International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Integration of Statistical Based Texture and Color Feature for Medical Image Retrieval

A.Saravanan, S.Sathiamoorthy

Abstract*: Today, the common problem in the doma n of computer v sion and pattern recogniti n is content bas d image retrieval (CBIR). In this paper, a novel image retrieval method using the geometric details based on the correlation among edgels and correlation between pixels has been introduced. The autocorrelation based choridiogram descriptor has been extracted from the image to obtain geometric, texture and spatial information. Color autocorrelogram has been computed to obtain color, texture and spatial information. The proposed method is tested on benchmark heterogeneous medical image database and LIDC-IDRI-CT and VIA/I-ELCAP-CT databases and results are compared with typical CBIR syste for medical image retrieval.

Keywords*: Autocorrelation based chordiogram image descriptor, color autocorrelogram, correlation, Manhattan measure.

I. INTRODUCTION

The development in the image capturing and storage technology is the root cause for the growth of image databases. The intention of image retrieval system is retrieving desired images from huge image collection and is one among the hot research subject and is useful in different domains like digital libraries, disease diagnosis, agriculture, engineering and many more. Particularly, CBIR for medical images is more supportive for physicians owing to extraordinary increase in the usage of medical images for disease diagnosis. Today, most of the hospitals used DICOM format to store medical images. The DICOM attributes are described using text and is more expensive and error prone [1]. The limitations with text-based approach are addressed by CBIR in which visual cues are described conveniently by linking the human visual perception and the structure, texture, shape and color in the image. Therefore, image representation is the key component of any CBIR system. Several research interests related to image representation have been suggested in the past.

Image retrieval system for CT and MRI brain images is reported by Chu et al., [2]. CBIR for PET images of brain is suggested by Cai et al., [3]. Lehmann et al. reported image Retrieval for Medical Applications i.e., IRMA [4]. In IRMA, local level features and spatial relationship among local level features are formalized using six semantic layers. Benign and malignant lesions in breast imaging is significantly differentiated by Gibbs and Turnbull [5]. Co-occurrence matrix is used by Felipe et al. [6] for the retrieval of CT and MRI images. Bag of words computed using SIFT descriptor is utilized for the classification of liver lesions in CT images [7]. CT images of lung is analyzed using textures computed through Riesz wavelets [8].Wavelet transformation based CBMIR is reported by Traina et al. [9]. Mixture of texture features and intensity is incorporated for dental images by Ramamurthy et al.[10]. For tissue identification, co-occurrence based gradient texture is presented by Felipe et al. [11]. Texture features based on uniformity of structure and brightness is computed by Peng et al. [12] for CT chest images.

Ridgelet and curvelet based texture analysis is reported in [13] for CBMIR. For mammogram images [14] used texture features. LBP is used for mammogram images in [15] and to classify cell phenotype images in [16]. Rahman et al. reported number of CBMIR systems [17-20] for mixed medical image collections. Murala et al., [21] reported directional binary wavelet pattern for CBMIR. Murala and Wu introduced co-occurrence of similar ternary edges for CT and MRI image retrieval [22]. Local diagonal extrema pattern is reported in [23] for CT images. Subsequently Dubey et al., [24] introduced local wavelet pattern for CT image retrieval. Seetharaman and Sathiamoorthy [25] introduced CBMIR for mixed medical image collections using full range Autoregressive model with Bayesian approach. Recently, Saravanan and Sathiamoorthy [26] introduced CBMIR using the combination of Chrordiogram image descriptor (CID), color and texture autocorrelaograms.

Even many CBMIR systems has been suggested by various researchers in the past, till the accuracy of existing CBMIR need to be concentrated much more for getting better disease diagnosis. Therefore, in this paper, we proposed the combination of autocorrelation-based CID and color autocorrelogram for medical image retrieval. The proposed system is tested on benchmark datasets and compared with existing typical CBMIR systems. The results proven that proposed CBMIR is superior in performance than the existing CBMIR systems [25, 26]. The rest of the manuscript is designed as follows: Proposed feature extraction strategy is explained in Section 2, Section 3 deals with Experimental results and the methodology used for measuring the similarity and performance. Section 4 concludes the manuscript finally.

