

基于商空间理论的遥感图像分割

徐 怡¹, 李龙澍², 李学俊¹

(1. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039;

2. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039)

摘 要: Sankar K. Pal 等最近提出了一种基于“粗糙熵”的图像分割算法, 主要是按照目标和背景这两大类对图像进行分割, 不足之处在于不能满足多类目标提取的需要。为此, 基于商空间的粒度分解和粒度合成原理, 综合粗糙集和聚类算法对之进行改进。通过对遥感图像进行分割处理, 证明了改进后算法的有效性。

关键词: 商空间; 粗糙集; 聚类; 遥感图像分割

中图分类号: TP18

文献标识码: A

文章编号: 1673-629X(2007)09-0058-03

Remote Sensing Image Segmentation Based on Quotient Space

XU Yi¹, LI Long-shu², LI Xue-jun¹

(1. Ministry of Education Key Lab. of IC & SP, Anhui University, Hefei 230039, China;

2. Department of Computer Science and Technology, Anhui University, Hefei 230039, China)

Abstract: An image segmentation algorithm based on “rough entropy” is proposed by Sankar K. Pal, etc. In recently, mainly according to two classes: object and background to segment an image. Defect of the algorithm is unsatisfactory for multi-class segmentation problem. Therefore, rough set and cluster methods are integrated based on granularity decompose and granularity synthesis theory of quotient space to improve this algorithm. The experimental result of remote sensing image segmentation demonstrate that the improved approach is valid.

Key words: quotient space; rough set; cluster; remote sensing image segmentation

0 引 言

图像分割是将图像中有意义的特征或区域提取出来的过程。分割的结果是把图像划分成若干互不相交的区域, 使各区域具有一致性, 而相邻区域间的属性特征有明显的差别。图像分割是从图像处理到图像分析的关键步骤, 是模式识别、图像理解、计算机视觉等领域的一个最重要基础环节, 为进一步对图像的分析 and 理解打下基础。目前图像分割的算法主要有阈值分割、聚类法、边缘提取分割法、区域生长法^[1]等。但是, 已有的这些算法都是独立的, 每一种算法都有其自身的局限性。在文献[2]中提出了图像分割的商空间粒度原理, 指出各种分割方法之间是有内在联系的, 它们都是利用像素特征数值分类或聚类技术实现对图像的

划分。图像分割实质与商空间粒度思想是一致的。用商空间模型来描述图像分割过程, 建立统一的图像分割理论框架是可行的。

基于这一原理, 针对文献[3]中算法的不足, 提出笔者的改进算法。在文献[3]中提出了一种基于粗糙熵的图像分割算法, 通过选取最佳的目标和背景分割阈值, 将目标从背景中分离出来。该算法的不足之处在于它对图像的分割仅仅是基于目标和背景这两大类来分割提取, 并没有考虑到多目标之间的不同细节, 在多目标提取的精确性方面还不够。为此, 一种可行的解决办法是, 先基于商空间的粒度分解原理, 用粗糙集和 K-均值聚类算法^[4], 对图像进行分割预处理, 把图像分成各个独立不相交的目标背景块; 然后再用文献[3]中基于粗糙熵的方法进行细化; 最后利用商空间的粒度合成原理, 对这两个方面进行合成, 从而得到更好的多目标提取效果。

我们把该算法用在遥感图像中, 通过实验对比, 证明该算法比单独使用粗糙集和 K-均值聚类算法或基于粗糙熵的方法, 在图像目标提取的性能方面有很大提高。

收稿日期: 2006-11-13

基金项目: 国家自然科学基金项目(60273043); 安徽省高校拔尖人才基金项目(05025102); 安徽省自然科学基金项目(050420204)

作者简介: 徐 怡(1981-), 女, 安徽凤阳人, 博士研究生, 研究方向为不精确信息处理; 李龙澍, 教授, 博导, 研究方向为智能软件、知识工程、软件体系结构。

1 商空间原理

1.1 商空间的定义

用一个三元组描述一个问题 (X, f, T) , X 表示问题的论域, f 表示论域的属性, T 是论域的结构, 指论域 X 中各元素的相互关系。 f 可用函数 $f: X \rightarrow Y$ 表示, 其中 Y 可以是实数集合, 可以是 n 维空间 R 中的集合, 也可以是更一般的空间。 f 可以是单值, 也可以是多值的。对论域中任一元素 $x \in X$, 有一个相应的 $f(x)$, 表示元素 x 的某些属性, 所以 f 又称属性函数。

1.2 粒度计算

粒度概念可以用子集来表示, 不同粒度的概念就体现为不同粒度的子集, 一簇概念就构成空间的一个划分, 即商空间(知识基), 不同的概念簇就构成不同的商空间。粒度计算就是研究在给定商空间上的各种子集合之间的关系和转换, 以及对同一问题取适当的粒度, 并在对不同粒度的研究中, 综合获取原问题的解。

