Multi-attribute Hierarchical Clustering for Product Family Division of Customized Wooden Doors

Na Zhang,[a] Wei Xu,[a,]* and Yong Tan [b]

To improve the production system for customized wooden doors and to gain research and development efficiency, this paper proposed the feasibility of using hierarchical clustering algorithms to cluster a company's customized wooden door products and its application to rational product family architecture. The particular use of multi-attribute feature data to locate products and the integration of image data into the database can make the original hierarchical clustering more compatible and adaptable for application to customized wooden doors. The preprocessed data was analyzed by clustering to obtain the clustering results and similarity relationships. Hierarchical clustering results were uneven and not entirely interpreted. However, the internal order structure of clusters and the clustering process could be clearly observed, and the distance hierarchical relationship between the products could be obtained, which was beneficial to the division of the product architecture. The results illustrated that processing using hierarchical clustering of multi-attribute data is feasible for optimizing customized wooden door product systems. In addition, the product architecture, product coding rules, and front-end development process were established to improve standardization and research and development efficiency. There is still great potential for developing the custom wooden door category in custom furniture companies.

DOI: 10.15376/biores.18.4.7889-7904

Keywords: Customized furniture; Product system optimization; Machine learning; Research and development

*Contact information: a: College of Furnishings and Industrial Design, Nanjing Forestry University, Nanjing 210037, China; b: Department of Product Research and Development, ZBOM Home Collection Co., Ltd., Hefei 230001, China; *Corresponding author: xuwei@njfu.edu.cn*

INTRODUCTION

Mass customization is a method of achieving personalization while reducing costs. This approach achieves standardization and modularity to reduce the internal diversity of a product. As an essential strategy for the manufacturing industry and a profitable business model, mass customization is widely adopted by furniture companies (Da Silveira *et al*. 2001; Lihra *et al*. 2012). China's mass-customized furniture market has been expanding in recent years, reaching $57.4 billion by 2021. Despite the severe environment, leading customized furniture companies still achieved solid growth, which was attributable to the mature customized cabinet and wardrobe categories (Xiong *et al*. 2022; Hu *et al*. 2023). Peng *et al*. (2019) built a library of functional modules for custom cabinet products based on consumers' daily behaviors and habits in the kitchen. Ren and Xu (2019) analyzed the existing standardized wardrobe parts platform and optimized the components platform for changing markets. Compared with custom cabinets and wardrobes, custom wooden doors have weak product systems. Customized wooden doors have chaotic product categories, insufficient functional diversity, low modularity and standardization, and complex research and development processes (Lv *et al*. 2008; Xu 2011; Zhu *et al*. 2013). Moreover, the mature development of custom cabinets and wardrobes has flattened the sales share trends from rapid growth. Therefore, customized wooden doors have received the attention of many enterprises, become an important direction in the development plan, and are expected to become the third pillar of the customized furniture enterprise (Feng and Wu 2020).

Product families are recognized as an effective means of mass customization based on product platforms with strong similarities and deriving different product families through slight differences (Jiao *et al*. 2007). Alizon *et al*. (2009) constructed a set of platforms that considered the inherent relationship between product architecture and processing activities. These platforms enabled similar components to be grouped and produced on a shared platform. Simpson and D'Souza (2004) proposed genetic algorithm modifications to resolve the inherent tradeoff between commonality and distinctiveness within a product family. Overall, exploring the similarities and diversity of product families is a common step in building product platforms (Jiao *et al*. 2007). Similarly, product family technology can sort and divide the customized wooden door product system to increase the level of mass customization, and similarities between customized wooden door product families must be examined. For manufacturers, the construction of product families facilitates product modularity, where similar components are used as a set of modules, and a large variety of products can be created by combining fewer modules. Designers can quickly design products, improving research and development (R&D) efficiency.

Data clustering identifies groups based on a similarity metric if the classification is unknown (Omran *et al*. 2007). The similarity of different products' functional, physical, and technical features is an important characteristic of product families. Clustering algorithms can be used to explore the similarities between custom wood door products and the internal structure of product families so that the product platform of custom wood doors can be constructed scientifically. Few studies have paid attention to the application of clustering to wooden doors, and this paper explores the feasibility of clustering to customized wooden doors in favor of the construction of product family architectures. Ren and Xiong (2022) proposed a multi-attribute overlapping clustering method for dividing a family of cabinet door parts. A total of 314 solid wood parts were divided into 17-part families, and a digital design platform and its vital technology were developed. When clustering is applied in furniture division, it is usually based on the clustering results of the parts separated by the product. However, as the custom wooden door is a whole product, it is not easy to consider it detachable, unlike custom cabinets and solid wood doors. Therefore, this study located a specific product through various attribute characteristics to achieve customized wooden door product clustering. In addition, the k-means and hierarchical clustering algorithms, standard in clustering methods, are suitable for analyzing mixed data. The k-means algorithm has high efficiency and poor anti-noise performance, and its clustering shape is spherical. The clustering shape of hierarchical clustering is arbitrary, and hierarchical clustering has good anti-noise performance but poor efficiency (Frades and Matthiesen 2010). The hierarchical clustering algorithm is more stable than k-means clustering when the amount of data is not large.

Inspired by these advances, this study used multiple attribute features of a company's customized wooden door products as a database. Hierarchical clustering was used to explore customized wooden door products' classification and similarity relationships. The similarity between different products was used to establish the rational product platform and product family division.

EXPERIMENTAL

Clustering Materials

The clustering environment was Intel(R) Core (TM) i7-7700HQ CPU @ 2.80GHz 2.81GHz with 8G RAM. MATLAB R2020b from MathWorks was used for data processing. The multi-attribute dataset of custom wood door products was derived from the product manual data of an enterprise and contained 444 products and 9 attribute features. The feature attributes included numerical features, categorical features, and image attributes. To protect the enterprise's information security, all the datasets used in this paper were desensitized.

