Journal of Network Intelligence
Taiwan Ubiquitous Information

Volume **9**, Number **4**, November **2024**

Optimization of Product Design Structure Based on Improved SOM Clustering

Jing Liu*, Hui Wei

College of Arts and Humanities
Guangxi Vocational University of Agriculture, Nanning 530000, P. R. China
Lj19830105@163.com, Vicky01220123@163.com

Allen Guo

College of Information Science
Mapúa University, Intramuros, 1002 Metro Manila, Philippines
gxnz2008@163.com

*Corresponding author: Jing Liu
Received December 16, 2023, revised March 20, 2024, accepted June 16, 2024.

Abstract. *Product design structure clustering hierarchical delineation is an important part of product module design, which can help the team better understand and organise the structure of the product and improve the development efficiency and quality. However, while existing methods are able to achieve clustering hierarchical delineation of product design structure, they are unable to optimise the digital model. Therefore, this work investigates the intelligent algorithm Self-Organising Feature Mapping (SOM) algorithm and improves it to achieve clustering hierarchical division of product design structures. Firstly, the network structure and learning mode of SOM are analysed to illustrate the advantages of the application of SOM in clustering hierarchical division of product design structures. Secondly, a clustering model for product design structure optimisation based on Design Structure Matrix (DSM) is constructed, including the representation of DSM and optimisation objectives. Then, in order to make up for the defects of SOM network that the convergence time is too long and the accurate clustering information cannot be obtained, a two-stage clustering combination algorithm that combines SOM and K-means is proposed. On the basis of DSM, a product design structure optimisation model based on improved SOM clustering is proposed. Finally, the product design example was tested by a certain model of household air humidifier. The results show that the optimisation model based on SOM-K-means clustering has a better structural clustering hierarchical effect.*
Keywords: Product design; Intelligent algorithms; SOM; DSM; K-means

---

1. **Introduction.** product design structure clustering Hierarchical segmentation is an important part of product module design [1, 2]. In the product design process, the use of hierarchical partitioning can help the team to clarify the organisational structure of the product and the relationship between modules for better design and development.

Product design structure clustering hierarchical division [3, 4] is to group and divide the modules of a product according to certain principles and criteria to form a hierarchical structure. This division can be based on factors such as function, performance, user needs, and technical requirements. Through hierarchical division, a complex product structure can be disassembled into more manageable and implementable modules, each of which is responsible for a specific function or task [5].

In hierarchical division, a top-down approach is usually adopted [6, 7], where the whole product is first divided into several high-level modules and then gradually subdivided into smaller sub-modules until reaching the lowest level of specific functional modules. This division allows team members to better understand the structure of the product and the relationship between individual modules [8], which helps in the division of labour and collaboration of tasks. Through the product design structure clustering hierarchical division, the modules of the product can be effectively managed and controlled to improve the development efficiency and quality. At the same time, this hierarchical division also helps in later maintenance and updating because each module can be developed and tested independently, reducing the scope of influence on the whole product.

In the process of product design, complex products are often decomposed into a number of functionally independent small-scale clustering modules, so that the design can be carried out relatively independently. In addition, we need to determine the hierarchical relationship between the cluster modules, and then integrate them after the design of each cluster module to complete the overall design of the product. Clustering hierarchy can reduce the complexity of the design process [9], and can improve the design efficiency and shorten the product design cycle through parallel design. In conclusion, product design structure clustering hierarchical division is an important part of the product module design process, which can help the team better understand and organise the structure of the product and improve the development efficiency and quality. However, although existing methods can achieve clustering hierarchical delineation of product design structure, they cannot optimise the digital model. Therefore, this work investigates the Intelligent Algorithm Self-Organising Feature Mapping (SOM) algorithm [10, 11] and improves it to achieve clustered hierarchical delineation of product design structures.

1.1. **Related Work.** Clustering can help design teams understand the commonalities and differences in product design structures, optimise the design process and improve efficiency. By clustering product design structures, manufacturing processes can be optimised and adjusted in a more targeted way to improve productivity.

Cardone et al. [12] proposed a similarity measure based on feature vectors to assess the similarity of products for effective product redesign. Hsiao and Liu [13] proposed a product structure modelling and clustering method based on component flow analysis, which exploits the flow relationship between components to model and cluster the product structure to better understand the components and interactions of products. Jeong et al. [14] proposed a new product structure clustering model to support the modular design of composite product structures. The method adopts modular design theory and QAP solution method to perform cluster analysis of product structure.

Qiao et al. [15] applied the concept of green design in the process of product design, and put forward the micro-theoretical analysis method to model the product parts in the design elements, and obtained the product clustering module. According to the relationship among design requirements, production relations, characteristic variables and parameter factors, this method establishes a fuzzy conceptual model, and then carries out clustering division to analyze the design parameters in the product design structure to meet customer needs. Wen et al. [16] put forward the DSM model of design structure matrix, which can automatically divide the hierarchy. By clustering the design elements, the cluster of design elements with hierarchical relationship can be automatically obtained. When analyzing the product design structure, the multi-domain mapping of product function and product structure is considered. Then, the tree graph network is used to analyze the relationship between design elements, and a dynamic model of product design structure clustering is established for collaborative design and development. In order to reduce the number of rework iterations caused by design changes, Yang et al. [17] proposed using DSM to cluster design elements, and proposed a multi-coupling clustering model to analyze the aggregation degree of complex products, and obtained the coupling clustering relationship, which improved the design efficiency of complex products and the aggregation degree of decomposed design models.

