9-2015

A survey on artificial intelligence-based modeling techniques for high speed milling processes

Amin Jahromi TORABI

Meng Joo ER

Xiang LI

Beng Siong LIM

Lianyin ZHAI

See next page for additional authors

DOI: https://doi.org/10.1109/JSYST.2013.2282479

Citation

TORABI, Amin Jahromi; ER, Meng Joo; LI, Xiang; LIM, Beng Siong; ZHAI, Lianyin; OENTARYO, Richard Jayadi; PEEN, Gan Oon; and ZURADA, Jacek M.. A survey on artificial intelligence-based modeling techniques for high speed milling processes. (2015). *IEEE Systems Journal*. 9, (3), 1069-1080. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/3248

Author
Amin Jahromi TORABI, Meng Joo ER, Xiang LI, Beng Siong LIM, Lianyin ZHAI, Richard Jayadi OENTARYO, Gan Oon PEEN, and Jacek M. ZURADA

Published in IEEE Systems Journal, 2015 September, Volume 9, Issue 3, Pages 1069-1080
https://doi.org/10.1109/JSYST.2013.2282479

A Survey on Artificial Intelligence-Based Modeling Techniques for High Speed Milling Processes

Amin Jahromi Torabi, Meng Joo Er, *Senior Member, IEEE*, Xiang Li, Beng Siong Lim, Lianyin Zhai, Richard J. Oentaryo, Gan Oon Peen, and Jacek M. Zurada, *Fellow, IEEE*

***Abstract*—The process of high speed milling is regarded as one of the most sophisticated and complicated manufacturing operations. In the past four decades, many investigations have been conducted on this process, aiming to better understand its nature and improve the surface quality of the products as well as extending tool life. To achieve these goals, it is necessary to form a general descriptive reference model of the milling process using experimental data, thermomechanical analysis, statistical or artificial intelligence (AI) models. Moreover, increasing demands for more efficient milling processes, qualified surface finishing, and modeling techniques have propelled the development of more effective modeling methods and approaches. In this paper, an extensive literature survey of the state-of-the-art modeling techniques of milling processes will be carried out, more specifically of recent advances and applications of AI-based modeling techniques. The comparative study of the available methods as well as the suitability of each method for corresponding types of experiments will be presented. In addition, the weaknesses of each method as well as open research challenges will be presented. Therefore, a comprehensive comparison of recent developments in the field will be a guideline for choosing the most suitable modeling technique for this process regarding its goals, conditions, and specifications.**

***Index Terms*—Artificial intelligence (AI), high speed machining (HSM), milling process, modeling techniques.**

Nomenclature

AISI	American Iron and Steel Institute.
AFPN	Adaptive fuzzy Petri net.
ANFIS	Adaptive neuro-fuzzy inference system.
ANN	Artificial neural network.
BCCD	Best cutting condition determination.
BN	Bayesian network.
BP	Back propagation.
CFFBP	Cascaded feedforward back propagation.
Dc	Depth of cut.
DoE	Design of experiment.
DEVS	Discrete event systems.
DWT	Discrete wavelet transform.
ESEM	Environment scanning electron microscopy.
Fc	Feed rate.
Fz	Feed per tooth.
FFT	Fast Fourier transform.
FL	Fuzzy logic.
FN	Fuzzy net.
FN-ASRC	Fuzzy-net adaptive surface roughness control.
FN-IPSRR	Fuzzy-net in-process surface roughness recognition.
FPN	Fuzzy Petri net.
EM	Expectation maximization.
GA	Genetic algorithm.
GONN	Genetic algorithm-optimized neural network.
GP	Genetic programming.
HMM	Hidden Markov model.
HTGLA	Hybrid Taguchi–genetic learning algorithm.
ISO	International Organization for Standardization.
LDA	Linear discriminant analysis.
MAE	Mean absolute error.
MAPE	Mean absolute percentage error.
MLP	Multilayer perceptron.
MSE	Mean square error.
NA	Not applicable.
NM	Not mentioned.
NN	Neural network.
OA	Orthogonal array.
PSO	Particle swarm optimization.
PSONN	Particle swarm optimized neural network.
PVD	Physical vapor deposition.
Ra	Roughness profile, arithmetic average.
Rq	Roughness profile, root mean squared.
Rv	Roughness profile, maximum valley depth.
Rp	Roughness profile, maximum peak height.
Rt	Roughness profile, maximum height of the profile.
RBF	Radial basis function.
SVM	Support vector machine.
SVR	Support vector regression.
TAN	Tree augmented naive.
TDBP	Time-delay back propagation.
Vc	Cutting speed.
VMC	Vertical machining center.
VQ	Vector quantization.

Manuscript received April 27, 2013; accepted August 31, 2013. Date of publication November 13, 2013; date of current version June 18, 2015. The project on modeling the HSM processes was supported by A*Star, Singapore.

A. J. Torabi, M. J. Er, and L. Zhai are with the Nanyang Technological University, Singapore 639798 (e-mail: Amin0005@ntu.edu.sg; Emjer@ntu.edu.sg; Lyzhai@ntu.edu.sg).

R. J. Oentaryo is with Living Analytics Research Centre, Singapore Management University, Singapore 178902 (e-mail: oentaryorj@ntu.edu.sg).

X. Li, B. S. Lim, and G. O. Peen are with the Singapore Institute of Manufacturing Technology, Singapore 638075 (e-mail: Xli@SIMTech.A-star.edu.sg; Bslim@SIMTech.A-star.edu.sg; Opgan@SIMTech.A-star.edu.sg).

J. M. Zurada is with the University of Louisville, Louisville, KY 40292 USA, and also with the Spoleczna Akademia Nauk, 90-011 Lodz, Poland (e-mail: Jmzura02@louisville.edu; j.zurada@ieee.org).

Digital Object Identifier 10.1109/JSYST.2013.2282479

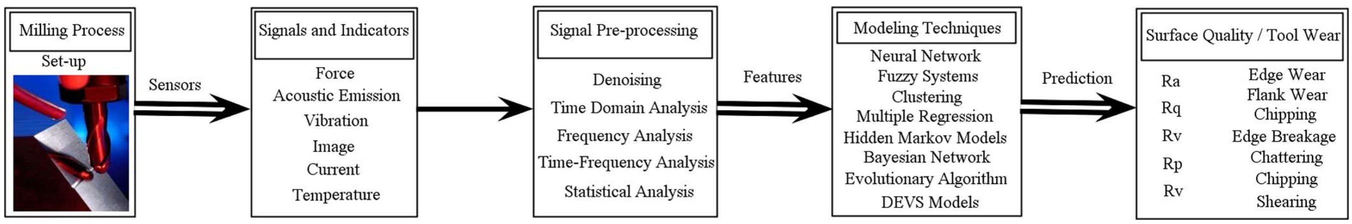


Fig. 1. Tool condition monitoring and surface roughness prediction [7].

I. Introduction

TODAY, high speed machining (HSM) is widely applied to fulfill the overwhelming and increasing demands for producing vital pieces for various industrial sectors, particularly in aerospace industries. The throughput of the machining process is a critical parameter for determining the quality of a production process. Large throughput, as well as the surface quality of the product, is directly related to the change in the total production rate and the overall gain. Early research in this area started in the late 1970s and early 1980s [1]. Afterward, many approaches have been proposed for the production process to improve performance and achieve the desired quality and final mass production.

A literature survey of the most popular information extraction and modeling techniques in this area is beneficial for clarifying the research issues and illustrating their weaknesses and achievements. The main goal of this paper is to consolidate the available knowledge on modeling techniques of milling processes. It facilitates the extraction of the inherent relationship between all the effective cutting parameters, sensor signals, and process results by choosing the most appropriate modeling technique [2]–[6]. As a result, it will be easier to choose the proper approach to a descriptive reference model.

There are numerous modeling methods to provide a reference model for milling processes. The classical methods in this field, as well as experiment setups and feature-extraction methods, were covered in our last paper [3]. Many of the state-of-the-art methodologies will be covered in the present paper. These methods are distinguished by their applied feature extraction and data preprocessing approaches. Another important factor for grouping modeling methods is the algorithm which they use. Numerous modeling methods are applied to provide a nonintrusive monitoring of the process. In this paper, artificial intelligence (AI)-based techniques are focused. Fig. 1 illustrates the different aspects of tool condition monitoring and surface roughness prediction on HSM processes.

Probabilistic modeling methods such as Bayesian networks (BNs) and hidden Markov models (HMMs) will be summarized in Section II-A and D. They apply probability rules and relations to form a model for milling process monitoring and prediction. However, these methods are not as common as the methods based on neural networks, fuzzy logic (FL), and their combinations, which are covered in Section II-B. Evolutionary approaches, such as genetic algorithms (GAs) and particle swarm optimization (PSO), are also applied in this field. As Section II-C presents, they are mostly used in combination with other methods for optimization purposes.

Different types of clustering methods and algorithms are also applied to the signal features as the first layer for signal interpretation. Categorization and grouping of distinct signal features and associating them with different cutting phenomena are also the goals of the research works summarized in Section II-E.

Finally, in Section III, discussions of available techniques and research issues and some suggestions for future studies will be presented. Conclusions will be drawn in Section IV.

II. AI-Based Analysis of High Speed Milling Processes

To provide an acceptable infrastructure for representing a general descriptive model, we have to note that milling processes have a nonlinear time-varying multivariable nature and the sensor signals and signal features are applied to represent roughly the state of the process. The following sections present the most commonly used AI techniques.

