



# (12) 发明专利

(10) 授权公告号 CN 110727793 B

(45) 授权公告日 2023.03.24

(21) 申请号 201810686490.3

G06F 16/33 (2019.01)

(22) 申请日 2018.06.28

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 110727793 A

CN 103605729 A, 2014.02.26

CN 103853700 A, 2014.06.11

CN 105701204 A, 2016.06.22

(43) 申请公布日 2020.01.24

US 2011231129 A1, 2011.09.22

(73) 专利权人 百度在线网络技术(北京)有限公司

CN 107103037 A, 2017.08.29

US 2013103697 A1, 2013.04.25

地址 100085 北京市海淀区上地十街10号  
百度大厦三层

US 2015095333 A1, 2015.04.02

US 2013041581 A1, 2013.02.14

(72) 发明人 黄晓莹 徐瑜 徐云峰

审查员 夏容

(74) 专利代理机构 北京英赛嘉华知识产权代理  
有限责任公司 11204

专利代理师 王达佐 马晓亚

(51) Int. Cl.

G06F 16/35 (2019.01)

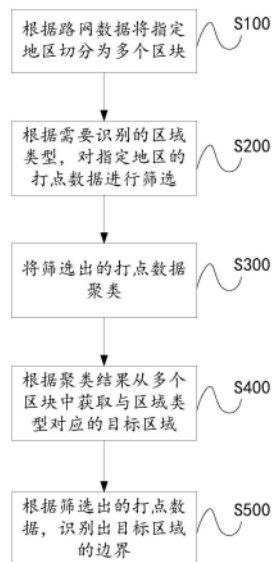
权利要求书2页 说明书7页 附图3页

## (54) 发明名称

区域识别的方法、装置、终端和计算机可读存储介质

## (57) 摘要

本发明实施例提出一种区域识别的方法、装置、终端和计算机可读存储介质,方法包括:根据路网数据将指定地区切分为多个区块;根据需要识别的区域类型,对指定地区的打点数据进行筛选;将筛选出的打点数据聚类;根据聚类结果从多个区块中获取与区域类型对应的目标区域;根据筛选出的打点数据,识别出目标区域的边界。本发明实施例由于能够根据实时路网数据和打点数据识别指定地区的区域类型和边界,因此能够保证当前时间段识别出的区域的类型和边界的时效性和准确度。



1. 一种区域识别的方法,其特征在于,包括:  
根据路网数据将指定地区切分为多个区块;  
根据需要识别的区域类型,对所述指定地区的打点数据进行筛选;  
将筛选出的所述打点数据聚类;  
根据聚类结果从多个所述区块中获取与所述区域类型对应的目标区域;  
根据筛选出的所述打点数据,识别出所述目标区域的边界;  
其中,根据路网数据将指定地区切分为多个区块,包括:  
从所述指定地区的路网数据中确定出多个切分道路;  
对各所述切分道路进行线段聚合,将所述指定地区切分成多个所述区块。
2. 如权利要求1所述的方法,其特征在于,将筛选出的所述打点数据聚类,包括:  
通过密度聚类算法对筛选出的所述打点数据进行聚类,并生成所述聚类结果。
3. 如权利要求1所述的方法,其特征在于,还包括:  
根据需要识别的所述区域类型,对所述指定地区的兴趣点进行筛选;  
对筛选出的所述兴趣点与所述目标区域进行关系判断和距离计算,得到与所述目标区域相关联的兴趣点。
4. 如权利要求3所述的方法,其特征在于,对筛选出的所述兴趣点与所述目标区域进行关系判断和距离计算,得到与所述目标区域相关联的兴趣点,包括:  
若一个兴趣点位于所述目标区域的边界内,则判定为所述相关联的兴趣点;  
若一个兴趣点位于所述目标区域的所述边界周围,则计算所述兴趣点与所述区域的所述边界之间的距离,若所述距离在阈值范围内,则判定为所述相关联的兴趣点。
5. 如权利要求3所述的方法,其特征在于,还包括:  
根据所述相关联的兴趣点,获取所述目标区域的区域信息。
6. 一种区域识别的装置,其特征在于,包括:  
切分模块,用于根据路网数据将指定地区切分为多个区块;  
第一筛选模块,用于根据需要识别的区域类型,对所述指定地区的打点数据进行筛选;  
聚类模块,用于将筛选出的所述打点数据聚类;  
区域类型识别模块,用于根据聚类结果从多个所述区块中获取所述区域类型对应的目标区域;  
区域边界识别模块,用于根据筛选出的所述打点数据,识别出所述目标区域的边界;  
所述切分模块包括:  
道路选择子模块,用于从所述指定地区的路网数据中确定出多个切分道路;  
切分子模块,用于对各所述切分道路进行线段聚合,将所述指定地区切分成多个所述区块。
7. 如权利要求6所述的装置,其特征在于,还包括:  
第二筛选模块,用于根据所述区域类型,对所述指定地区的兴趣点进行筛选;  
处理模块,用于对筛选出的所述兴趣点与所述目标区域进行关系判断和距离计算,得到与所述目标区域相关联的兴趣点。
8. 如权利要求7所述的装置,其特征在于,处理模块包括:  
判断子模块,用于判断若一个兴趣点位于所述目标区域的边界内,则判定为所述相关

