

## 基于相关向量机的高光谱图像噪声评估算法

王晓飞<sup>1,2</sup>, 侯传龙<sup>2</sup>, 阎秋静<sup>2</sup>, 张钧萍<sup>3</sup>, 汪爱华<sup>1,4</sup>

- (1. 北京二十一世纪科技发展有限公司, 北京 100096;
2. 黑龙江大学 黑龙江省普通高等电子工程重点实验室, 黑龙江 哈尔滨 150080;
3. 哈尔滨工业大学 信息工程系, 黑龙江 哈尔滨 150001;
4. 二十一世纪空间技术应用股份有限公司, 北京 100096)

**摘要:** 为了更准确的估计高光谱图像噪声强度, 提出了一种基于相关向量机(RVM)的高光谱图像噪声评估算法。对该算法所采用的 RVM 回归原理、残差与噪声的关系等进行了研究。首先, 介绍了高光谱图像噪声评估中应用较为广泛的空间/光谱维去相关法的特点及不足。接着, 对可有效进行非线性回归分析的 RVM 进行了介绍。然后, 针对传统的空间/光谱维去相关法在系统中存在较强的非线性关系时, 得到的残差将会过大这一问题, 提出利用 RVM 回归分析去除具有高相关性的信号, 利用得到的残差图像对噪声进行估算, 从而提高评估系统的稳定性。实验结果表明: 噪声强度估计精度优于 8%; 相比传统算法更有效。总体看, 该算法可以满足自动高光谱图像噪声评估的稳定可靠、精度高等要求。

**关键词:** 高光谱; 噪声估计; 相关向量机(RVM); 去相关法

**中图分类号:** TP751.1 **文献标志码:** A **文章编号:** 1007-2276(2014)12-4159-05

## Noise estimation algorithm based on relevance vector machine for hyperspectral imagery

Wang Xiaofei<sup>1,2</sup>, Hou Chuanlong<sup>2</sup>, Yan Qiujing<sup>2</sup>, Zhang Junping<sup>3</sup>, Wang Aihua<sup>1,4</sup>

- (1. Beijing Twenty-First Century Science&Technology Development Co. Ltd, Beijing 100096, China;
2. Key Laboratory of Electronic Engineering, College of Heilongjiang Province, Heilongjiang University, Harbin 150080, China;
3. Department of Information Engineering, Harbin Institute of Technology, Harbin 150001, China;
4. Twenty First Century Aerospace Technology Co. Ltd, Beijing 100096, China)

**Abstract:** In order to more accurately estimate noise intensity for hyperspectral imagery, the paper proposed a noise estimation algorithm based on relevance vector machine(RVM)for hyperspectral imagery. And the algorithm that used RVM regression, residuals and noise was studied. First of all, this paper introduced the characteristics and shortage of spatial/spectral dimension decorrelation in noise estimation that used widely nowadays for hyperspectral imagery. Then, the nonlinear regression analysis of RVM was introduced. And the residuals will be too large, when there was a strong nonlinear correlation in the system for spatial/spectral dimension decorrelation. To this problem, the paper proposed a new method that used RVM regression to remove strong signal correlation and used the residual images to estimate the

收稿日期: 2014-04-10; 修订日期: 2014-05-15

基金项目: 国家自然科学基金(61273148); 北京市博士后工作经费资助项目(2012ZZ-100); 北京市科技计划(Z121100006112032); 黑龙江省普通高等电子工程重点实验室项目(D22D20100018)

作者简介: 王晓飞(1977-), 男, 副教授, 博士, 主要从事高光谱图像处理及应用方面的研究。Email: [nk\\_wxf@hlju.edu.cn](mailto:nk_wxf@hlju.edu.cn)

noise, so as to improve the stability of the assessment system. Experimental results indicate that the precision of the noise intensity is better than 8%, and show that the method is more effective compared to the traditional method. It concludes that the RVM can satisfy the system requirements of higher precision and stabilization in noise estimation for automatic hyperspectral imagery.

