

文章编号: 0258-7025(2009)11-2983-06

基于 FPGA 的高光谱图像奇异值分解降维技术

何光林 彭林科

(北京理工大学机电工程与控制国家级重点实验室, 北京 100081)

摘要 为了解决高光谱图像维数高、数据量巨大、实时处理技术实现难的问题,提出了高光谱图像实时处理降维技术。采用奇异值分解(SVD)算法对高光谱图像进行降维,又在可编程门阵列(FPGA)芯片中针对这一算法划为自相关模块、特征求解模块、特征提取模块和降维实现模块 4 个模块进行编程实现、仿真和验证。仿真结果表明,高光谱图像降维后数据量为降维前的 1/3,而降维后的分类像素点误差为 0.2109%,证明了奇异值分解算法进行高光谱图像降维算法的有效性。

关键词 光谱学;高光谱图像;数据降维;奇异值分解;可编程门阵列

中图分类号 TP751.1 **文献标识码** A **doi:** 10.3788/CJL20093611.2983

FPGA Implement of SVD for Dimensionality Reduction in Hyperspectral Images

He Guanglin Peng Linke

(National Laboratory for Mechatronic Engineering and Control, Beijing Institute of Technology, Beijing 100081, China)

Abstract To solve hyperspectral image's problems of the high dimensionality, the huge amount of data, and the real-time solution and so on, a real-time hyperspectral dimensionality reduction method is brought forward. Based on singular value decomposition (SVD) method, hyperspectral dimensionality is reduction, and finish the design of the chip system with top-down method. The chip system is divided into autocorrelation module, SVD module, feature extraction module and dimensionality reduction module. It completes the design, simulation and verification of these modules. The results indicate that the hyperspectral image reduced to 1/3, classification error is only 0.2109 percent after the dimensionality reduction. All of this show, the SVD method for hyperspectral dimensionality reduction is effective.

Key words spectroscopy; hyperspectral image; data dimensionality reduction; singular value decomposition; field programmable gate array

1 引 言

高光谱图像具有波段多、光谱分辨率高、相邻波段相关性大、数据量庞大等特点,因而带来信息冗余度高、数据存储所需空间大、处理时间长^[1~3]等问题。因此,高光谱图像在处理、分析之前进行图像数据降维是非常重要的。L. O. Jimenez 等^[4]从理论上对数据降维的可行性进行了论证。

现有降维方法可分为 3 类:利用波段选择进行

降维,通过划分数据源进行降维和利用特征值提取进行降维^[5]。利用波段选择和划分数据源都达不到理想降维效果,因此近年一些采用特征提取方法的降维方式在不断发展。A. A. Green 等^[6~10]发展了主成分分析方法(PCA);Jia Xiuping 等^[3]提出了分段主成分变换(SPCT)的方法;L. O. Jimenez 等^[11,12]提出基于投形寻踪法(PP)和奇异值分解(SVD)算法等。这些算法各有特点,主成分分析法

收稿日期: 2008-11-18; 收到修改稿日期: 2009-01-14

基金项目: 国家部委项目(51305010102)资助课题。

作者简介: 何光林(1974—),男,博士,副教授,主要从事高光谱遥感图像处理及其应用、智能目标检测等方面的研究。

E-mail: heguanglin@bit.edu.cn

和奇异值分解法相对简单,分段主成分分析法和基于投影寻踪法的算法则很复杂,精度也稍高。本文利用可编程门阵列(FPGA)实现高光谱图像降维,则需选择便于硬件实现的奇异值分解算法。

高光谱图像的实时处理是近年来国内外研究的热点,国内还没有在高光谱图像实时降维处理领域有太多的研究。随着FPGA技术的发展,为开发数据吞吐量巨大的高光谱图像数据实时处理芯片创造了条件。田纳西大学的Hongtao Du等^[13]提出用FPGA实现高光谱的图像降维,但是没有实现实时处理的功能。本文结合现有降维方法和FPGA的特点,选取XILINX公司Virtex-5系列芯片,对基于奇异值分解算法进行模块化划分,利用VHDL语言编程实现高光谱图像数据的实时降维处理。

2 奇异值分解算法基本原理

奇异值分解算法采用集合总体信息的概念,以某种代数或集合准则最优化技术对一个数据阵的结构进行描述和简化。其原理^[5,14,15]为

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} = [\mathbf{x}_{(1)} \quad \mathbf{x}_{(2)} \quad \cdots \quad \mathbf{x}_{(p)}], \quad (1)$$

