ISSN: 2664-9640

Application of improved revolving door algorithm in FMS real-time data compression

Tao Zhang

School of Mechanical Engineering, Sichuan Light Chemical Engineering University, Yibin, 644000)

Zhang Tao, 1996 --), M. S. candidate, whose research interest covers machine tool data acquisition. E-mail: 1297262529@qq.com.

Abstract

For FMS in real-time data acquisition process takes up a lot of memory space problem, in order to make a lot of real-time process data for storage and management, puts forward a improved method based on the revolving door (SDT) of real-time data is compressed, the revolving door of compression algorithm is based on the data of wave to improve the tolerance, Delete part of the data while keeping the original data rule, which reduces the PC memory.

Keywords

SDT; Real-time process data; Tolerance improvement; Reduce memory space.

1. Introduction

With the rapid development of the manufacturing industry, it has gradually developed towards the direction of intelligent manufacturing. It has become an essential process to monitor the processing site of the underlying equipment, so it is necessary to collect the real-time data generated in the processing process. However, PLC will produce data in each scanning cycle, and only need to use the data at a certain moment. Other unused data is not lost but stored in the PC, which can take up a lot of memory space. In order to reduce the memory space, the unnecessary data is deleted without changing the law of the original data, and the revolving door compression algorithm is used to compress the data so as to reduce the memory space.

2. Process data processing

Due to the rapid development of intelligent manufacturing technology, all enterprises around the world are developing towards the direction of digitization and intelligence, and with it, more and more types and amounts of data are coming. For numerical control equipment, usually on the data acquisition requirements are higher, not only need to collect high efficiency but also do take up less space, and data acquisition.

The types of sets are also rich, such as: load, temperature, spindle, feed shaft, PLC, power, current, etc., the amount of data a day will reach the GB level, therefore, if all kinds of data are not processed directly stored, not only occupy a large memory space and easy to reduce the life of the machine tool. Firstly, a large number of process data are classified in a certain way, and then data compression is implemented for each module, so as to reduce the memory size and improve the collection efficiency.

2.1. Data compression algorithm

Data compression is a technology that reduces the amount of data without losing useful information so as to reduce the memory space, improve the efficiency of transmission, storage and processing, or reorganize the data according to a certain algorithm to reduce the miscellaneous data and storage space. With the intellectualization of the factory, more and more data are produced, and the problem of data storage is more and more serious, so the data compression technology has attracted attention. An important index to measure the efficiency of data compression is compression ratio (CR), which represents the ratio of the amount of data before compression to the amount of data after compression. The higher the compression ratio is, the better the compression effect will be and the less memory space will be occupied after compression. Another index is compression error (CE), which indicates the similarity between the recovered data after compression and the data before compression.

$$CR = \frac{n}{m}$$

$$CE = \sqrt{\frac{1}{n}\sum\left(y_n - y_m\right)^2}$$

Where, m is the amount of data after compression n , y_m is the amount of data before compression, y_n is the amount of data after decompression and recovery after compression, and is the actual data.

At present, there are two types of data compression: lossy compression and lossless compression. Lossy compression means that after the data volume is compressed, the data after decompression and recovery is different from the data before compression. Lossless compression means that after the data volume is compressed, there is no difference between the data before compression after decompression and recovery after compression[1]. Generally, data compression is divided into three categories in industry[2]: piecewise linear interpolation method, vector quantization method and signal change method. Among them, piecewise linear interpolation methods, including rectangular wave string method, reverse slope method, revolving door method (SDT) and piecewise linear trend method, are most widely used in the industrial field. Rotating door algorithm[3] is a fast linear fitting lossy compression algorithm, which has the characteristics of high efficiency, high compression ratio, simple implementation, small error, etc. Vector quantization method needs to spend a lot of time to calculate the number of codes, low efficiency; Signal transformation is to transform data from one form to another form of processing methods, usually including discrete cosine transform, wavelet transform, etc., because of the advantages of wavelet transform, become the current research hotspot.

3. Data processing based on revolving door data compression algorithm

Based on the complexity of the industrial field of machining, every moment can produce processing data, in view of the large amount of data and use PC memory problems, is now using the revolving door data compression algorithm to compress in the process of data processing, in keeping the law of the original data delete some data and makes occupy PC memory is reduced.

