



Technical note

A mel-frequency cepstral coefficient-based approach for surface roughness diagnosis in hard turning using acoustic signals and gaussian mixture models



Edielson P. Frigieri^{b,*}, Paulo H. S. Campos^a, Anderson P. Paiva^a, Pedro P. Balestrassi^a, João Roberto Ferreira^a, Carlos A. Ynoguti^b

^a Industrial Engineering Institute, Federal University of Itajuba, Minas Gerais, Brazil

^b Department of Computer Engineering, National Institute of Telecommunication, Minas Gerais, Brazil

ARTICLE INFO

Article history:

Received 10 May 2016

Received in revised form 28 June 2016

Accepted 29 June 2016

Available online 7 July 2016

Keywords:

Hard turning

Mel-frequency cepstral coefficients

Gaussian mixture models

Sound

Monitoring

ABSTRACT

During the last years, notable efforts have been made to develop reliable and industrially applicable machining monitoring systems based on different types of sensors, especially indirect methods that does not required to interrupt the machining process. As the main objective in machining processes is to produce a high-quality surface finish which, however, can be measured only at the end of the machining cycle, a more preferable method would be to monitor the quality during the cycle. Motivated by this premise, results of investigation on the relationship between audible sound emitted during process and the resulted surface finish are reported in this paper. Through experiments with AISI 52100 hardened steel, this work shows that such a correlation does exist between the surface roughness and the Mel-Frequency Cepstral Coefficients (MFCCs) and based on that correlation, a new quality monitoring method is proposed using Gaussian Mixture Models (GMM). Obtained results show that this method can identify three different levels of surface roughness with an average accuracy of 98.125%.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Customers who need parts machined place a premium on the quality of surface finish [1]. Thus to identify changes, failures, or tears in the machining process, researchers have studied and developed monitoring methods for many years, as can be seen in the overview presented by Teti et al. [2]. Monitoring methods can be classified as either direct or indirect, where direct methods analyze measurements of tool wear or surface roughness but, carrying out of such methods must either interrupt the machining process or wait until its completion [3]. Indirect methods however, can be carried out without stopping or interfering with the machining process, thereby increasing efficiency and allowing online adjustments [4]. To achieve this advantage it is necessary to make use of other sources of information, such as, cutting forces

[5], vibration signals [6], acoustic emission [7], electric current, image [8], and many other sources.

Among all source of information used indirect methods, the sound generated during machining process has recently attracted researchers attention [9–11]. Mainly because using the produced sound to gauge information about machining process has two major advantages: ease of installation [9] and implementation at a lower cost than other sensors [12].

Some examples of investigations into the use of sound emissions to determine the status of processes and structures include, but are not limited to, Lu and Wan [13] that analyzed the high-frequency sound signals (range between 20 and 80 kHz) generated in micro-milling process and proposed a tool wear monitoring method. Results indicated that the normalized sound signals can be potentially applied in monitoring methods with the proper selection of feature bandwidth and other parameters.

Also, airborne sound was found by Robben et al. [14] to be a valuable source of information in an ongoing machining process for the cut-off grinding of concrete. The authors highlighted that because of the very high sound emission of the machining process in a controlled environment, there was no total influence of environmental noise on the proposed monitoring method.

* Corresponding author at: Department of Computer Engineering, National Institute of Telecommunication, 510 João de Camargo Avenue, Santa Rita do Sapucaí, MG 37540-000, Brazil.

E-mail addresses: edielson@inatel.br (E.P. Frigieri), paulo.campos@unifei.edu.br (P.H.S. Campos), andersonppaiva@unifei.edu.br (A.P. Paiva), pedro@unifei.edu.br (P.P. Balestrassi), jorofe@unifei.edu.br (J.R. Ferreira), ynoguti@inatel.br (C.A. Ynoguti).

In the case of audible sound emissions in the frequency range spectrum of human hearing (between 0 and 20 kHz), Downey et al. [15] observed that it was possible to correlate the sound energy with tool wear in machining operations. Furthermore, audible acoustic spectra highlighted the possibility of identifying discrete phases in the cutting interface performance. This work was based in the premise that some experienced machinists are indeed capable of discerning between a good and a degraded machining process through variations in the audible sound emissions from the process.

Following this idea, the fundamental aim of this work is to determine if, besides detecting tool wear, it is possible to identify differences in audible acoustic emissions for different finishing surface roughness levels enabling its utilization in quality monitoring of machining process. One feature that can successfully model human auditory system is the Mel-Frequency Cepstral Coefficients (MFCC) [16], extensively used for speech recognition [17,18]. This feature has not been explored for monitoring machining process, what makes it the major contribution of this work.

Other important point to highlight is that most of presented work have focus on monitoring tool wear and analysis of data that is offered by the machining process from the sensor configurations. As the objective of a machining process is the final quality of the work-piece, actually many efforts are directed to find efficient ways to monitor the quality of the machining process, where surface roughness raised as an important parameter for quality monitoring [19,20]. Using the MFCC extracted from sound energy as acoustic spectrum features, a new surface roughness diagnosis approach based on Gaussian Mixture Models (GMM) is proposed.

The rest of the paper is organized as follows. Section 2 provides a detailed explanation of the methodology. In Section 3, the experiment is outlined showing all the steps for database creation. In Section 4, the obtained results are presented. Finally, conclusions and future work are given in Section 5.

