

分类号_____

密级_____

UDC _____

编号_____

中国科学院研究生院

硕士学位论文

基于 FISS 的岩矿多维高光谱数据库研究

覃环虎

指导教师 张立福 副研究员 中国科学院遥感应用研究所

童庆禧 院 士 中国科学院遥感应用研究所

申请学位级别 理学硕士 学科专业名称 地图学与地理信息系统

论文提交日期 2011 年 12 月 论文答辩日期 2012 年 2 月

培 养 单 位 中国科学院遥感应用研究所

学位授予单位 中国科学院研究生院

答辩委员会主席_____

M.S. Dissertation

**Research on Multidimensional Hyperspectral Database for Rock and
Mineral Based on FISS**

By

QIN Huanhu

Directed by: Assoc.Prof. ZHANG Lifu

Prof.TONG Qingxi

A dissertation submitted in fulfillment
of the requirements for the degree of
Master of Science
in
Cartography and Geographical Information System
February 2012

本论文的研究得到国家自然科学基金项目“高光谱遥感岩矿多维数据库关键技术研究”（41072248）的资助。

中国科学院遥感应用研究所

学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。论文中凡引用他人已经发表或未发表的成果、数据、观点等，均已明确注明出处。除文中已经注明引用的内容外，本论文不含任何其它个人或集体已经发表或撰写过的成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。

本声明的法律责任由本人承担。

论文作者签名： 日期： 年 月 日

学位论文使用授权说明

本人在导师指导下完成的论文，知识产权归属中国科学院遥感应用研究所。

本人完全了解中国科学院遥感应用研究所关于收集、保存、使用学位论文的规定，同意研究所保存或向国家有关部门及机构送交论文的印刷版和电子版，允许论文被查阅和借阅；同意授权研究所采用影印、缩印、数字化或其它复制手段保存和汇编本学位论文。

本人保证毕业后，以本论文数据和资料发表论文（著作）或使用论文工作成果时，第一署名单位为中国科学院遥感应用研究所，并征得导师同意。

保密论文在解密后遵守此规定。

论文作者签名：

导师签名：

日期： 年 月 日

摘 要

随着各种高光谱传感器的成功研制, 高光谱遥感技术正得到日益广泛的应用, 其所积累的数据量也与日俱增。如何有效的管理和共享海量高光谱数据, 并从海量数据中获取有益的信息, 是遥感科学工作者面临的一个重要挑战。建设高光谱数据库是解决数据有效管理和共享的方法, 而多维数据库技术和多维分析技术的发展, 将为从海量数据中获取有益信息提供技术基础。

本文以自主研发的地面成像光谱辐射测量系统 (Field Imaging Spectrometer System, FISS) 所采集到的高光谱数据为研究对象, 以建立岩矿多维高光谱数据库为最终目标, 开展了两个方面的研究工作: 一是分析、设计与开发了 FISS 高光谱数据库原型系统, 二是基于多维数据库的理论, 以 FISS 高光谱数据库存储的数据为数据源, 探讨了岩矿多维高光谱数据库的基本概念、数据库设计以及多维分析工具集设计。

论文的主要成果和结论如下:

1、提出了以 GeoRaster 对象为核心的 FISS 数据存储模型, 在该存储模型的基础上, 对基于存储过程的数据入库方法和基于动态数据表转换的光谱数据获取方法进行了深入研究。研究表明, 该存储模型和数据入库方法能够有效的支持高光谱数据图谱合一的存储。

2、在 Arc Engine 开发工具和 C#的支持下, 研究了 FISS 合成影像的显示方法和光谱曲线的显示方法。实验结果表明, 数据获取方法、影像显示方法和光谱曲线显示方法的完美结合, 有效的实现了高光谱数据图谱合一的显示。

3、在成功研究数据存取技术、影像和光谱显示方法的基础上, 开发实现了 FISS 高光谱数据库原型系统, 在系统中集成了一些具有高光谱特色的光谱分析方法。该系统基本具备了对高光谱数据进行管理、显示和分析的功能。

4、以多维数据库理论为基础, 研究了 FISS 岩矿多维高光谱数据库的基本概念, 初步设计了它的系统组成结构, 提出了嵌套的岩矿多维数据结构, 构建了岩矿多维数据的雪花模型, 研究了其多维分析工具集的组成, 认真设计了每一个多维分析操作的具体功能, 并展示了多维分析的效果。这些基础研究为下一步工作的开展奠定了坚实的基础。

关键词: 高光谱数据库 数据存储模型 多维数据库 FISS

Research on Multidimensional Hyperspectral Database for Rock and Mineral Based on FISS

QIN Huanhu (Cartography and Geographical Information System)

Directed by Assoc.Prof.ZHANF Lifu, Prof.TONG Qingxi

Abstract

Although hyperspectral remote sensing technology is advancing our understanding in various fields, it also brings us data redundancy. It is really a great challenge to manage and share hyperspectral remote sensing data effectively and obtain useful information from such massive data. We believe the establishment of a database for effective management and sharing of hyperspectral remote sensing data is really necessary. Meanwhile, the development of multidimensional database technology and multidimensional analysis technology will lay the foundation for obtaining useful information from massive data.

This study aims at the establishment of a multidimensional hyperspectral database for rock and mineral based on the hyperspectral data collected by the Field Imaging Spectrometer System (FISS). This study includes two parts: 1) design and develop the prototype system of FISS Hyperspectral Database (FHD); 2) probe into the basic concepts and design of the multidimensional hyperspectral database for rock and mineral (MDHDR&M) as well as the design of multidimensional analysis tools based on multidimensional database theory.

Main results and conclusions are summarized as follows:

1. FISS data storage model which key is GeoRaster is proposed. Extensive researches were carried out on data import method based on procedure and data extract method transformed through dynamic data table. The experimental results indicate that this model and data import method work well at the integrated storage of both image and spectrum of hyperspectral data.

2. Researches about the visualization of the FISS image and the visualization of FISS spectral curve are based on Arc Engine and C#. The result shows that flawless combination of data obtained method, image visualization method and spectral visualization method can support the integrated visualization of spectrum and image of hyperspectral data effectively.

3. A prototype system of FHD with several spectral analysis methods is developed. This prototype system is composed of such functions as management,

visualization and analysis.

4. This study redefines the basic concept of MDHDR&M and preliminarily designs its system structure. This model also brings in a multi-dimensional nested data structure to establish a snowflake schema. Besides, the compositions and detailed functions of multidimensional analysis tools are analyzed. Some results of multidimensional analysis are showed. All these lay the foundation for the further implementation of hyperspectral multi-dimensional database for rock and mineral.

Keywords: Hyperspectral database, Data storage model, Multidimensional database, FISS

目 录

第一章 引 言	1
1.1 研究背景及意义	1
1.2 研究内容与论文结构	2
第二章 FISS 高光谱数据库理论基础	4
2.1 地面成像高光谱辐射测量系统	4
2.1.1 地面成像高光谱辐射测量系统简介	4
2.1.2 地面成像高光谱辐射测量系统测量规范	5
2.1.3 地面成像高光谱辐射测量系统测数据预处理	6
2.2 高光谱数据库	8
2.2.1 高光谱数据库的概念及其重要意义	8
2.2.2 高光谱影像的几种管理方式	9
2.2.3 高光谱数据库的研究现状	10
2.3 多维数据库理论	11
2.3.1 多维数据库基本概念	11
2.3.2 多维数据库分析方法	12
2.3.3 多维数据库的存储方式	13
2.4 本章小结	14
第三章 FISS 高光谱数据库的分析与设计	15
3.1 功能需求分析	15
3.1.1 系统功能需求	15
3.1.2 系统业务流程分析	16
3.2 总体结构设计	17
3.2.1 体系结构设计	17
3.2.2 功能模块设计	19
3.3 系统开发环境	21
3.3.1 数据库管理系统	21
3.3.2 开发语言	22
3.3.3 ADO.NET 数据库接口	23
3.3.4 图形化数据库开发工具 SQL Developer	23
3.3.5 Arc Engine	24
3.4 FISS 数据存储模型设计	24
3.4.1 FISS 数据特点	24
3.4.2 FISS 数据存储模型	25
3.5 数据存储结构设计	28
3.6 本章小结	32
第四章 FISS 高光谱数据库关键技术研究及其实现	33
4.1 FISS 高光谱数据库数据存取技术的研究与实现	33
4.1.1 高光谱影像数据的存储设计	33
4.1.2 高光谱影像光谱数据的读取设计	35
4.1.3 基于存储过程的高光谱影像数据入库的实现	37
4.1.4 基于动态数据表转换的光谱数据读取的实现	39

4.2 影像及光谱显示方法研究与实现.....	40
4.2.1 影像显示方法研究与实现.....	40
4.2.2 光谱显示方法研究与实现.....	42
4.3 FISS 高光谱数据库的实现.....	44
4.4 基于 FISS 高光谱数据库的光谱分析功能.....	47
4.4.1 光谱平滑.....	47
4.4.2 光谱重采样.....	48
4.4.3 光谱匹配.....	49
4.4.4 包络线去除.....	52
4.5 本章小结.....	54
第五章 岩矿多维高光谱数据库初步研究.....	55
5.1 岩矿多维高光谱数据库基本概念.....	55
5.2.1 维.....	55
5.2.2 岩矿影像数据立方体.....	56
5.2.3 岩矿多立方体数据.....	57
5.2.4 维层次.....	57
5.2.5 维成员.....	57
5.2.6 度量.....	57
5.2 岩矿多维高光谱数据库设计.....	58
5.2.1 岩矿多维高光谱数据库功能设计.....	58
5.2.2 岩矿多维高光谱数据库系统结构设计.....	58
5.2.3 岩矿多维高光谱数据库概念模型设计.....	59
5.2.4 岩矿多维高光谱数据库逻辑模型设计.....	61
5.2.5 岩矿多维高光谱数据库物理模型设计.....	63
5.3 岩矿多维高光谱数据库多维分析工具集的设计.....	64
5.5.1 Oracle 多维数据库技术介绍.....	64
5.5.2 岩矿多维高光谱数据库多维分析工具集的组成.....	65
5.5.3 岩矿多维高光谱数据库多维分析工具的功能设计.....	65
5.4 本章小结.....	69
第六章 结论与展望.....	70
6.1 结论.....	70
6.2 展望.....	71
参考文献.....	72
发表文章目录.....	75
致 谢.....	77

第一章 引言

1.1 研究背景及意义

高光谱遥感技术是上世纪 80 年代出现的新型对地观测综合技术,始于成像光谱技术的发展。所谓高光谱是指光谱分辨率很高(达到 $10^{-2}\lambda$ 量级),在 400~2500nm 的波长范围内其光谱分辨率一般小于 10nm。高光谱遥感器,顾名思义,是指能够获取高光谱数据的遥感器。由于高光谱遥感器光谱分辨率高,往往在一定波长范围内(比如可见光—近红外、可见光—短波红外)相邻波段有光谱重叠区,也就是连续光谱成像,所以高光谱遥感器又称为成像光谱仪,即以获取大量窄波段连续光谱图像数据为目的的光谱采集设备^[1]。高光谱遥感技术的发展得益于卫星技术、传感器技术以及计算机技术的高速发展,随着国内外一系列空间计划的实施,一系列高光谱传感器得以成功研制。在星载成像光谱仪方面,有国外的星载中分辨率成像光谱仪 MODIS、高分辨率成像光谱仪 HIRIS、高光谱成像仪 Hyperion、紧密型高分辨率成像光谱仪 CHRIS 以及中国嫦娥一号搭载的干涉成像光谱仪、环境与减灾小卫星上搭载 HSI 高光谱成像仪等;在机载成像光谱仪方面,有国外的机载航空成像光谱仪 AIS 系列、航空可见光/红外成像光谱仪 AVIRIS、高光谱制图仪 HyMap 以及中国的模块化航空成像光谱仪 MAIS、推扫型成像光谱仪 PHI、航空成像光谱仪 OMIS 系列等;在地面成像光谱仪方面,国外很多公司和研究机构专门从事这种设备的研制,如芬兰的 Spectral Imaging Ltd. 公司,美国的 Resonon 和 Surface Optics Corporation 公司,取得了一系列的成果,我国自主研制的基于制冷面阵 CCD 地面成像光谱辐射测量系统 FISS(Field Imaging Spectrometer System, FISS)是国内第一台地面成像光谱仪^[1-6]。此外,国内外也有一系列的高光谱传感器的研究计划。这些已经投入使用的高光谱传感器以及即将使用的高光谱传感器,每天所获得的数据量将会以 TB 级甚至是 PB 级的增长。与多光谱影像不同,高光谱影像有自己的特点,主要表现在:(1)波段多——可以为每个像元提供几十、数百甚至上千个波段;(2)光谱范围窄——波段范围一般小于 10nm;(3)波段连续——有些传感器可以在 350~2500nm 的太阳光谱范围内提供几乎连续的地物光谱;(4)数据量大——随着波段数的增加,数据量成指数增加;(5)信息冗余增加——由于相邻波段高度相关,冗余信息也相对增加^[7]。如何根据高光谱数据的这些特点,对高光谱数据进行科学有效的管理和共享,是遥感工作者面临的严峻挑战。

数据库技术、存储技术等高技术的发展为海量高光谱数据的存储和管理提供了解决方案。

数据库技术的发展为海量高光谱数据的管理提供了有效的方法。数据库技术经历了第一代数据库(20 世纪 70 年代)的层次型、网状数据库系统发展到第二

代（20 世纪 80 年代）的关系型数据库，在很多领域都取得了巨大的成功，随着应用领域的不断扩展，关系数据库的限制和不足日益显现出来。随着面向对象和多维分析技术的出现，面向对象数据库系统和多维数据库成为数据库系统领域研究和发展的新方向。而随着面向对象、多维分析等方法和技术正逐步融入数据库，网络技术与数据库技术的融合使得数据库技术为高光谱数据的管理提供有益的技术支撑^[8, 9]。此外，在数据库恢复、并发控制、事务管理、完整性和安全性的实施、查询处理、优化技术等方面的研究也取得了长足的进展。多维数据库是数据库发展的最新阶段，其独特的多维数据模型，把数据组织成更利于分析的多维数据立方体，使得多维数据库需要更少的存储空间且具有更高效的检索速度。多维数据库的多维分析把一个实体的多项重要属性定义为多个维度，使得用户能方便的汇总数据集，简化了数据分析处理逻辑，并能对不同维度值的数据进行比较，从而发现数据的分布规律。

存储技术的发展为高光谱海量数据的存储提供了解决方案。目前的遥感影像管理系统对 TB 级的数据存储有了相对完善的解决方案，前述分析表明，众多成像光谱仪的使用将使高光谱数据的数据量很容易就达到 PB 级，因此研究 PB 级数据的存储管理系统已迫在眉睫。存储技术在近十几年内一直保持快速发展态势，在存储网络技术、存储系统架构、数据保护技术、新型存储介质、以及绿色存储等方向都取得了一系列重大进展。同时在寻求新的存储系统结构和管理方法、制定新的数据存储标准、不断推出数据存储系统、存储区域网络 SAN、网络附属存储 NAS、集群存储、对象存储系统等方面都进行了广泛的研究；国内企业如浪潮、华赛、圣桥等，也研制出了一系列存储产品，并取得了良好的市场业绩^[10-13]。

数据库技术和存储技术的发展使我们已经知道如何在外部存储设备上存储数据、如何使用各种复杂的存取方法、缓冲策略和索引技术访问外部存储设备上的数据。同时分布处理技术、并行计算技术、人工智能技术、多媒体技术的广泛应用也促进了数据库技术和存储技术的发展，它们使得管理海量高光谱数据并从中挖掘出有益信息在技术上成为了可能。

1.2 研究内容与论文结构

本研究的最终目标是建立岩矿多维高光谱数据库，围绕着这个目标，分两个阶段开展研究：第一个阶段是构建高效管理 FISS 高光谱数据的数据库并使之具有基本光谱分析功能，第二个阶段是以 FISS 高光谱数据库的数据为数据源，建立多维高光谱数据库。在本文中，主要进行了以下两个方面的工作：一是研究并建立 FISS 高光谱数据库原型系统，即第一个阶段的研究。在本阶段的研究中，继承前人提出的高光谱数据库的概念，对 FISS 高光谱数据库进行分析与设计，研究 FISS 数据存取的关键技术，设计实现 FISS 影像及光谱的显示方法，开发 FISS 高光谱数据库原型系统，在原型系统的基础上，集成几个具有高光谱特色

的光谱分析方法。二是结合多维数据库的理论，以岩矿多维高光谱数据库为研究对象，对多维高光谱数据库进行了初步研究，将从基本概念、数据库设计和多维分析工具集设计三个方面展开研究，为构建多维高光谱数据库提供了理论基础，是第二个阶段研究的序曲。

论文首先在引言中介绍了研究背景和意义，其次从 FISS、高光谱数据库、多维数据库三个方面介绍了本文的基础理论知识，然后进行了 FISS 高光谱数据库原型系统的设计与开发，接下来对岩矿多维高光谱数据库进行了初步研究。论文的整体结构如下图所示：

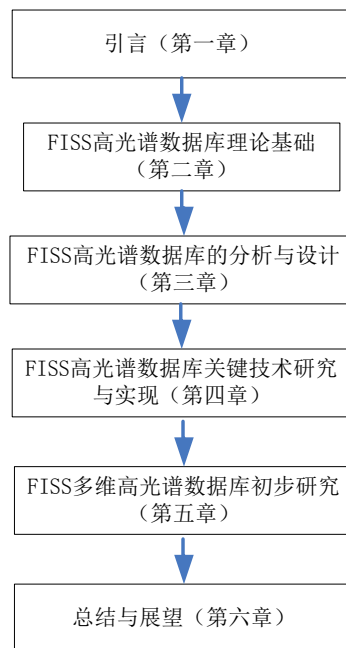


图 1-1 论文整体结构图

- 第一章， 主要介绍本文的研究背景和意义，据此引出本文的研究内容和结构。
- 第二章， 介绍 FISS 高光谱数据库的理论基础，包括：FISS 系统技术指标、测量规范及注意事项以及数据预处理方法；高光谱数据库的概念、高光谱影像的几种管理方式以及高光谱数据库的研究现状；多维数据库的基本概念、分析方法以及数据存储方式。
- 第三章， 分析 FISS 高光谱数据库的功能需求以及业务流程，设计其体系结构与功能模块，确定其开发环境，设计 FISS 数据存储模型，建立系统需要的数据表。
- 第四章， 研究高光谱数据的存取技术、影像和光谱显示方法，开发实现 FISS 高光谱数据库原型系统，研究简单的的光谱分析方法，并将它们集成到原型系统中。
- 第五章， 以岩矿多维高光谱数据库为研究对象，定义和描述岩矿多维数据库的基本概念，分析和设计多维高光谱数据库及其多维分析工具集。
- 第六章， 对全文进行总结，并对未来需要开展的工作进行了展望。

第二章 FISS 高光谱数据库理论基础

高光谱技术获取的高光谱影像，具有海量数据、图谱合一的特点，对数据的存储和管理提出了新的挑战。为迎接这一挑战，高光谱数据库应运而生，并取得了一定的研究成果。多维数据库是数据库发展的最新阶段。它把数据组织成更有利于分析的多维数据立方体，是面向决策支持的海量数据管理的关键因素^[14]。我国第一套地面成像光谱仪—地面成像光谱辐射测量系统（Field Imaging Spectrometer System, FISS）所获取的高光谱数据，是本文的数据源。而本文所研究的高光谱数据库也是专门存储和管理 FISS 数据的高光谱数据库。本章主要介绍三个方面的内容：一、FISS 的组成、技术指标、测量流程以及数据预处理方法；二、高光谱数据库的概念以及研究现状；三、多维数据库的基本理论、分析方法以及数据存储方式。

2.1 地面成像高光谱辐射测量系统

2.1.1 地面成像高光谱辐射测量系统简介

地面成像光谱辐射测量系统（Field Imaging Spectrometer System, FISS）是我国第一套地面成像光谱仪，由中国科学院遥感应用研究所和中国科学院上海技术物理研究所联合研制，其设计借鉴和参照了中国自行研制的航空推扫成像光谱仪 PHI 的设计原理，主要由三个系统即计算机系统、光机系统和电子学系统构成，其主要技术指标如表 2-1 所示。

表 2-1 FISS 主要技术参数指标和性能

波段数/个	344
光谱范围/nm	437-902
光谱分辨率/nm	优于 5nm
地面分辨率/mm	最高可优于 2
实验室辐射定标精度/%	优于 5
成像速率/（帧/s）	20
扫描视场/度	-20--20
量化位数/位	12
信噪比	60%波段>500
光谱采样间隔/nm	约 1.4



图 2-1 FISS 实际测量工作图

2.1.2 地面成像高光谱辐射测量系统测量规范

1、测量流程

1) 根据被测物体以及试验目的, 确定观测平台高度和视场范围, 并由此确定地面覆盖范围和空间分辨率。

2) 根据观测条件, 进行图像预扫描, 并观测实时光谱曲线来调整积分时间与光圈设置来确定合适的曝光量。由于成像光谱能量细分使得信噪比受限, 总的原则是应该让数据 DN 值处于 CCD 探元响应较好的线性响应范围内, 防止数据饱和或者曝光不足。

3) 布置好参考板, 并且最好使用与被测地物反射率相近的参考板, 尽可能保证参考板与被测地物测量条件一致。

4) 盖上镜头盖, 测量暗电流图像; 之后打开镜头盖, 获取地物与参考板数据; 测量时应尽量保证被测物体置于视场中心。如需要也可以对太阳直射光进行遮挡获取漫射光下的数据。

5) 实时观察图像和光谱曲线, 查看其清晰度、曲线形状等, 确保数据没有异常。

2、推荐参数设置

测量时需要设置的参数为积分时间、光圈等级、制冷温度、DN 动态范围、扫描帧速率、视场角等。

积分时间: 积分时间可以设置的范围为 1-100ms, 为了保证图像尽可能小的失真变形, 需要确定一个合适的积分时间。

光圈等级: F 值可以设置的范围为 2.8-22 之间。配合积分时间确定合适的曝光量。

制冷温度: FISS 系统可以实现比环境温度低约 20-25℃的制冷能力, 因此设置的制冷温度比当前环境温度低 20-25℃均可。

DN 动态范围: 可以设置为 0-4095, 根据所测地物类型而设置, 通常设置为最大值 4095, 若测量水体, 则应适当降低 DN 量程。

帧速率：地面成像光谱仪的帧速率通常取决于 CCD 相机的数据传输速率和数据采集模式。本系统中 CCD 相机采用 USB 接口进行数据传输，12 位数据采样，采用了拍照模式（照相机模式）进行数据采集，帧速率最大为 20 帧，一般推荐设置 10 帧。

视场角：视场角可以在正负 20° 之间变化，常规上设置在正负 15° 之内。

3、注意事项

1) 系统检验与标定

(1) 要每隔一段时间对 FISS 系统的进行一次光谱定标、辐射定标、几何定标等。

(2) 应该每半年对参考板进行一次重新标定，同时要配备多反射率等级的参考板。

2) 测量条件与环境

(1) 野外测量时，若无特殊需要，应选择地方时 9:30-15:30 之间、地面能见度不低于 10 公里、尽可能无云、无风的条件下测量。

(2) 室内测量时，要保证被测物体光照均匀一致，应使用卤素灯等光谱范围宽广的稳定光源，视样品需要可采用冷光源。

(3) 观测几何的要求：如果是为了配合或者模拟航天航空平台的观测，则要注意 FISS 观测几何条件与航空航天平台观测几何条件的一致性。

(5) 存放环境：要在干燥的环境中存放，湿度不大于 60%，温度保持在 $0-40^{\circ}\text{C}$ 之间。

3) 人员与记录

(1) 人员着装：为减少测量人员反射光对测量物体的影响，实验人员应着深色衣服。

(2) 人员位置：试验过程中，人员应该面向光源位于被测物体后，避免在目标区走动。

(3) 参数记录：记录 FISS 系统使用中的各种参数设置。

(4) 场景记录：对观测目标以及观测环境应该拍摄照片以记录真实情况。

4) 存储空间：由于 FISS 系统成像光谱图像数据量较大，每次应保证计算机有足够的存储空间。

2.1.3 地面成像高光谱辐射测量系统测数据预处理

在对 FISS 进行光谱定标、辐射定标之后，主要进行的预处理工作就是噪声去除和反射率的转换。刘波在使用 FISS 的过程中，总结了大量的野外和室内试验，较为全面的总结了 FISS 数据的预处理过程。针对 FISS 的噪声去除，有如下的几种推荐方法^[15]：

1、三维均值滤波处理，计算公式为：

$$s_{n_i, j, k} = \sum_{m, n, q} a_{m, n, q} \cdot s_{m, n, q} \quad (2-1)$$

其中 m, n, q 分别代表行列及波段序号。

2、三维中值滤波，中心像元值去其立方体窗口中数字中值。

3、先进行 PCA（或 MNF）变换，然后选取前面信息量大的成分进行反变换。

4、先进行 PCA（或 MNF）变换，对变换后的前面各个主成分采取空间与滤波，之后进行反变换，如下图 2-2 所示。

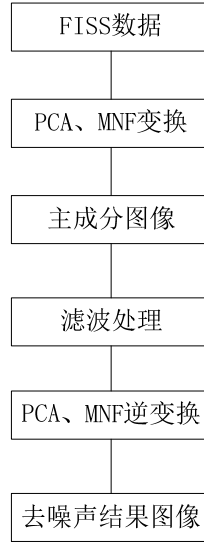


图 2-2 FISS 噪声处理流程^[15]

在 FISS 反射率转换方面，可以采用如下两种方法：

第一种方法类似于传统非成像光谱辐射计的反射率转换方法，利用如下公式：

$$R_{\lambda} = \frac{DN_{\lambda} - D}{W_{\lambda} - D} \cdot P_{\lambda} \quad (2-2)$$

其中 R_{λ} 为被测目标反射率， DN_{λ} 为被测目标 DN 值，D 为暗电流， W_{λ} 为参考板 DN 值， P_{λ} 为参考板反射率。当 FISS 参数设置合理的情况下，暗电流的 DN 值不超过 10，相对于被测目标 DN 值十分微小，可以忽略不计，此时上式可以进一步简化为：

$$R_{\lambda} = \frac{DN_{\lambda}}{W_{\lambda}} \cdot P_{\lambda} \quad (2-3)$$

此种方法要通过控制曝光将被测物体的 DN 值尽量落于 CCD 响应较好的量化范围内，并且使用的参考板反射率与被测物体接近。

第二种方法类似于经验线性法，假设图像 DN 数值与地物反射率 R 存在线性

关系：

$$DN_{\lambda} = a \cdot R_{\lambda} + b \quad (2-4)$$

通过获取已知光谱反射率的物体（通常情况下为参考板）的 DN 值，通过最小二乘法求解系数 a 和 b，进而可以求取反射率。可见，要采用这种方法最起码需要两种不同反射率的已知物体。

