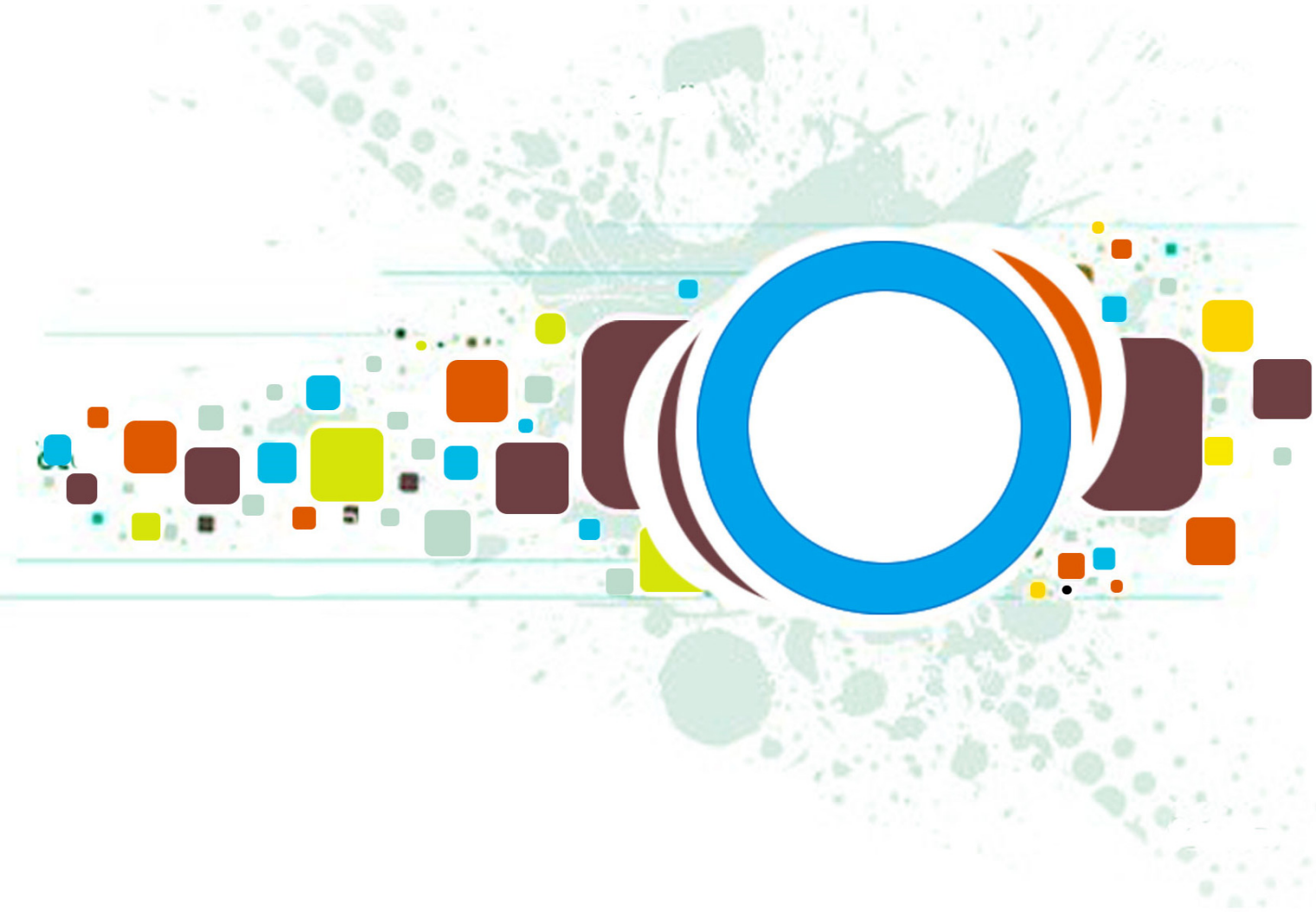
Volume 9 • Issue 4 • July / August 2015

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 4, 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 4, July / August 2015
Publishing Date: 31-08-2015
ISSN (Online): 1985-2304

IJIP Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation 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 *Fourth* 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. Started with Volume 9, 2015, IJIP appears 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

TABLE OF CONTENTS

Volume 9, Issue 4, July / August 2015

Pages

Development of A Smart Interface For Safety and Protection of Automotives

AkshayKumar Yalkar *akshaykumar.yalkar1@gmail.com*
Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

Harsha Vardhan Singh N *harshavardhansingh7@gmail.com*
Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

Jagadevan V *jagadevan.vijay.94@gmail.com*
Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

Nandini C *nandininandy3@gmail.com*
Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

K. Ezhilarasan *murali981983@gmail.com*
Assistant professor, Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

Pushpa Mala S *pushpasiddaraju@gmail.com*
Assistant professor, Dept. of ECE
Sambhram institute of technology
Bengaluru, Karanataka, India

Abstract

This paper is mainly directed towards the safety and protection of the human beings by synchronizing both the software and hardware modules. Automotive safety sensors are mainly streamed towards the application in automobiles. The safety and protection of the automobile driver is monitored and abnormalities are detected by these sensors. These abnormalities are highlighted and alerts are provided to the driver, by the combinational synchronization of hardware and software.

Keywords: Eye Blink, Heartbeat, Alcoholic Detection, Object Detection, Face Turn Detection.

1. INTRODUCTION

As we aware of the increase in population rate, there is a rapid increase in the usage of automobiles. This has led to the increase in the average percentage of accidents. This is due to various reasons. One of the main reason is due to human errors in driving method. This can be avoided by adopting technology in the automobile. A smart interface system for an automobile mainly deals with safety of passengers by continuously monitoring the activities of the driver. These monitoring activities can be performed by different modules involved in the smart interface

system such as Eye blink detection, Face turn, Heartbeat, Alcohol detection, GSM module, Object detection. Activities of driver are monitored by these modules and the outputs are interfaced with alerting section which generates audio and visual warnings to alert the driver in abnormal conditions to avoid accidents.

Several smart interface systems exist in literature. Boon-Giin Lee and Wan- Young Chang [1] proposed an eye blink detection using IR light. IR light is harmful to the human eye when it is used for long period. Lai and Liu [3] developed a fuzzy-control massage seat to keep drowsy drivers awake. Bergasa *et al.* [4] proposed a nonintrusive prototype of a computer vision system for monitoring driver's attentiveness in real-time. Kasukabe *et al*. [5] developed a system with visual, cognitive, and decision- making functions for elderly drivers. This system was able to recognize the objects encountered during driving.

Pauwelussen and Feenstra [6] developed a traffic-simulation model in which the vehicle is equipped with an adaptive cruise-control (ACC) and lane- departure warning (LDW) system to monitor the driver's behavior in a real traffic environment. Lee *et al*. [7] proposed a system with two fixed cameras to capture images of the driver and the road respectively. These images are mapped to the global coordinates to monitor the driver's line of sight. The authors found four distinctive driving patterns through analysis by a hidden Markov model (HMM). Zhao *et al*. [8] studied the reliability of steering wheel behavior to detect driver fatigue by multi wavelet packet energy spectrum using a support vector machine (SVM).

Lee and Chung [9] developed a video sensor- based eye-tracking and blink-detection system with Haar-like features and template matching for an automated drowsiness warning system. In addition, Yanget *et al*. [10] demonstrated that drowsiness has a greater effect on rule-based driving tasks than on skill-based tasks using a Bayesian network (BN) paradigm through simulator-based human-in-the-loop experiments. Wang and Gong [11] proposed system. This system adopted a latent variable to represent the attributes of individual drivers for recognizing the emotional state of drivers. Four sensors, each for respiration, skin conductance, temperature, and blood pressure is used. Shin *et al*. [2] proposed the design of an electrocardiograph (ECG) and photoplethysmography (PPG) sensor to measure the driver's metabolic condition. Eye blink was also detected using EEG signal from the neurons. The EEG signals is detected through the electrode placed on scalp of the head in the form of wearable helmet. This device is uncomfortable for the human and when used for a long time. In order to overcome this drawback, we use a webcam to detect the eye blink. ECG signal are used to monitor the heartbeat.

2. PROPOSED METHOD

This paper mainly concentrates on the safety of the automotive by continuously monitoring the driver activities. The complete architecture of the proposed method is shown in FIGURE.1. The complete Process flow is depicted in FIGURE.2. The main modules assisting in the monitoring action are:

2.1 Eye Blink Detection

The webcam is used to continuously monitor the eye blink. The eye blink rate is processed in two ways. Firstly, a normal eye blink rate represents that the driver is normal during driving activity. The other case is when the eye blink rate is not normal i.e. when the eye is closed for certain period of time. The latter represents abnormal eye blink during the driving activity. The abnormal condition is detected by the eye blink detection module. This may indicate that the driver is

drowsy and lead to accidents. This must be detected and the driver must be alerted. To avoid this situation the module generates an alert through a buzzer. The buzzer alerts the driver and he resumes normal eye blinking state. This concept is mainly directed towards the safety and protection of the automobile.

2.2 Heart Beat Detection

The heart beat detection module is used to continuously monitor the heart beat rate of the driver. Abnormal heart beat rate of the driver is detected. Under such conditions, two simultaneous actions take place. Firstly, the speed of the car is reduced. Secondly, an alert is generated through the buzzer and finally a visual alert is given by the LCD. This visual alert displays a message denoting a normal or abnormal heart beat rate to the driver.

On the other hand, if the heart beat rate of the driver is normal, the alert system remains off.

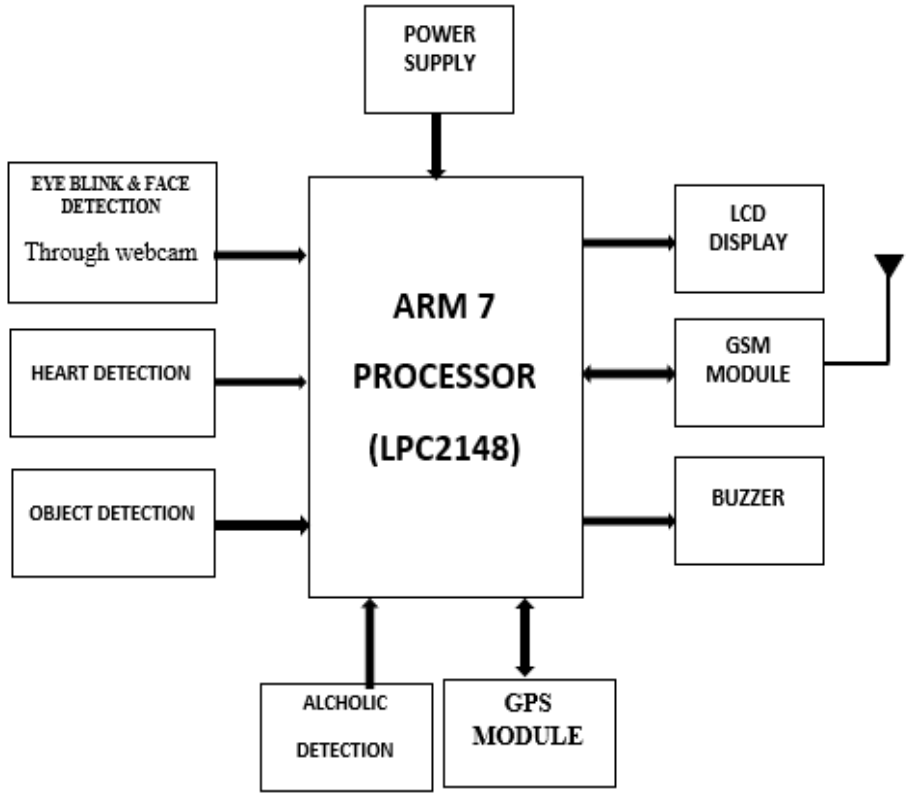


FIGURE 1: Complete Architecture of Proposed Method.

2.3 GSM Module

This module is mainly used for alerting the second party, when the first party (driver) is in an abnormal state. This is achieved by synchronizing the continuously monitored eye blink rate and the pulse rate. In other words, this indicates that eye is closed for a certain period of time (abnormal state) during driving activity and the pulse rate is in abnormal state. When both the condition are synchronized, this module generates an alert to the second party.

2.4 Face Detection

The face detection module is used for detecting the position of the drivers face. This indicates that the driver is visualizing in some direction with his face directed towards an abnormal position

than the normal driving position during the driving activity for a certain period of time. This generates an alert through a buzzer. This alerts the driver and he resumes normal visualization.

2.5 Obstacle Detection

This feature is used to detect the obstacle which is present behind the automobile. In some situations, this is normally not visible to the driver. This module detects the obstacle and then alerts the driver. This is achieved by using an LCD. IR transmitter and receiver is used for object detection.

2.6 Location Tracker

Location tracker is used to identify the location of the automobile in emergency situations.

2.7 Alcoholic Detection

This module works for two different conditions. Firstly, if the driver is drunk, before he starts the automobile, his attempts to start the vehicle fails. Secondly, if the driver consumes alcohol while driving the automobile, the vehicle slows down and finally stops. MQ3 Sensor is used for alcohol detection.

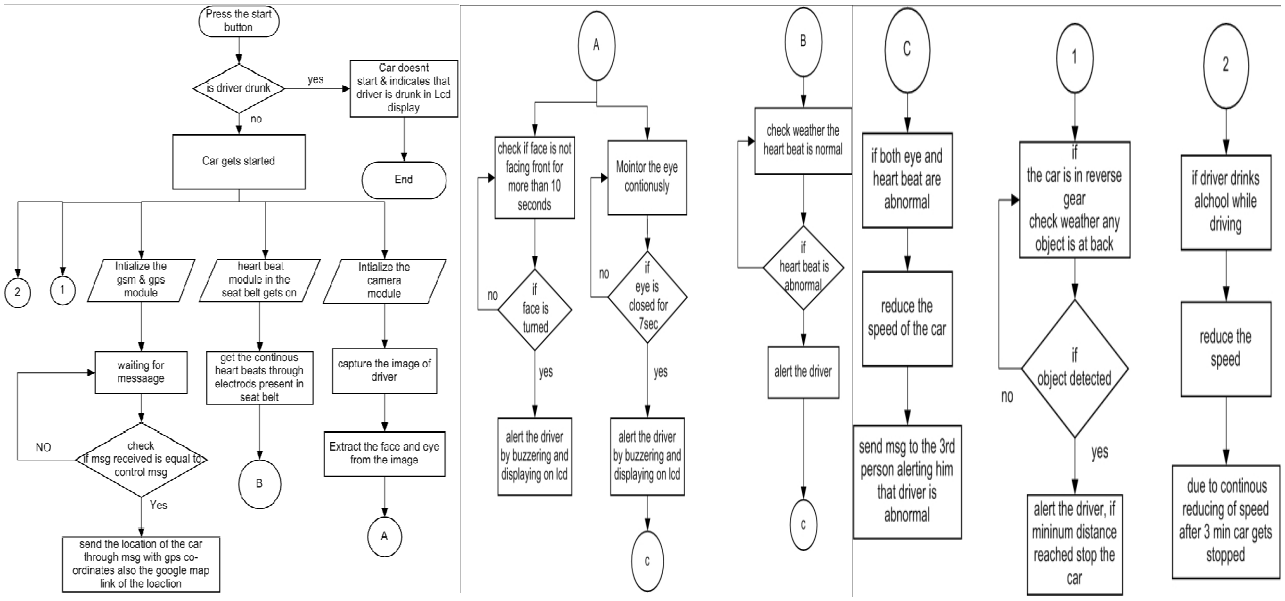


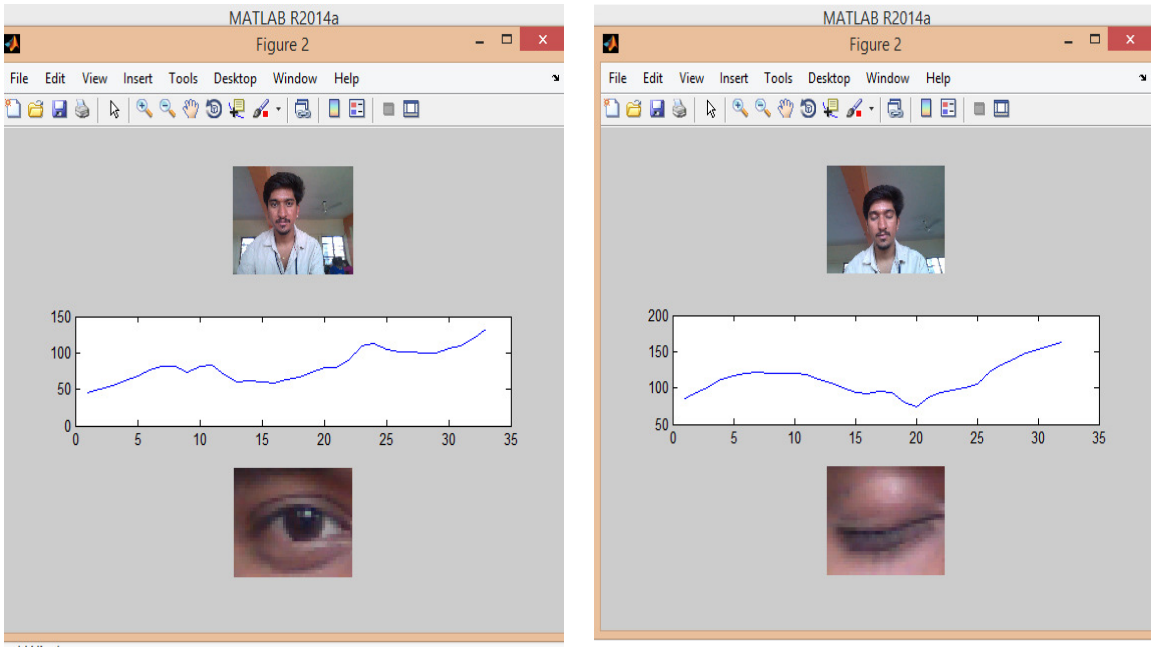
FIGURE 2: Flow Chart for the Proposed Method.

3. RESULTS AND DISCUSSIONS

Webcam is used to capture the images, Vision Cascade Tool in the Matlab Tool Box is used for the detection of eye and face in captured image. Detection is done based on Viola-Join Algorithm. Here, the retina of the eye is monitored. If retina of the eye is not detected, then it indicates that eye is closed. The simulated output images are given in FIGURE 3. Fig 3a shows the eye is open the corresponding graph is plot of pixel point vs pixel intensity, the two graphs show the diff between eye open and eye close.

Heart beat module detects the ECG signal from the heart. This module mainly consists of electrode, amplifier, and analog to digital converter. Since, ECG signals are weak signals, they are amplified using ECG amplifier. Later they are converted into digital form to interface with the processor/ controller. Digital value is compared with pre-programmed normal heart beat value. When this value doesn't match with later, the output goes high.

Alcoholic detector is basically a breath analyzer which detects the alcoholic content in the breath. In case alcohol is detected output of this module goes high. Figure 4 shows the location of various sensor in the vehicle. The normal and abnormal conditions is depicted in Table 1.



a) b)

FIGURE 3: Simulated Output Images a) Eye Open, b) Eye Close.

Fig 4a. Webcam & Alcohol

Fig 4b. Heart Beat Sensor.

Fig 4c. Complete Model.

FIGURE 4: Location of sensors in the Vehicle.

Table 1: Different parameters of proposed system.

Parameters	Normal	Abnormal
Eye close (drowsiness)	Less than 4sec	More than 4sec
Face turn	Less than 4sec	More than 4sec
Heart beat	More than 68 less than 75	Less than 68 more than 75
Alcohol	Less than 30%	More than 30%

4. CONCLUSION

This paper mainly highlights the factors responsible for the safety and protection of the automobile driver. The interrupts responsible for causing accidental acts is being detected leading to a safer pathway. In this paper an attempt is made to describe various modules used to avoid the causes of accidents caused by human errors during driving period such as drowsiness and face turn. This paper also concentrates on drivers violating traffic rules i.e. consumption of alcohol during driving that leads to accidents. We have also made an attempt to provide medical aid during abnormal conditions.

5. REFERENCES

[1] Boon-Giin Lee and Wan-Young Chung, "Driver Alertness Monitoring Using Fusion of Facial Features and Bio-Signals", IEEE SENSORS JOURNAL, VOL. 12, NO. 7, JULY 2012.

[2] H. S. Shin, S. J. Jung, J. J. Kim, and W. Y. Chung, "Real time car driver's condition monitoring system," in Proc. IEEE Sensors, Nov. 2010, pp. 951–954.

[3] R. L. Lai and C. L. Liu, "A fuzzy control massage seat for awaking drowsy drivers," in Proc. 7th Ind. Eng. Manage. Syst. Conf., Bangkok, Thailand, 2006, pp. 618–623.

[4] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," IEEE Trans. Intell. Transport. Syst., vol. 7, no. 1, pp. 63–77, Mar. 2006.

[5] T. Kasukabe, M. Hiraoka, O. Yamamoto, M. Yamada, and T. Nakano, "Development of system for comprehensively measuring driving ability for elderly safe driving," in Proc. Conf. Mach. Vis. Appl., Yokohama, Japan, May 2009, pp. 443–446.

[6] J. Pauwelussen and P. J. Feenstra, "Driver behavior analysis during ACC activation and deactivation in a real traffic environment," IEEE Trans. Intell. Transport. Syst., vol. 11, no. 2, pp. 329–338, Jun. 2010.

[7] J. D. Lee, J. D. Li, L. C. Liu, and C. M. Chen, "A novel driving pattern recognition and status monitoring system," in Proc. Pacific-Rim Symp. Image Video Technol., Hsinchu, Taiwan, Dec. 2006, pp. 504–512.

[8] S. F. Zhao, G. H. Xu, and T. F. Tao, "Detecting driver's drowsiness using multiwavelet packet energy spectrum," in Proc. Int. Congr. Image Signal Process. Tianjin, China, Oct. 2009, pp. 1–5.

