基于包含度的变精度模糊粗糙集

徐伟华 张先韬 王巧荣
(重庆理工大学数学与统计学院 重庆 400054)

摘 要 利用模糊集的一个强包含度,在弱模糊划分的基础上建立了基于该包含度的变精度模糊粗糙集模型,对其重要性质进行了深入研究,并给出了对应形式粗糙度的计算方法,进一步利用海明距离和欧几里得距离定义了该模型下模糊粗糙集的两个粗糙性度量。给出的变精度模糊粗糙集模型能够使模糊粗糙集的运算按照模糊集的运算实现,为变精度模糊粗糙集理论的研究和应用奠定了一定的理论基础。

关键词 包含度,模糊集,粗糙集,变精度模糊粗糙集,粗糙性度量
中图法分类号 TP18 **文献标识码** A

Variable Precision Fuzzy Rough Set Based on Inclusion Degree

XU Wei-hua ZHANG Xian-tao WANG Qiao-rong
(School of Mathematics and Statistics, Chongqing University of Technology, Chongqing 400054, China)

Abstract We studied a power inclusion degree on fuzzy sets and established the fuzzy rough set model based on it with the foundation of weak fuzzy partition. Some important properties of this model were approached and the method to compute the roughness correspondingly was illustrated. Furthermore, two roughness measures of fuzzy rough set in this model were defined according to the Hamming-distance and Euclid-distance. The fuzzy rough set model we proposed can carry the computing on fuzzy rough set by fuzzy sets operating properties. And this model can lay the theoretical foundation for research and applications of variable precision fuzzy rough set.

Keywords Inclusion degree, Fuzzy set, Rough set, Variable precision fuzzy rough set, Roughness measure

模糊集理论和粗糙集理论都是处理不确定性问题的数学理论和工具。模糊集理论是美国计算机与控制论专家 Zadeh L. A. 教授于1965年提出的一种研究模糊性或不确定性问题的理论[1],迄今已成为一个较为完善的学术分支。模糊集理论是对 Cantor 集合理论的有益的推广,在很多领域中获得了卓有成效的应用,取得了显著的成果和经济效益。粗糙集理论是波兰数学家 Pawlak Z. 于1982年提出的一种数据分析理论[2],是一种处理不精确、不确定、不协调知识和信息的软计算工具,该理论在数据的决策与分析、模式识别、机器学习与知识发现等方面已被成功应用,是信息科学理论最为活跃的研究领域之一。随着更多学者的研究,该理论已经渗透到其他学科中并取得了很多重要成果。

Pawlak 提出的粗糙集模型是按照等价关系对论域产生划分的经典粗糙集模型,所处理的对论域进行的分类是完全肯定的,只考虑包含或属于的问题,不存在包含或属于的程度问题。经典粗糙集模型处理的训练集对象是已知的,其理论仅适用于特定的对象集。在很多实际问题中,论域上的一般二元关系不是等价关系,这时 Pawlak 经典粗糙集模型受到限制。由此,必须将 Pawlak 经典粗糙集模型进行推广,研究者们提出了不同的粗糙集模型并进行了研究和拓展[3]。Dubois 和 Prade 提出了模糊粗糙集模型[4];Kuncheva 提出了基于弱模糊划分的模糊粗糙集模型[5];Ziarko 提出了变精度粗糙集模型[6];Bodjanova 在 Dubois 基础上提出了基于包含度的修正型模糊粗糙集模型[7];Zhang 等人在文献[8,9]中继续研究阐述了模糊粗糙集的理论;研究者们也对概念的粗糙性度量进行了推广研究[10-13];文献[15]中还介绍了以弱模糊划分为基础基于包含度的模糊粗糙集模型。

文献[7-9,15-18]利用包含度研究了模糊粗糙集,其弱点是当被近似的概念是模糊集时,包含度计算不能以隶属度的形式直接进行,这就使得模糊粗糙集在应用中可能产生更多的不精确性,得到的近似集不能很好地刻画和反映原有系统的不确定信息。文献[18]在模糊集的理论研究中提出了由隶属度定义的包含度,但没有将该包含度与粗糙集结合进行相关研究。

本文在上述研究的基础上,利用模糊集上一个由隶属度定义的强包含度,通过引入精度系数 k,在弱模糊划分基础上建立基于包含度的变精度模糊粗糙集模型,并对其重要性质作了讨论,进而给出该模型在精度系数 k 下的粗糙度的计算方法并加以研究,利用模糊集的距离定义两个新的粗糙性度量。本文所利用的包含度公式和粗糙性度量方法能够使模糊粗糙集的有关运算按照模糊集的运算实现,这将对该模型变精度模糊粗糙集的研究和应用提供一定的帮助。

到稿日期:2010-10-31 返修日期:2011-02-25 本文受国家自然科学基金(71071124,11001227)资助。
徐伟华(1979－),男,副教授,硕士生导师,主要研究方向为模糊集、粗糙集、人工智能的数学基础,E-mail:chxuwh@gmail.com;张先韬(1986－),男,硕士生,主要研究方向为粗糙集、人工智能的数学基础;王巧荣(1986－),女,硕士生,主要研究方向为粗糙集、人工智能的数学基础。

1 预备知识

为方便本文的论述和有关性质的研究，本节首先给出需要用到的一些基本概念和性质。

定义 1[15] 设 U 为论域，在 U 上给定一个映射

$$\widetilde{A}: U \to [0,1]$$
$$x \mapsto \widetilde{A}(x)$$

则称 $\widetilde{A}$ 为 U 上的一个模糊集，$\widetilde{A}(x)$ 称为模糊集 $\widetilde{A}$ 的隶属函数（或称每个 $\widetilde{A}(x)$ 的值为对象 x 对 $\widetilde{A}$ 的隶属度）。U 上模糊集全体记为 $\mathscr{F}(U)$，给定论域 U 上的一个模糊集时，必须给出每个对象的隶属度（或隶属函数）。

