Visual Analysis of Machine Tool Operation Mode Correlation Based on Parameter Category Coding

Jinxin Long[1], Lijuan Peng[1](✉), Xuejun Li[1], Kaiming Ma[1], and Feng Qiu[2]

[1] School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang, China
plj@swust.edu.cn

[2] Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang, China

Abstract. Aiming at the problem that the machine tool operation data has many dimensions, the parameters relationship is complex, and its abnormal patterns and hidden correlation information are difficult to fully excavate, this paper proposed a visual analysis method for the correlation of CNC machine tool operation mode. Firstly, the parameter category encoding is carried out from the two aspects of the sliding window and time point of the machine tool operation data, and then the multi-parameter category encoding combination is clustered and association rule mining is carried out to extract the machine tool operation mode and parameter state association mode, and establish visual map. Integrating ease of use, flexibility, and interpretability, the visual analysis system MachineVis is further constructed, and a variety of interactive methods are designed to support users in discovering abnormal patterns of machine tool data, analyzing the changes of various parameters of machine tool data and capturing the relationship between each parameter. Finally, the validity and practicability of the system are proved by case studies.

Keywords: symbolic aggregate approximation · association rules · machine tool operation data · visual analysis

1 Introduction

CNC machine tools play a vital role in the industrial field, which can control machine movements through digital commands to complete various complex parts processing tasks, and are widely used in military equipment, automobile manufacturing, mold processing, aerospace and other fields. However, breakdowns in the operation of machine tool equipment can cause multiple problems, including damage to the equipment itself, loss of productivity, economic losses and safety accidents. Therefore, abnormality detection and condition monitoring of machine tools are increasingly becoming a hot research topic. Most of the existing anomaly detection methods are constructed based on machine learning models, which are largely applied to feature extraction of sensor signals and

B. Sheng et al. (Eds.): CGI 2023, LNCS 14497, pp. 238–249, 2024.
https://doi.org/10.1007/978-3-031-50075-6_19

equipment fault diagnosis, with little comprehensive analysis and deep exploration of machine tool anomaly data, while providing very limited knowledge in the interpretation of anomaly causes.

With the development of digital promotion work, a variety of sensors and recorders and other equipment can record and detect machine tool production operation status in real time, and accumulate a large amount of machine tool operation data. This data contains valuable domain knowledge such as the changing patterns and correlation characteristics of the machine's operating conditions. In-depth analysis and utilization of equipment operation history data will help to control, analyze and make decisions on the operation status of machine tools. However, machine operation data has multiple dimensions, operation status has multiple categories, and is time-dynamic, which makes machine operation data analysis very difficult.

The rapid development of visualization and visual analytics technology provides an effective way to solve the above problems. Use visualization tools and techniques to quickly visualize machine operation data for more intuitive observation and analysis of changes in machine operating conditions. In addition, visual analytics allows users to interactively select production segments of interest and drill down, layer by layer, to analyze and explore the impact of each parameter in the segment on equipment operation and the relationships between parameters, which not only helps to understand abnormal machine problems, but also helps to uncover potential correlations and patterns in machine operation data to better prevent failures and improve machine productivity and quality.

2 Related Works

The analysis of machine tool operation data is actually the analysis of time series data. Symbolization technology can refer to complex and variable time series data with simpler symbols, weaken the influence of local data fluctuations, and is widely used in time series data analysis. The symbolic aggregate approximation (SAX) [1] is used more, which divides the time series into segments based on the piecewise aggregation approximation method (PAA) [2], and calculates mean of each segment and mapped to different signs on the standard normal distribution. Georgoulas et al. used the method of symbol aggregation approximation and related intelligent icon representation to extract the features of the bearing vibration records, and input the features into the classifier to realize the fault diagnosis [3]. Park et al. apply symbol aggregation approximation to represent time series and use association rule mining to discover frequent rules between symbols of anomalous events [4]. Shi et al. used a symbolic aggregation approximation method to represent the consumer's load curve, and clustered the load curve, and finally analyzed and mined the typical power consumption behavior of the user [5]. Yin et al. proposed an improved Lempel-Ziv method based on a symbolic aggregation approximation method, which can realize early fault diagnosis of bearings [6]. Li et al. proposed an equiprobable association rule mining method on the basis of symbol aggregation approximation, and used it for fault classification and defect severity identification [7]. In terms of correlation modeling, Lv et al. design a region-based adaptive association learning framework that well improves the performance of scene recognition [8]. Gao et al. proposed a dynamic correlation model and an improved fuzzy clustering algorithm for solving the inaccuracy problem in complex image segmentation [9]. Xie et al. proposed a new extensive attention graph fusion network to efficiently perform complex higher-order feature interactions at the granularity of feature dimensions and verified its effectiveness [10].

