

<http://doi.org/10.24867/IJIEM-2018-3-167>

# Tool Condition Monitoring System Based on a Texture Descriptors

**Aco Antić**

Associate Professor, University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad Serbia, [antica@uns.ac.rs](mailto:antica@uns.ac.rs);

**Milan Zeljković**

Professor, University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad Serbia, [milanz@uns.ac.rs](mailto:milanz@uns.ac.rs);

**Nicolae Ungureanu**

Professor, Technical University of Cluj-Napoca, Department of Engineering and Technologic Management, Dr. V. Babes 62/A, 430083 Baia Mare, Romania, [unicu@ubm.ro](mailto:unicu@ubm.ro);

**Ivan Kuric**

Professor, University of Žilina, Dpt. of Machining and Automation, Univerzitna 1, 01026 Zilina, Slovakia, [ivan.kuric@fstroj.uniza.sk](mailto:ivan.kuric@fstroj.uniza.sk);

Received (21.02.2018.); Revised (29.08.2018.); Accepted (04.09.2018.)

## Abstract

All state-of-the-art Tool condition Monitoring systems (TCM), especially those that use vibration sensors, in the tool wear recognition task, heavily depend on the choice of descriptors that contain information concerning the tool wear state, which are extracted from the particular sensor signals. All other post-processing techniques do not manage to increase the recognition precision if those descriptors are not discriminative enough. In this work, we propose tool wear monitoring strategy, which relies on the novel texture based descriptors. We consider the module of the Short Term Discrete Fourier Transform (STDFT) spectra obtained from the particular vibration sensors signal utterance, as the 2D textured image. This is done by identifying the time scale of STDFT as the first dimension, and the frequency scale as the second dimension of the particular textured image. The obtained textured image is then divided into particular 2D texture patches, covering part of the frequency range of interest. After applying the appropriate filter bank, for each predefined frequency band 2D textures are extracted. From those, for each band of interest, by averaging in time, we extract information regarding the Probability Density Function (PDF) of those textures in the form of lower order moments, thus obtaining the robust tool wear state descriptors. We validate the proposed features by the experiments conducted on the real TCM system, obtaining the high recognition accuracy.

**Key words:** *tool wear, texture descriptors, vibration signal*

## 1. INTRODUCTION

Tool Condition Monitoring (TCM) is very important in manufacturing processes and has been of great interest to many academic and practical researches. Signal processing and information technology has resulted in the use of multiple sensors for the effective monitoring of tool wear conditions, which is the most crucial feedback information to the process control and tool wear prediction [1]. Tool failure can be prevented by efficiently monitoring conditional changes in the tool. Cho et al, in their basic work, divided tool conditions into following categories: tool breakage, tool chipping, and tool wear [2]. A key issue for an unattended and automated machining system is the development of reliable and robust TCM systems.

There are many different ways to gather information about the tool failure, by the usage of adequate sensors, and thus corresponding signals used in TCMs. Those

can be divided into two groups: direct (consist of laser, optical, and ultra-sonic sensors which provide direct measurement, for example [3] and indirect, based on sensors which infer the machining state by sensing cutting forces, vibrations, temperatures, current consumption, etc.

All Intelligent TCM systems, and especially those that use vibration sensors, heavily depend on the choice of descriptors, i.e., features extracted from the particular sensor signals. This is due to the fact that if the descriptors do not describe the signal adequately, other techniques, such as feature extraction or feature selection, as well as recognition methodology, fail to be efficient as well. Bahr et al in [4], and Tsai et al in [5], were the first to apply the descriptors obtained from vibration signals in the TCM tasks. Actually, they used RMS and/or the mean of the vibration sensor signal, in order to detect an increase in vibration magnitude, which corresponds to the increase in cutting energy generated

due to flank wear. Also in [6], the mean and peak values of vibration sensor signals were used in the TCM task. In [7], Dimla analyzed the correlation between the vibration signal features and the cutting-tool wear, both in time and frequency domains, during turning operations. Time domain features were deemed to be more sensitive to the cutting condition than tool wear, whereas certain peak values in the frequency domain correlated well with the measured wear values. In [2], authors proposed descriptors extracted from multiple sensors in time and frequency domain and used them in their multisensor fusion-based tool condition monitoring system which is applied in the end-milling task. In [8], [9], [10] authors presented a tool wear monitoring strategy based on a large number of signal features extracted from time domain signals, as well as from their frequency domain representation and also their wavelet coefficients. Also, in [8], time-domain based features are combined with wavelet based, and feature selection using logistic regression is applied in the task of determining whether or not the cutting tool is reliable (two class problems. In [11], authors were among the first to use wavelet analysis and lower order moments, such as average value, standard deviation, power value, kurtosis value, harmonics frequency, skew value, etc., in order to express workpiece and spindle vibrations in the X, Y and Z directions. Nevertheless, concerning the usage of vibration based signals, those systems, in most cases use standard, time-domain and/or frequency-domain extracted features or wavelet based features, mostly developed in previous works, some of those already mentioned. In [12], the same authors proposed an adaptive network-based fuzzy inference based method as their actual decision making system, using the same, previously mentioned classical frequency-based descriptors.

