

基于离散小波变换的高光谱特征 提取中分解尺度的确定方法^{*}

李 军 李培军^{**} 郭建聪

北京大学地球与空间学院 遥感与地理信息系统研究所, 北京 100871

摘要 近年来, 小波变换被成功地运用于高光谱数据的特征提取. 其中的关键环节之一是如何确定合适的分解尺度, 以获得较高的分类精度. 文中提出一种改进的基于相关系数的小波分解最佳尺度的确定方法. 该方法首先对高光谱影像进行多尺度小波分解, 提取特征图像, 然后, 计算各个尺度的小波近似系数与原始光谱信号之间的相关系数, 根据相关系数的变化趋势确定最佳分解尺度. 通过两个高光谱数据验证所提出的方法, 并与现有方法进行对比. 实验结果表明, 所提出的方法可较准确地确定最佳的分解尺度.

关键词 特征提取 小波变换 分解尺度 高光谱影像

特征提取 (feature extraction) 是高光谱遥感信息处理中的重要内容之一. 一般来说, 特征提取是指从多波段遥感影像中提取有限的特征波段, 而这些波段的分类精度与原始影像的分类精度相比没有显著降低^[1,2]. 目前, 已有不少有效的特征提取方法, 如主成分分析方法 (PCA)、最小噪声分量 (MNF)、决策边界特征提取 (DBFE) 等. 近年来, 小波变换被成功地运用到高光谱遥感数据的特征提取中. Bruce 等^[3] 提出基于连续和离散小波变换的特征提取方法; Pu 和 Gong^[4] 将其中的离散小波特征提取方法用于森林叶面积指数估算. Kaewpijit 等^[5] 提出了另一种基于离散小波变换的特征提取方法, 并分别用多种分类方法对新得到的特征影像进行分类和比较^[5,6]. 刘正军等^[7] 将离散小波变换和遗传算法结合, 进行高光谱数据的特征提取.

在基于小波变换的高光谱数据特征提取中, 小波分解尺度的确定是一个关键问题. 分解尺度过大或过小, 都不能得到合适的特征图像, 也无法获得高的分类精度, 因此, 最合适(或最佳)分解尺度的

选择, 直接影响特征提取的结果. 这里的最佳分解尺度, 是指获得最高分类精度的分解尺度. 目前, 主要有两种最佳分解尺度的确定方法: (1) 根据母小波函数及光谱信号的长度确定最佳分解尺度^[3,4]; (2) 计算重构信号与原始信号之间的相关系数, 设定阈值来确定最佳分解尺度^[5,8]. 尽管上述两种方法在基于小波变换的特征提取中得到了广泛的应用, 并与一些经典的特征提取方法(如主成分分析)进行了对比, 但还没有相关的研究对它们进行相互比较. 本文通过实际的高光谱数据对两种方法进行比较, 提出了一种改进方法, 并与上述两种方法进行了对比.

1 方法

1.1 基于小波变换的特征提取方法

小波变换是一种较新的信号分析工具, 它提供了一种分析多尺度或多分辨率信号的方法, 已被应用于许多的领域. 小波变换在遥感领域中的应用相当广泛, 如数据压缩、图像纹理分析、图像边缘检

2007-01-23 收稿, 2007-04-25 收修改稿

^{*} 北京大学工程研究院项目资助

^{**} 通信作者, pili@pku.edu.cn

©1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

测、图像降噪以及图像融合等^[9,10]。利用连续小波变换(CWT),可以在连续尺度上对信号进行分析,包括多维信号,比如高光谱影像。离散小波变换(DWT)则在一系列离散的尺度上分析信号。通常,离散尺度都是二阶的(2, 4, 6, 8, 16, ...),且小波变换可以利用多种快速运算方法来实现^[11-13]。

近年来,小波变换被运用于高光谱数据的特征提取^[3-5],其中两种方法得到广泛应用^[3-8]。第一种方法是,对高光谱影像进行多尺度的离散小波变换,在每一尺度的小波分解后,特征数都将减半。取最佳分解尺度的小波近似(approximate)系数作为最终的特征影像。由于小波近似系数和原始信号具有较高的相似性,因此该方法得到的特征影像可以获得较好的分类结果^[5]。第二种方法是对高光谱影像进行多尺度的离散小波变换后,同时取最佳分解尺度的近似系数和所有细节(detail)系数,分别计算它们的能量值,作为最终的特征影像。该方法为信息量无损变换,所以,该方法也获得了较好的分类结果^[3,4]。当然,在上述两种方法中所采用的小波分解方法不同。