As an important technology for understanding complex data, visualization is frequently introduced into industrial data analysis [11], and it can efficiently help researchers analyze problems in many industrial application scenarios. Liu et al. designed a novel interactive system ECoalVis, which enables experts to intuitively analyze the control strategies of coal-fired power plants extracted from historical sensors [12]. Eirich et al. proposed the visual analysis system IRVINE, which can help analyze acoustic data in order to detect and understand unknown errors in motor manufacturing [13]. The visualization system designed by Musleh et al. supports local and global exploration of multivariate time-series data generated by blow molding machine sensors [14]. Wu et al. proposed a visual analysis method for factory equipment condition monitoring [15]. The visual analysis system designed by Zhang et al. supports interactive exploration of fault propagation patterns in power grid simulation data [16]. Li et al. designed an automatic household waste detection and sorting system that helps to improve the efficiency of waste recycling [17].

Therefore, this paper proposes a CNC machine tool operation mode association visual analysis method, using sliding window technology, symbolic aggregation approximation, association rule mining and clustering methods to extract the association modes and operation modes in the machine tool operation data, and designing an interactive visual analysis system MachineVis that contains five views: control panel view, global view, local view, parameter state association view, and contrast view.

3 Task Analysis

In order to fully explore the machine operation data, assist users to analyze the machine operation data from the whole to the details, and realize the all-round visual presentation of the operation data. Accordingly, the design tasks for the MachineVis interactive visual analysis system were derived as follows:

- T1: Supports visual exploration of anomalous patterns in historical machine tool operating data. It also helps users to deeply investigate the abnormal pattern of machine operation data.
- T2: Supports visual analysis of the relationships between machine parameters. Users can better understand the relationship between parameters in different modes.
- T3: Supports users in efficiently exploring historical machine operation data and provides flexible interaction.

In this paper, we encode parameter categories and extract symbolic features for machine operation data in terms of sliding windows and time points, and then clustering and association rules are mined for the combination of multiple parameter category codes to extract machine operation patterns and parameter state association patterns (T1, T2), and an interactive visual analysis system MachineVis (T1, T2, T3) is further constructed, as shown in Fig. 1.

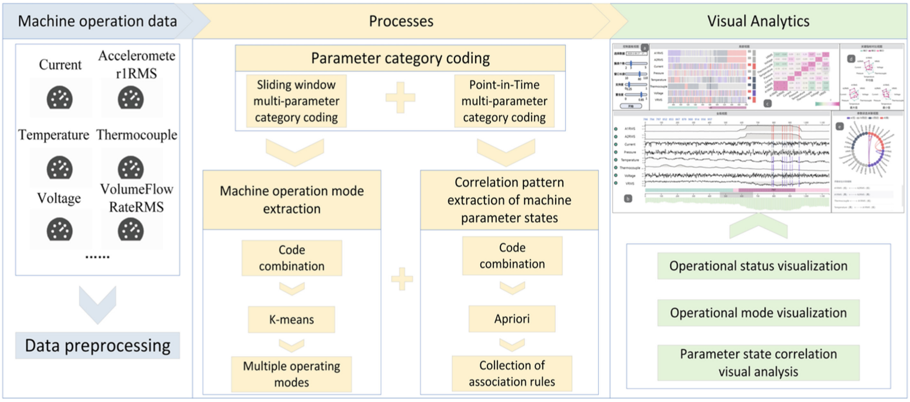


Fig. 1. MachineVis interactive visual analysis system architecture

4 Methodology Overview

4.1 Machine Operation Mode Extraction

In order to detect changes and trends in the data in a timely manner, the SAX method was used to categorize and code the sliding window multiparameter data. A matrix consisting of multidimensional time series can be obtained by first normalizing multiple subsequences of machine tool operation data using z-score. The multidimensional time series matrix is segmented according to the segmentation method of fixed sliding window, and the mean value of each sliding window is extracted for symbol coding, and the coding matrix E is obtained according to SAX coding, denoted as:

$$E = \left\{ \begin{matrix} e_{11} & \cdots & e_{1(n-w+1)} \\ \vdots & e_{ij} & \vdots \\ e_{m1} & \cdots & e_{m(n-w+1)} \end{matrix} \right\} (1 \le i \le m, 1 \le j \le n - w + 1) \tag{1}$$

where e_{ij} is the symbolic encoding of the mean, n is the length of the time series, m is the number of machine parameters, and w is the sliding window length. According to SAX, it can be obtained that when the number of letters is 4, there are three breakpoints of -0.67, 0 and 0.67 which are divided into 4 regions. For the values of these 4 regions, the letters a, b, c and d are used from lowest to highest, respectively.

By extracting the multi-parameter class coding combinations at the same moment and performing cluster analysis on the multi-parameter class coding combinations at different moments. Thus, the similar multi-parameter category codes are combined into the same cluster, and finally the machine operation mode is recognized. In this paper, we use the K-means clustering algorithm (K-means) [18]. Specifically, the multi-parameter category coding combinations in each column of the coding matrix sequence E are extracted and de-duplicated as data samples, and the samples are clustered using the K-means algorithm.

4.2 Machine Parameter State Correlation Pattern Extraction

After normalizing the machine operation data, multiple parameter categories are coded for each time point using the SAX method. The coding matrix is denoted as F:

$$F = \begin{Bmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & f_{ij} & \vdots \\ f_{m1} & \cdots & f_{mn} \end{Bmatrix} (1 \le i \le m, 1 \le j \le n) \tag{2}$$

where m represents the number of machine parameters, n represents the length of the time series, and f_{ij} is the symbolic encoding of the time points. For the coding of multi-parameter categories at time points, the letters representing the different regional divisions and the numbers representing the different parameters are used for symbolization. For example, the code combination in the first column $\{f_{11}, f_{21}, f_{31}, f_{41}\} = \{d_1, c_2, b_3, a_4\}$, it indicates that at the first second, the first parameter is in state d; the second and third parameters are in states c and b, respectively; and the fourth parameter is in state a.

The Apriori [19] algorithm is a common association rule mining algorithm. For each column in the encoding matrix sequence F with a combination of multi-parameter class encoding at time points, the Apriori association rule algorithm is applied to generate the set of parameter state association rules. It consists of two components, which are support and confidence. Users can set their own support and confidence thresholds according to actual needs and data characteristics.

5 Interactive Visual Analytics System Design

This section introduces the five key views of the MachineVis interactive visual analysis system (Fig. 2), which are the control panel view, the global view, the local view, the contrast view and the parameter state association view.

5.1 Control Panel View

In Fig. 2(a), three options are included from top to bottom: selection data, support level, and confidence level. The selection data is used to determine the machine operation data to be analyzed. The settings of support and confidence are related to the extraction of correlation patterns of machine parameter states.

5.2 Global View

In Fig. 2(b), line graphs are used to show the machine operation data, and the shaded areas of the line graphs represent the fluctuation size and duration of the data. The background color of the mouse position represents the current running state of the machine tool. When the user clicks on the state of a certain time point, the global view can search and highlight other time points that are the same as the current machine running state.

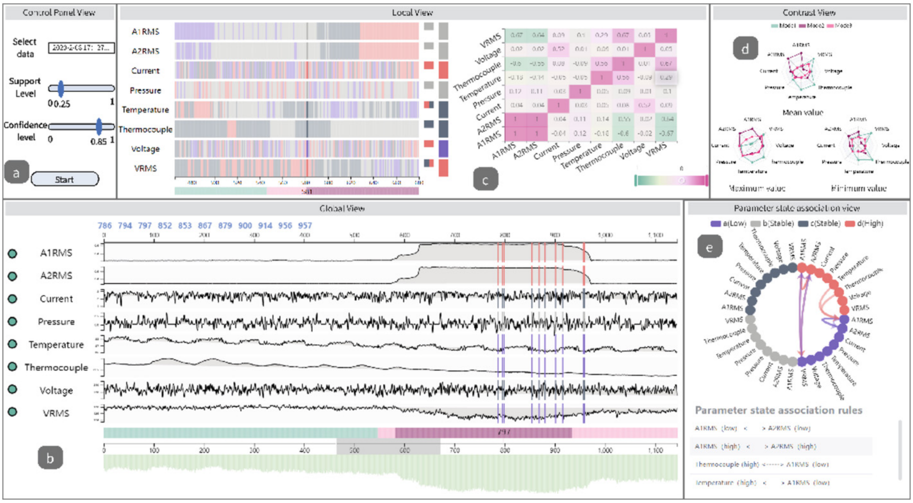


Fig. 2. MachineVis interactive visual analysis system

According to the clustering results, different clusters can represent different operating modes in the machine tool work, and different clusters are mapped with different colors. Below the line graph shows the change of the machine operating mode.

