

高光谱数据降维及 压缩技术

陈雨时 编著

HEUP 哈尔滨工程大学出版社

高光谱数据降维及压缩技术

陈雨时 编著

 哈尔滨工程大学出版社

内 容 简 介

高光谱遥感能够同时获取目标的空间信息及光谱信息,从而以更高的可信度辨别和区分地物目标。但与此同时,高光谱数据带来了大数据量和高数据维。本书针对高光谱数据的降维及压缩展开论述,改进了高光谱数据压缩的几种方法。本书可供从事遥感信息处理领域的相关人员参考。

图书在版编目(CIP)数据

高光谱数据降维及压缩技术/陈雨时编著. —哈尔滨:
哈尔滨工程大学出版社,2014. 11
ISBN 978 - 7 - 5661 - 0937 - 8

I. ①高… II. ①陈… III. ①谱分辨率 - 光学遥感 -
研究 IV. ①TP72

中国版本图书馆 CIP 数据核字(2014)第 258258 号

出版发行 哈尔滨工程大学出版社
社 址 哈尔滨市南岗区东大直街 124 号
邮政编码 150001
发行电话 0451 - 82519328
传 真 0451 - 82519699
经 销 新华书店
印 刷 黑龙江省地质测绘印制中心印刷厂
开 本 787mm × 960mm 1/16
印 张 7.5
字 数 136 千字
版 次 2014 年 11 月第 1 版
印 次 2014 年 11 月第 1 次印刷
定 价 16.90 元
<http://www.hrbeupress.com>
E-mail: heupress@hrbeu.edu.cn

前 言

遥感技术自从问世以来,就得到了迅猛的发展,遥感技术的理论发展及其广泛应用是 20 世纪具有标志性的科学成就之一。在遥感技术的发展中,有一个重要的发展方向——更高的分辨率。这包括:更高的空间分辨率、更高的光谱分辨率和更高的时间分辨率,而高光谱遥感就是更高光谱分辨率的一种实现。

高光谱遥感能够在纳米级的光谱分辨率上就待观测的地物同时进行上百个谱带的成像,这种遥感方式实现了地物空间信息和光谱信息的同步获取,具有重大的理论意义及应用价值。高光谱遥感为每个空间像素提供数百个近似连续的窄谱带,从而形成一条完整而近似连续的光谱曲线,精确地光谱信息可以以较高的可信度辨别和区分地物目标。因此,高光谱遥感在环境监测、精准农业、洪涝灾害监测、国土资源调查,尤其是在侦察、反伪装、战场态势评估、侦察打击一体化系统中具有独特的优势。但与此同时,上百个谱带的成像也带来了海量数据,而这给存储和传输带来了很大的压力。高光谱数据本身的特点决定了其不同于普通数据,因此,如何针对高光谱数据本身的特点,进行高光谱数据的降维与压缩是遥感信息处理领域中的一个重要研究内容,而本书针对这一问题展开相关的论述。

本书的出版可以给广大读者提供关于高光谱遥感中的降维以及压缩等数据处理手段的一些基本性概念,掌握高光谱遥感降维以及压缩的基本方法。期望本书的出版能为高光谱遥感的数据处理增加一些知识及技术,促进高光谱遥感技术的交流与发展。

本专著由国家自然科学基金资助(61301206,61471148),同时本书在撰写过程中参考或引用国内外一些专家学者的论著,在此一并表示感谢!

由于时间及水平有限,本书的内容在理论、应用等方面存在着不足,错误之处在所难免,望读者能够不吝赐教。

陈雨时

2014 年 9 月 1 日

目 录

第1章 绪论	1
1.1 高光谱遥感技术简介	2
1.2 高光谱数据压缩技术发展综述	5
第2章 高光谱数据特性分析	9
2.1 高光谱数据介绍	9
2.2 高光谱数据的空间相关性分析	13
2.3 高光谱数据的谱间相关性分析	16
2.4 高光谱数据的数据维分析	19
2.5 高光谱数据的信息量分析	21
2.6 高光谱数据的非线性分析	23
第3章 高光谱数据降维技术	26
3.1 引言	26
3.2 基于主成分分析的高光谱数据降维	29
3.3 基于子空间分解的高光谱数据降维	31
3.4 高光谱图像波段选择技术	35
3.5 基于核的非线性降维技术	43
第4章 基于预测的高光谱数据压缩	54
4.1 应用双向预测与自回归双向预测压缩高光谱数据	54
4.2 基于线性模型的高光谱数据最优预测	60
4.3 应用基于线性模型的最优预测压缩高光谱数据	64
第5章 基于矢量量化的高光谱数据压缩	71
5.1 矢量量化技术及其在高光谱数据压缩中的应用	71
5.2 应用基于特征选择的矢量量化技术压缩高光谱数据	78
5.3 应用基于特征变换的矢量量化技术压缩高光谱数据	86
第6章 基于感兴趣信息的高光谱数据压缩技术	93
6.1 基于感兴趣区域的高光谱数据压缩	94
6.2 基于感兴趣谱带的高光谱数据压缩	100
参考文献	109

