Int J Adv Manuf Technol (2008) 36:1072–1079
DOI 10.1007/s00170-006-0922-7

ORIGINAL ARTICLE

A variable precision rough set based modeling method for pulsed GTAW

W. H. Li · S. B. Chen · B. Wang

Received: 11 May 2006 / Accepted: 20 December 2006 / Published online: 24 March 2007

Abstract Modeling is an important step both for quality and shaping control of the arc welding process. Current modeling methods have made great advances in the field of arc welding, however they all posses certain limitations. It is due to these limitations that we created the variable precision rough set (VPRS) based modeling method. The VPRS modeling has been shown to be both a more efficient and reliable modeling method for the arc welding process due to its ability to account for the character of the welding media. The method was used to produce a dynamic predictive model for pulsed gas tungsten arc welding (GTAW). Results showed that the VPRS modeling method was able to sufficiently acquire knowledge during welding practices. In addition, comparison of VPRS model with classic rough set model and BP neural network model showed that VPRS model was more stable and could predict the unseen data better than classic RS model. Moreover, the VPRS model owns similar precision with neural network model, but has better understandability.

Keywords GTAW · Modeling · Variable precision rough set (VPRS) · Welding process

1 Introduction

Modeling of pulsed gas tungsten arc welding (GTAW) is important for the automation of this type of welding. A model that can predict the backside width of a welding pool is critical to get the desired quality and shape of a welding seam. However, it is difficult to obtain a precise mathematical model from the classical modeling methods, because the welding process is inherently nonlinear, time-delayed, has a strong coupling in its input/output relationships and involves many uncertain factors (such as metallurgy, heat transfer, chemical reaction, arc physics, and magnetization.)

Recently, intelligent modeling methods, such as neural network (NN), fuzzy set (FS) and rough set (RS) methods [1–7], have attracted significant attention in welding for their adaptability for complex system. However, the NN model is hard to modify, and has to be continually rebuilt whenever new samples are added. Network structures as well as transfer functions need to be determined prior to building a NN model. This is a specialized task that needs experienced operators to perform and often the results are flawed by finding local minima. For the FS model, trained operators which are scarce and usually hard to obtain in welding field, are needed. Furthermore, if the FS model's inputs, outputs and linguistic variables are too many, rule explosion will occur. Contrary to older models, the RS theory can easily acquire knowledge from data even when the operator has limited prior knowledge. Additionally, the model has the ability to reduce superfluous variables, is easily commanded with 'IF THEN' statements and can simply be modified. Therefore, the theory of RS modeling will be useful for the future of welding where skilled welders' may be in limited supply. However, the fact that the classic RS model is sensitive to noise seriously reduced its reliability and efficiency and may harm its applications [8].

To take the advantages of RS model and overcome its shortages, a generalized rough set-variable precision rough set (VPRS) [8] is introduced to welding because of its better noise suppression merits. In this paper, the RS model

W. H. Li (✉) · S. B. Chen · B. Wang
School of Materials Science and Engineering,
Shanghai Jiao Tong University,
Shanghai 200030, People's Republic of China
e-mail: lwh-ikl@sjtu.edu.cn

proposed by Pawlak [9] is called classic RS. This paper's contents are organized as shown below: In Sect. 2, the basic concepts and background of VPRS is provided. In Sect. 3, the procedure of the VRPS modeling method is introduced and the key reduction algorithms in the modeling method are addressed in detail. In Sect. 4, a dynamic welding process model is obtained by the VPRS modeling method. Lastly in Sects. 5 and 6, the VPRS model is compared with the classic RS model and NN model respectively. Conclusions are drawn accordingly in the end of this paper.

2 Brief description of VPRS

Both VPRS and classic RS operate on what can be described as a *decision table* or *information system*. As illustrated in Table 1 [10], a set of *objects* U (o_1,..., o_7) are contained in the rows of the table. The columns denote *condition attributes* C (c_1,..., c_6) of these objects and a related *decision attribute* D (d). A value denoting the nature of an attribute to an object is called a *descriptor*. A VPRS data requirement is that it must be in discrete or categorical form. The table shows that the objects have been classified into one of these decision values, which are also referred to as *concepts*.

