



(12) 发明专利申请

(10) 申请公布号 CN 103049388 A

(43) 申请公布日 2013. 04. 17

(21) 申请号 201210519408. 0

(22) 申请日 2012. 12. 06

(71) 申请人 深圳市江波龙电子有限公司  
地址 518000 广东省深圳市南山区科发路 8 号金融服务技术创新基地 1 栋 8 楼 A、B、C、D、E、F1

(72) 发明人 郭丹 梁小庆

(74) 专利代理机构 深圳中一专利商标事务所  
44237

代理人 张全文

(51) Int. Cl.  
G06F 12/02 (2006. 01)  
G06F 11/10 (2006. 01)

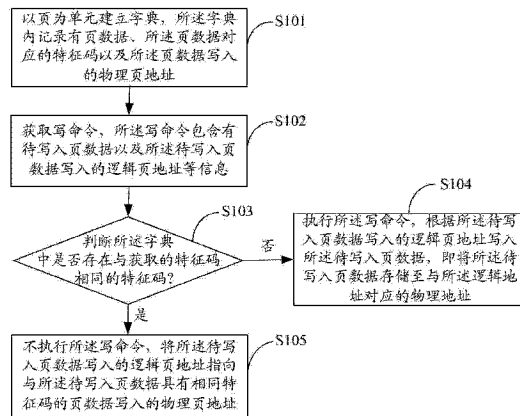
权利要求书 4 页 说明书 13 页 附图 4 页

(54) 发明名称

一种分页存储器件的压缩管理方法及装置

(57) 摘要

本发明适用于存储器技术领域,提供了一种分页存储器件的压缩管理方法及装置,所述方法包括:以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址;获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址等信息;获取待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。通过本发明可有效解决以页为最小存储单元的存储器件的管理问题。



1. 一种分页存储器件的压缩管理方法,其特征在于,所述方法包括:

步骤A、以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址;

步骤B、获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址;获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

2. 如权利要求1所述的方法,其特征在于,所述字典为静态字典,所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和/或用户关心的数据。

3. 如权利要求1所述的方法,其特征在于,所述字典为动态字典;

所述方法在获取写命令的步骤之前还包括:

设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

在所述执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据步骤之后还包括:

在所述待写入页数据的写入次数达到所述第二阈值时,将该待写入页数据更新到所述动态字典中。

4. 如权利要求1所述的方法,其特征在于,所述字典包含静态字典和动态字典;所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和/或用户关心的数据;

所述方法在获取写命令的步骤之前还包括:

设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

所述获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码具体包括:

获取所述待写入页数据的特征码,判断所述静态字典中是否存在与所获取的特征码相同的特征码,若否,根据所述待写入页数据写入分页存储器件的次数判断所述待写入页数据是否需要更新到所述动态字典。

5. 如权利要求3或4所述的方法,其特征在于,所述方法还包括:

当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中无效的条目或者在预设时间内未使用的条目,所述无效的条目为所述动态字典中存储的页数据的写入地址已经写入了其他数据,导致该页数据无效;所述预设时间内未使用的条目为该条目对应的页数据在预设时间内未写入过分页存储器件。

6. 如权利要求3或4所述的方法,其特征在于,所述方法还包括:

在所述页数据的写入次数达到所述第二阈值之后继续对所述页数据的写入次数进行计数,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中写入次数小于第四阈值的页数据。

7. 如权利要求1所述的方法,其特征在于,所述步骤B具体包括:

在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入

的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述字典中记录有页数据多次消息摘录后的特征码；

或者，在所述分页存储器件支持一种哈希算法时，对当前写命令中的待写入页数据进行多次不同偏移的消息摘录，获取每次不同偏移消息摘录后的特征码，判断所述字典中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码，若否，执行当前写命令，根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码；

或者，在所述分页存储器件支持错误检查和纠正 ECC 码时，获取当前写命令中待写入页数据的 ECC 码和消息摘录后的特征码，判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码，若否，执行当前写命令，根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码；

或者，获取当前写命令中待写入页数据的特征码，在所述字典中存在与所获取的特征码相同的特征码时，判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同，若否，执行当前写命令，根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数。

8. 如权利要求 1 所述的方法，其特征在于，所述步骤 B 具体包括：

获取当前写命令中待写入页数据的特征码，对所述特征码进行消息摘录获得二次摘录值，判断所述字典中是否存在所述二次摘录值，若存在，再判断所述字典中是否存在所述特征码，若否，执行当前写命令，根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述字典中记录有页数据对应特征码的二次摘录值。

9. 如权利要求 1 至 8 任一项所述的方法，其特征在于，所述获取所述待写入页数据的特征码具体包括：

通过哈希算法获取所述待写入页数据的特征码。

10. 一种分页存储器件的压缩管理装置，其特征在于，所述装置包括：

字典建立单元，用于以页为单元建立字典，所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址；

处理单元，用于获取写命令，所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址；获取所述待写入页数据的特征码，判断所述字典中是否存在与所获取的特征码相同的特征码；若否，执行所述写命令，根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将所述待写入页数据写入的逻辑页地址

指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

11. 如权利要求 10 所述的装置,其特征在于,所述字典为静态字典,所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和 / 或用户关心的页数据。

12. 如权利要求 10 所述的装置,其特征在于,所述字典为动态字典;

所述处理单元在获取写命令之前还用于,设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

所述处理单元在执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据之后,还用于在所述页数据的写入次数达到所述第二阈值时,将该页数据更新到所述动态字典中。

