Research on product appearance patent spatial shape recognition for multi-image feature fusion

Abstract

In the process of patent retrieval, the traditional content-based single image retrieval method mainly has the following two reasons: a) semantic deviation caused by text description, b) the similarity of a single pixel in the image is high but the whole is inconsistent. Low accuracy leads to unsatisfactory retrieval results, which makes it difficult to obtain product design information timely and effectively and reduces design efficiency. How to obtain data quickly and accurately has become a challenging problem. In this paper, by analyzing the problems existing in Locarno classification method, combined with the characteristics of the patent image, a new improved method is proposed. Firstly, the structural features of product parts are extracted through segmentation. Subsequently, combine it with multi-view image fusion to determine the spatial shape of product parts jointly. Finally, the spatial shape of key structures is confirmed to refine the specific search range as well as improve the search accuracy. The feasibility and effectiveness of the proposed method are verified by taking a shower appearance patent as an example.

1 Introduction

Since the 1990s, Content-Based Image Retrieval [24] (CBIR) has been widely used in search engines, e-commerce and other fields. This technology starts from the underlying visual features of the image, establishes a global feature index of the image (such as texture, color, shape, etc.), and achieves image recognition and retrieval by calculating the similarity between different images. However, the performance of patent image data search is unsatisfactory, partly due to factors such as uneven quality of patent images, obvious noise or low clarity, embedded text information, and image shooting angle. In addition, traditional image search generally retrieves images based on metadata information such as titles and tags around them. Nonetheless, textual information cannot provide a complete overview of the information conveyed by the image. For example, it cannot meet the needs of users in terms of concretization when searching for a specific shape of a product, and the search results may be unreliable. The problem of the "semantic gap" widely exists in CBIR systems, which cannot fully meet users' needs for content retrieval, especially in appearance product patents. Therefore, the improvements and optimizations of search range and accuracy are needed. In fact, the traditional retrieval method based on single image input can no longer meet the needs of users. To solve the problem of obtaining input images, feature-based image retrieval combined with annotation-based image retrieval has become an important trend [1].

The extraction and recognition technology based on image features has been a hot research topic in recent years. At present, most of the existing research on image retrieval of appearance product design patents is based on classic image processing methods to extract features, which using a single underlying feature or fused underlying features as search conditions. Dai et al. [21] proposed a retrieval method to extract texture, shape, and color features from appearance patent images, the shape features extracted by serial Hu moments and the texture features extracted by grayscale co-occurrence matrix are used as retrieval inputs to provides a new solution for optimizing the appearance patent image retrieval system. Later, Fang et al. [8] improved the retrieval method of appearance patent images based on texture and shape features, taking advantage of chain codes and Hu invariant moments to extract shape features, and designed a query system suitable for combining appearance patent images and text. Zhou et al. [25] introduced semantic features in similar patent retrieval, which combined it with the underlying visual features of the image as search criteria. By analyzing the characteristics of appearance patents, Zheng et al. [10] proposed a retrieval method that combines multi-view features and semi-automatic keyword annotation, aiming to overcome the shortcomings of harnessing a single image information or keyword retrieval system. Subsequently, Sun et al. [13] proposed a method using color, texture, and shape features to extract information from appearance design patent images, then completed clustering by adjusting weights to calculate patent similarity, finally the research team has developed a design patent analysis software which has functions such as searching for patent development trajectories and regional differences. Furthermore, He et al. [23] utilized the invariant moment shape feature vectors by using improved weighted Euclidean distance as retrieval input, fused it with the similarity values of the image's internal structure texture features to achieve better improvement. Li et al. [28] extracted four features of patent images (namely shape, color, edge, and texture), and normalized and synthesized them for similarity-matching retrieval, which achieved certain results. Meanwhile, Wang et al. [22] extracted the shape and edge features from appearance patent images, then introduced support vector machines to classify design patent images. In addition, Li et al. [20] proposed a method of integrating the underlayer features of patent images and patent issue features to achieve multi-modal retrieval of appearance patents.

Different with normal images, unique information in patent images is described by color, geometry, and spatial arrangement. To capture the content in patent images, color, shape information, geometric structure and the spatial relationships between geometric structures will be the target information to be explored. Huet [7] developed a proprietary image retrieval system based on image content, the results depend on the size and complexity of the images which used for comparison. Zeng [31] proposed a method for appearance patent image retrieval based on a contour description matrix which uses Euclidean distance for similarity calculation, and then introduces a parameter to prioritize and weigh radial and angular information.

Csurka [3] proposed a multimodal patent retrieval system that combines the proposed visual and text-based approaches by using post-weighted fusion techniques, but the overall gain is small as for MAP (Mean Accuracy) and NDCG (Normalized Discounted Cumulative Gain) metrics. Meanwhile, Cai [2] extracted multiple underlying visual features of an image, set different weights for each feature, and achieved retrieval of design patent images by fusing these features. Li [12] performed frequency domain filtering on the image to extract local texture features and global features of the image, respectively to achieve multi-feature fusion retrieval. Notwithstanding, the method of extracting only the underlying visual features of an image for retrieval is limited and cannot meet the retrieval needs of people. A large number of design patent image recognition and classification methods mentioned above focus on visual feature extraction of design patent images, such as extraction of shape, texture, color, and edge features, or the fusion of multiple features mentioned above.

With the development of deep learning, some researchers focus on deep learning to extract image features to improve the performance of retrieval functions. For instance, Jiang [9] extracted image features of design patents under the massive patent image by training a convolutional neural network to optimize the performance of the retrieval system. Besides, an automatic vectorization using a novel convolutional neural network architecture, dual visual geometry group (VGG), was proposed in Lu [15]. Li Ming [11] chosen Deep Convolutional Neural Network (DCNN) for feature extraction. A region suggestion network (RPN) was added to the DCNN to improve the recognition accuracy of patent graphics and to reduce the probability of image feature loss, which reduces the number of parameters and shortens the retrieval time.

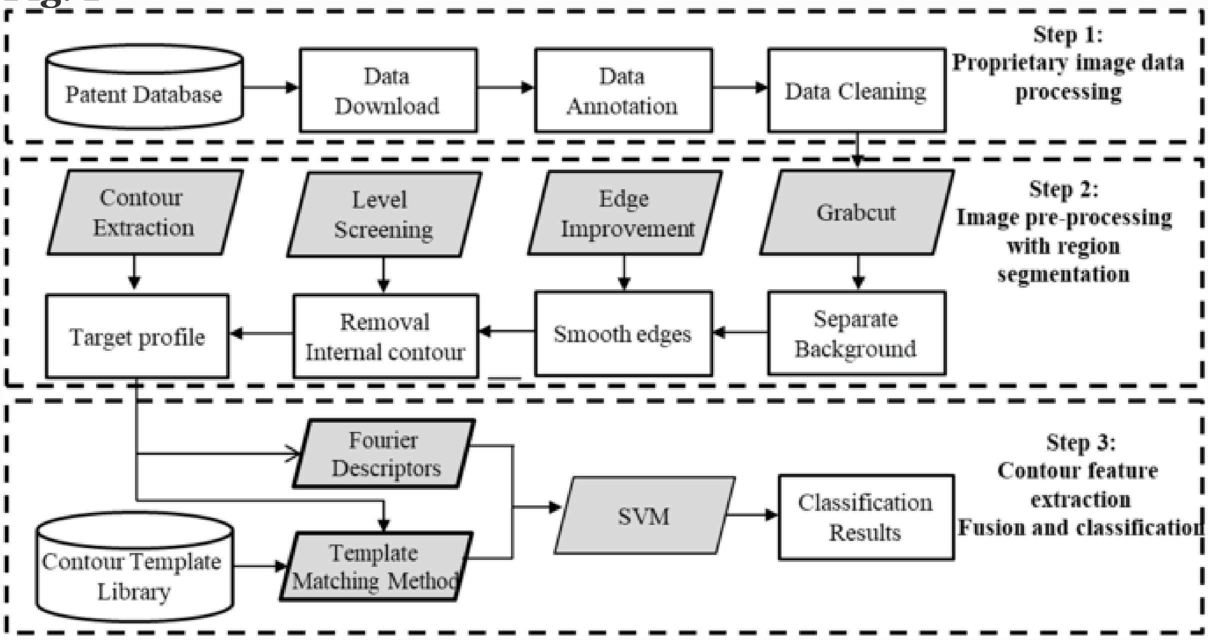
In summary, to fill the gap in existing research, this paper comprehensively combines the characteristics of appearance patent images with the current research status at home and abroad. Guided by the spatial shape of product appearance patents, it analyzes the design concept of product appearance patents, clarifies the product design concept and constructs innovative product design models, finally studies the positive effects of product innovation methods on product design. Firstly, multi-perspective image data of appearance patents is taken as the research object to address the impact of factors such as multi-perspective, uneven image quality, and no fixed specifications on appearance patent images. Secondly, traditional digital image processing and extraction of feature vectors from different perspectives in patent images are analyzed and compared, supplemented by multi-image feature fusion rules to explore product design diversity and meet people's search intentions. Moreover, the retrieval system is constructed for the convenience of management, examination, and application of design patents, and the case of shower patents is applied to verify the feasibility of the method finally.

