

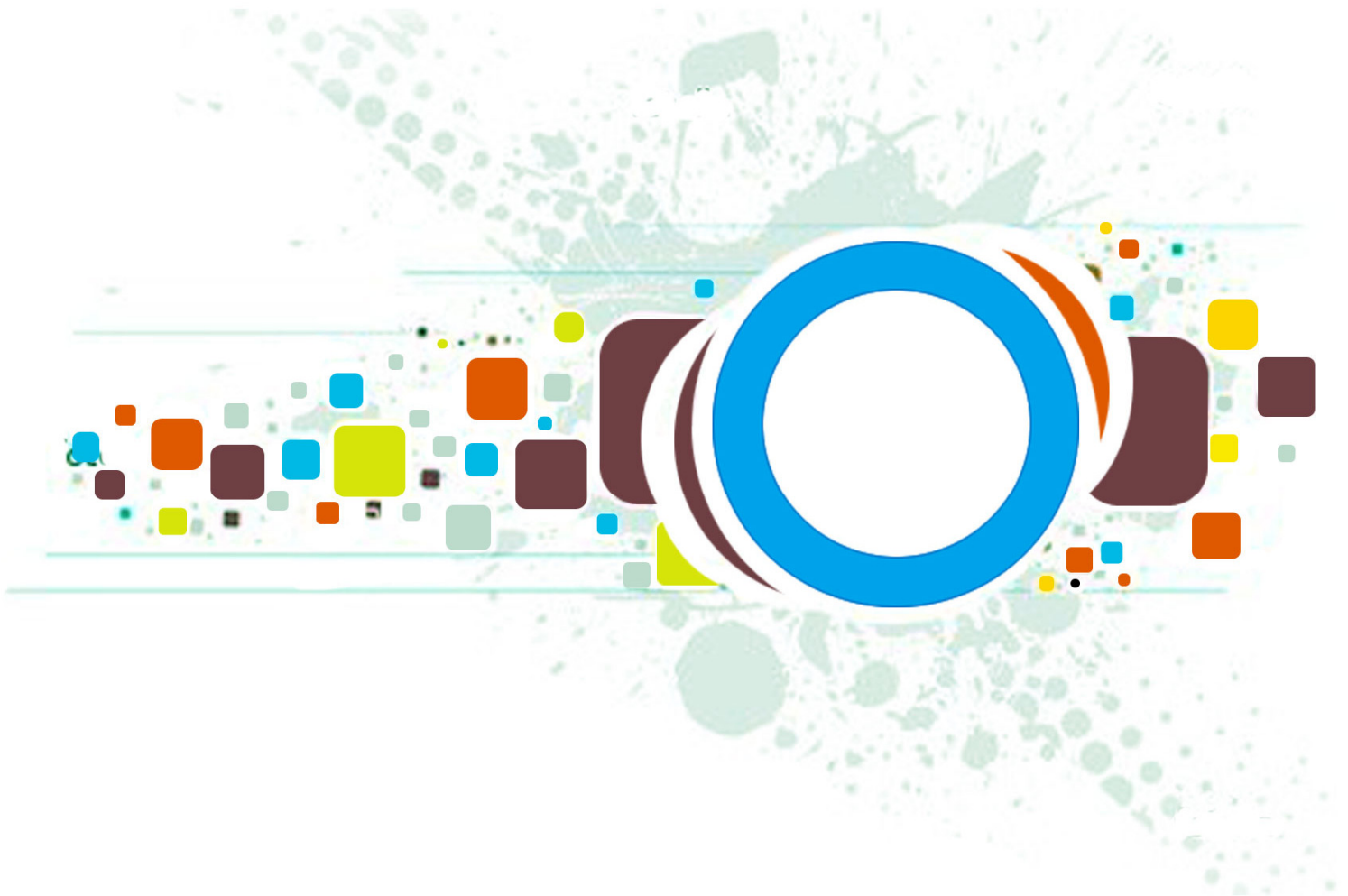
Volume 6 • Issue 5 • October 2012

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>

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

VOLUME 6, ISSUE 5, 2012

**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 6, Issue 5, October 2012

Publishing Date: 24-10- 2012

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, 2012

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 fifth issue of volume six 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 will be appearing 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. Iftekharruddin
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 6, Issue 5, October 2012

Pages

- 273 - 282 A Hybrid SVD Method Using Interpolation Algorithms for Image Compression
J. Izadian, A. Hosaini, Maryam Jalili
- 283 - 296 Learning Based Single Frame Image Super-resolution Using Fast Discrete Curvelet Coefficients
Anil A. Patil, Jyoti Singhai
- 297 - 305 Comprehensive Infrared Image Edge detection Algorithm
Abhishek Gudipalli, Ramashri Tirumala
- 306 - 316 Image Restoration Using Particle Filters By Improving The Scale Of Texture With MRF
Sree Anna Saro Vijendran, Bobby Lukose
- 317 - 325 Fast Segmentation of Sub-cellular Organelles
Dilip Kumar Prasad, Chai Quek, Maylor K. H. Leung
- 326 - 333 Assessing Error Bound For Dominant Point Detection
Dilip K. Prasad
- 334 - 348 Gray Coded Grayscale Image Steganography using Huffman Encoding
Nithyanandam Pandian, Ravichadran Thangavel
- 349 - 356 Walsh, Sine, Haar & Cosine Transform With Various Color Spaces for 'Color to Gray and Back'
H. B. Kekre, Sudeep D. Thepade, Ratnesh Chaturvedi, Saurabh Gupta
- 357 - 362 A Novel Method to Improve Measurement Results of Flame Photometry Using Image Change Detection
Balachandran Ganesan, Sujeevan Kumar Agir

- 363 - 372 Color Image Watermarking using Cycle Spinning based Sharp Frequency Localized
Contourlet Transform and Principal Component Analysis
K.Kishore Kumar, Movva Pavani, V. Seshu Babu
- 373 - 379 Adaptive Multiscale Stereo Images Matching Based on Wavelet Transform Modulus Maxima
Abdelhak EZZINE, Youssef Yazidi Alaoui, Achraf Cohen
- 380 - 388 Robust Block-Matching Motion Estimation of Flotation Froth Using Mutual Information
Anthony Amankwah , Chris Aldrich
- 389 - 396 Comparison of Some Motion Detection Methods in cases of Single and Multiple Moving
Objects
Shamir Alavi

A Hybrid SVD Method Using Interpolation Algorithms for Image Compression

J. Izadian , A. Hosaini

*Department of Mathematics, Faculty of Sciences,
Mashhad Branch, Islamic Azad University,
Mashhad, Iran*

Jalal_Izadian@yahoo.com

M. Jalili

*Department of Mathematics,
Neyshabur Branch, Islamic Azad University,
Neyshabur, Iran*

Jalili.maryam@yahoo.com

Abstract

In this paper the standard SVD method is used for image processing and is combined with some interpolation methods as linear and quadratic interpolation for reconstruction of compressed image. The main idea of the proposed method is to select a particular submatrix of main image matrix and compress it with SVD method, then reconstruct an approximation of original image by interpolation method. The numerical experiments illustrate the performance and efficiency of proposed methods.

Keywords: SVD Method, Interpolation Method, Image Reconstruction, Compressing Image, Lossless Compression, Lossy Compression.

1. INTRODUCTION

Image compression techniques plays an important role in transmission and storage of image information in computer science and related domains. The goal of image compression is to obtain a representation that minimizes the bits volume while still maintaining the important meaning and the intrinsic structure of the original image. Image compression techniques can be classified into two groups. Lossless compression and lossy compression.

In lossless compression the reconstructed image is identical to the original one and deduces a low compression ratio while lossy compression methods allow a loss in the actual image data. So the original image cannot be created exactly from the compressed image. But these methods deduce high compression ratio. There is many ways to compress depending on the application field. One popular method for compressing an image is the wavelet technique [1]. Wavelet functions form an orthonormal basis on which it is possible to project each data set. Another important method is the Singular Value Decomposition (SVD). This technique is based on the factorization of the real matrix of image in three matrices that can be used to reconstruct the main image or an approximation of it, for more information see [2,3,4].

In this paper the SVD method is combined with two interpolation processes, that are linear 2D interpolation (triangular interpolation), and bilinear interpolation, to decrease the volume of transferred image . This paper is organized as follows. Section 2 is specified to describe the basic concepts of proposed method. Section 3 presents the numerical experiments . Finally section 4 concludes the paper with discussion.

2. THE BASIC CONCEPTS OF METHOD

Suppose that a $m \times n$ pixel gray-scale image is given. Each pixel having some level of black and white given by some integers that can be selected between 0 and 255 or a real number

between zero and one. In the case of integer, each integer requires approximately one byte to store. Then the resulting image has approximately $m \times n$ bytes volume. If the image is coloured in (RGB) system the image contains three $m \times n$ matrices for red, green, and blue colours, then $3 \times m \times n$ pixels image. The Singular Value Decomposition (SVD) for a gray-scale image for compressing $m \times n$ pixels image matrix is based on the following decomposition of A ,

$$A = USV^T.$$

Actually, recalling from advanced linear algebra one has the following theorem, that is demonstrated in advanced linear algebra (for example see [5,6]).

Theorem 2.1 Any nonzero real $m \times n$ matrix A with rank $r > 0$, there are an orthogonal $m \times m$ matrix U , and an orthogonal $n \times n$ matrix V such that $U^T A V = S$ is an $m \times n$ "diagonal matrix" of the following form

$$S = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

$$D = \text{diag}(\sigma_1, \dots, \sigma_r), \quad \sigma_1 \geq \sigma_2 \geq \dots \sigma_r > 0,$$

where $\sigma_1, \sigma_2, \dots, \sigma_r$ are non-zero singular values of A . This factorization is called the Singular Value Decomposition (SVD) of A .

By using above theorem and assuming that

$$U = [u_1 \ u_2 \ \dots \ u_r \ u_{r+1} \ \dots \ u_m], \quad i = 1, 2, \dots, m,$$

where $u_i, i = 1, \dots, m$ are the column of U , by considering the orthogonality of U , one has

$$u_i^T u_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Similarly for orthogonal $n \times n$ matrix V , by considering.

$$V^T = \begin{bmatrix} v_1^T \\ v_2^T \\ \dots \\ v_n^T \end{bmatrix},$$

where v_i are for $i = 1, \dots, n$ are columns of V , it yields

$$v_i^T v_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Here, S is $m \times n$ diagonal matrix with singular values of A on the diagonal of D in rectangular matrix S . The matrix S can be presented by the following matrix:

$$S = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 \\ 0 & \sigma_2 & & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix}$$

For $i = 1, 2, \dots, n$, σ_i are singular values of matrix A . It is proved in Theorem(2.1) that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$. For $i = 1, 2, \dots, m$, σ_i are called singular values of

matrix A^T . The vector v_i for $i = 1, 2, \dots, n$ are called the right singular vector of A , and u_i for $i = 1, 2, \dots, m$ are the left singular vectors of A (see[4,5]). then

$$A = USV^T = [u_1 \quad u_2 \quad \dots \quad u_{m-1} \quad u_m] S \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 \sigma_1 v_1^T + u_2 \sigma_2 v_2^T + \dots + u_r \sigma_r v_r^T.$$

When compressing the image, the sum is not performed to the very last. The singular values with small enough values are dropped. A matrix of rank k is obtained by truncating these sums after the first k terms, denoted by

$$A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_k u_k v_k^T, \quad (1)$$

is chosen as approximation of A . The total storage volume for A_k will be $k(m+n+1)$. This matrix is an approximation of A , that can be used as an approximation of compressed image A . For having a criterion for storage volume decrease, the compression ratio is defined as follows:

$$C_R = \frac{mn}{k(m+n+1)}.$$

We also use the Schur norm to measure the quality of obtained image A_k , called Mean Square Error (MSE), which is introduced by following expression:

$$MSE = \frac{1}{mn} \|A - A_k\|_2 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - a_{ij}^{(k)})^2.$$

where $A = (a_{ij})_{m \times n}$ and $A_k = (a_{ij}^{(k)})_{m \times n}$. An error value can be defined related to sum of σ_i by following relation:

$$e_k = 1 - \frac{\sum_{i=1}^k \sigma_i}{\sum_{i=1}^r \sigma_i},$$

A typical choice of k is so that the storage space required for A_k will be less than $\frac{1}{5}$. For having an superior compression ratio, one can utilize a submatrix of A for small k and use an interpolation process to reconstruct A_k .

Consider m and n are even, and \tilde{A} is a submatrix of order $\frac{m}{2} \times \frac{n}{2}$ of A , by omitting the element of A which are situated on rows or column of odd numbers. This matrix is approximated by *SVD* method with a suitable value of k . Then an interpolation method is used to reconstruct an approximation of A_k , that can be chosen as approximation of A_k as obtained image. In linear case, consider the chosen submatrix of A as:

$$\tilde{A}_k = (\tilde{a}_{ij}^{(k)})_{\frac{m}{2} \times \frac{n}{2}},$$

where m and n are even numbers and the elements of A_k can be determined from $A_k = (a_{ij}^{(k)})_{m \times n}$ by the following relations:

$$\begin{aligned} (\tilde{a}_{i,j}^{(k)}) &= a_{2(i-1)+1,2(j-1)+1}^{(k)}, & (\tilde{a}_{i,j+1}^{(k)}) &= a_{2(i-1)+1,2(j-1)+3}^{(k)}, \\ (\tilde{a}_{i+1,j}^{(k)}) &= a_{2(i-1)+3,2(j-1)+1}^{(k)}, & (\tilde{a}_{i+1,j+1}^{(k)}) &= a_{2(i-1)+3,2(j-1)+3}^{(k)}, \\ i &= 1,2,\dots,\frac{m}{2}, & j &= 1,2,\dots,\frac{n}{2}. \end{aligned}$$

The matrix \tilde{A}_k is of order $\frac{m}{2} \times \frac{n}{2}$, therefore it has $\frac{3}{4}$ less element than A_k . This matrix is chosen as the compression of A that results a compression ratio given by

$$C_R = \frac{mn}{k(\frac{m}{2} + \frac{n}{2} + 1)} = \frac{2mn}{k(m+n+2)},$$

Therefore, in this method the compression ratio is doubled in comparison with A_k . The matrix \tilde{A}_k is supposed as a base for interpolating method to reconstruct an approximation of A_k . In the linear case, this matrix is calculated with the following equations:

$$\begin{aligned} A_k^*; A_k, A_k^* &= (a_{ij}^*)_{m \times n}, \\ a_{2(i-1)+1,2(j-1)+1}^* &= \tilde{a}_{ij}^{(k)}, & a_{2(i-1)+3,2(j-1)+1}^* &= \tilde{a}_{i+1,j}^{(k)}, \\ a_{2(i-1)+1,2(j-1)+3}^* &= \tilde{a}_{i,j+1}^{(k)}, & a_{2(i-1)+3,2(j-1)+3}^* &= \tilde{a}_{i+1,j+1}^{(k)}, \\ a_{2(i-1)+2,2(j-1)+1}^* &= (\tilde{a}_{i+1,j}^{(k)} + \tilde{a}_{i,j}^{(k)})/2, & a_{2(i-1)+1,2(j-1)+2}^* &= (\tilde{a}_{i,j}^{(k)} + \tilde{a}_{i,j+1}^{(k)})/2, \\ a_{2(i-1)+3,2(j-1)+2}^* &= (\tilde{a}_{i+1,j}^{(k)} + \tilde{a}_{i+1,j+1}^{(k)})/2, & a_{2(i-1)+2,2(j-1)+3}^* &= (\tilde{a}_{i,j+1}^{(k)} + \tilde{a}_{i+1,j+1}^{(k)})/2, \\ a_{2(i-1)+2,2(j-1)+2}^* &= [\tilde{a}_{i+1,j}^{(k)} + \tilde{a}_{i,j+1}^{(k)} + \tilde{a}_{i,j}^{(k)} + \tilde{a}_{i+1,j+1}^{(k)}]/4. \end{aligned}$$

In order to describe the hybrid *SVD* using bilinear interpolation, the following submatrix of A_k is considered

$$\tilde{A}_k = (\tilde{a}_{ij}^{(k)})_{\frac{m}{3} \times \frac{n}{3}},$$

where A_k is considered a $m \times n$ matrix which m and n are the integers that are divisible by 3. Then \tilde{A}_k can be defined as follows

$$\begin{aligned} A_k &= (a_{ij}^{(k)})_{m \times n}, \\ (\tilde{a}_{i,j}^{(k)}) &= a_{3(i-1)+1,3(j-1)+1}^{(k)}, & (\tilde{a}_{i,j+1}^{(k)}) &= a_{3(i-1)+1,3(j-1)+4}^{(k)}, \\ (\tilde{a}_{i+1,j}^{(k)}) &= a_{3(i-1)+4,3(j-1)+1}^{(k)}, & (\tilde{a}_{i+1,j+1}^{(k)}) &= a_{3(i-1)+4,3(j-1)+4}^{(k)}, \end{aligned}$$

$$i = 1, 2, \dots, \frac{m}{3}, \quad j = 1, 2, \dots, \frac{n}{3}.$$

Since the matrix \tilde{A}_k is a matrix of order $\frac{m}{3} \times \frac{n}{3}$, the compression ratio is obtained by

$$C_R = \frac{mn}{n(\frac{m}{3} + \frac{n}{3} + 1)} = \frac{3mn}{k(m+n+3)},$$

Now for each 4×4 block of matrix A_k which is reduced to 2×2 block matrix \tilde{A}_k , the following bilinear interpolation is applied on $[0, 1]^2$,

$$\tilde{P}^{(i,j)}(x, y) = \sum_{r=1}^2 \sum_{s=1}^2 L_{r,s}(x, y) \tilde{a}_{i+(r-1), j+(s-1)}$$

where $L_{r,s}$ is defined by the following equations:

$$L_{11}(x, y) = \frac{(x-1)(y-1)}{(0-1)(0-1)}, \quad L_{12}(x, y) = \frac{(x-1)(y-0)}{(0-1)(1-0)},$$

$$L_{21}(x, y) = \frac{(x-0)(y-1)}{(1-0)(0-1)}, \quad L_{22}(x, y) = \frac{(x-0)(y-0)}{(1-0)(1-0)}.$$

or

$$L_{11}(x, y) = (x-1)(y-1), \quad L_{12}(x, y) = y(1-x),$$

$$L_{21}(x, y) = x(1-y), \quad L_{22}(x, y) = xy.$$

The matrix A_k^* of order $m \times n$ which is an approximation of A_k is given as follows:

$$A_k^* = (a_{ij}^*)_{m \times n} \cong (a_{ij}^{(k)})_{m \times n},$$

where a_{ij}^* can be determined by the following formulae

$$a_{3(i-1)+1, 3(j-1)+1}^* = \tilde{a}_{i,j}^{(k)}, \quad a_{3(i-1)+4, 3(j-1)+1}^* = \tilde{a}_{i+1,j}^{(k)},$$

$$a_{3(i-1)+1, 3(j-1)+4}^* = \tilde{a}_{i,j+1}^{(k)}, \quad a_{3(i-1)+4, 3(j-1)+4}^* = \tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+2, 3(j-1)+1}^* = \tilde{P}^{(i,j)}\left(\frac{1}{3}, 0\right) = \frac{2}{3} \tilde{a}_{ij}^{(k)} + \frac{1}{3} \tilde{a}_{i+1,j}^{(k)},$$

$$a_{3(i-1)+3, 3(j-1)+1}^* = \tilde{P}^{(i,j)}\left(\frac{2}{3}, 0\right) = \frac{1}{3} \tilde{a}_{ij}^{(k)} + \frac{2}{3} \tilde{a}_{i+1,j}^{(k)},$$

$$a_{3(i-1)+1, 3(j-1)+2}^* = \tilde{P}^{(i,j)}\left(0, \frac{1}{3}\right) = \frac{2}{3} \tilde{a}_{ij}^{(k)} + \frac{1}{3} \tilde{a}_{i,j+1}^{(k)},$$

$$a_{3(i-1)+1, 3(j-1)+3}^* = \tilde{P}^{(i,j)}\left(0, \frac{2}{3}\right) = \frac{1}{3} \tilde{a}_{ij}^{(k)} + \frac{2}{3} \tilde{a}_{i,j+1}^{(k)},$$

$$a_{3(i-1)+4,3(j-1)+2}^* = \tilde{P}^{(i,j)}\left(1, \frac{1}{3}\right) = +\frac{2}{3}\tilde{a}_{i+1,j}^{(k)} + \frac{1}{3}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+4,3(j-1)+3}^* = \tilde{P}^{(i,j)}\left(1, \frac{2}{3}\right) = +\frac{1}{3}\tilde{a}_{i+1,j}^{(k)} + \frac{2}{3}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+2,3(j-1)+4}^* = \tilde{P}^{(i,j)}\left(\frac{1}{3}, 1\right) = +\frac{2}{3}\tilde{a}_{i,j+1}^{(k)} + \frac{1}{3}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+3,3(j-1)+4}^* = \tilde{P}^{(i,j)}\left(\frac{2}{3}, 1\right) = +\frac{1}{3}\tilde{a}_{i,j+1}^{(k)} + \frac{2}{3}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+2,3(j-1)+2}^* = \tilde{P}^{(i,j)}\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{4}{9}\tilde{a}_{ij}^{(k)} + \frac{2}{9}\tilde{a}_{i,j+1}^{(k)} + \frac{2}{9}\tilde{a}_{i+1,j}^{(k)} + \frac{1}{9}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+2,3(j-1)+3}^* = \tilde{P}^{(i,j)}\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{2}{9}\tilde{a}_{ij}^{(k)} + \frac{4}{9}\tilde{a}_{i,j+1}^{(k)} + \frac{1}{9}\tilde{a}_{i+1,j}^{(k)} + \frac{2}{9}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+3,3(j-1)+2}^* = \tilde{P}^{(i,j)}\left(\frac{2}{3}, \frac{1}{3}\right) = \frac{2}{9}\tilde{a}_{ij}^{(k)} + \frac{1}{9}\tilde{a}_{i,j+1}^{(k)} + \frac{4}{9}\tilde{a}_{i+1,j}^{(k)} + \frac{2}{9}\tilde{a}_{i+1,j+1}^{(k)},$$

$$a_{3(i-1)+3,3(j-1)+3}^* = \tilde{P}^{(i,j)}\left(\frac{2}{3}, \frac{2}{3}\right) = \frac{1}{9}\tilde{a}_{ij}^{(k)} + \frac{2}{9}\tilde{a}_{i,j+1}^{(k)} + \frac{2}{9}\tilde{a}_{i+1,j}^{(k)} + \frac{4}{9}\tilde{a}_{i+1,j+1}^{(k)},$$

In fact, one has

$$a_{3(i-1)+(1+r),3(j-1)+(1+s)}^* = \tilde{P}^{(i,j)}\left(\frac{r}{3}, \frac{s}{3}\right) \quad r, s = 0, 1, 2, 3$$

Then, A^* is obtained as suitable approximation of A_k .

3. NUMERICAL EXPERIMENTS

In this section some test examples are presented. In these examples we compare visually and numerically the compressing. The reconstructed image and the compression ratio and, the mean square error (MSE) in each case are presented. In the tables, L_I , and BI , are used for linear and bilinear interpolation, respectively.

3.1 Example

In this example we consider a 144×300 pixels black and white image from MATLAB gallery. The initial and final images are given in Figure 1.(a), and (b), the results for simple SVD linear interpolation and bilinear interpolation are presented in Figures 2. and 3, respectively. In Table 3.1. numerical results for three methods are given.

<i>method</i>	<i>k</i>	<i>e</i>	<i>CR</i>	<i>MSE</i>	<i>cpu time</i>
SVD	9	0.4304	10.78658	1.2458×10^{-4}	8.7622
SVD	16	0.3384	6.0674	8.0855×10^{-5}	8.6436
SVD	25	0.258	3.8831	5.5209×10^{-5}	8.0967
SVD	36	0.1889	2.6966	3.9410×10^{-5}	9.4412
SVD,LI	9	0.3873	10.7143	2.5198×10^{-4}	8.1953
SVD,LI	16	0.2860	6.00268	1.6349×10^{-4}	8.4375
SVD,LI	25	0.1985	3.8571	1.1551×10^{-4}	8.7658
SVD,LI	36	0.1228	2.6786	8.1363×10^{-5}	11.4910
SVD,BI	9	0.3335	95.3642	4.001×10^{-4}	7.3140
SVD,BI	16	0.2308	53.6424	2.5966×10^{-4}	7.9940
SVD,BI	25	0.1579	34.3311	1.6774×10^{-4}	11.3992
SVD,BI	36	0.0454	23.8411	1.0217×10^{-4}	12.6754

TABLE 3.1: The results for black and white image.



FIGURE 1: The results of SVD method for k=36.



FIGURE 2: The results of SVD and linear interpolation method for k=36.



FIGURE 3: The results of SVD and bilinear interpolation method for k=36.

3.2 Example

In this example a $3 \times 384 \times 512$ pixel colored image from MATLAB gallery is selected. The initial and final images are given in Figure 4. and the results for SVD linear interpolation , and bilinear interpolation are shown in Figure 4-5, respectively. The numerical results are presented in Table 3.2.

<i>method</i>	<i>M</i>	<i>e</i>	<i>CR</i>	<i>MSE</i>	<i>time</i>
SVD	9	0.4913	24.3538	5.9974×10^{-5}	25.9324
SVD	16	0.3969	13.6990	3.3733×10^{-5}	26.5473
SVD	25	0.32242	8.7674	1.9898×10^{-5}	26.3382
SVD	36	0.2685	6.0884	1.3075×10^{-5}	28.4017
SVD,LI	9	0.7508	24.2726	2.2499×10^{-4}	11.3439
SVD,LI	16	0.7158	13.6533	2.0673×10^{-4}	11.8093
SVD,LI	25	0.7084	8.7381	2.0631×10^{-4}	12.4191
SVD,LI	36	0.6945	6.0681	2.0611×10^{-4}	14.2736
SVD,BI	9	0.7370	217.0066	7.0313×10^{-4}	8.7288
SVD,BI	16	0.7001	122.0662	6.9817×10^{-4}	10.8634
SVD,BI	25	0.6922	78.1224	6.9801×10^{-4}	12.1687
SVD,BI	36	0.6535	54.2517	6.978×10^{-4}	13.1736

TABLE 3.2: The results for black and white.



FIGURE 4: The results for SVD method for k=36.



FIGURE 5: The results for SVD and linear interpolation method.



FIGURE 6: The results for SVD and bilinear interpolation method.

3.3 Example

In this example an other colored image is chosen from MATLAB gallery which has $3 \times 318 \times 318$ pixels. The initial and final images are given in Figures 7-9. and the results for simple SVD linear

interpolation, and bilinear interpolation are shown in the same Figures, respectively. The numerical results are presented in Table 3.3.

<i>method</i>	<i>M</i>	<i>m_i</i>	<i>n_i</i>	<i>e</i>	<i>CR</i>	<i>MSE</i>	<i>time</i>
SVD	9	3	3	0.4304	17.6389	8.8662×10^{-5}	17.8987
SVD	16	4	4	0.3384	9.9212	4.4780×10^{-5}	18.5721
SVD	25	5	5	0.2580	6.3500	2.9478×10^{-5}	25.36016
SVD	36	6	5	0.1889	4.4097	1.968×10^{-5}	32.3716
SVD,LI	9	3	3	0.7493	17.5562	3.5891×10^{-4}	10.3739
SVD,LI	16	4	4	0.7192	9.8754	3.5681×10^{-4}	11.7131
SVD,LI	25	5	5	0.6943	6.3202	3.5630×10^{-4}	13.90725
SVD,LI	36	6	5	0.6738	4.3891	3.5615×10^{-4}	17.5310
SVD,BI	9	3	3	0.7330	155.7969	0.0016	9.1011
SVD,BI	16	4	4	0.6995	87.6358	0.0016	10.1484
SVD,BI	25	5	5	0.6720	56.0869	0.0016	12.1943
SVD,BI	36	6	5	0.6535	38.9491	0.0016	13.17828

Table 3.3: The results for colour image .



FIGURE 7: The results for SVD method for k=36.



FIGURE 8: The results for SVD and linear interpolation method for k=36



FIGURE 9: The results of SVD and bilinear interpolation method for $k=36$

4. CONCLUSION

In this paper two hybrid *SVD* methods, using linear interpolation, and bilinear interpolation are presented. The results show the preference of bilinear interpolation combined with *SVD*. The linear method is also economically acceptable.

5. REFERENCES

- [1] J. Stoer, R. Bulirsch. Introduction to Numerical Analysis. translated by R.Bartels , W.Gautschi. C.Witzgall. Springer-Verlag, second edition, (1991).
- [2] H. Abdi. "Singular Value Decomposition (*SVD*) and Generalized Singular Value Decomposition (*GSVD*).". In N.J. Salkind (Ed.): Encyclopedia of Measurement and Statistics. Thousand Oaks (CA): Sage. pp. 907-912, 2007.
- [3] R. Ashino, A. Morimoto, M. Nagase, R.Vaillancourt. "Image Compression with Multiresolution Singular Value Decomposition and Other Methods." Mathematical and Computer Modelling. Vol. 41(6-7). pp. 773-790, 2005.
- [4] P.Dostert. "An application of linear algebra to image compression." The university of arizona. July 2009.
- [5] M.Sing Song. "Wavelet Image Compression." *Mathematical subject classification*, 1991.
- [6] M.Sing Song. "Wavelet Image Compression", Southern Illinois University Edwardsville, Edwardsville, IL62026, USA, 2004.

Learning Based Single Frame Image Super-resolution Using Fast Discrete Curvelet Coefficients

Anil A. Patil

*Department of Electronics & Telecommunication
COE, Malegaon(Bk),
Pune, India.*

panil21mail@gmail.com

Jyoti Singhai

*Department of Electronics & Communication Engg.
MANIT, Bhopal, India.*

j.singhai@gmail.com

Abstract

High-resolution (HR) images play a vital role in all imaging applications as they offer more details. The images captured by the camera system are of degraded quality due to the imaging system and are low-resolution (LR) images. Image super-resolution (SR) is a process, where HR image is obtained from combining one or multiple LR images of same scene. In this paper, learning based single frame image super-resolution technique is proposed by using Fast Discrete Curvelet Transform (FDCT) coefficients. FDCT is an extension to Cartesian wavelets having anisotropic scaling with many directions and positions, which forms tight wedges. Such wedges allow FDCT to capture the smooth curves and fine edges at multiresolution level. The finer scale curvelet coefficients of LR image are learnt locally from a set of high-resolution training images. The super-resolved image is reconstructed by inverse Fast Discrete Curvelet Transform (IFDCT). This technique represents fine edges of reconstructed HR image by extrapolating the FDCT coefficients from the high-resolution training images. Experimentation based results show appropriate improvements in MSE and PSNR.

Keywords: Super-resolution, Multiresolution, Learning Method, Fast Discrete Curvelet Transform

1. INTRODUCTION

High resolution (HR) images are essential in almost all imaging applications. HR images provide additional details that are essential, to the success of various applications that require accurate image analysis such as image/video resolution enhancement, medical imaging, remote sensing and video surveillance, etc. The image captured by current camera system is of degraded quality because of the resolution of a digital imaging device which is mainly limited by the number of pixels on the sensor and the optical system. Such captured images are acknowledged as low-resolution images (LR). The spatial resolution of the low resolution image can be improved by increasing the pixel density or by growing the chip-size. Both techniques enforce certain limitations. The spatial resolution of a single image (frame) can be enhanced by the traditional interpolation techniques. It has also limited application since aliasing is present. The inherent limitations of the current camera system are prevailed over by the technique known as Super-resolution (SR) based on digital signal processing. SR is a technique where one or many low resolution images (frames) are combined to obtain a high spatial resolution image. In this process low resolution image is upsampled by recovering the missing high frequency details and degradations are impassive. Super-resolution reconstruction technique is basically classified in two categories as reconstruction based and learning based methods. In reconstruction based methods high-resolution image is obtained from several low-resolution observations by proper fusion of a series of accurately registered aliased images. Most of the literature available on super-resolution is for multi-frame and majority of them are based on the motion as cue. The super-resolution idea was introduced by Tsai and Hung, where a pure translation motion has been considered [1]. In such methods the quality of reconstructed SR image obtained from a set

of LR images depends upon the registration accuracy of the LR images and some prior knowledge of imaging system [2, 3]. The ideal sub-pixel displacement errors in the observation with noise can be used to obtain a high resolution image by constrained total least square algorithm [4] and their effect on the convergence rate of iterative approach is discussed by Ng and Bose [5]. All these iterative methods are computationally complex. In many resolution enhancement applications the point spread function (PSF) of the imaging system is unknown which reflects in blur. Nguyen et al. propose a technique for parametric blur identification and regularization based on generalized cross-validation (GVC) theory where circulant block preconditioners are used to accelerate the conjugate gradient decent (CG) method for solving the Tikhonor-regularization super-resolution problem [6]. Single observed image expansion by interpolation or super-resolution smooth the image data at edge regions. A maximum a posteriori (MAP) estimator with Huber-Markove random field (MRF) prior can be used for image expansion [7]. In recent work blur, defocus, zoom are used as cues for super-resolving the low resolution image [8, 9]. The SR image problem can be solved by combining, maximum likelihood (ML) estimator, MAP estimator and projection onto convex sets (POCS) to find unified approach [10]. Learning methods becomes more useful, when only a single observation is available and several other high resolution images are present in the data set. All high resolution images from data set will act as training images. In many realistic applications, such as in biometry, criminal surveillance a single observation is available. Nearly all SR reconstruction algorithms are based on the fundamental constraints that provide less useful information as the magnification factor increases. Baker and Kanade found these limitations and developed a SR algorithm by modifying the prior term in cost to include the result of a set of recognition called as recognition based super-resolution or hallucination [11]. A similar approach is used for multiple views using learned image models through use of principal component analysis (PCA) Capel and Zisserman [12]. A fast and simple one pass example based super-resolution algorithm is proposed by Freeman et al. which is resolution independent [13]. Zoomed observations have been used to reconstruct super-resolved image by learning the high resolution image model through most zoomed observations [14]. In SR, handling of data at different resolution level is tedious. The multiresolution technique can handle data well. In wavelet transform based learning technique high frequency coefficients of the unknown high-resolution image are learned locally from a set of training images based on best match by Jiji et al. [15]. Chang et. al. have proposed method where the generation of the high resolution image patch depends simultaneously on multiple nearest neighbors in the training set [16]. Face is represented by linear combination of prototypes of shapes and texture in top-down learning of LR facial images by Park and Lee [17]. An image hallucination approach based on primal sketch prior with reconstruction constraint to improve the quality is proposed by Sun et al. [18]. Multiscale, directional Contourlet transform is proposed by M.N.Do and M.Vetterli. It decomposes image in scales with directional subbands thus captures the smoothness along the contours [19]. Jiji and Chaudhuri have proposed a single frame image super-resolution through Contourlet learning which learns the best edge primitives from the HR training set in contourlet domain [20]. Recently a frame work to combine power of reconstruction and learning based methods is proposed, where redundancy within single image is used. It requires registration at subpixel accuracy and is computationally complex [24]. New learning based SR technique using DWT and IGMRF prior improves the quality of reconstructed image [23]. By introducing the statistics to learning-based super-resolution with global and local constraints to obtain high quality reconstructed image is proposed by Kim et al. [25]. A SR problem can be resolved by kernel ridge regression (KRR) with a prior model of a generic image class is suggested by Kim and Kwon [26]. In this paper a geometric curvelet transform (FDCT) [21, 22] is proposed to learn the salient features from the data set while upsampling the input image. FDCT gives sparse representation of C^2 function away from edges along piecewise smooth curves. The super-resolved image represents the learnt features of low resolution oriented edges from the high resolution data set. This leads to significant improvement in quality of the reconstructed image.

This paper is organized as follows: imaging model is shown in section 2. Section 3 discusses the Fast discrete curvelet transform. Procedure to learn the fast discrete curvelet coefficients is

discussed in section 4. Experimentation results on different gray and color images are discussed in section 5. Paper concludes with section 6.

2. THE IMAGING MODEL

In this paper, the following linear image formation model is adapted to obtain observed low resolution image.

$$y = DBx + n \tag{1}$$

where x is a original high resolution image represented lexicographically ordered of $M^2 \times 1$ pixel size, D is decimation matrix and B is the blur matrix. The size of the decimation matrix depends upon the decimation factor and is not invertible. y is a low resolution observation obtained, and represented lexicographically in order of $N^2 \times 1$ pixel size and n is noise vector of size $M^2 \times 1$. Here noise is assumed to be zero mean independent identically distributed and blur to be an identity matrix. For a single observation y , high resolution image x is estimated. The imaging model is illustrated in Figure 1.

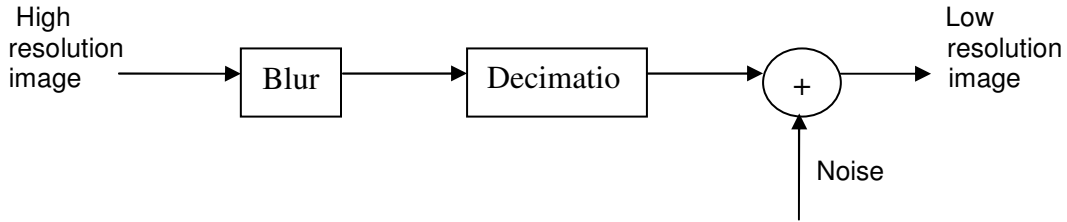


FIGURE 1: Illustration of low-resolution image formation model

3. FAST DISCRETE CURVELET TRANSFORM

Fast Discrete Curvelet transform (FDCT) gives local components at different frequencies for analysis and synthesis of digital image in multi-resolution analysis. FDCT is multi-scale geometric transform, which is a multi-scale pyramid with many directions and positions at each length scale. FDCT is basically 2D anisotropic extension to classical wavelet transform that has main direction associated with it. Analogous to wavelet, FDCT can be translated and dilated. The dilation is given by a scale index that controls the frequency content of the curvelet with the indexed position and direction can be changed through a rotation. This rotation is indexed by an angular index. Curvelet satisfy anisotropic scaling relation, which is generally referred as parabolic scaling. This anisotropic scaling relation associated with FDCT is a key ingredient to the proof that curvelet provides sparse representation of the C^2 function away from edges along piecewise smooth curves. FDCT is constructed by a radial window W and angular window V . The radial window W is expressed as

$$\tilde{W}_j(w) = \sqrt{\phi_{j+1}^2(w) - \phi_j^2(w)} \quad , \quad j \geq 0 \tag{2}$$

Where, j is scale and ϕ is defined as the product of low-pass one dimensional window and separate scales in Cartesian equivalents. The angular window V is defined as

$$V_j(w) = V(2^{\lfloor j/2 \rfloor} w_2 / w_1) \tag{3}$$

where, W_1 and W_2 are low pass one dimensional windows. The Cartesian window $\tilde{U}_{j,l}$ is constructed by combining radial window W and angular window V and is expressed as

$$\tilde{U}_{j,l}(w) = W_j(w) V_j(S_{\theta_l} w) \tag{4}$$

Where, the angle θ_l have same slope but are not equally spaced. S_{θ} is shear matrix,

$S_\theta = \begin{pmatrix} 1 & 0 \\ \tan \theta & 1 \end{pmatrix}$. Shear matrix S_θ is used to maintain the symmetry around the origin and rotation by $\pm \Pi/2$ radiance. The family $\tilde{U}_{j,l}$ implies a concentric tiling whose geometry is shown in Figure 2. The shaded region represents a wedge.

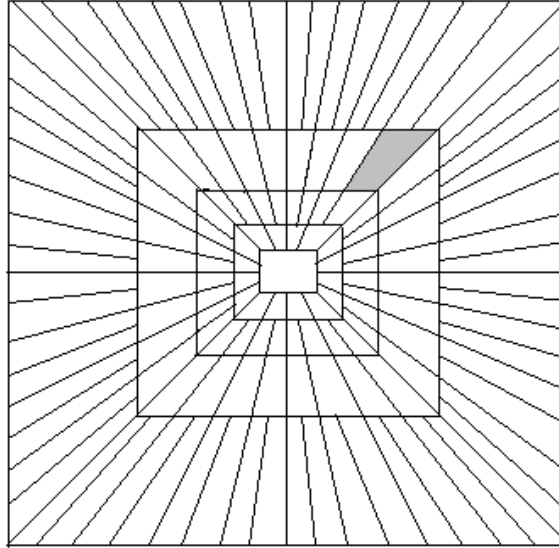


FIGURE 2: Basic digital tiling.

The above construction gives pseudopolar tiling, an alternative to ideal polar tiling. In FDCT via wrapping the curvelets at a wedge are wrapped for a given scale and angle by translating the curvelets on regular rectangular grid which is same for every angle within each quadrant with proper orientation. The frequency domain definition of digital curvelet is,

$$\overline{\phi_{j,l,k}^D}[t_1, t_2] = \hat{U}_j[t_1, t_2] e^{-i2\Pi[k_1 t_1 + k_2 t_2]} \quad (5)$$

where, $\hat{U}_j[t_1, t_2]$ is Cartesian window. Here the discrete localizing window $\hat{U}_{j,l}[n_1, n_2]$ does not fit in a rectangle, aligned with axes. At each scale j , there exist two constants $L_{1,j} \approx 2^j$ and $L_{2,j} \approx 2^{j/2}$ such that for every orientation θ_l , one can tile the two-dimensional plane which translates the respective rectangle by multiples of $L_{1,j}$ in the horizontal direction and $L_{2,j}$ in the vertical direction. The windowed data is wrapped around the origin. The correspondence between the wrapped and original indices is one to one where the wrapping transformation is reindexing of the data. Figure 3 illustrates the wrapping process.

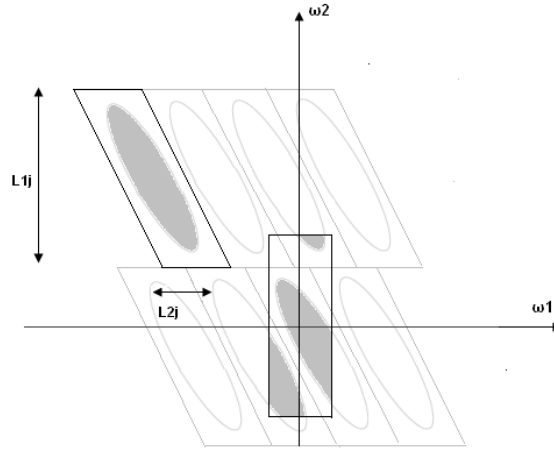


FIGURE 3: Illustrate the wrapping process

Discrete Curvelet transform is expressed as

$$c^D(j, l, k) = \sum_{0 \leq t_1, t_2 < n} f[t_1, t_2] \overline{\phi_{j,l,k}^D[t_1, t_2]} \quad (6)$$

where, $c^D(j, l, k)$ represents curvelet coefficients with j as scale parameter, l as orientation parameter and k as position parameter. $f[t_1, t_2]$ is an input of Cartesian arrays [16]. This transform is also invertible. To illustrate the FDCT decomposition, a standard Lena image is decomposed at three levels with eight orientations as shown in Figure 4. The coarser first level FDCT coefficients are low frequency components and are wavelet coefficients because they do not have orientation. Second finer scale gives eight different oriented wedges where as the third finest scale gives sixteen corresponding oriented wedges. The number of wedges in a subband increases by a factor of two only after every other scale. Here FDCT via wrapping is used as it is simple to implement.

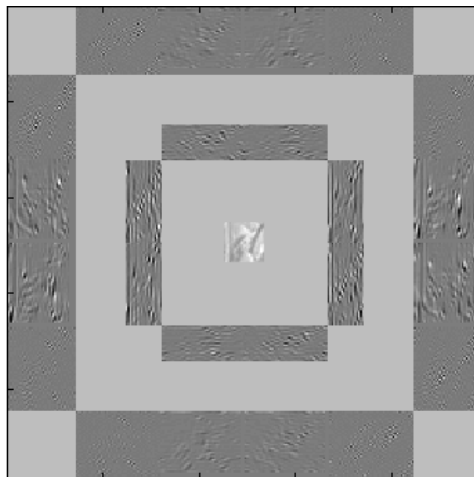


FIGURE 4: Decomposition Lena image at two scale with eight directions in FDCT domain

4. LEARNING FAST DISCRETE CURVELET COEFFICIENTS

The decomposition of input image in Fast Discrete Curvelet Transform (FDCT) domain gives coarser scale coefficients and finer scale coefficients. The coarser scale subband contains low frequency components which are wavelet components as they do not have orientation. The finer scale subbands contain directional high frequency components. The high frequency components at finer scale of low resolution test image are estimated by learning them from high resolution training images in dataset. The learning procedure is performed by finding the minimum Euclidean distance between coefficients of low-resolution test image and all high-resolution training images in particular subband. This gives best closeness of the coefficients. To represent an image fine edges are important. In this proposed work 4x4 pixel block is considered as edge primitive element localized in low-resolution test image and the corresponding 8x8 pixel block in the high resolution image. The low resolution test image is decomposed at two levels and all the training images in data set are decomposed at three levels in FDCT domain. The size of the test image is of $M \times M$ pixel where as all training images are of $2M \times 2M$ size. Figure 5 illustrates the learning of fast discrete curvelet coefficients at the finer scale from a set of N training images.