13. 如权利要求 10 所述的装置,其特征在于,所述字典包含静态字典和动态字典;所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和 / 或用户关心的数据;

所述处理单元在获取写命令之前还用于,设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

所述处理单元,用于获取所述待写入页数据的特征码,判断所述静态字典中是否存在与所获取的特征码相同的特征码,若否,根据所述待写入页数据写入分页存储器件的次数判断所述待写入页数据是否需要更新到所述动态字典。

14. 如权利要求 12 或 13 所述的装置,其特征在于,所述处理单元还用于,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中无效的条目或者在预设时间内未使用的条目,所述无效的条目为所述动态字典中存储的页数据的写入地址已经写入了其他数据,导致该页数据无效;所述预设时间内未使用的条目为该条目对应的页数据在预设时间内未写入过分页存储器件。

15. 如权利要求 12 或 13 所述的装置,其特征在于,所述处理单元还用于,在所述页数据的写入次数达到所述第二阈值之后继续对所述页数据的写入次数进行计数,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中写入次数小于第四阈值的页数据。

16. 如权利要求 10 所述的装置,其特征在于,所述处理单元具体用于,

在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次消息摘录后的特征码;

或者,在所述分页存储器件支持一种哈希算法时,对当前写命令中的待写入页数据进行多次不同偏移的消息摘录,获取每次不同偏移消息摘录后的特征码,判断所述字典中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码;

或者,在所述分页存储器件支持错误检查和纠正 ECC 码时,获取当前写命令中待写入

页数据的 ECC 码和消息摘录后的特征码,判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码;

或者,获取当前写命令中待写入页数据的特征码,在所述字典中存在与所获取的特征码相同的特征码时,判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数。

17. 如权利要求 10 所述的装置,其特征在于,所述处理单元具体用于,

获取当前写命令中待写入页数据的特征码,再对所述特征码进行消息摘录获得二次摘录值,判断所述字典中是否存在所述二次摘录值,若存在,再判断所述字典中是否存在所述特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据对应特征码的二次摘录值。

18. 如权利要求 10 至 17 任一项所述的装置,其特征在于,所述处理单元,用于通过哈希算法获取所述待写入页数据的特征码。

## 一种分页存储器件的压缩管理方法及装置

### 技术领域

[0001] 本发明属于存储器技术领域,尤其涉及一种分页存储器件的压缩管理方法及装置。

### 背景技术

[0002] 当前通用的数据压缩方法包括无损数据压缩和有损数据压缩,如基于统计模型的压缩技术和基于字典模型的压缩技术为无损数据压缩技术;多媒体数据压缩包括音频压缩、图像压缩、视频压缩等则大多为有损数据压缩技术;对存储设备而言,必须使用无损的数据压缩方式。

[0003] 现有的无损数据压缩方式包括:

[0004] 一、字典型数据压缩方法,如表 1 所示:

[0005]

Index	String
1	Prince
2	Love
3	Princess
4	Forever

[0006] 表 1

[0007] 例如需要压缩 Yes ! Prince love Princess forever 则根据表 1 提供的字典得到压缩结果为:

[0008] Yes ! &1 &2 &3 &4 (其中 & 是特殊符号,如果需要输出 & 则压缩数据为 &&)

[0009] 缺点:字典的大小和好坏直接影响压缩结果。大的字典开销较大,小的字典无法满足压缩要求,例如通过表 1 提供的字典压缩 I am happy today! 则完全无法压缩。

[0010] 二、行程编码型数据压缩方法

[0011] 行程编码的基本原理是:通过一个符号值或串长代替具有相同值的连续符号(连续符号构成了一段连续的“行程”),使符号长度少于原始数据的长度。在各行或者各列数据的代码发生变化时,一次记录该代码及相同代码重复的个数,从而实现数据的压缩。

[0012] 例如:55555577777733322221111111

[0013] 行程编码为:(5,6)(7,5)(3,3)(2,4)(1,7)。可见,行程编码的位数远远少于原始字符串的位数。

[0014] 然而,并不是所有的行程编码都远远少于原始字符串的位数,例如:555555 是 6 个字符,而(5,6)是 5 个字符,这也存在压缩量的问题。另外,行程编码针对随机数据的压缩出来的编码可能比没压缩还大。

### [0015] 三、哈夫曼型数据压缩方法

[0016] 现有的哈夫曼型数据压缩方法在进行数据压缩时需要构建哈夫曼树,空间开销过大。而且,压缩后的长度是一个变化值,不便于以页为最小存储单元的存储器件,例如对 8k 的数据进行压缩时,压缩完可能是 7.5k,对于以页为最小存储单元的存储器件来说,如果其页的大小为 8k,则压缩后剩余的 0.5k 基本无法使用。

[0017] 综上所述,现有的数据压缩方法存在以下问题:(1)对运算开销过大或者对存储空间要求过高;(2)压缩的效果无法保证;(3)压缩结果不便于以页为最小存储单元的存储器件管理。

### 发明内容

[0018] 本发明实施例的目的在于提供一种分页存储器件的压缩管理方法,以解决现有的压缩管理方法对运算开销过大或者对存储空间要求过高、压缩效果无法保证、且压缩结果不便于以页为最小存储单元的存储器件管理的问题。

[0019] 本发明实施例是这样实现的,一种分页存储器件的压缩管理方法,所述方法包括:

[0020] 步骤 A、以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址;

[0021] 步骤 B、获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址;获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0022] 本发明实施例的另一目的在于提供一种分页存储器件的压缩管理装置,所述装置包括:

[0023] 字典建立单元,用于以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址;