联的兴趣点;若一个兴趣点位于所述目标区域的所述边界周围,则计算所述兴趣点与所述区域的所述边界之间的距离,若所述距离在阈值范围内,则判定为所述相关联的兴趣点。

9.一种区域识别的终端,其特征在于,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1至5中任一项所述的方法。

10.一种计算机可读存储介质,其存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1至5中任一项所述的方法。

## 区域识别的方法、装置、终端和计算机可读存储介质

### 技术领域

[0001] 本发明涉及地理信息技术领域,尤其涉及一种区域识别的方法、装置、终端和计算机可读存储介质。

### 背景技术

[0002] 传统的区域类型和边界识别方案包括两种,一种是通过在地图上进行格栅分解,找出人群聚集的格栅标注为某一区域类型。另一种是通过卫星图形进行图像识别。其中,格栅分解的方式不够准确,由于线下区域不规则,利用格栅表示的形状可能与实际差别较大。而卫星图形识别的方式则存在依赖图像清晰程度的问题,且通过图像识别出的区域准确度不高,不能包括全部区域或面积圈选过大。另一方面,线下的区域类型和边界会随着时间变化,因此现有方式获取的区域类型和边界不能满足数据的时效性。

[0003] 在背景技术中公开的上述信息仅用于加强对本发明的背景的理解,因此其可能包含没有形成本领域普通技术人员所知晓的现有技术的信息。

### 发明内容

[0004] 本发明实施例提供一种区域识别的方法、装置、终端和计算机可读存储介质,以解决现有技术中的一个或多个技术问题。

[0005] 第一方面,本发明实施例提供了一种区域识别的方法,包括:

[0006] 从查询语句中提取查询语义信息;

[0007] 在键值库中查找与所述查询语义信息相关的多个键值对;

[0008] 获取与所述查询语义信息相关的辅助信息;

[0009] 将所述辅助信息与查找到的各所述键值对进行加权,得到答复语义信息;

[0010] 将所述答复语义信息转换为答复语句。

[0011] 结合第一方面,本发明实施例在第一方面的第一种实现方式中,在键值库中查找与所述查询语义信息相关的多个所述键值对,包括:

[0012] 从所述查询语义信息中提取知识语义信息;

[0013] 从知识库中获取与所述知识语义信息相关的知识语料;

[0014] 将所述知识语义信息与所述知识语料组成所述键值对。

[0015] 结合第一方面的第一种实现方式,本发明实施例在第一方面的第二种实现方式中,将所述知识语义信息与所述知识语料组成所述键值对,包括:

[0016] 通过神经网络将所述知识语义信息与所述知识语料组成所述键值对。

[0017] 结合第一方面,本发明实施例在第一方面的第三种实现方式中,还包括:

[0018] 将所述答复语义信息与所述查询语义信息组成问答键值对;