Overall, these papers demonstrate the effectiveness of various clustering methods for product design structure analysis and optimization. The proposed methods can effectively partition the product architecture into reasonable modules, identify the functional combination-oriented modules, and optimise the product design structure. proposed methods can effectively partition the product architecture into reasonable modules, identify the functional combination-oriented modules of a product, and cluster Kansei needs adjectives in product emotional design. The proposed methods can effectively partition the product architecture into reasonable modules, identify the functional combination-oriented modules of a product, and cluster Kansei needs adjectives in product emotional design. The research on product design structure clustering is an active area, and researchers are exploring new methods and applications to improve the efficiency and effectiveness of product design.

SOM can automatically identify and cluster data with similar characteristics without providing category information in advance. In product design, when faced with a large amount of data and modules, SOM can automatically perform clustering based on the similarity of the modules and place similar modules in neighbouring positions, thus forming a hierarchical structure. Younus et al. [18] proposed a measurement model of product design attributes based on SOM. By calculating the similarity between the design parameters and the weight vector of SOM neurons, the model evaluates whether the design scheme meets the specific attributes. The research shows that this method can effectively evaluate whether the design scheme is consistent with the expected design attributes. Vahidnia et al. [19] uses SOM to process the design graphic data of different products to support the design decision-making process of product development. Through SOM visualization technology, the research results show that designers can analyze the design topology information more intuitively and effectively, and provide useful support for product development.

Both of these documents apply SOM neural network technology to support the product development process from different angles, especially in the design scheme and design topology information processing. These studies provide a typical example for the practical application of SOM in engineering design and product development. Compared with traditional methods, SOM neural network technology can process high-dimensional design data more efficiently and automatically, and provide intelligent support for complex product development.

1.2. **Motivation and contribution.** SOM can map complex multidimensional data into a low-dimensional space and help design teams understand the relationships and patterns between the data by visualising them on a 2D or 3D grid. This enables designers to better identify different clusters and patterns and to segment product structures based on this information. This work improves the SOM and implements a product design structure clustering hierarchy based on the improved SOM. The main innovations and contributions of this work include:

(1) A clustering model for product design structure optimisation based on DSM [20, 21] is proposed. Through the optimisation algorithm and division method, the interaction dependence and coupling between different clusters are minimised as much as possible, and the mutual constraints between clusters are reduced. The hierarchical relationship

between the clustering blocks is divided as much as possible into the upper level clustering module on the lower level clustering module influence, and there is a design time sequence in the product design process to prevent the number of rework iterations.

(2) In order to make up for the defects of SOM network that the convergence time is too long and the accurate clustering information cannot be obtained, a two-stage clustering combination algorithm that combines SOM and K-means is proposed. On the basis of DSM, a product design structure optimisation model based on improved SOM clustering is proposed.

2. Self-Organising Map (SOM).

2.1. Network structure. Scientists have obtained the biological basis of the competitive mechanism in the SOM network through research in biology and human brain science: the orderly arrangement of neurons and the continuous mapping of information to external stimuli.SOM is an unsupervised competitive neural network learning algorithm.

The neurons on the competing layers of the SOM network will be arranged in an orderly fashion, with neurons with similar roles close to each other and neurons with very different roles far away from each other [22].The competing layers of the SOM network are also divided into different regions, which produce different responses to the same input samples, and this process is not interfered with by human beings, but is processed by the algorithm on its own.The structure of the SOM neural network As shown in Figure 1, it consists of an input layer and an output layer. The input layer consists of input neurons and the output layer, also known as the competition layer, consists of output neurons [23].

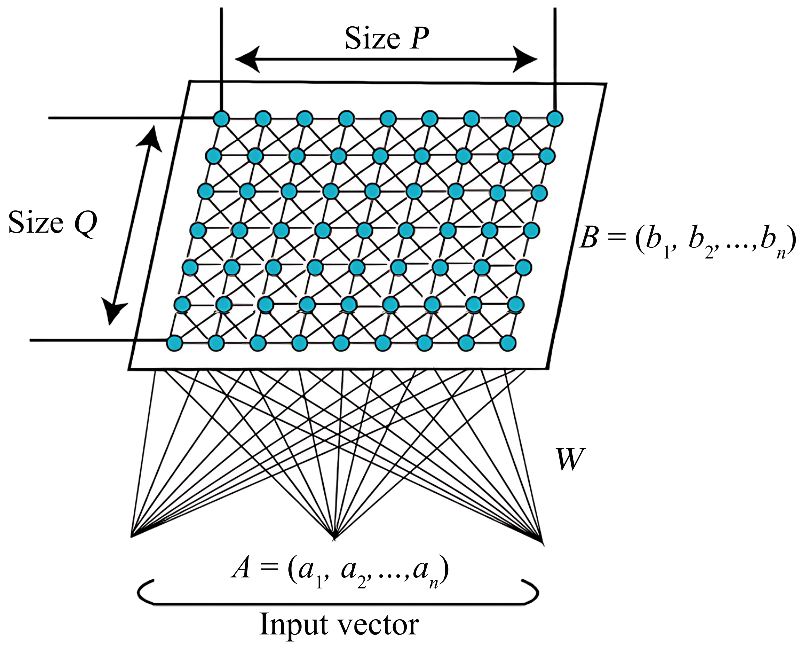


Figure 1. Structure of the SOM neural network

$A = (a_1, a_2, \ldots, a_n)$ is the n input components of the input layer, and $B = (b_1, b_2, \ldots, b_n)$ is the classification corresponding to the input samples. w_{ij} is the connection weight between the i-th input unit and the j-th output unit. Assuming X is the input sample of the SOM network and w is the connection weight between j and X, the output of output unit j is:

$$y_j = W_j X \tag{1}$$

And the output unit k of the actual response of the network obtained through the competitive mechanism is:

$$y_k = \max\{y_j\} \tag{2}$$

With appropriate corrections, more accurate Equation (1) and Equation (2) can be obtained.

$$y_j = \alpha\left[\sum_{i=1}^{m} w_{ij}x_i + \sum_{t\in R_l} g_k y_l\right] \tag{3}$$

$$y_k = \max_j \{y_j\} - \varepsilon \tag{4}$$

$$\alpha(t) = \begin{cases} 0 & t < 0 \\ \alpha(t) & 0 \le t \le A \\ A & t > A \end{cases} \tag{5}$$

where x_i is the output of input cell i, ϵ is a very small positive number, and g_k is a coefficient associated with the weights of the lateral connections.