由于 FISS 为便携式地面成像光谱系统，在测量的时候可以很方便地获取同等观测条件下的一个或数个参考板数据，用于反射率转换。由此可见，上述两种反射率转换方法对于 FISS 都是简单易行的。

2.2 高光谱数据库

2.2.1 高光谱数据库的概念及其重要意义

高光谱遥感技术与应用的发展对高光谱数据存储与管理提出了新的要求，而数据库技术的不断发展，将能有效的解决在新领域面临的新问题。张雄飞根据这些新的要求，提出了高光谱数据库的概念，即高光谱数据库系统是专门面向高光谱数据，体现图谱合一特性，综合了光谱数据库、光谱分析功能和数据挖掘功能于一体的专用数据库系统^[16]。李兴在这个基础上，对高光谱数据库的概念进行了进一步的补充和发展：高光谱数据库不仅存储室内或野外光谱辐射计获取的目标光谱数据，它还存储以图像立方体形式存在的高光谱图像光谱数据。它能够将高光谱遥感应用的图像、光谱有机的融合起来，形成对高光谱影像的综合分析与应用^[17]。

高光谱数据库在面向星-空-地-地下各级高光谱数据的管理和行业应用上具有非常重要的意义。

(1)地面成像光谱辐射测量系统FISS的出现，弥补了以前光谱辐射计均采用点测量模式只获得测量对象的光谱曲线的缺陷。

尺度问题是任何对地观测手段都面临的问题，在传感器器视场角不变的情况下，不同的测量高度对应不同的象元地面大小。在目前成像光谱遥感科学与技术发展中，无论是航空还是航天成像光谱仪，其图像反映的是像元尺度上的光谱信息。然而以前基于点测量的野外光谱辐射计只适合点状地物的光谱测量，难以做到地面视场的精确设定。使得不同平台的高光谱仪器在空间尺度上衔接的不够紧密，因此在像元组分光谱分析、混合像元分解等问题上缺乏创新性和精确的测量手段。中科院遥感所牵头研制的FISS系统，保留了地面光谱测量的特点，也具有精确设定视场范围和高地面分解力的优势，尤其适合于混合光谱研究和可变尺度下的像元光谱组分分析。

但是目前的光谱数据库还没有考虑到应对这种集图像与光谱为一体的地面成像光谱辐射测量系统数据，FISS高光谱数据库系统的建设将可有效解决这个问题。

(2) 另外, 高光谱数据库在地质矿产资源的调查应用中也具有十分重要的意义。

中国地质调查局的大型科研项目“星空地一体化光谱关键技术与设备研发”以及国土资源部的公益项目“高光谱多维矿物填图技术与应用示范”, 将会得到一套航空高光谱遥感系统、一套地面成像光谱辐射测量系统以及一套岩芯成分探测与编录系统, 加上已经发射和即将发射的高光谱卫星, 将会得到星-空-地-地下一体的高光谱数据。通过卫星光谱探测技术与手段, 可以满足国内及全球尺度的陆表地质填图需要, 掌握全球矿产资源分布现状; 航空光谱探测技术及装备研制, 可以满足高分辨率的矿物填图和矿物靶圈圈定, 实现间接找矿; 地面光谱探测技术与装备研制, 将实现矿物精细识别与探测、精细矿产开采, 通过矿物含量的光谱定量反演技术, 为矿物冶炼提供精细的技术服务; 岩芯光谱探测技术及装备的研制, 将满足我国钻探岩芯的光谱自动化编录技术与水平。因此, 制定高光谱数据的处理标准与入库规范, 建立多源高光谱数据管理系统, 对即将获得的星空地地下高光谱数据进行有效管理, 将会极大的提高高光谱数据的利用率, 为深部找矿提供支持。本文将专门研究管理地面成像光谱辐射测量系统FISS数据的高光谱数据库, 并希望通过FISS高光谱数据库的研究, 为星-空-地-地下一体的高光谱数据库的建设提供一定的借鉴。

从以上两点可以看出, 不论是在一系列新型光谱测量仪器的出现还是在行业应用上, 高光谱数据库都具有十分重要的意义。

2.2.2 高光谱影像的几种管理方式

遥感影像数据的管理模式经历了文件模式、文件与关系数据库混合模式、纯关系数据库模式、对象关系数据库模式和面向对象数据库的管理方式^[18]。

1、纯文件管理模式

这种管理方式将遥感影像数据存放在操作系统的目录下, 根据具体的需要设计文件的存储结构, 来存储影像数据和相应的元数据。如果数据结构设计合理, 这种方式将是效率最高的一种数据管理方式。但这种方式也存在很多缺点, 如实现多用户之间的数据共享和并发控制比较困难, 数据安全性很难保证, 而且当数据文件多时会引起管理上的混乱, 同时备份也很烦琐。

2、文件与关系数据库混合管理模式

混合管理模式是指影像数据本身用文件系统管理, 而与其相关的属性数据(元数据)用关系数据库来管理, 它们之间用关键字进行连接。在这种管理模式中, 影像数据与属性数据除关键码连接外, 在组织、管理与检索上几乎是独立的。采用文件与关系数据库管理系统的混合管理模式, 还不能说建立了真正意义上的影像数据库管理系统, 因为文件管理系统的功能较弱, 特别是在数据的安全性、一致性、完整性、并发控制以及数据损坏后的恢复方面缺少基本的功能^[19]。

3、纯关系数据库管理模式

纯关系型空间数据库管理系统是指影像和属性数据都用现有的关系数据库进行管理。这种管理方式是目前遥感影像管理常用的方法之一，它将影像数据以大二进制对象（BLOB）的方式存储在数据库的一个字段中，有时也根据需要将特别大的影像数据分割成较小的数据块以多个记录的形式存放在数据库中。目前，Oracle、SQL SERVER 等大型商业数据库均提供了 BLOB 类型，其最大可存储 4G 的数据。这种方式对数据集中控制，数据冗余度小，数据独立性强，并发控制容易实现，数据库易恢复，数据安全性和完整性较好，也可实现多用户访问，元数据可以与影像数据无缝的集成管理。

4、对象关系数据库管理模式

对象关系数据库管理模式将影像数据作为对象来进行管理，并提供一系列方法。对象关系数据库管理模式既具有关系数据库的特点，又能有效地管理复杂数据，是当前影像数据库研究的一个热点。有两种方式，一种是用户自定义对象，另一种是数据库厂商提供的现成的对象。如 Oracle 数据库的 GeoRaster 模块就是在关系数据库的基础上，引入对象的概念，对关系数据库的数据类型进行扩展，使之能存储影像数据。同时像 Arc SDE 和 SuperMap SDX+空间数据库引擎，是利用关系数据库提供的各种接口而开发的专用程序，将影像数据分解成关系数据库的一个或多个数据表，实现对遥感影像的存储和管理，也可以算是对象关系数据库的管理模式。

5、面向对象数据库的管理模式

该模式用纯的面向对象的数据库来管理影像数据，可以支持对象的语义查询和基于内容的查询等。但由于影像数据本身的复杂性和系统开发的复杂性，使得系统开发很困难。面向对象数据库目前仍处于研究阶段，还没有比较完善的面向对象影像数据库管理系统，在技术上还很不成熟。

2.2.3 高光谱数据库的研究现状

张雄飞总结了前人的研究成果，研究了目前光谱数据库的发展现状，根据高光谱数据应用的需要，在国际和国内首次提出了高光谱数据库的概念。分析了系统组成，建成了一个基于网络的高光谱数据库原型系统。通过分析高光谱数据的特点，结合数据库开发实践，提出了高光谱数据集包括图像、光谱、属性等在关系数据库中的存储规范^[16, 20]。李兴在构建高光谱数据库基础之上，提出了针对数据库存储方案的光谱数据模型和影像数据模型，通过将必要的高光谱分析方法和分析模型整合到数据库之中，实现了数据、方法、模型在高光谱数据库中的集成与统一；将多源数据统一到一个数据库平台之中，将原有的大表结构和表群结构作转换，设计了以二元数据为核心的星型数据库概念结构，建立了通用的高光谱数据库建库模式和模型设计线路^[17]，同时他还研究了高光谱环境数据库组成和机

制^[21]以及面向对象的高光谱数据库及其应用^[22]。董广军等针对高光谱数据特性以及处理和管理的需要,提出了基于 Oracle 数据库的面向对象的高光谱数据管理模型,从而便于高光谱数据的存储、管理、查询;并基于高光谱数据库进行了 SAM 分类实验,验证了该模型的有效性^[23]。满旺等提出了将高光谱遥感数据按像元顺序或者波段顺序分割,按大型对象存储的技术思路,设计了高光谱遥感数据的存储结构,并开发了相应的数据管理平台。实际应用表明,该平台能够较好地完成高光谱数据的存储、管理与查询,取得了理想的应用效果^[24]。孙婷等提出了基于 Oracle 数据库平台的高光谱影像存储和管理的解决方案。基于对象关系模型设计了光谱数据表组+属性数据表组的数据存储规范以及波段独立顺列式、波段集中整合式以及表单位式的三种存储模式。采用地形瓦片和影像金字塔的快速索引结构设计和建立了高光谱影像数据库的原型^[20, 25]。向娟等利用成熟的关系数据库和程序设计语言,开发了一个方便实用的高光谱遥感影像管理系统以管理高光谱遥感影像和其它遥感信息,提高影像管理效率。系统已初步实现了多景高光谱遥感影像检索,可以任意加载、导入高光谱遥感影像,运行效果良好^[26]。

2.3 多维数据库理论

多维数据库 (Multidimensional Database, MDB) 可以简单地理解为:将数据存放在一个 n 维数组中,而不是像关系数据库那样以记录的形式存放。一般地, MDB 是面向查询对象 (类), 由其度量属性集、构成多维空间的维的集合以及在多维访问中的各点的值的集合共同描述^[27]。

2.3.1 多维数据库基本概念

(1) 数据立方体

数据立方体采用多维立体数据存储方式来取代传统的平面数据存储方式,为进行多维数据分析提供根本保证,是多维数据库中组织和存储数据的重要手段。多维数据库由一系列相关的数据立方体集组成。图 2-3 所示为最常见的销售数据立方体。该立方体只有三维,但是在理论上,数据立方体可以具有任意数量的维数。

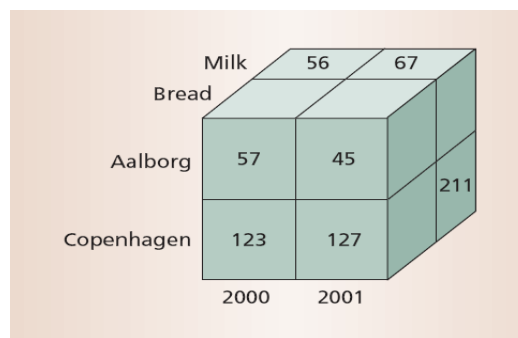


图 2-3 销售数据立方体

(2) 维

维是人们观察问题的角度，是多维数据库中至关重要却又与众不同的概念。多维数据模型的一个主要目标就是利用维为事实的访问提供尽可能多的途径，每个维代表一个统一的访问多维数据库的信息路径。如遥感工作者关心同一种地物在不同地点的光谱曲线的变化情况，则地点是一个维（空间维）。

（3）维层次

在实际中数据立方体的维常常具有多个属性，这些属性可按细节程度组织成层次形式。如空间维的层次可以分为国家、省、市、县、乡。

（4）维成员

维的 1 个取值称为该维的 1 个成员，若维是多层次的，则维成员是在不同维层次中取值的组合。

（5）事实

事实代表着要进行分析的且我们感兴趣的模式或事件，在大多数多维数据模型中，事实由其组合在一起的维值的隐式定义，只有当特定维值的组合没有造成空穴时，一事实才会存在，然而有些模型将事实看成具有独特个性的一级对象。大多数多维模型也要求将每件事实与每维最低级别的一个维值建立联系，但是有些模型放宽了这种要求。

（6）度量

度量包括两部分内容：一是事实的数量属性，如产品销售中的价格或利润；另一个是公式，通常是 1 个简单的综合函数，比如 sum，它能够将几个度量值组合为 1 个。例如在商品交易中，一次交易就是一个事实，交易的数量和价格就是度量。

2.3.2 多维数据库分析方法

多维数据库的基本分析包括钻取、切片、切块和旋转。其中，钻取包括上钻和下翻，是对多维数据进行概括分析的方法。切片和切块是对多维数据进行局部分析的方法。旋转是对多维数据进行整体分析的方法。在此简要介绍各个操作的概念。

（1）上钻

上钻和下翻主要是通过改变多维数据的维数或维的层次来实现。上钻包括维上钻和维层次上钻。维上钻是在 n 维数据立方体的基础上，综合一个或更多的维（维₁，维₂，…，维 _{m} ，其中 $m < n$ ），聚集出相应度量值的过程。维层次上钻就是在多维数据立方体的基础上，不综合掉任何一个或更多的维，在某维的层次链上按照其维的成员从低等级到高等级进行详细到概括的操作，聚集出相应度量值的过程。

（2）下翻

下翻包括维下翻和维层次下翻。维下翻是维上钻操作的逆操作，基于 n 维数

据立方体,在已上钻数据立方体(维1, 维2, ..., 维 m)的基础上(其中 $m < n$),再额外产生一个或更多的维(维1, 维2, ..., 维 k),使这一个或更多的维具有成员值,其中 $m < k < n$,并计算出相应度量的过程。维层次下翻是维层次上钻操作的逆操作,基于 n 维数据立方体,在维层次上钻的基础上,不增加任何一个或更多的维,在某维的层次链上,其维的成员从高等级到低等级,并计算出相应度量的过程。

(3) 切片和切块

切片和切块是对多维数据进行局部分析的方法,是在多维数据立方体一部分维上选定成员值后,即固定一部分维,关心度量数据在剩余维上的分布。如果所选定的部分维只有一个,则是切片;如果所选定的部分维有两个或两个以上,则是切块。上钻或下钻可以和切片或切块同时进行。

(4) 旋转

旋转就是变换数据立方体维的角度,从不同的方向关心度量数据在维上的分布。在(维1, 维2, ..., 维 n , 度量)空间数据立方体的基础上,每个维彼此在垂直方向固定,为了变化人们的观察视点,需改变数据立方体维的方向,对整个数据立方体进行方向变换,这个过程称之为旋转。数据立方体的旋转并不改变度量的值,只是变换了人们观察分析的角度。

2.3.3 多维数据库的存储方式

多维数据库的数据存储方式可分为基于关系数据库的存储方式、基于多维数据库的存储方式和混合存储方式三种。

(1) 基于关系数据库的存储方式

将数据存储的关系数据库的表结构中,用星型模型和雪花模型来组织数据,在元数据的管理下,可以实现多维数据库的功能,但是它的底层数据库是关系数据库。

这种方式的优点是灵活性强,用户可以动态定义统计或计算方式,不需额外的存储空间存放信息,利用既有的资料存储空间即可;缺点是在多维数据模型定义好后,从数据库中抽取数据往往需要编制复杂的程序,反应时间较慢,较适合应用在不常使用且量大的情况,通用性差,很难维护。

(2) 基于多维数据库的存储方式

利用一个专有的多维数据库来存储数据,数据以多维方式存储,有维索引及相应的元数据管理文件与数据相对应,并以多维视图方式显示。在这种方式中,数据在被存入多维数据库时,将根据它们所属的维进行一系列的预处理操作,并把结果按一定的层次结构存入多维数据库中。

这种方式的优点为在进行多维分析操作时反应时间较快,缺点为存储空间需求大,适合应用在使用频率高、量小且反应时间要求快的情况。

(3) 混合存储方式

在数据存储时采用了多维数据库和关系数据库混合型架构, 有两种实现方法: 第一种方法是, 对于常用的维度和维层次, 采用多维数据表来记录, 对于不常用的维度和数据, 采用类似于关系存储方式的星型结构来存储。适用于在大量数据下仍要求快速反应的情况。

第二种方法是, 利用一个多维数据库存储高级别的综合数据, 同时用关系型数据库管理系统 (RDBMS) 存储细节数据。这种方法是目前实现混合存储结构较理想的方法, 它结合了关系数据库方式和多维数据库方式的优点。

2.4 本章小结

本章主要从地面成像光谱辐射测量系统 FISS、高光谱数据库以及多维数据库三个方面介绍了本文开展研究的相关理论基础。首先简单介绍了 FISS 的主要参数、测量流程、参数设置、注意事项以及数据预处理的内容以及方法。其次阐述了高光谱数据库概念的提出及其发展过程, 总结了影像数据的五种管理方式, 介绍了高光谱数据库的国内外研究现状。最后介绍了多维数据库有别于关系数据库的基本概念、基本操作以及多维数据库的三种数据存储方式。这些基本理论、知识为本文的后续研究奠定了理论基础。

第三章 FISS 高光谱数据库的分析与设计

随着高光谱数据在地质、农业、林业、城市土地利用、环境保护等诸多领域日益广泛的应用，光谱数据库也如雨后春笋般建立起来。FISS 系统自投入使用以来，已经进行了大量的室内和野外试验，在食品安全、杂草识别、植物叶片生化信息反演、近海环境调查等应用中取得了有益的研究成果，获取了大批数据。如何综合考虑各方面的应用需求，结合 FISS 数据特有的特点，进行面向数据管理和初步分析应用的 FISS 高光谱数据库设计，是亟待解决的一个重要问题。

本章将针对 FISS 高光谱数据库的功能需求，分析系统的业务流程，设计系统的体系结构和功能模块，简单介绍系统的开发环境，提出 FISS 数据存储模型，并在 Oracle 数据库里设计各个数据表。

3.1 功能需求分析

3.1.1 系统功能需求

建立 FISS 高光谱数据库，把测量到的大量 FISS 数据进行有效管理和长期的安全存放，实现快速检索和有效管理，对于推广 FISS 数据应用，促进科学研究来说，具有非常重要的意义。考虑到 FISS 数据特有的特点并结合其它高光谱遥感应用的特点，FISS 高光谱数据库要满足以下的功能需求。

1、海量数据存储

高光谱数据最大的特点就是其海量的数据和众多的波段，每景数据根据数值类型的不同，动辄就是一个 GB 乃至两个 GB 以上，如果有一千景数据，其数据量很容易就达到 TB 级，可以预见，在不久的将来，随着 FISS 应用的不断推广，其获取的数据将会越来越多，极有可能达到 PB 级，这就需要一个海量的存储系统。而且在实际应用中，可能需要对 FISS 数据进行空间截取或者光谱波段截取等更加详细的要求，这就需要考虑一个高效的存储结构，以提高检索和读取效率。可见，在设计 FISS 高光谱数据库时，既要考虑大容量数据存储的高效性，另一方面也要考虑到进行操作时的灵活性。

2、FISS 数据的完整性

从广义的角度来理解，高光谱数据不应只包括影像数据，同时还应包括配套地面测量光谱数据、配套测量参数还有针对应用的属性参数等等。对于 FISS，它是地面成像光谱仪，在测量时积分时间、光圈等级、制冷温度、DN 动态范围、扫描帧速率、视场角等参数都是可以设置的，为便于后续数据处理和理解，需要记录下这些参数并连同 FISS 数据一到存入数据库中。除了测量参数外，还有其它的面向应用的属性，如测量目标、测量时间、测量地点、经纬度信息等。如果将这些信息整合到一起，保证数据的完整性，需要对数据存储模型以及数据存储结构进行精细的研究。

3、数据的快速查询检索

FISS 数据与其它的高光谱数据略有不同，其辅助数据和属性没有那么复杂，因此其查询检索也显得相对简单，但是，随着数据记录的增加，也要考虑一个合适的索引结构，来提高其查询效率。同时，为了充分体现 FISS 数据高空间高光谱分辨率的优势，除了使用影像属性信息进行检索以外，可以考虑基于图像内容的检索，以实现数据的综合信息检索，能够快速高效的检索到用户需要的数据。

4、图谱合一的存储与显示

高光谱遥感的一个重要特点是从平面图像空间能够提取出像元的光谱曲线，直观的提供地物目标的光谱信息，即所谓的图谱合一特性，高光谱数据库应当继承这一特点。因此，在数据库存储时，应当考虑如何将变化性很强的影像光谱维数据与空间维数据很好的结合在一起，同时，在显示的时候，也要在显示影像空间维信息的同时，能够提取出地物的光谱曲线，实现高光谱数据图谱合一的存储与显示，这就需要我们选择一个合适的存储方式和数据模型。

5、光谱分析功能

FISS 高光谱数据库，不应仅仅是能存储、查询和显示高光谱数据，为了发挥 FISS 数据高空间高光谱分辨率的优势，在光谱曲线显示之后，还应当具有一些光谱分析功能。比如，光谱平滑、光谱重采样、包络线去除以及光谱匹配等，以便帮助用户更好的辨别当前影像有可能包含了那些地物，有利于数据的进一步选择。

6、系统的扩展性

该系统目前主要存储的是 FISS 数据，所以要从以下两个方面来考虑系统的扩展性：一个是纵向上的扩展，即各种功能的增加和完善，如更多的分析功能或者对数据模型的一些扩展；一个是横向上的扩展，即增加对其它类型成像光谱数据的管理以及增加针对这些成像光谱数据的数据模型，从而使 FISS 高光谱数据库系统变成完全的高光谱数据库。所以在系统设计时，考虑到系统的延展性，可扩展性，设计好接口。

3.1.2 系统业务流程分析

高光谱数据库的业务流程一定程度上决定了高光谱数据库的系统结构和功能模块的划分。传统的遥感数据库的业务流程主要是基于数据管理的目的，其核心在于数据存储和发布，重点在于数据库，而不是高光谱。在分析 FISS 高光谱数据的功能需求后，将 FISS 高光谱数据库的重点稍微向高光谱转移，使该数据库既有数据管理功能也具有一些高光谱特色的功能。为此，设计了如图 3-1 所示的系统业务流程图。

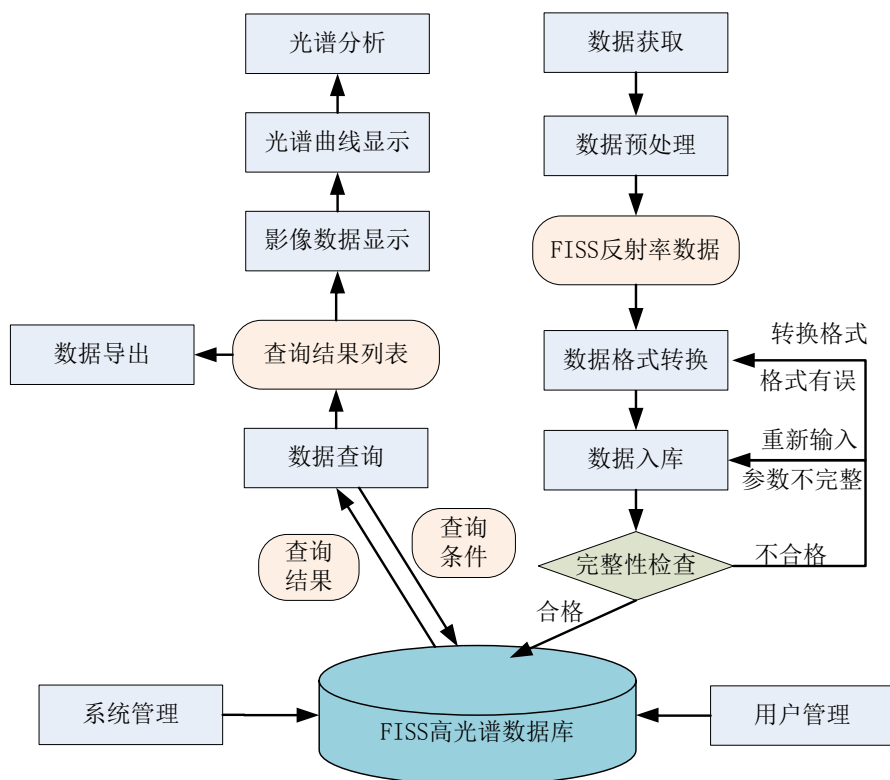


图 3-1 FISS 高光谱数据库业务流程图

3.2 总体结构设计

3.2.1 体系结构设计

1、系统结构设计的原则

高光谱数据库系统结构设计应该遵循数据库设计的一般原则，即一致性、完整性、安全性、可靠性和效率原则，每个原则的具体内涵如下：

(1) 一致性原则 (Consistency)

一致性是指如果多个用户同时、以同样方式、对同一数据查询，数据库的回答结果是一样的，那么称数据库对多用户具有一致性。一致性考虑的参数有：修改数据的方法；修改的响应范围；数据修改的算法；执行强制修改前等待时间；执行强制修改时其它活动用户的等待时间等；同时要求发生修改同一数据的用户数目（并发度）。

(2) 完整性原则 (Integrity)

完整性是指数据库数据的正确性和一致性，数据库中的数据值应满足指定的约束，且对数据库进行更新后仍然保持这种性质。完整性考虑的参数有：参照完整性、实体完整性、域完整性；强制和有效执行约束——触发条件、测试出违反完整性约束时采取的措施（动作）等。

(3) 安全性原则 (Security)

数据库安全性是指对数据库有意或无意的泄漏、修改或丢失的保护能力。设

计数据库安全性的主要目的是以最小的代价防止对数据的非法的使用。实现方法是控制对数据库数据的访问，DBMS 提供控制访问的功能，如创建子模式或存储过程、授权/收权、授予角色、用户确认、审计等功能。

（4）可靠性原则（Reliability）

数据库的可靠性包括：故障发生的可恢复性（Recoverable）、故障恢复所需要的时间和故障发生的频率。当数据库发生故障时，具有恢复数据库完整性的能力，称数据库具有可恢复性。可恢复性的设计过程包括建立一个检查系统，防止事务和数据的丢失；当故障发生时，在合理时间内把事务和数据的状态恢复到故障发生前的情况。一般数据库管理系统都具有这一功能，如数据库备份、镜像、日志等。

（5）效率原则（efficiency）

主要是指计算机资源的利用和系统的响应时间。效率原则的目标包括三个部分：模式设计的合理性，使每个应用的执行时间最小，联机响应时间不超过设计的要求值；查询优化，最小的数据传送量和通信次数、最优处理顺序、最少I/O 次数和操作量、最优访问路径；合理的数据冗余，处理时间/存储与执行空间的最优选择。例如在模式中保留一定的冗余（属性的冗余）可以减少检索的响应时间，但它增加了新的开销和一致性问题，这根据数据库应用的具体需求而定。

2、FISS 高光谱数据库体系结构设计

体系结构设计是数据库系统设计的重要一步，高光谱数据库的结构设计，要能体现高光谱数据库的扩展性能，网络性能，和跨平台性能，这就要求必须要选择一个合理的系统结构。当前比较流行的系统结构是三层的浏览器/服务器（Browse/Server）结构，客户机/服务器（Client/Server）结构虽然在某些方面有所不足，但是在系统结构设计、开发等方面的难度和复杂性较低，也在被广泛的使用。FISS 高光谱数据库主要存储和管理 FISS 影像及其配套参数，其长远的发展规划是用 B/S 结构和 C/S 结构相结合的模式，即网络版和桌面版。本文只讨论桌面版的体系结构、详细设计和及其实现。桌面版 FISS 高光谱数据库系统体系结构由三层组成，包括用户表现层、应用接口层和数据库服务层。

