

# 开放红外光谱分析系统的架构设计<sup>①</sup>

王 帅, 冯新泸, 杨 岚, 周 俊, 崔 健

(中国人民解放军后勤工程学院 军事油料应用与管理工程系, 重庆 401331)

**摘 要:** 目前国内的红外光谱软件缺乏统一开放的软件架构模型, 使得红外光谱软件的开发缺乏合理的设计和规划, 由此产生了诸如软件稳定性、数据格式通用性、软件扩展性等一系列问题。针对这一现状, 提出了一种具备灵活性、复用性的红外光谱系统架构设计方案。在软件数据结构方面, 借鉴 JCAMP-DX 国际光谱数据格式标准对光谱数据进行抽象, 重新设计了光谱软件的底层数据结构, 以满足光谱数据处理的特殊需要; 在系统设计方面: 将系统划分为相互独立的三层, 即: UI 界面表示层、光谱数据处理逻辑层、底层仪器控制层。系统整体设计采用模块化设计思想, 可根据具体需求进行功能定制, 光谱数据处理器的设计采用多核心设计, 便于计量学算法的扩充和更新。

**关键词:** 红外光谱; 分析软件; 架构设计; 软件开发; JCAMP-DX

## Architecture Design of Infrared Spectroscopy Software

WANG Shuai, FENG Xin-Lu, YANG Lan, ZHOU Jun, CUI Jian

(Postgraduate Institute, PLA Logistical Engineering University, Chongqing 401331, China)

**Abstract:** The IR analysis software lacks a unified open software architecture model currently, this resulted in a series of problems, software stability, data format interoperability and software scalability was poor. The article aimed at this situation, proposed a flexible and reusable IR system architecture design. In the software data structure, referenced the JCAMP-DX standard spectral data format, re-designed the software underlying data structure to meet the spectral data processing needs. In the system design, the system was divided into three separate layers: UI layer, spectral data processing logic layer and instrument control layer. The IR analysis systems adopted the modular design, which can be customized according to functional requirements. Spectral data processor use multi-core design, this makes it is easy to extend and update the chemometrics algorithms.

**Keywords:** infrared spectroscopy; analysis software; architecture design; software development; JCAMP-DX

红外光谱反映的是一种间接的物质结构信息, 因此在获得红外光谱图之后, 我们不得直接得到结论, 需要专家对谱图进行解析之后, 才能得到结论。但是红外谱图是一种典型的多维数据信息, 一条谱线一般包含上千个点的数据, 人工直接阅读、分析数据是不现实的。

目前, 对于谱图的分析主要采用计算机谱图软件, 来进行谱图的分析、计算。但是看到, 目前的光谱分析软件都是被几个大的光谱仪器公司所控制, 例如

PE、惠普、安捷伦、尼高立等公司, 软件均是商业软件, 单独购买价格昂贵, 一套动辄上万。且各公司的软件往往还需要与公司光谱仪器连接后, 才能运行。另外, 各个公司的软件之间数据也存在交换障碍问题。尽管国际原子与分子光谱联合委员会建立了一个标准的光谱数据交换格式, 即: JCAMP/DX 文件格式。但是, 实际上各个光谱公司均有自己的文件格式, 例如: 德国布鲁克、日本岛津公司的仪器数据就是其自身的格式, 他们可以利用各自软件转换为 JCAMP/DX 文件

<sup>①</sup> 收稿时间:2010-06-28;收到修改稿时间:2010-09-07

格式，但是往往会出现光谱文件格式繁多，导致数据格式不兼容的问题。

当前国内各光谱仪器公司的软件缺乏科学、统一的软件架构标准，软件稳定性存在很大问题，同时，统一公司的软件风格差异很大，导致用户学习成本的升高。因此，开发一种开放的红外光谱软件具有很大的必要性。首先，软件应该专门针对光谱分析，与仪器平台脱离；其次，软件应该具有最大的光谱数据格式兼容性，在保护知识产权的前提下，广泛兼容各公司的光谱数据格式；第三，软件采用开放架构设计，兼容数据格式，以及算法处理器均可以组件形式进行扩展。

