

Research Article

Medical Image Retrieval Approach by Texture Features Fusion Based on Hausdorff Distance

Sun Xiaoming , Zhang Ning, Wu Haibin, Yu Xiaoyang, Wu Xue, and Yu Shuang

The Higher Educational Key Laboratory for Measuring & Control Technology and Instrumentations, Harbin University of Science and Technology, Heilongjiang Province, Harbin 150080, China

Correspondence should be addressed to Sun Xiaoming; xiaoming_66881982@163.com

Received 26 January 2018; Revised 7 May 2018; Accepted 21 May 2018; Published 27 August 2018

Academic Editor: Seungik Baek

Copyright © 2018 Sun Xiaoming 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.

Medical images play an important role in the hospital diagnosis and treatment, which include a lot of valuable medical information. Manually annotated viewing is obviously not effective in managing large amounts of medical imaging data. Hence it is an important task to establish an efficient and accurate medical image retrieval system. In this paper, a medical image retrieval approach based on Hausdorff distance combining Tamura texture features and wavelet transform algorithm is proposed. The combination of Tamura texture features and wavelet transform features can extract the texture features of medical images more effectively, and Hausdorff distance can reflect the overall similarity of medical image feature set. In this paper, 6 group experiments of brain MRI database and the lung CT database were conducted separately. Experiments show that the proposed approach has higher accuracy than a single feature texture algorithm and is also higher than the approach of Tamura texture features and wavelet transform features combined with Euclidean distance.

1. Introduction

With the development of big data technology and hospital information technology, medical images play an extremely important role in the hospital diagnosis and treatment. Each hospital generates a large number of medical images every day, and a single imaging equipment daily products data up to more than 10G. Thus the vast amount of data includes a lot of valuable medical information; it is unable to effectively manage the massive medical imaging data only by the manual annotation way [1]. Therefore, it is the important task to establish the efficient and accurate medical image retrieval system. The research of medical image retrieval is of a great significance to clinical medicine.

Due to the particularity of imaging modes and imaging characteristics of medical images, the majority of images (except B-images and pathological sections) are gray scale images. Thus color is not obvious for medical image to distinguish. Shape features are often associated with the target. In the same place, the images have high similarity, so the shape is not obvious for the retrieval of similar symptoms [2]. Hence, the texture feature is important for medical image

analysis, which is the main basis to determine pathology. Particularly in the pathological diagnosis, the textures of the surface of normal organs and diseased organ are very different, and the textures also reflect the spatial homogeneity of the pixels in the image. In this paper, we will focus on the texture features.

Texture features are used to capture the granularity and recurring patterns of the image surface. For color images and irregularly shaped images, researchers propose the image retrieval method based on multifeature fusion [3–6]. Liu Wanjun et al. calculated the Euclidean distance of the color histogram and the texture feature of the image, respectively. Finally, the Euclidean distances are fused to get the similarity. Geng Yanping extracted color features in HSV space and GLCM feature, and GLCM combined Tamura features to form more abundant texture features. Li Bing extracted the third-order color characteristics, three texture features by Tamura, and shape features by Fourier, which were fused to propose the image retrieval method based on multifeature fusion. Rehan Ashraf combined the edges of color histograms with discrete wavelet transforms to enrich performance of content-based retrieval.

Extracted texture features by wavelet transforms and Gray Level Concurrence Matrix (GLCM) were also often used in image retrieval [7–9]. The texture feature vectors of images were constituted by mean and standard variance of each subimages after wavelet transform decomposition and second moment, contrast, correlation coefficient, and entropy of GLCM. Then similarity measurement methods were used, for example, the different weights Euclidean distances or Canberra distance, as the measurement standard of the retrieval result. Kumar and Singh [10], P Srivastava [11], and Vijay Kumar Nath [12] combined the Discrete Wavelet Transform (DWT) with Local Binary Pattern. They retrieved the ear images, Corel-1K, Olivia-2688, Corel-5K, Corel-10K, and GHIM-10K images and biomedical images. M Hare [13] obtained the LBP descriptors of the two-dimensional gray image, calculated the DWT coefficients and completed the feature vector construction through the GLCM.

Zhou Jianlong extracted texture features by GLCM and Tamura from the DDSM database for medical image retrieval [14]. MA Anran et al. show that, according to the frequency characteristics of medical images, characteristics of frequency distribution and amplitude were obtained through analysis of medical images. According to the different frequency characteristics, an image retrieval method based on spectral similarity was proposed [15]. Guo Jing conducted the same experiments on the human brain and lungs. Tamura algorithm was more suitable for the feature extraction of human brain images under the same setting. The GLCM algorithm was more suitable for the human heart images [16]. Lu Xiaoqi finished image preprocessing on the ITK platform. Tamura algorithm was used to extract the texture features of the segmented images and was verified through heart CT images [17]. Gao Yuan calculated the iris structure density using Tamura texture features, and the relationship was determined between iris fiber structure density and selected features [18]. Mei Jun combined the directionality of Tamura, morphological operation with Gaussian blurring, and successfully generated the latitude and longitude organization point distribution map to realize the fabric organization identification [19].

M Kumar created image database of eyes, extracted the texture and intensity of eye images to form feature vectors, and used Euclidean distances to retrieve eye images [20]. X Ou described and extracted texture by Gabor wavelet transform to retrieve skin images [21]. SM Anwar used fast wavelet transform to extract the detail coefficients of images. The standard deviation and kurtosis of images were calculated, and the Euclidean distance was used to calculate the distance between the features of the query image and each images in the database [22]. Amin Khatami used wavelet transform to decompose and analyze medical images [23]. Zhang Chunyan et al. focused on the security of medical image retrieval, combined with the characteristics of Henon mapping to perform frequency domain cryptographic operations on the image, and then decomposed the encrypted image by wavelet, and an algorithm of encrypted medical image retrieval based on discrete wavelet and perceptual hash was proposed [24]. In order to distinguish between benign tuberculosis and malignant tuberculosis, Guohui Wei used three methods (LBP,

Gabor, and GLCM) to extract texture features and used SVM, ELM, Semantic, TSCBIR, and Mahalanobis distances for classification [25]. Xu Juanjuan used the wavelet transform to extract texture features and used Restricted Boltzmann Machine (RBM) to reduce the dimension to reduce the amount of computation [26]. Zhang Jianxun proposed a new medical image method based on Hausdorff distance combined with wavelet transform. The images were decomposed by wavelet transform, the feature points were detected by gradient vector flow, and the Hausdorff distance was used to match the set of feature points [27].

In summary, texture features are particularly significant for medical image retrieval. However, for current medical image retrieval methods, GLCM is mainly used for texture feature extraction of images, which still cannot directly reflect the texture features of diseased organs. Tamura texture features are more intuitive and visually advantageous than GLCM. Therefore, this paper combines wavelet and Tamura texture features to extract texture features.

In addition, similarity measurement is another key step in image retrieval accuracy. At present, while Euclidean distance is the commonly used measurement method, it has some shortcomings such as incomplete consideration of the nature of space, lack of clear physical meaning, and sensitivity to deformation of graphics. However, for other measurement method, Manhattan distance, it needs to rely on the system's rotation of the coordinates when calculating, rather than coordinate rotation and mapping. Canberra distance is greatly sensitive to change in value close to 0 (greater than or equal to 0), and the higher the variable, the stronger the sensitivity in the high-dimensional space. Mahalanobis distance is as sensitive as all features and cannot prominent the influence of major factors. In contrast, Hausdorff distance does not emphasize matching point in the image, but the maximal-minimum distance, which reflects the overall similarity of medical image feature set.

