

Research on the Application of CBR Technology in Intelligent Process Design System

Junli Liu, Hui Lu *, Guanhui Cui, Xibin An

School of Mechanical and Power Engineering, Henan Polytechnic University, Jiaozuo, Henan, China

* Corresponding author: Hui Lu (Email: 212105020011@home.hpu.edu.cn)

Abstract: A Case-Based Reasoning (CBR) intelligent process design system is developed through Visual Studio development tools to improve the processing efficiency of mechanical parts and the recurrence rate of corporate knowledge. The key factor in improving the accuracy of case matching in the CBR system is the similarity calculation of parts. In this paper, similarity calculation models for different attribute types are presented by combining the nearest neighbor method. And the improved AHP method and matrix calculation of MATLAB are used to determine the weighting coefficient. The most similar cases are matched according to the overall similarity of the cases and the set threshold, and the method is applied to the intelligent process design of shafts. The results show that this method is conducive to shortening the development cycle and quickly responding to the market, which provides a reference for intelligent manufacturing of mechanical parts.

Keywords: Intelligent Process; Case-Based Reasoning; Similarity; Improved Analytic Hierarchy Process; Weight Coefficient.

1. Introduction

In mechanical product design, process design is the bridge that connects designers and manufacturing floors. The design efficiency of the part process directly affects the production cycle of the product. The design efficiency of the part process directly affects the production cycle of the product. In the face of complex process systems, the reuse of similar processes based on product three-dimensional models and enterprise knowledge management has become the mainstream means of design of many enterprises. Case-Based Reasoning (CBR) is a similar problem-solving method that uses the existing knowledge and cases of enterprises, and is undoubtedly the preferred method for dealing with complex problems and multi-attribute decisions[1].

based on case reasoning was developed. According to the overall similarity of the cases and the set threshold, the most similar cases are matched, which is used to guide the design of new cases, and provides a fast process design method for mechanical parts, thereby improving the knowledge reuse rate and process design efficiency of mechanical parts in enterprises.

2. Establishment and Solution of CBR Model

In the life cycle of CBR, case retrieval is an important part of case reasoning, and CBR is based on the idea that "similar problems have similar solutions"[3]. Therefore, in CBR calculation, the key to case retrieval lies in the calculation of similarity, and the case similarity calculation process mainly includes two aspects: ①Evaluate the similarity between the target case and the historical case on each feature item, that is, the local similarity; ②Evaluate overall similarity on the basis of local similarity[4].

$$Sim(X, C) = \sum_{j=1}^n f(X_j, C_j)W_j \quad (1)$$

In the formula, X represents the target case; C represents the mature case; n indicates the number of feature attributes contained in the case; f represents the similarity calculation function between case C and target case X ; W_j represents the weight of the feature attribute J . According to Equation (1), it can be seen that attribute similarity calculation and weight assignment will affect the matching accuracy of the case and the reliability of the final decision[5]. Figure 2 shows the overall framework of the similarity calculation solution. Different similarity calculation models are established according to the accurate numerical type, text type and feature set type. By using Delphi method combined with the improved AHP method, the weight of each attribute is calculated. The global similarity of each case is obtained, and the best similar case is finally matched.

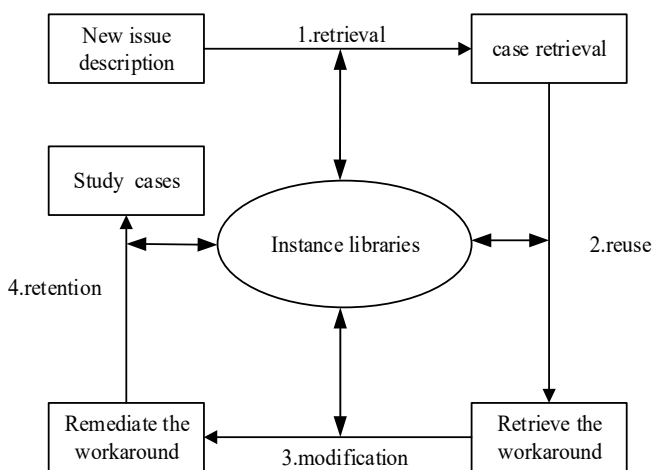


Figure 1. Case-Based Reasoning diagram

CBR technology solves current problems by comparing the differences between new and old problems, modifying and adjusting the cases and reusing the knowledge and information of old cases[2]. A complete CBR solution should include retrieval, reuse, modification, and retention of cases, as shown in Figure 1. Taking the belt conveyor roller shaft parts as an example, an intelligent process design system