A. BNs

1) Methodologies and Applications: A BN is a probabilistic graphical model which represents a set of random variables and their probabilistic dependences. It is one of the most famous decision-making methods based on the statistical behavior of the process [8], [12], [36]–[38]. A BN was used in [9] to present the surface-finishing results of a milling process. Naive and tree-augmented naive (TAN) classifiers were used as the learning paradigm. After validation and comparing the confusion matrices, it was shown that, in many cases, TAN-trained BNs are superior to naive-trained BN [9]. Another similar report applies both naive and TAN and compares their performances with artificial neural network (ANN). Since the complex structure of ANN is not opaque comprehensive, Correa *et al.* [10] suggest a BN over ANN. They propose a model for the surface roughness prediction where the correlations between the variables are clearly visualized.

Combined with support vector regression (SVR), BN was applied in [12] to detect tool wear, and its performance is compared to another BN–multilayer perceptron combination. Force features are used as the inputs to both the networks, and they were compared in terms of their prediction accuracy. It is shown that the former model is more accurate [12]. Also, in [8], a BN was used for studying acoustic emission and spindle power metrics. Face milling and drilling processes were investigated, and the applicability of BN to the prediction of their surface-finishing results was compared. As a result, the root causes of many changes in the signal during the process were correlated to the available cutting conditions.

TABLE I
BNs AND HIDDEN MARKOV MODELING APPROACHES TO MILLING PROCESSES

Papers	Material	Condition	Cutting Conditions	Center	Signals	Analysis Target	Preprocessing Technique
Bayesian networks							
[8]	AISI 4140	NM*	NM	NM	AE*, Spindle power	Tool wear, WH*	Feature extraction
[9]	F114 steel	NM	Vc Fc Dc Cn Fz Flutes Td*	Kondia HS1000	NA*	Ra*	K-means discretizer
[10]	Aluminium		Cutting conditions, Geometry	Kondia HS1000	Force	Ra	TAN* algorithm
[11]	Steel St 52-3	NM	Dc Vc Fc	Spinner VC 560	NA	Ra	NA
[12]	ASSAB718HH	NM	NA	Makino CNC	Force	Tool wear	Support vector machine
Hidden Markov models							
[13]	Inconel 718	Dry	Fc Dc Vc	Ruder HSM	Vibration, Force, AE	Tool wear	Continious wavelet transform
[14]	NM	NM	NM	NM	Vibration	Failure detection	Modulus maxima wavelet
[15]	Aluminium	NM	Vc Cn Fc Dc Td	HS-1000 Kondia	AE, Force, Vibration	Tool life	DFT*
[16]	NM	NM	NA	NM	Vibration	Tool monitoring	Spectral feature extraction
[17]	AISI 4340 steel	Water-soluble	Dc Vc Fc	MAZAK H800	Vibration, AE	Tool wear	Classification

* See the Abbreviation Appendix

2) Pros and Cons: Table I consolidates the recent research works applying BNs for high speed milling processes. BNs have been extensively used in the best cutting condition determination problem. Using proper signal features as inputs, there are many applications where BN is used for modeling. Compared to other AI techniques, statistical models need more data for training to achieve the same level of accuracy which is considered a negative aspect. However, since it is graphically representable, using a transition probability matrix makes all the significant and insignificant parameters in the process easily recognizable for the researcher.

B. FL, Neural, and FNN-Based Methods

1) Methodologies and Applications: ANNs, FL, and their combinations such as fuzzy nets (FNs) are widely used in modeling HSM processes. They have also been shown to be capable of modeling not only end milling but also other kinds of machining processes, providing an accurate approximation of the surface finishing [39]–[45]. Each report applies ANN with a different algorithm. However, choosing the best structure is still an open problem. In order to model the machining process, the feedforward-back-propagation algorithm has been used extensively in many articles. The details of the structure and connections between inputs/outputs, e.g., the number of hidden layers and their neurons, are also considered as an important issue. Some discussions on the optimum modeling structure can be found in [18], [21], and [22].

In [31], tool wear and surface roughness are correlated with cutting conditions and force features using a back propagation neural network structure. However, since there are many choices for ANN modeling, there is still an issue in choosing the best method and structure. For example, in [25]–[27], the radial basis functions (RBFs), back-propagation methods, and dynamic models are compared to find the best structure. Using only one hidden layer and proper design of experiment, a model was presented that had the ability to capture the characteristics of the force signal given the cutting conditions. Then, Lu [26] obtained an approximation to the surface profile while Briceno *et al.* [25] showed that RBF is superior in the sense of a presented cost function in the prediction of force features.

Given the fact that the wavelet coefficients of the force signal carry different patterns in normal and a broken tool, an ART2-type self-learning neural network was designed to detect signs of tool failure from the force signal [23].

In addition to the cutting conditions and vibration signal, to predict the output surface profile, the fractal geometry and self-similarity properties of the surface were used as a reference building block for all surface patterns and for determining fractal parameters in [24].

Overfitting and slow learning are also important challenges in applying ANN models. The support vector machine (SVM) method has been developed to overcome such issues by minimizing the generalization error as well as by maximizing the separation margin rather than the training error.

As described in [46], there are comparatively few parameters to be set in SVM methods. With their benefits, SVM and SVR have been used for force, power, and spindle displacement signals to classify broken tools [11], [32]–[35].

FL-based tool wear monitoring was suggested in [47]. To predict flank wear, it utilizes the maximum cutting force with other cutting conditions. Forming its rule base according to experimental and expert knowledge, it is able to estimate the existing flank wear. In [48], an FL-based controller was applied on feed current signal to increase the metal removal rate (and lessen the production time) while maintaining a constant cutting force.

Fuzzy-neural network (FNN) can also be applied to many machining processes as a condition monitoring system [49]. For example, the hybrid Taguchi–genetic learning algorithm was used in [40] to fit a nonlinear model to the R_a values of a best cutting condition determination experiment. The learning data are identical to that used in [50]. The aim is to compare the results of different choices for membership functions which are used in the adaptive neuro-fuzzy inference system. As a complete example of a combined monitoring and control system, fuzzy-neuro adaptive surface roughness control (FN-ASRC) was applied in [39], where FN-ASRC is divided into two distinct parts. One is the fuzzy-neuro in-process surface roughness recognition, which predicts the surface roughness, and the other subsystem is the FN adaptive feed-rate control (FN-AFRC), which suggests appropriate modifications to the cutting conditions in order to achieve a determined surface roughness set point. Fig. 2 illustrates the framework of the FN-AFRC method introduced [39].

In order to develop the whole monitoring and control system, two distinct five-layer FNs were used. The layers are the input, feature-extraction, relations, combination, and defuzzification layers. The fuzzy rules for identification and control are defined, and conflicting rules are moved out of the rule base. The process is stopped halfway for the fuzzy neuro-in process surface roughness prediction (FN-IPSRP) system to predict the surface roughness for the rest of the path. Then, in order to improve the surface roughness, a feed-rate modification is

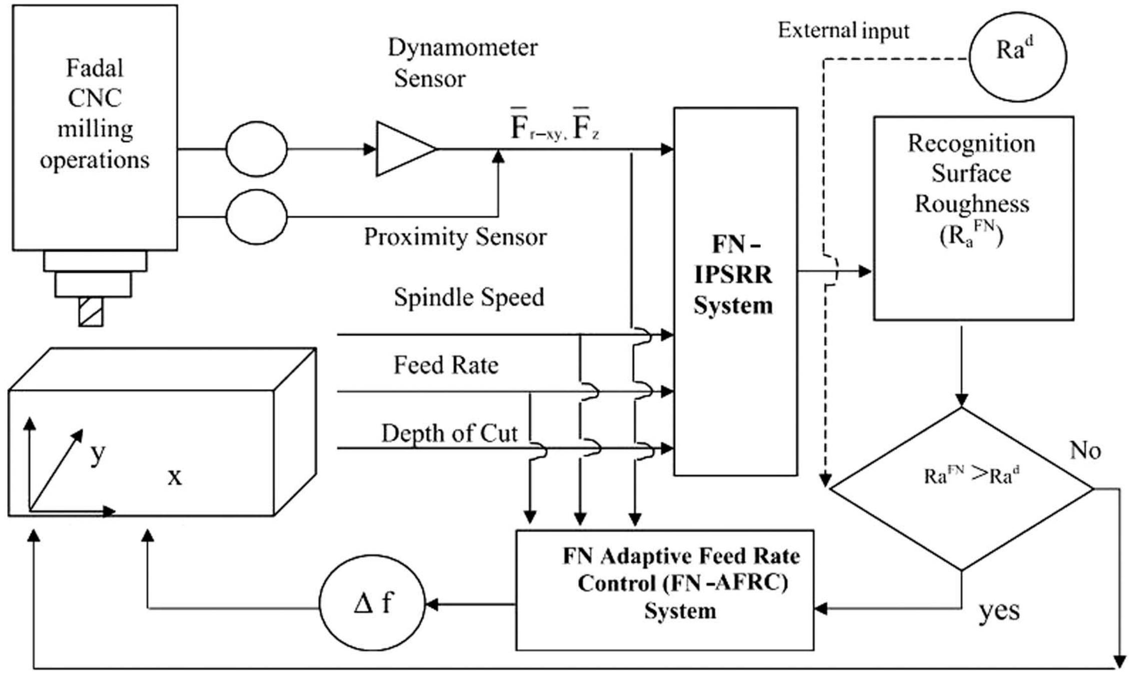


Fig. 2. FN system proposed by [39] to adaptively control the surface roughness according the predicted values for surface finishing.