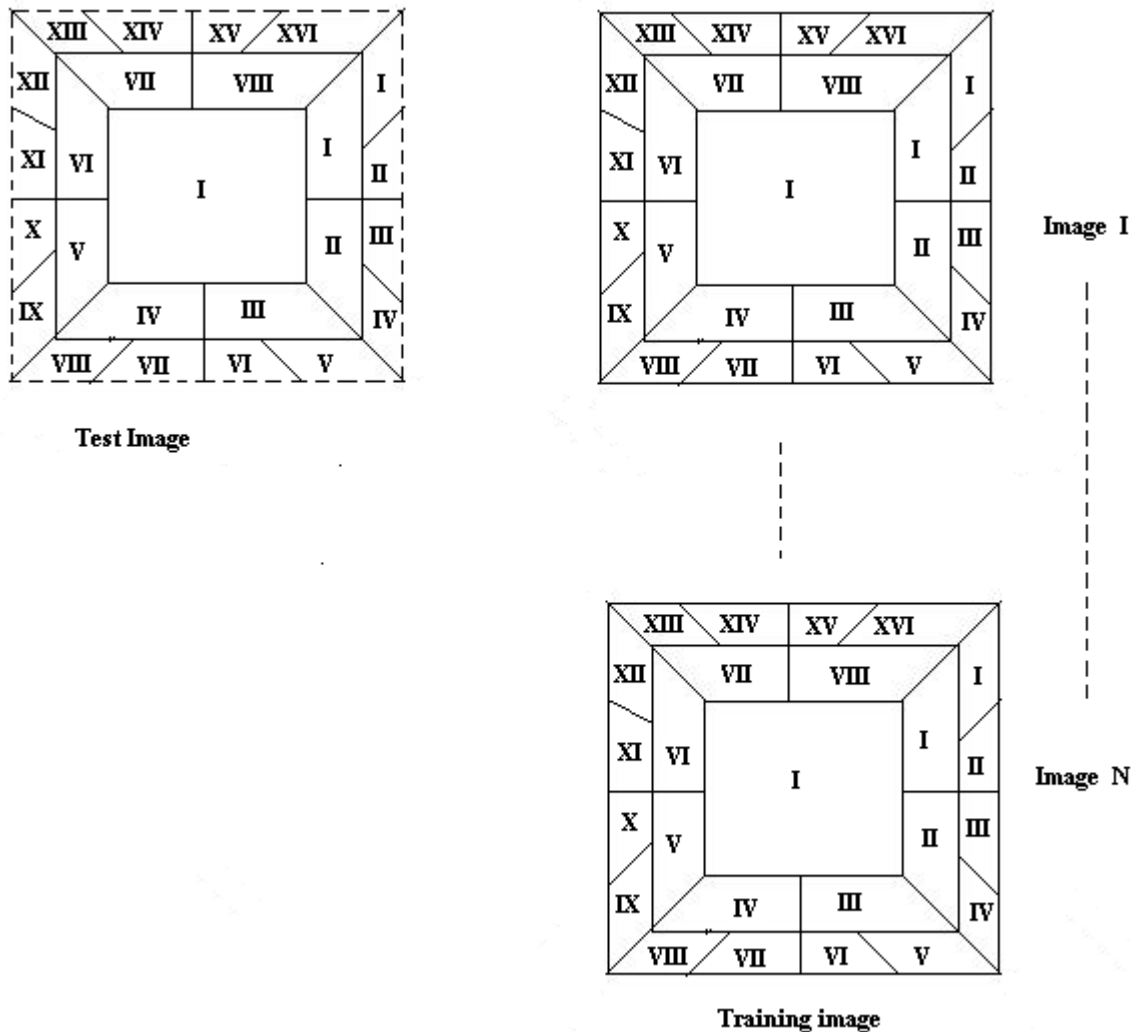


FIGURE 5: Illustrates the learning process in FDCT domain

The minimum Euclidean distance is computed by comparing low-resolution test image edge primitive element with corresponding high-resolution training images at second coarser scale. The minimum Euclidean distance gives best match of edge primitive element. The finer scale

coefficients of low-resolution test image at particular subband are learnt from the corresponding third scale coefficients of high-resolution training images in dataset. Zero tree concept is applied to learn the curvelet coefficients at finer scale of the low-resolution test image. To find the best match, minimum Euclidean distance of curvelet coefficients at $I-VIII$ subbands of low-resolution test image and the corresponding subbands of the high-resolution training images are considered. The location (l, k) as $d(l, k)$ at $I-VIII$ subbands of the test image is searched for best match in all the training images in particular subband. The best matched curvelet coefficients at the finest scale of a training image for the given location are copied to the corresponding subbands $I-XVI$ of low-resolution test image. The formula used to calculate the minimum Euclidean distance is as follows:

$$\begin{aligned}
 m(\hat{p}, \hat{q}) = \arg \min_{m,p,q} & \left[\left| d_{IS_2}(l, k) - d_{I(m)S_2}(p, q) \right|^2 + \left| d_{IIS_2}(l, k) - d_{II(m)S_2}(p, q) \right|^2 \right. \\
 & + \left| d_{IIIS_2}(l, k) - d_{III(m)S_2}(p, q) \right|^2 + \left| d_{IIVS_2}(l, k) - d_{IV(m)S_2}(p, q) \right|^2 + \left| d_{IVS_2}(l, k) - d_{V(m)S_2}(p, q) \right|^2 \\
 & \left. + \left| d_{VIS_2}(l, k) - d_{VIS_2}(p, q) \right|^2 + \left| d_{VIIS_2}(l, k) - d_{VII(m)S_2}(p, q) \right|^2 + \left| d_{VIIS_2}(l, k) - d_{VIII(m)S_2}(p, q) \right|^2 \right]
 \end{aligned} \tag{7}$$

Here $d_J(m)S$ denotes the curvelet coefficients for the m th training image at the J th subband, at scale S and $m=1, 2, \dots, K$. The best match of the test image at (l, k) th location is (\hat{p}, \hat{q}) th location of m th training image at scale S . The best match of 8×8 pixel area is copied at the corresponding subbands of the test image. So the 8×8 pixel area in the low resolution test image is learnt from different training images independently. When 4×4 block at second level does not find a good match in the training data set minimum distance is too large. It shows that high resolution representation is not available in the dataset which introduces artifacts in the reconstructed image. Such artifacts can be avoided by choosing certain threshold for minimum distance. First scale curvelet coefficients are lowpass coefficients and are not considered while calculating minimum Euclidean distance.

The learning algorithm is given in steps as below:

- 1) Perform two-level FDCT decomposition on the low-resolution test image of size $M \times M$.
- 2) Perform three-level FDCT decomposition on all the training images each of size $2M \times 2M$.
- 3) Find minimum Euclidean distance by considering 4×4 pixel block of curvelet coefficients in subbands $I-VIII$ at second level of low resolution test image and all high resolution images in training data set.
- 4) If minimum Euclidean distance $<$ threshold, copy 8×8 block of curvelet coefficients in subband $I-XVI$ of high resolution training image at third level to the corresponding $I-XVI$ subband of low-resolution test image else set them zero.
- 5) Repeat steps 3-4 for every 4×4 block of curvelet coefficients in $I-VIII$ subbands at second level of low resolution image.
- 6) Perform inverse FDCT transform to obtain the high-resolution image of the given test image.

In post processing histogram specification technique is used for image enhancement. The finest level curvelet coefficients of the low resolution test image are learnt from the high-resolution training images in dataset. In SR, during upsampling process edges get blurred. The edge primitive element is used to learn the LR edge from its HR representation locally. Here 4×4 pixel block is selected as edge primitive element in the LR image and the corresponding 8×8 pixel area in the HR image. Each local region is learned independently from HR data set. This edge primitive gives the localization in the particular subband of the image. Large edge primitive gives

poor localization and better matching where as small edge primitive gives better localization but more fails matching.

5. EXPERIMENTAL RESULTS

Experiments are carried out on different types of images to find the efficiency of the proposed algorithm. Arbitrary high resolution gray images of different objects are downloaded from the internet to form the training data set. The training images in the data set considered are of 100. A high resolution gray image which does not belong to training data set is considered to obtain a low resolution image. This low resolution image is obtained through imaging model and upsampled to obtain a super-resolution image. Figure 6 (a) shows a low-resolution image of old man and Figure 6 (b) is high resolution image. Bicubic interpolated image is shown in Figure 6(c). The interpolated image is blurred one and cannot get the details of eye and lines on the forehead. The image shown in Figure 6 (d) is super-resolved by using Contourlet transform. The image shows the fine details of the eye and the lines on the forehead face. Image shown in Figure 6(e) is super-resolved by proposed method. The super-resolved image shows every detail of the eyes, lines on the forehead and face. The details of the eye brows can be visualized in the image and image is sharp.

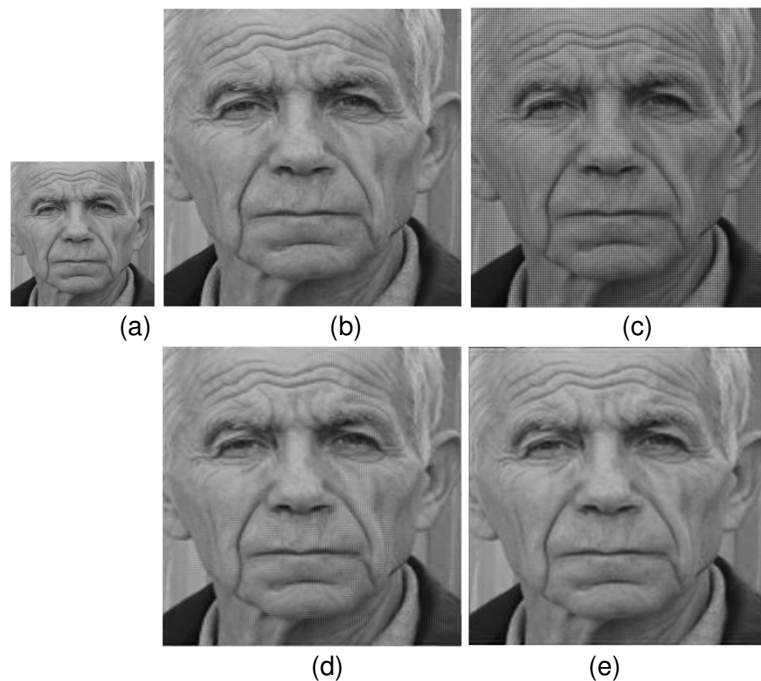


FIGURE 6: a) Low resolution image b) Original High resolution image c) Bicubic interpolated image d) Super-resolved image by using Contourlet transform e) Super-resolved image by using proposed method.

Image shown in Figure 7 (a) is low resolution image of a girl and original high resolution image is shown in figure 7(b). Image shown in figure 7(c) is bicubic interpolated image which is blurred one. The details of texture on the transparent hat and, the shadow on the face appear to be blurred. The details of eye have lost and the hairs found to be total undistinguishable. Figure 7 (d) shows the super-resolved image by using Contourlet transform. The texture on hat, shadow on face and the details of eyes are sharp and details have been retained. Image shown in figure 7(e) is of the proposed method. The particulars of the texture on hat have been reconstructed without blur. Black Strips on the hat have been reconstructed without blur and edge of hat is artifact free. The details of the eye balls and eye brow can be observed clearly. The hair strands can be distinguishable. Image looks sharper. Image shown in figure 8(a) is of low resolution Barbara image, cropped from the original image and figure 8(b) is the high resolution version of it. Bicubic interpolated image is shown in figure 8(c) where the blurred image have been observed at edges.

Strips on scarf appeared to be blurred. Image shown in figure 8(d) is super-resolved by using the Contourlet transform. Image shows sharp strips on the scarf with minor details on palm and, image is not blurred. The image shown in figure 8(e) is of proposed method. Here every strip on the scarf is distinguishable and has been reconstructed well. The texture is retained in the reconstructed image without blur. Details on palm are visible and image is sharper.

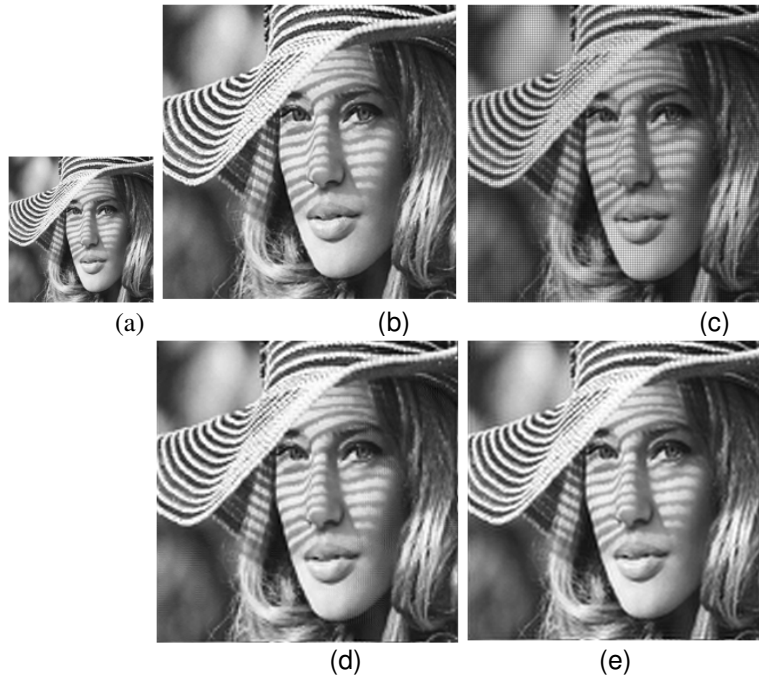


FIGURE 7: a) Low resolution image b) Original High resolution image c) Bicubic interpolated image d) Super-resolved image by using Contourlet transform e) Super-resolved image by using proposed method

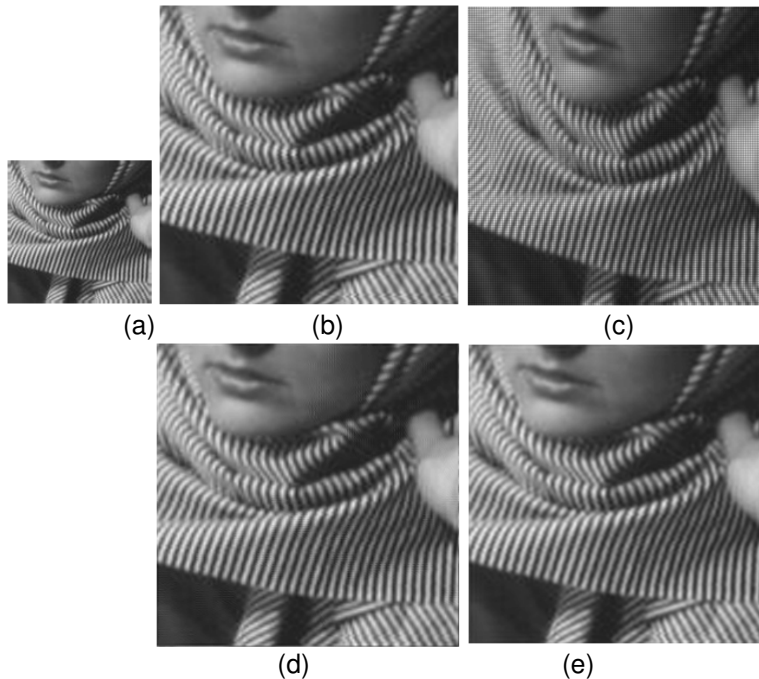


FIGURE 8: a) Low resolution image b) Original High resolution image c) Bicubic interpolated image d) Super-resolved image by using Contourlet transform e) Super-resolved image by using proposed method.

The reconstructed image is analyzed by computing mean squared error (MSE), peak signal-to-noise ratio (PSNR) in db and correlation coefficient (CC). Table 1 shows the comparison of the proposed method over bicubic interpolation and Contourlet learning algorithm for gray images.

Image	MSE			PSNR			Correlation		
	Bicubic	Contourlet Learning	Proposed Method	Bicubic	Contourlet Learning	Proposed Method	Bicubic	Contourlet learning	Proposed method
Old Man	0.0361	0.0026	0.0016	20.3618	31.8079	33.7574	0.8950	0.9815	0.9882
Girl	0.0472	0.0151	0.0137	18.2783	23.2210	23.6505	0.9333	0.9630	0.9675
Barbara	0.0443	0.0113	0.0095	19.9407	26.3643	27.1431	0.9283	0.9677	0.9729

TABLE 1: Comparison of MSE, PSNR in db and correlation coefficient.

Experiments on color images also have been conducted to recognize the performance of the proposed method. The training data set of high quality 100 color images is considered for the experimentation. The low resolution image is obtained from the same imaging model and it does not belong to data set. The input low resolution and all the high resolution images from data set are converted into $Y - C_b - C_r$ format. The Y plane which represents luminance is used to learn the curvelet coefficients of low resolution image. The other two planes C_b and C_r are interpolated with bicubic interpolation. The high resolution image is then reconstructed by combining these three planes.

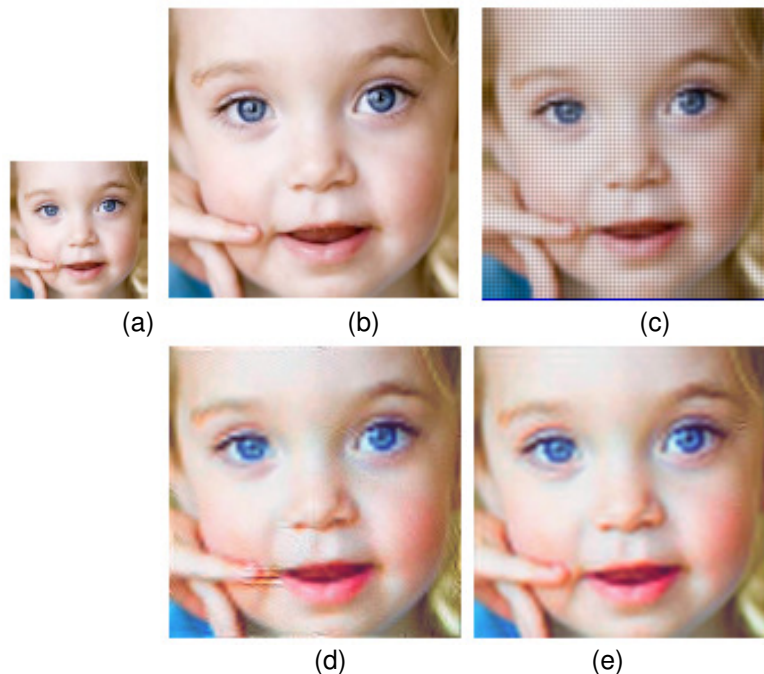


FIGURE 9: a) Low resolution image b) Original High resolution image c) Bicubic interpolated image d) Super-resolved image by using Contourlet transform e) Super-resolved image by using proposed technique

Figure 9 (a) shows the low resolution image of a child and figure 9(b) is its high resolution version. The image shown in figure 9(c) is bicubic interpolated image. The result shows blurred

image with less details. Figure shown in 9(d) is super-resolved image by using Contourlet transform. The reconstructed image shows more details such as eyes, eye balls and eye brows. The image shown in figure 9(e) is super-resolved by the proposed method. All the details of eye and eye brow have been preserved in the reconstructed image without blur. Figure 10 (a) shows the low resolution image of face and figure 10(b) is original high resolution image. The image shown in figure 10(c) is bicubic interpolated image. The result shows blurred image with few details. Image shown in figure 10(d) is super-resolved by using Contourlet transform. Almost all the details are visible and the image is sharp. The image shown in figure 10 (e) is super-resolved by proposed method. The result shows sharp image with all details such as eye balls, eye brow. The eyelids are also visible clearly.

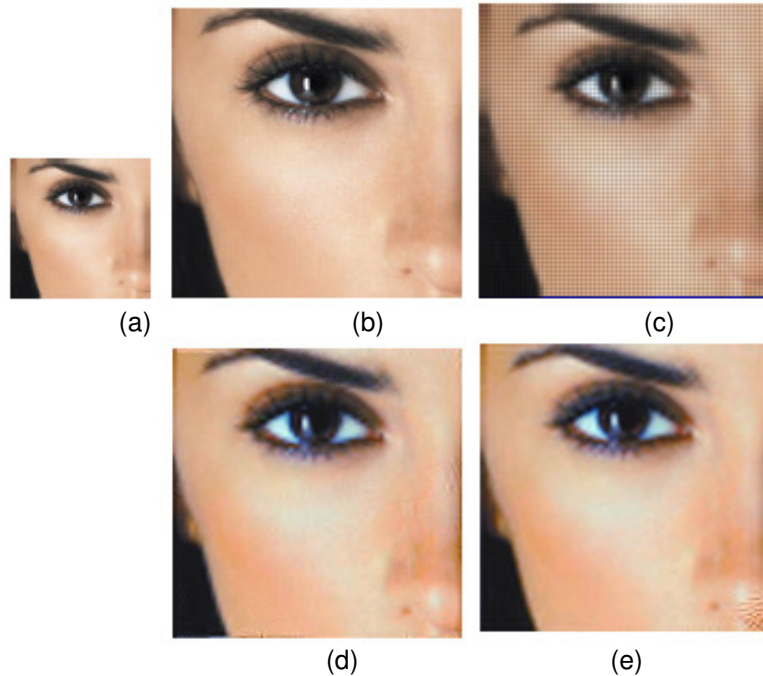


FIGURE 10: a) Low resolution image b) Original High resolution image c) Bicubic interpolated image d) Super-resolved image by using Contourlet transform e) Super-resolved image by using proposed technique

Table 2 shows the comparison of the proposed method over bicubic interpolation and Contourlet transform algorithm for color images in terms of MSE and PSNR in db. The mean squared error between the original image and reconstructed super-resolved image is expressed as:

$$MSE = \frac{\sum_{a,b} (S(a,b) - \hat{S}(a,b))^2}{\sum_{a,b} (S(a,b))^2} \quad (8)$$

Where $S(a,b)$ original high resolution image and $\hat{S}(a,b)$ is reconstructed super resolved image. The peak signal to noise ratio in db is defined as:

$$PSNR = 10 \log \left(\frac{255^2}{MSE} \right) \quad (9)$$

Correlation coefficient between original high-resolution image and reconstructed super-resolved image is computed as

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right)\left(\sum_m \sum_n (B_{mn} - \bar{B})^2\right)}} \quad (10)$$

where $\bar{A} = \text{mean}(A)$, and $\bar{B} = \text{mean}(B)$

Image	MSE			PSNR		
	Bicubic	Contourlet Learning	Proposed method	Bicubic	Contourlet Learning	Proposed Method
Child	0.0225	0.0053	0.0049	19.2070	25.5049	25.8248
Face	0.0213	0.0038	0.0030	20.0667	27.5502	28.5381

TABLE 2: Comparison of MSE and PSNR in db

For all the images there is appropriate improvement in the MSE, PSNR and correlation coefficient. The PSNR of the proposed method for the textured image has been improved almost by 6 db over bicubic interpolation and almost 1% improvement over the Contourlet learning method has been observed. An appropriate improvement in correlation coefficient is also observed with the proposed method.

6. CONCLUSION

The super-resolved image is reconstructed by estimating the finer scale FDCT coefficients of low-resolution test image by learning them from a set of high resolution training images. In this proposed method, FDCT captures the smoothness along the curves and high-resolution edges are learned from a training set at multiresolution level. This technique is useful when a single observation is available. The reconstructed result depends upon the high-resolution training images in dataset and the method is resolution dependent. Bicubic interpolation technique gives blurred reconstructed images whereas compatible results have been observed for Contourlet learning method. The appropriate improvement in MSE, PSNR and correlation coefficient is observed with the proposed method. Computational complexity is less than Contourlet learning method.

7. REFERENCES

- [1] R.Y.Tsai and T.S.Hung, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, vol.1,chapter 7,pp.317-339.JAI press,greenwich,Conn,USA,1984.
- [2] M.Irani and S. Peleg." Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3 pp. 231-239, 1991.
- [3] M. Irani and S. Peleg," Motion analysis for image enhancement: resolution, occlusion, and transparency," *Journal of Visual Communication and Image Representation*, vol. 4, no. 4, pp. 324-335,1993.
- [4] M. K. Ng, J. Koo, and N. K. Bose." Constraint total least-square computations for high-resolution image reconstruction with multisensors," *International Journal Of Imaging Systems and Technology*, vol. 12, no. 1, pp. 35-42, 2002.

- [5] M. K. Ng and N. K. Bose," Analysis of displacement errors in high-resolution image reconstruction with multisensors," IEEE Transactions on Circuits and Systems Part I, vol. 49, no. 6, pp. 806-813,2002.
- [6] N. Nguyen, P. Milanfar, and G. Golub," A computationally efficient super-resolution image reconstruction algorithm," IEEE transaction on Image Processing, vol. 10, no. 4, pp. 573-583, 2001
- [7] R.R. Schuitz and R. L. Stevenson," A Bayesian approach to image expansion for improved definition," IEEE Transactions on Image Processing, vol. 3, no. 3, pp. 233-242,1994
- [8] D. Rajan and S.Chaudhuri," An MRF-based approach to generation of super-resolution images from blurred observations," Journal of mathematical Imaging and Vision, vol. 16, no. 1, pp. 5-15, 2002.
- [9] D. Rajan and S. Chaudhuri," Simultaneous estimation of super-resolved scene and depth map from low resolution defocused observations," IEEE transctions on Pattern Analysis and Machine Intelligence, vol. 25,no. 9, pp. 1102-1117, 2003
- [10] M. Elad and A. Feuer," Restoration of a single super-resolution image from several blurred, noisy and under sampled measured images," IEEE transactions on Image Processing, vol. 6. no.12. pp. 1646-1658, 1997
- [11] S. Baker and T. Kanade," Limits on super-resolution and how to break them," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1167-1183, 2002
- [12] D. Capel and A. Zisserman," Super-resolution from multiple views using learnt image models," In Proceedings of IEEE Computer Society Conference on Computer Vision and pattern Recognition (CVPR'01), vol. 2, pp. II-627-II-634, Kauai, Hawaii, USA, December 2001
- [13] W. T. Freeman, T. R. Jones, and E. C. pasztor," Example-based super-resolution," IEEE Computer Graphics and Applications, vol. 22, no. 2, pp. 56-65, 2002
- [14] M.V. Joshi and S. Chaudhuri," Alearning-based method for image super-resolution from zoomed observations," In Proceedings of 5th International Conference on Advances In Pattern Recognition (ICAPR'03), pp. 179-182, Culcutta, India, December 2003
- [15] C.V. Jiji, M. V. Joshi, and S. Chaudhuri," Single –frame image super-resolution using learned wavelet coefficients," International Journal of Imaging Systems and Technology, vol. 14, no. 3, pp. 105-112,2004
- [16] H. Chang, D. Y. Yeung, and Y. Xiong," Super-resolution trough neighbor embedding," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), vol.1, pp. I-275-I-282, Washington, DC, USA, June-July 2004.
- [17] J. S. Park and S. W. Lee." Enhancing low-resolution facial images using error back-projection for human identification at a distance," in Proceedings of 17th IEEE International Conference on Pattern Recognition (ICPR'04), vol.1, pp. 346-349, Cambridge, UK, August 2004
- [18] J.Sun, N. N. Zheng, H. Tao, and H. Y. Shum," image hallucination with primal sketch priors," In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03), vol. 2, pp. II-729-II-736, Madison, Wis, USA, June 2003

- [19] M.N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2005
- [20] C.V.Jiji and S.Chaudhuri, "Single-frame image Super-resolution through Contourlet Learning," *EURASIP Journal on Applied Signal Processing*, vol.2006, Article ID 73767, pp.1-11,2006
- [21] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," *IEEE International Conference on Computer Vision (ICCV)*, pp. 349-356,2009
- [22] P. P. Gajjar and M. V. Joshi, "New learning based super-resolution use of DWT and IGMRF prior," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1201-1213, May 2010.
- [23] C. Kim, K. Choi, H. Lee, K. Hwang, and J. B. Ra, "Robust Learning-Based Super-resolution," in *Proceedings of International Conference on Image Processing (ICIP'10)*, pp. 2017-2020, Hong Kong, Sept. 2010.
- [24] K. I. Kim and Y. Kwon, "Single-image Super-Resolution Using Sparse regression and Natural Image prior," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1127-1133, 2010.
- [25] E. J. Candes and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise- C^2 singularities," *Comm. on Pure and Appl. Math.* Vol.57, pp. 219–266, 2004.
- [26] E.J.Candes, D.L.Donoho and L.Ying, "Fast Discrete Curvelet Transform.," *Journal of Multiscale modeling & simulation*, vol.5, no.3, pp.861—899, 2006.

Comprehensive Infrared Image Edge Detection Algorithm

Abhishek Gudipalli

*Assistant Professor/School of Electrical Engineering
VIT University
Vellore,632014, India*

abhishek.g@vit.ac.in

Dr.Ramashri Tirumala

*Associate Professor/Dept of Electronics & Communication Engineering
Sri Venkateswara University College of Engineering
Tirupati,517501,India*

rama.jaypee@gmail.com

Abstract

Edge detection is one of the most powerful image analysis tools for enhancing and detecting edges. Indeed, identifying and localizing edges are a low level task in a variety of applications such as 3-D reconstruction, shape recognition, image compression, enhancement, and restoration. This paper introduces a new algorithm for detecting edges based on color space models. In this RGB image is taken as an input image and transforming the RGB image to color models such as YUV, YCbCr and XYZ. The edges have been detected for each component in color models separately and compared with the original image of that particular model. In order to measure the quality assessment between images, SSIM (Structural Similarity Index Method) and VIF (Visual Information Fidelity) has been calculated. The results have shown that XYZ color model is having high SSIM value and VIF value. In the previous papers, edge detection based on RGB color model has low SSIM and VIF values. So by converting the images into different color models shows a significant improvement in detection of edges.

Keywords: Edge detection, Color models, SSIM, VIF.

1. INTRODUCTION

The important image processing tool that is used as a fundamental pre-processing step in many image processing applications is edge detection. Edges map have a significant role in application such as image categorization, image registration, feature extraction, and pattern recognition. An edge detector can be defined as a mathematical operator that responds to the spatial change and discontinuities in gray levels of pixels set in an image. Each industry will use its suitable color model. For example, the RGB color model is used in computer graphics, YUV is used in video systems, PhotoCD production and so on. Edge detection in color images requires several approaches of different complexity already exist. In image processing edge detection is an important process, color images provides more detailed edge information than gray value images, and color edge detection becomes vital for edge based image segmentation or edge-based stereo matching. Edges will not be detected in gray value images when neighboring objects have different hues but equal intensities. Additionally, the common shortcomings of the RGB image edge detection arithmetic are the low speed and the color losses after the each component of the image is processed. Thus, color edge detection is proposed based on changing the domain of the edge detection. The RGB image is transformed to YUV, YCbCr and XYZ color models. The quality assessment metrics such that SSIM and VIF have been applied to the above models and found that in XYZ color model is providing more detailed edge information than the other color models.

2. DIFFERENT COLOR SPACES

2.1 RGB color model

In the RGB model, each color appears as a combination of red, green, and blue. This model is called additive, and the colors are called primary colors. The primary colors can be added to produce the secondary colors of light (see Figure "Primary and Secondary Colors for RGB ") - magenta (red plus blue), cyan (green plus blue), and yellow (red plus green). The combination of red, green, and blue at full intensities makes white. The color subspace of interest is a cube shown in Figure "RGB Color Model" (RGB values are normalized to 0..1)[1], in which RGB values are at three corners; cyan, magenta, and yellow are the three other corners, black is at their origin; and white is at the corner farthest from the origin. The gray scale extends from black to white along the diagonal joining these two points. The colors are the points on or inside the cube, defined by vectors extending from the origin. Thus, images in the RGB color model consist of three independent image planes, one for each primary color. The importance of the RGB color model is that it relates very closely to the way that the human eye perceives color. RGB is a basic color model for computer graphics because color displays use red, green, and blue to create the desired color. Therefore, the choice of the RGB color space simplifies the architecture and design of the system. Besides, a system that is designed using the RGB color space can take advantage of a large number of existing software routines, because this color space has been around for a number of years. However, RGB is not very efficient when dealing with real-world images. To generate any color within the RGB color cube, all three RGB components need to be of equal pixel depth and display resolution. Also, any modification of the image requires modification of all three planes[1].

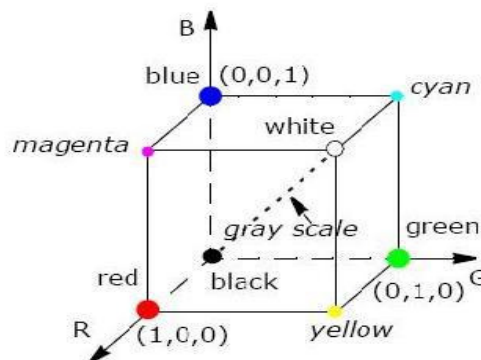


FIGURE 1: RGB Colour model.

2.2 YCbCr and YUV color model

In the YCbCr color space is used for component digital video is a scaled and offset version of the YUV color space. The YUV color model is the basic color model used in analogue color TV broadcasting. Initially YUV is the re-coding of RGB for transmission efficiency (minimizing bandwidth) and for downward compatibility with black-and white television. The YUV color space is "derived" from the RGB space. It comprises the *luminance* (Y) and two color difference (U, V) components. The luminance can be computed as a weighted sum of red, green and blue components; the color difference, or *chrominance*, components are formed by subtracting luminance from blue and from red. The principal advantage of the YUV model in image processing is decoupling of luminance and color information. The importance of this decoupling is that the luminance component of an image can be processed without affecting its color component. For example, the histogram equalization of the color image in the YUV format may be performed simply by applying histogram equalization to its Y component. There are many combinations of YUV values from nominal ranges that result in invalid RGB values, because the possible RGB colors occupy only part of the YUV space limited by these ranges. For example,

the histogram equalization of the color image in the YUV format may be performed simply by applying histogram equalization to its Y component. There are many combinations of YUV values from nominal ranges that result in invalid RGB values, because the possible RGB colors occupy only part of the YUV space limited by these ranges. Figure " YUV Color Model" shows the valid color block in the YUV space that corresponds to the RGB color cube RGB values are normalized to [0..1][1].

The Y'U'V' notation means that the components are derived from gamma-corrected R'G'B'. Weighted sum of these non-linear components forms a signal representative of luminance that is called *luma* Y'. (*Luma* is often loosely referred to as *luminance*, so you need to be careful to determine whether a particular author assigns a linear or non-linear interpretation to the term *luminance*)[1].The YCbCr color space is used for component digital video is a scaled and offset version of the YUV color space. The position of the block of RGB-representable colors in the YCbCr space is shown in Figure2.1 RGB Colors Cube in the YCbCr Color Model [1].

Conversion between RGB and YCbCr models:

$$Y' = 0.257 * R' + 0.504 * G' + 0.098 * B' + 16$$

$$Cb' = -0.148 * R' - 0.291 * G' + 0.439 * B' + 128$$

$$Cr' = 0.439 * R' - 0.368 * G' - 0.071 * B' + 128$$

Conversion between RGB and YUV models:

$$Y' = 0.299 * R' + 0.587 * G' + 0.114 * B'$$

$$U' = -0.147 * R' - 0.289 * G' + 0.436 * B' = 0.492 * (B' - Y')$$

$$V' = 0.615 * R' - 0.515 * G' - 0.100 * B' = 0.877 * (R' - Y')$$

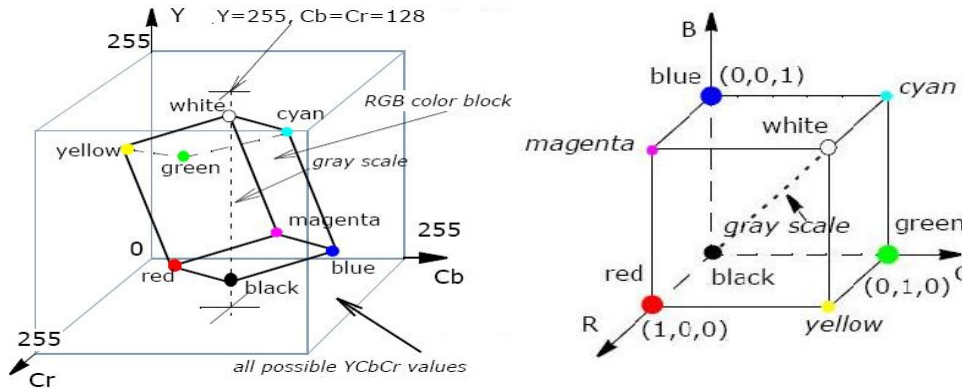


FIGURE 2: YCbCr and YUV color models.

2.3 XYZ color model

The XYZ color space is an international standard developed by the CIE (Commission Internationale de l'Eclairage). This model is based on three hypothetical primaries, XYZ, and all visible colors can be represented by using only positive values of X, Y, and Z. The CIE XYZ primaries are hypothetical because they do not correspond to any real light wavelengths. The Y primary is intentionally defined to match closely to luminance, while X and Z primaries give color information. The main advantage of the CIE XYZ space (and any color space based on it) is that this space is completely device-independent. The position of the block of RGB-representable colors in the XYZ space is shown in Figure 3 XYZ Color model [1].

Conversion between RGB and XYZ models:

$$X = 0.412453 * R' + 0.35758 * G' + 0.180423 * B'$$

$$Y = 0.212671 \cdot R' + 0.71516 \cdot G' + 0.072169 \cdot B'$$

$$Z = 0.019334 \cdot R' + 0.119193 \cdot G' + 0.950227 \cdot B'$$

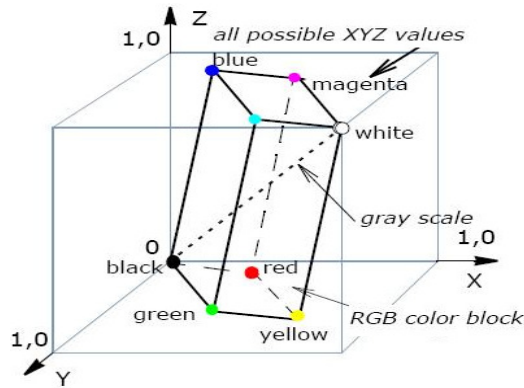


FIGURE 3: XYZ color model.

2.4 Infrared Imaging

Infrared (IR) light is electromagnetic radiation with a wavelength longer than that of visible light, measured from the nominal edge of visible red light at $0.74\mu\text{m}$, and extending conventionally to $300\mu\text{m}$. These wavelengths correspond to a frequency range of approximately 1 to 400 THz and include most of the thermal radiation emitted by objects near room temperature. Microscopically, IR light is typically emitted or absorbed by molecules when they change their rotational and vibration movements. Infrared imaging is used extensively for military and civilian purposes. Military applications include target acquisition, surveillance, night vision, homing and tracking. Non-military uses include thermal efficiency analysis, environmental monitoring, industrial facility inspections, remote temperature sensing, short-ranged wireless communication, spectroscopy, and weather forecasting[2]. Infrared astronomy uses sensor-equipped telescopes to penetrate dusty regions of space, such as molecular clouds; detect objects such as planets, and to view highly red-shifted objects from the early days of the universe. Humans at normal body temperature radiate chiefly at wavelengths around $12\mu\text{m}$, at the atomic level infrared energy elicits vibration modes in a molecule through a change in the dipole moment, making it a useful frequency range for study of these energy states for molecules of the proper symmetry. Infrared spectroscopy examines absorption and transmission of photons in the infrared energy range, based on their frequency and intensity [2].

3. EVALUATION METRICS

In image processing applications, the measurement of image quality plays main role. Image quality assessment algorithms are classified into three categories: FullReference (FR), Reduced-Reference (RR), and No-Reference (NR) algorithms. True No Reference algorithms are difficult to design and little progress has been made [3]. Full Reference algorithms are easier to design and The SSIM index is a full reference metric. In this, the measurement of image quality is based on reference image of perfect quality. SSIM is designed to improve Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE), which is proved to be inconsistent with human eye perception[4]. However, in RR or NR quality assessment, partial or no reference information is available. The SSIM index is defined as[4]:

$$SSIM(x,y) = \frac{\sigma_{xy} + C_1}{\sigma_x \sigma_y + C_1} \cdot \frac{2\mu_x \mu_y + C_2}{\mu_x^2 + \mu_y^2 + C_2} \cdot \frac{2\sigma_x \sigma_y + C_3}{\sigma_x^2 + \sigma_y^2 + C_3}$$

Let x and y be the two discrete non-negative signals extracted from the same spatial location from two images being compared, respectively μ_x, σ_x^2 and σ_{xy} be the mean of x , the variance of x and the covariance of x and y , respectively. μ_x and σ_x gives the information on luminance and contrast of x . σ_{xy} measures the structural similarity. where C_1, C_2 and C_3 are small constants given by $C_1 = (K_1 L)^2$; $C_2 = (K_2 L)^2$ and $C_3 = C_2 / 2$; respectively. L is the dynamic range of the pixel

values ($L = 255$ for 8 bits/pixel gray scale images), and $K_1 < 1$ and $K_2 < 1$ are two scalar constants [4].

Sheikh and Bovik (2006) developed a visual information fidelity (VIF) index for Full Reference measurement of quality of image. VIF is calculated between the reference image and its copy[5]. For ideal image, VIF is *exactly* unity. For distorted image types, VIF lies in between interval [0, 1]. Let $e=c+n$ be the reference image, and n zero-mean normal distribution $N(0, \sigma_n^2 I)$ noise. Also, let $f=d+n' = gc+v' + n'$ be the test image, where g represents the blur, v' the additive zero-mean Gaussian white noise with covariance $\sigma_v^2 I$, and n' the zero-mean normal distribution $N(0, \sigma_{n'}^2 I)$ noise[3]. Then, VIF can be computed as the ratio of the mutual information between c and f , and the mutual information between c and e for all wavelet subbands except the lowest approximation subband[4].

$$VIF = \frac{\Sigma I(c;f|z)}{\Sigma I(c;e|z)}$$

4. PROPOSED EDGE DETECTION ALGORITHM

The proposed algorithm can be explained in seven steps

- Step 1: The RGB image is sub divided into R, G and B layers of the image.
- Step 2: A 3X3 Laplacian mask is convolved with the R component of the image.
- Step 3: The edge detected R and the G, B layers of the image are concatenated to obtain edge detected image
- Step 4: SSIM and VIF values are calculated between the R edge detected image and RGB image.
- Step 5: Repeat steps 2 to 4 to calculate the SSIM and VIF values between G edge detected image and RGB image
- Step 6: Repeat steps 2 to 4 to calculate the SSIM and VIF values between B edge detected image and RGB image
- Step 7: The SSIM and VIF values of individual components are averaged.
- Step 8: R, G and B values of the image are transformed into its YCbCr, YUV and XYZ Intensity values using the conversion formulas.
- Step 9: Repeat steps 1 to 8 to calculate SSIM and VIF values for YCbCr, YUV and XYZ images.

5. EXPERIMENTAL RESULTS

The Proposed algorithm has been applied to RGB, YCbCr, YUV and XYZ images and SSIM & VIF values are computed for a set of edge detected images and dataset is formed and tabulated in Table 1. The property of infrared images is that intensity value depends on temperature, object surface, surface direction, wavelength, etc. based up on these, for RGB color model, The SSIM value range is around 0.57 and VIF value range is around 0.21. For YCbCr color model, The SSIM value range is around 0.58 and VIF value range is around 0.13. For YUV color model, The SSIM value range is around 0.49 and VIF value range is around 0.16. For XYZ color model, The SSIM value range is around 0.61 and VIF value range is around 0.33. from the dataset, XYZ model shows better quality of edge detection than the other color models. The original and edge detected RGB, YCbCr, YUV and XYZ images are shown below.