式中 \mathbf{X} 为高光谱图像数据矩阵, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 表示 n 个像素点在不同波段下的光谱特性, $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(p)}$ 表示 p 个波段下的不同图像,奇异值分解的算法步骤为:

1) 计算矩阵 $\mathbf{G} = \mathbf{X}^T \mathbf{X}$;

2) 求出 \mathbf{G} 的全部特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ 和对应的标准正交特征向量 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ 为数据阵 \mathbf{X} 在 \mathbf{R}^p (p 维实数空间)上的主轴(主分量);

3) 计算 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 的主成分 $\mathbf{y}_{(1)} = \mathbf{X}\mathbf{u}_1, \mathbf{y}_{(2)} = \mathbf{X}\mathbf{u}_2, \dots, \mathbf{y}_{(p)} = \mathbf{X}\mathbf{u}_p$ 它们是 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 分别第1,第2,……,第 p 个主轴上的坐标;

4) 对于给定的主成分变差贡献比例 q_0 ,求满足条件的正整数 l

$$\frac{\sum_{j=1}^l \lambda_j}{\sum_{j=1}^p \lambda_j} < q_0 \leq \frac{\sum_{j=1}^{l+1} \lambda_j}{\sum_{j=1}^p \lambda_j}, \quad (2)$$

则取前 l 个主成分就足以描述原始数据阵;

5) 求出用前 l 个主成分近似表示原始数据阵的系数矩阵 $\mathbf{C} = (c_{ij})$,其中 \mathbf{C} 的第 i 行是用前 l 个主成分逼近 $\mathbf{x}_{(i)}$ 的线性组合系数。

用数据拟合的观点来看奇异值分解算法,就是在使信息损失尽可能少的前提下,降低原始数据阵

\mathbf{X} 的维数。采用雅可比过关法实现奇异值分解算法的特征提取。

雅可比法的基本思想^[16]为:对于任意实对称矩阵 \mathbf{A} ,只要能求得正交矩阵 \mathbf{U} ,使得 $\mathbf{U}^T \mathbf{A} \mathbf{U}$ 为对角矩阵 \mathbf{D} ,就可得到 \mathbf{A} 的所有特征值和特征向量。但在雅可比法中,每次变换前需遍历所有非对角线元素,取绝对值最大的元素,耗时多,不利于实时处理。雅可比过关法对此做了改进。

首先计算对称矩阵 \mathbf{A} 的所有非对角线元素平方和的平方根,即

$$E = \sqrt{2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij}^2}, \quad (3)$$

然后设置第一道关口 $r_1 = E/n$,对 \mathbf{A} 中非对角线元素进行逐行(逐列)扫描。若 $|a_{ij}| < r_1$,则让其过关,否则旋转变换将 a_{ij} 化为零。矩阵 \mathbf{A} 中所有非对角线元素都过了第一道关口后,再设置第二道关口 $r_2 = r_1/n = E/n^2$,重复上述过程。经一系列关口 r_1, r_2, \dots ,直到某个关口满足条件为 $r_n < \epsilon$ 止,其中 ϵ 为预先设定的精度要求。

3 基于FPGA的奇异值分解降维技术

结合奇异值分解数据降维算法原理,把算法按模块化划分法分别进行编程实现、仿真和验证。基于FPGA芯片的奇异值分解降维技术分为4个模块:自相关模块、特征求解模块、特征提取模块和降维实现模块。

3.1 自相关模块设计

自相关模块主要工作为矩阵乘法运算。以机载可见及红外成像光谱仪(AVIRSI)的标准高光谱图像为例,其数据量大小为 $224 \times 512 \times 512 \times 12$ bit,若一次输入整幅图像数据,因FPGA芯片内部资源有限,不能实时完成自相关运算。考虑矩阵的自相关运算特性

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mm} \end{bmatrix}, \quad (4)$$

式中 $b_{ij} = a_{i1} \times a_{1j} + a_{i2} \times a_{2j} + \dots + a_{im} \times a_{mj}$, m 为图像波段数, n 为图像空间大小。在此降维运算中,高光谱图像数据通过网口传输,图像数据阵按列传输。由矩阵运算特性和数据的传输特性,可对数据矩阵按列进行分块运算,考虑实时处理要求,通过网口输入的数据需利用BlockRAM进行缓存。数据接收端用两个BlockRAM进行数据的协调缓存。