本文采用第二种方法进行特征提取。首先对影像中每个像元的光谱曲线进行多尺度的一维离散小波变换,在每一尺度,可得到与原始信号相关性较高的近似系数和包含光谱曲线细节信息的高频系数(细节系数),然后对近似系数再进行下一尺度的小波分解,依此重复,这样将得到所有分解尺度的细节系数以及最终分解尺度的近似系数。假设对影像进行 n 尺度小波分解,可得到 n 个细节系数和 1 个近似系数,分别计算细节系数以及近似系数的能量值,则最终可得到 $n+1$ 个特征波段^[3,4,14]。这样得到的 $1 \times (n+1)$ 维特征向量 $F = \{F_i\}$ 由下式得到:

$$F_i = \sqrt{\frac{1}{K} \sum_{k=1}^K W_{ik}^2}$$

其中 K 表示分解尺度为 i 时的小波系数个数, W_{ik} 表示分解尺度为 i 时的第 k 个小波系数, n 为小波分解的最大尺度。向量 F 的大小为 $(n+1)$, 其中包括 n 个小波细节系数和最终分解尺度的小波近似系数。 $(n+1)$ 个特征向量代表了每个像元的新的光谱信息。

1.2 分解尺度的确定

运用小波变换进行特征提取的关键环节之一是如何确定最佳分解尺度,以获得最高的分类精度。目前主要有两种分解尺度的确定方法。一种是根据母小波函数和光谱信号的长度共同确定^[3,4]。一方面,小波分解的最佳尺度 n 可由方程: $n = 1b(N)$ 确定,其中 N 为高光谱信号的长度。例如,根据上式计算,具有 220 个波段的 AVIRIS 数据的最佳分解尺度 n 应为 8。另一方面, n 也可以根据母小波的滤波器长度来确定。如利用 Haar 小波进行多尺度分解时, n 值为 8;若母小波为 db4,则 n 值为 4。由于由这两个条件得到的最佳分解尺度可能不同,通常取两者之间的最大值作为最终的分解尺度^[3,4]。尽管该方法有确切的理论依据,但并没有考虑具体数据的特点,而且由此方法确定的最佳分解尺度只是一个较为概略的估计,可能并不准确。例如,当高光谱数据的波段数在 128 (2^7) 与 256 (2^8) 之间时,根据以上的确定方法,计算得到的最佳分解尺度可能都为 8。

另一种最佳分解尺度的确定方法,是利用相关系数作为判断条件^[5]。这里的相关系数为小波近似系数的重构信号与原始信号的相关性,即重构的高光谱信号与原始信号之间的相关性。假设原始光谱信号为 x , 近似系数的重构信号为 y , 则二者间的相关系数可由下式计算^[5]:

$$\rho(x, y) = \frac{\sum x_i y_i - \left(\frac{1}{N}\right) \sum x_i \sum y_i}{\sqrt{\left[\sum x_i^2 - \left(\frac{1}{N}\right) \left(\sum x_i\right)^2\right] \left[\sum y_i^2 - \left(\frac{1}{N}\right) \left(\sum y_i\right)^2\right]}} \quad (1)$$

其中 N 为原始信号的维数,即原始高光谱影像的波段数, i 表示高光谱影像的第 i 波段。

利用这一相关系数确定小波分解尺度的具体方法为^[5]:首先,对每一分解尺度,利用公式(1)计算每个像元重构信号与其对应的原始信号间的相关系数,这样,每个像元就得到与分解尺度数相同的多个相关系数;然后,设定一个相关系数阈值,可得到每个像元满足该阈值条件的相关系数所处的最大分解尺度,对不同的像元来说,这个最大尺度

可能不同;最后,对影像中所有的最大分解尺度进行统计,计算这些不同的最大尺度的百分比,取其中占多数的最低尺度作为最佳分解尺度.例如,在 Kaewpjit 等^[5]的研究中,当相关系数阈值为 0.99 时,满足阈值条件的像元中有 98.86% 的分解尺度为 1,而只有 1.13% 的像元分解尺度为 2,因此,当阈值条件为 0.99 时,最佳尺度为 1;而当阈值为 0.97 时,满足阈值条件的像元中 77.67% 的分解尺度为 2,22.28% 的分解尺度为 3,所以,当阈值为 0.97 时,可以确定最佳分解尺度为 2.因此可以看出,这种确定最佳分解尺度的方法存在以下不足:

(1) 相关系数阈值的确定没有统一的规则,即阈值的确定具有较强的主观性,这样难以确定一个最佳阈值以得到最佳的分解尺度.

(2) 过程繁琐,计算量大.因为该方法不仅需要计算整个图像的各个尺度的重构影像,而且还要计算每个像元不同尺度的重构信号与原始信号之间的相关系数,同时还要通过大量实验来确定相关系数阈值.

