

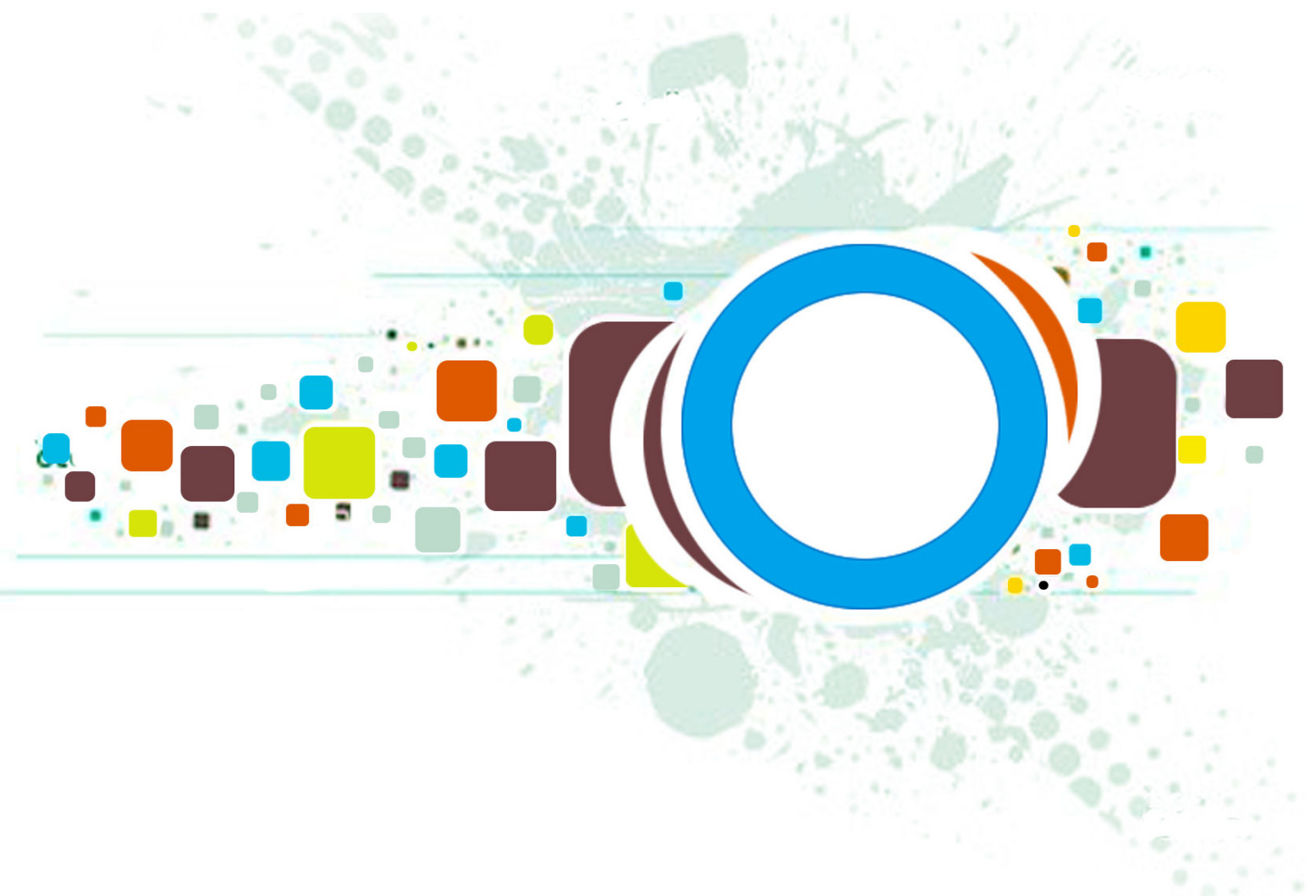
Volume 7 • Issue 3 • June 2013

Editor-in-Chief
Professor Hu, Yu-Chen

INTERNATIONAL JOURNAL OF
IMAGE PROCESSING (IJIP)

ISSN : 1985-2304

Publication Frequency: 6 Issues Per Year



CSC PUBLISHERS
<http://www.cscjournals.org>

Copyrights © 2013 Computer Science Journals. All rights reserved.

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

VOLUME 7, ISSUE 3, 2013

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2304

International Journal of Image Processing (IJIP) is published both in traditional paper form and in Internet. This journal is published at the website <http://www.cscjournals.org>, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJIP Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

Book: Volume 7, Issue 3, June 2013

Publishing Date: 30-06-2013

ISSN (Online): 1985-2304

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

IJIP Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJIP Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2013

EDITORIAL PREFACE

The International Journal of Image Processing (IJIP) is an effective medium for interchange of high quality theoretical and applied research in the Image Processing domain from theoretical research to application development. This is the Third Issue of Volume Seven of IJIP. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJIP emphasizes on efficient and effective image technologies, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing. IJIP comprehensively cover the system, processing and application aspects of image processing. Some of the important topics are architecture of imaging and vision systems, chemical and spectral sensitization, coding and transmission, generation and display, image processing: coding analysis and recognition, photopolymers, visual inspection etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 7, 2013, IJIP appears in more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJIP gives an opportunity to scientists, researchers, engineers and vendors from different disciplines of image processing to share the ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development. This journal is helpful for the researchers and R&D engineers, scientists all those persons who are involve in image processing in any shape.

Highly professional scholars give their efforts, valuable time, expertise and motivation to IJIP as Editorial board members. All submissions are evaluated by the International Editorial Board. The International Editorial Board ensures that significant developments in image processing from around the world are reflected in the IJIP publications.

IJIP editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJIP. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJIP provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

International Journal of Image Processing (IJIP)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Professor Hu, Yu-Chen
Providence University (Taiwan)

ASSOCIATE EDITORS (AEiCs)

Professor. Khan M. Iftakharuddin
University of Memphis
United States of America

Assistant Professor M. Emre Celebi
Louisiana State University in Shreveport
United States of America

Assistant Professor Yufang Tracy Bao
Fayetteville State University
United States of America

Professor. Ryszard S. Choras
University of Technology & Life Sciences
Poland

Professor Yen-Wei Chen
Ritsumeikan University
Japan

Associate Professor Tao Gao
Tianjin University
China

Dr Choi, Hyung Il
Soongsil University
South Korea

EDITORIAL BOARD MEMBERS (EBMs)

Dr C. Saravanan
National Institute of Technology, Durgapur West Benga
India

Dr Ghassan Adnan Hamid Al-Kindi
Sohar University
Oman

Dr Cho Siu Yeung David

Nanyang Technological University
Singapore

Dr. E. Sreenivasa Reddy

Vasireddy Venkatadri Institute of Technology
India

Dr Khalid Mohamed Hosny

Zagazig University
Egypt

Dr Chin-Feng Lee

Chaoyang University of Technology
Taiwan

Professor Santhosh.P.Mathew

Mahatma Gandhi University
India

Dr Hong (Vicky) Zhao

Univ. of Alberta
Canada

Professor Yongping Zhang

Ningbo University of Technology
China

Assistant Professor Humaira Nisar

University Tunku Abdul Rahman
Malaysia

Dr M.Munir Ahamed Rabbani

Qassim University
India

Dr Yanhui Guo

University of Michigan
United States of America

Associate Professor András Hajdu

University of Debrecen
Hungary

Assistant Professor Ahmed Ayoub

Shaqra University
Egypt

Dr Irwan Prasetya Gunawan

Bakrie University
Indonesia

Assistant Professor Concetto Spampinato

University of Catania
Italy

Associate Professor João M.F. Rodrigues

University of the Algarve
Portugal

Dr Anthony Amankwah

University of Witswatersrand
South Africa

Dr Chuan Qin

University of Shanghai for Science and Technology
China

AssociateProfessor Vania Vieira Estrela

Fluminense Federal University (Universidade Federal Fluminense-UFF)
Brazil

Dr Zayde Alcicek

firat university
Turkey

Dr Irwan Prasetya Gunawan

Bakrie University
Indonesia

TABLE OF CONTENTS

Volume 7, Issue 3, June 2013

Pages

- | | |
|-----------|--|
| 227 - 236 | Skin Cancer Prognosis Based Pigment Processing <i>Muthana H. Hamd, Kadhum A. Essa</i> |
| 237 - 247 | Content Based Video Retrieval in Transformed Domain using Fractional Coefficients <i>H.B.Kekre, Sudeep D. Thepade, Saurabh Gupta</i> |
| 248 - 265 | Reversible Data Hiding Using Contrast Enhancement Approach <i>Amjed S. Al-Fahoum, Mohammed Yaser</i> |
| 266 - 277 | Satellite Imaging System <i>AA Somaie</i> |
| 278 - 285 | Enhanced Spectral Reflectance Reconstruction Using Pseudo-Inverse Estimation Method <i>Ibrahim Anis El-Rifai, Hend Mahgoub, Mennat-Allah Magdy, Jay Arre Toque, Ari Ide-Ektessabi</i> |
| 286 - 301 | A Novel Multiple-kernel based Fuzzy c-means Algorithm with Spatial Information for Medical Image Segmentation <i>Venu Nookala, B.Anuradha</i> |
| 302 - 313 | Local Phase Oriented Structure Tensor To Segment Texture Images With Intensity Inhomogeneity <i>Hiren K Mewada, Suprava Patnaik</i> |

Skin Cancer Prognosis Based Pigment Processing

Muthana H. Hamd

Computer & SW Engineering
Al Mustansiriya
Baghdad, Iraq

muth700@yahoo.com

Kadhum A. Essa

Computer & SW Engineering
Al Mustansiriya
Baghdad, Iraq

kadhum2@yahoo.com

Abstract

This paper develops a new computerized vision of skin cancer prognosis based on symmetry and color matching for lesion pigments. Initially, the lesion/tumor edge is detected and segmented. Then, the symmetrization is computed for all images to isolate benign (mole) tumor. The even symmetry parameter is introduced here to improve the symmetrization computations. The suspicious images would be nominated into one of three categories: melanoma, Basal Cell Carcinoma (BCC), or Squamous Cell Carcinoma (SCC) tumor depending on the symmetrization and pigment-color matching score table. Two matching procedures have been developed for nominating the suspicious images. First procedure matches pigment values with artificial spectrums of Reddish, Yellowish, Brownish, and Blackish. The second procedure matches pigment values with true malignancy/benign pigment database. The results of two procedures are compared over 40 pre-classified images. With Mean Squared Error (MSE) value equals to 0.003, procedure#1 satisfied 80% true classification while 92.5% for procedure#2. These results could be improved if lesion segmentation and/or spectrums/pigment-database are increased.

Keywords: Melanoma, SCC, BCC, Skin Cancer, Color Matching.

1. INTRODUCTION

Skin cancer is mainly divided into two types: melanoma or non-melanoma. The non-melanoma tumors have two sub-types: Squamous Cell Carcinoma (SCC or SqCC) and Basal Cell Carcinoma (BCC). So, computer-aided diagnosis software is essential in this case to assist the early detection of the malignant tumors. Skin as a large organ of the body it consists of three layers: Outer epidermis layer; dermis layer; and deep sub-cutis layer. The melanocyte which is located in the epidermis, synthesizes melanin that determines the pigment of the skin (skin color) [2, 3]. Studies results found that melanoma is responsible for 80% of skin cancer deaths and cases of non-melanoma skin cancer have been raised an average of 4.2% a year. The clinical criteria that pathologist applies them for tissue diagnosis includes: the type of cell that is proliferating, its histological grade, genetic abnormalities, and other features of the tumor relating to its pigments. Together, this information is useful to evaluate the prognosis of the patient and to choose the best treatment [4, 5, and 6]. The ABCD is a traditional diagnostic rule of dermatoscopy which is made by Stolz's [7], it is based on the four main criteria or lesion parameters: Asymmetry, Border, Colour and Diameter. The "E" in Elevation is added later to describe the uneven surface of tumour. ***This rule applies a semi-quantitative score system to make a decision*** [8, 9]. The ABCD method is improved by "ELM 7"; which is a computerized scanning method that based on polarized light surface microscopes. G. Di Leo, A. Paolillo, et al., developed "ELM 7 point checklist" diagnostic method in 2010. It provides same accuracy of traditional "ABCD" criteria where, it defines a set of seven features, based on colour and texture parameters that describe the lesion malignancy [10]. The comprehensive research that is made

by [11] in 2009, provides clear view about such tools that are mostly intended for assisting and supporting the decision systems at early stages. If a comparison is made between these tools and the medical experts in the field, even with the best diagnostic results, the system depicts relatively lower performance in terms of accuracy and confidence. However, the physicians admitted that the computerized diagnostic tools are very useful in producing quantified results, recording patient follow-ups, and monitoring the therapeutic and healing progress. These tools would not be in any case, used for replacing the physicians, but just to serve as early diagnostic adjuncts.

This work develops a new computerized vision of skin cancer diagnosis. It adopts spatial processing for pigment-color matching and symmetry computations to detect early the malignant tumor. This paper is organized as follows: the next section II, presents the three types of skin cancer and their signs; symptoms; and lesion pigments. Section III illustrates the lesion boundary detection (segmentation). Section IV and V illustrate the two matching procedures which they have been developed for approximating the lesion color values (pigments). One procedure uses color spectrums for matching operations, while second procedure applies pigment database for that matching. Section VI is a conclusion summarization for the two procedures results and evaluation that have been tabulated in eight tables.

2. SKIN CANCER TYPES

Skin cancer can be defined as an abnormal growth up of skin cells due to defective in the DNA. Oftenly, not necessarily it may be happened because of the sunshine, tanning beds, or genetic defective. This section presents three kinds of malignant tumor.

2.1 Melanoma

Melanoma is the least common type of skin cancer, but it is the most deadly one. It can be quickly spread to other body parts causing a secondary cancer. The signs and symptoms caused by this most dangerous disease are [12,13]:

- Appears as a mole, dark spot, or freckle anywhere on the body with changing in shape, color, and size
- The border is smudgy (blurred) and irregular
- Lesion pigment has more than one color like: red, brown, black, white and/or light grey

The most three common diagnostic rules are:

- ABCDE" which is based on semi-quantitative analysis
- The ELM 7-point checklist scoring diagnosis analysis which defining only seven standard ELM criteria.
- Pattern analysis, which is based on the "expert" qualitative assessment of numerous individual ELM criteria;

2.2 Squamous Cell Carcinoma

SCC is the second most common form of skin cancer. it is mainly caused by cumulative ultraviolet exposure over long time. It is less risk than melanoma, but more dangerous than BCC. The SCC signs and symptoms are:

- Typically occur everywhere on the body including the mucous membranes and genitals. The exposing body parts like ear, neck, arm, etc are the most common affected areas.

- It looks like tender scaly, scaly red patches, open sores or warts, ulcerated lump they may be crust or bleed easily.
- Usually it presents with one or more dry or crusted red or brown patches.

In 2002 almost all people with SCC were aged 40 years and over 1138,000 new cases of SCC were estimated to have been diagnosed in 2008 [13].

2.3 Basal Cell Carcinoma

BCC is by far the most common form of skin cancer. It grows from cells in the lower part of the upper layer of the skin, taking a period of months to years. The BCC signs and symptoms are:

- BCC is very difficult to recognized, only a specialist in diseases of the skin, can decide for sure. A persistent, non-healing sore is a very common sign of an early BCC.
- An open sore, reddish patch, pink growth, scars, or irritated area that commonly occurring on the face, chest, shoulders, arms, or legs.
- A shiny bump or nodule which is pearly and is often pink, red, or white. It can be confused with mole because of the rolled border. The bump can also be tan, black, or brown.
- A scar-like area that is white, yellow or waxy with poorly defined borders; the skin itself appears shiny and taut.

In 2002, 96% of people with BCC were aged 40 years or older. 1296,000 new cases of BCC were estimated to have been diagnosed in 2008. BCC is more easily treated in its early stages, but if it lefts untreated it can grows, erodes and destroys adjoining structures [12, 13, and14].

3. LESION BOUNDARY DETECTION

Lesion boundary can be detected and extracted by finding the orthogonal and perpendicular diagonal line values. These intersected values are used to find the symmetry, even symmetry, semi-symmetry, or asymmetry of lesion [15, 16, and 17]. See Fig. 1.

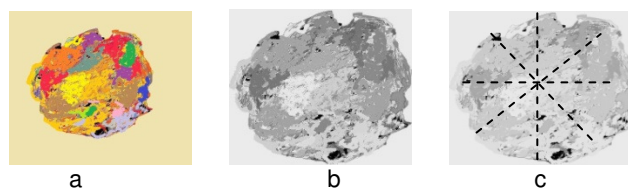


FIGURE 1: a) Example of Colored Lesion Image, b) Lesion Boundary Detection, c) Symmetry/Asymmetry Computations.

4. MATCHING PROCEDURE#1: PIGMENT vs. SPECTRUMS

This procedure attempts finding a true color approximation for all sub-image lesions, where the four new sub-images are segmented from the original one with size $(m/2 \times n/2)$: where, m & n is the row and column respectively). Fig. 1 explains clearly the segmented operations. Matching operation starts finding the minimum Mean Squared Error (MSE) between sub-image pigments and the 140 artificial spectrums of reddish, brownish, yellowish, and blackish, see Fig. 2. The results of this procedure would be stored in image profile that contains score of matching events. The MSE between pigments and the artificial colors could be calculated as follows:

$$MSE = (m - \mu)^2$$

(1)

The necessary procedure steps for matching operations can be detailed below:

Procedure #1 (suspicious pigments, spectrums)

for all suspicious images

for all artificial colors

error = (mean (sub-image pigment) - spectrum value)²

if error ≤ threshold value (□)

matching is true; event=event+1

else matching is false

end if;

next spectrum; next suspicious image

end for; end for

Table 1, 2, 3, and 4 represent procedure#1 matching results that have been applied on 40 suspicious images. Those images were already classified, and the procedure is just succeeded reclassifying benign (80%), melanoma (90%), BCC (60%), and SCC (90%) with overall succeeded predicting equals to 80%. The MSE or threshold value (□) is chosen to be 0.003. Procedure#1 segments the suspicious image into four sub-images. So, each quarter tries matching it pigment mean value with 120 spectrum values. That means the total number of operations are:

4 (no. of quarters) × 40 (no. of images) × 120 (no. of spectrum) = 19,200 matching operations

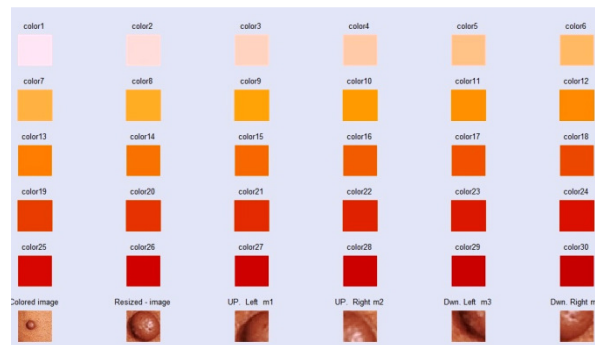


FIGURE 2.: Example of Matching Operations between Benign Image Pigment and Reddish Spectrums.

5. MATCHING PROCEDURE#2: PIGMENT vs. PIGMENT DATABASE

This procedure is similar to the first one except it matches the suspicious sub-image pigments with malignant pigment database. So, the procedure#2 necessary steps would be:

Procedure #2 (suspicious pigments, pigment database)

for all suspicious images

for all malignant image database

error = (mean (sub-image pigment) - malignant pigment)²

if error ≤ threshold value (□)

matching is true; event=event+1

else matching is false

end if;

next malignant image; next suspicious image

end for; end for

Table 5, 6, 7 and 8 represent procedure#2 matching results that have been applied on the same 40 images. According to the results of these tables, procedure#2 succeeded re-classifying 90% of

benign images, melanoma (90%), BCC (90%), and SCC (100%). The same MSE value (0.003) is considered as a threshold value. The overall predicting result would be 92.5%. With this procedure, the suspicious image is segmented into four sub-images so, 4 quarters goes to full, 3 quarters, half, and quarter. Clearly, from two procedure results, procedure#2 satisfies better matching operations because of the true color values (lesion pigments) that have been used for matching process. Fig. 3 is an example of procedure#2 pigment matching operation.



FIGURE 3.a: Example of Procedure#2 Matching Operation, MSE Equals to 0.003, No Matching Event.

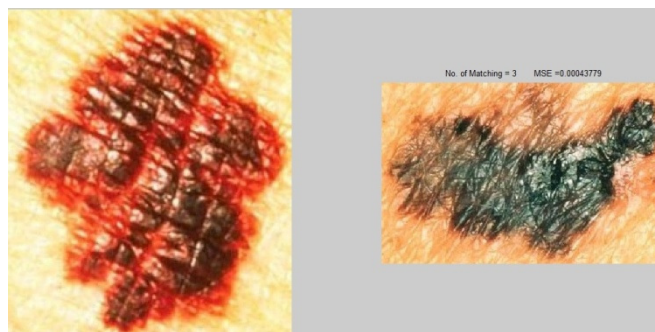


FIGURE 3.b: Example of Procedure#2 Matching Operation with MSE Equals to 0.003 and 3 Quarters Matching Events.



FIGURE 3.c: Example of Procedure#2 Matching Operation with MSE Equals to 0.003 and 4 Quarters (full) Matching Events.

Procedure#2 segments the suspicious image into 16 sub-images and only the interested lesion area is considered (the inner 4 sub-images). So, each quarter tries matching itself mostly once with pigment database. That means the no. of matching operations for each kind of malignant or benign image are less than:

4 (no. of quarters) \times 10 (no. of suspicious images) \times 40 (no. of sub-image database) = 1, 600 matching operations

| Image No. | Symmetrisation | | Matching spectrums | | | | Predicted results |
|-----------|----------------|---------------|--------------------|--------|---------|-------|-------------------|
| | Symmetry | Even Symmetry | Red | Yellow | Brown | Black | |
| Ben-1 | semi | yes | √,√ | - | √ | - | -ve |
| Ben-2 | semi | yes | √ | - | - | √ | -ve |
| Ben-3 | semi | yes | √ | - | - | √ | -ve |
| Ben-4 | semi | yes | - | - | - | - | -ve |
| Ben-5 | no | yes | √ | - | √ | √ | +ve |
| Ben-6 | semi | yes | - | - | - | - | -ve |
| Ben-7 | yes | yes | √, √ | - | √, √, √ | - | -ve |
| Ben-8 | no | yes | √ | - | √ | - | -ve |
| Ben-9 | no | yes | √, √ | - | √,√ | √, √ | +ve |
| Ben10 | yes | yes | √, √ | - | - | √, √ | -ve |

TABLE 1: Procedure#1 Benign Matching Results.

| Image No. | Symmetrisation | | Matching spectrums | | | | Predicted results |
|-----------|----------------|---------------|--------------------|--------|-------|-------|-------------------|
| | Symmetry | Even Symmetry | Red | Yellow | Brown | Black | |
| Mel-1 | no | yes | √,√,√ | √ | - | √ | +ve |
| Mel-2 | no | no | √,√ | - | √ | - | +ve |
| Mel-3 | no | no | - | √,√ | - | √ | +ve |
| Mel-4 | no | no | √ | √ | √,√,√ | √ | +ve |
| Mel-5 | no | no | √ | √,√ | - | √,√ | +ve |
| Mel-6 | no | no | - | - | √ | - | -ve |
| Mel-7 | no | no | √ | √ | √,√ | - | +ve |
| Mel-8 | no | no | √ | √,√ | - | - | +ve |
| Mel-9 | semi | no | √ | √ | √,√ | √,√ | +ve |
| Mel 10 | no | no | √,√ | √ | √,√ | - | +ve |

TABLE 2: Procedure#1 Melanoma Matching Results.

| Image No. | Symmetrisation | | Matching spectrums | | | | Predicted results |
|-----------|----------------|---------------|--------------------|--------|-------|-------|-------------------|
| | Symmetry | Even Symmetry | Red | Yellow | Brown | Black | |
| BCC1 | no | no | - | - | √,√ | - | suspicious |
| BCC2 | no | no | √,√,√ | - | - | √,√ | +ve |
| BCC3 | no | no | √ | - | - | √ | +ve |
| BCC4 | no | no | √ | - | √ | √ | +ve |
| BCC5 | no | no | √ | - | - | - | suspicious |
| BCC6 | yes | no | - | - | √,√ | √ | suspicious |
| BCC7 | no | no | - | - | √ | - | suspicious |
| BCC8 | no | no | √ | - | √,√ | √,√ | +ve |
| BCC9 | no | no | √ | - | √ | - | +ve |
| BCC10 | no | no | √,√,√ | - | √,√ | √ | +ve |

TABLE 3: Procedure#1 BCC Matching Results.

| Image No. | Symmetrisation | | Matching spectrums | | | | Predicted results |
|-----------|----------------|---------------|--------------------|--------|-------|-------|-------------------|
| | Symmetry | Even Symmetry | Red | Yellow | Brown | Black | |
| SCC1 | no | no | - | √ | √ | √,√ | +ve |
| SCC2 | no | no | √,√ | √,√ | √ | √ | +ve |
| SCC3 | no | no | √ | - | - | √ | +ve |
| SCC4 | no | no | √ | √,√ | √,√ | √,√ | +ve |
| SCC5 | no | no | - | √,√ | - | √,√,√ | +ve |
| SCC6 | no | no | √ | √ | - | √ | +ve |
| SCC7 | no | yes | √ | - | √ | - | suspicious |
| SCC8 | no | no | - | √,√ | √,√,√ | √,√ | +ve |
| SCC9 | no | no | - | √ | √ | √ | +ve |
| SCC10 | no | no | √,√ | √,√ | √ | √ | +ve |

TABLE 4: Procedure#1 SCC Matching Results.

| Image No. | Symmetrisation | | Matching Database (pigments) | | | | Predicted results |
|-----------|----------------|---------------|------------------------------|-----------|------|---------|-------------------|
| | Symmetry | Even Symmetry | Full | 3quarters | half | quarter | |
| Ben-1 | semi | yes | √√ | √√ | √ | √√ | +ve |
| Ben-2 | semi | yes | √√√√ | √ | √ | √√ | +ve |
| Ben-3 | semi | yes | √√√√√ | √ | - | √ | +ve |
| Ben-4 | semi | yes | √√ | √√ | - | √ | +ve |
| Ben-5 | no | yes | √√√ | √√ | √ | √√ | +ve |
| Ben-6 | semi | yes | √√√ | - | √√ | √√ | +ve |
| Ben-7 | yes | yes | √ | √ | √√√ | √ | Suspicious |
| Ben-8 | no | yes | √√√√ | - | √ | - | +ve |
| Ben-9 | no | yes | √√ | √√√ | √√√ | √√ | +ve |
| Ben10 | yes | yes | √√√ | √√ | √√ | √ | +ve |

TABLE 5: Procedure#2 Benign Matching Results.

| Image No. | Symmetrisation | | Matching Database (pigments) | | | | Predicted results |
|-----------|----------------|---------------|------------------------------|-----------|--------|---------|-------------------|
| | Symmetry | Even Symmetry | Full | 3quarters | half | quarter | |
| Mel-1 | no | yes | √ | - | √√ | √√√√√* | -ve |
| Mel-2 | no | no | √√ | - | √√√√√* | - | -ve |
| Mel-3 | no | no | √ | - | √ | - | +ve |
| Mel-4 | no | no | √√√√ | √√√√√ | - | - | -ve |
| Mel-5 | no | no | √√√√√* | - | √√ | - | -ve |
| Mel-6 | no | no | √√√ | √√ | √ | √√√ | -ve |
| Mel-7 | no | no | √√√√√ | √√√√ | - | - | -ve |
| Mel-8 | no | no | √√√√ | √√√√ | √ | - | -ve |
| Mel-9 | semi | no | √√√√ | √ | √√ | √ | -ve |
| Mel 10 | no | no | √√√√√ | √√ | √ | √ | -ve |

TABLE 6: Procedure#2 Melanoma Matching Results.

| Image No. | Symmetrisation | | Matching Database (pigments) | | | | Predicted results |
|-----------|----------------|---------------|------------------------------|-----------|------|---------|-------------------|
| | Symmetry | Even Symmetry | Full | 3quarters | half | quarter | |
| BCC1 | no | no | √√√ | √√√√ | √ | - | +ve |
| BCC2 | no | no | √√√√* | - | √√ | - | +ve |
| BCC3 | no | no | √√ | √√√√√ | √ | √ | +ve |
| BCC4 | no | no | √√ | √√√√ | √√ | - | +ve |
| BCC5 | no | no | √√√ | - | - | - | +ve |
| BCC6 | yes | no | √√√ | √√√ | - | √√√ | +ve |
| BCC7 | no | no | √ | - | - | √ | -ve |
| BCC8 | no | no | √√√√ | √√√ | √√ | √ | +ve |
| BCC9 | no | no | √√ | √√√√√ | - | - | +ve |
| BCC10 | no | no | √√√√√* | √√ | - | - | +ve |

TABLE 7: Procedure#2 BCC Matching Results.

| Image No. | Symmetrisation | | Matching Database (pigments) | | | | Predicted results |
|-----------|----------------|---------------|------------------------------|-----------|------|---------|-------------------|
| | Symmetry | Even Symmetry | Full | 3quarters | half | quarter | |
| SCC1 | no | no | √√√ | √√√√√ | √ | - | +ve |
| SCC2 | no | no | √√ | - | √√√ | √√ | +ve |
| SCC3 | no | no | √ | √√√ | √√ | √√ | +ve |
| SCC4 | no | no | √√ | √ | √√√ | - | +ve |
| SCC5 | no | no | √√ | √√√√ | √√ | √ | +ve |
| SCC6 | no | no | √√√√√ | √√√ | √√ | - | +ve |
| SCC7 | no | yes | √√√√ | √ | √√ | √ | +ve |
| SCC8 | no | no | √ | √ | √ | √√√√ | +ve |
| SCC9 | no | no | √√√√√ | √√ | √ | √ | +ve |
| SCC10 | no | no | √√√√√ | √√ | √ | √ | +ve |

TABLE 8: Procedure#2 SCC Matching Results.

6. CONCLUSION

This research develops new computerized vision for early detection of most three dangerous kind of skin cancer where, the spatial processing of lesion pigment is applied to find the features like “A” in Asymmetric and “C” in Color instead of using mathematical rules of the “ABCD” which is relatively more complicated comparing with this approach. The even symmetry has been added in symmetry computations as an extra parameter to improve the calculations. The prognosis approach depends completely on score table which is completely depending on the constructed image profile. The precise evaluation of this work requires comparing it with “ABCD” and/or “ELM-7” by running these three methods on the same suspicious images and comparing the results. However, such comparison is planning to take a place in future research, especially when the three color planes (red, green, blue) of lesion pigment are considered for pigment matching operations instead of the equivalent gray-scaled value. For this work, the two matching procedures have been implemented to re-classify 40 suspicious images. These 40 images are already classified into mole/benign, melanoma, BCC, and SCC image. Procedure#1 uses 120 spectrum of reddish, yellowish, brownish, and blackish to match lesion pigment with these spectrums. While procedure#2 apply pigment vs. pigment as it uses the stored database for matching operations. For MSE value equals to 0.003, procedure#1 succeeded re-classifying 80% of suspicious images with 19,200 matching operations, while procedure#2 improved this result when it re-classified 92.5% of them with matching operations were less than 6,400.