3. Calculation Model for Attribute Similarity

In this paper, K-Nearest Neibhdor (K-NN) is used to describe the similarity between case attributes. It determines similarity by calculating the spatial distance of attribute features. and provide highly similar cases to designers for reuse[6]. In this paper, the process parameters such as the category, blank, material, heat treatment method, basic parameter set (shaft length, maximum outer diameter, roughness, main machining accuracy) and shape feature set (number of main features and number of auxiliary features) of shaft parts are described. Since the process attribute information involves types mainly include numeric, text and set. Therefore, their similarity calculation function $f(X_j, C_j)$ is also different. The following will establish the similarity calculation model for each type of attribute in the shaft part.

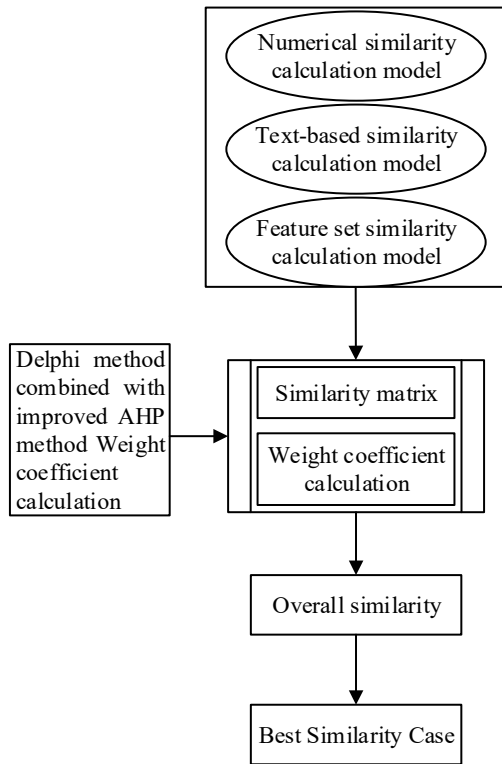


Figure 2. Intelligent design architecture of mechanical products

3.1. Similarity Calculation of Numeric Types

When the attribute is a deterministic numeric type, such as the basic process parameters of the axis class (shaft length, maximum outer diameter, roughness, main machining accuracy), the spatial distance between the two is usually used as the similarity calculation method.

$$Sim(X, C_i) = 1 - Dis(X, C_i) = 1 - \left[\sum_{j=1}^n w_j Dis(x_j, c_{ij}) \right]^{\frac{1}{r}} \quad (2)$$

In the formula, when $r = 1$, $Dis(X, C_i)$ represents the Hamming distance; When $r = 2$, $Dis(X, C_i)$ represents the Euler distance.

Where $Dis(X, C_i)$ is expressed as:

$$Dis(x_j, c_{ij}) = \frac{|x_j - c_{ij}|}{x_{max} - x_{min}}, x_j, c_{ij} \in [x_{min}, x_{max}] \quad (3)$$

In the formula, x_{max} represents the maximum value of the case attribute j , and x_{min} represents the minimum value of the

case attribute. When $Dis(X, C_i)=0$, it means that the properties j of the two cases are the same, and the similarity is 1; Otherwise, the two properties are completely different and have a similarity of 0.

3.2. Similarity Calculation of Text Types

In this system, some properties of the case are described by text, such as part name, stock type, heat treatment method and material, collectively referred to as text type. For text class shapes, the similarity is calculated by the following formula:

$$Sim(x_j, c_{ij}) = \begin{cases} 0, & x_j \neq c_{ij} \\ 1, & x_j = c_{ij} \end{cases} \quad (4)$$

3.3. Similarity Calculation of Feature Set Types

Due to the complex geometric features of the structure of shaft parts, the structure will involve shape features such as bosses, slots, and chamfers. Therefore, the geometric features are described as sets such as primary features and auxiliary features, as shown in Figure 3.