定义 2[14] 设 U 为论域，$\widetilde{A},\widetilde{B}\in\mathscr{F}(U)$，则模糊集的并、交、补运算和包含关系定义如下

(1) $(\widetilde{A}\cup\widetilde{B})(x)=\widetilde{A}(x)\vee\widetilde{B}(x)$

(2) $(\widetilde{A}\cap\widetilde{B})(x)=\widetilde{A}(x)\wedge\widetilde{B}(x)$

(3) $\sim\widetilde{A}:(\sim A)(x)=1-\widetilde{A}(x)$

(4) $\widetilde{A}\subseteq\widetilde{B}\Leftrightarrow\widetilde{A}(x)\leqslant\widetilde{B}(x),\forall x\in U$

其中，$\vee$ 为实数的取大运算，$\wedge$ 为实数的取小运算，$\sim$ 为补运算。

由定义可见，模糊集的运算性质除互补律外，其他运算性质与 Cantor 集相同，详见文献[14]。

定义 3[15] 设 U 为论域，对于任意 $\widetilde{A},\widetilde{B}\in\mathscr{F}(U)$ 有实数 $D(\widetilde{B}/\widetilde{A})$ 对应，且满足

(1) $0\leqslant D(\widetilde{B}/\widetilde{A})\leqslant 1$；

(2) $\forall\widetilde{A},\widetilde{B}\in\mathscr{F}(U),\widetilde{A}\subseteq\widetilde{B}\Rightarrow D(\widetilde{B}/\widetilde{A})=1$；

(3) $\forall\widetilde{A},\widetilde{B},\widetilde{C}\in\mathscr{F}(U),\widetilde{A}\subseteq\widetilde{B}\subseteq\widetilde{C}\Rightarrow D(\widetilde{A}/\widetilde{C})\leqslant D(\widetilde{A}/\widetilde{B})$；

则称 D 为 $\mathscr{F}(U)$ 上的包含度；

若还有

(4) $\forall\widetilde{A},\widetilde{B},\widetilde{C}\in\mathscr{F}(U),\widetilde{A}\subseteq\widetilde{B}\Rightarrow D(\widetilde{A}/\widetilde{C})\leqslant D(\widetilde{B}/\widetilde{C})$

则称 D 为 $\mathscr{F}(U)$ 上的强包含度。

定义 4[15] 设 U 为有限非空论域。R 为 U 上的任意等价关系，称 (U,R) 为 Pawlak 近似空间，对于对象 x 称 $[x]_R=\{y|xRy\}$ 为 x 在关系 R 下的等价类。对于任意 $X\subseteq U$，X 关于 Pawlak 近似空间 (U,R) 的下近似 $\underline{R}(X)$ 和上近似 $\overline{R}(X)$ 分别定义为

$$\underline{R}(X)=\{x\in U|[x]_R\subseteq X\}$$
$$\overline{R}(X)=\{x\in U|[x]_R\cap X\neq\varnothing\}$$

当 $\underline{R}(X)=\overline{R}(X)$ 时，称 X 为 R 精确集；否则，称 X 为 R 粗糙集。

定义 5[15] 设 U 为有限非空论域，$X\subseteq U$，X 关于 Pawlak 近似空间 (U,R) 的下近似和上近似分别为 $\underline{R}(X)$ 和 $\overline{R}(X)$，称

$$\rho_R(X)=1-\frac{|\underline{R}(X)|}{|\overline{R}(X)|}$$

为 X 关于 Pawlak 近似空间 (U,R) 的粗糙度。

粗糙集理论的更多知识可参阅文献[15,16]，不再赘述。

2 $\mathscr{F}(U)$ 上的一个强包含度

上一节给出了有关的概念和性质，本节叙述文献[18]中介绍的 $\mathscr{F}(U)$ 上一个由隶属度定义的强包含度。

依定义 3，满足定义并且有实际意义的包含度不唯一，文献[15]中提到的 $\mathscr{F}(U)$ 上的一个包含度为

$$D(\widetilde{B}/\widetilde{A})=\frac{|n(\widetilde{B},\widetilde{A})|}{|\mathrm{supp}\widetilde{A}|}$$

式中，$n(\widetilde{B},\widetilde{A})=\{x\in\mathrm{supp}\widetilde{A}\,|\,\widetilde{B}(x)\leqslant\widetilde{A}(x)\}$，$|n(\widetilde{B},\widetilde{A})|$ 和 $|\mathrm{supp}\widetilde{A}|$ 为对应 Cantor 集合的基。这个包含度采用转化为 Cantor 集合基的形式来定义。为了将模糊粗糙集运算与模糊集的运算统一起来，下面按照我们的研究需要重新表述文献[18]中介绍的 $\mathscr{F}(U)$ 上由隶属度定义的一个强包含度。

定义 6 设 U 为有限非空论域，对于任意 $\widetilde{A},\widetilde{B}\in\mathscr{F}(U)$，记

$$D^*(\widetilde{B}/\widetilde{A})=\frac{\sum_{x\in U}\widetilde{B}(x)\wedge\widetilde{A}(x)}{\sum_{x\in U}\widetilde{A}(x)}$$

约定当 $\widetilde{A}=\varnothing$ 时，$D*(\widetilde{B}/\widetilde{A})=1$，则 $D*$ 为 $\mathscr{F}(U)$ 上的强包含度。

