



(12)发明专利



(10)授权公告号 CN 104657370 B

(45)授权公告日 2018.09.04

(21)申请号 201310587065.6

(22)申请日 2013.11.19

(65)同一申请的已公布的文献号
申请公布号 CN 104657370 A

(43)申请公布日 2015.05.27

(73)专利权人 中国移动通信集团天津有限公司
地址 300300 天津市东丽区空港经济加工
区中央大道99号

(72)发明人 刘玉娜

(74)专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 张颖玲 张振伟

(51)Int.Cl.

G06F 17/30(2006.01)

(56)对比文件

CN 1726495 A,2006.01.25,

CN 1564160 A,2005.01.12,

CN 101799830 A,2010.08.11,

CN 102207940 A,2011.10.05,

CN 102737123 A,2012.10.17,

US 2004139061 A1,2004.07.15,

CN 101178732 A,2008.05.14,

US 2006116976 A1,2006.06.01,

US 2012317137 A1,2012.12.13,

李建中等.一种数据仓库的多维数据模型.

《软件学报》.2000,第11卷(第7期),第908-917
页.

刘红敏.数据立方体生成技术研究.《中国优秀硕士学位论文全文数据库 信息科技辑》
.2007,(第05期),第I138-987页.

审查员 王明芳

权利要求书1页 说明书7页 附图2页

(54)发明名称

一种实现多维立方体关联的方法和装置

(57)摘要

本发明公开了一种实现多维立方体(Cube)关联的方法,确定需要关联的Cube没有共同维度;对Cube的源表进行追踪,获取明细表;根据获得的明细表,生成动态Cube。本发明同时还公开了一种实现Cube关联的装置。通过本发明的方案,能够实现无公共维度的Cube关联,为用户提供更加便捷的数据分析。

101、服务器确定需要关联的Cube没有共同
维度

102、服务器对所述Cube的源表进行追踪,获
取明细表

103.服务器根据所述明细表,生成动态Cube

1. 一种实现多维立方体Cube关联的方法,其特征在于,所述方法包括:

服务器确定需要关联的Cube没有共同维度,对所述Cube的源表进行追踪,获取并解析明细表,获得相同的参数信息,通过所述参数信息,对所述Cube的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态Cube。

2. 根据权利要求1所述的方法,其特征在于,所述对所述Cube的源表进行追踪,获取并解析明细表,包括:

所述服务器根据需要关联的Cube的元数据,查询需要关联的Cube的生成过程,获取生成需要关联的Cube的源表,根据所述源表得到明细表。

3. 根据权利要求2所述的方法,其特征在于,所述获取并解析明细表之后,还包括:

记录组成所述明细表的参数信息;其中,所述参数信息包括主键、维度和度量信息。

4. 根据权利要求3所述的方法,其特征在于,所述对所述Cube的明细表进行关联包括:

服务器通过相同的主键信息,对生成需要关联的Cube的明细表进行关联。

5. 根据权利要求3或4所述的方法,其特征在于,所述根据所述虚拟明细表生成动态Cube包括:通过任务Task将虚拟明细表生成事实表,再使用结构化查询语言SQL语句将所述事实表的维度信息和度量信息结合生成动态Cube。

6. 一种实现多维立方体Cube关联的装置,其特征在于,所述装置包括:判断模块、源表分析模块和生成模块;其中,

判断模块,用于确定需要关联的Cube没有共同维度;

源表分析模块,用于对所述Cube的源表进行追踪,获取明细表;

生成模块,用于解析所述明细表,获得相同的参数信息,通过所述参数信息,对Cube的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表,生成动态Cube。

7. 根据权利要求6所述的装置,其特征在于,所述判断模块,具体用于比较需要关联的Cube的维度信息,在没有相同的维度信息时,确定所述需要关联的Cube没有共同维度,并通知所述源表分析模块。

8. 根据权利要求6所述的装置,其特征在于,所述源表分析模块,具体用于根据需要关联的Cube的元数据,查询所述Cube的生成过程,获取生成所述Cube的源表,根据所述源表得到明细表。

9. 根据权利要求6所述的装置,其特征在于,所述生成模块,具体用于对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的Cube的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态Cube;其中,所述参数信息包括主键、维度和度量信息。