TABLE II
ANNs AND SVM MODELING APPROACHES TO MACHINING PROCESSES

Papers	Material	Condition	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Artificial neural network							
[18]	Titanium	NM*	Vc Fc Dc*	NM	NA*	Ra*	NA
[19]	Stainless steel 304L	Dry	Vc Fc Dc	CNC lathe	NA	Surface profile	NA
[20]	Steel	NM	Dc Vc Fc	VDF lathe	Flank wear	Tool life	PSO*
[10]	Aluminium	NM	Cutting conditions, Geometry	Kondia HS1000	Force	Ra	NA
[15]	Aluminium	NM	Vc Cn Fc Dc Td	HS-1000 Kondia	AE*, Force, Vibration	Tool life	DFT*
[21]	Steel	NM	Vc Fc Dc Cn Fz Td	CNC lathe,	AE, Force, Vibration	Ra	NA
[18]	Titanium	NM	Vc Fc Rake angel	NM	NA	Ra	NA
[22]	AISI 1030 steel	Dry	Dc Vc Fc	CNC lathe	NA	Ra	NA
[23]	Kistler 9257A	NM	NA	NM	Force	Tool failure	Wavelet transformation
[24]	NM	NM	Vc Cn Fc Dc Td Initial tool wear	NM	Vibration	Surface profile	Fractal geometry approach
[25]	NM	NM	Dc Vc Fc	VMC-3016L	Force	Force features	Feature extraction
[26]	Stainless steel 304L	Dry	Dc Vc Fc	NM	AE, Vibration	Surface profile	FFT*
[27]	NM	NM	Dc Vc Fc	NM	Force	Flank wear	Normalization
[28]	AISI 1020 steel	NM	Dc Vc Fc	NM	NA	Tool wear	Comparison to MRM*
[29]	1040 carbon steel	Dry	NA	Bridgeport	Vibration	Tool wear	FFT features
[30]	16MnCrSi5 XM steel	NM	Geometry Vc Fc Material	Heller	force	force	Actual Value
[31]	AISI 1040	dry	Vc Fc Dc	Taksan	force	Tool Wear, Ra	ANOVA
Support vector machine							
[32]	6061 Aluminium	NM	Dc Vc Fc	Fadal CNC	NA	Ra	PSO
[11]	Steel St 52-3	NM	Dc Vc Fc	Spinner VC 560	NA	Ra	NA
[33]	Cast-Iron Aluminium	NM	Vc Cn Fc Dc Td	Fadal VMC-40	Force	Tool breakage	Low pass filter
[34]	steel AISI-1018	NM	Dc Vc Fc	Okuma ES-V3016)	Force, Power	Tool breakage	Support vector regression
[35]	7075 Aluminium	NM	NA	NM	Displacement, Power	Tool breakage	NA

* See the Abbreviation Appendix

suggested by FN-ASRC based on the predicted results of the FN-IPSRP [39].

FNs have also been applied to model the milling process [42], [51], [52]. In this method, a number of membership functions are assigned to the input space and are fine-tuned in order to obtain the most accurate input/output model. Then, combinations of these membership functions are considered as possible associative rules in the model's rule base. After all, only the rules with more occurrences and no conflicts will remain. For performance verification, several designs have been tested, and the FN method performs acceptably in its surface roughness predictions. The previously discussed papers are summarized in Tables II and IV.

2) Pros and Cons: This section has mentioned methods that have been applied due to their ability to model nonlinear processes which are applicable to experiments to determine the best cutting conditions as well as destructive tests. Since the literature has been established on these techniques, there have been plenty of different implementations of these methods concerning prediction and control of milling processes. There are also reports that claim the repeatability of the models. However, none have claimed to be a universal reference model for milling processes, and there is no consistency with regard to their requirements for inputs and outputs which opens the doors for more investigations. Also, many types of models have yet to be developed and tested, such as the combination of neuro-fuzzy algorithms with other AI methods and dynamic fuzzy models [53] for offline and online monitoring systems. However, the capability of these models to capture the nonlinear time-varying nature of the process is an advantage of such methods,

TABLE III
GAs, FUZZY PETRI NET, PSO MODELING APPROACHES TO MACHINING PROCESSES

Papers	Material	Condition	Cutting conditions	Center	Signals	Analysis target	Preprocessing technique
Genetic algorithm							
[54]	6061 Aluminium	NM*	Vc Cn Dc Fc*	NM	Vibration	Ra*	Genetic programming
[55]	AISI 1018 steel	NM	Fc	NM	Cutting force, Feed force	Tool wear	Fuzzification
[56]	Aluminium 6061-T8	NM	Dc Vc Fc	VMC550	NA*	Ra	NA
[57]	NM	NM	Fc Vc Power	NM	NA	Fc Vc Power	Simulated annealing
[58]	NM	NM	Dc Vc Fc	NA	NA	Dc Vc Fc	NA
[59]	T20 Grey cast	NM	Dc Vc	NM	NA	Fc	NA
[60]	X20Cr13 Steel	NM	Dc Vc Fc Cutter engagement	Fil Fresatrici	Flank wear	Ra	ANN*, Compared with PSO*
Particle swarm optimization							
[61]	Aluminium	Dry	Dc Vc Fc	Fadal VMC	NA	Ra	NA
[32]	6061 Aluminium	NM	Dc Vc Fc	Fadal CNC	NA	Ra	SVM*
[20]	Steel	NM	Dc Vc Fc	VDF lathe	Flank wear	Tool life	ANN
[60]	X20Cr13 Steel	NM	Dc Vc Fc Cutter engagement	Fil Fresatrici	Flank wear	Ra	ANN, Compared with GA*

* See the Abbreviation Appendix

and their inflexible and complex structure is a disadvantage. In addition, there are very few online prediction and control research works available in this field, which leaves space for more investigations on universal online models.

C. Evolutionary Algorithms, GAs, GP, and PSO

1) Methodologies and Applications: The method of GAs is an optimization method based on evolutionary searching of the solution space. The idea was based on the works introduced in [62]. GA is used in the machining technology field for modeling issues wherever optimization is concerned. In [58] and [59] for example, it was used for best cutting condition determination. With proper economic justifications for the cost of the process and modified limitations on the variables, many cost functions are defined for the process with some cutting conditions as optimization variables [58], [59]. The same problem is solved with a combination of simulated annealing (SA) and GA in [57]. However, since GA training requires random measurements on surface roughness and tool wear, it is not easy to be generalized in the available form or to be used for online analysis and prediction.

The genetic programming (GP) method was first introduced in the early 1990s by Koza [63]. Basically, it is an evolutionary algorithm that makes the program perform better in evolving and producing an optimal model that matches the data. Theoretically, they are represented in the form of recursively evaluated and evolved tree structures. Every tree node has an operator function, and every terminal node has an operand, making mathematical expressions easy to evolve. There are several implementations of this method for milling process modeling [54], [55], [64]. A general review of these methods can be found in [64]. The method was used in [54] and [56] to represent the surface roughness in its dependence on the cutting conditions and the vibration signal. According to this method, an evolutionary algorithm investigates the best match for the experimental data by evolving the tree of operators and operands as modeling functions for the milling process using simple function genes and terminal genes. In [60], a GA-optimized neural network was applied to tool condition monitoring where GA was applied to fine-tune the neural network parameters. The performance of this model was also compared with that of PSO-based neural network. In both the GA- and PSO-based approaches, these optimization methods are applied for determining the neural network parameters.

The PSO method is a famous optimization procedure based on a direct search method which imitates social behavior in the presence of objectives. It was first introduced by Kennedy *et al.* [65] and was used in several applications. It uses an iterative formula for the swarms to approach global maxima

$$\begin{aligned} v_{i,j} &= c_0 v_{i,j} + c_1 r_1 (globalbest_j - \ldots x_{i,j}) \\ &\quad + c_2 r_2 (localbest_{i,j} - x_{i,j}) \\ &\quad + \ldots c_3 r_3 (neighbourhoodbest_j - x_{i,j}) \\ x_{i,j} &= x_{i,j} + v_{i,j}. \end{aligned} \tag{1}$$

Due to its ability to search for the global optimum, $globalbest$, proportional to the local optimum, $localbest$, and nearest optimum, $neighbourhoodbest$, it has been mostly applied in milling processes to optimize the cutting conditions. PSO was used for the first time in the machining literature where, to find the best matching parameters for a proposed surface roughness model [61], [66]

$$R_a = \frac{10 a D_c^b F_c^c}{V_c^d} \tag{2}$$

where R_a is the surface roughness, D_c is the radial depth of the cut, F_c represents the feed factor, V_c is the spindle speed, and a, b, c, and d are unknown parameters.

In [32] and [61], PSO was applied to the results of an SVM. The SVM determines the unknown parameters of the model in (2). Then, PSO was used to find the optimal cutting conditions [32].

Because of its ability to find the optimal solution for most nonlinear objective functions, there is no specific limitation on using any predefined model for the process. For example, Cus *et al.* [66] use an ANN model for the force versus surface roughness, and a PSO algorithm was applied to find the optimum cutting conditions. In [20], an ANN was applied to model the tool life dependent on the cutting conditions and flank wear. PSO is also utilized to optimize the ANN parameters.

