

## Research Article

# Ancient Chinese Character Image Retrieval Based on Dual Hesitant Fuzzy Sets

Songbo Du , Fang Yang , and Xuedong Tian 

*School of Cyber Security and Computer, Hebei University, Baoding 071002, China*

Correspondence should be addressed to Xuedong Tian; [xuedong\\_tian@126.com](mailto:xuedong_tian@126.com)

Received 14 December 2020; Revised 16 January 2021; Accepted 29 January 2021; Published 22 February 2021

Academic Editor: Shah Nazir

Copyright © 2021 Songbo Du et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The complex and changeable structures of ancient Chinese characters result in the decreasing accuracy of their image retrieval. To resolve this problem, a new retrieval method based on dual hesitant fuzzy sets is proposed. Dual hesitation fuzzy sets that can express uncertain information more comprehensively are employed in the feature extraction process of directional line elements. The multiattribute evaluation index of adjacent grids for the current grid and its corresponding membership and nonmembership functions are established, and the weight of each attribute is calculated by the dual hesitation fuzzy entropy, such that the proposed features can fully reflect the topological structure of ancient Chinese characters. Using the dual hesitation fuzzy correlation coefficient to measure the similarity between the ancient Chinese character images to be retrieved and the candidate images, the retrieval of ancient Chinese character images is realized. Experiments show that when the threshold value of the correlation coefficient is 0.9, the average retrieval accuracy is 90.4%.

## 1. Introduction

Image retrieval of ancient Chinese characters can assist researchers in tracing similar glyphs in the research process. It is an effective tool for the study of Chinese characters in ancient books.

The theoretical basis of image retrieval of Chinese characters in ancient books is content-based image retrieval [1, 2]. The key link is feature extraction and matching, that is, color, texture, shape, and other features [3–5] or their fusion features [6] are extracted through mathematical descriptions, and similarity calculation is conducted based on image features to identify images that are similar. Traditional Chinese character feature extraction methods mainly adopt structural features [7] and statistical features [8]. The fuzzy set theory has a good ability to represent uncertain problems; thus, it is applied to the recognition and retrieval of handwritten Chinese characters. Wei and Guo [9] extracted image features using dual-elastic grid technology and the correlation fuzziness between grid words. Ran et al. [10] proposed a normalized overlapped fuzzy bielastic grid, which was used to improve the effectiveness of the proposed

features. Li et al. [11] used fuzzy entropy to classify Chinese characters with high accuracy and improved the recognition rate of Chinese characters. Liu and Meng [12] improved the membership function of a fuzzy support vector machine and improved the ability of text classification and recognition efficiency.

The fuzzy set theory improves the efficiency of Chinese character recognition and retrieval to some extent. This theory expresses the singularity of fuzzy information, which causes information loss. Therefore, a hesitant fuzzy set [13] is proposed and applied to deal with multiattribute decision-making (MADM) problems. The dual hesitant fuzzy set (DHFS) [14] is an extended form of the hesitant fuzzy set that combines the membership and nonmembership degrees of multiattribute evaluation information to express the uncertainty problem more comprehensively. Su et al. [15] proposed some distance measures and similarity measures and illustrated their applications in pattern recognition. Wang et al. [16] proposed a new dual hesitation fuzzy distance metric for the multiattribute decision-making problem with completely unknown attribute weights. Zhang [17] proposed several distance measurements and entropy

measurement methods for dual hesitation fuzzy sets, which avoided the process of data expansion and overcame the problem of information loss to a certain extent. Combining the DHFS, Wei et al. [18] proposed some dual hesitant Pythagorean fuzzy Hamy mean aggregation operators, which are more valid for dealing with the MADM problem. Wang et al. [19] extended the q-rung orthopair fuzzy sets (q-ROFSs) to a dual hesitant fuzzy environment and presented the dual hesitant q-ROFSs. Tang and Wei [20] defined some dual hesitant Pythagorean fuzzy generalized Bonferroni mean operators which are utilized to design some methods to handle the MADM problems. Tang and Wei [21] investigated MADM problems based on Bonferroni mean operators with dual Pythagorean hesitant fuzzy numbers.

In view of the complex and changeable structure of ancient Chinese characters, this study proposes an image retrieval method for ancient Chinese books based on dual hesitant fuzzy sets. The membership and nonmembership degrees between adjacent elastic grids are calculated under the set attribute index, and the obtained evaluation values are placed in a dual hesitation fuzzy set, which can be used to extract the features of directional line elements. The dual hesitant fuzzy correlation coefficient is used to evaluate the similarity between the query image and the candidate images to realize the image retrieval of ancient Chinese characters.

## 2. Dual Hesitant Fuzzy Attribute Index

*2.1. Elastic Grid Division of Chinese Characters in Ancient Books.* Ordinary elastic mesh [22] does not evenly distribute the pixel density in each grid and has a low tolerance for feature mutations caused by stroke offset. In this study, the general elastic mesh is improved and a new elastic mesh division method is designed. The Chinese character image is divided into elastic grids according to the pixel density in a certain direction. Subsequently, each layer is divided twice according to the pixel density. The grid lines after the second division may not be connected to straight lines, as shown in Figure 1. The grid in Figure 1(a) is used to extract features of horizontal strokes of Chinese characters in ancient books. The grid in Figure 1(b) is used to extract features of vertical strokes. The grid in Figure 1(c) is used to extract features of apostrophe strokes. The grid in Figure 1(d) is used to extract features of downstrokes.

