Research on Efficient Retrieval Technology of Optical Characteristics of Image Resources in Digital Library

Guotao Zhao¹, Jie Ding²,*, Honglai Li¹, Fang Xu¹
¹Hubei Engineering University, Xiaogan, China
²College of Technology, Hubei Engineering University, Xiaogan, China

*Corresponding author’s e-mail: dingjie@hbeu.edu.cn

Abstract. In order to improve the intelligence of digital library image resources, this paper proposes an efficient retrieval technology of digital library image resources based on deep learning and semantic feature extraction, which uses infrared and radio frequency transverse scanning technology to scan the front cover and back cover of digital library image resources, and divides the text and image regions of scanned digital library image resources infrared and radio frequency images. In the infrared and radio frequency transmission space, the collected scanned image resources of digital library are subjected to semantic diversification feature decomposition and feature separation, and the binary semantic ontology information features of digital library image resources are extracted. Combined with the image semantic feature extraction method, the convergence control in the process of extracting the structural features of digital library image resources is realized, and the digital library image resources are retrieved based on the extracted library resource information feature components. Simulation results show that this method has higher precision and shorter retrieval time, which improves the retrieval efficiency of digital library.

Keywords. digital library; Image resources; Semantic features; Search; Deep learning

1. Introduction
With the development of library information management technology, the degree of modernization of digital library is getting higher and higher. Digital library is an information processing technology that realizes image resource management, information retrieval and resource sharing of digital library with digital processing and storage technology, so as to facilitate cross-regional and object-oriented network query and dissemination [1]. The core of digital library's image resources utilization lies in the storage, retrieval, transmission, processing and utilization of books, so as to realize seamless linking and intelligent retrieval across libraries. Studying the efficient retrieval method of image resources in digital library is of great significance to improve the information management degree of library and realize the virtual construction and knowledge sharing of digital library [2].

At present, the retrieval methods of digital library image resources mainly include semantic retrieval method, manual annotation method, text detection method and image feature extraction method, etc. By extracting the semantic information feature quantity of digital library image resources,
combined with database construction and big data analysis method, the retrieval ability of digital library image resources is achieved, and the retrieval methods of related literatures are studied [3]. Literature [4] puts forward a digital book retrieval method based on semantic directional feature clustering, which adaptively marks the image frame sequence of the book cover with overlapping texts, and uses the edge contour feature extraction method to locate the image resources of the digital library, thus improving the retrieval accuracy. However, with the expansion of the image resources of the digital library, the retrieval efficiency decreases and the calculation overhead is too high. Literature [5] puts forward a multimedia book retrieval method based on Gaussian marginal fusion. The constructed book retrieval system is essentially a distributed information system made of multimedia, which can realize co-construction and sharing based on network environment and improve the accuracy of retrieval [6]; Literature proposes a multimedia digital book retrieval method based on semantic crawler recognition, which uses graph model recognition and edge pixel segmentation to identify and locate image resources of digital library, and improves the ability of accurate location and recognition of image resources of digital library. However, this retrieval method still has some problems, such as excessive computational overhead and poor retrieval timeliness [7-9].

With the development of infrared and radio frequency image processing technology, it has a certain application prospect to use the infrared and radio frequency image scanning and processing method of book title page for book retrieval. Aiming at the problems existing in traditional methods, this paper proposes an efficient retrieval technology of digital library image resources based on deep learning and semantic feature extraction. Firstly, infrared and radio frequency horizontal scanning technology is used to scan the front cover and back cover of digital library image resources. Text and image region segmentation is carried out for scanned digital library image resources infrared and radio frequency images. Semantic diversity feature decomposition and feature separation are carried out for collected digital library image resources infrared and radio frequency transmission space, and then binary semantic ontology information features of digital library image resources are extracted. Combined with image semantic feature extraction method, digital library image resources information fusion processing is carried out, and the word structure features of digital library image resources are extracted. Finally, the simulation experiment is carried out, and the validity conclusion is obtained.

2. Collection and pretreatment of image resources in digital library

2.1. Digital library image resources collection

In order to realize the retrieval of digital library image resources, the word structure feature extraction method is used to collect images [8]. Firstly, 3D infrared and radio frequency point scanning and transverse scanning methods are used to render and collect the optical features of digital library image resources, and the grid area is divided to establish an image analysis model of 3D infrared and radio frequency scanning optical features of digital library image resources:

$$I(x) = J(x)t(x) + A(1 - t(x))$$

Wherein, $A$ is the pixel value of semantic features of digital library image resources in the direction, $t(x)$ is the statistical distribution of semantic features of digital library image resources in different semantic projection directions, and $J(x)t(x)$ is the correlation coefficient of semantic features of digital library image resources in two-dimensional coordinate plane.

