



(12) 发明专利

(10) 授权公告号 CN 102591864 B

(45) 授权公告日 2015. 03. 25

(21) 申请号 201110001856. 7

CN 101350030 A, 2009. 01. 21, 全文.

(22) 申请日 2011. 01. 06

US 2008201140 A1, 2008. 08. 21, 全文.

(73) 专利权人 上海银晨智能识别科技有限公司
地址 201203 上海市浦东新区张江高科技园
区郭守敬路 498 号浦东软件园 9 号楼 4
楼

审查员 王骞

(72) 发明人 苗春晖

(74) 专利代理机构 北京集佳知识产权代理有限
公司 11227

代理人 骆苏华

(51) Int. Cl.

G06F 17/30(2006. 01)

(56) 对比文件

CN 101110088 A, 2008. 01. 23, 说明书第 1-7
页.

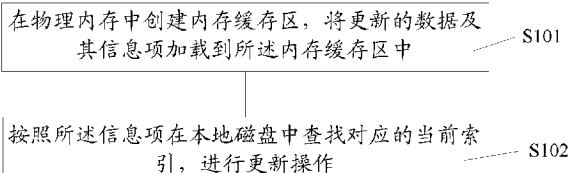
权利要求书3页 说明书8页 附图1页

(54) 发明名称

比对系统中的数据更新方法及装置

(57) 摘要

一种比对系统中的数据更新方法,包括:在物理内存中创建内存缓存区,将更新的数据及其信息项加载到所述内存缓存区中;按照所述信息项在本地磁盘中查找对应的当前索引,进行更新操作。为实现所述比对系统中的数据更新方法,本发明还提供了一种比对系统中的数据更新装置,包括:创建单元,用以在物理内存中创建内存缓存区;加载单元,用以将更新的数据及其信息项加载到所述内存缓存区中;搜索单元,用以按照所述信息项在本地磁盘中查找对应的当前索引;更新单元,用以进行更新操作。所述比对系统中的数据更新方法和所述比对系统中的数据更新装置,避免大规模数据更新的延迟,使更新后的数据能够及时参与比对,获得更为准确的比对结果。



1. 一种比对系统中的数据更新方法,其特征在于,包括:
在物理内存中创建内存缓存区;
当所述比对系统存在更新的数据时,执行下列步骤:
根据更新的数据建立信息项,其中,所述信息项与所述比对系统的比对条件相对应;
将更新的数据及其信息项加载到所述内存缓存区中;
当进行更新操作时,执行下列步骤:
按照所述信息项在本地磁盘中查找对应的当前索引,进行本地磁盘的更新操作,其中,所述当前索引根据信息项的内容建立。
2. 如权利要求 1 所述的数据更新方法,其特征在于,所述更新的数据为数据库中新增的特征数据,则所述更新操作为将所述内存缓存区中的数据及其信息项加载到所述本地磁盘中成为特征数据,且重新创建当前索引。
3. 如权利要求 1 所述的数据更新方法,其特征在于,所述更新的数据为所述本地磁盘中修改后的特征数据,则所述更新操作为将所述内存缓存区中的数据替代所述本地磁盘中的原特征数据。
4. 如权利要求 1 所述的数据更新方法,其特征在于,所述更新的数据为所述本地磁盘中待删除的特征数据,则所述更新操作为删除所述本地磁盘中待删除的特征数据及其信息项,且重新创建当前索引。
5. 如权利要求 1-4 任一项所述的数据更新方法,其特征在于,所述当前索引根据下述步骤创建:
将数据的信息项转换为数值型信息数据;
根据所述数值型信息数据对所述数据进行排序;
创建对应排序后的数据的索引表为当前索引。
6. 如权利要求 2 所述的数据更新方法,其特征在于,所述重新创建当前索引包括:
将所述新增的特征数据的信息项转换为数值型信息数据;
将所述数值型信息数据插入当前索引中。
7. 如权利要求 4 所述的数据更新方法,其特征在于,所述重新创建当前索引包括:
将所述待删除的特征数据的信息项转换为数值型信息数据;
在所述当前索引中查找所述数值型信息数据对应的键值,且删除当前索引中对应的所述键值。
8. 如权利要求 1-4 任一项所述的数据更新方法,其特征在于,所述更新操作定时进行一次。
9. 如权利要求 1-4 任一项所述的数据更新方法,其特征在于,所述更新数据的数量达到预定值,进行一次所述更新操作。
10. 如权利要求 1 所述的数据更新方法,其特征在于,还包括在所述本地磁盘中创建磁盘缓存区,将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。
11. 如权利要求 1 所述的数据更新方法,其特征在于,所述本地磁盘为比对服务器的硬盘,所述物理内存为所述比对服务器的内存。
12. 如权利要求 1 所述的数据更新方法,其特征在于,所述数据为人脸特征数据或指纹

特征数据,所述数据的信息项包括性别、年龄和区域。

13. 一种比对系统中的数据更新装置,其特征在于,包括:

