

插值算法在高光谱数据中的应用

张 军, 王茂芝, 陈 聆, 张 涛

(数学地质四川省重点实验室, 成都 610059)

摘 要:针对高光谱遥感数据具有波段多,各光谱通道间连续,含数据信息量大,像元光谱曲线交叉并不甚多等特点,就分析光谱曲线特点之后,选择性的运用波段插值方法进行光谱重建,利用 MATLAB 平台对多种插值算法进行了拟合实现。通过对比实验插值结果与原始数据,以及各插值算法插值精度,均方差和峰值信噪比等图像参数相互比较,显示用二次拟合的方法能够得到较好的拟合效果。

关键词:高光谱;插值;峰值信噪比

中图分类号:TB115

文献标识码:A

引 言

电磁波是地物信息的载体,而高光谱遥感则通过多达至数百个光谱通道,用很窄而连续的光谱通道对地物持续成像,更加丰富了地物信息,提高了分辨率精度。

但是由于光谱仪或者外界因素的影响,造成某些波段分辨率低,无法正常使用。本文主要就 MATLAB 平台,对高光谱数据立方体各缺省波段进行多种插值算法拟合,最后再对插值结果通过峰值信噪比,均误差等参量加以评价,得出结论。为了能够更好的选择插值算法,下面是对波谱曲线就形态上加以初步分析。

(1)如图 1 光谱曲线示例所示,横向表示波段数,共 158 个波段,纵向表示像元点在某波段处(放大后)的反射率值,假设“—”线 p 表示当前我们需要插值获取的像元点 P 对应的光谱曲线,“·”线 yyp 表示像元点 P 下方第二个像元点对应光谱曲线,“—”线 yp 表示像元点 P 下方第一个像元点对应的光谱曲线,“—”线 py 为像元点 P 上方第一个像元点对应的光谱曲线,“—”线 pyy 为像元点 P 上方第二个像元点对应的光谱曲线。可见在前约 30 个波段,他们相邻像元点反射率值是非常接近的,理论上存在多种插值算法能够较好的拟合,但是由于过多的像元点(本文中数据约 10 万个),单纯的一种插值算法也很难保证其整体有较高的精度,所以本文采用多种插值算法,比较精度,对每一个

波段选择最优的插值算法组合,得出精度最好的插值结果。

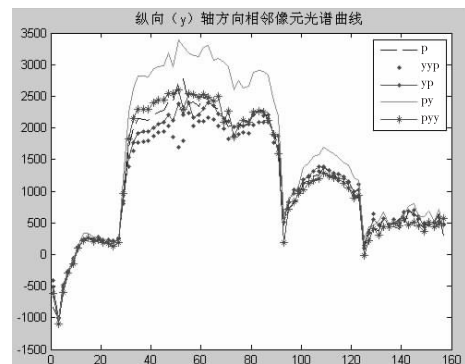


图 1 光谱曲线示例 1

(2)如图 1 所示,在 40 到 80 之间的波段,相邻像元点反射率值具有较大波动,而且并不是我们预想的,距离越近,像元点反射率值越接近。图中“—”线 p 为我们所求波段,y 方向上离它最近的像元点对应光谱曲线为“—”线 yp 和“—”线 py,而图中显示在这连续 40 多波段中,“—”线 yp 为离所求光谱曲线最远者,而“—”线 pyy 和“—”线 yp 更为接近我们所求,所以针对以上波段不适合于用反距离权重类似的插值算法。

(3)如图 2 光谱曲线 2 示例,横坐标为波段数,纵坐标为光谱反射率值(放大后)。针对其连续性和较平缓

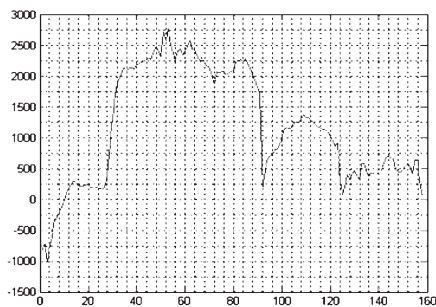


图2 光谱曲线示例2

特点,我们可以通过直线拟合、抛物线拟合和3次拟合来对曲线逼近。考虑到抛物线、3次曲线分别只有一个和2个波峰或波谷,而光谱曲线并没有单一的趋势走向,所以如果将整条曲线都作为插值参量势必会有较大误差。本文中我们选择与所差值波段邻近的前后两个共4个波段进行多次拟合,计算各自峰值信噪比参量,得出最优拟合方法。

