

# Vague 集的三维表示及一组距离测度<sup>\*</sup>

黄国顺<sup>1,2</sup> 刘云生<sup>1</sup>

(华中科技大学计算机学院 武汉430074)<sup>1</sup> (佛山科技学院理学院 佛山528000)<sup>2</sup>

**摘要** Gau 和 Buehrer 提出 Vague 集以来,为刻画两个 Vague 集(值)之间的相似程度,一些 Vague 集的相似度量方法被提出来,但在许多情况下,求相似度量往往和距离测度有关。本文通过建立 Vague 集隶属函数的三维表示方法,证明了一个不等式,在此基础上给出了一组 Vague 集(值)之间的距离测度公式,并证明它们满足距离测度的有关性质,最后给出一个它们的实例应用。

**关键词** Fuzzy 集, Vague 集, 距离测度

## A Three-Dimension Expression of Vague Sets and a Group of Distance Measures between Them

HUANG Guo-Shun<sup>1,2</sup> LIU Yun-Sheng<sup>2</sup>

(College of Computer, Huazhong University of Science and Technology, Wuhan 430074)<sup>1</sup>

(School of Science, Foshan College of Science and Technology, Foshan 528000)<sup>2</sup>

**Abstract** Since the Vague set was proposed by Gan and Buehrer, some measures of similarity between Vague sets (values) have been developed to describe the degree of similarity between them, but similarity measures are related with distance measures in many cases. A three-dimension expression of Vague set membership degrees is proposed in this paper. We prove a inequality and then a group of distance measures between Vague sets (values) are given. Some properties of them are proved and a example is presented at last.

**Keywords** Fuzzy set, Vague set, Distance measure

## 1 引言

1993年 Gau 和 Buehrer 提出了 Vague 集的概念<sup>[1]</sup>。在一个 Vague 集  $V$  中,用一个真隶属函数  $t_v(x)$  和一个假隶属函数  $f_v(x)$  来描述其隶属度的边界,这两个界是  $[0,1]$  的一个子区间  $[t_v(x), 1-f_v(x)]$ 。Vague 集的特点是它同时给出了支持和反对的证据,目前 Vague 集已被国内外成功地应用于模糊控制、决策分析、模式识别等领域<sup>[2~4]</sup>,而其中关键的技术之一就是怎样用相似度量去度量两个集合的相似程度,但在许多情况下求相似度量都和距离有关<sup>[5~7]</sup>,文<sup>[8]</sup>虽然定义了一套 Vague 集(值)之间的距离测度,但方法远不止这一种,本文从 Vague 集的三维表示入手建立三维度量空间,证明了一个不等式并以此为基础建立一组距离测度公式。

## 2 预备知识

本节先回顾一下 Vague 集的定义及距离的有关性质。

### 2.1 Vague 集的定义和基本标记符号

**定义1** 令  $U$  是一个点(对象)的空间,其中任意一个元素用  $x$  表示, $U$  中的一个 Vague 集  $V$  用真隶属函数  $t_v$  和一个假隶属函数  $f_v$  数表示, $t_v(x)$  是从支持  $x$  的证据导出的  $x$  的隶属度下界, $f_v(x)$  是从反对  $x$  的证据所导出的  $x$  的否定隶属度的下界, $t_v(x)$  和  $f_v(x)$  将区间  $[0,1]$  中的实数与  $U$  中每一点联系起来,即

