

Book Retrieval Method Based on QR Code and CBIR Technology

Qiuyan Wang^{1,*} and Haibing Dong²

Abstract: It is the development trend of library information management, which applies the mature and cutting-edge information technology to library information retrieval. In order to realize the rapid retrieval of massive book information, this paper proposes a book retrieval method combining QR code with image retrieval technology. This method analyzes the visual features of book images, design a book image retrieval method based on boundary contour and regional pixel distribution features, and realizes the association retrieval of book information combined with the QR code, so as to improve the efficiency of book retrieval. The experimental results show that, the books can be retrieved effectively through the boundary contour and regional pixel distribution features, the book information can be displayed through QR code, readers can be provided with fast and intelligent massive book retrieval services.

Keywords: Book retrieval, image retrieval, QR code, visual features.

1 Introduction

Under the background of massive data, realize the rapid retrieval of massive book information has become the urgent demand of book informatization management. At present, There are two methods for massive book retrieval: the label classification retrieval and the instrument scanning retrieval [Tang (2019)]. The label classification retrieval classifies the book catalogue of manual, tags the book information incorporate tag, and facilitates readers to understand the book information through the label. The instrument scanning retrieval relates the book information with bar-code, and obtains book information through the special scanning equipment. Although the retrieval accuracy of the two methods is higher, the retrieval time is longer, and the establishment of labels and pasting bar codes requires higher costs, later maintenance is troublesome, more labor is paid.

The development of the computer information technology has brought new ways of life and cultural ideas to people. Book retrieval is also affected by informatization gradually, the traditional retrieval method is broken gradually, it is developing forward to intelligent direction [Jin (2018)]. For example: Multi-category information classification of massive books is realized by computer technology modeling [gao (2017)], The classification probability of book retrieval information is realized by

¹ School of Electrical and Information Engineering, Hunan Institute of Technology, Hengyang, China.

* Corresponding Author: Qiuyan Wang. Email: 2015001948@hnit.edu.cn.

combining the principles of machine learning and probability statistics such as support vector machine [Pu and Qiu (2015)], The retrieval optimization is realized by analyzing the relevance of book retrieval through fuzzy target determination model [Zhang (2017)], and so on. The use of these technologies has improved the retrieval efficiency of mass books, and promoted the progress of book information technology, but it also has the disadvantages of poor expansibility, large classification error and low retrieval accuracy. Therefore, intelligent retrieval of mass books is realized based on computer technology. It is still necessary to introduce modern information retrieval technology and design efficient, accurate and convenient retrieval methods based on the needs of readers.

2 Design of book retrieval method based on QR code and CBIR technology

Content Based Image Retrieval(CBIR) [Munish, Paval and Naresh (2018)] is using image visual content (such as color, texture, shape, semantic features, etc.) set up the index, it is used to retrieve the required object information in the database through image matching technology. The book retrieval method designed in this paper takes the book image as the medium, uses the sample book image, realizes the retrieval in the book database according to the demand. In the process of retrieval, the paper analyzes the two kinds of features of the book image: the boundary contour and the regional pixel distribution, and used the weighted euclidean distance method to measure the similarity, so as to realize the book image retrieval which is similar to the sample. QR code is called quick response matrix code [Yuan, Zhao and Cui (2018)]. It has the characteristics of ultra-high speed reading, which can facilitate the transmission of information. In order to better facilitate readers to enjoy quality book retrieval services and Get the detailed book information, the paper uses QR codes to introduce the detailed information of book information, When the reader retrieves the corresponding book image, the corresponding book information can be retrieved by QR code, so as to better realize the retrieval of book information. The specific retrieval method mainly includes two modules: the book retrieval library creation and the book retrieval query, as shown in Fig. 1.