Key words: hyperspectral imagery; noise estimation; RVM; decorrelation

## 0 引言

自 20 世纪 80 年代成像光谱仪问世后,人们不仅可以获得地物目标二维空间景物信息,同时也可以获得更详尽的表征其物理属性的光谱信息,从而实现了图谱合一<sup>[1]</sup>。目前高光谱图像的光谱分辨率已经达到 10 nm 甚至更高,可以解决许多原先全色或多光谱不能解决的难题。因此,国内外很多学者都对高光谱图像的处理与应用给予了足够的关注,并得到了很多的研究成果。这些研究成果对目前人类迫切需要解决的一些重点问题如环境保护、能源危机、矿产勘探、农业发展等都具有重要的研究价值和应用意义<sup>[2]</sup>。

遥感图像噪声强度是衡量遥感器性能的重要指标之一<sup>[3]</sup>,同时,高光谱图像因成像光谱仪波段通道窄而造成光成像能量低,使其更容易受噪声影响,因此,针对高光谱图像的噪声评估对其应用而言具有重要的意义。

针对遥感图像进行噪声评估主要有方法有<sup>[3]</sup>:实验室法、暗电流法和图像法。其中实验室法和暗电流法都是在传感器设计和制造过程中经常采用的,需要特定的试验条件,在对遥感图像数据处理中较难实现。而图像法是通过获取的图像进行分析来估算噪声强度的方法,因其更利于实际应用而得到了广泛的研究。图像法中常用的有均匀区域法、地学统计法<sup>[4]</sup>、局部均值与局部标准差法<sup>[5]</sup>和空间/光谱维去相关法<sup>[6]</sup>等。这些算法及其改进算法都是通过不同的方式,利用遥感图像的空间性实现了噪声的估计的,而对于高光谱遥感而言,由于其独特的高光谱分辨率,可以得到一条近似连续的光谱曲线,相邻波段间具有很高的相关性。同时,其作为图像,其在空间域上的也具有连续性,因此,其每个波段的相邻像素值也具有较强相关性。空间/光谱维去相关法(Spectral and Spatial De-correlation Method, SSDC)就

是利用了这种空间维和光谱维存在高相关性的特点,通过多元线性回归去除具有高相关性的信号,利用得到的残差图像对噪声进行估算。使计算结果更接近真实情况,该方法受地物覆盖类型影响小,并且可以自动执行,而成为目前较为稳定的高光谱图像噪声评估方法<sup>[7]</sup>。但是,传统的 SSDC 法采用的是多元线性回归法进行相关性去除,而实际相邻像素间的关系往往并不是线性的<sup>[8]</sup>,这样必然会引入较大的回归误差。为了解决非线性回归的问题,文中引入了在分类和回归分析中都发挥重要作用的相关向量机来解决相关性去除的问题<sup>[9]</sup>。

## 1 相关向量机回归原理

对于一维目标函数的回归问题,假定训练集为  $\{x_n, t_n\}_{n=1}^N$ , 目标函数是来自带有附加噪声的模型的样本,即:

$$t_n = y(X_n; w) + \varepsilon_n \quad (1)$$

这里  $\varepsilon_n$  可以认为是一个独立分布的噪声分量,一般都假设其是高斯分布,即  $\varepsilon_n \sim N(0, \sigma^2)$ , 这样有  $p(t_n | X) \sim N(t_n | Y(x_n), \sigma^2)$ , 即  $t_n$  也是一个高斯分布,其均值为  $y(X_n)$ , 方差为  $\sigma^2$ 。则相关向量机最后的回归形式可表示为:

$$y(X; w) = \sum_{i=1}^N w_i K(X, X_i) + w_0 \quad (2)$$

式中:  $K(X, X_i)$  为一个核函数;  $w_i$  是回归系数。对应系数不为零的向量称为相关向量。

可以合理地假设  $\{t_n\}_{n=1}^N$  是相互独立的随机变量,结合公式(1),则训练集的概率分布函数为:

$$p(t | w, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} \|t - \Phi w\|^2\right\} \quad (3)$$

这里  $t = (t_1 \cdots t_N)^T$ ,  $w = (w_1 \cdots w_N)^T$ ,  $\Phi$  是一个大小为  $N \times (N+1)$  的设计矩阵,其中  $\phi_{m1} = 1$ ,  $\phi_{mn} = K(X_m, X_{n-1})$ 。由前面的概率估计,可求得条件概率为:

$$p(t^* | t) = \int_{\theta} p(t^* | w, \delta^2) p(w, \delta^2 | t) dw d\sigma^2 = \int_{\theta} p(t^* | w, \delta^2) \frac{p(t | w, \delta^2) p(w, \delta^2)}{p(t)} dw d\sigma^2 \quad (4)$$

