



(12) 发明专利

(10) 授权公告号 CN 102890719 B

(45) 授权公告日 2015. 12. 16

(21) 申请号 201210388174. 0

审查员 欧阳琦

(22) 申请日 2012. 10. 12

(73) 专利权人 浙江宇视科技有限公司

地址 310053 浙江省杭州市滨江区东信大道
66 号 C 厂房二、三层 A 区

(72) 发明人 王江柱

(74) 专利代理机构 北京博思佳知识产权代理有
限公司 11415

代理人 林祥

(51) Int. Cl.

G06F 17/30(2006. 01)

(56) 对比文件

CN 101025750 A, 2007. 08. 29, 全文 .

CN 101079032 A, 2007. 11. 28, 全文 .

US 2008/0111718 A1, 2008. 05. 15, 全文 .

US 2008/0275837 A1, 2008. 11. 06, 全文 .

US 2011/0066631 A1, 2011. 03. 17, 全文 .

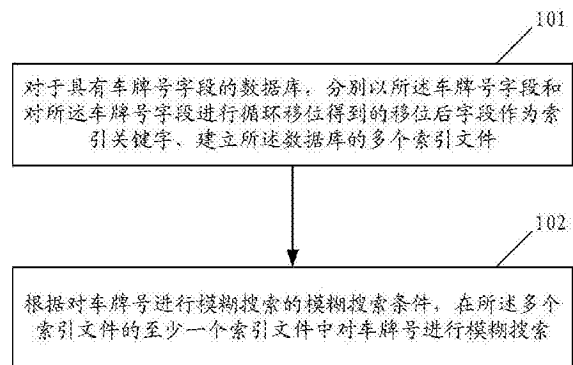
权利要求书2页 说明书6页 附图3页

(54) 发明名称

一种对车牌号进行模糊搜索的方法及装置

(57) 摘要

本发明提供一种对车牌号进行模糊搜索的方法及装置,属于智能识别技术领域。所述方法包括如下步骤:A、对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,并且每个索引文件与一种循环移位对应;B、根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索。本发明能够提高对车牌号进行模糊搜索的速度。



1. 一种对车牌号进行模糊搜索的方法,其特征在于,包括如下步骤:

A、对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,并且每个索引文件与一种循环移位对应;

B、根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索;

所述步骤 B 中进行一位模糊搜索的具体步骤包括:

B1、将所述模糊搜索条件转换为一个以上的子搜索条件的组合,其中,每个子搜索条件对应一个所述索引文件;

B2、依次根据每个子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

B3、对所有子搜索结果进行组合,得到车牌号的模糊搜索结果;

所述步骤 B 中进行多位模糊搜索的具体步骤包括:

B01、将所述模糊搜索条件转换为一个以上的子搜索条件的组合,所述子搜索条件包括第一类型子搜索条件和第二类型子搜索条件,其中,每个第一类型子搜索条件对应一个所述索引文件;

B02、依次根据每个第一类型子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

B03、对所有子搜索结果进行组合,并根据第二类型子搜索条件在组合结果中进行搜索,得到车牌号的模糊搜索结果。

2. 如权利要求 1 所述的方法,其特征在于,所述步骤 B3 具体包括:

对所有子搜索结果取并集或者取交集,得到车牌号的模糊搜索结果。

3. 如权利要求 1 所述的方法,其特征在于,所述移位后字段为:

所述车牌号字段的冗余索引字段,或者,所述数据库的冗余字段,或者,所述数据库的外键。

4. 如权利要求 1 所述的方法,其特征在于,所述多个索引文件的数量小于或等于车牌号的位数。

5. 一种对车牌号进行模糊搜索的装置,其特征在于,包括:

索引模块,用于对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,并且每个索引文件与一种循环移位对应;

搜索模块,用于根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索;

其中,所述搜索模块具体用于进行一位模糊搜索时,包括:

转换模块,用于将所述模糊搜索条件转换为一个以上的子搜索条件的组合,其中,每个子搜索条件对应一个所述索引文件;

子搜索模块,用于依次根据每个子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

组合模块,用于对所有子搜索结果进行组合,得到车牌号的模糊搜索结果;

