



(12)发明专利申请

(10)申请公布号 CN 105843859 A

(43)申请公布日 2016.08.10

(21)申请号 201610152630.X

(22)申请日 2016.03.17

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 冷继南 丹尼斯·帕尔霍维科

牛进保 沈建强 王工艺

伊万·马祖连科

(74)专利代理机构 北京龙双利达知识产权代理有限公司 11329

代理人 毛威 时林

(51)Int.Cl.

G06F 17/30(2006.01)

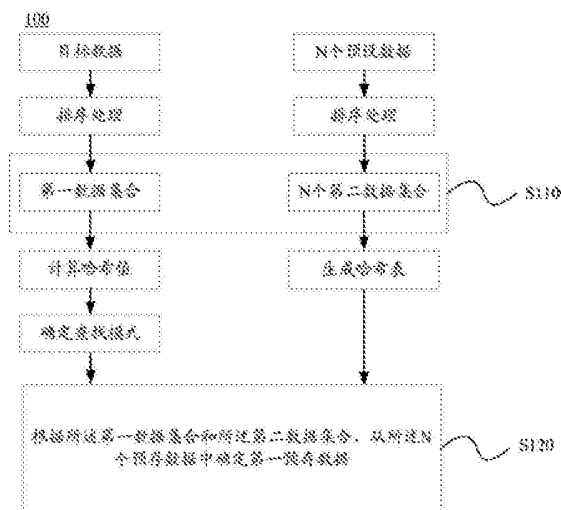
权利要求书4页 说明书18页 附图6页

(54)发明名称

数据处理的方法、装置和设备

(57)摘要

一种数据处理的方法、装置和设备,该方法包括:根据相似度阈值j和目标数据,生成第一数据集合,目标数据包括T1个第一比特组,第一数据集合包括M1个第一数据,M1个第一数据与从T1个第一比特组中选取j个第一比特组时的M1种组合方式一一对应;根据j和N个预存数据,生成N个第二数据集合;N个预存数据与N个第二数据集合一一对应,每个预存数据包括T2个第二比特组,每个第二数据集合包括M2个第二数据,第i个第二数据集合中的每个第二数据包括第i个预存数据中的T2个第二比特组,第i个第二数据集合中的M2个第二数据与从T2个第二比特组中选取j个第二比特组时的M2种组合方式一一对应;根据第一数据集合和第二数据集合,从N个预存数据中确定第一预存数据,能够降低相似数据查找过程的复杂度。



1. 一种数据处理的方法,其特征在于,所述方法包括:

根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,所述目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,所述第一数据集合包括 M_1 个第一数据,所述 M_1 个第一数据与从所述 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2$, $k \geq 1$, $T_1 > j \geq 1$;

根据所述相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,所述 N 个预存数据与所述 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从所述 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1$, $T_2 \geq T_1$, $i \in [1, N]$;

根据所述第一数据集合和所述第二数据集合,从所述 N 个预存数据中确定第一预存数据,所述第一预存数据和所述目标数据之间相似度与所述相似度阈值 j 相对应,其中,所述第一预存数据所对应的第二数据集合与所述第一数据集合之间包括至少一个相同的数据。

2. 根据权利要求1所述的方法,其特征在于,在根据预设的相似度阈值 j 和目标数据,生成第一数据集合之前,所述方法还包括:

根据预设规则,对所述目标数据中的 T_1 个第一比特组进行排序;以及

在根据预设的相似度阈值 j 和目标数据,生成第一数据集合之前,所述方法还包括:

根据所述预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

3. 根据权利要求1或2所述的方法,其特征在于,所述根据预设的相似度阈值 j 和目标数据,生成第一数据集合,包括:

根据预设的相似度阈值 j 和所述第一比特组的数量 T_1 ,确定第一生成矩阵,所述第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,所述 T_1 个元素包括 j 个“1”元素和 $T_1 - j$ 个“0”元素,任意两个第一向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

根据所述第一生成矩阵和目标数据,生成第一数据集合。

4. 根据权利要求1至3中任一项所述的方法,其特征在于,所述根据所述相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,包括:

根据预设的相似度阈值 j 和所述第二比特组的数量 T_2 ,确定第二生成矩阵,所述第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,所述 T_2 个元素包括 j 个“1”元素和 $T_2 - j$ 个“0”元素,任意两个第二向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

根据所述第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

5. 根据权利要求1至4中任一项所述的方法,其特征在于,所述根据所述第一数据集合和所述第二数据集合,从所述 N 个预存数据中确定第一预存数据,包括:

从备选模式中确定查找模式,所述备选模式包括快速模式和完整模式;

当所述查找模式为所述快速模式时,基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,并在首次查找到与所述第一数据相同的数据时,将与

所述第一数据相同的数据所属于的第二数据集合所对应的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$;

当所述查找模式为所述完整模式时,基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,如果首次查找到的与所述第一数据相同的数据属于至少两个第二数据集合,则将所述至少两个第二数据集合所对应的预存数据中与所述目标数据之间存在相同的比特组的数量最多的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$ 。

6.一种数据处理的装置,其特征在于,所述装置包括:

生成单元,用于根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,所述目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,所述第一数据集合包括 M_1 个第一数据,所述 M_1 个第一数据与从所述 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2, k \geq 1, T_1 > j \geq 1$;

所示生成单元还用于,根据所述相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,所述 N 个预存数据与所述 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从所述 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;

确定单元,用于根据所述第一数据集合和所述第二数据集合,从所述 N 个预存数据中确定第一预存数据,所述第一预存数据和所述目标数据之间相似度与所述相似度阈值 j 相对应,其中,所述第一预存数据所对应的第二数据集合与所述第一数据集合之间包括至少一个相同的数据。

7.根据权利要求6所述的装置,其特征在于,所述装置还包括:

排序单元,用于根据预设规则,对所述目标数据中的 T_1 个第一比特组进行排序;

用于根据所述预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

8.根据权利要求6或7所述的装置,其特征在于,所述生成单元具体用于根据预设的相似度阈值 j 和所述第一比特组的数量 T_1 ,确定第一生成矩阵,所述第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,所述 T_1 个元素包括 j 个“1”元素和 $T_1 - j$ 个“0”元素,任意两个第一向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

用于根据所述第一生成矩阵和目标数据,生成第一数据集合。

9.根据权利要求6至8中任一项所述的装置,其特征在于,所述生成单元具体用于根据预设的相似度阈值 j 和所述第二比特组的数量 T_2 ,确定第二生成矩阵,所述第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,所述 T_2 个元素包括 j 个“1”元素和 $T_2 - j$ 个“0”元素,任意两个第二向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

用于根据所述第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

10. 根据权利要求6至9中任一项所述的装置,其特征在于,所述确定单元具体用于从备选模式中确定查找模式,所述备选模式包括快速模式和完整模式;

当所述查找模式为所述快速模式时,所述确定单元具体用于基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,并在首次查找到与所述第一数据相同的数据时,将与所述第一数据相同的数据所属于的第二数据集合所对应的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$;

当所述查找模式为所述完整模式时,所述确定单元具体用于基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,如果首次查找到的与所述第一数据相同的数据属于至少两个第二数据集合,则将所述至少两个第二数据集合所对应的预存数据中与所述目标数据之间存在相同的比特组的数量最多的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$ 。

11. 一种数据处理的设备,其特征在于,所述设备包括:

总线;

与所述总线相连的存储器;

与所述总线相连的处理器;

所述处理器用于经由所述总线调用并执行所述存储器中的程序,以用于根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,所述目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,所述第一数据集合包括 M_1 个第一数据,所述 M_1 个第一数据与从所述 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2, k \geq 1, T_1 > j \geq 1$;

用于根据所述相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,所述 N 个预存数据与所述 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从所述 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;

用于根据所述第一数据集合和所述第二数据集合,从所述 N 个预存数据中确定第一预存数据,所述第一预存数据和所述目标数据之间相似度与所述相似度阈值 j 相对应,其中,所述第一预存数据所对应的第二数据集合与所述第一数据集合之间包括至少一个相同的数据。

12. 根据权利要求11所述的设备,其特征在于,所述处理器具体用于根据预设规则,对所述目标数据中的 T_1 个第一比特组进行排序;

用于根据所述预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

13. 根据权利要求11或12所述的设备,其特征在于,所述处理器具体用于根据预设的相似度阈值 j 和所述第一比特组的数量 T_1 ,确定第一生成矩阵,所述第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,所

述 T_1 个元素包括 j 个“1”元素和 $T_1 - j$ 个“0”元素,任意两个第一向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

用于根据所述第一生成矩阵和目标数据,生成第一数据集合。

14.根据权利要求11至13中任一项所述的设备,其特征在于,所述处理器具体用于根据预设的相似度阈值 j 和所述第二比特组的数量 T_2 ,确定第二生成矩阵,所述第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,所述 T_2 个元素包括 j 个“1”元素和 $T_2 - j$ 个“0”元素,任意两个第二向量彼此之间至少存在一个在所述第二维度方向上的排列位置不同的“0”元素或“1”元素;

用于根据所述第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

15.根据权利要求11至14中任一项所述的设备,其特征在于,所述处理器具体用于从备选模式中确定查找模式,所述备选模式包括快速模式和完整模式;

用于当所述查找模式为所述快速模式时,基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,并在首次查找到与所述第一数据相同的数据时,将与所述第一数据相同的数据所属于的第二数据集合所对应的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$;

用于当所述查找模式为所述完整模式时,基于各所述第一数据在所述第一数据集合中的排序,按预设顺序进行 y 次查找处理,所述 y 次查找处理与所述第一数据集合中基于所述预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定所述 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,如果首次查找到的与所述第一数据相同的数据属于至少两个第二数据集合,则将所述至少两个第二数据集合所对应的预存数据中与所述目标数据之间存在相同的比特组的数量最多的预存数据作为所述第一预存数据,其中, $x \in [1, y], y \leq M_1$ 。

数据处理的方法、装置和设备

技术领域

[0001] 本发明涉及数据信息技术领域,并且更具体地,涉及数据处理的方法、装置和设备。

背景技术

[0002] 相似检测技术广泛应用于互联网,图像识别,大数据分析和数据缩减等信息技术(IT,Information Technology)领域。相似数据查找是相似检测技术中的重要环节。

[0003] 随着对查找精度和智能性的要求,目前,该相似数据查找的输出结果需要是“相似”的数据,即,假设所处理的数据包括 α 个字节(Byte),则所输出的“相似”的数据之间有 β ($\beta < \alpha$)个字节相同。其中, β 可以是管理员或者系统规定的相似度阈值。

