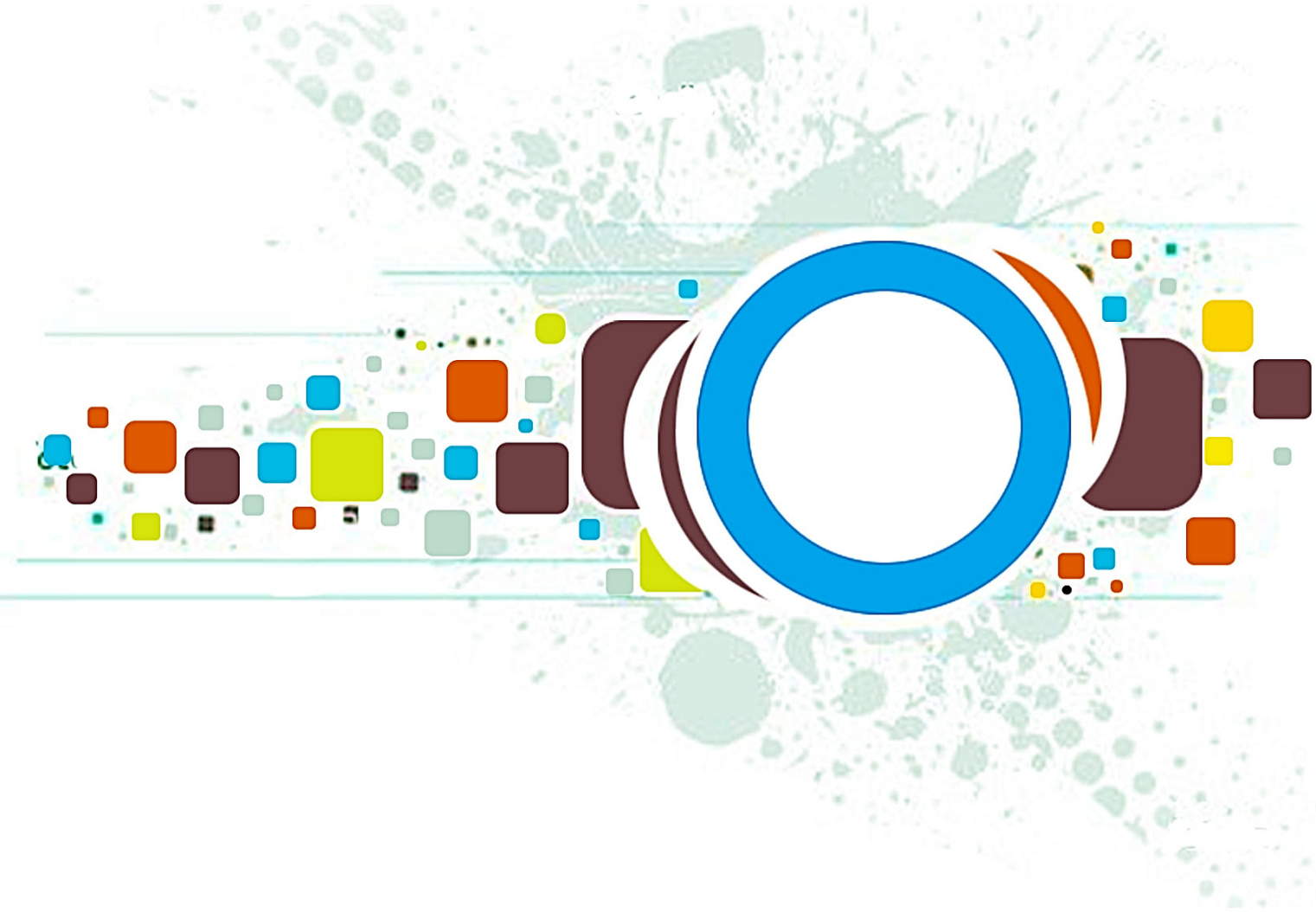
Volume 8 • Issue 4 • July 2014

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 8, ISSUE 4, 2014

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 8, Issue 4, July 2014
Publishing Date: 01-07-2014
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, 2014

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 *Eight* of IJIP. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJIP emphasizes on efficient and effective image technologies, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing. IJIP comprehensively cover the system, processing and application aspects of image processing. Some of the important topics are architecture of imaging and vision systems, chemical and spectral sensitization, coding and transmission, generation and display, image processing: coding analysis and recognition, photopolymers, visual inspection etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 8, 2014, IJIP appears in more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

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

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

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

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

Editorial Board Members
International Journal of Image Processing (IJIP)

EDITORIAL BOARD

TABLE OF CONTENTS

Volume 8, Issue 4, July 2014

Pages

Feature Extraction and Analysis on Xinjiang High Morbidity of Kazak Esophageal Cancer by Using Comprehensive Feature

Murat Hamit *murat.hamit@xjmu.edu.cn*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Fang Yang *1090466411@qq.com*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Abdugheni Kutluk *akutluk@hotmail.co.jp*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Chuanbo Yan *ycbsky@163.com*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Elzat Alip *elzat003@qq.com*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Weikang Yuan *454069946@qq.com*
College of Medical Engineering Technology
Xinjiang Medical University
Urumqi, 830011, China

Abstract

Image feature extraction technology has been widely applied in image data mining, pattern recognition and classification. Esophageal cancer is a common digestive malignant tumor, China is one of the world's highest incidence and mortality rates of esophageal cancer among the countries, Xinjiang Uygur Autonomous Region is a high incidence area of esophageal cancer, and the kazak is the esophageal cancer high-risk groups. In this paper, we selected 60 advanced esophageal X-ray barium images, half of them are constricted esophagus and the rest are ulcerous esophagus. Firstly, image preprocessing approaches were used to preprocess images. Secondly, extracting the gray-scale histogram features and the GLCM features of the images, then composing the two features into comprehensive feature. Finally, using Bayes discriminant analysis to verify the classification ability of the comprehensive feature. The classification accuracy for constricted esophagus was 86.7%, for ulcerous esophagus was 93.3%.

Keywords : Xinjiang High Morbidity of Kazak, Esophageal Cancer, Comprehensive Feature, Feature Extraction, Image Classification.

1. INTRODUCTION

Esophageal cancer is a common digestive malignant tumor and its worldwide morbidity and mortality in a common cancer were ranked in the sixth and eighth, respectively [1]. The geography difference on its pathogenesis is the most obviously among all cancers, China is one of the world's highest incidence and mortality rates of esophageal cancer among the countries[2], the world's annual increase of 300 thousand patients with esophageal cancer, about half occurred in China[3]. Esophageal cancer is one of the most common cancers in China, the incidence and mortality are above all cancers in the world, esophageal cancer has become a seriously threat to people's life and health, it has been proved to be one of the focus cancer of research [4]. Xinjiang Uygur Autonomous Region is a high incidence area of esophageal cancer, the kazak is the esophageal cancer high-risk groups, whose esophageal cancer mortality rate up to 155.9/106, higher than the average level 15.23/106 in China, so this disease is the regional focus malignant tumors to prevent [5].

CAD can assist the clinician to discover lesions and improve diagnostic accuracy through medical image processing technology as well as other possible physiological and biochemical methods, combining with computer analysis and calculation technology. Image feature extraction technology a cross discipline, has been widely applied in image data mining, pattern recognition and classification. It's not only included in the computer vision technology, but also involved in the image processing, the aim is to extract the image invariant features through computer analysis and solve practical problems [6]. The gray level histogram method and the GLCM method are usually used in the image feature extraction. There is no related research on image feature extraction of high morbidity of kazak esophageal cancer in Xinjiang. Therefore, this study means a lot to kazak esophageal cancer in Xinjiang.

2. METHODLOGY

In this paper, we selected two kinds of advanced esophageal X-ray barium images, which were acquired from First Affiliated Hospital, Xinjiang Medical University of China. Classifying them under the clinician's guidance, Selecting 60 images, half of them are constricted esophagus and the rest are ulcerous esophagus. For these selected images, firstly, image preprocessing approaches were used to preprocess images before the analysis algorithm was applied in order to keep the useful image information under different conditions, converting RGB image to the grayscale intensity image by eliminating the hue and saturation information while retaining the luminance, through median filter to remove the image noise and using histogram equalization to enhance the contrast; secondly, to extract the gray-scale histogram features and the GLCM features of the images, then composing the two features into comprehensive feature; Finally, using Bayes discriminant analysis to verify the classification ability of the comprehensive feature.

2.1 Image Preprocessing

Medical images obtained from the hospital not only contain valid information, but there are also some noise and the variation caused by the rotation and translation. So, the images cannot be used directly. The purpose of image preprocessing is to process the image obtained from the medical apparatus, removing the noise due to external disturbance, enhancing the contrast between the normal tissue and the pathological tissue, getting the interest region of strong visuality and contrast, thus providing the better support for the doctor's clinical diagnosis and the subsequent image processing[7]. In this paper, for the purpose of image feature extraction, the original image needs to be preprocessed to convert color, remove noise and enhance the contrast, reducing the influence of the factors mentioned above.

After the image preprocessing, the quality of X-ray barium image was improved significantly, the contrast of normal tissue and pathological tissue was enhanced apparently and the details of image were more clearly (see figure 1 and 2 lower right), which laying a good foundation for the subsequent image extraction of high morbidity of kazak esophageal cancer in Xinjiang, furthermore, improving the subsequent image classification accuracy.

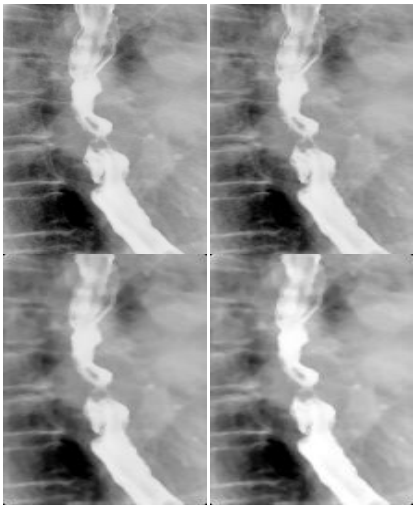


FIGURE1: Preprocessing results of constricted esophageal X-ray. Upper left: Original image; Upper right: Color transformed; Lower left: Removed noise; Lower right: Enhanced.

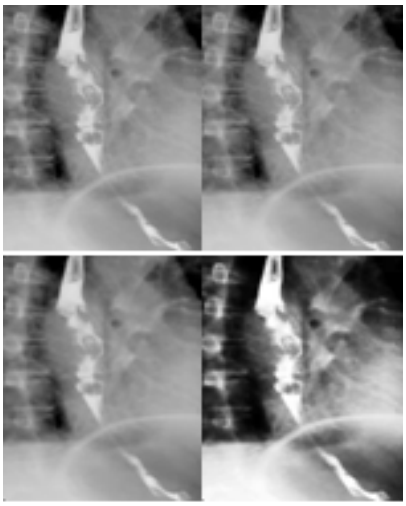


FIGURE 2: Preprocessing results of ulcerous esophageal X-ray. Upper left: Original image; Upper right: Color transformed; Lower left: Removed noise; Lower right: Enhanced.

