

基于 UML 技术的空间基础数据库设计与实现

李 龙, 刘 澎, 张永东, 张可佳

(东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要: 设计合理的空间基础数据库不仅能提高整个空间数据操作时的性能和效率, 而且还可以减少后期的维护修复工作, 使整个空间基础数据运行的更加快捷, 需要对空间基础数据建立数据库; 当前的空间基础数据库设计方法采用 ArcSDE 数据引擎对空间基础数据进行不断地更新调整, 再利用多源空间数据格式转换的技术对空间基础数据库进行设计, 存在空间基础数据运行时速度缓慢, 计算精度低的问题; 为此, 提出了一种基于 UML 技术的空间基础数据库设计方法; 该方法首先在空间基础数据库设计中建立空间基础数据索引结构, 利用 R-树族构建空间基础数据索引树, 依据空间基础数据索引树, 扫描索引空间基础数据, 过滤掉不满足查询条件的空间基础数据对象, 使空间数据查询结果可以在额定时间内获得, 然后采用空间基础数据点、数据线、以及由数据线组成的区域、一组区域、空间基础数据网络的详细存储方式对空间基础数据进行存储, 利用 MongoDB 驱动程序对矢量空间数据进行存储, 最后通过对空间基础数据库索引、查询、存储等设计实现了空间基础数据库的建立; 仿真实验结果证明, 所提方法提高了空间基础数据的建库速度, 减少了数据运行的时间, 提升了空间基础数据的利用率。

关键词: UML 技术; 空间基础数据库; 设计与实现

Based on UML Technology of Space Database Design and Implementation

Li long, Liu Peng, Zhang Yongdong, Zhang Kejia

(Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: Design and reasonable basis of the database can not only improve performance and efficiency of the whole spatial data operation, but also can reduce the late maintenance repair work, make whole space based data run faster, need to create a database of space based data. The current spatial database design method using ArcSDE data engine to adjust constantly updated space based data reuse technology of multi-source spatial data format conversion based on space database design, there is space based data run-time speed is slow, the problem of low precision. To this end, this paper proposes a space based database design method based on UML technology. This method first based on space database design based spatial data index structure, using R-based data index tree tree building space, on the basis of space based data index tree, scanning index space based data, filter based data object does not satisfy the query conditions of space, make the space data query result can be gained in the rated time and then USES the data points in space, data line, and area of cable, a set of detailed area, space based data network based data storage model of space for storage, use mongo drivers to store of vector spatial data, finally through to the space foundation database index, query, storage design has realized the space such as the establishment of the underlying database. The simulation experimental results show that the proposed method improves the speed of space of the construction of basic data, reduced the data of the running time, improve the utilization of space based data.

Keywords: UML technology; space the underlying database; design and implementation

0 引言

近年来, 随着计算机网络的不断发展, 空间基础数据库的应用受到了很多领域的高度重视, 这些领域通过空间基础数据库管理系统的功能来辅助其工作与任务的完成, 使空间基础数据库的设计与实现成为了当今的热门话题^[1-2]。由于空间基础数据的运行具有复杂性, 无序性, 不安全性等特点, 需要对其设计数据库方便其运行和处理。多数空间基础数据库的设计方法无法对空间基础数据库进行安全、灵活、全面的设计, 导致空间基础数据运行或处理时出现操作延迟, 数据丢失等情况^[3-4]。在这种情况下, 如何提高空间基础数据库设计与实现的精度成为了急需解决的问题。而基于 UML 技术的空间基础数据库设计方法可以对空间基础数据的索引查询和存储进行设计。是解决上述问题的有效途径^[5-6], 成为了空间基础数据库