[0024] 处理单元,用于获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址;获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0025] 从上述技术方案可以看出,本发明实施例中数据压缩的最小单位是页,页内的数据不进行压缩和解压缩,数据的写入和读出也是以页为单位,因此本发明实施例提供的压缩管理方法非常适合采用运算能力不足或者成本低廉的分页存储器件。而且本发明实施例提供的压缩管理方法,在当前写命令中待写入的数据已经存入字典的话,则不写入数据,因此能有效提高分页存储器件的数据写入效率,减少数据的写入次数,进而减少分页存储器件的磨损。

### 附图说明

[0026] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中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性的前提下,还可以根据这些附图获得其他的附图。

[0027] 图 1 是本发明实施例一提供的分页存储器件压缩管理方法的实现流程图;

[0028] 图 2 是本发明实施例二提供的分页存储器件压缩管理方法的实现流程图;

[0029] 图 3 是本发明实施例三提供的分页存储器件压缩管理方法的实现流程图;

[0030] 图 4 是本发明实施例四提供的分页存储器件压缩管理装置的组成结构图。

## 具体实施方式

[0031] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0032] 为了说明本发明所述的技术方案,下面通过具体实施例来进行说明。

[0033] 实施例一:

[0034] 图 1 示出了本发明实施例一提供的分页存储器件压缩管理方法的实现流程,该方法过程详述如下:

[0035] 在步骤 S101 中,以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址等信息。

[0036] 在本实施例中,所述分页存储器件为以页为最小存储单元的存储器件,即按块擦除、按页读写的存储器件,例如闪存存储器件等。所述页数据为页内存储的数据。

[0037] 在步骤 S102 中,获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址等信息。

[0038] 在步骤 S103 中,获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码,若判断结果为“是”,则执行步骤 S105,若判断结果为“否”,则执行步骤 S104。

[0039] 在本实施例中,获取所述待写入页数据的特征码,将获取的特征码与所述字典中的所有特征码进行比较,判断所述字典中是否存在与所获取的特征码相同的特征码,若是,则表示当前写命令中的待写入页数据已存入所述字典,执行步骤 S105;若否,执行步骤 S104。

[0040] 本实施例中所述待写入页数据的特征码以及所述字典中页数据的特征码都可采用相同的算法获取。优选的是,通过哈希算法获取。当然也可以采用其他算法获取,在此并不用以限制本发明。

[0041] 哈希(Hash)算法属于对一个大的数据进行消息摘录,得到该段数据的特征码,主要用于密码验证和数据完整性验证,同时也可以用于快速判断两段数据是否相同。

[0042] 以 MD5 哈希算法为例,例如表 2 中如下的几段数据:

[0043]



页数据	MD5 值
1234567890abcdefg	7206ddfa511b6ab05734b603c1b88be6
2234567890abcdefg	628cb0947d31075fa305699029ddfec
0234567890abcdefg	df9e94f848b6608dd2af317b693744ba
1234567890abcdefgh	ad91a28a12e83ce4999b23260f7785c3
1234567890abcdef	996ce17f6abc9fe126b57aa5f1d8c92c
1234567890abcdefg	7206ddfa511b6ab05734b603c1b88be6

[0044] 表 2

[0045] 从表 2 可以看到,即使页数据只改变一点点,经哈希算法计算后的特征码都会产生大幅度变动,因此采用哈希算法计算特征码,能够降低重码率。

[0046] 在步骤 S104 中,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据,即将所述待写入页数据存储至与所述逻辑地址对应的物理地址。

[0047] 在步骤 S105 中,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0048] 在本实施例中,在所述写命令中的待写入页数据之前已写入过,即已存在于分页存储器件中时,不执行所述写指令,只将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,并提醒用户该数据已写入过,因此能有效提高分页存储器件的数据写入效率,减少数据的写入次数,进而减少分页存储器件的磨损。

[0049] 需要说明的是,本实施例所述字典中记录的页数据地址为物理页地址,但也可以为逻辑页地址,在写命令中的待写入页数据之前已写入过时,只需将当前写命令中待写入页数据的写入逻辑页地址指向具有相同特征码的原始数据写入的逻辑页地址即可。另外,当记录的是逻辑页地址时,在读出数据时需要将逻辑页地址转换为物理页地址。

[0050] 作为本发明的另一优选实施例,为了降低重码率,所述获取写命令中待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括:

[0051] 在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次消息摘录后的特征码;

[0052] 或者,在所述分页存储器件支持一种哈希算法时,对当前写命令中的待写入页数据

据进行多次不同偏移的消息摘录,获取每次不同偏移消息摘录后的特征码,判断所述字典中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码;

[0053] 或者,在所述分页存储器件支持错误检查和纠正 ECC 码时,获取当前写命令中待写入页数据的 ECC 码和消息摘录后的特征码,判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码;

[0054] 或者,获取当前写命令中待写入页数据的特征码,在所述字典中存在与所获取的特征码相同的特征码时,判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数,即如果所述分页存储器中页的大小为 8k 字节,则所述 N 小于或等于 8k 字节。需要说明的是,为了不影响数据压缩效率,上述比较过程可以采用硬件比较器实现。

[0055] 优选的是,为了保证在字典较大的情况下,能进一步提高数据压缩的效率,所述所述获取写命令中待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括:

[0056] 获取当前写命令中待写入页数据的特征码,对所述特征码进行消息摘录获得二次摘录值,判断所述字典中是否存在所述二次摘录值,若存在,再判断所述字典中是否存在所述特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据对应特征码的二次摘录值。其中,所述二次摘录值小于特征码的大小且在重码率允许的范围內。