<i>Color model</i>	<i>SSIM</i>	<i>VIF</i>
RGB	0.5730	0.2147
XYZ	0.6156	0.3306
YCbCr	0.5790	0.1316
YUV	0.4907	0.1675

TABLE 1: SSIM and VIF Values for RGB, XYZ, YCbCr and YUV images.

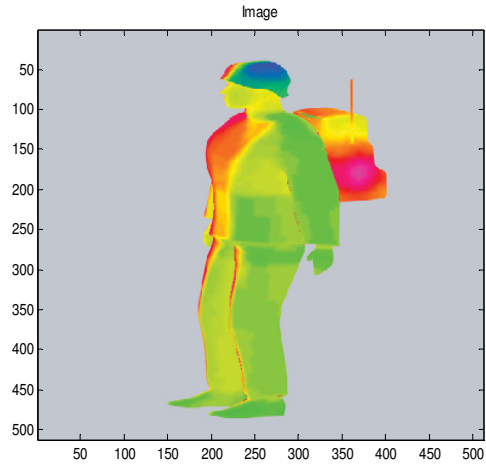


FIGURE 4: RGB Image

RGB image after Red component edge detection



FIGURE 5: RGB edge detected image

XYZ image

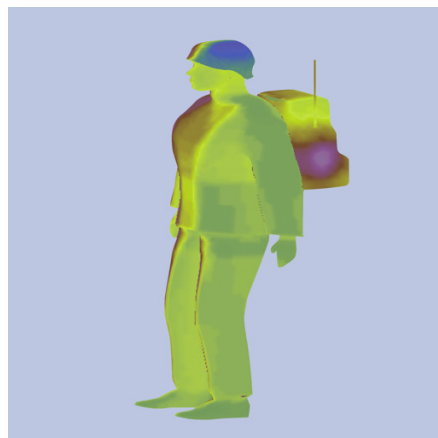


FIGURE 6: XYZ image

XYZ image after X component edge detection



FIGURE 7: XYZ edge detected image

YCbCr image



FIGURE 8: YCbCr image

YCbCr image after Y component edge detection



FIGURE 9: YCbCr edge detected image

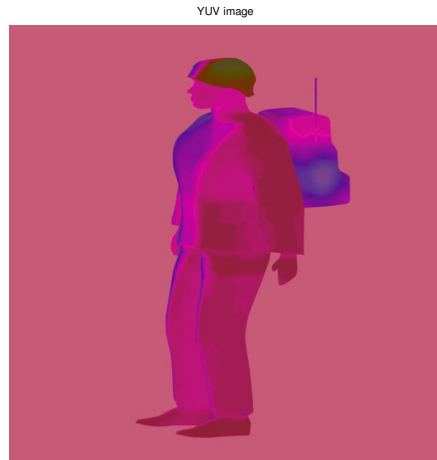


FIGURE 10: YUV image

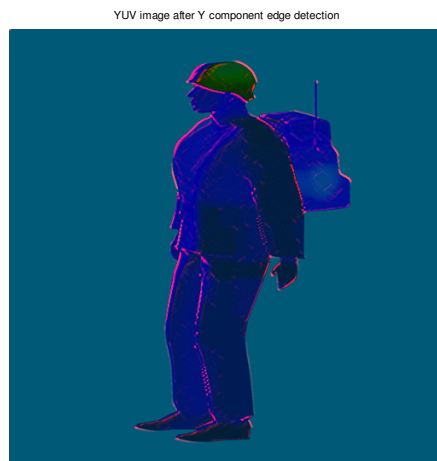


FIGURE 11: YUV edge detected image

6. CONCLUSION

The proposed approach has a potential for various applications to detect edges of Infrared images used extensively for military and civilian purposes. Humans at normal body temperature radiate chiefly at wavelengths around $12 \mu\text{m}$ a comprehensive method can be developed for pattern recognition models based on edge detection algorithms. The method shown is a new approach to detect edges in different color space models. The algorithm was developed based on RGB colour space and the significant features extracted by converting it into YUV, YCbCr and XYZ models. Among these XYZ model shows better quality of edge detection.

7. REFERENCES

1. Image Color Conversion, "http://software.intel.com/sites/products/documentation/hpc/ipp/ippi/ippi_ch6/ch6_Intro.html".
2. Dong Wang and Jingzhou Zhang "Infrared image edge detection algorithm based on sobel and ant colony algorithm", *2011 International conference on multimedia technology*, pp.4944 – 4947.

3. Sheikh, H. R., Bovik, A. C. and Cormack, L., "Noreference quality assessment using natural scene statistics: JPEG2000", *IEEE Transactions on Image Processing*, 14(11), 2005, pp. 1918-1927.
4. S Qian, G Chen "Four reduced-reference metrics for measuring hyperspectral images after spatial resolution enhancement" In: Wagner W., Székely, B. (eds.): *ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS, Vol. XXXVIII, Part 7A*
5. Sheikh, H. R. and Bovik, A. C., "Image information and visual quality", *IEEE Transactions on Image Processing*, 15(2), 2006, pp. 430-444.
6. Chen, G. Y. and Qian, S. E., "Evaluation and comparison of dimensionality reduction methods and band selection", *Canadian Journal of Remote Sensing*, 34(1), 2008a, pp. 26-32.
7. Li, Q. and Wang, Z., "Reduced-reference image quality assessment using divisive normalization-based image representation", *IEEE Journal of Selected Topics in Signal Processing, Special issue on Visual Media Quality Assessment*, 3(2), 2009, pp. 202-211.
8. S K Naik , C A Murthy, "Standardization of edge magnitude in color images", *IEEE Transactions on Image Processing* , 2006 , 15(9) : 2588 —2595.
9. C L Novak, S A Shafer "Color edge detection", *Proc of DARPA Image Understanding Workshop*. 1987, pp:35 -37
10. R C Gonzalez, R E Woods, "Digital Image Processing" (Second Edition). New York : Prentice Hall , 2003 pp.300-450.
11. P E Trahanias , A N Venetsanopoulos, "Color Edge Detection Using Vector Order Statistics", *IEEE Transactions on Image Processing* , 1993 , 2 (2) : 259 -264.
12. J Fan , W GAref , M Hacid , et al. "An improved automatic isotropic color edge detection technique", *Pattern Recognition Letters*, 2001, 22(3):1419-1429
13. A. Koschan, "A comparative study on color edge detection," in Proceedings of the 2nd Asian Conference on Computer Vision ACCV'95, pp. 574-578, Singapore, 1995.
14. A. N. Evans and X. U. Liu, "A morphological gradient approach to color edge detection," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp.1454-1463, 2006.
15. X. W. Li and X. R. Zhang, "A perceptual color edge detection algorithm," *International Conference on Computer Science and Software Engineering*, vol. 1, pp.297-300, 2008.
16. Yajun Fang, Keiichi Yamada, Yoshiki Ninomiya, Berthold Horn, Ichiro Masaki "Comparison between infrared image based and visible image based approaches for pedestrian detection", *IEEE Intelligent Vehicles Symposium*, pp.505-510, 2003.
17. Kenji Omasa, Fumiki Hosoi and Atsumi Konishi, "3D lidar imaging for detecting and understanding Plant responses and canopy structure", *Journal of Experimental Botany*, Vol. 58, No. 4, pp. 881–898, 2007.

Image Restoration Using Particle Filters By Improving The Scale Of Texture WithMRF

Dr. Anna Saro Vijendran

Director &Head ,
Department of Computer Applications,
S.N.R Sons College, Coimbatore

saroviji@rediffmail.com

Bobby Lukose

Assistant Professor
Department of Computer Science,
Hindusthan College of arts & Science College, Coimbatore

bobbylukose74@gmail.com

Abstract

Traditional techniques are based on restoring image values based on local smoothness constraints within fixed bandwidth windows where image structure is not considered. A common problem for such methods is how to choose the most appropriate bandwidth and the most suitable set of neighboring pixels to guide the reconstruction process. The present work proposes a denoising technique based on particle filtering using MRF (Markov Random Field). It is an automatic technique to capture the scale of the texture. The contribution of our method is the selection of an appropriate window in the image domain. For this we first construct a set containing all occurrences then the conditional pdf can be estimated with a histogram of all center pixel values. Our method explores multiple neighbors' sets that can be used for pixel denoising, through a particle filtering approach. This technique associates weights for each hypothesis according to its relevance and its contribution in the denoising process.

Keywords:Additive Guassian Noise, MC, MRF, Transition, Particle Filters

1. INTRODUCTION

A digital image could get corrupted easily due to various types of noise during transmission and acquisition. A noise is any unwanted signal/pixel that may be added or subtracted during transmission. These unwanted signals/pixels decrease the image quality. The sources of noise in digital images arise during image acquisition and/or transmission with unavoidable shot noise of an ideal photon detector.

1.1 Additive and Multiplicative Noises

Noise is undesired information that contaminates the image. Typical images are corrupted with noise modeled with either a Gaussian, uniform, or salt and pepper distribution. Another typical noise is a speckle noise, which is multiplicative in nature.

An additive noise follows the rule $w(x, y) = s(x, y) + n(x, y)$, while the multiplicative noise satisfies $w(x, y) = s(x, y) \times n(x, y)$, where $s(x, y)$ is the original signal, and $n(x, y)$ denotes the noise introduced into the signal to produce the corrupted image $w(x, y)$, and (x, y) represents the pixel location.

Image noise is the random variation of brightness or color information in images produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is generally regarded as an undesirable by-product of image capture. Although these unwanted fluctuations became known as "noise" by analogy with unwanted sound they are inaudible and such as dithering. The types of Noise are following :

- Amplifier noise (Gaussian noise)
- Salt-and-pepper noise
- Shot noise (Poisson noise)
- Speckle noise

Amplifier noise / Gaussian Noise

Noise is a statistical noise that has its PDF (probability density function) equal to that of the normal distribution, which is also known as the Gaussian distribution. Gaussian noise is most commonly known as additive white Gaussian noise. The term "white Gaussian noise" could be precise. Gaussian noise is evenly distributed over the signal. This means that each pixel in the noisy image is the sum of the true pixel value and a random Gaussian distributed noise value [16]. The standard model of amplifier noise is additive, Gaussian, independently at each pixel and independent of the signal intensity. Amplifier noise is a major part of the "read noise" of an image sensor, that is, of the constant noise level in dark areas of the image.

Salt-and-pepper noise

An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. This type of noise can be caused by dead pixels, analog-to-digital converter errors, bit errors in transmission, etc. This can be eliminated in large part by using dark frame subtraction and by interpolating around dark/bright pixels.

Poisson noise

Poisson noise or shot noise is a type of electronic noise that occurs when the finite number of particles that carry energy, such as electrons in an electronic circuit or photons in an optical device, is small enough to give rise to detectable statistical fluctuations in a measurement.

Speckle noise

Speckle noise is a granular noise that inherently exists in and degrades the quality of the active radar and synthetic aperture radar (SAR) images. Speckle noise in conventional radar results from random fluctuations in the return signal from an object that is no bigger than a single image-processing element. It increases the mean gray level of a local area. Speckle noise in SAR is generally more serious, causing difficulties for image interpretation. It is caused by coherent processing of backscattered signals from multiple distributed targets. In SAR oceanography, for example, speckle noise is caused by signals from elementary scatters, the gravity-capillary ripples, and manifests as a pedestal image, beneath the image of the sea waves.

Image de-noising is a common procedure in digital image processing aiming at the suppression of different types of noises without losing much detail contained in an image. This procedure is traditionally performed in the spatial-domain or transform-domain by filtering. To reduce the noise from images, various images de-noising filters are used. Image denoising still remains a challenge for researchers because noise removal introduces artifacts, blurring of the images, and the noise remaining in the image edges.

A traditional way to remove noise from image data is spatial filters. Spatial filters are a low pass filter. It can be further classified into non-linear and linear filters. Linear filters, which consist of convolving the image with a constant matrix to obtain a linear combination of neighborhood values, have been widely used for noise elimination in the presence of additive noise. Linear filters destroy lines and other fine image details, also it produce a blurred and smoothed image with poor feature localization and incomplete noise suppression. Variety of nonlinear median type filters such as weighted median, rank conditioned rank selection has been developed to overcome this drawback.

The transform domain filtering methods can be subdivided as data adaptive and non-adaptive. Nonadaptive transforms are discussed first since they are more popular. The conventional Fast Fourier Transform (FFT) based image denoising method is a low pass filtering technique in which edge is not as sharp. The edge information is spread across frequencies because of the FFT basis functions, which are not being localized in time or space.

Wavelet Analysis, a new form of signal analysis is more efficient than Fourier analysis. Wavelet transforms are classified into discrete wavelet transforms (DWTs) and continuous wavelet transforms (CWTs). Both DWT and CWT are continuous-time (analog) transforms. But, the localized nature of the wavelet transforms both in time and space results in denoising with edge preservation. Wavelet Transform is the nonlinear coefficient thresholding based methods. It enables the separation of signal from noise. The procedure in which small coefficients are removed while others are left untouched is called Hard Thresholding. But the method generates spurious blips, better known as artifacts. To overcome the demerits of hard thresholding, wavelet transform using soft thresholding. In wavelet based denoising methods, the noise is estimated and wavelet coefficients are threshold to separate

signal and noise. Denoising of images using VisuShrink, SureShrink and Bayes Shrink, all these methods are based on the application of wavelet transforms. Bayes shrinks wavelet denoising has been widely used for image denoising. The focus was shifted from the Spatial and Fourier domain to the Wavelet transform domain.

Edge-Preserving Smoothing Filters (Neighborhood filters) are Bilateral filter, sigma filter, mean sigma filter. They are used to solve the HALO (artifacts) and noise. The extensions of Adaptive range and domain filters is Bilateral Filter which performs weighted averaging in both range and domain [28]. It smooths noisy images while preserving edges using neighboring pixels. Bilateral filtering is a local, nonlinear, and a non-iterative technique which considers both gray level and color similarities and geometric closeness of the neighboring pixels.

The NL mean filter achieves the best results in terms of small detail preserving since noise contains less image information. The disadvantage is slow in terms of computation time. The NL-mean algorithm assumes fixed size with respect to the local filtering window (window centered at the origin pixel). Then based on the similarity between the center patch and the candidate patches it performs filtering [5] [19].

The sigma filter identifies impulse noise from noisy gray scale images by utilizing the standard deviation measure [20]. And the mean shift filter does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. The main advantage of this method is computational efficiency, but it is constrained by the amount of information present at the considered window.

In the particle filter, the posterior probability density is approximated as a set of particles [1]. When the particles are properly placed, weighted and propagated, posteriors can be estimated sequentially over time. The density of particles represents the probability of posterior function. In this set of candidate pixels not fixed and changes per pixel location according to local pixel properties. The disadvantage is even with a large number of particles, there are no particles in the vicinity of the correct state. This is called the particle deprivation problem. In this paper presented an efficient particle filtering algorithm using MRF which overcomes the above disadvantages and to remove low to high density noise for several standard images. This denoising algorithm should be able to extract the most important correlations of local structure of the entire image domain. Gaussian kernels are the most common selection of such an approach. Sequential Monte Carlo is a well known technique evolving densities to the different hypotheses.

The remainder of this document is organized as follows: In section II discuss about the MRF of image structure learning. Application of particle filter image denoising is presented in section 3, experimental results and comparisons are presented in section 4, and tables and figures are presented in section 5. Finally, we conclude in section 6.

2. PRESENT WORK

Capturing the geometric structure of the image is the important process in image restoration. Such a process involves two steps, (i) a learning stage where the image structure is modeled, and (ii) a reconstruction step. Our aim is to introduce a strategy that allows a best possible selection of the pixels contributing to the reconstruction process driven by the observed image geometry. Using this to retrieve similar pixels. The issues are (i) the selection of the trajectory, (ii) and the evaluation of the trajectory appropriateness. Each walk is composed of a number of possible neighboring sites/pixels in the image which are determined according to the observed image structure. To overcome the issues optimizing the selection of candidate pixels within a walk as well as the overall performance of the method image structure at local scale is considered as a learning stage. It computes a probability density function that describes the spatial relation between similar image patches in a local scale. Here to improve the scale of texture by using MRF.

2.1 Automatic capturing of scale of texture using Markov Random field (MRF)

MRF models have been used in image restoration, region segmentation, and texture synthesis. In image processing, texture may be defined in terms of spatial interactions between pixel gray levels within a digital image. The aim of texture analysis is to capture the visual characteristics of a texture by mathematically modeling these spatial interactions. Markov Random Fields (MRFs) are widely used probabilistic models for regularization. The probability density function (pdf) defined by the MRF is the normalization constant. Maximum Likelihood (ML), probably the most common and popular

method of probabilistic parameter estimation, require the pdf to be normalized. Sampling techniques, such as Markov chain Monte Carlo (MCMC) used in this model [11] [12]. Our aim is to preserve the local structure of the texture as much as possible. To achieve this to define a strategy to generate neighborhood candidate windows that takes into account the image content. After that to determine the most appropriate window for estimating the image intensity in a given position.

2.2 Steps for synthesizing one pixel

1. Let I be an image that is being synthesized from a texture sample image $I_{smp} \subset I_{real}$ where I_{real} is the real infinite texture.
2. Let $p \in I$ be a pixel and let $\omega(p) \subset I$ be a square image patch of width ω centered at p .
3. Let $d(\omega_1, \omega_2)$ denotes some perceptual distance between two patches.
4. Assume that all pixels in I except for pair known.
5. To synthesize the value of p we first construct an approximation to the conditional probability distribution $P(p|\omega(p))$ and then sample from it.
7. Based on our MRF model we assume that p is independent of $I \setminus \omega(p)$ given $\omega(p)$.
8. If we define a set $\Omega(p) = \{\omega \subset I_{real} : d(\omega, \omega(p)) = 0\}$
9. containing all occurrences of $\omega(p)$ in I_{real} , then the conditional pdf of p can be estimated with a
10. histogram of all center pixel values in $\Omega(p)$.
11. Then a variation of the nearest neighbor technique is used for finding the closest match.
12. If the closest match

$$\omega_{best} = \underset{\omega \subset I_{smp}}{\operatorname{argmin}} d(\omega(p), \omega) \tag{1}$$

is found, and all image patches ω with $D(\omega(p), \omega) < (1 + \epsilon)d(\omega(p), \omega_{best})$ is included in $\Omega'(p)$, where $\epsilon = 0.1$.

13. The center pixel values of patches in $\Omega'(p)$ give us a histogram for p , which can then be sampled, either uniformly or weighted by d .
14. Then to find a suitable distance d by using normalized sum of squared differences metric d_{SSD} .
15. This metric gives the same weight to any mismatched pixel, whether near the center or at the edge of the window.
16. To generate neighborhood candidate windows that take into account the image content.
17. Finally to determine the most appropriate window to estimate the image intensity in a given position.

3. PARTICLE FILTER IMAGE DENOISING FRAMEWORK

The probability density function (pdf) that aims to find a spatial representation of different structures through the computation of the relative position of similar patches. Given this density, our aim is to determine the most appropriate set of neighbors to estimate the noise-free intensity of a given pixel. This is done through particle filtering technique. The particle filter is a special version of the Bayes filter based on Monte Carlo sampling [7]. The particle filter algorithm consists of three steps: sampling, calculation of the importance weight and resampling. In the sampling step, samples are generated according to pdf. In the step of importance weighting, the importance weight is computed for each particle. In the resampling step, the particles with different weights are sampled again with replacement according to their weights and the particles with different weights are replaced by the new particles with equal weights. The particles with larger weights are more likely to be selected than the particles with smaller weights.

3.1 Bayesian Tracking

Filtering is to determine an estimation of the state vector. In the Bayesian framework to compute the pdf of a system, based on observations [18]. The recursive computation of the prior and the posterior pdf leads to the exact computation of the posterior density. Particle filters, which are sequential Monte-Carlo techniques, estimate the Bayesian posterior probability density function (pdf) with a set of samples. The posterior pdf is computed using the equation

$$P(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{P(z_k | z_{1:k-1})} \tag{2}$$

Where x_k is the state vector and $(z_{1:k})$ is set of all the available observations. Similarly, the prior pdf is also computed using the equation

$$p(x_k | z_{1:k}) = \int p(x_k | x_{1:k-1}) p(x_{k-1} | z_{1:k-1}) dx_{k-1} \tag{3}$$

3.2 Sampling Importance Resampling Filter

The sequential Importance Sampling (SIS) algorithm is a Monte Carlo (MC) method that forms the basis for most sequential MC filters developed over the past decades. The key idea is to represent the required posterior density function by a set of random samples with associated weights and to compute estimates based on these samples and weights. The posterior pdf can be approximated by a discrete weighted sum

$$P(x_k | z_{1:k}) \approx \sum_{m=1}^{N_p} \omega_k^m \delta(x_k - x_k^m) \text{ and } \sum_{m=1}^{N_p} \omega_k^m = 1 \tag{4}$$

where δ is the Kronecker delta function and N_p is the random state samples. The samples are generated through the principles of Importance Sampling. Then compute the particle weights iteratively according to

$$\omega_k^m \propto \omega_{k-1}^m p(z_k | x_k^m) \tag{5}$$

A common problem is samples degeneration, where many particles carry less information and the variance in the weight increases. So that many particles have their weight close to zero. To overcome this problem a resampling step is necessary. The resampling scheme used to eliminate particles with smaller weights.

3.3 Transition Model

The transition according to the probability density function means that a maximum number of particles explore sites that are similar to the original pixel [26].

3.4 Likelihood measure

Measuring similarities between image patches is an important thing. In the case of denoising, filtering approach should consider pixels with the exact same value. Parallel to that, each particle corresponds to a random walk where a certain number of pixels have been selected and contribute a new element to the denoising process. The random walk is the set of the sites contributing to the reconstruction of a given pixel is determined [2]. This measure evaluates the new candidate pixel position using the formula

$$D_{sk} = \frac{\sum_{v \in [-w, w]^2} |I(x_0 + v) - I(x_k + v)|^2}{(2w+1)^2} \tag{6}$$

where D_{sk} is the similarity measure, x_0 is the original pixel and x_k is the current particle position and w is the bandwidth. The observation of the walk variance with respect to the origin value

$$D_{vk} = \frac{1}{k} \sum_{p=0}^{k-1} (I(x_p) - I(x_0))^2 \tag{7}$$

The patch around the original pixel is defined as an exponential function of the two metrics.

$$\omega_k = e^{-\left(D_{sk}/2\sigma_g^2 + D_{vk}/2\sigma_v^2 \right)} \tag{8}$$

where σ_g and σ_v are constants that determine the bandwidth of the weight computation function.

3.5 Intensity Reconstruction

For each pixel of the image, generate particles by applying perturbations starting from the initial position according to the transition law. Repeating the process for each particle number of times. To reconstruct the original pixel according to the walk of the particle using the formula

$$k$$

$$\hat{U}_k^m(x) = \left(\frac{1}{N_p}\right) \sum_{p=0}^{N_p} I(x^m_p) \tag{9}$$

where m is the particle. And the reconstructed value is a weighted average of the mean intensity of each walk which is defined as

$$\hat{U}_k(x) = \sum_{m=0} \omega^m_k \hat{U}_k^m(x) \tag{10}$$

The whole process such as transition, weight computation and resampling is repeated a number of times.

3.6 Denoising Algorithm

1. Read the Gaussian noisy image
2. Split the image into different candidate windows
3. Calculate pdf for each windows
4. calculate the mean value for each window
5. choose minimum mean value window using the equation (1)
6. apply particle filter to this window
7. For each pixel repeat the steps 8 to 14
8. Generate particle according to the pdf
9. Compute intensity for each walk using equation (9)
10. Compute the weight of each particle using equation (8)
11. Normalize the weight of the particle
12. Compute the estimated intensity
13. Perform resampling
14. Select the best pixel position
15. Final pixel intensity estimation = weighted mean of all filtered values of different random walks using equation (10).

4. EXPERIMENTAL RESULTS

The proposed algorithm is tested using 256 X 256 8-bits/pixel standardgray scaleimages. There are 20 imagetaken from the Berkely Segmentation Dataset & Benchmark database. The performance of the proposed algorithm is tested with different noise levels.Each time the test image is corrupted by different additive white Gaussian noise standard deviation ranging from 10 to 50 with an increment of 10. These noisy images are denoised by two algorithms and the performedifference between the particle filter and the proposedapproach measured by the parameters PSNR and MSE. All the filters are implemented inMatlab 10.

A quantitative measure of comparison likePeak signal to noise ratio (PSNR), mean square error (MSE) is used in this work.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \tag{12}$$

$$MSE = \frac{1}{\|\Omega\|} \sum_{x \in \Omega} (U(x) - \hat{U}(x))^2 \tag{13}$$

4.1 Algorithm for Peak Signal to Noise ratio (PSNR)

Step1: Difference of noisy images and noiseless image iscalculated.

Step2: Size of the matrix obtains in step 1 is calculated.

Step3: Each of the pixels in the matrix obtained in step2 issquared.

Step4: Sum of all the pixels in the matrix obtained instep3 is calculated.

Step5: (MSE) is obtained by taking the ratio of valueobtained in step 4 to the value obtained in the Step2.

Step6: (RMSE) is calculated by taking the square rootof thevalue obtained in Step5.

Step7: Dividing 255 with RMSE, taking log base 10 andmultiplying by 20 gives the value of PSNR.

4.2 Algorithm for Root Mean Square Error (RMSE)

- Step1: Difference of noisy images and noiseless image is calculated.
- Step2: Size of the matrix obtains in Step1 is calculated.
- Step3: Each of the pixels in the matrix obtained in step 2 is squared.
- Step4: Sum of all the pixels in the matrix obtained in step 3 is *calculated*.
- Step5: (MSE) is obtained by taking the ratio of value obtained in Step4 to the value obtained in the step 2.
- Step6: (RMSE) is calculated by taking the square rootof the *value* obtained in Step5.

In Table 1, provide a PSNR value of restored images for the particle filter and the MRF particle filter. As seen the results of Table I the MRF particle Filter method produces very good results. The PSNR values for particle filter and the MRF particle filter for 20 images at different Gaussian levels are displayed in Fig. 1.(a) and (b) respectively. We can make several observations from these plots. In Fig. 2. (a) and (b) the PSNR values for particle filter and MRF particle filter for 20 images at Gaussian noise $\sigma=20$ and MSE values for particle filter and MRF particle filter for 20 images at Gaussian noise $\sigma=20$ are displayed respectively. The visual quality results are presented in Fig. 3. Noise free image, Gaussian noise image, restored image using particle filter, restored image using MRF particle filters for 20 images at Gaussian noise $\sigma=20$ as shown in Fig. 3.(a), (b), (c) and (d) respectively. In all graphs the x-axis values are represented as 1,2,3 etc. which denotes 1 for baboon, 2 for Barbara respectively that are presented in Table 1. The visual quality and quantitative results clearly show the MRF particle filter perform much better thana particle filter in terms of PSNR and MSE.

4.3 Tables And Figures

Images	Particle filter					Particle filter using MRF				
	$\sigma =10$	$\sigma =20$	$\sigma =30$	$\sigma =40$	$\sigma =50$	$\sigma =10$	$\sigma =20$	$\sigma =30$	$\sigma =40$	$\sigma =50$
Baboon	33.72	39.16	42.44	44.79	46.58	37.17	42.76	46.06	48.42	49.98
Barbara	34.99	40.34	43.43	45.76	47.47	38.38	43.82	47.05	49.25	51.00
Boathouse	36.23	41.39	45.10	46.65	49.05	39.52	45.55	47.70	50.50	51.67
Bridge	36.49	41.75	44.97	47.19	48.97	39.82	45.31	48.39	50.76	52.09
Building	39.90	44.56	47.49	49.49	50.99	42.86	47.79	50.54	52.29	53.80
Cameraman	38.46	43.36	46.38	48.53	50.11	41.57	46.72	49.64	51.69	53.18
Capsicum	37.08	41.99	45.12	47.20	49.00	40.17	45.45	48.40	50.43	51.54
Flyover	36.64	41.71	44.87	47.11	48.88	39.84	45.19	48.33	50.59	51.85
Girl	37.99	42.71	45.74	47.80	49.47	40.95	46.07	48.92	50.88	52.14
Helicopter	42.77	46.87	49.48	51.29	52.84	45.35	49.73	52.36	54.50	56.46
Hills	38.17	43.05	46.08	48.25	49.95	41.24	46.38	49.45	51.61	52.92
Lena	38.49	42.90	45.74	47.84	49.40	41.24	46.02	48.95	50.91	52.33
Mechanic	37.14	42.16	45.30	47.43	49.20	40.32	45.63	48.61	50.86	52.06
Monkey	39.71	44.28	47.21	49.14	50.63	42.58	47.50	50.17	52.08	53.82
Nature	38.17	43.15	46.30	48.56	50.36	41.28	46.62	49.78	52.19	53.37
OldWomen	36.13	41.38	44.59	46.85	48.52	39.47	44.90	48.00	50.22	51.52
Owl	34.08	39.57	42.88	45.24	47.01	37.57	43.22	46.44	48.80	50.07
Pelican	38.11	43.27	46.44	48.72	50.46	41.37	46.75	49.93	52.24	53.52
Starfish	37.12	41.96	45.00	47.20	48.85	40.16	45.30	48.36	50.42	51.67
Water plant	40.51	44.75	47.47	49.46	50.89	43.18	47.74	50.46	52.34	53.85

TABLE 1: PSNR Values For The Denoised Images At Different Guassian Noise Levels

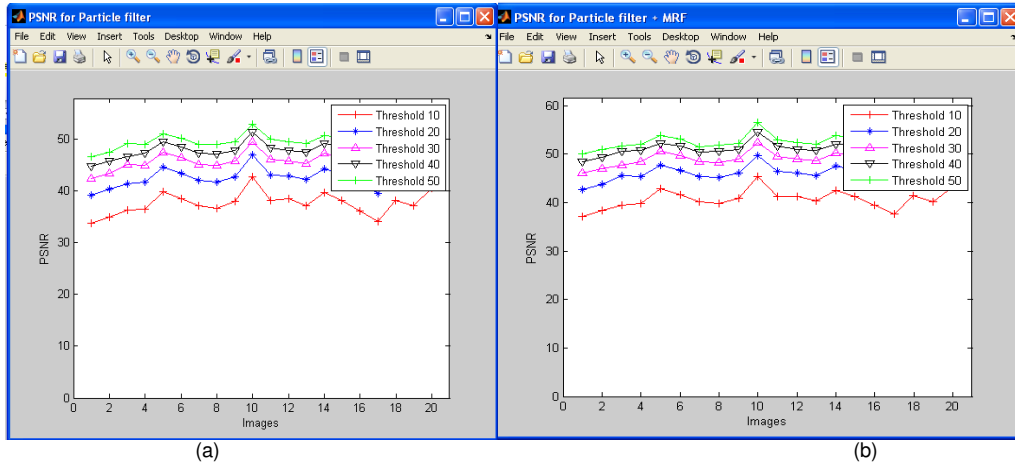


FIGURE 1:(a) PSNR of particle filter for 20 images at different Gaussian noise levels
 (b) PSNR of MRF particle filter for 20 images at different Gaussian noise levels

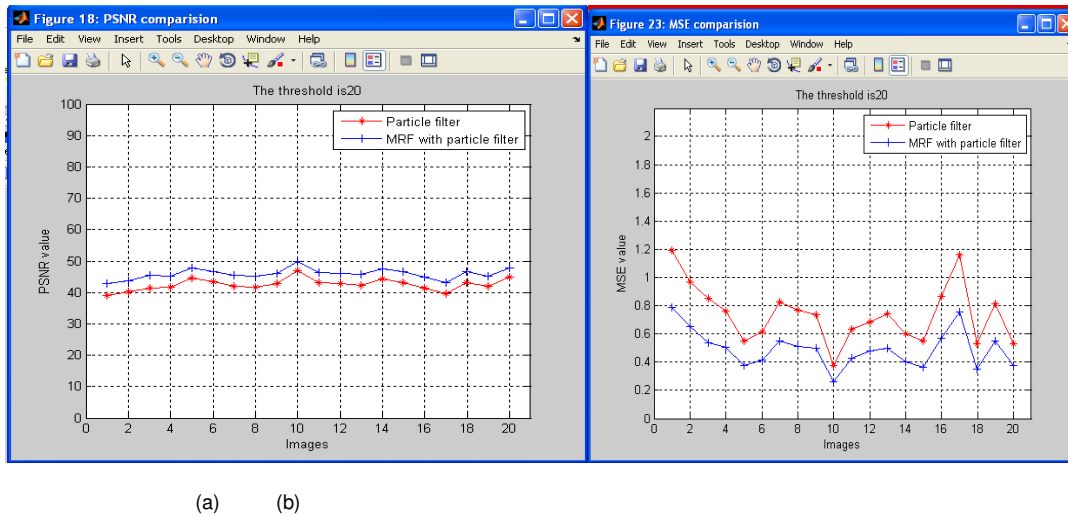
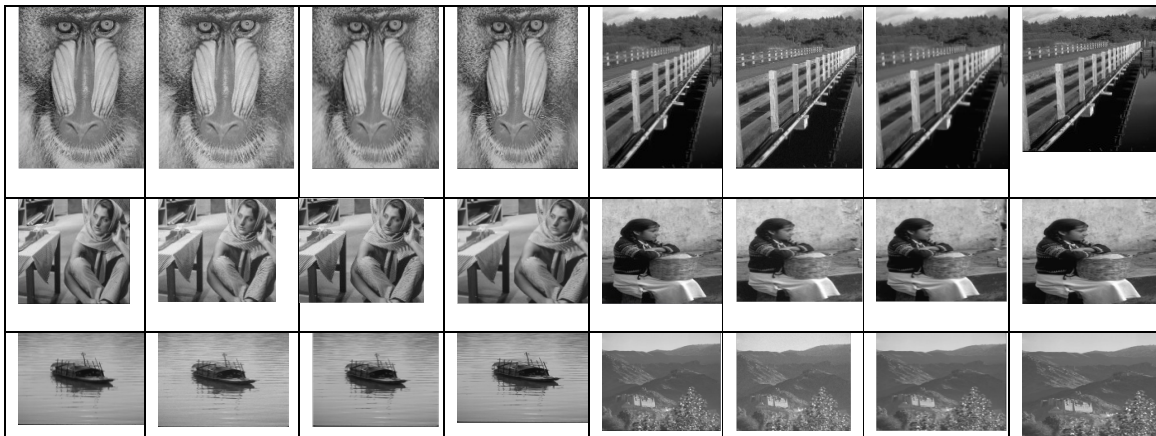
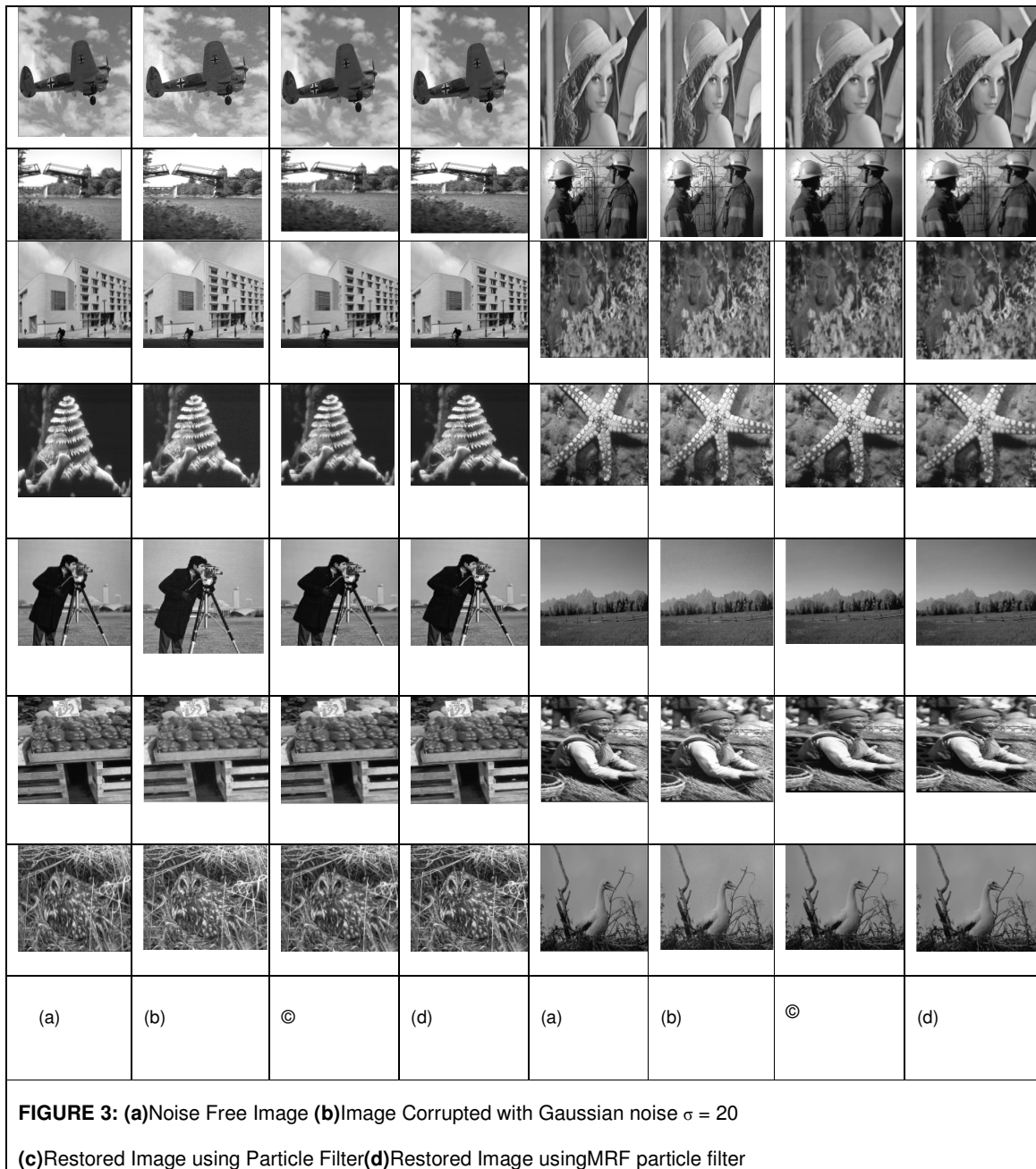


FIGURE 2. (a) MSE of particle filter & MRF particle filter for 20 images at Gaussian noise level $\sigma=20$
 (b) PSNR of particle filter & MRF particle filter for 20 images at Gaussian noise level $\sigma=20$





5. CONCLUSION

In this paper, an efficient algorithm is proposed for removing noise from corrupted image. This was achieved by particle filtering using MRF which is an automatic technique to capture the scale of the texture. The contribution of our method is the selection of an appropriate window in the image domain. For this we first construct a set containing all occurrences then the conditional pdf can be estimated with a histogram of all center pixel values. Particle evolution is controlled by the image structure leading to a filtering window adapted to the image content. To demonstrate the superior performance of the proposed method, extensive experiments have been conducted on several standard test images. The proposed MRF particle filter performs better than particle filter both in PSNR and visually. Promising experimental results demonstrate the potentials of our approach.

The limitation of this work is Computational complexity. And the performance of the system is reduced since each particle is iterated separately.

Future work is First,improving the learning stage of the image structural modeland guiding the particles to the most appropriate directionscould be a step toward increasing the efficiency of particletransitions. Next, thelikelihood measure could be also modified to be more specificfor each noise distribution and more robust. Last, the ability to perform the process in parallel for all pixelsand benefit from the reconstructed values might improve theperformance of the method.

6. REFERENCES

- [1]M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesiantracking," *IEEE Trans. Signal Process.*, vol. 50, pp. 174–188, 2002.
- [2] N. Azzabou, N. Paragios, and F. Guichard, "Random walks, constrained multiple hypothesis testing and image enhancement," in*Proc.Eur. Conf. ComputerVision*, 2006, pp. 379–390.
- [3] N. Azzabou, N. Paragios, F. Guichard, and F. Cao, "Variable bandwidth image denoising using image-based noise models," in *Proc. IEEE Int.Conf. Computer Vision and Pattern Recognition*, 2007, pp. 1–7.
- [4] J. S. D. Bonet. Multiresolution sampling procedure for analysis and synthesis of texture images. In *SIGGRAPH '97*, pages 361–368, 1997.
- [5] A. Buades, B. Coll, and J.-M.Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2005, pp. 60–65.
- [6] J. Carpenter, P. Clifford, and P. Fearnhead, "Improved particle filter for nonlinear problems," *Proc. Inst. Elect. Eng., Radar, Sonar, Navig.*, 1999.
- [7] T. Clapp and S. Godsill, "Improvement strategies for Monte Carlo particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G.deFreitas, and N. J. Gordon, Eds. New York: Springer-Verlag, 2001.
- [8]P. Del Moral, "Non-linear filtering: Interacting particle solution," *Markov Processes Related Fields*, vol. 2, no. 4, pp. 555–580.
- [9] A. Doucet, J. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*.New York: Springer-Verlag, 2001.
- [10]A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans.Signal Processing*, vol. 49, pp. 613–624, Mar. 2001.
- [11]A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. Int. Conf. Computer Vision*, 1999, pp. 1033–1038.
- [12]D. Geman, "Random fields and inverse problems in imaging," in *Lecture Notes in Mathematics*, vol. 1427, pp. 113–193. Springer–Verlag, 1991.
- [13]S. Geman and C. Graffigne, "Markov random field image models and their applications to computer vision," *Proceedings of theInternational Congress of Mathematicians*, pp. 1496–1517, 1986.
- [14] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol.SMC-6, pp. 610–621, 1973.
- [15] R. E. Helmick, D. Blair, and S. A. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Trans. Inform.Theory*, vol. 41, pp.1845–1855, Nov. 1995.
- [16]J. Huang and D. Mumford, "Statistics of natural images and models," in *Proc. IEEE Int. Conf. Computer Vision and PatternRecognition*, 1999, pp. 541–547.
- [17]John Moussouris, "Gibbs and Markov random systems with constraints," *Journal of Statistical*

Physics, vol. 10, no. 1, pp. 11–33, 1974.

[18]G. Kitagawa, “Monte carlo filter and smoother for non-Gaussian nonlinear state space models,” *J. Comput.Graph.Statist.*, vol. 5, pp.1–25, 1996.

[19]A. Lee, K. Pedersen, and D. Mumford, “The nonlinear statistics of high-contrast patches in natural images,” *Int. J. Comput.Vis.*, pp.83–103, 2003.

[20]S. Lee, “Digital image smoothing and the sigma filter,” *CVGIP*, vol. 24, no. 2, pp. 255–269, Nov. 1983.

[21] J. S. Liu and R. Chen, “Sequential Monte Carlo methods for dynamical systems,” *J. Amer. Statist.Assoc.*, vol. 93, pp. 1032–1044,1998.

[22]M. Mahmoudi and G. Sapiro, “Fast image and video denoising via nonlocal means of similar neighborhoods,” *IEEE SignalProcess.Lett.*, vol. 12, pp. 839–842, 2005.

[23]S.Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 674–693,1989.

[24]J. Polzehl and V. Spokoiny, “Adaptive weights smoothing with applications to image restoration,” *J. Roy.Statist.Soc.B*, vol. 62, pp.335–354, 2000.

[25] Richard C. Dubes and Anil K. Jain, “Random field models in image analysis,” *Journal of Applied Statistics*, vol. 16, no. 2, pp. 131–164, 1989.

[26] B. Smolka and K. Wojciechowski, “Random walk approach to image enhancement,” *Signal Process.*, vol. 81, pp. 465–482, 2001.

[27] R. L. Streit and R. F. Barrett, “Frequency line tracking using hidden Markov models,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 586–598, Apr. 1990. 188 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 50, NO. 2,FEBRUARY 2002.

[28]C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. Int. Conf. Computer Vision*, 1998, pp. 839–846.

Fast Segmentation of Sub-Cellular Organelles

Dilip K. Prasad

*School of Computer Engineering
Nanyang Technological University
Singapore, 639798*

dilipprasad@gmail.com

Chai Quek

*School of Computer Engineering
Nanyang Technological University
Singapore, 639798*

ashcquek@ntu.edu.sg

Maylor K. H. Leung

*Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman(Kampar)
Malaysia*

asmkleung@gmail.com

Abstract

Segmentation and counting sub-cellular structure of size range $0.1\mu\text{m}$ to $1\mu\text{m}$ is a very challenging problem even for medical experts. Image acquisition time for the state-of-the-art fluorescence microscopes is about 1-5 seconds. A fast and efficient method for segmentation and counting of sub-cellular structure is proposed for concurrent post-processing of fluorescence microscopy images. The proposed method uses a hybrid combination of several image processing techniques and is effective in segmenting the sub-cellular structures in a fast and effective manner.