10. 根据权利要求9所述的装置,其特征在于,所述生成模块包括:虚拟明细表生成模块以及Cube生成模块;其中,

虚拟明细表生成模块,用于将明细表进行关联,生成虚拟明细表;

Cube生成模块,用于根据所述虚拟明细表,生成动态Cube。

一种实现多维立方体关联的方法和装置

技术领域

[0001] 本发明涉及分布式统计技术,尤其是涉及一种实现多维立方体(Cube)关联的方法和装置。

背景技术

[0002] 联机分析处理(OLAP)是由关系数据库之父E.F.Codd于1993年提出的,Codd认为联机事务处理(OLTP)已不能满足终端用户对数据库查询分析的要求,结构化查询语言(SQL)对大数据库的简单查询也不能满足用户分析的需求。用户的决策分析需要对关系数据库进行大量计算才能得到结果,而查询的结果并不能满足决策者提出的需求。因此,Codd提出了多维数据库和多维分析的概念,即OLAP。OLAP委员会对联机分析处理的定义为:从原始数据中转化出来的、能够真正为用户所理解的、并真实反映企业多维特性的数据称为信息数据,使分析人员、管理人员或执行人员能够从多种角度对信息数据进行快速、一致、交互地存取,从而获得对数据的更深入了解的一类软件技术。

[0003] 现有的多维分析方案主要有以下两种:

[0004] ROLAP:基于关系数据库的OLAP实现(Relational OLAP);以关系数据库为核心,以关系型结构进行多维数据的表示和存储。ROLAP将多维数据库的多维结构划分为两类表:一类是事实表,用来存储数据和维关键字;另一类是维表,即对每个维至少使用一个表来存放维的层次、成员类别等维的描述信息。维表和事实表通过主关键字和外关键字联系在一起,形成了“星型模式”。

[0005] MOLAP:基于多维数据组织的OLAP实现(Multidimensional OLAP);以多维数据组织方式为核心,也就是说,MOLAP使用多维数组存储数据。多维数据在存储中将形成“Cube”的结构,在MOLAP中对“Cube”的“旋转”、“切块”、“切片”是产生多维数据报表的主要技术。

[0006] 随着数据量的增大,上述方案中生成的Cube已不能满足用户的需求,这就需要两个或者多个Cube进行关联,以得到用户所需要的信息。

[0007] 图1示出了Cube的生成过程:先由接口表通过任务(Task)生成汇总表,汇总表通过Task生成明细表,明细表再通过Task生成事实表,最后事实表通过Task生成Cube;当需要对两个或多个具有共同维度的Cube进行关联时,会形成大数据运算,例如将两个Cube的数据汇总到一个Cube中,会通过两个Cube的共同维度进行关联,然后通过关联后的数据,计算各个Cube的明细表。

[0008] 因此,现有的两个方案中,在对两个Cube实现关联分析时都存在有共同的缺点:实现两个Cube的关联分析,必须要求这两个Cube具有公共维度,无公共维度的Cube无法实现关联分析。