2 Research procedure

Figure 1 shows the extraction, detection, and fusion of design patent image features, which are divided into three steps as follows:

- Step 1: Patent downloading, screening, and cleaning.

Fig. 1



Process of extraction, recognition, and fusion of design patent image features

Full size image

Search for relevant design patents; download patent images in bulk; sort by view and shape while considering patent image diversity; clean data to ensure diversity and category balance.

- Step 2: Image pre-processing and region segmentation.

Pre-processing and image segmentation are performed on the filtered appearance patent images. Due to the varying quality of the patent images in terms of appearance, it is especially important to extract smooth contour lines. Therefore, it is necessary to segment the main body first, segment the structure of the study subject, and preprocess the segmentation results to facilitate subsequent feature extraction operations.

- Step 3: Image contour feature extraction, screening and classification

Multi-view image contour features are extracted for each structure to build a proprietary library of contour templates, and image matching coefficients are calculated. The contour features are then normalized and used as input to a continuously fused SVM to train the model for classification.

3 Key technologies

Checking the visual content contained in a patent becomes extremely important in order to confirm the novelty of a newly filed patent application. And graphics in patent applications, as key descriptive documents, include technical drawings, graphs, charts, data flow diagrams, flowcharts, drawings and diagrams.

- A single scene, clear classification, etc. but no uniform color background
- some images have embedded view text information.
- There are so many different types of images that the main method of searching is text-based image retrieval.
- Users may input images with complex backgrounds when using the system for queries, which may have an impact on recommending similar images.
- Multi-view representation

3.1 Image pre-processing and region segmentation
Considering the inevitable background interference in the process of acquiring the appearance patent image, it may have an impact on the sampling quality of the appearance patent image. Before extracting the appearance patent features of the shower, image pre-processing is required. These interference factors include background color, shooting angle, light source angle, electromagnetic wave interference, etc. To cope with these noises, image processing with OpenCV is required, including grayscale, filtering, threshold segmentation, edge detection, and other operations to extract the complete connected domain shape. In addition, partial

morphological processing of the patented image of the shower appearance is required to ensure the accuracy of the extracted feature parameters.

Firstly, the GrabCut algorithm is used to achieve target region segmentation. The GrabCut algorithm is an interactive target segmentation method, whose main idea is to map the image into an s-t network graph [16]. The foreground and background are divided based on the manually marked rectangular box of the region of interest (ROI), with the background outside the rectangle and the background and foreground inside the rectangle. Utilizing texture (color) and edge (contrast) information in the image, a Gaussian mixture model was used to model the foreground and background, and the classification energy function was applied to pixels within the ROI until convergence [19]. For grayscale images, it can be used as a grayscale value matrix for segmenting region R Use one "transparency" value parameter for each pixel within $\upgamma =({\upgamma }_{1},{\upgamma }_{2}\dots {\upgamma }_{\mathrm{n}})$. Among them $({\upgamma }_{\mathrm{j}}\in (\mathrm{0,1})$, final output $({\upgamma }_{\mathrm{n}}=1$, it is determined that it may be the foreground part, and the segmented foreground background image is obtained. The implementation steps are as follows:

- Step 1: First, define a rectangular box containing the object to be segmented and define a target area.
- Step 2: Model the foreground and background using a Gaussian mixture model and classify pixels with uncertainty.
- Step 3: Connect image pixels to foreground or background nodes. Connect adjacent pixels using virtual edges and assign probabilities of belonging to foreground or background on these edges based on the similarity of colors.
- Step 4: For interconnected nodes, if they belong to different terminals, cut the edges between them to achieve image segmentation.

As shown in Fig. 2, the results of image segmentation before the shower and background are presented. This algorithm has better contour recognition performance than other segmentation algorithms. Such as the Otsu threshold segmentation algorithm, which is sensitive to noise, not significantly different in grayscale, and does not have obvious overlapping segmentation of different target grayscale values. Its robustness is not high [18].

Fig. 2

Comparison of entity extraction for shower

Full size image

3.1.1 Improved edge smoothing algorithm

The contour is crucial for shape recognition, and the smoothness of the contour curve will affect the accuracy of subsequent shape feature descriptions. The contour extraction algorithms based on edge segmentation in the past were prone to missed segmentation and over-segmentation, and the curves were not smooth enough. As shown in Fig. 3, it is evident that the unprocessed image edges do not perform well, with redundant backgrounds being judged as solid and edge optimization not presenting a stepped shape. After segmentation of the target area, some edges of the image appear stepped or are misjudged as redundant backgrounds.

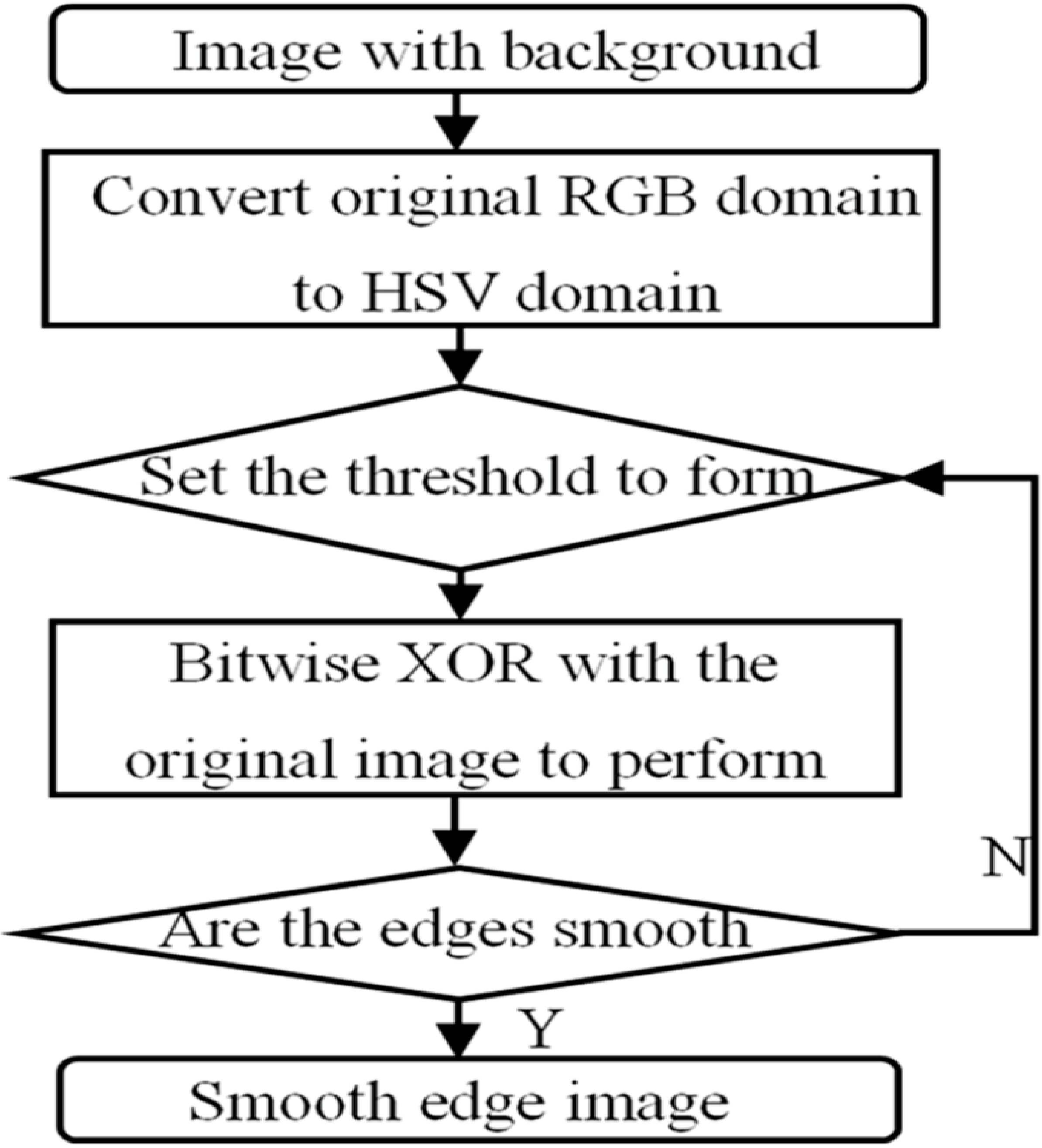
Fig. 3

Unprocessed Shower Edge Image

Full size image

Therefore, the image edges need to be smoothed to eliminate the segmented edge defects to obtain a better target segmented image. Figure 4 shows the specific steps, as follows:

Fig. 4



The edge smoothing improvement process

The improved edge smoothness of the shower is shown in Fig. 5, which removes the mistakenly considered foreground parts and achieves the effect of softening the contour. Figure 6 shows the extraction and recognition of the target shower contour.