Keywords: Segmentation, Bio-cell Organelles, Fluorescence Microscopy, Ellipse Detection.

1. INTRODUCTION

Cell and sub-cellular segmentation in biomedical images is helpful in diagnosis and cell biology research. Often manual segmentation and classification is slow since one image may contain numerous cells or sub-cellular structures. Automatic segmentation software can be used for this purpose but the accuracy is often low and is often unable to filter away the artifacts on its own.

Here, a segmentation algorithm for cells or sub-cellular structures that can be modeled as elliptic shapes is proposed. Examples of such datasets can be found in [1-4]. The dataset considered in this paper is a dataset of images of mixed cell organelle types (mitochondria and lysosomes) [4]. Fluorescence confocal microscope is used to generate the images. Preprocessing of the images leads to cleaner images in which the cells appear in the foreground (for example as shown in FIGURE 3(a)). These images are used as input images for the proposed algorithm. It is highlighted that in several of the images, no sub-cellular structures are seen. Thus, images that contain sub-cellular features are selected manually. There are a total of 444 such images, each of size 1349×1030 pixels.

The proposed algorithm employs three simple blocks and the core of the algorithm is a least squares ellipse fitting method. The algorithm is tested on a dataset of images of two types of sub-cellular structures [4] and the algorithm shows a good performance. Further it is quite fast and easily parallelizable. Thus, with some code optimization, it can be made real time.

The proposed algorithm is presented in section 2. The dataset for testing the proposed algorithm and the control parameters chosen for this dataset are also discussed in section 2. The results are presented in section 3. The paper is concluded in section 4.

2. PROPOSED ALGORITHM

The proposed algorithm employs three simple blocks, viz., pre-processing, ellipse fitting, and ellipse selection. The flowchart of the proposed algorithm is shown in FIGURE 1.

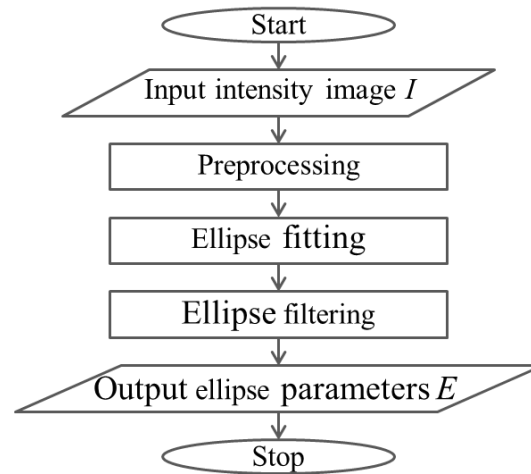


FIGURE 1: Flowchart of the proposed method.

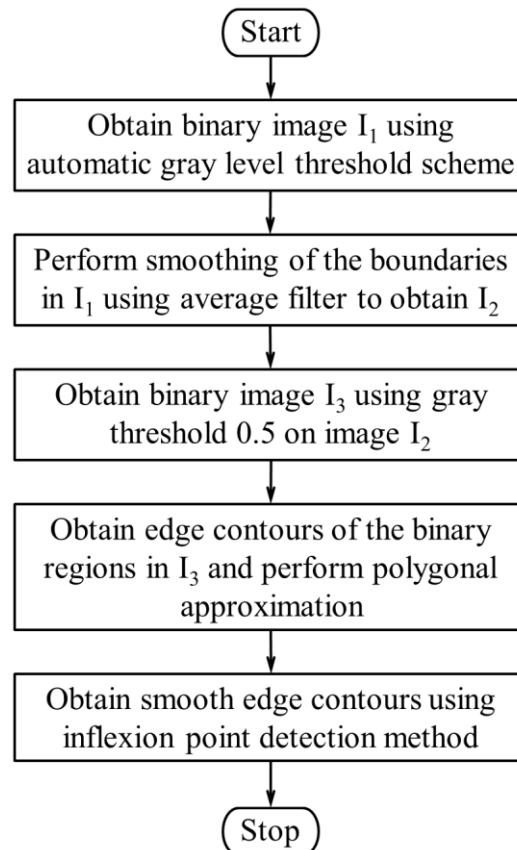


FIGURE 2: Flowchart for preprocessing block.

2.1 Preprocessing Block

It is assumed that the input is the intensity image I (gray scale 0-255). The flowchart of this block is presented in FIGURE 2. One of the critical aspects is the choice of the threshold level for obtaining the binary images. The images in the biomedical dataset obtained using microscopy can suffer from several issues. Examples include low contrast, bleaching (background illumination), noise, scattering from irrelevant organelles, etc. However, assuming that the same instrument and measurement setup is used to generate the images in a particular dataset, a suitable threshold level t_1 can be determined *a priori* for binarizing the image (step 1 in FIGURE 2), where the binary image is being referred to as I_1 .

In order to deal with extremely low contrast features, it is preferable to use a low value of t_1 . Consider for example the image in FIGURE 3(a). The highlighted circle shows a region that contains a cell but is invisible due to extremely low contrast. As a consequence, this cell is present in the binarized image when $t_1 = 5$ is used (FIGURE 3(b)) but absent when $t_1 = 15$ is used (FIGURE 3(c)). It is also notable that artifacts due to noise appear when low threshold is used.

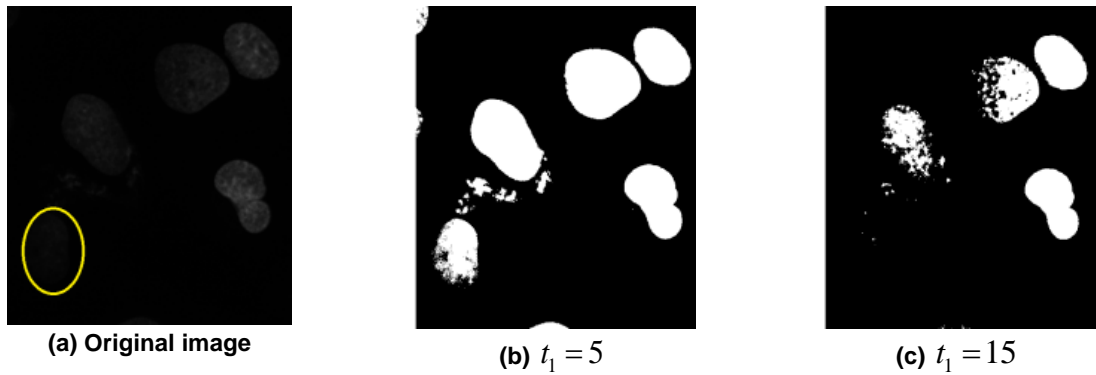


FIGURE 3: Example of the effect of threshold value t_1 on the binary image I_1 .

Thus, a suitable selection of the threshold is quite important. A statistical scheme is proposed here for choosing the value of threshold t_1 . For determining the suitable value of t_1 , the histogram of each image in the dataset is generated for gray levels 0 – 255. Let the histogram count for a gray level $g \in \{0, 255\}$ for an image I be denoted as $h(g, I)$. Cumulative histogram is generated for each gray level and image is computed as:

$$\bar{C}(g, I) = \sum_{g'=0}^g h(g', I) \quad (1)$$

Then, the normalized cumulative histogram is computed as follows:

$$C(g, I) = \frac{\bar{C}(g, I)}{\max(\bar{C}(g, I) | \forall g)} \quad (2)$$

The values of the normalized cumulative histogram for a given gray level g are averaged for all the images and plotted in FIGURE 4(a). It is seen that the images have low intensity since only lower gray levels (till $g = 100$) have contribution in the images. A zoom-in of FIGURE 4(a) is provided in FIGURE 4 (b) where only $g = 0$ to 31 is considered). It is seen that $g = 4$ is sufficient for more than 80% of the cumulative histogram of the images. Thus, $t_1 = 4$ is chosen as the threshold for binarizing the images (step 1 of FIGURE 2).

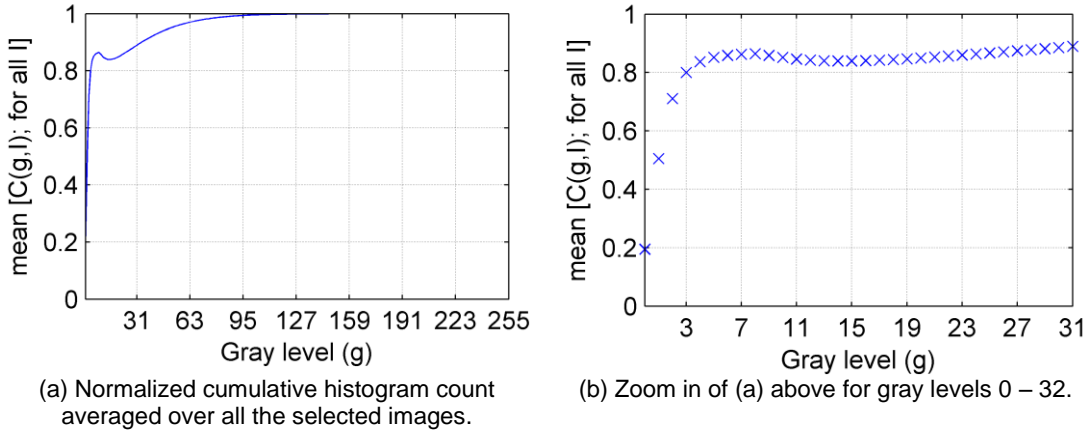


FIGURE 4: Normalized cumulative histogram averaged over all the images used for choosing the value of threshold t_1 .

Yet, as shown in FIGURE 3, the boundaries of the cells may not be well defined due the binarization. Thus, a spatial mean filter of size 5 is applied on the binary image I_1 to get a gray scale image I_2 with smoother features (step 2 in FIGURE 2). The image I_2 is binarized again using a threshold value of 0.5, since the image I_2 obtained after applying average filter on the binary image I_1 (step 3 in FIGURE 2) is a gray scale image. The final binary image used for further processing is denoted as I_3 .

The edge contours of the white regions in I_3 are extracted (step 4 in FIGURE 2) and are denoted by index $e=1$ to \hat{e} . For each edge e , polygonal approximation of the edge is derived using RDP-mod proposed in [5] (step 4 in FIGURE 2). Other polygonal approximation approaches may be considered as well [5-15]. The polygonal approximation of the edge contour is consequently used to remove the inflexion points and obtain smooth edges $e'=1$ to \hat{e}' using the algorithm for removing inflexion points appearing in section 2 of [16] (step 5 in FIGURE 2), see also [17-19].

2.2 Ellipse Fitting Block

This block calls a least squares based ellipse fitting method for each edge $e'=1$ to \hat{e}' . If the method generates a valid ellipse, the geometric parameters of the ellipse (length of semi-major axis a , length of semi-minor axis b , x-coordinate of the center x_0 , y-coordinate of the center y_0 , and the angle made by semi-major axis with the x -axis α of the fitted ellipse) is appended in the array containing the parameters of ellipses E .

The choice of the ellipse fitting method has a significant impact on the overall performance of the method. Hough transform based methods have several problems like a huge number of samples are required to obtain robust results, five-dimensional parameter space of ellipses is difficult to deal with computationally, and the whole process can be significantly time consuming. Hybrid methods for high end applications [16, 20, 21], which generally give high accuracy, are slow and computationally expensive. They also include other very sensitive steps like tangent estimation [22-27]. So, least squares based methods for fitting ellipses can be used. In our numerical results, it shall be shown that the geometry based least squares fitting method presented in [28] performs better than other least squares methods.

For the sake of completeness, the algorithmic structure of the geometry based least squares fitting method presented in [28] is reproduced here. For a sequence of pixels $\{P_i(x'_i, y'_i); i=1$ to $N\}$, we intend to find the parameters a, b, θ_c, x_c , and y_c , such that the residue

given in eqn. (3) is minimized. Here, $|\cdot|$ denotes the absolute value in the case of scalars and Euclidean norm in the case of vectors.

$$r'_i = |y'_i - m'_i x'_i - c'_i|. \quad (3)$$

Step 1: Compute $\tilde{x}_i = x'_i - \tilde{x}$, $\tilde{y}_i = y'_i - \tilde{y}$ where (\tilde{x}, \tilde{y}) are given by eqn. (4).

$$\tilde{x} = \frac{1}{N} \sum_{\forall i} x'_i; \quad \tilde{y} = \frac{1}{N} \sum_{\forall i} y'_i \quad (4)$$

Step 2: Form the matrix \mathbf{X} and the vector \bar{Y} as given in eqns. (5) and (6).

$$\mathbf{X} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{x}_i^2 & -\tilde{x}_i \tilde{y}_i & \tilde{x}_i & \tilde{y}_i & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (5)$$

$$\bar{Y} = [\tilde{y}_1^2 \quad \tilde{y}_2^2 \quad \dots \quad \tilde{y}_N^2]^T, \quad \bar{Y} \in \mathbb{Z}^N \quad (6)$$

Step 3: Compute $\bar{\Phi}$ using eqn. (7).

$$\bar{\Phi} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \bar{Y} \quad (7)$$

Step 4: Compute $a, b, \theta_c, \tilde{x}_c$, and \tilde{y}_c using eqns. (8) – (12).

$$a = 2 \frac{\sqrt{\phi_2 \phi_3 \phi_4 - \phi_4^2 \phi_1 - \phi_3^2 - \phi_5 (\phi_2^2 - 4\phi_1)}}{\sqrt{(\phi_2^2 - 4\phi_1) \left((1 + \phi_1) - \sqrt{(1 - \phi_1)^2 + \phi_2^2} \right)}} \quad (8)$$

$$b = 2 \frac{\sqrt{\phi_2 \phi_3 \phi_4 - \phi_4^2 \phi_1 - \phi_3^2 - \phi_5 (\phi_2^2 - 4\phi_1)}}{\sqrt{(\phi_2^2 - 4\phi_1) \left((1 + \phi_1) + \sqrt{(1 - \phi_1)^2 + \phi_2^2} \right)}} \quad (9)$$

$$\theta_c = -0.5 \tan^{-1}(-\phi_2 / (1 - \phi_1)) \quad (10)$$

$$\tilde{x}_c = (\phi_2 \phi_4 - 2\phi_3) / (\phi_2^2 - 4\phi_1) \quad (11)$$

$$\tilde{y}_c = (\phi_2 \phi_3 - 2\phi_1 \phi_4) / (\phi_2^2 - 4\phi_1) \quad (12)$$

Step 5: Compute $x_c = \tilde{x}_c + \tilde{x}$ and $y_c = \tilde{y}_c + \tilde{y}$.

2.3 Ellipse Filtering Block

After fitting ellipses on each edge contour, available *a priori* information about the dataset can be used to filter/remove some unreasonable ellipses. The filtering criteria depend upon the dataset and *a priori* information known about it. For example the imaging resolution and the CCD grid size can be used to determine the size range of the cells and thus the bounds on the lengths of semi-major axis a and semi-minor axis b may be generated. Also, the biological information about the cells can be used to generate an estimate of the maximum ratio of the semi-major and semi-minor axes a/b . In the proposed algorithm, the following filtering criteria have been used:

$$a/b \leq t_2 \quad (13)$$

$$t_3 \leq b \leq t_4 \tag{14}$$

where t_2 , t_3 , and t_4 are the thresholds determined based on the apriori knowledge of the cells and the imaging system. The ellipses satisfying eqns. (13) and (14) are retained. It is known that the size of the lysosomes and mitochondria varies in the range $0.1 \mu\text{m}$ to $1 \mu\text{m}$. Thus, the value of the threshold t_2 is chosen as the ratio of the maximum to the minimum size of these sub-cellular structures, i.e., $t_2 = 10$. Further, based on the general size of the sub-cellular structures in the dataset, the threshold values of t_3 and t_4 are chosen as 10 and 100 respectively.

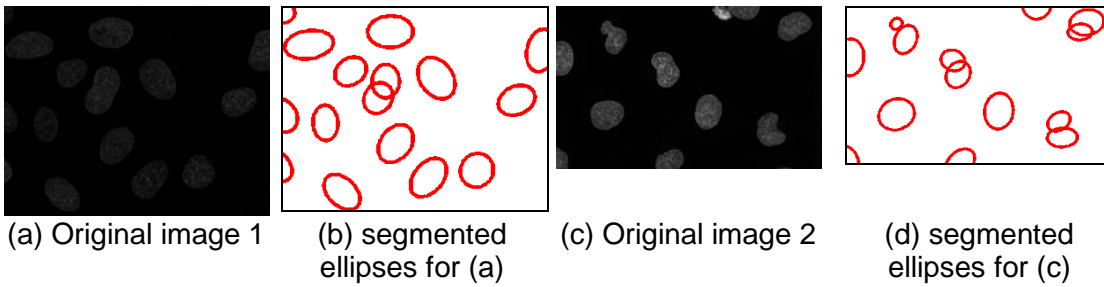


FIGURE 5: Examples of images and ellipses segmented by the proposed method.

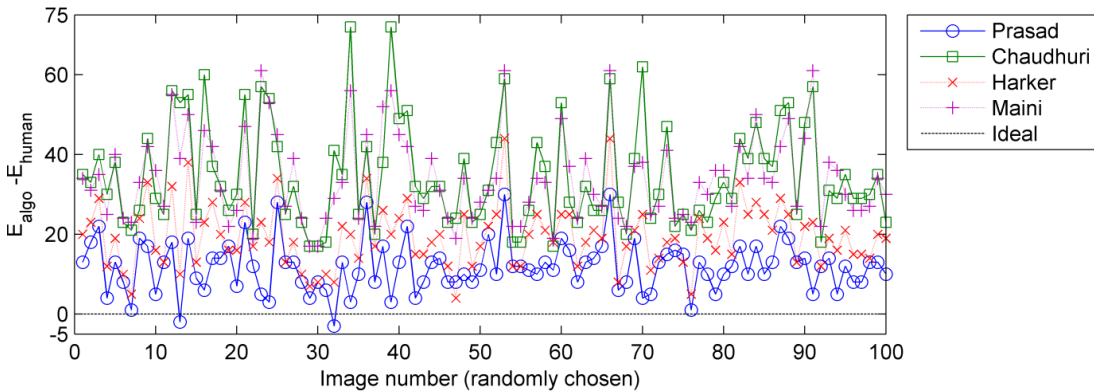


FIGURE 6: Comparison of the performance for various least squares methods. Prasad here represents the unconstrained least squares method presented in section 2.2.

3. NUMERICAL RESULTS

Two examples of the result of the proposed algorithm are shown in FIGURE 5. It is clearly seen that the proposed algorithm can segment the sub-cellular structures very well for low contrast as well as high contrast images, even when some of the sub-cellular structures may be occluded by other structures. The number of ellipses found using the proposed algorithm is denoted using E_{algo} and the number of ellipses found by a human (occluded or otherwise) is denoted by E_{human} . The quantity $E_{\text{algo}} - E_{\text{human}}$ is plotted for 100 randomly chosen images (since it is difficult to collectively present and compare the results for all 444 images) in FIGURE 6. As highlighted in section 2.2, the method chosen for least squares ellipse fitting has a significant impact on the performance of the algorithm. Thus, four methods (Prasad [28] – used in the proposed algorithm and presented in section 2.2, Chaudhuri [29], Harker [30], and Maini [31]) are considered and their performances are compared in FIGURE 6. It is seen that Prasad gives the best performance, thus proving to be the best choice among the four least squares ellipse fitting methods.

	Average time (s)	Standard Deviation
Preprocessing	0.52	0.43
Ellipse fitting (Prasad)	7.86	2.04
Ellipse fitting (Chaudhuri)	9.42	2.93
Ellipse fitting (Harker)	11.27	3.38
Ellipse fitting (Maini)	9.46	3.01
Ellipse filtering	0.02	0.02

TABLE 1: Statistics of computation time without parallel processing

	Average time (s)	Standard Deviation
Preprocessing	0.52	0.43
Ellipse fitting (Prasad)	0.99	0.30
Ellipse fitting (Chaudhuri)	1.21	0.38
Ellipse fitting (Harker)	1.42	0.43
Ellipse fitting (Maini)	1.23	0.39
Ellipse filtering	0.02	0.02

TABLE 2: Statistics of computation time with parallel processing
(parallelization of step 2 of
FIGURE 2 is performed using 8 parallel cores)

The time comparison of the four methods is presented in TABLE 1 (without any parallelization). It is seen that among the four least squares ellipse fitting methods, Prasad takes the least computation time as well. It is also noted that the preprocessing and ellipse filtering steps take very little time. In fact, the computation time for each image can be easily reduced below 1 second by parallelizing the ellipse fitting block. This is illustrated in TABLE 2, where it is shown that the proposed method with Prasad's unconstrained least squares method takes less than 1 second for the most time consuming portion when 8 cores are used. The time performance can be further improved by increasing the number of cores and more effective programming. Thus, the proposed algorithm is capable of providing computation time less than the image acquisition time of a typical fluorescence microscope.

4. CONCLUSION

A fast algorithm is proposed to segment the sub-cellular structures in a dataset of fluorescence microscopy images. The three blocks in the proposed algorithm are simple, fast and effective. The core of the algorithm is the least squares fitting method used in the ellipse fitting block. It is shown that [28] is superior choice than three other least squares fitting methods.

REFERENCES

- [1] S. Kumar, S. H. Ong, S. Ranganath, and F. T. Chew, "Invariant texture classification for biomedical cell specimens via non-linear polar map filtering," *Computer Vision and Image Understanding*, vol. 114, pp. 44-53, Jan 2010.
- [2] G. Dong, N. Ray, and S. T. Acton, "Intravital leukocyte detection using the gradient inverse coefficient of variation," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 910-924, 2005.
- [3] X. Z. Bai, C. M. Sun, and F. G. Zhou, "Splitting touching cells based on concave points and ellipse fitting," *Pattern Recognition*, vol. 42, pp. 2434-2446, Nov 2009.
- [4] T. Peng, G. M. C. Bonamy, E. Glory-Afshar, D. R. Rines, S. K. Chanda, and R. F. Murphy, "Determining the distribution of probes between different subcellular locations through

- automated unmixing of subcellular patterns," *Proceedings of the National Academy of Sciences of the USA*, vol. 107, pp. 2944-2949, Feb 2010.
- [5] D. K. Prasad, M. K. H. Leung, C. Quek, and S.-Y. Cho, "A novel framework for making dominant point detection methods non-parametric," *Image and Vision Computing*, 2012.
- [6] D. K. Prasad and M. K. H. Leung, "Reliability/Precision Uncertainty in Shape Fitting Problems," in *IEEE International Conference on Image Processing*, Hong Kong, 2010, pp. 4277-4280.
- [7] D. K. Prasad and M. K. H. Leung, "Polygonal representation of digital curves," in *Digital Image Processing*, S. G. Stanciu, Ed., ed: InTech, 2012, pp. 71-90.
- [8] D. K. Prasad, C. Quek, and M. K. Leung, "A non-heuristic dominant point detection based on suppression of break points," in *Image Analysis and Recognition*. vol. 7324, A. Campilho and M. Kamel, Eds., ed Aveiro, Portugal: Springer Berlin Heidelberg, 2012, pp. 269-276.
- [9] D. K. Prasad, C. Quek, M. K. H. Leung, and S. Y. Cho, "A parameter independent line fitting method," in *Asian Conference on Pattern Recognition (ACPR)*, Beijing, China, 2011, pp. 441-445.
- [10] D. K. Prasad, "Assessing error bound for dominant point detection," *International Journal of Image Processing*, vol. 6, 2012.
- [11] N. Barnes, G. Loy, and D. Shaw, "The regular polygon detector," *Pattern Recognition*, vol. 43, pp. 592-602, 2010.
- [12] A. Carmona-Poyato, R. Medina-Carnicer, F. J. Madrid-Cuevas, R. Muoz-Salinas, and N. L. Fernandez-Garca, "A new measurement for assessing polygonal approximation of curves," *Pattern Recognition*, vol. 44, pp. 45-54, 2011.
- [13] P. C. Chung, C. T. Tsai, E. L. Chen, and Y. N. Sun, "Polygonal approximation using a competitive Hopfield neural network," *Pattern Recognition*, vol. 27, pp. 1505-1512, 1994.
- [14] T. M. Cronin, "A boundary concavity code to support dominant point detection," *Pattern Recognition Letters*, vol. 20, pp. 617-634, 1999.
- [15] F. Feschet, "Canonical representations of discrete curves," *Pattern Analysis and Applications*, vol. 8, pp. 84-94, 2005.
- [16] D. K. Prasad, M. K. H. Leung, and S. Y. Cho, "Edge curvature and convexity based ellipse detection method," *Pattern Recognition*, vol. 45, pp. 3204-3221, 2012.
- [17] D. K. Prasad and M. K. H. Leung, "An ellipse detection method for real images," in *25th International Conference of Image and Vision Computing New Zealand (IVCNZ 2010)*, Queenstown, New Zealand, 2010, pp. 1-8.
- [18] D. K. Prasad and M. K. H. Leung, "Methods for ellipse detection from edge maps of real images," in *Machine Vision - Applications and Systems*, F. Solari, M. Chessa, and S. Sabatini, Eds., ed: InTech, 2012, pp. 135-162.
- [19] X. Bai, C. Sun, and F. Zhou, "Splitting touching cells based on concave points and ellipse fitting," *Pattern Recognition*, vol. 42, pp. 2434-2446, 2009.

- [20] D. K. Prasad, "Adaptive traffic signal control system with cloud computing based online learning," in *8th International Conference on Information, Communications, and Signal Processing (ICICS 2011)*, Singapore, 2011.
- [21] G. Heitz, G. Elidan, B. Packer, and D. Koller, "Shape-based object localization for descriptive classification," *International Journal of Computer Vision*, vol. 84, pp. 40-62, 2009.
- [22] D. K. Prasad and M. K. H. Leung, "Error analysis of geometric ellipse detection methods due to quantization," in *Fourth Pacific-Rim Symposium on Image and Video Technology (PSIVT 2010)*, Singapore, 2010, pp. 58 - 63.
- [23] D. K. Prasad, M. K. H. Leung, and C. Quek, "DEB: Definite error bounded tangent estimator for digital curves," *Image and Vision Computing*, 2012(under review).
- [24] D. K. Prasad, R. K. Gupta, and M. K. H. Leung, "An Error Bounded Tangent Estimator for Digitized Elliptic Curves," in *Discrete Geometry for Computer Imagery*. vol. 6607, ed: Springer Berlin / Heidelberg, 2011, pp. 272-283.
- [25] I. M. Anderson and J. C. Bezdek, "Curvature and tangential deflection of discrete arcs: a theory based on the commutator of scatter matrix pairs and its application to vertex detection in planar shape data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 27-40, 1984.
- [26] D. M. Tsai and M. F. Chen, "Curve fitting approach for tangent angle and curvature measurements," *Pattern Recognition*, vol. 27, pp. 699-711, 1994.
- [27] B. Dubuc and S. W. Zucker, "Complexity, confusion, and perceptual grouping. Part I: The curve-like representation," *Journal of Mathematical Imaging and Vision*, vol. 15, pp. 55-82, 2001.
- [28] D. K. Prasad, C. Quek, and M. K. Leung, "A precise ellipse fitting method for noisy data," in *Image Analysis and Recognition*, A. Campilho and M. Kamel, Eds., ed Aveiro, Portugal: Springer Berlin Heidelberg, 2012, pp. 253-260.
- [29] D. Chaudhuri, "A simple least squares method for fitting of ellipses and circles depends on border points of a two-tone image and their 3-D extensions," *Pattern Recognition Letters*, vol. 31, pp. 818-829, Jul 2010.
- [30] M. Harker, P. O'Leary, and P. Zsombor-Murray, "Direct type-specific conic fitting and eigenvalue bias correction," *Image and Vision Computing*, vol. 26, pp. 372-381, 2008.
- [31] E. S. Maini, "Enhanced direct least square fitting of ellipses," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 20, pp. 939-953, 2006.

Assessing Error Bound For Dominant Point Detection

Dilip K. Prasad

*School of Computer Engineering
Nanyang Technological University
Singapore, 639798*

dilipprasad@gmail.com

Abstract

This paper compares the error bounds of two classes of dominant point detection methods, viz. methods based on reducing a distance metric and methods based on digital straight segments. We highlight using two methods in each class that the error bounds within the same class may vary depending upon the fine details and control parameters of the methods but the essential error bound can be determined analytically for each class. The assessment presented in this paper will enable the users of dominant point detection methods to understand the nature of error in the method of their choice and help them to make better decision about the methods and their control parameters.

Keywords: Dominant point detection, digital straight segment, error bound, comparison, discrete geometry.

1: INTRODUCTION

Dominant point detection in digital curves is a preliminary but important step in various image processing applications like shape extraction, object detection, etc. [1-13]. Despite being a very old problem of interest, this problem attracts significant attention even today in the research community. Some of the recent PA methods are proposed by Masood [14, 15], Carmona-Poyato [16], Wu [17], Kolesnikov [18, 19], Chung [20], Ngyuen [21], and Bhowmick [22] while few older ones are found in [23-33]. These algorithms can be generally classified based upon the approach taken by them. Often, algorithmic approach is used for classification. For example, some used dynamic programming [18, 19, 23], while others used splitting [27-29], merging [24], tree search [21, 22, 34], suppression of break points [14-16, 35, 36], etc.. However, this is not the focus of the current work.

The focus of the current work is the basic discrete geometry approach used in the methods since the geometric concept used in the method determines the achievable accuracy or the inherent error bound of the dominant point detection methods. In the sense of geometric approach, there are three major categories – methods based on minimization of a distance metric (like maximum deviation, integral square error, etc.) [14-20, 27-33, 37], based on the concept of digital straight segments [21, 22, 38], or based on curvature and convexity [17, 20, 25, 26, 30-32]. We highlight that the methods based on curvature and convexity often use k-cosine term for studying the convexity and curvature changes and the decisive factor in the selection of dominant points is often based on some distance metric or another. Thus, the error analysis of methods based on curvature and convexity is considered redundant with the error analysis of methods based on distance metric presented in this paper.

The outline of the paper is as follows. The error bound for the methods based on distance metrics is presented in section 2. The error bound for the methods based on digital straight segments is presented in section 3. The paper is concluded in section 4.

2: METHODS BASED ON MINIMIZING THE MAXIMUM DEVIATION

Two most famous and classic methods from this class of dominant point detection methods are considered here, Ramer-Douglas-Peucker method [28, 29] (RDP) and Lowe's method [27].

These methods are classical methods which use splitting the digital curve based on maximum deviation. They have laid foundation for dominant point detection in terms of the algorithm, the idea of using distance measures, and the support region of the dominant points. Thus, they have been the basis of several later works on dominant point detection methods. Notwithstanding the later work, their computational efficiency and effectiveness in representing the digital curves has ensured the popularity and use of these methods in several image processing applications. These methods and their error metrics are discussed in the subsequent sub-sections.

2.1 Ramer-Douglas-Peucker method

Ramer-Douglas Peucker (RDP) method was first proposed in [28, 29]. The method is briefly described as follows. Consider a digital curve $S = P_1 P_2 \dots P_N$, where P_i is the i th edge pixel in the digital curve e . By default, the start and end points P_1, P_N are included in the list of dominant points. If the digital curve is a close loop, then only one of them is included. For the straight line joining the points P_1, P_N , the deviation d_i of a pixel $P_i(x_i, y_i) \in S$ is computed. Accordingly, the pixel with maximum deviation (MD) can be found. Let it be denoted as P_{\max} . Then considering the pairs P_1, P_{\max} and P_{\max}, P_N , two new pixels of maximum deviations are found from S . This process is repeated until the condition in inequality eq. (1) is satisfied by all the line segments. The algorithm is terminated when the condition in eq. (1) is satisfied.

$$\max(d_i) < d_{\text{tol}}, \quad (1)$$

where d_{tol} is the chosen threshold and its value is typically a few pixels.

Thus, the error bound of the RDP method is determined by the theoretical value of the maximum deviation $\max(d_i)$ of the pixels from the polygon formed by the dominant points or the control parameter d_{tol} . In general, researchers heuristically choose the value of d_{tol} in the range 1, 2.

It is notable that the maximum deviation is used as optimization goal or termination condition in several methods [14, 15, 20, 36, 37]. Further, several other methods use the integral square error (ISE) as the optimization goal or termination condition [18-20, 32, 33]. Since the maximum value of the integral square error is upper bounded by $N \max(d_i)^2$ where N is the number of pixels, thus $\max(d_i)$ serves as the indicator of the upper bound for these methods as well.

2.2 Lowe's method

While the algorithmic structure of the Lowe's method proposed in [27] is essentially similar to RDP, its termination condition is different from RDP and more useful than using only the maximum deviation. Lowe considers two distances, first the maximum deviation $\max(d_i)$ of the pixels in the digital curve spanned by two consecutive dominant points (just like RDP method), second the distance between the two dominant points (say s). He defines the significance ratio:

$$r = s / \max(d_i), \quad (2)$$

and uses it as the basis for the decision to retain a dominant point. For three consecutive dominant points DP_{j-1} , DP_j , and DP_{j+1} , if $\max(r_{j-1,j}, r_{j,j+1}) > r_{j-1,j+1}$, then the dominant point DP_j is retained. This is done for all the dominant points except the start and end points. Then the points with $r_j < 4$ are also deleted.

It is notable that Lowe assumes that the maximum deviation is always at least 2 pixels. It was shown in [37, 39] that the maximum deviation is less than 2 pixels in most cases. Further, it is

evident that the error bound of Lowe's method is determined by the inverse of the significance ratio $\max(d_i)/s$.

It is also notable that such similar significance ratios were used by several subsequent researchers as well [11, 16, 17, 20, 31-35], sometimes as a constraint and other times as the decision determinant. Most notably, such a ratio has served as the criterion of the support region of the dominant points.

2.3 Comparison of the methods based on error bound of the maximum deviation

It was shown in [37, 39, 40] that if a continuous line segment is digitized, the maximum distance of the pixels of the digital line from the continuous line segment is given by:

$$d_{\max} \approx \max \left\{ s \tan^{-1} \left\{ \frac{1}{s} \left| \sin \phi \pm \cos \phi \right| \sqrt{1 - t_{\max} + t_{\max}^2} \right\} \right\} \tag{3}$$

where s is the length of the continuous line segment, $\tan \phi$ is the slope of the line segment and t_{\max} is given by:

$$t_{\max} = \left(\frac{1}{s} \right) \left(|\cos \phi| + |\sin \phi| \right) \tag{4}$$

While the error bound d_{\max} directly applies to the methods based on maximum deviation and integral square error, such as discussed in section 2.1, the ratio d_{\max}/s applies to the methods based on significance ratio such as discussed in section 2.2. The expression for this is given by:

$$\delta \phi_{\max} = \frac{d_{\max}}{s} \approx \max \left\{ \tan^{-1} \left\{ \frac{1}{s} \left| \sin \phi \pm \cos \phi \right| \sqrt{1 - t_{\max} + t_{\max}^2} \right\} \right\} \tag{5}$$

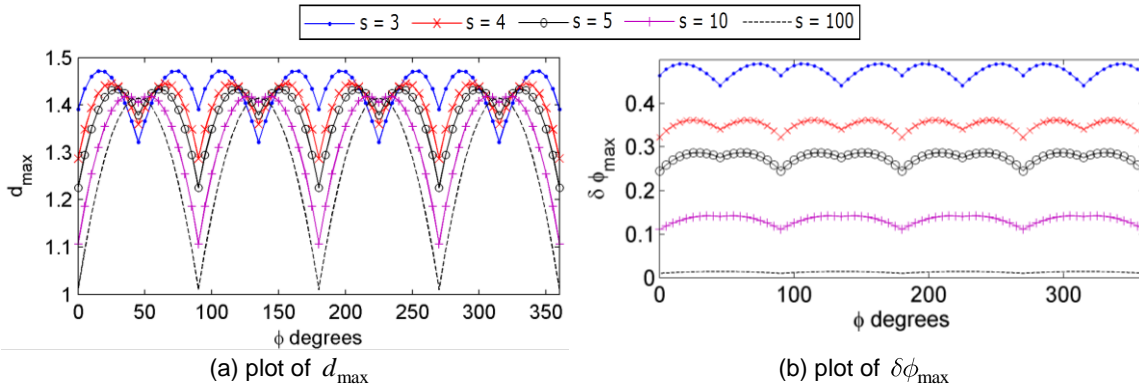


Figure 1: Illustration of the error bounds of the distance based methods.

These two errors bounds are plotted in Figure 1. Figure 1(a) shows the theoretical error bound for RDP method. Based on the analysis, it is seen that the typical range of $d_{\text{tol}} \in 1,2$ used by most researchers incorporates the theoretical error bound. This is also the basis of the non-parametric framework presented in [37]. Figure 1(a) also indicates that Lowe's assumption that the maximum deviation is at least two pixels is incorrect and the maximum deviation is typically less than two pixels. Figure 1(b) shows the theoretical error bound for Lowe's method. It is seen that the maximum value of d_{\max}/s is close to 0.5, which indicates that the constraint on the value of the

significance ratio can be $r \geq 2$, as opposed to Lowe's criterion $r \geq 4$. Similar analysis can be made regarding the error bounds of other distance based methods also using eqn. (3).

3: METHODS BASED ON DIGITAL STRAIGHT SEGMENTS

The concept of the digital straight segments (DSS) is a mathematical concept of discrete geometry [41] which specifically discusses a continuous line segment and a digital line segment and its properties. Evidently, it should serve as an important concept for dominant point detection methods. However, the simplicity and effectiveness of already popular distance based methods and the mathematical rigor of the concepts of DSS have restricted the researchers' interest in using DSS for dominant point detection. Nevertheless, DSS based methods are a very important class of dominant point detection methods, especially for extremely noisy digital curves for which the distance based metrics force over-fitting of the dominant points. Here, we consider two recent methods based on DSS, Nguyen's method [21] and Bhowmick's method [22].

3.1 Nguyen's method of blurred digital straight segments

Nguyen [21] uses the concept of maximally blurred segments for determining the dominant points on a noisy digital curve. The concept of the blurred segments in turn is based upon the concept of DSS which is presented here briefly. A digital curve is called a digital straight segment $D(a, b, \mu, \omega)$; $a, b, \mu, \omega \in \mathbb{Z}^{(+)}$ if the points on the digital curve satisfy the equation below:

$$\mu \leq ax - by \leq \mu + \omega \quad (6)$$

The digital straight segment is called maximal if $\omega = |a| + |b|$ and blurred segment of width ν if $\omega - 1 / \max(a, b) < \nu$ [21]. Thus, the error bound of Nguyen's method is determined by the value of the control parameter ν . The value of ν used in [21] varies from 0.7 for noiseless digital curves to 9 for noisy digital curves.

3.2 Bhowmick's method of approximate digital straight segments

Bhowmick's method [22] is based on approximate digital straight segments (ADSS), which is also based upon the concept of DSS. However, as compared to several usual works on DSS, Bhowmick uses the properties of the DSS derived from the Freeman chain code in [41]. Out of the four properties of DSS (R1-R4 in [22]), only two (R1 and R2) are used for defining ADSS and two additional conditions (c1 and c2 in [22]) are imposed on the digital curve to be concluded as ADSS. It is highlighted that the isothetic error bound of the ADSS was presented in [22] but is reconsidered for comparison with other methods. According to [22], the isothetic error bound is given by:

$$\varepsilon \leq 1 + \frac{d}{p+1} \quad (7)$$

where $d = \left\lfloor \frac{p+1}{2} \right\rfloor$ and p is the minimum intermediate run-lengths of the freeman chain code of the digital curve (see [22] for details). Thus, the maximum value of the isothetic error is $\varepsilon \leq 1.5$. However, the error bound of the polygonal approximation in [22] is the product of both the isothetic error and the control parameter τ selected heuristically. Thus, the net error bound of Bhowmick's method is given by 1.5τ . The value of τ in [22] varies from 1 to 14.

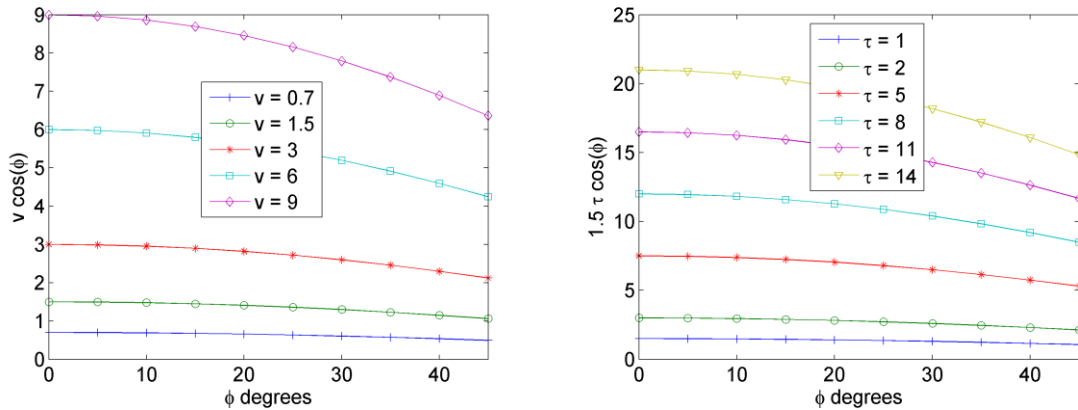
3.3 Comparison of the methods based on digital straight segments

The methods based on DSS can be compared with each other directly based on the maximum isothetic (vertical/maximum distance of the pixels from the continuous line segment) distance of the digital curve from the line segments formed by dominant points. Thus, from the control parameters' values and the error bounds of Nguyen's and Bhowmick's methods, it can be seen

that the maximum values of the upper error bounds of these methods are 9 and 21 respectively, while the minimum are 0.7 and 1.5 respectively.

If the DSS based methods should be compared with distance based methods, the maximum perpendicular distance of the pixels from the digital curves should be determined for DSS based methods also. Such distance can be easily computed by taking the projection of the maximum isothetic distance on the direction normal to the line segments. If the isothetic distance is denoted by d_{iso} , then the desired distance for comparison is computed as $|d_{iso} \cos \phi|$ where $\tan \phi$ is the slope of the line segment joining the dominant points. It is notable that in both [21] and [22], it is assumed that $0 < a < b$, which implies that $\tan \phi \in (0, 1)$.

Accordingly, the upper bound of the maximum deviation of Nguyen’s method (for width $v \in (0.7, 9)$) varies as shown in Figure 2(a) and Bhowmick’s method (for $\tau \in (1, 14)$) varies as shown in Figure 2(b). From both the upper bounds, it is seen that DSS based methods allow for a large value of maximum deviation, which is especially suitable for extremely noisy curves. As a trade-off, the quality of fitting in the DSS based methods is severely dependent upon the choice of the control parameters.



(a) Maximum deviation for Nguyen’s method for $v \in (0.7, 9)$ (b) Maximum deviation for Bhowmick’s method for $\tau \in (1, 14)$

Figure 2: Illustration of the error bounds of the maximum deviations of Nguyen’s and Bhowmick’s methods.

It is also worth considering the error bound when both the blurred segments of Nguyen [21] and the ADSS of Bhowmick [22] are both forced to be the maximal straight segments (which is a well-defined control parameter independent) quantity. In such situation, considering the equation of a line in eqn. (8):

$$ax - by = c \tag{8}$$

where a and b correspond to a maximal digital line segment $D(a, b, \mu)$ while the points $P(x, y)$ belong to the continuous two-dimensional space. For the pixels $P'(x', y')$ belonging to the digital straight segment a, b, μ , if they are to satisfy eqn. (8), then c has to satisfy inequality (9):

$$\mu < c < \mu + |a| + |b| \tag{9}$$

Using the above, the perpendicular distance (deviation) of the pixels in the maximal DSS from the line given in eqn. (8) satisfies eqn. (10) below:

$$0 < d_i < |\sin \phi| + |\cos \phi| \quad (10)$$

where $\phi = \tan^{-1} -a/b$. This error bound $d_{\text{DSS}} = |\sin \phi| + |\cos \phi|$ is plotted in Figure 3. It can be seen that the maximum deviation for the DSS based methods (assuming no blurring or approximation of DSS) is about $\sqrt{2}$ pixels.

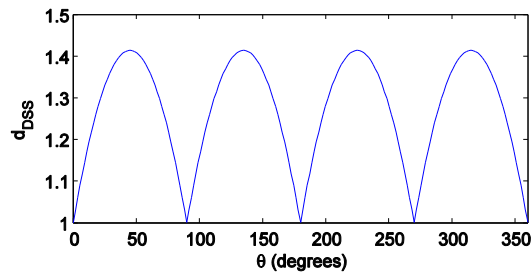


Figure 3: Illustration of the error bounds of the methods based on digital straight segments.