For the condition attributes in this example, all of the objects (U) can be placed in five groups: $\mathbf{X_1}=\{o_1, o_4, o_6\}$, $\mathbf{X_2}=\{o_2\}$, $\mathbf{X_3}=\{o_3\}$, $\mathbf{X_4}=\{o_5\}$ and $\mathbf{X_5}=\{o_7\}$. The objects within a group are indiscernible from each other. For example objects o_1, o_4 and o_6 in X_1 have the same descriptor values for each of the condition attributes. These groups of objects are referred to as *equivalence classes* or *conditional classes* for their specific attributes. The equivalence classes for the decision attribute are: $\mathbf{Y_1}=\{o_1, o_2, o_3\}$ and $\mathbf{Y_2}=\{o_4, o_5, o_6, o_7\}$. The abbreviation of the set of equivalence classes for the conditional attributes C is denoted by $\mathbf{E(C)}=\{\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \mathbf{X_4}, \mathbf{X_5}\}$ and for the decision attribute it is defined $\mathbf{E(D)}=\{\mathbf{Y_1}, \mathbf{Y_2}\}$.

VPRS measurement is based on ratios of elements contained in various sets. A case in point is the conditional probability of a concept given a particular set of objects (a condition class). For example: Pr $(\mathbf{Y_1}|\mathbf{X_1})$ = Pr $(\{o_1, o_2, o_3\}|\ \{o_1, o_4, o_6\}) = |$ Pr $(\{o_1, o_2, o_3\} \cap \{o_1, o_4, o_6\}/|\{o_1, o_4, o_6\})|$=0.333.

Table 1 An example of decision table

Object	Condition attributes(C)						Decision attributes(D)
	c_1	c_2	c_3	c_4	c_5	c_6	d
O_1	1	0	1	1	0	1	0
O_2	1	0	0	0	0	0	0
O_3	0	0	1	0	0	0	0
O_4	1	0	1	1	0	1	1
O_5	0	0	0	0	1	1	1
O_6	1	0	1	1	0	1	1
O_7	0	0	0	0	1	0	1

It follows that this measures the accuracy of the allocation of the conditional class $\mathbf{X_1}$ to the decision class $\mathbf{Y_1}$. Hence for a given probability value, the β-positive region corresponding to a concept delineated as the set of objects with conditional probabilities of allocation at least equal to or larger than β. More formally, β-positive region of the set $\mathbf{Z}\subseteq\mathbf{U}$

$$POS_P^{\beta}(Z) = \bigcup_{p_r(Z|Xi)\geq\beta} \{X_i \in E(P)\}, \text{ with } P \subseteq C. \tag{1}$$

Following An et al. [11], β is defined to be within range of 0.5 to 1. Hence for the current example, the condition equivalence class $\mathbf{X_1}=\{o_1, o_4, o_6\}$ have a majority inclusion (with at least 60% majority needed, i.e., β=0.6) in $\mathbf{Y_2}$, in that most objects (2 out of 3) in $\mathbf{X_1}$ belong in $\mathbf{Y_2}$. Hence $\mathbf{X_1}$ is in $\mathrm{POS}_C^{0.6}$ $(\mathbf{Y_2})$. It follows $\mathrm{POS}_C^{0.6}$ $(\mathbf{Y_2})=\{o_1, o_4, o_5, o_6, o_7\}$.

Corresponding expressions for the β-boundary and β-negative regions are given by Ziarko [8], as follows: β-boundary region of the set $\mathbf{Z}\subseteq\mathbf{U}$

$$BND_P^{\beta}(Z) = \bigcup_{1-\beta<p_r(Z|Xi)<\beta} \{X_i \in E(P)\}, \text{ with P} \subseteq \text{C}, \tag{2}$$

β-negative region of the set $\mathbf{Z} \subseteq \mathbf{U}$

$$NEG_P^{\beta}(Z) = \bigcup_{p_r(Z|Xi)\leq\beta} \{X_i \in E(P)\}, \text{ with P} \subseteq \text{C}. \tag{3}$$

Using $\mathbf{P}$ and $\mathbf{Z}$ from the previous example, if β=0.6, then $BND_C^{0.6}(\mathbf{Y_2}) = \varnothing$ (empty set) and $NEG_C^{0.6}(\mathbf{Y_2}) = \{o_2, o_3\}$ can be obtained. Similarly, for the decision class $\mathbf{Y_1}$, it follows that $\mathrm{POS}_C^{0.6}(\mathbf{Y_1}) = \{o_2, o_3\}$, $BND_C^{0.6}(\mathbf{Y_1}) = \varnothing$ and $NEG_C^{0.6}(\mathbf{Y_1}) = \{o_1, o_4, o_5, o_6, o_7\}$.

VPRS applies these concepts by firstly seeking subsets of the attributes, which are capable (via construction of decision rules) of explaining allocations given by the whole set of condition attributes. These subsets of attributes are termed β - *reducts* or *approximate reducts*, and the process is called *attribute reduction*. Ziarko [8] states that a β -reduct, a subset $\mathbf{P}$ of the set of conditional attributes $\mathbf{C}$ with respect to a set of decision attributes $\mathbf{D}$, must satisfy the following conditions: (i) the subset $\mathbf{P}$ offers the same quality of classification (subject to the same β value) as the whole set of condition attributes $\mathbf{C}$; and (ii) no attribute can be eliminated from the subset $\mathbf{P}$ without affecting the quality of the classification (subject to the same β value).

