European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 07, Issue 9, 2020

Count of Equivalent Pixel (CEP): A Novel Algorithm to Extract Most Relevant Images from the Medical Image Database

DNVSLS Indira[1], Kalyanapu Srinivas[2], Kavitha Chaduvula[3], J N V R Swarup Kumar[4]

[1,2]Assocaite Professor, [3]Professor and Head, [4]Assistant Professor

[1,3]Dept. of Information Technology, [2,4]Dept. of Computer Science and Engineering

Gudlavalleru Engineering College, Gudlavalleru, AP - 521356

[1]indiragamini@gmail.com, [2]kalyanapusrinivascse@gmail.com, [3]kavithachaduvula12@gmail.com, [4]swarupjnvr@gecgudlavalleru.ac.in

Abstract

Now a day images or pictures play a prominent role in the world. Content-based image retrieval (CBIR) systems are used to retrieve similar images based on query image. This system is capable of identifying similarities between query image and the set of images placed in the database. Even though it is an important research area from the last two decades, still there is scope for new technologies and algorithms to manipulate large amount of image databases in different fields like medical, social media and space images. Image contents are colors, texture and shape play significant role for image retrieval. All most all image retrieval systems are based on efficient feature extraction methods which are well-organized. This paper mainly concentrates to extract exact or most relevant images from the image database, here we compare query image with database image using pixels present in intensity vector. To accomplish this, the images are converted into gray scale. To achieve this, we are proposing a novel algorithm called as CEP (Count of Equivalent Pixel). This paper compares the existing sum of values of local histograms results with proposed system. The proposed system gives best results in retrieval of medical images in large database.

Keywords: CBIR, Local Histogram, Bins, Count of Equivalent Pixel, Medical Database

I. Introduction

Content-based Image Retrial has been a functioning exploration zone lately. The interest in this exploration zone has persuaded me to look and well oversee enormous volumes of Multimedia and Medical data. CBIR removes low-level highlights which are inbuilt in the pictures to introduce the substance of pictures. Each picture has Visual highlights, for example, ordered into three primary classes: shading, surface, and shape highlights [3]. Shading is a significant picture highlight, for example, utilized in Content-Based Image Retrieval [1]. These highlights can possibly distinguish protests and recover comparative pictures based on their substance. These strategies accomplish work proficient in item acknowledgment and Web looking and Medical Images recovery [13]. A proficient and viable inquiry reformulation is basic for finding the pertinent pictures from the information base.

In this paper we proposed new calculation called Count of Equivalent Pixel (CEP) to recover practically significant pictures from the picture information base. Here likewise the shading picture is changed over into dim scale picture. The dark scale picture is resized to 256x256 power framework. On the off chance that the inquiry picture is coordinated precisely with the information base picture, at that point the CEP is equivalent to 65536.

We contrasted the proposed calculation and existing one. Our calculation gives the best results. The exhibition of a test picture recovery framework is assessed on an information base of around 770 pictures. The test results show that CEP is powerful in depicting elevated level picture content and can give adaptable picture portrayal and effective picture recovery execution.

II. Earlier Research

II. 1 Literature Survey

Reference No.	Year	Features/ Methods Discussed
[1]	2010	Extraction Techniques using K-Means and Hierarchical Algorithms
[2]	2011	Features Extraction Techniques and Metric Data Structures
[3]	2011	CBMIR
[4]	2015	Classification Using Color Feature Extraction
[5]	2015	SOM, DWT, Feature Vector, Texture Vector
[6]	2015	Machine Learning Algorithms
[7]	2016	Gabor Wavelet, Color auto-Correlogram and Wavelet Transform
[8]	2017	CCM, Textual Features
[9]	2018	SIFT and ORB, LPP dimensionality reduction methods and K-Means clustering
[10]	2018	(CBIR) framework, Deep Belief Network (DBN)
[11]	2019	Segment IRF, Retinal Layers and The Fluid.
[12]	2019	Histogram of Oriented Gradient (HOG) Features
[13]	2019	Dimensional Auto encoders and Convolutional Neural Network methods
[14]	2019	Manifold Indexing Technique

Table 1 : Various features discussed in different articles.