RVM 是将每一个参数定义了一个先验概率分布来做限制。然后引进超参数,使每个权系数都有一个对应的超参数独立地对其进行限制和约束,为  $w$  加上先决条件:它们的概率分布是落在 0 周围的正态分布<sup>[10]</sup>:  $p(w_i | \alpha_i) = N(w_i | 0, \alpha_i^{-1})$ ,  $\alpha = [\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N]^T$ 。则

$$p(w | \alpha) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}) = \prod_{i=0}^N \frac{\alpha_i}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_i w_i^2}{2}\right) \quad (5)$$

这样,就为每一个权值设置了独立的超参数。

对于给定一个新的检验点  $x^*$ , 其相应的目标预测  $t^*$  的分布:

$$p(t^* | t) = \int p(t^* | w, \alpha, \delta^2) p(w, \alpha, \delta^2 | t) dw d\alpha d\delta^2 \quad (6)$$

对于  $p(t^* | w, \alpha, \delta^2)$  中,因为  $t^*$  只和  $w$  与  $\sigma^2$  直接相关,由马尔可夫性质得到:

$$p(t^* | w, \alpha, \sigma^2) = p(t^* | w, \sigma^2) = N(t^* | y(x^*; w), \delta^2) \quad (7)$$

有了先验之后,由贝叶斯准则可以得到所有未知参数的后验概率:

$$p(w, \alpha, \sigma^2 | t) = \frac{p(t | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2)}{p(t)} \quad (8)$$

由于公式(8)不易直接计算,将后验分解如下:

$$p(w, \alpha, \sigma^2 | t) = p(w | t, \alpha, \sigma^2) p(\alpha, \sigma^2 | t) \quad (9)$$

对于公式(9)等号右边第一项,可直接计算:

$$p(w | t, \alpha, \sigma^2) = \frac{p(t | w, \sigma^2) p(w | \alpha)}{p(t | \alpha, \sigma^2)} = (2\pi)^{-(N+1)/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(w-\mu)^T \Sigma^{-1}(w-\mu)\right\} \quad (10)$$

$$p(t | \alpha, \sigma^2) = (2\pi)^{-N/2} |\Omega|^{-1/2} \exp\left\{-\frac{t^T \Omega^{-1} t}{2}\right\} \quad (11)$$

其中后验分布的均值  $\mu = \sigma^{-2} \Sigma \Phi^T t$ ; 方差  $\Sigma = (\sigma^{-2} \Phi^T \Phi + A)^{-1}$ ,  $A = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$ 。

## 2 RVM 的噪声评估算法

文中算法其实现过程如下:

(1)首先对高光谱图像进行分块处理,分成大小

相同的矩形,连续而且不得重叠。这样是为了得到相对均一地物覆盖的地块。需要注意的是,虽然地块越小,地物覆盖会越均匀,但为了保证回归分析中有足够的训练样本用于对回归参数进行估计,地块不能过小。

(2)对每个地块分别进行多元 RVM 回归分析,设地块大小为  $m \times n$ ,  $x_{ij}^k$  代表在第  $k$  个波段中位于  $(i, j)$  点的像元值,则训练集为  $\{X_{ij}^k, t_{ij}^k\}$ , 其中  $t_{ij}^k = x_{ij}^k$ ,  $X_{ij}^k =$

$$\{x_{ij}^{k-1}, x_{ij}^{k+1}, x_p^k\}, 1 \leq i \leq m, 1 \leq j \leq n, x_p^k = \begin{cases} x_{i-1,j}^k & i > 1, j = 1 \\ x_{i,j-1}^k & j > 1 \\ \text{无意义} & i = 1, j = 1 \end{cases}$$

利用训练集对 RVM 回归机进行训练后,可得到相应参数  $w$ 。

(3)采用得到的回归参数  $w$  后,利用公式(2)求取该点的灰度拟合值  $\hat{x}_{ij}^k$ :