In this work, we propose the tool wear monitoring strategy which includes novel texture based descriptors, to be applied in the TCM problems utilizing vibration sensor signals. Actually, the texture based approach to the descriptor construction, by the means of exploring the texture structure of the time-frequency representation of vibration sensor signal, is to the knowledge of authors completely novel in the field of TCM. The proposed descriptors obtained from vibration sensor signals are spectral domain-based. They rely on considering the module of the Short Term Discrete Furrier Transform (STDFT) spectra obtained from the particular sensor signal, as the 2D textured "image". We identify the time scale as the first dimension, and the frequency scale as the second dimension of module of the STDFT of the particular sensor signal of interest. The 2D textured image obtained in such a way is then divided into particular disjoint narrow 2D texture patches, covering the part of the frequency range of interest. Furthermore, by applying an appropriate filter bank, for each predefined frequency band (where those bands form the partition of the whole frequency range), we extract 2D textons [13], i.e., low dimensional feature vectors in filter response space. Our aim is to exploit the mentioned filter response of the corresponding textons, in order to encode fine differences in the texture structure

of mentioned texture patches. The goal is to increase the discriminativity of the obtained descriptors in the task of tool wear classification. Moreover, we also tend to represent, in the form of low dimensional features, the significant details of the texture patches in order to gain on robustness of the description which corresponds to the tool wear states. Our approach lies on the assumption, which we later confirm in the experiments, that the underlying physical process of tool cutting, generated by different tool wear states, is closely related to the structure of mentioned textons. Nevertheless, the main problem that we face, when trying to use the texture based approach, is that most features, i.e., descriptors used in texture recognition, are inappropriate for this particular application. The reason is that discriminativity between texture patches, belonging to different tool wear condition classes, is contained in vertical lines spread across the entire frequency range, and many of the used techniques in texture analysis could not take that into account. Thus, the core of in our approach is the selection of the appropriate filter bank that is able to efficiently extract the information on the presence of mentioned lines, for each frequency band. In this paper, we apply the particular filter bank, proposed in [14], used in the problems of texture recognition, which has the mentioned capability. We apply it on 2D texture patches, obtained from the STDFT of the vibration sensor signal, in order to efficiently extract the information concerning the mentioned lines, in the form of textons, i.e. features in filter response space. We additionally model probability distributions, for each frequency band, which describe every particular utterance in filter response space, and use them as the low dimensional, robust descriptors. This is obtained by using the first four moments (time averaging is applied), and obtain final descriptors that represent the particular utterance to be used in the training or recognition process. Thus, the information on the texton probability distributions corresponding to each frequency band is extracted and contained in a small feature vector, in order to be efficiently used in training and recognition phase of TCM task.

## 2. APPARATUS USED IN EXPERIMENTS

Machining experiments were performed on CNC GU 600 lathe manufactured by INDEX and installed in a laboratory of the Faculty of Technical Sciences in Novi Sad. Investigation of tool wear process encompassed the monitoring of the dominant wear mechanism through following parameters: wear band (VB), crater wear (KT) and tool life. In the course of the turning process, vibration signals were registered at the tool shank. For each tool pass, the generated chip segments were sampled. The setup of tool sensors, as well as the dimension of the workpiece used in this experiment, is shown in Figure1. During the experiment two cutting speeds were employed, 180 to 250m/min, in conjunction with 0.15 and 0.3 mm/rev feed rates. Cross section of the tool shank used in the experiment was 20x20mm. The machining was performed with P25 tool inserts designated TNMM 110408. Accelerometer Kistler 8002 was fixed onto the tool holder, and used to measure acceleration of vibrations. This signal was sampled at

625 kHz, using A/D converter NI 625 USB, National Instruments. Workpiece material was 42CrMo4, 310 HB hardness and 950MPa respectively with guaranteed mechanical and chemical properties

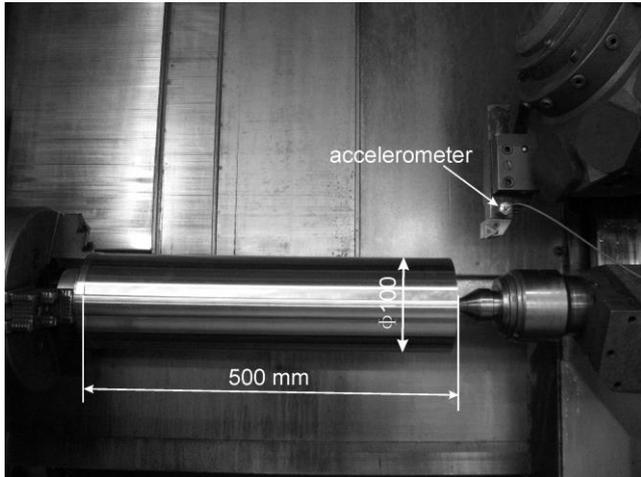


Figure 1. Experimental setup

### 3. TEXTURE BASED FEATURES

The utterance  $u$  from which we extract descriptors and consequently the features that we use in the training, as well as in the recognition phase of the classifier, is the signal, of various lengths, obtained from the vibration sensor. In this work, we view the problem of tool wear state classification from the texture recognition perspective. The most challenging difficulties inherent to the real texture recognition problem with 3D variations, coming from the variability of those textures, i.e., the fact that the classical texture is primarily a function of the following variables: the texture surface, its albedo (i.e., the reflection factor of the corresponding surface), the illumination, the camera and its viewing position. It should be noted, that the texture that corresponds to the STDFT spectrogram of the vibration sensor signal is viewed as 2D texture, and does not exhibits that kind of variability, [15]. Nevertheless, the presence of vertical lines in the STDFT of interest (viewed as the texture), as well as our goal to efficiently extract them and to obtain the reliable information of their presence, imposes the necessity for texture feature extraction that utilizes appropriate and specific filter bank, capable to capture those. In order to extract relevant features to be used in the actual tool state recognition task, we first extract relevant descriptors in the form of texton features obtained in an appropriate filter output space, and then model their probability distribution for each frequency band. The modelling is obtained by using first four moments of those outputs, as the final, compressed, and thus robust descriptors.

#### 3.1 Forming the texton based descriptors

Actually, we identify the discrete time frame  $k$  of the STDFT spectrogram of the utterance of interest, as the discrete  $x$ -axis of the textured image, and the discrete frequency  $\omega$ , as the discrete  $y$ -axis of the corresponding textured 2D image. The idea, later confirmed as the

hypothesis in our experiments, is that considering the STDFT spectrogram of the particular utterance, the discrimination between different tool wear states is mostly contained in the characteristics (number, position, structure) of the (approximately) vertical lines appearing on the corresponding texture (see Figure 2). Actually, the condition of the tool wear states is strongly correlated with the characteristics of those lines. Those lines, actually present sudden jumps in frequency of the signal obtained from the particular sensor, on various frequency scales, induced by the underlying tool cutting process. Obtaining information considering those lines is essential, and determines our choice of the filter bank to be used for that particular task.