The partition algorithm of Figure 1(a) is shown in Algorithm 1. Similarly, other grid division algorithms can be obtained.

*2.2. Dual Hesitant Fuzzy Attribute Index Setting.* Considering the correlation between adjacent grids and the grid to be calculated, the corresponding membership and nonmembership functions are provided under three attribute indexes. Using the extraction of horizontal stroke features as an example, the definition process of attribute indexes and their membership and nonmembership degrees is illustrated in Figure 2. In Figure 2, the six neighborhoods extracted from the left and right sides of any grid are  $G_k^{(i)}$  ( $i = 1, 2, \dots, 6$ ).

*2.2.1. Pixel Distance Index.* In Figure 3,  $a$  is the pixel point in any horizontal stroke in the grid:

$$\mu(d_a) = \exp\left(-\frac{(3d_a/W_k^{(2)})^2}{2}\right). \quad (1)$$

The standard normal distribution is used to characterize the membership and nonmembership of point  $a$  to  $G_k$  [9]. Subsequently, the membership and nonmembership functions under the pixel distance index are

$$h_1(G_k, G_k^{(2)}) = \frac{1}{n} \sum_{i=0}^n \mu(d_i), \quad (2)$$

$$g_1(G_k, G_k^{(2)}) = \frac{1}{n} \sum_{i=0}^n \mu(d_i - W_k^{(2)}), \quad (3)$$

where  $W_k^{(2)}$  is the width of  $G_k^{(2)}$ ,  $d_a$  is the distance between pixel point  $a$  and the left boundary of  $G_k^{(2)}$ , and  $n$  is the number of horizontal stroke pixel points in  $G_k^{(2)}$ .

*2.2.2. Stroke Position Index.* In Figure 4, stroke  $b$  intersects with the left edge of  $G_k$ , and stroke  $c$  is separated from the left edge of  $G_k$ . The membership and nonmembership functions under the stroke position index are

$$h_2(G_k, G_k^{(2)}) = \frac{n_1}{n_1 + n_2}, \quad (4)$$

$$g_2(G_k, G_k^{(2)}) = \frac{n_2}{n_1 + n_2}, \quad (5)$$

where  $n_1$  is the number of strokes intersecting with  $G_k$  and  $G_k^{(2)}$ , and  $n_2$  is the number of strokes that do not intersect with  $G_k$ .

*2.2.3. Grid Position Index.* In Figure 5, the left boundary of  $G_k$  and the right boundary of  $G_k^{(2)}$  overlap, and the left boundary of  $G_k$  and the right boundary of  $G_k^{(3)}$  partially overlap. The membership and nonmembership functions under the grid position index are

$$h_3(G_k, G_k^{(3)}) = \frac{H_o}{H_k^{(3)}}, \quad (6)$$

$$g_3(G_k, G_k^{(3)}) = \frac{H_k^{(3)} - H_o}{H_k^{(3)}}, \quad (7)$$

where  $H_o$  is the overlap height between the right edge of  $G_k^{(3)}$  and the left edge of  $G_k$  and  $H_k$  is the height of  $G_k$ .

Under each attribute index, the membership and nonmembership degrees of each grid are calculated by the aforementioned function, and the different numbers of Chinese character images are statistically analyzed with the average membership and average nonmembership degrees of three indexes.

In Figure 6,  $hi$  ( $i = 1, 2, 3$ ) are the average membership of three indexes and  $gj$  ( $j = 1, 2, 3$ ) are the average nonmembership of three indexes. As can be seen from Figure 6,  $h3$  is