2.2. **Learning approach of SOM.** The SOM algorithm automatically finds out the similarity between the input data and configures similar input data close to each other on the network. The main ideas and principles of SOM are as follows [24]:

(1) Construct an intermediate layer connecting the input and output layers, whose neurons are arranged in a two-dimensional matrix.

(2) When randomly sampling the input vectors, the neurons of the intermediate layer compete with each other, and the one that encodes the most similarity (e.g., the smallest Euclidean distance) to the input vectors is activated. Calculate the Euclidean metric between X and the connection weight vector of each neuron in the competing layer, where the distance between the j-th neuron and X is:

$$l_j = \|X - W_j\| = \sqrt{\sum_{i=1}^{m} (x_i(t) - w_{ij}(t))^2} \tag{6}$$

(3) Activate not only the mediator neuron, but also its nearby neurons through the "neighbourhood function", so that they have similar coding. As the learning process proceeds, the neighbourhood range decreases. By calculating and selecting the minimum distance l_j, a winner neuron j' is generated in the competitive layer of the SOM network. For any competitive layer neuron j, the unit k can be determined such that there is $l_k = \min(l_j)$, up to which the set of neighbouring neurons can be obtained.

(4) By iteratively learning the set of input vectors, the parameters of the neural network are gradually adjusted, and the topological positions of the neurons in the mediator layer reflect the similarity between different encoded vectors. The connection weights between neuron j' and its neighbouring neurons are updated.

$$x_i = \eta h(j, j')x_i w_{ij} \tag{7}$$

where $0 < \eta < 1$ and $h(j, j')$ is a valid domain mapping function.

$$h(j, j') = \exp\left(-\frac{|j - j'|^2}{\sigma^2}\right) \tag{8}$$

(5) After the final network convergence, the similar input samples are mapped to the neighbouring intermediate layer neurons, completing the unsupervised dimensionality reduction and feature extraction of the input feature space to the intermediate neuron matrix. Calculate the network output y_k as follows:

$$y_k = \alpha \left(\min_j \|X - W_j\| \right) \tag{9}$$

SOM can quickly generate prototypes and models and support rapid iteration and feedback. The design team can quickly modify and iterate based on the results of the SOM to further optimise the product design and clustering hierarchy. This iteration and feedback loop ensures that the product design is more accurate and aligned with user needs.

3. DSM-based clustering model for product design structure optimisation.

3.1. **Design structure matrix.** A product is a collection of several design units such as functional components, sub-functional components, assembly accessories, individual parts, etc., combined according to certain rules, which can be used as a system to show the characteristics and behaviours that the whole product has. In 2000, two famous professors of MIT, Ulrich and Eppinger, described product design structure as the planning of product functions to product parts based on the clustering module division technique and the interaction requirements between clustering modules. Product design structure is used to analyse the relationship between the constituent elements and the constituent elements, and the overall function of the product is accomplished by the joint action of these two to achieve. The product design structure is the organisational framework for the product design process and describes the design attributes.

Product design is a complex dynamic system. This system contains many interrelated design elements, such as design requirements, design constraints, design goals, process variables and design parameters. The interrelationships between these elements constitute the structural framework of product design. In order to effectively control and optimise such a complex system, it is crucial to establish a scientific product design structural model. The first step in building a structural model of product design is to correctly describe the design elements of the system and their interrelationships. This requires attention to the mutual constraints, mutual influences and interdependence between the elements. Clarifying the internal logic of these complex relationships is the basis for model building. A scientific and reasonable product design structure model will bring many benefits: (1) improve the parallelism of the design process, which helps to optimise the overall structure of the product; (2) enable the designer to consider the interactions between the elements more clearly, so as to improve the efficiency of the design and product quality; (3) provide a basis for the evaluation and improvement of the design, reduce the cost of the product, and shorten the design cycle.

The product design structural model considers the product as a whole, including its components and their interrelationships. It focuses not only on the performance of individual components, but also on the interface and matching between components. Product design structural models usually adopt a multi-level structure to represent the structural characteristics of complex products. From the whole to the local, from abstract to concrete step by step decomposition, each level describes one aspect of the product structure. The product design structure model integrates the theories and methods of many disciplines, including systems engineering, control theory, optimisation theory, etc. They are used to analyse and optimise the product structure. They provide a systematic analysis tool for analysing and optimising the product structure. The three basic types of product design structure models are shown in Figure 2.

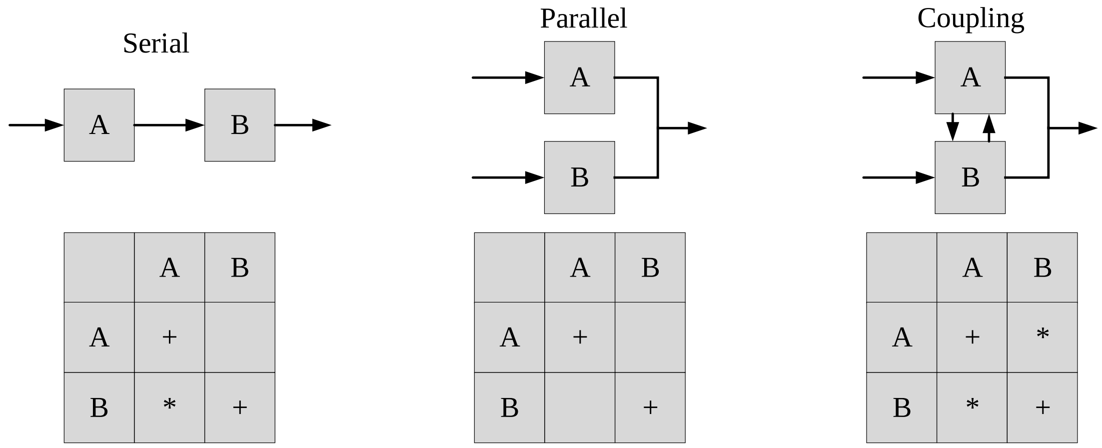


Figure 2. Three basic types of structural modelling for product design

3.2. **Representation of product design structure.** A product design structure model can represent the interdependencies, clustering and hierarchical relationships between product design elements and design elements. In general, the interrelationships between product design elements can be expressed and transformed using tree diagrams, directed graphs, and matrices.