Fig. 5

Shower edge image after improved edge smoothness

Full size image

Fig. 6

Extracting the contour of a shower

Full size image

3.2 Contour feature extraction and matching

After extracting the target edges by the improved edge smoothing algorithm, the images may be panned, rotated, or scaled considering the placement, angle, and camera height due to the placement of the image when it was captured. Hence the extracted features [17] should have panning, rotating and scaling characteristics. Therefore, Hu moments and Fourier descriptors are applied to detect the contours of the pre-processed image.

3.2.1 Hu moment feature extraction

Image can be recognized fast by features composed of Hu moments. However, this method has the disadvantage that the regions cannot be properly framed for contours of complex textures [29], thus resulting in a relatively low matching rate. Hu moments are generally used to recognize large objects in images. The better the shape description, the higher the matching rate.

A digital image I of size M $\times$ N has a gray value of $f\left(x,y\right)$ at point $\left(x,y\right)$, and is ($p+q$) order moment ${\mathrm{m}}_{\mathrm{pq}}$ and center moment ${\upmu}_{\mathrm{pq}}$ are:

$${m}_{pq}=\sum_{x=1}^{M}\sum_{y=1}^{N}{x}^{p}{y}^{q}f\left(x,y\right)$$

(1)

$${u}_{pq}=\sum_{x=1}^{M}\sum_{y=1}^{N}{\left(x-\overline{x}\right)}^{p}{\left(y-\overline{y}\right)}^{p}f\left(x,y\right)$$

(2)

where $f\left(x,y\right)$ is the gray value of the image at point $\left(x,y\right)$, $\overline{x}={m}_{10}/{m}_{00}$, $\overline{y}={m}_{01}/{m}_{00}$. After normalization, the center distance is

$${\upeta }_{\mathrm{pq}}=\frac{{u}_{\mathrm{pq}}}{{{u}^{r}}_{00}},\mathrm{r}=\frac{\mathrm{p}+\mathrm{q}}{2}+1$$

(3)

Seven invariant moments are constructed by the second-order and third-order normalized central moments. The invariant moments are highly concentrated image features with translation, grayscale, scale, and rotation invariant for continuous images [4]. The feature value of the Hu moments is defined as below:

$$\begin{array}{l}{h}_{1}={\eta }_{20}+{\eta }_{02}\\ {h}_{2}={\left({\eta }_{20}+{\eta }_{02}\right)}^{2}+{4{\eta }_{11}}^{2}\\ \begin{array}{l}{h}_{3}={\left({\eta }_{30}-3{\eta }_{12}\right)}^{2}+{\left(3{\eta }_{21}-{\eta }_{03}\right)}^{2}\\ {h}_{4}={\left({\eta }_{30}+{\eta }_{12}\right)}^{2}+{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\\ \begin{array}{l}{h}_{5}=\left({\eta }_{30}-3{\eta }_{12}\right)\left({\eta }_{30}+{\eta }_{12}\right)\left[{\left({\eta }_{30}+{\eta }_{12}\right)}^{2}-\left.3{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\right] \right.\\ +\left(3{\eta }_{21}-{\eta }_{03}\right)\left({\eta }_{21}+{\eta }_{03}\right)\left[3{\left({\eta }_{30}+{\eta }_{12}\right)}^{2}-{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\right]\\ \begin{array}{l}{\mathrm{h}}_{6}=\left({\eta }_{20}-{\eta }_{02}\right)\left[{\left({\eta }_{30}+{\eta }_{12}\right)}^{2}-{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\right]+{4\eta }_{30}({\eta }_{30}+{\eta }_{12})({\eta }_{21}+{\eta }_{03})\\ {h}_{7}=\left(3{\eta }_{21}-{\eta }_{03}\right)\left({\eta }_{21}+{\eta }_{30}\right)\left[{\left({\eta }_{30}+{\eta }_{12}\right)}^{2}-\left.3{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\right] \right.\\ -\left({\eta }_{30}-{3\eta }_{21}\right)\left({\eta }_{21}+{\eta }_{03}\right)\left[{3\left({\eta }_{30}+{\eta }_{12}\right)}^{2}-\left.{\left({\eta }_{21}+{\eta }_{03}\right)}^{2}\right]\right.\end{array}\end{array}\end{array}\end{array}$$

(4)

3.2.2 Feature extraction of Fourier descriptor

Fourier descriptors can provide a good description of contour features, where only a few descriptors are needed to roughly represent the entire contour. After a simple normalization, the descriptor exhibits translation, rotation, and scale invariant, which makes the descriptor immune from the position, angle, and scaling of the contour in the images. Therefore, it is an image feature with good robustness [30].

The Fourier descriptor is a characteristic parameter describing the contour features of an image. The basic idea is to use the Fourier transform of the object boundary information as the shape feature and transform the contour feature from the space domain to the frequency domain, to extract the frequency domain information as the feature vectors of the image. The contour is digitized with a vector representation, where different contours can be better distinguished, and the object can be recognized [14].

For the contour from any point $\left({\mathrm{x}}_{0},{\mathrm{y}}_{0}\right)$, advancing counterclockwise will encounter $\left({\mathrm{x}}_{1},{\mathrm{y}}_{1}\right),\left({\mathrm{x}}_{2}{\mathrm{y}}_{2}\right)$, in the spatial domain describing the contour $D\left(k\right)=\left[x\left(k\right),y\left(k\right)\right]$, $k=\mathrm{0,2}\dots \mathrm{K}-1$, the set of its boundary coordinates. In addition, treat each coordinate point as a complex number as shown in Eq. (5). Under the complex number domain, the two-dimensional problem is simplified to a one-dimensional problem, but the essence of the boundary remains unchanged.

$$D\left(k\right)=x\left(k\right)+j\mathrm{y}\left(k\right)$$

(5)

The discrete Fourier transform of the coordinate sequence of the contour curve is shown in formula (6).

$$\mathrm{a}\left(u\right)=\sum_{k=0}^{\mathrm{K}-1}D\left(k\right){e}^{-j2\pi uk/K}$$

(6)

Where $u=\mathrm{0,1},2\dots \mathrm{K}-1$. The complex coefficient is $\mathrm{a}\left(u\right)$ means Fourier descriptor of the boundary. The coefficients ${\mathrm{K}}_{\mathrm{fd}}\left(\mathrm{i}\right)$ of the Fourier descriptor are normalized by dividing each magnitude of the coefficient by the 2-norm of a(1), $\left|\left|\mathrm{a}\left(1\right)\right|\right|$, which gives:

$${K}_{fd}\left(i\right)=\frac{\left|\left|\mathrm{a}\left(u\right)\right|\right|}{\left|\left|\mathrm{a}\left(1\right)\right|\right|}$$

(7)

3.2.3 SVM support vector machine

The SVM is a binary classification model whose basic model is the largest margin linear classifier defined on the feature space [6]. The basic idea of SVM is to solve the separation hyperplane that can correctly divide the training data set and has the largest geometric margin. For linearly separable data sets, there are infinitely many such hyperplanes (i.e., Perceptrons), but the separation hyperplane with the largest geometric margin is unique [27]. For the known data, the SVM maps data into the high-dimensional feature space by nonlinear mapping and uses the linear problem-solving method to find the optimal regression function in that space.

