Multi-Class Classification Approach based on Fuzzy-Filtering for Condition Monitoring

Hammoud Aljoumaa and Dirk Söffker

***Abstract*—Analytical model-based methods have been developed during the last decades to achieve the goals of fault diagnosis of systems. One of the drawbacks of these methods, is the necessity of precise models of the considered system in order to design an appropriate fault detection/diagnosis system. This strong assumption can not be fulfilled for all cases. Additionally, different models have to be defined for distinguishing different states of machines operation. Qualitative model-based methods and also signal-based methods avoid this problem due to their different principle concepts of modeling. This contribution deals with the idea of combining qualitative model-based methods using fuzzy logic and statistical methods describing signal properties. The desired goal is to design a condition monitoring system based on suitable and available signals, related measurements experiments, and classifying information of the system to be monitored. The key idea of this contribution is the generation of a set of features to distinguish related different states of the system. For validation of the developed method, experimental data are used from an experimental studyed friction and wear processes of a metal surface allowing the distinction of different wear states. The developed method shows good ability to distinct the related states of wear.**

***Index Terms*—condition monitoring, pattern recognition, machine learning, feature extraction, fuzzy logic, fault diagnosis.**

I. Introduction

THE increased attention of monitoring and classification systems rests upon a goal to define operational states of a system. Therefore the following tasks/sub goals have to be achieved:

1) Prediction of the faults and failures during an operation of the system before its occurrence, in order to reduce the downtime of system and to decrease the possibility of production losses,
2) Improvement of the control of quality of products, which are closely linked to the operation conditions of the system,
3) Reduction of maintenance costs, and
4) Avoidance of unplanned downtimes of the system [13], [14], [15], [16].

Assuming measurements as data and related classifications of available states a suitable knowledge base has to be built connecting measurements and classifying states. Therefore, a theoretical framework is needed to enable the possibilities of linguistic/gradual expression of states in mathematical and logical terms. This can be achieved by using *membership functions* introduced by fuzzy logic. The membership function allows linguistically to express and to distinguish different states, its related numerical ranges and also gradually to define an amount of occurrence of any event in each individual states [1], [5], [4]. Many signal/machine learning-based approaches have been introduced in the last decades such as artificial neural networks (ANNs), support vector machine (SVM), and the related extensions to connect data and states [13], [14], [15], [17], [18], [19], [20], [21].
Despite the advantages of learning, generalization, fast convergence, and high accuracy provided by the mentioned methods and newest research; the artificial neural networks suffer from a complexity problem due to their internal structure [2], [1], [6], [11], [7], [8]. Whereby the number of processing units in each of the input and output layers are defined according to the number the variables/features in training datasets and dependant on the number of predictors for the considered issue. The number of processing units to be used in the hidden layer depends on the character of the application and system to be modeled. The number of the processing units in the hidden layer is an essential factor to define the ability of the designed neural network to model the system and its complex states. Increasing the number of these units causes increasing of model's ability the more complex states. The number of the processing units in the hidden layer must be adjusted carefully. The inaccuracy of process's adjustment leads to:

1) *Overfitting/overtraining* problem indicating that the neural network is designed more complicated than necessary, and
2) *Underfitting/undertraining* can be appeared if the design of network is simplier or not satisfactory ; therefore, the designed network will not be able to detect the full dataset expressing the complex situation of the considered system [6], [11].

For overcoming the problem of the overfitting/overtraining, several techniques are suggested such as "Jitter" technique based on the using the artificial noise beside the input data/signal to train the network and then to get smoother mapping in input-output space or faster convergence that is based on the stopping of iterations by using the cross validation, split-sample, or bootstrapping during the training phase in suitable time to get the satisfactory network without the reaching to overfitting time [6], [11], [7]. These techniques will lead to time consuming training because more time is needed to adjust/optimize the structure of ANNs model [6], [11].
The disadvantages of the SVMs methods can be summarized as follows:

- Multi-class applications and non separable data issues realted to SVMs are indirect. Because these classifiers are binary, only for separable classification problems and a direct decision function used in their structure.
- Problem *slowness* exists during the training phase. Because principle of this process is finding of solution

Hammoud Aljoumaa and Dirk Söffker, Chair of Dynamics and Control, University of Duisburg-Essen, Campus Duisburg, Engineering Faculty, Lotharstr. 1-21 47057, Duisburg, Germany, Email:{hammoud.aljoumaa;soeffker}@uni-due.de.

(Advance online publication: 10 February 2011)

of the quadratic programming/optimization problem for the large number of variables to get the optimal values of SVs [2], [22], [23], [11], [7], [12].

According to the above mentioned aspects, the developed approach tries to provide suitable solutions based on the following ideas:

1) The structure of the suggested classifying approach is based on a suitable number of features, which are generated/selected as a compact set. The resulting model is not based on any parameters. The training phase appears as a selection process of the useful features (later defined as main set) from all possible features (later defined as initial set), but not an optimization process of the parameters. This advantage of the proposed approach is a main property and different to known approaches.
2) The proposed approach uses the suggested fuzzy mapping method achieved to classify the considered samples into multi-classes (including binary-classes) results.

II. Fuzzy-filter approach

A. Basic ideas

1) Requirements of basic ideas: The requirements of the suggested approach are as follows:

- Measurements of suitable signals generated from a considered system are available.
- Related available information and expertise gained/generated by/from a human classification and/or related previous observations/monitoring are assumed asavailable.

