比例阈值包括接收对于所述第一主观元素指定所述一个或多个比例阈值的用户输入。

4. 根据权利要求 3 所述的装置，其中，所述处理器可执行指令还被配置为使得所述至少一个处理器将由所述用户指定的所述一个或多个比例阈值存储在与所述用户相关联的比例阈值的数据库中。

5. 根据权利要求 1 所述的装置，其中，所述处理器可执行指令还被配置为使得所述至少一个处理器执行以下处理：

接收用于修改所选择的数值特征的范围的用户输入；以及

基于修改后的数值特征的范围来执行对内容的第三搜索。

6. 根据权利要求 1 所述的装置，其中，所述搜索查询包括由用户提供的自然语言查询，并且其中，所述第一主观元素包括所述自然语言查询中的一个或多个词。

7. 一种方法，包括：

接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；

基于所述一个或多个对象元素执行对内容的第一搜索；

利用至少一个处理器来确定从所述第一搜索返回的结果的数值特征的分布，所述数值特征对应于所述一个或多个主观元素中的第一主观元素；

确定对于所述第一主观元素的一个或多个比例阈值；

通过对所述分布应用所述一个或多个比例阈值来选择数值特征的范围；以及

基于所选择的数值特征的范围来执行对内容的第二搜索。

8. 根据权利要求 7 所述的方法，其中，确定所述一个或多个比例阈值包括访问将所述一个或多个比例阈值与所述第一主观元素相关联的数据库。

9. 根据权利要求 7 所述的方法，其中，确定所述一个或多个比例阈值包括接收对于所述第一主观元素指定所述一个或多个比例阈值的用户输入。

10. 根据权利要求 9 所述的方法，还包括将由所述用户指定的所述一个或多个比例阈值存储在与所述用户相关联的比例阈值的数据库中。

11. 根据权利要求 7 所述的方法，还包括：

接收用于修改所选择的数值特征的范围的用户输入；以及
基于修改后的数值特征的范围来执行对内容的第三搜索。

12. 根据权利要求 7 所述的方法，其中，所述搜索查询包括由用户提供的自然语言查询，并且其中，所述第一主观元素包括所述自然语言查询中的一个或多个词。

13. 编码有计算机可执行指令的至少一个计算机可读存储介质，所述计算机可执行指令被执行时，实现包括以下步骤的方法：

接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；

基于所述一个或多个对象元素执行对内容的第一搜索；

确定从所述第一搜索返回的结果的数值特征的分布，所述数值特征对应于所述一个或多个主观元素中的第一主观元素；

确定对于所述第一主观元素的一个或多个比例阈值；

通过对所述分布应用所述一个或多个比例阈值来选择数值特征的范围；以及

基于所选择的数值特征的范围来执行对内容的第二搜索。

14. 根据权利要求 13 所述的至少一个计算机可读存储介质，其中，确定所述一个或多个比例阈值包括访问将所述一个或多个比例阈值与所述第一主观元素相关联的数据库。

15. 根据权利要求 13 所述的至少一个计算机可读存储介质，其中，确定所述一个或多个比例阈值包括接收对于所述第一主观元素指定所述一个或多个比例阈值的用户输入。

16. 根据权利要求 15 所述的至少一个计算机可读存储介质，其中，所述方法还包括将由所述用户指定的所述一个或多个比例阈值存储在与所述用户相关联的比例阈值的数据
库中。

17. 根据权利要求 13 所述的至少一个计算机可读存储介质，其中，所述方法还包括：

接收用于修改所选择的数值特征的范围的用户输入；以及

基于修改后的数值特征的范围来执行对内容的第三搜索。

18. 根据权利要求 13 所述的至少一个计算机可读存储介质，其中，所述搜索查询包括由用户提供的自然语言查询，并且其中，所述第一主观元素包括所述自然语言查询中的一个或多个词。

19. 一种装置，包括：

用于接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询的装置；

用于基于所述一个或多个对象元素执行对内容的第一搜索的装置；

用于确定从所述第一搜索返回的结果的数值特征的分布的装置，所述数值特征对应于所述一个或多个主观元素中的第一主观元素；

用于确定对于所述第一主观元素的一个或多个比例阈值的装置；

用于通过对所述分布应用所述一个或多个比例阈值来所述数值特征的范围的装置；以
及

用于基于所选择的所述数值特征的范围来执行对内容的第二搜索的装置。

20. 一种装置，包括：

至少一个处理器；以及

至少一个存储器，其存储处理器可执行指令，所述处理器可执行指令被配置为使得所述至少一个处理器执行以下处理：

接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；

显示从基于所述一个或多个对象元素的对内容的第一搜索返回的结果的数值特征的范围，所述数值特征对应于所述一个或多个主观元素中的第一主观元素；

接收用于修改数值特征的范围的用户输入；以及

显示基于修改后的数值特征的范围的、对内容的第二搜索的结果。

搜索装置、搜索方法和程序

技术领域

[0001] 本发明涉及搜索装置、搜索方法和程序。

背景技术

[0002] 近年来,例如使用互联网对各种信息搜索的可能性增加。例如,对交通工具的换乘信息、到达目的地的路径信息等的搜索正被经常执行。此外,利用诸如便携式电话或个人数字助理等的便携式设备对街道上的商店信息或者票或优惠券信息的搜索也正被经常执行。因此,近年来,信息搜索的几率迅速增加。此外,随着信息搜索服务的用户增加,已经提供了信息搜索系统的基础设施。结果,近年来,能够例如利用便携设备即时搜索各种信息。

[0003] 假设使用难以输入搜索条件的便携设备,已经设计了能够通过语音输入搜索条件的语音搜索服务。此外,还已经设计了能够以自然语句“给我看在银座便宜的寿司餐厅”的形式来输入搜索条件的信息搜索服务。结果,能够迅速并且简单地输入搜索条件,从而极大地提高可用性。然而,在大量的信息搜索服务中,即使当搜索条件中包含主观表达“便宜”和“好”时,也只有包含用语“便宜”和“好”的信息被抽取。

[0004] 例如,即使当使用自然语句“给我看在银座便宜的寿司餐厅”作为搜索条件时,只有含有关键词“便宜”的寿司餐厅信息被抽取并且一些低价的寿司餐厅可能不一定被抽取。因此,想要得到关于便宜的寿司餐厅信息的用户必须使用含有他们认为便宜的价格范围的搜索条件来执行搜索,或者必须一条一条的确认所抽取的寿司餐厅信息并找到便宜的寿司餐厅。作为这样的问题的解决方案,在日本专利特许公开 No. 2002-24212 中公开了如下方法,其中预先准备与数值范围相关联的主观表达数据库,并且自动地将包含在作为搜索条件输入的自然语句中的主观表达转换为数值范围来执行搜索。

发明内容

[0005] 如果应用在日本专利特许公开 No. 2002-24212 中公开的方法,肯定希望利用与用户使用主观表达所指定的期望的搜索条件近似的条件来执行搜索,并且抽取反映用户的意图的搜索结果。然而,在该方法中,如果除了主观表达以外的搜索条件改变,则没有考虑用户使用主观表达所指定的期望的数值范围的变化。例如,由搜索条件“给我看在银座便宜的寿司餐厅”中的“便宜”所指示的价格范围,与由搜索条件“给我看在银座便宜的荞麦面餐馆”中的“便宜”所指示的价格范围不同。然而,在该方法中,以相同的数值范围来取代在两个搜索条件中所包含的主观表达“便宜”。因此,用户需要参考搜索结果、再次执行指定搜索条件的操作。

[0006] 本公开被做出,以提供新颖的且改进的、并且能够根据含有主观表达的自然语句的内容将主观表达转换为适合用户意图的数值范围的搜索装置、搜索方法和程序。

[0007] 一些实施例涉及一种装置,其包括:至少一个处理器;以及至少一个存储器,其存储处理器可执行指令,处理器可执行指令被配置为使得至少一个处理器进行以下动作:接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询;基于一个或多个对象

元素执行对内容的第一搜索；确定从第一搜索返回的结果的数值特征的分布，数值特征对应于一个或多个主观元素中的第一主观元素；确定对于第一主观元素的一个或多个比例阈值；通过对分布应用一个或多个比例阈值来选择数值特征的范围；以及基于所选择的数值特征的范围来执行对内容的第二搜索。

[0008] 一些实施例涉及一种方法，该方法包括以下步骤：接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；基于一个或多个对象元素执行对内容的第一搜索；利用至少一个处理器来确定从第一搜索返回的结果的数值特征的分布，所述数值特征对应于一个或多个主观元素中的第一主观元素；确定对于第一主观元素的一个或多个比例阈值；通过对分布应用一个或多个比例阈值来选择数值特征的范围；以及基于所选择的数值特征的范围来执行对内容的第二搜索。

[0009] 一些实施例涉及编码有所述计算机可执行指令的至少一个计算机可读存储介质，其中当计算机可执行指令被执行时，实现包括以下步骤的方法：接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；基于一个或多个对象元素执行对内容的第一搜索；确定从第一搜索返回的结果的数值特征的分布，数值特征对应于一个或多个主观元素中的第一主观元素；确定对于第一主观元素的一个或多个比例阈值；通过对分布应用一个或多个比例阈值来选择所述数值特征的范围；以及基于所选择的数值特征的范围来执行对内容的第二搜索。

[0010] 一些实施例涉及一种装置，其包括：用于接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询的设备；用于基于一个或多个对象元素执行对内容的第一搜索的设备；用于确定从第一搜索返回的结果的数值特征的分布的设备，数值特征对应于一个或多个主观元素中的第一主观元素；用于确定对于第一主观元素的一个或多个比例阈值的设备；用于通过对分布应用一个或多个比例阈值来选择数值特征的范围的设备；以及用于基于所选择的数值特征的范围来执行对内容的第二搜索的设备。

[0011] 一些实施例涉及一种装置，其包括：至少一个处理器；以及至少一个存储器，其存储处理器可执行指令，处理器可执行指令被配置为使得至少一个处理器进行以下动作：接收包括至少一个或多个主观元素和一个或多个对象元素的搜索查询；显示从基于一个或多个对象元素的对内容的第一搜索返回的结果的数值特征的范围，数值特征对应于一个或多个主观元素中的第一主观元素；接收用于修改数值特征的范围的用户输入；以及显示基于修改后的数值特征的范围的、对内容的第二搜索的结果。

附图说明

[0012] 图 1 是图示了根据本公开的实施例的信息搜索方法的概述的说明图。

[0013] 图 2 是图示了一般信息搜索系统的问题（课题 1：如果对主观表达指派了数值范围，实际上不能得到即时搜索结果）的说明图。

[0014] 图 3 是图示了一般信息搜索系统的问题（课题 2：如果用数据库管理详细的组合，则数据库变得超出实际得大）的说明图。

[0015] 图 4 是图示了能够实现根据实施例的信息搜索方法的搜索装置的功能配置的说明图。

[0016] 图 5 是图示了在根据实施例的信息搜索方法中使用的查询生成词典数据库的结

构的说明图。

[0017] 图 6 是图示了在根据实施例的信息搜索方法中使用的标准相对值数据库的结构的说明图。

[0018] 图 7 是图示了在根据实施例的信息搜索方法中使用的个人相对值数据库的结构的说明图。

[0019] 图 8 是图示了在根据实施例的信息搜索方法中包括的使主观表达数值化的方法的说明图。

[0020] 图 9 是图示了在根据实施例的信息搜索方法中包括的使主观表达数值化的方法的说明图。

[0021] 图 10 是图示了在根据实施例的信息搜索方法中使用的内容元数据数据库的结构的说明图。

[0022] 图 11 是图示了在根据实施例的信息搜索方法中包括的、在搜索条件修改步骤中使用的条件修改屏幕的显示配置的说明图。

[0023] 图 12 是图示了为实现根据实施例的信息搜索方法而执行的处理的流程的说明图。

[0024] 图 13 是图示了为实现根据实施例的信息搜索方法而执行的处理的流程的说明图。

[0025] 图 14 是图示了能够实现根据实施例的第一变形例的信息搜索方法的信息搜索装置的功能配置的说明图。

[0026] 图 15 是图示了在根据实施例的第一变形例的信息搜索方法中使用的语义概念词典数据库的配置的说明图。

[0027] 图 16 是图示了为实现根据实施例的第一变形例的信息搜索方法而执行的处理的流程的说明图。

[0028] 图 17 是图示了为实现根据实施例的第一变形例的信息搜索方法而执行的处理的流程的说明图。

[0029] 图 18 是图示了能够实现根据实施例的第二变形例的信息搜索方法的信息搜索装置的功能配置的说明图。

[0030] 图 19 是图示了为实现根据实施例的第二变形例的信息搜索方法而执行的处理的流程的说明图。

[0031] 图 20 是图示了为实现根据实施例的第二变形例的信息搜索方法而执行的处理的流程的说明图。

[0032] 图 21 是说明了能够实现根据实施例的搜索装置的功能的信息处理装置的硬件配置的说明图。

具体实施方式

[0033] 下文，将参考附图详细描述本公开的优选实施例。注意，在本说明书和附图中，利用相同的附图标记来指示具有相同功能和结构的结构元件，并且省略这些结构元件的重复说明。

[0034] 【说明流程】

[0035] 这里,将简要地描述下文中所描述的本公开的实施例的说明流程。首先,将参考图 1 描述根据本实施例的信息搜索方法的概述。其次,将参考图 2 和图 3 描述一般信息搜索系统的问题。

[0036] 然后,将参考图 4 描述能够实现根据本实施例的信息搜索方法的搜索装置 100 的功能配置。在该描述中,将参考图 5 至图 11 详细描述在根据本实施例的信息搜索方法中使用的各种数据库的结构、将主观表达数据化的方法或者修改搜索条件的方法。此外,将参考图 12 和图 13 描述为实现根据本实施例的信息搜索方法、由搜索装置 100 执行的处理的流程。

[0037] 接下来,将参考图 14 和图 15 描述根据本实施例的第一变形例的搜索装置 130 的功能配置。此外,将参考图 16 和图 17 描述为实现根据本实施例的第一变形例的信息搜索方法、由搜索装置 130 执行的处理的流程。然后,将参考图 18 描述根据本实施例的第二变形例的搜索装置 150 的功能配置。此外,将参考图 19 和图 20 描述为实现根据本实施例的第二变形例的信息搜索方法、由搜索装置 150 执行的处理的流程。

[0038] 接下来,将参考图 21 描述能够实现搜索装置 100、130 和 150 的功能的信息处理装置的硬件配置。最后,将总结本实施例的技术精神,并简要地描述能够从本实施例的技术精神得到的效果。

[0039] (说明项目)

[0040] 1 : 简介

[0041] 1-1 : 实施例的概述

[0042] 1-2 : 信息搜索系统的主题

[0043] 2 : 实施例

[0044] 2-1 : 搜索装置 100 的功能配置

[0045] 2-2 : 信息搜索方法 (搜索装置 100 的操作)

[0046] 3 : 第一变形例

[0047] 3-1 : 搜索装置 130 的功能配置

[0048] 3-2 : 信息搜索方法 (搜索装置 130 的操作)

[0049] 4 : 第二变形例

[0050] 4-1 : 搜索装置 150 的功能配置

[0051] 4-2 : 信息搜索方法 (搜索装置 150 的操作)

[0052] 5 : 硬件配置

[0053] 6 : 总结

[0054] <1. 简介>

[0055] 在详细描述根据本公开的实施例之前,将简要说明根据实施例的技术概述以及一般信息搜索系统的问题。

[0056] [1-1 : 实施例的概述]

[0057] 将参考图 1 描述根据本实施例的技术的概述。根据本实施例的技术涉及信息搜索系统。具体而言,根据本实施例的技术涉及,当自然语句被作为搜索条件输入时、基于搜索条件输出适合于用户意图的搜索结果的方法。此外,根据本实施例的技术涉及,在含有将被转换为数值范围的主观表达的自然语句被输入的情况下、用于将主观表达转换为适当的数

值范围并且有效地利用含有数值范围的搜索条件来执行搜索的机制。

[0058] 例如,将考虑如图 1 所示、自然语言查询“给我看在银座便宜并且美味的寿司餐厅”被输入的情况。在这种情况下,“便宜”和“美味”是将被转换为数值范围的主观表达。在一般信息搜索系统中,主观表达被识别为单纯的关键词,并且含有用语“便宜”或“美味”的信息被作为搜索结果输出。然而,在根据本实施例的技术中,“便宜”被转换为适当的价格范围并且“美味”被转换为适当的评价值范围。通过主观表达被转换为价格范围或评价值范围的搜索条件来搜索信息。

[0059] 此外,根据本实施例的技术包括用于使用户能够确认或修改与主观表达对应的数值范围的机制。例如,如图 1 所示,如果主观表达“便宜”被转换为数值范围“预算 :0 到 5000 日圆”,则将该数值范围呈现给用户,使得用户可改变该数值范围。类似地,如果主观表达“美味”被转换为数值范围“评价值 :2.5 到 5.0”,则将该数值范围呈现给用户,使得用户能够修改该数值范围。此外,搜索条件的一部分可以被删除或者每个搜索条件可以被加重。