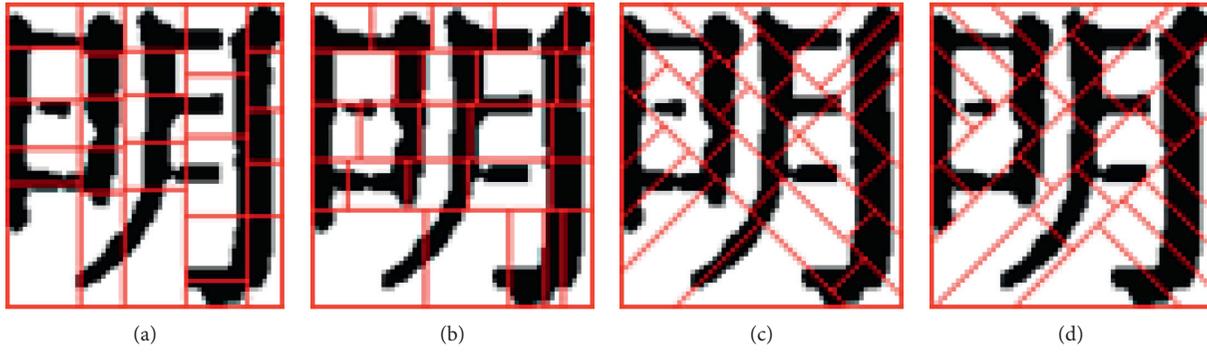


FIGURE 1: 5 × 5 elastic grid.

```

Input: Chinese character image (size:  $M \times N$ )
Output: vertical and horizontal grid lines ( $J_k$  and  $I_{k_z}$ )
(1). for  $i=0$  to  $M-1$ //traverse the whole image pixels
(2).   for  $j=0$  to  $N-1$ 
(3).      $DV = DV + Density[i][j]$ //accumulate pixel density
(4).     if  $DV > Sum * (k + 1)/5$  then//the image is divided into 5 areas in the horizontal direction,  $Sum$  is the total number of pixels
(5).        $J_k = j$ ;  $k = k + 1$ //the initial value of  $k$  is 0
(6).     end if
(7).   end for
(8). end for
(9). for  $i=0$  to  $M-1$ //traverse the pixels of each area of the above partition
(10).   for  $j = J_k$  to  $J_{k+1}$ 
(11).      $DH = DH + Density[i][j]$ //accumulate pixel density
(12).     if  $DH > Sum * (z + 1)/25$  then//the image is divided into 25 areas
(13).        $I_{k_z} = i$ ;  $l = l + 1$ ;  $z = z + 1$ //the initial values of  $l$  and  $z$  are 0
(14).     end if
(15).   end for
(16). end for
(17). return  $J_k, I_{k_z}$ 
    
```

ALGORITHM 1: Elastic grid division algorithm.

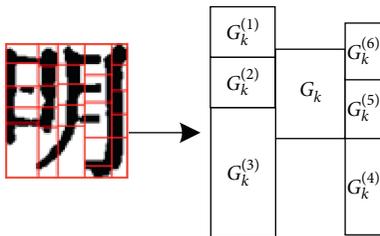


FIGURE 2: Schematic diagram of six neighborhoods of  $G_k$ .

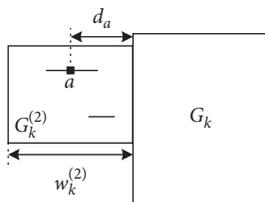


FIGURE 3: Schematic diagram of pixel distance index.

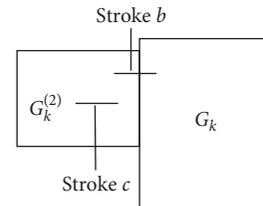


FIGURE 4: Schematic diagram of stroke position index.

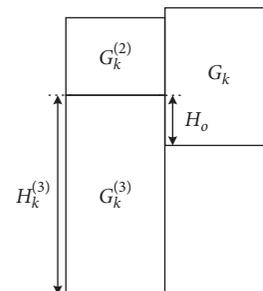


FIGURE 5: Schematic diagram of grid position index.

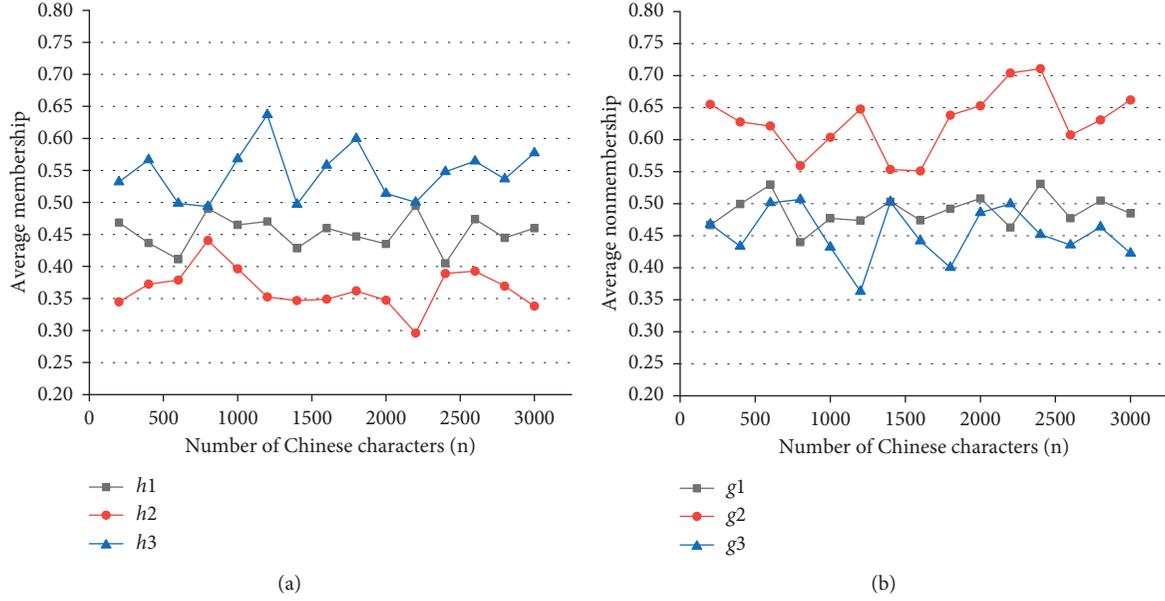


FIGURE 6: Average (a) membership degree and (b) nonmembership degree of different numbers of Chinese characters.

higher than  $h1$  and  $h2$ , and, from Figure 6(b),  $g2$  is higher than  $g1$  and  $g3$ . Through the aforementioned analysis, it can be concluded that the evaluation information of the grid has considerable differences under different indexes, and the three attribute indexes cannot be granted equal weight when the membership and the nonmembership degrees are calculated. In this study, the weight of the grid attribute index is calculated using dual hesitant fuzzy entropy, which improves the authenticity of the evaluation information.