证明：由定义 3，验证定义即证。

设 U 为有限非空论域，若任意 $A,B\in P(U)$，即 A,B 都为 Cantor 集合时，则有

$$\frac{\sum_{x\in U}\widetilde{B}(x)\wedge\widetilde{A}(x)}{\sum_{x\in U}\widetilde{A}(x)}=\frac{|n(B,A)|}{|\mathrm{supp}A|}=\frac{|B\cap A|}{|A|}$$

式中，$\frac{|B\cap A|}{|A|}$ 为 $P(U)$ 上的强包含度。通过上式可以发现，包含度 D^* 可退化为 Cantor 集合的包含度，而且是文献[15]中定义的更广泛的形式，实现了与模糊集运算的统一。

例 1 设论域 $U=\{x_1,x_2,\cdots,x_8\}$。$\mathscr{A}=\{\widetilde{A}_1,\widetilde{A}_2,\cdots,\widetilde{A}_6\}$，$\widetilde{A}_i\in\mathscr{F}(U)$，$(i=1,\cdots,6)$ 为 U 上的一个模糊子集类，$\widetilde{X}_j\in\mathscr{F}(U)$，$(j=1,2,3,4)$。各隶属度如表 1 所列。

表 1

U	$\widetilde{A}_1$	$\widetilde{A}_2$	$\widetilde{A}_3$	$\widetilde{A}_4$	$\widetilde{A}_5$	$\widetilde{A}_6$	$\widetilde{X}_1$	$\widetilde{X}_2$	$\widetilde{X}_3$	$\widetilde{X}_4$
x_1	0	0.8	0.3	0.7	0.2	0	0.5	0.7	0	0.7
x_2	1	0	0	0	0.1	0.8	0	0	1	1
x_3	0.6	0.3	0	0	1	0	0	0.5	0.7	1
x_4	0	0.5	0.6	0	0	1	0.6	0.5	0	1
x_5	0	1	0	0	0.7	0.5	0	1	0	0.7
x_6	0	0.1	0.8	1	0	0	0.8	0.4	0	1
x_7	0	0	1	0.4	0.6	0	1	0	0	0.6
x_8	1	0	0	0.5	0.5	0	0	0	1	1

计算 $D^*(\widetilde{X}_j/\widetilde{A}_i)$ 和 $D(\widetilde{X}_j/\widetilde{A}_i)$ 并分别用 D^*_{ij} 和 D_{ij} 表示，如表 2 所列。

表 2

$\mathscr{A}$	D^*_{i1}	D_{i1}	D^*_{i2}	D_{i2}	D^*_{i3}	D_{i3}	D^*_{i4}	D_{i4}
$\widetilde{A}_1$	0	0	0.17	0	1	1	1	1
$\widetilde{A}_2$	0.4	0.4	0.96	0.8	0.1	0.2	0.85	0.6
$\widetilde{A}_3$	1	1	0.44	0.25	0	0	1	1
$\widetilde{A}_4$	0.65	0.25	0.42	0.25	0.19	0.25	1	1
$\widetilde{A}_5$	0.26	0.33	0.39	0.33	0.36	0.33	1	1
$\widetilde{A}_6$	0.35	0	0.44	0.33	0.35	0.33	1	1

从表 2 中数据 D^*_{32}，D^*_{42}，D_{32}，D_{42}，D^*_{52} D^*_{62}，D_{52}，D_{62}，D^*_{53}，D^*_{63}，D_{53}，D_{63} 可以看出，当文献[15]中的包含度不能区分一个集合对另外两个的包含情况时，D^* 可以对其确定。在表 2 中将 D^*_{ij} 和 D_{ij} 分别作比较，可以看到本文提出的包含度 D^* 能够反映模糊集的包含情况。在使用时可以将这两个包含度结合起来，更好地考查模糊集的包含情况。

3 基于包含度的变精度模糊粗糙集

本节在上节的基础上讨论基于包含度 D^* 的变精度模糊粗糙集模型，这个模型以论域上的模糊关系产生的弱模糊划分为基础，基于模糊集的包含度 D^*，通过引入一个精度系数 $k\in(0.5,1]$ 定义论域上模糊概念的下近似和上近似算子来刻画模糊粗糙集。

定义 7[15] 设 U 为有限非空论域，$\mathscr{A}=\{\tilde{A}_1,\tilde{A}_2,\cdots,\tilde{A}_m\}$ 为 U 的一个模糊子集类，其中，$\tilde{A}_i\in\mathscr{F}(U)$，$i=1,\cdots,m$。若 $\bigcup_{i=1}^{m}\mathrm{supp}\tilde{A}_i=U$，则称 $\mathscr{A}$ 为 U 的一个弱模糊划分。

定义 8 设 U 为有限非空论域，$\mathscr{A}=\{\tilde{A}_1,\tilde{A}_2,\cdots,\tilde{A}_m\}$ 为 U 的一个弱模糊划分，D^* 为 $\mathscr{F}(U)$ 上的强包含度，则称 $(U,\mathscr{A},D^*)$ 为模糊强包含近似空间。

定义 9 设 U 为有限非空论域，对任意 $\tilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，$\tilde{X}$ 关于 $(U,\mathscr{A},D^*)$ 的 k-下近似 $\underline{R}_k(\tilde{X})$ 和 k-上近似 $\overline{R}_k(\tilde{X})$ 分别定义为

$$\underline{R}_k(\tilde{X})=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)\geqslant k\}$$

$$\overline{R}_k(\tilde{X})=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)>1-k\}$$