Manuscript published on 30 September 2019
* Correspondence Author
A.Saravanan,* Department of Computer and Information Science, Annamalai University, Annamalai Nagar, India, sarancseau@gmail.com.
S.Sathiamoorthy*, Tamil Virtual Academy, Chennai, India, ks_sathia@yahoo.com.

II. PROPOSED FEATURE EXTRACTION METHOD

A. Autocorrelation based CID

We introduced ACID in [27], which computes geometric details of edgels based on matching among center and its neighborhood edgel values in a 3×3 sub image and thus ACID exploits correlation as well as geometric details of edgels which includes orientations of center and its identical neighbor edgel, the degree of angle between the horizontal axis and the line segment connecting the center and its identical neighbor edgel. The aforesaid procedure is repeated for all the 3×3 sub images of image I and for each quantized edgel value. The computation of ACID is explained as in [27]

Let I be an n×n image, an edgels in I are $e_1, e_2, \ldots, e_m$. For an edgel $E = (x, y) \in I$, let $I(E)$ represent its edge value. Let $I_e \triangleq \{E | I(E) = e\}$. Hence, the notation $E \in I_e$ is synonymous with $E \in I, I(E) = e$. For convenience, [27] use the L_∞ norm to assess the distance among edgels that is for edgels $E_1 = (x_1, y_1), E_2 = (x_2, y_2)$, [27] define $|E_1 - E_2| \triangleq max\{|x_1 - x_2|, |y_1 - y_2|\}$. [27] denote the set $\{1,2,\ldots,n\}$ by $|n|$.

Definitions. The histogram h of I is expressed for $i \in |m|$ by

$$h_{e_i}(I) \triangleq n^2 \Pr_{E\in I} \left|E \in I_{e_i}\right| \quad (1)$$

For any edgel in image $h_{e_i} {(I)}/{n^2}$ gives the probability that value of edgel is e_i.

Let a distance $d \in |n|$ be fixed a priori. Then, correlation of I is expressed for $i, j \in |m|, k \in |d|$ as

$$\gamma^{(k)}_{e_i,e_j} \triangleq \Pr_{E_1\in I_{e_i}, E_2 \in I} [E_2 \in I_{e_j} \; ||E_1 - E_2| = k] \quad (2)$$

Given any edgel of value e_i in the image, $\gamma^{(k)}_{e_i,e_j}(I)$ gives the probability that an edgel at distance k away from given edgel is of value e_j. The autocorrelation of I exploits spatial correlation among identical edgels only, and defined as [27]

$$\alpha^{(k)}_c \triangleq \gamma^{(k)}_{e,e}(I) \quad (3)$$

For example, computation of geometric details based on autocorrelation is illustrated in Fig.1. In Fig.1, geometric detail between the center edgel and identical neighbor edgel is computed.

B. Color autocorrelogram

Color is a predominant low-level feature because images or its parts can be easily and immediately recognisable. Among the many color features reported in literature, CA is outperforming by exploiting texture and spatial details more effectively and thus it is most effectively and frequently used for CBIR. The CA exploits color details based on correlation between center and neighbourhood pixels based on matching in color values in a 3×3 sub image and this procedure repeated for all the 3×3 sub images and for each quantized color value. Therefore, CA is computed as described in [25, 26].

Let I is an image of size N X N. Assume that colors in I is quantized into m color say $c_1, c_2, \ldots, c_m$. Let us consider the pixels p1 and p2 where p1=(x1, y1) and p2=(x2, y2). Let a distance between p1 and p2 is d and $d \in N$. The correlogram for image I is defined in equation(2) as in [25, 26].

$$\varphi^{d}_{c_i c_j}(I) \triangleq \Pr_{p_1\in i_{c_i}, p_2 \in i_{c_j}} [p_2 \in I_{c_i} \, | \, |p_1 - p_2| = k] \quad (4)$$