**2.3. Determination of the Attribute Index Weight.** Dual hesitant fuzzy entropy can effectively describe the degree of uncertainty of dual hesitant fuzzy elements [23]; it is defined as

$$E_d = \frac{2\phi(d)\varphi(d) + \pi(d)}{\phi(d)^2 + \varphi(d)^2 + \pi(d)}, \quad (8)$$

where  $\phi(d)$ ,  $\varphi(d)$ , and  $d$  refer to the membership, nonmembership, and hesitancy of any dual hesitant fuzzy element, respectively.

The entropy theory can describe the degree of uncertainty of each attribute index to determine the attribute weight. Equation (9) is used to calculate the entropy  $E_j$  of the attribute  $\omega_j$  ( $j = 1, 2, 3$ ) of the grid  $G_k^{(i)}$ :

$$E_j = \frac{1}{m} \sum_{i=1}^m E_{d_{ij}}, \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, n). \quad (9)$$

At this point,  $E_j$  and  $\omega_j$  are inversely proportional. The larger the  $E_j$ , the smaller the  $\omega_j$ ; the smaller the  $E_j$ , the larger the  $\omega_j$ . Therefore, equation (10) can reasonably describe the relationship between  $E_j$  and  $\omega_j$ :

$$\omega_j = \frac{1 - E_j}{n - \sum_{j=1}^n E_j}. \quad (10)$$

### 3. Dual Hesitant Fuzzy Direction Line Element

**3.1. Dual Hesitant Fuzzy Set.** Based on the hesitant fuzzy set, the dual hesitant fuzzy set fuses the membership and nonmembership degrees of multiple attribute indexes to improve its ability to express uncertain problems in decision-making. Its definition is as follows:

$$D = \{ \langle x, h_{(x)}, g_{(x)} \rangle | x \in X \}, \quad (11)$$

where  $h_{(x)}$  and  $g_{(x)}$  are sets of some numbers in  $[0, 1]$ , respectively, representing the membership and nonmembership degrees of element  $x$  in the nonempty set  $X$  under  $D$  [12].

**3.2. Feature Extraction of Dual Hesitant Fuzzy Direction Line Element.** The traditional feature vector of the directional line element only involves the set of directional attributes of strokes in a single grid; it does not consider the correlation between strokes in adjacent grids and in the grid to be computed, thus, affecting the stability of the proposed feature. Through equations (2)–(6), the evaluation value of  $G_k^{(i)}$  ( $i = 1, 2, \dots, 6$ ) to  $G_k$  under each attribute index is calculated, and the dual hesitation fuzzy set  $\mu_{G_k^{(i)}}$  is constructed. The weighted correlation coefficient between  $G_k^{(i)}$  ( $i = 1, 2, \dots, 6$ ) and the ideal grid at the corresponding position is calculated [24], as shown in equation (12), which represents the degree of influence of adjacent grids on the current grid:

$$\rho(\mu_{G_k^{(i)}}, \mu_{G_k^{(*)}}) = \frac{C(\mu_{G_k^{(i)}}, \mu_{G_k^{(*)}})}{[C(\mu_{G_k^{(i)}}) \cdot C(\mu_{G_k^{(*)}})]} = \frac{\sum_{j=1}^n \omega_j \left( (1/k_j) \sum_{s=1}^{k_j} \gamma_{ij}^s \right)}{\sqrt{\sum_{j=1}^n \omega_j \left( (1/k_j) \sum_{s=1}^{k_j} (\gamma_{ij}^s)^2 + (l/l_j) \sum_{s=1}^{l_j} (\eta_{ij}^s)^2 \right)}} \quad (12)$$

where  $\mu_{G_k^{(+)}}$  is the dual hesitant fuzzy set of the ideal mesh pair and the set element is  $d_i^+ = \{\{1\}, \{0\}\}$ ,  $C(\mu_{G_k^{(i)}}, \mu_{G_k^{(+)}})$  is the covariance of  $\mu_{G_k^{(i)}}$  and  $\mu_{G_k^{(+)}}$ ,  $\gamma_{ij}^s$  and  $\eta_{ij}^s$  represent the  $s$ th largest element of the membership degree and nonmembership degree of  $G_k^{(i)}$  under each attribute [25].

The dual hesitant fuzzy direction line element characteristics of horizontal strokes in  $G_k$  are as follows:

$$F_{H_{G_k}} = \frac{\text{card}(H) + \sum_{i=1}^n \rho_{G_k^{(i)}} \cdot \text{card}(H_{G_k^{(i)}})}{\text{sum\_H}}, \quad (13)$$

where  $F_{H_{G_k}}$  represents the dual hesitant fuzzy direction line element feature of horizontal strokes in the grid of Chinese characters,  $\rho_{G_k^{(i)}}$  is the weighted correlation coefficient between  $G_k^{(i)}$  and the ideal grid,  $\text{card}(H_{G_k^{(i)}})$  is the number of elements in the horizontal stroke set  $H_{G_k^{(i)}}$  in  $G_k^{(i)}$ , and  $\text{sum\_H}$  is the sum of the horizontal stroke pixel points in the entire image.