当 $\underline{R}_k(\tilde{X})=\overline{R}_k(\tilde{X})$ 时，称 $\tilde{X}$ 关于 $(U,\mathscr{A},D^*)$ 为 k-模糊粗糙可定义集；否则，称 $\tilde{X}$ 关于 $(U,\mathscr{A},D^*)$ 为 k-模糊粗糙集。当 $\overline{R}_k(\tilde{X})$ 和 $\underline{R}_k(\tilde{X})$ 越贴近时，称 $\tilde{X}$ 越精细；反之，称 $\tilde{X}$ 越粗糙。

定理 1 设 U 为有限非空论域，对任意 $\tilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，$\tilde{X}$ 关于 $(U,\mathscr{A},D^*)$ 的 k-下近似 $\underline{R}_k(\tilde{X})$ 和 k-上近似 $\overline{R}_k(\tilde{X})$ 满足下列性质。

(1) $\underline{R}_k(\emptyset)=\emptyset=\overline{R}_k(\emptyset)$；$\underline{R}_k(U)=\overline{R}_k(U)$。

(2) $\underline{R}_k(\tilde{X})\subseteq\overline{R}_k(\tilde{X})$。

(3) $k_1\leqslant k_2\Rightarrow\underline{R}_{k_1}(\tilde{X})\supseteq\underline{R}_{k_2}(\tilde{X})$；$\overline{R}_{k_1}(\tilde{X})\subseteq\overline{R}_{k_2}(\tilde{X})$。

(4) $\tilde{X}\subseteq\tilde{Y}\Rightarrow\underline{R}_k(\tilde{X})\subseteq\underline{R}_k(\tilde{Y})$；$\overline{R}_k(\tilde{X})\subseteq\overline{R}_k(\tilde{Y})$。

(5) $\underline{R}_k(\tilde{X}\cap\tilde{Y})\subseteq\underline{R}_k(\tilde{X})\cap\underline{R}_k(\tilde{Y})$；$\overline{R}_k(\tilde{X}\cup\tilde{Y})\supseteq\overline{R}_k(\tilde{X})\cup\overline{R}_k(\tilde{Y})$。

(6) $\underline{R}_k(\tilde{X}\cup\tilde{Y})\supseteq\underline{R}_k(\tilde{X})\cup\underline{R}_k(\tilde{Y})$；$\overline{R}_k(\tilde{X}\cap\tilde{Y})\subseteq\overline{R}_k(\tilde{X})\cap\overline{R}_k(\tilde{Y})$。

证明：(1)由定义 6 有

$$D^*(\emptyset/\tilde{A}_i)=0,D^*(U/\tilde{A}_i)=1。$$

再由定义 9，$\forall k\in(0.5,1]$有

$$\underline{R}_k(\emptyset)=\bigcup\{\tilde{A}_i\mid D^*=0\geqslant k\}=\emptyset$$

$$\overline{R}_k(\emptyset)=\bigcup\{\tilde{A}_i\mid D^*=0>1-k\}=\emptyset\Rightarrow\underline{R}_k(\emptyset)=\emptyset=\overline{R}_k(\emptyset)$$

$$\underline{R}_k(U)=\bigcup\{\tilde{A}_i\mid D^*=1\geqslant k\}=\bigcup_{i=1}^{m}\tilde{A}_i$$

$$\overline{R}_k(U)=\bigcup\{\tilde{A}_i\mid D^*=1>1-k\}=\bigcup_{i=1}^{m}\tilde{A}_i\Rightarrow\underline{R}_k(U)=\overline{R}_k(U)$$

(2)(3) 由定义 9 可证。

(4) $\tilde{X}\subseteq\tilde{Y}\Rightarrow D^*(\tilde{Y}/\tilde{A}_i)\geqslant D^*(\tilde{X}/\tilde{A}_i)$

$\forall\tilde{A}_i\subseteq\underline{R}_k(\tilde{X})\Rightarrow D^*(\tilde{Y}/\tilde{A}_i)\geqslant D^*(\tilde{X}/\tilde{A}_i)\geqslant k$

$\Rightarrow\tilde{A}_i\subseteq\underline{R}_k(\tilde{Y})$

$\Rightarrow\underline{R}_k(\tilde{X})\subseteq\underline{R}_k(\tilde{Y})$

同理可证 $\overline{R}_k(\tilde{X})\subseteq\overline{R}_k(\tilde{Y})$。

(5)(6) 由性质(4)可以证得。

例 2(续例 1) 由定义 7，对 A 有 $\bigcup_{i=1}^{6}\mathrm{supp}\,A_i=U$，故 $\mathscr{A}$ 为论域 U 的一个弱模糊划分。取 $k_1=0.6$，$k_2=0.7$，$k_3=0.9$，分别计算 $\underline{R}_k(\tilde{X}_j)$ 和 $\overline{R}_k(\tilde{X}_j)$，$j=1,2,3,4$ 得到表 3（为简便起见，记 $\tilde{A}_{rs}=\tilde{A}_r\cup\tilde{A}_s$）。

表 3

	k_1	k_2	k_3
$\underline{R}_k(\tilde{X}_1)$	$\tilde{A}_{34}$	$\tilde{A}_3$	$\tilde{A}_3$
$\overline{R}_k(\tilde{X}_1)$	$\tilde{A}_{34}$	$\tilde{A}_{2346}$	$\tilde{A}_{23456}$
$\tilde{R}_k(\tilde{X}_2)$	$\tilde{A}_2$	$\tilde{A}_2$	$\tilde{A}_2$
$\overline{R}_k(\tilde{X}_2)$	$\tilde{A}_{2346}$	$\tilde{A}_{23456}$	$\tilde{A}_{23456}$
$\underline{R}_k(\tilde{X}_3)$	$\tilde{A}_1$	$\tilde{A}_1$	$\tilde{A}_1$
$\overline{R}_k(\tilde{X}_3)$	$\tilde{A}_1$	$\tilde{A}_{156}$	$\tilde{A}_{156}$
$\underline{R}_k(\tilde{X}_4)$	$\tilde{A}_{123456}$	$\tilde{A}_{123456}$	$\tilde{A}_{13456}$
$\overline{R}_k(\tilde{X}_4)$	$\tilde{A}_{123456}$	$\tilde{A}_{123456}$	$\tilde{A}_{123456}$