2.2 Gray-scale Histogram Feature Extraction

A digital image histogram is a gray-scale discrete function, the following formula represents the image histogram definition [8].

$$H(i) = \frac{n_i}{N}, i = 0, 1, \cdots, L - 1$$

i represents gray level, L represents the number of gray level types, *ni* represents the number of *i*, *N* represents the total number of pixels in an image. The equation describes the percentage of the number of i account the total number of pixels in an image. The abscissa is the gray level, the vertical axis is the frequency of the gradation. Calculating the following statistics to reflect the image characteristic values [9] on the basis of the gray-scale histogram.

(1) Average value: Mean reflects the average gray value of an image.

$$u = \sum_{i=0}^{L-1} iH(i)$$

(2) Variance: Variance, which is a measure for the width of the histogram, reflects the discrete distribution of a gray-scale image numerically. That is, the difference between the gray level and the average.

$$\sigma^2 = \sum_{i=0}^{L-1} (i-\mu)^2 H(i)$$

(3) Skewness: Skewness reflects the degree of asymmetry in the histogram distribution, the greater skewness represents the histogram distribution is more asymmetric.

$$\mu_s = \frac{1}{\sigma^3} \sum_{i=0}^{L-1} (i-\mu)^3 H(i)$$

(4) Kurtosis: Kurtosis, which measures whether the distribution of gray-scale image is very focused on the average gray nearby, reflects the image gray-scale distribution when it closes to the mean. The smaller Kurtosis represents the histogram distribution is more concentrative.

$$\mu_k = \frac{1}{\sigma^4} \sum_{i=0}^{L-1} (i-\mu)^4 H(i)$$

(5) Energy: Energy reflects the uniform degree of gray-scale distribution, the more uniform gray-scale, the larger energy.

$$\mu_N = \sum_{i=0}^{L-1} H(i)^2$$

2.3 GLCM Feature Extraction

GLCM reflects the microtexture of an image area. It described the gray correlation of pixel pairs by a certain spatial relationship [10-11]. GLCM is defined as that starting at the gray level of pixel i from the image, calculating the probability of the pixel j from the distance d and the angleθ, reflecting the spatial correlation of gradation between two points in the image.

The space GLCM usually can not be used as texture analysis feature due to its complex computation, so we extracted the texture feature on the basis of the GLCM. Haralick [12] proposed 14 kinds of GLCM texture quantitative methods on the basis of the character of texture. We calculated the following statistic characters:

(1) Angular second moment: Also known as energy, reflecting the uniformity of gray, the rougher texture moments, the greater energy.

$$ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [P(i,j,d,\theta)]^2$$

(2) Entropy: Using entropy to detect the complexity of the image space and internal uniformity, the thinner texture, the greater entropy.

$$ENT = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P(i,j,d,\theta)\log P(i,j,d,\theta)$$

(3) Inertia Moment: Also known as Contrast, which represents the total change of gray in a small image area. This parameter reflects the degree that high value matrixes away from the diagonal line.

$$CON = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2 P(i,j,d,\theta)$$

(4) Correlation: Correlation described the similarity of GLCM between rows and columns, which reflects the extend length of a certain gray value along a certain direction, the longer extension, the greater correlation value.

$$COR = \frac{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} i\cdot j\cdot P(i,j,d,\theta) - \mu_x\mu_y}{\sigma_x\sigma_y}$$

μ_x represents the mean of gray value, μ_y represents the smooth mean,
σ_x^2 represents the grayscale variance, σ_y^2 represents the smooth variance.

$$\mu_x = \sum_{i=0}^{L-1} i\sum_{j=0}^{L-1} P(i,j,d,\theta)$$

$$\mu_y = \sum_{i=0}^{L-1} j\sum_{j=0}^{L-1} P(i,j,d,\theta)$$

$$\sigma_x^2 = \sum_{i=0}^{L-1}(i-\mu_x)^2\sum_{j=0}^{L-1} P(i,j,d,\theta)$$

$$\sigma_y^2 = \sum_{i=0}^{L-1}(i-\mu_y)^2\sum_{j=0}^{L-1} P(i,j,d,\theta)$$

(5) Inverse difference moment: Reflecting the concentration of high value matrixes on the main diagonal line, the greater value, the higher concentration.

$$IDM = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{P(i,j,d,\theta)}{[1+(i-j)^2]}$$

3. RESULTS AND DISCUSSION

3.1 Gray-scale Histogram Feature Extraction Result

Extracting the average vaule (avg), variance (var), skewness (sk), kurtosis (kur), energy (ene) of all the X-ray images selected, and composing them into vector component(see table 1).

Image type	avg	var	sk	kur	ene
constricted esophagus	147.8329761	1501.590717	-0.82923649	-0.20165276	0.01113704
	141.8383579	2327.279287	-0.91476189	0.38049406	0.00748258
	145.4243755	4301.323428	-0.33738664	-0.5308574	0.00683868
	…	…	…	…	…
	159.1537156	2545.407865	-0.27509711	-1.37506984	0.00992652
ulcerous esophagus	151.0811004	732.2059987	-0.50390497	1.49034815	0.01210474
	153.7212397	778.7894774	-0.2101274	-0.18883096	0.01077037
	158.4681136	780.5992869	-0.60767563	0.45453657	0.01107275
	…	…	…	…	…
	140.3558579	2590.974543	-0.68579449	-0.10677265	0.00725628

TABLE 1: Two kinds of image features extracted by gray-scale histogram method.

3.2 GLCM Feature Extraction Result

Extracting the GLCM features of all the images selected, electing the Angular Second Moment, the Entropy, the Inertia Moment, Correlation and Inverse difference moment from 4 directions of pixel distance, d = 1, θ = {0 °, 45 °, 90 °, 135 °}, calculating the average value and variance of each feature at four directions and 10 features were obtained, then composing them into vector component (see table 2).

Image type	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
constricted esophagus	0.0386	0.0035	3.6072	0.1279	0.9848	0.4915	0.0468	0.0006	0.8676	0.0267
	0.0325	0.0042	3.7719	0.1494	0.8085	0.2943	0.0465	0.0002	0.8200	0.0357
	0.0304	0.0030	3.8440	0.1210	0.9388	0.3269	0.0462	0.0003	0.8039	0.0295
	…	…	…	…	…	…				
	0.0351	0.0056	3.6790	0.1975	0.7982	0.3763	0.0465	0.0002	0.8424	0.0429
ulcerous esophagus	0.0257	0.0043	4.0761	0.1926	1.8209	0.6707	0.0450	0.0008	0.7459	0.0483
	0.0272	0.0041	3.9627	0.1707	1.2833	0.474	0.0456	0.0006	0.7672	0.0436
	0.0342	0.0034	3.7483	0.1347	1.1431	0.5102	0.0463	0.0005	0.8326	0.0305
	…	…	…	…	…	…				
	0.0245	0.0047	4.1009	0.2099	1.6412	0.6626	0.0452	0.0007	0.7247	0.0570

TABLE 2: Two kinds of image features extracted by GLCM method.

3.3 Bayes Discriminant Analysis

Discriminant analysis is a statistical method, which can estimate the discriminate objects' category according to their observation results. It has been widely used in the medical filed. Robin Hanson used Bayes analysis method of unsupervised models' classification in different complexity.

Using the gray-scale histogram features and GLCM features extracted to classify the images through Bayes discriminant analysis. The classification accuracy for constricted esophagus was 86.7%, for ulcerous esophagus was 93.3% (see table 3).

Image Type	Bayes Discriminant Classification		total	accuracuy rate(%)
	correct	error		
Constricted Esophagus	26	4	30	86.7
Ulcerous Esophagus	2	28	30	93.3

TABLE 3: Discriminant analysis between two kinds of image by comprehensive features.

4. CONCLUSION

According to the difference between histogram distribution and texture distribution in different types of high morbidity of kazak esophageal cancer in Xinjiang, combining with the characteristics of esophageal cancer, we proposed the feature extraction based on histogram and GLCM method, combining the extracted gray-scale histogram features and GLCM features into the comprehensive feature, and then evaluated the feature's classification ability through Bayes discriminant analysis. Experimental results show that using comprehensive feature for image classification has a high accuracy, and feature classification ability is different when classifying different images, which provides a new direction for the research of computer aided diagnosis system for the high incidence of kazak esophageal cancer in Xinjiang Uygur Autonomous Region.

5. ABBREVIATIONS

CAD: computer aided diagnosis; GLCM: gray level co-occurrence matrix

6. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China, Xinjiang, 30960097 and 81160182.

7. REFERENCES

[1] Parkin DM, Bray FI, Devesa SS. "The Cancer burden in the year 2000. The global picture". Eur J Cancer, vol.37, pp.S4 - S66, Aug. 2001.

[2] ZHAO Fengjuan, YUN Miaoying, ZHANG Yan, "XU Yu.Research progress on Xinjiang kazak esophageal cancer". Journal of central university for nationalities, vol.18, pp.85-89, Mar. 2009.

[3] CHEN Xiangchuan, PANG LiJuan, LI Feng. "Research progress on Xinjiang kazak esophageal cancer". Agricultural reclamation of medicine, vol.28, pp.384-387, May. 2006.

[4] ZHANG Huixia, CHEN Yan, YIN Dong, DENG Yanchao, MA Yanqing, JuLaiDi. "Discussion on risk factor of Xinjiang kazak esophageal cancer". Modern preventive medicine, vol.36, pp.1804-1806, Oct. 2009.

[5] GUO Hui, DING Jianbing, ZHANG Wei, ZHANG Tong: "Gene research progress on Xinjiang kazak esophageal cancer". Basic medicine and clinical, vol.30, pp.428-430, Apr. 2010.

[6] WANG Zhirui, YAN Cailiang: "Review on image feature extraction method". journal of jishou university, vol.30, pp.52-56, May. 2011.

[7] Murat Hamit, ZHOU Jingjing, YAN ChuanBo, LI Li, CHEN Jianjun, HU Yanting, KONG Dewei. "Feature extraction and analysis of Xinjiang local liver hydatid CT images based on gray-scale histogram". Tech review, vol.30, pp.79-83, May. 2010.

[8] JIN Hua. "The research of segmentation and feature extraction method for medical images based on density clustering". M.A. thesis.: Jiangsu university, Zhenjiang 2005.

[9] WANG Shuqin. “Research on feature selection and extraction of Liver CT aided diagnosis system”. M.A. thesis, Shanghai jiaotong university, Shanghai 2010.

[10] TONG Longzheng, WANG Lei, CHEN Hairong etc. “Gray Level Co-occurrence Matrix Analysis on liver fibrosis images”. capital medical university, vol.24, pp.240-242, Mar. 2003.

