Journal of Machine Construction and Maintenance
QUARTERLY 3/2018(110)
ISSN 1232-9312
p. 7–22

Wojciech BATKO*, Jędrzej BLAUT, Tomasz KORBIEL
AGH University of Science and Technology, Cracow, Poland
* Corresponding author: batko@agh.edu.pl

PREDICTION OF HYDRODYNAMIC STABILITY LOSS IN WATER-LUBRICATED PLAIN BEARINGS BASED ON THE LYAPUNOV EXPONENT

Key words: vibration diagnostics, non-linear methods of analysis of diagnostic signals, stability of hydrodynamic bearing nodes, vibration monitoring.

Abstract: The article proposes a new algorithm for data analysis for diagnostics and the prediction of the loss of hydrodynamic stability in plain bearings associated with the appearance of oil vortexes in the lubricant. Their formation is related to the increase of the vibration energy of the spigot, which affects the change of the monitored movement of the shaft in the bearings. These vibrations are controlled by a measuring instrumentation installed on the bearings. The proposed solution was related to the research on the usefulness of nonlinear analysis of diagnostic signals represented by a time series. These signals are the results of measurements of the displacement of the shaft of the rotor machine in relation to the bearing of the plain bearing. The operation of the proposed method is aimed at recognizing the loss of hydrodynamic bearing stability. In particular, we analysed the usefulness of changes in the value of the Lyapunov exponent of monitored signals to early detection of occurring disturbances in the work of a plain bearing. Observations of the increase of its value, before a significant increase in the value of monitored vibrations, confirmed its usefulness as a valuable diagnostic symptom for forecasting the loss of hydrodynamic stability of hydrodynamic plain bearings.

Analiza przydatności analizy wykładnika Lapunova w prognozowaniu utraty hydrodynamicznej stabilności w łożysku ślizgowym

Słowa kluczowe: diagnostyka drganiowa, nieliniowe metody analizy sygnałów diagnostycznych, stabilność pracy hydrodynamicznych węzłów łożyskowych, monitoring drganiowy.

Streszczenie: W artykule zaproponowano nowy algorytm analizy danych do diagnostyki i prognozy utraty hydrodynamicznej stabilności w łożysku ślizgowym, związanej z pojawianiem się wirów olejowych w czynniku smarnym. Ich powstawanie jest związane ze wzrostem energii drgań czopa, co wpływa na zmianę monitorowanego ruchu wału w panwi łożyska. Drgania te są kontrolowane przez zainstalowaną na łożyskach instrumentalizację pomiarową. Zaproponowane rozwiązanie związano z badaniami przydatności metod nieliniowej analizy sygnałów diagnostycznych reprezentowanych szeregami czasowymi. Sygnały te są wynikami pomiarów przemieszczeń czopa wału maszyny wirnikowej względem panwi łożyska ślizgowego. Działanie zaproponowanej metody ukierunkowane jest na rozpoznawanie utraty hydrodynamicznej stabilności łożyska. W szczególności poddano analizie przydatności zmian wartości wykładnika Lapunowa monitorowanych sygnałów do wczesnego wykrycia pojawiających się zaburzeń w pracy łożyska ślizgowego. Obserwacje wzrostu jego wartości przed istotnym wzrostem wartości monitorowanych drgań potwierdziły jego przydatność jako wartościowego symptomu diagnostycznego do prognozowania utraty hydrodynamicznej stabilności pracy hydrodynamicznych łożysk ślizgowych.

Introduction

In today's operation of machines based on hydrodynamic bearing nodes, diagnostic monitoring systems using advanced analyses and algorithms of monitored diagnostic signals for their work are increasingly used. These tools are used to assess their technical condition and the operating conditions. In the construction of bearings, dynamic evaluation is also used. New constructions, characterized by other lubricants than oil and significant rotations, require an innovative approach to monitoring their condition. Examples are gas

film bearings [1] and bearings lubricated with water [2]. Due to the strong non-linearities in the diagnosis of plain bearings, attention is paid to various solutions, including methods based on the analysis of residual processes [3] or qualitative analysis of shaft displacement signals relative to the pan [4–5].

Implementation of a number of them in operational systems of the continuous supervision of correct operation of hydrodynamic bearing units is, however, limited, due to the high costs of the measurement systems required in them, due to the complexity of application conditions, or the lack of automatic state identification. For example, computational algorithms used, for example, in acoustic cameras used in diagnostic surveillance processes can analyse signals from several hundred microphones in real time, but usually only in limited frequency ranges.

Thus, classic diagnostic solutions, based on the analysis of the properties of monitored diagnostic signals using different algorithms, i.e. fast Fourier transform FFT, SFT, wavelet analysis, correlation or synchronous averaging, are still the basic solutions dedicated to the diagnostics of the state of rotor machines. Recently, solutions derived from nonlinear modelling methods [6] and nonlinear signal analysis [7] also appear in use, including studies on the identification of the presence of deterministic chaos in monitored signals. An example may be the use of entropy to detect early signs of wear in a rolling bearing of a wind power plant [8–9].

In connection with the above, it is justified to conduct further research on the usefulness of non-linear signal analysis methods, in terms of their practical application in various monitoring systems. Therefore, the authors of the work have drawn attention to this path of the recognition of disturbances in the proper functioning of hydrodynamic plain bearings lubricated with water. In particular, they analysed the behaviour of changes in the value of the Lyapunov exponent, analysed in relation to monitored diagnostic signals, monitoring the condition of correct hydrodynamic plain bearings operation. This discriminant describes the same properties of diagnostic signals as the metric entropy with the difference that metric entropy describes the entire phase stream, and the Lyapunov exponent operates on single phase trajectories.

The LLE (Largest Lyapunov exponent) algorithm was used to carry out research experiments dedicated to the diagnosis of floating bearings lubricated with water, which allows studying the behaviour of the Lyapunov exponent in relation to the problem. The research incentives for its use were the results presented in the paper [10]. They demonstrated the high resistance of this solution to the variability of input parameters of the diagnosed object, documented by the results of diagnostic tests of low-speed slewing bearings [11]. In the assessment of the usefulness of this method of diagnostic diagnoses, other references to non-linear analysis of diagnostic signals, including the evaluation of functional properties of such methods, e.g., the fractal dimension analysis or the energy operator Teiser Kaiser, have been used in diagnosis of bearings [2, 12].

1. Monitoring of rotating machines based on hydrodynamic bearing nodes

Diagnostic forecasting processes have an undeniable value for the smooth operation and management of machine maintenance. Both at the macro level, from the point of view of the policy of controlling production processes at the plant, and at the micro level referring to a specific diagnosis object. Correct prediction of controlled changes in the condition of the diagnosed facility over a given time horizon is crucial for making decisions about maintaining machines. They are inseparably connected with operational control processes and constitute the basis for the maintenance practice of technical facilities related to making specific decisions regarding the exclusion of an object from operation, e.g., completion of certain inspections, exchanges, repairs, and overhauls. Helpful solutions in the implementation of such tasks in relation to rotating machines are increasingly monitoring systems that monitor the vibrations of their bearing nodes. One of the typical elements of such solutions that are widely present in industry is a module for controlling the relative vibrations of the shaft and shaft movements in the hydrodynamic bearing of the plain bearing.

Their implementation is determined by the continuous measurement of the displacement of the shaft spigot in the bearing shaft on two mutually perpendicular directions, carried out by eddy current sensors.

The process of monitoring the change in the condition of the bearing node is carried out by statistical analysis of the signal or frequency analysis. In the statistical analysis (Figure 1), the state estimator is based on the measurement of the maximum vibration radius of a spigot on its trajectory or the value of the higher peak-to-peak amplitude of vibration displacements from both controlled signals. This estimator is compared with the criterion threshold values. Frequency analysis is based on estimators related to rotational frequency and its harmonics and subharmonics [13].