7. REFERENCES

- [1] M. H. Hamd and L. M. Mohammad. “Skin Cancer Prognosis Based Color Matching and Segmentation of Pigmented Skin Lesion.” *journal of engineering and technology*, submitted, 2012

- [2] P. Shetty. "Melanoma Decision Support System for Dermatologist," International Conference on Recent Trends in Information Technology and Computer Science (IRCTITCS),2011, pp.101-105.
- [3] L. DANIEL. "American Family Physician." volume 70, No. 8, October 15 2004, www.aafp.org/aafp.
- [4] S. L. Phung and A. Bouzerdoun. "Skin segmentation using color pixel classification: analysis and comparison." *IEEE Transactions on Pattern Analysis and Machine Intelligence* , Vol. 27, No. 1, Jan. 2005.
- [5] R. G. Gonzalez and E. R. Woods, "Digital Image Processing Using MATLAB." A John Wiley & Sons, Inc., Publication 4th Edition , 2004, pp. 220-222.
- [6] G. Blanchet and C Maurice "Digital Signal and Image Processing using MATLAB.", ISBN 10: 1-905209-13-4, 2006, pp. 43-46.
- [7] M. Thorsten. "Functional Infrared Imaging for Skin-Cancer Screenin," Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.
- [8] W. Stolz, D. Hölzel, A. Riemann. "Multivariate analysis of criteria given by dermatoscopy for recognition of melanocytic lesions." Book of Abstracts, Fiftieth Meeting of the American Academy of Dermatology, Dallas, Texas, December 1991.
- [9] R. A. Fiorini, G. Dacquino and G. Laguteta. "A New Melanoma Diagnosis Active Support System," Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, September 1-5, 2004, 3206-3209.
- [10] G. Di. Leo and A. P. Sommella. "Automatic Diagnosis of Melanoma: a Software System based on the 7-Point Check-List," Proceedings of the 43rd Hawaii International Conference on System Sciences, 2010.
- [11] M. Ilias and C. N. Doukas, "Overview of Advanced Computer Vision Systems for Skin Lesions Characterization.", *IEEE Transaction on Information Technology in Biomedicine*, Vol. 13, NO. 5, September 2009.
- [12] M. M. Rahman and B. C. Desai , "Image Retrieval-Based Decision Support System for Dermatoscopic Images," Proceedings of the 19th IEEE, Symposium on Computer-Based Medical Systems, 0-7695-2517-1/06, 2006.
- [13] C. Fatichah, "Skin Lesion Detection using Fuzzy Region Growing and ABCD Feature Extraction for Melanoma Skin Cancer Diagnosis." *Journal of computing and informatics technology*, 2010, www.cs.ui.ac.id/files/icacsis2009/pdf/43.pdf.
- [14] X. Yuan. "SVM-based Texture Classification and Application to Early Melanoma Detection," Proceedings of the 28th IEEE EMBS Annual, International Conference New York City, USA, Aug 30-Sept 3, 2006.
- [15] Y. Zhou and M. Smith. "Segmentation of Clinical Lesion Images Using Normalized Cut ", IEEE conference, 978-1-4244-3610-1/09/, 2009.
- [16] R. Cláudio and R. Jung. "Sharpening Dermatological Color Images in the Wavelet Domain." *IEEE journal of Selected Topics in Signal Processing*, Vol. 3, No.1, February 2009.

- [17] M. Sadeghi. "Automated Detection and Analysis of Dermoscopic Structures on Dermoscopy Images," Congress on Image and Signal Processing, 2008.
- [18] H. Yu-Chen. International Journal of Image Processing, Book, 2009, Vol. 3, Issue 5, ISSN(online):1985-2304, pp.229-245.

Content Based Video Retrieval in Transformed Domain using Fractional Coefficients

Dr. H. B. Kekre

hbkekre@yahoo.com

Senior Professor Computer Engineering Department
Mukesh Patel School of Technology, Management & Engineering
NMIMS University
Mumbai, India

Dr. Sudeep D. Thepade

sudeepthepade@gmail.com

Professor & Dean (R&D),
Pimpri Chinchwad College of Engineering,
University of Pune,
Pune, India

Saurabh Gupta

saurabh.gupta761@gmail.com

M.Tech (Comp. Engg.) Student
Mukesh Patel School of Technology, Management & Engineering
NMIMS University
Mumbai, India

Abstract

With the development of multimedia and growing database there is huge demand of video retrieval systems. Due to this, there is a shift from text based retrieval systems to content based retrieval systems. Selection of extracted features play an important role in content based video retrieval. Good features selection also allows the time and space costs of the retrieval process to be reduced. Different methods[1,2,3] have been proposed to develop video retrievals systems to achieve better performance in terms of accuracy.

The proposed technique uses transforms to extract the features. The used transforms are Discrete Cosine, Walsh, Haar, Kekre, Discrete Sine, Slant and Discrete Hartley transforms. The benefit of energy compaction of transforms in higher coefficients is taken to reduce the feature vector size by taking fractional coefficients[5] of transformed frames of video. Smaller feature vector size results in less time for comparison of feature vectors resulting in faster retrieval of images. The feature vectors are extracted and coefficients sets are considered as feature vectors (100%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012%, 0.006% and 0.003% of complete transformed coefficients). The database consists of 500 videos spread across 10 categories.

Keywords: Key Frame, Feature Extraction, Similarity Measures, Orthogonal Transforms.

1. INTRODUCTION

The amount of video content being uploaded to the internet is increasing day by day. Hence, extracting specific videos from the massive amount of multimedia data has been the focus of attention over the past few years. The traditional text-based search has drawbacks like manual annotations are time consuming and costly to implement [1]. With the increase in the number of media in a database, the complexities in determining the required information also increases.

1.1 Content Based Video Retrieval (CBVR)

Content-Based Video Retrieval System (CBVR) is defined as the search which will retrieve video from database based on contents. Content relates to color, shapes, textures, or any other information that can be obtained from the video directly.

CBVR mainly contains three stages. Firstly key frames are extracted from video. Key-frames are still images extracted from original video data that best represent the content of video [9]. Secondly, feature extraction (FE) is done, where a set of features, called feature vector, is generated to accurately represent the content of each video in the database. After this, similarity measurement is done where a distance between the query video and each video in the database using their feature vectors is used to retrieve the “closest” videos. Features such as Motion features [2,3], Color features and edge using histograms [2] and DCT transforms[3]. Video Retrieval Based on Textual Queries [10] presented an approach that enables search based on the textual information present in the video. Regions of textual information are indented within the frames of the video.

Here transforms are used to extract features as they give high energy compaction in transformed domain, hence frames from video in transformed domain are used for feature extraction in CBVR. The energy compaction is in few elements, so large number of the coefficients of transformed image can be neglected to reduce the size of feature vector [4].

In this paper, Discrete Cosine, Walsh, Haar, Kekre, Discrete Sine, Slant and Discrete Hartley transforms are used with reduced size feature vector using fractional coefficients of transformed frames. Mean Squared Error(MSE) is used as similarity measure as shown in equation(1.1)

$$MSE(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad \dots\dots\dots(1.1)$$

2. ORTHOGONAL TRANSFORMS

2.1 DCT(Discrete Cosine Transform) [6]

The NxN cosine transform matrix $C=\{c(k,n)\}$, also called the Discrete Cosine Transform(DCT), is defined as

$$c(k,n) = \begin{cases} \frac{1}{\sqrt{N}} & k=0, 0 \leq n \leq N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{\Pi(2n+1)k}{2N} & 1 \leq k \leq N-1, 0 \leq n \leq N-1 \end{cases} \quad \dots\dots\dots(2.1.1)$$

The one-dimensional DCT of a sequence $\{u(n), 0 \leq n \leq N-1\}$ is defined as

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos \left[\frac{\Pi(2n+1)k}{2N} \right] \quad 0 \leq k \leq N-1 \quad \dots\dots\dots(2.1.2)$$

Where,

$$\alpha(0) = \frac{1}{\sqrt{N}}, \alpha(k) = \sqrt{\frac{2}{N}} \text{ for } 1 \leq k \leq N-1$$

The inverse transformation is given by

$$u(n) = \sum_{k=0}^{N-1} \alpha(k) v(k) \cos \left[\frac{\Pi(2n+1)k}{2N} \right], 0 \leq n \leq N-1 \quad \dots\dots\dots(2.1.3)$$

2.2 DST(Discrete Sine Transform)

The NxN sine transform matrix $\Psi = \{\Psi(k, n)\}$, also called the Discrete Sine Transform(DST), is defined as

$$\Psi(k, n) = \sqrt{\frac{2}{N+1}} \sin \frac{\Pi(k+1)(n+1)}{N+1} \quad 0 \leq k, n \leq N-1 \quad \dots\dots\dots(2.2.1)$$

The sine transform pair of one-dimensional sequences is defined as

$$v(k) = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} u(n) \sin \frac{\Pi(k+1)(n+1)}{N+1} \quad 0 \leq k, n \leq N-1 \quad \dots\dots\dots (2.2.2)$$

The inverse transformation is given by

$$u(n) = \sqrt{\frac{2}{N+1}} \sum_{k=0}^{N-1} v(k) \sin \frac{\Pi(k+1)(n+1)}{N+1} \quad 0 \leq n \leq N-1 \quad \dots\dots\dots(2.2.3)$$

2.3 Haar Transform

The Haar wavelet's mother wavelet function $\varphi(t)$ can be described as:

$$\varphi(t) = \begin{cases} 1, & 0 \leq t \leq \frac{1}{2} \\ -1, & \frac{1}{2} \leq t \leq 1 \\ 0, & \text{Otherwise} \end{cases} \quad \dots\dots\dots (2.3.1)$$

And its scaling function $\varphi(t)$ can be described as,

$$\varphi(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 0, & \text{Otherwise} \end{cases} \quad \dots\dots\dots(2.3.2)$$

2.4 Walsh Transform

Walsh transform matrix is defined as a set of N rows, denoted W_j , for $j = 0, 1, \dots, N-1$, which have the following properties:

- W_j takes on the values +1 and -1.
- $W_j[0] = 1$ for all j .
- $W_j \times W_{KT} = 0$, for $j \neq k$ and $W_j \times W_{KT}$ W_j has exactly j zero crossings, for $j = 0, 1, \dots, N-1$.
- Each row W_j is even or odd with respect to its midpoint.

Walsh transform matrix is defined using a Hadamard matrix of order N. The Walsh transform matrix row is the row of the Hadamard matrix specified by the Walsh code index, which must be an integer in the range [0... N-1]. For the Walsh code index equal to an integer j , the respective Hadamard output code has exactly j zero crossings, for $j = 0, 1, \dots, N-1$.

2.5 Kekre Transform

Kekre Transform matrix can be of any size NxN, which need not have to be in powers of 2. All upper diagonal and diagonal values of Kekre's transform matrix are one, while the lower diagonal part

except the values just below diagonal is zero. Generalized NxN Kekre Transform matrix can be given as

$$K_{NxN} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ -N+1 & 1 & 1 & \dots & 1 & 1 \\ 0 & -N+2 & 1 & \dots & : & : \\ : & 0 & \dots & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & -N+(N-1) & 1 \end{bmatrix} \dots\dots\dots(2.5.1)$$

The formula for generating the term K_{xy} of Kekre Transform matrix is

$$K_{x,y} = \begin{cases} 1, & x \leq y \\ -N+(x+1), & x = y+1 \\ 0, & x > y+1 \end{cases} \dots\dots\dots(2.5.2)$$

2.6 Hartley Transform [7]

The Discrete Cosine Transform(DCT) utilizes cosine basis functions, while Discrete Sine Transform(DST) uses sine basis function. The Hartley transform utilizes both sine and cosine basis functions. The discrete 2-dimensional Hartley Transform is defined as,

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \text{Cas} \left\{ \frac{2\pi}{N} (ux + vy) \right\} \dots\dots\dots(2.6.1)$$

Inverse discrete 2-dimensional Hartley Transform is,

$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) \text{Cas} \left\{ \frac{2\pi}{N} (ux + vy) \right\} \dots\dots\dots(2.6.2)$$

where, $\text{Cas} \theta = \cos \theta + \sin \theta$

2.7 Slant Transform [8]

The Slant transform is a member of the orthogonal transforms. It has a constant function for the first row, and has a second row which is a linear (slant) function of the column index. The matrices are formed by an iterative construction that exhibits the matrices as products of sparse matrices, which in turn leads to a fast transform algorithm.

The Slant transform matrix of order two is given by

$$S_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \dots\dots\dots(2.7.1)$$

$$S_N = \frac{1}{2^{1/2}} \begin{bmatrix} 1 & 0 & \vdots & \vdots & 1 & 0 & \vdots & \vdots \\ a_N & b_N & \vdots & 0 & -a_N & b_N & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & 0 & \vdots & I_{(n/2)-2} & & 0 & \vdots & I_{(n/2)-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 1 & \vdots & \vdots & 0 & -1 & \vdots & \vdots \\ -b_N & a_N & \vdots & 0 & b_N & a_N & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & 0 & \vdots & I_{(n/2)-2} & & 0 & \vdots & -I_{(n/2)-2} \end{bmatrix} \begin{bmatrix} S_{N/2} & \vdots & 0 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \vdots & S_{N/2} \end{bmatrix}$$

The matrix $I_{(n/2)-2}$ is the identity matrix of dimension $(N/2)-2$. The constants a_N , b_N may be computed by the formula

$$a_{2N} = \left(\frac{3 N^2}{4 N^2 - 1} \right)^{(1/2)}, \quad b_{2N} = \left(\frac{N^2 - 1}{4 N^2 - 1} \right)^{1/2} \quad \dots\dots\dots(2.7.2)$$

3. FRACTIONAL ENERGY

3.1 Feature Vector Database 'Fractional T-RGB' [4,5]



Frame of Size (NxN)



'T' transformed frame of Size (NxN)

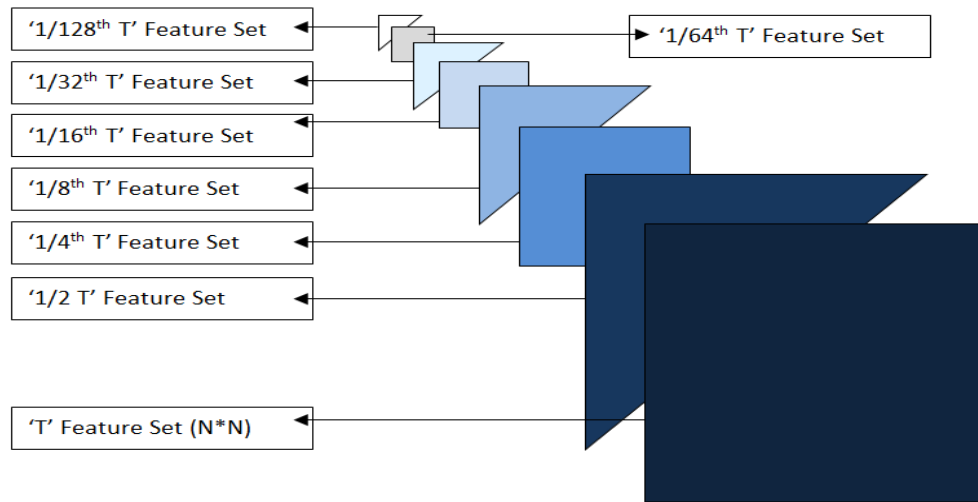


FIGURE 1: Proposed CBVR Techniques using fractional Coefficients of Transformed Frames.

The fractional coefficients of transformed frame as shown in Figure 1, are considered to form 'fractional T' feature vector databases. Here first 50% of coefficients from upper triangular part of feature vector 'T' are considered to prepare the feature vector database '50%' for every frame. Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT feature databases are used to obtain new feature vector databases as 50%-DCT, 50%-Walsh, 50%-Haar, 50%-Kekre, 50%-DST, 50%-Slant, 50%-DHT respectively. Then for each frame in the database, fractional feature vector set for DCT, Walsh, Haar, Kekre, DST, Slant, DHT using 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012%, 0.006% and 0.003% of total coefficients are formed.

4. PROPOSED CBVR TECHNIQUES

4.1 Extraction of Key Frames

A Video is read from database. Then, every 10th frame of each video is stored in database as a key frame in RGB color space. Up to 100th frame of every video is stored in the database.

4.2 Feature Extraction for feature vector 'Fractional T-RGB'

Here the feature vector space of the frame of size $N \times N \times 3$ has $N \times N \times 3$ number of elements. This is obtained using following steps of T-RGB:

- i. Read Videos from Database.
- ii. Read every 10th frame of Videos up to 100th frames.
- iii. Separate Red, Green and Blue planes of the color frame.
- iv. Apply the Transform 'T'[4] on individual color planes of frame.
- v. Combine Red, Green and Blue planes of the color frame.
- vi. Extract features of color planes of frames.
- vii. Concatenate feature of frames of Video.
- viii. The result is stored as the complete feature vector 'T-RGB' for the respective video.

4.3 Query Execution

Query Execution is done as follows:

- i. Read Query Videos from Database.
- ii. Read $N \times N \times 3$ every 10th frame of Query Videos up to 100th frames.
- iii. Separate Red, Green and Blue planes of the color frame.
- iv. Apply the Transform 'T'[4] on individual color planes of frame.
- v. Combine Red, Green and Blue planes of the color frame.
- vi. Extract features of color planes of frames.

- vii. Concatenate feature of frames of Query Video.
- viii. This feature set is compared with other feature sets in feature database using Mean Squared Distance as similarity measure.

5. IMPLEMENTATION

Database contains 50 videos of each category Mountains, Elephant, Desert, Zoozoo, Ocean, Firework, Dance, Lecture, Obama and Lawn Tennis. Hence, total of 500 videos are stored in the database. Precision and recall are used as statistical comparison parameters[4,5] for the proposed CBVR technique. The standard definitions of these two measures are given by following equations.

$$Precision = \frac{Number_of_relevant_videos_retrieved}{Total_number_of_videos_retrieved}$$

$$Recall = \frac{Number_of_relevant_videos_retrieved}{Total_number_of_relevant_videos_in_database}$$

6. RESULT AND DISCUSSION

The performance of each proposed CBVR technique is tested by firing 500 queries per technique on the database of 500 videos. Average Mean Square Error distance is used as similarity measure. The average precision and average recall are computed by grouping the number of retrieved videos sorted according to ascending average Mean Square Error distances with the query video. In all transforms, the average precision and average recall values for CBVR using fractional coefficients are higher than CBVR using full set of coefficients.

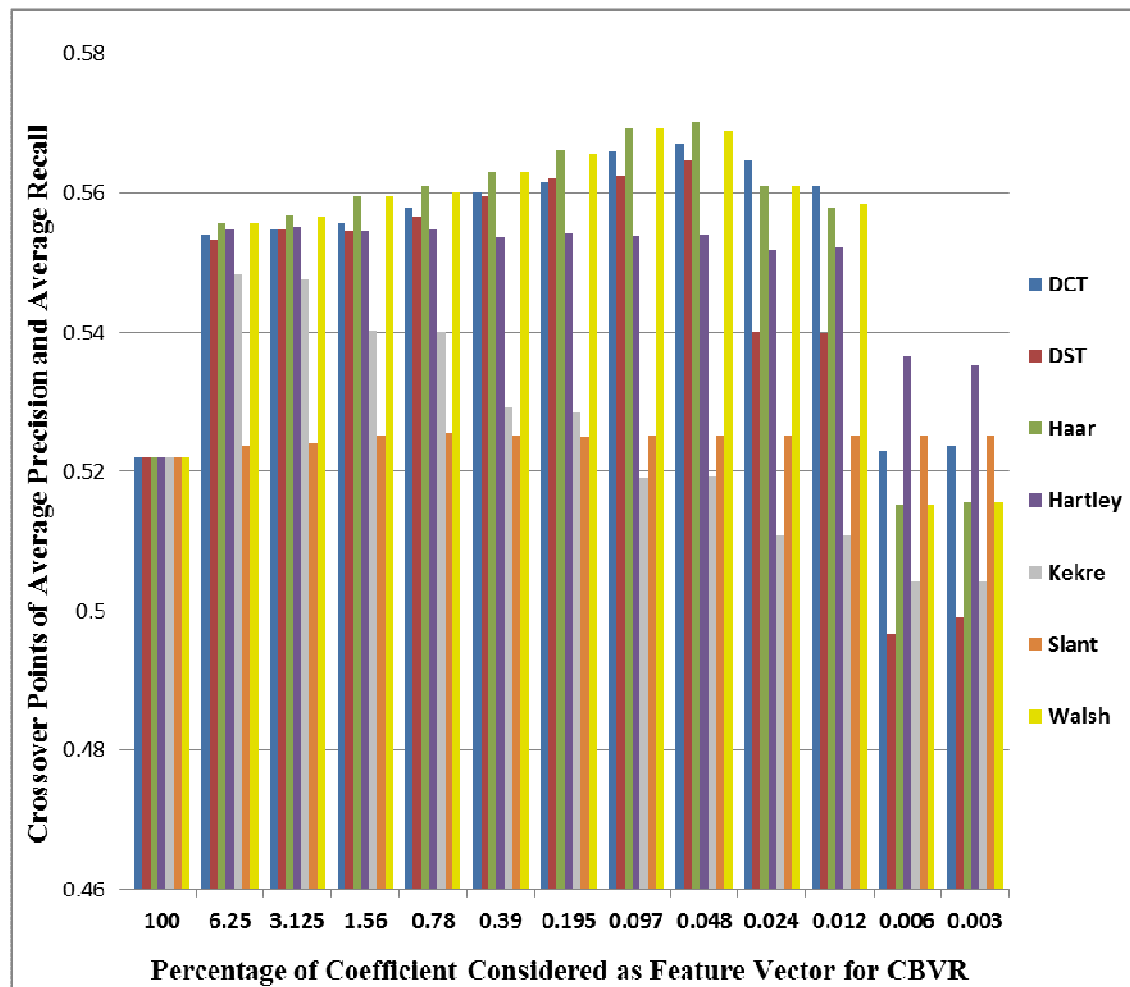


FIGURE 2: Performance Comparison Using Fractional Coefficients across Transforms.

Figure 2 shows performance comparison using Fractional Coefficients across Transforms. As shown in figure 2, Haar transform performs better than all discussed transforms till 0.024% of coefficients. Walsh transform performs second best compared to all discussed transforms till 0.024% of coefficients. DCT Transform performs best result when 0.012% coefficient is considered. But, Hartley Transform performs best result when 0.006% and 0.003% coefficient is considered.

In all, Haar Transform gives highest average precision and average recall showing it outperforms all transforms.

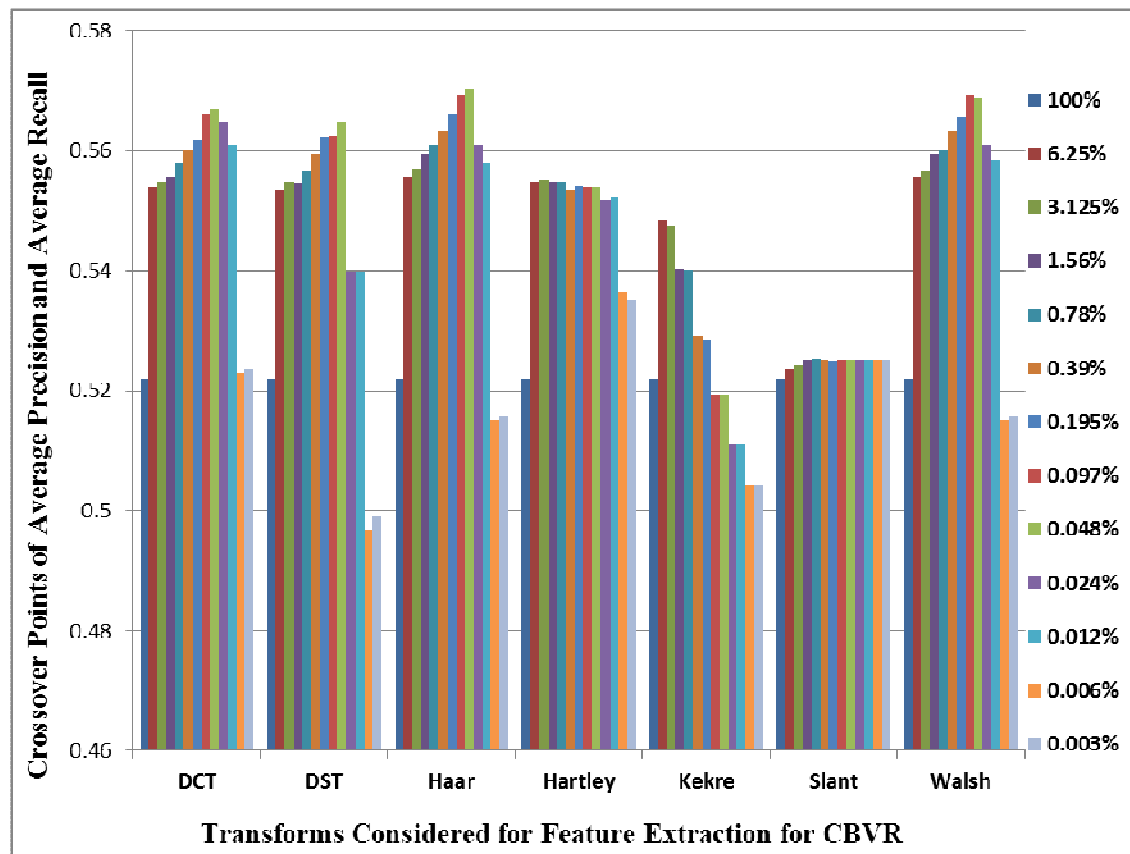


FIGURE 3: Performance Comparison of Transforms w.r.t Percentage of Coefficient Taken into Consideration.

As shown in Figure 3, when 0.048% coefficient is taken in consideration DCT, DST, Haar and Walsh transforms gives best performance as compared to all percentage coefficients taken. DCT, DST, Haar and Walsh transforms shows continuous increase in performance from 100% to 0.048%. Performance of Hartley and Slant does not show much variation with reduction of number of coefficients.

7. COMPARATIVE ANALYSIS

The following are the comparative analysis of the proposed method with other researches:

- In the current trend, every first frame of the shot of a video is taken as a key frame. To compute first frame of every shot of a video, firstly a video is divided into shots. This is done by taking the difference between every two consecutive frames. Whenever the difference is large then, those consecutive frames are kept in separate shot. When a video is divided into shots then, every first frame of the shot is taken as a key frame. This leads to increase in time complexity.

This is avoided in the proposed method by taking every 10th frame of a video and hence reducing the time complexity.

- In other researches, feature vector is formed by extracting Textures, comparing blocks of frames to get motion feature, color histograms to get color features, which increases computational complexity because size of the feature vector is large.
- In the proposed method, feature vector is formed by taking frames in Transformed domain with variable sizes of coefficients and hence reducing computational and as well time complexity by reducing feature vector size.

- In others researches, only 5 videos are taken into consideration and with only 1 category of video, whereas in the proposed method, results are displayed using 500 videos with 10 different categories.
- In the proposed method, the highest crossover point of average precision and recall is 0.5702 using Haar Transform when 0.048% coefficient is taken into consideration. This means that accuracy is 57.02% with reduction in time complexity by 2048 times compared to full feature set.

8. CONCLUSION

In the information age where the size of video databases is growing exponentially, more precise retrieval techniques are needed, for finding relatively similar videos. Computational complexity and retrieval efficiency are the key objectives in the video retrieval system.

Average Precision and Recall Crossover points is taken as performance index since its values varies between 0 and 1. When precision and recall is 1, then all the videos, similar to the query video are fetched from the database. But, when precision and recall is 0, then none of the videos which are similar to the query video are fetched from the database.

Here, the performance of video retrieval is improved using fractional coefficients of transformed frames of video at reduced computational complexity. In all transforms (DCT, DST, Haar, Hartley, Kekre, Slant and Walsh), the average precision and average recall values for CBVR using fractional coefficients are higher than using full set of coefficients. Hence, the feature vector size for video retrieval could be greatly reduced, which ultimately will result in faster query execution in CBVR with better performance.

Haar Transform performs best compared to all other transforms that is DCT, DST, Haar, Kekre, Slant and Walsh. Crossover points of Precision and Recall of Kekre reduces as size of coefficients reduces. Hartley and Slant does not show any variation as size of coefficients decreases whereas for other transforms increases due to energy compaction. When 0.048% coefficient is taken then DCT, DST, Haar and Walsh transforms give best performance as compared to all percentage coefficients.

In future, wavelets generated from these transforms can be implemented and compare results with the transforms.

9. REFERENCES

- [1] B.V.Patel and B.B.meshram, "Content based Video Retrieval Systems", IJU, vol.3, No.2, April 2012.
- [2] T.N.Shanmugam and Priya Rajendran, "An Enhanced Content-Based Video Retrieval System Based On Query Clip", International Journal of Research and Reviews in Applied Sciences, ISSN: 2076-734X, EISSN: 2076-7366 ,vol.1, Issue 3(December 2009).
- [3] Kalpana Thakre, Archana Rajurkar and Ramchandra Manthalkar, "An effective CBVR system based on motion, quantized color and edge density features", IJCSIT, vol.3, No 2, April 2011.
- [4] H.B.Kekre, Sudeep D. Thepade, "Improving the Performance of Image Retrieval using Partial Coefficients of Transformed Image", International Journal of Information Retrieval (IJIR), Serials Publications, vol. 2, Issue 1, 2009, pp. 72-79(ISSN: 0974-6285)
- [5] Dr.H.B.Kekre, Dr. Sudeep D. Thepade and Akshay Maloo, "Comprehensive Performance Comparison of Cosine, Walsh, Haar, Kekre, Sine, Slant and Hartley Transforms for CBIR with Fractional Coefficients of Transformed Image", IJIP, vol.5, Issue (3) : 2011
- [6] Ahmed, N.; Natarajan, T. ; Rao, K.R. "Discrete Cosine Transform", IEEE TRANSACTIONS ON COMPUTERS, vol.C-23, Issue: 1, pp 90 – 93, Jan. 1974.
- [7] R. N. Bracewell, "Discrete Hartley transform," Journal of the Optical Society of America, vol.73, Issue 12, pp 1832-1835, Dec. 1, 1983.

- [8] Maurence M. Angush and Ralph R. Martin, "A Truncation Method for Computing Slant Transforms with Applications to Image Processing", IEEE TRANSACTIONS ON COMMUNICATIONS, vol.43, No.6, June 1995
- [9] P.Geetha and Vasumathi Narayan, "A Survey of Content Based Video Retrieval", Journal of Computer Science, vol. 4 (6),pp 474-486, 2008
- [10] C.V.J Jawahar, Balakrishna Chennupati, Balamanohar Paluri, Nataraj Jammalamadaka, "Video Retrieval Based on Textual Queries", International Conference on Advanced Computing and Communication, 2005

Reversible Data Hiding Using Contrast Enhancement Approach

A. S. Al-Fahoum

*Hijjawi Faculty for Eng. Tech./Biomedical Sys. & Inform. Eng. Dept.
Yarmouk University
Irbid, 21163, Jordan*

afahoum@yu.edu.jo

M. Yaser

*Hijjawi Faculty for Eng. Tech./Biomedical Sys. & Inform. Eng. Dept.
Yarmouk University
Irbid, 21163, Jordan*

amjed.alfahoum.rcap@gmail.com

Abstract

Reverse Data Hiding is a technique used to hide the object's data details. This technique is used to ensure the security and to protect the integrity of the object from any modification by preventing intended and unintended changes. Digital watermarking is a key ingredient to multimedia protection. However, most existing techniques distort the original content as a side effect of image protection. As a way to overcome such distortion, reversible data embedding has recently been introduced and is growing rapidly. In reversible data embedding, the original content can be completely restored after the removal of the watermark. Therefore, it is very practical to protect legal, medical, or other important imagery. In this paper a novel removable (lossless) data hiding technique is proposed. This technique is based on the histogram modification to produce extra space for embedding, and the redundancy in digital images is exploited to achieve a very high embedding capacity. This method has been applied to various standard images. The experimental results have demonstrated a promising outcome and the proposed technique achieved satisfactory and stable performance both on embedding capacity and visual quality. The proposed method capacity is up to 129K bits with PSNR between 42-45dB. The performance is hence better than most existing reversible data hiding algorithms.

Keywords: Reversible Data Hiding, Histogram, Stenography, Performance, Watermarking.

1. INTRODUCTION

As multimedia becomes an important form of information exchange, a large number of digital products are created and transmitted via the Internet. One of the characteristics of digital products is that they are very easy to create, to store, to duplicate, to transmit, and to modify. These results in a serious problem because unauthorized use, copy, or modify of the products will also become very easy. How to protect the products in various aspects presents a problem that challenges the academia and the business. Several technologies have been proposed for intellectual property right (IPR) protection [1]. One is encryption. However, conventional encryption and copy protection mechanisms do not fully solve this issue in some applications. Recently, data hiding (watermarking) is proposed as a hopeful method for authentication, fingerprint, security, data mining, and copyright protection [2]. In general, the watermarking algorithm attempts to imperceptibly embed one signal into another. However, even if the presence of the watermark is imperceptible, the embedding process usually introduces irreversible degradation of the original media. The distortion might be very slight, but zooming in on the image to emphasize the details would result in visual perception below what's acceptable [2-9].