Clustering Theory and Formulas

Hierarchical clustering

Hierarchical clustering algorithms are classified as agglomerative and divisive, and the order of agglomerative hierarchical clustering algorithms is bottom-up. As shown in Eq. 1, the dataset X is a matrix n by m, where n is the number of sample points and m is the number of dimensional features. The strategy of agglomerative hierarchical clustering is to treat each sample point as a cluster. By calculating the distance between each cluster, the two clusters with the smallest distance are merged into a new cluster, and the calculation and merging are repeated until the cluster containing all the points is produced.

$$\boldsymbol{X} = [x_{ij}]_{n\times m} = \begin{bmatrix} x_{11} & x_{12} \cdots & x_{1m} \\ x_{21} & x_{22} \cdots & x_{2m} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} \cdots & x_{nm} \end{bmatrix} \tag{1}$$

Euclidean distance

One of the important elements of clustering algorithms is calculating the distance between sample points to measure the degree of affinity. Euclidean distance is usually used to calculate the distance between different sample points in hierarchical clustering. Euclidean distance is the distance between two sample points in m-dimensional space. The distance between sample a and sample b is calculated by Eq. 2.

$$dist(a,b) = \sqrt{\sum_{i=1}^{m}(a_i - b_i)^2} \tag{2}$$

Average linkage

When a cluster contains multiple sample points, the distance between a new cluster and other clusters can be calculated by different methods. The average linkage method considers the distance between clusters as the average distance between all pairs of points. It is also called UPGMA (Unweighted Pair Group Method with Arithmetic Mean). The average distance method is more computationally intensive, but the results are more reasonable than others. The distance between cluster C_p and cluster C_q is calculated by Eq. 3, where i is the sample points in cluster C_p, and j is the sample points in cluster C_q.

$$dist(C_p, C_q) = \frac{1}{|C_p|\cdot|C_q|}\sum_{i\in C_p}\sum_{j\in C_q} dist(i,j) \tag{3}$$

Procedure

As shown in Fig. 1, the data preprocessing procedure was set before cluster processing.

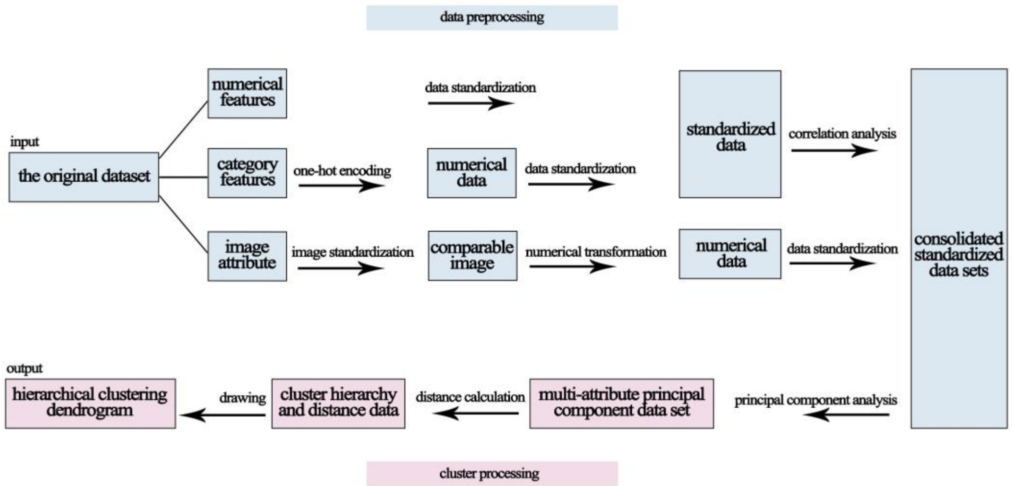


Fig. 1. Multi-attribute hierarchical clustering procedure

Data preprocessing

Before using machine learning to cluster data, real data sets in practical applications are mostly in diverse formats due to different sources. The data sets are prone to phenomena such as missing, noise, and inability to effectively identify by computer, which leads to poor quality of machine learning results if the data is directly used in data analysis work. Therefore, it is necessary to preprocess the initial data so that machine learning can better understand, process, and output. The original dataset of the enterprise's customized wooden door products had 444 sample data with 9 product features. The features of the product were "style", "craft door type", "material", "type", "number of doors", "price", "sales", "craft", and "appearance". The "number of doors", "price" and "sales" are numerical features, whereas "style", "craft door type", "material", "type" and "craft" are category features, and "appearance" is an image attribute. The original dataset is large and has missing values, and the data of category features and image attributes cannot be recognized effectively by the clustering algorithm.

One-hot encoding: Since the values of the category features in the dataset are discrete rather than continuous, distance calculations cannot be performed. However, one-hot encoding can extend discrete features to Euclidean space, and it operates by converting each of the discrete features to a state and generating a set of binary vectors. It can be interpreted that the m values of a feature are converted into m binary features. The binary feature corresponding to the state is 1, and the rest are 0. The "style" features in the custom wood door product dataset were subjected to a one-hot encoding process, then the "classical style" category was converted to [1,0,0,0,0,0], the "austere European style" to [0,1,0,0,0,0], the "minimalist style" to [0,0,1,0,0], the " austere Chinese style" to [0,0,0,1,0], and the "light luxury style" to [0,0,0,0,0,1]. The "style" features were converted from category-based to 5-dimensional numerical features. Similarly, the "craft door type" was converted to 10-dimensional features, the "material" was converted to 8-dimensional features, the "type" was converted to 8-dimensional features, and the "craft" was converted to 15-dimensional features. It can be seen that categorical data were converted to numerical data with a large dimension increase. The increase in data dimensions would lead to feature redundancy, increase in computation, and collinearity problems.

Data standardization: Multi-attribute indicators need to be compared both horizontally and vertically. Due to the nature differences in the attributes, the dimensions and orders of magnitude differ. If the original data is used to measure the distance while the difference between the attributes is large, it will amplify the distance difference between the attributes. Thus, the reliability of the clustering results is affected. Therefore, it is necessary to convert the data into a standard format. Z-score standardization is a method commonly used in data standardization. After solving the mean (μ) and standard deviation (σ) of the attributes, the score of observation x was calculated by $(x-\mu)/\sigma$ to achieve the Z-score standardization. The missing values in the dataset of custom wood door products might lead to inaccuracies in the standardization process. Standardized treatments are generally performed after missing values have been added, so we complemented the missing values with 0 in this dataset. The 49-dimensional characteristic data of style, craft door type, material, type, door number, price, sales volume, and craft were standardized by z-score to transform to standard normal distribution.

