

文章编号: 1007-4619(2007)04-0568-06

基于证据理论的遥感图像分类方法探讨

邓文胜, 邵晓莉, 刘 海, 万浩方, 许 亮

(湖北大学 资源环境学院, 湖北 武汉 430062)

摘 要: 遥感图像分类是一项十分重要而且复杂的问题。传统的图像分类方法多数是基于贝叶斯主观概率理论的图像分类方法, 由于其在解决不确定性问题上存在诸多缺陷, 近年来, 将数学的证据理论应用于遥感图像分类已成为新的发展趋势。本文首先介绍了证据理论与贝叶斯主观概率理论对于解决不确定性问题的差异, 证据理论的主要定义和算法。文中还重点介绍了证据理论用于模式识别的支持度函数, 并将支持度应用于图像分类, 得到像元级按大类(如土地利用的城镇用地、耕地、林地和水体)划分的支持度表面, 然后进行硬分类叠加, 得到初次分类结果, 再进行精度评估。如果不满足精度要求, 再对各类支持度小于某域值的像元进行二次分类, 如此下去, 直到达到分类所要求的精度。该方法的主要优点是可以进行分类后的再分类, 且精度非常高, 而贝叶斯分类不可以进行分类后的再分类, 只能重新训练样本进行整体分类, 效率低, 精度也难提高。该方法理论可靠, 实用性强, 易操作, 有研究潜力。

关键词: 证据理论; 模式识别; 支持度; 遥感图像分类

中图分类号: TP79 **文献标识码:** A

Discussion of Remote Sensing Image Classification Method Based on Evidence Theory

DENG Wen-sheng SHAO Xiao-li LIU Hai WAN Gao-fang XU Liang

(School of Resource and Environmental Science Hubei University Hubei Wuhan 430062, China)

Abstract Remote sensing image classification is an important and complex problem. Conventional remote sensing image classification methods are mostly based on Bayes' subjective probability theory. Because there are many defects, a new tendency is that the mathematical theory of evidence has been applied to remote sensing image classification. At first, this paper introduces differences between Dempster-Shafer's evidence theory and Bayes' subjective probability theory and main definitions and algorithms on D-S evidence theory. Especially degree of belief, degree of plausibility and degree of support are the bridges that D-S evidence theory is used in other fields. It emphatically introduced Support function that D-S evidence theory is used on pattern recognition, and degree of support is applied to classification. We acquire degree of support surfaces according to large classes such as urban land, farm land, forest land, and water, then use "hard classification" to gain initial result. If initial classification accuracy is fitted to the acquirement, reclassification for degree of support surfaces of less than threshold is conducted until final classification result reaches accuracy. Main advantages of this method are that it can perform reclassification after classification and its classification accuracy is very high. This method has dependable theory, intensive application, easy operation and research potential.

Key words D-S evidence theory; pattern recognizes; degree of support; remote sensing image classification

收稿日期: 2006-09-01; **修订日期:** 2006-10-30

基金项目: 湖北省科技厅自然科学类项目 (编号: 2006ABA013), 湖北省教育厅自然科学类重点项目 (编号: D200510006), 武汉市科技局软科学类项目 (编号: 20054005062-28-20)。

作者简介: 邓文胜 (1964-), 男, 副教授, 2005年6月武汉大学摄影测量与遥感专业工学博士毕业, 现主要从事遥感、地理信息系统的理论与应用研究, 已发表论文 10 多篇, 参编著作两部。

1 引言

随着人类活动的加深,地表土地利用覆盖的快速变化,新的信息的获取成为人们十分关注的问题之一。随着卫星遥感数据的大量涌现,为这一问题的解决提供了条件,因此遥感图像分类成为十分重要且复杂的问题。图像分类也称图像识别,是指对图像信息进行处理、分辨、检测,确定图像上物体的属性或特征,对图像进行分类或识别^[1]。语义与非语义信息的自动提取、快速发现地表的变化以及 GIS 数据库的更新都离不开图像分类^[2]。近年来,随着计算机技术的飞速发展,计算机识别分类成了遥感应用的一个重要组成部分^[3]。