According to these requirements, the suggested approach (figure 1) can be understood as a *signal-based modeling* process and also as a *machine learning* task. Dealing with the considered system to be monitored; the related information, expertise, and knowledge about related states of the considered system are used to generate a data matrix to present m related states. The term *state* here is understood as a linguistic expression summarizing the conditions operation/related states of the considered system. The linguistic expression is usually defined according to the human classification. So ”healthy” or ”fault-free” expressions can be used to represent the situation of regular operation of the system and the status of operation in abnormal state can be expressed by ”unhealthy” or ”faulty”. The distinction of states allows a linguistic division of the related states, namely (”fault-free” and ”faulty”), into sub-states such as ”healthy 1”, ”healthy 2”, etc., or ”faulty 1”, ”faulty 2”, etc. Usually, the generated data matrix can be characterised by *raw* data that can not give direct information able to distinguish the related states of the system. Thus the *rawness* property should be removed by transformation of this data from its original form to a new and more useful form represented by a new quantity so-called *feature*. In this context, the term feature is understood as a synonym of an input variable/attribute/quantity to be able to highlight important relationships/underlying representations inside the raw signals with related states of the considered system in a best possible form of interpretation [1], [9].

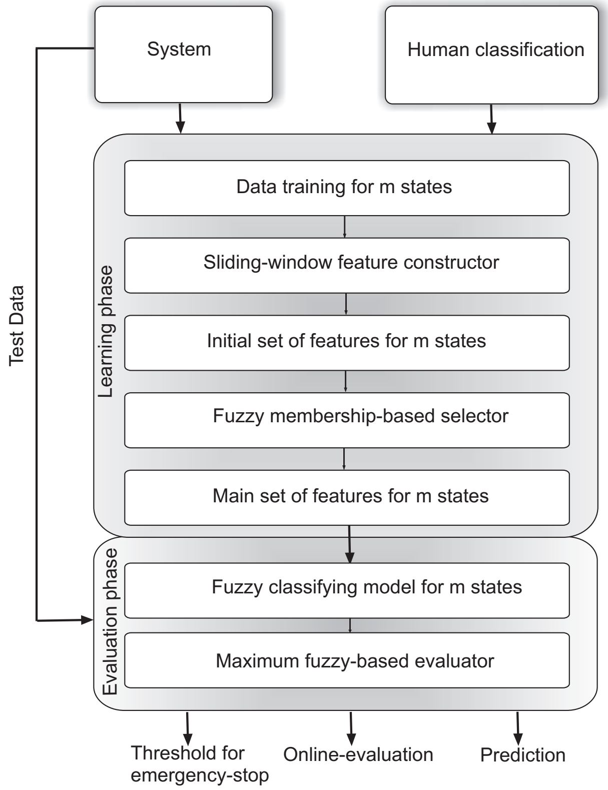

Fig. 1. General structure of proposed fuzzy-filter approach

2) Defining a suitable feature space: The transformation process is usually called *feature extraction*, which can be divided into two steps; *feature construction* and *feature selection* processes [1]. In the context of this contribution two sets of features are used; the first set called *initial set*, which consists of n possible relevant and irrelevant features, and the second set called *main set*, which consists of p relevant features. Relevant and irrelevant terms are identified according to the ability of the feature to distinguish the related states of the considered system. A distinguishing ability ω_{AD} is proposed as an index of the relevance and irrelevance of the feature. Suggested types of the used features for the proposed approach can be statistical, mathematical, geometrical, or any other type to be related to the nature of the considered application. The suggested feature construction process in the proposed approach is called *feature constructor* using *sliding-window* principle to generate the initial set of the features for the m related states. The distinguishing ability ω_{AD} of each feature in the initial set is calculated by using *fuzzy membership-based selector* to keep the relevant features and to delete the irrelevant features.

3) Building the classifying model: The output of the fuzzy membership-based selector is the main set of the features for the m related states. The features of the main set and the values of its related distinguishing abilities are used to build the *fuzzy classifying model* .

4) Evaluation process: For evaluation process, the fuzzy classifying model is used to calculate the p*m membership values for the considered data. These membership values are used as input of a *maximum fuzzy-based evaluator* to indicate the related state of the data, so that this information can be helpful to trigger the tasks as thresholds for emergency-stop, online-evaluation, and also for prediction.

In the next subsections these ideas and concepts will be explained in detail.

B. Feature constructor

The proposed mechanism of feature construction (figure 2), is used to build a vector of n features. The key element within this procedure is a sliding window with s elements to be able to scan a defined data vector of f elements length continuously by shifting one value at equal times. Once the sliding window shifts each time, the n features vector, so-called *Hybrid system state vector* [9], is generated. The term of *"Hybrid"* denotes that the vector consists of several types/different quantities of features. The *Hybrid system state vector* description helps to view the system from a problem/task-oriented side. The used principle of sliding window with s elements to construct the hybrid system state vector guarantees to generate the values of the features representing different areas of the changes the signal.

C. Fuzzy membership-based feature selector

The proposed approach uses only the main set of the features to build the fuzzy classifying model, therefore the separation between relevant features and others of the initial set of the features is needed. For separation the following procedure (figure 3) is proposed:

- Building of a suitable fuzzy classifier for each feature in the initial set,
- Selection of feature with certain value of ω_{AD}, which is defined according to predefined hypotheses space, and adding it to the main set,
- Fuzzy classifiers of the features of the main set and its related values of ω_{AD} are used to build the fuzzy classifying model.