4: CONCLUSION

The error bounds of various methods falling into two categories of dominant point detection methods are assessed and compared in this paper. It is shown that the analytical bound on the maximum deviation can be computed for both distance based and DSS based dominant point detection methods. It is observed in each analysis that the error bound depends upon the orientation of the line segment and the control parameter of the algorithm. The assessment also gives clues on assessing the error bound of other methods as well. Thus, this work shall help researchers on studying the effect of the control parameters and the error bounds of the dominant point detection methods. A well understood choice of dominant point detection method shall in turn result into better performance for their higher level applications as well [1-13, 42, 43].

REFERENCES

- [1] S. Lavalée and R. Szeliski, "Recovering the position and orientation of free-form objects from image contours using 3D distance maps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 378-390, 1995.
- [2] D. K. Prasad and M. K. H. Leung, "A hybrid approach for ellipse detection in real images," in *2nd International Conference on Digital Image Processing*, Singapore, 2010, pp. 754601-6.
- [3] J. H. Elder and R. M. Goldberg, "Image editing in the contour domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 291-296, 2001.
- [4] D. K. Prasad and M. K. H. Leung, "Reliability/Precision Uncertainty in Shape Fitting Problems," in *IEEE International Conference on Image Processing*, Hong Kong, 2010, pp. 4277-4280.
- [5] D. Brunner and P. Soille, "Iterative area filtering of multichannel images," *Image and Vision Computing*, vol. 25, pp. 1352-1364, 2007.
- [6] D. K. Prasad and M. K. H. Leung, "Error analysis of geometric ellipse detection methods due to quantization," in *Fourth Pacific-Rim Symposium on Image and Video Technology (PSIVT 2010)*, Singapore, 2010, pp. 58 - 63.
- [7] D. K. Prasad and M. K. H. Leung, "An ellipse detection method for real images," in *25th International Conference of Image and Vision Computing New Zealand (IVCNZ 2010)*, Queenstown, New Zealand, 2010, pp. 1-8.
- [8] R. Yang and Z. Zhang, "Eye gaze correction with stereovision for video-teleconferencing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 956-960, 2004.

- [9] D. K. Prasad and M. K. H. Leung, "Methods for ellipse detection from edge maps of real images," in *Machine Vision - Applications and Systems*, F. Solari, M. Chessa, and S. Sabatini, Eds., ed: InTech, 2012, pp. 135-162.
- [10] D. K. Prasad, C. Quek, and M. K. H. Leung, "Fast segmentation of sub-cellular organelles," *International Journal of Image Processing*, vol. 6, 2012.
- [11] A. Kolesnikov and P. Fränti, "Data reduction of large vector graphics," *Pattern Recognition*, vol. 38, pp. 381-394, 2005.
- [12] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 34-43, 1986.
- [13] D. K. Prasad, M. K. H. Leung, and S. Y. Cho, "Edge curvature and convexity based ellipse detection method," *Pattern Recognition*, vol. 45, pp. 3204-3221, 2012.
- [14] A. Masood, "Dominant point detection by reverse polygonization of digital curves," *Image and Vision Computing*, vol. 26, pp. 702-715, 2008.
- [15] A. Masood and S. A. Haq, "A novel approach to polygonal approximation of digital curves," *Journal of Visual Communication and Image Representation*, vol. 18, pp. 264-274, 2007.
- [16] A. Carmona-Poyato, F. J. Madrid-Cuevas, R. Medina-Carnicer, and R. Muñoz-Salinas, "Polygonal approximation of digital planar curves through break point suppression," *Pattern Recognition*, vol. 43, pp. 14-25, 2010.
- [17] W. Y. Wu, "An adaptive method for detecting dominant points," *Pattern Recognition*, vol. 36, pp. 2231-2237, 2003.
- [18] A. Kolesnikov and P. Fränti, "Reduced-search dynamic programming for approximation of polygonal curves," *Pattern Recognition Letters*, vol. 24, pp. 2243-2254, 2003.
- [19] A. Kolesnikov and P. Fränti, "Polygonal approximation of closed discrete curves," *Pattern Recognition*, vol. 40, pp. 1282-1293, 2007.
- [20] K. L. Chung, P. H. Liao, and J. M. Chang, "Novel efficient two-pass algorithm for closed polygonal approximation based on LISE and curvature constraint criteria," *Journal of Visual Communication and Image Representation*, vol. 19, pp. 219-230, 2008.
- [21] T. P. Nguyen and I. Debled-Rennesson, "A discrete geometry approach for dominant point detection," *Pattern Recognition*, vol. 44, pp. 32-44, 2011.
- [22] P. Bhowmick and B. B. Bhattacharya, "Fast polygonal approximation of digital curves using relaxed straightness properties," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1590-1602, 2007.
- [23] J. C. Perez and E. Vidal, "Optimum polygonal approximation of digitized curves," *Pattern Recognition Letters*, vol. 15, pp. 743-750, 1994.
- [24] L. J. Latecki and R. Lakämper, "Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution," *Computer Vision and Image Understanding*, vol. 73, pp. 441-454, 1999.
- [25] B. K. Ray and K. S. Ray, "An algorithm for detection of dominant points and polygonal approximation of digitized curves," *Pattern Recognition Letters*, vol. 13, pp. 849-856, 1992.
- [26] P. V. Sankar and C. U. Sharma, "A parallel procedure for the detection of dominant points on a digital curve," *Computer Graphics and Image Processing*, vol. 7, pp. 403-412, 1978.
- [27] D. G. Lowe, "Three-dimensional object recognition from single two-dimensional images," *Artificial Intelligence*, vol. 31, pp. 355-395, 1987.
- [28] U. Ramer, "An iterative procedure for the polygonal approximation of plane curves," *Computer Graphics and Image Processing*, vol. 1, pp. 244-256, 1972.
- [29] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, pp. 112-122, 1973.
- [30] C.-H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 859-872, 1989.
- [31] N. Ansari and K. W. Huang, "Non-parametric dominant point detection," *Pattern Recognition*, vol. 24, pp. 849-862, 1991.
- [32] T. M. Cronin, "A boundary concavity code to support dominant point detection," *Pattern Recognition Letters*, vol. 20, pp. 617-634, 1999.

- [33] M. Salotti, "Optimal polygonal approximation of digitized curves using the sum of square deviations criterion," *Pattern Recognition*, vol. 35, pp. 435-443, 2002.
- [34] B. Sarkar, S. Roy, and D. Sarkar, "Hierarchical representation of digitized curves through dominant point detection," *Pattern Recognition Letters*, vol. 24, pp. 2869-2882, 2003.
- [35] M. Marji and P. Siy, "Polygonal representation of digital planar curves through dominant point detection - A nonparametric algorithm," *Pattern Recognition*, vol. 37, pp. 2113-2130, 2004.
- [36] D. K. Prasad, C. Quek, and M. K. H. Leung, "A non-heuristic dominant point detection based on suppression of break points," in *Image Analysis and Recognition*. vol. 7324, A. Campilho and M. Kamel, Eds., ed Aveiro, Portugal: Springer Berlin Heidelberg, 2012, pp. 269-276.
- [37] D. K. Prasad, M. K. H. Leung, C. Quek, and S.-Y. Cho, "A novel framework for making dominant point detection methods non-parametric," *Image and Vision Computing*, 2012.
- [38] G. Damiand and D. Coeurjolly, "A generic and parallel algorithm for 2D digital curve polygonal approximation," *Journal of Real-Time Image Processing*, vol. 6, pp. 145-157, 2011.
- [39] D. K. Prasad, C. Quek, M. K. H. Leung, and S. Y. Cho, "A parameter independent line fitting method," in *Asian Conference on Pattern Recognition (ACPR)*, Beijing, China, 2011, pp. 441-445.
- [40] D. K. Prasad and M. K. H. Leung, "Polygonal representation of digital curves," in *Digital Image Processing*, S. G. Stanciu, Ed., ed: InTech, 2012, pp. 71-90.
- [41] A. Rosenfeld, "Digital straight line segments," *IEEE Transactions on Computers*, vol. C-23, pp. 1264-1269, 1974.
- [42] D. K. Prasad, "Adaptive traffic signal control system with cloud computing based online learning," in *8th International Conference on Information, Communications, and Signal Processing (ICICS 2011)*, Singapore, 2011.
- [43] D. K. Prasad, R. K. Gupta, and M. K. H. Leung, "An Error Bounded Tangent Estimator for Digitized Elliptic Curves," in *Discrete Geometry for Computer Imagery*. vol. 6607, ed: Springer Berlin / Heidelberg, 2011, pp. 272-283.

Gray Coded Grayscale Image Steganography using Huffman Encoding

Nithyanandam Pandian[#]

*Assistant Professor Department of Computer Applications,
SSN College of Engineering,
Kanchipuram Dt. 603110, India*

nithyanandamp@ssn.edu.in

Ravichandran Thangavel

*Principal
Hindustan Institute of Technology,
Coimbatore 641032,India*

dr.t.ravichandran@gmail.com

Corresponding Author

Abstract

Steganography is an art of hiding secret information on a cover medium through imperceptible methodology. The cover medium can be any digital entity ranging from an image, audio, video to any object which can be digitally represented. Embedding capacity, imperceptibility and robustness are the primary goal of steganography. However the requirements of the above said goal vary from application to application. The notion of the proposed work is to focus on improving embedding capacity and bring down the distortion occurring to the stego image. For accomplishing higher embedding capacity and undetectability it is not essential to rely only on the cover image and efficient embedding technique but also on the refinement of the algorithm. One such attempt has been taken in the proposed work by applying gray coding technique to the payload (secret image). The gray coded binary data is bit planed and runlength encoded, prior to the submission of Huffman encoding technique results in supporting higher embedding capacity. The experimental results shows that the assessment metrics such as peak signal to noise ratio, histogram and embedding capacity are highly improved.

Keywords: Steganography, Graycoding, Huffman Encoding, Peak Signal to Noise ratio, distortion, redundancy.

1. INTRODUCTION

Steganography is a secret communication technique, in which the event of communication taking place itself is concealed. Cryptography and watermarking are the close related techniques to steganography. Fundamentally both cryptography and steganography are information securing technique but they differ in their implementation. Cryptography makes secret data unreadable by a third party, whereas steganography hides secret data from a third party. Both of their notions [1] remain the same. The cover medium suitable for a steganography [2] can be any entity that can be digitally represented such as a text file, image, audio, video and unused portion of TCP/IP packet headers. Steganography [3] and watermarking [3] are information hiding technique but the notion between them is different. Watermarking is claiming/proving the ownership of any digital entity, whereas steganography is a secret/hidden communication accomplished in a digital entity. The impact of embedding secret data on the cover medium should be lighter in order to achieve undetectability. The payload is relatively small in watermarking on converse with steganography. Because steganography is a secret communication technique, whereas watermarking is a authentication certificate made up of watermark merely take size of few bytes. In watermarking the watermarks are basically a metadata which help to prevent copyrights and other cyber related issues. Unlike watermarking, steganography does not have a constraint in choosing the cover medium to establish secret communication. Application of steganography is vast and in each

application the degree of imperceptibility, embedding capacity and robustness [2] does not remain same. For example one application may focus on higher imperceptibility compromising embedding capacity whereas in other case lesser importance to robustness but imperceptibility and embedding capacity should be the best. All the above said parameters are interdependent and it is a tradeoff [29] in controlling each other.

LSB embedding [4], optimum pixel adjustment process [5][6], pixel value differencing [7][8][9] and matrix embedding [1][3][4][10] were some of the spatial domain embedding technique. All the spatial domain techniques are flexible to support fixed bits, variable bits, and adaptive embedding on their pixel intensities. F5 [3], Outguess[3], JP Hide and Seek[11] were some of the frequency domain technique. The limitation of frequency domain is its inflexibility of supporting [12] low payload to embed. Therefore ideally, spatial domain embedding technique [17] is preferred over frequency domain which helps to attain higher embedding capacity.

Supporting a higher payload on a cover image primarily depends upon the embedding mechanism (algorithm). But it also can be viewed in another direction of compressing the payload prior to embedding on the cover image. With respect to an image, there exists several redundancies such as [13] coding redundancy, interpixel redundancy and psychovisual redundancy. The intensity values in the two dimensional matrix of an image are represented by n bits, which is called as bit depth of an image. Bit depth vary from 2 to 32 bits depends upon the support offered by the hardware and software. Among the redundancy, coding redundancy [13] is the one which can be confidently used to attain error-free or lossless compression. Huffman encoding [14] is a variable length lossless compression technique, can be applied to any entity which can be digitally represented. Several coding techniques can be used in conjunction with Huffman encoding to compress the payload. Integrating all the above said entity under a single strand ends up with higher compression results in high embedding capacity of a steganography algorithm. This is discussed detail in background section.

The objective of the proposed work is to embed a grayscale secret image on a 24 bit RGB cover image using various coding technique. The gray coded binary data of secret image is bit planed and runlength encoded, prior it is submitted to Huffman encoding technique results in supporting higher embedding capacity and peak signal to noise ratio. The rest of the paper is organized as follows: Section 2 discusses the background study, Section 3 covers the related works. The proposed work is exhibited in Section 4, experimental results and discussions were given in Section 5. Finally, the concluding remark and future direction are given.

2. BACKGROUND STUDY

Different coding techniques involved in the proposed work should be studied individually helps to understand the role carried by each of them in the proposed work. Coding technique such as gray coding, bit plane coding, run length coding and Huffman coding are discussed in detail which helps in understanding the attainment of high embedding capacity and less distortion.

2.1 Run Length Coding

Run length coding [15] is a simple form of data compression. Image intensity is represented by a 8 bit binary magnitude in a gray scale image. It is obvious that alike pixel intensities are get repeated across the rows and columns of the image (interpixel redundancy [13]). It is unwise to code every repeated intensity with same 8 bit binary magnitude. Instead of coding all the repeated intensities (values) with fixed bits just encode the binary runs of each consecutive 0's and 1's. For example consider 129,129,129,129,129 are the five consecutive gray scale pixel intensities. Their 8 bit binary magnitude representation (10000001 10000001 10000001 10000001 10000001) sums to 40 bits. If the same is run length encoded it can be represented as follows: 1,6,2,6,2,6,2,6,2,6,1. Here the runs refer to consecutive number of 0's and 1's. In a binary model, 1 can be represented in single bit and 2 can be represented in 2 bit and 6 can be represented in 3 bits. Therefore two 1's, four 2's and five 6's ($2 \times 1 + 4 \times 2 + 5 \times 3$) can be represented using 25 bits. The bit rate is $25/5=5$ bits per pixel in lieu of 8 bits per pixel accounts

compression on various information coding method for the 4 bit depth 10 x 10 image is presented in Table 2.

Image of 4 bit depth Method	Huffman encoding (size in bits)	Huffman table (size in bits)	Huffman compression (size in bits)
Run length	312	89	401
Bit Plane and Run length	34	161	195
Gray code ,Bit Plane and Run length	22	142	164

TABLE 2: RESULTS OF HUFFMAN COMPRESSION ON VARIOUS CODING METHODS

3. RELATED WORKS

A stenographic technique proposed in [18] which is based on LSB replacement technique. Varying lengths of secret bits get embedded in every pixel. In method1 green and blue are embedding channels keeping red channel as indicator channel. In method2 an option is provided for choosing the indicator channel among the three channels. Once chosen, the remaining two channel act as embedding channel. In method3 the indicator channel is chosen by rotation scheme across all the pixels. In the first pixel red channel is indicator; green channel is the indicator in second pixel and in third channel blue act as indicator. Once indicator is finalized the remaining two channels will be used for embedding. This scheme is repeated for the consecutive pixels. The Mean Square Error (MSE) and PSNR is calculated for all channel and the average number of bits get embedded in every pixel is shown in their results.

A stenographic technique proposed in [7] is based on edge adaptive scheme. The absolute difference between two adjacent pixels is the primary criteria in identifying the region for embedding secret message. They use LSBMR (Least Significant Bit Matching Revisited) as their data hiding algorithm. Only 2 secret bits can be embedded in each embedding unit and threshold T is used in identifying the embedding region. Region selection and cover image hiding capacity is determined through trial and error process. The sharper edge regions of cover image alone are used for embedding. Even though the embedding capacity is lesser it withstand against statistical attack and they had proved that RS steganalysis is ineffective in detecting stego work.

In the steganography scheme adopted in [19], the embedding efficiency is improved by adopting matrix embedding method. ME-RA (matrix embedding repeat accumulate) is the data hiding algorithm used to hide the secret data. The reason to choose matrix embedding is to less adulterate the cover image, at the same time the secret data bits should get embedded. Here, a hamming code matrix is employed in attaining the goal. In the proposed work, instead of hamming code (for matrix embedding) a simple XOR operation is performed on the host image bits to check its coincidence against the secret bits. The host image bit is adjusted accordingly to suit the secret bits.

A novel image steganography technique [20] was discussed in which the cover image's spatial value is transformed in to Discrete Cosine Transformation (DCT); its LSB is modified to match the secret message. The secret message is Huffman encoded prior to the embedding scheme which achieves a significant compression rate. A higher embedding capacity and PSNR is obtained using this technique. This technique is superior to the method proposed in [23].

A stenographic technique [21] based on wavelet transformation on the images is proposed. Discrete Wavelet Transformation (DWT) converts the spatial domain of cover image into frequency domain. Huffman compression is applied for the stream of secret bits before

overlaying them on the cover image. A high PSNR and very high embedding capacity is achieved. A higher level of security is obtained because the Huffman table and encoding rules is black box to the intruder.

A Least Significant Bit [22] steganographic scheme technique in which the secret image is bit plane coded, runlength encoded and finally Huffman compressed yielded a high embedding capacity. The payload (secret image) get embedded is just 4 bits out of every 8 bit pixels. Remaining 4 bits of every pixel is artificially constructed in the destination during the process of making the secret image from the stego image. The secret image retrieved in the destination is a lossy one, but sill PSNR above 30dB is acceptable.

4. PROPOSED METHOD

Applying coding technique on the payload will compress its size and further Huffman encoding reduce the coding redundancy present in an image. This is the central theme behind the proposed work. In the proposed method of [22], coding technique such as bitplane,runlength and Huffman encoding were applied. Gray coding can be attempted in addition prior to all the above mentioned coding technique which will improve the performance. This is the central theme behind the proposed work.

4.1 Least Significant Embedding

Least Bit (LSB) embedding [4][6][22][24][25][26][30] is one of the renowned spatial domain steganography techniques. The process of LSB embedding carried on a cover medium is explained below. Suppose we want to encode the letter A (**ASCII 65 or binary 01000001**) in the following 8 bytes (pixel intensities) of an image cover medium, it can be done as follows:

93	208	28	172	231	135	107	227
01011101	11010000	00011100	10101100	11100111	10000111	01101011	11100011

becomes

92	209	28	172	230	134	106	227
0101110 <u>0</u>	1101000 <u>1</u>	0001110 <u>0</u>	1010110 <u>0</u>	1110011 <u>0</u>	1000011 <u>0</u>	0110101 <u>0</u>	1110001 <u>1</u>

4.2 Matrix Embedding

In matrix embedding technique, the LSB’s of all the channels of every pixels are taken to embed two bits of secret image.

4.2.1 Embedding

To encode bit b_1 and b_2 in the LSB of three channels of a pixel: say X_1 , X_2 and X_3 . Equation3 must be satisfied.

$$\begin{aligned}
 b_1 &= LSB(x_1) \oplus LSB(x_2) \\
 b_2 &= LSB(x_2) \oplus LSB(x_3) \dots \dots \dots (3)
 \end{aligned}$$

If equation3 is not satisfied, a minor modification will be done on X_1 , or X_2 or X_3 to satisfy the same. If the first one is satisfied but not the second one, simply flip the LSB of x_3 . If the second one is satisfied but not the first one, flip the LSB of x_1 . If neither is satisfied, flip LSB of x_2 .

4.2.2 Extraction

Every consecutive 3 bytes of stego image are extracted to construct 2 bits of secret image. This process is repeated until the required secret image bits are constructed. To extract bit b_1 and b_2

from LSB of three channels of a pixel: say X_1 , X_2 and X_3 equation 4 is applied on the stego image pixels.

$$\begin{aligned} LSB(x_1) \oplus LSB(x_2) &= b_1 \\ LSB(x_2) \oplus LSB(x_3) &= b_2 \end{aligned} \quad \dots \dots \dots (4)$$

These collected bit streams are grouped and submitted for further operation of constructing the secret image.

4.3 Embedding Process

Fig.1 shows the embedding process carried on the sender side. After embedding the sender will send the cover image which has secret image embedded in it. The steps carried on the sender side are given below.

- Step1: From each pixel of the secret image, remove the last four LSB's. Apply gray coding followed by bit plane slicing in the remaining four MSB's.
- Step2: Apply run length encoding for the output obtained from Step1.
- Step3: Apply Huffman encoding for the output obtained from Step2 which results in Huffman table and Huffman encoded secret image bit streams.
- Step4: Embed dimensions of secret image and the resultant component (Huffman table and Huffman encoded secret image bit streams) obtained from Step3 into the cover image using matrix/LSB embedding technique.
- Step5: Send the stego image obtained from Step5 to the receiver.

4.4 Extraction Process

Fig.2 shows the extraction process carried on the receiver side. Upon receiving the stego image, the receiver should extract the Huffman table, Huffman encoded bit streams, the four LSBs equivalent decimal value (which was computed during embedding on the sender side) and secret image dimension from the stego image. The steps carried on the receiver side are given below.

- Step 1: For retrieving the secret image from the stego image, extract the LSB's from every pixel, apply matrix/LSB extraction technique and construct Huffman table and Huffman encodings using the LSB's bit streams.
- Step 2: The Huffman table and Huffman encodings obtained in Step1 is used in Huffman decoding process. As a result a run length pattern is produced.
- Step 3: Apply run length decoding to the result obtained in Step2 which results in producing corresponding binary runs consist of 0's and 1's.
- Step 4: Finally, the image is bit plane, gray decoded and later the extracted decimal value of LSB four bits are converted to binary and appended to the existing 4 MSB's to make an 8 bit gray scale value. At last the image is constructed using all the pixels which will reveal the secret image.

The four bit decimal value which is substituted during secret image construction process in the receiver side is, precomputed and stored in the cover image by the sender. This precomputation is explained in detail. Three types of decimal value are computed for the 4 bit LSB, such as maximum value, minimum value and random value and any one among them is get substituted.

4.4.1 Maximum value: The 4 LSB's of all the pixels are separated and its decimal value is determined. The decimal value range will be from 0- 15, since it is just 4 bits. The individual count of every decimal value is identified and the decimal value which has higher count become the substitution value for the 4 LSB in the destination side. This value is stored in the stego image in addition to its size,Huffman encodings and Huffman table.

4.4.2 Minimum value: The individual count of every decimal value is identified and the decimal value which has lower count become the substitution value for the 4 LSB in the destination side.

4.4.3 Random value: A random value ranges from 0 -15 is computed for every 4 MSB and this random 4 bit value is padded to get 8 bit gray pixel value.

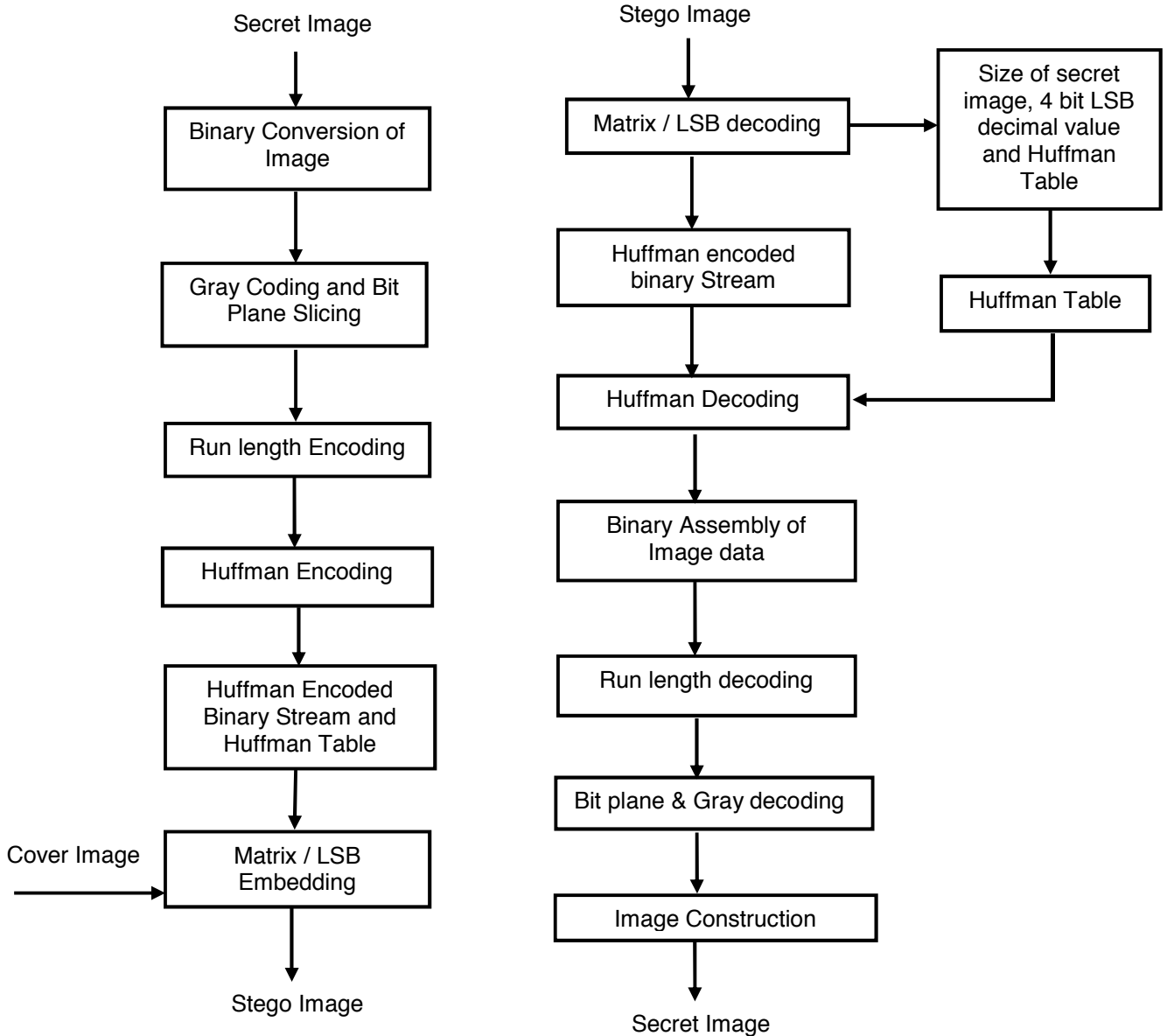


Figure 1: Embedding Process

Figure 2: Extraction Process

$$MSE = \frac{1}{m * n} \sum_{i=1}^m \sum_{j=1}^n (A_{ij} - B_{ij})^2 \quad \dots \dots \dots (5)$$

All the above three substitution mechanism has attempted and their result (PSNR) while constructing the secret image is presented in Table5. The rate of adulteration between the cover matrix and stego matrix should be measured. To assess this, mean square error and Peak Signal to Noise ratio between the matrixes should be determined. The equation for MSE and PSNR are followed:

where A_{ij} represents pixel in the cover image and B_{ij} represents pixel in the stego image; m , n represents the height and width of the image respectively.

$$PSNR = 10 * \log_{10} \left(\frac{Max^2}{MSE} \right) \dots \dots \dots (6)$$

here max denote maximum color intensity of grayscale (255).PSNR is measured in decibels (dB).

5. EXPERIMENTAL RESULTS

Java 2.0 and MATLAB R2009a are the programming tools used to implement the proposed method. Standard 24 bit cover images of size 512 x 512 such as lena, baboon, flight and boat were used. Fig. 3a-3c shows the secret image and Fig 4a-4d shows the cover images used in the experiments. Fig.5a-5d shows the stego image obtained for LSB embedding. Fig. 6a-6d and 7a-7d shows the histogram of the cover and stego image respectively for LSB embedding technique. Similarly Fig. 8a-8d and 9a-9d shows the histogram of the cover and stego image respectively for Matrix embedding technique. PSNR and embedding Capacity are the metrics taken here to consolidate the strength of proposed method. Histogram of the stego image shows that the distortion between cover and stego image is minimum. Gray coding is used in the proposed to improve the performance of Huffman coding. The performance of Huffman encoding depends upon unique symbol present in the secret image. Here symbol refers to various unique pixel intensity or grayscale value. Table3 shows the experimental results compared against [21],[22] and [27]. There is a significant improvement in PSNR and embedding capacity through the proposed method. The improvement obtained for LSB and matrix embedding is exhibited in both Table3 and Table4. The maximum embedding capacity of [22] is experimented in the proposed method and the revised result using the proposed method is given in Table4. The maximum embedding capacity of the proposed method given in Table4. The secret image constructed in the destination is not as exact as the original secret image. The 4 LSB decimal value required for every pixel of secret image are unanimously replaced by either one of the following three entities: a maximum value, minimum value or random value which is explained in detail section 4.4. The quality of the secret image which is retrieved from the stego image is assessed. The experimental result is exhibited for various secret images in Table5. It is inferred from Table5 that PSNR of secret image is influenced by the substitution value get replaced while constructing the secret image in the destination. For Fig.3a, random value substitution gives good result whereas in the case of Fig.3b it is maximum and finally for Fig.3c it is minimum value. PSNR above 30dB is acceptable [28] but high quality recovering (secret) image should strive for 40dB and above.

The last 64 pixel in cover image is reserved for storing the technical details, which will be used in the receiver side to extract the secret data from the stego image. This 64 pixel (64x3=192 bytes) should be excluded while computing the maximum hiding capacity of cover image. Secret data of more unique symbol with any size can be hidden through our proposed method, provided it meets the above said condition.

5.1 Discussion

In the proposed method, gray coding is done prior to other coding technique which has enhanced the embedding capacity when comparing [22]. The reason is due to the principle that any two adjacent [13][16] decimal value when represented in binary pattern may have a huge difference between each other in their patterns. In the case of gray coding the binary pattern between any two adjacent gray coding value there will be a maximum of 1 bit change between each other. This nature is utilized to get a high throughput.

We quite often found that a secret image which is richer and whose dimension is lesser than Cameraman 256 X 248 shown in Fig. 3 (image used as secret image in the proposed method) cannot be embedded in this 512 x 512 cover image Lena. In contrast, a secret image which is not richer whose dimension is higher than cameraman 256 X 248 can be embedded in the cover image. This makes us to finalize that the embedding capacity of our proposed technique depends on Huffman encoding. Any image, whose Huffman compression is less, fits in the cover image irrespective of its size and richness. To discuss on security side, the proposed technique is robust enough; because extracting a data without knowing the architecture of the proposed technique is

difficult, moreover data is Huffman encoded. Based on how the symbols are treated, the implementation of Huffman may vary. But however the basic rule that should be adhered is that Huffman implementation should have prefix free coding. Even if the intruder collects the LSBs from the cover image, the intruder should separate the Huffman table and actual Huffman encodings. And also, only four bits are transmitted, remaining four bits are constructed. Working-out this idea is another level of difficulty. Moreover, the Huffman output should be run length decoded, biplane decoded and finally gray decoded to construct the actual image. It is cumbersome to attack all the stages of our proposed algorithm.

Cover Image 512×512	Capacity (bits)	PSNR(dB)							
		DWT base [9]	DWTand Huffman Method [11]	Bitplane, Runlenth and Huffman Method [22]		Proposed Method			
				LSB	Matrix	LSB	Improved PSNR	Matrix	Improved PSNR
Lena	507856	46.08	54.93	53.54	54.79	54.14	0.60	55.39	0.60
Airplane	507856	45.99	54.67	53.54	54.77	54.14	0.60	55.37	0.60
Baboon	507670	46.19	55.11	53.52	54.78	54.13	0.61	55.40	0.62
Boat	507867	46.13	54.80	53.51	54.77	54.15	0.64	55.40	0.63

TABLE 3: Comparison of PSNR results for the Proposed Method against [9] , [11] and [22]

Cover Image 512×512	Capacity (bits)	PSNR(dB)					
		Bitplane, Runlenth and Huffman method [22]		Proposed Method			
		LSB	Matrix	LSB	Improved PSNR	Matrix	Improved PSNR
Barbora 270 x270	583200 (71 KB)	52.96	54.22	54.03	1.07	55.26	1.04
Barbora 380 x 380	1155200 (141 KB)	---		52.94	-----	54.19	----

TABLE 4: PSNR results for Maximum Embedding Capacity for Barbora in LSB and Matrix embedding

Methods	Minimum	Maximum	Random
Obtained Secret Images' PSNR			
256 x 248 Cameraman	33.20	31.79	33.20
270 x 270 Elaine	32.80	29.20	31.79
380 x 380 Barbora	30.07	33.89	30.07

TABLE 5: Different Secret Images' PSNR, embedded on Lena cover image of size 512 x 512

CONCLUSION

We had proposed an image steganography algorithm which brings a better PSNR than [22]. Histogram of stego image and cover image are almost equal which emphasize on the result that distortion between cover and stego image is minimum. Capacity improvement and distortion reduction has been addressed in this proposed technique. In the proposed work, the embedding capacity of the cover image is increased, at the same time the PSNR also controlled. The proposed technique is not robust against any geometrical distortion such as rotation, translation, scaling, cropping etc., induced on the stego image.

FUTURE WORK

The secret images which is constructed at the destination is resembling to be exactly the original image sent from sender side; but still it is lossy picture generated by constructing the 4 LSB (max,min and random) and 4 MSB received. The proposed work should be refined such that the PSNR of secret image should go beyond 40dB. The proposed algorithm should be customized to support embedding in the frequency domain. It should be enhanced to support color images and withstand geometrical distortion induced on the image.



Secret Images

FIGURE 3a: Cameraman

FIGURE 3b: Barbora

FIGURE 3c: Elaine



Cover Images

FIGURE 4a: Lena

FIGURE 4b: Baboon

FIGURE 4c: Flight

FIGURE 4d: Boat



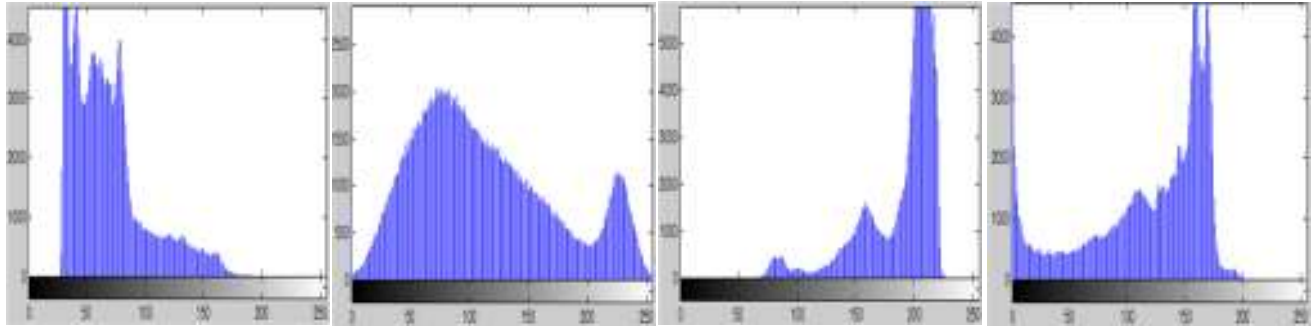
Stego Images: Least Significant Embedding

FIGURE 5a: Lena

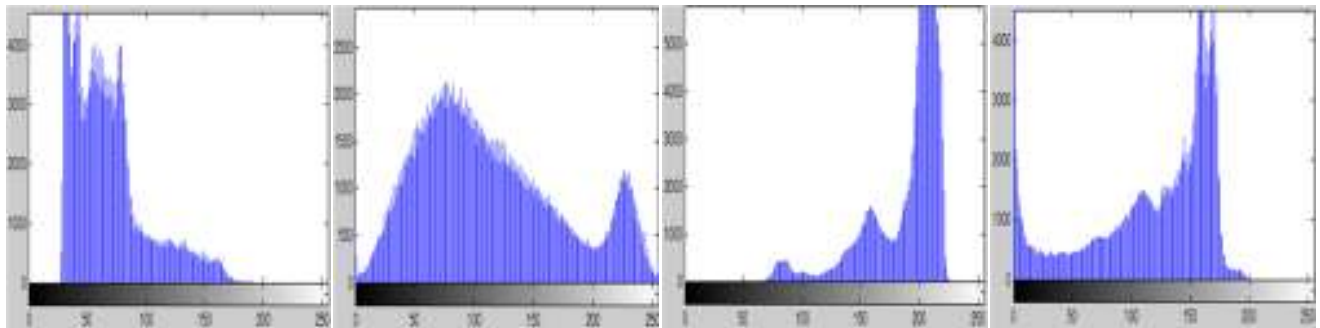
FIGURE 5b: Baboon

FIGURE 5c: Flight

FIGURE 5d: Boat



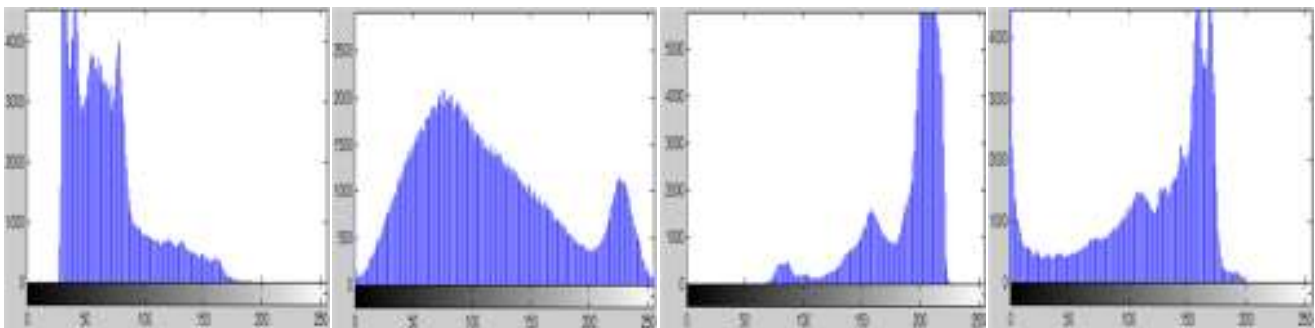
Histogram Images: Cover Images
FIGURE 6a: Lena **FIGURE 6b:** Baboon **FIGURE 6c:** Flight **FIGURE 6d:** Boat



Histogram Images: Least Significant Bit Embedding
FIGURE 7a: Lena **FIGURE 7b:** Baboon **FIGURE 7c:** Flight **FIGURE 7d:** Boat



Stego Images: Matrix Embedding
FIGURE 8a: Lena **FIGURE 8b:** Baboon **FIGURE 8c:** Flight **FIGURE 8d:** Boat



Histogram Images: Matrix Embedding
FIGURE 9a: Lena **FIGURE 9b:** Baboon **FIGURE 9c:** Flight **FIGURE 9d:** Boat

REFERENCES

- [1] Mao,J-F. Ru,Z.,Niu,X-X.,Yang,Y-X.,and Zhou,L-N "Research of Spatial Domain Image Digital Watermarking Payload", *EURASIP Journal on Information Security*, pp.1-12, Mar. 2011,DOI:10.1155/2011/502748.
- [2] Tayana Morkel, Jan H P Eloff and Martin S Olivier "An Overview of Image te ganography", *Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005)*, Sandton, South Africa, Jul. 2005.
- [3] Injemar J. Cox,Matthew L. Miller, Jeffrey A. Bloom, Jessica Fridrich and Ton Kalker. "*Digital Watermarking and Steganography*", Burlington, MA 01803, USA: Morgan Kaufmann, Second Edition,2008.
- [4] Sutaone, M. S., and Khandare, M.V. "Image Based Steganography Using LSB Insertion Technique" in *IET International Conference on Wireless, Mobile and Multimedia Networks*(Beijing,China) IET,2008, pp.146-151.
- [5] Amirtharajan, R., Behera, S.K., Swarup, M.A., Ashfaaq, M.K. and Rayappan, J.B.B. "Colour guided colour image steganography", *Universal Journal of Computer Science and Engineering Technology*, 1(1),16–23,Oct.2010.
- [6] Chan,C-K., and Cheng, L.M. 2004. "Hiding data in images by simple LSB substitution",*Pattern Recogn.* 37(3),pp.469 – 474.DOI:10.1016/j.patcog.2003.08.007.
- [7] Luo,W.,Huang,F.,and Huang,J. "Edge Adaptive Image Steganography Based on LSB Matching Revisited" *IEEE T INF FOREN SEC*, 5(2), pp.201-214,2010, DOI:10.1109/TIFS.2010.2041812.
- [8] Wu,D-C.,and Tsai,W-H. "A steganographic method for images by pixel-value differencing".*Pattern Recogn.Lett.* 24(9),pp.1613–1626,2003. DOI:10.1016/S0167-8655(02)00402-6.
- [9] Wu,H. C.,Wu,N.I.,Tsai,C.S.,and Hwang,M.S. "Image steganographic scheme based on pixel-value differencing and LSB replacement methods". *IEE P-VIS IMAGE SIGN*, 152(5),pp.611–615,2005. DOI:10.1049/ip-vis:20059022.
- [10] Fridrich,J.,and Soukal, D. "Matrix Embedding for Large Payloads", *IEEE T INF FOREN SEC*,1(3),pp.390-395,2006. DOI: 10.1109/TIFS.2006.879281.
- [11] Kipper,G., "*Investigator's Guide to Steganography*", Virginia:Auerbach Publications,2003.
- [12] Kharrazil,M.,Sencar,H.T.,and Memon,N. 2004. "WSPC Lecture Notes Series:ImageSteganography:concepts and practice", [On-line] <http://iwearshorts.com/Mike/uploads/2011/06/10.1.1.62.8194.pdf> [Jun. 1, 2010]
- [13] Gonzalez, R.C. and Woods, R.E. "*Digital Image Processing*",3rd Edition , New Delhi, India:PHI, 2008.
- [14] David A. Huffman "A method for the construction of minimum-redundancy codes". *Proceedings of the Institute of Radio Engineers*, 40(9):1098–1101, Sep. 1952.
- [15] Data Compression Project. [Online] Available at <http://www.binaryessence.com> [Feb. 20, 2011]
- [16] Darko Dimitrov, Tom´aš Dvořák, Petr Gregor, and Risteř Skrekovski , "Gray Code Compression" *Combinatorial Algorithms Lecture Notes in Computer Science, Springer, Volume 5874, pp 183-193 2009.*
- [17] Anil K. Maini ,"*Digital Electronics, Principal, Devices and Applications*" , Hoboken, New Jersey: John wiley & Sons, 2007.
- [18] Amirtharajan R, Sandeep Kumar Beher, Motamarri Abhilash Swarup, Mohamed Ashfaaq K and John Bosco Balaguru Rayappan. "Colour Guided Colour Image Steganography", *Universal Journal of Computer Science and Engineering Technology*, Volume 1, pp.16 – 23, Oct . 2010.
- [19] Sarkar,A.,Madhow,U.,and Manjunath B.S. "Matrix Embedding With Pseudorandom Coefficient Selection and Error Correction for Robust and Secure Steganography",*IEEE T INF FOREN SEC*,5(2),pp.225-239,2010,DOI:10.1109/TIFS.2010.2046218.