As the serial feature fusion method is simple and lite in computation, it is a relatively good choice for feature fusion where the sum of dimensions of multiple types of features is not large [26]. Assuming that the image has an n-dimensional feature value $\mathrm{\alpha }$ a and -dimensional feature value of $\upbeta$, the dimension of the new serial feature $\upomega$ is (n + m) dimension. The Hu moments and Fourier descriptors of the patent image are serially fused, normalized, and used as input to the SVM classification model to determine the accuracy of the shape recognition of the design patent target. The serial feature fusion is shown in Eq. (8) below,

$$K\left(I\right)=\left[\begin{array}{c}{\varepsilon }_{1}{K}_{hu}\left({I}_{\mathrm{hu}}\right)\\ {\varepsilon }_{2}{K}_{fd}\left({I}_{\mathrm{fd}}\right)\end{array}\right]$$

(8)

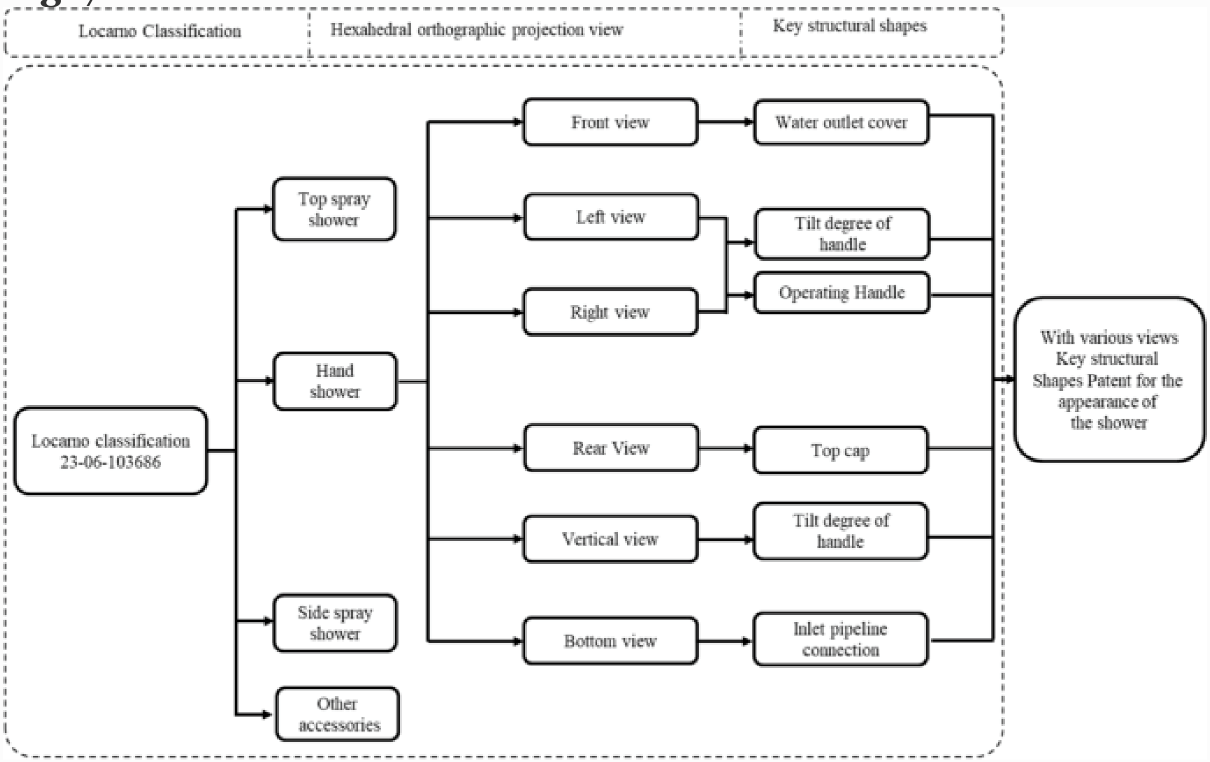
Where ${I}$ represents the image to be evaluated, represents the Hu moments feature of the image, and ${I}_{\mathrm{fd}}$ represents the Fourier descriptor feature of the image. The final fusion features are obtained by adjusting the weighting parameters ${\upvarepsilon }_{1}$ and ${\upvarepsilon }_{2}$.

3.3 Multi-image fusion

Although the Locarno Classification (LOC) has become a widely accepted product design classification standard in most countries, there are still some unreasonable issues with this classification standard. Specifically, this standard only sets a single and relatively small number of categories, while there are some cases where the boundaries of large categories are blurred and the range of small categories is too wide. Therefore, this article aims to redefine the classification standards for product appearance patent data, to improve the classification accuracy of product appearance patent images.

Based on the overall shape similarity of the same type of appearance patent, different shapes of key parts, and the use of multiple images to express the same object, product shape judgment, recognition, and verification are performed on multiple perspective images of the same patent to obtain accurate experimental data. Taking the appearance patent of a shower as an example, the perspective classification in this article is based on the fusion of multi-perspective image features of the shower appearance patent, by refining and classifying the key structural features of the multi-perspective image. Figures 7 show the appearance patent of a shower based on perspective classification. The principle of classification is to follow the key structural shapes in the perspective map and corresponding perspective maps. By classifying the key structural shapes in various perspectives of the patented image of the appearance of the shower, the differences between different patented images of the appearance of the shower can be effectively distinguished, which is conducive to the development of subsequent classification research.

Fig. 7



Patent for Shower Appearance Based on Perspective Classification

Full size image

4 Case analysis

4.1 Data collection

The research environment setup for this case study is Intel(R) Core (TM) i7-6700 CPU@3.40 GHz, 32.00 GB memory, Windows 10 operating system, and Python 3.8 programming software. The images in the sample library were selected from the Smart Buds patent database. From January 1, 2011, to December 31, 2020, 200 sets of shower design patents from some enterprises in China are chosen with a total of 8 types of showers with different shapes. Taking the six-sided views of the shower design patent (front view, back view, left view, right view, top view, bottom view) as the basic data set, the front view, top view, and bottom view were selected as the research objects.

Let I be a set of shower design patent images, and its spatial shape is ${\mathrm{X}}_{i}{\mathrm{Y}}_{i}{\mathrm{Z}}_{i}$. In the front view of the shower design patent, ${\mathrm{X}}_{i}$ is ${\mathrm{C}}_{\mathrm{C}}$ when the shower cover is round, and ${\mathrm{S}}_{\mathrm{S}}$ when it is square. The side view shape ${\mathrm{Y}}_{i}$, ${\mathrm{S}}_{\mathrm{R}}$ and ${\mathrm{S}}_{\mathrm{LC}}$ are used to indicate whether the side handle of the shower has obvious curvature. Among them,${\mathrm{S}}_{\mathrm{R}}$ is straight and ${\mathrm{S}}_{\mathrm{LC}}$ is curved. The bottom view shape ${\mathrm{Z}}_{\mathrm{i}}$ is ${\mathrm{C}}_{\mathrm{C}}$ and ${\mathrm{S}}_{\mathrm{Rc}}$, respectively indicating that the bottom view shower's bottom shape is a round or Rounded rectangle. As shown in Fig. 8, the shower samples from left to right are, the round cover and straight cylindrical handle shower ${\mathrm{C}}_{\mathrm{C}}{\mathrm{S}}_{\mathrm{R}}{\mathrm{C}}_{\mathrm{C}}$; the square cover straight cuboid handle shower ${\mathrm{S}}_{s}{\mathrm{S}}_{\mathrm{R}}{\mathrm{S}}_{\mathrm{Rc}}$; the trapezoidal cover curved cylindrical handle shower ${\mathrm{C}}_{\mathrm{C}}{\mathrm{S}}_{\mathrm{LC}}{\mathrm{C}}_{\mathrm{C}}$; the round cover straight cuboid handle shower ${\mathrm{C}}_{\mathrm{C}}{\mathrm{S}}_{\mathrm{R}}{\mathrm{S}}_{\mathrm{Rc}}$.

Fig. 8

Shower sample. **a** ${C}_{C}{S}_{R}{C}_{C}$. **b** ${\mathrm{S}}_{S}{S}_{R}{S}_{Rc}$. **c** ${\mathrm{C}}_{\mathrm{C}}{\mathrm{S}}_{\mathrm{LC}}{\mathrm{C}}_{\mathrm{C}}$. **d** ${C}_{C}{S}_{R}{S}_{Rc}$

Full size image

4.2 Feature extraction, matching, and recognition of shower images

There are two steps in feature extraction and matching: eigenvalue extraction and feature matching. In this paper template matching method based on Hu's moments and shape recognition method based on Fourier descriptors are used.

4.2.1 Hu moment feature extraction and matching

First, five images are randomly collected to build a contour template library from view-based classifications of each type of shower. The Hu moments of the three-view image of the shower are extracted with similarity checked with the corresponding viewing template of the template library. The results are shown in Table 1.