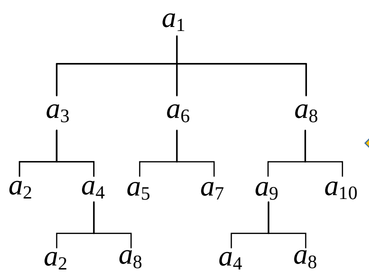
(1) Tree diagrams.

Tree representation of product design structure is a way to show the hierarchical relationship of product structure intuitively and effectively. The top of the tree structure is the conceptual node of the whole product, which is the starting point of the product design. In other words, the root node is the starting point of the whole product structure. The second level is the decomposition of the whole product into several major component sub-systems, each sub-system completes an independent function. Starting from the third layer, the tree structure expands down to each subsystem, and each subsystem is further decomposed into its component parts. Layers are broken down until they reach the basic components and key parts, which can no longer be meaningfully broken down. The tree structure also requires lines to identify the interfaces between the nodes at each level, highlighting the information and energy flow connections between components.

The product design structure tree can be interchanged with the matrix, as shown in Figure 3. Through the combination of tree level-by-level decomposition and interface expression, the product structure hierarchy from the whole to the local level, as well as the interdependence and constraint relationship between components can be clearly expressed.

In Figure 3, a_1 denotes the product, and the rest denotes the clustering module (component) or part. When performing the matrix transformation, the upper level has an influence on the lower level, indicating a hierarchical relationship. In the corresponding matrix, a connection between clustering modules is represented by the number 1, and no connection is represented by 0. Note that 0 has been omitted here.

3.3. **Directed graphs.** The directed graph expression clearly shows the interdependence between the components in the product structure and uses arrows to identify the transfer paths of the directionality of the information flow and energy flow. This approach helps to determine the critical links of the product structure through control path analysis,



$$
\begin{array}{c}
\\
a_1\\ a_2\\ a_3\\ a_4\\ a_5\\ a_6\\ a_7\\ a_8\\ a_9\\ a_{10}
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} \end{array}\\
\left[\begin{array}{cccccccccc}
 & & & & & & & & & \\
 & & 1 & 1 & & & & & & \\
1 & & & & & & & & & \\
 & & 1 & & & & & & 1 & \\
 & & & & & 1 & & & & \\
1 & & & & & & & & & \\
 & & & & & 1 & & & & \\
1 & & & 1 & & & & & 1 & \\
 & & & & & & & 1 & & \\
 & & & & & & & 1 & &
\end{array}\right]
\end{array}
$$

Figure 3. Product structure tree and matrix transformation

which plays an important role in improving the coordination of the product and reducing system risk.

Suppose the directed graph is $G = \langle V, E\rangle$, where $V = \{V_1, V_2, \ldots, V_3\}$ denotes the set of design elements, and $E = \{e_1, e_2, \ldots, e_n\}$ denotes the set of relationships between design elements. An example of a directed graphical representation of a product design structure is shown in Figure 4. The numbers represent the design elements, and the directed line segments represent the relationships that exist between the design elements.

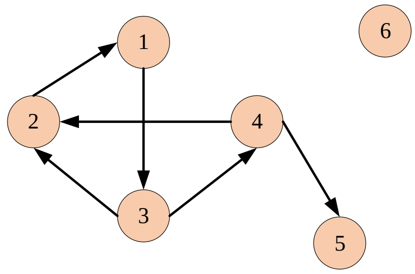


Figure 4. Example of a directed graph of a product design structure.

(3) Matrix representation.

Matrix expression quantifies the complexity and coupling of the interfaces between components in a product structure by establishing a mapping matrix of the characteristics of the interfaces between components, identifying the most complex and critical interfaces in the structure, and providing a quantitative evaluation and analysis method for optimising the interfaces and improving the compatibility and maintainability of the structure. The matrix representation of product design structure mainly includes adjacency matrix, reachability matrix and DSM.

Tree diagrams, directed graphs and matrix representations are three ways of representing product design structures, which essentially describe the structural characteristics of the same design system, but show different characteristics of the design structure from different perspectives, and there is a close correspondence between them. Tree diagrams, directed graphs and matrix representations can refer to each other to increase the systematicness and comprehensiveness of structural analyses.

3.4. **DSM-based optimisation objective.** Compared with the adjacency matrix and reachability matrix, DSM not only reflects whether there are interdependencies between components, but also pays more attention to the direction and strength of such dependencies [25]. DSM matrix is easy to be updated and iterated, and can dynamically reflect the changes of component relationships. Therefore, DSM is chosen to represent the product design structure in this work.

A DSM matrix is a square matrix consisting of a series of row and column elements arranged in the same order. The row and column elements in the matrix represent the design elements in the design process [26]. The non-diagonal cells in the matrix denote the link between the design elements and the numbers in the cells denote the linkage weights. The DMM matrix is an $m \times n$ matrix of different elements which is a matching mapping of two different DSM elements where m denotes the number of elements in the first DSM matrix and n denotes the number of elements in the second DSM matrix.The DSM matrix schematic is shown in Figure 5.

