

粗糙集理论在图像处理中的应用

李孟歆^{1,2}, 吴成东²

(沈阳建筑大学 信息与控制工程学院, 辽宁 沈阳 110168;

2. 东北大学 信息科学与工程学院, 辽宁 沈阳 110004)

摘要:将粗糙集理论应用到图像处理中,在很多场合具有比传统方法和精确算法更好的效果。介绍了粗糙集理论的基本概念和模型的发展状况,描述了粗糙集用于图像处理的特点,详细论述了粗糙集理论在图像滤波和平滑、图像增强、图像分割以及图像特征提取与分类的研究方法与应用。研究表明粗糙集理论与其他智能方法的结合在图像处理中具有良好的应用前景。大数据集、高效约简算法、并行计算以及混合智能算法研究等问题仍是粗糙集研究的主要方向。

关键词:粗糙集;图像处理;模型;混合算法

中图分类号:TN911.73

文献标识码:A

文章编号:1673-629X(2009)03-0208-04

Rough Set Theory and Its Applications in Image Processing

LI Meng-xin^{1,2}, WU Cheng-dong²

(1. School of Information and Control Engineering, Shenyang Jianzhu University, Shenyang 110168, China;

2. School of Information Science & Engineering, Northeastern University, Shenyang 110004, China)

Abstract: In order to reach better results than traditional methods and accurate algorithms, rough sets is induced in image processing. The basic concepts in rough sets and development of rough sets models are introduced. The characteristics are suited for image processing. The applications of rough sets for image filter, image enhancement, image segmentation, feature extraction and classification are described detailed. It shows the hybrid methods combined rough sets and other intelligent methods are used to result in better performance in image processing field. The research for large data sets, highly effective reducts algorithms, parallel algorithms and hybrid intelligent methods is still main direction for rough sets.

Key words: rough sets; image processing; model; hybrid algorithm

0 引言

粗糙集理论是波兰学者 Pawlak 提出的一种刻划不完整性和不确定性的软计算数学工具^[1],它为信息科学和认知科学提供了新的科学逻辑和研究方法。它能有效地分析不精确、不一致、不完整等各种不完备信息,对数据进行分析 and 推理,从中发现隐含的知识,揭示潜在的规律。该理论处理不确定信息的优势是无需提供任何先验信息,对不确定性的描述或处理比较客观。

目前,粗糙集已成为人工智能领域中的学术热点,在机器学习、知识获取、决策分析、过程控制、模式识别等许多领域中得到了广泛的应用^[2]。

1 粗糙集理论基本概念

1.1 知识表达系统

知识表达在智能数据处理中占有十分重要的地位。 $S = (U, A, V, f)$ 是一个知识表达系统,其中, U 为对象的非空有限集合,称为论域; A 为属性的非空有限集合; $V = \bigcup_{a \in A} V_a$, V_a 是属性 a 的值域; $f: U \times A \rightarrow V$ 是一个信息函数,它为每个对象的每个属性赋予一个信息值,即: $a \in A, x \in U, f(x, a) \in V_a$ 。如果 $A = C \cup D, C \cap D = \emptyset$, C 称为条件属性集, D 称为决策属性集。具有条件属性和决策属性的知识表达系统称为决策系统。知识表达系统通常也用 $S = (U, A)$ 来简化表示。

1.2 不可分辨关系

对于一个知识表达系统,令 $P \subseteq A$, 定义属性集 P 的不可区分关系 $\text{ind}(P)$ 为 $\text{ind}(P) = \{(x, y) \in U \times U \mid \forall a \in P, f(x, a) = f(y, a)\}$ 。如果 $(x, y) \in \text{ind}(P)$ 则称 x 和 y 是 P 不可分辨的。若 $P \subseteq A, X \subseteq U$, 称 $PX = \{x \in U \mid [x]_{\text{ind}(P)} \subseteq X\}$, $\underline{PX} = \{x \in U \mid$

收稿日期:2008-07-16

基金项目: 建设部项目(2007-K4-9); 辽宁省教育科研项目(05L362)

