IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 37-41
www.iosrjournals.org

Malarial positive image retrieval using Content Based Retrieval Systems

Jasdeep Kaur[1]
[1](*Department of Computer Science, Guru Nanak College for girls, Muktsar, India)*

ABSTRACT: *CBIR Systems retrieve the relevant images by using various perceptions and mathematical calculations for example size, shape and other features of various objects in the image. These systems are being largely used in various fields. Here we use them for Retrieval of malarial positive images. In this paper retrieval could be done by using image or text as a query. To achieve best performance in order to help doctors we have proposed a methodology that does annotation of images by use of ontology. Further, more the number of objects in image larger is the need to do annotation. This methodology focuses on decrease in semantic gap by use of ontology and hence leading to increase in recall and precision. The results obtained are compared with previously defined algorithms and are found to be better than them.*
Keywords – *CBIR, Thresholding, Ontology, Similarity comparison.*

I. Introduction

Content based image retrieval systems are important part of digital image processing. These emerged in early 1990's. By content based image retrieval systems we mean getting the image similar to that we have queried. The query could be done either by image or specific property of image. Various properties of each image called feature vector are stored in database. These may be color, texture, shape, etc. image is retrieved that have similar feature vectors. This similarity is calculated by various methods. And we try to retrieve the exact image that was queried, i.e. we minimize the semantic gap [1]. The detailed structure of CBIR is explained in [2] and medical CBIR in [3].

II. Problem Formulation

In previous papers, various algorithms were proposed to retrieve the malaria positive images from various samples. But the different types of parasites were not identified. In our algorithm we have used ontology based annotation in order to decrease the semantic gap. Also this made our expert system highly accurate and reliable. The time that our system takes to retrieve the relevant image is also noted down in [4].

III. Related Work

The Content based image retrieval systems use different approaches. Need of CBIR is increasing due to increase in Image database day by day. A brief description of work done by various researchers is elaborated in below.

Table 1. Comparative analysis of work done on CBIR by various researchers.

	Paper[5]	Paper[6]	Paper[7]	Paper[8]	Paper[9]
Method for color features	Color histogram(Euclidean distance)	BIC method, gray level histogram	1st, 2nd & 3rd moments of color histogram.	K-mean (NFA) Color histogram(UM)	HSV, fuzzy color histogram.
Method for texture features	Wavelet transform (Euclidean distance & correlation)	Gabor transform, Haralick	GLCM	2-haar wavelet(NFA)GLCM(UM)	GLCM
Method for shape feature			Segmentation by 8- conn algorithm.	Shape descriptor & composite descriptor (GPF)	Canny operator, double threshold.
Method for Similarity measure	Euclidean distance method was best	Euclidean distance.		GPF have better results as it used diff. methods for diff. features.	Euclidean distance.

Ontology					PROTÉGÉ
Precession	33.3% best Euclidean distance	Best for gabor transform	Best for co-occurrence method.	NFA -76%	
Recall	14.28% best using Euclidean method	Best for gabor transform			
Advantages	Color information retrained and better performance.	Gabor filter gave better results.	Independent of architecture as java used.	NFA good results as it used local & global features.	Proposed system 70 – 100 % efficient

Some other researchers have also given significant contribution in developing CBIR systems [11-14]. To improve the efficiency of these systems ontology was added in some research works [15-16]. Method was proposed for detection of malaria parasite in [17-22].

IV. Proposed Algorithm

The Proposed work is done to fulfill the following objectives (a) to develop ontology of the image content retrieval system for annotation of each feature extracted from the image. (b) To evaluate performance of the system in terms of evaluation gap, ground truth, recall and precession. These factors lead to increase in efficiency of system. Therefore the methodology of the proposed work has been made keeping in mind the above considerations. Dataset of images is taken from [10]. We have two phases of our work. Fig.1 shows the Phase one and Fig.2 shows phase two of our work.