[0057] 在本实施例中,预先建立特征码表和二次摘录值表,所述特征码表和二次摘录值表分别用于记录页数据的特征码以及所述特征码的二次摘录值。由于刚开始只需匹配数据量小的二次摘录值表,只有在出现相同的二次摘录值的情况下,才需要比较数据量大的特征码表,因此,能够大大提高匹配效率,进而提高数据压缩效率。

[0058] 以下通过举例来详细说明:

[0059] 在字典较大的情况下,例如有 1024 个页,每页数据采用哈希算法进行消息摘录后生成的特征码为 32 字节,则生成的特征码表有  $1024 \times 32$  字节,在这样的范围内进行搜索对

系统硬件要求较高。所以对所述特征码表进行二次摘录,该二次摘录后的值远小于 32 字节,且在重码率允许的范围内,如表 3 所示。

[0060]

原始数据	特征码值	二次摘录值
55555555555555555555	130668ae34911a14e4288367b071447a	130668ae
aaaaaaaaaaaaaaaaaaaa	22d42eb002cefa81e9ad604ea57bc01d	22d42eb0
000000000000000000	3ea032bf79e8c116b05f4698d5a8e044	3ea032bf
ffffffffffffffffffffffff	c78068577ec5dc1d06605f291a79174f	c7806857

[0061] 表 3

[0062] 表 3 中取特征码值(MD5 值)前 4 个字节为二次哈希值(二次摘录值),用于快速搜索特征码表中是否有匹配的二次摘录值,如果字典中存在与当前写命令中待写入页数据的二次摘录值相同的值,则再进行特征码值的匹配。

[0063] 本实施例中数据压缩的最小单位是页,页内的数据不进行压缩和解压缩,数据的写入和读出也是以页为单位,因此本发明实施例提供的压缩管理方法非常适合采用运算能力不足或者成本低的分页存储器件。而且本发明实施例提供的压缩管理方法,在当前写命令中待写入的数据已经存入字典的话,则不写入数据,因此能有效提高分页存储器件的数据写入效率,减少数据的写入次数,进而减少分页存储器件的磨损。

[0064] 实施例二:

[0065] 图 2 示出了本发明实施例二提供的分页存储器件压缩管理方法的实现流程,该方法过程详述如下:

[0066] 在步骤 S201 中,以页为单元建立静态字典,所述静态字典内记录有写入次数大于或者等于第一阈值的页数据和 / 或用户关心的页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址等信息。

[0067] 在本实施例中,所述静态字典是固定在所述分页存储器件内,不允许更新的字典,所述静态字典可以根据以下条件创建:

[0068] 1、根据所述分页存储器件的存储容量,设定满足一定比例(如 1%)的字典大小(存储容量),例如当分页存储器件的存储容量为 100M 时,字典的大小(存储容量)为 1M;

[0069] 2、根据所述分页存储器件的主要应用范围进行大量采样,统计分析以后得出最佳字典;例如,将写入次数大于或者等于第一阈值(10 次)的数据作为字典中的页数据;

[0070] 3、根据用户选择自身需要(即用户关心的数据),设定字典内容。

[0071] 在步骤 S202 中,获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址等信息;

[0072] 在步骤 S203 中,获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码,若判断结果为“是”,则执行步骤 S205,若判断结果为“否”,则执行步骤 S204;

[0073] 优选的是,判断所述静态字典中是否存在与所获取的特征码相同的特征码,若判断结果为“是”,则执行步骤 S205,若判断结果为“否”,则执行步骤 S204;

[0074] 在步骤 S204 中,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据,即将所述待写入页数据存储至与所述逻辑地址对应的物理地址;

[0075] 在步骤 S205 中,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0076] 作为本发明的另一优选实施例,为了降低重码率,所述获取写命令中待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括:

[0077] 在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次消息摘录后的特征码;

[0078] 或者,在所述分页存储器件支持一种哈希算法时,对当前写命令中的待写入页数据进行多次不同偏移的消息摘录,获取每次不同偏移消息摘录后的特征码,判断所述字典中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码;

[0079] 或者,在所述分页存储器件支持错误检查和纠正 ECC 码时,获取当前写命令中待写入页数据的 ECC 码和消息摘录后的特征码,判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码;

[0080] 或者,获取当前写命令中待写入页数据的特征码,在所述字典中存在与所获取的特征码相同的特征码时,判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数,即如果所述分页存储器中页的大小为 8k 字节,则所述 N 小于或等于 8k 字节。需要说明的是,为了不影响数据压缩效率,上述比较过程可以采用硬件比较器实现。

[0081] 优选的是,为了保证在字典较大的情况下,能进一步提高数据压缩的效率,所述所

述获取写命令中待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括:

[0082] 获取当前写命令中待写入页数据的特征码,对所述特征码进行消息摘录获得二次摘录值,判断所述字典中是否存在所述二次摘录值,若存在,再判断所述字典中是否存在所述特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据对应特征码的二次摘录值。其中,所述二次摘录值小于特征码的大小且在重码率允许的范围。

[0083] 在本实施例中,预先建立特征码表和二次摘录值表,所述特征码表和二次摘录值表分别用于记录页数据的特征码以及所述特征码的二次摘录值。由于刚开始只需匹配数据量小的二次摘录值表,只有在出现相同的二次摘录值的情况下,才需要比较数据量大的特征码表,因此,能够大大提高匹配效率,进而提高数据压缩效率。