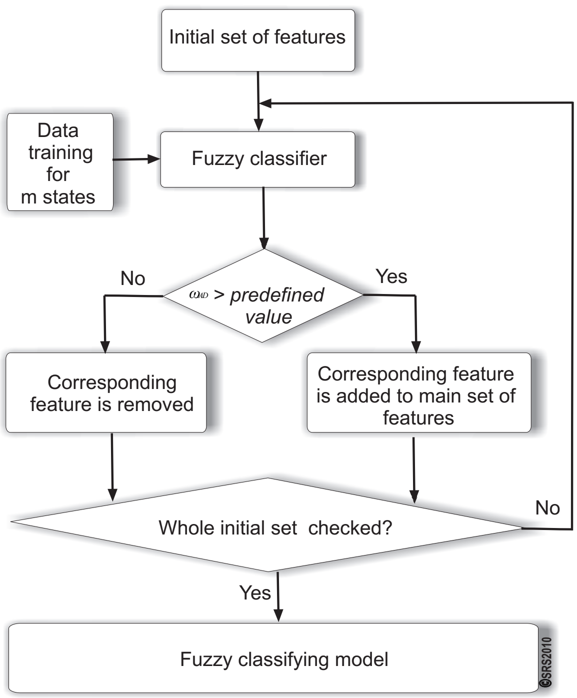

Fig. 3. Fuzzy membership-based feature selector

D. Fuzzy mapping

The proposed approach generates basically feature and state spaces that should be mapped to achieve the goal of classification. The proposed fuzzy-based mapping/fuzzification process is based on the following function (figure 4)

$$\mu_k\left(x_{ji}\right) = \left|1 - \left(\left(x_{ji} - \overline{x_{jk}}\right) * \alpha_d\right)\right| \tag{1}$$

Here $\mu_k\left(x_{ji}\right)$ denotes the membership value of value feature j of sample i for the state k, $\overline{x_{jk}}$ the value of mean characteristic of range of feature j for the state k. The value of $\overline{x_{jk}}$ represents the center of membership function, which is denoted by C_{S_k}, of the state k, $i : 1 \rightarrow f - s$, $j : 1 \rightarrow p$, α_d denotes to the factor the decrease in the membership value according to distance between the considered sample and the center of corresponding state. The values of features at points A, B and C, as illustrated in figure (4), are presented as the example for the using of the proposed mapping procedure to calculate its membership values as follows:

1) The membership values of the point A are calculated by using the equation (1) only for the states 2 and m; the membership values are zeros for other states.
2) The membership value of the point B is calculated by using the equation (1) only for the state 2; the membership values are zeros for other states.
3) The membership value of the point C is calculated by using the equation (1) only for the state $m - 1$; the membership values are zeros for other states.

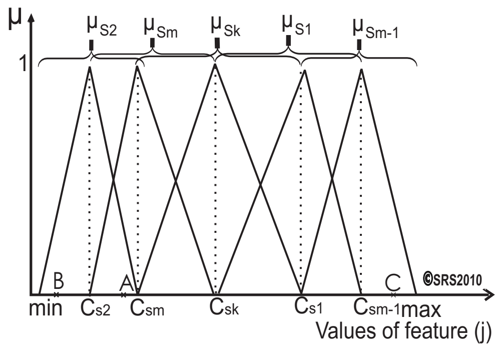

Fig. 4. Fuzzification process

The sequence of membership functions of states with respect to the feature axis in figure (4) is according to the ascending order of the values of mean characteristic of states in the range of feature j.

E. Maximum fuzzy-based evaluator

By analysis of the samples the considered data using the fuzzy classifying model, the p*m membership values are generated. The desired goal is to fuse these values to one value expressing the state. Therefore the procedure of the fuzzy-based evaluator using the maximum operator principle is developed by the following steps (as illustrated in 5):

- Calculation of the p*m membership values of the related sample by using fuzzification process as described in subsection (II-D),