	A	B	C	D	E	F
A				1		
B						
C						1
D	1					
E		1				
F						

Figure 5. Schematic diagram of the DSM matrix

The optimisation objectives of the DSM-based product design structure clustering model include three main points:

(1) Minimise the interaction complexity between clusters. By optimising the algorithm and division method, the interactive dependence and coupling relationship between different clusters can be minimised as much as possible, and the mutual constraints between clusters can be reduced.

(2) Maximise internal consistency of clusters. As far as possible, components with high functional relevance and frequent interactions are put together in a cluster, so that the consistency of interactions and cooperation within the cluster can be maximised.

(3) Optimise the level of modularity. The hierarchical relationship between the clustering blocks should divide the upper-level clustering module from the lower-level clustering module influence. In the product design process, there is a design of the time sequentiality to prevent the number of rework iterations.

4. A structural optimisation model for product design based on improved SOM clustering.

4.1. **SOM-K-means.** The SOM algorithm is an unsupervised learning model that maps data from a high-dimensional space onto a low-dimensional space, identifies the main statistical features of multidimensional data through dimensionality reduction, and automatically divides the data into different classes based on the similarity between the data and obtains the number of clusters and the centre of clusters.

However, the SOM algorithm cannot provide specific clustering information after classification. Moreover, during network training, the network convergence time of the SOM algorithm is too long, i.e., the clustering efficiency is low. The k-means algorithm has the characteristics of high efficiency and accuracy, high scalability, and can provide the exact clustering information after classification [27, 28]. Based on the analysis of the advantages and disadvantages of the two algorithms, this paper proposes a two-stage clustering combination algorithm that combines SOM and k-means, whose basic idea is to obtain the initial number of clusters and various types of centroids according to the advantages of automatic clustering of the SOM algorithm, and then use this as the initial input to the k-means algorithm for further clustering, so as to obtain the accurate clustering information.

This combined clustering algorithm combines the self-organisation characteristics of SOM network and the advantages of K-means algorithm which is efficient and accurate, and at the same time makes up for the defects of SOM network that the convergence time is too long and the accurate clustering information cannot be obtained. The execution process of this combined algorithm is divided into the following two stages:

Stage 1: Initial clustering by SOM to derive the number of clusters K and the centroids of each category.

Firstly, the SOM algorithm is executed, and the data objects to be clustered are input to the SOM network for training, the SOM network considers the feature vectors with similar features as belonging to the same class, and the sample data are thus clustered into different classes, and the number of clusters K and the centroid of each class are derived.

Stage 2: K-means uses the results of the first stage (number of clusters K and centroids of each category) as an initial value input and further clustering to form the final clustering results.

In summary, the flow of the proposed SOM-K-means algorithm is shown in Figure 6.

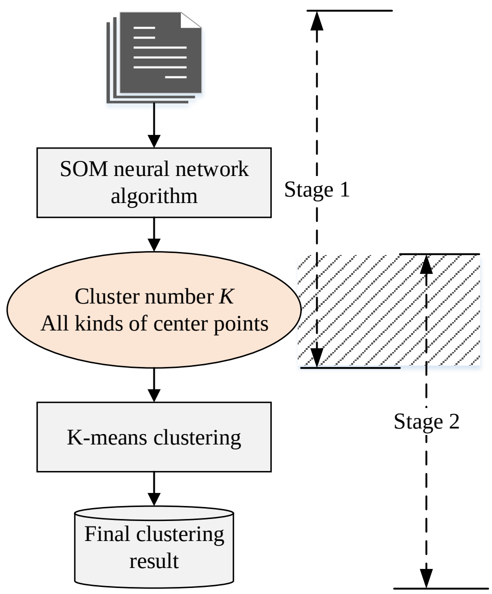


Figure 6. Flow of the SOM-K-means algorithm

The network plane of the SOM output consists of m neurons, and each output neuron has a connection weight with the input vector. There are three network structures of the SOM [29], which are rectangular layer structure, hexagonal layer structure and random layer structure. There are also various ways to calculate the distance between neurons, different cases are calculated based on different distance functions, in this work, the hexagonal layer structure is chosen to facilitate the improvement of convergence efficiency, as shown in Figure 7.

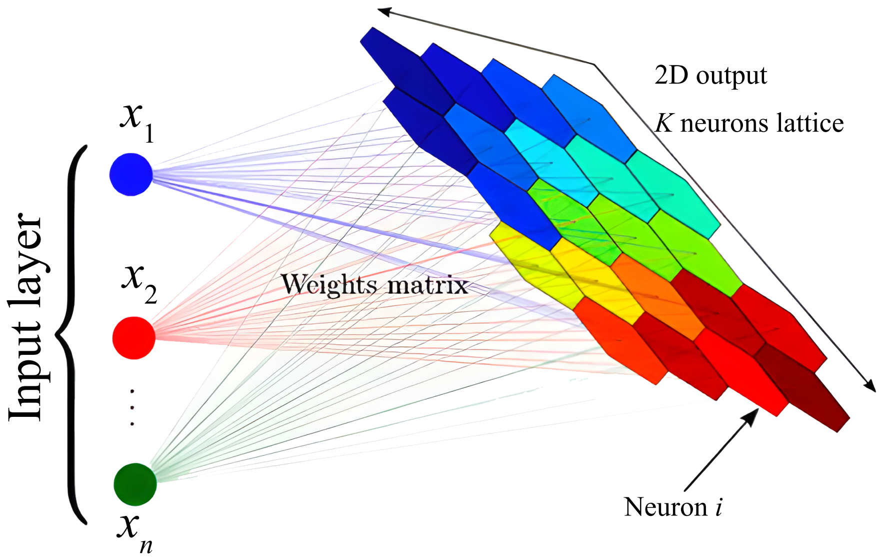


Figure 7. SOM Hexagonal Layer Structure

4.2. **Evaluation methods for clustering models.** When dividing the product design structure clustering module, it is required that the connection between design elements should be as large as possible within the clustering module and as few as possible between the clustering blocks. Therefore, two indicators are used in this paper to evaluate the effectiveness of the clustering model.