第1章 绪 论

遥感技术是从20世纪60年代发展起来的综合性科学技术,遥感技术应用各种传感器(如可见光、红外、微波等)对目标对象进行观测,接收并记录电磁波信号,根据电磁波与地表物体的作用机理及对探测目标的电磁特性进行分析,进而获取物体特征性质及其变化^[1]。遥感技术与电子技术、光学技术、计算机科学、信息科学、地球科学等众多科学技术紧密相关,是现代科学技术的一个重要组成部分。遥感技术提供了新的数据源和新的探测方式,随着遥感技术的发展,遥感技术在地理学、地质学、生态学、环境科学、大气科学、海洋学及军事侦察等学科领域得到了广泛应用^[2,3]。

遥感技术的飞速发展,使人类能够获取更高空间分辨率、更高光谱分辨率、更高时间分辨率的遥感数据。与此同时,人类也能够不同平台上以更加广阔的电磁波谱范围对地观测,从全色遥感到多光谱以及高光谱遥感,从光学遥感到红外和微波遥感^[4]。探测信息的手段极其多样化,使得获取的多维/多尺度/多时相的数据信息空前丰富。近年来,遥感探测技术得到了长足的发展,在数据获取、辐射定标、光谱重建、数据处理等技术方面都有了突破。目前,已经呈现出“三高”(高空间分辨率、高光谱分辨率和高时间分辨率)和“三多”(多平台、多传感器和多角度)的局面,一个多层次、多角度、全天候、多分辨率互补的全球对地观测系统正在逐步形成。遥感对地观测技术的这些新特点为后续的进一步应用带来了更多的可能性,使其原来许多悬而未决的问题得以解决^[4]。

在遥感技术的发展过程中,更高的光谱分辨率是一个重要趋势。高光谱数据能够在纳米级的光谱分辨率上对地物进行上百个谱带的成像,实现了地物空间信息和光谱信息的同步获取,是遥感技术在光谱分辨率上的一个重大飞跃^[5]。更高的光谱分辨率对于高光谱数据的应用至关重要,但与此同时也带来了海量数据,而这给存储和传输带来了很大的压力。高光谱数据本身的特点决定了其不同于普通数据,因此如何针对高光谱数据本身的特点,进行高光谱数据的降维与压缩是遥感信息处理领域中的一个重要研究内容,而本书针对这一问题将展开相关的论述。

1.1 高光谱遥感技术简介

遥感技术的广泛应用促使着遥感技术的进一步发展,而日益广泛地应用也给遥感技术的发展方向提出了新的要求。在 20 世纪 80 年代,随着光学技术与传感器技术的不断进步,具有高光谱分辨率的成像光谱仪出现了。高光谱遥感技术是一种新型的遥感技术,其突出特点是将地物的光谱和相应的空间数据有机结合,形成了新的遥感系统,它的问世是遥感技术的一个重大飞跃^[6]。在图 1-1 中展示了高光谱数据的基本概念。高光谱遥感技术能够在电磁波谱的紫外、可见光、近红外和中红外区域,在对目标的空间特征成像的同时,对每个空间像元经过色散形成数十个甚至数百个窄波段以进行近似连续的光谱覆盖,获得许多非常窄且光谱近似连续的数据,具有获得高光谱分辨率数据的能力。高光谱数据的光谱分辨率在 10 nm 左右,与通常的多光谱数据相比,高光谱数据的光谱分辨率要高得多。高光谱遥感为每个单元提供数十个至数百个窄波段(通常波段宽度约为 10 nm)的光谱信息,能产生一条完整而近似连续的光谱曲线。因此,高光谱遥感可以同时处理高光谱数据中目标的空间特征和光谱特征,从而以较高的可信度辨别和区分地物目标。在环境监测、精准农业、洪涝灾害监测、国土资源调查,尤其是在侦察、反伪装、战场态势评估、侦察打击一体化系统中具有独特的优势,也有着广泛的应用前景。