设计研究者的主要研究方向, 受到了该领域学者的广泛关注, 同时也取得了很多优秀的成果^[7]。

文献 [8] 提出了一种基于 SQLServer 的空间基础数据库设计方法。该方法首先利用面向对象技术法以及 C/S 体系通过对图形数据的分解, 进行空间基础数据和属性数据统一存储, 然后采用分布式数据库管理系统中的 GeoView 系统完善关系数据库管理系统, 从而完成对空间基础数据库的设计。该方法虽然较为简单, 但是实现过程中存在精度低的问题。文献 [9] 提出了一种基于 SDB 的空间基础数据库设计方法。该方法首先采用 SDB 数据引擎对空间基础数据库进行构建与管理, 为此建立专有的数据转换模块, 然后完善空间基础数据库属性建设, 修改空间数据库属性结构, 编辑空间数据对象的属性值, 为使用者提供灵活便捷的空间数据基础库功能, 对空间基础数据库进行最大优化程度的构建与管理。该方法所设计的空间基础数据库较为完善, 且利用率较高, 但是存在消耗时间过长的的问题。文献 [10] 提出了一种基于 Oracle Spatial 的空间基础数据库设计方法。该方法先将 Oracle Spatial 中的空间基础数据对象类型设为结构体, 并使其可供 OCI 操作, 然后根

收稿日期: 2017-03-31; 修回日期: 2017-04-13。

作者简介: 李 龙 (1966-), 男, 山东莱阳人, 博士, 副教授, 硕士生导师, 主要从事数据挖掘、智能系统方向的研究。

据 OCI 功能进行空间基础数据中异构数据的转换, 将格式不同的空间基础数据在 Oracle 中实现集成, 最后设计并实现一个基于 Oracle Spatial 的空间基础数据库。该方法对空间基础数据库的设计与实现速度快, 效果好, 但是进行空间基础数据中异构数据转换过程比较繁琐。

针对上述产生的问题, 提出了一种基于 UML 技术的空间基础数据库设计方法。该方法对空间基础数据库的索引、查询和存储进行了设计, 仿真实验证明, 所提方法可以迅速全面地对空间基础数据库进行设计, 实验中的各项有利数据也证明了基于 UML 技术的空间基础数据库设计方法是可实现的。

1 基于 UML 技术的空间基础数据库设计方法

1.1 空间基础数据库索引查询设计

在空间基础数据库设计中建立空间基础数据索引结构, 可利用扫描索引空间基础数据, 过滤掉部分不满足查询条件的空间基础数据对象, 从而减少空间系统负载。一般采用的空间基础数据索引为 R-树族, 依据空间基础数据索引树, 使查询结果可以在额定时间内获得。在 R-树族算法中, 空间基础数据区域查询是最有代表性的查询方法, 评估 R-树族的性能一般以空间基础数据区域查询的性能来表示。空间基础数据区域查询的响应时间主要由读取与查询区域节点所需时间和 CPU 处理区域节点所需时间组成。因为访问一个空间基础数据节点页面的时间开销很大, 所以基本可以忽略, 以空间基础数据区域查询所需的访问节点数来表示空间基础数据库索引查询性能。

假设 N_d 为空间基础数据集中主引导记录的数目, N 为 R-树族中空间基础数据节点数目。空间基础数据集中所有 MBR 归一化在单位正方形 U 中。假如空间基础数据查询区域为 Q , 大小为 $q_x \times q_y$, 空间基础数据查询区域均匀的分布在单位正方形 U 中。空间基础数据点查询表示空间基础数据区域查询在 $q_x = q_y = 0$ 时的一个特例。如果 R-树族中空间基础数据节点 n_{ij} 的主引导记录为 $x_{ij} \times y_{ij}$, 假设在一次空间基础数据点查询中该节点被访问的概率为 $NB(x_{ij}, y_{ij})$, 则:

$$U = [0, 2] \times [0, 2] \quad (1)$$

$$NB(x_{ij}, y_{ij}) = x_{ij} \times y_{ij} \quad (2)$$

其中: $NB()$ 代表在一个任意空间基础数据点查询中 R-树族中空间基础数据节点访问次数, x_{ij} 和 y_{ij} 分别代表空间基础数据节点的横纵坐标中第 i 个值和第 j 个值。

对空间基础数据给定区域 $q_x \times q_y$ 的查询, 可能引起的空间基础数据节点访问次数为 $p(q_x, q_y)$:

$$p(q_x, q_y) = \sum_{i=0}^H \sum_{j=1}^{N_i} x_{ij} \times y_{ij} + q_y \times \sum_{i=0}^H \sum_{j=1}^{N_i} x_{ij} + q_x \times \sum_{i=0}^H \sum_{j=1}^{N_i} y_{ij} + N \times q_x \times q_y \quad (3)$$

其中:

$$p(q_x, q_y) = \text{空间基础数据节点面积} + q_y * L_x + q_x * L_y + N * q_x * q_y \quad (4)$$

