Role of Hidden-Markov Models
for Autonomous Diagnostics of Cutting Tools

Akhilesh Kumar
Industrial & Systems Engineering
Wayne State University
Detroit, MI 48202 USA
a.kumar@wayne.edu

Fling Tseng
Advanced Engineering
Ford Motor Company
Dearborn, MI 48121 USA
ftseng@ford.com

Ratna Babu Chinnam,Ph.D.
Industrial & Systems Engineering
Wayne State University
Detroit, MI 48202 USA
r_chinnam@wayne.edu

Abstract— **Despite considerable advances in sensing instrumentation and IT infrastructure, monitoring and diagnostics technology has not yet found its place in health management of mainstream machinery and equipment. The fundamental reason for this being the mismatch between the growing diversity and complexity of machinery and equipment employed in industry and the historical reliance on "point-solution" diagnostic systems that necessitate extensive characterization of the failure modes and mechanisms. While these point solutions have a role to play, in particular for monitoring highly-critical assets, generic yet adaptive solutions, could facilitate large-scale deployment of diagnostic and prognostic technology. We present the role of hidden-Markov models for autonomous diagnostics. The proposed methods have been tested on a CNC machining test-bed outfitted with thrust-force and torque sensors for monitoring drill-bits.**

Keywords- autonomous diagnostics; hidden Markov model; sequential clustering

I. Introduction

Condition Based Maintenance (CBM) is a philosophy that applies sensors to equipment for the purposes of monitoring, diagnostics, and prognostics, to facilitate optimal maintenance. CBM has the potential to greatly reduce costs by helping to avoid catastrophic failures and by more efficiently determining the intervals required for maintenance schedules [1]. The economic ramifications of CBM are many fold since they affect labor requirements, replacement part costs, routine maintenance scheduling, increased capacity, enhanced logistics, and supply chain performance [1,2]. Despite considerable advances over the last few decades in sensing instrumentation, signal processing algorithms, and information technology infrastructure, monitoring and diagnostics technology has not yet found its place in health management of mainstream machinery and equipment [6].

Diagnostics is the process of identifying, localizing, determining, and classifying the severity of equipment failure, whereas prognostics is the process of predicting the remaining-useful-life (RUL) [4]. Diagnostics is not only important but is a prerequisite for effective prognostics. The primary challenge is to achieve high degree of accuracy in reasoning out the health-state of the equipment given the sensory signals. The major technical challenges with effective diagnostics are as follows [1]: 1) Sensory signal statistics tend to be quasi-stationary and vary as a function of operating conditions and ambient conditions, 2) Machine character can be quite variable due to differences in machining, part-size variations, fastener tightness, wear variations, replacement-part variations, and aging, and 3) Features indicative of machine health can be obscured by signals from other sources, multitude of transmission paths, and by noise. In addition, historical datasets and cases when available for building diagnostic algorithms tend to be limited and not "labeled" in terms of fault progression and severity. CBM diagnostic techniques must be robust and effective under these conditions.

The traditional methods for diagnostics, leading to "point solutions", can be broadly grouped into two categories: *physics (or mechanistic) based* and *empirical based* [2]. The physics based methods involve extensive characterization of equipment to understand the different failure modes and their mechanisms, something tedious and resource intensive. Sensor selection, mounting, and feature selection are equally important and demanding issues. Physics based methods are economically justified when dealing with equipment that is pervasive (e.g., motors, pumps, generators, gear boxes and so on) and/or mission critical. The empirical methods often involve tracking of few critical features (based on failure mode) combined with simplistic thresholds set from experience. The extant literature is vast and reports good success in developing these point solutions [5, 6]. However, we need cost-effective technologies for monitoring a wide-array of equipment that is neither mission-critical nor pervasive. A further complicating factor is that industry, partly attributable to a growing push for mass customization, is building and employing more and more "custom" equipment, mostly ruling out the traditional "point solution" method. The goal then is to develop "generic" diagnostic and prognostic algorithms that are rapidly configurable, and adaptive (i.e., learn on-line using unsupervised learning algorithms) to facilitate effective and efficient large-scale deployment of CBM technology for a wide variety of equipment/assets. A study by NIST concluded that the availability of "generic" methods for effective diagnostics and prognostics and their reliability is a prerequisite for widespread deployment of CBM technology [7].