Exceeding the permissible levels of selected estimators is connected with the possibility of the loss of the stability of the equilibrium position and the creation of self-excited vibrations of high amplitude, which may cause the breaking of the lubricant carrier layer and the destruction of the bearing. The increasing presence of monitoring systems performing vibration monitoring of bearing units of rotating machines with the use of eddy current sensors creates a wide platform for possible analyses of the correct operation of the bearing node [15] and related management solutions for their operation. Mention should be made of the conditions of the appearance of significant disturbances in the

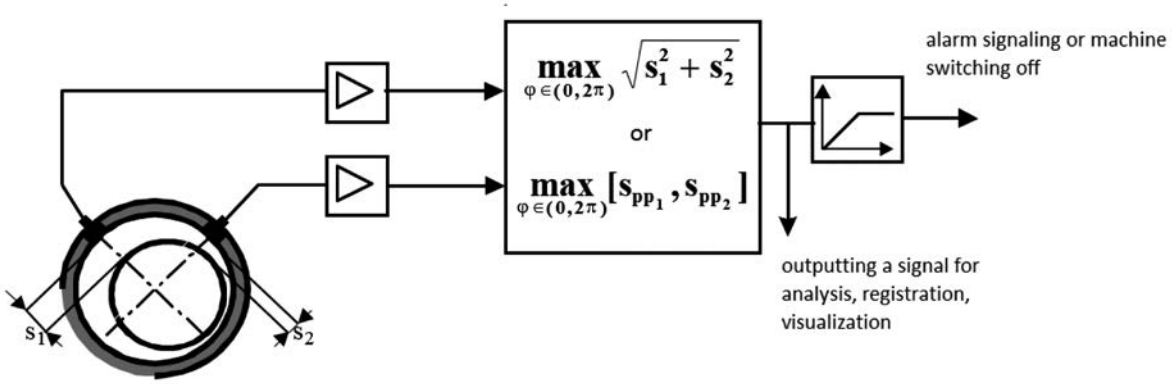

Fig. 1. Vibration monitoring system of sliding bearing nodes [14]

correct functioning of the vibration supervision system of bearing hubs. They are related to the high sensitivity of eddy current sensors to the inhomogeneities of ferromagnetic surfaces of the monitored shaft as well as various types of damage to it, for example, scratches, scratches, deformations, or deposits occurring on their surfaces. The mentioned impacts are referred to as electrical runout and mechanical runout. Their presence leads to significant measurement distortions. To minimize them, solutions are needed to filter out the interference they generate [14].

The above-mentioned method of the operation of rotor machines based on hydrodynamic slide bearings supported by vibration supervision was the inspiration to search for new methods of diagnosing malfunctions, which may lead to the abrasion of the slide bearings and, consequently, to machine downtime.

2. Detection of hydrodynamic stability in the slide bearing using the analysis of changes in the value of the Lyapunov exponent

The vibration monitoring system of the rotor machine, which is synthetically described in the previous item, is one of the basic elements of the selection of the conditions of its operation process and the level of risk during its use. The requirement of safe operation generates the need for early identification of the occurrence of instability in the lubricant operation and related changes in vibration signals that determine the monitored geometric trajectories of the movement of the shaft spigot in the slide bearing.

To solve this problem, the usefulness of nonlinear signal analysis methods was considered in relation to geometric trajectories monitored by the vibration monitoring system of the machine. In particular, the subject of the research was the analysis of the usefulness of changes in the value of the Lyapunov exponent.

The occurrence of an emergency condition, manifested by significant changes in the geometric trajectory describing the location of the centre of the shaft journal relative to the pan is a consequence of the growing instability of the lubricating wedge. Therefore, the analysis of the stability of this wedge can be a tool allowing one to predict the bearing condition in a short time horizon. The Lyapunov exponent can be used as a tool to study the rate of infinitesimal separation of individual geometric trajectories.

This exponent describes the evolution of the state x_0 with the function $f(x_0)$. Iterative function $f^i(x_0)$ allows determining the value of x_i after the time corresponding to i-iteration, i.e.: $x_i = f^i(x_0)$.

The Lyapunov exponent is defined as the following boundary:

$$\left(x_0\right) = \lim_{i\to\infty}\left(\frac{1}{i}\right)\ln\left|\frac{df^i\left(x_0\right)}{dx_0}\right| \tag{1}$$

This signal discriminator allows one to examine relations between neighbouring phase trajectories. For systems exhibiting chaotic behaviour, distances between consecutive trajectories increase exponentially.

As an inverse function, the exponent form of (1) can be written as follows:

$$e^{i\lambda(x_0)} = \left|f^i\left(x_0+\varepsilon\right) - f^i\left(x_0\right)\right| \tag{2}$$

Which, in the boundary $\varepsilon \to 0$, a $i \to \infty$ gives the exact expression for $\lambda(x_0)$:

$$\lambda\left(x_0\right) = \lim_{i\to 0}\lim_{\varepsilon\to 0}\left(\frac{1}{i}\right)\ln\left|\frac{f^i\left(\mathrm{x}_0+\varepsilon\right) - f^i\left(\mathrm{x}_0\right)}{\varepsilon}\right|$$

$$\lambda\left(x_0\right) = \lim_{i\to\infty}\left(\frac{1}{i}\right)\ln\left|\frac{df^i\left(x_0\right)}{dx}\right| \tag{3}$$

Thus, the distance of two closely related points under the influence of one reproduction iteration increases by an average of $e^{\lambda(x_0)}$ (Fig. 2).

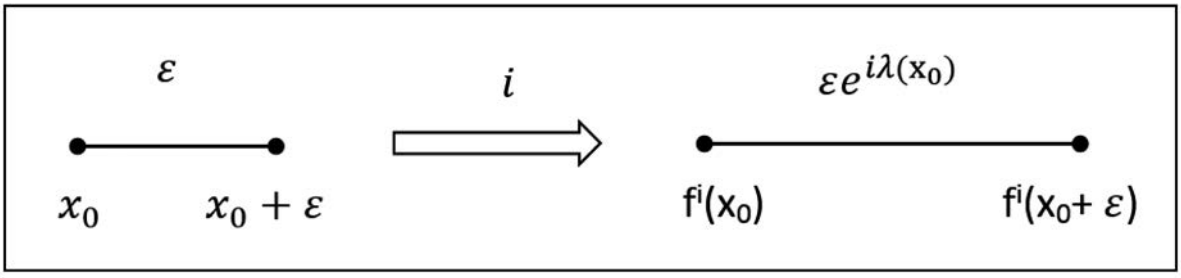

Fig. 2. Graphical interpretation of the Lyapunov exponent

Let us assume the theoretical evolution of the system with trajectories within the sphere (in the n-dimensional phase space), which tends to deform the phase space into the shape of an ellipsoid. Axes of ellipsoids change their length. The axis in which the greatest elongation occurs determines the direction of the highest instability. The Lyapunov exponent is described as follows:

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} ln\left[\frac{r_i(t)}{r_i(0)}\right] \tag{4}$$

where
$r_i(0)$ for i = 1, 2, 3,…, n is the radius of the initial trajectory,
$r_i(t)$ for i = 1, 2, 3,…, n is the i-th major axis of the ellipsoid at time.

Each layout dimension is matched by one Lyapunov exponent. In practice, only the largest Lyapunov exponent is considered: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_n$. The algorithm for estimating the Lyapunov exponent was presented in the paper [10]. The algorithm allows one to determine the largest exponential of Lyapunov. The procedure first reconstructs the data using a delayed method, so that each successive value of the time series x_i is saved as a vector

$$X_i = \left[x_i, x_{i+T}, x_{i+2T} \ldots, x_{i+nT}\right] \tag{5}$$

where T is the delay between consecutive data points of the time series.