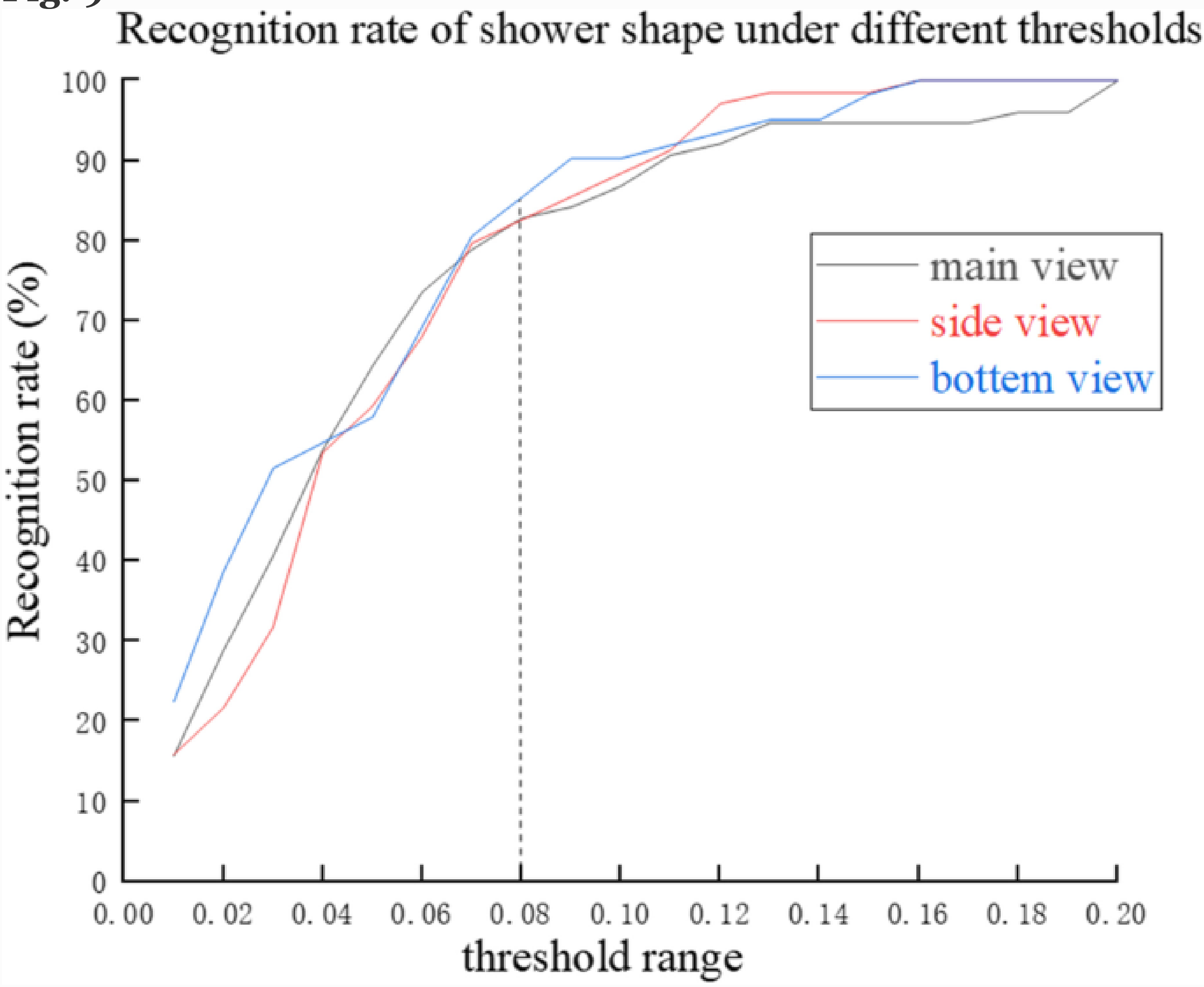
Table 1 Image matching coefficient

Full size table

Where $\mathrm{X}_{\mathrm{i}}, \mathrm{Y}_{\mathrm{i}}, \mathrm{Z}_{\mathrm{i}}$ represents the shape of the main view, side view, and top view. When the image matching coefficient is smaller, the contours are more similar; On the contrary, the larger the matching coefficients, the lower the matching degree.

The threshold value of the image matching coefficient is set to 0.01 ~ 0.2, and the recognition results of the three views of the shower are analyzed to calculate the number of correct and incorrect classifications of the three views. The recognition is considered a failure if the matching coefficient is greater than the threshold, and success if it is smaller than the threshold. The shape recognition rate is shown in Fig. 9.

Fig. 9



The recognition rate of shower shape under different thresholds

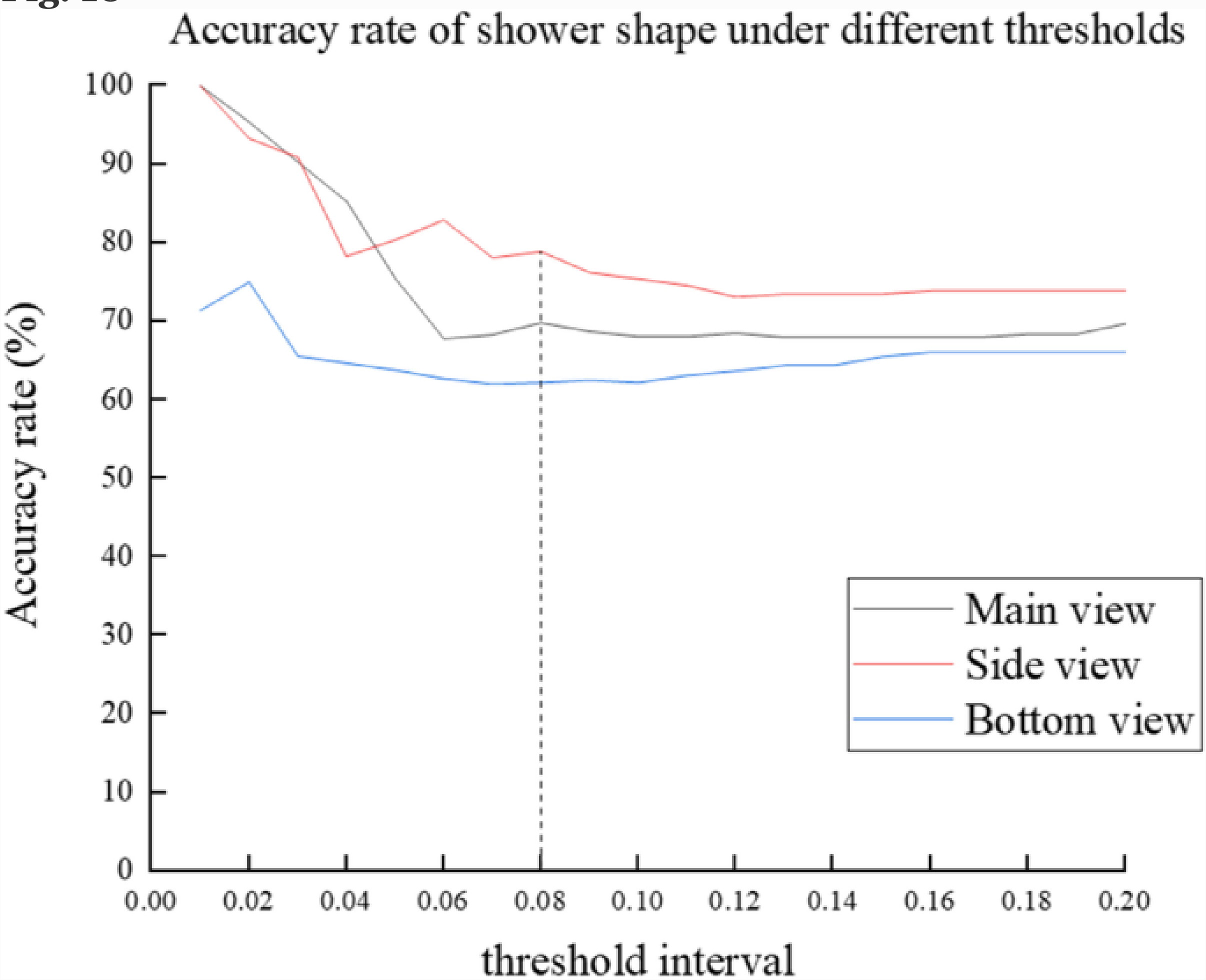
Full size image

The recognition accuracy is marked as 0 if it belongs to the first category in the template library; otherwise, it is marked as 1 for the second category with the first category set as negative samples. The TP, TN, FN, and FP values are recorded for the three-view images. The evaluation criteria of the experiment results are shown in Table 2 below. The recognition accuracy of shower shapes with different thresholds are shown in Fig. 10.