Texture analysis is a common method in pattern recognition and image processing. The wavelet transform can effectively separate and extract the low-frequency approximate information and the high-frequency detail texture information of medical images; the latter can display the details of the medical image and obtain more useful information for retrieval. Most of the primitives on the surface of diseased organs are irregular, so it is feasible to analyze the texture by Tamura, which is commonly used in statistical analysis. Therefore, we will combine the Tamura texture feature and the average coefficients obtained by the wavelet decomposition to form a new feature vector group in this paper. And, the Hausdorff distance of matching point set will be used in the medical image retrieval. Hence, in this paper, medical image retrieval approach by Hausdorff distance based on the Tamura texture features and wavelet transform features was proposed.

The main work of this paper is given as follows. Sections 2 and 3 demonstrate the proposed method. Section 2 gives the image preprocessing steps and the feature extraction method, that is, using Tamura texture features and wavelet transform algorithm to extract the effective texture features of medical

image data. Section 3 describes similarity measurement standard; i.e., we use Hausdorff distance to measure the medical image data. Section 4 demonstrates the proposed method through experiments. Section 5 summarizes this paper.

2. Research Methods

2.1. Image Preprocessing. Due to imaging devices, the medical images are different about the storage format and size and data formats. Medical images are also subject to noise, light, or human influence in the process of collection and transmission, resulting in blurred images and quality decline. In order to restore the sharpness of original image and improve the quality of the image, the image preprocessing is also an essential step [28]. The process of medical image preprocessing is described with Figure 1. First, since the medical images have different sizes, we use uniform quantization method to normalize the images. Then owing to the existence of noise produced by collection and transmission, the images are filtered by two-dimensional statistical sequence. This filtering method does not affect image quality and image feature extraction. Furthermore, to let grayscale distribution of medical images be more uniform and the detail description be clearer, the nonlinear transformation is used to enhance the images in the spatial domain.

2.2. Texture Feature Extraction

2.2.1. Tamura Feature Extraction. Tamura et al. proposed that there are six components of the Tamura texture features, which are named coarseness, contrast, directionality, linelikeness, regularity, and roughness, respectively. In this paper, three important features, coarseness, contrast, and directionality, are used.

(1) *Coarseness.* Computational process of coarseness is as follows:

(i) According to formula (1), in the active window of size $2^k \times 2^k$ in the medical image, the mean of the brightness of each pixel is calculated.

$$A_k(x, y) = \frac{\sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i, j)}{2^{2k}} \quad (1)$$

where (x, y) represents the position of the region of the selected medical image, $g(i, j)$ represents the mean of the brightness in the i^{th} , j^{th} points in the selected region, and k determines the range of the pixel.

(ii) According to formulae (2) and (3), the mean intensity difference between no overlapping activity windows in the horizontal and vertical directions is calculated.

$$E_{k,h} = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (2)$$

$$E_{k,v} = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})| \quad (3)$$

where $E_{k,h}$ represents the horizontal difference of this pixel and $E_{k,v}$ represents the vertical difference of this pixel.

(iii) Then, according to formula (4), the maximum of E is calculated and the optimum size of each pixel S_{best} is calculated according to formula (5).

$$E_k = E_{\max} = \max(E_1, E_2, \dots, E_n) \quad (4)$$

$$S_{best}(x, y) = 2^k \quad (5)$$

(iv) According to the formula (6), the mean value of S_{best} of all pixels in the entire image is calculated, which is called the coarseness F_{crs} .

$$F_{crs} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j) \quad (6)$$

where m is the length of the image and n is the width of the image.

(2) *Contrast.* Generally, the image contrast refers to the brightness level between the darkest black and the brightest white in an image. According to formula (7), the contrast of the medical image is calculated.

$$F_{con} = \frac{\sigma}{\alpha_4^{1/4}} \quad (7)$$

where σ is the standard deviation of gray value of the medical image and α_4 is the kurtosis of gray value of the medical image.

(3) *Directionality.* The directionality is particularly important in medical images, which reflects the texture direction of the human muscle or tissue. The calculation process is as follows:

(i) According to formulae (8) and (9), the modulus of the gradient vector and the local edge direction of each pixel in medical image are calculated.

$$|\Delta G| = \frac{(|\Delta H| + |\Delta V|)}{2} \quad (8)$$

$$\theta = \tan^{-1} \left(\frac{\Delta V}{\Delta H} \right) + \frac{\pi}{2} \quad (9)$$

Formulae (8) and (9) can be achieved by convolving a 3×3 rectangular area around a medical image pixel point with two 3×3 masks as shown in Figure 2.

(ii) Firstly, divide the region $0 - \pi$ into 16 equal parts and obtain the angle ϕ corresponding to the largest mode of the gradient vector in each equal interval. Then calculate pixel number n_p when $|\Delta G|$ in each region corresponding to the angle θ is greater than the threshold. Secondly calculate the number of gradient vectors of all pixels to construct the histogram H_D , and discrete the range values of this histogram, then the peak position of H_D is denoted by ϕ_p . Finally, according to formula (10), the overall direction of the medical image can be calculated by the sharpness of the peak in the histogram.

$$F_{dir} = \sum_P \sum_{\phi \in \omega_p}^{n_p} (\phi - \phi_p)^2 H_d(\phi) \quad (10)$$

where p is the peak and ω_p is the range of the peak between each valley.

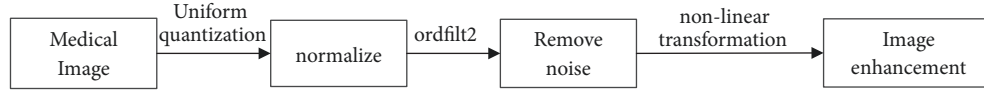


FIGURE 1: Flowchart of the image preprocessing.

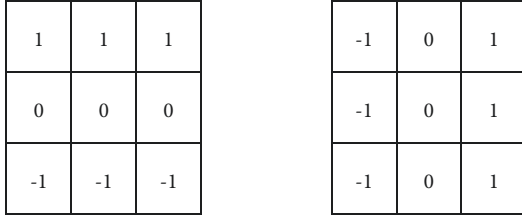
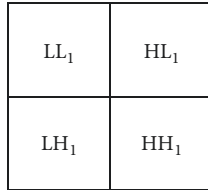
FIGURE 2: Two 3×3 masks.

FIGURE 3: Structure of 1-level discrete wavelet.

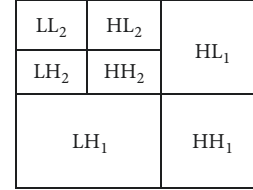


FIGURE 4: Structure of 2-level discrete wavelet.

TABLE 1: *ER* and *PSNR* of each wavelet.

Wavelet function	38.bmp		41.bmp	
	<i>ER</i>	<i>PSNR</i>	<i>ER</i>	<i>PSNR</i>
Dmey	0.9491	21.8719	0.9517	19.5713
Haar	0.9086	22.0671	0.8953	21.4771
Bior3.7	0.9783	22.0532	0.9757	20.0996
Rbio6.8	0.9789	22.0666	0.9793	19.6148
Coif4	0.9852	22.2586	0.9794	21.5337
Db4	0.9733	22.2749	0.9786	22.2923
Sym4	0.9856	24.7927	0.9847	23.5570

2.2.2. Calculate Texture Features by Wavelet Transform Algorithm

(1) *The Introduction of Wavelet Decomposition.* In this paper, the texture information of medical image is extracted after 2 times wavelet decomposition to form the eigenvector of image retrieval. The medical image is decomposed by 1-level wavelet to obtain 4 subimages, of which size is $1/4$ of the original medical image. As show in Figure 3, the upper left corner LL_1 is the low-frequency information of the image and the remaining is high-frequency information of the image, where LH_1 is the vertical component of the image, HL_1 is the horizontal component of the image and HH_1 is the diagonal component of the image. The 2-level wavelet decomposition of medical images is performed once again on the low-frequency component LL_1 obtained after the 1-level wavelet decomposition. As shown in Figure 4, we can see the 1 low-frequency component LL_2 and 6 high-frequency components. In addition, in this paper, as shown in Figures 5 and 6, the schematic of the brain MRI images and lung CT images by 1-level and 2-level discrete wavelet decomposition is also given.