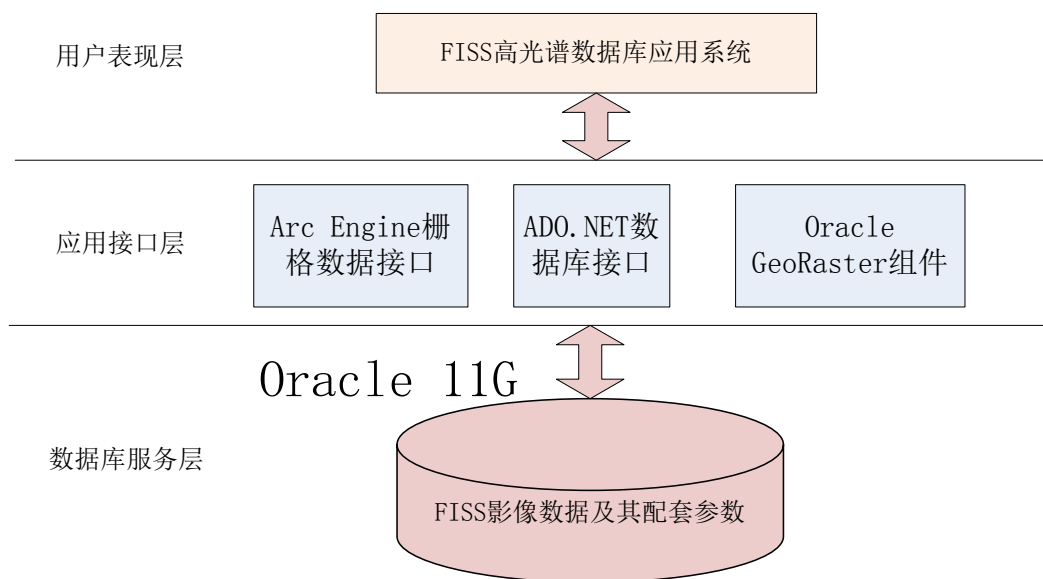


图3-2 高光谱数据库系统结构图

用户表现层：用户表现层是连接用户和应用层的桥梁，它把用户的请求转化为应用逻辑，提交给应用接口层，并处理返回的结果。本研究用户表现层主要是 FISS 高光谱数据库应用系统，实现用户应用需求的输入以及处理结果的展示，包括输入输出、查询检索、可视化显示、数据更新、用户管理和系统管理六个功能。

应用接口层：主要是实现 C# 和 Oracle 进行通信的 ADO.NET 数据库接口，以及用来支持栅格影像显示的 Arc Engine 的栅格数据类，以及类似于中间件的 Oracle GeoRaster 对象的相关接口。这些接口在整个系统的开发中发挥非常重要的作用。

数据库服务层：主要包括 FISS 高光谱数据及其配套参数库。高光谱数据的特殊性决定了其数据访问方式和检索方式需要重新审视，不仅要考虑海量数据在数据库中的存储，还要考虑对大对象的读取效率，更要考虑对于将来存储空间的扩展和分布式数据库的拓展。为了降低应用层与数据库的耦合，方便数据库的移植，需要我们认真进行数据库的存储结构设计，以便改善性能，降低系统开销，提高处理效率，方便系统维护和功能扩展，对于高光谱数据库系统而言，也是非常有益的。

3.2.2 功能模块设计

在系统功能需求和业务流程分析的基础上，根据系统的结构设计，将 FISS 高光谱数据库划分为如图 3-3 所示的几大功能模块。

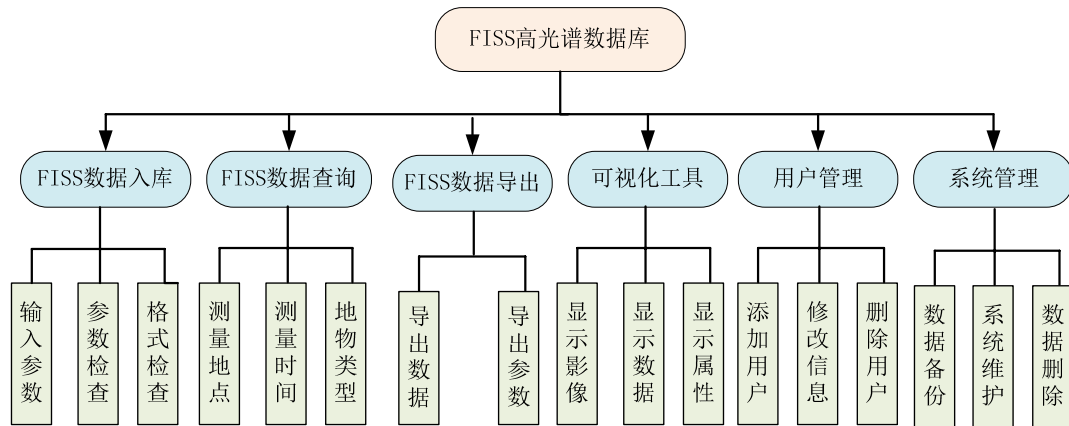


图 3-3 FISS 高光谱数据库功能模块组成图

1、数据入库

主要实现 FISS 数据入库的功能。FISS 数据在入库之前，需要进行一系列的预处理操作，得到反射率数据，然后还要转换成以 BIP 格式存储的 TIFF 类型数据。入库时，进行配套参数的完整性检查以及数据格式检查。如合格的，将数据入库，若不合格，则重新进行格式转换或者重新输入参数，直到符合数据入库要求为止。

2、数据查询

主要实现对 FISS 数据的查询。因为 FISS 是地面成像光谱仪，所以它的每一次测量都会有比较明确的目标，这与机载或星载的高光谱数据有明显的不同。可以按照影像拍摄时间、采集地点、主要测量目标名称三个要素来查询。测量时间查询可以指定某个特定的时间也可以选择某一个时间范围，采集地点查询可以按照行政区划的不同级别层层查询，如按照省、市、县不同级别来查询。

3、数据导出

当查询出符合条件的数据，需要对该数据进行进一步研究时，可以把该数据导出。亦或者是在对数据进行影像可视化或者光谱可视化或者是进行简单的光谱分析之后，想对该数据进行进一步的研究，则导出该数据。导出时可以选择只导出影像还是同时导出影像及其配套的参数，包括测量参数和属性参数。

4、可视化工具

在可视化部分，首先要能显示满足查询条件的各个影像的属性信息，其次才是显示 RGB 合成的彩色影像，同时还能显示每一个像元的光谱曲线。在显示影像时，可以对影像进行放大、缩小、漫游等操作。在显示光谱曲线时，可以对曲线进行光谱平滑、光谱重采样、光谱匹配、包络线去除等操作。

5、用户管理和系统管理

主要是实现用户信息添加、修改、删除等操作。系统管理是指数据安全、数据库维护、数据扩充等。安全管理是指建立完整的用户管理体系，对数据进行严格的权限设置，保证存储在数据库中的数据可被正确、合理地使用。同时还需

要对数据进行维护和修改、删除等管理操作。

3.3 系统开发环境

系统的开发环境如表 3-1 所示：

表 3-1 系统开发环境

计算机配置	内存为 4.0G CPU 为 Intel(R) Core(TM)2 Quad Q9550 主频为 2.83GHZ 操作系统为 Windows 7
数据库管理系统	Oracle 11G
开发工具	VS2008 (C#语言) ADO.NET (数据库接口) Oracle SQL Developer SQL 语言 Arc Engine

3.3.1 数据库管理系统

本次研究的数据库管理系统为功能强大的 Oracle 数据库，Oracle 是一个成熟的、大型的数据库应用开发软件，它突破传统的关系数据库理论，引入面向对象技术，使得对各种复杂数据类型的管理成为了可能。从 1978 年 Oracle 1 诞生至今，Oracle 已经发展到了 11G，它具有以下主要特点^[28]：

- 1、遵守数据存取语言、操作系统、用户接口和网络通信协议的工业标准。
- 2、支持大数据库、多用户的高性能事务处理。

Oracle 支持最大数据库（如几百 TB 等），可充分利用硬件设备。支持大量用户同时在同一数据上执行各种应用，并使数据争用最小，保证数据一致性。

- 3、更好的安全性和完整控制。

Oracle 有用户鉴别、特权（甚至到列级）、角色、触发器、日志、备份等功能，有效地保证了数据存取的安全性与完整性以及并以控制和数据的恢复。对数据安全性要求较高的系统（如银行系统）大都选 Oracle 作数据库管理系统（DataBase Management System, DBMS）。

- 4、硬件环境独立。

Oracle 具有良好的硬件环境独立性，已经用于美国十大计算机公司，而且支持各种类型大型和中型机、小型机和微机系统。它也用于国内的联想等系列微机上。

- 5、具有可移植性、可兼容性和可连接性。

Oracle 不仅可以在不同型号的机器上运行，而且可以在同一厂家的不同操作系统支持下运行。具有操作系统的独立性。

- 6、支持分布式数据库和分布处理。

Oracle 允许将处理分为数据库服务器和客户应用程序处理,所有共享的数据管理由具有数据库管理系统的计算机处理,而运行数据应用的工作站侧重于解释和显示数据。通过网络(甚至远程广域网)连接的计算环境,Oracle 将存放在多台计算机上的数据组合成一个逻辑数据库,可由全部网络用户存取。分布式系统像集中式数据库一样具有用户透明性和数据一致性。

7、支持对空间数据包括栅格影像数据的处理。

Oracle Spatial 是 Oracle 公司推出的空间数据库组件,通过 Oracle 数据库系统存储和管理空间数据,空间数据提供了较为完备的支持。GeoRaster 是 Oracle 数据库从 10G 开始引入的 Oracle Spatial 的一个特性,它允许存储、索引、查询、分析和传送 GeoRaster 数据,即影像和网格化栅格数据及其相关元数据。GeoRaster 提供了 Oracle Spatial 数据类型和一个对象关系模式,可以使用 GeoRaster 来存储多维的网格化数据和栅格层,而这些数据和栅格层可以参照到地球表面或者本地地球坐标系统中的坐标位置。如果数据是地理参照数据,则可以找到栅格单元所对应的地理位置;而如果给定地球上的位置,则可以找到与该位置相关的栅格层的单元。GeoRaster 为大型影像和栅格数据处理解决方案提供企业级数据管理功能。

3.3.2 开发语言

本次开发采用的语言是 VS2008 平台的 C#语言。C# (C Sharp)是微软(Microsoft)为 .NET Framework 量身订做的程序语言,C#拥有 C/C++的强大功能以及 Visual Basic 简易使用的特性,是第一个组件导向(Component-oriented)的程序语言,和 C++与 Java 一样亦为对象导向(Object-oriented)程序语言。这从根本上保证了 C#与 .NET 框架的完美结合。在 .NET 运行库的支持下,.NET 框架的各种优点在 C#中表现得淋漓尽致。C#具有以下突出的特点:

1、简洁的语法: C#对 C++语法中的冗余此进行了简化,只保留了常见的形式,同时没有了指针,并用真正的关键字换掉了那些把活动模板库(Active Template Library, ALT)和 COM 搞得乱糟糟的伪关键字。

2、精心地面向对象设计: C#具有面向对象的语言所应有的一切特性:封装、继承与多态性。在 C#的类型系统中,每种类型都可以看作一个对象。C#提供了一个叫做装箱(boxing)与拆箱(unboxing)的机制来完成这种操作。

3、与 Web 的紧密结合: C#对 SOAP 的使用使得与 WEB 标准相统一的大规模深层次的分布式开发从此成为可能。C#组件将能够方便地为 Web 服务,并允许它们通过 Internet 被运行在任何操作系统上的任何语言所调用。

4、完整的安全性及错误处理: C#提供了包括类型安全在内的完整的安全性, .NET 运行库提供了代码访问安全特性。C#中不能使用未初始化的变量; C#不支持不安全的指向; 不能将整数指向引用类型。

5、版本处理技术：C#提供内置的版本支持来减少开发费用，使用 C#将会使开发人员更加轻易地开发和维护各种商业用户。

6、灵活性与兼容性：在简化语法的同时，C#并没有失去灵活性。同时 C#允许将某些类或者类的某些方法声明为非安全的，将能够使用指针、结构和静态数组，并且调用这些非安全代码不会带来任何其它的问题，体现了较好的兼容性。

3.3.3 ADO.NET 数据库接口

ADO.NET 是 Microsoft .NET 应用程序的数据访问模型。它由 ADO 技术发展而成，在某种程度上，ADO.NET 代表了最新版本的 ADO 技术。ADO.NET 由两个核心组件组成：数据提供程序（Data Provider）和数据集（DataSet）。Data Provider 是一组包括 Connection, Command, DataReader 和 DataAdapter 对象在内的组件。它实现了对数据库的连接、操作和快速、只读的访问。DataSet 是 ADO.NET 的非连接结构的核心，它的设计目的是为了实现在任何数据源的数据访问，在本地内存中实现一个数据缓存^[29]。其结构示意图如图 3-4 所示。

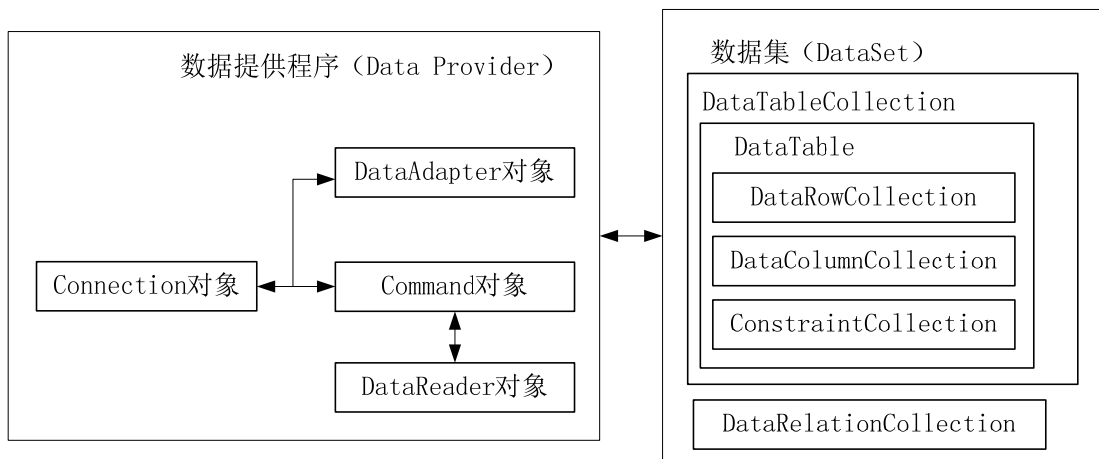


图 3-4 ADO.NET 结构示意图

由于 ADO.NET 建构于 .NET 框架之内，它的建立和管理都是基于 CLR（Common Language Runtime，通用语言运行时），所以直接或间接地得益于 .NET 框架在内存管理、类型转换、对象池等方面的技术改善和优化。

3.3.4 图形化数据库开发工具 SQL Developer

Oracle SQL Developer 是一个图形化的数据库开发工具，用户通过它可以浏览数据库，运行 SQL、PL/SQL 和 SQL*Plus 语句，执行 SQL 脚本，以及编辑和调试 PL/SQL 语句。它具有以下特点：（1）以 JAVA 编写而成，能够提供跨平台工具，可以运行在 Windows、Linux 和 MAC OS X 上。（2）使用瘦 JDBC 驱动程序作为默认连接，意味着无需安装 Oracle 客户端也可连接数据库。（3）使用的扩展框架借用了 JDeveloper，提供了一个包含示例和文档的扩展开发工具包，可以方便的进行扩展。

3.3.5 Arc Engine

本次开发采用 Arc Engine 来支持 FISS 影像的放大、缩小以及导航等功能。Arc Engine 是一个简单的、独立于应用程序的 Arc Objects 编程环境，开发人员用于建立自定义应用程序的嵌入式 GIS 组件的一个完成类库。Arc Engine 是有一个软件开发包和一个可以重新分发的为 ARCGIS 应用程序提供平台的运行时 (Runtime) 组成。Arc Engine 的功能层次由以下 5 个部分组成^[30]。

- 1、基本服务：由 GIS 核心 Arc Objects 构成，如要素几何体和显示。
- 2、数据存取：Arc Engine 可以对许多栅格和矢量格式进行存取，包括强大而灵活的地理数据库。
- 3、地图表达：包括用于创建和显示带有符号体系和标注功能的地图的 Arc Objects，以及包括创建自定义应用程序的专题图功能的 Arc Objects。
- 4、开发组件：用于快速应用程序开发的高级用户接口控件和高效开发的一个综合帮助系统。
- 5、运行时选项：Arc Engine 运行时可以与标准功能或其它高级功能一起部署。

Arc Engine 包含了 System、SystemUI、Geometry、Display、Server、Output、Geodatabase、GISClient、DataSourcesFile、DataSourcesGDB、DataSourcesOleDB、DataSourcesRaster、GeoDatabaseDistributed、Carto、Location、NetworkAnalyst、Controls、GeoAnalyst、3Danalyst、GlobeCore、SpatialAnalyst 等 21 个类库，用来支持各种对矢量数据和栅格数据的分析和处理。

3.4 FISS 数据存储模型设计

3.4.1 FISS 数据特点

FISS 数据是典型的高光谱影像，具有与高光谱影像一样的各种数据结构。高光谱影像数据通常以三种格式排列：BIP（波段按像元交叉）、BIL（波段按行交叉）和BSQ（按波段顺序）格式。设P、L、B 分别表示像元维、扫描行和波段维，以三维数组表示图像D 有：BIP 格式为D (P, L, B)，其中波段维B 为最低维，扫描行L 为最高维；BIL 格式为D (P, B, L)；BSQ 格式为D (L, B, P)^[31]。一般成像光谱仪是以BIP 格式获取数据的，也有些面阵成像光谱仪以BIL 格式获得数据，如FISS在获取数据时就是BIL格式的。

FISS 数据拥有 344 个波段，每个数据的数量也很大，一般都是 1 个 G 以上（不同的数据类型数据量大小有所不同），数据的存储格式将会极大的影响数据的处理效率。例如，要提取某个像元的光谱曲线，分析它的光谱特性，就需要访问该像元所有波段的值，这时以 BIP 格式存储比较合适。如果影像是以 BSQ 或者 BIL 格式存储，则需要从不同的地方读取数据，由于 FISS 数据的众多波段和海量数

据，将需要大量的时间。如果是进行空间分析，则是以 BSQ 格式存储比较合适。BIL 格式介于 BIP 和 BSQ 之间，既不利于空间分析也不利于光谱分析。由于 FISS 高光谱数据库侧重于对 FISS 光谱信息的显示和分析，因此，采用 BIP 格式作为 FISS 数据的存储方式。

FISS 是地面成像光谱仪，具有便携灵活的特点，与点辐射计一样，它对目标也有多种测量方案，包括实地测量，将目标采集带回实验室测量或者采集后择机进行室外测量。同时，FISS 在测量时可以设置其配置参数，所以，FISS 数据的配套参数比较复杂，如何准确清晰的描述 FISS 的配套参数是一个需要充分考虑的问题。

3.4.2 FISS 数据存储模型

1、FISS 数据模型的核心—Georaster 对象

GeoRaster 是一个基于组件的、逻辑分层并且多维的通用栅格数据模型。它是在 Oracle Spatial 选件的基础上，采用面向对象关系技术，来实现对遥感影像数据的存储和管理。其存储结构示意图如图 3-5 所示。影像数据通过 SDO_GEORASTER 数据类型存储到数据库的影像数据表中，该表存储与该影像数据相关的信息，如获取时间、经纬度等，并不真正存储栅格数据，是影像数据的逻辑实体。因为 SDO_GEORASTER 没有存储实际的栅格数据，因此常规表操作的响应速度不会受影响，而对表中栅格数据的操作则可以通过 SOD_GEORASTER 对象提供的方法很方便的实现。栅格数据以分块的形式存储在栅格数据表中，每个块存储为栅格数据表的一条记录。

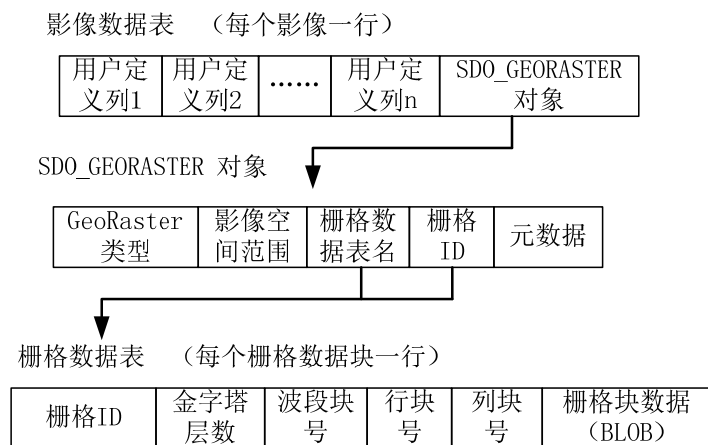


图 3-5 GeoRaster 模型的物理存储结构示意图

在物理存储上，GeoRaster 数据使用两种对象类型：SDO_GEORASTER 和 SDO_RASTER。SDO_GEORASTER 包含了栅格数据集和相关的元数据，这种对象类型的表称为 GeoRaster 表；SDO_RASTER 包含了栅格数据的每一个块信息，它使用 BLOB 对象存储块得栅格项数据。这种对象类型的表称为栅格数据表（RDT）。

SDO_GEORASTER 对象类型的定义是：

```
CREATE TYPE sdo_georaster AS OBJECT (
```

```
rasterType NUMBER,  
sSpatialExtent SDO_GEOMETRY,  
rasterDataTable VARCHAR2(32),  
rasterID NUMBER,  
metadata XMLType );
```

其中,

(1) rasterType 属性必须由五位整数组成, 用来表示栅格数据类型。第一位数字表示栅格数据的维数。第二位数字表示波段数或数据层数, 0 表示一个波段, 1 表示一个以上波段。需要注意的是这个数字不是数据的所有波段。第四位提供将来使用目前默认为 0, 后两位目表示 GEORASTER 的类型, 目前统一使用 01。

(2) SpatialExtent 是 Oracle Spatial 中定义的空间矢量数据对象类型 SDO_GEOMETRY, 用来表示栅格数据的空间范围。

(3) rasterDataTable 属性标识了存放栅格数据块的栅格数据表名称。栅格数据表是 SDO_RASTER 的对象表, 具有 SDO_RASTER 所有的数据段。

(4) rasterID 属性存储在栅格数据表的数据行中用来标识该行属于哪个 GeoRaster 对象。rasterID 属性与 rasterDataTable 属性组合在一起便唯一标识了整个数据库中的一个 GeoRaster 对象。

(5) metadata 属性是存储栅格数据对象元数据的 XMLType 对象。GeoRaster 采用 XMLSchema 定义栅格对象的元数据, 并提供获取和设置 XML 元数据值的函数以及验证其准确性的函数。

SDO_RASTER 对象是实际存储栅格数据的数据对象。栅格数据表是根据此对象建立的对象表。SDO_RASTER 对象类型的定义是:

```
CREATE TYPE sdo_raster AS OBJECT(  
rasterID NUMBER,  
PyramidLevel NUMBER,  
bandBlockNumber NUMBER,  
rowBlockNumber NUMBER,  
columnBlockNumber NUMBER,  
blockMBR SDO_GEOMETRY,  
rasterBlock BLOB);
```

(1) rasterID 属性, 与 SDO_GEORASTER 中的 rasterID 对应, 代表栅格数据对象的标示码。

(2) PyramidLevel 属性表示当前数据块的金字塔层数。“0”代表存储的是原始数据。

(3) bandBlockNumber, rowBlockNumber, columnBlockNumber 三个参数分

别表示当前数据块的波段编号、行编号和列编号，它们三个可以唯一确定当前数据块在原始栅格图像中的位置。

(4) blockMBR 参数存储当前数据块的最小外包矩形，与栅格数据的空间延展属性一样，采用 SDO_GEOMETRY 字段存储，方便空间计算。

(5) rasterBlock 存储实际的栅格数据，以 LBOB 对象的形式存储在数据库中。

GeoRaster 模块还提供了可以对 GeoRaster 数据和元数据进行操作和查询的语言包，包括 SDO_GEOR、SDO_GEOR_UTL 和 SDO_GEOR_ADMIN，这些语言包里包含了很多函数和存储过程，如影像分块、建立空间索引、建立金字塔、像元值获取、波段数获取等，这些函数可以在 PL/SQL 中使用。同时 GeoRaster 还提供 OCI、OCCI、JAVA 等方式进行开发。

2、FISS 数据存储模型

在 FISS 高光谱数据库系统中，FISS 数据是以 FISSIMG 对象类型存储在数据库中的。为继承高光谱数据的数据特点和应用，参考李兴提出的高光谱影像数据模型^[17]，在 FISS 数据存储模型的设计中，把 FISS 数据的主要应用参数集成到模型中，使得原始影像数据与影像参数集成在一起，大大提高数据应用的效率。

FISS_IMG
IMG_ID:number OBJ_TYPE:varchar IMG_LOC:varchar IMG_DATE:date IMG_TIME:varchar IMG_HEIGHT:number IMG_LONG:number(7,4) IMG_LAT:number(6,4) IMG_DATA:sdo_georaster

图 3-6 FISS 数据存储模型

FISSIMG 的属性主要包括了对 FISS 数据本身的描述，包括在数据库中的唯一标识，观测目标名称、观测目标采集地点、测量日期、测量时间、测量高度、目标采集时的经纬度等。在此需要说明的是，FISSIMG 模型的观测目标采集地点和经纬度属性指的是观测目标采集时的地点以及经纬度，而不是测量时的地点和经纬度。

3、FISS 数据入库的格式

在 Oracle 11g 中，GeoRaster 只支持以下有限的影像数据类型：TIFF、GeoTiff、JPEG、BMP、GIF、PNG 和 JP2^[32]。在这些影像数据类型中，只有 TIFF 和 GeoTIFF 格式可以成为高光谱影像的数据类型。

3.3.1 节的分析表明，以 BIP 格式存储数据更有利于读取光谱曲线，因此，

在数据入库之前, 需要将高光谱影像转换成以 BIP 格式存储的 TIFF (GeoTIFF) 类型。

3.5 数据存储结构设计

数据存储结构设计就是数据表和数据表空间的设计。数据表是数据存储的基本单位, 各种数据根据自身的特点以及互相之间的联系, 按照 Oracle 规定的一定的数据类型, 设定表的字段设置、确定主键、确定连接方式^[25]。数据结构设计是整个数据库的根本所在, 系统能否高效稳定的运行在很大程度上取决于这部分实现的优劣。对于图片数据等数据量比较大的字段, 则需要分配好数据表空间, 以保证查询时不必反复的读取这一庞大的空间, 从而提高查询效率。充分考虑系统的需求, 并结合 FISS 数据的特点以及所设计的数据存储模型, 本次研究为 FISS 高光谱数据库设计了五个数据表, 分别是 FISS_GeoRaster 数据表、FISS_Measure 数据表、FISS_RGB 数据表、FISS_RASTER 数据表和 FISS_USER 表。