所述搜索模块具体用于进行多位模糊搜索时,包括:

转换模块,用于将所述模糊搜索条件转换为一个以上的子搜索条件的组合,所述子搜索条件包括第一类型子搜索条件和第二类型子搜索条件,其中,每个第一类型子搜索条件对应一个所述索引文件;

子搜索模块,用于依次根据每个第一类型子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

组合模块,用于对所有子搜索结果进行组合,并根据第二类型子搜索条件在组合结果中进行搜索,得到车牌号的模糊搜索结果。

6. 如权利要求 5 所述的装置,其特征在于,所述组合模块具体用于:

对所有子搜索结果取并集或者取交集,得到车牌号的模糊搜索结果。

7. 如权利要求 5 所述的装置,其特征在于,所述移位后字段为:

所述车牌号字段的冗余索引字段,或者,所述数据库的冗余字段,或者,所述数据库的外键。

8. 如权利要求 5 所述的装置,其特征在于,所述多个索引文件的数量小于或等于车牌号的位数。

一种对车牌号进行模糊搜索的方法及装置

技术领域

[0001] 本发明涉及智能识别技术,尤其涉及一种对车牌号进行模糊搜索的方法及装置。

背景技术

[0002] 随着智能识别技术的普及,车牌识别技术已经应用到了实际交通领域。随着卡口、电警项目的普及与推广,数据库中存储的车牌记录也急剧增加,为后期车牌检索带来了一些性能瓶颈,特别是在进行数据库中车牌号模糊搜索时,采用现有的数据库匹配算法性能较差。其中,车牌号模糊搜索是指,在数据库中搜索满足预定条件的车牌号,所述预定条件为车牌号的一位或多位是不确定值。

[0003] 现有技术进行车牌号模糊搜索的方法为,以车牌号字段作为索引关键字为数据库建立索引文件,然后,在索引文件中搜索满足模糊搜索条件的车牌号。例如,对车牌号“浙 A1 2345”进行一位的模糊搜索(即查询与该车牌号最多有 1 位不相同的车牌号)时,需要查询下面的车牌号:

[0004] “浙 A1234×”

[0005] “浙 A123×5”

[0006] “浙 A12×45”

[0007] “浙 A1×345”

[0008] “浙 A×2345”

[0009] “浙 ×12345”

[0010] “×A1 2345”,其中“×”表示为任意可能值。

[0011] 采用现有方法的搜索性能较低,例如数据库中含有的过车记录为 1000 万条,对过车记录中的车牌号字段建索引文件,在该索引文件中存在 1000 万条索引信息。为了加快数据库搜索速度,一般该索引会按照某种排序进行存储,如按照车牌号的字符串大小进行排序。

[0012] 在进行车牌号搜索时,“浙 A1234×”的搜索时间较小,由于索引已经排序,因此只需要在“浙 A12340”到“浙 A12350”的索引间检索“浙 A1234×”,该检索速度较快。但是在搜索“×A12345”时需要在整个索引内进行检索,需要耗费较长的搜索时间,特别是在过车记录较多的卡口与电警系统中,这个搜索速度用户较难容忍。另外,现有的方法在进行多位模糊搜索(即车牌号中的多位为不确定值)时,也存在严重的性能问题。

发明内容

[0013] 有鉴于此,本发明的目的提供一种对车牌号进行模糊搜索的方法及装置,以提高对车牌号进行模糊搜索的速度。

[0014] 为实现上述目的,本发明提供技术方案如下:

[0015] 一种对车牌号进行模糊搜索的方法,包括如下步骤:

[0016] A、对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进

行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,并且每个索引文件与一种循环移位对应;

[0017] B、根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索。

[0018] 一种对车牌号进行模糊搜索的装置,包括:

[0019] 索引模块,用于对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,并且每个索引文件与一种循环移位对应;

[0020] 搜索模块,用于根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索。

[0021] 本发明通过以车牌号字段和对车牌号字段进行循环移位得到的移位后字段作为索引关键字,为数据库建立多个索引文件,可以根据具体的模糊搜索条件,选择合适的索引文件对车牌号进行模糊搜索,从而能够显著提高对车牌号进行模糊搜索的速度。