在对图像进行自相关运算时,选取 32 个乘法 IP 核并行进行乘法累加运算,不断进行乘法相关并累加前一列相关结果,最后便可得到自相关图像矩阵。奇异值分解算法模块所需数据为浮点型,但是 FPGA 芯片内部逻辑资源有限,选取浮点数据为单

精度型。

自相关模块又可分为几个基本的小模块,其模块关系如图 1 所示。其中 RAMC 用来保存相关运算的中间结果,RAMD 用来保存整幅图像自相关结果转换成浮点型后的矩阵数据。

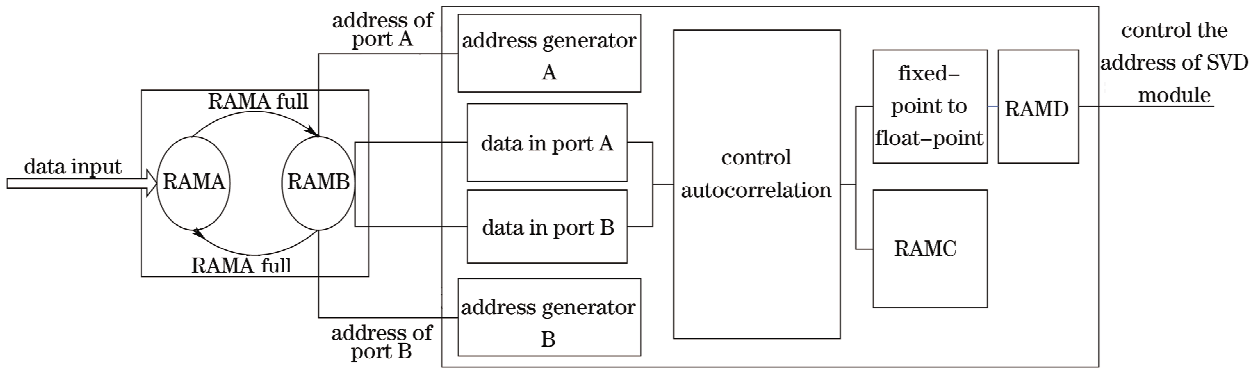


图 1 自相关模块结构图

Fig. 1 Autocorrelation module chart

3.2 特征求解模块设计

特征值求解模块即奇异值分解模块,利用雅可比过关法进行奇异值分解,雅可比过关法又可分为几个基本的子模块,模块关系如图 2 所示。

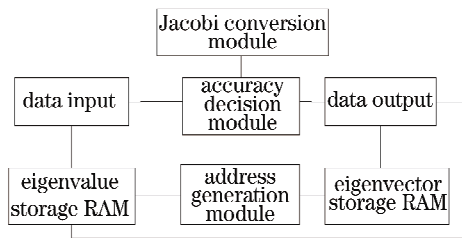


图 2 奇异值分解子模块图

Fig. 2 Singular value decomposition (SVD) module chart

其中特征值存储 RAM 原始数据为自相关模块中数据输出结果矩阵,特征向量存储矩阵 RAM 中原始数据为单位矩阵。精度判定矩阵模块判断特征值矩阵是否达到精度要求,若达到要求则奇异值分

解完成;若没达到精度要求,则继续变换到满足精度要求。雅可比旋转模块用来变换进行奇异值分解。整个模块程序流程如图 3 所示。

3.3 特征提取模块设计

特征提取模块主要功能是特征值排序,并找出对应的特征向量。在此模块中不需对整个矩阵排序,只需提取特征值矩阵的几个大值,并找出与之对应的特征向量即可。整个模块的程序流程如图 4 所示。

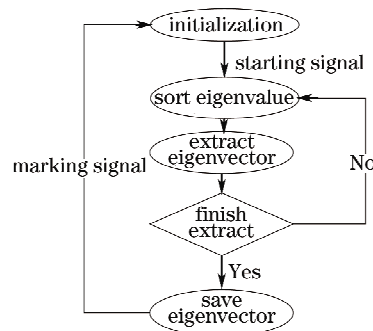


图 4 特征值提取模块程序流程图

Fig. 4 Feature extraction module flow chart

3.4 降维实现模块设计

高光谱图像数据的降维结果由原始高光谱图像数据矩阵与特征提取模块所得结果矩阵相乘得到。降维结果的实现与自相关矩阵的求解过程相似,因而降维实现模块的设计基本与自相关模块的设计相同。