1 算法简介

1.1 最近邻法

若图像中两相邻点的距离为1,即行间距为1,列间距为1,取与所计算点 (x,y) 周围相邻的4个点,比较他们与被计算点的距离,取最近的点的光谱值作为 (x,y) 点的光谱值^[1-6]。

1.2 双线性内插值法

以实际位置邻近的4个像元值,根据以下公式确定所需像元的灰度值:

$$g(m,n) = \frac{p_1g_1 + p_2g_2 + p_3g_3 + p_4g_4}{p_1 + p_2 + p_3 + p_4} = \frac{\sum_1^4 p_i g_i}{\sum_1^4 p_i} \quad (1)$$

其中, $g(m,n)$ 为4个点 g_i 之间某点输出像元灰度值, g_i 为邻近点 i 的灰度值, p_i 为邻近点对投影点的权重,即距离的倒数,最近者权重最大。

双线性内插法^[1,4-6]计算量大,但是缩放后图像质量较高,不会出现不连续的情况。另外由于双线性插值具有低通滤波器的性质,使高频分量受损,所以可能会使图像轮廓在一定程度上变得模糊。

与此类似,如果取邻近4个像元值拟合一个三次函数,求的所需位置的像元值就是三次立方插值;如果取邻近的16个点作双线性类似处理就是三次卷积法插值,其中三次卷积法基本能够克服轮廓模糊的问题,但是其计算量会大大增加,所以本文就双线性插值法进行拟合。

1.3 十点插值

针对高光谱数据立方体,假设中间第 i 个波段不能使用,为了使所得的光谱值网格面在各个方向上更加平缓,我们将第 $i-1$ 和 $i+1$ 个波段对应各方向,水平方向、竖直方向的平均值的均值作为第 i 个波段 (m,n) 像元点的值。

首先,我们假设各个波段之间的间距是相等的。水平方向,考虑距离远近的权重关系,离第 i 波段 (m,n) 像元点最近的是第 $i-1, i+1$ 波段的 (m,n) 像元点,分别取它们的权重系数为 $1/4$,对于离第 i 波段 (m,n) 像元点的8个次近点,分别取它们的权重系数为 $1/16$,即所求像元值:

$$p(m,n)_i = \frac{1}{4} * [p(m,n)_{i-1} + p(m,n)_{i+1}] + \frac{1}{16} * [p(m-1,n)_{i-1} + p(m+1,n)_{i-1} + p(m,n-1)_{i-1} + p(m,n+1)_{i-1} + p(m-1,n)_{i+1} + p(m+1,n)_{i+1} + p(m,n+1)_{i+1} + p(m,n-1)_{i+1}] \quad (2)$$

其中 p 分别为对应像元值;若所需插值波段为边界波段或者边界行列,某些方向不存在对应的像元值,如插值波段为第一波段,则不存在前一波段像元值。在这种情况下我们只取存在像元值的对应方向的加权平均,权重系数设定同上。

1.4 拟合插值

针对高光谱光谱曲线特点,如图2所示,其具有连续性,以及除个别地方外,具有局部平缓性,所以对于缺省波段处,我们可以通过邻波段像素值对未知点进行一次函数,二次函数以及更高次函数对其进行局部拟合,通过最小二乘法择出最优拟合函数。

本文考虑到插值波段的随机性,就所需插值波段前、后各2个共4个波段对应像素值分别用一次、二次、三次对若干波段进行了拟合,并取得了较好结果,因为4个点总会模拟出一个简单的趋势。

如图3插值模型分析所示,针对左侧两点呈递增趋势的几种分布形式为例(反之类似),分析几种默认假设。

(1)如果4个点呈模型一和模型二所示,中间圆圈点为我们插值的点,实点相对对称的分布在插值点的两侧。此两种情形都是左侧两点递增,右侧两点递减且第三点小于等于第二点的值。由图我们可以知道前者用二次抛物线可以较好拟合,而后者为3次拟合,它与二次拟合曲线也非常接近,故以上两种情况我们认为用二次或三次拟合均可。