[0004] 如何降低相似数据查找过程的复杂度,成为业界亟需解决的问题。

发明内容

[0005] 本发明实施例提供一种数据处理的方法、装置和设备,能够降低相似数据查找过程的复杂度,减少相似数据查找的处理时间,改善用户体验。

[0006] 第一方面,提供了一种数据处理的方法,该方法包括:根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,该目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,该第一数据集合包括 M_1 个第一数据,该 M_1 个第一数据与从该 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2, k \geq 1, T_1 > j \geq 1$;根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,该 N 个预存数据与该 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从该 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;根据该第一数据集合和该第二数据集合,从该 N 个预存数据中确定第一预存数据,该第一预存数据和该目标数据之间相似度与该相似度阈值 j 相对应,其中,该第一预存数据所对应的第二数据集合与该第一数据集合之间包括至少一个相同的数据。

[0007] 根据本发明实施例的数据处理的方法,通过根据目标数据确定包括 M_1 个第一数据的第一数据集合,并根据 N 个预存数据确定 N 个第二数据集合,其中,第一数据集合中的 M_1 个第一数据与从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的 M_1 种组合方式一一对应,每个第二数据集合中的 M_2 个第二数据与从包括 T_2 个第二比特组的预存数据中选择 j 个第二比特组时的 M_2 种组合方式一一对应,其中, j 为预设的相似度阈值,从而,在一个第二数据集合与该第一数据集合之间包括至少一个相同的数据时,能够将该第二数据集合所对应的预存数据作为与该目标数据之间的相似度满足该相似度阈值 j 所对应的相似度要求的相似数据,即,能够将相似数据查找过程转化为相同数据的判定过程,从而,能够降低相似数据查找的复杂度,减少相似数据查找的处理时间,改善用户体验。

[0008] 结合第一方面,在第一方面的第一种实现方式中,该根据预设的相似度阈值 j 和目标数据,生成第一数据集合,包括:根据目标数据生成 M_3 个子目标数据,其中,该 M_3 个子目标数据与 T_1 个第一比特组的所有可能的排列方式一一对应;根据预设的相似度阈值 j 和该 M_3 个子目标数据,生成第一数据集合,其中,该 M_1 个第一数据与从该 M_3 个子目标数据中的每个子目标数据中选取 j 个第一比特组时的 M_1 种组合方式一一对应;以及,根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,包括:根据第 i 个预存数据生成 M_4 个子预存数据,其中,该 M_4 个子目标数据与第 i 个预存数据的 T_2 个第二比特组的所有可能的排列方式一一对应;根据预设的相似度阈值 j 和每个预存数据所对应的 M_4 个子预存数据,生成第二数据集合,其中,第 i 个第二数据集合中的 M_2 个第二数据与从该第 i 个预存数据所对应的 M_4 个子预存数据中的每个子预存数据中选取 j 个第二比特组时的 M_2 种组合方式一一对应。

[0009] 根据本发明实施例的数据处理的方法,通过确定目标数据中的各第一比特组所有可能的排列方式,并确定各预存数据中的各第二比特组的所有可能的排列方式,从而能够使该第一数据集合中的第一数据对应目标数据的各第一比特组的所有可能排列方式下从该 T_1 个第一比特组中选取 j 个第一比特组时的组合方式,使每个第二数据集合中的第二数据对应在所对应的预存数据的各第二比特组的所有可能排列方式下从该 T_2 个第二比特组中选取 j 个第二比特组时的组合方式,从而,能够提高的相似数据查找的可靠性和准确性。

[0010] 结合第一方面及其上述实现方式,在第一方面的第二种实现方式中,在根据预设的相似度阈值 j 和目标数据,生成第一数据集合之前,该方法还包括:根据预设规则,对该目标数据中的 T_1 个第一比特组进行排序;以及在根据预设的相似度阈值 j 和目标数据,生成第一数据集合之前,该方法还包括:根据该预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

[0011] 根据本发明实施例的数据处理的方法,通过在生成通过基于相同的预设规则对目标数据和预存数据中的各比特组进行排序,能够确保所确定的第一数据集合和第二数据集合中包括相同的比特组的数据中各比特组的位置也相同,从而,能够在确保相似数据查找的可靠性和准确性的前提下,能够进一步降低相似数据查找的复杂度,减少相似数据查找的处理时间。

[0012] 结合第一方面及其上述实现方式,在第一方面的第三种实现方式中,该根据预设的相似度阈值 j 和目标数据,生成第一数据集合,包括:根据预设的相似度阈值 j 和该第一比特组的数量 T_1 ,确定第一生成矩阵,该第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,该 T_1 个元素包括 j 个“1”元素和 T_1-j 个“0”元素,任意两个第一向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;根据该第一生成矩阵和目标数据,生成第一数据集合。

[0013] 结合第一方面及其上述实现方式,在第一方面的第四种实现方式中,该根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,包括:根据预设的相似度阈值 j 和该第二比特组的数量 T_2 ,确定第二生成矩阵,该第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,该 T_2 个元素包括 j 个“1”元素和 T_2-j 个“0”元素,任意两个第二向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;根据该第二生成矩阵和 N 个预存数据,生成 N 个第二数据集

合。

[0014] 结合第一方面及其上述实现方式,在第一方面的第五种实现方式中,该根据该第一数据集合和该第二数据集合,从该N个预存数据中确定第一预存数据,包括:从备选模式中确定查找模式,该备选模式包括快速模式和完整模式;当该查找模式为该快速模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行y次查找处理,该y次查找处理与该第一数据集合中基于该预设顺序的前y个第一数据一一对应,其中,第x次查找处理用于确定该N个第二数据集合中是否存在与第x个第一数据相同的数据,并在首次查找到与该第一数据相同的数据时,将与该第一数据相同的数据所属于的第二数据集合所对应的预存数据作为该第一预存数据,其中, $x \in [1, y]$, $y \leq M_1$;当该查找模式为该完整模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行y次查找处理,该y次查找处理与该第一数据集合中基于该预设顺序的前y个第一数据一一对应,其中,第x次查找处理用于确定该N个第二数据集合中是否存在与第x个第一数据相同的数据,如果首次查找到的与该第一数据相同的数据属于至少两个第二数据集合,则将该至少两个第二数据集合所对应的预存数据中与该目标数据之间存在相同的比特组的数量最多的预存数据作为该第一预存数据。

[0015] 根据本发明实施例的数据处理的方法,通过设置快速模式和完整模式,并在快速模式下输出预存数据中所有相似度满足相似度阈值j所对应的要求的数据,在完整模式输出预存数据中相似度满足相似度阈值j所对应的要求且与目标数据之间的相似度最高的数据,能够灵活应对不同的用户需求,进一步提高用户体验。

[0016] 结合第一方面及其上述实现方式,在第一方面的第六种实现方式中,该目标数据和该预存数据为指纹数据。

[0017] 结合第一方面及其上述实现方式,在第一方面的第七种实现方式中, $A(T_1, j) \geq M_1 \geq C(T_1, j)$, $A(T_2, j) \geq M_2 \geq C(T_2, j)$ 。

[0018] 第二方面,提供了一种数据处理的装置,包括用于执行上述第一方面以及第一方面的各实现方式中的各步骤的单元或模块。当本发明第一方面提供的方法通过软件模块实现时,本发明提供的数据处理的装置可以指示软件模块或软件包。

[0019] 第三方面,提供了一种数据处理的设备,包括存储器和处理器,该存储器用于存储计算机程序,该处理器用于从存储器中调用并运行该计算机程序,使得数据数据处理的设备执行上述第一方面,及其各种实现方式中的任一种数据处理的的方法。

[0020] 第四方面,提供了一种计算机可读存储介质,所述计算机可读存储介质存储有程序,当存储在所述计算机可读存储介质中的程序被计算机设备运行时,使得所述计算机设备执行上述第一方面,及其各种实现方式中的任一种数据传输的方法。

附图说明

[0021] 为了更清楚地说明本发明实施例的技术方案,下面将对本发明实施例中所需要使用的附图作简单地介绍,显而易见地,下面所描述的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0022] 图1是根据本发明实施例的数据处理的的方法的示意性流程图。

- [0023] 图2是本发明实施例中目标数据和预存数据的一例的示意图。
- [0024] 图3是本发明实施例中经过排序处理后的目标数据和预存数据的一例的示意图。
- [0025] 图4是根据本发明实施例的生成数据集合的过程的一例的示意图。
- [0026] 图5是根据本发明实施例的相似数据查找过程的示意图。
- [0027] 图6是根据本发明实施例的哈希表的生成方式的示意图。
- [0028] 图7是根据本发明实施例的哈希表的一例的示意图。
- [0029] 图8是根据本发明实施例的快速模式下的查找过程的示意性流程图。
- [0030] 图9是根据本发明实施例的数据处理的装置的示意性框图。
- [0031] 图10是根据本发明实施例的数据处理的设备的示意性结构图。

具体实施方式

[0032] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0033] 本发明实施例提供的数据处理的方法、装置和设备,可以应用于计算机上,该计算机包括硬件层、运行在硬件层之上的操作系统层,以及运行在操作系统层上的应用层。该硬件层包括CPU、存储器管理单元(MMU,Memory Management Unit)和内存(也称为存储器)等硬件。该操作系统可以是任意一种或多种通过进程实现业务处理的计算机操作系统,例如, Linux系统、Unix系统、Android系统、iOS系统或windows系统等。该应用层包含浏览器、通讯录、文字处理软件、即时通信软件等应用。并且,在本发明实施例中,该计算机可以是智能手机等手持设备,也可以是个人计算机等终端设备,本发明并未特别限定,只要能够通过运行记录有本发明实施例的数据处理的方法的代码的程序,以根据本发明实施例的数据处理的方法对数据进行处理即可。本发明实施例的数据处理的方法的执行主体可以是计算机设备,或者,本发明实施例的数据处理的方法的执行主体可以是计算机设备中能够调用程序并执行程序的功能模块。

[0034] 此外,本发明的各个方面或特征可以实现成方法、装置或使用标准编程和/或工程技术的制品。本申请中使用的术语“制品”涵盖可从任何计算机可读器件、载体或介质访问的计算机程序。例如,计算机可读介质可以包括,但不限于:磁存储器件(例如,硬盘、软盘或磁带等),光盘(例如,CD(Compact Disk,压缩盘)、DVD(Digital Versatile Disk,数字通用盘)等),智能卡和闪存器件(例如,EPROM(Erasable Programmable Read-Only Memory,可擦写可编程只读存储器)、卡、棒或钥匙驱动器等)。另外,本文描述的各种存储介质可代表用于存储信息的一个或多个设备和/或其它机器可读介质。术语“机器可读介质”可包括但不限于,无线信道和能够存储、包含和/或承载指令和/或数据的各种其它介质。