由于自然环境的复杂性,以及自然环境与遥感波谱相互作用的复杂性,从传感器记录的光谱信号中提取的关于地表的信息中,总是存在不确定性,但人们总是希望从遥感数据中提取的信息完全客观准确地反映实际情况。目前最常用的遥感图像分类方法是以贝叶斯(Bayes)的主观概率理论为代表的最大似然法,它的前提条件要求图像数据呈正态分布,而且最主要的问题是反映不出对知识的不知道信息^[4]。新的图像分类方法在不断涌现,如基于归纳学习的空间知识发现,应用到图像分类中^[5]。

Dempster-Shafer 提出的证据理论(A mathematical Theory of Evidence)是能处理各种不同来源的数据的一种科学方法^[6,7],它是研究不确定性问题的另一种理论。从诞生到现在虽然仅仅经过二十几年时间,但是理论上却取得了丰硕的成果,得到了国际学术界的普遍重视。基于证据理论的不确定性问题的研究正在进行之中^[8,9]。本文正是基于证据理论的进行遥感图像分类方法的探讨。

2 证据理论

Shafer^[10,11]认为,对于概率推断的理解,不仅要强调证据的客观性而且也要强调证据估计的主观性。数字化的概率并没有独立于人类判决的客观属性,在人思考之前也不会人的头脑中存在,但是人可以在客观证据的基础上构造出这样一个数字化的概率来。这种概率(Probability)是某人在证据的基础上构造出的他对一命题真的信任程度,简称信度。而这种理论也不称为是一种概率理论,而称为证据的数学理论或简称为证据理论^[12],通常称 D-S 证据理论。

2.1 证据理论与贝叶斯主观概率理论的主要区别

D-S 证据理论是区分不确定性和忽略性的工具,它用于图像分类比 Bayes 方法产生更详细的描述并可以替代模糊分类方法^[13]。

D-S 证据理论的信度函数是集函数而不是点函数,其取值通常是一个范围,而不是一个固定值,它反映了不确定性程度。贝叶斯理论认为一个命题为真,有且仅有一个确定的概率,是一个确定值。

D-S 证据理论定义了不确定性度量,对于命题集的信度不满足加法法则。在 D-S 证据理论中,一个命题的信度与它的补集的信度之和通常情况下小于 1。即如果一个命题的信度为 0.8 并不能推出它的否命题的信度为 0.2 它可能小于 0.2 而在 Bayes 概率理论中,它的否命题的概率一定为 0.2。

D-S 证据理论认为多个证据(命题)的信度可以合成得到一个新证据(命题),而概率理论没有相对应的计算。在 Bayes 模型中,概率被分配给感兴趣的所有命题,新的概率分配出现,旧的概率分配就被取代。在 D-S 证据理论中,只要是证据支持的命题都可以分配权重值,都对新证据产生影响。

2.2 证据理论的描述

这里以一个简单的例子来对证据理论进行一下回顾。假设有一幅遥感图像,根据其光谱特征判断大概可以分成以下 4 类:水体、城镇、林地、耕地用地,分别以 A、B、C、D 代表,即 $A = \{\text{Water}\}$, $B = \{\text{Urban}\}$, $C = \{\text{Forest}\}$, $D = \{\text{Fam}\}$ 。

设 $\Theta = \{A, B, C, D\} = \{\text{Water, Urban, Forest, Fam}\}$, 则可能存在的子集包括 $\{A, B, C, D\}$, $\{A, B, C\}$, $\{A, C, D\}$, $\{B, C, D\}$, $\{A, B, D\}$, $\{A, B\}$, $\{A, C\}$, $\{A, D\}$, $\{B, C\}$, $\{B, D\}$, $\{C, D\}$, $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$, $\{\Phi\}$, 共有 2^4 个子集。识别框架 Θ 就是所有可能的集合。确立了识别框架 Θ , 就把对于命题的研究转化成了对集合的研究。

