

冠层光谱的多项式分析及生化参量反演*

颜春燕** 刘 强 牛 铮 王纪华 黄文江 刘良云

(中国科学院遥感信息科学国家重点实验室, 北京 100101; 北京农业信息技术研究中心, 北京 100089)

摘要 提出了一种冠层光谱的多项式表达模型和分解方法, 并将它用于反演冠层多角度光谱数据获取生化参量. 首先叙述了多项式表达的基本公式, 并分析了多项式的 n 阶项及系数的物理意义, 在此基础上给出完整的多项式表达模型, 以及反演多项式系数的分解方法. 通过对 SAILH 模型模拟的冠层光谱的分解, 多项式表达不仅能很好地拟合冠层光谱, 而且揭示了冠层的每一次散射对总反射的贡献. 以一次散射系数 a_{10} 和 a_{01} 为例, 可以看到多项式系数很好地反映了冠层光谱中的热点现象及观测角度、叶面积指数和叶倾角的影响. 利用多项式表达与叶片模型 PROSPECT 的耦合, 建立了冠层生化参量反演模型. 分别利用两种不同的冠层模型(SAILH 模型和 4-SCALE 模型)模拟了冠层多角度光谱数据, 然后以多项式表达+PROSPECT 模型反演, 以提取生化参量, 都取得了理想的结果. 在此基础之上, 利用实测的玉米冠层多角度光谱数据, 也获得了较为准确的叶绿素含量反演结果. 说明多项式分析提供了不依赖于特定冠层模型而反演生化参量的一条新思路.

关键词 多项式表达 冠层光谱 分解 生化组分含量 反演

1 引言

植被是生态系统中最重要的重要组成部分之一, 正确估计植被的生化组分含量为了解不同尺度的生态系统功能提供了非常有用的帮助. 利用测量的叶片光谱提取叶片生化参量的方法已经相对比较成熟. 但是在冠层尺度上, 因为冠层光谱受到叶片生化参量、土壤背景、冠层结构等因素的共同影响, 诸多未知因素限制了生化参量的反演精度, 所以从冠层光

谱中提取生化参量仍然是具有挑战性的研究. 目前提取植被冠层生化参量的方法可分为经验半经验的统计回归方法和冠层模型耦合叶片模型的反演方法两类. 前者物理意义不明确, 建立的统计关系难以推广; 而后者受到模型的准确性与复杂性这一对矛盾的制约, 同时受敏感的冠层参数的限制, 因此目前也难以推广应用^[1].

遥感研究的地表通常是多种组分的复合结构,

2003-11-26 收稿, 2004-10-27 收修稿

* 国家重点基础研究发展规划项目(G2000077900)、中国科学院知识创新工程重大项目——中国陆地和近海生态系统碳收支研究(KZCX1-SW-01)和国家自然科学基金资助项目(40271086)

** E-mail: yanchunyan@263.net

其反射率是太阳/观测角度、组分光谱和像元结构的复杂函数。研究者提出了许多物理模型用于模拟复杂像元的光谱和方向特征,然而这些模型通常都非常复杂,以至于不能或不易准确地反演。多角度传感器的发展为重建像元结构提供了可能,但是由于许多模型都是在理想状态下建立的,而模型的假设经常与真实自然景观不太符合,因此冠层物理模型的应用受到了限制。

新一代的高光谱传感器可以提供更好的光谱分辨率和定标精度,甚至可以提供同一目标的多个观测角度下的光谱。但是,高光谱数据的分析方法仍然有待研究。线性混合像元分解仍然是当前许多应用中的主要手段^[2,3]。线性分解对于有三维结构的地表显然不是合适的方法,例如,多次散射在植被近红外波段是不可忽略的。

多角度与高光谱遥感数据提供了与冠层结构和生化参量有关的信息,但是如何提取这些信息则依赖于分析数据的方法。本文中利用高阶多项式来描述光在冠层中的多次散射过程,分析冠层光谱数据,在线性混合像元分解和复杂的物理模型之外,提供了解释组分光谱和冠层光谱间关系的另一条途径。多项式表达可以耦合叶片模型用于反演生化参量,与其他冠层物理模型与叶片模型的耦合相比,多项式表达很少依赖于对冠层结构的特定假设,对于均匀或是非均匀的冠层都能适用。

2 冠层光谱的多项式表达模型和分解方法

2.1 多项式表达的前向模型

当太阳/观测状态和结构参数确定的时候,冠层反射率就是组分光谱特性的函数。本文中只考虑叶片和土壤两个组分,它们是植被冠层中的主要光线散射体。组分光谱都具有二向反射特性,但是在冠层尺度的研究中,由于组分的倾角分布具有随机性,通常认为可以对组分光谱作朗伯假设。例如现有的多数冠层辐射传输模型都假设叶片的反射率和透过率分别是朗伯的,于是叶片光谱特性可由测量的叶片方向-半球反射率(ρ)和方向-半球透过率(τ)表述。对于多数植物而言, ρ 和 τ 的光谱形状相似,数值接近,是两个相关性非常强的参数,因此在模型中直接

使用 ρ 和 τ 并不是理想的参数化方法。本文中用 $\rho+\tau$ 和 $\rho-\tau$ 来参数化叶片光谱特性,其实质与 ρ 和 τ 是等价的,但是相互间更为独立一些。 $\rho+\tau$ 的物理意义是叶片对光线的总散射率,是决定冠层光谱的主要参数; $\rho-\tau$ 是叶片反射率和透过率之间的差异,也直接影响到冠层反射率的方向性,但是其作用主要表现在一次散射之中。对于土壤光谱特性,在朗伯假设下可用土壤反射率 r_s 来参数化。