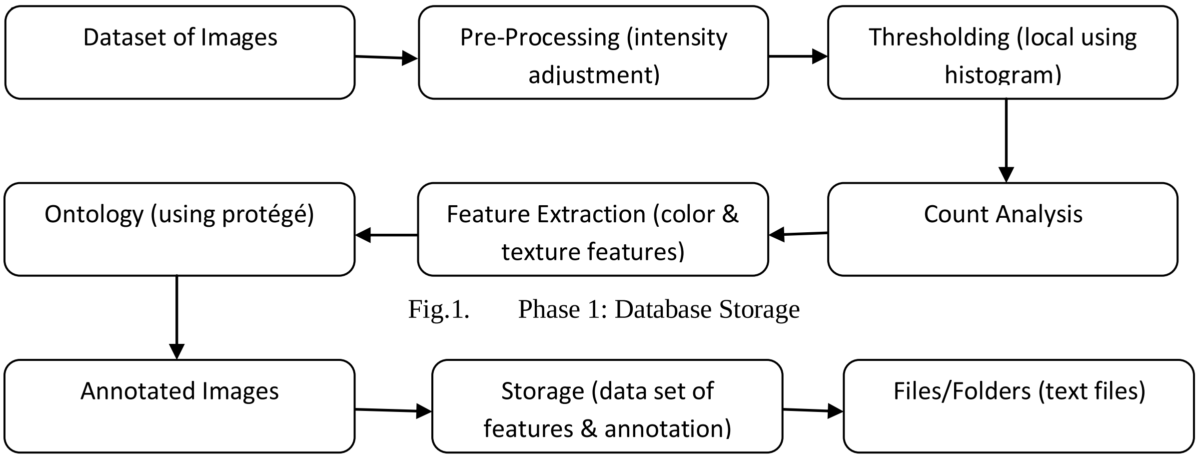

Fig.1. Phase 1: Database Storage

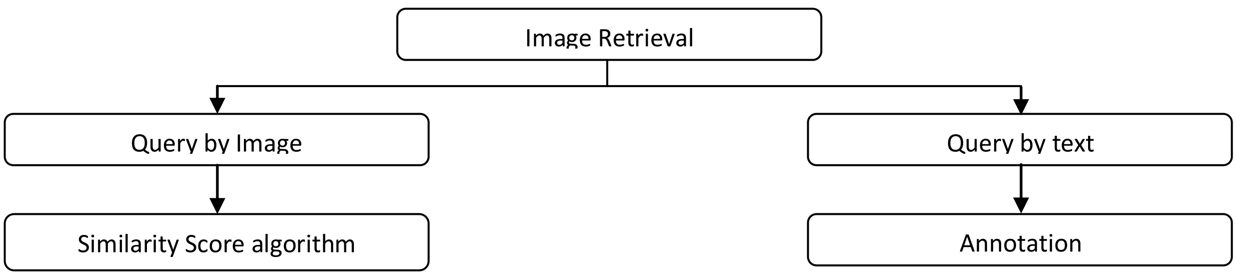

Fig.2. Phase 2: Image retrieval

4.1. Database storage phase:

In this phase, we have following steps:

4.1.1 Dataset of images: We have dataset of 129 images of malarial positive images containing different types of parasites obtained from http://www.cdc.gov/dpdx/malaria/gallery.html [9]

4.1.2 Pre-processing: Pre-processing is done to remove the noise from images so that we can get better results. In our proposed method we have focused on filtering and intensity adjustment. Filtering is done by median filter using salt and pepper noise. Morphological filter is not used because it can remove some parasite detail, so this is used where parasite information is not important. Intensity is adjusted by using imadjust. [10]

4.1.3 Segmentation: We require segmentation in order to separate out parasite and erythrocyte from background. Threshold is calculated manually from the histogram. First deepest valley is considered to be the threshold value for parasite. Second deepest valley is chosen as threshold for RBCs. This local thresholding provides better results than otsu's thresholding. Histogram of image is shown in Fig.3.

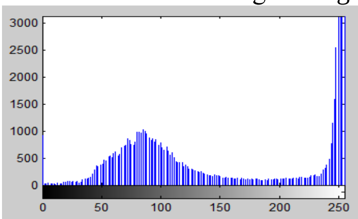

Fig.3. Image Histogram

4.1.4 Feature extraction: Mainly we extract color and texture features from the given images. Color features: firstly we calculate the values of hsv histogram. After that color auto-correlogram is calculated. It is extended histogram that carries spatial information in addition to color distribution. Texture features are calculated using Gray level co-occurrence matrix. From here we calculate the value of 13 features. These features are used to differentiate between different types of parasites.