[0019] 将所述问答键值对与所述键值库中存储的已有键值对进行匹配,若所述键值库的所述已有键值对中不存在所述问答键值对,则将所述问答键值对存入所述键值库中。

[0020] 结合实施例在第一方面的第三种实现方式,本发明实施例在第一方面的第四种实

现方式中,将所述问答键值对与所述键值库中存储的已有键值对进行匹配,若所述键值库的所述已有键值对中不存在所述问答键值对,则将所述问答键值对存入所述键值库中,具体步骤包括:

- [0021] 将所述问答键值对暂存到所述键值库的短期记忆区中;
- [0022] 通过评估模型对所述问答键值对进行评估,若评估所述问答键值对的阈值超出所述已有键值对的阈值,则将所述问答键值对存入到所述键值库的长期记忆区中。
- [0023] 结合第一方面,本发明实施例在第一方面的第五种实现方式中,还包括:
- [0024] 通过随机变量调整所述答复语句的语言表述方式。
- [0025] 第二方面,本发明实施例提供了一种区域识别的装置,包括:
- [0026] 提取模块,用于从查询语句中提取查询语义信息;
- [0027] 获取模块,用于在键值库中查找与所述查询语义信息相关的多个键值对;
- [0028] 辅助信息模块,用于获取与所述查询语义信息相关的辅助信息;
- [0029] 处理模块,用于将所述辅助信息与查找到的各所述键值对进行加权,得到答复语义信息;
- [0030] 转换模块,用于将所述答复语义信息转换为答复语句。
- [0031] 在一个可能的设计中,所述获取模块包括:
- [0032] 提取子模块,用于从所述查询语义信息中提取知识语义信息;
- [0033] 获取子模块,用于从知识库中获取与所述知识语义信息相关的知识语料;
- [0034] 键值对子模块,用于将所述知识语义信息与所述知识语料构建成所述键值对。
- [0035] 在一个可能的设计中,还包括:
- [0036] 配对模块,用于将所述答复语义信息与所述查询语义信息组成问答键值对;
- [0037] 扩充模块,用于将所述问答键值对与所述键值库中存储的已有键值对进行匹配,当所述键值库的所述已有键值对中不存在所述问答键值对,则将所述问答键值对存入所述键值库中。
- [0038] 在一个可能的设计中,所述扩充模块包括:
- [0039] 短记忆子模块,用于将所述问答键值对暂存到所述键值库的短期记忆区中;
- [0040] 长记忆子模块,用于通过评估模型对所述问答键值对进行评估,若评估所述问答键值对的阈值超出所述已有键值对的阈值,则将所述问答键值对存入到所述键值库的长期记忆区中。
- [0041] 第三方面,本发明实施例提供了一种区域识别的终端,包括:
- [0042] 终端的功能可以通过硬件实现,也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。
- [0043] 在一个可能的设计中,区域识别的终端的结构中包括处理器和存储器,所述存储器用于存储支持区域识别的终端执行上述第一方面中区域识别的方法的程序,所述处理器被配置为用于执行所述存储器中存储的程序。区域识别的终端还可以包括通信接口,用于区域识别的终端与其他设备或通信网络通信。
- [0044] 第四方面,本发明实施例提供了一种计算机可读存储介质,用于存储区域识别的终端所用的计算机软件指令,其包括用于执行上述第一方面中区域识别的方法为区域识别的终端所涉及的程序。

[0045] 上述技术方案中的一个技术方案具有如下优点或有益效果:由于能够根据实时路网数据和打点数据识别指定地区的区域类型和边界,因此能够保证当前时间段识别出的区域的类型和边界的时效性和准确度。

[0046] 上述概述仅仅是为了说明书的目的,并不意图以任何方式进行限制。除上述描述的示意性的方面、实施方式和特征之外,通过参考附图和以下的详细描述,本发明进一步的方面、实施方式和特征将会是容易明白的。

### 附图说明

[0047] 在附图中,除非另外规定,否则贯穿多个附图相同的附图标记表示相同或相似的部件或元素。这些附图不一定是按照比例绘制的。应该理解,这些附图仅描绘了根据本发明公开的一些实施方式,而不应将其视为是对本发明范围的限制。