(1) Intra-cluster variance.

A major criterion for internal metrics is the error squared and minimum variance criterion, also known as intra-cluster variance [30]. By calculating the error (versus the distance from the cluster centre) for each data point, the sum of the variance errors is then derived.

$$V(C) = \sum_{c_i \in C_i} \delta(i, u_k)^2 \tag{10}$$

where C denotes all the clusters, u_k denotes the clustering centre of cluster C_k, and $\delta(\cdot)$ denotes the distance function. The minimum value of intra-cluster variance depends on the data and the number of clusters, and the smaller its value, the better the clustering result.

(2) Overall Cluster Quality (QCQ).

QCQ refers to a comprehensive evaluation method using a combination of cluster denseness and cluster proximity [?]. Cluster denseness is a measure of variance within the cluster in question. Cluster denseness is defined as follows:

$$cmp = \frac{1}{k} \frac{1}{\sum_{i=1}^{k}} \left[\frac{V(C_i)}{V(X)} \right] \tag{11}$$

where X denotes a given dataset and k is the number of clusters.

Cluster proximity is defined as follows:

$$prox = \frac{1}{k(k-1)} \sum_{i=1}^{k} \sum_{j=1, j \neq i}^{k} \exp\left[\frac{-d^2(X_{ci}, X_{cj})}{2\delta^2} \right] \tag{12}$$

where δ denotes the Gaussian constant, X_{ci} denotes the clustering centre, and $d(X_{ci}, X_{cj})$ is the distance between the clustering centres.

The combined evaluation value of the clusters is defined:

$$ocq(\epsilon) = 1 - [\epsilon \times cmp + (1 - \epsilon) \times prox] \tag{13}$$

where ϵ is an adjustment weight to balance cluster denseness and cluster proximity, generally $\epsilon = 0.5$ means that the two evaluation methods have equal weights. Obviously, the larger the integrated evaluation value of clustering, the better the smaller the intra-cluster variance and the larger the QCQ, the better the output of the product design structure optimisation clustering model.

4.3. **Input layer matrix for SOM-K-means based on DSM.** Modelling with DSM (0-1 matrix), when transformed into the input matrix, the linkage at the diagonal position is marked as 1. In the product design structure, the relationship between the design elements can be expressed as a numerical value to indicate the degree of linkage, with the larger value indicating the closer linkage between the design elements and the greater the mutual influence. For simple 2D data, SOM-K-means clustering division is applied to establish the input layer as follows:

$$\begin{bmatrix} v_1, v_2, \cdots v_i \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_i & y_i \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{14}$$

where i denotes the number of input data and $[x, y]$ denotes a two-dimensional data distribution.

In the product design process, the clustering division is to study the coupled modular relationship of design elements, and each design element can only be represented by one dimension of the SOM-K-means input matrix. The design structure matrix DSM model is established based on the relationship between design elements, n is the number of design elements, then the corresponding SOM-K-means input layer matrix has n dimensions. Where, in the DSM model, the value 1 indicates the relationship between the design elements and the design elements, and the number 0 (blank space, omitted here) indicates that there is no connection between the design elements.

For hierarchical delineation of the product design structure, the current training vector and the best weight matching weight vector are calculated.

$$j^* = \arg\min_j \|x_i(t) - w_j(t)\| \tag{15}$$

Adjusting the weight vector w_j of the winning neuron only, within the neighbourhood range of the clustering block $N_j(t)$, the adjusted weight vector is shown as follows:

$$w_j(t+1) = \begin{cases} w_j(t) + \eta(t)(x_i(t) - w_j(t)), & j \in N_j^*(t) \\ w_j(t), & j \notin N_j^*(t) \end{cases} \tag{16}$$

where $\eta(t)$ is the learning rate parameter.

5. **Experimental results and analyses.**

5.1. **Evaluation of clustering effect.** In order to verify the advantages of the proposed SOM-K-means clustering algorithm even further and to make its test results more fully supported by data, this paper uses the datasets IRIS, WINE and YEAST from the UCI (University of California in Irvine) machine learning repository for the test.

The IRIS dataset has 150 sample data of 4 feature dimensions, the WINE dataset has 178 sample data of 13 feature dimensions, and the YEAST dataset has 1484 sample data of 9 feature dimensions. The performance of the three clustering algorithms is verified using intra-cluster variance and OCQ, and the test results are shown in Table 1.

From the test results, it can be seen that the proposed two-stage clustering combination algorithm has lower intra-cluster variance and higher OCQ values than the single clustering method. Therefore, the effectiveness of the SOM-K-means combined clustering algorithm is verified.

Table 1. Data set test results.

	Intra-cluster variance			**OCQ**		
	IRIS	**WINE**	**YEAST**	**IRIS**	**WINE**	**YEAST**
SOM	670.502	2132.622	5818.142	0.592	0.613	0.599
K-means	626.312	1835.212	5610.302	0.584	0.59	0.592
SOM-K-means	454.352	958.802	1993.452	0.632	0.644	0.656

In order to further analyse the stability of each clustering algorithm, this paper each time from the total sample data randomly selected sample proportion of 10%, 30%, 50%, 70%, 90% of the samples were tested, each test repeated 10 times, and then take its average as the clustering results, plotted as a graph to compare its stability, the results are shown in Figure 8.