$$\hat{x}_{ij}^k = \sum_{i=1}^N w_i K(X, x_{ij}^k) + w_0 \quad (12)$$

则残差值  $r_{ij}^k = x_{ij}^k - \hat{x}_{ij}^k$ 。

(4)对求取残差值,计算其方差的无偏估计,将其作为第  $k$  波段该块的噪声方差  $\sigma_k^2 = (M-4)^{-1} S^2$ , 其中

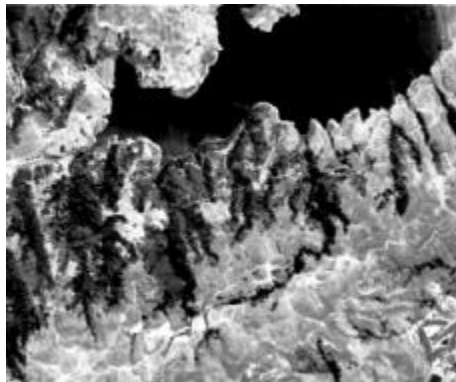
$$M = (m-1) \times (n-1) - 1, S^2 = \sum_{i=2}^{m-1} \sum_{j=2}^{n-1} (r_{ij}^k)^2$$

(5)与 SSDC 类似,这样也将第  $k$  波段所有地块残差的标准差的平均值作为第  $k$  波段的噪声估计,其中为了减小某些奇异地块的影响,通常截取中间 80% 的数值为有效值,并取平均值<sup>[7]</sup>。

## 3 实验及分析

这里选用了两幅 AVIRIS 图像及其人工加噪图像作为实验图像对算法进行验证。AVIRIS 图像是由美国喷气推进实验室研制的覆盖 400~2 500 波长范围的成像光谱仪,共 224 个波段,光谱采样间隔为 10 nm。

所选择的两幅实验图像如图 1(a)、(b)所示,为同一次航空实验获得,地点为美国 Moffett Field 地区。其中图(a)为数据 1,以山地水体为主,纹理简单,而图(b)为数据 2,以城市建筑为主,相对于图(a)纹理复杂。由于这两幅图像为同一次航空实验获得,可以认为包含噪声的强度相同。同时,为了验证算法的有



(a) 山地和水体为主图像

(a) Image dominated by mountains and water



(b) 城市为主的图像

(b) Image dominated by city

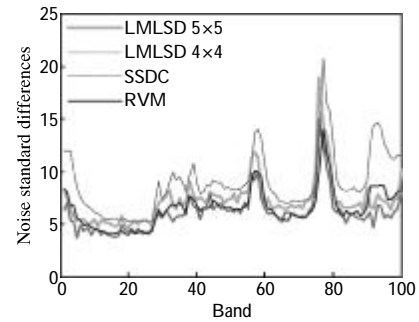
图 1 实验用 AVIRIS 数据

Fig.1 AVIRIS data used in the experiment

效性,还分别对这两幅图像叠加了标准差为 30、60 和 90 的高斯噪声。

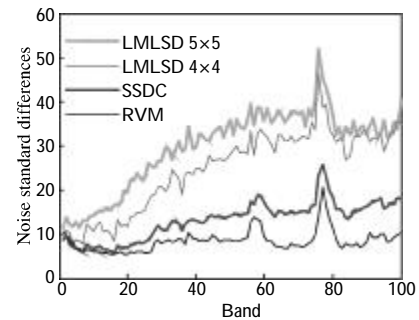
对以上数据分别采用 LMLSD 法、SSDC 法和文中方法(简称 RVM 法)进行噪声评估实验。其中 LMLSD 法模板大小分别取  $4 \times 4$  和  $5 \times 5$ ,SSDC 和 RVM 法模板取  $15 \times 15$ ,相关向量机的核函数采用高斯核函数。图 2 是对两组真实数据进行噪声评估结果。对于地物覆盖的图像,采用不同模板的 LMLSD 法、SSDC 法和文中方法都能得到相似的结果。但是当物覆盖不均匀或包含较多纹理时,LMLSD 的模板稍作改变,结果就发生了较大的变化,而 SSDC 法结果也有一定程度的改变,而文中方法受纹理等因素的影响较小。

同时,为了测试文中方法的有效性和估计精度,又分别对仿真数据进行了测试。对于添加不同强度噪声的数据,采用 RVM 法进行了噪声评估实验。图 3 为实验结果,其中实线为数据 1 及其加噪数据评估



