Advances in Intelligent Systems and Computing 249

Suresh Chandra Satapathy
P. S. Avadhani
Siba K. Udgata
Sadasivuni Lakshminarayana *Editors*

# ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India - Volume II

Hosted by CSI Vishakapatnam Chapter

Springer

# Advances in Intelligent Systems and Computing

Volume 249

*Series Editor*

Janusz Kacprzyk, Warsaw, Poland

For further volumes:
http://www.springer.com/series/11156

Suresh Chandra Satapathy · P.S. Avadhani
Siba K. Udgata · Sadasivuni Lakshminarayana
Editors

# ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India - Volume II

Hosted by CSI Vishakapatnam Chapter

Springer

*Editors*
Suresh Chandra Satapathy
Anil Neerukonda Institute of Technology
and Sciences, Sangivalasa
(Affilated to Andhra University)
Vishakapatnam
Andhra Pradesh
India

P.S. Avadhani
College of Engineering (A)
Andhra University
Vishakapatnam
India

Siba K. Udgata
University of Hyderabad
Hyderabad
Andhra Pradesh
India

Sadasivuni Lakshminarayana
CSIR-National Institute of Oceanography
Vishakapatnam
India

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-03094-4 ISBN 978-3-319-03095-1 (eBook)
DOI 10.1007/978-3-319-03095-1
Springer Cham Heidelberg New York Dordrecht London



Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

This AISC volume contains the papers presented at the $48^{th}$ Annual Convention of Computer Society of India (CSI 2013) with theme 'ICT and Critical Infrastructure' held during 13th –15th December 2013 at Hotel Novotel Varun Beach, Visakhapatnam and hosted by Computer Society of India, Vishakhapatnam Chapter in association with Vishakhapatnam Steel Plant, the flagship company of RINL, India.

Computer society of India (CSI) was established in 1965 with a view to increase information and technological awareness among Indian society, and to make forum to exchange and share the IT- related issues. The headquarters of the CSI is situated in Mumbai with a full-fledged office setup and is coordinating the individual chapter activities. It has 70 chapters and 418 students' branches operating in different cities of India. The total strength of CSI is above 90000 members.

CSI Vishakhapatnam Chapter deems it a big pride to host this prestigious $48^{th}$ Anuual Convention after successfully organizing various events like INDIA-2012, eCOG-2011, 28th National Student convention, and AP State Student Convention in the past.

CSI 2013 is targeted to bring researchers and practitioners from academia and industry to report, deliberate and review the latest progresses in the cutting-edge research pertaining to emerging technologies.

Research submissions in various advanced technology areas were received and after a rigorous peer-review process with the help of program committee members and external reviewer, 173 ( Vol-I: 88, Vol-II: 85) papers were accepted with an acceptance ratio of 0.43.

The conference featured many distinguished personalities like Dr. V.K. Saraswat, Former Director General, DRDO, Prof. Rajeev Sangal, Director, IIT-BHU, Mr. Ajit Balakrishnan, Founder & CEO Rediff.com, Prof. L.M. Patnaik, Former Vice Chancellor, IISc, Bangalore, Prof. Kesav Nori, IIIT-H & IIT-H, Mr. Rajesh Uppal, Executive Director & CIO, Maruti Suzuki-India, Prof. D. Krishna Sundar, IISc, Bangalore, Dr. Dejan Milojicic, Senior Researcher and Director of the. Open Cirrus Cloud Computing, HP Labs, USA & President Elect 2013, IEEE Computer Society, Dr. San Murugesan, Director, BRITE Professional Services, Sydney, Australia, Dr. Gautam Shroff, VP and Chief Scientist, Tata Consultancy Services, Mr. P. Krishna Sastry, TCS, Ms. Angela R. Burgess Executive Director, IEEE Computer Society, USA, Mr. Sriram Raghavan, Security & Digital Forensics Consultant, Secure Cyber Space, and Dr. P. Bhanu Prasad, Vision Specialist, Matrix Vision GmbH, Germany among many others.

Four special sessions were offered respectively by Dr. Vipin Tyagi, Jaypee University of Engg. & Tech., Prof. J.K. Mandal, University of Kalyani, Dr. Dharam Singh, CTAE, Udaipur, Dr. Suma V., Dean, Research, Dayananda Sagar Institutions, Bengaluru. Separate Invited talks were organized in industrial and academia tracks in both days. The conference also hosted few tutorials and workshops for the benefit of participants.

We are indebted to Andhra University, JNTU-Kakinada and Visakhapatnam Steel plant for their immense support to make this convention possible in such a grand scale. CSI 2013 is proud to be hosted by Visakhapatnam Steel Plant (VSP), which is a Govt. of India Undertaking under the corporate entity of Rashtriya Ispat Nigam Ltd. It is the first shore-based integrated steel plant in India. The plant with a capacity of 3 mtpa was established in the early nineties and is a market leader in long steel products. The Plant is almost doubling its capacity to a level of 6.3 mtpa of liquid steel at a cost of around 2500 million USD. RINL-VSP is the first integrated steel plant in India to be accredited with all four international standards, viz. ISO 9001, ISO 14001, ISO 50001 and OHSAS 18001. It is also the first Steel Plant to be certified with CMMI level-3 certificate and BS EN 16001 standard.