Similarly, the dual hesitant fuzzy direction line element features of vertical, apostrophe, and downstroke in the grid can be calculated, and the four stroke feature combinations in all the grids can be used as Chinese character feature vectors.

#### 4. Image Retrieval Algorithm for Chinese Characters in Ancient Books

By introducing the dual hesitant fuzzy direction line element feature and using it to calculate the similarity between images of Chinese characters in ancient books, the retrieval results of images of Chinese characters in ancient books are obtained and the output is well ordered in Algorithm 2.

The correlation coefficient was extended to the similarity measure between Chinese images. Equation (14) is used to calculate the correlation coefficient  $\rho_{R-i}$  between the image  $T_I$  to be retrieved and the image  $T_{R-i}$  in the database. Multiple threshold values of correlation coefficients  $\rho_\theta$  are set to control the number of image outputs:

$$\rho_{R-i} = \frac{C(F_I, F_{R-i})}{[C(F_I) \cdot C(F_{R-i})]^{1/2}}, \quad i = 1, 2, \dots, N-1, \quad (14)$$

where  $F_I$  is the feature vector of the Chinese character image to be retrieved,  $F_{R-i}$  is the feature vector of the image in the database,  $N$  is the number of images in the database, and  $C(F_I, F_{R-i})$  is the covariance of  $F_I$  and  $F_{R-i}$ .

### 5. Results and Discussion

**5.1. Experimental Parameter Setting.** The image samples of Chinese characters in ancient books were collected from ‘‘Si ku Quan shu,’’ an important document recognized in the study of Chinese characters in ancient books. The images were marked according to information such as the cabinet, ministry, and book to which they belong. Owing to the absence of a public retrieval dataset, an experimental dataset for image retrieval of Chinese characters in ancient books was established based on the Chinese character samples collected earlier. VS2017 and SQLSEVER2017 were used to realize the image retrieval system of ancient Chinese books. The retrieval experiment was conducted on 92840 ancient Chinese book samples in the dataset.

**Definition 1.** Precision ( $P$ ) is the ratio of the number of similar Chinese images in the output result to the total number of Chinese images in the output result:

$$P = \frac{N_s}{N_A} \times 100\%. \quad (15)$$

**Definition 2.** Recall rate ( $R$ ) is the ratio of the number of images of similar Chinese characters in the retrieval results to the number of images of all similar Chinese characters in the experimental data:

$$R = \frac{N_s}{N_{J-A}} \times 100\%. \quad (16)$$

**Definition 3.**  $F$ -Measure ( $F_\alpha$ ) is the weighted harmonic mean of  $P$  and  $R$ :

$$F_\alpha = \frac{(1 + \alpha^2)P \times R}{\alpha^2 P + R}, \quad (17)$$

where  $\alpha = 1, P$ , and  $R$  have the same weight.  $F_1$  combines the results of  $P$  and  $R$ ; the higher the  $F_1$ , the better the retrieval performance.

#### 5.2. Analysis of Experimental Results

**5.2.1. Retrieval Performance Analysis.** Four groups were selected according to the structure of Chinese characters and 10 samples in each group were retrieved for analysis, as shown in Table 1.

The feature extraction methods in [9, 10] were used for the comparative tests. We use the same dataset and similarity measure to compare these algorithms. The retrieval results of the aforementioned experimental samples were counted and analyzed under the set threshold of the correlation coefficient, and the recall, precision, and  $F_1$  were calculated. The final result was obtained from the average of 40 samples.

According to the experimental results, when  $\rho_\theta$  is 0.7, the average recall rate of three experiments falls below 60%. To ensure the validity of the experiment, the threshold value was selected from [0.7, 0.9] and the interval was set as 0.05.

It can be seen from Tables 2 to 4 that, under the set correlation thresholds, the retrieval method in this study is superior to the comparison tests in terms of recall rate and precision. This indicated the effectiveness of this method in the image retrieval of ancient Chinese characters.

Figure 7 shows the contrast line chart of  $F_1$ , where  $E1$  refers to the algorithm of this study,  $E2$  refers to the algorithm in [9], and  $E3$  refers to the algorithm in [10]. From Figure 7, the performance of the algorithm in this study is higher than other algorithms. The proposed algorithm evaluated the influence of adjacent grids on the current grid and improved the robustness of the features. Therefore, the algorithm in this study is applicable in the retrieval of ancient Chinese character images.

Figure 8 shows the average recall and precision of the four groups of samples in this study.