Correlation analysis: The increase in dimensionality caused by one-hot encoding might lead to feature redundancy and collinearity problems. Therefore, feature selection in feature engineering was used to remove redundant features, reduce computational cost, and avoid overfitting issues. Correlation analysis could measure the degree of correlation and closeness between features, and redundant features with over-high correlation could be removed by calculating the Pearson correlation coefficient. The correlation coefficients of 49-dimensional feature attributes were analyzed. The results are shown in Fig. 2.

Among the characteristic attributes, the correlation coefficient between "craft door type_buckle line" and "craft_buckle line" features, "craft door type_insertion" and "craft_insertion" features, "craft door type_colour patchwork" and "craft_hand patchwork" features, "craft door type_procurement", "material_alloy", "craft_procurement" and "craft_basic" features, "material_PVC membrane" and "craft_plastic suction" features, "material_water-based paint" and "craft_spray paint" features, "material_glass", "craft_Frame cutting" and "craft_glass fastening" features, "type_mute" and "craft_bottom slot milling" features were greater than 0.9. "Type_mute" for door products with a mute function achieved by attaching mute strips, mute locks, and bottom parts to the door to achieve sound insulation. This result indicated an extremely linear correlation or even collinear, so only the former was retained, and the remaining redundant features were removed. After feature selection, the feature dimensions of the new data were reduced from 49 to 39, including 5 dimensions for style, 10 dimensions for craft door type, 7 dimensions for material, 8 dimensions for type, 1 dimension for number of doors, 1 dimension for price, 1 dimension for sales and 6 dimensions for craft.

Image processing

The appearance attributes in the dataset about custom wood door products were derived from CAD files containing 444 drawings of the appearance of custom wooden door products, which could be output as images. Each greyscale image is stored in a two-dimensional array, where each array element, namely pixel, is a byte [0,255]. In the paper, image features in clustering were used to aid in the product description instead of image segmentation with clustering within a single image. Therefore, the image data must be processed and stretched into a standardized uniform and comparable one-dimensional array. It could be interpreted as an image with m pixels long and n pixels wide, stored in the computer as a two-dimensional array of m columns and n rows. It needed to be stretched into a one-dimensional array containing m by n data, which could be regarded as product features with m by n dimensions to characterize the appearance of that product.

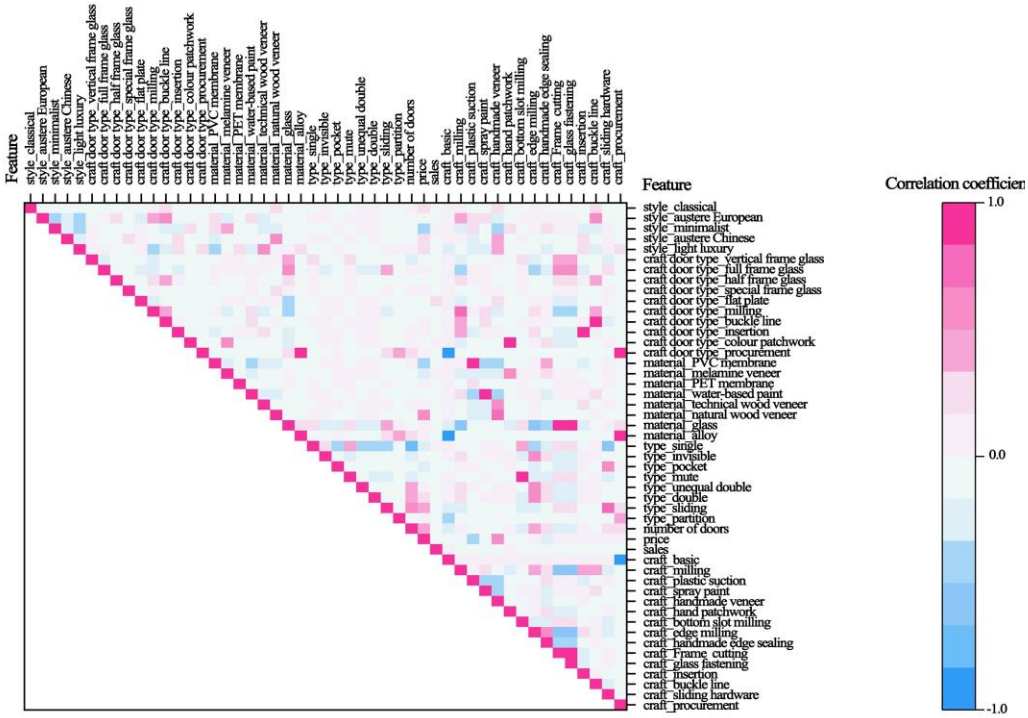


Fig. 2. Correlation coefficient heatmap

Image standardization: The image of m by n pixels would be converted into features with m by n dimensions by image processing. To avoid the problem of increased computation caused by large dimensions, minimizing the pixels while improving the efficiency of image representation was necessary. The drawings of the customized wooden door products were unified with a single-door aspect ratio of 2.3 to 0.83. The lower part of the drawings, sufficient to express the differences between the products, was intercepted and output as an image. The images were uniformly compressed to 50 by 33 pixels using Photoshop. It was observed that the differences in the product could not be clearly expressed if the images were too small. On the contrary, the data of images could not be stored in Excel files if the images were too large.

Numerical transformation: The image was presented as a 33 x 50 x 3 uint8 array after importing into MATLAB, and the grey scale image was converted to a binary image to form a 33 by 50 logical value (threshold was set to 0.8). After that, the logical values were converted to double precision values to facilitate calculations. The values were sorted into columns to convert the two-dimensional array into a one-dimensional array. Then the image was converted to a 1650-dimensional data set and exported to the corresponding data set of products.