Where $k \in d$ and $i, j \in m$. For any pixel of color c_i in the image, $\varphi^{d}_{c_i c_j}(I)$ gives the probability that a pixel at distance k away from the given pixel is of color. c_j. The autocorrelogram of I captures spatial correlation between identical colors only and is defined in equation (3) as in [25, 26]

$$\Phi^{k}_{c}(I) \triangleq \varphi^{d}_{c_i c_j}(I) \quad (5)$$

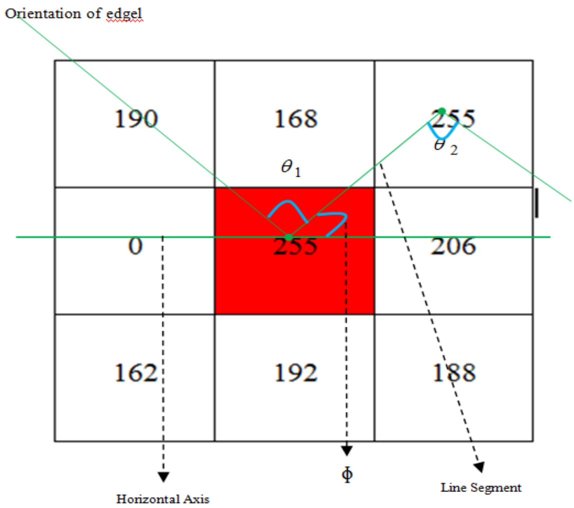


Fig.1. Computation of orientation of edgels with value 255 and angle between the horizontal axis and line segment of edgels 255 and 255 at distance 1

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Datasets

We used benchmark heterogeneous image database of [25, 26]. The database has 83 classes of images. The images of various body parts and modalities like Microscopy, CT, Mammogram, X-ray, ultrasound, MRI, Endoscopy are available in image database with ground truth information. We have shown some sample images of benchmark databases in Fig.2.The LIDC-IDRI [28] have 84 lung CT images cases in DICOM format and each case contains 100 to 400 images. The VIA/I-ELCAP-CT [29] consists of lung CT images of size 512 x 512 in DICOM format with the annotations of physicians in XML file.

B. Similarity and evaluation measurement

The aim of similarity metric is computing the distance between query and DB images. In the proposed CBMIR, we used Manhattan distance to measure the distance among two features in Euclidean space. The Manhatan measure is given as[30].

$$S(I^q, I^t) = \sum_{i=1}^{k} (I_i^q - I_i^t) \quad (6)$$

where, I_i^q and I_i^t - query and DB image feature vectors respectively and k-dimension of proposed feature vector.

Retrieval Number: C5567098319/2019©BEIESP
DOI:10.35940/ijrte.C5567.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In the proposed CBMIR, each DB image is used as query and matched with remaining DB images. The evaluation criteria used in the proposed CBMIR are precision(þ)

