

人工神经网络计算大气点扩展函数

王海东^{1,2} 马晓珊¹ 杨震¹ 李立钢¹

¹中国科学院国家空间科学中心复杂航天系统电子信息技术重点实验室, 北京 100190

²中国科学院大学, 北京 100049

摘要 大气点扩展函数(PSF)是开展光学遥感邻近效应研究和校正的有效方法。基于蒙特卡罗模拟获得的大气PSF,设计足够多带有 Sigmoid 函数的隐藏神经元和线性输出神经元的两层前馈神经网络,采用 Levenberg-Marquardt 反向传播算法,获得了大气、光谱和观测几何等输入参数与大气 PSF 之间的关系。模拟结果证明该方法能够在相对较短的时间内,以 95% 的计算精度产生预期的大气 PSF 的近似值。

关键词 遥感; 邻近效应; 大气点扩展函数; 逆向蒙特卡罗法; 人工神经网络

中图分类号 TP751 **文献标识码** A

doi: 10.3788/LOP53.102801

Computing the Atmospheric Point Spread Function by Artificial Neural Networks

Wang Haidong^{1,2} Ma Xiaoshan¹ Yang Zhen¹ Li Ligang¹

¹Key Laboratory of Electronics and Information Technology for Space System, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

²University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Atmospheric point spread function (PSF) is an effective research and correction method for the adjacency effect of the optical remote sensing. Based on the atmospheric PSF acquired by Monte Carlo simulation, a two-layer feed-forward neural network which has enough hidden neurons with Sigmoid function and linear output neurons is designed and implemented. By means of Levenberg-Marquardt back-propagation algorithm, the relationship between the atmospheric PSF and its influence factors, such as atmosphere condition, spectral range and observation geometry is obtained. The results obtained show that our neural network can estimate the atmosphere PSF with 95% accuracy within relatively short time.

Key words remote sensing; adjacency effect; atmospheric point spread function; backward Monte Carlo method; artificial neural network

OCIS codes 280.4788; 010.1285; 200.4260

1 引言

光学遥感的邻近效应是指遥感中非目标像元对目标像元辐射亮度的贡献,又称为大气的交叉辐射效应^[1-2]。邻近效应一般增加了暗像元的辐射值,减少了亮像元的辐射值,使得遥感图像边缘模糊、反差降低,影响卫星对地面目标的遥感能力^[3]。随着成像空间分辨率的提高,这种影响越来越显著。Lyapustin 等^[4]的研究表明:当传感器的空间分辨率高于 1 km 时必须考虑邻近效应的影响。随着光学遥感空间分辨率和应用遥感信息定量化程度要求的不断提高,开展邻近效应机理及其校正方法研究已成为亟待解决的关键技术问题。

邻近效应可以看作是地表辐射场和大气点扩展函数(PSF)的卷积^[5],大气 PSF 的物理意义是它描述了邻近像元对传感器总辐射亮度贡献率的空间分布。对邻近效应研究的核心是对大气 PSF 的求解。传统的

收稿日期: 2016-06-03; **收到修改稿日期:** 2016-06-06; **网络出版日期:** 2016-09-20

基金项目: 国家自然科学基金(61308098)

作者简介: 王海东(1992—),男,硕士研究生,主要从事邻近效应方面的研究。E-mail: nongfugengxia@163.com

导师简介: 马晓珊(1978—),女,博士,副研究员,硕士生导师,主要从事光学遥感成像模拟、大气光学特征方面的研究。

E-mail: maxs@nssc.ac.cn(通信联系人)

大气 PSF 的获取方法主要有辐射传输方程的解析近似^[6-8]和蒙特卡罗模拟^[9-10]两种。辐射传输方程的解析近似抛弃传输方程的多次散射项,和实际大气状况存在差距;蒙特卡罗模拟以最大的逼真度描述了光子在大气中的传输过程,既保留了多次散射过程,又回避了求解辐射传输方程的困难,是研究邻近效应的有效方法^[11]。但蒙特卡罗方法存在的问题是不能获得简单明晰的解析表达式,其规律只能靠对大量的统计数据进行分析获得。只有当模拟的次数足够高时,才能获得稳定的、有代表性的结果,有时需要几十万次到几百万次的数值模拟实验^[12]。