The quality of the classification is defined as the proportion of the objects consisted of the union of the β -positive regions of all the decision equivalence classes based on the condition equivalence classes for a subset **P** of the condition attributes **C**. Associated with each conditional class is an upper bound on the β value above which there is no opportunity for majority inclusion and hence not in a β -positive region: in the previous example $\Pr(\mathbf{Y_1}|\mathbf{X_1})$=0.333 and $\Pr(\mathbf{Y_2}|\mathbf{X_1})$=0.667, hence if β=0.7 then $\mathbf{X_1}$ is not in the associated β-positive region, since the upper bound on β (for majority inclusion) is 0.667. The lowest of these upper bounds (amongst the condition classes) on the β values is defined $\beta_{\min}$ and relates to the overall level of confidence in classification by a particular β-reduct. The quality of classification associated with all the condition attributes is dependent on the $\beta_{\min}$ value. Hence for $\beta_{\min}\in(0.500, 0.667)$ the quality of classification is one $\left(\text{with } \mathbf{X_1} \text{ in } POS_C^\beta(\mathbf{Y_1})\right)$, for $\beta_{\min}\in$ (0.667, 1) the quality of classification is 0.571 (i.e., classifying four of the seven objects, with $\mathbf{X_1}$ in BND_C^β $(\mathbf{Y_2})$). Table 2 provides examples of four such β-reducts based on the condition attributes in Table 1.

After the attribute reduction, *attribute value reduction* and *rule reduction* are then to be computed. Attribute value reduction is to remove the superfluous condition attribute in each object for a given β -reduct while preserving the consistency of classifications. After the attribute value reduction, rule reduction is implemented to construct a minimal rule set, which also preserves the consistency of classifications of original decision table.

3 VPRS modeling method

When VPRS was applied into welding, the specific characteristics of the welding process and VPRS theory should be considered. To begin this section, the procedure of VPRS modeling method was first described. Then, the reduction algorithm in the modeling method was addressed.

3.1 Procedure of the VPRS modeling method

Figure 1 shows the procedure of VPRS modeling method. The VPRS based modeling method begins with acquiring raw data, which could be the results of experiments, history data, etc. Usually, the raw data cannot be directly treated by VPRS, and corresponding preprocessing should be adopted to improve the quality of data. Furthermore, if data in decision table is continuous, discretization is necessary.

Table 2 Examples of β-reducts from data shown in Table 1

$\beta_{\min}$	β-reduct	Classification quality
0.571	$\{c_2\}$	1
0.667	$\{c_1,c_3\}$	1
0.667	$\{c_1,c_4\}$	1
1.000	$\{c_4,c_5\}$	0.571

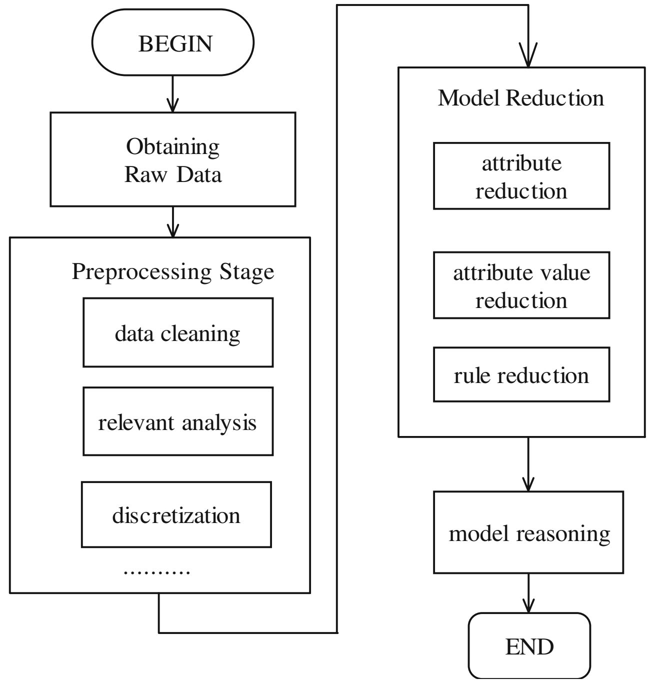

Fig. 1 Procedure of the VPRS modeling method

After the pre-processing, a decision table could be obtained, where condition attributes are input variables of system and decision attributes are output variables of system. It followed by attribute reduction, during which computation of attribute value reduction and rule reduction are conducted, the most important part in the modeling method. After that a VPRS model, made up of "IF THEN" rules, could be obtained. At last, the VPRS model could be used to predict the unseen objects (samples).