<p><b>Input:</b> image <math>T_I</math> to be retrieved</p> <p><b>Output:</b> images <math>T_{R_i}</math> of Chinese characters within the threshold of the correlation coefficient</p> <ol style="list-style-type: none"> <li>(1). Open the image <math>T_I</math> to be retrieved</li> <li>(2). Preprocess <math>T_I</math></li> <li>(3). Elastic grid division</li> <li>(4). The dual hesitant fuzzy direction line element characteristics of <math>T_I</math> were extracted</li> <li>(5). <b>while</b> <math>i &lt; N</math> <b>do</b>//traversal database</li> <li>(6).     Calculate the correlation coefficient <math>\rho_{R_i}</math> between <math>T_I</math> and <math>T_{R_i}</math></li> <li>(7).     <b>if</b> <math>\rho_{R_i} &lt; \rho_\theta</math> <b>then</b></li> <li>(8).         <math>T_{R_i}</math> is added to the result data table <math>R(id, \rho)</math></li> <li>(9).     <b>else</b></li> <li>(10).         <math>i = i + 1</math></li> <li>(11).     <b>end if</b></li> <li>(12). <b>end while</b></li> <li>(13). <b>return</b> <math>R(id, \rho)</math></li> </ol>
---

ALGORITHM 2: Image retrieval algorithm of Chinese characters in ancient books based on the dual hesitant fuzzy direction line element.

TABLE 1: Partial sample of retrieval set.

Set number	Chinese characters structure	Samples of Chinese characters
1	Left and right structure	類錄 GJHZ_0000010020133 GJHZ_0000010020117
		讀雜 GJHZ_0000010070072 GJHZ_0000010180061
2	Upper and lower structure	書賞 GJHZ_0000010030093 GJHZ_0000010050136
		熙鑒 GJHZ_0000010170205 GJHZ_0000010150181
3	Surrounding structure	進國 GJHZ_0000010040044 GJHZ_0000010140071
		閣歷 GJHZ_0000010240190 GJHZ_0000010360216
4	Monomeric structure	而之 GJHZ_0000010200180 GJHZ_0000010050314
		臣月 GJHZ_0000010070269 GJHZ_0000010070285

TABLE 2: Average precision, average recall, and  $F_1$  of the algorithm in this study.

$\rho_\theta$ (%)	0.9	0.85	0.8	0.75	0.7
$P_1$	90.4	83.47	75.25	65.85	58.54
$R_1$	58.21	65.98	73.8	81.26	87.2
$F_1^{(1)}$	70.82	73.7	74.52	72.75	70.05

TABLE 3: Average precision, average recall, and  $F_1$  of the algorithm in [9].

$\rho_\theta$ (%)	0.9	0.85	0.8	0.75	0.7
$P_2$	82.9	74.34	68.73	56.56	50.43
$R_2$	51.23	58.36	64.58	70.71	78.58
$F_1^{(2)}$	63.33	65.39	66.59	62.85	61.43

TABLE 4: Average precision, average recall, and  $F_1$  of the algorithm in [10].

$\rho_\theta$ (%)	0.9	0.85	0.8	0.75	0.7
$P_3$	87.96	81.05	72.59	61.83	55.03
$R_3$	56.07	63.41	70.13	78.04	84.5
$F_1^{(3)}$	68.48	70.66	71.34	69	66.65

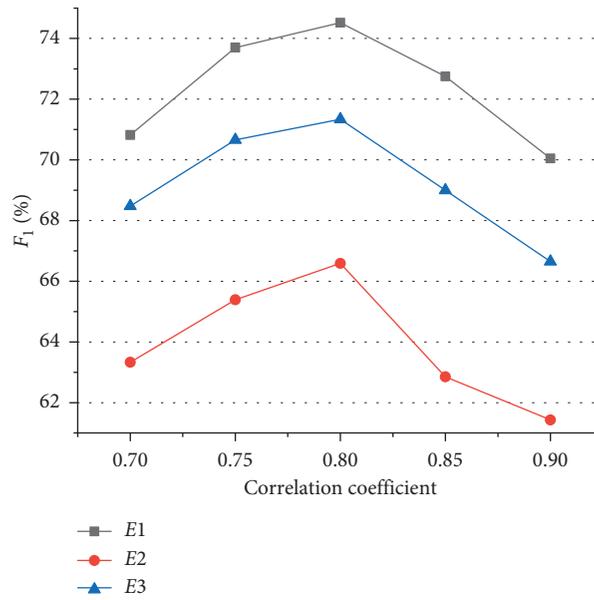


FIGURE 7:  $F_1$  of three experiments.

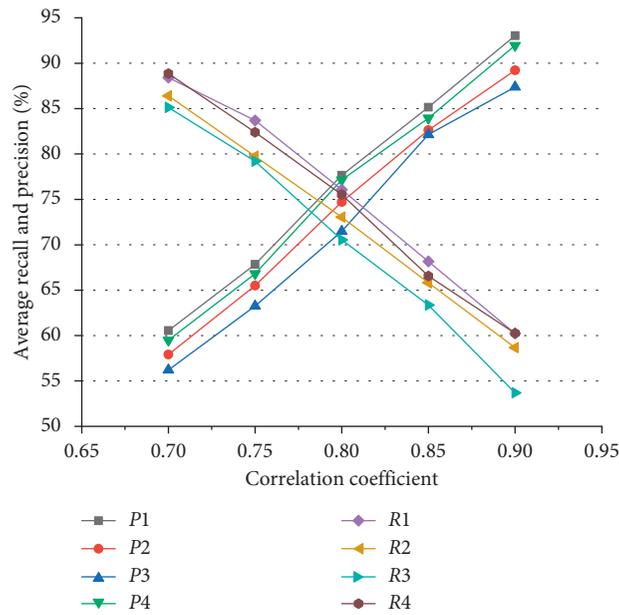


FIGURE 8: Average recall and precision of the 4 groups.