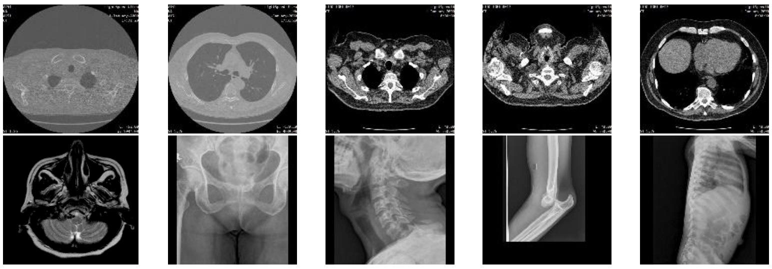


Fig-2.Example images from experimental databases

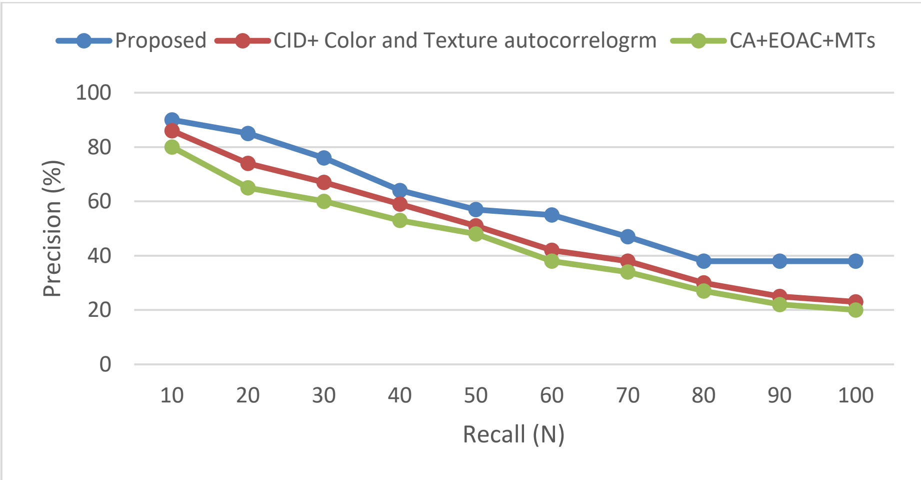


Fig.3.Precision(þ) Vss Recall(ɍ) for proposed andd existing aapproaches for LIDC-IDRI-CT

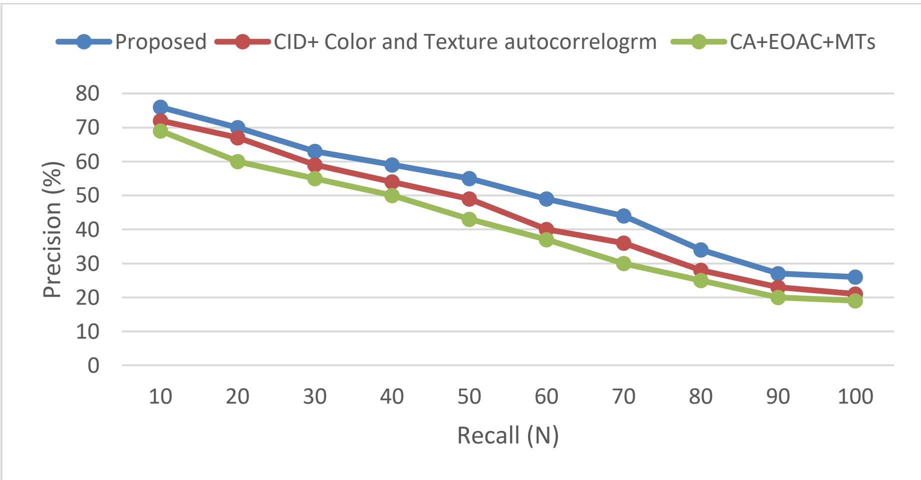


Fig.4.Precision(þ) Vss Recall(ɍ) for proposed andd existing aapproaches for VIA/I-ELCAP-CT

Table 1. Accuracy (%) of proposed and existing approaches

Datasets	Proposed	CA+CID+TA	CA+EOAC+MTs
LIDC-IDRI-CT	67.83	66.98	65.34
VIA/I-ELCAP-CT	65.78	65.02	65.14
Heterogeneous image collection	61.74	61.09	60.67