1、FISS_GeoRaster 表, 用来存储跟 FISS 高光谱数据库应用紧密相关的参数以及 FISS 数据逻辑存储实体即 GeoRaster 对象, 其字段名、数据类型和说明如表 3-2 所示。

表 3-2 FISS_GeoRaster 表

字段名	数据类型	说明
IMG_ID	NUMBER	影像唯一编码, 主键
OBJ_TYPE	VARCHAR2(4000)	测量目标名称
IMG_LOC	VARCHAR2(4000)	测量目标采集地点
IMG_DATE	DATE	测量日期
IMG_TIME	VARCHAR2	测量时间
IMG_HEIGHT	NUMBER	测量时仪器高度
IMG_LONG	NUMBER	目标采集时的经度
IMG_LAT	NUMBER	目标采集时的纬度
IMG_DATA	SDO_GEORASTER	栅格数据模型

使用 SQL 语言在 SQL Developer 工具创建该表的代码如下:

```
CREATE TABLE FISS_GeoRaster
(
  IMG_ID NUMBER NOT NULL,
  OBJ_TYPE VARCHAR2(4000) NOT NULL,
  IMG_LOC VARCHAR2(4000) NOT NULL,
  IMG_DATE VARCHAR2(4000) NOT NULL,
```

```

IMG_TIME VARCHAR2(4000),
IMG_HEIGHT VARCHAR2(4000),
IMG_LONG NUMBER,
IMG_LAT NUMBER,
IMG_GEOR MDSYS.SDO_GEORASTER
CONSTRAINT FMD_PK PRIMARY KEY
(
    IMG_ID
);

```

2、FISS_Measure 表，用来存储测量时 FISS 仪器的参数，包括积分时间、光圈等级、制冷温度、DN 动态范围、扫描帧速率、视场角等。同时，还要提供一个说明字段，对在测量时发生的其它情况进行说明。

表 3-3 FISS_MEASURE 表

字段名	数据类型	说明
IMG_ID	NUMBER	影像唯一编码，主键
IMG_INTEGRAL	NUMBER	积分时间
IMG_APETURE	NUMBER	光圈等级
IMG_TEMPERATURE	NUMBER	制冷温度
IMG_DNRANGE	NUMBER	DN 动态范围
IMG_SCANRATE	NUMBER	扫描帧速率
IMG_SACNFIELD	NUMBER	视场角
IMG_NOTE	VARCHAR2(4000)	采集时其它情况说明

使用 SQL 语言在 SQL Developer 工具创建该表的代码如下：

```

CREATE TABLE FISS_MEASURE
(
    IMG_ID NUMBER NOT NULL,
    IMG_INTEGRAL NUMBER,
    IMG_APETURE NUMBER,
    IMG_TEMPERATURE NUMBER,
    IMG_DNRANGE NUMBER,
    IMG_SCANRATE NUMBER,
    IMG_SACNFIELD NUMBER,
    IMG_NOTE VARCHAR2(4000),

```

```

CONSTRAINT FMD_PK PRIMARY KEY
(
    IMG_ID
);

```

3、FISS_RGB 表，用来存储 FISS 数据的彩色合成图像，当需要显示影像时，采用该表提取数据。

表 3-4 FISS_RGB 表

字段名	数据类型	说明
IMG_ID	NUMBER	影像唯一编码，主键
IMG_PIC	BLOB	RGB 合成图像

使用 SQL 语言在 SQL Developer 工具创建该表的代码如下：

```

CREATE TABLE FISS_RGB
(
    IMG_ID NUMBER NOT NULL,
    IMG_PIC BLOB NOT NULL,
    CONSTRAINT FISS_GEORASTER_PK PRIMARY KEY
    (
        IMG_ID
    )
);

```

4、FISS_RASTER 表，用来存储实际的栅格数据块，是上文所介绍的 SDO_RASTER 的实例化。该表通过 RASTERID 字段与 FISS_GeoRaster 表中的 GEORASTER 对象相关联。

表 3-5 FISS_RASTER 表

字段名	数据类型	说明
rasterId	NUMBER	影像唯一编码，主键
pyramidLevel	NUMBER	金字塔级数
bandBlockNumber	NUMBER	波段块号
rowBlockNumber	NUMBER	行块号
columnBlockNumber	NUMBER	列块号
BLOCKMBR	SDO_Geometry	定界矩形
RASTERBLOCK	BLOB	存储实际数据

使用 SQL 语言在 SQL Developer 工具创建该表的代码如下：

```

create table FISS_RASTER of mdsys.sdo_raster

```



```
(primary key(rasterId, pyramidLevel, bandBlockNumber,
rowBlockNumber, columnBlockNumber))
lob(rasterblock)store as(
chunk 32768
cache reads
nologging
pctversion 0
storage ( pctincrease 0));
```

5、FISS_USER 表，用于存放用户的属性数据,包括用户 ID、用户名称、用户密码、用户权限、用户单位、用户职位、电话和 Email 等 8 个字段，其中用户字段为关键字段。

表 3-6 FISS_USER 表

字段名	数据类型	说明
USER_ID	NUMBER	用户 ID
USER_NAME	VARCHAR2(100)	用户名称
USER_PWD	VARCHAR2(100)	用户密码
USER_AUTHORITY	NUMBER	用户权限
USER_UNIT	VARCHAR2(1000)	用户单位
USER_POSITION	VARCHAR2(100)	用户职位
USER_PHONE	NUMBER	电话
USER_EMAIL	VARCHAR2(100)	电子邮件

使用 SQL 语言在 SQL Developer 工具创建该表的代码如下：

```
CREATE TABLE FISS_USER
(
  USER_ID NUMBER NOT NULL,
  USER_NAME VARCHAR2(100),
  USER_PWD VARCHAR2(100),
  USER_AUTHORITY NUMBER,
  USER_UNIT VARCHAR2(1000),
  USER_POSITION VARCHAR2(100),
  USER_PHONE NUMBER,
  USER_EMAIL VARCHAR2(100),
  CONSTRAINT FMD_PK PRIMARY KEY
  (
    IMG_ID
```

)
);

3.6 本章小结

本章主要针对 FISS 高光谱数据库的功能需求以及业务流程进行了分析，根据系统结构设计的原则，提出了 FISS 高光谱数据库的体系结构并进行了阐述。采用模块化的思想，将系统功能划分为数据导入、数据查询检索、数据输出、数据可视化、用户管理、系统管理六大功能模块，详细论述了每个模块的功能。介绍了系统开发时所采用的 Oracle 数据库、C#语言、ADO.NET 数据接口、SQL Developer 开发工具以及用来支持影像可视化的 Arc Engine 工具。详细分析了 FISS 数据的特点，对现有 FISS 数据情况进行了分析，详细介绍了 GeoRaster 模型，提出了 FISS 数据存储模型，最后详细设计了各个数据表的结构，并给出了创建各个表的 SQL 语句。

第四章 FISS 高光谱数据库关键技术研究及其实现

对于高光谱数据库来说,如何在数据库层面上实现高光谱数据的图谱合一是需要研究的问题。为了解决这个问题,本章对 FISS 高光谱数据库的数据存取技术、影像和光谱的显示方法进行了研究并在确定的开发平台上实现了这些技术。基于这些关键技术,开发实现了 FISS 高光谱数据库系统原型。在原型系统之上,,集成了一些能体现高光谱特色的常见的光谱分析技术(包括光谱平滑、光谱重采样、光谱匹配、包络线去除等)。

4.1 FISS 高光谱数据库数据存取技术的研究与实现

在内存为 4.0G、CPU 为 Intel(R) Core(TM)2 Quad Q9550、操作系统为 Windows 7 的 PC 机上,采用 VS2008 中的 C#作为开发语言,采用 Oracle 11g 作为数据库管理系统,采用 Oracle SQL Developer 作为存储过程的开发调试工具,对高光谱影像的存取进行了研究。

在 Oracle 数据库中建立 IMG_GeoRaster 表存储栅格数据属性信息,其中 IMG_GEOR 字段存放 SDO_GeoRaster 对象,建立栅格数据表 IMG_ RASTER 存储栅格数据。

4.1.1 高光谱影像数据的存储设计

1 存储思路

高光谱影像的海量数据把如何提高数据入库效率提到了一个更为重要的高度。提高数据入库效率,要从硬件、操作系统平台、应用程序服务器、客户端和数据库等方面来考虑。研究平台确定之后,各种硬件及软件就随之确定,此时影像入库的效率就取决于应用程序和数据库的交互性能。宋国兵等研究表明,在存储大二进制对象(BLOB)数据时,使用存储过程的效率比单独使用 SQL 语句高^[33]。GeoRaste 模块提供了 SDO_GEOR.importFrom()方法实现数据的导入,该方法需要在 SQL 语言中使用。同时在向数据表里添加数据之前,需要为 GeoRaster 对象创建一个唯一的主码,如果使用单独的 SQL 语句实现这些功能,必然会降低系统的效率。为实现最佳系统性能,提高数据入库效率,设计基于存储过程的数据入库方法:创建一个存储过程,先在里面创建主码,然后再进行 GEORSATER 对象的插入,以及数据分块、建立空间索引和金字塔等操作。这样可以有效的减少应用程序与数据库的交互,提高运行效率。

创建完毕后,就可以调用存储过程来实现数据的入库。因为 Oracle 权限管理的原因,需要使用系统管理员的权限赋予 MDSYS 用户和普通用户读取文件的权限。因此,整个流程要分为两个部分,第一部分是赋予用户读取文件的权限,第二部分是调用存储过程将得到授权的文件导入数据库中,如图 4-1 所示。

2 数据存储格式设计

在 3.4.1 节已经提到，高光谱影像数据通常以三种格式排列：BIP（波段按像元交叉）、BIL（波段按行交叉）和 BSQ（按波段顺序）。因为波段很多，数据量很大，数据的存储格式对数据访问的效率有很大的影响。本研究的目标是要从 GeoRaster 对象中提取出像元的光谱曲线，需要访问图像某个像元所有波段的数值，因此，采用 BIP 格式作为高光谱影像的存储方式。在 Oracle 11g 中，GeoRaster 只支持以下有限的影像数据类型：TIFF、GeoTiff、JPEG、BMP、GIF、PNG 和 JP2^[32]。在这些影像数据类型中，只有 TIFF 和 GeoTIFF 格式可以成为高光谱影像的存储数据类型。

因此，在数据入库之前，需要将高光谱影像转换成以 BIP 格式存储的 TIFF（GeoTIFF）类型。

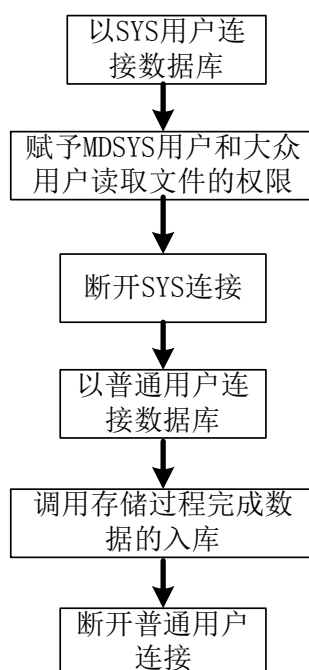


图 4-1 高光谱影像入库流程

3 SQL 存储过程设计

高光谱影像入库的存储过程首先要为每个影像建立唯一的主码，然后通过初始化、导入等步骤，将影像导入数据库中。其输入参数为高光谱影像的路径 p_imgdir 以及栅格数据表名 p_rtb。其核心实现代码如下：

```

create or replace procedure Import_Img_Data (p_imgdir in varchar2,
p_rtb in varchar2) as
    geor MDSYS.SDO_GeoRaster;
    newid number;
    oldid number;
begin

```

```

--创建主码：将 IMG_GeoRaster 表中最大的 img_id 加 1；
.....（此处往下，省略号表示省略了一部分代码）
--初始化 GeoRaster 对象
insert into Img_GeoRaster values (newid, p_Imgdir , MDSYS.SDO_
GEOR.init(p_rtb));
--取出 IMG_GEOR 字段，放入 geor 变量中
SELECT img_geor INTO geor FROM Img_GeoRaster WHERE img_id = newid FOR
UPDATE;
--调用 importFrom 方法，写入 p_Fissdir 路径对应的 tif 文件
mdsys.sdo_geor.importFrom(geor, null, 'TIFF', 'file',p_Imgdir);
--更新 IMG_GEOR 字段
UPDATE Img_GeoRaster SET img_geor=geor WHERE img_id = newid;
commit;
end;

```

4.1.2 高光谱影像光谱数据的读取设计

1 读取思路

GeoRaster 模块提供的 `sdo_geor.getCellValue()` 方法可以从 GeoRaster 对象里面获取单个像元的数值，通过不断的循环，便可获得每个像元所有波段的像素值。因为 VS2008 中的 `System. Data. OracleClient` 数据提供程序所提供的 Oracle 数据类型 (`OracleType`) 里面不包括数组 (`Array`) 类型，因此就无法从存储过程里返回光谱曲线的数组。为了获取像元的光谱曲线，同时发挥 Oracle 数据库在处理数据表上的优势，本文提出了通过动态数据表转换的光谱数据获取方法，其基本思路是：以数据表作为转换平台，在存储过程里，动态创建数据表，循环利用 `getCellValue()` 方法获取某像元各个波段的像素值，将这些值存到动态创建的数据表里，然后利用应用程序查询该数据表来获取像元的光谱数值。

2 动态转换数据表设计

根据高光谱影像的特点以及 Oracle 中数据表的创建方法，动态数据表有两种创建形式，一种是将每个波段作为数据表的一个属性（列），一种是将每个波段作为数据表的一条记录（行），如图 4-2 所示。

波段号	波段 1	波段 2	波段 3	波段 n
像元值					

(a) 一个波段作为一个属性

波段号	像元值
波段 1	

波段 2	
.....	
波段 n	

(b) 一个波段作为一条记录

图 4-2 数据表的两种创建方式

为选择一种合适的数据表创建方式，分别为这两种方式编写了对应的存储过程，并分析它们的效率。输入参数为某个像元的行号和列号，通过 C#程序调用存储过程创建数据表，并将数据存储到数据表中，再通过 C#程序查询所创建的数据表来获取数据，并将数据转换成 C#语言可以处理的数组形式。从参数输入时开始计时，到获取数组时停止计时。当波段数不同时，两种方式所需的时间如表 4-1 所示。

表 4-1 两种建表方式获取数据的时间

波段数	以波段作为属性 需要的时间 (s)	以波段作为记录 需要的时间 (s)
50	0.585	0.199
100	1.320	0.343
150	2.328	0.474
200	3.744	0.615
250	5.298	0.752
300	7.067	0.886

通过表 4-1 可以发现，以波段作为记录从数据库里获取光谱数据的速度明显比以其作为属性的速度快，而且波段数越多，相对速度越快。因此，在通过数据表转换的数据获取方法中，采用以波段作为记录的数据表创建方式。

3 SQL 存储过程设计

光谱数据读取的存储过程要能动态创建存储光谱数据的数据表，查找相应的 GeoRaster 对象，从中获取像元的像素值。其输入参数为 GeoRaster 对象的 ID 码 p_rid，像元的行列号 pixelRow 和 pixelColumn，返回的参数为 GeoRaster 对象的波段数 bandNum，供后面的 C#应用程序使用。因为 getCellValue() 函数里面的行列号和波段号是从 0 开始计数，而传递进去的行列号是从 1 开始计数的，所以相关参数都进行了减 1 的处理。其实现代码如下：

```
create or replace procedure createPixelSpecTable
(p_rid in number,pixelRow in number,pixelColumn in number,bandNum
out integer) As
-- 定义变量： number 类型的 v_ref、 pixelRow_1、 pixelCol_1、
```

bandNum_1, integer 类型的 i 以及 MDSYS.SDO_GeoRaster 类型的 geor。

```

.....
begin
--动态创建数据表
    execute immediate 'drop table pixelspectable';
    execute immediate 'create table pixelspectable (bandno number,
reflectance number)';
--查找主码为 p_rid 的 GeoRaster 对象, 并传给 geor 变量
    select IMG_GEOR into geor from IMG_GeoRaster where img_id=p_rid;
    pixelRow_1:=pixelRow-1;
    pixelCol_1:=pixelColumn-1;
--获取 GeoRaster 对象包含的波段数
    bandNum:=sdo_geor.GETBANDDIMSIZE(geor);
--循环获取对应像元各个波段的像素值, 并将其插入所创建的数据表中
    for i in 1..bandNum loop
        bandNum_1:=i-1;
        v_ref:=sdo_geor.getCellValue(geor,0,pixelRow_1,pixelCol_1,bandNum
_1);
        insert into pixelspectable (bandno,reflectance) values (i,v_ref);
    end loop;
    commit;
end;
```

4.1.3 基于存储过程的高光谱影像数据入库的实现

1 赋予 MDSYS 用户和大众用户读取文件权限的实现

①以管理员身份连接数据库并打开 OracleConnection 连接。

```

string          strConnAsSystem="Data          Source=ORCL;User
ID=system;password=admin";
    OracleConnection connAsSystem = new
OracleConnection(strConnAsSystem);
    connAsSystem.Open();
```

②赋予 MDSYS 用户读取文件的权限。

```

string          sqlPermission1 = "Call
dbms_java.grant_permission(' MDSYS','SYS: java.io.FilePermission','" +
strFissDataDir + "','read')";
```

③赋予普通用户读取文件的权限, 将②中代码的 MDSYS 改成 PUBLIC 即可。

④为上述两条语句分别创建 OracleCommand 对象。

⑤调用 ExecuteNonQuery() 方法分别执行这两条语句，即可完成赋予权限的功能。

⑥关闭数据库连接

2 调用存储过程进行数据入库的实现方法

①以用户身份连接数据库，连接对象为 conFissImport，User ID 和 password 都是用户的，并打开连接。

②创建调用存储过程的 OracleCommand 对象，并指定其命令类型为存储过程。存储过程 Import_Fiss_Data 已在上文定义。

```
OracleCommand comFissImport = new OracleCommand("Import_Fiss_Data",
conFissImport);
```

```
comFissImport.CommandType = CommandType.StoredProcedure;
```

③创建参数数组，参数赋值并指定参数传递方向，并传递参数。其中 strFissDataDir 对要入库的 IMG 数据的路径，strRasterTab 为栅格数据表，名称为 IMG_RASTER。

```
OracleParameter[] parameters = {
    New                                OracleParameter("p_Fissdir",
OracleType.VarChar, strFissDataDir.Length),
    new                                OracleParameter("p_rtb",
OracleType.VarChar, strRasterTab.Length)
};//创建参数数组
parameters[0].Value = strFissDataDir; //参数赋值
parameters[1].Value = strRasterTab; //参数赋值
parameters[0].Direction = ParameterDirection.Input; //指定参数
传递方向
parameters[1].Direction = ParameterDirection.Input; //指定参数传递
方向
```

```
foreach ( OracleParameter parameter in parameters)
{
    comFissImport.Parameters.Add(parameter); //传递参数
}
```

④ 调用 OracleCommand 对象的 ExecuteNonQuery() 执行存储过程，使用 try……catch 结构检测是否遇到异常，若遇到异常则抛出异常。

⑤最后关闭数据库连接

以上是在 VS2008 的 C#语言和 Oracle 11g 数据库平台下实现的 tiff 格式高光谱影像入库的基本步骤和方法，除此之外，在入库过程中还可以进行影像分

块、建立空间参照、建立金字塔等操作，GeoRaster 模块均提供了相应的函数。其中，影像分块方式通过设置 mdsys.sdo_geor.importFrom() 函数的 storageParam 参数的 blocking 和 blocksize 属性来实现，建立空间参照可以使用 sdo_geor.georeferenced() 函数，建立金字塔可以使用 sdo_geor.generatePyramid() 方法。由于篇幅有限，在此不一一介绍其实现方法。

4.1.4 基于动态数据表转换的光谱数据读取的实现

①以用户身份连接数据库，连接对象 conCol，具体实现方法与上文相应内容一致，打开连接。

②创建调用存储过程的 OracleCommand 对象，并指定其命令类型为存储过程。存储过程 createPixelSpecTable 已在上文定义。

③创建参数数组，参数赋值并指定参数传递方向，并传递参数。参数数组的成员为 createPixelSpecTable 存储过程的参数，需要注意的是，参数数组里的第四个参数 bandNum 的传递方向为 output。

④调用 ExecuteNonQuery() 执行存储过程，并检测是否遇到异常，若遇到异常则抛出异常。

⑤获取波段数并创建光谱数组，本次试验的数据类型为 float 型。

```
int bandNum =(int) parameters[3].Value;
float[] specData = new float[bandNum];
```

⑥使用 ExecuteReader() 从创建的光谱数据表里查询并取出数据，将数据存放到 OracleDataReader 对象中。

```
string sqlFromCol = "select * from pixelspectable";
OracleCommand comGetData = new OracleCommand(sqlFromCol,
conCol);

OracleDataReader reader = comGetData.ExecuteReader();
```

⑦使用 Read() 方法将 OracleDataReader 对象的数据读出来并将其转化成数组，该数组就是以 pixelRow 和 pixelColumn 为行列号的像元的光谱曲线数组。

```
int i = 0;
while (reader.Read() && i < bandNum)
{
    specData[i] = reader.GetFloat(1);
    i++;
}
```

⑧关闭数据库连接

以上是从存储在数据库里的 GeoRaster 对象中提取像元光谱曲线的实现过

程,可以使用得到的光谱数组,进行光谱匹配、包络线去除以及光谱尺度转换等数据分析操作。

4.2 影像及光谱显示方法研究与实现

4.2.1 影像显示方法研究与实现

FISS 高光谱数据库不仅仅是能够存储和管理 FISS 高光谱影像,同时还要能实现 FISS 数据的显示以及各像元光谱曲线的显示。目前,GeoRaster 对象只能通过以下两种方式来显示:一是调用 SDO_GEOR.exportTo 函数将 GeoRaster 对象导出并存成某一影像文件,然后使用其它影像显示工具(如 ENVI、ERDAS 等)来显示该影像;二是使用独立的 GeoRaster 显示工具,如 GeoRasterViewer 工具。这两种方式都可以显示 FISS 的彩色合成影像,但是又各有不足,第一种方式可以显示像元的光谱曲线,但是在数据导出的过程中需要花费大量的时间,在本系统的开发环境中,经研究发现,使用 SDO_GEOR.exportTo 方法导出 1GB 的数据需要 5 分钟左右。第二种方式不需要将数据导出,但是又无法支持像元光谱曲线的显示,同时也没有提供 C# 的开发接口。因此,这两种方式都不适合用来进行 FISS 高光谱数据的显示。考虑到对 FISS 数据的分析更多的是侧重于其光谱信息,对于彩色合成影像的显示不是很看重,因此本研究采用一个折中的方案来进行 FISS 数据的显示。即在存储 FISS 数据的同时,以 BLOB 类型的形式存储其 BMP 或 JPG 格式的彩色合成影像。在显示影像时,先将对应的 BLOB 数据导出成对应的 BMP 或者 JPG 格式,临时存储在某个安全的路径下面。因为 JPG 或者 BMP 格式只是包含了三个波段的数据量,与 FISS 数据 344 波段的数据量相比,其导出时间肯定大大减少,从而节约了时间。临时文件存储完毕后,使用 ARC ENGINE 的相关方法来显示该文件。在 ARC ENGINE 中,在访问栅格数据之前,必须先打开一个工作空间。工作空间必须通过工作空间工厂来创建,可以通过 RasterWorkspaceFactory 对象来创建。RasterWorkspaceFactory 对象通常用于生成一个 RasterWorkspace 对象,RasterWorkspace 对象就是一个栅格数据集容器,RasterWorkspaceFactory 类实现了 IWorkspaceFactory 和 RasterWorkspace2 接口。RasterWorkspace 对象实现了 IRasterWorkspace 和 IRasterWorkspace2 接口,这两个接口定义了用于打开和设置一个数据集对象。

为了便于对显示的影像进行放大、缩小或者漫游操作,采用了导航图和鹰眼图的功能来显示 FISS 数据的彩色合成影像。在具体实现的时候,建立了 NavImage 类来实现导航图的功能,建立了 PanImage 类来实现鹰眼图的功能。

NavImage 类中显示临时彩色合成影像的关键代码如下:

```
int Index = pFileName.LastIndexOf('\\');
string pFolder = pFileName.Substring(0, Index);
string pFileName1 = pFileName.Substring(Index + 1);
```

```
IWorkspaceFactory pWorkspaceFactory = new
RasterWorkspaceFactoryClass();
IRasterWorkspace pRasterWorkspace =
(IRasterWorkspace)pWorkspaceFactory.OpenFromFile(pFolder, 0);
IRasterDataset pRasterDataset =
(IRasterDataset)pRasterWorkspace.OpenRasterDataset(pFileName1);
IRasterLayer pRasterLayRGB = new RasterLayerClass();
pRasterLayRGB.CreateFromDataset(pRasterDataset);
IRaster ClipRaster = pRasterLayRGB.Raster;
this.axMapControll.AddLayer(pRasterLayRGB);
IRasterProps rasterProps = (IRasterProps)ClipRaster;
int dHeight = rasterProps.Height;
int dWidth = rasterProps.Width;
float ration;
if (dWidth >= dHeight)
{
    if (dWidth >= 300)
    {
        ration = 300 / ((float)dWidth);
        dWidth = 300;
    }
    else
    {
        ration = 1.0f;
    }
    dHeight = (int)(dHeight * ration);
}
else
{
    if (dHeight >= 300)
    {
        ration = 300 / ((float)dHeight);
        if (dWidth * ration < 123)
        {
            ration = 123 / ((float)dWidth);

```

```

        dWidth = 123;
        dHeight = (int)(dHeight * ration);
    }
    else
        dWidth = (int)(dWidth * ration);
    }
}

axMapControl1.Width = dWidth;
axMapControl1.Height = dHeight;