II.2 Existing Sum of Values of Local Histograms to Retrieve Images

Color is one of best reliable feature in image retrieval. Calculation of Histograms on color images is the existing technology. Shading histogram is the most well-known procedure for separating the shading highlights of hued pictures. Color histogram tells the worldwide appropriation of tones in the pictures. It includes low calculation cost, and it is obtuse toward little varieties in the picture structure. They cannot completely oblige the spatial data, and they are not novel and vigorous. Two disparate pictures with comparative color appropriation produce fundamentally the same as histograms. Also, comparative pictures of same perspective conveying distinctive lighting conditions make divergent histograms.

This existing technique goes through various steps:

1. Pre-processing of data- Converting original images into histograms.

Figure 1: An image selected at random as a sample from a collection of images.

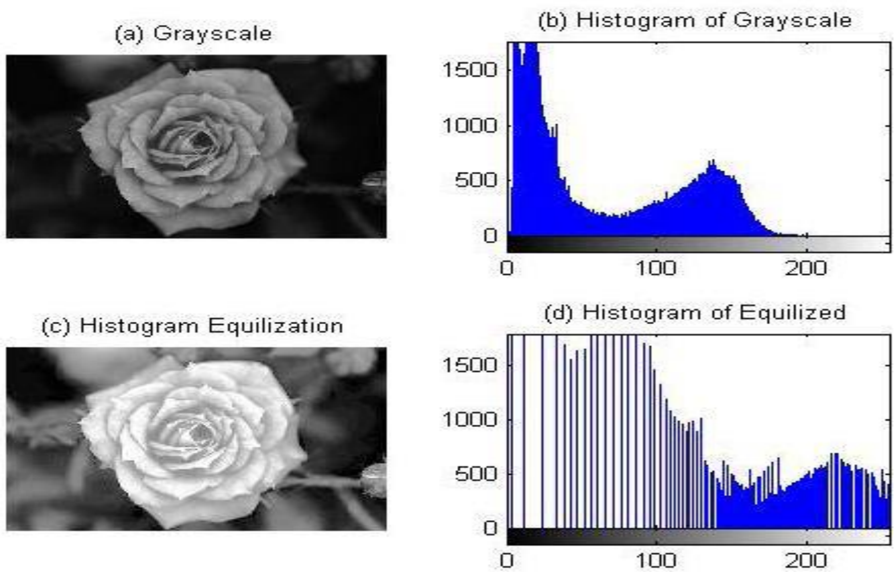

Figure 2: (a) Grayscale (intensity) representation of image in figure 1. (b) Histogram of the grayscale image. (c) Grayscale image after histogram equalization. (d) Histogram of the flat (equalized) image.

2. Splitting Histogram Values by Fixed Frequency Range

The histogram adjusted picture is part into four fixed containers to remove more particular data from it. The frequencies of 256 estimations of dark scale are part into sixteen (16) bins conveying 16 qualities each (0~15, 16~31, 32~ 47, 48~63, etc) [12].

3. Storing the obtained information in feature vector.

Bin#	1	2	3	4	5	6	7	….	16
Sum	122	334	22	1	334	5	67	….	231

Table 2: Sum of values listed against bins

III. Proposed Work

III. 1 CEP (Count of Equivalent Pixel)

The proposed system uses an affluent feature called pixel to make an analogy with dataset images irrespective of image contents like color, shape and texture which are generally used to retrieve the image from image database. In this algorithm we are using intensity matrices instead of histogram equivalent matrix and bins feature vector. In the proposed algorithm the query image is captured and converted into gray scale and the image is resized. It is stored in 256x256 intensity matrix. The database image is also converted into gray scale and resized into 256x256 intensity matrix (im).