高光谱遥感在许多相互邻接的狭窄谱带收集地面的数据。这样,不仅可以得到多个狭窄谱带的地面数据,而且还可以得到数据上每个像素的光谱曲线数据。根据获得的每个谱带的数据信息,可以研究其空间结构;根据获得的光谱信息,可以研究其物理特性。在地面数据二维空间信息的基础上,增加第三维(光谱)信息,大大提高了从遥感数据判别与获取地面信息的能力。空间分辨率的提高可以使数据看起来更加清晰,有助于对数据的直接解译;而远程探测得到的光谱信息作为描述地物特征的另一种方式,可以实现对地物的辨识。而且光谱信息直接与目标物质结构和性质有关,特别是对于矿物识别、植被的精细分类、海洋水色定量监测以及军事上对伪装的辨认等,从光谱的角度远比空间的角度更适合。图 1-1 右侧的几条曲线分别表示了土壤、水、植被特征的光谱曲线,可以看出不同物质的光谱曲线差别很大。根据目标的光谱曲线可以对其进行分类和识别。许多物质在单波段的遥感数据中灰度值是一样的,无法区分,但在高光谱数据中可以根据其光谱曲线的不同将其分开,这也是成像光谱仪的优势所在。更进一步的研究表明,许多地表物质的吸收特征在吸收峰深度一般处的宽度为 20 ~ 40 nm。而成像光谱系统获得的波段宽度一般都在 10 nm 以内,因此这种数据能以足够高的光谱分辨率区分出