The concept of *autonomous diagnostics* is based on

unsupervised techniques. The term "unsupervised" implies ability to learn on-line without human supervision. Autonomous diagnostic methods learn gradually from the system onto which they are deployed. If developed, they can be deployed onto a variety of systems with ease, without requiring much equipment specific fine-tuning. This paper presents a practical framework for autonomous diagnostics based on *Hidden Markov Models* (HMMs). HMM is a finite-state machine that is also a doubly stochastic process involving at least two levels of uncertainty: a random process associated with each state, and a Markov chain that characterizes the probabilistic relationship among the states in terms of how likely one state is to follow another [8]. HMMs are known for their application in temporal pattern recognition such as automatic speech recognition [9, 10], handwriting, gesture recognition, economic and financial series analysis [11], and bioinformatics (e.g., EEG time-series clustering [12] and gene expression clustering [13]).

Given the success of HMMs with these applications, in particular with speech recognition that has a lot of similarity to machine diagnostics, there is hope that they will be equally effective in diagnostic applications. Two side-benefits are the existence of computationally efficient methods for system identification and computing of likelihoods using HMMs. They can also be used to build data-driven models of machines relieving somewhat the need to identify specific features in data to be used as health indicators [1]. We should however note that there are some notable differences between speech recognition and machine diagnostics [1]. For example, in speech processing the number of phonemes is a relatively small finite set. Furthermore, words which are constructed as sequences of phonemes also represent a finite (although large) set. In spite of this, the literature reports good results from application of HMMs for machine monitoring and diagnostics [1-2, 14-15]. However, almost all of this literature treats the task of developing diagnostics models as one of building classification models (i.e., supervised learning) relying on labeled training histories/datasets (in terms of fault progression and severity). On the contrary, the goal here is to build HMM models for diagnostics while working with unlabeled datasets, necessitating a "clustering" approach. However, a critical challenge in working with many clustering algorithms is of knowing in advance the target # of clusters.

To overcome aforementioned problem with clustering methods, this paper reports two HMM-based clustering methods with varying level of complexity. These methods 'generate' an effective labeling scheme for sensor signals, in turn promoting autonomy of the diagnostics engine. First method employs a competitive learning framework [14] whereas, the second method exploits inherent gradual deterioration of machine health and uses a sequential clustering [15] approach to drive the HMM-based time-series clustering. The underlying assumption is that the sensor signals are in the form of time-series (univariate or multivariate). While the manuscript focuses on monitoring cutting tools used by CNC machines (in particular drill bits), using a dynamometer to monitor thrust-force and torque on the cutting tool, the framework is relevant for monitoring a wide variety of equipment (e.g., rotary equipment employing vibration sensors) but might involve significant signal pre-processing and/or feature selection. While the monitored unit could be a component, sub-system, or a whole piece of equipment, the rest of this manuscript generally refers to the monitored unit as an asset or equipment.

Rest of the paper is organized as follows: section 2 briefly presents the background of HMM. Methods are discussed in section 3. Experimentation and results have been presented in section 4. Finally, conclusion and future research in section 5.

II. BACKGROUND OF HIDDEN-MARKOV MODELS

Hidden Markov model (HMM) is a finite-state machine that is also a doubly stochastic process involving at least two levels of uncertainty: a random *observation* process associated with each *hidden-state*, and a Markov chain, which characterizes the probabilistic relationship among the states in terms of how likely one state is to follow another. Note here that the "hidden state" of a HMM is not the same as the "health-state" of an equipment under diagnosis. In fact, in the proposed approach, a complete HMM will model a single health-state. In working with HMMs, the objective is to either characterize the hidden-states given the observation sequence or calculate the likelihood of the sequence given the HMM.

Let X_t denote the hidden state at time t and O_t the corresponding observation. Assuming that there are k possible states, we have $X_t \in \{1,...,k\}$. Let $\{o_1, o_2, ...o_T\}$ denote the observation sequence of random variable O_t. Note that o_i can be univariate or multivariate and that the length of the observation sequence, T, can be arbitrary. Characterization of a HMM is done through its parameters, $\lambda = (\pi, A, B)$. The parameters for a basic (first-order) HMM are the initial state distribution $\pi(i) = P(X_1 = i)$, the transition model $A = \{\alpha_{i,j}\} = P(X_t = j \mid X_{t-1} = i)$, and the observation model $b_j(o_t) = P(O_t = o_t \mid X_t = j)$, which is the probability of a particular observation vector at a particular time t for state j. The complete collection of parameters for all observation distributions is represented by $B = \{b_j(\cdot)\}$. A flexible representation of $P(O_t \mid X_t)$ as a mixture of Gaussians for observation vectors in $\Re^L$ is [16]:

$$P(O_t = o_t \mid X_t = i) =$$
$$\sum_{m=1}^{M} P(M_t = m \mid X_t = i) N(o_t; \boldsymbol{\mu}_{m,i}, \boldsymbol{\Sigma}_{m,i})$$

where $N(o_i; \boldsymbol{\mu}_{m,i}, \boldsymbol{\Sigma}_{m,i})$ is the Gaussian density with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$, M_t the hidden variable that specifies the mixture components, and $P(M_t = m \mid X_t = i) = c_{i,m}$ the conditional prior weight of each mixture component.