发明内容

[0009] 本发明实施例主要提供一种实现Cube关联的方法和装置,能够实现无公共维度的Cube关联,提高多Cube关联分析的效率。

- [0010] 本发明的技术方案是这样实现的：
- [0011] 本发明提供一种实现Cube关联的方法，所述方法包括：
- [0012] 服务器确定需要关联的Cube没有共同维度，对所述Cube的源表进行追踪，获取明细表，根据所述明细表生成动态Cube。
- [0013] 上述方案中，所述服务器对所述Cube的源表进行追踪，获取明细表，包括：
- [0014] 所述服务器根据需要关联的Cube的元数据，查询需要关联的Cube的生成过程，获取生成需要关联的Cube的源表，根据所述源表得到明细表。
- [0015] 上述方案中，所述根据所述明细表，生成动态Cube，包括：
- [0016] 服务器对所述明细表进行解析，记录组成所述明细表的参数信息，通过所述参数信息，对生成需要关联的Cube的明细表进行关联，得到虚拟明细表，根据所述虚拟明细表生成动态Cube；其中，所述参数信息包括主键、维度和度量信息。
- [0017] 上述方案中，所述对生成需要关联的Cube的明细表进行关联包括：
- [0018] 服务器通过相同的主键信息，对生成需要关联的Cube的明细表进行关联。
- [0019] 上述方案中，所述根据所述虚拟明细表生成动态Cube包括：通过任务Task将虚拟明细表生成事实表，再使用结构化查询语言SQL语句将所述事实表的维度信息和度量信息结合生成动态Cube。
- [0020] 本发明提供一种实现Cube关联的装置，所述装置包括：判断模块、源表分析模块和生成模块；其中，
- [0021] 判断模块，用于确定需要关联的Cube没有共同维度；
- [0022] 源表分析模块，用于对所述Cube的源表进行追踪，获取明细表；
- [0023] 生成模块，用于根据所述明细表，生成动态Cube。
- [0024] 上述方案中，所述判断模块，具体用于比较需要关联的Cube的维度信息，在没有相同的维度信息时，确定所述需要关联的Cube没有共同维度，并通知所述源表分析模块。
- [0025] 上述方案中，所述源表分析模块，具体用于根据需要关联的Cube的元数据，查询所述Cube的生成过程，获取生成所述Cube的源表，根据所述源表得到明细表。
- [0026] 上述方案中，所述生成模块，具体用于对所述明细表进行解析，记录组成所述明细表的参数信息，通过所述参数信息，对生成需要关联的Cube的明细表进行关联，得到虚拟明细表，根据所述虚拟明细表生成动态Cube；其中，所述参数信息包括主键、维度和度量信息。
- [0027] 上述方案中，所述生成模块包括：虚拟明细表生成模块以及Cube生成模块；其中，
- [0028] 虚拟明细表生成模块，用于将明细表进行关联，生成虚拟明细表；
- [0029] Cube生成模块，用于根据所述虚拟明细表，生成动态Cube。
- [0030] 本发明实施例提供一种实现Cube关联的方法和装置，当需要关联的Cube没有共同维度时，服务器通过对生成所述Cube的源表进行追踪，获取明细表；然后根据获取的明细表生成动态Cube；如此，能够实现无公共维度的Cube关联，提高多Cube关联分析的效率，为用户分析数据提供便利；并且通过追踪源表数据的方法，从而减小Cube关联时的系统性能消耗和执行时间。

附图说明

- [0031] 图1为现有技术中生成Cube的方法的流程示意图；

[0032] 图2为本发明实施例一提供的实现Cube关联的方法流程示意图；

[0033] 图3为本发明实施例二提供的实现Cube关联的装置的结构示意图。

具体实施方式

[0034] 本发明实施例中,当需要关联的Cube没有共同维度时,服务器对所述需要关联的Cube的源表进行追踪,获取明细表,然后对明细表进行解析,获得相同的参数信息,并通过相同的参数信息,生成动态Cube,从而实现无公共维度的Cube关联,解决了没有共同维度的Cube不能关联的问题。

[0035] 其中,所述需要关联的Cube,在生成过程中,具有共同的参数信息;所述参数信息包括主键信息。

[0036] 下面通过附图及具体实施例对本发明做进一步的详细说明。

[0037] 实施例一

[0038] 本发明实施例一实现Cube关联的方法,如图2所示,该方法包括以下几个步骤:

[0039] 步骤101:服务器确定需要关联的Cube没有共同维度。

[0040] 具体的,服务器比较需要关联的Cube的维度信息,在没有相同的维度信息时,确定所述需要关联的Cube没有共同维度;所述服务器可以是用于数据管理的服务器,如移动公司后台的数据管理服务器。

[0041] 在具有相同的维度信息时,所述服务器直接通过所述维度信息对所述需要关联的Cube进行关联。

[0042] 例如,对于表1中的Cube1和表2中的Cube2,当用户需要对各个地区的各个品牌的用户数量和收入进行统计查看时,就需要对Cube1和Cube2进行关联,生成一个具有品牌信息、收入信息、在网客户数信息和收入信息的新的Cube;通过查询表3所示的Cube1的事实表1和表4所示的Cube2的事实表2,比较Cube1和Cube2是否具有共同的维度信息,由于Cube1的事实表1的维度信息为品牌信息,Cube2的维度信息为地区信息,可知Cube1和Cube2没有共同维度,所以通过Cube1的维度信息或Cube2的维度信息无法实现Cube1和Cube2的直接关联。