2) Pros and Cons: Table III present the papers on these methods. Since GAs were not developed for dynamic training until recently, they were just used for offline best cutting condition determination. Since the basic idea is to reach an optimum point for an objective function, it can be properly used for building a best fitting model on offline raw data. However, there is no report that this method is capable of online adaptation. There are some studies that suggest merging dynamic learning with this method [67], so it might be applied in the future studies. Another issue that exists with GP is the complex formulations

TABLE IV
FNN MODELING APPROACHES TO MACHINING PROCESS

Papers	Material	Condition	Cutting conditions	Center	Signals	Analysis target	Preprocessing technique
Fuzzy neural network							
[40]	NM*	NM	Vc Dc Fc*	NA*	NA	Ra*	NA
[42]	6061-T6511 Aluminium	Dry	Vc Fc Dc	Storm CNC A50	Force, vibration	Ra	Fuzzification
[72]	Aluminium	NM	Dc Fc Vc	Beaver CNC	AE*, Force, spindle acceleration	TCM*	Taguchi's signal/sensor selection
[73]	Steel #45	Dry	Vc Cn Fc Dc Td	Makino-FNC74-A20	Feed current, Spindle current	TCM	Wavelet analysis
[40]	6061 Aluminium	NM	Vc Dc Fc		NA	Ra	Hybrid Taguchi–genetic learning algorithm
[44]	Alumic-79	NM	Vc Fc Dc flutes Td	NM	NA	Ra	Fuzzification
[45]	NM	NM	NA	NM	Force, vibration, AE	Tool wear	Sensor fusion
[50]	6061 Aluminium	NM	Vc Dc Fc	NM	NA	Ra	NA
[51]	6061 Aluminium	NM	Vc Dc Fc	Fa dal CNC	Vibration, proximity	Ra	Fuzzification
[52]	6061 Aluminium	NM	Vc Dc Fc	Fadal VMC-40	Vibration, Proximity	Ra	Fuzzification
[74, 75]	Inconel 718	semi-dry	NA	Roder	Vibration, force, acoustic	Tool wear	time domain features
[47]	Steel AISI-1018	NM	Dc Fc	Fadal VMC	Vibration, proximity	Tool wear	Fuzzification
[76]	Inconel 718	semi-dry	NA	Roder	Vibration, force, acoustic	Tool wear	Wavelet analysis
[48]	Aluminium	NM	Geometry Vc Fc material	ACE-V30	Spindle and feed current	Force control	Actual value
[47]	Steel	NM	Dc Fc Material	Fadal	Force	Flank wear (Vb)	Actual value

* See the Abbreviation Appendix

TABLE V
CLUSTERING MODELING APPROACHES TO MACHINING PROCESSES

Papers	Material	Condition	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Clustering and classification methods							
[79]	Steel	NM*	Cutting conditions, geometry	Makino FNC 74-A20	AE*	Tool wear	Wavelet packet transform
[80]	EN1A	Dry	Fc Dc Vc*	Cincinnati Sabre 500	Power, force	Flank wear	Signal feature extraction
[17]	AISI 4340 steel	Water-soluble	Dc Vc Fc	MAZAK H800	Vibration, AE	Tool wear	HMM* modeling
[81]	Inconel 718	semi-dry	NA*	Roder	Vibration, AE, force	Tool wear	Wavelet analysis
Self organizing maps							
[16]	NM	NM	NA	NM	Vibration	Tool monitoring	Spectral feature extraction
[82]	Inconel 718	HRC52	dry	Roder	Vibration, force, AE	Tool wear	Fuzzy regression model

* See the Abbreviation Appendix

and functions in the output. To make the modeling more meaningful, the output model has to reflect the mechanical nature of the process. This makes the model process more computationally intensive.

As an optimization method similar to GAs, PSO is also used to facilitate nonlinear model identification and parameter determination. Also, it can be used as a training method for other AI techniques to find the best fitting model for the milling process. Since it requires an existing nonlinear function, it might not be suitable for online data analysis and prediction. Perhaps with some modifications in the variable definitions, it might be able to work in real time as well as GA.

D. HMMs

1) Methodologies and Applications: HMMs were first introduced in [68] as "probabilistic functions of Markov chains." Afterward, several methods were introduced for modeling, and their application was summarized in [69]. To formulate an HMM model $\lambda = (\pi, A, B)$, usually, N distinct (hidden) states q_i for the system are considered. The Markov chain is defined by the connecting transitions between q_i states. These connections are completely defined by the state transition matrix $A = [a_{ij}]$ where each element a_{ij} represents the probability of the corresponding transition

$$a_{ij} = P(q_t = j | q_{t-1} = i), 1 \leq i, \qquad j \leq N. \tag{3}$$

Since the a_{ij}'s are probability values, the following axiomatic constraints are applied [70]:

$$a_{ij} \geq 0, \sum_{j=1}^{N} a_{ij} = 1 \qquad \forall i. \tag{4}$$

We may assume without loss of generality that the start time of the model is 0, at which point the model will have an initial condition. It is represented by the probability of each individual state at the initial time or the initial condition probability distribution, $\pi_i = P(q_0 = i)$, which is the ith element of π. Then, the probability of any chain of states will be

$$P(q|A, \pi) = \pi_{q_0} a_{q_0 q_1} \ldots a_{q_{T-1} q_T}. \tag{5}$$

Since the states of the system are not always observable, the only thing that is available about the system is the observation O_t which is according to the changes in the system states. The relation between these observations and the states is declared by another probability matrix which is called the emission matrix

$$B = \{b_i(O_t)\}_{i=1}^{N}, b_i(O_t) = P(O_t | q_t = i). \tag{6}$$

There are three major issues to be faced for developing an HMM structure to model a system. The first one is to compute the probability of an output event's happening in the available model λ (*Evaluation*). The second issue is to find the unknown parameters of the HMM model which best match the observations O (*Estimation*). The last problem is to find out the most probable sequence of states q, regarding the observations O (*Decoding*). Further details of the available solutions to these three problems and many other applications of HMM are discussed in [17], [69], and [71].

HMM has rarely been used in the literature to present a dynamical model of a process. Each paper has its own way to provide sequential data to HMM training algorithms such as Baum–Welch, known also as the expectation–maximization method. It is a maximum-likelihood-based method that finds the parameters of the state transition matrix and the output emission matrix from the internal states [15], [70]. Originally, it seemed that HMM is not as accurate as other models for a nonlinear system. However, its aggregation with classification and nonlinear methods can lead to better results [16].

To provide data for training an HMM model for a milling process, some papers applied the vector quantization (VQ) method based on the discrete wavelet decomposition of sensor signals which is briefly described in [3]. Applying the codebook of the worn or sharp tool, its status is predicted by applying the current state and emission matrix. There are classification methods other than VQ that have been used to generate the input/output sequence for an HMM to model. For example, Fish *et al.* [17] suggest a modified classifier for HMM training the status of the tool in probabilistic terms rather than in binary output, for example, worn/sharp states.

HMM was also used to correlate the observable changes in the energy content of different continious wavelet transform (CWT) scales of vibration signals with tool wear. Vibration signals are analyzed for some of their details, and for each detail, the changes in the energy are observed for a certain period of time [13]. After proper training, two distinct codebooks for the sharp/worn tool are developed. Simulations show that HMM models can successfully monitor and detect the internal status of a milling tool. Wavelet modulus maxima information was used in [14] to build a combined HMM model. It was shown that this feature has meaningful changes according to tool wear progress and it was applied to provide an accurate representation of machining condition. Therefore, three models for different states of the tool, i.e., normal, warning, and failure condition, were presented. The probability of each sequence of the data was estimated according to these models and the sensor signals, and finally, the highest probability is chosen as the real state of the system.

2) Pros and Cons: HMM has only been applied in a few studies in the literature (see Table I). Compared to BN, it has the benefit of being able to reflect the behavior of milling processes in the form of dynamic models rather than static models. It facilitates an estimation of the internal states of the system, needing only system outputs. As an essential issue, finding the probability distribution structure that fully describes the sequence of the signal features has been investigated in many research works. However, since it requires a large amount of data for training, it seems less appropriate for the modeling of the best cutting condition determination experiments. For destructive tests, however, it might be used the same way that it is used in speech processing and recognition [70], [77], [78] because of the availability of acoustic emission (AE) sensors [13], [16]. However, among the reports on the performance of HMM in milling processes, there are very few predictive accuracy comparisons with other AI techniques in the field.

E. Clustering and Classification Methods

1) Methodology and Applications: Clustering methods are meant to keep similar data together in clusters to facilitate a proper overview of the domain. There are two types of clustering which are commonly applied in milling process: hard/crisp clustering and fuzzy clustering. In the former, a datum can only belong to one cluster, but in fuzzy clustering methods, a datum can be a member of several clusters with a certain membership value. The process of assigning a datum to a cluster or some clusters depends on its distance, similarity, or connectivity to other data in that specific cluster [83], [84].

Fuzzy C-means clustering is a famous clustering technique [85]. It classifies the finite information into several classes based on some criteria. Given a finite set of data, the algorithm returns a list of cluster centers and a partition matrix. Each of its elements is a membership value of a datum that belongs to a specific cluster [85].

On the other hand, each datum is assigned to only one cluster in hard/crisp clustering, as in the k-means algorithm, where a datum is attributed to the cluster with the nearest center. The center of each cluster is the arithmetic mean of all its members. Crisp clustering, such as k-means and k-medoids, is applied in [86] to illustrate the applicability of such methods to modeling approaches.