Then, for the vector X_i the distance from the nearest neighbour vector X_j is sought. The distance of the trajectory with subsequent X_{i+nT} from X_{j+nT} in a chaotic system increases according to the formula $D_{i(k)} = c^{\lambda k}$, where λ is the approximation of the largest Lyapunov exponent. This dependence occurs because the increase of the radius along the axis associated with the largest exponent will quickly dominate the increase or decrease along the remaining axes.

To calculate the value of the approximate largest Lyapunov λ exponent, the following algorithm calculates the logarithm of the distance between trajectories:

$$log\left(d_{i(k)}\right) = \log(c) + \lambda \cdot k \tag{6}$$

A set of distances is obtained in this way (one for each i-th element) whose slope is the approximation of the largest Lyapunov exponent λ. W The method used determines the average logarithms of the trajectory distance d'(k) by the log mean value $\log(d_{ik})$ of all orbits of X_i vectors. Finally, the slope of the line in the diagram d'(k) from k makes it possible to determine the largest Lyapunov exponent λ.

3. An example of the usefulness analysis of the indicated method

The presented method first requires the separation procedure of individual geometric trajectories. Based on the FFT analysis, the basic rotational frequency is identified. Then, the appropriately prepared signals from the measurement system are divided in such a way as to create independent individual trajectories. Reconstruction of the trajectory function will allow determining the distance between particular points of subsequent trajectories. Therefore, calculated distances between consecutive trajectories make it impossible to find data points close to each other both on the phase space and in time. This allows the analysis of dynamic changes in the properties of the system, not time dependencies.

The number of samples per shaft rotation relative to the bearing shaft was determined by the possibility of an analogue digital card transducer used in the experiment. According to Rosensetin's recommendations, the algorithm should include sampling so that the autocorrelation function drops below $1-\frac{1}{e}$ of the maximum value.

The implementation of the algorithm for calculating the Lyapunov exponent includes the following steps [10]:

- Estimating individual trajectories relative to shaft revolutions based on FFT analysis,
- The reconstruction of the geometric trajectory function,
- The determination of the axis of the highest value of trajectory separation in individual revolutions,

- The calculation of the average value of separation, and
- The interpolation of the Lyapunov exponent function by the least squares method.

The implementation of this algorithm was carried out in a Labview programming environment, and its usefulness in the analysed diagnostic task was tested on the results of the experiment carried out on a constructed laboratory stand, allowing us to analyse the work of a slide bearing lubricated with water, shown in Fig. 3.

The test stand is equipped with an electric motor powered by a frequency converter that allows the speed to be regulated. The motor drives the shaft bearing (3) supports through the clutch. The two extreme supports are equipped with rolling bearings, while the middle support is the tested slide bearing. The bearing is lubricated with water and has a set of two magnetic relative displacement sensors installed. The test bearing lubricated with water is made of a rubber liner with longitudinal grooves, and this lining is embedded in a brass frame. The lubricant water is supplied and returned to the bearing from the reservoir in a system with a closed lubricant circuit. The slide bearing lubrication system allows the viscosity of the lubricant to be modified by changing the protein concentration in the water. The measuring system consists of two eddy current sensors and a phase marker (tachometer). The signal from the tachometer and eddy current sensors is recorded using a measuring card with a sampling rate of 4,200 Hz and a 16-bit analogue-to-digital converter. The proprietary software allows one to record measurement data (tachometer signal and x and y displacement), remove the runout phenomenon, and plot the rotation trajectories of the shaft axis. Signal analyses were carried out after the experiment in the software developed for this circumstance using the Labview programming environment.

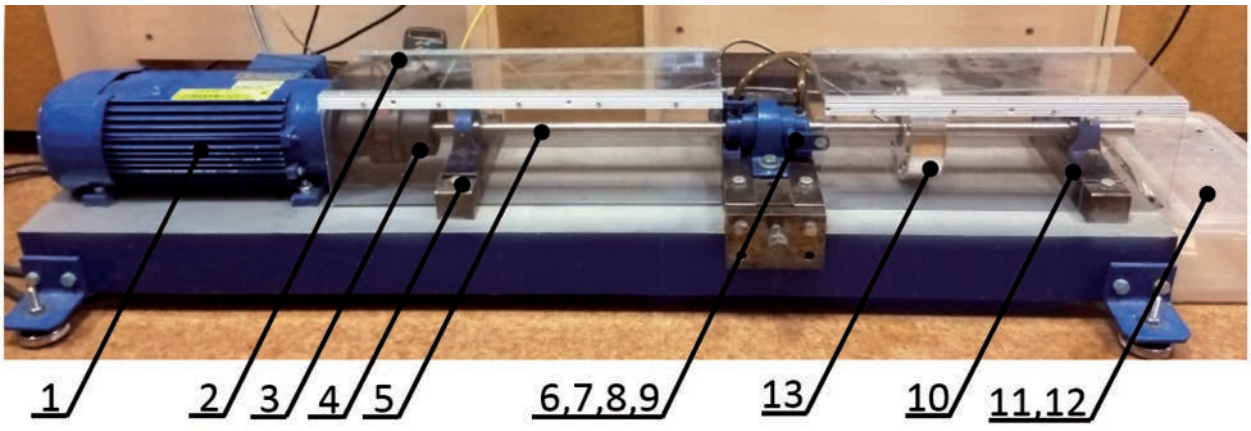

Fig. 3. Test stand, 1 – motor with speed control system, 2 – tachometer, 3 – coupler, 4, 10 – rolling bearing, 5 – shaft, 6 – slide bearing, 7 – bearing loading system, 8 – lubricant delivery and discharge valves, 9 – eddy current sensors, 12 – lubricant tank, 13 – rotor disc used for balancing

The planned active experiment consisted in changing the dynamic parameters of the lubricant during operation by diluting it with a 15% albumin solution. This caused the viscosity of the lubricant to change. During the experiment, there was a change in the noise level during the dilution of the lubricant.

Analysis of the signal during the test allowed observing an increase in the vibration amplitude during the dilution of the lubricant. The bearing operating parameters, such as rotational speed or bearing load, were constant during the measurement. The shaft speed was 2,200 RPM, and the load was caused by an unbalancing disk. The main resonance frequency of the shaft is 31 Hz. The signal has been filtered by a medium pass filter with limit frequencies of 10 and 120 Hz.

Figure 4 presents the displacement amplitude diagram for the x and y signals calculated according to Fig. 1. The marks in Fig. 4 (62.5 [s] and 63 [s]) indicate the time in which the rotational trajectories shown in Fig. 5 were determined.

Figure 5 shows the evolution of the rotational trajectory, the rotational trajectory is associated with the x and y signal. It represents the movement of the centre of the shaft axis with respect to the bearing pan. Phase markers are set every 360 degrees, which means that we observe the path circled by the centre of the shaft with respect to the pan during two full revolutions of the shaft. The first trajectory overlaps during two turns, and we observe stable operation. The second trajectory does not overlap anymore, and one can see the difference between the first and second considered turnover.