Image data hiding represents a class of processes used to embed the data of secret image into another image which is defined as cover images [10]. The hidden data usually can be a string of

binary bits (e.g., digital signature), a logo image, identification (ID) number, or any information that is useful [11]. In these approaches, the embedding process should be reversible [5]. In other words it is possible to completely restore the product content to its original state after watermark removal [4-9]. The basic requirements for a reversible data embedding technique can be summarized as follows:

1. **Reversibility:** Tian [10] has defined reversibility as “one can remove the embedded data to restore the original image.” It is the most important and essential property for reversible data embedding algorithms.
2. **Capacity:** The capacity for data embedding should be as large as possible. Too small payload size will restrict the range of application. The capacity is one of the important factors for measuring the performance of the algorithm.
3. **Fidelity:** Watermarking techniques characterized with high capacity might lead to low fidelity. The perceptual quality of the watermarked image should not be degraded severely after watermark embedding, although the original content is supposed to be restored completely.
4. **Robustness:** The embedded data can be retrieved if some irreversible distortions took place [12]. The techniques should be as robust as possible against common image processing operations.

Recently, reversible data embedding techniques have drawn more and more interest. Existing methods can be classified according to the techniques associated with restoration. Some of the approaches rely on lossless compression to exploit the redundant space created by the compression operation. The space thus created is used for data embedding and for preserving the original pixel values [13-22]. Fridrich [17] proposed a reversible watermarking system that losslessly compresses a full least-significant-bit plane. The compressed result is combined with the authentication code and the payload to produce the watermark, which is further encrypted for security considerations. The encrypted watermark is then embedded into the original image by replacing the least-significant-bit plane. Their method meets the requirements of reversibility and blind authentication, but the capacity is very limited (around 100 bits/image) and varies according to the characteristics of the image. Fridrich also developed a number of extended techniques for uncompressed images such as BMP or TIFF format, compressed images such as JPEG, and video such as MPEG-2. Celik et. al. [18] made some improvement on Fridrich's method to produce higher capacity and lower distortion on the image. The original image is quantized and the quantization residual is obtained by an L-level scalar quantization function. The residual is then losslessly compressed in order to create a room for the payload. The compressed residual and the payload data are then combined and embedded into the quantized image via G-LSB modification. The capacity depends on the characteristics of the host image i.e. images with many smooth regions will have larger capacity than those with irregular textures producing the same degree of fidelity. The resulting embedding capacity is about 1 and 0.17 bits/pixel for the best and the worst cases, respectively.

In general, the capacity with regard to compression-based algorithms depends on the nature of the image and is restricted to the performance of lossless compression techniques. In Kalker et. al. [20] the theoretical capacity bounds for reversible data embedding based on lossless compression was proven. As wavelet transform becomes very popular, some techniques used it to map integers to integers so that the basic property of reversibility can be achieved. Such special techniques are referred to as integer wavelet transform, which is used by many researchers to produce high embedding capacity [21]-[22]. Tian [10] used the integer Haar wavelet transform as a reversible transform and used difference expansion to create extra capacity for data embedding. As the pair of neighboring pixels of a natural image is usually highly correlated, the intensity difference between them is usually very small. Hence, Tian developed a method for creating high embedding capacity and applied it to the difference values of pairs of pixels of the host image. The resulting embedding capacity is 0.5 bits/pixel for the best case. One

of the drawbacks of Tian's approaches is that the original LSB of the difference values and the location map should also be embedded as part from the payload into the original image, thus also taking up the capacity since the pure payload is decreased due to the inserted information that is considered vital for image restoration. To circumvent, the location map is compressed before embedding. Invertible data hiding techniques utilized spatial and transform domains for maintaining good fidelity and higher payloads. Several techniques are applied with considerable improvements and shortcomings [21]-[33].

Xuan [21] for example used the integer wavelet transform to avoid the round-off error in forward and inverse translation procedures. Since lossless compression on the original wavelet coefficients is performed to create the embedding space, the capacity is only as high as 0.35 bits/pixel. Such a shortcoming is resolved by various high capacity lossless data embedding methods. Lin et. al. [34] used the coefficients varieties of Discrete Cosine Transformation (DCT) of an image by dividing the image into several bands and embedded the high frequency band with the secret data. Their DCT approach is based on integer mapping. The DCT coefficient distribution was very close to a Gaussian distribution with zero mean. Accordingly, they used the histogram shifting approach to shift the positive and negative coefficient around the zero to the right and left respectively leaving a space to the hidden data. However, if the marked image is loosely compressed high frequency details cannot be obtained leading to losing the hidden data.

Most existing watermarking techniques distort the original content as a side effect of image protection. To solve this problem, new technique using histogram statistics is proposed. Histogram manipulation techniques are evolved techniques that provided promising outcomes [7]. Techniques based on lossless compression or integer wavelet transform embed the original values concerning the host image to restore the exact original content. Such a problem may also be solved by histogram manipulation. De Vleeschouwer [15] used the method of patchwork histogram rotation, which is based on the concept of circular image histogram. The image is portioned into a given number of blocks. Each block is subdivided into 2 subsets A & B that are mapped into a unit circle. The two vectors that are related for A & B are computed. Embedding of the message bits can be released by rotation of both vectors in opposite ways. By slightly modifying the distribution of the pixel values of an image, the watermarked image is very similar to the original. As only one bit of the message is associated with a block, the capacity of their algorithm is very limited and depends on the number of blocks allocated in the image.

In [35] the histogram shifting techniques was used by using the prediction to explore the similarity of neighborhood pixels in the image and then hide the secret data by using the residual histogram of the predicted error, also they increased the hiding capacity using overlapping between the peak and zero pairs.

Hong et. al. [36] proposed an extension to [35] by employing set of basic pixels to predict the value of non basic pixels. This improves the prediction accuracy and increases the payload. The image quality was maintained by using a threshold to select the low variance blocks only. Chang et. al. [37] used an absolute difference technique of the neighboring pixels to determine whether the pixel is embedded or not. If the absolute difference is higher than the threshold then the pixel remains unchanged to achieve high image quality.

Zhicheng [4] utilized the redundancy of the image histogram and slightly modified the pixel values to embed the data. The algorithm is simple and efficient and can keep high visual quality for all images with a PSNR value guaranteed to be higher than 48 dB. However, the capacity only equals to the peak of the image histogram and is about 0.05 bits/pixel in average. Furthermore, their approach is not general enough for every image; namely, the performance of the algorithm varies from image to image. However, those previously proposed algorithms that are based on histogram manipulation are usually limited in the embedding capacity but they have an advantage that the original values about the host image need not necessarily be embedded into the host image [22]-[33]. This paper proposes a novel reversible data embedding method using histogram statistics to improve the embedding capacity and fidelity. The method enables exact recovery of

the original host image, achieves very high embedding capacity, and keeps the reversible distortion very low.

2. METHODOLOGY

In this work, a novel reversible data embedding algorithm using contrast stretching is proposed. The contrast stretching technique is used instead of lossless compression to explore the redundancy in the image histogram and to create greater embedding space. In addition, such a technique is designed to reversibly embed the payload into the image histograms without embedding the original values concerning the host image. In addition, the proposed technique achieved satisfactory and stable performance both on embedding capacity and visual quality. Figure 1 and 2 respectively summarize the proposed algorithm. Figure.1 shows a schematic diagram of the proposed reversible data embedding system that will be applied to the output image obtained using Zhicheng's method. First, the output image from zhicheng's method, I_o , is divided into equally, non-overlapping blocks, which are the inputs to the system. And then, the system computes the histogram statistics, H_o , of each image block, which includes the gray level with the most counts and the number of pixels used to calculate the histogram. The contrast of the histogram of each image block is then stretched to create an extra embedding space, with the reverse-aid information A produced as well. The watermark W thus consists of the data payload and the reverse-aid. Finally, the watermark is embedded into the generated space in the stretched image, I_s , producing the watermarked image, I_w .

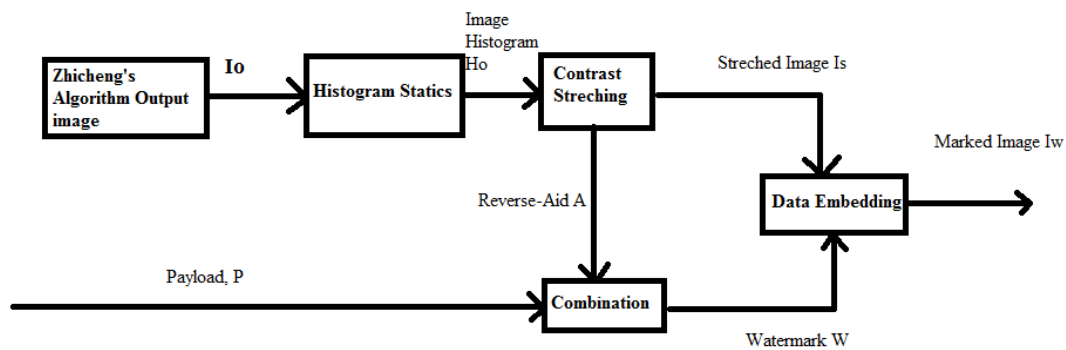


FIGURE 1: The Embedding Process of the Removable Data Using Proposed Method.

The data extraction and restoration process of the removable data embedding using proposed method is illustrated in Figure.2. The restoration process starts with the division of the received image I_w into equally sized $N \times N$ blocks. Next, the system computes the histogram statistics, H_w , of each image block and extracts the watermark, W' . During restoration, the embedded watermark is removed from the received image, thus restoring it to its original state.

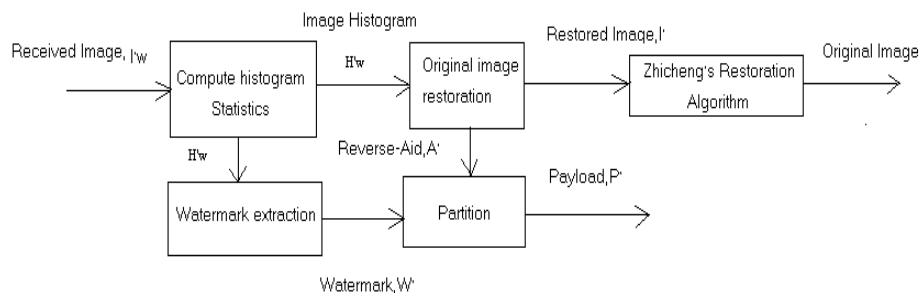


FIGURE 2: The Data Extraction and Restoration Process of the Removable Data Embedding using Proposed Method.

2.1 Zhicheng's Method [4]

Zhicheng et. al. utilized the redundancy of the image histogram and slightly modified the pixel values to embed the data. The algorithm is simple and efficient. Therefore, it can keep high visual quality for all images with a PSNR value guaranteed to be higher than 48 dB. However, the capacity only equals to the peak of the image histogram and is about 0.05 bits/pixel. This capacity belongs to the range of (5K-80K) bits for a 512x512x8 grayscale image. Furthermore, their approach cannot be generalized since the capacity variation is very high and may change dramatically from image to image.

2.1.1 An Embedding Algorithm Procedure

The "lena" image is used as an example to illustrate this algorithm. For a given grayscale image, say lena image (512x512x8), its histogram is first generated as shown in figure 3.

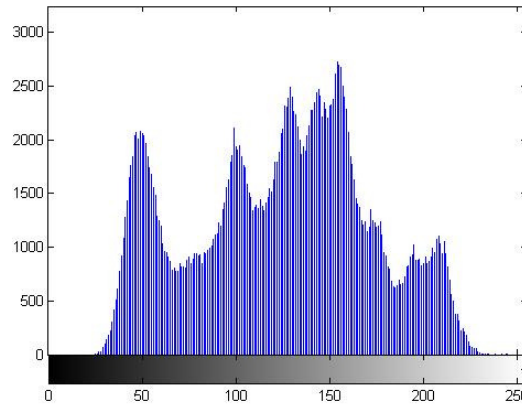


FIGURE 3: Lena Original Histogram.

An illustration of the Embedding Algorithm Using an Example With One Zero Point and One Peak Point is summarized as follows: In the histogram, we first find a zero point, and then a peak point. A zero point corresponds to the grayscale value which has zero frequency and belongs to the image domain, e.g., $h(255)$ as shown in Figure 3. A peak point corresponds to the grayscale value which has the maximum number of repeated pixels within the given image, e.g., $h(154)$ as shown in Figure 3. To illustrate the principle of the algorithm, and for the sake of notational simplicity, only one zero point and one peak point are used in this example. The objective of finding the peak point is to increase the embedding capacity to be as large as possible since in this algorithm, as shown below, the number of bits that can be embedded into an image equals to the number of pixels which are associated with the peak point.

The whole image is scanned in a sequential order, say, row-by-row, from top to bottom, or, column-by-column, from left to right. The grayscale value of pixels between 155 (including 155) and 254 (including 254) is incremented by "1." This step is equivalent to shifting the range of the histogram, [155 254], to the right-hand side by 1 unit, leaving the grayscale value 155 empty as shown in Figure 4.

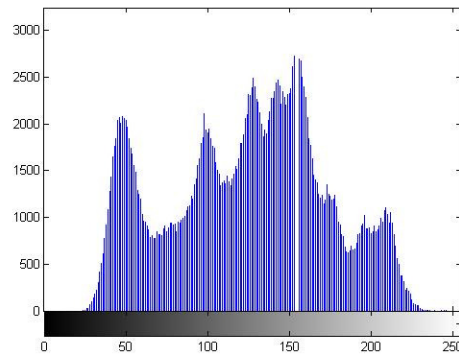


FIGURE 4: Lena Histogram after Shifting.

Now, the whole image is scanned once again in the same sequential order. Once a pixel with grayscale value of 154 is encountered, we check the to-be-embedded data sequence. If the corresponding to-be-embedded bit in the sequence is binary “1”, the pixel value is incremented by 1. Otherwise, the pixel value remains intact. (For illustration purposes, it is preferred to present the embedding algorithm in these three steps only.)

The above three steps complete the data embedding process. Now we can observe that the data embedding capacity of this algorithm, when only one pair of zero and peak points is used, equals to the number of pixels that assume the grayscale value of the peak point as mentioned above. The histogram of the marked Lena image is plotted in Figure 5. It is noteworthy that the peak at 154, which was shown in Figure 3 is no longer the peak value in Figure 5.

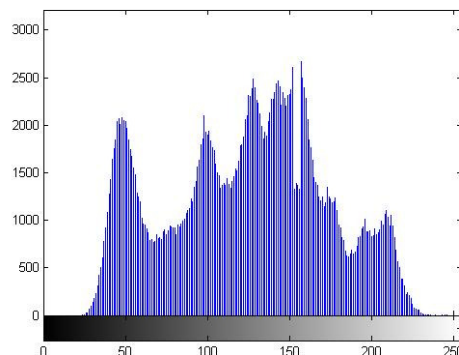


FIGURE 5: Lena Histogram after Zhicheng's et. el. Method.

In very rare cases, the zero point is not able to be found in a histogram. In such situations, the minimum point is used instead of the zero point. Then, the gray value and the coordinates of the minimum point are recorded as an overhead part of the embedded data. This book-keeping information will be used later to recover the minimum point data retrieval. If there are multiple pairs of zero points and peak points, it is possible to further increase the payload by adding complexity to this algorithm.

2.1.2 Data Extraction Procedure

Only the simple case of one pair of minimum point and maximum point is described here because, as shown above, the general cases of multiple pairs of maximum and minimum points can be decomposed as a few one pair cases.

Assuming the grayscale value of the maximum point and the minimum points are a and b , respectively. Without loss of generality, assume $a < b$. The marked image is of size $(M \times N)$, and

each pixel grayscale value x belongs to the interval between $[0,255]$ (assuming 8 bit pixel representation).

The next step is to scan the marked image in the same sequential order as that used in the embedding procedure. If a pixel with its grayscale value $(a+1)$ is encountered, a bit "1" is extracted. If a pixel with its value (a) is encountered, a bit "0" is extracted.

Scan the image again, for any pixel whose grayscale value between $[a,b]$, the pixel value is subtracted by 1.

If there is overhead bookkeeping information found in the extracted data, set the pixel grayscale value (whose coordinate (i,j) is saved in the overhead) as b .

In this way, the original image can be recovered without any distortion, but the total capacity is limited to the sum of peaks count that is used in embedding process which will not be able to give a capacity more than 0.05 bpp. To improve the capacity of this algorithm, the next section proposes an enhancement method that will provide better results compared to this method.

2.2 Contrast Stretching

Contrast stretching is one of the histogram manipulation techniques. It tries to extend the narrow range of image pixel values over a wider range. It is often used to improve the image quality by stretching the range of intensity values such that the image contrast could be enhanced. Contrast enhancement will produce a histogram which contains many gaps. These gaps represent a missing set of pixels of particular intensity values. This additional created empty space would be used for data embedding. Figure 6 illustrates the proposed method of contrast enhancement for reversible data embedding. In this method, every intensity value in the original histogram is stretched to a position either to the left or to the right of the peak value, hence creating the missing intensity values. As large gaps of missing intensity values together could produce a downgraded look; the width of a gap is defined as one grayscale value in this work. The embedding process is implemented as follows: the image block is scanned pixel by pixel. The pixel values associated with the histogram peak are kept intact for reversibility. The rest may be selected to embed the watermark. If the current watermark bit is "1," the pixel value is added or subtracted by 1, otherwise, they are kept intact. For example, Figures 6.a and 6.b show the original and stretched histograms of an image block, respectively. Pixels with gray level 57 are selected for embedding and gray level 58 is a missing intensity value. While scanning the image block, if the watermark bit is "1," the value 57 of the current pixel is incremented by 1, whereas if the watermark bit is "0," nothing is done. Figure 6.c shows the histogram after embedding the watermark. The total number of pixels of gray levels 57 and 58 is equal to the number of pixels of gray level 60 in the original histogram. The embedding capacity in this image block is the number of selected pixels, i.e., the number of pixels of gray levels 60, 61, 62, 64 and 65.

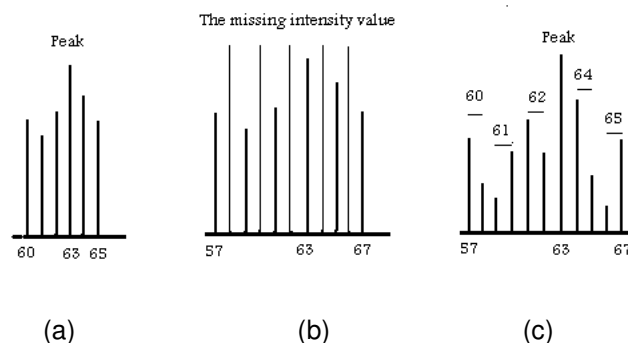


FIGURE 6: Removable data embedding by contrast stretching: (a) original histogram of the image block, (b) stretched histogram, (c) histogram after embedding.

Let m be the image block of size $N \times N$ and $m(i, j)$ denotes the original pixel value at position (i, j) , where $0 \leq i, j \leq N-1$. The contrast stretching is defined as

$$m'(i, j) = p + 2 * [m(i, j) - p] = 2m(i, j) - p, \quad (1)$$

where $m'(i, j)$ is the pixel value at i, j position after stretching and p is the gray level value with the highest count of pixels in the block's histogram. A threshold value τ is used to balance between the shift widths and the number of selected gray levels for embedding. In Figure 6.a, the pixels with gray values of 61, 62, 64 and 65 are selected to embed the watermark, therefore τ is selected to be 2. The original pixel values that satisfy $p - \tau \leq b(i, j) \leq p + \tau$ are mapped to the output value using the contrast stretching function in equation 1. Thus the range of $[p - 2\tau, p + 2\tau]$ is called the embedding area. Other pixel values will be shifted to the right or left by τ levels. This means

$$b'(i, j) = 2 * b(i, j) - p, \text{ if } p - \tau \leq b(i, j) \leq p + \tau \quad (2)$$

$$b'(i, j) = b(i, j) - \tau, \text{ if } b(i, j) < p - \tau \quad (3)$$

$$b'(i, j) = b(i, j) + \tau, \text{ if } b(i, j) > p + \tau \quad (4)$$

As grayscale values are bounded to the interval $[0, 255]$, then $0 \leq b'(i, j) \leq 255$, which is equivalent to $0 \leq 2 * b(i, j) - p \leq 255$. To overcome the over/underflow problems (the range of shifted pixel values to be less than 0 or greater than 255) and the extreme value problem (if intensity value 0 or 255 exists, there is no room for stretching). Accordingly, extra conditions should be applied to avoid such problems. The pixel values in an image block must satisfy the following condition:

$$p/2 < b(i, j) < (p + 255)/2 \quad (5)$$

Considering the threshold, the pixel values violating condition (3) might not cause any problem if they satisfy another constraint:

$$\tau < b(i, j) < 255 - \tau \quad (6)$$

In the cases where the value in an image block satisfies equation (5) & (6), then it could be stretched via equations (2), (3) & (4); otherwise, the following equations are used instead without causing the over flow or underflow problem:

$$b'(i, j) = 2 * b(i, j) - p, \text{ if } p - \varepsilon_1 \leq b(i, j) \leq p + \varepsilon_2 \quad (7)$$

$$b'(i, j) = b(i, j) - \tau, \text{ if } b(i, j) < p - \varepsilon_1 \quad (8)$$

$$b'(i, j) = b(i, j) + \tau, \text{ if } b(i, j) > p + \varepsilon_2 \quad (9)$$

where ε_1 and ε_2 are the numbers of empty bins in the histogram from 0 to the minimum gray level and from the maximum gray level to 255, respectively. The range of $[p - 2\varepsilon_1, p + 2\varepsilon_2]$ is also called the embedding area. In addition, the number of the gray levels to be stretched will be the number of the missing intensity values clustered at one end of

the histogram. If the range of grayscale values in the histogram covers the full possible set of values, straightforward contrast stretching will achieve nothing; in other words, the histogram should have at least one gap on the left or right side of the peak in order to provide room for stretching. Usually, the pixel values in an image block spread only within a small range so that they could be stretched.

2.3 Reversible Data Embedding

The reversible data embedding algorithm can be summarized in the following steps:

The original image is divided into equally-sized ($N \times N$), non-overlapping image blocks, and N is set to be less than 16 such that the full-coverage problem is avoided.

The histogram statistics is computed for each image block, including the gray level with the most counts, p , and the number of pixels. If there is more than one peak in a block, the peak is defined as the one with greater gray level. Figure 7 shows one sample for a block histogram before stretching.

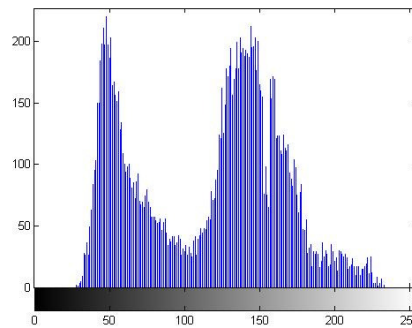


FIGURE 7: A Sample of Block Histogram before Stretching.

Contrast stretching is performed on each image block to create extra embedding space. Figure 8 shows the histogram of a block after contrast stretching.

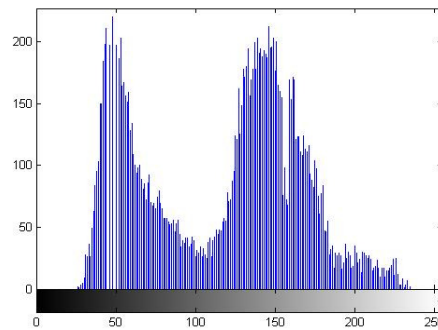


FIGURE 8: Block Histogram after Stretching.

Each block is divided into two sets of pixels, L (left of the histogram) and R (right of the histogram), according to the pixel values being less or greater than p . In some exceptional cases, if the peak is the minimum or maximum gray level, this block is only divided into R or L accordingly. If all of the pixels in the block have the same value, this block is discarded.

The watermark is embedded into the created space after contrast stretching and the watermarked image is obtained. The data embedding starts with scanning the pixel values, and pixel values in the embedding area are modified by either adding or subtracting one bit:

$$b''(i, j) = b'(i, j) - 1, \text{ if } b'(i, j) > p \quad (10)$$

$$b''(i, j) = b'(i, j) + 1, \text{ if } b'(i, j) < p \quad (11)$$

Where $b''(i, j)$ denotes the pixel value after data embedding. Figure 9 shows the block histogram after data embedding.

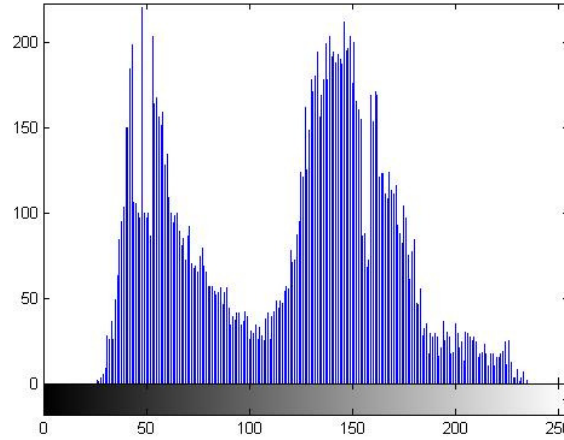


FIGURE 9: Block Histogram after Data Embedding.

2.4 Data Extraction and Image Restoration

To extract the watermark and restore the original image, the steps below are followed.

The received image is divided into image blocks by the same size $N \times N$, followed by finding the peak of the histogram of each block.

The watermark extraction starts with pixel scanning, and is extracted by

$$w_k = b''(i, j) + p \bmod_2 \quad (12)$$

where w_k is the watermarked image bit.

The image restorer implements the difference translation according to the pixel value in $L \{ b(i, j) < p \}$ or $R \{ b(i, j) > p \}$. For pixel values in the embedding area, then they are restored by the floor operation:

$$b(i, j) = \lfloor \frac{b''(i, j) + p}{2} \rfloor \quad (13)$$

For the pixels which are not in the embedding area, the pixels in L are restored by

$$b(i, j) = b''(i, j) + \tau \quad (14)$$

and the pixels in R are restored by

$$b(i, j) = b''(i, j) - \tau \quad (15)$$

The new overall image histogram is created and the Zhicheng's restoration procedure shown above in 2. 1 is executed. The embedded watermark is extracted and removed from the received image. Thus the original state of the image is restored without any kind of losses.

3. EXPERIMENTAL RESULTS

The proposed algorithm has been implemented and tested on various 512×512 standard grayscale images, including F-16, Peppers, Lena, Barbara, Boat, and Baboon. The selected images range from very smooth images (as F-16) to very rough images (as Baboon). The payload data are created from a pseudo random number generator. The relationship between the block size and the performance of the proposed algorithm is important. Whereas increasing the block size does not dramatically increase the capacity and a larger block size leads to lower image quality as shown in figure10. Since each image block is stretched individually without considering the distribution of adjacent blocks. The block effects are apparent and the image quality is degraded. For this reason, in this work the 4×4 image blocks is chosen, and τ is chosen as 2 to get a higher quantitative quality (PSNR) and to avoid any degradation that could affect the visual quality.

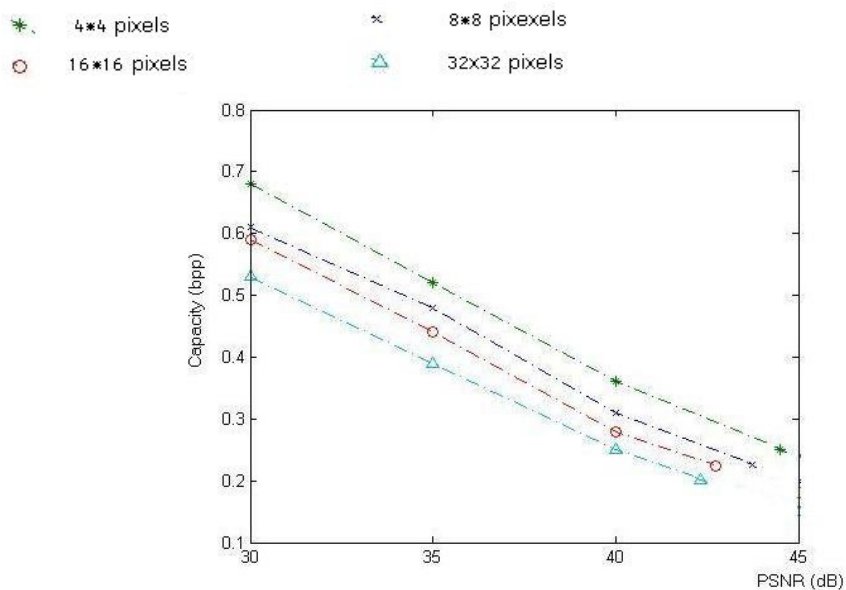


FIGURE10: The Effect of Different Block Size in Terms of Image Quality Versus Capacity.

In all conducted experiments, the PSNR of the all marked images is found to be above 42.01 dB. This cannot be theoretically proven, since in this algorithm, there are two shifting processes: one at the whole image intensity level; and the other at the block intensity level. Therefore, it is hard to find the real relation between these two shifting processes due to the possibility of having some kind of overlapping. In other words, the shift that may happen for the peaks at the image level can be adjusted during any contrast stretching for the block intensities. This reason leads to the conclusion that the PSNR may vary from image to image and block to block. However, this will make it very hard to theoretically calculate the exact range. Accordingly, the PSNR is calculated for each marked image independently.

The proposed reversible data hiding algorithm has been applied to many typical grayscale images. Figure 11 shows the original and marked Lena image (512×512×8), respectively. Figure 12 shows the histogram of the original and marked Lena image (512×512×8). The measured PSNR is found to be 44.39 dB.

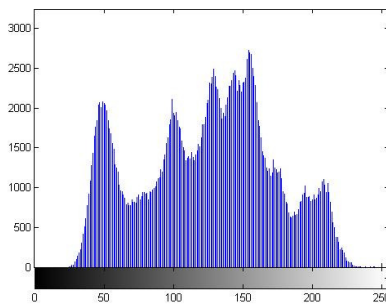


(a).Lena Original.

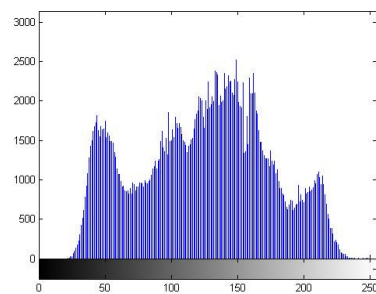


(b).Lena Marked (PSNR=44.39 dB)

FIGURE11: Lena Image.



(a). The Original Lena Histogram.



(b). The Final Histogram for Marked Lena Image.

FIGURE12: Lena Histograms.

The payload in Lena image is equal to the summation of counts for the maximum two peaks at the whole image histogram level plus the counts for the two peaks around the peak of each block which is equal to 5460 bits (Zhicheng's algorithm)+ 24098 bits (proposed algorithm)=29558 bits.

Figure 13 shows the experimental results for airplane. The pure payload is equal to 60,279. The measured PSNR is found to be 43.68 dB.

To demonstrate the efficacy of the proposed technique, the technique is applied on several images to investigate its influence on the values of both PSNR and Pay load. The results are shown in Table.1. Overall comparison between the existing reversible marking techniques and the proposed technique in terms of pure payload and the PSNR is presented in Table .2.



(a) Original.



(b) Marked (PSNR=43,68dB)

FIGURE 13: Air Plane Image.

| Images (512x512x8) | PSNR of marked image (dB) | Capacity (bits) |
|-------------------------------|--|----------------------------|
| Lena | 44.39 | 29,558 |
| Airplane | 43.68 | 60,279 |
| Baboon | 42.01 | 19,998 |
| House | 42.91 | 37,239 |
| Sail boat | 43.21 | 34,124 |
| Boat | 43.59 | 41,106 |
| Barbara | 43.89 | 20,864 |
| Blood | 44.74 | 129 |

TABLE 1: Experimental Results for the Proposed Algorithm.

| Methods | Pure payload in a 512x512x8 image (bits) | PSNR of marked image (dB) |
|-----------------------|--|---------------------------|
| Zhicheng's[4] | 5k-80k | 48.2 |
| Macq and Deweyand[12] | <2046 | Not mentioned |
| Vleeschouwer's[15] | <4096 | <35 |
| Honsinger's[16] | <1024 | Not mentioned |
| Fridrich's [17] | 1024 | Not mentioned |
| Celik's[18] | 15k-143k | 38 |
| Xuan's[21] | 15k-94k | 24-36 |
| Goljan's[24] | 3k-41k | 39 |
| Proposed | 20k-129k | 42-45 |

TABLE 2: The Comparative Evaluation between the Proposed Algorithm and 8 Reversible Methods.