[0084] 在本实施例中,根据预先建立的静态字典来判断所述待写入页数据是否之前已写入,由于该静态字典中包含经常性写入的数据,从而可有效提高判断的效率。

[0085] 实施例三:

[0086] 图3示出了本发明实施例三提供的分页存储器件压缩管理方法的实现流程,该方法过程详述如下:

[0087] 在步骤S301中,以页为单元建立动态字典,所述动态字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址等信息。

[0088] 在本实施例中,所述动态字典为允许被更新的字典。

[0089] 在步骤S302中,设定第二阈值(例如10次),对写入分页存储器件的页数据的写入次数进行计数。

[0090] 在步骤S303中,获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址。

[0091] 在步骤S304中,获取所述待写入页数据的特征码,判断所述动态字典中是否存在与所获取的特征码相同的特征码,如果判断结果为“是”,则执行步骤S305,如果判断结果为“否”,则执行步骤S306。

[0092] 在步骤S305中,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0093] 在步骤S306中,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据。

[0094] 在步骤S307中,判断所述待写入页数据的写入次数是否达到所述第二阈值,如果判断结果为“是”,则执行步骤S308,如果判断结果为“否”,则结束本次操作。

[0095] 在步骤S308中,将该待写入页数据更新到所述动态字典中。

[0096] 具体包括该待写入页数据、该待写入页数据对应的特征码以及该待写入页数据写入的物理页地址等信息更新到所述动态字典中。

[0097] 作为本发明的另一优选实施例,为了防止所述动态字典变得越来越大,影响数据压缩及搜索的效率,还包括:

[0098] 当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中无效的条目或者在预设时间内未使用的条目,所述无效的条目为所述动态字典中存储的页数据的写入地址已经写入了其他数据,导致该页数据无效;所述预设时间内未使用的条目为该条目对应的页数据在预设时间内未写入过分页存储器件。

[0099] 或者,在所述页数据的写入次数达到所述第二阈值之后继续对所述页数据的写入次数进行计数,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中写入次数小于第四阈值的页数据。例如全 0x8f 的页数据写入次数达到了 10 次,则该页数据被更新到所述动态字典,但所述动态字典内各页数据的计数仍在继续,当所述动态字典的存储容量达到第三阈值,全 0x8f 页数据的写入次数小于第四阈值(如 15 次)时,将全 0x8f 页数据从所述动态字典中删除。

[0100] 作为本发明的另一优选实施例,为了降低重码率,所述获取写命令中待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括:

[0101] 在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次消息摘录后的特征码;

[0102] 或者,在所述分页存储器件支持一种哈希算法时,对当前写命令中的待写入页数据进行多次不同偏移的消息摘录,获取每次不同偏移消息摘录后的特征码,判断所述字典中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码;

[0103] 或者,在所述分页存储器件支持错误检查和纠正 ECC 码时,获取当前写命令中待写入页数据的 ECC 码和消息摘录后的特征码,判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码;

[0104] 或者,获取当前写命令中待写入页数据的特征码,在所述字典中存在与所获取的特征码相同的特征码时,判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同,若否,执行当前写命令,根据当前写命令中待写入

页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数，即如果所述分页存储器中页的大小为 8k 字节，则所述 N 小于或等于 8k 字节。需要说明的是，为了不影响数据压缩效率，上述比较过程可以采用硬件比较器实现。

[0105] 优选的是，为了保证在字典较大的情况下，能进一步提高数据压缩的效率，所述所述获取写命令中待写入页数据的特征码，判断所述字典中是否存在与所获取的特征码相同的特征码；若否，执行所述写命令，根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址具体包括：

[0106] 获取当前写命令中待写入页数据的特征码，对所述特征码进行消息摘录获得二次摘录值，判断所述字典中是否存在所述二次摘录值，若存在，再判断所述字典中是否存在所述特征码，若否，执行当前写命令，根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据；若是，不执行所述写命令，将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址，其中所述字典中记录有页数据对应特征码的二次摘录值。其中，所述二次摘录值小于特征码的大小且在重码率允许的范围之内。

[0107] 在本实施例中，预先建立特征码表和二次摘录值表，所述特征码表和二次摘录值表分别用于记录页数据的特征码以及所述特征码的二次摘录值。由于刚开始只需匹配数据量小的二次摘录值表，只有在出现相同的二次摘录值的情况下，才需要比较数据量大的特征码表，因此，能够大大提高匹配效率，进而提高数据压缩效率。

[0108] 作为本发明的另一优选实施例，所述字典还可以同时包含静态字典和动态字典；所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和 / 或用户关心的数据；

[0109] 所述方法在获取写命令的步骤之前还包括：

[0110] 设定第二阈值，对写入分页存储器件的页数据的写入次数进行计数；

[0111] 所述获取所述待写入页数据的特征码，判断所述字典中是否存在与所获取的特征码相同的特征码具体包括：

[0112] 获取所述待写入页数据的特征码，判断所述静态字典中是否存在与所获取的特征码相同的特征码，若否，根据所述待写入页数据写入分页存储器件的次数判断所述待写入页数据是否需要更新到所述动态字典。

[0113] 实施例四：

[0114] 图 4 示出了本发明实施例四提供的分页存储器件压缩管理装置的组成结构，为了便于说明，仅示出了与本发明实施例相关的部分。

[0115] 该分页存储器件压缩管理装置可以应用于存储设备中，可以是运行于存储设备内的软件单元、硬件单元或者软硬件相结合的单元，也可以作为独立的挂件集成到存储设备中或者运行于存储设备的应用系统中。