3.1. Revolving door compression algorithm principle

The revolving door compression algorithm has the advantages of small computation, can track the process trend change and is suitable for rolling compression. Rotating door algorithm is based on the upper and lower two points of a certain height from the last storage point as the fulcrum to construct a parallelogram. When there is a data, the door closes. As the data increases, the door gradually opens, and once opened, it cannot be closed. As long as the Angle between the two doors is less than 180 degrees, the door will continue to open slowly. Only when the Angle is greater than 180 degrees, the operation will stop, and the current point will be stored, and the compression will continue with the current point. The principle is shown in

Figure 1.Point I is taken as the previous pre-storage point, and the two points above and below the point I as E are taken as the starting points of the construction of two doors. At point a, both doors are just able to reach point a; however, at point b, the lower door cannot reach point b, so the state stored at point a is as. With the increase of the number of points, when the door arrives at point C, the sum of the current Angle and the current Angle of the lower door is greater than 180 degrees, so save point C and continue to compress with point C as the starting point. As can be seen from Figure 1, the points that meet the conditions are c, g and I. Therefore, in the current period, the data points compressed by the revolving door are c, g and I, and the rest points are discarded[4][5][6][7].

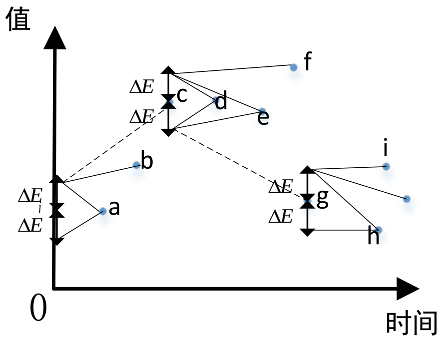

Figure 1 Revolving door compression algorithm principle

3.2. Improved revolving door compression algorithm

According to the principle of the revolving door algorithm, ΔE is the key factor that determines the effect of data compression [8]. If the value ΔE is larger, although the compression efficiency is high, the compression error is large, and the degree of data loss after decompression is large, and the compression effect is poor. On the contrary, if the value ΔE is small, then the compression efficiency is low. Although the compression error is small and the degree of data loss after decompression is small, the compression effect is poor and the purpose of compression cannot be achieved. Therefore, the value of ΔE determines the data compression effect.

3.2.1 Improved principle of revolving door compression algorithm

As a lossy compression algorithm, the revolving door compression algorithm will inevitably abandon some data. If extremely individual data points fluctuate greatly, in the process of data compression, if the abandonment of them will affect the compression accuracy and the compression error will increase. For example, when a certain ΔE is used as the tolerance of the revolving door, one point at a certain time differs greatly from other points at a certain time,If this point is the last point in the compression interval, it will be recorded as the starting point of the next compression interval, If the point is the point between the compression interval, then the point is discarded, The accuracy of the data after decompression will be much lower than that before decompression, Therefore, the value of ΔE determines the degree of fluctuation of the data interval.Mathematically, the standard deviation is used to describe the degree of fluctuation of a string of data, and the standard deviation of a certain time region is used to reflect the degree of fluctuation of data in this time region [9]. In addition, the degree of fluctuation of a linear curve tends to be stable. Therefore, the degree of fluctuation in the next period of time can be reflected according to the degree of fluctuation in the previous period. The value of tolerance ΔE is measured by the standard deviation of two adjacent time regions.

3.2.2 Rotating door compression algorithm improvement steps

In the improved revolving door compression algorithm, ΔE represents the tolerance, u represents the mean value, ΔE_{max} represents the maximum tolerance, ΔE_{min} represents the minimum tolerance, and σ represents the standard deviation.The basic improvement steps are as follows:

(1)During the first compression, data points of a certain time interval are taken and parameters are initialized at first.$\Delta E=(\Delta E_{max}+\Delta E_{min})/2$，$\sigma=\Delta E$.

(2)Standard SDT compression algorithm is adopted, with the most recently saved data point as the starting point and the current data point as the end point. The connection between the two points is the middle axis to construct a parallelogram, All data points in the time region are measured. The points inside the parallelogram are discarded, and the absent points are saved and repeated as the starting point of the new parallelogram.