Figure 6 shows the values of the Lyapunov exponent (multiplied by 100 times) during gradual dilution of the aqueous albumin solution. The

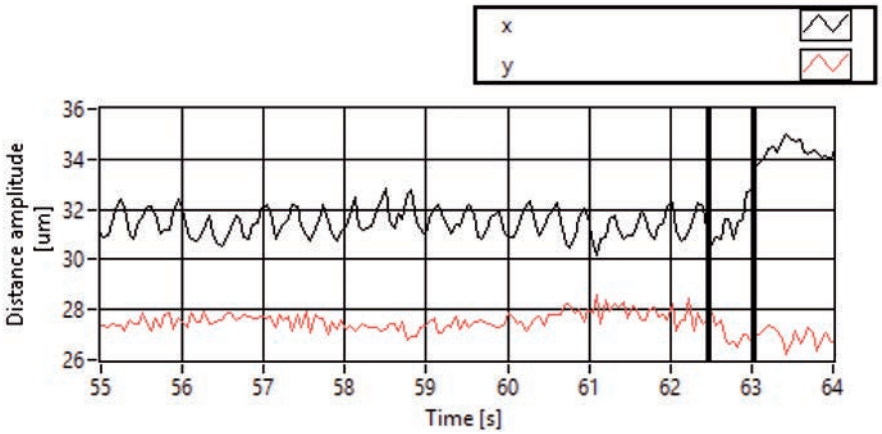

Fig. 4. Diagram of displacement amplitude for x and y axes for rotational speed: 2200RPM, x – black, y – red graph

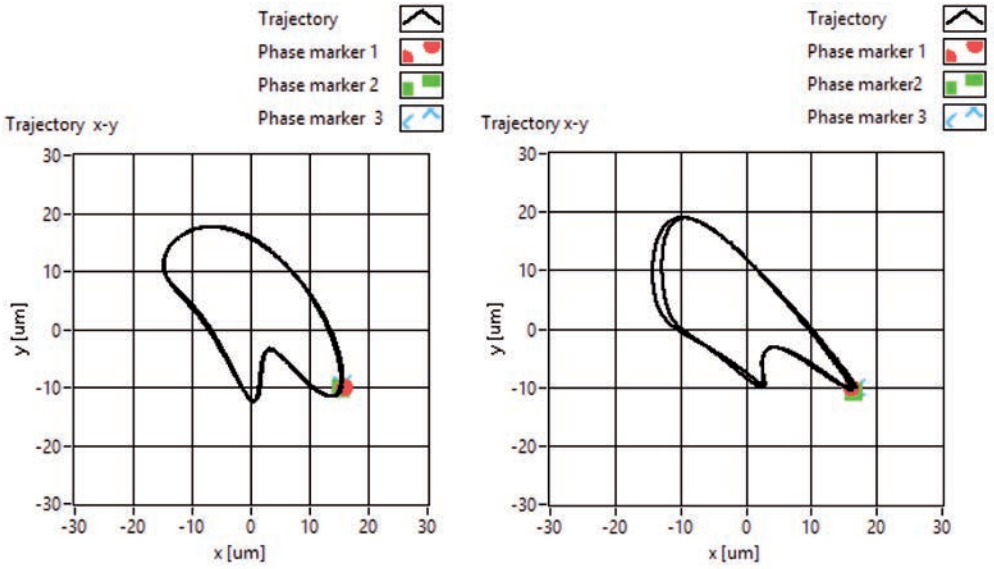

Fig. 5. Evolution of the trajectory, the first trajectory is determined in place of the first marker in Fig. 4, the second corresponds to the second mark. Trajectories are drawn for two rotations

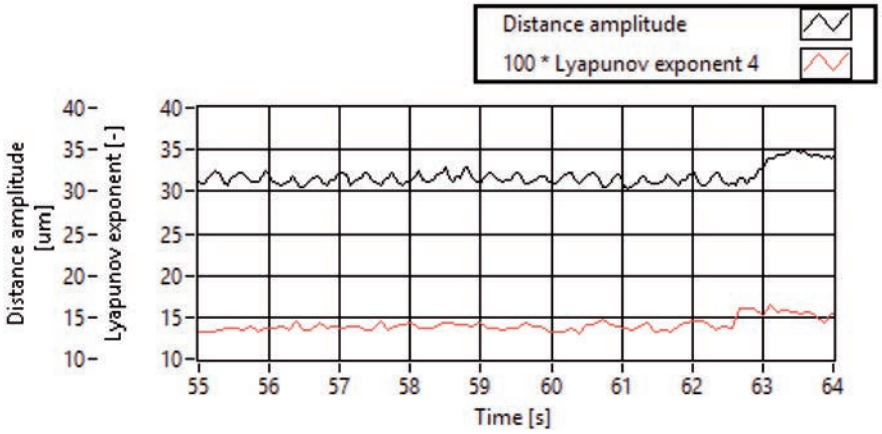

Fig. 6. The value of the displacement amplitude in the x-axis and the Lyapunov exponent

displacement amplitude is the value of the maximum amplitude of the displacement signal determined every 4 revolutions. The displacement amplitude is determined every 4 revolutions in order to compare its value with the value of the Lyapunov exponent. The Lyapunov exponent is determined by the LLE algorithm for a time series describing 4 revolutions, due to the possibility of determining the largest Lyapunov exponent for a series describing 4 adjacent rotary trajectories. The Lyapunov exponent, as a measure of the distance of points lying on the phase trajectory, allows one to analyse the "divergence" of the phase trajectory, which means that the signal should consist of at least two cycles.

It has been observed that a longer observation time smoothes out the graph of the value of the Lyapunov exponent facilitating the analysis, which caused the determination of values for every 4 revolutions. Extending the signal analysed in the adopted procedure reduces the time to make a decision. It means that, for the determination of the first value of the exponent, the Lyapunov signal must describe 4 revolutions. It is also possible to determine LLE for each subsequent collected sample by collecting a signal describing the first 4 turns and then adding each subsequent sample along with the removal of the "oldest" sample.

The place marked with the second marker indicates the moment of increase in the displacement amplitude in the x axis. This event was considered potentially dangerous, because sudden impulse increases in the displacement amplitude may lead to seizing the slide bearing. After this deflection, the vibrations in the x-axis are gradually decreased. The maximum value of the Lyapunov exponent occurs just before the amplitude increases. The observation made suggests that a sudden increase in the value of the Lyapunov exponent may predict the loss of the lubrication wedge and the consequent seizing of the bearing.

Summary

The implementation of the conducted research works and the analysis of its results allowed us to recognize the issue of the applicability in diagnostic practice of a new method of identifying disturbances in the correct functioning of hydrodynamic plain bearings. To this end, one of the indicated and analysed methods of non-linear analysis of diagnostic signals, i.e. the method of analysing the variability of the Lyapunov coefficient value determining the variability of monitored diagnostic signals, was used.

The results of the conducted experiments have shown that, when using the analysis of changes in the value of the Lyapunov exponent in the vibration signals of the shaft plug relative to the hydrodynamic bearing junction lubricated with water, the occurrence of lubricating wedge instability can be detected at an early stage. The examined changes in the value of the Lyapunov exponent, compared with the observed increases in the values of the monitored displacement of the shaft spigot in the bearing, made it possible to identify the occurrence of chaotic disturbances in the lubrication wedge.

This allows considering a wider application of the indicated method for early identification of the wedge instability phenomenon, i.e. disturbances in the correct functioning of the bearing node. The above conclusion was documented by the results of an active diagnostic experiment (on the test bench of a slide bearing lubricated with water) on which the properties of the lubricating wedge were changed by changing the concentration of the protein solution in water, which determines the change in viscosity of the lubricant. They showed the ability to react quickly to the value of the Lyapunov coefficient for these changes.

The experiments carried out have shown that the subtle changes in the properties of the lubrication wedge are accompanied by sudden increases in the monitored displacement amplitudes and their slow suppression. The observed disorder is preceded by a change in geometric trajectory, defining disturbances in the phase trajectory. The associated changes in the value of the Lyapunov's coefficient of monitored signals allow for a quick recognition of the possibility of loss of hydrodynamic stability of the bearing unit.

The current research stage confirmed the usefulness of the proposed algorithm to identify disturbances in the correct functioning of the bearing, resulting from changes in the properties of the lubricating wedge. This can successfully predict the broader possibilities of using the described method in the operation of sliding bearings, where temporary disturbances in the proper functioning of the bearing occur, resulting from changes in the properties of the lubricating wedge.

