



(12) 发明专利申请

(10) 申请公布号 CN 112384911 A

(43) 申请公布日 2021. 02. 19

(21) 申请号 201980045709.4

(74) 专利代理机构 永新专利商标代理有限公司
72002

(22) 申请日 2019.03.01

代理人 徐殿军

(30) 优先权数据

2018-131457 2018.07.11 JP

(51) Int.Cl.

G06F 16/908 (2006.01)

(85) PCT国际申请进入国家阶段日

2021.01.07

(86) PCT国际申请的申请数据

PCT/JP2019/008197 2019.03.01

(87) PCT国际申请的公布数据

W02020/012700 JA 2020.01.16

(71) 申请人 株式会社东芝

地址 日本东京都

申请人 东芝数字解决方案株式会社

(72) 发明人 近藤真晖

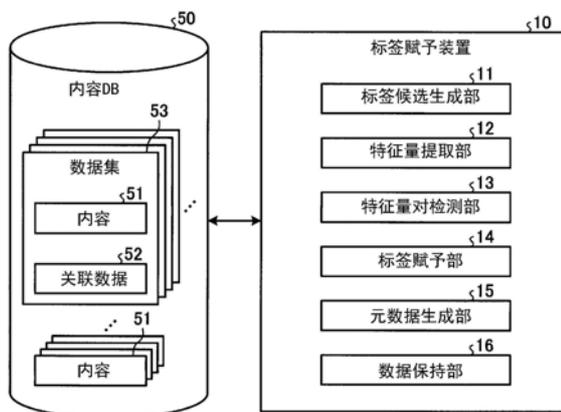
权利要求书2页 说明书18页 附图19页

(54) 发明名称

标签赋予装置、标签赋予方法以及程序

(57) 摘要

实施方式的标签赋予装置(10)具备标签候选生成部(11)、特征量对检测部(13)、以及标签赋予部(14)。标签候选生成部(11)从与内容(51)相关联的关联数据(52)生成标签候选。特征量对检测部(13)在从第一内容提取的特征量和从第二内容提取的特征量的组合中,检测特征量间的相似度最高的组合即特征量对。标签赋予部(14)将从与所述第一内容相关联的第一关联数据和与所述第二内容相关联的第二关联数据这两方生成的共同的标签候选作为标签,赋予给构成所述特征量对的各特征量。



1. 一种标签赋予装置,具备:
标签候选生成部,根据与内容相关联的关联数据生成标签候选;
特征量对检测部,在从第一内容提取的特征量和从第二内容提取的特征量的组合中检测作为特征量间的相似度最高的组合的特征量对;以及
标签赋予部,将根据与所述第一内容相关联的第一关联数据和与所述第二内容相关联的第二关联数据这两方生成的共同的标签候选作为标签,赋予给构成所述特征量对的各特征量。
2. 如权利要求1所述的标签赋予装置,
在从第三内容提取的特征量与构成所述特征量对的各特征量之间的距离为阈值以下的情况下,所述标签赋予部对从所述第三内容提取的特征量赋予与构成所述特征量对的各特征量共同的标签。
3. 如权利要求1或2所述的标签赋予装置,
所述标签赋予装置还具备特征量提取部,该特征量提取部从内容提取特征量。
4. 如权利要求3所述的标签赋予装置,
所述特征量提取部对从内容提取的特征量组进行聚类,生成每个聚类的代表特征量,
所述特征量对检测部在根据所述第一内容生成的代表特征量和根据所述第二内容生成的代表特征量的组合中,检测代表特征量间的相似度最高的组合作为所述特征量对。
5. 如权利要求1至4中任一项所述的标签赋予装置,
所述标签赋予装置还具备元数据生成部,该元数据生成部使用被赋予了标签的特征量来生成内容的元数据。
6. 如权利要求1至5中任一项所述的标签赋予装置,
所述第一内容与所述第二内容通过将内容分割而生成。
7. 如权利要求1至6中任一项所述的标签赋予装置,
所述标签赋予装置还具备标签候选修正部,该标签候选修正部生成所述标签候选的语义矢量,使语义矢量间的相似度为阈值以上的标签候选彼此共同化。
8. 如权利要求1至7中任一项所述的标签赋予装置,
所述标签赋予装置还具备:
示教对象选择部,在存在多个所述共同的标签候选、且通过决定一个标签候选与特征量的关系来决定其他标签候选与特征量的关系的情况下,选择所述一个标签候选作为示教对象;以及
操作受理部,受理对所述示教对象的标签候选与特征量的关系进行示教的用户操作。
9. 如权利要求1至8中任一项所述的标签赋予装置,
所述标签赋予装置还具备特征量对判定部,该特征量对判定部计算所述特征量对的错误得分,基于计算出的错误得分判定所述特征量对是否表示相同的对象。
10. 如权利要求9所述的标签赋予装置,
所述标签赋予装置还具备标签确认部,该标签确认部向用户确认是否对构成所述错误得分超过阈值的特征量对的各特征量、或者无法构成所述特征量对的特征量正确赋予了标签。
11. 如权利要求10所述的标签赋予装置,

所述标签赋予装置还具备确认对象筛选部,该确认对象筛选部在追加了新的内容的情
况下,使用从新的内容提取的特征量,筛选构成所述错误得分超过阈值的特征量对的各特
征量、或者无法构成所述特征量对的特征量中的需要确认是否被正确赋予了标签的特
征量。

12. 如权利要求1至11中任一项所述的标签赋予装置,
所述内容是动态图像、静态图像、语音、传感器数据中的某一个。

13. 如权利要求1至12中任一项所述的标签赋予装置,
所述关联数据是对所述内容预先赋予的文本数据、或者通过对所述内容进行规定的处
理而得的文本数据。

14. 一种标签赋予方法,包括:

根据与内容相关联的关联数据生成标签候选的步骤;

在从第一内容提取的特征量和从第二内容提取的特征量的组合中,检测作为特征量间
的相似度最高的组合的特征量对的步骤;以及

将根据与所述第一内容相关联的第一关联数据和与所述第二内容相关联的第二关联
数据这两方生成的共同的标签候选决定为构成所述特征量对的各特征量的标签的步骤。

15. 一种程序,用于使计算机实现如下功能:

根据与内容相关联的关联数据生成标签候选的功能;

在从第一内容提取的特征量和从第二内容提取的特征量的组合中,检测作为特征量间
的相似度最高的组合的特征量对的功能;以及

将根据与所述第一内容相关联的第一关联数据和与所述第二内容相关联的第二关联
数据这两方生成的共同的标签候选决定为构成所述特征量对的各特征量的标签的功能。

标签赋予装置、标签赋予方法以及程序

技术领域

[0001] 本发明的实施方式涉及标签赋予装置、标签赋予方法以及程序。

背景技术

[0002] 伴随着存储性能、线路技术的发展,从在本地的存储设备、网络上的数据库服务器等中保有的大量内容中检索并利用所希望的内容的机会增加。关于内容的检索,关键字检索最简单且便利性高。若对从内容提取的特征量赋予了相当于关键字的标签,则能够通过关键字检索高精度地检索所希望的内容。另外,若内容为动态图像内容等,则也能够进行例如找出由关键字指定的人物所出现的场景这样的场景检索。但是,通过人工对从内容提取的特征量赋予标签的作业非常繁杂,要求能够自动进行这样的标签赋予的技术。

[0003] 现有技术文献

[0004] 专利文献

[0005] 专利文献1:日本特开2016—81265号公报

[0006] 专利文献2:日本专利第3602765号公报

[0007] 非专利文献

[0008] 非专利文献1:Common Subspace for Model and Similarity:Phrase Learning for Sentence Generation from Images (ICCV 2015)

[0009] 非专利文献2:MeCab:Yet Another Part-of-Speech and Morphological Analyzer (<http://mecab.sourceforge.net/>)

[0010] 非专利文献3:Efficient estimation of word representations in vectorspace (ICIL 2013)

[0011] 非专利文献4:GloVe:Global Vectors for Word Representation (EMNLP 2014)

发明内容

[0012] 发明要解决的课题

[0013] 本发明要解决的课题在于,提供一种能够自动对从内容提取的特征量进行标签赋予的标签赋予装置、标签赋予方法以及程序。

[0014] 用来解决课题的手段

[0015] 实施方式的标签赋予装置具备标签候选生成部、特征量对检测部、以及标签赋予部。标签候选生成部从与内容相关联的关联数据生成标签候选。特征量对检测部在从第一内容提取的特征量和从第二内容提取的特征量的组合中,检测特征量间的相似度最高的组合即特征量对。标签赋予部将从与所述第一内容相关联的第一关联数据和与所述第二内容相关联的第二关联数据这两方生成的共同的标签候选作为标签,赋予给构成所述特征量对各特征量。

附图说明

- [0016] 图1是说明实施方式的基本概念的图。
- [0017] 图2是表示第一实施例的标签赋予装置的功能性构成例的框图。
- [0018] 图3是说明从动态图像中提取面部特征矢量而生成面部特征代表矢量的处理的图。
- [0019] 图4是说明从动态图像对检测矢量对并赋予标签的处理的图。
- [0020] 图5是说明从动态图像对检测矢量对并赋予标签的处理的图。
- [0021] 图6是说明从动态图像对检测矢量对并赋予标签的处理的图。
- [0022] 图7是说明生成动态图像的元数据的处理的图。
- [0023] 图8是表示第二实施例的标签赋予装置的功能性构成例的框图。
- [0024] 图9是表示第三实施例的标签赋予装置的功能性构成例的框图。
- [0025] 图10是说明第三实施例的处理的概要的图。
- [0026] 图11是说明第三实施例的处理的概要的图。
- [0027] 图12是说明第三实施例的处理的概要的图。
- [0028] 图13是表示第四实施例的标签赋予装置的功能性构成例的框图。
- [0029] 图14是表示第四实施例的处理步骤的一个例子的流程图。
- [0030] 图15是表示第四实施例的处理步骤的一个例子的流程图。
- [0031] 图16是表示第四实施例的处理步骤的一个例子的流程图。
- [0032] 图17是表示第四实施例的处理步骤的一个例子的流程图。
- [0033] 图18是表示第五实施例的标签赋予装置的功能性构成例的框图。
- [0034] 图19是表示标签赋予装置的硬件构成例的框图。

具体实施方式

[0035] 以下,参照附图,对实施方式的标签赋予装置、标签赋予方法以及程序进行详细说明。

[0036] <实施方式的概要>

[0037] 本实施方式自动进行对从内容提取的特征量的标签赋予。本实施方式中设想的内容例如包括动态图像、静态图像、语音、传感器数据等各种形式的数据。在本实施方式中,使用与这些内容相关联的关联数据,生成对从内容提取的特征量赋予的标签的候选。关联数据例如可以是如内容的说明文字、字幕信息等预先对内容赋予的文本数据,也可以是通过内容进行语音识别、文字识别、图像说明文字生成(captioning)等规定的处理而得的文本数据。

[0038] 若从内容提取的特征量和从关联数据生成的标签候选为各一个,则能够唯一地确定特征量与标签的对应关系。但是,这样的状况极其罕见,一般来说,多为从内容提取多个特征量、从关联数据也生成多个标签候选。在该情况下,由于无法唯一地确定哪个标签候选与哪个特征量对应,因此无法自动进行对特征量的标签赋予。因此,在本实施方式中,使用多个内容和关联数据的数据集,将数据集间的共同部分作为线索,确定特征量与标签候选的对应关系。

[0039] 图1是对实施方式的基本概念进行说明的图。如图1所示,作为关联有关联数据的

内容,假设存在内容A与内容B。这里,从内容A与内容B分别提取三个特征量。另外,从内容A的关联数据生成“a”、“b”、“c”这三个标签候选,从内容B的关联数据生成“a”、“d”、“e”这三个标签候选。即,在内容A及其关联数据的数据集和内容B及其关联数据的数据集中,从关联数据生成的标签候选仅一个是共同的。在本实施方式中,像这样选择从关联数据生成的标签候选仅有一个是共同的数据集的组合。

[0040] 在本实施方式中,首先,进行从内容A、B分别提取特征量的处理、以及从各个关联数据生成标签候选的处理。之后,将从内容A提取的特征量和从内容B提取的特征量映射到共同的特征量空间。然后,在从内容A提取的特征量和从内容B提取的特征量的组合中,检测特征量间的相似度最高的组合(例如,特征量间的距离最小的组合)作为特征量对。这样检测出的特征量对表示内容A与内容B这两方中所包含的共同的(相同的)对象的可能性高。因此,针对构成该特征量对各特征量,赋予从内容A的关联数据与内容B的关联数据这两方生成的共同的标签候选(在图1的例子中,为标签候选“a”)作为标签。

[0041] 在以上的处理中,从内容A与内容B中检测一个特征量对,对构成该特征量对各特征量赋予标签,因此内容A、B均残留未赋予标签的特征量。但是,内容A、B这两方均能够通过选择与其他数据集的组合反复进行以上的处理,对未赋予标签的特征量也同样地赋予标签。另外,若使用多个数据集反复进行以上的处理,对多个特征量赋予标签,则对于从未关联有关联数据的内容提取的特征量,也能够通过最近邻搜索来自动赋予标签。即,若从未关联有关联数据的内容提取的特征量足够靠近标签赋予完毕的特征量,则能够赋予相同的标签。

[0042] 以下,作为易于理解的例子,以内容为TV节目(动态图像)、关联数据为对TV节目的内容进行说明的节目说明文字的情况为例,对应用了本实施方式的概念的具体的实施例进行详细说明。另外,在以下的说明中,对具有相同的功能的构成要素标注相同的附图标记,并适当省略重复的说明。

[0043] <第一实施例>

[0044] TV节目一般成套地提供动态图像与节目说明文字。节目说明文字是对TV节目的内容进行说明的文本数据,大多记载有在TV节目出场的主要出场人物的姓名。在本实施例中,对从TV节目的动态图像得到的主要出场人物的面部特征代表矢量(代表特征量的一个例子)自动赋予从节目说明文字得到的主要出场人物的姓名作为标签。

[0045] 首先,对本实施例的处理的概要进行说明。在本实施例中,对TV节目的动态图像的各帧进行面部检测·面部特征矢量生成的处理。此时,由于TV节目的主要出场人物在该TV节目中出现较多,因此从动态图像提取较多的面部特征矢量。另外,从动态图像提取的同一人物的面部特征矢量在特征矢量空间中存在于附近。即,从动态图像提取的多个主要出场人物的面部特征矢量在特征矢量空间上大量存在,可认为是更密集的。另一方面,由于主要出场人物以外的人物(非主要出场人物)的出场次数较少,因此可认为在特征矢量空间上密集率低。因而,在将特征矢量空间上稀疏的面部特征矢量作为偏离值的基础上,若针对特征矢量空间上的面部特征矢量组设 K =主要出场人物数而进行 K -means聚类,则可得到主要出场人物的面部特征代表矢量。

[0046] 在多个TV节目中,有时存在主要出场人物为共同的节目。这里,将在多个TV节目中共同出场的主要出场人物称作共同出场人物。此时,通过适当地选择TV节目的组合,能够筛

选共同出场人物。例如,在存在人物a与人物b作为主要出场人物出场的TV节目X和人物a与人物c作为主要出场人物出场的TV节目Y时,由这些TV节目X与TV节目Y的组合而得的共同出场人物为a。

[0047] 在本实施例中,首先,通过对从各TV节目的动态图像提取的面部特征矢量进行上述的K-means聚类,按每个TV节目生成主要出场人物的面部特征代表矢量。此时,由于在一个TV节目中多个主要出场人物出现较多,因此无法将面部特征代表矢量与主要出场人物的姓名唯一地建立对应。因此,选择共同出场人物为1名的多个TV节目的组合,将从这些TV节目的动态图像分别生成的面部特征代表矢量映射到共同的特征矢量空间。例如,在将共同出场人物为1名的两个TV节目中的一方的TV节目的动态图像设为第一动态图像、将另一方的TV节目的动态图像设为第二动态图像时,将从第一动态图像生成的多个面部特征代表矢量、以及从第二动态图像生成的多个面部特征代表矢量映射到共同的特征矢量空间。