[9] Y. S. Lee and W. Y. Chung, "Video sensor based eye tracking and blink detection to automated drowsy driving warning system using image processing," in Proc. 13th Int. Meet. Chem. Sensors, Perth, Australia, Jul. 2010, p. 358.

[10] J. H. Yang, Z. H. Mao, L. Tijerina, T. Pilutti, J. F. Coughlin, and E. Feron, "Detection of driver fatigue caused by sleep deprivation," IEEE Trans. Syst. Man Cybern. Part A.: Syst. Humans, vol. 39, no. 4, pp. 697–705, Jul. 2009.

[11] J. Wang and Y. Gong, "Recognition of multiple drivers' emotional state," in Proc. 19th Int. Conf. Pattern Recognit., Tampa, FL, Dec. 2008, pp. 1–4.

High Dynamic Range Imaging- A Review

Nayana A. *nayana239@gmail.com*
Department of Electronics and Communication Engineering
Mar Baselios College of Engineering and Technology
Trivandrum, India.

Anoop K. Johnson *anoopkjohnson2340@gmail.com*
Department of Electronics and Communication Engineering
Mar Baselios College of Engineering and Technology
Trivandrum, India.

Abstract

The real world scenes have a very wide range of luminance levels. But in the field of photography, the ordinary cameras are not capable of capturing the true dynamic range of a natural scene. To enhance the dynamic range of the captured image, a technique known as High Dynamic Range (HDR) imaging is generally used. HDR imaging is the process of capturing scenes with larger intensity range than what conventional sensors can capture. It can faithfully capture the details in dark and bright part of the scene. In this paper HDR generation method such as multiple exposure fusion in image domain and radiance domain are reviewed. The main issues in HDR imaging using multiple exposure combination technique are Misalignment of input images, Noise in data sets and Ghosting artefacts. The removal of these artefacts is a major step in HDR reconstruction. Methods for removing misalignment and noise are discussed and detailed survey of ghost detection and removal techniques are given in this paper. Single shot HDR imaging is a recent technique in the field of HDR reconstruction. Here instead of taking multiple exposure input images, a single image is used for generating HDR image. Various methods for Single shot HDR imaging are also reviewed.

Keywords: High Dynamic Range Imaging, Multiple Exposure Fusion, Image Registration, Ghosting Artefacts, Single Shot Imaging.

1. INTRODUCTION

Dynamic range (DR) of a scene is defined as the range from the lowest to the highest light intensity occurring in the scene. It is also known as scene contrast. An image is said to be in high dynamic range if it has details in both the bright and dark regions in the image. The dynamic range of real-world scenes can be as high as in the ratio of 100,000:1. Human visual system can process and adapt to a dynamic range of about 50,000:1. The images captured by the cameras can only have dynamic ranges between 300:1 to 1,000:1. These images are therefore considered as low dynamic range (LDR) images. Longer exposures would capture details in the dark areas of a scene while shorter exposures would capture the bright areas of a scene. That is, the details in both the dark and bright areas of a scene are not visible in a single image. Such a type of image is known as LDR image.

High Dynamic Range (HDR) images capture the whole tonal range of real-world scenes, using 32-bit floating-point values to store each colour channel. The use of floating point values gives HDR images several advantages over LDR images. The areas that are too dark are clipped to black and areas that are too bright are clipped to white in an LDR image. Pixel values in an HDR image are normalized between 0.0 and 1.0. Dark and bright areas are assigned different values close to 0.0 and 1.0. Therefore, HDR images can preserve details of a scene having large dynamic range. Another advantage of HDR images is that they preserve optical phenomena such as reflections and refractions. In LDR images, the pixels representing all the bright light sources

in a scene are assigned to have the maximum possible integer value and the reflected light should have value less than the light source. In HDR images, the reflected light is assigned with values close but less than 1.0 and the bright source can assume values equal to 1.0. Therefore, HDR images are able to better represent scenes as perceived by human eyes.

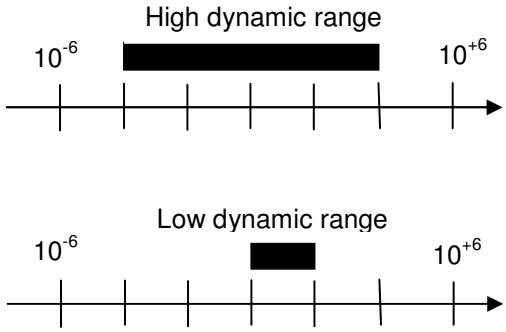


FIGURE1: High Dynamic vs. Low Dynamic Range.

2. HIGH DYNAMIC RANGE IMAGING

High dynamic range (HDR) imaging refers to a set of techniques that allow a greater dynamic range of luminance between the lightest and darkest areas of an image than standard digital imaging techniques. The wide dynamic range of HDR images allows representing the wide range of intensity levels found in real scenes. HDR has the following qualities:

(i) High contrast ratio: Bright and dark regions need to be captured.
(ii) High bit depth: To encode values with quantization levels as small as the just noticeable difference; so that no stepping is visible resulting in a smooth intensity gradient.
(iii) Details are preserved: There is no or little clipping due to over- or under-saturation.

To directly capture HDR content, specialized cameras are proposed but these devices are expensive. The approach commonly used is to take sequential LDR images at different exposure levels (known as bracketed exposures) and then merge them into an HDR image. This method is known as multiple exposure technique.

High dynamic range (HDR) imaging consists of:

(i) HDR recovery or acquisition: In HDR acquisition the true luminance of a scene is captured with a low dynamic range capturing device.
(ii) Tone mapping: To faithfully represent HDR information on a low dynamic range display device that is incapable of reproducing the actual luminance and the true colours.

2.1 Tone Mapping

One of the problems associated with high dynamic range imaging is the display of high dynamic range radiance maps on conventional reproduction media such as CRT or LCD monitors, and projectors. The dynamic range of the generated HDR image often spans more than five orders of magnitude (2^{16} = 65536). But a conventional display is able to visualize a maximum of two orders of magnitude (2^{8} = 256). A solution to this problem is to compress the dynamic range of the radiance maps such that the mapped image can be fitted into the dynamic range of the display devices. This method is called tone mapping [1,2]. Here one set of colors are mapped to another to approximate the appearance of high dynamic range images in a medium that has a limited dynamic range. Tone mapping operator reduces the scene radiance contrast to the displayable range while preserving the image details and color appearance.

There are several tone mapping methods and these methods can be divided into two broad categories. The global tone mapping techniques use a single spatially invariant mapping

function for all pixels in the image. It is much simpler to implement but tend to lose details. The local mapping techniques adapt the mapping functions to local pixel statistics. It is much more computationally intensive as there are a number of parameters which have to be set empirically. Depending on the particular requirements, one method will be better suited than others in some cases, or a combination of methods will be necessary in some other cases.

3. MULTIPLE EXPOSURE FUSION TECHNIQUE

The most common method of HDR generation is to take multiple images of the same scene with different exposure times, and combine them into a single HDR image. The basic concept is that each pixel will be properly exposed in at least one image. A single photo is insufficient to record all the details in the scene because, some areas in the photo may be over-exposed and other portions may appear under-exposed ie, details can hardly be seen (Figure 2). These limitations motivate the development of fusion techniques for multi-exposure images. The fusion of a set of LDR images can be classified into two main approaches: fusion in the radiance domain and fusion in the image domain.

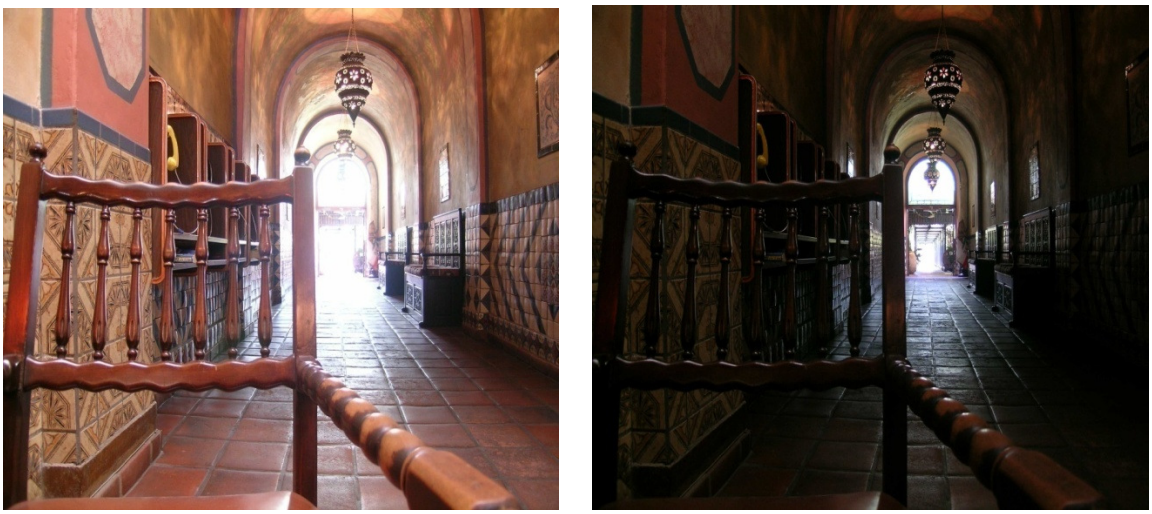


FIGURE 2: Multiple Exposure Images.

FIGURE 3: HDR Image.

4. HDR RECONSTRUCTION IN RADIANCE DOMAIN

HDR image generation process in radiance domain consists of three steps.

(i) Recover the camera response function to bring the pixel brightness values into radiance values. This function models the effect of non-linearities introduced in the image acquisition process. Different methods can be used for its estimation.
(ii) Then all radiance maps are combined into an HDR image.
(iii) To make the HDR image displayable on LDR monitors a tone mapping operator is used.

HDR image is computed as the weighted average of pixels values across exposures using the camera response function.

$$I_m^{hdr} = \frac{\sum_{n=1}^{N} w(I_m^n)[g^{-1}(I_m^n)/t_n]}{\sum_{n=1}^{N} w(I_m^n)}$$

where I_m^n is value of pixel *m* of the *n*-th exposure image, *N* is the number of images, t_n is the exposure time of the *n*-th image, and *w* is a weighting function. *w(.)* is assigned to reduce the influence of unreliable pixels. In order to display the obtained HDR image on a low dynamic range device, a tone mapping operator is applied. Tone mapping techniques can be classified into global and local methods.

In [3], an image with extended dynamic range is obtained from a sequence of exposures using the function

$$M = \alpha + \beta I^{\gamma}$$

to model the response curve *g* of the imaging system. α is the bias parameter and β is the scaling factor. α is obtained by taking an image with the lens covered and then performing a regression operation. Consider two images taken under different exposures with known ratio $R = e_1/e_2$. The measurement $M(I)$ in the first image (I is unknown) produces the measurement $M(RI)$ in the second. A pixel with brightness $M(RI)$ is sought in the first image that would then produce the brightness $M(R^2 1)$ in the second image. This search process is repeated to obtain the measurement series $M(I)$, $M(RI)$, ..., $M(R^n I)$. To these samples regression is applied to estimate the parameter y. In [4] also the radiometric response function is computed from a sequence of differently exposed images. This algorithm does not require precise estimates of the exposures used. Rough estimates of the ratios of the exposures are sufficient for accurate recovery of the response function as well as the actual exposure ratios. They improved [3] by assuming f^{-1} to be a polynomial and its coefficients are determined through a regression model. This response function is then used to fuse the multiple images into a single HDR radiance image. The response function of the imaging system is estimated using the principle of reciprocity in [5]. Using the recovered response function the multiple LDR images are fused into HDR radiance. They worked with logarithmic data and did not impose any restrictive analytic form to f^{-1}, yet they required it to be smooth by adding a penalization term proportional to the second derivative of f^{-1} to the optimization problem. This work is applicable in many areas of computer graphics and has become the de-facto standard in the field of HDR imaging with static conditions.

5. HDR RECONSTRUCTION IN IMAGE DOMAIN

In this category multiple exposures are combined directly without computing camera response function. These methods combine LDR images by preserving only the best parts of each exposure. Final HDR image is obtained as a weighted average of pixel values across exposures:

$$I_{ij}^{C} = \sum_{k=1}^{N} w(Z_{ij,k}) Z_{ij,k}$$

where I^C is the composite image.

The image domain technique of merging is mentioned in [6] which is based on image partitioning and merging. Differently exposed input images are firstly divided into some sub regions and a gradient method is implemented to judge the well exposed one. In another technique [7], an image with extended dynamic range is obtained simply by using quality measures like saturation, contrast and well exposedness. Input images in the stack may contain under- and overexposed regions. These regions should receive less weight while areas which contain color and brightness details should be preserved. Three quality measures are used to achieve this. To obtain the contrast measure C a Laplacian filter is applied to the gray scale version of each image and the absolute value of the filter response is taken. A saturation measure S is computed as the standard deviation within the R, G and B channel, at each pixel. To obtain the well exposed pixels E the intensities that are not near zero (underexposed) or one (overexposed) must be kept. Each intensity is weighted based on how close it is to 0.5 using a Gauss curve. There is no need for the estimation of camera response function. The quality of resulting image is comparable to existing tone mapping operators. As in [8] differently exposed raw images can also be used for HDR generation. The effective dynamic range can be extended to 256 times if five differently exposed raw images are fused. Edge detection iterations are used to extract the image details.

5.1 Comparison between Fusion in Radiance and Image domain

The two different HDR image generation processes are depicted in Figure 4. Methods that combine images in the radiance domain are highly relied on accurate estimation of the camera response function, which is sensitive to image noise and misalignment. These methods require tone mapping operators for HDR image reproduction. A true HDR radiance map is obtained in the combination step which contains the whole dynamic range of the captured scene. Methods that combine exposures in the image domain are more efficient since they avoid the estimation of the camera response function and do not require tone mapping. They directly produce a tonemapped-like HDR image.

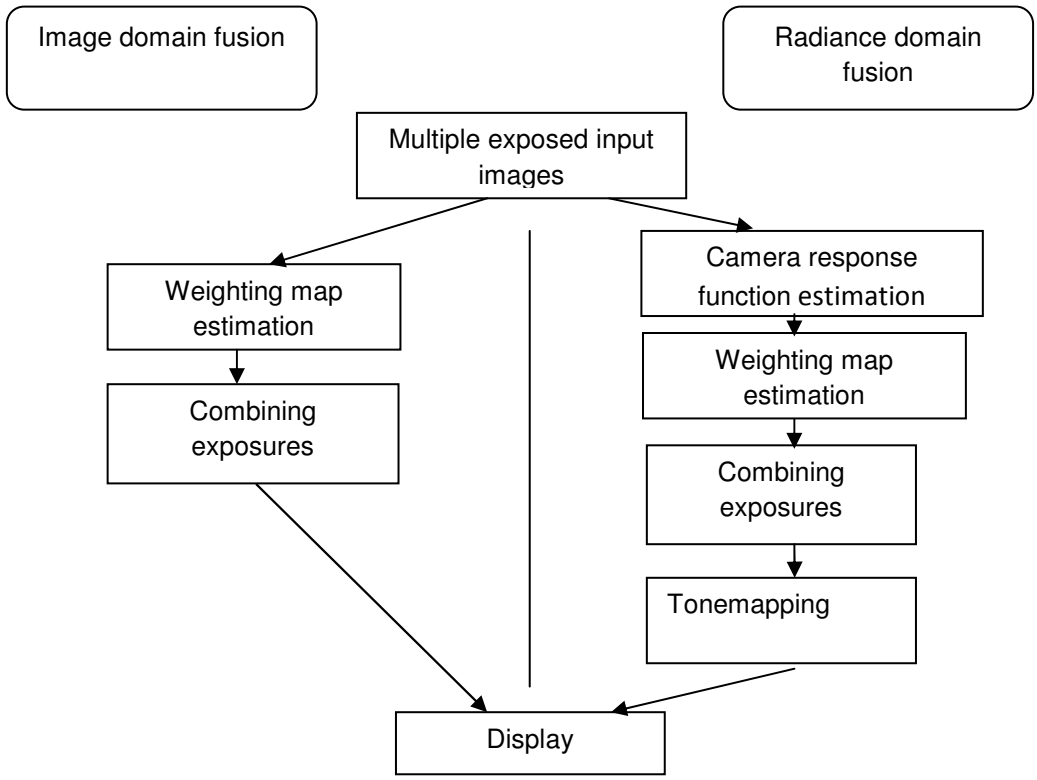


FIGURE 4: Comparison of Image and Radiance Domain Fusion.

6. ARTEFACTS IN HDR IMAGE

Multiple exposure technique suffers from three main problems:

(i) Misalignment: Global camera motion, from hand-held camera for instance, results in misaligned input images that cause the combined HDR image to look blurry. This problem can be solved by placing the camera on a tripod or by using an image registration method. Image registration is the process of transforming different sets of data into one coordinate system. In particular, the median threshold bitmap (MTB) technique [9,10] is an efficient solution. The translation alignment based on median threshold bitmaps [10] is extended with an additional rotation [9].This method is fast and can accurately recover the small displacements between images. Other registration methods based on key points extraction and matching can also be used. The most commonly used key points detectors are Harris corners and SIFT features.

(ii) Noise: The irradiance estimation that is used in multiple exposure technique is generally performed on a per-pixel basis. That is, for a given pixel, the estimation only relies on the observed values at the same location for all the available exposures. This estimation scheme does not take advantage of the redundancy present in most images. This problem can be solved by making use of a bilateral filtering during the frame fusion step [11]. However, such a local filtering cannot cope with large motions. A patch-based approach [12] is used in case of large motions, to find similar pixels that are then combined for the irradiance estimation thus making use of redundancy present in the input images.

(iii) Ghosting: Moving objects in the scene will appear in different locations in the combined HDR image, creating what are called ghosts or ghosting artefacts. Ghosting artefacts are the most severe limitation of the multiple exposures technique since moving object is unavoidable.

The Figure 5 denotes three level exposures of the scene with moving object [13]. Those exposures are captured in different time and using different shutter speed. In each image the sun (moving object) is located at a different location leading to ghosting artefact in the resultant HDR image.

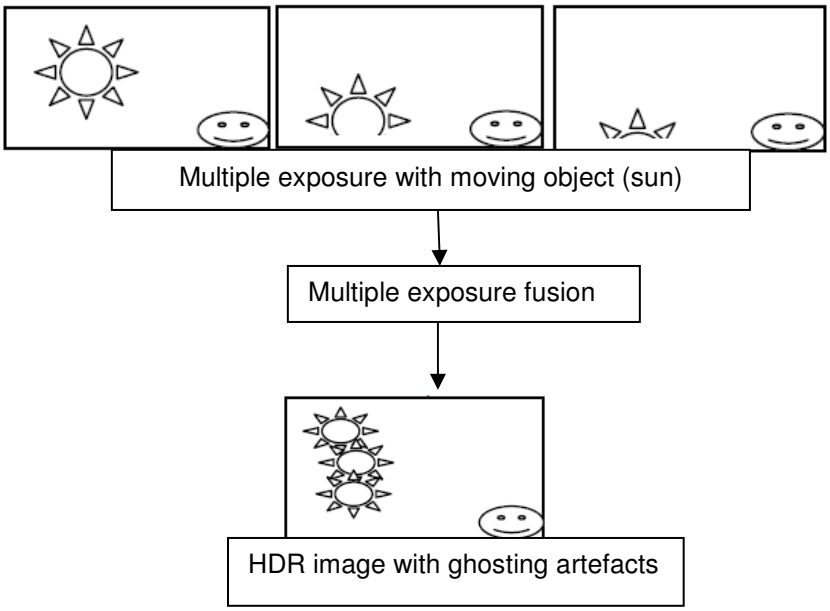


FIGURE 5: Framework for HDRI with Ghost Artefact.

FIGURE 6: An Example for Ghosting Artefact.

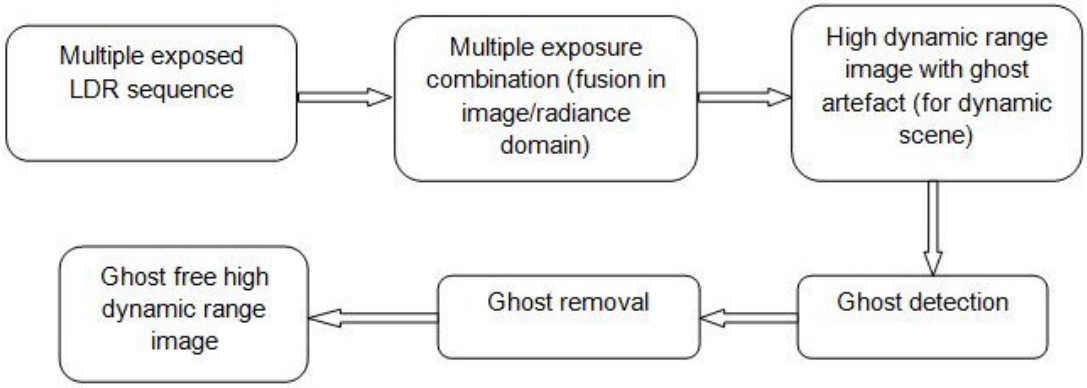


FIGURE 7: The Overall Process of HDRI.

7. GHOST DETECTION METHODS

Most of the HDR generation methods employ a two-step strategy: ghost detection and ghost removal. Ghost detection methods are based on motion detection. There are two types of motions in a dynamic scene:

(i) a moving object on a static background.
(ii) a moving background with static or dynamic objects. Some of the methods can detect only the first type of motion while others can detect both.

Weighted variance measure [14,15] can be used to detect ghost regions. The camera response function is first estimated using which the radiance maps are computed. By evaluating the variance of the radiance values a variance image (*VI*) is generated. In the *VI* image regions affected by movement exhibit high variance.