The front and back cover of digital library image resources are scanned by infrared and radio frequency transverse scanning technology [10]. Assuming that the associated pixel values of semantic features of digital library image resources scanned in each frame are the same, as $w \ast N$, information fusion is carried out according to the texture information features of digital library image resources, and information reconstruction of the front and back cover images of digital library image resources is carried out in three-dimensional infrared and radio frequency scanning space [11], and the correlation constraint function of images is obtained as follows:
\[ \begin{align*}
\text{minimize} & \quad f(\tilde{x}) = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \\
\text{subject to} & \quad g_j(\tilde{x}) \leq 0, j = 1, \ldots, l \\
& \quad h_j(\tilde{x}) = 0, j = l+1, \ldots, p
\end{align*} \] (2)

Wherin, \( x \in \Omega \) represents the feasible region of pixel traversal of semantic features of digital library image resources in affine invariant moment space domain. Assuming that the local binar fitting coefficient of digital library image resources after optical scanning is \( g = \{g(i), i \in \Omega\} \), the trajectory tracking of characteristic points of digital library image resources is carried out in a single pixel value distribution area \( I(i,j) \), and the trajectory tracking function is defined as:

\[ L = J(w,e) - \sum_{i=1}^{N} a_i \{w^T \varphi(x_i) + b + e_i - y_i\} \] (3)

By the above processing, the semantic feature attribute set of digital library image resources recorded by gray scale model is \( X = \{x_1, x_2, \ldots, x_n\} \), and the useful pixel parameter test set \( E_{i,j} = \{e_1, e_2, \ldots, e_n\} \) in digital library image resources is set, in which \( e_i \in [0,1] \) and \( \{E_i, E_j, d, t\} \) are expressed as the main feature quantities of the edge data set of digital library image resources, and the quad is used to express the main feature decision tree, \( E_i, E_j \) is the entity set of bifurcation nodes of digital library image resources distribution, which realizes the collection of digital library image resources and pixel separation, and carries out image feature reconstruction and information retrieval according to the extracted image pixels [12].

2.2. Regional segmentation of image resources in digital library

On the basis of the above-mentioned image acquisition, the scanned digital library image resources are divided into text and image regions, and the semantic diversification feature decomposition of digital library image resources is carried out by using the adaptive texture rendering method [13]. The decomposed gray pixel features are mapped in point space, and the optical highlight features of images are extracted. The edge segmentation scale of digital library image resources is as follows:

\[ H = \sum_{r=1}^{t} \sum_{p=1}^{k} (x_{r'} - x_{ijp'})(x_{r'} - x_{ijp'})^T A_{ijp} \] (4)

Gabor wavelet transform is used to obtain the bottom data distribution of semantic features of image resources in digital library, and the pixel level of scanning image resources profile in digital library can be simplified in the following form:

\[ J(W_i) = \sum_{r=1}^{t} \sum_{q=1}^{k} W_{ijr}^T x_{ijr} - W_{ijr}^T x_{ijr} B_{ijr} \]
\[ = tr(W_{ijr}^T \sum_{r=1}^{t} \sum_{q=1}^{k} (x_{r'} - x_{ijr'})(x_{r'} - x_{ijr'})^T B_{ijr} W_{ijr}) \]
\[ = tr(W_{ijr}^T H_2 W_{ijr}) \] (5)

Wherin

\[ H_2 = \sum_{r=1}^{t} \sum_{q=1}^{k} (x_{r'} - x_{ijr})(x_{r'} - x_{ijr'}) B_{ijr} \] (6)

According to the text information and image information on the cover of digital library image resources, the image high-order moment scale information in the semantic features of different digital library image resources is expressed. The location information of digital library image resources is extracted by \( 3 \times 3 \) topology, and the geometric structure of location distribution is as follows:

\[ W_i = (H_1 - H_2)\omega + \lambda \omega \] (7)