[0048] 这里,由于两个TV节目的共同出场人物为1名,因此从第一动态图像生成的面部特征代表矢量和从第二动态图像生成的面部特征代表矢量的组合中的面部特征代表矢量间的相似度最高的组合(特征量对的一个例子。以下,将其称作矢量对。)属于共同出场人物的可能性较高。因而,能够对构成矢量对的各面部特征代表矢量赋予共同出场人物的姓名作为标签。另外,在以下的说明中,列举了使用面部特征代表矢量间的距离(例如欧几里德距离)判定面部特征代表矢量间的相似度的例子,但也可以使用两个面部特征代表矢量的内积来判定两者的相似度。在使用面部特征代表矢量间的距离判定相似度的情况下,距离越小相似度越高,但在使用两个面部特征代表矢量的内积判定相似度的情况下,内积越大相似度越高。

[0049] 接下来,对本实施例的标签赋予装置的构成进行说明。图2是表示本实施例的标签赋予装置10的功能性构成例的框图。如图2所示,该标签赋予装置10具备标签候选生成部11、特征量提取部12、特征量对检测部13、标签赋予部14、元数据生成部15、以及数据保持部16。

[0050] 该标签赋予装置10从储存于内容DB50的内容51提取特征量,对提取的特征量赋予从关联数据52生成的标签候选中的某一个作为标签。在本实施例中,如上述那样,设想由内容51与关联数据52的组合构成的数据集53为TV节目、内容51为TV节目的动态图像、关联数据52为节目说明文字的情况。

[0051] 标签候选生成部11从与内容51相关联的关联数据52生成标签候选。例如,在关联数据52为TV节目的节目说明文字的情况下,标签候选生成部11对该节目说明文字进行词素分析,从其结果取得在TV节目中出场的主要出场人物的姓名并将其作为标签候选。另外,关联数据52只要是能够取得成为标签候选的主要出场人物的姓名的文本数据即可,例如可以是对与动态图像链接的语音进行语音识别处理而得的文本数据,也可以是通过构成动态图像的帧图像应用例如非专利文献1所示的图像说明文字生成的技术而得的文本数据。另外,也可以将对动态图像赋予的字幕信息的文本数据用作关联数据52。

[0052] 特征量提取部12从内容51提取特征量。例如,在内容51为TV节目的动态图像的情况下,特征量提取部12对动态图像的各帧进行面部检测·面部特征矢量生成的处理。由此,可网罗在TV节目出场的所有出场人物的面部特征矢量。另外,由于从图像中检测面部并提取其特征量(面部特征矢量)的技术被广泛知晓,因此这里省略详细的说明。

[0053] 另外,特征量提取部12在处理对象的内容51与关联数据52建立了关联的情况下,将从该内容51提取的特征量组聚类而生成每个聚类(cluster)的代表特征量。例如,在内容51为TV节目的动态图像、且节目说明文字与该动态图像建立了关联而作为关联数据52的情况下,特征量提取部12对从动态图像提取的面部特征矢量组进行K-means聚类,生成每个聚类的面部特征代表矢量。该K-means聚类的K的值为由标签候选生成部11生成的标签候选的数量、即在TV节目中出场的主要出场人物的数量。由此,可得到与主要出场人物的数量相同的数量的面部特征代表矢量。另外,在从内容51提取的特征量的数量与标签候选的数量相等的情况下,也可以不进行对特征量的聚类(代表特征量的生成)。

[0054] 由标签候选生成部11生成的标签候选(主要出场人物名)的一览以及由特征量提取部12生成的多个面部特征代表矢量(代表特征量)作为本实施例中的处理的中间数据,与内容51(TV节目的动态图像)的识别信息建立对应地保持于数据保持部16。

[0055] 特征量对检测部13使用储存于内容DB50的数据集53中的从关联数据52生成的标签候选仅有一个是共同的数据集53的组合,进行特征量对的检测。例如,假设储存于内容DB50的数据集53为TV节目、内容51为TV节目的动态图像、关联数据52为节目说明文字。而且,对于各数据集53,由标签候选生成部11生成主要出场人物的姓名的一览,作为标签候选的一览,由特征量提取部12生成与主要出场人物数量相同的数量的面部特征代表矢量,作为代表特征量。在该情况下,特征量对检测部13在作为数据集53而储存于内容DB50的TV节目中,选择主要出场人物仅有1名是共同的TV节目的组合。然后,当将一方的TV节目的动态图像设为第一动态图像,并将另一方的TV节目的动态图像设为第二动态图像时,特征量对检测部13在从第一动态图像生成的面部特征代表矢量和从第二动态图像生成的面部特征代表矢量的组合中,检测面部特征代表矢量间的距离最小(相似度最高)的面部特征代表矢量的组合作为矢量对。

[0056] 另外,在本实施例中,设想从一个内容51(TV节目的动态图像)中提取多个表示相同的对象(主要出场人物)的特征量(面部特征矢量),特征量提取部12对从内容51提取的特征量进行聚类。但是,在对从内容51提取的特征量分别赋予标签的情况下,不需要特征量的聚类(代表特征量的生成)。这样,在特征量提取部12不进行特征量的聚类的情况下,特征量对检测部13不针对代表特征量,而只要针对特征量的组合判定相似度,并检测特征量对即可。

[0057] 标签赋予部14对构成由特征量对检测部13检测出的特征量对各特征量,赋予从上述的数据集53的组合中包含的各关联数据52生成的共同的标签候选作为标签。例如,在由特征量对检测部13检测出上述的矢量对的情况下,对构成该矢量对各面部特征代表矢量赋予共同的主要出场人物的姓名作为标签。

[0058] 另外,标签赋予部14在从未关联有关联数据52的内容51提取的特征量(面部特征矢量)与构成由特征量对检测部13检测出的特征量对(矢量对)的各特征量(面部特征代表矢量)类似的情况下,能够对从未关联有关联数据52的内容51提取的特征量赋予与对构成特征量对各特征量赋予的标签相同的标签。该处理例如能够通过后述的最近邻搜索的方法实现。

[0059] 元数据生成部15使用被赋予了标签的特征量来生成内容51的元数据。例如,在特征量提取部12如上述那样从动态图像的各帧中提取了面部特征矢量的情况下,记录所提取

的面部特征矢量与帧编号的对应关系。之后,在由标签赋予部14对面部特征矢量赋予了标签的情况下,能够对面部特征矢量赋予的标签赋予给提取了该面部特征矢量的帧编号。赋予给该帧编号的标签成为用于检索从TV节目中检索所希望的主要出场人物出场的场景的元数据。

[0060] 接下来,对由本实施例的标签赋予装置10进行的处理的具体例进行说明。首先,参照图3,对生成TV节目的主要出场人物的面部特征代表矢量与主要出场人物一览的处理进行说明。图3是说明从动态图像提取面部特征矢量并生成面部特征代表矢量的处理的图。

[0061] 标签候选生成部11对TV节目的节目说明文字进行词素分析。然后,从该词素分析的结果取得在TV节目中出场的主要出场人物的姓名作为标签候选,生成主要出场人物一览。该主要出场人物一览中包含的主要出场人物的数量为标签候选数。在图3中,示出了从与动态图像A相关联的节目说明文字中取得人物a、人物b、人物c这3名主要出场人物名来作为标签候选的例子。另外,在对节目说明文字的词素分析中,例如能够使用非专利文献2所公开的MeCab等。

[0062] 如图3所示,特征量提取部12从TV节目的动态图像A的各帧中进行出场人物的面部检测,生成面部特征矢量。由此,从一个动态图像A生成多个人物以及多个面部特征矢量组。即,可网罗在TV节目中出场的人物的面部特征矢量。

[0063] 接下来,特征量提取部12对从动态图像A生成的面部特征矢量组进行K-means聚类,将各聚类的质心(centroid)作为面部特征代表矢量。此时的K的值是由标签候选生成部11生成的标签候选数,即主要出场人物一览中包含的主要出场人物名的数量。由此,可得到与TV节目中出场的主要出场人物的数量对应的数量的面部特征代表矢量。在图3中,可得到三个面部特征代表矢量。这些面部特征代表矢量被认为是表示主要出场人物一览中包含的任意一个主要出场人物的特征量。但是,在从动态图像A生成的面部特征矢量组中有可能也包含主要出场人物以外的人物(非主要出场人物)。因此,进行如下的处理,去除非主要出场人物的面部特征矢量。

[0064] 首先,计算与面部特征代表矢量属于同一聚类的所有面部特征矢量间的距离。这里,距离在统计上成为偏离值的矢量被认为是具有与面部特征代表矢量不同的标签的人物(非主要出场人物)。因此,将这样的矢量视为“偏离值矢量”。

[0065] 接下来,在去除了偏离值矢量的基础上,再次进行K-means聚类。

[0066] 反复进行该处理,直到聚类的结果收敛为止(直到偏离值矢量消失为止)。然后,将聚类的结果收敛时的聚类内最大距离与作为该聚类的质心的面部特征代表矢量建立对应地记录于数据保持部16。该值是在后述的元数据生成时的最近邻搜索中被用作阈值的值。

[0067] 通过以上的处理,可生成动态图像中的主要出场人物的面部特征代表矢量与主要出场人物一览。但是,在该状态下,除了面部特征代表矢量与主要出场人物仅各存在一个的特殊情况以外,无法将面部特征代表矢量与主要出场人物建立对应。因此,无法自动进行对面部特征矢量对的标签赋予。因此,通过组合使用多个数据集53,进行面部特征代表矢量与主要出场人物的对应关系的筛选。因此,对作为数据集53储存于内容DB50的其他TV节目也进行上述(1)~(3)的处理,生成动态图像中的主要出场人物的面部特征代表矢量与主要出场人物一览。然后,选择共同出场人物为1名的TV节目的组合。

[0068] 这里,假设从三个TV节目的动态图像A、B、C分别生成了三个面部特征代表矢量。另

外,假设从与动态图像A相关联的节目说明文字中作为标签候选生成了人物a、人物b、人物c,从与动态图像B相关联的节目说明文字中作为标签候选生成了人物a、人物d、人物e,从与动态图像C相关联的节目说明文字中作为标签候选生成了人物b、人物d、人物f。在该情况下,动态图像A与动态图像B的共同出场人物为a这1名,动态图像A与动态图像C的共同出场人物为b这1名,动态图像B与动态图像C的共同出场人物为d这1名。

[0069] 在本实施例中,使用这些共同出场人物为1名的TV节目的动态图像的组合,进行面部特征代表矢量与主要出场人物的对应关系的筛选。另外,这里对使用两个动态图像的组合(以下,将其称作动态图像对)的例子进行说明,但也可以使用三个以上的动态图像的组合。

[0070] 以下,参照图4至图6,对使用上述的动态图像对进行面部特征代表矢量与主要出场人物的对应关系的筛选的处理进行说明。图4至图6是说明从动态图像对检测矢量对并赋予标签的处理的图。

[0071] 特征量对检测部13首先如图4所示那样,选择由动态图像A与动态图像B构成的动态图像对,将从动态图像A生成的三个面部特征代表矢量和从动态图像B生成的三个面部特征代表矢量映射到共同的特征矢量空间。然后,在从动态图像A生成的面部特征代表矢量和从动态图像B生成的面部特征代表矢量的组合中检测面部特征代表矢量间的距离最小的组合即矢量对。一般来说,同一人物的面部特征矢量存在于附近。因而,可以推断,构成从由动态图像A与动态图像B构成的动态图像对检测出的矢量对的两个面部特征代表矢量为动态图像A与动态图像B的共同出场人物即人物a的面部特征代表矢量。因此,标签赋予部14对构成该矢量对的两个面部特征代表矢量赋予人物a的标签。

[0072] 特征量对检测部13接着如图5所示那样,选择由动态图像A与动态图像C构成的动态图像对,将从动态图像A生成的三个面部特征代表矢量、以及从动态图像C生成的三个面部特征代表矢量映射到共同的特征矢量空间。然后,在从动态图像A生成的面部特征代表矢量和从动态图像C生成的面部特征代表矢量的组合中检测面部特征代表矢量间的距离最小的组合即矢量对。可以推断,构成这样检测出的矢量对的两个面部特征代表矢量为动态图像A与动态图像C的共同出场人物即人物b的面部特征代表矢量。因此,标签赋予部14对构成该矢量对的两个面部特征代表矢量赋予人物b的标签。另外,关于动态图像A,由于对三个面部特征代表矢量中的两个分别赋予了人物a与人物b的标签,因此剩余的面部特征代表矢量能够推断为人物c的标签。因此,标签赋予部14对该剩余的面部特征代表矢量赋予人物c的标签。

[0073] 特征量对检测部13接着如图6所示那样,选择由动态图像B与动态图像C构成的动态图像对,将从动态图像B生成的三个面部特征代表矢量和从动态图像C生成的三个面部特征代表矢量映射到共同的特征矢量空间。然后,在从动态图像B生成的面部特征代表矢量和从动态图像C生成的面部特征代表矢量的组合中检测面部特征代表矢量间的距离最小的组合即矢量对。可以推断,构成这样检测出的矢量对的两个面部特征代表矢量为动态图像B与动态图像C的共同出场人物即人物d的面部特征代表矢量。因此,标签赋予部14对构成该矢量对的两个面部特征代表矢量赋予人物d的标签。另外,关于动态图像B,由于对三个面部特征代表矢量中的两个分别赋予了人物a与人物d的标签,因此剩余的面部特征代表矢量能够推断为人物e的标签。因此,标签赋予部14对该剩余的面部特征代表矢量赋予人物e的标签。

另外,关于动态图像C,由于对三个面部特征代表矢量中的两个分别赋予了人物b与人物d的标签,因此剩余的面部特征代表矢量能够推断为人物f的标签。因此,标签赋予部14对该剩余的面部特征代表矢量对赋予人物f的标签。

[0074] 如以上那样,通过使用共同出场人物为1名的动态图像对进行面部特征代表矢量与主要出场人物的对应关系的筛选,能够对从各动态图像生成的面部特征代表矢量自动赋予主要出场人物名的标签。另外,若对面部特征代表矢量赋予标签,则能够使用被赋予了该标签的面部特征代表矢量,生成例如动态图像的场景检索用的元数据。

[0075] 以下,参照图7,对使用被赋予了标签的面部特征代表矢量生成动态图像的元数据的处理的具体例进行说明。图7是说明生成动态图像的元数据的处理的图。

[0076] 元数据生成部15从内容DB50取出想要生成元数据的动态图像D并将其传递给特征量提取部12,请求面部特征矢量的生成。该动态图像D也可以是未关联有关联数据52的内容51。另外,动态图像D也可以是通过上述的处理进行了对面部特征代表矢量的标签赋予的动态图像(上述的动态图像A、B、C等)。如图7所示,特征量提取部12对从元数据生成部15传递来的动态图像D,与对上述的动态图像A的处理同样地进行生成面部特征矢量的处理。即,从动态图像D的各帧中进行出场人物的面部检测,生成面部特征矢量。此时,元数据生成部15将由特征量提取部12提取的面部特征矢量和提取了该面部特征矢量的场景的帧编号建立对应地记录在数据保持部16中。

[0077] 接下来,元数据生成部15将从动态图像D生成的所有面部特征矢量传递给标签赋予部14,并请求基于使用了被赋予标签的面部特征代表矢量的最近邻搜索的标签赋予。标签赋予部14根据来自元数据生成部15的请求,将从动态图像D生成的各面部特征矢量作为查询矢量,对被赋予了标签的面部特征代表矢量组进行最近邻搜索。最近邻搜索是从矢量组中搜索最接近查询矢量的矢量的处理。通过该最近邻搜索,针对从动态图像D生成的各面部特征矢量确定最接近的面部特征代表矢量。

[0078] 接下来,标签赋予部14针对从动态图像D生成的各面部特征矢量,计算与被确定为最邻近的面部特征代表矢量之间的距离。然后,在计算出的距离为上述的阈值以下的情况下、即面部特征代表矢量的聚类内最大距离以下的情况下,对面部特征矢量赋予与面部特征代表矢量相同的标签。由此,能够抑制对从动态图像D生成的面部特征矢量赋予错误的标签。

