Volume-3, Issue-11, Nov Special Issue -2014 • ISSN No 2277 - 8160

Research Paper Systems

A Refined Euclidean Distance Matching Technique for Improved Cbir Using Colour, Straight Line and Outline Sketch Features of the Image

L.JAYANTHI Assistant Professor, Periyar Maniammai University

Dr. K.LAKSHMI Dean, School of Computing Science and Engineering, Periyar Maniammai University

ABSTRACT

The vast growth of internet and digital contents of internet creates much interest on automated indexing and retrieval techniques. In this work, we address an improved Content Based Image Retrieval system (CBIR) using combined colour histograms for three regions (top, bottom and middle), straight line features and outline features of the image. Average precision is used as a metric to measure the performance of CBIR system . We used the SIMPLIcity database (1000 images of Semantics Sensitive Integrated Matching for Picture Libraries) using MAT Lab to evaluate the performance of the proposed improved CBIR system (ICSLOS) and compared the performance of our method with four other previous methods of CBIR using Integrated Region Matching(IRM), Fuzzy Integrated Region Matching (F-IRM), CBIR using colour salient points and shape features (CSP) and Interactive CBIR using Ripplet Transform and Fuzzy Relevance Feedback Mechanism (RT-FRFM) . In our previous work CSLOS, Euclidean distance method is used for comparing the similarity between the query image and the images in the database. Distance matrix is sorted and indexed. We have improved this by finding average signature of top N matching images and used that as the signatures of virtual query and repeated the search by finding euclidean distance between the virtual query and the database images. According to the arrived results, the performance of the proposed refined Euclidean distance matching technique for ICSLOS provided improvement in performance.

KEYWORDS : CBIR, Matching Technique, ICSLOS, Distance Matrix.

INTRODUCTION

Large numbers of images are created in various domains such as education, entertainment and commercial application, remote sensing, agriculture, crime prevention, publishing, medicine, architecture etc. So, it is necessary to develop a system to manage huge number of images in the image data base. Content Based Image Retrieval system is an automatic retrieval system which is used for browsing, searching and retrieving images from large image database as required by users by comparing the features of all the images in the data base such as colour , shape, texture with that of the query image.

Colour histogram is one of the most important techniques in Content Based Image Retrieval. The colour histogram technique is efficient to compute and effective in searching results. The weakness of this method is that the space information will be missing. It means, the colour histogram will not say where that colour was originally present in that image. It will not fully explain the nature of colour distribution in the image at various places which actually makes that image.[1]

Images are represented by set of regions based on objects. The objects are characterized by features such as colour, texture, shape and location. The features of query image is compared with all the regions of the images in the database. The similarity measure is computed as a weighted sum of similarity between the region pairs and the matched images are retrieved by CBIR using Integrated Region Matching (IRM) [2] This reduces the adverse effect of inaccurate segmentation. This system achieves more accurate retrieval at higher speed. In CBIR using Fuzzy Integrated Region Matching, image is represented by a set of segmented regions. Each region is characterized by a fuzzy set of features reflecting colour ,texture and shape properties. The similarity between the features of query image and the images in the database are measured by comparing the fuzzy set of features. This method greatly reduces the influence of inaccurate segmentation. [3]

In CBIR using colour boosted salient points and shape features, salient points are locations in an image where there is a significant variation with respect to a chosen image feature. These salient points represent objects in the image. They focus on shape. The colour and texture information around these points of interest serve as the local descriptors of the image. In addition, edge of the images computed using gradient vector flow fields. Invariant moments are then used to record shape features. [4].