3.2 Key algorithm in VPRS

As mentioned above, the redution algorithm which includes attribute redution, attribute value reduction and rule reduction is the most important part in the modeling method. There might be more than one reduction result. The attribute reduction with the minimal number of attributes for a given decision table is called the *minimal reduction*, and it is usually the optimal solution. However, in RS theory, the optimal solution is NP-hard [12]. Heuristic information has to be used to improve the efficiency of algorithm and only suboptimal solution can be obtained.

3.2.1 Attribute reduction algorithm

In RS theory, there are some condition attributes which appear in all the attribute reduction and they are called *reduction core*. It is obvious that the reduction core should be determined first to avoid redundancy in the obtained reduction, and the computing of reduction core can be completed according to its definition in VPRS. Then, other attributes should be added into reduction core (or removed out condition attributes) to get the β-attribute reduction that maintained the same discernibility as the origin decision table.

Here hieuristic information is used by computing attributes' significance, which decided the attribute reduction strategy, and usually is defined according to the roughness of attributes, the information entropy or attribute frequency function in discernibility matrix (DM) [13]. The entropy based attribute importance is adopted in the work because it is computationally cheap and simple. Here is the definition:

Let H(a/R) be the condition entropy of the attribute a for the attribute set R. Let D be the decision attribute. The attribute significance of decision table can be defined as.

$$SGF(a,R,D) = H(D/R) - H(D/R \cup \{a\}).$$

In VPRS theory, all the attribute reductions are the same as their discernibility ability for the decision table. However in practical application, different condition attributes have different accuracy or cost in measurement. To make more attributes that are easy to measure and have high accuracy appear in the attribute reduction, the hieuristic information is used according to the entropy based significance, accuracy and cost of the measurement of the attribute. The new formula is as follows:

$$SGF(a,R,D) = k_1 \times (H(D/R) - H(D/R \cup \{a\})) + k_2 \times cost(a) + k_3 \times precise(a), \tag{4}$$

where k_1, k_2, k_3 are the weight value coefficients for different parts. If an attribute is easier to measure, its $cost(a)$ value is high. If its value have higher accuracy, the $precise(a)$ value is higher. For an attribute with higher $SGF(a,R,D)$, it is easier to be selected in the attribute reduction.

3.2.2 Algorithms of attribute value reduction and rule reduction

It has been shown that there are conditional equivalence relation between attribute reduction, attribute value reduction, rule reduction and minimum set cover [14]. Therefore, the attribute value reduction and rule reduction could be constructed based on minimal set cover, and DM could be used here. In this work, only attribute value reduction algorithm is given, and rule reduction algorithm is similar to that.

Before introducing the algorithm, the following definition is introduced.

(i) In a decision table DT, let x be an object. $FS_{DT}(x)$ is the set of 'attribute set' which discerns object x with other objects in DT. For example, in Table 1, $FS_{DT}(o_1)=\{\{c_3, c_4,c_6\}\{c_1,c_4,c_6\}\{c_1,c_3,c_4,c_5\}\{c_1,c_3,c_4,c_5,c_6\}\}$. The first element can discern o_1 with o_2, the second can discern o_1 with o_3, the third can discern o_1 with o_5 and the last can discern o_1 with o_7.

(ii) Let A be set of attributes, eg A=$\{\{a, b, c\}\ \{b, d\}\ \{e\}\ \{f\}\}$, ES(A) is the union of sets, which have only one element in A. In this case, ES(A)=$\{e\}\cup\{f\}=\{e, f\}$.

(iii) Let A, B be the attribute set, REM_SET(B, A)={A': A'=A-B, and A'$\neq\varnothing$}. For example, if A=$\{\{a,b\}\{a,b,c\}\{a,b,f\}\{h\}\{i\}\}\}$ and B=$\{a,b\}$ then REM_SET(A, B)=$\{\{c\}\{f\}\{h\}\{i\}\}$.

(iv) Let A, B be the attribute set, EXC_SET(B, A)={B: B∩A$\neq\varnothing$}. Take the example in (iii), EXC_SET(A,B)= $\{\{h\}\{i\}\}$. Now the attribute value reduction algorithm is shown below.