3. Realization of semantic feature retrieval of image resources in digital library

3
3.1. Extraction of word structure features of library images

On the basis of scanning the front cover and back cover of digital library image resources with infrared and radio frequency transverse scanning technology, and segmenting the text and image regions of scanned digital library image resources with infrared and radio frequency images, this paper proposes an efficient retrieval technology of digital library image resources based on deep learning and semantic feature extraction. Semantically diversified feature decomposition and feature separation are carried out on collected digital library image resources in infrared and radio frequency transmission space, edge contour features of digital library image resources are extracted in each block area, the best matching block $\Psi$ of semantic feature area of digital library image resources is searched, and then the nearest neighbor domain, $i_1, i_2, \ldots, i_K$, optical flow vector field $D_i = \{x_i', x_i'', \ldots, x_i''\}$ of digital library image resources infrared and radio frequency infrared scanning are selected. The distance distribution from the semantic feature area of digital library image resources to the corresponding point of edge contour is $s_1, s_2, \ldots, s_K$, and the information fusion component of pixel feature points of digital library image resources in the retrieval information distribution space obtained by semantic feature analysis method is expressed as:

$$G_{new} = (1 + \mu T)(1 + \lambda T)G_{old}$$  \hspace{1cm} (8)

$$T(g_i) = \frac{1}{n} \sum_{i=1}^{n} \sum_{o_j} (g_i - g_j)$$  \hspace{1cm} (9)

Wherein, $G_{new}$ and $G_{old}$ are the color feature components of the semantic features of digital library image resources, respectively, and affine invariant mapping is described as $G \rightarrow \Phi$, which represents the text feature components of the semantic features of digital library image resources, and is expressed as follows:

$$F = \tilde{p}(x, y) = p(x, y) \left( \frac{v(x)}{v(y)} \right)^{1/2}$$  \hspace{1cm} (10)

Wherein

$$p(x, y) = \frac{k(x, y)}{v(x)}, \quad v(x) = \sum_{y} k(x, y)$$  \hspace{1cm} (11)

Color features and text features of scanned image resources of digital library are fused, $C([a, b], R)$ respectively represents the interframe invariant feature quantity and pixel intensity of image resources of digital library to be retrieved, pixel sequence reconstruction is carried out in the phase space after information fusion, and the vertical distribution vector set in image resource cabinet of digital library is described as $\|\phi\| = \sup \{\phi(\theta)\}$, $C([a, b], R)$ is the horizontal set of pixels and the horizontal distribution vector set of image resources of digital library. With edge contour feature value, $F$ and $C([a, b], R)$ are the text information matching matrix of image frame sequence and digital library image resource catalogue, which is written as:

$$f_{i}(z) = \begin{bmatrix} f_{i}(z) \\ f_{y}(z) \\ f_{x}(z) \\ h_{x} * f(z) \\ h_{y} * f(z) \end{bmatrix}$$  \hspace{1cm} (12)

Digital library image resources retrieve the autocorrelation information component of the object, so as to extract the word structure features of digital library image resources, extract the binary semantic ontology information features of digital library image resources, and combine the semantic feature extraction method of images to fuse the information of digital library image resources [14].

3.2 Semantic feature retrieval of image resources in digital library

The output feature sequence of the fused digital library image resources retrieval information is $r(n) = r(n \Delta t)$, $n = 0, 1, 2, \ldots, N - 1$, and the copy correlation sequence of the useful features in the digital library image resources is as follows:
Let the pixel set of the input digital library image resources be \( x(t), t = 0,1, \cdots, n - 1 \), and the affine moment invariants of the output image are as follows:

\[
J_j(nT_b) = A \cos(n \times 2\pi nT_b) - B \sin(n \times 2\pi nT_b) = C \cos(n \times 2\pi nT_b - \theta)
\]

Wherein

\[
A = \frac{2LT}{SN} \sin c(n \Delta t) \sum_{i=0}^{n} c \cos \left(2\pi nT_b \left( i + \frac{1}{2} \right) + \phi \right)
\]

\[
B = \frac{2LT}{SN} \sin c(n \Delta t) \sum_{i=0}^{n} c \sin \left(2\pi nT_b \left( i + \frac{1}{2} \right) + \phi \right)
\]

\[
C = \sqrt{A^2 + B^2}, \quad \theta = \arctan \left( \frac{B}{A} \right)
\]

Combining with the semantic feature extraction method of images, the digital library image resources information fusion processing is carried out, and the triangular function expansion is carried out on the digital library image resources, and the regional pixel set of the retrieval output is expressed as:

\[
R_k(k) = R_k(k) \exp(-j\alpha T_r / 2), \quad k = 0,1, \cdots, (N-3)/2
\]

\[
R_k(k) = A_k \exp(j\beta), \quad k = 0,1, \cdots, (N-3)/2
\]

The location distribution of digital library image resources is analyzed by the fusion method of semantic information and optical information, and the binary subsequences of retrieval output of digital library image resources are expressed as follows:

\[
r_i(n) = r_i(n) \exp(-j\alpha T_r / 2), \quad n = 0,1, \cdots, (N-3)/2
\]

\[
r_i(n) = A \exp(j\beta) \exp(j\alpha n T + \theta), \quad n = 0,1, \cdots, (N-3)/2
\]

To sum up, the semantic features of image resources in digital library are extracted, and the extracted feature components of library resources information are taken as the retrieval basis to realize the optimization of image resources retrieval in digital library [15].