[11] JIN Jing, SHI Li. Intelligent diagnosis of hepatocellular carcinoma based on the CT images[J]. system engineering research of China and clinical rehabilitation, vol.13, pp.5919-5922, Mar. 2009.

[12] Haralick R M. Statistical and structural approaches to texture. Proceedings of the IEEE, vol.67, pp.786-840, May. 1979.

An Ultrasound Image Despeckling Approach Based on Principle Component Analysis

Jawad F. Al-Asad *jalasad@pmu.edu.sa*
Department of Electrical Engineering
Prince Mohammad Bin Fahd University
Al Khobar, P.O. Box 1664, 31952, Saudi Arabia

Ali M. Reza *reza@uwm.edu*
Department of Electrical Engineering and Computer Science
University of Wisconsin-Milwaukee
Milwaukee, United States of America

Udomchai Techavipoo *udomchai@gmail.com*
National Electronics and Computer Technology Center
Pathumthani, Thailand

Abstract

An approach based on principle component analysis (PCA) to filter out multiplicative noise from ultrasound images is presented in this paper. An image with speckle noise is segmented into small dyadic lengths, depending on the original size of the image, and the global covariance matrix is found. A projection matrix is then formed by selecting the maximum eigenvectors of the global covariance matrix. This projection matrix is used to filter speckle noise by projecting each segment into the signal subspace. The approach is based on the assumption that the signal and noise are independent and that the signal subspace is spanned by a subset of few principal eigenvectors. When applied on simulated and real ultrasound images, the proposed approach has outperformed some popular nonlinear denoising techniques such as 2D wavelets, 2D total variation filtering, and 2D anisotropic diffusion filtering in terms of edge preservation and maximum cleaning of speckle noise. It has also showed lower sensitivity to outliers resulting from the log transformation of the multiplicative noise.

Keywords: Covariance Matrix, Denoising, Despeckling, Principle Component Analysis, Ultrasound Imaging.

1. INTRODUCTION

Ultrasound medical imaging is considered to be cost efficient and practically harmless to the human body. However, the quality of medical ultrasound images is degraded by the presence of speckle noise. Images acquired by pulse-echo ultrasound systems demonstrate the interaction between the incident pressure field and the spatial inhomogeneities of the medium [1]. The majority of such spatial inhomogeneities in soft tissue is highly concentrated with the dimensions much smaller than the wavelength of the incident pressure waves and can be modeled as diffuse scatterers, which radiate the incident acoustic energy in all directions. The large concentration of small scattering targets with sub-wavelength dimensions gives rise to a characteristic pseudo-random granular texture in the envelope-detected image known as speckle. Unlike other kinds of noise, speckle noise is not strictly random but is rather an intrinsic feature of soft tissue. Multiple scans of the same region that are taken in the same position with the same probe and under the same conditions will, in the absence of electrical noise, yield exactly the same speckle pattern each time [1].

Despite the negative effect on the quality of ultrasound images, speckle also carries clinically

important textural information that can be useful for tissue identification [2]. Therefore, methods for image restoration aimed at improving the diagnostic utility of ultrasound images need to enhance image resolution and improve image clarity while preserving the textural information present in the speckle.

In literature, various speckle noise reduction methods have been proposed [3]–[12]. Averaging filters and adaptive weighted median filters [5] are simple and effective noise reduction methods. By introducing weight coefficients to the well known median filter and adjusting the smoothing characteristics of the filter, it is possible to effectively suppress noise. However, such schemes seem to remove fine details being actually filters with a low-pass characteristic [6].

Wavelet denoising is applied for despeckling in medical ultrasound imaging [2], [4], [7], [11], [12]. It is based on multiscale decompositions. It consists of three main steps. First, the noisy signals are analyzed using a wavelet transform. Then the empirical wavelet coefficients are shrunk. Finally, denoised signals are synthesized from these shrunk coefficients through the inverse wavelet transform. These methods are generally referred to as wavelet shrinkage techniques [1]. It is found to be the best approach among many denoising methods in synthetic aperture radar (SAR) images [3]. In [4], logarithmic transformation is used to convert the multiplicative noise to an additive noise prior to wavelet denoising. These are referred to as the homomorphic wavelet despeckling (HWDS) methods [4]. In [11], a non-Gausian statistical model with an adaptive smoothing parameter is used in the wavelet transformed domain. It was shown that the HWDS does not improve the signal-to-noise ratio (SNR) [8] because the wavelet transformed speckle coefficients are larger than the threshold value, thus not suitable for removing the speckle noise in ultrasound images. Moreover, the log transformed multiplicative noise is spiky in nature, following Fisher-Tippett distribution [6]. HWDS tends to preserve such spikes unless it is properly treated. Two preprocessing stages have been proposed to solve this problem [6]. The first stage is to decorrelate the speckle noise samples. This stage requires point spread function estimation from the ultrasound envelope image. The second stage applies the median filter to normalize the distribution of the log transformed spiky noise. These stages improve the noise distribution to be more Gaussian that is suitable for most wavelet denoising schemes. In [7], adaptive decorrelation is used for further image enhancement.

Total variation filtering (TVF) [10] and anisotropic diffusion filtering (ADF) [9] are also used for despeckling. Similar results can be obtained via replacing the wavelet denoising step proposed in [6] by TVF and ADF. TVF is useful for recovering constant signals and it is considered to be among the most successful methods for image restoration and edge enhancement. It is mainly because of its capability of filtering out the noise without blurring the most universal and crucial features of image edges. Adversely, ADF takes advantage of the locality and anisotropy diffusion. It is capable of smoothing images with a decreased blur of the boundaries between their homogenous regions. ADF was shown to perform well for images corrupted by additive noise. However, in cases where images contain speckle noise, ADF enhances that noise instead of eliminating it [13].

Principle component analysis (PCA) is applied to reduce speckle noise in SAR images [14], [15] PCA is a result from linear algebra [16]. It arises from the eigen decomposition of the covariance matrix of the signals. It is used for dimensionality reduction and compression of multidimensional data. Unlike many despeckling techniques that trade spatial information for noise reduction, using PCA reduces speckle noise with minimal loss of spatial resolution. The first few principal components hold significant spatial information, while higher order components are dominated by speckle [14].

In this paper, a PCA-based approach for despeckling medical ultrasound images is explored and analyzed in detail. The approach is first introduced in [17]. PCA is applied through overlapped

segments of the ultrasound images. After log transformation, a global covariance matrix is formed by averaging the corresponding covariance matrices of the segments. A projection matrix is then calculated by selecting few principal eigenvectors of the global covariance matrix. This projection matrix is used to filter the speckle noise from the ultrasound images. The PCA-based approach is applied to simulated and real ultrasound images and the despeckling results are compared to those from HWDS, TVF, and ADF approaches in terms of the resolution, edges, and the signal to speckle noise ratio.

2. ULTRASOUND SIGNAL MODEL

A generalized model of a speckle noisy image is given by [18]

$$g(n,m) = f(n,m)u(n,m) + \xi(n,m) \qquad (1)$$

where g, f, u, and ξ stand for the observed envelope image, original image, and multiplicative and additive components of the speckle noise, respectively. The n and m respectively denote the axial and lateral indices of the image samples, or alternatively, the angular and radial indices for the sector images. This model has been successfully used both in ultrasound and SAR imaging. When applied to ultrasound images, this model can be simplified by disregarding the additive noise term. This leads to the following model:

$$g(n,m) \approx f(n,m)u(n,m) \qquad (2)$$

Consequently, adopting (2) as the basic model, it is assumed that the image g(n, m) is observed before the system processing is applied [6].

2.1 Proposed Algorithm

In the implementation of the proposed PCA-based approach to despeckle ultrasound envelope images, subtracting the mean before despeckling and adding it afterward is avoided. Through experimentations, it shows that subtracting the mean can be more beneficial for additive noise rather than a log transformed multiplicative noise. The non-linear logarithmic transformation for conversion of the multiplicative noise into additive noise is done globally at the beginning rather than for each segment (block). This method is found to have a positive impact on the quality of the denoised images compared to the method of transforming segments individually.

The following algorithm summarizes the steps implementing the concept of PCA-based approach in despeckling ultrasound envelope images. Assuming that the size of the ultrasound envelope images is $n \times m$, where n and m are the numbers of pixels in the axial and lateral directions, respectively.

Step 1: Apply logarithmic transformation to the given image.

Step 2: Segment the image into overlapping segments s_i, each of $q \times p$ size, where i is the index of the segment; and q and p are the numbers of pixels in the axial and lateral directions, respectively. The segmentation can be done laterally or axially across the 2D image. Segment s_{i+1} is shifted by one pixel from segment s_i.

Step 3: Reshape s_i into a column vector v_i of $q.p \times 1$ size and find a covariance matrix c_i of the vector v_i, such that $c_i = v_i \times v_i^T$, where T denotes transposition.

Step 4: Sum all c_i and average them by the number of segments of the image to get the overall covariance matrix (R) for the whole image. This method of averaging all covariance matrices of the segments serves to represent various speckle developments within the image, thus yielding a better estimation of the global (overall) covariance matrix.

Step 5: Calculate the eignevalues and eigenvectors of R. Select $k < q.p$ eigenvectors that correspond to the set of largest eigenvalues. The remaining $q.p - k$ eigenvectors are simply rejected or ignored. The largest eigenvalues and the corresponding eigenvectors contribute to the true signal in the presence of noise.

Step 6: Form a feature matrix F by sorting the k selected eigenvectors according to their corresponding eigenvalues in the descending order. The size of the feature matrix F is now $q.p \times k$. This sorting procedure enables the use of the eigenvectors that correspond to the true signal.

Step 7: Calculate the transformation or the global projection matrix $P = F \times F^T$.

Step 8: Calculate a denoised vector d_i by projecting v_i onto the projection matrix P, i.e., $d_i = P \times v_i$.

Step 9: Reshape the denoised vector d_i back to a segment of the size of $q \times p$.

Step10: Reconstruct the envelope image by averaging all the overlapping denoised segments according to their original locations and the number of pixels superimposed on each coordinate, and take the exponential.