从上述的分析可以认为,一个合适的最佳尺度选择方法,需要针对不同数据特点,而且具有相对可靠的制约条件,计算相对简便.基于这些考虑,本文提出一种新的最佳分解尺度判断方法.该方法仍采用 Kaewpjit 等^[5]的方法的相关系数作为最佳尺度的判断准则,但与 Kaewpjit 等^[5]的方法不同的是,该方法不需要计算影像中所有像元重构值与原始值之间的相关系数,而只计算每个地物类别训练样本中所有像元在各个尺度的重构值与原始值之间的相关系数,并取各个地物类别各尺度的相关系数的平均值,作为判断最佳尺度的条件.同时,最佳分解尺度的判断方法与 Kaewpjit 等^[5]的方法也不同,不是采用某一个阈值作为判断条件,而是根据相关系数的稳定性来确定最佳分解尺度,即根据相关系数随尺度变化而变化的特征来确定.具体说来,对每个地物类别来说,随着分解尺度的增大,当相邻两个尺度之间的相关系数的差值很小,并趋于一个常数时,则认为该相关系数达到稳定.以各类地物相关系数趋于稳定时的尺度作为该地物的最佳分解尺度.由于不同地物类别的最佳分解尺度可能不同,因此,把多数地物类别相关系数趋于稳定的尺度作为整个影像的最佳分解尺度.

1.3 评价方法

本文根据影像分类的精度来评价该方法的有效性.首先,分别运用现有的尺度确定方法和本文提出的方法确定最佳分解尺度,然后对获得的特征图像进行分类,根据总体分类精度来判断不同尺度确定方法的准确性.本文采用最大似然法对不同分解尺度得到的特征影像进行图像分类.最大似然法是一种经典有效的分类方法,已被广泛地运用在遥感影像的分类处理当中^[5,6,8].在分类过程中,各类采用等概率设置.本文采用 Daubechies 小波系列的 db4 作为母函数,对光谱信号进行小波分解,这一系列的小波在遥感图像处理中被广泛采用^[3-6,8].

2 实验数据

本文利用两个经典的高光谱数据,即 AVIRIS 数据和 HYDICE 数据,来评价比较本文提出的小波分解最佳尺度确定方法和两种现有方法.这两个数据被广泛用于高光谱处理方法的评价比较^[8,15-18].图像分类所采用的训练样本和检验样本也主要来自以往的研究数据^[15-18],同时根据实际情况增加了一些样本.

2.1 AVIRIS 数据

所采用的 AVIRIS 数据为 1992 年获取的,影像覆盖区域为美国加利福尼亚州的一处农业用地区^[15],影像大小为 145 * 145 像元,空间分辨率 20 m,包含 220 个波段,光谱范围从 0.4—2.04 μm,光谱分辨率

表 1 地物类别及训练样本和检验样本

类名	训练样本	检验样本
玉米 1	250	277
玉米 2	378	395
大豆 1	123	214
草/树	213	251
大豆 2	554	624
林地	451	481
草/牧草	119	155
大豆 3	266	258
干草	153	201

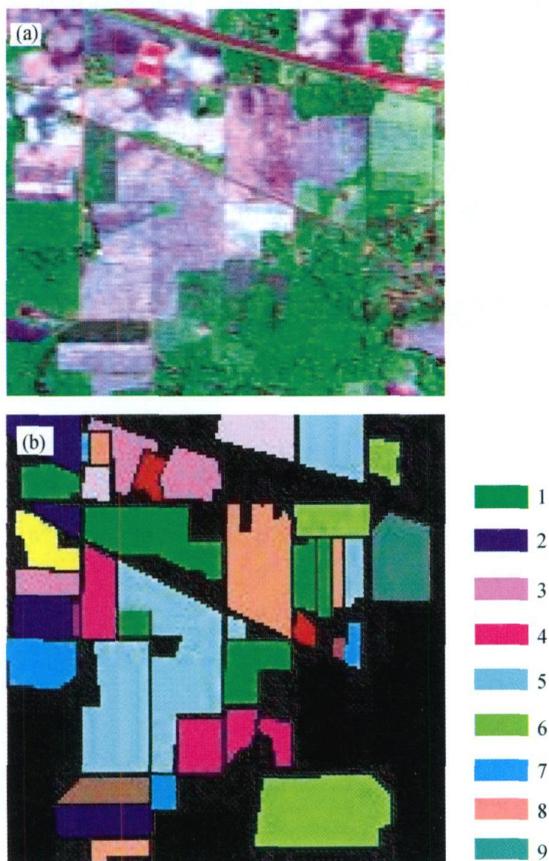


图 1 AVIRIS 高光谱影像及其地面参考数据图

(a) AVIRIS 假彩色合成影像, 第 10, 80, 180 波段分别为 R, G, B; (b) 地面真实样本数据
地物类别: 1 玉米 1; 2 林地; 3 大豆 1; 4 草/树; 5 大豆 2; 6 林地; 7 草/牧草; 8 大豆 3; 9 干草

10 nm. 实验区包括 16 个地物类别. 图 1 为 AVIRIS 高光谱影像及其地面参考数据. 表 1 为本实验所采用的训练样本和检验样本.

