Application of methods of multidimensional statistical analysis for the classification of signal fragments during cutting processing

A V Belov¹, N N Trufanov², D V Churikov²,³ and O V Kravchenko¹,²,⁴

¹Bauman Moscow State Technical University, 2 Baumanskaya Street, Moscow, 105005, Russian Federation
²Scientific and Technological Center of Unique Instrumentation of the Russian Academy of Sciences, 15 Butlerova Street, Moscow, 117342, Russian Federation
³Kotel'nikov Institute of Radio-Engineering and Electronics of RAS, 11-7 Mokhovaya Street, Moscow, 125009, Russian Federation
⁴Federal Research Center “Computer Science and Control” of RAS, 40 Vavilova Street, Moscow, 119333, Russian Federation

E-mail: trufan0vn@yandex.ru

Abstract. The preliminary processing of the signal coming from the vibration sensor is performed. Spectral analysis methods allow identifying the distinctive zones corresponding to three parameters: the average amplitude of vibrations, the average value of the spectrum, and the complex parameter of the vibration density obtained using interpolation methods. Hence, the vector contains the listed parameters as components produce a vector space, and the equations of separating surfaces in three-dimensional space and classification errors are calculated.

1. Introduction
The addition of industrial equipment with innovative computing resources allows for real-time equipment diagnostics for early detection and troubleshooting, which allows us to improve the quality of the production process by reducing unforeseen situations [1]. The timely performance of equipment maintenance, depending on its actual condition, minimizes downtime and equipment setup costs, but at the same time, it requires a large amount of data to build a model for analyzing the system as a whole [2]. A literature survey shows that an accurate system for monitoring the condition of industrial equipment can increase productivity by 10–50% [3–6].

During the experiment, the vibration sensor generates real-time data, which is then processed. The spectral analysis methods allow us to identify distinctive zones, each corresponding to three parameters: the average oscillation amplitude, the average value of the spectrum, and the complex parameter of the oscillation density obtained using interpolation methods. Static methods allow us to calculate classification errors. Namely, the equations of separating surfaces in three-dimensional space are found, considering the listed parameters.

Currently, the critical point in metalworking is controlling the flow of the technological process and its analysis. One of the essential links in this chain is non-destructive testing in general and
vibration diagnostics in particular. In this paper, we consider the problem of classifying a fragment of a signal coming from an accelerometer placed on the front panel of the Dobot Mooz 2 machine is shown in figure 1, by methods of multidimensional statistical analysis.

![General view of the experimental setup with stationary accelerometer board.](image1)

**Figure 1.** General view of the experimental setup with stationary accelerometer board.

The received signal undergoes preliminary processing, as a result of which it becomes dimensionless and also normalizes on a segment [0,1]. Hence, one may get the signal image shown in figure 2.

![CNC control program trajectory in the X-Y plane.](image2)

**Figure 2.** CNC control program trajectory in the X-Y plane.

Classification is the assignment of a sample to one of several pairwise non-intersecting sets. In our case, we divide the processing signal into distinctive zones (figure 3) and construct a separating plane using discriminant analysis methods. The zoning mentioned above allows one to highlight the distinctive parts in the signal [7].
2. Methodology

Let’s assign three parameters to each zone:

1. Average oscillation amplitude

\[ S_i = \frac{1}{n_i} \sum_{k=1}^{n_i} v_k, \]

where \( n_i \) is a number of points in the \( i \)-th zone and \( v_j \) is a fragment of the signal corresponding to the \( i \)-th zone [8].

2. The average amplitude of the Fourier transform

\[ F_i = \frac{1}{N_j} \sum_{k=0}^{N_j-1} \sum_{n=0}^{N_j-1} v_{i,n} \exp \left( -2 \pi i \frac{k n}{N_j} \right). \]

3. The average area between the envelopes.

For the initial signal, we calculate the upper and lower envelopes are introduced in figure 4.

![Figure 3](image1.png)

**Figure 3.** Distinctive parts in the signal.

Further, for each zone we define the parameter

\[ T_i = \frac{1}{n_i} \int_{t_i}^{t_{i+1}} (f_u - f_l) dt, \]

where \( n_i \) is a number of points in the \( i \)-th zone, \( f_u \) and \( f_l \) are upper and lower envelopes obtained by cubic interpolation, \( t_i \) and \( t_{i+1} \) are borders for the \( i \)-th zone.

![Figure 4](image2.png)

**Figure 4.** Envelopes for the signal.
Now, consider random variables
\[ \xi_i = \begin{pmatrix} S_i \\ F_i \\ T_i \end{pmatrix}, \eta_j = \begin{pmatrix} S_j \\ F_j \\ T_j \end{pmatrix}, \]
provided that the index \( i \) corresponds to points from the first set, and the index \( j \) corresponds to points from the second set. Then, from the assumptions of normality and homoscedasticity of the obtained vectors, one can obtain
\[ \xi \sim N(\mu_i, D), \eta \sim N(\mu_j, D). \]