根据 Cong^[13]的研究,可以尝试采用神经网络计算大气 PSF。利用人工神经网络的后向传输监督学习机制^[14],对大气 PSF 的蒙特卡罗模拟结果进行优化,获得大气特征和大气 PSF 之间的关系,该研究结果可用于光学遥感邻近效应的模拟与校正。

2 蒙特卡罗模拟

利用蒙特卡罗算法模拟大气 PSF 包括两个步骤:1) 建立大气模型,即构建光子的传输环境;2) 用逆向蒙特卡罗模拟光子的传输过程。

2.1 构建光子传输环境

光子传输的大气环境基于大气辐射传输模型 MODTRAN4 获得。预先计算不同高度的大气光学厚度(包括分子和气溶胶光学厚度),建立大气光学厚度和大气层高度的查找表,用于蒙特卡罗模拟时确定光子的空间位置。设置 MODTRAN4 的计算条件如下:1976 年美国标准大气,乡村型气溶胶类型,波长 $\lambda = 550 \text{ nm}$ 、能见度 $v = 23 \text{ km}$,计算大气光学厚度随观测距离的变化,如图 1 所示。

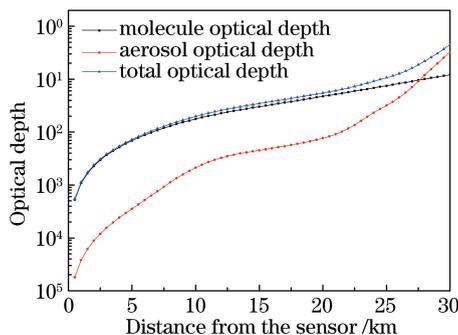


图 1 大气光学厚度随观测距离的变化

Fig. 1 Variation of the atmospheric optical depth with the distance

2.2 逆向蒙特卡罗模拟

由于粒子传输问题的互易性,大气 PSF 可以等同于一束光子从传感器向目标发射、经大气传输、最终在地面形成光子数的相对密度空间分布函数。利用逆向蒙特卡罗法可以对上述光子传输过程进行随机模拟,详细过程参看文献^[12]。针对 550 nm 波长,从 30 km 高度垂直向下发射 10^6 个光子,采用计算图 1 所需的大气条件以及图 1 所示的大气光学厚度廓线,通过逆向蒙特卡罗模拟获得大气 PSF 的一个样例,如图 2 所示。

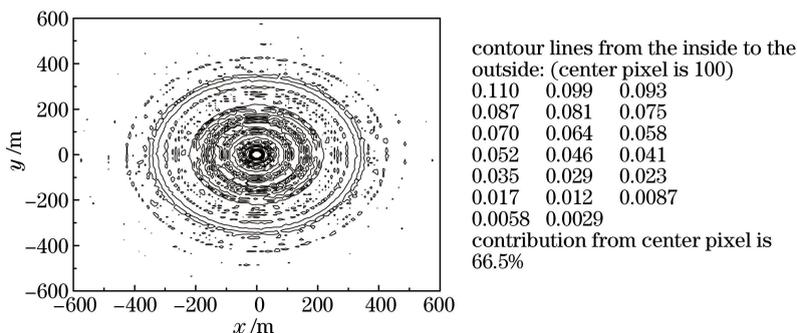


图 2 蒙特卡罗模拟获得的大气 PSF

Fig. 2 Atmospheric point spread function simulated by Monte Carlo method

3 建立人工神经网络模型

从图 2 的模拟结果可知,设置大气参数、光谱参数、观测几何等条件,运行蒙特卡罗模拟程序,可以获得该实验条件下一个特定的大气 PSF。上述任何一个参数发生变化,意味着实验条件的变化,进而获得的大气 PSF 的样式也会发生变化。