Table 3 shows the comparison between other reversible watermarking methods and the proposed method applied on two typically different images: Lena and baboon.

| Method | Baboon (512x512x8) | | Lena (512x512x8) | |
|-----------------------|---------------------|---------------|---------------------|---------------|
| | Pure payload (bits) | PSNR (dB) | Pure payload (bits) | PSNR (dB) |
| Zhicheng's[4] | 5,421 | 48.2 | 5,460 | 48.2 |
| Macq and Deweyand[12] | <2046 | Not mentioned | <2046 | Not mentioned |
| Vleeschouwer's[15] | 1024 | 29 | 1024 | 30 |
| Honsinger's[16] | <1024 | Not mentioned | <1024 | Not mentioned |
| Fridrich's[17] | 1024 | Not mentioned | 1024 | Not mentioned |
| Celik's[18] | 15,176 | 38 | 74,600 | 38 |
| Xuan's[21] | 14,916 | 32.8 | 85,507 | 36.6 |
| Goljan's[24] | 2,905 | 39 | 24,108 | 39 |
| Proposed | 19,998 | 42.0 | 29,558 | 44.4 |

TABLE 3: Comparison between other reversible marking methods and the proposed method on two typically different images: Lena and baboon.

Table 4 shows the comparison between the proposed method and zhicheng's method for the same images.

| Images (512x512x8) | PSNR of marked image (dB) | Zhicheng's methods Capacity (bits) | Proposed methods Capacity (bits) |
|--------------------|---------------------------|------------------------------------|----------------------------------|
| Lena | 44.39 | 5,460 | 29,558 |
| Airplane | 43.68 | 16,171 | 60,279 |
| Baboon | 42.01 | 5,421 | 19,998 |
| House | 42.91 | 14,310 | 37,239 |
| Sail boat | 43.21 | 7,310 | 34,124 |

TABLE 4: Experimental Summery Results for Both Zhicheng's Method and the Proposed Method.

4. DISCUSSION

Tables 1-4 show the superiority of the proposed algorithm and its robustness against payload variations. Moreover, it shows the ability of the proposed algorithm to achieve higher quantitative and qualitative quality measures. The proposed algorithm satisfied the highest payload with an acceptable PSNR. Moreover, the computational load of the proposed algorithm is easily implemented in real time without any need for external added processors. This light computational load with its high payload capacity makes this algorithm very promising. The algorithm doesn't need to apply any transform such as discrete cosine transform (DCT), discrete wavelet transform (DWT), and fast Fourier transform (FFT). All the processing is taking place in the spatial domain. The required processing mainly lies on generating histogram, determining minimum and maximum (and possibly sub minimum and sub maximum) points, scanning pixels, and adding or subtracting pixel grayscale values by one in the spatial domain. Hence, the execution time of the algorithm is rather short.

The reversible data embedding based on contrast stretching –that applied to the output image of Zhicheng's method- creates extra space by translation of the image pixel values from a narrow range to a wider range. In addition, this technique preserves the reversibility by enforcing the peak intact in the block histogram. The experimental results demonstrated a promising outcome and it is obvious that contrast stretching method that applied to the output image from Zhicheng's method produces high and stable embedding capacity as well as fidelity. However, some limitations of the proposed method should be discussed. The embedded watermark has a bit length of $|A|$ (Reverse-Aid) $+|P|$ (Payload) and the total embedding capacity will be $|Counts\ of\ the\ maximum\ two\ peaks\ at\ whole\ histogram\ level|$ (Zhicheng's method) $+|Counts\ of\ four\ peaks\ around\ maximum\ peak\ at\ the\ block\ level|$ (Contrast stretching method). To successfully embed data into the host image, the following conditions must hold: $|A|+|P| \leq |Zhicheng's\ method| + |Contrast\ Stretching\ method|$, thus the payload size is bounded by

$$|P| = |Zhicheng's\ method| + |Contrast\ stretching| - |A| \quad (16)$$

The extended size of the payload capacity is due to the Contrast Stretching. However, the value of the capacity depends also on the type of the image probability distribution and the histogram of individual blocks. The embedding capacity increases if the distribution is more or less Gaussian distribution. This is because most pixel values are located around the peak.

It should be noted also that the size of A must be less than or equal to the size of the capacity provided by Zhicheng's and Contrast Stretching method's. Otherwise, the data embedding process cannot be performed. However, it is worth mentioning that the ability of the algorithm to be unsuccessful in embedding data is almost rare when natural images are used. This is because the probability of having an extreme value distribution is very low. Furthermore, the capacity might be very small if the distribution of image intensity is extreme. This kind of images is not suitable for data embedding.

5. CONCLUSION

In this paper, a novel and fully reversible data embedding algorithm for digital images is proposed. The new algorithm utilizes the simplicity and efficiency of reversible data embedding algorithms and enhances their payload capacity. The histogram contrast stretching is proposed to enhance Zhicheng's method to produce extra space for data embedding without affecting the quality of restoration process. In addition, the performance both on embedding capacity and visual quality is satisfactory and stable. For extremely important images, such as those used in medical, legal or military environment, the technique is very useful.

The proposed reversible data hiding technique is able to embed about 129kb into a 512 x 512 x 8 grayscale image while guaranteeing the PSNR of the marked image versus the original image to be above 42 dB. In addition, this algorithm can be applied to virtually all types of images. In fact, it has been successfully applied to many frequently used images with better capacity and high

visual quality. Further analysis and modifications may lead to provide a better trade in between visual quality and payload capacity. Future work will concentrate on reversible data hiding schemes that will utilize variations in frequency contents of the image features. The image will be decomposed into several different frequencies, and the low amplitude frequency components will be embedded with secret data. A multilevel histogram shifting can be used to improve capacity and fidelity of the new technique.

6. Acknowledgement

The authors would like to acknowledge the support and help provided by both Yarmouk University Research Council and Hijjawi Foundation.

7. REFERENCES

- [1]. G. Braudway, K. Magerlin. And F. Mintezer. "protecting publicity available images with a visual image watermark," Proc. SPIE: optical security and Counterfeit Deterrence Techniques, Vol. 2659. pp.126-133, 1996.
- [2]. B. Macq and J. Quisquater. "Cryptology for Digital TV broadcasting," Proc. Of IEEE. Vol. 83. No.6, pp 944-957, June 1995.
- [3]. W. Zeng, "Digital watermarking and data hiding: technologies and applications," in Proc. Int. Conf. Inf. Syst., Anal. Synth., vol. 3, pp. 223–229, , 1998.
- [4]. N. Zhichenge, ,Y. Q. Shi,N. Ansari, and W. Su, "Reversible Data Hiding" IEEE Transactions on circuits and systems for video technology, vol.16, No. 3, March 2006.
- [5]. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673–1687, Dec. 1997.
- [6]. Z. Tirkel, C. F. Osborne, and R. G. Van Schyndel, "Image watermarking a spread spectrum application," in Proc. IEEE 4th Int. Symp. Spread Spectrum Techn. Applicat., vol. 2, pp. 785–789, Sep. 1996.
- [7]. A. van Leest, M. van der Veen, and F. Bruekers, "Reversible Image Watermarking," in Proc. Int. Conf. Image Processing, Vol. 2, pp. 731–734, 2003.
- [8]. J. Huang and Y. Q. Shi, "An adaptive image watermarking scheme based on visual masking," Electron. Lett., vol. 34, no. 8, pp. 748–750, 1998.
- [9]. J. Huang, Y. Q. Shi, and Y. Shi, "Embedding image watermarks in DC component," IEEE Trans. Circuits Syst.: Video Technol., vol. 10, no. 6, pp. 974–979, Sep. 2000.
- [10]. J. Tian, "Reversible Data Embedding Using a Difference Expansion," IEEE Trans. Circuits and Systems for Video Technology, Vol. 13, No. 8, pp. 890–896, 2003.
- [11]. J. Irvine and D. Harle, Data Communications and Networks: An Engineering Approach. New York: Wiley, 2002.
- [12]. B. Macq and F. Deweyand, "Trusted headers for medical images," presented at the DFG VIII-D II Watermarking Workshop, Erlangen, Germany, Oct. 1999.
- [13]. M. M. Yeung and F. C. Mintzer, "Invisible watermarking for image verification," Electron. Imag., vol. 7, no. 3, pp. 578–591, Jul. 1998.

- [14]. B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.
- [15]. C. De Vleeschouwer, J. F. Delaigle, and B. Macq, "Circular interpretation on histogram for reversible watermarking," in *IEEE Int. Multimedia Signal Process. Workshop*, France, pp. 345–350, Oct. 2001.
- [16]. C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel, "Lossless Recovery of an Original Image Containing Embedded Data," U.S. Patent 6 278 791 B1, Aug. 21, 2001.
- [17]. J. Fridrich, M. Goljan, and R. Du, "Invertible authentication," in *Proc. SPIE Security Watermarking Multimedia Contents*, San Jose, CA, pp. 197–208, Jan. 2001,.
- [18]. M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Reversible data hiding," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, Sep. 2002, pp. 157–160.
- [19]. F. Perez-Gonzalez and F. Balado, "Quantized projection data hiding," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, pp. 889–892, Sep. 2002.
- [20]. T. Kalker, and F.M.J Willems, "Capacity Bounds and Code Constructions for Reversible Data-Hiding," in *Proc. Electronic Imaging, Security, and Watermarking of Multimedia Contents V*, 2003.
- [21]. G. Xuan, J. Zhu, J. Chen, Y. Q. Shi, Z. Ni, and W. Su, "Distortion less data hiding based on integer wavelet transform," *IEE Electron. Lett.*, vol. 38, no. 25, pp. 1646–1648, Dec. 2002.
- [22]. W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3–4, pp. 313–336, 1996.
- [23]. Y. Q. Shi, Z. Ni, D. Zou, and C. Liang, "Lossless data hiding: fundamentals, algorithms and applications," in *IEEE Int. Symp. Circuits Syst.*, Vancouver, Canada, pp. 33–36, May 2004.
- [24]. M. Goljan, J. Fridrich, and R. Du, "Distortion-free data embedding," in *Proc. 4th Inf. Hiding Workshop*, Pittsburgh, PA, pp. 27–41, , Apr. 2001.
- [25]. R. Calderbank, I. Daubechies, W. Sweldens, and B. Yeo, "Wavelet transforms that map integers to integers," *Appl. Comput. Harmonic Anal.*, vol. 5, no. 3, pp. 332–369, 1998.
- [26]. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, pp. 247–269, 1998.
- [27]. Park, J. Y., Lim, J. H., Kim, G. S. and Won, C.S. "Invertible Semi-Fragile Watermarking Algorithm Distinguishing MPEG-2 Compression from Malicious Manipulation," in *Proc. Int. Conf. Consumer Electronics (ICCE '02)*, pp.18–19, 2002.
- [28]. Yang, Y. and Bao, F. "An Invertible Watermarking Scheme for Authentication of Electronic Clinical Brain Atlas," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '03)*, Vol. 3, pp. 533–536, 2003.
- [29]. Zou D., Wu, C. W., Xuan, G. and Shi, Y.Q. "A Content-Based Image Authentication System with Lossless Data Hiding," in *Proc. Int. Conf. Multimedia and Expo (ICME '03)*, Vol. 2, pp. 213–216, 2003.
- [30]. P. Roos, M. A. Viegervet, M.C.A Van Dijke, and J. A. Peters, "Reversible Intraframe of Medical Images," *IEEE Trans. Medical Image*, Vol. 7, pp. 328–336, 1998.

- [31]. J. Y. Park, J. H. Lim, G.S. Kim, and C. S. Won, "Invertible Semi-Fragile Watermarking Algorithm Distinguishing MPEG-2 Compression from Malicious Manipulation," in Proc. Int. Conf. Consumer Electronics (ICCE '02), pp.18–19, 2002.
- [32]. X. Luo, Q. Cheng, and J. Tan, "A Lossless Data Embedding Scheme for Medical Images in Application of E-Diagnosis," in Proc. IEEE 25th Annual Int. Conf. Engineering in Medicine and Biology Society, Vol. 1, pp. 852–855, 2003.
- [33]. C. T. Hsu and J. L. Wu: Hidden Digital Watermarks in Images, IEEE Transactions on Image Processing, 58-68, 1999.
- [34]. Y. K. Lin " High capacity reversible data hiding scheme based upon discrete cosine Transformation" , The Journal of System and Software, vol:85(10), pp:2395-2404, 2012.
- [35]. P. Tsai, Y. ChenHu, , H.L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting" Signal Processing , pp: 1129-1143, 2009.
- [36]. W. Hong, T. S. Chen, " A local variance-controlled reversible data hiding method using prediction and histogram-shifting" The Journal of System and Software, vol. 83 (12), pp: 2653-2663, 2010.
- [37]. Ch. Ch. Chang, Y. H. Huang, H. Y. Tsai, and Ch. Qin," Prediction-based reversible data hiding using the difference of neighboring pixels" International Journal of Electronics and Communication (AEU), vol. 66. issue 9, pp:758:766, 2012.

Satellite Imaging System

AA Somaie

Management Information Systems
University of Business and Technology UBT,
Jeddah, KSA

aasomaia@cba.edu.sa

Abstract

The aim of this paper is to present the essential elements of the electro-optical imaging system EOIS for space applications and how these elements can affect its function. After designing a spacecraft for low orbiting missions during day time, the design of an electro-imaging system becomes an important part in the satellite because the satellite will be able to take images of the regions of interest. An example of an electro-optical satellite imaging system will be presented through this paper where some restrictions have to be considered during the design process. Based on the optics principals and ray tracing techniques the dimensions of lenses and CCD (Charge Coupled Device) detector are changed matching the physical satellite requirements. However, many experiments were done in the physics lab to prove that the resizing of the electro optical elements of the imaging system does not affect the imaging mission configuration. The procedures used to measure the field of view and ground resolution will be discussed through this work. Examples of satellite images will be illustrated to show the ground resolution effects.

Keywords: Satellite Images, Digital Image Processing, Space Technology, Electro-optical Sensors.

1. INTRODUCTION

From many years, the researchers have been interested in the field of imaging detectors like solid-state detectors, CCDs and non-imaging sensors like photo-emissive detectors, image intensifiers. However, in reconnaissance applications a good imagery can be obtained with a photographic camera (panchromatic) under almost any condition clear enough for flying photographic mission at different altitudes. Most of these imaging systems are used with different platforms like airplanes and spacecrafts.

The difference between the photographic camera and the electro-optical imaging system is the light sensor and the storage media. The photographic films have the major advantage of simplicity and low cost, in comparison to other imaging detectors. They can be used for imaging a very large angle of view with very high resolution, so its capability is not yet matched by electronic detectors. However, the quantitative accuracy is not as good as that achieved with electronic sensors. This is because the blackening of the film is nonlinear with integrated radiation flux. In applications where quantitative accuracy and very high sensitivity are important, CCD sensors are used instead of photographic cameras.

Two-dimensional solid-state sensor arrays as CCDs can be used for visible, ultraviolet and x-ray detection, although the devices must be especially prepared for use in the wavelength range of interest. The CCDs arrays are available in format of 4096x4096 pixels and with pixel sizes ranging from about 30 microns to as small as 7 microns. For imaging satellites, the critical requirement for any optical system is to have an imaging system, which will capture images to the earth at different altitudes (150 to 800 kilometers) and during different times (day or day and night). The current generation of electro-optical imaging system uses large scale of framing focal plane arrays sensors. The new products of the imaging systems can be used for visible or infrared imaging alone, or for dual spectral band (visible and IR) operation. Although the imaging systems have better than 0.5

meter ground resolution during day missions, it has poorer ground resolution in the IR band (during night missions). Having this resolution, objects on earth like vehicles, roads, and buildings can be easily recognized in case of visible band. However, the IR images will need more digital processing to recognize the detected objects within each image.

Upgrading the surveying photographic camera type UMK 10/1318 to an electro-optical imaging system EOUMKS was a project in the computer engineering department in collaboration with the laboratories of Geomatics System, and Physics and Astronomy at the University of Calgary [1]. In this project, Somaie and his colleagues succeeded to upgrade the camera type UMK 10/1318 to an electronic one by replacing the film magazine and the shutter by a couple of lens and CCD detector of dimension 8.8x6.6 mm, 0.2 megapixels resolution, and all one chip of type TC241 produced by Texas Instruments for B/W NTSC TV applications [2].

The following section contains the design of the electro-optical imaging system. Section 2.1 includes experiments and results. Examples of satellite images are shown in section 2.2. Section 2.3 contains conclusions and discussions.

2. ELECTRO-OPTICAL IMAGING SYSTEM

In this section, we will present the steps to design a CCD camera in a visible electromagnetic spectrum range for satellite imaging system. The optical design is usually applied under some definitions or parameters and constraints. Most of these parameters are shown in FIGURE 1. However, the constraints of the optical design and CCD detector can be defined as follows:

- (1) The effective focal length should be big enough to give the sufficient magnification for one of the picture elements or pixel to be equivalent to the required target resolution.
- (2) The diameter of the lens aperture must be sufficiently large to meet the Raleigh principles for resolving the angular resolution.
- (3) The F-number of the lens should be sufficiently small to give a bright image.
- (4) Field of view should cover the dimension of the CCD detector.
- (5) The electro optical elements of the satellite imaging system should match the satellite sizing requirements.

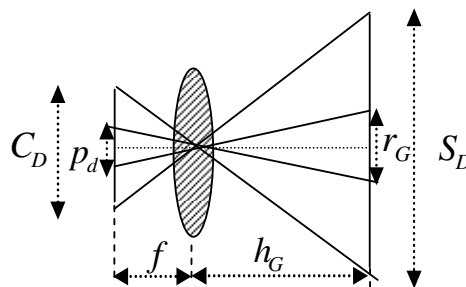


FIGURE 1: The Principal of the Camera.

The following algorithms will explain the design procedures of the electro-optical imaging system that match the requirements to a payload of low orbiting satellites. Assuming the path geometry of the satellite is an elliptical orbit defined by a preliminary data in TABLE 1. So we can start our design with the apogee or perigee region information after investigating the above constraints. The steps of next algorithm will be applied to design the CCD camera of the satellite imaging system.

| Satellite information | Specifications |
|-----------------------|---|
| Orbit | - Elliptical orbit has apogee and perigee points at distances $h_G = h_{apogee} = 400 \text{ km}$, and $h_G = h_{perigee} = 160 \text{ km}$ from the earth respectively. |
| Payload | - Ground resolution $r_G : \leq 1.0 \text{ [m]}$ - Spectral range $\lambda : 0.3\text{-}0.75 \text{ [}\mu\text{m]}$ - Pixel pitch $p_d : 7.5 \text{ [}\mu\text{m]}$ - CCD dimension: $4069 \times 4069 \text{ [pixels]}$ |

TABLE 1: An Example Shows the Preliminary Data of Electro-optical Imaging System.

- (1) The effective focal length of the electro-optical camera, f_0 is calculated as,

$$f_0 = h_G (p_d / r_G) , \quad (1)$$

Where h_G is the satellite altitude, p_d is the pixel pitch or the diagonal of the pixel, and r_G is the object ground resolution. In the apogee region $h_G = h_{apogee}$ and $r_G = r_{G-apogee}$, and in the perigee region $h_G = h_{perigee}$ and $r_G = r_{G-perigee}$.

- (2) The Raleigh criterion determines the value of the angular ground resolution θ_{ang} as,

$$\theta_{ang} = 2 \tan^{-1} (r_G / 2h_G) \quad (2)$$

- (3) Taking a value of the wavelength λ in the visible spectral range, the worst case is the imaging light at $\lambda = 0.75$ microns; then the diameter of the lens aperture ϕ_0 is calculated as,

$$\phi_0 = 1.22 (\lambda / \theta_{ang}) \quad (3)$$

- (4) The F-number of the camera lens or the focal ratio F is calculated as,

$$F = f_0 / \phi_0 , \quad (4)$$

Where the lens aperture ϕ_0 is modified to any standard value (e.g., $\phi_0 = 50 \text{ cm}$) that yield small F-number.

- (5) Using the value of the lens aperture ϕ_0 to be greater than 0.5 m will be positive from gathering light aspects but does not match the satellite dimensions. Based on the optics principals and ray tracing techniques [3], rays usually are concentrated into middle of the lens, then the expected diagonal of image plane or CCD detector is C_{D0} , and $C_{D0} = \phi_0 / 2$. In this case we will find that the diagonal of the CCD detector should equal 25 cm, which is not available yet.

At this stage of designing the satellite imaging system for low orbit reconnaissance purpose, the film screen or CCD detector will be placed at a distance equals f_0 from the original lens L_0 , which does not match the satellite dimensions. As shown from step (5), the expected diagonal of the required CCD detector should be reduced from about 25 cm to about 4.34 cm and moving location of the CCD detector close to the original lens L_0 along the optical axis will match the satellite and CCD detector sizing requirements.

To overcome those problems, two optical lenses L_1 and L_2 were used in addition to the original lens L_0 , and the new CCD detector with pixel resolution of about 17 megapixels, $P_d = 7.5 \mu m$ were lined up to keep the same optical axis for all the elements as shown in FIGURE 2.

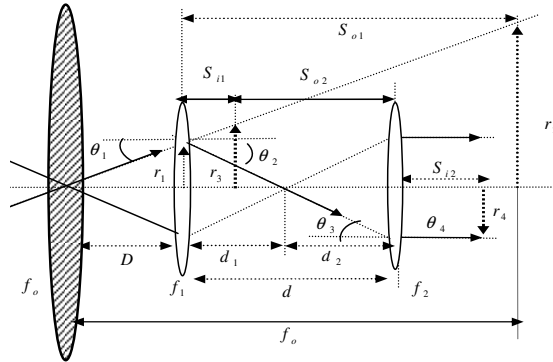


FIGURE 2: The original lens and the two added lenses.

- (6) Using the new CCD detector of diagonal $C_{DN} = 43.4 \text{ mm}$, the field of view θ_{FOV} of the electro-optical imaging system is computed as,

$$\theta_{FOV} = 2 \tan^{-1}(C_{DN} / 2f_0) \quad (5)$$

- (7) The field of view ($\theta_{FOV} = 0.83^\circ$) will cover a ground square area with a side length $S_L = S_D / \sqrt{2}$, where S_D the diagonal of square is calculated as,

$$S_D = 2h_G \tan(\theta_{FOV} / 2) \quad (6)$$

- (8) If a lens has a focal length f_0 of one meter, its refractive power equals to one diopter. It means that the refractive power (P) of the lens is measured in diopter and calculated as,

$$P = 1 / f_0(\text{meter}) \quad (7)$$

The refractive power of a converging and diverging lenses are positive and negative respectively.

- (9) The magnification power ratio of the lens (MPR) is a constant factor depending on how an object or image will appear, if it is examined from a distance of about 25.4 cm , the magnification power ratio of the lens is defined as,

$$MPR = 25.4 / f_0(\text{cm}) \quad (8)$$

- (10) The light gathering or image brightness (IB) is inversely proportional to the square of the F-number and the IB is determined by,

$$IB = (\phi_0 / f_0)^2 \quad (9)$$

So, image brightness is controlled by the focal length f_0 and the lens aperture ϕ_0 . That means the CCD detector will receive a light through an opening hole of radius about $\phi_0 / 2$.

Many experiments were done in the physics lab, where two He-Ne laser beams are considered as two incident rays coming from a far object. Although laboratory He-Ne laser has good beam quality for couple of meters (small divergence and dispersion), once optical components are placed and beam passes through, beam quality significantly drops such that was not good for our optical measurements. Therefore, we used the first diffraction order of beam passing through 100 microns pinhole as a guiding beam. Then with using a fine beam expander to reduce the size of this beam as shown in FIGURE 3. Many experimental trials were done with different lenses that have focal length smaller than the original one. The additional lenses L_1 or L_2 has lens aperture $\phi_1 = \phi_2 = \phi_0 / 4$, and they have the same focal length ($f_1 = f_2 = 70.5mm$). However, the optics and mathematical algorithms are designed to prove the match between the optical components L_0 , L_1 , L_2 and CCD detector.

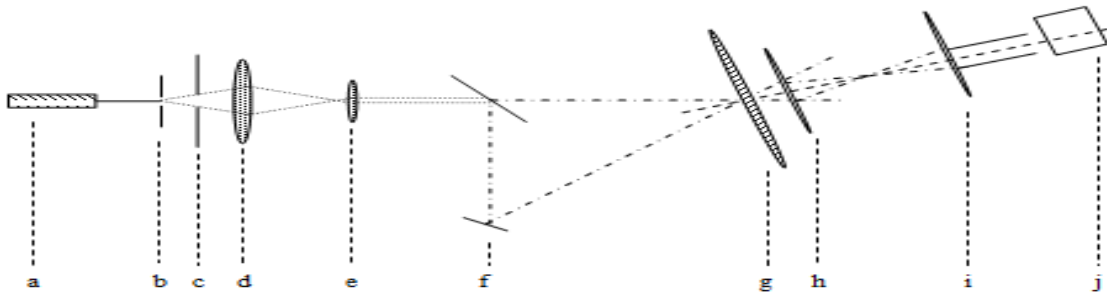


FIGURE 3: Optical layout of the experimental work. Components are: (a) He-Ne Laser source (b) Pinhole (c) Aperture (d and e) Beam Expander (f) Mirror and beam splitter (g) Lens f_0 (h) Lens f_1 (i) Lens f_2 (j) CCD detector.

The optics algorithms [3] that used to prove the match between the original lens L_0 and the two added lenses L_1 and L_2 , and CCD detector based on the ray tracing and the image creations techniques. Since gathering light in a small area is the goal of this work, the ray-tracing method seems to be more convenient way to explain the optical method; behind the practical work. The concept of ray tracing is to use a unique matrix for each optical element like lens, mirror and empty space between the optical elements. In this work, the final beam after the compound lens is assumed collimated beam. However, the incident ray has an angle of some degree, which has to be considered in the calculations. Based on the FIGURE 2, the ray tracing method [3] can be summarized in the following steps,

- (1) The ray of angle θ_1 incident to the lens L_0 of the satellite imaging system is passing to the image plane passing through the lenses L_1 and L_2 as shown in FIGURE 2. Because the half of the aperture of the first lens L_1 is about $\phi_1 / 2$, and $\phi_1 = \phi_0 / 4$ then for simplicity, safety, and the rays concentration $r_1 = \phi_1 / 4 = \phi_0 / 16$, then the distance D between the two lenses L_0 and L_1 is defined as,

$$D = f_0 (r_1 / r_2), \quad (11-a)$$

Where $r_2 = C_{D0} / 2$ and C_{D0} is the diagonal of the original CCD detector ($C_{D0} = \phi_0 / 2$). In this case the above equation can be rewritten as,

$$D = f_0 / 4 \quad (11-b)$$

(2) As shown in FIGURE 2, the minification M_1 of the lens L_1 is calculated as,

$$M_1 = r_1 / r_2 \quad (12)$$

(3) The total minification for the two lenses L_1 and L_2 is calculated as,

$$M_t = C_{DN} / C_{D0} , \quad (13-a)$$

$$M_t = M_1 \times M_2 \quad (13-b)$$

In this example $C_{DN} \approx \phi_0 / 12$, where C_{DN} is the diagonal of the new CCD detector, and M_2 is the minification of the lens L_2 .

(4) The distance d between the two lenses L_1 and L_2 is defined as,

$$d = f_2 (1 + 1/M_2) \quad (14)$$

(5) The equivalent focal length of the two lenses L_1 and L_2 is calculated as,

$$f_{eq} = f_1 f_2 / ((f_1 + f_2) - d) \quad (15)$$

The value of f_{eq} could be positive or negative depends on the separation between the two lenses.

The following steps describe the image algorithm which is mainly used to determine the position of the new CCD detector and to measure the expected minifications of the compound optical components.

(1) Assuming that with adjusting the lens L_0 , the image of objects placed from 4.2 m to ∞ is formed on the film screen at a distance equal f_0 behind the lens L_0 , and then the lens L_1 is placed at a distance D as shown in FIGURE 2. The distance S_{01} between the lens L_1 and the expected image plane of the lens L_0 is calculated as,

$$S_{01} = f_0 - D \quad (16)$$

(2) We considered an imaginary object at the image plane of the lens L_0 , and then the corresponding image shall be at distance S_{i1} from the lens L_1 . If that image is considered as an object to the lens L_2 and at distance S_{o2} from it, then the corresponding image is at distance S_{i2} from the other side of the lens L_2 as shown in FIGURE 2. These distances are calculated as,

$$S_{i1} = (1/f_1 - 1/S_{o1})^{-1} \quad (17-a)$$

$$S_{o2} = d - S_{i1} \quad (17-b)$$

$$S_{i2} = (1/f_2 - 1/S_{o2})^{-1} \quad (17-c)$$

(3) The magnifications \hat{M}_1 , \hat{M}_2 and \hat{M}_t of the lenses L_1 and L_2 are defined as,

$$\hat{M}_1 = S_{i1} / S_{o1} \quad (18-a)$$

$$\hat{M}_2 = S_{i2} / S_{o2} \quad (18-b)$$

$$\hat{M}_t = \hat{M}_1 \times \hat{M}_2 \quad (18-c)$$

The ray tracing model and the image algorithm were applied and the results are illustrated in TABLE 2.

| Ray tracing algorithm | | Image algorithm | |
|-----------------------|--------|------------------------|-----------|
| Parameters | Values | Parameters | Values |
| D [cm] | 75 | S_{o1} [cm] | 225 |
| d [cm] | 17.6 | S_{i1} [cm] | 7.3 |
| f_{eq} [cm] | -14.2 | S_{o2} [cm] | 10.3 |
| M_1 | 1/4 | S_{i2} [cm] | 22.3 |
| M_2 | 2/3 | \hat{M}_1, \hat{M}_2 | 0.03, 2.2 |
| M_t | 1/6 | \hat{M}_t | 0.07 |

TABLE 2: This Table Shows the Results of the Ray Tracing Technique, and the Image Algorithm.

From the above steps, one of the satellite imaging system requirements is to work in visible band, where the wave length $\lambda=0.3-0.75$ microns. The ground resolution r_G and coverage area will change, if the altitude of the satellite is changed and this is clear enough when the path geometry of the satellite is an elliptical orbit. In that case the satellite will go in counter clockwise direction through an elliptical path from perigee point (close to earth) to apogee point (far from earth) as shown in FIGURE 4. It means that the ground resolution of satellite imaging system will be 0.4 m, covering area (square) of length $S_L = 1.6 \text{ km}$ near perigee point rather than the correspondence values (1.0 m and 4.1 km) close to the apogee point. Based on the space model relations [4] given in TABLE 3, and the data of the example at hand defined by TABLE 1, the results are illustrated in TABLE 3. The satellite will fly through its path about 16.0 revolutions per day with orbital period $T_{orbit} = 89.97$ minutes and an average speed $V_{Sat} = 6.98$ km/sec as shown in TABLE 3. If the corresponding ground speed of satellite at perigee is 6.94 km/sec, so the satellite will cover 0.4 m on the ground in 58 μsec . That time represents the shutter rate of the satellite camera which is very short corresponding to the shutter time of the ordinary cameras (33.3 ms), so it puts severe constraints on the satellite imaging system. Related to the above discussion, the exposure time of the camera on board will vary from 58 μsec (around perigee) to 155 μsec (close to apogee) and this should require a very sensitive CCD detector. This requires improving in the quality the optical system (lenses and mirrors) and cooling the CCD sensor to increase its sensitivity which in turn reduces its noise. The quality of images of the camera system depends on gathering a large light through capture process and this gives field of view of the satellite imaging system is 0.83 degree.

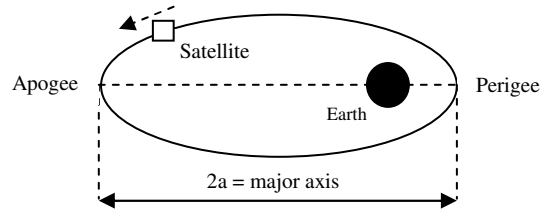


FIGURE 4: An elliptical orbit geometry.