其中: L_x 、 L_y 分别代表 R-树族中空间基础数据节点 X 轴和 Y 轴的边长, H 代表空间基础数据查询的响应时间。

由上式得知, R-树族中空间基础数据节点面积与空间基础数据节点周长都对 R-树族的空间基础数据库区域查询性能有影响, R-树族中空间基础数据节点数目越少, 区域查询效果就越好。所以空间基础数据库查询设计主要考虑以下内容:

(1) 确定空间基础数据库结构, 利用空间基础数据以及空间属性数据不同的管理方式, 采用混合模式实现空间基础数据库的结构设计, 混合模式是将空间基础数据存储于空间文件系统中。

(2) 空间基础数据分析, 空间基础数据分析包括确定空间基础数据范围中水平方向的空间基础数据范围以及垂直方向的空间基础数据范围以及空间基础数据分辨率等。基于 UML 技术的空间基础数据库设计方法依据的空间基础数据模型是基于对象的空间数据模型。

(3) 空间基础数据访问方法设计, 空间基础数据访问方法的组成部分是空间基础数据索引和在空间基础数据索引上的操作。

(4) 空间基础数据查询操作设计, 空间基础数据查询操作包括空间基础数据点查询和区域查询。

综上所述, 完成对空间基础数据库索引查询的设计。

1.2 空间基础数据库存储设计

在空间基础数据库中需要将空间基础数据对象进行存储, 以下是对空间基础数据对象存储的方式: (1) 将空间基础数据对象的特征信息存储到空间基础数据库中, 空间基础数据对象原来的几何图形保存到数据大对象中; (2) 将空间基础数据对象的原始几何图形进行离散化, 然后再进行存储。以下是空间基础数据对象的具体存储方式:

对于空间数据点的存储: 假设一个 n 维的空间基础数据点, 用 n 元数组 (x_1, x_2, \dots, x_n) 表示。

对于空间基础数据线 L 的存储: 用一组空间基础数据点表示为:

$$L = \{ (a_1^{(1)}, a_2^{(1)}, \dots, a_n^{(1)}), (a_1^{(2)}, a_2^{(2)}, \dots, a_n^{(2)}), \dots, (a_1^{(m)}, a_2^{(m)}, \dots, a_n^{(m)}) \} \quad (5)$$

其中: a 代表空间基础数据点, m 代表空间基础数据点顺序值。

对于空间基础数据区域 R 的存储: 把空间基础数据区域边界看成一条曲线, 将区域看作一条闭合的曲线, 用存储空间基础数据的方式将闭合的曲线存储, 空间基础数据区域的第一个点和最后一个点相同的情况下, 空间基础数据区域 R :

$$R = \left\{ (a_1^{(1)}, a_2^{(1)}, \dots, a_n^{(1)}), (a_1^{(2)}, a_2^{(2)}, \dots, a_n^{(2)}), \dots, (a_1^{(m)}, a_2^{(m)}, \dots, a_n^{(m)}), (a_1^{(1)}, a_2^{(1)}, \dots, a_n^{(1)}) \right\} \quad (6)$$

对于空间基础数据一组区域 P 的存储: 用空间基础数据存储区域的方式存储一组区域:

$$P = \{R_1, R_2, \dots, R_m\} \quad (7)$$

对空间基础数据网络的存储, 空间基础数据网络是由若干数据点以及若干数据线组成的, 则存储方式是点和线存储方式的结合。其存储方式表示为:

$$net = \{nodes, lines\} \quad (8)$$

其中:

$$node = \left\{ (b_1^{(1)}, b_2^{(1)}, \dots, b_n^{(1)}), (b_1^{(2)}, b_2^{(2)}, \dots, b_n^{(2)}), \dots, (b_1^{(m)}, b_2^{(m)}, \dots, b_n^{(m)}) \right\} \quad (9)$$

$$lines = \{L_1, L_2, \dots, L_k\} \quad (10)$$

其中: L_i 代表空间基础数据网络中的线, b 代表空间基础数据点的数量值。

以上是空间数据库点、线、区域、一组区域、网络空间基础数据的存储设计。而针对矢量空间基础数据类型和结构特点

的不同，创建不同的空间基础数据集满足空间基础数据的存储：(1) 空间基础数据集，用来存储空间基础数据；(2) 空间元数据项集合，用来存储空间基础数据中的元数据项。空间基础数据分为共享性数据和私人数据，针对空间共享性数据，利用过滤器的安全模式将空间基础数据存储于统一的空间基础数据库中，针对空间私人数据，每个私人用户利用权限控制的安全模式设计单独的空间存储数据库。以下是对矢量空间基础数据的具体存储方式。