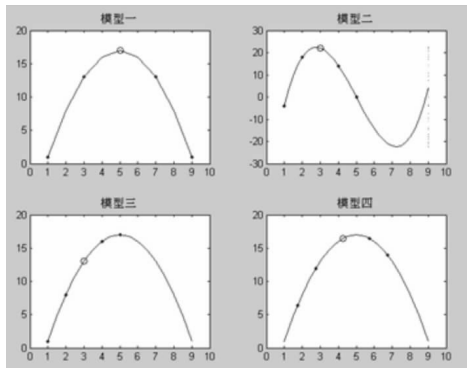


图 3 插值模型分析

(2)如图 3 插值模型分析中模型三所示,如果 4 个已知点呈单调递增形式,此时比较适合于二次抛物线拟合和线性拟合。

(3)相比模型三,模型四不同的是前三点依然呈递增趋势,而第 4 点呈下降趋势,此时我们用二次拟合效果比较好。

按照以上 3 种情况,我们把左侧递增的情况大致分为:四点递增(模型三),第四点开始递减(模型四),第三、四点递减(模型一、二);另外还有一种没列出来的情况就是,左侧两点递增,第三点值小于第二点值,第四点值大于第三点值,由于反射率值分布的随机性,很难判断用什么曲线拟合才是最好的,所以对此一种情况不妨也采用二次拟合,来减小整体多种情形下的误差,开口向上的情况类似处理。

1.5 拟合插值改进

同图 3 插值模型分析,设 p_1, p_2, p_3, p_4, p_5 依次为图中对应已知像元点反射率值, p_3 对应空心点。按图所示,模型一和模型二显然是用二次拟合比其他拟合的精度要高,而模型三和模型四用二次拟合、三次拟合或线性拟合都能有较好的结果,所以我们把最适合用二次拟合的情形用二次拟合,而其他情况用线性、均值或者三次拟合。

本文定义对称值均方差来区分以上两种情况。对称值均方差就是:将第三、四点的值关于第二点值做垂直对称,我们得到一个新的数组,此数组的方差称为对称均值方差。如图 3 插值模型分析所示,在模型一、二中,数组 $[p_1 p_2 p_4 p_5]$ 的方差明显小于对称翻转过后数组 $[p_1 p_2 (2 * p_2 - p_4) (2 * p_2 - p_5)]$ 的方差,甚至几倍关系。于是本算法中,将原方差小于若干倍对称值均方差的数据用二次拟合插值,而其他的情况下则选择线性、三次、均值等进行插值拟合。

2 算法应用

在本文中,算法选择是关键。为了查看以上插值算法重建高光谱数据立方体的精度,本文先后用最近邻法、双线性插值等对某地区高光谱遥感数据缺省波段进行处理。处理结果如表 1 波段插值误差表所示:

表中列举方差 1 至 3 倍,是因为对呈线性分布的 4 点按照假设(2)处理后,前者方差是后者的 1/3 左右。

如图 2 光谱曲线示例所示,本文选取了该遥感影像第 35,85,57,129 波段,分别分布在光谱曲线上较平缓段和震动较剧烈处作为样本,查看其各自误差精度,见表 1 波段误差表所示。

表 1 波段插值误差表

波段数	第 35 波段	第 85 波段	第 57 波段	第 129 波段
波段均值	1841	2050.6	2076.1	749.80
1 次拟合平均误差	43.73	29.83	38.90	84.42
2 次拟合平均误差	29.80	28.15	41.92	78.83
双线性平均误差	155.18	156.16	165.06	133.79
十点插值平均误差	77.07	72.88	78.29	89.67
改进的 1 倍方差	30.58	27.79	33.49	77.65
改进的 2 倍方差	32.00	28.06	33.77	78.99
改进的 3 倍方差	33.20	27.88	34.21	79.04