A high contrast movement occurs when the moving object is different from the background, can be detected using variance measure. While a low contrast movement occurs when the dynamic object and the background are similar in color and this type of motion can be detected using entropy measure [15]. For each input image an entropy map is computed. An Uncertainty Image (*UI*) is then derived from the weighted difference of the entropy images. This uncertainty image is used to find ghost regions based on thresholding. In another method, the deviation between the predicted intensity value of a pixel in first image and the actual intensity value in second image is used to decide whether that pixel is a ghost pixel using camera response function [9]. If there is

significant difference between predicted and the actual value then that pixel is a ghost pixel. Order relation [16] between pixel values in differently exposed images can be used to find ghost areas. Moving areas in the scene can also be detected based on multi-level threshold maps [17]. The multi-level threshold maps are computed by classifying the intensity values of the images into P levels using P thresholds. Then the ghost map estimate is generated using the multi-threshold maps. In [18], ghost regions are detected based on median bitmaps which impose relations between pixels in each single exposure. If a pixel is not affected by ghost, then its relation to the median intensity of the image must be the same in all LDR images. The ghost map is recovered from the median bitmaps. Patches of ghost regions can be detected using RANSAC method [19]. In this method the intensity values at any location (*u, v*) in any two input images *Lk* and *Ll* are related by:

$$\frac{Z_{uv}^{k}}{\Delta t_k} = \frac{Z_{uv}^{l}}{\Delta t_l}$$

The above equation deviates only at locations affected by ghosts, apart from saturated pixels. In another method of ghost detection [20], the second biggest singular values extracted over local spatio-temporal neighbourhoods can be used for ghost region detection. The second biggest singular value should be small or approximately equal to zero for a ghost free region. This is the best ghost detection method since it gives higher sensitivity and specificity values, i.e. it correctly detects almost all ghost pixels with very few false positives.

8. GHOST REMOVAL METHODS

Ghost removal methods can be classified into two. The first category is to remove ghosting artefacts while keeping a single occurrence of the moving object. The second method will completely remove the moving object in the image.

a) Keeping Single Occurrence of Moving Object

If the moving object is of interest, then it is desirable to keep it at a fixed location in the final HDR image rather than completely removing it.

Ghost removal techniques are based on the detected ghost map. The approach is to apply the standard multiple exposure fusion method in ghost-free regions while selecting a single reference exposure in ghost affected areas [14,15,9]. Another approach [19] is to determine the correct number of exposures to use in different ghost affected areas. This number is obtained as the number of images in which the patch does not deviate from the patch in the reference image. Then HDR image is built using different number of exposures on each detected ghost region. For a seamless composition of exposures, Pece and Kautz [18] and Mertens et al. [21] use a Laplacian pyramid blending at multiple resolutions [22]. To avoid boundary effects in the final HDR image [19] uses a gradient domain approach. Laplacian pyramid blending and Poisson editing frameworks are used to avoid boundary effects introduced by using a single reference exposure in ghost affected regions. A simpler method which produces good results is based on weights adaptation [17,14, 18, 11]. Another ghost-free HDR image generation method using gradient information is proposed in [23]. The gradient direction in stationary regions remains stable in different exposures and varies if the content changes due to object movement.

b) Completely Removing Moving Objects

To achieve this goal, the approach is to discard exposures that are affected by ghosting at each pixel location in the combination step. This idea is used in [16] where for each pixel location two sets of exposures are created. First exposure set contains ghosting at location (u, v), while the second represents exposures that do not contain ghosting. Combining only exposures in second set leads to a ghost-free HDR image. A similar algorithm proposed in [19] is based on image patch processing rather than working with pixels individually. Other methods [24, 25] directly remove ghosting by adjusting the weighting function used in the HDR image generation equation. These methods iteratively change pixels weights to minimize the number of visible artefacts. In [24] a kernel density estimation method is used that iteratively estimates the probability that a

pixel belongs to the static part of the scene. In [25] bandwidth matrices are estimated for computing the accurate probability that a pixel belongs to the background. The final probabilities are used as weights in the HDR generation.

9. SINGLE SHOT HIGH DYNAMIC RANGE IMAGING

Multiple exposure technique of HDR generation has several drawbacks, such as the need for alignment and motion estimation to avoid artefacts. In order to avoid these limitations, another technique where a large range of exposures are captured in only one shot (single shot) using spatially varying pixel exposures (SVE) is proposed in [26]. For controlling the amount of light that reaches each pixel an optical mask with spatially varying transmittance is used. This gives different exposure levels to the pixels according to the given transmittance pattern. The main drawback of the SVE acquisition is that the under and over exposed pixels are unknown. For the reconstruction of these unknown pixels the aggregation or the interpolation approaches can be performed [26]. Aggregation approach consists in averaging the local irradiance values produced by the correctly exposed pixels. The interpolation approach is based on bi-cubic interpolation [27]. To avoid aliasing problems spatially varying exposures in a non-regular (random) pattern is used [28]. The irradiance image is then reconstructed using a frequency selective extrapolation algorithm [29]. To reconstruct the unknown pixels and denoise the known ones another method using Gaussian Mixture Model (GMM) is proposed in [30]. Another method [31] utilizes weighted histogram separation (WHS) to estimate the threshold for histogram separation. Then it divides the histogram into two sub-histograms and generates differently exposed (over and under) LDR images from a single image. The drawback of this method is that utilizes a fixed weighting factor for histogram separation which is not suitable for images displaying different characteristics. In a recent method [32] over- and under-exposed images are generated from a single input image by making use of adaptive histogram separation. HDR image is constructed from the generated differently exposed LDR images by making use of a fuzzy based fusion approach.

10. CONCLUSION

In this paper various methods of high dynamic range image generation such as multiple exposure technique and single shot imaging are reviewed. Several limitations of multiple exposure technique such as misalignment of input images and ghosting artefacts are discussed. Detailed study of various fusion techniques in multiple exposure technique like fusion in image and radiance domain is performed. It is concluded that true radiance map can be obtained using fusion in radiance domain methods. Methods using fusion in image domain technique are more efficient since they avoid estimation of the camera response function and do not require tone mapping. Then a detailed survey of different artefacts found in HDR image obtained from multiple exposures are performed. The reason of occurrence of these artefacts, the various methods of detection and removal of these artefacts are also discussed. Ghosting artefact is the most severe limitation among other artefacts and it is concluded that SVD method of ghost detection is the best ghost detection method. The several limitations of multiple exposure technique can be avoided by using single shot method of HDR generation. Existing HDR generation methods in image domain can be extended to single shot method by generating multiple exposed images from single shot image using histogram separation or gamma correction. As a suggestion the performance of ghost detection can be improved by combining singular value decomposition and maximum a posterior probability (MAP) parameters.

11. REFERENCES

[1] R. K. Chaurasiya, Prof. K. R. Ramakrishnan, “High Dynamic Range Imaging”, *International Conference on Communication Systems and Network Technologies, 2013.*

[2] Y. Bando, G. Qiu, M. Okuda, S. Daly,T. Aach and Oscar C. AU, “Recent Advances in High Dynamic Range Imaging Technology”, *Proceedings of Sep. 2010 IEEE 17th International Conference on Image Processing, Hong Kong.*

[3] S. Mann , and R. W. Picard "On-being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures", *In Proceedings of IS&T 46th annual conference (May 1995), pp. 422–428.*

[4] T. Mitsunaga and S. Nayar, "Radiometric self calibration", *In Computer Vision and Pattern Recognition, IEEE Computer Society Conference(1999), volume 1, pp 374–380.*

[5] Paul E. Debevec, Jitendra Malik,"Recovering High Dynamic Range Radiance Maps from Photographs", *in: Proceedings of the SIGGRAPH 97 Conference, 1997, pp. 369–378.*

[6] Xiaojun Wu, Zhan Song, Gang Yu,*"*A Novel Multiple Exposure Merging Method for High Dynamic Range Image Generation*" 2010 2nd International Conference on Signal Processing Systems (ICSPS).*

[7] T. Mertens, J. Kautz and F. V. Reeth, "Exposure Fusion," *in IEEE Computer Society, pp. 382-390, 2007.*

[8] Wen-Chung Kao, "High Dynamic Range Imaging by Fusing Multiple Raw Images and Tone Reproduction", *IEEE Transactions on Consumer Electronics, Vol. 54, No. 1, Feb. 2008.*

[9] T. Grosch, " Fast and robust high dynamic range image generation with camera and object movement", *Proceedings of Vision, Modeling and Visualization Conference, 2006, pp. 277–284.*

[10] G. Ward, "Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures", *Journal of Graphics Tools, Volume 8/*2003, *pp. 17-30*.

[11] Y.-S.Heo, K.-M.Lee, S.-U.Lee, Y.Moon, J.Cha, "Ghost-free high dynamic range imaging*", in: Proceedings of the 10th Asian Conference on Computer Vision—ACCV, 2010, pp.486–500.*

[12] C. Aguerrebere, J. Delon, Y. Gousseau and P. Mus´e, "Simultaneous HDR image reconstruction and denoising for dynamic scenes", *2013 IEEE International Conference on Computational Photography (ICCP), 19-21 Apr. 2013, pp. 1 – 11.*

[13] A.S. Deepa and S. Muthumariammal, "An Overview of Ghost Artifact in HDRI", *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), Vol. 2, No. 1, Jan-Feb 2014.*

[14] E. Reinhard, G. Ward, S. Pattanaik and P. Debevec, "High Dynamic Range Imaging-Acquisition, Display, And Image-Based Lighting*", The Morgan Kaufmann Series In Computer Graphics, 2005.*

[15] K. Jacobs, C. Loscos, G. Ward, "Automatic high dynamic range image generation of dynamic environment*", IEEE Computer Graphics and Applications (2008), pp.84–93.*

[16] D.Sidibe, W.Puech, O.Strauss, "Ghost detection and removal in high dynamic range images", *Proceedings of the 17th European Signal Processing Conference, EUSIPCO, 2009, pp.2240–2244.*

[17] Tae-Hong Min, Rae-Hong Park, Soon Keun Chang, "Histogram Based Ghost Removal in High Dynamic Range Images", *Proceedings of the International Conference on Multimedia and Expo ICME, 2009, pp.530–533.*

[18] F. Pece J. Kautz, "Bitmap movement detection: HDR for dynamic scenes", *Proceedings of Visual Media Production (CVMP), 2010, pp. 1–8.*

[19] O. Gallo, N. Gelfand, W.C. Chen, M. Tico, K. Pulli," Artefact-free high dynamic range imaging", *Proceedings of the IEEE International Conference on Computational Photography (ICCP), 2009, pp. 1–7.*

[20] A. Srikantha, D´esir´e Sidib´e and F. M´eriaudeau, "An SVD-Based Approach for Ghost Detection and Removal in High Dynamic Range Images", *21st International Conference on Pattern Recognition (ICPR 2012),Tsukuba, Japan, Nov.11-15, 2012.*

[21] T. Mertens, J. Kautz, F.V. Reeth," Exposure fusion: a simple and practical alternative to high dynamic range photography", *Computer Graphics Forum 28 (2009), pp.161–171.*

[22] P. Burt, T. Adelson, "The Laplacian pyramid as a compact image code", *IEEE Transactions on Communication 31 (1983), pp. 532–540.*

[23] W. Zhang, W.-K. Cham, "Gradient-directed composition of multi- exposure images", *in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2010, pp. 530–536.*

[24] E. A. Khan, A. O. Akyiz and E. Reinhard, "Ghost Removal in High Dynamic Range Images", *2006 IEEE, International Conference on Image Processing.*

[25] M.Pedone, J.Heikkila," Constrain propagation for ghost removal in high dynamic range images", *in: Proceedings of the International Conference on Computer Vision Theory and Applications—VISAPP, 2008, pp.36–41.*

[26] S. Nayar ,T. Mitsunaga," High Dynamic Range Imaging:Spatially Varying Pixel Exposures", *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume1, pp.472–479, Jun 2000.*

[27] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, "Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum", *IEEE Transactions on Image Processing, 99, Mar 2010.*

[28] M. Sch¨oberl, A. Belz, J. Seiler, S. Foessel, and A. Kaup." High dynamic range video by spatially non-regular optical filtering", *Image Processing (ICIP), 2012 19th IEEE International Conference on, pp. 2757–2760.*

[29] J. Seiler and A. Kaup. "Complex-valued frequency selective extrapolation for fast image and video signal extrapolation", *Signal Processing Letters, IEEE, 17(11), pp.949–952, 2010.*

[30] C. Aguerrebere, A.Almansa, Y. Gousseau and J. Delon, "Single Shot High Dynamic Range Imaging Using Piecewise Linear Estimators", *Pablo Mus´e, IIE, Universidad de la Rep´ublica Herrera y Reissig 565, 11300 Uruguay.*

[31] J. Im, J. Jeon, M. Hayes, and J. Paik, "Single image-based ghost-free high dynamic range imaging using local histogram stretching and spatially-adaptive denoising," *IEEE Trans. Consumer Electronics, Nov. 2011, vol. 57, no. 4, pp.1478-1484.*

[32] A. T. Çelebi, R. Duvar and O. Urhan, "Fuzzy Fusion Based High Dynamic Range Imaging using Adaptive Histogram Separation", *IEEE Transactions on Consumer Electronics, Vol. 61, No. 1, Feb. 2015.*

Image Deblurring using L_0 Sparse and Directional Filters

Aparna Ashok *aparnaashok1988@gmail.com*
Department of Electronics and Communication Engineering
Mar Baselios College of Engineering and Technology
Trivandrum, India.

Deepa P. L. *deeparahul2022@gmail.com*
Department of Electronics and Communication Engineering
Mar Baselios College of Engineering and Technology
Trivandrum, India.

Abstract

Blind deconvolution refers to the process of recovering the original image from the blurred image when the blur kernel is unknown. This is an ill-posed problem which requires regularization to solve. The naive MAP approach for solving the blind deconvolution problem was found to favour no-blur solution which in turn led to its failure. It is noted that the success of the further developed successful MAP based deblurring methods is due to the intermediate steps in between, which produces an image containing only salient image structures. This intermediate image is essentially called the unnatural representation of the image. L_0 sparse expression can be used as the regularization term to effectively develop an efficient optimization method that generates unnatural representation of an image for kernel estimation. Further, the standard deblurring methods are affected by the presence of image noise. A directional filter incorporated as an initial step to the deblurring process makes the method efficient to be used for blurry as well as noisy images. Directional filtering along with L_0 sparse regularization gives a good kernel estimate in spite of the image being noisy. In the final image restoration step, a method to give a better result with lesser artifacts is incorporated. Experimental results show that the proposed method recovers a good quality image from a blurry and noisy image.

Keywords: Motion Blur, Blind Deconvolution, Deblurring, L_0 Sparsity, Directional Filtering, Image Restoration.

1. INTRODUCTION

Motion blur caused by camera shake is the most common artifact and one of the predominant source of degradation in digital photography. Photos taken in low light without flash requires higher exposure times and thereby the photos get affected by motion blur. Increasing the light sensitivity of the camera using a higher ISO setting may help in reducing the exposure time, but there is a trade off with noise levels. The lower exposure time comes at the cost of higher noise levels. Even then, the exposure time still remains high for handheld photography and camera shake is likely to happen without the use of a tripod. As a result, the photos end up being blurry and noisy. Recovering an unblurred, sharp image from a single motion blurred image is one of the major research areas in digital photography. This problem can be further divided into blind and non-blind cases. If the blur kernel that has degraded the image is known, then the only problem is the estimation of the unknown latent image by the deconvolution process. This is referred non-blind deblurring. If there is no information available about the kernel that has degraded the image, the original image has to be estimated from the blurred image using the mathematical model of the blurring process. The kernel and the image has to be estimated simultaneously and iteratively in this process. This is a much more ill-posed task and is referred to as blind deblurring. A large number of techniques have been proposed in the recent times to address the problem of blind image deblurring, by jointly estimating the latent deblurred image

while recovering the blur kernel which has degraded the image. Most of these methods assume an ideal condition with little image noise and demonstrates a fair level of success in such photos. However a significant amount of noise affect the performance of the existing standard blind deblurring algorithms. The presence of noise causes high frequency perturbations of the image values and this is not taken into consideration by the standard blind deblurring methods. In the blind deblurring process, the image noise manifests itself as noise in the estimated kernel and is further amplified by the deconvolution process. This produces artifacts in the deblurring result.

Single image blind deconvolution was extensively studied in the recent years and the field has reached considerable success with many milestones. The basic concept of blind deconvolution and different existing types of blind deconvolution algorithms are explained in detail by D. Kundur and D. Hatzinakos in their paper [1]. The classifications of blind deconvolution techniques are explained and its merits and demerits are discussed. MAP method is the most commonly used among the blind deblurring techniques because it does not require any apriori information about the PSF and also, it does not constrain the PSF to be of any specific parametric form. Naive MAP approach was seen to fail on natural images as it tends to favor no blur solution. Levin et al.[2] analyzed the source of MAP failure and demonstrated that marginalizing the process over the image x and MAP estimation of the kernel k alone proves successful and recovers an accurate kernel. As noted by Levin et al.[2] and Fergus et al.[3], blurry images have a lower cost compared to sharp images in MAP approach and as a result blurry images are favored. So, a number of more complex methods have been proposed that include marginalization over all possible images[2,3], dynamic adaptation of the cost function used[4],determining edge positions by the usage of shock filtering[5] and reweighting of image edges during optimization[6]. Many papers emphasized the usage of sparse prior in derivative domain to favor sharp images. But expected result has not been yielded by the direct application of this principle, as it required additional processes like marginalization across all possible images as demonstrated by Fergus et al.[3] spatially varying terms or solvers that vary optimization energy over time as shown by Shan et al.[4].Krishnan et al.[7] used l1/l2 norm as the sparsity measure which acts as a scale invariant regularizer. Xu et al.[8] introduced L_0 sparsity as regularization term ,which relies on the intermediate representation of image for its success. All of the mentioned deblurring methods generally work well when the image is noise free, but their performance deteriorates in the presence of noise, as the noise level increases[3,9,10]. It has been noticed that standard denoising methods like Wiener filtering, NLM filtering have negative effect on the kernel estimation process[11,12]. Zhong et al.[13] proposed an approach based on directional filtering followed by Radon transform for accurate kernel estimation in the presence of noise.

This paper proposes a new blind deconvolution algorithm for the deblurring of blurry and noisy images using L_0 sparse prior, which is equipped with appropriate noise handling using directional filters, and a method incorporated in final image restoration step to obtain a good quality latent image with lesser artifacts. It has been noted that the prior MAP based approaches that are successful can be roughly classified into two types i.e. those with explicit edge prediction steps like using a shock filter [5,6,14,15,16,17,18,19] and those which include implicit regularization process[7,8]. The common factor in these two is that they both include an intermediate step, which produces an image which contains only the salient structures while suppressing others. This intermediate image which contains only step like or high contrast structures is called the unnatural representation of the image. These image maps are the key for making motion deblurring process accomplishable in different successful MAP based methods. The L_0 sparse prior acts as a regularization term and enables accurate kernel estimation through the unnatural representation of the image. Directional filters are seen to remove noise in an image effectively, without affecting the blur kernel estimation process in a significant manner. Consequently, directional filters can be successfully incorporated in the deblurring technique for noise handling, without the problem of steering kernel estimation along the wrong direction. The incorporation of an additional process for reducing the artifacts while enhancing finer details ensures a good quality latent image.

2. PROPOSED WORK

Directional lowpass filters can be applied to an image to reduce its noise level without tampering with the process of kernel estimation on a significant level. The application of directional lowpass filter h_θ to an image I can be explained theoretically by the equation

$$I(p) * h_\theta = \frac{1}{r}\int_{-\infty}^{\infty} w(t) I(p + t u_\theta)\, dt \quad (1)$$

where p denotes the location of each pixel, r is the normalization factor given as $r = \int_{-\infty}^{\infty} w(t)dt$, t is the spatial distance from each pixel to p and u_θ is the unit vector in the direction of application of the filter, θ. We choose the directional filter to have a Gaussian profile and this is determined by the factor w(t) which can be given by the expression $w(t) = e^{-t^2/2\sigma^2}$ where σ is the factor which controls the strength of the Gaussian directional filtering process[20].

The image, after the directional filtering process, is to be now advanced to the kernel estimation step for the blind deblurring process. A regularization term which consists of a family of loss functions, that approximates L_0 cost, is incorporated into the objective function which has to be optimized. L_0 approximation enables a high sparsity pursuit regularization, which also leads to consistent energy minimization and fast convergence during the optimization process. Since only the salient structures of the image are retained in the unnatural representation, the method is faster than other implicit regularization methods. This family of loss functions that approximates L_0 cost implements graduate non-convexity into the optimization process and the significant edges guides the kernel estimation process in the right direction, thus quickly improving the estimation process in only a few iterations.

The image that has been obtained at the end of kernel estimation step is not our required latent image due to the lack of details, as it contains just high contrast edges. A final image restoration step has to be carried out for obtaining the latent image. An existing non-blind deconvolution can be used for this. From the study of the existing non-blind deblurring methods, it can be seen that non-blind deblurring using Laplacian prior shows a great preservation of finer details[21]. But at the same time, the result is seen to have considerably significant artifacts in the case of Laplacian prior. On the contrary, a restoration step using L_0 prior produces image with very less artifacts though the finer details are fewer[22]. Therefore, instead of simply performing non-blind deconvolution with the obtained kernel using hyper Laplacian prior, a simple method is proposed to get a latent image with finer details and fewer artifacts. First, latent image I_1 is estimated using the non-blind deconvolution method with hyper-Laplacian prior. After this, latent image I_0 is estimated using the L_0 prior scheme introduced in our paper. The difference map of these two estimated images is computed, followed by bilateral filtering to remove artifacts in the difference map obtained. Subtracting the result of bilateral filtering from I1 gives the desired final image output which contains finer details with fewer artifacts.

3. MATHEMATICAL FRAMEWORK

We denote the latent image by x, blurred and noisy image by y and the blur kernel by k. The blurring process can be represented generally as

$$y = k * x + \eta \quad (2)$$

where η represents the image noise. The blurred and noisy image y has to be first of all denoised, by the application of directional low pass filters with Gaussian profile in the desired orientations. The strength of the filtering process by directional filter h_θ along each orientation θ is denoted by its σ value. This value can be decided accordingly based on the image at hand and the noise infected along that particular direction. After deciding on the different orientations for the application of directional filters and the strength of Gaussian profile filtering along each direction,

denoising is carried out. This gives a noise-free image without affecting the blur of the image considerably in a way that will interfere with accurate kernel estimation.

