Journal of Engineering and Applied Sciences 12 (11): 2946-2950, 2017
ISSN: 1816-949X

Content Based Medical Image Retrieval System using SIFT and Hu-Moment for Hepatobiliary Images

[1]Manoj Kumar and [2]Kh. Manglem Singh
[1]Department of Electronics and Communication Engineering,
[2]Department of Computer Science Engineering,
National Institute of Technology Manipur, Imphal, India

Abstract: This study presents retrieval of hepatobiliary images based on content-based medical image retrieval technique. Hepatobiliary images are related to hepatobiliary diseases. Hepatobiliary diseases include a group of diseases related to liver and biliary system. For image retrieval, low-level feature is used. For feature extraction sift (scale invariant feature transform) and Hu-moments methods are used. Features extracted by these methods are combining to form a single feature vector. Feature extracted by query images and images which are stored in the database are compared using Euclidean distance method. Based on minimum distance images are retrieved. For performance measurement precision, recall and error rate are measured.

Key words: Scale Invariant Feature Transform (SIFT), Content Based Medical Image Retrieval (CBMIR), Difference of Gaussian (DOG), Hu-moment, Euclidean distance, precision, recall

INTRODUCTION

Content Based Medical Image Retrieval (CBMIR) system uses the low level feature for image retrieval. CBMIR is more efficient compare to text-based approach. Color, texture and shape are used as a low level feature. These features are derived from the image itself. SIFT (Scale Invariant Feature Transform) is invariant to image scale, rotation and translation. Now a days, sift are used to describe the content of images. Local features are presented by sift. Sift mainly detect keypoints and descriptors from the image. Then, these are used to find the similar image from the database compares to query image. Content-based image retrieval using sift for binary and grayscale images was presented in 2014 (Bakar *et al*., 2013). In this approach, researcher show the excellent experimental result for the image retrieval. A comparative study of sift and PCA for content based image retrieval was presented by Reddy and Narayana (2016). In this study researcher got better image retrieval performance compare to PCA approach. Mohammed Alkhawlani, Mohammed Elmogy and Hazem Elbakry use SIFT approach for content-based image retrieval (Alkhawlani *et al*., 2015). A survey research paper for content-based image retrieval system using SIFT was presented in 2014 (Velmurugan, 2014). Hu-moment is used for shape feature extraction. A comparative study on Feature extraction using texture and shape for content based image retrieval was presented in 2015 (Bagri and Johari, 2015). In this study researcher has used Hu-moment for shape feature shape feature extraction (Zhao *et al*., 2016). A basic block diagram for content-based image retrieval is shown in Fig. 1 (Ramamurthy *et al*., 2012). In this CBIR system, feature extracted from the database are compared with query feature. Similar images are retrieved based on minimum distance.

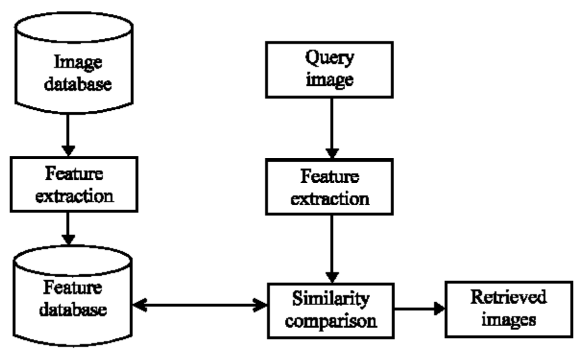

Fig. 1: General CBIR system architecture

MATERIALS AND METHODS

Proposed work: For feature extraction, SIFT and Hu moment algorithm are used. Block diagram for proposed work is shown in Fig. 2. Algorithm discussed stepwise are as follow:

Corresponding Author: Manoj Kumar, Department of Electronics and Communication Engineering, National Institute of Technology Manipur, Imphal, India