2.2 HYDICE 数据

本实验所采用的 HYDICE 高光谱数据, 大小为 307×1280 像素^[15], 包含 210 个波段, 光谱范围 $0.4-2.40 \mu\text{m}$. 去掉位于 $1.4 \mu\text{m}$ 及 $1.9 \mu\text{m}$ 附近由于水分强吸收而产生的噪声波段, 最后剩下 191 波段. 图 2 为 HYDICE 影像及各个地物类别的地面真实样本. 表 2 为本实验数据的训练样本和检验样本.

表 2 地物类别及训练样本和检验样本

类名	训练样本	检验样本
屋顶	1263	2966
窄路	183	304
宽路	651	1072
阴影	223	372
树	136	888
草地	970	3892
水体	1172	3020

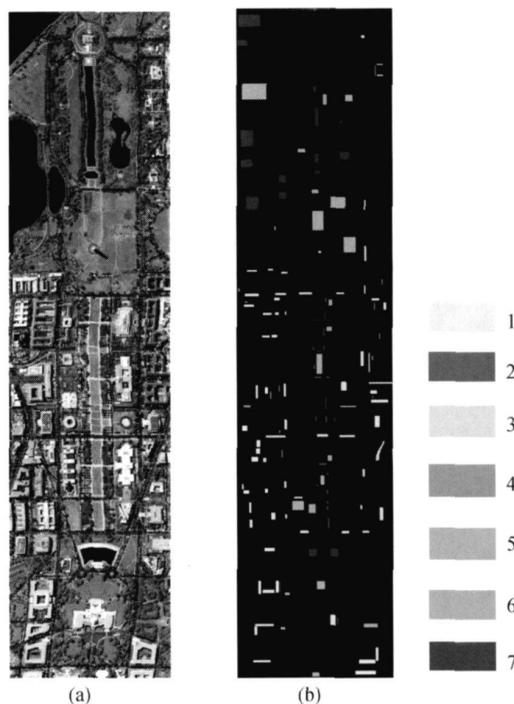


图 2 HYDICE 影像及各地物的地面真实样本图

(a) HYDICE 假彩色合成影像, 第 60, 27, 17 波段分别为 R, G, B; (b) 地面真实样本数据
地物类别: 1 屋顶; 2 窄路; 3 宽路; 4 阴影; 5 大树; 6 草地; 7 水体

3 实验结果及讨论

本研究采用 ENVI 4.3 和 MATLAB 7.0 作为实验软件环境. 其中, ENVI 4.3 用于影像分类、精度评价和影像显示, MATLAB 7.0 用于影像的小波分解、特征提取及相关系数计算.

3.1 AVIRIS 数据

由公式 $n = 1b(N)$ 可以得到, 220 个波段的 AVIRIS 影像的最佳分解尺度应为 8; 同时, 根据母小波函数 (db4) 可以得到最佳分解尺度为 4. 取两

者的最大值 8. 因此, 根据母小波函数与信号长度所确定的最佳分解尺度为 8.

根据 Kaewpigit 等^[5]的方法, 计算每个像元在每个分解尺度的重构信号与原始信号的相关系数. 表 3 为取不同的相关系数阈值时, AVIRIS 影像中满足相关系数阈值条件的像元所处的最大分解尺度的百分比. 从表 3 可以看出, 当阈值较大时, 尺度

分布比较集中, 而阈值较小时, 尺度的分布则比较分散. 当阈值小于 0.90 时, 满足相关系数阈值条件的最低分解尺度均为第 5 尺度. 因此, 根据 Kaewpigit 等^[5]的方法, 最佳尺度应为 5. 同时, 值得注意的是, 阈值为 0.80—0.83 之间, 满足阈值条件的最低尺度中, 第 10 尺度占有很大比例 (> 50%).