Principal component analysis

After image processing, the dimensionality of the data set increased obviously, which might lead to information overlap. Therefore, it was necessary to reduce the number of features and redundant features by dimensionality reduction to explain most of the original information with fewer features. The principal component analysis method is one of the most widely used data dimensionality reduction algorithms, which reflects most of the original information through the linear combination of the original features to achieve the purpose of dimensionality reduction. However, the meaning of the linearly combined principal components is often obscure and difficult to interpret. Therefore, in this paper, the feature attributes of processed data sets were grouped according to the original feature attributes, which meant the original feature attributes were unfolded for principal component analysis to reduce the fusion of components between different attributes. For instance, the 5-dimensional data of the style were subjected to principal component analysis, and the new 5-dimensional features after linear combination were obtained. The linearly combined 5-dimensional features were sorted in descending order of contribution, and 4-dimensional features were selected based on a cumulative contribution of up to 0.9. The scores of these 4 dimensions were exported as new feature data noted as style_ch1, style_ch2, style_ch3, and style_ch4, respectively. Similarly, after principal component analysis, the dimensionality of "craft door type" was reduced from 10 to 8, the dimensionality of "material" from 7 to 6, the dimensionality of "type" from 8 to 6, the dimensionality of "craft" from 6 to 5, and the dimensionality of "appearance" from 1650 to 26. Overall, the origin dataset of custom wooden door products became a dataset containing 444 samples and 58 features after data preprocessing.

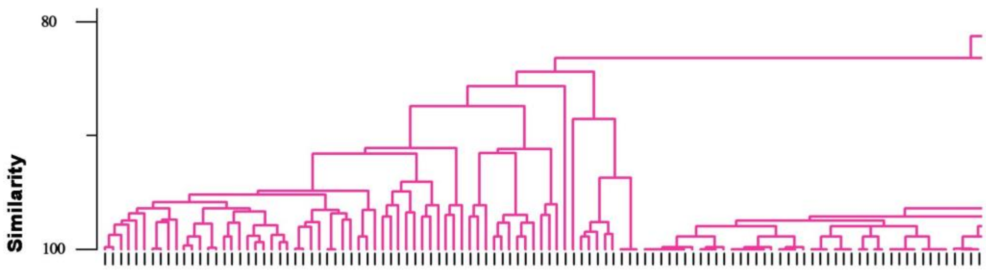
Cluster processing

The preprocessed custom wooden door product dataset with 444 sample points and 58-dimensional features was input to a 444 by 58 matrix. Due to the large dataset, the Euclidean distance between the row vectors, namely, between the 444 sample points, was measured by the *pdist* function in MATLAB, and a row vector containing 98,346 distance values was obtained. The *linkage* function in MATLAB was required to visualize the hierarchical clustering results to get the clustering tree diagram. After inputting the distance row vector, the distance between the new cluster and the other clusters was calculated according to the average distance method, and the 443 by 3 matrix containing the cluster tree information was output. The leaf nodes on the clustering tree were sample points of the original data from 1 to 444. The first and second columns of the output matrix represented the leaf nodes of the binary tree containing the new clusters generated by the combination of the initial leaf nodes, and the third column represented the distances between the leaf nodes in the first and second columns. The hierarchical clustering dendrogram of the custom wooden door product dataset was plotted with the *dendrogram* function based on clustering tree information to understand the clustering results' hierarchical structure better.

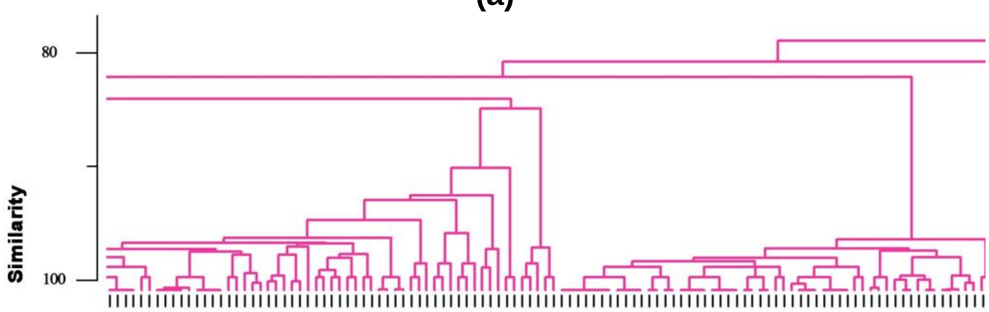
RESULTS AND DISCUSSION

Clustering Result

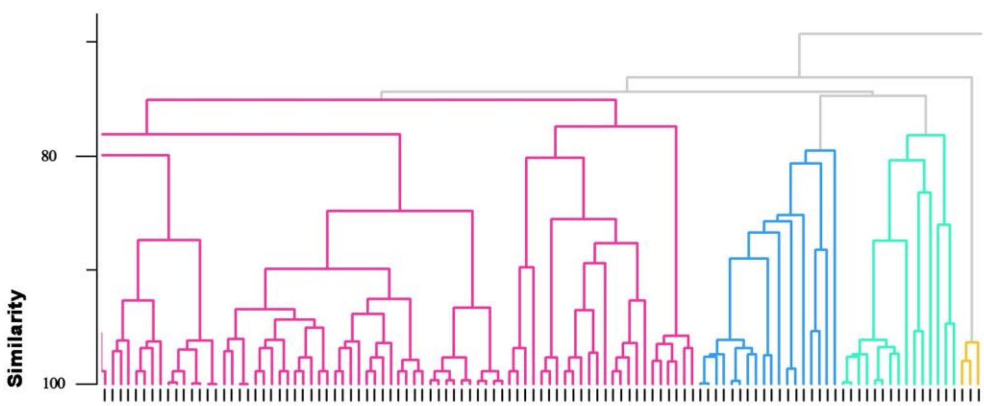
According to the hierarchical clustering dendrogram shown in Fig. 3, 444 samples were selected to be divided into 32 categories as the final clustering results. Category 1 (contained 297 samples): The category of single doors, which contained most of the single doors, included flat plate single doors, decorative single doors and glass single doors; Category 2 (18): The category of unequal double doors, which contained some of the unequal double doors, included flat plate unequal double doors and some of the decorative unequal double doors; Category 3 (15): The category of double doors, which contained some of the double doors, included flat plate double doors and some of the decorative double doors; Category 4 (3): The category of single doors decorated with double vertical lines on the left side; Category 5 (46): The category contained wooden frame glass sliding doors and some of the decorative double doors; Category 6 (1): The category contained the double doors decorated with two rectangular shape; Category 7 (4): The category contained carving decorative single doors and carving decorative unequal double doors; Category 8 (8): The category contained the unequal double doors decorated with vertical and horizontal lines and double doors decorated with vertical and horizontal lines; Category 9 (4)): The category of cross decorative glass doors; Category 10 (1): The category of cross decorative alloy doors; Category 11 (2): The category contained the double doors decorated with vertical and horizontal plate; Category 12 (9): The category contained the unequal double doors decorated with rectangular shape, two rectangular shape and H shape;