Now the CEP algorithm compares the QI(i,j) with DI(i,j) if both are equal then count(n) gets incremented by one ,where (i,j) delineates the ith row jth column in intensity matrix and n represents number of images(i.e.,1..1000) in image database. In similar manner, all the pixels are compared, if all are equal, the count gets incremented automatically. If both images are exact, then the count value is 65,536. Then it gives exact image retrieval. We used Wang's [15] dataset comprising of 10000 Corel images. This data set comprises of 10 categories of each with 1000 images. The images are of the size 256 x 384 or 384X256. But the images with 384X256 are resized to 256X384.

III. 2. Algorithm

1. Read Query Image
2. Convert RGBtoGray(QI)
3. im=imageresize(QI,256x256)
4. for m is 1 to 1000
    a. read Databse Image
    b. convert RGBtoGray(DI)
    c. imdi=imageresize(DI,256x256)
    d. set temp=0;
    e. for i is 1 to 256
        i. for j is 1 to 256
            1. if(im(i,j)==imdi(i,j))
                a. temp=temp+1

               end

          end

    end

    f. set count(m)=temp
    g. sort count in descending order along with index
    h. pic top 20 images from the database using count array.
5. End

III.3. Comparison of Existing system with Proposed:

The proposed algorithm is very simple and efficient to retrieve an exact image from the database. It retrieves most relevant pictures from the database when compare with existing one.

	1	**2**	**3**	**4**	**5**	**6**	**----**	**256**
1	120	40	12	65	12	45		125
2	1	90	11	1	13	95		254
3	5	60	10	6	9	67		1
4	110	11	140	9	3	46		6
5	6	9	98	125	130	123		87

256	23	35	45	12	11	5		10

Table 3: Query Image Intensity Matrix

	1	**2**	**3**	**4**	**5**	**6**	**----**	**256**
1	11	90	13	1	254	95		1
2	45	35	11	12	10	5		23
3	10	60	9	6	1	67		5
4	12	40	12	65	125	45		120
5	140	11	3	9	6	46		110

256	98	9	130	125	87	123		6

Table 4: Data base Image Intensity Matrix

Bin#	1	2	3	4	5	6	7	….	16
Sum	21	1	5	1	2	3	2	….	1

Table 5: using the information shown in the intensity tables when we calculated bins for both QI and DI

If the Database Image also has the feature vector like query Image then the difference of those two vectors is zero, the image is retrieved as top image in existing system. But the positions are different. If we use the existing algorithm the image is retrieved as the topmost retrieval image. But in the proposed algorithm the count of exact pixel position in both images is zero so it cannot retrieve the image. So, our algorithm is the best one to retrieve topmost exact or relevant image retrieval.