Cube1	
品牌	总收入
全球通	27490.672
动感地带	13878.082
神州行	23091.921

[0044] 表1

[0045]	Cube2	
	地区	在网客户数
	和平河西	327490
	南开红桥	113878
	河北河东	423091
	。 。 。 。	。 。 。 。

[0046] 表2

[0047]	事实表 1		
	属性	类型	说明
	品牌	varchar2 (10)	维度
	总收入	number	度量

[0048] 表3

[0049]	事实表 2		
	属性	类型	说明
	地区	varchar2 (10)	维度
	在网客户数	number	度量

[0050] 表4

[0051] 步骤102:服务器对所述Cube的源表进行追踪,获取明细表。

[0052] 具体的,所述服务器根据所述Cube的元数据,查询所述Cube的生成过程,获取生成所述Cube的源表,根据所述源表得到明细表;其中,所述源表包括明细表和事实表。

[0053] 所述元数据为记录Cube的生成过程以及底层数据逻辑关系的数据,通过所述元数据,可以追踪到生成Cube的源表,所述源表包括明细表和事实表;图1中示出了所述元数据记录的Cube的生成过程:所述Cube是由明细表通过Task生成事实表,再由所述事实表通过Task生成的;因此,通过所述元数据记录的Cube的生成过程,可以追踪到生成所述Cube的源表,从而获得明细表。

[0054] 对于步骤101中的Cube1和Cube2,本步骤可以通过元数据,由Cube1的事实表1追踪到表5所示的Cube1的明细表1,由Cube2的事实表2追踪到表6所示的Cube2的明细表2。

[0055]	明细表 1		
	属性	类型	说明
	用户标识	char (20)	主键
	品牌	varchar2 (10)	维度
	总收入	number	度量

[0056] 表5

明细表 2		
属性	类型	说明
用户标识	char (20)	主键
地区	varchar2 (10)	维度
在网客户数	number	度量

[0058] 表6

[0059] 步骤103:服务器根据所述明细表,生成动态Cube。

[0060] 具体的,服务器对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的Cube的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态Cube;其中,所述参数信息包括主键、维度和度量信息。

[0061] 通过所述参数信息,对生成需要关联的Cube的明细表进行关联,一般可以是服务器通过相同的主键信息,对生成需要关联的Cube的明细表进行关联;

[0062] 服务器通过Task将虚拟明细表生成事实表,使用SQL语句将所述事实表的维度信息和度量信息结合生成动态Cube。

[0063] 例如,服务器可以对步骤102中的Cube1的明细表1和Cube2的明细表2进行解析,记录明细表1和明细表2的参数信息,对于Cube1:记录主键信息为用户标识信息,记录维度信息为品牌信息,记录度量信息为总收入信息;对于Cube2:记录主键信息为用户标识信息,记录维度信息为地区信息,记录度量信息为在网客户数信息;通过比较明细表1和明细表2得知,明细表1和明细表2具有相同的主键信息(即用户标识),通过相同的主键信息,对明细表1和明细表2进行关联,生成表7所示的虚拟明细表1,将虚拟明细表1结合记录的维度信息(品牌和地区)和度量信息(总收入和在网客户数),生成表8所示的动态Cube3,表8中包含了各个地区的各个品牌的用户数量和收入,从而满足了用户的需求。

虚拟明细表 1		
属性	类型	说明
用户标识	char (20)	主键
品牌	varchar2(10)	维度
地区	varchar2(10)	维度
总收入	number	度量
在网客户数	number	度量

[0065] 表7

动态 Cube3			
品牌	地区	总收入	在网客户数
[0066] 全球通	和平河西	49.632	227
	南开红桥	21.56	878
	河北河东	48.091	391
[0067]		17.378	878
	和平河西	82.091	291
	南开红桥	71.878	178
	河北河东	30.91	482
	。。。。	23.67	765
	和平河西	23.9	123
	南开红桥	20.1	156
	河北河东	35.21	231
	。。。。	83.25	67

