



(12)发明专利申请

(10)申请公布号 CN 108732313 A  
(43)申请公布日 2018.11.02

(21)申请号 201810548447.0

(22)申请日 2018.05.31

(71)申请人 深圳市创艺工业技术有限公司  
地址 518000 广东省深圳市南山区南山街  
道深南大道10128号南山数字文化产  
业基地南山软件园西塔楼1708

(72)发明人 李健斌

(51) Int. Cl.  
G01N 33/00(2006.01)  
G01S 19/14(2010.01)

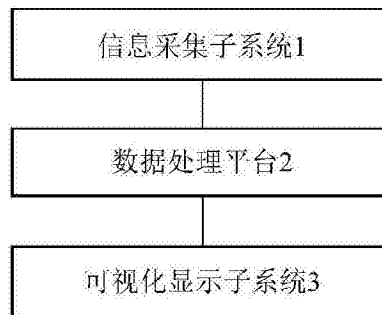
权利要求书2页 说明书4页 附图1页

(54)发明名称

城市空气污染物浓度智能观测系统

(57)摘要

本发明提供了城市空气污染物浓度智能观测系统,包括设于城市公交上的信息采集子系统、数据处理平台和可视化显示子系统;所述信息采集子系统包括北斗卫星导航系统、污染物浓度采集模块和数据传输模块;信息采集子系统通过北斗卫星导航系统获取观测位置信息,通过污染物浓度采集模块获取并预处理观测位置的污染物浓度数据,经数据传输模块传输至数据处理平台;其中污染物浓度采集模块包括用于获取观测位置的污染物浓度数据的传感监测装置;所述数据处理平台对污染物浓度数据进行数据异常检测和异常类型分析,并对检测出的异常数据按照异常类型进行相应处理,并将处理后的污染物浓度数据及观测位置信息上传至可视化显示子系统。



1. 城市空气污染物浓度智能观测系统,其特征是,包括设于城市公交上的信息采集子系统、数据处理平台和可视化显示子系统;

所述信息采集子系统包括北斗卫星导航系统、污染物浓度采集模块和数据传输模块;信息采集子系统通过北斗卫星导航系统获取观测位置信息,通过污染物浓度采集模块获取并预处理观测位置的污染物浓度数据,经数据传输模块传输至数据处理平台;其中污染物浓度采集模块包括用于获取观测位置的污染物浓度数据的传感监测装置;

所述数据处理平台对污染物浓度数据进行数据异常检测和异常类型分析,并对检测出的异常数据按照异常类型进行相应处理,并将处理后的污染物浓度数据及观测位置信息上传至可视化显示子系统。

2. 根据权利要求1所述的城市空气污染物浓度智能观测系统,其特征是,传感监测装置包括:一氧化碳浓度传感器、二氧化硫浓度传感器、一氧化氮浓度传感器、颗粒物浓度传感器。

3. 根据权利要求1所述的城市空气污染物浓度智能观测系统,其特征是,所述可视化显示子系统包括区域实时动态分布模块,所述区域实时动态分布模块用于显示城市空气污染物浓度的实时分布图以及固定地点污染物浓度随时间变化的趋势图。

4. 根据权利要求3所述的城市空气污染物浓度智能观测系统,其特征是,所述可视化显示子系统还包括预报分析模块以及数据结果发布模块;所述预报分析模块用于输出区域空气质量的评估、未来发展趋势的预报以及出行方面的建议;所述数据结果发布模块用于在客户端上发布观测数据结果。

5. 根据权利要求1-4任一项所述的城市空气污染物浓度智能观测系统,其特征是,数据处理平台对污染物浓度数据进行数据异常检测和异常类型分析,并对检测出的异常数据按照异常类型进行相应处理,具体为:

(1) 采用滑动窗口对传感器a的污染物浓度数据进行提取,设滑动窗口尺寸为b,滑动窗口当前提取的传感器a的污染物浓度数据集合为 $\{y_{t-b+1}, y_{t-b+2}, \dots, y_t\}$ ,其中 $y_t$ 为传感器a在第t时刻采集的污染物浓度数据,则滑动窗口在下次提取的传感器a的污染物浓度数据集合为 $\{y_{t+1}, y_{t+2}, \dots, y_{t+b}\}$ ;