```

其中, pFileName 为存储的临时文件名称。最终影像的显示效果如图 4-3 所示。

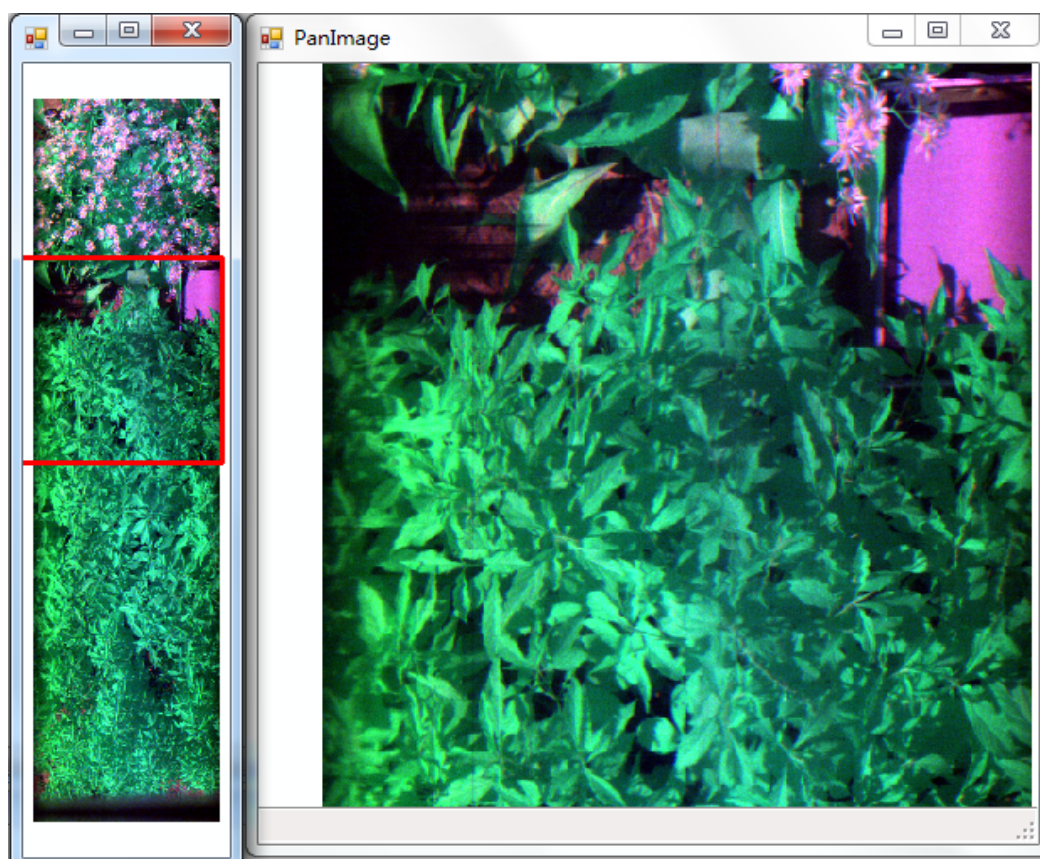


图 4-3 FISS 彩色合成影像导航图和放大图实现效果

4.2.2 光谱显示方法研究与实现

显示了 FISS 的彩色合成影像之后,为充分展现高光谱数据图谱合一的特点,特意设计了在 PanImage 窗口右键可以显示对应像元的光谱曲线。在 PanImage 类中,当右击鼠标,点击显示光谱曲线按钮时,将会返回当前像元的行列坐标,供数据获取程序使用。上文已经详细论述了光谱曲线数据的获取办法,在此不再赘述。获取光谱曲线数据之后,如何在 C#程序中显示,这是要关心的问题。在本次开发中,首先建立了一个 DRAW 类,该类包括了以下四个功能函数:

```

    public void DrawXY(Graphics g, Panel panSpec) {.....}
    public void SetYAxis(Graphics g, Panel panSpec, float
reflecMax) {.....}
    public void SetXAxis(Graphics g, Panel panSpec, float xBandMax, float
xBandMin) {.....}
    public PointF[] TransferToPoint(Panel panSpec, float[] waveLength,
float[] data, float xBandMax, float xBandMin) {.....}

```

其中, DrawXY() 函数负责绘制 X 轴和 Y 轴的基本部分, SetXAxis() 负责绘制 x 轴上的刻度, SetYAxis() 负责绘制 Y 轴上的刻度, TransferToPoint() 函数负责将输入的波长数组和反射率值数组转换成可直接用于绘制的 PointF[] 数组。

其次, 在 PanImage 类中, 使用 public void GetCoord(out int X, out int Y) 函数获取当前坐标, 使用 public float[] GetFissWavelength() 函数获取 FISS 数据的波长数组 fWavelength, 使用 public float[] GetFissRefData() 函数获取 FISS 数据的反射率信息数组 fRefData。

最后, 建立 plotFissSpec 类, 将获取的波长数据和反射率数据传给该类的绘图函数, 最终实现了从影像上获取光谱曲线的显示。其核心代码如下:

```

    float xBandMax = 950;//固定最短波长及最长波长
    float xBandMin = 400;
    Pen greenpen = new Pen(Color.Green, 0.5F);
    Draw drawxy = new Draw();
    Graphics g = panell.CreateGraphics();
    //调用DRAW类的方法, 实现绘制坐标轴以及坐标转换
    drawxy.DrawXY(g, panell);
    drawxy.SetXAxis(g, panell, xBandMax, xBandMin);
    drawxy.SetYAxis(g, panell, 100);
    PointF[] specPoints =
drawxy.TransferToPoint(panell, fWavelength, fRefData, xBandMax, xBandMin)
;

    if (specPoints == null)
    {
        Close();
        return;
    }
    else

```

```
{
    g.DrawCurve(greenpen, specPoints);
}
```

FISS 像元光谱曲线的显示效果如下图 4-4 所示。

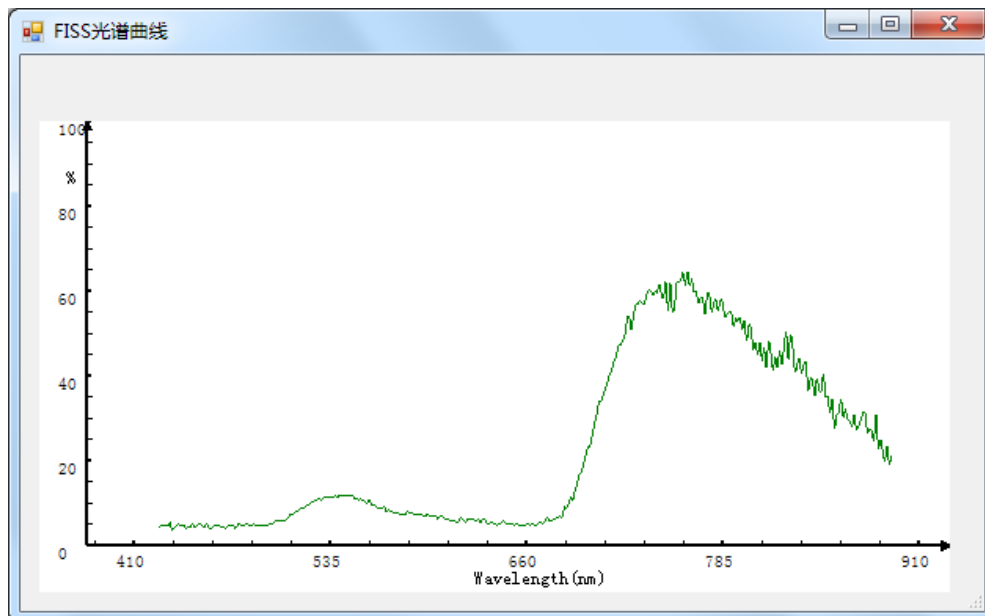


图 4-4 FISS 像元光谱曲线的显示效果图

4.3 FISS 高光谱数据库的实现

在进行了完整的分析、设计与关键技术突破的基础上，实现了 FISS 高光谱数据库的原型系统。系统提供了基本的数据库管理功能，实现了 FISS 数据的导入、查询检索以及合成影像和光谱曲线的显示。以下图片显示了原型系统的效果。

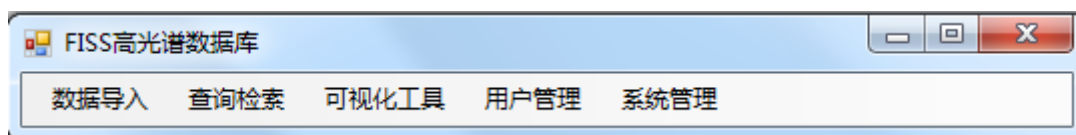


图 4-5 FISS 高光谱数据库界面

FISS数据入库

岩矿类型:

采集地点:

采集日期:

2011年12月17日

时间:

15:07:02

测量高度:

m

快视图:

...

FISS数据:

...

经 度:

纬 度:

配置参数文件:

...

测量情况说明:

确定

取消

图 4-6 参数输入界面

数据查询界面

FISS数据查询

波谱数据查询

影像数据查询

☒ 岩矿类型

花岗岩

☐ 采集地点

☐ 采集时间

2011年12月17日

至

2011年12月17日

查询

	ID	岩矿类型	测量地点	测量日期	测量时刻	测量高度	采集经度	采集纬度	显示影像	显示参数
▶	3	花岗岩	河北省承德市...	2011年12月16日	12:11:25	0.5	123.56	38.36	显示影像	显示参数
	1	花岗岩	北京市朝阳区	2011年11月8日	12:31:54	0.5	123.36	40.36	显示影像	显示参数
	2	花岗岩	北京市朝阳区	2011年11月8日	12:31:54	0.5	123.36	40.36	显示影像	显示参数
*										

图 4-7 数据查询界面

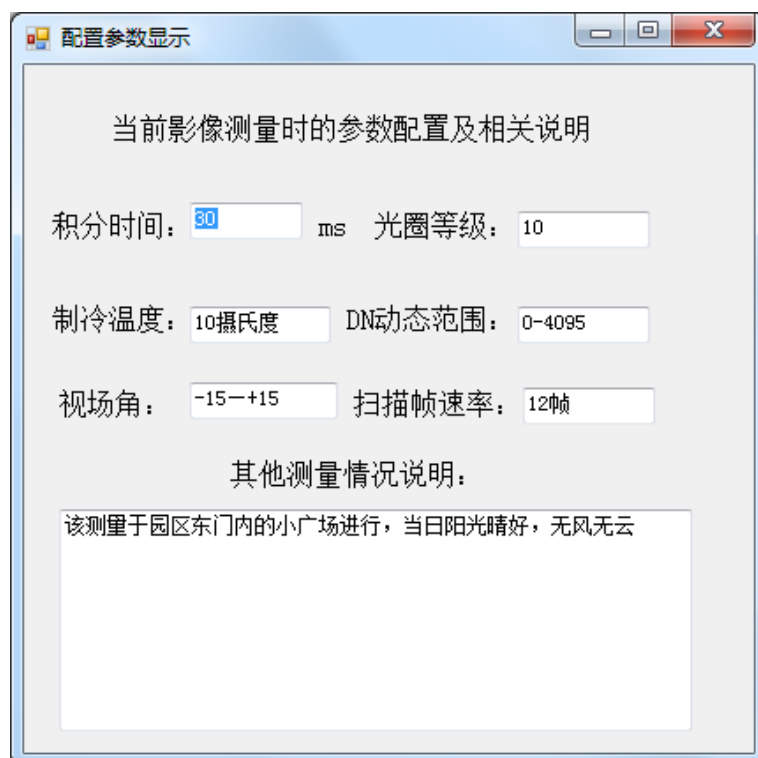


图 4-8 配置参数显示界面

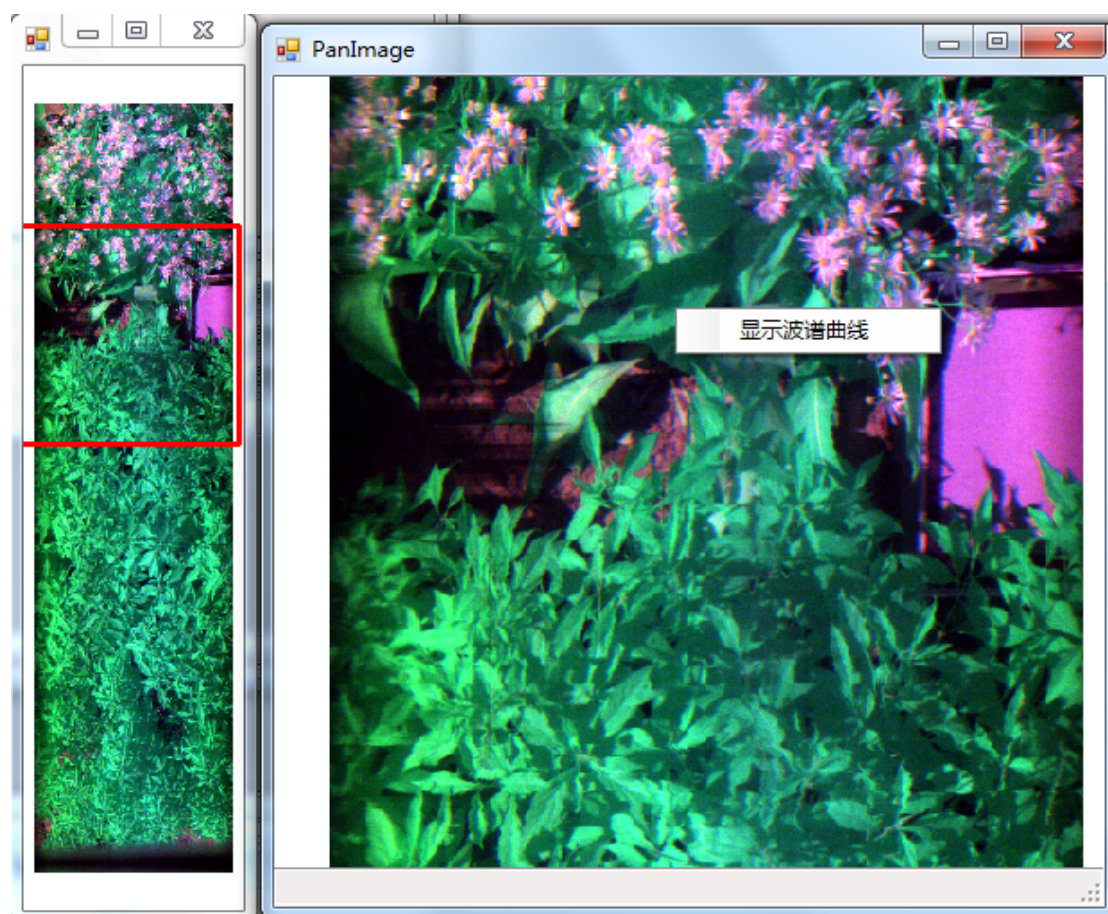


图 4-9 显示影像及显示波谱曲线按钮图

4.4 基于 FISS 高光谱数据库的光谱分析功能

根据高光谱数据库的概念,高光谱数据库在实现对高光谱数据进行管理的同时,还应具备一些基本的光谱分析功能。这对于 FISS 高光谱数据库来说尤为重要,因为 FISS 巨大的数据量,如果根据查询条件把有可能需要进一步研究的数据一幅幅导出,将会浪费大量的时间。而借助简单的光谱分析工具,将有助于确定哪些数据是需要进一步的研究,哪些不需要进行进一步的研究。基于已经开发完成的 FISS 高光谱数据库,将光谱平滑、光谱重采样、光谱匹配和包络线去除算法集成到了数据库中。以下将介绍这些算法的基本理论以及在基于数据库的实现效果。

4.4.1 光谱平滑

光谱平滑的本质是对光谱曲线进行低通滤波,去掉高频成分,保留有用低频信息、剥离背景信号、复原被扭曲的信号、提高信号质量等。基本思路是在平滑点的前后各取若干点来进行“平均”处理,以求得平滑点的最佳估计值,消除随机噪声。对光谱数据进行平滑处理的方法也叫数字滤波器^[34]。目前常常采用的平滑方法有均值滤波与中值滤波。这里主要介绍均值滤波的方式。均值滤波是一种典型的低通滤波器,传统的均值滤波是用一个有奇数点的滑动窗口在光谱曲线上滑动,将窗口中心点对应的光谱曲线波段的灰度值用窗口内的各个点的灰度值的平均值代替,如果滑动窗口规定了在取均值过程中窗口各个像素点所占的权重,这时候就称为加权均值滤波;传统均值滤波和加权均值滤波的表达式分别如下:

$$r'_k = \frac{\sum_{i=k-b}^{i=k+b} r_i}{(2b+1)} \quad (4-1)$$

$$r'_k = \frac{\sum_{i=k-b}^{i=k+b} a_i r_i}{\sum_{i=k-b}^{i=k+b} a_i} \quad (4-2)$$

式(4-1)为传统均值滤波方法,式(4-2)为加权均值滤波方法。其中, r_i 和 r'_k 分别为第k点(中心点)的平滑前、后的值,而 a_i 为平滑系数(或权重),“窗口”宽度为 $2b+1$ 个点,在窗口内进行加权平均,平均区段是逐点后移的。这种平滑处理方法可有效平滑高频噪声,提高信噪比。

在4.2.2节的FISS像元光谱曲线效果图中,我们可以发现,直接从FISS反射率文件中提取的光谱曲线,具有较强的抖动现象。为解决这一现象,采用了传统均值滤波方法,对光谱进行了平滑。经过多次试验,发现b值取5时,效果较

好。如图 4-10 所示。

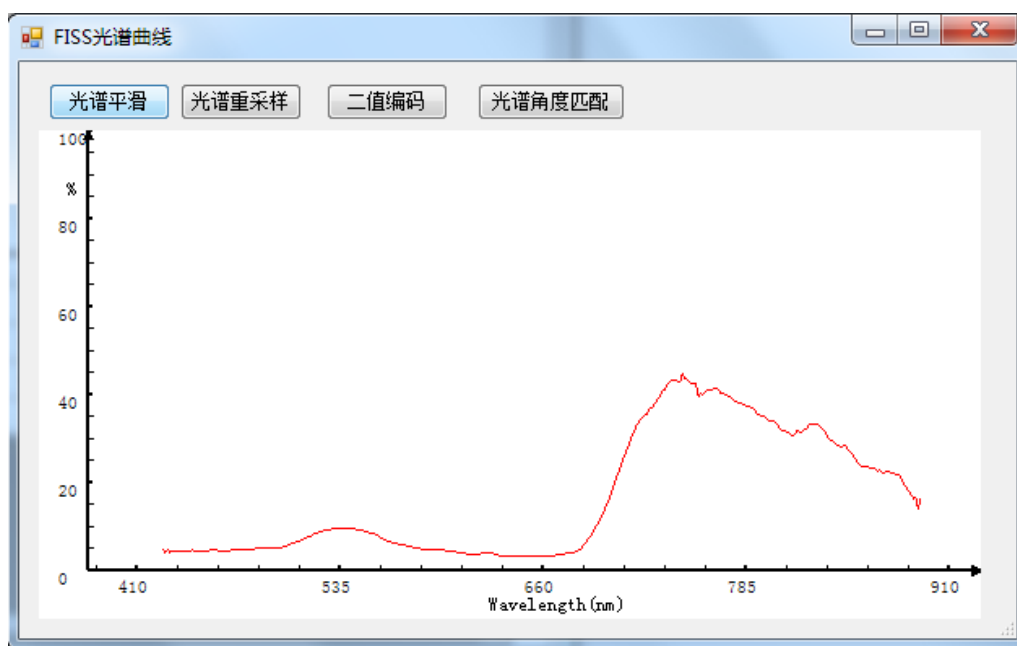


图 4-10 FISS 光谱曲线平滑效果图

4.4.2 光谱重采样

光谱重采样主要是在任意的波段间隔中，变换波段间隔重新采样，这样可以放大某部分区间，也可以模拟某个传感器的数据。

需要输入的数据有：起始波段：r1 结束波段：r2 间隔：b

计算公式为：

$$r'_i = \frac{\sum_{j=\frac{i-b}{2}}^{\frac{i+b}{2}} r_j}{n+1} \quad (4-3)$$

r'_i 为重采样后的 λ_i 反射率值， r_i 为重采样前的 λ_i 反射率值，如果 b 为偶数，n 取 b，如果 b 为奇数，则 n 取 b+1。

图 4-11 和 4-12 是重采样功能的实现过程。首先点击重采样按钮，在弹出的窗口中输入重采样的间隔，点击确定后，将在原窗口显示重采样后结果。

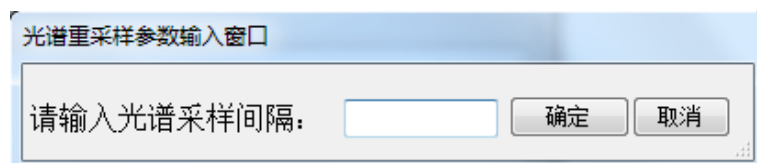


图 4-11 重采样参数输入窗口

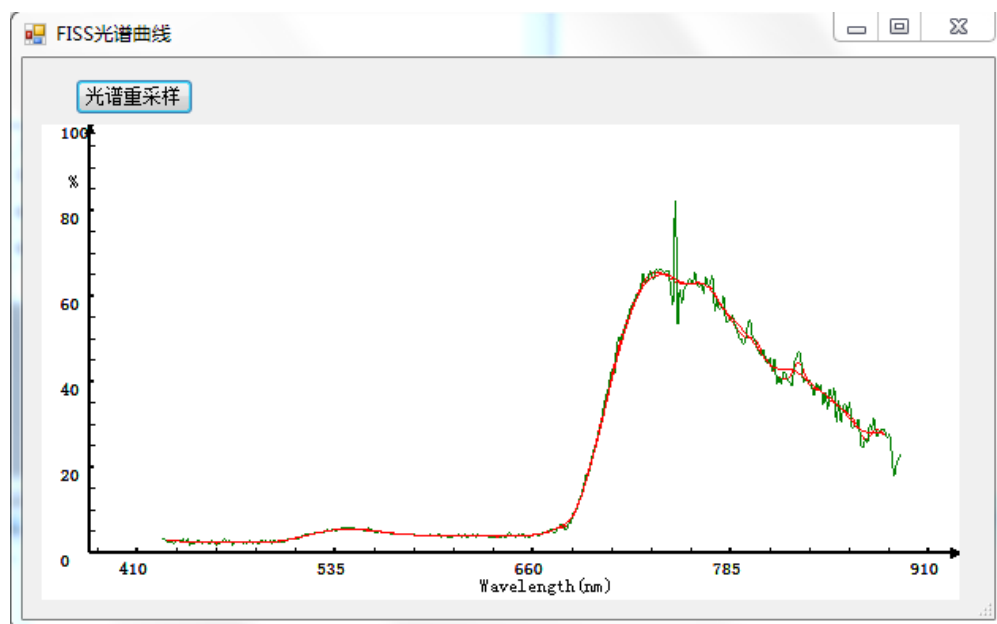


图 4-12 重采样效果图

4.4.3 光谱匹配

在成像光谱图像处理中，光谱匹配技术是成像光谱地物识别的关键技术之一。所谓光谱匹配是通过研究两个光谱曲线的相似度来判断地物的归属类别。它由已知地物类型的反射光谱，通过波形或特征匹配比较来达到识别地物类型的目的^[3]。在长期的研究中，人们逐渐认识了电磁波与地物的相互作用机理，收集了大量的实验室标准数据，建立了许多地物标准光谱数据库。在这些研究工作的基础上，已经具备了从图像直接识别对象的条件。从匹配源和匹配目标的不同出发，光谱匹配可以分为以下三种模式：

(1) 从图像的反射光谱出发，将像元光谱数据与光谱数据库中的标准光谱响应曲线进行比较搜索，并将像元归于与其最相似的标准光谱响应所对应的类别，这是一个查找过程。

(2) 利用光谱数据库中，将具有某种特征的地物标准光谱响应曲线当作模板与遥感图像像元光谱进行比较，找出最相似的像元并赋予该类标记，这是一个匹配过程。

(3) 根据像元之间的光谱响应曲线本身的相似度，将最相似的像元归并为一类，这是一种聚类过程。

在前两种运作模式中，解决问题的关键一是图像辐射亮度值到地物表面反射率的精确反演，使用经验线性法已经可以较好的对此加以解决；二是光谱匹配算法的研究。目前，常用的光谱匹配算法有二值编码匹配和光谱角度匹配，其具体算法如下所述。

1、二值编码匹配

所谓二值编码匹配，由 A. F. H. Goetz 在 1990 年针对成像光谱数据存在很

大程度的光谱数据冗余度和如何使光谱库的查找和匹配过程更加快速有效这两个问题而提出，主要考虑是在实施匹配的过程中，可能并不需要全部光谱数据的原始形式，从而提出了一系列对光谱进行二进制编码的建议^[35]，使得光谱可用简单的 0, 1 来表述。最简单的编码方法是：

$$\left. \begin{aligned} h(n) &= 0, x(n) < T \\ h(n) &= 1, x(n) \geq T \end{aligned} \right\} \quad (4-4)$$

其中 $x(n)$ 是像元第 n 通道的亮度值； $h(n)$ 是其编码； T 是选定的门限值，一般选为光谱的平均亮度，这样每个像元灰度值变为 1bit。但是有时这种编码不能提供合理的光谱可分性，也不能保证测量光谱与数据库里的光谱库相匹配，所以需要更复杂的编码方式。

(1) 分段编码

对编码方式的一个简单变形是将光谱通道分成几段进行二值编码，这种方法要求每段的边界在所有像元矢量都相同，对每一段来说，编码方式同上所示。为使编码更加有效，段的选择可以根据光谱特征进行，例如在找到所有的吸收区域以后，边界可以根据吸收区域来选择。

(2) 多门限编码

采用多个门限进行编码可以加强编码光谱的描述性能。将每个灰度值变为 2bit，或者将光谱范围划分为几个小的子区域，每个子区域独立编码。例如采用两个门限 T_a , T_b 可以将灰度划分为三个域：

$$\left. \begin{aligned} h(n) &= 00, x(n) < T_a \\ h(n) &= 01, T_a \leq x(n) < T_b \\ h(n) &= 11, x(n) \leq T_b \end{aligned} \right\} \quad (4-5)$$

其中， $n=1, 2, \dots, N$ ， N 为通道数。这样像元每个通道值编码为 2 位二进制数，像元的编码长度为通道数的两倍。事实上，两位码可以表达 4 个灰度范围，所以采用三个门限进行编码更加有效。

(3) 仅在一定波段进行编码

这个方法仅在最能区分不同地物覆盖类型的光谱区编码。如果不同波段的光谱行为是由不同的物理特征所主宰，那么我们可以仅选择这些波段进行编码，这样既能达到良好的分类目的，又能提高编码和匹配识别效率。

一旦完成编码，则可利用基于最小汉明距离的算法来进行匹配识别。人们根据匹配系数的大小来确定和提取位置图像上感兴趣的地物信息。这种编码匹配技术有助于提高成像光谱数据分析处理的效率。由于这种技术在处理编码过程中会丢失许多细部光谱信息，因此这种二值编码匹配技术适合较粗略地识别地物光谱。

2、光谱角度匹配

光谱角度匹配 (Spectral Angle Match: SAM) 通过计算一个测量光谱 (像元光谱) 于一个参考光谱之间的“角度”来确定他们两者之间的相似性。参考光谱可以是实验室光谱或野外测定光谱或是从图像上提取的像元光谱。这种方法假设图像数据已被缩减到“视反射率”，即所有暗辐射和路径辐射偏差已经去除。它被用于处理一个光谱维数等于波段数的光谱空间中的一个向量^[36, 37]。下面通过两波段 (二维) 的一个简单例子来说明参考光谱和测试光谱的关系，如图 4-13。