Our special thanks to Fellows, President, Vice President, Secretary, Treasurer, Regional VPs and Chairmen of Different Divisions, Heads of SIG Groups, National Student Coordinator, Regional and State Student Coordinators, OBs of different Chapters and Administration staff of CSI-India. Thanks to all CSI Student Branch coordinators, Administration & Management of Engineering Colleges under Visakhapatnam chapter for their continuous support to our chapter activities. Sincere thanks to CSI-Vizag members and other Chapter Members across India those who have supported CSI-Vizag activities directly or indirectly.

We take this opportunity to thank authors of all submitted papers for their hard work, adherence to the deadlines and patience with the review process. We express our thanks to all reviewers from India and abroad who have taken enormous pain to review the papers on time.

Our sincere thanks to all the chairs who have guided and supported us from the beginning. Our sincere thanks to senior life members, life members, associate life members and student members of CSI-India for their cooperation and support for all activities.

Our sincere thanks to all Sponsors, press, print & electronic media for their excellent coverage of this convention.

December 2013

Dr. Suresh Chandra Satapathy
Dr. P.S. Avadhani
Dr. Siba K. Udgata
Dr. Sadasivuni Lakshminarayana

# Organization

## Chief Patrons

Shri A.P. Choudhary, CMD, RINL
Prof. G.S.N. Raju, VC, Andhra University
Prof. G. Tulasi Ram Das, VC, JNTU-Kakinada

## Patrons

Sri Umesh Chandra, Director (Operations), RINL
Sri P. Madhusudan, Director (Finance), RINL
Sri Y.R. Reddy, Director (Personnel), RINL
Sri N.S. Rao, Director (Projects), RINL

## Apex (CSI) Committee

Prof. S.V. Raghavan, President
Shri H.R. Mohan, Vice President
Dr S. Ramanathan, Hon. Secretary
Shri Ranga Rajagopal, Hon. Treasurer
Sri Satish Babu, Immd. Past President
Shri Raju L. Kanchibhotla, RVP, Region-V

## Chief Advisor

Prof D.B.V. Sarma, Fellow, CSI

## Advisory Committee

Sri Anil Srivastav, IAS, Jt. Director General of Civil Aviation, GOI
Sri G.V.L. Satya Kumar, IRTS, Chairman I/C, VPT, Visakhapatnam

Sri N.K. Mishra, Rear Admiral IN (Retd), CMD, HSL, Visakhapatnam
Capt.D.K. Mohanty, CMD, DCIL, Visakhapatnam
Sri S.V. Ranga Rajan, Outstanding Scientist, NSTL, Visakhapatnam
Sri Balaji Iyengar, GM I/C, NTPC-Simhadri, Visakhapatnam
Prof P. Trimurthy, Past President, CSI
Sri M.D. Agarwal, Past President, CSI
Prof D.D. Sharma, Fellow, CSI
Sri Saurabh Sonawala, Hindtron, Mumbai
Sri R.V.S. Raju, President, RHI Clasil Ltd.
Prof. Gollapudi S.R., GITAM University & Convener, Advisory Board, CSI-2013

## International Technical Advisory Committee:

Dr. L.M. Patnaik, IISc, India
Dr. C. Krishna Mohan, IIT-H, India
Dr. K. RajaSekhar Rao, Dean, KLU, India
Dr. Chaoyang Zhang, USM, USA
Dr. Wei Ding, USA
Dr. Cheng-Chi Lee, Taiwan
Dr. Pramod Kumar Singh, ABV-IIITM, India
Dr. B. Biswal, GMRIT, India
Dr. G. Pradhan, BBSR, India
Dr. B.K. Panigrahi, IITD, India
Dr. S. Yenduri, USA
Dr. D. Janikiram, IIT-M
Dr. H.R. Vishwakarma, VIT, India
Dr. Deepak Garg, TU, Patiala
Dr. Hamid Arabnia, USA
Dr. Azah Kamilah Muda, Malaysia
Dr. Yun-Huoy Choo, Malaysia
Dr. Hongbo Liu, Dalian Maritime
Dr. B.N. Biswal, BEC, India
Dr. Saihanuman, GRIET, India
Dr. L. Perkin, USA
Dr. V. Shenbagraman, SRM, India and many others

