

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 7 月 4 日 (04.07.2019)



(10) 国际公布号
WO 2019/128409 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2018/111180
- (22) 国际申请日: 2018 年 10 月 22 日 (22.10.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201711455790.2 2017 年 12 月 28 日 (28.12.2017) CN
- (71) 申请人: 中国银联股份有限公司 (CHINA UNIONPAY CO., LTD.) [CN/CN]; 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。
- (72) 发明人: 何东杰 (HE, Dongjie); 中国上海市浦东新区含笑路 36 号银联大厦 7 楼, Shanghai 200135 (CN)。
- (74) 代理人: 中国专利代理 (香港) 有限公司 (CHINA PATENT AGENT (HK) LTD.); 中国香港特别行政区香港湾仔港湾道 23 号鹰君中心 22 字楼, Hong Kong (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告 (条约第 21 条 (3))。

(54) Title: DATA COMPRESSION AND STORAGE METHOD AND DATA COMPRESSION AND STORAGE DEVICE

(54) 发明名称: 一种数据压缩存储方法以及数据压缩存储装置

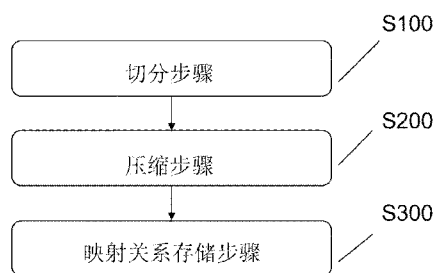


图 1

S100 Segmentation step
S200 Compression step
S300 Mapping relationship storage step

(57) Abstract: The present invention relates to a data compression and storage method and a data compression and storage device. The data compression method comprises the following steps: a segmentation step: segmenting original data into a plurality of fields; and a compression step: on the basis of different data contents, compressing different fields by using different compression policies, and storing the compressed data. According to the data compression and storage method and the data compression and storage device of the present invention, different compression methods can be used in consideration of different data contents, data compression efficiency can be effectively improved, and compared with common data compression tools such as GZIP and SNAPPY, data compression rate is significantly improved.

(57) 摘要: 本发明涉及数据压缩存储方法以及数据压缩存储装置。该数据压缩方法包括下述步骤: 切分步骤, 将原始数据切分成多个字段; 以及压缩步骤, 基于数据内容的不同, 对于不同字段采用不同的压缩策略进行压缩并存储压缩后的压缩数据。根据本发明的数据压缩存储方法以及数据压缩存储装置, 能够考虑到数据内容的不同而采用不同的压缩方法, 能够有效提高数据压缩效率, 相比通用的 GZIP、SNAPPY 等数据压缩工具, 在数据压缩率上有明显的提升。



WO 2019/128409 A1

一种数据压缩存储方法以及数据压缩存储装置

技术领域

[0001] 本发明涉及数据处理技术，具体涉及一种数据压缩存储方法以及数据压缩存储装置。

背景技术

[0002] 企业在进行数据存储时候，从节约存储空间以及提升读取效率方面考虑，一般都会对数据进行压缩存储。但是，通用的压缩工具针对的是所有的数据。

[0003] 再者，现有的常见的数据压缩工具包括 GZIP、SNAPPY 等，是针对通用数据进行压缩。

[0004] 但是，如上所说，目前企业进行数据存储时采用的压缩工具是对所有数据通用的工具，对企业来说，没有充分考虑到企业数据的特点。因此，数据的压缩效率并不是很高。

发明内容

[0005] 鉴于所述问题，本发明旨在提出一种进一步数据压缩存储方法以及数据压缩存储装置。

[0006] 本发明的数据压缩存储方法，其特征在于，包括下述步骤：

切分步骤，将原始数据切分成多个字段；以及

压缩步骤，基于数据内容的不同，对于不同字段采用不同的压缩策略进行压缩并存储压缩后的压缩数据。

[0007] 优选地，在所述压缩步骤中，判断字段之间的关联关系强弱，根据关联关系强弱，设定压缩策略。

[0008] 优选地，所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及

针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0009] 优选地，所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