设 Θ 为识别框架,如果集函数 $m: 2^\Theta \rightarrow [0, 1]$ (2^Θ 为 Θ 的幂集)满足:

$$m(\Phi) = 0 \quad (1)$$

$$\forall X \subseteq \Theta \quad 0 \leq m(X) \leq 1 \quad (2)$$

$$\sum_{X \subseteq \Theta} m(X) = 1 \quad (3)$$

$$\text{Bel}(X) = \sum_{Y \subseteq X} m(Y) \quad (4)$$

$$\text{Pl}(X) = 1 - \text{Bel}(X) = \sum_{X \cap Y \neq \Phi} m(Y) \quad (5)$$

则称 m 为识别框架 Θ 上的基本可信度分配 (Basic Probability Assignment), $m(X)$ 为 X 的基本可信数 (Basic Probability Number), Bel 为识别框架 Θ 上的信度函数, $Bel(X)$ 为命题 X 的信度, Pl 为识别框架 Θ 上的似真度函数, $Pl(X)$ 为命题 X 的似真度。 $Pl(X)$ 是比 $Bel(X)$ 对命题 X 更宽松的信度, 因此, 任何给定命题都有一个信度区间 $[Bel(X), Pl(X)]$ 。

对于贝叶斯概率理论所得的概率 (Probability) 可以看成是特殊的信度 (Belief)。如果 Θ 为一个识别框架, $Bel: 2^\Theta \rightarrow [0, 1]$ 满足

$$\begin{aligned} Bel(A) &= 0.15 & Pl(A) &= 1 - 0.35 - 0.2 - 0.1 = 0.35 & Pl(A) - Bel(A) &= 0.2 \\ Bel(B) &= 0.35 & Pl(B) &= 1 - 0.15 - 0.2 - 0.1 = 0.55 & Pl(B) - Bel(B) &= 0.2 \\ Bel(C) &= 0.2 & Pl(C) &= 1 - 0.15 - 0.35 - 0.1 = 0.4 & Pl(D) - Bel(D) &= 0.2 \\ Bel(D) &= 0.1 & Pl(D) &= 1 - 0.15 - 0.35 - 0.2 = 0.3 & Pl(D) - Bel(D) &= 0.2 \\ Bel(A \cup B \cup C) &= m(A) + m(B) + m(C) + m(A \cup B \cup C) = 0.15 + 0.35 + 0.2 + 0.1 = 0.8 \\ Pl(A \cup B \cup C) &= 1 - m(D) = 1 - 0.1 = 0.9 & Pl(A \cup B \cup C) - Bel(A \cup B \cup C) &= 0.1 \end{aligned}$$

按照贝叶斯概率理论, 满足可加性原理

$$\begin{aligned} Bel(A \cup B \cup C) &= Bel(A) + Bel(B) + Bel(C) \\ &= 0.15 + 0.35 + 0.2 = 0.7 \neq 0.8 \end{aligned}$$

因此, D-S 证据理论能够表达部分信度, 是解决不确定性问题理想模型。

3 证据理论在遥感图像分类中的应用

3.1 图像分类

遥感图像分类 (图像识别) 实质上就是确定某一像元与一系列模式中的哪一个模式相似的问题。若两个像元模式 x 与 y 相似, 即像元 x 与 y 的特征向量相似, 则两特征向量的距离最短。

设 x 与 y 分别为遥感图像的两个向量, 它们各包含 m 个特征, 即:

$$\begin{aligned} x &= (x_1, x_2, \dots, x_m)^T \\ y &= (y_1, y_2, \dots, y_m)^T \end{aligned}$$

则可用欧氏距离 $\|x - y\|$ 来描述 x 与 y 的相似程度

$$\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2}$$

若 $\|x - y\|$ 比较小, 则说明 x 与 y 的各个特征相差不多, 即 x 与 y 相似; 若 $\|x - y\|$ 比较大, 则说明 x 与 y 差异性显著而不相似。