[0035] 图1是根据本发明实施例的数据处理的方法100的示意性流程图。如图1所示,该方法100包括:

[0036] S110,根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,该目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,该第一数据集合包括 M_1 个第一数据,该 M_1 个第一数据与从该 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq$

$2, k \geq 1, T_1 > j \geq 1$;

[0037] 根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,该 N 个预存数据与该 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从该 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;

[0038] S120,根据该第一数据集合和该第二数据集合,从该 N 个预存数据中确定第一预存数据,该第一预存数据和该目标数据之间相似度与该相似度阈值 j 相对应,其中,该第一预存数据所对应的第二数据集合与该第一数据集合之间包括至少一个相同的数据。

[0039] 本发明实施例的数据处理的方法可以应用于从多个预存数据中查找与选定的目标数据之间的相似度满足预设要求的数据的过程。

[0040] 在本发明实施例中,“数据”可以包括至少两个比特组,每个比特组包括至少一个比特。

[0041] 作为示例而非限定,在本发明实施例中,可以将一个字节(Byte)作为一个比特组,即,此情况下,一个比特组包括8个比特。

[0042] 可选地,该目标数据和该预存数据为指纹数据。

[0043] 具体地说,本发明实施例的数据处理的方法可以应用于相似指纹数据的查找过程,即,可以建立指纹数据库,该指纹数据库包括一个或多个预先存储的指纹数据(即, N 个预存数据的一例),以下,为了便于理解和区分,称为“预存指纹数据”,基于本发明实施例的数据处理的方法,能够从该指纹数据库中查找与所选定的目标指纹数据(即,目标数据的一例)之间的相似度满足预设的相似度要求的预存指纹数据(即,第一预存数据的一例)。其中,指纹数据可以是包括多个字节(例如,8Byte)的数据,每个字节能够唯一地指示一种指纹特征。

[0044] 应理解,以上列举的本发明的处理对象仅为示例性说明,本发明并未限于此,本发明的数据处理的方法可以用于针对例如,图像数据或声音数据等各种数据的相似数据查找过程。

[0045] 在本发明实施例中,判定预存数据和目标数据之间的相似度是否相似度要求的过程,可以表示为判定预存数据和目标数据中相同的比特组(例如,字节)的数量是否大于或等于预设值(即,相似度阈值)的过程。

[0046] 该相似度阈值可以由系统规定,也可以是用户输入的数值,本发明并未特别限定。

[0047] 为了便于说明,以下,不失一般性,以指纹数据作为本发明实施例的数据处理的方法的处理对象,对本发明实施例的数据处理的方法进行详细说明。

[0048] 作为示例而非限定,假设指纹数据的包括8个比特组(例如,8个字节),相似度阈值为6,即,如果两个指纹数据中有6个字节相同,则可以确定两个指纹数据相似。

[0049] 例如,如图2所示,指纹数据FP#A和指纹数据FP#B分别包括8个字节。

[0050] 假设,FP#A所包括的8个字节依次为a、b、c、d、e、f、g、h。

[0051] FP#B所包括的8个字节依次为d、b、p、c、a、q、e、f。

[0052] 应理解,以上列举的指纹数据所包括的字节的数量和字节的具体值仅为示例性说明,本发明并未特别限定,例如FP#A和FP#B所包括的字节的数量也可以不同。

[0053] 在图2所示示例中,由于FP#A和FP#B所包括的相同的字节(即,a、b、c、d、e、f)的数量为6,等于预设的相似度阈值,因此,FP#A和FP#B是相似的。

[0054] 在本发明实施例中,具体是通过判定FP#A和FP#B是否具有相同的子数据来确定是FP#A和FP#B是否相似。

[0055] 下面,对“子数据”的概念和生成方式进行示例性说明。

[0056] 这里,一个子数据所包括的比特组(例如,字节)的数量为相似度阈值,并且,一个子数据所包括的比特组均来自父数据。

[0057] 例如,在本发明实施例中,可以根据该相似度阈值,确定指纹数据FP#A(即,父数据的一例)的多个子数据(即,第一数据集合的一例),具体地说,在本发明实施例中,可以将指纹数据FP#A所包括的8(即, T_1 的一例)个字节中选择6(即,相似度阈值 j 的一例)个字节的多种(例如, M_1 种)组和方式的数据作为指纹数据FP#A的子数据(即,第一数据的一例)。

[0058] 类似地,可以根据该相似度阈值,确定指纹数据FP#B(即,父数据的另一例)的多个子数据(即,第二数据集合的一例),具体地说,在本发明实施例中,可以将指纹数据FP#B所包括的8(即, T_2 的一例)个字节中选择6(即,相似度阈值 j 的一例)个字节的多种(例如, M_2 种)组和方式的数据作为指纹数据FP#B的子数据(即,第二数据的一例)。

[0059] 即,在本发明实施例中,目标数据可以作为第一数据集合中的各第一数据的父数据,第一数据集合中的各第一数据可以作为目标数据的子数据。并且,预存数据可以作为第二数据集合中的各第二数据的父数据,第二数据集合中的各第二数据可以作为目标数据的子数据。

[0060] 下面,对基于父数据所生成的子数据的方法和过程进行说明。

[0061] 在本发明实施例中,两个数据相同是指两个数据所包括的每个相同位置(例如,字节位置)上的比特组(例如,字节)均相同。

[0062] 例如,如图2所示,FP#A与FP#B中相同的字节为a、b、c、d、e、f。

[0063] 将FP#A的 M_1 个子数据中包括上述字节(a、b、c、d、e、f)的子数据记做:子数据#1,其中,该子数据#1可能是一个,也可能是多个,将子数据#1中上述6个字节(a、b、c、d、e、f)的排列记做:排列#1,其中,该排列#1可能是一个,也可能是多个。

[0064] 并且,将FP#B的 M_2 个子数据中包括上述字节(a、b、c、d、e、f)的子数据记做:子数据#2,其中,该子数据#2可能是一个,也可能是多个,将子数据#2中上述6个字节(a、b、c、d、e、f)的排列记做:排列#2,其中,该排列#2可能是一个,也可能是多个。

[0065] 则,如果排列#1与排列#2中存在相同的排列,则可以确定FP#A和FP#B包括相同的子数据,即,FP#A与FP#B相似;

[0066] 如果排列#1与排列#2中不存在相同的排列,则可以确定FP#A和FP#B不包括相同的子数据,即,FP#A与FP#B不相似。

[0067] 作为示例而非限定,在本发明的一种实施方式中,一个子数据所包括的各字节之间在该子数据中的排列顺序(或者说,位置关系)与该子数据所包括的各字节之间在该子数据的父数据中的排列顺序(或者说,位置关系)相对应,例如,相同。

[0068] 此情况下,如图1所示,由于上述排列#1为:a→b→c→d→e→f。上述排列#2为:d→b→c→a→e→f。即,排列#1与排列#2不同,因此,确定为FP#A与FP#B不相似。

[0069] 但是,如果FP#B所包括的8个字节依次为a、b、c、d、e、f、p、q。则上述排列#1为:a→b

→c→d→e→f。上述排列#2为:d→b→c→a→e→f。即,排列#1与排列#2相同,因此,确定为FP#A与FP#B相似。

[0070] 由此可见,在各字节之间的在子数据和父数据中的排列顺序相对应(例如,相同)时,目标数据和预存数据中各字节的排列顺序,对本发明的处理结果能够产生不同影响。

[0071] 对此,在本发明实施例中,为了确保处理结果的可靠性,可以采用方式1进行处理,即,可以对目标数据和预存数据中的个比特组进行排序处理,此情况下,该 $M_1=C(T_1, j)$, $M_2=C(T_2, j)$;或者,也可以采用方式2进行处理,即,确定目标数据中的各比特组的所有可能的排列方式,以及预存数据中各比特组的所有可能的排列方式,并使所生产的子数据对应上述各排列方式,此情况下, $M_1=A(T_1, j)$, $M_2=A(T_2, j)$ 。下面,分别对以上两种方式的处理进行详细说明。

[0072] 方式1

[0073] 可选地,在根据预设的相似度阈值j和目标数据,生成第一数据集合之前,所述方法还包括:

[0074] 根据预设规则,对所述目标数据中的 T_1 个第一比特组进行排序;以及

[0075] 在根据预设的相似度阈值j和目标数据,生成第一数据集合之前,所述方法还包括:

[0076] 根据所述预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

[0077] 具体地说,在本发明实施例中,在根据目标数据和预设数据(即,父数据),确定第一数据集合和第二数据集合(即,子数据)之前,可以按照预设规则,对目标数据和预设数据进行排序处理,从而,能够确保目标数据和预设数据中相同的各字节彼此之间,在目标数据和预设数据中的排列顺序相同,即,能够确保包括该目标数据和预设数据之间相同的各字节的第一数据和第二数据中,各字节的排列顺序相同,进而能够确保处理结果的可靠性。

[0078] 需要说明的是,上述预设规则可以根据需要任意确定,只要能够确保对目标数据和预设数据进行排序处理时使用的规则一致即可,例如,可以按照字节所对应的数值的大小关系,按由小到大或由大到小的顺序,对目标数据和预设数据进行排序。

[0079] 作为示例而非限定,如图2所示,FP#A所包括的8个字节依次为a、b、c、d、e、f、g、h。设a、b、c、d、e、f、g、h之间的大小关系(例如,可以是二进制的字节所对应的十进制的数值之间的大小关系)为 $a>b>c>d>e>f>g>h$,则如图3所示,按照由大到小的顺序(即,预设规则的一例)进行排序处理后的FP#A所包括的8个字节依次为a、b、c、d、e、f、g、h。

[0080] 类似的,如图2所示,FP#B所包括的8个字节依次为FP#B所包括的8个字节依次为d、b、p、c、a、q、e、f。设d、b、p、c、a、q、e、f之间的大小关系(例如,可以是二进制的字节所对应的十进制的数值之间的大小关系)为 $a>b>c>d>e>f>p>q$,则如图3所示,按照由大到小的顺序(即,预设规则的一例)进行排序处理后的FP#A所包括的8个字节依次为a、b、c、d、e、f、p、q。

[0081] 从而,如图3所示,上述排列#1为:a→b→c→d→e→f。上述排列#2为:a→b→c→d→e→f。即,排列#1与排列#2相同,因此,确定为FP#A与FP#B相似,进而能够确保判定结果的可靠性。