The framework of the proposed deblurring technique includes a loss function $\Phi_0(.)$ that approximates L_0 cost into the objective function during optimization process. The loss function for an image z can be defined as

$$\Phi_0(\partial_* \mathrm{z}) = \sum_{\mathrm{i}} \Phi(\partial_* z_i) \tag{3}$$

where,

$$\Phi(\partial_* z_i) = \begin{cases} \frac{1}{\varepsilon^2} |\partial_* z_i|^2 & , if\ |\partial_* z_i| \leq \varepsilon \\ 1 & otherwise \end{cases} \tag{4}$$

and $* \epsilon \{ h, v\}$ denoting the horizontal and vertical directions respectively for each pixel i ,for taking gradient of the image. This function is continuous when $|\partial_* z_i| \leq \varepsilon$,which is a necessary condition for being a loss function mathematically. This loss function is a very high sparsity pursuit one that approximates sparse L_0 function very closely.

The final objective function is obtained by incorporating the loss function $\Phi_0(.)$ in our method as a regularization term during the optimization process, which in turn seeks an intermediate sparse representation of the image containing only salient edges, which guides the kernel estimation in the right direction. The objective function for kernel estimation can be given as

$$\min_{(\tilde{x}, k)} \{\|k * \tilde{x} - y\|^2 + \lambda \sum_{*\in\{h,v\}} \Phi_0(\partial_* \tilde{x}) + \gamma \|k\|^2\} \tag{5}$$

where $\tilde{x}$ is the unnatural representation of the image during the process of optimization of the objective function and λ, γ are regularization weights .The first term of the objective function is the data fidelity term which enforces blur model constraint, second term is the loss function approximating sparse L_0 and the last term helps in reducing the kernel noise. It can be seen that the new regularization term is the key factor in guiding the kernel estimation process quickly and accurately in the right direction. If we consider the case of explicit edge prediction methods like employing a shock filter, it cannot be incorporated into the objective function during the optimization process. But the advantage in the case of using the L_0sparse regularization term is that, it can be efficiently incorporated into the objective function during the optimization process, which ensures the fact that the intermediate representation of the image contains only necessary strong edges that satisfy the constraints, regardless of the blur kernel.

The objective function can be solved by alternatively computing the intermediate image value and the kernel value in each iteration[23]. The computation process for $(t+1)^{th}$ iteration of optimization process can be given by Eq. 6 and Eq. 7.

$$\tilde{x}^{t+1} = \underset{\tilde{x}}{argmin} \left\{ \|k^t * \tilde{x} - y\|^2 + \lambda \sum_{*\in\{h,v\}} \Phi_0(\partial_* \tilde{x}) \right\} \tag{6}$$

$$k^{t+1} = \underset{k}{argmin} \{\|\tilde{x}^{t+1} * k - y\|^2 + \gamma \|k\|^2\} \tag{7}$$

Equation 4 for the loss function can be rewritten as follows with ε as a parameter, for the ease of the optimization process.

$$\Phi(\partial_* \mathrm{z_i}, \varepsilon) = \min_{\mathrm{l_{*i}}} \left\{ |\mathrm{l_* i}|^0 + \frac{1}{\varepsilon^2} (\partial_* \mathrm{z_i} - \mathrm{l_{*i}})^2 \right\} \tag{8}$$

where, $* \in \{h, v\}$ and

$$l_{*i} = \begin{cases} 0 & , |\partial_* \tilde{x}_i| \leq \varepsilon \\ \partial_* \tilde{x}_i & , \text{otherwise} \end{cases} \quad (9)$$

The above function can be proved to be mathematically equivalent to the previous one. A family of loss functions are obtained based on this equation by setting the value of ε differently. Figure 1 shows the family of loss functions for different values of ε starting from 1 to 1/8 and the plot approaches L_0 further as the value of ε keeps on decreasing.

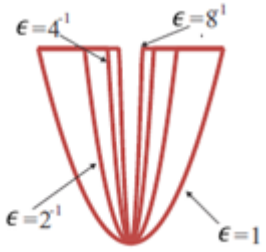


FIGURE 1: Plots of the loss function approximating L_0 for different values of ε.

Equation 6 for the optimization process computing $\tilde{x}$ can also be rewritten accordingly based on Eq.8 and Eq.9 .

$$\min_{\tilde{x}, l} \left\{ \frac{1}{\lambda} \|k * \tilde{x} - y\|^2 + \sum_{*\epsilon\{h,v\}} \sum_i \{ |l_{*i}|^0 + \frac{1}{\varepsilon^2} (\partial_* \tilde{x}_i - l_{*i})^2 \} \right\} \quad (10)$$

We alternate the process of computing intermediate image $\tilde{x}$ and updating the value of l_{*i} in the iterations for each of the loss function obtained for different values of ε. The whole optimization process can be speeded up by transforming the computation process into FFT domain. Using FFTs with the quadratic form enables fast kernel estimation process. The solution in FFT domain can be expressed by Eq.11 and Eq. 12.

$$\tilde{x}^{t+1} = F^{-1} \left\{ \frac{\overline{F(k^t)}.F(y) + \frac{\lambda}{\varepsilon^2} \left(\overline{F(\partial_h)}.F(l_h) + \overline{F(\partial_v)}.F(l_v) \right)}{\overline{F(k^t)}.F(k^t) + \frac{\lambda}{\varepsilon^2} (|F(\partial_h)|^2 + |F(\partial_v)|^2)} \right\} \quad (11)$$

$$k^{t+1} = F^{-1} \left\{ \frac{F(\tilde{x}^{t+1}).F(y)}{|F(\tilde{x}^{t+1})|^2 + \gamma} \right\} \quad (12)$$

where, F(.) and F(.) are the FFT operator and its conjugate respectively, F^{-1} is the inverse FFT operation whereas and are vectors concatenating and values for each pixel. F_D^2 denotes $|F(\partial h)|^2 + |F(\partial v)|^2$ .Multiplication and division operations are performed in an element wise manner on the complex vectors.

During the implementation process, we use a family of 4 loss functions with ε ϵ {1,1/2,1/4,1/8}. We start from ε = 1 and then proceed to the other values as shown in Figure 1. The number of iterations for different loss functions is set to be inversely proportional to the corresponding value of ε. This is because of the fact that large ε values cause the loss function to be more convex like

and hence makes it more easy to optimize. So it requires only a few iterations. The results obtained here is taken as an initialization for further refinement in loss functions with smaller ε values, as they are of more concave nature and difficult to optimize. Also, the blur kernel estimation process is carried out in a pyramid-like fashion, by convention. Kernels are estimated in a coarse to fine manner in an image pyramid. The estimate obtained in one image pyramid level is taken as the initialization of the next one. The optimization process in each iteration t+1 in its finest level is as explained by the equations Eq.11 and Eq.12 .Computation is similar in the coarser level for different iterations.

The algorithm for kernel estimation process in one image level can be given as follows

Input: Blurry and noisy image y

Output: Blur kernel k, deblurred image $\tilde{x}$

1 Apply N directional filters to the input image y, where each filter has a direction given by the expression $\theta=(i.\pi/N)$,i=1,2,...N and N is the number of directional filters. Choose the σ value for each direction accordingly depending on the image.

2 Initialize k from the kernel estimate of coarser scale

3 for t= 1:5

4 //update image

5 $\varepsilon \longrightarrow 1$

6 for i=1:4

7 for j=1: ε^{-1}

8 solve for l using equation (9)

9 solve for $\tilde{x}^{t+1}$ using equation (11)

10 end

11 $\varepsilon \longrightarrow \varepsilon/2$

12 end

13 //update kernel

14 solve for k^{t+1} using equation (12)

15 end

Desired kernel estimate is obtained at the end of algorithm execution for each image level, in a coarse to fine manner. Five iterations of alternative image and kernel estimation are required generally at each level. But the computed image $\tilde{x}$ at the end of the algorithm is not the final latent image because it contains only salient edges and lacks details. A non-blind deconvolution using hyper Laplacian prior can be used for latent image restoration in general case. But here, we go for a better method to remove artifacts and obtain a much better quality image. We take advantage of the fact that non-blind deconvolution with hyper Laplacian prior produces an image result with very fine details but a considerable amount of artifacts, whereas restoration using L_0 prior and the estimated kernel produces a result with very less artifacts though the result may not contain much finer details. A difference map between these two results is computed and it is

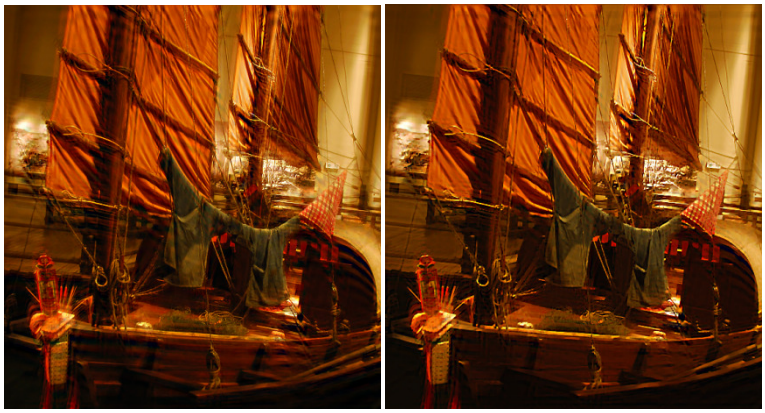
subjected to bilateral filtering. Further subtraction of the obtained filtered result from the result of non-blind deconvolution with hyper Laplacian prior produces a latent image with finer details and practically very less artifacts .It can be seen that this results in a much better quality result when compared to the result that would have been produced if we had gone for a normal non-blind deconvolution using estimated kernel obtained in the previous step. The latent image restoration step can also be accelerated by the use of FFT.

4. EXPERIMENTAL RESULTS

We experiment with data on different natural and synthetic images blurred by a set of 8 different blur kernels. We implemented the proposed method in Matlab on an Intel Core i7 CPU. The value of parameters are set to λ=2e-3 and γ = 40 for all experiments. Initially, we test the algorithm for images affected by just blurring by excluding the directional filtering part from our experimental process i.e. we estimate the kernel by L0 prior scheme and then use the method explained in the paper for latent image restoration along with artifacts removal. After that, we take blurry and noisy images and carry out the entire experimental process starting with directional filtering. Gaussian random noise with a sigma value of 7 has been used in our experiments. We applied directional filters along 24 regularly sampled directions i.e. one sample every 15^0 .We have taken σ value of the filter to be 1 or 2 depending on the image and we have tested with the same σ value in all directions ,though this can be varied for different directions if the image demands so. It can be seen that the final images in the case without noise are of excellent visual quality and shows a high improvement in PSNR values. Further, in the case of blurry and noisy images, the process produces a good visual quality image in spite of the presence of noise, unlike standard deblurring algorithms that produce deteriorated results in the presence of noise. Comparative studies have been perfomed with standard existing methods of deblurring and denoising. The results obtained using directional filtering for noise removal process can be found superior to the one obtained using Wiener filter, which has been the most commonly used filtering method in image processing field in the recent times.

First of all, experimental results obtained on simulation with a blurry image is shown to demonstrate the process. Fig. 2 shows the comparison of results between our method and the method proposed by Xu et al. [8] which has a standard non-blind deconvolution step using hyper-Laplacian priors [21].The intermediate step which is the unnatural representation of the image, the estimated kernel, latent image restored in the case of non-blind deconvolution applied directly[8],latent image restored by our method including a simple artifact removing step (and excluding directional filtering step) are shown in the figure. It can be seen that latent image restored by our method is of clearly of much better visual quality and even has a higher PSNR value compared to the results produced by Xu et al.[8] Some more results of our method (excluding directional filtering) on blurry images are shown in Fig. 3 which shows the blurry image, restored latent image and estimated kernel. The method clearly restores an excellent visual quality image.

(a)blurry input. (b)unnatural representation and estimated kernel.



(c)Restored latent image for Xu et al.[8]. PSNR= 28.52

(d) our restored latent image. PSNR=29.34

FIGURE 2: Demonstration of intermediate representation, estimated kernel and recovered latent image in our process with blurry image as input. Restored image for Xu et al. is provided for comparison process.

(a) blurry input Image 1. (b) our restored latent image and estimated kernel.

(c) blurry input Image 2. (d) our restored latent image and estimated kernel.

(e) blurry input Image 3. (f)our restored latent image and estimated kernel.

FIGURE 3: Experimental Results for Blurry Images.

Now, we go for the experimental results of images affected by both blur and noise, obtained by simulation based on the parameters and conditions defined in the beginning of this section. Wiener filtering has been a classic method used commonly in image denoising field [24]. The results by simulation of our method using directional filters for noise handling are given in Fig. 4, along with the results using Wiener filter proposed by Jin et al. [24], for comparison process. It can be seen that directional filtering produces much better visual quality images compared to Wiener filtering in all cases. It produces images with quality that is very comparable to that of the results obtained without noise, inspite of the presence of noise. Comparison of estimated kernels with ground truth kernels for two of the test images is shown in Fig. 5. It can be seen that the estimated kernel is indeed very much closer to the ground truth. Table 1 shows the PSNR values of the experimental results obtained for blurry images as well as blurry & noisy images equipped with noise handling. The last two columns of the table shows the comparative performance evaluation of results obtained by our method and those obtained by incorporating denoising proposed by Jin et al. [24] into the L_0 deblurring process with a standard non-blind deconvolution proposed by Xu et al. [8].Our method is seen to produce good visual quality results with high

improvement of PSNR values even in the presence of noise. Most of the existing methods fail to produce such high PSNR values without the use of explicit edge prediction steps like shock filter, in the presence of noise.

(a)restored latent image using Wiener filtering for blurry and noisy Image 1.

(b)restored latent image using directional filtering for blurry and noisy Image 1.

(c)restored latent image using Wiener filtering for blurry and noisy Image 2.

(d)restored latent image using directional filtering for blurry and noisy Image 2.

(e)restored latent image using Wiener filtering for blurry and noisy Image 3.

(f)restored latent image using directional filtering for blurry and noisy Image 3.

FIGURE 4: Experimental results of our method and Jin et al.[24] for blurry and noisy images - a comparison.

(a)Estimated kernels.

(b)Ground truths.

FIGURE 5: Comparison of estimated kernels with ground truth.

Image	Input blurry image PSNR(dB)	L0 deblurred output for blurry image PSNR (dB)	L0 deblurred output for blurry & noisy input after noise handling by Wiener filter(dB)	L0 deblurred output for blurry & noisy input after noise handling by directional filter(dB)
Image 1	20.33	30.20	28.12	29.25
Image 2	20.255	31.47	29.07	30.91
Image 3	27.33	32.81	30.51	32.61

TABLE 1: PSNR values of experimental results of our method for blurry images as well as for blurry & noisy images.Results of our method using directional filtering is compared to the results using Wiener filtering[24].

5. CONCLUSION

In this paper, we propose a new single image deblurring technique that is robust to noise unlike standard deblurring methods. Our method uses directional filtering for noise handling which does not affect the blur information and the kernel estimation process significantly in an adverse manner. A loss function approximating L0 cost is used as a prior for regularization in the objective function, which produces an unnatural sparse representation that benefits kernel estimation and optimization processes. A simple method is introduced in the final latent image restoration step to obtain a better quality image with much finer details and lesser artifacts. The method is very effective in handling blurry & noisy images compared to the existing deblurring techniques which deteriorates in performance when noise is present. As a future scope, the method can be extended to handle non-uniform blur in presence of noise.

6. REFERENCES

[1] D. Kundur and D. Hatzinakos. "Blind image deconvolution", IEEE Signal Processing Magazine, 1996.

[2] A. Levin, Y. Weiss, F. Durand and W. T. Freeman. "Understanding and evaluating blind deconvolution algorithms." In CVPR,2009, pp.1964–1971.

[3] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis and W. T. Freeman." Removing camera shake from a single photograph." ACM Transactions on Graphics (SIGGRAPH ASIA), vol. 28 no.5, p. article no. 787–794, 2006.

[4] Q. Shan, J. Jia, and A. Agarwala. "High-quality motion deblurring from a single image."ACM Transactions on Graphics (SIGGRAPH ASIA),vol. 27(3),2008.

[5] J. H. Money and S. H. Kang. "Total variation minimizing blind deconvolution with shock filter reference." Image and Vision Computing, 26(2):302–314, 2008.

[6] S. Cho and S. Lee. "Fast motion deblurring." ACM Transactions on Graphics (SIGGRAPH ASIA), vol. 28, no. 5, p. article no. 145, Dec. 2009.

[7] D. Krishnan, T. Tay, and R. Fergus. "Blind deconvolution using a normalized sparsity measure." In CVPR, 2011,pp. 233–240.

[8] L. Xu, S. Zheng and J. Jia. "Unnatural l0 sparse representation for natural image deblurring." In CVPR, 2013,pp. 1107–1114.

[9] R. Koehler, M. Hirsch, S. Harmeling, B. Mohler and B.Scholkopf. "Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database." In ECCV, 2012, pp. 27–40.

[10] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. "Efficient marginal likelihood optimization in blind deconvolution." In CVPR2011,pp. 2657–2664.

[11] A. Buades, B. coll and J. Morel." A non-local algorithm for image denoising." CVPR, 2005,pp. 60-65.

[12] L. Zhang, A. Deshpande and X. Chen. "Denoising vs. deblurring: Hdr imaging techniques using moving cameras." In CVPR,2010 pp. 522-529.

[13] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang. "Handling noise in single image deblurring using directional filters",In CVPR,2013, pp. 612–619.

[14] L. Xu and J. Jia." Two-phase kernel estimation for robust motion deblurring", ECCV, 2010, pp. 157–170.

[15] L. Xu, Q. Yan, Y. Xia and J. Jia. "Structure extraction from texture via relative total variation." ACM Transactions on Graphics (SIGGRAPH ASIA), vol. 31(6), 2012.

[16] J. Jia. "Single image motion deblurring using transparency." In CVPR, 2007, pp. 1-8.

[17] N. Joshi, R. Szeliski and D. J. Kriegman."PSF estimation using sharp edge prediction." In CVPR, 2008, pp. 1-8.

[18] T. S. Cho, S. Paris, B. K. P. Horn and W. T. Freeman. "Blur kernel estimation using the radon transform." In CVPR, 2011,pp. 241–248.

[19] M. Hirsch, C. J. Schuler, S. Harmeling and B. Scholkopf. "Fast removal of non-uniform camera shake." In ICCV, 2011,pp. 463–470.

[20] A. Buades, C. B. and J.-M. Morel. "The stair casing effect in neighborhood filters and its solution." IEEE Transaction on Image Processing,vol.15,pp. 1499-1505, 2006.

[21] D. Krishnan and R. Fergus."Fast image deconvolution using hyper-laplacian priors." In NIPS, pp. 1033-1041,2009.

[22] Z.Hu, M.H.Yang, J.Pan and Z.Su. "Deblurring text images via L0 regularized intensity and gradient prior." In CVPR,2014, pp. 2901-2908.

[23] L. Xu, C. Lu, Y. Xu and J. Jia. "Image smoothing via l_0 gradient minimization."ACM Transactions on Graphics (SIGGRAPH ASIA),vol. 30(6), 2011.

[24] F. Jin, P. Fieguth, L. Winger and E. Jernigan. "Adaptive Wiener filtering of noisy images and image sequences." International Conference for Image Processing, IEEE, 2003, pp. 349-352.

A Novel Approach To Detection and Evaluation of Resampled Tampered Images

Amrit Hanuman *amrit2025@hotmail.com*
Department of Electrical and Computer Engineering
The University of the West Indies
St. Augustine, Trinidad

Azim Abdool *azim.abdool@sta.uwi.edu*
Department of Electrical and Computer Engineering
The University of the West Indies
St. Augustine, Trinidad

Akash Pooransingh *akash.pooransingh@sta.uwi.edu*
Department of Electrical and Computer Engineering
The University of the West Indies
St. Augustine, Trinidad

Aniel Maharajh *admaharajh@outlook.com*
Department of Electrical and Computer Engineering
The University of the West Indies
St. Augustine, Trinidad

Abstract

Most digital forgeries use an interpolation function, affecting the underlying statistical distribution of the image pixel values, that when detected, can be used as evidence of tampering. This paper provides a comparison of interpolation techniques, similar to Lehmann [1], using analyses of the Fourier transform of the image signal, and a quantitative assessment of the interpolation quality after applying selected interpolation functions, alongside an appraisal of computational performance using runtime measurements. A novel algorithm is proposed for detecting locally tampered regions, taking the averaged discrete Fourier transform of the zero-crossing of the second difference of the resampled signal (ADZ). The algorithm was contrasted using precision, recall and specificity metrics against those found in the literature, with comparable results. The interpolation comparison results were similar to that of [1]. The results of the detection algorithm showed that it performed well for determining authentic images, and better than previously proposed algorithms for determining tampered regions.

Keywords: Image Forgery, Resampling, Interpolation, Passive Detection.

1. INTRODUCTION

The availability of affordable high resolution digital cameras and powerful image editing software has increased the ease of digital image tampering [2]. Digital forgeries can be mistaken for authentic images as they seldom leave visual clues of the alteration. People are confronted with digital forgeries recurrently through mainstream media outlets [3]. These forgeries are created to stir controversy, invoke emotion or sell products. In resampled images, interpolating functions are directly responsible for the output image quality when altering them. Image editing tools are continuously being developed and improved to hide any sign of tampering. Consequently, the validity of images used as evidence in the judicial system, news media and scientific journals can be questioned [2]. It is essential to verify the integrity of images to avoid deception.

There are three main classifications of digital forgeries: copy-move, splicing and retouching [4]. In copy-move forgery (Figure 1), part of a source image is copied and pasted at another location in the same image. Image splicing (Figure 2) creates a new image using composites of two or more images, while retouching (Figure 3) deals with changes to image features such as color, textures or blurs. Each forgery type can be detected using different methods.