- [20] Amitava Nag, S. Biswas, D. Sarkar, P.P. Sarkar. "A Novel Technique for Image Steganography Based on Block-DCT and Huffman Encoding" , *International Journal of Computer Science and Information Technology*, Volume 2, Number 3, pp. 103-112,June 2010.
- [21] Amitava Nag, S. Biswas, D. Sarkar, P.P. Sarkar. "A Novel Technique for Image Steganography Based on DWT and Huffman Encoding" , *International Journal of Computer Science and Security*, Volume 4, Issue 6, pp. 561-570, Feb. 2011.
- [22] P. Nithyanandam , T. Ravichandran, E. Priyadarshini and N.M. Santron."A Image Steganography Technique on Spatial Domain Using Matrix and LSB Embedding based on Huffman Encoding", *Journal of Future Engineering and Technology*, vol 6, no.3, Feb 2011.
- [23] Chang, C.C and Tseng, H.W. "A Steganographic method for digital images using side match",*Pattern Recognition Letters*, 25: pp.1431 – 1437, Jun. 2004.
- [24] Nan-I Wu and Min-Shiang Hwang, "Data Hiding: Current Status and Key Issues", *International Journal of Network Security*, Vol.4, No.1, PP.1–9, Jan. 2007
- [25] Nithyanandam, P., Ravichandran, T.,Priyadarshini, E. and Santron, N.M. (2012) "An image steganography for colour images using lossless compression technique", *Int. J. Computational Science and Engineering*, Vol. 7, No. 3,pp.194–205.
- [26] P.Nithyanandam, T.Ravichandran, N.M.Santron and E.Priyadarshini.(2011) 'A Spatial Domain Image Steganography Technique based on Matrix Embedding and Huffman Encoding' ,*International Journal of Computer Science and Security*, Volume 5, Issue 5,pp.456-468, Dec.2011.
- [27] Chen, P.Y. and Wu, W.E. "A DWT Based Approach for Image Steganography", *International Journal of Applied Science and Engineering*, Volume 4, Issue 3: pp. 275 – 290, 2006.
- [28] Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt "Digital Image steganography:Survey and analysis of current methods" *Signal Processing 90* pp. 727-752, 2010.
- [29] Abbas Cheddad. "Steganoflage: A New Image Steganography Algorithm",Ph.D. Thesis, School of Computing & Intelligent Systems Faculty of Computing & Engineering University of Ulster, United Kingdom,2009.
- [30] Nithyanandam, P., Ravichandran, T. (2012) "Geometrical Rotation Addressed Steganograhly Technique using Matrix Embedding and Huffman Encoding", *European Journal of Scientific Research*, Vol. 87, No. 1,pp.31–45.

Walsh, Sine, Haar& Cosine Transform With Various Color Spaces for 'Color to Gray and Back'

Dr. H. B. Kekre

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

hbkekre@yahoo.com

Dr. Sudeep D. Thepade

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

sudeepthepade@gmail.com

Ratnesh N. Chaturvedi

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

ratneshnc@gmail.com

Saurabh Gupta

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

saurabh.gupta761@gmail.com

Abstract

The paper shows performance comparison of various color spaces with Image transforms alias Walsh, Cosine, Sine&Haar for 'Color to Gray and Back'. The color information of the image is embedded into its gray scale version/equivalent [1][2][3]. Instead of using the original color image for storage and transmission, matted gray (Gray scale version with embedded color information) can be used, resulting into better bandwidth or storage utilization. Total twenty-eight variations of the algorithm for 'Color to Gray and Back' are proposed and evaluated for qualitative performance using four image transforms and seven color spaces (RGB, YCbCr, YCgCb, YUV, YIQ, XYZ and Kekre's LUV) [4]. Among all considered image transforms and color spaces, Discreet Cosine Transform (DCT) gives better performance with YCbCr color space in 'Color to gray and Back'.

Key Words: Color Embedding, Color-to-Gray Conversion, Transforms, Color Spaces, Compression.

1. INTRODUCTION

In recent years, there is increase in the size of databases because of color images. There is need to reduce the size of data. To reduce the size of color images, information from all individual color components (color planes) is embedded into a single plane by which matted gray image is obtained [2]. This also reduces the bandwidth required to transmit the image over the network. Matted gray image, which is obtained from color image, can be printed using a black-and-white printer or transmitted using a conventional fax machine. This matted gray image then can be used to retrieve its original color image.

In earlier researches, this has been done in YCbCr[2] and Kekre's-LUV [1] color spaces using Haar Wavelet Transform. Further, In [3] Haar, Kekre& Walsh Wavelet transforms are employed for 'Color to Gray and Back'.

Here the better performance of 'Color to Gray and Back' using Haar [5], Walsh[6], Discrete Cosine Transform(DCT) [7] and Discrete Sine Transform(DST) [8] in various color spaces like RGB, YCbCr, YCgCb, YUV, YIQ, XYZ and Kekre's LUV[4].

The paper is organized as follows. Section 2 describes various color spaces. Section 3 presents method to convert color-to-matted gray image. Section 4 presents method to recover color image. Section 5 describes experimental results and finally the concluding remark are given in section 6.

2. COLOR SPACES

Here along with RGB six other color space alias YCbCr, YCgCb, YUV, YIQ, XYZ and Kekre's LUV are also employed for 'Color-to-Gray and Back'.

2.1 Kekre's LUV Color Space(K-LUV)

Kekre's LUV color space [4] is special form of Kekre Transform, where L is luminance and U and V are chromaticity value of color image. RGB to LUV conversion matrix is given in equation 1.

$$\begin{pmatrix} L \\ U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{-----(1)}$$

The LUV to RGB conversion matrix is given in equation 2.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \bullet \begin{pmatrix} L / 3 \\ U / 6 \\ V / 2 \end{pmatrix} \text{-----(2)}$$

2.2 YCbCr Color Space

In YCbCr [4], Y is luminance and Cb and Cr are chromaticity value of color image. To get YCbCr components, convert RGB to YCbCr components. The RGB to YCbCr conversion matrix is given in equation 3.

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.1688 & -0.3312 & 0.5000 \\ 0.5000 & -0.4184 & -0.0816 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{-----(3)}$$

The YCbCr to RGB conversion matrix is given in equation 4.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & -0.0010 & 1.4020 \\ 1 & -0.3441 & -0.7140 \\ 1 & 1.7718 & 0.0010 \end{pmatrix} \bullet \begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} \text{-----(4)}$$

2.3 YUV Color Space

The YUV color model [4] is used in PAL, NTSC, and SECAM composition color video standard. Where Y is luminance and U and V are chromaticity value of color image. To get YUV

components, convert RGB to YUV components. The RGB to YUV conversion matrix is given in equation 5.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.144 \\ -0.14713 & -0.22472 & 0.436 \\ 0.615 & -0.51498 & 0.10001 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{----(5)}$$

The YUV to RGB conversion matrix is given in equation 6.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 0.7492 & -0.50901 & 1.1398 \\ 1.0836 & -0.22472 & -0.5876 \\ 0.97086 & 1.9729 & -0.000015 \end{pmatrix} \bullet \begin{pmatrix} Y \\ U \\ V \end{pmatrix} \text{----(6)}$$

2.4 YIQ Color Space

The YIQ color space [4][11] is derived from YUV color space and is optionally used by NTSC composite color video standard. The `I` stands for phase and `Q` for quadrature which is the modulation method used to transmit the color information. RGB to YIQ conversion matrix is given in equation 7.

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.144 \\ 0.595716 & -0.274453 & -0.32126 \\ 0.211456 & -0.522591 & 0.31135 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{----(7)}$$

The YIQ to RGB conversion matrix is given in equation 8.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0.9563 & 0.6210 \\ 1 & -0.2721 & -0.6474 \\ 1 & -1.107 & 1.7046 \end{pmatrix} \bullet \begin{pmatrix} Y \\ I \\ Q \end{pmatrix} \text{----(8)}$$

2.5 YCgCb Color Space

To get YCgCb [4][10] components, convert RGB to YCgCb components. The RGB to YCgCb conversion matrix is given in equation 9.

$$\begin{pmatrix} Y \\ Cg \\ Cb \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{-----(9)}$$

The YCgCb to RGB conversion matrix is given in equation 10.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 0 \\ 1 & 0 & -2 \end{pmatrix} \bullet \begin{pmatrix} Y / 3 \\ Cg / 3 \\ Cb / 3 \end{pmatrix} \text{-----(10)}$$

2.6 XYZ Color Space

The RGB to XYZ [4][9] conversion matrix is given in equation 11.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.71160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \bullet \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{-----(11)}$$

The XYZ to RGB conversion matrix is given in equation 12.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 3.240479 & -1.53750 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{pmatrix} \bullet \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \text{-----(12)}$$

3. CONVERSION OF COLOR TO MATTED GRAY IMAGE

The ‘Color to Gray and Back’ has two steps as Conversion of Color to Matted Gray Image with color embedding into gray image & Recovery of Color image back. Here the transform-based mapping method is elaborated as per the following steps.[1][2][3]

- I. Image to be converted to desired color space of size N x N i.e. K-LUV, YIQ, YUV, XYZ, YCbCr and YCgCb or kept in RGB.
- II. First color component is kept as it is and second & third color component are resized to N/2 x N/2.
- III. Transform i.e. DCT, DST, Haar or Walsh to be applied to all the components of image.
- IV. First component to be divided into four subbands: corresponding to the low pass [LL], vertical [LH], horizontal [HL], and diagonal [HH] subbands, respectively.
- V. LH to be replaced by second color component, HL to replace by third color component and HH by zero.
- VI. Inverse Transform to be applied to obtain Matted Gray image of size N x N.

4. Recovery of Color Image

One nice feature of the proposed embedding method is the ability to recover the color from the matted gray image (gray scale version with embedded color information).For that, reverse all steps in the color-to-gray mapping. [1][2][3]

- I. Transform to be applied on Matted Gray image of size N x N to obtain four subbands as LL, LH, HL and HH.
- II. Retrieve LL as first color component, LH as second color component and HL as third color component of size N/2 x N/2.
- III. Inverse Transform to be applied on all three color component.
- IV. All three color component are resized to N x N.
- V. All three color component are merged to obtain Recovered Color Image.
- VI. If not in RGB, convert recovered color image to RGB color space.

5. RESULTS & DISCUSSION

The quality of ‘Color to Gray and Back’ is measured using Mean Squared Error (MSE) of original color image with that of recovered color image, also the difference between original gray image and matted gray image (where color information is embedded) gives an important insight through user acceptance of the methodology. This is the experimental result taken on 10 different images of different category as shown in Fig 1. Fig 2 shows the conversion of sample original color image to its matted gray equivalent having colors embedded into it, and matted gray to recovered color image.



FIGURE 1: Test bed of Image used for experimentation.

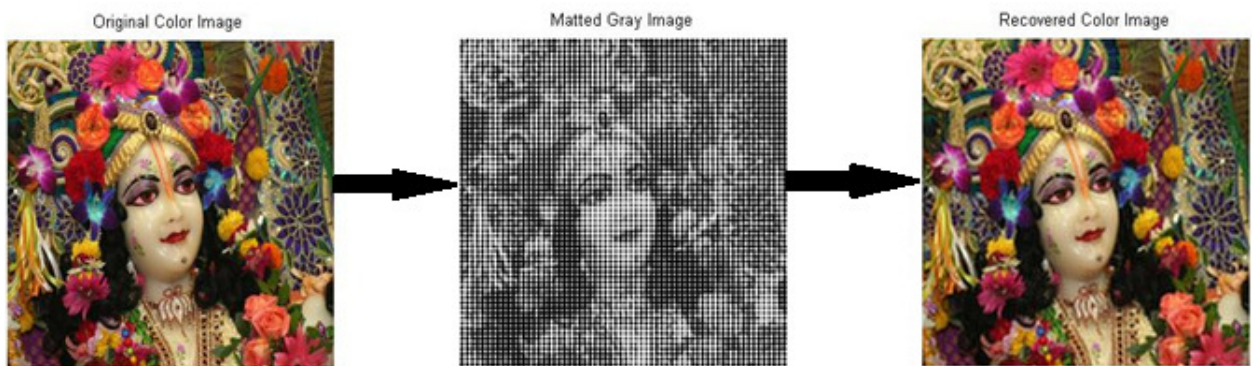


FIGURE 2: Conversion of Original Color to Matted Gray and Matted Gray to Recovered Color Image

It is observed in Fig 3 that YCbCr color space shows the least MSE between Original Color Image and the Recovered Color Image for the Transforms [DCT, DST, Haar, Walsh]. As observed in Fig 4, YUV and YIQ color spaces show the least MSE between Original Gray Image and the Matted Gray Image for the Transforms [DCT, DST, Haar, Walsh], indicating better performance as compared to other considered color spaces.

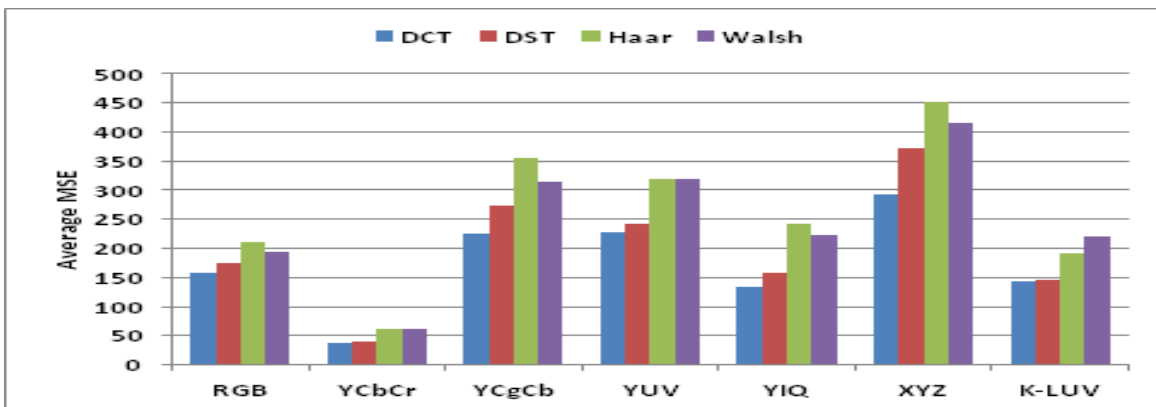


FIGURE 3: MSE between Original Color-Recovered Color Image

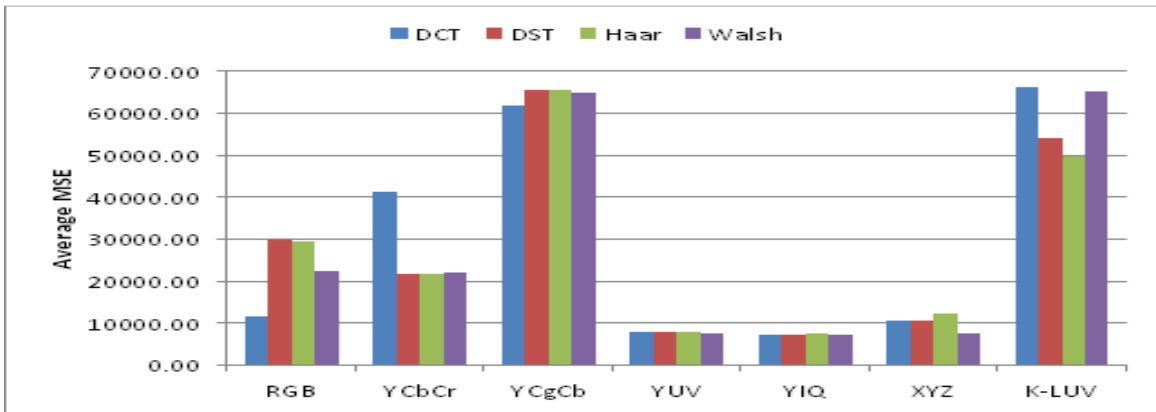


FIGURE 4 : MSE between Original Gray-Matted Gray Image

It is observed in Fig 5 that DCT gives least MSE between Original Color Image and the Recovered Color Image in any of the color space of the image. Among all considered image transforms, DCT gives best results. And in Fig 6 it is observed that Haar gives least MSE between Original Gray Image and the Matted Gray Image in any of the color space of the image. Among all considered image transforms, less distortion in Gray Scale image after information embedding is observed for Haar Transform.

The quality of the matted gray is not an issue, just the quality of the recovered color image matters. This can be observed that when DCT is applied on YCbCr color space the recovered color image is of best quality as compared to other image transforms and color spaces.

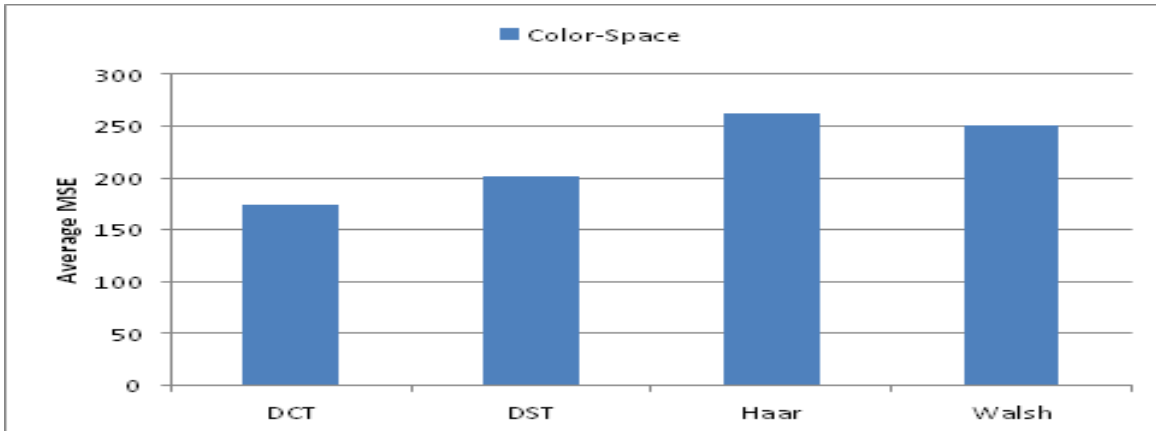


FIGURE 5 : Average MSE's of Original Color-Recovered Color Image of Color Spaces w.r.t Transform

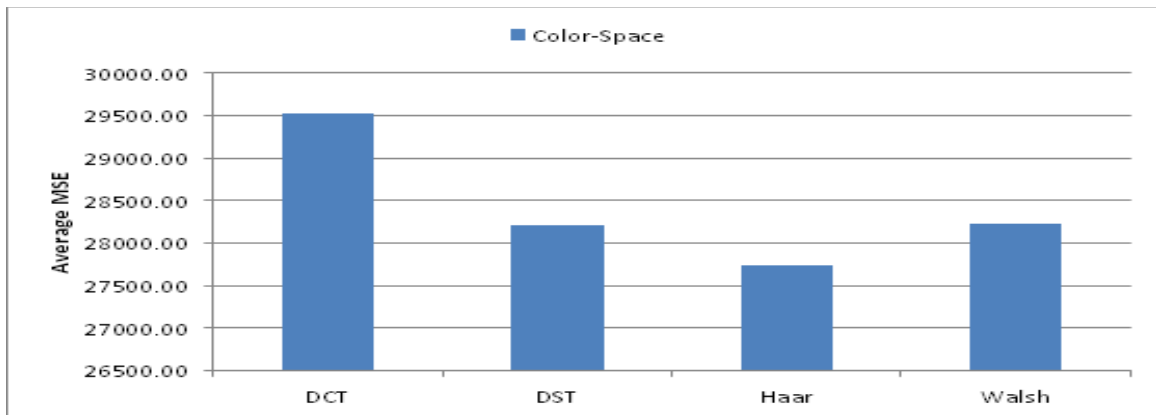


FIGURE 6 : Average MSE's of Original Gray-Matted Gray Image of Color Spaces w.r.t Transform

6. CONCLUSION

This paper have presented a method to convert image to matted gray with color information embedding into it and method of retrieving color information from matted gray image. The method allows one to send color images through regular black and white fax systems, by embedding the color information in a gray image. The method is based on transforms i.e DCT, DST, Haar, Walsh and color spaces alias RGB, YCbCr, YCgCb, YUV, YIQ, XYZ and Kekre's LUV. The YCbCr color space is proved to be better with DCT for 'Color-to-Gray and Back' Our next research step could be to test hybrid transforms, wavelet transforms and hybrid wavelets for 'Color-to-Gray and Back'.

7. REFERENCES

- [1] H.B. Kekre, Sudeep D. Thepade, "Improving `Color to Gray and Back` using Kekre's LUV Color Space." IEEE International Advance Computing Conference 2009, (IACC 2009),Thapar University, Patiala,pp 1218-1223.
- [2] Ricardo L. de Queiroz,Ricardo L. de Queiroz, Karen M. Braun, "Color to Gray and Back: Color Embedding into Textured Gray Images" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 15, NO. 6, JUNE 2006, pp 1464- 1470.
- [3] H.B. Kekre, Sudeep D. Thepade, AdibParkar, "An Extended Performance Comparison of Colour to Grey and Back using the Haar, Walsh, and Kekre Wavelet Transforms" International Journal of Advanced Computer Science and Applications (IJACSA) , 0(3), 2011, pp 92 - 99.
- [4] H.B. Kekre, Sudeep D. Thepade, VarunLodha, PoojaLuthra, Ajoy Joseph, ChitragadaNemani,"Augmentation of Block Truncation Coding basedImage Retrieval by using Even and Odd Images with Sundry Colour Spaces", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010, pp2535-2544.
- [5] Alfred Haar, "ZurTheorie der orthogonalenFunktionensysteme" (German), MathematischeAnnalen, Volume 69, No 3, pp 331 – 371, 1910.
- [6] J Walsh, "A closed set of normal orthogonal functions", American Journal of Mathematics, Volume 45, No 1, pp 5 – 24, 1923.
- [7] Dr. H. B. Kekre, DrTanuja K. Sarode, Sudeep D. Thepade, Ms.SonalShroff, "Instigation of Orthogonal Wavelet Transforms using Walsh, Cosine, Hartley, Kekre Transforms and their use in Image Compression", International Journal of Computer Science and Information Security, Volume 9, No 6,pp 125-133, 2011.

- [8] H.B. Kekre, Sudeep D. Thepade, Adib Parkar, "Performance Analysis of Kekre's Median Fast Search, Kekre's Centroid Fast Search and Exhaustive Search used for Colouring a Greyscale Image", International Journal of Computer Theory and Engineering (IJCTE), Volume 2, Number 4, August 2010, pp 507-513 Available at <http://www.ijcte.org/papers/193-H044.pdf>
- [9] H.B. Kekre, Sudeep D. Thepade, Adib Parkar, "A Comparison of Kekre's Fast Search and Exhaustive Search for various Grid Sizes used for Colouring a Greyscale Image", 2nd International Conference on Signal Acquisition and Processing (ICSAP 2010), IACSIT, Bangalore, pp. 53-57, 9-10 Feb 2010, The paper is uploaded on online IEEE Xplore.
- [10] H.B. Kekre, Sudeep D. Thepade, Karan Dhamejani, Sanchit Khandelwal, Adnan Azmi, "Performance Comparison of Assorted Color Spaces for Multilevel Block Truncation Coding based Face Recognition ", International Journal of Computer Science and Information Security (IJCSIS), Volume 10, Number 3, March 2012
- [11] H. B. Kekre, Sudeep Thepade, Rik Kamal Kumar Das, Saurav Ghosh, "Image Classification using Block Truncation Coding with Assorted Color Spaces", International Journal of Computer Applications (IJCA), Volume 44,, Number 6, April 2012. pp 9-14.

A Novel Method to Improve Measurement Results of Flame Photometry Using Image Change Detection

Balachandran Ganesan

*Jaya Prakash Narayan College of Engineering
Jawaharlal Nehru Technological University
Mahabubnagar, 509001, India*

balachandranganesan1@gmail.com

Sujeevan Kumar Agir

*Jaya Prakash Narayan College of Engineering
Jawaharlal Nehru Technological University
Mahabubnagar, 509001, India*

sujeevankumaragir@gmail.com

Abstract

The estimation of alkali metals in clinical chemistry has many parameters which can influence the result. A novel method to improve the measurement parameters of the Flame Photometer is presented. The improvement of accuracy and reliability is achieved through image processing Change Detection technique using Wiener Filter. The proposed method can be used in low cost Medical Equipment with improved measurement parameters performance.

Keywords: Flame Photometer, Image Processing, Change Detection, Wiener Filter, Medical Equipment.

1. INTRODUCTION

The measurement of the alkali metals (like Na, K) in blood samples is one of the important factors for the critically ill patient related to kidney function in Intensive Care Unit. The two major techniques used for these measurements are Flame Photometry and Atomic Absorption. Though the Flame Photometry method is comparatively economic, it suffers from lot of measurement errors. On the other hand, the Atomic Absorption method produces reliable results at significant higher cost. This work aims to provide a novel method which can improve the measurement values of the Flame Photometer with cost advantage. Though this paper focuses on medical application, it can be used in various other fields of science and technology such as environmental protection and electric energy production.

2. MEASUREMENT

Generally there are two methods exist in estimating the alkali metals using Flame Photometry:

- i) Optical sensor method and
- ii) Digital Camera method

2.1 Optical Sensor Method

Principles of flame photometer operations are based on the measurement of the radiating power of light emitted by atoms of known chemical elements, which are excited in a flame. The measurement of radiating power in the photometer goes through two optical sensors with interference filters, which have a maximum transmission for wavelengths of 589 nm for sodium and 768 nm for potassium. Burner is supplied with a mixture of air sample and propane-butane gas. This mixture is prepared in two-chamber nebulizer. Constant pressure, hot air sucks and atomizes the aqueous sample into the first chamber on a principle jet pump (ejector effect). So obtained aerosol is mixed with gas. [1]

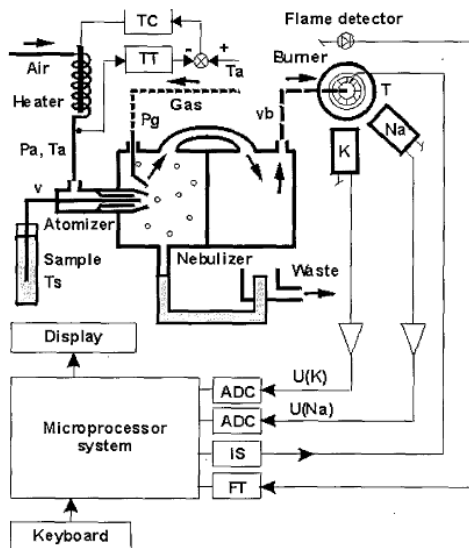


Figure1: Flame Photometer Block Diagram [3]

Next it flows intrinsically into the second chamber for additional mixing, continuing burning in the burner. Block diagram of apparatus is presented in **Figure1: Flame Photometer Block Diagram** [3]

2.2 Digital Camera method

Digital cameras have gained significant popularity in recent years. When sample is burnt in flame, colour of flame changes with respect to the concentration. So if images of the source are captured and post processing is carried out, it is possible to predict the amount of electrolytes in a flame by correlating it with colour values [3]. Anagha Panditrao et al have proposed a low cost non-contact temperature measurement technique using consumer grade digital still camera. The images of various visible heat sources are captured. Using colour image segmentation, source zones are identified. The colour temperature correlation is established by applying various analytical techniques [2].

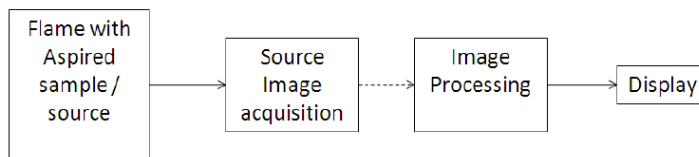


Figure2: Direct Reading Flame Photometer [3]

The basic block schematic of the proposed technique is shown in **Figure2: Direct Reading Flame Photometer** [3]

3. EARLIER WORKS

There are two important papers on Flame Photometry [1] and [3].

W. Holiczer et al introduced an error model of the system and chosen two erroneous situations: sucking of air bubbles and partial or full clogging liquid sucking pipe. However, it presented the event of perturbation “air bubble” (U_{pow}) which is more general than “clogging” and proposes an adaptive filter AMGM [1]. The estimator structure consists of the basic Kalman filter matched to the undisturbed model and bank of M filters with models adequate to the expected specific changes in successive time moments. The practical results of AMGM are compared against

various different methods like Direct Measurement (DM), Transversal Filter (TF100) and Kalman Filter (KF) [1]. It summarizes the AMG improves metrological features of measurement signal in stationary states of measurement system and it is immune to anomalous disturbances.

Anagha Panditrao et al proposed new technique to use digital camera and image processing and establish correlation between colour and concentration [3]. This technique improves the accuracy by the construction/process wherein there is no moving parts and minimal human interface with the total digital process. This technique has cost advantage too and has flexibility of continuous monitoring. Though this paper has image processing algorithm like Segmentation but it is used primarily for co-relation of the colour and concentration and doesn't address the improvement of the measurements results using image processing techniques.

In this paper, we propose a new method to improve the measurement results of [3]. In this novel method, the event of perturbation "air bubble" case is considered as disturbance and image processing technique to identify the same being evaluated.

4. PROPOSED METHOD

Detecting regions of change in images of the same scene taken at different times is of widespread interest due to a large number of applications in diverse disciplines [4]. Important applications of change detection include video surveillance, remote sensing medical diagnosis and treatment, civil infrastructure, underwater sensing and driver assistance systems.

The proposed method is to detect the change in frame caused due to bubble disturbance in sample flow path is shown in **Figure3**: Proposed Method to Improve Measurement Parameters. We have a set of images of the same scene taken at several different times. The method is to identify the set of pixels that are "significantly different" between the last image of the sequence and the previous images; these pixels comprise the change mask. For this paper case, it is changes in brightness or color.

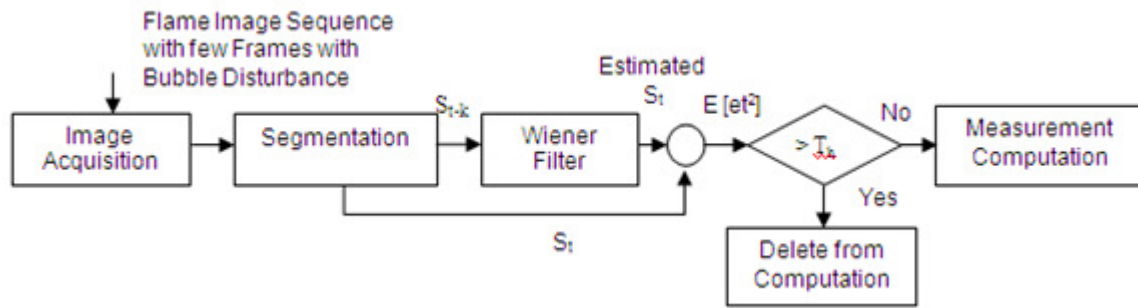


Figure3: Proposed Method to Improve Measurement Parameters

To make the change detection problem more precise, let $\{I_1, I_2, \dots, I_M\}$ be an image sequence. The acquired image is segmented based on the expected range of intensity values. The intensity of the pixels are computed on the segmented image frames.

The intensity of each pixels are calculated as per ITU-R BT.601 Standard

$$Y' = 0.299r + 0.587g + 0.114b$$

Where r, g, b are each pixel S_{rgb} co-ordinates values [5]. The average intensity of the each frame is then computed. The change detection is done using one step wiener prediction filter on the averaged intensity values of each frame.

This frame-level algorithm makes probabilistic predictions about what average frame intensity values are expected in the next live image using a one-step Wiener prediction filter. The Wiener filter is a linear predictor based on a recent history of values. Any frame intensity that deviates significantly from its predicted value is declared bubble frame. For a given frame, the linear prediction of its next value in time is

$$S_t = - \sum_{k=1}^p a_k S_{t-k}$$

where S_t is the predicted average intensity value of the image frame at frame t , the S_{t-k} is a past frame value, and the a_k are the prediction coefficients. The filter uses p past frame values to make its prediction. The expected squared prediction error,

$$E[e_t^2] = E[S_t^2] + \sum_{k=1}^p a_k E[S_t S_{t-k}]$$

The a_k are computed from the sample covariance values of the S_n . In this implementation, we used the past 60 values of S_n to compute $p=30$ prediction coefficients. The average intensity of the frame is calculated. If the actual value of the frame average intensity value differs by more than $4.0 \cdot \sqrt{E[e_t^2]}$ (Threshold T_h) from its predicted value [6], the frame is considered to be bubble error frame. The measurement is computed using only the frames which are not tagged as bubble error frames. The threshold will be chosen based on the allowable performance limits of the system.

5. PROCESSING OF REAL MEASUREMENT DATA

The experiment was conducted on 100/3 m Eq./l Sodium/Potassium solution and the video was captured using webcam at 15 frames/per second at resolution of 160 x 120pixels. The visibly different flames during no samples, with sample and Bubble disturbances are given in **Figure4: Captured Flame Details**. The captured video file (.avi format, approximately 400 frames) is passed through wiener filter algorithm after segmentation.



Figure4: Captured Flame Details

The algorithm separates the bubble frame and normal frame using change detection technique. If the difference between the actual values and wiener predicted values are more than threshold, either we can skip the frames from computation or use the last best values of the frames.

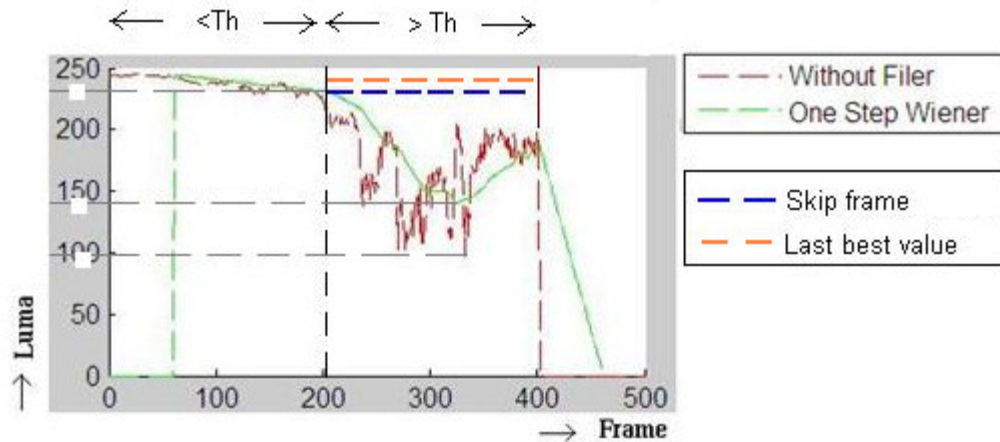


Figure5: System Response in Direct Digital Method

The four different results are compared to arrive at the best result to improve measurement parameter, variation in intensity of the photo flame meter. Refer **Figure5:** System Response in Direct Digital Method.

Filter/Method Type	*Error % (Variation)
Direct Measurement	> 58%
Wiener Filter	> 39%
Wiener Filter and Skip frames	< 5%
Wiener Filter and Last Best Value	< 3%

*255 values Luma steps, each 1% variation corresponds to 2.55 Luma

TABLE 1: Measurement Results

Basic statistic parameters of sodium and potassium concentration values in terms of Luma before and after disturbance of the bubble frames were presented in **TABLE 1:** Measurement Results

The following results are derived from the analysis of the results

- Wiener filter predicted values with skipping frames or using last best values provides better results.
- Variation in intensity level due to bubble frame can be filtered out by using image change detection technique

6. CONCLUSION

The Change Detection using one-step Wiener Prediction Algorithm to improve the measurement errors originated from instantaneous changes of image intensity due to bubble in sample flow path had been designed. The experimental observations of the MATLAB results are provided comparing the performance of with and without wiener filter. Based on threshold, the actual and predicted values are compared and decision to skip the frame or use the last best values is derived. The proposed new Digital Image Processing algorithm has significant potential to improve measurement results variations of low cost medical equipment. The implementation of the designed Wiener Filter makes it possible to eliminate or replace the bad frames in the measurements computation what is especially important in medical equipments. In future, image processing results are converted to concentration values and it will be compared against the Transversal Filer (TF), Kalman filter (KM), Adaptive filters of [1].

7. REFERENCES

- [1] W. Holiczer, Yu. Grishin, D. Janczak, E. Olech, "A New Data Processing Algorithm for Enhancing the Measurement Accuracy of Monovalent Alkaline Metals Microconcentration", *Proceedings of IEEE, vol.1, pp.63-66, 12-15, May 2004. MELECON 2004.*
- [2] Anagha M. Panditrao, Priti P. Rege, "Temperature Estimation of Heat Sources by Digital Photography and Image Processing", *IEEE Transaction on Instrumentation and Measurement vol.59, pp.1167-1174, 5-7, May 2010.*
- [3] Anagha M. Panditrao, "Direct Reading Flame Photometer Using Digital Photography and Image Processing", *2011 International Conference on Instrumentation, Communication, Information Technology and Biomedical Engineering 8-9 November 2011, Bandung, Indonesia*
- [4] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image Change Detection Algorithms: A Systematic Survey," *IEEE Trans. Image Processing, vol. 14, no. 3, pp. 294–303, March. 2005*
- [5] S.Bezryadin, P.Bourov*, D.Ilinih, "Brightness Calculation in Digital Image Processing", *KWE Int.Inc., San Francisco, CA, USA; *UniqueIC's, Saratov, Russia*
- [6] Toyama, Kentaro, Krumm, John, Brumitt, Barry, Meyers, Brian. "Wallflower: Principles and Practice of Background Maintenance," *IEEE Conference on Computer Vision, 1999.*

Color Image Watermarking using Cycle Spinning based Sharp Frequency Localized Contourlet Transform and Principal Component Analysis

K.Kishore Kumar

Assistant Professor & HOD, ECE
Faculty of Science & Technology,
Icfai Foundation of Higher Education University,
Hyderabad, 501203, India

kishorekamarajugadda@gmail.com

Movva Pavani

Associate Professor, ECE
Aurora's Technological & Research Institute
Hyderabad, 500039, India

pavanimovvap@gmail.com

V.Seshu Babu

Research Engineer,
UURMI Systems, Hyderabad. India

seshubabuv@uurmi.com

Abstract

This paper describes a new approach for color image watermarking using **Cycle Spinning based Sharp Frequency Localized Contourlet Transform and Principal Component Analysis**. The approach starts with decomposition of images into various subbands using Contourlet Transform(CT) successively for all the color spaces of both host and watermark images. Then principal components of middle band(x bands) are considered for inserting operation. The ordinary contourlet transform suffers from lack of frequency localization. The localization being the most important criterion for watermarking, the conventional CT is not very suitable for watermarking. This problem of CT is over come by Sharp Frequency Localized Contourlet, but this lacks of translation invariance. Hence the cycle spinning based sharp frequency localized contourlet chosen for watermarking. Embedding at middle level sub bands(x band) preserves the curve nature of edges in the host image hence less disturbance is observed when host and watermark images are compared. This result in very good Peak Signal to Noise Ratio (PSNR) instead of directly adding of mid frequency components of watermark and host images the principal components are only added. Likewise the amount of payload to be added is reduced hence host images get very less distortion. Usage of principal components also helps in fruitful extraction of watermark information from host image hence gives good correlation between input watermark and extracted one. This technique has shown a very high robustness under various intentional and non intentional attacks.

Keywords: Color Image Watermarking, Contourlet Transform (CT), Principal Component Analysis(PCA), Cycle Spinning, Frequency Localization.

1. INTRODUCTION

Currently, all multimedia production and distribution is digital. The advantages of digital media for creation, processing and distribution of productions are well known: easy modification and possibility of software processing rather than the more expensive hardware alternative. Maybe the most important advantage is the possibility of unlimited copying of digital data without any loss of quality. This latter advantage is not desirable at all to the media producers and content providers. In fact, it is perceived as a major threat, because it may cause them considerable financial loss. Digital watermarks have been proposed as a way to tackle this problem. This digital signature could discourage copyright violation, and may help determine the authenticity and ownership of an image. Digital watermarking [1] emerged as a solution for protecting the

multimedia data. Digital watermarking is the process of hiding or embedding an imperceptible signal (data) into the given signal (data). This imperceptible signal (data) is called watermark or metadata and the given signal (data) is called cover work.

Watermarking techniques can be broadly classified into two categories: Spatial and Transform domain methods. Spatial domain methods [2,3] are less complex and not robust against various attacks. Transform domain methods [4-6] are robust compared to spatial domain methods. In transform domain methods when image is inverse transformed, watermark is distributed irregularly over the image, making the attacker difficult to get the knowledge of presence of watermark. The localization is the most essential criterion to adapt a technique in watermarking. Wavelet is one of the transform domain technique that features localization. Barni et al. [7] proposed a wavelet domain based method which gives better invisibility with respect to (HVS). Based on the texture and the luminance content of all image sub bands, a mask is accomplished pixel by pixel. Dawei et al. [8] proposed a new type of technique in which the authors used the wavelet transform applied locally, based on the chaotic logistic map. This technique shows very good robustness to geometric attacks but it is sensitive to common attacks like filtering and sharpening.

Kundur et al. [4] proposed the use of gray scale logos as watermark. They addressed a multi-resolution fusion based watermarking method for embedding gray scale logos into wavelet transformed images. The logo undergoes 1-level decomposition for watermarking. Each sub band of the host image is divided into block of size equal to the size of sub band of the logo. Four sub bands of the logo corresponding to different orientations are added to the blocks of the same orientation. For fusion, the watermark is scaled by salience factor computed on a block by block basis. Reddy et al. [9] proposed a method in which the authors used a gray scale logo as watermark. To embed watermark, HVS characteristics were used to select the significant coefficients and watermark is added to these selected coefficients. Further, they used the model of Barni et al. [7] to calculate the weight factors for wavelet coefficients of the host image. They extracted watermark from the distorted image by taking into consideration the distortion caused by the attacks.

However, by traditional two dimensional wavelet it is hard to represent sharp image transitions [10] and smoothness along the contours [12]. Hence, bandelet [11] with adaptation to the geometric structure and contourlet [12] with anisotropy scaling law and directionality are presented to sparsely represent natural images. However, the computation of geometry in bandelet is in high complexity thus it is not possible to use in watermarking. Contourlet [12] proposed by Minh N. Do and Martin Vetterli is utilized to capture intrinsic geometrical structure and offer flexible multi scale and directional expansion form images. Because of the nearly critical sampling and fast iterated filter bank algorithm, contourlet is in lower complexity than bandelet. However, non-ideal filter are used in the original contourlet result in significant amount of aliasing components showing up at location far away from the desired support [13] and exhibit some fuzzy artifacts along the main image ridges. Yue Lu [13] proposes a new construction of the contourlet, called sharp frequency localization contourlet transform (SFLCT) and alleviates the non-localization problem even with the same redundancy of the original contourlet. Unfortunately, due to the down samplers and up samplers presented in the directional filter banks of SFLCT, SFLCT is not shift invariant, which is important in image watermarking and easily causes pseudo-Gibbs phenomena around singularities [14]. In this paper, we apply cycle spinning [14] to compensate for the lack of translation invariance property of SFLCT and successfully employed in image watermarking. Experimental results demonstrate that our proposed method outperforms the original contourlet (CT), SFLCT and cycle spinning-based contourlet (CS-CT) in terms of PSNR and visual effect.

2. SHARP FREQUENCY LOCALIZED CONTOURLET TRANSFORM

The original contourlet [12] is constructed by the combination of laplacian pyramid, which is first used to capture the point discontinuities and the directional filter banks (DFB), which is used to link point discontinuities into linear structure. In the frequency domain, the laplacian pyramid

iteratively decompose a two dimensional image into low pass and high pass sub bands and the DFB divide the high pass subbands into directional sub bands as shown in Fig.1.

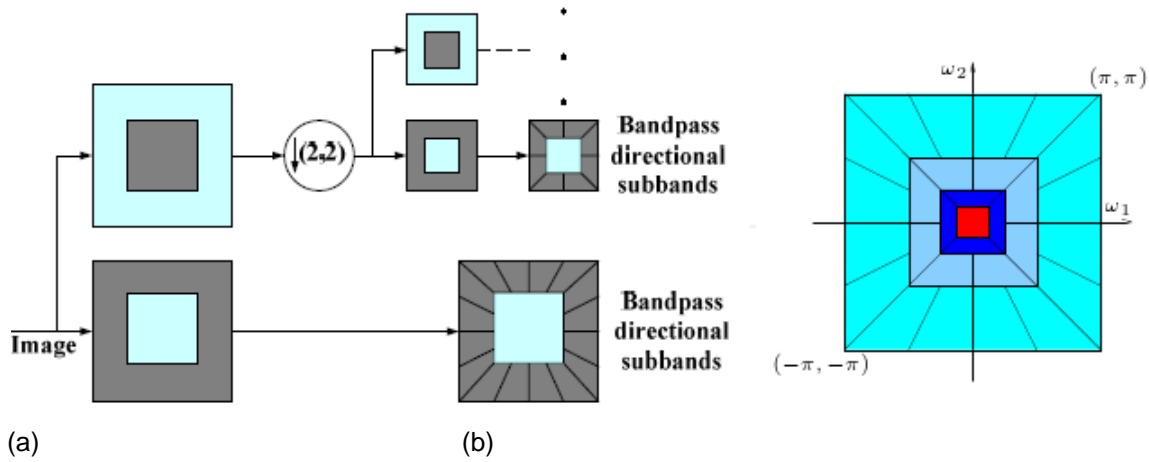


FIGURE 1: The original contourlet transform (a) Block diagram (b) Resulting frequency division

However, the frequency division in Fig.1 (b) is obtained by ideal filters. When non ideal filters are combined with laplacian pyramid, and suppose only one direction is extracted, the aliasing frequency spectrum will concentrate along two parallel lines $2\omega = \pm\pi$ as shown in Fig.2 (a). Furthermore, if the directional filters are up sampled by 2 along each dimension, the aliasing components will be folded towards the low pass regions, as patterned in Fig.2 (b), and concentrated mostly along two lines $2\omega = \pm\pi$. When the directional filters are combined with bandpass filter in laplacian pyramid, the contourlets are not localized in frequency, with substantial amount of aliasing components outside of the desired trapezoid shaped support [13] as the gray region shown in Fig.2 (d).

