



(12)发明专利申请

(10)申请公布号 CN 106874386 A

(43)申请公布日 2017.06.20

(21)申请号 201710020706.8

(22)申请日 2017.01.12

(71)申请人 深圳市汉云科技有限公司

地址 518000 广东省深圳市南山区粤海街
道科技南十二路方大大厦203

(72)发明人 陈晓攀 熊志强 胡伟 梁嘉

(74)专利代理机构 深圳中一专利商标事务所

44237

代理人 阳开亮

(51)Int.Cl.

G06F 17/30(2006.01)

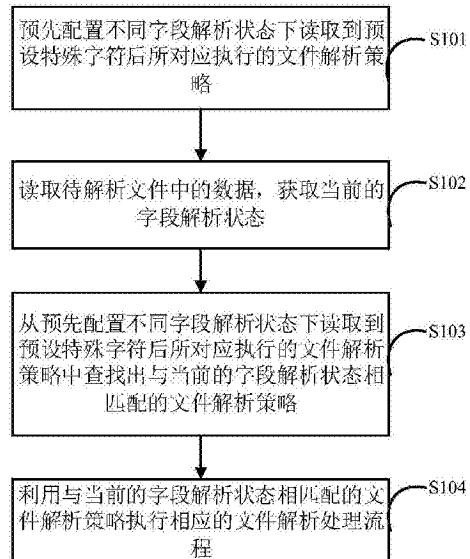
权利要求书3页 说明书6页 附图5页

(54)发明名称

文件解析方法及装置

(57)摘要

本发明适用于数据解析技术领域，提供了一种文件解析方法及装置，所述解析方法包括：预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略；读取待解析文件中的数据，获取当前的字段解析状态；从所述预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与所述当前的字段解析状态相匹配的文件解析策略；利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。本发明可以快速对文件进行解析，保证解析的高效性，并且具有高容错性，能够处理不规范的数据。



1. 一种文件解析方法,其特征在于,所述解析方法包括:

预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略;

读取待解析文件中的数据,获取当前的字段解析状态;

从所述预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与所述当前的字段解析状态相匹配的文件解析策略;

利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

2. 如权利要求1所述的解析方法,其特征在于,所述利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程包括:

当所述当前的字段解析状态为行结束状态时,读取下一个字符,若所述下一个字符为字段包括符,则所述当前的字段解析状态由所述行结束状态转变成字段开始状态;

当所述当前的字段解析状态为所述字段开始状态时,若接下来读取到的字符不是所述字段包括符,则所述当前的字段解析状态由所述字段开始状态转变成字段匹配状态;

当所述当前的字段解析状态为所述字段匹配状态时,根据与所述字段匹配状态相匹配的文件解析策略读取所述待解析文件中的数据,若读取到当前字符为所述字段包括符,则所述当前的字段解析状态由所述字段匹配状态转变成字段结束状态;

当所述当前的字段解析状态为所述字段结束状态时,读取下一个字符,若所述下一个字符不是字段分隔符或行结束符,则对所述当前字符之后的字符进行容错处理。

3. 如权利要求2所述的解析方法,其特征在于,所述当所述当前的字段解析状态为所述字段结束状态时,读取下一个字符,若所述下一个字符不是字段分隔符或行结束符,则对所述当前字符之后的字符进行容错处理包括:

当所述当前的字段解析状态为所述字段结束状态时,读取所述下一个字符,若所述下一个字符为所述字段包括符,检测所述待解析文件中是否有出错标签;

若有所述出错标签,则关闭容错标记功能,结束容错处理;

若没有所述出错标签,则打开所述容错标记功能,在文件解析出错的地方设置出错标签,且所述当前的字段解析状态由所述字段结束状态转变成所述字段匹配状态。

4. 如权利要求2所述的解析方法,其特征在于,所述当所述当前的字段解析状态为所述字段结束状态时,读取下一个字符,若所述下一个字符不是字段分隔符或行结束符,则对所述当前字符之后的字符进行容错处理还包括:

当所述当前的字段解析状态为所述字段结束状态时,读取所述下一个字符,若所述下一个字符为转义字符或普通字符,则检测所述待解析文件中是否有出错标签;

若有所述出错标签,则延长所述出错标签的标志范围;

若没有所述出错标签,则打开容错标记功能,在文件解析出错的地方设置所述出错标签,且所述当前的字段解析状态由所述字段结束状态转变成所述字段匹配状态。

5. 如权利要求2所述的解析方法,其特征在于,所述解析方法还包括:

当所述当前的字段解析状态为所述字段结束状态时,若所述下一个字符为所述字段分隔符或所述行结束符,检测所述待解析文件中是否有出错标签;

若没有所述出错标签,则所述当前的字段解析状态由所述字段结束状态转变成所述字段分割状态或所述行结束状态;

若有所述出错标签，则检测容错标记功能是否打开；
若打开，则延长所述出错标签的标志范围；
若没有打开，则所述当前的字段解析状态由所述字段结束状态转变成字段分割状态或所述行结束状态，并结束出错标签标记功能。