$$t_v: u \rightarrow [0,1] \quad f_v: u \rightarrow [0,1]$$

其中  $t_v(x) + f_v(x) \leq 1$ 。记  $\pi_v(x) = 1 - t_v(x) - f_v(x)$ ,表示

$x$  在  $V$  中的不确定程度。当论域  $U$  是离散(连续)的时候,则 Vague 集  $V$  可写成

$$V = \sum_{i=1}^n [t_v(x_i), 1 - f_v(x_i)] / x_i \quad x_i \in U$$

$$(V = \int [t_v(x), 1 - f_v(x)] / x_i \quad x \in U)$$

为简单起见,下文分别用  $t_x, f_x, \pi_x$  代替  $t_v(x), f_v(x), \pi_v(x)$ 。

### 2.2 距离的有关性质

设  $d(x, y)$  是某度量空间规范化的距离公式,当且仅当它满足如下性质: (P1)  $d(x, y) = 0$  当且仅当  $x = y$ ; (P2)  $0 \leq d(x, y) \leq 1$ ; (P3)  $d(x, y) = d(y, x)$ ; (P4) 对任意的  $x, y, z \in U, d(x, z) \leq d(x, y) + d(y, z)$ 。

## 3 Vague 值的三维表示及其值间的距离测度

由于 Vague 集  $V$  中的任何一个元素  $x$  都有

$$t_x + f_x + \pi_x = 1 \quad (1)$$

其中  $0 \leq t_x \leq 1, 0 \leq f_x \leq 1, 0 \leq \pi_x \leq 1$ 。

这使我们想到三维空间  $R^3$  中的平面方程式,因而我们借助于三维空间来形象刻画  $x$ , 引进笛卡尔坐标  $(t_x, f_x, \pi_x)$ , 其中  $t_x$  表示  $x$  的真隶属度,  $f_x$  表假隶属度,  $\pi_x$  表示未知隶属度,那么根据(1)式及条件有:它表示的是一个有限截平面,其图如图1所示,那么对于 Vague 集  $V$  中的任何一点都应该落在图1的  $\triangle ABC$  上。

$\triangle ABC$  是边长为  $\sqrt{2}$  的等边三角边,由  $R^3$  空间的性质,

<sup>\*</sup> 本文得到广东省自然科学基金(034071)资助。黄国顺 博士生,主要研究领域为信息融合,数据挖掘,模糊信息处理。刘云生 教授,博士生导师,主要研究领域为现代数据库理论与技术及其实现,数据库与信息系统开发及实时应用,信息融合。

可定义 \$V\$ 中任何两元素 \$x, y\$ 之间的距离并规范化后有

$$d_1(x, y) = \sqrt{\frac{(t_x - t_y)^2 + (f_x - f_y)^2 + (\pi_x - \pi_y)^2}{2}} \quad (2)$$

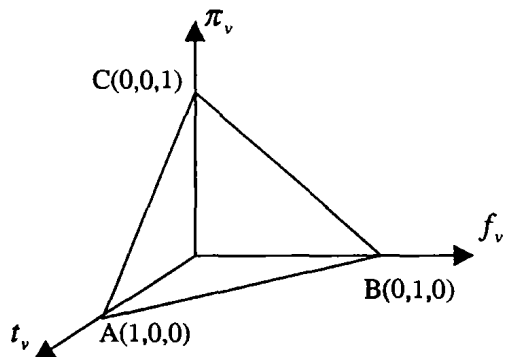


图1

由 \$R^3\$ 空间的性质及图1中任意两点间的距离不超过 \$\sqrt{2}\$ 知 \$d\_1(x, y)\$ 满足性质(P1)~(P4)。

显然, 当 \$\triangle ABC\$ 上任意两点间距离越大, 表示它们的相似程度越小, 反之, 距离越小则表示它们的相似程度越大。特别地, 当 \$d\_1(x, y) = 0\$ 则表明 \$x\$ 和 \$y\$ 完全一样。

设 \$x = [t\_x, 1 - f\_x], y = [t\_y, 1 - f\_y],\$ 其中, \$0 \le t\_x \le 1 - f\_x, 0 \le t\_y \le 1 - f\_y,\$ 则有如下不等式成立

$$\frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{2} \le 1 \quad (3)$$

证明: 当 \$(t\_x - t\_y)(f\_x - f\_y) < 0\$ 时, 因为 \$(1 - (t\_x - t\_y))^2 (1 - (f\_x - f\_y))^2 \ge 0\$ 从而有 \$1 + (t\_x - t\_y)^2 (f\_x - f\_y)^2 \ge (t\_x - t\_y)^2 + (f\_x - f\_y)^2\$