将存在相关关系的多个字段进行组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0010] 优选地，所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及
针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且，另一方面，也将存在相关关系的多个字段进行组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0011] 优选地，作为压缩策略，对于枚举型字符串字段使用二进制进行压缩存储、对于字符串数值转换为整数或者浮点进行压缩存储。

[0012] 优选地，作为压缩策略，将短字段进行组合后再进行压缩存储、将近似字段压缩冗余信息后进行压缩存储、对于字段之间信息逆序的仅存储其中一个、对于字段之间存在包含关系的压缩冗余信息后存储。

[0013] 优选地，作为压缩策略，对于枚举型字符串字段使用二进制进行压缩存储、将字符串数值转换为整数或者浮点进行压缩存储、将短字段进行组合后再进行压缩存储、将近似字段压缩冗余信息后进行压缩存储、对于字段之间信息逆序的仅存储其中一个、对于字段之间存在包含关系的压缩冗余信息后存储。

[0014] 优选地，进一步具备：

映射关系存储步骤，建立所述原始数据与所述压缩数据之间的映射关系并存储。

[0015] 本发明的数据压缩存储装置，其特征在于，具备：

切分模块，用于将原始数据切分成多个字段；以及

压缩模块，用于基于数据内容的不同，对于不同字段采用不同的压缩策略进行压缩并存储压缩后的压缩数据。

[0016] 优选地，所述压缩模块判断字段之间的关联关系强弱，根据关联关系强弱，设定压缩策略。

[0017] 优选地，进一步具备：

映射关系存储模块，用于建立所述原始数据与所述压缩数据之间的映射关系并存储。

[0018] 优选地，所述压缩步骤具备下述子模块：

内容分析子模块，对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0019] 优选地，所述压缩模块具备：

内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

压缩子模块，将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0020] 优选地，所述压缩模块具备：

内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且也将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0021] 本发明的计算机可读介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时实现上述的数据压缩存储方法。

[0022] 本发明的计算机设备，包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现上述的数据压缩存储方法的步骤。

[0023] 如上所述，根据本发明的数据压缩存储方法以及数据压缩存储装置，提出了一种面向企业数据特点的高效数据压缩方案，通过针对企业数据特点进行分析，针对不同数据字段使用相应的高效压缩算法，从而提高数据压缩效率。相比通用的 GZIP、SNAPPY 等数据压缩工具，在数据压缩率上有明显的提升。

[0024]

附图说明

[0025] 图 1 是表示本发明的数据压缩存储方法的流程图。

[0026] 图 2 是表示本发明的数据压缩存储装置的结构示意图。

具体实施方式

[0027] 下面介绍的是本发明的多个实施例中的一些，旨在提供对本发明的基本了解。并不旨在确认本发明的关键或决定性的要素或限定所要保护的范围。

[0028] 本发明的数据压缩存储方法以及数据压缩存储装置的主要构思在于，针对企业数据内容进行分析，建立数据分布图，结合数据内容和数据分布，对每个信息采用对应的优化压缩算法。比如说对于枚举型字符串字段采用二进制编码进行压缩，将存储数字的字符串转换为整数或者浮点类型等。再者，对于数据内容中相关联的字段，可以进行合并后再进行压缩。再者，有些字段来自于其他字段的组合，则可以仅存储其中一份数据，有些字段是其他字段的倒序，则仅存储其中一份数据。

[0029] 接着，对于本发明的数据压缩存储方法进行说明。

[0030] 图 1 是表示本发明的数据压缩存储方法的流程图。

[0031] 如图 1 所示，本发明的数据压缩存储方法包括下述步骤：

切分步骤 S100：将原始数据切分成多个字段；以及

压缩步骤 S200：基于数据内容的不同，对于不同字段采用不同的优化的压缩策略进行压缩并存储压缩后的压缩数据，其中，判断字段之间的关联关系强弱，根据关联关系强弱，设定不同的压缩策略。

[0032] 根据切分步骤 S100 和压缩步骤 S200，通过数据内容的分析，建立数据表字段、字段之间的关联关系，再采用对应的优化压缩算法进行压缩，从而能够达到提高数据压缩率的效果。

