Available online at www.sciencedirect.com

ScienceDirect

Journal of Computational Design and Engineering 2 (2015) 137–147

www.elsevier.com/locate/jcde

Hybrid Type II fuzzy system & data mining approach for surface finish

Tzu-Liang (Bill) Tseng[a], Fuhua Jiang[b], Yongjin (James) Kwon[c,*]

[a]Department of Industrial, Manufacturing and Systems Engineering, The University of Texas at El Paso, El Paso, TX 79968, USA
[b]Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA
[c]Department of Industrial Engineering, Ajou University, Suwon 443-749, Republic of Korea

Received 28 November 2014; received in revised form 21 February 2015; accepted 22 February 2015
Available online 5 March 2015

Abstract

In this study, a new methodology in predicting a system output has been investigated by applying a data mining technique and a hybrid type II fuzzy system in CNC turning operations. The purpose was to generate a supplemental control function under the dynamic machining environment, where unforeseeable changes may occur frequently. Two different types of membership functions were developed for the fuzzy logic systems and also by combining the two types, a hybrid system was generated. Genetic algorithm was used for fuzzy adaptation in the control system. Fuzzy rules are automatically modified in the process of genetic algorithm training. The computational results showed that the hybrid system with a genetic adaptation generated a far better accuracy. The hybrid fuzzy system with genetic algorithm training demonstrated more effective prediction capability and a strong potential for the implementation into existing control functions.

Keywords: Data mining; Fuzzy set theory; Surface roughness; Metal cutting; Machining quality characteristics

1. Introduction

In this study, a data mining technique (i.e., a new heuristic algorithm) for reduct selection in the Rough Set Theory (RST) is applied to select significant factors (features). Literature review suggests that the RST has not been widely applied to metal cutting problems thus making this research novel [1,2]. In the RST, features characterize each object, and it discovers the dependencies between them. Compared to the usual statistical tools that use population-based approach, the RST uses an individual, object-model based approach that makes a very good tool for analyzing quality control problems [3]. The RST is also able to identify "defective" and "significant factors" simultaneously, which is unique and useful in solving quality control problems. After significant factors are identified, a Fuzzy Logic Theory (FLT) is used to construct the approach that adapts and predicts the surface finish because of the following reasons: (1) using a FLT enables fast and easy synthesis and modification of the control rule base; (2) if a rapid adaptation, using only a few data points, with good accuracy is obtained, the process can respond to the changes; and (3) adaptation is more suitable for today's machining environment because the adaptive approach can be integrated into the CNC controller to compensate for process variations. The applied Fuzzy Logic System (FLS) is a combined system of Type I and Type II. Different types of system express different strengths to handle heterogeneous factors as well as variables in the process. For example, Type I is effective to deal with "crisp" type of membership function, while Type II is adequate to handle "uncertain" type of membership. The Type II FLS has not been widely used to solve machining process problems thus making this study unique. Finally, the Genetic Algorithm (GA) is incorporated to the FLS for fuzzy adaptation. The combination of the unique strength in each domain is expected to provide a better solution space.

In current practice, setting machining parameters are usually conducted by the experience of skilled engineers. Once set, the parameters are usually unchanged during machining, unless prominent anomalies are present. The proposed scheme can be incorporated into the intelligent CNC controllers, and used as constant monitoring device as the machining

*Corresponding author. Tel.: +82 31 219 2418; fax: +82 31 219 1610.
E-mail address: yk73@ajou.ac.kr (Y. Kwon).

http://dx.doi.org/10.1016/j.jcde.2015.02.002
2288-4300/

operations are carried out. Such practice can significantly improve the machining efficiency as well as the perceived quality of the machined parts. It is expected that the proposed approach will help compensate for the unforeseeable variations in the machining process, hence ultimately affects the overall quality of the CNC machining operations.