We invoke novel texture based descriptors in the task of TCM and call them the Texture Based Tool Condition Descriptors (TBTCD). The TBTCD descriptors are obtained by considering the STDFT spectrogram of the windowed time frames delivered from the previously mentioned utterance, as the 2D textured image.

We proceed as follows: Common to the most texture recognition approaches, we divide the STDFT 2D textured image in to texture patches spread across different filter bands. Then, by using the appropriate filter bank, we extract texton based descriptors as the feature vectors in the low dimensional filter response space. It is crucial to select the appropriate filter bank that is able to efficiently extract the information on the presence of mentioned lines, for each frequency bend. The most promising [14, 16] for the mentioned tasks are response spatially invariant filter banks, i.e., filter banks based on the Maximum Response filter set such as (BFS, MR8, MR4, and MRS4). Those are previously reported and discussed in material classification tasks [12, 13]. In our application to TCM task, we have chosen to use MR8 filter bank, presented in Figure 2, but to exclude the last two isotropic components, as they do not extract relevant texture features.

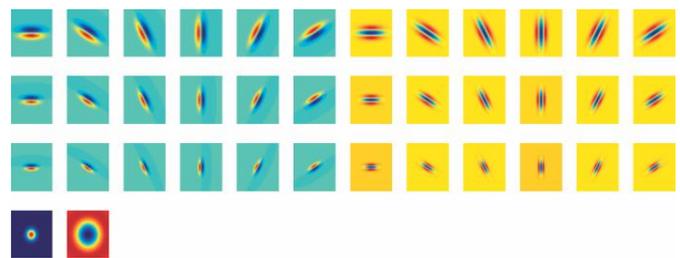


Figure 2. MR8 filter bank: It consists of an oriented edge filter at 6 orientations and 3 scales, a bar filter at the same set of orientations and scales, an isotropic Gaussian filter and a Laplacian of Gaussian filter

Let  $F(x,y)$  be the 2D textured image that corresponds, i.e., is identified with the STDFT spectrogram of the vibration sensor signal of the particular utterance  $s$ . Recall, that the STDFT spectrogram of signal  $s$  is defined as the square module of the spectral density, i.e.,  $|S(k, \omega)|^2$ , where

$$S(k, \omega) = \sum_{n=-\frac{K}{2}}^{\frac{K}{2}} s(n)w(k-n)e^{-ikn} \tag{1}$$

represents the STDFT of the signal  $s$ . In (1), we denote the discrete time frame by  $k$  and the discrete time frequency i.e., discrete frequency beam, by  $\omega$ . Term  $w$  is the windowing sequence that we use (we use Hamming window in the experiments), with the fixed length  $K$ . We identified  $x$  axes of the textured image with the Discreet time frame  $k$  of the mentioned STDFT, for  $k = 0, \dots, k_{\max}$ , so that it holds  $x_{\max} = k_{\max}$ . Actually, for every discrete time frame  $k$  in STDFT, there is a corresponding discrete index  $x$  of the texture patch in the textured image  $F$ . Also, the discreet frequencies  $\omega = 0, \dots, \omega_{\max}$  are identified with discreety axes of the textured image  $F$ , so that it holds  $y_{\max} = \omega_{\max}$ . We now have the following interpretation of STDFT in the form of the textured image:

$$F(x, y) = |S(k, \omega)|^2 \quad (2)$$

$$x = 0, \dots, x_{\max}, y = 0, \dots, y_{\max}$$

$$k = 0, \dots, k_{\max}, \omega = 0, \dots, \omega_{\max}$$

We further perform analysis on the textured image  $F(x, y)$  obtained as previously explained, where for the ease of modeling, without loss of generalization, we consider the continuous case  $x \in [0, x_{\max}]$  and  $y \in [0, y_{\max}]$ . We briefly describe the filter that we use: Let

$$G(\sigma_x, \sigma_y, \theta, x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{(A(\theta)^T [x \ y]) \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} (A(\theta) \begin{bmatrix} x \\ y \end{bmatrix})} \quad (3)$$

be the directed anisotropic Gaussian kernel. The fixed terms  $\sigma_x > 0$  and  $\sigma_y > 0$  denote the scales in  $x \in [0, x_{\max}]$ ,  $y \in [0, y_{\max}]$  direction, respectively, while  $\theta$  and  $A(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ , for  $\theta \in [0, 2\pi]$ , denote the orientation of the kernel (3) and its 2D rotational matrix, respectively. The MR8 filter bank that we use in our application consists of 18 Gaussian filters of the form (3), at 6 different orientations  $\theta_j \in \{j\pi/6 \mid j = 0, \dots, 5\}$ , and 3 different scales

$$(\sigma_x, \sigma_y) \in \{(1,3), (2,6), (4,12)\},$$

and also of 18 Laplacians of Gaussians, defined as  $\Delta G(\sigma_x, \sigma_y, \theta, x, y)$ . Term  $\Delta = \partial_x^2 + \partial_y^2$  is the Laplace operator of differentiation, evaluated at the same orientations and scales (for more details see [14], [15]). The filter outputs are then collapsed, by taking only the maximum values across all orientations (for Gaussians and Laplacians of Gaussians separately), consequently obtaining 6 different filter responses. Being isotropic, two additional responses are unable to extract characteristic lines in the textured image  $F$ , so they are therefore omitted.