Estimation of parameters $\lambda = (\pi, A, B)$ is normally carried out through an iterative learning process. A-priori values of π, A and B are assumed and observation sequences are

presented iteratively to the model for parameter estimation. All the results reported in this paper are based on the use of the Baum-Welch's Expectation Maximization algorithm [17]. Estimation of $p(\boldsymbol{O}|\lambda)$ from given $\boldsymbol{O} = (o_1,...,o_T)$ is obtained using either the Forward procedure or the Backward procedure of the FB-algorithm [18]. Estimation of $p(\boldsymbol{O}|\lambda)$ is essential in building a HMM-based classifier.

III. AUTONOMOUS DIAGNOSTICS THROUGH HMM BASED CLUSTERING

As noted earlier, our objective is to develop effective diagnostic methods for tracking incipient equipment failures given unlabeled historical datasets (i.e., sensor signal histories). Given signal histories from identical assets that have undergone a particular type of incipient failure (i.e., distinct failure mode), the goal is to characterize the distinct health-states of the asset during the degradation from a state of perfect health to a state of total failure. This is critical to facilitate timely and optimal condition-based maintenance.

Given that there exists no labeled target outputs within the historical degradation datasets, we rule out the possibility of building a diagnostic classifier through supervised learning. Instead, we shall rely on unsupervised learning, in particular, clustering. We can however do better than pure clustering. While the historical datasets do not have labeled target outputs during the degradation process, they do provide us with the knowledge that the equipment was perfectly healthy at the very beginning and that the equipment has failed at the end. Given this information, we are recommending two model-based clustering approaches to develop the diagnostic model.

A. Clustering of Historical Data

Let us suppose that sensor signal histories are available for N identical assets subject to a common failure mode $(\boldsymbol{O}_1, \boldsymbol{O}_2,...,\boldsymbol{O}_N)$. The sensor signal history from asset n is of the form $\boldsymbol{O}_n = \{o_{n,1}, o_{n,2},...o_{n,T_n}\}$, where $o_{n,1}, o_{n,2},...o_{n,T_n}$ constitutes time-series segments sampled from the asset at regular intervals of time or usage (number of segments available from different assets until failure could be different depending the rate of degradation). Thus, $o_{n,1}$ is the first time-series historical segment from asset n. Note again that these time-series segments can be univariate or multivariate depending on the number of sensors employed and the signal processing and feature extraction procedures employed.

A. 1. Competitive Learning Approach

Step 1- Generation: We *randomly generate* a pre-specified number (say K) of HMMs (denoted $\{\lambda_1, \lambda_2,...\lambda_K\}$) with a pre-specified configuration (i.e., the number of hidden states and the observation model representation). During this step, all the competing HMMs start with uniform a-priori values for π, A and B. Subsequent steps involve initialization as well as iterative fine tuning of these parameters through competitive learning.

Step 2- Initialization: Each of the HMMs is *initialized* using a random sequence out of the N available sequences through very limited training (say, 2 or 3 iterations). Mathematically, this is equivalent to adjusting the HMM model parameters $\lambda = (\pi, A, B)$ towards maximizing $p(\boldsymbol{O}|\lambda)$. This initialization process better '*locates*' the HMMs so as to properly span the observation space, and in turn, dramatically improves convergence during the learning process.

Step 3- Ordering Phase: All temporal sequences available in the dataset are presented in a random sequence (constitutes one epoch) to the HMM pool for *competition* that involves calculation of log-likelihoods (i.e., $\ln P(O_i | HMM_j) \; \forall \; i, j$). HMM that wins the competition is allowed one iteration of *learning* (i.e., π, A, B adjustment) using EM algorithm.

Step 4- Consistency: Check for consistency or else go back to Step-3. Consistency is declared if every sequence is won by the same HMM for two consecutive epochs.

Step 5- Convergence Phase: Further update and fine tune the parameters of HMMs using the EM algorithm and its won sequences until no significant further improvements in log-likelihood are witnessed in successive epochs.