6. 一种文件解析装置，其特征在于，所述解析装置包括：

配置模块，用于预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略；

获取模块，用于读取待解析文件中的数据，获取当前的字段解析状态；

查找模块，用于从所述预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与所述当前的字段解析状态相匹配的文件解析策略；

执行模块，用于利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

7. 如权利要求6所述的解析装置，其特征在于，所述执行模块包括：

第一字段解析状态处理单元，用于当所述当前的字段解析状态为行结束状态时，读取下一个字符，若所述下一个字符为字段包括符，则所述当前的字段解析状态由所述行结束状态转变成字段开始状态；

第二字段解析状态处理单元，用于当所述当前的字段解析状态为所述字段开始状态时，若接下来读取到的字符不是所述字段包括符，则所述当前的字段解析状态由所述字段开始状态转变成字段匹配状态；

第三字段解析状态处理单元，用于当所述当前的字段解析状态为所述字段匹配状态时，根据与所述字段匹配状态相匹配的文件解析策略读取所述待解析文件中的数据，若读取到当前字符为所述字段包括符，则所述当前的字段解析状态由所述字段匹配状态转变成字段结束状态；

第四字段解析状态处理单元，用于当所述当前的字段解析状态为所述字段结束状态时，读取下一个字符，若所述下一个字符不是字段分隔符或行结束符，则对所述当前字符之后的字符进行容错处理。

8. 如权利要求7所述的解析装置，其特征在于，所述第四字段解析状态处理具体用于：

当所述当前的字段解析状态为所述字段结束状态时，读取所述下一个字符，若所述下一个字符为所述字段包括符，检测所述待解析文件中是否有出错标签；

若有所述出错标签，则关闭容错标记功能，结束容错处理；

若没有所述出错标签，则打开所述容错标记功能，在文件解析出错的地方设置所述出错标签，且所述当前的字段解析状态由所述字段结束状态转变成所述字段匹配状态。

9. 如权利要求7所述的解析装置，其特征在于，所述第四字段解析状态处理单元还用于：

当所述当前的字段解析状态为所述字段结束状态时，读取所述下一个字符，若所述下一个字符不为所述预设特殊字符，则检测所述待解析文件中是否有出错标签；

若有所述出错标签，则延长所述出错标签的标志范围；

若没有所述出错标签，则打开容错标记功能，在文件解析出错的地方设置所述出错标签，且所述当前的字段解析状态由所述字段结束状态转变成所述字段匹配状态。

10. 如权利要求7所述的解析装置,其特征在于,所述第四字段解析状态处理单元还用于:

当所述当前的字段解析状态为所述字段结束状态时,若所述下一个字符为所述字段分隔符或所述行结束符,检测所述待解析文件中是否有出错标签;

若没有所述出错标签,则所述当前的字段解析状态由所述行结束状态转变成所述字段分割状态或所述行结束状态;

若有所述出错标签,则检测容错标签是否打开;

若打开,则延长所述出错标签的标志范围;

若没有打开,则所述当前的字段解析状态由所述行结束状态转变成字段分割状态或所述行结束状态,并结束出错标签标记功能。

文件解析方法及装置

技术领域

[0001] 本发明属于数据解析技术领域,尤其涉及文件解析方法及装置。

背景技术

[0002] 现今,不同文本的格式规范各不相同,常见的CSV解析方法严重依赖与CSV文件的格式规范,即只能处理常规CSV文件,当CSV文件稍有不规范的地方,就会导致整个文件解析失败,容错性低,而为了提高容错性,会对一行数据进行多次扫描分析,降低解析效率。因此,现有技术中存在着解析文件时无法处理不规范数据,容错性低,且解析效率低的问题。

发明内容

[0003] 本发明实施例提供一种文件解析方法及装置,旨在解决现有技术中解析文件时无法处理不规范数据,容错性低,且解析效率低的问题。

[0004] 第一方面,本发明实施例提供了一种文件解析方法,所述解析方法包括:

[0005] 预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略;

[0006] 读取待解析文件中的数据,获取当前的字段解析状态;

[0007] 从所述预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与所述当前的字段解析状态相匹配的文件解析策略;

[0008] 利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

[0009] 第二方面,本发明实施例提供了一种文件解析装置,所述解析装置包括:

[0010] 配置模块,用于预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略;

[0011] 获取模块,用于读取待解析文件中的数据,获取当前的字段解析状态;

[0012] 查找模块,用于从所述预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与所述当前的字段解析状态相匹配的文件解析策略;

[0013] 执行模块,用于利用与所述当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

[0014] 本发明实施例与现有技术相比存在的有益效果是:本发明实施例通过预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略;读取待解析文件中的数据,获取当前的字段解析状态;从预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与当前的字段解析状态相匹配的文件解析策略;利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程,从而可以快速对文件进行解析,保证解析的高效性,并且具有高容错性,能够处理不规范的数据。