即 \$[1 + (t\_x - t\_y)(f\_x - f\_y)]^2 \ge [(t\_x - t\_y) + (f\_x - f\_y)]^2\$  
故有 \$1 + (t\_x - t\_y)(f\_x - f\_y) \ge |(t\_x + f\_x) - (t\_y + f\_y)|\$ (4)

当 \$(t\_x - t\_y)(f\_x - f\_y) \ge 0\$ 时, 由于 \$|(t\_x + f\_x) - (t\_y + f\_y)| \le 1,\$ (4)式显然成立。

将(4)式两边乘2再加上 \$(t\_x - t\_y)^2 + (f\_x - f\_y)^2\$ 有  
\$(|t\_x - t\_y| + |f\_x - f\_y|)^2 \le (2 - |(t\_x + f\_x) - (t\_y + f\_y)|)^2\$  
即有 \$|t\_x - t\_y| + |f\_x - f\_y| \le 2 - |(t\_x + f\_x) - (t\_y + f\_y)|\$  
故有

$$\frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{2} \le 1$$

根据不等式(3)可定义第二种距离测度  
 $d_2(x, y) = \frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{2} \quad (5)$

根据(5)式知它满足(P1)~(P4), 其中(P1)~(P3)结论易知, 下面只证明(P4)

证明: \$d\_2(x, y) + d\_2(y, z) = \frac{|t\_x - t\_y| + |f\_x - f\_y| + |\pi\_x - \pi\_y|}{2} + \frac{|t\_y - t\_z| + |f\_y - f\_z| + |\pi\_y - \pi\_z|}{2} \ge \frac{|t\_x - t\_y + t\_y - t\_z| + |f\_x - f\_y + f\_y - f\_z| + |\pi\_x - \pi\_y + \pi\_y - \pi\_z|}{2} = d\_2(x, z)\$

更一般地, 有如下两点间距离测度公式  
 $d_3(x, y) = \sqrt{\frac{|t_x - t_y|^p + |f_x - f_y|^p + |\pi_x - \pi_y|^p}{2}} \quad (p \ge 1) \quad (6)$

当 \$p = 1\$ 时 \$d\_3(x, y)\$ 变成(5)式, 当 \$p = 2\$ 时变成(2)式, 由于  
 $\frac{|t_x - t_y|^p + |f_x - f_y|^p + |\pi_x - \pi_y|^p}{2} \le 1$

$$\frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{2} \le 1$$

故 \$d\_3(x, y) \in [0, 1],\$ 满足(P2), 显然(6)式也满足(P1)和(P3), 由 Minkowski 不等式知它满足(P4)。

在文[8]中, 卢和王给出了一对 Vague 值 \$x, y\$ 之间的距离测度公式

$$L(x, y) = \frac{|S(x) - S(y)|}{4} + \frac{|t_x - t_y| + |f_x - f_y|}{4} \quad (7)$$

下面我们给出一个类似的公式

$$d_4(x, y) = \frac{|S(x) - S(y)|}{4} + \frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{4} \quad (8)$$

其中, \$S(x) = t\_x - f\_x, S(y) = t\_y - f\_y,\$ 根据不等式(4)知 \$0 \le d\_4(x, y) \le \frac{2}{4} + \frac{2}{4} = 1,\$ 满足(P2), 同样(8)式满足(P1)、(P3)和(P4), 下面只证明它满足(P4)