对问题 (X, f, T) , 称从不同的粒度(角度、层次)考察问题 (X, f, T) , 是指给定 X 的一个等价关系 R , 并由 R 产生一商集 $[X]$, 然后研究相应问题 $([X], [f], [T])$, 其中

* $[X]$: 定义 $[X]$ 为对应于等价关系 R 的商集。

* $[T]$: 设 T 是拓扑, 则定义商拓扑 $[T] = \{u \mid p^{-1}(u) \in T, u \in X\}$, $p: X \rightarrow [X]$ 是自然投影。

* $[f]$: 设属性函数 $f: X \rightarrow Y$, 定义 $[f]: [X] \rightarrow Y$ 。

这样就得到一个新的问题 $([X], [f], [T])$, 称 $([X], [f], [T])$ 是问题 (X, f, T) 对应于等价关系 R 的商空间(商问题), 其中 R 对应于一定的粒度。 X 的所有不同的商集及其对应的商空间, 就构成问题 (X, f, T) 的不同粒度世界。

1.3 等价关系与划分

等价关系是拓扑学的基本概念, 是商空间粒度计算理论的关键。

定义: 设 X 是一集合, R 是 X 上的一个关系, 并满足

(1) 自反性, xRx

(2) 对称性, 若 xRy , 则 yRx

(3) 传递性, 若 xRy, yRz , 则 xRz

则称 R 是 X 上的一个等价关系, 对等价关系 R , 通常将 xRy 记为 $x \sim y$ 表示为 $[x]_R = \{y \mid yRx, y \in X\}$, $[x]_R$ 是 X 的一个划分。令 $[X] = \{[x]_R \mid x \in X\}$, 称 $[X]$ 是 X 关于 R 的商集。故商集是将等价类 $[x]_R$ 看成元素而构成的新空间。

1.4 商空间关系推理

商空间理论就是研究不同粒度之间的关系, 粒度的合成、综合、分解和在商空间中的推理。推理过程中

用到两个重要的原理^[5]:

(1) 保假原理: 若一个命题在粗粒度空间中为假, 则该命题在比它细的空间中一定也无解。

(2) 保真原理: 若一个命题在两个较粗粒度的商空间中为真, 则(在一定条件下)在其合成的商空间中对应的问题也是真的。

这两个原理在商空间模型的推理中起到很重要作用。对一个问题进行求解, 当问题十分复杂时, 常先进行初步分析, 将问题化成在较粗粒度空间上的问题, 然后进行求解, 若得出该问题在粗粒度空间中是无解, 则由“保假原理”, 可以判断原问题是无解。由于粗粒度的空间规模小, 故计算量也少, 这样就可以用很少的计算量得出所要的结果。反过来对于一个复杂问题, 可以把它分解为两个较细空间上的子问题, 如果在两个较细空间上进行求解, 得出对应的问题有解。利用“保真原理”可得, 在其合成的空间上问题也有解。从而将复杂性从“相乘”降为“相加”, 同样降低了求解的复杂度。

2 改进的遥感影像分割算法

设初始遥感图像用商空间方法表示为: $([X], [F], [T])$, 具体算法可分为如下几个步骤:

第一步: 把粗糙集理论和K-均值聚类算法相结合对遥感图像进行分割。

具体算法见文献[4], 下面介绍的是主要算法思想。在K-均值聚类算法中, 第一步初始聚类中心的选择很重要。一般是随机选取待聚类样本集的 K 个样本, 然而聚类的性能与初始聚类中心的选取有关, 聚类的结果与样本的位置有极大的相关性, 一旦这 K 个样本选取不合理, 将会增加运算的复杂程度, 误导聚类过程, 得到一个不合理的聚类结果。对此可以通过粗糙集理论的近似分类功能, 提供K-均值聚类所需要的初始类的个数和均值, 提高聚类的效率和分类的精度。初始聚类中心选取后, 计算各个样本到聚类中心的距离, 把样本归到离它最近的那个聚类中心所在的类, 对调整后的新类计算新的聚类中心。如果相邻两次的聚类中心没有任何变化, 说明样本调整结束, 聚类准则函数 J_C 已经收敛^[6]。当所有的初始中心都已经收敛时, 整个聚类算法结束。从而得到图像的分割结果, 用商空间理论可表示为 $([X]_1, [F]_1, [T]_1)$ 。