(2) 对滑动窗口当前提取的污染物浓度数据集合中的污染物浓度数据进行异常检测,设检测出的异常数据的数目为U,根据U值判定异常数据的类型,并根据异常数据的类型对异常数据进行相应处理:

当 $U=1$ 时,判定异常数据为点异常数据,利用与该点异常数据相邻的两个污染物浓度数据的平均值作为修复值,采用修复值对点异常数据进行替换;

当 $1 < U \leq K$ 时,判定异常数据为候选异常数据;对候选异常数据进行连续性分析,若检测为连续性异常数据序列时,将所有候选异常数据标定为真正异常数据;否则标定为点异常数据;

当 $U > K$ 时,判定异常数据为待定异常数据;当与传感器a同类型的邻居传感器采集的污染物浓度数据中,在同一滑动窗口提取的污染物浓度数据也被检测出待定异常数据时,将传感器a及其同类型的邻居传感器的待定异常数据标定为正常数据;否则标定为真正异常数据;

对所有真正异常数据进行剔除,其中,K为设定的异常数据个数上限;

(3) 移动滑动窗口,对新提取的污染物浓度数据集合中的污染物浓度数据按照(2)进行异常检测。

6. 根据权利要求5所述的城市空气污染物浓度智能观测系统,其特征是,所述对滑动窗口当前提取的污染物浓度数据集合中的污染物浓度数据进行异常检测,具体包括:

(1) 计算滑动窗口当前提取的污染物浓度数据集合的污染物浓度数据均值 $\bar{y}$ 、中值 $y_{med}$ 和标准差 $g$ ,若所述当前提取的污染物浓度数据集合中的污染物浓度数据 $y_i$ 不满足下列公式,则将该污染物浓度数据 $y_i$ 标识为可能异常数据:

$$\frac{\bar{y} + y_{med}}{2} - 3g \leq y_i \leq \frac{\bar{y} + y_{med}}{2} + 3g$$

(2) 计算滑动窗口当前提取的污染物浓度数据集合中任意相邻两个污染物浓度数据之间的邻差,该邻差即当前污染物浓度数据和前一个污染物浓度数据之间的差值,对计算出的各邻差进行检测,将大于设定阈值的邻差判定为异常邻差;

(3) 保存第一次检测到的异常邻差 $\Delta y_0$ ,若下一次检测到的异常邻差 $|y_j - y_{j-1}|$ 与所述异常邻差 $\Delta y_0$ 的乘积为负数,且 $y_j$ 已被标识为可能异常数据,则将 $y_j$ 判定为异常数据。

## 城市空气污染物浓度智能观测系统

### 技术领域

[0001] 本发明涉及空气污染监测技术领域,具体涉及城市空气污染物浓度智能观测系统。

### 背景技术

[0002] 有针对性的进行空气污染防控与治理,需要对城市内空气污染时空分布进行监控。而从当前城市空气污染物浓度监控现状来看,整体处于测站密度过低、难以实时测量的状况。

[0003] 现有技术中,空气污染物浓度观测主要以地面气象站观测和应急移动观测为主。对于地面气象站观测来说,建设成本高且受环境限制,分布密度过低,无法适应城市内空气污染物浓度分布实时监控的需求;而对于移动观测来说,为获得尽可能同时次的多的数据,需要大量的移动观测车辆,这样观测的成本也将非常高,在业务研究污染物浓度实时分布中也失去了实际意义。

### 发明内容

[0004] 针对上述问题,本发明提供城市空气污染物浓度智能观测系统。

[0005] 本发明的目的采用以下技术方案来实现:

[0006] 提供了城市空气污染物浓度智能观测系统,包括设于城市公交上的信息采集子系统、数据处理平台和可视化显示子系统;