Manish Chowdhury in [5] has analyzed the performance of interactive CBIR using ripplet transform and fuzzy relevance feedback (F-RFM). Multi scale Geometric Analysis Tool called ripplet transform type I is used to represent the images in the database in different scales and directions . The images in the database are transformed into YCbCr colour space. RT decomposition over the intensity plane characterizes the texture information. RT decomposition over chromaticity planes characterize colour. Retrieval result is obtained by measuring the Euclidean distance between the RT based features of the query image and the images stored in the database. If the user is satisfied with the retrieved results without using relevance feedback mechanism, the retrieval process is ended in the first pass. If the user is not satisfied, user can select top query based relevant based images as positive feedback and the remaining as negative examples. Fuzzy features of positive examples are compared with the query image. The performance of CBIR systems is improved through continuous learning and interaction with the end users. But in all these cases time complexity per iteration are high and the accuracy of the relevant images are low.

The rest of this article is organized as follows. Section II describes the algorithm of CBIR system using combined three region colour histogram and straight line and outline signatures of the image. Section III details three region colour histogram and Euclidean distance method of finding the images similar to the query image. The combination of those histograms will indirectly include the space information within it. Section IV addresses the extraction of straight line signatures to improve performance in retrieving a particular category of image and in section V CBIR using outline sketch signatures is explained. Section VI addresses the improved CBIR by finding the Euclidean distance between the query image and the images of the database for the second time using virtual query. We demonstrated the result of average precision of the proposed system and analyzed with the average precision of other system in section VI.

Precision of our systems are compared with the precision of previous existing CBIR system and the performance of our system is good. Recall, False positive rates are calculated and time taken for the retrieval of images is determined and analysed and shown as tables in section VII.3

THE DESIGN OF CBIR SYSTEM

Improved 3 Region Colour, Straight-Line and Outline Sketch based CBIR Algorithm (ICSLOS)

- Let I be the input query image of size m x n.
- Apply median filter on the input image
- Convert the high colour input image in to a 256 colour image
- Separate the top, middle and bottom regions R1, R2, R3, of the image
- Find the histogram of the regions R1, R2, and R3 as
- Use that combined 3 region colour histograms as the signature S1 of input image

$$S1 = h_i(r_1, r_2, r_3)$$

- Find all the long straight lines in the the image above the threshold and use it as Staight line signature S2 of the image.
- Detect the outlines of the input image.
- Find the FFT of the one dimentional representation of the binary outline Image . This will give the outline signature S3 of the image
- Form the combined signature S ={ S1, S2, S3} of input image
- Let M be the image signatures of all the images in the dataset.
- Find Euclidean Distance between the input image signature and the signature of the all the images of the database
- The distance D1=Euclidean distance (M,S)
- Sort the distance matrix D in ascending order
- Find the index of Top N ranked minimum distances of D
- Find the average image signature of the top N matching images
- Use that average image signature V as a virtual query image and repeat the search.
- Again, find Euclidean Distance between the virtual input image signature V and the signature of the all the images of the database
- The distance D2= Euclidean Distance(M,V)
- Find the index of Top N ranked minimum distances of D2
- Display the Top N ranked Images from the image database using the index.

The following diagram shows the Indexing Phase of the CBIR system. Figure 1 represents the indexing phase and figure 2 represents the CBIR system.

