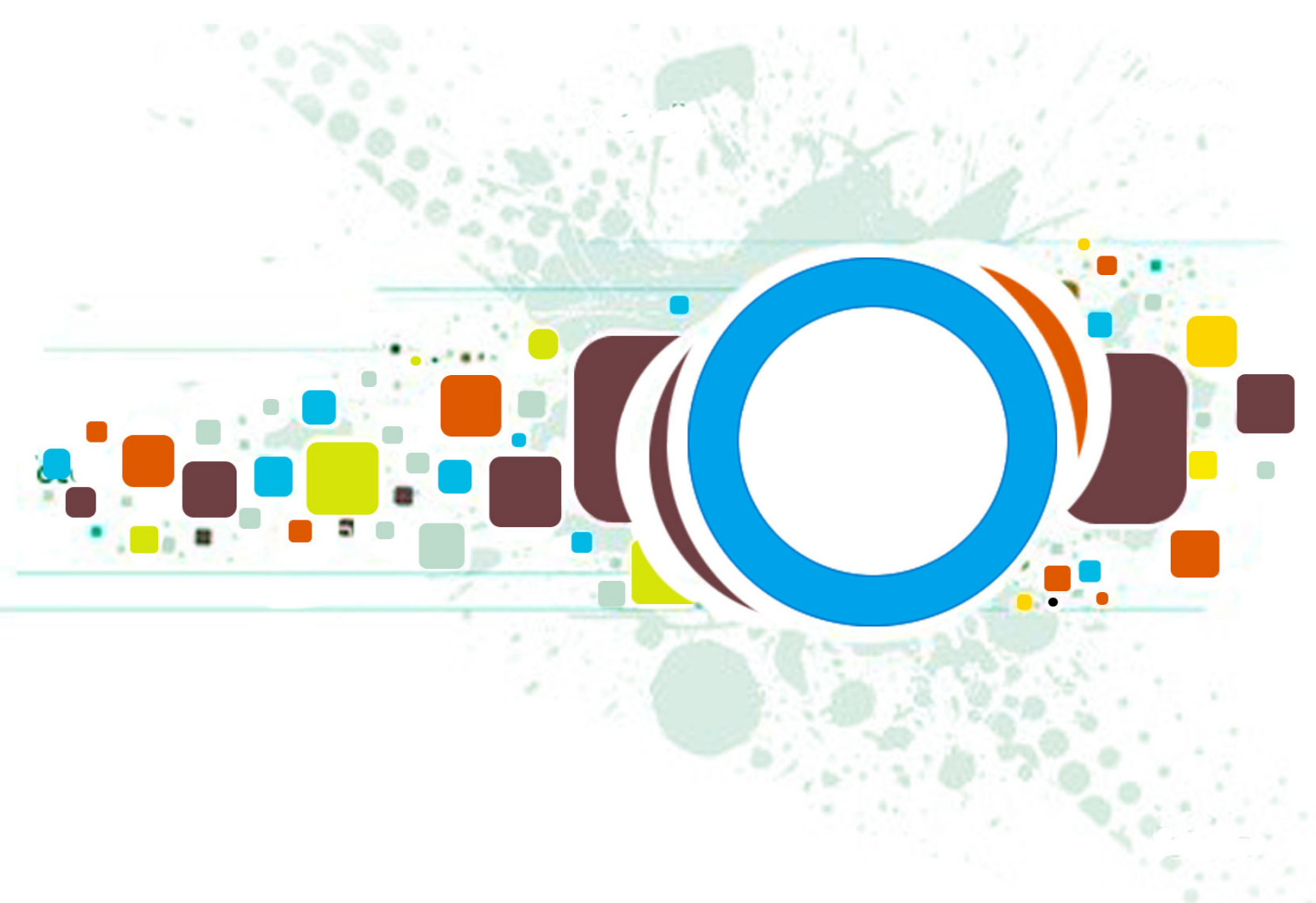


Volume 9 • Issue 6 • November / December 2015

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 9, ISSUE 6, 2015**

**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 9, Issue 6, November / December 2015

Publishing Date: 31-12-2015

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, 2015**

## 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 *Sixth* Issue of Volume *Nine* 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 10, 2016, IJIP will be appearing with more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

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

### ASSOCIATE EDITORS (AEiCs)

---

**Professor. Khan M. Iftekharuddin**  
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 9, Issue 6, November / December 2015

### Pages

- 304 - 310 Denoising Process Based on Arbitrarily Shaped Windows  
*Huda Al-Ghaib, Reza Adhami*
- 311 - 319 Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM  
*Abrham Debasu Mengistu, Dagnachew Melesew Alemayehu*
- 320 - 334 Textural Feature Extraction of Natural Objects for Image Classification  
*Vishal Krishna, Ayush Kumar, Kishore Bhamidipadi*
- 335 - 345 Novel Approach to Use HU Moments with Image Processing Techniques for Real Time Sign Language Communication  
*Matheesha Fernando, Janaka Indrajith Wijjayanayake*



# Denoising Process Based on Arbitrarily Shaped Windows

**Huda Al-Ghaib**

*Assistant Professor/Computer Science/Technology and Computing  
Utah Valley University  
Orem, 84058, USA*

*Huda.Ghaib@uvu.edu*

**Reza Adhami**

*Professor/Electrical and Computer Engineering  
University of Alabama in Huntsville  
Huntsville, 35899, USA*

*adhamir@uah.edu*

---

## Abstract

Many factors, such as moving objects, introduce noise in digital images. The presence of noise affects image quality. The image denoising process works on reconstructing a noiseless image and improving its quality. When an image has an additive white Gaussian noise (AWGN) then denoising becomes a challenging process. In our research, we present an improved algorithm for image denoising in the wavelet domain. Homogenous regions for an input image are estimated using a region merging algorithm. The local variance and wavelet shrinkage algorithm are applied to denoise each image patch. Experimental results based on peak signal to noise ratio (PSNR) measurements showed that our algorithm provided better results compared with a denoising algorithm based on a minimum mean square error (MMSE) estimator.

**Keywords:** Region Merging, Wavelet Transform, Image Denoising, Noise Estimation, Wavelet Shrinkage Process.

---

## 1. INTRODUCTION

Digital images have numerous applications in areas such as medical imaging, biometrics, robotics, and image navigation [1]. Some examples of medical imaging are computed tomography scan (CT), magnetic resonance imaging (MRI), ultrasound, X-ray, myocardial perfusion, and mammography. Medical imaging devices have helped physicians diagnose different diseases at early stages. In some cases, medical images may be noisy due to factors such as patient movement during the imaging process [2]-[4]. Noise presented in a digital image affects its quality. Noise occurs during image acquisition and/ or transmission. In most cases, noise originates from an unknown source and location. In this case, the noise is assumed to be an additive white Gaussian noise (AWGN) of unknown mean and variance. Accurate estimation of noise parameters is a preliminary step in a successful denoising process.

The main goal of the presented research is to provide a method to denoise digital images using accurate estimations of noise variance. The denoising process is applied on local image patches. First, an input image is subdivided into a number of patches. For each patch, the local variance is estimated and used to denoise the image using a wavelet shrinkage denoising algorithm. Noise estimation and elimination are executed in the wavelet domain.

## 2. MATERIALS and METHODS

### 2.1 The Proposed Algorithm

AWGN presented in digital images is of an additive nature. Differentiation of image information and noise is a challenging procedure. AWGN is characterized by its mean and variance. In most cases, AWGN is assumed to possess a mean of zero. In our research, wavelet transform is

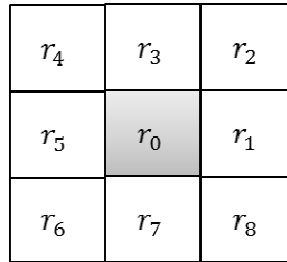
applied to decompose an input image of size  $M \times N$  into high and low frequency components. The magnitude of the vertical and horizontal details is computed to produce an image,  $MAG$ , of size  $M \times N$ .  $MAG$  is subdivided into a number of sub images. For each sub image, region merging is computed to estimate the local variance using only the homogenous regions [5]. The region merging algorithm is explained in detail in Section 2.2. A wavelet shrinkage process, explained in Section 2.3, is applied to denoise the sub image using the estimated variance [6].

**2.2 Local Noise Estimation Using The Region-Merging Algorithm**

An image  $I$  is subdivided into  $M \times M$  windows. Next,  $R$  window is subdivided into  $Q$  sub windows as shown in Figure 1. Where  $r_0, \dots, r_{Q-1}$  are the sub windows each of size  $m \times m$ .  $r_0$  is the sub window at the center in  $R$ . For each sub window in  $R$ , the following two conditions must be satisfied:

1.  $r_i \cap r_j = \emptyset$ , for  $i \neq j$ .
2.  $\cup_i r_i = R$ .

where  $M = 9$ ,  $m = 3$ , and  $Q = 9$ .



**FIGURE 1:** Region  $R$  of size  $9 \times 9$ .

The variance for the pixels within  $r_0$  is  $\sigma_0^2$  and is considered as the seed sub window in  $R$ . Other adjacent sub windows are merged with  $r_0$  if they pass the homogeneity test. The homogeneity test is computed as:

$$h_q = \frac{|\sigma_q^2 - \sigma_0^2|}{\sigma_0^2}, q = 0, 1, \dots, Q - 1 \tag{1}$$

where  $\sigma_q^2$  is the variance of  $r_q$ . Each sub window is assumed to be of zero mean and the variance is computed as:

$$\sigma_q^2 = \frac{1}{|c_q|} \sum_{y_m \in c_q} y_m^2 \tag{2}$$

where  $c_q$  is the coefficient within  $r_q$ . The following condition is applied to test  $h_q$ :

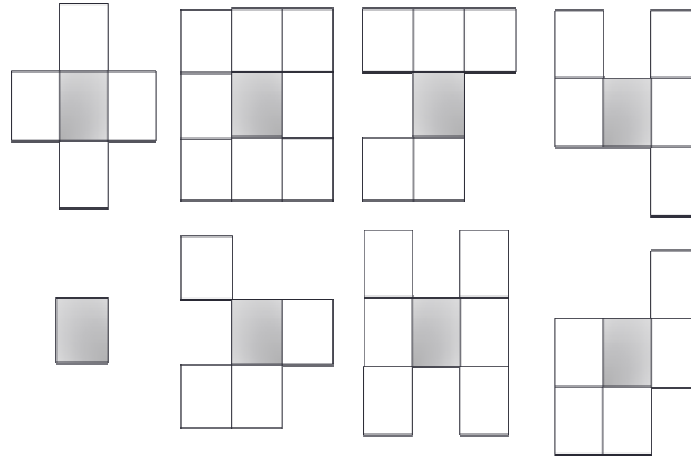
$$m_q = \begin{cases} 1, & \text{if } h_q < t \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where  $t = 0.2$ . If  $m_q = 1$  for a given  $r_q$ , then  $r_q$  is merged with  $r_0$ , otherwise it is discarded. The homogeneity test is performed on every sub window in  $R$ . As a result, final  $R$  is of an arbitrary shape. When  $Q = 9$ , there are  $2^{Q-1} = 256$  different configurations for  $R$ . Figure 2 shows some possible configurations for  $R$ .

The local noise variance for the merged sub windows is computed as:

$$\hat{\sigma}_k^2 = \left( \frac{\sum_{q=0}^{Q-1} \sigma_{k,q}^2 \cdot m_{k,q}}{\sum_{q=0}^{Q-1} m_{k,q}} - \sigma_n^2 \right) \quad (4)$$

where  $\sigma_n^2$  is the global noise variance estimated using a variation-adaptive evolutionary approach [5]. Equation (4) is applied on each image patch.



**FIGURE 2:** Local windows with arbitrary size and shape;  $r_o$  is the gray-scale region.

### 2.3 The Denoising Algorithm

A wavelet shrinkage denoising operator can be defined as [24]:

$$C(x) = \text{sign}(x) \begin{cases} |x| - T_n, & |x| \geq T_n \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Since  $x$  is the magnitude of the detail coefficients and is always  $\geq 0$ , equation (5) can be rewritten as:

$$C(M) = \begin{cases} M - T_n, & M \geq T_n \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

The function  $C(M)$  must satisfy the following two conditions:

1. Being a piece wise linear function.
2. Being a monotonically non-decreasing function.

$T_n$  is the threshold value. An accurate estimation for  $T_n$  is needed to have an efficient denoising process. Reference [8] suggested the following mathematical model to compute  $T_n$  as:

$$T_n = \sqrt{2 \ln(N)} \sigma / \sqrt{N} \quad (7)$$

$N$  is the signal length and  $\sigma$  is the standard deviation of the wavelet coefficients. However, equation (7) cannot be applied to our algorithm for the following reasons:

1. It uses a non-orthogonal UWT.
2. The shrinkage operation is applied to the magnitudes of the gradient coefficients instead of the wavelet coefficients.

For a white Gaussian noise, the probability distribution function of the magnitude of gradients is characterized by the Rayleigh distribution as [8]:

$$Pr_{||\Delta f||}(m) = \begin{cases} \frac{m}{\eta^2} e^{-\frac{(m)^2}{\eta^2}} / 2, & m \geq 0 \\ 0, & m < 0 \end{cases} \quad (8)$$

As a result, there is a direct relationship between  $\sigma$  and  $\eta$ , where  $\sigma$  and  $\eta$  are the standard deviations for the Gaussian and Rayleigh distributions respectively. Thus, equation (7) can be rewritten as:

$$T_n = \eta \sqrt{-2 \ln(1 - p)} \quad (9)$$

$p$  is the probability of noise removal for a particular threshold  $T_n$  and is computed as:

$$p = \frac{\int_0^{T_n} Pr_{||\Delta f||}(m) dm}{\int_0^{\infty} Pr_{||\Delta f||}(m) dm} \quad (10)$$

Based on Equation (10),  $T_n = 3.7169\eta$  for  $p = 0.999$ , and  $T_n = 4.5\eta$  for  $p = 0.99996$ . In [6] and [8]  $\eta$  is computed using the histogram of  $||\Delta f||$  and an iterative curve fitting function. In our paper we applied a novel approach based on ACO to achieve an accurate estimation of  $\eta$  [7].