由表 3 可以验证、分析定理 1 的性质。

对比 Pawlak 粗糙集模型，$\underline{R}_k(U)=\overline{R}_k(U)=U$ 在此模型中不一定成立。因为 $\mathscr{A}$ 为 U 的弱模糊划分，被近似的概念是 U 时，可以得到 $\underline{R}_k(U)=\bigcup_{i=1}^{m}\tilde{A}_i=\overline{R}_k(U)$ 成立，根据模糊集的并运算可以知道，$\bigcup_{i=1}^{m}\tilde{A}_i$ 并不一定等于论域 U。

从定义 9 和例 2 可以看到，精度系数 k 越小，$\underline{R}_k(\tilde{X})$ 和 $\overline{R}_k(\tilde{X})$ 越贴近，$\tilde{X}$ 则相对越精确。

Pawlak 经典粗糙集模型定义了正域、负域和边界来理解粗糙集，下面定义给出基于包含度 D^* 的模糊粗糙集模型的 k-正域、k-负域和 k-边界的定义。

定义 10 设 U 为有限非空论域，对任意 $\tilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，$\tilde{X}$ 关于模糊包含近似空间 $(U,\mathscr{A},D^*)$ 的 k-正域 $pos(\tilde{X},k)$、k-负域 $neg(\tilde{X},k)$ 和 k-边界 $bn(\tilde{X},k)$ 分别为

$$pos(\tilde{X},k)=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)\geqslant k\}$$

$$neg(\tilde{X},k)=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)\leqslant 1-k\}$$

$$bn(\tilde{X},k)=\bigcup\{\tilde{A}_i\mid 1-k<D^*(\tilde{X}/\tilde{A}_i)<k\}$$

粗糙集理论中概念不确定性的一个原因来自给定近似空间的粗糙集的边界，当边界为空集时没有不确定性，边界越大就越粗糙或越模糊，下面定理反映了基于包含度 D^* 的变精度模糊粗糙集模型在边界不确定时的理解。

定理 2 设 U 为有限非空论域，$\tilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，下述结论成立。

(1)$\tilde{X}$ 是 k-模糊粗糙可定义集当且仅当 $bn(\tilde{X},k)=\emptyset$；

(2)$\tilde{X}$ 是 k-模糊粗糙集当且仅当 $bn(\tilde{X},k)\neq\emptyset$。

证明：(1) $\tilde{X}$ 是 k-模糊粗糙可定义集

$\Leftrightarrow\underline{R}_k(\tilde{X})=\overline{R}_k(\tilde{X})\Leftrightarrow bn(\tilde{X},k)=\emptyset$；

(2) $\tilde{X}$ 是 k-模糊粗糙集

$\Leftrightarrow\underline{R}_k(\tilde{X})\neq\overline{R}_k(\tilde{X})\Leftrightarrow bn(\tilde{X},k)\neq\emptyset$。

对于不同的精度系数 k，下面性质成立。

定理 3 设 U 为有限非空论域，对任意 $\tilde{X}\in\mathscr{F}(U)$，$k,k_1,k_2\in(0.5,1]$，则有

(1)若 $k_1<k$，则当 $\tilde{X}$ 是 k-模糊粗糙可定义集时也是 k_1-模糊粗糙可定义集；

(2)若 $k_2>k$，则当 $\tilde{X}$ 是 k-模糊粗糙集时也是 k_2-模糊粗糙集。

证明：由定义 9 和定理 1 性质(1)(3)易证。

定理 3 的结论在例 2 中是显然的。

以弱模糊划分为基础，基于包含度 D^* 的变精度粗糙集模型是 Pawlak 经典粗糙集的推广。事实上，当 $\tilde{A}_i\in\mathscr{A}$ 为精典集且 $\mathscr{A}=\{\tilde{A}_1,\tilde{A}_2,\cdots,\tilde{A}_m\}$ 构成论域的划分时，$\tilde{A}_i=[x]_R$ 就是一个 R 等价类，对于任意的 $\tilde{X}\in\mathscr{F}(U)$，在本文提出的包含度 D^* 下，取 $k=1$，则

$$\begin{aligned}\underline{R}_k(\tilde{X})&=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)\geqslant 1\}\\&=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)=1\}\\&=\bigcup\{\tilde{A}_i\mid\tilde{A}_i\subseteq\tilde{X}\}\\&=\bigcup\{[x_i]_R\mid[x_i]_R\subseteq\tilde{X}\}\\&=\{x\in U\mid[x]_R\subseteq\tilde{X}\}\end{aligned}$$

$$\begin{aligned}\overline{R}_k(\tilde{X})&=\bigcup\{\tilde{A}_i\mid D^*(\tilde{X}/\tilde{A}_i)>0\}\\&=\bigcup\{\tilde{A}_i\mid\tilde{A}_i\cap\tilde{X}\neq\emptyset\}\\&=\bigcup\{[x_i]_R\mid[x_i]_R\cap\tilde{X}\neq\emptyset\}\\&=\{x\in U\mid[x]_R\cap\tilde{X}\neq\emptyset\}\end{aligned}$$

若 $\widetilde{X}\in p(U)$，即 $\widetilde{X}$ 为 Cantor 集合时，k-下近似和 k-上近似完全退化为 Pawlak 意义下的下近似和上近似。