附图说明

[0022] 图1是本发明实施例的对车牌号进行模糊搜索的方法的流程图;

[0023] 图2是本发明实施例中为数据库建立多个索引文件的流程图;

[0024] 图3是本发明实施例中对车牌号进行模糊搜索的具体流程图;

[0025] 图4是本发明实施例中以车牌号字段作为索引关键字建立的索引文件示意图;

[0026] 图5是本发明实施例中以车牌号字段循环移动1位作为索引关键字建立的索引文件示意图;

[0027] 图6是本发明实施例的对车牌号进行模糊搜索的装置的结构图。

具体实施方式

[0028] 以下结合附图对本发明进行详细描述。

[0029] 针对现有技术中采用单一的车牌号索引对车牌号进行模糊搜索的方法存在的搜索速度慢的问题,本发明实施例通过采用车牌号循环冗余方式建立多个索引文件,来提高对车牌号的模糊搜索速度。

[0030] 参照图1,本发明实施例的对车牌号进行模糊搜索的方法,可以包括如下步骤:

[0031] 步骤101、对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,所述多个索引文件的数量小于或等于车牌号的位数,并且每个索引文件与一种循环移位对应;步骤102、根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索。

[0032] 步骤101中,用户在进行过车记录数据库存储时,可以将原先的车牌号字段进行冗余备份或对单一车牌号字段建立多个索引。如对原车牌号“浙A12345”进行循环移位转换后为:

[0033] “浙A12345”原车牌号字段

[0034] “5浙A1234”冗余备份

[0035] “45 浙 A123”冗余备份

[0036] “345 浙 A12”冗余备份

[0037] “2345 浙 A1”冗余备份

[0038] “12345 浙 A”冗余备份

[0039] “A12345 浙”冗余备份

[0040] 其中“冗余备份”信息可以以原车牌号字段的索引形式存在,也可以作为一个在数据库中新增的字段,或者,以数据库的外键等其他形式存在。

[0041] 参照图 2,为数据库建立多个索引文件的过程可以包括如下步骤:

[0042] 步骤 201,创建数据库,所述数据库中具有车牌号字段;

[0043] 步骤 202,以车牌号字段作为索引关键字,为数据库建立索引文件;

[0044] 步骤 203,将车牌号字段进行循环移动 1 位,以移位后得到的字段作为索引关键字,为数据库建立另外一个索引文件;

[0045] 步骤 204,判断是否完成所有车牌位的移位(即判断原车牌号的第 1 位是否已经移动至最后 1 位),若是,结束本流程,否则,返回步骤 203。

[0046] 可见,在本发明实施例中,为数据库创建的索引文件的数目可以等于车牌号字段的总位数,每个索引文件对应于车牌号字段的一种循环移位。例如,车牌号“浙 A12345”的总位数为 7,则可以建立 7 个索引文件。

[0047] 参照图 3,所述步骤 102 可以具体包括如下步骤:

[0048] 步骤 301,用户在进行车牌号模糊搜索时,先将模糊搜索条件转换为一个或者一个以上的子搜索条件的组合,其中,每个子搜索条件对应一个所述索引文件;

[0049] 步骤 302,依次根据每个子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

[0050] 步骤 303,对所有子搜索结果进行组合,得到车牌号的模糊搜索结果。

[0051] 其中,所述对所有子搜索结果进行组合为对所有子搜索结果取并集或者取交集。

[0052] 例如,需要搜索车牌号为“浙 A1×345”的数据库记录信息,可以在“345 浙 A12”冗余备份中搜索,需要搜索的搜索区间为“345 浙 A10”到“345 浙 A20”。

[0053] 用户如果需要搜索跟“浙 A12345”相差一位的所有车牌,则可以依次在原车牌号字段对应的索引文件以及 6 个冗余备份索引中进行搜索,搜索完成后对 7 次搜索结果进行取并集。