Further we proceed with the procedure of texton extraction. We divide the interval  $[0, y_{\max}]$  where texture image  $F$  have its values into  $M$  sub-intervals (bands)  $\{[y_{q-1}, y_q] \mid q = 1, \dots, M\}$ ,  $y_0 = 0$ ,  $y_M = y_{\max}$  where for every index  $x = 0, \dots, x_{\max}$ , the sub-patch of  $F(x, y)$ , corresponding to the  $q$ -th band, is defined as a matrix of pixels bounded by  $[y_{q-1}, y_q]$  on yaccess and by  $[x, x + T]$  on xaccess. Here,  $T$  is predefined width of the patch. We note, that the interval  $[y_{q-1}, y_q]$  of the textured image  $F$ , corresponds to the frequency band  $[\omega_{q-1}, \omega_q]$  of the

STDFT spectrogram of the particular utterance. We denote such a described textured patch as  $P_{x,q}$ . The texton, in the form of a filter response, is then formed for every band  $q$  and discrete coordinate  $x$ . Actually, for each  $x$  and each different band  $q$ , texton  $v_{x,q}$  is composed of  $d = 6$  components of the MR8 filter bank. Thus, we have:

$$v_{x,q}^l = [v_{x,q}^{l(1)} \mid v_{x,q}^{l(2)} \mid l \in L]. \quad (4)$$

where

$$v_{x,q}^{l(1)} = \max_{\theta} G(\sigma_x, \sigma_y, \theta, x, y) * P_{x,q}$$

$$v_{x,q}^{l(2)} = \max_{\theta} \Delta G(\sigma_x, \sigma_y, \theta, x, y) * P_{x,q}$$

$$\theta \in \{j\pi/6 \mid j = 0, \dots, 5\}$$

and

$$L = \{(\sigma_x, \sigma_y) \mid (1,3), (2,6), (4,12)\}.$$

In (4), for simplicity we enumerate:  $l = 1, 2, 3$ .

### 3.2 Using moments to describe the probability distribution in the filter response space

We model the probability distribution of textons described in the previous section, for every band  $q$  separately. As a final result, we obtain compressed information about the mentioned distributions in the form of robust features. We proceed as follows:

For the fixed utterance  $u$  and fixed band  $q = 1, \dots, M$ , we have the set of observation features, i.e., texton descriptor vectors

$$\{v_{x,q}^u \mid x = 0, \dots, x_{\max}^u\} \quad (5)$$

where  $v_{x,q}^u$  given by (4). It is obtained for all patches  $P_{x,q}$ . It holds that  $x_{\max}^u = k_{\max}^u$  (see (2)) is the number of time samples in the utterance  $u$ , which, can vary, depending on the utterance. We consider  $v_{x,q}$  to be the observations, i.e., realizations of the random variable  $V_q^u$  with the probability distribution function (pdf)  $p_{V_q^u}$ . For the simplicity, we consider all components of  $V_q^u$  (i.e.,  $v_{x,q}$ ) to be uncorrelated. As, for the fixed band  $q$ , there is a unique correspondence between the probability distribution  $p_{V_q^u}$  of  $V_q^u$  and its characteristic function, given by

$$k_{V_q^u}(t) = E[e^{iV_q^u t}] = \sum_{j=0}^{\infty} \frac{(it)^j}{j!} E\left(\left(V_q^u\right)^j\right) \quad (6)$$

where one could observe that (5) is given as an expansion which depends only on moments  $m_q^{u,j} = E\left(\left(V_q^u\right)^j\right)$ ,  $j = 1, 2, \dots$ , where the averaging is conducted over  $x$ . It should be noted, that all moments  $m_q^{u,j}$  are actually the  $d = 6$  dimensional vectors, as we considered components of  $V_q^u$  to be uncorrelated. By taking only the first few  $\{j = 1, \dots, P\}$  moments into account we manage to (roughly) approximate pdf  $p_{V_q^u}$ , and yet obtain robust features, i.e., descriptors which can be used in the recognition part of TCM task. We note that the robustness of the features is especially useful in the situations where the number of training samples is not

particularly big, as is the case in most applications concerning TCM.

The final feature vector assigned to each utterance  $u$  is obtained as follows: In the first case, we consider each component of moments  $m_q^{u,j}$  separately, and gather them into one single vector (for each component), thus obtaining  $d = 6$  independent PM dimensional vectors, for each utterance  $u$ . We call this "P moments per scale" case, as described in Section 4. In that case, in the actual classification phase, we use  $d = 6$  different classifiers and average their results. In the second case, which we call "all scales together" case, we consider all scales together and obtain one PMd dimensional vector corresponding to each utterance. As described in section 4, in that particular case, we use one single classifier in the recognition phase.

For the first case, we obtain  $d = 6$ , separate feature vectors given by:

$$v_i^u = [m_1^{u,1,(i)} \dots m_1^{u,P,(i)} | \dots | m_M^{u,1,(i)} \dots m_M^{u,P,(i)}], i = 1, \dots, d. \tag{7}$$

where  $m_q^{u,j,(i)}$  denotes the  $i$ -th component of moment  $m_q^{u,j}$  which is  $d = 6$  a dimensional vector, obtained as previously explained.

For the second case, a vector containing PMd components is obtained as  $v^u = [v_1^u | \dots | v_d^u]$ , i.e.,  $v_i^u$ , given by (7), are concatenated into a single vector.

In Figure 3, we give the flow chart of the whole TCM procedure used. Also, in the sequel, we give the pseudocode for the final procedure of extracting of TBTCd descriptors:

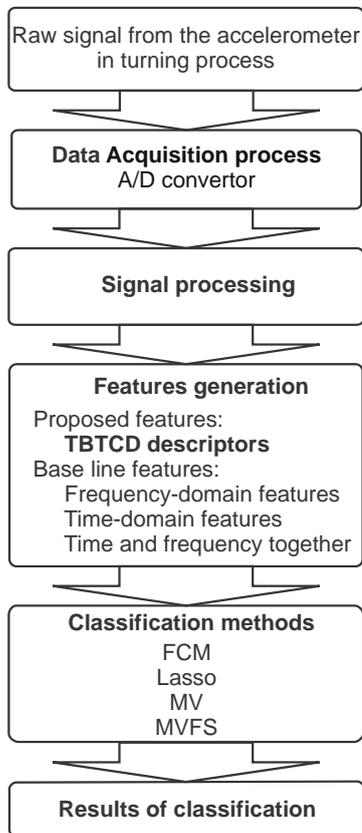


Figure 3. Flow chart describing the TCM procedure used

For every signal  $s$  obtained from utterance  $u$ , do:

Begin

- Calculate the spectral density  $|S(k, \omega)|^2$  of the STDFT defined by (1),  $k = 0, \dots, k_{max}, \omega = 0, \dots, \omega_{max}$

- Form the target textured image as  $F(x, y) = |S(k, \omega)|^2, x = 0, \dots, x_{max}, y = 0, \dots, y_{max}$

- Divide the interval  $[0, y_{max}]$  where texture image  $F$  has its values into  $M$  sub-intervals (bands)  $\{[y_{q-1}, y_q] | q = 1, \dots, M\}, y_0 = 0, y_M = y_{max}$  and extract sub-patches  $P_{x,q}$  of  $F(x, y)$ , for  $x = 0, \dots, x_{max}$ , corresponding to the  $q$ -th band

- Form textons  $v_{x,q}$  for  $x = 0, \dots, x_{max}$ , and  $q = 1, \dots, M$ , by using the  $d = 6$  components of MR8 filter bank, as follows:

$$v_{x,q}^1 = [v_{x,q}^{1,(1)} | v_{x,q}^{1,(2)} | l \in L],$$

$$v_{x,q}^{l,(1)} = \max_{\theta} G(\sigma_x, \sigma_y, \theta, x, y) * P_{x,q}, v_{x,q}^{l,(2)} = \max_{\theta} \Delta G(\sigma_x, \sigma_y, \theta, x, y) * P_{x,q}$$

$$\theta \in \{j\pi/6 | j = 0, \dots, 5\}, L = \{(\sigma_x, \sigma_y) | (1,3), (2,6), (4,12)\},$$

where  $G$  is defined by (3).

- Form separate  $v_i^u$ , given by (7), for  $i = 1, \dots, d$ , where  $m_q^{u,j,(i)}$  denotes the  $i$ -th component of the moment  $m_q^{u,j}$ ,  $\dim(m_q^{u,j}) = d = 6$ ,  $m_q^{u,j} = E((V_q^u)^j), j \in \{1, \dots, P\}$ , where  $P$  is the number of moments used.

End

#### 4. DESCRIPTORS AND EXPERIMENTAL METHOD

Tool wear state classification is conducted indirectly by monitoring vibration signals during the cutting process. The experiments have been conducted over 150 vibration signals, classified into 3 states (50 signals per each state), based on the level of abrasion (new tool insert, i.e., no wear, small wear, i.e., up to 0.25mm, and a large wear, i.e. above 0.5mm).

Each state represents one class in the recognition task. Bearing in mind that the tool wears condition changes continually, the borders between classes could not be ideally sharp.

The analyses of the results are provided by using 5-fold cross validation: The total number of 120 vibration signals used for the training of the classifier, and another 30 vibration signals used for the test phase, where each of the 3 categories were equally represented in both sets.

All 5 results from the folds have been averaged to produce the final average classification accuracy.

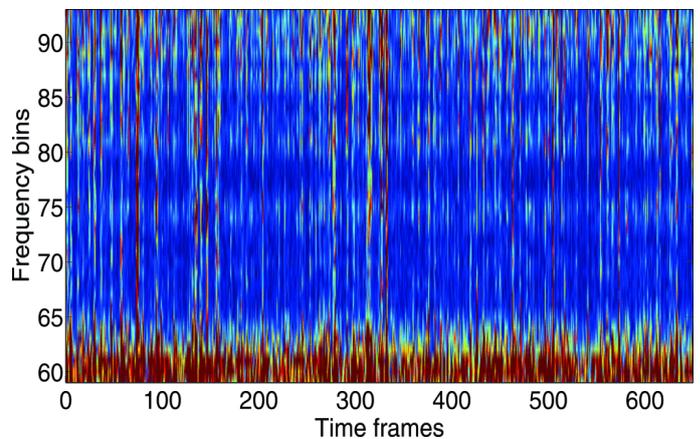
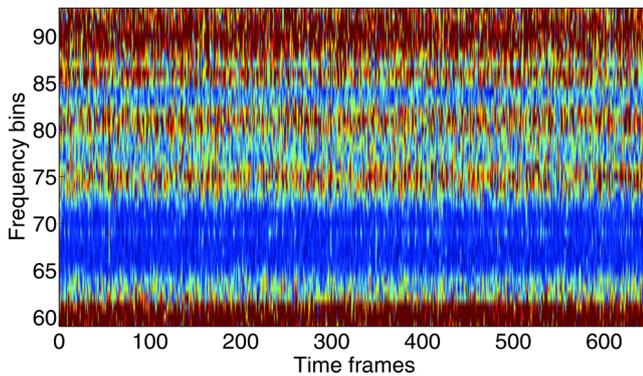


Figure 4. Vibration signal spectrogram with a new tool insert



**Figure 5.** Vibration signal spectrogram with a Wear tool insert where the number of characteristically vertical lines is high

Signal transformation, for both sets, begins by computing the spectrum of vibration signals, using a discrete-time Short-Time Fourier Transform (STFT), where we use  $M = 256$  frequency points and the Hamming window of the equal length  $M$ , with  $M/4$  overlap between the segments (example is given in Fig. 4 and 5). The signals were 0.2 seconds long in time, with the sampling frequency of 625 kHz. The result is a matrix containing 129 frequencies ( $M/2 + 1$  for real signals and even number of frequency points, although we used a limited frequency range for subsequent transformations) and 650 time frames (the number of segments). Each segment represents the estimation of a short-time, time-localized frequency content of input vibration signals. Each spectrogram is centralized in order to exclude the DC component. The convolution is conducted between the spectrum and the appropriate filter belonging to the MR8 filter bank described in 3.1 and presented on Figure 2.