#### 2.4 Error Rate for $\sigma$ Estimation

A set of images was used as experimental input to test the noise variance estimation algorithms. These images are: Lena, cameraman, barbara, kodim05, kodim06, kodim07, kodim08, kodim21, and kodim24, Figure 3 shows the input images. AWGN with different standard deviation values were added to the input images, i.e.,  $\sigma = 80.6, 71.4, 57, 51$  and 25.5 respectively. This produced images of high noise density. Noise variance is estimated using a variation-adaptive evolutionary approach [7]. Table I displays the averaged normalized values for the noise variance, i.e., added and estimated. Table II displays the averaged error rate for the estimated noise variance for the input images. Even though the added noise possesses high standard deviations, it is obvious from Tables I and II that the estimation process is of low error rate.



FIGURE 3: Input Images.

**TABLE 1:** Noise Variance Estimation.

$\sigma^2$ (Added)	0.1	0.078	0.05	0.04	0.01
$\sigma^2$ (Estim.) ACO	0.082	0.068	0.059	0.048	0.028

**TABLE 2:** Mean and variance of error rate: Comparison for different noise estimation algorithms.

$\sigma^2$	Mean of error rate	Variance of error rate
0.1	0.0197	1.2419e-005
0.0784	0.0108	1.2535e-005
0.05	0.0034	1.1162e-005
0.04	0.0055	3.6346e-005
0.01	0.0035	5.3444e-006

### 3. EXPERIMENTAL RESULTLS

The noisy images, along with the estimated noise variance, are used as inputs for the denoising algorithm. Wavelet transform is applied to decompose the input image. The algorithm is performed on the magnitude of the horizontal and vertical details. The algorithm explained in Section 2.3 is implemented to estimate the local variance for each image patch. The local denoising process is applied using the denoising algorithm explained earlier in Section 2.3. Denoising is performed twice using a minimum mean square error (MMSE) estimator and wavelet shrinkage algorithm respectively. The peak signal to noise ratio (PSNR) is computed for the noisy and denoised images, and the results are presented in Table III. Notations (1) and (2) in Table III illustrate wavelet-based denoising based on a MMSE estimator and our algorithm respectively. From Table III it is obvious that our improved algorithm provided better results compared with the original algorithm based on using a MMSE estimator [5]. These results showed that MMSE provided slightly better results in only 10 cases compared with our algorithm. Our algorithm accuracy is 78.78%.

### 4. CONCLUSION

This research provided an improved denoising algorithm based on wavelet shrinkage operation. A region merging algorithm is developed in the wavelet domain to locate the homogenous regions. The local homogenous regions are applied to estimate the local variance to denoise the region using the region merging algorithm. Experimental results based on PSNR illustrated in Table III showed that our improved algorithm provided better results compared with minimum mean square error (MMSE) in 78.78% of the cases.

### 5. REFERENCES

- [1] H. Al-Ghaib and R. Adhami, *An E-learning interactive course for teaching digital image processing at the undergraduate level in engineering*, Interactive Collaborative Learning (ICL), 2012 15th International Conference on, vol.1, no.5, pp.26-28, September 2012.
- [2] R. Eisenberg and A. Margulis. *A Patient's Guide To Medical Imaging*. New York, Oxford University Press, 2011.
- [3] U. Bhowmik, M. Iqbal, and R. Adhami, Mitigating motion artifact in FDK based 3D Cone-beam Brain Imaging System using markers, *Central European Journal of Engineering*, vol.2, no.3, pp.369-382.
- [4] U. Bhowmik and R. Adhami, A Novel Technique for Mitigating Motion Artifacts in 3D Brain Imaging System, *Scientific Iranica*, vol.20, no.3, pp.746-759, March 2013.

- [5] I. Kyu and Y. Kim, Wavelet-based denoising with nearly arbitrarily shaped windows, *Signal Processing Letters, IEEE*, vol.11, no.12, pp.937-940, December 2004.
- [6] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing, *Instrumentation and Measurement, IEEE Transactions on*, vol.57, no.7, pp.1422-1430, July 2008.
- [7] J. Tian and L. Chen, Image Noise Estimation Using A Variation-Adaptive Evolutionary Approach, *Signal Processing Letters, IEEE* , vol.19, no.7, pp.395-398, July 2012.
- [8] J. Fan and A. Laine, Contrast enhancement by multiscale and nonlinear operators, *Wavelets in Medicine and Biology.*, pp. 163-192. Boca Raton, FL: CRC, March 1996.

**TABLE 3:** PSNR for Noisy and Denoised Images.

$\sigma^2$ (Added)	0.1	0.078	0.05	0.04	0.01
Estimated $\sigma^2$ , Lena	0.08	0.07	0.046	0.039	<b>0.014</b>
PSNR -noisy Lena	14.25	15.04	16.58	17.41	23.07
PSNR- denoised Lena(1)	23.02	23.09	25.94	25.31	31.5
PSNR- denoised Lena(2)	20.93	20.15	20.22	22.99	27.83
Estimated $\sigma^2$ , cameraman	0.082	0.067	<b>0.05</b>	0.056	0.022
PSNR- noisy cameraman	14.34	15.2	16.61	17.42	23.05
PSNR- denoised cameraman (1)	22.08	22.47	24.69	23.94	29.25
PSNR- denoised cameraman (2)	19.17	20.06	22.04	22.04	26.36
Estimated $\sigma^2$ , barbara	0.08	0.07	0.05	<b>0.04</b>	0.02
PSNR- noisy barbara	14	14.78	16.4	17.27	23.03
PSNR- denoised barbara (1)	23.17	23.6	25.16	25.93	<b>32.58</b>
PSNR- denoised barbara (2)	20.91	22.64	21.25	23.12	28.32
Estimated $\sigma^2$ , kodim05	0.08	0.067	0.047	0.05	0.037
PSNR-noisy kodim05	13.53	14.47	16.28	17.17	23.08
PSNR-denoised kodim05 (1)	<b>25.02</b>	<b>24.61</b>	<b>26.62</b>	<b>27.1</b>	27.86
PSNR-denoised kodim05 (2)	23.27	22.95	24.76	24.32	27.51
Estimated $\sigma^2$ , kodim06	0.09	0.07	0.054	0.04	0.02
PSNR- noisy kodim06	14.67	15.5	17.7	17.95	23.56
PSNR-denoised kodim06 (1)	19.03	20.45	21.95	21.43	23.56
PSNR-denoised kodim06 (2)	16.89	18.64	19.03	19.95	21.66
Estimated $\sigma^2$ , kodim07	0.08	0.068	0.048	0.041	0.016
PSNR- noisy kodim07	13.83	14.6	16.25	17.15	23.03
PSNR- denoised kodim07 (1)	23.55	26.89	26.4	27.1	28.04
PSNR- denoised kodim07 (2)	21.99	21.84	22	24.48	27.51
Estimated $\sigma^2$ , kodim08	<b>0.09</b>	0.07	0.06	0.06	0.06
PSNR- noisy kodim08	23.42	15.28	16.92	17.8	23.45
PSNR- denoised kodim08 (1)	23.21	20.21	20.4	21.85	22.73
PSNR- denoised kodim08 (2)	17.74	18.84	19.1	18.55	21.37
Estimated $\sigma^2$ , kodim21	0.08	<b>0.07</b>	0.05	0.04	0.03
PSNR- noisy kodim21	13.96	14.77	16.39	17.2	23.1
PSNR- denoised kodim21 (1)	23.42	23.54	23.76	25.17	27.5
PSNR- denoised kodim21 (2)	20.5	20.96	22.7	22.96	23.57
Estimated $\sigma^2$ , kodim24	0.08	0.07	0.05	0.04	0.025
PSNR-noisy kodim24	13.92	14.79	16.53	17.43	23.35
PSNR- denoised kodim24 (1)	21.55	21.84	22.17	22.41	23.43
PSNR- denoised kodim24 (2)	19.01	20.4	19.77	20.63	22.12

# Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM

**Abrham Debasu Mengistu**

*Bahir Dar University, Bahir Dar institute of technology  
Faculty of computing  
Bahir Dar, Ethiopia*

*abrhamd@bdu.edu.et*

**Dagnachew Melesew Alemayehu**

*Bahir Dar University, Bahir Dar institute of technology  
Faculty of computing  
Bahir Dar, Ethiopia*

*dagnachewm@bdu.edu.et*

---

## Abstract

Human skin is the largest organ in our body which provides protection against heat, light, infections and injury. It also stores water, fat, and vitamin. Cancer is the leading cause of death in economically developed countries and the second leading cause of death in developing countries. Skin cancer is the most commonly diagnosed type of cancer among men and women. Exposure to UV rays, modernize diets, smoking, alcohol and nicotine are the main cause. Cancer is increasingly recognized as a critical public health problem in Ethiopia. There are three type of skin cancer and they are recognized based on their own properties. In view of this, a digital image processing technique is proposed to recognize and predict the different types of skin cancers using digital image processing techniques. Sample skin cancer image were taken from American cancer society research center and DERMOFIT which are popular and widely focuses on skin cancer research. The classification system was supervised corresponding to the predefined classes of the type of skin cancer. Combining Self organizing map (SOM) and radial basis function (RBF) for recognition and diagnosis of skin cancer is by far better than KNN, Naïve Bayes and ANN classifier. It was also showed that the discrimination power of morphology and color features was better than texture features but when morphology, texture and color features were used together the classification accuracy was increased. The best classification accuracy (88%, 96.15% and 95.45% for Basal cell carcinoma, Melanoma and Squamous cell carcinoma respectively) were obtained using combining SOM and RBF. The overall classification accuracy was 93.15%.

**Keywords:** SOM, RBF, KNN, Digital Image Processing, Dermofit.

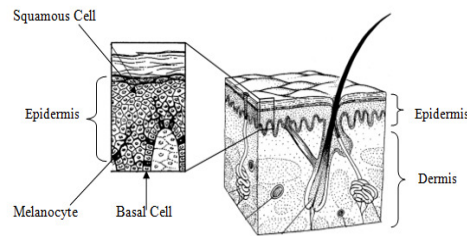
---

## 1. INTRODUCTION

Human Skin Cancer is a disease that appears on the outer layers of the skin which are caused when the skin cells are dead or damaged due to over exposure to Sun's ultraviolet radiation. But Skin cancer can also occur on areas of one's skin not ordinarily exposed to sunlight. The human skin is the largest organ in our body which provides protection against heat, light, infections and injury. It also stores water, fat, and vitamin D. [1] The Human skin is composed of two major layers called epidermis and dermis. The top or the outer layer of the skin which is called the epidermis composed of three types of cells flat and scaly cells on the surface called SQUAMOUS cells, round cells called BASAL cells and MELANOCYTES, cells that provide skin its color and protect against skin damage. The inner layer of the skin known as the dermis is the layer that contains the nerves, blood vessels, and sweat glands. There are three type of skin cancer Melanoma, Basal cell carcinoma and Squamous cell carcinoma. Skin cancer is diagnosed by physical examination and biopsy. Biopsy is a quick and simple procedure where part or all of the spot is removed and sent to a laboratory. It may be done by doctor or you can be referred to a



dermatologist or surgeon. Results may take about a week to be ready [1]. Dermatology imaging research believes that diagnosis of skin cancer can be automated based on a certain physical feature and color information which is the characteristics of different category of skin cancer.



**FIGURE 1:** Layers of Skin.

## 2. LITERATURE REVIEW

Regarding with the benefits of early detection of human skin cancer, several dermatologists, medical professionals, medical industries, clinicians, computer scientists, academic researchers and technical experts have dedicated time and efforts to improve the early detection of human skin cancers. And some researchers have been done on the application on neural classifiers to skin injury classification purposes on this research paper presents a computer vision approach based on image processing algorithms and supervised learning techniques to help detecting and classifying wound tissue types which play an important role in wound diagnosis. The system proposed involves the use of the k-means clustering algorithm for image segmentation and a standard multilayer perceptron neural network to classify effectively each segmented region as the appropriate tissue type. [11]. Common classification method like statistical and rule based ones were applied in the researches to describe the diagnostic performance of Solar Scan, an automated instrument for the diagnosis of primary melanoma Images from a data set of 2430 lesions (382 were melanomas) were divided into a training set and an independent test set Solar Scan is a robust diagnostic instrument for pigmented or partially pigmented melanocytic lesions of the skin. [11].

More advanced techniques such as Neural Network were presented in the research [12] The aim of the study was to provide mathematical descriptors for the border of pigmented skin lesion images and to assess their efficacy for distinction among different lesion groups. New descriptors such as lesion slope and lesion slope regularity are introduced and mathematically defined descriptors was tested on a data set of 510 pigmented skin lesions, composed by 85 melanomas and 425 nevi, by employing statistical methods for discrimination between the two populations.

K-nearest neighborhood as another classification method was employed in the research of [12] on this research paper presents an algorithm for classification of non melanoma skin lesions based on a novel hierarchical K- Nearest Neighbors (KNN) classifier. The KNN classifier here, skin lesions are characterized by their color and texture. Finally, towards identification of human skin cancer uses the following common steps is image acquisitions, preprocessing, Segmentation, feature extraction, classification and finally the result will display to the user. [13] In the recent years computational vision based diagnostic systems for dermatology have demonstrated significant progress. In this work, they review these systems by firstly presenting the visual features used for skin lesion classification and methods for defining them. Then they describe how to extract these features through digital image processing methods, i.e., segmentation, registration, border detection, color and texture processing) [14]. However, these imaging technologies are still expensive and may require specialized training to read the resulting images. Dermoscopy is the methodology for the examination of skin injuries based on imaging. This method provides a good and detailed view of the injuries. The imaging equipment used for taking the images is called Dermatoscope. It is handheld equipments which is compact and easy

to use. An oil film is placed between the lens of dermoscope and skin injuries. Main purpose of placing oil film is to obtain the magnified view of skin tissues [15].

