Information Entropy Measurement In Uncertainty Analysis of Rough Control Rules For Vision-based Navigation of Outdoor Mobile Robot

Gui-Sheng Chen, Bing Li , Ke-Zhong He and Pei-Fa Jia
Department of Computer Science and Technology, Tsinghua University
State Key Lab of Intelligent Technology and System, Beijing, China , 100084
Phone:+86-10-6278 1770 Email: cgs00@mails.tsinghua.edu.cn

1. Introduction

Tsinghua Mobile Robot (THMR-5) is designed as an outdoor mobile robot with the function of both autonomous navigation and remote operation, and is shown in figure 1.Now the emphasis of research includes smooth and stable path tracking based on machine vision for lane detection on high way .One of the relative issues is about the intelligent decision of driving heading. In this paper, we approach to acquire driving rule set from human driving experience (data sample) through rough set theory, this is quite different from fuzzy inferring system, in which the membership function extremely depends upon subjective ideas of experts. Especially the uncertainty analysis of rough control rules based on information entropy will be discussed firstly.

Figure 1. the fifth generation of outdoor mobile robot THMR-V

Rough set theory , which is set up by Pawlak Z. at the early 1980s, is an effective mathematic tool to deal with the uncertain or fuzzy knowledge [1][2].Its main idea includes acquisition of rough decision rules by reduction of data under the condition of keeping an invariability of classification capacity. And it basically represents a concept by a pair of approximation sets according to the partition on the given data based on the indiscernibility relation. In the applications about classification, the rough classifying rules can then be generated from these approximations and used to classify the new cases. Rough sets theory is well suited to deal with inconsistency of the rough approximations..

We assume that $T=(U,C,D,V_a,f_a)|_{a\in C\cup D}$ is a decision table, where $C\bigcap D=\phi, C\bigcup D=A$, C is condition attribute set, D is decision attribute set. Let X_i represents an equivalence class under the indiscernibility relation of condition attribute C,and Y_j represents an equivalence class under the indiscernibility relation of decision attribute D, $M_i=\{Y_j : X_i\bigcap Y_j\neq\phi\}$ stands for all the sets in the decision classes, where $M_i\neq\phi$, $1\le|M_i|\le t$,then we may obtain rough rule as followed:

IF Des($x\in X_i$) THEN Des($\vee_{1\le k\le m}(x\in Y_{jk})$),

where Des($x\in X_i$) means the language description of $x\in X_i$ .and it is proved that the C Upper approximation of D, which is

$$\sum_{j\in U/Ind(D)}\bar{C}Y_j ,$$

may generates all possible rules. In general, the accuracy of rule $X_i\to Y_j$ is defined as $\alpha_{i,j}=|X_i\bigcap Y_j|/|X_i|$,which represents the inconsistency of rule set. The coverage of rule $X_i\to Y_j$ is defined as $\lambda_{i,j}=|X_i\bigcap Y_j|/|Y_i|$ .and $\lambda_{i,j}$ represents the ratio of the cardinal number of intersect of X_i and Y_j over that of the decision attribute equivalent Y_j .it means the randomness of rule set..

The uncertainty coming up from the granularity of the indiscernibility-based partition includes not only inconsistency, but also randomness, we need to construct another criterion to evaluate both the two aspects of uncertainty. In the following discussion, we will define an information entropy based uncertainty measure ,which may deal with the two aspects of the uncertainty

2. Information Entropy Criterion In Uncertainty Analysis of Rough Rule Set

An indiscernibility relation on U corresponds with a random variable ,which has the value relative with the equivalent classes $\{X_1,\ldots,X_k\}$ by partition ,and the information entropy corresponding to partition *P(R)* is:

$$H(R)=\sum_{i=1}^{k}\frac{|X_i|}{|U|}\log_2\frac{|U|}{|X_i|}$$

In general, for two indiscernibility relations R_1 and R_2,If $P(R_1)\Rightarrow P(R_2)$,then $H(R_1)\geq H(R_2)$.It means that the less the granularity of partition is, the larger the entropy we obtain. so, *H(R)* may be taken to measure the uncertainty coming from the partition granularity.