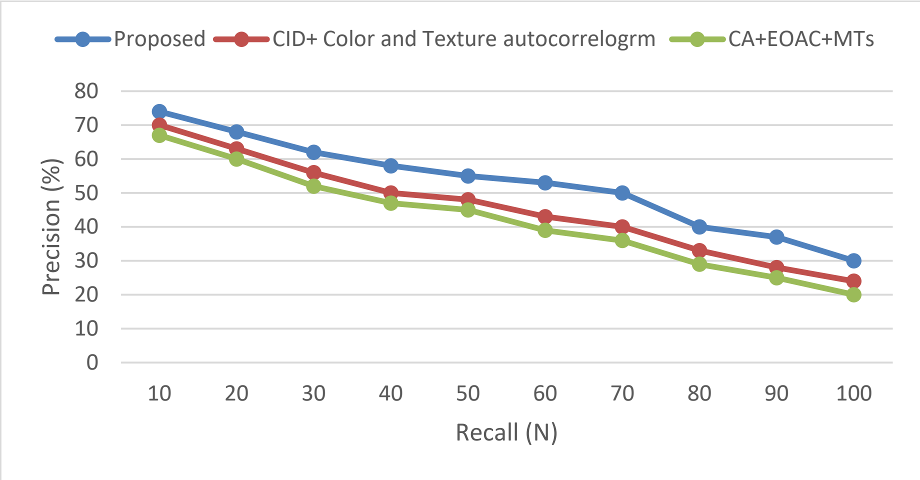
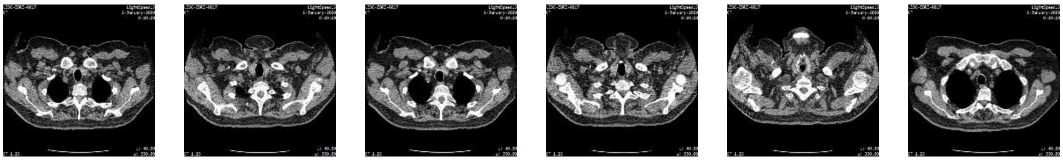


Fig.5.Precision(þ) Vss Recall(ɍ) for proposed andd existing aapproaches for heterogeneouss image collection



(a) (b) (c) (d) (e) (f)

Fig.6 (a). Example input image;(b)-(f).Top 5 retrieval rresults using pproposed approach

and recall(ɍ). We computed the mean average þ and mean average ɍ from the average þ and ɍ respectively. The precision(þ) and recall(ɍ)is given as in [25-27]

$$þ = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of images retrieved}} \quad (7)$$

$$ɍ = \frac{\text{No. of releveant images retrieved}}{\text{Total No. of releveant images in database}} \quad (8)$$

C. Results and discussion

The performance of proposed CBMIR system is analysed in comprehensive manner. For each image in benchmark database, we calculated the ACID and CA. We compared the proposed CBMIR with the approach of Saravanan et al., [26] in which color autocorrelogram (CA), texture autocorrelogram (TA) and CID is utilized for heterogeneous medical image retrieval. The proposed approach is also compared with the approach of [25] in which CA, micro-textures (MT) and edge orienatation autocorrelogram (EOAC) is used for CBMIR for heterogeneous image collection. In the experiment, we randomly chosen the query image. For each query image, the proposed CBMIR selects N number of images having shortest distance values from the DB. The plots for precision(þ) and recall(ɍ) for the proposed and existing approaches [25, 26] is shown in Fig.3, Fig.4 and Fig.5 respectively. The precision(þ) and recall(ɍ) graphs obviously illustrated that the proposed CBMIR using ACID and CA performs better than the existing approaches. The Table.1 reports the accuracy of proposed CBMIR and existing CBMIR approaches. From Table. 1, it is clearly seen that accuracy of proposed approach is significantly outperforms the existing CBMIR systems.

IV. CONCLUSION

In this paper, a new combination of feature vectors namely ACID and CA is proposed for medical image retrieval. The ACID exploits texture, hape and geometric information from the edge identified image. The CA computes texture and color information. The robustness and efficiency of proposed CBMIR is tested on benchmark medical mage databases. the detailed examination of the results depict a considerable improvement f proposed CBMIR in erms of accuracy as compared to CA+EOAC+MTs and CA+CID+TA on LIDC-IDRI-CT, VIA/I-ELCAP-CT and heterogeneous image collections. In future, the proposed CBMIR can be combined with some another image representation techniques and machine earning algorithms for accuracy enhancement.