从数学的角度看,任何连续函数都有它唯一的多项式近似,因此冠层反射率可通过多项式表达式来近似,多项式的自变量为组分光谱特性。我们以 x 代表叶片总散射率 $\rho+\tau$,以 y 代表土壤反射率 r_s ,在第一步近似中,我们忽略叶片反射率和透过率之间的差异 $\rho-\tau$,则冠层反射率 R 可用组分光谱特性 x, y 的多项式函数来表达:

$$R = P_1 + P_2 + \dots + P_{N-1} + P_N + \dots, \quad (1)$$

其中 R 代表冠层反射率, $P_n (n=1, 2, \dots, N, \dots)$ 代表多项式表达中的第 n 阶项,

$$\begin{aligned} P_1 &= a_{1,0}x + a_{0,1}y \\ P_2 &= a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 \\ &\dots \\ P_n &= a_{n,0}x^n + a_{n-1,1}x^{n-1}y + \dots + a_{1,n-1}xy^{n-1} + a_{0,n}y^n, \end{aligned} \quad (2)$$

a_{ij} 表示多项式的系数。

(1)式给出多项式表达的基本形式,在进一步修正该表达式之前,我们先来探讨一下它的物理意义。通过追踪光子散射过程,多项式表达式的物理意义不难理解。假设一个光子最多经历三次散射(更高次散射原理类推),一个单位能量光子到达冠层并经反射后到达传感器的所有可能列举如下:

(1) 以概率 p_1 碰到叶片(参考图 1);

如果碰到叶片,则经过第一次散射,其能量减为 x ,然后它可能:

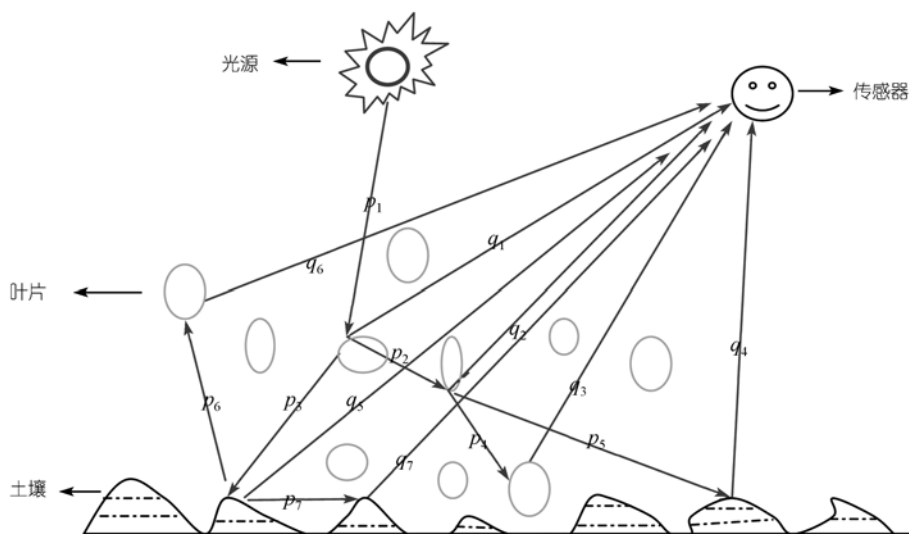
· 以概率 q_1 进入传感器;

· 以概率 p_2 再次碰到叶片;

如果再次碰到叶片,经过第二次散射,其能量减为 x^2 ,然后它可能:

() 以概率 q_2 进入传感器;

() 以概率 p_4 再次碰到叶片,经过第三次散射,

图1 单位能量光子如果以概率 p_1 碰到叶片后的可能去向

其能量减为 x^3 , 最后以概率 q_3 进入传感器;

() 以概率 p_5 碰到土壤, 经过第三次散射, 其能量减为 x^2y , 最后以概率 q_4 进入传感器;

. 以概率 p_3 碰到土壤;

如果碰到土壤, 经过第二次散射, 其能量减为 xy , 然后它可能:

() 以概率 q_5 进入传感器;

() 以概率 p_6 再次碰到叶片, 经过第三次散射, 其能量减为 x^2y , 最后以概率 q_6 进入传感器;

() 以概率 p_7 再次碰到土壤, 经过第三次散射, 其能量减为 xy^2 , 最后以概率 q_7 进入传感器;

(2) 以概率 p'_1 碰到土壤;

如果碰到土壤, 则经过第一次散射, 其能量减为 y , 然后它可能:

. 以概率 q'_1 进入传感器;

. 以概率 p'_2 再次碰到土壤;

如果再次碰到土壤, 经过第二次散射, 其能量减为 y^2 , 然后它可能:

() 以概率 q'_2 进入传感器;

() 以概率 p'_4 再次碰到土壤, 经过第三次散射, 其能量减为 y^3 , 最后以概率 q'_3 进入传感器;

() 以概率 p'_5 碰到叶片, 经过第三次散射, 其能量减为 xy^2 , 最后以概率 q'_4 进入传感器;

. 以概率 p'_3 碰到叶片

如果碰到叶片, 经过第二次散射, 其能量减为 xy , 然后它可能:

() 以概率 q'_5 进入传感器;

() 以概率 p'_6 再次碰到土壤, 经过第三次散射, 其能量减为 xy^2 , 最后以概率 q'_6 进入传感器;

() 以概率 p'_7 再次碰到叶片, 经过第三次散射, 其能量减为 x^2y , 最后以概率 q'_7 进入传感器;

综上所述, 经过三次散射后, 最终到达传感器的能量为

$$Q = p_1 x (q_1 + p_2 x (q_2 + p_4 x q_3 + p_5 y q_4) + p_3 y (q_5 + p_6 x q_6 + p_7 y q_7)) + p_1' y (q_1' + p_2' y (q_2' + p_4' y q_3' + p_5' x q_4') + p_3' x (q_5' + p_6' y q_6' + p_7' x q_7'))$$