It means the satellite has always needed to point its cone of view (0.415°) towards a target precisely to capture a ground area ($1.6 \times 1.6 \text{ km}^2$) within each snapshot around the perigee region. This requires detailed information about the attitude determination and control system (ADCS) to achieve high degree of accuracy about position of the satellite within its orbital path, and to minimize the dead time of the imaging mission.

| Relations | Values |
|---|--------------|
| If $r_{\text{earth}} = 6371 \text{ [km]}$, then $r_{\text{apogee}} = r_{\text{earth}} + 400 \text{ [km]}$ | 6771 |
| $r_{\text{perigee}} = r_{\text{earth}} + 160 \text{ [km]}$ | 6531 |
| The semi major axis of ellipse: $a = (r_{\text{apogee}} + r_{\text{perigee}}) / 2 \text{ [km]}$ | 6651 |
| If $\mu = 3.986 \times 10^5 \text{ km}^3 / \text{sec}^2$, the orbital period $T_{\text{orbit}} = 2\pi \sqrt{a^3 / \mu} \text{ [minutes]}$ | 89.97 |
| $n = 24 \times 60 / T_{\text{orbit}} \text{ [revolution s/day]}$ | 16.00 |
| $V_{\text{apogee}} = \sqrt{\mu[(2/r_{\text{apogee}}) - (1/a)]} \text{ [km/sec]}$ | 6.85 |
| $V_{\text{perigee}} = \sqrt{\mu[(2/r_{\text{perigee}}) - (1/a)]} \text{ [km/sec]}$ | 7.11 |
| $V_{\text{Sat}} = (V_{\text{apogee}} + V_{\text{perigee}}) / 2 \text{ [km/sec]}$ | 6.98 |
| $V_{G-\text{apogee}} = V_{\text{apogee}} (r_{\text{earth}} / r_{\text{apogee}}) \text{ [km/sec]}$ | 6.45 |
| $V_{G-\text{perigee}} = V_{\text{perigee}} (r_{\text{earth}} / r_{\text{perigee}}) \text{ [km/sec]}$ | 6.94 |
| The time of capture ground square area: $T_{\text{perigee}} = S_L / V_{G-\text{perigee}} \text{ [sec]}$ $T_{\text{apogee}} = S_L / V_{G-\text{apogee}} \text{ [sec]}$ | 0.23 0.64 |

TABLE 3: This Table Shows the Relationships of Satellite's Parameters and the Corresponding Values.

It's apparent that ADCS will be sophisticated to control satellite to fly over the target area in about 0.23 sec in case of perigee region and 0.64 sec close apogee point (TABLE 3), so the satellite would need to be 3-axis controlled to achieve stable platform for the camera on board [5].

2.1 Experiments and Results

Many experiments were done in the physics lab, and outdoor, and the samples of snapshots of the EOIS are shown in FIGURE 5. The most important parameters of the imaging system are the coverage area and the target resolution, both of which depend on the altitude of the satellite. The coverage area can be measured by the field of view, which is the angle between two rays passing through the perspective center (rear nodal point) of a camera lens to the two opposite corners of the CCD detector. The field of view of the camera system used in this example, can be measured easily at the lab in two directions, the horizontal direction (from the left to the right of the frame), $\theta_{\text{FOV}} = 0.83^\circ$ as shown in FIGURE 6, and the same in the vertical direction. Digital image resolution is divided into two basic categories: (1) pixel count resolution; (2) spatial resolution.

Much of the confusion over digital image resolution can be settled by understanding the difference between the two.

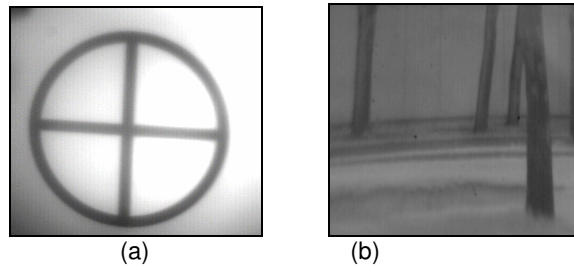


FIGURE 5: (a) and (b) Show Samples of the Lab and Outdoor Images Respectively by using the EOIS.

Spatial resolution relates to the number of pixels in a spatial measurement of a physical image (pixels per inch), while pixel count resolution is the pixel dimensions (in microns) and the number of pixels within the CCD detector. Pixel count resolution and spatial resolution are the two factors that determine the output image size. In the example at hand, the pixel resolution of the camera system is about 17 mega pixels with pixel pitch equals 7.5 microns, and new CCD dimensions $(31 \times 31) \text{ mm}^2$.

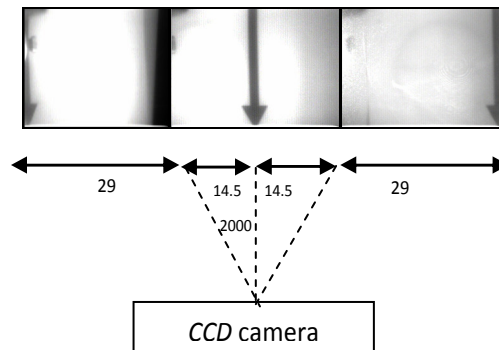


FIGURE 6: Three Snapshots of the Arrow Move in a Horizontal Plane, (all numbers are in cm).

The two terms target resolution and angular resolution, which are basically two equivalent parameters to measure the ground resolution, are another cause for confusion. The target resolution is typically specified in meter or cm, and is equivalent to the picture element or the pixel. The angular resolution or the angle of view is the angle between two rays passing through the perspective center (rear nodal point) of a camera lens to the two opposite corners of the pixel. In practice, if we want to measure the target resolution of a CCD camera at a certain altitude, let us say 400 km, this camera has to capture an image, such as a train track for example. The standard distance between the two rails of a train track is 1 m, so if the image of that train track appears as two lines, it means the target resolution of this camera is one meter at that altitude. It should be made clear that the target resolution differs from the pixel resolution. In other words, the pixel resolution determines whether the pixel dimension (in order of microns) is enough to cover the required target resolution. In the digital image processing lab a resolution pattern chart was designed as three vertical pairs of lines and the same in the horizontal plane as shown in FIGURE 7. The distance between the lines of each pair is 0.022 mm and the resolution pattern chart is fixed in front of the camera with a distance of about 9 m along its optical axis to keep the chart within its field of view. Different snapshots have to be captured to measure the target resolution of the camera system at different distances. If each pair of lines appears as two lines in both directions (vertical and horizontal), then the resolution of the imaging system is identified

correctly and it is said that the pixel resolution of the camera system is about 17 mega pixels with angular target resolution $\theta_{ang} = 2.5 \times 10^{-6}$ radian as shown in FIGURE 7.

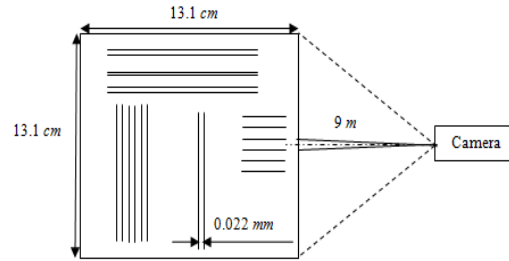


FIGURE 7: The Resolution Pattern Chart is Fixed in Front of the Camera by a Distance 9m.

The above configurations of the EOIS are compared with the specifications of the new electro-optical camera type UMK 10/1318 for the surveying purpose [2] and the results are illustrated in TABLE 4. The ground resolution of the EOIS is superior to the EOUMKS by about 40 percent. However, the first imaging system is expected to operate in space within an elliptical orbit having apogee altitude up to 400 km, and the second imaging system is working on air with altitude up to 4 km. It is obvious that the EOIS has a wide coverage imaging area equal 2.56 km², and 16.81 km² at perigee and apogee regions respectively, and the EOUMKS has a coverage imaging area of about 0.06 km², and 1.04 km² at altitude of 1 km and 4 km in respective. In addition to the parameters shown in TABLE 4, the components of the EOIS is resized and tested to match the space configuration necessities.

| Parameters | EOIS | EOUMKS [2] |
|----------------------------------|--|--|
| Altitude | - (160 ÷ 400) km | - 6.2 m ÷ 4 km |
| Ground resolution | - 0.40 m (at 160 km) - 1.00 m (at 400 km) | - 0.62 m (at 1 km) - 2.50 m (at 4 km) |
| Spectral range | - 0.75 μm | - 0.70 μm |
| Pixel pitch | - 7.5 μm | - 27 μm (v) by 11.5 μm (H) |
| CCD dimension | - 31 by 31 mm - (4069x4069) pixels | - 8.8 by 6.6 mm - (780x244) pixels |
| Focal length of the main lens | - 3.0 m | - 0.2 m |
| Focal length of the added lenses | - 70.5 mm , 70.5 mm | - 49.3 mm , 49.3 mm |
| Field of view | - 0.83 ° | - 12 . 10 ° (V), 17 . 8 ° (H) |
| Coverage area | - (1.6x1.6) km ² (at 160 km) - (4.1x4.1) km ² (at 400 km) | - (0.3x0.2) km ² (at 1 km) - (1.3x0.8) km ² (at 4 km) |
| Purpose | - Reconnaissance | - Survey |

TABLE 4: Illustrates the Comparisons between the EOIS and the EOUMKS.

2.2 Satellite Images

The satellite images have been used as source of data in different fields or applications as: (1) military field as reconnaissance and obtaining data about the specified areas of the earth surface; (2) agriculture as crop area measurements, harvest estimation, soil status definition, fire damage estimation; (3) water and coastal resources as slime boundaries, coastline and shoals definition, floods and inundations overflows mapping, lakes and flood areas registration, pollution and oil spillage monitoring; (4) environment as minerals extraction and land recultivation monitoring, water pollution mapping and monitoring, electric power stations and industrial facilities locality definition; (5) others fields. If the satellite images have satisfied the requirements of the

end user from coverage area to target resolution, the size of these images is still a big challenge to researchers in this field. The solution was to compress the satellite images on board to adapt the bandwidth of the communication link and decompress it when the download process is done successfully. The author presented a satellite image compression network using a back-propagation technique that satisfied a compression ratio 4.2:1 with goodness of fit more than 95% [6]. An example of satellite image that captured by IKONOS to the Pyramids and Sphinx at Egypt at November 11, 1999 as shown in FIGURE 8. FIGURE 9 shows the first image of Egyptian satellite called Egypt Sat-1 that shot Cairo city in July 23, 2007. The image resolution of IKONOS satellite is obviously better than the Egyptian satellite.

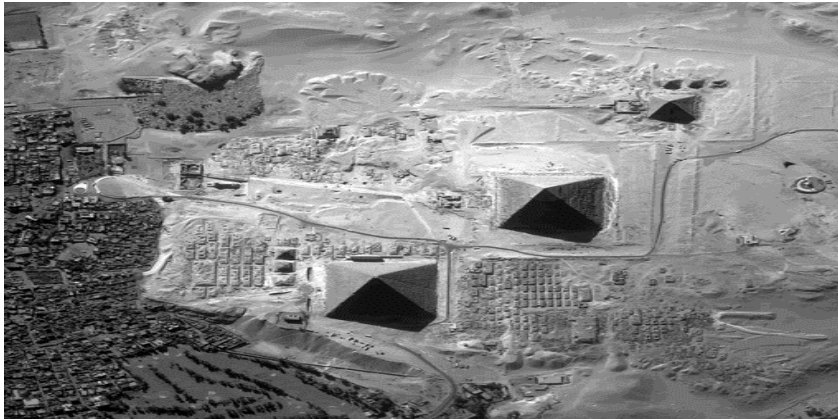


FIGURE 8: IKONOS Shots of the Pyramids and Sphinx at Giza-Egypt with Resolution One Meter at November 11, 1999.



FIGURE 9: Shows the First Image of Egypt Sat-1, Shot of Cairo City in July 23, 2007.

In general, each satellite has a foot print showing its trajectories when it flies within its orbital path. However, the foot print of a satellite should be matched to its imaging mission to take shots of the regions of interest. It is important to mention here that this work illustrates the design and experiments of the EOIS in the lab and outdoor environments but not in board.

2.3 Conclusions and Discussions

The design of the satellite imaging system is not a simple task, since it is considered the main core of the satellite payload. Based on some constraints, a mathematical model to design an electro-optical imaging system is presented for the satellite application. It is important to mention that the average speed of the low orbit satellite is about 9.98 km/sec and it flies around earth

about 16.0 revolutions per day. The sun-synchronicity has an important effect on the imaging satellite since the satellite visits the region of imaging at least once daily. As a result of satellite movement from high to low orbits (as part of its elliptical orbit), the target resolution of the imaging system will be inversely proportional to the altitude of the satellite and it could be one meter or better.

The field of view that covers the imaging area for the region of interest and satisfies the ground resolution is greater than that one covers the standard CCD detectors. In this work, the ray tracing technique was exploited to resize the image plane of the EOIS to match the standard CCD sensors of different megapixel resolution.

Regarding the design procedures that produce the EOIS having a lens of a focal length 3_m matching the reconnaissance requirements, this work still has a proof that the resizing of the electro optical components of the imaging system does not affect the space mission configurations. As shown from the above results, the ground resolution of the EOIS is better rather than that getting from the EOUMKS by about 40 percent. In consequence, the coverage imaging area of the EOIS equals about 16.81 km^2 within the apogee region. However, the coverage imaging area of the EOUMKS is about 1.04 km^2 at altitude up to 4 km . Examples of satellite images are listed to show the effect of the payload upon the low orbiting imaging satellites which varies from one to the other according to the satellite configuration and budget.

Finally, the study presented in this paper belongs to the research category because the objective is to explain the effects of the critical elements and the main parameters on the design of the electro-optical imaging systems in different environments and within multiple of constraints. In this context, it will be great, if more investigations and experiments are applied to this model getting a real system.

2.4 ACKNOWLEDGMENT

I would like to thank the University of Business and Technology UBT, College of Business Administration CBA, Jeddah, KSA, and Dr. Dina AA Somaie for supporting this research.

3. REFERENCES

- [1] AA Somaie. "Optical Sensors and Imaging Systems." Post Doctoral Fellow Program. University of Calgary, Department of Electrical and Computer Engineering, Calgary, Canada, 2005.
- [2] AA Somaie, Naser El-Sheimy, Ziad Abusara, and Mehdi Dehghani. "From the Photographic Camera to an Electro-Optical Imaging System." 5th International Conference on Electrical Engineering, Military Technical College, Cairo, Egypt, May, 2006.
- [3] Eugene Hecht. Optics. Addison Wesley, Fourth Edition, 2002.
- [4] Davidoff M. The Satellite Experimenter's Handbook. New York: McGraw Hill, Second Edition, 1997.
- [5] Sellers JJ. Understanding Space: An Introduction to Astronautics. New York: McGraw Hill, 1994.
- [6] AA Somaie, M. B. Raid, and M. A. El-Bahtity. "Satellite Image Compression Using a Bottleneck Network." Proceedings of CIE International Conference on Radar, 2001, Article No. 374.

Enhanced Spectral Reflectance Reconstruction Using Pseudo-Inverse Estimation Method

Ibrahim El-Rifai
Cairo, 11577, Egypt

ibrahimeg@yahoo.com

Hend Mahgoub
Cairo, Egypt

eng_hend_fci@yahoo.com

Mennat-Allah Magdy
Cairo, Egypt

eng_menna2009@yahoo.com

Jay Arre Toque
Kyoto, Japan

jayarre81@gmail.com

Ari Ide-Ektessabi
Kyoto, Japan

ektessabi@gmail.com

Abstract

This paper will present an enhanced approach for the reconstruction of spectral reflectance by the combination between two methods, the Pseudo-Inverse (PI) as the base formula, whilst adaptively selecting the training samples as performed in the Adaptive Wiener estimation method proposed by Shen and Xin for the estimation of the spectral reflectance. This enhancement will be referred to as Adaptive Pseudo-Inverse (API) through this research.

Training and verification datasets have been prepared from GretagMacbeth ColorChecker CC chart, Kodak Color Chart and a specially designed palette of Japanese organic and inorganic mineral pigments to test and compare the estimation results, using the Pseudo-Inverse and Adaptive Pseudo-Inverse method. The performance of spectral reconstruction methods will be presented in terms of spectral and colorimetric error for the estimation accuracy. The experimental results showed that the proposed method achieved better performance and noticeable decline in spectral estimation error.

Keywords: Adaptive Pseudo-Inverse, Spectral Reflectance Reconstruction, Colorimetry.

1. INTRODUCTION

Some of the main aims of a multispectral system are the efficient extraction of spectral and colorimetric information, and in this regard several algorithms have been developed for best estimation of spectral reflectance and the reproduction of color information.

In this research we will focus on the estimation of spectral reflectance using Pseudo-Inverse (PI) method [1:3]. PI is known to be a traditional and straightforward method while it lack of accuracy and efficiency. PI estimates spectral reflectance without priori knowledge of acquisition system, depends on system responses and reflectance of training samples to get best transformation matrix aiming to minimize the spectral error between the actual reflectance measurement and the estimated one.

More accurate estimation can be extracted from other methods like Wiener estimation method [4:8] which depends on spectral responsivity, spectral reflectance and imaging noise. While the original calculations of Wiener depends on all available training samples, Shen and Xin proposed a modified approach of the original Wiener called Adaptive Wiener [9] which can estimate

spectral reflectance - without priori knowledge about the spectral characteristics of the verification sample - depending on the idea of adaptively selecting the training samples which had a strong impact on the color accuracy of the reflectance reconstruction.

This can be achieved by, firstly, selecting the closest responses of training samples to the verification sample, then calculating the weights of each and finally, recalculating the transformation matrix according to the selected training samples for the verification sample.

As in Shen and Xin approach we will adopt the adaptive selection of training samples to be integrated with Pseudo-Inverse method for the estimation of spectral reflectance which will be examined for estimation accuracy using some datasets.

2. FORMULATION OF MULTISPECTRAL IMAGING

The response of a digital camera can be formulated by this equation [9-10]:

$$V_i(x, y) = \int t_i(\lambda) E(\lambda) S(\lambda) r(x, y; \lambda) d\lambda + n_i(x, y) \quad (1)$$

$i = 1, \dots, m,$

Where $V_i(x, y)$ is the response of the camera in (x, y) coordinate with i th color filters, $t_i(\lambda)$ is the transmission of the i th filter, $E(\lambda)$ is the spectral power distribution of the light, $S(\lambda)$ is the sensitivity of the camera, $r(x, y; \lambda)$ is the reflectance in (x, y) coordinates, and n_i is the additive noise for each channel which is ignored for simplicity and m is the number of channels.

$t_i(\lambda)$, $E(\lambda)$ and $S(\lambda)$ are the unknown factors where they are merged in $F_i(\lambda)$ which is known as spectral responsivity.

Using vector-matrix notation, this equation can be written as follows:

$$V_{(m \times 1)} = F \cdot r_{(K \times 1)} \quad (2)$$

For each pixel in the image, v is the vector of camera response and r is the vector of the reflectance spectrum.

The estimation of reflectance spectra can be obtained by

$$r_{est} = G \cdot V_{verification} \quad (3)$$

where v is the response of the camera and G is the estimation matrix which aims to reduce the minimum square error between original r and estimated r_{est} according to the used estimation method.

Using the traditional Pseudo-Inverse [1-3] to get the estimation matrix G_{PI}

$$G_{PI} = r_{training} \cdot Pinv(V_{training}) \quad (4)$$

$$Pinv(V) = V^t \cdot (V \cdot V^t)^{-1}$$

$r_{training}$ and $V_{training}$ are known from the measured training samples and $Pinv(V_{training})$ is the Pseudo-Inverse of the camera response $v_{training}$. By solving Eq(4) and getting G_{PI} (Estimation Matrix based on Pseudo-Inverse method) we can substitute it in Eq(3) using $V_{verification}$ as system response of verification sample to get the spectral reflectance r_{est} of this sample.

From this point in the research we will integrate the original Pseudo-Inverse with the adaptive approach of Shen and Xin [9]; after getting the estimated spectral reflectance r_{est} of the verification sample based on all training samples Eq(3), we start adaptively to select the training samples according to their spectral similarity to the r_{est} which is calculated from the following equation:

$$d_k = \alpha \text{ mean } \left\{ \left| \frac{r_k}{\|r_k\|} - \frac{r_{est}}{\|r_{est}\|} \right| \right\} + (1 - \alpha) \text{ max } \left\{ \left| \frac{r_k}{\|r_k\|} - \frac{r_{est}}{\|r_{est}\|} \right| \right\}, \quad (5)$$

Where d_k is the spectral similarity between each training sample and the verification sample, α is a scaling factor (= 0.5 in this research), $|x|$ means the absolute values of elements in vector x and finally r_k is the measured spectral reflectance of each training sample. All reflectances are normalized from [0 - 1] and the mean and max spectral distances between two similar reflectances shouldn't be large.

By sorting the training samples in ascending order according to their spectral similarity and selecting L samples of it where $d_1 < d_2 < \dots < d_L$ we get the L samples which are the more close samples to the verification sample.

Recalculating the Eq(4) and getting the new estimation matrix based on the selected L training sample as follows,

$$G_{API} = r_{training}(L) \cdot \text{Pinv} \left(V_{training}(L) \right) \quad (6)$$

Then substituting G_{API} in Eq(3) to get the new spectral reflectance of verification sample r_{est} .

3. EXPERIMENTAL

In this research, the multispectral imaging system comprises Niji Scanner [14] of Kyoto University (Advanced Imaging Technology Laboratory) with a monochrome line camera, led light system of known spectral power distribution and set of 8 band pass and sharp cut filters in the range of 420nm to 700nm. IR cut filter were used throughout the full scanning process which has been done in a dark room. The GretagMacbeth ColorChecker CC (24 patches), Kodak Color Chart (18 patches) and a specially designed palette of Japanese organic and inorganic mineral pigments (173 patches) [15] were collected and prepared to be used as the primary training and verification datasets. The reflectance of color targets were measured by X-Rite spectrophotometer [16] and re-sampled in the range of 400-700nm at 10nm interval. In this experiment, each target is tested separately for training and testing and the spectra has been reconstructed for each patch using different L (number of selected training samples) to see the effect of that on the color and estimation accuracy. Comparisons have been conducted between the PI and the API methods. The results of the comparison have been concluded according to the spectral and colorimetric error, using the mean squared error equation for spectral error between the actual and the estimated spectral reflectance and the equation of ΔE_{00} obtained from the formula of CIEDE2000[11:13] under D65 as standard illuminant for colorimetric error.

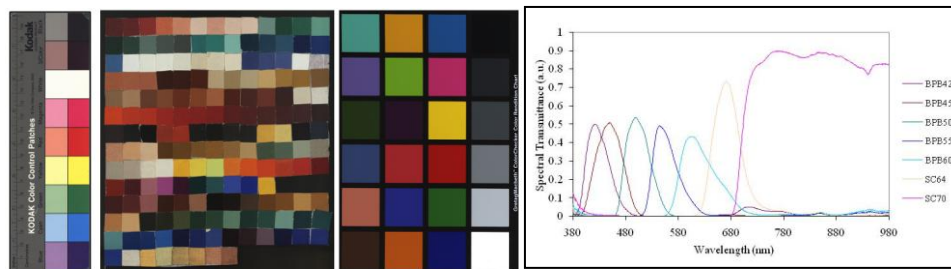


FIGURE 1: Color Targets Left and Spectral Transmittance of the Band Pass Filters Right.

Figure 1 shows the three color targets that have been used; The Kodak Color Chart, the Japanese pigments palette and the GretagMacbeth ColorChecker CC, in addition to the spectral transmittance of the 7 band pass filters Fujifilm BPB42, BPB45, BPB50, BPB55, BPB60, SC64, SC70 in the range from 380 to 980nm.

Each dataset has been investigated with different number of selected training samples (L) for Eq(6) as presented in Figure 2. In case of Kodak and Macbeth targets, which have less number of patches between 18 and 24, L value = 5 was the most appropriate number with the least spectral and colorimetric errors but in the Japanese palette case with 173 patches L value = 10 was the point with the least errors for both spectral and colorimetric. In general datasets shows that both errors are changing in an increase monotonic trend with L .

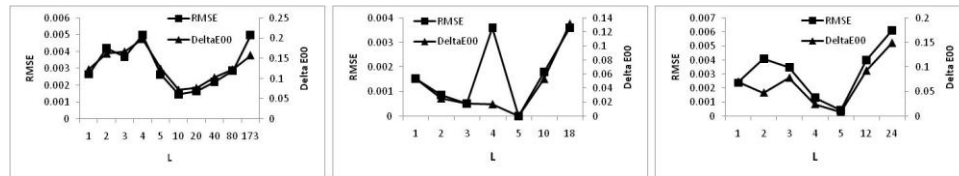


FIGURE 2:..Japanese pigments palette (left) , Kodak chart (middle), Macbeth CC patches (right). The effect of using different L values on the spectral and colorimetric errors for the three datasets.

The spectral rms and colorimetric errors are shown in the Table1 that includes the mean, standard deviation and maximum of the tested methods for the three target datasets. Results of API method are presented according to the most appropriate L value. It is noticed that the results of the proposed method API is showing noticeable improvement over the PI method after applying the adaptive selection on the training samples. Except for the Japanese palette, although the rmse of the API method is less than the PI, the max value of the colorimetric error is showing higher value than the PI method, this is probably the influence of the duplicates of dark and shiny patches. Samples from best and worst cases for the spectral and colorimetric errors are presented in Figure (3a:3c) for each dataset.

A comparison have been presented in Figure 4 showing the colorimetric error between Munsell [18], BabelColor [17] and Estimated API spectral Data for Macbeth ColorChecker CC which are less than 1 and within acceptable error range [23].

| | | | API | PI |
|--------------------------------------|---------------------------------|------|-------------|-------------|
| Kodak Chart (18 patches) | RMSE | Mean | 3.60592E-16 | 0.003603183 |
| | | STD | 5.56468E-16 | 0.001756709 |
| | | Max | 2.35017E-15 | 0.007424878 |
| | ΔE_{00} under D65 | Mean | 1.42027E-14 | 0.128468081 |
| | | STD | 1.54E-14 | 0.070764767 |
| | | Max | 6.27972E-14 | 0.238846758 |
| Macbeth CC (24 patches) | RMSE | Mean | 0.000427617 | 0.006133283 |
| | | STD | 0.001984025 | 0.006493265 |
| | | Max | 0.009728257 | 0.033024449 |
| | ΔE_{00} under D65 | Mean | 0.008956815 | 0.163237335 |
| | | STD | 0.038608483 | 0.192726688 |
| | | Max | 0.188428881 | 0.976090041 |
| Japanese Palette (173 patches) | RMSE | Mean | 0.001460306 | 0.004998975 |
| | | STD | 0.003768597 | 0.00461258 |
| | | Max | 0.028263637 | 0.032854126 |
| | ΔE_{00} under D65 | Mean | 0.071885883 | 0.128116604 |
| | | STD | 0.236286633 | 0.123883428 |
| | | Max | 2.248539644 | 0.680269435 |

TABLE 1: Spectral and Colorimetric Errors for PI and API Methods.

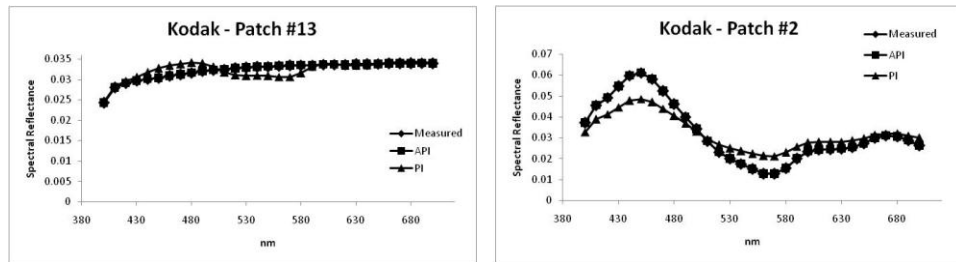


FIGURE 3a: Kodak Chart, Measured and reconstructed spectral reflectance of the API and PI methods (left) best case with spectral error = $2.34\text{E-}17$ and colorimetric error = $6.37\text{E-}17$, (right) worst case with spectral error = $2.35\text{E-}15$ and colorimetric error = $6.28\text{E-}14$.

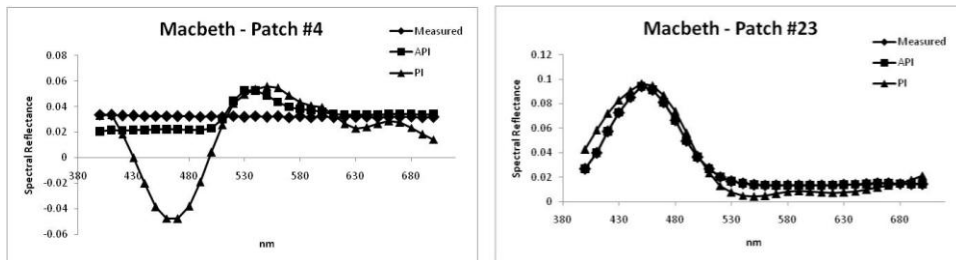


FIGURE 3b: Macbeth Chart, Measured and reconstructed spectral reflectance of the API and PI methods (left) best case with spectral error = $2.22\text{E-}17$ and colorimetric error = $1.59\text{E-}17$, (right) worst case with spectral error = 0.09728 and colorimetric error = 0.188429 .

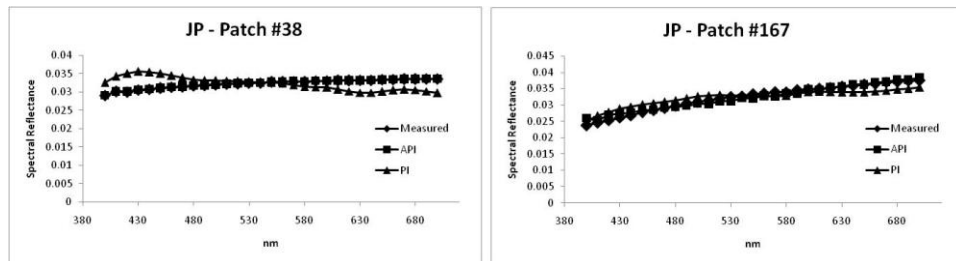


FIGURE 3c: Japanese Palette, Measured and reconstructed spectral reflectance of the API and PI methods (left) best case with spectral error = $1.42\text{E-}05$ and colorimetric error = 0.000713 , (right) worst case with spectral error = 0.028264 and colorimetric error = 2.24854 .

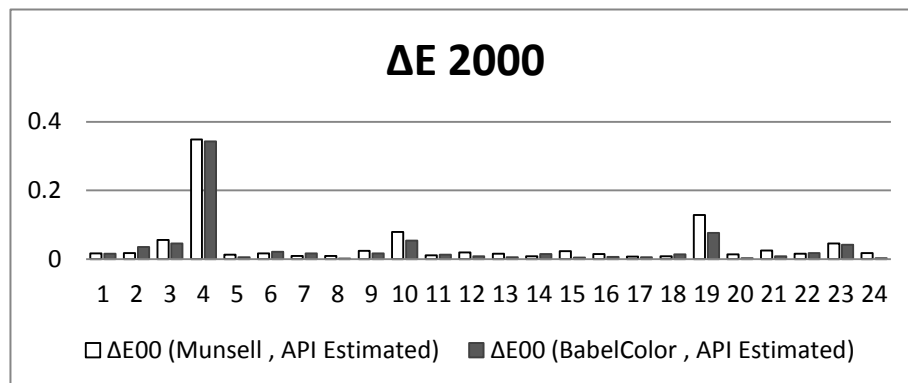


FIGURE 4: Comparison between Munsell, BabelColor and Estimated API spectral Data for Macbeth ColorChecker CC target (24 Patches) using Delta E2000 – colorimetric error.

4. DISCUSSION

Since the acquisition of good quality spectral information is one of the main concerns of color systems, exhaustive research has been conducted for the acquisition, processing and estimation and even evaluation of spectral information. Spectral estimation methods as API depend on the accuracy of the spectral data in the reference database which depends on the spectral properties of the system calibration target [19].