针对矢量空间基础数据的存储采用的是 Mongo DB 驱动程序，Mongo DB 驱动程序利用文档形式对空间基础数据进行存储，每个文档由空间基础数据对应值组成，可以存储复杂的空间基础数据。空间基础数据存储到数据库时，将每个空间基础数据对象转换成一个 json 数据对象，采用 key/value 方式存储空间基础数据对象的空间属性，实现了对矢量空间基础数据的存储，由此完成对空间基础数据库存储的设计。

2 仿真实验结果与分析

为了证明基于 UML 技术的空间基础数据库设计方法的有效性，需要进行一次仿真实验。实验仿真平台建立在由 10 台物理机构建的空间基础数据集环境下，实验数据取自于 Shapefile 文件中的若干空间基础数据，由此验证基于 UML 技术的空间基础数据库设计方法的整体有效性和可实现性。表 1 是基于 UML 技术的空间基础数据库设计方法中空间基础数据库查询数据量（万个）与查询响应时间（s）的关系。

表 1 空间基础数据库查询数据量与查询响应时间关系

查询数据量/万个	查询响应时间/s
1000	0.2
2000	0.4
3000	0.6
4000	0.8
5000	0.9

由表 1 可以明显看出，基于 UML 技术的空间基础数据库设计方法中空间数据库查询系统的响应时间随着空间数据量的变化而变化，在变化过程中随着空间基础数据量的增加查询响应时间也随之增加，但是幅度较小，查询数据量在 1000 万个—4000 万个时，空间基础数据库查询响应时间以 0.2 s 的速度逐渐增长，查询数据量在 5000 万个以上时，查询响应时间以 0.1 s 的速度增长，说明了两点内容：(1) 查询数据量的增加没有给查询系统的响应时间带来过分影响；(2) 查询系统响应时间随着查询数据量的增加反而越来越少，这证明了基于 UML 技术的空间基础数据库设计方法具有良好的实用性，也证明了该方法是可实现的。表 2 是空间基础数据库存储速度（万个/s）与数据量（万个）的关系。

表 2 空间基础数据库存储速度与数据量的关系

数据存储速度万个/s	数据量/万个
300	3000
400	4000
500	5000
600	6000
700	7000

表 2 中的数据显示了不同数量的大规模空间基础数据存储时的速度，以每秒数万个的速度进行存储，有效提高了空间基础数据的运行效率。主要是因为利用基于 UML 技术的空间基础数据库设计方法进行空间数据存储时细化了存储模式，采用点、线、区域、组区域、网络等细化模式对空间基础数据进行存储，增强了存储的速度，提高了空间基础数据的存储精度。进一步证明了基于 UML 技术的空间基础数据库设计方法的整体性能和可行性。图 1 是本文方法对空间基础数据存储精度（%）的描述。

图 1 是对基于 UML 技术的空间基础数据库设计方法中空间基础数据量（万个）与存储精度（%）关系的描述。当空间基础数据刚开始进行存储时存储精度在 87% 左右，随着数据量的不断增加存储精度反而处于上升的趋势，且平稳处于 90% 以上，证明了基于 UML 技术的空间基础数据库设计方法存储时的安全稳定性。为空间基础数据库的实现提供了可靠依据。图 2 是文献 [8] 与本文空间基础数据库查询效率（%）的对比。

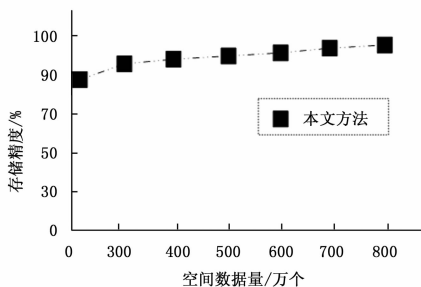


图 1 空间基础数据库存储精度

图 2 中显示的是文献 [8] 所提方法与本文方法空间基础数据库查询效率（%）的对比。在该对比中明显本文方法空间基础数据库查询效率（%）较高，图中本文方法随着查询数据点的个数变化处于平稳的波动中，且波动是持续上升的，证明了基于 UML 技术的空间基础数据库设计方法在查询数据点（个）时的查询效率（%）很高，有利于基于 UML 技术的空间基础数据库设计方法的实现。