A. 2. Sequential Clustering

Step1- Initialization: The proposed algorithm initiates with construction of an HMM λ_1 with a pre-specified configuration, with the configuration to be "optimized" using cross-validation procedures outlined later. The parameters of HMM λ_1 (i.e., π, A and B) are initialized using infancy signal history from a random asset i (e.g., $o_{i,1}$ or $o_{i,2}$).

Step2- Characterize "Excellent" Health-state: We train HMM λ_1 to adjust (π, A, B) towards maximizing $p(O|\lambda)$. For training, signals from earliest stages of the asset's life are considered, e.g. $o_{i,1}$ or $o_{i,2}$ of all N assets. Because of such factors as "wear-in", our research suggests that $o_{i,2}$ or $o_{i,3}$ can be more reliable than $o_{i,1}$ for characterizing the "excellent" health-state (depending of course on such factors as time-between two consecutive samples and typical life of the asset). Once trained, all N sequences are evaluated with this trained HMM and similarity is observed based on log-likelihood values, in particular, their distribution. The log-likelihood similarity threshold or cutoff for characterizing the "next" health-state is set at $\mu - k\sigma$, where μ and σ denote the mean and standard deviation of the log-likelihood value distribution, respectively, and k an integer. The higher the value of k, the lower the resolution of characterization of health-states.

Step 3- Identify Candidate Signals for Characterizing "Next" Health-State: Segments from each asset i that just missed the log-likelihood threshold of the previous health-state are identified as candidate training signals for the "Next" health-state. A new HMM λ_2 is constructed and trained with these newly identified training signal segments.

Step 4- Termination: Step 3 is repeated until all sensor signal segments of each asset have been characterized.

B. Labeling Health-states based on HMM Clusters

Let us suppose that the clustering process yielded *M* distinct health-state HMMs, representing the distinct temporal dynamics of the sequences that make up the respective clusters. Thus, each HMM represents one health-state sequentially from "Excellent" to "Failure".

C. On-line Diagnostics using HMM-Based Classifier

Once the distinct health-sates are characterized, during on-line diagnostics, given the sensor signal segment, determination of "current" health-state involves calculation of the log-likelihood with respect to all characterized HMMs. The health-state corresponding to the HMM with the largest log-likelihood is the estimated health-state.

IV. CASE STUDY: DRILLING PROCESS DIAGNOSTICS

A. Experimental Setup

Drilling process, a commonly used machining process, is selected here as a test-bed for validating the proposed autonomous diagnostics framework. The diagnostics objective is to assess the health or well-being of the drill-bit during the machining process by utilizing thrust-force and torque signals captured by a dynamometer during the drilling cycle. Tests were conducted on HAAS VF-1 CNC Machining Center with Kistler 9257B piezo-dynamometer at 250Hz to drill holes in ¼ inch stainless steel bars. High-speed twist drill-bits with two flutes were operated at feed rate of 4.5 inch/minute and spindle-speed at 800 rpm without coolant. Each drill-bit was used until it reached a state of physical failure either due to excessive wear or due to gross plastic deformation of the tool tip due to excessive temperature (from excessive wear). Fourteen drill-bits were used to generate the signal histories necessary for building the diagnostics model.

B. Observation Sequence

A sequence is defined here as one that covers an individual hole. Due to bit wear and non-uniformity of the work piece surface, the actual time necessary to drill a hole varies. This results in sequences of different lengths. Thrust-force (Newtons) and torque (Newton-meters) signal amplitudes are usually quite different. To improve the convergence properties of the EM algorithm used for training the HMMs, the observational sequences are all normalized to mean zero and standard deviation of unity for both thrust-force and torque. Observation sequences that are presented to HMMs are not subjected to any transformation other than this normalization. Figure 1 illustrates a joint plot of normalized thrust-force and torque signals during a particular hole.

C. Labeling Health-states

[14] subjected 10 HMMs to competitive learning, only 3 survived the learning process (the remaining 7 HMMs did not win the competition for a single sequence). These 3 HMMs were able to cluster the thrust-force and torque observation sequences from individual drill-bits into 3 health states. HMM-1 ended up representing the "good", HMM-10 represented the "medium", and HMM-9 the "bad".
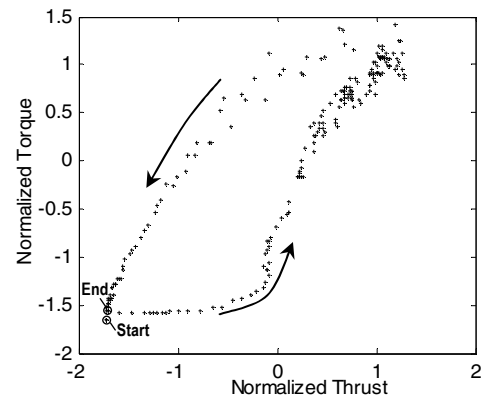

Fig. 1. Joint plot of normalized thrust-force and torque signals during a particular hole. Note the overall shape in spite of local randomness.