4 k-模糊粗糙度

前面讨论了在弱模糊划分基础上基于包含度 D^* 的变精度粗糙集模型，本节首先给出该粗糙集模型从经典粗糙集模型中推广形式的粗糙度的定义。

对于任意的 $\widetilde{X}\in\mathscr{F}(U)$，$x\in U$，记

$$\alpha_k(\widetilde{X})=\frac{\sum_{x\in U}\underline{R}_k(\widetilde{X})(x)}{\sum_{x\in U}\overline{R}_k(\widetilde{X})(x)}$$

显然 $\alpha_k(\widetilde{X})\in[0,1]$。

对于 k-模糊粗糙集 $\widetilde{X}$，$\widetilde{X}$ 越精细，那么 $\underline{R}_k(\widetilde{X})$ 和 $\overline{R}_k(\widetilde{X})$ 就越贴近，$\alpha_k(\widetilde{X})$ 也就越靠近 1，当且仅当 $\underline{R}_k(\widetilde{X})=\overline{R}_k(\widetilde{X})$ 时为 1；反之，$\widetilde{X}$ 越粗糙 $\alpha_k(\widetilde{X})$ 越靠近 0，当且仅当 $\underline{R}_k(\widetilde{X})=\emptyset$ 时为 0；当 $\widetilde{X}=\emptyset$ 时，$\underline{R}_k(\widetilde{X})=\overline{R}_k(\widetilde{X})=\emptyset$ 为精确集，定义 $\alpha_k(\emptyset)=1$。经分析可知 $\alpha_k(\widetilde{X})$ 反映了 $\widetilde{X}$ 的精细程度。

由上分析，所研究的 k-模糊粗糙集的粗糙度定义如下。

定义 11 设 U 为有限非空论域，$\widetilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，$\widetilde{X}$ 基于包含度的 k-下近似和 k-上近似分别为 $\underline{R}_k(\widetilde{X})$ 和 $\overline{R}_k(\widetilde{X})$，记

$$\rho_k(\widetilde{X})=1-\alpha_k(\widetilde{X})=1-\frac{\sum_{x\in U}\underline{R}_k(\widetilde{X})(x)}{\sum_{x\in U}\overline{R}_k(\widetilde{X})(x)}$$

约定 $\rho_k(\emptyset)=0$，则称 $\rho_k(\widetilde{X})$ 为 $\widetilde{X}$ 关于模糊包含近似空间 $(U,\mathscr{A},D^*)$ 的 k-模糊粗糙度。

$\rho_k(\widetilde{X})$ 取值的大小能够反映 k-模糊粗糙集的粗糙程度，我们有下面定理成立。

定理 4 设 U 为有限非空论域，对任意 $\widetilde{X}_1,\widetilde{X}_2\in\mathscr{F}(U)$，$k\in(0.5,1]$。$\widetilde{X}_1$ 比 $\widetilde{X}_2$ 粗糙的充要条件是 $\rho_k(\widetilde{X}_1)\geqslant\rho_k(\widetilde{X}_2)$。

证明：由 $\alpha_k(\widetilde{X})$ 的分析和定义 11 即证。

对于不同的精度系数 $k\in(0.5,1]$，所表示的 k-模糊粗糙集 $\widetilde{X}$ 的粗糙程度不同，有如下定理。

定理 5 设 U 为有限非空论域，$\widetilde{X}\in\mathscr{F}(U)$，$k_1,k_2\in(0.5,1]$，若 $k_1\geqslant k_2$，则 $\rho_{k_1}(\widetilde{X})\geqslant\rho_{k_2}(\widetilde{X})$。

证明：由定理 1 性质(3)和定义 11 易证。

例 3(续例 2) 计算 $\rho_{k_l}(\widetilde{X}_j)$，$j=1,2,3,4$，$k_1=0.6$，$k_2=0.7$，$k_3=0.9$，并作比较，验证定理 4 和定理 5。

按照定义 11，计算 $\rho_{k_l}(\widetilde{X}_j)$ 如表 4 所列。

表 4

	k_1	k_2	k_3
$\rho_k(\widetilde{X}_1)$	0	0.58	0.62
$\rho_k(\widetilde{X}_2)$	0.58	0.62	0.65
$\rho_k(\widetilde{X}_3)$	0	0.53	0.53
$\rho_k(\widetilde{X}_4)$	0	0	0.04

从表 4 可以知道，对于不同的精度系数 k 都有 $\rho_k(\widetilde{X}_2)\geqslant\rho_k(\widetilde{X}_1)$ 成立。由于 k-下近似集和 k-上近似集都是 k-模糊粗糙可定义集，是精确的模糊集，为了比较不同的 X 的 k-下近似集和 k-上近似集的靠近程度，我们使用模糊集的贴近度来比较，这里采用公式

$$N(\widetilde{A},\widetilde{B})=1-[\sum_{i=1}^{n}|\widetilde{A}(x_i)-\widetilde{B}(x_i)|]/n$$

进行计算。取 $k_2=0.7$，有

$$N_1=N(\underline{R}_{0.7}(\widetilde{X}_1),\overline{R}_{0.7}(\widetilde{X}_1))=0.54$$

$$N_2=N(\underline{R}_{0.7}(\widetilde{X}_2),\overline{R}_{0.7}(\widetilde{X}_2))=0.45$$

由 $N_1>N_2$ 知，$k_2=0.7$ 时，$\widetilde{X}_1$ 的近似集的贴近度比 $\widetilde{X}_2$ 的近似集的贴近度大，即 $\widetilde{X}_1$ 比 $\widetilde{X}_2$ 精细，从而在精度系数 $k_2=0.7$ 时，$\widetilde{X}_2$ 比 $\widetilde{X}_1$ 粗糙。对于不同的精度系数 k 都有 $\widetilde{X}_2$ 比 $\widetilde{X}_1$ 粗糙。故 $\widetilde{X}_2$ 是比 $\widetilde{X}_1$ 粗糙的。$\rho_k(\widetilde{X})$ 取值的大小就反映了 $\widetilde{X}$ 的粗糙程度。