[0082] 在方式1下,在根据目标数据确定第一数据集合时,可以使第一数据(即,子数据)中的各第一比特组(例如,字节)的排列顺序与该第一比特组在目标数据(即,父数据)中的

排列顺序一致。并且,第一数据的数量 M_1 可以为从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的所有组合方式的数量,即 $M_1=C(T_1, j)$ 。

[0083] 类似地,在方式1下,在根据预存数据确定第二数据集合时,可以使第二数据(即,子数据)中的各第二比特组(例如,字节)的排列顺序与该第二比特组在目标数据(即,父数据)中的排列顺序一致。并且,第一数据的数量 M_2 可以为从包括 T_2 个第二比特组的目标数据中选择 j 个第二比特组时的所有组合方式的数量,即 $M_2=C(T_2, j)$ 。

[0084] 通过在生成通过基于相同的预设规则对目标数据和预存数据中的各比特组进行排序,能够确保所确定的第一数据集合和第二数据集合中包括相同的比特组的数据中各比特组的位置也相同,并且,经过排序处理后,能够使所生成的第一数据集合所包括的第一数据的数量为 $M_1=C(T_1, j)$,能够使所生成的第二数据集合所包括的第二数据的数量为 $M_2=C(T_2, j)$,从而,能够在确保相似数据查找的可靠性和准确性的前提下,减少所需要对比为数据的数量,即能够进一步降低相似数据查找的复杂度,减少相似数据查找的处理时间。

[0085] 方式2

[0086] 可选地,该根据预设的相似度阈值 j 和目标数据,生成第一数据集合,包括:根据目标数据生成 M_3 个子目标数据,其中,该 M_3 个子目标数据与 T_1 个第一比特组的所有可能的排列方式一一对应;根据预设的相似度阈值 j 和该 M_3 个子目标数据,生成第一数据集合,其中,该 M_1 个第一数据与从该 M_3 个子目标数据中的每个子目标数据中选取 j 个第一比特组时的 M_1 种组合方式一一对应;以及,根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,包括:根据第 i 个预存数据生成 M_4 个子预存数据,其中,该 M_4 个子目标数据与第 i 个预存数据的 T_2 个第二比特组的所有可能的排列方式一一对应;根据预设的相似度阈值 j 和每个预存数据所对应的 M_4 个子预存数据,生成第二数据集合,其中,第 i 个第二数据集合中的 M_2 个第二数据与从该第 i 个预存数据所对应的 M_4 个子预存数据中的每个子预存数据中选取 j 个第二比特组时的 M_2 种组合方式一一对应。

[0087] 具体地说,在本发明实施例中,可以确定目标数据所包括的各第一比特组之间所有可能的排列方式,并且,以子数据中的各比特组(例如,字节)的排列顺序与该比特组在父数据中的排列顺序一致的方式,分别确定针对每种可能的排列方式下从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的所有组合方式的数量,即 $M_1=A(T_1, j)$ 。

[0088] 类似地,可以确定预设数据所包括的各第二比特组之间所有可能的排列方式,并且,以子数据中的各比特组(例如,字节)的排列顺序与该比特组在父数据中的排列顺序一致的方式,分别确定针对每种可能的排列方式下从包括 T_2 个第二比特组的目标数据中选择 j 个第二比特组时的所有组合方式的数量,即 $M_2=A(T_2, j)$ 。

[0089] 以下表1示出了包括8个字节的指纹数据中各字节(A1~A8)之间所有可能的排列方式。

[0090] 表1

[0091]

指纹数据中字节位置	位置 #1	位置 #2	位置 #3	位置 #4	位置 #5	位置 #6	位置 #7	位置 #8
排列方式#1	A1	A2	A3	A4	A5	A6	A7	A8
排列方式#2	A2	A1	A3	A4	A5	A6	A7	A8
排列方式#3	A2	A3	A1	A4	A5	A6	A7	A8
排列方式#4	A2	A3	A4	A1	A5	A6	A7	A8
排列方式#5	A2	A3	A4	A5	A1	A6	A7	A8
排列方式#6	A2	A3	A4	A5	A6	A1	A7	A8
排列方式#7	A2	A3	A4	A5	A6	A7	A1	A8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
排列方式#8!	A8	A7	A6	A5	A4	A3	A2	A1

[0092] 如图2所示,FP#A与FP#B中相同的字节为a、b、c、d、e、f。根据方式2的处理,能够确保第一数据集合包括a、b、c、d、e、f之间所有可能的排列方式,并且,能够确保第二数据集合包括a、b、c、d、e、f之间所有可能的排列方式,从而能够确保根据本发明的数据处理的方法所判定的,FP#A与FP#B之间的关系为相似,从而,能够提高的相似数据查找的可靠性和准确性。

[0093] 需要说明的是,在上述方式2中,在如上所述确定了第一数据集合和第二数据集合之后,还可以按照预设规则(例如,按照由大到小顺序)对第一数据集合和第二数据集合中的各数据中的比特组(例如,字节)进行排序,并且,对于经过上述排序处理后的第一数据集合中发生重复的数据,可以仅保留一个,类似地,对于经过上述排序处理后的第二数据集合中发生重复的数据,也可以仅保留一个,从而,经过上述排序处理后,能够使第一数据集合所包括的第一数据的数量 M_1 从 $M_1=A(T_1, j)$ 下降至 $M_1=C(T_1, j)$,并且,使第二数据集合所包括的第二数据的数量 M_2 从 $M_2=A(T_2, j)$ 下降至 $M_2=C(T_2, j)$ 。

[0094] 根据本发明实施例的数据处理的方法,通过确定目标数据中的各第一比特组所有可能的排列方式,并确定各预存数据中的各第二比特组的所有可能的排列方式,从而能够使该第一数据集合中的第一数据对应目标数据的各第一比特组的所有可能排列方式下从该 T_1 个第一比特组中选取 j 个第一比特组时的组合方式,使每个第二数据集合中的第二数据对应所对应的预存数据的各第二比特组的所有可能排列方式下从该 T_2 个第二比特组中选取 j 个第二比特组时的组合方式,从而,能够提高的相似数据查找的可靠性和准确性。

[0095] 应理解,以上列举的方式1和方式2仅为确定第一数据集合和第二数据集合的示例性说明,本发明并未限于此,例如,在目标数据和预存数据所包括的相同的比特组之间的排序天然相同的情况下(例如,目标数据中的各第一比特组天然按某种顺序排列,且预设数

据中的各第二比特组也天然按该顺序排列),确保该 M_1 个第一数据与从该 T_1 个第一比特组中选取 j 个第一比特组时的所有组合方式一一对应即可,即, $M_1=C(T_1, j)$;并且,确保该 M_2 个第二数据与从该 T_2 个第二比特组中选取 j 个第二比特组时的所有组合方式一一对应即可,即, $M_2=C(T_2, j)$ 。

[0096] 可选地,该根据预设的相似度阈值 j 和目标数据,生成第一数据集合,包括:

[0097] 根据预设的相似度阈值 j 和该第一比特组的数量 T_1 ,确定第一生成矩阵,该第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,该 T_1 个元素包括 j 个“1”元素和 T_1-j 个“0”元素,任意两个第一向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0098] 根据该第一生成矩阵和目标数据,生成第一数据集合。

[0099] 并且,可选地,该根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,包括:

[0100] 根据预设的相似度阈值 j 和该第二比特组的数量 T_2 ,确定第二生成矩阵,该第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,该 T_2 个元素包括 j 个“1”元素和 T_2-j 个“0”元素,任意两个第二向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0101] 根据该第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

[0102] 具体地说,在本发明实施例中,可以使子数据中的各字节之间的排列顺序与该字节在父数据中的排列顺序一致,作为示例而非限定,以下表2示出了从包括8个字节的父数据中选择6个字节时所有组合方式的子数据所包括的父数据中的字节。

[0103] 表2

[0104]

父数据中的字节位置	位置 #1	位置 #2	位置 #3	位置 #4	位置 #5	位置 #6	位置 #7	位置 #8
子数据#1	0	0	1	1	1	1	1	1
子数据#2	0	1	0	1	1	1	1	1
子数据#3	0	1	1	0	1	1	1	1
子数据#4	0	1	1	1	0	1	1	1
子数据#5	0	1	1	1	1	0	1	1
子数据#6	0	1	1	1	1	1	0	1
子数据#7	0	1	1	1	1	1	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

[0105]

子数据# $C(8,6)$	1	1	1	1	1	1	0	0
---------------	---	---	---	---	---	---	---	---

[0106] 表2中的“0”表示该“0”所处行的子数据不包括父数据中该“0”所处列的位置上的字节,表2中的“1”表示该“1”所处行的子数据包括父数据中该“1”所处列的位置上的字节。

[0107] 如图4所示,在本发明实施例中,可以根据父数据所包括的字节的数量M和相似度阈值j确定生成矩阵,该生成矩阵由“0”元素和“1”元素构成。

[0108] 作为示例而非限定,该生成矩阵中的列(第一维度方向的一例)数为该父数据所包括的字节数M,该生成矩阵中的行(第二维度方向的一例)数为从该M个字节中选择j个字节的所有可能的方式的数量,即, $C(M, j)$,其中,任意两个行之间至少存在一个在行方向上的排列位置不同的“0”元素或“1”元素。

[0109] 作为示例而非限定,例如,在 $M=8, j=6$ 时,该生成矩阵可以表示为:

[0110]

11111100
11111010
11110110
11101110
11011110
10111110
01111110
⋮
01111101
01111011
01110111
01101111
01011111
00111111

[0111] 并且,在本发明实施例中,可以将父数据视为一维向量,该一维向量包括延上述“行”方向(第二维度方向的一例)排列的M个元素(即,M个字符)。

[0112] 从而,将该生成矩阵与父指纹数进行乘处理(或者说,与处理)后,能够得到 $C(M, j)$ 个子数据。

[0113] 需要说明的是,在本发明实施例中,该父数据可以是进行排序处理之后的目标数据,此情况下该生成矩阵可以作为上述第一生成矩阵, $M=T_1$,子数据的数量为 $M_1=C(T_1, j)$ 。

[0114] 并且,该父数据可以是进行排序处理之后的每个预存数据,此情况下该生成矩阵可以作为上述第二生成矩阵, $M=T_2$,每个预存数据的子数据的数量为 $M_2=C(T_2, j)$ 。

[0115] 或者,该父数据可以是目标数据的 T_1 个第一比特组的所有排列方式的数据,并且,该父数据也可以是每个预存数据的 T_2 个第一比特组的所有排列方式的数据。