Table 2 Test result evaluation standard table

Full size table

Fig. 10



Accuracy of shower shape discrimination under different thresholds

Full size image

To verify the effectiveness of the method, patented images of shower designs from January 1, 2019, to December 31, 2020, were selected as test data. In the case of relatively ideal contour extraction, the number of successful identifications increased with more mismatches as the threshold value increased. When the threshold value is too small, the recognition success rate is low but the accuracy rate is high, and the number of successfully recognized graphics is low. Based on Figs. 9 and 10, 0.08 was chosen as the threshold value as a trade-off between recognition rate and accuracy. Accuracy P, recall R, and F-score were used to evaluate the performance of the model. The results are shown in Table 3.

Table 3 Recognition performance evaluation statistics

Full size table

$$P=\frac{TP}{TP+FP}$$
(9)

$$R=\frac{TP}{TP+FN}$$
(10)

$$F=\frac{2\times P\times R}{P+R}$$
(11)

From Table 4, it can be observed that the method based on Hu moment variable with the shower contour as the feature can retrieve the feature information of the showers and match relatively accurately. The best recognition is obtained for the side view and relative worse recognition is observed for the bottom view. The segmentation boundary of the side view is clearer, and the overall shape is used for matching; due to the positioning error and pre-processing of the bottom of the shower in the bottom view, the distinction is weaker compared with the other two views.

Table 4 Fourier descriptor after normalization

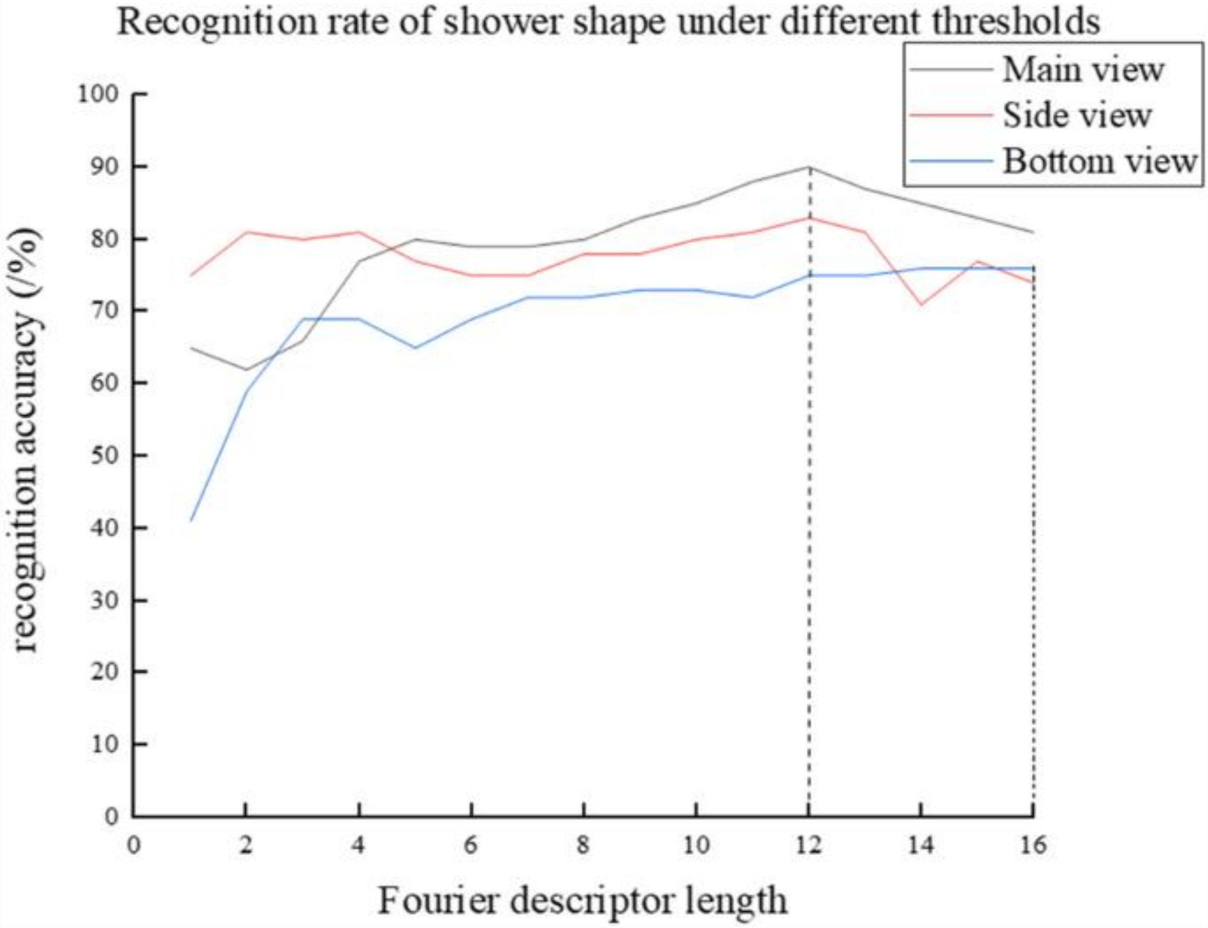
Full size table

4.2.2 Fourier descriptor recognition

When Fourier coefficients are used to describe the shape, they are invariant to rotation and translation and are independent of the choice of the starting point of the profile. The low-frequency component of the Fourier descriptor can better reflect the overall shape of the shower, while the high-frequency component can better reflect the details of the shower [5]. An appropriate length needs to be determined when the feature vector is being identified. The Fourier coefficients can be normalized according to these characteristics, with the normalized results shown in Table 4.

Out of 200 sets of sample shower design patent data, 150 sets are randomly selected as training data, with the remaining 50 sets being test data. To reflect the classification of Fourier descriptors for different views of the patent images, the Fourier descriptors of different lengths are used as inputs to the SVM for recognition with accuracy shown in Fig. 11.

Fig. 11



Recognition accuracy of Fourier descriptors with different feature vector lengths

Full size image

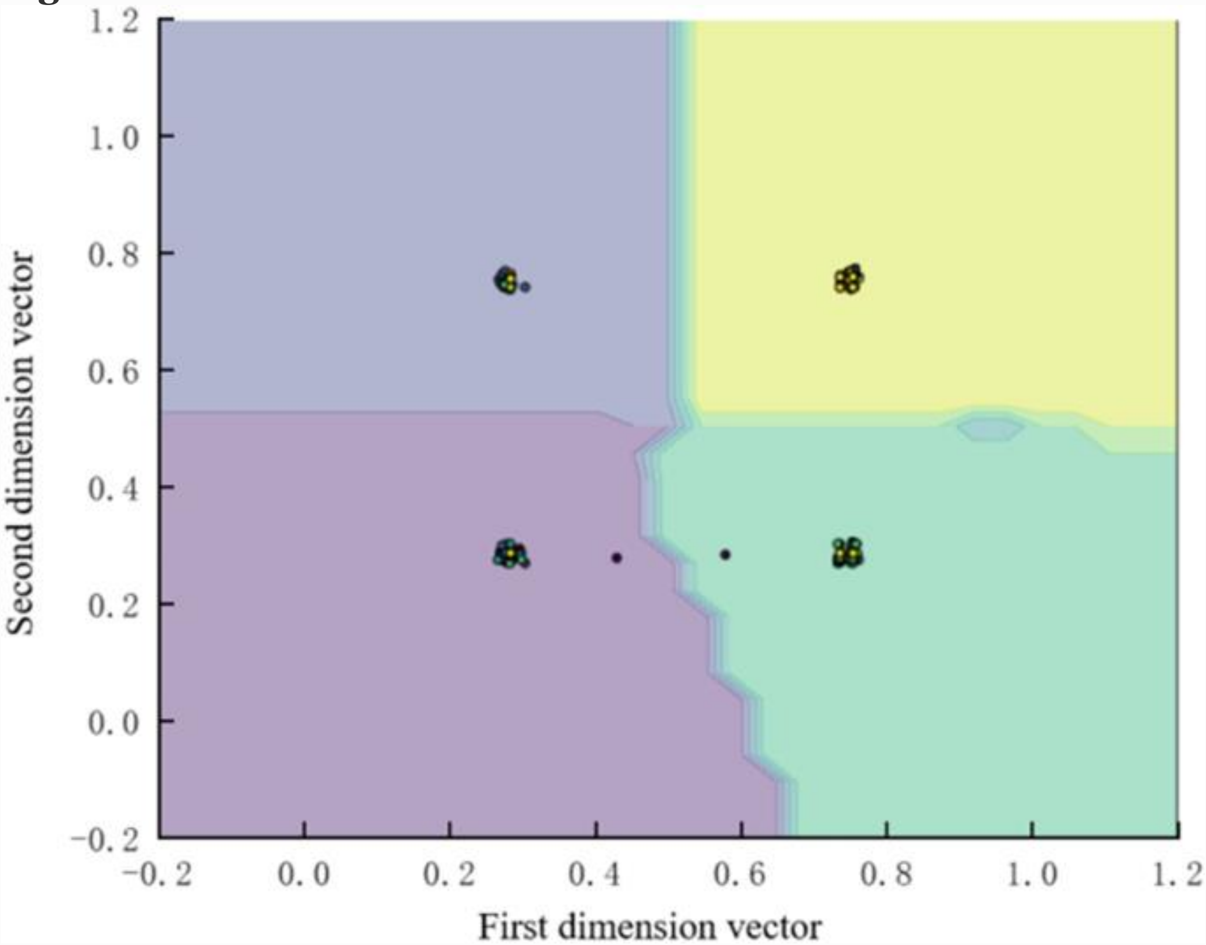
It can be seen from Fig. 11 that the recognition accuracy peaks when the Fourier descriptor lengths are set as 12, 12, and 16, respectively, for the front view, side view, and bottom view. Therefore, the front view and the side view of the Fourier descriptor length are selected as 12, and the bottom view is selected as 16.