附图说明

- [0015] 图1为本发明实施例提供的一种文件解析方法的流程示意图；
- [0016] 图2为图1所示实施例中的步骤S104的细化步骤流程示意图；
- [0017] 图3为图2所示实施例中的步骤S204的细化步骤流程示意图；
- [0018] 图4为图2所示实施例中的步骤S204的另一细化步骤流程示意图；
- [0019] 图5为图1所示实施例中的步骤S104的另一细化步骤流程示意图；
- [0020] 图6为本发明实施例提供的一种文件解析装置的结构示意图；
- [0021] 图7为图6所示实施例中的执行模块604的细化结构示意图。

具体实施方式

[0022] 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。

[0023] 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。

[0024] 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。

[0025] 请参阅图1，图1为本发明实施例提供的一种文件解析方法的流程示意图。本实施例的执行主体为解析装置。本实施例提供的文件解析方法包括以下步骤：

[0026] 步骤S101、预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略。

[0027] 在本发明实施例中，预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略，并将该文件解析策略保存于解析装置中。

[0028] 其中，文件解析策略为对待解析文件中的数据进行读取、容错处理、保存等处理，并对字段解析状态进行状态转变的文件解析策略。

[0029] 步骤S102、读取待解析文件中的数据，获取当前的字段解析状态。

[0030] 在本发明实施例中，解析装置在执行文件解析的过程中，自动获取当前处于何种文件解析状态，在文件解析状态下读取到预设特殊字符之后，根据文件解析状态所对应的文件解析策略解析待解析的文件。

[0031] 步骤S103、从预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与当前的字段解析状态相匹配的文件解析策略。

[0032] 在本发明实施例中，在某个字段解析状态下，若解析装置读取到某个预设特殊字符，则从预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与上述某个字段解析状态相匹配的文件解析策略。

[0033] 步骤S104、利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

[0034] 在本发明实施例中，在某个字段解析状态下，解析装置利用与该字段解析状态相匹配的文件解析策略对待解析文件中的数据进行读取、容错处理、保存等处理，以及将该字段解析状态转变成其他字段解析状态。

[0035] 进一步地，图2示出了步骤S104中利用与当前的字段解析状态相匹配的文件解析

策略执行相应的文件解析处理流程的具体实现流程,详述如下:

[0036] 步骤S201、当当前的字段解析状态为行结束状态时,读取下一个字符,若下一个字符为字段包括符,则根据与行结束状态相匹配的文件解析策略确定当前的字段解析状态由行结束状态转变成字段开始状态。

[0037] 在本发明实施例中,当当前的字段解析状态为行结束状态时,读取下一个字符,并判断该字符是否为字段包括符,若是字段包括符,则根据与行结束状态相匹配的文件解析策略确定当前的字段解析状态由行结束状态转变成字段开始状态。若不是字段包括符,则根据与行结束状态相匹配的文件解析策略确定当前的字段解析状态由行结束状态转变成字段匹配状态。

[0038] 步骤S202、当当前的字段解析状态为字段开始状态时,若接下来读取到的字符不是字段包括符,则根据与字段开始状态相匹配的文件解析策略确定当前的字段解析状态由字段开始状态转变成字段匹配状态。

[0039] 在本发明实施例中,当当前的字段解析状态为字段开始状态时,若接下来读取到的字符是字段包括符,则根据与字段开始状态相匹配的文件解析策略确定当前的字段解析状态由字段开始状态转变成字段结束状态。若接下来读取到的字符不是字段包括符,则根据与字段开始状态相匹配的文件解析策略确定当前的字段解析状态由字段开始状态转变成字段匹配状态。

[0040] 步骤S203、当当前的字段解析状态为字段匹配状态时,根据与字段匹配状态相匹配的文件解析策略读取待解析文件中的数据,若读取到当前字符为字段包括符,则根据与字段匹配状态相匹配的文件解析策略确定当前的字段解析状态由字段匹配状态转变成字段结束状态。

[0041] 在本发明实施例中,当当前的字段解析状态为字段匹配状态时,根据与字段匹配状态相匹配的文件解析策略读取待解析文件中的数据,若读取到当前字符为字段包括符,则根据与字段匹配状态相匹配的文件解析策略确定当前的字段解析状态由字段匹配状态转变成字段结束状态。若读取到当前字符不为字段包括符,则根据与字段匹配状态相匹配的文件解析策略继续读取待解析文件中的数据,且不转变当前的字段解析状态。

[0042] 步骤S204、当当前的字段解析状态为字段结束状态时,读取下一个字符,若下一个字符不是字段分隔符或行结束符,则根据与字段结束状态相匹配的文件解析策略对当前字符之后的字符进行容错处理。

[0043] 进一步地,图3示出了步骤S204中当当前的字段解析状态为字段结束状态时,读取下一个字符,若下一个字符不是字段分隔符或行结束符,则根据与字段结束状态相匹配的文件解析策略对当前字符之后的字符进行容错处理的具体实现流程,详述如下:

[0044] 步骤S301、当当前的字段解析状态为字段结束状态时,读取下一个字符,若下一个字符为字段包括符,检测待解析文件中是否有出错标签。

[0045] 步骤S302、若有出错标签,则关闭容错标记功能,结束容错处理。

[0046] 步骤S303、若没有出错标签,则打开容错标记功能,在文件解析出错的地方设置出错标签,且当前的字段解析状态由字段结束状态转变成字段匹配状态。

[0047] 进一步地,图4示出了步骤S204中当当前的字段解析状态为字段结束状态时,读取下一个字符,若下一个字符不是字段分隔符或行结束符,则根据与字段结束状态相匹配的

文件解析策略对当前字符之后的字符进行容错处理的另一个具体实现流程，详述如下：

[0048] 步骤S401、当当前的字段解析状态为字段结束状态时，读取下一个字符，若下一个字符为转义字符或普通字符，则检测待解析文件中是否有出错标签。

[0049] 步骤S402、若有出错标签，则延长出错标签的标志范围。

[0050] 步骤S403、若没有出错标签，则打开容错标记功能，在文件解析出错的地方设置出错标签，且当前的字段解析状态由字段结束状态转变成字段匹配状态。

[0051] 进一步地，图5示出了步骤S104中利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程的另一具体实现流程，详述如下：

[0052] 步骤S501、当当前的字段解析状态为行结束状态时，读取下一个字符，若下一个字符为字段包括符，则当前的字段解析状态由行结束状态转变成字段开始状态。

[0053] 步骤S502、当当前的字段解析状态为字段开始状态时，若接下来读取到的字符不是字段包括符，则当前的字段解析状态由字段开始状态转变成字段匹配状态。

[0054] 步骤S503、当当前的字段解析状态为字段匹配状态时，根据与字段匹配状态相匹配的文件解析策略读取待解析文件中的数据，若读取到当前字符为字段包括符，则当前的字段解析状态由字段匹配状态转变成字段结束状态。

[0055] 步骤S504、当当前的字段解析状态为字段结束状态时，若下一个字符为字段分隔符或行结束符，检测待解析文件中是否有出错标签。

[0056] 步骤S505、若没有出错标签，则当前的字段解析状态由字段结束状态转变成字段分割状态或行结束状态。

[0057] 在本发明实施例中，若下一个字符为字段分隔符，且没有检测到出错标签，则当前的字段解析状态由字段结束状态转变成字段分割状态。若下一个字符为行结束符，且没有检测到出错标签，则当前的字段解析状态由字段结束状态转变成行结束状态。

[0058] 步骤S506、若有出错标签，则检测容错标记功能是否打开。

[0059] 步骤S507、若打开，则延长出错标签的标志范围。

[0060] 步骤S508、若没有打开，则当前的字段解析状态由字段结束状态转变成字段分割状态或行结束状态，并结束出错标签标记功能。

[0061] 在本发明实施例中，若下一个字符为字段分隔符，且检测到出错标签，则检测容错标签是否打开，若没有打开，则当前的字段解析状态由字段结束状态转变成字段分割状态，并结束出错标签标记功能。若下一个字符为行结束符，且检测到出错标签，则检测容错标签标记功能是否打开，若没有打开，则当前的字段解析状态由字段结束状态转变成行结束状态，并结束出错标签标记功能。

[0062] 进一步地，在转变成行结束状态，且结束出错标签后，结束当前行解析任务，保存当前行解析任务的解析结果及当前行解析任务的解析结果在解析文件中的位置。保存当前行解析任务中的原始数据及原始数据在原始文件中的位置。

[0063] 在本发明实施例中，通过上述实施方案，可以对出错的文件进行容错解析处理，避免对出错的数据进行多次解析。

[0064] 以上可以看出，本实施例提供的一种文件解析方法由于通过预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略；读取待解析文件中的数据，获取当前的字段解析状态；从预先配置不同字段解析状态下读取到预设特殊字符后所对应

执行的文件解析策略中查找出与当前的字段解析状态相匹配的文件解析策略；利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程，从而可以快速对文件进行解析，保证解析的高效性，并且具有高容错性，能够处理不规范的数据。

[0065] 请参阅图6，图6为本发明实施例提供的一种文件解析装置的结构示意图。本实施例的解析装置包括的各模块用于执行图1对应的实施例中的各步骤，具体请参阅图1以及图1对应的实施例中的相关描述，此处不赘述。本实施例的解析装置包括：配置模块601、获取模块602、查找模块603、执行模块604。具体地：

[0066] 配置模块601，用于预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略。

[0067] 获取模块602，用于读取待解析文件中的数据，获取当前的字段解析状态。

[0068] 查找模块603，用于从预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与当前的字段解析状态相匹配的文件解析策略。