2. The methodology

Fig. 1 presents an overview of the presented methodology. First, a windowing approach is used over the sound signal providing a frame-by-frame basis. Second, mel-frequency cepstral coefficients are extracted from the windowed sound signal. Next, extracted

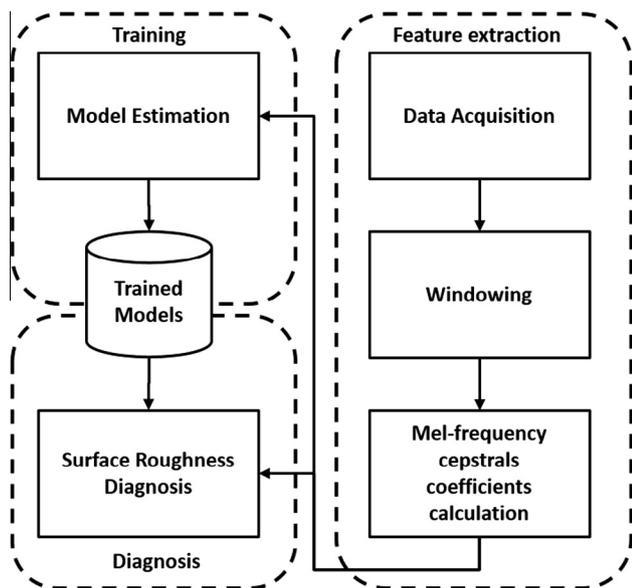


Fig. 1. Overview of the methodology.

coefficients can be used for training and estimate GMM models or, if all models are already estimated, used for surface roughness diagnosis. The methodology will be presented in 4 sections. Section 2.1 discusses how sound signal is divided in frames and the power spectral density is obtained for each frame. Section 2.2 shows the feature extraction technique used to calculate the mel-frequency cepstral coefficients. This is an important step since it is used in both system training and surface roughness diagnosis, as detailed in the next sections. Then, in Section 2.3 is presented a brief review on Gaussian Mixture Models and also the models estimation procedure is detailed. Finally, Section 2.4 presents the diagnosis procedure.

2.1. Signal processing and windowing in sound signals

The sound that comes from the machining process is a dynamic signal. It varies with time due to small differences in the machined material, vibration, fluctuations in rotating speed, and so forth. Therefore, it is reasonable to update the acoustic information on a frame-by-frame basis where a frame consists of F samples of the signal corresponding to a period of t_f seconds. This corresponds to the length of time (in seconds) over which a set of parameters is valid. Then the analysis moves t_w samples forward for a new frame, where $W \leq F$, resulting in superposition between adjacent frames. The amount of overlap to some extent controls how quickly parameters can change from frame to frame, which normally corresponds to an overlap of 50% [16].

Each frame is multiplied by a windowing function in order to smooth quick variations and increase continuity between frames [18,21]. For this work, it was used a Hamming function that can be calculated by

$$w(n) = 0.54 - 0.46 \cos(2\pi n / (F - 1)), 0 \leq n \leq F - 1$$

where F represents the frame size. The windowed frame can be expressed as

$$\tilde{x}_k(n) = x(n)w(n), 0 \leq n \leq F - 1$$

where F is the frame size. The total number of frames depends on the selected frame interval F and time shift W besides the total signal length.

Finally, for each k windowed frame is calculated the spectral energy $S_k = |X_k|^2$, where X_k is the Discrete Fourier Transform (DFT) of the windowed frame \tilde{x}_k , calculated through the FFT algorithm by

$$X_k = \sum_{n=0}^{N-1} \tilde{x}_k(n) e^{2\pi i k n / N}$$

S_k is composed by $N/2 + 1$ values of energy where N corresponds to the number of FFT points.

2.2. Mel-frequency cepstral coefficients extraction

Depending on the number of FFT points N , a correlation analysis of all $N/2 + 1$ energy values with each machining parameter and also finishing surface roughness may not be satisfactory since many components tend to be correlated. Also, for successful classification, features extracted from signal must eliminate as much irrelevant information as possible and retain only the significant information. Thus the feature vectors should attend the following requirements: first, be of low dimensionality to allow a reliable estimation of the free parameters of the recognition system; second, be independent of the recording environment, and finally be characteristic for each machining setup, to allow an optimal discrimination between the different acoustic models.

One such parameter that has achieved a great success in this task is the Mel-Frequency Cepstral Coefficients (MFCC), first proposed for the speech recognition problems [22]. Based on filter bank, can be regarded as a crude model of the initial stages of transduction in the human auditory system [16]. The main motivations for this kind of representation are: first, the position of maximum displacement along the basilar membrane for stimuli, such as pure tones, is proportional to the logarithm of the frequency of the tone; second, experiments in human perception have shown that frequencies of a complex sound within a certain bandwidth of some nominal frequency cannot be individually identified. When one of the components of this sound falls outside this bandwidth, it can be individually distinguished. We refer to this bandwidth as the critical bandwidth, which is nominally 10–20% of the center frequency of the sound [16].