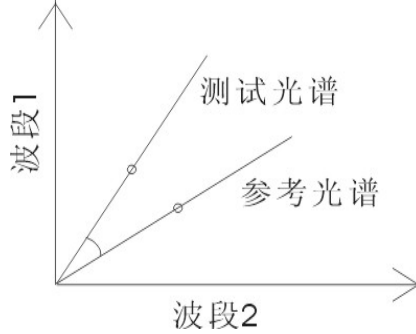


图 4-13 参考光谱和测试光谱在二维空间里光谱角度关系

它们的位置可考虑是二维空间中的两个光谱点。各个光谱点连到原点可以代替所有不同照度的物质。照度低的像元比起具有相同光谱特征但照度高的像元往往集中在原点附近 (暗点)。SAM 通过下式确定测试光谱 t_i 与一个参考光谱 r_i 的相似性：

$$\alpha = \cos^{-1} \left[\frac{\sum_{i=1}^{n_b} t_i r_i}{\left(\sum_{i=1}^{n_b} t_i^2 \right)^{1/2} \left(\sum_{i=1}^{n_b} r_i^2 \right)^{1/2}} \right] \quad (4-6)$$

式中 n_b 等于波段数。这种两个光谱之间相似性度量并不受增益因素影响，因为两个向量之间的角度不受向量本身长度的影响。这一点在光谱分类上可以减弱地形对照度的影响 (它的影响反映在同一方向直线的不同位置上)。若只是研究两条光谱之间的匹配程度，SAM 流程分为以下三步：

- (1) 从光谱数据库中选择感兴趣的“最终成分光谱”；
- (2) 对“最终成分光谱”做重采样，因为图像光谱分辨率通常要低于地面测量的光谱分辨率，使两者光谱分辨率一致。
- (3) 计算最终成分光谱与图像像元光谱两个光谱向量之间的角度 α (广义夹角余弦)，以评价此两光谱向量相似性。 α 值域为 $0 \sim \pi/2$ ，当 $\alpha=0$ 时表示两个光谱完全相似，而 $\alpha=\pi/2$ 时则两个光谱完全不同。

有时，我们也想知道一幅图像上的所有光谱与参考光谱的匹配情况，所以也可以直接用参考光谱来与遥感图像上逐个点的反射率光谱比较而达到光谱识别的目的。从而有 SAM 流程的第 4 步：

- (4) 计算成像光谱图上每个像元光谱与每个最终成分光谱的 α_i ，从而实现

对图像光谱的匹配和分类。具体匹配分类时，对于一个像元光谱 x ，计算它与第 i 个最终成分光谱的广义夹角 α_i ($i=1,2,\dots,c$; c 为光谱库中第五类别数)。假如， $\alpha_i = \{\alpha_j\}_{\min}$, ($j=1, 2, \dots, c$; $j \neq i$)，则 x 被判为第 i 最终成分光谱。如果只是为了突出感兴趣的一类最终成分光谱，那么 SAM 的输出是一幅灰度图，其上低值代表相似性高，即和目标光谱有较高的吻合性。

在 FISS 高光谱数据库中，我们采用公式 4-4 所示的二值编码方式，以光谱曲线的平均值为阈值，对经过平滑的 FISS 曲线进行的二值编码，其效果如图 4-14 所示。

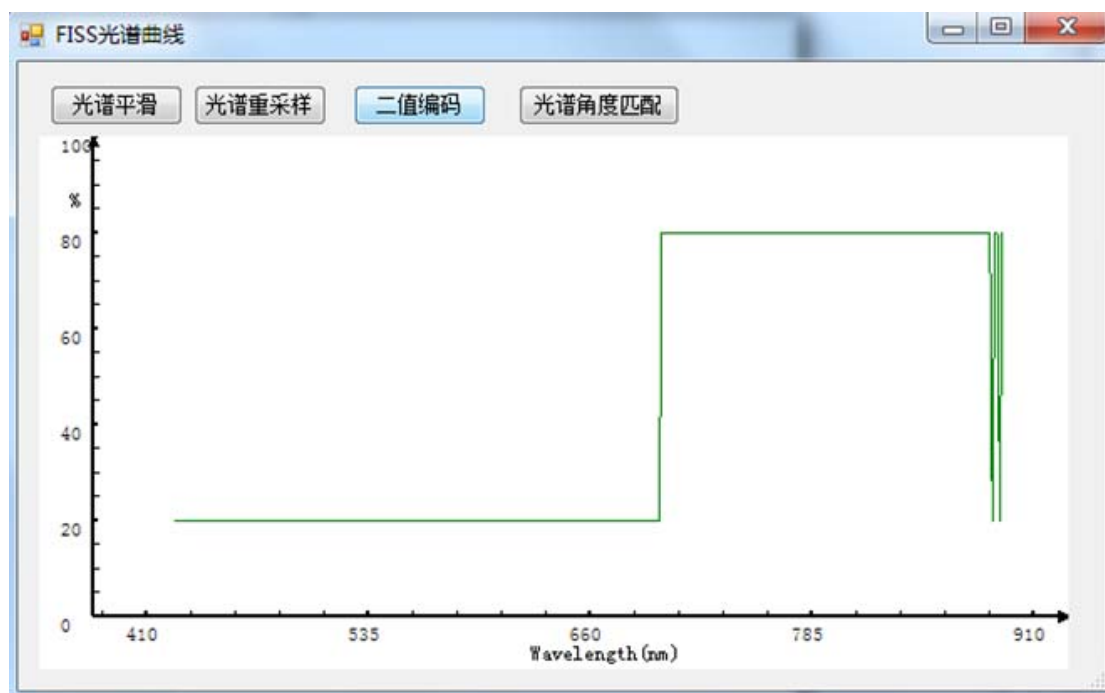


图 4-14 二值编码效果图

同时，按照公式 4-6，基于数据库实现了光谱角度匹配的算法，其实现步骤如下：

- 1、输入符合系统格式数据的参考光谱；
- 2、运行光谱匹配算法，该算法已集成在数据库中。
- 3、弹出显示匹配运行的结果。

实现效果如图 4-15 和图 4-16 所示。

4.4.4 包络线去除

包络线消除法是一种常用的光谱分析方法，它可以有效的突出光谱曲线的吸收和反射特征，并且将其归一到一个一致的光谱背景上，有利于和其它光谱曲线进行特征数值的比较，从而提取出特征波段以供分类识别。

光谱曲线的包络线从直观上来看，相当于光谱曲线的“外壳”，因为实际的光谱曲线由离散的样点组成，所以我们用连续的折线段来近似光谱曲线的包络

线。包络线去除 (Continuum Remove) 方法是利用包络线来对图像的光谱维进行归一化。

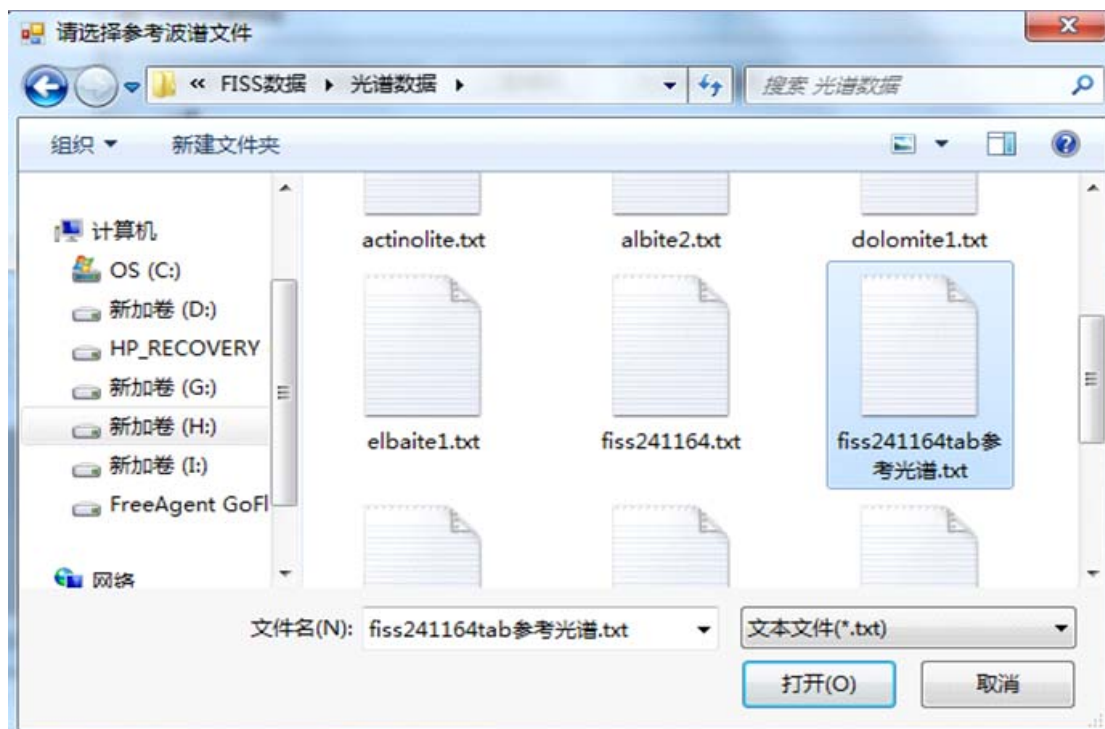


图 4-15 选择参考波谱文件

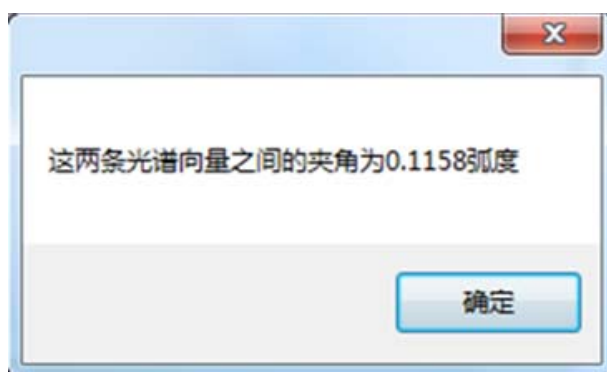


图 4-16 光谱匹配结果

求光谱曲线包络线的算法描述如下：

设有反射率曲线样点数组： $r(i), i=0, 1, \dots, k-1$ ；波长数组： $w(i), i=0, 1, \dots, k-1$ ；

(1) $i:=0$ ，将 $r(i), w(i)$ ，加入到包络线节点表中；

(2) 求新的包络节点。如 $i=k-1$ 则结束，否则 $j:=i+1$ ；

(3) 连接 i, j ；检查 (i, j) 直线与反射率曲线的交点，如果 $j=k-1$ ，则结束，将 $w(j), r(j)$ 加入到包络线节点表中，否则：

①. $m:=j+1$ ；

②. 若 $m=k-1$ 则完成检查， j 是包络线上的点，将 $w(j), r(j)$ 加入到包络线节点表中， $i=j$ ，转到(2)；

③. 否则, 求 i, j 与 $w(m)$ 的交点 $r_1(m)$ 。

④. 如果 $r(m) < r_1(m)$, 则 j 不是包络线上的点, $j := i+1$, 转到(3); 如果 $r(m) \leq r_1(m)$, 则 i, j 与光谱曲线最多有一交点, $m := m+1$, 转到②。

(4) 得到包络线节点表后, 将相邻的节点用直线段依次相连, 求出 $w(i), i = 0, 1, \dots, k-1$ 所对应的折线段上的点的函数值 $h(i), i = 0, 1, \dots, k-1$; 从而得到该光谱曲线的包络线。显然有: $h(i) \geq r(i)$;

(5) 求出包络线后对光谱曲线进行包络线消除: $r'(i) = r(i)/h(i), i = 0, 1, \dots, k-1$ 。

按照以上算法, 在 FISS 数据库原型系统上实现了包络线去除算法, 其效果如图 4-17 所示。

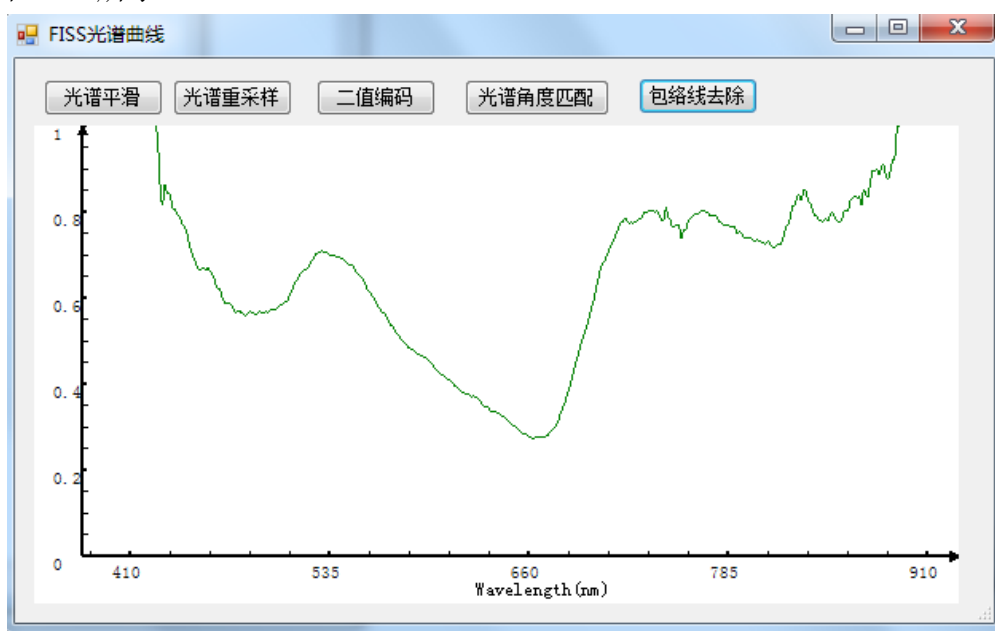


图 4-17 包络线去除效果图

4.5 本章小结

本章主要研究了高光谱数据库的数据存取技术、影像及光谱的显示方法以及常用的光谱分析方法。提出了基于存储过程的数据入库方法和基于动态数据表转换的数据获取方法, 通过使用临时文件的方式, 实现了 FISS 数据合成影像的显示, 同时研究了基于数据库的像元光谱曲线显示方法, 通过这些方法, 实现了高光谱数据在数据库层面上图谱合一的存储、获取和显示。在选择的开发环境里, 实现了 FISS 高光谱数据库原型系统。将常用的高光谱数据分析方法集成到了数据库中, 从而可以在数据库系统上进行初步的光谱分析功能。

第五章 岩矿多维高光谱数据库初步研究

随着遥感器的日益发展,遥感所获得的数据量也与日俱增。数据库技术在管理海量遥感数据方面发挥了重要的作用,也取得了丰硕的成果。高光谱数据库的提出在数据库层面上有效的实现了高光谱数据的图谱合一特点。

多维数据库是数据库发展的最新阶段,它把数据组织成更利于分析的多维数据立方体,是面向决策支持的海量数据处理的关键因素^[14]。其独特的多维数据模型,使得多维数据库需要更少的存储空间且具有更高效的检索速度^[38]。多维分析把一个实体的多项重要属性定义为多个维度,使得用户能方便的汇总数据集,简化了数据分析处理逻辑,并能对不同维度值的数据进行比较^[39],从而发现数据的分布规律。自多维数据库提出以来,一个高度抽象的、独立于任何执行的普遍适用的概念模型并没有被设计出来^[40],但是它仍然引起国内外学者极大的研究兴趣。在理论方面,Pedersen 等详细描述了多维数据库有别于关系数据库的概念和查询分析方法^[14],Papadias 等针对空间数据立方体,将数据立方体的分析操作分为概括分析、局部分析、全局分析和旋转分析^[41-43]。在应用方面,不同学者对海洋数据^[44]、铁路货运数据^[45]、医疗信息数据^[46]、图书馆数据^[47]与油气勘探数据^[48]进行多维数据建模以及实现开展了研究。多维数据库也被应用到电子文献组织^[49]、集装箱空箱调运优化^[50]等研究领域。

在遥感数据库方面,除中国典型地物波谱库运用星型模型进行地面观测波谱数据的存储架构设计外^[51],运用多维数据库和多维分析的理念来研究具有图谱合一特点的高光谱数据库的设计和构建工作开展得相对较少。高光谱遥感获得的影像立方体数据与多维数据库的核心——数据立方体具有相类似的结构,可以为多维空间分析带来方便,提高空间分析的速度。

本章以 FISS 测量的岩矿高光谱数据为数据源和研究对象,立足于建立岩矿多维高光谱数据库这一目标,定义和描述了岩矿多维数据库的基本概念,对多维高光谱数据库及其多维分析工具集进行了分析和设计。力求通过这些设计,最终能通过岩矿多维高光谱数据库的多维分析,可以满足人们对不同空间分辨率、不同光谱分辨率光谱信息的需要这一目标。

5.1 岩矿多维高光谱数据库基本概念

本文将多维数据库技术和多维分析的理念引入遥感数据库建设这个领域。因此,有必要将多维数据库的基本概念、高光谱遥感的相关概念与构建岩矿多维高光谱数据库这个目标结合起来,定义和描述岩矿多维高光谱数据库的基本概念。

5.2.1 维

维是人们观察岩矿多维数据的角度。岩矿多维数据库的一个主要目标就是利用维为多维分析提供尽可能多的途径,每个维代表一个统一访问多维数据库的信

息路径。根据时间、空间等各种因素对地物光谱的影响以及对岩矿多维数据研究的需要,确定岩矿多维数据的维度包括岩矿类型维、时间维、采集地点维、行维、列维、光谱维六个维度。

根据各个维度不同的特点,可以把维度分为非影像维和影像维。

1、非影像维

非影像维是一种仅包含非影像数据的维。在岩矿多立方体数据中,将与影像关联的各种属性信息作为非影像维,如岩矿类型、时间、采集地点等,可以统一称为非影像维。由此可见,岩矿多立方体数据中非影像维是多个的(非影像维 1, 非影像维 2, ..., 非影像维 n),它们共同构成了岩矿多立方体的非影像维世界。

2、影像维

影像维是一种直接跟影像数据相关联的维,如与影像的行、列、光谱相对应的行维、列维和光谱维可以统称为影像维。可以发现,对于高光谱数据来说,影像维有且只有三个,即行维、列维和光谱维。

非影像维和影像维在岩矿多立方数据中所起的作用不同,主要体现在,通过非影像维的各种组合,就能确定某个具体的影像立方体,通过影像维的各种组合,能够确定该影像立方体的各种反射率值。

5.2.2 岩矿影像数据立方体

简称影像立方体,指岩矿高光谱影像数据,由地面成像光谱辐射测量系统(Field Imaging Spectrometer System, FISS)测量得到。FISS 高光谱数据具有很高的分辨率(空间分辨率可达厘米甚至毫米级),是对目标进行结构光谱分析、混合光谱分解、纯像元提取、空间尺度转换和光谱尺度转换研究的有利数据^[52]。FISS 数据是图像数据和光谱数据结合的统一体,具有行、列和光谱三个维度,如图 5-1 所示。它本身具有统一完整性,与其它影像立方体互相独立。在存储时,影像立方体作为一个整体存储,因此,在没有确定某一个具体的影像立方体前,无法对影像立方体本身进行上钻、下钻、切片、切块和旋转等操作;只有确定某个影像立方体后,才能在影像立方体的行维、列维或光谱维上进行钻取、切片、切块或旋转操作。

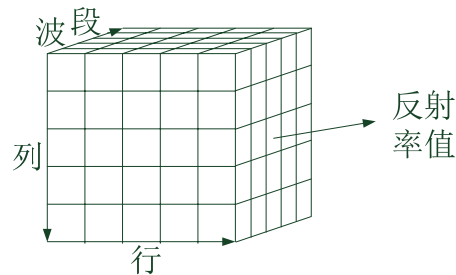


图 5-1 高光谱数据立方体

5.2.3 岩矿多立方体数据

简称岩矿多立方体,是指由岩矿多维数据六个维度和其它属性数据组成的数据集,采用层层嵌套的结构。由于行维、列维和光谱维组成了相对独立的影像立方体,所以,岩矿多立方体在表现上有三个维度,即岩矿类型维、时间维和采集地点维,但在本质上具有六个维度,即岩矿类型维、时间维、采集地点维、行维、列维和光谱维。岩矿多立方结构将在下文进行详细介绍。

5.2.4 维层次

维层次是人们观察岩矿数据时的细节程度。根据维属性的取值,可对维成员进行分组形成多个层次 $h(i)=\{h_1, h_2, \dots, h_n\}$, $n>0$, 各层次的值域是维的值域的子集。根据相关研究成果和本次研究需要,将岩矿类型维分为五个层次,采集地点维分为五个层次,时间维分为三个层次,具体分层将在下文进行介绍。对于行维、列维和光谱维,没有确定的维层次,所分的层次应根据要对影像立方体所进行的多维分析操作来确定。

5.2.5 维成员

维的 1 个取值称为该维的 1 个成员,若维是多层次的,则维成员是在不同维层次中取值的组合。如采集地点维有五个层次,在每个层次上取一个值就组成了采集地点维的一个维成员“某国某省某市某县某乡”。

5.2.6 度量

度量是由全部维或若干维确定的某一个对象的属性值,度量的测量值均发生在维的交叉点上,度量往往不局限于一个,(度量 1, 度量 2, ..., 度量 n) 共同构成了度量世界。一般数据立方体只有数字度量。空间数据立方体具有两种度量,即数字度量和空间度量^[53]。对于高光谱遥感应用来说,我们更为关注的是某一地物的光谱曲线,这就使得了岩矿多立方体数据不同于一般的数据立方体和空间数据立方体,需要不同的度量。综合其它数据立方体和高光谱遥感应用的特点,岩矿多立方体有三种类型的度量,分别是数字度量、曲线度量、影像度量。

1、数字度量

在岩矿多立方体数据中,全部影像维或部分影像维确定的某一个对象往往是一组反射率值的集合,描述该集合就得用数据度量来表述。数字度量是一种包含数字值的度量,可以分为分布式的、代数的、整体的数字度量^[41]。一个数字度量是分布式的当且仅当该数字度量值能够通过分布式聚集函数作用于该数据立方体所属子立方体的度量聚集计算得到,如 `count()`, `sum()`; 一个数字度量是代数的当且仅当该数字度量值能够通过代数聚集函数作用于该数据立方体所属子立方体的度量聚集计算得到,如 `avg()` 可以由 `sum()/count()` 计算; 否则,该数字度量就是整体的^[54]。

2、曲线度量

曲线度量是由影像立方体的行维和列维共同确定得到的某一个点所有波段的反射率值的组合，是一组数字度量的集合。但是该组合又不是数字度量的简单集合，它是按照波长从短到长顺序存放的一系列反射率值的集合。如果暂时不考虑“同物异谱”或者“异物同谱”的影像，该度量可以唯一的代表某种地物。曲线度量是具有高光谱特色的度量方式。

3、影像度量

在岩矿多立方体数据中，由全部或部分非影像维确定的某一个对象往往就是一个或数个影像对象，描述这一影像对象的聚集结果就得使用影像度量来表示。例如，研究人员要关心某年某月某日在某个采集地点有多少景某个特定地物的影像时，就要将这些结果用影像度量来表示。

5.2 岩矿多维高光谱数据库设计

上述定义的岩矿多维高光谱数据库的基本概念为构建岩矿多维高光谱数据库提供了具有针对性、专门性的理论基础。根据数据库构建的过程，岩矿多维高光谱数据库的构建需要经过功能设计、系统结构设计、概念模型设计、逻辑模型设计和物理模型设计等几个阶段。本节将围绕这几个方面逐一进行研究。

5.2.1 岩矿多维高光谱数据库功能设计

岩矿多维高光谱数据库要能够根据一个统一的多维数据模型，将大量的、现实的岩矿数据按照相应的主题转换成统一的格式，实现以多维方式组织和显示数据，从而实现复杂岩矿数据的集成、存储和管理，供人们从不同角度、多方位的认识岩矿数据。其最重要的功能应该是在有效组织数据的基础上进行多维分析。在具体分析时，能够将岩矿数据按照维层次关系进行聚集，了解岩矿在各个维度上的分布情况：在非影像维上，通过对时间维层次或采集地点维层次的聚集，可以知道岩矿类型在不同时间、不同地点的分布情况等；在影像维上，通过行维、列维和光谱维的层次聚集操作，得到相应的光谱曲线或反射率值。能够支持进行空间尺度和光谱尺度的转换，得到不同空间分辨率和光谱分辨率的数据等。

5.2.2 岩矿多维高光谱数据库系统结构设计

基于关系数据库的数据存储方式是当前多维数据库数据存储采用较多的形式，所以岩矿多维高光谱数据库也需要关系型数据库管理系统（RDBMS）来存储数据，并做为主要数据源之一。所以，结合高光谱应用的特点，确定多维高光谱数据库系统主要由四个部分组成：数据源、构建模型及工具、多维高光谱数据库、分析工具，如图 5-2 所示。

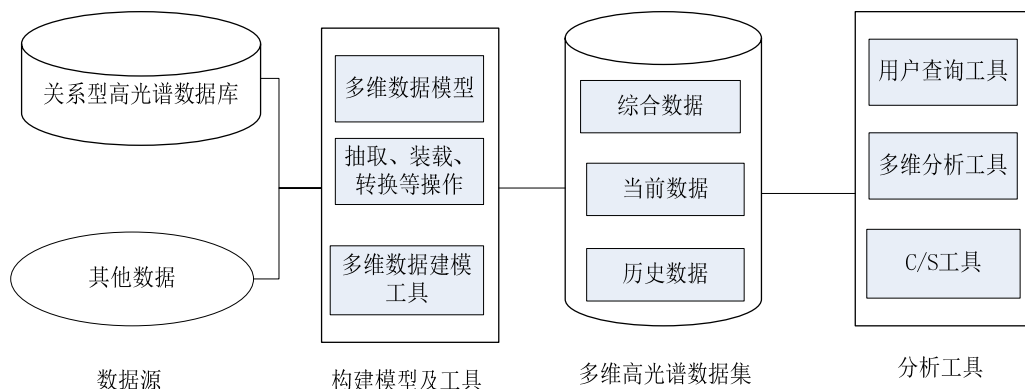


图 5-2 多维高光谱数据库系统结构图

多维高光谱数据库的源数据主要是存储在关系数据库里面的各类高光谱数据，例如地面测量高光谱数据及其元数据、各种高光谱影像及其元数据以及其它相关数据。

确定了信息需求以后，下一步要进行数据建模，通过对数据的抽取、清理和转换完成从源数据到多维数据的转换，再进行维度划分和确定多维数据库的物理存储结构。多维高光谱数据库的核心是多维数据建模工具，主要用来存储数据模型，定义数据结构、数据转换、多维数据集构建和控制信息等。

数据经过多维建模、创建数据集，并完成数据加载之后，形成了多维数据集市。然后使用多维分析工具，便可进行多维分析。

5.2.3 岩矿多维高光谱数据库概念模型设计

概念模型设计主要进行岩矿多维数据结构的设计、事实表的构建以及维层次关系的划分。

1、岩矿多维数据结构设计

多维数据结构是指多维数据在多维空间里的分布，主要有超立方和多立方两种结构^[55]。超立方结构是指用 3 维或更多的维数来描述一个对象，每个维彼此垂直。数据的测量值发生在维的交叉点上，数据空间的各个部分都有相同的维属性。多立方结构是将大的数据结构分成多个多维结构。这些多维结构是大数据维数的子集，面向某一特定应用对维进行分割，它具有很强的灵活性，提高了数据（特别是稀疏数据）的分析效率。