The conventional diagnosing method for skin cancer is biopsy. It is a painful and time consuming technique. By incorporating Artificial intelligence and Digital Image Processing for skin cancer detection, it is possible to do the diagnosis without any physical contact with the skin. This can be implemented in a computer with the help of some software. Skin cancer detection system implemented using computer and software is known as Computer Aided Detection. The detection system is mainly based on Artificial intelligence and Digital Image Processing. Artificial intelligence has proven to be very efficient in decision making and pattern recognition applications. In this paper, the ANN Classifier is implemented in MATLAB software for skin cancer detection [16]. First stage in the skin cancer detection system is the input image. Dermoscopic image in digital format is given as input to the system. Dermoscopic image in digital format is given as input to the system. Next stage is the noise removal. The image contains hairs and other noises. These noises cause errors in classification. The noises are removed by filtering. Filtering method implemented here is the Median Filtering [16].

### **3. PROBLEM STATEMENT**

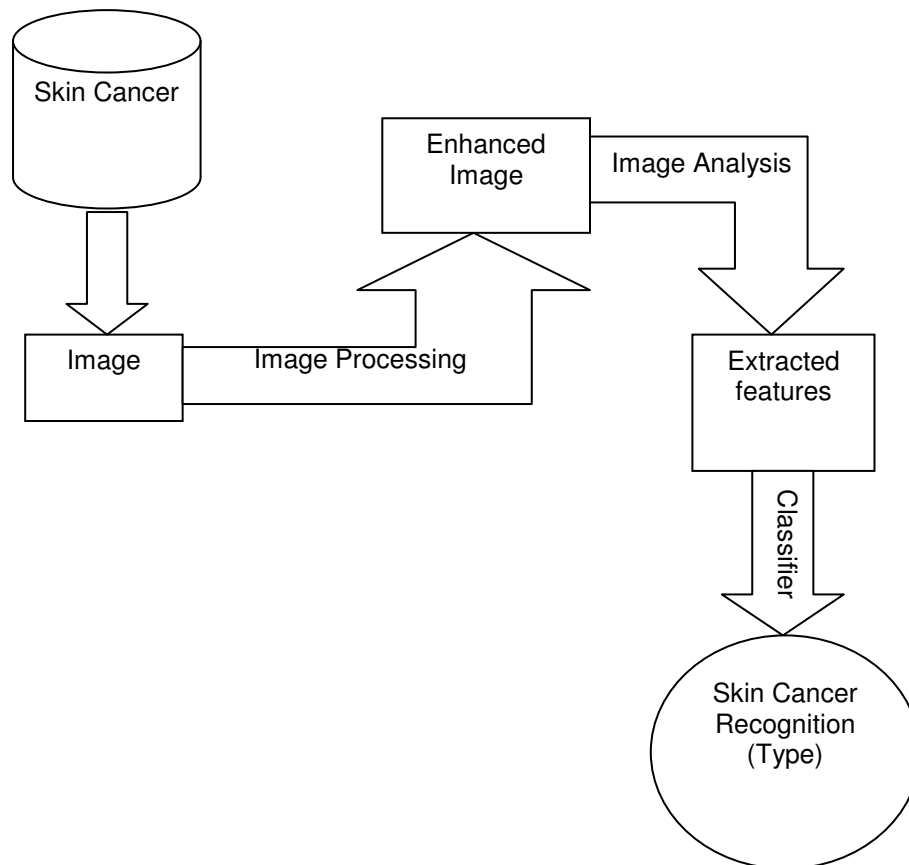
On the previous researches there is a scope for the design of classifier to detect the type of cancer this provide a better and more reliable results for the patients, so that more patients can be diagnosed and cured. In line with this, human skin cancer identification is very useful in encouraging good quality in skin cancer diagnosis. There is a need for automated in recognition of human skin cancer systems so that the abuses during diagnosis and treatment can be minimized. Therefore, this thesis work will initiate a model for human skin cancer recognition which is consistent, efficient and cost effective by exploring the technology of image processing techniques. The ultimate goal is to ease the doctor's role in the recognition of skin cancer mentioned above by providing better and more reliable results, so that more patients can be diagnosed. The work on classifier design to detect the type of cancer will be taken in future [31]. Skin cancer is diagnosed by physical examination and biopsy. In case of physical examination the doctors will try to see the physical properties of skin cancers. When we see biopsy it is the procedure that the dermatologist takes some part or all of the spot and sent to a laboratory. It may be done by doctor or you can be referred to a dermatologist or surgeon. Results may take about a week to be ready [1]. To this end this study answers the following research question:

- ✓ What appropriate image processing techniques used for human skin cancer recognition?
- ✓ To what extent recognition effectiveness is registered for the human skin cancer?
- ✓ What are the features that distinguish the three type of skin cancer?
- ✓ What are the common features that the three type of skin cancer shares?
- ✓ How to develop an automatic skin cancer recognition system based on image processing techniques?

### **4. DESIGN OF HUMAN SKIN CANCER RECOGNITION**

The task of recognition occurs in wide range of human activity. The problem of recognition is concerned with the construction of a procedure that will be applied to differentiate items, in which each new item must be assigned to one of a set of predefined classes on the basis of observed attributes or features.

Accordingly, image analysis or computer vision is used in the recognition of human skin cancer to predefined classes. The predefined classes are the feature or attributes are computed from skin cancer images. These observed features of skin cancer were used to decide the class or the type of skin cancer. Hence, in this research the main interest is to differentiate the type of skin cancer varieties by using image analysis technique this is because of in order to maximize the curability of the disease if we identify the type of skin cancer where it belongs to it is very simple to cure the disease otherwise it is difficult.



**FIGURE 2:** Skin Cancer Recognition Process Model.

As shown in the above figure 2, classification of human skin cancer involves the following activities:

Image acquisition of human skin cancer, an image processing techniques is applied on the acquired image to enhance the quality of image so as to reduce noises, appropriate feature are extracted from the enhanced image by using image analysis techniques that is used to classify dermoscopy images of skin cancer, the classification model of training and testing data of dermoscopy images of skin cancer will be developed, finally suitable pattern classifiers are selected to classify dermoscopy images of skin cancer to the predefined classes of skin cancer.

#### **A. Image Acquisition**

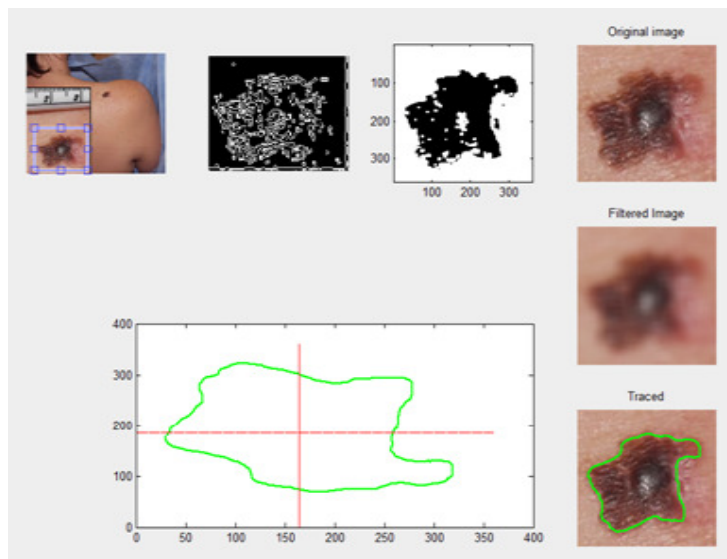
Image analysis starts with image acquisition this involves all aspects that have to be addressed in order to obtain dermoscopy image of human skin cancer the selection of radiation (light) sources and sensors (such as cameras) has to be considered very carefully. For this study, images have been taken from the following websites:

- ✓ <https://www.dermquest.com/image-library/>
- ✓ <http://www.dermnet.com/images/>

#### **B. Image Processing**

Image processing is a mechanism that focuses on the manipulation of images in different ways in order to enhance the image quality. Images are taken as the input and output for image processing techniques it is the analysis of image to image transformation which is used for the enhancement of image i.e. to increase the contrast for the input image and also restoration for geometrical distortion. [10]. Image segmentation is one of the most important tasks in image processing. It is the process of dividing an image into different homogeneous regions such that

the pixels in each partitioned region possess an identical set of properties or attributes [10]. The result of segmentation is a number of homogeneous regions, each having a unique label. Image segmentation is basically used to isolate region of interest from the background noise. For image processing techniques, we have used Matlab R2013a in which MATLAB is a high performance language for technical computing. MATLAB (MATRIX LABORATORY) is an interactive system for matrix based computation designed for scientific and engineering use. It is good for many forms of numeric computation and visualization. Hence, MATLAB was used for image processing tasks of Human skin cancer images to enhance the quality of image and to change images to binary for feature extraction purposes. From the original skin cancer images, the image is filtered in order to avoid other noises that are formed due to illumination effects as shown figure below.



**FIGURE 3:** Image Processing.

### C. Median Filtering

The Dermoscopic Image in Digital format is subjected to various Digital Image Processing Techniques. The standard image size is taken as 360x360 pixels [27]. Usually the image consists of noises in the form of hairs, bubbles etc. These noises cause inaccuracy in classification. In order to avoid that, images are subjected to various image processing techniques. One of the key element in image processing is filtering of image pre processing is done to removes the noise, fine hair and bubbles in the image. For smoothing image from noise, median filtering is used. Median filtering is a common step in image processing. Median filtering is used for minimizing the influence of small structures like thin hairs and isolated islands of pixels like small air bubbles [30].

### D. Feature Extraction

Feature extraction is the method by which unique features of skin lesion images are extracted. This method reduces the complexity in classification problems. The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. We have the following three groups of features:

GLCM (Texture features of skin cancer): GLCM is a powerful tool for image feature extraction by mapping the grey level co occurrence probabilities based on spatial relations of pixels in different angular directions.

Morphological features: Morphology is the geometric property of a given image, in our case it is the size and shape characteristics of human skin cancer image.

Color features: Color is one of the features of skin cancer, they have different color variation of each cancer type and color analysis computed by taking the mean value of RGBs (Red, Green and Blue) components and the mean value of HSIs (Hue, Saturation and Intensity) components. Therefore, to compute the mean value of each component of these color spaces, we use matlab 2013 to split each component because matlab has built in function to convert to HIS or RGB color spaces. By using the built in function of matlab we can find RGB, the RGB color image stack is split to red, green and blue components. Hence, the color features are extracted by computing the mean values of RGBs and HSIs of Dermoscopy skin cancer images. That is, the mean value of red, mean value of green, mean value of blue, mean value of hue, mean value of saturation and mean value of intensity are computed from each component.

### 5. EXPERIMENTAL RESULTS

We extract 15 features (four morphology, five GLCM and six color features) were identified; hence, the total input features were fifteen. These features were used to classify different skin cancer image of human body. In line with this, we have designed experimental scenarios to test the classification performance by taking the extracted features of cancer images. The classifications were tested by four different algorithms namely ANN (Artificial Neural Network), KNN (Nearest Neighbor classification) , Naive Bayes and combining SOM and RBF classifiers in order to get a more accurate result. In order to train the classifiers, a set of training skin cancer image was required, and the class label where it belongs to, 235 skin cancer image were taken from American skin cancer society and DERMOFIT from the predefined three types of skin cancer i.e. Melanoma, Basal cell carcinoma and Squamous cell carcinoma.

There are two basic phases of pattern classification. They are training and testing phases. In the training phase, data is repeatedly presented to the classifier, in order to obtain a desired response. In testing phase, the trained system is applied to data that it has never seen to check the performance of the classification. Hence, we need to design the classifier by partitioning the total data set into training and testing data set. From the total data set of each skin cancer type 70% was used to build training and the remaining 30% of the total was used for testing data. From the total of 235 data sets, 162 were used for training and 73 were used for testing. In general, a classifier has some input features based on the scenario of the designed experiment and some output features. In this study, there were three output classes, because the predefined human skin cancer images were three. The total numbers of examples or patterns were 235. This examples were normalized with mean 0 and variance 1.

	KNN			ANN				Naïve				SOM			
	Basal cell	Melanoma	Squamous cell	Basal cell	Melanoma	Squamous cell	Basal cell	Melanoma	Squamous cell	Basal cell	Melanoma	Squamous cell	Basal cell	Melanoma	Squamous cell
Basal cell	18	3	4	Basal cell	15	5	5	Basal cell	14	4	7	Basal cell	22	2	1
Melanoma	3	20	3	Melanoma	6	18	2	Melanoma	8	14	4	Melanoma	1	25	0
Squamous cell	6	2	14	Squamous cell	7	2	13	Squamous cell	7	2	13	Squamous cell	0	1	21
Total	73			Total	73			Total	73			Total	73		
correct	52			correct	46			correct	41			correct	68		
not correct	21			not correct	27			not correct	32			not correct	5		
%	71.232877			%	63.013699			%	56.164384			%	93.150685		

FIGURE 4: Summary result of KNN, ANN, Naïve and SOM classifier using extracted features.

As we have presented in detail in the previous section, the experiments were conducted under four scenarios by using feature sets of morphology, texture and color separately, and finally combining the three feature sets. Then, the experiment results were compared the performance of the Naïve Bayes classifier, KNN, neural network and SOM classification over the three

scenarios. The total number of data sets is 235. Out of these, 70% were used for training and the remaining 30% were used for testing. In general, the overall result showed that morphology and color features have more discriminating power than texture features and the classification performance of SOM is by far better than Naïve Bayes , artificial neural network classifier and KNN.