事实上，红外光谱的检测机理很简单，红外光源发出光，通过样品后，经 CCD 检测器检测，然后将信号传回电路，最终将数据送给上位机(如图 1)。

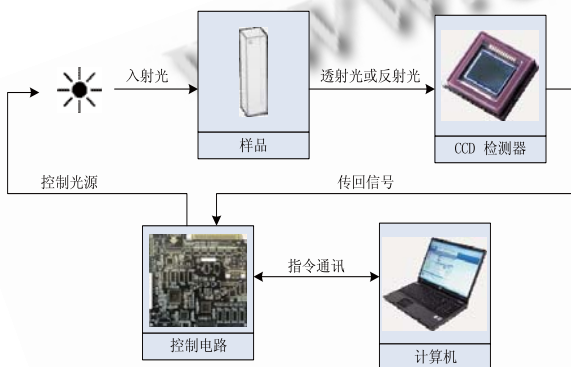


图 1 红外光谱仪各功能模块结构图

以下是傅里叶光谱仪的基本原理图(图 2)：

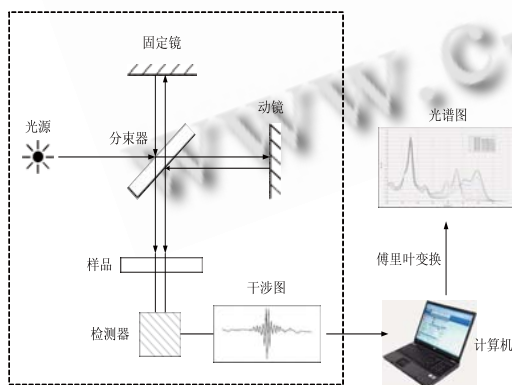


图 2 傅里叶光谱仪基本原理图

根据以上检测过程，可知红外光谱分析系统主

要包括两个大的部分：仪器控制部分和光谱数据分析处理部分，部分红外光谱软件侧重于数据处理分析，因此可省去仪器控制功能。软件系统的整体架构如图 3 所示。下以 VB 语言为例，进行软件架构设计的说明。

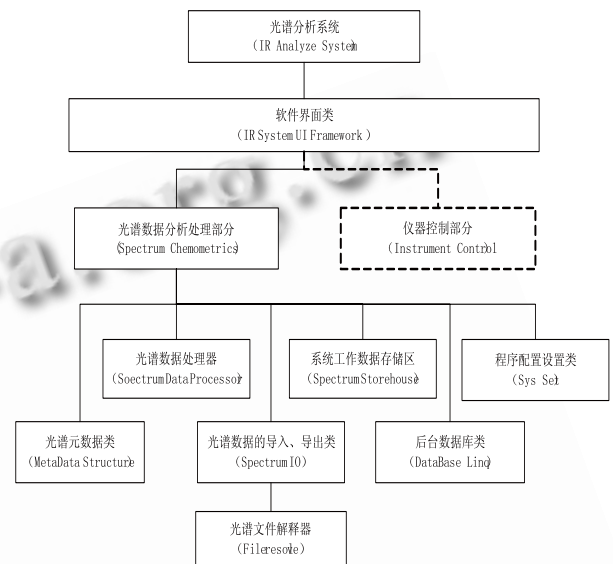


图 3 红外光谱分析系统整体架构图

### 1 光谱元数据的设计：光谱的基本数据结构

光谱软件开发的最基础、最终要的工作是光谱数据模型的建立。光谱的数据文件一般是由头文件和数据文件两部分构成，头文件主要是对光谱数据进行描述，可包含试验条件及仪器参数；数据部分则是光谱文件的主体部分(表 1)。

如图 4 示，由各个类聚合，最终构成一个油品光谱文件类<sup>[1]</sup>。

表 1 光谱元数据中包含的类

序号	类名	功能说明
1	CSpectrumHeader	光谱数据头
2	CSpectrumData	光谱数据体
3	CSpectrumLine	由光谱数据头、光谱数据体组合成的光谱数据文件类
4	COilNorm...	各种指标类(例如油品指标、煤炭指标等，可以细化)
5	COilSpecFile...	由光谱数据文件类和指标类组合成的光谱文件类