创建单元,用以在物理内存中创建内存缓存区;

加载单元,用以当所述比对系统存在更新的数据时,根据更新的数据建立信息项,将更新的数据及其信息项加载到所述内存缓存区中,其中,所述信息项与所述比对系统的比对条件相对应;

搜索单元,用以当进行更新操作时,按照所述信息项在本地磁盘中查找对应的当前索引,其中,所述当前索引根据信息项的内容建立;

更新单元,用以当进行更新操作时,进行本地磁盘的更新操作。

14. 如权利要求 13 所述的数据更新装置,其特征在于,所述更新的数据为数据库中新增的特征数据,则所述更新单元将所述内存缓存区中的数据及其信息项加载到所述本地磁盘中成为特征数据。

15. 如权利要求 13 所述的数据更新装置,其特征在于,所述更新的数据为所述本地磁盘中修改后的特征数据,则所述更新单元将所述内存缓存区中的数据替代所述本地磁盘中的原特征数据。

16. 如权利要求 13 所述的数据更新装置,其特征在于,所述更新的数据为所述本地磁盘中待删除的特征数据,则所述更新单元删除所述本地磁盘中待删除的特征数据及其信息项。

17. 如权利要求 13-16 任一项所述的数据更新装置,其特征在于,还包括:

转换单元,用以将数据的信息项转换为数值型信息数据;

排序单元,用以根据所述数值型信息数据对所述数据进行排序;

操作单元,用以创建对应排序后的数据的索引表为当前索引。

18. 如权利要求 17 所述的数据更新装置,其特征在于,所述更新的数据为数据库中新增的特征数据,则:

所述转换单元,还用以将所述新增的特征数据的信息项转换为数值型信息数据;

所述操作单元,还用以将所述数值型信息数据插入当前索引中。

19. 如权利要求 17 所述的数据更新装置,其特征在于,所述更新的数据为所述本地磁盘中待删除的特征数据,则:

所述转换单元,还用以将所述待删除的特征数据的信息项转换为数值型信息数据;

所述操作单元,还用以在所述当前索引中查找所述数值型信息数据对应的键值,且删除当前索引中对应的所述键值。

20. 如权利要求 13-16 任一项所述的数据更新装置,其特征在于,还包括控制单元,用以控制所述更新单元定时进行一次所述更新操作。

21. 如权利要求 13-16 任一项所述的数据更新装置,其特征在于,还包括控制单元,用以在所述更新数据的数量达到预定值时,控制所述更新单元进行一次所述更新操作。

22. 如权利要求 13 所述的数据更新装置,其特征在于,还包括在所述本地磁盘中创建磁盘缓存区,所述加载单元将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。

23. 如权利要求 13 所述的数据更新装置,其特征在于,所述本地磁盘为比对服务器的

硬盘,所述物理内存为所述比对服务器的内存。

24. 如权利要求 13 所述的数据更新装置,其特征在于,所述数据为人脸特征数据或指纹特征数据,所述数据的信息项包括性别、年龄和区域。

比对系统中的数据更新方法及装置

技术领域

[0001] 本发明涉及比对系统,尤其涉及比对系统中的数据更新方法及装置。

背景技术

[0002] 随着技术的进步,社会的需要,比对系统逐渐在各行业领域中广泛的应用。由于比对系统主要基于数据的支持,因此目前的比对系统多为大规模比对系统甚至是超大规模的比对系统,如人像识别系统中的人脸特征比对系统、指纹系统中的指纹特征比对系统。所述比对系统的数据通常储存在数据库中,存储量一般可达到百万甚至数千万、上亿,每条数据的大小一般在数 KB 至数十 KB 之间,因此比对系统中的数据总量一般可达数 GB 至数百 GB。

[0003] 为提高比对结果的精度,所述数据库中会不断添加新的数据,且还有可能对已有数据进行修改、删除等更新操作。如此海量的数据若同时进行更新操作,不仅会对比对操作造成影响,减低比对结果的精度,而且还会延长更新的时间,造成更新延迟。

[0004] 申请号为 200710096870.3 的发明专利申请,公开了一种数据更新的方法,应用于客户端与服务器存储数据的同步更新。但是如何避免大规模数据更新的延迟,使更新后的数据能够及时参与比对,获得更为准确的比对结果成为本领域技术人员亟待解决的问题之一。

发明内容

[0005] 本发明解决的问题是提供一种比对系统中的数据更新方法及装置,以避免大规模数据更新的延迟,使更新后的数据能够及时参与比对,获得更为准确的比对结果。

[0006] 为解决上述问题本发明提供了一种比对系统中的数据更新方法,包括:

[0007] 在物理内存中创建内存缓存区,将更新的数据及其信息项加载到所述内存缓存区中;

[0008] 按照所述信息项在本地磁盘中查找对应的当前索引,进行更新操作。

[0009] 可选的,所述更新的数据为数据库中新增的特征数据,则所述更新操作为将所述内存缓存区中的数据及其信息项加载到所述本地磁盘中成为特征数据,且重新创建当前索引。