IV Experimental Results

Figure 3 - Query Image

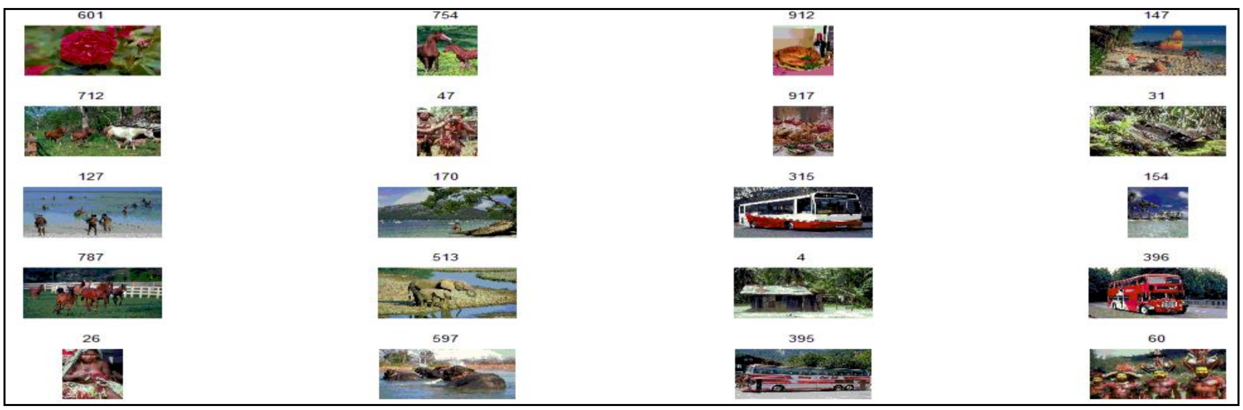

Figure 4 - Images retrieved by existing system - Top 20

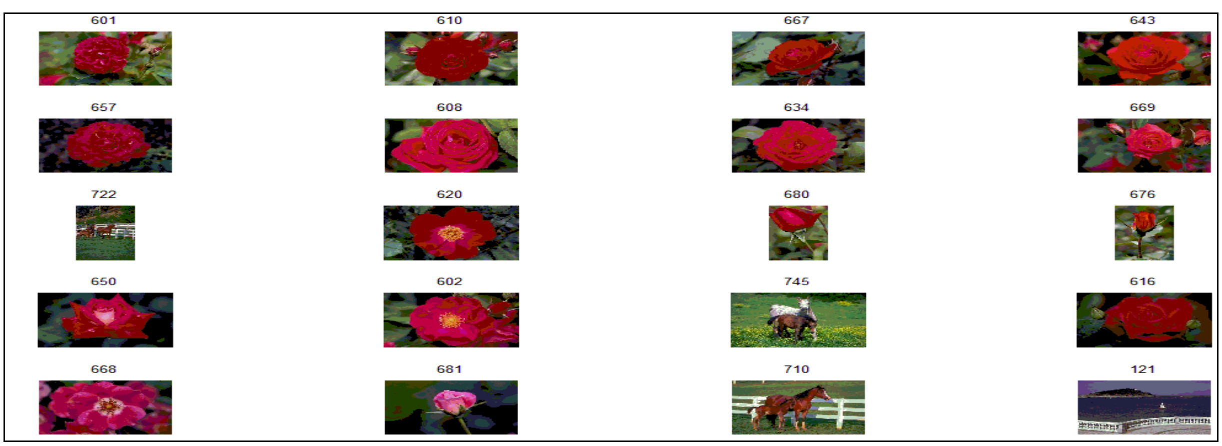

Figure 5 - Images retrieved by CEP - Top 20

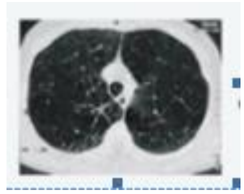

Figure 6: Medical Query Image

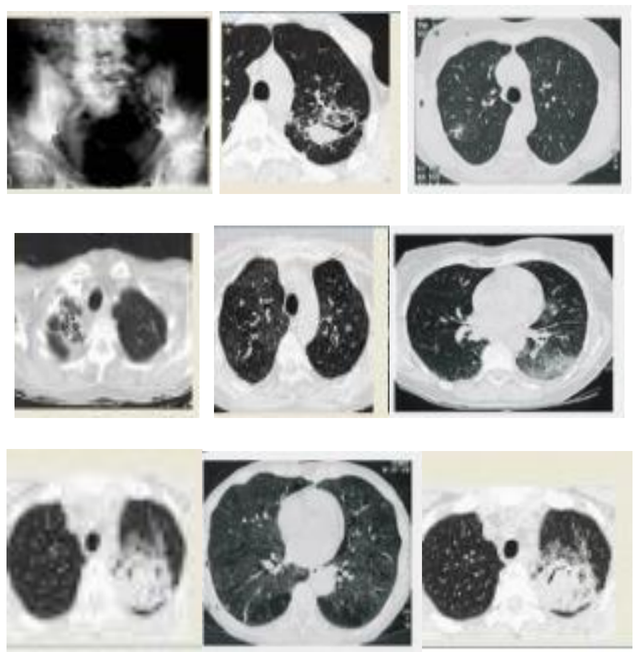

Figure 7 – Relevant Images retrieved from the medical dataset by CEP

In our paper the database consisting of 7770 images. These images are categorized into ten groups of images, each group consisting of hundred relevant images. We need to evaluate the performance of the Existing system with Proposed Algorithm in terms of precision and recall. The performance of any image processing applications can be measured in terms of precision and recall. Here in my work also we used those two measuring parameters to calculate the difference between existing algorithm with the proposed algorithm.