The fuzzy C-means clustering method was used in [79] on wavelet packet features of AE sensor signals and in [81] and [86] on the energy contents of different scales of CWT of the force and vibration signals. Power consumption and vertical force are also clustered in [80]. Since the rms value of each frequency band in an AE signal changes with different tool conditions [82], this signal feature is indicative of tool wear and surface roughness. As reported in [79], four states for the tool wear, with seven features each, compose the codebook of the clustering method. Fuzzy clustering on continuous and discrete wavelet analysis of ac servomotor current signals of the spindle and feeder was used in [73] for tool breakage detection and tool wear monitoring.

Classification methods have also been applied for milling condition detection purposes. From the experimental knowledge, Elbestawi *et al.* [87] suppose five different classes for feature patterns of sensor signals, applying this knowledge with linear discrimination classification techniques.

Self-organizing maps (SOMs) can be considered as another clustering technique to reduce the dimensionality of the data. The new dimension depends on how the new sets of vectors are ordered. For example, for 2-D SOM, the code vectors are ordered in 2-D and referred to by a code vector index. To train the SOM, each training sample of the high-dimensional space is mapped to its nearest code vector member and hence belongs to the corresponding class. Then, the code vector is updated by moving toward the training vector. Therefore, in the learning procedure, all code vectors move toward the training vector depending on the iteration number and distance from the vector under which the last training vector was classified. SOM was used in [16] to reduce the dimension of the feature space of the time–frequency blueprint of time windowed signals. It was also applied as a part of the rule generation procedure in [82] in combination with a dynamic fuzzy regression modeling system.

2) Pros and Cons: The clustering of the available data of the process will lead to the generalization of the model as the clusters are easier to associate with the tool status than were the pure signals (see Table V). This issue mostly appears when there are different cutters involved. In addition, there are many uninvestigated and unclassified features to be studied, which leaves space for more research. Among them, time–frequency analysis features can be mentioned. These features can be applied in a more methodical way when clustering methods are involved. Moreover, there are quite a number of classification methods that have not been applied to the field of intelligent machining. Also, the combination of clustering methods with AI techniques remains to be investigated more extensively in the field so that the contribution of clustering methods can be clarified. Clustering methods can also be used to investigate the similarities between different signal features. Finding these similarities, other AI techniques can be applied to map different classes to the different respective conditions of the tool and milling process. However, the number of classes and the structure of the classification method may determine its accuracy, and they are open issues for further investigations.

III. Discussion

The survey presented in the previous sections shows that there is no lack of good ideas in modeling milling processes. However, there are some open issues that need to be addressed in future investigations. One of these issues is that the predictions resulting from these approaches must be accurate and repeatable. It has been shown experimentally and mathematically that AI-based methods are more accurate than other classical methods. It is also clear that each one of these state-of-the-art modeling, inference, and decision-making methods is able to predict surface roughness and tool wear in a nonintrusive manner. As such, any theoretical development in one of these methods results in a more informative, accurate, and repeatable reference model. However, from the industrial point of view, any approach developed must be easy to implement. The learning speed and simplicity of the model structure dealing with changes in the system are the challenges for the future. One of the beneficial characteristics that a future research in this field has to address is an insightful comparison between methods. The majority of the available papers concern only one method and its capabilities of dealing with the process. Referring to the different sections of this paper, it is obvious that, although many AI techniques have been utilized for tool wear detection or modeling surface roughness, there are many methods yet to be investigated. For example, not all of these AI techniques have been studied as to finding the most appropriate configuration, algorithm, and structure. Many of the proposed methods have yet to be tuned in some of their parameters, and they vary from one experiment design to another. Moreover, there are no dynamic and intelligent methods in the field that can be applied without unnecessary initializations. Other methods, such as BNs [8], [9] and Petri nets [41], have been applied to tool status and surface-finishing predictions. However, the justification of event-based models needs more study, and many advanced and intelligent event-based models such as that in [88] have not yet been investigated in this field.

To summarize the discussion, there are some obvious research gaps in the field that need to be addressed.

1) One challenging area is to take better and more descriptive features out of the collected signals using more suitable signal processing schemes and feature selection methods.
2) Unavoidable frequency drift of the signals and changes in their shape during their lifetime due to mechanical parameter imperfections have not been extensively investigated. These frequency drifts are different from those due to tool aging.
3) Changes in machine dynamics during long-term running, which can lead to undesirable inaccuracy of the reference system model, are another issue to be focused on in monitoring systems.
4) The lack of proper investigation of the data preprocessing methods is also obvious in the field of machining. For example, wavelet analysis and other state-of-the-art pattern decomposition and extraction methods have only just recently been utilized for milling process signals, but they seem to be appropriate approaches for the extraction of the different properties of the signal.
5) There are not many reports on the interpolation of the results from one type of cutter to another. Therefore, no matter how the cutters differ in their diameter or edge-preparation methods, for every new cutter, the modeling must be repeated, which is expensive and time consuming.
6) The effect of some production parameters of the cutters, such as edge-preparation methods, grinding quality, initial surface roughness on the cutting edge, geometrical cutting-edge design angles, and various coatings, has not yet been investigated.
7) There is apparent lack of investigation of the use of clustering, classification, and grouping methods in this field. One possible reason is the direct use of cutting conditions and signal features instead of clustering data in AI-based models.
8) In the literature, there are very few papers that pay attention to the changes in the shape of the signal due to tool degradation, aging, or tool wear. These methods do not quantitatively investigate such changes. Mostly, they are limited to the use of frequency- and time-domain features and not the cross-correlation of the shape of the signals with corresponding tool edge phenomena.
9) There are many AI techniques that have not been used in the field of modeling of machining processes. As an example, syntactic classification and modeling can be considered. This is a knowledge-based pattern recognition method. To model a sophisticated pattern, it provides more simple patterns, called *primitives*, which are composed to make that complex one. Therefore, a hierarchical model is presented for any similar pattern, or simply, any major pattern is decomposed to appropriate primitives as its building blocks [89]. It has been used in some articles to find the prespecified shapes in the signals [90]. This method was extensively used in speech processing [91] and can be used for fault diagnostic and automatic failure sign distillation for tool condition detection. Another example could be the extreme learning method [92]. This method has proved to be applicable in online sequential learning. Similarly, other similar state-of-the-art techniques that have not yet been used in the field can be applied.

TABLE VI
ADVANTAGES AND DISADVANTAGES OF THE AI METHODS

Method	Advantages	Disadvantages	Target variable	Preprocessing methods	Signals
Mathematical modeling and numerical difference equation solution methods	A less costly method, detailed analysis, many simulation runs before actual running of the system	Can't deal with surface and tool imperfections, Incapable of real-time simulation and analysis	Ra*, tool breakage, force, chip formation, temperature, strain, crack formation, surface profile	Low pass filter, Hilbert transform, finite element techniques, grey relational analysis	Feed-motor current, force, vibration
Statistical and experimental evaluation	Outline of similarities and differences, easy to visualize	No mathematical model, never can be used on-line	Flank wear, tool wear, tool breakage, Ra, surface profile, force	FFT*, time frequency analysis, power spectrum, NM*, Grey relational analysis, Taguchi design procedure, wavelet packets, discrete wavelet transform, force dynamical model	AE*, force, temperature, tool wear, power, vibration
Multiple regression	Predetermined structure, not accurate enough	Simple to be used, simple and few computations, trained model can be used on-line	Tool wear, Ra	Comparison to ANN*, wavelet transform	Force, cutting conditions
Genetic algorithm and genetic programming	Optimization of the model, Suggesting new models	Time consuming, is not yet used for on-line machining, needs a pre-determined model to optimize	Tool wear, Ra	Genetic programming, fuzzification	Vibration, cutting force, feed force
Particle swarm optimization (PSO)	Parameter optimization of the model, can be combined with other methods, simple to implement	Needs a predetermined model, slow, cannot be used on-line	Tool life, Ra	Combined with SVM*, ANN	Flank wear
Bayesian networks and HMMs*	Presents hidden relations, visually representable, BNN* is superior to other statistical methods, good for generalization	Needs a lot of data, slow training, HMM is not reported accurate	Failure detection, workpiece hardness, Ra, tool wear	Feature extraction, k-means discretizer, TAN* algorithm, support vector machine, discrete wavelet transform, modulus maxima wavelet, DFT*, spectral feature extraction, classification	Spindle power, vector quantized vibration, AE, force, vibration, torque, current
Fuzzy-logic and neural network based methods	Good for nonlinear models, well-established theory, high accuracy, fast in evaluation, easy to generalize	Overtraining problems, slow in training, ANN is not easily generalized to other cutters, fixed structure of model	Tool wear, flank wear, surface profile, force features, tool failure, Ra, tool condition monitoring	Fuzzification, Taguchi's Signal/sensor selection, wavelet analysis, sensor fusion, hybrid Taguchi-genetic learning algorithm, PSO, fractal geometry approach, FFT, normalization, compared to MRM*, FFT* features, DFT	Flank wear, AE, force, vibration, power, spindle acceleration, feed current, spindle current, vibration, proximity
Discrete-event based intelligent methods	Can be combined with other AI*, visually representable, new in machining field	Needs clear justification for event-based model, needs reasons for events, vague dealing with fixed parameters, definition of states of DEVS*	Ra	Fuzzification	Force, vibration, AE, spindle speed, feed rate
Clustering methods	Easy to generalize, lowers the data dimensions, can be combined with AI techniques, new in the machining field, easy to correlate the distinct classes with cutting phenomena	Low accuracy, some of the techniques have fixed and huge structure, slow training	Tool wear, Flank wear, Ra, Tool breakage	HMM modeling, Signal feature extraction, Wavelet packet transform, Support vector regression, PSO, Low pass filter	Force, power, displacement power, AE, vibration

* See the Abbreviation Appendix

10) In addition, many variable structure AI techniques, such as that in [53], have not been studied yet. They might replace the fixed structure of many of the mentioned structures and facilitate the generalization of those methods. The fixed structures of the available methods prevent them from being easily generalized and from being used online.
11) Some papers provide a solution for one specific experimental design in such a way that the results cannot easily be generalized to other design issues and conditions. As a result, many experiments are needed for modeling a new experimental design. The ability of the models to remain descriptive and useful in different scenarios is a critical issue.
12) Monitoring and prediction cover only one part of the mission. The resulting reference models ought to be applied in forming model-based controllers to adjust the cutting parameters according to the demands of the end user.
13) The overall structures for generalizable monitoring and prediction system using the available modeling methods have not been considered in the field. This would seem to be a big gap needing to be covered in future studies. To cover this area, the entire structure of the monitoring system has to be investigated for the best techniques to be applied in each part and their interconnectivity and reasonable places in the structure.