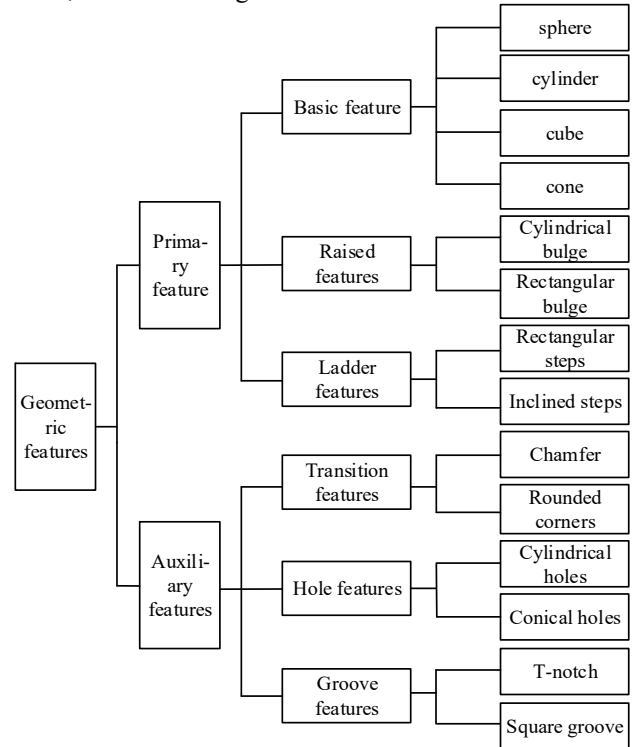


Figure 3. Classification of shape features

The similarity of the shape feature set is judged and calculated by the primary and auxiliary features of the part. The part shape feature set similarity formula defined by this system is:

$$Sim_x = 0.7 Sim_{zx} + 0.3 Sim_{fx} \quad (5)$$

In the formula, $0.7 Sim_{zx}$ represents the similarity coefficient of the primary feature of the current part and the case part. $0.3 Sim_{fx}$ indicates the similarity coefficient of the auxiliary features.

If the main feature set of the target part is O , the primary feature set of the case part is P , and $O=[o_1, o_2, \dots, o_n]$, $P=[p_1, p_2, \dots, p_m]$, $A=O \cap P$, then the primary feature similarity of the two parts is:

$$Sim_{zx} = \frac{2Q_{ZA}}{Q_{ZO} + Q_{ZP}} \quad (6)$$

In the formula, Q_{ZO} represents the number of major features of the target part, Q_{ZP} represents the number of major features of the case part, and Q_{ZA} is the number of intersections of the two. Similarly, the formula for calculating auxiliary features Sim_{fx} can be obtained, which will not be repeated here.

4. Establishment of Attribute Weight Coefficient Calculation Model

The size of the weight coefficient reflects the importance of each feature attribute of the case, which has a direct impact on the result of CBR matching. The Analytic Hierarchy Process (AHP) is a decision-making method that decomposes complex problems into multiple levels and conducts qualitative and quantitative analysis. This method evaluates the importance of attributes by constructing a judgment matrix, which provides a selection basis for decision-making. However, the calculation is more complex due to the need for consistency testing in traditional AHP. Therefore, it can improve the objectivity of judgment and omit consistency testing by introducing optimal transfer matrix to improve traditional AHP. The calculation steps to improve the AHP are as follows:

4.1. Constructing the Judgment Matrix

The judgment matrix is the core of the entire operation and directly affects the evaluation results[7]. The questionnaire was administered to P experts and the questionnaire data were synthesized into a judgment matrix A , by judging the importance of the guidelines according to the events as in Table 1.

$$A_{ij} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} \quad (7)$$

$$a_{ij} = \left(\prod_{r=1}^p a_{ij}^r \right)^{\frac{1}{p}} \quad (8)$$

Table 1. The importance of the event judgment criterion

The importance of attribute i over attribute j	Pairwise comparison standard
extremely important	9
more important	7
significantly important	5
slightly more important	3
equally important	1
slightly less important	1/3
significantly less important	1/5
Strongly unimportant	1/7
Extremely unimportant	1/9
The importance level falls between the two	8,6,4,2,1/2,1/4,1/6,1/8

In the formula, a_{ij} indicates the results of the fusion of multiple expert opinions, and $a_{ij} = 1/a_{ji}$

a_{ij}^r is the r -th expert to determine the importance value of the i -th attribute relative to the j -th attribute. P is the number of experts and n is the number of case attributes.