The main idea behind Mel-scale is therefore a mapping between the real frequency scale (Hz) and the perceived frequency scale (Mels). Frequencies between 0 Hz and 1 kHz are linearly approximated and then a logarithmic scale is used for frequencies beyond 1 kHz. It is possible to define a mapping from the actual frequency f to a perceptual frequency scale by

$$f_{mel} = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \quad (1)$$

The process of extracting the MFCC can be outlined as follow. The first step is to separate the signal into K frames of size F , and calculate the energy spectrum S_k for each frame $k = 1, \dots, K$ [18]. This first step was detailed in Section 2.1. Next, each energy frame is filtered using a triangular filter bank, which center frequencies can be calculated by (1), resulting

$$\tilde{s}_k(l) = \sum_{n=0}^{N/2} S_k(n) M_l(n)$$

where N is the number of FFT points, L is the number of Mel-scale filters and M_l is the l th-filter [8,23]. Finally, a Discrete Cosine Transform (DCT) is applied to the natural logarithm of the Mel spectrum, resulting in the Mel-frequency Cepstral Coefficients, as follows

$$c_k(m) = \sum_{l=0}^{L-1} \log(\tilde{s}_k(l)) \cos \left(\frac{\pi m}{2L} (2l+1) \right), \quad \forall k = 1, \dots, K \quad (2)$$

where $m = 1, 2, \dots, C$ and C is the number of desired coefficients. Since most of the signal information is represented by the first coefficients, normally C is chosen between 12 and 20 [24,21].

2.3. Estimation of surface roughness models

Gaussian mixture model (GMM) was adopted in order to perform surface roughness diagnosis, as it provides a probabilistic model of the underlying sound. The use of GMM is present in many applications where features extracted from audible sound are used for some kind of classification [25,17]. According to [26], for a D dimensional feature vector \vec{x} , a Gaussian mixture density is given by

$$p(\vec{x}|\lambda) = \sum_i^M w_i b_i(\vec{x}) \quad (3)$$

The density $p(\vec{x}|\lambda)$ is a weighted linear combination of $i = 1, \dots, M$ component densities $b_i(\vec{x})$, each parameterized by a $D \times 1$ mean vector, $\vec{\mu}_i$, a $D \times D$ co-variance matrix, Σ_i , and D -mixture weights w_i that satisfy the constraint that $\sum_{i=1}^M w_i = 1$ [23,24]. Each component density $b_i(\vec{x})$ is a D -variate Gaussian function of the form

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_i)' \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right\}$$

In summary, the complete Gaussian mixture density λ can be parameterized by $\lambda = (w_i, \vec{\mu}_i, \Sigma_i)$. The number of mixture components is empirically chosen for a given data set.

In order to generate reliable and accurate results, the GMM-based pattern recognition technique, as with other techniques (neural network, HMM, etc.), needs a training procedure prior to diagnosis.

After extracting features from signals (see details in Sections 2.1 and 2.2), it is necessary to find the best cluster separation that can accurately represent the different desired classes. Once defining such clusters, models parameters $(w_i, \vec{\mu}_i, \Sigma_i)$ can be estimated using the expectation–maximization (EM) algorithm [27]. For a sequence of T training vectors $X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_T)$, the GMM likelihood can be calculated as $p(X|\lambda) = \prod_{t=1}^T p(\vec{x}_t|\lambda)$. So, the parameters are adjusted on each iteration, increasing the likelihood of the estimated model, that is, for iterations k and $k+1$, $p(X|\lambda^{k+1}) > p(X|\lambda^k)$. On each iteration, the mixture weights, means, and variances are re-estimated using the respective formulas

$$\bar{w}_i = \frac{1}{T} \sum_{t=1}^T p(i|\vec{x}_t, \lambda)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T p(i|\vec{x}_t, \lambda) \vec{x}_t}{\sum_{t=1}^T p(i|\vec{x}_t, \lambda)}$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T p(i|\vec{x}_t, \lambda) \vec{x}_t \vec{x}_t' - \bar{\mu}_i \bar{\mu}_i'}{\sum_{t=1}^T p(i|\vec{x}_t, \lambda)}$$

The *a posteriori* probability for a class i can be calculated as

$$p(i|\vec{x}_t, \lambda) = \frac{w_i b_i(\vec{x}_t)}{\sum_{k=1}^M w_k b_k(\vec{x}_t)}$$

2.4. Surface roughness diagnosis

Supposing that a group of S surface roughness clusters obtained in the training procedure are represented by the GMM models $\lambda_1, \lambda_2, \dots, \lambda_S$. Surface roughness diagnosis in the process of finding the model λ with the maximum likelihood representation, calculating

$$\hat{S} = \arg \max_{1 \leq k \leq S} p(X|\lambda_k)$$

where $X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_T)$ is the feature vector sequence under analysis. Assuming that the feature vectors of X are independent, the log-likelihood of a model λ can be computed by

$$\hat{S} = \arg \max_{1 \leq k \leq S} \sum_{t=1}^T \log p(\vec{x}_t|\lambda_k)$$

where $p(\vec{x}_t|\lambda_k)$ is calculated by (3).

3. Experimental outline

3.1. Experimental setup

The experimental setup is based on a CNC lathe with a maximum rotational speed of 4000 rpm and power of 5.5 kW and using Wiper mixed ceramic (Al₂O₃ + TiC) inserts (ISO code CNGA 120408S01525WH) coated with a very thin layer of titanium

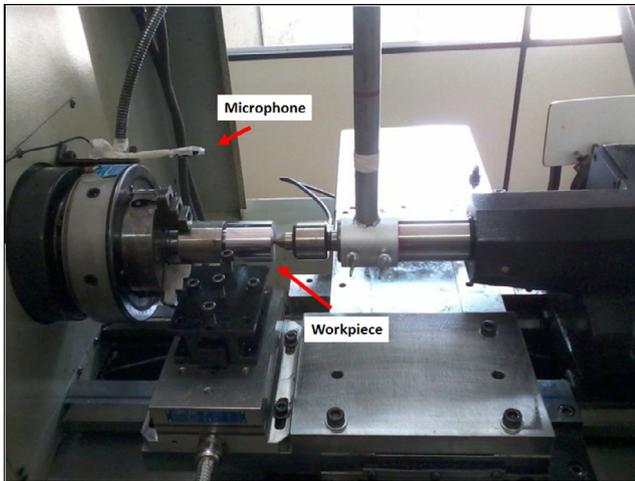


Fig. 2. Machining setup.

nitride (TiN; Sandvik-Coromant GC 6050). The work-pieces were made of AISI 52100 steel with dimensions of 49 mm 50 mm and a hardness between 49 and 52 HRC, up to a depth of 3 mm below the surface. Details related to the experimental setup such as the location of microphone in relation to cutting interface and the tool holding configuration can be seen in Fig. 2.