3.5 处理时间管理

要达到实时要求,需严格管理各模块处理时间。当数据传输率为 400 Mb/s 时,一幅 113 波段 160×160 大小高光谱图像需时间为 0.086784 s。

自相关模块运算时间取决于乘法运算时间,输入

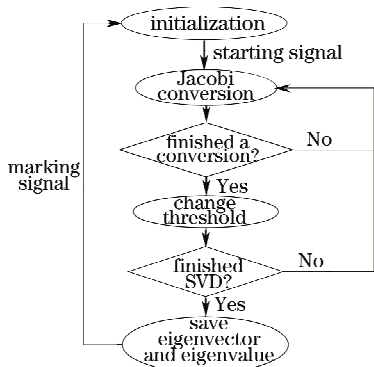


图 3 奇异值分解模块程序流程图

Fig. 3 SVD module flow chart

图像数据矩阵大小为 $113 \times 160 \times 160$, 需乘法 326886400 次。当 Virtex-5 芯片工作于 200 MHz 的时钟频率时, 选用乘法 IP 核 32 个, 其运算时间为 0.051076 s, 远小于图像传输时间, 满足实时性要求。

特征求解模块的运行时间取决于数值解法的旋转次数和比较次数。自相关运算得到矩阵大小为 113×113 。用雅可比过关法对其进行运算时, 选取 4 个乘法 IP 核进行变换, 则完成特征值运算共需要周期数大概为 1.16×10^7 , 其运算时间为 0.058 s, 也满足实时性要求。

特征值提取模块只需简单排序运算, 所花时钟周期较少, 可满足实时性要求。降维实现模块的运算次数少于自相关模块的运算次数, 亦可满足实时性要求。

经分析, 降维芯片可在数据传输速率不高于 400 Mb/s 的情况下达到实时降维的要求。

4 仿真及结果分析

高光谱图像数据来源于澳大利亚机载成像光谱仪(HYMAP)。该成像系统是由澳大利亚集成光电

公司(ISPL)研制生产的, 可获取 126 个连续的光谱波段图像, 光谱范围为 400~2500 nm, 瞬时视场角(IFOV)为 215 mrad。图像大小取 160 pixel \times 160 pixel, 空间分辨率为 4.5 m, 如图 5 所示。



图 5 HYMAP 红树林地区第 5 波段图像及选取区域

Fig. 5 HYMAP mangrove No. 5-band image and the select region

图 6 分别为部分单波段图像, 分别位于第 5, 60, 118 个波段, 对应波长分别为 464, 1314 和 2358nm。

采用单精度浮点型数据进行奇异值分解, 提取高光谱图像的 40 个主成分进行降维, 分别对降维前后的图像信息利用 K 均值法进行分类, 分 4 类后得到的分类结果伪彩色图像如图 7 所示。

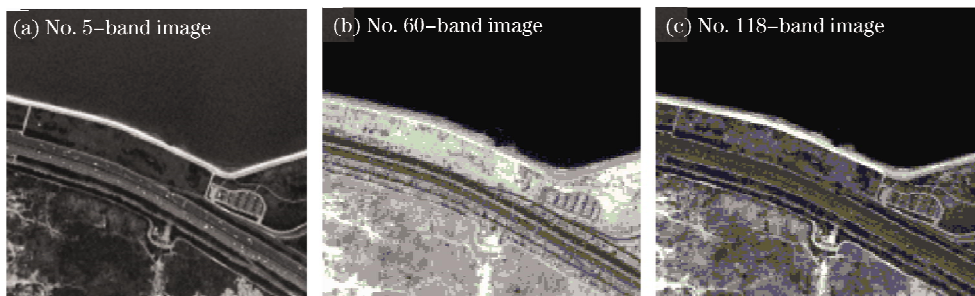


图 6 HYMAP 选取区域单波段图像

Fig. 6 HYMAP single-band image of the select region

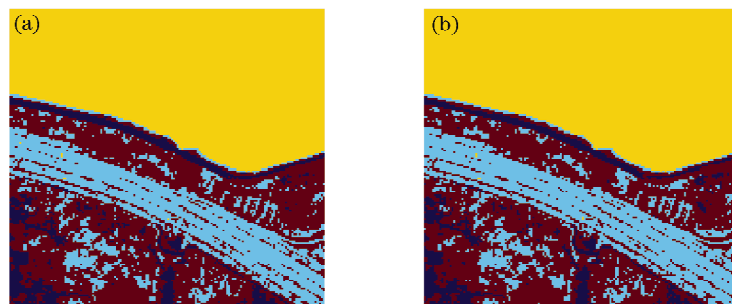


图 7 降维前(a)后(b)分类结果比较

Fig. 7 Results of classification before (a) and after (b) the dimensionality reduction