(3)After the compression of this time interval is completed, the data in this region is decompressed, and the decompressed data is denoted as $x_1, x_2, x_3, \cdots, x_n$,then the maximum error $\sigma_{max}=\max(|x_i-y_i|)$,Among them i=1,2,⋯,n. if$\sigma_{max}\geq\sigma$,then save the original data corresponding to σ_{max}.

(4)After the completion of the compression in this time interval, the data in this region were decompressed and the average value of the original data was calculated

$$u=\frac{1}{n}\sum_{i=1}^{n}y_i$$

Therefore, the standard deviation of the original data in this time interval is

$$\sigma=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-u)^2}$$

(5)Calculate the values of σ in two adjacent time intervals, denoted as $\sigma_{前}$ and $\sigma_{后}$,if $\sigma=0$,Represents no fluctuation in the interval, then ΔE remains unchanged, if $k=\sigma_{前}/\sigma_{后}\geq 1$,means it fluctuates a lot, then $\Delta E=\Delta E/k$; if $k=\sigma_{前}/\sigma_{后}\leq 1$,means the fluctuation is small, then $\Delta E=\Delta E\times k$.

(6)Update the standard deviation of the compression interval, let $\sigma_{后}=\sigma$.

The process of the improved algorithm is shown in Figure 2. Compared with the previous improved revolving door compression algorithm, the improved revolving door compression algorithm can set an optimal tolerance to improve the compression efficiency[10].

4. Experimental analysis and verification

In order to verify the effect of data compression, it is now tested. In this test, the current position value of the FMS exchange car is taken as the test data. The current position changes with time, so data will be generated all the time. In this test, the compression effect is predicted by comparing and analyzing the change trend of data after compression with that before compression. In the non-real-time operating system Windows, the real-time communication of EtherCAT fieldbus is guaranteed by the real-time kernel of TWINCAT master software. The PLC control cycle set by it is 10ms, that is, the master station sends a message every 10ms. Therefore, this test collects and exchanges the data of the current position within 10s of the car. That is, a total of 1000 data were collected (due to the large amount of data, only part of the data was recorded). The initial value of compression tolerance E was set as 0.1 to perform the data compression test. The partial results of data compression are shown in Table 1, and the curve of compression trend is drawn according to the data, as shown in Figure 3).