[0116] 该分页存储器件压缩管理装置包括字典建立单元 41 以及处理单元 42。其中，各单元的具体功能如下：

[0117] 字典建立单元 41,用于以页为单元建立字典,所述字典内记录有页数据、所述页数据对应的特征码以及所述页数据写入的物理页地址;

[0118] 处理单元 42,用于获取写命令,所述写命令包含有待写入页数据以及所述待写入页数据写入的逻辑页地址;获取所述待写入页数据的特征码,判断所述字典中是否存在与所获取的特征码相同的特征码;若否,执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将所述待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址。

[0119] 优选的是,所述字典为静态字典,所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和/或用户关心的页数据。

[0120] 优选的是,所述字典为动态字典;

[0121] 所述处理单元 42 在获取写命令之前还用于,设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

[0122] 所述处理单元 42 在执行所述写命令,根据所述待写入页数据写入的逻辑页地址写入所述待写入页数据之后,还用于在所述页数据的写入次数达到所述第二阈值时,将该页数据更新到所述动态字典中。

[0123] 优选的是,所述字典包含静态字典和动态字典;所述静态字典中的页数据为写入次数大于或者等于第一阈值的数据和/或用户关心的数据;

[0124] 所述处理单元 42 在获取写命令之前还用于,设定第二阈值,对写入分页存储器件的页数据的写入次数进行计数;

[0125] 所述处理单元 42,用于获取所述待写入页数据的特征码,判断所述静态字典中是否存在与所获取的特征码相同的特征码,若否,根据所述待写入页数据写入分页存储器件的次数判断所述待写入页数据是否需要更新到所述动态字典。

[0126] 进一步的是,所述处理单元 42 还用于,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中无效的条目或者在预设时间内未使用的条目,所述无效的条目为所述动态字典中存储的页数据的写入地址已经写入了其他数据,导致该页数据无效;所述预设时间内未使用的条目为该条目对应的页数据在预设时间内未写入过分页存储器件;或者,

[0127] 在所述页数据的写入次数达到所述第二阈值之后继续对所述页数据的写入次数进行计数,当所述动态字典的存储容量达到第三阈值时,删除所述动态字典中写入次数小于第四阈值的页数据。

[0128] 进一步的,所述处理单元 42 具体用于,

[0129] 在所述分页存储器件支持多种哈希算法时,对当前写命令中的待写入页数据进行多次消息摘录,获取每次消息摘录后的特征码,判断所述字典中是否存在与所述每次消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次消息摘录后的特征码;

[0130] 或者,在所述分页存储器件支持一种哈希算法时,对当前写命令中的待写入页数据进行多次不同偏移的消息摘录,获取每次不同偏移消息摘录后的特征码,判断所述字典



中是否存在与所述每次不同偏移消息摘录后的特征码都相同的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据多次不同偏移消息摘录后的特征码;

[0131] 或者,在所述分页存储器件支持错误检查和纠正 ECC 码时,获取当前写命令中待写入页数据的 ECC 码和消息摘录后的特征码,判断所述字典中是否同时存在该 ECC 码和消息摘录后的特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据的 ECC 码和消息摘录后的特征码;

[0132] 或者,获取当前写命令中待写入页数据的特征码,在所述字典中存在与所获取的特征码相同的特征码时,判断所述待写入页数据中的 N 个字节与所述字典中相同特征码对应的页数据中的 N 个字节是否全部相同,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述 N 为大于零且小于或者等于所述分页存储器中页的大小的整数。

[0133] 进一步的,所述处理单元 42 具体用于,

[0134] 获取当前写命令中待写入页数据的特征码,再对所述特征码进行消息摘录获得二次摘录值,判断所述字典中是否存在所述二次摘录值,若存在,再判断所述字典中是否存在所述特征码,若否,执行当前写命令,根据当前写命令中待写入页数据写入的逻辑页地址写入所述待写入页数据;若是,不执行所述写命令,将当前写命令中待写入页数据写入的逻辑页地址指向与所述待写入页数据具有相同特征码的页数据写入的物理页地址,其中所述字典中记录有页数据对应特征码的二次摘录值。

[0135] 在本实施例中,所述页数据的特征码可以通过哈希算法获取。

[0136] 本实施例提供的分页存储器件压缩管理装置可以使用在前述对应的分页存储器件压缩管理方法,详情参见上述分页存储器件压缩管理方法实施例一至三的相关描述,在此不再赘述。

[0137] 本领域普通技术人员可以理解为上述实施例四所包括的各个单元只是按照功能逻辑进行划分的,但并不局限于上述的划分,只要能够实现相应的功能即可;另外,各功能单元的具体名称也只是为了便于相互区分,并不用于限制本发明的保护范围。

[0138] 综上所述,本发明实施例中数据压缩的最小单位是页,页内的数据不进行压缩和解压缩,数据的写入和读出也是以页为单位,因此本发明实施例提供的压缩管理方法非常适合采用运算能力不足或者成本低廉的分页存储器件。而且本发明实施例提供的压缩管理方法,在当前写命令中待写入的数据已经存入字典的话,则不写入数据,因此能有效提高分页存储器件的数据写入效率,减少数据的写入次数,进而减少分页存储器件的磨损。另外,本发明实施例通过多种比较方式,可有效降低比较过程中的重码率,并能提高匹配的效率和数据压缩效率,具有较强的易用性和实用性。