Further confirmation of the concepts presented in the work will be a premise for the construction of rapid displacement detectors in the correct functioning of hydrodynamic slide bearings based on the observation of changes in the Lyapunov coefficient value and their wider implementation into solutions of commonly used systems monitoring the state of rotor machines based on hydrodynamic bearing nodes.

References

1. Bagiński P., Żywica G.: Operational Problems Of High-Speed Foil Bearings Tested Under Laboratory Conditions. Journal of Machine Construction and Maintenance. Problemy Eksploatacji, 2017, 105(2), pp. 19–24.
2. Blaut J., Korbiel T., Batko W.: Application of the Teager-Kaiser energy operator to detect instability of a plain bearing. Diagnostyka, 2016, 17(4), pp. 99–105.

3. Fiebig W., Wróbel J.: Identyfikacja źródeł hałasu na przykładzie agregatu hydraulicznego. Inżynieria Maszyn, 2014, 19(2), pp. 92–97 (in Polish).
4. Bielawski P.: Diagnostics of marine propeller shafts. Journal of Polish CIMAC, 2011, 6(2), pp. 31–40.
5. Kiciński J.: Dynamika wirników i łożysk ślizgowych. Wydawnictwo IMP PAN, 2005 (in Polish).
6. Peruń G., Łazarz B.: Modelowanie uszkodzeń łożysk tocznych przekładni zębatych stanowiska mocy krążącej. Zeszyty Naukowe Politechniki Śląskiej, Seria: TRANSPORT, 2008, 64, pp. 201––208 (in Polish).
7. Pennacchi P., Vania A., Chatterton S.: Nonlinear effects caused by coupling misalignment in rotors equipped with journal bearings. Mechanical Systems and Signal Processing, 2012, 30, pp. 306–322.
8. Lu X., Zhang J., Ma L., Lin J., Wang J., Dai H.: Effects of misalignment on the nonlinear dynamics of a two-shaft rotor-bearing-gear coupling system with rub-impact fault. Journal of Vibroengineering, 2017, 19(8), pp. 5960–5977.
9. Qing N., Feng K., Kesheng W., Binyuan Y., Wang Y.: A case study of sample entropy analysis to the fault detection of bearing in wind turbine. Case Studies in Engineering Failure Analysis, 2017, 9, pp. 99–111.
10. Rosenstein M., Collins J., De Luca C.: A practical method for calculating largest Lyapunov exponents from small data sets. Physica D: Nonlinear Phenomena, 1993, 65(1–2), pp. 117–134.
11. Caesarendra W., Kosasih B., Tieu A., Moodie C.: Application of the largest Lyapunov exponent algorithm for feature extraction in low speed slew bearing condition monitoring. Mechanical Systems and Signal Processing, 2015, 50–51, pp. 116–138.
12. Janjarasjitta S., Ocak H., Loparo K.: Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal. Journal of Sound and Vibration, 2008, 317(1–2), pp. 112–126.
13. Miąskowski W., Kiciński J.: Modelowanie uszkodzeń powierzchni panwi łożyskowych. Biuletyn WAT, 2007, 56(1), pp. 232–244 (in Polish).
14. Batko W., Banek, T.: Estymacja zaburzeń w systemach monitorujących. Kraków: Wydawnictwo AGO, 1997 (in Polish).
15. Bilosova A., Bilos J.: Vibration Diagnostics. Ostrava: ESF, 2012.
16. Blaut J., Korbiel T., Uliński A.: Analiza parametrów smarnych łożyska hydrodynamicznego smarowanego wodą o wybranych parametrach technologicznych. Przegląd Mechaniczny, 2015, pp. 6–39 (in Polish).
17. Korbiel T., Blaut J.: Ocena sprawności przekładni mechanicznej w oparciu o ocenę dyssypacji energii. Diagnostyka, 2014, 15(4), pp. 15–20 (in Polish).
18. Kosiński R.A.: Sztuczne sieci neuronowe, dynamika nieliniowa i chaos. Warszawa: WNT, 2014 (in Polish).
19. Morel J.: Drgania maszyn i diagnostyka ich stanu technicznego. Polskie Towarzystwo Diagnostyki Technicznej, 1995 (in Polish).

Journal of Machine Construction and Maintenance
QUARTERLY 3/2018(110)
ISSN 1232-9312
p. 15–22

Antonin VITECEK[a], Andrzej SIOMA[b,*], Piotr SULIGA[b], Janusz KOWAL[b]
[a] Faculty of Mechanical Engineering, VSB Technical University of Ostrava, Czech Republic
[b] Faculty of Mechanical Engineering and Robotics, AGH University of Science and Technology, Cracow, Poland
[*] Corresponding author: sioma@agh.edu.pl

AUTOMATIZATION OF SCREW THREAD PROFILE MEASUREMENT USING A 3D VISION SYSTEM

Key words: 3D image, 3D image analysis, 3D vision system, thread, thread parameters.

Abstract: The purpose of this study is to perform automatic measurement of a screw thread profile acquired with a 3D vision system. Measurements are carried out using a laser triangulation method. The triangulation method requires a calibration procedure to be performed on a measurement stand. The first stage of the calibration procedure allows one to compensate for lens and perspective distortion. Calculated projection model coefficients are also used to transform the image coordinate system into the world coordinate system. Next, a specially designed calibration target with known geometric dimensions is used to position the object's rotation axis on the measurement plane and prepare it for the measurement procedure. This paper presents the algorithms used to perform a measurement for a single cross-section of a screw thread. For the chosen screw thread, the pitch, flank angles, and the radiuses at the crest and root are measured.

Automatyzacja pomiaru wartości profile wkrętaków z wykorzystaniem system wizyjnego 3D

Słowa kluczowe: obraz 3D, analiza obrazu 3D, systemy wizyjne 3D, gwint, parametry gwintu.

Streszczenie: Celem przedstawionych w artykule badań jest przeprowadzenie zautomatyzowanego pomiaru elementu gwintowanego. Pomiar profilu gwintu wykonano przy pomocy systemu wizyjnego 3D opartego na zasadzie triangulacji laserowej. W artykule przedstawiono opis kalibracji stanowiska pomiarowego. Pierwszy etap procedury kalibracyjnej pozwala na kompensację dystorsji układu optycznego oraz zniekształcenia perspektywicznego. Wyznaczone na tym etapie współczynniki macierzy projekcji wykorzystywane są również do transformacji układu współrzędnych obrazu do układu współrzędnych świata. Na kolejnym etapie specjalnie przygotowany cel kalibracyjny, o znanych parametrach geometrycznych, wykorzystany został do wypozycjonowania osi obrotu elementu na płaszczyźnie pomiarowej – wyznaczonej przez wiązkę laserową. W artykule przedstawiono algorytmy wykorzystane podczas wykonywania pomiarów na obrazie pojedynczego przekroju poprzecznego elementu gwintowanego. Przedstawione metody pozwalają na wyznaczenie skoku, kąta zarysu oraz promieni wierzchołka i dna zarysu gwintu.

Introduction

One of the characteristic features of modern manufacturing processes is sustainable quality. The benefits of incorporating quality control into the production process of a screw thread – the most widely used mechanical element – are numerous. Due to its complex shape, narrow geometrical tolerances, and large-scale production, quality control is not a straight forward process. Over the years, this task has been the focus of much research. Finding a method that is able to guarantee both sufficient resolution and repeatability while being fast enough to control every manufactured element has been the goal of numerous workers in the scientific community. Conventional methods of controlling a screw thread are performed using mechanical gauges [1], which can quickly differentiate whether the screw is satisfactory or unsatisfactory. Unfortunately, gauges are subject to wear and tear, and in most cases, it is difficult to determine the source of a problem. To get more in-depth, we can use a coordinate measuring machine (CMM), equipped with a needle-like probe [2]. The information provided allows for a detailed measurement of the thread parameters.