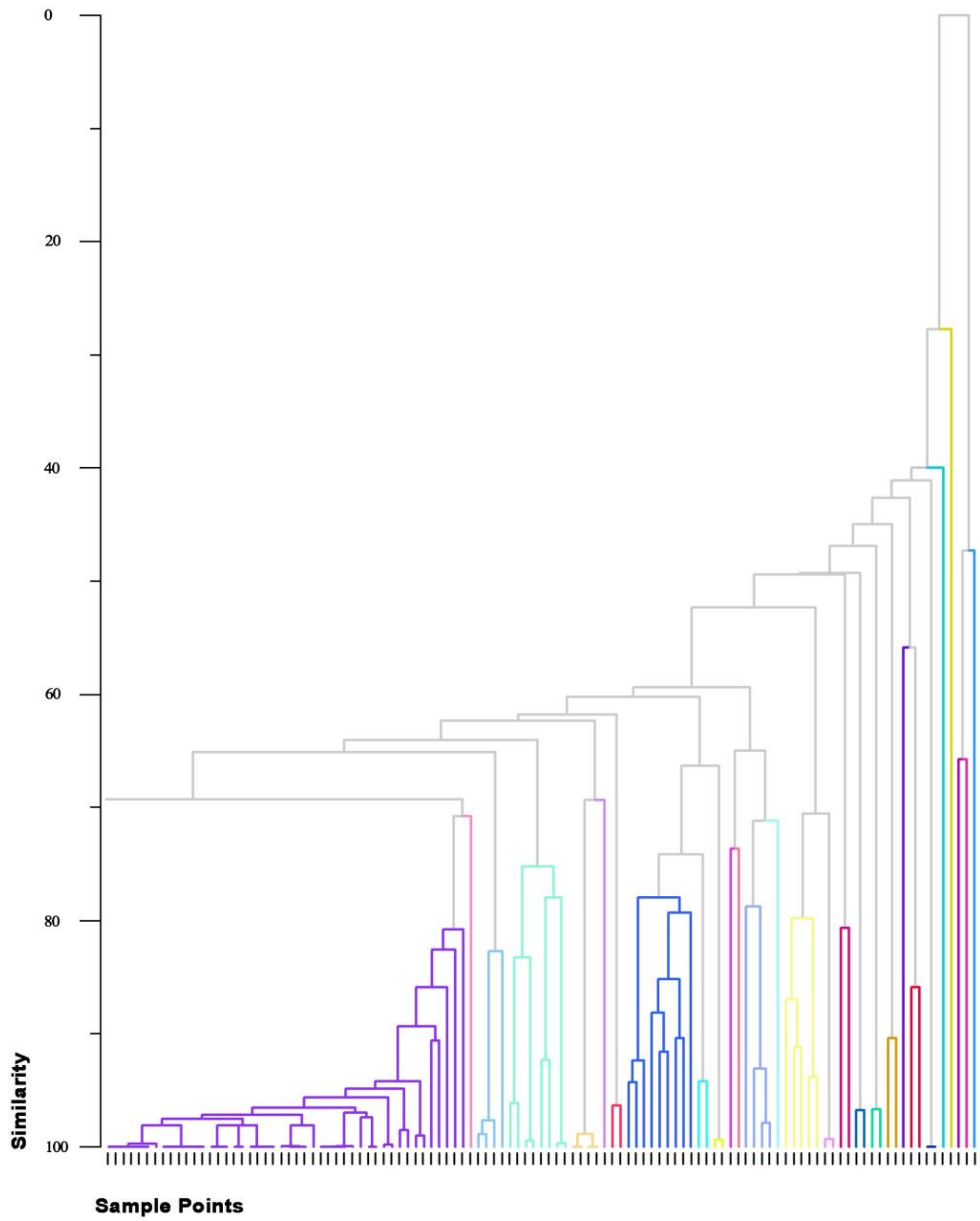
(a)



(b)



(c)



(d)

Fig. 3. Hierarchical clustering dendrogram

Category 13 (2): The category of the unequal double doors decorated with two rectangular shape; Category 14 (2): The category of the unequal double doors decorated with complex two rectangular shape; Category 15 (1): The category of the unequal double doors decorated with double vertical lines; Category 16 (1): The category of the unequal double doors decorated with rectangular shape on the outer side; Category 17 (4): The category contained the double doors decorated with double vertical lines ,the double doors decorated with rectangular shape on the outer side and the alloy sliding doors decorated with rectangular shape on the outer side; Category 18 (1): The category of the double door decorated with insertion on the outer side; Category 19 (5): The category contained some of the double doors decorated with two rectangular shape; Category 20 (2): The category of the double door decorated with complex two rectangular shape; Category 21 (2): The category contained the unequal double doors decorated with double vertical lines on the left side and the double doors decorated with double vertical lines on the left side; Category 22 (2): The category of the unequal double doors decorated with multi-vertical lines; Category 23 (2): The category of the double doors decorated with multi-vertical lines; Category 24 (2): Grille pocket doors and partitions; Category 25 (1): The category of the unequal double door decorated with rectangular shape on the outer side; Category 26 (2): The category contained the double door decorated with rectangular shape on the outer side and the wooden frame glass sliding doors decorated with insertion on the outer side on the outer side; Category 27 (2): The category of cross decorative wooden frame glass sliding doors; Category 28 (1): The category of four-leaf alloy sliding door; Category 29 (1): The category of carving decorative double door; Category 30 (1): The category of grille alloy double sliding door; Category 31 (1): The category of grille alloy three-leaf sliding door; Category 32 (1): The category of grille alloy four-leaf sliding door.