[0069] 执行模块604，用于利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程。

[0070] 在本发明实施例中，在某个字段解析状态下，解析装置利用与该字段解析状态相匹配的文件解析策略对待解析文件中的数据进行读取、容错处理、保存等处理，以及将该字段解析状态转变成其他字段解析状态。

[0071] 进一步地，图7示出了执行模块604的具体结构示意图，执行模块604包括的各单元用于执行图2对应的实施例中的各步骤，具体请参阅图2以及图2对应的实施例中的相关描述，此处不赘述。本实施例的执行模块604包括：第一字段解析状态处理单元701、第二字段解析状态处理单元702、第三字段解析状态处理单元703、第四字段解析状态处理单元704。具体地：

[0072] 第一字段解析状态处理单元701、用于当当前的字段解析状态为行结束状态时，读取下一个字符，若下一个字符为字段包括符，则当前的字段解析状态由行结束状态转变成字段开始状态。

[0073] 第二字段解析状态处理单元702、用于当当前的字段解析状态为字段开始状态时，若接下来读取到的字符不是字段包括符，则当前的字段解析状态由字段开始状态转变成字段匹配状态。

[0074] 第三字段解析状态处理单元703、当当前的字段解析状态为字段匹配状态时，根据与字段匹配状态相匹配的文件解析策略读取待解析文件中的数据，若读取到当前字符为字段包括符，则当前的字段解析状态由字段匹配状态转变成字段结束状态。

[0075] 第四字段解析状态处理单元704、当当前的字段解析状态为字段结束状态时，读取下一个字符，若下一个字符不是字段分隔符或行结束符，则对当前字符之后的字符进行容错处理。

[0076] 进一步地，上述第四字段解析状态处理单元704具体用于：

[0077] 当当前的字段解析状态为字段结束状态时，读取下一个字符，若下一个字符为字段包括符，检测待解析文件中是否有出错标签。

[0078] 若有出错标签，则关闭容错标记功能，结束容错处理。

[0079] 若没有出错标签，则打开容错标记功能，在文件解析出错的地方设置出错标签，且

当前的字段解析状态由字段结束状态转变成字段匹配状态。

[0080] 进一步地,上述第四字段解析状态处理单元704,还用于:

[0081] 当当前的字段解析状态为字段结束状态时,读取下一个字符,若下一个字符为转义字符或普通字符,检测待解析文件中是否有出错标签。

[0082] 若有出错标签,则延长出错标签的标志范围。

[0083] 若没有出错标签,则打开容错标记功能,在文件解析出错的地方设置出错标签,且当前的字段解析状态由字段结束状态转变成字段匹配状态。

[0084] 进一步地,上述第四字段解析状态处理单元704,还用于:

[0085] 当当前的字段解析状态为字段结束状态时,若下一个字符为字段分隔符或行结束符,检测待解析文件中是否有出错标签。

[0086] 若没有出错标签,则当前的字段解析状态由行结束状态转变成字段分割状态或行结束状态。

[0087] 若有出错标签,则检测容错标签是否打开。

[0088] 若打开,则延长出错标签的标志范围。

[0089] 若没有打开,则当前的字段解析状态由行结束状态转变成字段分割状态或行结束状态,并结束出错标签标记功能。

[0090] 在本发明实施例中,通过上述实施方案,可以对出错的文件进行容错解析处理,避免对出错的数据进行多次解析。

[0091] 以上可以看出,本实施例提供的一种文件解析装置由于通过预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略;读取待解析文件中的数据,获取当前的字段解析状态;从预先配置不同字段解析状态下读取到预设特殊字符后所对应执行的文件解析策略中查找出与当前的字段解析状态相匹配的文件解析策略;利用与当前的字段解析状态相匹配的文件解析策略执行相应的文件解析处理流程,从而可以快速对文件进行解析,保证解析的高效性,并且具有高容错性,能够处理不规范的数据。

[0092] 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作的任何修改、等同替换和改进等,均应包含在本发明的保护范围之内。