[0048] 图1为本发明实施方式提供的区域识别的方法流程图。

[0049] 图2为本发明实施方式提供的获取目标区域的区域信息的方法流程图。

[0050] 图3为本发明实施方式提供的区域识别的装置的结构示意图。

[0051] 图4为本发明实施方式提供的区域识别的终端的结构示意图。

### 具体实施方式

[0052] 在下文中,仅简单地描述了某些示例性实施例。正如本领域技术人员可认识到的那样,在不脱离本发明的精神或范围的情况下,可通过各种不同方式修改所描述的实施例。因此,附图和描述被认为本质上是示例性的而非限制性的。

[0053] 本发明实施例提供了一种区域识别的方法,特别是区域识别类型和边界的方法,如图1所示,包括以下步骤:

[0054] S100:根据路网数据将指定地区切分为多个区块。指定地区可以包括但不限于城市、区县、小区或是面积较大的自然地理地块等。

[0055] S200:根据需要识别的区域类型,对指定地区的打点数据进行筛选。区域类型可以包括但不限于居民区(住宅区)、商业区、工业区、行政区等。打点数据为线下定位的GPS点,根据数据类型的不同,同一区域中可含有若干种类型的打点数据。

[0056] S300:将筛选出的打点数据聚类。

[0057] S400:根据聚类结果从多个区块中获取与区域类型对应的目标区域。

[0058] S500:根据筛选出的打点数据,识别出目标区域的边界。边界可以理解为是该区域的地理边界或是占地范围。

[0059] 在一个实施例中,根据路网数据将指定地区切分为多个区块,包括:从指定地区的路网数据中确定出多个切分道路。对各切分道路进行线段聚合,将指定地区切分成多个区块。进一步的,为了更准确的切分,还可以引入指定地区的相关区域数据,通过指定地区的相关区域数据与路网数据结合,将指定地区切分为多个区块。

[0060] 需要说明的是,路网数据可以包括但不限于道路的类型、等级、名称和限速等多种数据。例如,道路的类型包括高速公路、普通马路、林荫道路、公有道路和私有道路等。道路等级包括机动车辆主路、机动车辆辅路、人行道和自行车道等。道路限速包括起始段限速和终止段限速等。

[0061] 在一个具体的实施方式中,可以根据预先设定的规则切分道路。例如,设定规则为道路为公有道路,道路的起始限速不低于30km/h,根据该规则对指定地区内的道路进行筛选。若切分出的区块与原指定地区的公有边界出现交错,则以根据切分道路切分出的区块的边界为准。

[0062] 在一个实施例中,将筛选出的打点数据聚类,包括:通过密度聚类算法对筛选出的打点数据进行聚类,并生成聚类结果。

[0063] 在一个具体的实施方式中,若需要识别的区域为居民区,则可以从指定区域的打点数据中筛选出与居民区有关的人群打点数据。例如,将人群夜间常驻位置的打点数据和/或人群家常驻位置的打点数据作为聚类特征,并通过密度聚类算法进行聚类挖掘聚集区域,进而识别出属于居民区类型的区块。其中,聚类特征可以仅包含一种与居民区类型有关的打点数据,也可以包含多种与居民区类型有关的打点数据。

[0064] 类似的,若需要识别的区域类型为商业区、行政区或工业区时,则从指定区域的打点数据中筛选出与指定区域类型相关的打点数据即可。具体筛选的数据可根据需要进行选择,具体的识别方式参照上述实施例,在此不再一一举例。

[0065] 在一个实施例中,为了能够对筛选出的类型区块做进一步的分析,还可引入兴趣点。如图2所示,该方法还可以包括:

[0066] 步骤S600:根据需要识别的区域类型,对指定地区的兴趣点进行筛选。

[0067] 步骤S700:对筛选出的兴趣点与目标区域进行关系判断和距离计算,得到与目标区域相关联的兴趣点。

[0068] 步骤S800:根据相关联的兴趣点,获取目标区域的区域信息。

[0069] 兴趣点可以包括目标区域所在位置的任何信息及数据。例如,当需要识别的区域类型为居民区时,则筛选出的兴趣点可以包含所在位置的小区名称、街道名称、小区的面积、人口及居住人群类型等任何与居民区有关的区域信息。