将该多项式展开并按一次项, 二次项, 三次项合并得

$$Q = (p_1 q_1 x + p_1' q_1' y) + (p_1 p_2 q_2 x^2 + (p_1 p_3 q_5 + p_1' p_3' q_5') y x + p_1' p_2' q_2' y^2) + (p_1 p_2 p_4 q_3 x^3 + (p_1 p_2 p_5 q_4 + p_1 p_3 p_6 q_6 + p_1' p_3' p_7' q_7') x^2 y + (p_1 p_3 p_7 q_7 + p_1' p_2' p_5' q_4' + p_1' p_3' p_6' q_6') x y^2 + p_1' p_2' p_4' q_3' y^3) \quad (3)$$

按照方向反射率的定义, 冠层的方向反射率 R 等

于冠层辐射传输后进入传感器的能量 Q 与相同光照条件下被理想漫反射体反射进入传感器的能量的比. 对比(3)和(2)式可知: 多项式的 n 阶项表示 n 次散射对冠层反射率的贡献, 例如一阶项 P_1 代表冠层的一次散射反射率; 多项式系数 a_{ij} 正比于光线经过 $i+j$

$$\begin{aligned} a_{1,0} &\propto p_1 q_1 & a_{0,1} &\propto p_1' q_1' \\ a_{2,0} &\propto p_1 p_2 q_2 & a_{1,1} &\propto p_1 p_3 q_3 + p_1' p_3' q_3' & a_{0,2} &\propto p_1' p_2' q_2' \\ a_{3,0} &\propto p_1 p_2 p_4 q_3 & \cdots & & a_{0,3} &\propto p_1' p_2' p_4' q_3'. \end{aligned} \quad (4)$$

事实上多项式表达不可能考虑到无穷高阶项, 随着阶数的增大, 多项式的系数越多, 多项式也越复杂. 通常我们假设存在着某个阶数 N , 高于 N 阶的项可以忽略、或可作简化处理. 本文中我们把大于或等于 N 阶的项合并表示为

$$P'_N = P_N + P_{N+1} + \cdots. \quad (5)$$

为了推导 P'_N 的近似表达式, 我们假设经过 N 次散射后, 光线可认为是各向同性的, 这时 $n+1$ 次与 n 次散射的比恒定, 即 $P_{n+1}/P_n = S$ (当 $n \geq N$). 若令 s_1 和 s_2 分别代表各向同性光在冠层中被叶片和土壤截取的概率, 根据对多项式物理意义的分析, S 可通过下式计算:

$$S = s_1 x + s_2 y, \quad (6)$$

那么所有大于或等于 N 阶项的和为

$$P'_N = P_N / (1-S). \quad (7)$$

另外, 前面的讨论中忽略了叶片反射率和透射率之间的差异 $\rho - \tau$. 如前所述, $\rho - \tau$ 在数值上比较小, 其作用主要表现在光线的一次散射中, 因此多项式表达中只考虑其一次项 $\delta(\rho - \tau)$, δ 也是一个多项式系数, 忽略高次项. 于是多项式表达的完整形式为

$$R = P_1 + P_2 + \cdots + P_{N-1} + P'_N + \delta(\rho - \tau). \quad (8)$$

2.2 多项式分解方法

多项式表达的系数虽然都有其物理意义, 但是目前缺乏正演计算这些系数的快速方法, 而用微分分析法(例如光线跟踪)去计算数目众多的多项式系数是不现实的. 本文更倾向于将多项式表达作为一个反演模型使用, 即所有多项式系数都由光谱数据反演得出. 假设冠层光谱和组分光谱(即 R, ρ, τ, r_s)都可以由测量得到, 对于有 K 个波段的高光谱数据, (8)式

次散射后进入传感器的总概率, 亦即从光源出发, 被叶片散射 i 次且被土壤散射 j 次后到达传感器的所有微观路径的概率之和. 因为这个比例系数恒定, 所以在不引起混淆的情况下, 我们认为多项式系数就代表了这个概率.

为有 K 个方程的方程组, 其中有 $N(N+3)/2+1$ 个未知的多项式系数, N 为多项式分解的最高阶数. 在反演多项式系数时, 我们把 s_1, s_2 固定在某个估计值, 则公(8)式对于多项式系数来说是线性方程, 可用矩阵方程表示为

$$R = CA, \quad (9)$$

式中 R 为冠层光谱, 为 K 维矢量, R_k 表示第 k 个波段的冠层反射率值; A 为 $N(N+3)/2+1$ 维的多项式系数矢量, A_1 到 A_N 分别为 1 阶项 P_1 到 N 阶项 P_N 的系数矢量. 组分光谱矩阵 C 是一个 K 行和 $N(N+3)/2+1$ 列的矩阵:

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_K \end{bmatrix}, A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \\ \delta \end{bmatrix}, A_1 = \begin{bmatrix} a_{1,0} \\ a_{0,1} \end{bmatrix}, \dots, A_N = \begin{bmatrix} a_{N,0} \\ a_{N-1,1} \\ \vdots \\ a_{1,N-1} \\ a_{0,N} \end{bmatrix},$$

$$C = \begin{bmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,N-1} & C_{1,N}/S_1 & \rho_1 - \tau_1 \\ C_{2,1} & C_{2,2} & \cdots & C_{2,N-1} & C_{2,N}/S_2 & \rho_2 - \tau_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ C_{K,1} & C_{K,2} & \cdots & C_{K,N-1} & C_{K,N}/S_K & \rho_K - \tau_K \end{bmatrix},$$

其中

$$\begin{aligned} C_{k,1} &= [x_k \quad y_k]; \\ C_{k,2} &= \begin{bmatrix} x_k^2 & x_k y_k & y_k^2 \end{bmatrix} \\ &\cdots \\ C_{k,N} &= \begin{bmatrix} x_k^N & x_k^{N-1} y_k & \cdots & x_k y_k^{N-1} & y_k^N \end{bmatrix} \\ S_k &= s_1 x_k + s_2 y_k, \quad (k = 1, \dots, K). \end{aligned}$$