分析降维前后的分类结果, 提取各类的像素点数目, 像素点数列于表 1 所示。

由表 1 知降维后光谱图像由 113 维降为 40 维, 分类错误像素点数为 54, 其分类误差为 0.2109%。由分类结果可知, 利用本方法进行高光谱图像数据

的降维, 图像信息有少量的损失。而此高光谱图像数据降维前数据量为 34.713 M, 降维后数据量为 12.288 M, 降维后数据量减少了 22.425 M。分析实验结果可知, 利用奇异值分解法进行高光谱图像数据降维可以实现图像数据大幅度降低, 便于高光

谱图像数据信息的后续利用和数据的存储与处理。

表 1 降维前后分类像素点数目表

Table 1 Results of classification before and after the dimensionality reduction

Sort	1	2	3	4
Before dimensionality reduction	5219	1679	7770	10932
After dimensionality reduction	5207	1664	7797	10932

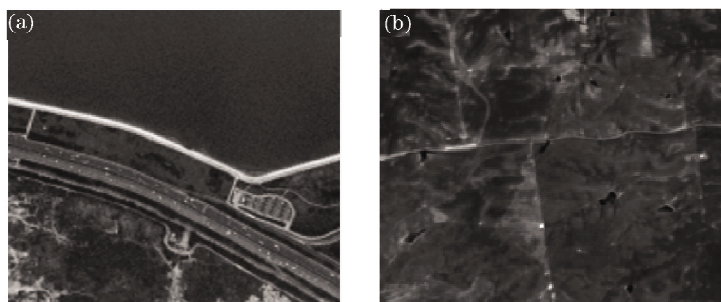


图 8 复杂场景(a)及简单场景(b)的原始高光谱图像

Fig. 8 Original hyperspectral image of complex scene (a) and simple scene (b)

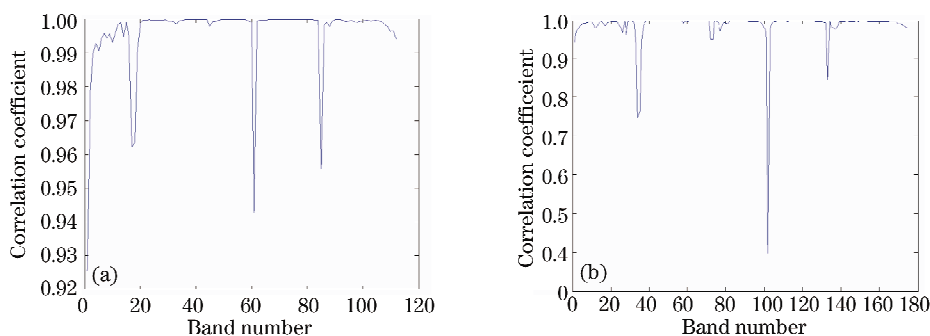


图 9 复杂场景(a)及简单场景(b)的波段间相关系数

Fig. 9 Correlative coefficient between bands in complex scene (a) and simple scene(b)

5 结 论

基于 FPGA 的高光谱图像数据奇异值分解算法,选取雅可比过关法进行奇异值分解,数据类型为单精度浮点型,完成基于 FPGA 的实时高光谱图像降维芯片的设计和仿真。仿真结果表明,高光谱图像降维后数据量为降维前的 1/3,而降维后的分类像素点误差为 0.2109%,证明了奇异值分解算法进行高光谱图像降维算法的有效性。

参 考 文 献

- 1 Wang Su, Chang Chien-L. Variable-size variable-band selection for spectral feature characterization in hyperspectral data [C]. *SPIE*, 2006, **6378**: 63781B-1~63781B-11
- 2 Lü Dongya, Huang Puming, Sun Xianpu. Data feature and compression of hyperspectral images [J]. *Space Electronic Technology*, 2005, **20**(1): 15~22
吕东亚, 黄普明, 孙献璞. 高光谱图像的数据特征及压缩技术 [J]. *空间电子技术*, 2005, **20**(1): 15~22
- 3 Jia Xiuping, J. A. Richards. Segmented principal components