[0010] 可选的,所述更新的数据为所述本地磁盘中修改后的特征数据,则所述更新操作为将所述内存缓存区中的数据替代所述本地磁盘中的原特征数据。

[0011] 可选的,所述更新的数据为所述本地磁盘中待删除的特征数据,则所述更新操作为删除所述本地磁盘中待删除的特征数据及其信息项,且重新创建当前索引。

[0012] 可选的,所述当前索引根据下述步骤创建:

[0013] 将数据的信息项转换为数值型信息数据;

[0014] 根据所述数值型信息数据对所述数据进行排序;

[0015] 创建对应排序后的数据的索引表为当前索引。

[0016] 可选的,所述重新创建当前索引包括:

- [0017] 将所述新增的特征数据的信息项转换为数值型信息数据；
- [0018] 将所述数值型信息数据插入当前索引中。
- [0019] 可选的,所述重新创建当前索引包括：
- [0020] 将所述待删除的特征数据的信息项转换为数值型信息数据；
- [0021] 在所述当前索引中查找所述数值型信息数据对应的键值,且删除当前索引中对应的所述键值。
- [0022] 可选的,所述更新操作定时进行一次。
- [0023] 可选的,所述更新数据的数量达到预定值,进行一次所述更新操作。
- [0024] 可选的,所述数据更新方法还包括在所述本地磁盘中创建磁盘缓存区,将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。
- [0025] 可选的,所述本地磁盘为比对服务器的硬盘,所述物理内存为所述比对服务器的内存。
- [0026] 可选的,所述数据为人脸特征数据或指纹特征数据,所述数据的信息项包括性别、年龄和区域。
- [0027] 为解决上述问题,本发明还提供了一种比对系统中的数据更新装置,包括：
- [0028] 创建单元,用以在物理内存中创建内存缓存区；
- [0029] 加载单元,用以将更新的数据及其信息项加载到所述内存缓存区中；
- [0030] 搜索单元,用以按照所述信息项在本地磁盘中查找对应的当前索引；
- [0031] 更新单元,用以进行更新操作。
- [0032] 可选的,所述更新的数据为数据库中新增的特征数据,则所述更新单元将所述内存缓存区中的数据及其信息项加载到所述本地磁盘中成为特征数据。
- [0033] 可选的,所述更新的数据为所述本地磁盘中修改后的特征数据,则所述更新单元将所述内存缓存区中的数据替代所述本地磁盘中的原特征数据。
- [0034] 可选的,所述更新的数据为所述本地磁盘中待删除的特征数据,则所述更新单元删除所述本地磁盘中待删除的特征数据及其信息项。
- [0035] 可选的,所述数据更新装置还包括：
- [0036] 转换单元,用以将数据的信息项转换为数值型信息数据；
- [0037] 排序单元,用以根据所述数值型信息数据对所述数据进行排序；
- [0038] 操作单元,用以创建对应排序后的数据的索引表为当前索引。
- [0039] 可选的,所述更新的数据为数据库中新增的特征数据,则：
- [0040] 所述转换单元,还用以将所述新增的特征数据的信息项转换为数值型信息数据；
- [0041] 所述操作单元,还用以将所述数值型信息数据插入当前索引中。
- [0042] 可选的,所述更新的数据为所述本地磁盘中待删除的特征数据,则：
- [0043] 所述转换单元,还用以将所述待删除的特征数据的信息项转换为数值型信息数据；
- [0044] 所述操作单元,还用以在所述当前索引中查找所述数值型信息数据对应的键值,且删除当前索引中对应的所述键值。
- [0045] 可选的,所述数据更新装置还包括控制单元,用以控制所述更新单元定时进行一

次所述更新操作。

[0046] 可选的,所述数据更新装置还包括控制单元,用以在所述更新数据的数量达到预定值时,控制所述更新单元进行一次所述更新操作。

[0047] 可选的,所述数据更新装置还包括在所述本地磁盘中创建磁盘缓存区,所述加载单元将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。

[0048] 可选的,所述本地磁盘为比对服务器的硬盘,所述物理内存为所述比对服务器的内存。

[0049] 可选的,所述数据为人脸特征数据或指纹特征数据,所述数据的信息项包括性别、年龄和区域。

[0050] 与现有技术相比,本发明具有以下优点:

[0051] 在物理内存中创建内存缓存区,将更新的数据及其信息项加载到所述内存缓存区中,这可以加速读取所述数据的速度,缩短比对时间。按照所述信息项在本地磁盘中查找对应的当前索引,进行更新操作,这不仅可以保护更新数据不被破坏,而且方便数据的更新。

[0052] 对不同类型的更新数据,采取不同的更新操作:新增的数据,采取插入式的更新操作;修改后的数据,采取替代式的更新操作;待删除的数据,采取删除式的更新操作。这样不仅可以迅速的完成更新操作,而且不会破坏原数据的完整性。