作者简介: 李孟歆(1972-),女,博士,副教授,硕士研究生导师,从事图像处理、模式识别研究。

$\{x\} \text{ind}(P) \cap X \neq \emptyset$ 分别为 X 的 P 下近似和 P 上近似。 $\text{POSP}(X) = PX$, $\text{NEGP}(X) = U - PX$, $\text{BNP}(X) = PX - \underline{PX}$ 分别称为 X 在 P 下的正域、负域和边界。 $\underline{P}(X)$ 是 U 中根据现有的信息(不可分辨关系)可确定划归入 X 的不可分辨的元素集合; $\text{NEGP}(X)$ 是肯定不属于 X 的不可分辨类的元素集合; BNP 是根据现有的信息,这些元素可能属于 X ,也可能不属于 X ;而 PX 则是所有可能属于 X (包括确定属于 X ,即 PX) 的不可分辨类的元素集合。可以看出,粗糙集理论与人的认知特性很相似。

1.3 约简

知识约简是粗糙集理论的核心内容之一。知识约简,就是在保持知识库分类能力不变的条件下,删除其中不相关或不重要的知识。对于任一属性 $a \in C$,如果 $\text{ind}(C) \neq \text{ind}(C - \{a\})$,则称 a 为 C 中不可省略的;否则, a 为 C 中可省略的。决策规则的约简是利用决策逻辑分别消去每一个决策规则的不必要条件,它不是整体上的简化属性,而是针对每一个决策规则,去除表达该规则的冗余属性值,以便进一步简化决策规则。令 $F = \{X_1, X_2, \dots, X_n\}$ 为一属性值集合,对于任一子集 $X_i \in F$,如果 $\bigcap (F - \{X_i\}) \neq \bigcap F$,则称 X_i 为 F 中不可省略的;否则 X_i 为 F 中可省略的。

2 粗糙集模型

为了使粗糙集理论更能适应不确定信息的处理,不少研究对传统粗糙集进行拓展,如为处理含有噪声的决策表,Ziarko 提出了变精度粗糙集^[3];为处理信息不完备的决策表,Kryszkiewicz 提出基于容差关系的拓展粗糙集^[4],Stefanowski 提出基于量化容差关系的拓展粗糙集^[5],王国胤提出基于限制容差关系的拓展粗糙集^[6];为处理决策表中的模糊信息,有学者提出模糊粗糙集和粗糙模糊集等^[7];为处理含有偏好信息的多属性问题,Greco 提出优势关系粗糙集理论^[8],用优势关系代替不可分辨关系。菅利荣、刘思峰提出了偏好多属性概率决策的扩展粗糙集方法^[9],这些模型为解决不确定信息问题提供了理论基础,也拓宽了粗糙集的应用范围。但粗糙集理论的实用性还不够完善,值得深入研究。

3 粗糙集理论用于图像处理

由于图像信息具有较强的空间复杂性、相关性,处理过程中会遇到不完整性和不确定性问题,因此将粗糙集理论应用到图像处理中,在很多场合具有比传统方法和精确算法更好的效果。目前,粗糙集已经应用于图像增强、图像分割、图像滤波、图像修复等领域。

3.1 图像滤波和平滑

图像滤波涉及到噪声滤除和图像细节保留的兼顾和平衡,传统方法在有效去除噪声的同时,会使图像细节受损。粗糙集理论应用于图像滤波的基本思想是将每一幅图像表达的信息作为一个知识系统,利用粗糙集中的近似集合概念检测并处理图像中的噪声像素。

Zbigniew 和 Smart 提出基于粗糙集的图像滤波方法,使用上近似检查窗体中央的连续性,该方法可以去除噪声,加强边缘。杨阳等提出了一种基于粗糙集的图像中值滤波算法^[10]。根据图像灰度和邻域灰度均值相似程度,建立粗糙集等价关系和相应划分集合,然后利用它判断图像中的噪声。胡旺等针对图像滤波方法对噪声类型具有偏好性的不足,提出了一种基于粗糙集的多向差分图像去噪方法^[11]。根据多向差分图像提出了描述像素平稳性的方向数和描述噪声性的方向极值数的概念,采用粗糙集理论按照像素的平稳性和噪声性将图像划分为平稳噪声子图、脉冲噪声子图以及正常像素子图,并对两类噪声子图分别设计了不同的自适应滤波算法。陈铁民和王卫星提出一种基于粗糙集阴影边缘方法^[12],根据粗糙集理论、梯度、最大邻域差及噪声的条件属性,将一幅图像划分为不同的子图像,然后对子图像分别进行处理,得到阴影边缘点,再对边缘点进行细化和跟踪,删除那些假边缘点,最后得到阴影边缘图像。胡小梅和李众立提出粗糙集彩色图像滤波算法^[13],能有效去除彩色图像的椒盐噪声,并较好地保持了图像的边缘细节信息。郭海涛等人利用粗糙集理论进行声纳图像平滑^[14],具有较强的保护图像细节和边缘的能力。吕振肃等人提出了改进的基于粗糙集的图像平滑方法^[15],通过像素点的粗糙集近似相似度来判断像素所属的掩模,对图像进行平滑处理,既去除了噪声,保持了细节,且增加了图像的对比度。