[0070] 在一个具体的实施方式中,对筛选出的兴趣点与目标区域进行关系判断和距离计算,得到与目标区域相关联的兴趣点,包括:

[0071] 若一个兴趣点位于目标区域的边界内,则判定为相关联的兴趣点。

[0072] 若一个兴趣点位于目标区域的边界周围,则计算兴趣点与区域的边界之间的距离,若距离在阈值范围内,则判定为相关联的兴趣点。

[0073] 其中,若一个兴趣点位于多个目标区域中时,选取距离最近的目标区域作为相关联的兴趣点。

[0074] 例如,当筛选出多个类型为居民区的区块后,根据兴趣点可以进一步的判断每个居民区的具体小区名字、小区面积、居住人群的类型分布等区域信息。

[0075] 在一个实施例中,区域识别方法还包括:将识别出的区块、区块边界以及区域信息进行整合,形成完整的关联数据。

[0076] 本发明实施例提供了一种区域识别的装置,如图3所示,包括:

[0077] 切分模块10,用于根据路网数据将指定地区切分为多个区块。

[0078] 第一筛选模块20,用于根据需要识别的区域类型,对指定地区的打点数据进行筛选。

[0079] 聚类模块30,用于将筛选出的打点数据聚类。

- [0080] 区域类型识别模块40,用于根据聚类结果从多个区块中获取区域类型对应的目标区域。
- [0081] 区域边界识别模块50,用于根据筛选出的打点数据,识别出目标区域的边界。
- [0082] 在一个实施例中,切分模块10包括:
- [0083] 道路选择子模块,用于从指定地区的路网数据中确定出多个切分道路。
- [0084] 切分子模块,用于对各切分道路进行线段聚合,将指定地区切分成多个区块。
- [0085] 在一个实施例中,如图3所示,区域识别的装置还包括:
- [0086] 第二筛选模块60,用于根据区域类型,对指定地区的兴趣点进行筛选。
- [0087] 处理模块70,用于对筛选出的兴趣点与目标区域进行关系判断和距离计算,得到与目标区域相关联的兴趣点。
- [0088] 在一个实施例中,处理模块70包括:
- [0089] 判断子模块,用于判断若一个兴趣点位于目标区域的边界内,则判定为相关联的兴趣点。若一个兴趣点位于目标区域的边界周围,则计算兴趣点与区域的边界之间的距离,若距离在阈值范围内,则判定为相关联的兴趣点。
- [0090] 在一个实施例中,还包括数据产出模块80,用于将识别出的区块、区块边界以及区域信息进行整合,形成完整的关联数据。
- [0091] 本发明实施例提供了一种区域识别的终端,如图4所示,包括:
- [0092] 存储器910和处理器920,存储器910内存储有可在处理器920上运行的计算机程序。处理器920执行计算机程序时实现上述实施例中的目标区域类型及边界的方法。存储器910和处理器920的数量可以为一个或多个。
- [0093] 通信接口930,用于存储器910和处理器920与外部进行通信。
- [0094] 存储器910可能包含高速RAM存储器,也可能还包括非易失性存储器(non-volatile memory),例如至少一个磁盘存储器。
- [0095] 如果存储器910、处理器920、以及通信接口930独立实现,则存储器910、处理器920以及通信接口930可以通过总线相互连接并完成相互间的通信。总线可以是工业标准体系结构(ISA,Industry Standard Architecture)总线、外部设备互连(PCI,Peripheral Component)总线或扩展工业标准体系结构(EISA,Extended Industry Standard Component)总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示,图4中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。
- [0096] 可选的,在具体实现上,如果存储器910、处理器920以及通信接口930集成在一块芯片上,则存储器910、处理器920及通信接口930可以通过内部接口完成相互间的通信。
- [0097] 本发明实施例提供了一种计算机可读存储介质,其存储有计算机程序,该程序被处理器执行时实现如实施例一包括的任一所述的区域识别的方法。
- [0098] 本发明实施例由于能够根据实时路网数据和打点数据识别指定地区的区域类型和边界,因此能够保证当前时间段识别出的区域的类型和边界的时效性和准确度。在此基础上,通过与兴趣点进行关联,能够在识别出的类型区域基础上,对该区域进行更多的商业分析,应用范围广泛。例如,应用到大数据营销、选址、销量预估等各种不同需求中。
- [0099] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特