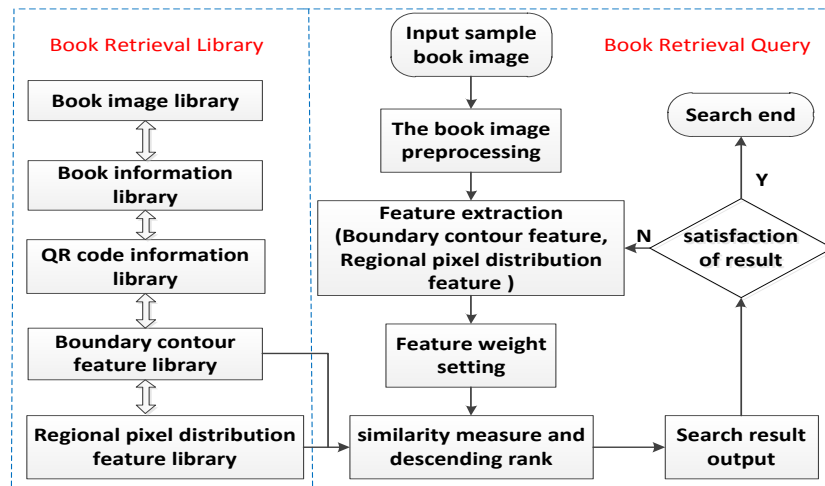


Figure 1: Book retrieval method based on QR code and CBIR technology

Book retrieval library is the basis of book retrieval. The book retrieval library designed in this paper contains five sub-libraries, and there is a one-to-one correspondence between the books in the five sub-libraries according to the ported number. The feature library of boundary contour and regional pixel distribution is built according to the feature analysis method designed in this paper, which is mainly used to measure the similarity with the features of sample book images. Book information library and QR code library is mainly used for book information display, simple book information (such as the name of the book, the author information, publishing information, etc.) by the book information library in the form of text intuitive display, detailed book information (such as content abstract illustration, chapter distribution, layout, for example, etc.) by QR code information library displayed in the form of QR code, readers can get more detailed book information by scanning the QR code, as shown in Fig. 2.

The book retrieval query is the reader realizes the book retrieval process. Readers can search out the book information and similar book information by providing the required book sample image. In order to facilitate more intelligent and convenient book retrieval for readers, the paper also designs the setting of feature weight, multiple searches based on requirements. It not only realizes the retrieval function of sample book information, but also realizes the recommendation function of similar books.



Figure 2: A example of obtaining book information through the QR code

2.1 Book image preprocessing

The book retrieval method designed in this paper is mainly based on the book cover image, the acquisition of book images has direct shooting, network download, website screenshots and other methods usually. The book retrieval method designed in this paper is mainly based on the book cover image, the acquisition of book images has direct shooting, network download, website screenshots and other methods usually. The image obtained by different ways usually has changed size, and inevitably produces noise in the process of collection and transmission. Therefore, in order to improve the retrieval efficiency, the paper uses image scaling, image denoising, format conversion and other methods to complete the book image normalization [Kyung and Eui (2017)], normalize the book image as a 380*540 JPG image, and establish the standard book image library. The effect of book image preprocessing is shown in Fig. 3.



(a) Original image

(b) Preprocessed image

Figure 3: A example of book image preprocessing effect

2.2 Book image feature extraction

Book is one of the important tools for the dissemination of cultural knowledge, the type of book is wide, the book image is variety, but most books are dominated by rectangles, the cover designs of books in the same series are often similar in layout. Therefore, this paper analyzes the boundary contour feature and regional pixel distribution feature of the book cover image.

(1) Boundary contour feature

Book cover image as a kind of artificial design image, its boundary contour information is relatively clear usually. This paper adopts the canny operator [Jiang, Zhou, Shen et al. (2015)] to obtain the boundary contour information of book image, Realizes the interpolation of contour column points based on B-spline interpolation [Yue, Liu and Duan (2018)], represents the boundary contour feature of the book image with Fourier descriptor [Emir and Samim (2016)], the specific method is shown in Algorithm 1.