Thus, the robust features, i.e., descriptors are obtained as it is described in Sections 3.1 and 3.2. We treated only the one simple band case,  $M = 1$ , where we first excluded all relative frequencies lower than  $\omega_{low} = 0.4535$  and greater than  $\omega_{high} = 0.7256$  as it is presented on Figure 4 and 5. Those frequencies were not discriminative enough for our recognition task. We used  $P = 3$  moments, namely variance, skewness and kurtosis, while the first moment, i.e., the mean is excluded from our consideration, since we treated only centralized spectrograms, i.e., the corresponding centralized textured images.

The recognition is performed by using the following classifiers: The first one is a fuzzy classifier over the selected set of features. The classifier generates a Fuzzy Inference System (FIS) structure of the Mamdani type from input coefficients, using Fuzzy C-Means (FCM) clustering algorithm, by extracting a set of rules that models the data behavior. Each input signal is represented by a vector of  $P$  ("P moments per scale" case, see Section 3.2) or  $Pd$  coefficients ("all scales together" case) and is classified into one of the 3 output states, marked by numerical values 1, 2 and 3, depending on the wear state (the level of abrasion) of a steel plate. Note that in all experiments, as previously explained, we use  $P = 3$  and  $d = 6$ . In "P moments per scale" case, we train  $d$  different classifiers and obtain the final result, by averaging their output. In "all scales

together" case, we use one single classifier. Rather than being attached to only one cluster, each vector has a degree of belonging to the clusters, depending on the distance from the center of a cluster. In FCM clustering, the centroid of a cluster is the mean of all vectors, weighted by their degree of belonging to the cluster, related inversely to the distance from the cluster center. During initialization, each vector is attached randomly to the clusters. The levels of attachment change between consecutive iterations. The centroids are computed repeatedly until convergence. The rule extraction method uses the FCM function to determine the number of rules and membership functions for the antecedents and consequents. The second classifier is based on Lasso regression. We recall that the object of Lasso is to solve the following minimization problem, presented in the compact matrix form:

$$\alpha^* = \operatorname{argmin}_{\alpha} \|Y - X\alpha\|_2^2 + \mu \|\alpha\|_1, Y = [y_1 \dots y_N]^T, X = [1, x_1 \dots x_N]^T, \alpha = [\alpha_0, \dots, \alpha_p] \quad (8)$$

where  $Y = [y_1 \dots y_N]^T$  are the soft data labels, each corresponding to particular class, i.e.,  $y_j \in \{1, 2, 3\}$ , where each number represents a distinctive class identifier, i.e., the label, and at the same time, the value attached to the particular data instance. Next,  $X$  is the matrix of an impute features data vector, where  $x_j \in \mathbb{R}^p$  represents a particular feature corresponding to  $j$ -th instance  $j = 1, \dots, N$ , where  $N$  is a number of training data instances, while  $1 = [1, \dots, 1]$  and  $p$  is a number of different features. The term  $\mu > 0$  adjusts the trade-off between the error term  $\|Y - X\alpha\|_2^2$  in (8) and the sparse regularizer  $\mu \|\alpha\|_1$ , where  $\|\cdot\|_1$  is the  $l_1$  norm of the coefficients  $\alpha$ , thus penalizing a large number of nonzero  $\alpha_j$  coefficients. In all experiments presented in Section 5, we set  $\mu = 0.2$ .

The third is the classifier based on the classical multivariate regression, which is obtained from (8), by setting  $\mu = 0$ . We applied two different versions. In the first version, we first apply Lasso based feature selection by solving (8), on some initial set of features. Those  $j$ -th features with  $\alpha_j > t$ , with some predefined threshold  $t > 0$  are then chosen in the set of features on which we obtain classical multivariate regression based classification, using data labels as previously described. We call it a MVFS classifier. As it is described in Section 5, in our experiments, we use the baseline, as well as novel descriptors that we propose and apply Lasso feature selection, in order to obtain the most discriminant feature set. In the second version, we apply classical multivariate regression on the complete given set of features. We call it a MV classifier.

## 5. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, experimental results are presented which involve the proposed novel TBTCDD descriptors in comparison to classical time-domain, frequency-domain based features, referenced in the Section 4, and also the combination of those two. We use the following classifiers, all described in Section 4: a FCM based classifier, a Lasso classifier, a multivariate regression

based MV classifier and a Lasso selection multivariate regression based MVFS classifier.

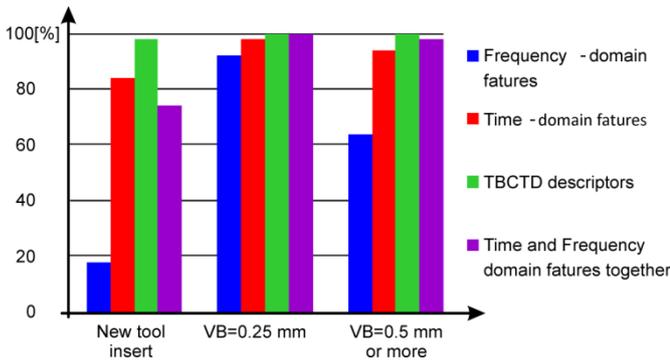


Figure 6. Recognition accuracy for each class separately for a FCM based classifier

On Figures 6 to 9, recognition accuracies are presented, for the mentioned types of classifiers, for each class of data separately. It can be seen that the best accuracies overall, are obtained using the FCM based classifier (results presented on Figure 6), where the proposed TBCTD descriptors have the best score in comparison to all other baseline feature sets. Moreover, although other classifiers, namely Lasso, MV and MVFS obtained much lower recognition accuracies (Figures 7, 8 and 9, respectively) for all features used, in comparison to FCM-based classifier, our TBCTD descriptors obtain overall higher recognition accuracies, when averaged over all 3 tool wear classes.