[0054] 另外,用户还可以进行多位模糊匹配搜索,可以充分利用冗余备份信息,灵活组合完成用户输入的模糊搜索条件。

[0055] 此时,所述步骤 102 可以具体包括如下步骤:

[0056] 步骤 S 1、将所述模糊搜索条件转换为一个或者一个以上的子搜索条件的组合,所述子搜索条件包括第一类型子搜索条件和第二类型子搜索条件,其中,每个第一类型子搜索条件对应一个所述索引文件;

[0057] 本步骤中,所述第二类型子搜索条件可以与任何索引文件相对应。

[0058] 步骤 S2、依次根据每个第一类型子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

[0059] 步骤 S3、对所有子搜索结果进行组合,并根据第二类型子搜索条件在组合结果中

进行搜索,得到车牌号的模糊搜索结果。

[0060] 本领域技术人员应当明白,如果第一类型子搜索条件的数目为1,则子搜索结果的数目也为1,此时,对所有子搜索结果进行组合对应的组合结果是指该子搜索结果自身。

[0061] 例如,用户需要搜索“浙 A1×3×5”的车牌,用户只需要在冗余备份“5 浙 A1234”中进行“5 浙 A1×××”搜索,在搜索获得的结果中再进行“×××××3×”的搜索。因为第一步已经将搜索范围降低了好几个量级,因此搜索速度会提高很多。而原先传统的搜索算法需要完成 X 次的“浙 A1B3×5”搜索(X 为 B 位值可能存在的字符可选值,B 为当前检索选择的确定值)。

[0062] 以下结合索引的实现原理对本发明实施例的技术效果进行说明。

[0063] 索引是建立记录间有规律排列的主要方式之一,可显著提高文件操作速度。当文件中记录的数目和数据量很大时,顺序查找速度会明显下降。为了采用随机查询提高查找速度,可以对文件建立索引。索引文件由两部分组成:索引和数据文件。由于数据文件记录多、数据量大,并且占据大量物理块,因此在数据文件中查找时速度很慢。如果对记录建立索引,那么相对数据文件而言,索引空间小,因而查找速度就快。

[0064] 可以将过车记录中的车牌号字段作为关键字段建立索引,在建立索引时可以采用 B+ 树、稀疏索引等形式。例如,以车牌号字段建立的稀疏索引后索引文件的结构如图 4 所示,其中索引的层数根据实际数据库与记录信息数量可以为多层。

[0065] 通过上面的索引建立后,再按照车牌号进行精确查找的速度会很快,同时也会加快较低位进行模糊匹配查找的速度。例如在查询“浙 A1234×”的车牌时速度也会大幅提高。本发明实施例在建立一个索引的基础上,为了提高车牌号模糊搜索的速度,采用循环移位方式建立了多个类似索引。例如,进行一位移位后的数据库索引文件如图 5 所示。

[0066] 本发明实施例的上述对车牌号进行模糊搜索的方法,通过适当的增加数据库中的冗余信息,可以提高车牌号搜索的速度,使得车牌号搜索的模糊匹配能够在实际中应用,满足用户对以车牌号为检索条件的实时性要求。

[0067] 现有的车牌号搜索算法在进行车牌号模糊匹配时,如在进行“浙 A12345”的一位模糊匹配时,搜索范围如下:

[0068]

搜索值	搜索范围
“浙 A1234×”	“浙 A12340” - “浙 A12350”
“浙 A123×5”	“浙 A12300” - “浙 A12400”

[0069]

“浙 A12×45”	“浙 A12000” - “浙 A13000”
“浙 A1×345”	“浙 A10000” - “浙 A20000”
“浙 A×2345”	“浙 A00000” - “浙 B00000”
“浙×12345”	“浙”的所有项记录
“×A12345”	整个搜索项

[0070] 从上面的统计来看,要完成“浙 A12345”的一位模糊匹配,需要对整个车牌索引区域进行一遍以上的数据比对操作。

[0071] 而采用本发明实施例的方法,每次搜索都在相应的索引文件中检索,每次的检索范围与“浙 A1234X”的基本相同,即为“浙 A12340”-“浙 A12350”。因此完成“浙 A12345”的一位模糊匹配,索引范围为整个搜索范围的 $1/(7*X^6)$,X 为每位车牌可能存在的 X 种选择字符。