Sound emissions were collected using an audio microphone connected to the sound card of a Dell Vostro laptop, recorded through the software Audacity [28] using a sample frequency of 44.1 kHz and 16 bits resolution. Respective surface roughness was measured using a Mitutoyo portable roughness meter model SurfTest SJ 201, fixed to a cut-off length of 0.25 mm.

3.2. Experimental methodology

In order to determine the machining setup parameters, a Central Composite Design (CCD) was used to create a sequential set of experimental runs [29]. The number of tools available for the experiment led us to choose 10 sets of machining setups which were selected based on D-optimality criteria [30]. Table 1 shows the resulted parameter values for each experimental run considering cutting speed (CS m/min), feed rate (F mm/rev), depth of cut (D mm) and material removal rate (MRR).

Each experimental run was replicated 15 times to guarantee the statistical reliability of the experiment, resulting a total of 150 experimental runs. During the experiment, it was guaranteed that there were no other machining operations taking place in the vicinity, which might have added interfering acoustic components and compromise the signals being detected. For each execution, sound signal was recorded and the following surface roughness

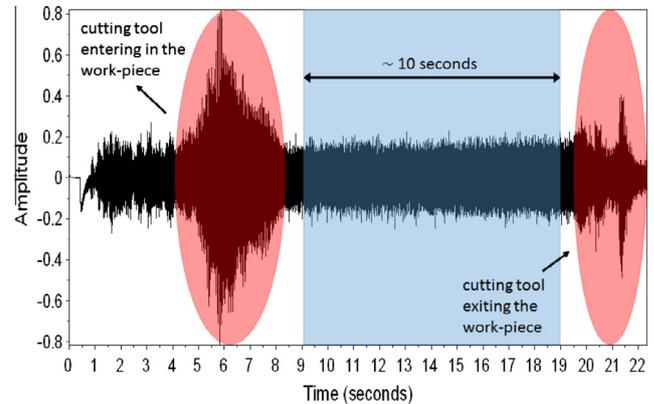


Fig. 3. Full audio recorded during machining ($CS = 200$, $F = 0.10$ and $D = 0.10$).

parameters were measured: arithmetic average surface roughness (R_a), maximum surface roughness (R_y), root mean square roughness (R_q), 10-point height (R_z), and maximum peak to valley (R_t), measured three times at four positions in the work-piece middle, as proposed by Paiva et al. [29]. Finally, mean and variance were calculated for each surface roughness obtained resulting the values presented in Table 1. For each experimental cycle, the only change was the replacement of the cutting tool insert in order to guarantee the same tool wear condition.

3.3. Feature extraction

As outlined, measurements of the surface finish of the work-piece were taken regularly during the experimental runs. Samples in the range of 7–10 s duration were taken from each recorded audio data, for all experimental runs. The sample time range was chosen considering just the stable cutting period avoiding moments when cutting tool enters and exits of the work-piece [15], as illustrated by Fig. 3. Feature vectors were extracted, according to Section 2.2, using $N = 1024$ FFT points, a frame length $t_f = 20$ ms and frame shift $t_w = 10$ ms, that are common values used in speech analysis [23] and also applied in machining monitoring methods [11]. On average, a total of $K = 900$ frames was obtained from each data file, each one composed by $C = 12$ Mel-frequency Cepstral Coefficients, calculated using $L = 31$ filters.

4. Results

4.1. Correlation analysis

Fig. 4 represents the spectrogram of two different machining setups (Ms_1 and Ms_7), where it is possible to note that power spectral density has a stochastic behavior along time. As the MFCC feature

Table 1
Surface roughness average for all machining setups.

Machining setup	Parameters				Surface roughness					
	CS	F	D	MRR	R_a	R_y	R_z	R_q	R_t	R_{sm}
Ms_1	200.00	0.10	0.10	2.00	0.15	1.05	0.75	0.18	1.44	85.60
Ms_2	240.00	0.10	0.10	2.40	0.16	1.17	0.81	0.20	1.56	91.30
Ms_3	200.00	0.20	0.10	4.00	0.44	2.39	1.76	0.53	3.09	169.72
Ms_4	200.00	0.10	0.20	4.00	0.19	1.36	1.00	0.24	1.80	58.91
Ms_5	240.00	0.10	0.20	4.80	0.18	1.32	0.89	0.23	1.89	100.67
Ms_6	240.00	0.20	0.20	9.60	0.52	2.90	2.24	0.63	3.41	195.46
Ms_7	186.36	0.15	0.15	4.19	0.26	1.87	1.25	0.33	2.71	104.31
Ms_8	220.00	0.23	0.15	7.72	0.50	2.62	2.06	0.60	3.10	212.34
Ms_9	220.00	0.15	0.23	7.72	0.26	1.90	1.29	0.33	2.70	73.04
Ms_{10}	220.00	0.15	0.15	4.95	0.19	1.42	0.97	0.24	2.10	88.14