## Organizing Committee

**Chair**
Sri T.K. Chand, Director (Commercial), RINL & Chairman, CSI-Vizag

**Co-Chairmen**
Sri P.C. Mohapatra, ED (Projects), Vizag Steel
Sri P. Ramudu, ED (Auto & IT), Vizag Steel

**Vice-Chairman**
Sri K.V.S.S. Rajeswara Rao, GM (IT), Vizag Steel

**Addl. Vice-Chairman:**
Sri Suman Das, DGM (IT) & Secretary, CSI-Vizag

**Convener**
Sri Paramata Satyanarayana, Sr. Manager (IT), Vizag Steel

**Co-Convener**
Sri C.K. Padhi, AGM (IT), Vizag Steel

**Web Portal Convener**
Sri S.S. Choudhary, AGM(IT)

**Advisor (Press & Publicity)**
Sri B.S. Satyendra, AGM (CC)

**Convener (Press & Publicity)**
Sri Dwaram Swamy, AGM (Con)

**Co-Convener (Press & Publicity)**
Sri A.P. Sahu, AGM (IT)

## Program Committee

**Chairman:**
Prof P.S. Avadhani, Vice Principal, AUCE (A)

**Co Chairmen**
Prof D.V.L.N. Somayajulu, NIT-W
Dr S. Lakshmi Narayana, Scientist E, NIO-Vizag

**Conveners:**
Sri Pulle Chandra Sekhar, DGM (IT), Vizag Steel – Industry
Prof. S.C. Satapathy, HOD (CSE), ANITS – Academics
Dr. S.K. Udgata, UoH, Hyderabad

## Finance Committee

**Chairman**
Sri G.N. Murthy, ED (F&A), Vizag Steel

**Co-Chairmen**
Sri G.J. Rao, GM (Marketing)-Project Sales
Sri Y. Sudhakar Rao, GM (Marketing)-Retail Sales

**Convener**
Sri J.V. Rao, AGM (Con)

**Co-Conveners**
Sri P. Srinivasulu, DGM (RMD)
Sri D.V.G.A.R.G. Varma, AGM (Con)

Sri T.N. Sanyasi Rao,
Sr.Mgr (IT)

**Members** Sri V.R. Sanyasi, Rao
Sri P. Sesha Srinivas

## Convention Committee

**Chair** Sri D.N. Rao, ED(Operations), Vizag Steel

## Vice-Chairs

Sri Y. Arjun Kumar
Sri S. Raja
Dr. B. Govardhana Reddy
Sri ThyagaRaju Guturu
Sri Bulusu Gopi Kumar
Sri Narla Anand

## Conveners

Sri S.K. Mishra
Sri V.D. Awasthi
Sri M. Srinivasa Babu
Sri G.V. Ramesh
Sri A. Bapuji
Sri B. Ranganath
Sri D. Satayanarayana
Sri P.M. Divecha
Sri Y.N. Reddy
Sri J.P. Dash
Sri K. Muralikrishna

## Co-Conveners

Sri Y. Madhusudan Rao
Sri S. Gopal
Sri Phani Gopal
Sri M.K. Chakravarty
Sri P. Krishna Rao
Sri P. Balaramu
Sri B.V. Vijay Kumar
Mrs M. Madhu Bindu
Sri P. Janardhana
Sri G.V. Saradhi

## Members:

Sri Y. Satyanarayana
Sri Shailendra Kumar
Sri V.H. Sundara Rao
Mrs K. Sarala
Sri V. Srinivas
Sri G. Vijay Kumar
Mrs. V.V. Vijaya Lakshmi
Sri D. Ramesh
Shri K. Pratap
Sri P. Srinivasa Rao
Sri S. Adinarayana
Sri B. Ganesh
Sri Hanumantha Naik
Sri D.G.V. Saya

Sri V.L.P. Lal
Sri U.V.V. Janardhana
Sri S. Arun Kumar
Sri K. Raviram
Sri N. Prabhakar Ram
Sri BH.B.V.K. Raju
Sri K. Srinivasa Rao
Mrs T. Kalavathi
Mrs A. Lakshmi
Sri N. Pradeep
Sri K. Dilip
Sri K.S.S. Chandra Rao
Sri Vamshee Ram
Ms Sriya Basumallik
Sri Kunche Satyanarayana
Sri Shrirama Murthy

# Contents

## Session 3: Network Security, Digital Forensics and Cyber Crime

# Content Based Retrieval of Malaria Positive Images from a Clinical Database VIA Recognition in RGB Colour Space

Somen Ghosh and Ajay Ghosh

Department of Applied Optics and Photonics, University of Calcutta
92, APC Road, Kolkata 700009
ghosh_somen@hotmail.com, aghosh.cu@gmail.com

**Abstract.** Modern hospitals are trying to create a database of patients' diagnostic history that also contains multiple images taken during different clinical tests on a patient. This has lead to a demand for easy retrieval of images matching a query condition, so that this database can be used as a clinical decision support system. This paper presents a technique for retrieval of malaria positive images, matching a specific query condition, from a clinical image database. The candidate image is segmented in RGB colour space, and a pseudo-colour is imparted to the non-region of interest pixels. The technique additionally retains the full features of the chromosomes, and hence the modified image can be used for further studies on the chromosomes. The algorithm utilizes 4-connected labeled region map property of images to analyze and modify the image, i.e., delete unwanted artifacts, etc. This property is also used to count the number of RBCs.

**Keywords:** Malaria, Segmentation, RGB space, cell counting, labeled regions.

## 1 Introduction

Modern hospitals are attempting to create a database of patients' diagnostic history. More and more patient records in this database contain data in form of images. This has lead to a demand for an application for easy retrieval of images and associated diagnostic details so that it can be used as a clinical decision support system.