本部分的主要目标是设计和实现一个描述由众多物理参数组成的实验条件和大气 PSF 之间关系的神经网络模型。

3.1 输入层

可以作为神经网络输入条件的参数有:波长、大气类型、气溶胶类型、大气能见度、观测天顶角和方位角、观测高度、太阳天顶角等。输入参数可选择项越多,越有助于了解更多的参数对大气 PSF 的影响及其程度,但是增加了神经网络的设计和训练难度。初步计划选择波长、能见度、大气模式作为神经网络的输入参数开展算法实验。

波长选择范围为 500~600 nm,每隔 10 nm 作为一个波长的实验条件,总共 11 个不同的波长。

选择 4 种典型模式作为能见度的取值:发散型雾(能见度为 0.5 km)、乡村气溶胶模式一(能见度为 5 km)、乡村气溶胶模式二(能见度为 23 km)、对流层模式(能见度为 50 km)。

大气模式共有 6 种:热带(北纬 15°)大气模型(记为整数值 1);中纬度夏季(北纬 45°,7 月)大气模型(记为 2);中纬度冬季(北纬 45°,1 月)大气模型(记为 3);亚北极区夏季(北纬 60°,7 月)大气模型(记为 4);亚北极区冬季(北纬 60°,1 月)大气模型(记为 5);1976 年美国标准大气(记为 6)。

上述 3 种条件共有 $11 \times 4 \times 6 = 264$ 种不同组合,组成了 264 种大气光子传输环境下的 264 例样本数据,作为神经网络训练样本数据输入。由输入条件的不同属性,波长和能见度可以取小数值,而大气模式的类型只能取 1~6 之间的整数值,比如,三元值(512.3, 23.5, 6)是有效的,而(512, 23.5, 5.5)是无效的。

3.2 隐含层

隐含层神经元采用 Sigmoid 函数和线性输出神经元,给定与隐含层一致的训练数据和足够的神经元,就能够很好地拟合任意的多维映射问题。一般隐含层选择 10 个神经元,如果训练效果不佳(均方误差较大)则增加神经元数目,重新训练神经网络。

3.3 输出层

模拟获得的大气 PSF 将作为神经网络的输出。根据输入层的描述,共有 264 种不同的输入组合,相应地,就有 264 个大气 PSF 作为网络输出。

通过蒙特卡罗模拟获得光子在地表的分布情况。以目标像元为中心,一定范围内均匀分为若干个间隔,记录每个间隔内接收到的光子数目,即可获得当前实验条件下一个特定样式的大气 PSF。对于垂直观测,大量模拟结果显示,地表接收到的光子数呈对称分布,故可以仅考虑正半轴上的光子分布。

根据图 2 的模拟结果,大多数光子集中在中心像元,周围的光子数急剧下降,以观测目标为中心的边长为 1024 m 的范围外的光子数目可以暂时忽略不计。在正半轴 0~512 m 的范围内,每隔 1 m 放置一个接收光子的容器,记录每个容器接收到的光子数目,以所得到的 512 个光子数目值作为神经网络的输出,则对应于一次实验的输入条件,神经网络将有 512 个输出。

神经网络通常无法承受如此高维的输出。考虑把一个具有 $N(N=512)$ 个输出参数的神经网络划分成 512 个只有一维输出参数的小网络,这些小网络的输入都相同。至此,大气 PSF 与 512 个间隔 1 m 的容器所接收到的光子数目形成了一个映射,每一个大气 PSF 的值表示对应容器中接收到的光子数目。这种做法既解决了高维输出的问题,也有利于在很多工作站或是电脑上同时运行以加速训练。

3.4 数据划分

训练样本采用随机划分的方式:70%训练集,15%验证集,15%测试集。神经网络使用训练集来训练神经网络,根据它的误差来进行调整;验证集用来测量神经网络的泛化能力,当验证样本的均方误差上升时(表明泛化能力提升停止),训练将会被挂起;测试集对于训练没有影响,用于度量神经网络的性能。