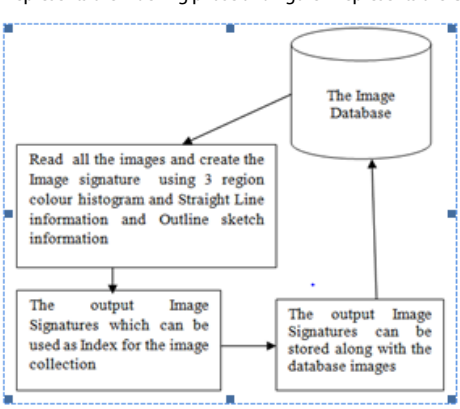


Figure.1 Indexing Phase

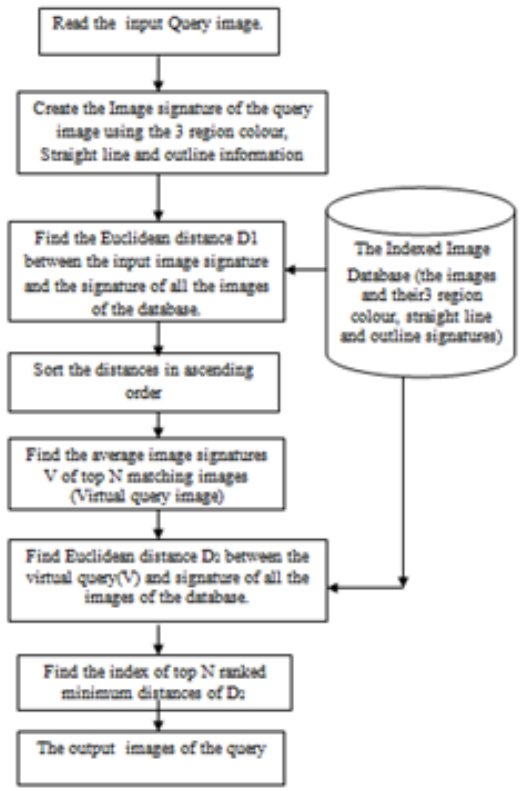


Figure. 2 CBIR system

CBIR USING COMBINED THREE REGION COLOUR HISTOGRAM SIGNATURES [3RCS]

Figure 3 represents the colour signature from 3 regions of the image. For an m*n image I, the colours in that image are quantized to C_1, C_2, ……, C_k. The colour histogram H(I)={h_1, h_2, ……, h_k}, where hi represents the number of pixels in colour C_i.

C_i is the i^{th} colour index, h_i is the number of pixels with that colour index. m x n is the total number of pixels in the image.

i = 0,1,2,….,L-1. (where L = 256).

Let the image is decomposed in to three regions R_1, R_2, R_3.

I_1={ R_1, R_2, R_3 }

Let h_1 and h_2 represent two colour histogram of three regions R_1, R_2, R_3, of two images I_1 and I_2.

Now, we can represent the combined histograms of the image I_1 as

$$h_1 = h_1(r_1, r_2, r_3)$$
$$h_2 = h_2(r_1, r_2, r_3)$$

The Histogram Euclidean Distance between the colour histograms h1 and h2 can be computed as:

Combination of histograms for different regions of the image will indirectly include the space information within it. [6]

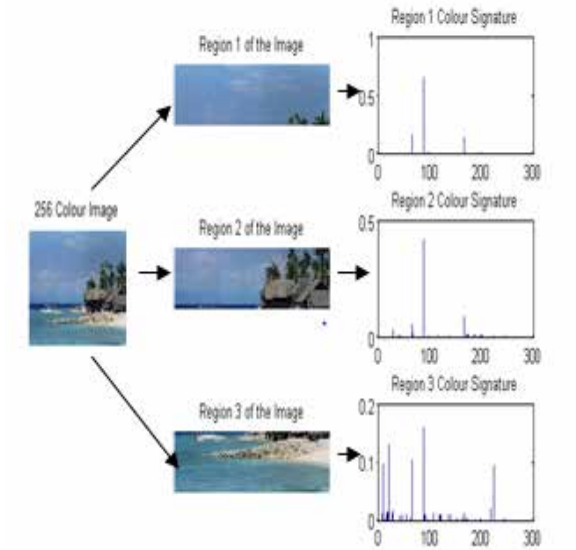


Figure.3.Creating Colour Signature from 3 Regions of the Image

CBIR USING IMAGE STRAIGHT LINE SIGNATURES [CSLS]

A line in the image space can be expressed with two variables (r, q) in polar coordinate system and a line equation can be written as

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right)$$

$$r = x\cos\theta + y\sin\theta$$

In general for each point (x_o, y_o), we can define the family of lines that goes through that point as:

$$r_o = x_o \cos\theta + y_o \sin\theta$$

A sinusoid q - r plane shown in Figure 4. represents the line passing through the point (x0 , y0).We consider only points such that r>0 and $0 < \Theta < 2\pi$. We can do the same operation for all the points in the image.If the curves of two different points intersect in theq - r plane, then two points belong to the same line.