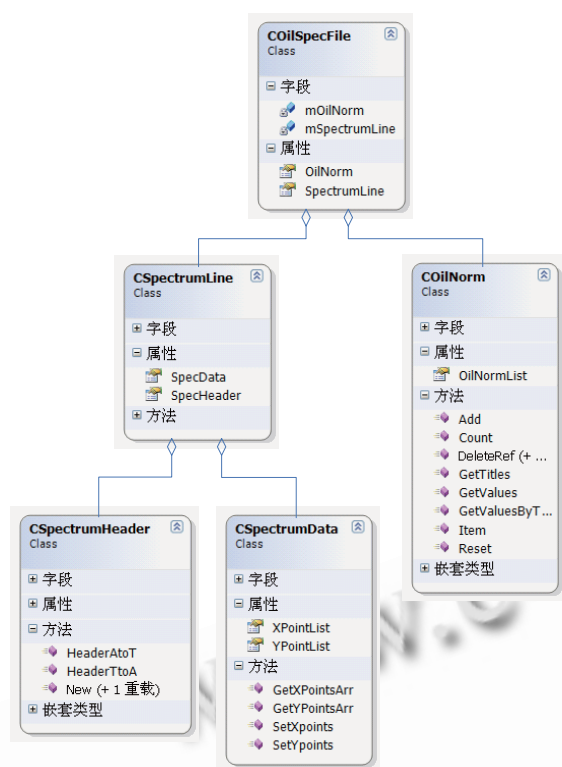


图 4 光谱元数据中各个类的聚合关系图

光谱数据头 CSpectrumHeader 类是一个包含了光谱文件信息类，包含以下信息：光谱文件名、x 轴格式、y 轴格式、文件时间、数据点数、x 轴起点、x 轴终点、文件说明、文件版本、y 轴的放大倍数、光谱数据的内部索引 ID(针对重名数据)、光谱处理的信息(括求导、平滑、归一、PCA、PLS 等)

光谱数据体 CSpectrumData 类的设计，应考虑到计算机的处理效率，一方面要保证光谱数据增减的灵活性，另一方面又应该保证数据的运算效率。因此应使用 List 泛型技术实现，它比 ArrayList 链表具有更加优良的性能。这样可以避免数据对象的“装箱”与“拆箱”，而降低运算速度<sup>[2]</sup>。泛型结构的实现如下：

```
Public Class CSpectrumData
    '这里采用list泛型
    'x轴点数组
    Private mXPointList As New List(Of Single)
    'y轴点数组
    Private mYPointList As New List(Of Single)
    .....
```

End Class

2 文件解释器设计

由于不同公司的光谱数据文件是不同的，文件解释器负责将外部数据格式转化为统一的内部数据格式，即 COilSpecFile 类的对象(图 5)。只有这样，在软件内部，才能够以统一的标准算法处理分析。

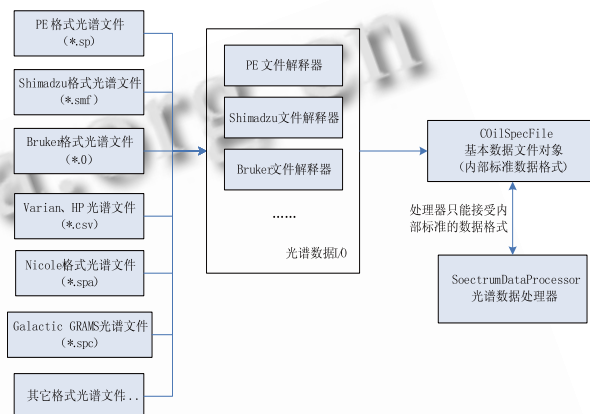


图 5 文件解释器结构图

实时上，就是根据不同文件的后缀名称，采用不同的数据解析读取规则，将光谱文件数据读入计算机内存中。各种文件解释器可利用接口进行调用，并且可以在必要时进行解释器扩充，向软件中添加新的文件解释器，以使软件支持更多格式的光谱文件。

Public Interface CfileInterface

```
'序号
ReadOnly Property _ID() As String
'解释器名称
ReadOnly Property _ModeName() As String
'调用解释器解释文件
Function Resolve(ByVal sPath As String, _
                ByVal Suffix As String) _
                As CCommonFile
'调用解释器存储文件
Function Restore(ByVal sPath As String, _
                ByVal sName As String, _
                ByVal CComFileObj As
                CCommonFile) _
                As Boolean
```

.....