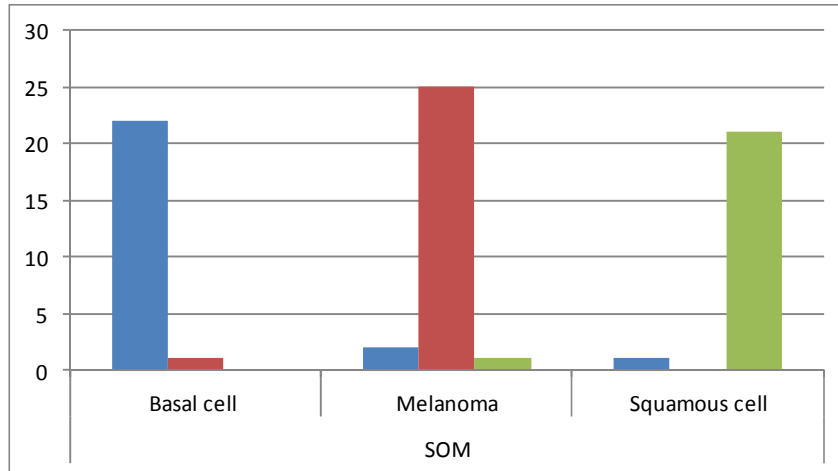


FIGURE 5: Overall Performance.

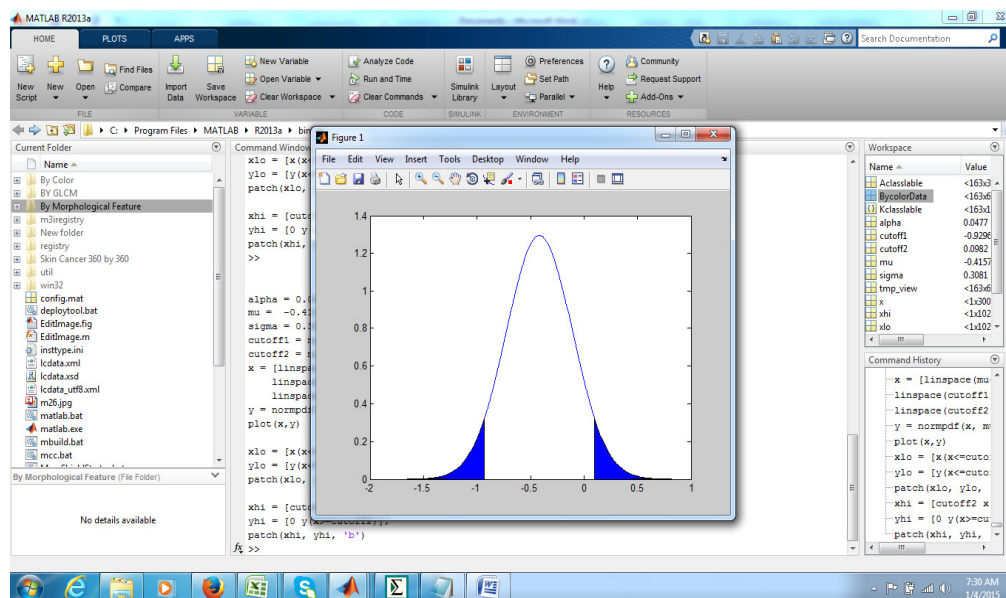


FIGURE 6: Confidence Interval of Detection.

## 6. CONCLUSION AND FUTURE WORK

In the classification problem of skin cancer recognition, morphological, GLCM and color features were extracted from a skin cancer images taken from three type of skin cancer Basal cell carcinoma, Melanoma and squamous cell carcinoma by using image analysis techniques. These selected features were used as input to the classification model. The experiment was conducted under four scenarios of the features data set such as GLCM, Morphology, Color and combining the three features. The result of the experimentation showed that the three varieties of Human skin cancer have been classified more accurately by SOM than using Naïve Bayes , ANN and KNN classifier. The image analysis for the recognition of the type of skin cancer can be further

investigated. The work can also be seen in depth and researched by the different characteristics of its physical and chemical in connection to image technology.

In light with this, the following recommendations are made for further research and improvements.

Identification of skin cancer type by exploring more features, Skin cancer recognition by levels of injuries using image analysis, implementing skin cancer recognition on mobile to make simplified for doctors, Computer vision for Recognition of leprosy and skin cancer.

## 7. REFERENCES

- [1] <http://www.cancer.org/cancer/skincancer>
- [2] Cancer in Africa, World Health Organization, International Agency for Research on cancer; 2009.
- [3] <http://www.skincancer.org/skin-cancer-information/skin-cancer-facts>
- [4] <http://www.cdc.gov/cancer/skin/statistics/>
- [5] <http://www.skincancer.org/skin-cancer-information/skin-cancer-facts>
- [6] World Health Organization, the global burden of disease: 2008 Update Geneva: World Health Organization,2008
- [7] A Morrone, Skin Cancer in Ethiopia,21st world congress of dermatology, 2007
- [8] <http://www.skincancer.org/skin-cancer-information/skin-cancer-facts>
- [9] William K. Pratt: Digital image processing, PIKS Scientific inside, John Wiley, 4th Edition, 2007
- [10] John Breneman towards early stages of malignant melanoma detection Using Consumer Mobile Devices
- [11] Luís Filipe Caeiro M argalho Guerra Rosado, Automatic System for Diagnosis of Skin Lesions Based on Dermoscopic Images
- [12] T Y Satheesha,, Dr. D Sathya Narayana, Dr. M N Giriprasad, review on early detection of melanoma:2012
- [13] Ilias Maglogiannis, Dimitrios I. Kosmopoulos, Computational vision systems for the detection of malignant melanoma
- [14] Ioana Dumitrache, Alina Elena Sultana, and Radu Dogaru Automatic Detection of Skin Melanoma from Images using Natural Computing Approaches
- [15] Aswin.R.B, J. Abdul Jaleel, Sibi Salim, Implementation of ANN Classifier using MATLAB for Skin Cancer Detection,2013
- [16] Ioana Dumitrache, Alina Elena Sultana, and Radu Dogaru Automatic Detection of Skin Melanoma from Images using Natural Computing Approaches
- [17] Nilkamal S. Ramteke and Shweta V. Jain ABCD rule based automatic computer-aided skin cancer detection using MATLAB
- [18] Santosh Achakanalli and G. Sadashivappa Skin Cancer Detection and Diagnosis Using Image Processing and Implementation Using Neural Networks and ABCD Parameters

- [19] Mariam A.Sheha and Mai S.Mabrouk Automatic Detection of Melanoma Skin Cancer using Texture Analysis
- [20] Catarina Barata, Margarida Ruela, Mariana Francisco, Teresa Mendonça, and Jorge S. Marques Detection of Melanomas in Dermoscopy Images Using Texture and Color Features
- [21] Santosh Achakanalli & G. Sadashivappa, Skin Cancer Detection And Diagnosis Using Image Processing And Implementation Using Neural Networks And ABCD Parameters
- [22] Luís Filipe Caeiro M argalho Guerra Rosado, Automatic System for Diagnosis of Skin Lesions Based on Dermoscopic Images
- [23] Mariam A. Sheha , Mai S. Mabrouk , Amr Sharawy, Automatic Detection of Melanoma Skin Cancer using Texture Analysis
- [24] Lin Li, Qizhi Zhang, Yihua Ding, Huabei Jiang, Bruce H Thiersand, Automatic diagnosis of melanoma using machine learning methods
- [25] Sarika Choudhari, Seema Biday, Artificial Neural Network for SkinCancer Detection
- [26] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B, Diagnosis and Detection of Skin Cancer Using Artificial Intelligence
- [27] Peyman Sabouri, Hamid GholamHosseini, John Collins, Border Detection of Mela noma Skin Lesions, 2013
- [28] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B, Artificial Neural Network Based Detection of Skin Cancer,2012
- [29] Snehal Salunke, Survey on Skin lesion segmentation and classification,2014
- [30] Nandini M.N., M.S.Mallikarjunaswamy, Detection of Melanoma Skin Disease using Dermoscopy Images
- [31] Tinku Acharya and Ajoy K. Ray, Image Processing Principles and Applications, Jhon Wiley, 2005.
- [32] William K. Pratt: Digital image processing, PIKS Scientific inside, John Wiley, 4th Edition, 2007.



# Textural Feature Extraction of Natural Objects for Image Classification

**Vishal Krishna**

*Computer Science  
Georgia Institute of Technology  
Atlanta – 30332, US*

*vkrishna7@gatech.edu*

**Ayush Kumar**

*Computer science  
BITS Pilani, Goa Campus  
Goa – 403726, India*

*f2010029@goa.bits-pilani.ac.in*

**Kishore Bhamidipati**

*Computer Science Engineering  
Manipal Institute of Technology  
Manipal – 576104, India*

*Kishore.b@manipal.edu*

---

## Abstract

The field of digital image processing has been growing in scope in the recent years. A digital image is represented as a two-dimensional array of pixels, where each pixel has the intensity and location information. Analysis of digital images involves extraction of meaningful information from them, based on certain requirements. Digital Image Analysis requires the extraction of features, transforms the data in the high-dimensional space to a space of fewer dimensions. Feature vectors are n-dimensional vectors of numerical features used to represent an object. We have used Haralick features to classify various images using different classification algorithms like Support Vector Machines (SVM), Logistic Classifier, Random Forests Multi Layer Perception and Naïve Bayes Classifier. Then we used cross validation to assess how well a classifier works for a generalized data set, as compared to the classifications obtained during training.

**.Keywords:** Feature Extraction, Haralick, Classifiers, Cross-Validation.

---

## 1. INTRODUCTION

Texture is an important feature for many types of analysis of images and identification of regions of interest. Texture analysis has a wide array of applications, including industrial and biomedical monitoring, classification and segmentation of satellite or aerial photos, identification of ground relief, and many others. [1] Various methods have been proposed via research over the years for identifying and discriminating the textures. Measures like angular second moment, contrast, mean, correlation, entropy, inverse difference moment, etc. have been typically used by researchers for obtaining feature vectors, which are then manipulated to obtain textural features. One of the most popular approaches to texture analysis is based on the co-occurrence matrix obtained from images, proposed by Robert M. Haralick in 1973, which forms the basis of this paper.

Image classification is one of the most important part of digital image analysis. Classification is a computational procedure that sorts images into subsets according to their similarities. [4] Contextual image classification, as the name suggests, is a method of classification based on the contextual information in images, i.e. the relationship amongst neighbouring pixels. [2].

For classification, we used the WEKA (“Waikato Environment for Knowledge Analysis”) tool, which is an open source machine-learning software suite developed using Java, by the University

of Waikato, New Zealand.[6] It contains set of tools for different data analysis and modelling techniques such as: pre-processing, classification, clustering, segmentation, association rules and visualization. It implements many artificial intelligence algorithms like decision trees, neural networks, Particle Swarm Optimization etc.).[5]

## 2. LITERATURE SURVEY

The classification of images can be done either on the basis of a single resolution cell or on a collection of resolution cells. When a block of cells are used, the challenge is to define a set of features to represent the information given by those cells, which can be used for classification of the images.

Human perception of images is based on three major classes of features: spectral, textural and contextual. Spectral features are obtained as the average variation of tone across various bands of the electromagnetic spectrum. Textural features, on the other hand, provide information about the variation of tone within a single band. Information from portions of image surrounding the part under analysis constitute the contextual features. In gray-scale photographs, tone represents the varying gray levels in resolution cells, while the statistical distribution of the gray levels is interpreted as texture. Tone and texture form an intrinsic part of any image, though one can get precedence over the other according to the nature of the image. Simply stated, the relation between the two is: tone is dominant when the sample under consideration shows only small range of variation of gray levels, while gray levels spread over a wide range in a similar sample indicate the dominance of texture.

Haralick's work is based on the assumption that information regarding the texture of any image can be obtained from calculating the average spatial relation of the gray tones of the image with each other. The procedure for calculating the Haralick textural features is based on a set of gray-tone spatial-dependence probability distribution matrices (also termed as Gray-Level Co-occurrence Matrices or GLCM, or gray-level spatial dependence matrix), computed for various angles at fixed distances. From each such matrix, fourteen features can be calculated, which provide information in terms of homogeneity, contrast, linear variation of gray tone, nature and number of boundaries etc.

**Co-occurrence Matrix:** A co-occurrence matrix, P, is used to describe the relationships between neighbouring (at a distance, d) pixels in an image. 4 co-occurrence matrices, each calculated for a different angle, can be defined. A co-occurrence matrix, termed as  $P^0$ , describes pixels that are adjacent to one another horizontally (at angle  $0^\circ$ ). Similarly, co-occurrence matrices are defined for the vertical direction ( $90^\circ$ ) and both diagonals ( $45^\circ$  and  $135^\circ$ ). These matrices are called  $P^{90}$ ,  $P^{45}$  and  $P^{135}$  respectively. [3]

$$x = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 \end{pmatrix} \quad p_0 = \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$p_{45} = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad p_{90} = \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

$$p_{135} = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

There are 4 pairs of (0,0) in angular 0, thus  $P^0(0,0)=4$ , there are 2 pairs of (0,1), thus  $P^0(0,1)=2$ . Similarly all the four matrices are computed.

Based on the co-occurrence matrices calculated as above, the thirteen texture features as proposed by Haralick are defined below:

Notation:

$N_g$  : Number of distinct gray levels in quantized image

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j)$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j)$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j), |i - j| = k \text{ and } k = 0, 1, \dots, N_g - 1$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j), i + j = k \text{ and } k = 2, 3, \dots, 2N_g$$

a) Angular Second Moment

$$f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)^2$$

b) Contrast

$$f_2 = \sum_{k=0}^{N_g-1} k^2 p_{x-y}(k)$$

c) Correlation

$$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Where  $\mu_x, \mu_y, \sigma_x, \sigma_y$  are mean of x, y and standard deviation of x, y respectively.

d) Sum of Squares: Variance

$$f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p(i, j)$$

e) Inverse Difference Moment

$$f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} p(i, j)$$

f) Sum Average

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$$

g) Sum Variance

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i)$$

h) Sum Entropy

$$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log (p_{x+y}(i))$$

i) Entropy

$$f_9 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log (p(i, j))$$

j) Difference Variance

$$f_{10} = \text{variance of } p_{x-y}.$$

k) Difference Entropy

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log (p_{x-y}(i))$$

l) Information measures of correlation

$$f_{12} = \frac{f_9 - HXY1}{\max (HX, HY)}$$

$$f_{13} = \sqrt{1 - \exp^{-2(HXY2-f_9)}}.$$

where HX and HY are entropies of  $p_x$  and  $p_y$ .