证明: \$d\_4(x, y) + d\_4(y, z) = \frac{|S(x) - S(y)|}{4} + \frac{|t\_x - t\_y| + |f\_x - f\_y| + |\pi\_x - \pi\_y|}{4} + \frac{|S(y) - S(z)|}{4} + \frac{|t\_y - t\_z| + |f\_y - f\_z| + |\pi\_y - \pi\_z|}{4} \ge \frac{|S(x) - S(y) + S(y) - S(z)|}{4} + \frac{|t\_x - t\_y + t\_y - t\_z| + |f\_x - f\_y + f\_y - f\_z| + |\pi\_x - \pi\_y + \pi\_y - \pi\_z|}{4} = \frac{|S(x) - S(z)|}{4} + \frac{|t\_x - t\_z| + |f\_x - f\_z| + |\pi\_x - \pi\_z|}{4} = d\_4(x, z)\$

对于 \$d\_4(x, y)\$ 还可以改写成下式

$$d_4(x, y) = \frac{|S(x) - S(y)|}{4} + \frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{4} = \frac{3}{4} \max(|t_x - t_y|, |f_x - f_y|) + \frac{1}{4} \min(|t_x - t_y|, |f_x - f_y|)$$

故引入第5种距离测度公式

$$d_5(x, y) = \lambda_1 \max(|t_x - t_y|, |f_x - f_y|) + \lambda_2 \min(|t_x - t_y|, |f_x - f_y|) \quad (9)$$

其中, \$0 < \lambda\_1 < 1, 0 < \lambda\_2 < 1, \lambda\_1 + \lambda\_2 = 1,\$ 且 \$\lambda\_1 \ge \lambda\_2\$

由(9)式知, \$d\_5(x, y)\$ 满足(P1)和(P3), 又因

$$\lambda_1 \max(|t_x - t_y|, |f_x - f_y|) + \lambda_2 \min(|t_x - t_y|, |f_x - f_y|) \le \lambda_1 + \lambda_2 = 1$$

故(9)式满足(P2), 下面证明它一样满足(P4)

证明: \$\lambda\_1 \max(|t\_x - t\_y|, |f\_x - f\_y|) + \lambda\_2 \min(|t\_x - t\_y|, |f\_x - f\_y|) + \lambda\_1 \max(|t\_y - t\_z|, |f\_y - f\_z|) + \lambda\_2 \min(|t\_y - t\_z|, |f\_y - f\_z|) = (\lambda\_1 - \lambda\_2) \max(|t\_x - t\_y|, |f\_x - f\_y|) + \lambda\_2 (|t\_x - t\_y| + |f\_x - f\_y|) + (\lambda\_1 - \lambda\_2) \max(|t\_y - t\_z|, |f\_y - f\_z|) + \lambda\_2 (|t\_y - t\_z| + |f\_y - f\_z|) = (\lambda\_1 - \lambda\_2) (\max(|t\_x - t\_y|, |f\_x - f\_y|) + \max(|t\_y - t\_z|, |f\_y - f\_z|)) + \lambda\_2 (|t\_x - t\_y| + |f\_x - f\_y| + |t\_y - t\_z| + |f\_y - f\_z|) \ge (\lambda\_1 - \lambda\_2) \max(|t\_x - t\_y| + |t\_y - t\_z|, |f\_x - f\_y| + |f\_y - f\_z|) + \lambda\_2 (|t\_x - t\_y| + |f\_x - f\_y| + |t\_y - t\_z| + |f\_y - f\_z|) = (\lambda\_1 - \lambda\_2) \max(|t\_x - t\_z|, |f\_x - f\_z|) + \lambda\_2 \max(|t\_x - t\_z|, |f\_x - f\_z|) + \min(|t\_x - t\_z|, |f\_x - f\_z|) = \lambda\_1 \max(|t\_x - t\_z|, |f\_x - f\_z|) + \lambda\_2 \min(|t\_x - t\_z|, |f\_x - f\_z|) = d\_5(x, z)\$

对于 \$d\_5(x, y)\$, 若 \$\lambda\_1 = \lambda\_2 = 1/2,\$ 则是 \$d\_2(x, y)\$ 的前两项之和, 而若 \$\lambda\_1 = \frac{3}{4}, \lambda\_2 = \frac{1}{4},\$ 则 \$d\_5(x, y)\$ 变成 \$d\_4(x, y)\$。