Let's introduce the notation
\[ n_p(\gamma, \mu, D) = \frac{1}{(\sqrt{2\pi})^p \cdot \sqrt{D}} e^{-\frac{1}{2}(\gamma - \mu)^T D^{-1} (\gamma - \mu)}. \]
is the normal multidimensional density, where \( p \) is the dimension of the vector. Then the classification is based on the likelihood ratio. If the inequality is met
\[ \frac{n_p(\gamma, \mu_1, D)}{n_p(\gamma, \mu_2, D)} > 1, \]
then one can refer the vector \( \gamma \) to the first set, otherwise to the second. By transforming the likelihood ratio, we obtain the equation of the separating plane as follows [9,10]. Let the inequality is satisfied
\[ -\frac{1}{2}(\gamma - \mu_1)^T D^{-1} (\gamma - \mu_1) + \frac{1}{2}(\gamma - \mu_2)^T D^{-1} (\gamma - \mu_2) > 0. \]

Opening the brackets one may get
\[ -\frac{1}{2}(\gamma - \mu_1)^T D^{-1} (\gamma - \mu_1) + \frac{1}{2}(\gamma - \mu_2)^T D^{-1} (\gamma - \mu_2) = -\frac{1}{2} \gamma^T D^{-1} \gamma + \frac{1}{2} \mu_1^T D^{-1} \mu_1 + \frac{1}{2} \mu_2^T D^{-1} \mu_2 = \gamma^T D^{-1} \mu_1 - \gamma^T D^{-1} \mu_2 - \frac{1}{2} \mu_1^T D^{-1} \mu_1 + \frac{1}{2} \mu_2^T D^{-1} \mu_2. \]

After some algebra, one may achieve the following inequality
\[ \gamma^T D^{-1} (\mu_1 - \mu_2) - \frac{1}{2} \mu_1^T D^{-1} \mu_1 + \frac{1}{2} \mu_2^T D^{-1} \mu_2 - \frac{1}{2} \mu_1^T D^{-1} \mu_2 = \gamma^T D^{-1} (\mu_1 - \mu_2) - \frac{1}{2} (\mu_1 - \mu_2)^T D^{-1} (\mu_1 + \mu_2) > 0. \]

Then the equation of the separating surface is as follows
\[ \gamma^T D^{-1} (\mu_1 - \mu_2) - \frac{1}{2} (\mu_1 - \mu_2)^T D^{-1} (\mu_1 + \mu_2) = 0. \]

3. Simulation results
Let's construct separating planes for different clusters are shown in figure 5.
Figure 5. Separating planes. The green color shows the points S1, red is S2, blue is S3, gray is S4.

The classifying function is considered in the form of

$$y = a_1 x_1 + a_2 x_2 + a_3 x_3 + c.$$  

where in our case $x_1 = S_i$, $x_2 = F_i$, $x_3 = T_i$. The main goal of the discriminate analysis is to find such a linear combination of discriminant variables that would optimally divide the groups under consideration, and therefore maximize the ratio of inter-group variation to intra-group variation. With the help of the STATISTICA package, the corresponding classification functions $f_1, f_2, f_3, f_4$ were calculated, which are shown in figure 6.
Thus, having calculated the value of the classification function for each vector, we refer it to the set whose function has the most significant value. With the help of this model, one may get the opportunity to determine exactly what process is taking place on the considered signal fragment, which in turn allows us to analyze the technological process.

4. Conclusion

The conducted statistical analysis allows us to make a classification model based on the selected features. Due to averaged parameters $S$, $F$ and $T$, it becomes possible to apply this model to signal fragments of different lengths. The resulting data clusters correspond to different physical processes: there is a separation by the type of milling and by the input and output of the tool into the workpiece in different positions. Despite the simple shape of the trajectory of the control program, the analysis makes it possible to generalize the results obtained to the selection of specified features when processing on more complex trajectories. The expansion of the feature space and the verification of the model on actual industrial signals is the subject of further research.

This work was performed using the equipment of the Centre for Collective Use of the Scientific and Technological Centre of Unique Instrumentation of the Russian Academy of Sciences (STC UI RAS) [http://ckp.ntcup.ru].

References

[1] Hsu C-Y and Liu W-C 2021 Multiple time-series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing J. Intell. Manuf. 32 823–836
[2] Li Y, Mou W, Li J, Liu C, and Gao J 2021 An automatic and accurate method for tool wear inspection using grayscale image probability algorithm based on bayesian inference RCIM 68 102079
[3] Qin A, Guo L, You Z, Gao H, Wu X, and Xiang S 2020 Research on automatic monitoring method of face milling cutter wear based on dynamic image sequence Int. J. Adv. Manuf. Technol. 110 3365–3376
[4] Masalimov K A, Munasypov R A and Kudoyarov R G 2021 Russian Engineering Research 41 3 252–256
[5] Masalimov K A and Munasypov R A 2020 Diagnostic of the state of the cutting tool of metal-cutting machines using bidirectional recurrent neural networks with a long short-term memory Russian Engineering Research 1093–1097
[6] Rozhkov S V, Trushin N N and Shadsky G V 2016 Monitoring of the technical condition of spindle assemblies of metal-cutting machines Technical sciences 8 (2) 188–193
[7] Jiang J-S R and Sun C-T 1995 Neuro-fuzzy modeling and control Proc. of the IEEE 83 3 378–406

[8] Kravchenko V F, Fedorov I V, and Churikov D V 2005 Rvachev’s V.L. R-Functions and the atomic functions in problems of complex-shaped contour objects description and digital image processing EW&ES 10 7 70–80. In Russian.

[9] Anderson T W 1963 An Introduction to Multivariate Statistical Analysis (New Jersey: Wiley)

[10] Ivchenko T I 2019 Mathematical Statistics (Moscow: KD Libkom)