3.5 训练算法

采用 Levenberg-Marquardt(LM)反向传播算法^[15],类似于类牛顿法,LM 算法不用计算 Hessian 矩阵

而达到二阶训练速度,当内存不足时,采用共轭梯度下降法来训练神经网络。

反馈网络训练的性能函数采用平方和的形式,其中 Hessian 矩阵近似为

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}, \quad (1)$$

并且梯度 \mathbf{g} 为

$$\mathbf{g} = \mathbf{J}^T \mathbf{e}, \quad (2)$$

其中 \mathbf{J} 是包含神经网络误差关于权重和偏置的一阶偏导的雅可比矩阵, \mathbf{e} 是网络的误差向量。雅可比矩阵能通过比计算 Hessian 矩阵简单得多的标准反向传播技术计算出来。LM 算法使用 Hessian 矩阵的近似,类似牛顿法更新为

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]^{-1} \mathbf{J}^T \mathbf{e}, \quad (3)$$

其中, k 表示第 k 次迭代, $k+1$ 表示第 $k+1$ 次迭代。标量 $\mu=0$ 时,就是牛顿法,使用的是 Hessian 矩阵的近似值。当 μ 比较大时,变成带有小步长的梯度下降。牛顿法更快并且接近误差最小,因此算法设计的目标是尽快地向牛顿法转移。因此,每次成功迭代一次(性能函数减小), μ 就减小;只有当某次迭代使性能函数增加时, μ 才增加。用这种方式,算法每次迭代,性能函数总是在减小。这种算法对于训练中等规模的前馈神经网络是最快的方法。

3.6 训练神经网络

训练神经网络需要足够数量的样本。根据输入层和输出层的设计分析,共生成 264 例训练样本。每一例训练样本,包含 3 个输入参数和 512 个输出参数。

按照神经网络输出层的设计,本实验中共有 512 个神经网络需要训练,第 i 个神经网络的输出为距光子分布中心第 i 个光子容器中接收到的光子数目($i=1\sim 512$)。根据对网络输入层的设计,所有 512 个神经网络,每个网络都有 264 个不同的输入条件。针对 264 组输入条件,基于蒙特卡罗模拟的大气 PSF 数据,以 512 个神经网络进行训练,当神经网络的正确率高于 95% 时停止训练,得到训练之后的神经网络,该网络对于训练样本能够产生超过 95% 的正确输出。

4 模拟结果

4.1 已有条件下的计算结果对比

利用训练得到的神经网络,任意选择 264 个训练样本中的任意一种输入条件,得到相对应大气条件下径向方向接收到的光子数目。比如大气模式为 1976 年美国标准大气能见度为 23 km,波长为 550 nm,可得到对应大气条件下的径向方向接收到的光子数目,并与相同条件下蒙特卡罗模拟所得到的光子数目进行比较,如图 3 所示。图 3 中仅绘出了由于大气作用偏离目标像元的光子。

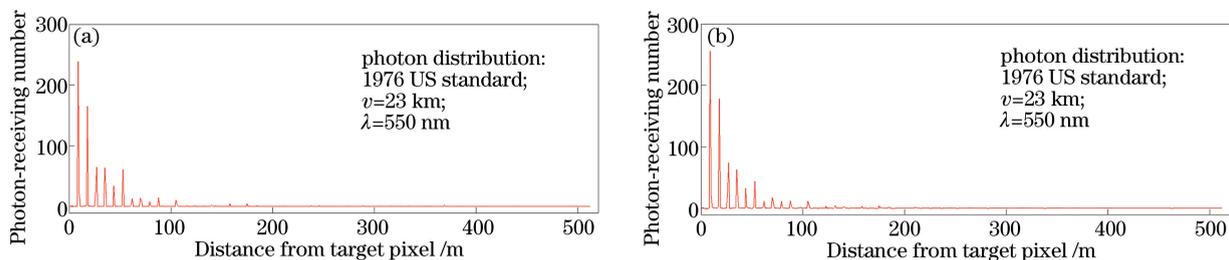


图 3 相同输入条件下采用不同方法所得结果的对比。(a)蒙特卡罗法;(b)神经网络法

Fig. 3 Results comparison between (a) Monte Carlo method and (b) neural network under the same input condition