[0079] 元数据生成部15在如以上那样对从动态图像D生成的面部特征矢量赋予标签时,参照记录在数据保持部16中的面部特征矢量与帧编号的对应关系,对提取了面部特征矢量的帧编号赋予对该面部特征矢量赋予的标签。对该帧编号赋予的标签成为用于从TV节目中检索所希望的主要出场人物出场的场景的元数据。

[0080] 如以上那样,根据本实施例的标签赋予装置10,通过组合使用多个包含内容51与关联数据52的数据集53,进行将特征量与标签候选的对应关系筛选处理,能够对从内容51提取的特征量自动赋予标签。另外,根据本实施例的标签赋予装置10,通过进行使用了被赋予标签的特征量的简单的最近邻搜索,能够对从其他内容51提取的特征量自动赋予标签,例如能够自动生成内容51的场景检索用的元数据等。

[0081] 另外,以上的说明以存在多个数据集53为前提,但通过分割使用一个内容51,也能够利用与上述处理相同的处理来实现对特征量的自动标签赋予。例如,将一个动态图像分

割为多个,将分割后的动态图像的每一个视为单独的内容51。此时分割后的各动态图像需要分别与关联数据52建立关联,但例如通过采用对动态图像赋予的字幕信息、对与动态图像链接的语音的语音识别结果等文本数据作为关联数据52,能够将关联数据52与分割后的动态图像分别建立关联。这是因为,字幕信息、语音识别结果的文本数据与动态图像的各场景建立了对应(具有时间信息),在分割动态图像时这些信息也被自动分割。

[0082] 另外,以上是设想TV节目的动态图像为内容51的情况下的说明,即使处理对象的内容51是其他动态图像,也能够同样地对特征量进行标签的自动赋予、元数据的自动生成。作为TV节目以外的应用例,例如有动态图像分发服务。另外,例如,通过使用工厂的进出记录作为关联数据52,也能够对从由设置于工厂内的相机拍摄到的工厂内的影像中提取的作业者的特征量自动赋予作业者的姓名、雇员编号等作为标签。另外,在麦克风、相机这样的具备多个模态(modality)的家用机器人中,例如通过组合对由麦克风取得的语音的语音识别结果的文本数据和由相机拍摄的影像,能够对从影像提取的人物的特征量自动赋予(所谓的概念习得)从语音识别结果的文本数据生成的标签候选作为标签。

[0083] 另外,在以上的说明中,例示了内容51为动态图像的情况,但即使内容51为静态图像、语音、传感器数据,也能够通过与上述的例子相同的处理来实现对从内容51提取的特征量的自动标签赋予。以下,列举其例子。

[0084] 首先,在静态图像的情况下,例如,将拍摄有多个人物的照片作为内容51,将照片的标题作为关联数据52。在该情况下,通过对照片进行面部检测以及面部特征矢量的生成,可得到多个面部特征矢量。另外,根据照片的标题可得到多个标签候选。通过使用这些,能够与上述的例子同样地实现针对面部特征矢量的自动标签赋予。

[0085] 接下来,在语音的情况下,例如,将广播剧的语音作为内容51,将剧的说明文字作为关联数据52。在该情况下,根据语音数据的波形生成确定讲话者的特征矢量,并根据剧的说明文字生成讲话者的姓名等标签候选。通过使用这些,能够与上述的例子同样地实现针对确定讲话者的特征矢量的自动标签赋予。

[0086] 接下来,在传感器数据的情况下,例如将安装于监视对象的马达的旋转传感器输出的马达转速的时间序列数据作为内容51。另外,设想存在用户观测马达的状态并将其记录的文章,将该记录文字作为关联数据52。在该情况下,从记录文字中取出例如过度旋转、过小旋转、正常旋转等关键字并将它们作为标签候选。另外,通过从旋转传感器输出的马达转速的时间序列数据提取马达转速作为特征量,并对所提取的马达转速以关键字数进行K-means聚类,由此可得到过度旋转的代表特征量、过小旋转的代表特征量、正常旋转的代表特征量。通过使用这些,能够与上述的例子同样地对马达转速自动赋予过度旋转、过小旋转、正常旋转等标签。

[0087] <第二实施例>

[0088] 接下来,对第二实施例进行说明。本实施例相对于上述的第一实施例,附加了吸收标签候选的表述不一致的功能。由于其他功能与上述的第一实施例相同,因此以下仅对本实施例特有的功能进行说明。

[0089] 上述的第一实施例以针对TV节目中出场的相同的主要出场人物生成相同的标签候选为前提。但是,实际上,即使是同一人物,也存在以本名与艺名等不同的表述记载于节目说明文字中的情况。在该情况下,由于根据节目说明文字生成的标签候选中产生表述不

一致,因此需要将其吸收。因此,在本实施例中,通过将标签候选的单词语义矢量化,并使语义矢量类似的标签候选彼此共同化,作为表述不一致的对策。

[0090] 图8是表示第二实施例的标签赋予装置10的功能性构成例的框图。本实施例的标签赋予装置10是相对于图2所示的第一实施例的构成追加了标签候选修正部17的构成。

[0091] 标签候选修正部17生成由标签候选生成部11从关联数据52生成的标签候选的语义矢量,使语义矢量间的相似度为阈值以上的标签候选彼此共同化。语义矢量是表示单词的含义的矢量,例如已知有使用单词间的同现信息来表示含义的方法等。

[0092] 标签候选修正部17首先进行用于生成标签候选的语义矢量的模型的学习。即,标签候选修正部17首先收集储存于内容DB50的所有关联数据52(文本数据),对收集到的文本数据,通过使用了上述的Mecab等的词素分析,进行单词分割处理。然后,标签候选修正部17使用通过单词分割而得的单词组,学习用于生成语义矢量的模型。在该学习中,例如能够利用非专利文献3所公开的Word2Vec等。另外,也可以使用非专利文献4所公开的GloVe等。

[0093] 之后,标签候选修正部17在由标签候选生成部11从关联数据52生成主要出场人物的姓名作为标签候选时,使用学习过的模型,对主要出场人物进行语义矢量化。然后,标签候选修正部17针对语义矢量的全部组合,计算语义矢量间的相似度。此时,认为相似度为阈值以上的语义矢量的组合具有类似的含义。因此,提取这样的语义矢量的组合。另外,阈值为超参数。另外,语义矢量的组合并不一定是两个语义矢量的组合,也可以组合三个以上的语义矢量。

[0094] 标签候选修正部17对如以上那样提取的语义矢量的组合进行使标签候选共同化的处理。所谓标签候选的共同化,是指使标签候选的表述为相同的表述。例如,能够以上述的单词分割处理的结果为基础,计算每个标签候选的单词出现数。标签候选修正部17使用该单词出现数最多的标签候选的表述,使语义矢量的相似度为阈值以上的标签候选彼此共同化。

[0095] 如以上说明那样,根据本实施例,由于能够通过标签候选修正部17吸收标签候选的表述不一致,因此在同一人物以不同的表述记载于节目说明文字中的情况下等,即使在从关联数据52生成的标签候选中产生了表述不一致,也能够将其吸收的基础上,进行与上述的第一实施例相同的处理。由此,能够消除例如尽管是同一人物但被赋予不同的标签的问题。

[0096] <第三实施例>

[0097] 接下来,对第三实施例进行说明。本实施例相对于上述的第一实施例,附加了用于对无法自动赋予标签的特征量高效地进行基于手动的标签赋予的功能。由于其他功能与上述的第一实施例相同,因此以下仅对本实施例特有的功能进行说明。

[0098] 上述的第一实施例使用共同出场人物为1名的动态图像对,对被推断为表示共同出场人物的面部特征代表矢量,自动赋予共同出场人物的姓名作为标签。但是,在内容DB50内的动态图像中,设想存在即使尝试了所有与其他动态图像的组合也无法使共同出场人物为1名的动态图像。在这样的情况下,用户需要手动进行标签赋予。但是,对无法自动进行标签赋予的所有模式通过手动赋予标签的作业是费事的。因此,在本实施例中,若对一个面部特征代表矢量手动地赋予标签,则选择能够自动进行对其他面部特征代表矢量的标签赋予的标签候选作为示教对象,并提示给用户。具体而言,通过使用多个具有多个共同出场人物

的动态图像组合,对这些多个动态图像组合中的共同出场人物的重复进行检查,确定一旦进行了示教后就通过消除法也自动决定其他人物的标签那样的人物。然后,通过使用该人物优先地对该人物进行示教操作,从而以较少的劳力实现标签赋予。

[0099] 图9是表示第三实施例的标签赋予装置10的功能性构成例的框图。本实施例的标签赋予装置10是相对于图2所示的第一实施例的构成追加了示教对象选择部18和操作受理部19的构成。

[0100] 示教对象选择部18在存在多个共同的标签候选,并通过决定标签候选与特征量的关系来决定其他标签候选与特征量的关系的情况下,选择一个标签候选作为示教对象。例如,示教对象选择部18通过使用多个存在多个共同出场人物的动态图像组合,并检查这些多个动态图像组合中的共同出场人物的重复,来确定成为示教对象的人物。

[0101] 操作受理部19受理对由示教对象选择部18选择的示教对象的标签候选与特征量的关系进行示教的用户操作。例如操作受理部19将作为示教对象的标签候选的人物名与从构成动态图像组合的各动态图像生成面部特征代表矢量时的面部动态图像一起提示给用户,受理选择与作为示教对象的标签候选的人物名一致的面部图像的用户操作。

[0102] 以下,参照图10至图12,对本实施例的处理的具体例进行说明。图10至图12是说明本实施例的处理的概要的图。这里,以图10所示的使用三个动态图像对 α 、 β 、 γ 的情况为例进行说明。动态图像对 α 的共同出场人物为人物a、人物b、人物c这3名,动态图像对 β 的共同出场人物为人物b、人物c这2名,动态图像对 γ 的共同出场人物为人物a、人物c、人物d这3名。

[0103] 示教对象选择部18首先根据共同出场人物的数量对所构建的动态图像对进行分类。在上述的例子中,分为共同出场人物为2名的动态图像对 β 、以及共同出场人物为3名的动态图像对 α 、 γ 。

[0104] 接下来,示教对象选择部18针对共同出场人物少的动态图像对与共同出场人物多的动态图像对的组合,检查共同出场人物的重复。然后,选择不重复的共同出场人物的数量少的动态图像对的组合、即共同出场人物的否定逻辑积(NOT AND)的结果少的动态图像对的组合。在上述的例子中,如图10所示,在动态图像对 α 与动态图像对 β 的组合中不重复的共同出场人物为1名(人物a)、在动态图像对 α 与动态图像对 γ 的组合中不重复的共同出场人物为3名(人物a、人物b、人物d),因此选择动态图像对 α 与动态图像对 β 的组合。

[0105] 接下来,示教对象选择部18在所选择的动态图像对的组合中选择共同出场人物的人数最少的动态图像对。在上述的例子中,在动态图像对 α 与动态图像对 β 的组合中,动态图像对 α 的共同出场人物为人物a、人物b、人物c这3名、动态图像对 β 的共同出场人物为人物b、人物c这2名,因此选择动态图像对 β 。

[0106] 接下来,示教对象选择部18针对所选择的动态图像对的共同出场人物,适当地选择一个按人数存在的标签候选。该标签候选是作为示教对象而从主要出场人物一览中选择的主要出场人物的姓名。然后,操作受理部19将由示教对象选择部18选择的标签候选和与按人数存在面部特征代表矢量的每一个对应的面部图像的一览一起提示给用户。在上述的例子中,作为动态图像对 β 的共同出场人物的人物b或者人物c的姓名与面部图像一览一起被提示给用户。

[0107] 用户在由操作受理部19提示标签候选与面部图像一览时,从面部图像一览中选择

与标签候选一致的面部图像。该操作为对所提示的标签候选与共同出场人物的面部特征代表矢量的关系进行示教的操作,当该示教操作被操作受理部19受理时,标签赋予部14对该面部特征代表矢量赋予标签候选的人物名作为标签。

[0108] 这里,如果动态图像对的共同出场人物为2名,则若针对一方的共同出场人物确定了标签候选与面部特征代表矢量的对应关系的话,那么针对另一方的共同出场人物也确定了标签候选与面部特征代表矢量的对应关系。因而,能够对未赋予标签的面部特征代表矢量自动赋予标签。另外,若动态图像对的共同出场人物为3名以上,则通过反复进行示教对象的标签候选的选择、以及将所选择的标签候选与面部图像一览一起提示并受理用户的示教操作的处理,能够针对动态图像对的所有共同出场人物确定标签候选与面部特征代表矢量的对应关系,从而对各个面部特征代表矢量赋予标签。另外,通过在未赋予标签的共同出场人物为1名的阶段应用上述的第一实施例的处理,也能够对未赋予标签的面部特征代表矢量自动赋予标签。

[0109] 通过以上的步骤,对所选择的动态图像对的组合中的一方的动态图像对完成共同出场人物的标签赋予。在上述的例子中,在动态图像对 α 与动态图像对 β 的组合中,完成作为动态图像对 β 的共同出场人物的人物b与人物c的标签赋予。即,如图11所示,通过对作为动态图像对 β 的一方的共同出场人物的人物b手动进行标签赋予,从而对作为另一方的共同出场人物的人物c自动进行标签赋予。

[0110] 接下来,考虑所选择的动态图像对的组合中的另一方的动态图像对。首先,删除另一方的动态图像对的共同出场人物中的与已完成了标签赋予的一方的动态图像对重复的共同出场人物。在上述的例子中,删除动态图像对 α 的共同出场人物即人物a、人物b、人物c中的完成了标签赋予的人物b与人物c。这里,若未被删除而残留的共同出场人物为1名,则由于标签候选与面部特征代表矢量的对应关系确定,因此能够自动进行标签赋予。在上述的例子中,由于动态图像对 α 的共同出场人物中的未被删除而残留的共同出场人物仅为人物a,因此如图11所示,对作为动态图像对 α 的共同出场人物的人物a自动进行标签赋予。另外,假设在动态图像对 α 的共同出场人物中的未被删除而残留的共同出场人物为2名以上的情况下,通过与对上述的动态图像对 β 的处理同样地进行示教对象的标签候选的选择、以及将所选择的标签候选与面部图像一览一起提示并受理用户的示教操作的处理,能够对未赋予标签的共同出场人物进行标签赋予。

[0111] 另外,对于不包含在所选择的动态图像对的组合中的未处理的动态图像对,也同样地删除标签赋予已完成的共同出场人物。在上述的例子中,删除作为动态图像对 γ 的共同出场人物的人物a、人物c、人物d中的标签赋予已完成的人物a与人物c。这里,若未被删除而残留的共同出场人物为1名,则标签候选与面部特征代表矢量的对应关系确定,因此能够自动进行标签赋予。在上述的例子中,由于动态图像对 γ 的共同出场人物中的未被删除而残留的共同出场人物仅为人物d,因此如图12所示,对作为动态图像对 γ 的共同出场人物的人物d自动进行标签赋予。另外,假设在动态图像对 γ 的共同出场人物中的未被删除而残留的共同出场人物为2名以上的情况下,能够与对上述的动态图像对 β 的处理同样地,通过进行示教对象的标签候选的选择、以及将所选择的标签候选与面部图像一览一起提示并受理用户的示教操作的处理,能够对未赋予标签的共同出场人物进行标签赋予。

[0112] 如以上说明那样,根据本实施例,在用户需要手动进行标签赋予的情况下,选择能

够通过消除法高效地进行标签赋予的标签候选作为示教对象,优先地进行对该标签候选的示教操作,因此能够减轻手动进行标签赋予时的用户的劳力,高效地进行标签赋予。

[0113] <第四实施例>

[0114] 接下来,对第四实施例进行说明。本实施例相对于上述的第一实施例,附加了如下功能:确认是否按照期待处理了对特征量的标签的自动赋予,在有可能没有按照期待进行处理的情况下,进行用户的手动确认以及修正。由于其他功能与上述的第一实施例相同,因此以下仅对本实施例特有的功能进行说明。