### 3. METHODOLOGY

For any value of  $d$ , as mentioned before, 4 matrices are calculated for each of the thirteen features detailed above. The mean and range of each set of four values give a set 28 values which are then passed to the classifier. Out of the input features, some share a strong correlation, so a feature-selection procedure can identify a subset of features in order to give good results in classification.

The test data has a total of 25 classes, which are known Apriori. We use this knowledge to calculate the effectiveness of various classification algorithms available, on the Haralick features. The classification algorithms used are:

1. Naïve Bayes Classifier (NB) - A Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. [7]
2. Logistic Classifier (Log) - Logistic regression is a probabilistic statistical classification model. It measures the relationship between the categorical dependent variable and one or more independent variables, which are usually (but not necessarily) continuous, by using probability scores as the predicted values of the dependent variable.[8]

$$HXY1 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log (p_x(i)p_y(j))$$

$$HXY2 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log (p_x(i)p_y(j))$$

3. Multilayer Perception Classifier (MP) – In conventional MLP, components of feature vectors are made to take crisp binary values, and the pattern is classified according to highest activation reached. [9]

4. Random Forest Classifier (RF) - Random forests operate by constructing a number of decision trees training data and classifying data according to the mode of the obtained.[10]
5. Sequential Minimal Optimization – The algorithm is used to train support vector machines for classification. [11]

The parameters on which the effectiveness of each of the above algorithms are:

1. True Positive Rate (TP) – it is the number of items correctly labelled as belonging to the particular class divided by the total number of elements labelled as belonging to that class
2. False Positive Rate (FP) – it is the number of items incorrectly labelled as belonging to the particular class divided by the total number of elements labelled as belonging to that class
3. Precision - it is the fraction of retrieved instances that are relevant
4. Recall - it is the fraction of relevant instances that are retrieved
5. F-Measure – it is a measure that combines precision and recall, calculated as the harmonic mean of precision and recall
6. ROC Area - receiver operating characteristic (ROC) is a plot of the performance of a binary classifier system. The area under the curve is treated as a measure of accuracy of the classifier.

A second set of experiments are carried out, using the same test data, algorithms and parameters, but with the added constraint of using cross validation factor of 10.

#### 4. RESULTS AND ANALYSIS

Each algorithm is first run on the data set and all six parameters are measured and compared. The results obtained are given below.

Class	TP Rate				
	NB	Log	MP	SMO	RF
1	0.525	1.000	0.950	0.675	1
2	0.750	1.000	0.975	0.675	1
3	0.850	1.000	1.000	0.875	0.975
4	0.775	0.975	0.975	0.800	1
5	0.900	1.000	1.000	0.975	1
6	0.825	1.000	0.975	0.800	1
7	0.900	1.000	1.000	0.950	1
8	0.775	1.000	0.925	0.700	1
9	0.850	0.975	0.825	0.675	1
10	0.800	1.000	1.000	0.925	1
11	0.725	1.000	0.975	0.775	1
12	0.750	1.000	0.950	0.825	1
13	0.800	1.000	1.000	0.900	1
14	0.650	0.975	0.975	0.850	1
15	0.725	0.975	1.000	0.850	0.975
16	0.850	1.000	0.975	0.800	0.975
17	0.975	0.975	1.000	0.800	1
18	0.900	0.975	1.000	0.975	1
19	0.600	0.975	0.975	0.800	1
20	1.000	1.000	1.000	1.000	1
21	0.250	1.000	0.925	0.725	0.975
22	0.750	1.000	0.975	0.900	1
23	0.525	1.000	0.950	0.775	0.9

24	1.000	1.000	1.000	0.975	1
25	0.850	1.000	1.000	0.975	1

FIGURE 1.1: Values of TP Rate of each class for different classification methods.

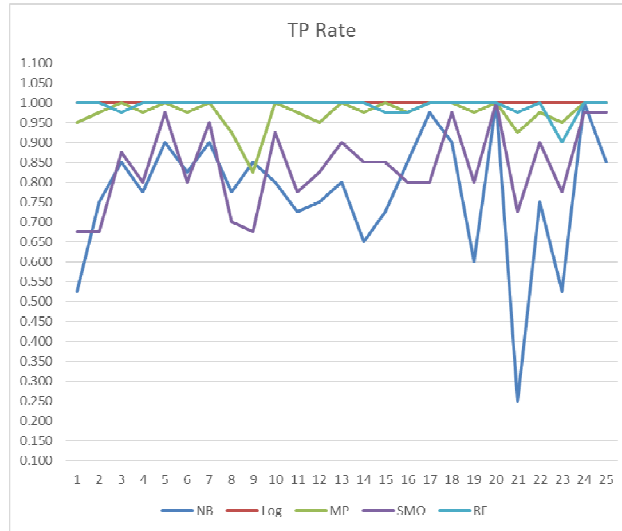
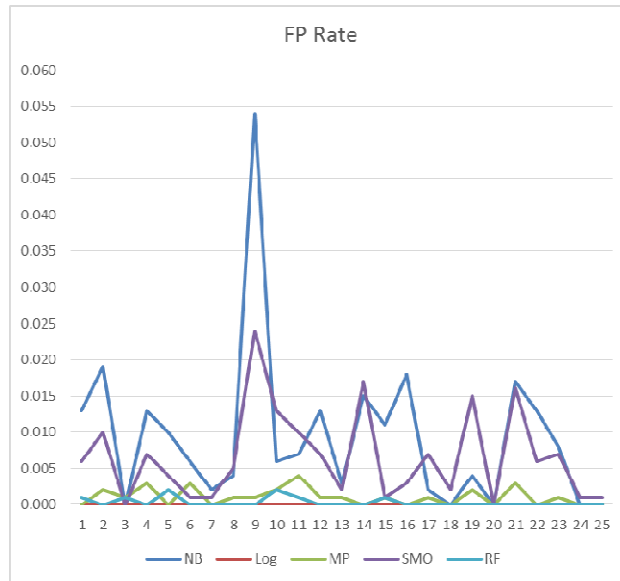


FIGURE 1.2: Graphical representation of TP Rate values.

Class	NB	Log	MP	SMO	RF
1	0.013	0.000	0.000	0.01	0.001
2	0.019	0.000	0.002	0.01	0
3	0.000	0.000	0.001	0	0.001
4	0.013	0.001	0.003	0.01	0
5	0.010	0.000	0.000	0	0.002
6	0.006	0.000	0.003	0	0
7	0.002	0.000	0.000	0	0
8	0.004	0.000	0.001	0.01	0
9	0.054	0.001	0.001	0.02	0
10	0.006	0.000	0.002	0.01	0.002
11	0.007	0.000	0.004	0.01	0.001
12	0.013	0.000	0.001	0.01	0
13	0.003	0.000	0.001	0	0
14	0.015	0.000	0.000	0.02	0
15	0.011	0.000	0.001	0	0.001
16	0.018	0.000	0.000	0	0
17	0.002	0.001	0.001	0.01	0
18	0.000	0.001	0.000	0	0
19	0.004	0.000	0.002	0.02	0
20	0.000	0.000	0.000	0	0
21	0.017	0.000	0.003	0.02	0
22	0.013	0.000	0.000	0.01	0
23	0.008	0.000	0.001	0.01	0
24	0.000	0.000	0.000	0	0
25	0.000	0.000	0.000	0	0

FIGURE 2.1: Values of FP Rate of each class for different classification methods.

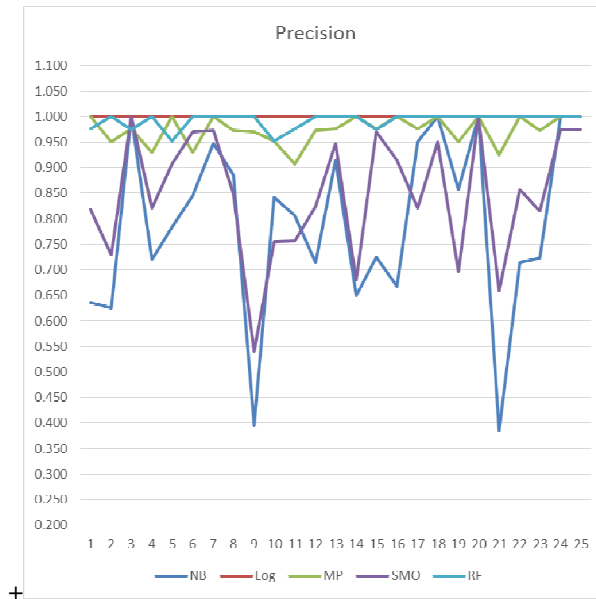


**FIGURE 2.2:** Graphical representation of FP Rate values.

Class	NB	Log	MP	SMO	RF
1	0.636	1.000	1.000	0.82	0.976
2	0.625	1.000	0.951	0.73	1
3	1.000	1.000	0.976	1	0.975
4	0.721	0.974	0.929	0.82	1
5	0.783	1.000	1.000	0.91	0.952
6	0.846	1.000	0.929	0.97	1
7	0.947	1.000	1.000	0.97	1
8	0.886	1.000	0.974	0.85	1
9	0.395	0.983	0.971	0.54	1
10	0.842	1.000	0.952	0.76	0.952
11	0.806	1.000	0.907	0.76	0.976
12	0.714	1.000	0.974	0.83	1
13	0.914	1.000	0.976	0.95	1
14	0.650	0.994	1.000	0.68	1
15	0.725	0.992	0.976	0.97	0.975
16	0.667	1.000	1.000	0.91	1
17	0.951	0.978	0.976	0.82	1
18	1.000	0.984	1.000	0.95	1
19	0.857	0.993	0.951	0.7	1
20	1.000	1.000	1.000	1	1
21	0.385	1.000	0.925	0.66	1
22	0.714	1.000	1.000	0.86	1
23	0.724	1.000	0.974	0.82	1
24	1.000	1.000	1.000	0.98	1

25	1.000	1.000	1.000	0.98	1
----	-------	-------	-------	------	---

**FIGURE 3.1:** Values of Precision of each class for different classification methods.



**FIGURE 3.2:** Graphical representation of Precision values.

Class	NB	Log	MP	SMO	RF
1	0.525	1.000	0.950	0.68	1
2	0.750	1.000	0.975	0.68	1
3	0.850	1.000	1.000	0.88	0.975
4	0.775	0.975	0.975	0.8	1
5	0.900	1.000	1.000	0.98	1
6	0.825	1.000	0.975	0.8	1
7	0.900	1.000	1.000	0.95	1
8	0.775	1.000	0.925	0.7	1
9	0.850	0.975	0.825	0.68	1
10	0.800	1.000	1.000	0.93	1
11	0.725	1.000	0.975	0.78	1
12	0.750	1.000	0.950	0.83	1
13	0.800	1.000	1.000	0.9	1
14	0.650	0.975	0.975	0.85	1
15	0.725	0.975	1.000	0.85	0.975
16	0.850	1.000	0.975	0.8	0.975
17	0.975	0.975	1.000	0.8	1
18	0.900	0.975	1.000	0.98	1
19	0.600	0.975	0.975	0.8	1
20	1.000	1.000	1.000	1	1
21	0.250	1.000	0.925	0.73	0.975



22	0.750	1.000	0.975	0.9	1
23	0.525	1.000	0.950	0.78	0.9
24	1.000	1.000	1.000	0.98	1
25	0.850	1.000	1.000	0.98	1

FIGURE 4.1: Values of Recall of each class for different classification methods.

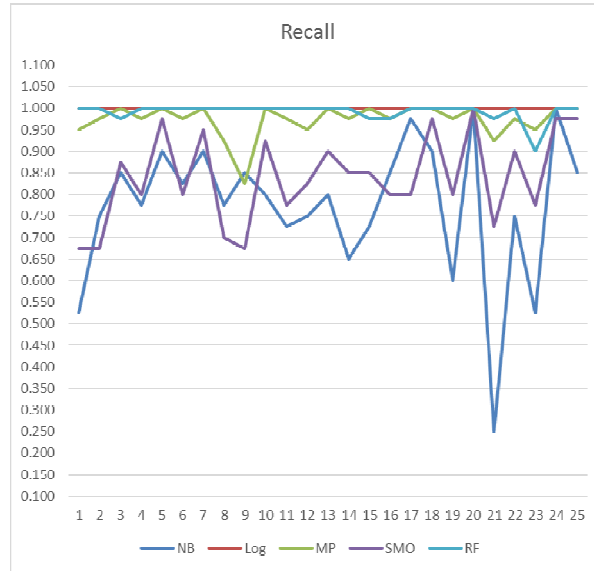
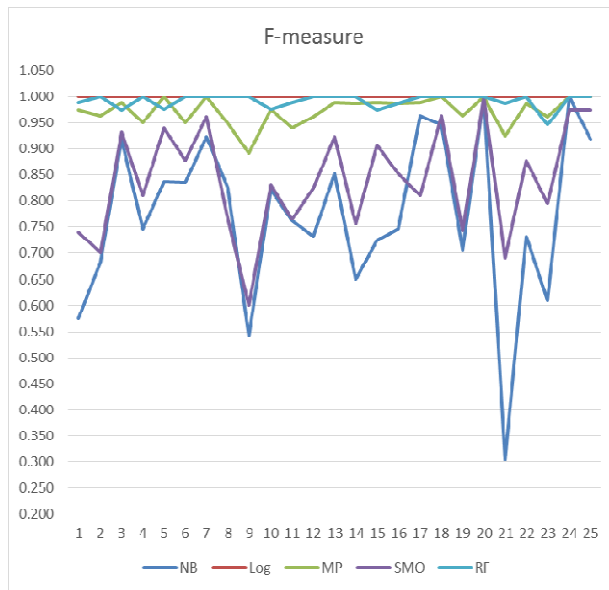


FIGURE 4.2: Graphical Representation of values of Recall.