表 3 相关系数阈值从 0.99—0.80 时 AVIRIS 影像各像元满足阈值条件的相关系数所处的分解尺度(1—10)所占的百分比

	1	2	3	4	5	6	7	8	9	10
0.99	0	99.62	0.38	0	0	0	0	0	0	0
0.98	0	0.43	99.36	0.21	0	0	0	0	0	0
0.97	0	0	97.69	2.24	0.06	0	0	0	0	0
0.96	0	0	44.23	54.92	0.33	0.52	0	0	0	0
0.95	0	0	0	97.66	0.91	1.43	0	0	0	0
0.94	0	0	0	71.36	26.34	2.3	0	0	0	0
0.93	0	0	0	47.97	35.6	16.43	0	0	0	0
0.92	0	0	0	31.76	32.22	36.01	0	0	0	0
0.91	0	0	0	17.52	39.58	41.3	1.6	0	0	0
0.90	0	0	0	0	53.52	8.55	37.93	0	0	0
0.89	0	0	0	0	48.92	4.9	46.17	0	0	0
0.88	0	0	0	0	44.61	3.48	51.81	0	0	0.1
0.87	0	0	0	0	40.01	3.45	51.15	0	0	5.39
0.86	0	0	0	0	35.4	3.06	37.43	0	0	24.11
0.85	0	0	0	0	31.78	2.01	25.12	0	0	41.08
0.84	0	0	0	0	28.34	1.68	22.96	0	0	47.02
0.83	0	0	0	0	25.13	1.56	21.46	0	0	51.85
0.82	0	0	0	0	22.13	1.43	20.51	0	0	55.92
0.81	0	0	0	0	18.4	1.92	19.42	0	0	60.26
0.80	0	0	0	0	13.66	2.15	19.34	0	0	64.83

表 4 为 AVIRIS 影像中各类地物类别训练样本的不同分解尺度的重构信号与原始信号的相关系数. 可以看出, 从第 1 尺度到第 9 尺度, 所有地物各尺度之间的相关系数变化都较大, 但第 9 尺度和第 10 尺度之间变化较小, 到第 10 尺度以后, 所有地物类别的相关系数变化都很小, 趋于稳定. 例如林地类的第 4 尺度和第 5 尺度的相关系数差值为 0.17, 第 10 尺度和第 11 尺度之间的相关系数差为 0.0013; 草/牧草类第 5 尺度和第 6 尺度的相关系数差值为 0.13, 而第 10 尺度和第 11 尺度之间的相关系数差为 0.0039. 同时可以看出, 到第 10 尺度后, 几乎所有地物的每两个尺度之间的相关系数的差值

都小于 0.005, 由此, 相关系数可以视为稳定. 为了更清楚地表示相关系数随着尺度的变化情况, 将相邻两个尺度的相关系数相减, 来观察差值的变化. 图 3 为各类地物相邻两尺度相关系数之间的差值(绝对值)的变化曲线, 其中, 横轴尺度中的 1 表示第 1 尺度和第 2 尺度之间的相关系数差值, 2 表示第 2 尺度和第 3 尺度之间的差值, 依此类推. 从此图中可以清楚地看出, 当分解尺度为 10 以后, 各类地物的相关系数均变化很小, 趋于稳定. 因此, 根据各类地物相关系数稳定性条件, 可以确定最佳分解尺度为 10.

表 4 AVIRIS 数据各地物不同尺度的重构信号与原始信号之间的相关系数表

分解尺度	玉米 1	玉米 2	大豆 1	草/树	大豆 2	林地	草/牧草	大豆 3	干草
1	0.998	0.998	0.9979	0.9974	0.998	0.99741	0.9974	0.9981	0.99773
2	0.9954	0.9952	0.995	0.9933	0.9952	0.9936	0.9933	0.9955	0.99444
3	0.9854	0.9842	0.9837	0.9815	0.9841	0.98149	0.9815	0.985	0.98081
4	0.9682	0.9644	0.9637	0.9561	0.9641	0.95634	0.9561	0.9663	0.95628
5	0.9475	0.9405	0.9414	0.908	0.9387	0.78248	0.908	0.9437	0.92923
6	0.936	0.9273	0.9228	0.7804	0.9251	0.77792	0.7804	0.9334	0.89109
7	0.9057	0.9053	0.9032	0.7759	0.9053	0.71745	0.7759	0.9073	0.88614
8	0.8606	0.8621	0.8556	0.714	0.8642	0.705	0.714	0.8658	0.83793
9	0.9089	0.9022	0.8923	0.7011	0.902	0.71907	0.7011	0.9111	0.84976
10	0.9057	0.9018	0.8934	0.7154	0.9019	0.72507	0.7154	0.9084	0.85771
11	0.901	0.8988	0.8913	0.7215	0.8989	0.72637	0.7215	0.9041	0.85975
12	0.8966	0.8953	0.8882	0.7228	0.8956	0.7272	0.7228	0.8999	0.85889
13	0.8939	0.8931	0.8864	0.7237	0.8935	0.72733	0.7237	0.8974	0.85839

黑色粗体标示相关系数的变化趋于稳定的尺度

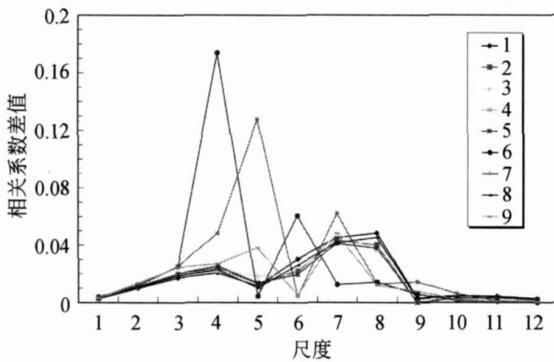


图 3 AVIRIS 数据各地物相邻两个尺度的相关系数间的差值(绝对值)图
地物类别: 1 玉米 1; 2 玉米 2; 3 大豆 1; 4 草/树; 5 大豆 2; 6 林地; 7 草/牧草; 8 大豆 3; 9 干草