如果一个像元 x 属于类 A 则可以看成像元 x 与类 A 的中心的均值向量相似。

设 $A = \{x_1, x_2, \dots, x_n\}$, 每一个 x_i 又都有 m 个特征, 设为

$$x_i = (x_{i1}, x_{i2}, \dots, x_{im})^T \quad (i=1, 2, \dots, n)$$

令

$$(1) \quad Bel(\phi) = 0$$

$$(2) \quad Bel(\Theta) = 1$$

$$(3) \quad \text{若 } X, Y \subset \Theta \text{ 且 } X \cap Y = \phi, \text{ 那么 } Bel(X \cup Y) = Bel(X) + Bel(Y)$$

则称 Bel 为贝叶斯信度函数。

假设统计分析得到 4 类的基本可信度分配为

$$\begin{aligned} m(A) &= 0.15; \quad m(B) = 0.35; \quad m(C) = 0.2; \\ m(D) &= 0.1; \quad m(A \cup B \cup C) = 0.1 \end{aligned}$$

用 Shafer 的证据理论, 计算信度、似真度及其差值分别为:

$$y^j = \frac{1}{n} \sum_{i=1}^n x_i^j \quad (j = 1, 2, \dots, m)$$

则可构造出类 A 的均值向量 $y = (y^1, y^2, \dots, y^m)$ 。像元 x 属于类 A 也即 x 与 A 的均值向量 y 的欧氏距离 $\|x - y\|$ 最短。这只是表示了 x 与类 A 的相似程度, 然而是否还有可能与类 B 的距离相当, 则无法判断, 并没有反映 x 属于 A 的信度或似真度或支持度。

3.2 基于证据理论的图像分类算法

设用 θ_i 代表 $x \in A_i (i=1, 2, \dots, n)$, 则可用 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ 作为识别框架。识别证据的作用就是在框架 Θ 上产生一个似真函数 Pl (段新生, 1993)。此处约定

$$Pl(\{\theta_i\}) = \frac{C}{\|x - A_i\|} \quad (C \text{ 为常数}, i=1, 2, \dots, n)$$

$$\begin{aligned} \forall A \subset \Theta \quad Pl(A) &= \max_{\theta_i \in A} Pl(\{\theta_i\}) = \max_{\theta_i \in A} \frac{C}{\|x - A_{\theta_i}\|} \\ &= \frac{m \cdot \min_{\theta_i \in A} \|x - A_{\theta_i}\|}{\min_{\theta_i \in A} \|x - A_{\theta_i}\|} \end{aligned}$$

由此即得到识别证据在识别框架 Θ 上产生的支持函数 S

$$S(A) = 1 - \frac{m \cdot \min_{\theta_i \in A} \|x - A_{\theta_i}\|}{\min_{\theta_i \in A} \|x - A_{\theta_i}\|}$$

得到支持函数 S 以后, 即可以根据该支持函数进行识别。可以取使 $S(A)$ 最大的 θ_i 将 x 归属入 A_i 类。如果这样识别的话, 则实际上就是说 x 与哪个类的代表集整体上更相似, 则 x 就归属于哪个类。

3.3 应用举例

以 2005 年孝感 Quickbird 图像为例,假设分成 4 种土地覆盖类型(水体用地、城镇用地、林地、耕地),以这 4 种地类作为变量来运用 D-S 证据理论对图像进行分类。按下面 5 步进行:

(1) 原始图像(xing)数据预处理。选择能显示最有用信息的光谱波段组合,以便获得一幅比较清晰、对比度强、位置准确的图像以提高分类精度;

(2) 训练区的选择。从待处理数据中抽取具有普遍性、代表性的数据作为训练样本。训练区选择得好与否,训练样本数是否足够,关系到分类精度的高低;

(3) 特征提取。研究感兴趣区域的光谱特征,统计最大值、最小值和平均值,作为证据,派生出确定量;

(4) 确定判别函数,图像分类运算。按照相似性原则即最大似真度和最大支持度,得到按大类划分的信度表面,对各大类信度表面二值化叠加得到初始分类结果;