从表 4，对不同精度系数 k 有如下式子

$$\rho_{0.6}(\widetilde{X}_j)\leqslant\rho_{0.7}(\widetilde{X}_j)\leqslant\rho_{0.9}(\widetilde{X}_j),\ j=1,2,3,4$$

精度系数 k 越接近 0.5，$\widetilde{X}$ 越精细，$\rho_k(\widetilde{X})$ 值越小，这与表 3 中 $\widetilde{X}_j$ 在不同精度系数下的近似集的关系体现的 $\widetilde{X}_j$ 的粗糙程度是一致的。

由近似集定义知，下近似和上近似都是模糊集，如定义中所述，下近似和上近似越接近，那么精细程度越高，对应的粗糙度就越低，如果能够合理地度量这两个模糊集之间的距离，即可以用来度量粗糙集的粗糙性。根据模糊集的距离的意义，结合上面例题对定理 4 的验证，可以从模糊集的距离的角度考察上下近似，从而度量集合的粗糙性。

定义 12 设 $U=\{x_1,x_2,\cdots,x_n\}$ 为论域，对任意 $\widetilde{X}\in\mathscr{F}(U)$，$k\in(0.5,1]$，$\widetilde{X}$ 的 k-下近似和 k-上近似分别为 $\underline{R}_k(\widetilde{X})$ 和 $\overline{R}_k(\widetilde{X})$，则可以通过模糊集的距离来度量 $\widetilde{X}$ 的粗糙性。

根据海明距离我们定义一种粗糙性度量为

$$\rho_k^H(\widetilde{X})=[\sum_{i=1}^{n}(\overline{R}(\widetilde{X})(x_i)-\underline{R}(\widetilde{X})(x_i))]/n$$

根据欧几里得距离我们定义一种粗糙性度量为

$$\rho_k^E(\widetilde{X})=\sqrt{[\sum_{i=1}^{n}(\overline{R}(\widetilde{X})(x_i)-\underline{R}(\widetilde{X})(x_i))^2]/n}$$

由定义 11 和定义 12 知，在精度系数 k 下，当 $\underline{R}_k(\widetilde{X})=\emptyset$ 时，$\rho_k(\widetilde{X})=1$，从这个度量来看，此时 $\widetilde{X}$ 是最粗糙的；但事实上从边界域的角度来考虑，最粗糙的情形是边界域最大时，也就是上近似和下近似之差最大时，$\widetilde{X}$ 是最粗糙的，当 $\underline{R}_k(\widetilde{X})=\emptyset$ 时，$\overline{R}_k(\widetilde{X})=U$ 不一定成立，从边界域的意义上，此时不一定是最粗糙的，在这种情况下 $\rho_k(\widetilde{X})$ 对于模糊集 $\widetilde{X}$ 的粗糙性的度量受到限制。

在利用粗糙集方法解决实际应用问题时，对于训练集的选取，可以参考粗糙性度量，选取保证所有决策类的粗糙性最小的一个训练集来进行问题的处理。而且实际问题中，在所有可能的训练集中，模糊决策类的下近似总是等于空集的情形是可能存在的，在这种情形下，可能决策对我们还是有应用价值的，这个时候选择保证决策类粗糙性最小的训练集是有最大价值的，此时可以通过定义 12 中的粗糙性度量来进行选择。

定理 4 和定理 5 的结论对于定义 12 中的粗糙性度量成立，通过上面例题经计算可以验证，这里不再具体验证。在使用时对于粗糙性度量的选取，需要根据具体的情形选择，当单纯的一种粗糙性度量失效时，需要将定义 11 和定义 12 的粗糙性度量结合起来使用。

结束语 本文利用由隶属度定义的强包含度，通过引入精度系数 $k\in(0.5,1]$，在弱模糊划分基础上，建立了基于包含度的变精度模糊粗糙集模型，研究了该模型的一些性质，给出了该模型在精度系数 k 下从经典粗糙集模型推广得到的粗糙度计算公式，并定义了两种粗糙性度量方法，将这些度量集合起来使用，可补充经典粗糙集中推广的粗糙度在模糊粗糙集模型中的不足。文中运用例题对所讨论的模型进行了分

(下转第 252 页)

仿真图像检验法是目前国际上常用的亚像素定位算法的验证方法。本文使用 Matlab 生成棋盘格图像，并跟据文献中仿真图像制作方法进行处理获得测试仿真图。分别使用工具箱中 Harris 角点检测算法和本文检测算法对仿真图像进行控制点检测，以获得观测数据。仍然使用观测值与理论坐标的标准差作为评价定位算法的性能指标。从实验中发现，在不加入噪声的情况下，两种算法均能够准确检测出控制点坐标。在加入方差为 0.01 的高斯噪声后，Harris 检测算法的观测值与理论值的标准差为[0.0341 0.0362]，本文算法的观测值与理论值的标准差为[0.0281 0.0262]。从仿真图像检验实验中可以看出，本文算法具有更好的噪声稳定性。

4.3 时间性能分析

从本文算法的检测结果来看，本文算法较已有算法具有较好的有效性和检测精确，但其在 SUSAN 原有算法的基础上增加了两个约束算子，因此在时间复杂度上较 SUSAN 算法要高。但由于本文算法增加的两个算子仅在 SUSAN 结果集中计算约束条件，因此增加的计算量不高，对 SUSAN 算法整体的时间复杂度增加不大。表 2 是本文算法两个约束算子与原有 SUSAN 算法的时间效率对比。