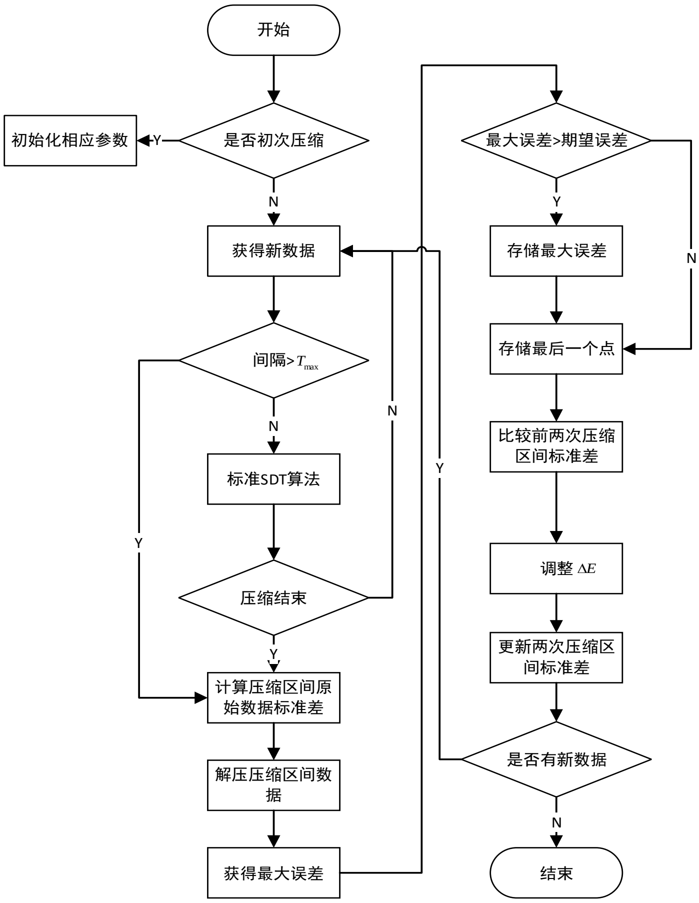

Figure 2 Improved algorithm flow

Table 1 partial compression results of SDT

	Standard revolving door algorithm (unit: mm)			Improved revolving door algorithm (unit: mm)		
Time (s)	Original data	Compressed data	Decompress data	Original data	Compressed data	Decompress data
1.08	0.0118		0	0.0118		0
1.09	0.032600001		0	0.032600001		0
1.1	0.071599998		0	0.071599998		0
1.11	0.1338	0.1338	0.1338	0.1338	0.1338	0.1338
1.12	0.247199997		0	0.247199997		0
1.13	0.399399996		1	0.399399996		0
1.14	0.612999976		1	0.612999976		1
1.15	0.859399974		1	0.859399974	0.8594	0.8594

1.16	1.14760005	1.1476	1.14760005	1.14760005		1
1.17	1.48179996		1	1.48179996		2
1.18	1.75839996		2	1.75839996		2
1.19	2.07179999		2	2.07179999		2

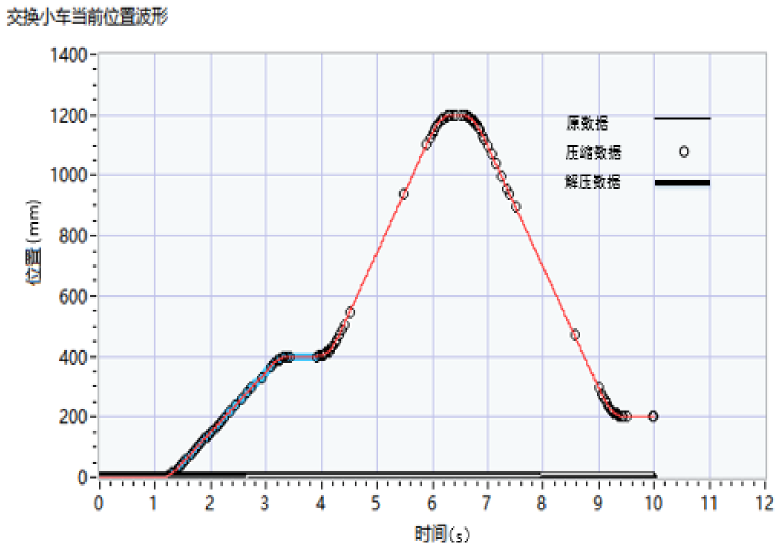

Figure 3 Data compression comparison curve of the revolving door algorithm

As shown in Figure 3, the changing trend of the positions of original data, compressed data and decompressed data over time within 10 seconds is recorded respectively. There are 1000 original data in total. After the revolving door data compression, the data has 112,and the compressed data can be roughly close to the original data change curve, At the same time, the decompressed data can ensure the change trend of the original data(In order to facilitate the comparison of trends, only trends within the first 4s are recorded in the decompressed data),this greatly reduces the memory space. But there are too few compressed data points between 4s and 6s and between 7s and 9s,it may be impossible to guarantee the trend comparison between the original data and the compressed data in this period, because the value of tolerance E is not appropriate, resulting in a certain error. Therefore, the algorithm is improved to be closer to the original curve.

Improved revolving door algorithm is based on the initial tolerance in a certain way to modify the tolerance to determine. According to the formula, the mean value of the original data is 497.3029017, and the standard deviation is 384.1294977. The standard deviation of decompressed data is 385.4856576, and the revised tolerance E is obtained to be 0.09964819446. The data compression effect obtained according to the tolerance is shown in Figure 4.

As shown in Fig. 4, the improved revolving door algorithm retains the variation trend of the original curve of the standard revolving door algorithm, while abandoning some non-critical information of slight disturbance. Moreover, compression data points are added between 4S and 6S and between 7S and 9S to ensure that it is closer to the original curve and more in line with the compression target. In order to further explain the advantages of the improved revolving door algorithm compared with the standard revolving door algorithm, some parameters are calculated and explained. The compression error Ce and compression ratio Cr are obtained from the formula. The parameters are shown in Table 2.