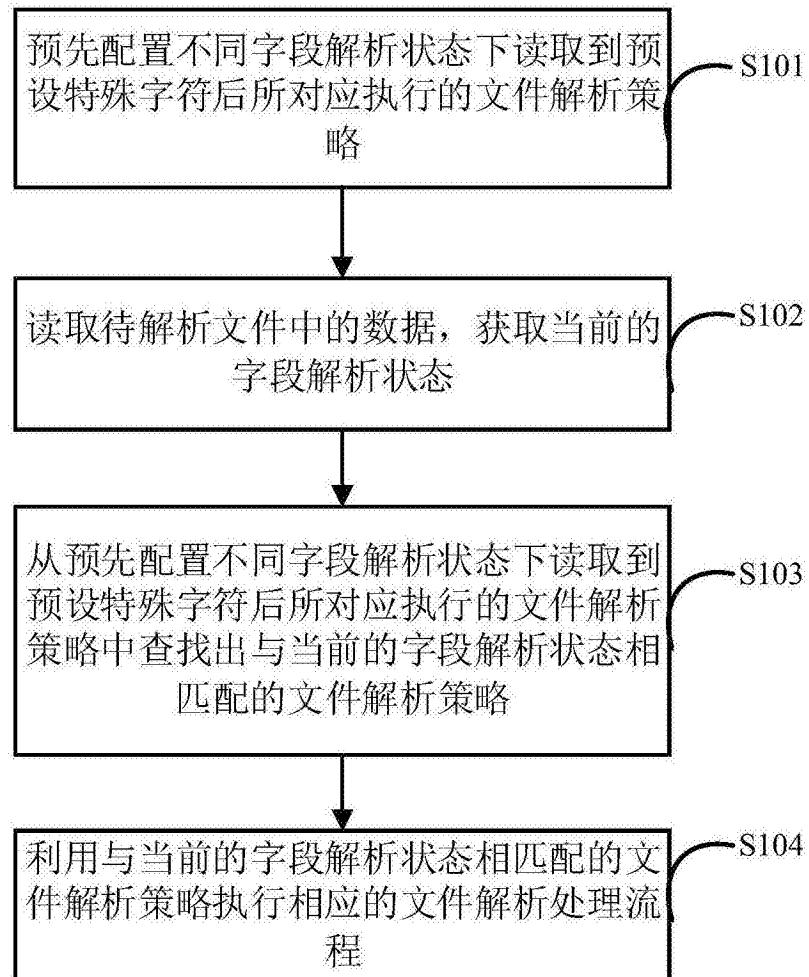


图1

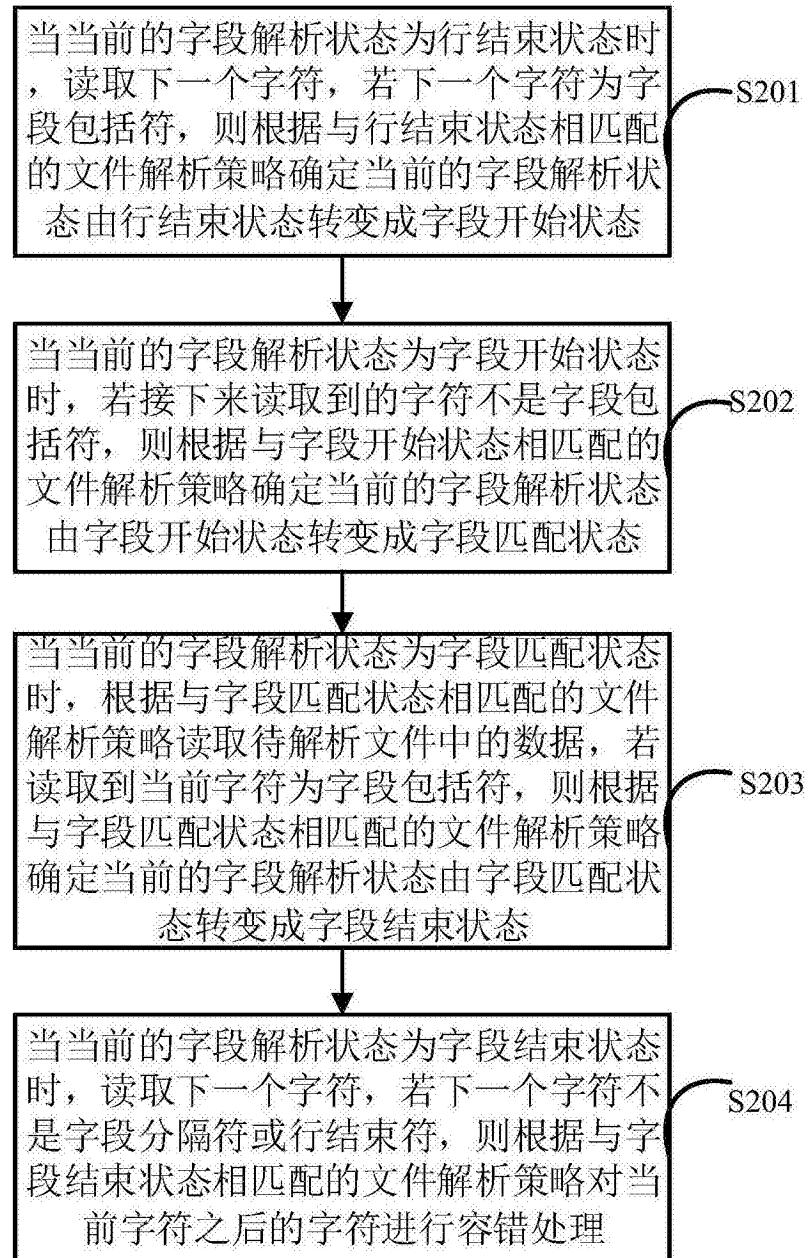


图2