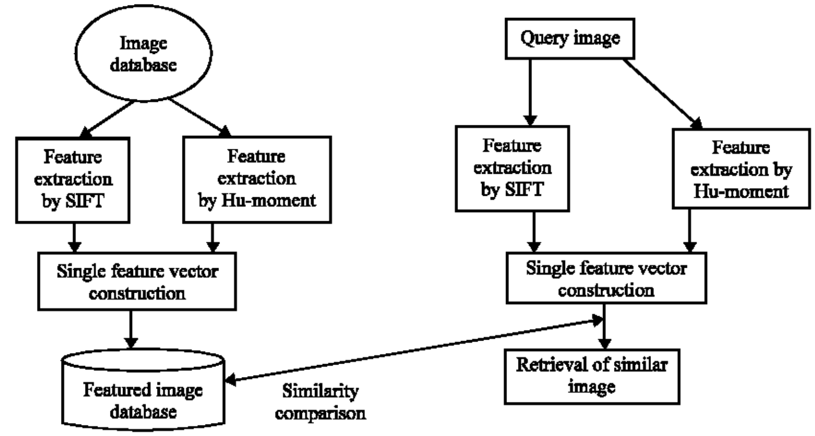

Fig. 2: System flow diagram

- Step 1: create a database containing various hepatobiliary images
- Step 2: extract features using SIFT and Hu-moment.
- Step 3: both these feature vector are combined to form a single feature vector
- Step 4: distance between the feature vector of query images and the database images are calculated
- Step 5: based on minimum distance images are retrieved

Feature extraction by SIFT: Steps for SIFT algorithm are discussed.

Step 1 (scale-space extrema detection): Difference of Gaussian (DOG) function is used to detect extrema points:

$$L(x, y, \sigma) = G(x, y, \sigma)\times I(x, y) \quad (1)$$

Where:

$L(x, y, \sigma)$ = Represents the scale space of an input image $(I(x, y))$
$G(x, y, \sigma)$ = Variable scale gussian function
$*$ = The convolution operator

$$G(x,y, \sigma) = 1/2\pi\sigma^2 \ e \ (x^2+y^2)/2\sigma^2 \quad (2)$$

Stable keypoints are detected by using the equation given (Eq. 3):

$$D(x, y, \sigma) = L(x, y, \underline{k}\sigma) - L(x, y, \sigma) \quad (3)$$

where, k is a constant multiplicative factor (Lowe, 2004).

Step 2 (keypoint location:) Taylor expansion series are used to localize the stable keypoints:

$$D(x) = D+\frac{dD^{t}X}{dx}+\frac{1}{2}XT\frac{d^2DX}{dx^2} \quad (4)$$

where, D is the difference of Gaussian. Then, the extrema X is determined by:

$$X = \frac{-d^2D^{-1}}{dx^2}\frac{dD}{dx^2}$$

Hession matrix is used to find local maxima:

$$H = D_{xx}\ D_{xy}, D_{xy}\ D_{yy}$$

Where:

D_{xx} = D(x, y-1)-2 D(x, y)+D(x, y+1)
D_{yy} = D(x-1, y)-2 D(x,y)+D(x+1, y)
D_{xy} = D(x, y -1)-2 D(x, y) D (x-1, y)-D (x-1, y-1)

$$Tr(H) = D_{xx}+D_{yy} = a+B;$$
$$Det(H) = D_{xx}\ D_{yy}-(D_{yy})^2 = a\times B;$$
$$Ratio = (a+B)^2/a\times B$$

Hession matrix is calculated for each keypoint selected by second order derivative of D. Keypoints are removed which having a value less than ratio. Figure 3 represents keypoints for Fig. 4.

Step 3 (orientation assignment): Based on local image gradient directions one or more orientation is assigned to each keypoint (Lowe, 2004).

Step 4 (keypoint descriptor): At each keypoint local image gradients are measured at selected scale (Lowe, 2004).

Fig. 3: Hepatobiliary diseases image

Fig. 4: Key point for figure 3

Feature extraction by Hu-moment: Hu-moment gives seven-moments. These moments are invariant to rotation, scaling and translation. Central moment is given by equation:

$$\mu_{p,q} = \sum (x - x_c)^p \, (y - y_c)^q \, f(x, y)$$

where, center of the object is given by (x_c, y_c). Normalized central moment is given by equation:

$$\eta_{p,q} = \mu_{p,q} / \mu_{00}^{\gamma}$$

$$\gamma = \left[\frac{p+q}{2}\right] + 1$$

Seven-moment equation is given:

$$\varnothing 1 = \mu_{2,0} + \mu_{0,2}$$

$$\varnothing 2 = (\mu_{2,0} - \mu_{0,2})^2 + 4\mu^{1}1, 1$$

$$\varnothing 3 = (\mu_{3,0} - 3\mu_{1,2})^2 + (\mu_{3,0} - 3\mu_{2,1})^2$$

$$\varnothing 4 = (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{0,3} + \mu_{2,1})^2$$

$$\varnothing 5 = (\mu_{3,0} - \mu_{1,2})(\mu_{3,0} + \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2] + (3\mu_{2,1} - \mu_{0,3})(\mu_{2,1} + \mu_{0,3})[3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2$$

$$\varnothing 6 = (\mu_{2,0} - \mu_{0,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2] + 4\mu_{1}, 1(\mu)(\mu_{2,1} + \mu_{0,3})$$

$$\varnothing 7 = (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2] - (\mu_{3,0} - \mu_{1,2})(\mu_{2,1} + \mu_{0,3})[3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2]$$

These seven moments are used to extract a feature from an image.