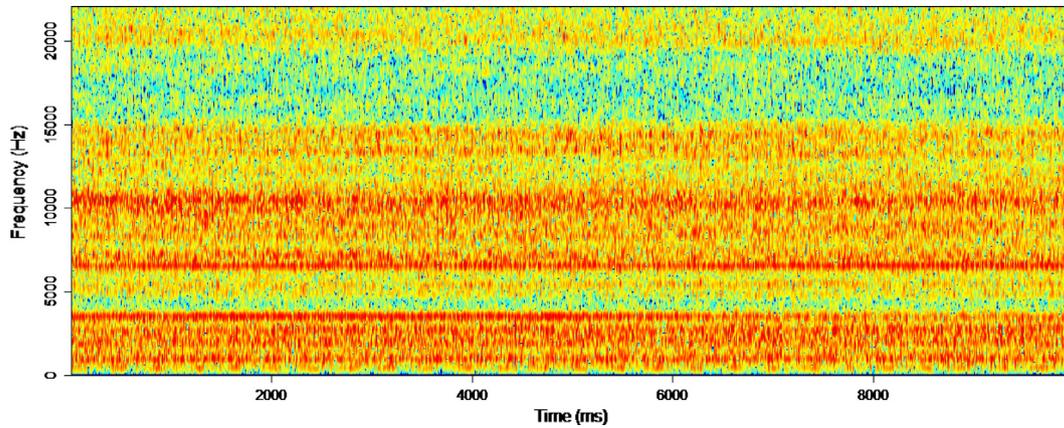
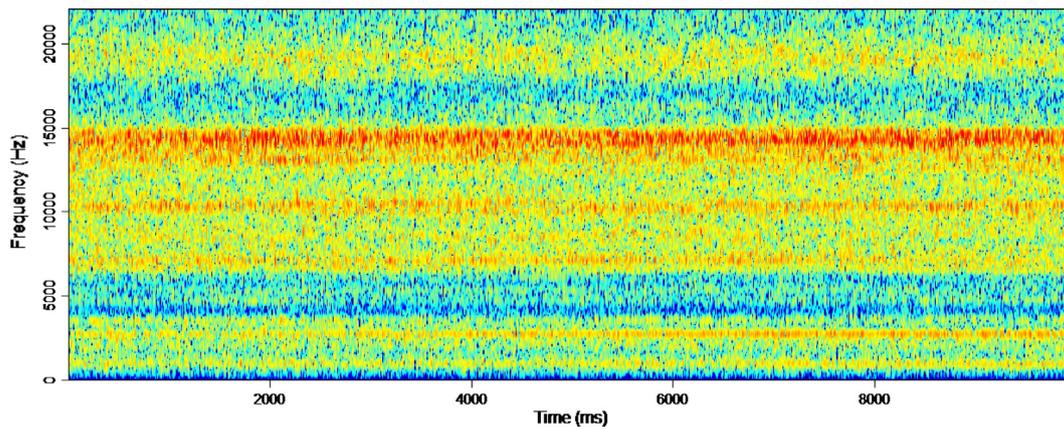
(a) Machining setup 1 (Ms_1).(b) Machining setup 7 (Ms_7).

Fig. 4. Resulted spectrogram for different machining setups.

Table 2

Correlation analysis between mel-frequency cepstral coefficients, machining setup parameter and surface roughness.

MFCC	CS	F	D	MRR	Ra	Ry	Rz	Rq	Rt	Rsm
c_1	0.061 ^a	0.594^A	0.563	0.781	0.595	0.665	0.629	0.607	0.685	0.419
c_2	-0.069	-0.703	-0.506	-0.802	-0.649	-0.722	-0.671	-0.662	-0.752	-0.507
c_3	0.123	0.430	0.210	0.414	0.228	0.265	0.220	0.235	0.318	0.285
c_4	0.440	-0.240	0.334	0.204	-0.040	-0.053	-0.003	-0.048	-0.158	-0.097
c_5	0.673	0.029	0.437	0.432	0.003	0.037	0.037	0.006	0.018	0.021
c_6	0.566	0.001	-0.098	0.058	-0.052	-0.126	-0.066	-0.063	-0.216	0.093
c_7	-0.069	0.029	0.077	0.106	-0.081	-0.058	-0.039	-0.072	-0.037	-0.135
c_8	-0.607	0.207	-0.228	-0.173	0.063	0.076	0.051	0.068	0.165	-0.031
c_9	-0.562	-0.021	-0.361	-0.296	-0.032	-0.066	-0.069	-0.036	-0.043	0.010
c_{10}	0.013	0.239	-0.032	0.110	0.291	0.237	0.237	0.280	0.206	0.346
c_{11}	0.616	-0.037	0.271	0.324	0.062	0.062	0.065	0.059	0.014	0.159
c_{12}	0.614	0.052	0.334	0.363	-0.020	0.025	0.012	-0.013	0.024	0.006

^A Bold values represent correlations that are statistically significant (P -Value < 5%).^a Pearson correlation.

vector describes the power spectral envelope of a single frame, this stochastic behavior means that there is no information in the trajectories of the MFCC coefficients over time (dynamics), and time based features as delta or delta-delta are not applicable. Based on that information, the MFCC of all K frames can be substituted by one vector containing their average, i.e., the MFCC of each machining setup can be represented by $c_m = 1/K \sum_{k=1}^K c_k(m)$, $\forall m = 1, 2, \dots, C$ and $c_k(m)$ obtained by (2) for each frame k .