[15] initialized "excellent" health-state using signal segments from the second hole of the drill-bit set. Continuing with the steps outlined in Section III.A.2, the rest of the health-states are characterized as well, using for establishing the log-likelihood threshold. Overall, sequential clustering has yielded three health-states representing "good", "medium" and "bad".

While different drill-bits have spent different number of holes in the different health-states, all drill-bits have gone through all the three health-states prior to failure. A "dynamic" policy of replacing the drill-bit once entering the final health-state would not have resulted in any failures. Thus, enhancing the value of tracking the health-state of the drill-bit on-line for efficient replacement.

D. Cross-Validation and Testing Process

While the results reported in Section IV.C are very good, they are not necessarily reproducible. This is attributable to the fact that data from all the drill-bits were employed for building the diagnostic model as well as testing its performance. To evaluate the generalization performance of the proposed method, we now report results from cross-validation and testing experiments.

For competitive learning [14], the overall classification accuracy was 80% for the testing data sets (calculated over the 14 cross validation passes). The process for calculating these results is as follows. When the model was presented with data from an individual drill-bit, as long as the model did not suggest reverse jumps (i.e., a jump from a 'medium' state to a 'good' state), the accuracy was considered to be 100%. The higher the number of reverse jumps, the lower the accuracy. While two hidden-states were quite inadequate based on the observed log-likelihoods, three hidden-states were acceptable, however, four hidden-states provided very good performance and achieved nearly the same performance achieved by HMMs with nearly double the number of hidden states. [14] used four hidden-states for all the HMMs.

[15] proposed more stringent criteria for convergence evaluation criteria to evaluate the performance of the model.
1) No "reverse jumps", meaning that the drill-bit cannot enter a state and then revert to a previous state. If 2 or more

reverse jumps are noticed in the validation cycle, the model building run will be terminated.

2) At least two health-states have to be detected. If during the sequential clustering process no more than 1 HMM model was obtained, the run would be terminated.

Performance Criteria: The performance is judged based on the 4 drill-bits in the testing group, based on three criteria:

1) Fraction number of runs that did not produce any reverse jumps- $R(Q)$.

2) Fraction number of drill-bits that did enter all the identified health-states characterizing the degradation from a state of perfect health to a state of total failure- $M(Q)$.
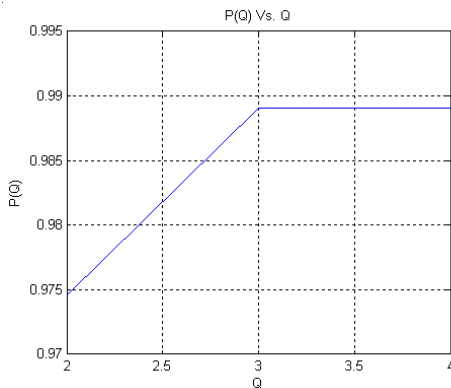

Fig. 2. Plot of Q vs. $P(Q)$.

Given the number of hidden-states, Q, combining these performance metrics (i.e, $R(Q)$ and $M(Q)$) yields an overall performance measure. While it is certainly possible to increase the importance of certain measures over others, we chose the following multiplicative model:

$$P(Q)=R(Q)*M(Q)$$

In our tests, we reasonably varied the number of hidden-states, Q, from 2 to 8 for full $^{14}C_{10}$ fold cross-validation.

Figure 2 clearly suggests that for this application, three hidden-states are best and adequate for configuring the HMMs for health-state characterization. The overall performance is quite impressive at nearly 99%, even under testing. In theory, it is also possible to vary Q for the different health-state HMM models.

Certainly, sequential clustering is least demanding and more promising. Further, the competitive learning process is tedious in particular with HMMs and there are issues with convergence and initialization.