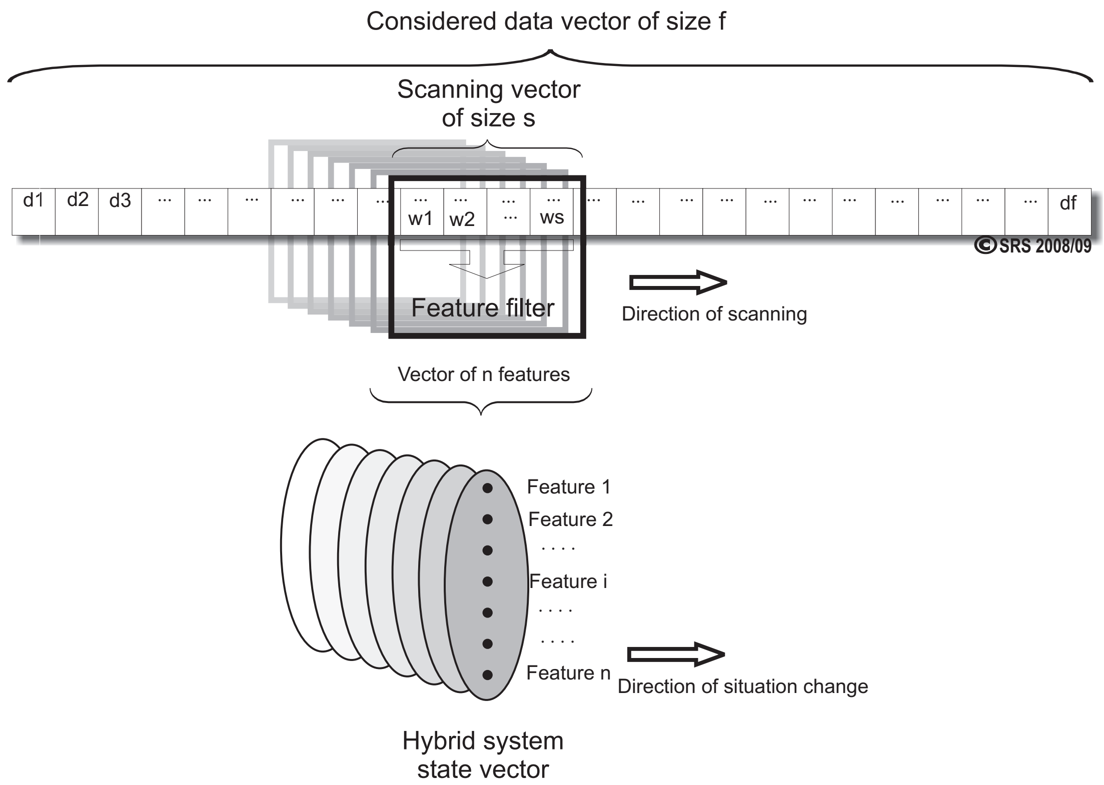

Fig. 2. Sliding-window feature constructor

- Calculation of a final membership value of the considered sample for each state using the following equation

$$\mu_{Final}\left(x_{ik}\right)=\sum_{j=1}^{p}\mu_{k}\left(x_{ji}\right)*\omega_{ADj}, \tag{2}$$

where ω_{ADj} is a weight of the j^{th} feature, and
- Definition of a final state of considered sample using the equation

$$S_{Final}\left(x_{i}\right)=S(max\left(\mu_{Final}\left(x_{ik}\right)\right)). \tag{3}$$

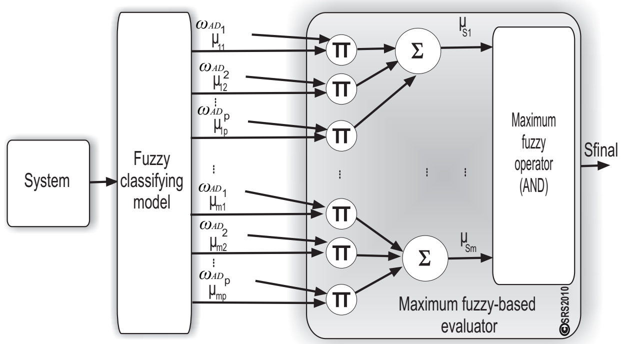

Fig. 5. Maximum fuzzy-based evaluator as decision unit

III. Experimental results

A test rig is used for studying/analysis friction and wear processes of metal surface (see figure (6)) related to changes of conditions of operation such as changes of lubrication and temperature, etc. based on analysis of the pressure, force, and acceleration signals generating from this system. According to the results, the best signal being is determined and used to build the classifying model. Therefore temperature, pressures, and force as well as acceleration signals generated from the system can be used to be analyzed.

A. Learning phase

In previous examinations it has been shown that the pressure signal is the suitable signal to be used for further examinations [9], [10]. Based on the evaluation of experiments, the five related states (m=5) of surface conditions should be distinguished as follows:

1) Regular 1: Indicating the stable operation, represented by green color.
2) Regular 2: Indicating the stable operation with minor changes, represented by cyan color.
3) Regular 3: Indicating the stable operation with acceptable changes, represented by blue color.
4) Abnormal 1: Indicating the abrupt changing surface conditions with acceptable changes, represented by yellow color.
5) Abnormal 2: Indicating the Abrupt changing surface conditions with non-acceptable changes, represented by red color.

The related five data sets using raw pressure signal as training data are used according to the subsection (II-A1). An initial set consisting of (n=17) statistical features is used to build the feature vector/the hybrid system state vector as an output of the sliding window-based feature constructor as explained in procedure (II-B). This initial set is used as input of the

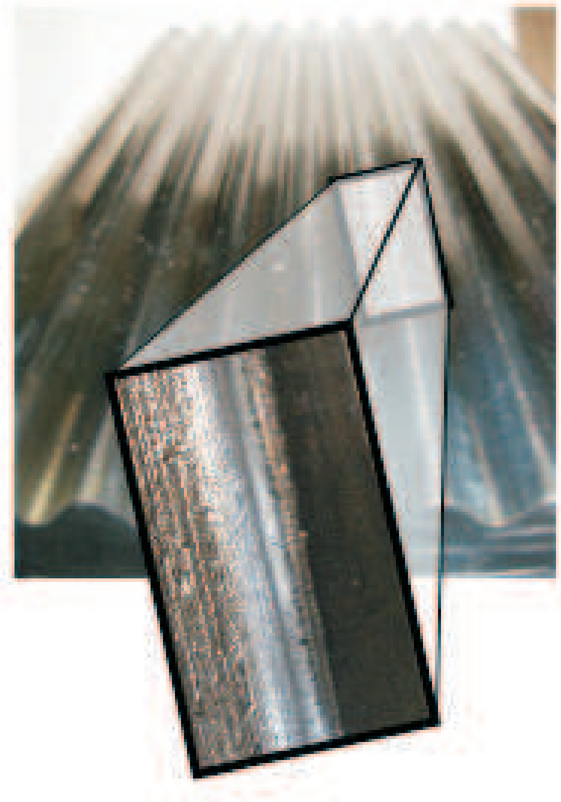

Fig. 6. Metal surface to be analysed depeding on the changes of conditions of operation