那些具有诊断性光谱特征的地表物质,这也是高光谱技术用于探测甚至识别地物的物理基础。

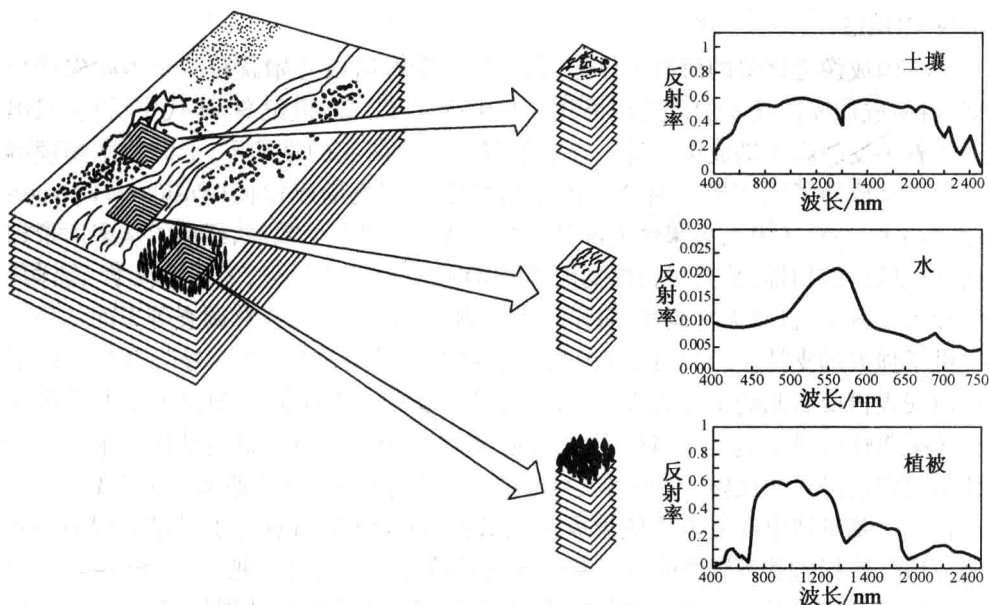


图 1-1 高光谱数据的概念示意图

具有“图谱合一”特性的成像光谱技术是遥感领域迈出的崭新一步,也是遥感技术发展到新阶段的重要标志。正是因为高光谱在众多领域的独特优势,世界各国竞相发展高光谱遥感技术。世界上越来越多的国家竞相投入到成像光谱仪的研制和应用中来。自从美国喷气推进实验室(Jet Propulsion Laboratory, JPL)在 NASA (National Aeronautics and Space Administration) 支持下首先对成像光谱仪进行概念设计以来,高光谱遥感技术得到了极大的关注。世界上的很多国家都投入了大量的人力物力进行成像光谱仪的研制和应用。从 20 世纪 80 年代至今已经研制了三代高光谱成像光谱仪,其中最为成功的当属 AVIRIS (Airborne Visible and InfraRed Imaging Spectrometer)。它能同时收集从可见光到红外谱带范围内的光谱信息,将波长从 400 nm 到 2 500 nm 的光谱信息表示为 224 个近似连续的谱带,光谱分辨率大约为 10 nm,其空间分辨率为 20 m × 20 m。经过 20 世纪 80 年代的起步和 90 年代的发展,一系列高光谱成像系统在国际上研制成功并获得广泛地应用。目前,高光谱遥感一方面由实验研究阶段逐步转向实际应用阶段,另一方面则由以航空应用为主开始转向以航空和航天应用相结合的阶段,同时星载成像光谱仪也已研制

成功。美国发射的 EO-1 地球观测卫星上搭载的高光谱成像仪 HYPERION 代表着航天成像光谱仪的先进水平。除此之外,欧空局研制了紧凑型高分辨率成像光谱仪 CHRIS。

我国成像光谱仪的研制工作起步较早,从跟踪阶段开始就跻身于国际先进行列,目前成像光谱技术的研究成果几乎与国际同步。20 世纪 80 年代中后期,我国开始着手发展高光谱成像系统。在自然科学基金研究的推动下,上海技术物理研究所研制成功了我国第一台 224 谱带的推帚式高光谱成像仪(Pushbroom Hyperspectral Imager, PHI),效果令人满意。我国成像光谱仪的发展,经历了从多谱带扫描到成像光谱扫描,从光机扫描到面阵 CCD 扫描的发展过程,为陆地科学、海洋科学和大气科学、自然灾害监测、海洋环境监测、农作物估产、森林调查等方面的研究提供了强大的支持。2007 年,我国发射了首颗探月卫星“嫦娥一号”,其中搭载了由西安光机所研制的干涉式成像光谱仪,其主要技术指标幅宽 25.6 km,月表像元分辨率 200 m,光谱范围 0.48 ~ 0.96 μm ,谱段数为 32。这标志着我国在成像光谱仪的研究已经日趋成熟。“神州三号”飞船搭载的 CMODIS,“嫦娥一号”探月卫星上搭载了基于傅里叶变换的航天成像光谱仪,以及环境与减灾小卫星搭载的高光谱成像仪都标志着我国高光谱遥感技术已经跻身国际先进行列。

高光谱遥感要真正获得实际应用,必须发展相关的数据处理技术。因此,国际上对于高光谱遥感数据处理的研究日益关注。IEEE 的《地球科学与遥感》《信号处理》都集中出版了关于高光谱数据分析与应用方面的专集,包括高光谱的诸多热门课题,如数据压缩、分类聚类、检测识别、数据融合和仿真建模等^[7]。鉴于高光谱技术的重要性,美国、日本、加拿大、澳大利亚等国家对高光谱数据处理投入了相当大的人力物力,取得了一定的成绩,这也给我国的科研工作者提出了一个挑战。我国在硬件方面已经取得了很大成绩,但是国内的高光谱数据理论研究和应用开发相对滞后。国内高光谱数据处理方面大多集中于成像光谱仪的定标及一些辐射校正研究,直到最近几年才有一些学者将高光谱遥感应用于矿物识别、植被监测研究,又相继进行了基于神经网络的分类方法研究。高光谱成像技术在军事上有很高的应用价值,因而军用卫星上采用这种遥感器的趋势正在快速增长。因此,当前的紧要任务是进行相应的理论研究和应用开发,使得耗费巨资获得的高光谱数据具有真正的实用价值,为我国经济建设和国防建设服务。

由于成像光谱仪的光谱通道数多,光谱分辨率高,因而它的数据量是巨大的。例如,标准的 AVIRIS 一个谱带需要 $512 \times 614 \times 2$ 字节,每幅数据有 224 个谱带。这样一来,一幅标准的 AVIRIS 需要占用约 140 M 字节。高光谱的海量数据为存储和传输带来巨大的压力。

随着成像光谱仪的发展,数据的空间分辨率、光谱分辨率和时间分辨率越来越高,这使得数据急剧增长。因此,如何对高光谱的海量数据进行有效地压缩以适应遥感技术的发展就成了一个突出的问题。由于成像光谱数据的特殊性,利用常规的数据压缩处理方法效果不是很理想。有必要研究一种基于成像光谱特征的数据压缩方法,在压缩过程中最大限度地利用高光谱数据的特征,最终提高高光谱数据的压缩效果。