(5) 分类结果精度评估。与原始影像对照,进行精度检验。一般一次分类很难满足要求,往往要进行再分类,对各大类的信度表面再确定一定的域值,对各大类在域值以下的像元重新进行第 2 步至第 5 步,直到所有的像元分类都达到满意为止。

选取波段 3、4 和 5 提供最有用的亮度值,而且没有太大的重叠和数据冗余,3 个波段结合用来监督训练样地。根据各类在该区域中所占比例不同,选择样地数目和像元数不同。训练样地要分散选择且要考虑代表性,能够反映这一区域的变化特征。训练样地均值统计结果如图 1 所示。

如果按传统的“硬分类”方法,按均值分类必然会产生较大误差,因为均值只代表平均状况,实际上有许多像元在最小值与最大值区间内难以区分。在这种情况下,要么重新训练样地,研究满意的光谱响应特征,要么通过人工编辑修改样地的最小值和最大值,以减少重叠。在任何一种情况下,都不能保证成功。因为重新训练样地不一定产生更好的结果,调整训练样地统计值可能造成信息损失,而产生大量未分类像元。

与其修改训练样地或光谱特征值,不如把数据直接应用到 Dempster-Shafer 算法中。按照 3.2 节中的基于证据理论的图像分类算法,用 ERDAS 软件进行建模处理,找出各像元与上述 4 类模式中的哪个模式最接近,即似真度最大,支持度最高。根据以上算法

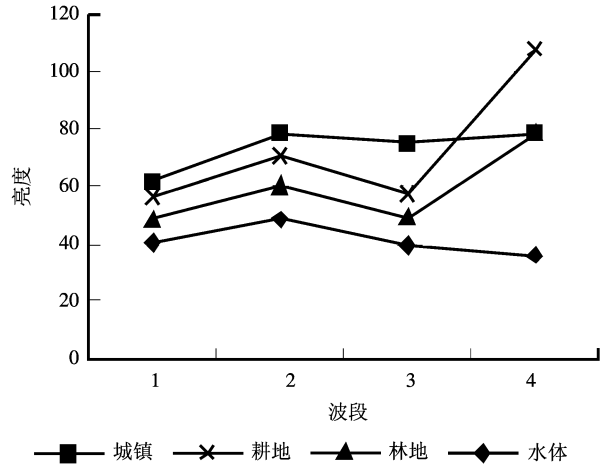


图 1 各类训练样地的平均亮度

Fig 1 Average brightness values of all kinds of training sites

得到的基于证据处理的分类图像如图 2 所示。

4 分类结果评价

图 2(a)~(d)显示了 4 种类型的基于证据理论分类的支持度表面,与分类图或原始影像对比显示,表面上看图 2(a)水体用地和图 2(d)耕地用地的分类精度较高,图 2(b)城镇用地和图 2(c)林地分类精度较低。如果把各类对支持度分类结果进行二值化“硬分类”处理,支持度为 0 的划分为一类,支持度不为 0 的划分为另一类,水体、城镇、林地和耕地分别赋值 1、2、3、4,然后将 4 类二值化图又叠加,即得 4 类合成结果,如图 3(a)所示。

图 3(a)的分类结果,按比例随机选取 270 个点,与参考图或原始影像目视解译类值进行对比,得到分类精度评估报告(表 1)。

表 1 分类混淆矩阵及精度评估报告

Table 1 Classification error matrix and the accuracy assessment report

类型	水体	城镇	林地	耕地	类样本数	分类精度
水体	71	2	1	1	75	0.9467
城镇	6	39	0	0	45	0.8667
林地	0	21	69	30	120	0.5750
耕地	0	0	3	27	30	0.9000