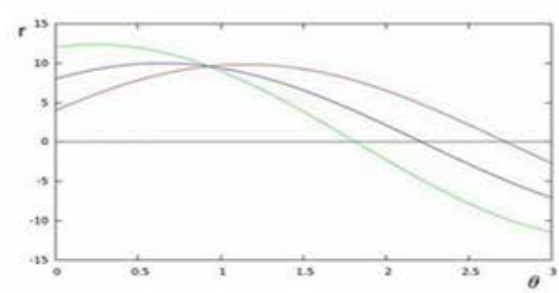


Figure .4 Illustrates the intersection of three plots in one single point (0.925, 9,6).

T

he three plots which intersect in one single point is shown in Figure 4. It represents line in which (xo,yo) (x1, y1) and (x2,y2)lay.Hough Line Transform keeps track of the intersection between curves of every point in the image.In general we can define a threshold of mimumum number of intersections needed to detect a line.If the number of intersections is above some threshold,then it declares it as a line with the parameters (q - r) of the intersection point.[7]

CBIR USING OUTLINE SIGNATURE [CSLOS]

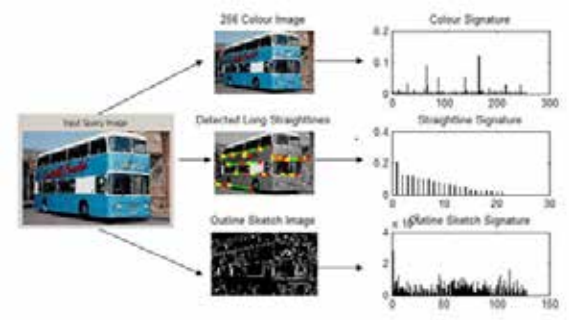


Figure 5: Colour, Straight line and Outline Sketch Signature from image.

Figure 5 represents the combine colour signature, straight line signature and outline signature.Edge detection is done by identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensity which characterize boundaries of objects in a scene. Edge detection is implemented by convolving the image with an operator (2D filter).This is sensitive to large gradients in the image while returning zero values in uniform regions. Edges occurring in images should not be missed. It should not respond to non edges. The distance between the edge pixels found by the detector and the actual edge should be minimum. It should give only one response to a single edge. The canny edge detector first smoothes the image to eliminate noise. It then finds the image gradient to highlight regions with high spatial derivatives. The algorithm then tracks along these regions and suppresses any pixel that is not at the maximum. If the magnitude is above high threshold, it is made an edge. After detecting the outline sketch image using the canny edge detection operation, the Outline Sketch signature is created using fast Fourier transform. For that, the two dimensional sketch image is converted in to a one dimensional array and FFT operation is applied. Since the outputs of one half of the FFT will be the mirror image of the another half, only 50% of the output is used as the signature. So, the 128 frequency values will be used as the outline signature of the image.[8]

IMPROVED 3 REGION COLOUR, STRAIGHT LINE AND OUTLINE SKETCH BASED CBIR (ICSLOS)

Combined 3 region colour histograms is represented as the signature S1 of input image. All the long straight lines in the the image above the threshold is used as Staight line signature S2 of the image. Outlines of the input image are detected.Then FFT of the one dimentional representation of the binary outline Image is determined . The signature is represented as outline signature S3 of the image.The combined signature S ={ S1, S2, S3} of input image is then formed. The Euclidean Distance between the input image signature and the signature of the all the images of the database is then calculated. Distances of the images are sorted in ascending order. The average image signature of the top N matching images is then determined. This average image signature V is considered as the image signature of the virtual query image and the search is repeated. Euclidean distance between the virtual input image signature V and the signature of the all the images of the database is again found. Top N ranked Images corresponding to the index from the image database are displayed.