Splicing can be considered the most malicious of the forgery types where resampling is used to change a foreign image's dimension or angle such that it looks natural on the host image.

FIGURE 1: An Example of Copy-Move Forgery [4].

FIGURE 2: An Example of Splicing Forgery [4].

FIGURE 3: An Example of Retouching Forgery [4].

Resampling is required since the new image's coordinate points will not always line up with the old points. Resampling is defined as “the process of transforming a discrete image which is

defined at one set of coordinate locations to a new set of coordinate points" [5]. There are two steps to resampling (Figure 4): interpolation and sampling.

Interpolation uses known data to estimate values at unknown points. It is the process of fitting a continuous function to the discrete points in the digital image [5]. The discrete data is convolved with a continuous interpolating function to create a continuous signal. This continuous signal is then sampled to obtain the values at the new coordinate points [5].

Image interpolation has many applications ranging from simple viewing and editing online images, to medical image processing [6, 1, 7, 8]. In image editing software, the three commonly used interpolating functions are nearest-neighbor, bilinear and bicubic interpolation. Additionally, some include the lanczos3 function (see Figure 5). These are categorized as non-adaptive algorithms since all pixels are treated equally. Adaptive algorithms rely on the intrinsic image features and contents such as sharp edges or smooth texture [6]. The choice of interpolation function produces different results, and successive interpolation diminishes image quality except for rotations that are multiples of 90° or shifts over integer pixel values [1].

Interpolation affects the underlying statistics of the image by introducing specific correlations of pixel values into the image [2]. The detected correlation can be used as evidence of tampering. Farid [3] defined pixel-based forensic tools as those that detect statistical anomalies introduced at the pixel level. Detection methods are classified into two types: active and passive. Active methods require prior knowledge of the original image whereas passive methods do not [4].

An example of a current active detection method is digital watermarking. A digital code (a watermark) is embedded into the image at the time of recording through specially equipped devices. If the image is tampered with, this code will be altered. The image is authentic if the extracted code matches the inserted code, after transmission. Passive or blind methods are favored over active methods since the original image is often unavailable. These are the methods that detect the statistical anomalies in the image [9].

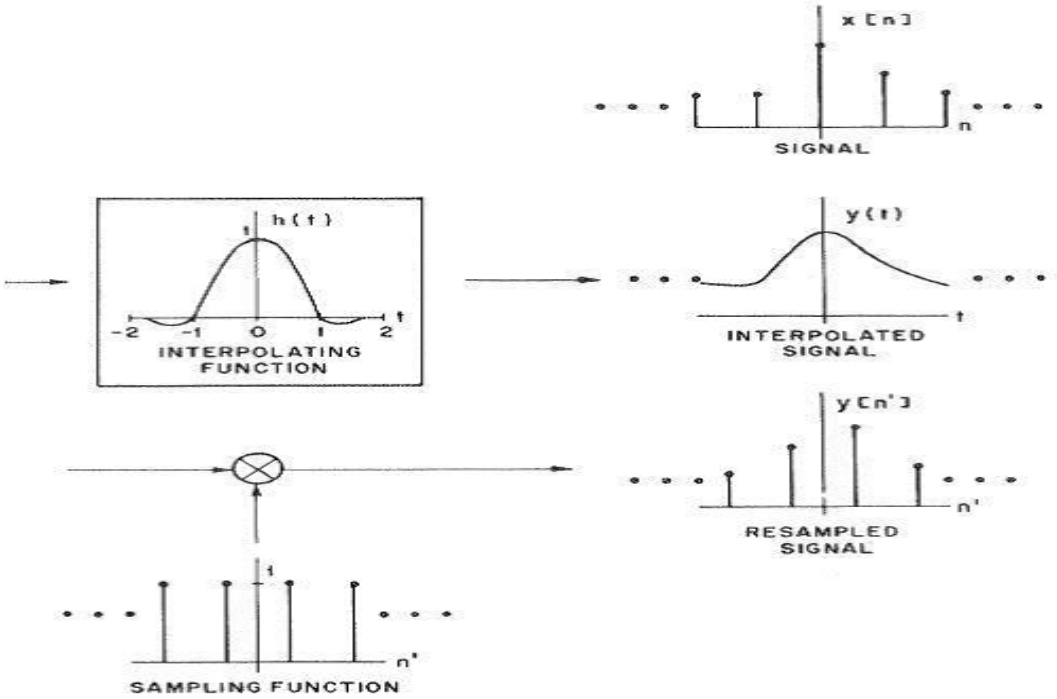


FIGURE 4: The Resampling Process: A discrete signal is fitted with a continuous function (interpolated) to create a continuous signal that is sampled to get the new discrete values. [5].

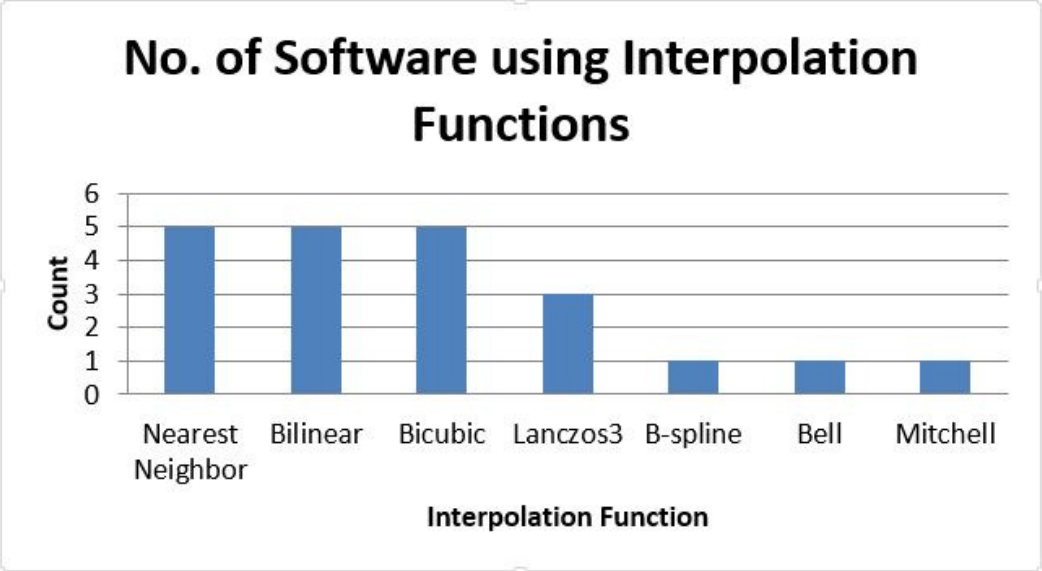


FIGURE 5: Interpolation Functions Found in 5 Popular Image Editing Software.

This paper presents a brief review of common interpolation functions. A novel algorithm is proposed for automatic detection of tampered images. This paper does not include estimation of resampling factors, but presents an evaluation method to determine the detection efficiency. The organization of this paper is as follows. Section 2 reviews previous works done on comparison and detection of interpolation. The proposed detection algorithm is described in Section 3. The details of the implementation are provided in Section 4, with results in Section 5. Section 6 discusses the results and future works, followed by the concluding remarks in Section 7.

2. RELATED WORK

2.1 Comparison of Interpolation Techniques

A comparison of different interpolation methods, examining the frequency response of the interpolating functions, was done in [5]. It was noted by Lehmann, Gonner and Spitzer [1] that the appearance of the image after resampling should also be taken into account. They presented a comprehensive survey of interpolation methods, investigating the Fourier analysis, qualitative and quantitative error, computational complexity and run time measurements. Their comparison was extended to other interpolation methods in [10]. Following is a brief review of the different interpolation techniques.

2.1.1 Ideal Interpolation

A continuous band-limited signal can be sufficiently reconstructed from samples if it is sampled at the Nyquist frequency or higher. [5]. However, sampling produces infinite replicas of the continuous spectrum in the frequency domain. To recover the band-limited signal, only one of these replicas is needed. The interpolating function used should be an ideal low-pass filter. Multiplying the continuous spectrum with the rectangle function in the frequency domain is equivalent to convolving the discrete data with a sinc function in the spatial domain [1]. The function in the spatial domain is given by Equation (1).

$$h_{ideal}(x) = \frac{\sin(\pi x)}{\pi x} = sinc(x) \tag{1}$$

2.1.2 Nearest Neighbor Interpolation

Nearest neighbor is the simplest interpolation technique, requiring the least computational processing time. Only one pixel is considered for each new coordinate point. The value of the

pixel at the new coordinate point of the output image is equal to that of the pixel nearest its corresponding point in the source image. The interpolation is done by convolving the sampled image with a rectangle function in the spatial domain described by Equation (2) [1]. In the frequency domain, this is a sinc function, a poor low pass filter. The resultant image is pixelated or blocky [1].

$$h_{nearest(x)} = \begin{cases} 1 & 0 \leq |x| < 0.5 \\ 0 & elsewhere \end{cases} \tag{2}$$

2.1.3 Bilinear Interpolation

To estimate the value at the new point, bilinear interpolation considers the closest 2x2 neighborhood of known pixels. This value is a weighted average, where the points are weighted based on the distances to the diagonally opposite point [6, 1]. It is a combination of linear interpolation along the horizontal and vertical axes. Bilinear interpolation is a convolution of the discrete data with a triangle function expressed in Equation (3). There is attenuation of high frequency components and the aliasing of data beyond the cutoff point into the low frequencies. Bilinear interpolation produces a smoother output image [5].

$$h_{bilinear}(x) = \begin{cases} 1 - |x| & 0 \leq |x| \leq 1 \\ 0 & elsewhere \end{cases} \tag{3}$$

2.1.4 Bicubic Interpolation

Bicubic interpolation considers the closest 4x4 neighborhood of known pixels for estimation. Closer pixels are given higher weightings. It assumes that the image follows a polynomial in this neighborhood. The interpolation function is described in Equation (4). Lehmann, Gonner and Spitzer [1] suggest using an optimal value of a = 0.5. Bicubic interpolation performs better than bilinear interpolation [7].

$$h_{bicubic}(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & 0 \leq |x| \leq 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & 1 \leq |x| \leq 2 \\ 0 & elsewhere \end{cases} \tag{4}$$

2.1.5 Lanczos3 Interpolation

Lanczos3 is a windowed sinc function equivalent to the multiplication of the sinc function by a rectangular function in the spatial domain [1]. It is described by Equation (5) [6]. The truncation introduces ringing artifacts. These artifacts are reduced if the edges of the window coincide with a pair of the sinc function's zero-crossings [7].

$$h_{lanczos3}(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \times \frac{\sin(\pi \frac{x}{3})}{\pi \frac{x}{3}} & |x| < 3 \\ 0 & elsewhere \end{cases} \tag{5}$$

2.2 Detection of Interpolation Techniques

Two of the more common detection methods used were the Expectation-Maximization method [2] and the properties of the second difference statistics method [11, 12]. Popescu and Farid [2] found that specific samples in a resampled sequence can be written as a linear combination of their neighboring samples. The Expectation-Maximization method was applied to determine the probability that the samples were correlated and find the specific form of the correlation. This was because the probabilities were arranged in periodic patterns, shown in its Fourier transform, which depended on the resampling rate.

Gallagher [11] introduced a novel algorithm for interpolation detection and estimation in JPEG images. The second derivative signal of interpolated images was proven to have a periodicity related to the resampling factor. The algorithm worked by detecting patterns in the pixel-wise difference. First, the second derivative of each row of the image was computed using Equation (6) followed by the average across all rows.

$$Sp(i,j) = 2p(i,j) - p(i,j+1) - p(i,j-1) \tag{6}$$

The DFT magnitudes of this trace were then plotted. Distinct peaks would be displayed if the image was interpolated. If no peaks occurred, the image was considered authentic. Where peaks occurred, the resampling factors could then be estimated using Equation (7) if the peak occurred at a normalized frequency greater than or equal to 0.5, or Equation (8) for normalized frequency less than 0.5.

$$\text{Resampling Factor} = \frac{1}{(\text{Normalized Peak Frequency})} \tag{7}$$

$$\text{Resampling Factor} = \frac{1}{(\text{1-Normalized Peak Frequency})} \tag{8}$$

Gallagher [11] noted that the algorithm's performance was better for low order interpolators such as bilinear and bicubic, but decreased as the order of interpolator increased. Gallagher [11] was unable to detect resampling factors of 2.0 since phase was preserved. Due to JPEG compression, peaks occurred at normalized frequencies of 1/8, 1/4, 3/8, 5/8, 3/4 and 7/8. These were ignored for all images, however if the image was resampled at a rate corresponding to those frequencies, it was difficult to determine.

Similar techniques to [11] were described in [13] for resampling detection using the second difference properties. The concept follows that when a sequence of P samples is resampled by a factor of M / N, there are now PM / N samples. If $M / N \geq 2$, there is at least one interpolated sample between a pair of original samples. Taking the second difference of the sample between the original pairs will produce a zero at that location. These zeros are periodic and do not occur in untampered sequences. Prasad and Ramakrishnan [13] generated a binary sequence from the second differences where the result was true if the second difference equaled zero as given by Equation (9). The DFT of this sequence would show distinct peaks indicating the presence of periodic zeros.

$$p[k] = \begin{cases} 1 & if|x''[k]| = 1 \\ 0 & otherwise \end{cases} \tag{9}$$

Alternatively, the zero-crossings of the second difference of the image was found before taking the DFT using Equation (10).

$$p[k] = \begin{cases} 1 & if x"[k] > 0 \text{ and } x"[k+1] \leq 0 \\ 1 & if x"[k] < 0 \text{ and } x"[k+1] \geq 0 \\ 0 & otherwise \end{cases} \tag{10}$$

A major weakness in both the algorithms in [11] and [13] is that they cannot detect interpolation if the image was rotated. Mahdian and Saic [14] proposed a new algorithm which was able to detect rotation and skewing using the radon transform. They selected a region of interest by dividing the image into overlapping blocks before applying the method.

Inspired by [11], Wei et al. [15] proposed an algorithm for rotation angle estimation. Nevertheless, it was their efforts in detection that was of interest. They compared the plots generated in [11], where the average of the second difference of each row was taken before computing the DFT, to one where the DFT was computed before taking the average. They found that when an image was scaled then rotated, the curve generated by the latter produced distinct peaks while the former did not. This was just a change in steps from [11] to achieve rotation tolerance. Wei et al. [15] also divided the image into overlapping blocks before applying their method.

Birajdar and Mankar [9] worked towards improving the algorithm in [13] by including estimation as done in [11]. Their input image was the Y component of an YCbCr image but tests were done using other color spaces. Their results showed that detection was better for the green component and grayscale image than it was for the Y component. Additionally, their algorithm was able to detect resampling factors of 2.0 unlike those by [11, 14].

3. THE PROPOSED DETECTION ALGORITHM: ADZ

The detection algorithm implemented aimed at detecting locally tampered regions, taking the Averaged Discrete Fourier transform of the Zero-crossing of the second difference of the resampled signal (ADZ). To achieve local tamper detection, the image was divided into blocks. The input image was the green plane of the tampered image, since [9] showed this plane worked better than other color spaces in resampling detection. Image forgeries often involve scaling the alien image to match the size of the objects in the image being tampered. Rotation is done for alignment. The technique which found the DFT then average of the image block used in [15] was chosen for its rotation tolerance. The zero-crossings of the second difference were incorporated as it was shown in [9] to detect resampling factors of 2.0. The algorithm's steps are outlined as follows for an $M \times N$ image:

- Compute the second difference: The green component G of the RGB image was taken and the second difference of each row was computed.

- Generate the binary image: Using (10), the zero-crossings of the second difference was found.

- Divide into blocks: The binary image was divided into overlapping blocks, each sized 64×64 with overlapping area of 32 pixels between each pair of blocks, row-wise and column-wise.

- Block-wise forgery detection: For each block, the DFT of all rows was generated. The average of the DFT magnitudes across the rows was found.

- Peak Detection: A threshold-based peak detector was used to find all peaks in the averaged spectrum that were 1.625 times the average in that array. If peaks existed in the averaged spectrum, the message "Interpolation Detected" and output image showing the interpolated block was displayed. If no peaks existed, the message "No Interpolation Detected" was displayed.

4. METHODOLOGY

4.1. Comparison of Interpolation Techniques

The manuals of five popular image editing software were found and the most commonly used interpolating functions were determined (see Figure 5). For the chosen interpolation techniques in the investigation, the methods of comparison as done in [1] were followed.

4.1.1 Frequency Analysis

Plots of the different interpolating functions truncated within the interval $-3 \leq x \leq 3$ were implemented using Matlab. The Fourier transformations were performed in order to study the frequency domain characteristics over the frequency band $-12 \leq f \leq 12$. The analysis looked at the

passband for deviation from the ideal, the cut-off point for the amplitude of the slope and the stopband for the occurrence and amplitudes of ripples and sidelobes. Attenuation in the passband leads to blurring while amplification will improve the transformed images sharpness as well as image noise. Aliasing effects are caused by small slopes having high cut-off amplitudes, and the presence of sidelobes and ripples aliases the continuous spectrum replicas into the passband [1].

4.1.2 Interpolation Quality

Taking an original image after histogram equalization a forward and backward transformation was done using the different interpolation techniques to produce the transformed image. The forward transformation was a magnification of 2.0 and the backward transformation was a magnification of 0.5. Thus the image returned to its original size after the transformations. The images within a 25-pixel border from both the original and transformed images were taken (to avoid border effects) for the following sequence of operations:

- The pixel-wise, absolute difference of the original and transformed images was found.
- A threshold was created for the resultant image to show true for all pixels equaling zero. This produced a black and white image where white represented all pixels that remained unchanged after the transformation and black represented otherwise.
- The image quality was measured using the normalized cross-correlation coefficient as shown in Equation (11) [1]. This compares the original and transformed images, and returns a value between 0 and 1, where 1 represents perfect similarity [16].

For the interpolation quality test, classic *512 x 512* grayscale “Lena” and “Living Room” TIFF images were used.

$$C = \left| \frac{\sum_{k,l} s(k,l)r(k,l) - KL\bar{s}\bar{r}}{\sqrt{(\sum_{k,l} s^2(k,l) - KL\bar{s}^2(\sum_{k,l} r^2(k,l) - KL\bar{r}^2)}} \right| \quad (11)$$

4.1.3 Runtime Measurements

The average time taken to transform an image using the different interpolation techniques was done in the Matlab environment. The transformation was a scale by 2.0. The time taken to run the transformation was found in seconds.

4.2. Detection of Interpolation Techniques

Preliminary tests were done to verify the best color space of the input image, determine a suitable block size and peak detection threshold. Global detection of an upsampled image was a good starting point since blocks indicating the tampered region should appear over the entire image. The green plane of the RGB image, grayscale image and Y component of the Ycbcr image were used as the input image to detect global resampling factors of 1.2, 2.0, 2.5 and 5. The green plane was chosen since it produced the most blocks. The threshold value was found empirically. The value of 1.625 maintained a good true negative rate on authentic images while detecting the most blocks on the globally scaled image. Values at the beginning and end of the averaged DFT trace were eliminated from the peak search to remove the high DC component which caused false positives. The block size was required to be large enough to detect periodicity in the pixel values, while not being so large that the tampered region was less than half the area. A greater overlapping area would increase the accuracy of the method but it would be computationally expensive. As recommended in [15], a block size of *64 x 64* pixels with 32 pixel overlap in each direction worked efficiently.

90 digital forgeries were created in Adobe Photoshop CS6 using uncompressed TIFF images from the UCID database [17]. Subsets of 30 images had tampered portions that were in

equivalence classes of: scaled only, rotated only, and scaled and rotated. These were resampled at various factors greater than 1.0. For each subset, bilinear and bicubic interpolation were used on 15 images each, with no post-processing done on the tampered regions. Nearest neighbor and Lanczos3 were not considered. This was because nearest neighbor is known to be the poorest interpolator and therefore less likely to be used in tampering. Lanczos3 was not included since it was known that detection accuracy reduces for higher order interpolators. This was not a problem since it is not widely used and this work sought to address interpolating functions found in commonly available image editing software. An additional 30 images were left untampered, giving a total of 120 images to test the algorithm. All images were a fixed size of *384 x 512* pixels. These test images were restricted to TIFF format to avoid the detection and elimination of peaks caused by JPEG compression.

The following terms were used in the evaluation of the ADZ algorithm: true positive, true negative, false positive and false negative. Where an interpolation-detected block fell on a tampered region, it was a true positive. If the block fell on a region that had not been resampled, it was a false positive. A false negative is counted for every block that should have fell on a tampered region, while a true negative is counted for every block that did not indicate interpolation on authentic areas. To measure the performance of the detection method, the precision and recall metrics were used on tampered images, and the specificity metric was used on authentic images. Precision is a measure of accuracy. That is, how well the algorithm determined a true positive region. Recall corresponds to the true positive rate, while specificity corresponds to the true negative rate [18].

$$\text{Precision} = \frac{TP}{(\text{TP} + \text{FP})} \tag{12}$$

$$\text{Recall} = \frac{TP}{(\text{TP} + \text{FN})} \tag{13}$$

$$\text{Specificity} = \frac{TN}{(\text{TN} + \text{FP})} \tag{14}$$

Where: TP - True Positive; FP - False Positive; TN - True Negative; FN - False Negative.

Since algorithms in previous works were not evaluated in this way, algorithms for [11, 9, and 15] were implemented using a common platform for comparison. The color space, block size and peak threshold identified in the respective papers were used. Preliminary tests were done to determine a range of the averaged DFT searched for peaks, to reduce false positives. For algorithms that considered only global detection, the entire image was used as a single block. The grayscale image was used as the input image in [11] as the color space was not specified. The peak threshold value used in [9] was also not specified but thought to be 10 since the algorithm followed [11], which used the same value. Only 15 bilinear interpolated images, and 15 authentic images were tested.