2.2 Image Quality Measures

To assess the denoising capabilities, we apply five image quality measures: α, β, SNR, S-SNR, and PSNR. These measures are defined as follows. Let I_{org} be the noise free original image and I_{est} be the estimated image or the image after denoising. First, the measure α is the ratio of the number of pixels of the estimated image's autocorrelation function ($R_{Iest\ Iest}(\tau_n, \tau_m)$, where τ_n and τ_m are the lags in n and m indices, respectively) that exceeds 75% of its maximum value to the total number of pixels. This α is mostly used to evaluate the resolution in ultrasound imaging [6]. Lower α usually implies better image resolution. Second, the measure β is used to assess the ability of the despeckling methods to preserve sharp details of the images. It is given by [19] as

$$\beta = E\left\{ \frac{\left\langle \Delta I_{org}, \Delta I_{est} \right\rangle}{\left\| \Delta I_{org} \right\|_F \left\| \Delta I_{est} \right\|_F} \right\} \qquad (3)$$

where Δ is the Laplacians operator; $\langle \cdot , \cdot \rangle$ denotes the standard inner product; E{ . } is the expectation; and || . ||F is the Frobenius matrix norm, i.e., $\|I(n,m)\|_F = \sqrt{\sum_n \sum_m |a_{nm}|^2}$. The β closer to 1 indicates the better despeckling ability in preserving image edges. Third, the SNR is the conventional signal to noise ratio which is defined as $E\left\{ \|I_{org}\|_F / \|I_{org} - I_{est}\|_F \right\}$. Fourth, the S-SNR is the speckle signal to noise ratio. It is defined as the mean to the standard deviation of the estimated image I_{est}. S-SNR is an indicator of the degree of contamination of the image by the speckle noise. Finally, the PSNR is the peak signal to noise ratio. It is the most commonly used measure of quality in image denoising. The PSNR is given by

$$PSNR = E\left\{ \max |I_{org}|^2 / \|I_{org} - I_{est}\|_F \right\} \qquad (4)$$

For simulations, the β, SNR, S-SNR, and PSNR are used. Due to the absence of the original image in case of real ultrasound data, only α and S-SNR can be applied. The SNR, S-SNR, and PSNR can be reported in dB unit by taking 10 $\log_{10}$(.) of their values.

3. RESULTS, CRITICAL DISCUSSION AND COMPARATIVE EVALUATION

3.1 Simulated Phantom Image

The Field II Program [20], [21] is used to simulate ultrasound envelope images in this paper. Although the Field II program does not account for nonlinear propagation of the ultrasound wave with higher harmonics, which in return improve the quality of the image, the simulated image from the program is used to compare the despeckling approaches. To test denoising efficiency, sharp edges and curves are included within the simulated image as shown in Fig. 1. The image also includes low, medium, and high scattering areas, as represented by dark, white, and background areas, respectively. We simulate 128 radiofrequency (RF) echo signals using 5625 scatterers per cm^2 with a lateral resolution of 0.156 mm. This number of scatterers is chosen in accordance with the simulations made in [20], [22]. The envelopes of the RF echo signals are obtained by taking the absolute value of the Hilbert transformation of the RF echo signals [23]. All of the envelopes are rearranged side-by-side to form a 2D envelope image. The envelope image is decimated in the axial direction from 1024 to 256 to yield a 256×128 image, as shown in Fig. 1-a. The corresponding speckle noisy version, shown in Fig. 1-b, is created by corrupting the undecimated envelope image by noise according to equation 2.

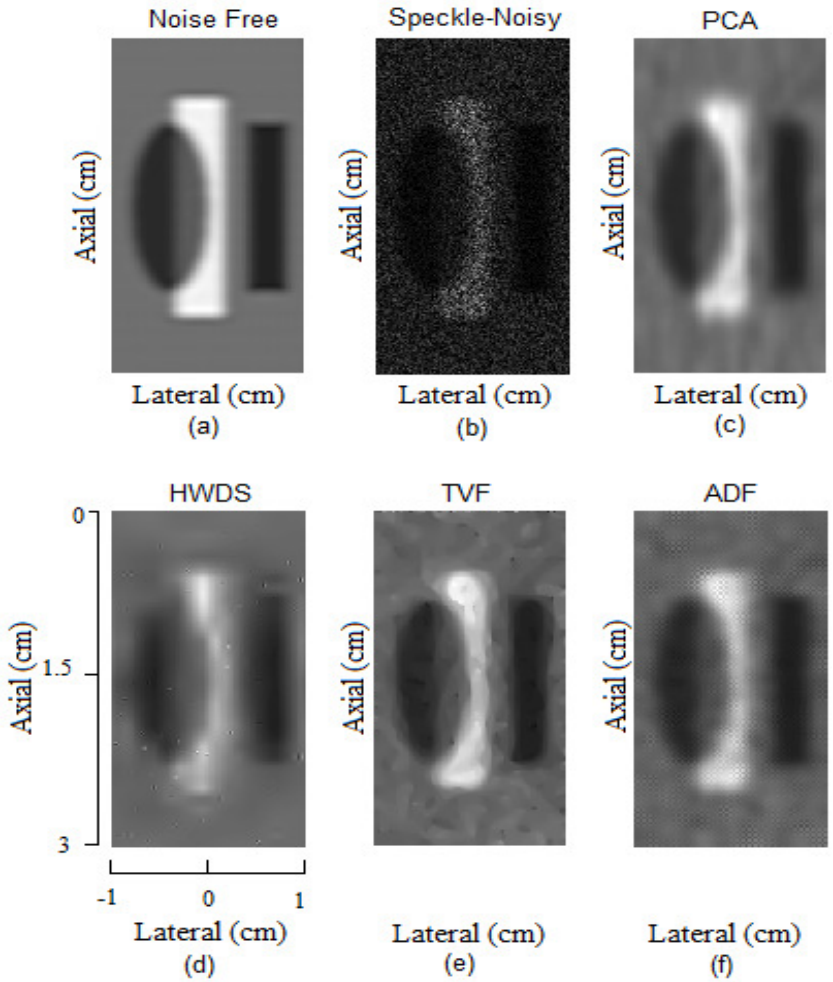


FIGURE 1: Simulated phantom images: a noise-free envelope image (a), a speckle noisy image (b), and the images after despeckling using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. Note that the horizontal and vertical axes represent the axial and lateral axes, respectively (the images are transposed).

The PCA-based approach is compared to three nonlinear despeckling schemes, HWDS, TVF, and ADF. For HWDS, the WaveLab® package (Department of Statistics, Stanford University) is

used. HWDS is applied with four-level wavelet decomposition and Daubechies with four vanishing moments. Since the objective in this paper is not to compare various thresholding schemes, the original approach in [24] is used. The noise variance is estimated by assuming that most empirical wavelet coefficients at the finest level of decomposition are induced by the noise. TVF is also used to include the best empirically adjusted parameters for yielding the best visual results and for avoiding image blurring. The regularization parameter (λ) controls how much smoothing is performed, large noise levels call for large λ. TVF uses 100 iterations and λ = 400. ADF uses 35 iterations and a conduction coefficient of 25. Speed of diffusion controlled by λ is set to the maximum (λ = 0.25). The option of favoring wide regions over smaller ones is chosen for the ADF. The proposed PCA-based approach is used with a segment size of 16 $\times$ 8 and one eigenvector or written in short as (16/8-1). The images after being despeckled using PCA, HWDS, TVF, and ADF methods are shown in Fig. 1-c to Fig. 1-f, respectively. According to the plots in Fig. 1, the edge preservation and maximum cleaning of the speckle noise can be seen through the proposed approach as compared to the other denoising schemes.

The plots in Fig. 2 compare the performance measures averaged over 100 independent trials. As can be seen, the proposed PCA-based approach provides higher β, SNR, and PSNR than the other methods. However, the S-SNR from the proposed approach comes in the second place after that from the HWDS.

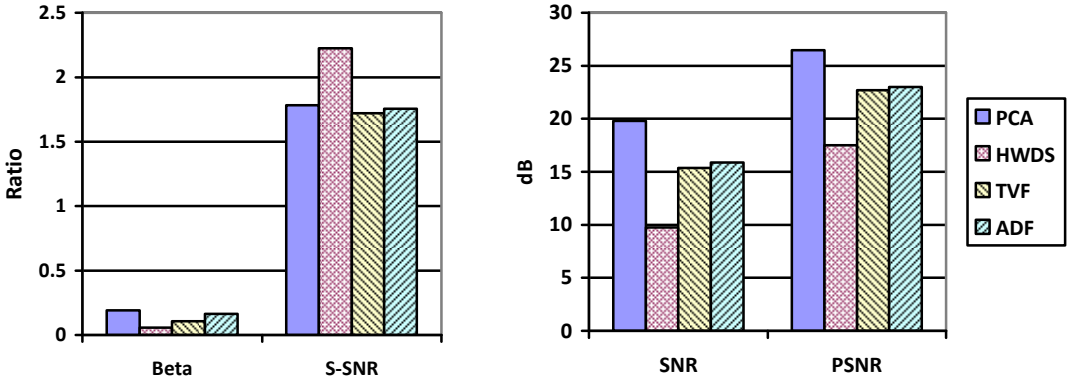


FIGURE 2: Performance comparison of denoising schemes when applied to the simulated phantom image. Vertical axis represents the magnitude in dB for SNR and PSNR while the magnitude is of no unit for Beta and S-SNR.

The comparisons of the axial and lateral profiles are shown in Fig. 3 and Fig. 4, respectively. These profiles are taken from the centers of the images in Fig. 1. The profiles plotted in Fig. 3 and Fig. 4 are from the original (a) and noisy (b) images, and the images after despeckled using the PCA (c), HWDS (d), TVF (e), and ADF (f) methods. The profile from the original image in (a) is also plotted as dotted lines on the same axes of the profiles from the despeckled images in (c) to (f). It is clear from these figures that our PCA-based approach has provided the closest profiles to the original ones. In agreement with the results provided in (Yu and Acton 2002), Fig. 3-f and Fig.4-f also show that the ADF method is less efficient in removing the speckle noise. Notice a remaining spike in the profile around the 70th pixel in Fig. 3-d, a drop-off in the profile around the 120th pixel in Fig. 3-e, and noisy profiles in Fig. 3-f and Fig. 4-f, compared to a smooth profile that follows the original profile closely in Fig. 3-c and Fig. 4-c. From these figures, it implies that the PCA-based approach is insensitive to spikes. The insensitivity of the PCA-based approach to the log transformed spiky noise is also observed by comparing Fig. 1-c with Fig. 1-d.

To further investigate the number of principle components used in our PCA-based approach, the parameters (16/8-1) are changed to (16/8-2), (16/8-3), and (16/8-4). Fig. 5 displays the results. It

shows that the higher the number of eigenvectors, the more the speckle noise becoming dominant.