[0053] 创建所述当前索引的方法一般是:将所述信息项转换为数值型信息数据;根据所述数值型信息数据对所述数据进行排序;创建对应排序后的数据的索引表为当前索引。根据上述方法建立的当前索引不仅创建的时间短,而且所述当前索引排列有序,方便查找数据。

[0054] 为了缩短更新操作的时间和提高更新数据的速度,针对不同的更新数据,重新创建当前索引的方式略有不同。新增的数据采取以下方式:将所述数据的信息项转换为数值型信息数据,将所述数值型信息数据插入原当前索引中。这种方式仅需要重建部分当前索引即可创建新的当前索引,节省了时间。待删除的数据采用以下方式:将所述数据的信息项转换为数值型信息数据;在所述原当前索引中查找所述数值型信息数据对应的键值,且删除原当前索引中对应的所述键值。这种方式也是局部修改当前索引形成新的当前索引,节省时间,减少资源浪费。

[0055] 所述更新操作定时进行一次,或者所述更新数据的数量达到某一预定值,进行一次。这不仅可以防止数据堵塞,避免实时更新数据影响比对结果,而且提高了物理内存空间的利用率,防止数据冗余。

[0056] 为保护数据安全,在所述本地磁盘中创建磁盘缓存区,将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。这保证了更新数据的备份,防止数据丢失,造成难以寻回的后果。

[0057] 所述本地磁盘为比对服务器的硬盘,所述物理内存为所述比对服务器的内存。这可以方便比对系统控制,数据储存在对比服务器硬盘上,比对系统可以直接管理数据,不需要进行数据转换。所述物理内存为所述比对服务器的内存,比对系统直接读取内存中的数据,缩短比对时间,提高效率。

[0058] 所述比对系统一般为人像比对系统或指纹比对系统,因此所述特征数据为人脸特

征数据或指纹特征数据,所述特征数据的信息项包括性别、年龄和区域。

[0059] 本发明还提供了一种比对系统中的数据更新装置,实现了上述比对系统中的数据更新方法。通过所述数据更新装置,避免大规模数据更新的延迟,使更新后的数据能够及时参与比对,获得更为准确的比对结果。

附图说明

[0060] 图 1 是本发明提供的数据更新方法的一种实施方式的流程图;

[0061] 图 2 是本发明提供的数据更新方法中建立索引的一种实施方式的流程图;

[0062] 图 3 是本发明提供的数据更新装置的一种实施方式的结构示意图。

具体实施方式

[0063] 本发明提供了一种比对系统中的数据更新方法及其装置,达到了避免大规模数据更新的延迟,使更新后的数据能够及时参与比对,获得更为准确的比对结果。

[0064] 为使本发明的上述目的、特征和优点能够更为明显易懂,下面结合附图对本发明的具体实施方式做详细的说明。

[0065] 在以下描述中阐述了具体细节以便于充分理解本发明。但是本发明能够以多种不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本发明内涵的情况下做类似推广。因此本发明不受下面公开的具体实施方式的限制。

[0066] 由于比对系统需要基于大量数据的支持,才能更好的完成比对任务,达到较高的比对结果。因此目前的比对系统多为大规模比对系统甚至是超大规模的比对系统。所述比对系统的数据存储量一般可达到百万甚至数千万、上亿,每条数据的大小一般在数 KB 至数十 KB 之间,因此比对系统中的数据总量一般可达数 GB 至数百 GB。

[0067] 为了提高比对结果的精度,比对系统中数据的更新频率较快,若不能将更新的数据及时参与到比对中,容易出现比对结果不准的情况。鉴于上述情况本发明提供了一种比对系统中的数据更新方法,图 1 是本发明提供的数据更新方法的一种实施方式的流程图,下面结合如图 1 详细说明。

[0068] 步骤 S101,在物理内存中创建内存缓存区,将更新的数据及其信息项加载到所述内存缓存区中。

[0069] 步骤 S102,按照所述信息项在本地磁盘中查找对应的当前索引,进行更新操作。

[0070] 由于比对系统大都安装在大型计算机或数据处理器等设备上,通过中央处理器等硬件获取数据,完成比对请求。因此此处以安装在计算机上的人像识别系统为例,具体说明本实施例中数据更新的方法,但不限于下述举例内容的描述,可做适应性修改。

[0071] 比对系统一般安装在某一外部主服务器上,例如性能较优的大型计算机上。所述主服务器再与一个或多个比对服务器相连接,实现并行操作完成比对请求。由上述可知比对服务器大都为计算机,所述本地磁盘一般指比对服务器的硬盘,所述物理内存一般为所述比对服务器的内存。

[0072] 由于比对系统中数据量庞大,实时更新数据会与比对请求发生冲突,这样很容易对比对结果造成不良影响。若数据更新的操作在某一时间段内统一进行,则容易造成更新延迟,也就是说,先更新数据再进行比对,这需要经过一段时间,才能反映在比对结果中。步