While it is easy to do a query on structured information available in a database, image data being random and statistical in nature, query on them is not trivial. In the past many content based image retrieval systems have been proposed using different visual features. Shape is one key feature proposed in literature [1,4,6,9,14] . However, the shape of real objects varies in images depending on the pose, illumination, etc. Hence such techniques have limited applications. Recently a few alternative methods have been proposed that use texture, colour, etc. [2,3,8,10,11,13,15]. However these methods suffer from poor segmentation quality. None of these methods can however be directly applied to biomedical images. Such images are characterized by the occurrence of multiple objects in the same image, and are usually stained with

S.C. Satapathy et al. (eds.), *ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of CSI - Volume II*, Advances in Intelligent Systems and Computing 249,
DOI: 10.1007/978-3-319-03095-1_1, 

objective specific stains. Hence such images need segmentation techniques that search for specific signatures of the object being searched.

In this paper we attempt to create an algorithm for identifying the presence of malaria, the type of smear (thick or thin) whose image is being studied, and the degree of disease. The method segments the chromosomes of the parasites from the rest of the image, and, at the same time retains its complete colour (rgb) details. A pseudo-coloring technique is used to impart a specific colour to the pixels that do not represent the chromosomes. The method utilizes 4-connected labeled region properties of digital images to count the RBCs within the image so that the degree of the disease can be estimated.

Clinical databases usually have the following types of images: (1)x-ray, (2)ultra-sonographs and echo-cardiographs, (3)ECG (4)CT and MRI (5)pathological and histo-pathological and (6) nuclear medicine (bone density/vessel maps, etc.). The application is expected to successfully identify malaria positive images from a database containing these diverse types of images.

## 2 Signature of Malaria

The presence of the colored chromatin dots is a sufficient signature to indicate the presence of parasite in the digital images of stained blood smears. Both thin and thick smears carry this signature. The manual clinical test process involves visual discrimination between the colours of chromatin dots and the RBCs under a microscope. These chromatin dots are usually located within the RBCs but in advanced disease they can exit and RBC. Besides, a diseased RBC can have more than one chromatin dots, but they are considered as single infection. Thus we can say that the presence of chromatin dots is a sufficient signature to indicate the presence of malaria parasite in the digital image of stained blood smears.

Extracts from malaria positive images are shown in figure 1.

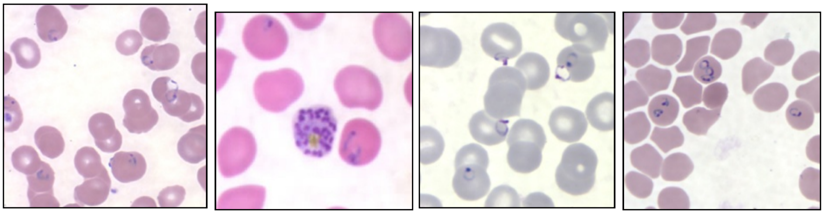

**Fig. 1.** Extracts from malaria-positive images

## 3 Algorithm for Recognition of Malaria

To meet the objective of the segmentation, our approach would be: 1)work in rgb colour space, 2) pixels other than those representing chromosomes should be assigned a specific colour, and then, 3)the presence of the chromosomes should be tested in the modified image.

In the first step, some knowledge about images found in a clinical database is used to eliminate certain category of images from consideration. Since malaria positive images are RGB images, hence gray images (single layered or 3-layered) are eliminated from consideration.

The colour that is to be assigned to the non-Region of Interest (non-RoI) pixels must satisfy the following conditions: 1)have a colour distinct from that rendered by the stain to the chromatin dots, and 2)have an intensity that is less than the intensity of the RBCs. The color is defined in the HSI space(H and I can be defined separately) and then the equivalent r, g and b values calculated. A hue of green was chosen to be rendered to the non-RoI pixels, since it has maximum contrast with the hue of the stained chromatin dots. A hue value of $227^0$ was selected that is located within the range of green hues ($210^0$ – $269^0$) in the hue-scale of HSI space. To arrive at the intensity value, the candidate RGB image was converted to a gray scale image, and then the average intensity of the pixels representing the RBCs was measured. An intensity value that was 50% of the average intensity existing within the RBCs was selected. The saturation value (S) was taken at 0.2. These HSI values (H=$227^0$,S=0.2 and I=as calculated) were converted to the corresponding r, g, and b values for RGB space. This calculated r, g and b values were the pseudo-colour that was assigned to the non-RoI pixels.

Identification and pseudo-coloring of non-RoI pixels is a two step process. In the first step we make use of the fact that in the gray scale version of the image (fig. 2a), the background has an intensity distinct from that of the pixels representing RBCs. Hence this image is thresholded at an intensity value given by Otsu's formula. This results in a binary image that has the background pixels identified as 1(white). This binary image is used as a template image to identify the location of the background pixels in the original RGB image, and these pixels are assigned the new defined pseudo-colour (fig. 4). The modified image still has the RBCs and the chromosomes retained in their original colour.