4.1.5 Adding meta info knowledge to image: This is done using ontology. Ontology is used to capture knowledge. Various tools are present but we use protégé for this purpose.

4.2 Image retrieval:

We can retrieve images by two methods: query by image or query by text.

1.2.1 Query by image: here query is given by image and the relative images are given as result. Here Euclidean distance is calculated between the features of images in the dataset and query image. The images with minimum distance are retrieved.

1.2.2 Query by text: here the images are annotated by use of ontology by using protégé. The text is compared with the annotated database and the relevant images are retrieved.

V. RESULTS

Here the results are presented for query by text. In Fig 4. Set of retrieved images for Perimeter = 4 are shown. In Fig 5. annotation for each image is shown.

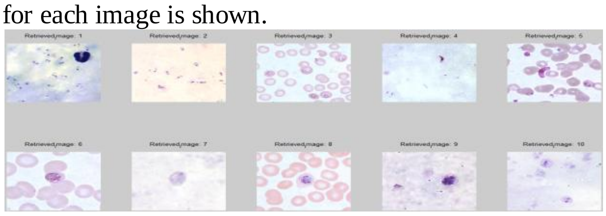

Fig.4. Retrieved image for parameter 4

```
Image_Number1threshold1=85threshold2=105Roundness=0.3Area=9.1Perimeter=4d
Image_Number2threshold1=85threshold2=105Roundness=0.3Area=5.8Perimeter=4d
Image_Number3threshold1=85threshold2=105Roundness=0.3Area=6Perimeter=4deg
Image_Number4threshold1=85threshold2=105Roundness=0.3Area=5.2Perimeter=4d
Image_Number5threshold1=85threshold2=105Roundness=0.3Area=6.5Perimeter=4d
Image_Number6threshold1=85threshold2=105Roundness=0.2Area=16.8Perimeter=4
Image_Number7threshold1=85threshold2=105Roundness=0.3Area=9.9Perimeter=4d
Image_Number8threshold1=85threshold2=105Roundness=0.3Area=6.3Perimeter=4d
Image_Number9threshold1=85threshold2=105Roundness=0.3Area=10.2Perimeter=4
Image_Number10threshold1=85threshold2=105Roundness=0.3Area=7.3Perimeter=4
```

Fig.5. Annotation for retrieved image

Precision and recall can be mathematically it can be described as

$$\text{Precision} = \frac{\text{no.of relevant items retrieved}}{\text{no.of items retrieved}} \quad \dots (1)$$

$$\text{recall} = \frac{\textit{no.of relevant items retrieved}}{\textit{no.of relevant items}} \quad \dots (2)$$

Table.2 Results of Recall, Precision and Computation time.

Query no:	Relevant items retrieved	Relevant items	Items retrieved	Recall	Precession	Computation Time
1	40	40	41	100	97.56	5.098
2	19	19	19	100	100	3.234
3	65	66	66	98.48	98.48	6.346
4	50	50	50	100	100	10.324
5	54	55	55	98.18	98.18	4.545
6	49	49	49	100	100	4.876
7	45	45	46	100	97.83	6.444
8	40	40	40	100	97.56	3.098
9	62	63	63	98.41	98.41	4.454
10	66	66	67	100	98.51	7.054

The proposed system uses ontology due to which we acquire increase in precision of system and decrease of computation time.

Table.3 Comparison of Proposed Method with Previous Works

Method	Computation time (sec)	Precision (%)
Ghosh and Ghosh [2]	8-15	95.00
Proposed algorithm	4-10	97.00

VI. Conclusion

Proposed Algorithm gave better results previous algorithm in terms of computation time and precision. This was due to use of ontology for purpose of annotation thus helping in reducing semantic gap. In order to make the system more efficient we could do further research reducing the semantic gap further between the required and retrieved image. Also we could use these techniques for other images than microscopic images.