Through correlation analysis was possible to identify which Mel-Frequency Cepstral Coefficients were strongly correlated with machining parameters and surface roughness (see Table 2). Each Mel coefficient presented higher correlation with different machining parameter, i.e., coefficient c_2 is correlated with material remove rate (MRR) and also with all finishing surface roughness parameter (Ra , Ry , Rz , Rq , Rt and Rsm), where maximum peak to valley surface roughness (Rt) presented highest correlation. Already c_5 coefficient

presented stronger correlation with cutting speed (CS) machining parameter. Fig. 5 illustrates how second coefficient (c_2) is strongly negative correlated with surface roughness measures. When surface roughness level increases, c_2 has the same behavior but in the opposite way, which explains the negative correlation.

With the aim of highlight such a correlation, response surface analysis was applied in two replicas randomly selected resulting the regression coefficients, R-Sq (adj.) and Anderson–Darling normality test for residuals of each response, presented in Table 3. Results have presented not only adequate coefficients of determination (above 95%, exception for $R_t = 86.13\%$), but also enough evidence to affirm that residuals are following normal distributions. As can be observed in Table 3, feed rate F was the most important factor explaining the average behavior of surface roughness parameters and was also one of the most import factors explaining c_2 . According to Tekner and Yeşlyurt [31], surface roughness is strongly correlated with feed rate (F) so this similarity might be one of the reasons why strong correlations between surface roughness parameters and MFCC were observed.

The presented analysis showed that the cutting specific energy (related to the MRR) applied to the system is probably transformed in other types of energy like vibration, heat (tool edge, work-piece, chip and air), sound and others. In this case, the part of the energy transferred to the audible sound can be identified by MFCC, proving that this feature can be applied in monitoring methods.

4.2. Cluster criteria definition

To recognize the patterns of the acoustic signals and correlate them with the surface roughness measured, it is necessary to evaluate in how many distinct clusters the measured surface roughness can be separated. Using the statistical technique of cluster hierarchical analysis based on the Ward method [32], the machining setups were separated into three different clusters: $S_1 = \{Ms_1, Ms_2\}$, $S_2 = \{Ms_3, Ms_4, Ms_5, Ms_7, Ms_{10}\}$ and $S_3 = \{Ms_6, Ms_8, Ms_9\}$. The similarity among the machining setups was evaluated regarding the behavior of the average surface roughness Ra and the material removal rate MRR . As can be observed, some machining setups resulted a surface roughness Ra not sufficiently different to support the hypothesis that it is possible to separate the 10 machining setups. In this way, some different machining setups can result similar surface roughness levels.

4.3. Surface roughness diagnosis

Based on the obtained clusters, two training and testing sets were separated using the 150 total sound data files. A second

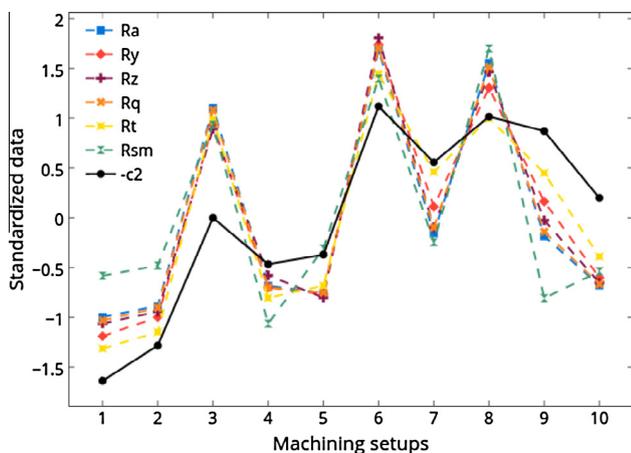


Fig. 5. Line plots for standardized surface roughness parameters and Mel-frequency cepstral coefficient ($-c_2$).

Table 3 Regression coefficients and coefficients of determination.

	Ra	Ry	Rz	Rq	Rt	c_2
Constant	0.185^A	1.387	0.958	0.236	2.133	-32.234
CS	0.074	0.202	0.24	0.079	0.074	2.765
F	0.169	0.746	0.616	0.194	0.773	-9.278
D	-0.053	-0.035	-0.111	-0.051	0.075	-10.073
CS ²	0.075	0.33	0.267	0.084	0.201	0.091
F ²	0.014	0.002	0.044	0.015	-0.123	0.78
D ²	0.057	0.198	0.172	0.062	0.162	3.079
CS × F	0.077	0.197	0.266	0.081	0.043	5.743
CS × D	-0.01	-0.051	-0.052	-0.011	-0.009	0.743
F × D	-0.063	-0.081	-0.159	-0.063	0.046	-5.759
S	0.013	0.135	0.085	0.018	0.262	1.283
R-Sq(adj)	99.27%	95.79%	97.62%	98.80%	86.13%	98.70%
Normality test	0.433 ^a	0.268	0.117	0.118	0.328	0.382
For residuals	0.273 ^b	0.646	0.988	0.988	0.492	0.365

^A Bold values represent the significant terms in the models (P -Value < 5%).

^a Anderson–Darling statistic test.

^b P -Value.