运用各分解尺度所得到的特征图像进行分类实验, 结果表明, AVIRIS 影像的分类精度, 从第 1 尺度开始, 分类精度一直呈上升趋势, 到第 10 尺度, 精度达到最高, 为 83.77%, 然后随着尺度的增大, 精度略有下降. 因此, 第 10 尺度应为最佳分解尺度. 三种不同的最佳分解尺度确定方法得到的分解尺度的分类精度分别为: 由本文方法所确定的最佳分解尺度为第 10 尺度, 也正是分类精度最高的尺度; 根据信号长度和母小波函数共同确定的最佳尺度为第 8 尺度, 相应的分类精度为 71.18%, 而由 Kaewpigit 等^[5]

的方法确定的最佳尺度为第 5 尺度, 相应的分类精度仅为 64.60%. 所以, 本文所提出的分解尺度判断方法取得了最好的分类结果.

3.2 HYDICE 数据

根据公式 $n = 1b(N)$ 可以得到, 191 个波段的 HYDICE 影像的最佳分解尺度应该为 8, 而根据母小波函数 (db4) 可以得到最佳分解尺度为 4, 取两者的最大值为 8. 因此, 根据母小波函数与信号长度所确定的最佳分解尺度应为 8.

由 Kaewpigit 等^[5]的方法, 计算每个尺度每个像元的重构信号与原始信号的相关系数. 表 5 为取不同的相关系数阈值时, HYDICE 影像中满足相关系数阈值条件的像元所处的最大分解尺度的百分比. 可以看出, 在所有阈值条件下, 最大分解尺度的分布都比较分散: 当阈值为 0.98—0.99 时, 最佳尺度为 2; 当阈值为 0.92—0.97 时, 最佳尺度为 3; 当阈值为 0.91—0.86 时, 最佳尺度为 4; 当阈值为 0.80—0.85 时, 最佳尺度为 5. 因此, 该方法有多种选择, 阈值不同选择不同, 确定的最佳尺度也不同. 在此, 本文选择较小阈值, 最佳分解尺度为 5. 同时可以注意到, 阈值为 0.80—0.90 之间, 满足阈值条件的最低尺度中, 第 10 尺度占有很大比例(表 5).

表 5 相关系数阈值从 0.99—0.80 时 HYDICE 影像各像元满足阈值条件的
相关系数所处的分解尺度 (1—10) 所占的百分比

	1	2	3	4	5	6	7	8	9	10
0.99	0.02	49.47	30	5.64	12.14	2.73	0	0	0	0
0.98	0	34.05	25.9	8.25	9.74	22.06	0	0	0	0
0.97	0	0.03	49.87	6.05	15.51	28.54	0	0	0	0
0.96	0	0.01	45.03	4.97	16.56	33.43	0	0	0	0
0.95	0	0	40.06	6.67	15.47	37.79	0	0	0	0
0.94	0	0	32.32	11.53	14.14	41.97	0.04	0	0	0
0.93	0	0	23.42	17.26	13.59	43.48	2.21	0	0	0.03
0.92	0	0	14.39	21.9	14.78	36.67	10.14	0.01	0	2.12
0.91	0	0	4.59	26.53	16.91	25.87	14.47	0.11	0	11.52
0.90	0	0	0.01	25.76	19.52	17.44	16.33	0.44	0	20.49
0.89	0	0	0	20.85	21.86	14.26	15.46	1.32	0	26.25
0.88	0	0	0	16.13	24.05	13.09	13.03	2.51	0	31.19
0.87	0	0	0	11.3	26.09	12.86	9.85	4.48	0	35.42
0.86	0	0	0	6.26	28.03	13.01	7.48	6.04	0	39.17
0.85	0	0	0	1.73	29.34	12.03	6.16	7.1	0	43.65
0.84	0	0	0	0.08	27.76	7.85	5.7	7.61	0	51
0.83	0	0	0	0.01	22.02	9.27	5.82	7.82	0	53.2
0.82	0	0	0	0.01	24.77	8.59	5.59	7.82	0	55.06
0.81	0	0	0	0.01	19.44	9.66	6.06	8.2	0	56.64
0.80	0	0	0	0	16.95	9.97	6.2	8.8	0	58.07