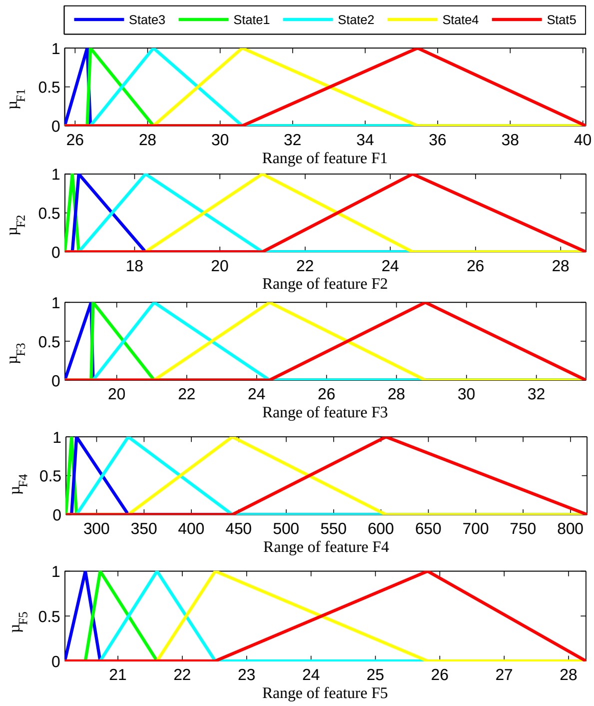

Fig. 7. Fuzzy classifiers/membership functions of the features of the main set (p=5)

fuzzy membership-based selector (as explained in procedure (II-C)) to generate the main set , which consists of the (F1, F2, F3, F4, and F5) features. The selection of the features (F1, F2, F3, F4, and F5) is chosen such that the related distinguishing ability is greater than a given, suitably chosen predefined value, like 0.7 as example. Based on the procedure explained in the subsection (II-D), the fuzzy classifier (as illustrated in Figure (7)) of each feature in the main set is built. The related values of the distinguishing ability each of five features (F1, F2, F3, F4, and F5) are illustrated in (I).

TABLE I
WEIGHTS OF THE P FEATURES

$Features$	$F1$	$F2$	$F3$	$F4$	$F5$
$Weights$	0.76	0.72	0.86	0.72	0.82

B. Validation phase

For validation of the proposed algorithm, the data set is used as input of the fuzzy classifying model and by using the procedure as described in subsection (II-E). The result of the evaluation is illustrated in figure (8). The result is consistent with the human classification to the changes of the states for the related data. According to the human classification, the erosion rate increases due to the increase of damage of metal surface combined with the continuous operation. This classification is illustrated through a gradual disappearance of colors, namely green, blue, and cyan colors, presenting the different the states of regular 1, regular 2, and regular 3 of erosion rate and a gradual appearance of colors, namely yellow and red colors, representing the states of abnormal 1 and abnormal 2 of erosion rate with the progress of run-time (see first five rows of figure (8)). The human classification for the state changing of the data used for validation can be seen based on the results of automated evaluation through an approximate transition of the changing state from states 1 to state 2, and 3 to states 4 and 5 (see last row in figure (8)).

C. Testing phase

To test the suggested algorithm two datasets are used. According to the human classification the classifications resulting from these datasets are as follows:

1) In the first one, the erosion rate increases due to the increase of damages of the metal surface combined with the continued operation, and
2) In the second one, the erosion rate is within regular operation range.

- Testing based on dataset number 1
  The results of the evaluation using the first dataset are illustrated in figure (9). As it can be seen from the first five rows of figure (9), a gradual disappearance of the related colors, namely green, blue, and cyan colors, presenting the states of regular 1, regular 2, and regular 3 of erosion rate and a gradual appearance of colors, namely yellow and red colors, representing the states of abnormal 1 and abnormal 2 of erosion rate changes with the progress of run-time. Additionally an approximate and continuously transition of the states from states 1, 2, and 3 to states 4 and 5, as it can be seen from the last row in figure (9). This result is consistent with the human classification to the changes of the states for the related data.
- Testing based on dataset number 2
  The results of evaluation of the second test dataset are