ISSN: 2664-9640

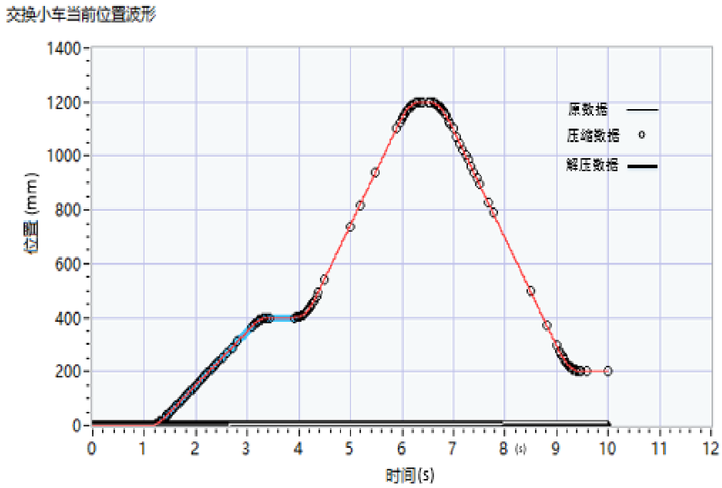

Figure 4 Data compression comparison curve of the improved revolving door algorithm

Table 2 Comparison of parameters between standard revolving door algorithm and improved revolving door algorithm

Algorithm	Raw data volume	The amount of compressed data	Compression ratio(CR)	Compression error(CE)	Time/ms
Standard revolving door algorithm	1000	112	8.928571428	0.4001963802	31
Improved revolving door algorithm	1000	127	7.874015748	0.3070063616	15

As shown in Table 2, the compression ratio of the improved revolving door algorithm is slightly lower than that of the standard revolving door algorithm under the same compression of 1000 data, the reason is that the improved algorithm retains some data points with large errors, but reduces the compression error, at the same time, the compression time is less than the standard revolving door algorithm. Therefore, with the same compression error, the improved algorithm has better compression performance.

5. Conclusion

The paper aiming at the problem of large amount of real-time data, a revolving door compression algorithm is proposed to compress data and reduce the memory space.the algorithm is improved by changing the tolerance by dynamic fluctuation. First, the collected data was compressed using the standard revolving door algorithm and the trend curve was drawn. Through analysis, it is found that the position of the compressed curve can not express the original law, so as to test the improved revolving door algorithm. Through the comparison between the standard revolving door algorithm and the improved revolving door algorithm, the improved revolving door algorithm reduces the compression error and the compression time is shorter. Therefore, the application of this algorithm in industry can achieve better results, save system storage space, reduce storage costs, and improve the efficiency of data access.

ISSN: 2664-9640

References

[1] Wangwang, Xinchu Chen, Dingxing Lu. Improved Research and Application of Process Data Compression Algorithm SDT [J]. Industrial Control Computer,2009,22(08):1-3+6.

[2] Feng Dong.Improved SDT Process Data Compression Algorithm[J].High Technology Letters, 2003(02):91-96.

[3] Yang Xin, Chen Shu. Research on joint coding technology of source and channel based on integer wavelet [J]. Digital Technology and Application,2010(01):114-117.

[4] Wangwang Chen, Xinchu Chen, Dingxing Lu. Improved Research and Application of Process Data Compression Algorithm SDT [J]. Industrial Control Computer,2009,22(08):1-3+6.

[5] Xu Feng. Research on Compression Algorithm in SCADA System [D]. Southwest Jiaotong University,2012.

[6] Qiu Hongkai. Research on Key Technology of Embedded Real-time Database for Electricity Consumption and Energy Efficiency Information Acquisition System [D]. North China Electric Power University,2015.

[7] Yang Guanjie, Li Yue. Analysis of Industrial Real-time Data Compression Algorithm [J]. Automation & Instrumentation,2016(08):214-215.

[8] R S H. Mah et al. Process trending with piecewise linear smoothing[J]. Computers and Chemical Engineerning,1995,19(2):129-137.

[9] Zhou Jing, Ni Jing, Meng Zhen, Chen Xing, Dong Yunlong. Research on Synthetic Test Method of Joint Servo System Based on Ether CAT [J]. Mechanical and Electrical Engineering,2020,37(12):1492-1497.

[10] Liu Jinsong. Research and Application of Intelligent Monitoring Technology for CNC Machine Tools [D]. University of Chinese Academy of Sciences (Shenyang Institute of Computing Technology, Chinese Academy of Sciences),2017.