Additionally, the count of the background pixels in the binary image is used to decide whether the smear was thick or thin. Only images of thin smears are analysed.

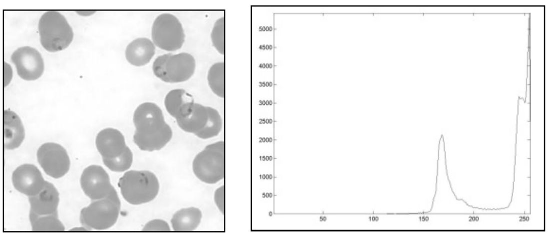

**Fig. 2.** (a) Gray scale image of 1(a) and (b) its histogram

A histogram of the gray scale version of the modified image shows three distinct regions (fig. 5): 1) the intensity assigned to the background pixels, this is also the lowest intensity, 2) an intermediate intensity possessed by a low number of pixels, and 3) a higher intensity possessed by pixels representing the RBCs. Otsu's formula returns an intensity value that is between regions 2 and 3 indicated above. However, for this to be successful, the volume of background pixels should be high. The gray scale image is thresholded at this threshold value to create another binary mask image. This new template image is used to identify the foreground pixels and they are also assigned the pseudo-colour defined.

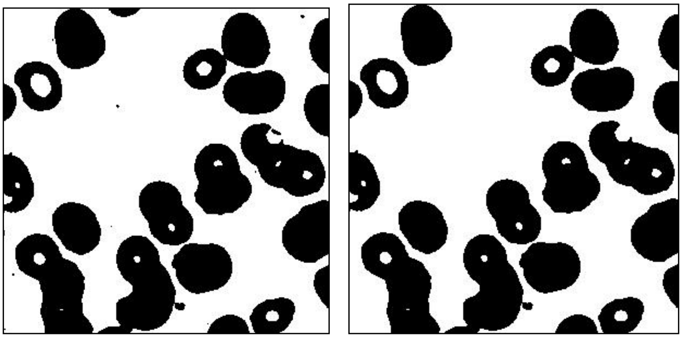

**Fig. 3.** (a),(b) Binary mask images of fig 1(a), before and after removing artifacts from background

Note that the first binary mask image (fig. 3a) clearly shows artifacts in the background region. These are black regions on a white background. To remove the artifacts, a digital negative of the binary image is converted into a 4-connected region labeled image. The population of each labeled region is an indication of area occupied by each region. Any region with an area less than 0.3% of the total image size was erased and the corresponding pixels in the binary mask image were marked as background pixels.

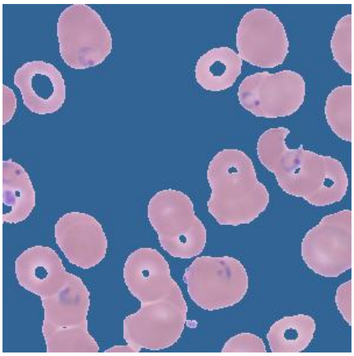

**Fig. 4.** Image of 1(a) with background pixels pseudo-colored

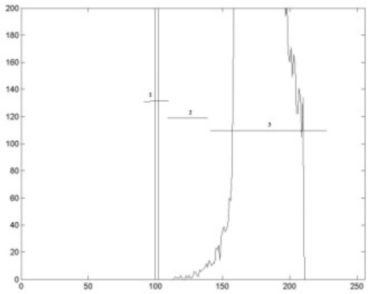


**Fig. 5.** Zoomed histogram of modified image

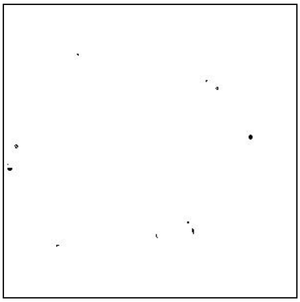

**Fig. 6.** Chromatin dots identified for 1(a) – binary image used for easy demonstration

# 4 Algorithm for Counting of RBCs

The degree of disease is estimated for thin smears only. The degree of the disease is given by the formula

[No. of affected RBCs / Total number of RBCs in view ] x 100

Hence we need to count the number of RBCs and affected RBCs in the field of view. For this we have used the final binary mask prepared previously. The mask image is converted into a 4-connected, labeled region map array. The centroids, and major and minor axis of all disjoint regions are measured. The number of disjoint regions is essentially the number of cells in the field of view. However the count needs to be corrected to account for two factors: 1) partially visible cells located at the boundary of the image and 2) overlapped cells.

The ratio of the major axis to minor axis is used to decide whether the cells are free standing or are overlapped. Free standing cells have a ratio ~1. The minor axis of cells with ratio ≈ 1 is taken as the diameter of a RBC. If the major or the minor axis of any region is less than the radius of a free standing cell, then the cell is partially visible. For cells having ratio > 1, if their major or minor axis is ≥ diameter of the cell, then cells are partially overlapped.