Attribute value reduction algorithm:

(1) Assign objects universal U to X;
(2) If X=∅, the program ends;
(3) Take a object x from X, X⇐X-$\{x\}$, computer $FS_{DT}(x)$;
(4) i⇐0;
(5) If i>$Kind$, go to (2);
(6) P⇐∅, Q⇐∅, T⇐∅;
(7) A⇐ $FS_{DT}(x)$;
(8) If A=∅, then i⇐ i+1 and go to (13), else go to (9);
(9) Q⇐ES(A);
(10) If Q=∅, then go to (11), else go to step (12);
(11) Randomly select an attribute a from C−(P∪T), A⇐ REM_SET(A,{a}), T⇐ T∪ $\{a\}$, go to (8);
(12) A⇐ EXC_SET(A, P), P⇐P∪Q, go to (8);
(13) if P have been obtained go to step (2), else note down P and go to step (5).

After the program, P noted down each time is the attribute value reduction. For any object x, its attribute value reduction is less than the value of $Kind$ in the algorithm.

4 Obtain VPRS model in welding

In arc welding the welding arc melts the workpiece and welding wire which generates a welding pool under the welding arc. With the movement of the welding arc, the former welding pool cools down and become solidified followed by a new welding pool appearing. The consecutive solidified welding pool will form the welding seam.

In welding quality control, fully fusion penetration was usually necessary for a valid welding seam. In butt welding, the backside width of welding seam or pool directly reflected the fusion penetration. Therefore, obtaining a forecasting model for backside width of welding seam is significant. In the following section, the VPRS modeling method will be used to build such a model for pulsed GTAW.

4.1 Obtaining raw date

Before raw date could be obtained, the input variables (condition attributes) and output variables (decision attributes) of the predictive model should be selected. According to the character of GTAW process, welding current (I), welding voltage, welding wire feed speed (S_F), welding speed (S_W) and topside shape of welding pool have relation with the backside width (W_B) of the welding pool. In GTAW, welding voltage is decided by the welding arc's length which is approximately equal to the distance between electrode and workpiece. In the experiment, we made the distance between electrode and workpiece the same and, the welding voltage value was assumed to be the same to simplify the VPRS model. Thus, the model's input variables were I, S_F, S_W and the shape of topside welding pool, which was denoted by W_T, L_T and W_B introduced below.

Here, a custom designed optical sensor [15] is used to obtain the topside and backside images of the welding pool in pulsed GTAW. After some image processing, the topside width, topside length and backside width of the welding pool represented by W_T, L_T and W_B respectively, could be obtained. Figure 2 is an example of obtained images and its feather after image process. The area of topside pool's area $A(=W_T \times L_T)$ is also added into the decision table. To make the model better reflect the real process, I, S_W and S_F are randomly changed within specified range. Considering the pure delay of welding process, the tth, (t−1)th,(t−2)th,(t−3)th time's I, S_W, S_F, A, W_T, L_T and W_B were used to predict the tth time's W_B. At last, a decision table which included 1013 samples (objects), with 27 condition attributes and 1 decision attribute are obtained. The parameters of pulsed GTAW are shown in Table 3. Part of raw data is shown in Table 4, where C denotes condition attributes and D denotes decision attribute.

4.2 Preprocessing

It is known that discretization for attributes (including decision attribute sometimes) in decision table, is usually necessary if their values are continuous. In this case, the equal width method is used to discrete the decision attribute, and each interval is 1 mm in width. Then, MDLP [16], an entropy based method, is used to discrete the condition attributes since the MDLP method has been demonstrated as better than other common algorithms in welding experiments [17].

4.3 Model reduction

The algorithms introduced in Sect. 3 are used to obtain the model. In attribute reduction, when computing the importance according Eq. (4). It is set that k_I=0.5, k_I=0.25, k_I=0.25 and the cost or precision of attributes are shown in Table 5. The larger the value SGF was, the more important the attribute is. In the experiment, the attribute reduction is "S_F^{t-3}, S_W^{t-1}, S_W^{t-2}, S_W^{t-3}, I^{t-1}, I^{t-2}, I^{t-3}, W_T^t, W_T^{t-3}, L_T^t, L_T^{t-1}, L_T^{t-2}, L_T^{t-3}, A^t, A^{t-1}, A^{t-2}, A^{t-3}, W_B^{t-1}, W_B^{t-2}, W_B^{t-3}".

In attribute value reduction, the most reduction for an object is $kind$=5. After attribute reduction and rule reduction, the VPRS model has 262 rules and the average length of rules is 5.25. It is obvious that the VPRS model has good reduction ability.