Discrete wavelet transform is multilevel decomposition of low-frequency components of the original medical image to form wavelet decomposition tree, and finally the low-frequency approximation information and the high-frequency detail texture information of the medical image are effectively separated and extracted. Since high-frequency subimages reflect the detail features of the texture image, the

low-frequency subimages reflect the overall features of the texture image [29].

(2) *The Choice of Wavelet Basis.* In normal experiment, there are kinds of commonly used 15 wavelet, of which 7 kinds of discrete wavelet can be used in medical image feature extraction, namely, Haar, Daubechies (db N), Biorthogonal (biorNr.Nd), Coiflet (Coif-N), Symlets (symN), Dmeyer, and ReverseBior (rbioNr.Nd). By decomposing medical images, according to formula (11), formula (12), and formula (13), the energy ratio (*ER*) and peak signal to noise ratio (*PSNR*) of brain MRI images and lung CT images are calculated. As can be seen in the Table 1.

$$ER = \frac{E_{LL}}{E} \quad (11)$$

$$PSNR = \frac{-10 \lg MSE}{(255 \times 255)} \quad (12)$$

$$MSE = \frac{1}{M \times N} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} (f_{ij} - \tilde{f}_{ij})^2 \quad (13)$$

where *ER* is the low-frequency energy of image after wavelet transform, *E* is the total energy of image. *MSE* is the mean-square error, *M* and *N* are the Width and height of image, and f_{ij} and \tilde{f}_{ij} represent the pixel values of the original image and the reconstructed image at the coordinates (i, j) , respectively.

It can be seen from Table 1 that, under similar conditions, the sym4 has the highest value *PSNR* when processing the

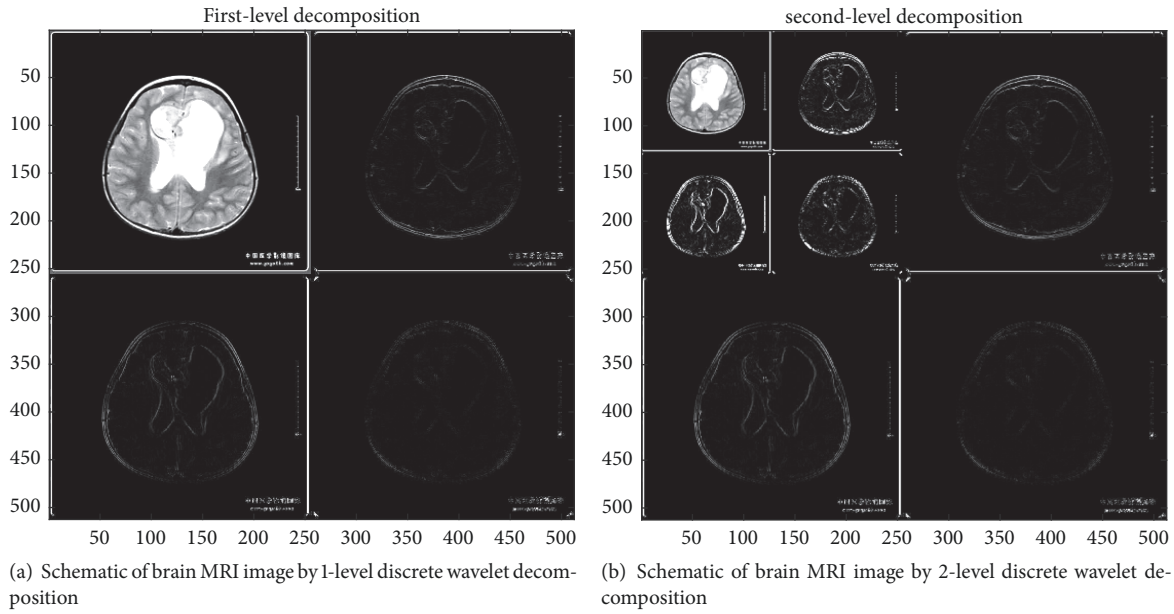


FIGURE 5: Schematic of brain MRI image by discrete wavelet decomposition.

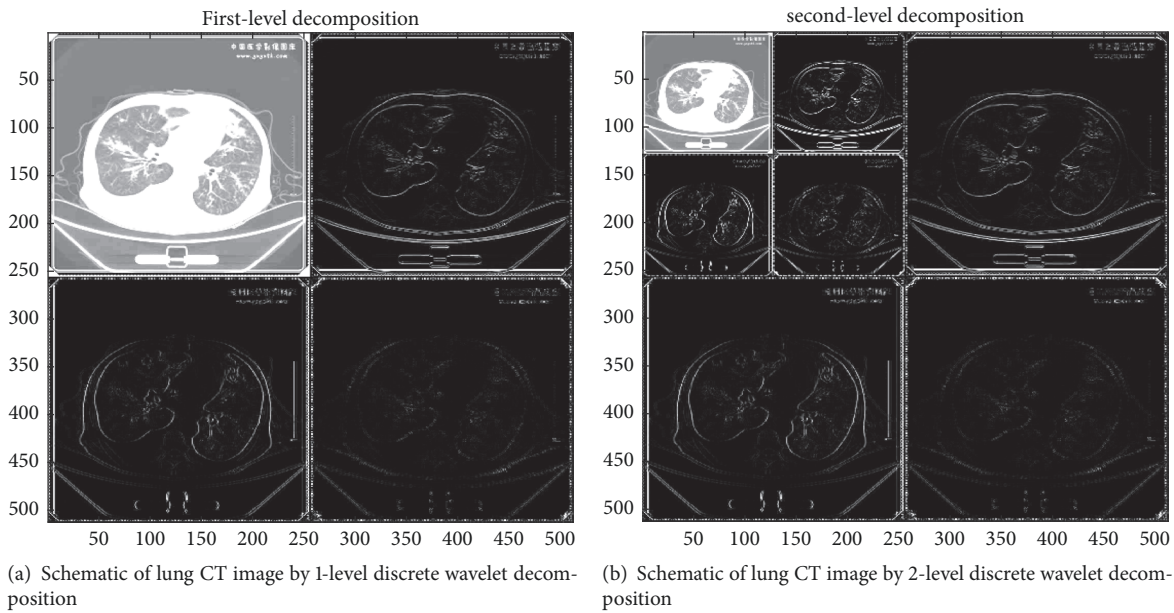


FIGURE 6: Schematic of Lung CT image by discrete wavelet decomposition.

two medical images. Symlet is the improved wavelet basis based on the Daubechies. It is better than Daubechies in a certain performance. During the experiment process, Coif-N wavelet needs setting the threshold by manual adjustment. In this paper, we choose the wavelet function sym4 as the medical image decomposition. Considering the existence of certain redundancy between the information and computer running speed, the image is decomposed by 2 times wavelet transform algorithm to get six high-frequency subimages. We choose the average coefficients of six high-frequency subimages, which are the average coefficient of the 2-level

horizontal component, the average coefficient of the 2-level vertical component, the average coefficient of the 2-level diagonal component, the average coefficient of the 1-level horizontal component, the average coefficient of the 1-level vertical component, and the average coefficient of the 1-level diagonal component, respectively. These coefficients are expressed as mcH2, mcV2, mcD2, mcH1, mcV1, and mcD1. Finally, a new set of eigenvectors with nine feature vectors is composed of the three features extracted from the Tamura texture features and the six average coefficients from the wavelet decomposition.