[0060] 此外,根据本实施例的技术包括,用于使某个搜索结果的数值范围的修改能够被反映到另一个搜索结果的机制。例如,在以自然语句“给我看在银座便宜并且美味的寿司餐厅”的输入作为搜索条件的搜索结果中,如果针对主观表达“便宜”的数值范围“预算 :0 至 5000 日圆”被修改,则修改内容也被反映在以自然语句“给我看在银座便宜的荞麦面餐厅”的输入作为搜索条件的搜索结果上。即,根据本实施例的技术具有用于使用户能够消除在用户的“便宜”的概念或用户所认为的“便宜”的感觉与搜索结果所包含的“便宜”的概念或感觉之间的差别的机制。

[0061] 因此,根据本实施例的技术涉及这样的信息搜索系统,其用于将主观表达转换为适当的数值范围以产生搜索条件并且基于该搜索条件有效地执行搜索。此外,根据本实施例的技术提供了能够在不使用庞大的数据库的条件下、根据搜索条件将主观表达转换为适当的数值范围的机制。此外,根据本实施例的技术使用户能够修改用于将主观表达转换为数值范围的信息,并且调整由主观表达所表示的概念或感觉。根据本实施例的技术使得搜索机制能够基于主观表达个人化。

[0062] 已经简要地描述了根据本实施例的技术的概述。稍后将描述根据本实施例的信息搜索方法或将主观表达数值化的方法。

[0063] [1-2 : 信息搜索系统的问题]

[0064] 下面,将简要地描述一般信息搜索系统的问题。大多数一般信息搜索系统将包含在自然语句中的主观表达识别为关键字,并且输出含有关键字的信息作为搜索结果。然而,含有主观表达的信息可能不一定是反映主观表达的意图的信息。因此,这样的信息搜索系统的大多数搜索结果没有反映用户的意图。已经发明了只有与数值范围相关联的主观表达的数据库并且利用数据库将主观表达转换为数值范围以重构搜索条件的方法。

[0065] 在该方法的情况下,例如,如图 2 所示,即使当输入自然语句“给我看在银座便宜的寿司餐厅”时或输入自然语句“给我看在涩谷便宜的寿司餐厅”时,以相同的数值范围(在图 2 的示例中,“预算 :0 到 5000 日圆”)取代与主观表达“便宜”对应的数值范围。现在,假定在银座寿司的市场价格是 8000 日圆并且假设在涩谷寿司的市场价格是 4000 日圆。在该示例中,响应于自然语句“给我看在银座便宜的寿司餐厅”的输入,在银座与“预算 :0 到 5000 日圆”对应的寿司餐厅的信息被作为搜索结果输出。即,实际上输出了在银座“便宜”

的寿司餐厅的信息。

[0066] 同时,响应于自然语句“给我看在涩谷便宜的寿司餐厅”的输入,在涩谷与“预算:0到5000日圆”对应的寿司餐厅的信息也被作为搜索结果输出。即,与“预算:4000到5000日圆”对应的寿司餐厅信息也被作为搜索结果输出。因此,实际上,也输出了在涩谷具有“昂贵”的寿司的寿司餐厅的信息。此外,当输入“给我看在银座便宜的荞麦面餐厅”时,同样得到不适当的搜索结果。为了防止这样的不适当的输出,需要持有其中地点或目标元素(影响主观表达的分析的元素)以及主观表达和数值范围相关联的数据库,如图3所示。

[0067] 然而,从图3所示的数据库的示例中可见,存在大量对主观表达的分析有影响的数据以及大量与各个主观表达保持相关的记录。结果,由于找到想要的记录需要时间,因此降低了搜索效率。此外,由于需要具有大存储容量的存储器来保持大量数据,因此成本大大增加。此外,由于很难将庞大的数据库提供给各个用户,因此很难使得主观表达与数值范围之间的关联关系个人化。

[0068] 因此,利用主观表达与数值范围相关联的数据库来重构搜索条件的方法涉及如上上述问题。本发明人已经发明了用于解决该问题并且提高以含有主观表达的自然语句为搜索条件的搜索的效率的机制。下文,将详细描述该机制。

[0069] <2:实施例>

[0070] 将描述本公开的实施例。本公开涉及在不使用庞大的数据库的情况下、将主观表达转换为适当的数值范围以执行搜索的信息搜索方法。

[0071] [2-1:搜索装置100的功能配置]

[0072] 首先,将参考图4描述能够实现根据本实施例的信息搜索方法的搜索装置100的功能配置。图4是图示了根据本实施例的搜索装置100的功能配置的说明图。

[0073] 如图4所示,搜索装置100包括语音识别单元101、自然语言分析单元102、语言分析词典数据库103、查询生成单元104以及查询生成词典数据库105。此外,搜索装置100还包括内容搜索单元106、内容元数据数据库107、数值化单元108、标准相对值数据库109以及个人相对值数据库110。此外,搜索装置100包括搜索信息输出单元111、搜索条件修改单元112和相对值反向计算单元113。

[0074] 然而,假设语言分析词典数据库103、查询生成词典数据库105、内容元数据数据库107、标准相对值数据库109以及个人相对值数据库110被存储在未示出的存储设备中。或者,语言分析词典数据库103、查询生成词典数据库105、内容元数据数据库107、标准相对值数据库109以及个人相对值数据库110可以设置在搜索装置100的外部。此外,搜索装置100可具有未示出的输入设备和输出设备。

[0075] 输入设备包括例如键盘、鼠标、开关、触摸屏、触摸板或麦克风。此外,输出设备包括例如显示器或扬声器。在下文中,虽然将以示例的方式描述通过语音输入搜索条件的方法,但是输入搜索条件的方法根据包括在搜索装置100中的输入设备的类型适当地改变。此外,虽然为了说明假设了显示并输出搜索结果或搜索条件的方法,但是显示并输出搜索结果或搜索条件的方法可根据包括在搜索装置100中的输出设备的类型改变为例如语音输出。应该明白,这样的改变包括在本实施例的技术范围内。

[0076] 在下文中,将以从通过语音的搜索条件输入到搜索结果输出的流程来描述搜索装置100的功能和操作。

[0077] 首先,由用户利用输入设备输入语音。在这种情况下,用户像对朋友说话一样输入搜索条件。例如,用户通过语音输入自然语言查询“给我看在银座便宜的寿司餐厅”作为搜索条件。如果输入了语音,则语音信号从输入设备中输出并且被输入到语音识别单元 101。如果输入了语音信号,则语音识别单元 101 将语音信号转换为文本。语音识别单元 101 将通过转换语音信号获得的文本输入到自然语言分析单元 102。

[0078] 如果输入了文本,则自然语言分析单元 102 执行诸如形态分析之类的自然语言分析,以将被输入的文本转换为诸如词类、属性、短语和从属关系等的信息。例如,自然语言分析单元 102 将自然语句“给我看在银座便宜的寿司餐厅”分成短语“在银座 / 便宜的 / 寿司餐厅 / 给我看”。自然语言分析单元 102 输出诸如“银座 : 专有名词, 地点名称”“寿司 : 名词, 食物名称”(“词表示 : 词类, 属性”)之类的分析结果。当执行这些自然语言分析时,自然语言分析单元 102 使用语言分析词典数据库 103。

[0079] 例如,将词与概念连接的信息或者指示词的分类或使用的信息存储在语言分析词典数据库 103 中。在后面的说明中,将由自然语言分析单元 102 分开的单独的短语(例如,一个或多个词)称作元素。

[0080] 由自然语言分析单元 102 从自然语句抽取的元素被输入到查询生成单元 104。如果输入了构成自然语句的元素,则查询生成单元 104 利用除了与将被转换为数值范围的主观表达对应的那些元素以外的元素、产生用于执行暂定搜索的查询(下文成为暂定查询)。首先,查询生成单元 104 抽取被包含在搜索条件中的自然语句的元素。例如,如果搜索餐厅信息,则查询生成单元 104 抽取诸如“地点”、“类型”和“预算”之类的适合于搜索条件的自然语句的元素。

[0081] 在这种情况下,查询生成单元 104 考虑预先在查询生成词典数据库 105(例如,见图 5)中登记的“词表示,词类,属性”情况、从属关系等的映射信息,抽取适合于搜索条件的自然语句的元素。例如,如果自然语句“给我看在银座便宜的寿司餐厅”被输入为搜索条件,则查询生成单元 104 抽取“银座”作为适合于搜索条件“地点”的元素,并且抽取“寿司”作为适合于搜索条件“类型”的元素。此外,查询生成单元 104 抽取“便宜”作为适合于搜索条件“预算”的元素。在这样的情况下,查询生成单元 104 确认主观元素“便宜”是将被转换为数值范围的主观表达。

[0082] 如果抽取了适合各个搜索条件的元素,查询生成单元 104 将除了主观表达以外的元素组合以产生暂定查询。例如,查询生成单元 104 以客观元素“地点 : 银座”和“类型 : 寿司”为搜索条件来产生暂定查询。由查询生成单元 104 产生的暂定查询被输入到内容搜索单元 106。如果输入了暂定查询,内容搜索单元 106 从内容元数据数据库 107(例如,见图 10)抽取适合于被输入的暂定查询的信息。例如,如果暂定查询使用“地点 : 银座”和“类型 : 寿司”作为搜索条件,则内容搜索单元 106 可以抽取如图 9 所示的搜索结果。图 9 示出了基于暂定查询和数值化方法的搜索结果的示例。

[0083] 利用暂定查询由内容搜索单元 106 抽取的搜索结果被输入到数值化单元 108。此外,将被转换为数值范围的主观表达被从查询生成单元 104 输入到数值化单元 108。如果输入了基于暂定查询的搜索结果和主观表达,则数值化单元 108 基于所输入搜索结果、将所输入的主观表达转换为数值范围。在这种情况下,数值化单元 108 利用存储在标准相对值数据库 109 或个人相对值数据库 110 中的信息、将主观表达转换为数值范围。由数值化单

元 108 从主观表达转换来的数值范围被输入到查询生成单元 104。

[0084] 这里,将更详细地描述将主观表达转换为数值范围的方法。

[0085] 首先,将参考图 6 和图 7 描述标准相对值数据库 109 和个人相对值数据库 110 的配置。如图 6 所示,标准相对值数据库 109 是以主观表达与相对值相关联的数据库。此外,标准相对值数据库 109 的记录被假设为预先设置。例如,主观表达“便宜”与相对值 0% 至 30% 相关联。此外,如图 7 所示,个人相对值数据库 110 是指示各个不同用户的主观表达与相对值之间的关联关系的数据库。此外,如果相对值被用户修改,则个人相对值数据库 110 的记录被增加或修改。

[0086] 相对值表示用户利用主观表达想要表达的感觉(例如,合理价格的感觉)。例如,与主观表达“便宜”对应的相对值 0% 至 30%,意味着用户感觉在一定条件下抽取的价格范围的列表中较低的 0 至 30 百分位的价格范围是“便宜”。即,标准相对值数据库 109 和个人相对值数据库 110 利用相对值定量地表示主观表达的用户感觉。如果使用该相对值,则即使在对于相同主观表达的用户感觉根据搜索条件发生变化时,主观表达也能被转换为适合于用户感觉的数值范围。

[0087] 图 8 示出了利用相对值和分布使主观表达数值化的方法:相对值 [0%, 10%] 的情况。具体而言,如图 8 所示,执行了基于暂定查询从搜索结果获取数值特征(在图 8 的示例中,价格范围)的分布、并且从该分布中确定适合于相对值的数值范围的处理。例如,如果自然语句中包含主观表达“非常便宜”,则数值化单元 108 从标准相对值数据库 109(见图 6)中抽取与该主观表达对应的相对值(0 至 10%)。接下来,数值化单元 108 基于暂定查询从搜索结果获取价格范围的分布。如果获取了价格范围的分布,则数值化单元 108 确定与在所获取的价格范围分布中从便宜的一个顺次累加的搜索结果的数目达到 10% 处的界限对应的价格(在图 8 的示例中,3200 日圆)。通过对分布应用 10% 比例阈值,数值化单元 108 确定与主观表达对应的数值范围是从 0 到 3200 日圆。

[0088] 因此,数值化单元 108 基于暂定查询、利用搜索结果和相对值来确定与主观表达对应的数值范围。数值化单元 108 将主观表达转换为数值范围。虽然这里示出了使用标准相对值数据库 109 的示例,但是如果对应于主观表达的记录在个人相对值数据库 110 中,则使用个人相对值数据库 110。

[0089] 现在,如果通过数值化数据库 108 将对应于主观表达数值范围输入到查询生成单元 104,则查询生成单元 104 产生将所输入的数值范围添加到搜索条件的最终查询。即,最终查询包含与搜索条件对应的数值范围以及包含在暂定查询中的主观表达。由查询生成单元 104 产生的最终查询被输入到内容搜索单元 106 中。如果输入了最终查询,则内容搜索单元 106 基于所输入的最终查询搜索内容元数据数据库 107 的信息。内容搜索单元 106 将基于最终查询的搜索结果和搜索条件输入到搜索信息输出单元 111 中。

[0090] 如果输入了基于最终查询的搜索结果和搜索条件,则搜索信息输出单元 111 输出所输入的搜索结果和搜索条件。例如,搜索信息输出单元 111 显示基于最终查询的搜索结果,如图 1 所示。如果在输出了基于最终查询的搜索结果以后、用户指示转换到搜索条件修改屏幕,则搜索信息输出单元 111 显示搜索条件。具体而言,搜索信息输出单元 111 显示与主观表达对应的数值范围。此外,搜索信息输出单元 111 显示用于操作的操作对象,使得与主观表达对应的数值范围可被修改,如图 11 所示。

[0091] 在图 11 的示例中,显示了三个区域 A1、A2 和 A3。在区域 A1 中显示作为搜索条件输入的自然语句“银座车站附近便宜的寿司餐厅”。此外,在区域 A1 中显示用于基于修改后的搜索条件执行重新搜索的按钮 Obj1。如果用户在修改搜索条件后按压该按键 Obj1,则可基于修改后的搜索条件执行搜索。

[0092] 此外,在区域 A2 和区域 A3 中显示,用于修改作为主观表达的对象的短语、与主观表达对应的数值范围以及搜索条件的对象 Obj2 和 Obj3。对象 Obj2 是用于切换搜索条件的有效 / 无效的复选框。例如,如果在区域 A2 中的对象 Obj2 未选中,则在区域 A2 中显示的搜索条件变为无效。同时,对象 Obj3 是用于修改数值范围的滑动条。例如,如果滑动在区域 A2 中的对象 Obj3 的条,则数值范围“距离”改变。或者,对象 Obj2 和 Obj3 可以分别不是复选框和滑动条。

[0093] 如果利用这样的用户界面修改了搜索条件,则修改的内容被输入到搜索条件修改单元 112。例如,如果与主观表达对应的数值范围被修改,则搜索条件修改单元 112 将修改后的数值范围输入到相对值反向计算单元 113。如果输入了修改后的数值范围,则相对值反向计算单元 113 从数值化数据库 108 获取从基于暂定查询的搜索结果获取的数值范围的分布,并且基于所获取的分布计算与修改后的数值范围对应的相对值。例如,如果在图 8 的示例中确定的 0 到 3200 日圆的数值范围被修改为 0 至 2500 日圆,则相对值反向计算单元 113 计算在分布中价格对应于 0 至 2500 日圆的搜索结果的数目的百分比,并且基于该百分比计算相对值。从而,将上限从 3200 日圆变为 2500 日圆的用户输入将与“便宜”相关联的比例阈值从一个相对值(例如,百分比)改变到另一个。

[0094] 虽然这里所述的示例示出了修改数值范围的上限,但是也可以修改数值范围的下限。或者,数值范围的上限和下限都可以被修改。例如,用户可以将 0 至 3200 日圆的数值范围修改为 1500 至 3000 日圆。在这种情况下,相对值反向计算单元 113 计算在分布中价格对应 0 至 1500 日圆的搜索结果的数目的百分比 Q1,以及在分布中价格对应 0 至 3000 日圆的搜索结果的数目的百分比 Q2,并且确定相对值 Q1 和 Q2。因此,可以由用户修改和 / 或指定多个比例阈值,并且可任意地扩大或缩小搜索条件的数值范围。

[0095] 如果基于修改后的数值范围计算了相对值,则相对值反向计算单元 113 利用所计算的相对值更新个人相对值数据库 110 的记录。然而,在个人相对值数据库 110 中没有对应于所计算的相对值的主观表达的记录,则相对值反向计算单元 113 将该记录添加到个人相对值数据库 110 中。在这种情况下,相对值反向计算单元 113 基于用户 ID 单独地识别用户并且添加或更新记录,如图 7 所示。此外,如果用户指示重新搜索,则基于存储在个人相对值数据库 110 中的更新后的记录、通过数值范围再次产生最终查询,并且基于最终查询搜索信息。

[0096] 已经以从搜索条件输入到搜索结果输出的处理的流程、描述了搜索装置 100 的功能和操作。如上所述,搜索装置 100 基于相对值将主观表达转换为数值范围。结果,可得到适当地考虑了随搜索条件变化的主观表达的分析的搜索结果。此外,由于通过相对值代表主观表达,因此数据库的大小较小并且用于将主观表达转换为数值范围的信息可切实地个性化。