表 6 为 HYDICE 影像各类地物类别训练样本的不同尺度的重构信号与原始信号的相关系数。可以看出, 从第 1 尺度到第 12 尺度之间, 相关系数曲线的变化较大, 尤其是从第 4 尺度到第 11 尺度的相关系数曲线, 变化更明显; 但第 12 尺度和第 13 尺度之间变化较小, 到第 13 尺度以后, 所有地物类别的相关系数都趋于稳定。如水体类, 第 6 尺度和第 7 尺度之间的相关系数差为 0.14, 而到第 12 尺度后, 相邻两尺度间的相关系数差都小于 0.003; 树

类第 5 尺度和第 6 尺度之间的相关系数差值为 0.1, 第 12 尺度和第 13 尺度之间的相关系数差值为 0.0016; 屋顶类第 5 尺度和第 6 尺度之间的相关系数差值接近 0.8, 从第 11 尺度开始, 相邻两尺度的相关系数都小于 0.005。图 4 为各类地物相邻两尺度相关系数之间的差值 (绝对值) 的变化曲线。从图中可清楚地看出, 当分解尺度大于 13 时, 相关系数变化很小, 趋于稳定。因此, 根据本文提出的分解尺度确定方法, 最佳分解尺度应为 13。

表 6 HYDICE 数据各地物不同尺度的重构信号与原始信号之间的相关系数表

分解尺度	屋顶	窄路	宽路	阴影	树	草地	水体
1	0.9997	0.9993	0.9996	0.9998	0.9989	0.99889	0.99982
2	0.999	0.9968	0.9986	0.9992	0.9934	0.99395	0.99932
3	0.9976	0.9911	0.9969	0.998	0.9738	0.98052	0.99863
4	0.9921	0.9769	0.9927	0.9948	0.9137	0.94703	0.9969
5	0.9635	0.971	0.9914	0.9938	0.8629	0.92549	0.99529
6	0.9591	0.8941	0.9902	0.9801	0.7621	0.88455	0.98814
7	0.8446	0.8825	0.9082	0.8681	0.7112	0.86675	0.84837
8	0.8359	0.8412	0.9126	0.89	0.6835	0.83645	0.87305
9	0.8985	0.7868	0.9628	0.9424	0.6269	0.75889	0.92888
10	0.89	0.8091	0.9492	0.9139	0.6479	0.78578	0.89952
11	0.8809	0.8209	0.9376	0.8944	0.6606	0.80121	0.87906
12	0.8752	0.8257	0.9305	0.8832	0.6655	0.80762	0.86758
13	0.8713	0.8284	0.9258	0.876	0.6687	0.81368	0.85709
14	0.8689	0.8299	0.9228	0.8716	0.6703	0.81508	0.85556
15	0.8673	0.8308	0.9208	0.8687	0.6715	0.81588	0.85254
16	0.8662	0.8313	0.9196	0.8669	0.6721	0.81593	0.85064

黑色粗体标示相关系数的变化趋于稳定的尺度

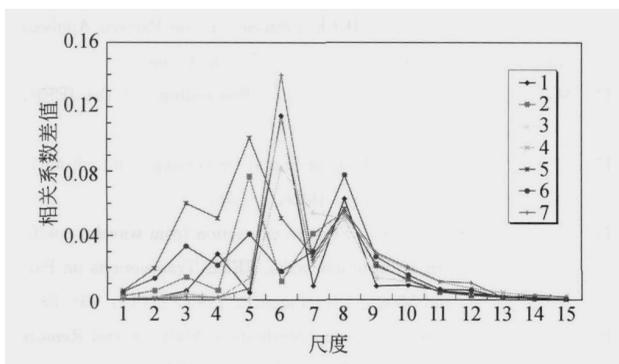


图 4 HYDICE 数据各地物相邻两个尺度的相关系数间的差值(绝对值)图
地物类别: 1 屋顶; 2 窄路; 3 宽路;
4 阴影; 5 大树; 6 草地; 7 水体

运用各分解尺度所得到的特征图像进行分类实验, 分类精度随分解尺度变化的趋势为, 从第 1 尺度开始, 分类精度呈上升趋势, 到第 13 尺度精度达到最高; 然后随着尺度的增大, 分类精度基本不变. 因此, 第 13 尺度应为最佳分解尺度. 三种最佳尺度确定方法的总体分类精度分别为, 由本文方法所确定的最佳分解尺度为第 13 尺度, 分类精度为 98.70%, 也是分类精度最高的尺度; 根据信号长度和母小波函数共同确定的最佳尺度为第 8 尺度, 对应的分类精度为 97.16%, 而由 Kaewpjit 等^[5]的方法确定的最佳尺度为第 5 尺度, 其分类精度为 95.34%. 由此看出, 本文的分解尺度确定方法取得了最好的分类结果.

3.3 讨论

通过上述两个数据对三种小波分解尺度确定方法的比较, 可以得到以下认识.