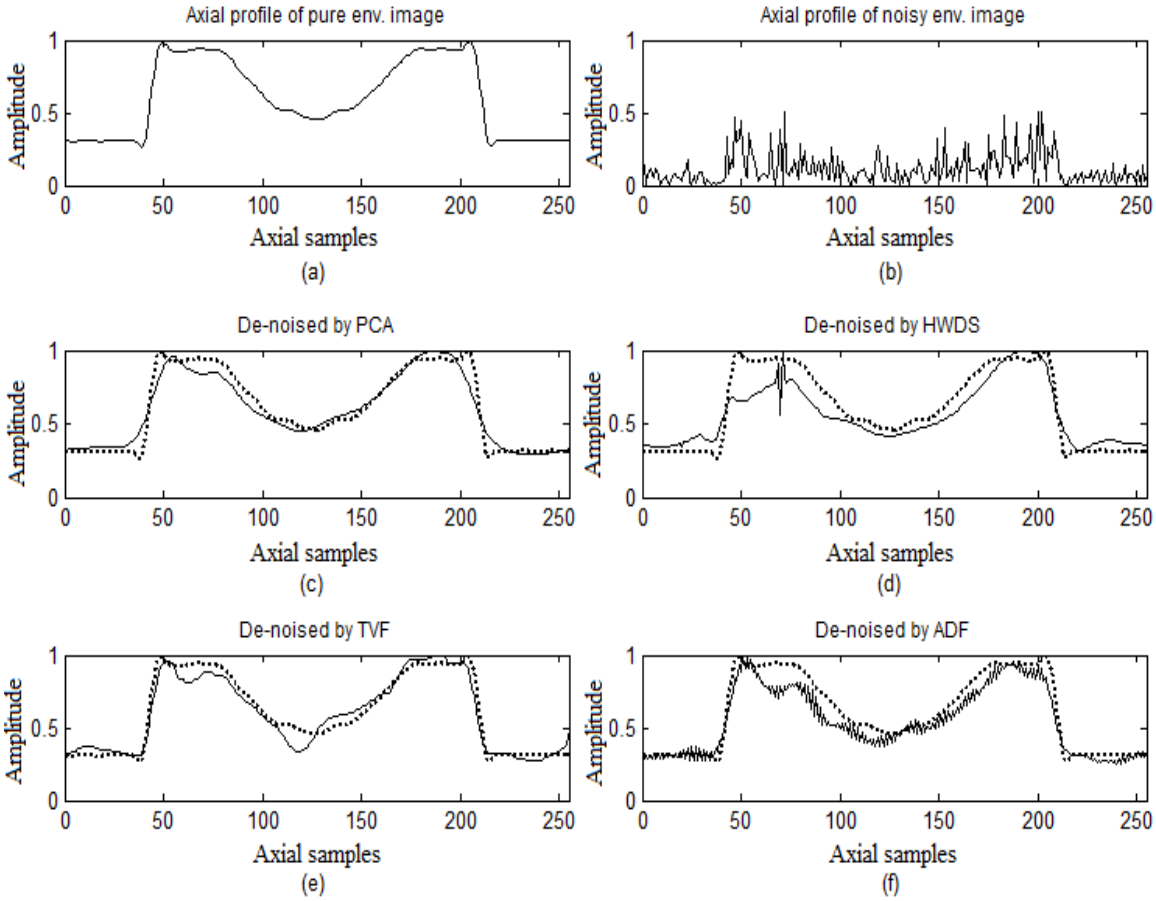


FIGURE 3: Comparison of the axial profiles of the simulated phantom images taken from the centers of the original (a) and noisy (b) images and from the centers of the despeckled images using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. The dotted lines in (c) to (f) are the profile in (a).

3.2 Tissue Mimicked Image

For this simulation, the amplitude profile of the scatterers is weighted by a white Gaussian noise field in order to mimic the tissue reflectivity function of the scatterers. The difference between this simulation and the previous one is that the multiplicative noise is interred into the image to mimic the reflectivity function rather than interred to corrupt the undecimated pure 2D envelope image. In addition, a single tumor is also simulated as shown by the circular white region at the center of the image of Fig. 6. For this simulation, the Field II Program with a linear array transducer of 3.5 MHz center frequency and 100 MHz sampling frequency was used. The scatterer map consisted of 3750 scatterers per cm^2. We generate 64 scan lines with a axial resolution of 1024 samples per line and a lateral resolution of 0.312 mm. Each scan line is decimated from 1024 to 128 to yield a 128×64 image, as shown in Fig. 6. Note that the size of the image is different from that in the previous simulation. HWDS is used with 8 Daubechies vanishing moments and 8 decomposition levels. For the TVF, λ is set to 400 and the number of iterations is also set to 300. The number of iteration above 300 would blur the image. For the ADF, λ is set to the maximum of 0.25 with the number of iterations set to 50, and the conduction coefficient is set to 30. The option of favoring wide regions over small ones is chosen for the ADF. The PCA-based proposed approach is used with (16/8-1), denoting a segment size of 16×8 with one eigenvector.

The noise free phantom image before weighted by the white Gaussian noise and the speckle noisy version after weighted are shown in Fig. 6-a and b, respectively. Note that to display the image in the correct spatial aspect ratio, the images are interpolated from 128×64 to 128×256. Fig. 6 also demonstrates a visual comparison between the denoising techniques. As can be seen from Fig. 6-d, the image after despeckled using the HWDS method still contains speckle noise.

The result from the TVF method in Fig. 6-e shows that the size of the inclusion seems to be smaller than the true size, especially in the axial direction (the vertical axis). In contrast to the result from the ADF method in Fig. 6-f, the size of the inclusion seems to be larger than the true size, especially in the lateral direction (the horizontal axis). Furthermore, the background of the despeckled image from the ADF method contain high and elongated hills along the lateral direction that usually connected to the inclusion.

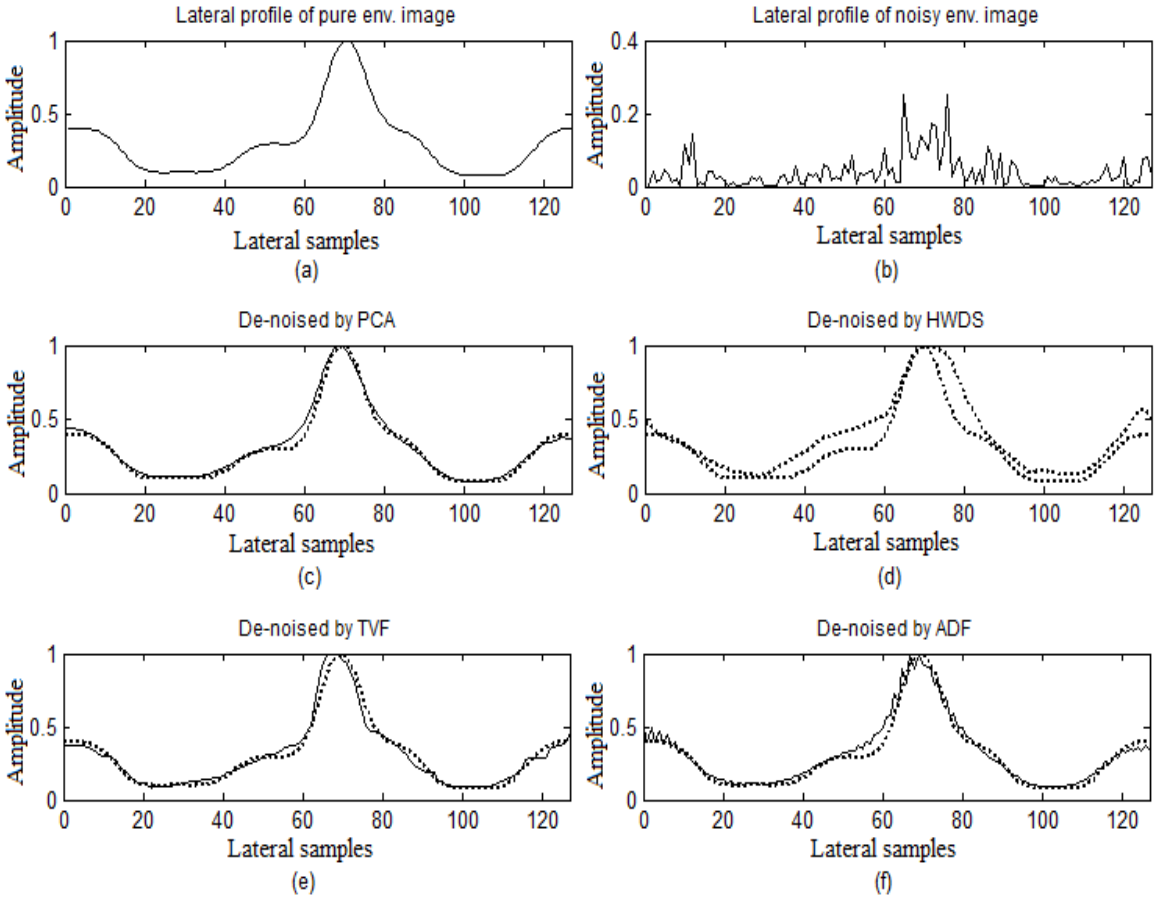


FIGURE 4: Comparison of the lateral profiles of the simulated phantom images taken from the centers of the original (a) and noisy (b) images and from the centers of the despeckled images using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. The dotted lines in (c) to (f) are the profile in (a).

Fig. 7 plots a performance comparison between the denoising techniques. The proposed PCA method provides higher β, SNR, and PSNR than other methods. However, the S-SNR from the proposed approach comes in the third place after that from the TVF and the ADF methods. In agreement with the results obtained for the simulated phantom image, through the plots in Fig. 6-c and Fig. 7, the edge preservation and maximum cleaning of speckle noise obtained can be seen through the proposed PCA-based approach as compared to the other denoising schemes.

The comparisons of the axial and lateral profiles are also shown in Fig. 8 and Fig. 9, respectively. These profiles are taken from the centers of the original and noisy images in Fig. 6-a and b and

from the centers of the despeckled images using PCA, HWDS, TVF, and ADF methods in Fig. 6-c to d. The profile from the original image in (a) is also plotted as dotted lines on the same axes of the profiles from the despeckled images in (c) to (f). In general, the PCA-based approach has provided closer profiles to the original ones except for the lateral profile provided by TVF in Fig. 9-e where the lateral length of the simulated tumor is better preserved. However, in Fig. 8-c the PCA has preserved the axial length of the simulated tumor better than the TVF in Fig. 8-e. In agreement with the results provided in [8], Fig. 8-d and Fig. 9-d also show that HWDS is less efficient in removing the speckle noise. While the ADF performance looks close to the PCA performance in Fig. 8 and in Fig. 9, the calculation of the mean square error of the corresponding profiles has showed that PCA profiles are closer to the original ones than the ADF profiles.