An overview on the available techniques and their application in milling process modeling can be found in Table VI. It summarizes the works mentioned in this paper and presents their advantages and disadvantages. Advances in AI, dynamic structure modeling techniques, and clustering methods as well as data preprocessing schemes should be considered to affect the future of the investigations and provide better solutions for industry, such as better quality and more productivity.

IV. CONCLUSION

This paper has investigated several commonly used methods for surface-finishing quality modeling of high speed milling processes. It covered many AI methods as well as classical ones. The simpler methods are typically used for the simple presentation of the behavior of the process while AI-based methods are applied for modeling, online monitoring, and predictive control. Based on these two categories, we investigated the state-of-the-art methods which are commonly used for both modeling and control. Since the nature of the process is multivariable and nonlinear, most of these modeling approaches are found to be able to model such systems. BNs, fuzzy Petri nets, HMMs, and dynamic FNNs have proved to be the most suitable modeling techniques. On the other hand, there are many research gaps that need to be addressed in this field. Moreover, there are very few research reports on the tool-production methods; tool attributes; and their effects and correlations with the sensor signals, surface roughness, and tool degradation. Also, there is an obvious research gap as to presenting a single general model for milling processes where the available models are not expandable to other cutters (even those with similar attributes). Before obtaining a general descriptive model for milling processes, the field keeps being updated by new ideas based on fresh AI techniques and different features of the sensor signals.

In this paper, many of the available modeling methods were discussed. Their benefits and disadvantages were presented, and many research gaps in this field were identified. This survey paper will facilitate the selection of an appropriate modeling technique for different research purposes concerning milling processes. Also, some weaknesses from the research point of view in the field of machining technology were made clear.

ACKNOWLEDGMENT

The authors would like to thank S. Javidi and Dr. S. Dehghani for their kind help and guidance and Singapore's SIMTech staff for their contributions and sincere help. The authors would also like to thank the anonymous reviewers of this paper whose suggestions helped much in the improvement of its quality.

REFERENCES

[1] W. A. Kline, R. E. DeVor, and J. R. Lindberg, "The prediction of cutting forces in end milling with application to cornering cuts," *Int. J. Mach. Tool Des. Res.*, vol. 22, no. 1, pp. 7–22, 1982.
[2] S. Y. Liang, R. L. Hecker, and R. G. Landers, "Machining process monitoring and control: The state-of-the-art," *J. Manuf. Sci. Eng.*, vol. 126, no. 2, pp. 297–310, Jul. 2004.
[3] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, L. Zhai, S. J. Phua, J. Zhou, S. Lin, S. Huang, and J. T. T. Tijo, "A survey on artificial intelligence technologies in modeling of high speed end-milling processes," in *Proc. IEEE/ASME Int. Conf. AIM*, 2009, pp. 320–325.
[4] P. G. Benardos and G. C. Vosniakos, "Predicting surface roughness in machining: A review," *Int. J. Mach. Tools Manuf.*, vol. 43, no. 8, pp. 833–844, Jun. 2003.
[5] P. W. Prickett and C. Johns, "An overview of approaches to end milling tool monitoring," *Int. J. Mach. Tools Manuf.*, vol. 39, no. 1, pp. 105–122, Jan. 1999.
[6] R. Teti, K. Jemielniak, G. O'Donnell, and D. Dornfeld, "Advanced monitoring of machining operations," *CIRP Ann.—Manuf. Technol.*, vol. 59, no. 2, pp. 717–739, 2010.
[7] K. P. Zhu, Y. S. Wong, and G. S. Hong, "Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results," *Int. J. Mach. Tools Manuf.*, vol. 49, no. 7/8, pp. 537–553, Jun. 2009.
[8] S. Dey and J. A. Stori, "A Bayesian network approach to root cause diagnosis of process variations," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 1, pp. 75–91, Jan. 2005.
[9] M. Correa, C. Bielza, M. D. J. Ramirez, and J. R. Alique, "A Bayesian network model for surface roughness prediction in the machining process," *Int. J. Syst. Sci.*, vol. 39, no. 12, pp. 1181–1192, Dec. 2008.
[10] M. Correa, C. Bielza, and J. Pamies-Teixeira, "Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process," *Expert Syst. Appl.*, vol. 36, pt. 2, no. 3, pp. 7270–7279, Apr. 2009.
[11] B. Lela, D. Baji, and S. Jozi, "Regression analysis, support vector machines, and Bayesian neural network approaches to modeling surface roughness in face milling," *Int. J. Adv. Manuf. Technol.*, vol. 42, no. 11/12, pp. 1082–1088, Jun. 2009.
[12] J. Dong, K. Subrahmanyam, Y. Wong, G. Hong, and A. Mohanty, "Bayesian-inference-based neural networks for tool wear estimation," *Int. J. Adv. Manuf. Technol.*, vol. 30, no. 9/10, pp. 797–807, Oct. 2006.
[13] A. J. Torabi, E. M. Joo, L. Xiang, Z. Lianyin, and S. Linn, "Hidden Markov model for ball-nose tool condition monitoring," in *Proc. Postgraduate Student Conf. AOTULE*, Bandung, Indonesia, Nov. 2010.
[14] Q. Miao and V. Makis, "Condition monitoring and classification of rotating machinery using wavelets and hidden Markov models," *Mech. Syst. Signal Process.*, vol. 21, no. 2, pp. 840–855, Feb. 2007.
[15] A. J. Vallejo, R. Morales-Menndez, and J. R. Alique, *On-Line Cutting Tool Condition Monitoring in Machining Processes Using Artificial Intelligence*. Rijeka, Croatia: InTech, 2008.
[16] L. M. D. Owsley, L. E. Atlas, and G. D. Bernard, "Self-organizing feature maps and hidden Markov models for machine-tool monitoring," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2787–2798, Nov. 1997.
[17] R. K. Fish, M. Ostendorf, G. D. Bernard, and D. A. Castanon, "Multilevel classification of milling tool wear with confidence estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 75–85, 2003.
[18] A. M. Zain, H. Haron, and S. Sharif, "Prediction of surface roughness in the end milling machining using artificial neural network," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1755–1768, Mar. 2010.
[19] Y. Tang and H. Guo, "HSM strategy study for hardened die and mold steels manufacturing based on the mechanical and thermal load reduction strategy," *J. Univ. Sci. Technol. Beijing, Miner., Metallurgy, Mater.*, vol. 15, no. 6, pp. 723–728, Dec. 2008.
[20] U. Natarajan, V. Periasamy, and R. Saravanan, "Application of particle swarm optimisation in artificial neural network for the prediction of tool life," *Int. J. Adv. Manuf. Technol.*, vol. 31, no. 9/10, pp. 871–876, Jan. 2007.
[21] U. Zuperl and F. Cus, "Optimization of cutting conditions during cutting by using neural networks," *Robot. Comput.-Integr. Manuf.*, vol. 19, no. 1/2, pp. 189–199, Feb.–Apr. 2003.
[22] M. Nalbant, A. AltIn, and H. Gokkaya, "The effect of cutting speed and cutting tool geometry on machinability properties of nickel-base inconel 718 super alloys," *Mater. Des.*, vol. 28, no. 4, pp. 1334–1338, 2007.
[23] I. N. Tansel, C. Mekdeci, and C. Mclaughlin, "Detection of tool failure in end milling with wavelet transformations and neural networks (WT-NN)," *Int. J. Mach. Tools Manuf.*, vol. 35, no. 8, pp. 1137–1147, Aug. 1995.
[24] I. A. El-Sonbaty, U. A. Khashaba, A. I. Selmy, and A. I. Ali, "Prediction of surface roughness profiles for milled surfaces using an artificial neural network and fractal geometry approach," *J. Mater. Process. Technol.*, vol. 200, no. 1–3, pp. 271–278, May 2008.
[25] J. F. Briceno, H. El-Mounayri, and S. Mukhopadhyay, "Selecting an artificial neural network for efficient modeling and accurate simulation of the milling process," *Int. J. Mach. Tools Manuf.*, vol. 42, no. 6, pp. 663–674, May 2002.
[26] C. Lu, "Study on prediction of surface quality in machining process," *J. Mater. Process. Tech.*, vol. 205, no. 1–3, pp. 439–450, Aug. 2008.
[27] K. Venkatesh, M. Zhou, and R. J. Caudill, "Design of artificial neural networks for tool wear monitoring," *J. Intell. Manuf.*, vol. 8, no. 3, pp. 215–226, May 1997.
[28] P. Palanisamy, I. Rajendran, and S. Shanmugasundaram, "Prediction of tool wear using regression and ANN models in end-milling operation," *Int. J. Adv. Manuf. Technol.*, vol. 37, no. 1/2, pp. 29–41, Apr. 2008.
[29] I. Abu-Mahfouz, "Drilling wear detection and classification using vibration signals and artificial neural network," *Int. J. Mach. Tools Manuf.*, vol. 43, no. 7, pp. 707–720, May 2003.
[30] U. Zuperl, F. Cus, and M. Reibenschuh, "Neural control strategy of constant cutting force system in end milling," *Robot. Comput.-Integr. Manuf.*, vol. 27, no. 3, pp. 485–493, Jun. 2011.
[31] H. Saglam and A. Unuvar, "Tool condition monitoring in milling based on cutting forces by a neural network," *Int. J. Product. Res.*, vol. 41, no. 7, pp. 1519–1532, Jan. 2003.
[32] C. Prakasvudhisarn, S. Kunnapapdeelert, and P. Yenradee, "Optimal cutting condition determination for desired surface roughness in end milling," *Int. J. Adv. Manuf. Technol.*, vol. 41, no. 5, pp. 440–451, Mar. 2009.
[33] Y.-W. Hsueh and C.-Y. Yang, "Prediction of tool breakage in face milling using support vector machine," *Int. J. Adv. Manuf. Technol.*, vol. 37, no. 9/10, pp. 872–880, Jun. 2008.
[34] S. Cho, S. Asfour, A. Onar, and N. Kaundinya, "Tool breakage detection using support vector machine learning in a milling process," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 3, pp. 241–249, Mar. 2005.
[35] Y.-W. Hsueh and C.-Y. Yang, "Tool breakage diagnosis in face milling by support vector machine," *J. Mater. Process. Technol.*, vol. 209, no. 1, pp. 145–152, Jan. 2009.
[36] T. A. Stephenson, *An Introduction to Bayesian Networks Theory and Usage*, Feb. 2000. [Online]. Available: ftp://ftp.idiap.ch/pub/reports/2000/rr00-03.pdf
[37] O. Cetin, M. Ostendorf, and G. D. Bernard, "Multirate coupled hidden Markov models and their application to machining tool-wear classification," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2885–2896, Jun. 2007.
[38] A. Kumar, F. Tseng, Y. Guo, and R. B. Chinnam, "Hidden-Markov model based sequential clustering for autonomous diagnostics," in *Proc. IEEE IJCNN (IEEE World Congr. Comput. Intell.)*, 2008, pp. 3345–3351.