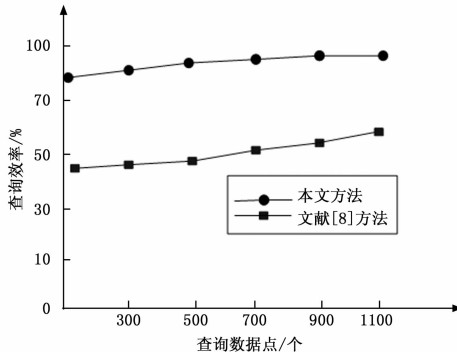


图 2 空间基础数据库查询效率对比

仿真实验证明，基于 UML 技术的空间基础数据库设计方法对空间基础数据库的索引，查询以及存储进行了详细的设计，可以高效，安全，准确地对空间基础数据库进行设计与实现，是切实可行的数据库设计与实现方法。

3 结束语

采用当前方法对空间基础数据库进行设计与实现时, 无法安全精确地对空间基础数据库进行设计与实现, 存在数据库设计有误差, 空间基础数据运行过程复杂的问题。提出了一种基于 UML 技术的空间基础数据库设计方法。并通过仿真实验证明, 所提方法可以安全可靠地对空间基础数据库进行设计与实现, 是该领域研究发展的重点参考对象, 具有良好的应用价值。

参考文献:

- [1] 李运兴, 马 骏. 基于 Hadoop 高性能查询的空间数据库设计与实现 [J]. 中原工学院学报, 2014, 25 (4): 58-63.
- [2] 王 琰, 邓永刚. 基于 MapGIS 的区域水文地质图空间数据库设计与实现 [J]. 测绘技术装备, 2016, 18 (1): 62-63.
- [3] 马海燕, 李文欣. 城市基础地理空间数据库设计与实现研究 [J]. 科技资讯, 2015, 13 (1): 24-25.
- [4] 陈达伦, 陈荣国, 谢 炯. 基于 MPP 架构的并行空间数据库原型

(上接第 164 页)

1) 人为因素。本实验的测试设备主要为示波器。整个操作过程, 尤其是测量幅值的过程中, 波形放大后观察判断数据难以做到准确判断。这是造成测试结果存在误差的一个主要原因。

2) 硬件因素。本系统基带信号的产生和调制都是由软件设计完成的, 调制好的数字信号在经过数/模转换之后, 会有低通滤波器对模拟信号低通滤波, 低通滤波器的相位特性决定了输出信号的相位特性, 这一过程可能会对方位信号的相位造成一定的延迟。

3) 程序原因。本系统脉冲的产生与编码都是软件完成的。GNU Radio 软件中脉冲的生成是对矩形脉冲整形滤波得到的。主、辅基准脉冲群的调制在 vector source 中是以采样点的形式进行的, 主、辅基准脉冲群在周期上不是绝对的 9 倍关系, 随着发射时间的延长, 误差会逐渐累积, 这也是造成误差的原因之一。但是根据对示波器接收波形的较长时间观察, 基准脉冲群与方位包络信号相位关系基本相对不变, 所以这个因素影响应该不是很大。

综上所述, 对信号模块的测试, 验证了信号参数及编码符合塔康导航信号国军标要求。虽然方位测试有略大的误差, 但基本在预期之内, 实验过程及结果基本验证系统的功能。

6 结束语

论文采用软件无线电思想设计实现了塔康系统地面台的信号模拟系统, 通过实验验证了信号模拟系统设计的正确性。既满足了飞机对塔康信号的测试要求, 也符合软件无线电技术思想, 为塔康系统的测试提供了一种新的平台, 为塔康系统设备验收及在生命周期中正常工作提供了重要的保障手段。也为解决飞机航电设备在综合测试和总装在线测试中无线电信号的模拟问题奠定了一定基础; 同时, 这种基于软件无线电技术构建的信号模拟系统, 稍加改造, 也能用于未来新研制或增加机载设备。针对新研制或增加机载设备的开发更多地转到相应的软件的开发和设计上来, 可大大地降低成本。

参考文献:

- [1] 徐新丽, 万会兵. 先进通用飞机航电系统概述 [J]. 航空电子技术, 2015 (2): 32-35.