骤 S101 中,将更新的数据放入缓存区则改善了上述情况。

[0073] 缓存区是数据交换的缓冲区,当计算机的某一硬件要读取数据时,会首先从缓存区中查找需要的数据,如果找到了则直接执行,找不到的话则从内存中找,若内存中找不到,进入硬盘或其他存储介质中寻找。也就是说,将更新的数据储存在缓存区内,计算机会先读取缓存区的更新数据,再读取内存中其他的数据,这样不仅不会延误比对时间,而且更新的数据也会及时的参与比对过程,不会造成比对误差。由此可知,缓存区中的数据是最先被读取的,步骤 S101 中,在内存中创建缓存区存放更新的数据及其信息项,可以大幅度的提高数据的读取速度,缩短比对时间,提高比对效率,而且由于更新的数据可以及时参与比对,不会影响比对结果,反而提高比对结果的精准度。

[0074] 为保证比对系统中的海量数据的信息安全,防止数据遗失,所述数据一般储存在外部数据库中,为方便说明下述内容中出现的的数据除特别说明外,均为特征数据。为提高数据的利用率,减少数据冗余,一般存入数据库的数据为特征数据。对于人脸特征比对系统来讲,所述特征数据一般是指人脸特征数据,例如脸型、眼型、唇形等五官的形状。所述特征数据还可以包括其他明显的区别特征,例如是否有痣、痣的位置等信息。

[0075] 所述特征数据储存在数据库中,通常分类存储,也就是说,将具有相同特点的所述特征数据归纳为一类,并将上述相同特点作为所述特征数据的信息项,共同存入数据库中。一般来讲,所述信息项为数据库中的字段名,可以是性别、年龄、区域等,所述区域通常指出生地、居住地等地域信息。比对系统一般均具有相应的比对条件,所述比对条件与所述信息项大都对应设置,也就是说,所述比对条件与所述信息项基本相同,此处所述比对条件为性别、年龄、区域等。

[0076] 不同的比对系统,所述特征数据、所述信息项、所述比对条件一般不尽相同,对于指纹比对系统来讲,所述特征数据为指纹特征数据,例如指纹纹型。由于指纹比对系统与人脸特征比对系统都是以人为参照物的,所以指纹比对系统中的所述信息项、所述比对条件可以为性别、年龄、区域等信息。此处仅是举例说明,针对不同的比对系统,参照比对系统的功能确定上述特征数据、信息项和比对条件。

[0077] 所述数据库中的特征数据通常加载到本地磁盘中。这样一方面保证所述数据库和所述本地磁盘可以互为备份数据,防止数据丢失,若所述数据库或所述本地磁盘发生故障,造成数据缺失,可以相互弥补;另一方面,保护数据的安全,防止所述特征数据被篡改。

[0078] 所述本地磁盘加载所述特征数据时,通常会一并加载所述特征数据的信息项,且会按照所述信息项创建索引储存。将所述特征数据连同所述信息项一起加载,一方面防止数据缺失,另一方面方便制作索引。所述特征数据和所述信息项,按照一定的组织方式,一般是一对一的映射关系,组合成一个或多个文件,加载到所述本地磁盘中。按照如上方式组合的具有所述信息项的文件,按照所述信息项的内容形成当前索引。

[0079] 步骤 S101 中更新的数据及其信息项加载到所述内存缓存区中,在步骤 S102 中,借助所述信息项在本地磁盘中查找对应的当前索引,进行更新操作。由上述可知,本地磁盘中应的当前索引是根据数据的信息项建立的,因此借助所述内存缓存区中数据的信息项,不仅可以方便的找到对应的所述当前索引,节省时间,提高速度,便于数据更新的操作,而且信息项对称,减少查找数据错误的情况,降低误差的出现几率。

[0080] 所述更新操作一般分为增加新数据,修改数据,删除数据三种,根据更新数据的类

型,进行不同的操作。

[0081] 若所述更新的数据为数据库中新增的特征数据,则所述更新操作为将所述内存缓存区中的数据及其信息项加载到所述本地磁盘中成为特征数据,且重新创建当前索引。

[0082] 若所述更新的数据为所述本地磁盘中修改后的特征数据,则所述更新操作为将所述内存缓存区中的数据替代所述本地磁盘中的原特征数据。

[0083] 若所述更新的数据为所述本地磁盘中待删除的特征数据,则所述更新操作为删除所述本地磁盘中待删除的特征数据及其信息项,且重新创建当前索引。

[0084] 上述更新操作都是在比对服务器中进行的,也就是说在本地磁盘中进行。数据比对过程中,有时会对数据进行如上所述的三种形式的修改,擅自改动所述特征数据,会对后续的比对造成影响,因此所述数据库中原始数据需要存档,将所述特征数据加载到所述本地磁盘中,若发生篡改,可以与所述数据库中的原始数据比对,保证数据的正确性。