[0116] 由此,能够确定第一数据集合以及N个第二数据集合。

[0117] 图5是根据本发明实施例的相似数据查找方案的示意图,如图5所示,在如上所示确定的第一数据集合和N个第二数据集合之后,可以确定第一数据集合和第i个第二数据集合之间是否存在相同的数据,如果存在,则可以确定目标数据和第i个预存数据相似;如果不存在,则可以确定目标数据和第i个预存数据不相似。

[0118] 作为示例而非限定,在本发明实施例中,可以将N个第二数据集合中的各第二数据

保存至哈希表,具体地说,是哈希表各行(bucket)。

[0119] 图6是根据本发明实施例的哈希表的维护方式的示意图。如图6所示,在本发明实施例中,在需要将第*i*个预存数据所对应的第*i*个第二数据集合中的各第二数据插入哈希表中时,可以计算各第二数据的哈希值,并根据所确定的哈希值,将该第*i*个预存数据所对应的各第二数据保存至哈希表中与各哈希值相对应的索引位置。

[0120] 另外,在需要将第*i*个预存数据所对应的第*i*个第二数据集合中的各第二数据从哈希表中删除时,可以计算各第二数据的哈希值,并根据所确定的哈希值,将哈希表中与各哈希值相对应的索引位置的数据(即,第*i*个第二数据集合中的各第二数据)删除。

[0121] 需要说明的是,在本发明实施例中,*N*个预存数据中可能存在多个与目标数据相似的数据,即,*N*个预存数据中可能存在多个包括目标数据中的*j*个第一比特组的数据,此情况下,哈希表中的某些bucket中可能存在多个数据,即,哈希表中的同一索引位置上可能需要保存多个数据。

[0122] 此情况下,本发明实施例可以提供多种哈希表的bucket结构。

[0123] 例如,如果需要保存在同一bucket内的数据的数量小于或等于预设的数量阈值(例如,5),则可以将需要保存在同一bucket内的数据组成数据链,进行保存。即,如果当哈希表中填充比例不高,各个bucket中元素个数均衡。直接将元素组成链。

[0124] 再例如,如果需要保存在同一bucket内的数据的数量大于预设的数量阈值(例如,5),则可以将需要保存在同一bucket内的数据组成制高点(VP, Vantage Point)树,进行保存。即,当哈希表总填充比例超过预设值,各个bucket中元素个数不均衡,将元素按照VP树结构放置。

[0125] 再例如,如果需要保存在同一bucket内的数据的数量大于预设的数量阈值(例如,5),则可以将需要保存在同一bucket内的数据组成套嵌的哈希表,进行保存。即,当哈希表总填充比例超过预设值,各个bucket中元素个数不均衡,将元素按照套嵌的哈希表放置。

[0126] 需要说明的是,在本发明实施例中,哈希表的各bucket的结构相同,即,如果某一bucket采用链结构,则其他bucket也采用链结构;如果某一bucket采用树结构,则其他bucket也采用树结构;如果某一bucket采用哈希表结构,则其他bucket也采用哈希表结构。

[0127] 另外,在本发明实施例中,哈希表的各bucket的所采用的结构可以基于同一bucket内的数据的数量确定,也可以基于系统要求或管理员的设置来确定,本发明并未特别限定。

[0128] 图7是根据本发明实施例的哈希表的一例的示意图。如图7所示,假设预设数据包括指纹数据FP(B)和指纹数据FP(C),本发明实施例的哈希表的每个bucket具有预设的索引(即,哈希值),在对FP(B)和FP(C)的各子数据进行哈希值计算而确定各子数据的哈希值之后,可以将各子数据保存至哈希表中索引值相同的bucket中,并且,可以将各子数据的父数据的指示信息,例如,指示该子数据的父数据为FP(B)或FP(C)的信息一同保存至哈希表中。

[0129] 从而,在查找目标数据(例如,指纹数据FP(A))的相似数据时,可以按照如上所示方式确定FP(A)的各子数据,并确定各子数据的哈希值,并将哈希表中索引的值与FP(A)的任一子数据的哈希值相同的bucket中存储的数据的父数据,作为FP(A)的相似数据。

[0130] 在本发明实施例中,针对一个目标数据,在*N*个预设数据中可能存在多个相似的数据。对此,本发明提供了多种查找方案。

[0131] 例如,本发明实施例的查找方案可以包括快速模式和完整模式。下面,分别对上述两种模式下的查找过程进行详细说明。

[0132] A. 快速模式

[0133] 可选地,该根据该第一数据集合和该第二数据集合,从该N个预存数据中确定第一预存数据,包括:

[0134] 从备选模式中确定查找模式,该备选模式包括快速模式和完整模式;

[0135] 当该查找模式为该快速模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行y次查找处理,该y次查找处理与该第一数据集合中基于该预设顺序的前y个第一数据一一对应,其中,第x次查找处理用于确定该N个第二数据集合中是否存在与第x个第一数据相同的数据,并在首次查找到与该第一数据相同的数据时,将与该第一数据相同的数据所属于的第二数据集合所对应的预存数据作为该第一预存数据,其中, $x \in [1, y]$, $y \leq M_1$ 。

[0136] 具体地说,图8是根据本发明实施例的快速模式下的查找过程的示意性流程图,如图8所示,首先,可以确定目标指纹的各子数据(即,第一数据);其后,可以令 $i=1$,并确定第i个子数据的哈希值;其后,可以判定在哈希表中索引的值与该第i个子数据的哈希值相同的bucket中是否保存有数据。

[0137] 如果判定为是,则可以将该bucket中所保存的数据的父数据作为与该模板指纹相似的数据(即,第一预存数据),并立即,结束查找。

[0138] 如果判定为否,则可以令 $i=i+1$,并继续查找,如果 $i=M_1$ 时仍然未查找到相似的数据,则可以认为N个预存数据中不存在与该目标数据相似的数据。

[0139] 从而,能够快速地查找到满足预设的相似条件(即,与相似度阈值j相对应)的预设数据。

[0140] B. 完整模式

[0141] 可选地,该根据该第一数据集合和该第二数据集合,从该N个预存数据中确定第一预存数据,包括:

[0142] 从备选模式中确定查找模式,该备选模式包括快速模式和完整模式;

[0143] 当该查找模式为该完整模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行y次查找处理,该y次查找处理与该第一数据集合中基于该预设顺序的前y个第一数据一一对应,其中,第x次查找处理用于确定该N个第二数据集合中是否存在与第x个第一数据相同的数据,如果首次查找到的与该第一数据相同的数据属于至少两个第二数据集合,则将至少两个第二数据集合所对应的预存数据中与该目标数据之间存在相同的比特组的数量最多的预存数据作为该第一预存数据。

[0144] 具体地说,在上述快速模式的基础上,如果哈希表中索引的值与该第i个子数据的哈希值相同的bucket中保存有两个或两个以上的数据,则可以进一步判定该两个或两个以上的数据的父数据中与该目标数据所包括的相同的比特组最多的数据,作为该第一预设数据。

[0145] 应理解,以上列举的查找模式仅为示例性说明,本发明并未限于此,例如,还可以采用以下模式进行查找:

[0146] 依次进行针对目标数据的 M_1 个子数据中的每个子数据的查找处理,即,确定第i个

子数据的哈希值,其后,可以判定在哈希表中索引的值与该第 i 个子数据的哈希值相同的bucket中是否保存有数据。如果判定为是,则可以将该bucket中所保存的数据的父数据作为与该模板指纹相似的数据(即,第一预存数据),并保存该数据。

[0147] 并且,在查找到多个满足预设的相似条件(即,与相似度阈值 j 相对应)的预设数据后,可以进一步确定该多个预设数据中与该目标数据之间存在的相同的比特组最多的数据,作为该第一预设数据,从而能够得到 N 个预存数据中与目标数据最相似的数据。

[0148] 需要说明的是,上述查找模式的使用可以根据本发明实施例的数据处理方法的执行主体(例如,处理器)当前的负载确定,例如,如果处理器的负载较低,则可以使用完整模式;如果处理器的负载较大,则可以使用快速模式。

[0149] 根据本发明实施例的数据处理的方法,通过设置快速模式和完整模式,并在快速模式下输出预存数据中所有相似度满足相似度阈值 j 所对应的要求的数据,在完整模式输出预存数据中相似度满足相似度阈值 j 所对应的要求且与目标数据之间的相似度最高的数据,能够灵活应对不同的用户需求,进一步提高用户体验。

[0150] 应理解,以上列举的查找模式的确定方式仅为示例性说明,本发明并未限定于此,例如,上述查找模式的使用可以根据用户要求确定。

[0151] 根据本发明实施例的数据处理的方法,通过根据目标数据确定包括 M_1 个第一数据的第一数据集合,并根据 N 个预存数据确定 N 个第二数据集合,其中,第一数据集合中的 M_1 个第一数据与从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的 M_1 种组合方式一一对应,每个第二数据集合中的 M_2 个第二数据与从包括 T_2 个第二比特组的预存数据中选择 j 个第二比特组时的 M_2 种组合方式一一对应,其中, j 为预设的相似度阈值,从而,在一个第二数据集合与该第一数据集合之间包括至少一个相同的数据时,能够将该第二数据集合所对应的预存数据作为与该目标数据之间的相似度满足该相似度阈值 j 所对应的相似度要求的相似数据,即,能够将相似数据查找过程转化为相同数据的判定过程,从而,能够降低相似数据查找的复杂度,减少相似数据查找的处理时间,改善用户体验。

[0152] 图9是根据本发明实施例的数据处理的装置200的示意性框图。该装置200可以通过数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等实现。该装置200也可以指示软件模块或软件包。如图9所示,该装置200包括:

[0153] 生成单元210,用于根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,该目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,该第一数据集合包括 M_1 个第一数据,该 M_1 个第一数据与从该 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2, k \geq 1, T_1 > j \geq 1$;

[0154] 所示生成单元210还用于,根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,该 N 个预存数据与该 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从该 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;

[0155] 确定单元220,用于根据该第一数据集合和该第二数据集合,从该 N 个预存数据中

确定第一预存数据,该第一预存数据和该目标数据之间相似度与该相似度阈值 j 相对应,其中,该第一预存数据所对应的第二数据集合与该第一数据集合之间包括至少一个相同的数据。

[0156] 可选地,该装置200还包括:

[0157] 排序单元230,用于根据预设规则,对该目标数据中的 T_1 个第一比特组进行排序;

[0158] 用于根据该预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

[0159] 可选地,该生成单元210具体用于根据预设的相似度阈值 j 和该第一比特组的数量 T_1 ,确定第一生成矩阵,该第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,该 T_1 个元素包括 j 个“1”元素和 $T_1 - j$ 个“0”元素,任意两个第一向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0160] 用于根据该第一生成矩阵和目标数据,生成第一数据集合。