[0068] 表8

[0069] 实施例二

[0070] 本发明实施例二提供了一种实现Cube关联的装置,如图3所示,所述装置包括:判断模块301、源表分析模块302以及生成模块303;其中,

[0071] 判断模块301,用于确定需要关联的Cube没有共同维度;

[0072] 源表分析模块302,用于对所述Cube的源表进行追踪,获取明细表;

[0073] 生成模块303,用于根据所述明细表,生成动态Cube;

[0074] 其中,所述需要关联的Cube为:在生成过程中,具有相同的参数信息;所述参数信息包括主键、维度和度量信息;

[0075] 具体的,所述判断模块301比较需要关联的Cube的维度信息,在没有相同的维度信息时,确定所述需要关联的Cube没有共同维度,并通知所述源表分析模块302;

[0076] 其中,在具有相同的维度信息时,直接通过所述维度信息对所述需要关联的Cube进行关联;

[0077] 所述源表分析模块302根据需要关联的Cube的元数据,查询所述Cube的生成过程,获取生成所述Cube的源表,根据所述源表得到明细表;其中,所述源表包括明细表和事实表;

[0078] 其中,所述元数据为记录Cube的生成过程以及底层数据逻辑关系的数据,通过所述元数据,可以追踪到生成Cube的明细表。

[0079] 所述生成模块303,具体用于对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的Cube的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态Cube;其中,所述参数信息包括主键、维度和度量信息;

[0080] 具体的,所述生成模块303通过相同的主键信息,对生成需要关联的Cube的明细表进行关联,得到虚拟明细表,通过Task将虚拟明细表生成事实表,使用SQL语句将所述事实

表的维度信息和度量信息结合生成动态Cube。

[0081] 其中,所述判断模块301可以由服务器的CPU进行实现,或者由对比电路进行实现;

[0082] 所述源表分析模块302可以由服务器的CPU和内存或硬盘进行实现;

[0083] 所述生成模块303可以由服务器的CPU和显示器进行实现。

[0084] 所述生成模块303包括:虚拟明细表生成模块401以及Cube生成模块402;其中,

[0085] 虚拟明细表生成模块401,用于将明细表进行关联,生成虚拟明细表;

[0086] Cube生成模块402,用于根据所述虚拟明细表,生成动态Cube;

[0087] 具体的,所述虚拟明细表生成模块401通过相同的主键信息,对生成需要关联的Cube的明细表进行关联,得到虚拟明细表;所述Cube生成模块402通过Task将虚拟明细表生成事实表,使用SQL语句将所述事实表的维度信息和度量信息结合生成动态Cube。

[0088] 通过本发明所提供的方法和装置,能够对无公共维度的Cube进行关联;通过追踪到生成所述Cube的明细表,查询相同的主键信息,通过所述主键进行关联合并生成虚拟明细表,再将所述虚拟明细表生成动态Cube,实现对无公共维度的Cube的关联,提高系统的运行效率,为用户提供所需要关联的相关信息。

[0089] 以上所述,仅为本发明的较佳实施例而已,并非用于限定本发明的保护范围,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。