(9)式虽然繁琐, 其实质为关于多项式系数矢量 A 的线性方程. 这样, 多项式系数的反演问题简化为线性方程组求解. 本文讨论的应用都是针对高光谱

数据的,可保证方程个数大于未知数个数,即 $K > N(N+3)/2+1$, 因此可以用线性最小二乘法求解多项式系数。

s_1 和 s_2 只影响到高阶项的估算,是模型中的不敏感参数,可以固定在一个比较合理的典型值。不过,为了提高反演多项式系数的精度,本文的工作采用迭代逐步逼近的方法估算 s_1 和 s_2 , 步骤如下:

() 给 s_1 和 s_2 赋初值(应该满足 $s_1+s_2 < 1$);

() 将 s_1 和 s_2 代入(9)式,利用最小二乘法求解出多项式系数;

() 计算 P_N/P_{N-1} 得到 S , 然后根据(6)式估算新的 s_1 和 s_2 ;

() 重复步骤(), ()直到收敛或最大叠代次数达到。

多项式系数的反演是一个多未知数的方程求解问题,如果在反演时单纯依靠对数据的拟合,则当数据中存在噪声时可能会因过度拟合数据反而得到不符合物理意义的解。特别是在使用迭代逐步逼近估计 s_1 和 s_2 时,更增加了反演结果的不稳定性。

根据多项式系数的物理意义,多项式系数均为小于1的非负数, δ 则在-1和1之间。因此,在实际求解时,我们利用 Matlab 提供的带约束的线性最小二乘法工具(Lsqlin 函数),将(9)式的解约束在其值域区间内,并在迭代逐步逼近时将 s_1 和 s_2 也约束在0和1之间。使用约束后,反演结果增加了对噪声的抵抗能力。

至此,列出了多项式模型的公式表达和反演多项式系数的分解方法。下面用模拟的冠层光谱数据反演多项式系数,以检验多项式分解算法的稳定性,并结合其物理意义分析反演得到多项式系数与冠层结构、观测角度的关系。

2.3 多项式分解结果的检验和分析

冠层BRDF的辐射传输模型SAIL^[4]描述了在水平均匀冠层中直射和上行、下行散射光通量的辐射传输过程。它是使用最为广泛的冠层二向反射物理模型之一,模型参数包括:叶片反射率 $\rho(\lambda)$ 和透过率 $\tau(\lambda)$,叶面积指数LAI,平均叶倾角ALA,土壤反射率 $r_s(\lambda)$ 以及太阳辐射的散射分量skyl(通过水平能见

度vis来计算)。光线入射和观测的几何关系通过太阳天顶角 θ_s 、太阳方位角 φ_s 、观测天顶角 θ_v 和观测方位角 φ_v 来描述。Kusck^[5]对SAIL模型做了改进,考虑了热点效应,引进了描述热点形状的参数 s ,改进后的模型随之称为SAILH。

我们利用SAILH模型(以及随模型代码给出的典型叶片和土壤光谱)模拟了一系列冠层光谱,然后对这些冠层光谱做多项式分解。模拟采用的缺省参数设置为: $LAI = 1.5^\circ$, $ALA = 45^\circ$, $\theta_v=0$, $\theta_s=30^\circ$, $\theta_v=0^\circ$, $\varphi_s=0^\circ$, $s = 0.1$, $vis = 50 \text{ km}$ 。当分析多项式系数与冠层参数的关系时,LAI、平均叶倾角和观测角度分别从0.3~8, $5^\circ \sim 85^\circ$ 以及主平面上 $-80^\circ \sim 80^\circ$ 变化,其他参数都使用缺省值。

首先我们分析了多项式表达对冠层光谱的拟合能力:在前面描述的参数变化范围内随机生成了1000条冠层光谱曲线,不考虑噪声的情况下,使用 $N=5$ 的多项式表达式进行分解,统计得到拟合残差(RMSE)的平均值是0.00035。可见多项式表达的数据拟合能力非常强。

接下来分析多项式各阶项的意义。按照前面的分析,一阶项表示一次散射对冠层反射率的贡献,(8)式中的一阶项为 $P_1 + \delta(\rho - \tau)$,高阶项 P_n 则表示 n 次散射对冠层反射率的贡献。反演多项式系数之后,就可直接计算冠层反射率中的一次乃至多次散射分量。例如图2是缺省参数设置下冠层反射率中的各次散射分量,黑点为SAILH模拟的冠层反射率,短虚线为多项式拟合的曲线,实线条1~4依次为一次至四次散射分量,长虚线为大于或等于五次的散射分量。可以看到:多项式拟合线条与SAILH模拟的反射率光谱基本重合,一次散射占了冠层光谱的大部分,二次散射次之,高次散射项的贡献依次减少。

多项式表达中的一阶项系数代表了光在冠层中的一次散射,是冠层反射率中最重要的部分,同时也为简单清晰起见,下面以一次散射系数(也即 $a_{1,0}$ 和 $a_{0,1}$ 和 δ)为例来分析多项式系数与冠层结构参数及观测角度间的关系,如前所述, $a_{1,0}$ 代表光线首先到达叶片并被反射入传感器的概率, $a_{0,1}$ 代表光线首先到达土壤并被反射入传感器的概率, δ 则反映了经过叶片

一次散射到达传感器的光线中, 反射的光线比透射光线多的量.

图 3 描绘了主平面内一阶项系数随观测角度的变化状况. 图中多项式系数 $a_{1,0}$ 和 $a_{0,1}$ 随着观测角度接近热点而迅速增大, 说明热点附近叶片和土壤的一次散射都增多; 而同时 $a_{1,0}$ 随着观测天顶角增加而增大, 说明倾斜观测中叶片的一次散射分量增加, 而 $a_{0,1}$ 则表现相反; δ 在后向和前向观测天顶角较小时都比较大, 说明被叶片反射的光线占经叶片一次散射到达传感器的光线中的主要部分, 而在前向大天顶角时 δ 下降很快, 说明经叶片一次透射而到达传感器的光线比例增大了.