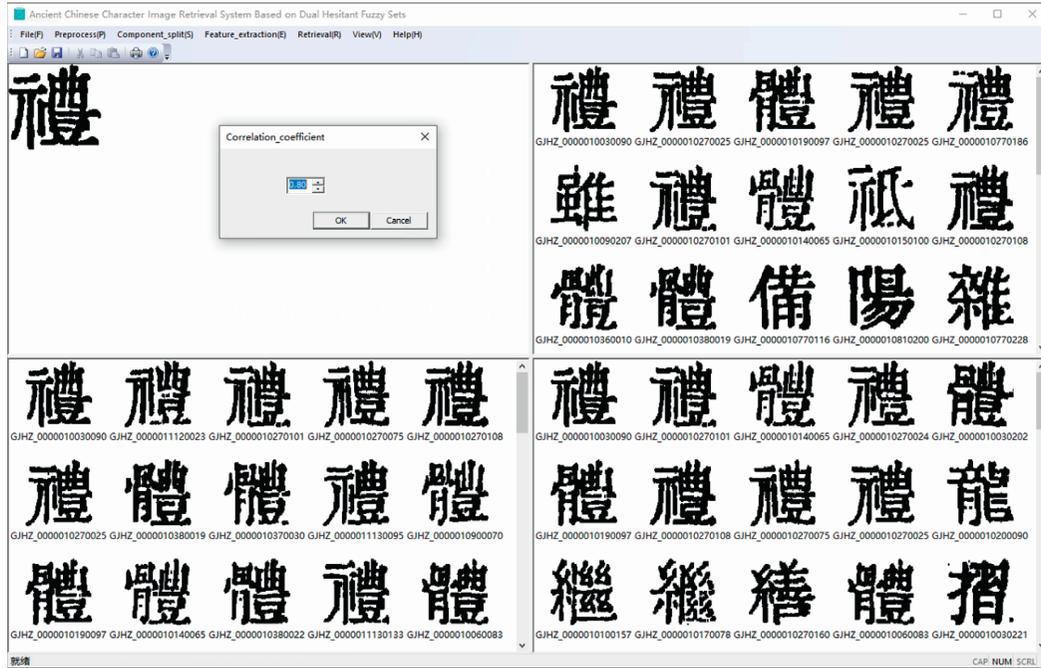


FIGURE 9: Retrieval results of Chinese character image.

$P_j$  ( $j=1,2,3,4$ ) are the average precisions and  $R_i$  ( $i=1,2,3,4$ ) are the average recalls. As can be seen from Figure 8, the retrieval accuracy of this method for Chinese characters with left and right structures and a single structure is higher than that of other structures.

**5.2.2. Retrieval Result Analysis.** A Chinese character image was randomly selected to compare the retrieval results of the three experiments. As shown in Figure 9, the threshold value of the correlation coefficient is set as 0.8. The lower left view shows the first 15 images of the retrieval results of this experiment. The upper right view shows the first 15 images of retrieval results of the method in [9] and the lower left view shows the first 15 images of retrieval results of the method in [10].

As can be seen from Figure 9, the first 15 images in this experiment have a higher similarity than the comparison experiments. The similarity of the output images in the comparison experiments is significantly lower than that in this experiment. This indicates that the retrieval method in this study has a high accuracy.

## 6. Conclusion

This study proposes a method of image retrieval for Chinese characters in ancient books based on a dual hesitant fuzzy set. Dual hesitant fuzzy sets have the advantage of expressing uncertain information more comprehensively. They introduce the information into feature extraction of directional line elements, calculate the comprehensive evaluation value of adjacent grids to the current grid under multiple attributes, extract more complex and robust image features of Chinese characters in ancient books, and improve the retrieval performance. The experimental results

show that the average precision and average recall of this method are 1–4 percentage points higher than those of the comparison methods under multiple correlation coefficient thresholds.

The follow-up work will be mainly conducted with the following two aspects: (1) improving the attribute index according to the topological structure of Chinese characters in ancient books and (2) optimizing the algorithm to improve the retrieval efficiency because the time complexity of the algorithm is relatively high as the membership and nonmembership should be calculated under multiple attributes.

## Data Availability

The data used to support the findings of this study have been deposited in the [https://github.com/ningmengweidexiaotanke/AccientCC\\_DHFS](https://github.com/ningmengweidexiaotanke/AccientCC_DHFS).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was funded by the National Natural Science Foundation of China (Grant no. 61375075), the Natural Science Foundation of Hebei Province of China (Grant no. F2019201329), and the Key Project of the Science and Technology Research Program in University of Hebei Province of China (the Science and Technology Project of Hebei Education Department) (Grant nos. ZD2017208 and ZD2019131).