Analysis of Phenomena and Errors

The clustering results showed that the categories were numerous and uneven. The largest category included 297 samples, and the smallest included only one sample. The clustering results were not entirely interpretable. For example, category 5 in the clustering results contained wooden frame glass sliding doors and some decorative double doors. Category 5 could not be entirely interpreted by wooden frame glass sliding doors. The distinction between categories was not clear enough. For example, category 2 contained a majority of unequal double doors, and category 12 also contained unequal double doors. It wasn't easy to find the critical descriptions that distinguish them. The clustering dendrogram was divided into 32 categories. If the number of categories was reduced, then the categories would be merged so excessively that the categories were more unbalanced and inexplicable. In contrast, if the number of categories were increased, then the categories would be too trivial to achieve a better effect of clustering and merging. In the 32 categories, the sample size of the single doors category, the unequal double doors category, the double doors category, and the wooden frame glass sliding doors category were far larger than the other categories, and it could be viewed as four large and concentrated categories and other stray categories. In the classification, most wooden frame glass sliding doors and a small proportion of decorative double doors were clustered together in category 5. The errors that appeared in Category 5 could be interpreted.

The position of the glass outer frame lines overlapped with the position of the decorative frame lines, which resulted in a similar appearance plan, but there was a large proportion of 26 appearance variable features, which resulted in a high degree of similarity between the wooden frame glass sliding doors and some of the decorative double doors. The errors that stray categories were not clustered into concentrated categories or became categories of their own could be explained. Although there were many identical features between the stray categories and some concentrated categories, the image conversion process based on pixel points and the large number of appearance features still resulted in many different variables. Because a pixel point difference in the pattern was a variable difference, the similarity and difference variables canceled out to form the free and independent categories. To minimize the error, distinguishing similar images and adjusting the percentage of image dimensions can be considered in future studies.

Analysis of Similarity

The clustering dendrogram was clustered bottom-up from the closest and most similar products. There were some samples whose positions were exactly overlapping and were clustered into one category first. This indicated that there were precisely the same products among the 444 products, with only naming differences. These duplicated and redundant products should be optimized and merged into one product to avoid any misalignment in the production process and sales records. Secondly, the closest were the doors with mute function and the same product without mute function, such as the basic single door with mute function and the basic single door, the glass door with mute function and the glass door. There was only one difference between the functional components, so they were the most similar products. In addition, the position of products with the same door but made of different materials was very close. They were made of different materials, and their processes were slightly different, but most were the same. From the customer's perspective, they were the same door products with varying properties of the material, such as full frame glass doors with PVC membrane, water-based paint, melamine veneer or wood veneer, flat plate basic single doors with different materials, vertical frame glass doors with different materials, and so on. The distance between invisible doors and pocket doors was also close. The invisible door and the pocket door were clustered into the same basic single door category, and their appearance was consistent, but the structure and process were quite different. Finally, the single door, unequal double door, double door, and wooden frame glass sliding door were clustered together.

In a word, the results of using hierarchical clustering to cluster customized wooden door products based on the style, craft door type, material, type, number of doors, price, sales, craft, and appearance features were uneven and not entirely interpretable. However, the internal order structure of clusters and the clustering process could be clearly observed, and the distance hierarchical relationship between the products could be obtained, which was beneficial to the division of the product architecture.

Product Architecture of Customized Wooden Doors

Traditional research and development (R & D) involves a separate development for one product, independent between products. The traditional R & D has no high efficiency in promptly responding to the agile demand development. However, the construction of product family architecture can precisely develop a modular component for agility requirements. It can be extended by developing a range of products in different combinations within the product platform. In order to optimize the product system architecture, based on the distance hierarchy between products derived from the clustering of custom wood door products, the custom wood door product family was scientifically and reasonably established, which improved R&D efficiency to cope with agile demand.

As shown in Fig. 4, the single door is regarded as a product platform. Different structural combinations of the product platform can constitute single doors, unequal double doors, double doors, sliding doors, and other basic types to satisfy the customer's demands for door size and opening methods. Personalized attributes such as structure, appearance, and function can be added to the product platform to enrich the product family generated by the combination. Among them, the structure attributes contain basic, pocket, invisible, fully frame glass, vertical frame glass, and alloy glass doors. R&D in the direction of the existing product platform structure attributes might require increased production lines and factory scale.

Table 1. Product Coding Rules of Customized Wooden Doors

Coding Digits Number	Classification Name	Descriptions
1	1-product code 2-part code ……	product code is distinguished from other codes
2-3	01- customized wooden door 02- customized wardrobe 03- customized cabinet ……	customized wooden door is distinguished from the company's other categories
4-5	01-single doors 02- unequal double doors 03- double doors 04-sliding doors ……	the basic type of doors
6-7	01-basic doors 02- pocket doors 03-invisible doors 04-full frame glass doors 05-vertical frame glass doors 06-alloy glass doors ……	structural attribute
8-9	01-PVC membrane 02- melamine veneer 03- water-based paint 04-pet membrane ……	the material information in the appearance attribute
10-11	01-image01 02- image02 ……	the color images of the material
12-13	01-insertion 02- milling 03- buckle line 04-hand patchwork ……	the decoration process information in the appearance attributes
14-15	01- vertical lines on the left side 02- H shape 03-two rectangular shape ……	the decoration pattern information in the appearance attribute
16-17	01-mute function 02-Air purifier ……	the function module