系统的设计与实现 [J]. 地球信息科学学报, 2016, 18 (2): 151-159.

- [5] 刘一明, 胡卓玮, 赵文吉, 等. 基于 Geodatabase 模型的扶贫开发空间数据库的设计与实现 [J]. 工程勘察, 2014, 42 (7): 44-49.
- [6] 史华林, 何 飞, 王云全, 等. 基于 ArcSDE 和 SQL Sever 城市规划空间数据库的设计与实现 [J]. 江西科学, 2014, 32 (3): 357-360.
- [7] 魏振华, 刘志锋, 李金萍, 等. 基于要素扩展管理的海量地质空间数据存储模型的设计与实现 [J]. 计算机应用与软件, 2014, 31 (7): 36-39.
- [8] 杜久升, 王 莉, 侯 争, 等. 露天矿卡车调度系统数据库设计与实现 [J]. 金属矿山, 2015, 44 (9): 105-111.
- [9] 吴 琰, 唐小明. 基于 HBase 的分布式空间数据库技术 [J]. 吉林大学学报: 理学版, 2016, 14 (6): 1355-1360.
- [10] 陈 辉, 赵晓华. 学生上机实训考核系统中数据库设计与实现 [J]. 电子设计工程, 2016, 24 (18): 4-6.
- [2] 李 冰. 导航信号模拟源的研究与实现 [D]. 西安: 西安电子科技大学, 2013.
- [3] 刘 祁. 某型飞机航电系统通用测试设备的研制 [J]. 机电工程技术, 2016 (s1): 1-6.
- [4] 李 骁. 航电导航设备一体化模拟器研究与设计 [D]. 北京: 北京理工大学, 2015.
- [5] 杨小牛, 楼才义, 徐建良. 软件无线电技术与应用 [M]. 北京: 北京理工大学出版社, 2014.
- [6] Unik Univ, Grad. Software Defined Radio: Challenges and Opportunities [J]. IEEE Communication Surveys Tutorials, 2010, 5 (10): 531-550.
- [7] 李 骏. 基于软件无线电信号模拟技术的研究 [J]. 硅谷, 2014 (18): 40-41.
- [8] Nambissan T J, Nikhil T V, Vinodkumar V. A VHF Radio for Software Defined Radio Applications [J]. Procedia Technology, 2016, 24: 820-826.
- [9] 梁 晶, 张 允, 彭德强. 塔康信号模拟器的设计 [J]. 舰船电子对抗, 2014 (1): 109-112, 120.
- [10] 周振国. 塔康系统关键技术与塔康测位的实现 [D]. 西安: 西安电子科技大学, 2012.
- [11] 吴 苗, 朱 涛, 李方能, 等. 无线电导航原理及应用 [M]. 北京: 国防工业出版社, 2008.
- [12] 焦守波, 李 乐. DME 系统在飞机导航中的应用 [J]. 中国科技信息, 2016 (16): 40-41.
- [13] Li S, Ni Y, Cai N. Optimal strategy of DME beacon distribution for DME/DME area navigation [J]. 2012, 3: 2036-2039.
- [14] GJB 914-90, 塔康系统信号要求和测试方法 [S]. 中华人民共和国国家军用标准 1991.
- [15] 陈通海, 李景春, 郑 娜, 等. 基于 GNURadio 和 USRP2 的未知信号检测 [J]. 无线电工程, 2012, 42 (12): 16-19.
- [16] 张 鹏, 孙福余, 徐亚明, 等. 基于 USRP 平台的 GPS 软件接收机设计与实现 [J]. 测绘通报, 2012 (11): 55-58.
- [17] Gandhiraj R, Soman K P. Modern analog and digital communication systems development using GNU Radio with USRP [J]. Telecommunication Systems, 2014, 56 (3): 367-381.
- [18] 曹健辉. 基于软件无线电的多通道宽带数字信号处理系统的实现 [D]. 成都: 电子科技大学, 2014.
- [19] 胡广书. 数字信号处理——理论、算法与实现 [D]. 北京: 清华大学出版社, 2003.