Class	NB	Log	MP	SMO	RF
1	0.575	1.000	0.974	0.74	0.988
2	0.682	1.000	0.963	0.7	1
3	0.919	1.000	0.988	0.93	0.975
4	0.747	0.976	0.951	0.81	1
5	0.837	1.000	1.000	0.94	0.976
6	0.835	1.000	0.951	0.88	1
7	0.923	1.000	1.000	0.96	1
8	0.827	1.000	0.949	0.77	1
9	0.540	0.979	0.892	0.6	1
10	0.821	1.000	0.976	0.83	0.976
11	0.763	1.000	0.940	0.77	0.988
12	0.732	1.000	0.962	0.83	1
13	0.853	1.000	0.988	0.92	1
14	0.650	0.982	0.987	0.76	1
15	0.725	0.986	0.988	0.91	0.975
16	0.747	1.000	0.987	0.85	0.987
17	0.963	0.976	0.988	0.81	1
18	0.947	0.993	1.000	0.96	1

19	0.706	0.981	0.963	0.74	1
20	1.000	1.000	1.000	1	1
21	0.303	1.000	0.925	0.69	0.987
22	0.732	1.000	0.987	0.88	1
23	0.609	1.000	0.962	0.8	0.947
24	1.000	1.000	1.000	0.98	1
25	0.919	1.000	1.000	0.98	1

**FIGURE 5.1:** Values of F-measure of each class for different classification methods.

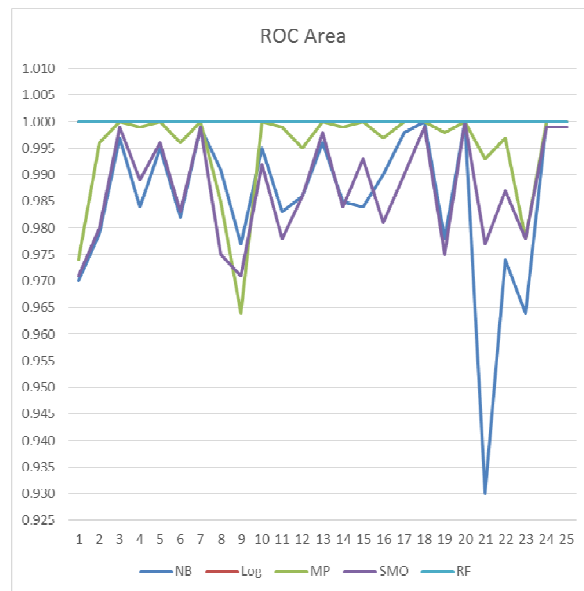


**FIGURE 5.2:** Graphical representation of F-measure values.

Class	NB	Log	MP	SMO	RF
1	0.970	1.000	0.974	0.97	1
2	0.979	1.000	0.996	0.98	1
3	0.997	1.000	1.000	1	1
4	0.984	1.000	0.999	0.99	1
5	0.995	1.000	1.000	1	1
6	0.982	1.000	0.996	0.98	1
7	0.999	1.000	1.000	1	1
8	0.991	1.000	0.985	0.98	1
9	0.977	1.000	0.964	0.97	1
10	0.995	1.000	1.000	0.99	1
11	0.983	1.000	0.999	0.98	1
12	0.986	1.000	0.995	0.99	1
13	0.996	1.000	1.000	1	1
14	0.985	1.000	0.999	0.98	1
15	0.984	1.000	1.000	0.99	1

16	0.990	1.000	0.997	0.98	1
17	0.998	1.000	1.000	0.99	1
18	1.000	1.000	1.000	1	1
19	0.978	1.000	0.998	0.98	1
20	1.000	1.000	1.000	1	1
21	0.930	1.000	0.993	0.98	1
22	0.974	1.000	0.997	0.99	1
23	0.964	1.000	0.978	0.98	1
24	1.000	1.000	1.000	1	1
25	1.000	1.000	1.000	1	1

**FIGURE 6.1:** Values of ROC Area of each class for different classification methods.



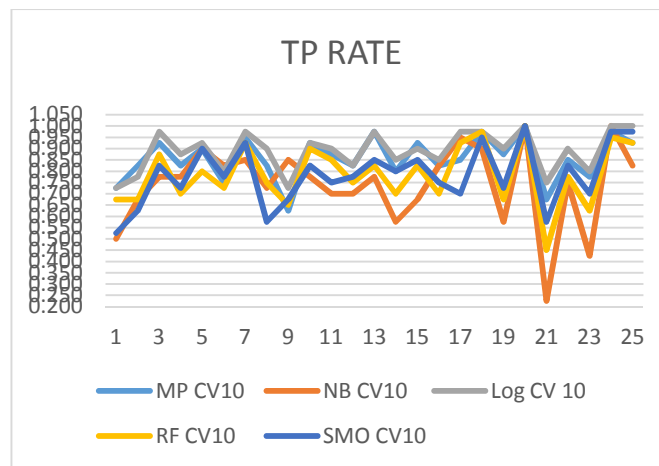
**FIGURE 6.2:** Graphical representation of ROC Area Values.

The following tables and diagrams pertain to the second set of experiments, i.e. with a cross validation factor of 10 in each case.

Class	MP CV10	NB CV10	Log CV 10	RF CV10	SMO CV10
1	0.725	0.5	0.725	0.675	0.525
2	0.825	0.675	0.775	0.675	0.625
3	0.925	0.775	0.975	0.875	0.825
4	0.825	0.775	0.875	0.7	0.725
5	0.900	0.9	0.925	0.8	0.9
6	0.750	0.825	0.8	0.725	0.775
7	0.950	0.85	0.975	0.925	0.925
8	0.825	0.725	0.9	0.75	0.575
9	0.625	0.85	0.725	0.65	0.675

10	0.925	0.775	0.925	0.9	0.825
11	0.875	0.7	0.9	0.85	0.75
12	0.825	0.7	0.825	0.75	0.775
13	0.975	0.775	0.975	0.825	0.85
14	0.800	0.575	0.85	0.7	0.8
15	0.925	0.675	0.9	0.825	0.85
16	0.825	0.825	0.85	0.7	0.75
17	0.850	0.95	0.975	0.925	0.7
18	0.975	0.9	0.975	0.975	0.95
19	0.875	0.575	0.9	0.675	0.725
20	1.000	1	1	1	1
21	0.675	0.225	0.75	0.45	0.575
22	0.850	0.75	0.9	0.775	0.825
23	0.775	0.425	0.8	0.625	0.7
24	0.975	1	1	0.95	0.975
25	0.925	0.825	1	0.925	0.975

**FIGURE 7.1:** Values of TP Rate of each class for different classification methods with cross validation 10.



**FIGURE 7.2:** Graphical representation of TP Rate values with Cross Validation.

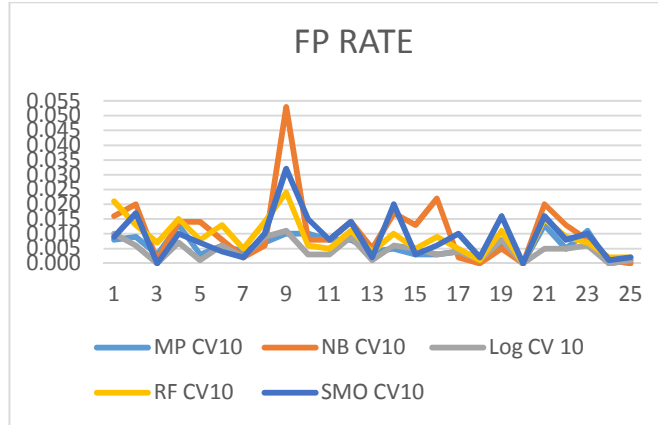


FIGURE 8.2: Graphical Representation of FP Rate values with cross validation 10.

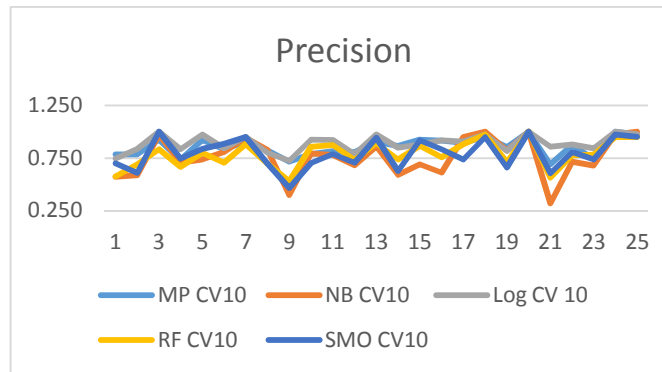


FIGURE 9.2: Graphical Representation of Precision values with cross validation 10.

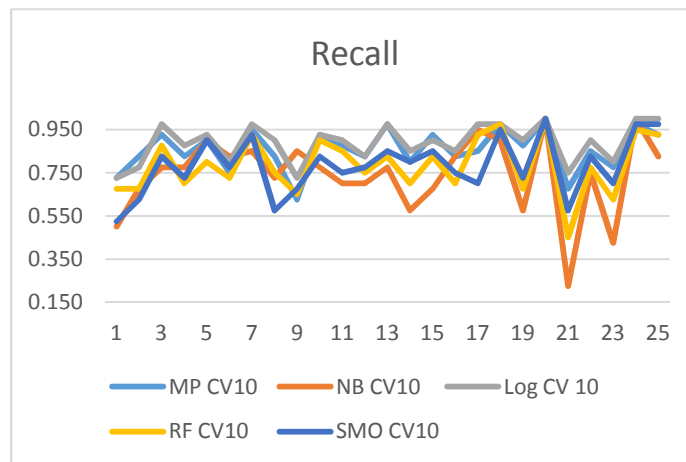
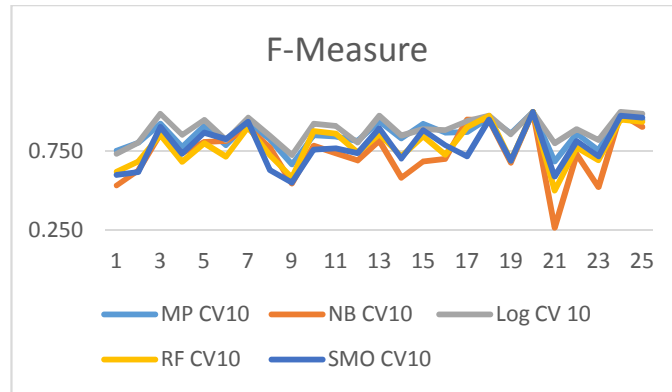
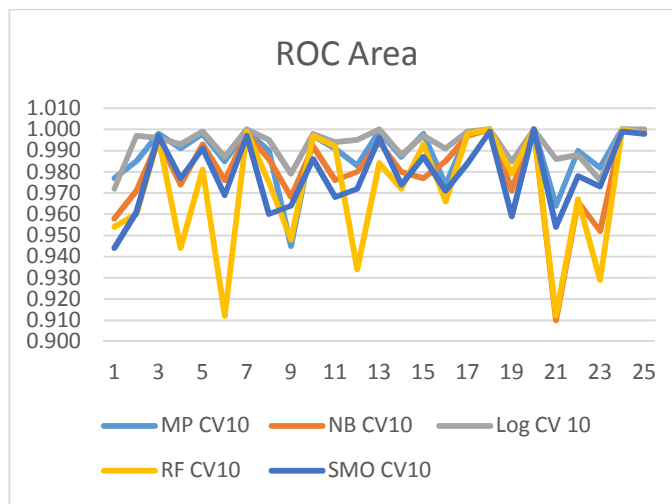


FIGURE 10.2: Graphical Representation of Recall values with cross validation 10.



**FIGURE 11.2:** Graphical Representation of F-measure values with cross validation 10.



**FIGURE 12.2:** Graphical Representation of ROC Area values with cross validation 10.

The overall accuracy of each algorithm, considering all classes is depicted below.

Log	MP	NB	RF	SMO
99.7	97.3	77.2	99.2	83.9

**FIGURE 13.1:** Overall accuracy values of all classes.

## 5. CONCLUSION

Our comparative study provides a comprehensive analysis to Haralick features and its use in the well-known classification models. From the analysis, we can see Logistic classifier performs extremely well under all parameters, which is reflected in the combined accuracy values. It has a 99.7 percent accuracy for the trained parameters across all the classes. Random Forest Classifier performs second with respect to the rest of the classifiers. It successfully predicted all the values for most of the classes. Native Bayes performs the worst, especially with certain classes, which brings down the total accuracy achieved.

On applying cross validation with a factor of 10, we see that the accuracy decreases across all the classifiers. The different classifiers perform similarly with respect to each other as they did without cross validation. However, it can be seen that MultiLayer Perception Classifier performs slightly better than Random Forest Classifier in this case.

Apart from Native Bayes, all other methods had an accuracy of over 80 percent. Logistical and Random Forest scored above 99 percent in its accuracy. This demonstrates the power of Haralick features and its efficiency in image classification using standard classification models.

## 6. REFERENCES

- [1] Timo Ojala, Matti Pietikainen and David Harwood, A comparative study of texture measures with classification based on feature distributions.
- [2] M. Pietikainen, T. Ojala, Z. Xu; Rotation-invariant texture classification using feature distributions
- [3] Eizan Miyamoto and Thomas Merryman Jr; FAST CALCULATION OF HARALICK TEXTURE FEATURES.
- [4] Frank, J. (1990) Quart. Rev. Biophys. 23, 281-329.
- [5] Baharak Goli and Geetha Govindan ; WEKA – A powerful free software for implementing Bio-inspired Algorithms; State Inter University Centre of Excellence in Bioinformatics, University of Kerala).
- [6] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.
- [7] Tom M. Mitchell; Machine Learning; McGraw Hill, 2010.
- [8] David A. Freedman; Statistical Models: Theory and Practice; Cambridge University Press, 2009, p. 128.
- [9] Shankar Pal, Shushmita Mitra; Multilayer Perceptron, Fuzzy Sets and Classification; IEEE Transactions on Neural Networks, Vol 3, September 1992.
- [10] Leo Breiman ; "Random Forests". Machine Learning 45 (1); 2001.
- [11] John Platt; Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines; 1998.