Unfortunately, these measurements are time-consuming, and due to the size of the probe, it is not always possible to detect surface level damage. Recently, attention has been given to a technique utilizing CT scans to acquire a three-dimensional cloud of points which represent the thread [3]. Despite high resolutions and the possibility of measuring internal threads, this method has some serious disadvantages, such as the high cost of equipment and long execution times, making it suited to a laboratory rather than on the production line. Another interesting alternative is the use of optical methods [4]. They have all the advantages of non-contact methods, while maintaining high speed and resolution, and they can be based on a 2D or 3D image. In the case of a 3D image, local surface height variations are given directly and every point on the image is assigned with a height value rather than an intensity value. There are a couple of possible approaches on creating a 3D image including stereovision, photometric stereo, shape from focus, shape from texture, and laser triangulation [8]. The last of the listed methods fulfils all the conditions to become an ultimate solution for thread measurement problems. The first part of this paper concentrates on performing the calibration of a measurement stand in order to acquire a thread surface 3D image. In optical methods, acquiring a high quality image is only the first step in the pursuit of a good final result. Properly chosen image analysis techniques are equally significant. According to the National Physics Laboratory, there are seven thread parameters which require control: pitch diameter, major and minor diameter, pitch, flank angles, radius at crest, and radius at root [1]. The second part of this paper includes a detailed explanation of an algorithm for performing measurements of the last four parameters.

1. Performing a calibration of a measurement stand

The calibration procedure is an essential part of preparing every measurement. When it refers to machine vision, it usually consists of three major tasks: The first is to compensate for a lens distortion; the second is to eliminate the perspective, and the third is to scale the image coordinate system into the world coordinate system. If mistakes are made on any of the calibration stages, it leads to erroneous values at the end of the analysis. In this paper, the acquisition of a thread 3D image was performed using a Sick Ranger 3D camera. The calibration procedure of this device involves using a calibration target with one flat side and one sawtooth-shaped side and moving it across the entire camera field of view on the laser plane. Projecting a laser beam on the flat side of the target allows compensation for possible lens distortion. Eliminating perspective and the scaling of the coordinate system is made possible by using the sawtooth-shaped side.

Acquiring a single image from the camera allows the calculation of the height of a single object's cross-section. The element has to be moved with reference to the camera-laser arrangement in order to perform a full surface scanning. There are two possible approaches for creating a mutual motion using either transitional or rotational element motion. Since getting an image of the entire thread surface is only possible with the second method, it was chosen for this paper. Unfortunately, it has its own disadvantages to be dealt with when building a laser triangulation stand. The threaded element has to be rotated around a fixed axis, positioned precisely on the plane defined by the laser beam or camera axis (depending on the camera-laser arrangement). Two examples of erroneous axis alignments are presented in Fig. 1.

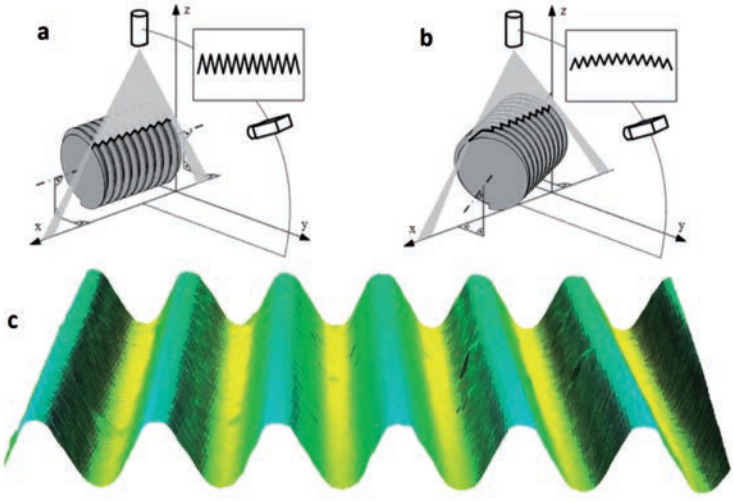

Fig. 1. Measuring stand and picture: a – correct position configuration, b – wrong configuration, c – correct surface image of the thread

In Fig. 1a, an entire thread axis is shifted in one direction with reference to the laser plane. An object profile, visible on the camera image, will be elongated. Performing a correct thread parameter measurement is not possible and will lead to incorrect results. Fig. 1b shows what happens to the camera image when the thread axis is not parallel to the laser beam. The entire object profile is curved and deformed, making it impossible to correctly assess any of the parameters.

One possible solution for this situation is to perform an additional calibration procedure involving a cylinder-shaped calibration target with known geometric dimensions. It should be manufactured with sufficient precision (depending on the task) and then measured with a CMM. The resolution of such a measurement should be at least an order of magnitude higher than the resolution required by a 3D stand. The calibration target has to have two different diameter values (d_1, d_2) in order to distinguish a bad axis alignment from a variable distance between the laser and the object. In order to perform a calibration, the cylinder-shaped target has to be mounted in lathe chucks or with another positioning mechanism, which will be used later to set a screw thread in rotational motion. Fig. 2a presents the process of the calibration of the measurement stand.

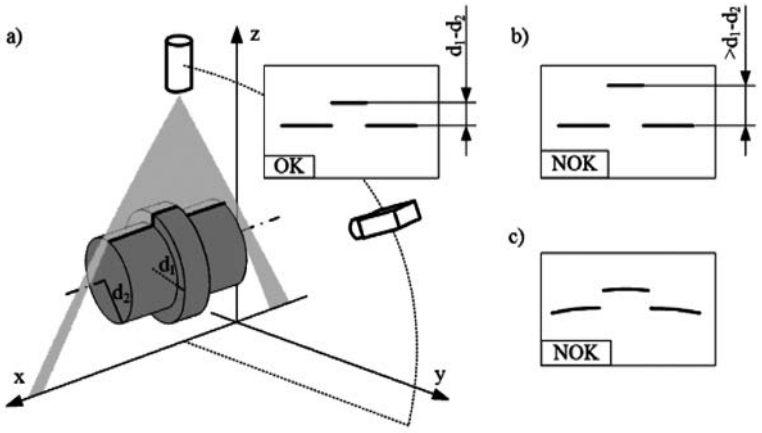

Fig. 2. Performing calibration with cylinder-shaped target

The calibration target has been designed to supply information about the relative position of the rotation axis and the measurement plane. If the calibration target is incorrectly positioned in relation to the vision system, it will be represented on its 3D profile. If the source of an error is analogous to one presented in Fig. 1a, then the height difference between the two parallel lines on the camera image will not match the real values (measured with a CMM) Fig. 2b. An axis rotated in relation to the laser beam plane (as in Fig. 1b) will result in curved lines representing the profile (Fig. 2c). There are a couple of possible approaches to detect these flaws on a mathematical level. One possible solution is to use linear least-squares fitting. This method will be presented in detail in this paper. Our method can be split into three steps. First, the image from the camera has to be processed using line-find algorithms in order to obtain a set of points which represent the profile. This is an extensive topic and will not be covered here. We used a centre of gravity algorithm supplied by the camera software producer. The second step is to segment obtained points into two subsets – one representing the surface of a bigger cylinder d_1 and one representing the surface of a smaller cylinder d_2. This was presented in Fig. 3a, where circles represent points of the first data subset and crosses represent points of the second data subset. When this is done, it is possible to perform a least-squares line fitting for both subsets. Vertical linear fit proceeds by finding the sum of the squares of the vertical deviations R^2 of a set of n data points. The minimized functions take the following forms:

$$R_1^2(a_1, b_1) = \sum [y_{i1} - (a_1 + b_1 x_{i1})] \quad (1)$$

$$R_2^2(a_2, b_2) = \sum [y_{j2} - (a_2 + b_1 x_{i1})] \quad (2)$$

where
$R_1^2(a_1, b_1)$ – is a function minimized for a subset of points (x_{i1}, y_{i1}) representing a greater cylinder;
$R_2^2(a_2, b_2)$ – is a function minimized for a subset of points (x_{j2}, y_{j2}) representing a smaller cylinder.