REFERENCES

1. Mustra, M.,Delac, K., Grgic, M.,2008. Overview of the DICOM standard. 50th Int Symp. ELMAR 1, 10–12.
2. W.W.Chu,C.C.Hsu,A.F.Cardenas,R.K.Tairaa,Knowledgebasedimager etrieval with spatialand temporal constructs,IEEE Trans. Knowl.Data Eng.10 (December (6))(1998)872–888.
3. W.Cai, D.Feng,R.Fulton, Content-basedretrieval of dynamicPET functional images, IEEE Trans.Info.Tech.Biomed.(2000) 152–158.
4. Lehmann, T.Mm.,Gu¨ld, M.O.,Deselaerss, T., Keysers,D.,Schubertt, H., Spitzer, K., Neyy,H., Wein, B.B.,2005. Automatic categorizationof medicalimages for content-basedretrieval and data mining, Computerized MedicalImaging and Graphics. 29:143–155.
5. P.Gibbs,L.W.Turnbull, Textural analysis ofcontrast-enhanced MR imagesof thebreast, Magn. Reson. Med. 50(1)(2003)92–98.
6. C.Felipe, A.J.M.Traina,C.J.Traina, Retrieval bycontent ofmedical images using texturefor tissue identification, in: Proceedingsof the16th IEEE Symposiumon Comp.-BasedMed. Systems,New York,USA, 2003, pp.175–180.

Retrieval Number: C5567098319/2019©BEIESP
DOI:10.35940/ijrte.C5567.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

7. W.Yang,Z.Lu, Yu.M,M.Huang, Q.Feng,W.Chen, Content-basedretrieval of focalliver lesionsusing Bag-off-Visual-Wordsrepresentations of single-andmultiphase contrast enhanced CTimages, J.Dig.Img, 25(2011)708–719.
8. Depeursinge Adrien,Foncubierta Rodriguez Antonio,Van dee Ville Dimitri,Muller Henning, Multiscale lung texturesignature learning usingthe Riesz transform,Medi.Image Comp.-Assist.Interv.-MICCAI(2004)517–524.
9. Traina, C.Castanon,C.J.Traina, Multiwavemed: a systemfor medical imageretrieval throughwavelets transformations,in: Proceedingsofthe 16th IEEESymposium onComp.-Based Med. Sys,New York,USA, 2003, pp.150–155.
10. Ramamurthy,K.R.Chandran,V.R.Meenakshi,V.Shilpa, CBMIR: content based medicalimage retrievalsystem usingtexture and intensityfor dental images, Com.Comp.Info.Sci.305(2012)125–134.
11. J. C. Felipe,A. J. M. Traina,and C. Traina Jr., “Retrieval bycontent of medicalimages usingtexture fortissue identification,” in Proc.16th IEEE Symp.Comp.-Based Med.Sys., 2003, pp.175-180.
12. S. Peng, D. Kim,S. Lee,M. Lim, “Texture featureextraction on uniformityestimation for local brightnessand structure in chestCT images,” J. Comp.Biol.Medi.,vol.40, pp.931-942, 2010.
13. Dettori Lucia,Lindsay Semler, A ccomparison of wavelet, ridgelet,and curvelet- based texture classification algorithms in computed tomography, Computatioanl. Biol. Medi.37(4)(2007) 486–498.
14. K.Vaidehi, T.S.Subashini, An intelligent content based image retrieval system for mammogram image analysis, Journal. Engg. Scie. Technol. 10 (11) (2015) 1453–1464.
15. Oliver, X.Lladó, J.Freixenet, J.Martí, False positive reduction in mammographic mass detection using local binarypatterns, in: Proceedings of the MedicalImage Computing and Computer Assisted Intervention (MICCAI 2007), Brisbane, Australia: Springer, Lecture Notes in Computer Science (LNCS) 4791, pp. 286–293, 2007.
16. L.Nanni, A.Lumini, A reliable method for cell phenotype image classification, Artif. Intell. Med. 43 (2) (2008) 87−97.
17. Mahmudur Rahman,Md., Bipin C.Desai, PrabirBhattacharya, 2008. Medical image retrieval withprobabilistic multi-class supportvector machine classifiers and adaptivesimilarity fusion. Computerized Medical Imaging and Graphics. 32, 95-108.
18. Mahmudur Rahman,Md., ParbirBhattacharya and Bipin C.Desai, 2009. A unified image retrieval frameworkon local visual andsemantic concept based feature space. J. Vis. Communication. Image R. 20(7), 450-462.
19. Mahmudur Rahman,Md., Sameer K.Antani, and George R.Thoma, 2011. A learning-based similarityfusion and filtering approachfor biomedical image retrievalusing svm classification and relevancefeedback. IEEE transact on information techn in biomedicine. 15(4), 640-646.
20. Mahmudur Rahman,Md., Daekeun You,Matthew S.Simpson, Dina Demner-Fushman, Sameer K.Antani, and George R.Thoma, 2013. Multimodal biomedicalimage retrieval using hierarchicalclassification and modality fusion. Int J Multimed Info Retr. 2, 159-173.
21. S.Murala, R.P.Maheshwari, R.Balasubramanian, Directional binary wavelet patterns for biomedicalimage indexing andretrieval, Journ. Med. System. 36 (5) (2012) 2865–2879.
22. Subrahmanyam Murala, Q.M.Jonathan Wu, (2013). Local ternary co-occurrence patterns: A new featuredescriptor for MRIand CT image retrieval, Neurocomputing, 119: 399–412.
23. S.R.Dubey, S.K.Singh, R.K.Singh, 2015. Local diagonal extrema pattern: a new and efficientfeature descriptor for CTimage retrieval, IEEE Sig. Processi. Let. 22 (9) 1215–1219.
24. S.R.Dubey, S.K.Singh, R.K.Singh, 2015. Local wavelet pattern: a new feature descriptor forimage retrieval in medicalCT databases, IEEE Transa. Image Processing. 24 (12) 5892–5903.
25. K.Seetharaman, S.Sathiamoorthy, 2016, A Unified Learning Framework for Content Based Medical Image Retrieval Using a Statistical Model, J. of King Saud Univ - Computer and Info. Sciences, 28 (1): 110-124.
26. A.Saravanan, M.Natarajan, S.Sathiamoorthy, A Novel Hybrid Framework for Medical Image Retrieval, Asian Journal of Engineering and Applied Technology, Vol. 7 No. 2, 2018, pp.37-41.
27. A.Saravanan, S.Sathiamoorthy, Autocorrelation based Chordiogram Image Descriptor for ImageRetrieval, 4th International Conference on Communication and Electronics Systems (ICCES 2019), July 17-19, 2019, PPG Institute of Technology, Coimbatore.
28. ftp://medical.nema.org/medical/Dicom/Multiframe/
29. http://www.via.cornell.edu/-databases/lungdb.html.
30. Fazal Malik, BaharumBaharudin, (2013). Analysis of distance metrics in content-based imageretrieval using statistical quantizedhistogram texture features in the DCTdomain, Journal of King Saud Uni. - Computer andInformation Sciences, 25(2):207-218.