End Interface

所有解释器均支持 CFileInterface 接口, 解释器中提供对于\_ID 属性、\_ModeName 属性、Resolve 方法和 Restore 方法的支持, 这样软件中可以相同的形式调用文件解释器。

3 光谱数据处理器的设计

光谱数据处理器设计主要依托化学计量学理论[3]。光谱数据处理器数学运算功能主要有三个方面: 光谱数据预处理、模式识别(定性分析)和回归校正(定量分析)[4,5]。

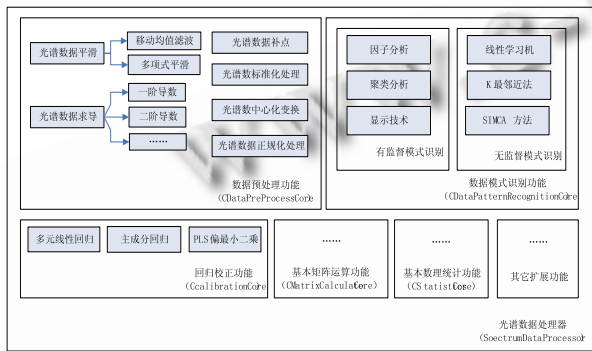


图6 光谱数据处理器结构图

如图6所示, 为光谱数据处理器的基本结构, 包括数据预处理功能、数据模式识别功能、回归校正功能等, 以及基本的矩阵运算和数理统计功能。

光谱数据处理器的工作任务分为两种: 一是建模, 二是预测。

建模过程中, 对于进入处理器的训练集数据具有以下要求: 数据预处理功能核心、数据模式识别功能核心只接收不带指标的光谱数据, 即: CSpectrumLine 类的对象。

回归校正功能核心需要指标数据进行定量计算, 因此接收带有指标的光谱数据, 即: COilSpecFile 类的对象, 如图7所示。

预测过程中, 对于进入处理器的测试集数据要求同上。

一般在实践中, 处理器的算法开发工作量较大, 为了缩短系统开发的周期, 可考虑借助第三方的数学

函数库实现, 同时也可以保证算法的权威性和高效性。开源库可考虑 Math.Net 或者 GUN-GSL 均可。商业库首选考虑调用 Matlab 函数库, Matlab 2006 及其以上的版本, 都具有支持.Net、COM 开发功能, 借助 MATLAB Builder NE 可将函数封装成\*.dll 形式, 以供调用[6]。但是注意系统中需要加入 MCR(Matlab Compiler Runtime)运行库文件, 以支持 Matlab 函数库的运行。

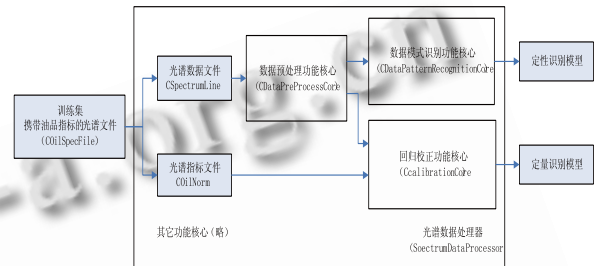


图7 光谱数据处理器的工作过程

4 系统工作数据存储区的设计

系统工作数据存储区是软件进行数据分析处理的场所, 它包含两个主要功能: 一是为当前光谱数据的处理分析提供数据容器 CDataWorkRegion; 二是为光谱数据的处理操作提供历史记录容器 CWorkRegionHistory, 即实时记录保存各步数据分析结果。

图8所示是一个数据处理过程, 只涉及了光谱数据预处理和数据模式识别, 如果要进行回归校正, 还需要将数据库中的标准数据送入数据处理器中。

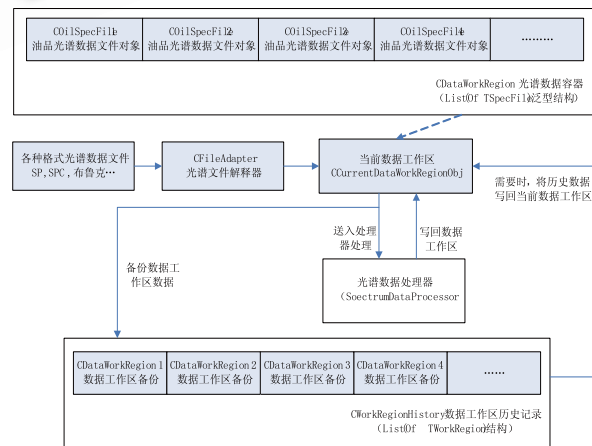


图8 光谱数据的存储、备份及恢复流程图

数据工作区 CDataWorkRegion 实现如下:

```
Public Class CDataWorkRegion(Of TSpecFile)
    定义以光谱文件为元素的链表
    Private mSpecFilesList As New List(Of TSpecFile)
    .....
End Class
```