[0085] 由于数据发生更改,因此索引也会做适应性调整,图2是本发明提供的数据更新方法中建立索引的一种实施方式的流程图,下面结合图2详细说明具体过程。

[0086] 步骤S1,将数据的信息项转换为数值型信息数据;

[0087] 步骤S2,根据所述数值型信息数据对所述数据进行排序;

[0088] 步骤S3,创建对应排序后的数据的索引表为当前索引。

[0089] 众所周知,计算机读取数值型的数据速度最快,因此根据所述数据的信息项创建当前索引时,将所述信息项转换为数值型信息数据。另外,数值型的数据排序最为简单,因此根据所述数值型信息数据对特征数据进行排序,创建对应排序后的特征数据的索引表,能够较大幅度的提高数据加载的速度。所述数值型一般为整型、无符号整型或枚举型,数值型为优选方案,也可以是计算机可读的其他类型的数据。

[0090] 上述索引表的创建流程适用于大部分的索引创建,也就是说,图2所示的索引创建方式不仅适用于更新数据过程中,还用于其他数据加载过程,是一种使用较为广泛的索引创建方式。由于更新数据的信息项对索引表的影响不大,因此为了提高工作效率,缩短创建索引的时间,还可以根据不同的更新方式可以采取不同的创建方式。

[0091] 若所述更新的数据为数据库中新增的特征数据,则将所述数据的信息项转换为数值型信息数据;将所述数值型信息数据插入原当前索引中。一般情况下,新增的特征数据在所述本地磁盘中具有对应的索引信息,因此找到对应的信息数据,将新增特征数据的信息项改变数据类型后,插入原当前索引就可以形成新的索引,且不会对其他数据造成影响。

[0092] 若所述更新的数据为所述本地磁盘中待删除的特征数据,则将所述数据的信息项转换为数值型信息数据;在所述原当前索引中查找所述数值型信息数据对应的键值,且删除原当前索引中对应的所述键值。由于删除所述本地磁盘中的特征数据,其信息项会一并删除,对其他的数据没有影响,因此在所述原当前索引中查找该特征数据的信息项对应的信息,且删除原当前索引中对应的所述信息创建索引。

[0093] 若所述更新的数据为所述本地磁盘中修改后的特征数据,由于仅是对所述本地磁盘中的特征数据做修改,并不影响其信息项,所以一般不会影响索引的改变。若修改所述本地磁盘中的特征数据涉及到其信息项的修改,可以按照图2所示的流程创建新的索引,也可以按照新增特征数据创建索引的方式创建,还可以通过删除的方式创建。上述创建索引的方式,在实际设计中可以根据需要进行适应性修改,不限于上述描述内容。

[0094] 比对服务器通常情况下会长时间的开机运转,因此有可能会发生掉电或故障的情况。若发生上述情况,内存中的数据经常会丢失,为防止数据丢失,还在所述本地磁盘中创建磁盘缓存区,将更新的数据及其信息项加载到所述磁盘缓存区中,所述磁盘缓存区中的数据与所述内存缓存区中的数据相同。这样即使比对服务器掉电或者发生故障,还能保证硬盘中存有备份数据,保证数据安全。

[0095] 由于大规模的数据更新会带来时间延迟,因此所述更新操作定时进行一次,或者所述更新数据的数量达到预定值,进行一次所述更新操作。更新时间一般 24 小时,所述预定值范围一般为 1 万条至 10 万条。上述更新时间和预定值可根据需要进行更改,不受上述内容限制。

[0096] 为实现所述比对系统中的数据更新方法,本发明还提供了一种比对系统中的数据更新装置。图 3 是本发明提供的数据更新装置的一种实施方式的结构示意图,包括:

[0097] 创建单元 A1,用以在物理内存中创建内存缓存区;

[0098] 加载单元 A2,用以将更新的数据及其信息项加载到所述内存缓存区中;

[0099] 搜索单元 A3,用以按照所述信息项在本地磁盘中查找对应的当前索引;

[0100] 更新单元 A4,用以进行更新操作。

[0101] 所述更新单元 A4 定时,或者所述更新数据的数量达到预定值,进行一次所述更新操作。所述预定值范围一般在一万至十万条,所述更新周期和所述预定值可以预先设置在所述更新单元 A4 中,也可以人工根据需要随时设置,不受上述内容的限制。

[0102] 其中创建单元 A1,加载单元 A2,搜索单元 A3,更新单元 A4 依次连接。由于每种比对系统的功能和作用不尽相同,因此上述列举的处理装置只是一种实施方式,根据不同的比对系统可做适应性修改,例如增加计时单元,信号处理单元,而且装置的数量也不受限制。