[0161] 可选地,该生成单元210具体用于根据预设的相似度阈值 j 和该第二比特组的数量 T_2 ,确定第二生成矩阵,该第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,该 T_2 个元素包括 j 个“1”元素和 $T_2 - j$ 个“0”元素,任意两个第二向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0162] 用于根据该第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

[0163] 可选地,该确定单元220具体用于从备选模式中确定查找模式,该备选模式包括快速模式和完整模式;

[0164] 当该查找模式为该快速模式时,该确定单元具体用于基于各该第一数据在该第一数据集合中的排序,按预设顺序进行 y 次查找处理,该 y 次查找处理与该第一数据集合中基于该预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定该 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,并在首次查找到与该第一数据相同的数据时,将与该第一数据相同的数据所属于的第二数据集合所对应的预存数据作为该第一预存数据,其中, $x \in [1, y]$, $y \leq M_1$;

[0165] 当该查找模式为该完整模式时,该确定单元具体用于基于各该第一数据在该第一数据集合中的排序,按预设顺序进行 y 次查找处理,该 y 次查找处理与该第一数据集合中基于该预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定该 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,如果首次查找到的与该第一数据相同的数据属于至少两个第二数据集合,则将至少两个第二数据集合所对应的预存数据中与该目标数据之间存在相同的比特组的数量最多的预存数据作为该第一预存数据,其中, $x \in [1, y]$, $y \leq M_1$ 。

[0166] 该装置200中的各单元或模块分别用于执行上述方法100中的动作和功能,这里为了避免赘述,省略其详细说明。

[0167] 根据本发明实施例的数据处理的装置,通过根据目标数据确定包括 M_1 个第一数据的第一数据集合,并根据 N 个预存数据确定 N 个第二数据集合,其中,第一数据集合中的 M_1 个第一数据与从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的 M_1 种组合方式一一对应,每个第二数据集合中的 M_2 个第二数据与从包括 T_2 个第二比特组的预存数据中选择 j

个第二比特组时的 M_2 种组合方式一一对应,其中, j 为预设的相似度阈值,从而,在一个第二数据集合与该第一数据集合之间包括至少一个相同的数据时,能够将该第二数据集合所对应的预存数据作为与该目标数据之间的相似度满足该相似度阈值 j 所对应的相似度要求的相似数据,即,能够将相似数据查找过程转化为相同数据的判定过程,从而,能够降低相似数据查找的复杂度,减少相似数据查找的处理时间,改善用户体验。

[0168] 图10是根据本发明实施例的数据处理的设备300的示意性结构图。如图10所示,该设备300包括:

[0169] 总线310;

[0170] 与该总线310相连的存储器320;

[0171] 与该总线310相连的处理器330;

[0172] 该处理器330用于经由该总线310调用并执行该存储器320中的程序,以用于根据预设的相似度阈值 j 和目标数据,生成第一数据集合,其中,该目标数据包括 T_1 个第一比特组,每个第一比特组包括 k 比特,该第一数据集合包括 M_1 个第一数据,该 M_1 个第一数据与从该 T_1 个第一比特组中选取 j 个第一比特组时的 M_1 种组合方式一一对应, $T_1 \geq 2, k \geq 1, T_1 > j \geq 1$;

[0173] 用于根据该相似度阈值 j 和 N 个预存数据,生成 N 个第二数据集合,其中,该 N 个预存数据与该 N 个第二数据集合一一对应,每个预存数据包括 T_2 个第二比特组,每个第二比特组包括 k 比特,每个第二数据集合包括 M_2 个第二数据,第 i 个第二数据集合中的每个第二数据包括第 i 个预存数据中的 T_2 个第二比特组,第 i 个第二数据集合中的 M_2 个第二数据与从该 T_2 个第二比特组中选取 j 个第二比特组时的 M_2 种组合方式一一对应, $N \geq 1, T_2 \geq T_1, i \in [1, N]$;

[0174] 用于根据该第一数据集合和该第二数据集合,从该 N 个预存数据中确定第一预存数据,该第一预存数据和该目标数据之间相似度与该相似度阈值 j 相对应,其中,该第一预存数据所对应的第二数据集合与该第一数据集合之间包括至少一个相同的数据。

[0175] 可选地,该处理器330具体用于根据预设规则,对该目标数据中的 T_1 个第一比特组进行排序;

[0176] 用于根据该预设规则,对每个预存数据中的 T_2 个第二比特组进行排序。

[0177] 可选地,该处理器330具体用于根据预设的相似度阈值 j 和该第一比特组的数量 T_1 ,确定第一生成矩阵,该第一生成矩阵包括在第一维度方向上排列的 M_1 个第一向量,每个第一向量包括在第二维度方向上排列的 T_1 个元素,该 T_1 个元素包括 j 个“1”元素和 $T_1 - j$ 个“0”元素,任意两个第一向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0178] 用于根据该第一生成矩阵和目标数据,生成第一数据集合。

[0179] 可选地,该处理器330具体用于根据预设的相似度阈值 j 和该第二比特组的数量 T_2 ,确定第二生成矩阵,该第二生成矩阵包括在第一维度方向上排列的 M_2 个第二向量,每个第二向量包括在第二维度方向上排列的 T_2 个元素,该 T_2 个元素包括 j 个“1”元素和 $T_2 - j$ 个“0”元素,任意两个第二向量彼此之间至少存在一个在该第二维度方向上的排列位置不同的“0”元素或“1”元素;

[0180] 用于根据该第二生成矩阵和 N 个预存数据,生成 N 个第二数据集合。

[0181] 可选地,该处理器330具体用于从备选模式中确定查找模式,该备选模式包括快速模式和完整模式;

[0182] 用于当该查找模式为该快速模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行 y 次查找处理,该 y 次查找处理与该第一数据集合中基于该预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定该 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,并在首次查找到与该第一数据相同的数据时,将与该第一数据相同的数据所属于的第二数据集合所对应的预存数据作为该第一预存数据,其中, $x \in [1, y], y \leq M_1$;

[0183] 用于当该查找模式为该完整模式时,基于各该第一数据在该第一数据集合中的排序,按预设顺序进行 y 次查找处理,该 y 次查找处理与该第一数据集合中基于该预设顺序的前 y 个第一数据一一对应,其中,第 x 次查找处理用于确定该 N 个第二数据集合中是否存在与第 x 个第一数据相同的数据,如果首次查找到的与该第一数据相同的数据属于至少两个第二数据集合,则将至少两个第二数据集合所对应的预存数据中与该目标数据之间存在相同的比特组的数量最多的预存数据作为该第一预存数据,其中, $x \in [1, y], y \leq M_1$ 。

[0184] 应理解,在本发明实施例中,该处理器330可以是中央处理单元(Central Processing Unit,简称为“CPU”)。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0185] 该存储器320可以包括只读存储器和随机存取存储器,并向处理器330提供指令和数据。存储器320的一部分还可以包括非易失性随机存取存储器。例如,存储器320还可以存储设备类型的信息。

[0186] 该总线310除包括数据总线之外,还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见,在图中将各种总线都标为总线310。

[0187] 在实现过程中,上述方法的各步骤可以通过处理器330中的硬件的集成逻辑电路或者软件形式的指令完成。结合本发明实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成,或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器320,处理器330读取存储器320中的信息,结合其硬件完成上述方法的步骤。为避免重复,这里不再详细描述。

[0188] 其中,该设备300用于执行上述方法100中的动作和功能,这里为了避免赘述,省略其详细说明。

[0189] 根据本发明实施例的数据处理的设备,通过根据目标数据确定包括 M_1 个第一数据的第一数据集合,并根据 N 个预存数据确定 N 个第二数据集合,其中,第一数据集合中的 M_1 个第一数据与从包括 T_1 个第一比特组的目标数据中选择 j 个第一比特组时的 M_1 种组合方式一一对应,每个第二数据集合中的 M_2 个第二数据与从包括 T_2 个第二比特组的预存数据中选择 j 个第二比特组时的 M_2 种组合方式一一对应,其中, j 为预设的相似度阈值,从而,在一个第二数据集合与该第一数据集合之间包括至少一个相同的数据时,能够将该第二数据集合所对应的预存数据作为与该目标数据之间的相似度满足该相似度阈值 j 所对应的相似度要求的相似数据,即,能够将相似数据查找过程转化为相同数据的判定过程,从而,能够降低相似数据查找的复杂度,减少相似数据查找的处理时间,改善用户体验。

[0190] 应理解,本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。

另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。

[0191] 应理解,在本发明的各种实施例中,上述各过程的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不应对本发明实施例的实施过程构成任何限定。

[0192] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0193] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0194] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0195] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0196] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0197] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0198] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