# Novel Approach to Use HU Moments with Image Processing Techniques for Real Time Sign Language Communication

**Matheesha Fernando**

*Department of Industrial Management,  
University of Kelaniya  
Sri Lanka*

*matheemit@gmail.com*

**Janaka Wijayanayake**

*Department of Industrial Management,  
University of Kelaniya  
Sri Lanka*

*janaka@kln.ac.lk*

---

## Abstract

Sign language is the fundamental communication method among people who suffer from speech and hearing defects. The rest of the world doesn't have a clear idea of sign language. "Sign Language Communicator" (SLC) is designed to solve the language barrier between the sign language users and the rest of the world. The main objective of this research is to provide a low cost affordable method of sign language interpretation. This system will also be very useful to the sign language learners as they can practice the sign language. During the research available human computer interaction techniques in posture recognition was tested and evaluated. A series of image processing techniques with Hu-moment classification was identified as the best approach. To improve the accuracy of the system, a new approach; height to width ratio filtration was implemented along with Hu-moments. System is able to recognize selected Sign Language signs with the accuracy of 84% without a controlled background with small light adjustments.

**Keywords:** Sign Language Recognition, Height to Width Ratio, Hu-moments, YCrCb Color Space.

---

## 1. INTRODUCTION

More than 40 million of the world population suffers from speech impairments. Sign Language is the mode of communication of speech and auditory impaired people. In today's world communication facilities provided for such people are very less. Generally people in the society don't know about sign language. An ordinary person to communicate with speech and auditory impaired person, a translator is usually needed to convert the sign language into natural language and vice versa. This is an inflexible way of communication and restricts communication between speech and hearing impaired people and ordinary people.

The objective of the research is to develop a technology that can act as an intermediate flexible medium for speech and hearing impaired people to communicate amongst themselves and with other individuals to enhance their level of learning / education.

The research outcome will facilitate speech and hearing impaired people to communicate with the world in a more comfortable way where they can easily get what they need from the society and also can contribute to the well-being of the society. Furthermore, another expectation is to use the research outcome as a learning tool of sign language where learners can practice signs.

The rest of this paper is organized as follows. Section 2 briefly introduces the area of study with related work in sign language recognition and human computer interaction. In section 3 the proposed methodology for feature extraction is described in detail. Section 4 describes the



proposed methodology for classification with the experiments results. Section 5 reveals the limitations and drawbacks of the proposed approach. In section 6 possible future improvements are introduced and in section 7 the conclusion of the research findings is presented.

## 2. RELATED WORK

With the development of information technology new areas of human computer-interaction are emerging. There, human gesture plays a major role in the field of human computer interaction. Hand gestures provide a separate complementary modality to speech for expressing ones ideas. As sign language is a collection of gestures and postures any effort in sign language recognition falls in the field of human computer interaction and image processing.

There are two types of approaches commonly used to interpret gestures for human computer interaction. First category, Data Glove based approach relies on electromechanical devices attached to a glove for digitizing hand and finger motions into multi-parametric data. The major problems with that approach are; it requires wearing the devices which are expensive and also cause less natural behaviors [1]. The second category, Vision based approach is further divided into two categories as 3D hand model based approach and appearance based approach. 3D hand model based approach relies on the 3D kinematic hand model and tries to estimate the hand parameters by comparison between the input images and the possible 2D appearance projected by the 3D hand model. The appearance based approach uses image features to model the visual appearance of the hand and compare these parameters with the extracted image features from the video input [1, 2].

There have been a number of research efforts on appearance based methods in recent years. Freeman and Roth [3] presented a method to recognize hand gestures based on pattern recognition technique employing histograms of local orientation. There are some different types of gestures with similar orientation histograms which cannot be separately recognized by this method. This is one of the major drawbacks of Freeman and Roth [3] proposed method. Orientation histogram method is most appropriate for close-ups of the hand and also it doesn't provide a method to recognize hand itself.

Roth and Tanaka [4] have presented a way of classifying images based on their moments with more concern on recognizing shapes of the static postures. Serge and Malik [5] have discussed an approach to measure similarity between shapes and exploit it for object recognition where they attach a descriptor – shape context to each point. The corresponding points in similar shapes will have similar shape contexts. By that they compute the sum of matching errors between corresponding points. This requires more computational power. Flusser [6] presented a survey on Moment Invariants in Image Analysis on object recognition /classification methods based on image moments. The paper presents a general theory to construct invariant classification methods. Nianjun Liu and Brian C. Lovell [7] have presented an approach to Hand Gesture Extraction by Active Shape Models where a set of feature vectors will be normalized and aligned and then trained by Principle Component Analysis (PCA). Chen Chang [8] came up with an approach for Static Gesture Recognition by recognizing static gestures based on Zernike moments (ZMs) and pseudo-Zernike moments (PZMs). This approach takes a step toward extracting reliable features for static gesture recognition. Zhang et al. [9] suggested a Fast Convex Hull Algorithm for Binary Image for pattern recognition. The recognition is achieved by computing the extreme points, dividing the binary image into several regions, scanning the region's existing vertices dynamically, calculating the monotone segments, and merging these calculated segments. Deng and Jason, [10] came up with an idea of shape context based matching with cost matrix for real time Hand Gesture Recognition. There, they translate the edge elements of image shape to a group of feature points with N value.

Erdem Yörük et al. [11] have suggested a method for person recognition and verification based on their hands. There at preprocessing stage of the algorithm, the silhouettes of hand images are registered to a fixed pose, which involves both rotation and translation of the hand and the

individual fingers separately. Two feature sets, Hausdorff distance of the hand contours and independent component features of the hand silhouette images have been comparatively assessed. The system is not implemented in real time as it is tested based on the images of the hands captured by a scanner.

In appearance based method, feature extraction and classification are the major components that take part in. Under feature extraction Hand contour, Multi scale color features, Scale Invariant Feature Transform (SIFT), Speeded Up Robust Feature(SURF), Haar-like features, Histograms of Oriented Gradients (HOG), Local Orientation Histogram, Hough Transform (HT) are some of available research findings in efficient feature extraction. Qing Chen [12] has used a combined approach for recognition using two levels. Haar-like features were used to hand posture detection and tracking. AdaBoost learning algorithm was used to construct an accurate cascade of classifiers by combining a sequence of weak classifiers and to recognize different hand postures with parallel cascades. Yikai Fang et al. [13] have suggested a combined approach of co-training haar and HOG features simultaneously reduces the required training samples.

### 3. METHODOLOGY

As shown in Figure 1, this research is focused on a system that can convert a video signal (processed as sequence of images) into a sequence of written words (text) and voice in real time. Further, in the other way round it can provide a sign demonstration for a given text.

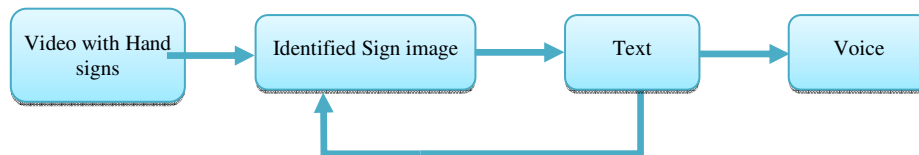


FIGURE 1: The Idea.

Robustness, Computational Efficiency and User’s Tolerance were the important challenges need to be considering in conducting this research project. Overall methodology followed for the sign language classification is summarized in figure 2.

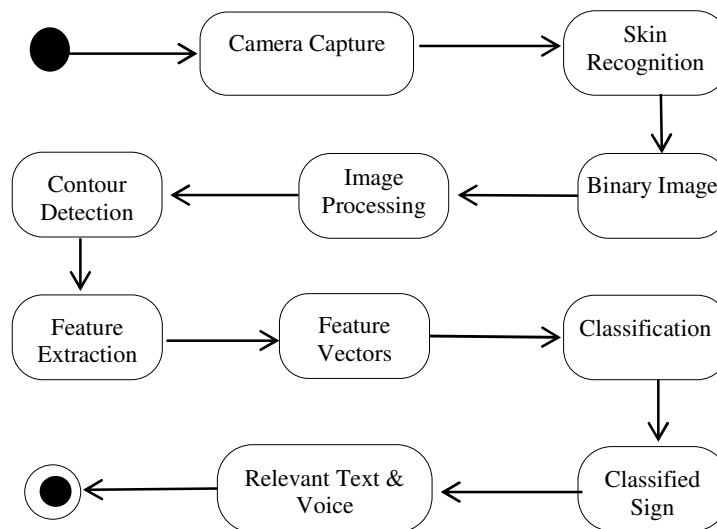


FIGURE 2: Classification Process.

In this approach the movement of the hand is recorded by a camera and the input video is decomposed into a set of features by taking individual frames into account. The video frames contain background pixels other than the hand, as a hand will never fill a perfect square. These pixels have to be removed as they contain arbitrary values that introduce noise into the process.

System captures images containing the hand signs from a video first and then identifies the skin colored area and gets the hand separated out from the rest of the background. Thereafter, a binary mask for skin pixels is constructed. Then the hands are isolated from other body parts as well as background objects. The basic idea is to get the real time image and then extract the predefined features and compare the feature vectors against the features extracted from stored template sign images. Sign database has the text and voice relevant to the meaning of each sign. When the real time sign is matched with a stored sign, the system will provide the relevant text and voice.

### 3.1. Image Capture Phase

Only a web camera was used to make this system a low cost, affordable application and that made the capturing process simple. As the normal web camera images are very poor in quality it requires more image processing techniques to deal with the issue.

### 3.2. Hand Segmentation

The hand must be localized in the image and segmented from the background before recognition. Color is the selected feature due to various reasons such as its computational simplicity, its invariant properties regarding the hand shape configurations and due to the human skin-color characteristic values. The method is based on a more rule based model of the skin-color pixels distribution. Since people have different hand colors it is incorrect to rely on the intensity. Therefore, it is required to convert the RGB representation to an intensity free color space model. Hue-Saturation-Value (HSV) representation and YCrCb color space are the best options available [14]. In YCrCb, Y is the luminance component and CB and CR are the blue-difference and red-difference chroma components. YCC is simply created from RGB [15].

The advantage of YCC over RGB is that it can reduce resolution of the color channels without altering the apparent resolution of the image, since the main Y (luminance) signal is untouched. YCrCb derived from the corresponding RGB space is as follows,

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

$$Cb = 0.564(B - Y) + 1/2 \text{ full scale}$$

$$Cr = 0.713(R - Y) + 1/2 \text{ full scale}$$

After a number of iterations of experiment, the suitable color ranges for human skin in both HSV and YCrCb were figured out as follows,

$$\text{hsv\_min} = (0, 30, 60)$$

$$\text{hsv\_max} = (30, 150, 255)$$

$$\text{YCrCb\_min} = (0, 131, 80)$$

$$\text{YCrCb\_max} = (255, 185, 135)$$

In a color image there is more information to be processed. The processing part is more complicated and time consuming. Therefore, binarizing (converting the image to black and white) is more desirable, since it reduces the amount of information contained in the source image. Thresholding can be used to binarize the image. Figure 3 shows the effect of binarizing.



Source Image



Binary Image with threshold value 75

**FIGURE 3:** Binary Image.

### 3.3. Image Pre-processing Phase

Image processing techniques were used to improve the quality of the image. Morphological transformations were used to get a more clear hand image. The basic morphological transformations are called dilation and erosion, and they arise in a wide variety of contexts such as removing noise, isolating individual elements, and joining disparate elements in an image. The skin recognized image was processed with both dilation and erosion to get the image clearer as shown in figure 4.

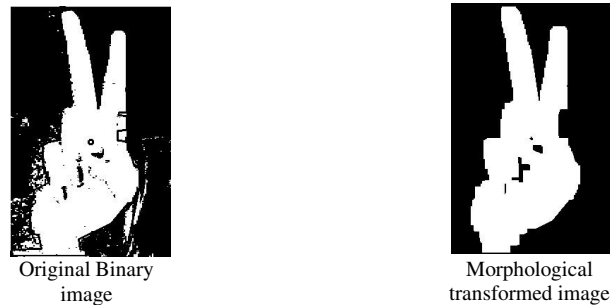


FIGURE 4: Morphological Transformation.

### 3.4. Contour Detection

Detecting contours of a hand image means finding edges that have high contrast pixels than its neighbors. In the real world a hand image may have dark places and shadows not only at the ends but also at the middle of the hand. It will result in a set of contours in various lengths. Filtering the outer contour is done under feature extraction phase.

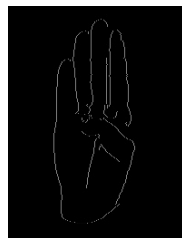


FIGURE 5: Contour Detected.

### 3.5. Feature Extraction

In the feature extraction phase what is most important is to get possible precise features as output. Features selected for classification are hand contour, orientation histogram, convex hull, convexity defects and hu moments.

The convex hull for the hand shape was computed and the poly line minimum area covering box and the rectangle was identified. Then the center of the box was computed. The Region of Interest (ROI) of the image was set to the minimum rectangle. In the algorithm Region of Interest (ROI) plays a major role as it is the area subjected to matching.

## 4. CLASSIFICATION

At the first stage of research, practicing and testing on image processing techniques were done and through that the investigations were carried out to identify the best methods and algorithms for sign language recognition. These attempts were focused to find out the best matching algorithm for sign language recognition. Different researched methods and algorithms in the literature of human computer interaction such as template matching and shape context matching were implemented and tested with modifications to apply them into sign language recognition domain. Normalized template matching was tested with equation,

$$G(x, y) = \sqrt{\sum_{x', y'} T(x', y') \cdot \sum_{x', y'} I(x + x', y + y')^2} \quad (1)$$

Contour matching was tested where it compared the number of points match at the sequences. To make the contour features more powerful, filtration of hand out line contour was used. The biggest contour by the area that each contour covers was computed for filtering process. As the experiment results shows the biggest contour is very much close to hand out line.

Shape match with contour was introduced considering the outer contour as a shape and did the shape matching. Then defects computation for the given real time image was done, where it calculates the total area which is differed from the template, and that was used for classification.