下面引入另一种距离测度公式

理,逐一地匹配训练得到的规则集合的前提条件,从而可以识别出图像的病害类型。对于图4所示的图像,因为满足 Rule03 的前提条件: holeORnot ≤ 0,bscale ≤ 0, areamedian < 95,所以判断为横裂,得到如图6所示结果。

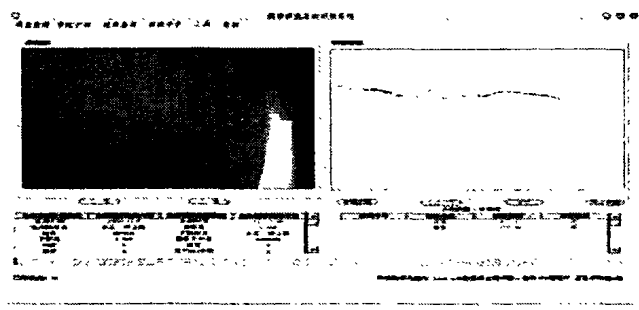


图6 识别结果

(6)显示识别结果,并将该结果存入对应的数据库中,以

(上接第118页)

$$d_6(x,y) = \frac{2(|t_x - t_y| + |f_x - f_y|)}{2 + |t_x - t_y| + |f_x - f_y|} \quad (10)$$

由  $0 \leq |t_x - t_y| \leq 1, 0 \leq |f_x - f_y| \leq 1$ , 及  $d_6(x,y)$  定义式易知它满足 (P1)~(P3), 下证它满足 (P4)

由于函数  $\frac{2x}{2+x}$  是关于  $x$  的增函数且有  
 $|t_x - t_y| + |f_x - f_y| + |t_y - t_z| + |f_y - f_z|$   
 $\geq |t_x - t_z| + |f_x - f_z|$

故

$$\begin{aligned} d_6(x,y) + d_6(y,z) &= \frac{2(|t_x - t_y| + |f_x - f_y|)}{2 + |t_x - t_y| + |f_x - f_y|} + \frac{2(|t_y - t_z| + |f_y - f_z|)}{2 + |t_y - t_z| + |f_y - f_z|} \\ &\geq \frac{2(|t_x - t_y| + |f_x - f_y| + |t_y - t_z| + |f_y - f_z|)}{2 + |t_x - t_y| + |f_x - f_y| + |t_y - t_z| + |f_y - f_z|} \\ &\geq \frac{2(|t_x - t_z| + |f_x - f_z|)}{2 + |t_x - t_z| + |f_x - f_z|} = d_6(x,z) \end{aligned}$$

#### 4 Vague 集之间的距离测度

假设  $A$  和  $B$  是论域  $U = \{u_1, u_2, \dots, u_n\}$  上的两个 Vague 集, 其中

$$A = \sum_{i=1}^n t_A(u_i), 1 - f_A(u_i) / u_i \quad (11)$$

$$B = \sum_{i=1}^n t_B(u_i), 1 - f_B(u_i) / u_i \quad (12)$$

$V_A(u_i) = [t_A(u_i), 1 - f_A(u_i)]$  表示 Vague 集  $A$  中  $u_i$  的隶属值;  $V_B(u_i) = [t_B(u_i), 1 - f_B(u_i)]$  表示 Vague 集  $B$  中  $u_i$  的隶属值,  $i = 1, 2, \dots, n$ , 则 Vague 集  $A$  和  $B$  的距离测度由下式计得:

$$R_j(A,B) = \frac{1}{n} \sum_{i=1}^n d_j(V_A(u_i), V_B(u_i)) \quad j = 1 \sim 6$$

其中,  $d_j(x,y)$  是上节定义的两点之间的各距离测度公式, 由它们都满足 (P1)~(P4) 知  $R_j(A,B)$  有如下性质

- 性质1  $R_j(A,B) = 0$  当且仅当  $A = B$
- 性质2  $R_j(A,B) \in [0, 1]$
- 性质3  $R_j(A,B) = R_j(B,A)$
- 性质4  $R_j(A,B) + R_j(A,C) \geq R_j(B,C)$