[0115] 上述的第一实施例基于主要出场人物在动态图像中出场较多、非主要出场人物在动态图像中几乎没有出场的假设进行了处理。因而,在该假设不成立的事例中,存在未进行期待的处理的隐患。即,未生成主要出场人物的面部特征代表矢量而生成非主要出场人物的面部特征代表矢量的情况。因此,在本实施例中,针对上述的矢量对计算错误得分,基于该错误得分检查构成该矢量对的面部特征代表矢量是否属于同一人物。关于检查的结果,在面部特征代表矢量不属于同一人物的情况下,生成了非主要出场人物的面部特征代表矢量的可能性高。因此,进行用户的手动检查以及修正。由此,在由于生成了非主要出场人物的面部特征代表矢量的而对面部特征代表矢量赋予了错误的标签的情况下,能够修正该错误。

[0116] 图13是表示第四实施例的标签赋予装置的功能性构成例的框图。本实施例的标签赋予装置10是相对于图2所示的第一实施例的构成追加了特征量对判定部20和标签确认部21的构成。另外,在本实施例中,在数据保持部16内,构建有未检查特征量保持部16a、检查完毕特征量保持部16b、需要手动检查特征量对保持部16c、需要手动检查特征量保持部16d以及需要手动指定特征量保持部16e。

[0117] 未检查特征量保持部16a保持无法确认正确进行了标签赋予的面部特征代表矢量、以及与其相关联的信息(标签、面部图像、矢量对构成信息中的某一个)。另外,矢量对构成信息是与该面部特征代表矢量一起表示构成矢量对的其他面部特征代表矢量的信息。

[0118] 检查完毕特征量保持部16b保持能够确认正确进行了标签赋予的面部特征代表矢量、以及与其相关联的信息(标签、面部图像)。

[0119] 需要手动检查特征量对保持部16c保持需要由用户进行的手动检查的构成矢量对的面部特征代表矢量、以及与其相关联的信息(标签、面部图像、矢量对构成信息)。

[0120] 需要手动检查特征量保持部16d保持需要由用户进行的手动检查的面部特征代表矢量、以及与其相关联的信息(标签、面部图像)。

[0121] 需要手动指定特征量保持部16e保持需要用户手动指定标签的面部特征代表矢量、以及与其相关联的信息(标签、面部图像)。

[0122] 特征量对判定部20将从未检查特征量保持部16a取出的面部特征代表矢量中的能够与其他面部特征代表矢量构成矢量对的面部特征代表矢量、即具有矢量对构成信息的面部特征代表矢量作为对象,计算矢量对的错误得分。然后,特征量对判定部20基于计算出的错误得分,判定矢量对是否属于同一人物。例如,若矢量对的错误得分为规定的阈值以下,则判定为矢量对属于同一人物,在矢量对的错误得分超过阈值的情况下,判定为矢量对不属于同一人物。若构成矢量对的两个面部特征代表矢量属于同一人物,则矢量对的错误得分低,若属于不同的人物,则矢量对的错误得分变高,例如能够使用两个面部特征代表矢量

间的距离(欧几里德距离)、或者从1中减去两个面部特征代表矢量的内积的值而得的值等。

[0123] 在矢量对属于同一人物的情况下,适当地生成了面部特征代表矢量、即成为主要出场人物的面部特征代表矢量的可能性非常高。这是因为,在以主要出场人物重复的方式选择了动态图像组合时,认为在各个动态图像中非主要出场人物比主要出场人物出场多,并且相同的非主要出场人物在多个动态图像中重复出场的可能性极低。

[0124] 标签确认部21向用户确认是否对构成错误得分超过阈值的矢量对的各面部特征代表矢量、或者无法构成矢量对的面部特征代表矢量赋予了正确的标签。例如,标签确认部21向用户提示与确认对象的面部特征代表矢量相关联的面部图像与标签(人物名),并取得两者是否一致的确认结果。

[0125] 以下,参照图14至图17,对本实施例的处理的流程进行说明。图14至图17是表示本实施例的处理步骤的一个例子的流程图。本实施例的处理包括以下的处理(1)与处理(2)这两个处理。图14示出了处理(1)的步骤,图15至图17示出了处理(2)的步骤。

[0126] 处理(1):自动确认是否以高概率适当地生成了面部特征代表矢量,将需要用户的手动检查的面部特征代表矢量储存于需要手动检查特征量对保持部16c或者需要手动检查特征量保持部16d。

[0127] 处理(2):通过用户的手动检查确认是否对储存于需要手动检查特征量对保持部16c与需要手动检查特征量保持部16d的面部特征代表矢量适当地赋予了标签,在有需要时,受理由用户进行的手动的修正。

[0128] 首先,使用图14的流程图对处理(1)进行说明。当开始处理(1)时,特征量对判定部20从未检查特征量保持部16a取出一个面部特征代表矢量(步骤S101)。然后,特征量对判定部20判定在步骤S101中取出的面部特征代表矢量是否是能够构成矢量对的面部特征代表矢量(步骤S102)。关于面部特征代表矢量是否能够构成矢量对,例如能够通过该面部特征代表矢量是否具有矢量对构成信息来判定。

[0129] 这里,在判定为步骤S101中取出的面部特征代表矢量无法构成矢量对的情况下(步骤S102:否),特征量对判定部20将该面部特征代表矢量储存于需要手动检查特征量保持部16d(步骤S103)。另一方面,在判定为步骤S101中取出的面部特征代表矢量能够构成矢量对的情况下(步骤S102:是),特征量对判定部20计算该面部特征代表矢量构成的矢量对的错误得分(步骤S104),判定计算出的错误得分是否为阈值以下(步骤S105)。

[0130] 在步骤S105的判定中使用的阈值例如被设定为所有矢量对的平均错误得分的2倍。这是因为,在图像数据充分存在时,矢量对的大半为同一人物的矢量,平均值周边被认为是正确地建立了对应的矢量对。另外,这里为了方便设为2倍,但也可以通过调整(tuning)来决定阈值。另外,也可以不使用平均值而使用中央值等。

[0131] 这里,若矢量对的错误得分为阈值以下(步骤S105:是),则视为该矢量对由同一人物的面部特征代表矢量构成。在该情况下,认为构成矢量对的各面部特征代表矢量被正确赋予了标签,因此特征量对判定部20将构成该矢量对的各面部特征代表矢量储存于检查完毕特征量保持部16b(步骤S106)。

[0132] 另一方面,在矢量对的错误得分超过了阈值的情况下(步骤S105:否),特征量对判定部20判定在步骤S101中取出的面部特征代表矢量是否能够构成另一矢量对(步骤S107),在能够构成另一矢量对的情况下(步骤S107:是),返回到步骤S104,反复进行该矢量对的错

误得分计算与阈值判定。即,在步骤S101中取出的面部特征代表矢量能够通过变更动态图像组合来构成另一矢量对的情况下,确认该矢量对的错误得分是否为阈值以下。然后,在即使尝试了所有动态图像组合也未找到错误得分为阈值以下的矢量对的情况下(步骤S107:否),将错误得分最小的矢量对储存于需要手动检查特征量对保持部16c(步骤S108)。

[0133] 之后,特征量对判定部20判定在未检查特征量保持部16a中是否残留有未处理的面部特征代表矢量(步骤S109)。然后,若在未检查特征量保持部16a中残留有未处理的面部特征代表矢量(步骤S109:是),则返回到步骤S101,反复进行相同的处理。然后,当对保持于未检查特征量保持部16a的所有面部特征代表矢量的处理结束时(步骤S109:否),处理(1)结束。

[0134] 接下来,对处理(2)进行说明。处理(2)包括:对储存于需要手动检查特征量对保持部16c的矢量对的处理、对储存于需要手动检查特征量保持部16d的面部特征代表矢量对的处理、以及对储存于需要手动指定特征量保持部16e的面部特征代表矢量的处理。

[0135] 首先,使用图15的流程图,说明对储存于需要手动检查特征量对保持部16c的矢量对的处理。当开始该处理时,标签确认部21从需要手动检查特征量对保持部16c取出一个矢量对(步骤S201)。然后,标签确认部21向用户提示分别与构成矢量对的两个面部特征代表矢量相关联的面部图像和共同的标签,确认是否对各个面部特征代表矢量赋予了正确的标签(步骤S202)。例如,用户检查所提示的面部图像是否属于同一人物,若属于同一人物,则检查共同的标签是否是该人物的姓名。另外,若不属于同一人物,则检查共同的标签是哪个人物的姓名。

[0136] 这里,在判定为对构成矢量对的各面部特征代表矢量赋予了正确的标签的情况下(步骤S202:是),标签确认部21将构成矢量对的各面部特征代表矢量储存于检查完毕特征量保持部16b(步骤S203)。另一方面,在判定为未对构成矢量对的面部特征代表矢量的至少一方赋予正确地标签的情况下(步骤S202:否),标签确认部21将未被赋予正确的标签的面部特征代表矢量储存于需要手动检查特征量保持部16d(步骤S204)。

[0137] 之后,标签确认部21判定在需要手动检查特征量对保持部16c中是否残留有未处理的矢量对(步骤S205)。然后,若在需要手动检查特征量对保持部16c中残留有未处理的矢量对(步骤S205:是),则返回到步骤S201,反复进行相同的处理。然后,当对保持于需要手动检查特征量对保持部16c的所有矢量对的处理结束时(步骤S205:否),结束一系列的处理。

[0138] 接下来,使用图16的流程图,说明对储存于需要手动检查特征量保持部16d的面部特征代表矢量的处理。当开始该处理时,标签确认部21从需要手动检查特征量保持部16d取出一个面部特征代表矢量(步骤S301)。然后,标签确认部21向用户提示与面部特征代表矢量相关联的面部图像和标签,确认是否对该面部特征代表矢量赋予了正确的标签(步骤S302)。例如,用户检查所提示的标签是否是面部图像的人物的姓名。

[0139] 这里,在判定为对面部特征代表矢量赋予了正确的标签的情况下(步骤S302:是),标签确认部21将该面部特征代表矢量储存于检查完毕特征量保持部16b(步骤S305)。另一方面,在判定为未对面部特征代表矢量正确赋予标签的情况下(步骤S302:否),标签确认部21向用户提示与生成该面部特征代表矢量的动态图像对应的主要出场人物的一览,确认是否生成了与该面部特征代表矢量对应的标签候选(步骤S303)。例如,用户检查所提示的面部图像的人物名是否包含在主要出场人物一览中。然后,若面部图像的人物名包含在主要

出场人物一览中,则用户进行选择该人物名的操作。

[0140] 这里,在生成了与面部特征代表矢量对应的标签候选的情况下,即在所提示的面部图像的人物名包含在主要出场人物一览中,且用户进行了从主要出场人物一览中选择面部图像的人物名的操作的情况下(步骤S303:是),标签确认部21根据该用户操作变更面部特征代表矢量的标签(步骤S304),将该面部特征代表矢量储存于检查完毕特征量保持部16b(步骤S305)。另一方面,在未生成与面部特征代表矢量对应的标签候选的情况下(步骤S303:否),标签确认部21将该面部特征代表矢量储存于需要手动指定特征量保持部16e(步骤S306)。

[0141] 之后,标签确认部21判定在需要手动检查特征量保持部16d中是否残留有未处理的面部特征代表矢量(步骤S307)。然后,若在需要手动检查特征量保持部16d中残留有未处理的面部特征代表矢量(步骤S307:是),则返回到步骤S301,反复进行相同的处理。然后,当对保持于需要手动检查特征量保持部16d的所有面部特征代表矢量的处理结束时(步骤S307:否),结束一系列的处理。

[0142] 接下来,使用图17的流程图,说明对储存于需要手动指定特征量保持部16e的面部特征代表矢量的处理。当开始该处理时,标签确认部21从需要手动指定特征量保持部16e取出一个面部特征代表矢量(步骤S401)。然后,标签确认部21向用户提示与面部特征代表矢量相关联的标签、以及与在生成该面部特征代表矢量时作为“偏离值矢量”从聚类的对象中排除的面部特征矢量对应的面部图像的一览,根据用户操作变更与标签对应的面部特征矢量(步骤S402)。

[0143] 即,标签确认部21将生成处理对象的面部特征代表矢量时作为“偏离值矢量”的面部特征矢量作为修正候选,将与修正候选的面部特征矢量对应的面部图像的一览和对处理对象的面部特征代表矢量赋予的标签一起提示给用户。此时,面部图像一览也可以以面部图像的大小、面部朝向正面的程度等、在帧图像上越是显眼的人物的面部图像越靠上位的方式排序。用户进行从面部图像一览中选择与所提示的标签的人物名一致的人物的面部图像的操作。标签确认部21根据该用户操作,变更与向用户提示的标签对应的面部特征矢量。然后,标签确认部21将变更后的面部特征矢量储存于检查完毕特征量保持部16b(步骤S403)。

[0144] 之后,标签确认部21判定在需要手动指定特征量保持部16e中是否残留有未处理的面部特征代表矢量(步骤S404)。然后,若在需要手动指定特征量保持部16e中残留有未处理的面部特征代表矢量(步骤S404:是),则返回到步骤S401,反复进行相同的处理。然后,当对保持于需要手动指定特征量保持部16e的所有面部特征代表矢量的处理结束时(步骤S404:否),结束一系列的处理。

[0145] 如以上说明那样,根据本实施例,确认是否按照期待进行了针对特征量的标签的自动赋予的处理,在有可能未按照期待进行处理的情况下,进行用户的手动确认以及修正,因此能够更高精度地对特征量进行标签赋予。

[0146] <第五实施例>

[0147] 接下来,对第五实施例进行说明。本实施例相对于上述的第四实施例附加了用于高效地进行用户的手动确认以及修正的功能。由于其他功能与上述的第四实施例相同,因此以下仅对本实施例特有的功能进行说明。

[0148] 在上述的实施例4中,在计算矢量对的错误得分的结果是判断为面部特征代表矢量不是主要出场人物的面部特征代表矢量的可能性较高的情况下,进行用户的手动确认以及修正。但是,存在通过在内容DB50中追加新的内容51(动态图像)而能够构建错误得分为阈值以下的新的矢量对的情况,在该情况下,不需要追加新的内容51之前所需的用户的手动确认以及修正。

[0149] 因而,在追加了新的内容51时,再次进行相同的处理是有效的,但若以内容DB50内的所有内容51为对象再次进行相同的处理,则处理时间变长。因此,在本实施例中,通过以从新的内容51生成的特征量和需要用户的手动确认以及修正的特征量为对象再次进行实施例4的处理,由此以较短的处理时间筛选需要用户的手动确认以及修正的特征量。

[0150] 图18是表示第五实施例的标签赋予装置10的功能性构成例的框图。本实施例的标签赋予装置10是相对于图13所示的第四实施例的构成追加了确认对象筛选部22的构成。

[0151] 确认对象筛选部22在对内容DB50追加了新的内容51的情况下,使用从新的内容51提取的特征量,筛选构成错误得分超过阈值的特征量对各特征量、或者无法构成特征量对的特征量中的需要确认是否被正确赋予了标签的特征量。

[0152] 以下,对本实施例中的处理的概要进行说明。在本实施例中,首先,特征量对判定部20执行上述的第四实施例的处理(1),将需要用户的手动确认以及修正的面部特征代表矢量储存于需要手动检查特征量对保持部16c、需要手动检查特征量保持部16d。另外,对作为新的内容51而追加到内容DB50的动态图像执行上述的第一实施例的处理,生成新的面部特征代表矢量。

[0153] 接下来,确认对象筛选部22将从作为新的内容51追加的动态图像生成的新的面部特征代表矢量、保持于需要手动检查特征量对保持部16c的面部特征代表矢量、保持于需要手动检查特征量保持部16d的面部特征代表矢量全部储存于未检查特征量保持部16a。之后,特征量对判定部20通过再次执行上述的第四实施例的处理(1),将储存于需要手动检查特征量对保持部16c、需要手动检查特征量保持部16d的面部特征代表矢量、即需要用户的手动确认以及修正的面部特征代表矢量筛选。