Histogram based classification was implemented with binary image, as histograms can be used to represent the color distribution of an object, an edge gradient template of an object etc. [3]. Here the histogram was generated for the given binary images and the histograms were normalized before matching. As the whole image is used for histogram construction the computational efficiency is very less and the real time concept didn't actually work with this model. As a solution, histograms were computed for contours considering the drawbacks of the previous method, with the idea to reduce the image data for histogram construction. Histogram for a contour is shown in figure 6.

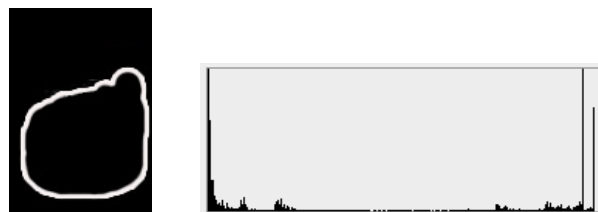


FIGURE 6: Histograms and Contour.

Then the moment based recognition was implemented and tested. Moment based recognition showed more accuracy in identifying the signs than all other attempts mentioned before. In general, (p, q) moment of a contour can be defined as,

$$m_{p,q} = \sum_{i=1}^n I(x_i, y_i) x_i^p y_i^q \quad (2)$$

Here p is the x-order and q is the y-order, whereby order means the power to which the corresponding component is taken in the sum just displayed. The summation is over all of the pixels of the contour boundary (denoted by n in the equation) [16].

#### 4.1. Hu Moments

Hu moments use the central moments where it is basically the same as the moments just described except that the values of x and y used in the formulas are displaced by the mean values.

$$\mu_{p,q} = \sum_{i=0}^n I(x, y) (x - x_{avg})^p (y - y_{avg})^q \quad (3)$$

Where  $x_{avg} = m_{10}/m_{00}$  and  $y_{avg} = m_{01}/m_{00}$

The idea here is that, by combining the different normalized central moments, it is possible to create invariant functions representing different aspects of the image in a way that is invariant to scale and rotation. Seven hu moment calculations by Hu (1962) are shown below [16].

$$\begin{aligned}
 I_1 &= \eta_{20} + \eta_{02} \\
 I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
 I_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
 I_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
 I_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
 &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
 I_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} - \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
 I_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\
 &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
 \end{aligned}$$

As hu moments are rotation invariant the conflicts is aroused between the shapes that have similar shapes in different angles as shown in the figure 7.

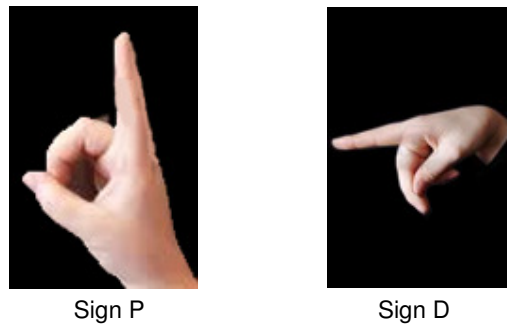


FIGURE 7: Problem of Rotation Invariant.

If sign D is 90 degrees rotated anti clockwise or sign P is 90 degrees rotated clockwise then these two signs have a very similar shape. Therefore, for a rotation invariant algorithm such as hu moments or histograms, these shapes are the same. In [17] the hidden problem behind the unsuccessfulness of Hu moment can be reason out with above discovery. To overcome this problem it requires a new algorithm or any other method.

The solution suggested by researchers is hu moments with height to width ratio filtration. Considering the drawbacks of using Hu-moments, this proposed method works fairly well. As the research found out, problems occur between the signs that are same when rotated. The method proposed by this research gives better performance in recognizing uniqueness between the signs that are rotationally equal.

**4.2. Height to Width Ratio**

First for a given sign the Region of Interest (ROI) is identified using image processing and skin recognition methods. Identification of ROI is shown in figure 8. Thereafter, for the identified hand area the height to width ratio is calculated as shown in table 1.

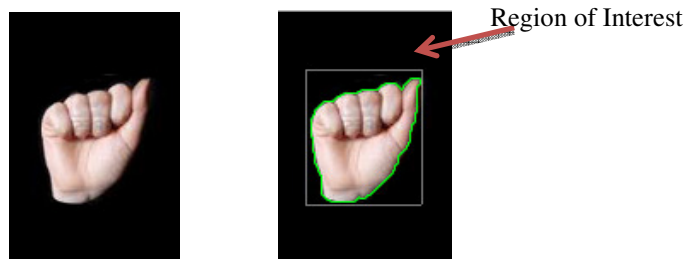







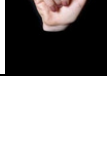


FIGURE 8: Region of Interest Identification.

Sign	Image	Height/Width	Ratio
A		73 / 68	1.07352941176471
B		230 / 98	2.3469387755102
C		124 / 94	1.31914893617021
D		152 / 108	1.40740740740741
L		247 / 245	1.00816326530612
P		169 / 235	0.719148936170213
V		141 / 63	2.23809523809524
Y		55 / 57	0.964912280701754

**TABLE 1:** Height to Width Ratio.

When the ratio is considered, total algorithm remains to be scale invariant and also the rotation is invariant up to certain point. That means as the consideration is on whether the ratio  $\Rightarrow 1$  or ratio  $< 10^\circ$  to  $30^\circ$  deviation of the sign won't affect much on the height to width ratio. However, when the rotation is more than  $30^\circ$ , the new method can recognize the sign correctly. Hu moments classification with height to width ratio is found out to be the best matching method for sign language recognition.

### 5. RESULTS AND DISCUSSION

In finding the actual output of hu moment matching, table 2 display some real time generated hu moment matching result values.





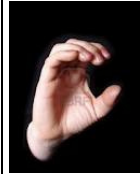





Sign	Template image	Real time image	Result values
A			A - 0.08956746514951 V - 0.27577140714475 L - 0.34207054932742 B - 0.24181056482163 D - 0.37176152384964 C - 0.18130743044722 P - 0.25942242245184 Y - 0.43210990521291
B			A - 0.54202431872088 V - 0.20274582657002 L - 0.77298229967616 B - 0.08136118898609 D - 0.48160990559653 C - 0.32445697847339 P - 0.38891345418583 Y - 0.53140628754324
C			A - 0.22470657026369 V - 0.26110625710516 L - 0.40813802027182 B - 0.27948067338527 D - 0.27798597325585 C - 0.05910328442404 P - 0.14887830150346 Y - 0.34734240316219
D			A - 0.49651217137801 V - 0.34664529950186 L - 0.72065353243200 B - 0.55167658949923 D - 0.24146058292755 C - 0.32495304940430 P - 0.25311353806331 Y - 0.42539378543313
V			A - 0.77538997758076 V - 0.15136323806656 L - 1.0413004951077 B - 0.32666217539152 D - 0.47280905471127 C - 0.52194655877205 P - 0.48663355955186 Y - 0.69101854910652

TABLE 2: Hu Moment Calculation Results.

Hu moments alone was invariant to scale and rotation and succeeded in identifying signs that are not identical in shape than other classification algorithms. But when the sign collection increased with different signs that has same shape with different orientation, Hu moment based classification tend to mislead.



Testing was conducted with 50 signs from 10 users where 5 signs (A, B, C, D and V) from each individual. 8 signs were stored as templates including testing 5 signs and 3 other signs (A, B, C, D, L, P, V and Y). When the test was conducted without the height to width ratio addition to the hu algorithm the accuracy was 76% as Sign D was conflicted with Sign P. Exact similar behavior has been reported in the research by Rodriguez et al. [17]. When the testing was conducted with suggested approach, out of 50 signs only 8 signs failed to identify clearly. That gives an accuracy level of 84%. This research clearly identifies that hu moments can be used to classify sign language recognition by incorporating controlled rotation variance.

## 6. LIMITATIONS AND DRAWBACKS

As this is a real time application, process is highly depends on the user's behavior. Being an appearance based solution, the system is very sensitive to light condition of the environment. In this research only the postures of sign language signs were considered based on the idea that postures are fair enough to identify a certain level of sign language. The spatial information for the hand signs was not considered for classification. There can be some occasions where the same sign posture is presented for different meanings with spatial differences.

## 7. FUTURE IMPROVEMENT AND EXTENSIONS POSSIBLE

Sign language is a very vast area with its grammar and literature component. For the research mainly ASL (American Sign Language) alphabet is used as the test cases. In future, this project can be advanced to convert words, phrases and simple sentences. Further, this project only looks at the hand postures as it is the major part of sign language. In future research, this can be extended to identify hand gestures with spatial information. There a state machine for gesture recognition can be implemented with a machine learning algorithm where it will result in an intelligent sign language recognition system. Movements of the hand sign can be detected by taking the absolute difference of the center of ROI box. As a further improvement, facial expressions can be recognized to get more sophisticated combined idea.

## 8. CONCLUSION

The proposed research has dealt with the sign language gesture/posture and identified the sign and converted that sign into text and voice. The design phase and development phase were conducted together using a more agile software architecture where it improves with more added functionalities and algorithm developments with the evolution of system design. Main part of the design was to find a best fitting algorithm for sign language posture recognition. A series of image processing techniques with Hu-moment classification was identified as the best approach. To improve the accuracy of the system, a new approach; height to width ratio filtration was implemented along with Hu-moments. System is able to recognize selected Sign Language signs with the accuracy of 84% without a controlled background with small light adjustments.

## 9. REFERENCES

- [1] C. Harshith, R.S. Karthik, M. Ravindran, M.V.V.N.S Srikanth, N. Lakshmikhanth, "Survey on various gesture recognition Techniques for interfacing machines based on ambient intelligence", *International Journal of Computer Science & Engineering Survey (IJCSES)* Vol.1, No.2, November, 2010
- [2] Q. Chen, N.D. Georganas, E.M. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features", *Instrumentation and Measurement Technology Conference*, 2007
- [3] W. T. Freeman and M. Roth, "Orientation histograms for hand gesture recognition", *IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition*, Zurich, June, 1995
- [4] M. Roth, K. Tanaka, C. ssman, W. Yerazunis, "Computer Vision for Interactive Computer Graphics", *IEEE Computer Graphics and Applications*, May-June, 1998, pp. 42-53

- [5] S. Belongie, J. Malik, "Shape Matching and Object Recognition Using Shape Contexts", *IEEE Transactions on pattern analysis and machine intelligence*, Vol.24, No.24, 2002
- [6] J. Flusser, "Moment Invariants in Image Analysis", *World Academy of Science, Engineering and Technology*, 2005
- [7] N. Liu, B. C. Lovell, "Hand Gesture Extraction by Active Shape Models", *Digital Image Computing: Techniques and Applications, DICTA '05. Proceedings*, 2005
- [8] C. chang, J. chen, W. Tai and C. Han, "New Approach for Static Gesture Recognition", *Journal of information science and engineering*, 2006
- [9] X. Zhang and Z. Tang, J. Yu, M. Guo, "A Fast Convex Hull Algorithm for Binary Image", *Informatica Oct2010*, Vol. 34 Issue 3, pp.369, 2010
- [10] L. Y. Deng, J. C. Hung, H. Keh, K. Lin, Y. Liu, and N. Huang, "Real-time Hand Gesture Recognition by Shape Context Based Matching and Cost Matrix", *Journal of Networks*, Vol. 6, No. 5, May 2011
- [11] E. Yörük, E. Konukoğlu, B. Sankur, "Shape-Based Hand Recognition", *IEEE transactions on image processing*, vol. 15, no. 7, July 2006, 2006
- [12] Q. Chen, "Real-Time Vision-Based Hand Tracking and Gesture Recognition", *Doctoral Dissertation, University of Ottawa*, 2008
- [13] Y. Fang, J. Cheng, J. Wang, K. Wang, J. Liu, H. Lu, "Hand Posture Recognition with Co-Training", *19th International Conference on Pattern Recognition*, 2008
- [14] G. Kukharev, A. Nowosielski, "Visitor Identification - Elaborating Real Time Face Recognition System", *WSCG Short Communication papers proceedings*, 2004, pp. 157-164
- [15] X. Zabulis, H. Baltzakisy, A. Argyroszy, "Vision-based Hand Gesture Recognition for Human-Computer Interaction", *World Academy of Science, Engineering and Technology*, 2006
- [16] M.K. Hu, "Visual pattern recognition by moment invariants", *Information Theory, IRE Transactions*, 1962, pp. 179-187
- [17] K.C.O. Rodriguez, G.C. Chavez, D. Menotti, "Hu and Zernike Moments for Sign Language Recognition", *Computing Department, Federal University of Ouro Preto*, Brazil, 2012

## 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 10, 2016, IJIP will be appearing with more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

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

### LIST OF TOPICS

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

- Architecture of imaging and vision systems
- Character and handwritten text recognition
- Chemistry of photosensitive materials
- Coding and transmission
- Color imaging
- Data fusion from multiple sensor inputs
- Document image understanding
- Holography
- Image capturing, databases
- Image processing applications
- Image representation, sensing
- Implementation and architectures
- Materials for electro-photography
- New visual services over ATM/packet network
- Object modeling and knowledge acquisition
- 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

- Photographic emulsions
- Prepress and printing technologies
- Remote image sensing
- Storage and transmission

- Photopolymers
- Protocols for packet video
- Retrieval and multimedia
- Video coding algorithms and technologies for ATM/p

## **CALL FOR PAPERS**

---

**Volume: 10 - Issue: 1**

**i. Submission Deadline :** January 31, 2016

**ii. Author Notification:** February 28, 2016

**iii. Issue Publication:** March 2016

## **CONTACT INFORMATION**

### **Computer Science Journals Sdn Bhd**

B-5-8 Plaza Mont Kiara, Mont Kiara  
50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627

Fax: 006 03 6204 5628

Email: [cscpress@cscjournals.org](mailto:cscpress@cscjournals.org)

CSC PUBLISHERS © 2015  
COMPUTER SCIENCE JOURNALS SDN BHD  
B-5-8 PLAZA MONT KIARA  
MONT KIARA  
50480, KUALA LUMPUR  
MALAYSIA

PHONE: 006 03 6204 5627

FAX: 006 03 6204 5628

EMAIL: [cscpress@cscjournals.org](mailto:cscpress@cscjournals.org)