3. Similarity Measurement

Feature extraction is an important step of image retrieval. Similarity measurement is another key step in image retrieval. At present, distance and correlation are commonly used to measure the similarity. The shorter the distance between the features points of two images, the more similar the two images. The most commonly used methods are Euclidean distance, Hausdorff distance, Manhattan distance, and EMD distance.

Euclidean distance is the simplest similarity measurement method. It is easy to understand, but it is sensitive to the deformation of the image. Manhattan distance needs to rely on the system's rotation of the coordinates when calculating, rather than coordinate rotation and mapping. Hausdorff distance is not the corresponding distance between point and point, but the maximal-minimum distance. It belongs to the fuzzy matching between point sets, which reflects the overall similarity of medical image feature set.

In the following, we calculate the Hausdorff distance of two medical images. For a query medical image set $X = \{x_1, x_2, \dots, x_N\}$ and a certain image set $Y = \{y_1, y_2, \dots, y_M\}$ in medical image database. We use (14), (15), and (16) to measure the Hausdorff distance $H(X, Y)$ of similarity between two medical images.

$$H(X, Y) = \max \{h(X, Y), h(Y, X)\} \quad (14)$$

where

$$h(X, Y) = \max_{x \in X} \min_{y \in Y} \|x - y\| \quad (15)$$

$$h(Y, X) = \max_{y \in Y} \min_{x \in X} \|y - x\| \quad (16)$$

where $\|\bullet\|$ is the distance between feature points x and y in the medical image.

$h(X, Y)$ represents the forward Hausdorff distance from point x to point y . It calculates the distance about a certain point x of set X of a medical image to be retrieved and all points y of a certain image set Y in medical image database to get the smallest of all distances. All points x in the retrieved medical image set X repeat the above process, and each point x gets a corresponding minimum distance, then the maximum of the minimum distances is chosen as a result. $h(Y, X)$ is called the backward Hausdorff distance; its calculation process is the same as $h(X, Y)$. Finally the maximum between $h(X, Y)$ and $h(Y, X)$ is chosen as the Hausdorff distance of a queried medical image and a certain image in medical image database.

The above method only describes the Hausdorff distance between the queried medical image and one certain image in medical image database. As shown in Figure 7, according to the above method, the Hausdorff distance between the queried medical image and all images in medical image database are calculated. All Hausdorff distances are ranked from small to large, and the images in database corresponding to the first 10 Hausdorff distances are chosen in order as retrieval result.

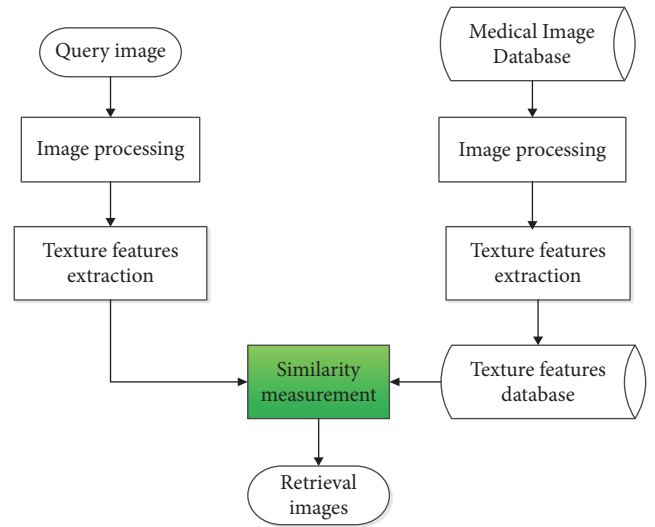


FIGURE 7: Flowchart of the retrieval process.

4. Experimental Results and Analysis

4.1. Image Database. In order to verify the performance of the proposed algorithm, we build a medical image database, including brain MRI image database and lung CT image database. China Medical Imaging Gallery provides the classification of diseases, for example, including brain tumor or tumor-like lesions and lung lesions. In our self-built brain image database, there are 122 brain images with jpg type, 512 * 512 size, and 96 dpi horizontal and vertical resolution. This brain database includes 38 ependymomas, 48 glioblastomas, and 36 hairy cell type astrocytomas. Also, for our self-built lung image database, there are 35 infective lesions, 25 lung Candida infections, and 33 inflammatory pseudotumor. For convince, S, J, X, H, Z, and Y are used to represent ependymomas, glioblastomas, hairy cell type astrocytomas, infective lesions, lung Candida infections, and inflammatory pseudotumor.

The experimental computer hardware configuration is Intel(R) Core(TM)i5-3230M CPU @ 2.60GHZ 2.60GHZ, the memory size is 4GB, and the hard disk size is 750GB. The software disposition is Windows7.0 operating system, MATLAB programming language.

4.2. Results of Image Preprocessing. The preprocessing results of brain MRI and lung CT image are shown in Figures 8 and 9. After normalization, noise removal, and enhancement, as shown in the histogram in Figures 8 and 9, the contrast and sharpness have been improved of these images.

4.3. Standards of Experimental Measurement. According to formula (17), we use the accuracy of image retrieval as the measurement standard of the retrieval result.

$$P_n = \frac{M}{N} \quad (17)$$

where N is the number of similar images and M is the number of same type images of queried medical image.

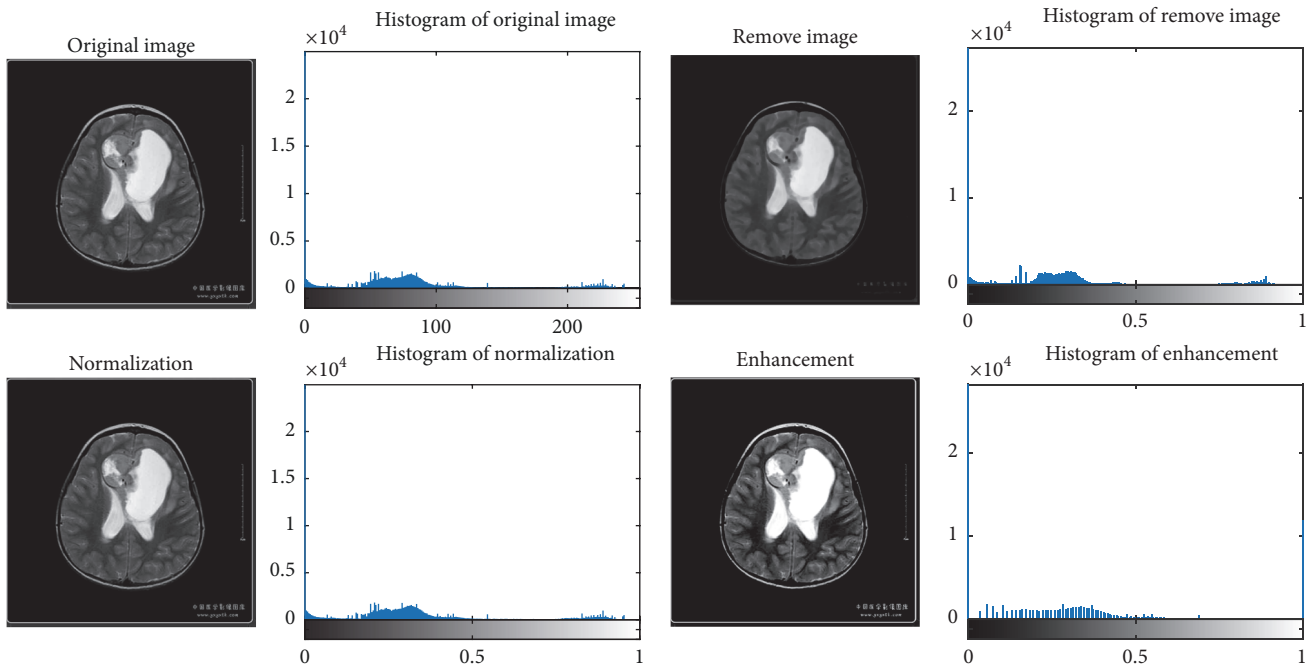


FIGURE 8: Results of brain MRI image preprocessing.