Then, we may construct a certain kind of partition, and then evaluate the uncertainty by calculating the corresponding entropy of the rule set obtained., there are many methods to construct a partition of U and the corresponding entropy criterions, such as local attribute based partition H^{loc} and positive region based partition $H^{\det}$ [3] [4].The key problem is how to deal with the boundary of rule set. $H^{\det}$ takes the positive region and the boundary region as completely different parts, and then over-evaluate the weakness of inconsistency. and H^{loc} takes no account of the difference between positive region and boundary region, so it dos not care for the problem of consistency. Here we define a particular part ,which is called β near boundary of the normal boundary region, referring to the conception of "edge softening" in image processing. It has a certain tolerance for the inconsistency of the data so that it can be used in the noisy cases to protect some useful rules with a little degrees of inconsistency.

Let $C(X_i,Y_j)$ represents the Majority Inclusion Relation :

$$C(Xi,Yj)=1-{|X_i\cap Y_j|}\big/{|X_i|}$$

then: $V_1=\bigcup\{X_i\cap Y_j:0<C(X_i,Y_j)<\beta\}$ is defined as the β near boundary of rule set $B\rightarrow d$.

Let V_0,V_1 is the positive region and near boundary region of rule set $B\rightarrow d$ respectively,

$$P(R^\beta)=\{X_i:X_i\in V_0\}\bigcup\{X_i\cap V_1:X_i\cap V\neq\phi)\}$$

$$\bigcup\{\{x\}:x\in U\setminus(V_0+V_1)\}$$

is called a partition of U based on V_0 and V_1.

So, $P(R^\beta)$ may essentially divided into three parts :

1.the positive region V_0;

2.the intersect of V_0 and near boundary V_1;

3.$U\setminus(V_0+V_1)$,where every one datum object construct a single equivalent class $\{x\}$;

and the probability distribution of every equivalent class is:

$$\rho_i=\begin{cases}{|Xi|}\big/{|U|},\ \text{for}\ X_i\in V_0\\ {|Xi\cap V_1|}\big/{|U|},\ for\ X_i\cap V_1\neq\phi\\ {1}\big/{|U|},\ \text{the others}\end{cases}$$

and the information entropy evaluation of rule set $B\rightarrow d$ in accordance with partition $P(R^\beta)$ of *U* is:

$$H^\beta\ (B\rightarrow d)=H(R^\beta)=\sum_i\rho_i\log_2\frac{1}{\rho_i}=$$

$$\sum_{X_i\in V_0}\frac{|Xi|}{|U|}\log_2\frac{|U|}{|X_i|}+\frac{|U\setminus(V_0+V_1)|}{|U|}\log_2|U|+$$

$$\sum_{X_i\cap V_1\neq\phi}\frac{|X_i\cap V_1|}{|U|}\log_2\frac{|U|}{|X_i\cap V_1|}$$

where

$\sum_{X_i\in V_0}\frac{|X_i|}{|U|}\log_2\frac{|U|}{|X_i|}=H^{con}$ represents the evaluation of consistency of rules coming from the positive region.

$\sum_{X_i\cap V_1\neq\phi}\frac{|X_i\cap V_1|}{|U|}\log_2\frac{|U|}{|X_i\cap V_1|}=H^{ncon}$ represents the evaluation of weak-consistency of rules coming from the near boundary region.

$\frac{|U\setminus(V_0+V_1)|}{|U|}\log_2|U| = H^{inc}$ represents the evaluation of inconsistency of rules .

So, for the rule set $B \rightarrow d$,the rules in positive region satisfy: $X_i \cap Y_j \neq \phi$ and $X_i \subseteq Y_j$,and rules in the near boundary satisfy: $X_i \cap Y_j \neq \phi$ and $C(X_i, Y_j) < \beta$ .Rules in positive region are consistent rules(here the coverage of rules is not taken into account),and rules in near boundary have weak inconsistency, which represents the tolerance of rule set to the data noise. Here it is particularly noted that the cardinality of near boundary is that of the intersect of X_i and Y_j , not of X_i itself, where $X_i \cap Y_j \neq \phi$ and $C(X_i, Y_j) < \beta$ .In other words, even for those weak consistent rules, the weakness of inconsistency is not the same. Consequently, the β near boundary may reach these two needs.