5. RESULTS

5.1. Comparison of Interpolation Techniques

The results of the Fourier transformations of the interpolation equations are shown in Figure 6. For each interpolation function, the graph shows its shape in the spatial domain (left graph), and the frequency domain (right graph). In the frequency domain, the deviation from the ideal interpolator, which is a rectangular function, is observed. The lanczos3 filter had the closest resemblance to the ideal, producing sharp images, while nearest neighbor showed the least with its high sidelobes, producing blurred images. Judging by the slopes at the cutoff point, and the sidelobes, the bilinear filter resulted in more aliasing than the bicubic filter.

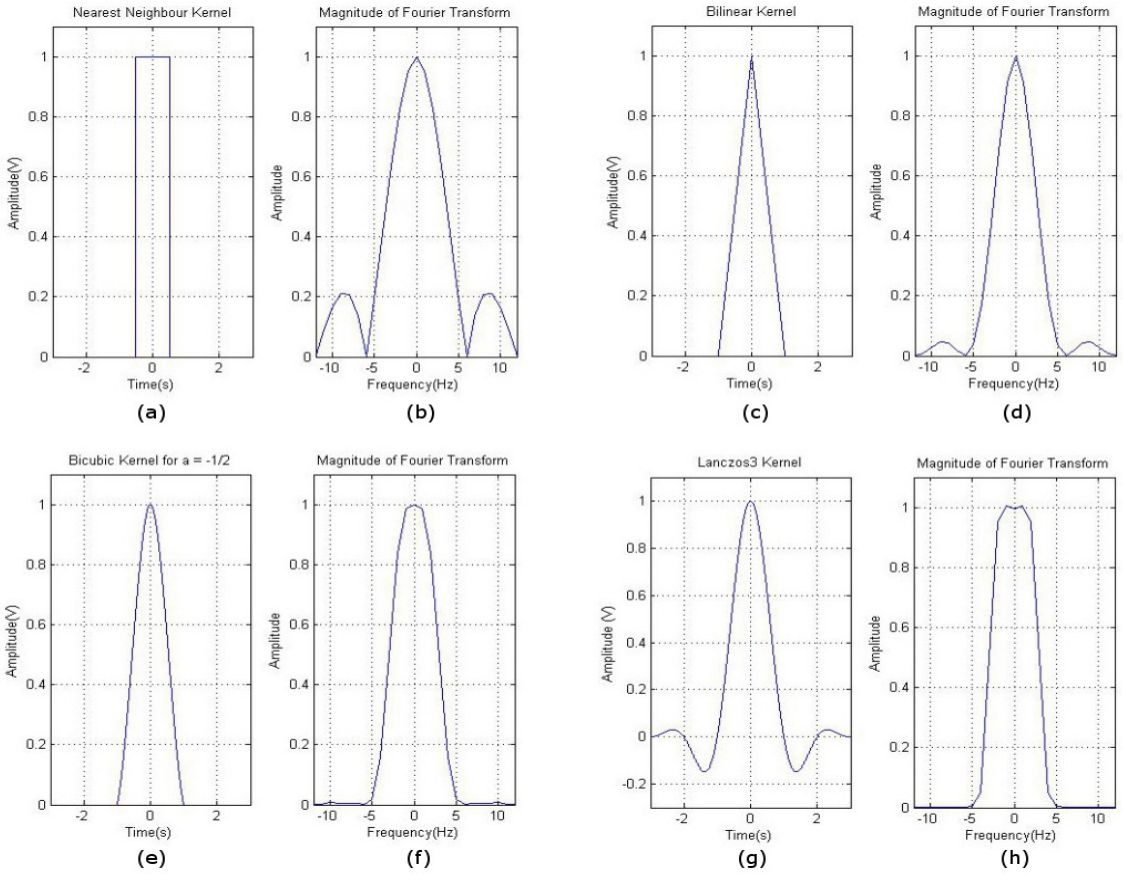


FIGURE 6: Fourier Analysis: (a) Nearest neighbor function in the spatial domain (b) Magnitude of nearest neighbour Fourier transform (c) Bilinear function in the spatial domain (d) Magnitude of bilinear Fourier transform (e) Bicubic function in the spatial domain (f) Magnitude of bicubic Fourier transform (g) Lanczos3 function in the spatial domain (h) Magnitude of lanczos3 Fourier transform.

The interpolation quality test results are shown in Figure 7. The binary images show those pixels that remained the same after the forward and backward transformations in white, and those that were changed in black. The greater the number of white pixels there are, the better the quality of the image. In both test cases, the nearest neighbor interpolator returned the image exactly. The result of bilinear interpolation showed significant loss of edge information, having more pixel values changed after transformation than bicubic interpolation. Lanczos3 interpolation showed the least error, excluding nearest neighbor.

FIGURE 7: Interpolation Image Error: (a) Original Lena image (b) Original Living room image (c) Lena image nearest neighbor error (d) Lena image bilinear error (e) Lena image bicubic error (f) Lena image lanczos3 error (g) Living room image nearest neighbor error (h) Living room image bilinear error (i) Living room image bicubic error (j) Living room image lanczos3 error.

TIFF Image	Nearest	Bilinear	Bicubic	Lanczos3
Lena	1.0	0.9973	0.9994	0.9997
Living room	1.0	0.9918	0.9979	0.999
Average	1.0	0.9946	0.9987	0.9994

TABLE 1: Normalized Cross-Correlation Similarity Measure.

TIFF Image	Nearest	Bilinear	Bicubic	Lanczos3
Lena	0.0101	0.0198	0.0296	0.10
Living room	0.0047	0.0135	0.0208	0.0240
Average	0.0074	0.01665	0.0252	0.062

TABLE 2: Average Time Taken to Perform a Transformation.

This visual quality assessment is supported by the quantitative measurement. Table 1 shows the results of the normalized cross-correlation coefficient. In both cases, the nearest neighbor function had perfect similarity, 1.0. On average, the transformation done using bilinear interpolation had the lowest similarity of 0.9946. The similarity increased with the bicubic function producing a value of 0.9987, and 0.9994 had the lanczos3 function.

A measurement of computational time (in seconds) to perform a transformation using the various interpolation techniques was performed. The results are shown in Table 2. On average, using nearest neighbor interpolation took the least time of 0.0074s. The time taken for bilinear, bicubic and lanczos3 interpolation increased with each one. These values were 0.01665s, 0.0252s and 0.062s respectively.

5.2. Detection of Interpolation Techniques

Figure 8 presents the precision, recall and specificity values. The overall precision is the percentage accuracy of the proposed ADZ algorithm to detect true positive regions, for all 30 images in each tampered subset. The same goes for the overall recall which shows the true positive rate. The bilinear and bicubic precision and recall values are those for the 15 images interpolated by the respective functions, in each subset. The specificity is the average true negative rate found for the 30 authentic images.

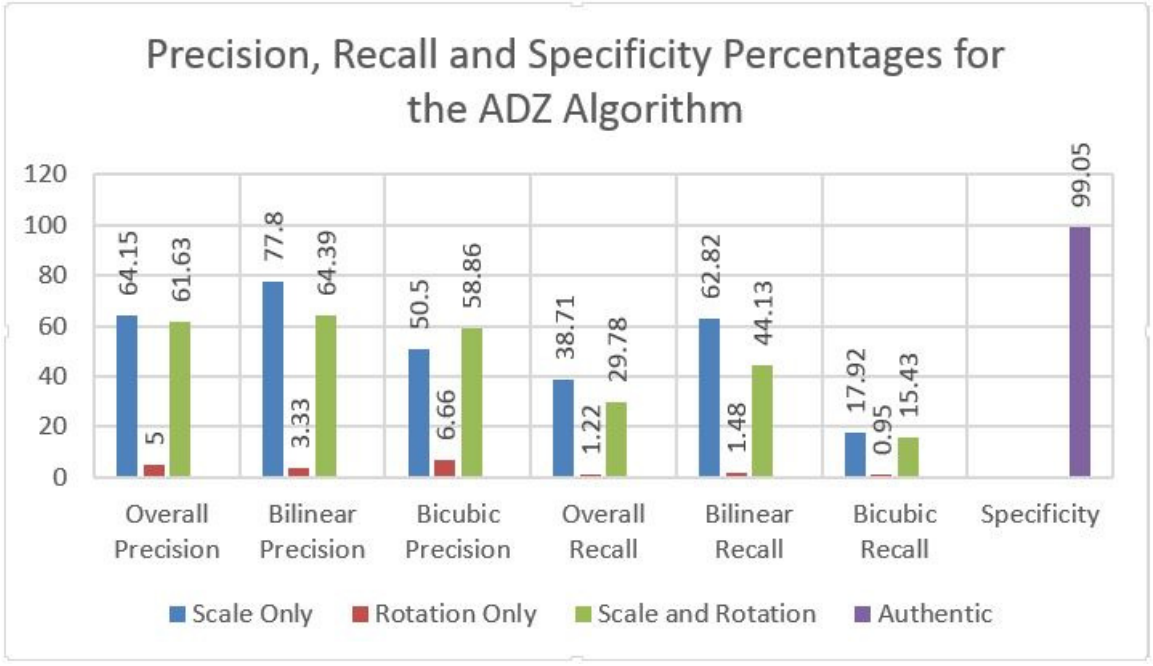


FIGURE 8: Precision, Recall and Specificity Percentages for the ADZ Detection Algorithm.

Figure 8 shows that on average, the ADZ algorithm implemented correctly identified 99.05% of blocks per authentic image as not resampled. That leaves an approximated 1% false positive

rate. The overall precision showed that for tampered, scaled images, only 64.15% of blocks detected were correctly identified as interpolated. This was closely followed at 61.63% in tampered images that had scaled and rotated regions. The algorithm performed poorly in detection of tampered portions that were rotated, with an accuracy of 5%. To be clear, the remainder of these percentages was for false positive blocks. The trend followed for the overall recall values. That is, the algorithm identified 38.71%, 29.78% and 1.22% of the tampered region for scaled only, scaled and rotated, and rotated only images respectively. The remainder of these percentages was for false negative blocks. For all subsets, the precision and recall values were higher for bilinear interpolation than bicubic. The only exception was in the precision of rotated only tampered regions, where bicubic interpolation had 6.66% accuracy and bilinear had 3.33%.

The comparison of the proposed ADZ algorithm to previous algorithms is shown in Figure 9. The proposed algorithm had 77.8% precision in scaled only tampered images, followed by 64.39% in scaled and rotated images, and 3.33% in rotated images. Likewise, the recall values were 62.82%, 44.13% and 1.48% for scaled only, scaled and rotated, and rotated only tampered images. The algorithm had a 99.5% specificity. The algorithms in [11] and [9] both had a specificity of 100%. However, the algorithm in [9] outperformed that in [11] for scaled only tampered images, having precision and recall values of 33.3% as compared to 6.67%. They both failed to detect rotated only, or scaled and rotated tampered regions, having 0%. The proposed ADZ algorithm is better matched by that in [15], which had its highest precision value of 58.67% in scaled and rotated tampered images, followed by 46.67%, and 6.67% in scaled only and rotated only tampered images respectfully. The recall values were 46.67% for scaled and rotated tampered images, followed by 18.32% and 0.83% in scaled only and rotated only tampered images. The specificity value for the 15 authentic images was 98.06%. The proposed ADZ algorithm performed the best as described in Section 6.

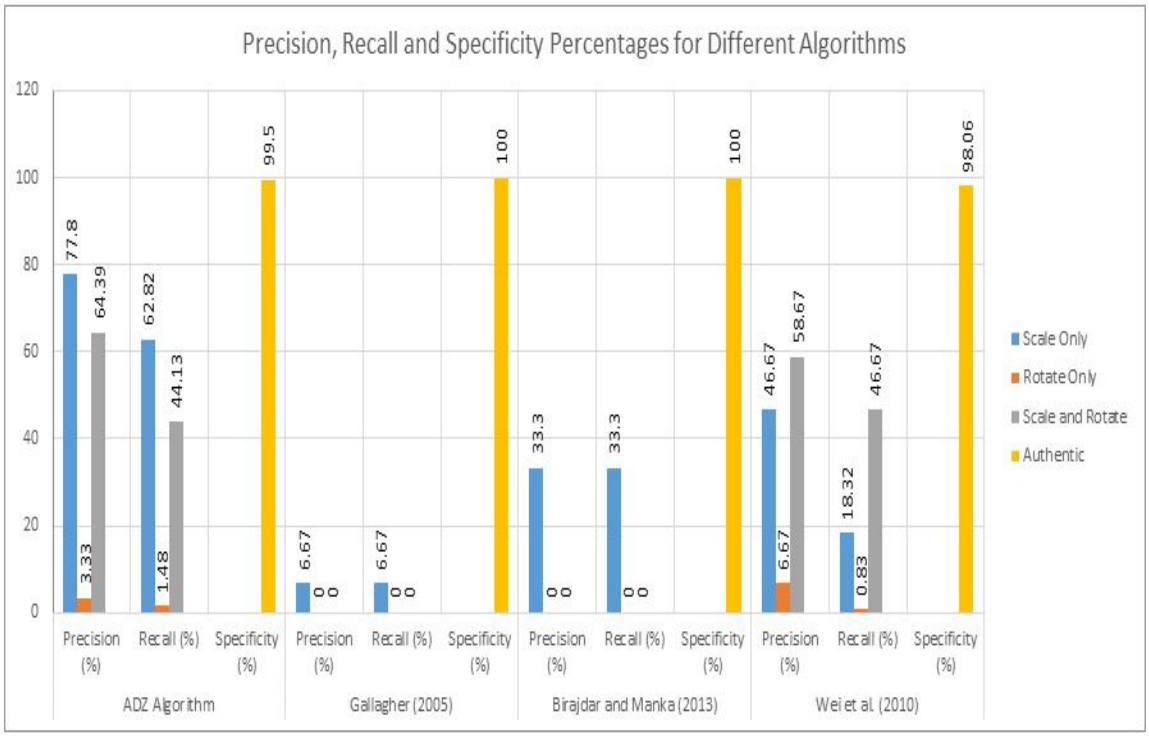


FIGURE 9: Precision, Recall and Specificity Percentages for Different Algorithms.

One of the objectives of the paper was to summarize where on the image the tampered region was detected. Figure 10 shows overlapping blocks in green indicating the detected regions. This is an example of a 100% accurate detection. It was noticed that on authentic images, where a

block fell on an all-white or textured region, false positives were detected. These are shown in Figure 11 and Figure 12 respectfully.

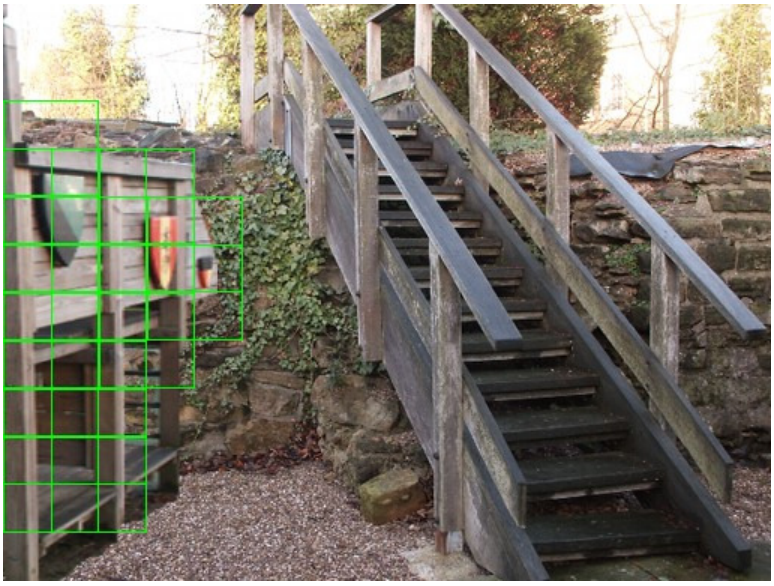


FIGURE 10: Detected Tampered Regions Indicated by the Green Boxes.

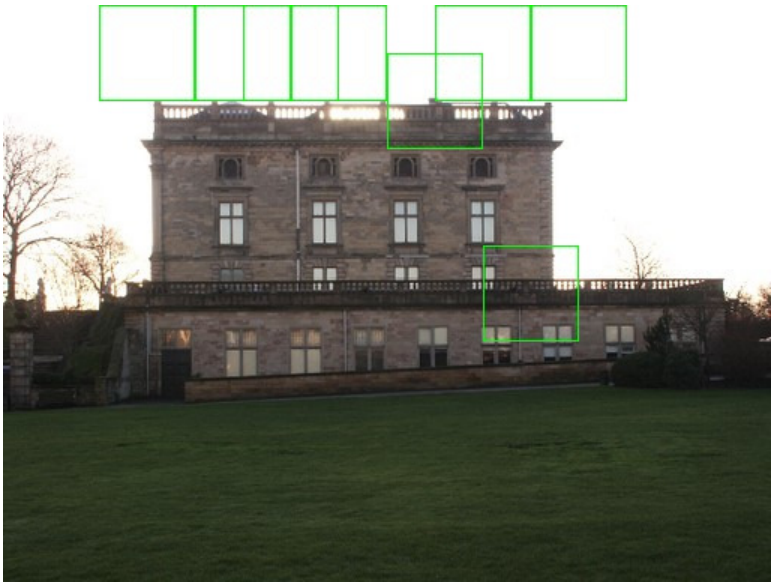


FIGURE 11: False Detected Regions of Solid White Indicated by the Green Boxes.

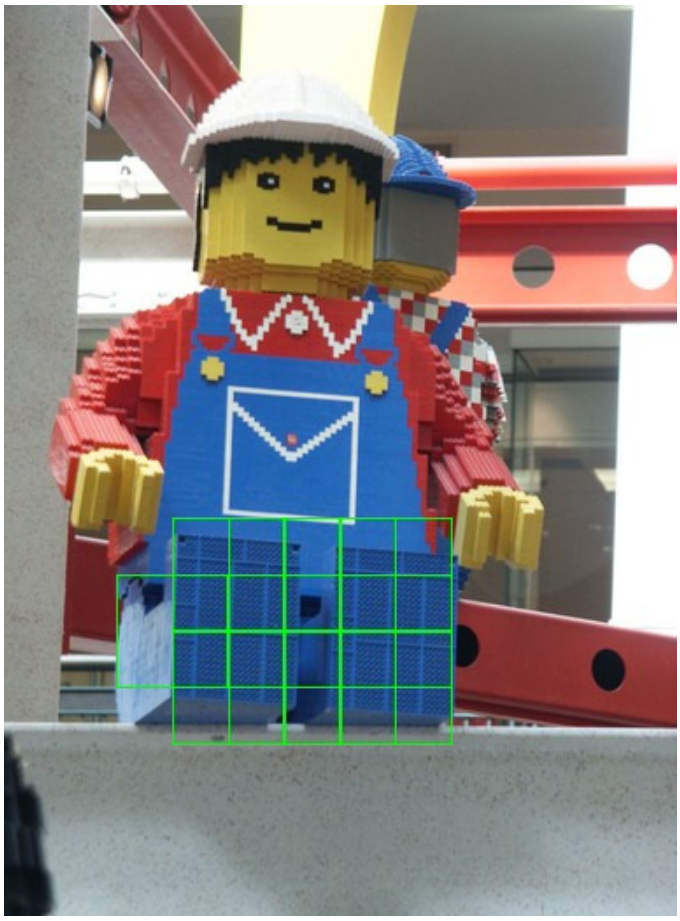


FIGURE 12: False Detected Textured Regions Indicated by the Green Boxes.

6. ANALYSIS

The results of the comparison of interpolation techniques were supported by the literature [1, 19]. The interpolation functions in order of increasing complexity were: nearest neighbor, bilinear, bicubic and lanczos3. The Fourier transformation plots were similar to those produced in [1]. The visual quality tests showed a deviation from what was expected where the nearest neighbor function reproduced an exact image. This was because the backward transform compensated the error introduced by the forward transform. In this case, the quality of nearest neighbor interpolation compared to others was not best represented since nearest neighbor performs the worst. The Fourier plots showed that nearest neighbor interpolation was not an effective filter. The prominent sidelobes would alias the continuous spectrum in the frequency domain, producing blocky images. The nearest neighbor was the simplest interpolator, considering only one pixel for its output, while bilinear considered four and bicubic considered sixteen. The result of the runtime measurement shows that the time taken to perform the same transformation increased for the more complex interpolation techniques. Therefore a better quality image is produced using increasingly complex interpolation techniques at the cost of computational time.

In the detection of authentic images, two cases were found to be incorrectly classified as interpolated; the image consisted of regions that were all-white within the block or had heavy texture. Detecting the white regions can be explained using Equation (6). Since the pixels values were the same, both the first and second difference of neighboring pixels would be zero. The algorithm searched for periodic zeros in the sequence to determine if it was tampered. This would be the same for all blocks of uniform color. The detected block can be ignored by human judgment. Alternatively, the algorithm may be modified to include a pre-processing phase to determine if the detected block had a majority of uniform color pixels using a threshold and if so,

ignore it. However, the computational time to detect the tampered block in the latter would increase. The false detection of textured regions was encountered in [14]. This occurred because textured regions displayed the same periodic patterns as resampled regions. Human judgment determined if the block should be ignored.

The ADZ algorithm worked best at detecting tampered regions that were scaled only, followed by regions that were scaled and rotated, but did not perform well for those that were rotated only. It was explained in [20] that in a rotated image, the periodicity that would exist in a row was skewed onto other rows. When each row was summed, the noise would be strong and no periodicity would be detected [21]. This explains why scaled only regions were detected better than those rotated. To overcome this issue, Qian et al. [20] proposed an algorithm that eliminated rows that were not factors of a constant, before searching for periodicity.

The results of the precision and recall values for bilinear and bicubic interpolation were as expected. Several sources [11, 14] indicated that bilinear interpolation produced larger peaks than bicubic interpolation. When the number of pixels considered producing the interpolated value increased, detection accuracy decreased. Currently, the use of higher order interpolators is not common so this detection is not an issue.