[0097] (补充说明 1:产生暂定查询的方法)

[0098] 在说明中,已经介绍了将除了主观表达以外的自然语句的元素组合以产生暂定查

询的方法。然而,如果在自然语句中包含多个主观表达,则可利用下面的方法产生暂定查询。例如,考虑如下方法,其中如图 6 所示准备用于针对主观表达的暂定查询的数值范围,并且用于暂定查询的该数值范围包含在构成暂定查询的搜索条件中。如果使用该方法,则通过排出认为不必要的搜索结果,可抑制基于暂定查询的搜索结果的数目。结果,提高了当基于暂定查询从搜索结果获取分布时的处理效率。作为利用用于暂定查询的数值范围来选择主观表达的方法,例如,可以考虑优选选择后面将描述的具有较低权重的主观表达。

[0099] (补充说明 2: 权重设置方法)

[0100] 此外,上面已经讨论了将用户的修改动作反映到搜索结果的方法。然而,将频繁修改的搜索条件认为是对用户具有相对较高的重要性的搜索条件。同时,将无效的搜索条件认为是对用户具有相对较低的重要性的搜索条件。随着将搜索条件的重要性反映在搜索结果中,预期可实现更精确的搜索。将介绍当添加或修改个人相对值数据库 110 的记录时、根据用户的修改动作来分配权重的方法。

[0101] 例如,如果基于暂定查询的搜索结果的数目大于给定数目,则用户试图通过调整搜索条件来限制搜索结果的数目。在这种情况下,认为用户改变具有相对较高的重要性的搜索条件。即,认为被改变以减小搜索结果的数目的搜索条件是具有相对较高的重要性的搜索条件。例如,认为被改变以减小数值范围的宽度的搜索条件是具有相对较高的重要性的搜索条件。因此,将这样的搜索条件的权重设定为较大值。

[0102] 另一方面,如果基于暂定查询的搜索结果的数目小于给定数目,则用户试图通过调整搜索条件或使搜索条件中的一些作废、来增大搜索结果的数目。在这种情况下,认为用户改变或废止具有相对较低重要性的搜索条件。例如,用户试图通过扩大搜索条件的数值范围或使搜索条件作废来去除限制。这样的搜索条件的改变可以说是用户对条件的宽松妥协的结果。即,被改变以增大搜索结果的数目的搜索条件可以说是那些被认为用户允许妥协的搜索条件。因此,将这样的搜索条件的权重设定为较小值。

[0103] 此外,认为具有较高变化频率的搜索条件是被用户注意的搜索条件。因此,将具有较高变化频率的搜索条件的权重设定为相对较大值,或者修改为具有较大值。此外,认为具有较大改变量的搜索条件是对用户具有相对较低重要性的搜索条件。即,可以认为可被极大地改变的搜索条件是用户不太强烈需要的搜索条件。因此,将具有较大改变量的搜索条件的权重设定为相对较小值或修改为具有较小值。从而,根据用户改变搜索条件的方法来设置与各个搜索条件对应的权重。

[0104] 该权重被认为是当执行搜索时各个搜索条件的重量。结果,能够得到对用户重要的极大地考虑了搜索条件的搜索结果。权重设置方法不限于此,而是可使用各种设置方法,只要将权重设置为反映用户所认为的重要性即可。

[0105] (补充说明 3: 更新相对值的方法)

[0106] 在上述说明中,已经介绍了直接在个人相对值数据库 110 中登记由相对值反向计算单元 113 计算出的相对值的方法。然而,如果相对值由于用户对数值范围的修改而极大地改变,在个人相对值数据库 110 中登记的相对值可能受到限制。即,由于认为主观表达的分析不会突然极大地改变,因此,相对值的极端变化没有反映在个人相对值数据库 110 中。例如,可以累积给定数量的日志,并且计算日志的平均值并将其登记在个人相对值数据库 110 中。

[0107] 此外,由用户输入搜索满意度,并且基于所输入的搜索满意度、经过权重平均的相对值可登记在个人相对值数据库 110 中。例如,如果第一次改变后的相对值是 0% 至 23% (搜索满意度是 50),第二次改变后的相对值是 5% 至 23% (搜索满意度是 70),并且第三次改变后的相对值是 5% 至 30% (搜索满意度是 80),则权重的平均值是 $(0*50+5*70+5*80)/200-(23*50+23*70+30*80)/200$ 。该值被登记在个人相对值数据库 110 中,从而防止相对值被极大地改变。

[0108] 此外,搜索满意度可由用户手动地输入,或可基于例如屏幕保留时间自动地计算。作为用户手动输入搜索满意度的方法,考虑了将搜索满意度作为数值输入的方法,以及使得指示用户对单个搜索结果是否满意的复选框被选中并且根据选中结果计算搜索满意度的方法。此外,作为屏幕保留时间,例如考虑详细的搜索内容已被参考的时间,或者诸如电影或音乐等内容已被使用的数据。通过使用这样的方法可以得到搜索满意度。

[0109] 上面已经描述了根据本实施例的搜索装置 100 的功能配置。

[0110] [2-2 :信息搜索方法 (搜索装置 100 的操作)]

[0111] 下面,将参考图 12 和图 13 描述根据本实施例的信息搜索方法。图 12 和图 13 是示出了为实现根据本实施例的信息搜索方法而执行的处理的流程的流程图。

[0112] 首先,参考图 12。如图 12 所示,首先,由用户通过语音输入搜索条件 (S101)。如果搜索条件由用户通过语音输入,则搜索装置 100 利用语音识别单元 101 的功能将所输入的语音转换为文本 (S102)。搜索装置 100 然后利用自然语言分析单元 102 的功能分析被转换为文本的搜索条件 (例如,自然语句查询) (S103)。搜索装置 100 然后利用查询生成单元 104 的功能、组合由步骤 S103 中的分析得到的自然语句的元素以产生暂定查询 (S104)。在这种情况下,查询生成单元 104 组合除了将被转换为数值范围的主观表达以外的自然语句的元素 (例如,对象元素),以产生暂定查询。

[0113] 接下来,搜索装置 100 利用内容搜索单元 106 的功能,基于暂定查询执行搜索,并且基于暂定查询获取搜索结果。搜索装置 100 然后利用数值化单元 108 的功能、确认与将被转换为数值范围的主观表达对应的记录是否已登记在个人相对值数据库 110 中 (S105)。如果该记录已经登记在个人相对值数据库 110 中,则在搜索装置 100 中的处理前进到步骤 S106 中。另一方面,如果该记录还没有登记在个人相对值数据库 110 中,则在搜索装置 100 中的处理前进到步骤 S107。

[0114] 如果处理前进到步骤 S106,则搜索装置 100 从个人相对值数据库 110 获取对应于主观表达的相对值 (S106)。另一方面,如果处理前进到步骤 S107,则搜索装置 100 从标准相对值数据库 109 获取对应于主观表达的相对值 (S107)。已经获取了相对值的搜索装置 100 利用数值化单元 108 的功能、从基于暂定查询的搜索结果获取数值特征的分布,并且基于该分布确定对应于相对值的数值范围 (例如,通过应用基于相对值相对于分布的一个或多个比例阈值)。搜索装置 100 利用查询生成单元 104 的功能、产生将所确定的数值范围添加到搜索条件的最终查询 (S108),并且处理前进到步骤 A。

[0115] 下面,参考图 13。如图 13 所示,其中处理已经前进到步骤 A 的搜索装置 100 利用内容搜索单元 106 的功能基于在步骤 S108 中产生的最终查询执行搜索 (S109)。搜索装置 100 然后利用搜索信息输出单元 111 的功能输出基于最终查询的搜索结果 (S110)。此外,搜索装置 100 利用搜索信息输出单元 111 的功能、显示最终查询所包含的搜索条件以及用

于修改搜索条件的界面，并且接收诸如用户操作等的用户输入 (S111)。搜索装置 100 然后判断搜索条件是否已被用户修改 (S112)。

[0116] 如果搜索条件已经被修改，则在搜索装置 100 中的处理前进到步骤 S113。另一方面，如果搜索条件还没有被修改，则搜索装置 100 结束一系列处理。如果处理前进到步骤 S113，则搜索装置 100 利用相对值反向计算单元 113 的功能、从修改后的数值范围计算相对值 (S113)。在这种情况下，相对值反向计算单元 113 利用从基于暂定查询的搜索结果获取的分布计算相对值。

[0117] 接下来，搜索装置 100 使用相对值反向计算单元 113 的功能、利用在步骤 S113 中计算的相对值来更新个人相对值数据库 110 (S114)。然而，如果在个人相对值数据库 110 中没有对应的记录，则相对值反向计算单元 113 将与所计算的相对值对应的主观表达的记录添加到个人相对值数据库 110 中。

[0118] 在搜索装置 100 结束步骤 S114 的处理以后，处理前进到步骤 B。即，搜索装置 100 中的处理返回到图 12 所示的步骤 S108，并且再次从步骤 S108 执行处理。搜索装置 100 反复地执行步骤 S108 到 S114，直到由用户进行的搜索条件修改完成为止，并且在搜索条件修改结束以后结束一系列处理。此外，如果在作为搜索条件输入的自然语句中含有多个主观元素，则对每个主观表达执行步骤 S104 以后的处理。

[0119] 上面已经描述了根据本实施例的信息搜索方法。

[0120] <3 : 第一变形例 >

[0121] 下面，将描述本实施例的第一变形例。本变形例涉及在搜索结果上反映用户的精细条件的方法。

[0122] <3-1 : 搜索装置 130 的功能配置 >

[0123] 首先，将参考图 14 描述本变形例的搜索装置 130 的功能配置。图 14 是图示了根据本变形例的搜索装置 130 的功能配置的说明图。与搜索装置 100 的那些具有基本相同的功能的组件被分配了相同的附图标记并且将省略其详细描述。

[0124] 如图 14 所示，搜索装置 130 包括语音识别单元 101、自然语言分析单元 102、语言分析词典数据库 103、查询生成单元 104 和查询生成词典数据库 105。此外，搜索装置 130 包括内容搜索单元 106、内容元数据数据库 107、数值化单元 131、标准相对值数据库 109 和个人相对值数据库 110。此外，搜索装置 130 包括搜索信息输出单元 111、搜索条件修改单元 112、相对值反向计算单元 113 和语义概念词典数据库 132。

[0125] 与搜索装置 100 的区别是数值化单元 131 的功能以及语义概念词典数据库 132 的出现。将仅详细描述数值化单元 131 和语义概念词典数据库 132 的配置。

[0126] 如上所述，本变形例的目的是在搜索结果上反映用户的精细条件。例如，假设某个用户认为“对于寿司味道重要而对于酒馆低价格重要”。假设用户已经输入了自然语句“给我看在银座便宜并且美味的寿司餐厅”作为搜索条件，并且已经得到了通常的搜索结果。此外，如果基于搜索结果修改了搜索条件，则可以认为与主观表达“便宜”和“美味”对应的相对值是基于“寿司餐厅”的条件被修改的。然后，如果自然语句“给我看在银座便宜并且美味的酒馆”被输入并且执行了搜索，则搜索是基于反映了针对“寿司餐厅”的条件的相对值来执行的。

[0127] 因此，如果存在用户的精细条件，则可能不能得到适合于用户的主观表达的分析

的搜索结果。当然,如果修改搜索条件,则可以获得适当的搜索结果,但是一个一个地修改搜索条件是很复杂的。在本变形例中,提出了通过群集(cluster)主观表达的对象(自然语句的元素)而在搜索结果上反映这样的用户条件的方法。用于实现该方法的组件是数值化单元131和语义概念词典数据库132。

[0128] 数值化单元131利用具有如图15所示的结构(数据结构的概念图)的语义概念词典数据库132来群集对象,并且根据对象识别概念的区别(主观表达的分析)。首先,数值化单元131利用从属性分析来从自然语句中识别对象。例如,对于自然语句“给我看在银座便宜并且美味的寿司餐厅”,数值化单元131识别出“寿司餐厅”是对象。接下来,数值化单元131在个人相对值数据库110中存储对象、主观表达和相对值的组合。在该示例中,具有寿司餐厅、便宜和0至35%的组合的记录被存储在个人相对值数据库110中。

[0129] 然后,如果自然语句“给我看在涩谷便宜并且美味的酒馆”被用户作为搜索条件输入,假设基于该搜索条件执行搜索。在这种情况下,数值化单元131通过参考存储在个人相对值数据库110中的具有寿司餐厅、便宜和0至35%的组合的记录、利用与主观表达“便宜”对应的相对值0至35%、将主观表达转换为相对值。然后,如果相对值被用户极大地修改(例如,相对值被修改为0至20%),则数值化单元131在个人相对值数据库110中存储具有酒馆、便宜和0至20%的记录。

[0130] 相反,如果相对值没有被用户极大地修改(例如,相对值被修改为0至32%),则利用语义概念词典数据库132的树状结构、使用以寿司餐厅和酒馆两者为子节点的父节点的词来更新关于存储在个人相对值数据库110中的“寿司餐厅”的记录。例如,数值化单元131利用与“寿司餐厅”和“酒馆”的父节点对应的词“餐馆”、用具有餐馆、便宜和0至33.5%的组合的记录来更新原来的记录。数值33.5%是35%和32%的平均值。

[0131] 此外,如果对于一个主观表达存在如上所述的多个记录(例如“酒馆、便宜和0至20%”和“寿司餐厅、便宜和0至35%”),假设以另一个搜索条件(例如,自然语句“给我看在六本木便宜的酒吧”)。在这种情况下,数值化单元131使用记录“酒馆、便宜和0至20%”。这是因为,基于语义概念词典数据库132判断“酒馆”和“酒吧”比“寿司餐厅”和“酒吧”更接近。该判断是基于到作为主题的两个词共同的节点(父节点)的距离来作出的。在该示例的情况下,如果“酒馆”和“酒吧”被对象“饮酒的店”结合在一起,则它们可被判断为比“寿司餐厅”和“酒吧”更接近。

[0132] 因此,数值化单元131利用语义概念词典数据库132来群集对主观表达的分析有影响的对象,或者管理对象或群集单元的相对值。该配置能够实现不依赖于对象的自适应搜索和依赖于对象的具体搜索,从而实现反映用户条件的搜索。

[0133] 上面已经描述了根据本实施例的第一变形例的搜索装置130的功能配置。

[0134] [3-2:信息搜索方法(搜索装置130的操作)]

[0135] 下面,将参考图16和图17描述根据本实施例的第一变形例的信息搜索方法。图16和图17是示出了为实现根据本实施例的第一变形例的信息搜索方法而执行的处理的流程的流程图。

[0136] 首先,参考图16。首先,如图16所示,由用户通过语音输入搜索条件(S131)。如果搜索条件由用户通过语音输入,则搜索装置130利用语音识别单元101的功能将所输入的语音转换为文本(S132)。搜索装置130然后利用自然语言分析单元102的功能分析被

转换为文本的搜索条件（自然语句）(S133)。搜索装置 130 然后利用查询生成单元 104 的功能、组合由步骤 S133 中的分析得到的自然语句的元素以产生暂定查询 (S134)。在这种情况下，查询生成单元 104 组合除了将被转换为数值范围的主观表达以外的自然语句的元素，以产生暂定查询。

[0137] 接下来，搜索装置 130 利用内容搜索单元 106 的功能，基于暂定查询执行搜索，并且基于暂定查询获取搜索结果。搜索装置 130 然后利用数值化单元 131 的功能、确认与将被转换为数值范围的主观表达对应的记录是否已登记在个人相对值数据库 110 中 (S135)。如果该记录已经登记在个人相对值数据库 110 中，则在搜索装置 130 中的处理前进到步骤 S136 中。另一方面，如果该记录还没有登记在个人相对值数据库 110 中，则在搜索装置 130 中的处理前进到步骤 S137。

[0138] 如果处理前进到步骤 S136，则搜索装置 130 从个人相对值数据库 110 获取对应于主观表达的相对值 (S136) 利用数值化单元 131 的功能、通过参考语义概念词典数据库 132、从个人相对值数据库 110 获取含有与包含在搜索条件中的对象最接近的词的记录中所记录的相对值。另一方面，如果处理前进到步骤 S137，则搜索装置 130 从标准相对值数据库 109 获取对应于主观表达的相对值 (S137)。已经获取了相对值的搜索装置 130 利用数值化单元 131 的功能、从基于暂定查询的搜索结果获取数值范围的分布，并且基于该分布确定对应于相对值的数值范围。搜索装置 130 利用查询生成单元 104 的功能、产生将所确定的数值范围添加到搜索条件的最终查询 (S138)，并且处理前进到步骤 A。

[0139] 下面，参考图 17。如图 17 所示，其中处理已经前进到步骤 A 的搜索装置 130 利用内容搜索单元 106 的功能基于在步骤 S138 中产生的最终查询执行搜索 (S139)。接下来，搜索装置 130 基于最终查询、利用搜索信息输出单元 111 的功能输出搜索结果 (S140)。此外，搜索装置 130 利用搜索信息输出单元 111 的功能、显示最终查询所包含的搜索条件以及用于修改搜索条件的界面，并且接收用户操作 (S141)。接下来，搜索装置 130 判断搜索条件是否已被用户修改 (S142)。