由于高光谱遥感技术在当前乃至今后一段时间都将处于遥感的前沿,因此国内外对它的研究都极为重视。如何在压缩中尽可能地保留高光谱数据丰富的光谱信息和空间信息,研究高光谱数据的压缩新方法,是高光谱遥感的重要研究内容之一。

1.2 高光谱数据压缩技术发展综述

经过数十年的发展,数据压缩取得了相当大的进展并日趋成熟,也形成了相应的标准^[8]。但是众多的数据压缩算法都是针对二维数据的,对于高光谱这样的特殊的三维数据并不能完全适用。另外,自从成像光谱仪研制成功以来,高光谱数据的海量数据已经成为制约高光谱遥感发展的一个因素。因此高光谱数据压缩作为数据压缩领域内的一个新问题,没有形成成熟或标准的压缩方法,因此该问题有着一定的研究价值和较高的应用价值。

高光谱数据作为一种新型的遥感数据有着本身的特点,其中最重要的就是在获得空间信息的同时所获得的光谱信息,这样一来,高光谱数据不仅有空间的相关性,而且具有光谱之间的相关性。因此,高光谱数据的去相关过程大致分为两个过程,即光谱去相关和空间去相关,如图1-2所示。关于数据的空间相关性研究得比较多,形成了众多的算法和标准。而对于光谱之间的相关性研究得比较少,是高光谱数据压缩的一个特别的地方,也是高光谱数据压缩的关键所在。

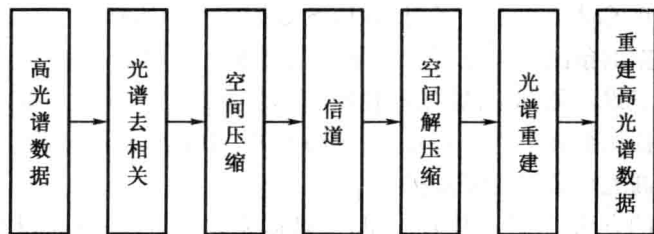


图1-2 高光谱数据压缩与重建框图

近年来关于高光谱数据压缩的研究在国际上不断增多,许多针对高光谱数据的压缩技术也相继出现,鉴于高光谱数据的光谱去相关是高光谱数据压缩的特色所在,接下来以此为线索对高光谱压缩技术进行简要综述。

1.2.1 基于预测技术的高光谱数据压缩

基于预测的压缩方法提出的时间较早,获得了广泛地应用,已经成为无损压缩编码的一个重要方向。基于预测的方法特别适用于数据连续变化的自然物理过程,该方法直接探索光谱维上谱带与谱带之间的相关性。其基本原理是:任何一个像素可以由其周围的像素(包括空间相邻和谱带相邻)预测,而由预测产生的残余误差比较容易压缩^[9]。

基于预测技术的高光谱数据压缩先在光谱维上进行去相关,然后再进行空间去相关。高光谱数据在光谱维上有着相当强的相关性,一个谱带可以由其相邻的谱带预测,而其产生的去相关以后的残余误差往往较小,比较容易压缩。基于预测技术的压缩是高光谱数据压缩的主流方法之一^[10,11]。

基于差分脉冲调制的预测是最基本的预测方法,通过对预测系数的选择,使预测器输出的预测值与当前的实际值的误差最小。DPCM 的预测系数由数据的统计特征决定。DPCM 的改进算法较多,如最优线性预测、自适应预测、双向预测、三维组合预测等。目前常用的方法是用谱间 DPCM 预测方法与谱内压缩算法相结合进行高光谱数据的压缩^[12]。

假设 $x_{i,j,n-1}$ 是第 $n-1$ 个谱带数据在 (i, j) 处的像素灰度值, $x'_{i,j,n|n-1}$ 是由它预测得到的第 n 个谱带数据在同一空间位置的像素灰度值。将这个预测器应用到数据上的每一个像素,就可以由第 $n-1$ 个谱带数据预测第 n 个谱带数据,然后对残余误差数据 $\varepsilon_n = x_n - x'_{n|n-1}$ 进行进一步的处理便可实现高光谱数据的压缩。可以使用自适应的 DPCM 进行谱带间的去相关,然后对残余误差数据采用变长度压缩 VLC (Variable Length Coding),实现了对 AVIRIS 数据的无损压缩。