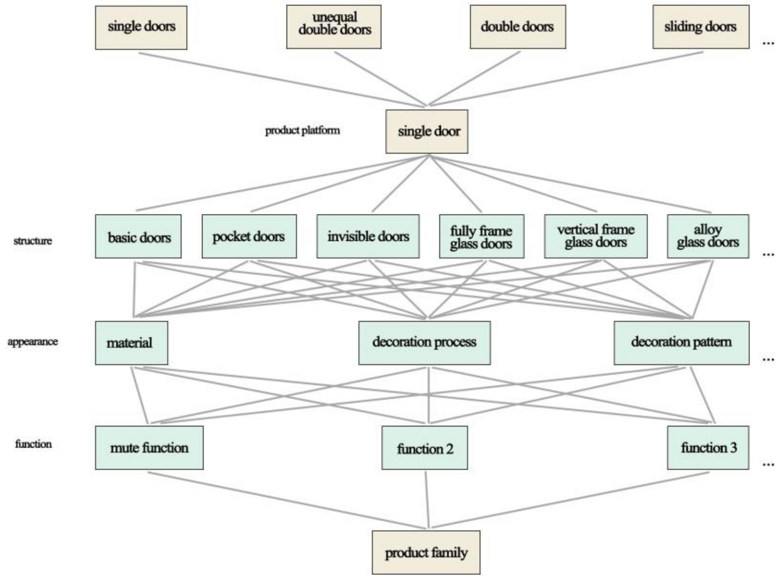


Fig. 4. Product architecture of customized wooden doors

Appearance attributes include material, decoration process, and decoration pattern: The existing material includes PVC membrane, melamine veneer, water-based paint, pet membrane, and other materials. The raw materials offered by the suppliers are abundant and are constantly being updated in search of scratch-resistant, dirt-resistant skin textures and new aesthetic patterns. The existing decoration process includes the milling, buckle line, insertion, and other decorative methods, and the new R&D designs combine two or more of the above. The existing decoration patterns are diverse, and minimalist patterns are usually prevalent in the modern market. In contrast, complex retro patterns are gradually being optimized and eliminated due to the complexity and high cost of the process. Existing functional attributes include only the mute function, which achieves additional value-added effects through modular replacement of small parts. Most of the enterprise R & D updates in the structure and appearance attributes and the modular functional attributes are rarely concerned about. The development of functional modules and module standard interfaces for customized wooden doors is lacking.

Product Coding Rules of Customized Wooden Doors

Product coding has evolved from serial number coding to a combination of letters and serial numbers. It is a conscious effort to include product information while keeping the coding clear and concise (Xiong *et al*. 2008, 2009). During coding iteration, the simultaneous emergence of old and new codes results in unclear coding meanings, misrepresentation of financial statistics, and errors caused by the inability of employees to adapt to the new regulations promptly. In recent years, major enterprises have implemented information management systems, such as Product Lifecycle Management. Considering the convenience of coding information entry, coding can only use numeric characters (Xiong *et al*. 2011; Qi *et al*. 2021). Therefore, the product coding needs to be revamped completely, and the impact of new products on coding changes should be considered to minimize changes in coding information due to product replacement. Coding rules that refer to product family division can complete the product information in the code. In this way, when a new component is developed, only the information of the component needs to be entered instead of all the product information corresponding to the serial number.

As shown in Table 1, the 1st-bit identification code contains the information to which the code belongs, *e.g.*, the product code is distinguished from the part code. The 2nd to 3rd digit are the product category code, which contains information that distinguishes the custom wood door category from the company's other categories. Bits 4-17 are the feature code. Among them, bits 4-5 contain information about the basic type of doors, such as single, unequal double, double, and so on. Bits 6-7 contain structural attribute information, such as basic doors, pocket doors, full frame glass doors, *etc*. Bits 8-9 contain the material information in the appearance attribute, such as PVC membrane, melamine veneer, water-based paint, *etc*. Bits 10-11 contain the color code of the material. Bits 12-13 contain the decoration process information in the appearance attributes, such as insertion, buckle line, hand patchwork, *etc*. Bits 14-15 contain the decoration pattern information in the appearance attribute, such as vertical lines on the left side, H shape, two rectangular shape, *etc*. Bits 16-17 contain information about the function module, such as the mute function.

Front-End Development Process of Customized Wooden Doors

The product development process is the system of information acquisition, processing, and transmission. Many companies use modern information systems to manage the entire product life cycle, and the combination of online and offline is used for product development. Rapid decision-making is achieved to shorten the R&D cycle, R&D processes are standardized and simplified, and internal operational efficiency is improved (Xiong *et al*. 2015; Zhao and Zhong 2022). The development and design of customized wooden door products begins with market research and the needs expressed by the marketing department. When the market research and demand are passed to the R&D department, the development engineers and the head of R&D decide whether to accept the demand. After accepting the requirements, the development engineers design the solution and generate models and sample drawings. The head of the R&D department decides whether the solution is approved or needs to be re-optimized. The Proofing department trial-produces samples and sends them to the laboratory for testing. If the sample works well, then internal and external people are organized to review the sample offline. The reviewers include the marketing department to assess the sales feasibility of the components to be put into the market, the production department to determine the risk and impact of the quality of production in the factory, the person in charge of the supply chain to confirm the procurement of materials for the components, the pricing specialist to assess the price of the components to be shipped from the factory, the person in charge of product management to determine the possibility of the components to be included in the catalog, the quality department of the standards department to review the quality of the components and the drawings, and the person in charge of information technology to assess the degree of information technology according to the type of parts to delete and add reviewers. If the review score is passed, the engineer submits the online application for inclusion, and the information required for inclusion is submitted by different persons in charge and reviewed. Then, the samples are sealed and recorded in the information system.

CONCLUSIONS

1. The results of using hierarchical clustering to cluster customized wooden door products based on the style, craft door type, material, type, number of doors, price, sales, craft, and appearance features were found to be uneven and not entirely interpretable.

However, the internal order structure of clusters and the clustering process could be clearly observed, and the distance hierarchical relationship between the products could be obtained, which was beneficial to the division of the product architecture.