[0072] 与上述方法相对应,本发明实施例还提供一种对车牌号进行模糊搜索装置。参照图 6,所述装置可以包括:

[0073] 索引模块 10,用于对于具有车牌号字段的数据库,分别以所述车牌号字段和对所述车牌号字段进行循环移位得到的移位后字段作为索引关键字,建立所述数据库的多个索引文件,所述多个索引文件的数量小于或等于车牌号的位数,并且每个索引文件与一种循环移位对应;所述移位后字段可以为:所述车牌号字段的冗余索引字段,或者,所述数据库的冗余字段,或者,所述数据库的外键;

[0074] 搜索模块 20,用于根据对车牌号进行模糊搜索的模糊搜索条件,在所述多个索引文件的至少一个索引文件中对车牌号进行模糊搜索。

[0075] 其中,所述搜索模块可以包括:

[0076] 转换模块 21,用于将所述模糊搜索条件转换为一个或者一个以上的子搜索条件的组合,其中,每个子搜索条件对应一个所述索引文件;

[0077] 子搜索模块 22,用于依次根据每个子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

[0078] 组合模块 23,用于对所有子搜索结果进行组合,得到车牌号的模糊搜索结果。所述组合模块 23 具体可以对所有子搜索结果取并集或者取交集,来得到车牌号的模糊搜索结果。

[0079] 如前所述,用户还可以进行多位模糊匹配搜索,可以充分利用冗余备份信息,灵活组合完成用户输入的模糊搜索条件。

[0080] 此时,所述转换模块 21 用于:将所述模糊搜索条件转换为一个或者一个以上的子搜索条件的组合,所述子搜索条件包括第一类型子搜索条件和第二类型子搜索条件,其中,每个第一类型子搜索条件对应一个所述索引文件,,所述第二类型子搜索条件可以与任何索引文件相对应;

[0081] 所述子搜索模块 22 用于:依次根据每个第一类型子搜索条件,在对应的索引文件中进行搜索,得到对应的子搜索结果;

[0082] 所述组合模块用于:对所有子搜索结果进行组合,并根据第二类型子搜索条件在组合结果中进行搜索,得到车牌号的模糊搜索结果。

[0083] 本领域技术人员应当明白,如果第一类型子搜索条件的数目为 1,则子搜索结果的数目也为 1,此时,对所有子搜索结果进行组合对应的组合结果是指该子搜索结果自身。

[0084] 综上所述,本发明通过以车牌号字段和对车牌号字段进行循环移位得到的移位后字段作为索引关键字,为数据库建立多个索引文件,可以根据具体的模糊搜索条件,选择合适的索引文件对车牌号进行模糊搜索,从而能够显著提高对车牌号进行模糊搜索的速度。