3.2 图像增强

图像增强处理是改善图像质量的重要技术,它是将图像中感兴趣的信息有选择地突出、强化,比如突出轮廓、细节,消除噪声等,目的是增加图像的对比度,改善灰度层次,提高图像的“可懂度”。徐立中等提出一种粗糙集图像增强方法^[16]属性把图像分成不同子图,对子图进行对比度增强。叶灵伟等按属性对图像作粗糙分类^[17]增强变换,得到增强输出图像,粗糙集增强的图像在灰度层次和边缘清晰程度方面均优于常规的图像增强方法。徐宁和杜明辉将粗糙集图像增强用于 CT 图像处理中^[18]。

3.3 图像分割

图像分割是将图像表示为有具体物理意义的连通

区域集合。通过对图像的不同特征如边缘、纹理、颜色、灰度等的分析达到图像分割的目的,分割得到的图像区域在特征描述上具有相似性。

S. Shek 将图像看成知识系统^[19],基于不可分辨关系分割成不同区域。罗诗途等提出了一种基于粗糙集的图像分割智能决策方法^[20],选取若干具有代表性的分割算法构成算法库,用其对各种样本图像进行分割,然后利用从样本图像中提取出来的各种数值特征以及根据图像分割质量评价标准确定的各样本图像的最优分割算法来生成决策信息表,最后应用粗糙集理论对决策信息表进行离散化处理和属性约简,以生成图像分割算法选取的决策规则。Liu 等人提出粗糙集图像分割算法^[21],利用粗糙集相似性关系解决分割中聚类问题。

3.4 图像特征提取与分类

粗糙集方法用于模式识别主要是基于知识系统的表达和属性的简化方法,目的是求出每个模式的最小描述和对应的决策算法。陈敏针对遥感图像波段选择传统手工方法的不足,应用粗糙集理论,通过核属性的方法计算达到简化决策表中的条件属性的目的,得到图像分类的关键波段和一些次关键波段,实现了波段的自动选择^[22]。肖迪等提出根据肺癌细胞的每个特征属性对粗糙集下、上近似集的影响程度作为属性约简的依据^[23],根据约简的结果,再采用扩展的相近关系粗糙集对肺癌细胞进行识别诊断。利用下近似集中的结果进行判断可以提高识别的准确率,利用上近似集中的结果进行判断可以降低肺癌细胞识别的漏诊率。李孟歆等提出了基于粗糙集的神经网络分类器^[24],用粗糙集约简的样本集进行胶合板缺陷的神经网络分类研究。吴成东等采用粗糙集方法对墙地砖的原始特征进行优化^[25],去除冗余特征,提高了分类精度,在一定程度上提高了系统实时性。

粗糙集模式分类可以获得简洁的 IF-THEN 规则,这种规则接近人类语言,易于理解。武广富提出了一种基于粗糙集理论的火焰状态识别方法^[26]。该方法利用粗糙集理论的分类原则,得到火焰燃烧的高温区域,结合其它火焰图像的特征量,用基于粗糙集的知识决策系统建立火焰状态识别规则,构建全炉膛燃烧状态识别的基本模型,判断燃烧状态。该方法简化了计算量,提高反应速度。

粗糙集方法能有效地改进识别速度,但在容错性等方面需要做进一步的研究。另外,知识表达系统的顺序问题对结果也有一定的影响。

3.5 基于粗糙集的图像处理混合方法

粗糙集理论的进一步推广主要表现在和其他方法

的有效结合上,以弥补单一理论的局限性。

在图像预处理方面:王英健等提出了基于模糊粗糙集的视觉特征图像增强方法^[27],取得了较好的效果。张东波等提出了模糊粗神经网络的图像脉冲噪声滤波方法^[28],建立了一种兼有模糊化神经元和不可微粗神经元的模糊粗神经网络,并采用了融合爬山法的遗传算法,增强网络学习过程中的搜索能力。在图像分割方面:李孟歆和庞玉红提出了基于塔形小波和特征加权的墙地砖纹理图像分割方法^[29],改进原有的图像分割算法,对原始纹理图像进行高阶小波分解,在有效提取近似信息的基础上提高了系统的实时性。