In addition, in the global view, you can adjust the range of the local view by dragging the slider. The light green histogram at the bottom of the view shows the one-dimensional data characteristics of the machine tool operation data after PCA dimensionality reduction. The higher the height of the bars in the bar chart represents the more obvious changes in the data at the current moment, which helps to observe the changes in the overall machine operation status.

5.3 Local View

The left part of the partial view (Fig. 2(c)) shows the detailed changes of each parameter data in the selected time period. If the number of symbol letters is 4, symbol *a* is represented by red, which means the current parameter data is high; symbol *d* is represented by blue, which represents the current parameter data is low; symbol *b* and symbol *c* are represented by light gray and dark gray, representing that the current data is in a stable range.

This view provides a comparison of the state at a certain point in time with the previous state. Put the mouse at a certain time point, the two small rectangles on the left will display the running status of the previous second and the previous two seconds, and the large rectangle represents the current running status of the mouse positioning. Additionally, a heatmap represents the correlation between different parameters over a selected time period.

5.4 Contrast View

In the contrast view (Fig. 2(d)) the comparison of the characteristic values of each parameter of the machine tool in different operation modes is shown in radar plots, the comparison of the characteristic values including average, maximum and minimum values.

5.5 Parameter State Association View

This view (Fig. 2(e)) uses a circular relationship diagram and a table to show the correlation between the states of each parameter. Each circle represents a state, and the color of the circle represents a different state, so the user can quickly distinguish between different parameter states. The line between the circles represents an association rule. When the user clicks on a state, the association between the different states of each parameter and the time when the state association appears can be analyzed and discovered in conjunction with the global view.

6 Case Studies and User Evaluations

6.1 Data Sets and Data Preprocessing

This paper uses the publicly available dataset SKAB, which is a multivariate time series collected from sensors mounted on a motor test bench and contains eight parameters, namely Accelerometer1RMS, Accelerometer2RMS, Current, Pressure, Temperature, Thermocouple, Voltage and VolumeFlowRateRMS. The data is recorded for each second from 2020-02-08 17:27:19 to 2020-02-08 17:47:18 and is 1144 in length.

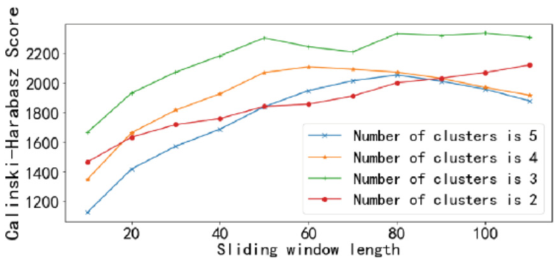


Fig. 3. Cluster comparison of different values

The data first needs to be z-score normalized, using a number of letters of 4. For the number of clusters and different sliding window lengths, this paper uses the Calinski-Harabasz Score clustering model evaluation metric to evaluate, with larger metrics indicating better clustering results [20]. In our experiments, we consider the case where the number of clusters is between [2–5] and the sliding window length is between [10–110] with a step size of 10. As shown in Fig. 3, it can be found that the evaluation metric is greatest when the sliding window length is taken as 80 and the number of clusters is 3.

Therefore, in this paper, the sliding window length of 80 and the number of clusters of 3 are used to extract the machine operation mode. In order to find the rules of strong correlation of parameter states in machine operation data, the support is set to 0.25 and the confidence level is set to 0.85.

6.2 Anomaly Analysis

In Fig. 4, the color rectangle bar below the global view shows that the rectangle colors are mapped to three different colors, so the machine operation mode is divided into three modes.