RESULTS AND ANALYSIS

	[1]	[2]	[3]	[4]	Proposed methods			
Class	IRM	FIRM	CSP	RT F-RFM	3RCS	CSLS	CSLOS	ICSLOS
Africans	0.48	0.47	0.48	0.49	0.66	0.63	0.63	0.667
Beach	0.32	0.35	0.34	0.40	0.37	0.40	0.39	0.376
Buildings	0.35	0.35	0.33	0.39	0.48	0.47	0.47	0.476
Buses	0.36	0.60	0.52	0.58	0.67	0.70	0.70	0.739
Dinosaur	0.93	0.95	0.95	0.96	0.98	0.99	0.99	0.988
Elephant	0.38	0.25	0.40	0.50	0.43	0.43	0.43	0.424
Flower	0.42	0.65	0.60	0.75	0.68	0.76	0.75	0.774
Horse	0.72	0.65	0.70	0.80	0.90	0.88	0.88	0.905
Mountain	0.35	0.30	0.36	0.40	0.44	0.48	0.48	0.478
Food	0.38	0.48	0.46	0.51	0.64	0.59	0.59	0.603
Average	**0.471**	**0.505**	**0.514**	**0.578**	**0.625**	**0.634**	**0.634**	**0.643**

Table 1 . The Average Precision of various CBIR .
In the table 1, the first four columns are the results from 4 reference works. And the last four column are the results from our work.



Table2 . Precision, recall and false positive rate of 3RCS, CSLS, CSLOS and ICSLOS

CONCLUSION

The proposed CBIR system has been successfully implemented using Mat Lab and the performance of the system in terms of precision, recall and false positive rate was validated with images from SIMPLIcity. The arrived results were compared with previous methods IRM, FIRM, CSP and F-RFM and the performance of the proposed improved CBIR using three region colour histogram signature, straight line signature and outline sketch signature was significantly better than all other compared methods. The performance of the proposed algorithm was overall good .This method will increase the storage cost considerably. Future works may address the ways to reduce the size of the signature by using suitable feature reduction techniques.

REFERENCES

[1] A.Ramesh Kumar, D.Saravanan, "Content Based Image Retrieval Using Color Histogram" | in the International Journal of Computer Science and Information Technologies (IJCSIT), | Vol.4, Issue 2,pp.242 –245,2013. | [2] J.Jia Li, James ,Z.Wang and Gio Wiederhold,"IRM: Integrated Region Matching for Image | Retrieval," in the Proceedings of the 8th ACM International Conference on Multimedia, | pp.147-156, 2000. | [3] Y.Chen and James Z Wang, "A Region Based Fuzzy feature Matching Approach to Content | Based Image Retrieval," in IEEE Transaction onPAMI, Vol. 24,No.9,pp.1252-1267,2002. | [4] P.S.Hiremath and Jagadeesh Pujari,"Content Based Image Retrieval Using Color Boosted | Salient Points and Shape Features of an Image" in the International Journal of Image | Processing (IJIP),Volume 2,Issue 1, pp.10-17,2008. | [5] Manish Chowdhury,Sudeb Das and Malay Kumar Mundu, " Interactive Content Based | Image Retrieval using Ripplet Transform and Fuzzy Relevance Feedback" Per Min2012, | LLNCS 7143, pp.243–251,2012. | [6] L.Jayanthi,K.Lakshmi, "An Enhanced Content Based Image Retrieval using Three Region | Colour Histogram of the Image " in the International Journal of Scientific Research, | Volume 3, Issue 4, April 2014-03-28,ISSN No.2277-8179, pp. 73- 76 | [7] L.Jayanthi, K.Lakshmi, "An improved Content Based Image Retrieval using Three Region | Colour and Straight line Signatures" in the IEEE International Conference on | Communication and Signal Processing(ICCSP - 2014). | [8] L.Jayanthi, K.Lakshmi, "An improved CBIR using Three region colour, Straight line and Outline Sketch Signatures " in the IEEE International Conference on Information, Communication and Embedded Systems – ICICES- 2014. |