点包含于本发明的至少一个实施例或示例中。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0100] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或隐含地包括至少一个该特征。在本发明的描述中,“多个”的含义是两个或两个以上,除非另有明确具体的限定。

[0101] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本发明的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本发明的实施例所属技术领域的技术人员所理解。

[0102] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为是用于实现逻辑功能的可执行指令的定序列列表,可以具体实现在任何计算机可读介质中,以供指令执行装置、装置或设备(如基于计算机的装置、包括处理器的装置或其他可以从指令执行装置、装置或设备取指令并执行指令的装置)使用,或结合这些指令执行装置、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行装置、装置或设备或结合这些指令执行装置、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPR0M或闪速存储器),光纤装置,以及便携式只读存储器(CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0103] 应当理解,本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行装置执行的软件或固件来实现。例如,如果用硬件来实现,和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0104] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0105] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读存储介质中。所述存储介质可以是只读存储器,磁盘或光盘等。

[0106] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到其各种变化或替换,这些都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

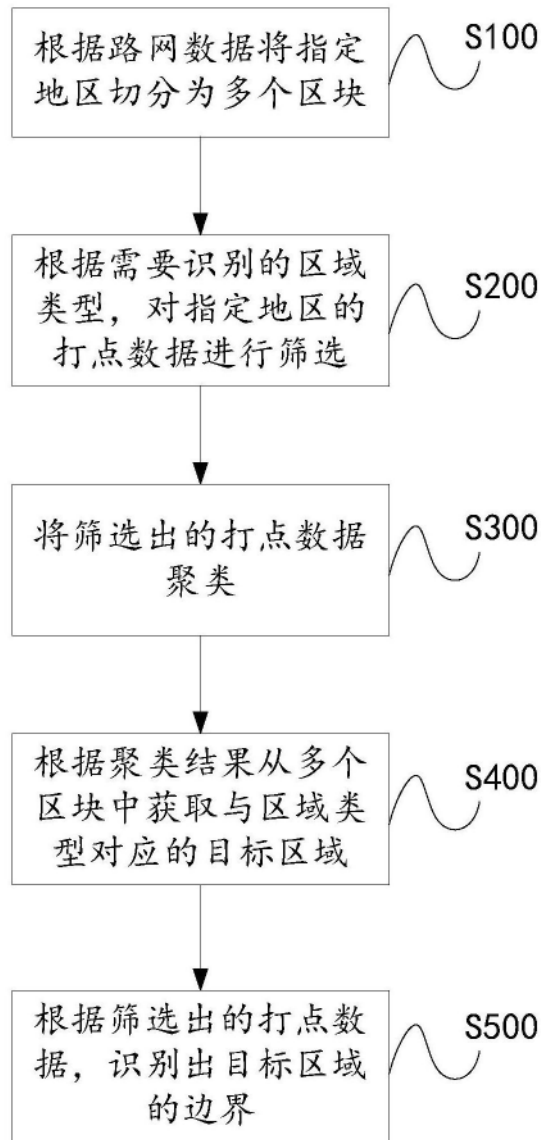


图1

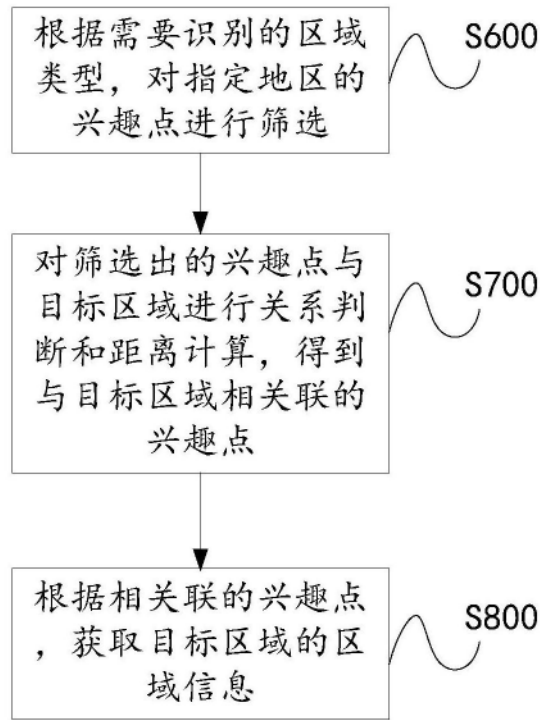


图2

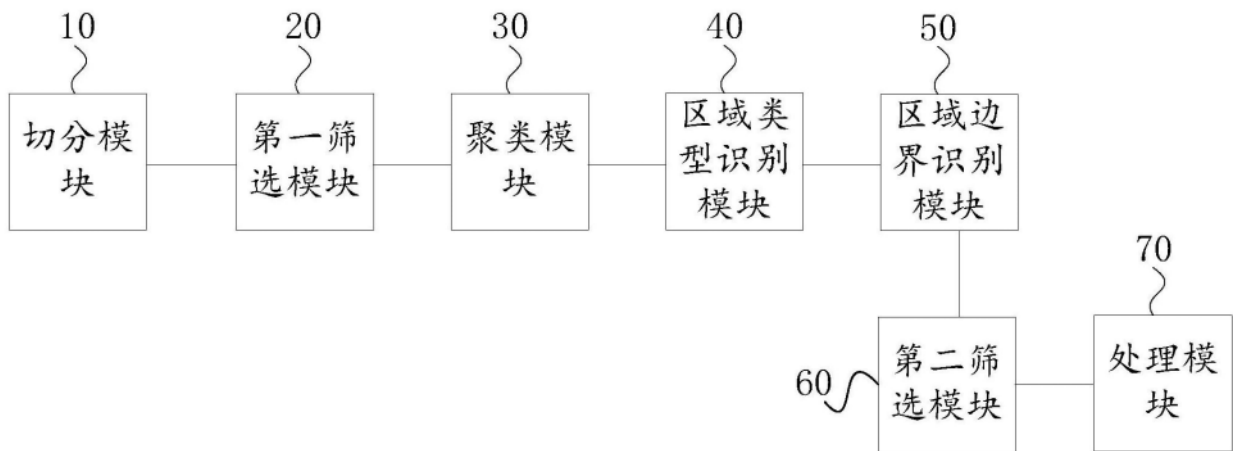


图3

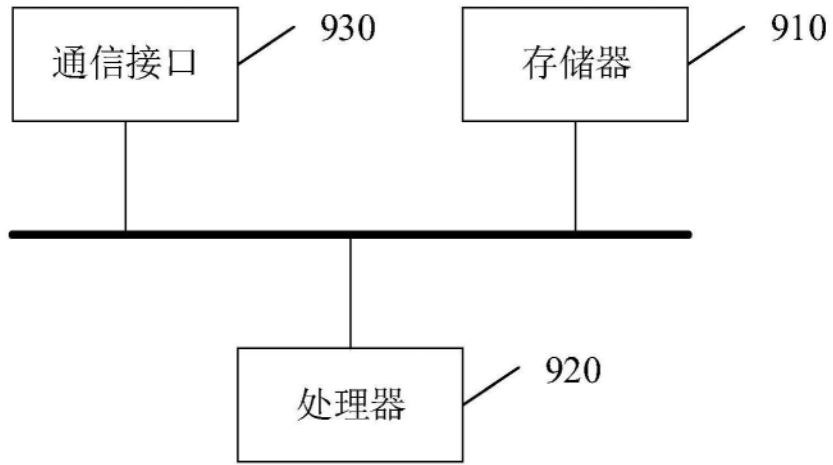


图4