总检验样本: 270 正确分类样本: 206 总分类精度: 0.7629

表 1 的分类精度评价结果表明,总分类精度 0.7629,其分类精度跟最大似然法分类(经过 6 次分类试验)精度相当,后者的分类总精度为 0.7415,显

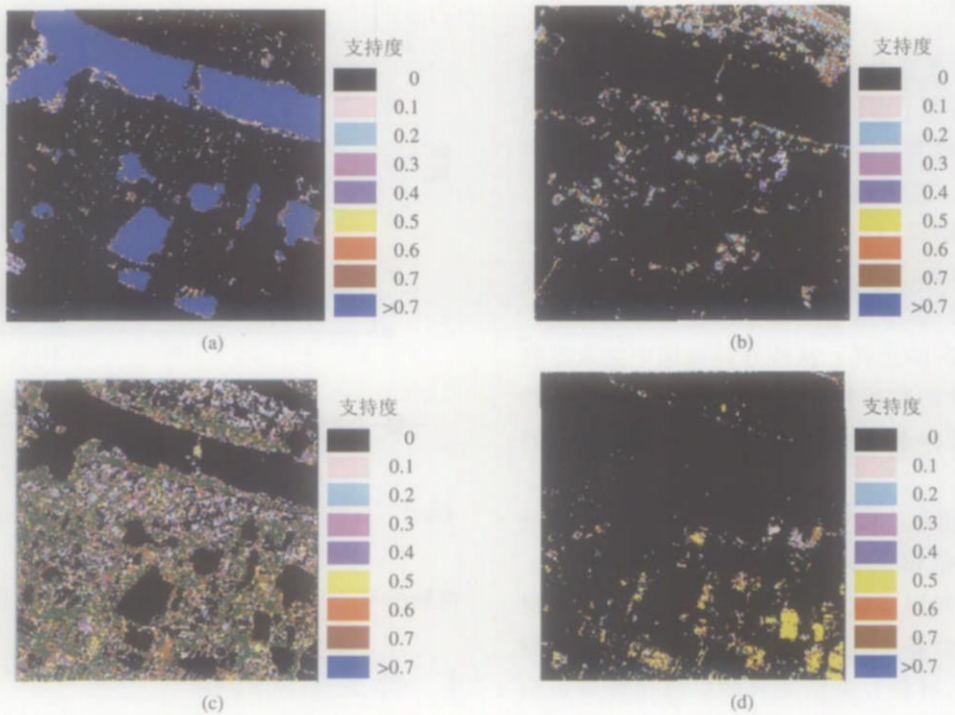


图 2 证据支持度表面
 (a) 水体覆盖; (b) 城镇覆盖; (c) 林地; (d) 耕地
 Fig.2 Surface of support degree
 (a) Water area; (b) Urban area; (c) Woodland; (d) Farmland

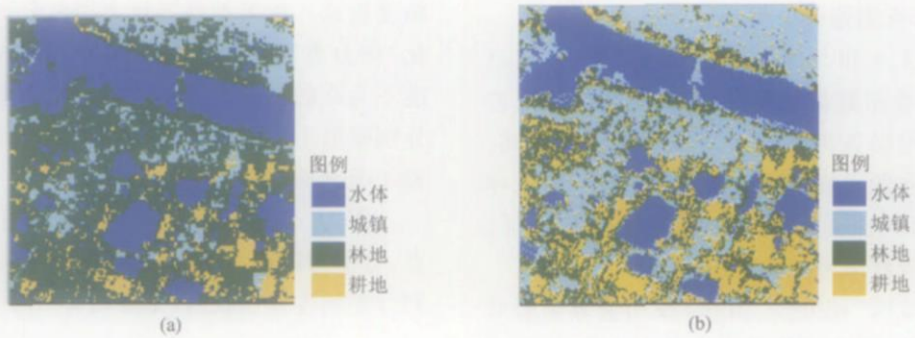


图 3 基于证据理论的图像分类结果图
 (a) 初始分类; (b) 二次分类
 Fig.3 The result of image classification based on evidence theory
 (a) The first classification; (b) The second classification

然前者比后者更符合分类要求,但其中林地分类精度仍然太低,前者可以设定域值进行二次分类,但后者不能。考虑到耕地在该区域所占总面积大,经过多次修改域值进行试验,确定对水体取支持度小于 0.2,城镇取支持度小于 0.4,林地取支持度小于 0.7,耕地取支持度小于 0.3的这部分区域重新进行特征提取,按上述方法再进行二次分类,结果如图 3(b)所示。选择同样的 270 个样本(参考点),