When the minimization is performed, we obtain a linear equation for the greater cylinder surface ($y_{i1} = a_1 x_{i1} + b_1$) and a linear equation for the smaller cylinder surface ($y_{j2} = a_2 x_{j2} + b_2$) Fig. 3a.

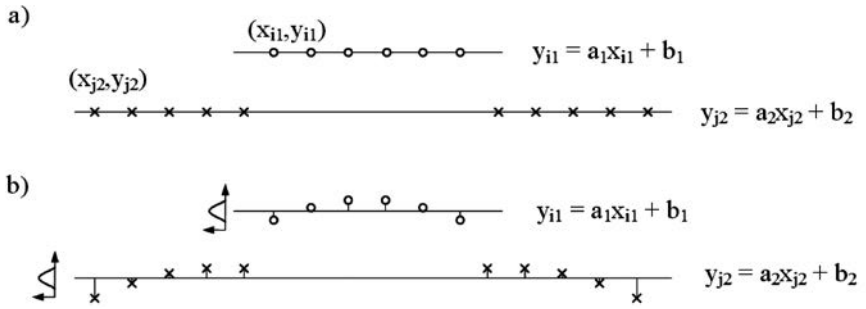

Fig. 3. Detecting an axis misalignment using the cylinder-shaped calibration target image

If the calibration target is perfectly aligned, then "a" parameters for both cylinders should be equal to zero:

$$a_1 = 0, \quad a_2 = 0 \tag{3}$$

The difference between "b" factors should be equal to half the cylinder diameter difference:

$$b_1 - b_2 = ½ (d_1 - d_2) \tag{4}$$

The last condition involves assessing the normal deviation of the vertical distance between subset points and the line calculated using this subset (Fig. 3b). This was done using the following equations:

$$\sigma_1 = \sqrt{\frac{\sum\left(y_{i1} - \overline{y_1}\right)^2}{n-1}} \tag{5}$$

$$\sigma_2 = \sqrt{\frac{\sum\left(y_{i2} - \overline{y_2}\right)^2}{m-1}} \tag{6}$$

In an ideal situation, both normal deviation values should be equal to zero. In the real world, all measurements are subject to noise and thus values will differ from zero. The difficulty is to distinguish errors led by noise from those being the result of an axis misalignment. Assessment of impact made by noise δ_σ was done by repeatedly scanning the same profile of the calibration target surface and calculating a normal deviation value for every point along the profile. This has to be done again for every specific surface, camera-laser arrangement, optical setup, and laser power configuration, since the values can have a significant impact on the results obtained in every specific case. The mean value of calculated normal deviations serves as a threshold for detecting axis misalignment. If one of the values σ_1 σ_2 is greater than δ_σ, axis alignment should be corrected.

$$\sigma_1 \leq \delta_\sigma, \ \sigma_2 \leq \delta_\sigma \tag{7}$$

When all the conditions given by Eq. 3, Eq. 4, and Eq. 7 are met, the element rotation axis is well positioned with respect to camera-laser arrangement. Now it is possible to mount the threaded element in the same lathe chucks or other positioning mechanism and perform a measurement.

2. Measurement and parameter assessment of a single profile

There are several thread forms in use. Each one of them is characterized by a certain set of parameters. In this paper, we present a method of calculating thread pitch, flank angles, radius at crest, and radius at root for ISO metric screw thread. Different kinds of thread have some common features that make it possible to adapt the presented method to other types of screw threads as well.

Measuring thread parameters involves finding some of its characteristic features. The method presented in this paper can be split into four main stages. The first step requires calculating the positions of the thread roots and crests. There are two possible solutions that may be used depending on a thread type.

If the thread in question has rounded roots or crests (Whitworth thread, knuckle thread), they can be efficiently found using first derivative of the profiles. Since the profile is defined by discrete points rather than a continuous function, the central difference approximation of the first derivative can be used.

$$f'(x_i) = f(x_{i-1}) - f(x_{i+1}) / 2h \tag{8}$$

Of course, performing a numerical differentiation in such a way amplifies high-frequency noises. Using a median filter before performing differentiation should help to cut out noise without affecting data quality. If the noise value is high compared to a useful signal, we recommend reducing the noise by controlling the environment. Possible causes of high noise values are vibrations of an object or measurement stand, external light sources, too wide laser lines being visible on the

camera sensor, or specular reflections on the object. If introducing any of these improvements do not lead to sufficient data quality, we recommend taking a look at some more advanced methods of performing numerical differentiation as presented in work of Chartrand [5]. The roots and crest of a thread can now be easily found by assessing the intersection point between the profile derivative and the coordinate system x-axis (Fig. 4).

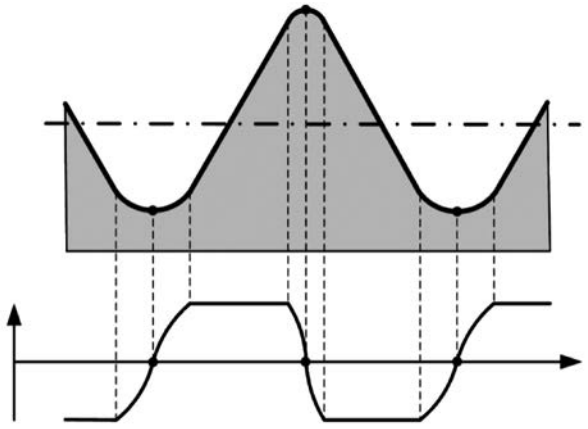

Fig. 4. A thread profile with rounded roots and crest and its derivative

If the crests or roots of a thread are flat, (square thread, Buttress thread), a different approach has to be applied. This involves calculating a centre of gravity of two areas defined by a thread profile and a straight line, and the position of which is calculated using a mean difference value between a crest's heights and a root's heights (Fig. 5). The line position is dependent on a thread type, but it has minor importance. For an ISO screw thread, we can assume that the line is located at $2/5H_N$ from the thread root.

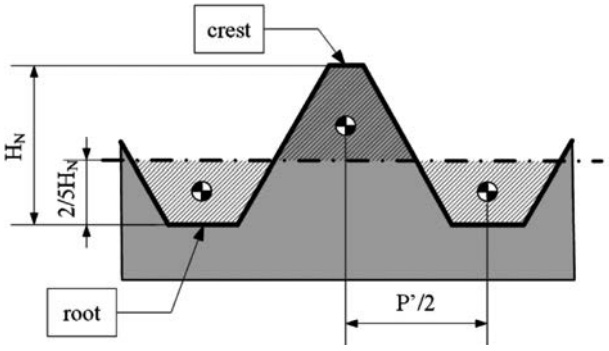

Fig. 5. A thread profile with rounded crests and roots and calculated geometrical centres of selected areas

Regardless of which method was applied in a specific case, a set of coordinates is obtained that represent the positions of roots and crests along a profile (Fig. 6).

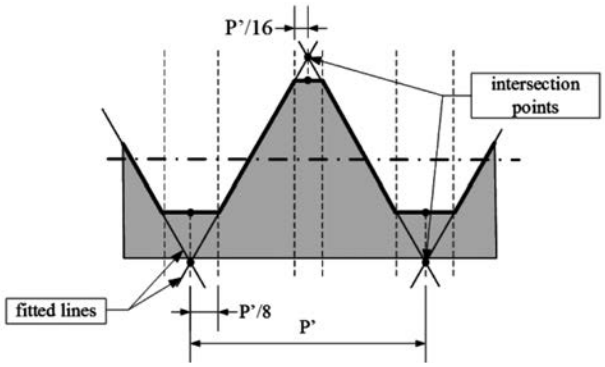

Fig. 6. Algorithm result of line fitting and finding intersection points

Based on the distance between the consecutive roots or crest, we can calculate an approximate value of thread pitch P'. It will be necessary to establish the start and end points of each tooth flank. For an ISO metric thread, each flank should start at approximately P'/16 from the profile crest and end at P'/8 from the root point. The next step involves fitting straight lines in areas between each starting and ending point as presented in Fig. 7.