(a) 数据 1 测试结果

(a) Test result of data 1



(b) 数据 2 测试结果

(b) Test result of data 2

图 2 两组数据不同方法测试结果

Fig.2 Test result of two set of data using different method

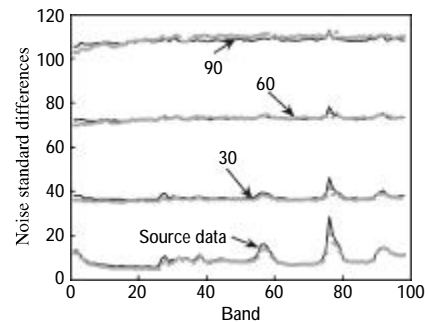


图 3 采用 RVM 法对额外添加不同强度噪声图像评估结果

Fig.3 Assessment results of adding different intensity noise using RVM method

结果,虚线为数据 2 及其加噪数据评估结果。从图中可以看出,不同噪声强度下,文中方法都能输出稳定的结果。同时,实验结果和图像原始噪声叠加上高斯噪声的强度也是基本吻合的。各波段的估计误差不超过 8%。表明该方法不仅能输出稳定的结果,精度也能达到实际应用的要求。

#### 4 结论

文中针对空间/光谱维去相关(SSDC)法当图像



中存在明显的特征吸收或者复杂的纹理时,其光谱维或空间维的线性关系变弱,利用线性回归分析方法来计算噪声的误差较大这一问题,提出基于 RVM 的高光谱图像噪声评估算法。首先,分析了空间/光谱维去相关法的特点。接着,介绍了 RVM 回归分析的工作原理。然后,提出利用 RVM 回归分析去除具有高相关性的信号,利用得到的残差图像对噪声进行估算。实验结果证明:文中算法在对于存在明显的特征吸收,或者复杂的纹理的高光谱图像,进行噪声评估时,也能输出稳定的结果。并且精度优于 8%,基本满足了自动高光谱图像噪声评估的要求,具有一定实际应用价值。

#### 参考文献:

- [1] Liu Danfeng, Wang Ligu. Color display of hyperspectral data in three levels [J]. *Infrared and Laser Engineering*, 2012, 41(9): 2527-2533. (in Chinese)  
刘丹凤, 王立国. 高光谱数据三级彩色显示方法 [J]. *红外与激光工程*, 2012, 41(9): 2527-2533.
- [2] Cheng Xin, Zhang Bao, Hong Yongfeng, et al. Optical design of an airborne dual-wavelength imaging spectrometer with high throughput [J]. *Infrared and Laser Engineering*, 2012, 41(3): 690-695. (in Chinese)  
程欣, 张葆, 洪永丰, 等. 机载高光通量双波段成像光谱仪的设计 [J]. *红外与激光工程*, 2012, 41(3): 690-695.
- [3] Zhang Bing, Gao Lianru. *Hyperspectral Image Classification and Target Detection* [M]. Beijing: Science Press, 2011. (in Chinese)  
张兵, 高连如. *高光谱图像分类与目标探测* [M]. 北京: 科学出版社, 2011.
- [4] Curran P J, Dungan J L. Estimation of signal-to-noise: a new procedure applied to AVIRIS data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1989, 27(5): 620-628.
- [5] Gao B C. An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers [J]. *Remote Sensing of Environment*, 1993, 43(1): 23-33.
- [6] Roger R E, Arnold J F. Reliably estimating the noise in AVIRIS hyperspectral images [J]. *International Journal of Remote Sensing*, 1996, 17(10): 1951-1962.
- [7] Liu Xiang. *Target detection on hyperspectral imagery based on transformation of spectral dimensions* [D]. Beijing: Graduate University of the Chinese Academy of Sciences, 2008. (in Chinese)  
刘翔. *基于光谱维变换的高光谱图像目标探测研究* [D]. 北京: 中国科学院研究生院, 2008.
- [8] Wei Feng, He Mingyi, Mei Shaohui. Hyperspectral data feature extraction using spatial coherence based neighborhood preserving embedding [J]. *Infrared and Laser Engineering*, 2012, 41(5): 1249-1254. (in Chinese)  
魏峰, 何明一, 梅少辉. 空间一致性邻域保留嵌入的高光谱数据特征提取 [J]. *红外与激光工程*, 2012, 41(5): 1249-1254.
- [9] Shutin D, Buchgraber T, Kulkarni S R, et al. Fast variational sparse bayesian learning with automatic relevance determination for superimposed signals [J]. *IEEE Transactions on Signal Processing*, 2011, 59(12): 6257-6261.
- [10] Tipping M. Sparse Bayesian learning and the relevance vector machine [J]. *Journal of Machine Learning Research*, 2001 (1): 211-244.