4. Experimental analysis

In order to test the application performance of this method in realizing the efficient retrieval of digital library image resources, a simulation experiment was carried out. In the experiment, C++ and Matlab 7 mixed programming were used to design the digital library image resources processing and retrieval algorithm. The digital library image resources came from the stacks of the National Library. In the parameter setting, the maximum scale of global search \( D = 12 \), the semantic feature distribution dimension \( G_{max} = 30 \), Search iteration step \( g \), maximum iteration times \( c = 3 \), maximum gradient difference \( m = 2 \), area search scale coefficient, area pixel set dimension \( f_i = 2.1 \), vector quantization coding rate of semantic features as \( f_i = 0.23 \) characters, initial sampling frequency \( f_i = 2.1 \) Hz and termination frequency \( f_i = 0.23 \) Hz of digital library image resources, pixel size of digital library scanning image resources acquisition as \( 512 \times 512 \times 8 \) bit, compression ratio of image compression coding as \( ICR = 64 \times 8/10 = 51.2 \), design accuracy and recall of digital library image resources retrieval as test indexes, which are defined as follows:

\[
\text{SNR} = -10 \log_{10} \frac{\sigma^2}{D}
\]

\[
\text{PSNR} = -10 \log_{10} \frac{255^2}{D}
\]

The \( \sigma^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})^2 \) represents the number of books that retrieve the image resources of the
digital library, the \( D = \frac{1}{N} \sum_{i=0}^{N-1} (x_0 - \hat{x}_i)^2 \) is the catalogue of the image resources of the digital library in the library, and the \( \bar{x} = \frac{1}{N} \sum_{i=0}^{N-1} x_i \), \( \hat{x}_i \) is the average scanning rate of the semantic feature image. The normalized error of digital library image resource retrieval is defined as:

\[
E = (1 - \frac{N_{\text{correct}}}{N}) \times 100\%
\]

(23)

Where, \( N_{\text{correct}} \) represents the probability of accurately retrieving the target book.

According to the above simulation environment and parameter setting, the semantic feature retrieval of digital library image resources is carried out, and different versions of Digital signal processing textbook books are taken as the retrieval object. First, the digital library image resources are scanned by digital library image resources. The scanned original image is shown in Figure 1.
Taking the digital library image resources collected in figure 1 as the research object, the image segmentation and text scanning are carried out to realize the character structure feature extraction of the digital library image resources, and to extract the binary semantic ontology information features of the digital library image resources as figure 2.
By analyzing the results in Figure 2, we know that the word structure feature extraction of digital library image resources by using this method can accurately mine the text information and cover image information of digital library image resources directory, and use this as the retrieval basis to carry out book retrieval, and the retrieval results are output as shown in Figure 3.
By analyzing the results in Figure 3, it can be known that using this method to retrieve digital library image resources can accurately region different versions of book catalogues with the same title, because using image recognition method to retrieve, this method can effectively remove the interference of text features of related bibliographies, improve the target and object-oriented of retrieval, and the retrieval results of retrieval output are accurate and reliable. In order to compare the performance of different retrieval methods in improving retrieval accuracy, this method is compared with traditional semantic retrieval methods and text retrieval methods, and the comparison result of retrieval precision is shown in Figure 4. After statistical analysis of retrieval performance, the results are shown in Table I. The time cost of book retrieval by this method is 3.23s, which is obviously lower than that of traditional methods (12.34s and 9.76s). The average precision rate and recall rate are higher, and the required prior sample size is 50% and 12% of that of traditional methods. Therefore, the retrieval performance of this method is obviously better than that of traditional methods, and the retrieval efficiency is improved.
Figure 4. Comparison of precision rate of image resources in digital library

Table 1. Comparison of retrieval parameters of digital library image resources

|                  | This method | Semantic retrieval | Text retrieval |
|------------------|-------------|-------------------|---------------|
| Time cost/s      | 3.45        | 12.24             | 9.56          |
| Average precision rate /% | 99.34        | 84.56             | 85.53         |
| Average recall rate /%  | 98.56        | 85.43             | 83.56         |
| Prior sample size  | 101         | 214               | 1200          |