REFERENCES

[1]. J. Sanghavi, and D. Kayande, Content based Image Retrieval (CBIR) System for Diagnosis of Blood Related Diseases, National Conference on Innovative Paradigms in Engineering & Technology (NCIPET), 2013, 11-15.
[2]. H.Müller, N.Michoux, D.Bandon and A.Geissbuhler, A Review of Content- Based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Direction, International Journal of Medical Informatics, 73, 2004 , 1-23.
[3]. T. M. Lehmann, B. B.Wein, J.Dahmen, J.Bredno, F.Vogelsang and M.Kohnen, Content-based image retrieval in medical applications: A novel multistep approach, In Proc. of IS&T/SPIE Conference on Storage and Retrieval for Media Databases, 43(4), 2004 , 312-320.
[4]. S. Ghosh, and A. Ghosh, Content Based Retrival Of Malaria Positive Images From A Clinical Database, Proceedings of the 2013 IEEE Second International Conference on Image Information Processing, 2013 , 313-318.
[5]. W.C.Seng, and S.H.Mirisaee, A Content-Based Retrieval System for Blood Cells Images, International Conference on Future Computer and Communications , Kuala Lumpur, 2009 , 412-415.
[6]. H.Chuctaya, C.Portugal, C.Beltran, and J.Gutierrez, M-CBIR: A medical content-based image retrieval system using metric data-structures, Procedings of the 30th International Conference of the Chilean Computer Science Society, 2011 , 135-141.
[7]. Bhagat, and M.Atique, Design and Development of systems for image segmentation and Content Based Image Retrieval,proceedings CISP 2012, 2012 , 109-113.

[8]. S. Khan, A. Hussain, and I.Alshaikhli, Comparitive study on content based image retrieval(CBIR), International conference on advanced computer science applications and technology, 2012 , 61-66.
[9]. D. Gowsikhaa, S.Abirami and R. Baskaran, Construction of image ontology using low level features for image retrieval, International conference on comport communication and informatics, 2012 , 1-7.
[10]. Laboratory identification of parasites of public health concern," Centers for Disease Control Prevention, National Center for Infectious Diseases, Division of Parasitic Diseases. http://www.cdc.gov/dpdx/malaria/gallery.html
[11]. P.H.Bugatti, A.J.M.Traina, and C.Traina, Improving Content-Based Retrieval of Medical Images through Dynamic Distance on Relevance Feedback, Computer based medical systems, 2011 , 1-6.
[12]. T.M. Lehmann, M.O. Guld, C.Thies, B.A. Plodowski, D.B.Keysers, B.C.Ott, H.Schubert, C.Thies, B.Plodowski and B.Ott, IRMA – Content-Based Image Retrieval in Medical Applications,Proceedings of 14th World Congr. Med. Inf. (Medinfo), 2(1), 2004 , 842-846.
[13]. Y.Liu, D.Zhang, G.Lu, and W.Y.Ma, A survey of content-based image retrieval with high-level semantics,The journal of pattern recognition society, 2007 , 262-282.
[14]. N.Doulamis, and A.Doulamis, Evaluation of relevance feedback schemes in content based in retrieval systems, Signal processing:Image communication, 21, 2006 , 334-357.
[15]. V. Mezaris, I. Kompatsiaris, and M.G. Strintzis, An ontology approach to object based image retrieval, in proceedings IEEE international conference on image processing, Barcelona, Spain, 2, 2003 , 511-514.
[16]. T.W. Chang, Y.P.Huang, F.E. Sandnes, An Ontology Oriented Region Based Image Retrieval Strategy, IEEE International Conference on Systems, Man and Cybernetics, 2008 , 2671 – 2676.
[17]. M. Khan, and B. Acharya, Content Based Image Retrieval Approaches for Detection of Malarial Parasite in Blood Images, International Journal of Biometrics and Bioinformatics (IJBB), 5, 2011 , 97-110.
[18]. G. Diaz, A.Gonzalez, and E. Romero, A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images, Journal of Biomedical Informatics, 2009 , 296-307.
[19]. W.C.Seng, and S.H.Mirisaee, A Content-Based Retrieval System for Blood Cells Images, International Conference on Future Computer and Communications (ICFCC 2009), Kuala Lumpur, 2009 , 412-415.
[20]. P. Suradkar, Detection of Malarial Parasite in Blood Using Image Processing, International Journal of Engineering and Innovative Technology (IJEIT), 2, 2013 , 124-126.
[21]. Tek, and A. Dempster, and I.Kale, Computer vision for microscopy diagnosis of malaria, Malaria Journal, 8, 2009 , 1-14.
[22]. Tek, and A. Dempster, and I.Kale , Parasite detection and identification for automated thin blood film malaria diagnosis, 114(1), 2010, 21-32.