[0085] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精

神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

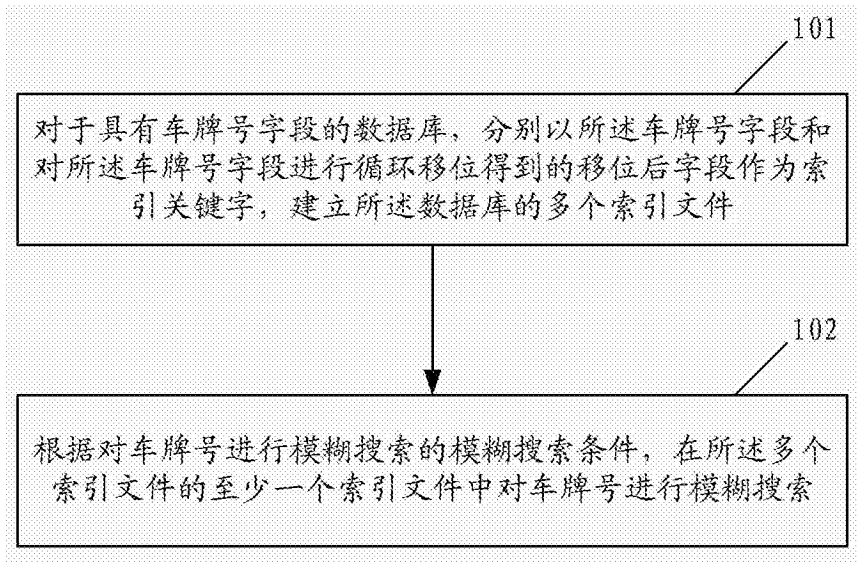


图 1

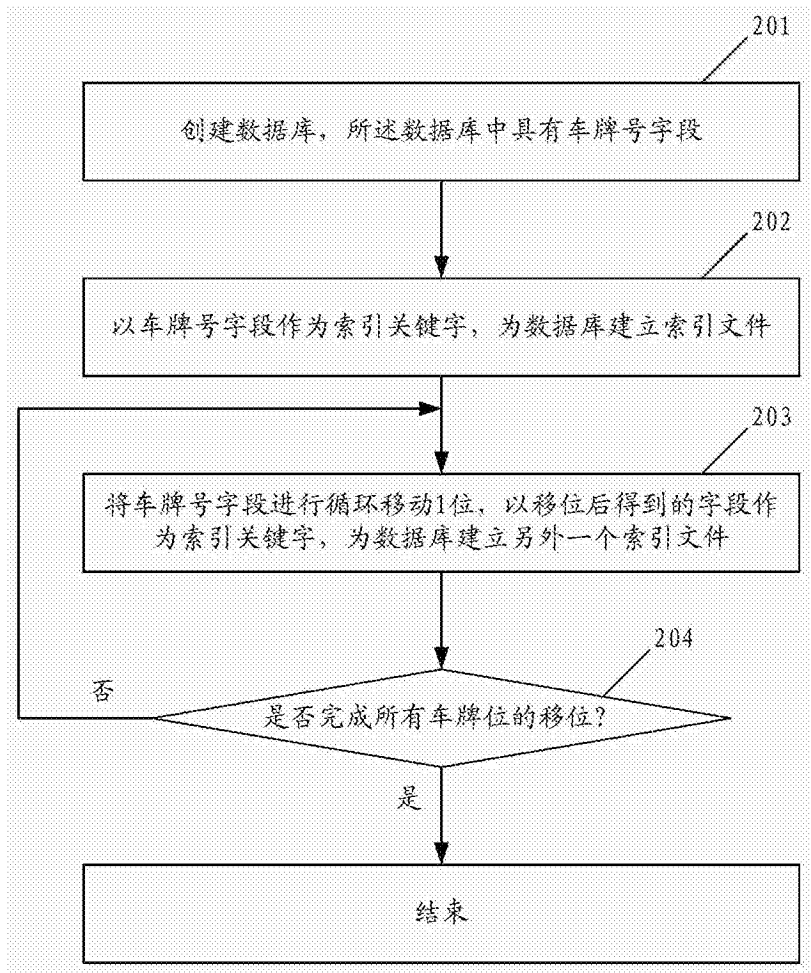


图 2

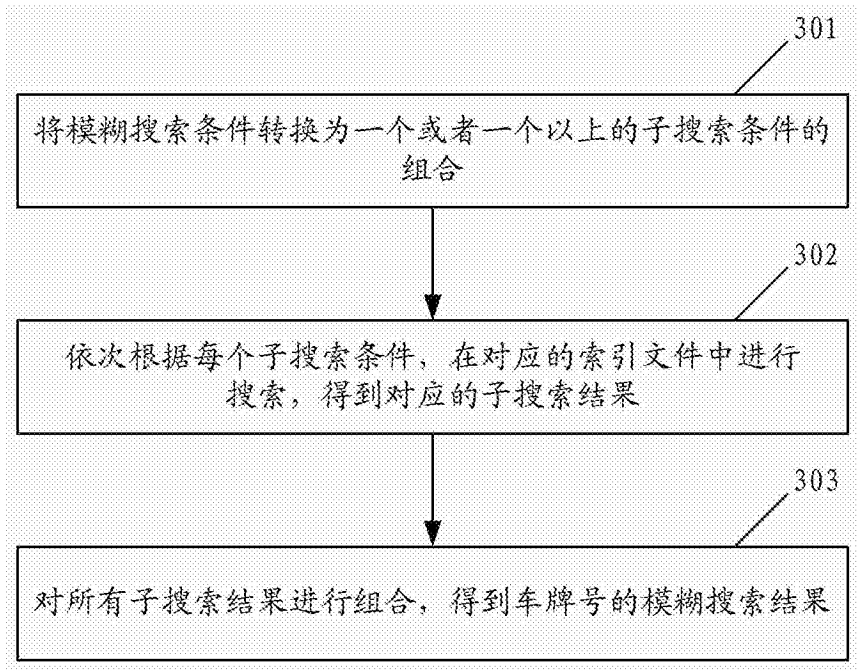


图 3