5. Conclusions
In this paper, a new retrieval method of digital library image resources is studied, and an efficient retrieval technology of digital library image resources based on deep learning and semantic feature extraction is proposed. The collected digital library image resources are divided into grid regions, and an image analysis model of semantic features of digital library image resources scanned by 3D infrared and radio frequency is established. Adaptive texture rendering method is used to decompose the semantic diversity features of digital library image resources. The decomposed gray pixel features are mapped in point space, the optical highlights of images are extracted, and the digital library image resource information fusion process is carried out, so as to realize the structural feature extraction and retrieval optimization of digital library image resource words. The research shows that the retrieval method in this paper has higher precision, less time cost and superior overall performance.

Acknowledgment
This work was supported by the Innovation Fund for Production, Study and Research of Chinese Universities of the Ministry of Education, Innovation Project for New Generation's Information Technology (No.2019ITA03027), Science and Technology Research Project of the Education Department of Hubei Province (No.B2019149). Hubei Provincial Department of Education Outstanding Youth Scientific Innovation Team Support Foundation (T2020017).

References
[1] V.S. BHADOURIA, D. GHOSHAL, and A.H. SIDDIQI, A new approach for high density saturated impulse noise removal using decision-based coupled window median filter. Signal, Image and Video Processing, vol. 8, no. 1. pp. 71-84. 2014.
[2] S.K. MEHER, Recursive and noise-exclusive fuzzy switching median filter for impulse noise reduction. Engineering Applications of Artificial Intelligence, vol. 30, pp. 145-154. 2014.
[3] W. DI, MM. CRAWFORD, View generation for multiview maximum disagreement based active learning for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 5. pp. 1942-1954. 2012.

[4] J. PAN, R. LIU, Z. SU, et al, Kernel estimation from salient structure for robust motion deblurring. Signal Processing: Image Communciation, vol. 28, no. 9. pp. 1156-1170. 2013.

[5] X. NIE, W. LIU and W. WU, Ship detection based on enhanced YOLOv3 under complex environments. Journal of Computer Applications, vol. 40, no. 9. pp. 2561-2570. 2020.

[6] J. REDMON, S. DIVVALA, R. GIRSHICK, et al, You only look once: unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, pp. 779-788. 2016.

[7] YP. WANG, Y. YANG and Y. YAO, Single shot multibox detector for ships detection in inland waterway. Journal of Harbin Engineering University, vol. 40, no. 7. pp. 1258-1262. 2019.

[8] M. HUANG, L. WANG and HC. ZHANG, Face Recognition Based on Gabor Wavelet Transform and K-L Gaussian Riemannian Manifold Discriminant. Computer Engineering, vol. 42, no. 9. pp. 208-213. 2016.

[9] K. ZHU and E. HOSSAIN, Virtualization of 5G cellular networks as a hierarchical combinatorial auction. IEEE Transactions on Mobile Computing, vol. 15, no. 10. pp. 2640–2654. 2016.

[10] O. AL-KHATIB, W. HARDJAWANA and B. VUCETIC, Spectrum sharing in multi-tenant 5G Cellular networks: Modeling and planning. IEEE Access, vol. 7. pp. 1602–1616. 2018.

[11] R. YIN, C. ZHONG, G. YU, et al., Joint spectrum and power allocation for D2D communications underlaying cellular networks. IEEE Transactions on Vehicular Technology, vol. 65, no. 4. pp. 2185–2195. 2016.

[12] P.A. Bliman, G. Ferrari-Trecate, Average consensus problems in networks of agents with delayed communications. Automatica, vol. 44, no. 8. pp. 1985-1995. 2013.

[13] N. TIM, E. ZORITA and B. HÜNICKE, Decadal variability and trends of the Benguela upwelling system as simulated in a high-resolution ocean simulation. Ocean Science, vol. 11, no. 3. pp. 483-502. 2015.

[14] G. HUANG, Y. SUN, Z. LIU, et al., Deep networks with stochastic depth. Proceedings of the 2016 European Conference on Computer Vision, LNCS 9908. Cham: Springer, pp. 646-661. 2014.

[15] T.Y. LIN, M. MAIRE, S. BELONGIE, et al., Microsoft COCO: common objects in context. Proceedings of the 2014 European Conference on Computer Vision, LNCS 8693. Cham: Springer, pp. 740-755. 2014.