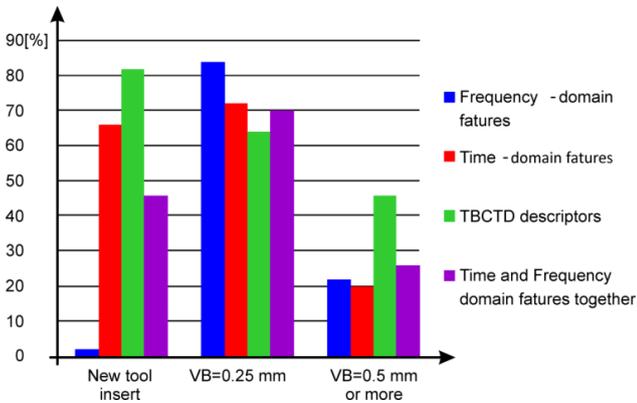


Figure 7. Recognition accuracy for each class separately for a Lasso classifier

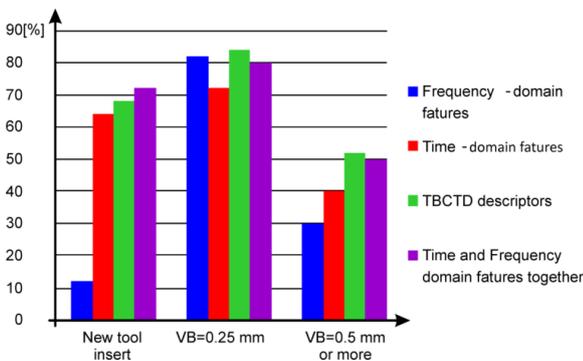


Figure 8. Recognition accuracy for each class separately for a MV classifier

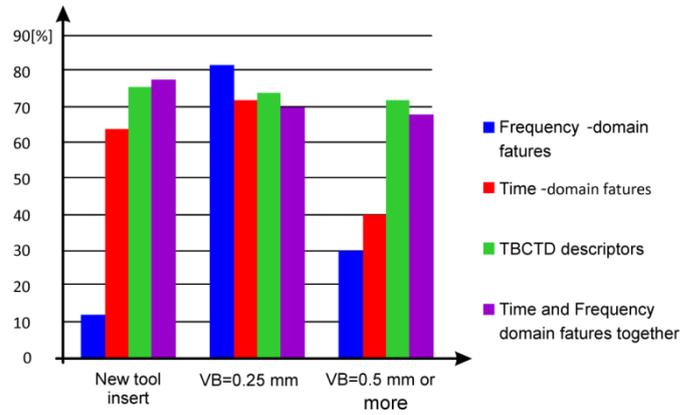


Figure 9. Recognition accuracy for each class separately for a MVFS classifier

The overall recognition accuracies of the proposed TBCTD descriptors in comparison to other baseline feature sets used are presented in Figure 10.

From the computational efficiency point of view, the feature extraction method that utilizes the proposed TBCTD descriptors, adds additional 18M operations (number of multiplications are taken into account, as it is commonly adopted complexity measure) per time frame, in comparison to the classical frequency based feature extraction that uses STDFT. Also, we can say that the additional computational load introduced by the moments extraction phase is the same, for the proposed and the baseline feature extraction methods. Moreover, it is clear that the additional memory requirements can be neglected.

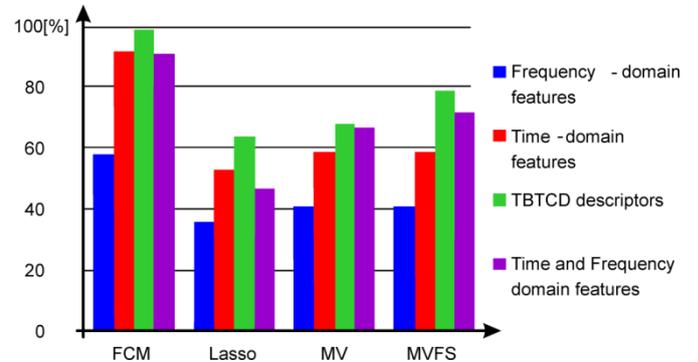


Figure 10. The overall recognition accuracies

## 6. CONCLUSION

In this work, we propose the Tool Wear Monitoring Strategy, which relies on the novel texture based descriptors. We use the module of the STDFT spectra obtained from the particular vibration sensors signal utterance and consider it as a 2D textured image, by identifying the time scale of STDFT as the x dimension, and the frequency scale as the y dimension of the image. For each band of interest, we extract information regarding the PDFs corresponding to those textons in the form of lower-order moments, where the averaging was

done in x dimension of the 2D textured image. Thus, we obtain discriminative and robust tool wear state descriptors. The experiments conducted on the real TCM data, show that the proposed descriptors, together with the fuzzy classifiers provide very high recognition accuracy in the tool wear state recognition task, higher than with the usage of classical time-domain and frequency-domain (or combined together) features. Moreover, the proposed TBTCDD descriptors obtain higher average recognition accuracies in comparison to the mentioned baseline features, when Lasso, MV and MVFS classifiers are used. Also, the additional computational and memory load introduced by the proposed features is fully acceptable from the point of view of modern embedded system abilities. Thus, we conclude, that the proposed TBTCDD descriptors are suitable for application in efficient TCM systems.