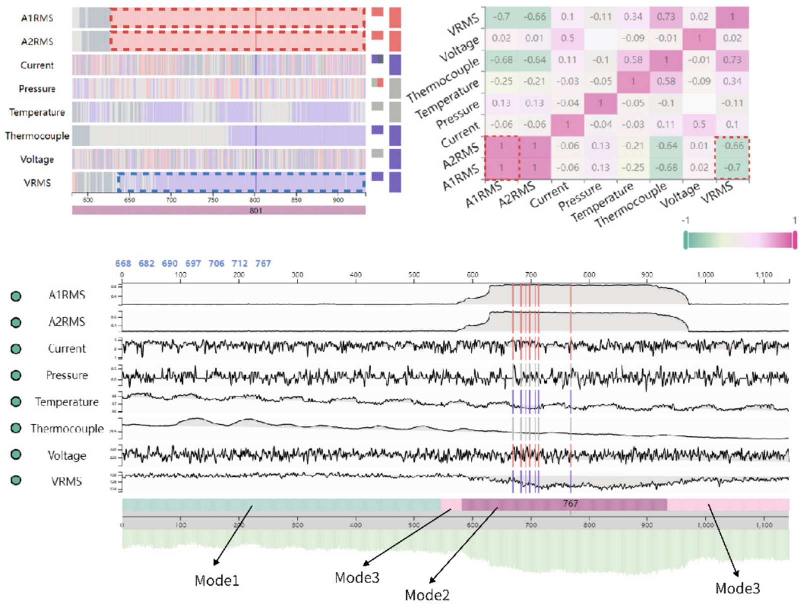


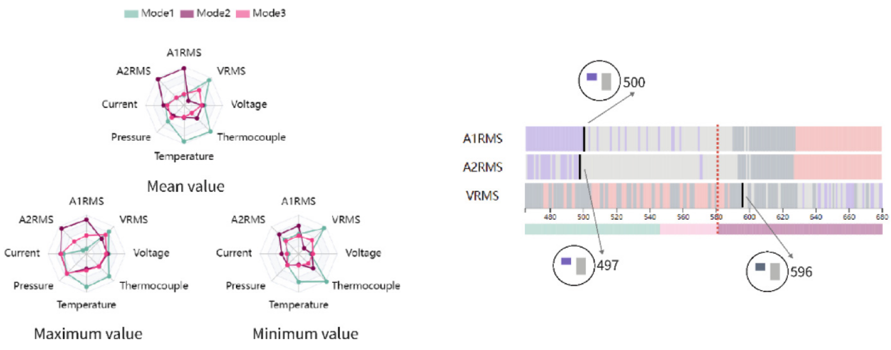
Fig. 4. Anomaly analysis

In the view you can see that the light green bar at the bottom rises in height and lasts longer in the middle and late stages, representing a greater change in parameter status from the previous one. Also in operation Mode 2, it is clearly observed that the parameters A1RMS (abbreviation for Accelerometer1RM1), A2RMS (abbreviation for Accelerometer1RM2) and VRMS (abbreviation for VolumeFlowRateRMS) have a large shadow area, indicating that the parameters fluctuate greatly during this time period and that an abnormal condition is likely to have occurred. Therefore, special attention needs to be paid to the distribution of these parameter states during this time period. By dragging the time brush in the global view and positioning the time range in operation mode 2, it can be seen by looking at the local view that the red area of A1RMS and A2RMS covers a wider area and the blue area of VRMS occupies most of the area. This indicates that most of the data for A1RMS and A2RMS were relatively high during this time period, while most of the data for VRMS were relatively low, allowing for greater certainty that A1RMS, A2RMS, and VRMS were anomalous during the Mode 2.

By looking at the correlation heat map on the right side of the local view, a positive correlation is found between A1RMS and A2RMS in operation mode 2, and there is also a high negative correlation between A1RMS and A2RMS and VRMS. Where when clicking on the state at the 767th second in the anomaly mode, it was found that the state also appeared at the moments 668, 682, 690, 697, 706 and 712, indicating that a similar anomalous state appeared at these moments.

6.3 Sensitive Parametric Analysis

The radar plot comparison reveals that the characteristic values of Current, Voltage and Pressure are relatively close in the three operation modes (Fig. 5(a)). It shows that the Current, Voltage and Pressure are in a stable state during the whole working process. However, A1RMS, A2RMS, VRMS, Temperature and Thermocouple have large disparities in multiple eigenvalues in different modes of operation. The A1RMS, A2RMS and VRMS are probably due to a lack of stability in the machine's operating conditions or to abnormal factors, as judged by the global view in Fig. 4. And Temperature and Thermocouple show an overall decreasing trend, so leading to a large difference between the eigenvalues in the different modes.