## References

- [1] J. M. Guo, H. Prasetyo, and J. H. Chen, "Content-based image retrieval using error diffusion block truncation coding features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 466–481, 2015.
- [2] R. Ashraf, M. Ahmed, and S. Jabbar, "Content based image retrieval by using color descriptor and discrete wavelet transform," *Journal of Medical Systems*, vol. 42, no. 3, pp. 1–12, 2018.
- [3] X. X. Zhang, M. Jiang, and Q. Q. Kou, "Color image retrieval algorithm fusing color and principal curvatures information," *IEEE Access*, vol. 8, pp. 184945–184954, 2020.
- [4] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," *Information Fusion*, vol. 5, pp. 76–99, 2019.
- [5] M. W. Jian, Y. L. Yin, and J. Y. Dong, "Content-based image retrieval via a hierarchical-local-feature extraction scheme," *Multimedia Tools and Applications*, vol. 77, no. 21, pp. 29099–29117, 2018.
- [6] S. Unar, X. Y. Wang, and C. P. Wang, "A decisive content based image retrieval approach for feature fusion in visual and textual images," *Knowledge-Based Systems*, vol. 179, pp. 8–20, 2019.
- [7] R. Chen and P. S. Wang, "Chinese handwriting identification method based on keyword extraction," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 31, no. 11, 2017.
- [8] R. Y. Li, H. Wang, and K. F. Ji, "Feature extraction and identification of handwritten characters," in *Proceedings of 8th International Conference on Intelligent Networks and Intelligent Systems*, pp. 193–196, IEEE Press, Washington, DC, USA, November 2015.
- [9] W. Wei and X. D. Guo, "Study of handwritten Chinese characters' double elastic mesh and fuzzy Feature," *Control Engineering of China*, vol. 19, no. 6, pp. 1025–1028, 2012.
- [10] G. Ran, S. Huang, and Z. H. He, "Standardized elastic dual-mesh Chinese character feature extraction based on overlap and fuzzy technology," *Computer Engineering and Design*, vol. 37, no. 1, pp. 211–215, 2016.
- [11] F. Li, Q. Shen, Y. Li, and N. M. Parthaláin, "Handwritten Chinese character recognition using fuzzy image alignment," *Soft Computing*, vol. 20, no. 8, pp. 2939–2949, 2016.
- [12] X. C. Liu and Y. Y. Meng, "Tangut character recognition based on HOG feature extraction and fuzzy support vector machine," *Journal of Northwest Normal University (Natural Science)*, vol. 55, no. 5, pp. 39–43, 2019.
- [13] V. Torra, "Hesitant fuzzy sets," *International Journal of Intelligent Systems*, vol. 25, no. 6, pp. 529–539, 2010.
- [14] B. Zhu, Z. Xu, and M. Xia, "Dual hesitant fuzzy sets," *Journal of Applied Mathematics*, vol. 2012, pp. 1–13, 2012.
- [15] Z. Su, Z. Xu, H. Liu, and S. Liu, "Distance and similarity measures for dual hesitant fuzzy sets and their applications in pattern recognition," *Journal of Intelligent & Fuzzy Systems*, vol. 29, no. 2, pp. 731–745, 2015.
- [16] R. G. Wang, W. M. Li, and T. Zhang, "New distance measures for dual hesitant fuzzy sets and their application to multiple attribute decision making," *Symmetry-Basel*, vol. 12, no. 2, 2020.
- [17] H. M. Zhang, "Distance and entropy measures for dual hesitant fuzzy sets," *Computational and Applied Mathematics*, vol. 39, no. 2, pp. 82–99, 2020.
- [18] G. W. Wei, J. Wang, and C. Wei, "Dual hesitant pythagorean fuzzy Hamy mean operators in multiple attribute decision making," *IEEE Access*, vol. 7, pp. 86697–86716, 2019.
- [19] J. Wang, G. W. Wei, and C. Wei, "Dual hesitant q-rung orthopair fuzzy muirhead mean operators in multiple attribute decision making," *IEEE Access*, vol. 7, pp. 67139–67166, 2019.
- [20] X. Y. Tang and G. W. Wei, "Dual hesitant Pythagorean fuzzy Bonferroni mean operators in multi-attribute decision making," *Archives of Control Sciences*, vol. 29, no. 2, pp. 339–386, 2019.
- [21] X. Y. Tang and G. W. Wei, "Multiple attribute decision-making with dual hesitant pythagorean fuzzy information," *Cognitive Computation*, vol. 11, no. 2, pp. 193–211, 2019.
- [22] L. W. Jin and B. Z. Xu, "Directional cellular feature extraction with elastic meshing for handwritten Chinese character recognition," *Journal of Circuits and Systems*, vol. 2, no. 3, pp. 8–13, 1997.
- [23] F. F. Yan, C. P. Wei, and S. N. Ge, "Entropy measures for dual hesitant fuzzy sets and their application of multi-attribute decision-making," *Journal of Circuits and Systems*, vol. 48, no. 17, pp. 131–139, 2018.
- [24] J. Ye, "Correlation coefficient of dual hesitant fuzzy sets and its application to multiple attribute decision making," *Applied Mathematical Modelling*, vol. 38, no. 2, pp. 659–666, 2014.
- [25] X. Li and X. H. Zhang, "On the correlation coefficients of dual hesitant fuzzy sets and their applications," *Fuzzy Systems and Mathematics*, vol. 32, no. 5, pp. 121–129, 2018.