AUTHORS PROFILE

A.Saravanan received B.Sc degree in Computer Science as well as M.C.A degree from Bharathidasan University and M.Phil degree from Annamalai University, India. He joined the Department of Computer Science and Engineering, Annamalai University in the year 2001. At present he is with the Department of Computer and Information Science, Annamalai University. He has almost 19 years of teaching experience. He is currently pursuing Ph.D degree in the Department of Computer and Information Science, Annamalai University. His research interests include Image processing, Medical image analysis and Content-based image retrieval.

S.Sathiamoorthy received B.Sc degree in Physics from University of Madras, M.C.A degree from Bharathidasan University and M.Phil as well as Ph.D. from Annamalai University, India. In the year 2001, he joined in the Department of Computer Science and Engineering, Annamalai University and also served in the Computer Science and Engineering Wing of Directorate of Distance Education and Department of Computer and Information Science, Annamalai University. At present, he is Assistant Director (Controller of Examinations), Tamil Virtual Academy, Information Technology Department of Tamil Nadu Government, India. He has almost 19 years of teaching experience. Currently he is working on Pattern recognition, Pattern analysis, Image retrieval and classification. His research interests include Machine learning algorithms and Medical image analysis. He has published more than 25 papers in various International journals and more than 15 papers in national and international conferences. He has been on the reviewing and editorial board of many reputed journals.