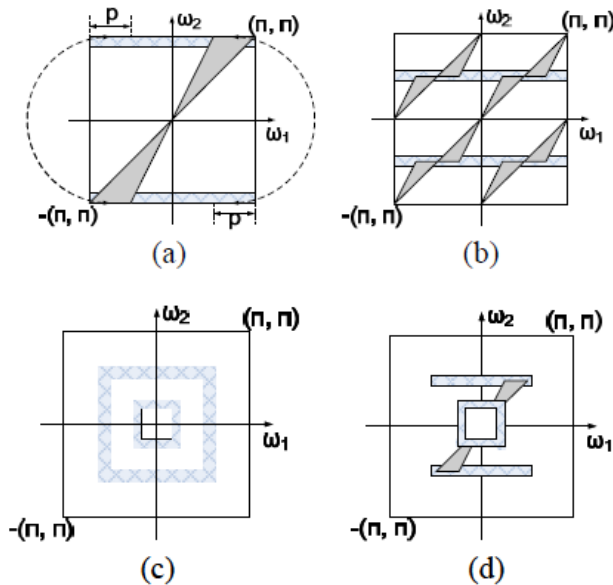


FIGURE 2: Spectrum aliasing of the original contourlet. Gray regions represent the ideal pass band support. Patterned regions represent the aliasing components or transition bands. (a) One directional filter, (b) The directional filter up sampled by 2, (c) A band pass filter from the laplacian pyramid, (d) The resulting contourlet subband.

To solve this problem, Yue Lu [13] proposes a new construction of sharp frequency localization contourlet (SFLCT). Since the combination of laplacian pyramid and directional filters banks make the aliasing problem serious, new multi scale pyramid with different set of low pass and high pass filters for the first level and all other levels are employed. Though SFLCT is sharply localized in the frequency domain and improves the watermarking performance [13], down samplers and up samplers presented in the directional filter banks of SFLCT make it lack shift invariance, which could easily produce artifacts around the singularities, e.g. Edges. Thus, cycle spinning is employed here to compensate for the lack of translation invariance. It is a simple yet efficient method to improve the quality of watermarking for a shift variant transform. In order to show the way of effective representation of input image textures in output watermarked and extracted images, the proposed Cycle Spinning based SFLCT (CS-SFLCT), it is compared with the original SFLCT [13] and Contourlet (CT) [12] as well as the Cycle Spinning based Contourlet (CS-CT) [15], by using '9-7' and 'pkva' filters are used in pyramidal and directional decomposition and inverse of the transform is applied on them and the figure below shows the presence of artifacts in the processed peppers image.

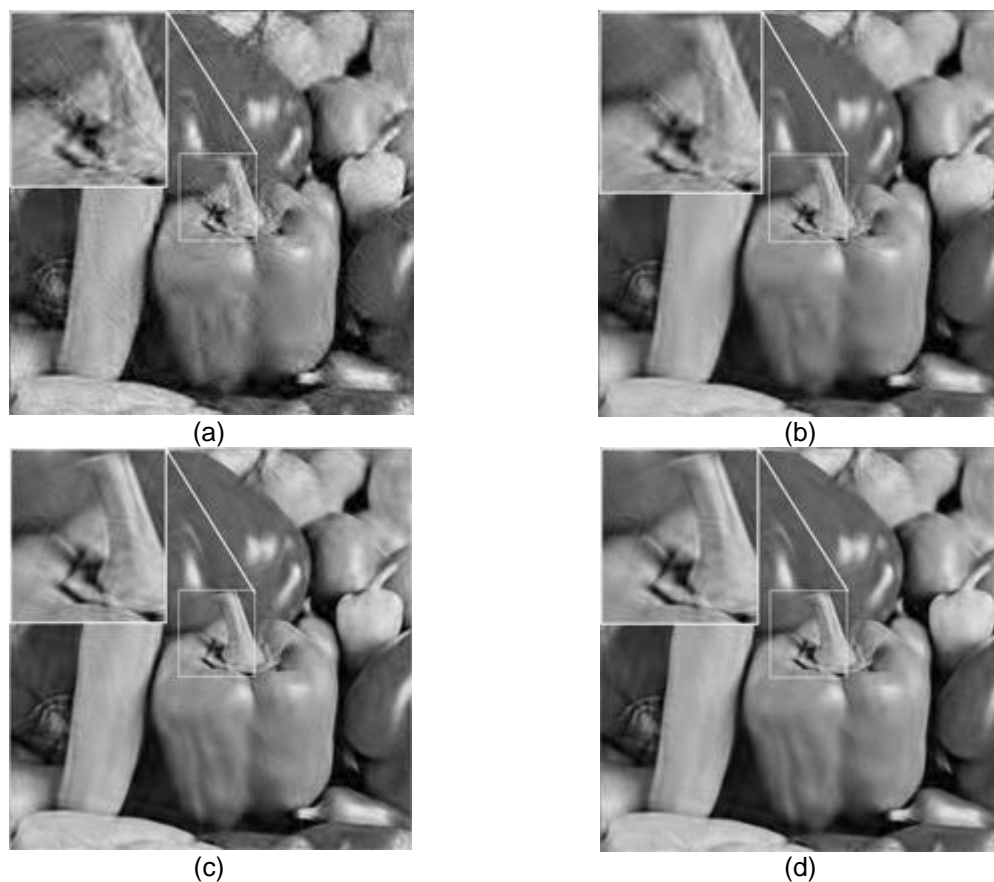


FIGURE 3: (a)-(d) are the processed images using CT, CS-CT, SFLCT, CS-SFLCT and their inverses respectively

3. PRINCIPAL COMPONENT ANALYSIS

An effective way to suppress redundant information and provide only one or two composite data most of the information from the initial data is also called principal component analysis. This plays vital role in our proposed approach by reducing redundant information, which acts as additional payload in the Meta data (the information to be embedded)

EXISTING SYSTEM

The existing non blind watermarking algorithms include wavelet transforms that are failing to preserve the edge with curvy structure. This makes lot of degradation in the quality difference of host and watermarked Images. This is becoming a serious issue. Adding watermark directly to wavelet coefficients is also not a good practice. The Peak Signal to Noise Ratio is less for the above reasons. Some people have used contourlet transform instead of wavelet. This has improved the performance in better representing curves than wavelets. But non ideal filters are used in the original contourlet transform, especially when combined with laplacian pyramid, which results in pseudo-Gibbs phenomena. This causes aliasing of frequency components introducing some unwanted artifacts and results distortions in watermarked image

PROPOSED ALGORITHM

We are using Cycle Spinning-based Sharp Frequency Localized Contourlet Transform in place of wavelets or contourlets which gives better representation of curves without introducing unnecessary artifacts as mentioned above. Instead of adding watermark information to *CS-SFLCT* coefficients we are applying *CS-SFLCT* to watermark and embedding action is taking place at principal components instead at contourlet coefficients. The additive technique is also varied in accordance with the amplitudes of principal components of host images, This increases still more distortions and resulting watermarked image will perfectly resembles with host images according to Human Visual System.

The Host and Watermark images we are using of true color images and the watermarking coefficient is varying from 0.001 to 0.25 in steps of 0.005 (a total of 50 values are used).

Our watermarking system consists of three main parts namely

- 1.Embedding at Transmitter Part.
- 2.Channel effects and attacks
3. Extraction at Receiver Part.

4.1 EMBEDDING AT TRANSMITTER PART

The multimedia data to be sent/broadcast is processed to Embedding process. The embedding process typically involves the following procedure.

Step1: Apply *CS-SFLCT* to the R, G, B planes of true color host image and watermark image.

Step2: Consider the mid frequency subband of each color space and apply PCA to each of them.

Step3: As the watermark image is also a true color image apply the steps 1& 2 to watermark image.

Step4: Principal components obtained in steps 2 and 3 are processed by the following procedure

max=maximum(PCA_MSB_HI);

min=minimum(PCA_MSB_HI);

Range=(max-min)/Count;

temp=(PCA_MSB_HI[i,j]-min)/Range;

alp=0.001+(temp-1)*0.005;

PCA_MSB_Wmkd[i,j]=PCA_MSB_HI[i,j]+alp*PCA_MSB_WI[i,j];

Where PCA_MSB_HI: Principal components of Middle Sub band of Host Image at one of the three color Spaces

PCA_MSB_WI: Principal Components of Middle Sub band of Host Image at one of the three color spaces.

PCA_MSB_Wmkd: Resulting Principal Components of Watermarked image.

Alp is the watermarking Coefficient varies from 0.001 to 0.25 (takes a total of 50 values).

Step 5: The PCA_MSB_Wmkd and covariance matrix obtained in step 2(during the calculation of PC's of Host image sub band) are multiplied and resulting will be the middle sub band coefficients of watermarked image. This is repeated for R,G,B color spaces.

Step 6: The coefficients obtained in step5 and coefficients of host images are fed to inverse *CS-SFLCT* to get the watermarked image. This is repeated for R,G,B color spaces.

Step7: The watermarked image obtained will converted to unsigned integer 8 representation from double to be checked for invisibility using PSNR.

4.2. CHANNEL EFFECTS AND ATTACKS:

When the watermarked images are broad casted or transmitted through various channels it will undergo with some properties of channel such as noise and filtering effects. Generally when a signal is traveling through a channel, the channel output will be the convolution of channel transfer function and input signal. The transfer function of channel is dependence on some of its physical features and mainly on its band width. So the final output from the channel is filtered.

At this stage one do not have the knowledge of channel that is going to be used,so our algorithm should be tested with respect to the channel affects. That is why, we have processed the watermarked image with low pass filtering, Median filtering, Gaussian noise and salt & pepper noise etc.

When a multimedia data is made to transfer over a public channel, always we have the threat of hackers and malware issues. Now a days there are marvelous number of software and malware over World Wide Web to introduce significant changes in Multimedia Data. That is why our algorithm is tested with respect to some more attacks like Histogram equalization, scaling,rotation,Gamma Correction, Cropping, row or column removal, Wiener filtering, Blurring, Dilation, Color Space Conversions, shearing etc.

4.3. EXTRACTION AT RECEIVER PART:

The direct watermarked image or channel processed and received image is checked for authentication. This is done at receiver side. Our watermarking system is non blind in which the host image is required for extraction of hidden watermark in the watermarked image.

The Procedure of Extraction as follows:

Step1: Apply *CS-SFLCT* to the R, G,B planes of true color host image and watermarked image.

Step2: Consider the mid frequency sub band of each color space and apply PCA to each of them.(both for host and watermarked images)

Step3: Principal components obtained in steps 2 are processed by the following procedure.

```
max=maximum(PCA_MSB_Wmkd);
min=minimum(PCA_MSB_Wmkd);
Range=(max-min)/Count;
temp=(PCA_MSB_Wmkd[i,j]-min)/Range;
alp=0.001+(temp-1)*0.005;
PCA_MSB_Extd[i,j]=(PCA_MSB_Wmkd[i,j]-PCA_MSB_HI[i,j])/alp;
```

Where PCA_MSB_Extd : Resulting principal components of Extracted image.

Alp is the watermarking Coefficient varies from 0.001 to 0.25 (takes a total of 50 values).and this is separated for R,G,B color spaces.

Step 4: As an authentication key, the extractor should carry the partial information of watermark image.That will be the Covariance matrix and other sub bands of Watermark image.They are used to apply inverse PCA and Inverse *CS-SFLCT* to get back the watermark image by doing same for R,G and B color spaces.

Step5: Extracted image obtained will converted to unsigned integer 8 representation from double to be check for Robustness by comparing with original watermark image using Normalized Correlation Coefficient.

5. INTERVISIBILITY AND ROBUSTNESS MEASURES

As our total concept wanders around Data Security issues, keeping the knowledge of watermarking makes an added advantage. Likewise the watermarked image and Host image should perfectly resemble each other with respect to HVS.The degree of disagreement can be measured in the MSE (Mean Square Error) or PSNR (Peak Signal to Noise Ratio).The more the PSNR, more the invisibility.

$$MSE = \sum_{y=1}^M \sum_{x=1}^N [I(x, y) - I'(x, y)]^2$$

$$PSNR = 20 \times \text{Log}_{10} \left(\frac{255}{\sqrt{MSE}} \right)$$

Robustness is also one of the quality measure of watermarking system with respect to watermark image and extracted image. The robustness gives the quality measure of watermarking system with respect to intentional and unintentional attacks and channel actions. After undergoing all these, if the watermarked image is driven to extraction process at receiver part, the extracted image should match with the original watermark image. Then only the system is said to have Robustness. The quality of match can be measured by Bit Error Rate and Normalized Correlation Coefficient.

$$NCC = \frac{\sum_{x=1}^m \sum_{y=1}^n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left[\sum_{x=1}^m \sum_{y=1}^n (A_{mn} - \bar{A})^2 \right] \left[\sum_{x=1}^m \sum_{y=1}^n (B_{mn} - \bar{B})^2 \right]}}$$

$$BER = \frac{\text{No.of .Error.Bits}}{\text{Total.No.of .Bits}}$$

6. EXPERIMENTAL RESULTS

True color images barbara.pgm and lena.pgm of sizes 256X256 as host and watermarked images. The experiment is repeated for different ranges of alpha(the watermarking coefficient).The prominent value with good invisibility is observed at alpha 0.001 to 0.25.Under ideal channel without any attacks we have observed PSNR of 63.72dB and NCC of 0.9999, BER=0.0002.The NCC less than one is because the PCA is not completely reversible technique. Experiment is repeated for different intentional and unintentional malicious attacks and results are stated below.



Host Image
Barbara.pgm





Watermark Image
Lena.pgm



Watermarked Image
PSNR=63.72dB



Extracted Image
NCC=0.9999, BER=0.1319

<p>Salt & Pepper Noise PSNR=53.6935,NCC= 0.9858 BER=0.3586</p>	<p>JPEG Compression PSNR=42.9998,NCC= 0.9861 BER= 0.3105</p>	<p>Gaussian Blur PSNR= 26.8958,NCC = 0.9899 BER = 0.2952</p>
		
<p>JPEG2000 Compression PSNR=53.5443,NCC=0.9995 BER=0.1658</p>	<p>Low Pass Filtering PSNR= 51.0915,NCC= 0.9893 BER=0.2966</p>	<p>Dilation PSNR=35.3595,NCC = 0.9851 BER= 0.6405</p>
		
<p>Rotation PSNR=40.2362, NCC=0.9868 BER= 0.3055</p>	<p>Histogram Equalization PSNR= 26.8958, NCC= 0.9807 BER= 0.3512</p>	<p>Shearing Attack PSNR=25.856, NCC = 0.9893 BER = 0.2969</p>
		
<p>Median Filtering PSNR= 36.9651, NCC= 0.9852 ,BER= 0.3298</p>	<p>Contrast Adjustment PSNR= 38.8907, NCC= 0.9998 , BER= 0.1058</p>	<p>Color to Gray scale conversion PSNR= 25.0829, NCC = 0.9441 BER = 0.3775</p>
		
<p>Wiener Filtering PSNR=48.3950, NCC=0.9989 BER =0.2158</p>	<p>Gamma Correction PSNR=56.2008, NCC = 0.9996, BER = 0.0728</p>	<p>Row Column removal PSNR= 35.5668, NCC = 0.9977 BER = 0.1931</p>
		
<p>Gaussian Noise PSNR=38.4451, NCC = 0.7266 BER = 0.5851</p>	<p>Sharpening PSNR=40.6818, NCC = 0.8448 BER = 0.5177</p>	<p>Row column copying PSNR=38.9696, NCC = 0.9895 BER= 0.3208</p>
		
<p>Rescaling (150%) PSNR=48.9696, NCC = 0.9861 BER = 0.6408</p>	<p>Automatic cropping (85%) PSNR= 28.9606, NCC= 0.9936 BER=0.2852</p>	<p>Speckle Noise PSNR=38.5177, NCC= 0.7031 BER= 0.5683</p>
		

7. CONCLUSION AND FUTURE SCOPE

Our hybrid approach "Image watermarking using CS-SFLCT and PCA" is showing good performance when compared with DWT and PCA, CT and SVD, CT and PCA in both invisibility and robustness. The usage of CS-SFLCT is reducing aliasing effects and by means of which reducing unexpected artifacts like in contourlet transform. Usage of PCA as hybrid improves still further improving the quality of watermarking system. But our approach being non blind requires host image and some partial information of watermark also at the receiving end. Usage of security key still further improves the authentication. This system is best suitable for medical images especially because the action of unexpected artifacts in CT will change the information it contains. This work can be extended to blind technique (at the extraction stage only watermarked image is sufficient) using regression techniques. It is also observed that usage of Cycle Spinning-based form of Sharp Frequency Localized Contourlet greatly reduces extra time taking of watermarking using non subsampled Contourlet Transform.

8. REFERENCES

- [1] I.J. Cox, M.L. Miller, J.A. Bloom, Digital watermarking, Morgan Kaufmann, 2001.
- [2] A.G. Bors, I. Pitas, Image watermarking using DCT domain constraints, Proceedings of IEEE International Conference on Image Processing, vol. 3, 1996, pp. 231–234.
- [3] R.G.V. Schyndle, A.Z. Tirkel, C.F. Osbrone, A digital watermark, Proceedings of IEEE International Conference on Image Processing, vol. 2, 1994, pp. 86–90.
- [4] D. Kundur, D. Hatzinakos, Towards robust logo watermarking using multiresolution image fusion, IEEE Transactions on Multimedia 6 (2004) 185–197.
- [5] J. Ohnishi, K. Matsui, Embedding a seal in to a picture under orthogonal wavelet transform, Proceedings of IEEE International Conference on Multimedia and Computing system, 1996, pp. 514–521, IEEE, Hiroshima, Japan.
- [6] P. Meerwald, A. Uhl, A survey on wavelet domain watermarking algorithms, Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents III, vol. 4314, 2001, pp. 505–516, SPIE, CA, USA. wise masking, IEEE Transactions on Image Processing 10 (2001) 783–791.
- [8] Z. Dawei, C. Guanrong, L. Wenbo, A chaos based robust wavelet domain watermarking algorithm, Chaos, Solitons, and Fractals 22 (2004) 47–54.
- [9] A.A. Reddy, B.N. Chatterji, A new wavelet based logo-watermarking scheme, Pattern Recognition Letters 26 (2005) 1019–1027.
- [10] U.Majer Ali, E.Vinoth Kumar, Digital Watermarking using DWT and SVD, NCACSA 2012, IJCA 2012
- [11] Pennec E, Mallat S, "Sparse geometric image representation with bandelets", *IEEE Transaction on Image Processing*, 2005, 14(4): 423-438.
- [12] MN Do, M Vetterli, "The contourlet transform: an efficient directional multiresolution image representation", *IEEE Transaction on Image Processing*, 2005, 14(12): 2091-2106.
- [13] Yue Lu, MN Do, "A New Contourlet Transform with Sharp Frequency Localization", *Proceedings of 2006 IEEE International Conference on Image Processing*, IEEE, Atlanta, USA, 2006, 1629-1632.

[14] R. R. Coifman ,D. L. Donoho, "Translation invariant de-noising," *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds. New York: Springer-Verlag, 1995,125–150.

[15] Eslami R , Radha H, "The contourlet transform for image denoising using cycle spinning", *Proceedings of Asilomar Conference on Signals ,Systems, and Computers*, 2003,1982-1986.

Adaptive Multiscale Stereo Images Matching Based on Wavelet Transform Modulus Maxima

Abdelhak EZZINE
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

aezzine@uae.ac.ma

Youssef YazidiAlaoui
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

yazidi17@gmail.com

Achraf Cohen
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

cohenachraf@gmail.com

Abstract

In this paper we propose a multiscale stereo correspondence matching method based on wavelets transform modulus maxima. Exploitation of maxima modulus chains has given us the opportunity to refine the search for corresponding. Based on the wavelet transform we construct maps of modules and phases for different scales, then extracted the maxima and then we build chains of maxima. Points constituents maxima modulus chains will be considered as points of interest in matching processes. The availability of all its multiscale information, allows searching under geometric constraints, for each point of interest in the left image corresponding one of the best points of constituent chains of the right image. The experiment results demonstrate that the number of corresponding has a very clear decrease when the scale increases. In several tests we obtained the uniqueness of the corresponding by browsing through the fine to coarse scales and calculations remain very reasonable.

Keywords:Maxima modulus, stereo matching, wavelet transform, maxima modulus chains.

1 INTRODUCTION

Matching the extracted visual indexes of stereoscopic images is a very significant step within the process of 3D version. It is made up of bringing together the primitive extracts of right and left images so as to realize a 3D reconstruction of the scene. This is based on estimating the offset between the positions of matched primitives and building a disparity map showing the scene by highlighting the depth of scene objects. The disparity map generated from the correspondence matching process, along with a stereo camera are used to estimate the depth map and produce the 3D view of the scene.

This problematic is largely treated by a variety of researchers yet still there is no general method [6,8,10,26], many approaches exist in the field and exploit generally the primitives such as reference points, contours or areas but generally via the same characteristics which are either photometric, geometric or either morphologic.

This technique remains difficult due to the multiplicity of essential parameters, mainly the pertinence of considerable attributes, in the process of matching and the difficulty to find the good

corresponding primitive is subject to a number of potential problems in vision process like occlusion, ambiguity, illuminative variations and radial distortion and all algorithms proposed are adapted to control at least some of these problems. We find many class of stereo vision algorithms, the first class is named global algorithms (GA) who deal with the correspondence estimation process as global cost function optimization problem and the second one is the local algorithms (LA) and are categorized into either area based or feature based algorithms [23]. The area-based algorithm used the correlation of the two image functions over locally defined regions and the feature-based algorithms establish correspondence between the selected primitives of the image.

Our method is one of the new class of algorithms [5,6], it is based on the concept of wavelets because the matching is the result of a multi-scale operation of the modulus and the phase of wavelet coefficients. This class is considered in the middle way between the local and global algorithm, these combine the best features of the LA and GA techniques and uses the multiresolution concept to involves the matching of two images at different scales.

This method is based on the wavelets modulus maxima chain (CMI) and involves matching of the different points forming the chain (CMI) to find the best corresponding in the right image for each extremity of one chain in left image. The search of the best one is done by estimating a similarity test between the points of each left chain and points of the right chains satisfying the geometric constraint like epipolar and orientation. The disparity map generated has a multiscale character as it is built based on the simultaneous treatment of space and scale. The multiscale character will permit the refinement of the disparity map and possibility to refine the selection of the potential candidates using a multiscale criterion based on normalized correlation and geometric refinement. As proved in Mallat[7,22,12], the extremity of the chain coincides with a singularity of the image characterized by its coefficient of Lipchitz, thus with a good choice for constructing maxima chains, the proposed matching algorithm not only provides the precise results but it also converges rapidly.

Firstly, we are going to recapitulate the main principles of wavelet transform modulus maxima (MMW). Then, we move to present the method and its particularities. Finally, we synthesis the major results of the method we intend to compare with other attributes.

2. WAVELETS TRANSFORM MODULUS MAXIMA

Based on the idea that in the image processing, visual data is put in a hierarchic way via the scale, wavelets transform decompose image into elementary blocs that are well localized both on scale and space. As a result of this decomposition, the image can be represented by the approximation to the coarse resolution and detail to all intermediate scales.

Among the main interests of representation in the form of wavelets we state the coefficient wavelet length exam of an image can inform us about it analysis. Indeed, in [2, 9], it is proved that the regularity accentuates the decay of wavelet coefficients; we are talking about the local regularity related to the analyzing wavelet support [17, 14].

Thereby, if the image is continuously regular everywhere except at a few isolated points the estimated wavelet coefficients will be affected only if the support of the analyzing wavelet contains these points or edges. In this sense, multiscale representations are better suited to focus the information in an image whose regularity is not homogeneous because the wavelet coefficients above a certain threshold focus only near the singularities (edges, single points, etc.). One of the multiscale decompositions that exploits well the wavelet properties which permit making a study, in a different levels, about the discontinuity points present in an image is wavelet transform modulus maximas proposed by Mallat and Zhong[1,7].

The principle of wavelet modulus maxima can give the image of characteristics to identify it, they transmit obvious features of the image depending on well defined models and well determined

directions. For more details about wavelets modulus maxima we refer to [2, 13,14, 21] and present bellow the essential terms.

Suppose two wavelets ψ^x and ψ^y such that $\psi^x = -\frac{\partial G}{\partial x}$ and $\psi^y = -\frac{\partial G}{\partial y}$ and $\int_{-\infty}^{+\infty} G(t)dt \neq 0$. For this work we choose G as a Gaussian wavelet to ensure both a valorization of wavelet maxima and continuity of the wavelet maxima chains [4, 7]. Thus wavelet transform of an image I can be written as a multiscale differential operator

$$\begin{pmatrix} W^x I(u, v, j) \\ W^y I(u, v, j) \end{pmatrix} = \begin{pmatrix} I * \psi_j^x(-u, -v) \\ I * \psi_j^y(-u, -v) \end{pmatrix} = 2^j \vec{\nabla} (I * G_j(-u, -v)) \quad (1)$$

The polarized representation of this gradient vector offers the wavelet transform modulus

$$MI(u, v, j) = \sqrt{|W^x I(u, v, j)|^2 + |W^y I(u, v, j)|^2} \quad (2)$$

and its direction is defined by the angle $\theta I(u, v, j) = \tan^{-1} \left(\frac{W^y I(u, v, j)}{W^x I(u, v, j)} \right)$. (3)

Then the local maximum of the wavelet transform modulus $MMI(u, v, j)$ can be found by solving $\partial MI(u, v, j) = 0$.

Otherwise on some scale j , the points (u_0, v_0) such that $MMI(u, v, j)$ is a local maximum, according the L_∞ norm, on a neighborhood of (u_0, v_0) in the direction $\theta MI(u, v, j)$. We note that for image color we consider the maxima according DiZenzo approach[11].

3. MULTILEVEL MATCHING METHOD

In works of Mallat and Zhong[7] is prove that a wavelet coefficient can be influenced by a singularity resides within support of the analyzing wavelet; we talk about the cone of influence to express the set of points where the wavelet coefficients may be influenced by a singularity at a point (u_0, v_0) . This singularity will generate local maxima of wavelet coefficient upon the cone and till a particular scale. In order to characterize this singularity, we will observe the behavior of local maxima in terms of scale; this will be much easier to follow if the maxima, between two successive scales, are connected together to form a curve in the plane scale space so-called line or chain maxima.

In practice, we start with an image of size $(N, N), N = 2^J$. We obtain the wavelet transform according dyadic approach [3,11,18] in which scale parameter is expressed by the power of 2, $2^j, j = J \dots L$, L is the coarsest scale.

This method generates on each scale, a wavelet transform with the same number of pixels as the image I ; in spite of that the support of the wavelet is increasingly large when scales grow. After that we calculate the wavelets transform modulus and directions by equations (2) and (3). The results are two matrix, the maxima modulus matrix $MMI(\dots, j)$ and the angle matrix $\theta MI(\dots, j)$ for each scale $j = J - 1 \dots L$. The image I is then represented by $(MMI(\dots, j))_{j=J-1 \dots L}$, positions and values of the wavelet coefficients at each scale j when the module is a local maximum in the direction θ .

We construct the line or chain of maxima by a chaining (x_j, y_j) , the location of each element of $MMI(\dots, j)$, with (x_{j-1}, y_{j-1}) , the location of its successor in $MMI(\dots, j-1)$ [21,23]. The search for these successors is limited to the neighborhood defined by the cone of influence and every chain can be indexed by $(x_j, y_j)_{j=J \dots m}$ and noted $CMI(x_j^l, y_j^l, m)$ [14,11]. The integer m indicate the length of the chain, and is defined such that the last point of the chain is ones of maxima of the

wavelet transform at scale $J-m$. We note that each chain is composed of maxima and is characterized by its extremity $MMI(x_j, y_j, J)$ at the level J that coincides with a singular point of the image.

3.1 The matching approach

Given two rectified color images, from a vision system calibrated. we associate to each one its maxima modulus chains representation $(x_j^i, y_j^i)_{j=1..m}$, $i = g$ for the left image and $i = d$ for the right one. With different representations defined above, we are faced with three possible methods to mapping. The first is to match directly the maxima modulus transforms $MMI_i(\dots, j)$ of the two images at each scale $j = J - 1 \dots L$. The second performs the matching of the extremity of each chain $CMG(x_j^g, y_j^g, m)$ with an extremity of a chain $CMD(x_j^d, y_j^d, m')$ by a non-linear algorithm. The third method proposes to consider the maxima chains as primitives. In this paper we develop the second method and we will give an idea about the treatment of the third, we note that the first one looks approximately like the work of Bhatti and we invite the reader to discover in this work [5].

To chain $CMG(x_j^g, y_j^g, m)$ of the left image, we will search for all Chain of the right image such that their extremities respect the epipolar constraint with the pixel (x_j^g, y_j^g) ; all maxima of the left that have no match in the right image verifying the epipolar constraint are dropped. We then perform a test of similarity between $MMI(x_j^g, y_j^g, J)$ and each one of the maxima modulus on extremities of the rights chains selected before. We note $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$ the set of extremity of the right chains validating this test. If $c = 1$, we conclude that (x_j^g, y_j^g) is matched with $(x_{j,1}^d, y_{j,1}^d)$. If $c > 1$, we move to the next scale and we redo the estimating test between the successors, i.e. (x_{j-1}^g, y_{j-1}^g) and $\{(x_{j-1,k}^d, y_{j-1,k}^d)_{k=1..c}\}$. This step is repeated, recursively on the scale j , until the end of the chain and it will result necessarily matching of (x_j^g, y_j^g) and ones of $\{(x_{j,k}^d, y_{j,k}^d)_{k=1..c}\}$ points.

The similarity is evaluated, under the orientation constraint, by using ones off different similarity measure available in the literature. Major similarity measures used in matching methods are measures of distortion (SAD) and (SSD) and the normalized cross correlation (NCC)[15]. It is known that SAD and SSD are computationally fast then NCC who is more accurate and it allowed us to refine the similarity test.

At fine scales, there are many edge points created by the image noise in the wavelet transform maxima modulus representation. By reason of the presence of these wrong maxima and because it is well known that the SSD and SAD are justified when the additive noise distribution is Gaussian or exponential, we are led to combine the NCC and SAD to the similarity test for the fine scales. We note that at larges scales the smoothing process removes most of the wrong maxima and the NCC is sufficient.

To find the correspondence between pixels of $MMG(x_j^g, y_j^g, j)$ and $MMD(x_j^d, y_j^d, j)$, one can maximize a correlation or minimize the distortion measure. For each point (x_j^g, y_j^g) , we search for a point (x_j^d, y_j^d) in $CMD(x_j^d, y_j^d, m')$, whose a neighborhood of a size s has a maximum correlation or maximum distortion with the neighborhood of the point (x_j^g, y_j^g) in $CMG(x_j^g, y_j^g, m)$. Due to the epipolar constraint we have $y_j^d = y_j^g$ and the search is reduced to points (x_j^d, y_j^g) on the epipolar line.

4. EXPERIMENT RESULTS

The algorithm presented in this paper is tested on the images from web site [10,10b] www.middlebury.edu/stereo/. These images are created for the stereo matching algorithms we present the results obtained for four of them with the same illumination. In order to give visual results, the performance of the proposed nonlinear algorithm is presented for different

scales j with the dyadic approach wavelet transform so we can follow the behavior of a maximum on different scales.

The experimental results were generated using the Gaussian wavelets. In the first column of the figure 1, we present the left images cones, teddy and tsukuba. In the fifth column we present the right view of the same images. In the second and fourth column we present the maxima modulus on scale $J=4$.

In the third column are presented the number (Z axis) of maxima of the right image validating the similarity test with each one of the maxima (X axis) of the left image. We present these results of the simulation only for 50 points (on the X axis) represents the great maxima of each one of the tree images. The third axis (Y axis) represents the scale and for each maxima of the left image we represent the behavior of the number of corresponding maxima of the right image on each scale j .

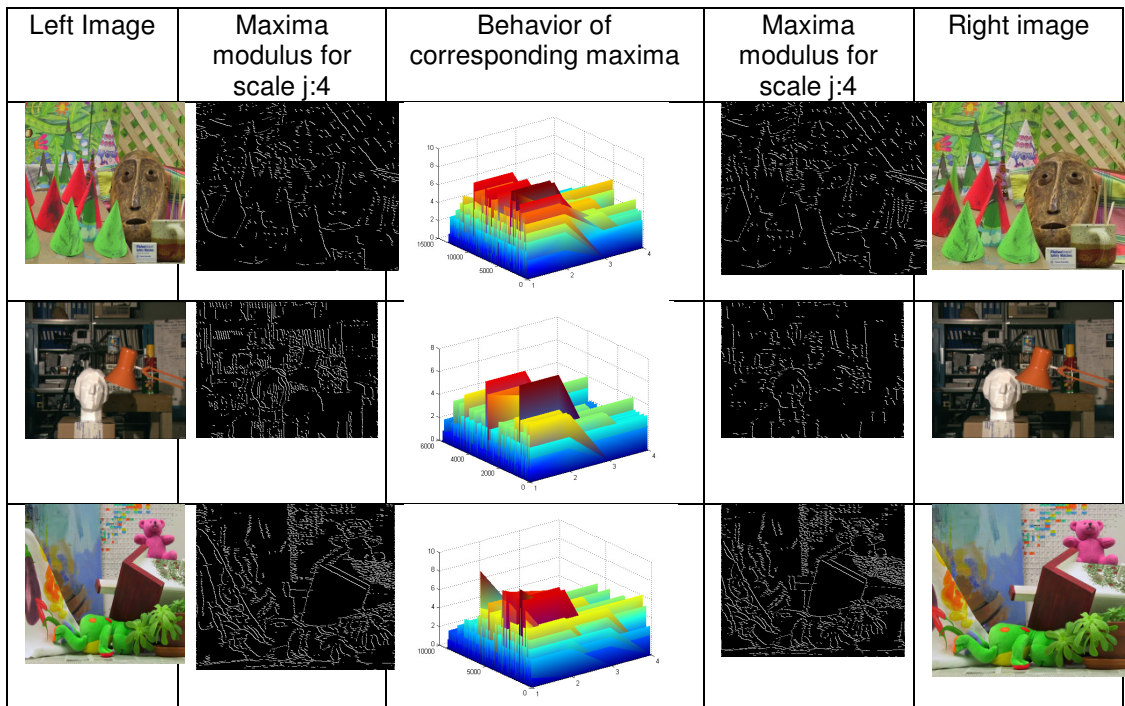


Figure 1: in third column we observe the decreasing behavior of number of corresponding maxima in right image at scale $j:4$ for everyone of maxima in left image (x axis :scale; y axis: maxima and z axis: number of corresponding)

We observe that the number of corresponding decrease clearly when the scale increases. There are more than 6 corresponding in the first scale, our method allows the reduction of this number to 1 for several maximas. We observe also that maxima corresponding to the strong singularity require further exploration scales.

This observation is more important and we are working to prove the relationship between the category of singularities and the speed of convergence of this method. To compare the results of this method with other work and especially by platform Middleburry[10,10b], we need to validate the disparity map based on the points of interest of our method and this is one of the future direction of our research.

5. CONCLUSION

This paper introduced a nonlinear algorithm matching based on wavelet transform maxima modulus. The maxima modulus chains are used to find the best corresponding maxima in the right image of each maximum in left image. This process exploits intelligently the multiscale character and offers an accurate and fast algorithm. We note also that the first results of the construction of disparity maps are very encouraging.

Following this work is, firstly, affine the disparity map and compare it to techniques benchmarked and published in the Middlebury database; and secondly integer the Lipschitz regularity of pixels [2,16,17] in the similarity process as an additive information, because it was proven that there is a strong relationship between maxima and this regularity.

This will allow us to exploit the nature of the maxima as edge points in the image and propose a non-linear feature-based algorithm. The next future direction on this work is to explore different wavelets basis and observe the behavior of the nonlinear algorithm of this paper according the choice of wavelets basis.

6. REFERENCES

- [1] S. Zhong, "Edge Representation from Wavelet Transform Maxima". PhD thesis, New York University, September 1990.
- [2] Y. Meyer, "Ondelettes et opérateurs", vol. I. Paris: Hermann, 1990.
- [3] M. Shensa, "Wedding the à trous and Mallat algorithms," IEEE Trans. Signal Processing, vol. 40, no. 10, pp. 2464-2482, 1992.
- [4] J. Canny, "computational approach to edge detection", IEEE Trans. on PAMI, vol. 8, no. 6, pp. 679-698, 1986.
- [5] A. Bhatti, S. Nahavandi "Stereo Image matching Using Wavelet Scale-Space Representation", Proceeding of CGIV'06, pp 267-272, 2006
- [6] P. B. Zadeh, C. V. Serdean "A Hierarchical multiwavelet based stereo correspondence matching techniques", EURASIP'2011, pp 393-397
- [7] S. Mallat and S. Zhong, "Characterization of signals from multi-scale edges", IEEE Trans. on PAMI, vol. 14, pp. 710-732, July 1992.
- [8] G. Qu, D. Zhang and P. Yan "Medical image fusion by wavelet transform modulus maxima", Optics Express Vol 9, No 4 pp 184-190, 2001.
- [9] A. Cohen and A. Ezzine "Quadratures singulières et fonctions d'échelle", CRAS, t.323, Serie I, 829-834, 1996.
- [10] D. Scharstein and R. Szeliski, www.Middlebury.edu/stereo/,
- [10b] D. Scharstein and R. Szeliski, Stereo matching with nonlinear diffusion. Int. J. of Computer vision, 28(2), 1998.
- [11] S. Dizenzo, "A note on the gradient of a multi-image", CVGIP, vol. 33, pp. 116-125, 1986.
- [12] S. Mallat and W. Hwang, "Singularity detection and processing with wavelets", IEEE Trans. Information Theory, vol. 38, no. 2, pp.617-643, 1992.

- [13] P. Carré, " Maxima d'ondelettes couleur, colorwavelet maxima ", traitement du signal 2004 vol 21 special L'image numérique couleur 623-632.
- [14] S. Mallat, " Wavelets for a vision ", Processing of IEEE, vol. 84, NO 4, pp. 604-614, April 1996.
- [15] L. Distefano, Stefano Mattocia, " Fast template matching using bounded partial correlation ", Machine vision and application, NO 13, pp. 213-221, 2003.
- [16] S. Jaffard, " Exposants de Holder en des points données et coefficients d'ondelettes", CRAS France,308ser I ,pp79-81,1989.
- [17] I. Daubechies, Ten lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
- [18] S. Mallat, a wavelet tour of signal processing, Academic Press, second edition, 1998.
- [19] J. Canny A computational approach to edge detection, IEEE Trans. Patt. Anal. and Mach. Intell.,vol PAMI-8 no. 6, pp. 679-698, Novembre 1986.
- [20] M. Holschneider, R. Kronland-Martinet, J. Morlet, P. Tchamitchian Wavelets, Time-Frequency Methods, and Phase Space, Springer-Verlag, Berlin, 1989.
- [21] O. Lecadet Méthodes d'ondelettes pour la segmentation d'images. Applications à l'imagerie médicale et au tatouage d'images, these de doctorat, 2003.
- [22] S. Meignen, S. Achard, Time Localization of Transients with Wavelet Maxima Lines, IEEE. Trans. sign. Proc., accepted, 2004.
- [23] Y. Ruichek, Perception de l'environnement par stereovision , Application à la sécurité dans les systèmes de transports terrestres, Phdthesis, USTL, France 2006.

Robust Block-Matching Motion Estimation of Flotation Froth Using Mutual Information

Anthony Amankwah
*School of Computer Science
University of Witwatersrand
Johannesburg, 2050, South Africa*

anthony.amankwah@wits.ac.za

Chris Aldrich
*Western Australian School of Mines
Curtin University of University, GPO Box U1987
Perth, WA, 6845, Australia*

chris.aldrich@curtin.edu.au

Abstract

In this paper, we propose a new method for the motion estimation of flotation froth using mutual information with a bin size of two as the block matching similarity metric. We also use three-step search and new-three-step-search as a search strategy. Mean sum of absolute difference (MAD) is widely considered in blocked based motion estimation. The minimum bin size selection of the proposed similarity metric also makes the computational cost of mutual information similar to MAD. Experimental results show that the proposed motion estimation technique improves the motion estimation accuracy in terms of peak signal-to-noise ratio of the reconstructed frame. The computational cost of the proposed method is almost the same as the standard machine vision methods used for the motion estimation of flotation froth.

Keywords: Mutual information, Mean sum of absolute difference, Motion Estimation, Froth Flotation.

1. INTRODUCTION

Froth flotation is one of the most widely used separation methods in the mineral processing industries. Despite being in use for more than a century, it is still not fully understood and remains comparatively inefficient. Conventionally, the flotation process is controlled by trained operators who visually inspect the froth and adjust the process control parameters based on their training. A problematic situation is that different operators tend to interpret the visual characteristics of the froth differently.

However, during the last couple of decades, froth image analysis has been studied widely as a means to obviate these problems, as summarized in a recent review by Aldrich et al. [1]. Froth images are a potentially rich source of information with regard to the state of the flotation process and can be used as a basis for the development of advanced online control systems. Apart from the appearance of the froth, the motion of flotation froth is also an important variable related to the optimal recovery of minerals on industrial plants and has been used in industry among other as a basis for mass pull control strategies [2].

The estimation of motion of froth using image processing techniques is not easy, since bubbles collapse and merge, resulting in bubble deformations that may be difficult to track. Moreover, although the froth as a whole has an average motion, it moves rapidly at the edges of the cell at the launder overflow, while remaining almost stagnant at the centre of the cell [1].

Several approaches to froth motion estimation have been investigated, including bubble tracking [3], cluster matching [4], and pixel tracing [5], but the most popular motion estimation technique is

block matching [6-7] using the mean absolute difference (MAD) between successive images of parts thereof [2]. In the block matching algorithm, the image frame is typically divided into nonoverlapping rectangular blocks [8]. The best match to a given block of pixels is subsequently searched in the previous frame of the video sequence within a search area about the location of the current block. In principle, the best solution can be obtained with a full search, which exhaustively searches for the best matched block within all locations of the search window. However, this approach is computationally expensive and several fast search techniques have been proposed to reduce the high computational cost of full search methods. These fast search techniques select a subgroup of possible search candidate locations and thus reduce the number of matching calculations per block image.

Examples of such fast search techniques are three-step-search (TSS) [9], new-three-step-search (NTSS) [10], four-step-search [11], 2-D logarithmic search [12] and the diamond search [13]. The mean square error (MSE) or mean absolute difference (MAD) are considered to be among the best similarity metrics for motion estimation. In this paper, a new method is proposed for estimating the motion of flotation froth using mutual information. As will be shown, estimation of mutual information with a bin size of 2 as the block matching criterion improves the accuracy of motion estimation in terms of the peak signal-to-noise ratio of the reconstructed frame, while the computational complexity of the proposed similarity metric is also similar to MAD.

2. SIMILARITY METRICS

2.1 Mutual Information with Bin Size of Two

If X and Y are two image blocks to be matched, the mutual information (MI) can be defined by

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \tag{1}$$

where $H(X)$ and $H(Y)$ are the Shannon entropies [14] of X and Y respectively, and $H(X, Y)$ is the Shannon entropy of the joint distribution of X and Y . If $p_X(x)$ and $p_Y(y)$ are the marginal probabilities, and $p_{XY}(x, y)$ is defined as the joint probability density of X and Y . Then MI is defined as

$$MI(X, Y) = \sum_x \sum_y p_{XY}(x, y) * \log \left(\frac{p_{XY}(x, y)}{p_X(x) * p_Y(y)} \right) \tag{2}$$

In this work, we use the histogram method to estimate MI . Let $H_X(x)$ and $H_Y(y)$ be the histograms of X and Y respectively and $H_{XY}(x, y)$ their joint histogram. then MI can be defined as

$$MI(X, Y) = \frac{1}{E} \sum_x \sum_y H(x, y) * \log \left(\frac{EH_{XY}(x, y)}{H_X(x) * H_Y(y)} \right) \tag{3}$$

where E is the number of entries [15-17].