图 4 描绘了一阶项系数随叶面积指数的变化, 可

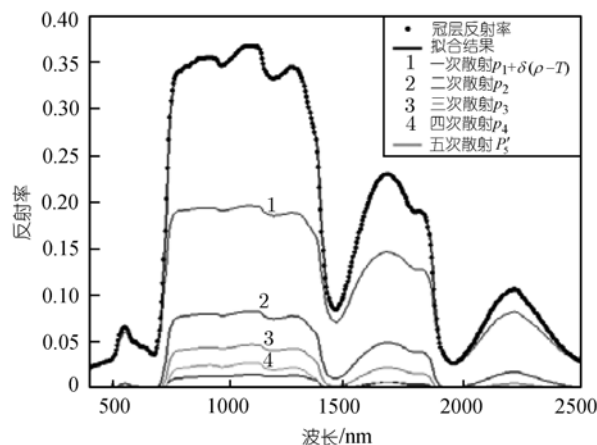


图 2 冠层光谱的多项式拟合及分解

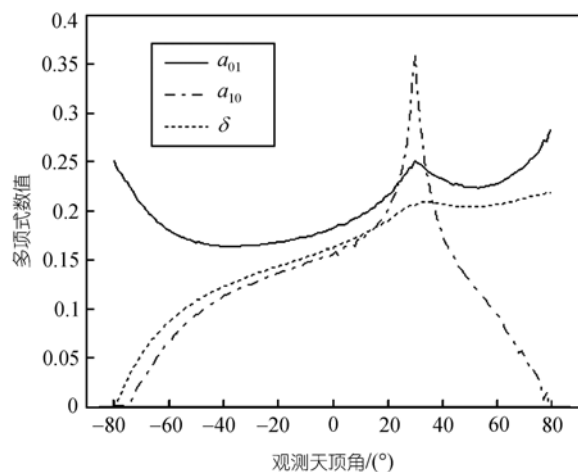


图 3 主平面内一次散射系数随观测角度的变化

以看到, $a_{1,0}$ 随 LAI 的增加而逐渐增大, 但开始时增加迅速, 随后减慢并有饱和趋势; 而 $a_{0,1}$ 随 LAI 增大而减小直至为零. 说明 LAI 的增加使得被叶片散射的光线增多, 被土壤散射的光线减少, 当 LAI 增加到一定程度, 视场内全部为叶片覆盖, 这时光线到达土壤并被反射入传感器的概率为零, 而被叶片反射的概率也趋于最大值.

图 5 描绘了 $a_{1,0}$ 和 $a_{0,1}$ 随平均叶倾角 ALA 的变化状况. 可以看到, $a_{1,0}$ 随 ALA 增加而逐渐减小, 而 $a_{0,1}$ 情况的相反. ALA 的增大表明叶片的直立倾向, 因此视场内叶面积减小, 土壤面积增大, 光线到达叶片并被反射入传感器的概率减小, 土壤反之.

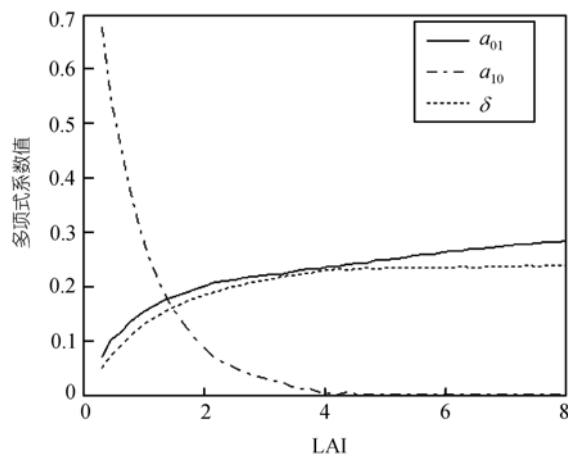


图 4 一次散射系数随 LAI 的变化

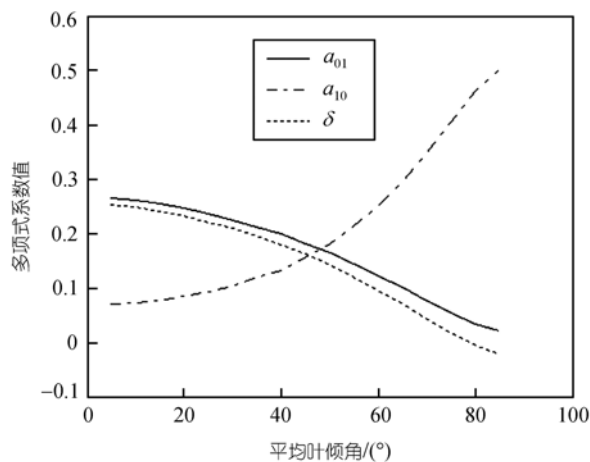


图 5 一次散射系数随平均叶倾角的变化

在上述几例分析中，反演多项式系数的表现都是与常理吻合的。需要说明的是，多项式分解在例子中扮演的只是反演模型的角色，实际上是 SAILH 模型模拟了冠层方向反射率与各个模型参数间的关系，而多项式分解只是如实地将这种关系分解并反映到多项式系数上。事实上如果 SAILH 模型存在不足，多项式分解也将如实地反映。这里的分析的目的在于说明多项式表达的前向模型能够兼容 SAILH 模型，而分解算法成功地解释了 SAILH 模型模拟的结果。

因此我们认为对于 SAILH 模型模拟的冠层光谱，在没有添加噪声的情况下，用多项式分解方法反演得到的多项式系数是准确的。在下面的稳定性分析中，我们对模拟的冠层光谱添加标准差为 0.1%到 10%的乘性高斯噪声，然后比较反演结果与准确值的差异。表 1 是对 1000 个随机样本比较的结果，我们对一阶项到五阶项的系数分别统计均方根误差(RMSE)。从表中看，多项式 n 阶项系数的均值随 n 的增加下降很快，虽然随机噪声引起各阶多项式系数的 RMSE 都基本上在一个数量级上，但是从相对误差来说，高阶多项式系数对数据中噪声是非常敏感的，而低阶项系数却相对稳定。对于 0.1%的随机噪声，基本上各阶项的系数都能准确的反演，而当噪声水平达到 1%时，一阶项系数仍然能够比较准确的反演，但是高阶项系数的误差已经接近信号的数量级。