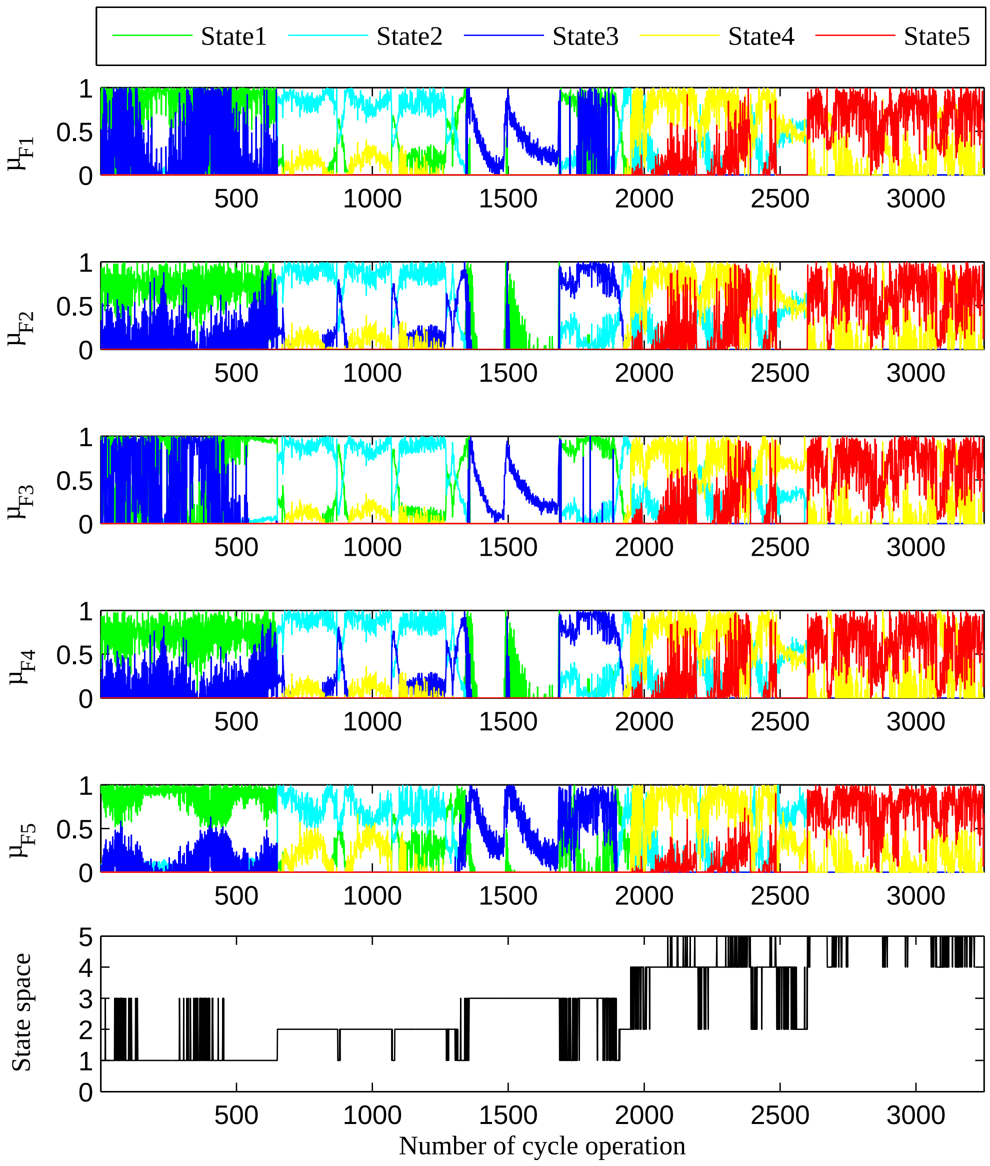

Fig. 8. Results of the evaluation for the validation dataset based on the proposed algorithm

illustrated in figure (10). This results are also consistent with the human classification to the changes of states for this data, which can also be seen by the complete appearance of the colors green, blue, and cyan colors, represented the states of regular 1, regular 2, and regular 3 of erosion rate as in the five rows in figure (10). For the state space, namely the last row of figure (10), it is observed that the change of state is approximately between the states 1, 2, and 3.

IV. Conclusions and Summary

This contribution introduces in detail the development of a new adaptive classifying algorithm. The proposed algorithm uses the combination of qualitative model-based methods based on fuzzy logic and signal-based methods using statistical methods. The goal of this combination is to convert automatically raw data into a new presentation achieving automatic classification tasks. For initial training of the algorithm an originary related multi-class classification of the training data is necessary. The generation of classifying features from this is automated by the proposed approach. Therefore the approach provides a new definition (equation 1) for automated setting of triangular fuzzy membership function parameters depending on the suggested parameter α_d. The proposed definition differs from the usual one based on three numerical values to be calculated manually. The proposed definition the value α_d can be automatically calculated during the learning phase according to the range of the considered state of each feature's range and also to the relations between the ranges of the m related states for the range of the considered feature. It works therefore as an important algorithmic element for automatic data-driven modeling. With the use of the Maximum fuzzy-based evaluator as decision unit; a pragmatic and systematic decision of the overall state based on the automated generated model can be also automatically achieved. The used research example is used to distinguish surface conditions of sliding metal parts. A macroscopic variable (here the hydraulic pressure used for