Algorithm 1 Obtaining boundary contour feature of book image

Input: the book image $f(x, y)$

Output: the boundary contour feature T_1

1. The book image binarization treatment;
 2. The canny operator is used to obtain the boundary contour curve $g(x, y)$ of book image;
 3. According to the nth B-spline recursive function expressed in formula (1), b-spline interpolation is performed on the acquired image contour column points, so that the number of column points is 2^n , and the interpolated contour information $U(k)$ is obtained;
 4. According to formula (3), the Fourier descriptor $FFT[U(k)]$ of the graph boundary contour curve $g(x, y)$ is obtained;
 5. Take the real part vector of $FFT[U(k)]$ as the boundary contour feature T_1 ;
 6. Get the boundary contour feature $T_1 = T_1 / \text{Max}(T_1)$.
-

$$b^n(x) = \frac{\left[\frac{n+1}{2} + x \right] b^{n-1} \left[\frac{1}{2} + x \right] + \left[\frac{n+1}{2} - x \right] b^{n-1} \left[\frac{1}{2} + x \right]}{n} \quad (1)$$

$$U(k) = x_k + jy_k, k = 0, 1, 2, \dots, 2^n \quad (2)$$

$$F(\mu) = FFT[U(k)] = \sum_{k=0}^{2^n-1} U(k) e^{-j2\pi\mu k / 2^n} \quad (3)$$

In the formulas, $U(k)$ is the complex variable of the boundary contour column point coordinate value (x is the real part, y is the imaginary part). $n=10$ in this paper, so the sampling points of the boundary contour are 1024, the fast Fourier transform of $U(k)$ can obtain its frequency information parameter $F(\mu)$.

According to Algorithm 1, the contour information of the book image shown, as shown in Fig. 4. The feature vector of boundary contour is $T1=[1.000000 \ 0.688331 \ 0.350670 \ 0.237626 \ 0.180517 \ 0.145823 \ 0.122227 \ 0.104945 \ 0.091613 \ 0.081042 \ 0.072518 \ 0.065563 \ 0.059872 \ 0.055193 \ 0.051295]$ (according to the analysis of experimental results, the first 15 values are taken as effective eigenvalues).

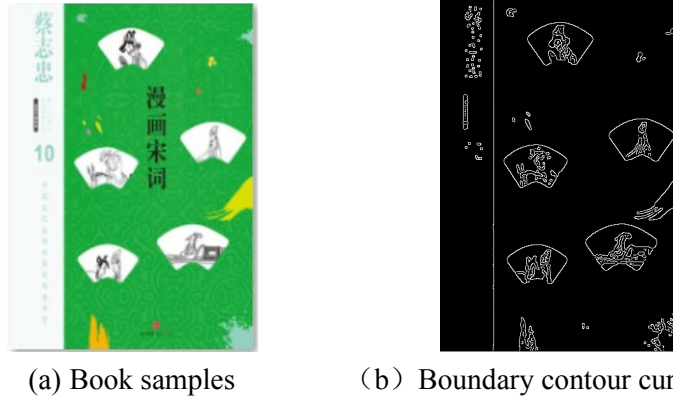


Figure 4: A example of book image boundary contour effect

(2) Regional pixel distribution feature

In general, the book image contains not only the backgrounds element, but also the character element that reflects the book information. For character elements, the distribution characteristics of regional pixels [Yang, Liu and Zhou (2015)] are obvious, therefore, the paper designs that the regional pixel distribution features of image are represented by the statistics of the sum of pixel points in the image region, the specific method is shown in Algorithm 2.