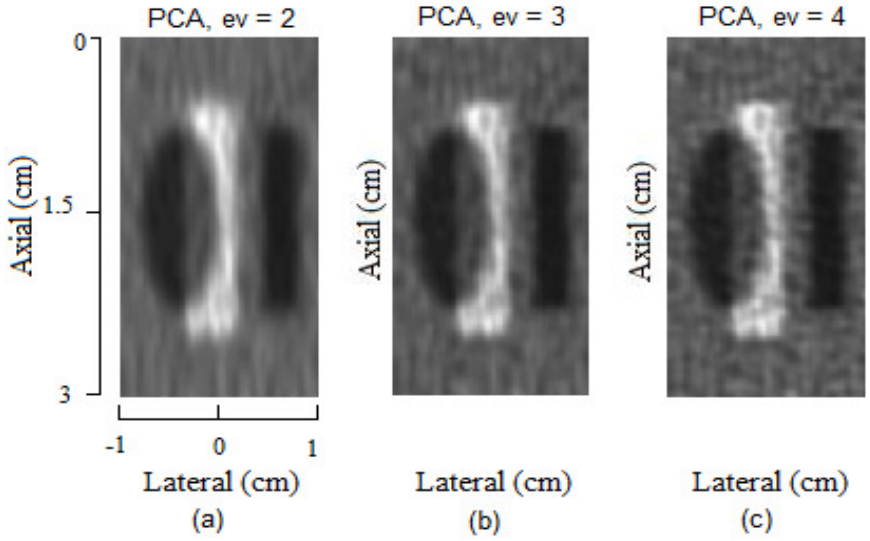


FIGURE 5: Comparison of the despeckled images using the PCA method with different numbers of principle components: ev = 2 (a), 3 (b), and 4 (c). The segment size is 16 × 8. Note that the horizontal and vertical axes represent the axial and lateral axes, respectively (the images are transposed).

The effective number of the principle components used in our PCA method is also investigated for the tissue mimicked image. The number of principle eigenvectors is changed to 2, 3, and 4 eigenvectors. The size of the segment is 16 × 8, similar to the previous experiment. The images after despeckling are shown in Fig. 10. In agreement with the results obtained for the simulated phantom image, the higher the number of eigenvectors the more dominant the speckle noise becomes. More spatial information is provided by the fewer number of eigenvectors.

To show the effect of our PCA-based approach parameters (*q*/*p*-*ev*) on the image quality measures (α, β, S-SNR, SNR, and PSNR), the simulated tissue mimicking image is despeckled with every possible combination of the parameters (*q*/*p*-*ev*), where *q* = {8, 16, 32}, *p* = {8, 16, 32}, and *ev* = {1, 2, 3}. The image size of 128 × 64 as used in the previous experiment is chosen.

Table I lists the PCA-based approach parameters versus the image quality measures. The ranks are also shown in the parentheses following the image quality measures. The summations of these ranks are also calculated and put in the last column with their rank in ascending order in the parentheses. Better α values (or lower number), which indicate higher resolution, seem to come from the parameters that have smaller segment sizes and more principle eigenvectors used (as shown in the second column of Table 1).

For β values, where a larger β indicates a better edge preservation of the denoising method, the segment size, both *q* and *p* parameters, seems to affect this quality measure. Larger segment sizes in the axial direction seem to provide better β. In contrast, smaller segment sizes in the

lateral direction seem to provide better β. This can be seen from the results of the parameters (32/8-1) that rank 4th in β from 27 sets of the parameters and the parameters (8/32-1) that rank 23rd. Using one principle eigenvectors seems to provide better β if the segment size is set as recommended above. This can be seen from the results of the parameters (32/8-1), (32/8-2), and (32/8-3) have rank 4th, 12th, and 9th.

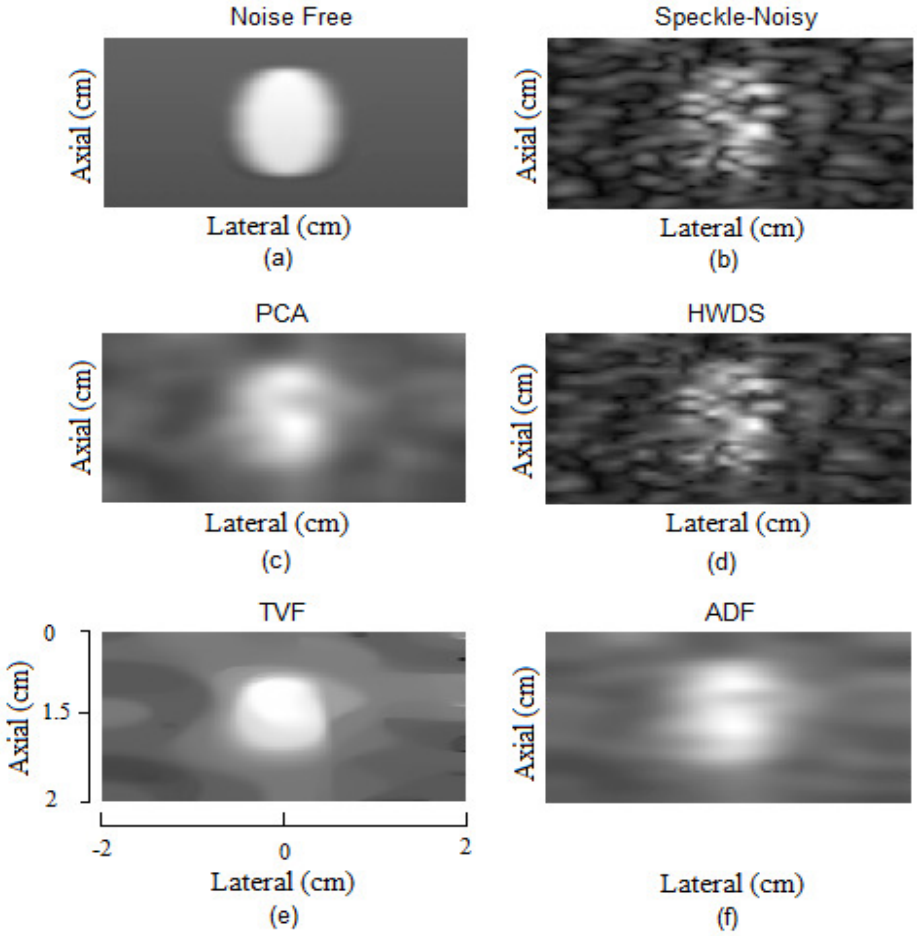


FIGURE 6: Tissue mimicking images: a noise-free envelope image (a), a speckle noisy image (b), and the images after despeckling using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. Note that the horizontal and vertical axes represent the lateral and axial axes, respectively, and their representations are different from those in Fig. 1.

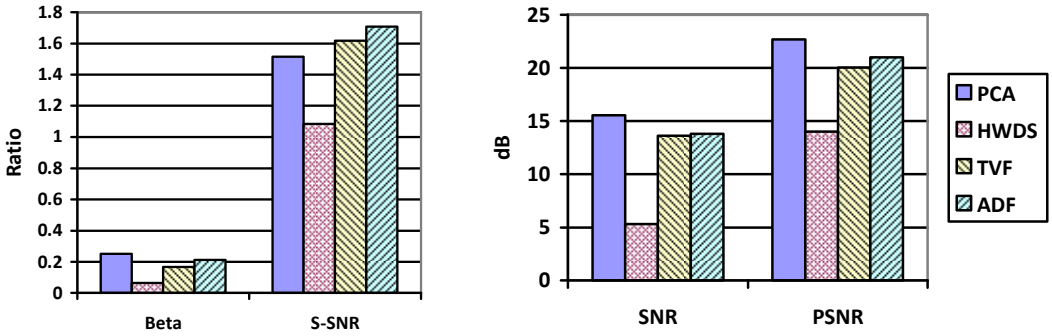


FIGURE 7: Performance comparison of denoising schemes when applied to the tissue mimicked image. Vertical axis represents the magnitude in dB for SNR and PSNR while the magnitude is of no unit for Beta and S-SNR.

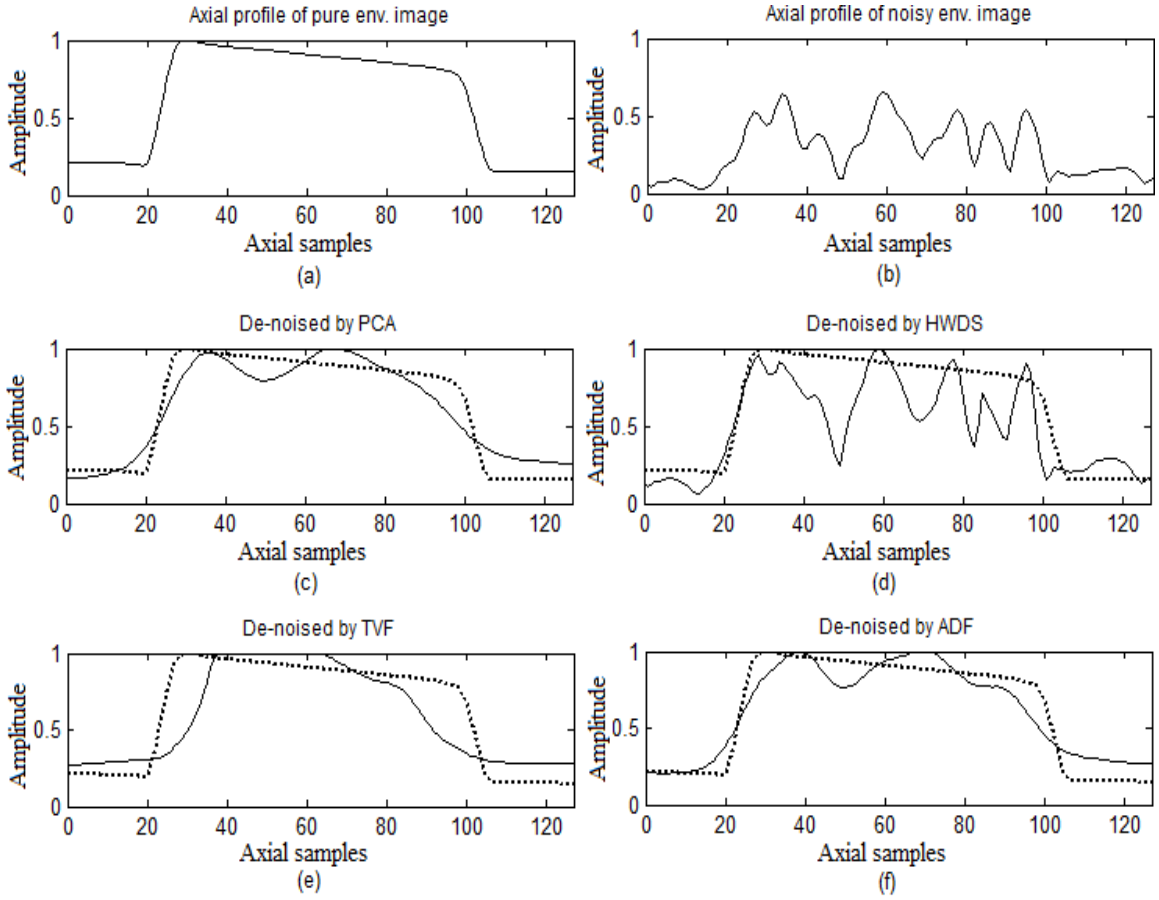


FIGURE 8: Comparison of the axial profiles of the tissue mimicking images taken from the centers of the original (a) and noisy (b) images and from the centers of the despeckled images using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. The dotted lines in (c) to (f) are the profile in (a).

For S-SNR, larger segment sizes and smaller number of principle components used seem to provide better S-SNR values, as can be seen from the results of the parameters (32/32-1), (32/32-2), and (32/32-3), that rank 1st, 7th, and 11th, respectively.