通过上文的分析可以看出，影像立方体是超立方结构。为简化数据结构，提高存储性能，将岩矿多维数据设计成嵌套的结构关系，在岩矿多维数据的多立方结构里面嵌套着一个个超立方结构的影像立方体，如图 5-3 所示。

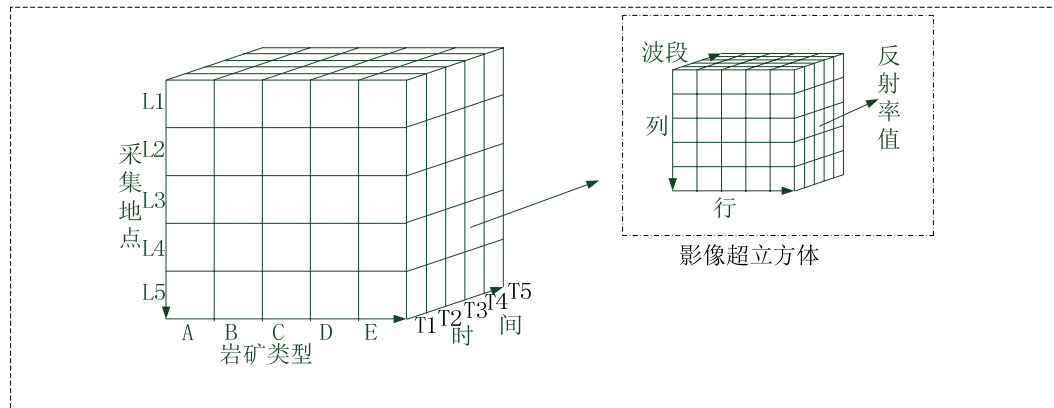


图 5-3 嵌套的岩矿多立方数据结构示意图

在多立方结构中，时间维、采集地点维、岩矿类型维组织的立方体下嵌套着行维、列维、光谱维组织的影像立方体。其中，反射率值在抽象结构上为最底层的行维、列维、光谱维立方体中的数据。

2、事实表的构建

在岩矿多维高光谱数据库里，岩矿影像数据作为一个整体存储。人们首先关心的是影像立方体在岩矿类型、时间和采集地点的任意一种组合下的分布情况。其次人们才能从影像立方体里通过钻取、切片、切块等操作得到相应的反射率值信息。因此，每个影像立方体和它的属性数据是一个事实，它存在于岩矿多维数据库中；每个反射率值和它对应的行、列和光谱信息也是一个事实，它存在于影像立方体中。对于整个数据库来说，反射率事实表以影像立方体的形式存在于影像事实表中。如图5-4所示，其中，(a) 是影像事实表，(b) 是反射率事实表，岩矿类型、时间、采集地点做为影像事实表的维度，影像的行、列、光谱作为反射率事实表的维度，采集地点具体的经纬度信息作为属性数据存储于影像事实表中。

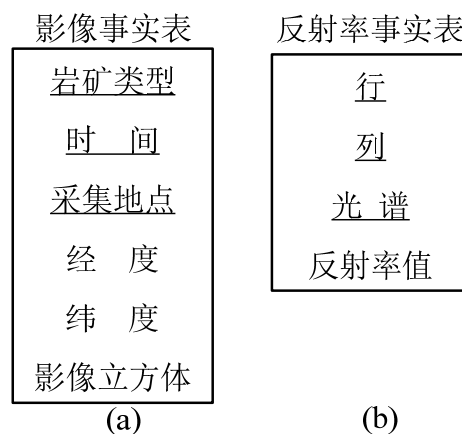


图 5-4 岩矿影像多维数据库事实表

3、维层次设计

维层次是多维数据建模中一个非常重要的概念。若没有维的层次关系，就无

法进行维层次上钻、下翻的分析操作。根据对岩矿高光谱数据分析的需要,对岩矿类型维、采集地点维和时间维的维层次划分如图5-5所示。其中,岩石和矿物的各级分类参照《中国典型地物波谱知识库》第五章第一节“典型岩矿的分类体系与特性描述”进行,在此不再赘述。

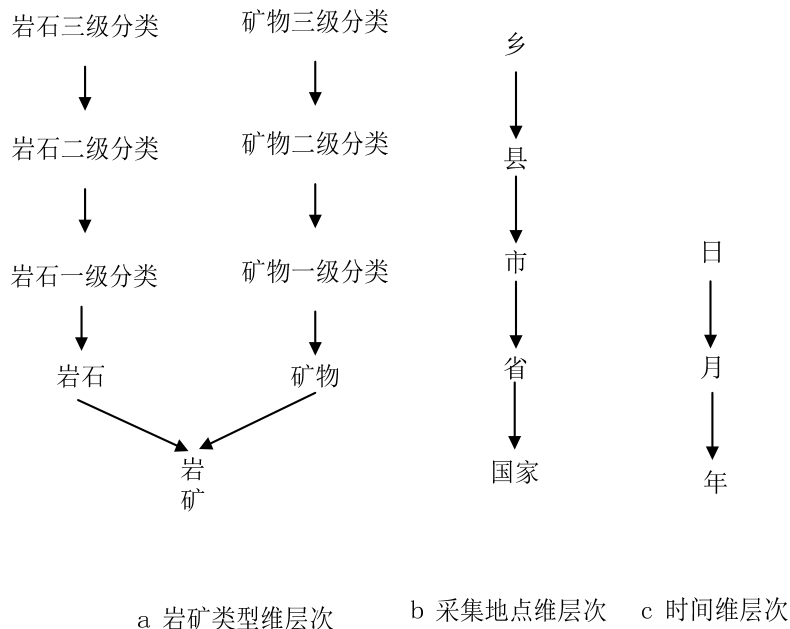


图 5-5 维层次划分

对于影像立方体的行维、列维和光谱维,不设定具体的维层次,在确定影像立方体之后,根据需要对行、列和光谱进行操作。

5.2.4 岩矿多维高光谱数据库逻辑模型设计

在关系型数据库中,对于逻辑层一般采用关系表和视图,而多维数据库采用的是星型模型和雪花模型。

(1) 星型模型

大多数多维数据库都采用星型模型来表示多维数据模型,数据库中包括一个“事实表”,对于每一维都有一个“维表”。事实表中的每条元组都包含有指向各个维表的外键和一些相应的测量数据,维表中记录的是有关这一维的属性。事实表中的每一个元组只是包含了一些指针(外建),而对应的主键分别放在不同的维表中。每个指针指向一个维表,这就构成了数据库的多维结构。在每个维表中除包含每一维的主键外,还有说明该维的一些其它属性字段。

星型模型是大多数多维数据库的集成方案所采用的方式,但这种数据组织方式存在数据冗余、多维操作速度慢等特点。

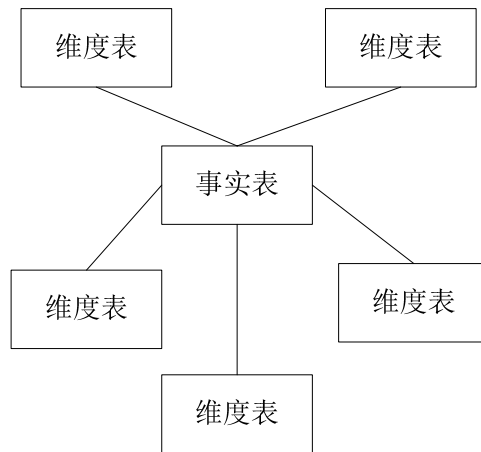


图 5-6 星型模型结构示意图

(2) 雪花模型

雪花模型是对星型模型的扩展。如果某个维表有多个归类层次，就可以形成雪花模型。它是对星型模型维表的进一步层次化，形成了一些局部的“层次”区域。雪花模型增加了用户必须处理的表数量，增加了某些查询的复杂性，但这种方式可以使系统进一步专业化和实用化，同时降低了系统的通用程度。雪花模型的优点是可以通过最大限度地减少数据存储量以及联合较小的维表来改善查询性能。

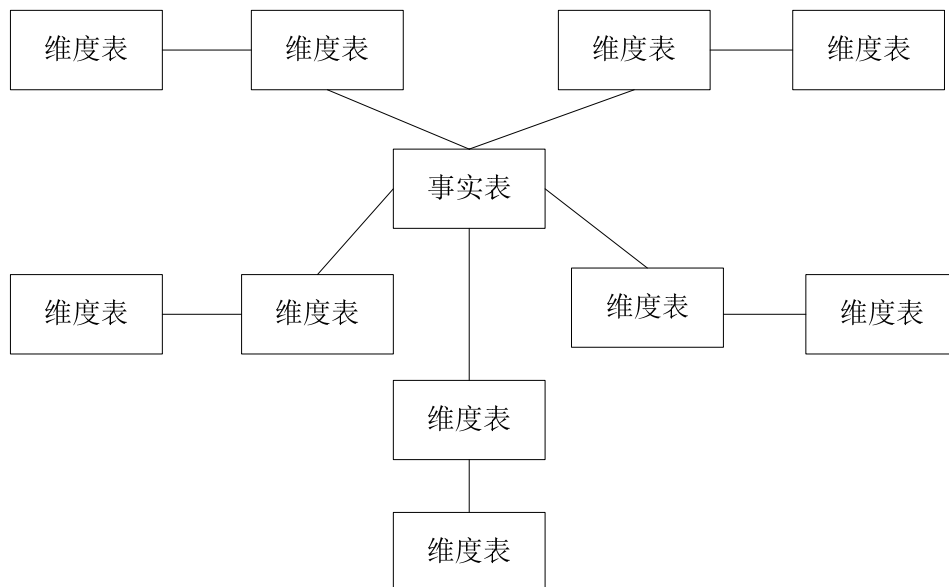


图 5-7 雪花模型结构示意图

在多维数据库的逻辑设计中，除了定义关系模式外，还包括数据粒度的选择、为表增加时间字段、进行表的分割、合理化表的划分等方面。

多维高光谱数据库逻辑模型设计主要是针对岩矿多立方数据，研究合适的多维数据模型，对岩矿多立方数据进行逻辑表达。以上论述表明，当前技术较成熟、灵活性较强的多维数据模型主要是雪花模型和星形模型。以往的研究发现，星形模型中维表不能有效地表达维的层次结构，雪花模型中当维层次过多时又会增加

数据查询的复杂度，人们有时使用折衷方案的混合模型来解决这些矛盾^[44]。岩矿多维数据与一般数据立方体的区别在于它的嵌套结构，在它的多立方结构里嵌套了一个影像超立方结构。在岩矿多维数据中，岩矿类型维有5个层次，采集地点维有5个层次，时间维有3个层次。维的层次不是很多，同时岩矿多维数据模型应能清楚的表达地物类型和时间的层次。所以可以选用雪花模型来构建其多维数据模型，如图5-8所示。

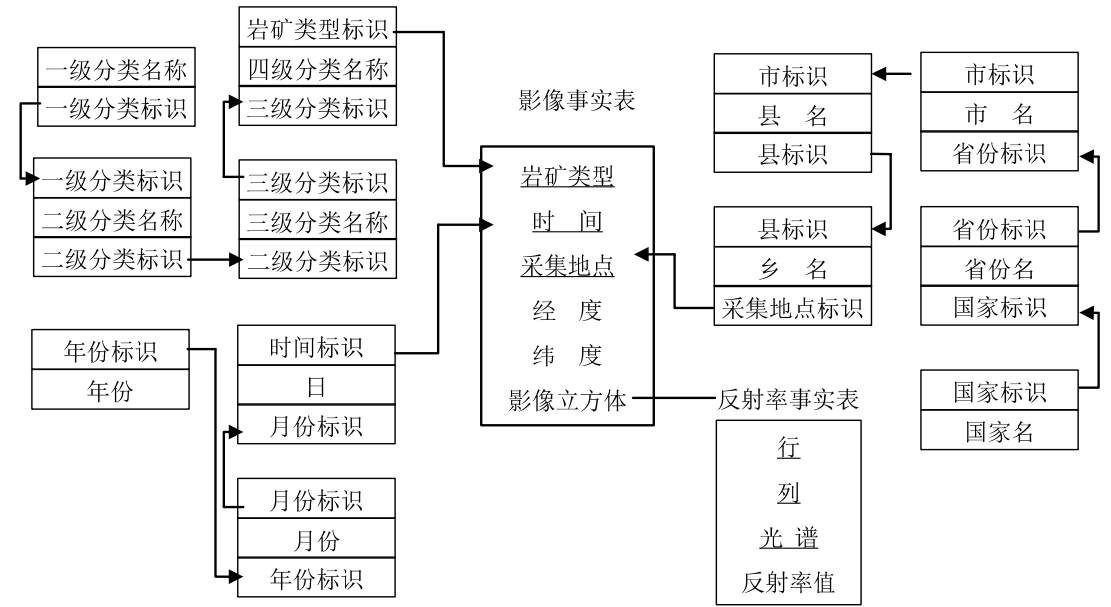


图 5-8 岩矿多维数据雪花模型

5.2.5 岩矿多维高光谱数据库物理模型设计

多维数据库的物理模型设计，是指针对特定的数据库管理系统，如 Oracle、SQL Server 等，确定数据的存储结构、确定多维数据库的索引策略、确定数据的存储策略、对表进行归并、优化存储分配等。物理数据模型显示了数据表的所有结构，包括列名、列数据类型、列约束、是否主键、是否外键以及和其它表之间的关系等。物理模型与逻辑模型的区别在于：物理模型中实体名都是数据表的名字，实体的各个属性都是列的名字，每一列的数据类型都已经指定。

表 5-1 影像数据表

IMG_TABLE		
PK	IMG_ID	NUMBER
FK	TYPE_ID	NUMBER
FK	TIME_ID	NUMBER
FK	LOC_ID	NUMBER
	LONG	NUMBER (7, 4)

	LAT	NUMBER (6, 4)
	IMG_DATA	SDO_GEORASTER

表 5-2 反射率数据表

REF_TABLE		
	ROW	NUMBER
	COL	NUMBER
	BAND	NUMBER
	REF	NUMBER

以上只是按照多维数据模型给出了影像事实表和反射率事实表的物理模型设计，对于岩矿类型维、采集地点维、采集时间维，不一一给出物理模型，但是可以参照以下原则：标识字段的数据类型为 NUMBER，名称字段数据类型为 VARCHAR。

5.3 岩矿多维高光谱数据库多维分析工具集的设计

5.5.1 Oracle 多维数据库技术介绍

Oracle 数据库内嵌入了多维数据库技术，除了支持 ROLAP (Rational Online Analytical Processing)，也支持 MOLAP (Multidimensional Online Analytical Processing)。Oracle 数据库中集成了多维引擎和数据类型，因此具备多维分析能力。Oracle 数据库的 MOLAP 技术简称 Oracle OLAP 选件，它的核心部分是多维引擎、多维数据类型和 OLAP API。OLAP 多维引擎在 Oracle 数据库内运行，它对维模型、多维数据类型和种类繁多的计算提供支持。

OLAP 分析引擎：支持在数据库中进行多维数据的分析查询和计算。多维引擎和数据类型支持复杂的多维计算和规划功能。多维引擎支持各种不同的功能，例如非附加型聚合方法、时序计算、索引和统计功能；还支持用户定义的分析功能。多维引擎利用基于数组的数据结构，对于数据存储称为变量。这些变量是存储在 Oracle 数据文件中的真正多维数据类型，它们在数据存储和查询性能方面十分有效。

分析工作间 (Analytic Workspace)：存储多维格式数据的空间。分析工作区以表的形式存储在 Oracle 数据库中，表的前缀是 AW\$。每个分析工作区只有一个 AW\$表。该表使用 BLOB 数据类型将多维数据类型的数据存储在分析工作区中。AW\$表只能通过 Oracle 数据库的 OLAP 选件的多维引擎读取和写入。

分析工作间管理 (Analytic Workspace Manager)：创建和维护分析工作间的图形化界面，可以创建维度、立方等 MOLAP 对象。负责以下工作：简化维模型、实施物理存储模型、映射关系源、管理分析工作区生命周期、用 XML 文件保存模型设计。

OLAP 选件支持多种聚合函数,如合计、平均、最大/最小、分层加权平均和比例总和。不同于其它技术,OLAP 选件支持随维变化的聚合方法。例如,总人数度量,在时间维上可能以平均方式聚合,在机构维上以合计方式进行聚合。OLAP 选件同时也支持非可加方式聚合。

5.5.2 岩矿多维高光谱数据库多维分析工具集的组成

通过以上对 Oracle 多维数据库技术的介绍,可以发现,Oracle 的多维数据库技术更多的是支持在商业智能上的应用或者是对具有可加性数据的应用,与高光谱应用领域需要的多维分析略有不同,主要体现在:1、高光谱应用领域的每个度量值,包括三个层次,第一个层次是每个影像,第二个层次是每个像元的光谱曲线,第三个层次是组成每个波段的各个像元值;2、高光谱应用需要的多维分析,不仅仅是各个像元值的简单聚合,更何况影像是无法进行聚合的,还要能进行空间维或者光谱维上的分析。所以要为高光谱数据库设计专门的多维分析工具,这些工具同时要继承多维数据库的理念,同时还要适应高光谱应用的特点。图 5-9 所示为高光谱数据库多维分析工具集的组成,它继承了多维数据库特有的钻取、切片、切块和旋转操作,同时,在光谱维、行维、列维、采集地点维、时间维和岩矿类型维上都可以进行这些操作。

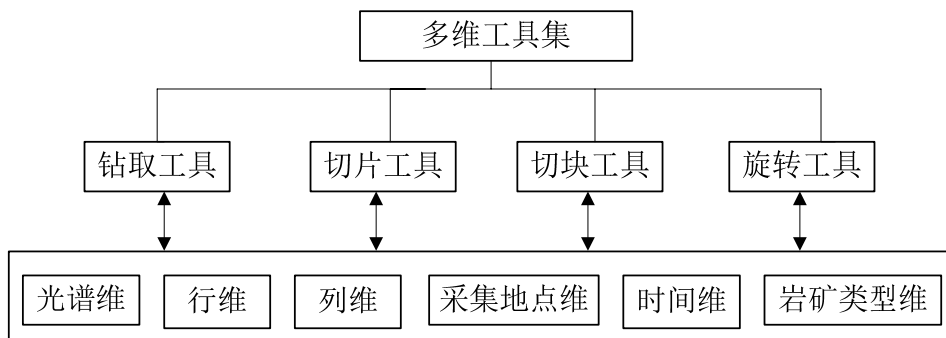


图 5-9 多维分析工具集功能组成图

5.5.3 岩矿多维高光谱数据库多维分析工具的功能设计

多维高光谱数据库的多维分析功能,是在借鉴多维数据库基本分析操作的基础上,结合高光谱数据库的实际情况,进行再进一步的论述、分析和设计。具体来说,岩矿多维高光谱数据库的多维分析工具最终要能实现以下所述的钻取、切片、切块和旋转等分析操作功能。

1、钻取

岩矿多维数据的钻取主要在维层次上进行,包括维层次上钻和维层次下翻。维层次上钻是在岩矿多立方体某个维的层次链上按照其维成员从低等级到高等级进行概括分析操作,聚集出相应度量值的过程^[56]。维层次下翻是上钻的逆操作。

根据岩矿多立方体的特点，将钻取分为两种类型，非影像钻取和影像钻取。非影像钻取是指在岩矿多立方体的地物类型维、时间维、采集地点维上进行的钻取操作，可以分为非影像上钻和非影像下翻。非影像钻取的结果是得到一系列满足钻取条件的影像立方体。比如，按照岩矿的四级分类体系，沿着岩矿类型维进行上钻分析，从第四级分类上钻到第三级分类时，将会得到第三级分类所包含的所有第四级分类的数据；沿着时间维日、月、年的维层次进行上钻分析，当上钻到月这个层次时，将会得到这个月所有的影像数据；沿着采集地点维进行上钻操作，如果从县级钻到市级，那么将得到隶属于该市的所有县份的影像数据。各维的维层次下翻分析是上钻分析的逆操作，如当从省级下翻是市级时，将会得到每个市的影像数据，其它维的下翻以此类推。影像钻取是指在影像立方体的行维、列维和光谱维上进行的钻取，可以分为影像上钻和影像下翻，影像钻取的结果是得到一系列满足钻取条件的反射率曲线或反射率值的组合。影像的行和列代表了影像在空间上的特性，因此单独对行维或列维进行钻取分析没有什么太大的意义，更多的时候是同时在影像的行维和列维上进行钻取分析。在行维和列维同时进行上钻可以理解为类似于为影像建立金字塔的过程，若是在影像立方体的所有波段对行维和列维进行上钻，那么得到的就是一条条的光谱曲线，如果对部分波段进行行维和列维的上钻，那么得到的就是一组反射率值的集合。对光谱维的上钻类似于将光谱曲线从高分辨率采样到低分辨率的过程，下翻类似于从低分辨率模拟到高分辨率的过程。

2、 切片

切片与钻取类似，分为非影像切片和影像切片。当固定岩矿多立方体的任一非影像维，观察影像立方体在其余非影像维上的分布情况，就是对岩矿多立方体的非影像切片操作，非影像切片操作得到的是一系列满足切片条件的影像立方体；当固定影像立方体的任一影像维，观察影像像元数据在其它影像维上的分布情况，就是对影像立方体的切片操作，影像切片的结果是得到一系列满足切片条件的反射率值的组合，这些组合有可能是一组反射率曲线，也有可能是影像某个波段的数据。图5-10为非影像切片示意图，其中，(a)为岩矿多立方体，(b)以岩矿类型C对岩矿多立方体进行切片，(c)以采集地点L3进行切片，(d)以时间T2进行切片。

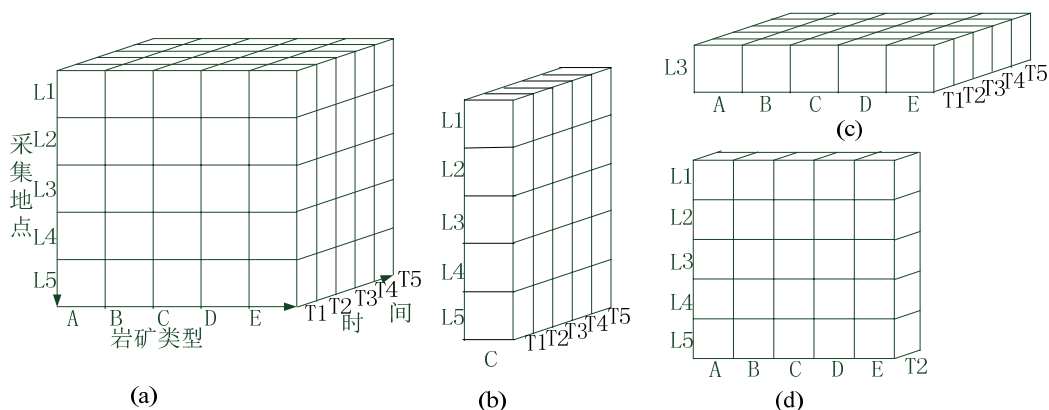


图 5-10非影像切片示意图

影像立方体的影像维切片分析，可以分别在行维、列维和光谱维上进行切片。行维上的切片，得到该行号所对应的列上所有像元的光谱曲线；列维上的切片，得到该列号所对应的行上所有像元的光谱曲线；光谱维上的切片，则得到该波段号所对应的波段影像。图5-11是影像立方体的切片示意图，其中，左下角为（a）是影像立方体效果图，右下角为（b）是行维上切片效果图，顶部为（c）是列维上切片效果图。

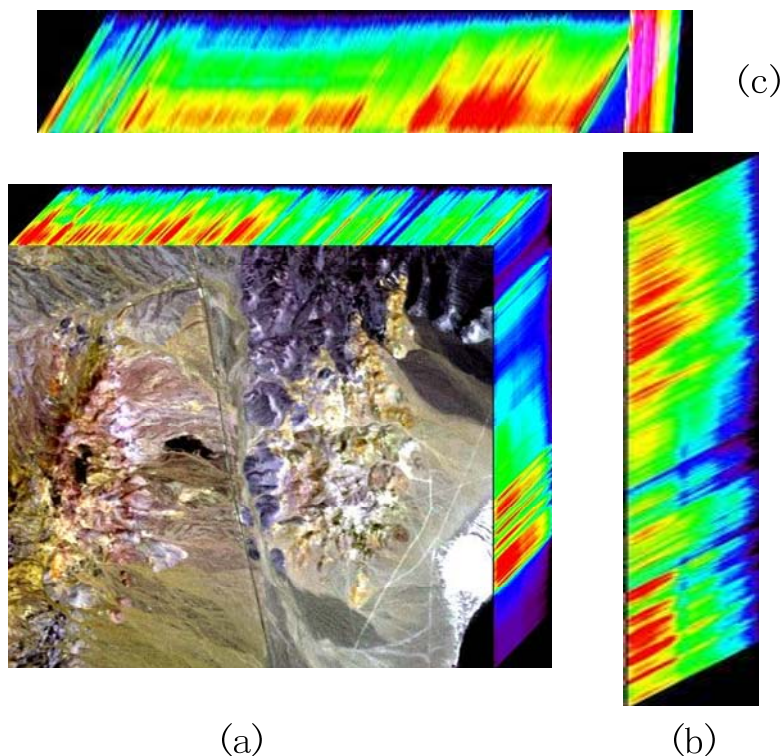


图5-11 影像立方体切片效果图

3、切块

切块与切片类似，也分为非影像切块和影像切块。当固定岩矿多立方体的任意两个维或三个维，观察影像立方体其它非影像维上的分布情况，就是对岩矿多

立方体的非影像切块，非影像切块的结果是得到一系列满足切块条件的影像立方体。图5-12为非影像切块示意图，其中，(a)为岩矿多立方体，(b)为固定采集地点维(L1, L2)和时间维(T1, T2, T3)对岩矿多立方数据进行切块，(c)为固定采集地点维(L1, L2)、时间维(T1, T2, T3)和岩矿类型维(C、D、E)三个维度进行切块。

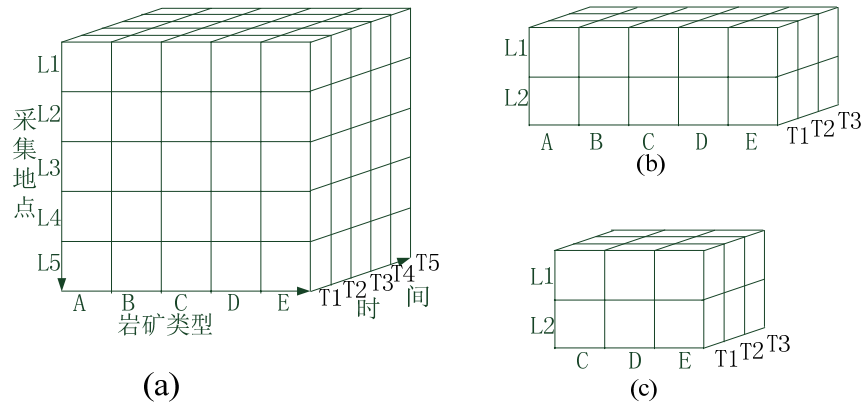


图 5-12 非影像切块示意图

当固定影像立方体的任意两个维或三个维，观察影像像元数据在其它影像维上的分布情况，就是对影像立方体的切块操作，影像切块的结果是得到一系列满足切片条件的反射率值。图5-13是影像立方体切块效果图，其中，(a)是影像立方体，(b)是同时在行维、列维和光谱维上切块的效果图，(c)是在行维和列维上切块的效果图。