Feature vector construction: Feature extracted by SIFT and Hu-moment are combined to form a single feature vector. For combining one or more feature vector fusion methods are used. Sift features are fused to Hu-moment features to generate combined feature vector (Kavitha *et al*., 2011).

RESULTS AND DISCUSSION

Image database is of three types 1; Radiology 2 and 3; Histlogy; Radiology database contains images related to hepatobiliary diseases which are taken by radiology imaging technique. Radiology database contains >200 images. Pathology database contains images related to hepatobiliary diseases which are taken by pathology imaging technique. Pathology database contains around 300 images. Histology database contains hepatobiliary diseases which are taken by histology imaging technique. Histology database contains >200 images. The proposed method has been implemented using Matlab (Fig. 5-10).

Image retrieval using euclidean method: The distance between the query image and database image are calculated by using Euclidean method (Veni and Narayanankutty, 2010). Threshold value is being set to 0.1 (Chandrakar *et al*., 2011). By using this threshold value image is retrieved correctly. Euclidean distance method formula is given:

$$d(A^1, A^Q)\sqrt{\sum_{i=1}^{n} \left(A_i^1 - A_i^Q\right)^2}$$

Where:

A^1 = The image in the database
A^Q = The query image for retrival

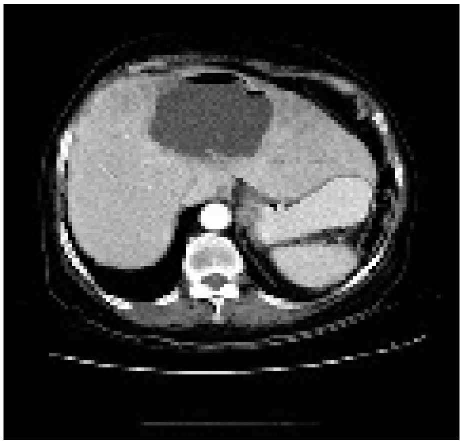

Fig. 5: Query image for the radiology database

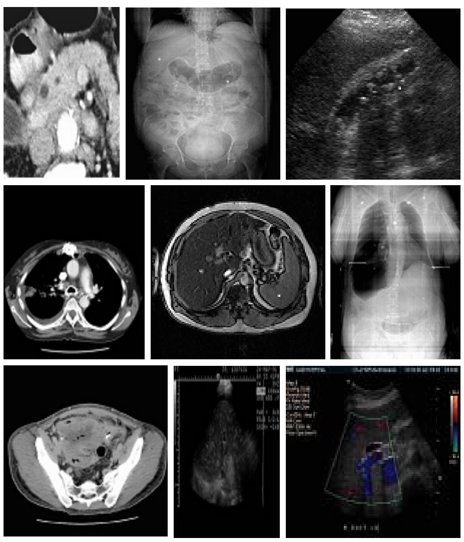

Fig. 6: Images in the radiology database

Fig. 7: Query image for the pathology database

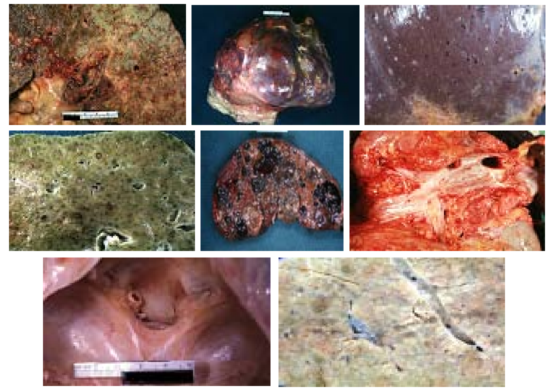

Fig. 8: Images in the pathology database

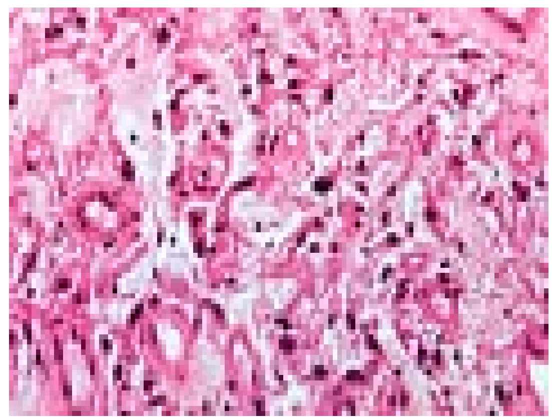

Fig. 9: Query image for the histology database

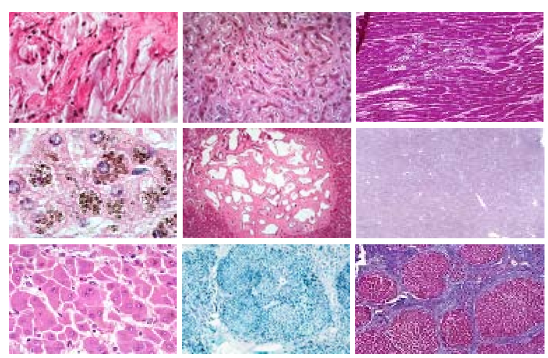

Fig. 10: Images in the histology database

Retrieval efficiency: Precision, recall and error-rate are measured for retrieval efficiency. Results are shown in Table 1-3. For performance measurement retrieval time are measured for all databases:

$$\text{Precision} = \frac{\text{No. of relevent images retrieved}}{\text{Total number of images retrieved}}$$

Table 1: Experimental results for radiology database

Database type	Precision (%)	Recall (%)	Error rate (%)	Time taken for retrieval in seconds
Radiology	94.71	100	5.28	3.6118

Table 2: Experimental results for pathology database

Database type	Precision (%)	Recall (%)	Error rate (%)	Time taken for retrieval in seconds
Pathology	95.04	100	4.95	4.109085

Table 3: Experimental results for histology database

Database type	Precision (%)	Recall (%)	Error rate (%)	Time taken for retrieval in seconds
Histology	94.58	100	5.41	3.6969

$$\text{Recall} = \frac{\text{No.of relevent images retrieved}}{\text{Total number of relevent images in the database}}$$

$$\text{The databse error rate} = \frac{\text{No. of non-relevant images retrieved}}{\text{Total number of images retrieved}}$$