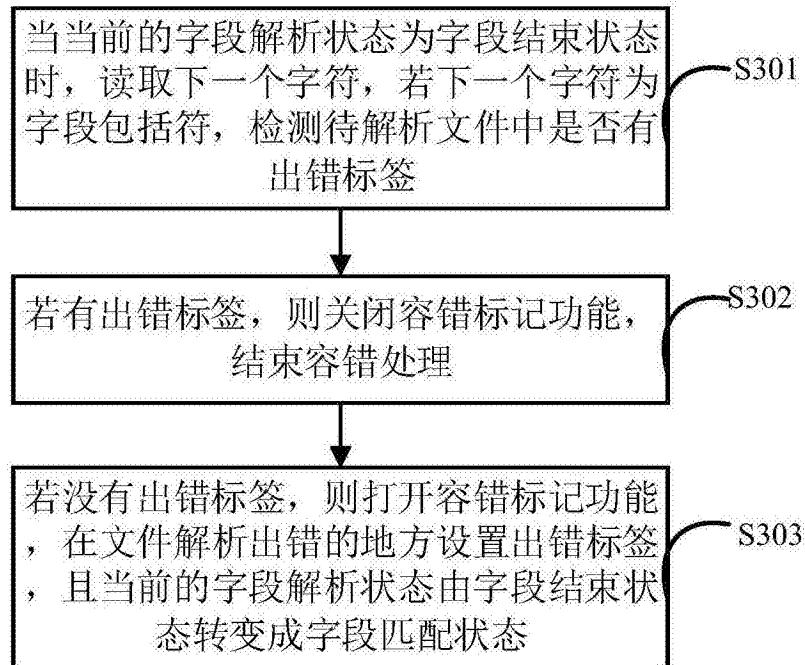


图3

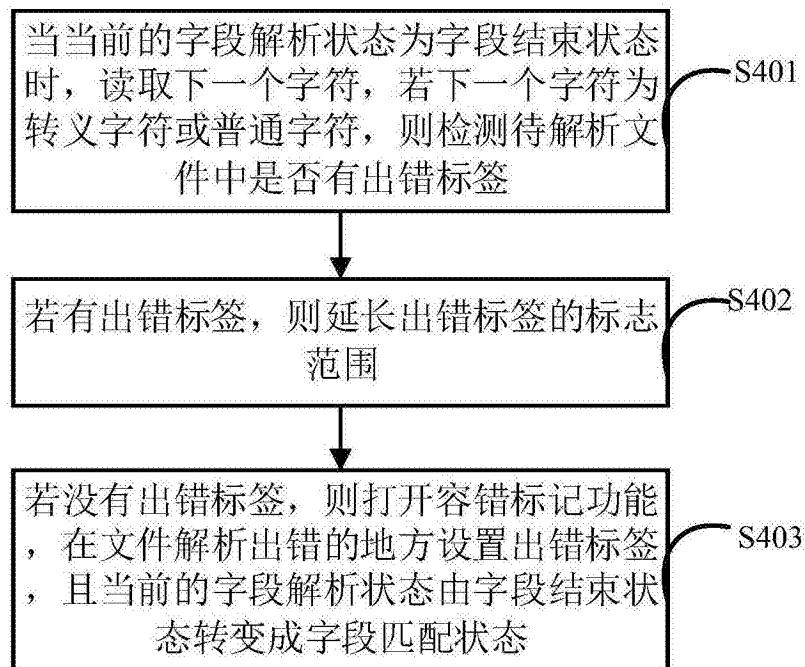


图4

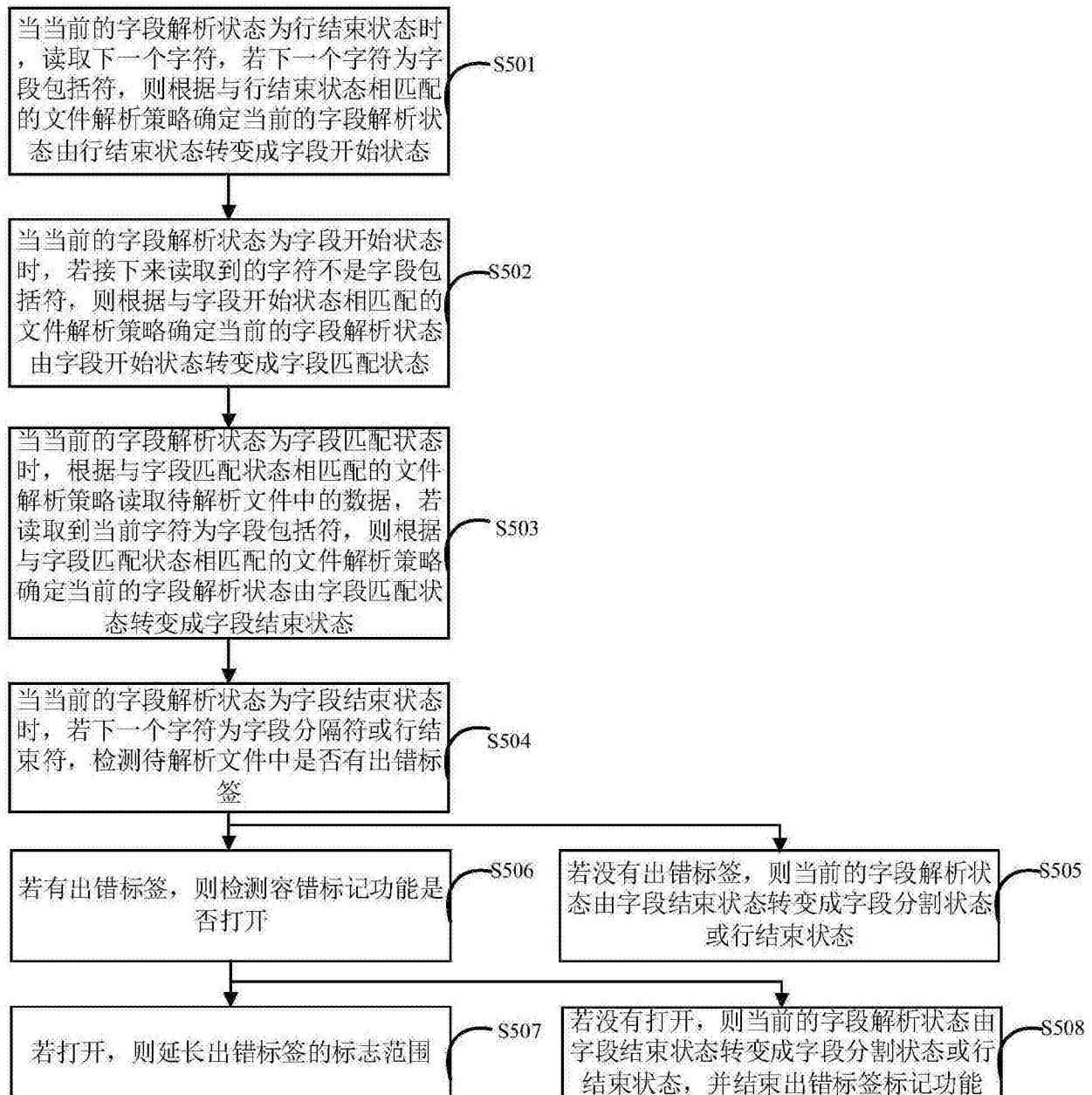