[0154] 如以上说明那样,根据本实施例,在追加了新的内容51的情况下,使用从新的内容提取的特征量来筛选需要用户的手动确认以及修正的特征量,因此能够高效地进行用户的手动确认以及修正。

[0155] <补充说明>

[0156] 作为一个例子,上述的各实施例的标签赋予装置10能够通过在使用了作为一般的计算机的硬件的执行环境下动作的程序进行安装。在该情况下,标签赋予装置10中的上述的各功能性构成要素(标签候选生成部11、特征量提取部12、特征量对检测部13、标签赋予部14、元数据生成部15、数据保持部16、标签候选修正部17、示教对象选择部18、操作受理部19、特征量对判定部20、标签确认部21、确认对象筛选部22)通过硬件与软件(程序)的协作而实现。

[0157] 图19是表示标签赋予装置10的硬件构成例的框图。例如,如图19所示,标签赋予装置10能够采用具备CPU(Central Processing Unit)101等处理器电路、ROM(Read Only Memory)102、RAM(Random Access Memory)103等存储装置、供显示面板、各种操作设备连接的输入输出I/F104、与网络连接而进行通信的通信I/F105、连接各部的总线106等的、使用

了一般计算机的硬件构成。

[0158] 另外,在上述构成的硬件上执行的程序例如以能够安装的形式或者能够执行的形式记录于CD-ROM(Compact Disk Read Only Memory,光盘只读存储器)、软盘(FD)、CD-R(Compact Disk Recordable,可录光盘)、DVD(Digital Versatile Disc,数字多功能光盘)等能够由计算机读取的记录介质而作为计算机程序产品来提供。另外,也可以构成,将在上述构成的硬件上执行的程序储存于与因特网等网络连接的计算机上,通过经由网络下载来提供。另外,也可以构成,经由因特网等网络提供或者发布在上述构成的硬件上执行的程序。另外,也可以构成,将在上述构成的硬件上执行的程序预先组装于ROM102等中来提供。

[0159] 在上述构成的硬件上执行的程序成为包含标签赋予装置10的各功能性构成要素的模块构成,例如通过CPU101(处理器电路)从上述记录介质读出程序并执行,上述各部被加载到RAM103(主存储)上,并被生成在RAM103(主存储)上。另外,标签赋予装置10的各功能性构成要素也可以是跨多个计算机而实现的构成。另外,也能够使用ASIC(Application Specific Integrated Circuit,应用型专用集成电路)、FPGA(Field-Programmable Gate Array,可编程逻辑门阵列)等专用的硬件来实现上述的功能性构成要素的一部分或者全部。

[0160] 根据以上所述的至少一个实施方式,能够自动进行对从内容提取的特征量的标签赋予。

[0161] 以上,对本发明的实施方式进行了说明,但这里说明的实施方式是作为例子而提出的,并不意图限定发明的范围。这里说明的新的实施方式能够以其他方式实施,在不脱离发明的主旨范围内,能够进行各种省略、替换、变更。这里说明的实施方式及其变形包含在发明的范围或主旨中,并且包含在权利要求书所记载的发明及其等效的范围中。