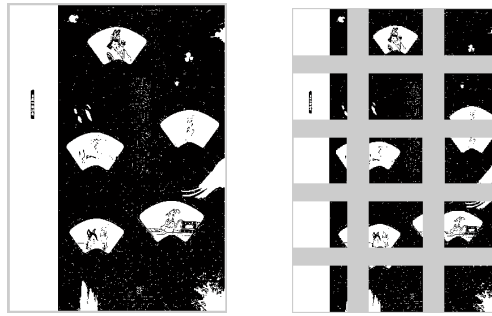
Algorithm 2 Regional pixel distribution feature of book image

Input: the book image $f(x, y)$

Output: the regional pixel distribution feature T_2

1. The book image binarization treatment;
 2. The book image is partitioned into 5*3 sub-images;
 3. Calculate the sum of the target pixels of each sub-image $Sum(i)$;
 4. Get the regional pixel distribution feature $T_2 = Sum. / Max(Sum)$.
-

Image partitioning is performed according to Algorithm 2, as showing in Fig. 5. According to the sub-image of the divided region, the pixel distribution feature vector of the book image region is obtained $T_2 = [0.321420 \ 0.700718 \ 0.989498 \ 0.342048 \ 1.000000 \ 0.840725 \ 0.268757 \ 0.739101 \ 0.736230 \ 0.289989 \ 0.733434 \ 0.673819 \ 0.326709 \ 0.919154 \ 0.878806]$ (The sub-images are sorted from left to right and from top to bottom).



(a) The binary image (b) Subimage division

Figure 5: A example of book image partition effect

2.3 Similarity measure

The book retrieval method designed in this paper is mainly based on the boundary contour of the book image and the distribution of regional pixels. In order to facilitate users to search for books more conveniently according to their need, the paper adopts weighted Euclidean distance [Soumva Maitreyee and Patrick (2018)] to measure the similarity, the specific method is shown in formula (4). The emphasis of retrieval direction can be adjusted by the feature weights of two kinds of features, the similarity can be calculated according to the size of D. The first N books similar to the sample books can be retrieved by calculating the similarity between the sample books and the library books and sorting them in descending order (N=10 in this paper).

$$D = \omega_1 \sqrt{\sum_{i=1}^{15} [T_1^A(i) - T_1^B(i)]^2} + \omega_2 \sqrt{\sum_{i=1}^{15} [T_2^A(i) - T_2^B(i)]^2} \quad (4)$$

In the formulas, T_1^A and T_1^B represent the boundary contour features of two different images respectively; T_2^A and T_2^B represent the regional pixel distribution features of two different images respectively; ω_1 and ω_2 respectively represent the feature weights of the boundary contour feature and the regional pixel distribution feature ($\omega_1 + \omega_2 = 1$).

3 Experimental results and analysis

In the book retrieval, it is the key to search the required books quickly and accurately. To test the performance of the retrieval method designed in this paper, this paper uses 500 different books to set up the book retrieval library, and 10 sample books for performance testing, each sample book is associated with 10 similar books in the library, as shown in Fig. 7.



Figure 6: Example of test the sample books



Figure 7: Example of similar books

This paper uses MATLAB R2014A to write programs and design GUI interface [Vitucci, Minniti and Tremsim (2018)] to display the results of book retrieval, each retrieval output four pages of 10 book information, as showed in Fig. 8. Q is used to represent recall rate [Lv and Duan (2018)] of similar book retrieval; P is used to indicate whether the sample books can be retrieved accurately, P=1 is used to indicate that they can be accurately retrieved, and P=0 is used to indicate that they cannot be accurately retrieved. The retrieval performance of the method under different feature weight is analyzed by setting the feature weight coefficients, as showed in Fig. 7 and Tab. 1.

As showed in Fig. 8, it takes about 2 s to retrieve the sample books in a library with a stock of 500. Therefore, the book retrieval method designed in this paper can quickly realize the book retrieval, and can display the book information in two ways to respond to the different retrieval needs of readers. As showed in Tab. 1, under the different feature weight, the book retrieval method designed by this paper can accurately retrieve the same books as the sample books basically. As shown in Fig. 9, when the value of feature weight is 0.7 and 0.3 respectively, the book retrieval performance designed by combining the boundary contour feature and the regional pixel distribution feature is the best, and the recall rate is higher than that realized by using the boundary contour feature or the regional pixel distribution feature alone. In conclusion, the book retrieval method based on QR code and CBIR technology designed in this paper can not only realize the quick and accurate book retrieval, but also retrieve similar books of the same series recommended and samples.