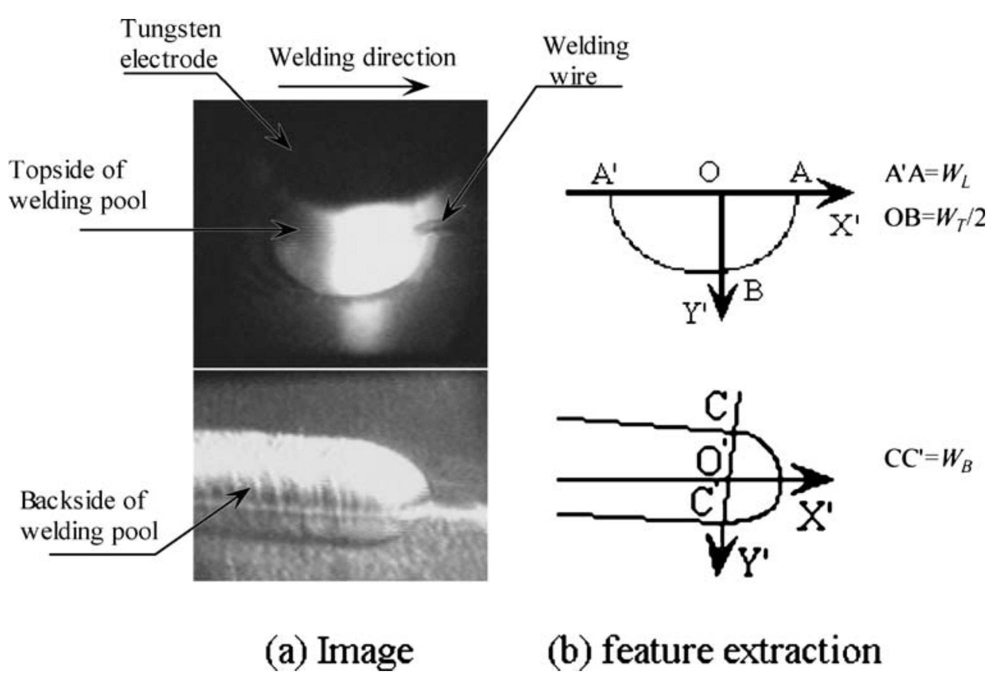

Fig. 2 An example of obtained picture

Table 3 Welding parameter of GTAW

Parameters	
Pulse:	Duty ratio=0.5
	Frequency=50 Hz
Workpiece:	A3 steel
	280 mm×80 mm×2 mm
	Closed flat butt joint
Tungsten electrode:	Diameter=3.2 mm
Welding current:	Peak value: variable
	Base value: 100A
Welding wire:	Wire feed speed: Variable
	Diameter=0.8 mm
Welding speed:	Variable
Protective atmosphere:	Argon gas (99%)
Arc length:	3.0 mm

4.4 Model reasoning

Approximate reasoning is implemented when the VPRS model is use to predict unseen data. Because the welding process is very complex, there will be few rules that match the unseen data. In the reasoning, the most matched five rules are selected, and the minimum *matching rate* of the model is 70%. Furthermore in rough set method, the longer a rule is, the better its prediction ability can be predicted. Therefore, a longer rule is prior to other rules which have the same matching rate. One of the obtained VPRS model contained 267 rules and the average length of a rule (including conclusion part) is 6. It indicates that the VPRS model can greatly reduce the redundant information of the origin decision table.

Table 4 Part data of soft steel mulit-variable decision table

C						D
S_F	S_W	I	W_T	L_T	A	W_B
9	2	132	7.47	10.72	80.0784	6.5
6	2.2	136	7.97	11.6	92.452	5.96
6	2.2	136	6.97	11.04	76.9488	5.07
5	1.6	140	8.13	9.6	78.048	5.25
5	1.6	140	7.47	9.28	69.3216	5.61
5	2.2	168	6.97	8.88	61.8936	6.41
5	2.2	168	7.14	10.32	73.6848	6.5
7	2	144	6.47	10.32	66.7704	6.5
7	2	144	8.13	10.32	83.9016	6.32
9	1.9	150	7.97	10.32	82.2504	6.14
9	1.9	150	7.97	10.4	82.888	5.7
8	1.8	132	7.3	11.84	86.432	5.61
8	1.8	132	7.8	9.84	76.752	5.79
7	2.3	146	7.14	8.88	63.4032	5.34

Table 5 The cost and precise value of attributes

	Cost	Precise
I	0.8	1
S_F	0.8	0.5
S_W	0.8	0.6
A	0.3	0.3
W_T	0.3	0.3
L_T	0.3	0.3
W_B	0.1	0.4

5 Comparison between VPRS and classic RS

Generally, the procedure of classic RS modeling method [2] is similar to that of VPRS. To compare their predictive ability, the two modeling method should use the same raw data, take the same preprocessing procedure, adopt the same reasoning method, and test obtained models on the same testing data. Because of the complexity of welding process, there are usually many noisy data in raw data. Moreover, RS model usually is obtained based on a small part of the sample space and is used to predict more unseen samples in practice. Therefore in the comparison, a small part of raw data is used for modeling and the larger for testing. The raw data is randomly split into 3 parts (2:1:1). The first 2/4 part is used as testing data, and the other 1/4 data is used to obtained classic RS or VPRS model. To better show the comparative result, following parameters are introduced.