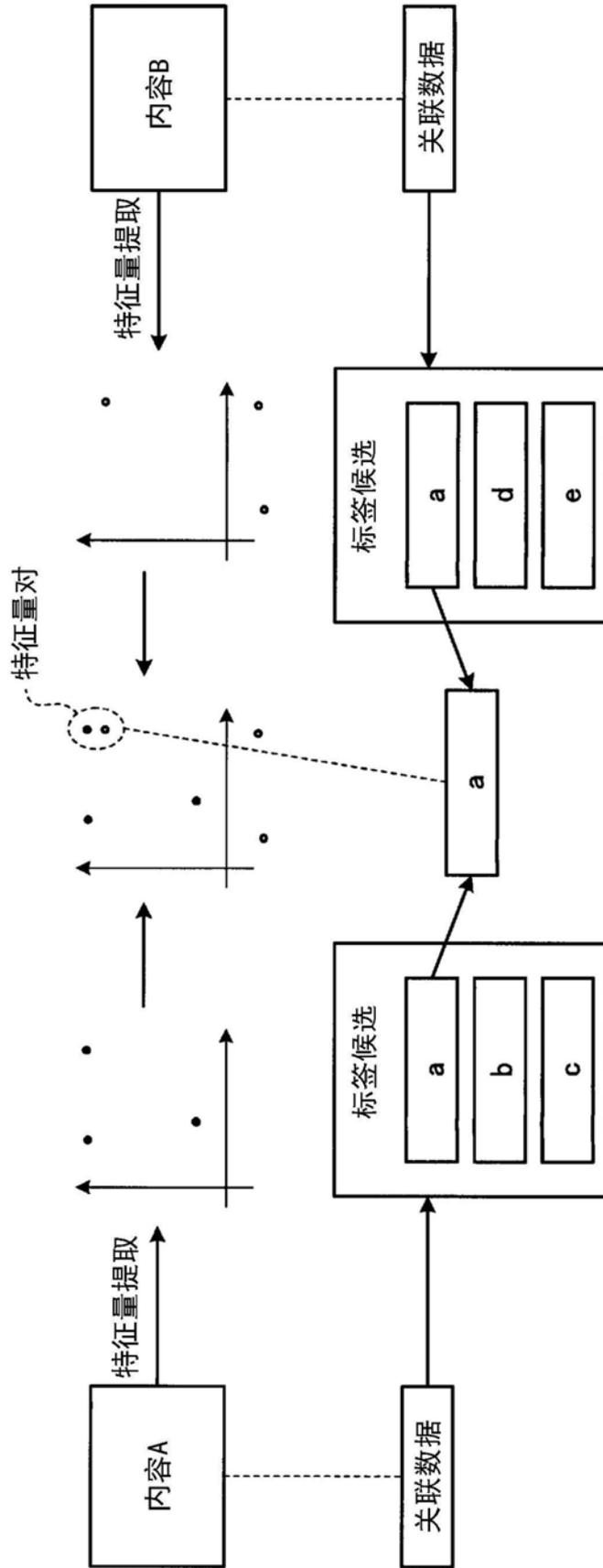


图1

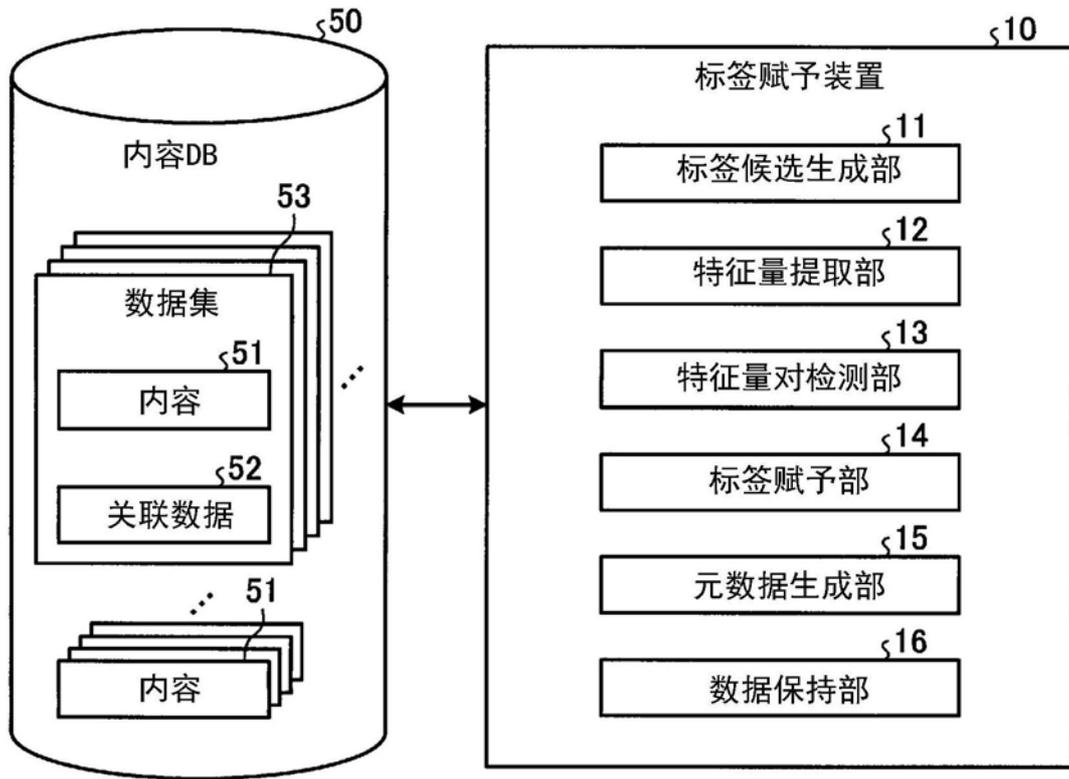


图2

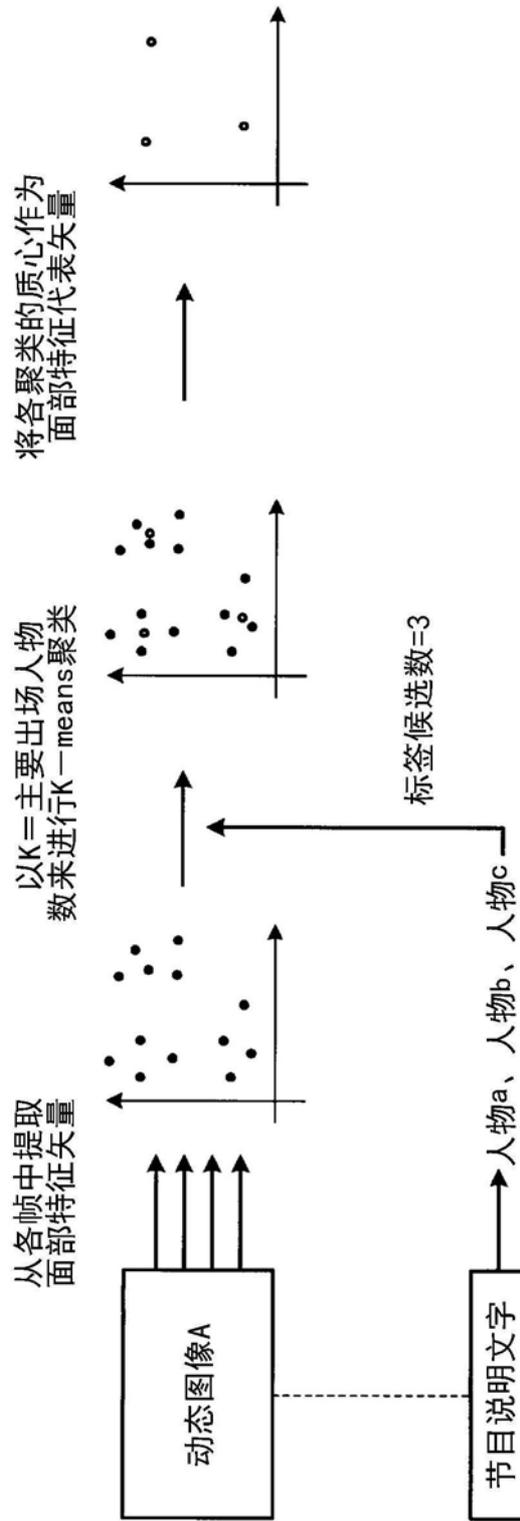


图3

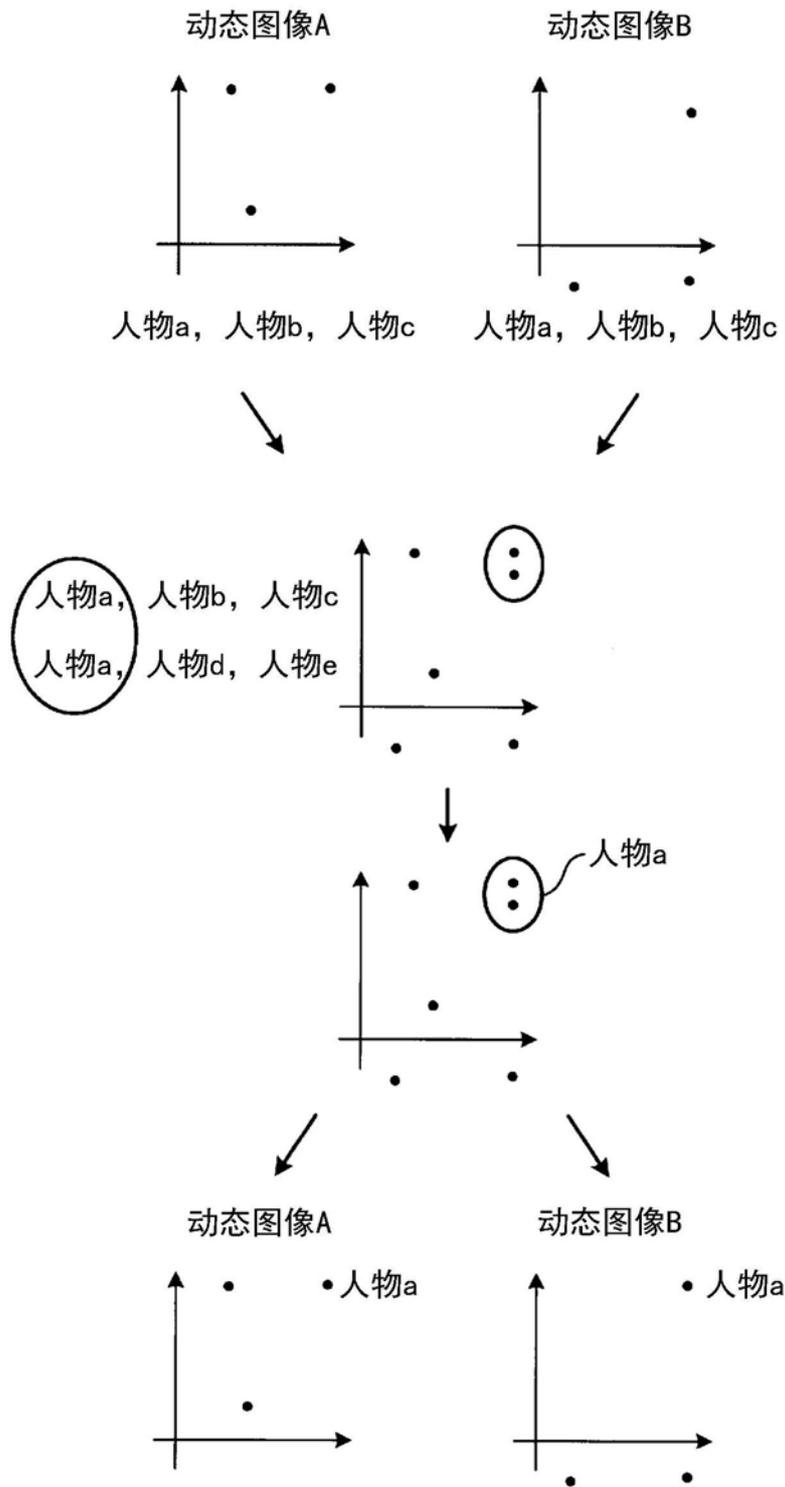


图4

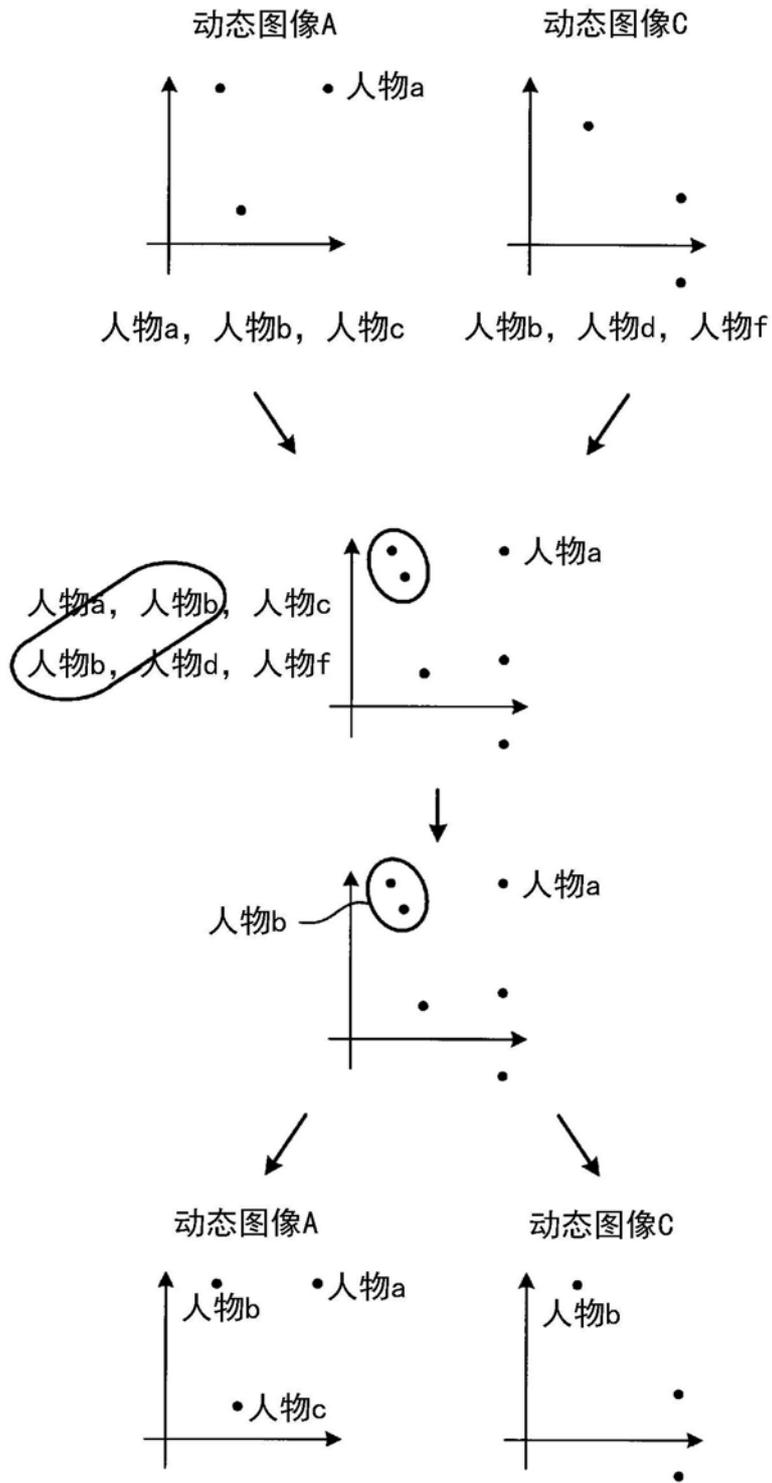


图5

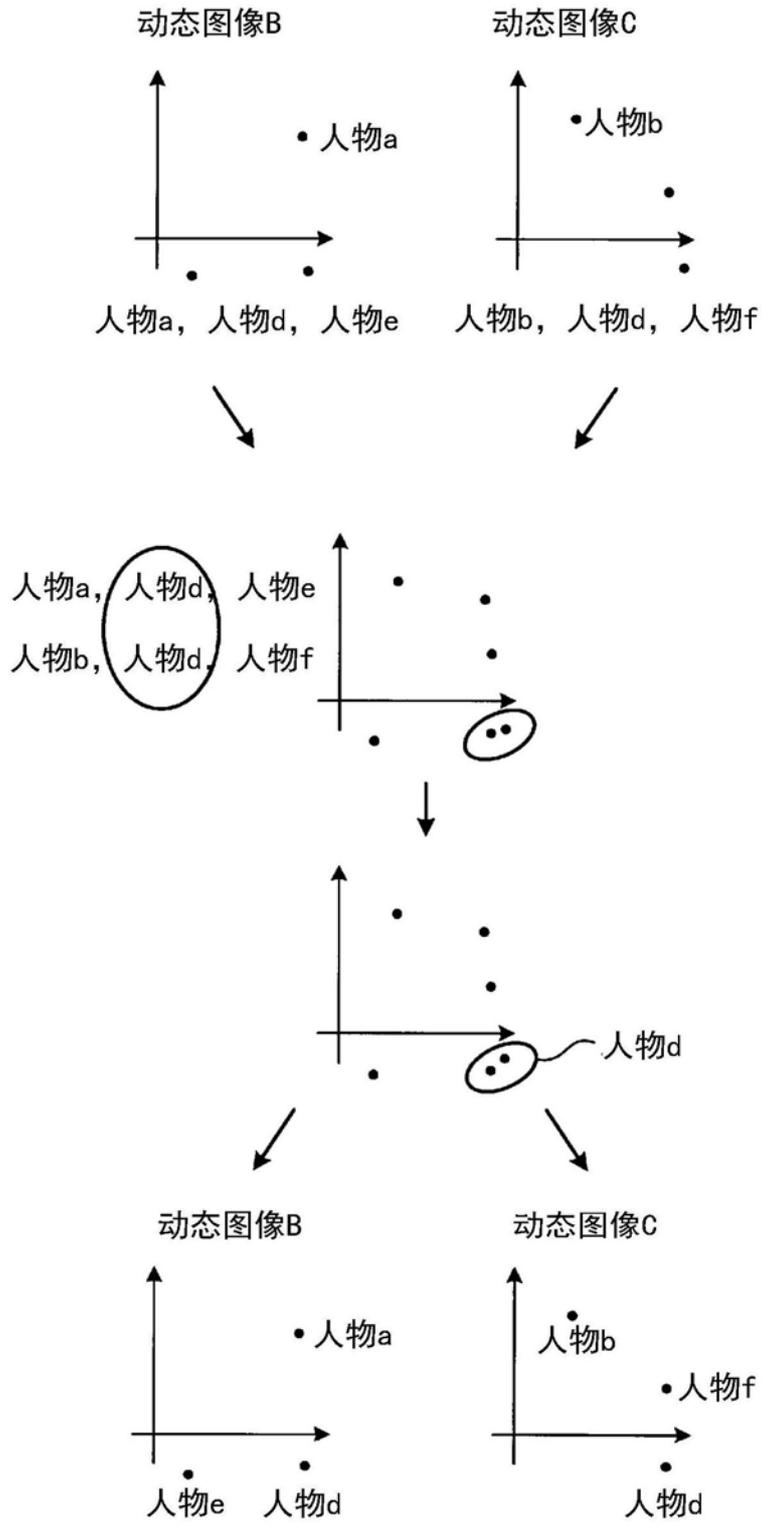


图6