[0103] 加载单元 A2 主要是将数据加载入所述本地磁盘,因此加载单元 A2 还可以与所述本地磁盘 A100 连接,一般加载单元 A2 通过串口线与所述本地磁盘 A100 连接,也可以采用其他的方式。搜索单元 A3 主要是按照所述信息项在本地磁盘 A100 中查找对应的当前索引;更新单元 A4 主要是在所述本地磁盘 A100 中进行更新操作,因此所述本地磁盘 A100 还分别与搜索单元 A3、更新单元 A4 相连接。在计算机中,磁盘和内存之间存在着数据通信,因此所述本地磁盘 A100 与物理内存 A200 之间也存在连接关系。所述物理内存 A200 还分别与所述创建单元 A1、所述更新单元 A4 相连接。

[0104] 另外,本实施例中所述特征数据是定期或者定量的加载到所述本地地磁 A100 中,因此所述数据更新装置还包括控制单元(图中未标示),与更新单元 A4 相连接,用以定时控制所述更新单元 A4 进行一次更新操作;或者,用以在更新数据的数量达到预定值时,控制所述更新单元 A4 进行一次更新操作。

[0105] 定时加载数据时,还可以增加计时单元(图中未标示),用以计时,累计时间满 24 小时,发出加载信号或消息到所述控制单元,再由所述控制单元控制更新单元 A4 进行一次更新操作;定量加载数据时,还可以增加计量单元(图中未标示),用以计算更新数据的数量,达到预定值时,发出加载信号或消息到所述控制单元,再由所述控制单元控制加载单元 A2 进行一次更新操作。所述预定值范围一般在一万至十万条,这个数目可以更改且不会影响加载效果。

[0106] 在数据加载和数据更新的过程中,大都会涉及创建索引的步骤,因此上述装置还

包括：

[0107] 转换单元（图中未标示），用以将数据的信息项转换为数值型信息数据；

[0108] 排序单元（图中未标示），用以根据所述数值型信息数据对所述数据进行排序；

[0109] 操作单元（图中未标示），用以创建对应排序后的数据的索引表为当前索引。

[0110] 上述单元之间一般顺次连接，由于功能需要也可以按照其他方式连接。由前述内容可知，上述单元基于数据的信息项创建索引，且所述数据不限于更新的数据，也可以是原始数据或其他数据。上述单元具有通用性，不限于特定的数据类型，也不限于特定的使用环境，普通的数据加载过程中也可以用以创建索引。

[0111] 由于所述本地磁盘定期或定量进行一次更新操作，因此每次更新的数据有可能不尽相同。由于每次更新数据，都会重新创建所述索引表，因此创建索引的过程是动态的。也就是说，每次更新数据创建的所述当前索引不一定会相同。若数据简单，变化不大，还有可能会创建一级索引；若数据复杂，信息量大，则有可能创建多级索引，而且每次创建的索引级数也不一定相同。这样可以适应所述特征数据的形式，设置不同的索引，方便后续的储存和排列，以及比对工作。

[0112] 转换单元，转换数据的信息项的数据类型，排序单元，对所述数据进行排序，操作单元，对应排序后的数据的索引表为当前索引。上述创建索引的方法与图 2 中流程图相同，此处不再多做赘述。由于不同数据类型，创建索引的方式略有不同，因此上述转换单元，排序单元，操作单元可根据需要进行增减，例如去掉排序单元，保留转换单元和操作单元。

[0113] 当所述更新的数据为数据库中新增的特征数据，则：所述转换单元，还用以将所述新增的特征数据的信息项转换为数值型信息数据；所述操作单元，还用以将所述数值型信息数据插入当前索引中。

[0114] 当所述更新的数据为所述本地磁盘中待删除的特征数据，则：所述转换单元，还用以将所述待删除的数据的信息项转换为数值型信息数据；所述操作单元，还用以在所述当前索引中查找所述数值型信息数据对应的键值，且删除当前索引中对应的所述键值。

[0115] 由上述内容可知，所述转换单元负责数据类型的转换，所述操作单元进行插入操作或删除操作，在具体设计时可做适应性调整。这样不仅可以提高索引创建的步骤，节省时间，而且减少设备耗损，节约成本。

[0116] 上述转换单元、排序单元和操作单元的工作也可以由所述加载单元完成，即所述加载单元在进行数据加载操作时，还会进行创建索引的操作，并且每次加载数据，都可以根据数据类型进行相应的索引创建操作。在更新数据时，则可以由所述更新单元通知所述加载单元重新创建当前索引。

[0117] 上述装置的连接关系不受举例内容的限制，装置之间一般通过线路连接，有时可能无线连接，或者通过网络远程控制，信号控制。所述装置的数量和种类也不受上述内容的限制，根据实际情况，可以灵活设置。

[0118] 本发明虽然已以具体实施例公开如上，但其并不是用来限定本发明，任何本领域技术人员在不脱离本发明的精神和范围内，都可以利用上述揭示的方法和技术内容对本发明技术方案做出可能的变动和修改，因此，凡是未脱离本发明技术方案的内容，依据本发明的技术实质对以上实施例所作的任何简单修改、等同变化及修饰，均属于本发明技术方案的保护范围。