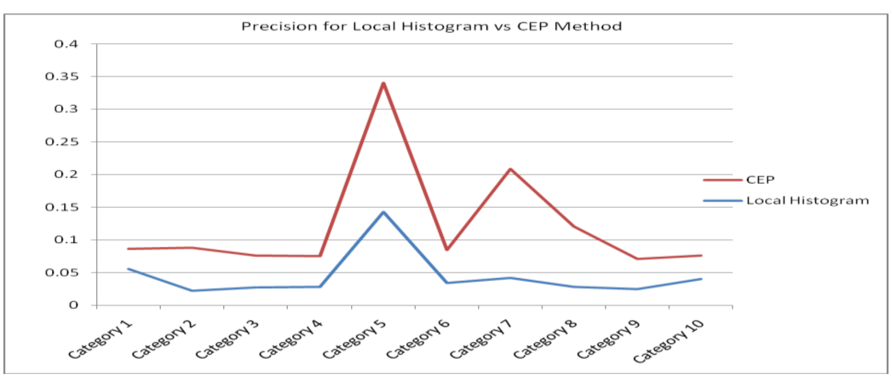

Figure 8: Precision for Local Histogram vs CEP Method

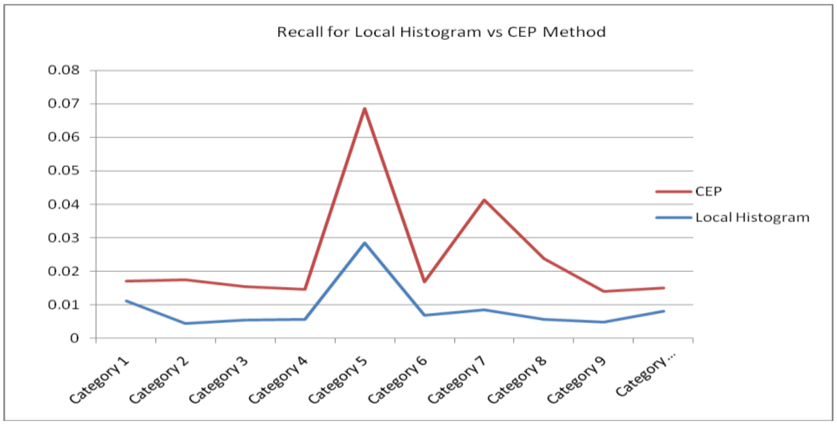

Figure 9: Recall for Local Histogram vs CEP Method

CBIR utilizes histogram-based techniques for picture ordering. Histograms depict the worldwide force dispersion of pictures. They are exceptionally simple to figure and are harsh toward little changes in item interpretations and turns. Be that as it may, they are not strong to huge appearance changes, and they may give comparable outcomes for various types of pictures if the conveyances of tones are same in the pictures. However, the disadvantage of histogram-based recovery is, it recovers pictures those are superfluous with question picture even though the component vector distinction is zero. Our examination centers essentially around the picture pixel correlation whereas the neighborhood histogram strategy utilizes containers (histogram esteem divisions by recurrence) partition method followed by ascertaining the amount of qualities and utilizing them as picture nearby highlights. From the start, the histogram is determined for a picture, later it is partitioned into sixteen equivalent receptacles and the amount of neighborhood esteems is processed and put away.