Once the spectral information is acquired either measured by spectrometers, published by standard color targets manufacturers or estimated by spectral estimation methods as in this research, due attention has to be paid for the evaluation of the spectra. Delta E2000 [11:13,20] has been used for the evaluation and comparison of different estimation methods as it is a quantitative evaluation of color differences in LAB color space which matches with the future aims of this research in the reconstruction of color [21].

The performance of the proposed method (Adaptive Pseudo-Inverse) was tested by three different datasets and the results of the spectral estimation have been compared for the original and proposed methods using spectral and colorimetric errors (Table 1). Results showed noticeable improvement of the estimation accuracy. Moreover, the resulted spectra have been compared with other different Spectral data for GretagMacbeth Colorchecker CC according to the colorimetric error (Figure 4) which showed a potential improvement in the quality of the color reconstruction.

5. CONCLUSION AND FURTHER WORK

This research introduced integrated approach for the reconstruction of spectral reflectance by the combination between the original Pseudo-Inverse method and the adaptive selection of samples as stated in Adaptive Wiener introduced by Xin and Shen. By the adaptive selection of training samples according to their spectral similarity to the verification sample, a new transformation matrix has been calculated for the estimation of spectral reflectance which improved the accuracy of the spectral estimation.

Further work would be developed for the comparison and evaluation of API method with other spectral estimation methods such as wiener, adaptive wiener...etc. Also, it is worth mentioning that more investigation about the appropriate L value (number of selected training samples) should be performed to achieve least spectral and colorimetric errors which consequently lead to better accuracy to be used in further research in the development of a portable multi-spectral system for the pigment identification and color reproduction [22].

6. REFERENCES

- [1] J. Barata and M. Hussein. "The Moore-Penrose Pseudoinverse: A Tutorial Review of the Theory" Brazilian Journal of Physics, vol. 42, Issue 1-2, pp. 146-165, Apr. 2012.
- [2] J.Y. Hardeberg. "Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches.". Parkland, Florida: Universal Publishers/dissertation.com, 2001, pp. 304-305.
- [3] V. Cheung, S. Westland, C. Li, J. Hardeberg, and D. Connah. "Characterization of trichromatic color cameras by using a new multispectral imaging technique" Journal of the Optical Society of America A, vol. 22, Issue 7, pp.1231-1240, 2005.
- [4] N. Shimano. "Recovery of spectral reflectances of objects being imaged without prior knowledge." IEEE Transactions on Image Processing, vol.15, Issue 7, pp.1848-1856, Jul 2006.

- [5] H.L. Shen, J.H. Xin and S.J. Shao. "Improved reflectance reconstruction for multispectral imaging by combining different techniques." *Optics Express*, vol. 15, Issue 9, pp. 5531-5536, 2007.
- [6] Y. Murakami, T. Obi, M. Yamaguchi and N. Ohyama. "Nonlinear Estimation of Spectral Reflectance Based on Gaussian Mixture Distribution for Color Image Reproduction." *Applied Optics*, vol. 41, Issue 23, pp. 4840-4847, 2002.
- [7] H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura and Y. Miyake. "System Design for Accurately Estimating the Spectral Reflectance of Art Paintings." *Applied Optics*, vol. 39, Issue 35, pp. 6621-6632, 2000.
- [8] P. Stigell, K. Miyata and M. Hauta-Kasari. "Wiener estimation method in estimating of spectral reflectance from RGB images". *Pattern Recognition and Image Analysis*, vol. 17, Issue 2, p.233-242, Jun 2007.
- [9] H.L. Shen, P.Q. Cai, S.J. Shao and J. Xin. "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation." *Optics Express*, vol.15, Issue 23, pp.15545-15554, 2007.
- [10] H.L. Shen and J.H. Xin. "Spectral characterization of a color scanner based on optimized adaptive estimation." *Journal of the Optical Society of America A*, vol. 23, Issue 7, pp.1566-1569, 2006.
- [11] ASTM International. "Standard Practice for Computing the Colors of Objects by Using the CIE System", ASTM Standard E308-08, West Conshohocken, United States, 2008. www.astm.org.
- [12] H.C. Lee. "Introduction to color imaging science", Cambridge, UK: Cambridge University Press, Feb. 2005.
- [13] M.R. Luo, G. Cui and B. Rigg. "The development of the CIE 2000 colour-difference formula: CIEDE2000." *Color Research and Application (Special Issue on Color Difference)*, vol. 26, Issue 5, PP 340-350, Oct. 2001.
- [14] A. Ide-Ektessabi. "A System for High resolution Digitizing, Non-Destructive Analyses and Secure-Dynamic Display of Cultural Heritage" Microsoft Research Asia, E-Heritage, 2008.
- [15] J.A. Toque. "Materials Investigation through High-resolution Analytical Imaging and Spectrometry" Ph.D. Thesis, Faculty of Engineering, Kyoto university, 2010.
- [16] X-Rite, Incorporated. SP60 Series, Sphere Spectrophotometer-Getting Started, X-Rite, Incorporated, Germany, 2011.
- [17] BabelColor. "The ColorChecker" Internet: www.babelcolor.com/main_level/ColorChecker.htm#ColorChecker_data, Apr., 2012 [Jun. 15, 2013].
- [18] Munsell Color Science Laboratory. "Spectral Data for Commonly Used Color Products: Macbeth ColorChecker" Internet: www.cis.rit.edu/research/mcsl2/online/CIE/MacbethColorChecker.xls, 2010 [Jun. 15, 2013].
- [19] R.S. Berns and F.H. Imai. "Pigment Identification of Artist Materials Via Multi-Spectral Imaging" in *Proc. Color Imaging Conference*, 2001, pp. 85-90.

- [20] L. Lovisolo and R.C.C. de Souza. "Improvement of Objective Image Quality Evaluation, Applying Colour Differences in the CIELAB Colour Space". Computer Science Journals, International Journal of Image Processing (IJIP), Vol.5, Issue 3, 2011.
- [21] Bruce Justin Lindbloom. "Useful Color Equations" Internet: www.brucelindbloom.com, Apr. 20, 2003 [Jun. 15, 2013].
- [22] J.A. Toque, Y. Murayama and A. Ide-Ektessabi. "Pigment Identification based on spectral reflectance reconstructed from RGB images for cultural heritage investigations" in Proc. SPIE 7531, Computer Vision and Image Analysis of Art, 2010.
- [23] HunterLab. "The basics of color perception and measurements: What is an acceptable color difference?" Internet: www.hunterlab.com/pdf/color.pdf, 2001 [Jun. 15, 2013].

A Novel Multiple-kernel based Fuzzy c-means Algorithm with Spatial Information for Medical Image Segmentation

Nookala Venu

Research Scholar

Department of Electronics and Communication Engineering

Sri Venkateswara University College of Engineering

Sri Venkateswara University, Tirupati-517502, India

venunookala@gmail.com

Dr. B. Anuradha

Professor

Department of Electronics and Communication Engineering

Sri Venkateswara University College of Engineering

Sri Venkateswara University, Tirupati-517502, India

anubhuma@yahoo.com

Abstract

Fuzzy c-means (FCM) algorithm has proved its effectiveness for image segmentation. However, still it lacks in getting robustness to noise and outliers, especially in the absence of prior knowledge of the noise. To overcome this problem, a generalized a novel multiple-kernel fuzzy c-means (FCM) (NMKFCM) methodology with spatial information is introduced as a framework for image-segmentation problem. The algorithm utilizes the spatial neighborhood membership values in the standard kernels are used in the kernel FCM (KFCM) algorithm and modifies the membership weighting of each cluster. The proposed NMKFCM algorithm provides a new flexibility to utilize different pixel information in image-segmentation problem. The proposed algorithm is applied to brain MRI which degraded by Gaussian noise and Salt-Pepper noise. The proposed algorithm performs more robust to noise than other existing image segmentation algorithms from FCM family.

Keywords: FCM, Image Segmentation, Gaussian Kernal, Fuzzy, Multiple-Kernal.

1. INTRODUCTION

Image segmentation is one of the first and most important tasks in image analysis and computer vision. In the literature, various methods have been proposed for object segmentation and feature extraction, described in [1] and [2]. However, the design of robust and efficient segmentation algorithms is still a very challenging research topic, due to the variety and complexity of images. Image segmentation is defined as the partitioning of an image into non overlapped, consistent regions which are homogeneous in respect to some characteristics such as intensity, color, tone, texture, etc. The image segmentation can be divided into four categories: thresholding, clustering, edge detection and region extraction. In this report, a clustering method for image segmentation will be considered.

The application of image processing techniques has rapidly increased in recent years. Nowadays, capturing and storing of medical images are done digitally. Image segmentation is to partition image to different regions based on given criteria for future process. Medical image segmentation is a key task in many medical applications. There are lots of methods for automatic and semi automatic image segmentation, though, most of them fail in unknown noise, poor image contrast, and weak boundaries that are usual in medical images. Medical images mostly contain complicated structures and their precise segmentation is necessary for clinical diagnosis [3].

Magnetic resonance imaging (MRI) is a very popular medical imaging technique, mainly because of its high resolution and contrast, which represent great advantage above other diagnostic imaging modalities. Besides all these good properties, MRI also suffers from three Considerable obstacles: noises (mixture of Gaussian and impulse noises), artifacts, and intensity in homogeneity [4].

In recent literatures, several approaches are there for medical image segmentation. The available segmentation methods in literature for medical images are: thresholding approaches, clustering approaches, classifiers, region growing approaches, Artificial Neural Networks (ANNs), deformable models; Markov Random Field (MRF) models atlas-guided approaches and so on. Amongst the above said methods, in medical imaging research clustering based approaches perceived a great focus of interest.

Clustering is a process for classifying objects or patterns in such a way that samples of the same cluster are more similar to one another than samples belonging to different clusters. There are two main clustering strategies: the hard clustering scheme and the fuzzy clustering scheme. Forgy and MacQueen [5] proposed K-means clustering algorithm. K-means is one of the hard clustering method. The conventional hard clustering methods classify each point of the data set just to one cluster. As a consequence, the results are often very crisp, i.e., in image clustering each pixel of the image belongs just to one cluster. However, in many real situations, issues such as limited spatial resolution, poor contrast, overlapping intensities, noise and intensity in homogeneities reduce the effectiveness of hard (crisp) clustering methods. Fuzzy set theory [7] has introduced the idea of partial membership, described by a membership function. Fuzzy clustering, as a soft segmentation method, has been widely studied and successfully applied in image clustering and segmentation [8]–[13]. Among the fuzzy clustering methods, fuzzy c-means (FCM) algorithm [14] is the most popular method used in image segmentation because it has robust characteristics for ambiguity and can retain much more information than hard segmentation methods [15]. Although the conventional FCM algorithm works well on most noise-free images, it is very sensitive to noise and other imaging artifacts, since it does not consider any information about spatial context.

To compensate this drawback of FCM, a pre-processing image smoothing step has been proposed in [13], [16], and [17]. However, by using smoothing filters important image details can be lost, especially boundaries or edges. Moreover, there is no way to control the trade-off between smoothing and clustering. Thus, many researchers have incorporated local spatial information into the original FCM algorithm to improve the performance of image segmentation [9], [15], [18].

Tolias and Panas [9] developed a fuzzy rule-based scheme called the ruled-based neighborhood enhancement system to impose spatial constraints by post processing the FCM clustering results. Noordam et al. [10] proposed a geometrically guided FCM (GG-FCM) algorithm, a semi-supervised FCM technique, where a geometrical condition is used determined by taking into account the local neighborhood of each pixel.

Pham [19] modified the FCM objective function by including spatial penalty on the membership functions. The penalty term leads to an iterative algorithm, which is very similar to the original FCM and allows the estimation of spatially smooth membership functions.

Ahmed et al. [13] proposed FCM_S where the objective function of the classical FCM is modified in order to compensate the intensity in homogeneity and allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood. One disadvantage of FCM_S is that the neighborhood labeling is computed in each iteration step, something that is very time-consuming. Chen and Zhang [16] proposed FCM_S1 and FCM_S2, two variants of FCM_S algorithm in order to reduce the computational time. These two algorithms introduced the extra mean and median-filtered image, respectively, which can be computed in advance, to replace the neighborhood

term of FCM_S. Thus, the execution times of both FCM_S1 and FCM_S2 are considerably reduced.

Afterward, Chen and Zhang [16] improved the FCM_S objective function to more likely reveal inherent non-Euclidean structures in data and more robustness to noise. They then replaced the Euclidean distance by a kernel-induced distance and proposed kernel versions of FCM with spatial constraints, called KFCM_S1 and KFCM_S2. However, the main drawback of FCM_S and its variants FCM_S1 and FCM_S2 and KFCM_S1 and KFCM_S2 is that their parameters heavily affect the final clustering results.

Szilagyi et al. [17] proposed the enhanced FCM (EnFCM) algorithm to accelerate the image segmentation process. The structure of the EnFCM is different from that of FCM_S and its variants. First, a linearly-weighted sum image is formed from both original image and each pixel's local neighborhood average gray level. Then clustering is performed on the basis of the gray level histogram instead of pixels of the summed image. Since, the number of gray levels in an image is generally much smaller than the number of its pixels, the computational time of EnFCM algorithm is reduced, while the quality of the segmented image is comparable to that of FCM_S [17].

Cai et al. [20] proposed the fast generalized FCM algorithm (FGFCM) which incorporates the spatial information, the intensity of the local pixel neighborhood and the number of gray levels in an image. This algorithm forms a nonlinearly-weighted sum image from both original image and its local spatial and gray level neighborhood. The computational time of FGFCM is very small, since clustering is performed on the basis of the gray level histogram. The quality of the segmented image is well enhanced [20].

In this paper, an multi-kernel based FCM (MKFCM) algorithm with spatial information has been proposed using two Gaussian kernels in place of single kernel. Further, the membership values are modified by using their neighbors. The modified membership values are more robust noise images. The effectiveness of the proposed method is tested on four sample MRI brain images under different noise conditions and proved that the proposed algorithm is more robust as compared to FCM family algorithms.

The organization paper is: In section I, a brief review of image segmentation given. A concise review of FCM, KFCM, GKFCM and MKFCM is visualized in section II. The proposed MKFCM with spatial biasing is presented in section III. Further, experimental results and discussions to support the algorithm can be seen in section IV. Conclusions are derived in section V.

2. FUZZY C-MEAN ALGORITHMS (FCM)

2.1 Standard FCM Algorithm

A fuzzy set-theoretic model provides a mechanism to represent and manipulate uncertainty within an image. The concept of fuzzy sets in which imprecise knowledge can be used to define an event. A number of fuzzy approaches for image segmentation are available. Fuzzy C-means is one of the well-known clustering techniques.

Fuzzy c-means clustering algorithm a generalization of the hard c-means algorithm yields extremely good results in image region clustering and object classification. As in hard k-means algorithm, Fuzzy C-means algorithm is based on the minimization of a criterion function.

Suppose a matrix of n data elements (image pixels), each of size $s(s=1)$ is represented as $X = (x_1, x_2, \dots, x_n)$. FCM establishes the clustering by iteratively minimizing the objective function given in Eq. (1).

$$\text{Objective function: } O_m(U, C) = \sum_{i=1}^c \sum_{j=1}^n U_{ij}^m D^2(x_j, C_i) \quad (1)$$

$$\text{Constraint: } \sum_{i=1}^c U_{ij} = 1; \quad \forall j$$

Where, U_{ij} is membership of the j^{th} data in the i^{th} cluster C_i , m is fuzziness of the system ($m=2$) and D is the distance between the cluster center and pixel.

FCM algorithm

Figure 1 shows the flow chart of FCM algorithm and the implementation steps are given below:

Input: Raw image; Output: Segmented image;

- Initialize the cluster centers C_i ($c = 3$ clusters).
- Calculate the distance D between the cluster center and pixel by using eq. (2).

$$D^2(x_j, C_i) = \|x_j - C_i\|^2 \quad (2)$$

- Calculate the membership values by using Eq. (3).

$$U_{ij} = \frac{(D(x_j, C_i))^{-1/(m-1)}}{\sum_{k=1}^c (D(x_j, C_k))^{-1/(m-1)}} \quad (3)$$

- Update the cluster centers using Eq. (4).

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m x_j}{\sum_{j=1}^n U_{ij}^m} \quad (4)$$

- The iterative process starts:
 1. Update the membership values U_{ij} by using Eq. (3).
 2. Update the cluster centers C_i by using Eq. (4).
 3. Update the distance D using Eq. (2).
 4. If $|C_{\text{new}} - C_{\text{old}}| > \epsilon$; ($\epsilon = 0.001$) then go to step1
 5. Else stop

Assign each pixel to a specific cluster for which the membership is maximal

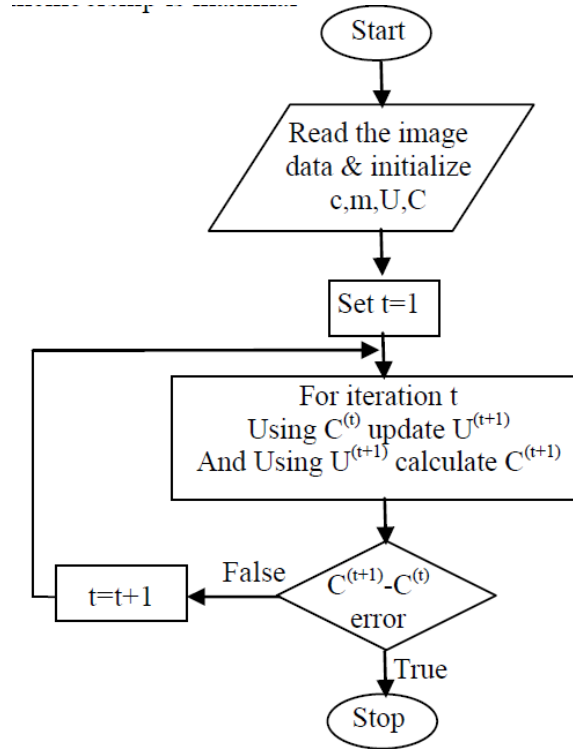


FIGURE 1: Standard FCM Flowchart.

2.2 Kernel Based FCM

Kernel version of the FCM algorithm and its objective function with the mapping as follows:

$$\text{Objective function: } O_m(U, C) = \sum_{i=1}^c \sum_{j=1}^n U_{ij}^m (1 - K(x_j, C_i)) \quad (5)$$

Thus, the update equations for the necessary conditions for minimizing $O_m(U, C)$ are as follows:

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m K(x_j, C_i) x_j}{\sum_{j=1}^n U_{ij}^m K(x_j, C_i)}; \quad i = 1, 2, \dots, C \quad (6)$$

$$U_{ij} = \frac{(1 - K(x_j, C_i))^{-1/(m-1)}}{\sum_{k=1}^c (1 - K(x_j, C_k))^{-1/(m-1)}}; \quad i = 1, 2, \dots, C \quad (7)$$

We point out that the necessary conditions for minimizing $O_m(U, C)$ are update Eqs. (6) and (7) only when the kernel function K is chosen to be the Gaussian function with $K(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma^2)$. Different kernels can be chosen by replacing the Euclidean distance for different purposes. However, a Gaussian kernel is suitable for clustering in which it can actually induce the necessary conditions. The above proposed KFCM algorithm is very sensitive to the noise. To address this problem Chen and Zhang [16] have proposed the KFCM_S1 and KFCM_S2 algorithms which are utilized the spatial neigh pixel information by introduce α parameter.

2.3 Gaussian Kernal FCM (GKFCM)

It is mentioned that the parameter α is used to control the effect of the neighbors for adjusting the spatial bias correction term. In fact, the parameter α heavily affects the clustering results of KFCM_S1 and KFCM_S2. Intuitively, it would be better if we can adjust each spatial bias correction term separately for each cluster i . That is, the overall parameter α is better replaced with $\hat{\eta}_i$ that is correlated to each cluster i . In this sense, Miin-Shen and Hsu-Shen [21] have considered the following modified objective function $O_m^G(U, C)$ with

$$O_m^G(U, C) = \sum_{i=1}^c \sum_{j=1}^n U_{ij}^m (1 - K(x_j, C_i)) + \sum_{i=1}^c \sum_{j=1}^n \eta_i U_{ij}^m (1 - K(\bar{x}_j, C_i)) \quad (8)$$

where $K(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma^2)$, \bar{x}_j is the mean of the neighbor pixels, σ^2 is the variance of the total image.

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m (K(x_j, C_i)x_j + \eta_i K(\bar{x}_j, C_i)\bar{x}_j)}{\sum_{j=1}^n U_{ij}^m (K(x_j, C_i) + \eta_i K(\bar{x}_j, C_i))}; \quad i = 1, 2, \dots, C \quad (9)$$

$$U_{ij} = \frac{\left((1 - K(x_j, C_i)) + \eta_i (1 - K(\bar{x}_j, C_i)) \right)^{-1/(m-1)}}{\sum_{k=1}^c \left((1 - K(x_j, C_k)) + \eta_k (1 - K(\bar{x}_j, C_k)) \right)^{-1/(m-1)}}; \quad i = 1, 2, \dots, C, \quad j = 1, 2, \dots, n \quad (10)$$

2.4 Multiple-Kernal Based FCM (MKFCM)

KFCM_S1, KFCM_S2 and GKFCM methods are utilized only one kernel (Gaussian) function but the multiple-kernel methods provide us a great tool to fuse information from different sources [22]. To clarify that, Long et. al [23] used the term “multiple kernel” in a wider sense than the one used in machine learning community. In the machine learning community, “multiple-kernel learning” refers to the learning using an ensemble of basis kernels (usually a linear ensemble), whose combination is optimized in the learning process. The Eq. (11) and (12) are modified as follows.

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m K_M(x_j, C_i)x_j}{\sum_{j=1}^n U_{ij}^m K_M(x_j, C_i)}; \quad i = 1, 2, \dots, C \quad (11)$$

$$U_{ij} = \frac{\left(1 - K_M(x_j, C_i) \right)^{-1/(m-1)}}{\sum_{k=1}^c \left(1 - K_M(x_j, C_k) \right)^{-1/(m-1)}}; \quad i = 1, 2, \dots, C, \quad j = 1, 2, \dots, n \quad (12)$$

Where $K_M(x_j, C_i) = K_1(x_j, C_i) \times K_2(x_j, C_i)$,

$$K_1(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma_1^2)$$

$$K_2(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma_2^2).$$

3. MKFCM WITH SPATIAL BIASING

The proposed MKFCM does not consider the spatial neighbor pixel information. Hence, MKFCM is very sensitive for the noise image segmentation. To address the effect of noise in image

segmentation, in this paper, a generalized a novel multiple-kernel fuzzy c-means (FCM) (NMKFCM) methodology with spatial information is introduced. In this paper, NMKFCM is represented as MKFCM_S1 and MKFCM_S2 for reader's clarity (see in Figure 2). The objective function, cluster centers and membership functions for the proposed method are given bellow.

$$O_m^G(U, C) = \sum_{i=1}^c \sum_{j=1}^n U_{ij}^m (1 - K_M(x_j, C_i)) + \sum_{i=1}^c \sum_{j=1}^n \eta_i U_{ij}^m (1 - K_M(\bar{x}_j, C_i)) \quad (13)$$

Where $K_M(x_j, C_i) = K_1(x_j, C_i) \times K_2(x_j, C_i)$,

$$K_1(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma_1^2),$$

$$K_2(x_j, C_i) = \exp(-\|x_j - C_i\|^2 / \sigma_2^2).$$

\bar{x}_j is the mean for MKFCM_S1 and median for MKFCM_S2 of the neighbor pixels, σ_1^2 , σ_2^2 are the variances.

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m (K_M(x_j, C_i)x_j + \eta_i K_M(\bar{x}_j, C_i)\bar{x}_j)}{\sum_{j=1}^n U_{ij}^m (K_M(x_j, C_i) + \eta_i K_M(\bar{x}_j, C_i))}; \quad i = 1, 2, \dots, C \quad (14)$$

$$U_{ij} = \frac{\left((1 - K_M(x_j, C_i)) + \eta_i (1 - K_M(\bar{x}_j, C_i)) \right)^{-1/(m-1)}}{\sum_{k=1}^c \left((1 - K_M(x_j, C_k)) + \eta_i (1 - K_M(\bar{x}_j, C_k)) \right)^{-1/(m-1)}}, \quad i = 1, 2, \dots, C, \quad j = 1, 2, \dots, n \quad (15)$$

$$\eta_i = \frac{\min_{i \neq l} (1 - K(C_i, C_l))}{\min_k (1 - K(C_k, x))}; i = 1, 2, \dots, C \quad (16)$$

The proposed method is robust to noise for image segmentation application and the same has been proved from experimental results and discussion in section 4.

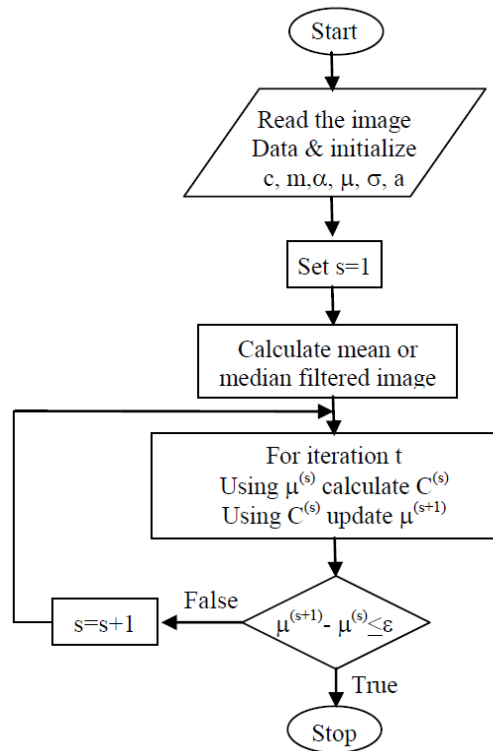


FIGURE 2: Flowchart of MKFCM with spatial biasing.

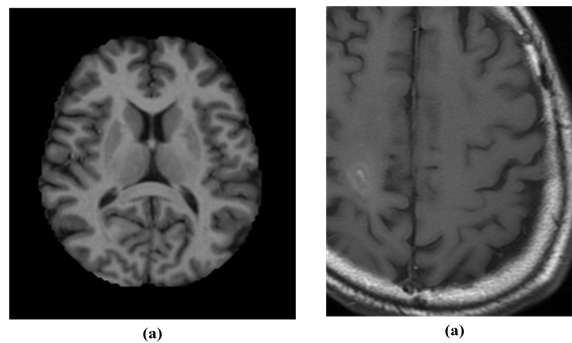


FIGURE 3: Sample images used for experiments

TABLE 1: MRI data acquisition details [26]

| Sequence | MP-RAGE |
|---------------------|----------|
| TR (msec) | 9.7 |
| TE (msec) | 4.0 |
| Flip angle (o) | 10 |
| TI (msec) | 20 |
| TD (msec) | 200 |
| Orientation | Sagittal |
| Thickness, gap (mm) | 1.25, 0 |
| Resolution (pixels) | 176×208 |

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to verify the effectiveness of the proposed algorithm, experiments were conducted on four brain MRIs [24] to compare the performance of the proposed algorithm with other existing methods.

The Open Access Series of Imaging Studies (OASIS) [24] is a series of magnetic resonance imaging (MRI) dataset that is publicly available for study and analysis. This dataset consists of a cross-sectional collection of 421 subjects aged 18 to 96 years. The MRI acquisition details are given in Table 1. The performance of the proposed method is evaluated in terms of score, number of iterations and time. Figure 3 illustrates the sample images selected for experimentation.

4.1. Score Calculation

For comparing segmentation results of different algorithms with a quantitative measure, we use the comparison score defined in [25] and [26]. The comparison score S_{ik} was defined as:

$$S_{ik} = \frac{A_{ik} \cap A_{refk}}{A_{ik} \cup A_{refk}} \quad (17)$$

Where A_{ik} represents the set of pixels belonging to the k^{th} class found by the i^{th} algorithm and A_{refk} represents the set of pixels belonging to the k^{th} class in the reference segmented image.

4.2. Experimental Results with Gaussian Noise

Figure 4 and 6 are the segmentation results obtained with 20% Gaussian noise. In Figure 4, the full size of sample images is used and in Figure 6, the region of interest (ROI) based image segmentation. Tables 2-5 summarize the performance of various methods with different Gaussian and salt & pepper noise. The performance of the proposed methods (MKFCM_S1 and MKFCM_S2) is compared to KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM. The performance of the methods is evaluated on the basis of score. From Tables 2-5 and Figures 3 and 6, MKFCM_S1 is showing better performance as compared to other existing methods (KFCM, GKFCM and MKFCM) in terms of score.

TABLE 2: Comparison of various technics interms of score on Figure (a) at differnt Gaussian noise
noise
C_i: Cluster

| | Gaussian Noise (%) | | | | | | | | | | | |
|-----------------|--------------------|------|------|------|------|------|------|------|------|------|------|------|
| | 5% | | | 10% | | | 15% | | | 20% | | |
| | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 |
| KFCM-S1 | 0.50 | 0.69 | 0.81 | 0.50 | 0.66 | 0.84 | 0.50 | 0.63 | 0.82 | 0.48 | 0.59 | 0.79 |
| KFCM-S2 | 0.45 | 0.66 | 0.86 | 0.47 | 0.62 | 0.83 | 0.45 | 0.58 | 0.80 | 0.44 | 0.55 | 0.77 |
| GKFCM-S1 | 0.42 | 0.60 | 0.83 | 0.41 | 0.53 | 0.77 | 0.44 | 0.49 | 0.73 | 0.42 | 0.46 | 0.71 |
| GKFCM-S2 | 0.40 | 0.59 | 0.82 | 0.41 | 0.51 | 0.76 | 0.41 | 0.47 | 0.72 | 0.41 | 0.44 | 0.69 |
| MKFCM | 0.45 | 0.51 | 0.75 | 0.39 | 0.41 | 0.67 | 0.36 | 0.36 | 0.63 | 0.34 | 0.33 | 0.58 |
| MKFCM-S1 | 0.54 | 0.69 | 0.88 | 0.55 | 0.66 | 0.84 | 0.54 | 0.64 | 0.83 | 0.56 | 0.61 | 0.81 |
| MKFCM-S2 | 0.46 | 0.66 | 0.86 | 0.46 | 0.62 | 0.82 | 0.48 | 0.59 | 0.81 | 0.48 | 0.56 | 0.78 |
| | | | | | | | | | | | | |

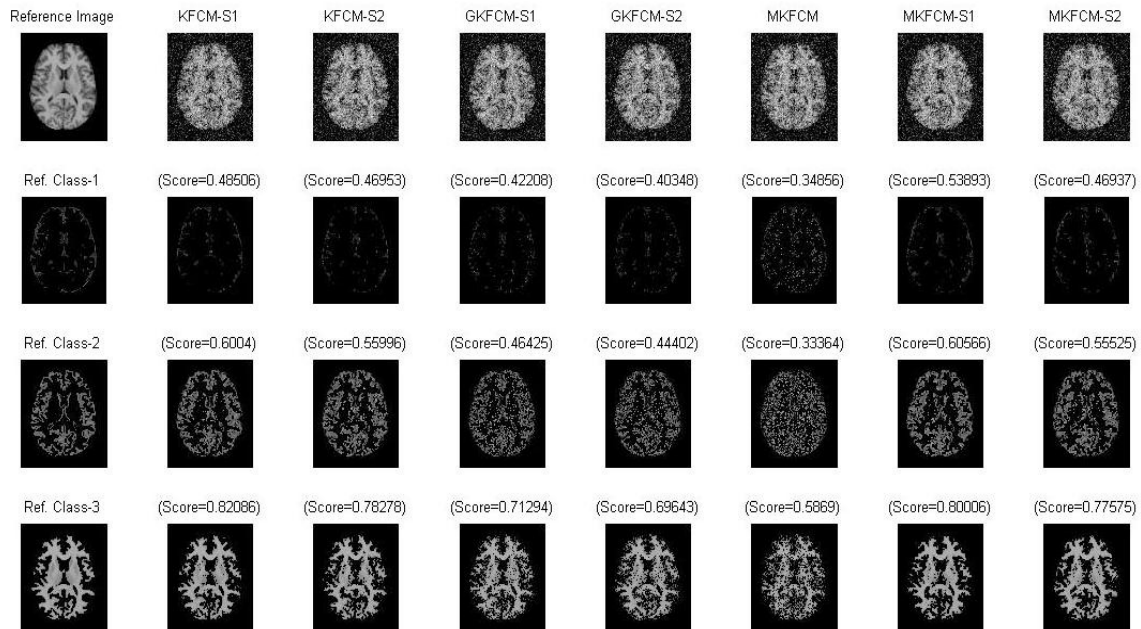


FIGURE 4: Comparison of proposed methods (MKFCM_S1 and MKFCM_S2) with other existing methods (KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM) in terms of score. The original images are corrupted by Gaussian noise.

