Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Robotics and Computer-Integrated Manufacturing 21 (2005) 559–567

Robotics and Computer-Integrated Manufacturing

www.elsevier.com/locate/rcim

Feature-based rule induction in machining operation using rough set theory for quality assurance

Tzu-Liang (Bill) Tseng[a,*], Yongjin Kwon[b], Yalcin M. Ertekin[c]

[a]Department of Mechanical and Industrial Engineering, The University of Texas at El Paso, El Paso, TX 79968, USA
[b]Applied Engineering Technology, Goodwin College of Professional Studies, Drexel University, Philadelphia, PA 19104, USA
[c]Department of Engineering Technology, Tri-State University, Angola, IN 46703, USA

Received 1 December 2004

Abstract

There have been many studies, mainly by the use of statistical modeling techniques, as to predicting quality characteristics in machining operations where a large number of process variables need to be considered. In conventional metal removal processes, however, an exact prediction of surface roughness is not possible or very difficult to achieve, due to the stochastic nature of machining processes. In this paper, a novel approach is proposed to solve the quality assurance problem in predicting the acceptance of computer numerical control (CNC) machined parts, rather than focusing on the prediction of precise surface roughness values. One of the data mining techniques, called rough set theory, is applied to derive rules for the process variables that contribute to the surface roughness. The proposed rule-composing algorithm and rule-validation procedure have been tested with the historical data the company has collected over the years. The results indicate a higher accuracy over the statistical approaches in terms of predicting acceptance level of surface roughness.

Keywords: Quality assurance; Data mining; Rough set theory; CNC machining; Surface roughness

1. Introduction

Traditionally, the quality of a machined product has been measured based on the specifications, once the machining process has been completed. However, this post-process inspection has several shortcomings: (1) the manufacturing cost has already been incurred when a non-conformance is detected, (2) it is difficult to isolate the causes of the defect, (3) there could be a significant time lag between discovery of the defect and corrective action, and (4) rework of any scope increases manufacturing cost and can be very difficult to accomplish. Today, efforts of manufacturers are shifting from the post-process inspection to improved monitoring of the manufacturing process, utilizing sensors and other measurement devices to effectively control the process. Improvements in machining precision can only be accomplished by the development of manufacturing systems that are capable of monitoring processes. Process monitoring reduces scrap, rework, lead-time, and conventional non-value-added inspection activities, and thereby increases the system's productivity. However, monitoring has to be based on sound, reliable process control algorithms. On the human side, an expert machinist can detect and/or predict deteriorating cutting conditions through the use of his/her senses so that appropriate corrective actions can be taken before the part quality is lost. This level of expertise requires many years of experience, yet does not guarantee that conditions will be monitored in effective and consistent manner, even for the most experienced operator. On the mechanical side, computer numerical control (CNC) of machine tools can produce consistent part quality, yet most cases, do not utilize sensor data to compensate for

*Corresponding author. Tel.: +1 915 747 7990; fax: +1 915 747 5019.
E-mail address: btseng@utep.edu (T.-L. (Bill) Tseng).

0736-5845/$ - see front matter
doi:10.1016/j.rcim.2005.01.001

anomalies generated by the cutting process (e.g., tool wear, chatter, incorrect machine setup, etc.). If sensors such as cutting force, vibration and spindle motor current were integrated into CNC machine tools, the control functions should be able to interpret and respond to sensory data as process continues. However, when many process variables need to be considered, it is rather difficult to predict quality attribute in machining (i.e., surface roughness).

This study, unlike the conventional statistical regression modeling approach, uses a data mining technique called the rough set theory (RST) [1] to identify variables affecting the quality characteristic of CNC machining operations. Instead of predicting exact surface roughness values, the focus is on the prediction of quality acceptance in machined parts. RST is a viable approach for extracting meaningful knowledge and making predictions for an individual data object, rather than a population of objects [2]. RST is introduced as an extension of the set theory for the study of intelligent systems characterized by incomplete information to classify imprecise, uncertain, or incomplete information or knowledge expressed in terms of data. RST is an effective tool for multi-attribute classification problems. This can be instrumental in constructing an intelligent control system, especially when a clear delineation within variables as to how they affect the surface roughness is difficult to achieve.

2. Theoretical background of data mining and rough set theory

Data mining is a process of extracting and refining knowledge from a large database [3–5]. The extracted information can be used to predict, classify, model, and characterize the data being mined. RST is a fundamental theory of data mining. This theory was originated by [1] and was developed to classify imprecise, uncertain, or incomplete information or knowledge expressed in terms of data acquired from experience or historical facts. Therefore, it complements the fuzzy set theory [6]. The rough set approach is also suitable for processing qualitative information that is difficult to analyze by standard statistical techniques [7]. It helps integrate learning-from-example techniques, extract rules from a data set of interest, and discover data regularities [8].

RST has been applied to address a variety of problems [9], including (1) representation of uncertain or imprecise knowledge, (2) empirical learning and knowledge acquisition from experience, (3) knowledge analysis, (4) analysis of conflicting data, (5) quality evaluation of the available information with respect to its consistency and presence or absence of repetitive data patterns, (6) identification and evaluation of data dependencies, and (7) approximation of pattern classification. In RST, data are 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 [10] in the form of

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

where U is a finite set of objects, Q is finite set of attributes, $V = \bigcup_{q \in Q} V_q$ and V_q is domain of the attribute q, and $f : UxQ \to V$ is the total decision function such that $f(x, q) \in V_q$ for every $q \in Q$, $x \in U$.

The main theme of RST is concerned with measuring what may be described as the "ambiguity" inherent in the data. In RST, the essential distinction is made between objects that may definitely be classified to 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 for 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 of their sets of attributes, but can be characterized in terms of lower or upper approximations. 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 dependencies between attributes and detecting main attributes is of primary importance in RST. 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. For example, let us consider the five objects in Table 1, each with four input features ($F1$–$F4$) and an output feature.

To derive reducts, 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 $[x][F1 = 0] = \{1, 2, 3, 5\}$. For the objects in set {1, 5}, the output feature is $O = 2$, for object 3, the

Table 1
Example data set

Object no.	Features				Output
	$F1$	$F2$	$F3$	$F4$	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
5	0	0	0	1	2

0: Not applicable, 1: low, 2: medium, 3: high.