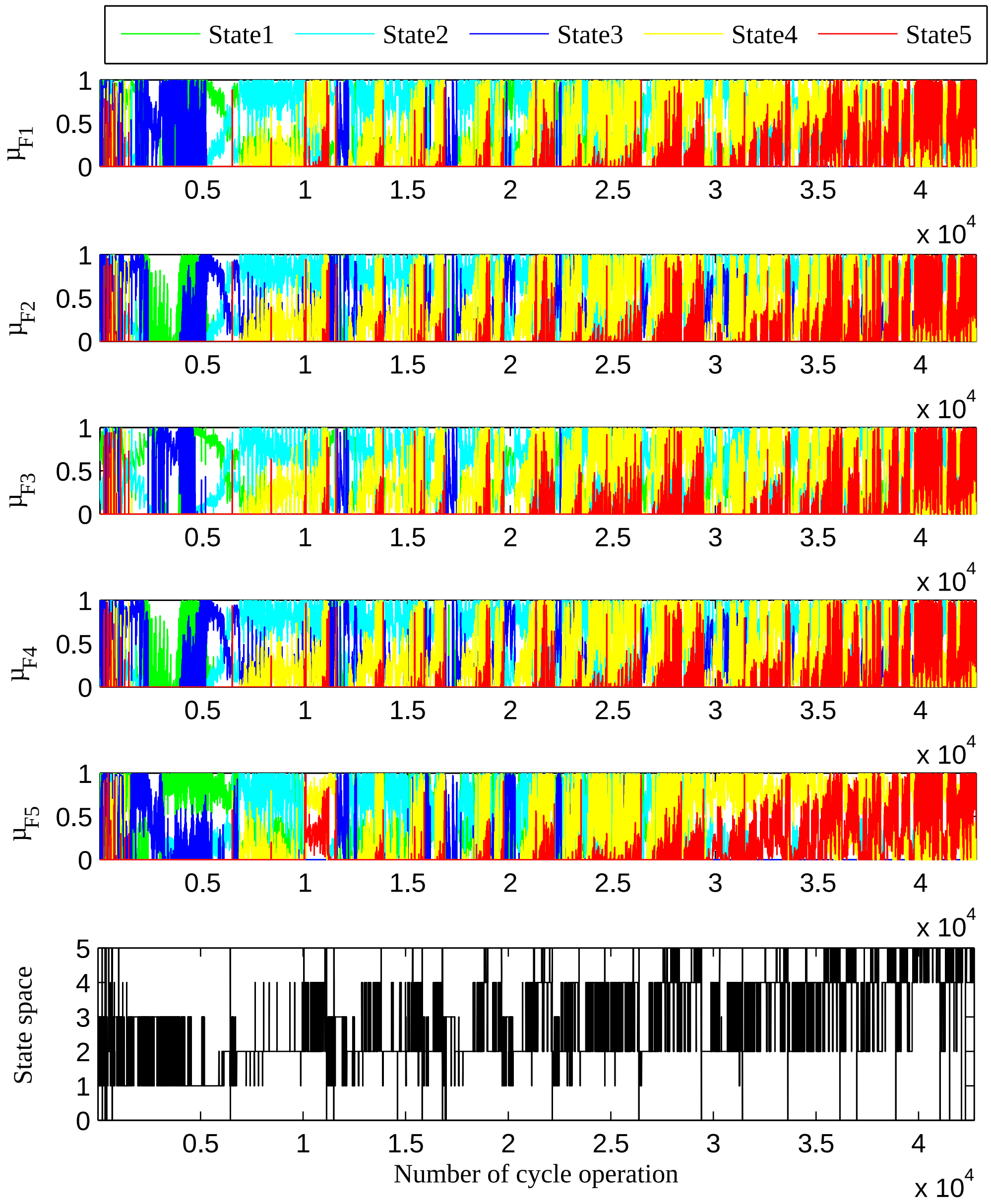

Fig. 9. Results of the evaluation for the test dataset 1 based on the proposed algorithm

realizing the tangential force for sliding of the metal parts) is used an analyzed. The application shows that the proposed approach can be used to classify complex material processes, here indicating the state of the surface erosion rate.

References

[1] I. Guyon, S. Gunn, M. Nikravesh and L. A. Zadeh, Feature Extraction, Foundations and Applications, Berlin Heidelberg, Springer-Verlag, 2006.

[2] S. Abe, Support Vector Machine Pattern Classification, Springer-Verlag London Limilted, 2005.

[3] R. Patton, P. Frank, and R. Clark, Fault Diagnosis in Dynamic Systems, Theory and Applications, Hemel, Hempstead, 1989.

[4] A. Ibrahim, FUZZY LOGIC for Embedded Systems Applications, MA, USA: Elsevier Science, 2004.

[5] F. McNeill and E. Thro, FUZZY LOGIC: APRACTICAL APPROACH, MA, USA: Academic Press, Inc., 1994.

[6] R. Abrahart, P. Kneale and L. See, Neural Networks for Hydrological Modelling, London, UK: Taylor & Francis Group plc, 2004.

[7] S. Abe, Pattern Classification: Neuro-fuzzy Methods and Their Comparison, London, Great Britain: Springer-Verlag London Limilted, 2001.

[8] F. Heijden, R. Duin, D. Ridder, and D. Tax, Classification, Parameter Estimation and State Estimation: An Engineering Approach using MATLAB, England: John Wiley & Sons Ltd, 2004.

[9] D. Söffker, H. Aljoumaa, "Signal-based modeling a new method for classifying system," *7th Int. Workshop on Structural Health Monitoring*, Stanford University, Stanford, CA, USA, vol. 2, pp. 1519-1527, Sep. 2009.

[10] H. Aljoumaa, D. Söffker, Condition Monitoring and Classification Approach based on Fuzzy-Filtering, Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2010, WCECS 2010, 20-22 October, 2010, San Francisco, USA, pp 503-508.

[11] J. Abrahart and S. White, "Modeling Sediment Transfer in Malawi: comparing backpropagation neural network solutions against a multiple linear regression benchmark using small data sets," *Physics and Chemistry of the Earth Part B: Hydrology, Oceans and Atmosphere*, vol. 26, no. 1, pp. 19-24, 2001.

[12] S. Yella, N. Gupta and M. Dougherty, "Condition monitoring using pattern recognition techniques on data from acoustic emissions," *Proceedings of 5th International Conference on Machine Learning and Applications(ICMLA'06)*, 2006.

[13] A. Malhi and R. X. Gao, "PCA-Based Feature Selection Scheme for Machine Defect Classification," *IEEE Transaction on Instrumentation and Measurement*, vol. 53, no. 6, pp. 1517-1525, Dec. 2004.