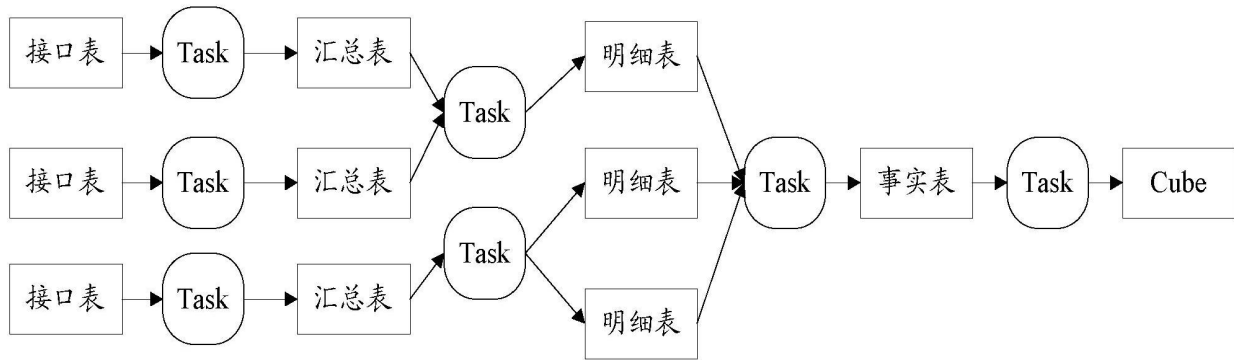


图1

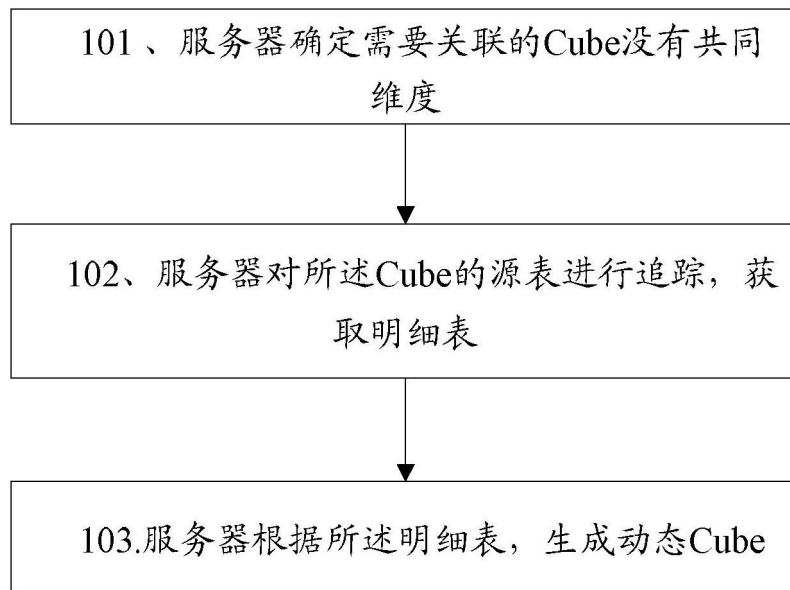


图2

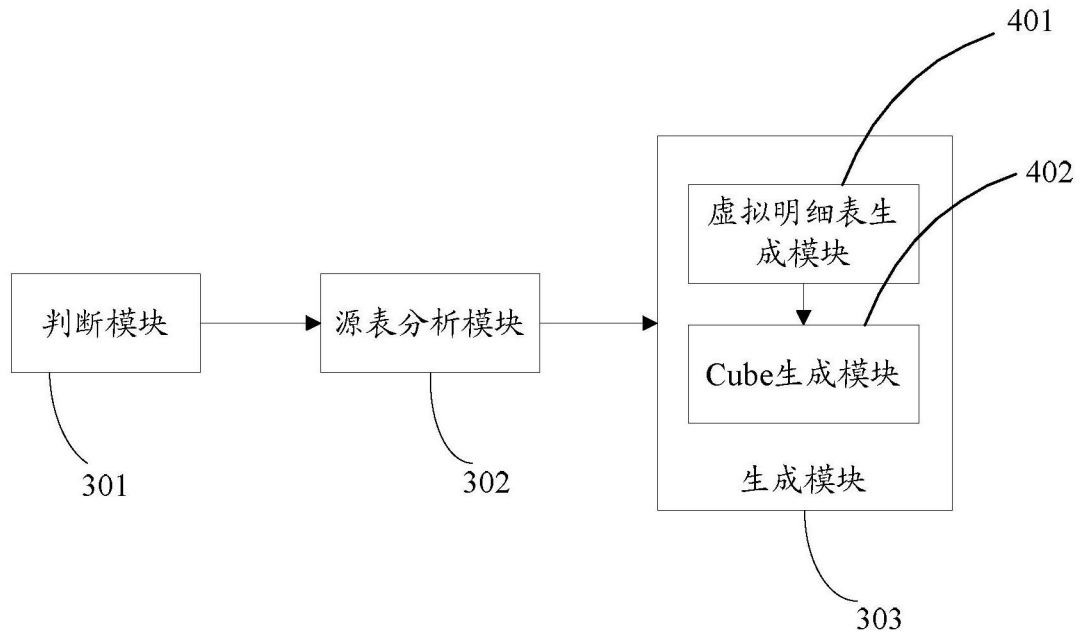


图3