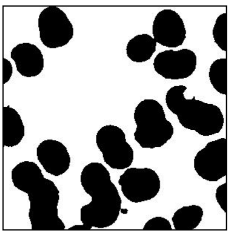

**Fig. 7.** Binary template with holes removed

Partially visible cells are identified as those cells having their centroids at a distance less than the radius of the cell from the boundary of the image. These cells are removed from the count. Overlapped cells are identified by their ratio. All regions having ratio greater than a threshold were considered for resolving. The major axis was divided by the diameter of the cells to find out how many cells were overlapped. A similar study was done on minor axis. However, for cases where the ratio was higher than the threshold, but the minor axis was less than 80% of the diameter of an RBC, the resolution was not done. These additional numbers of cells were added to arrive at the total number of cells.

## 5 Experimental Results

The training set of images was sourced from [5]. The prototype was developed using MATLAB.

The program fetches one image at a time and rejects images that do not match the filtering condition. A candidate image is analysed and if found to be malaria- positive, the information obtained ( thick or thin smear, degree of infection) is written to the database.

The image in figure 1(a) exhibited an average intensity of 227 within the RBCs. Hence the I value was taken as 110 for calculation of R,G and B values. The HSI values of H=$227^0$, S=0.2 and I=110 translates to R=85, G=99 and B=146 in RGB space. However, since MATLAB was used for the initial study, it calculates intensity using the formula 0.299R+0.587G+0.114B and not 1/3[R+G+B]. Thus, the resultant histogram shows the intensity value as 100 for the pseudo-colored pixels, and not 110 as expected.

The following table gives details of the calculations to arrive at the number of RBCs.

**Table 1.** Calculation of number of RBCs

| Reg. | Centroid 1 | Centroid 2 | Axes 1 | Axes 2 | Ratio | RBC |
|---|---|---|---|---|---|---|
| 1 | 5.52 | 86.80 | 33.55 | 12.72 | 2.64 | 0 |
| 2 | 12.52 | 163.29 | 56.16 | 29.15 | 1.93 | 0 |
| 3 | 33.47 | 71.96 | 49.27 | 39.95 | 1.23 | 1 |
| 4 | 47.82 | 255.71 | 109.94 | 48.03 | 2.29 | 3 |
| 5 | 68.11 | 203.18 | 51.59 | 40.69 | 1.27 | 1 |
| 6 | 69.65 | 29.17 | 57.14 | 45.83 | 1.25 | 1 |
| 7 | 129.56 | 259.08 | 97.33 | 58.52 | 1.66 | 3 |
| 8 | 150.40 | 190.05 | 65.77 | 41.59 | 1.58 | 2 |
| 9 | 184.17 | 239.85 | 55.53 | 47.23 | 1.18 | 1 |
| 10 | 208.85 | 40.92 | 99.00 | 46.17 | 2.14 | 2 |
| 11 | 200.47 | 159.80 | 68.37 | 50.14 | 1.36 | 1 |

**Table 1.** (*continued*)

| | | | | | | |
|---|---|---|---|---|---|---|
| 12 | 208.61 | 293.39 | 37.92 | 17.88 | 2.12 | 0 |
| 13 | 232.52 | 77.74 | 59.72 | 41.15 | 1.45 | 1 |
| 14 | 258.23 | 143.24 | 93.01 | 44.27 | 2.10 | 2 |
| 15 | 248.86 | 284.61 | 43.97 | 31.52 | 1.39 | 0 |
| 16 | 281.67 | 208.17 | 58.92 | 43.33 | 1.36 | 0 |
| 17 | 290.78 | 93.40 | 49.14 | 24.01 | 2.05 | 0 |
| 18 | 293.19 | 34.54 | 50.15 | 18.51 | 2.71 | 0 |
| 19 | 294.86 | 256.08 | 37.23 | 14.50 | 2.57 | 0 |

Diameter of free standing cell = 40.
Following regions are partially visible and are not counted: 1,2,12,15,16,17,18,19.
Partially overlapped cells that needed to be resolved are : 2,5,12.
Number of regions = 19
Partially visible cells to be removed = 8
Overlapped cells to be added = 7
Total cells = 19 - 8 + 7 = 18.

An analysis of the results obtained shows: (1) Malaria positive images available in the database in gray scale are not retrieved. These images are however not normal for clinical databases. (2) Images of malaria positive slides taken under very high magnification could not be processed. This is because a distinct background could not be identified in such images. These images too are not normal for a clinical database. (3) Images of thick smears were retrieved irrespective of the degree of the disease. This is because the algorithm does not do counting of cells in images of thick smears.

## 6 Conclusion

Using the algorithm described, we were able to successfully demonstrate that it is possible to build a system to identify malaria positive images in a clinical database, especially for images of thin smears. Thus we are able to convert an image database to a 'fetch-able' information repository. However, to develop a fully functional decision support system, separate query services need to be written for different categories of images : X-rays, MRI, ECG, etc. A key advantage of the algorithm is that is automatically adapts to the different stains used in different laboratories. However the algorithm does not yet identify the type of parasite. This will be included in a future development.