高光谱数据的预测方法也不仅仅限于光谱维上的预测,国内外的学者提出了包含空间维的三维预测方法,譬如张晓玲和孙兰荪提出了一种三维自适应预测的高光谱数据无损压缩算法^[13], Aiazzi 和 Alba 应用 3D 模糊预测处理,对预测误差采取算术编码的高光谱数据压缩方法^[14]。

基于预测的方法,原理上比较简单,实现起来比较方便,因而得到了广泛地应用。但是相对于其他方法,在取得同样压缩效果的情况下采用该方法压缩比较低。

1.2.2 基于变换的高光谱数据压缩

另一类高光谱数据压缩方案是由基于正交变换(Orthogonal Transform)的技术组成的,其主要原理是在光谱维上通过变换重新组织数据,使数据能量相对集中在较少的几个谱带,而其他的谱带只具有很小的能量,然后进行空间的正交变换。在变换之后进行相应的量化,通过抑制能量小的系数获得原始数据的稀疏表示,从而实现数据的压缩^[15,16]。

在光谱维上典型的方法是利用 K-L(Karhuen-Loene)变换、DCT 变换和小波变换以及它们的改进方案,其中 K-L 变换在理论上是最佳的^[17]。首先在谱带上利用离散 K-L 变换去除谱带间的相关性,然后再利用标准技术在每个谱带内对残余误差数据在空间域去相关继续压缩。鉴于 JPEG 2000 在二维数据压缩中的优点,该压缩方法也可和基于 K-L 变换的光谱去相关方法结合起来^[15]。

K-L 变换的主要缺点是变换的基函数需要计算与原始数据相关的协方差矩阵和特征矢量,这些计算在应用中通常是较难实现的。因此,人们往往利用 DCT 变换来替代 K-L 变换实现谱带的去相关,而利用 DCT 的主要优点是其基函数固定不变,不用考虑其原始数据。

考虑到小波变换的优点,人们也利用小波变换来压缩高光谱数据。基于整数小波变换的压缩可以实现高光谱数据的无损压缩。基于整数小波变换的压缩可以实现高光谱图像的无损压缩,赵春晖^[18]、罗欣等^[19]分别提出了基于提升格式和基于可逆整数变换的高光谱图像压缩方法。

高光谱图像是一种三维图像,因此基于三维变换的压缩方法具有天然的优势^[20]。Wu 提出了三维小波变换和嵌入零树编码^[21]。Kim^[22]和 Yan^[23]分别提出了基于低比特率和改进的三维 SPIHT 算法进行高光谱图像压缩,试验表明合适的有损压缩也是可行的。不同谱带间的相关性是不同的,基于分组的压缩就显得更加合理,张培强等提出了波段分组的 3D-SPIHT 高光谱图像无损压缩算法^[24]。另外,为了满足高光谱图像的渐进传输需要,Barbara Penna 提出了一种基于 JPEG 2000 的三维渐进编码^[25]。高光谱图像和立体视频图像存在一定的相似性,因此视频图像的压缩算法可以作为借鉴^[26]。

基于变换的压缩一直是图像压缩领域内的主流方法,基于变换的压缩方法在高光谱图像压缩中也有着很重要的地位。

1.2.3 基于矢量量化的高光谱数据压缩

第三类高光谱数据的压缩方法主要由基于矢量量化 VQ (Vector Quantization) 的技术及其改进方案组成。在矢量量化中, 每条光谱曲线被当作一个输入矢量。VQ 的原理是直接对光谱曲线进行量化, 从而达到去相关的目的^[27]。

基于 VQ 的方案理论上是以信息的高阶熵为下限, 能在高压缩率和平均最小失真间获得最佳折中, 压缩性能接近理论上的数据最小压缩熵^[27]。

VQ 的主要问题是其较高的压缩复杂性, 尤其是码书产生的过程, 其计算量随着矢量维数的增加而呈指数增长。因此, 在保持较高压缩性能的前提下, 人们也试图在减小 VQ 复杂性, 提出相应的快速算法。例如 Qian 提出了一种快速的矢量量化算法来提高码书的产生效率, 该算法无需进行全搜索, 大大降低了运算的复杂度^[28]。

近年来, 加拿大航天局的 Qian 提出了两种很有代表性的基于矢量量化的基于位的超光谱遥感数据无损压缩算法——SAMVQ (Successive Approximation Multi-stage Vector Quantization) 和 HSOCVQ (Hierarchical Self-Organizing Cluster Vector Quantization), 取得了很好的压缩效果^[29]。基于矢量量化的压缩算法一般是有损压缩, 以 Qian 为代表的学者详细分析了基于 VQ 压缩算法对高光谱数据应用的影响, 实验表明合理的基于 VQ 的压缩算法对后续应用的影响有限^[30]。