V. CONCLUSIONS AND FUTURE RESEARCH

In the context of condition-based maintenance, the reported HMM diagnostic models allow us to overcome the tedious and often impossible task of "labeling" dataset health-states, and hence, improve autonomy of techniques for diagnostics. On the contrary, traditional HMM based diagnostics frameworks often employ a "classification" framework that strictly requires labeling. The results from the drilling process case study are extremely satisfactory. Both models were able to successfully cluster and recognize variable length and bivariate time-series sequences that are non-stationary in nature. It is not clear at this point if these models will yield satisfactory results when spectral properties are of more importance over temporal aspects of the sensory signals. However, the speech processing community has successfully employed HMMs for modeling dynamics of spectral features derived from speech signals. Future research will study other types of rotary equipment as well as different sensor settings. Future research will also consider equipment prognostics.

REFERENCES

[1] C. Bunks, D. McCarthy, and T. Al-Ani, "Condition-based maintenance of machines using hidden Markov models," Mechanical Systems & Signal Processing, vol. 14,(2000) 597-612.

[2] F. Camci, *Autonomous, Process Monitoring, Diagnostics, and Prognostics Using Support Vector Machines and Hidden Markov Models*, Ph.D. Dissertation, Industrial and Manufacturing Engineering, Wayne State University, Detroit, MI, USA, 2005.

[3] Harbor Research Pervasive Internet Report, Approaching Zero Downtime: The Center for Intelligent Maintenance Systems, April 2003.

[4] K. Maynard, C. S. Byington, G. W. Nickerson, and M. V. Dyke, Validation of Helicopter Nominal and Faulted Conditions Using Fleet Data sets, Proc. of the Intnl. Conf.on Condition Monitoring, UK, (1999) 129 –141.

[5] Begg, C., Merdes, T., Byington, C.S, and Maynard, K.P., Mechanical System Modeling for Failure Diagnostics and Prognosis, *Proc. of Maintainability and Reliability Conference (MARCON 99)*, Gatlinburg, Tennessee, May 10-12, 1999.

[6] G. J. Kacprzynski and M. J. Roemer, "Health Management Strategies for 21st Century Condition-Based Maintenance Systems," *13th International Congress on COMADEM: Condition Monitoring and Diagnostic Engineering Management,* Houston, TX, December 3-8, 2000. (http://www.ascturbo.com/download/HMStrategies.PDF)

[7] NIST-ATP, "NIST-ATP Workshop on Condition-Based Maintenance," NIST-ATP CBM Workshop Report, Atlanta, GA November 1998.

[8] Hidden Markov Model. McGraw-Hill Dictionary of Scientific and Technical Terms. Retrieved December 10, 2007, from Answers.com Web site: http://www.answers.com/topic/hidden-markov-model

[9] I. Sanches, "Noise-compensated hidden Markov models," IEEE Transactions on Speech and Audio Processing, vol. 8, (2000) 533-40.

[10] B. D. Womack and J. H. L. Hansen, "N-channel hidden Markov models for combined stressed speech classification and recognition," IEEE Transactions on Speech and Audio Processing, vol. 7, (1999) 668-77.

[11] S. M. Focardi, "Clustering economic and financial time-series: Exploring the existence of stable correlation conditions," The Intertek Group 2005.

[12] Geva, A. B. and Kerem, D. H., Brain state identification and forecasting of acute pathology using unsupervised fuzzy clustering of EEG temporal patterns, in *Fuzzy and Neuro-Fuzzy Systems in Medicine*, edited by H Teodorescu, A Kandel and L C. Jain, (1998) 57–93 (CRC Press: USA).

[13] Joseph, Z., Gerber, G., Gifford, D. and Jaakkola, T., A new approach to analyzing gene expression time-series data, in *The Sixth Annual International Conference in esearch in Computational Molecular Biology*, 2002.

[14] Chinnam, R.B. and Baruah, P., "Autonomous diagnostics and prognostics in machining processes through competitive learning-driven HMM-based clustering," International Journal of Production Research, 47:23, 6739- 6758, 2009.

[15] Kumar, A., Tseng, F., Guo, Y., and Chinnam, R.B., "Hidden-Markov Model Based Sequential Clustering for Autonomous Diagnostics," Proc. of International Joint Conference on Neural Networks, Hong Kong, June 1-6, 2008.

[16] Murphy, K., *Dynamic Bayesian Network: Representation, Inference, and Learning,* PhD. Dissertation, University of California, Berkeley, 2002

[17] Poritz, A. B., Hidden Markov Models: A Guided Tour, in ICASSP, 1988.

[18] Blimes, J., A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models, Technical Report, University of Berkeley, USA, 1988.