The proposed ADZ algorithm performed much better in detecting tampered regions than those in previous works, as it was able to detect scales of 2.0 and rotated tampered regions. The algorithm in [9] worked better than that in [11] since it was able to detect scales of 2.0, which [11] could not. These algorithms could not detect tampered regions that were rotated. The algorithm in [15] could detect rotated tampered regions, but not scales of 2.0. The color space used for the input image also played a role in the detection, as noted in [9].

The peak threshold and block size had an impact on the results. For undetected blocks, with or without rotated regions, peaks may have existed but the magnitudes fell below the threshold value. The block size determines the amount of data provided for peak detection, apart from restricting the range of values searched for peaks. The blocks were required to be large enough to detect the periodic patterns but not so large that the tampered region occupied less than half the area, leaving insufficient data to detect the patterns. Using a fixed block size would limit the applicability. Overlapping blocks were used to increase the accuracy of detection by shifting the blocks to cover more of the tampered regions however the block size may still be inappropriate to detect the periodicity. Furthermore, the entire image may not be searched for tampering if there were not enough columns or rows lefts to make a block. Increasing the overlap by having a one pixel shift would solve this issue at a cost of computational time. Therefore, improvements include development of optimal peak detection and adaptable block size algorithms. The adaptable block size algorithm can also ease evaluation where counting blocks is concerned.

In the investigation of the detection methods, it was seen that most algorithms incorporated estimation of the resampling factor. The evaluation of these techniques would tend towards the accuracy of the detected resampling factor after a region of interest was selected or discovered. In the algorithm proposed in this paper, estimation was not a concern. The evaluation focused more on the coverage of tampered regions. That is, how many blocks were detected? How many fell on tampered regions? How many did not fall on tampered regions? How many more should have been detected to cover the tampered regions? This meant that the creator of the tampered images for testing would be required to take the measurements since other persons would be unable to identify the tampered region, unless ground truth data was provided.

Counting blocks was a task done manually and therefore human error existed. The low precision and recall values recorded did not mean that the algorithm failed to detect tampered images, but that it did not detect every area of the tampered region. In fact, the specificity value showed that the detection algorithm had a near perfect rate of determining that authentic images were not interpolated. Implementing estimation and evaluating the detection accuracy for various resampling factors would give more comparable statistics to that in the literature. Doing so will

determine the range of resampling factors within the algorithm's limits. Furthermore, a proper investigation of different image formats and input image color spaces can be done as in [9]. Testing can be done on more complex forgeries made with post-processing like blurred edges or regions.

7. CONCLUSION

Digital forgeries often require an image to be scaled and/or rotated. The transformations involve re-sampling, which employs interpolation. Interpolating functions conceal the areas of tampering by varying degrees. While these changes may not be seen with the naked eye, interpolation introduces correlations in the underlying statistics of the image. This paper compared interpolation techniques used in digital image editing software and presented a novel algorithm (ADZ) for its automatic detection.

The comparison of the interpolation techniques followed the methods outlined in [1]. The interpolating functions considered were nearest neighbor, bilinear, bicubic and lanczos3 interpolation. The investigations looked at the Fourier transformations of the interpolating functions, the qualitative and quantitative image quality produced by the different functions, and runtime measurements. The results were comparable to that in the literature. Lanczos3 interpolation was the closest to ideal interpolation. Nearest neighbor performed the worst, producing pixelated images. Bilinear interpolation produced a smoother image with artifacts, while bicubic interpolation produced a sharper image with fewer artifacts. Computational time increased as the complexity of the interpolator increased.

This paper proposed the ADZ algorithm that was able to identify local tampered regions by dividing the image into blocks and searching for periodicity in the zero-crossings of the second derivative of the resampled signal. This method was chosen over the Expectation-Maximization method as it was faster and based on a simpler principle. The algorithm was tested on upsampled images only, as suggested in [22], on authentic and tampered images, where the tampered regions were scaled, rotated or scaled and rotated using bilinear and bicubic interpolation. The performance was measured using the specificity metric for authentic images, and the precision and recall metrics for tampered images. The algorithm worked well to determine that authentic images were not resampled. However, the accuracy of detecting tampered regions was satisfactory for those that were scaled only, and scaled and rotated, but failed for rotation only regions. It was found that bilinear interpolation was detected better than bicubic. The algorithm performed better than those in previous works.

Several improvements can be made in the detection of interpolation as discussed in section 6. Therefore, future work includes:

- Estimation of resampling factors
- Investigation and application of techniques to improve rotation detection
- Investigation and application of optimal peak detection strategies
- Development of an adaptable block size algorithm
- Experiments using JPEG compressed Tampered images
- Experiments using different input image color spaces

This work shows that there may be a great impact on the detection of locally tampered images using resampling methods. It also outlines performance metrics-driven methodology which can be used in assessing any future work in tampered image detection, in order to have a similar

benchmark across the varying techniques. The Interpolation techniques are always evolving, so research into detection methods must continue [23].

8. ACKNOWLEDGMENT

The authors would like to thank the Department of Electrical and Computer Engineering at The University of the West Indies for supporting this research.

9. REFERENCES

[1] T. M. Lehmann, C. Gonner, and K. Spitzer, “Survey: Interpolation methods in medical image processing,” *Medical Imaging, IEEE Transactions on*, vol. 18, no. 11, pp. 1049-1075, 1999.

[2] A. Popescu and H. Farid, “Exposing digital forgeries by detecting traces of resampling," *Signal Processing, IEEE Transactions on*, vol. 53, no. 2, pp. 758-767, Feb 2005.

[3] H. Farid, “Image forgery detection," *Signal Processing Magazine, IEEE*, vol. 26, no. 2, pp. 16-25, March 2009.

[4] T. Qazi, K. Hayat, S. U. Khan, S. A. Madani, I. A. Khan, J. Ko lodziej, H. Li, W. Lin, K. C. Yow, and C.-z. Xu, “Survey on blind image forgery detection," *Image Processing, IET,* vol. 7, no. 7, pp. 660-670, 2013.

[5] J. A. Parker, R. V. Kenyon, and D. Troxel, “Comparison of interpolating methods for image resampling*," Medical Imaging, IEEE Transactions on,* vol. 2, no. 1, pp. 31-39, 1983.

[6] T. Acharya and P.-S. Tsai, “Computational foundations of image interpolation algorithms," *ACM Ubiquity,* vol. 8, no. 42, 2007.

[7] P. Thevenaz, T. Blu, and M. Unser, “Image interpolation and resampling," *Handbook of medical imaging, processing and analysis*, pp. 393-420, 2000.

[8] K. Turkowski, “Filters for common resampling tasks," in Graphics Gems. Academic Press Professional, Inc., 1990, pp. 147-165.

[9] G. K. Birajdar and V. H. Mankar, “Blind authentication of resampled images and rescaling factor estimation," in *Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), 2013 International Conference on. IEEE*, 2013, pp. 112-116.

[10] T. M. Lehmann, C. Gonner, and K. Spitzer, “Addendum: B-spline interpolation in medical image processing," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 7, pp. 660-665, 2001.

[11] A. C. Gallagher, “Detection of linear and cubic interpolation in jpeg compressed images," in *Computer and Robot Vision, 2005. Proceedings. The 2nd Canadian Conference on*. IEEE, 2005, pp. 65-72.

[12] D. Zhu and Z. Zhou, “Resampling tamper detection based on jpeg double compression," in *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*. IEEE, 2014, pp. 914-918.

[13] S. Prasad and K. Ramakrishnan, “On resampling detection and its application to detect image tampering," in *Multimedia and Expo, 2006 IEEE International Conference on.* IEEE, 2006, pp. 1325-1328.

[14] B. Mahdian and S. Saic, “Blind authentication using periodic properties of interpolation," *Information Forensics and Security, IEEE Transactions on*, vol. 3, no. 3, pp. 529-538, 2008.

[15] W. Wei, S. Wang, X. Zhang, and Z. Tang, “Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery," *Information Forensics and Security, IEEE Transactions on,* vol. 5, no. 3, pp. 507-517, 2010.

[16] T. Lehmann, A. Sovakar, W. Schmiti, and R. Repges, “A comparison of similarity measures for digital subtraction radiography," *Computers in Biology and Medicine,* vol. 27, no. 2, pp. 151-167, 1997.

[17] G. Schaefer and M. Stich, “Ucid: an uncompressed color image database," in *Electronic Imaging 2004.* International Society for Optics and Photonics, 2003, pp. 472-480.

[18] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks*," Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.

[19] E. Maeland, “On the Comparison of Interpolation Methods," *Medical Imaging, IEEE Transactions on,* vol. 7, no. 3, pp. 213-217, 1988.

[20] R. Qian, W. Li, N. Yu, and Z. Hao, “Image forensics with rotation-tolerant resampling detection," in *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on*. IEEE, 2012, pp. 61-66.

[21] B. Mahdian and S. Saic, “Detection of Resampling Supplemented with Noise Inconsistencies Analysis for Image Forensics," in *Computational Sciences and Its Applications (ICCSA 2008), International Conference on*. IEEE, 2008, pp. 546-556.

[22] Y.T. Kao, H.J. Lin, C.W. Wang and Y.C. Pai, “Effective Detection for Linear Up-Sampling by a Factor of Fraction," *Image Processing, IEEE Transactions on,* vol. 21, no. 8, pp. 3443-3453, 2012.

[23] S. Math and R.C. Tripathi, “Digital Forgeries: Problems and Challenges," *International Journal of Computer Applications,* vol. 5, no. 12, pp. 9-12, 2010.

Image Contrast Enhancement for Brightness Preservation Based on Dynamic Stretching

M.A. Rahman *md.arifur.rahman056@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

Shilong Liu *shilong.liu@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

S.C.F. Lin *stephen.lin@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

C.Y. Wong *chin.wong@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

G. Jiang *guannan.jiang@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

Ngaiming Kwok *nmkwok@unsw.edu.au*
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia

Abstract

Histogram equalization is an efficient process often employed in consumer electronic systems for image contrast enhancement. In addition to an increase in contrast, it is also required to preserve the mean brightness of an image in order to convey the true scene information to the viewer. A conventional approach is to separate the image into sub-images and then process independently by histogram equalization towards a modified profile. However, due to the variations in image contents, the histogram separation threshold greatly influences the level of shift in mean brightness with respect to the uniform histogram in the equalization process. Therefore, the choice of a proper threshold, to separate the input image into sub-images, is very critical in order to preserve the mean brightness of the output image. In this research work, a dynamic range stretching approach is adopted to reduce the shift in output image mean brightness. Moreover, the computationally efficient golden section search algorithm is applied to obtain a proper separation into sub-images to preserve the mean brightness. Experiments were carried out on a large number of color images of natural scenes. Results, as compared to current available approaches, showed that the proposed method performed satisfactorily in terms of mean brightness preservation and enhancement in image contrast.

Keywords: Image Contrast Enhancement, Histogram Equalization, Brightness Preservation, Golden Section Search.

1. INTRODUCTION

Image processing techniques have been applied in a wide range of engineering applications such as consumer electronic devices [1], robotics [2], machine condition monitoring [3], and remote sensing [4]. In order to obtain satisfactory performance in these image based applications, the enhancement of image contrast is a fundamental requirement. Moreover, it is necessary to preserve the mean brightness of the processed image by which a better perception of the scene can be delivered to the viewer. To this end, the histogram equalization technique is an attractive candidate approach for contrast improvement because of its implementation tractability and simplicity.

Histogram equalization (HE) based image contrast enhancement is able to provide contrast improvements; however, the method tends to shift the output image brightness to the middle of the permitted intensity range [5]. The need to preserve the mean brightness of an image had thus motivated a large number of research works. A class of methods had been developed which were based on separating the image into high and low intensity sub-images and then apply equalization to each component independently. An early work had initiated schemes dividing the input image with the mean brightness in the bi-histogram equalization method (BBHE) [6]. This idea was later modified to incorporate the median value as the divider threshold in the process called dualistic sub-image histogram equalization (DSIHE) [7]. However, due to the fact that HE does not have control on the enhancement level, the output image might suffer from over-enhancement, especially if the input image contains significantly brighter or darker regions than other parts of the image. As the sub-images are processed using histogram equalization, these algorithms cannot ensure the mitigation of over-enhancement or an exact match of the input and output mean brightness values.

In order to overcome such limitation, an alternative approach of clipping the peaks of each sub-image histogram is employed to match the target histogram for achieving equalization. A pioneer example is the bi-histogram equalization with plateau level (BHEPL) where the image was separated by the mean brightness value and each histogram of a sub-image was clipped using the mean of the sub-image histogram [8]. This approach was later modified to utilize the median of sub-image histogram for clipping (BHEPLD) [9]. A recent work, adopting the bi-histogram equalization principle, adaptive image enhancement based on bi-histogram equalization with a clipping limit (AIEBHE), was reported in [10]. The clipping limit imposed on each sub-image histogram depends on the minimum of the mean and median values obtained from the histogram. These variations had resulted in the mitigation of over-enhancements, but the preservation of mean brightness in the output image was not purposefully considered. On the other hand, the authors in [11] adopted an exhaustive search approach to find a separation threshold. However, with the histogram equalization on sub-images, an exact match of input and output mean brightness still cannot be guaranteed.

In addition to dividing the input image into two sub-images to process them independently, researchers also suggested dividing the image into multiple sub-images. For instance, the dynamic quadrants histogram equalization plateau limit (DQHEPL) method had divided the image into four sub-images [9]. The clipping limits were obtained from the mean values of each sub-image histograms. The concept of using quadratic sub-images was further carried forward in the process called median-mean based sub-image clipped histogram equalization (MMSICHE) [12]. Specifically, the input image was first separated by the mean brightness. Each sub-image was further divided according to the mean brightness of the corresponding sub-image. The peaks of the sub-image histograms were clipped with the median values. The implementation of these algorithms were made more complicated, however, there was no assurance that the mean brightness of the output image can be preserved.

On the other hand, the method, proposed in [13], emulates the image histogram using the Gaussian mixture model and separates it using the intersection points of the Gaussian components. The sub-images are then separately stretched using the estimated Gaussian parameters. Global contrast enhancement is achieved by applying this technique on both color

and depth images [14]. The homomorphic filtering is applied on the modified histogram in order to preserve mean brightness in [15].

In this work, instead of employing histogram equalization as the main contrast enhancement processor, the dynamic stretching strategy is adopted. The concept of using mean or median value as the separation threshold is abandoned. Instead, the efficient golden section search approach is employed to find an optimal threshold such that the preservation of mean brightness in the output image is accomplished. Experiments were conducted on a large number of color images and statistics were collected to verify the proposed method.

The rest of this paper is organized as follows. In Section 2, related backgrounds are reviewed. The proposed brightness preservation method using dynamic stretching is described in Section 3. Experiments and results are presented in Section 4. Section 5 contains the conclusion.

2. RELATED BACKGROUND

There are classes of reported works that aim at image contrast enhancement by equalizing the image histograms of separated sub-images and the clipping of histograms. The choice of the separation threshold and the clipping limit, such as the mean or median value, has not been unified. In this section, several representative approaches are reviewed.

2.1 Conventional Histogram Equalization

This is the basis of many aforementioned image contrast enhancement algorithms. Given an intensity image, which may be converted from a color image, the pixel value or brightness can be described as

$$I(u,v) \in [0 \;\; L-1], \tag{1}$$

where (u,v) is the pixel coordinate in the image, $L-1$ is the maximum allowed intensity level. For example, if the image is represented as an 8-bit digital signal, then $L = 2^8 = 256$. In general, we have $u = 1,\ldots,U$ and $v = 1,\ldots,V$; where the image is of the size $N = U \times V$ in width-by-height.

In histogram equalizations [5], at first the statistics of the image is collected, giving the histogram

$$\mathbf{h} = [n(i)], \tag{2}$$

where i is the intensity index, $n(i)$ is the number of pixels having intensity i. The histogram is often normalized to produce the probability density $p(i)$ by diving $\mathbf{h}$ with the total number of pixels N in the image, such that

$$\sum_{i=0}^{L-1} p(i) = 1.$$

The cumulative density is then constructed from

$$\begin{aligned} c(i) &= \sum_{j=0}^{i} p(j) \text{ and} \\ \max\{c(i)\} &= 1. \end{aligned} \tag{3}$$

The enhanced or output image is obtained from mapping the original intensity to a desired level as

$$I_{enh}(i) = I_{\min} + (I_{\max} - I_{\min}) \times c(i), \tag{4}$$

where $I_{\min}$ and $I_{\max}$ are the desired minimum and maximum intensity levels of the enhanced image.

2.2 Sub-image Based Histogram Equalization

The conventional histogram method can be regarded as a global approach where all pixels are processed. Conversely, there are algorithms that separate the input into sub-images and manipulate the pixel intensities using different procedural settings for each sub-image.

In the bi-histogram equalization algorithm, the input image is separated into two sub-images to produce two histograms and each is processed independently [6]. First, the mean intensity or brightness of the input image is calculated from

$$\mu_m = \sum_{i=0}^{L-1} (i \times p(i)). \tag{5}$$

Two sub-images $\mathbf{I}_{hi}$ and $\mathbf{I}_{lo}$ are created, where

$$\begin{aligned} \mathbf{I}_{hi} &= \{I(i) | i > \mu_m\}, \\ \mathbf{I}_{lo} &= \{I(j) | j \le \mu_m\}. \end{aligned} \tag{6}$$

Furthermore, two histograms $\mathbf{h}_{hi}$ and $\mathbf{h}_{lo}$ are formed and cumulative densities are computed. Finally, each sub-image is equalized using intensity ranges $[0 \ \mu_m]$ and $[\mu_{m+1} \ L-1]$, giving two enhanced sub-images $\mathbf{I}_{enh,hi}$ and $\mathbf{I}_{enh,lo}$. The final enhanced image is the aggregation of these sub-images, computed as,

$$\mathbf{I}_{enh} = \mathbf{I}_{enh,hi} \cup \mathbf{I}_{enh,lo} . \tag{7}$$

When the histogram equalization processes are carried out using a uniform target histogram, the mean brightness values of the enhanced sub-images are shifted to the middle of the corresponding ranges, that is

$$\mu_{enh,hi} = 0.5 \times (\mu_m + L), \ \mu_{enh,lo} = 0.5 \times \mu_m . \tag{8}$$

The output image mean brightness, $\mu_{enh,m}$ is the sum of these two mean intensities weighted by the respective number of pixel used from original image. We have

$$\mu_{enh,m} = w_{hi} \times \mu_{enh,hi} + w_{lo} \times \mu_{enh,lo} , \tag{9}$$

where $w_{hi} = n(i)/N$, $i = 0, \ldots, \mu_m$; and $w_{lo} = n(j)/N$, $j = \mu_{m+1}, \ldots, L-1$. It can be observed that the output image mean brightness depends on the features of the image. The limitation of this method is that only when the two sub-images have the same number of pixels or when two

weights are at $w_{hi} = w_{lo} = 0.5$ and $\mu_m = L/2$, the brightness preservation could be accomplished.

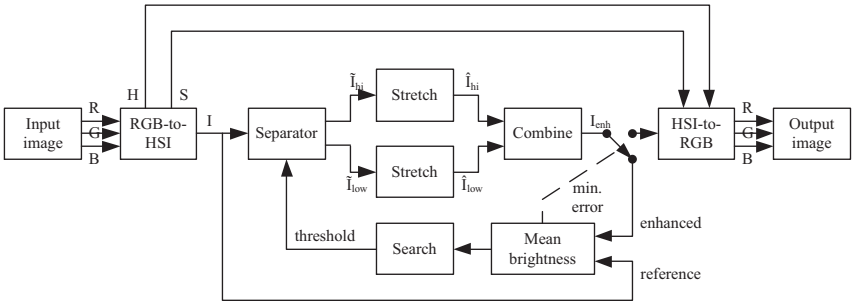


FIGURE 1: Block diagram of the proposed brightness preserving dynamic stretching approach.

2.3 Clipped Histogram Equalization

In addition to the shift in mean intensity, the conventional histogram equalization method also suffers from the generation of artifacts where abrupt intensity changes appear in the output image. This drawback arises where pixel intensities tend to concentrate in a narrow range giving rises to peaks in the histogram. In order to mitigate this problem, the histogram peaks are clipped prior to the calculation of the cumulative density [8]. Similar initial procedures in bi-histogram equalization are adopted where two sub-images are obtained and two histograms are constructed. Then two clipping limits are computed from

$$T_{hi} = \sum_{i=0}^{\mu_m} (i \times p(i))$$
$$T_{lo} = \sum_{j=\mu_m+1}^{L-1} (j \times p(j)), \tag{10}$$

which are effectively the mean values of the individual sub-histograms. The histograms are clipped such that

$$\mathbf{h}_{hi} = \begin{cases} h_{hi}(i), & h_{hi}(i) < T_{hi} \\ T_{hi}, & \text{otherwise} \end{cases}$$
$$\mathbf{h}_{lo} = \begin{cases} h_{lo}(i), & h_{lo}(i) < T_{lo} \\ T_{lo}, & \text{otherwise}. \end{cases} \tag{11}$$

The subsequent processes follow the conventional histogram equalization procedure where each sub-image is manipulated independently as found in the bi-histogram equalization algorithm.

The problem associated with this clipping action is that the resultant image mean brightness does not normally have a closed-form solution. This is due to the variations in image contents and the difficulties in choosing proper clipping limits.

3. DYNAMIC STRETCHING BASED BRIGHTNESS PRESERVATION

In order to enhance the image contrast and to preserve the output image mean brightness, an approach based on dynamic range stretching is proposed and coined as the dynamic stretching based brightness preservation (DSBP) method. The method adopts the sub-image separation principle; in addition, the separation threshold is obtained from the efficient golden section search approach [16].