[0140] 如果搜索条件已经被修改，则在搜索装置 130 中的处理前进到步骤 S143。另一方面，如果搜索条件还没有被修改，则搜索装置 130 结束一系列处理。如果处理前进到步骤 S143，则搜索装置 130 利用相对值反向计算单元 113 的功能、从修改后的数值范围计算相对值 (S143)。在这种情况下，相对值反向计算单元 113 利用从基于暂定查询的搜索结果获取的分布计算相对值。接下来，搜索装置 130 利用数值化单元 131 的功能判断相对值的修改宽度是否小于给定值 (S144)。如果相对值的修改宽度小于给定值，则搜索装置 130 中的处理前进到 S145。另一方面，如果相对值的修改宽度大于给定值，则搜索装置 130 中的处理前进到 S146。

[0141] 如果处理前进到步骤 S145，则搜索装置 130 利用数值化单元 131 的功能、通过参考语义概念词典数据库 132 产生结合有对象的记录 (S145)。另一方面，如果处理前进到步骤 S146，则搜索装置 130 利用数值化单元 131 的功能产生分开地登记对象的记录 (S146)。处理已经前进到步骤 S147 的搜索装置 130 使用数值化单元 131 的功能、利用在步骤 S145 或步骤 S146 中产生的记录来更新个人相对值数据库 110。

[0142] 已经结束了步骤 S147 的处理的搜索装置 130 中的处理，前进到步骤 B。即，在搜索装置 130 中的处理返回到图 16 所示的步骤 S138，并且再次从步骤 S138 执行处理。搜索装

置 130 反复地执行步骤 S138 到 S147, 直到由用户进行的搜索条件修改完成为止, 并且在搜索条件修改结束以后结束一系列处理。此外, 如果在作为搜索条件输入的自然语句中含有多个主观表达, 则对每个主观表达执行从步骤 S134 开始的处理。

[0143] 上面已经描述了根据本实施例的第一变形例的信息搜索方法。

[0144] <4 :第二变形例>

[0145] 下面, 将描述本实施例的第二变形例。本变形例涉及如果搜索结果的数量太大或太小, 则基于搜索条件的权重自动地调整针对主观表达的搜索条件的方法。

[0146] [4-1 :搜索装置 150 的功能配置]

[0147] 首先, 将参考图 18 描述本变形例的搜索装置 150 的功能配置。图 18 是图示了根据本变形例的搜索装置 150 的功能配置的说明图。与搜索装置 100 的那些具有基本相同的功能的组件被分配了相同的附图标记并且将省略其详细描述。

[0148] 如图 18 所示, 搜索装置 150 包括语音识别单元 101、自然语言分析单元 102、语言分析词典数据库 103、查询生成单元 104 和查询生成词典数据库 105。此外, 搜索装置 150 还包括内容搜索单元 106、内容元数据数据库 107、数值化单元 151、标准相对值数据库 109 和个人相对值数据库 110。搜索装置 150 还包括搜索信息输出单元 111、搜索条件修改单元 112 和相对值反向计算单元 113。

[0149] 与搜索装置 100 的区别在于数值化单元 151。将仅详细描述数值化单元 151 的配置。

[0150] 如上所述, 本变形例的目的是当搜索结果的数目太大或太小时、考虑由用户所保持的概念的权重 (分配给各个主观表达的权重; 见图 7) 来确定要改变的搜索条件和改变量。例如, 如果自然语句“在银座小于等于 2000 日圆的美味的寿司餐厅”被作为搜索条件输入, 则由于包含小于等于 2000 日圆的严格限制, 因此可能不能得到足够数目的搜索结果。

[0151] 同时, 如果自然语句“在涩谷的酒馆”被作为搜索条件输入, 则由于存在较小数目的元素来减小搜索结果的数目, 因此, 很有可能得到用户不能全部选择的很大数目的搜索结果。因此, 如果搜索结果的数目太大或太小, 则重新审查搜索条件。本变形例提供了自动地执行这样的搜索条件的重新审查的机制。

[0152] 如图 7 所示, 主观表达的权重被存储在个人相对值数据库 110 中。数值化单元 151 通过参考个人相对值数据库 110, 将与具有较小权重的主观表达对应的相对值调整得较大, 或者使与这些主观表达对应的搜索条件无效。

[0153] 例如, 考虑严格地考虑地点并且粗略地考虑预算的用户。此外, 假设用户将“地点”的权重设置为 160 并且将“预算”的权重设置为 75。如果不能充分地得到对于某个自然语句的输入的搜索结果, 则用户可以改变“地点”的数值范围或“预算”的数值范围。

[0154] 如上所述, 用户严格地考虑地点。因此, 用户可能不会极大地改变地点的数值范围。另一方面, 由于用户粗略地考虑预算, 因此, 可允许对预算的数值范围的较大的变化。为此, 数值化单元 151 例如基于下面的算法来更新相对值。

[0155] // 前提条件

[0156] “便宜”的权重 :0.75

[0157] “便宜”的相对值 Y :较低的 30 个百分点 ;

[0158] 当前点击数目 a :5

- [0159] 适当的点击数目范围 b :10 到 50(给定值)
- [0160] // 用于参数调整的变量 Z 的计算(适用于所有搜索条件)
- [0161] $Z = f(a, b) = f(5, [10:50])$
- [0162] $= -100*(10-5)/10$
- [0163] $= -50$
- [0164] $Z' = F(X, Z) = F(0.75, -50)$
- [0165] $= -50/0.75$
- [0166] $= -67$
- [0167] // 修改后的“便宜”的相对值 Y 的计算(对于“预算”)
- [0168] 因数= $(100-Z')/100$
- [0169] $= 1.67$
- [0170] $Y' = 30*1.67$
- [0171] = 较低的 50.1 个百分点
- [0172] 已经描述了根据本实施例的第二变形例的搜索装置 150 的功能配置。
- [0173] [4-2 :信息搜索方法(搜索装置 150 的操作)]
- [0174] 下面,将参考图 19 和图 20 描述根据本实施例的第二变形例的信息搜索方法。图 19 和图 20 是示出了在根据本实施例的第二变形例的信息搜索方法的处理的流程的流程图。
- [0175] 首先,参考图 19。如图 19 所示,首先,输入用于识别用户的标识符(用户 ID)(S151)。下面,由用户通过语音输入搜索条件(S152)。如果搜索条件由用户通过语音输入,则搜索装置 150 利用语音识别单元 101 的功能将所输入的语音转换为文本(S153)。搜索装置 150 然后利用自然语言分析单元 102 的功能分析被转换为文本的搜索条件(自然语句)(S154)。搜索装置 150 然后利用查询生成单元 104 的功能、组合由步骤 S154 的分析得到的自然语句的元素以产生暂定查询(S155)。在这种情况下,查询生成单元 104 组合除了将被转换为数值范围的主观表达以外的自然语句的元素,以产生暂定查询。
- [0176] 接下来,搜索装置 150 利用内容搜索单元 106 的功能,基于暂定查询执行搜索,并且基于暂定查询获取搜索结果。搜索装置 150 然后利用数值化单元 151 的功能、确认标识符是否已经在个人相对值数据库 110 中注册(S156)。如果标识符已经被注册,则在搜索装置 150 中的处理前进到步骤 S157 中。另一方面,如果标识符还没有被注册,则在搜索装置 150 中的处理前进到步骤 S159。
- [0177] 如果处理前进到步骤 S157,则搜索装置 150 利用数值化单元 151 的功能、确认与将被转换为数值范围的主观表达对应的记录是否已经登记在个人相对值数据库 110 中(S157)。如果该记录已经登记在个人相对值数据库 110 中,则搜索装置 150 中的处理前进到步骤 S158。另一方面,如果该记录还没有登记在个人相对值数据库 110 中,则搜索装置 150 中的处理前进到步骤 S159。
- [0178] 如果处理前进到步骤 S158,则搜索装置 150 利用数值化单元 151 的功能从个人相对值数据库 110 获取相对值(S158)。同时,如果处理前进到步骤 S159,则搜索装置 150 利用数值化单元 151 的功能从标准相对值数据库 109 获取与主观表达对应的相对值(S159)。已经获取了相对值的搜索装置 150 利用数值化单元 151 的功能、从基于暂定查询的搜索结果获取数值范围的分布,并且基于该分布确定对应于相对值的数值范围。搜索装置 150 利用

查询生成单元 104 的功能、产生将所确定的数值范围添加到搜索条件的最终查询 (S160)，并且处理前进到步骤 A。

[0179] 下面，参考图 20 所示，其中处理已经前进到步骤 A 的搜索装置 150 利用内容搜索单元 106 的功能、基于在步骤 S160 中产生的最终查询执行搜索 (S161)。搜索装置 150 然后利用数值化单元 151 的功能判断搜索结果的数目是否适合 (S162)。如果搜索结果的数目适合，则搜索装置 150 中的处理前进到步骤 S163。另一方面，如果搜索结果的数目不适合，则搜索装置 150 中的处理前进到步骤 S164。

[0180] 如果处理前进到步骤 S163，则搜索装置 150 利用数值化单元 151 的功能给搜索结果评分 (步骤 S163)，并且处理前进到步骤 S165。同时，如果处理前进到步骤 S164，则搜索装置 150 利用数值化单元 151 的功能、基于权重来调整相对值并且更新个人相对值数据库 110(S164)。已经结束了步骤 S164 的处理的搜索装置 150 中的处理前进到步骤 B。即，搜索装置 150 中的处理返回到图 19 所示的步骤 S160，并且搜索装置 150 再次从步骤 S160 执行处理。

[0181] 其中处理已经前进到步骤 S165 的搜索装置 150，利用搜索信息输出单元 111 的功能输出基于基于最终查询的搜索结果 (S165)。此外，搜索装置 150 利用搜索信息输出单元 111 的功能、显示最终查询所包含的搜索条件以及用于修改搜索条件的界面，并且接收用户操作 (S166)。搜索装置 150 然后判断搜索条件是否已被用户修改 (S167)。

[0182] 如果搜索条件已经被修改，则在搜索装置 150 中的处理前进到步骤 S168。另一方面，如果搜索条件还没有被修改，则搜索装置 150 结束一系列处理。如果处理前进到步骤 S168，则搜索装置 150 利用相对值反向计算单元 113 的功能、从修改后的数值范围计算相对值 (S168)。在这种情况下，相对值反向计算单元 113 利用从基于暂定查询的搜索结果获取的分布计算相对值。搜索装置 150 然后使用相对值反向计算单元 113 的功能、利用在步骤 S168 中计算的相对值来更新个人相对值数据库 110(S169)。

[0183] 已经结束了在步骤 S169 中的处理的搜索装置 150 的处理，前进到步骤 B。即，搜索装置 150 中的处理返回到图 19 所示的步骤 S160，并且再次从步骤 S160 执行处理。搜索装置 150 反复地执行步骤 S160 到 S169，直到由用户进行的搜索条件修改完成为止，并且在搜索条件修改结束以后结束一系列处理。此外，如果在作为搜索条件输入的自然语句中含有多个主观元素，则对每个主观表达执行步骤 S155 以后的处理。

[0184] 上面已经描述了根据本实施例的第二变形例的信息搜索方法。

[0185] <5 :第三变形例 >

[0186] 在某些情况下，在由用户输入的搜索结果中的主观表达可能是不明确的，并且搜索装置可执行处理以消除预期的搜索条件的歧义。例如，用户的主观表达“便宜”可以意味着“不贵”或者“低质量”。在某些实施例中，搜索装置可以使用词库或者词或关联的其他数据库来确定用户的主观表达是不明确的。在这种情况下，搜索装置可以提示用户消除搜索条件的歧义。例如，在某些情况下搜索装置可以呈现给用户多个选项（例如，“不贵”和“低质量”），并且允许用户来选择最好地对应主观表达的预期含义的一个。

[0187] <6 :硬件配置 >

[0188] 上述搜索装置 100、130 和 150 的各个结构元件的功能可通过使用例如图 21 所示的信息处理装置的硬件配置来执行。即，各个结构元件的功能可通过利用计算机程序来控

制图 21 所示的硬件来实现。此外，该硬件的模式是任意的，并且可以是个人计算机、诸如移动电话、PHS 或 PDA、等的移动信息终端、游戏机或者其他类型的信息。此外，PHS 是个人手持电话的缩写。此外，PDA 是个人数字助理的缩写。

[0189] 如图 21 所示，该硬件主要包括诸如 CPU 902 等的至少一个处理器、ROM 904、RAM 906、主机总线 908 和桥接器 910。此外，该硬件包括外部总线 912、接口 914、输入单元 916、输出单元 918、存储单元 920、驱动器 922、连接端口 924 和通信单元 926。而且，CPU 是中央处理器的缩写。此外，ROM 是只读存储器的缩写。此外，RAM 是随机存储器的缩写。

[0190] CPU 902 例如用作运算处理单元或控制单元，并且基于记录在 ROM 904、RAM 906、存储单元 920 或可移除记录介质 928 中的各种程序来控制各个结构元件的全部操作或一部分操作。ROM 904 是用于存储例如被下载到 CPU 902 上的程序或在运算操作中使用的数据等的装置。RAM 906 临时或永久地存储例如下载到 CPU 902 上的程序或者在程序的执行中任意改变的各种参数。

[0191] 这些结构元件通过例如能够执行高速数据传输的主机总线 908 相互连接。就其本身而言，例如，主机总线 908 通过桥接器 910 连接到数据传输速度相对较慢的外部总线 912。此外，输入单元 916 例如是鼠标、键盘、触摸屏、按钮、开关或控制杆。此外，输入单元 916 可以是利用红外线或其他无线电波来传输控制信号的远程控制器。

[0192] 输出单元 918 例如是能够可视地或听觉上通知用户所获取的信息的、诸如 CRT、LCD、PDP 或 ELD 等的显示设备、诸如扬声器或耳机等声音输出设备、打印机、移动电话或传真机。而且，CRT 是阴极射线管的缩写。LCD 是液晶显示器的缩写。PDP 是等离子显示板的缩写。此外，ELD 是电子发光显示器的缩写。

[0193] 存储单元 920 是用于存储各种数据的设备。存储单元 920 例如是诸如硬盘驱动器 (HDD) 等的磁存储设备、半导体存储设备、光学存储设备或磁光存储设备。HDD 是硬盘驱动器的缩写。

[0194] 驱动器 922 是读取记录在诸如磁盘、光盘、磁光盘或半导体记忆体等的可移除记录介质 928 上的信息，或者将信息写入可移除记录介质 928 的设备。可移除记录介质 928 例如是 DVD 介质、蓝光介质、HD-DVD 介质、各种类型的半导体存储介质等。当然，可移除记录介质 928 可以是例如其上安装了非接触 IC 的电子设备或 IC 卡。IC 是集成电路的缩写。

[0195] 连接端口 924 是诸如 USB 端口、IEEE1394 端口、SCSI、RS-232C 端口或诸如光纤音频端子等的用于连接外部连接设备 930 的端口之类的端口。外部连接设备 930 例如是打印机、移动音乐播放器、数码相机、数码摄像机或 IC 录音机。而且，USB 是通用串行总线的缩写。此外，SCSI 是小型计算机系统接口的缩写。

[0196] 通信单元 926 是连接到网络 932 的通信设备，并且例如是用于有线或无线 LAN 的通信卡、蓝牙（注册商标）、或 WUSB、光通信路由器、ADSL 路由器，或用于不同类型的通信的调制解调器。连接到通信单元 926 的网络 932 由有线连接或无线连接的网络构成，并且例如是互联网、家用 LAN、红外通信、可见光通信、广播或卫星通信。而且，LAN 是局域网的缩写。此外，WUSB 是无线 USB 的缩写。此外，ADSL 是非对称数字用户线路的缩写。

[0197] 某些实施例可包括编码有一个或多个程序的计算机可读存储介质（或多个计算机可读介质）（例如，计算机存储器、一个或多个软盘、光碟（CD）、光盘、数字视频光盘（DVD）、磁带、闪存、在现场可编程门阵列电路或其他半导体器件中的电路配置，或其他有形

的计算机存储介质),其中当所述一个或多个程序在一个或多个计算机或其他处理器上执行时,执行实现上面所讨论的各种实施例的方法。从前面的示例中可以清楚,计算机可读存储介质可以保留信息足够的时间来以非暂时性的形式提供计算机可读指令。这样的计算机可读介质或媒介是可传输的,使得存储在其上的一个或多个程序能够被下载到一个或多个不同的计算机或其他处理器、以实现如上所讨论的本发明的各个方面。如这里所使用的,术语“计算机可读存储介质”仅包括能够被认为是产品(即,制品)或机器的计算机可读介质。替换地或额外地,本发明可实施为诸如传播信号等的除了计算机可读存储介质之以外的计算机可读介质。

[0198] <7 :总结>

[0199] 最后,将简要地总结根据本公开的实施例的技术内容。这里所述的技术内容可应用于例如诸如PC、便携电话、便携式游戏设备、个人数字助理、信息设备和汽车导航系统等的若跟信息处理装置。