高光谱图像相邻波段相关性越高,数据冗余度越大。场景越复杂的图像其相邻波段相关性越小,对比高光谱图像 8(a),(b),知图 8(a)场景比图 8(b)的复杂,得到它们相邻波段相关系数如图 9,知图 8(a)的数据冗余度小,利用奇异值分解算法降维能降低 2/3 的数据量,对于场景复杂度小的图像则也都能达到理想的降维效果。

transformation for efficient hyperspectral remote-sensing image display and classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(1): 538~542

- 4 L. O. Jimenez, R-M. L. Jorge. Integration of spatial and spectral information homogenous by means of unsupervised extraction and classification for objects applied to multispectral and hyperspectral data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, **43**(4): 844~851
- 5 Wang Xiupeng. Hyperspectral Dimensionality reduction based on projection pursuit [D]. Shaanxi: Northwestern Polytechnical University, 2006. 3
王秀朋. 基于投影寻踪的高光谱图像降维算法研究 [D]. 陕西: 西北工业大学, 2006. 3
- 6 A. A. Green, M. Berman, P. Switzer. A transformation for ordering multispectral data in terms of image quality with implications for noise removal [J]. *IEEE Transactions Geoscience and Remote Sensing*, 1988, **26**(1): 65~74
- 7 Xiaoqun Wang, I. H Sloan. Brownian bridge and principal component analysis: towards removing the curse of dimensionality [J]. *IMA Journal of Numerical Analysis*, 2007, **27**(4): 631~654
- 8 D. Manolakis, D. Marden. Dimensionality reduction of hyperspectral imaging data using local principal components

- transforms[C]. *SPIE*, 2004, **5425**: 393~401
- 9 D. F. Michael, R. M. Mersereau. On the impact of PCA dimension reduction for hyperspectral detection of difficult targets [J]. *IEEE Geoscience and Remote Sensing Letters*, 2005, **2**(2): 192~195
- 10 Yan Jingwen, Qu Xiaobo, Chen Jiazhen. A new quasi-lossless compression of partitioned KLT and IWT multispectral images [J]. *Acta Optica Sinica*, 2007, **27**(10): 1740~1744
闫敬文, 屈小波, 陈嘉臻. 分组 Karhunen-Loeve 变换/整数小波变换高光谱影像无损压缩新方法[J]. *光学学报*, 2007, **27**(10): 1740~1744
- 11 L. O. Jimenez, L. A. David. Hyperspectral data analysis and supervised feature reduction via projection pursuit [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(6): 2653~2667
- 12 Xun Lina, Fang Yonghua, Li Xin. A small target detection approach based on endmember extraction in hyperspectral image [J]. *Acta Optica Sinica*, 2007, **27**(7): 1178~1182
寻丽娜, 方勇华, 李 新. 高光谱图像中基于端元提取的小目标检测算法[J]. *光学学报*, 2007, **27**(7): 1178~1182
- 13 Du Hongtao, Qi Hairong. An FPGA implementation of parallel ICA for dimensionality reduction in hyperspectral images [J]. *IEEE International Geoscience and Remote Sensing*, 2004, **5**(5): 3257~3260
- 14 Zhang Junping, Zhang Hua. State-of-arts and analysis on hyperspectral image classification in imaging spectral technique [J]. *Chinese Space Science and Technology*, 2001, **2**(1): 37~44
张钧萍, 张 晔. 成像光谱技术超谱图像分类研究现状与分析 [J]. *中国空间科学技术*, 2001, **2**(1): 37~44
- 15 Tang Qingtao, Rao Yunjiang, Zhu Tao *et al.*. High-resolution signal demodulation method for fiber-optic Fabry-Perot multiplexing system [J]. *Chinese J. Lasers*, 2007, **34**(10): 1353~1357
唐庆涛, 饶云江, 朱 涛 等. 光纤法-珀传感系统高分辨率复用信号解调方法[J]. *中国激光*, 2007, **34**(10): 1353~1357
- 16 Xu Shiliang. *Common Algorithm Procedure Assembly* [M]. Beijing: Tsinghua University Press, 2004
徐士良. *常用算法程序集*[M]. 北京:清华大学出版社, 2004