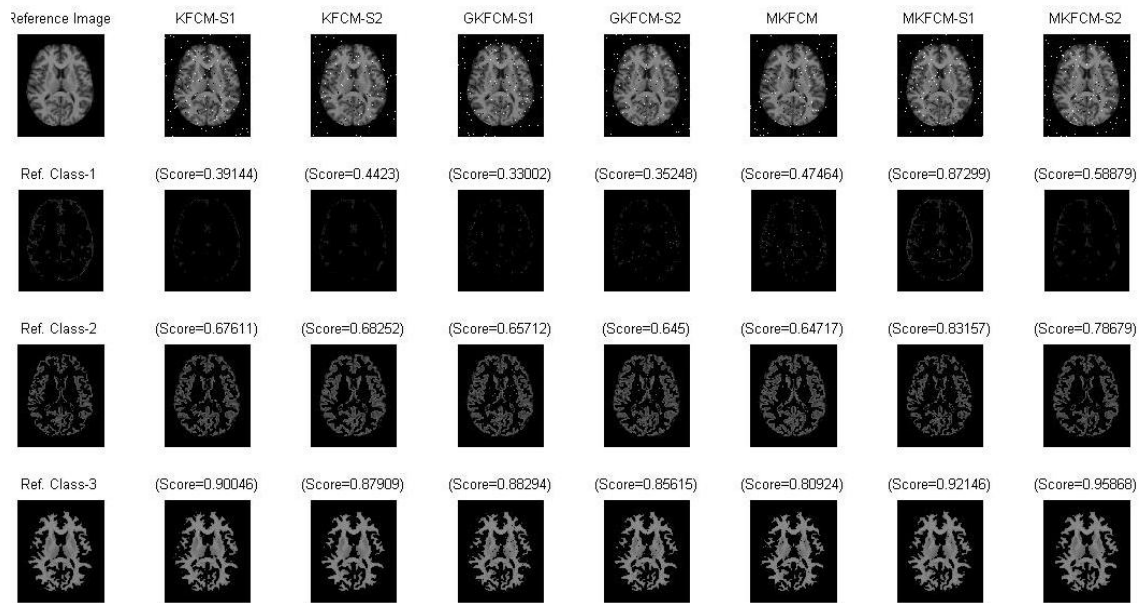


FIGURE 5: Comparison of proposed methods (MKFCM_S1 and MKFCM_S2) with other existing methods (KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM) in terms of score. The original images are corrupted by salt & Pepper noise.

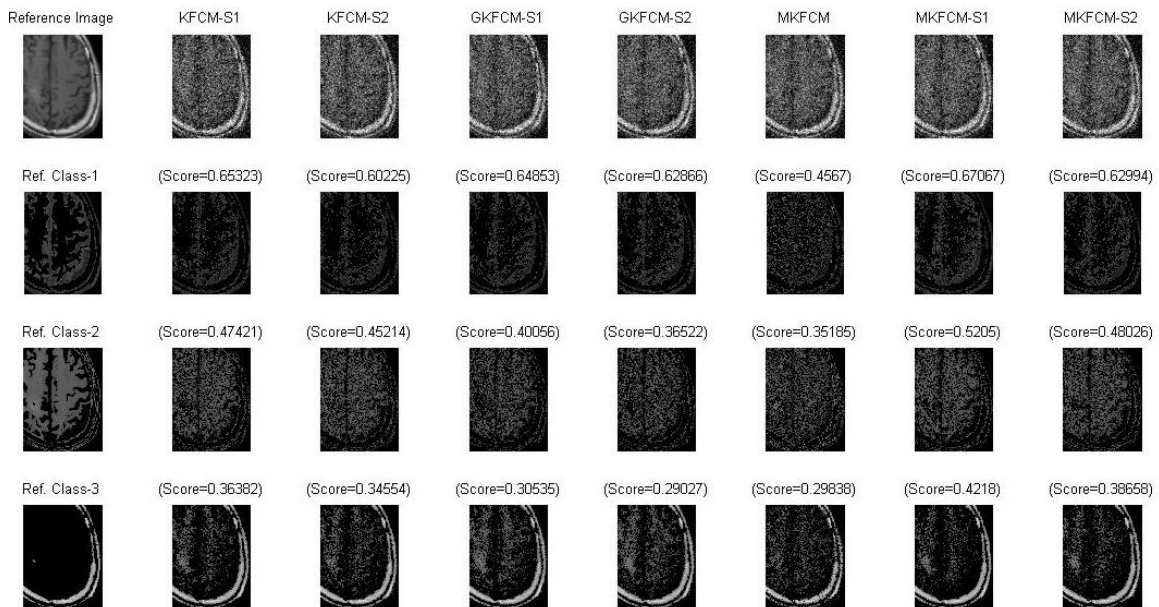


FIGURE 6: Comparison of proposed methods (MKFCM_S1 and MKFCM_S2) with other existing methods (KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM) in terms of score. The original images are corrupted by Gaussian noise.

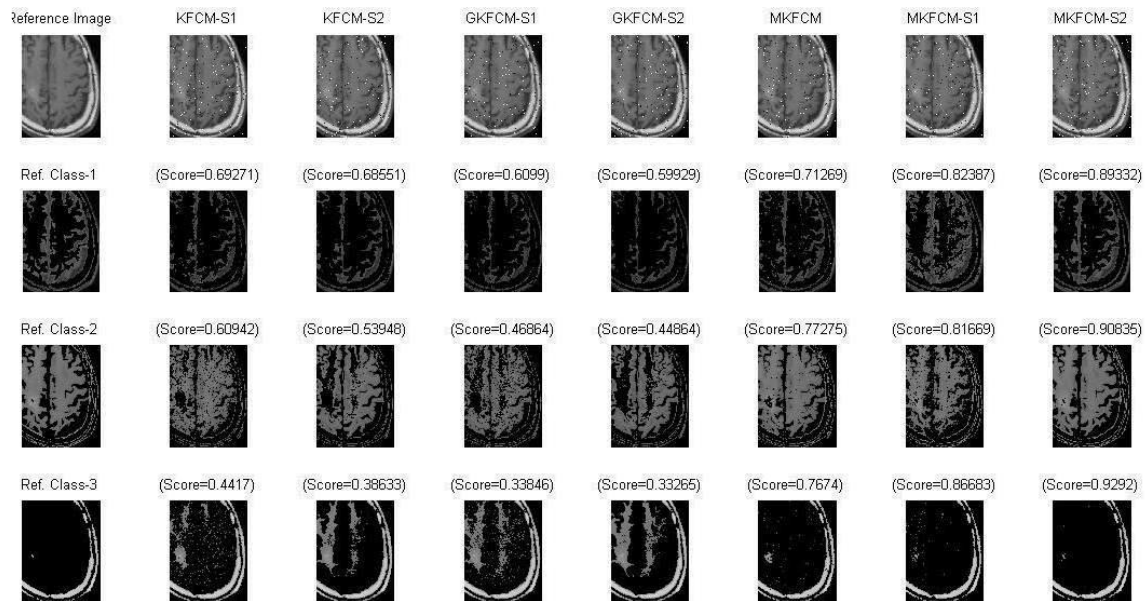


FIGURE 7: Comparison of proposed methods (MKFCM_S1 and MKFCM_S2) with other existing methods (KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM) in terms of score. The original images are corrupted by salt & pepper noise.

TABLE 3: Comparison of various techniques in terms of score on Figure (b) at different Gaussian noise
C_i: Cluster.

| | Gaussian Noise (%) | | | | | | | | | | | |
|----------|--------------------|------|------|------|------|------|------|------|------|------|------|------|
| | 5% | | | 10% | | | 15% | | | 20% | | |
| | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 |
| KFCM-S1 | 0.70 | 0.52 | 0.37 | 0.69 | 0.50 | 0.36 | 0.66 | 0.48 | 0.36 | 0.64 | 0.47 | 0.36 |
| KFCM-S2 | 0.68 | 0.52 | 0.37 | 0.65 | 0.48 | 0.35 | 0.62 | 0.46 | 0.35 | 0.60 | 0.45 | 0.34 |
| GKFCM-S1 | 0.64 | 0.43 | 0.31 | 0.68 | 0.43 | 0.31 | 0.67 | 0.41 | 0.30 | 0.66 | 0.40 | 0.30 |
| GKFCM-S2 | 0.64 | 0.41 | 0.30 | 0.66 | 0.40 | 0.30 | 0.65 | 0.38 | 0.29 | 0.62 | 0.37 | 0.29 |
| MKFCM | 0.58 | 0.45 | 0.36 | 0.51 | 0.39 | 0.33 | 0.47 | 0.36 | 0.30 | 0.45 | 0.35 | 0.29 |
| MKFCM-S1 | 0.79 | 0.67 | 0.51 | 0.73 | 0.60 | 0.46 | 0.70 | 0.55 | 0.43 | 0.67 | 0.52 | 0.42 |
| MKFCM-S2 | 0.77 | 0.68 | 0.47 | 0.70 | 0.55 | 0.42 | 0.66 | 0.50 | 0.40 | 0.64 | 0.47 | 0.38 |

TABLE 4: Comparison of various techniques in terms of score on Figure (a) at different Salt & Pepper Noise
C_i: Cluster.

| | Salt & Pepper Noise (%) | | | | | | | | | | | |
|----------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|
| | 5% | | | 10% | | | 15% | | | 20% | | |
| | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 |
| KFCM-S1 | 0.41 | 0.69 | 0.9 | 0.40 | 0.69 | 0.91 | 0.40 | 0.67 | 0.90 | 0.39 | 0.67 | 0.90 |
| KFCM-S2 | 0.44 | 0.68 | 0.88 | 0.44 | 0.68 | 0.88 | 0.44 | 0.68 | 0.88 | 0.44 | 0.68 | 0.87 |
| GKFCM-S1 | 0.34 | 0.67 | 0.89 | 0.33 | 0.66 | 0.88 | 0.31 | 0.66 | 0.88 | 0.33 | 0.65 | 0.88 |
| GKFCM-S2 | 0.35 | 0.66 | 0.88 | 0.35 | 0.65 | 0.87 | 0.35 | 0.65 | 0.86 | 0.35 | 0.64 | 0.85 |
| MKFCM | 0.47 | 0.66 | 0.85 | 0.48 | 0.67 | 0.84 | 0.47 | 0.65 | 0.82 | 0.47 | 0.64 | 0.80 |
| MKFCM-S1 | 0.95 | 0.95 | 0.98 | 0.90 | 0.90 | 0.96 | 0.86 | 0.85 | 0.94 | 0.87 | 0.83 | 0.92 |
| MKFCM-S2 | 0.81 | 0.87 | 0.96 | 0.71 | 0.83 | 0.97 | 0.63 | 0.79 | 0.96 | 0.59 | 0.79 | 0.96 |

TABLE 5: Comparison of various techniques in terms of score on Figure (b) at differnt Salt & Pepper Noise
C_i: Cluster

| | Salt & Pepper Noise (%) | | | | | | | | | | | |
|-----------------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|
| | 5% | | | 10% | | | 15% | | | 20% | | |
| | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 | CI-1 | CI-2 | CI-3 |
| KFCM-S1 | 0.70 | 0.56 | 0.40 | 0.69 | 0.59 | 0.43 | 0.69 | 0.61 | 0.45 | 0.69 | 0.60 | 0.44 |
| KFCM-S2 | 0.68 | 0.53 | 0.38 | 0.68 | 0.51 | 0.36 | 0.68 | 0.51 | 0.36 | 0.68 | 0.53 | 0.38 |
| GKFCM-S1 | 0.59 | 0.42 | 0.31 | 0.59 | 0.43 | 0.31 | 0.60 | 0.44 | 0.32 | 0.60 | 0.46 | 0.33 |
| GKFCM-S2 | 0.58 | 0.42 | 0.31 | 0.59 | 0.43 | 0.32 | 0.59 | 0.44 | 0.32 | 0.59 | 0.44 | 0.33 |
| MKFCM | 0.72 | 0.74 | 0.69 | 0.71 | 0.76 | 0.73 | 0.71 | 0.76 | 0.75 | 0.71 | 0.77 | 0.76 |
| MKFCM-S1 | 0.94 | 0.94 | 0.97 | 0.88 | 0.88 | 0.94 | 0.88 | 0.87 | 0.89 | 0.82 | 0.81 | 0.87 |
| MKFCM-S2 | 0.91 | 0.93 | 0.96 | 0.89 | 0.91 | 0.95 | 0.88 | 0.90 | 0.93 | 0.89 | 0.91 | 0.93 |

4.3. Experimental Results with Salt & Pepper Noise

Figure 5 and 7 are the segmentation results obtained with salt & pepper noise of density 20%. In Figure 5, the full size of sample images is used and in Figure 7, the region of interest (ROI) based image segmentation. The performance of the proposed methods (MKFCM_S1 and MKFCM_S2) is compared to KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM. The performance of the methods is evaluated on the basis of score. From Figures 5 and 7, MKFCM_S2 is showing better performance as compared to other existing methods (KFCM, GKFCM and MKFCM).

The performance of various methods is also evaluated in terms of number of iterations for convergence of algorithm and time requirement for the algorithm (see in Table 6-9). For time measurement, the evaluation is conducted on Core2Duo computer with speed of 2.66 GHz. From 6-9, it is clear that the proposed methods (MKFCM_S1 and MKFCM_S2) are taking less time as compared to KFCM_S1, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM. Similarly, proposed methods (MKFCM_S1 and MKFCM_S2) are taking less iteration for convergence of the algorithm which are very less as compared to other existing methods, KFCM_S2, GKFCM_S1, GKFCM_S2 and MKFCM. Figure 8 illustrates the comparison of proposed methods with other existing methods. From Figures 4–8, Table 2-9 and above observations, it is clear that the proposed algorithm shows a significant improvement in terms of score, number of iteration and time as compared to other existing methods.

TABLE 6: Comparison of various techniques in terms of number of iterations and execution time at differnt Gaussian noise
NI: Number of iterations; TM: Execution Time (Sec.)

| | Gaussian Noise | | | | | | | |
|-----------------|----------------|------|-----|------|-----|------|-----|------|
| | 5% | | 10% | | 15% | | 20% | |
| | NI | TM | NI | TM | NI | TM | NI | TM |
| KFCM-S1 | 25 | 0.59 | 30 | 0.69 | 24 | 0.56 | 29 | 0.68 |
| KFCM-S2 | 23 | 0.62 | 27 | 0.73 | 30 | 0.83 | 29 | 0.80 |
| GKFCM-S1 | 28 | 0.65 | 30 | 0.68 | 24 | 0.50 | 23 | 0.48 |
| GKFCM-S2 | 22 | 0.57 | 23 | 0.65 | 23 | 0.60 | 30 | 0.81 |
| MKFCM | 35 | 0.96 | 42 | 1.08 | 27 | 0.71 | 28 | 0.69 |
| MKFCM-S1 | 14 | 0.41 | 17 | 0.45 | 22 | 0.54 | 21 | 0.55 |
| MKFCM-S2 | 16 | 0.54 | 20 | 0.64 | 24 | 0.73 | 22 | 0.67 |

TABLE 7: Comparison of various techniques in terms of number of iterations and execution time at differnt Gaussian noise
NI: Number of iterations; TM: Execution Time (Sec.)

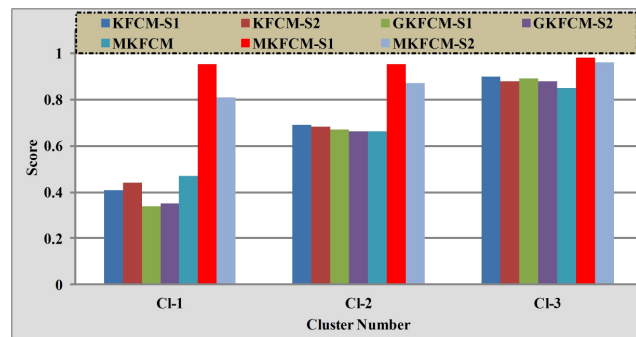
| | Gaussian Noise | | | | | | | |
|-----------------|----------------|-----|-----|-----|-----|-----|-----|-----|
| | 5% | | 10% | | 15% | | 20% | |
| | NI | TM | NI | TM | NI | TM | NI | TM |
| KFCM-S1 | 40 | 2.0 | 44 | 2.3 | 36 | 1.8 | 41 | 2.1 |
| KFCM-S2 | 46 | 2.8 | 31 | 1.8 | 32 | 1.9 | 34 | 2.1 |
| GKFCM-S1 | 35 | 1.7 | 36 | 1.8 | 39 | 2.0 | 32 | 1.6 |
| GKFCM-S2 | 35 | 2.1 | 33 | 2.0 | 32 | 1.8 | 38 | 2.3 |
| MKFCM | 27 | 1.6 | 25 | 1.5 | 28 | 1.6 | 26 | 1.5 |
| MKFCM-S1 | 61 | 3.0 | 42 | 2.2 | 32 | 1.7 | 25 | 1.3 |
| MKFCM-S2 | 50 | 2.9 | 34 | 2.1 | 31 | 1.9 | 27 | 1.7 |

TABLE 8: Comparison of various techniques in terms of number of iterations and execution time at different Salt & Pepper Noise
NI: Number of iterations; TM: Execution Time (Sec.)

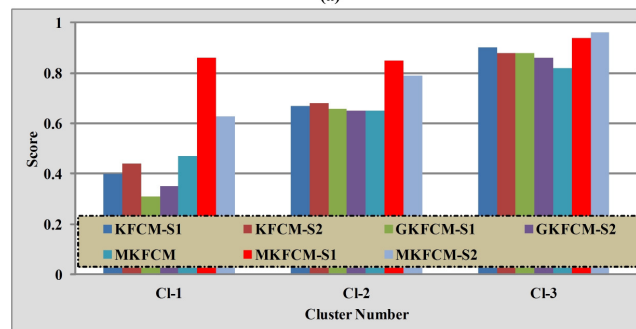
| | Salt & Pepper Noise | | | | | | | |
|-----------------|---------------------|------|-----|------|-----|------|-----|------|
| | 5% | | 10% | | 15% | | 20% | |
| | NI | TM | NI | TM | NI | TM | NI | TM |
| KFCM-S1 | 23 | 0.54 | 28 | 0.62 | 20 | 0.44 | 23 | 0.51 |
| KFCM-S2 | 24 | 0.60 | 23 | 0.58 | 22 | 0.55 | 21 | 0.52 |
| GKFCM-S1 | 18 | 0.36 | 26 | 0.55 | 24 | 0.51 | 26 | 0.56 |
| GKFCM-S2 | 22 | 0.53 | 23 | 0.54 | 25 | 0.61 | 20 | 0.48 |
| MKFCM | 24 | 0.60 | 20 | 0.49 | 21 | 0.51 | 21 | 0.52 |
| MKFCM-S1 | 20 | 0.51 | 19 | 0.49 | 18 | 0.44 | 18 | 0.50 |
| MKFCM-S2 | 27 | 0.75 | 25 | 0.69 | 26 | 0.73 | 19 | 0.58 |

TABLE 9: Comparison of various techniques in terms of number of iterations and execution time at different Salt & Pepper Noise
NI: Number of iterations; TM: Execution Time (Sec.)

| | Salt & Pepper Noise | | | | | | | |
|-----------------|---------------------|-----|-----|-----|-----|-----|-----|-----|
| | 5% | | 10% | | 15% | | 20% | |
| | NI | TM | NI | TM | NI | TM | NI | TM |
| KFCM-S1 | 61 | 3.2 | 58 | 3.0 | 52 | 2.8 | 51 | 2.7 |
| KFCM-S2 | 78 | 5.0 | 46 | 2.8 | 76 | 4.8 | 53 | 3.3 |
| GKFCM-S1 | 42 | 2.1 | 38 | 1.9 | 62 | 3.3 | 64 | 3.3 |
| GKFCM-S2 | 69 | 4.3 | 43 | 2.6 | 43 | 2.7 | 88 | 5.5 |
| MKFCM | 139 | 9.4 | 80 | 5.2 | 83 | 5.4 | 45 | 2.8 |
| MKFCM-S1 | 47 | 2.4 | 45 | 2.3 | 31 | 1.7 | 43 | 2.2 |
| MKFCM-S2 | 33 | 2.0 | 33 | 2.0 | 31 | 1.9 | 23 | 1.5 |



(a)



(b)

FIGURE 8: Comparison of proposed methods with other existing methods with Salt & Pepper noise: (a) 5% and (b) 15%.

5. CONCLUSIONS

In this paper, new image segmentation algorithms (MKFCM_S1 and MKFCM_S2) which are increasing the performance and decreasing the computational complexity. The algorithm utilizes the spatial neighborhood membership values in the standard kernels are used in the kernel FCM (KFCM) algorithm and modifies the membership weighting of each cluster. The proposed algorithm is applied on brain MRI which degraded by Gaussian noise and Salt-Pepper noise demonstrates that the proposed algorithm performs more robust to noise than other existing image segmentation algorithms from FCM family.

6. REFERENCES

- [1]. X. Munoz, J. Freixenet, X. Cufi, and J. Marti, "Strategies for image segmentation combining region and boundary information," *Pattern Recognition Letters*, vol. 24, no. 1, pp. 375–392, 2003.
- [2]. D. Pham, C. Xu, and J. Prince, "A survey of current methods in medical image segmentation," In *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, 2000.
- [3]. Mohammad Ali Balafar, Abd.Rahman Ramli, M.Iqbal Saripan, Syamsiah Mashohor, "Medical Image Segmentation Using Fuzzy C-Mean (Fcm), Bayesian Method And User Interaction," *Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition*, pp. 68-73, Aug. 2008.
- [4]. László Szilágyi, Sándor M. Szilágyi, Balázs Benyó and Zoltán Benyó, "Application of Hybrid c-Means Clustering Models in Inhomogeneity Compensation and MR Brain Image Segmentation," *5th International Symposium on Applied Computational Intelligence and Informatics*, pp.105-110, May. 2009.
- [5]. MacQueen,J.B. "Some Methods for classification and Analysis of Multivariate Observations,"*Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, pp. 281–297, 1967.
- [6]. Arthur D,Vassilvitskii S,"How Slow is the k-means Method?," *Proceedings of the 2006 Symposium on Computational Geometry* , June. 2006.
- [7]. L. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, pp. 338–353, 1965.
- [8]. J. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithm and applications in image segmentation," *Graphical Models and Image Processing*, vol. 58, no. 3, pp. 246–261, 1996.
- [9]. Y. Tolias and S. Panas, "Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 28, no. 3, pp. 359–369, Mar.1998.
- [10]. J. Noordam, W. van den Broek, and L. Buydens, "Geometrically guided fuzzy C-means clustering for multivariate image segmentation," in *Proceedings of the International Conference on Pattern Recognition*, 2000, vol. 1, pp.462–465.
- [11]. M. Yang, Y. J. Hu, K. Lin, and C. C. Lin, "Segmentation techniques for tissue differentiation in MRI of ophthalmology using fuzzy clustering algorithms," *Magnetic Resonance Imaging*, vol. 20, no. 2, pp. 173–179, 2002.
- [12]. G. Karmakar and L. Dooley, "A generic fuzzy rule based image segmentation algorithm," *Pattern Recognition Letters.*, vol. 23, no. 10, pp.1215–1227, 2002.
- [13]. M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 21, no. 3, pp. 193–199, 2002.
- [14]. J. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," Kluwer Academic Publishers, New York: Plenum, 1981.

- [15]. D. Pham, "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognition Letters*, vol. 20, pp. 57–68, 1999.
- [16]. S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, pp. 1907–1916, 2004.
- [17]. L. Szilagyi, Z. Benyo, S. Szilagyii, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proceedings of the 25th Annual International Conference of the IEEE EMBS*, pp. 17–21, 2003.
- [18]. M. Krinidis and I. Pitas, "Color texture segmentation based-on the modal energy of deformable surfaces," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1613–1622, Jul. 2009.
- [19]. D. Pham, "Fuzzy clustering with spatial constraints," in *Proceedings of International Conference on Image Processing*, New York, 2002, vol. II, pp. 65–68.
- [20]. W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825–838, Mar. 2007.
- [21]. Miin-Shen Yang, Hsu-Shen Tsai, "A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction," *Pattern Recognition Letters*, vol. 29, pp. 1713–1725, May 2008.
- [22]. G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
- [23]. Long Chen, C. L. Philip Chen, and Mingzhu Lu, "A Multiple-Kernel Fuzzy C-Means Algorithm for Image Segmentation," *IEEE Trans. Systems, Man, And Cybernetics—Part B: Cybernetics*, vol. 41, No. 5, pp. 1263 – 1274, February 9, 2011.
- [24]. D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (OASIS): Crosssectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.*, 19 (9) 1498–1507, 2007.
- [25]. Masulli, F., Schenone, A., 1999. A fuzzy clustering based segmentation system as support to diagnosis in medical imaging. *Artif. Intell. Med.* 16, 129–147.
- [26]. Zhang, D.Q., Chen, S.C., 2004. A novel kernelized fuzzy c-means algorithm with application in medical image segmentation. *Artif. Intell. Med.* 32, 37–50.

Local Phase Oriented Structure Tensor To Segment Texture Images With Intensity Inhomogeneity

Hiren Mewada

*Department of Electronics and Communication
Charotar University of Science and Technology
At Post Changa, Anand-388421, India*

mewadahiren@gmail.com

Suprava Patnaik

*Department of Electronics and Communication
S.V National Institute of Technology
Surat – 395009, India*

suprava_patnaik@yahoo.com

Abstract

This paper proposed the active contour based texture image segmentation scheme using the linear structure tensor and tensor oriented steerable Quadrature filter. Linear Structure tensor (LST) is a popular method for the unsupervised texture image segmentation where LST contains only horizontal and vertical orientation information but lack in other orientation information and also in the image intensity information on which active contour is dependent. Therefore in this paper, LST is modified by adding intensity information from tensor oriented structure tensor to enhance the orientation information. In the proposed model, these phases oriented features are utilized as an external force in the region based active contour model (ACM) to segment the texture images having intensity inhomogeneity and noisy images. To validate the results of the proposed model, quantitative analysis is also shown in terms of accuracy using a Berkeley image database.

Keywords: Linear Structure Tensor, Quadrature filter, Active contour, Image Segmentation

1. INTRODUCTION

Image segmentation is primary steps for the point of interpretation. This segmentation process relies on the extraction of required features from the image that discriminates the region of interest from the given image. The proposed model is toward the segmentation of texture images having intensity inhomogeneity. Many algorithms have been suggested in various literatures for texture image segmentation some of them are [10, 18, 11, 9]. These segmentation models include Gabor filter based approach, Markov random field based approach and multi-resolution approach. One of the most familiar technique is a Gabor filter to extract the features of the texture image. Another approach is the use of active contour model [16], which is utilized in this paper. The proposed model in this paper utilized the region based geometric active contour model for the texture image segmentation.

It is well known that region based CV model [16] failed to segment the images having either intensity inhomogeneity and texture patterns. Therefore to handle the intensity inhomogeneity, Lee et al [8] presented local kernel based active contour model. But limitation is still not able to segment the texture images and again it is highly sensitive to the size of the kernel. To segment texture images, Wang et al [17] modified the CV model [16] by guiding the active contour with histogram based information. Similarly Bhattacharya distance [12, 20] also proposed as the texture image segmentation model using histogram calculation. Then Kangyu Ni et. al. [19] proposed the similar histogram based level set function where Wasserstein distance was utilized as a distance vector between the histogram in comparison to the Bhattacharya distance. The most common disadvantage of all these models is the need of histogram calculation which incorporates

computational complexity. To extract the orientation information, earlier gradient based local structure tensor (LST) method was suggested in [1]. Gradient based LST is a simple method and gives good local orientation localization in two directions. Therefore LST is largely utilized to get the orientation of texture pattern. Another method to extract the orientation features of images is an estimation method based on Quadrature filter set after representation by Knutsson and Andersson [7]. This method has more advantage in terms of orientation localization because Quadrature filter based method can estimate more orientation information as compared to gradient based LST. But filter based approach is utilized to segment the medical images having intensity inhomogeneity [4] and not applied to texture images. Therefore the proposed model combines the advantages/features of a filter based structure into the LST and this information is utilized to guide the active contour model (ACM) for the image segmentation. This makes LST model to segment the texture images having intensity inhomogeneity.

The outline of the paper is as follows: Section 2 introduces the tensor oriented Quadrature filter based structure and the gradient based LST for texture image segmentation. Section 3 demonstrates the proposed model based on section 2 and corresponding results are discussed along with quantitative analysis. Finally in section 4 results are concluded.

2. BACKGROUND

2.1 Tensor oriented Quadrature filter

It is well known that image intensity is highly sensitive to the lighting condition. This variation in contrast introduced by lighting conditions cause difficulties in segmentation. The information carried by the image local phases is invariant to intensity variation. At the same time the phase information allows us to detect the lines and edges simultaneously; which make us to utilize this local phase information as an image feature for the image segmentation. One of the techniques is the use of quadrature filters to extract the phase information. To obtain the Quadrature filter, one radial function R and one directional function D are defined in Fourier domain. Hence,

$$F(u\hat{x}, v\hat{y}) = R(\rho)D(u\hat{x}, v\hat{y}) \quad (1)$$

Where $\rho = \sqrt{u^2 + v^2}$ and u, v are dummy variables in frequency domain and \hat{x}, \hat{y} represents the orientation vector. Component in R is symmetrical in nature, and hence in order to obtain the Quadrature property of F, D has to be Quadrature. i.e.

$$D(u\hat{x}, v\hat{y}) = 0 \text{ if } u\hat{x} \leq 0, v\hat{y} \leq 0 \quad (2)$$

Both R and D are assumed to be real valued. Radial function (R) utilized in the paper is given by

$$R(\rho) = \exp\left(\frac{-4}{B^2 \ln 2}\right) \ln^2(\rho / \rho_i) \quad (3)$$

Where B is the bandwidth and ρ_i is the center frequency.

If this filter is applied to the signal, the filter output is calculated by the product of the Fourier transform of the filter and the Fourier transform of that signal around the point. Let the filter output is denoted by q' which gives,

$$q'(u, v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(u, v) F(u\hat{x}, v\hat{y}) du dv \quad (4)$$

Where $S(u, v)$ is the Fourier transform of an arbitrary 2-D $s(x, y)$ in spatial domain. In above equation 4, q' represents the filtered signal oriented by directional component D. In this paper, three Quadrature filters are utilized.

Utilization of these Quadrature filters eliminates the negative frequency components. Hans Knutsson and Mats Andersson [7] introduce the orientation tensor T to represent the signal orientation. To generate the orientation tensor, following preference axis is utilized [6].

$$\hat{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad (5)$$

Where α depends on chosen coordinate systems. Three orientations 0, $2\pi/3$ and $-2\pi/3$ are utilized to calculate the tensor in this paper.

Based on these orientation corresponding orientation tensor is given by

$$T = \begin{pmatrix} T_{\alpha,11} & T_{\alpha,12} \\ T_{\alpha,21} & T_{\alpha,22} \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{pmatrix} = \begin{pmatrix} \cos^2 \alpha & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & \sin^2 \alpha \end{pmatrix} \quad (6)$$

To obtain the phase map of filtered image, linear combination of filter output and tensor z (from equation 6) is utilized. Therefore resulting phase map in respective direction can be written as

$$Z_\alpha = T_{\alpha,11}q_1 + T_{\alpha,12}q_2 + T_{\alpha,13}q_3 \quad (7)$$

The texture images contain repeated pattern within the region. And the task of tensor based Quadrature filter is to extract the local phase information. Therefore, it is not suitable to apply these filter pairs to the texture images. Therefore proposed model in this paper combines this orientation information to LST to obtain the boundary of objects.