[0033] 作为优选的方式，进一步可设置映射关系存储步骤 S300，在该映射关系存储步骤 S300 中在元数据中建立原始数据与压缩数据之间的映射关系并存储，这样，从外部访问数据时能够根据该映射关系顺利解析出原始数据。

[0034] 接着，对于本发明的数据压缩存储装置进行简单说明。图 2 是表示本发明的数据压缩存储装置的结构示意图。

[0035] 如图 2 所示，本发明的数据压缩存储装置具备：

切分模块 100，用于将原始数据切分成多个字段；以及

压缩模块 200，用于基于数据内容的不同，对于不同字段采用不同的优化压缩策

略进行压缩并存储压缩后的压缩数据，其中，所述压缩模块 200 判断字段之间的关联关系强弱，根据关联关系强弱，设定压缩策略。

[0036] 作为优选的方式，进一步还可以设置映射关系存储模块 300，在该映射关系存储模块 300 中建立原始数据与压缩数据之间的映射关系并存储，这样，从外部访问数据时能够根据该映射关系顺利解析出原始数据。另外，映射关系存储模块 300 并不是本发明的数据压缩存储装置所必须的构造单元，而只是优选具备的一个模块。

[0037] 接着对于本发明的数据压缩存储方法以及数据压缩存储装置的具体实施方式进行说明。

[0038]

第一实施方式

第一实施方式涉及的是针对每个字段使用优化的压缩策略进行压缩存储的实施方式。

[0039] 首先，对于第一实施方式的数据压缩存储方法进行说明。

[0040] 第一实施方式的数据压缩存储方法包括切分步骤和压缩步骤，其中切分步骤与上述切分步骤 S100 相同，而压缩步骤具体包括下述子步骤：

对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0041] 例如，作为压缩策略，对不同的数据内容采用相关优化的数据压缩存储方式；比如枚举型使用二进制存储、字符串数值转换为整数或者浮点存储等等。

[0042] 作为优选的方式，进一步可设置映射关系存储步骤，在该映射关系存储步骤中建立所述原始数据与所述压缩数据之间的映射关系并存储，这样，从外部访问数据时能够根据该映射关系顺利解析出原始数据。

[0043] 再者，简单说明第一实施方式的数据压缩存储装置。

[0044] 第一实施方式的数据压缩存储装置具备：切分模块以及压缩模块。第一实施方式中的切分模块的功能与上述切分模块 100 相同，第一实施方式中的压缩模块具体地具备下述子模块：

内容分析子模块，对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0045] 另外，也可以选择性地设置映射关系存储模块，在该映射关系存储模块中建立原始数据与压缩数据之间的映射关系并存储。

[0046]

第二实施方式

第二实施方式涉及的是针对多个字段使用优化的压缩策略的实施方式。

[0047] 首先，对于第二实施方式的数据压缩存储方法进行说明。

[0048] 第二实施方式的数据压缩存储方法包括切分步骤和压缩步骤，其中切分步骤与上述切分步骤 S100 相同，而压缩步骤具体包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0049] 例如，作为压缩策略，对多个字段进行组合，并采用相关优化的数据压缩存储方式；比如短字段进行组合后再压缩存储、近似字段压缩冗余信息后压缩存储、字段之间信息逆序的仅存储其中一个、字段之间存在包含关系的压缩冗余信息后存储等等。

[0050] 作为优选的方式，进一步可设置映射关系存储步骤，在该映射关系存储步骤中建立所述原始数据与所述压缩数据之间的映射关系并存储，这样，从外部访问数据时能够根据该映射关系顺利解析出原始数据。

[0051] 再者，简单说明第二实施方式的数据压缩存储装置。

[0052] 第二实施方式的数据压缩存储装置具备：切分模块以及压缩模块。第一实施方式中的切分模块的功能与上述切分模块 100 相同，第一实施方式中的压缩模块具体地具备下述子模块：