4.3 Object shape classification and multi-image fusion

The SVM is a binary classifier, which is a "one-to-many" method to achieve multi-target classification based on OVR(One-vs-Rest). During each training, one category is chosen as the positive sample, with the remaining categories being negative samples, and for the above operation iteration. [26] The grid search method is used to obtain the optimal hyperparameters of the model, the error penalty coefficient c, and the kernel function parameter r. Then, a balanced estimate of the current hyperparameters is performed by dividing the training set into different divisions through k-fold cross-validation.

The SVM is used to train the shower design patent image samples to obtain a classification model for the shape of the shower patent images for different views. Figure 12 shows the prediction of the SVM model under the two-dimensional Hu moments coefficient vector.

Fig. 12



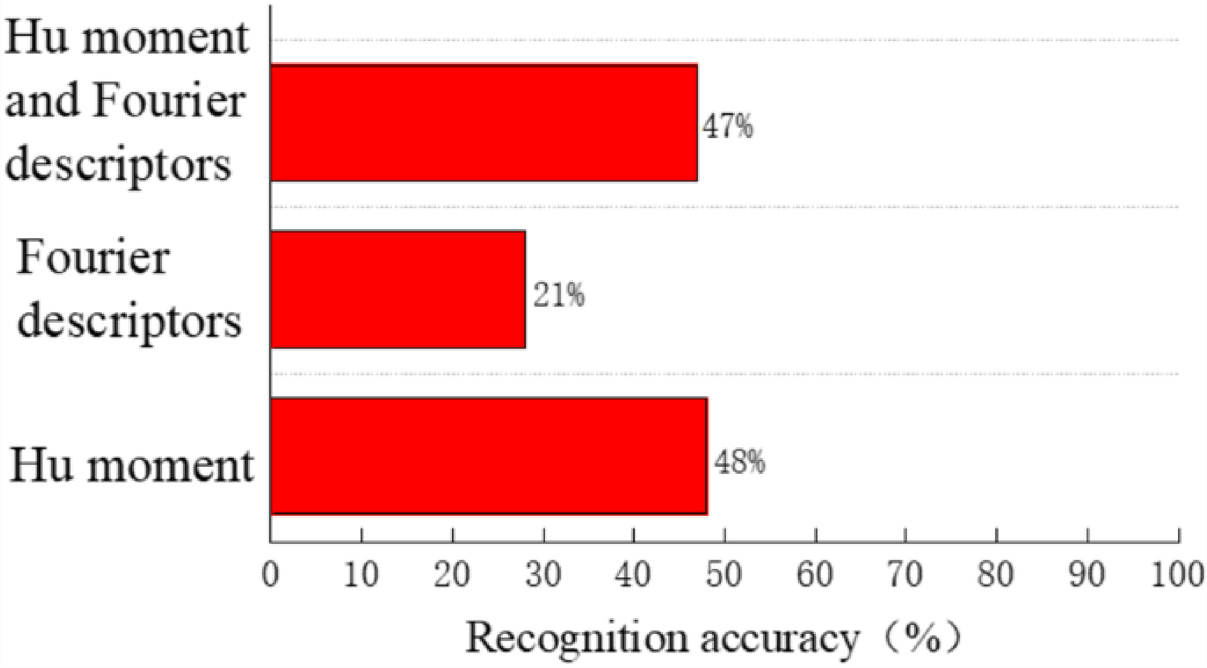
The prediction of the SVM classification model under the two-dimensional Hu moments matching coefficient vector

Full size image

To determine the reasonable input features of the SVM model, a comparison experiment of different feature recognition effects is performed. The extracted Hu moments feature, Fourier descriptors, and Hu moments + Fourier descriptors serial fusion vector are used as input to obtain the prediction results. Finally, the spatial shape of the shower is obtained by combining the shapes of the patent images from different views using multi-image fusion rules and the SVM classification results.

Figure 13 shows the recognition accuracy of the Hu moments feature can reach a highest of 48%.

Fig. 13



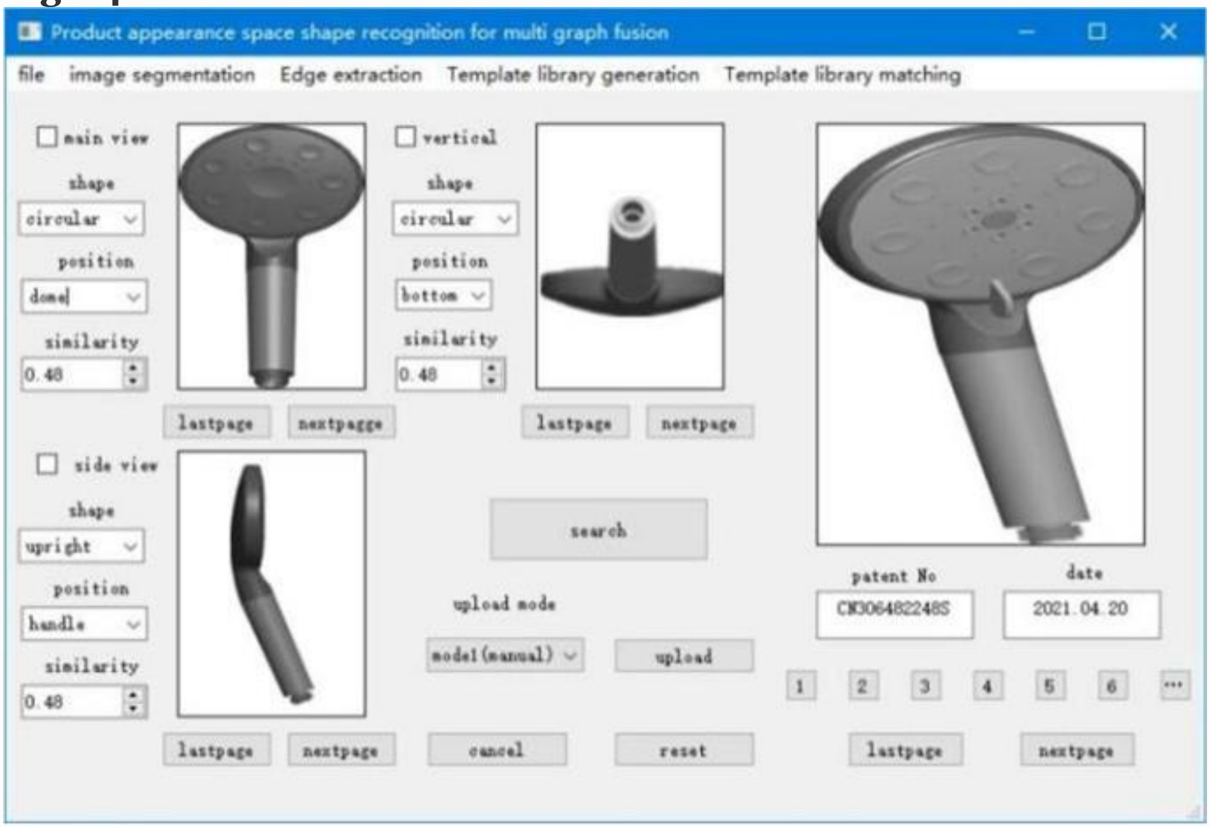
Multi-image fusion prediction situation of shower patent images

Full size image

4.4 Patent recommendation

The main program of the design patent search system is visualized based on the QT platform of the Windows system, and the graphical user interface (GUI) is shown in Fig. 14 below.