2. Literature survey

2.1. Fundamental of data mining and rough set theory

Data mining is the process of extracting and refining knowledge from large database [4–6]. The extracted information is used to predict, classify, model, and summarize the data being analyzed. The RST is a fundamental theory of data mining. This theory was originated by Pawlak [7] and was developed to classify imprecise, uncertain, or incomplete information or knowledge expressed in terms of data acquired from experience. Therefore, it complements the FST [8]. The rough set approach is suitable for processing qualitative information that is difficult to analyze by standard statistical techniques [9]. It integrates learning-from-example techniques, extracts rules from a data set of interest, and discovers data regularities [10]. The RST has been applied to address a variety of problems [11], including (1) representation of uncertain or imprecise knowledge; (2) empirical learning and knowledge acquisition from experience; (3) knowledge analysis; (4) analysis of conflicting; (5) evaluation of the quality of the available information with respect to its consistency and the presence or absence of repetitive data patterns; (6) identification and evaluation of data dependencies; and (7) approximate pattern classification. The RST is introduced as an extension of set theory for the study of intelligent systems characterized by using incomplete information to classify imprecise, uncertain, or incomplete information or knowledge expressed in terms of data. Indeed, the RST is an effective tool for multi-attribute classification problems. In RST, data is expressed in a decision table in which each row represents an object and each column represents an attribute. Formally, the decision table is represented by an information function [12]:

$$S = \langle U, Q, V, f \rangle \tag{1}$$

where U is a finite set of objects, Q is a finite set of attributes, $V = \bigcup_{q \in Q} V_q$ QUOTE and V_q is a domain of the attribute q, and $f: U \times Q \to V$ is the total decision function such that $f(x,q) \in V_q$ for every $q \in Q$, $x \in U$ QUOTE . The main theme of RST is concerned with measuring what may be described as the "ambiguity" inherent in the data. The essential distinction is made between objects that may definitely be classified into a certain category, and those that may possibly be classified. Considering all decision classifications yields to what is referred to as the "quality of approximation" that measures the proportion of all objects from which definite classification may be achieved. A rough set can be described as a collection of objects that in general cannot be precisely characterized in terms of their values or sets of attributes, but can be characterized in the form of lower or upper approximations [13,14]. The upper approximation includes all objects that possibly belong to the concept, while the lower approximation contains all objects that definitely belong to the concept. As each object is characterized with attributes, discovering the dependencies between attributes and detecting the main attributes is of primary importance. Attribute reduction is one unique aspect of the rough set approach. A reduct is a minimal sufficient subset of attributes, which provides the same quality of discriminating concepts as the original set of attributes.

Let us consider the five objects in Table 1, each with four input features and an output feature (outcome). To derive the reduct, consider the first feature $F1$. The set of objects corresponding to the feature value F1 = 0 is {1, 2, 3, 5}. This set {1, 2, 3, 5} cannot be further classified solely using the relation F1 = 0. It is discernible over the constraint F1 = 0, which is expressed as $[x][\mathrm{F1}=0]=\{1, 2, 3, 5\}$. For the objects in set {1, 5}, the output feature is $O = 2$. For object 3, the output feature is $O = 1$ and for object 2, the output feature is $O = 0$. Therefore, additional features are needed to differentiate between $O = 0$, 1, or 2. Applying this concept, the classification power of each feature can be evaluated. For instance, the feature value F1 = 1 is specific to $O = 1$. This discernible relation can be extended to multiple features, e.g., $[x][\mathrm{F1}=0] \wedge [\mathrm{F2}=1]=\{1, 3\}$ and $[x][\mathrm{F1}=0] \vee [\mathrm{F2}=1]=\{1, 2, 3, 5\}$, where $\wedge$ and $\vee$ refers to "or" and "and", respectively.

2.1.1. Reduct generation

Most of the rough set based approaches may generate more than one reduct for an object. This paper adapts the reduct generation procedure proposed by Pawlak [12] and presents it in the form of the reduct generation procedure as illustrated in Fig. 1. The reduct generation procedure enumerates all

Table 1
Example data set.

Object no.	*F1*	*F2*	*F3*	*F4*	*O*
1	0	1	0	2	2
2	0	0	1	3	0
3	0	1	1	1	1
4	1	2	2	0	1

O: Not Applicable, 1: Low, 2: Medium, 3: High.

Step 0: Set object number i = 1;
Step 1: Select object i and find a set of reducts with one to m - 1 features;
Step 2: Set i = i +1. If all objects have been considered, go to Step 3; otherwise go to Step 1;
Step 3: Terminate the algorithm and output the result.

Fig. 1. Reduct generation procedure.