(a) (b)

Fig. 5. Discovering sensitive features and sensitive state changes

The three anomalous sequences A1RMS, A2RMS and VRMS are analyzed in detail in the local view (Fig. 5(b)). it is found that before the anomaly occurred in the machine (before the red dashed line in the figure), it was found that the color combination of the rectangular bars of A1RMS and A2RMS changed considerably in the first period. Combined with the change in the size rectangle and the positioning of the mouse, it was found that the state of A2RMS and A1RMS changed considerably from the 497th and 500th seconds respectively. We can then deduce that the persistent elevation of A1RMS and A2RMS may have been the cause of the abnormal occurrence of the machine to the extent that it caused the VRMS to become abnormal.

6.4 Parameter State Correlation Analysis

In the parameter state association view (Fig. 6), it is found that A1RMS(d), A2RMS(d) and VRMS(a) are associated with each other and that a mouse clicks on one of the states shows the location of the individual state appearances in the global view. We found that the regions where the correlation states of A1RMS(d), A2RMS(d), and VRMS(a) occur are mostly within the anomaly interval, i.e., we can infer that A1RMS, A2RMS, and VRMS usually produce anomalies together and that VRMS is lower when A1RMS and A2RMS are higher.

Where A1RMS(a) and A2RMS(a) are also correlated with each other, indicating that when A1RMS is lower in Operation Mode 1 and Operation Mode 3, A2RMS is also lower. It was also found that both Temperature(d) and Thermocouple(d) point to A1RMS(a), indicating that within Operation Mode 1, when Temperature and Thermocouple are higher, A1RMS is also typically lower.

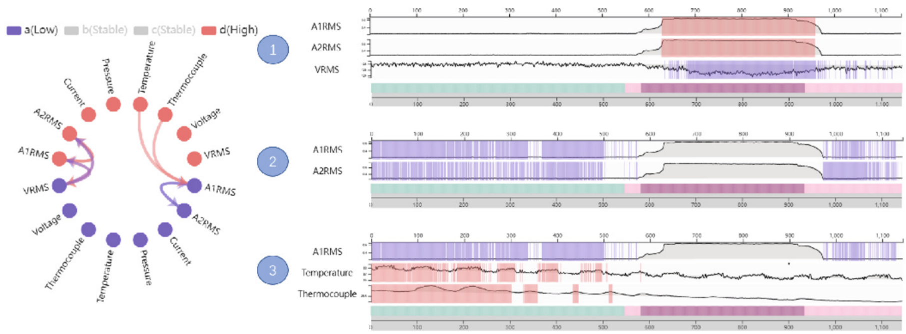


Fig. 6. Parameter state association pattern analysis

6.5 User Evaluation and Feedback

To further verify the effectiveness and usability of the system, eight volunteers were invited to evaluate the system. Two of the experts are from the field of machine automation research, three from the field of visualization and three from the field of industrial data analysis. We first designed a questionnaire:

- Q1: The system is able to distinguish between different machine operation modes.
- Q2: The system is able to observe machine operation in detail and quickly identify abnormalities.
- Q3: The system is able to clearly show machine parameter state correlations.
- Q4: The system offers flexible ways to help users explore machine operation data.

The volunteers were then explained how the system works, and finally a Likert scale was used, with volunteers scoring from 1 (strongly disagree) to 10 (strongly agree).

The results of the evaluation (Fig. 7) show that for each question the average score is above 8.5, indicating that our system is a good solution to the user's needs. After completing the evaluation, a return visit was made to experts in the field of machine automation research to discuss the usability and visualization design of the system, and the following are the valuable optimization suggestions made by the experts in the field:

1. Further comparative analysis of abnormal patterns can be achieved for multiple possible abnormal patterns.
2. The presentation of parameter status correlation information can be further enriched, allowing the user to quickly access parameter correlation information in different operating modes from multiple perspectives.