4.2. Optimization Consistency Matrix

In the traditional hierarchical analysis method, the consistency index CI , the average random consistency index RI and the consistency ratio CR must be used to test the consistency of the matrix, in order to satisfy the coordination between the elements of the judgment matrix and avoid the inconsistency of the importance between the elements that affect the decision results[8],[9].

However, with the optimization of the calculation method, the consistency check process can be replaced by constructing a consistency matrix. that is, the judgment matrix A belongs to the antisymmetric matrix, $B = \lg A$; If matrix C is an optimal transfer matrix of A , then $A^* = 10^C$ is the quasi optimal consistent matrix of judgment matrix A . Due to the consistency of this matrix, the traditional consistency verification process is omitted.

$$B = [b_{ij}]_{n \times n} = \lg A_{ij} \quad (9)$$

$$C = [c_{ij}]_{n \times n} = \frac{1}{n} \sum_{k=1}^n (b_{ik} - b_{jk}) \quad (10)$$

$$A^* = 10^C \quad (11)$$

4.3. Determination of Weighting Coefficients

The methods to calculate the weights include sum method, product method and eigenvalue method, and this paper adopts the sum method to solve for the weight coefficients. The specific steps are as follows.

Firstly, normalize each column element of the quasi-optimal uniform matrix, denoted as: $D = [d_{ij}]_{n \times n}$

$$D = \frac{a_{ij}^*}{\sum_{i=1}^n a_{ij}^*} \quad (12)$$

Then, sum the elements of matrix D in rows, denoted as: vector group: $C = [c_1, c_2, \dots, c_n]^T$

$$c_i = \sum_{j=1}^n d_{ij}, i = (1, 2, \dots, n) \quad (13)$$

Finally, the column elements in the vector group C are each divided by n to obtain the weight vector W_i .

$$W_i = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}^*}{\sum_{i=1}^n a_{ij}^*} \quad (14)$$

$W_i = [w_1, w_2, \dots, w_n]$, $i = (1, 2, \dots, n)$, W are the weight coefficients of each attribute.

4.4. Implementation of Attribute Weight Coefficients

Based on the model structure of the improved analytic hierarchy method described above, the weight coefficient of case attributes is calculated by using MATLAB software. Taking the decision matrix given by four experts as an example, according to Equations (7) and (8), the calculated similarity matrix A is as follows.

5. Application Cases

$$A = \begin{bmatrix} 1 & 1 & 0.4518 & 0.4083 & 0.2686 & 0.2272 \\ 1 & 1 & 0.4083 & 0.4083 & 0.3102 & 0.2272 \\ 2.2134 & 2.4495 & 1 & 1 & 0.4518 & 0.25 \\ 2 & 2 & 1 & 1 & 0.4518 & 0.2887 \\ 3.7224 & 3.2237 & 2.2134 & 2.2134 & 1 & 0.4518 \\ 4.4006 & 4.4006 & 4 & 3.4641 & 2.2134 & 1 \end{bmatrix}$$

According to the weight calculation formulas (9) to (14) of the improved AHP method, combined with MATLAB programming, the weight coefficients of the case properties can be obtained: shape feature set (0.3887), basic parameter set (0.2316), heat treatment method (0.1287), material (0.1236), roughcast (0.0639) and part category (0.0635). $W_i = [0.3887 \ 0.2316 \ 0.1287 \ 0.1236 \ 0.0639 \ 0.0635]$

This system is mainly based on the secondary development of SolidWorks API and Visual Studio software. By using feature recognition technology and CBR inference technology, it can meet the rapid design of the process. Finally, an intelligent process design system based on CBR is developed, which realizes the information extraction of 3D process models and intelligent matching of process cases. The overall operation interface of the system is shown in Figure 4 and 5.