3 冠层生化参量反演

冠层生化参量的反演通常是利用冠层模型耦合叶片模型，通过最小化预测光谱和实测光谱的差别来估算叶片模型中的生化参量。近年研究比较普遍并在部分地区得到尝试性应用的如反演SAILH模型

+ PROSPETC模型提取生化参量^[6,7]，此外也见到其他模型耦合PROSPECT反演生化参量的研究^[8]。目前利用反演冠层与叶片耦合模型提取植被生化参量存在的主要问题是：(1)前向模型的准确性和可反演性是一对矛盾：准确的模型往往非常复杂(如真实景观模型)，是不可反演的，而可反演的模型往往依赖于对冠层结构的某种假定而简化了辐射传输过程，其适用范围受到所依赖的假定的限制；(2)未知的冠层参数限制了生化参数的反演精度：叶片模型中的生化参数相比较多数冠层参数都是不敏感参数，已有研究表明敏感的冠层参数对生化参数反演的影响不可忽视；(3)在生化参量反演中使用多角度信息的工作还少见报道。

本文提出的冠层光谱的多项式表达对冠层结构只做了很少的假定，因此对多种不同结构的冠层均能适用，经验证，作为一个反演模型，多项式表达与常见的几个 BRDF 物理模型都是兼容的。另一方面，虽然说多项式表达中有大量未知参数，但是因为多项式表达是一个线性模型，而多项式系数又有明确的物理意义，所以我们可以根据系数的物理意义对其进行约束，增加了反演结果抵抗噪声的鲁棒性。因此，利用多项式表达耦合叶片模型，既能够有效利用遥感数据中的多角度和高光谱信息，又不依赖于对冠层结构的特定假设，可以直接应用于冠层生化参量的反演。

3.1 多项式表达与叶片模型的耦合的反演方法

PROSPECT^[9]是一个基于平板模型的叶片光谱模型，它表述了植株叶片从 400~2500 nm的反射和透过程。它通过下式将总吸收系数 K 与生化组分

表 1 光谱数据中的随机噪声对反演多项式系数的影响

反演多项式系数的误差 (RMSE)						
	系数均值	0.1%噪声	0.5%噪声	1%噪声	5%噪声	10%噪声
一阶项	0.1555	0.0012	0.0041	0.0096	0.0539	0.0690
二阶项	0.0590	0.0069	0.0226	0.0255	0.0325	0.0345
三阶项	0.0338	0.0052	0.0168	0.0192	0.0260	0.0278
四阶项	0.0228	0.0034	0.0119	0.0136	0.0181	0.0192
五阶项	0.0156	0.0025	0.0099	0.0111	0.0136	0.0147

含量 C_i 相联系.

$$K = k_e + \sum_i \frac{C_i \cdot k_i}{VAI}, \quad (10)$$

式中, k_e 为叶片的白化吸收, 可以看作常量; k_i 为组分吸收系数, 也可通过实验室光谱库获得; C_i 为组分含量, 本文考虑两种生化组分: 叶绿素含量 cab ($\mu\text{g}/\text{cm}^2$) 和水分含量 cw (cm). VAI 为叶片结构参数, 通常作为未知参数反演.

土壤参数不是本文讨论的重点, 我们对土壤光谱模型做简单处理, 假定下垫面光谱是干土光谱和湿土光谱的线性混合, 而实验区的干土光谱和湿土光谱可由测量得到, 混合比 α 也作为未知数反演.

将多项式表达模型与叶片和土壤光谱模型耦合, 就在冠层光谱与组分生化理化参数之间建立了模型描述. 在这个耦合模型中, 多项式系数是由冠层结构和观测条件共同决定的, 所以, 不同观测角度的多项式系数不同, 假设有 M 个角度的冠层光谱数据, 那么每一个角度有 $N(N+3)/2+1$ 个多项式系数. 与多项式系数不同, 叶片模型和土壤模型的参数与观测角度无关的; 另外, 根据 s_1 和 s_2 的物理意义, 它们也是与观测角度无关的. 所以共有 6 个与观测角度无关的未知参数, 分别是: VAI , cab , cw , α 以及 s_1 和 s_2 . 所以耦合模型总共有 $M(N(N+3)/2+1)+6$ 个未知参数, 假设观测数据有 K 个波段, 则方程数是 MK 个. 为了求解这个多未知数的方程组, 我们采用优化代价函数的反演方法, 代价函数由两部分构成: 一是预测光谱拟合实测光谱的残差, 见(11)式; 二是对参数的约束, 包括对生化参量的均值和标准差的先验知识的约束和对多项式系数的边界约束, 前者通过贝叶斯推论可写出表达式, 后者则通常依靠算法程序中的特殊设计来实现^[10].

$COST$

$$= \sum_k \sum_m \left(R_{m,k}^{obs} - R_{m,k}(A_m, cab, cw, VAI, \alpha) \right)^2 + \text{约束项}, \quad (11)$$

式中 $R_{m,k}^{obs}$ 表示多项式拟合光谱; A_m 表示第 m 个角度的多项式系数矢量; 其他参数定义如前.

因为当未知数个数很多时, 求解非线性优化问题非常耗时, 所以我们不能简单地直接使用通用非线性优化工具, 而必须针对问题的特点进行求解. 注意到大多数未知数都是多项式系数, 而根据我们前面的讨论, 多项式分解只需使用线性优化, 所以我们设计了嵌套的优化方法: 外层优化针对 6 个不随观测角度变化的参数, 是非线性的, 采用 MATLAB 提供的 `fminunc` 函数, 它实际上使用了 quasi-Newton^[11] 算法; 内层优化针对每一个观测角度的多项式系数, 为线性优化, 采用 MATLAB 提供的 `lsqlin` 函数.