第二步: 把 $[X]_1$ 作为研究对象, 利用基于粗糙集的算法对各个子目标背景块进行分析。

算法主要思想是, 首先将图像分成一定大小的粒度, 即窗口。窗口的大小是依据整幅图像的亮度直方图, 计算直方图中对应各峰值的区域宽度, 选择最小宽

度的一半作为窗口的长和宽。然后利用粗糙集的上近似和下近似概念,对图像不同的灰度级别 T ,分别计算目标和背景的上下近似集合,可表示为: $O_T^-, O_{T-}, B_T^-, B_{T-}$,在此基础上,分别定义目标和背景的粗糙程度如下: $R_{OT} = 1 - |O_{T-}| / |O_T^-|, R_{BT} = 1 - |B_{T-}| / |B_T^-|$,其中 $||$ 表示集合的基。利用粗糙程度就可以定义粗糙熵为: $RE_T = -e/2[R_{OT} \log_e(R_{OT}) + R_{BT} \log_e(R_{BT})]$,最后选取粗糙熵最大值,即 $T^* = \arg \max RE_T$,作为最佳的目标和背景分割阈值,将目标从背景中分离出来。按照上面的方法对每一个子目标背景块进行分析后,可得到图像的另一个分割结果,用商空间可表示为 $([X]_2, [F]_2, [T]_2)$ 。值得注意的是,在文献[3]算法中,窗口大小的选取是基于整幅图像的灰度直方图的峰值区域宽度,在算法中,只要基于特定子目标和背景的灰度直方图进行选取就可以了,从这个意义上说,也简化了算法复杂性,提高了算法效率。

第三步:利用商空间的粒度合成原理,对这两步的处理结果进行论域合成。

定义合成原则是求 $([X]_1, [F]_1, [T]_1)$ 和 $([X]_2, [F]_2, [T]_2)$ 的交集,定义 $[X]_3 = \{A_i \cap B_j \mid A_i \in [X]_1, B_j \in [X]_2\}$,分别取论域 $[X]_1$ 和 $[X]_2$ 中像素集合,按区域 A_i 与 B_j 中像素位置相同则同属原集合,位置不同则产生新集合的原则,进行运算,新产生的集合用 C_k 表示。对集合 C_k 中的每个像素,根据像素位置,计算每个像素相邻区域内的特征均值,并与该像素特征值进行比较,取差值最小的邻域作为该像素的区域,对 C_k 中所有像素进行重新分类。根据 C_k 的分类结果对 $[X]_3$ 中的集合更新,从而得到最终的分割结果。

3 实验结果

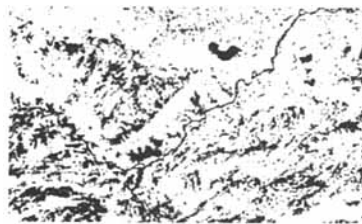
为了证明上述算法的有效性,对一幅遥感图像分别采用基于粗糙集和 K-均值聚类的算法、基于粗糙熵的算法和文中提出的基于商空间的改进算法进行处理。为了便于比较,对于粗糙集和 K-均值聚类算法的处理结果也进行了目标和背景的阈值化处理,最终结果如图 1 所示。



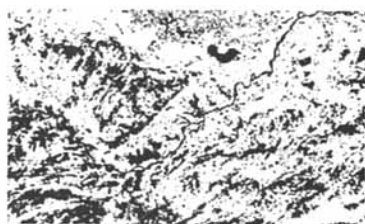
(a) 原始的遥感图像



(b) 基于粗糙集和 K-均值聚类算法处理后的遥感图像



(c) 基于粗糙熵的算法处理后的遥感图像



(d) 基于商空间的改进算法处理后的遥感图像

图 1 各算法结果比较

通过实验结果可以看出,基于商空间的改进算法比单独使用粗糙集和 K-均值聚类算法或基于粗糙熵的方法,在图像目标提取的性能方面有很大提高。

4 结论

商空间理论作为研究不同粒度世界的一种新的数学工具,具有较完备的理论基础。文中基于商空间的粒度分解和粒度合成原理,对已有的图象分割算法进行改进,通过实验可以看出该算法对于低对比度、多层次变化背景的遥感图像有很好的目标分类效果,从而证明了改进后算法的有效性。目前关于商空间理论在实际应用中的研究还很少,有待更多学者的进一步关注。

参考文献:

- [1] 杨晖,曲秀杰. 图像分割方法综述[J]. 电脑开发与应用, 2005, 18(3): 21-23.
- [2] 刘仁金,黄贤武. 图像分割的商空间粒度原理[J]. 计算机学报, 2005, 28(10): 1680-1685.
- [3] Pal S K, Shankar B U, Mitra P. Granular computing, rough entropy and object extraction[J]. Pattern Recognition Letters, 2005, 26: 2509-2517.
- [4] 邵锐,巫兆聪,钟世明. 粗糙集理论在遥感影像分割中的应用[J]. 地理空间信息, 2005, 3(5): 26-28.
- [5] 张钹,张铃. 问题求解理论及应用[M]. 北京:清华大学出版社, 1990.
- [6] 边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社, 2000.