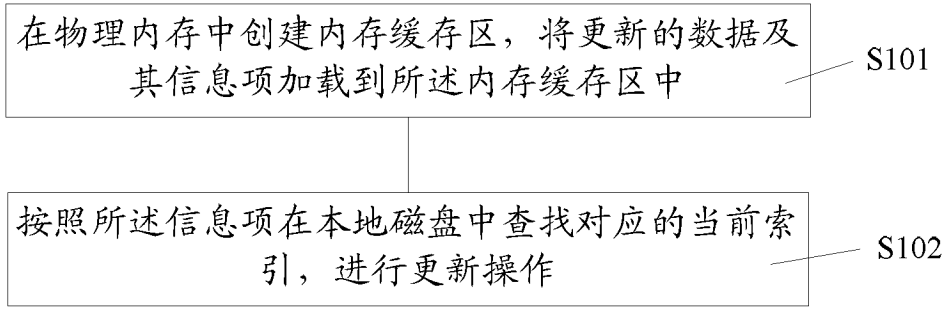


图 1

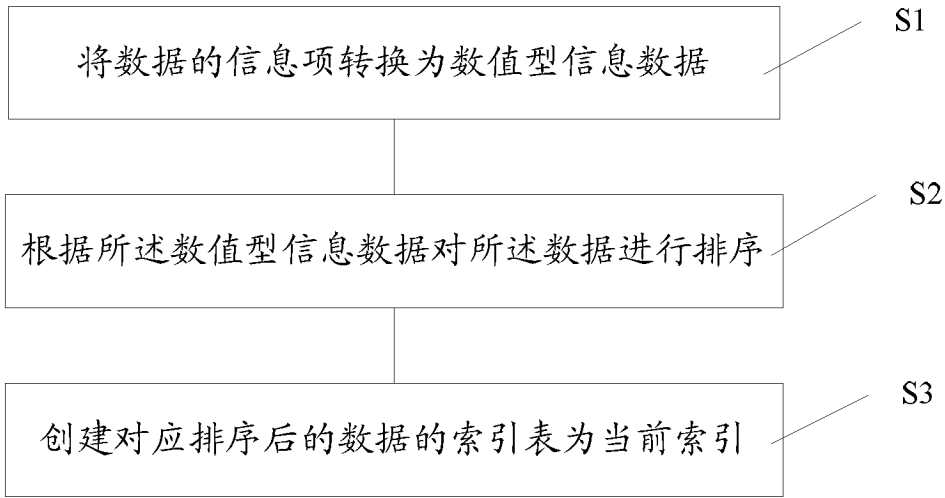


图 2

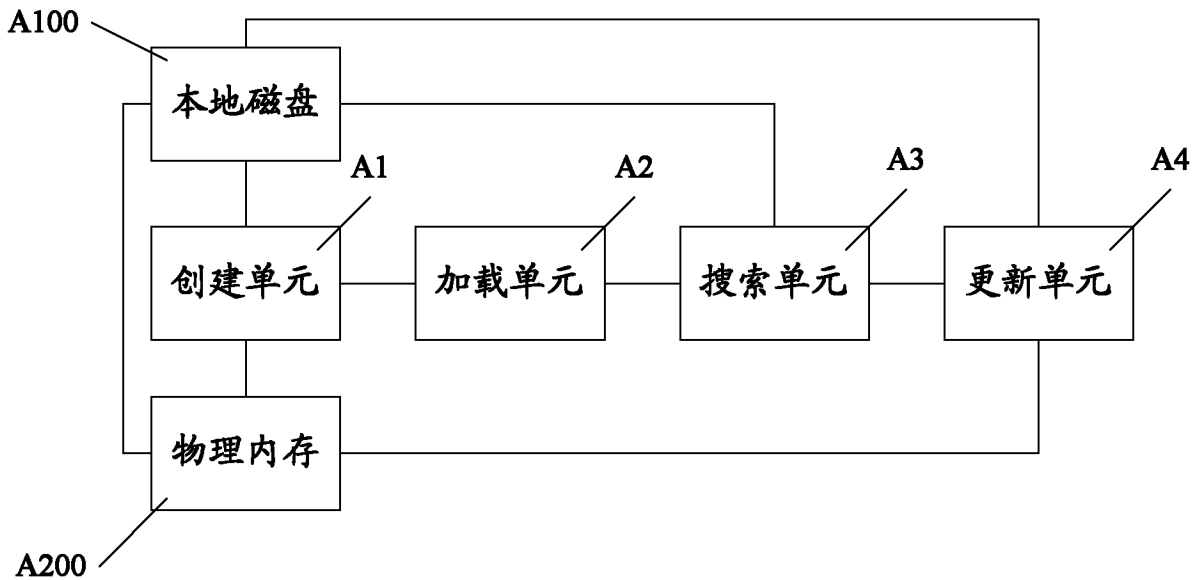


图 3