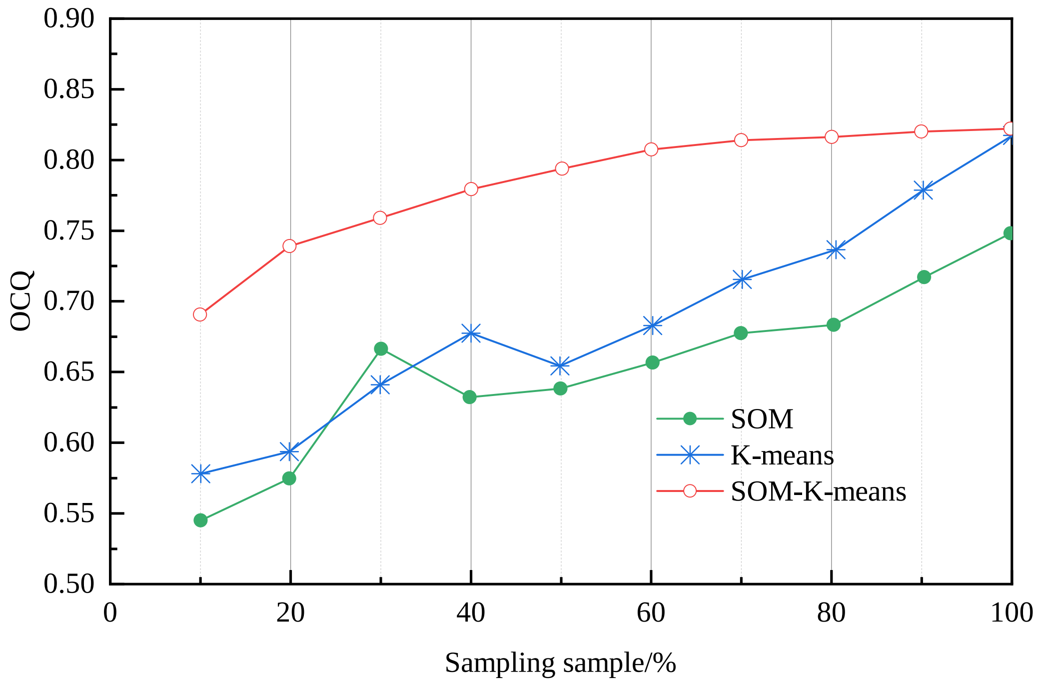


Figure 8. Stability curves for different clustering methods

It can be seen that the clustering accuracy of the three clustering algorithms is gradually increasing in the overall trend as the proportion of sampling samples increases.

Among them, the SOM-K-means two-stage clustering combination algorithm, its clustering accuracy is rising with the increase of the sample, there is no large fluctuation in the middle, with good stability; while the SOM and K-means single clustering algorithm there are large fluctuations, appearing to rise and then fall and rise again. This shows that the SOM-K-means combined clustering algorithm has better stability than the single clustering method.

5.2. **Product design example.** This example uses a certain model of household air humidifier to verify the feasibility of the SOM-K-means clustering method for product design structure clustering hierarchical division, according to the relationship between the car body parts, using the toolbox provided by MATLAB for its clustering division, and the application of MATLAB program to achieve the hierarchical division between clustered blocks. The calculation results of the SOM-K-means are shown in Figure 9.

	10	13	14	17	1	2	3	4	9	5	6	7	12	11	16	18	20	8	12	15	19
10		1																			
13			1	1											1						
14	1			1								1					1				
17		1	1																		
1						1			1							1				1	
2							1	1	1												
3					1									1							
4					1	1	1		1									1			
9					1	1		1													
5			1								1	1									
6										1		1	1								
7	1									1			1					1			
12										1	1										
11					1	1										1					
16														1		1	1				
18			1											1	1		1				
20										1				1		1					
8		1			1														1	1	1
12																		1		1	1
15											1							1	1		
19																		1		1	

Figure 9. SOM-K-means calculation results

The results of dividing each clustering module within a hierarchy are shown as follow:
Level 1: Cluster I (10, 13, 14, 17)
Level 2: Cluster II (1, 2, 3, 4, 9)
Level 3: Cluster III (5, 6, 7, 12)
Level 4: Cluster IV (11, 16, 18, 20)
Level 5: Cluster V (8, 12, 15, 19)

As can be seen from the clustering results of the 20 parts, the results obtained from the SOM-K-means calculation satisfy the expected requirements, and can themselves meet the design requirements by modifying the relevant parameters. The optimisation model based on SOM-K-means clustering has a better product design structure clustering hierarchical effect, which can both reduce the number of iterations of the design process and improve the design efficiency.

6. **Conclusion.** In this work, a product design structure optimisation model based on improved SOM clustering is proposed. The network structure and learning mode of SOM are analysed to illustrate the advantages of the application of SOM in the clustering hierarchical division of product design structure. Secondly, a product design structure optimisation clustering model based on DSM is constructed, including the representation of DSM and optimisation objectives. Then, in order to make up for the defects of SOM network that the convergence time is too long and the accurate clustering information cannot be obtained, a two-stage clustering combination algorithm that combines SOM and K-means is proposed. The test results of product design examples show that the optimisation model based on SOM-K-means clustering has a better effect of product design structure clustering hierarchy, which significantly improves the design efficiency. In addition, the SOM-K-means combined clustering algorithm has better stability than the single clustering method.

REFERENCES

[1] D. Tang, R. Zhu, J. Tang, R. Xu, and R. He, "Product design knowledge management based on design structure matrix," *Advanced Engineering Informatics*, vol. 24, no. 2, pp. 159-166, 2010.

[2] C.-H. Chu, Y.-P. Luh, T.-C. Li, and H. Chen, "Economical green product design based on simplified computer-aided product structure variation," *Computers in Industry*, vol. 60, no. 7, pp. 485-500, 2009.

[3] T. R. Browning, "Applying the design structure matrix to system decomposition and integration problems: a review and new directions," *IEEE Transactions on Engineering Management*, vol. 48, no. 3, pp. 292-306, 2001.