经检验得分类总精度为 0.9111,都较第一次分类结果和最大似然法分类结果大大提高,各类分类精度都超过 80%,分别为 0.9600, 0.8750, 0.9000, 0.9500,完全达到分类要求。

5 结 论

基于证据理论的遥感图像分类,一开始并不产生

一幅分类专题地图,而是产生按大类分类后的信度或似真度或支持度表面,在实践中是有用的和可行的。

证据理论进行分类属于“软分类”,能直接反映分类的不确定性程度,经过初次分类后,对于某些区域不确定性程度太低,可以对局部区域重复进行以上的分类步骤,直到满意为止。但贝叶斯主观概率分类属于“硬分类”,一次分类不满意,就要重新选择训练样地或者增加分类数重新分类,效率太低,且不能保证精度一定会提高。

正因为如此,证据理论是用近似推理来模拟主观判断,它在遥感图像分类中能使主观判断与人类专家知识结合,不需要过多的假设条件来使数据按照人们主观需要的方向进行分类,它为遥感图像分类提供了一个新的研究领域。

参考文献 (References)

- [1] Guan Z Q. Study on the Theory of Remotely Sensed Imagery Analysis Based on Quotient Space Description [D]. PhD Dissertation. Wuhan: Wuhan Technical University of Surveying and Mapping, 1995. [关泽群. 商空间下的遥感图像分析理论探讨 [D]. 博士论文,武汉:武汉测绘科技大学,1995.]
- [2] Li D R, Guan Z Q. Integrated Approach and Realization of Spatial Data [M]. Wuhan: Wuhan Technical University of Surveying and Mapping Press, 2000. [李德仁,关泽群. 空间信息系统的集成与实现 [M]. 武汉:武汉测绘科技大学出版社,2000.]
- [3] Sun J B. Remote Sensing Principle and Method [M]. Wuhan: Wuhan University Press, 2003. [孙家柄. 遥感原理与方法 [M]. 武汉:武汉大学出版社,2003.]
- [4] Richards A. Remote Sensing Digital Image Analysis: an Introduction [M]. 2nd Ed., Berlin: Springer-Verlag, 1993.
- [5] Di K C, Li D R, Li D Y. Remote Sensing Image Classification Method Study on Spatial Data Mining [J]. Journal of Wuhan Technical University of Surveying and Mapping, 2000, 25 (1): 43-47. [邱凯昌,李德仁,李德毅. 基于空间数据挖掘和的遥感图像分类方法研究 [J]. 武汉测绘科技大学学报, 2000, 25 (1): 43-47.]
- [6] Yager R J, Kacprzyk J, Fedrizzi M. Advances in the Dempster-Shafer Theory of Evidence [M]. New York: John Wiley and Sons, 1994.
- [7] Lein J. Environmental Decision Making: an Information Technology Approach [M]. Malden, MA: Blackwell Science, 1997.
- [8] Sarabjot S Anand, David A Bell, John G Hughes. EDM: a General Framework for Data Mining Based on Evidence Theory [J]. Data & Knowledge, 1996, 18: 189-223.
- [9] Mertikas P, Zervakis M E. Exemplifying the Theory of Evidence in Remote Sensing Image Classification [J]. Int J Remote Sensing, 2001, 22 (6): 1081-1095.
- [10] Shafer G. Mathematical Theory of Evidence [M]. Princeton, New Jersey: Princeton University Press, 1976.
- [11] Shafer G. Probability Judgment in Artificial Intelligence and Expert System [M]. Working Paper No. 165, School of Business, The University of Kansas, Lawrence, 1984.
- [12] Duan X S. Evidence Theory and Decision and Artificial Intelligence [M]. Beijing: China Renmin University Press, 1993. [段新生. 证据理论与决策、人工智能 [M]. 北京:中国人民大学出版社,1993.]
- [13] Lein J K. Applying Evidential Reasoning Methods to Agricultural Land Cover Classification [J]. Int J Remote Sensing, 2003, 24 (21): 4161-4180.