图5

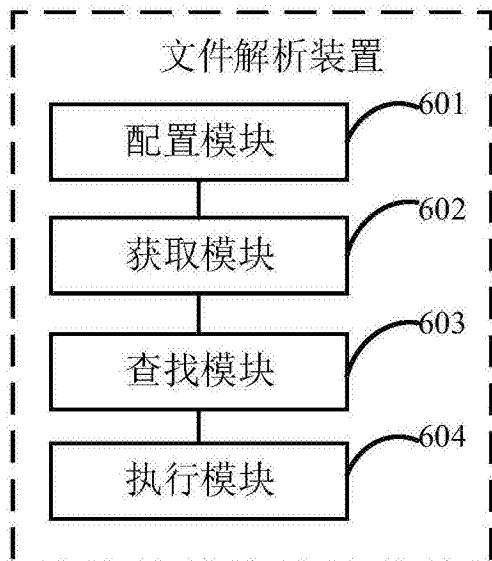


图6

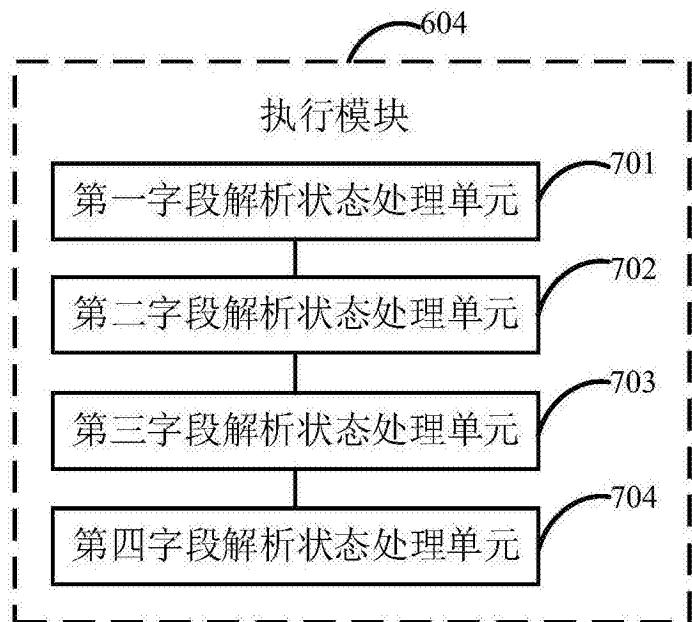


图7