1) Accurate rate R_{right}: the rate of right predict samples in testing data
2) Mean error:

$$E_{mean} = \frac{\sum_{i=1}^{N} \left| x_{i,\exp ect} - x_{i,pre} \right|}{N} \tag{5}$$

3) Standard error

$$E_{sd} = \sqrt{\frac{\sum_{i=1}^{N} \left(x_{i,\exp ect} - x_{i,pre} \right)^2}{N}} \tag{6}$$

In Eqs. 2 and 3, N is the number of testing samples, $x_{i,expect}$ is the expected value of sample x, and $x_{i,pre}$ is the predicted value of sample x.

The results are shown in Table 6. When β=1, VPRS became classic RS.

It shows that when the classic RS model (β=1) is used, the results accuracy varies within a relative large range, while the results of VPRS are stable. Furthermore, when

Table 6 Comparison between VPRS and classic RS model

β	The first 1/4			The second 1/4 part		
	R_{right}	E_{mean}	E_{sd}	R_{right}	E_{mean}	E_{sd}
1	55.6%	0.48	0.79	60.6%	0.42	0.72
0.9	62.1%	0.41	0.72	60.0%	0.42	0.72
0.8	57.7%	0.45	0.75	58.5%	0.46	0.77
0.7	58.5%	0.45	0.75	59.2%	0.45	0.75
0.6	61.1%	0.42	0.72	61.5%	0.41	0.71

proper β value is selected, the VPRS model can perform better than classic model.

6 Comparison between VPRS and NN

The VPRS method is also compared with the BP neural network method. A standard three-layer BP neural network is used as a benchmark. There are 27 nodes in the input layer: "I^{t-3}, S_F^{t-3}, S_W^{t-3}, L_T^{t-3}, W_T^{t-3}, A^{t-3}, W_B^{t-3}, I^{t-2}, S_F^{t-2}, S_W^{t-2}, L_T^{t-2}, W_T^{t-2}, A^{t-2}, W_B^{t-2}, I^{t-1}, S_F^{t-1}, S_W^{t-1}, L_T^{t-1}, W_T^{t-1}, A^{t-1}, W_B^{t-1}, I^t, S_F^t, S_W^t, L_T^t, W_T^t, A^t", where the superscript is sampling time. And there is one node (W_B^t) in the output layer. The number of hidden nodes and the learning rate are determined based on the validation set. The hidden nodes use the 'tansig' transfer function and the output node uses the 'logsig' transfer function. The 'Levenberg-Marquardt' method is used for training the BP neural network, which could give better performance than the batch training for large and non-stationary data sets. Standard Matlab toolboxes are used for the calculations of the BP neural network. For VPRS β is equal to 0.8, entropy based method discretization method is used for continuous condition attributes, and equal width discretization is used for decision attributes. When reasoning, the most matched 5 rules were selected where the minimal rule match rate is '0.7'.

To get a stable comparison result, ten fold cross validation method is used for testing, and the result is shown in Table 7, where the predict value is close to expectation. Part validation result is shown in Fig. 3. Comparison result between VPRS and BP network is shown in Table 8. It is clear that the VPRS model's predictive ability is close to that of NN while the complexity of VPRS is much lower than that of NN.

Table 7 Predicting result of soft steel multi-variable VPRS model

Expected value/mm	Predicted value/mm					
	3	4	5	6	7	Not matched
3	39	12	4	1	0	5
4	10	20	25	4	0	4
5	6	29	98	74	2	10
6	0	5	99	366	49	13
7	0	1	3	40	92	2