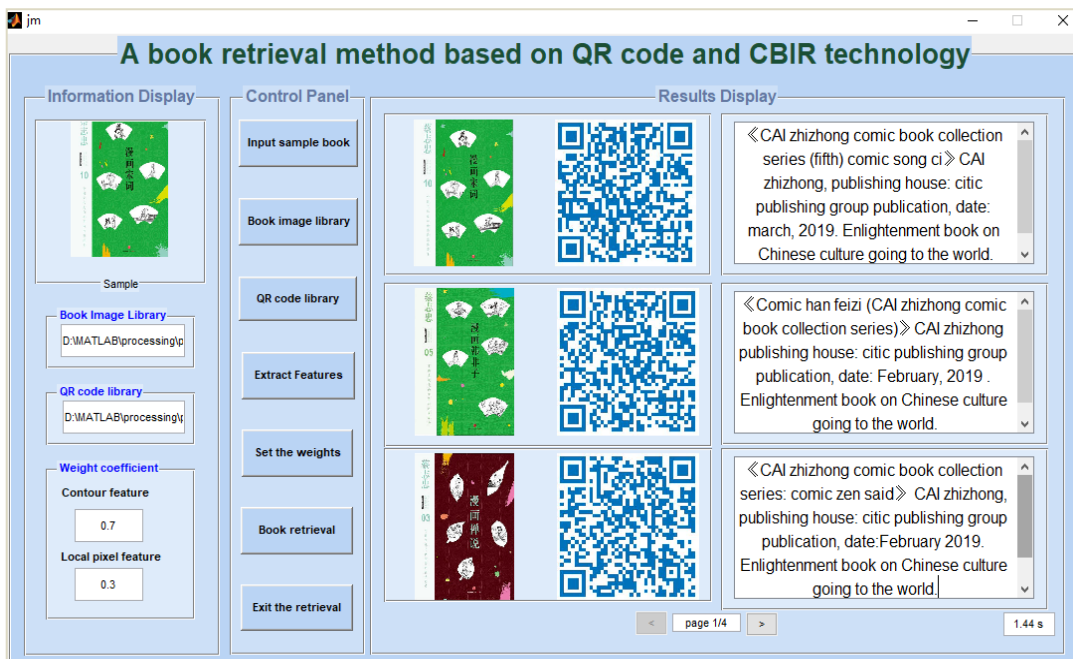


Figure 8: Example of book search results

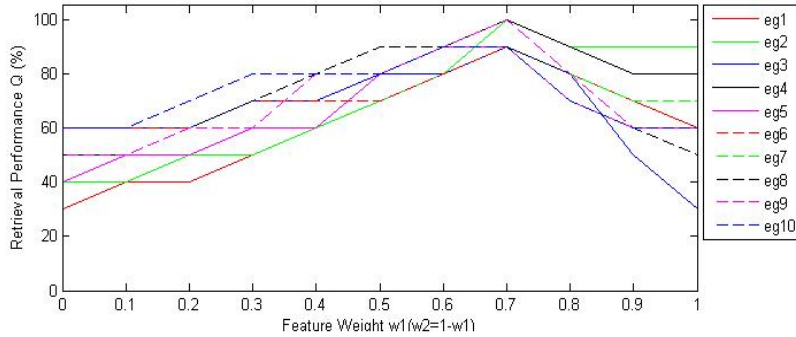


Figure 9: The book retrieval performance Q under the different weight coefficients

Table 1: The book retrieval performance P under the different weight coefficients