[39] L.-D. Yang, J. C. Chen, H.-M. Chow, and C.-T. Lin, "Fuzzy-nets-based in-process surface roughness adaptive control system in end-milling operations," *Int. J. Adv. Manuf. Technol.*, vol. 28, no. 3/4, pp. 236–248, Mar. 2006.

[40] W.-H. Ho, J.-T. Tsai, B.-T. Lin, and J.-H. Chou, "Adaptive network-based fuzzy inference system for prediction of surface roughness in end milling process using hybrid Taguchi-genetic learning algorithm," *Expert Syst. Appl.*, vol. 36, pt. 2, no. 2, pp. 3216–3222, Mar. 2009.

[41] Z. Kasirolvalad, M. R. J. Motlagh, and M. A. Shadmani, "An intelligent modeling system to improve the machining process quality in CNC machine tools using adaptive fuzzy petri nets," *Int. J. Adv. Manuf. Technol.*, vol. 29, no. 9/10, pp. 1050–1061, Jul. 2006.

[42] E. D. Kirby, J. C. Chen, and J. Z. Zhang, "Development of a fuzzy-nets-based in-process surface roughness adaptive control system in turning operations," *Expert Syst. Appl.*, vol. 30, no. 4, pp. 592–604, May 2006.

[43] M. M. Hanna, A. Buck, and R. Smith, "Fuzzy petri nets with neural networks to model products quality from a CNC-milling machining centre," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 26, no. 5, pp. 638–645, Sep. 1996.

[44] F. Dweiri, M. Al-Jarrah, and H. Al-Wedyan, "Fuzzy surface roughness modeling of CNC down milling of alumic-79," *J. Mater. Process. Technol.*, vol. 133, no. 3, pp. 266–275, Feb. 2003.

[45] C. Aliustaoglu, H. M. Ertunc, and H. Ocak, "Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system," *Mech. Syst. Signal Process.*, vol. 23, no. 2, pp. 539–546, Feb. 2009.

[46] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 2000.

[47] J. C. Chen and V. Susanto, "Fuzzy logic based in-process tool-wear monitoring system in face milling operations," *Int. J. Adv. Manuf. Technol.*, vol. 21, no. 3, pp. 186–192, Mar. 2003.

[48] D. Kim and D. Jeon, "Fuzzy-logic control of cutting forces in CNC milling processes using motor currents as indirect force sensors," *Precision Eng.*, vol. 35, no. 1, pp. 143–152, Jan. 2011.

[49] M. Elbestawi and M. Dumitrescu, "Tool condition monitoring in machining—Neural networks," in *Information Technology for Balanced Manufacturing Systems*. New York, NY, USA: Springer-Verlag, 2006, pp. 5–16.

[50] S.-P. Lo, "An adaptive-network based fuzzy inference system for prediction of workpiece surface roughness in end milling," *J. Mater. Process. Technol.*, vol. 142, no. 3, pp. 665–675, Dec. 2003.

[51] J. C. Chen and M. S. Lou, "Fuzzy-nets based approach to using an accelerometer for an in-process surface roughness prediction system in milling operations," *Int. J. Comput. Integr. Manuf.*, vol. 13, no. 4, p. 358, 2000.

[52] J. C. Chen and M. Savage, "A fuzzy-net-based multilevel in-process surface roughness recognition system in milling operations," *Int. J. Adv. Manuf. Technol.*, vol. 17, no. 9, pp. 670–676, May 2001.

[53] S. Wu and M. J. Er, "Dynamic fuzzy neural networks—A novel approach to function approximation," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 30, no. 2, pp. 358–364, Apr. 2000.

[54] M. Brezocnik, M. Kovacic, and M. Ficko, "Prediction of surface roughness with genetic programming," *J. Mater. Process. Technol.*, vol. 157/158, pp. 28–36, Dec. 2004.

[55] S. Achiche, M. Balazinski, L. Baron, and K. Jemielniak, "Tool wear monitoring using genetically-generated fuzzy knowledge bases," *Eng. Appl. Artif. Intell.*, vol. 15, no. 3/4, pp. 303–314, Jun.–Aug. 2002.

[56] O. Çolak, C. Kurbanoglu, and C. Kayacan, "Milling surface roughness prediction using evolutionary programming methods," *Mater. Des.*, vol. 28, no. 2, pp. 657–666, 2007.

[57] Z. G. Wang, M. Rahman, Y. S. Wong, and J. Sun, "Optimization of multi-pass milling using parallel genetic algorithm and parallel genetic simulated annealing," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 15, pp. 1726–1734, Dec. 2005.

[58] M. S. Shunmugam, S. V. Bhaskara Reddy, and T. T. Narendran, "Selection of optimal conditions in multi-pass face-milling using a genetic algorithm," *Int. J. Mach. Tools Manuf.*, vol. 40, no. 3, pp. 401–414, Feb. 2000.

[59] Y. Liu and C. Wang, "A modified genetic algorithm based optimisation of milling parameters," *Int. J. Adv. Manuf. Technol.*, vol. 15, no. 11, pp. 796–799, Sep. 1999.

[60] M. R. Razfar, M. Asadnia, M. Haghshenas, and M. Farahnakian, "Optimum surface roughness prediction in face milling X20Cr13 using particle swarm optimization algorithm," in *Proc. Inst. Mech. Eng., Part B, J. Eng. Manuf.*, Nov. 2010, vol. 224, no. 11, pp. 1645–1653.

[61] H. El-Mounayri, Z. Dugla, and H. Deng, "Prediction of surface roughness in end milling using swarm intelligence," in *Proc. IEEE SIS*, 2003, pp. 220–227.

[62] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Cambridge, MA, USA: MIT Press, 1992.

[63] J. R. Koza, *Genetic Programming*. Cambridge, MA, USA: MIT Press, 1992.

[64] A. M. Zain, H. Haron, and S. Sharif, "An overview of GA technique for surface roughness optimization in milling process," in *Proc. ITSim*, 2008, vol. 4, pp. 1–6.

[65] J. Kennedy, R. C. Eberhart, and Y. Shi, *Swarm Intelligence*. New York, NY, USA: Springer-Verlag, 2001.

[66] F. Cus, U. Zuperl, and V. Gecevska, "High speed end-milling optimisation using particle swarm intelligence," *J. Achievements Mater. Manuf. Eng.*, vol. 22, no. 2, pp. 75–78, Jun. 2007.

[67] A. Milani, "Online genetic algorithms," *Int. J. Inf. Theories Appl.*, vol. 11, pp. 20–28, 2004.