Table 4 Gaussian mixture model for surface roughness cluster S_1 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.12	-18.42	-17	24.56	27.1
2	0.12	-8.45	-8.81	36.79	17.75
3	0.13	-9.44	0.43	5.71	5.2
4	0.11	-6.92	-5.04	19.67	6.2
5	0.21	-6.04	4.42	5.08	5.56
6	0.14	-1.52	8.8	8.73	6.93
7	0.13	2.24	-4.22	19.27	7.5
8	0.04	11.11	-1.28	6.67	5.43

Table 5 Gaussian mixture model for surface roughness cluster S_2 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.05	-0.85	-31.66	25.84	43.28
2	0.06	6.6	-36.51	29.29	14.68
3	0.12	-2.28	-8.48	12.27	14.09
4	0.06	-0.87	-26.41	33.89	15.11
5	0.13	-10.75	-1.44	7.32	17.53
6	0.17	3.81	-2.14	11.12	9.21
7	0.26	-6.73	4.45	7.42	13.8
8	0.14	-1.49	7.99	13.17	10.26

Table 6 Gaussian mixture model for surface roughness cluster S_3 .

Density	w	μ_{c1}	μ_{c2}	σ_{c1}	σ_{c2}
1	0.01	3.13	-35.85	17.85	14.03
2	0.04	17.39	-38.88	18.77	10.02
3	0.06	15.61	-45.36	46.75	18.78
4	0.05	27.25	-45.45	39.84	15.5
5	0.1	-0.19	-9.03	8.48	10.13
6	0.13	4.09	-3.99	6.11	7.09
7	0.01	-17.38	-10.67	83.47	17.05
8	0.61	-3.69	3.87	22.1	16.3

training and testing set was used as a confirmation experiment to ensure the statistical accuracy of the proposed system. For each machining setup, eight of the 15 sound data files were separated for training and seven for testing. According to the number of machining setups in the cluster, the number of training and testing data files can vary from 16 to 40 and from 14 to 35, respectively, considering that each machining setup has 15 different replicas.

Table 7
Surface roughness diagnosis accuracy.

Cluster No.	Cutting speed, feed rate, depth of cut	Testing set	No. of replicas	Replicas	Accuracy (%)
2	186 m/min, 0.15 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
3	220 m/min, 0.23 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	85.7
3	220 m/min, 0.15 mm/rev, 0.23 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
1	240 m/min, 0.10 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	240 m/min, 0.10 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	220 m/min, 0.15 mm/rev, 0.15 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
3	240 m/min, 0.20 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	85.7
1	200 m/min, 0.10 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	200 m/min, 0.20 mm/rev, 0.10 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	200 m/min, 0.10 mm/rev, 0.20 mm	1	7	2, 4, 6, 8, 10, 12, 14	100
2	186 m/min, 0.15 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
3	220 m/min, 0.23 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	85.7
3	220 m/min, 0.15 mm/rev, 0.23 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
1	240 m/min, 0.10 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	240 m/min, 0.10 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	220 m/min, 0.15 mm/rev, 0.15 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
3	240 m/min, 0.20 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
1	200 m/min, 0.10 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	200 m/min, 0.20 mm/rev, 0.10 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
2	200 m/min, 0.10 mm/rev, 0.20 mm	2	7	3, 4, 7, 8, 11, 12, 15	100
Average					98.125

GMM model estimation algorithm was implemented in C programming language according to the method described in Section 2.3. Each surface roughness model, identified as S_1, S_2, S_3 , was estimated using $M = 8$ component densities. Observations sequence used during model estimation were composed only by the first and second Mel-frequency cepstral coefficients, as they presented the highest correlation with surface roughness parameters (see Section 4.1). The resulted GMM models for each surface roughness cluster can be seen on Tables 4–6.

Surface roughness diagnosis procedure was also developed using C programming language according to the method described in Section 2.4. Observations sequence, composed by the first and second mel-frequency cepstral coefficients, were used to find the model with the maximum likelihood. The testing set was used to access the accuracy of the proposed system calculating the number of sound signals correctly classified divided by the total number of sound signals tested. The proposed method achieved an average accuracy rate of 98.125%, as presented in Table 7. As the testing set for each surface roughness cluster is composed by different machining setups, this percentage indicates that the method is very robust and is not sensitive to cutting speed, feed rate, or depth-of-cut parameters.

5. Conclusions

Machining parameters were varied based on a Central Composite Design (CCD) in order to characterize the audible sound energy signals emitted by different cutting conditions during turning of AISI-5210 hardened steel. The corresponding acoustic signals were detected and processed in the frequency domain by extracting the Mel-Frequency Cepstral Coefficients.

The linear relationship between MFCC levels and machining parameters were evaluated through Pearson correlation analysis where it was possible to detect strong correlations such as c_1 and c_2 with material removal rate (MRR) and all surface roughness, and also c_5 with cutting speed (CS). To better understand such relationship, graphical analysis was used to demonstrate the linear relationship since changes in the machining parameters resulted in proportional changes in the MFCC levels.

Also, a response surface analysis was applied in the surface roughness parameters and MFCC. Feed rate (F) was the most

important factor explaining the average behavior of surface roughness parameters and was also one of the most important factors explaining second Mel coefficient (c_2), which highlights the strong correlations between surface roughness parameters and MFCC. Finally, using the MFCC as sound feature, this paper proposed a quality monitoring method based on Gaussian Mixture Models. Presented results validated the method and the effectiveness was evidenced by a diagnosis accuracy of 98.125%.

For future study, the proposed method could be analyzed using other machining processes like milling, drilling, welding and others. Also for future work, the MFCC applied with ANN or even SVM could be tested.

Acknowledgment

This research was supported by CAPES under process number BEX 3203/15-8. Thanks also to the Brazilian governmental agencies of CNPq and FAPEMIG.