使用余弦来衡量原始光子分布和对应条件下神经网络输出向量的差异性,即

$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} = \frac{\sum_{n=1}^{512} x_{1i} x_{2i}}{\sqrt{\sum_{n=1}^{512} x_{1i}^2} \sqrt{\sum_{n=1}^{512} x_{2i}^2}}, \quad (4)$$

式中向量 \mathbf{a} 是蒙特卡罗模拟所得到径向光子数目所组成的 512 维向量,向量 \mathbf{b} 是相同条件下神经网络输出所组成的 512 维向量, θ 是向量 \mathbf{a} 和向量 \mathbf{b} 之间的角度, x_{1i} 是向量 \mathbf{a} 的各个分量, x_{2i} 是向量 \mathbf{b} 的各个分量(这里的 $i=1\sim 512$),余弦的取值范围为 $[-1, 1]$ 。夹角余弦越大表示两个向量的夹角越小,反之则表示两

个向量的夹角越大。当两个向量的方向重合,即如果原始光子分布和神经网络获得的光子分布所组成的向量完全相同时,夹角余弦取最大值 1;当两个向量的方向完全相反时,夹角余弦取最小值-1。针对图 3 的输出结果,在相同输入条件下,蒙特卡罗模拟和神经网络输出的两个光子分布,其向量的余弦值为 0.9968,表示两个光子分布相似程度非常高,说明神经网络输出的正确率很高。计算所有 264 例样本,蒙特卡罗模拟和神经网络输出的两个光子分布,其向量的余弦值不小于 0.9447。

4.2 未知条件下的大气 PSF 预测

运用神经网络研究了诸多物理条件和 PSF 之间的关系,能很好地预测出训练样本中没有的条件下的光子分布。在输入样本的合理范围之内,随机生成 100 个与训练样本不同的输入条件作为神经网络的输入,通过训练后的神经网络获得光子分布,与蒙特卡罗模拟直接获得的光子分布进行对比,如图 4 所示,图 4(a)随机生成的神经网络输入是大气模型 2,能见度为 23.4 km,波长为 567 nm,图 4(b)随机生成的神经网络输入是大气模型 6,能见度为 5.5 km,波长为 543.2 nm。

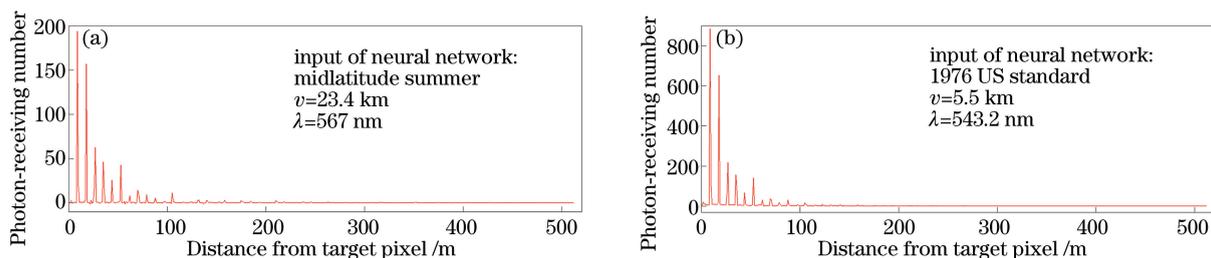


图 4 不同模型下训练数据中不包含的条件作为神经网络输入所得到的光子数。(a)模型 2;(b)模型 6

Fig. 4 Photon-receiving numbers under the exclusive conditions from training data for different models. (a) Model 2; (b) model 6