内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

压缩子模块，将存在相关关系的多个字段组合，针对组合后的字段，对于不同的

数据内容采用不同的压缩策略进行压缩存储。

[0053] 另外，也可以选择性地设置映射关系存储模块，在该映射关系存储模块中建立原始数据与压缩数据之间的映射关系并存储。

[0054]

第三实施方式

第三实施方式涉及的是针对单个字段和多个字段组合均使用优化的压缩策略的实施方式。

[0055] 首先，对于第三实施方式的数据压缩存储方法进行说明。

[0056] 第三实施方式的数据压缩存储方法包括切分步骤和压缩步骤，其中切分步骤与上述切分步骤 S100 相同，而压缩步骤具体包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且也将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0057] 作为压缩策略，对多个字段进行组合，针对单字段和多字段均采用相关优化的数据压缩存储方式；比如枚举型使用二进制存储、字符串数值转换为整数或者浮点存储、短字段进行组合后再压缩存储、近似字段压缩冗余信息后压缩存储、字段之间信息逆序的仅存储其中一个、字段之间存在包含关系的压缩冗余信息后存储等等。

[0058] 作为优选的方式，进一步可设置映射关系存储步骤，在该映射关系存储步骤中建立所述原始数据与所述压缩数据之间的映射关系并存储，这样，从外部访问数据时能够根据该映射关系顺利解析出原始数据。

[0059] 再者，简单说明第三实施方式的数据压缩存储装置。

[0060] 第三实施方式的数据压缩存储装置具备：切分模块以及压缩模块。第一实施方式中的切分模块的功能与上述切分模块 100 相同，第一实施方式中的压缩模块具体地具备下述子模块：

内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关

关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且也将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

[0061] 另外，也可以选择性地设置映射关系存储模块，在该映射关系存储模块中建立原始数据与压缩数据之间的映射关系并存储。

[0062] 本发明还提供一种计算机可读介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时实现权利要求上述各实施方式数据压缩存储方法。

[0063] 本发明还提供一种计算机设备，包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现上述的数据压缩存储方法的步骤。

[0064] 如上所述，根据本发明的数据压缩存储方法以及数据压缩存储装置，能够考虑到数据内容的不同而采用不同的压缩方法，能够有效提高数据压缩效率，相比通用的 GZIP、SNAPPY 等数据压缩工具，在数据压缩率上有明显的提升。

[0065] 以上例子主要说明了本发明的数据压缩存储方法以及数据压缩存储装置。尽管只对其中一些本发明的具体实施方式进行了描述，但是本领域普通技术人员应当了解，本发明可以在不偏离其主旨与范围内以许多其他的形式实施。因此，所展示的例子与实施方式被视为示意性的而非限制性的，在不脱离如所附各权利要求所定义的本发明精神及范围的情况下，本发明可能涵盖各种的修改与替换。

权利要求书

1. 一种数据压缩存储方法，其特征在于，包括下述步骤：

切分步骤，将原始数据切分成多个字段；以及

压缩步骤，基于数据内容的不同，对于不同字段采用不同的压缩策略进行压缩并存储压缩后的压缩数据。

2. 如权利要求 1 所述的数据压缩存储方法，其特征在于，

在所述压缩步骤中，判断字段之间的关联关系强弱，根据关联关系强弱，设定压缩策略。

3. 如权利要求 1 所述的数据压缩存储方法，其特征在于，

所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及

针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

4. 如权利要求 1 所述的数据压缩存储方法，其特征在于，所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

将存在相关关系的多个字段进行组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

5. 如权利要求 1 所述的数据压缩存储方法，其特征在于，所述压缩步骤包括下述子步骤：

对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且，

另一方面，也将存在相关关系的多个字段进行组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

6. 如权利要求 3 所述的数据压缩存储方法，其特征在于，

作为压缩策略，对于枚举型字符串字段使用二进制进行压缩存储、对于字符串数值转换为整数或者浮点进行压缩存储。

7. 如权利要求 4 所述的数据压缩存储方法，其特征在于，

作为压缩策略，将短字段进行组合后再进行压缩存储、将近似字段压缩冗余信息后进行压缩存储、对于字段之间信息逆序的仅存储其中一个、对于字段之间存在包含关系的压缩冗余信息后存储。