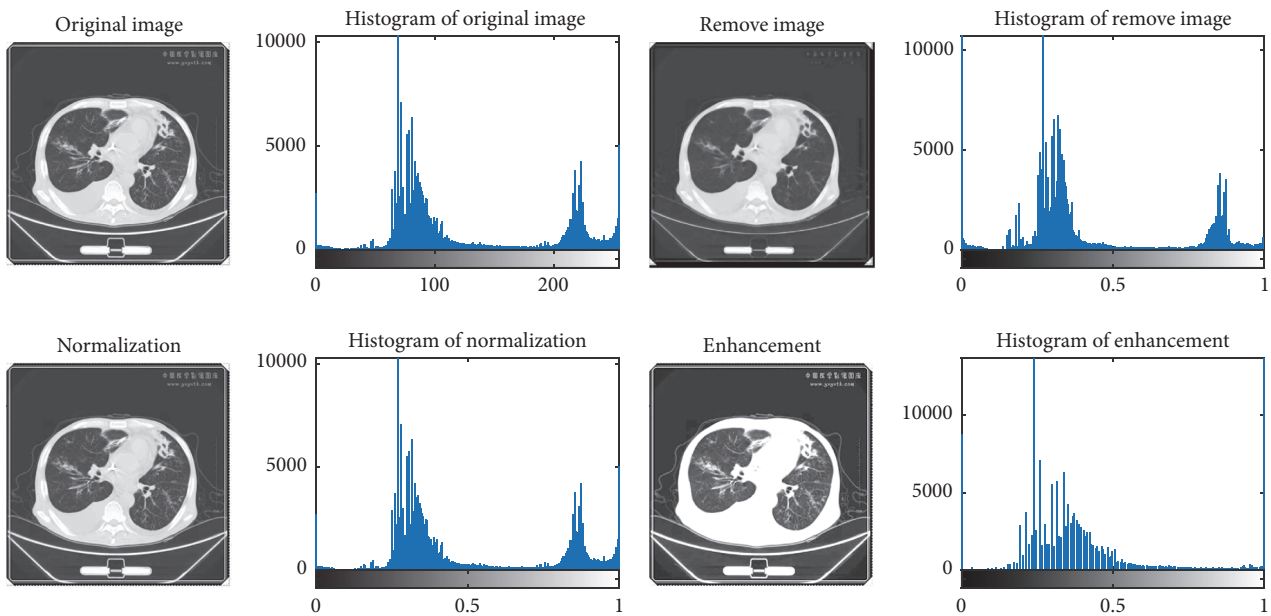


FIGURE 9: Results of lung CT image preprocessing.

For the 3 kinds of brain tumors, 5 medical images are randomly selected as the queried medical images in each kind, and the total number of queried medical images is 15. The retrieval accuracy of each queried medical images is p_n , where $n = 1, 2, \dots, 15$, the average retrieval accuracy is $p = \sum_{n=1}^{15} p_n / 15$.

For the 3 kinds of lung disease, 5 medical images are randomly selected as the queried medical images in each kind, and the total number of queried medical images is 15. The retrieval accuracy of each queried medical images is p_n ,

where, $n = 1, 2, \dots, 15$, the average retrieval accuracy is $p = \sum_{n=1}^{15} p_n / 15$.

4.4. Experimental Results and Analysis. In this paper, the medical image retrieval system is designed. The upper left of interface is the input image and the lower is the retrieval results. Click the “Browse” button in the image system; the image will be displayed in the upper left of the system. Click the “Retrieval” button in the image system, the top 10

medical images are displayed in the lower part of the medical image retrieval system interface. The Hausdorff distance or the Euclidean distance can be selected through the similarity drop-down list; Tamura texture, wavelet transform algorithm or the combination of Tamura texture, and wavelet transform algorithm can be selected through the texture feature extraction drop-down list. Click the “View-Retrieval-results” button in the image system, we can see the retrieved image category. In this paper, six experimental groups are carried out, which are Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance, Tamura texture feature and wavelet transform algorithm combined with Euclidean distance, wavelet transform algorithm combined with Hausdorff distance, Tamura texture feature combined with Hausdorff distance, wavelet transform algorithm combined with Euclidean distance, and Tamura texture feature combined with Euclidean distance. Taking the brain MRI image and lung CT image as example, the experimental results are as follows.

(i) *Experimental Results of Brain MRI Images.* The retrieved image is the 38th image in the brain MRI database. Figure 10 shows the retrieval results of Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance. Figure 11 shows the results of the Tamura texture feature and wavelet transform algorithm combined with Euclidean distance. Figure 12 is the result of wavelet transform algorithm combined with Hausdorff distance, and Figure 13 is the result of Tamura texture feature combined with Hausdorff distance, Figure 14 is the result of wavelet transform algorithm combined with Euclidean distance, and Figure 15 is the result of Tamura texture feature combined with Euclidean distance. The retrieved results of the 10th, 8th, 6th, 5th, 5th, and 4th image in the brain MRI database are, respectively, given in Figures 10, 11, 12, 13, 14, and 15. According to the experimental results, it is obvious that the Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance is priority. The Tamura texture feature and wavelet transform algorithm combined with Euclidean distance is lower. The wavelet transform algorithm combining Hausdorff distance and wavelet transform algorithm combined with Euclidean distance is follow. Tamura texture feature combining Hausdorff distance and Tamura texture feature combined with Euclidean distance are the lowest.

(ii) *Experimental Results of Lung CT Images.* The retrieval image is the 41th image in the lung CT database. Figure 16 shows the retrieval results of Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance, Figure 17 shows the results of Tamura texture feature and wavelet transform algorithm combined with Euclidean distance, Figure 18 is the result of wavelet transform algorithm combining Hausdorff distance, Figure 19 is the result of Tamura texture feature combining Hausdorff distance, Figure 20 is the result of wavelet transform algorithm combined with Euclidean distance, and Figure 21 is the result of Tamura texture feature combined with Euclidean distance. The retrieved results of the 10th, 9th, 8th, 6th, 5th, and 5th

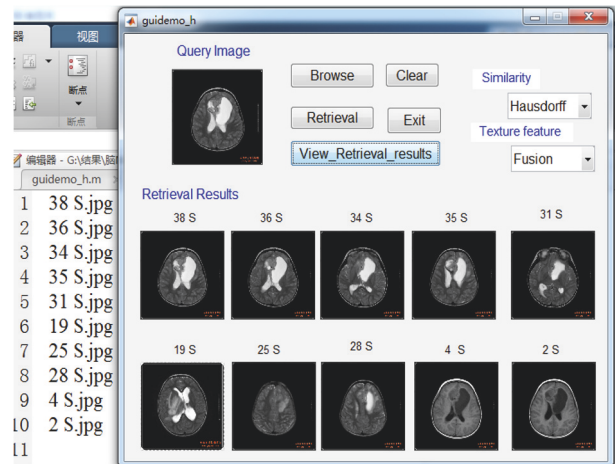


FIGURE 10: Retrieval results of Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance.