3.2 模型模拟数据的反演

为了模拟不同结构的冠层, 我们选取了两个不同的前向冠层模型: 辐射传输模型 SAILH 代表水平均匀冠层; 几何光学模型 4-Scale^[12] 代表离散的森林冠层. 分别用这两个模型耦合 PROSPECT, 并设置了几种不同的叶面积指数, 代表稀疏和浓密的冠层. 观测角度考虑了主平面内前向 20° , 40° , 60° , 后向 20° , 40° , 60° , 垂直 0° 共 7 个角度. 冠层模型的其它参数设在典型值. PROSPECT 模型中, 叶绿素含量从 $20 \mu\text{g}/\text{cm}^2$ 变化到 $80 \mu\text{g}/\text{cm}^2$, 水分含量从 0.005 cm 变化到 0.040 cm , 共有 21 个样本. 这 21 个叶片光谱与冠层模型耦合, 模拟了冠层二向反射光谱, 并且加入 0.1% 的乘性随机噪声.

利用本文的多项式分解耦合 PROSPECT 模型, 利用(11)式反演这些模拟数据得到的叶绿素和水分含量如表 2. 本文所考虑的重点是叶片生化参量, 因此其它参数的反演结果这里没有列出. 从表 2 可见无论对于利用 SAILH 模型, 还是利用 4-scale 模型模拟得到的冠层多角度 + 高光谱数据, 对于不同覆盖度的冠层, 利用多项式表达来反演冠层叶绿素和水分含量都取得了理想的结果, 平均准确率大于 90%.

3.3 实测数据的反演

2003 年 7 月到 9 月间在北京市农林科学院院内进行了夏玉米的光谱和生化参量观测系列实验, 具体测量的内容和方法¹⁾. 其中于 8 月 29 日和 9 月 8

1) 颜春燕. 高光谱遥感提取植被生化信息方法与模型研究. 中国科学院博士学位论文. 2003

表 2 利用 SAILH 和 4-scale 模拟的冠层多角度数据反演的叶绿素和水分含量结果

No.	真实值		反演 SAILH 模拟数据				反演 4-scale 模拟数据			
			参数 1(LAI=1.5)		参数 2(LAI=3)		参数 1(LAI=1.6)		参数 2(LAI=4.8)	
	叶绿素	水分	叶绿素	水分	叶绿素	水分	叶绿素	水分	叶绿素	水分
1	20	0.0115	20.27	0.0116	19.43	0.0113	22.86	0.0131	19.83	0.0117
2	25	0.0115	25.30	0.0116	24.25	0.0112	28.51	0.0131	24.99	0.0117
3	30	0.0115	30.44	0.0116	29.10	0.0112	33.92	0.0130	29.91	0.0116
4	35	0.0115	35.61	0.0116	33.98	0.0113	39.01	0.0128	34.74	0.0116
5	40	0.0115	40.86	0.0116	38.91	0.0113	43.90	0.0126	39.47	0.0115
6	45	0.0115	46.13	0.0117	43.78	0.0113	48.73	0.0125	44.32	0.0115
7	50	0.0115	51.49	0.0117	48.78	0.0113	53.69	0.0124	49.15	0.0115
8	55	0.0115	56.79	0.0117	53.88	0.0113	58.63	0.0123	53.99	0.0114
9	60	0.0115	62.15	0.0118	58.94	0.0114	63.64	0.0123	58.93	0.0114
10	65	0.0115	67.46	0.0118	64.01	0.0114	68.62	0.0122	63.93	0.0114
11	70	0.0115	59.89	0.0103	69.16	0.0114	73.75	0.0122	68.90	0.0114
12	75	0.0115	78.07	0.0118	74.23	0.0114	78.87	0.0122	73.83	0.0114
13	80	0.0115	83.40	0.0118	79.30	0.0114	83.99	0.0121	78.80	0.0114
14	48.6	0.005	49.64	0.0051	47.59	0.0050	51.08	0.0053	46.66	0.0049
15	48.6	0.01	49.92	0.0102	47.41	0.0098	51.91	0.0107	47.46	0.0099
16	48.6	0.015	49.84	0.0152	47.31	0.0147	53.31	0.0164	48.59	0.0152
17	48.6	0.02	49.34	0.0201	46.96	0.0194	54.34	0.0221	49.79	0.0206
18	48.6	0.025	49.09	0.0251	46.44	0.0240	54.55	0.0276	51.25	0.0264
19	48.6	0.03	49.06	0.0301	46.41	0.0289	54.62	0.0329	52.24	0.0322
20	48.6	0.035	48.89	0.0350	46.55	0.0338	54.31	0.0380	52.62	0.0376
21	48.6	0.04	48.98	0.0401	45.93	0.0405	51.10	0.0405	49.47	0.0402
RMSE / 均值			0.054	0.020	0.028	0.031	0.084	0.095	0.032	0.052

日选择了五个田块进行了冠层的多角度光谱测量和准同步的叶绿素和水分含量测量。光谱测量使用美国 ASD 公司产的 FieldSpec 光谱仪, 测量时分别在主平面内选择了 5 个角度, 主平面的垂直平面选择了 4 个角度共 9 个角度。因为是手持式观测, 具体的观测角度数只是一个参考值。实际上, 本文提出的多项式分解方法并不需要知道观测角度的具体数值, 它需要的只是多角度 + 高光谱数据本身提供的多维信息。于是我们利用这些多角度光谱数据, 去除水汽吸收波段后, 根据(11)式反演了这 10 个观测点处的冠层叶绿素含量和水分含量。

反演结果中发现水分含量反演的结果与实测值相比有系统偏差, 比实测值总体偏高。我们认为这可

能是将植株从户外取样后失水导致, 这是我们今后测量中值得注意的地方。表 3 列出了叶绿素含量的反演结果。从表 3 可以看到反演的叶绿素含量与真实值的误差大部分在可以接受的水平, 平均相对误差为 0.1478, 结果还是比较理想的。