VQ 的主要问题是其较高的压缩复杂性, 尤其是码书产生的过程。因此, 在保持较高压缩性能的前提下, 人们也试图在减小 VQ 复杂性, 提出相应的快速算法, 使其朝着实用化的方向发展。矢量量化的快速算法这一领域直到现在还是很活跃的, 不断地有新的快速算法出现。

考虑到高光谱数据本身的特点, 高光谱数据的压缩一般从光谱维去相关入手, 在去除高光谱数据谱带间相关性的基础上, 采用目前的数据压缩技术实现空间的压缩。根据光谱维上去相关方式的不同, 压缩分为基于预测的压缩、基于变换的压缩和基于矢量量化的压缩三大类。高光谱数据作为一种新型的遥感数据源在越来越多的领域得到应用, 针对高光谱数据压缩的各种方法至今还在不断地被提出。

第2章 高光谱数据特性分析

作为一种新型遥感方式的高光谱数据通常包含上百个谱带,这使得高光谱数据有别于普通的遥感数据,也有别于多光谱数据。高光谱数据的高数据维和大数据量的特点使高光谱数据有其本身的特殊性,由此带来高光谱数据在处理上的特殊性,也使得以往成功的遥感数据压缩方法不再适用。因此,必须对高光谱数据进行全面的分析,充分了解高光谱数据的特性,以便找到针对高光谱数据的合适的压缩方法。

2.1 高光谱数据介绍

在进行高光谱数据特性分析之前,先简单介绍一下在本书中所使用的高光谱数据。在众多的高光谱成像光谱仪中,其中最为成功的当属 NASA/喷气推进研究室(JPL)推出的机载可见光/红外成像光谱仪 AVIRIS。

AVIRIS 由 NASA/JPL 实验室开发,装载于 NASA ER-2 飞机上,飞行高度大约在海平面以上 20 km,视场 30° ,瞬时视场 1 毫弧度,地面分辨率 $20\text{ m} \times 20\text{ m}$ 。这是一种可以同时采集由可见光到红外(400 ~ 2 500 nm)范围内 224 个谱段信息的推帚式成像传感器,光谱分辨率约为 10 nm。四个偏轴双通施密特光谱仪用光纤从前光学系统接收到入射光,四个线性阵分别在 $0.4 \sim 0.7\text{ }\mu\text{m}$, $0.7 \sim 1.2\text{ }\mu\text{m}$, $1.2 \sim 1.8\text{ }\mu\text{m}$, $1.8 \sim 2.5\text{ }\mu\text{m}$ 范围内提供高灵敏度探测结果。1998 年以来,它也工作在低空 Twin Otter 飞机上,提供 2 ~ 4 m 空间分辨率。

AVIRIS 可以用“图像立方体”来形象描述,其中二维表征空间位置,另外一维表征光谱波长,因此高光谱数据中包含了丰富的空间和光谱信息。在目前的应用中,典型的高光谱数据是由 AVIRIS 获取的。一幅标准的 AVIRIS 图像数据占据 512 行,每行 614 个像素,共计 224 个谱段的图像立方体,如图 2-1 所示。每个谱带的宽度(光谱分辨率)在 9.7 ~ 12.0 nm 之间。