[4] A. Povilionis, and A. Bargelis, "Structural optimization in product design process," *Mechanics*, vol. 81, no. 1, pp. 66-70, 2010.

[5] T. AlGeddawy, and H. ElMaraghy, "Optimum granularity level of modular product design architecture," *CIRP Annals*, vol. 62, no. 1, pp. 151-154, 2013.

[6] R. Srinivasan, G. L. Lilien, A. Rangaswamy, G. M. Pingitore, and D. Seldin, "The total product design concept and an application to the auto market," *Journal of Product Innovation Management*, vol. 29, pp. 3-20, 2012.

[7] C.-M. Chen, S. Lv, J. Ning, and J. M.-T. Wu, "A Genetic Algorithm for the Waitable Time-Varying Multi-Depot Green Vehicle Routing Problem," *Symmetry*, vol. 15, no. 1, 124, 2023.

[8] L. Kang, R.-S. Chen, N. Xiong, Y.-C. Chen, Y.-X. Hu, and C.-M. Chen, "Selecting Hyper-Parameters of Gaussian Process Regression Based on Non-Inertial Particle Swarm Optimization in Internet of Things," *IEEE Access*, vol. 7, pp. 59504-59513, 2019.

[9] F. Zhang, T.-Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2020.

[10] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[11] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[12] A. Cardone, S. K. Gupta, and M. Karnik, "A survey of shape similarity assessment algorithms for product design and manufacturing applications," *Journal of Computing and Information Science in Engineering*, vol. 3, no. 2, pp. 109-118, 2003.

[13] S.-W. Hsiao, and E. Liu, "A structural component-based approach for designing product family," *Computers in Industry*, vol. 56, no. 1, pp. 13-28, 2005.

[14] S.-B. Jeong, Y. H. Shin, S. R. Koo, and H.-S. Yoon, "A market segmentation scheme based on customer information and QAP correlation between product networks," *Journal of the Korea Society for Simulation*, vol. 24, no. 4, pp. 97-106, 2015.

[15] L. Qiao, M. Efatmaneshnik, M. Ryan, and S. Shoval, "Product modular analysis with design structure matrix using a hybrid approach based on MDS and clustering," *Journal of Engineering Design*, vol. 28, no. 6, pp. 433-456, 2017.

[16] L. Weng, Y. Hu, and Y.-M. Deng, "Functional combination-oriented module identification for adaptable-function mechanical product design," *The International Journal of Advanced Manufacturing Technology*, vol. 116, no. 1-2, pp. 523-536, 2021.

[17] Y.-P. Yang, D.-K. Chen, R. Gu, Y.-F. Gu, and S.-H. Yu, "Consumers' Kansei needs clustering method for product emotional design based on numerical design structure matrix and genetic algorithms," *Computational Intelligence and Neuroscience*, vol. 2016, 5083213, 2016.

[18] Z. S. Younus, D. Mohamad, T. Saba, M. H. Alkawaz, A. Rehman, M. Al-Rodhaan, and A. Al-Dhelaan, "Content-based image retrieval using PSO and k-means clustering algorithm," *Arabian Journal of Geosciences*, vol. 8, pp. 6211-6224, 2015.

[19] S. Vahidnia, A. Abbasi, and H. A. Abbass, "Embedding-based detection and extraction of research topics from academic documents using deep clustering," *Journal of Data and Information Science*, vol. 6, no. 3, pp. 99-122, 2021.

[20] C. Liang, Z. Gu, Y. Zhang, Z. Ma, H. Qiu, and J. Gu, "Structural design strategies of polymer matrix composites for electromagnetic interference shielding: a review," *Nano-Micro Letters*, vol. 13, no. 1, 181, 2021.

[21] S. Clark, S. A. Sisson, and A. Sharma, "Tools for enhancing the application of self-organizing maps in water resources research and engineering," *Advances in Water Resources*, vol. 143, pp. 103676, 2020.

[22] P. Melin, J. C. Monica, D. Sanchez, and O. Castillo, "Analysis of spatial spread relationships of coronavirus (COVID-19) pandemic in the world using self organizing maps," *Chaos, Solitons & Fractals*, vol. 138, 109917, 2020.

[23] X. Qu, L. Yang, K. Guo, L. Ma, M. Sun, M. Ke, and M. Li, "A survey on the development of self-organizing maps for unsupervised intrusion detection," *Mobile Networks and Applications*, vol. 26, pp. 808-829, 2021.

[24] K. Sinha, S. Y. Han, and E. S. Suh, "Design structure matrix-based modularization approach for complex systems with multiple design constraints," *Systems Engineering*, vol. 23, no. 2, pp. 211-220, 2020.

[25] X. Zhang, S. Ma, and S. Chen, "Healthcare process modularization using design structure matrix," *Advanced Engineering Informatics*, vol. 39, pp. 320-330, 2019.

[26] K. P. Sinaga, and M.-S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716-80727, 2020.

[27] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Information Sciences*, vol. 622, pp. 178-210, 2023.

[28] Q. Xiang, H. Yu, H. Chu, M. Hu, T. Xu, X. Xu, and Z. He, "The potential ecological risk assessment of soil heavy metals using self-organizing map," *Science of the Total Environment*, vol. 843, pp. 156978, 2022.

[29] Z. Wu, W. Xue, H. Xu, D. Yan, H. Wang, and W. Qi, "Urban Flood Risk Assessment in Zhengzhou, China, based on a D-Number-Improved Analytic Hierarchy Process and a Self-Organizing Map Algorithm," *Remote Sensing*, vol. 14, no. 19, 4777, 2022.

[30] D. Nagar, P. Ramu, and K. Deb, "Visualization and analysis of Pareto-optimal fronts using interpretable self-organizing map (iSOM)," *Swarm and Evolutionary Computation*, vol. 76, 101202, 2023.