由图 4 可直观可以看出,由神经网络得到的光子数目随着离开目标像元距离的增加而减少,且径向光子数目的分布服从类似于一阶贝塞尔函数。对比图 4(a)和 4(b),随着能见度显著下降(v 从 23.4 km 变至 5.5 km),相同位置处所接收到的光子数目明显增多,说明能见度越低,由于大气散射作用偏离目标像元的光子数越多,符合经验判断。综上所述,可以看出神经网络有较好的泛化能力,能有效地计算并估计大气 PSF。

4.3 时间效率的对比

所使用的实验平台为台式计算机(4 CPU: Intel(R) Core(TM) i7-4770 CPU@3.40 GHz,内存: 8 G RAM,操作系统: Microsoft Windows 7 Ultimate)。实验显示,蒙特卡罗模拟、神经网络训练、神经网络计算所需的时间分别为 46.472, 96.130, 8.511 s。蒙特卡罗模拟是针对一种输入条件、发射 10^6 个光子获得大气 PSF 所需的模拟时间,如果提高模拟精度,则需要增加光子的数目,即相应地将会增加模拟时间;神经网络训练的时间是训练所有样本(这里是 264 个样本)生成模型的时间;神经网络计算的时间是给定一种输入条件生成 PSF 所用的时间。虽然神经网络训练所需的时间大于一次蒙特卡罗模拟的时间,但一旦网络训练完成,基于网络获得一次大气 PSF 的估计所需要的时间远远小于蒙特卡罗模拟的时间。

5 结 论

蒙特卡罗法是模拟大气 PSF 的常用算法。但该方法存在问题:每次改变条件都需要重新模拟且模拟时间过长,无法满足软件系统对算法的实时性要求;得到正确解的概率依赖模拟所用的时间,单次模拟结果没有规律性可言。基于人工神经网络对蒙特卡罗模拟方法进行改进,设计并实现一个多层前馈神经网络,获得了大气 PSF(由蒙特卡罗模拟获得)与其影响参数(通常包括大气参数、光谱参数和观测几何等众多物理量)之间的关系。训练之后的神经网络模型对于任意输入的条件参数,都可以得到相应的大气 PSF,对于与训练样本不同的输入条件,也能获得合理的大气 PSF 的估计。多次实验结果显示,这种优化方法的计算精度和效率均令人满意。

目前只利用一百万个光子,针对不同大气类型、在一定波长和能见度范围内模拟光子传输过程,仍有许多提升的空间。比如增加蒙特卡罗模拟的光子数目以提高模拟精度;减小波长和能见度的取值间隔、拓展波长范围等以扩大训练样本,训练出更加准确、泛化能力更强的神经网络模型。另外由于影响大气 PSF 的因

素很多,只考虑了影响程度较大的三种条件,其他因素(比如气溶胶类型、观测高度、观测天顶角、观测方位角等)都会对大气 PSF 产生不同程度的影响,需逐步进行考虑。尤其是当倾斜观测时,PSF 不再以目标像元为中心对称分布,这就要求对神经网络的输入和输出重新进行设计,后续将重点开展这方面的研究工作。