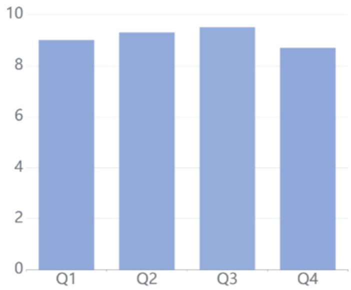


Fig. 7. Questionnaire result

7 Conclusion

This paper's interactive visual analysis system MachineVis, is used in a case study on the dataset. Through the visual mapping of the correlations among machine operating modes, machine operating states, and machine parameter states and the interactive linkage of multiple views, it is demonstrated that MachineVis can easily detect and analyze anomalies, and can fully explore the interactions, correlations, and their hidden correlation information among machine parameters. In future usage scenarios, we hope that the methods in this paper can be applied to different types of industrial systems to provide references for deeper mining and analysis of their anomaly patterns.

References

1. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 2–11 (2003)
2. Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S.: Dimensionality reduction for fast similarity search in large time series databases. Knowl. Inf. Syst. **3**, 263–286 (2001)
3. Georgoulas, G., Karvelis, P., Loutas, T., Stylios, C.D.: Rolling element bearings diagnostics using the Symbolic Aggregate approXimation. Mech. Syst. Signal Process. **60**, 229–242 (2015)

4. Park, H., Jung, J.-Y.: SAX-ARM: Deviant event pattern discovery from multivariate time series using symbolic aggregate approximation and association rule mining. Expert Syst. Appl. **141**, 112950 (2020)
5. Shi, Y., Yu, T., Liu, Q., Zhu, H., Li, F., Wu, Y.: An approach of electrical load profile analysis based on time series data mining. IEEE Access. **8**, 209915–209925 (2020)
6. Yin, J., Xu, M., Zheng, H.: Fault diagnosis of bearing based on symbolic aggregate approXimation and Lempel-Ziv. Measurement **138**, 206–216 (2019)
7. Li, Y., et al.: Association rule-based feature mining for automated fault diagnosis of rolling bearing. Shock Vibr. **2019**, 1–12 (2019)
8. Lv, G., Dong, L., Zhang, W., Xu, W.: Region-based adaptive association learning for robust image scene recognition. Vis. Comput. **39**, 1629–1649 (2023)
9. Gao, X., Zhang, Y., Wang, H., Sun, Y., Zhao, F., Zhang, X.: A modified fuzzy clustering algorithm based on dynamic relatedness model for image segmentation. Vis. Comput. **39**, 1583–1596 (2023)
10. Xie, Z., Zhang, W., Sheng, B., Li, P., Chen, C.P.: BaGFN: broad attentive graph fusion network for high-order feature interactions. IEEE Trans. Neural Netw. Learn. Syst. **34**, 4499–4513 (2021)
11. Zhou, F., et al.: A survey of visualization for smart manufacturing. J. Vis. **22**, 419–435 (2019)
12. Liu, S., et al.: ECoalVis: visual analysis of control strategies in coal-fired power plants. IEEE Trans. Visual Comput. Graphics **29**, 1091–1101 (2022)
13. Eirich, J., et al.: IRVINE: a design study on analyzing correlation patterns of electrical engines. IEEE Trans. Visual Comput. Graphics **28**, 11–21 (2021)
14. Musleh, M., Chatzimparmpas, A., Jusufi, I.: Visual analysis of blow molding machine multivariate time series data. J. Vis. **25**, 1329–1342 (2022)
15. Wu, W., Zheng, Y., Chen, K., Wang, X., Cao, N.: A visual analytics approach for equipment condition monitoring in smart factories of process industry. In: 2018 IEEE Pacific Visualization Symposium (PacificVis), pp. 140–149. IEEE (2018)
16. Zhang, T., Chen, Z., Zhao, Z., Luo, X., Zheng, W., Chen, W.: FaultTracer: interactive visual exploration of fault propagation patterns in power grid simulation data. J. Vis. **24**, 1051–1064 (2021)
17. Li, J., et al.: Automatic detection and classification system of domestic waste via multimodel cascaded convolutional neural network. IEEE Trans. Industr. Inf. **18**, 163–173 (2021)
18. Hartigan, J.A., Wong, M.A.: Algorithm AS 136: A k-means clustering algorithm. J. Roy. Stat. Soc. Ser. C (Appl. Stat.) **28**, 100–108 (1979)
19. Agrawal, R., et al.: Fast algorithms for mining association rules. In: Proceedings of the 20th International Conference very large data bases, VLDB, pp. 487–499. Santiago, Chile (1994)
20. Caliński, T., Harabasz, J.: A dendrite method for cluster analysis. Commun. Stat.-Theory Methods **3**, 1–27 (1974)