For SNR and PSNR, the effects of the despeckling parameters on these image quality measures seem to be the same. The ranks 1st, 2nd, and 3rd for each measure are from the same parameters (16/32-3), (16/16-2), and (16/8-1), respectively. The large segment sizes in the lateral direction need more principle eigenvectors to obtain better SNR and PSNR. The segment size of 16 pixels in the axial direction ($q = 16$) seem to optimize these image quality measures.

The tissue mimicking images after despeckling using the proposed PCA method with some sets of parameters in Table I are shown in Fig. 11. As shown, the despeckled images from the parameter (16/32-3), (16/16-2), and (16/8-1), in the middle row of the images in Fig. 11 are the best in cleaning the speckles and also in preserving the edge of the inclusion.

3.3 Experimental Results

Real RF data is obtained for the experimental scanner RASMUS (Department of Electrical Engineering, Technical University of Denmark) -- also described in [22] -- at the time of peak systole of a carotid artery for a healthy 30 year old male. The imaging protocol used to collect the human data was approved by the Ethics committee on Biomedical Experiments for Copenhagen and the subject provided informed consent. The transducer is a B-K 8812 linear array transducer with 6.2 MHz linear array probe, 40 MHz sampling frequency and 5 MHz center frequency. The RF data are composed of 64 RF lines with 1024 samples per line. For displaying in a correct spatial aspect ratio, the image is decimated in the axial direction and is interpolated in the lateral direction to yield a 256×256 image size as shown in Fig. 12.

In order to investigate the denoising efficiency of the proposed approach, as well as to show the effect of decimation in removing parts of the noise and consequently its effect on the performance of the denoising schemes, 3 different image sizes of 512×64, 256×64, and 128×64, resulting from decimation by factors of 2, 4, and 8 in the axial directions, are despeckled using the PCA, HWDS, TVF, and ADF methods. Their parameters are shown in Table II. The visual performance of these despeckling methods is shown in Fig. 13. Note that for the displaying purpose, the image lateral dimensions are interpolated to match their axial dimensions, i.e., 512×512, 256×256, and 128×128 image sizes. The plots in Fig. 14 show the numerical performance of these despeckling methods.

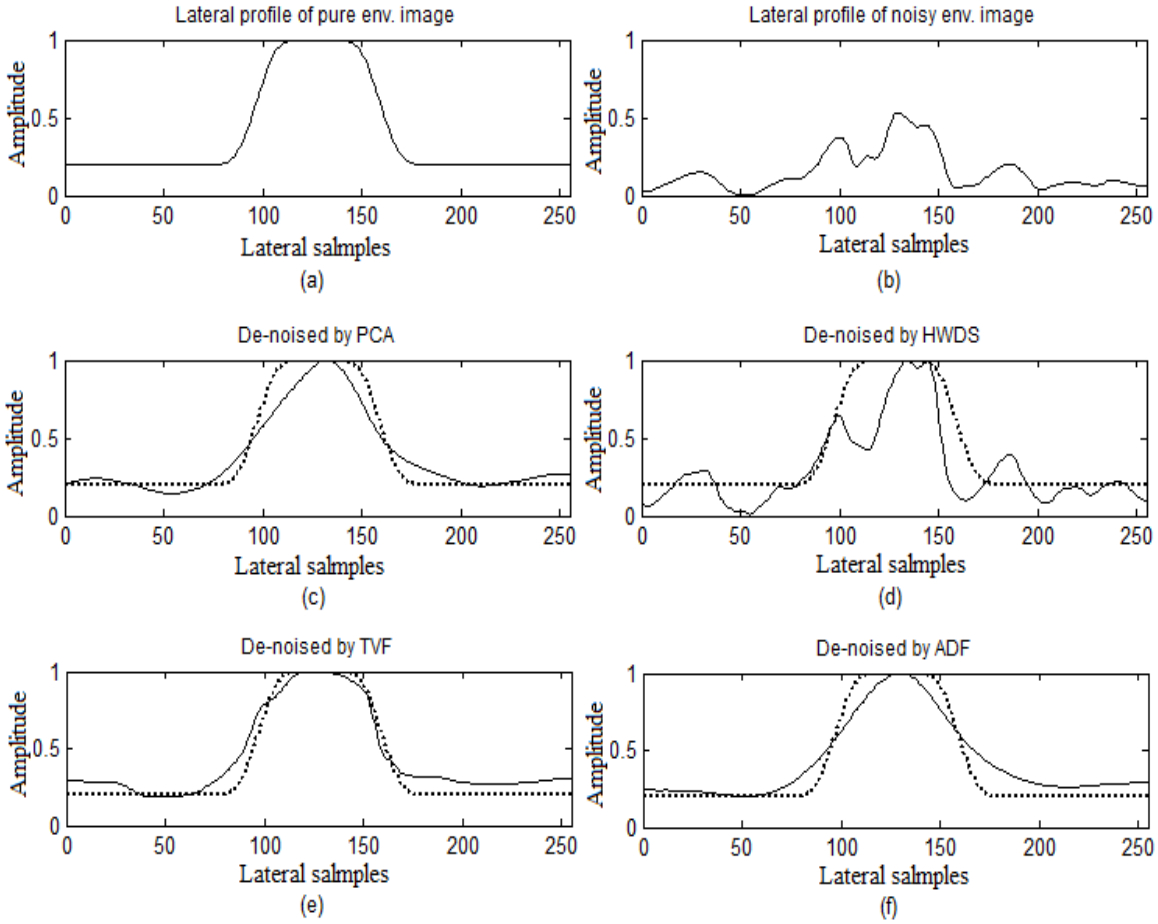


FIGURE 9: Comparison of the lateral profiles of the tissue mimicking images taken from the centers of the original (a) and noisy (b) images and from the centers of the despeckled images using PCA (c), HWDS (d), TVF (e), and ADF (f) methods. The dotted lines in (c) to (f) are the profile in (a).

Through these different decimation examples, it can be clearly observed how the decimation of the ultrasound image has a direct impact on the performance of the denoising scheme. Effect of decimation on the performance of the HWDS, TVF and ADF methods can be visually compared in Fig. 13. It can also be noticed in Fig. 14 in terms of the increasing values of α while increasing the decimation factors from 2 to 8. This implies that higher decimation factors generally reduce the despeckled image resolution. In contrast, the S-SNR values are very similar for these image sizes, except for the HWDS that the S-SNR is gradually increasing. Therefore, the decimation factor does not impact the S-SNR of the images despeckled using the PCA, TVF, and ADF method; however, higher decimation factors help the HWDS method to improve the S-SNR results.

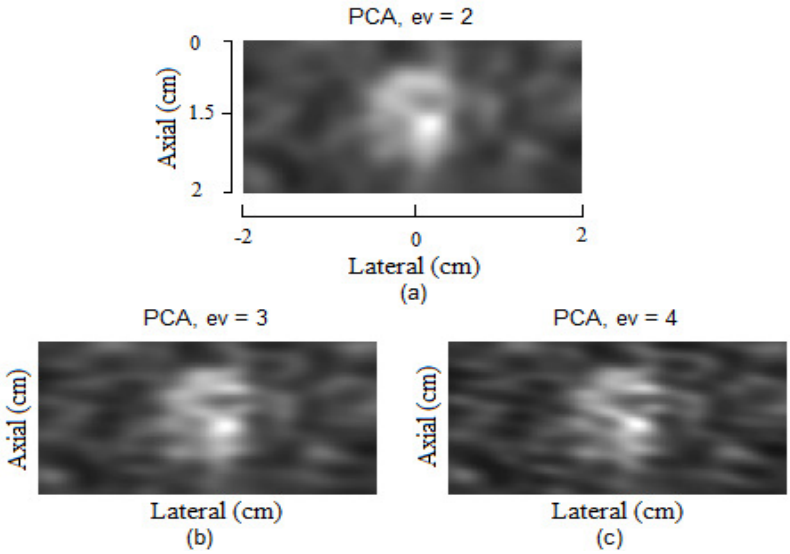


FIGURE 10: Comparison of the despeckled images using the PCA method with different numbers of principle components: *ev* = 2 (a), 3 (b), and 4 (c). The segment size is 16 × 8. Note that the horizontal and vertical axes represent the lateral and axial axes, respectively.

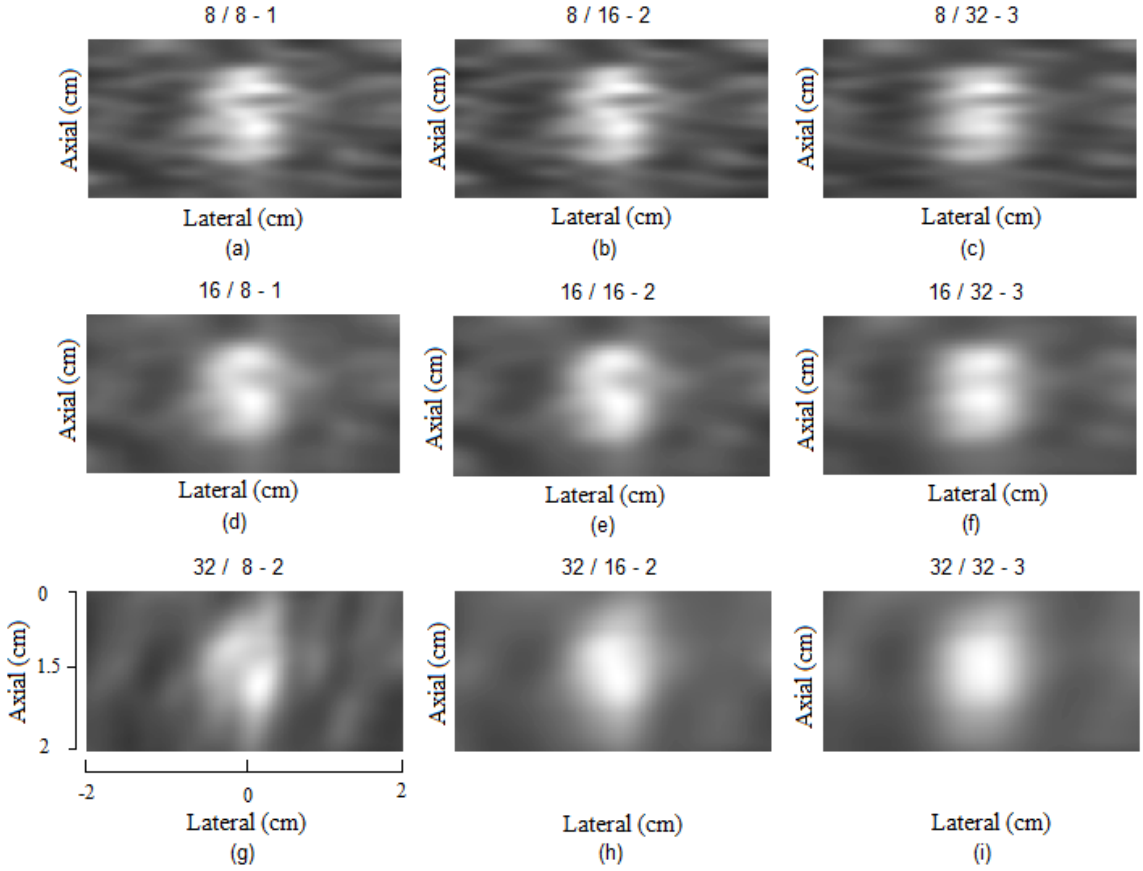


FIGURE 11: The denoising performance of the PCA-based approach for different parameters (*q*/*p*-*ev*), where *q* is the segment size in the axial direction, *p* is the segment size in the lateral direction, and *ev* is the number of the principle eigenvectors used.