The system block diagram is shown in Fig. 1. Given a color image in the red-green-blue format, it is first converted to the hue-saturation-intensity (HSI) space where the intensity is extracted by adopting the method described in [5]. This signal is often called the gray-level image. The I-channel is then separated into the low and high groups, $\mathbf{I}_{hi}$ and $\mathbf{I}_{lo}$, by the threshold to be determined with golden section search method. These two groups are stretched to occupy their corresponding complete intensity range and combined again to produce the enhanced image. The overall mean intensity or brightness of this image is calculated and compared with the input image mean brightness. The difference, namely brightness error, is used to drive the search for an optimal threshold. The separation, stretching and brightness error calculation processes are iterated until a threshold is found that minimizes the brightness error. Finally, when the iteration terminates, the enhanced intensity component is combined with the original hue and saturation to give the output image through the HSI to RGB conversion.

To start the Golden Section search process, a trial threshold $\tilde{I}_m \in [0 \;\; L-1]$ is determined to separate the image into two sub-images $\tilde{\mathbf{I}}_{hi}$ and $\tilde{\mathbf{I}}_{lo}$, where

$$\begin{aligned} \tilde{\mathbf{I}}_{hi} &= \{I(i) \mid i > \tilde{I}_m\}, \\ \tilde{\mathbf{I}}_{lo} &= \{I(j) \mid j \le \tilde{I}_m\}. \end{aligned} \tag{12}$$

The dynamic range of each sub-image is then stretched up to the permitted intensities. That is

$$\hat{\mathbf{I}}_{hi} = \tilde{I}_{m+1} + \frac{\tilde{\mathbf{I}}_{hi} - \min\{\tilde{\mathbf{I}}_{hi}\}}{\max\{\tilde{\mathbf{I}}_{hi}\} - \min\{\tilde{\mathbf{I}}_{hi}\}} \times \left(L - 1 - \tilde{I}_{m+1}\right) \tag{13}$$

$$\hat{\mathbf{I}}_{lo} = \frac{\tilde{\mathbf{I}}_{lo} - \min\{\tilde{\mathbf{I}}_{lo}\}}{\max\{\tilde{\mathbf{I}}_{lo}\} - \min\{\tilde{\mathbf{I}}_{lo}\}} \times \tilde{I}_m \,. \tag{14}$$

The two stretched sub-images are aggregated according to Eq. 7 and the mean brightness is calculated with Eq. 5.

Furthermore, a second trial threshold is selected to obtain another mean intensity. Hence, for these two thresholds $\tilde{I}_{m,1}$ and $\tilde{I}_{m,2}$, we have mean brightness $\hat{\mu}_1$ and $\hat{\mu}_2$. When the trial threshold cannot produce the desired mean brightness preservation, a search procedure is invoked to find a proper threshold by using the efficient golden search algorithm. A search ratio $\rho = \left(\sqrt{5}-1\right)/2$ as well as a threshold range $\Delta_m = \tilde{I}_{m,2} - \tilde{I}_{m,1}$ is defined; assuming $\tilde{I}_{m,2} > \tilde{I}_{m,1}$. For this work, $\tilde{I}_{m,1}$ and $\tilde{I}_{m,2}$ are set to the allowed intensity limit, that is, $\tilde{I}_{m,1} = 0$ and $\tilde{I}_{m,2} = L-1$. The function that is used for the golden section search is

$$f(\mu) = \mid \mu - \mu_m \mid, \tag{15}$$

which reaches the minima when the mean brightness of the enhanced image μ is equal to the mean brightness of the input image μ_m, hence make the function unimodal which is a requirement of the golden section search algorithm. The proposed process then proceeds as given in Algorithm 1 below.

Algorithm 1 Dynamic Stretching based Brightness Preservation Algorithm

Input: $\rho, \Delta_m, \mu_m, \tilde{I}_{m,1}, \tilde{I}_{m,2}, \hat{\mu}_1, \hat{\mu}_2$

Output: enhanced image $\mathbf{I}_{enh}$

Define convergence tolerance $\tau < 1/L, \Delta_m$

while $\Delta_m > \tau$

set $\Delta\tilde{I}_{m,1} = | \hat{\mu}_1 - \mu_m |, \Delta\tilde{I}_{m,2} = | \hat{\mu}_2 - \mu_m |$

if $\Delta\tilde{I}_{m,1} > \Delta\tilde{I}_{m,2}$ **then**

set $\tilde{I}_{m,1} \leftarrow \tilde{I}_{m,1} + \Delta_m \times \rho$

else

set $\tilde{I}_{m,2} \leftarrow \tilde{I}_{m,1} + \Delta_m \times (1-\rho)$

end if

Do image separating and stretching using $\tilde{I}_{m,1}$ and $\tilde{I}_{m,2}$ to obtain $\hat{\mu}_1$ and $\hat{\mu}_2$

Calculate $\Delta_m = \tilde{I}_{m,2} - \tilde{I}_{m,1}$

end while

Aggregate stretched sub-images to give $\mathbf{I}_{enh}$

4. EXPERIMENTS

A collection of 220 color images, including 20 publicly available standard test images from http://r0k.us/graphics/ko-dak/, captured in natural environment such as clear day, night time and poor conditions with overcast weather, were used in the experiment to verify the performance of the proposed method. The image sizes are 400×300 pixels width-by-height and are stored in 24-bit RGB colored JPG format. The qualities of the enhanced image are assessed and compared in terms of metrics including brightness error, contrast, entropy, and gradient as defined below. The statistics of the metrics for all the images are plotted in box plots for further comparison of the proposed dynamic stretching based brightness preservation (DSBP) method to other available methods namely, conventional uniform histogram equalization (UNFHE), BBHE, DSIHE, BHEPL, BHEPLD, AIEBHE, and MMSICHE.

Brightness Error:

The deviation in mean brightness while approximating the enhanced image in comparison to that of the input image is referred as the brightness error [6] and expressed as following

$$\mu_{err} = \mu_{enh} - \mu_m \,. \tag{16}$$

From Eq. 16, it is clear that if the average brightness is preserved in the output image, the error will be as minimum as zero.

Contrast:

The contrast depends not only on the average intensity but also on the dispersion around a central pixel [10], as expressed as,

$$\mathcal{C} = \frac{1}{N}\sum I_{enh}^2(x,y) - \left(\frac{1}{N} I_{enh}(x,y)\right)^2, \tag{17}$$

where (x, y) refers to the pixel location. A higher value of $\mathcal{C}$ denotes better contrast.

Entropy:
It is a measure of the information content of the image and defined as

$$\mathbf{H} = -\sum_{i=0}^{L-1} p(i)\log_2 p(i), \tag{18}$$

where a higher value is desirable.

Gradient:
According to [5], the gradient is defined as

$$\mathcal{G} = \frac{1}{N}\sum_{x,y}\left(\Delta x^2 + \Delta y^2\right), \tag{19}$$

where, $\Delta x = I_{enh}(x,y) - I_{enh}(x+1,y)$ and $\Delta y = I_{enh}(x,y) - I_{enh}(x,y+1)$. The larger the gradient value is, the sharper the image is perceived by a viewer.

4.1 Qualitative Evaluation

For the visual assessment, two sample images and their outputs from different approaches are shown in Fig. 2 and Fig. 3.

The standard benchmark image, shown in Fig. 2(a), contains a natural outdoor scene under daylight. Figures 2(b), 2(c), and 2(d) are obtained from the methods UNFHE, BBHE, DSIHE respectively, which are based on image separation and equalization towards a uniform distribution. As these methods do not provide any restriction on the processed brightness, the output images from these methods are observed to be brighter than the input image. Enhancements from the implementations based on clipping of histogram peak are shown in Figures 2(e), 2(f), 2(g), 2(h), and 2(i). For these methods, though the output brightness is close to the input, significant artifacts can be observed in the homogeneous region of the sky. On the contrary, the output from the proposed method, shown in Fig. 2(j), does not suffer from such limitation. The image is enhanced with better color tone and higher contrast, thus makes it more appealing to the human viewer. Moreover, the change in brightness is not noticeable.

Figure 3 depicts another example image sampled from the dataset along with the generated outputs from all the approaches compared. The image is captured in poor weather condition; hence the color is not vivid and the contrast is low. It can be observed that results from UNFHE, BBHE, DSIHE tend to generate viewing artifacts where there is abrupt change in intensity. Moreover, they produce patches that are either too dark or too bright, such that feature details are lost in the resultant image. For the cases of BHEPL, BHEPLD, AIEBHE, and MMSICHE, the artifacts are mitigated, but the improvement in contrast is not significant; especially the output from BHEPL gives the least contrast as observed by a human viewer. On the other hand, the result from DQHEPL has such a large deviation from the original mean brightness that the sea region in the image lost the fine details. The enhanced image from the proposed method is shown in Fig. 3(j). It is evident from the image that the proposed method not only maintains the mean brightness as close to that of input image, but also increases the image contrast in order to improve the fine details of the image. With more vivid color and improved information content, the output of the proposed method achieves higher perceptual performance than the other methods in comparison.

FIGURE 2: Sample experiment result - 1. Enhancement methods: (a) input image, (b) UNFHE, (c) BBHE, (d) DSIHE, (e) BHEPL, (f) BHEPLD, (g) AIEBHE, (h) DQHEPL, (i) MMSICHE, (j) DSBP.

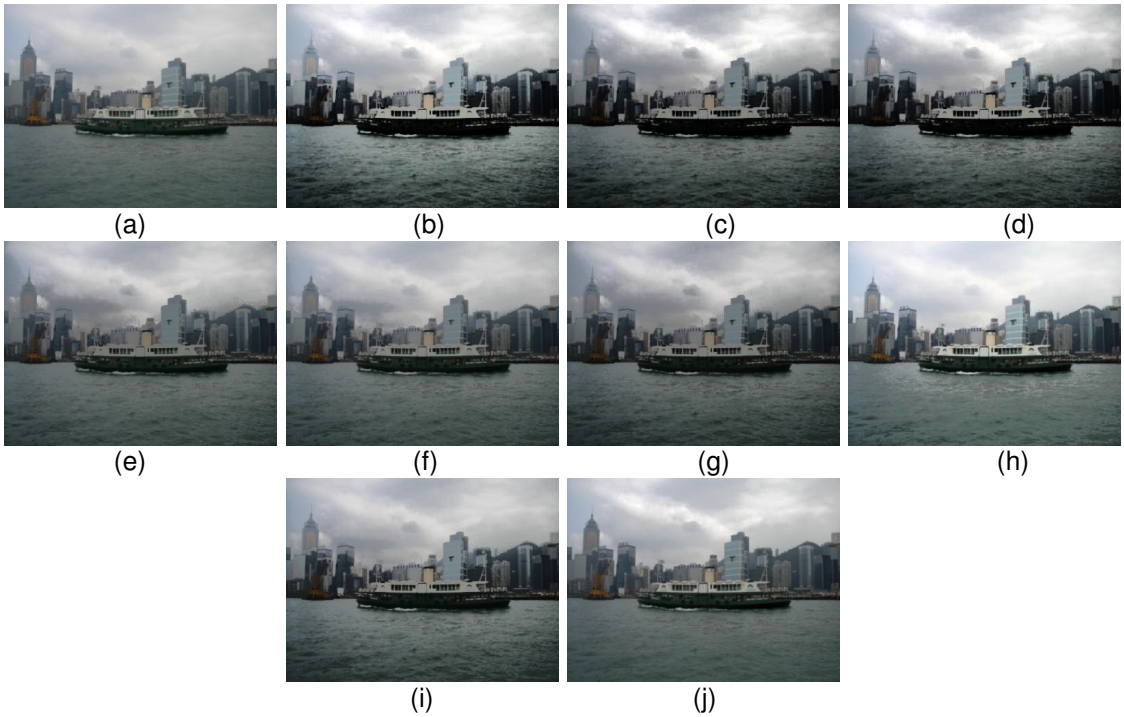


FIGURE 3: Sample experiment result - 2. Enhancement methods: (a) input image, (b) UNFHE, (c) BBHE, (d) DSIHE, (e) BHEPL, (f) BHEPLD, (g) AIEBHE, (h) DQHEPL, (i) MMSICHE, (j) DSBP.

4.2 Quantitative Evaluation

Statistics of the performance metrics were collected over the 220 test images and are displayed as box plots in Fig. 4 to illustrate the performance of the test methods.

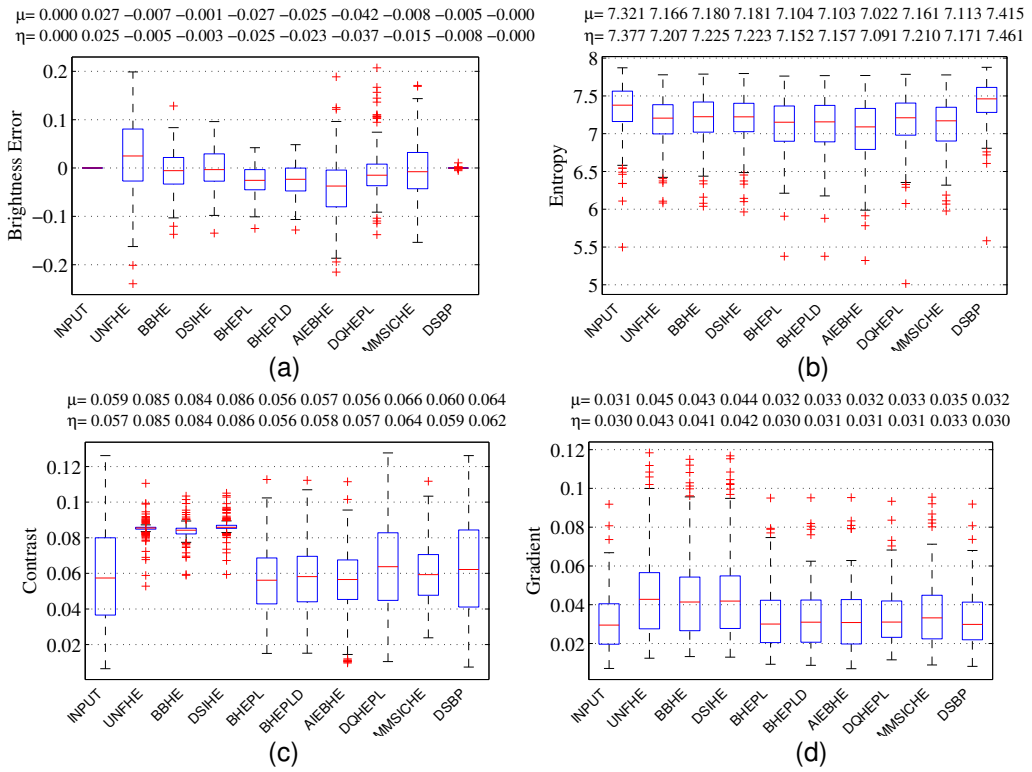


FIGURE 4: Box plots of result statistics for methods compared: (a) brightness error, (b) entropy, (c) contrast and (d) gradient[1].

Figure 4(a) shows the statistics of the brightness error between the input and the enhanced images from different methods. The input is the reference, hence has zero error, as shown in the left-most box. In the case of UNFHE method, the spread is largest as there is no mechanism to preserve the brightness. For the other methods, whether based on histogram equalization towards a uniform histogram or clipping of the histogram, there is significant brightness error with wide spreads. On the other hand, the proposed method produces the smallest spread with zero average brightness error among all the test methods.

Plots of information content, entropy, from the resultant images are illustrated in Fig. 4(b). It is shown that results from methods under comparison produce lower entropy values. This is due to the use of histogram equalization as the enhancement technique. With histogram equalization, the output image intensity may not cover all possible intensity levels, thus gives a reduction in entropy. On the other hand, the proposed DSBP approach, making use of dynamic range stretching, is able to occupy the permitted intensity levels and gives higher average entropy value of 7.415 against the input images of 7.321. Moreover, from the plot, it is clear that the first and third quartiles are close to each other, meaning that the entropy increment is consistent in the proposed method.

The performance metric in output image contrast is depicted in Fig. 4(c). The UNFHE, BBHE, DSIHE methods are expected to attain highest contrast as there is no unrealistic intensity limiting technique functioning in these methods. This often tends to produce undesirable viewing artifacts

[1] μ and η represent mean and median respectively.

as observed in the qualitative analysis. On the other hand, the smoothing based approaches are unsuccessful in preserving the brightness, though next higher level of contrast can be achieved with these methods. The output from the proposed method gives an average contrast of 0.064 as compared to the input at 0.059 which is better than the other methods by employing a marginal trade-off between brightness preservation and mitigation of artifacts.

Figure 4(d) shows the performance measure of image gradient. Similar to the case of contrast, approaches employing histogram equalization without clipping produce high gradient measures at the expense of unwanted artifacts. The methods that adopt histogram clipping reduces artifacts by compromising gradient to a lower measure. The weakness of the smoothing based methods is the limited capability in brightness preservation, though they give higher gradient metric. On the other hand, the proposed method not only preserves the brightness but also improves artifacts reduction while ensuring the increment in gradient to 0.032 against the input at 0.031.

In addition to the boxplot analysis, hypothetical investigation was carried out to evaluate the performance of the proposed DSBP method. The T-test was performed to determine if the brightness of the enhanced image from different methods is significantly different from that of the original image. For the analysis, significance level was set at 0.05 to test the null hypothesis, which is defined as the difference of brightness between the input and enhanced image has a mean value of zero. Here, $h = 1$ indicates the rejection of the null hypothesis at the mentioned significance level, whereas $h = 0$ indicates the opposite. The $p-$value (or probability value) is the probability of observing the outcome given the null hypothesis is true.

Method	**Mean** (μ)	h	$p-$ **value**
Input	0.000	N/A	N/A
UNFHE	0.027	1	0.000
BBHE	-0.007	0	0.316
DSIHE	-0.001	0	0.925
BHEPL	-0.027	1	0.000
BHEPLD	-0.025	1	0.001
AIBEHE	-0.042	1	0.000
DQHEPL	-0.008	0	0.336
MMSICHE	-0.005	0	0.370
DSBP	-0.000	0	0.997

TABLE 1: Comparison of brightness error.

From Table 1, it can be observed that for the UNFHE, BHEPL, BHEPLD, and AIBEHE methods, the null hypothesis is rejected, that is, the mean of the brightness error is significantly different from zero. Among the different methods that were failed to reject $h = 0$, the proposed DSBP has the maximum $p-$value of 0.997. Therefore it can be concluded that, for DSBP, there is a higher chance that the brightness is preserved in the output image.

In summary, the visual assessment and numerical evaluation demonstrate that the proposed method maintains the brightness level to avoid artifacts generation while ensuring sufficient contrast improvement, thus achieves a balanced trade-off between the two conflicting objectives in image processing.

5. CONCLUSION

An effective approach for image contrast enhancement with mean brightness preservation has been presented. Based on the principle of bi-histogram equalization, the integration of dynamic range stretching and the golden search process had produced enhanced images with accurate preservation of the output image mean brightness with respect to the input. Experiments carried out had verified that the proposed method outperforms other current methods in terms of absolute brightness error, information content, contrast and gradient.

6. REFERENCES

[1] Y.-J. Lin, C.-T. Shen, C.-C. Lin, and H.-C. Yen, "A fast image fusion algorithm for image stabilization on hand-held consumer electronics," in Consumer Electronics, 2009. ICCE '09. Digest of Technical Papers International Conference on, pp. 1–2, Jan 2009.

[2] F. Su, G. Fang, and N. M. Kwok, "Shadow removal using background reconstruction," in Image and Signal Processing (CISP), 2012 5th International Congress on, pp. 154–158, Oct 2012.

[3] T. Wu, H. Wu, Y. Du, N. Kwok, and Z. Peng, "Imaged wear debris separation for on-line monitoring using gray level and integrated morphological features," Wear, vol. 316, no. 1-2, pp. 19 – 29, 2014.

[4] G. Zhang, X. Jia, and N. Kwok, "Super pixel based remote sensing image classification with histogram descriptors on spectral and spatial data," in Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, pp. 4335–4338, July 2012.

[5] R. C. Gonzalez and R. E. Woods, Digital Image Processing (3rd Edition). Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

[6] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," Consumer Electronics, IEEE Transactions on, vol. 43, pp. 1–8, Feb 1997.

[7] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," Consumer Electronics, IEEE Transactions on, vol. 45, pp. 68–75, Feb 1999.

[8] C. H. Ooi, N. Kong, and H. Ibrahim, "Bi-histogram equalization with a plateau limit for digital image enhancement," Consumer Electronics, IEEE Transactions on, vol. 55, pp. 2072–2080, Nov 2009.

[9] C. H. Ooi and N. Isa, "Adaptive contrast enhancement methods with brightness preserving," Consumer Electronics, IEEE Transactions on, vol. 56, pp. 2543–2551, Nov 2010.

[10] J. R. Tang and N. A. M. Isa, "Adaptive image enhancement based on bi-histogram equalization with a clipping limit," Computers and Electrical Engineering, vol. 40, no. 8, pp. 86 – 103, 2014.

[11] S.-D. Chen and A. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," Consumer Electronics, IEEE Transactions on, vol. 49, pp. 1310–1319, Nov 2003.

[12] K. Singh and R. Kapoor, "Image enhancement via median-mean based sub-image-clipped histogram equalization," Optik - International Journal for Light and Electron Optics, vol. 125, no. 17, pp. 4646 – 4651, 2014.

[13] T. Celik and T. Tjahjadi, "Automatic image equalization and contrast enhancement using gaussian mixture modeling," Image Processing, IEEE Transactions on, vol. 21, no. 1, pp. 145–156, 2012.

[14] S.-W. Jung, "Image contrast enhancement using color and depth histograms," IEEE Signal Processing Letters, no. 21, pp. 382–385, 2014.

[15] T. K. Agarwal, M. Tiwari, and S. S. Lamba, "Modified histogram based contrast enhancement using homomorphic filtering for medical images," in Advance Computing Conference (IACC), 2014 IEEE International, pp. 964–968, IEEE, 2014.

[16] Y.-C. Chang, "N-dimension golden section search: Its variants and limitations," in Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on, pp. 1–6, IEEE, 2009.

INSTRUCTIONS TO CONTRIBUTORS

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

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

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

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 9, 2015, 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: 9 - **Issue: 5**

i. Submission Deadline : September 30, 2015 **ii. Author Notification:** October 31, 2015

iii. Issue Publication: November 2015

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

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