2.2 Local structure tensor

Covariance matrix based LST is strongly utilized as a texture descriptor. The advantage of LST is that its dimensionality is small and simplicity in the calculation. Therefore it is widely utilized in corner detection, texture analysis and texture segmentation [2, 5]. Consider a grayscale image . The matrix field of the structure tensor is [2, 3] represented by

$$J_0 = \nabla h \nabla h^T = \begin{bmatrix} h_x^2 & h_x h_y \\ h_x h_y & h_y^2 \end{bmatrix} \quad (8)$$

Where $\nabla h = \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{bmatrix}^T$, $h_x = \frac{\partial h}{\partial x}$, $h_y = \frac{\partial h}{\partial y}$ and T = Matrix transpose.

In order to avoid cancellation of the opposite signed gradient when direct integration is performed, the gradient is considered as the form of its outer product. In order to make the matrix field more robust to noise and other artifacts, smoothing is usually performed by convolution of the of the matrix components with a Gaussian kernel K_ρ with standard deviation ρ ,

$$J_\rho = K_\rho * (\nabla h \nabla h^T) \quad (9)$$

Where * indicate convolution operation.

Above LST model contains the orientation and magnitude of texture features and therefore it can be utilized to segment the texture images. However as it lacks the intensity information and thus failed to segment the texture images having intensity inhomogeneity. Again problem with this model is that it can not be combined with active contour model (ACM) due to the dependency of its energy minimization criterion on image intensity rather than texture pattern. Therefore this paper suggests the solution to this problem with a combination of filter based tensor values to the LST. This proposed model is explained in following section.

2.3 Region based Active Contour model

The CV Model is based on piecewise constant approximation of the Mumford Shah energy function [16]. For an image $I(x, y)$ in the image domain Ω , they proposed the energy minimization formula as

$$E(c_1, c_2, C) = \mu \text{Length}(C) + v \text{Area}(C) + \lambda_1 \int_{\text{inside}(C)} |I(x, y) - c_1|^2 dx dy + \lambda_2 \int_{\text{outside}(C)} |I(x, y) - c_2|^2 dx dy \quad (10)$$

where c_1 and c_2 are the average intensities inside and outside of the contour, respectively. λ_1 and λ_2 are positive constants. Generally we take λ_1 and λ_2 equal to one. As the last two terms in equation (10) is the global fitting energy, Chan-Vese model is global region based model. Here the contour is represented by the zero level set function ϕ [14]. Keeping c_1 and c_2 constant during the evolution of contour, the energy minimization equation was obtained as

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) [\mu k - v - \lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2] \quad (11)$$

where k is the Euclidean curvature of the contour C , δ is the smoothed Dirac delta function given by

$$\delta(x) = \frac{\epsilon}{\pi(\epsilon^2 + x^2)} \quad (12)$$

The limitation of the CV model is not able to segment the images having either intensity inhomogeneity or texture patterns. Therefore this paper proposed the solution using the advantage of LST and the steerable Quadrature filter which is discussed in following section 3.

3. PROPOSED MODEL FOR TEXTURE IMAGE SEGMENTATION

As the local phase map obtained by the tensor oriented Quadrature filter is invariant to intensity inhomogeneity, in the proposed model, local phase map is obtained from texture image using tensor oriented Quadrature filter. Here, three Quadrature filters are utilized having a bandwidth of 3 dB and center frequency of $\pi/5$. Each Quadrature filter has a different orientation and consists of real and imaginary part where the real part is able to detect the line in image and imaginary part detect the edge in image. To enhance local phase, it is required to have a rotation invariant phase map. Therefore we integrated all phase maps. Now during the integration of features from all directions, there may be chance of cancellation of features due to negation of phase at opposite direction. This may cause a discontinuity or flickering in phase map generated by the filter. Therefore we flip the phase along the real axis for a filter with direction opposing the other direction. Thus we obtain the phase maps in three filter orientation. These flipped phase map from all directions are added to obtain the final phase map as shown in equation 13.

$$q = \sum_{k=1}^3 \|q_k\| \quad (13)$$

This flipped phase map is driven by the tensor to form the dual tensor structure using equation 6 as follow.

$$Z_k = T_{k,11}q + T_{k,12}q + T_{k,22}q \quad k = 1, 2, 3 \quad (14)$$

From this Eigenvalues are calculated using following two equations:

$$l_1 = \frac{Z_1 + Z_2}{2} + \sqrt{\left(\frac{Z_1 - Z_2}{2}\right)^2 + Z_3^2} \quad l_2 = \frac{Z_1 + Z_2}{2} - \sqrt{\left(\frac{Z_1 - Z_2}{2}\right)^2 + Z_3^2} \quad (15)$$

These Eigenvalues can be considered as equivalent to the horizontal and vertical edge information with enhancement. To extract the perceptual visualized based shape information on the texture object from the given image, here this tensor oriented Quadrature filter output is combined with LST. For that, in LST's equation 8, the horizontal and vertical gradients are replaced by the two Eigenvalues calculated from the equation 15 and then after the local tensor is calculated using equation 9. Thus corresponding LST is defined as

$$L_{QF} = \begin{bmatrix} l_1^2 & l_1 l_2 \\ l_1 l_2 & l_2^2 \end{bmatrix} \quad (16)$$

Localized structure calculated using above equation 16 smooth out the foreground and background texture patterns while keeping enhanced edge or boundary of the object. Thus the present approach enhances the weak edges and solves the intensity inhomogeneity problem with smoothing of texture pattern in the background and foreground pattern as shown in figure 1 (d).

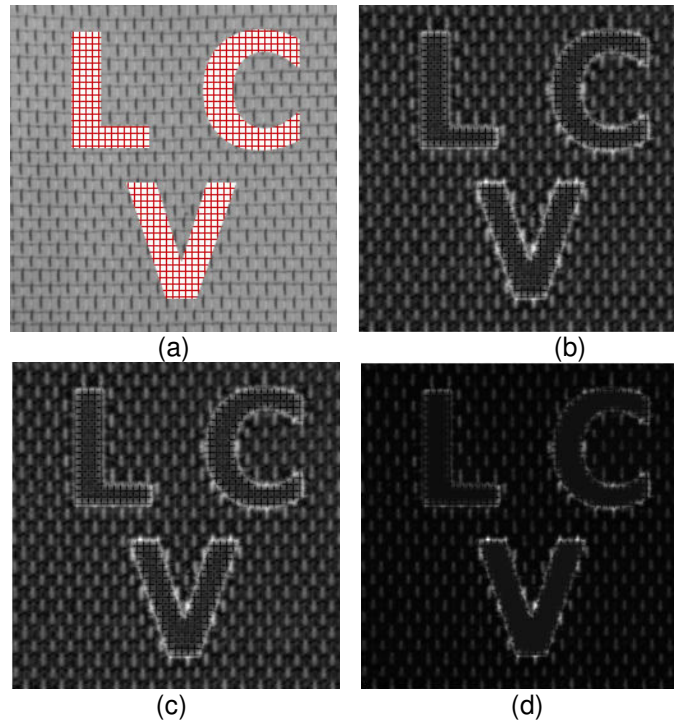


FIGURE 1: (a) Original Image, (b) and (c) Eigen image after tensor oriented Quadrature filter (d) LST based output using equation 16.

Figure 1 shows that use of the phase map helps in identifying the edge maps even in the inhomogeneous region, but it is not applicable to a texture object. Therefore to get the boundary of the texture object, eigenvalues of that phase map are calculated. These eigenvalues remove the texture pattern while keeping the object boundaries. Thus the proposed model is able to segment the texture images having intensity inhomogeneity.

To obtain the segmentation of a given image, this LST based output is utilized as the original image in active region based model (see equation 17). Thus the final energy minimization equation is given by

$$\frac{\partial \phi}{\partial t} = \delta_{\epsilon}(\phi) \left[\mu k - v - \lambda_1 (L_{QF} - c_1)^2 + \lambda_2 (L_{QF} - c_2)^2 \right] \quad (17)$$

4. EXPERIMENT RESULTS

To validate the model, the proposed model is applied to a number of texture images involving different types of complexity. From the point of complexity, here specific images are considered as follows: Like in the first row of figure 7, the head in left middle part and the legs of leopard has similar intensity range as the background. This makes difficult to segment the images, and an individual approach of either LST (due to intensity inhomogeneity) or tensor oriented Quadrature filter (due to texture region) or active contour based model (due to texture region) failed to segment the image (see figure 2), while the proposed model is able to segment it successfully as inhomogeneity is handled by the phase map obtained from the Quadrature filter and texture pattern is smoothen out using the eigen based structure tensor.

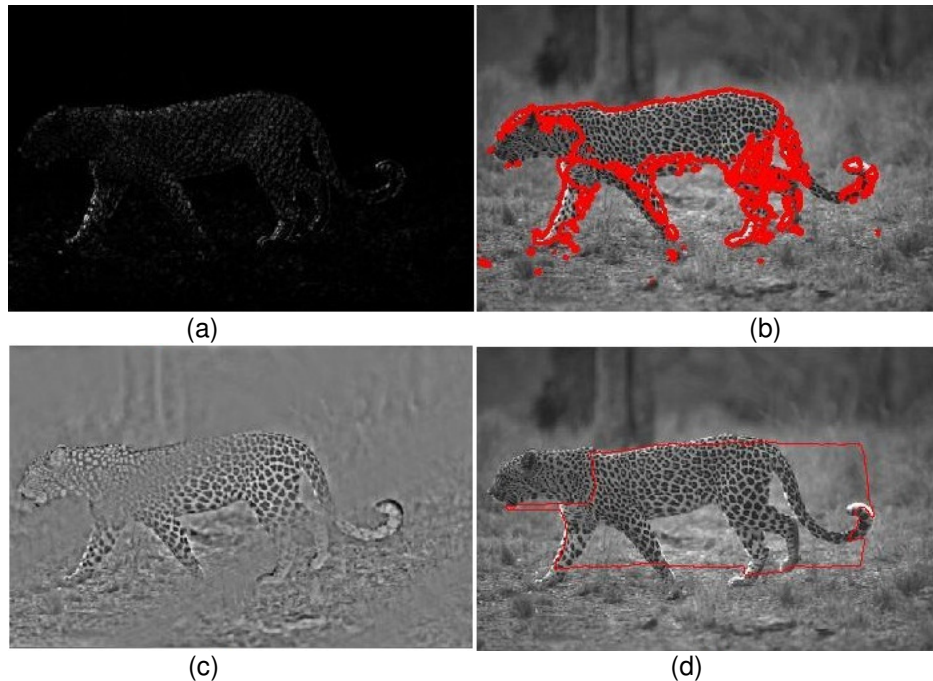


FIGURE 2: (a) LST of leopard image (b) Corresponding CV based segmentation (c) Phase map from Quadrature filter (d) Corresponding CV based segmentation.

Similarly in the tiger image (figure 3) due to inhomogeneity, the boundary is not easily distinguishable, while using tensor oriented Quadrature filter it can be enhanced but the presence of texture region cause failure in segmentation using the CV model. At the same time, LST removes the texture region but it is not able to enhance the boundary due to intensity inhomogeneity (figure 3 b). Using combine approach this part got enhanced (figure 3 c) and make proposed model to segment it accurately (see figure 7).

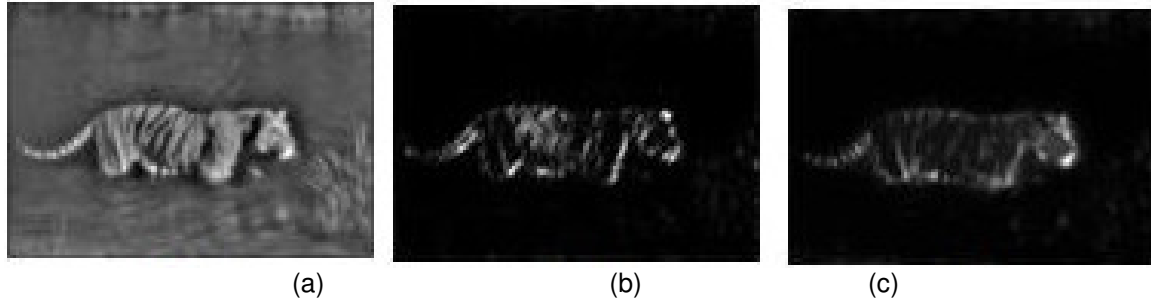


FIGURE 3: (a) Tensor oriented Quadrature filter output (b) LST oriented output (c) using combine approach as suggested in the proposed model.

In monkey image (figure 4), lower left (white shaded region) part is also segmented by region based model. But combined approach segments it perfectly. In last synthetic images are shown to highlight the capability of the proposed model to segment images with multiple texture patterns and/or images with multiple objects with similar texture patterns.

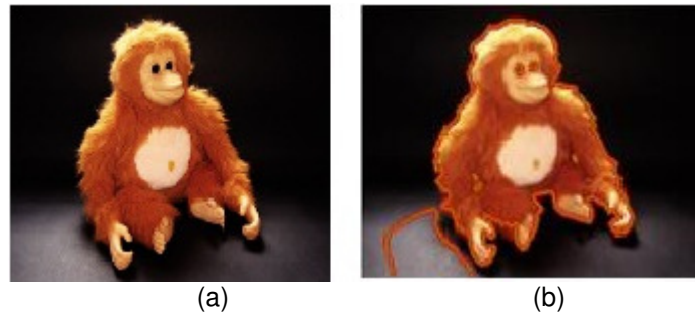


FIGURE 4: CV Model based Segmentation (a) Original Image (b) Segmented Image.

There are so many algorithms are proposed to segment the texture images using region based active contour model. Some of them are [10, 18, 11, 9, 15, and 12]. All these models utilized either Gabor based approach or histogram method to extract the texture region and accordingly the CV model is applied to segment the images. However the Gabor based approach [15] requires too many features channels and induces lots of redundancies. Another utilized method is based on histogram of an image is Bhattacharya distance based ACM [12]. In [12], probability density inside the contour and outside the contour are estimated. The distance between this two probability densities is calculated using the Bhattacharya coefficients given by

$$B(\phi(x)) = \int \sqrt{P_{in}(z | \phi(x)) P_{out}(z) | \phi(x)} dz \quad (18)$$

In this model, the authors utilized the color histogram which is computationally hard. Therefore an independent component analysis (ICA) was utilized to reduce the feature space. Finally isotropic diffusion function was utilized to minimize the energy function in the active contour model. Some of successful results obtained using this model is shown in figure 5 below.

However the Bhattacharya model has some limitations. First one is not able to segment complex geometries as it utilizes the anisotropic diffusion function. In figure 5 (b), it can be seen that the tail part of tiger image is not segmented properly due to its complex geometries. A second limitation is not able to segment images with intensity inhomogeneity. Thus the Bhattacharya based ACM fails to segment the image shown in the first row of figure 7.

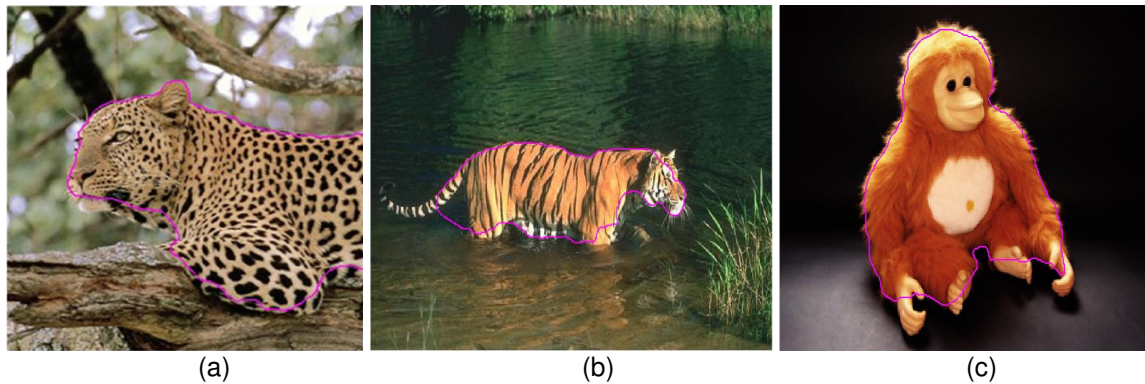


FIGURE 5: Segmentation Results Obtained using Bhattacharya Gradient Flow [12].

Third limitation is not able to segment multiple objects (figure 6) as splitting of contour is not supported in this model. Figure 6 shows the images having multiple object. In figure 6 “LCV” is the image containing a similar texture pattern inside the object and in the background with different scale while in the image “OK” both the foreground and the background have different texture pattern. Thus figure 6 shows that in comparison to the Bhattacharya model [12], the proposed model is able to segment multiple objects too.

Figure 7 shows the segmentation results obtained while applying the proposed model on the texture images having above said complexities. The ground truth images are also shown to compare the obtained segmentation results visually.

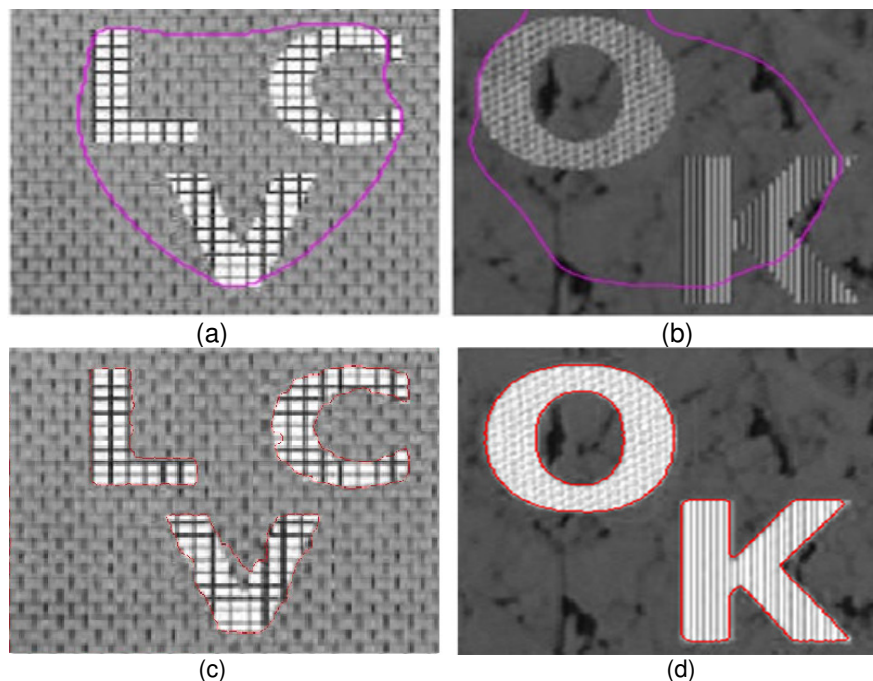


FIGURE 6: (a) and (b) Failer of the Bhattacharya based Gradient flow model [12] to segment multiple object (c) and (d) successful segmentation using the proposed model

4.1 Quantitative Analysis

In this subsection we also prove the accuracy and efficiency of the proposed model analytically. We generated ground truth using ImageJ tool for the images not available in the Berkely image database. Then we computed the Dice similarity coefficient (DSC) which represents the spatial

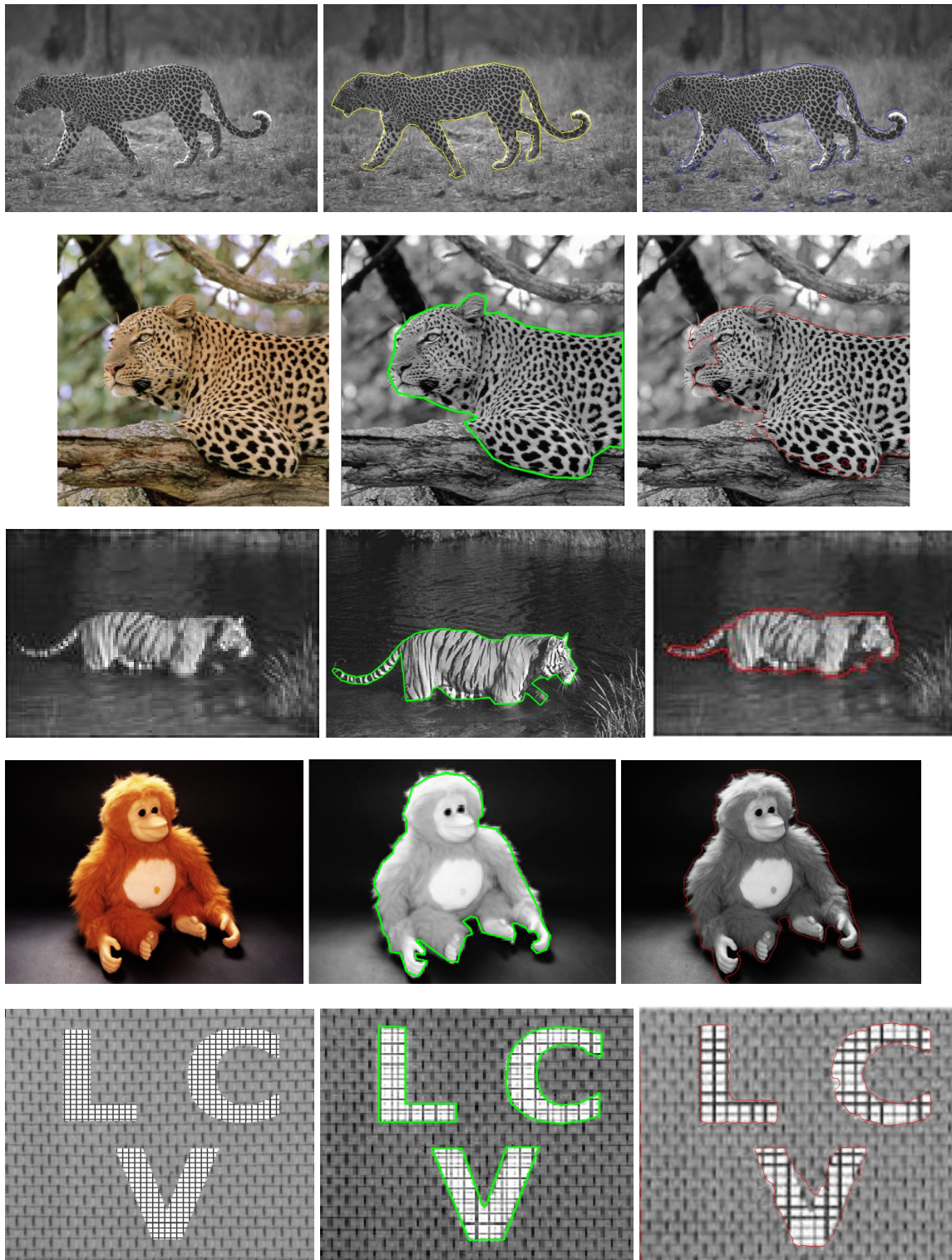


FIGURE 7: Left: Original Image, Middle: Ground truth image and Right: Segmented output using the proposed model.

overlap between the segmented image and ground truth image. Figure 8 shows the two sets: Resultant set (segmented image) and truth set (ground truth image). Where true positive represents the number of pixels detected correctly. False positive represents the number of pixels

in segmented image which are actually not a part of the ground truth image, and false negative represents the number of pixels which are actually part of the ground truth image but not detected in segmentation result using the proposed model.

Using these, DSC can be defined as

$$DSC = \frac{2TP}{(FP + TP) + (TP + FN)} \quad (19)$$

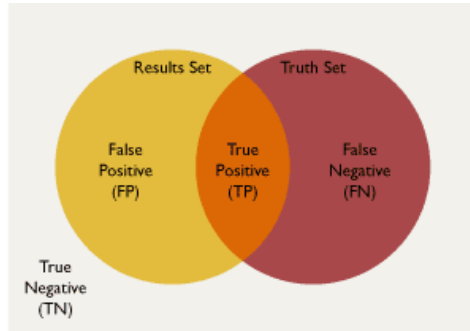


FIGURE 8: Dice Similarity Coefficients Calculation Parameter.

Thus for accurate segmentation (more similarity between the ground truth image and segmented image using the proposed model), DSC should be towards to one. The DSC for figure 7 are presented in table 1. In table 1, NA represents that the model is not able to segment image and therefore corresponding DSC value is not available. For the image "tiger_bhatta" (second row of figure 7), both the models have approximately equal accuracy, but visually Bhattacharya based model have better segmentation in comparison to the proposed model.

| Image | DSC using the proposed model | DSC using Bhattacharya model [12] |
|----------------|------------------------------|-----------------------------------|
| Leopard | 0.95 | NA |
| Tiger | 0.96 | 0.92 |
| Leopard_bhatta | 0.96 | 0.97 |
| Monkey | 0.99 | 0.90 |
| LCV | 0.96 | NA |
| OK | 0.99 | NA |

TABLE 1 TPR (%) Values of the Proposed Model and [12] On Above Shown Test Images.

5. CONCLUSION AND FUTURE WORK

The major limitation of region based active contour model is not able to segment images having texture region and intensity inhomogeneity. To obtain the texture features LST is widely utilized method having the advantage of discrimination of texture pattern. The limitation of LST is containing horizontal and vertical direction information only. At the same time, to tackle the problem of intensity inhomogeneity, tensor oriented Quadrature filter is the one of the best solution as it utilized the intensity invariant local phase features. Therefore they are widely utilized for medical images and limitation is not able to segment the texture image.

Therefore this paper proposed an approach to integrate the advantages of both LST and tensor oriented Quadrature filter which overcome the limitation of each individual method. This integrated approach is utilized to overcome the limitation of the region based CV model. In the proposed model, Eigen values are calculated from the complex output of the tensor oriented quadrature filter. The Eigenvalues represent the intensity invariant local phase map features.

Then a horizontal and vertical gradients in LST are replaced by these Eigenvalues, which contain all directional information. Thus the integration of Eigenvalues in LST overcomes the limitation of two directions information of texture pattern only. Finally this modified local structure tensor based image is utilized in the region based active contour model for the segmentation purpose. The promising results shown in paper conclude that the proposed model removes the limitation of CV model and make it able to segment the images having intensity inhomogeneity and the texture images too. The results are tested for a number of images and comparison is established with the Bhattacharya based Gradient decent flow model. The quantitative analysis also concludes that the proposed model achieves approximately 100% accuracy.

Still model is sensitive to the initialization of the contour of the image domain and different initialization may lead to different results due to non convex nature of active contour. Thus the model can be improved in which is insensitive to the initialization. Again in the proposed model, instead of Quadrature filter, the Cauchy filter [3] can be utilized to improve the accuracy of segmentation.

6. REFERENCES

- [1] J. Big Aijn, G. H. Granlund, and J. Wiklund. "Multidimensional orientation estimation with applications to texture analysis and optical flow" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.13(8), pp.775 -790, 1991.
- [2] T. Brox. "From pixels to regions: Partial differential equations in image analysis", PhD Thesis, Mathematical Image Analysis Group, Department of Mathematics and Computer Science Saarland University, Germany, 2005.
- [3] T. Brox, J. Weickert, B. Burgeth, and P. MrAazek. "Nonlinear structure tensors. Image and Vision Computing", Vol.24(1), pp.41-55, 2006.
- [4] R. Estepar. "Local structure tensor for multidimensional signal processing: Application to medical image analysis", Ph D Thesis, universitaires de Louvain, 2007.
- [5] C. Feddern, J. Weickert, and B. Burgeth. "Level-set methods for tensor valued images", Proc. Second IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision, pp. 65-72, 2003.
- [6] G. H. Granlund. "In search of a general picture processing operator". Computer Graphics and Image Processing, Vol.8(2), pp.155-173, 1978.
- [7] H. Knutsson and M. Andersson. "Loglets: Generalized Quadrature and phase for local spatio-temporal structure estimation". 13th Scandinavian Conference, SCIA-2003 Halmstad, Sweden, July 2003, pp.741 -748.
- [8] C. Li, C. Kao, J. Gore, and Z. Ding. "Implicit active contour driven by local binary fitting energy". IEEE Conf on Computer Vision and Pattern Recognition, 2007, pp. 1-7.
- [9] S. Li, J. T. Kwok, H. Zhu, and Y. Wang. Texture classification using the support vector machines. Pattern Recognition, Vol.36(12), pp.2883 - 2893, 2003.
- [10] A. Lorette, X. Descombes, and J. Zerubia. "Texture analysis through a markovian modelling and fuzzy classification: application to urban area extraction from satellite images". International Journal of Computer Vision, Vol. 36(5), pp. 221-236, 2002.
- [11] H. Lu, Y. Liu, Z. Sun, and Y. Chen. "An active contours method based on intensity and reduced Gabor features for texture segmentation". Intelligent Control and Information Processing (ICICIP), pp. 1369 -137, Nov 2009.

- [12] O. Michailovich, Y. Rathi and Tannenbaum. "Image segmentation using active contours driven by the Bhattacharyya gradient flow". IEEE Transactions On Image Processing, Vol.16(11), pp.2787 -2801, 2007.
- [13] J. Ning, L. Zhang, D. Zhang, and C. Wu. "Interactive image segmentation by maximal similarity based region merging". Journal of Pattern Recognition, Vol.43(11), pp. 445-456, 2010.
- [14] S. Osher and J. Sethian. "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations". Journal of Computational Physics, Vol. 79, pp. 12-49, 1988.
- [15] B. Sandberg, T. Chan, and L. Vese. "A level-set and Gabor-based active contour algorithm for segmenting textured images". Technical Report 39, Mathematical Department, UCLA, Los Angeles, 2002.
- [16] T.Chan and L. Vese. "Active contour without edges". IEEE Transactions on Image Processing, Vol.10(2), pp. 266 -277, 2001.
- [17] Y. Wang, Y. Xiong, L. Lv, H. Zhang, Z. Cao, and D. Zhang. "Vector-valued chan-vese model driven by local histogram for texture segmentation". 17th IEEE International Conference on Image Processing (ICIP), , Sept 2010, pp.645 -648.
- [18] D. Yang, T. Deng, C. Yang, and J. Bian. "Interactive graph cut method based on improved Gabor features for image segmentation". Intelligent Control and Information Processing (ICICIP), Vol.1(2), pp.267 - 270, July 2011.
- [19] Kangyu Ni, Xavier Bresson, Tony Chan and Selim Esedoglu, "Local histogram based segmentation using the Wasserstein distance", Internation Journal of Computer Vision, Vol. 84, pp. 97-111, April 2009.
- [20] C.C. Reyes-Aldasoroa, A. Bhalerao, "The Bhattacharyya space for feature selection and its application to texture segmentation" Internation Journal of Pattern Recognition Vol .39 pp. 812 – 826, 2006

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Image Processing (IJIP)* aims to be an effective forum for interchange of high quality theoretical and applied research in the Image Processing domain from basic research to application development. It emphasizes on efficient and effective image technologies, and provides a central forum for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing.

We welcome scientists, researchers, engineers and vendors from different disciplines to exchange ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJIP.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with volume 7, 2013, IJIP is appearing with more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS

The realm of International Journal of Image Processing (IJIP) extends, but not limited, to the following:

- Architecture of imaging and vision systems
- Character and handwritten text recognition
- Chemistry of photosensitive materials
- Coding and transmission
- Color imaging
- Data fusion from multiple sensor inputs
- Document image understanding
- Holography
- Image capturing, databases
- Image processing applications
- Image representation, sensing
- Implementation and architectures
- Materials for electro-photography
- New visual services over ATM/packet network
- Object modeling and knowledge acquisition
- Photographic emulsions
- Prepress and printing technologies
- Remote image sensing
- Autonomous vehicles
- Chemical and spectral sensitization
- Coating technologies
- Cognitive aspects of image understanding
- Communication of visual data
- Display and printing
- Generation and display
- Image analysis and interpretation
- Image generation, manipulation, permanence
- Image processing: coding analysis and recognition
- Imaging systems and image scanning
- Latent image
- Network architecture for real-time video transport
- Non-impact printing technologies
- Photoconductors
- Photopolymers
- Protocols for packet video
- Retrieval and multimedia

- Storage and transmission

- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 7 - Issue: 5

i. Paper Submission: November 30, 2013

ii. Author Notification: December 25, 2013

iii. Issue Publication: December 2013

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607

006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2013
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org