Taking the design of the roller shaft as an example, the similarity calculation model and weight calculation method are applied to this design scheme. Table 2 shows the contents of the case library of the roller shaft, and the attribute types include text, numeric, and feature types, so it can be solved using the global similarity model. By storing mature cases in the database, the numbers C_1-C_5 are partial cases, $C = \{c_1, c_2, \dots, c_5\}$, and the case characteristic properties are described by $U = \{u_1, u_2, \dots, u_{10}\}$

Table 2. Historical Case Feature Attribute Values

Cases	Part category	roughcast	material	heat treatment	Shaft length	Maximum outer diameter	roughness	Machining accuracy	Number of main features	number of auxiliary features
C ₁	shaft	Bar stock	45	tempering	2350	140	0.8	7	13	11
C ₂	shaft	Bar stock	40cr	quench	100	50	0.4	6	8	6
C ₃	shaft	Bar stock	45	tempering	2258	146	3.2	9	11	12
C ₄	shaft	Bar stock	35	anneal	4300	450	6.3	5	18	12
C ₅	shaft	Bar stock	45	anneal	2105	160	3.2	7	10	9

(1) Target part information extraction and output. By the function interfaces of Feature, GetFirstAnnotation2 and GetCustomProperty, the feature traversal and information extraction of the eight process parameters of the target part are carried out in the SolidWorks software. That is, click the "Information Extraction" button, the system extracts the attribute information of the category, roughcast, material, heat treatment, shaft length, maximum outer diameter, roughness, machining accuracy, number of main features, and number of auxiliary features of the target part. The information is automatically output to the system application interface, as shown in Figure 4, for the next intelligent process retrieval by the designer.

(2) Retrieval and matching of similar process cases. Based

on the similarity calculation model and attribute weight calculation method studied above, the similarity calculation between the extracted target case attribute information and the case in the database is calculated. By clicking the "Retrieval" button and setting a threshold, match the cases that meet the requirements and sort them by overall similarity, as shown in Figure 5.

(3) Modification of the case and output of the process card. From the system interface, it can be seen that the most similar process case (97.18%) is reused. After modification and adjustment, the final output is a process card that matches the current target part. This enables the reuse of similar process cases and the rapid design of part processes.

Figure 4. Information extraction of the target part

	Overall similarity	Part category similarity	roughcast similarity	Material similarity	Heat treatment similarity	Shaft length similarity	Maximum outer diameter
▶	97.18%	100%	100%	100%	100%	99.48%	86.44%
	92.86%	100%	100%	100%	100%	98.33%	97.00%
	87.87%	100%	100%	100%	0	94.71%	99.50%

retrieval criteria

Lower limit on similarity

The number of similar cases

Figure 5. The cases retrieval results

6. Conclusion

This paper studies the problems of mechanical parts that cannot effectively use enterprise resources and insufficient rapid response ability to the market when designing the process. Taking the roller shaft of belt conveyor as an example, the description and expression of different attribute information in shaft process cases are given, and a global similarity calculation model is established. By combining CBR technology and feature recognition technology, an intelligent process design system was developed. The feasibility of applying this method to the process system is verified by case studies, and the retrieval and reuse of cases are realized. Thus, the process design efficiency of shaft parts is greatly improved, and it provides a reference for the research of intelligent process design of mechanical parts.

References

- [1] Ren Gongchang, Tang Yali. Green products innovative design process model based on CBR[J]. Machinery Design & Manufacture, 2019, 341(07):254-257.
- [2] Chen Jianjun, Ge Ruhai, LIU Defang, et al. Applications of case retrieval based on improved AHP and grey relational analysis [J]. Manufacturing Automation, 2012, 34(11):85-88.
- [3] Jiang Zhansi, Chen Liping, Luo Nianmeng. Similarity analysis in nearest-neighbor case retrieval[J]. Computer Integrated Manufacturing Systems, 2007, (06):1165-1168.
- [4] Liu Yu, Ben Kerong. A new approach based on similarity rough set for determining case feature weights[J]. Information and Control, 2012, 18(6):1230-1235.
- [5] Yang Baohua, Gu Lichuan, Li Shaowen. Improved algorithm for case feature weight based on sensitivity analysis [J]. Computer Science, 2010, 37(5):194-196.
- [6] Liu Fengshan. Approximate nearest neighbor search algorithms and their application [D]. Nanjing: Nanjing University, 2021.
- [7] Zhang Chi. Index weight scheme design based on improved analytic hierarchy process[J]. New Technology & New Products of China, 2022, 472(18):139-141.
- [8] Li Bo, Huang Xinjing. Application of matlab software algorithm in analytic hierarchy process[J]. Techniques of Automation and Applications, 2018, 37(12):35-38.
- [9] Ren Shengbing, Feng Di, Chen Xiaonan. Grey analytic hierarchy process using optimal transfer matrix[J]. Computer Engineering and Applications, 2017, 53(18):44-50.