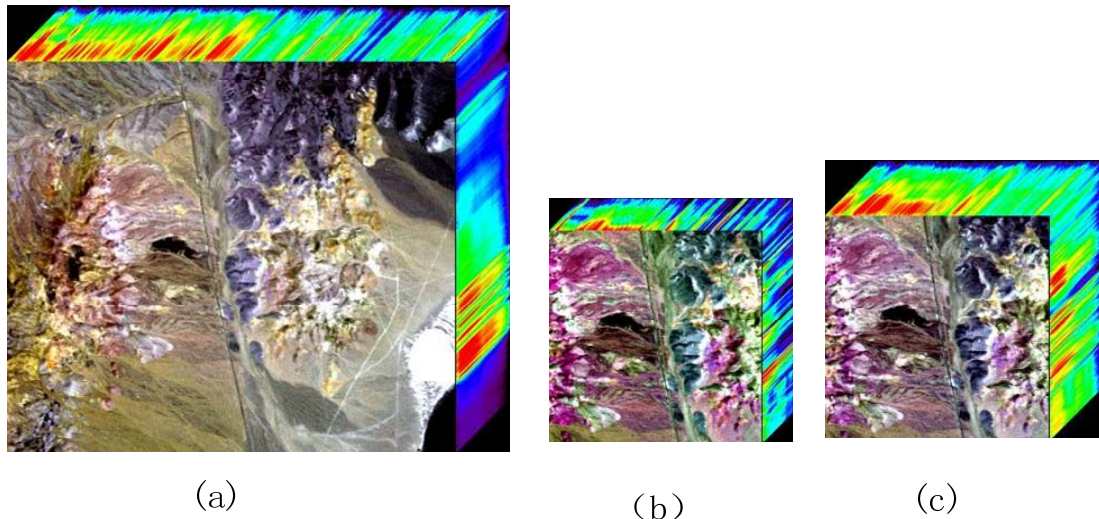


图 5-13 影像立方体切块示意图

4、旋转

旋转就是变换岩矿多立方体维的角度，从不同的方向关心岩矿类型在非影像维上的分布。在(维1, 维2, ..., 维n, 度量)数据立方体的基础上，每个维彼此在垂直方向固定，为了变化人们的观察视点，需改变数据立方体维的方向，

对整个数据立方体进行方向变换,这个过程称之为旋转。数据立方体的旋转并不改变度量的值,只是变换了人们观察分析的角度^[57]。图5-14为旋转示意图,其中,(a)为岩矿多立方体,(b)为以时间维为轴,将(a)顺时针旋转九十度,(c)为以岩矿类型维为轴,将(b)朝观察者方向旋转九十度。

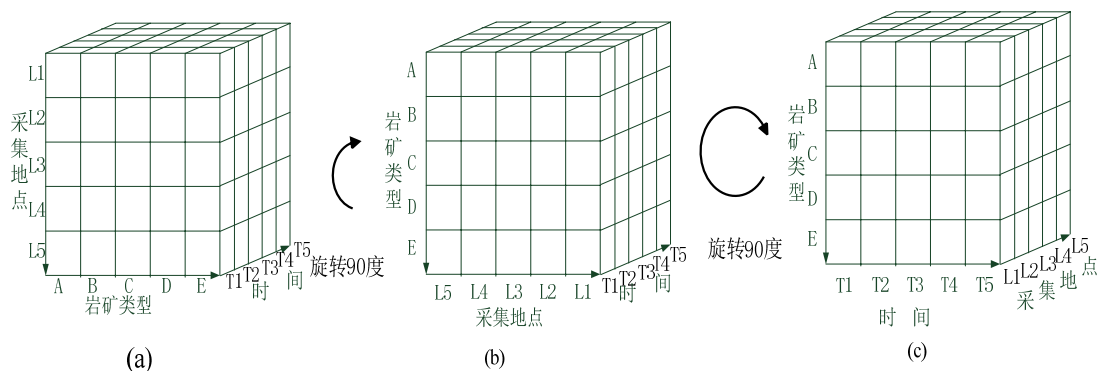


图 5-14 旋转示意图

5.4 本章小结

本章从构建岩矿多维高光谱数据库出发,结合多维数据库的理论、高光谱遥感的相关概念以及岩矿多维高光谱数据库的构建目标,定义并描述了岩矿多维高光谱数据库的基本概念。在岩矿多维高光谱数据库的设计过程中,设计了岩矿多维数据库应实现的功能,研究了岩矿多维高光谱数据库的系统结构。在概念模型设计部分,提出了嵌套的岩矿多维数据结构,建立了岩矿多维数据的事实表,划分了岩矿多维数据各非影像维的维层次关系。在逻辑模型设计部分,介绍了星形模型和雪花模型,比较了它们的优缺点,采用雪花模型构建了岩矿多立方数据的数据模型。在物理模型设计部分,为每个事实表和维度表确定了数据存储表的存储结构。最后在介绍 Oracle 多维数据库技术的基础上,分析了其与多维高光谱数据库在应用上的不同之处,为岩矿多维高光谱数据库的多维分析工具集进行了功能设计。

第六章 结论与展望

建立高光谱数据库是对数据进行有效管理的主要手段,而高光谱数据库与一般的影像数据库又有所不同,它除了具备数据管理的功能,还能对数据进行一些简单的分析。本文以建立岩矿多维高光谱数据库为最终目标,按照两个阶段的研究任务,主要完成了以下两个工作:一是研究并建立 FISS 高光谱数据库原型系统,即第一阶段的研究。本文继承了前人提出的高光谱数据库的概念,对 FISS 高光谱数据库进行了分析与设计,研究了 FISS 数据存取的关键技术,设计实现了 FISS 影像及光谱的显示方法,开发了 FISS 高光谱数据库原型系统,在原型系统的基础上,集成了几个具有高光谱特色的光谱分析方法。二是结合多维数据库的理论,对多维高光谱数据库进行了初步研究,主要从基本概念、数据库设计和多维分析工具集设计三个方面展开了研究,是第二阶段研究的序幕,为构建多维高光谱数据库提供了一定的基础。

6.1 结论

1、在深入分析数据库的功能需求和业务流程、全面设计其体系结构和功能模块以及仔细选择开发环境的基础上,根据 FISS 数据的特点和系统应用的需求,提出了以 GeoRaster 对象为核心的 FISS 数据存储模型。研究表明,该存储模型能够为高光谱数据库系统的数据存储、显示以及光谱分析功能提供有效的支持。

2、基于 FISS 数据存储模型,提出了基于存储过程的数据入库方法,该方法实现了高光谱数据图谱合一的存储;提出基于动态数据表转换的数据获取方法,该方法实现了从数据库中获取像元的光谱数据。在 Arc Engine 开发工具和 C# 的支持下,研究了 FISS 合成影像的显示方法和光谱曲线的显示方法。实验结果表明,数据获取方法、影像显示方法和光谱曲线显示方法的完美结合,有效的实现了高光谱数据图谱合一的显示。

3、开发了 FISS 高光谱数据库原型系统,实现了数据的导入、查询、显示功能,将光谱平滑、光谱重采样、光谱匹配和包络线去除算法等常用的光谱分析方法集成到了系统中,实现了基于数据库的光谱分析功能,为进一步的数据筛选提供了支持。

4、以多维数据库理论为基础,对多维高光谱数据库进行了初步研究。进一步定义和论述了岩矿多维高光谱数据库的基本概念,设计了多维高光谱数据库的系统结构,提出了嵌套的岩矿多维数据结构,建立了岩矿多维数据的事实表,划分了岩矿多维数据各非影像维的维层次关系,构建了岩矿多维数据的雪花模型。设计了多维高光谱数据库多维分析工具集的组成,详细论述了每个分析工具应实现的功能,并以示意图的形式给出了部分实现效果。这些研究工作的开展,为岩

矿多维高光谱数据库的后续研究奠定了坚实的理论基础,指出了下一步应该努力的方向。

6.2 展望

虽然本文对 FISS 数据的高光谱数据库原型系统进行了比较深入的研究和实现,也基于多维数据库的基本理论对 FISS 多维高光谱数据库进行了研究。但是由于时间和条件的限制,FISS 高光谱数据库原型系统离一个功能完善、性能稳定的系统还有不小的差距,在多维高光谱数据库的研究上,本文仅仅是拉开了万里长征的序幕,还有很多关键的东西有待研究。展望未来,应该从以下方面进一步深入和改进本文的研究:

1、FISS 高光谱数据库功能的完善。本原型系统基本实现了 FISS 数据的入库、查询检索、影像与光谱的显示,尚有用户管理、系统管理和数据导出等功能尚未完成,在以后的研究中,应逐步完善这些功能。

2、FISS 高光谱数据库性能的改善。本原型系统在研发过程中,将侧重点放在了主要功能的实现上,对系统运行的性能如响应时间、入库效率、查找效率等虽有所考虑但并不全面,在后续的研究中,应进一步研究如何能使系统的性能有更大的改善。

3、数据存储系统的改善。未来将需要对大量的 FISS 数据进行存储,在后续的研究中,需要结合最新的存储技术、存储系统如分布式存储、集群存储、对象存储等,充分发挥 Oracle 数据库的优势,来对存储系统进行改进。

4、多维建模工具的设计与研发。高光谱数据的特殊性,决定了高光谱数据库需要一个新的多维建模工具,就像 Oracle 的分析工作空间管理器 (Analytic Workspace Manager, AWM) 那样,能够进行构建维模型、实施物理存储模型、映射关系源、数据加载、刷新和聚合等功能,同时还应能适应高光谱数据的特点。

5、多维分析工具的实现。本文只是对多维分析的功能进行了介绍,应进一步分析、明确多维高光谱数据库多维分析的需求,完善功能规划。在以后的研究中,需要花费一定的时间和精力,组织人员来开发该多维分析工具集,使之真正能实现多维分析的功能。

参考文献

- [1] 童庆禧, 张兵, 郑兰芬. 高光谱遥感—原理、技术与应用[M]. 北京: 高等教育出版社, 2006.
- [2] 叶荣华, 范文义, 龙晶. 高光谱遥感技术在荒漠化监测中的应用[M]. 北京: 中国林业出版社, 2001.
- [3] 张良培, 张立福. 高光谱遥感[M]. 第1版. 北京: 测绘出版社, 2011.
- [4] Lu S, Wu B, Yan N, et al. Water body mapping method with HJ-1A/B satellite imagery[J]. International Journal of Applied Earth Observation and Geoinformation, 2011,13(3):428-434.
- [5] 童庆禧, 薛永祺, 王晋年, 等. 地面成像光谱辐射测量系统及其应用[J]. 遥感学报, 2010,14(3):409-415.
- [6] 童庆禧. 高光谱遥感的现在与未来[J]. 遥感学报, 2003,7(增刊):1-12.
- [7] 张立福. 通用光谱模式分解算法及植被指数的建立[D]. 武汉: 武汉大学, 2005.
- [8] 汪晓航. 数据库技术的发展[J]. 企业导报, 2011(02):289.
- [9] 吕玲玲. 数据库技术的发展现状与趋势[J]. 信息与电脑(理论版), 2011(08):118-120.
- [10] Barroso L A, Holzle U. The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines: Morgan & Claypool Synthesis Series on Computer Architecture, 2009[C].
- [11] Malewicz G, Austern M H, Bik A J C, et al. Pregel: a system for large-scale graph processing: Proceedings of the 2010 international conference on Management of data, Indianapolis, Indiana, USA, 2010[C]. ACM.
- [12] Schroeder B, Pinheiro E, Weber W. DRAM errors in the wild: a large-scale field study: Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems, Seattle, WA, USA, 2009[C]. ACM.
- [13] 金崇英, 李小勇, 白英彩. 海量存储系统的发展与展望[J]. 计算机应用与软件, 2011(08):193-195.
- [14] Pedersen T B, Christian S J. Multidimensional database technology[J]. Computer, 2001,34(12):40-46.
- [15] 刘波. 基于地面成像光谱系统的植物信息提取研究[D]. 北京: 中国科学院遥感应用研究所, 2010.
- [16] 张雄飞. 网络环境下高光谱数据库构建及其应用实践[D]. 中国科学院研究生院(遥感应用研究所), 2003.
- [17] 李兴. 高光谱数据库及数据挖掘研究[D]. 中国科学院研究生院(遥感应用研究所), 2006.
- [18] 欧立业. 遥感影像数据库技术比较研究[J]. 江西测绘, 2007(S1):93-96.
- [19] 朱海勇. 基于Oracle Spatial空间数据库管理系统的设计与实现[D]. 吉林大学, 2007.
- [20] 张雄飞, 张兵, 张霞, 等. 高光谱数据在数据库中的高效存储技术研究[J]. 遥感学报, 2004(05):404-408.
- [21] Li X, Hu X, Zhang B, et al. A hyperspectral-environmental database in China, 2005[C]. SPIE.
- [22] Li X, Zhang B, Tong Q, et al. Demand-oriented hyperspectral database and its applications: IGRSS, 2005[C].
- [23] 董广军, 王鹏, 赵楠. 高光谱数据在Oracle中的存储及应用技术研究[J]. 测绘科学技术学报, 2006(03).
- [24] 满旺, 张杰林, 曹代勇. 基于Oracle的高光谱数据管理平台设计与应用[J]. 世界核地质科学,

- 2007(04).
- [25] 孙婷, 张立朝, 唐汉松, 等. 基于 Oracle 的高光谱影像数据库研究[J]. 北京测绘, 2007(02):1-4.
- [26] 向娟, 李钢. 高光谱遥感影像管理系统的设计与实现[J]. 测绘与空间地理信息, 2008(05).
- [27] 王焕宝. 多维数据库研究[J]. 芜湖职业技术学院学报, 2001,3(1):19-20.
- [28] 窦凤云. 基于 Oracle 数据库的海量空间信息系统的优化[D]. 天津大学, 2006.
- [29] 华国栋, 刘文予. 基于 ADO.NET 的数据库访问及其性能优化[J]. 计算机应用研究, 2004(06):215-218.
- [30] 邱洪刚, 张青莲, 陆绍强. ArcGIS Engine 开发从入门到精通[M]. 北京: 人民邮电出版社, 2010.
- [31] 白继伟. 基于高光谱数据库的光谱匹配技术研究[D]. 中国科学院研究生院(遥感应用研究所), 2002.
- [32] Oracle Corporation. Oracle Spatial GeoRaster Developer's Guide 11g release 1(11.1)[EB/OL].
- [33] 宋国兵, 陈奇. 文件数据的数据库 Blob 存储及效率分析[J]. 计算机工程与设计, 2010(21):4625-4627.
- [34] 田海清, 应义斌, 徐惠荣, 等. 运动西瓜可见/近红外光谱采集系统及品质检测试验研究[J]. 光谱学与光谱分析, 2009(06):1536-1540.
- [35] Goetz A F H, Vane G, Solomon J E, et al. Imaging Spectrometry for Earth Remote Sensing[J]. Science, 1985,228(4704):1147-1153.
- [36] Kruse F A, Lefkoff A B, Boardman J W, et al. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data[J]. Remote Sensing of Environment, 1993,44(2-3):145-163.
- [37] Baugh W M, Kruse F A, Atkinson W W. Quantitative Geochemical Mapping of Ammonium Minerals in the Southern Cedar Mountains, Nevada, Using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)[J]. Remote Sensing of Environment, 1998,65(3):292-308.
- [38] Colliat G. OLAP, Relational, and Multidimensional Database Systems[J]. SIGMOD Record, 1996,25(3):64-69.
- [39] 李兴, 张兵, 张霞, 等. 高光谱数据仓库模型设计[J]. 遥感学报, 2003,7(增刊):61-68.
- [40] Trujillo J, Palomar M. An object oriented approach to multidimensional database conceptual modeling (OOMD): Proceedings of the 1st ACM international workshop on Data warehousing and OLAP, Washington, D.C., United States, 1998[C]. ACM.
- [41] Papadias D, Kalnis P, Zhang J, et al. Efficient OLAP Operations in Spatial Data Warehouses[M]//JENSEN C, SCHNEIDER M, SEEGER B, et al. Advances in Spatial and Temporal Databases. Springer Berlin / Heidelberg, 2001:443.
- [42] Zhou X, Truffet D, Han J. Efficient Polygon Amalgamation Methods for Spatial OLAP and Spatial Data Mining: Lecture Notes in Computer Science[Z]. GÜTING R, PAPADIAS D, LOCHOVSKY F. Springer Berlin / Heidelberg, 1999: 1651, 167-187.
- [43] Stefanovic N, Han J, Koperski K. Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes[J]. IEEE Trans. on Knowl. and Data Eng., 2000,12:938-958.
- [44] 季民, 靳奉祥, 李婷, 等. 海洋多维数据仓库构建研究[J]. 海洋学报, 2009,31(6):48-53.
- [45] 林友芳, 黄厚宽, 田盛丰. 铁路货运数据仓库多维视图的组织及其物化策略[J]. 铁道学报, 2001,23(2):8-12.
- [46] Bréant C, Thurler G, Borst F, et al. Design of a Multi Dimensional Database for the Archimed

- DataWarehouse[J]. Studies in health technology and informatics, 2005,116:169-174.
- [47] 朱东妹, 潘杏仙, 付勇. 高校图书馆多维数据仓库模型的设计与实现[J]. 情报探索, 2010,2:99-101.
- [48] 李日容, 吴冲龙, 邵玉祥. 运用多维数据模型实现油气勘探数据集市[J]. 计算机工程与应用, 2006,14:229-232.
- [49] 赵霞琦. 以立方体结构为例探讨多维数据库在电子文献组织中的使用[J]. 图书情报工作, 2006,50(8):67-69.
- [50] 徐奇. 基于多维数据库的集装箱空箱调运优化研究[D]. 大连, 2009.
- [51] Wang J, Zhang L, Liu Q, et al. Spectral Database of Typical Objects in China[M]. Beijing: Science Press, 2009.
- [52] Tong Q, Xue Y, Wang J, et al. Development and application of the field imaging spectrometer system[J]. Journal of Remote Sensing, 2010,14(3):409-415.
- [53] Han J, Stefanovic N, Koperski K. Selective materialization: An efficient method for spatial data cube construction[M]//WU X, KOTAGIRI R, KORB K. Research and Development in Knowledge Discovery and Data Mining. Springer Berlin / Heidelberg, 1998:144.
- [54] 邹逸江, 杨晓平. 空间数据立方体的技术框架[J]. 计算机应用研究, 2006(10):27-29.
- [55] 王付山. OLAP技术研究与发展述评[J]. 软件导刊, 2007,19:11-13.
- [56] 邹逸江. 多维空间分析的关键技术_空间数据立方体[J]. 地理与地理信息科学, 2006,22(1):12-16.
- [57] 邹逸江, 李德仁, 王任享. 空间数据立方体分析操作原理[J]. 武汉大学学报(信息科学版), 2004,29(9):822-826.

发表文章目录

- 1) Qin Huanhu, Zhang Lifu. Design of hyperspectral multidimensional database for rock and mineral. 2011 International Conference on Remote Sensing, Environment and Transportation Engineering, RSETE 2011 – Proceedings, 2011, pp. 1092-1095. (EI: 20113614292078)
- 2) 覃环虎,张立福,刘凯.高光谱影像库的数据存取技术研究.计算机工程. (CSCD,已接收)
- 3) Ding Ling,Zhang Xia,Ji Min,Shuai Tong, Qin Huanhu,Zhang Xuewen. Spectral indices for estimating the Fractional cover of non-vegetation land-cover types in karst environment. 2011 International Conference on Remote Sensing, Environment and Transportation Engineering, RSETE 2011 – Proceedings,2011, pp. 2499-2502. (EI: 20113614292428)

致 谢

三年多以前,怀揣着一颗充满向往、崇敬和忐忑的心,我踏上了由大连开往北京的列车,来遥感所参加保研复试。时至今日,复试时的情景仍历历在目,恍如发生在昨日。可是今天,却蓦然发现,研究生毕业之期已然不远,不禁感慨光阴之易逝,岁月之如梭。一路走来,得到了众多师友的帮助与关怀。

衷心感谢我的恩师张立福老师。三年前,承蒙老师的关爱,我得以提前来所学习,在您的引导下,我逐渐完成了由计算机专业到高光谱遥感研究的转换,在您的指导下,我完成了自己的本科毕业论文。在遥感所求学的三年时光里,老师严谨的治学态度,踏实的工作作风,亲切的待人风格,活跃的学术思维以及爱生如子的情怀,给我留下了非常深刻的印象并产生了非常深远的影响。在我撰写毕业论文的过程中,老师多次主动关心并与我进行讨论,从论文的选题、提纲的拟定再到论文的定稿,倾注了老师大量的心血。大恩不言谢,学生唯有更加勤奋努力,才不会辜负您的栽培与期望。

衷心感谢我的导师童庆禧院士。童老师是我国高光谱遥感事业的开山鼻祖,他知识渊博、治学严谨、视野广阔、德艺双馨,从学业上和生活上关心高光谱研究室每个同学的发展与成长,从思想上、方法上引导我们走上通往科学殿堂的道路,是一名卓越的师者与学者。愿在您引领之下,祖国的高光谱遥感科学事业蒸蒸日上。

衷心感谢郑兰芬研究员和王晋年研究员。郑老师在日常学习和工作中给我细致入微的关怀,从思想、学习和生活上全方位的关心我的成长和进步。王晋年老师多次参加我们的学术交流会,他在工作、生活等方面给予的关怀和指导以及如何进行科学研究上给予的启发都对我产生了非常深远的影响。

衷心感谢张霞老师、方俊永老师和刘学师兄在学习上和生活上给与我的关心和帮助,跟他们都有过难忘的出差经历。霞老师,曾记否,金秋十月,喀斯特之行,我们三下古周,转战于环江、凤山、大化、都安?喀斯特地貌之美以及那里人们生活之艰辛都令我震撼。方老师、刘师兄,可记得,我们曾分别于阴风刺骨的寒冬以及骄阳似火的仲夏奋斗在广西梧州,等待天气进行航飞试验?运五飞机巨大的轰鸣声以及蓝天大酒店我们欢快的笑声经常回响在我的耳畔。这三次出差,扩大了我的见识,增加了我对人生的感悟,是我人生中难忘的美丽回忆。

感谢王树东老师、王潇师姐、赵冬师兄、张晓红师姐、吴太夏师兄、杨杭师兄、岑奕师姐所给予的关心和帮助。尤其感谢吴太夏师兄和岑奕师姐在论文撰写过程中的诸多建议和帮助,以及在工作中所给予的大量支持、帮助、关心和理解。

感谢和海霞、李庆亨、申茜、张浩、刘波、李儒、岳跃民等师兄师姐的帮助。尤其要感谢和海霞师姐和李庆亨师兄在学习和生活上的关心以及在我遇到困难、碰到挫折时给予的关心、支持和鼓励。

感谢高光谱遥感研究室的师兄师姐们、同学们以及师弟师妹们所给予的关心和帮助,三

年里，我们一路走来，因为有你们，我的研究生生活才会如此的丰富多彩，他们是：张学文、孙韬、李颖、宋鹏飞、宁艳玲、黄长平、陈小平、胡顺石、王倩、尚坤、李晶宜、吴松、帅通、谢一淞、张明、兰穹穹、姜海玲、孙雪剑、刘凯、赵恒谦、张旭凯、丰明博、朱海涛、余晓君和王晶。特别感谢张学文师兄和黄长平师兄在我碰到疑问的给予的耐心解答以及在论文写作过程中的指导和帮助。特别感谢尚坤和李晶宜在论文写作中给予的关心和支持。特别感谢刘凯和孙雪剑尤其是刘凯在工作中给予的帮助，论文的顺利完成有你们的功劳，感谢你们不辞辛劳的为我的论文校稿。同时也感谢赵恒谦师弟为我的论文进行校稿。

感谢人事处孙文新老师、研究生处吴晓清老师和刘戈平老师、所办陈雪老师和路遥老师在学习和工作中给予的支持和帮助。

感谢论文的两位评审老师的认真评阅以及提出的宝贵建议，感谢出席论文答辩会的评委老师们提出的宝贵建议。

特别感谢我的好朋友覃帮勇同学和李霞同学，帮勇热情、真诚而执着，霞姐漂亮、温柔而婉约，从你们的身上，我学到了很多待人处事的方法。感谢你们在日常生活中给予的支持和帮助，感谢你们在我生病时给予的关心，这让我感到无比的温暖！祝你们幸福！

特别感谢我的好朋友刘爽同学，你知识丰富、涉猎广泛，在很多方面给我启发，与你的交谈常常使我受益匪浅，三年的研究生生涯因为有你而更加充实和美丽。

特别感谢我的学妹蒙丽洁和卢燕香，在京求学的日子里，我们一起游玩，有许多共同的美好回忆。感谢你们的关心和鼓励，以及给我研究生生涯带来的无尽欢笑和喜悦。

特别感谢王静、刘一良、刘珠妹、彭菁菁、王国军，感谢遥感所 09 级的硕士同学，他们是：蔡丹路、曹彪、宋玉娟、段祺坤、曾甜、张婉春、王智慧、赵飞飞、司敬知等，别忘了，我们不仅仅是同学，更是曾经一起战斗过的牌友，用娟姐的话说，那叫“牌友的情谊深”。

在这里我想再次感谢尚坤和谢一淞，在我研究生生涯有着最重要影响的四个名词：高光谱遥感研究室、梧州出差、车库楼、升级，都有你们的身影，我们不仅是好同学、好朋友，同时也是好搭档、好牌友，感谢你们陪我一路走来！

感谢大学的兄弟姐妹们，我们互相勉励，共同奋进。他们是：徐世江、王学成、夏晋京、荣坤、徐瑞雪、董恩超、叶云鹏、孙欣和陈奥，尤其要感谢徐世江对论文提出的宝贵意见。

感谢韦美新老师、郑阳梅老师、吴小宁老师多年来给予的关心和指导，十多年来，因为有你们，我少走了很多弯路。

最后，我要衷心感谢我的亲人们，首先是我的父母，谁言寸草心，报得三春晖，父母的养育之恩，无私的关爱，无限的宽容和无尽的支持，是我多年来一直不断拼搏奋进的动力；感谢我的两个姐姐以及堂哥给予的支持、关心和鼓励；感谢我的舅舅和叔叔，谢谢你们在我遇到困难时所给予的巨大的帮助；感谢奶奶和外婆给予的呵护和慈爱。

感谢所有关心和帮助过我的人们！

衷心祝愿祖国的遥感事业取得更大的成就！