结果表明,二次拟合与方差 1 倍混合拟合效果最佳,而混合拟合可以看作在 2 次拟合的基础上改进而得到,所以下面主要就拟合改进方法对遥感数据进行整体处理,结果见图 4 误差曲线图与峰值信噪比图所示,横坐标为波段数,左图中最大值曲线表示对应波段最大值,最小值曲线表示对应波段最小值,平均值曲线表示对应波段均值,此处误差曲线以及右图中峰值信噪比曲线都有两条曲线,因为他们相差很小,所以看起来重叠在一起;其中误差曲线表示拟合方法与改进插值后波段的平均误差曲线,右图中峰值信噪比曲线也表示拟合方法和改进插值后对应波段的峰值信噪比曲线。可见,除了个别波段外,此两种方法插值精度均较小,对应峰值信噪比值也非常接近。

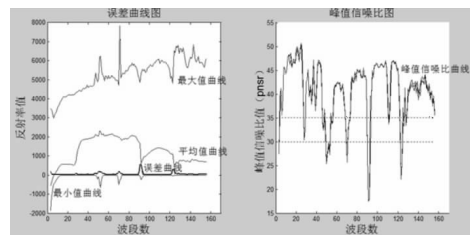


图 4 误差曲线图与峰值信噪比图

3 结 论

根据上图峰值信噪比^[7](*pn*sr)和平均误差(*u*)结果显示,对于绝大多数波段,混合拟合平均误差一致接近0,峰值信噪比(*pn*sr)值也只有少数的几个波段处于30以下,其大多数都处于35以上,由图5峰值信噪比与效果图可知,*pn*sr值与拟合图像效果呈正相关关系,而且上图中相当一部分波段对应*pn*sr值已经超过45,这与前面拟合误差较小的结论一致,可见此两种算法可以考虑用于波段插值拟合,而且只需用通过简单的数据处理就能得到我们需要的近似波段影像。

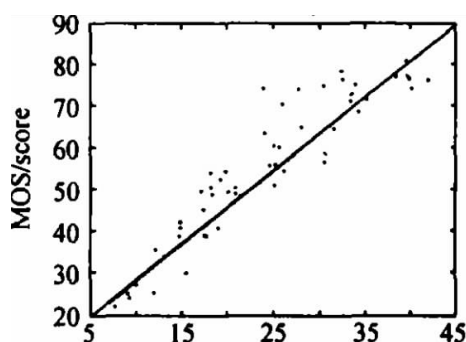


图5 峰值信噪比值 - 效果图

参 考 文 献:

- [1] 符祥,郭宝龙.图像插值技术综述[J].计算机工程与设计 2009,30(1):141-144.
- [2] 舒宁,苏俊英.高光谱影像光谱响应曲线分维计算[J].遥感应用,2009(1):23-26.
- [3] 宋一中,董祥军,赵志敏.光谱分析采样数据重建原始信号[J].光谱学与光谱分析,2008,28(7):1492-1495.
- [4] 任鹏远,廖宁放,柴冰华,等.基于多光谱成像的光谱反射率重建[J].光学技术,2005,31(3):427-433.
- [5] 杨萍,廖宁放,等.一种离散点插值算法用于重建光谱反射率[J].光学技术,2008,34(6):870-876.
- [6] 叶俊,苏跃斌.有限域上插值多项式的两种构造方法[J].四川理工学院学报:自然科学版,2010,23(5):521-523.
- [7] 黄小乔,石俊生,杨建,等.基于色差的均方差与峰值信噪比评价彩色图像质量研究[J].光子学报,2007,36(增刊):295-298.

Interpolation Algorithm in the Application of Hyperspectral Data

ZHANG Jun, WANG Mao-zhi, CHEN Ling, ZHANG Tao

(Geomathematics Key Laboratory of Sichuan Province, Chengdu 610059, China)

Abstract: The hyperspectral data have many bands, the spectrum lines between channel are continuous, much more data information is contained, and spectrum curve is not very much cross and so on. To the characteristics mentioned above, spectrum is reconstructed by selective band interpolation on Matlab after the spectrum curve characteristics analysed. At last, compare the results of each other, mean variance of each other and the peak value signal-to-noise ratio, come to the conclusion of the method is simple, and the two curve fitting with the better reconstruction precision.

Key words: hyperspectral; interpolation; peak signal-to-noise ratio