CONCLUSION

In this study, images are retrieved by using sift and Hu-moment feature vector. Both these features vectors are combined to form a single feature vector. The experiment is conducted for three types of image database. Precision, recall and error-rate are measured for retrieval efficiency. By using sift and Hu-moment performance of the system has been improved.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to the people whose assistance has been invaluable in this research. I would like to thank the Director of National Institute of Technology, Manipur for providing financial and infrastructure support for making this work possible.

REFERENCES

Alkhawlani, M., M. Elmogy and H. Elbakry, 2015. Content-based image retrieval using local features descriptors and bag-of-visual words. Int. J. Adv. Comput. Sci. Appl., 6: 212-219.

Bagri, N. and P.K. Johari, 2015. A comparative study on feature extraction using texture and shape for content based image retrieval. Intl. J. Adv. Sci. Technol., 80: 41-52.

Bakar, S.A., M.S. Hitam and W.N.J.H.W. Yussof, 2013. Content-Based image retrieval using SIFT for binary and greyscale images. Proceedings of the 2013 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), October 8-10, 2013, IEEE, Kuala Terengganu, Malaysia, ISBN: 978-1-4799-0269-9, pp: 83-88.

Chandrakar, A., A.S. Thoke and B.K. Singh, 2011. Indexing and retrieval of medical images using CBIR approach. Proceedings of the 1st International Conference on Parallel Distributed Computing Technologies and Applications, September 23-25, 2011, Springer, Berlin, Germany, pp: 393-403.

Kavitha, C., D.B.P. Rao and D.A. Govardhan, 2011. Image retrieval based on color and texture features of the image sub-blocks. Intl. J. Comput. Appl., 15: 33-37.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60: 91-110.

Ramamurthy, B., K.R. Chandran, V.R. Meenakshi and V. Shilpa, 2012. CBMIR: Content based medical image retrieval system using texture and intensity for dental images. Proceedings of the International Conference on Eco-friendly Computing and Communication Systems, August 9-11, 2012, Springer, Berlin, Germany, pp: 125-134.

Reddy, K.R. and D.M. Narayana, 2016. A comparative studyof sift and PCA for content based image retrieval. Intl. Refereed J. Eng. Sci., 5: 12-19.

Velmurugan, K., 2014. A survey of content-based image retrieval systems using scale-invariant feature transform (sift). Intl. J. Adv. Res. Comput. Sci. Software Eng., 4: 604-608.

Veni, S. and K.A. Narayanankutty, 2010. Image enhancement of medical images using gabor filter bank on hexagonal sampled grids. World Acad. Sci. Eng. Technol., 41: 816-821.

Zhao, Z., Q. Tian, H. Sun, X. Jin and J. Guo, 2016. Content based image retrieval scheme using color, texture and shape features. Intl. J. Signal Process. Image Pattern Recognit., 9: 203-212.