[0007] 所述信息采集子系统包括北斗卫星导航系统、污染物浓度采集模块和数据传输模块;信息采集子系统通过北斗卫星导航系统获取观测位置信息,通过污染物浓度采集模块获取并预处理观测位置的污染物浓度数据,经数据传输模块传输至数据处理平台;其中污染物浓度采集模块包括用于获取观测位置的污染物浓度数据的传感监测装置;

[0008] 所述数据处理平台对污染物浓度数据进行数据异常检测和异常类型分析,并对检测出的异常数据按照异常类型进行相应处理,并将处理后的污染物浓度数据及观测位置信息上传至可视化显示子系统。

[0009] 优选地,传感监测装置包括:一氧化碳浓度传感器、二氧化硫浓度传感器、一氧化氮浓度传感器、颗粒物浓度传感器。

[0010] 优选地,所述可视化显示子系统包括:区域实时动态分布模块、预报分析模块以及数据结果发布模块;

[0011] 所述区域实时动态分布模块用于显示城市空气污染物浓度的实时分布图以及固定地点污染物浓度随时间变化的趋势图;

[0012] 所述预报分析模块用于输出区域空气质量的评估、未来发展趋势的预报以及出行方面的建议;

[0013] 所述数据结果发布模块用于在客户端上发布观测数据结果。

[0014] 本发明的有益效果为:基于北斗卫星导航系统的实时定位功能,以城市公交系统

为平台,建立城市空气污染物浓度观测网,实现空气污染物浓度移动观测,观测结果更加及时、准确,为城市空气污染防控与治理提供有效支撑数据。

### 附图说明

[0015] 利用附图对本发明作进一步说明,但附图中的实施例不构成对本发明的任何限制,对于本领域的普通技术人员,在不付出创造性劳动的前提下,还可以根据以下附图获得其它的附图。

[0016] 图1是本发明一个示例性实施例的城市空气污染物浓度智能观测系统的结构示意图;

[0017] 图2是本发明一个示例性实施例的信息采集子系统的结构示意图。

[0018] 附图标记:

[0019] 信息采集子系统1、数据处理平台2、可视化显示子系统3、北斗卫星导航系统10、污染物浓度采集模块20、数据传输模块30。

### 具体实施方式

[0020] 结合以下实施例对本发明作进一步描述。

[0021] 参见图1、图2,本发明实施例提供了城市空气污染物浓度智能观测系统,包括设于城市公交上的信息采集子系统1、数据处理平台2和可视化显示子系统3;

[0022] 所述信息采集子系统1包括北斗卫星导航系统10、污染物浓度采集模块20和数据传输模块30;信息采集子系统1通过北斗卫星导航系统10获取观测位置信息,通过污染物浓度采集模块20获取并预处理观测位置的污染物浓度数据,经数据传输模块30传输至数据处理平台2;其中污染物浓度采集模块20包括用于获取观测位置的污染物浓度数据的传感监测装置;

[0023] 所述数据处理平台2对污染物浓度数据进行数据异常检测和异常类型分析,并对检测出的异常数据按照异常类型进行相应处理,并将处理后的污染物浓度数据及观测位置信息上传至可视化显示子系统3。

[0024] 传感监测装置包括:一氧化碳浓度传感器、二氧化硫浓度传感器、一氧化氮浓度传感器、颗粒物浓度传感器。

[0025] 其中,所述可视化显示子系统包括:区域实时动态分布模块、预报分析模块以及数据结果发布模块;

[0026] 所述区域实时动态分布模块用于显示城市空气污染物浓度的实时分布图以及固定地点污染物浓度随时间变化的趋势图;

[0027] 所述预报分析模块用于输出区域空气质量的评估、未来发展趋势的预报以及出行方面的建议;

[0028] 所述数据结果发布模块用于在客户端上发布观测数据结果。

[0029] 本发明上述实施例基于北斗卫星导航系统10的实时定位功能,以城市公交系统为平台,建立城市空气污染物浓度观测网,实现空气污染物浓度移动观测,观测结果更加及时、准确,为城市空气污染防控与治理提供有效支撑数据。

[0030] 在一个实施中,数据处理平台2对污染物浓度数据进行数据异常检测和异常类型

分析,并对检测出的异常数据按照异常类型进行相应处理,具体为:

[0031] (1) 采用滑动窗口对传感器a的污染物浓度数据进行提取,设滑动窗口尺寸为b,滑动窗口当前提取的传感器a的污染物浓度数据集合为 $\{y_{t-b+1}, y_{t-b+2}, \dots, y_t\}$ ,其中 $y_t$ 为传感器a在第t时刻采集的污染物浓度数据,则滑动窗口在下次提取的传感器a的污染物浓度数据集合为 $\{y_{t+1}, y_{t+2}, \dots, y_{t+b}\}$ ;

[0032] (2) 对滑动窗口当前提取的污染物浓度数据集合中的污染物浓度数据进行异常检测,设检测出的异常数据的数目为U,根据U值判定异常数据的类型,并根据异常数据的类型对异常数据进行相应处理:

[0033] 当 $U=1$ 时,判定异常数据为点异常数据,利用与该点异常数据相邻的两个污染物浓度数据的平均值作为修复值,采用修复值对点异常数据进行替换;

[0034] 当 $1 < U \leq K$ 时,判定异常数据为候选异常数据;对候选异常数据进行连续性分析,若检测为连续性异常数据序列时,将所有候选异常数据标定为真正异常数据;否则标定为点异常数据;

[0035] 当 $U > K$ 时,判定异常数据为待定异常数据;当与传感器a同类型的邻居传感器采集的污染物浓度数据中,在同一滑动窗口提取的污染物浓度数据也被检测出待定异常数据时,将传感器a及其同类型的邻居传感器的待定异常数据标定为正常数据;否则标定为真正异常数据;

[0036] 对所有真正异常数据进行剔除,其中,K为设定的异常数据个数上限;

[0037] (3) 移动滑动窗口,对新提取的污染物浓度数据集合中的污染物浓度数据按照(2)进行异常检测。

[0038] 数据异常的来源通常包括传感器错误、事件发生或网络攻击,由于污染物浓度数据异常的来源不同,相应对异常数据的处理方式也应该不同。本实施例设定了对异常数据进行分类的机制,该机制根据检测出的异常数据的数目来判定异常数据的类型,点异常数据通常来源于传感器错误,当异常数据的数目仅为1个时,可以直接将该异常数据作为点异常数据;当异常数据的数目为少数时,可能是源于传感器的多个连续的测量错误或者是多个点异常数据,此时需要进一步对异常数据发生的原因进行分析,基于此,本实施例设计对候选异常数据进行再次异常检测;当异常的数据大于设定数目阈值时,本实施例根据传感器的时空关联性对该异常数据进行分析。采用本实施例设定的分类机制对传感器的异常数据进行分类,并根据异常数据的类型对异常数据进行相应处理,能够降低异常检测处理的误报率,提高对污染浓度数据进行异常检测的精度和效率。

[0039] 在一个实施例中,所述对滑动窗口当前提取的污染物浓度数据集合中的污染物浓度数据进行异常检测,具体包括:

[0040] (1) 计算滑动窗口当前提取的污染物浓度数据集合的污染物浓度数据均值 $\bar{y}$ 、中值 $y_{med}$ 和标准差g,若所述当前提取的污染物浓度数据集合中的污染物浓度数据 $y_i$ 不满足下列公式,则将该污染物浓度数据 $y_i$ 标识为可能异常数据:

$$[0041] \quad \frac{\bar{y} + y_{med}}{2} - 3g \leq y_i \leq \frac{\bar{y} + y_{med}}{2} + 3g$$

[0042] (2) 计算滑动窗口当前提取的污染物浓度数据集合中任意相邻两个污染物浓度数据之间的邻差,该邻差即当前污染物浓度数据和前一个污染物浓度数据之间的差值,对计

算出的各邻差进行检测,将大于设定阈值的邻差判定为异常邻差;

[0043] (3) 保存第一次检测到的异常邻差  $\Delta y_0$ ,若下一次检测到的异常邻差  $|y_j - y_{j-1}|$  与前述异常邻差  $\Delta y_0$  的乘积为负数,且  $y_j$  已被标识为可能异常数据,则将  $y_j$  判定为异常数据。