[0200] 信息处理装置的功能配置可表示如下。信息处理装置包括下面的暂定搜索单元、分布获取单元、数值范围选择单元和当前搜索单元。如果输入了含有将被转换为数值范围的主观表达的自然语句,则暂定搜索单元使用构成自然语句的元素当中的除主观表达以外的元素作为搜索条件、来执行搜索。例如,如果输入了自然语句“给我看在银座便宜并且美味的寿司餐厅”,将被转换为数值范围的主观表达是指示价格范围的元素“便宜”。

[0201] 仅基于主观表达很难唯一地确定用户试图通过这样的主观表达来表达的数字范围。即,很难判断由主观表达“便宜”所指示的价格范围是0至3000日圆还是0至6000日圆。暂定查询利用自然语句的除该主观表达以外的元素暂定地搜索信息。利用该暂定搜索,可收集指示被用户认为是先决条件的条件的信息。例如,可收集关于“在银座的寿司餐厅”的信息。利用该信息,例如,可预测用户的市场展望。

[0202] 分布获取单元基于相对于暂定搜索单元的搜索结果、获取对应于主观表达的数值范围。基于上述示例,该分布获取单元从关于“在银座的寿司餐厅”收集的信息、获取价格范围的分布。如果能够获取价格范围的分布,则可识别“在银座的寿司餐厅”当中的“贵的寿司餐厅”和“便宜的寿司餐厅”。例如,与从低价格顺次X%对应的“寿司餐厅”可被判断为“便宜的寿司餐厅”。然而,X是由用户预先确定的值。此外,如果可识别出“便宜的寿司餐厅”,则可逆向地得到“不贵的寿司餐厅”的价格范围。

[0203] 数值范围选择单元基于预先与主观表达相关联的相对值、从由分布获取单元获取的数值的分布中选择数值范围。该相对值是与X对应的数值。如上所述,由于数值的分布可从由暂定搜索得到的信息获得,因此由利用数值的分布和相对值来选择由主观表达所指示的数值范围。例如,该数值范围选择单元可选择在“在银座的寿司餐厅”当中的对应于从低价格范围向上X%的“寿司餐厅”所对顶的价格范围(例如,0至5000日圆)。这里所选择的价格范围是反映“在银座的寿司餐厅”的市场价格的一个,并且满足用户所认为的“便宜”(X%)标准。

[0204] 本搜索单元利用构成自然语句的元素当中的除主观表达以外的元素、以及由数值范围选择单元所选择的数值范围作为搜索条件来执行搜索。由于如上所述主观表达被自动地转换为数值范围,因此可正确地得到适合于数值范围的搜索结果。

[0205] 此外,由于当选择数值范围时使用了相对值,例如,因此,适合于“在银座的寿司餐

厅”的市场价格的价格范围被用作搜索条件并且得到用户所期望的适当的搜索结果。如果搜索条件修改为“在涩谷的便宜的寿司餐厅”，则适合于“在涩谷的便宜的寿司餐厅”的市场价格的价格范围被用作搜索条件。因此，如果使用信息处理装置，则可自动地产生已经反映了被用户假设为先决条件的情况的搜索条件，并且基于该搜索条件得到适当的搜索结果。

[0206] 此外，在主观表达“便宜”与数值范围“0 至 6000 日圆”相关联并且以数值范围代替包含在自然语句中的主观表达来执行搜索的一般的信息搜索系统的情况下，偶然地得到对于“在银座便宜的寿司餐厅”的适当的搜索结果，但是不能得到对于“在涩谷便宜的寿司餐厅”的适当的搜索结果。信息处理装置能够适当地解决这样的问题。此外，信息处理装置能够自适应地应付作为主观表达的对象的任何元素之间的区别，以及地点差别。例如，已经反映了“在银座的寿司餐厅”的市场价格的价格范围是用于输入“在银座便宜的寿司餐厅”的搜索条件，而已经反映了“在银座的牛排馆”的市场价格的价格范围是用于输入“在银座便宜的牛排馆”的搜索条件。

[0207] 因此，如果应用根据本实施例的技术，则在由自然语句的除主观表达以外的元素确定的情形中、能够以适当的数值范围代替主观表达。结果，可以适当地反映随着在搜索条件中的情况变化的用户的主观性，并且可以提供用户期望的搜索结果。

[0208] 此外，由于诸如“银座”和“寿司餐厅”等的作为主观表达的对象的各个元素的组合以及各个主观表达不需要单独地与数值范围相关联，并且主观表达与相对值相关联，因此信息处理装置所持有的信息量较小。此外，由于作为搜索目标的信息量较小，因此提高了搜索效率。这样的效果不是通过应用暂定搜索单元、分布获取单元、数值范围选择单元和目前搜索单元的功能得到的。

[0209] (备注)

[0210] 查询生成单元 104 和内容搜索单元 106 是暂定搜索单元和当前搜索单元的示例。数值化单元 108、131 和 151 是分布获取单元和数值范围选择单元的示例。搜索信息输出单元 111 是输出单元的示例。个人相对值数据库 110 是相对值保持单元和元素特有的相对值保持单元的示例。语音识别单元 101 是语音 / 文本转换单元的示例。查询生成单元 104 是主观表达抽取单元的示例。数值化单元 131 是群集单元的示例。数值化单元 151 是权重设置单元和相对值调整单元的示例。

[0211] 本领域的技术人员应该明白，只要在所附权利要求或其等同物的范围内，根据设计需求和其他因素可发生各种修改、组合、子组合和变形例。

[0212] 例如，本技术可采用如下配置。

[0213] (1) 一种搜索装置包括：

[0214] 暂定搜索单元，其用于如果含有主观表达的自然语句被输入，则利用构成自然语句的元素当中除了将被转换为数值范围的主观表达以外的元素来执行搜索；

[0215] 分布获取单元，其用于相对于暂定搜索单元的搜索结果、获取与主观表达对应的数值的分布；

[0216] 数值范围选择单元，其用于基于预先与主观表达相关联的相对值、从由分布获取单元获取的数值的分布中选择数值范围；

[0217] 当前搜索单元，其用于利用构成自然语句的元素当中除了主观表达以外的元素以及由数值范围选择单元选择的数值范围作为搜索条件来执行搜索。

- [0218] (2) 根据 (1) 所述的搜索装置,还包括 :
- [0219] 输出单元,其用于至少输出由数值范围选择单元选择的数值范围;以及
- [0220] 输入单元,其用于改变包含在由当前搜索单元进行的搜索的条件中的数值范围。
- [0221] (3) 根据 (2) 所述的搜索装置,还包括 :
- [0222] 相对值反向计算单元,其用于如果数值范围通过输入单元被改变,则基于由分布获取单元获取的数值的分布计算与改变后的数值范围对应的相对值;以及
- [0223] 相对值保持单元,其用于保持将被与主观表达相关联的、由相对值反向计算单元计算的相对值。
- [0224] (4) 根据 (3) 所述的搜索装置,其中数值范围选择单元确认与主观表达对应的相对值是否已经被相对值保持单元保持,并且如果与主观表达对应的相对值已经被保持,则基于相对值保持单元所保持的相对值来选择数值范围。
- [0225] (5) 根据 (1) 至 (4) 中的任一项所述的搜索装置,还包括 :
- [0226] 语音输入单元,其用于通过语音输入自然语句;以及
- [0227] 语音 / 文本转换单元,其用于将输入到语音输入单元的语音转换为文本。
- [0228] (6) 根据 (1) 至 (5) 中的任一项所述的搜索装置,还包括 :
- [0229] 自然语言分析单元,其用于通过自然语音分析来抽取构成自然语句的各个元素;以及
- [0230] 主观表达抽取单元,其用于将由自然语言分析单元抽取的各个元素与给定的主观表达相比较,并且抽取在自然语句中所包含的主观表达,
- [0231] 其中暂定搜索单元和当前搜索单元利用由自然语言分析单元抽取的元素和由主观表达抽取单元抽取的主观表达来设置搜索的条件。
- [0232] (7) 根据 (2) 至 (6) 中的任一项所述的搜索装置,还包括 :
- [0233] 元素特定相对值保持单元,用于如果数字范围通过输入单元被改变,则从构成自然语句的元素当中抽取作为主观表达的对象的元素,并且保持将被与相对值反向计算单元所计算的相对值相关联的所抽取的元素;以及
- [0234] 群集单元,其用于群集相对于由元素特定相对值保持单元保持的元素具有相似的概念的元素,并且将各个群集与相对值相关联,
- [0235] 其中数值范围选择单元从构成自然语句的元素中抽取作为主观表达的对象的元素,并且基于与所抽取的元素具有相似概念的群集所对应的相对值来选择数值范围。
- [0236] (8) 根据 (1) 至 (7) 中的任一项所述的搜索装置,还包括 :
- [0237] 权重设置单元,其用于根据数值范围的改变频率来设置与数值范围对应的主观表达的权重;以及
- [0238] 相对值调整单元,其用于如果由当前搜索单元得到的搜索结果的数目小于第一给定数目或大于第二给定数目,则基于由权重设置单元设置的权重、来调整包含在由当前搜索单元进行的搜索的条件中的数值范围所对应的相对值。
- [0239] (9) 一种在搜索装置中的搜索方法,该方法包括以下步骤 :
- [0240] 如果输入了含有主观表达的自然语句,则利用构成自然语句的元素当中除了将被转换为数值范围的主观表达以外的元素来执行搜索;
- [0241] 相对于暂定搜索单元的搜索结果、获取与主观表达对应的数值的分布;

[0242] 基于预先与主观表达相关联的相对值、从在分布获取步骤中获取的数值的分布中选择数值范围；以及

[0243] 利用构成自然语句的元素当中除了主观表达以外的元素以及在数值范围选择步骤中选择的数值范围作为搜索条件来执行搜索。

[0244] (10) 一种用于使计算机实现以下功能的程序：

[0245] 暂定搜索功能，该功能在含有主观表达的自然语句被输入的情况下，利用构成自然语句的元素当中除了将被转换为数值范围的主观表达以外的元素来执行搜索；

[0246] 分布获取功能，该功能相对于通过暂定搜索功能的搜索的结果、获取与主观表达对应的数值的分布；

[0247] 数值范围选择功能，该功能基于预先与主观表达相关联的相对值、从通过分布获取功能获取的数值的分布中选择数值范围；

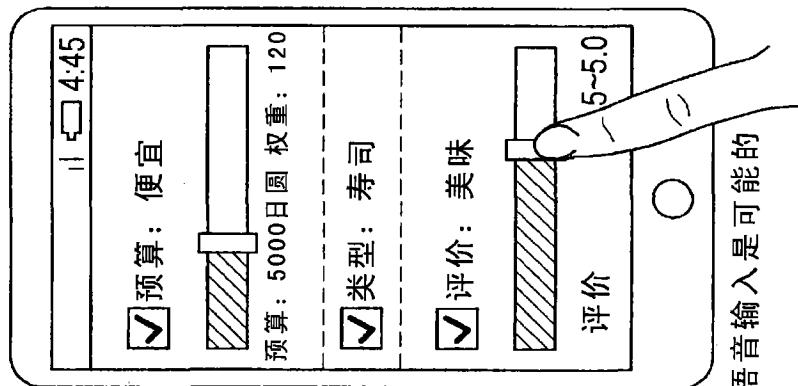
[0248] 当前搜索功能，该功能利用构成自然语句的元素当中除了主观表达以外的元素以及由数值范围选择功能选择的数值范围作为搜索条件来执行搜索。

[0249] 在权利要求中、为修改要求保护的元件的诸如“第一”、“第二”、“第三”等序词的使用，本身并不意味着任何优先权、优先级、或一个要求保护的元件相对于另一个的顺序或者执行方法的动作的时间顺序。这些术语仅用作将具有某个名称的一个要求保护的元件与具有相同名称（但除了序词的使用以外）的另一个元件区分开，以区别要求保护的元件。

[0250] 此外，这里使用的用语和术语是为了说明的目的并且不应该视为限制。这里的“包括”、“包含”或“具有”、“含有”、“涉及”及其变形例等使用意在包括其后所列的项目及其等同物以及其他项目的。

[0251] 本公开包含涉及在于 2011 年 1 月 11 日在日本专利局提交的日本优先专利申请 JP 2011-003005 中公开的主题，该申请的全部内容通过引用结合于此。

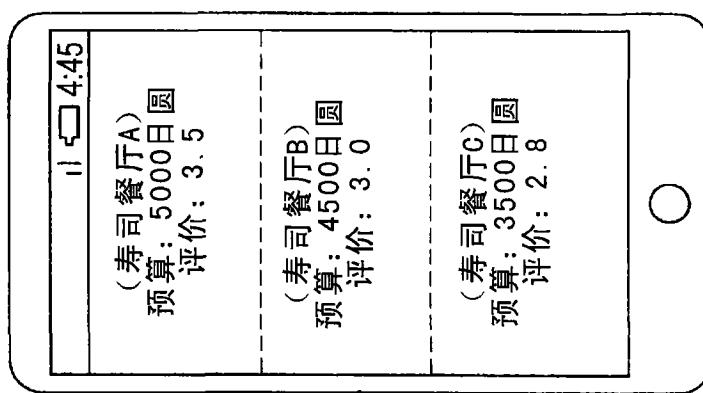
(搜索条件的确认/修改)



确认/修改

重新搜索

(搜索结果的显示)



语音或文本

自然语句输入
给我看在银座的
便宜并且美味的
寿司餐厅

图 1

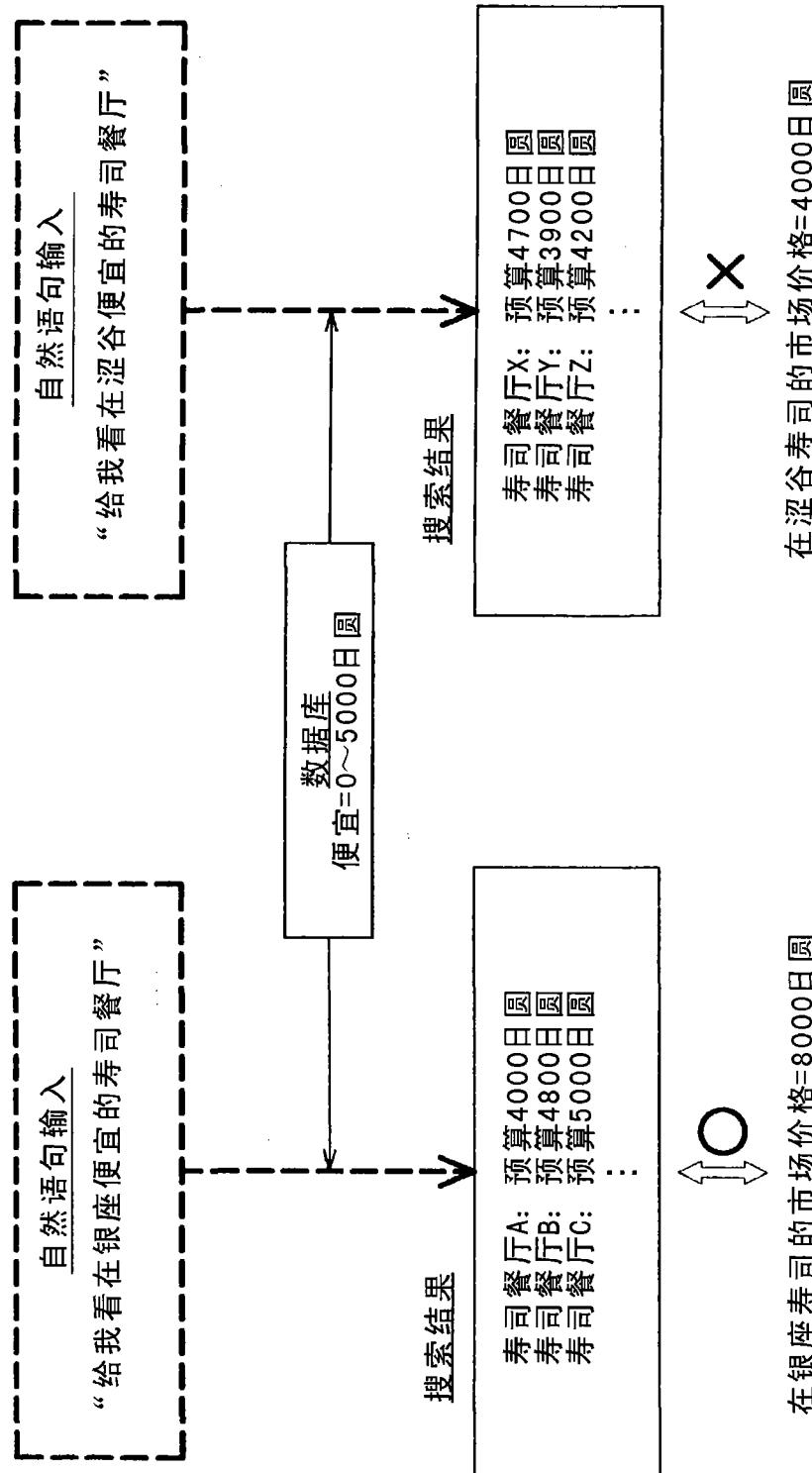
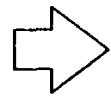