## ACKNOWLEDGMENTS

This paper presents a segment of the research on the project "Contemporary approaches in the development of special solutions bearing in mechanical engineering and medical prosthetics" project number TR 35025, financed by the Ministry of Education and Science of the Republic of Serbia

## 7. REFERENCES

- [1] Abellan-Nebot, J.V., Subirón, F.R. (2010). *A review of machining monitoring systems based on artificial intelligence process models*, International Journal of Advanced Manufacturing Technology, Vol. 47, pp. 237–257.
- [2] Cho, S., Binsaeid, S., Asfour, S. (2010). *Design of multisensor fusion-based tool condition monitoring system in end milling*, The International Journal of Advanced Manufacturing Technology, Vol. 46, pp. 681-694.
- [3] Sick, B. (2002). *On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research*, Mechanical Systems and Signal Processing, Vol. 16, pp. 487-546.
- [4] Bahr, B., Motavalli, S., Arfi, T. (1997). *Sensor fusion for monitoring machine tool conditions*, International Journal of Computer Integrated Manufacturing, Vol. 10, pp. 314-323.
- [5] Tsai, Y.-H., Chen, J.C., Lou, S.-J. (1999). *An in-process surface recognition system based on neural networks in end milling cutting operations*, International Journal of Machine Tools and Manufacture, Vol. 39, pp. 583-605.
- [6] Haber, R.E., Jiménez, J.E., Peres, C.R., Alique, J.R. (2004). *An investigation of tool-wear monitoring in a high-speed machining process*, Sensors and Actuators A: Physical, Vol. 116, pp. 539-545.
- [7] Stankovski, S., Ostojic, G., Senk, I., Rakic-Skokovic, M., Trivunovic, S., Kucevic, D. (2012). *Dairy cow monitoring by RFID*, Scientia Agricola, Vol. 69, (1), pp. 75-80.
- [8] Dimla, S. E. D., (2002). *The Correlation of Vibration Signal Features to Cutting Tool Wear in a Metal Turning Operation*, The International Journal of Advanced Manufacturing Technology, Vol. 19, pp. 705-713.
- [9] Jemielniak, K., Urbański, T., Kossakowska, J., Bombiński, S. (2012). *Tool condition monitoring based on numerous signal features*, The International Journal of Advanced Manufacturing Technology, Vol. 59, pp. 73-81.
- [10] Gajic, G., Stankovski, S., Ostojic, G., Tesic, Z., Miladinovic, L. (2014). *Method of evaluating the impact of ERP implementation critical success factors – a case study in oil and gas industries*, Enterprise Information Systems, Vol.8, (1), pp. 84-106.
- [11] Petru, J., Zlámal, T., Čep, R., Sadílek, M. i Stančeková, D. (2017). *The effect of feed rate on durability and wear of exchangeable cutting inserts during cutting Ni-625*, Tehnički vjesnik, 24 (Supplement 1), 1-6. <https://doi.org/10.17559/TV-20131221170237>
- [12] Mičietová, A., Neslušan, M., Čep, R., Ochodek, V., Mičiet, B. i Pagáč, M. (2017). *Detection of grinding burn through the high and low frequency Barkhausen noise*, Tehnički vjesnik, 24 (Supplement 1), 71-77. <https://doi.org/10.17559/TV-20140203083223>
- [13] Senk, I., Ostojic, G., Jovanovic, V., Tarjan, L., Stankovski, S. (2015). *Experiences in developing labs for a supervisory control and data acquisition course for undergraduate mechatronics education*, Computer Applications in Engineering Education, Vol. 23, (1), pp. 54-62.
- [14] Ostojic, G., Stankovski, S., Vukelic, D., (2011). Lazarevic, M., Hodolic, J., Tadic, B., Odri, S. Implementation of automatic identification technology in a process of fixture assembly/disassembly, Strojniski Vestnik/Journal of Mechanical Engineering, Vol. 57, (11), pp. 819-825.
- [15] Al-Habaibeh, A., Gindy, N. (2000). *A new approach for systematic design of condition monitoring systems for milling processes*, Journal of Materials Processing Technology, Vol. 107, pp. 243-251.
- [16] Rizal, M., Ghani, J.A., Nuawi, M.Z., Haron, C.H.C. (2013). *Online tool wear prediction system in the turning process using an adaptive neuro-fuzzy inference system*, Applied Soft Computing, Vol. 13, pp. 1960-1968.
- [17] Varma, M., Zisserman, A. (2009). *A Statistical Approach to Material Classification Using Image Patch Exemplars*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, pp. 2032-2047.
- [18] Varma, M., Zisserman, A. (2005). *A Statistical Approach to Texture Classification from Single Images*, International Journal of Computer Vision, Vol. 62, pp.61-81.

# Sistem za praćenje stanja alata zasnovan na teksturnim deskriptorima

Aco Antić, Milan Zeljković, Nicolae Ungureanu, Ivan Kuric

Primljen (21.02.2018.); Recenziran (29.08.2018.); Prihvaćen (04.09.2018.)

## Apstrakt

Svi savremeni sistemi za nadzor alata, posebno oni koji koriste senzore vibracije u zadatku prepoznavanja habanja alata, u velikoj meri zavise od izbora obeležja, izdvojenih iz određenih senzorskih signala, a koji sadrže informacije o stanju habanja alata. Većina primenjenih tehnika postprocesiranja ne uspeva povećati preciznost prepoznavanja ako ta obeležja nisu dovoljno diskriminativna. U ovom radu predlažemo strategiju nadzora habanja, koja se oslanja na izdvajanje novih obeležja baziranih na teksturi. Posmatran je segment kratkotrajnih diskretnih Furieovih transformacija (STDF) spektara vibracija, kao 2D teksture slike, dobijen od konkretnog senzorskog signala.

*Radi se identifikaciji vremenske skale STDFT kao prve dimenzije i skale frekvencije kao druga dimenzije definisane teksture slike. Dobijena tekstura slike se zatim deli na određene 2D segmente teksture, pokrivajući deo frekventnog opsega od interesa. Nakon primene odgovarajuće banke filtra, za svaki prethodno definisani frekventni opseg izdvajaju se 2D tekstoni. Dobijeni tekstoni, u posmatranom prosečnom vremenu, izdvajaju informacije o funkciji gustine verovatnoće (PDF) tih tekstona u obliku statističkih momenata nižeg reda, čime se dobijaju robusni deskriptori stanja habanja alata. Potvrda rada sistema sa predloženim obeležja je u realnim uslovima. Dobijeni rezultati daju visoku tačnost prepoznavanja.*

**Ključne reči:** *habanje alata, teksturni deskriptori, vibracioni signal*