**Acknowledgement.** Financial assistance from the Department of Applied Optics and Photonics, University of Calcutta, under the Technical Education Quality Improvement Program is acknowledged by the first author. Opinions expressed and conclusions drawn are of the authors and are not to be attributed to the Department of Applied Optics and Photonics.

## References

1. Adoram, M., Lew, M.S.: IRUS: Image Retrieval using Shape. In: IEEE International Conference on Multimedia Computing and Systems, vol. 2, pp. 597–602 (1999)
2. Avula, S.R., Tang, J., Acton, S.T.: An object based image retrieval system for digital libraries. Multimedia Systems Journal 11(3), 260–270 (2006)
3. Avula, S.R., Tang, J., Acton, S.T.: Image retrieval using Segmentation. In: Proceedings of IEEE 2003 Systems and Information Engineering Design Symposium, pp. 289–294 (2003)
4. Bartolini, I., Ciaccia, P., Patella, M.: WARP: Accurate Retrieval of shapes using phase of Fourier descriptors and time warping distance. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(1), 142–147 (2005)
5. Center for Disease Control and Prevention - malaria image library, `http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/Malaria_il.htm` (accessed July 2013)
6. Folkers, A., Samet, H.: Content based image retrieval using Fourier descriptors on a Logo database. In: IEEE Proceedings of 16th International Conference on Pattern Recognition 2002, vol. 3, pp. 521–524 (2002)
7. Ghosh, S., Kundu, S., Ghosh, A.: Rapid detection of malaria via digital processing of images of thin smears of peripheral Blood. In: Proceedings of International Conference on Trends in Optics and Photonics II, pp. 181–187 (2011)
8. Kutics, A., Nakajima, M., Ieki, T., Mukawa, N.: An object-based image retrieval system using an inhomogeneous diffusion model. In: Proceedings of International Conference on Image Processing 1999, ICIP 1999, vol. 2, pp. 590–594 (1999)
9. Milios, E., Petrakis, E.G.M.: Shape retrieval based on dynamic programming. IEEE Transactions on Image Processing 9(1), 141–147 (2000)
10. Natsev, A., Rastogi, R., Shim, K.: WALRUS: a similarity retrieval algorithm for image databases. IEEE Transactions on Knowledge and Data Engineering 16(3), 301–316 (2004)
11. Rajakumar, K., Muttan, S.: Medical image retrieval using energy efficient wavelet transform. In: International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–5 (2010)
12. Ross, N.E., Pritchard, C.J., Rubin, D.M., Duse, A.G.: Automatic Image processing method for the diagnosis and classification of malaria on thin blood smears. Med. Biol. Eng. Comput. 44, 427–436 (2006)
13. Sheikh, A.R., Lye, M.H., Mansor, S., Fauzi, M.F.A.: A content based image retrieval system for marine life images. IEEE 15th International Symposium on Consumer Electronics (ISCE), pp. 29–33 (2011)
14. Wang, Z., Chi, Z., Feng, D.: Shape based leaf image retrieval. In: IEEE Proceedings on Vision, Image and Signal Processing, vol. 150, pp. 34–43 (2003)
15. Huang, Y., Zhang, J., Zhao, Y., Ma, D.: Medical Image Retrieval with Query-Dependent Feature Fusion Based on One-Class SVM. In: IEEE 13th International Conference on Computational Science and Engineering (CSE), pp. 176–183 (2010)

# Modified Clustered Approach for Performance Escalation of Distributed Real-Time System

Urmani Kaushal[1] and Avanish Kumar[2]

[1] Department of Physical & Computer Sciences, FASC,
MITS Deemed University, Lakshmangarh-332311, Sikar, Raj., India
[2] Department of Maths, Stats & Computer Applications, Bundelkhand University,
Jhansi, U.P., India
{urmani10kaushal,dravanishkumar}@gmail.com

**Abstract.** In Distributed Real-Time Systems the execution of computational tasks must be completed within time else catastrophe may ensue. This constraint can be achieved by optimal task allocation in DRTS. In the process of task allocation, an appropriate and efficient load sharing strategies can improve the performance of distributed system. In this paper an efficient, two level clustering approach has been proposed. The initial level of clustering will be done on the basis of execution cost and inter task communication cost and second at the stage of task allocation. In the allocation of tasks a more scalable active approach is being followed which iteratively refines the performance of DRTS. The proposed model has been simulated in MATLAB 7.11.0. Two examples have been demonstrated to illustrate the model and algorithm for performance improvement in distributed system.

**Keywords:** Distributed Real-Time System, Task Allocation, Load Sharing, Clustering, Execution Cost.

## 1 Introduction

The availability of inexpensive high performance processors and memory chips has made it attractive to use Distributed Computing Systems for real time applications [1]. Distributed Real-Time Systems are characterized by their requirement that the execution of their computational tasks must be not only logically correct but also completed within time [2]. The task allocation in DRTS finds extensive applications in the faculties where large amount of the data is to be processed in relatively short periods of time, or real time computations are required [3]. Fields of research applications are: Meteorology, Cryptography, Image Analysis, Signal Processing, Solar and Radar Surveillance, simulation of VLSI circuits, Industrial Process Monitoring. All these applications require not only very fast computing speeds but different strategies involving DRTS, because in such application the quality of the output is directly proportional to the amount of the real time computations.