Fig. 14



Multi-image fusion-oriented design patent spatial shape recognition interface display

Full size image

The system provides the following function modules: image uploading module for retrieval, image segmentation module, edge extraction module, template library generation module, a template matching module, a sample retrieval module, etc.

Due to a large amount of shower design patents, additions, deletions, and changes are frequently needed while storing these design patents. The highly stable MySQL is used as the database management software, which is divided into two parts: one is the private database of all users, where users can store self-uploaded data; the other is the patent image data from the authoritative database, which can be used for client research or as the feature template library for shower comparisons.

An initiative foreground extraction capability is provided when users upload images. When searching for shower patent images, users can select the shape of a specified part and the corresponding views and adjust their similarity to narrow the search range to better meet their needs.

5 Conclusion

With the advancement of national patent laws and the growing legal awareness among individuals and companies, there has been a surge in the number of appearance patent applications. Patent agents are required to sift through vast databases to identify similar patents, review newly filed patents, and determine their potential for success. Similarly, designers must also consult patent databases to improve their designs and prevent duplication. Given the significance of images in design patents, text-based searches are inadequate, as different images may have the same textual description.

Spatial shape recognition for product design patents based on multi-image fusion is typically accomplished in three stages: pre-processing and image segmentation; feature vector extraction and recognition; and multi-image fusion.

- The characteristics of the patented image make the accurate extraction of contours very difficult, and it is very easy to miss segmentation, over-segmentation, and curves are not smooth enough. The method of masking the edge image with the original image can avoid similar situations and greatly improve the accuracy for the subsequent extraction of the outermost contours of the shower as features.

- The Fourier descriptors and Hu moments are used as SVM inputs to achieve multiple classifications of structures.
- By analyzing the characteristics of the patented shower appearance image, the six-view contour features of the patented shower appearance image are analyzed. Fourier descriptors and Hu variable moments are collected, a template library is constructed, and the accuracy of the single-view structure is verified.

To address the specificity of design patent images, a multi-feature perspective fusion algorithm for design patent image retrieval is proposed. The algorithm leverages Fourier descriptors and Hu variable moments to extract shape features from different views, incorporating shape contour features and view variations. The algorithm then employs a weighted fusion of the features extracted by both methods to serve as the retrieval features for design patent images.

With the integration of the multi-image feature fusion, the spatial shape of the product's appearance can be accurately determined. While this proposed method has shown significant improvement in recognizing the appearance of patent images of shower products, there is still a lot of potential for further enhancement in areas such as data set scale, feature, fusion techniques, contour selection, and machine learning algorithms. Further improvements will aim to improve the accuracy of appearance patent image recognition by refining our feature fusion method and increasing the quantity and quality of the data volume.

References

1. Arslan S, Yazıcı A, Saçan A, Toroslu IH, Acar E (2013) Comparison of feature-based and image registration-based retrieval of image data using multidimensional data access methods. Data Knowl Eng 86:124–145

   Article Google Scholar
2. Cai N, Zhang GH, Lou PX et al (2011) Image retrieval for a design patent base on shape features and texture features. Journal of Shandong University (Engineering Science) 4(2):1–4

   Google Scholar

3. Csurka G, Renders JM, Jacquet G (2011) XRCE's participation at patent image classification and image-based patent retrieval tasks of the Clef-IP 2011. In: Petras V, Forner P, Clough PD (eds.) CLEF (Notebook Papers/Labs/Workshop)

4. Duan W, Kuester F, Gaudiot J et al (2008) Automatic object and image alignment using Fourier Descriptors. Image Vis Comput 26(9):1196–1206

   Article Google Scholar

5. El-ghazal A, Basir O, Belkasim S et al (2012) Invariant curvature-based Fourier shape descriptors. J Vis Commun Image Represent 23(4):622–633

   Article Google Scholar

6. Gao J, Xu L, Huang F, Man Z, Huang G, Man Z, Huang G (2015) A spectral–textural kernel-based classification method of remotely sensed images. Neural Comput Appl 27(2):431–446

   Article Google Scholar

7. Huet B, Guarascio G, Kern NJ, Mérialdo B (2001) Relational skeletons for retrieval in patent drawing. In the Proceedings of the 2001 International Conference on Image Processing. IEEE, pp 737–740

8. Ji F, Qingyun D (2004) An Automatic Retrieval System for Appearance Patents Based on Image Content. Comput Eng Appl 34:209–2011

   Google Scholar

9. Jiang S, Luo J, Ruiz-Pava G, Hu J et al (2021) Deriving Design Feature Vectors for Patent Images Using Convolutional Neural Networks. J Mech Des 143(6):1–46

   Article Google Scholar

10. Junyong Z, Qi H, Shouqian S (2004) Research and Application of Content-Based Appearance Design Patent Retrieval Technology. In Proceedings of the 2004 International Conference on Industrial Design: Mechanical Industry Press, pp 154–158

11. Li M, Li H (2022) Application of deep convolutional neural network under region proposal network in patent graphic recognition and retrieval. IEEE Access 10:37829–37838. https://doi.org/10.1109/ACCESS.2021.3088757

Article Google Scholar

12. Li L, Liu Y (2015) Wavelet transforms image retival method based on content. Comput Sci 42(2):306–310

Google Scholar

13. Lingyun S, Shouqian S (2008) Patent Analysis Technology for Product Appearance Design Based on Patent Images. Comput Integr Manuf Syst 2:234–240

Google Scholar

14. Liu R, Zou H, Zhang L et al (2012) An improved Laplacian SVM algorithm for SAR image segmentation: an improved Laplacian SVM algorithm for SAR image segmentation. J Infrared Millimeter Waves 30(3):250–254

Article Google Scholar

15. Lu YW, Jiang JG, Qi MB et al (2017) Segmentation method for medical image based on improved GrabCut. Int J Imaging Syst Technol 27(4):383–390

Article Google Scholar

16. Mei W, Yujian Li, Xiaomei Q (2014) A Survey on Graph Theory Approaches of Image Segmentation. Comput Appl Softw 31(9):1–12

Google Scholar

17. Monge-Alvarez J, Hoyos-Barcelo C, Lesso P et al (2019) Robust Detection of Audio-Cough Events Using Local Hu Moments. IEEE J Biomed Health Inform 23(1):184–196

Article Google Scholar

18. Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst Man Cybern 9(1):62–66

Article Google Scholar

19. Qiliang Z, Shengyong C, Haigen H et al (2020) An instance segmentation scheme combining multiple image segmentation algorithms. Small Microcomput Syst 41(4):837–842

Google Scholar

20. Qingqing L, Changsheng Z, Xueqiang L et al (2016) Multimodal Image Retrieval Based on Appearance Design Patents. Comput Eng Des 37(9):2469–2474

Google Scholar

21. Qingyun D, Haipeng Li (2002) A retrieval method for appearance design patent images based on texture and shape features. Comput Eng Appl 3:27–29

Google Scholar

22. Senhong W (2013) Research on Classification Methods for Appearance Patent Images. Guangdong University of Technology, Guangzhou

Google Scholar

23. Shuangshuang H, Qingyun D (2013) Research on Appearance Patent Image Retrieval Algorithm Based on Texture and Shape Features. Microcomput Appl 32(4):42–44, 47

Google Scholar

24. Smeulders AWM, Worring M, Santini S et al (2000) Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Mach Intell 22(12):1349–1380

Article Google Scholar

25. Ting Z (2012) Research on image feature extraction technology in appearance patent search. Huazhong University of Science and Technology, Wuhan

Google Scholar

26. Yang A, Bai Y, Liu H et al (2022) Application of SVM and its improved model in image segmentation. Mobile Networks and Applications 27(3):851–861

Article Google Scholar

27. Yang Z, Wang C, Peng K (2021) Crankshaft intelligent recognition method based on deep support vector machine. Comput Integr Manuf Syst 27(6):1629–1640

Google Scholar

28. Li Y (2014) Research on intelligent retrieval algorithms for appearance design patents. Chinese Inventions and Patents 10:66–68

Google Scholar

29. Zhang B, Zhang Y, Liu J, Wang B (2021) FGFF Descriptor and Modified Hu Moment-Based Hand Gesture Recognition. Sensors (Basel, Switzerland) 21(19):6525

Article Google Scholar

30. Zheng Y, Meng F, Liu J, Guo B, Song Y, Zhang X, Wang L (2020) Fourier Transform to Group Feature on Generated Coarser Contours for Fast 2D Shape Matching. IEEE Access 8:90141–90152

Article Google Scholar

31. Zhiyuan Z, Juan Z, Bin X (2007) An outward-appearance patent-image retrieval approach based on the contour-description matrix. In: FCST, pp. 86–89. IEEE Computer Society, Washington, DC, USA