3. Rule Acquisition For Vision-based Navigation

The basic procedure of rough decision making for vision-based lane tracking of outdoor mobile robot includes:

1. lane check and environment understanding based on machine vision as shown in figure 2,figure 3 and figure 4.[6]

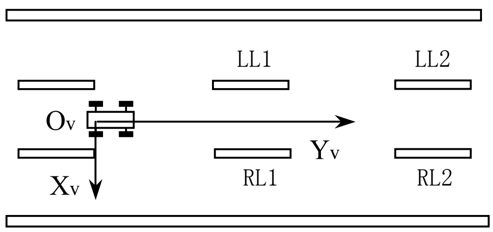

Figure 2 .the vehicle coordination

Figure 3. the image coordination

2. recording the driving parameters such as the heading bias Φ ,distance offset d and the driving command parameter from human experience as shown in

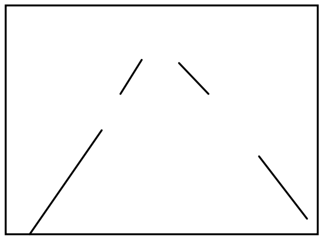

Figure 4. lane detect output

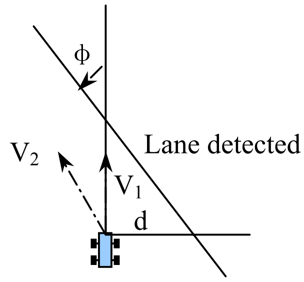

Figure 5. driving parameters

figure 5.

3. construction of the decision table of driving knowledge as shown in table 1.

Table 1: Driving control data sample from human driving experience of THMR-5 (vehicle-width=1.8 meter)

x	Heading Error	Distance Error	Steering Command
1	-52.6	3.97	-1
2	-51.7	3.66	1
3	-58.9	5.32	-2
4	-56.8	5.35	-2
...	......	......	......

we obtained a giant data sample of about 50000 records which represents the driving knowledge of as many typical cases as possible,

4.determination of the discrete interval of all the continuous attributes by method of Genetic Algorithms (GA), in which the fitness function is just the information entropy abovementioned and the chromosome is configured by the cutting points and coded by 0-1 string. In reality, we may obtain different entropy of partition for the corresponding discrete cutting points set. Consequently, we may finally obtain an optimized selection of discrete intervals.

5.generation of rough rule set. Finally we may obtain an optimized rough rule set, which has the least information entropy, according to the corresponding C upper-approximation of decision attribute D as described in section 1.

It is worth noting that the proposed procedure may need more computation time due to complicated process. Nevertheless, as shown in later experiment, the proposed algorithm can have better performance than other approaches do, Since the process is offline, the computation time is not so important.

4. Result of Experiment

In the experiment, we first deal with the recording data sample according to rough set theory, and delete those redundant or inconsistent records. Initially we set the cutting points set with 20 points of the condition attributes, distance error and heading error, and decision attributes of driving command. Also we initially set the tolerance coefficient beta=0.2.Through genetic iteration operation, we finally obtain 7 discrete intervals of driving command parameter, 7 discrete intervals of distance bias parameter and 9 discrete intervals of heading offset parameter. Meanwhile, the algorithm generates 18 consistent rules and 9 "weak consistent" rules, This rule set efficiently represents the human driving experience. Also the program generates a few inconsistent rules, which results from the data noise.

5. Conclusion

Rough set theory is an efficient tool to acquire decision rule from giant data sample. It plays an important role in dealing with uncertain knowledge of Artificial Intelligence. Information entropy based uncertainty analysis is more reasonable, objective and precise. It can protect "weak consistent rules" at certain accuracy by adjustment of β , So, this method avoids the weaknesses of overestimating the inconsistency caused by the noise and improperly rejecting some useful weak deterministic rules. Of course, in order to get more satisfactory results in some real decision tasks, both the priori knowledge of the experts and the biases of the decision makers need to be represented in the evaluation criteria. This will be considered in our future works.

References

1.Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning About Data [M]. Dordrecht: Kluwer Academic Publishers

2.Pawlak Z, Grzymala-BusseJ, SlowinskiR,et al.Rough Sets [J]. Communications of the ACM, 1995, 38(11): 88-95.

3.Düntsch I , Gediga G. Uncertainty measures of rough set prediction [J]. Artificial Intelligence, 1998, 106: 109-137.

4.Ziarko W. Variable precision rough set model [J]. Journal of Computer and System Sciences, 1993, 46: 39-59.

5.Yao Y Y, T.Y.Lin. Generalization of Rough Sets Using Modal Logics. intelligent Automation and soft computing, 1996 ,2(2), 103-120

6.Pengfei Zhang, Research on Visual Navigation Techniques of Outdoor High-speed Mobile Robot, PhD dissertation, Tsinghua University,Beijing,2002(In Chinese)