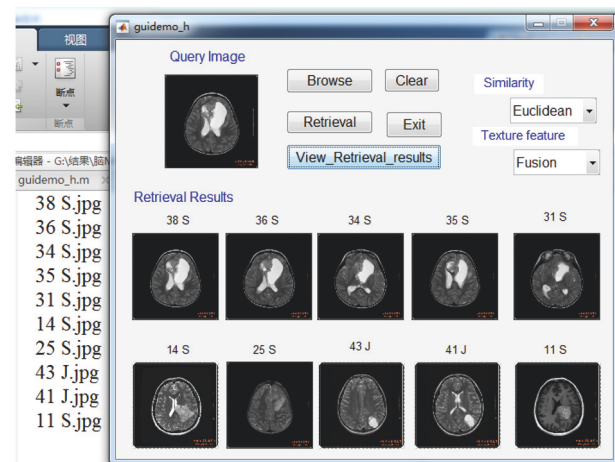


FIGURE 11: Retrieval results of Tamura texture feature and wavelet transform algorithm combined with Euclidean distance.

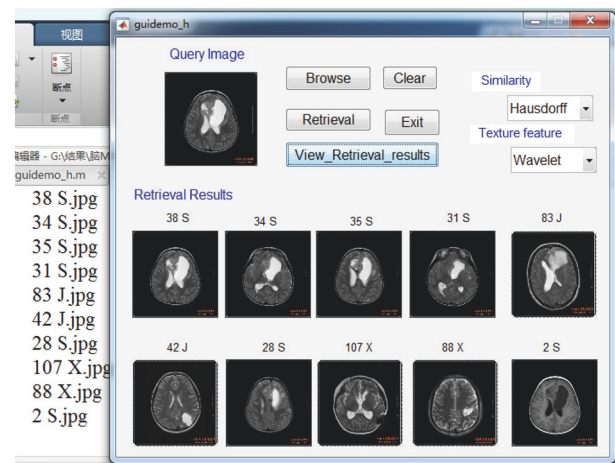


FIGURE 12: Retrieval results of wavelet transform algorithm combined with Hausdorff distance.

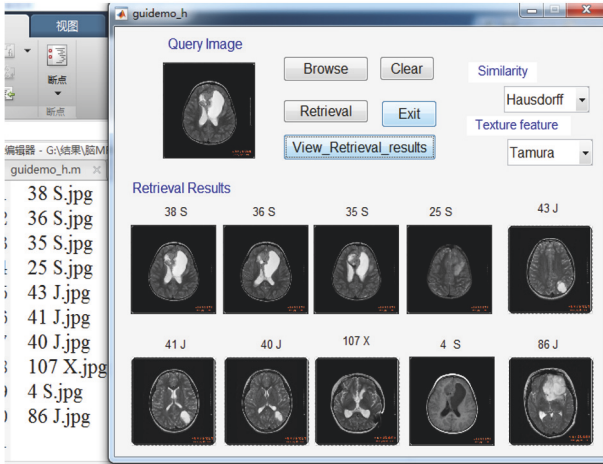


FIGURE 13: Retrieval results of Tamura texture feature combined with Hausdorff distance.

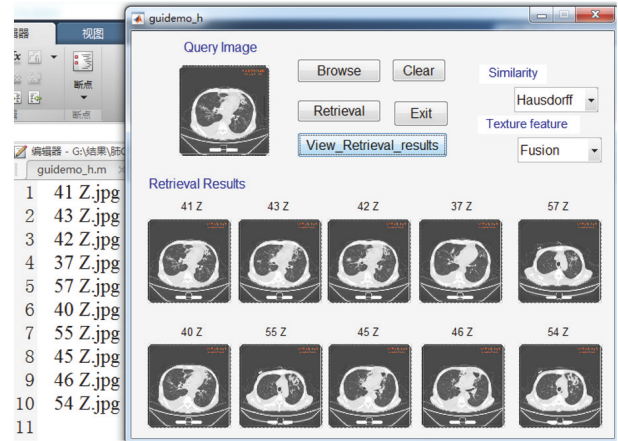


FIGURE 16: Retrieval results of Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance.

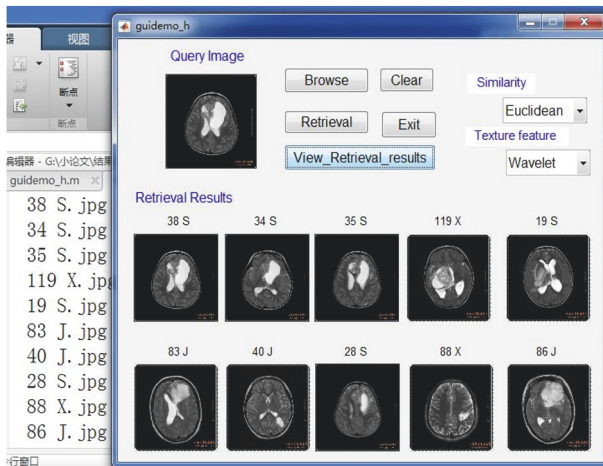


FIGURE 14: Retrieval results of wavelet transform algorithm combined with Euclidean distance.

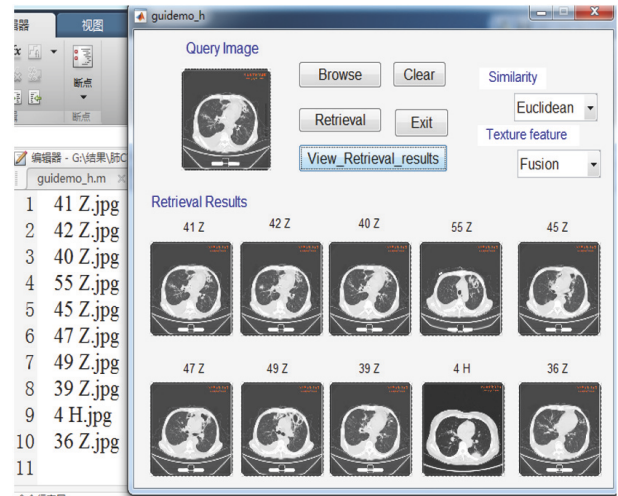


FIGURE 17: Retrieval results of Tamura texture feature and wavelet transform algorithm combined with Euclidean distance.

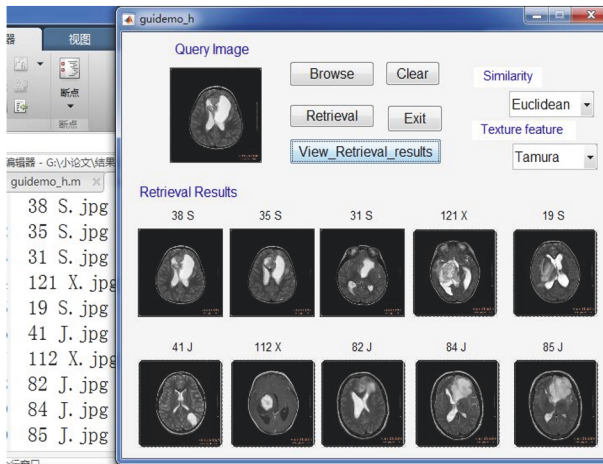


FIGURE 15: Retrieval results of Tamura texture feature combined with Euclidean distance.

image in the brain MRI database are, respectively, given in Figures 16, 17, 18, 19, 20, and 21. According to the experimental results, it is obvious that the Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance is priority. The Tamura texture feature and wavelet transform algorithm combined with Euclidean distance is lower. The wavelet transform algorithm combining Hausdorff distance wavelet transform algorithm combined with Euclidean distance follow. Tamura texture feature combined with Hausdorff distance and Tamura texture feature combined with Euclidean distance are the lowest.