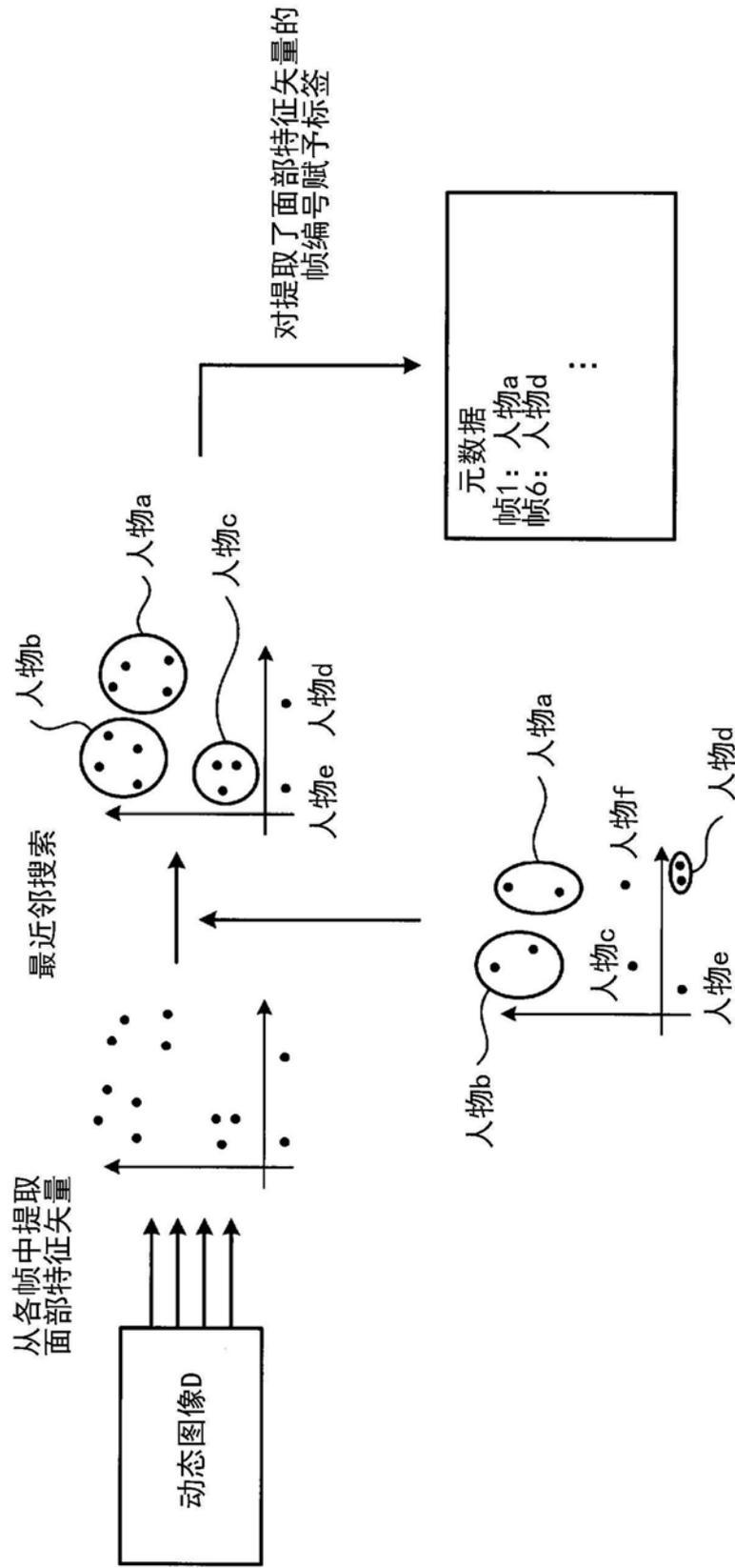


图7

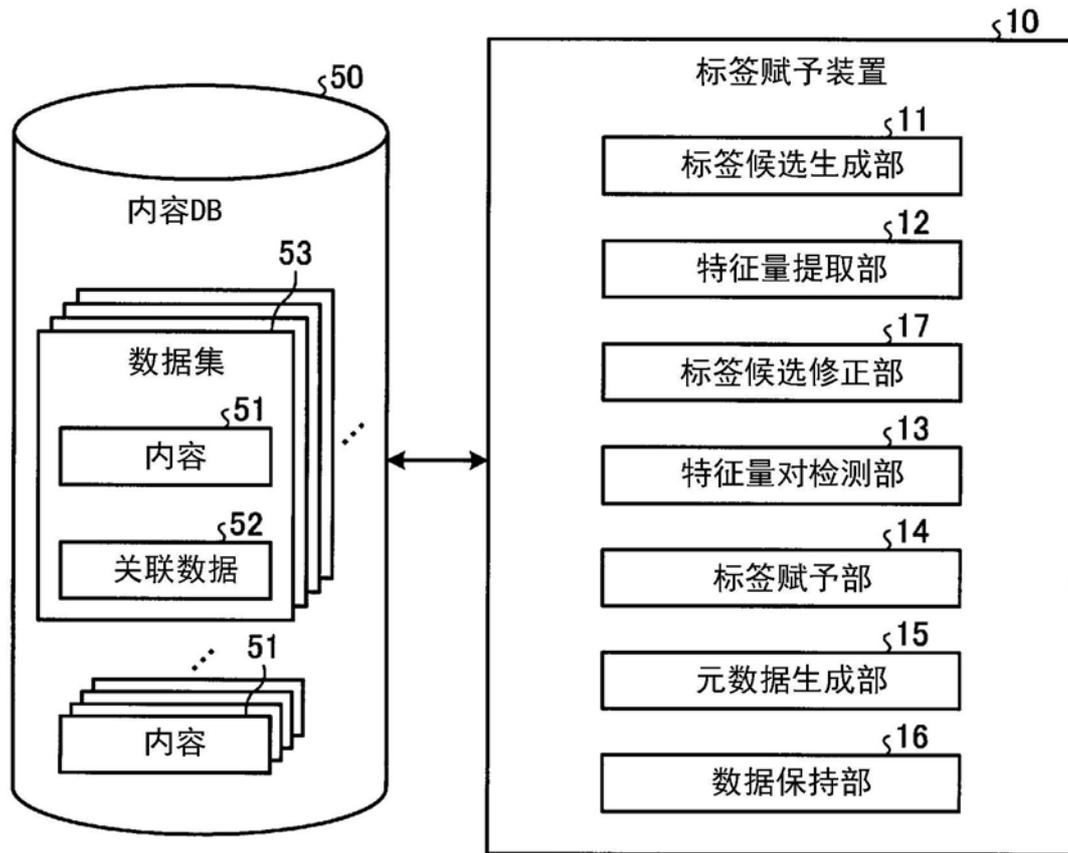


图8

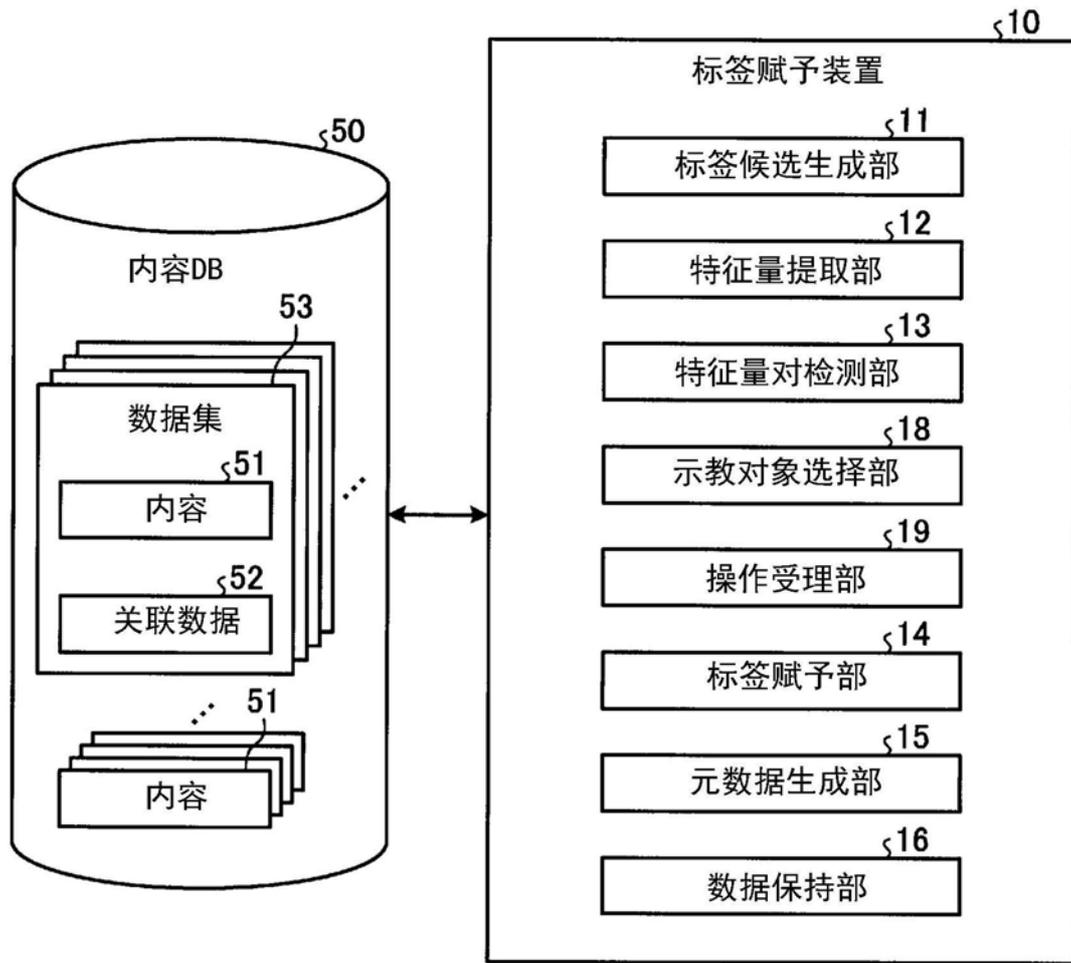


图9

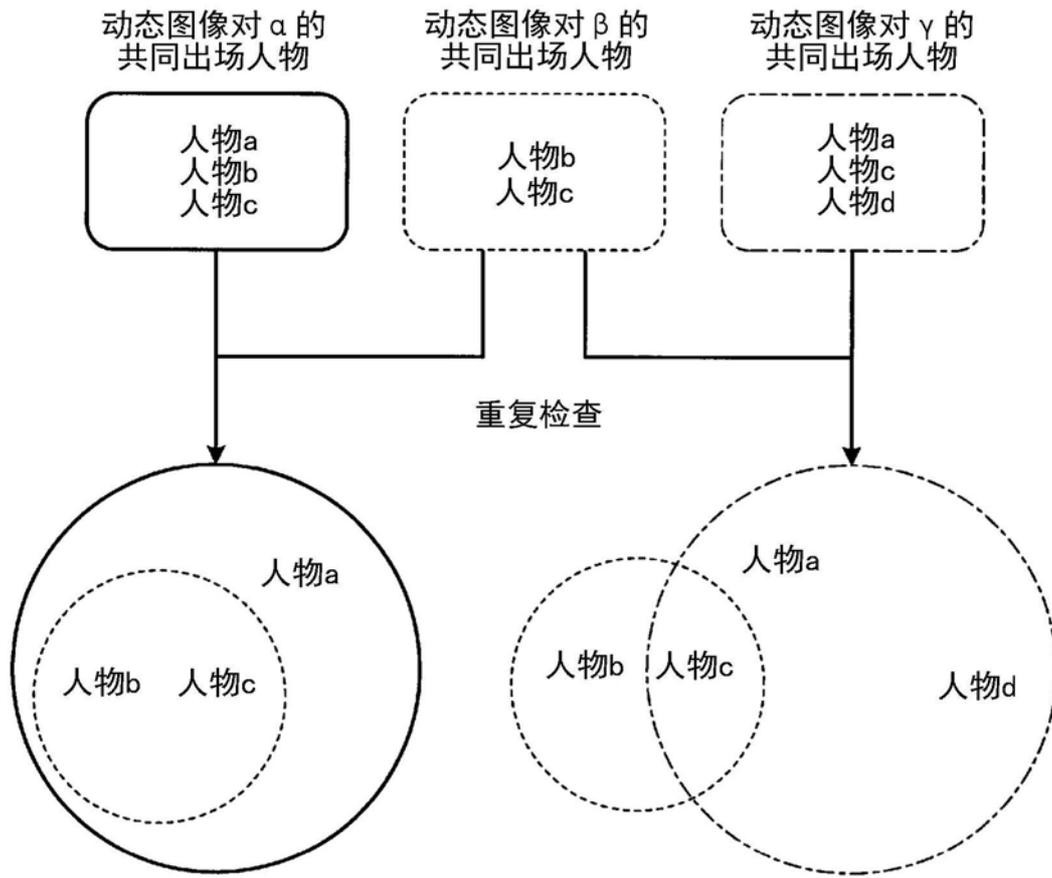


图10

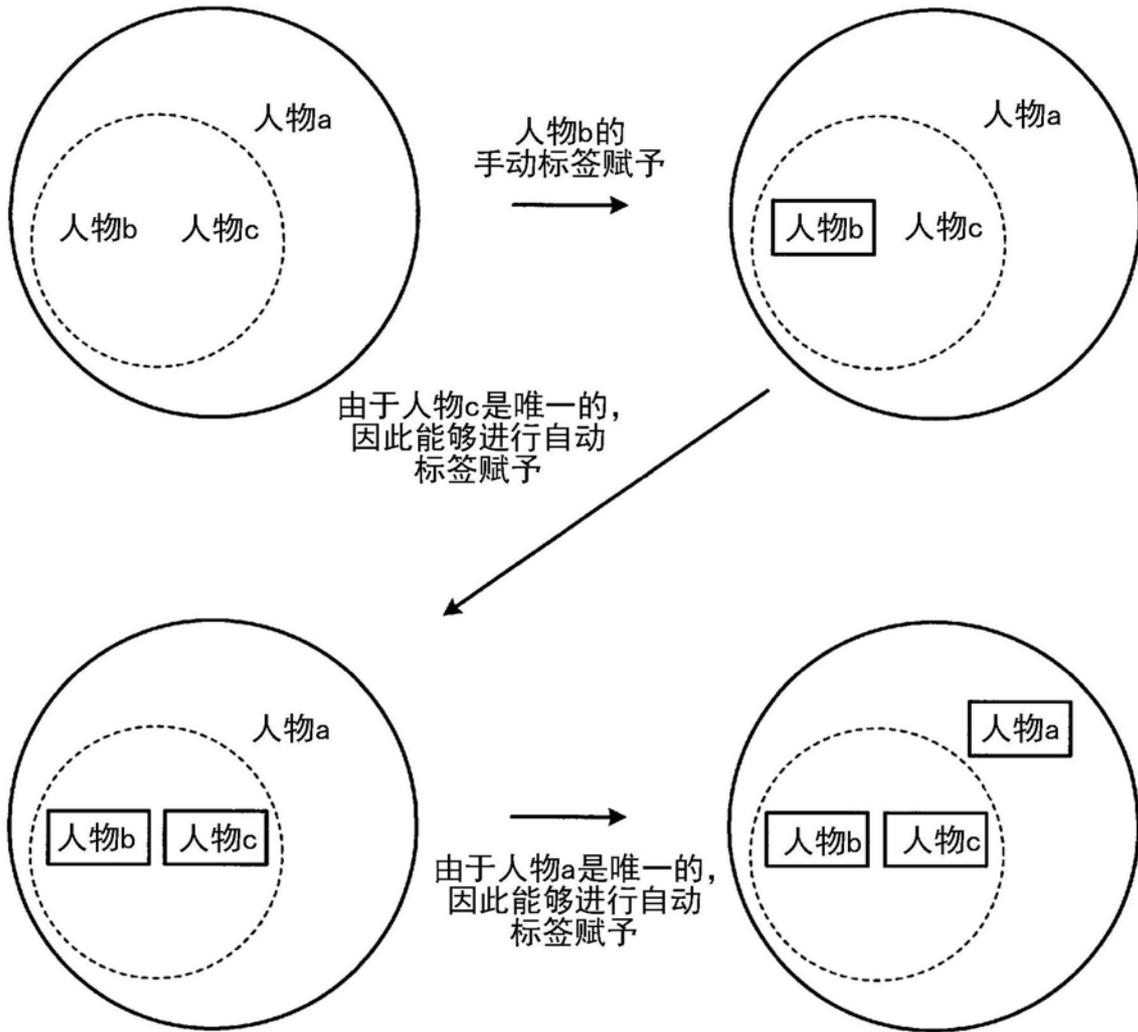


图11

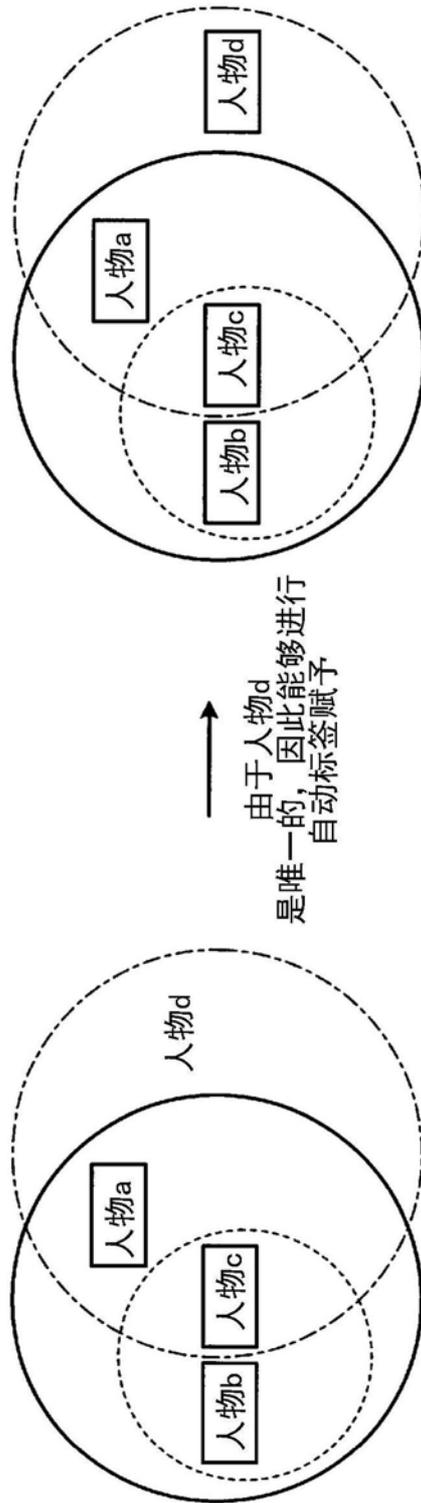


图12

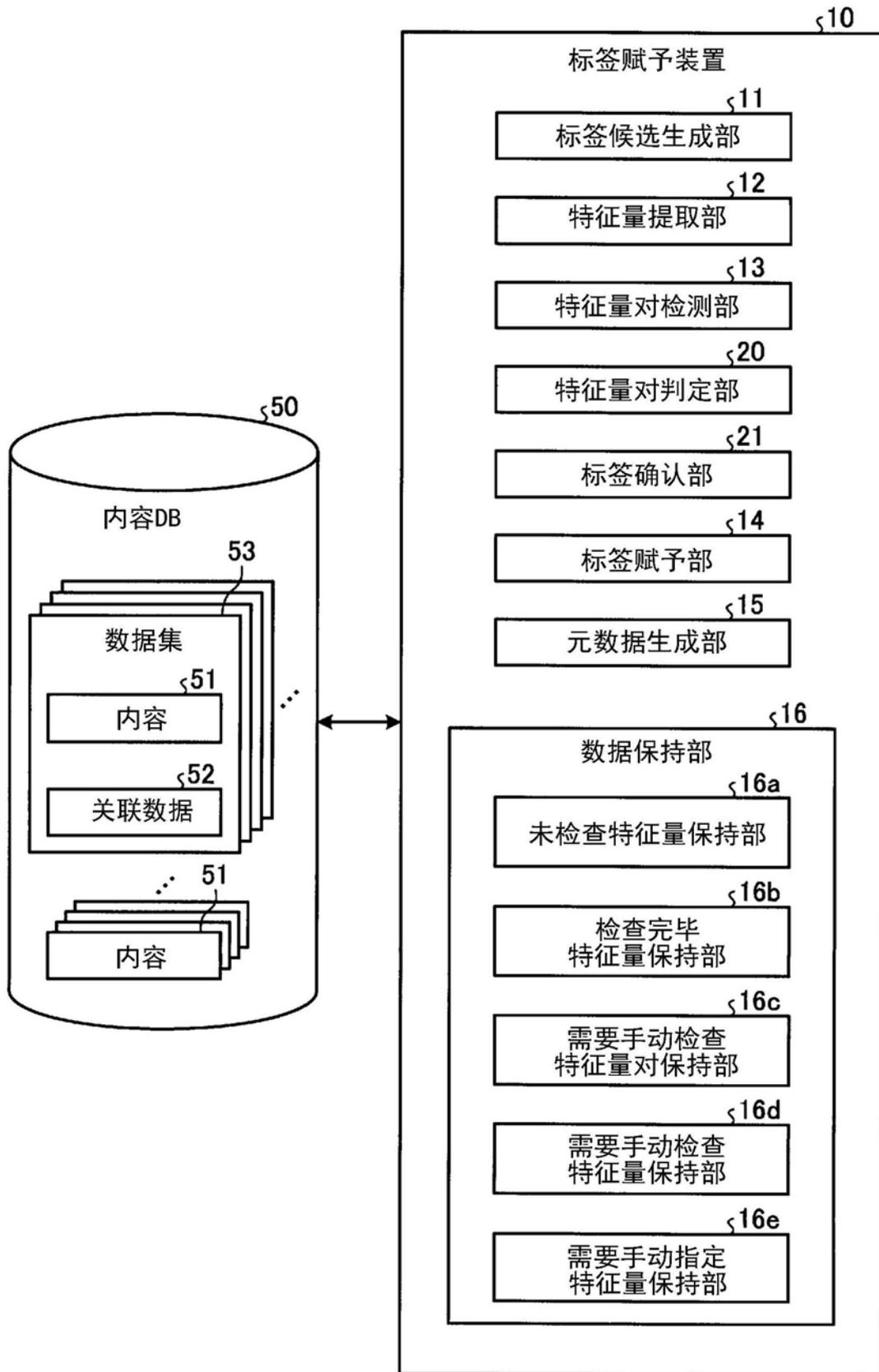