2. There were completely overlapping sample points in the products, and these duplicate products needed to be optimized. Differentiation of the closest sample is the presence or absence of mute function, which achieves additional value-added effects through modular replacement of small parts.
3. According to the internal similarity hierarchy, the customized wooden door product family was rationally divided based on the single door as the product platform. The product architecture, product coding rules, and front-end development process were established to improve standardization and research and development efficiency.
4. The results illustrated that processing using hierarchical clustering of multi-attribute data is feasible for optimizing customized wooden door product systems. To minimize the error, distinguishing similar images and adjusting the percentage of image dimensions can be considered in future studies. The current R&D on the structure and appearance of custom wooden doors is rich and varied, but the R&D on functional modules is weak. There is still great potential for developing the custom wooden door category in custom furniture companies.

ACKNOWLEDGMENTS

The authors are grateful for the support of the Qing Lan Project.

REFERENCES CITED

Alizon, F., Shooter, S. B., and Simpson, T. W. (2009). "Assessing and improving commonality and diversity within a product family," *Research in Engineering Design* 20(4), 241-253. DOI: 10.1007/s00163-009-0066-5

Da Silveira, G., Borenstein, D., and Fogliatto, F. S. (2001). "Mass customization: Literature review and research directions," *International Journal of Production Economics* 72(1), 1-13. DOI: 10.1016/S0925-5273(00)00079-7

Feng, X. H., and Wu, Z. H. (2020). "Research and analysis study on the development status of the domestic wardrobe and cabinet furniture market," *China Forestry Industry*

Frades, I., and Matthiesen, R. (2010). "Overview on techniques in cluster analysis," *Bioinformatics Methods in Clinical Research* 593, 81-107. DOI: 10.1007/978-1-60327-194-3_5

Hu, Z. P., and Xu, B. M. (2023). "Research on the promotion of customized furniture brands against the backdrop of the internet traffic economy," *Furniture Interior Design* 30(2), 79-83. DOI: 10.16771/j.cn43-1247/ts.2023.02.014

Jiao, J., Simpson, T. W., and Siddique, Z. (2007). "Product family design and platform-based product development: A state-of-the-art review," *Journal of Intelligent Manufacturing* 18(1), 5-29. DOI: 10.1007/s10845-007-0003-2

Lihra, T., Buehlmann, U., and Graf, R. (2012). "Customer preferences for customized household furniture," *J. For. Econ.* 18(2), 94-112. DOI: 10.1016/j.jfe.2011.11.001

Lv, B., Fu, Y. J., and Zhang, Y. P. (2008). "Status and development trends of wooden door industry in China," *Chinese Journal of Wood Science and Technology* 22(5), 17-20. DOI: 10.19455/j.mcgy.2008.05.006

Omran, M. G. H., Engelbrecht, A. P., and Salman, A. (2007). "An overview of clustering methods," *Intelligent Data Analysis* 11(6), 583-605. DOI: 10.3233/IDA-2007-11602

Peng, J. W., Zhang, J. J., and Zhang, Z. F. (2019). "Study on modular design based on customized cabinets," *Furniture Interior Design* 26(6), 20-21. DOI: 10.16771/j.cn43-1247/ts.2019.06.005

Qi, Y. Q., Xu, R., and Yu, D. F. (2021). "Digital transformation and intelligent manufacturing of furniture industry," *Furniture and Interior Design* 28(8), 68-71. DOI: 10.16771/j.cn43-1247/ts.2021.08.014

Ren, J., and Xiong, X. Q. (2022). "Digital design process and part family division of solid wood custom cabinet door based on multi-attribute overlapping clustering technology," *Bioresources* 17(3), 5393-5419. DOI: 10.15376/biores.17.3.5393-5419

Ren, J., and Xu, B. M. (2019). "Research on dynamic optimization of standardized parts platform for customized wardrobe," *China Forest Products Industry* 26(2), 64-67. DOI: 10.19531/j.issn1001-5299.201902013

Simpson, T. W., and D'Souza, B. S. (2004). "Assessing variable levels of platform commonality within a product family using a multiobjective genetic algorithm," *Concurrent Eng.: Res. Appl.* 12(2), 119-129. DOI: 10.1177/1063293X04044383

Xiong, X. Q., Huang, Q. T., Wu, Z. H., and Sun, H. N. (2015). "The building and application digital management platform in wood furniture processing," *China Forest Products Industry* 42(7), 40-43. DOI: 10.19531/j.issn1001-5299.2015.07.014

Xiong, X. Q., and Wu, Z. H. (2011). "Information collection and processing technique in mass customized furniture production," *Chinese Journal of Wood Science and Technology* 25(6), 17-20. DOI: 10.19455/j.mcgy.2011.06.006

Xiong, X. Q., Yue, X. Y., and Wu, Z. H. (2022). "Current status and development trends of Chinese intelligent furniture industry," *Journal of Renewable Materials* 11(3), 1353-1366. DOI: 10.32604/jrm.2022.023447

Xiong, X. Q., Zhong, S. L., and Wu, Z. H. (2008). "The application of coding in furniture enterprises," *Furniture* 29(4), 50-53. DOI: 10.16610/j.cnki.jiaju.2008.04.019

Xiong, X. Q., Zhong, S. L., and Wu, Z. H. (2009). "A coding method for identifying solid wood furniture manufacturing processes," *Chinese Journal of Wood Science and Technology* 23(6), 8-10. DOI: 10.19455/j.mcgy.2009.06.003

Xu, F. R. (2011). "Status and development trend of wooden door industry in China," *China For. Prod. Indust.* 38(2), 9-12. DOI: 10.19531/j.issn1001-5299.2011.02.003

Zhao, W. X., and Zhong, S. L. (2022). "Improving the information system of custom-furniture production," *Chinese Journal of Wood Science and Technology* 36(6), 41-46. DOI: 10.12326/j.2096-9694.2022099

Zhu, Y. J., Fu, Y. J., and Wang, C. X. (2013). "Current status and main restrictions of wooden door industry in China," *Chinese Journal of Wood Science and Technology* 27(1), 39-42. DOI: 10.19455/j.mcgy.2013.01.013

Article submitted: August 24, 2023; Peer review completed: September 16, 2023; Revised version received and accepted: September 26, 2023; Published: October 4, 2023.
DOI: 10.15376/biores.18.4.7889-7904