高光谱数据数据的不同表示方式强调了不同的信息,适合于不同的目的要求。高光谱数据所携带的信息一般有这样两种显示方式,如图 2-2 所示。

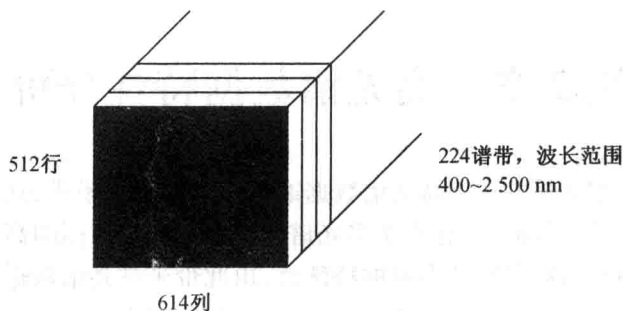


图 2-1 AVIRIS 图像立方体示意图

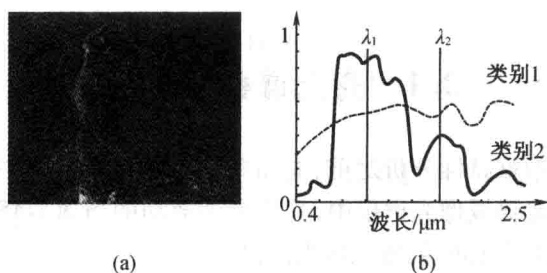


图 2-2 高光谱数据的两种显示方法示意图

(a) 图像空间; (b) 光谱空间

图像空间:对于人类的视觉系统而言,这是最自然、最直观的表达方式。对于每一固定的波长,相当于地面景物的照片。这种方式只能分别看到 n 幅灰度图像或者三个波段的彩色合成图像,波段之间的关系也不是很明显。但是图像表示对于纵览地物之间的相互位置关系是很有用的。

光谱空间:光谱响应作为波长的函数,反映了电磁波能量随波长的变化情况。如果测得的光谱响应曲线中包含了辨识地物所需的信息,那么光谱空间是一种简单而有效的表示方法,提供了直接用于解译的光谱信息。

在图 2-3 中给出了 AVIRIS 的 Cuprite 的一个谱带图像。在图 2-4 中给出了另一幅 AVIRIS 高光谱数据 Sandiego 的一个谱带图像。



图 2-3 高光谱数据 Cuprite 示意图



图 2-4 高光谱数据 Sandiego 示意图

图 2-3 中的 Cuprite 图像数据占据 512 行, 每行 614 个像素, 共计 224 个谱段, 图像的空间分辨率是 20 m。图 2-4 中给出的是另一幅 AVIRIS 高光谱数据 Sandiego。该图像数据占据 400 行, 每行 400 个像素, 共计 126 个谱段, 图像的空间分辨率是 3.5 m。

AVIRIS 作为“图像立方体”还可以从另外一个角度看。AVIRIS 在对目标的空

间特征成像的同时,对每个空间像元经过色散形成几十个乃至几百个窄波段以进行连续的光谱覆盖,从而形成谱分辨率达到纳米数量级的遥感数据。因此,每个空间像元都形成一个光谱矢量,如图 2-5 和图 2-6 所示。在图 2-5 和图 2-6 中,横坐标是谱带的编号,而纵坐标是像元反射率值。

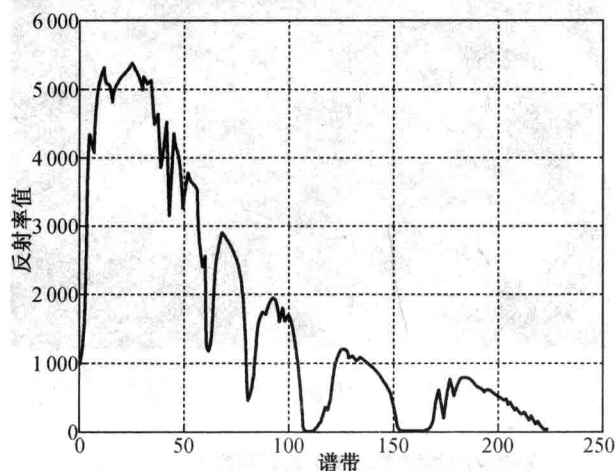


图 2-5 高光谱数据 Cuprite 光谱曲线图

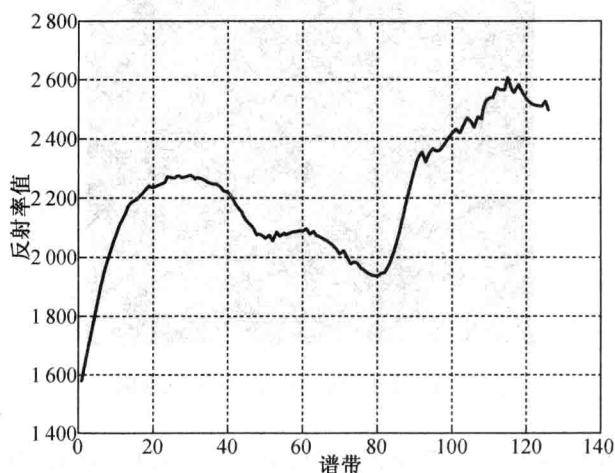


图 2-6 高光谱数据 Sandiego 光谱曲线图

在本小节中,介绍了书中用到的高光谱数据。为了能有直观的印象,分别显示了合成的高光谱数据和光谱曲线。