[0044] 相关技术中,常采用基于聚类的异常检测或者基于预测模型的异常检测方法对传感器采集的数据进行异常检测,但是这两种方法均对数据中的噪声点和异常点比较敏感,从而影响异常检测算法的检测结果。

[0045] 当污染物浓度数据发生异常时,相应污染物浓度数据对应的邻差值均有明显的改变。本实施例基于相邻两个污染物浓度数据的邻差,创新性地设计了对污染物浓度数据进行异常检测的机制,该机制相对于基于聚类的异常检测和基于预测模型的异常检测方法,能够避免噪声和异常点的影响,检测过程较为简单,且对污染物浓度数据进行异常检测的效果好。

[0046] 在一个实施例中,所述对候选异常数据进行连续性分析,具体包括:

[0047] (1) 提取第一个候选异常数据到最后一个候选异常数据之间的所有污染物浓度数据,作为进行连续性分析的数据样本;

[0048] (2) 设所述数据样本中的候选异常数据个数为  $\rho$ ,若数据样本符合下列连续性条件,判定该候选异常数据为连续性候选异常数据:

$$[0049] \left[ \left( \max_{b=1, \dots, r-1} H_{b, b+1} \right) - H_T \right] \times \left[ 0.6 \left( \max_{b=1, \dots, r-1} Z_{b, b+1} + \min_{b=1, \dots, r-1} Z_{b, b+1} \right) - Z_T \right] < 0$$

[0050] 式中,  $H_{b, b+1}$  为数据样本中第  $b$  个候选异常数据与第  $b+1$  个候选异常数据之间的污染物浓度数据个数,  $r$  为数据样本中的候选异常数据个数,  $H_T$  为设定的正常数据个数阈值;  $Z_{b, b+1}$  为所述第  $b$  个候选异常数据与第  $b+1$  个候选异常数据之间的差值绝对值,  $Z_T$  为设定的差值绝对值上限。

[0051] 本实施例对候选异常数据进行连续性分析时,根据候选异常数据及其之间的污染物浓度数据作为数据样本进行分析,并构建连续性条件,当数据样本符合连续性条件时,可以设定为异常源于传感器的多个连续的测量错误,否则将候选异常数据作为点异常数据,按照点异常数据的处理方式对候选异常数据进行处理,分析简单便捷,提高了异常检测处理的效率。

[0052] 最后应当说明的是,以上实施例仅用以说明本发明的技术方案,而非对本发明保护范围的限制,尽管参照较佳实施例对本发明作了详细地说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或者等同替换,而不脱离本发明技术方案的实质和范围。

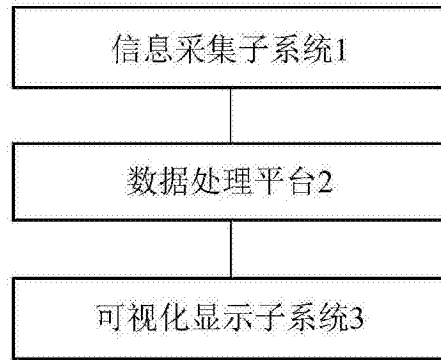


图1

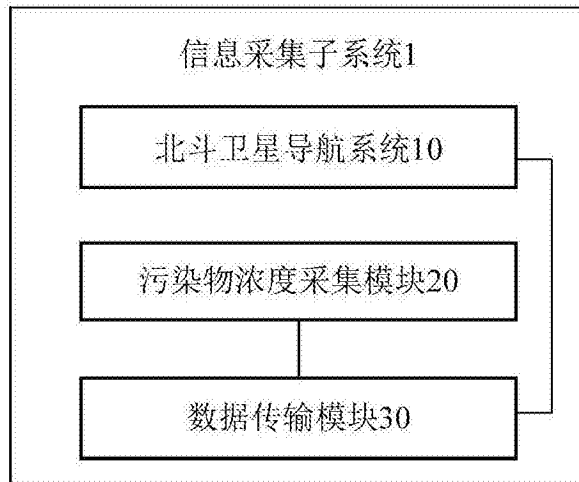


图2