参 考 文 献

- Li Haiwei, Zhang Hao, Chen Zhengchao, *et al.* MTF computing method of remote sensing image after removing the effects of atmospheric adjacency effect[J]. *Acta Optica Sinica*, 2015, 35(3): 0328002.
李海巍, 张 浩, 陈正超, 等. 去除大气邻近效应影响的遥感图像 MTF 计算方法[J]. *光学学报*, 2015, 35(3): 0328002.
- Wang Qian, Chen Xue, Ma Jianwen, *et al.* A comparative study of two remote sensing image adjacency effect correction algorithms base on SHDOM empirical equation and synchronized measured spectral data[J]. *Acta Optica Sinica*, 2010, 30(11): 3342-3348.
王 倩, 陈 雪, 马建文, 等. 基于 SHDOM 经验方程和基于同步实测光谱数据的遥感影像邻近效应校正算法对比研究[J]. *光学学报*, 2010, 30(11): 3342-3348.
- Richter R, Bachmann M, Dorigo W, *et al.* Influence of the adjacency effect on ground reflectance measurements[J]. *IEEE Geoscience & Remote Sensing Letters*, 2006, 3(4): 565-569.
- Lyapustin A I, Kaufman Y J. Role of adjacency effect in the remote sensing of aerosol[J]. *Journal of Geophysical Research Atmospheres*, 2001, 106(D11): 11909-11916.
- Reinersman P N, Carder K L. Monte Carlo simulation of the atmospheric point-spread function with an application to correction for the adjacency effect[J]. *Applied Optics*, 1995, 34(21): 4453-4471.
- Jia G R, Zhao H J, Li N. Simulation of hyperspectral scene with full adjacency effect[C]. *IEEE International Geoscience and Remote Sensing Symposium*, 2008(3): III724-III727.
- Yang Guijun, Liu Qinhuo, Liu Qiang, *et al.* Adjacency effect analysis in imaging simulation of high-resolution mid-infrared ($3\sim 5\ \mu\text{m}$) remote sensing[J]. *Journal of Infrared and Millimeter Waves*, 2008, 27(3): 233-240.
杨贵军, 柳钦火, 刘 强, 等. 高分辨率中红外遥感($3\sim 5\ \mu\text{m}$)成像模拟中邻近效应分析[J]. *红外与毫米波学报*, 2008, 27(3): 233-240.
- Liu Chengyu, Chen Chun, Zhang Shuqing, *et al.* Atmospheric adjacency effect correction of ETM images [J]. *Spectroscopy & Spectral Analysis*, 2010, 30(9): 2529-2532.
刘成玉, 陈 春, 张树清, 等. ETM 图像大气邻近效应校正[J]. *光谱学与光谱分析*, 2010, 30(9): 2529-2532.
- Li Hongshun, Liu Wei. Analysis of the adjacency effect in satellite remote sensing by using backward Monte Carlo method [J]. *Journal of Huazhong University of Science & Technology (Nature Science Edition)*, 2004, 32(11): 1-3.
李宏顺, 刘 伟. 用逆向蒙特卡罗法分析卫星遥感中的邻近效应[J]. *华中科技大学学报(自然科学版)*, 2004, 32(11): 1-3.
- Ma Xiaoshan, Guo Xiaoyong, Meng Xin, *et al.* Simulation and analysis of the adjacency effect in earth-imaging process of the optical remote sensing[J]. *Journal of Infrared and Millimeter Waves*, 2015, 34(2): 250-256.
马晓珊, 郭晓勇, 孟 新, 等. 光学遥感对地成像过程中的邻近效应模拟分析[J]. *红外与毫米波学报*, 2015, 34(2): 250-256.
- Modest M F. Backward Monte Carlo simulations in radiative heat transfer[J]. *Journal of Heat Transfer*, 2001, 125(1): 57-62.
- Xu Xiru, Wang Pingrong. Computing atmospheric point spread function by Monte-Carlo method[J]. *Journal of Remote Sensing*, 1999, 3(4): 268-278.
徐希孺, 王平荣. 用蒙特-卡罗方法计算大气点扩散函数[J]. *遥感学报*, 1999, 3(4): 268-278.
- Cong B. Encoding neural networks to compute the atmospheric point spread function[C]. *The International Conference on Information Technology: Coding and Computing*, 2000: 344-349.
- Huang Li. BP neural algorithm improvement and application research[D]. Chongqing: Chong Qing Normal University, 2008.
黄 丽. BP 神经网络算法改进及应用研究[D]. 重庆: 重庆师范大学, 2008.
- Hagan M T, Menhaj M B. Training feed forward networks with the Marquardt algorithm[J]. *IEEE Transactions on Neural Networks*, 1994, 5(6): 989-993.