V. Conclusion and Future Work

In this work we explored different avenues regarding the thoughts of Sum of Values of Local Histogram based Image Retrieval. We joined the Histogram Intersection measure technique to contrast the inquiry picture and information base pictures. A proportion of the general similitude between pictures, characterized by our methodology, consolidates all neighborhood properties of the surface histograms of the pictures. We demonstrated that our methodology is appropriate to recover most ideal outcomes. Examinations demonstrate that district explicit histogram properties can be extremely helpful, in light of the fact that, they add power to the histograms that, thus, add uniqueness of portrayal among a bunch of comparative pictures. We have tried the proposed calculation on a huge information base of pictures. This work is applied and tested on 7770 medical images collected from various hospitals which are categorized into MRI Scan images, high-resolution computed tomography (HRCT) lungs pictures of different part of the body from medical database. This work can be extended with the machine learning algorithms with different features.

VI. References

[1] J.N.V.R.Swarup Kumar et al. “Content Based Image Retrieval using Hierarchical and K-Means Algorithms”, International Journal of Engineering Science and Technology, Vol.2 (3), 2010, ISSN: 0975-5462.

[2] Herbert Chuctaya, Christian Portugal, César Beltrán, Juan Gutiérrez, Cristian López, Yván Túpac, “M-CBIR: A Medical Content-Based Image Retrieval System Using Metric Data-Structures”, 30th International Conference of the Chilean Computer Science Society, 2011.

[3] Abdol Hamid Pilevar, “CBMIR: Content-based Image Retrieval Algorithm for Medical Image Databases”, J Med Signals Sens. 2011 Jan-Apr; 1(1): 12–18.

[4] Sandhya R. Shinde, Experiments on Content Based Image Classification using Color Feature Extraction, IEEE, 2015.

[5] Ammar Huneiti, "Content Based Image Retrieval Using SOM and DWT", Journal of Software Engineering and Applications, 2015.

[6] S. Tunga, D. Jayadevappa and C. Gururaj, "A Comparative Study Of Content Based Image Retrieval Trends And Approaches", International Journal of Image Processing (IJIP), vol. 9, no. 3, pp. 127-155, 2015.

[7] A. Anandh, K. Mala and S. Suganya, "Content based image retrieval system based on semantic information using color texture and shape features", Computing Technologies and Intelligent Data Engineering (ICCTIDE) International Conference on, 2016.

[8] Parul Shrivas, Dr Umesh Kumar, Nitin Agrawal, “Genetic approach based image retrieval by using CCM and textual features”, 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), September 2017

[9] M. Kumar, P. Chhabra and N.K. Garg, "An efficient content based image retrieval system using BayesNet and K-NN",Multimedia Tools and Applications, pp. 1-14, 2018.

[10] B. Salafian, R. Kafieh, A. Rashno, M. Pourazizi and S. Sadri, Automatic segmentation of choroid layer in edi oct images using graph theory in neutrosophic space, 2018.

[11] G. N. Girish, B. Thakur, S. R. Chowdhury, A. R. Kothari and J. Rajan, "Segmentation of intra-retinal cysts from optical coherence tomography images using a fully convolutional neural network model",IEEE J. Biomed. Health Inform., vol. 23, no. 1, pp. 296-304, Jan. 2019.

[12] P. Das and A. Neelima, "Content-Based Medical Visual Information Retrieval" ,Studies in Computational Intelligence Hybrid Machine Intelligence for Medical Image Analysis, pp. 1-19, 2019.

[13] K. Kruthika, Rajeswari and H. Maheshappa, "CBIR system using Capsule Networks and 3D CNN for Alzheimers disease diagnosis",Informatics in Medicine Unlocked, vol. 14, pp. 59-68, 2019.

[14] M. S. Mirasadi and A. H. Foruzan, "Content-based medical image retrieval of CT images of liver lesions using manifold learning", International Journal of Multimedia Information Retrieval, vol. 8, no. 4, 2019.