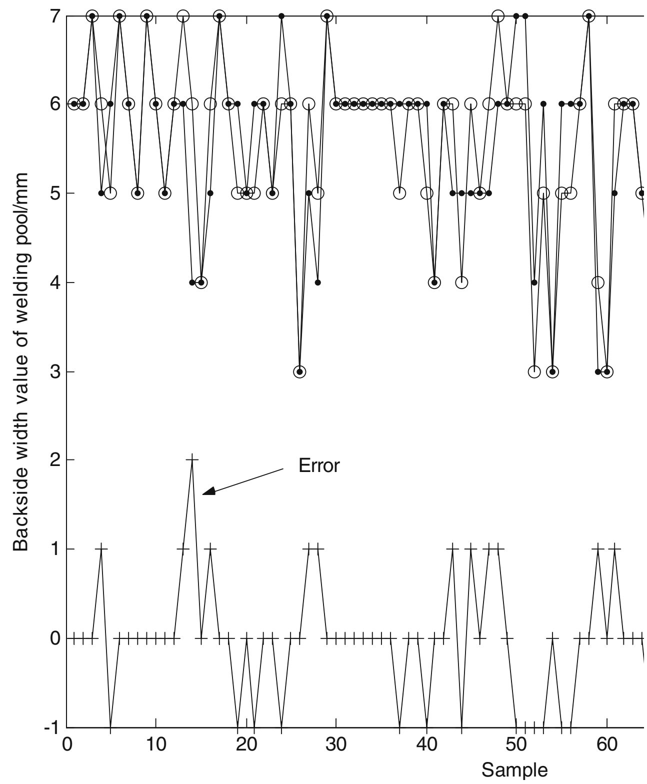

Fig. 3 Part validation result of multi-variables VPRS model of mild steel

7 Conclusions

Modeling is an important aspect for the design of a controller. For complex systems like welding process, a precise mathematical model, which is usually critical in traditional control, is hard to obtain. Intelligent controls have been given more attention because of their high adaptability for complex systems and their ability to control knowledge essentially. Accordingly, building an effective knowledge model is important. In contrast to other intelligent modeling methods the VPRS modeling method can acquire knowledge from experimental data more effectively, has good understandability while also improving its noisy suppression capability.

Table 8 Comparison between VPRS and NN

	VPRS	NN
R_{right}	62.1%	
E_{mean}	0.77	0.59
E_{sd}	0.44	0.45

In this paper, the VPRS method was introduced and applied to the welding process modeling. It was found to be able to discover knowledge in welding field with stable performance, and exhibit better noisy suppression capability than the classic RS model. The method shows similar precision on prediction when compared with BP or NN model, but is much simpler to use.

The VPRS modeling method can be applied for other similar complex systems to obtain knowledge models if there are enough effective raw data. Besides its application in modeling, it could be used in data mining or knowledge acquisition, and it is useful to experts system when building the expert knowledge base.

Acknowledgement The author greatly appreciate the support for this work from the National Natural Science Foundation (Grant No. 60474036.) The authors also like thank Dr. H. Zheng and E. Poli at University of California at Davis for preparing the manuscript.

References

1. Chen SB (2000) Intelligent methodology for sensing, modeling and control of pulsed GTAW: Part I - Bead-on-plate welding. Weld J 79(6):151S–163S
2. Wang B, Chen SB, Wang JJ (2005) Rough set based knowledge modeling for the aluminum alloy pulsed GTAW process. Int J Adv Manuf Technol 25(9–10):902–908
3. Di L (2001) "Neural-network-based self-organized fuzzy logic control for arc welding". Eng Appl Artif Intell 14(2):115–124
4. Zhang GJ, Chen SB, Wu L (2005) Neural network modeling and system simulating for the dynamic process of varied gap pulsed GTAW with wire filler. J Mater Sci Technol 21(4):515–520
5. Christensen KH, Sorensen T (2003) A neural network approach for GMA butt joint welding. Intern J Join Mater 15(4):1–7
6. Fernandes PE (2002) Artificial neural networks applied to the tubular electrode arc welding process: modeling of the weld bead geometry. Metal Mater ABM 58:9–13
7. Hirai A (2003) Neural network model of weld pool in pulsed MIG welding. Quarterly J Japan Weld Soc 21(2):227–233
8. Ziarko W (1993) Variable precision rough set model. J Comput Syst Sci 46(1):39–59
9. Pawlak Z (1991) Rough sets-Theoretical aspects of reasoning about data. Kluwer, Dordrecht
10. Beynon MJ, Peel MJ (2001) Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega-International J Manage Sci 29(6):561–576
11. An AJ (1996) Discovering rules for water demand prediction: An enhanced rough-set approach. Eng Appl Artif Intell 9(6): 645–653
12. Wong SKM, Ziarko W (1985) "On optimal rules in decision tables". Bull Pol Acad Sci, Math 33:693–696
13. Skowron A, Rauszer C (1992) The discernibility matrices and functions in information systems. Intelligent decision support-handbook of application and advances of the Rough Sets Theory. Kluwer, Dordrecht
14. Wang B, Chen S (2002) Reduction and minimal set cover. J Shanghai Jiaotong Univ 36:106–108
15. Chen SB (2004) Welding robotic systems with visual sensing and real-time control of dynamic weld pool during pulsed GTAW. Int J Robot Autom 19(1):28–35
16. Fayyad UM, Irani KB (1993) Multi-interval discretization of continuous-valued attributes for classification learning. Proc Thirteenth International Joint Conference on Artificial Intelligence 2
17. Li WH, Chen SB, Lin T (2006) The comparison of discretization method in rough set based modeling method for welding. J Shanghai Jiaotong Univ 40(7):1094–1097