Appropriate and efficient load sharing strategies can improve the performance of distributed system. Under load sharing scheme, the workload is to be distributed

S.C. Satapathy et al. (eds.), *ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of CSI - Volume II*, Advances in Intelligent Systems and Computing 249,
DOI: 10.1007/978-3-319-03095-1_2, 

among the nodes available to receive the workload in the system. In this paper static load sharing policy has been considered. In this policy system state information is not required to make load distribution decisions [4]. Once the load sharing decisions has been made, all the tasks should be placed to the appropriate node in the system. Task allocation over the distributed system must be performed in such a manner that the total system cost is minimized. Task allocation in distributed real time systems is an open research area. In task allocation time and space dimension and information requirement for decision making, makes it very complex and difficult. Task allocation can be viewed as a strategic factor which if not managed properly will affect the performance of the system [5]. Task allocation is a NP-complete problem [6][7][8]. Various models for task allocation have been proposed in literature [6][7][9][10][11][12][13][14][8]. Effective clustering technique for tasks of distributed application reduces the allocation search space [8]. In [12][14][6] clustering has been done only once at initial level.

Here in this model, clustering is done at two level, first at initial level on the basis of execution cost and inter task communication cost and second at the stage of task allocation. In the allocation of tasks a more scalable active approach is being followed which iteratively refines the performance of DRTS. The heterogeneity of the system has been considered at the time of allocation. Organization of the paper is as follows. In the next section problem formulation has been done. In the third section the problem has been stated. Fourth section describes the model and its components in detail. In the fifth section, the technique has been described and sixth section provides the algorithm for clustering and allocation of tasks. In next section the model has been simulated using MATLAB7.11.0. Here two examples have been taken to illustrate the proposed model. Results of the simulated experiments are collectively presented in the section 8.

## 2 Problem Formulation

DRTS is a useful platform for huge and complex real-time parallel application. The execution of a parallel application can be seen as the execution of multiple tasks (parallel application divided into number of tasks) over different processors in the system concurrently. The performance of parallel applications on DRTS basically depends on the arrangement of the tasks on various processors available in the system [15]. Methodical resource management is required to properly allocate tasks to achieve the constrained performance. Task allocation should be made in such a way that it can minimize the inter task communication and processor's capabilities must suit to the execution requirements of the task. Proposed model offer an optimal solution by assigning a set of "$m$" tasks of the parallel application to a set of "$n$" processors (*where,* $m > n$) in a DRTS with the goal to enhance the performance of DRTS. The objective of this problem is to enhance the performance of the distributed system by making optimal utilization of its processors and suitable allocation of tasks.

### 2.1 Notations

$T$ : the set of tasks of a parallel program to be executed.
$P$ : the set of processors in distributed system.

$n$ : the number of processors.
$m$: the number of tasks formed by parallel application .
$k$: the number of clusters.
$t_i$ : $i^{th}$ task of the given parallel application.
$P_l$: $l^{th}$ processor in $P$.
$ec_{il}$ : incurred execution cost (EC), if $i^{th}$ task is executed on $l^{th}$ processor.
$cc_{ij}$: incurred inter task communication cost between task $t_i$ and $t_j$ , if they are executed on separate processors.
$X$: an allocation matrix of order $m*n$, where the entry $x_{il} = 1$; *if* $i^{th}$ task is allocated to $l^{th}$ processor and 0; otherwise.
*CI* : cluster information vector.
*ECM ( , )* : execution cost matrix.
*ITCCM ( , )* : inter task communication cost matrix.

## 2.2 Definitions

### 2.2.1 Execution Cost (EC)

The execution cost $ec_{il}$ of a task $t_i$ , running on a processor $P_l$ is the amount of the total cost needed for the execution of $t_i$ on that processor during process execution. If a task is not executable on a particular processor, the corresponding execution cost is taken to be infinite (∞).

### 2.2.2 Communication Cost (CC)

The communication cost ($cc_{ij}$) incurred due to the inter task communication is the amount of total cost needed for exchanging data between $t_i$ and $t_j$ residing at separate processor during the execution process. If two tasks executed on the same processor then $cc_{ij} = 0$.

## 2.3 Assumptions

To allocate the tasks of a parallel program to processors in DRTS, following assumptions has been made:

**2.3.1.** The processors involved in the distributed system are heterogeneous and do not have any particular interconnection structure.

**2.3.2.** The parallel program is assumed to be the collection of $m$- tasks that are free in general, which are to be executed on a set of $n$- processors having different processor attributes.

**2.3.3.** Once the tasks are allocated to the processors they reside on those processors until the execution of the program has completed. At whatever time a cluster of tasks is assigned to the processor, the inter task communication cost (ITCC) between those tasks will be zero.

**2.3.4.** Data points for *k-mean clustering* will be collection of vectors which represents the execution cost of the task $t_m$ on each processor.