8. 如权利要求 5 所述的数据压缩存储方法，其特征在于，

作为压缩策略，对于枚举型字符串字段使用二进制进行压缩存储、将字符串数值转换为整数或者浮点进行压缩存储、将短字段进行组合后再进行压缩存储、将近似字段压缩冗余信息后进行压缩存储、对于字段之间信息逆序的仅存储其中一个、对于字段之间存在包含关系的压缩冗余信息后存储。

9. 如权利要求 1~8 任意一项所述的数据压缩存储方法，其特征在于，进一步具备：映射关系存储步骤，建立所述原始数据与所述压缩数据之间的映射关系并存储。

10. 一种数据压缩存储装置，其特征在于，具备：

切分模块，用于将原始数据切分成多个字段；以及

压缩模块，用于基于数据内容的不同，对于不同字段采用不同的压缩策略进行压缩并存储压缩后的压缩数据。

11. 如权利要求 10 所述的数据压缩存储装置，其特征在于，

所述压缩模块判断字段之间的关联关系强弱，根据关联关系强弱，设定压缩策略。

12. 如权利要求 10 所述的数据压缩存储装置，其特征在于，进一步具备：

映射关系存储模块，用于建立所述原始数据与所述压缩数据之间的映射关系并存储。

13. 如权利要求 10~12 任意一项所述的数据压缩存储装置，其特征在于，所述压缩步骤具备下述子模块：

内容分析子模块，对切成多个字段的数据进行内容分析，建立字段之间的关联关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

14. 如权利要求 1 所述的数据压缩存储装置，其特征在于，所述压缩模块具备：

内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

压缩子模块，将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

15. 如权利要求 1 所述的数据压缩存储装置，其特征在于，所述压缩模块具备：内容分析子模块，对切成多个字段的数据进行内容分析，建立数据分布图以及字段之间的关联关系图，基于数据分布图以及关联关系图识别数据字段之间的相关关系；以及

压缩子模块，针对单个字段，对于不同的数据内容采用不同的压缩策略进行压缩存储，并且也将存在相关关系的多个字段组合，针对组合后的字段，对于不同的数据内容采用不同的压缩策略进行压缩存储。

16. 一种计算机可读介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时实现权利要求 1~9 中任意一项所述的数据压缩存储方法。

17. 一种计算机设备，包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现权利要求 1~9 中任意一项所述的数据压缩存储方法的步骤。