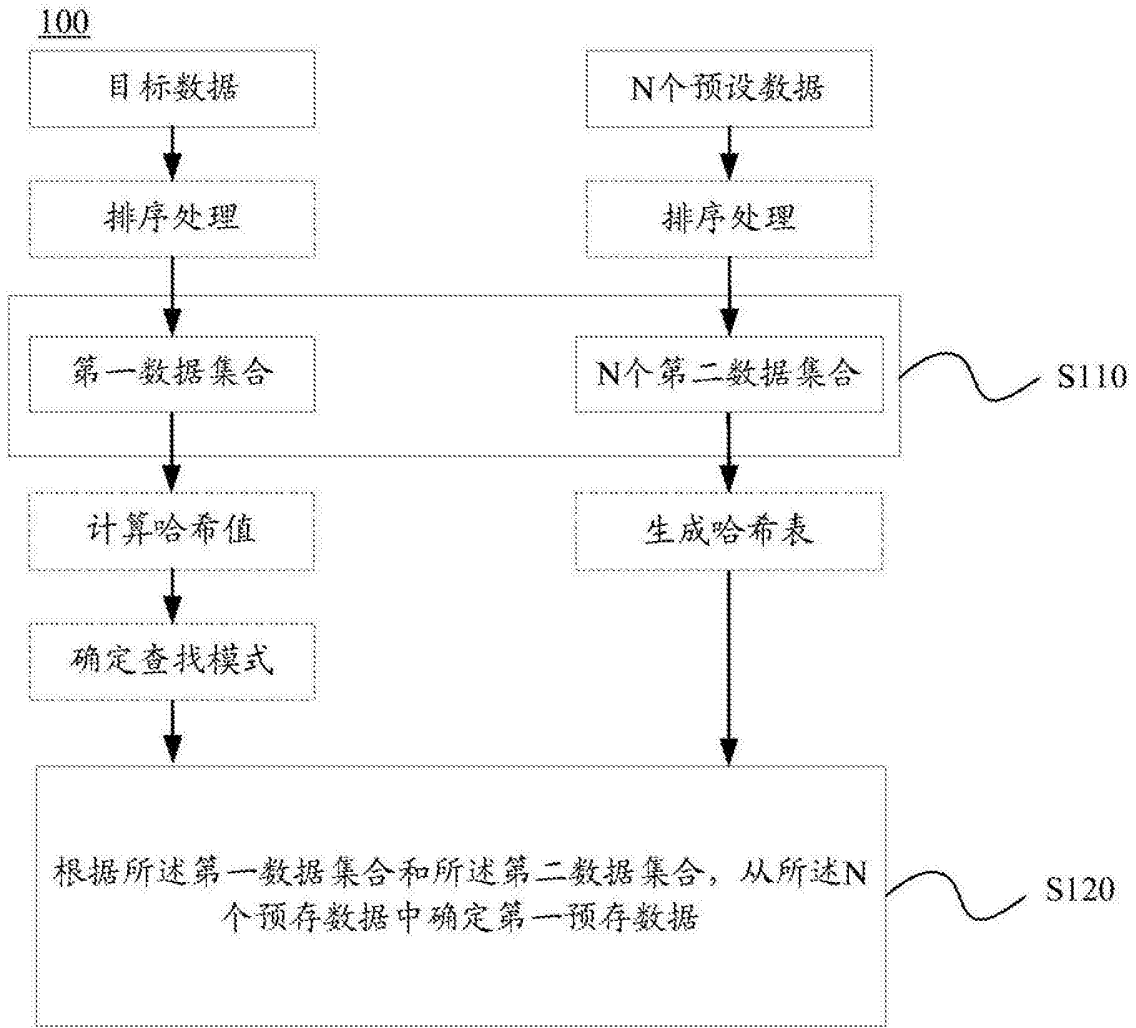


图1

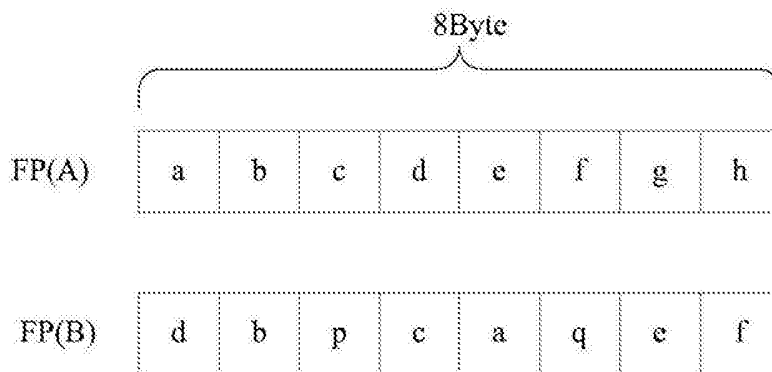


图2

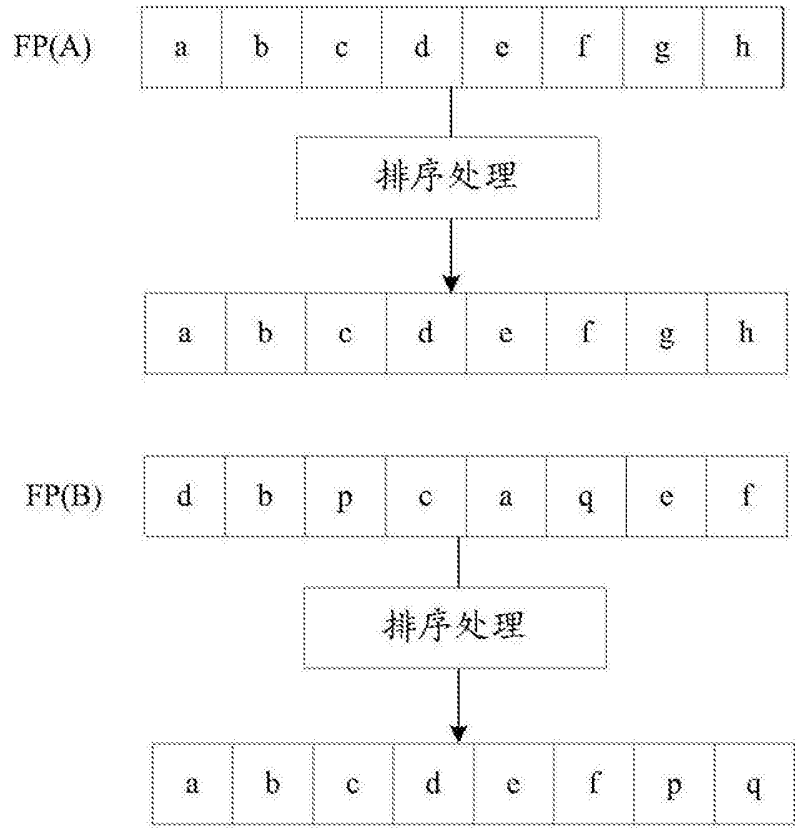


图3

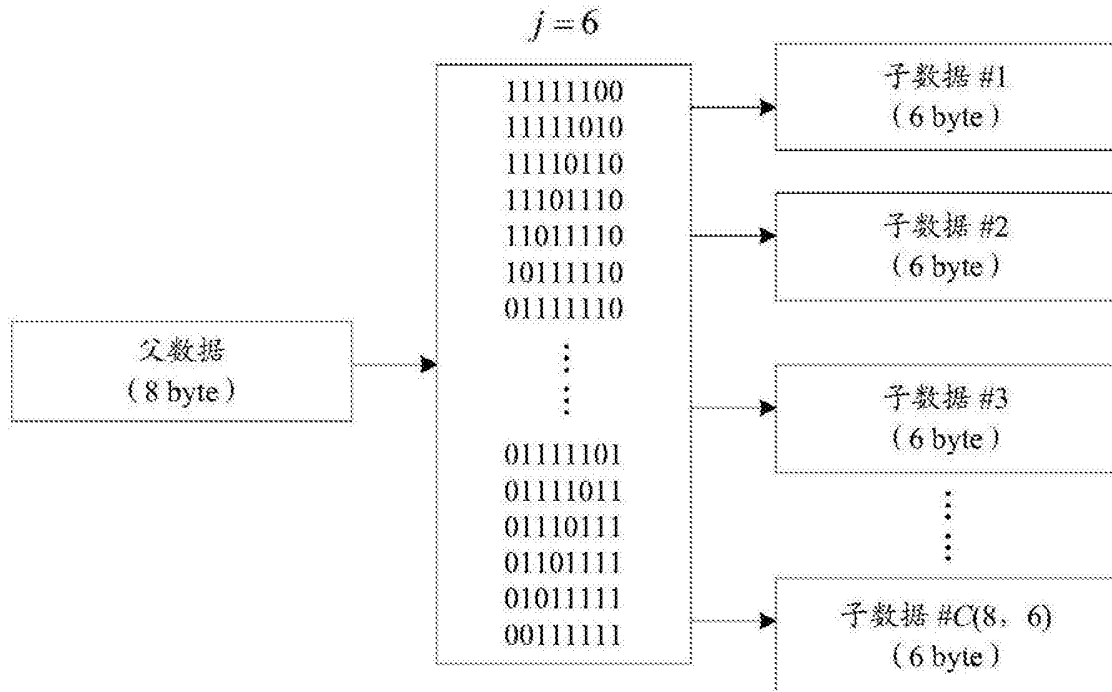


图4

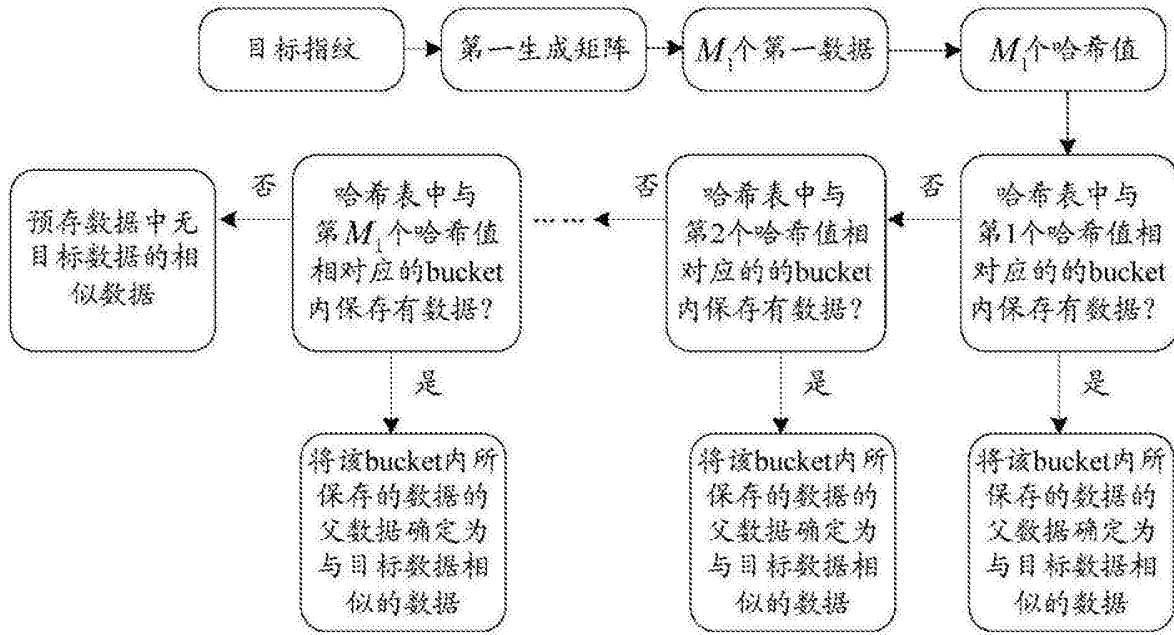


图5