[14] B. Samanta, KH. R. Al-Balushi and S. A. Al-Araimi, "Bearing Fault

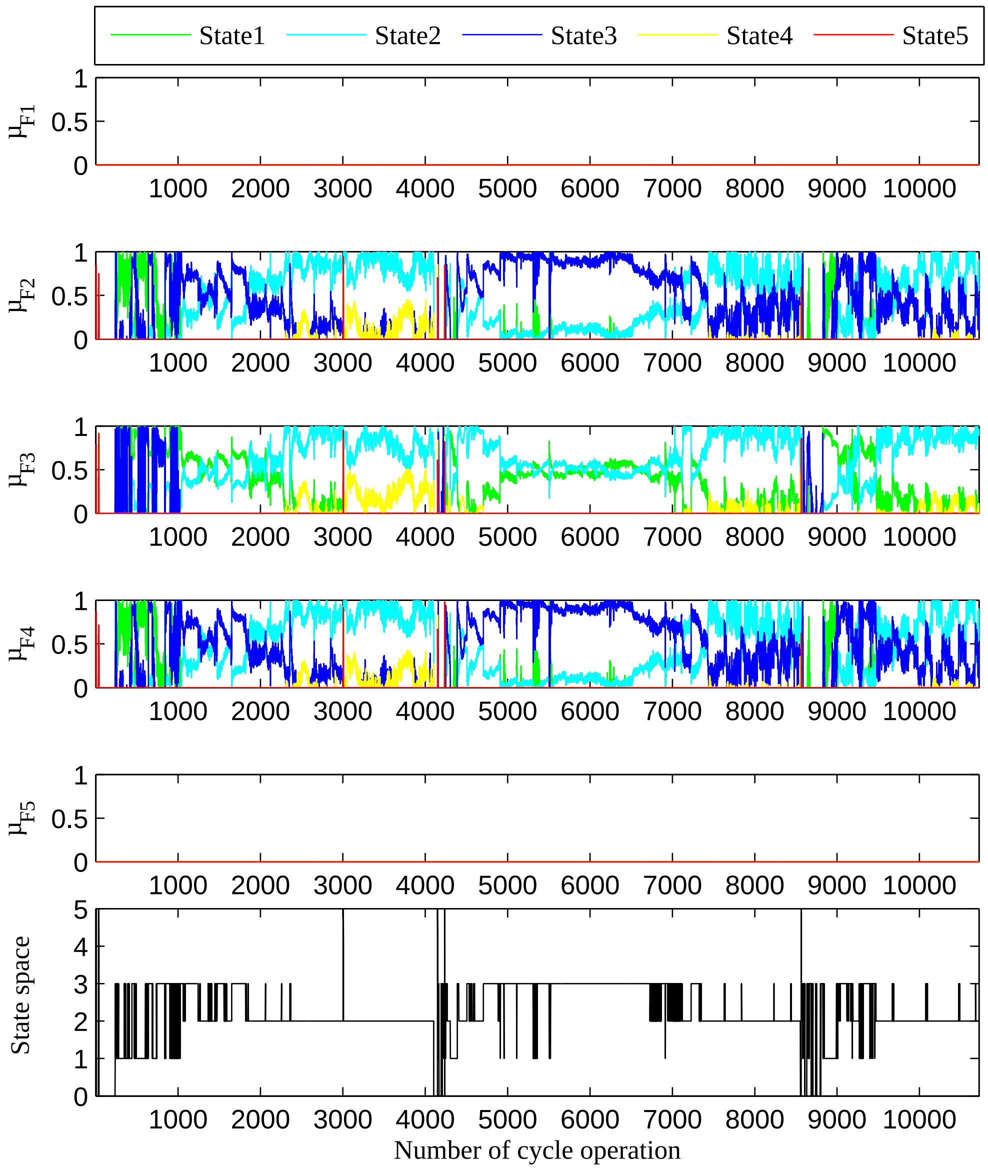

Fig. 10. Results of the evaluation for the test dataset 2 based on the proposed algorithm

Detection Using Artifical Neural Networksand Genetic Algorithm,“ *EURASIP Journal on Applied Signal Processing*, vol. 3, pp. 366-377, Mar. 2004.

[15] O. Ondel, E. Boutleux and G. Clerc, “A Method to Detect Broken Bars in Induction Machine Using Pattern Recognition Techniques,“ *IEEE Transaction on Industry Applications*, vol. 42, no .4, pp. 916-923, Jul-Aug. 2006.

[16] H. Guo, L. B. Jack and A. K. Nandi, “Feature Generation Using Genetic Programming With Application to Fault Classification,“ *IEEE Transaction on Systems, Man, And Cybernetics-Part B: Cybernetics*, vol. 35, no. 1, pp. 89-99, Feb. 2005.

[17] A. Joshi, N. Ramakrishman, N. Houstis, and R. Rice, “On Neurobiological, Neuro-Fuzzy, Machine Learning, and Statistical Pattern Recognition Techniques,“ *IEEE Transaction on Neural Networks*, vol. 8, no. 1, pp. 18-31, Jan. 1997.

[18] M. Reddy and D. K. Mohanta. “Adaptive-neuro-fuzzy inference system approach for transmission line fault classification and location incorporating effects of power swings,“ *IET Gener. Transm. Distrib*, vol. 2, no. 2, pp. 235-244, March. 2008.

[19] C. Fira and L. Gora, “An ECG Signals Compression Method and Its Validation Using NNs,“ *IEEE Transaction on Biomedical Engineering*, vol. 55, no. 4, pp. 1319-1326, April. 2008.

[20] L. Carro-Calvo, S. Salcedo-Sanz, R. Gil-Pita, A. Portilla-Figueras and M. Rosa-Zurera, “An evolutionary multiclass algorithm for automatic classification of high range resolution radar targets,“ *Integrated Computer-Aided Engineering*, vol. 16, pp. 51-60, Jan. 2009.

[21] S. Guan and F. Zhu, “An Incremental Approach to Genetic-Algorithms Based Classification,“ *IEEE Trans-Actions On Systems, Man ,And Cybernetics-Part B: Cybernetics*, vol. 35, no. 2, pp. 227-239, April. 2005.

[22] J. Cervantes, X. Li, and W. Yu, “SVM Classification for Large Data Sets by Considering Models of Classes Distribution,“ *IEEE Transaction on Computer Society*, pp. 51-60, Nov. 2007.

[23] H. Shuju, L. Jianlin and X. Honghua, “A SVM Method based on Active Area Vecto,“ *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, 2008*, pp. 2564-2568, April. 2008.