(iii) *Data Quantitative Analysis.* Take the brain MRI image as the image to be retrieved; the average accuracy of wavelet transform algorithm combined with Euclidean distance, the average accuracy of Tamura texture feature combined with Euclidean distance, the average accuracy of Tamura texture feature combining Hausdorff distance, the average accuracy of wavelet transform algorithm combining Hausdorff distance,

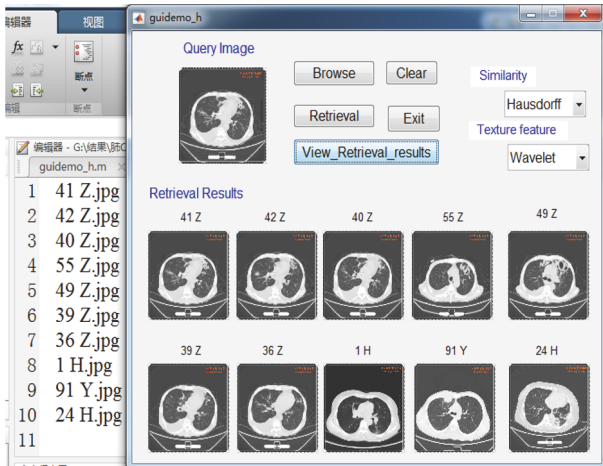


FIGURE 18: Retrieval results of wavelet transform algorithm combined with Hausdorff distance.

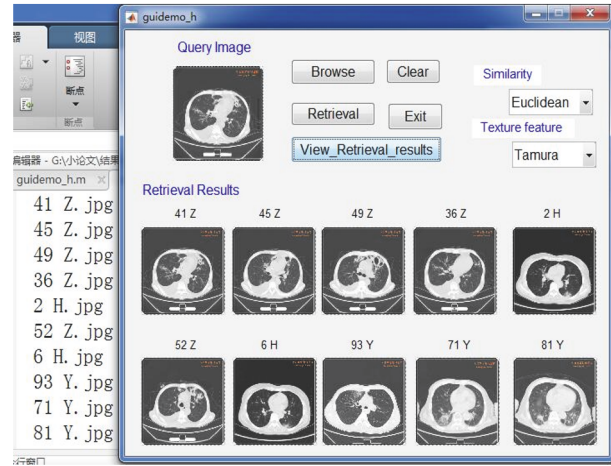


FIGURE 21: Retrieval results of Tamura texture feature combined with Euclidean distance.

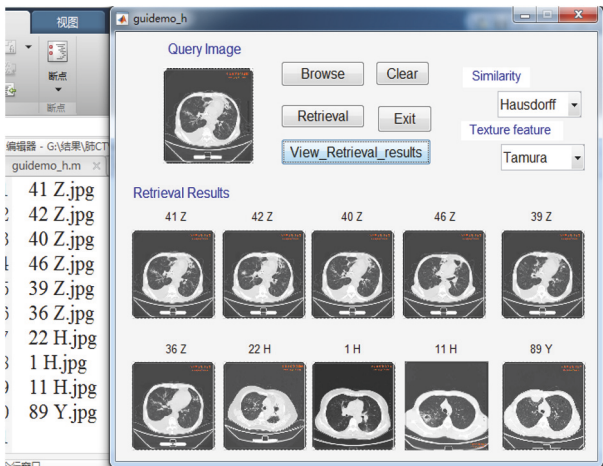


FIGURE 19: Retrieval results of Tamura texture feature combined with Hausdorff distance.

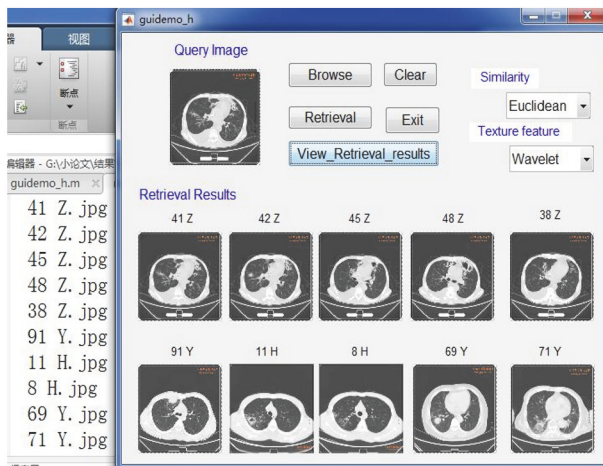


FIGURE 20: Retrieval results of wavelet transform algorithm combined with Euclidean distance.

TABLE 2: The average accuracy of brain MRI images each method.

Features	Similarity(%)	
	Euclidean distance	Hausdorff distance
Tamura feature	64	65.33
Wavelet transform	66	72
Tamura texture feature combined with wavelet transform	77.33	83.33

the average accuracy of Tamura texture feature and wavelet transform algorithm combined with Euclidean distance, and the average accuracy of Tamura texture feature and wavelet transform algorithm combined with Hausdorff distance are 64%, 66%, 65.33%, 72%, 77.33%, and 83.33%, respectively. Take the lung CT image as the image to be retrieved; the average accuracy of wavelet transform algorithm combined with Euclidean distance, the average accuracy of Tamura texture feature combined with Euclidean distance, the average accuracy of Tamura texture feature combining Hausdorff distance, the average accuracy of wavelet transform algorithm combining Hausdorff distance, the average accuracy of the Tamura texture feature and the wavelet transform algorithm combined with Euclidean distance, and the average accuracy of the Tamura texture feature and the wavelet transform algorithm combined with Hausdorff distance are 68%, 69.33%, 68.67%, 73.33%, 78.67%, and 85.33%, respectively, in Tables 2 and 3. According to the experimental data, it is obvious that the Hausdorff distance combined with Tamura texture feature and wavelet transform algorithm is priority. The Euclidean distance combined with Tamura texture feature and wavelet transform algorithm is lower; the wavelet transform algorithm combined with Hausdorff distance follows; Tamura texture feature combined with Hausdorff distance is the lowest.

TABLE 3: The average accuracy of lung CT images each method.

Features	Similarity(%)	
	Euclidean distance	Hausdorff distance
Tamura feature	68	68.67
Wavelet transform	69.33	73.33
Tamura texture feature combined with wavelet transform	78.67	85.33

5. Conclusion

With the development of hospital information technology, the effective management of massive medical images is very important. Retrieving useful images in the mass medical image database has become a worthy research problem. In order to solve this problem, a medical image retrieval approach based on Hausdorff distance combined with Tamura texture feature and wavelet transform algorithm is proposed in this paper. Texture is an important feature of medical images, and different types of disease have different characteristics in texture distribution. Thus, in order to increase the texture information, the proposed approach combines Tamura and wavelet transform algorithm to extract texture features. Similarity measurement is another important factor in image retrieval; however, European distance is a point to point distance, which cannot reflect the overall similarity of medical image feature set. Thus we use Hausdorff distance for similarity measurement in the proposed approach.

We build a medical image database, including brain MRI image database and lung CT image database. Experiments show that the proposed approach has higher accuracy than a Tamura feature texture and wavelet transform algorithm and is also higher than the approach of Tamura texture feature and wavelet transform algorithm combined with Euclidean distance. The average retrieval accuracy of the proposed approach is higher than other five approaches up to 19.33%, 17.33%, 18%, 11.33%, and 6% in the brain MRI experiments. For lung CT experiments, the average retrieval accuracy of proposed approach is higher than other three approaches up to more than 17.33%, 16%, 16.66%, 12%, and 6.66 %.