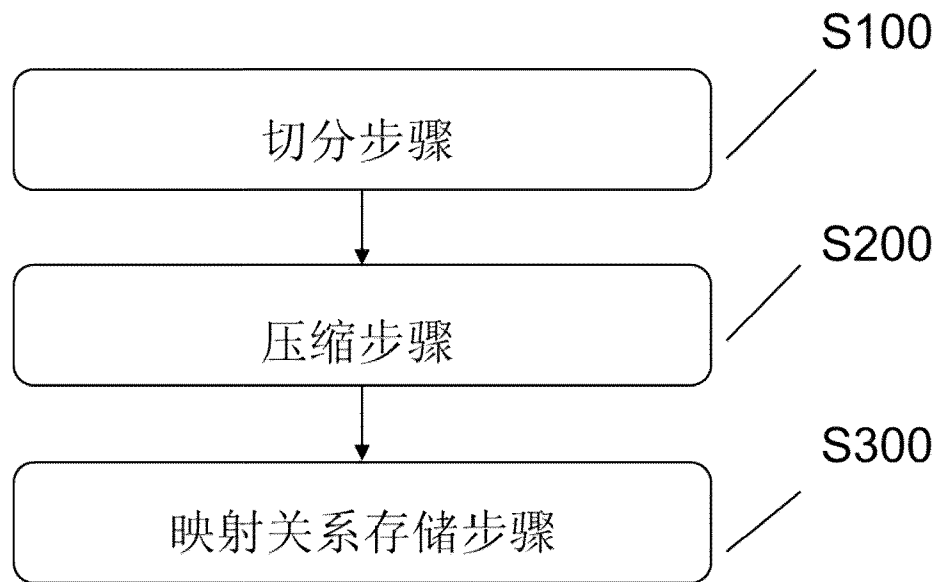


图 1

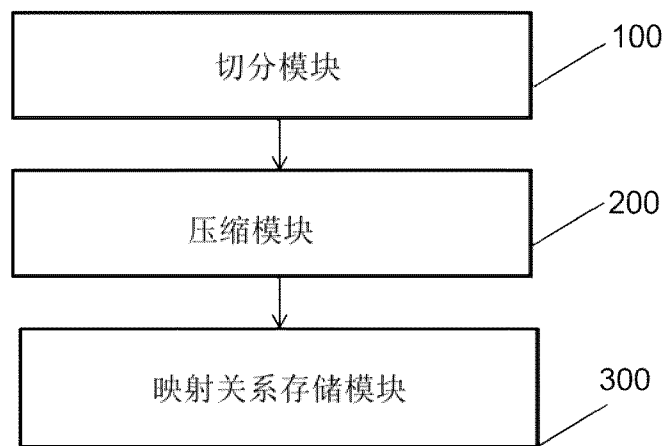


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/111180

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS; CNTXT; DWPI; SIPOABS: 数据, 压缩, 策略, 字段, 关联, 存储, 关系图, data, compress, strategytactic, field, associate, store, graph of relation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102708183 A (HUAWEI TECHNOLOGIES CO., LTD.) 03 October 2012 (2012-10-03) description, paragraphs [0048]-[0078]	1-17
Y	CN 105308589 A (LONGSAND LTD.) 03 February 2016 (2016-02-03) description, paragraphs [0007]-[0032]	1-17
Y	CN 106980639 A (CHINA UNIONPAY CO., LTD.) 25 July 2017 (2017-07-25) description, paragraphs [0015]-[0029]	2-9, 11-15
PX	CN 108304472 A (CHINA UNIONPAY CO., LTD.) 20 July 2018 (2018-07-20) claims and description	1-17



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

20 December 2019

Date of mailing of the international search report

28 December 2018

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/111180

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102708183	A	03 October 2012	CN	102708183	B	25 March 2015
CN	105308589	A	03 February 2016	US	2016085766	A1	24 March 2016
				WO	2014169956	A1	23 October 2014
				EP	2987092	A1	24 February 2016
CN	106980639	A	25 July 2017	None			
CN	108304472	A	20 July 2018	None			

国际检索报告

国际申请号

PCT/CN2018/111180

A. 主题的分类

G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F; H04L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CPRSABS; CNTXT; DWPI; SIPOABS; 数据, 压缩, 策略, 字段, 关联, 存储, 关系图, data, compress, strategytactic, field, associate, store, graph of relation

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 102708183 A (华为技术有限公司) 2012年 10月 3日 (2012 - 10 - 03) 说明书第[0048]-[0078] 段	1-17
Y	CN 105308589 A (朗桑有限公司) 2016年 2月 3日 (2016 - 02 - 03) 说明书第[0007]-[0032]段	1-17
Y	CN 106980639 A (中国银联股份有限公司) 2017年 7月 25日 (2017 - 07 - 25) 说明书第[0015]-[0029]段	2-9, 11-15
PX	CN 108304472 A (中国银联股份有限公司) 2018年 7月 20日 (2018 - 07 - 20) 权利要求书和说明书	1-17

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 12月 20日

国际检索报告邮寄日期

2018年 12月 28日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

唐田田

电话号码 62412065

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/111180

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	102708183	A	2012年 10月 3日	CN	102708183	B	2015年 3月 25日
CN	105308589	A	2016年 2月 3日	US	2016085766	A1	2016年 3月 24日
				WO	2014169956	A1	2014年 10月 23日
				EP	2987092	A1	2016年 2月 24日
CN	106980639	A	2017年 7月 25日	无			
CN	108304472	A	2018年 7月 20日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)