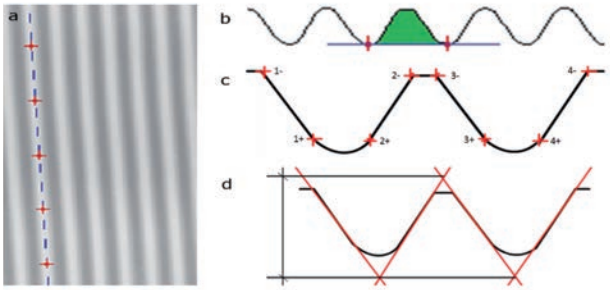

Fig. 7. Sample measurement of thread profile parameters in a 3D image

The minimized function is analogous to those presented in Eq. 1. The only difference is that the set of input points is smaller with each minimization working on a single thread flank. This must be done multiple times. When the calculation is finished, the screw thread flank angle can be easily computed from "a" parameter (linear function slope).

$$\alpha = \arctan(a) \tag{9}$$

Performing a minimization of a distance between real data points representing a flank and straight line can supply one additional piece of information about thread quality. Usually minimization algorithms work with a predefined end value. If the value calculated using Eq. 1 is below a certain limit, the minimization process finishes. If the selected limit value has to be selected from a relatively high value, the thread flanks may be damaged in some way (scratches, grooves, cavities).

The third step is to calculate the intersection points of the lines representing neighbouring flanks. In previous work done by Latypov [5], thread pitch was calculated by fitting a circle between profile flanks. This method is based on the widely used tactile thread measurement method of using three wires. Unfortunately, it requires some initial knowledge about the thread. The fitted circle diameter has to be known in advance based on expected pitch value. Based on our observations, measuring a distance between previously calculated intersection points does not give results that vary much from those calculated based on the Latypov method. The advantage is that no preliminary knowledge is necessary to perform a measurement, so the potential for human error is minimized in this case.

The fourth stage of our calculations allows us to assess the value of a radius at crest and root. In some kinds of threads, measuring these parameters does not lead to any meaningful information. An ISO metric thread often has flatted crests, but for other kinds of thread – for example round threads – both radius values are equally important in order to work properly. In order to obtain a radius value, we have to fit a circle into the section of thread representing the root or crest. Those sections are defined in a similar way to those used for line fitting. The circle fitting algorithm is based on the method described by Gander, Golub, and Strebel [6], if the requested circle is described in the following parametric form:

$$x = z_1 + r\cos\varphi,\ y = z_2 + r\sin\varphi \tag{10}$$

The points representing a root or a crest are given as $P_i(x_{i1}, x_{i2})$, so the function that has to be minimized and takes the following form:

$$\sum_{i=1}^{m}\left\{\min_{\varphi_i}\left[\left(x_{i1}-x(\varphi_i)\right)^2+\left(x_{i2}-y(\varphi_i)\right)^2\right]\right\} \tag{11}$$

As a result of the algorithm, we get a radius value for one selected root or crest. To get the parameters for the entire thread, the operation should be repeated.

The outline analysis allows for automatic detection of defects in the profile created during the thread forming process. Figure 7 presents a three-dimensional image with marked defect areas. Regardless of the type of the defect, the allowance for the loss of material during the measurement is visible as a disturbance of the correct profile. The visibility of the defects depends on the resolution of the 3D vision system, the thread size, and the size of the imaging field. Increasing the field of imaging results in a reduction in the size of defects and, consequently, in the possibility of detecting them.

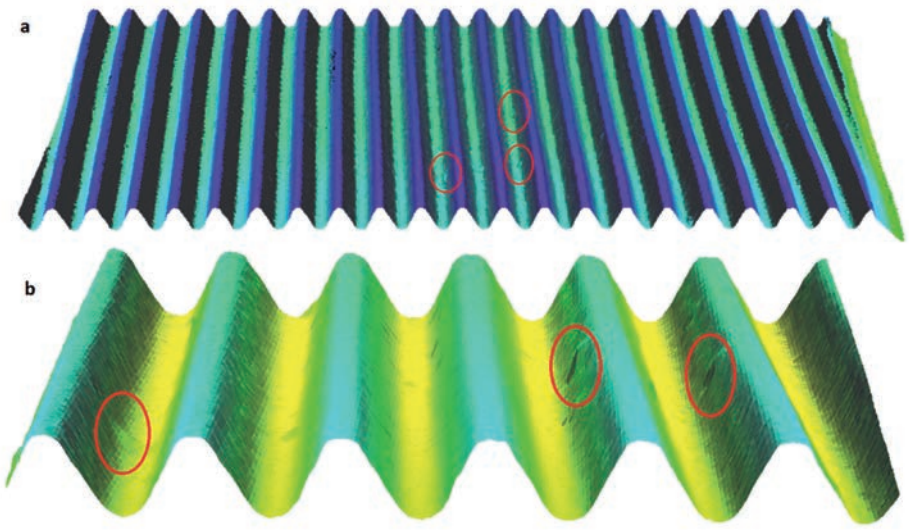

Fig. 8. Thread surface with marked areas of profile defects

Conclusion

The presented method is a good possible solution to solve the screw thread quality control conundrum. It can be implemented on a production line to control every produced threaded element and meets all requirements in terms of speed, accuracy, and overall capabilities. The presented calibration algorithm, using a cylinder-shaped calibration target, is straightforward but can easily ensure that the positioning of the measured object is correct with respect to the measurement unit.

Currently, screw thread quality control is only done on selected elements from a single batch. Moreover, the information obtained is purely binary (the element is either simply good or not good), and identifying

the source of the problem is difficult. Acquiring an image of a thread surface using laser triangulation and then assessing its parameters with the mathematical algorithms presented in this paper allows us to get more comprehensive information about the production process. Future research will involve comparing these obtained results with data from other measuring techniques such as CMM.

From the point of view of conducting quality control in the manufacturing of threads, there is an interesting possibility of combining the description of the thread profile parameters with evaluations of the thread surface. The thread may have the correct parameters of the profile, but it is still possible that it will have surface defects which disqualify its technical use.

References

1. NPL Notes on Screw Gauges. National Physical Laboratory. Great Britain, 2007.
2. Carmignato S., De Chiffre L.: A New Method for Thread Calibration on Coordinate Measuring Machines. CIRP Annals – Manufacturing Technology, 2003, 52(1), pp. 447–450.
3. Kosarevsky S., Latypov V.: Detection of Screw Threads in Computed Tomography 3D Density. Measurement Science Review, 2013, 13(6), pp. 292–297.
4. Mutambi J., Li-jun Y.: Application of Digital Image Analysis Method in Metric Screw Thread Metrology. Journal of Shanghai University, 20014, 8(2), pp. 208–212.
5. Kosarevsky S., Latypov V.: Development of an algorithm to detect screw threads in planar point clouds. Measurement Science Review, 2010, 10(4), pp. 136–141.
6. Chartrand R.: Numerical Differentiation of Noisy, Nonsmooth Data. ISRN Applied Mathematics, 2011, Article ID 164564.
7. Gander W., Golub G., Strebel R.: Least-Squares Fitting of Circles and Ellipses. BIT Numerical Mathematics, 1994, 34(4), pp. 558–578.
8. Sioma A.: The estimation of resolution in 3D range image system. In: International Carpathian Control Conference (ICCC), Rytro (Poland), May 2013. IEEE, 2013, pp. 346–349.