图 2

地点	主观表达	目标元素	数值范围
银座	便宜	寿司餐厅	0~7000 日圆
涩谷	便宜	寿司餐厅	0~3000 日圆
新宿	便宜	寿司餐厅	0~3500 日圆
品川	便宜	寿司餐厅	0~4000 日圆
练马	便宜	寿司餐厅	0~2800 日圆
⋮	⋮	⋮	⋮



数据库大小很大！
→ 降低搜索效率并且增加成本
→ 很难个人化

图 3

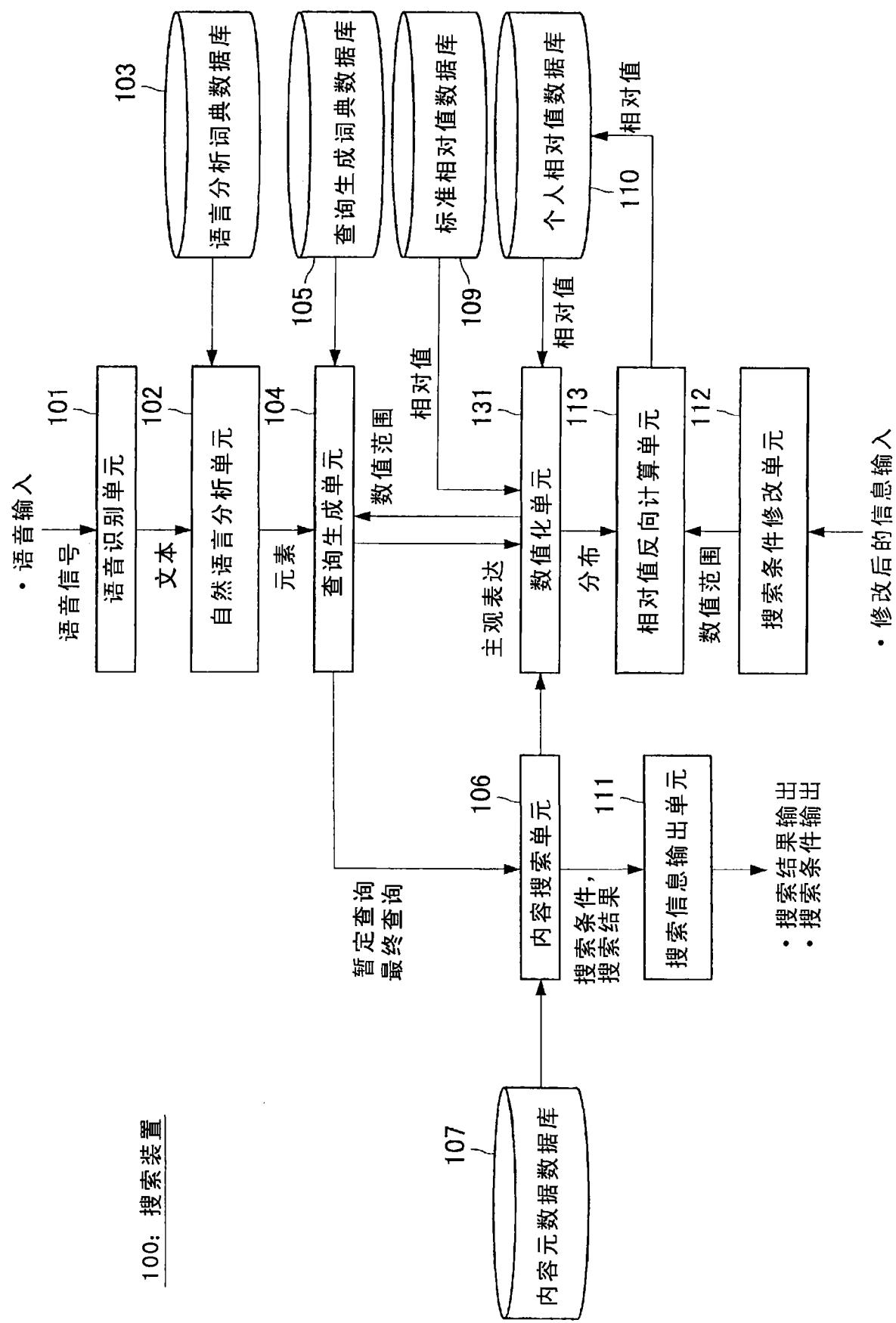


图 4

用语表示	词类/属性	类别	主观表达
非常便宜		价格	非常便宜
便宜		价格	便宜
:		:	:
高级		价格	昂贵
:		:	:
*	专有名称：地名	地点	*

图 5

类别	主观表达	相对值	暂定查询的数值范围
价格	非常便宜	[0%, 10%]	-
	便宜	[0%, 30%]	-
	昂贵	[70%, 100%]	-
评价	评价高	[0%, 35%]	3. 0以上
	厨艺好	[0%, 35%]	3. 0以上
	服务好	[0%, 35%]	3. 0以上

图 6

用户ID	类别	主观表达	相对值	权重值
A	价格	便宜	[0%, 23%]	160
A	价格	昂贵	[51%, 100%]	160
A	评价	评价高	[0%, 40%]	127
B	价格	非常便宜	[0%, 5%]	175
B	评价	评价高	[0%, 25%]	80
B	评价	厨艺好	[0%, 36%]	80
B	评价	服务好	[0%, 41%]	80
C	评价	服务好	[0%, 35%]	110

图 7

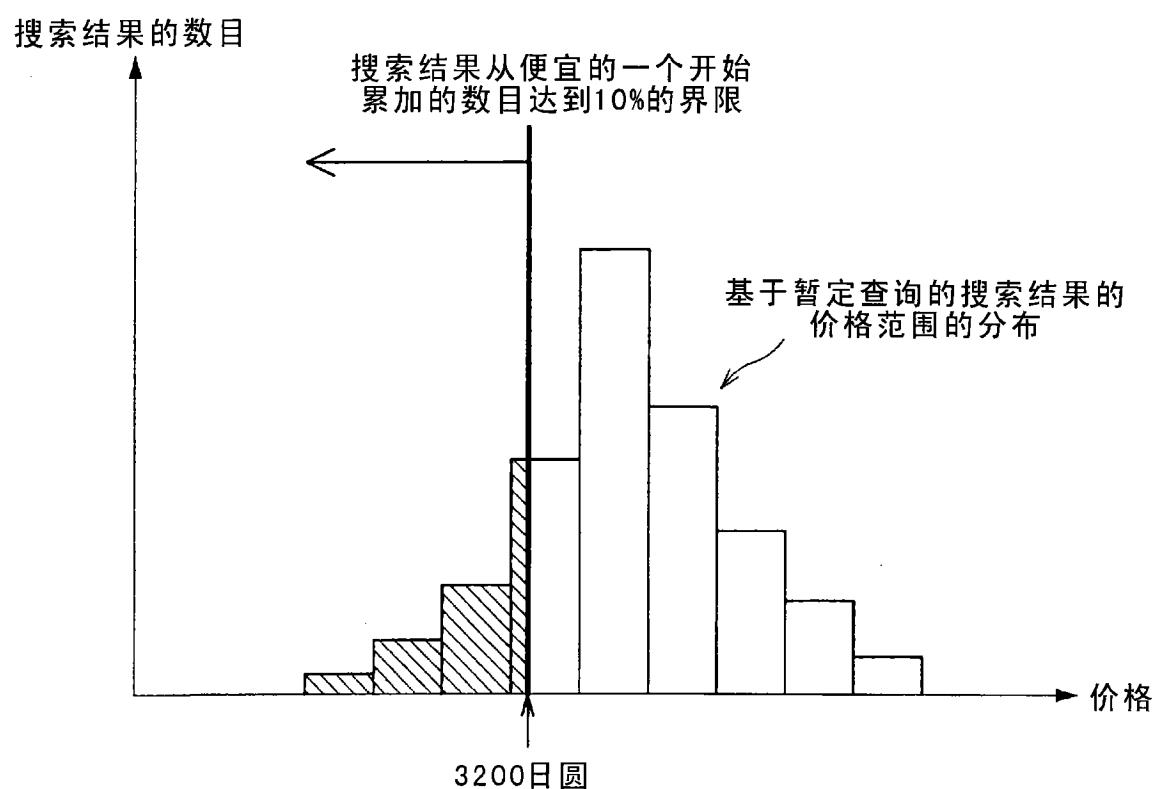


图 8

商店名称	地址	预算	评价
X寿司，银座总店	X-XX-1, 东京都中央区银座	3000	3.3
X寿司，银座2丁目店	2-XX-X, 东京都中央区银座	3800	3.0
Y寿司	1-11-X, 东京都中央区银座	4000	3.0
X寿司	X-1X-7, 东京都中央区银座	4500	3.4
江户前 玉寿司	3-XX-9, 东京都中央区银座	4500	3.1
柳寿司 总店	5-XX-2, 东京都中央区银座	5500	3.4
寿司 权太郎	X-11-1, 东京都中央区银座	6200	3.5
寿司 染次郎	4-1-X, 东京都中央区银座	6800	3.5
寿司 九兵卫	X-22-3, 东京都中央区银座	7800	3.8
:	:	:	:

图 9

店名	地址	最近的车站	(纬度, 经度)	类别ID	类别名称
京町料理 涩谷店	X-XX, 东京都涩谷区道玄坂	涩谷	(35.66, 139.70)	G001	日本料理
居酒屋 虎之助	XX-XXX, 东京都新宿区新宿	新宿	(35.69, 139.70)	G004	酒馆
宵之口 池袋店	X-XX-X, 东京都丰岛区池袋	池袋	(35.73, 139.71)	G004	酒馆
烤肉 炭	X-XX-X, 东京都港区港南	品川	(35.63, 139.74)	G008	烤肉
:	:	:	:	:	:

烹饪ID	烹饪类别	预算ID	预算额	评价	评价人数
R001	日本料理	B003	3001~4000 日圆	3.1	21
R004	串烧食物	B004	2001~3000 日圆	2.8	11
R006	砂锅菜	B002	4001~5000 日圆	3.8	8
R021	烤肉	B002	4001~5000 日圆	3.3	32
:	:	:	:	:	:

图 10

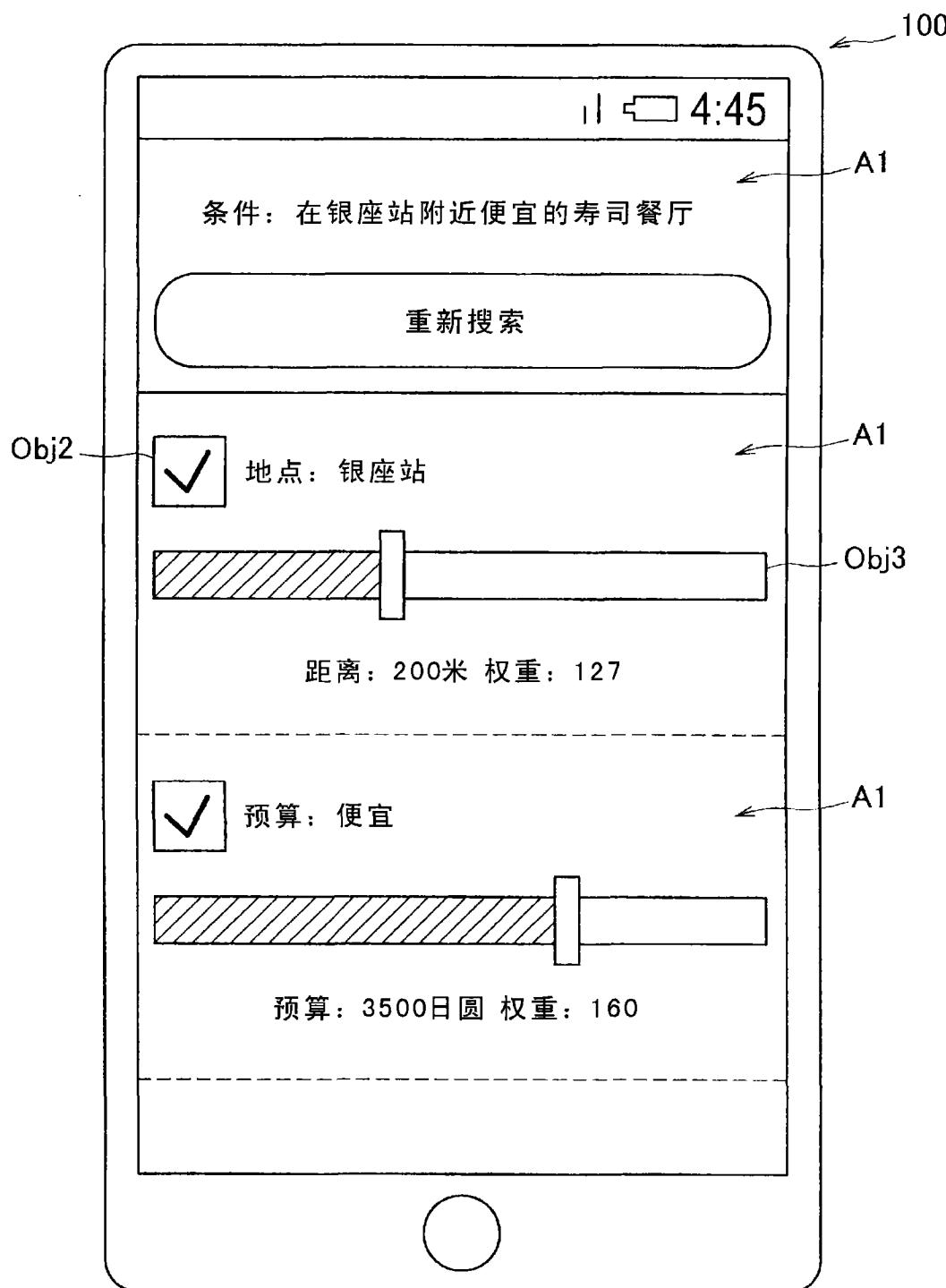


图 11

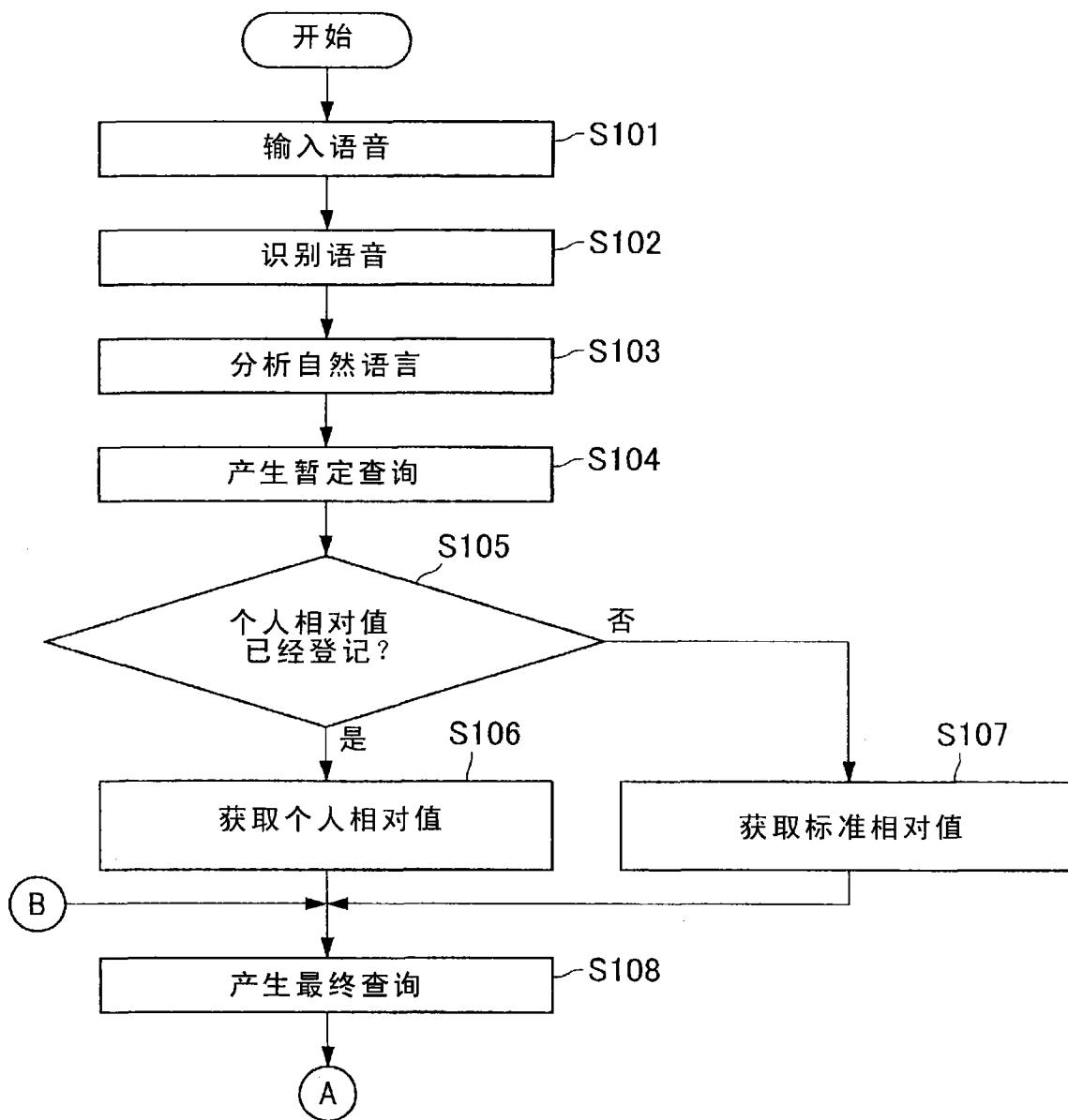


图 12

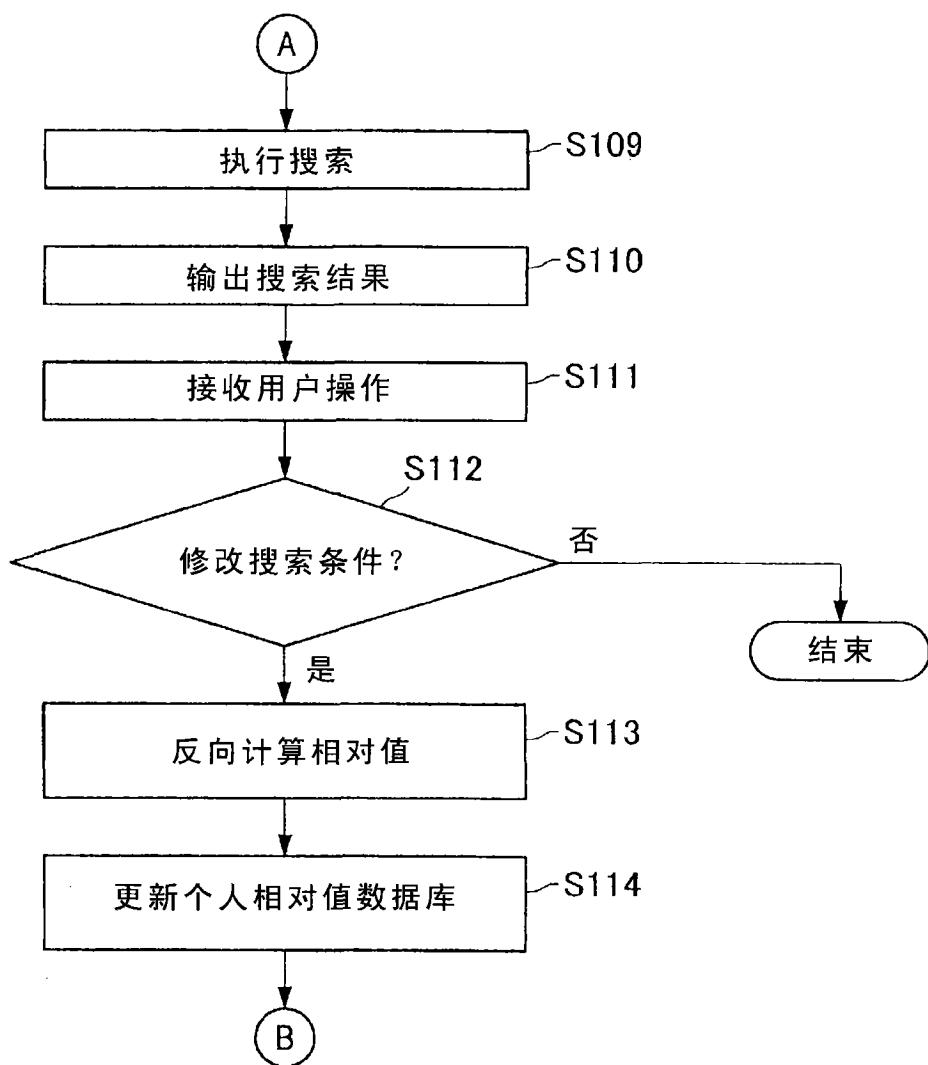


图 13

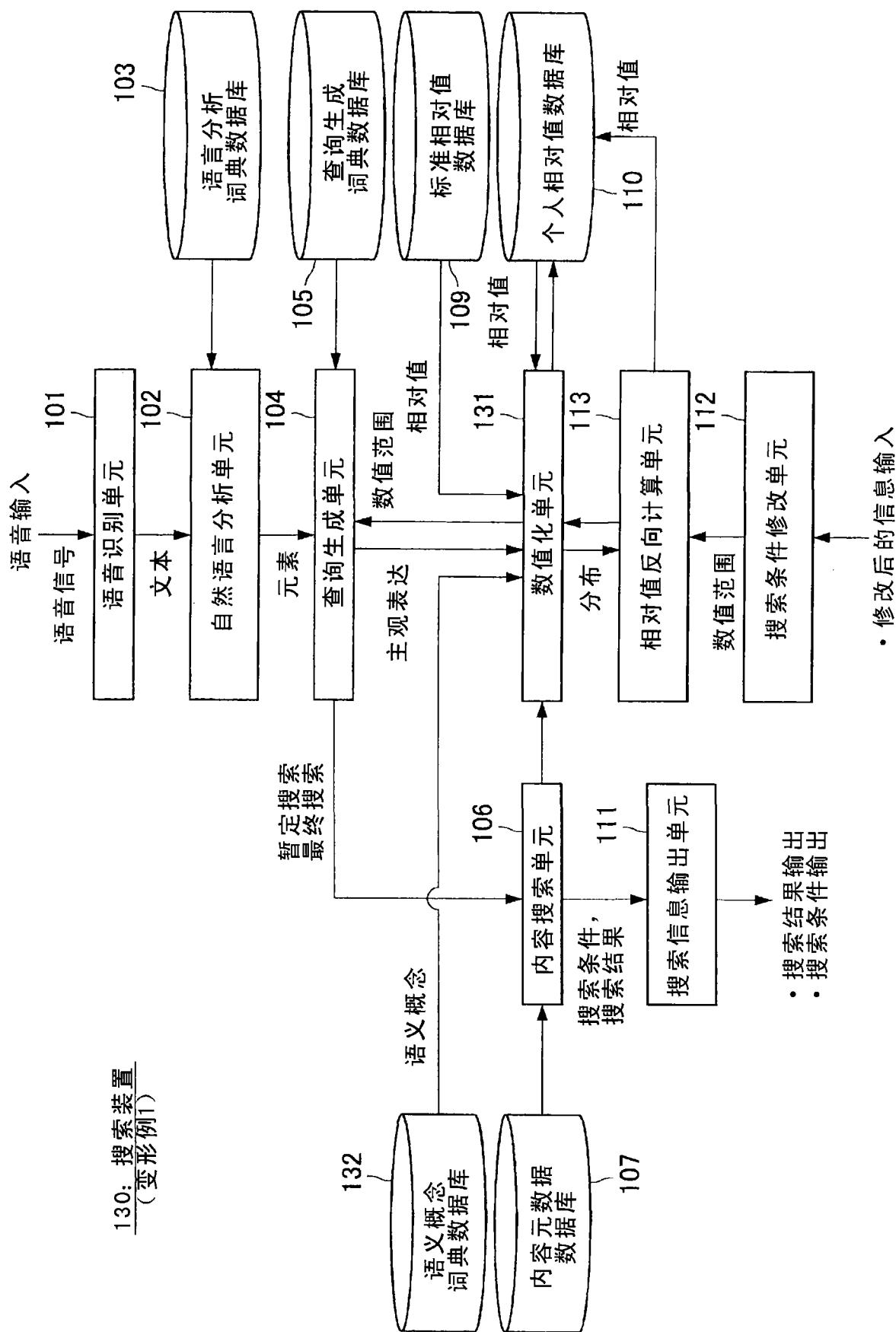


图 14

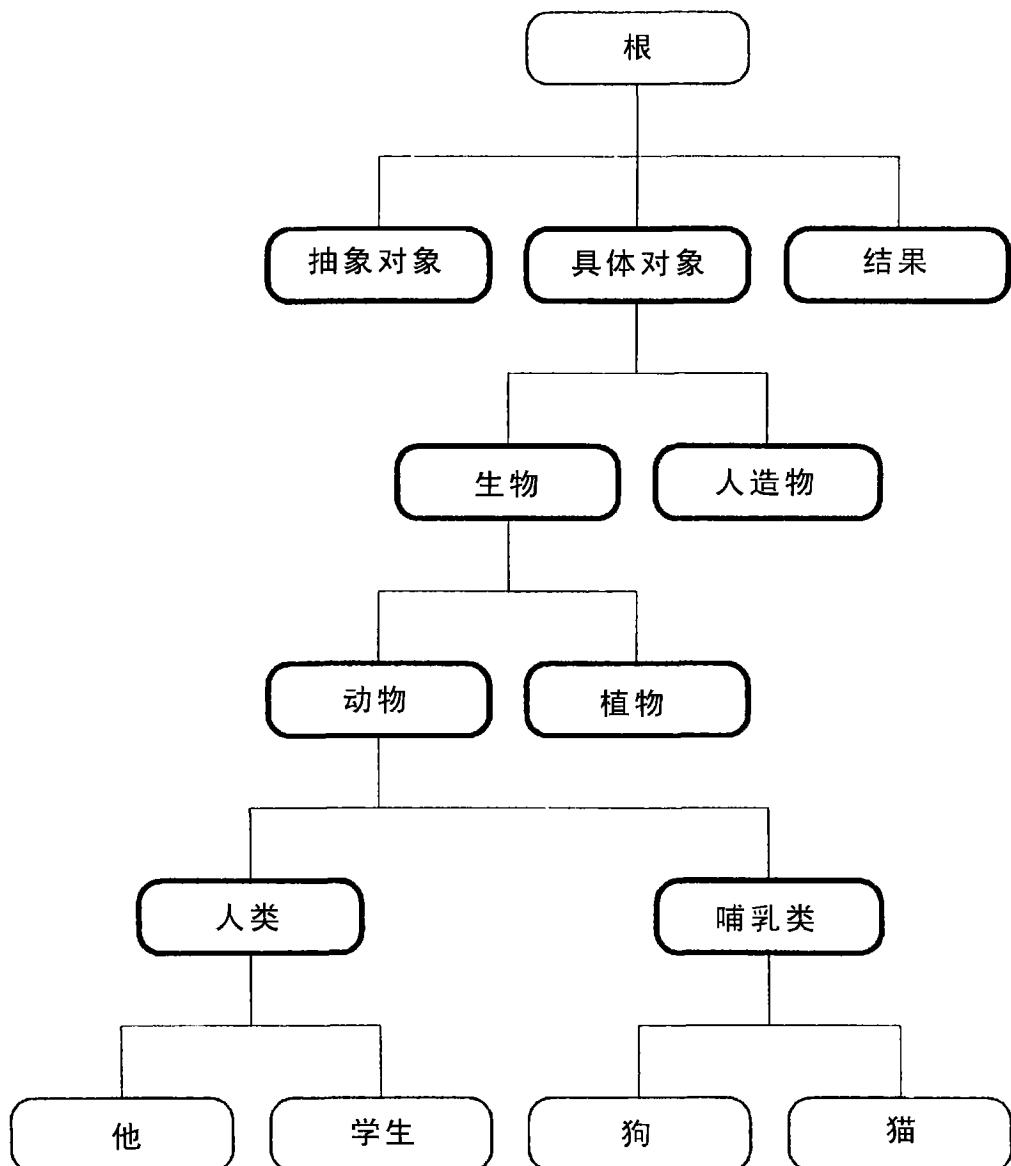


图 15

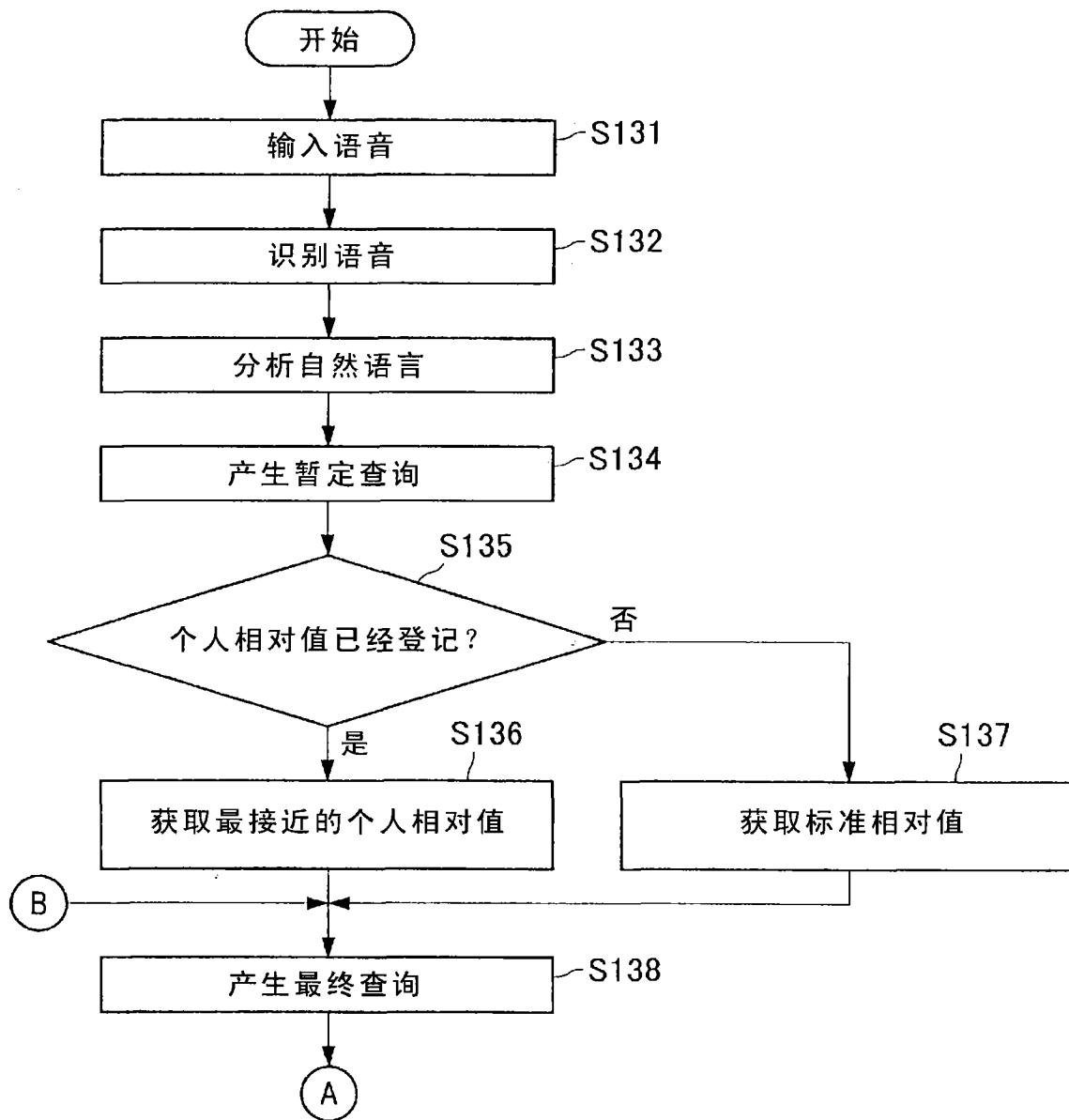