CDataWorkRegion 是一个光谱文件的链表, 用于存储打开的若干条光谱文件数据。

数据工作区 CWorkRegionHistory 实现如下:

```
Public Class CWorkRegionHistory(Of TWorkRegion)
    定义以工作区为元素的链表
    Private mWorkRegionList As New List(Of TWorkRegion)
    .....
End Class
```

CWorkRegionHistory 是一个 CDataWorkRegion 的链表, 用于将每一步操作处理的数据存储起来, 便于随时恢复到以前的数据。

### 5 数据库设计

系统数据库主要存储标准数据, 数据库的用途主要有以下两个方面:

一是保存带有指标数据的红外光谱两侧数据。其主要应用是: 种类识别和质量指标的计算。

二是保存化学结构的编码信息。其主要应用是: 物质化学结构的分析。

本系统的开发, 主要侧重第一方面的应用, 下面以油品种类识别为例进行说明<sup>[7]</sup>。油料指标主要存储油品的质量指标信息, 如辛烷值、烯芳烃含量等; 光谱数据表主要存储光谱数据信息, 如光谱数据, 光谱点数等(图 9)。数据表字段含义见表 2 和表 3。

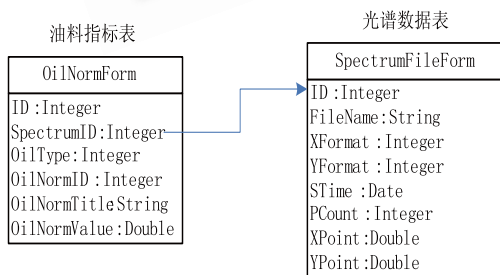


图 9 标表、光谱数据表对应关系图

表 2 光谱数据表(SpectrumFileForm)字段设计

字段名	说明
ID	编号(关键字段)
FileName	光谱文件名
XFormat	X 轴格式(波长或波数)
YFormat	Y 轴格式(吸光度或透过率)
STime	光谱采集时间
PCount	光谱数据点数
XPoint	X 轴数据(波数或波长数据)
YPoint	Y 轴数据(吸光度或透过率)

表 3 指标数据表(OilNormForm)字段设计

字段名	说明
ID	编号(关键字段)
SpectrumID	光谱 ID 编号
OilType	油品种类
OilNormID	油料指标 ID
OilNormTitle	指标名称
OilNormValue	指标值

### 6 红外光谱仪器的控制部分设计

仪器控制部分主要负责控制红外光谱仪的下位机部分, 与硬件电路之间通过 USB、IE488.2 GPIB 或者 RS232 接口进行通讯。当前, 部分的红外软件因为主要侧重化学计量学谱图分析处理功能, 而不涉及仪器控制功能。但是一个完整的开放红外光谱分析系统应该包括从仪器操控、数据采集到数据分析、光谱检索一系列完整的功能<sup>[8]</sup>。仪器控制部分功能主要有三个方面, 如图 10 所示。

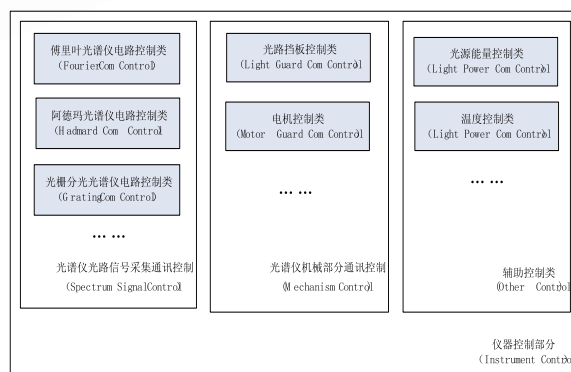


图 10 光谱仪通讯控制部分架构图

光谱仪器信号采集通讯控制(Spectrum Signal Control): 主要负责控制检测器电路, 检测光路信号,

并将其送给上位机,部分电路板底层即带有数据分析功能,但是目前的趋势是将数据处理任务越来越多的交给上位机程序来完成。

光谱仪机械部分通讯控制(Mechanism Control):主要控制光谱仪的机械部件,如光栅转动、光路挡板等。

辅助控制类(Other Control):主要负责光谱仪的一些周边辅助配件控制,如光源的能量、样品温度等。

当开放红外光谱分析系统框架建立后,只要为系统挂接不同的光谱仪器通讯控制类,系统便可完成对于光谱数据的采集任务。当然,这样做的前提是要获知特定光谱仪的底层硬件通信协议。