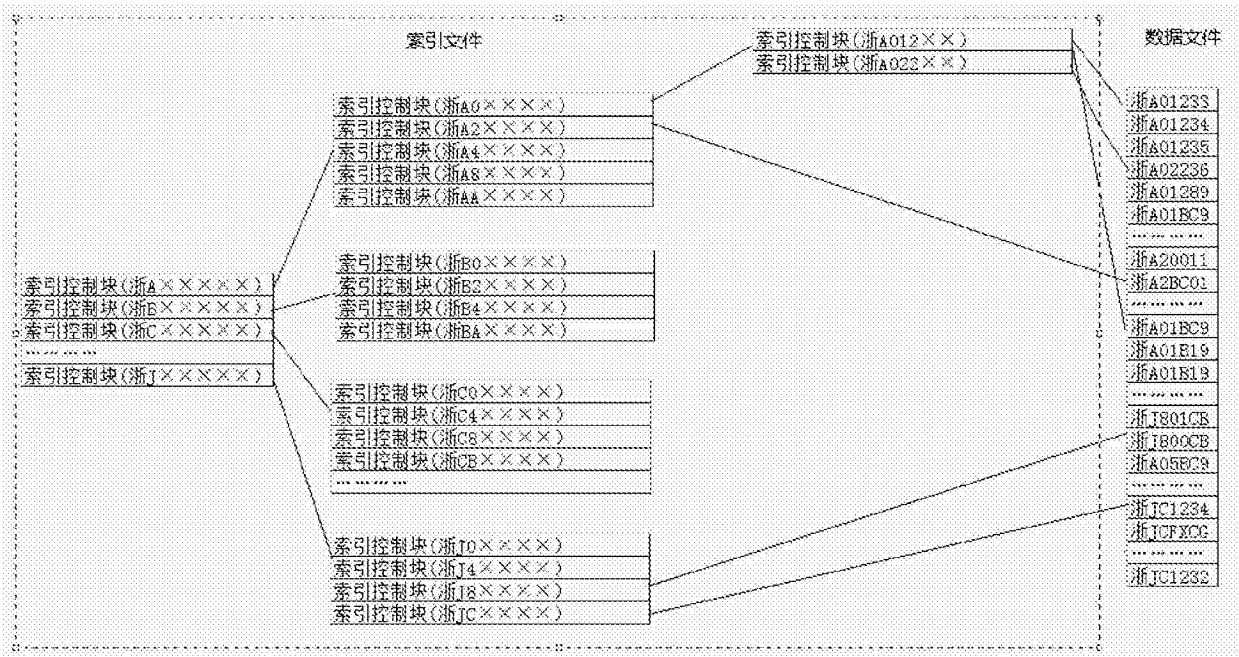


图 4

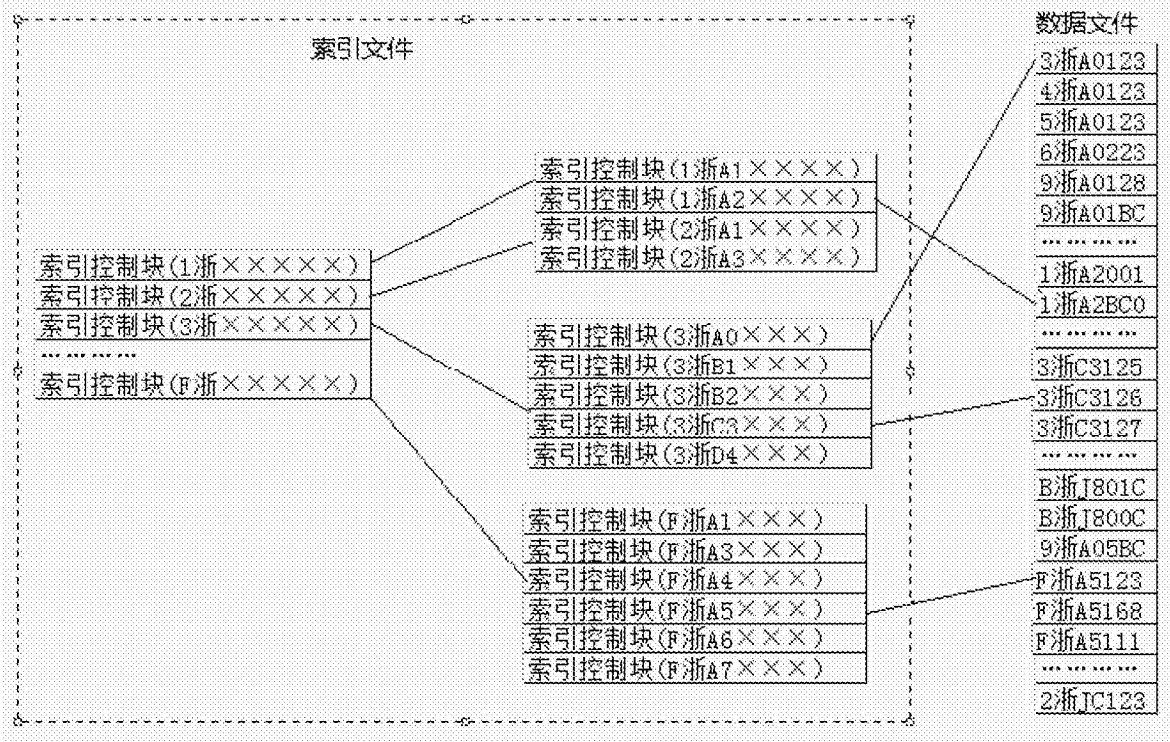


图 5

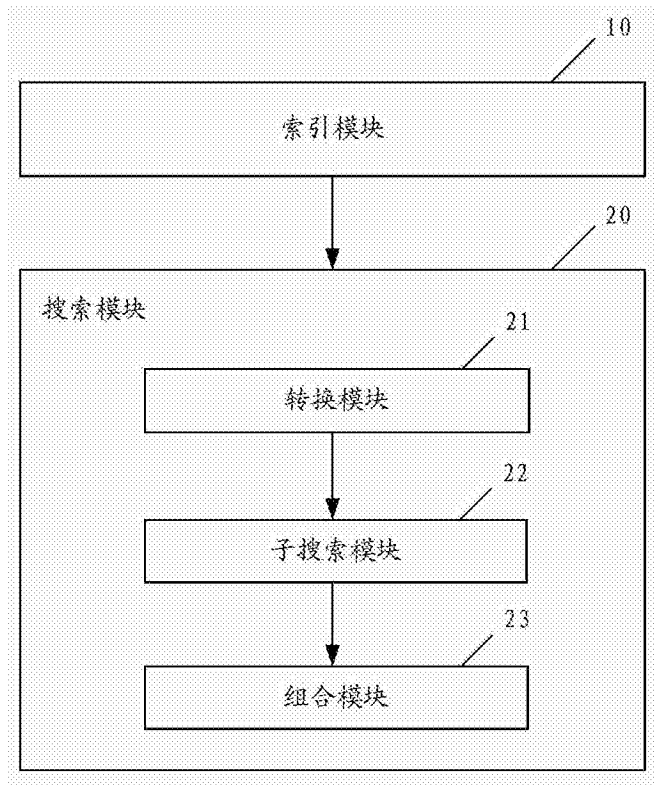


图 6