Figure 1 shows the joint grey value histograms and corresponding joint entropies of an image block with values shifted (a) 0 pixels, (b) 5 pixels, (c) 10 pixels and (d) 100 pixels. since the image pairs are identical, all corresponding grey values lie on the diagonal in Figure 1(a). From the joint histograms it can be deduced that i) diagonal features diminish with increasing dissimilarity. ii) the spatial dispersion pattern increases in areas with increasing dissimilarity, and iii) joint entropy increases with increasing similarity.

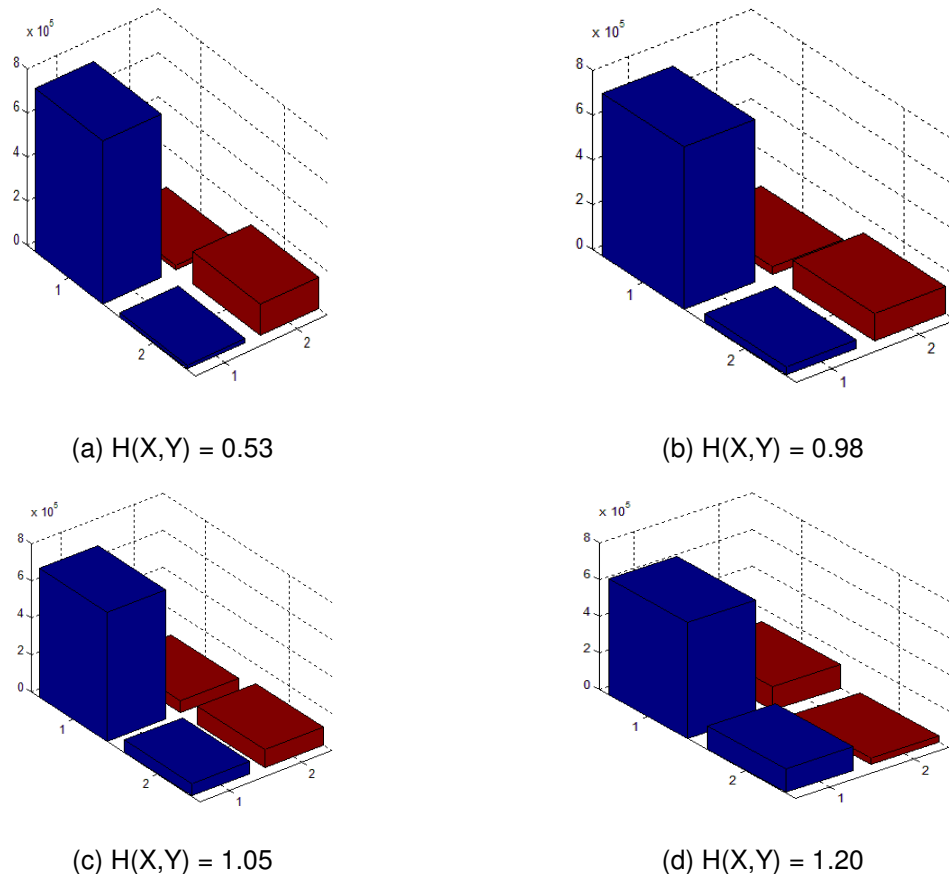


Figure 1. Joint grey values of a block image of a flotation froth and a shifted version of the image.

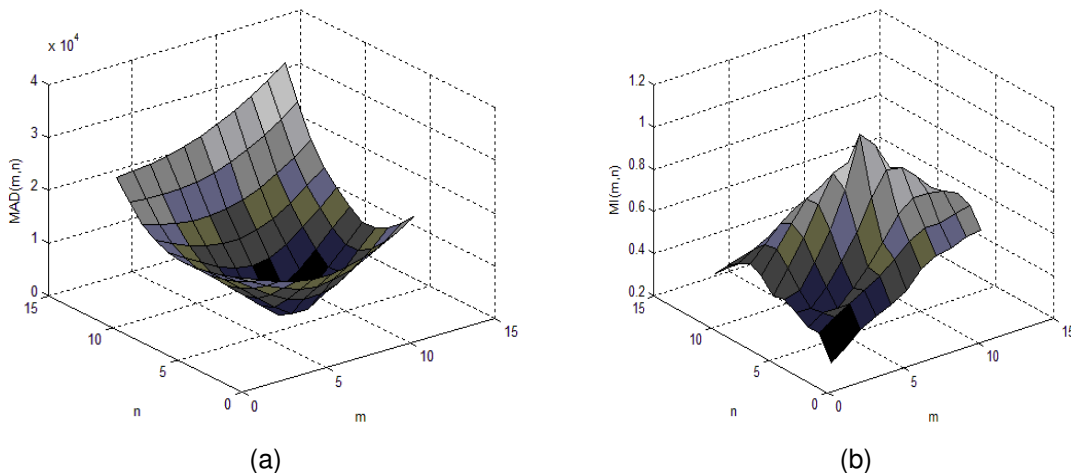


Figure 2. Example of (a) MAD surface of a block for a search area of 5 pixels (b) MI surface of the same block over a search area of 5 pixels

Mutual information contains the term $-H(A,B)$, which means that the minimization of the joint entropy will maximize the mutual information. Since generally joint entropy increases with

increasing image block mismatching, the mutual information decreases with increased mismatching, i.e. similarity corresponds with the maximization of the mutual information. That is, if images are aligned the amount of information they contain about each other is maximal.

The computational complexity between two block images depends not only on the number of pixels in each block image (M) but also the number of bins used the estimate MI . The computational cost of the histogram is $O(M)$. The computational cost relative to the number of histogram bins is $O(R^2)$, where R is the number of bins. In this work only two bins were used ($R = 2$), with a block size of 40 x 40 pixels, so that the computational cost of MAD and MI were approximately the same. The principle of maximization of mutual information is applied to the motion vector of blocks. The motion vector is obtained by equation

$$mv = \arg \max_{-p \leq m, n \leq p-1} MI(m, n) \quad (4)$$

Figures 2(a) and 2(b) respectively show the MAD surface and $MI2$ surface of same block over a search area of 5 pixels in each of the vertical and horizontal directions. The $MI2$ surface is sharper than the MAD surface, suggesting that $MI2$ can be more robust than MAD , since both surfaces are smooth.

2.2 Mean Absolute Difference (MAD)

The full MAD search of an image frame divided into $N \times N$ blocks, can be expressed as

$$MAD(m, n) = \sum_{i=0}^{i=N-1} \sum_{j=0}^{j=N-1} |F^t(i, j) - F^{t-1}(i + m, j + n)|, \quad -p \leq m, n \leq p \quad (5)$$

where $MAD(m, n)$ is the MAD computed for the displacement of (m, n) pixels. $F^t(i, j)$ and $F^{t-1}(i, j)$ are the grey level values of the current and previous frame respectively. The search area or the maximum displacement, p , allowed in the vertical and horizontal directions. The motion vector for a block is obtained by the equation

$$mv = \arg \min_{-p \leq m, n \leq p-1} MAD(m, n) \quad (6)$$

where mv is the motion vector of the vertical and horizontal displacements. The computational complexity of MAD is $O(M)$ additions, where M is the number of pixels in each frame

3. SEARCH METHODS

The three-step-search and the new-three-step-search algorithms were used in this investigation.

3.1 Three-Step-Search (TSS)

The TSS algorithm begins with search location considered as the centre and sets a step size depending on the search window. It searches eight locations around the centre. From the nine locations searched, the one with the best similarity metric is selected as the new centre. A new step size is selected, which is half of the original step size and a similar search is repeated for two more iterations. Figure 3(a) illustrates the TSS algorithm.

3.2 New-Three-Step-Search (NTSS)

The NTSS is a modified version of TSS and a centre biased search method. In the first step, 16 points are checked in addition to the search origin for the lowest or highest weight of the cost function. If the optimal cost is at the origin, the search is stopped and the motion vector is (0,0). If the optimal cost is at any of the 8 locations nearest to the origin, then the origin of the search is changed to that point and points adjacent to it are searched. Depending on which point it is, 5 points or 3 points might be checked, as shown in Figure 3(b). The location that gives the optimal cost is the best match and the motion vector is set to that location. On the other hand, if the optimal cost is associated with one of the 8 locations further away from the origin, then the TSS search method is followed. Thus although the NTSS method might need a minimum of 17 points for every macroblock in the image, it has a worst case scenario of 33 points that have to be searched.

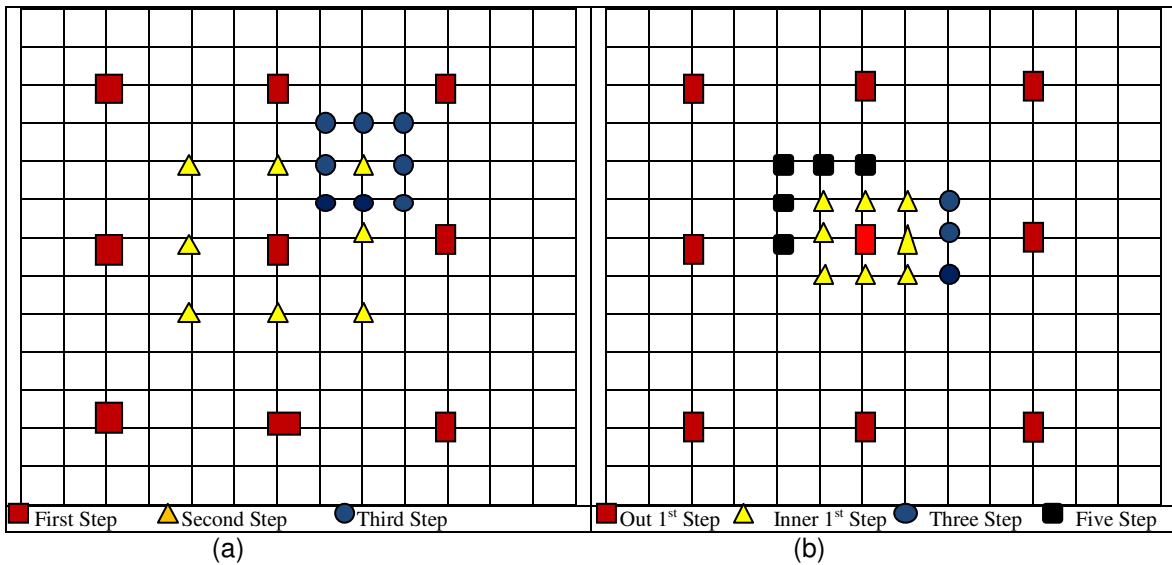


Figure 3(a). The Three Step Search algorithm with rectangles, triangles and circles representing first step search, second step search and third step search respectively, and (b) NTSS with rectangles as first outer step, triangle as inner first step, circles and squares as three step and five step respectively.

4. EXPERIMENTS AND RESULTS

The performance of the proposed similarity metric and MAD were evaluated using two video sequences “Flot1” and “Flot2” from the Anglo Platinum Centre of Excellence laboratory in Stellenbosch. The TSS and NTSS search strategies were used. The frame rate was 30 Hz and the resolution of both sequences was 720 x 1280 pixels. A 4 dB Gaussian was added to Flot2 to test the robustness of the algorithms, while a block size of 40 x 40 and a search area of 5 pixels were used in the vertical and horizontal directions. The peak signal to noise ratio (*PSNR*) of the image given in equation 6 is used to evaluate the performance of the algorithms

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad . (7)$$

Figures 4(a) and 4(b) show the motion vector estimate of frame 6 in the *Flot1* sequence using *MAD* with *TSS* and *MI2* with *TSS* respectively. Figures 4(c) and 4(d) also show the motion vector

estimate of frame 7 in the *Flot2* sequence using *MAD* with *TSS* and *MI2* with *TSS* respectively. Figure 5(a) and 5(b) show the comparisons of the *PSNR* of *MI2-with-NTSS*, *MI2-with-TSS*, *MAD-with-NTSS*, and *MAD-with-TSS* of the *Flot1* and *Flot2* sequences respectively. The performance of the algorithms in terms of the *PSNR* of the reconstructed frame for the *Flot1* sequence is almost the same. The performance of the *MI2* algorithms (*MI2-with-NTSS* and *MI2-with-TSS*) is better than the *MAD* algorithms (*MAD-with-NTSS* and *MAD-with-TSS*). The overall average *PSNR* of the experiments for the *MI2* algorithms and the *MAD* algorithms are 66.6 dB and 66.0 dB respectively.

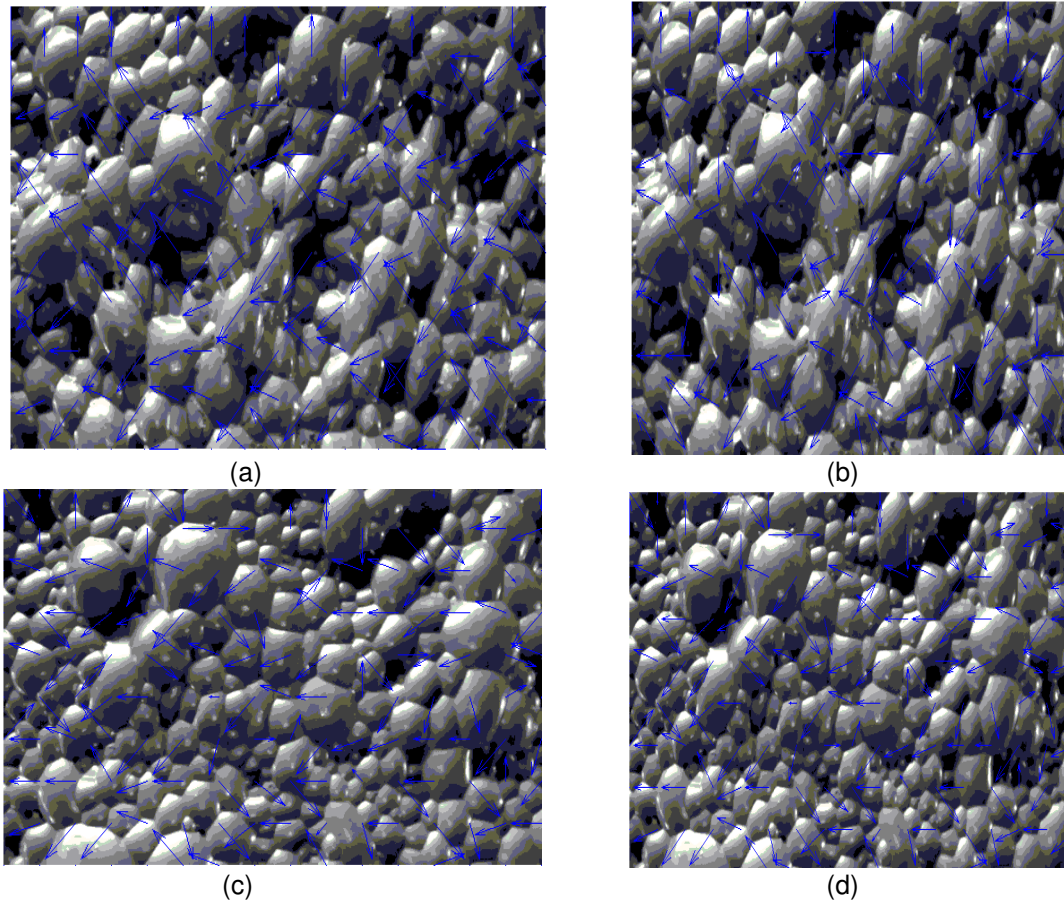


Figure 4(a). Motion vector of frame 6 in the *Flot1* sequence estimated by conventional *MAD* using *TSS*, (b) Motion vector of frame 6 in the *Flot1* sequence estimated by conventional *MI2* using *TSS* (c) Motion vector of frame 7 in the *Flot2* sequence estimated by the conventional *MAD* using *TSS*, and. (d) Motion vector of frame 2 in the *Flot2* sequence estimated by the conventional *MI2* using *TSS*.

Table.1 Average peak to signal noise ratio (*PSNR*) (dB) of the *Flot1* and *Flot2* sequences reconstructed by various motion estimation algorithms.

Method	Flot1	Flot2
MAD-with-TSS	67.34	64.67
MAD-with-NTSS	67.35	64.67
MI2-with-TSS	66.94	66.12
MI2-with-NTSS	67.00	66.24

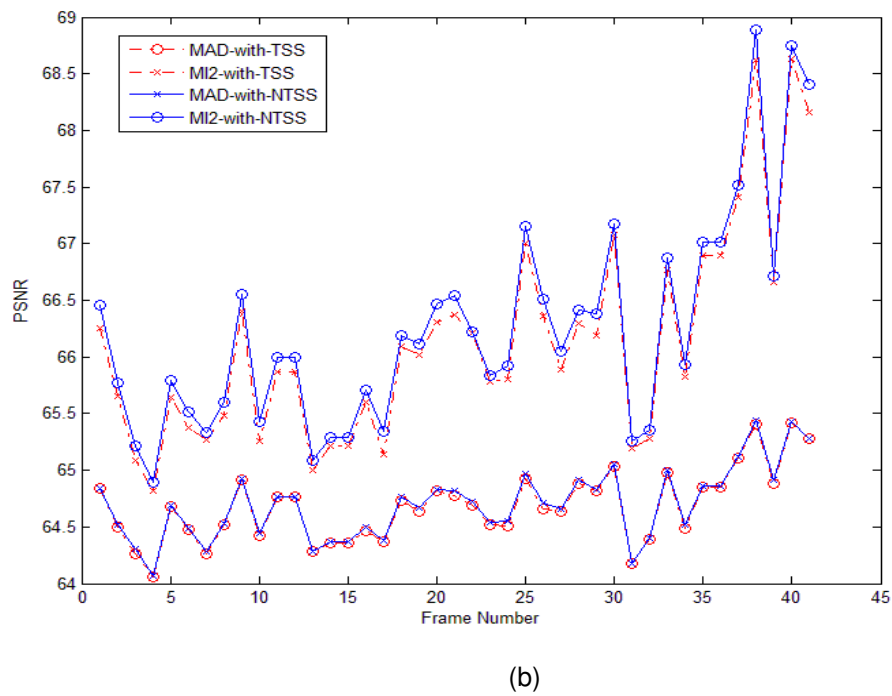
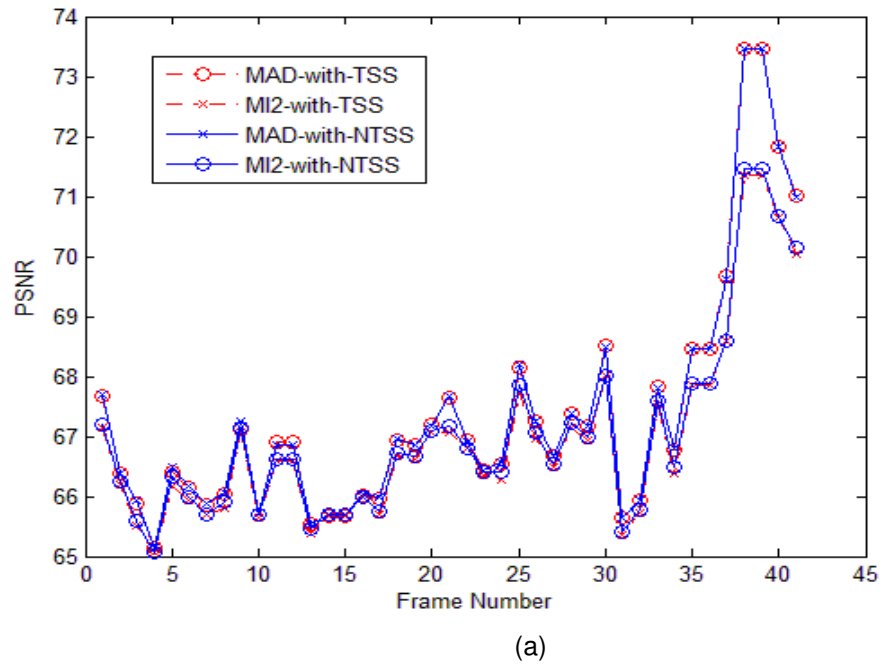


Figure 5(a). Comparisons of the peak signal to noise ratio (PSNR) of sequences reconstructed by various motion estimation algorithms. Using the Flot1 sequence with a resolution and frame rate of 720 x 1280 and 30Hz respectively, and (b) using the Flot2 sequence with a resolution and frame rate of 720 x 1280 and 30Hz respectively.

5. CONCLUSIONS

In this paper, a new motion estimation method for flotation froth is proposed by using mutual information with a bin size of two (*MI2*) as a similarity criterion. The computational cost is similar to the popular *MAD* approach. The three-step-search and the new-three-step-search algorithms were used to search the images. *MI2* yielded better results in terms of the PSNR of the reconstructed image frames than the classical minimum absolute difference (*MAD*). It reduces the amount of bad motion vectors especially for noisy images compared the *MAD* approach. In the future, the algorithm will tested with more video sequences in combination with different search strategies. We will also investigate optimization algorithms such as the Spall's algorithm [18] to reduce computational cost and the lightening systems for the experimental data.

6. REFERENCES

- [1] C. Aldrich, C. Marais, B.J. Shean and J.J. Cilliers. "Online monitoring and control of froth flotation systems with machine vision: A review". *International Journal of Mineral Processing*, 96(1-4), 1-13, 2010.
- [2] C. Marais and C. Aldrich. "Estimation of platinum grades from flotation froth images". *Minerals Engineering*, 24(5), 433-441, 2010.
- [3] C.P. Botha. 1999. An Online Machine Vision Froth Flotation Analysis Platform. M.Sc.thesis, University of Stellenbosch, Stellenbosch, South Africa.
- [4] D.P. Kottke and Y. Sun. "Motion estimation via cluster matching". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(11), 1128-1132, 1994.
- [5] N. Barbian, J.J. Cilliers, S.H. Morar, D.J. Bradshaw. "Froth imaging, air recovery and bubble loading to describe flotation bank performance". *International Journal of Mineral Processing* 84(1-4), 81–88, 2007.
- [6] K.K. Nguyen and A.J. Thornton. "The application of texture based image analysis techniques in froth flotation". In: Anthony Maeder & Brian Lovell (Eds.), *Proceedings of the DICTA-95, the 3rd Conference on Digital Imaging Computing Techniques and Applications*, 371–376, 1995, Brisbane, Australia, 6–8 December, St John's College, University of Queensland.
- [7] D.W. Moolman, C. Aldrich, J.S.J. van Deventer and W.W. Stange. "Digital image processing as a tool for on-line monitoring of froth in flotation plants". *Minerals Engineering* 7 (9), 1149–1164, 1994.
- [8] J.M. Jou, P.-Y. Chen, and J.-M. Sun, "The gray prediction search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 6, pp. 843–848, Sep. 1999.
- [9] T. Koga, K. Linuma, A. Hirano, Y. Lijima, and T. Ishiguro, "Motion-compensated interframe coding for video conferencing", in *Proc. NTC'81*, 1981, pp. G5.3.1–G5.3.5.
- [10] Renxiang Li, Bing Zeng, and Ming L. Liou, "A New Three-Step Search Algorithm for Block Motion Estimation", *IEEE Trans. Circuits And Systems For Video Technology*, vol 4., no. 4, pp. 438-442, August 1994.

- [11] L.M. Po and W.C. Ma, "A novel four-step search algorithm for fast block motion estimation", IEEE Trans. Circuits Syst. Video Technol., vol.6, no. 3, pp. 313–317, Jun. 1996.
- [12] J. Jain and A. Jain, "Displacement measurement and its application in internal image coding", IEEE Trans. Commun., vol. 29, no. COM–12, pp. 1799–1808, Dec. 1981.
- [13] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block matching motion estimation", IEEE Trans. Image Process., vol. 9, no. 2, pp. 287–290, Feb. 2000.
- [14] C.E. Shannon "A mathematical theory of communication," Bell System technical Journal, vol.27, pp 379-423, 1948.
- [15] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information", IEEE Trans. Med. Imag., vol. 16, Apr. 1997.
- [16] W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information", Med. Imag. Anal., vol. 1, pp. 35–51, 1996.
- [17] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration", IEEE Trans. Image Processing, vol. 9, pp. 2083–2099, Dec. 2000.
- [18] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," IEEE Trans. Automat. Contr., vol.37, no. 3, pp. 332–341, 1992.

Comparison of Some Motion Detection Methods in cases of Single and Multiple Moving Objects

Shamir Alavi

Electrical Engineering

National Institute of Technology Silchar

Silchar 788010 (Assam), India

alavi1223@hotmail.com

Abstract

Motion detection tells us whether there is a change in position of an object with respect to its surroundings or vice versa. It is applied to various domestic and commercial applications starting from simple motion detectors to high speed video surveillance systems. In this paper, results obtained from some simple motion detection algorithms, which use methods like image subtraction and edge detection, have been compared. The software used for this purpose was MATLAB 7.6.0 (R2008a). It has been observed that while image subtraction is sufficient to detect motion in a video stream, combining it with edge detection in different sequences yields different results in different scenarios.

Keywords: Motion Detection, Edge Detection, Canny, Sobel, Image Subtraction, Computer Vision

1. INTRODUCTION

Motion detection is a process of confirming a change in position of an object relative to its surroundings or the change in the surroundings relative to an object [1]. It has paramount importance in any vision based detection and tracking system. Throughout the last couple of decades, several techniques have been introduced to accomplish this task effectively. However, there is no perfect system or method which can overcome the various problems that are faced during detection. The difficulties are generally associated with lighting condition of the surrounding, illumination of the object itself which is to be detected, speed of its movement or the type of object [2].

Generally, motion detection is useful in real time or active video surveillance systems [2]. In this paper, the main focus is given to the processing of the captured video data to detect motion in it. Two main methods, image subtraction and edge detection have been used for detection. Three different cases have been considered in order to compare the results. One detects the motion by image subtraction only whereas the other two include edge detection in different sequence. Moreover, two different edge detection techniques, namely Sobel edge detection [3] and Canny edge detection [4] have also been taken into account while comparing the results. Finally, all of these are implemented on two different scenarios where the number of moving object is one or more. After a brief review of image subtraction and the two edge detection techniques, the motion detection algorithms used here along with comparisons in regard to the scenarios and edge detection techniques will follow accordingly.

2. Image Subtraction

Image subtraction is one of the popular techniques in image processing and computer vision. Basically image subtraction can be represented as:

$$\Delta I(i,j) = I_{Curr}(i,j) - I_{Prev}(i,j) \quad (1)$$

where:

$\Delta I(i,j)$ is the difference in image intensity between two consecutive frames. $I_{Curr}(i,j)$ and $I_{Prev}(i,j)$ represent image intensities for current and previous frames respectively [2].



FIGURE 1: (a) Frame 1 (b) Frame 2 (c) Image after subtraction

This is primarily done for one of two reasons – leveling uneven sections of an image such as half an image having a shadow on it, or detecting changes between two images. This detection of changes can be used to tell if something in the image moved [5].

3. Edge Detection

Edge detection refers to the process of identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensity which characterize boundaries of objects in a scene [6]. The purpose of edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing [7]. Classical methods of edge detection involve convolving the image with an operator (a 2-D filter), which is constructed to be sensitive to large gradients in the image while returning values of zero in uniform regions [6]. There are various ways to perform edge detection as various techniques have been introduced throughout the years. This work will compare two such techniques while detecting motion which are the Sobel operator or Sobel edge detection [3] and the Canny edge detection [4]. There are several other operators also such as Prewitt’s operator, Robert’s cross operator, Laplacian of Gaussian etc. but most of them (such as Prewitt’s and Roberts’s operator) work in a fashion similar to Sobel operator [6].

3.1 Sobel Operator

The operator consists of a pair of 3×3 convolution kernels as shown in Figure 1. One kernel is simply the other rotated by 90°.

-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

FIGURE 1: Masks used by Sobel Operator

These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (let us call these G_x and G_y). These can then be combined to find the absolute magnitude of the gradient at each point and the orientation of that gradient [3]. The gradient magnitude is given by:

$$|G| = \sqrt{G_x^2 + G_y^2} \tag{2}$$

Typically, an approximate magnitude is computed using:

$$|G| = |G_x| + |G_y| \quad (3)$$

which is much faster to compute.

The angle of orientation of the edge (relative to the pixel grid) giving rise to the spatial gradient is given by:

$$\theta = \arctan(G_y / G_x) \quad (4)$$

3.2 Canny Edge Detection

The algorithm runs in five separate steps [7]:

- i. **Smoothing:** Blurring the image to remove noise.
- ii. **Finding gradients:** The edges are marked where the gradients of the image has large magnitudes.
- iii. **Non-maximum suppression:** Only local maxima are marked as edges.
- iv. **Double thresholding:** Potential edges are determined by thresholding strong and weak edges.
- v. **Edge tracking by hysteresis:** Final edges are determined by suppressing all the edges that are not connected to a very certain (strong) edge.

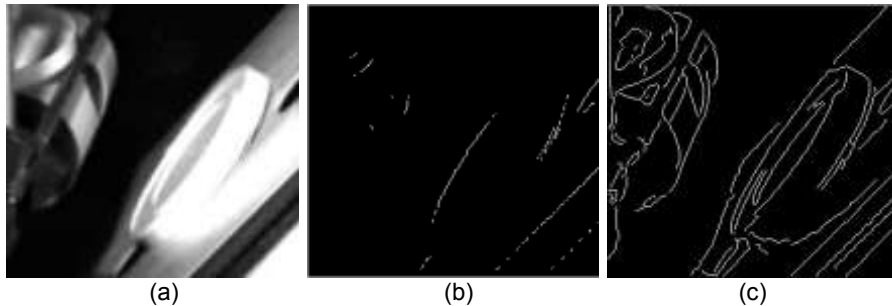


FIGURE 2: (a) Original image (b) Sobel edge (c) Canny edge

4. Algorithms Used for Motion Detection

Using image subtraction and edge detection as the main tool, following three algorithms have been used:

4.1 Image Subtraction Method

The first task is to extract frames from the continuous video stream so that they can be processed further for our next tasks. The steps for this algorithm are stated below:

- i. Extract frames from video stream.
- ii. Write the extracted frames as image files.
- iii. Subtract the previous image from current image as stated in (1).
- iv. Convert image to binary image.
- v. Label connected components.
- vi. Perform blob analysis (i.e. measure properties of each labeled image regions).
- vii. Calculate the centre of mass of each labeled region and label it (to detect as many moving elements as possible).
- viii. Play the labeled images as a continuous video stream to detect motion.

4.2 Edge Detection after Image Subtraction

In this algorithm, edge detection is performed after image subtraction. The algorithm is as follows:

- i. Extract frames from video stream.
- ii. Write the extracted frames as image files.
- iii. Subtract the previous image from current image.
- iv. Convert the image to grayscale.
- v. Detect edges.
- vi. Label connected components.
- vii. Perform blob analysis.
- viii. Calculate the centre of mass of each labeled region and label it.
- ix. Play the labeled images as a continuous video stream to detect motion.

4.3 Image Subtraction after Edge Detection

In this case, edge detection is performed before image subtraction. The algorithm is stated below:

- i. Extract frames from video stream.
- ii. Write the extracted frames as image files.
- iii. Detect edges in all the images.
- iv. Subtract the previous image from current image.
- v. Convert subtracted image to binary image.
- vi. Label connected components.
- vii. Perform blob analysis.
- viii. Calculate the centre of mass of each labeled region and label it.
- ix. Play the labeled images as a continuous video stream to detect motion.

5. Comparison of the Motion Detection Algorithms

5.1 Considered Scenarios

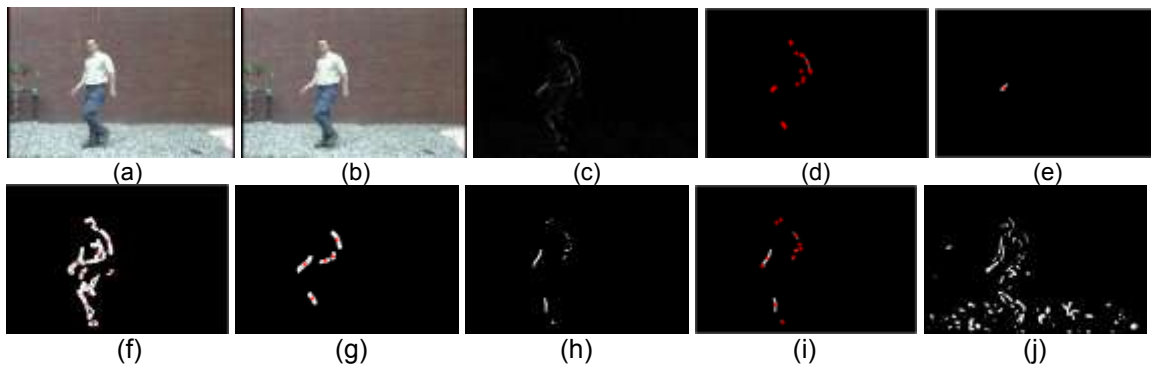
Two different scenarios have been considered for comparison:

Scenario 1: A man is walking while everything else is still. Therefore, we have a single moving object in this scenario (as in Figure 1)

Scenario 2: Plastic caps are being collected from one conveyor belt to another. Here, the caps as well as the belts are in motion. Hence, we have multiple moving objects in this scenario (as in Figure 2).

5.2 Visual Comparison

From the extracted frames, two consecutive frames have been taken for comparison and analysis:



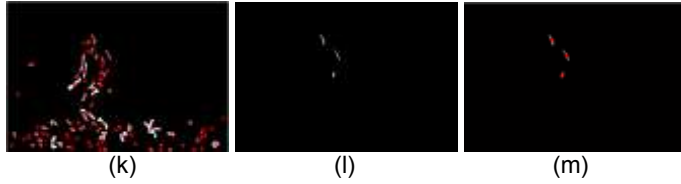


FIGURE 3: Comparison of motion detection algorithms on scenario 1, (a) and (b) two consecutive frames (c) subtracted image (d) algo 4.1 (e) algo 4.2 – Sobel (thresh = 0.17) (f) algo 4.2 – Canny (thresh = 0.17) (g) algo 4.2 – Canny (thresh = 0.55) (h) subtracted image of Sobel edges (i) algo 4.3 – Sobel (thresh = 0.17) (j) subtracted image of Canny edges (thresh = 0.17) (k) algo 4.3 – Canny (thresh = 0.17) (l) subtracted image of Canny edges (thresh = 0.55) (m) algo 4.3 – Canny (thresh = 0.55)
 algo = algorithm

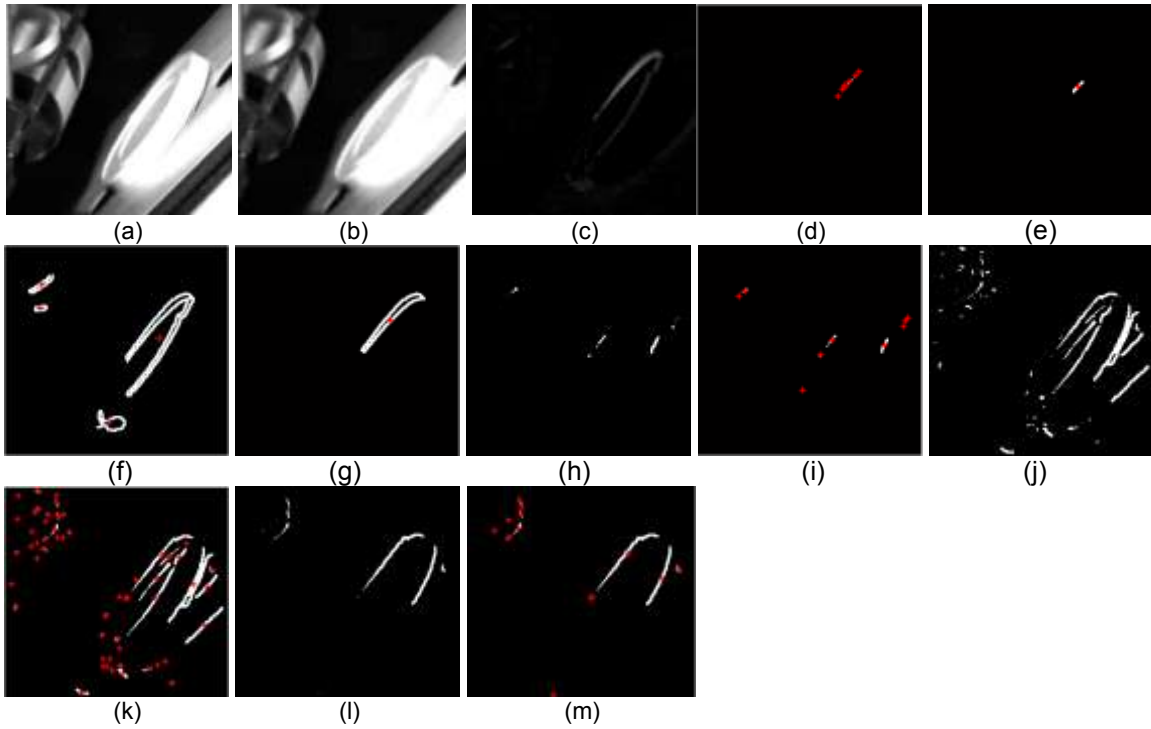


FIGURE 4: Comparison of motion detection algorithms on scenario 2, (a) and (b) two consecutive frames (c) subtracted image (d) algo 4.1 (e) algo 4.2 – Sobel (thresh = 0.17) (f) algo 4.2 – Canny (thresh = 0.17) (g) algo 4.2 – Canny (thresh = 0.55) (h) subtracted image of Sobel edges (i) algo 4.3 – Sobel (thresh = 0.17) (j) subtracted image of Canny edges (thresh = 0.17) (k) algo 4.3 – Canny (thresh = 0.17) (l) subtracted image of Canny edges (thresh = 0.55) (m) algo 4.3 – Canny (thresh = 0.55)

N.B.: For Canny edges, higher threshold = thresh, lower threshold = $0.4 * \text{thresh}$, standard deviation of the Gaussian filter = 1. For instance, in figure 4(j), higher threshold = 0.17, lower threshold = $0.4 * 0.17 = 0.068$, standard deviation of the Gaussian filter = 1.

5.3 Comparison Tables

The following table summarizes the visual comparisons shown above:

Algorithm	Edge	Scenario 1	Scenario 2
4.1		Detects motion quite well despite data loss due to binary conversion of the subtracted image - fig 3(d)	Partially detects motion of only one cap - fig 4(d).

4.2	Sobel	Minor visual detection – fig 3(e).	Minor visual detection – fig 4(e).
	Canny (thresh=0.17)	Great visual detection (detects movement of the whole body) – fig 3(f).	Unable to detect the motion of the belts (but better than Sobel as in fig 4(e)) – fig 4(f).
	Canny (thresh=0.55)	Detection of strong edges only (better than Sobel as in fig 3(e)) – fig 3(g).	Unable to detect the upper cap and the belts (still better than Sobel as in fig 4(e)) – fig 4(g).
4.3	Sobel	Detection quality almost similar to that of algo 4.1 -fig 3(i).	Partially detects both the caps but cannot detect movements of the belts – fig 4(i).
	Canny (thresh=0.17)	Detects movement of the man but erroneously detects several portions of the still pavement as moving objects – fig 3(k).	Great visual detection – detects movement of both the caps and the lower belt; slightly detects movement of the upper belt as well – fig 4(k).
	Canny (thresh=0.55)	Detection of strong edges only (no false detection) – fig 3(m). However, detection quality is lower than that of Sobel detection as in fig 3(i).	Unable to detect the motion of the belts – fig 4(m). However, detection quality is better than that of Sobel detection as in fig 4(i).

TABLE 1: Comparison of the motion detection algorithms under different scenarios

Average** time taken to obtain the above shown figures (only the ones in which motions are detected):

Algorithm	Edge	Scenario1***		Scenario 2****	
		Figure	Average time(sec)	Figure	Average time(sec)
4.1		3 (d)	3.0717644	4 (d)	4.1041660
4.2	Sobel	3 (e)	3.7338478	4 (e)	6.4473798
	Canny (thresh = 0.17)	3 (f)	4.0066358	4 (f)	6.4270532
	Canny (thresh = 0.55)	3 (g)	3.8619630	4 (g)	6.3946814
4.3	Sobel	3 (i)	7.5335376	4 (i)	13.8421282
	Canny (thresh = 0.17)	3 (k)	13.3262350	4 (k)	26.4577750
	Canny (thresh = 0.55)	3 (m)	12.6463506	4 (m)	25.0108116

TABLE 2: Comparison of average time taken by the algorithms to detect motion

** Each program has been run 5 times on the same system (Intel® Core(TM) i5 CPU M430 @ 2.27 GHz, 4GB DDR3 RAM, ATI Radeon HD 5400, Win 7 64-bit) and their average was considered.

*** Total number of frames in the video stream for this scenario = 80

**** Total number of frames in the video stream for this scenario = 249

6. Result and Discussion

From the comparisons shown above in table 1, it is clear that image subtraction only is good enough to detect motion in case of a single moving object in front of a still background. However, it fails to accomplish this task properly when there are multiple moving objects.

For a single moving object, the best result has been obtained by performing Canny edge detection after image subtraction where a low threshold value was used. Sobel edge detection could not perform as well as Canny edge detection. On the other hand, in case of multiple moving objects, the best result came from performing Canny edge detection before image subtraction. Here also, a low threshold value was used.

From table 2, we can see that 'Image Subtraction Method' is the fastest of the three motion detection methods that are compared here. Between the other two methods, performing edge detection after image subtraction is substantially faster. A careful look at the average times taken by this method also tells us that it is not that slow in comparison to the simpler image subtraction method (extra time taken by algo 4.2 as compared to 4.1 is below 1 second for scenario 1 and just above 2 seconds for scenario 2). For multiple moving objects, though Canny gave the best result, it came at the cost of high computation time (26.4577750 seconds).

7. Conclusion and Future Research

To sum up, edge detection, in addition to image subtraction is necessary to detect motion properly. Canny edge detection, in spite of being computationally slower and more expensive as a result, gives the best outcome under any circumstances. On top of it, 'Image Subtraction after Edge Detection' is the best method out of the three discussed above if we can compromise the higher computational time taken by it.

For further research, only Canny edge detection will be used as it has been deemed to work best among all other edge detection methods available at present [8, 9] and is referred to as a 'modern standard' [10]. It is also necessary to improve the higher computation times with Canny edge detection, especially in complex situations with multiple moving objects. This research work can be further extended by testing under different lighting conditions, differing the distance of the moving objects and finally going live with a faster algorithm and system.

8. References

- [1] Wikipedia. "Motion Detection." Internet: http://en.wikipedia.org/wiki/Motion_detection, May 20, 2012 [Jun. 16, 2012].
- [2] R. B. Wahyu, Tati R. Mengko, Bambang Pharmasetiawan and Andryan B. Suksmono. "Motion Detection Using Image Subtraction Edges Detection." *Risalah Lokakarya Komputasi dalam Sains dan Teknologi Nuklir XVII*, pp. 157-171, Aug. 2006.
- [3] J. Matthews. "An Introduction to Edge Detection: The Sobel Edge Detector." Internet: <http://www.generation5.org/content/2002/im01.asp>, 2002.
- [4] J. Canny. "A Computational Approach to Edge Detection." *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-697, Nov. 1986
- [5] Wikipedia. "Image subtraction." Internet: http://en.wikipedia.org/wiki/Image_subtraction, Sep. 27, 2011 [Jun. 16, 2012].
- [6] R. Maini and Dr. H. Aggarwal. "Study and Comparison of Various Image Edge Detection Techniques." *International Journal of Image Processing (IJIP)*, vol. 3, iss. 1, pp. 1-11, Feb. 28, 2009.
- [7] "Canny Edge Detection." Internet: http://www.cvmt.dk/education/teaching/f09/VGIS8/AIP/canny_09gr820.pdf, Mar. 23, 2009.

- [8] E. Nadernejad, S. Sharifzadeh and H. Hassanpour. "Edge Detection Techniques: Evaluations and Comparisons." *Applied Mathematical Sciences*, vol. 2, no. 31, pp. 1507 – 1520, 2008.
- [9] M. Juneja and P. S. Sandhu. "Performance Evaluation of Edge Detection Techniques for Images in Spatial Domain." *International Journal of Computer Theory and Engineering*, vol. 1, no. 5, pp. 614-621, Dec. 2009.
- [10] M. Heath, S. Sarkar, T. Sanocki and K. Bowyer. "Comparison of Edge Detectors: A Methodology and Initial Study." *Computer Vision And Image Understanding*, vol. 69, no. 1, pp. 38–54, Jan. 1998.

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. Starting with volume 7, 2013, IJIP will be appearing 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.

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: 2

i. Paper Submission: January 31, 2013

ii. Author Notification: March 15, 2013

iii. Issue Publication: April 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 © 2012
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

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