## 7 总结

以上介绍了一个开放红外光谱分析系统架构设计,开放的红外光谱平台具有以下优点:

① 光谱元数据采用多个类逐层组合构建而成,具有较好的兼容性,可以容纳不同公司的光谱数据格式,因此可以对不同来源的光谱数据进行分析处理,消除一般分析软件存在的数据格式不兼容问题;

② 有利于光谱系统的规范化,各个功能模块相互独立,便于根据不同的需求定制系统;

③ 光谱数据处理器的设计采用多核心设计,便于计量学算法的扩充和更新,同时,缩短新化学计量学算法开发时间;

(上接第4页)

部门共同协商,制定统一的标准,再由各个部门针对本部门所管理的图形对象分别进行建设、维护。这样既提高了数据的质量,又避免重复建设。对于专用专题空间数据,因仅主要涉及单个部门,所以这部分空间数据的建设和维护主要由该部门负责,可充分利用现有MIS中已有的数据资源。

基于铁路路网覆盖地域极为广阔、部门众多、业务复杂的特点,空间数据库的管理应采用“集中管理,分散维护”的原则。专题空间数据按业务特点分别由各业务部门负责维护,而在同一部门内部,按区域特点由基层单位进行数据更新和维护。同时,通过共同的平台、数据库标准和集中控制机制来统一管理空间数据库,实现资源共享。这样既保障大部分空间数据能够及时、有效地更新处理,也方便访问不同地区及不同部门的数据,将处理的高效性和高度的可访问性有机的结合起来。

④ 仪器控制部分允许添加和更换通讯控制类,因此开放红外光谱分析系统可与不同的光谱仪硬件进行集成,从而降低光谱仪器的研发成本。如果光谱仪硬件发生更新或变动,只需修改或更换相应的通讯控制类即可,光谱数据容器、数据处理器均不受影响。

## 参考文献

- 1 刘艺. Delphi 编程模式.北京:机械工业出版社,2005:25-73.
- 2 Balena F. Visual Basic 2005 技术内幕.北京:清华大学出版社,2006:83-146.
- 3 方利民,林敏. ICA 的近红外光谱分析软件的研制.中国计量学院学报,2010,21(1):42-46.
- 4 Beebe KR. Chemometrics: a practical guide. New York: John Wiley & Sons Ltd, 1998: 34-63.
- 5 Brereton RG. Chemometrics Data Analysis for the Laboratory and Chemical Plant. London: John Wiley & Sons Ltd, 2002:40-56.
- 6 褚小立,王艳斌. RIPP 化学计量学分析软件 3.0 的开发. 现代科学仪器,2009,(4):6-10.
- 7 杨岚. 近红外光谱专家系统的分析与设计[硕士学位论文]. 重庆:解放军后勤工程学院,2008.
- 8 段翔玉,许宝杰. 光谱测量系统应用软件设计. 北京信息科技大学学报,2009,24(2):75-77.

## 5 结论

铁路军事运输动态监控系统基础信息平台的实现方案主要是针对铁路军事运输信息化建设中统一编码、数据共享、数据传输等方面进行研究,实现基础数据的实时同步和更新。这样不仅为动态监控系统运输战备、运输计划、运输调度、运输统计四大功能模块提供可靠基础信息,同时在铁路军事运输信息化建设中起到了举足轻重的作用,成为整个铁路军事运输信息系统成败的关键。

## 参考文献

- 1 铁路军事运输信息系统需求分析报告.北京:中国人民解放军总后勤部军事交通部第一运输局,2007.
- 2 铁路军事运输信息系统概要设计报告.北京:中国人民解放军总后勤部军事交通部第一运输局,2007.
- 3 Robinson S.C#高级编程.北京:清华大学出版社,2005.