#### 5 实例应用

最后考虑文[5,6]中的一个应用: 设  $A$  是  $B$  论域  $U$  中的

提供给后期使用。

**结束语** 本系统在 Windows 2000 操作系统下用 VC++ 6.0 编程实现, 运行结果表明, 该系统所采用的图像处理方法、基于决策树的知识获取和推理方法性能良好, 并已被实际应用于江苏的沪宁高速公路的路面破损识别, 其效果很好。

#### 参考文献

- 1 李晋惠, 楼伟, 姜寿山. 基于 CCD 的公路路面病害检测技术研究. 西安工业学院学报, 2002, 22(2): 95~99
- 2 SconKa M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision (Second Edition). 北京: 人民邮电出版社, 2002
- 3 Han Jiawei, Kamber M. Data Mining Concepts and Techniques. Beijing: Higher Education Press, 2001
- 4 刘小虎, 李生. 决策树的优化算法. 软件学报, 1998, 9(10): 797~800
- 5 张敏灵, 周志华. 多示例学习的理论分析. 计算机科学, 2002, 29(9): 242~244

两个 Vague 集,  $U = \{u_1, u_2, u_3, u_4\}$ ,  $A$  和  $B$  的取值如表1。

表1

	A	B	C
$u_1$	[0.4, 0.6]	[0.6, 0.7]	[0, 1]
$u_2$	[0.8, 0.9]	[0.3, 0.5]	[0.4, 0.6]
$u_3$	[0.9, 0.9]	[0, 0]	[0.8, 0.9]
$u_4$	[0, 0]	[0.5, 0.8]	[1, 1]

则  $A$  和  $B$  之间的距离  $R_j(A,B)$  和相似程度  $S_j(A,B) = 1 - R_j(A,B)$  分别为

	1	2	3	4	5	6
$R_j(A,B)$	0.5579	0.6	0.5549	0.5687	0.5519	0.6541
$S_j(A,B)$	0.4421	0.4	0.4451	0.4313	0.4481	0.3459

注: 对于  $R_3(A,B)$  中取  $p = 3$ , 而  $R_5(A,B)$  中取  $\lambda_1 = \frac{5}{8}, \lambda_2 = \frac{3}{8}$

**结论** 本文通过建立 Vague 集的三维表示方法, 证明了一个不等式, 然后在此基础上提出了一组 Vague 集(值)之间的距离测度, 并证明它们满足距离的一些性质, 最后给出文[5,6]中的实例应用。

#### 参考文献

- 1 Gau Wen-Lung, Buechrer D J. Vague sets. IEEE Transaction Systems, man, and Cybernetics, 1993, 23(2): 610~614
- 2 Chen Shyi-Ming, Tan Jian-Mean. Handling multicriteria fuzzy decision-making, problems based on Vague set theory. Fuzzy sets and systems, 1994, 67(2): 163~172
- 3 马志锋, 邢汉承, 郑晓妹. 不完整 Vague 决策表中的近似学习方法. 计算机研究与发展, 2000, 37(9): 1050~1057
- 4 Liang Zhi-Zhen, Shi P F. Similarity measures on intuitionistic fuzzy sets. Pattern Recognition Letters, 2003, 24: 2687~2693
- 5 Chen S M. Measures of similarity between Vague sets and between elements. IEEE Transaction Systems, Man and Cybernetics-part B: Cybernetics, 1997, 27(1): 153~158
- 6 Hong D H, Kin C. A note on similarity measures between vague sets and between elements. Information sciences, 1999, 135: 83~96
- 7 李凡, 徐章艳. Vague 集之间的相似度量. 软件学报, 2001, 12(6): 922~926
- 8 卢正鼎, 王天江. Vague 集(值)之间的距离测度. 计算机科学, 2003, 30(7): 154~156