图 16

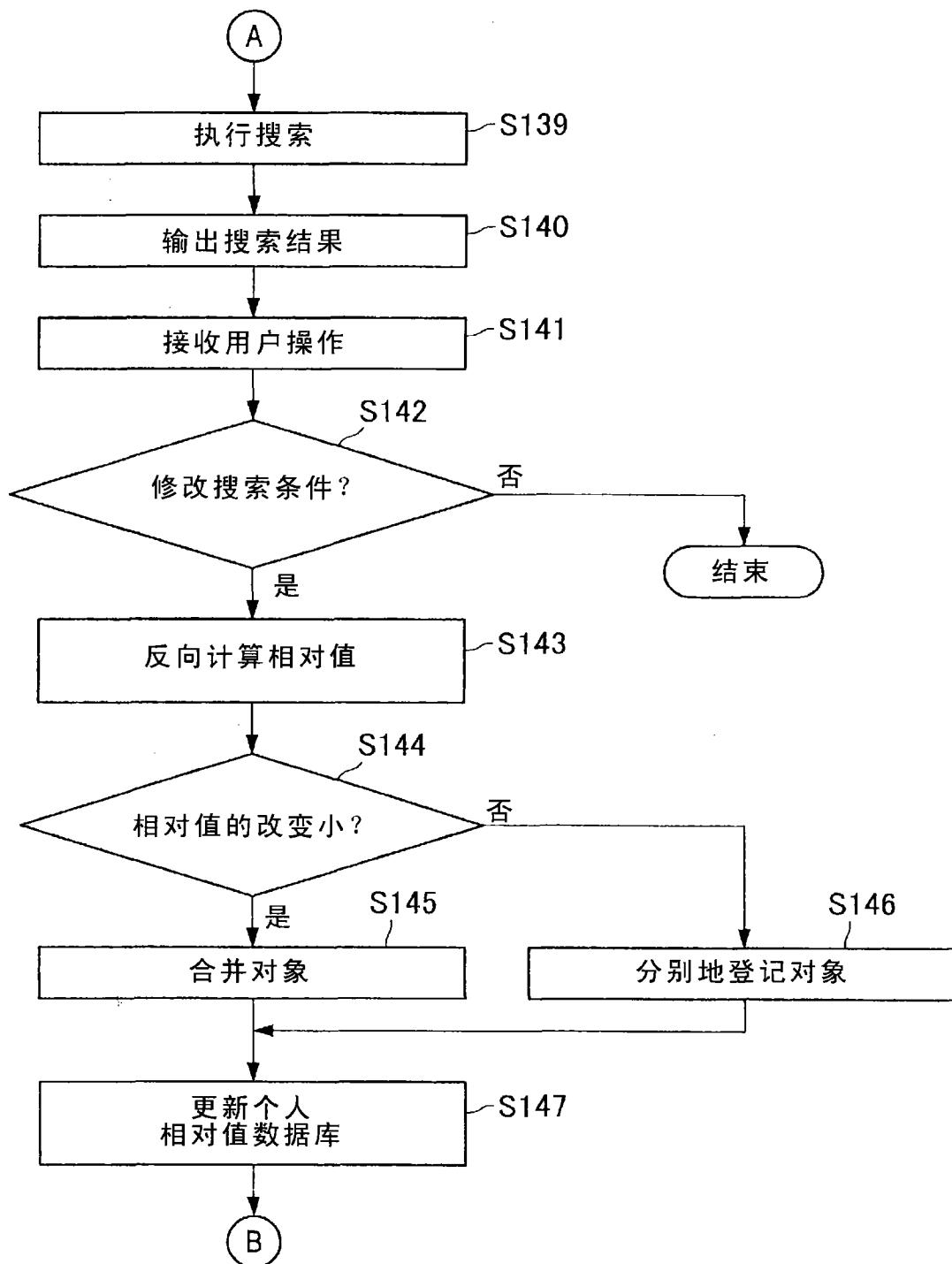


图 17

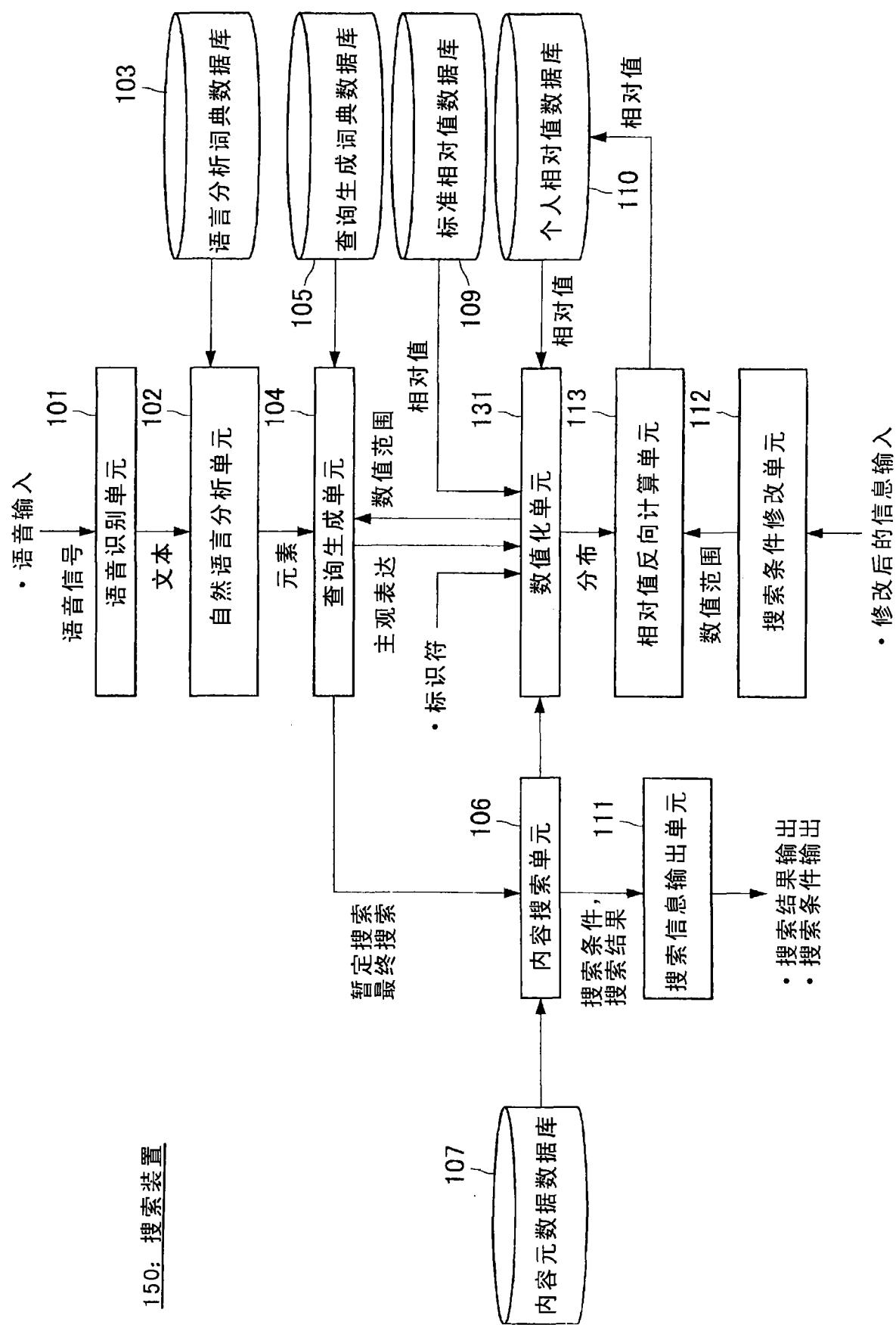


图 18

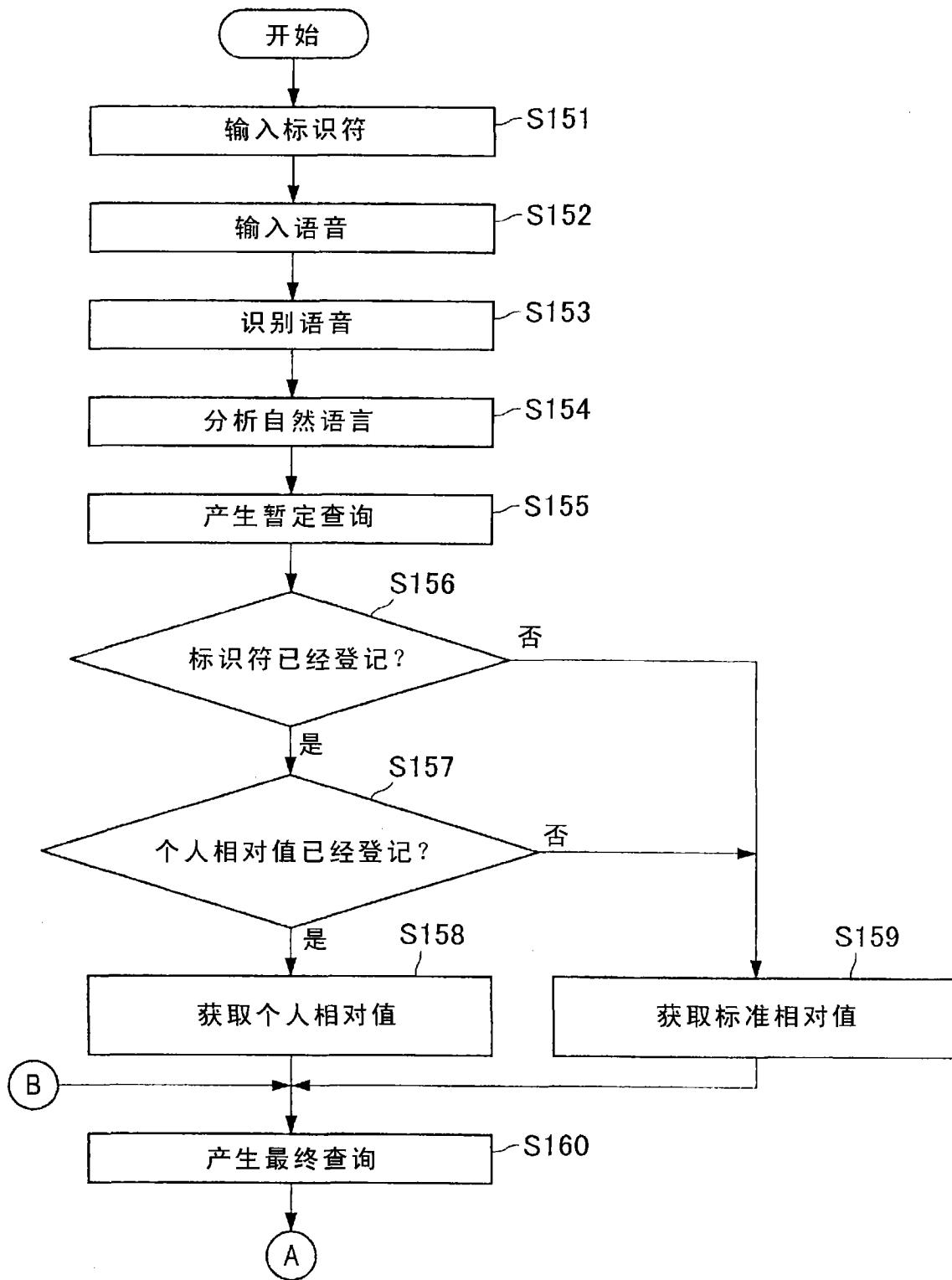


图 19

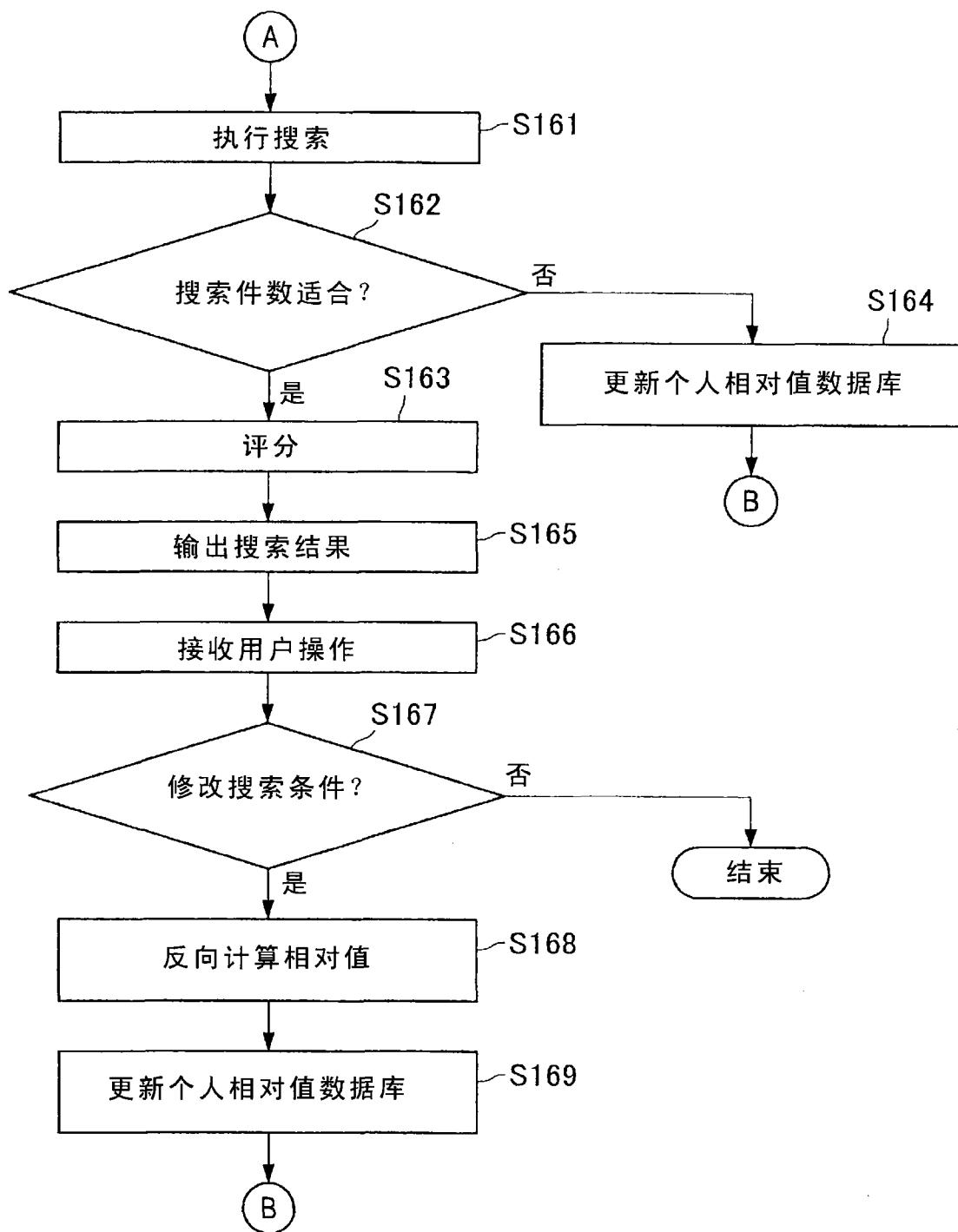


图 20

100,130,150

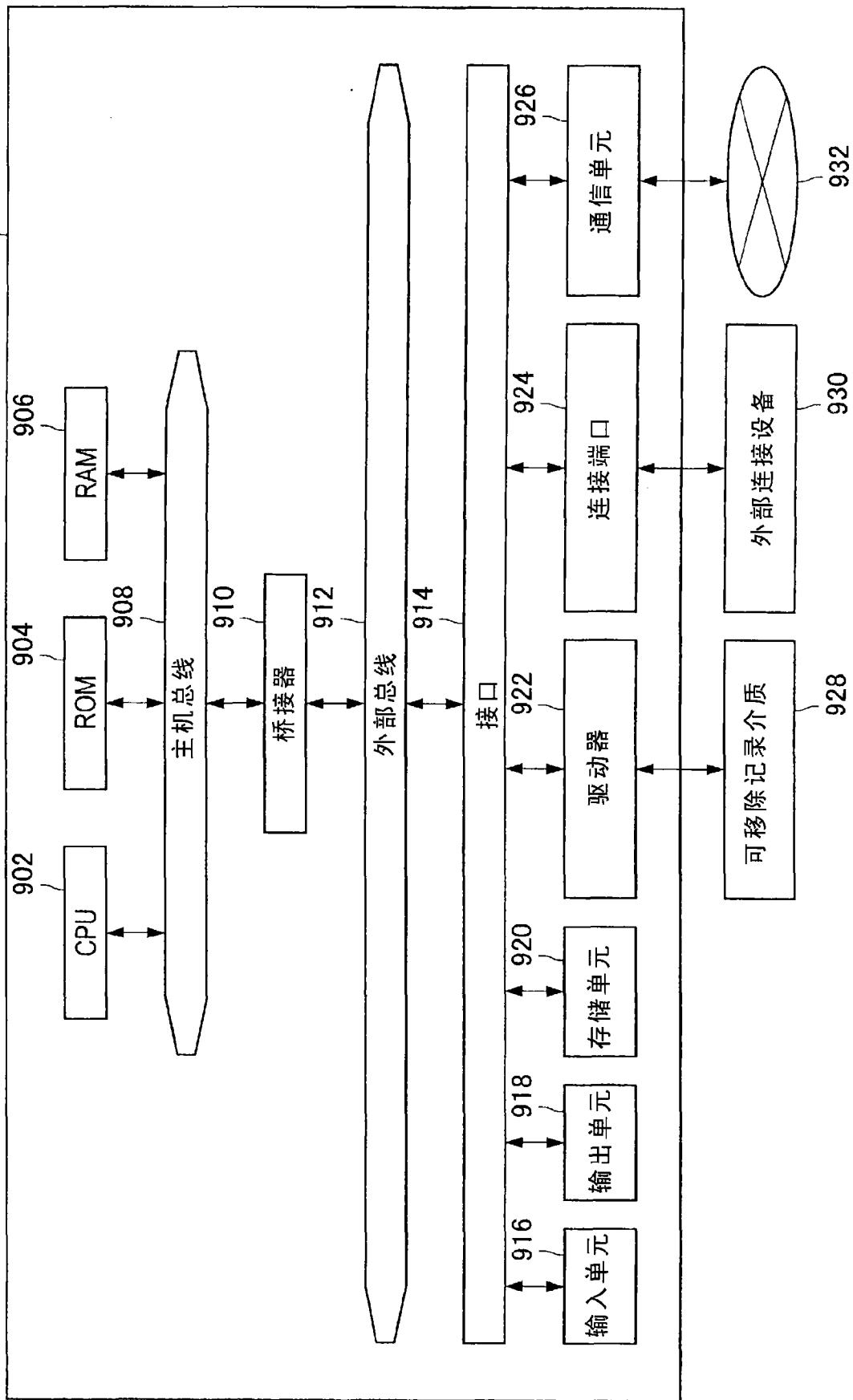


图 21