W1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
eg1	1	1	1	1	1	1	1	1	1	1	1
eg2	1	1	1	1	1	1	1	1	1	1	1
eg3	0	1	1	1	1	1	1	1	1	1	0
eg4	1	1	1	1	1	1	1	1	1	1	1
eg5	0	1	1	1	1	1	1	1	1	1	0
eg6	1	1	1	1	1	1	1	1	1	1	1
eg7	1	1	1	1	1	1	1	1	1	1	1
eg8	1	1	1	1	1	1	1	1	1	0	0
eg9	1	1	1	1	1	1	1	1	1	1	1
eg10	1	1	1	1	1	1	1	1	1	1	0

4 Conclusion

This paper proposes a book retrieval method combining QR code and CBIR technology, designs a book image retrieval method based on boundary contour feature and regional pixel distribution feature, and combines QR code to realize the introduction of book information. It is convenient for readers to understand book information more quickly and comprehensively, and provides intelligent and high-quality book retrieval service for readers. This image-based book retrieval method meets the needs of readers in the contemporary era of electronic reading.

Acknowledgement: This work is supported by the national innovation and entrepreneurship training program for college students (No. 201811528081); by university-level scientific research project of hunan institute of technology (No. 2018HY012); by scientific research project of hunan department of education (No. 18C0898).

References

- Emir, S.; Samim, K.** (2016): Phase preserving Fourier descriptor for shape-based image retrieval. *Signal Processing: Image Communication*, vol. 40, no. 8, pp. 53-61.
- Gao, Y.** (2017): Research on optimization design of rapid query system for massive book retrieval information. *Modern Electronic Technology*, vol. 40, no. 6, pp. 5-9.
- Jiang, W.; Zhou, H. Y.; Shen, Y.; Bo, L.; Fu, Z. J.** (2015): Image segmentation with pulse-coupled neural network and Canny operators. *Computers and Electrical Engineering*, vol. 46, no. 3, pp. 528-538.
- Jin, Z.** (2018): Research on chinese electronic book retrieval in university libraries. *2nd International Conference on Education Science and Economic Management*.
- Kyung, M.; Eui, Y.** (2017): Image recognition performance enhancements using image normalization. *Human-centric Computing and Information Sciences*, vol. 7, no. 1, pp. 86-92.
- Lv, X.; Duan, F.** (2018): Metric learning via feature weighting for scalable image retrieval. *Pattern Recognition Letters*, vol. 109, no. 1, pp. 131-137.
- Munish, K.; Payal, C.; Naresh, K.** (2018): An efficient content based image retrieval system using BayesNet and K-NN. *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 156-162.
- Pu, G.; Qiu, Y.** (2015): Image classification method based on sparse representation global dictionary learning. *Computer Applications*, vol. 35, no. 2, pp. 499-501, 514.
- Soumya, P.; Maitreyee, D.; Patrick, S.** (2018): Boosting content based image retrieval performance through integration of parametric & nonparametric approaches. *Journal of Visual Communication and Image Representation*, vol. 9, no. 3, pp. 65-73.
- Tang, H.** (2019): Research on management of university books and materials based on information retrieval technology. *Think Tank Era*, vol. 3, no. 10, pp. 125, 135.
- Vitucci, G.; Minniti, T.; Tremsin, A. S.; Kockelmann, W.; Gorini, G.** (2018): Investigation of image distortion due to MCP electronic readout misalignment and correction via customized GUI application. *Journal of Instrumentation*, vol. 13, no. 4, pp. 1748-1754.
- Yang, X.; Liu, T. S.; Zhou, D. K.** (2015): A multi-frame adaptive super-resolution method using double channel and regional pixel information. *Optik-International Journal for Light and Electron Optics*, vol. 126, no. 24, pp. 4789-5998.
- Yuan, B.; Zhao, X.; Cui, D.** (2018): The effective QR code technology development. *4th International Conference on Machinery, Materials and Compute*.
- Yue, Z.; Liu, C.; Duan, J.** (2018): Image interpolation via discontinuous B-spline on android platform. *International Conference on Digital Image Processing*.
- Zhang, Y.** (2017): Fuzzy target determination technology in mass book retrieval. *Modern Electronic Technology*, vol. 40, no. 13, pp. 137-140.