4 讨论与结论

从前面章节的理论分析和实验验证中可以看到, 本文提出的多项式表达利用高阶多项式来描述光在冠层中的多次散射过程, 不仅能够很好地拟合冠层光谱, 而且多项式系数具有明确的物理意义。通过对一次散射系数 $a_{1,0}$ 和 $a_{0,1}$ 的具体分析, 可以看到, 多项式系数对于传统冠层模型中的热点现象、观测角度、

表 3 实测多角度光谱数据反演叶绿素含量的结果

观测日期	2003 年 8 月 29 日					2003 年 9 月 8 日				
地块号	12	14	23	24	25	12	14	23	24	25
测量值	57.68	84.48	70.09	65.52	59.08	45.77	59.78	46.78	64.94	80.93
反演值	63.20	66.43	70.12	70.07	72.16	55.77	61.69	51.76	66.25	66.69

LAI 及叶片倾角的影响都能够如实地反映.

通过用两种不同的冠层模型——辐射传输模型 SAILH 和几何光学模型 4-scale, 耦合不同叶绿素和水分含量下的叶片光谱得到模拟的冠层光谱数据, 利用多项式表达耦合叶片模型 PROSPECT 反演冠层生化参量, 得到了满意的结果. 此外, 反演实测的玉米冠层多角度光谱, 也获得了比较理想的叶绿素含量反演结果. 说明多项式表达可以充分利用多角度 + 高光谱信息, 不仅物理意义明确, 避免了统计模型的任意性, 同时又回避传统方法反演冠层生化参量存在的问题及冠层模型中诸多敏感参数, 为生化参量反演提供了一条新思路.

需要说明的是, 虽然我们强调多项式表达相比其它冠层反射率物理模型的优点在于对冠层结构只做了很少的假设, 然而多项式表达仍然需要依赖一些基本假设, 它们是: (1) 冠层是由叶片和土壤两种组分构成的; (2) 土壤反射率可做朗伯假定, 叶片光谱特性可用 ρ 、 τ 表述, 而 ρ 与 τ 的差别只影响到光线的一次散射项; (3) 进入冠层的光线经过一定次数散射后, 可以认为是各向同性的. 多项式表达很容易扩展到两种以上的组分, 组分的光谱特性也可以用更多的变量来描述, 因而多项式表达从原理上可以摆脱上述第一、二个假设, 而在使用更高阶的多项式后第三个假设也是通常合理的. 但是从稳定性分析结果来看, 多项式分解作为一个多参数的反演算法, 其高阶项系数对数据中的噪声是相当敏感的, 反演 5 阶的多项式系数对数据噪声水平的要求大约是 0.5%~0.1%, 这已经接近现有的室外光谱测量方法的极限, 而扩展的多项式表达必然会对光谱数据的准确性有更高的要求. 所以说本文方法的应用目前还受到光谱数据测量精度的制约.

本文提出的生化参量反演方法需要高质量的多角度 + 高光谱数据, 由于现有传感器性能的限制, 该方法目前还不能用到遥感图像数据. 但是, 在地面实验测量中获取满足本文方法要求的数据是可行的, 用实测数据所做的反演证明了这一点. 另外, 高光谱和多角度是遥感数据获取技术的发展方向, 国际上已有计划研制实用化的星载多角度 + 高光谱传感器,

实际上一些试验系统已经存在, 例如欧洲 2001 年发射的 PROBA 卫星上的 CHRIS 传感器就能实现 5 个角度和多达 64 波段的观测.

总之, 本文介绍的只是一个系统工作的起步. 作为一种新的光谱分析工具, 多项式表达模型及分解方法还有很多待改进之处, 相应的也还有很多别的潜在应用, 比如提取冠层结构参数. 我们利用多项式分解方法提取冠层生化参量的工作只是一个初步尝试, 随着相关理论研究的深入展开和遥感数据获取技术的飞速发展, 本文提出的方法才能逐渐实用化.

参 考 文 献

- 1 Zarco-Tejada P J, Miller J R, Noland T L, et al. Scaling-Up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 2000, 39(7): 1491~1507[DOI]
- 2 Bayliss J, Gualtieri J A, Cromp R F. Analyzing hyperspectral data with independent component analysis. *Proc SPIE*, 1998, 3240: 133~143[DOI]
- 3 Keshava N, Mustard J F. Spectral unmixing. *IEEE Signal Processing*, 2002, 19(1): 44~57[DOI]
- 4 Verhoef W. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sensing of Environment*, 1984, 16: 125~141[DOI]
- 5 Kuusk A. The hot spot effect of a uniform vegetative cover. *Sov Journal of Remote Sensing*, 1985, 3: 645~658
- 6 Jacquemoud S. Inversion of the PROSPECT+SAIL canopy reflectance model from AVIRIS equivalent spectra: theoretical study. *Remote Sens Environ*, 1993, 44, 281~292[DOI]
- 7 Jacquemoud S, Baret F, Andrieu B, et al. Extraction of vegetation biophysical parameters by inversion of the PROSPECT+SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sens Environ*, 1995, 52: 163~172[DOI]
- 8 Demarez V, Gastellu-Etchegorry J P. A modeling approach for studying forest chlorophyll content. *Remote Sens Environ*, 2000, 71: 226~238[DOI]
- 9 Jacquemoud S, Baret F. PROSPECT: a model of leaf optical properties. *Remote Sens Environ*, 1990, 34: 75~91[DOI]
- 10 李小文, 王锦地, 胡宝新, 等. 论先验知识在遥感反演中的作用. *中国科学, D 辑*, 1998, 28(1): 67~72
- 11 Shanno D F. Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computing*, 1970, 24: 647~656
- 12 Chen J M, Leblanc S G. A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Transactions on Geoscience and Remote Sensing*, 1997, 35: 1316~1337[DOI]