References

- Özel T, Karpat Y. Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. *Int J Mach Tools Manuf* 2005;45:467–79.
- Teti R, Jemielniak K, O'Donnell G, Dornfeld D. Advanced monitoring of machining operations. *CIRP Ann – Manuf Technol* 2010;59:717–39.
- Sick B. On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mech Syst Signal Process* 2002;16:487–546.
- Chen B, Chen X, Li B, He Z, Cao H, Cai G. Reliability estimation for cutting tools based on logistic regression model using vibration signals. *Mech Syst Signal Process* 2011;25:2526–37.
- Dimla Sr DE, Lister PM. On-line metal cutting tool condition monitoring. I: force and vibration analyses. *Int J Mach Tools Manuf* 2000;40:739–68.
- Xiqing M, Chuangwen X. Tool wear monitoring of acoustic emission signals from milling processes. In: 2009 First international workshop on education technology and computer science. IEEE; 2009. p. 431–5.
- Van Hecke B, Yoon J, He D. Low speed bearing fault diagnosis using acoustic emission sensors. *Appl Acoust* 2016;105:35–44.
- Kasban H, Zahran O, Arafa H, El-Kordy M, Elaraby S, Abd El-Samie F. Welding defect detection from radiography images with a cepstral approach. *NDT & E Int* 2011;44:226–31.
- Rubio EM, Teti R. Process monitoring systems for machining using audible sound energy sensors. In: Aized T, editor. *Future manufacturing systems*. Sciyo; 2010. p. 217–35.
- Mannan M, Kassim Aa, Jing M. Application of image and sound analysis techniques to monitor the condition of cutting tools. *Pattern Recogn Lett* 2000;21:969–79.

- [11] Ai CS, Sun YJ, He GW, Ze XB, Li W, Mao K. The milling tool wear monitoring using the acoustic spectrum. *Int J Adv Manuf Technol* 2011;61:457–63.
- [12] Raja JE, Lim WS, Venkateshaiah C. Emitted sound analysis for tool flank wear monitoring using hilbert huang transform. *Int J Comput Electr Eng* 2012;4.
- [13] Lu M-C, Wan B-S. Study of high-frequency sound signals for tool wear monitoring in micromilling. *Int J Adv Manuf Technol* 2012;1785–92.
- [14] Robben L, Rahman S, Buhl JC, Denkena B, Konopatzki B. Airborne sound emission as a process monitoring tool in the cut-off grinding of concrete. *Appl Acoust* 2010;71:52–60.
- [15] Downey J, O'Leary P, Raghavendra R. Comparison and analysis of audible sound energy emissions during single point machining of HSTS with PVD TiCN cutter insert across full tool life. *Wear* 2014;313:53–62.
- [16] Picone J. Signal modeling techniques in speech recognition. *Proc IEEE* 1993;81:1215–47.
- [17] Lu Y, Wu Z. Maximum likelihood normalization for robust speech recognition. *Appl Acoust* 2013;74:640–6.
- [18] Koolagudi SG, Rastogi D, Rao KS. Identification of Language using Mel-Frequency Cepstral Coefficients (MFCC). *Procedia Eng* 2012;38:3391–8.
- [19] Lee SS, Chen JC. On-line surface roughness recognition system using artificial neural networks system in turning operations. *Int J Adv Manuf Technol* 2003;22:498–509.
- [20] Khorasani A, Yazdi MRS. Development of a dynamic surface roughness monitoring system based on artificial neural networks (ANN) in milling operation. *Int J Adv Manuf Technol* 2015.
- [21] Fahmy MM. Palmprint recognition based on mel frequency cepstral coefficients feature extraction. *Ain Shams Eng J* 2010;1:39–47.
- [22] Davis S, Mermelstein P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans Acoust Speech Signal Process* 1980;28:357–66.
- [23] Kumari RS, Nidhyanthan SS, G A. Fused mel feature sets based text-independent speaker identification using gaussian mixture model. *Procedia Eng* 2012;30:319–26.
- [24] Dhanalakshmi P, Palanivel S, Ramalingam V. Classification of audio signals using AANN and GMM. *Appl Soft Comput* 2011;11:716–23.
- [25] De Oliveira AG, Ventura TM, Ganchev TD, De Figueiredo JM, Jahn O, Marques MI, et al. Bird acoustic activity detection based on morphological filtering of the spectrogram. *Appl Acoust* 2015;98:34–42.
- [26] Reynolds DA, Rose RC. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Trans Speech Audio Process* 1995:72–83.
- [27] Dempster AP, Laird NM, Rubin D. Maximum likelihood from incompatible data via EM algorithm. *J Roy Stat Soc* 1977;39:1–38.
- [28] Dannenberg R. *Audacity*; 2013.
- [29] Paiva A, Campos P, Ferreira J, Lopes L, Paiva E, Balestrassi P. A multivariate robust parameter design approach for optimization of AISI 52100 hardened steel turning with wiper mixed ceramic tool. *Int J Refract Met Hard Mater* 2012;30:152–63.
- [30] Montgomery DC. *Design and analysis of experiments*. 8 ed. Hoboken: John Wiley and Sons; 2013.
- [31] Tekner Z, Yeşilyurt S. Investigation of the cutting parameters depending on process sound during turning of AISI 304 austenitic stainless steel. *Mater Des* 2004;25:507–13.
- [32] Johnson RA, Wichern DW. *Applied multivariate statistical analysis*, 5 ed., vol. 6. New Jersey: Prentice-Hall Inc.; 2007.