[0139] 以上内容是结合具体的优选实施方式对本发明所作的进一步详细说明,不能认定本发明的具体实施只局限于这些说明。对于本发明所属技术领域的普通技术人员来说,在不脱离本发明构思的前提下做出若干等同替代或明显变型,而且性能或用途相同,都应当视为属于本发明由所提交的权利要求书确定的专利保护范围。

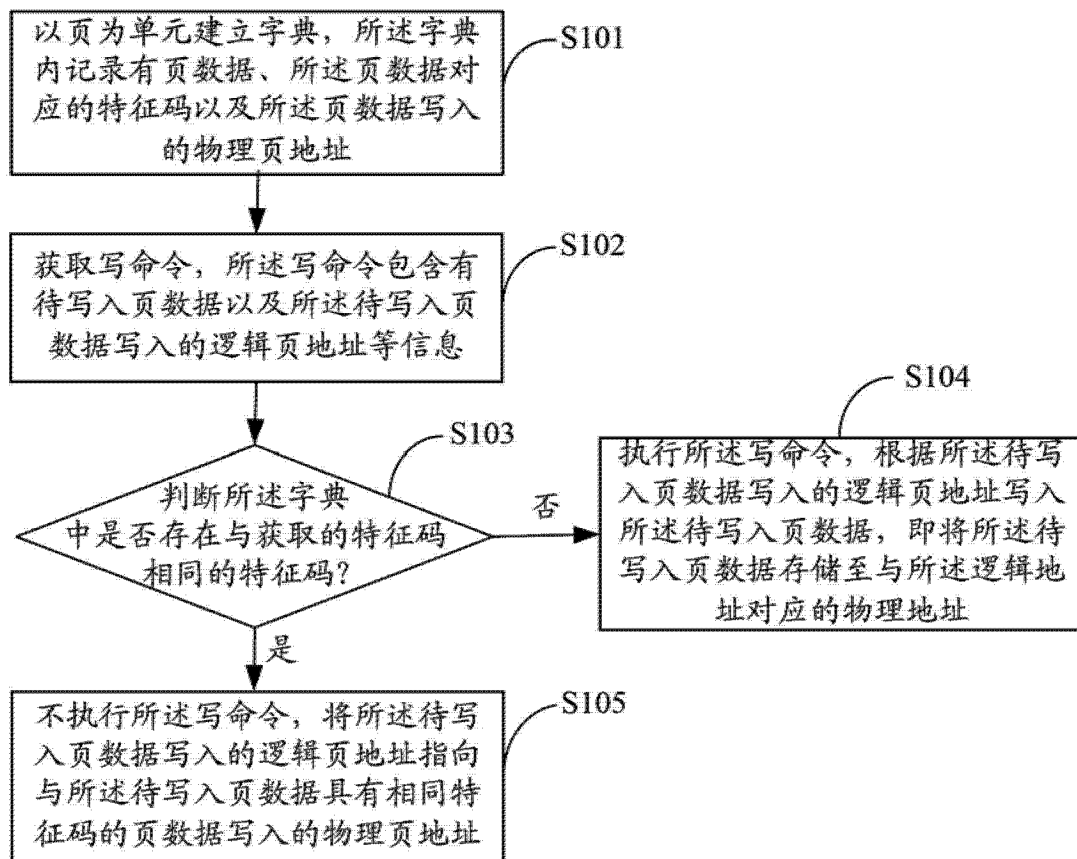


图 1

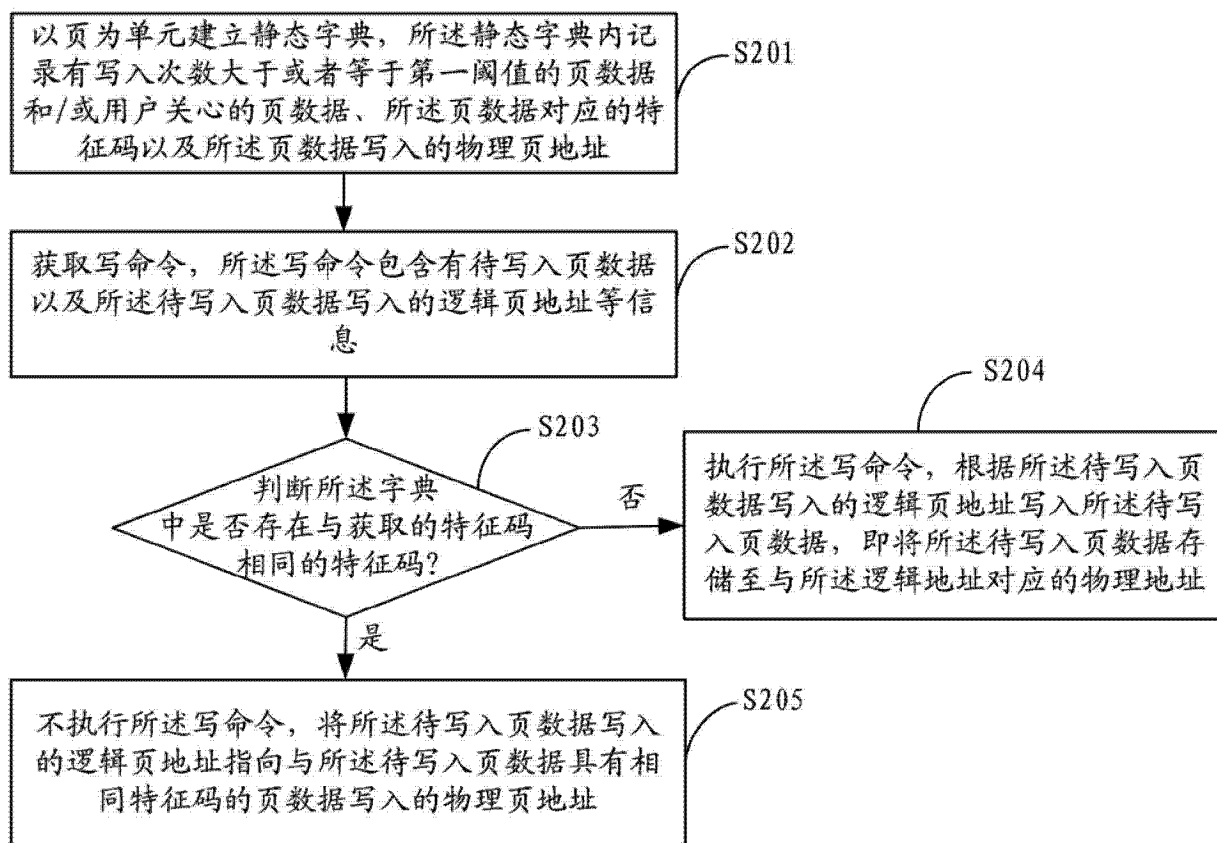


图 2

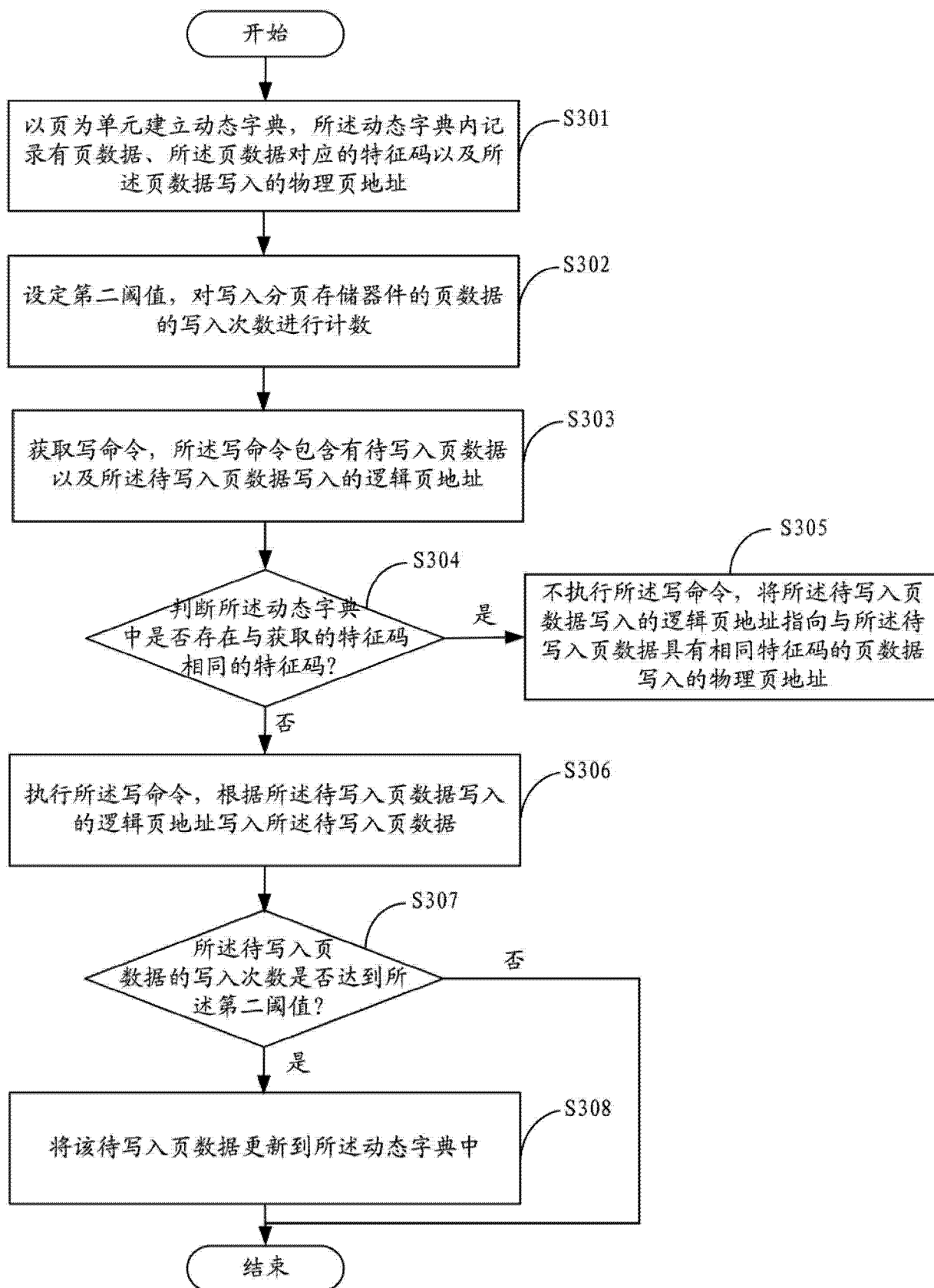


图 3

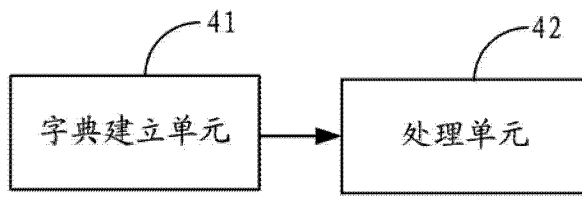


图 4