表 2 算法主要模块 CPU 时间比较

图像	算法时间(ms)			
	SUSAN	symmetry	uniformity	Total
HR	44.121	0.732	0.572	45.725
HS-L	13.723	0.535	0.330	14.832
HS-R	14.239	0.428	0.315	15.322

由上表数据可以看出，本文算法引入两个约束算子的计算量并没有大幅提升原有算法的时间消耗。

结束语 本文在 SUSAN 角点检测算法的基础上，通过对棋盘格标定图像中黑-黑格交点的特性分析，提出了一种基于多方向对称和匀质约束的 SUSAN 棋盘格角点检测算法。实验表明该算法具有较好的稳定性与定位精度，适合于摄像机标定应用中棋盘格角点检测应用。

参 考 文 献

[1] Kitchen L, Rosenfeld A. Gray level corner detector[J]. Pattern Recognition Letters, 1982, 3(1): 95-102

[2] Harris C, Stephens M. A combined corner and edge detector[C]// Proceedings of the 4th Alvey Vision Conference. 1988: 147-151

[3] Shi J, Tomasi C. Good Features to Track[C]// Proceedings of CVPR'94. 1994: 593-600

[4] Smith S M, Brady J M. SUSAN-A new approach to low level image processing[J]. International Journal of Computer Vision, 1997, 23(1): 45-78

[5] He Xiao-chen, Yung N H C. Corner detector based on global and local curvature properties[J]. Optical Engineering, 2008, 47(5): 1-12

[6] 王志衡. 基于局部方向分布的角点检测及亚像素定位[J]. 软件学报, 2008, 19(11): 2932-2942

[7] 刘阳成，朱枫. 一种新的棋盘格图像角点检测算法[J]. 中国图象图形学报, 2006, 11(5): 656-660

[8] 刘阳，王福利，常玉清，等. 黑白棋盘格角点检测算法[J]. 东北大学学报：自然科学版, 2007, 28(8): 1090-1093

[9] 罗钧，王莲，侯艳. 摄像机标定的棋盘格亚像素角点检测[J]. 重庆大学学报, 2008, 31(6): 615-618

[10] 于起峰，尚洋. 摄像测量学原理与应用研究[M]. 北京：科学出版社, 2009

[11] Camera Calibration Toolbox for Matlab [EB/OL]. http://www.vision.caltech.edu/bouguetj/calib_doc/, 2011-03-13

(上接第 233 页)

析，验证了相关结论。直接应用模糊集理论的相关知识来定义概念和分析结论，得到的结果使得基于包含度的模糊粗糙集模型可以直接利用对象的隶属度进行相关的计算和研究，为基于变精度模糊粗糙集的进一步研究和应用提供帮助。本文研究的模糊粗糙集模型，能够完善粗糙集理论，使概念的刻画更加准确，能在网络数据分类、文本识别、命名实体识别等实际应用中发挥粗糙集理论的优势，更好地挖掘有价值的信息。

参 考 文 献

[1] Zadeh L A. Fuzzy sets [J]. Information and Control, 1965(8): 338-353

[2] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Science, 1982(11): 341-356

[3] Huang Zheng-hua, Hu Bao-qing. The studies of fuzzy rough sets theory: A survey [J]. Fuzzy sets and Systems, 2004(19): 125-134

[4] Dubois D, Prade H. Rough fuzzy sets and fuzzy rough sets [J]. International Journal of General Systems, 1990(17): 191-208

[5] Ziarko W. Variable precision rough sets characterized by fuzzy sets [J]. Journal of Computer and Information Science, 1993 (46): 39-59

[6] Kuncheva L I. Fuzzy rough sets: Application to feature selection [J]. Fuzzy sets and System, 1992(51): 147-153

[7] Bodjanova S. Approximation of fuzzy concepts in decision making [J]. Fuzzy sets and Systems, 1997(85): 23-29

[8] Mi Ju-sheng, Zhang Wen-xiu. An axiomatic characterization of fuzzy generalization of rough sets [J]. Information Sciences, 2004(160): 235-249

[9] Wu Wei-zhi, Mi Ju-sheng, Zhang Wen-xiu. Generalized fuzzy rough sets [J]. Information Sciences, 2003(151): 263-282

[10] Zhang Xiao-yan, Xu Wei-hua. A novel approach to roughness measure in fuzzy rough sets [J]. Fuzzy Information and Engineering, 2007(40): 775-780

[11] Xu Wei-hua, Zhang Wen-xiu. Measuring roughness of generalized rough sets induced by a covering [J]. Fuzzy sets and Systems, 2007(158): 2443-2445

[12] Xu Wei-hua, Zhang Xiao-yan. Fuzziness in covering generalized rough sets [C]// Proceedings of the 26th Chinese Control Conference. 2007

[13] 解滨，李磊军，米据生. 基于知识粒度的粗糙集的不确定性度量[J]. 计算机科学, 2010, 37(9): 225-228

[14] 杨纶标，高英仪. 模糊数学原理及应用[M]. 广州：华南理工出版社, 2005

[15] 张文修，吴伟志，等. 粗糙集理论与方法[M]. 北京：科学出版社, 2001

[16] 张文修，梁怡，吴伟志. 信息系统与知识发现[M]. 北京：科学出版社, 2003

[17] 袁修久，张文修. 模糊粗糙集的包含度和相似度[J]. 模糊系统与数学, 2005, 19(1): 111-115

[18] 范九伦，吴茂成. 用于聚类有效性判定的包含度公式[J]. 模糊系统与数学, 2002, 16(1): 80-86