The future works are as follows. (1) The algorithm of this paper only validates the medical image database of brain MRI and lung CT, and we will increase the image database of chest and other medical images to further verify the accuracy of the algorithm. (2) Further we will study the Hausdorff distance and improve the Hausdorff distance to reduce the noise, occlusion, background, and other abnormal points of medical images and improve the accuracy of the similarity measurement.

Data Availability

The medical images, including brain tumor or tumor-like lesions and lung lesions, are available in the China Medical Imaging Gallery (<http://www.yxyxk.com/tuku.php>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (61671190 and 61401126), Natural Science Foundation of Heilongjiang Province of China (QC2015083), and University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2017086).

References

- [1] MD James Smirniotopoulos, "MedPix Medical Image Database", 2014.
- [2] *The Research of Medical Image Retrieval Based on Joint Index*, Jilin University, 2017.
- [3] L. Wanjuan and Z. Hongwei, "Image retrieval of improved color and texture features based on wavelet transform," *Computer Engineering and Applications*, vol. 52, no. 17, pp. 181–186, 2016.
- [4] G. Yanping, G. Hongbing, and R. Zhiying, "Image retrieval algorithm combining color feature and texture feature," *Wireless Internet Technology*, 2017.
- [5] L. Bing, "On technology of computer image retrieval based on multi feature fusion," *Journal of Southwest China Normal University (Natural Science)*, vol. 42, no. 1, pp. 54–59, 2017.
- [6] R. Ashraf, M. Ahmed, S. Jabbar et al., "Content Based Image Retrieval by Using Color Descriptor and Discrete Wavelet Transform," *Journal of Medical Systems*, vol. 42, no. 3, 2018.
- [7] Z. AoBo, W. Xianbin, and X. Zhang, "Retrieval Algorithm for Texture Image Based on Improved Dual Tree Complex Wavelet Transform and Gray Gradient Co-occurrence Matrix," *Computer Science*, vol. 44, no. 6, pp. 274–277, 2017.
- [8] J. Shiyong and M. Jiaoting, "Tire Pattern Retrieval Based on Wavelet Transform and Gray Level Co-occurrence Matrix," *Computer Measurement & Control*, pp. 210–213, 2016.
- [9] X. Kong, M. Hamit, C. Yan, J. Sun, and J. Yao, "Feature extraction on xinjiang high morbidity esophagus cancer based on wavelet transform," in *Proceedings of the 2016 International Conference on Biotechnology and Medical Science*, pp. 174–182, Nanjing, Jiangsu, China.
- [10] M. Kumar and K. M. Singh, "Content based medical image retrieval system using DWT and LBP for ear images," *International Science Press*, vol. 9, no. 40, pp. 353–358, 2016.
- [11] P. Srivastava and A. Khare, "Integration of wavelet transform, Local Binary Patterns and moments for content-based image retrieval," *Journal of Visual Communication and Image Representation*, vol. 42, pp. 78–103, 2017.
- [12] V. K. Nath, R. Hatibaruah, and D. Hazarika, "An Efficient Multi-scale Wavelet Local Binary Pattern for Biomedical Image Retrieval," in *Proceedings of the International Conference on Computing and Communication Systems*, vol. 24 of *Lecture Notes in Networks and Systems*, pp. 489–497, Springer Singapore, Singapore, 2018.
- [13] M. Khare, P. Srivastava, J. Gwak, and A. Khare, "A Multiresolution Approach for Content-Based Image Retrieval Using Wavelet Transform of Local Binary Pattern," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial*

- Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 10752, pp. 529–538, 2018.
- [14] J. Zhou, C. Feng, X. Liu, and J. Tang, “A texture features based medical image retrieval system for breast cancer,” in *Proceedings of the International Conference on Computing Convergence Technology*, vol. 8562, pp. 1010–1015, 2013.
- [15] M. Anran, R. Ning-ning, H. Li-bo, S. Yong, Y. SHAO, and Q. Jian-feng, “Image retrieval based on medical digital X-ray spectrum characteristics,” *Chinese Journal of Medical Physics*, vol. 33, no. 9, pp. 933–938, 2016.
- [16] G. Jinge, *The Research and application of medical image Retrieval technology based on texture feature*, Inner Mongolia University of Science Technology, 2012.
- [17] R. Luo, L. Rui, R. G. Liu, C. Ye, and Z. Z. Guan, “Total knee arthroplasty for the treatment of knee osteoarthritis caused by endemic skeletal fluorosis,” *Chinese Journal of Tissue Engineering Research*, vol. 16, pp. 1555–1557, 2012.
- [18] Y. Gao, H. Song, and Z. J. Zhang, “Calculation Method of Iris Structure Density Based on Tamura Features,” *Computer Technology & Development*, vol. 26, no. 3, pp. 36–39, 2016.
- [19] M. Jun, Z. Senlin, and F. Zhen, “Recognition Algorithm of Fabric Based on Tamura Texture Features,” *Light Industry Machinery*, vol. 35, no. 4, 2017.
- [20] M. Kumar, “Content based medical image retrieval using texture and intensity for eye images,” *International Journal of Scientific & Engineering Research*, vol. 7, no. 9, 2016.
- [21] X. Ou, W. Pan, X. Zhang, and P. Xiao, “Skin image retrieval using Gabor wavelet texture feature,” *International Journal of Cosmetic Science*, vol. 38, no. 6, pp. 607–614, 2016.
- [22] S. M. Anwar, F. Arshad, and M. Majid, “Fast wavelet based image characterization for content based medical image retrieval,” in *Proceedings of the 2017 International Conference on Communication, Computing and Digital Systems, C-CODE 2017*, pp. 351–356, Pakistan, March 2017.
- [23] A. Khatami, A. Khosravi, T. Nguyen, C. P. Lim, and S. Nahavandi, “Medical image analysis using wavelet transform and deep belief networks,” *Expert Systems with Applications*, vol. 86, pp. 190–198, 2017.
- [24] Z. Chunyan, L. Jingbing, and W. Shuangshuang, “Encrypted image retrieval algorithm based on discrete wavelet transform and perceptual hash,” *Journal of Computer Applications*, vol. 38, no. 2, pp. 539–544, 2018.
- [25] G. Wei, H. Cao, H. Ma, S. Qi, W. Qian, and Z. Ma, “Content-based image retrieval for Lung Nodule Classification Using Texture Features and Learned Distance Metric,” *Journal of Medical Systems*, vol. 42, no. 1, 2018.
- [26] J.-J. Xu, C. Chen, and H.-J. Yang, “Image retrieval method based on Gabor filter and deep learning,” *Shenyang Gongye Daxue Xuebao/Journal of Shenyang University of Technology*, vol. 39, no. 5, pp. 529–534, 2017.
- [27] J. Zhang and Y. Liu, “Medical Image Registration Based on Wavelet Transform Using Hausdorff Distance,” in *Transactions on Edutainment VII*, vol. 7145 of *Lecture Notes in Computer Science*, pp. 248–254, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [28] Wang. Hong-Jun, SHI. Nan, Hui. ZHAO, and YUE. You-Jun, “Image Preprocessing Technology Based on Improved Median Filter,” *Computer Systems & Applications*, vol. 24, no. 5, pp. 237–240, 2015.
- [29] D. Peng and C. University, “Wavelet Analysis Overview and its Application,” *Equipment Manufacturing Technology*, vol. 119, no. 1, pp. 74–79, 2014.