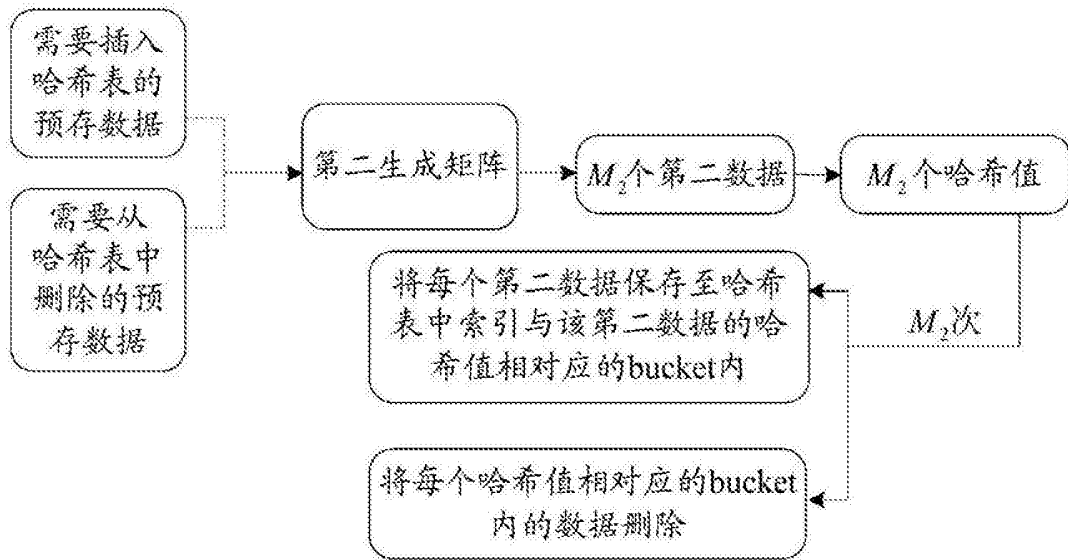


图6

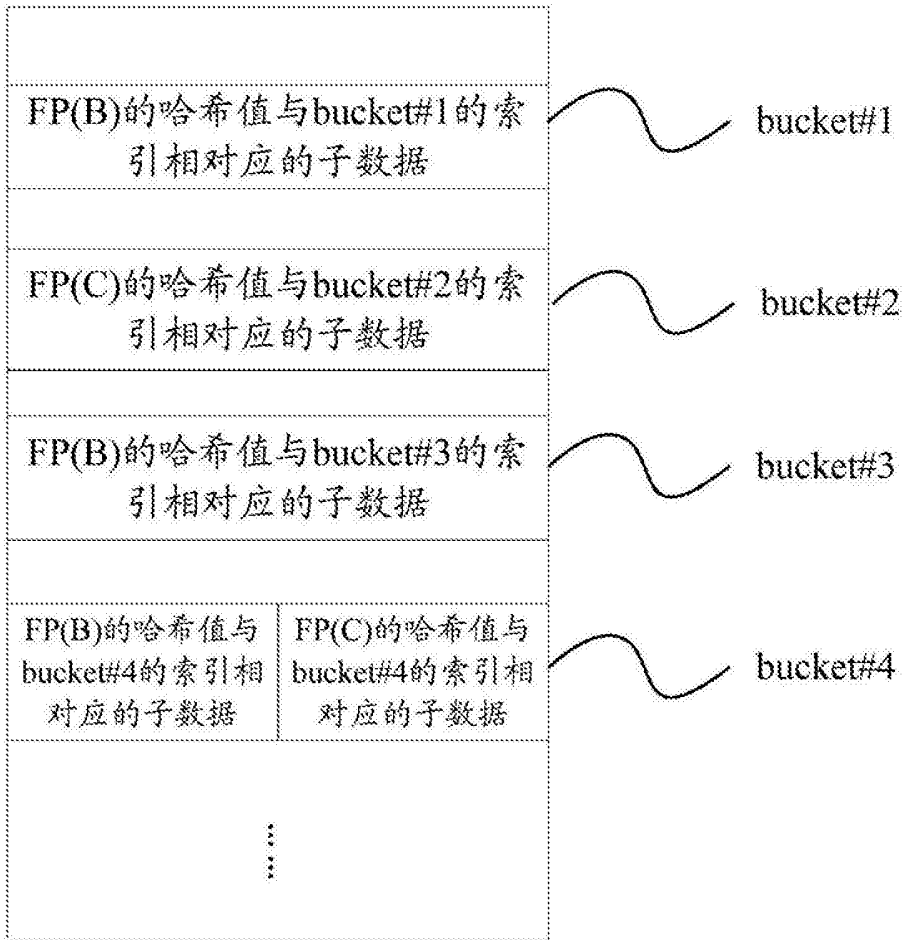


图7

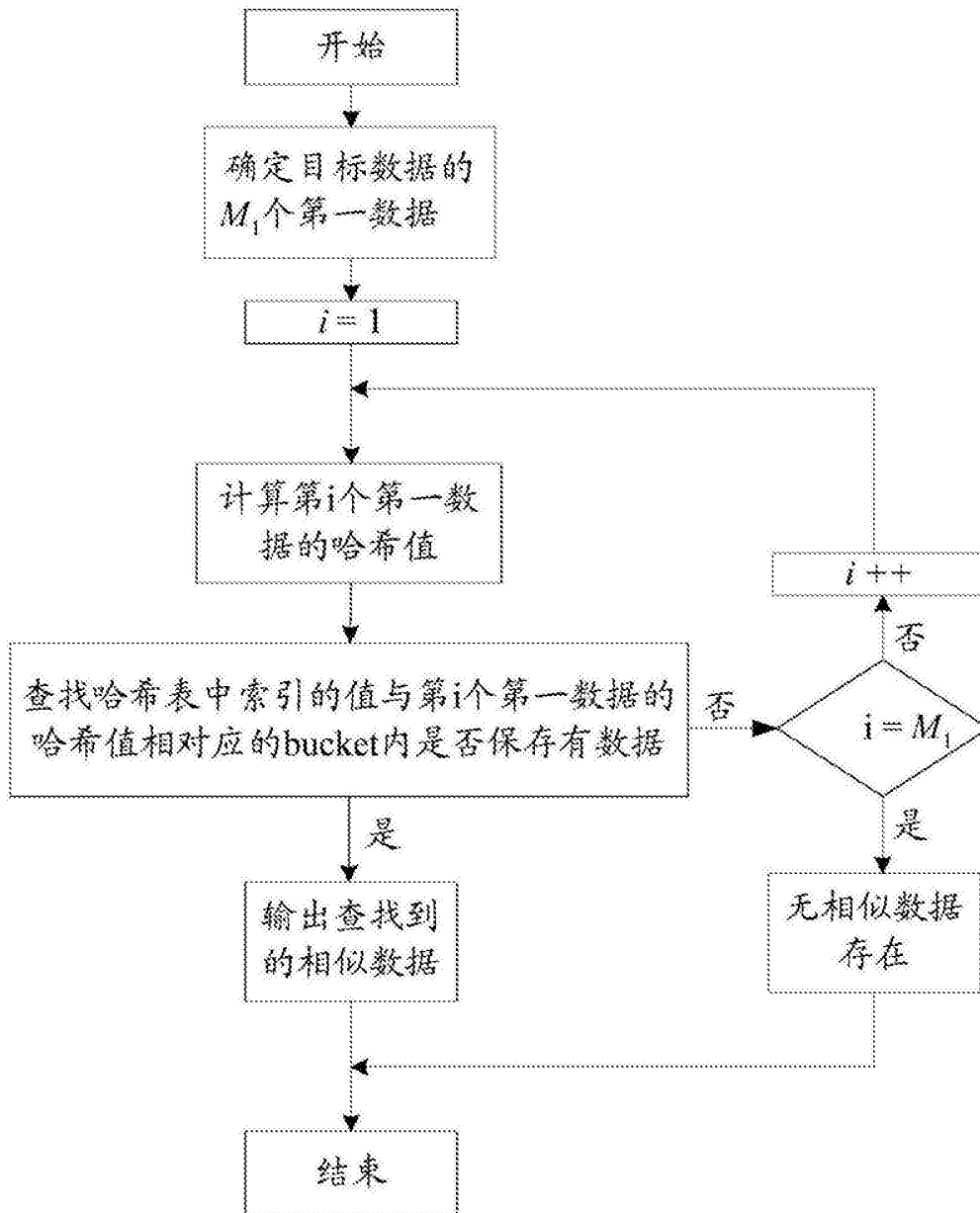


图8

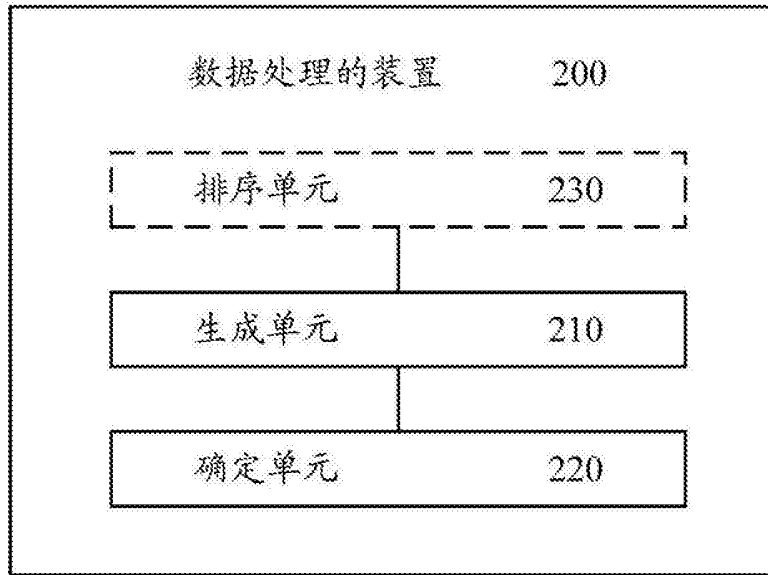


图9

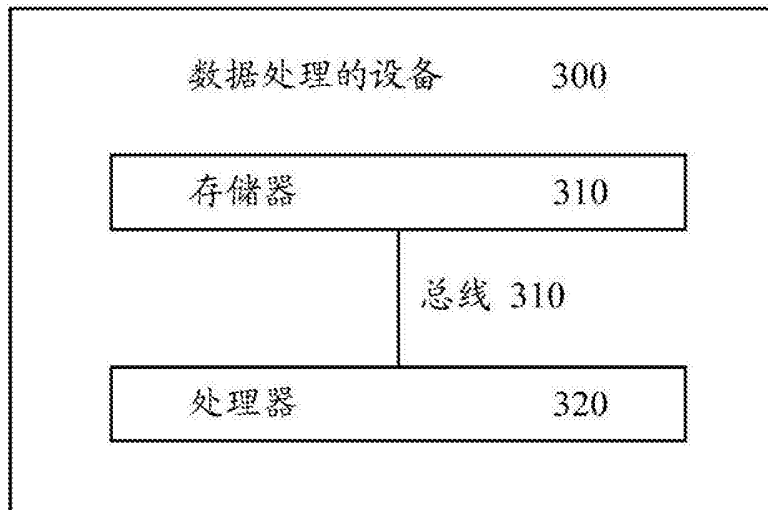


图10