由根据信号长度和母小波函数共同确定最佳分解尺度的方法, 所确定的最佳尺度与分类精度最高的尺度较为接近, 其分类精度与最高精度也比较接近. Kaewpjit 等^[5]的方法阈值难以确定合适的最佳分解尺度. 当阈值较大时, 确定的最佳尺度较小, 分类精度很低; 而当阈值较小时, 相关系数在各尺度分布较为分散, 尺度难以确定. 根据该方法确定的最佳尺度, 往往偏小.

根据本文提出的尺度确定方法得到的最佳分解尺度, 与分类精度最高的分解尺度相一致. 因此, 相关系数的稳定性可以较准确地确定小波分解尺

度. 相关系数之所以趋于稳定的原因, 主要是因为, 当分解达到一定的尺度时, 为满足小波分解条件, 会对信号进行一定程度的延伸, 小波分解得到的近似系数相似度会越来越高, 与原始信号的相关系数也就趋于稳定. 因此, 以相关系数作为最佳分解尺度判断标准, 可以得到可靠的分类结果.

本文提出的方法根据数据本身的特点来确定最佳分解尺度, 同时, 这种方法不需要计算整个影像所有尺度的重构影像以及各重构影像与原始影像之间的相关系数, 也大幅度地降低了计算量, 而且所确定的最佳分解尺度与分类精度最高的尺度相一致, 因此, 本文提出的最佳尺度确定方法比现有方法得到的结果更准确.

4 结论

在基于小波分解的高光谱数据特征提取中, 合适的分解尺度的确定是其中的关键环节之一. 本文提出了一种新的小波分解最佳尺度的判断方法, 用于高光谱数据的特征提取. 该方法计算各地物类别训练样本的重构信号与原始信号之间的相关系数, 并以相关系数趋于稳定的尺度作为判断最佳尺度. 与现有方法相比, 该方法可更准确地确定最佳分解尺度.

致谢 文中所采用的 HYDICE 数据来源于 Landgrebe 的“Signal theory methods in multispectral remote sensing”所附带的数字光盘^[15].

参 考 文 献

- 1 Benediktsson JA, Sveinsson JR. Feature extraction for multi-source data classification with artificial neural networks. *International Journal of Remote Sensing*, 1997, 18(4): 727-740
- 2 Scott DW. *Multivariate Density Estimation: Theory, Practice, and Visualization*. New York: Wiley, 1992, 1-336
- 3 Bruce LM, Morgan C, Larsen S. Automated detection of sub-pixel hyperspectral targets with continuous and discrete wavelet transforms. *IEEE Transaction on Geoscience and Remote Sensing*, 2001, 39(10): 2217-2226
- 4 Pu RL, Gong P. Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping. *Remote Sensing of Environment*, 2004, 91(2): 212-224
- 5 Kaewpjit S, Le moigne J, El-Ghazawi T. Automatic reduction of hyperspectral imagery using wavelet spectral analysis. *IEEE*

- Transactions on Geoscience and Remote Sensing, 2003, 41(4): 863—871
- 6 Kaewpijit S, Le Moigne J, El-Ghazawi T. Feature reduction of hyperspectral imagery using hybrid wavelet-principal component analysis. *Optical Engineering*, 2004, 43(2): 350—362
 - 7 Liu ZJ, Wang CY, Zhang JX. Feature extraction and feature selection based on wavelet and genetic algorithm. *Journal of Remote Sensing*, 2005, 9(2): 176—185
 - 8 Zheng J, Regentova E. Wavelet based feature reduction method for effective classification of hyperspectral data. *Information Technology: Coding and Computing [Computers and Communications]*, 2003. Proceedings. ITCC 2003. International Conference; 483—487
 - 9 Nunez J, Otazu X, Merino MT. A multiresolution-based method for the determination of the relative resolution between images: First application to remote sensing and medical images. *International Journal of Imaging Systems and Technology*, 2005, 15(5): 225—235
 - 10 Schmidt KS, Skidmore AK. Smoothing vegetation spectra with wavelets. *International Journal of Remote Sensing*, 2004, 25(6): 1167—1184
 - 11 Mallat SG. A theory for multiresolution wavelet transforms; theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(7): 674—693
 - 12 Mallat SG. Wavelets for a vision. *Proceedings of the IEEE*, 1996, 84(4): 604—614
 - 13 Mallat SG. *A Wavelet Tour of Signal Processing (2nd edition)*. San Diego: Academic press, 1999, 1—637
 - 14 Pittner S, Kamarthi SV. Feature extraction from wavelet coefficients for pattern recognition tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(1): 83—88
 - 15 Landgrebe DA. *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, N J: Wiley, 2003, 1—508
 - 16 Camps-Valls G, Gomez-Chova L, Munoz-Mari J, et al. Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2006, 3(1): 93—97
 - 17 Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778—1790
 - 18 Benediktsson JA, Palmason JA, Sveinsson JR. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(3): 480—491