在图像分割方面:张东波和王耀南结合 FCM 和粗糙集,提出了新的图像分割方法^[30],以不同聚类数情况下 FCM 的分割结果为依据构建属性值表,基于属性构成的不可分辨关系将图像分成多个小区域;然后,通过值约简获得各属性权值并以此为依据,计算各区域之间的差异度,进而通过差异度定义的等价关系,实现各区域相似度评价;最后,通过相似度定义的最终等价关系实现区域合并,完成图像分割。该方法在人工生成图像和大脑 CT 图像及 MRI 图像的分割中得到验证。吴冰等提出粗糙集神经网络用于图像分割的方法^[31],该方法充分发挥了粗糙集理论和神经网络的优势,组成网络具有学习速度快,容错能力强的特点,粗糙集神经网络在图像分割中的应用,明显优于神经网络。

在特征提取和分类方面:乔晓明等借助于粗糙集理论在处理大数据量、消除冗余信息等方面的优势,在进行属性约简的基础上,提出一种结合 FCM 的图像聚类方法^[32],该方法可以减少数据运算量,提高算法效率。蒋芸等将粗糙集理论中基于信息增益的约简方法和神经网络相结合,提出了粗糙神经网络算法 RNN^[33],将其应用于乳腺 X 光图像分类。该方法的分类精确度可达到 92.37%,所花费的时间也明显减少。Swiniarski 提出了基于小波变换和粗糙集的人脸识别算法^[34],识别精度能达到 99.24%,粗糙集方法起到了数据降维的重要作用。宋余庆等人提出基于主分量分析(PCA)和粗糙集的特征提取方法^[35],采用 PCA 对输入向量进行甄别,用粗糙集理论约简与分类无关或影响不大的向量。Swiniarski 和 Skowron 采用独立分量分析,主成分分析和粗糙集方法进行人脸识别^[36],粗糙集主要用于特征约简。

4 结束语

粗糙集理论处理不确定性信息的特点、相关信息提取能力,以及和其他智能方法的易融合性使得粗糙

集在图像处理领域具有良好的应用前景,不但可以用于构造新型的系统,而且它能够优化现有的许多算法。尽管如此,粗糙集在图像领域的应用仍处在探索阶段,新的方法、新的技术仍有待深入研究和探讨。大数据集、高效约简算法、并行计算以及混合智能算法研究等问题仍是粗糙集研究的主要方向。

参考文献:

- [1] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341 - 356.
- [2] 李孟歆, 吴成东. 粗糙集理论及应用[J]. 沈阳建筑工程学院学报, 2001, 17(3): 296 - 299.
- [3] Ziarko W. Variable precision rough set model[J]. Journal of Computer and System Science, 1993, 46(1): 39 - 59.
- [4] Kryszkiewicz M. Properties of incomplete information systems in the framework of rough sets [C]//In: Polkowski L, Skowron A. Rough Sets in Knowledge Discovery. Heidelberg: Physica, 1998: 422 - 450.
- [5] Stefanowski J, Tsoukias A. On the extension of rough sets under incomplete information [C]//In: Zhong N, Skowron A, Ohsuga S. Proceedings of the 7th International Workshop on New Directions in Rough Sets, Data Mining, and Granular - Soft Computing. Yamaguchi: Physica - Verlag, 1999: 73 - 81.
- [6] 王国胤. 粗糙集理论在不完备信息系统中的扩充[J]. 计算机研究与发展, 2002, 39(10): 1238 - 1243.
- [7] Dubois D, Prade H. Rough fuzzy sets and fuzzy rough sets[J]. International Journal of General System, 1990, 17: 191 - 208.
- [8] Greco S, Matarazzo B, Slowinski R. Rough sets theory for multicriteria decision analysis[J]. European Journal of Operational Research, 2001, 129(1): 41 - 47.
- [9] 普利荣, 刘思峰. 偏好多属性决策表概率决策的扩展粗糙集方法[J]. 南京航空航天大学学报, 1993, 46(1): 39 - 59.
- [10] 杨阳, 陈波, 许炳南. 基于粗糙集的图像中值滤波算法[J]. 哈尔滨工程大学学报, 2006, 27: 503 - 505.
- [11] 胡旺, 李志蜀, 黄奇. 基于粗糙集和多向差分图像的去噪方法[J]. 小型微型计算机系统, 2006, 27(12): 2341 - 2344.
- [12] 陈铁民, 王卫星. 粗糙集理论在阴影图像边缘检测中的应用[J]. 微计算机信息, 2007, 23: 297 - 298.
- [13] 胡小梅, 李众立. 一种基于粗糙集理论的彩色图像滤波算法[J]. 计算机与现代化, 2003, 9(Z1): 11 - 12.
- [14] 郭海涛, 田坦, 王连玉, 等. 利用粗糙集理论的声呐图像平滑[J]. 海洋技术, 2005, 24(2): 34 - 35.
- [15] 吕振肃, 魏弘博. 改进的基于粗糙集的图像平滑[J]. 中国图像图形学报, 2005, 10(7): 834 - 837.
- [16] 徐立中, 王慧斌, 杨锦堂. 基于粗糙集理论的图像增强方法[J]. 仪器仪表学报, 2000, 21(5): 514 - 515.
- [17] 叶灵伟, 张欣, 王勇. 基于粗糙集的图像增强方法[J]. 科技信息, 2007(17): 687 - 688.
- [18] 徐宁, 杜明辉. 基于边缘检测的CT图像粗糙集增强算法[J]. 医疗设备信息, 2003(9): 15 - 17.
- [19] Shek S, Lau Y. Image segmentation based on the indiscernibility relation [C]//RSKD - 93 Rough Sets and Knowledge Discovery. Banff: Springer - Verlag, 1993: 439 - 451.
- [20] 罗诗途, 张杞, 罗飞路, 等. 基于粗糙集理论的图像分割智能决策方法[J]. 中国图像图形学报, 2006, 11(1): 66 - 73.
- [21] Liu H J, Liu Y C. Method for image segmentation based on rough sets [J]. Journal of Infrared and Millimeter Waves, 2004, 23(6): 459 - 464.
- [22] 陈敏. 基于粗糙集核计算的遥感图像波段优选算法[J]. 宁德师专学报, 2006, 18(4): 378 - 380.
- [23] 肖迪, 张广明. 基于粗糙集理论的肺癌细胞图像识别[J]. 南京工业大学学报, 2007, 29(6): 87 - 90.
- [24] 李孟歆, 吴成东, 夏兴华. 基于粗糙集神经网络的自动视觉检测系统及其应用[J]. 计算机科学, 2003, 30(5): 201 - 202.
- [25] 吴成东, 周博, 李孟歆, 等. 基于粗糙集的神经网络缺陷分类方法的研究[J]. 沈阳建筑大学学报, 2007, 23(3): 525 - 528.
- [26] 武广富, 吕震中. 基于粗糙集理论的火焰图像处理与状态识别[J]. 热能动力工程, 2007, 22(3): 310 - 313.
- [27] 王英健, 史绍强, 唐贤英. 基于模糊粗糙集理论的视觉特性图像增强[J]. 长沙理工大学学报, 2004, 1(2): 79 - 82.
- [28] 张东波, 王耀南. 基于模糊神经网络的脉冲噪声滤除[J]. 计算机应用, 2005, 25(10): 2336 - 2338.
- [29] 李孟歆, 庞玉红. 基于塔形小波和特征加权的墙地砖纹理图像分割方法[J]. 沈阳建筑大学学报, 2008(3): 45 - 48.
- [30] 张东波, 王耀南. FCM聚类算法和粗糙集在医疗图像分割中的应用[J]. 仪器仪表学报, 2006, 27(12): 1683 - 1687.
- [31] 吴冰, 魏建, 刘艳昌, 等. 粗糙集神经网络在图像分割中的应用[J]. 河南理工大学学报, 2007, 26(5): 550 - 554.
- [32] 乔晓明, 刘有耀. 基于粗糙集理论和FCM的图像聚类方法[J]. 微计算机信息, 2007, 23: 283 - 285.
- [33] 蒋芸, 李战怀, 王勇, 等. 基于粗糙神经网络的医学图像分类新方法[J]. 计算机科学, 2006, 33(11): 151 - 153.
- [34] Swiniarski R. An Application of Rough Sets and Haar Wavelets to Face Recognition [C]//LNAI. Jin Yaochu, Wang Lipo, Hanoi, Vietnam: [s. n.], 2005: 561 - 568.
- [35] 宋余庆, 罗永刚, 孙志挥. 应用主分量分析与粗糙集处理的特征提取[J]. 计算机工程与应用, 2004, 40(22): 48 - 50.
- [36] Swiniarski R W, Skowron A. Independent component analysis, principal component analysis and rough sets in face recognition [M]//Transactions on Rough Sets I Lecture Notes In Computer Science. New York: Springer, 2004: 392 - 404.