图13

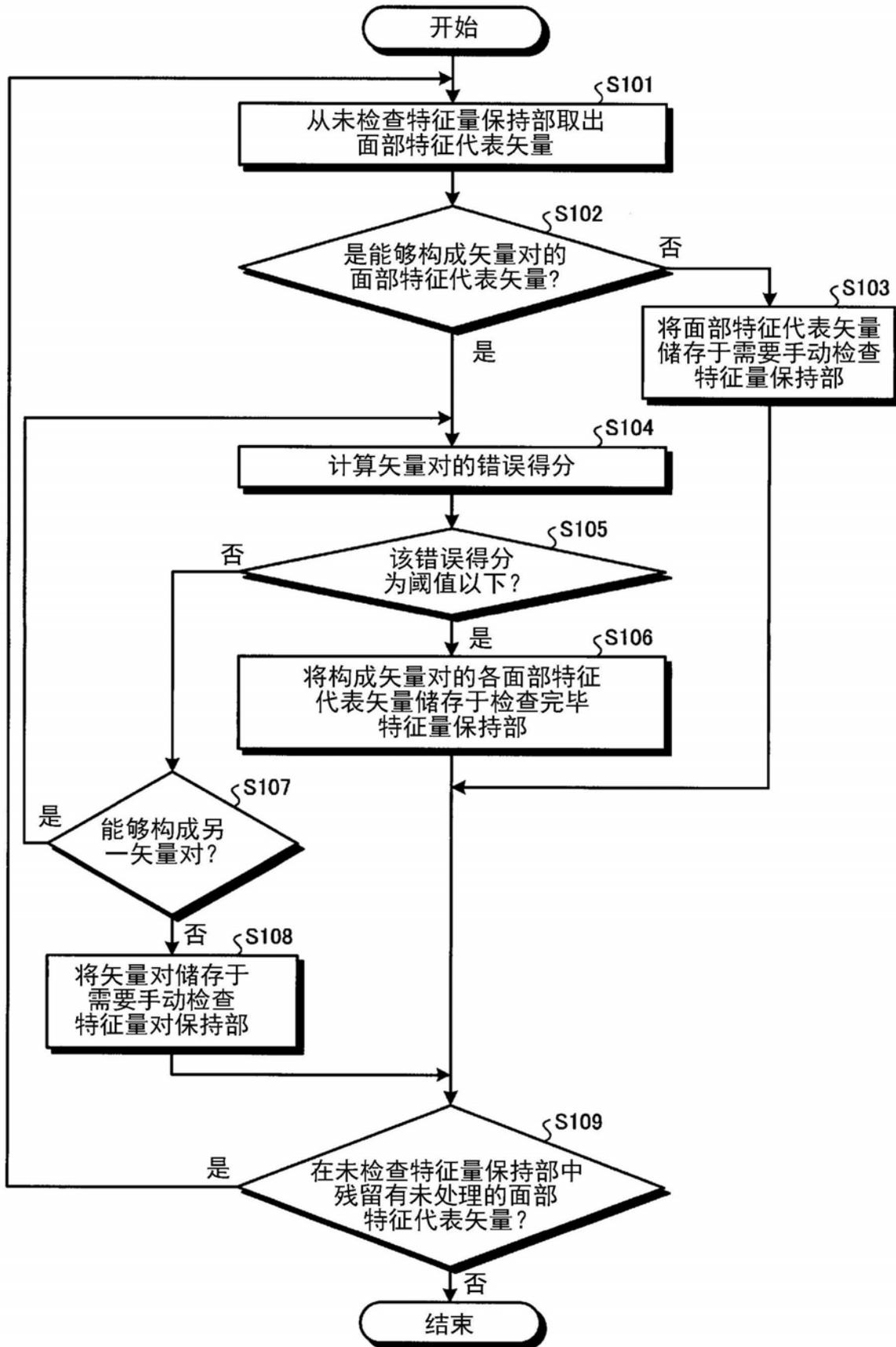


图14

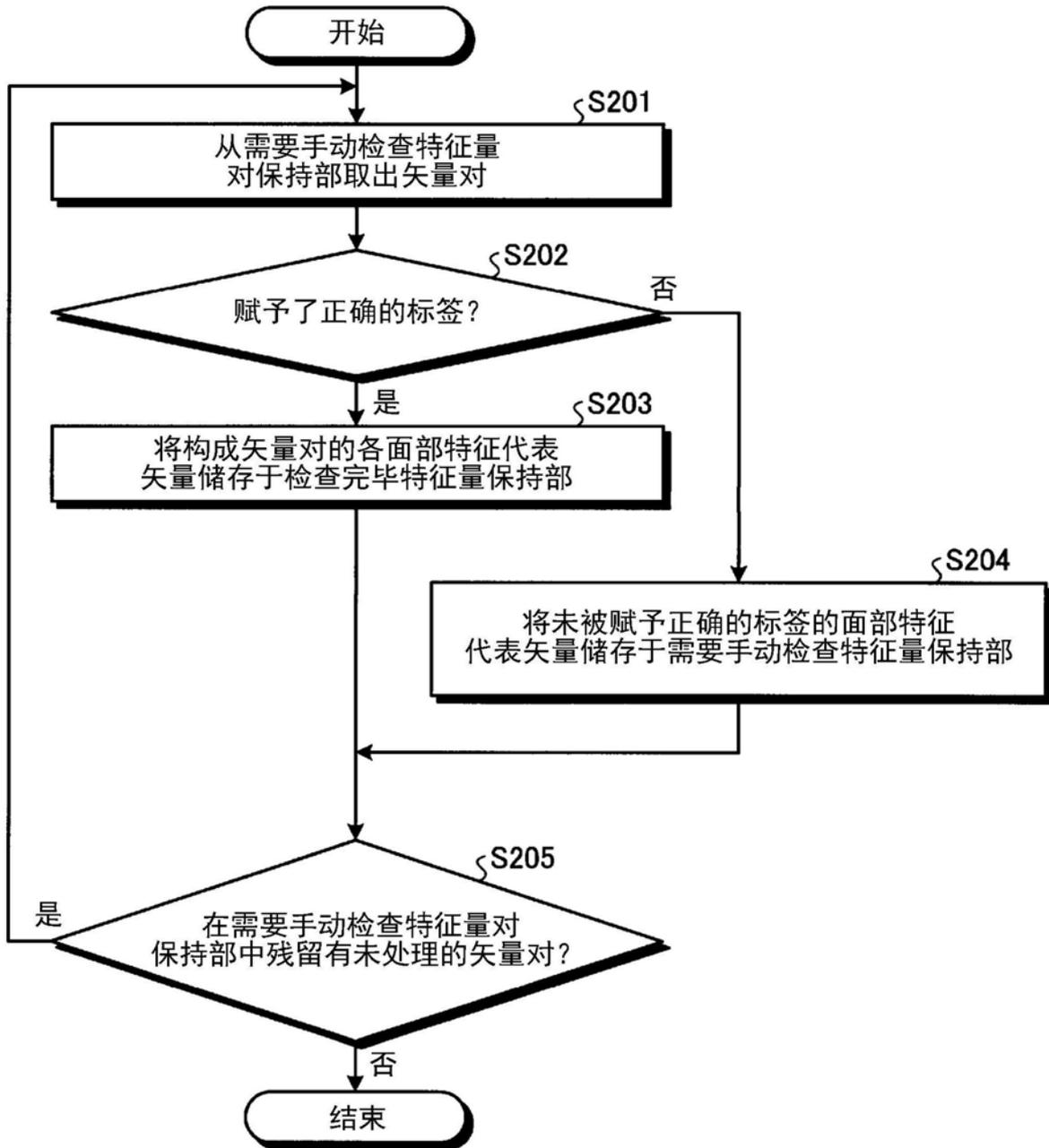


图15

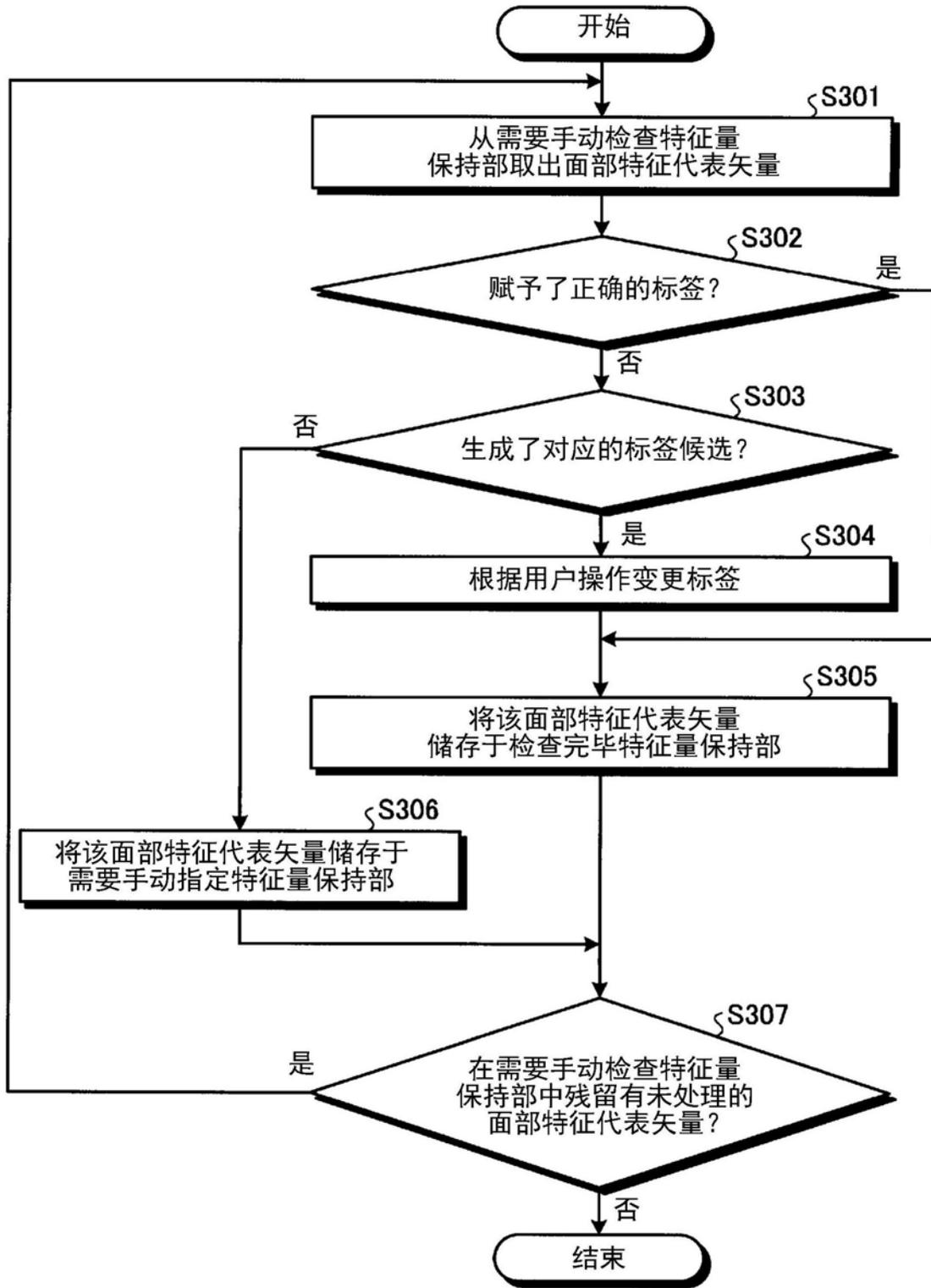


图16

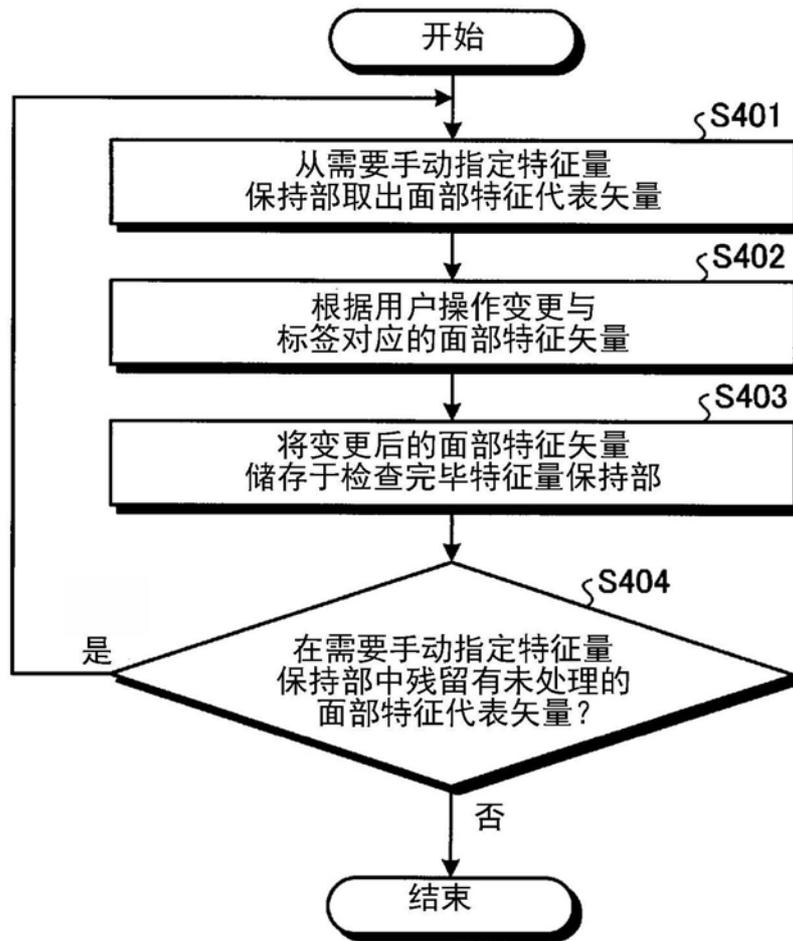


图17

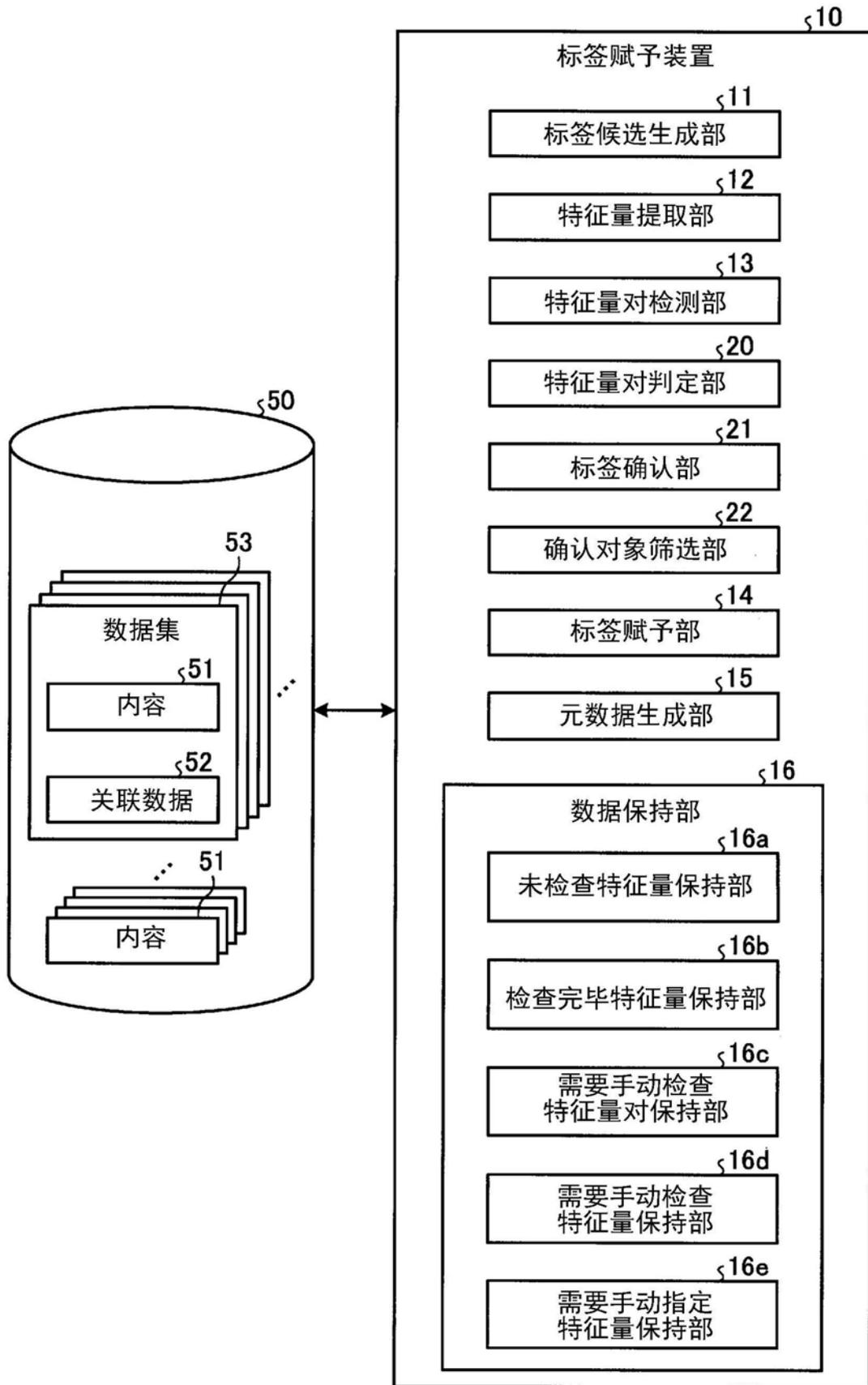


图18

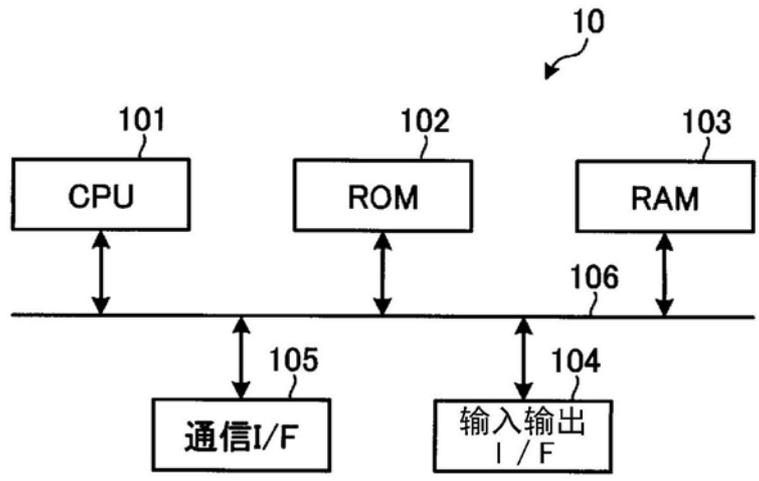


图19