[68] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *Annals Math. Stat.*, vol. 37, no. 6, pp. 1554–1563, Dec. 1966.

[69] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518–1569, Jun. 2002.

[70] B. H. Juang and L. R. Rabiner, "Hidden Markov models for speech recognition," *Technometrics*, vol. 33, no. 3, pp. 251–272, Aug. 1991.

[71] J. A. Bilmes, "What HMMs can do," *IEICE—Trans. Inf. Syst.*, vol. E89-D, no. 3, pp. 869–891, Mar. 2006.

[72] A. Al-Habaibeh and N. Gindy, "A new approach for systematic design of condition monitoring systems for milling processes," *J. Mater. Process. Technol.*, vol. 107, no. 1–3, pp. 243–251, Nov. 2000.

[73] X. Li, S. K. Tso, and J. Wang, "Real-time tool condition monitoring using wavelet transforms and fuzzy techniques," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 30, no. 3, pp. 352–357, Aug. 2000.

[74] M. Pratama, M. J. Er, X. Li, L. San, J. O. Richard, L.-Y. Zhai, A. Torabi, and I. Arifin, "Genetic dynamic fuzzy neural network (GDFNN) for nonlinear system identification," in *Proc. Adv. Neural Netw.—ISNN*, vol. 6676, *Lecture Notes in Computer Science*, Jan. 2011, pp. 525–534.

[75] X. Li, M. J. Er, H. Ge, O. P. Gan, S. Huang, L. Y. Zhai, S. Linn, and A. J. Torabi, "Adaptive network fuzzy inference system and support vector machine learning for tool wear estimation in high speed milling processes," in *Proc. 38th Annu. Conf. IEEE IECON*, 2012, pp. 2821–2826.

[76] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, L. Zhai, S. J. Phua, J. Zhou, S. Linn, S. Huang, O. Massol, and S. Raj, "Flute based analysis of ball-nose milling signals using continuous wavelet analysis features," in *Proc. 11th ICARCV*, 2010, pp. 1359–1364.

[77] A. P. Varga and R. K. Moore, "Hidden Markov model decomposition of speech and noise," in *Proc. ICASSP*, 1990, vol. 90, pp. 845–848.

[78] K. F. Lee, "Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 4, pp. 599–609, Apr. 1990.

[79] L. Xiaoli and Y. Zhejun, "Tool wear monitoring with wavelet packet transform–fuzzy clustering method," *Wear*, vol. 219, no. 2, pp. 145–154, Sep. 1998.

[80] Z. Wang, W. Lawrenz, R. B. K. N. Rao, and T. Hope, "Feature-filtered fuzzy clustering for condition monitoring of tool wear," *J. Intell. Manuf.*, vol. 7, no. 1, pp. 13–22, Feb. 1996.

[81] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, L. Y. Zhai, H. Sheng, S. Lin, and O. P. Gan, "Fuzzy clustering of wavelet features for tool condition monitoring in high speed milling process," in *Proc. Annu. Conf. Prognost. Health Manage. Soc.*, Portland, OR, USA, Oct. 2010, pp. 1–5.

[82] X. Li, M. J. Er, B. S. Lim, J. H. Zhou, O. P. Gan, and L. Rutkowski, "Fuzzy regression modeling for tool performance prediction and degradation detection," *Int. J. Neural Syst.*, vol. 20, no. 5, pp. 405–419, Oct. 2010.

[83] B. Farran, A. Ramanan, and M. Niranjan, "Sequential hierarchical pattern clustering," in *Proc. Pattern Recognit. Bioinformat.*, 2009, pp. 79–88.

[84] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping Multidimensional Data*. New York, NY, USA: Springer-Verlag, 2006, pp. 25–71.

[85] M.-S. Yang, "A survey of fuzzy clustering," *Math. Comput. Modell.*, vol. 18, no. 11, pp. 1–16, Dec. 1993.

[86] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, L. Zhai, S. Linn, G. O. Peen, and C. C. Teck, "Application of classical clustering methods for online tool condition monitoring in high speed milling processes," in *Proc. 7th IEEE Conf. ICIEA*, Jul. 2012, pp. 1249–1254.

[87] M. A. Elbestawi, J. Marks, and T. Papazafiriou, "Process monitoring in milling by pattern recognition," *Mech. Syst. Signal Process.*, vol. 3, no. 3, pp. 305–315, Jul. 1989.

[88] M. Doostfatemeh and S. C. Kremer, "New directions in fuzzy automata," *Int. J. Approximate Reason.*, vol. 38, no. 2, pp. 175–214, Feb. 2005.

[89] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[90] P. Trahanias and E. Skordalakis, "Syntactic pattern recognition of the ECG," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 648–657, Jul. 1990.

[91] D. D. Lewis, "Evaluation of phrasal and clustered representations on a text categorization task," in *Proc. 15th Annu. Int. ACM SIGIR Conf. Res. Dev. Inform. Retrieval*, 1992, pp. 37–50.

[92] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, Dec. 2006.

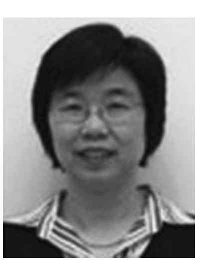
Amin Jahromi Torabi was born in Jahrom, Iran, in 1982. He received the B.Sc. and M.Sc. degrees in control engineering from Shiraz University, Shiraz, Iran, in 2004 and 2007, respectively. He is currently working toward the Ph.D. degree at Nanyang Technological University, Singapore.

He also serves as a Lecturer in the Persian Gulf University, Bushehr, Iran. He worked on many practical projects and industrial challenges during his academic studies in the Student Research Center of Shiraz University, "Radio Amatory Lab," and also SIMTech, Singapore. During his Ph.D. studies, he worked with some authors of this paper as an award winning team that secured the Institution of Engineers Singapore Prestigious Engineering Achievement Award 2011. His research interests include linear and nonlinear control theory and their applications, the application of intelligent systems, fuzzy logic and neural network-based systems, clustering algorithms, data mining, and data analysis, and the application of artificial intelligence theory in telecommunication systems.

Meng Joo Er (SM'07) received the Ph.D. degree from Australian National University, in 1992.

He is currently a Full Professor of electrical and electronic engineering with Nanyang Technological University, Singapore. He has authored five books, 16 book chapters, and more than 400 refereed journal and conference papers in his research areas of interest. His areas of research interests are computational intelligence, robotics and automation, sensor networks, biomedical engineering, and cognitive science.

Prof. Er was the recipient of the Institution of Engineers Singapore (IES) Prestigious Engineering Achievement Award 2011 in recognition of the significant and impactful contributions to Singapore's development by his research project entitled "Development of Intelligent Techniques for Modelling, Controlling and Optimizing Complex Manufacturing Systems." He is also the only dual winner in the Singapore IES Prestigious Publication Award in Application (1996) and IES Prestigious Publication Award in Theory (2001). He currently serves as the Editor-in-Chief of the International Journal of Electrical and Electronic Engineering and Telecommunications, an Area Editor of the International Journal of Intelligent Systems Science, an Associate Editor of 11 refereed international journals, and an editorial board member of the Electrical Engineering Times.



Xiang Li received the M.E. and B.E. degrees from Northeastern University, Shenyang, China, in 1987 and 1982, respectively, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2000.

She has more than 15 years of experience in research and applications of data mining, artificial intelligence, and statistical analysis, such as neural networks, fuzzy logic systems, data clustering, and multiple regression modeling. She is currently with SIMTech, Singapore.

Beng Siong Lim received the Ph.D. degree in the applications of computational intelligence for component design and tool engineering with a scholarship from the University of Nottingham, England.

He joined SIMTech, Singapore, in 1987, following the development of an aerospace flight simulator configuration and sales engineering system at Brighton University for Rediffusion Simulation at Crawley. His main research interest includes the development of evolutionary computation for performance degradation, characterization, and reference modeling.

Lianyin Zhai received the B.S. degree from Xi'an Jiaotong University, China, and the M.Eng. and Ph.D. degrees from Nanyang Technological University, Singapore.

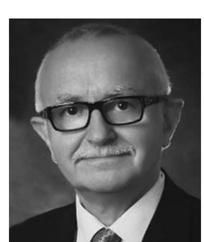
He is currently with Nanyang Technological University. His research interests include intelligent systems.

Richard J. Oentaryo received the B.E. (first-class honor) and Ph.D. degrees from the School of Computer Engineering, Nanyang Technological University, in 2004 and 2011, respectively.

He is currently a Research Fellow at the Living Analytics Research Centre, Singapore Management University. His research interests span neuro-fuzzy systems, social network mining, and brain-inspired architectures.

Gan Oon Peen received the Ph.D. degree from the National University of Singapore, Singapore, in 1997.

He is a Research Scientist and Group Manager at the Singapore Institute of Manufacturing Technology and a Technical Lead at the National Radion Frequency Identification Centre Singapore. His research interests are in the area of intelligent factory control and prognostic health management, particularly intelligent control, discrete event system modeling and control, data mining, operation research, and industrial automation.



Jacek M. Zurada (F'96) received the Ph.D. degree from the Technical University of Gdansk, Poland.

He is a Professor of electrical and computer engineering with the University of Louisville, Louisville, KY, USA. He served as Department Chair and is currently a University Scholar. He has authored several textbooks and over 360 publications in computational intelligence, image/signal processing, bioinformatics, and microelectronic systems that have resulted in about 7100 citations.