q/p-ev	α	β	S-SNR	SNR	PSNR	Sum Rank
8/8-1	0.3400 (10)	0.1091 (19)	1.4339 (16)	13.7227 (6)	21.0744 (5)	56 (9)
8/8-2	0.0831 (2)	0.1230 (18)	1.1923 (26)	6.8194 (22)	15.4807 (22)	90 (22)
8/8-3	0.0587 (1)	0.0939 (22)	1.1385 (27)	6.0404 (24)	14.7488 (24)	98 (25)
8/16-1	0.7478 (22)	0.1055 (20)	1.9680 (6)	10.4405 (15)	18.3838 (15)	78 (17)
8/16-2	0.2902 (8)	0.1306 (16)	1.3468 (20)	13.2858 (10)	20.6813 (9)	63 (12)
8/16-3	0.1217 (4)	0.1353 (15)	1.2243 (24)	8.5016 (19)	16.8057 (19)	81 (18)
8/32-1	0.6240 (16)	0.0554 (23)	2.9884 (3)	0.5456 (27)	8.7126 (25)	94 (24)
8/32-2	0.9285 (24)	0.0970 (21)	1.6538 (10)	10.9291 (13)	18.8886 (14)	82 (19)
8/32-3	0.4368 (13)	0.1967 (7)	1.3598 (19)	12.5929 (11)	20.4564 (10)	60 (10)
16/8-1	0.5542 (15)	0.2504 (6)	1.5137 (13)	15.5510 (3)	22.6982 (3)	40 (1)
16/8-2	0.1613 (5)	0.1674 (11)	1.2872 (22)	9.1980 (17)	17.5949 (17)	72 (15)
16/8-3	0.1207 (3)	0.1245 (17)	1.2114 (25)	7.5921 (20)	16.1847 (20)	85 (20)
16/16-1	0.9581 (25)	0.3465 (2)	2.0599 (5)	8.5137 (18)	17.0704 (18)	68 (14)
16/16-2	0.5110 (14)	0.2670 (5)	1.4173 (18)	16.2386 (2)	23.0961 (2)	41 (3)
16/16-3	0.2233 (6)	0.1800 (10)	1.3099 (21)	11.0315 (12)	19.0029 (12)	61 (11)
16/32-1	0.6902 (21)	0.0000 (25)	3.1351 (2)	1.6217 (26)	7.6727 (26)	100 (26)
16/32-2	1.0000 (26)	0.0448 (24)	1.7101 (8)	9.2392 (16)	17.7663 (16)	90 (22)
16/32-3	0.6472 (18)	0.3054 (3)	1.4263 (17)	16.5988 (1)	23.3503 (1)	40 (1)
32/8-1	0.6589 (19)	0.2673 (4)	1.6698 (9)	13.4351 (9)	20.4327 (11)	52 (5)
32/8-2	0.2911 (9)	0.1569 (12)	1.4372 (15)	13.4930 (8)	20.8241 (8)	52 (5)
32/8-3	0.2272 (7)	0.1879 (9)	1.2816 (23)	10.7855 (14)	18.8960 (13)	66 (13)
32/16-1	0.9260 (23)	0.6001 (1)	2.2816 (4)	6.7772 (23)	15.1996 (23)	74 (16)
32/16-2	0.6294 (17)	0.1419 (14)	1.5588 (12)	13.8496 (5)	21.1800 (4)	52 (5)
32/16-3	0.3795 (12)	0.1507 (13)	1.4617 (14)	13.6787 (7)	21.0017 (7)	53 (8)
32/32-1	0.3737 (11)	0.0000 (25)	3.8110 (1)	1.8763 (25)	7.3329 (27)	89 (21)
32/32-2	1.0000 (26)	0.0000 (25)	1.8563 (7)	7.5074 (21)	15.8430 (21)	100 (26)
32/32-3	0.6812 (20)	0.1948 (8)	1.5710 (11)	14.0827 (4)	21.0244 (6)	49 (4)

TABLE 1: Performance of PCA-based approach when applied to the simulated tissue mimicked image. Note: The numbers in the parentheses are the rank in the descending order, except for the first and the last column in the ascending order.

The increasing value of α can also imply less speckle noise remained after despeckling. This is clearly seen by comparing the despeckled images using the HWDS method in the 2nd column of Fig. 13 and their α values on the bar graphs in Fig. 14. As can be seen in the 2nd column of Fig. 13, the speckle noise is diminishing from the images while the α values are increasing. Therefore, higher α values could imply lower image resolution and less speckle noise remaining after despeckling.

From Fig. 13 and Fig. 14, the results from the PCA method can be compared to those from other methods in terms of image visual quality, α values, and S-SNR values. For the image visual quality, the results from the PCA method are comparable to those from the TVF and ADF methods. The TVF method seems to provide better edges of the carotid artery for the images of 512×64 and 256×64 image sizes. However, some speckle artifacts are left on the top border of the artery on the image of 128×64 image size (as shown in the 3rd row and the 3rd column of Fig. 13).

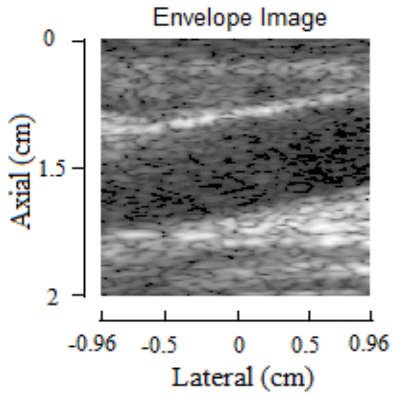


FIGURE 12: A carotid artery image before despeckling. Note that the horizontal and vertical axes are the lateral and axial directions, respectively.

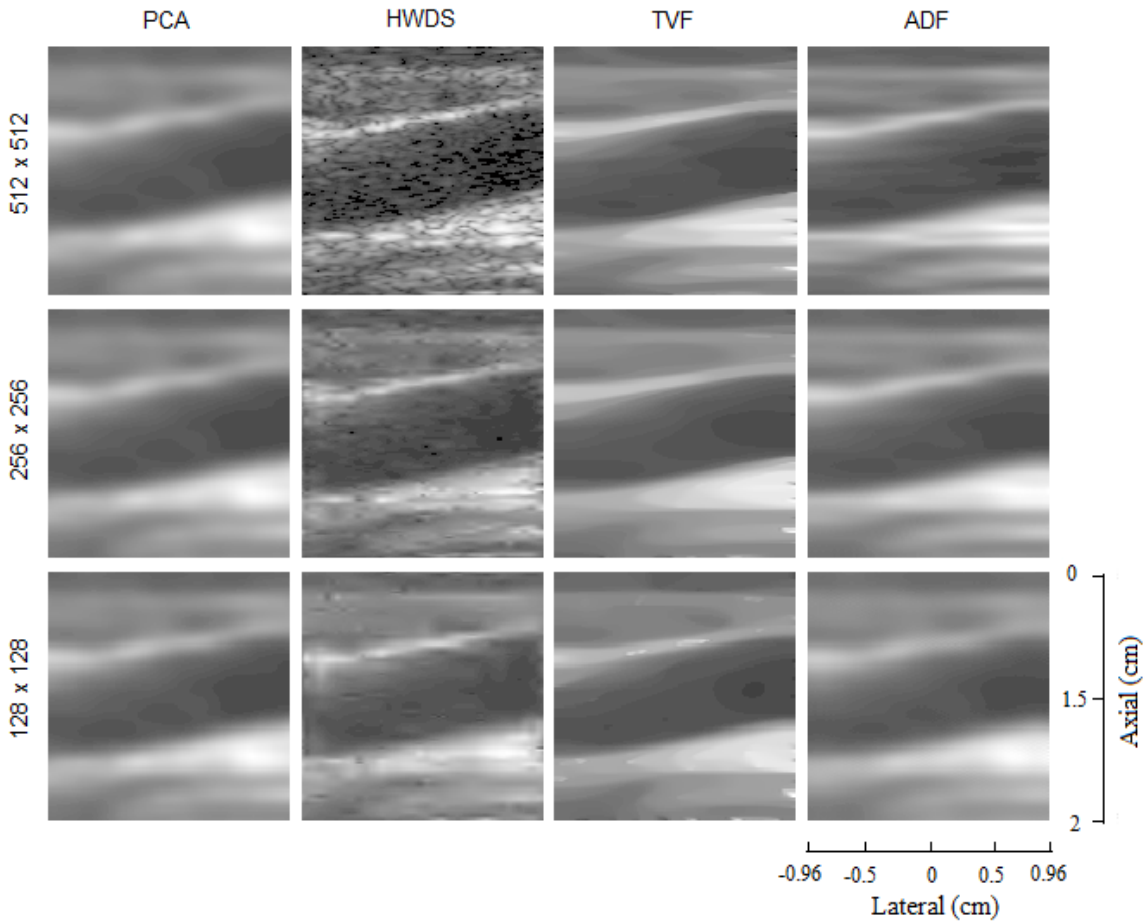


FIGURE 13: Comparison of despeckled carotid artery images using the PCA, HWDS, TVF, and ADF methods (respectively shown in 1st to 4th columns) applied on the images after decimation in the axial dimension to the sizes of 512×64, 256×64 and 128×64 (respectively shown in the 1st to 3rd rows). Note that for displaying purpose the images are interpolated in the lateral dimensions to match their axial dimensions. The horizontal and vertical axes are the lateral and axial directions, respectively.

Method	Parameters	Image size 512×64	Image size 256×64	Image size 128×64
PCA	Segment size	32×8	16×8	8×8
	Eigenvector used	1	1	1
HWDS	Wavelet function	Daubechies	Daubechies	Daubechies
	Vanishing moment	8	8	4
	Decomposition level	8	8	4
TVF	λ	500	500	400
	Iteration number	300	300	150
ADF	λ	0.25	0.25	0.25
	Iteration number	60	50	30
	Conduction Coefficient	30	25	25
	Favoring region	Wide	Wide	Wide

TABLE 2: Despeckling parameters used in the carotid artery images with different image sizes.

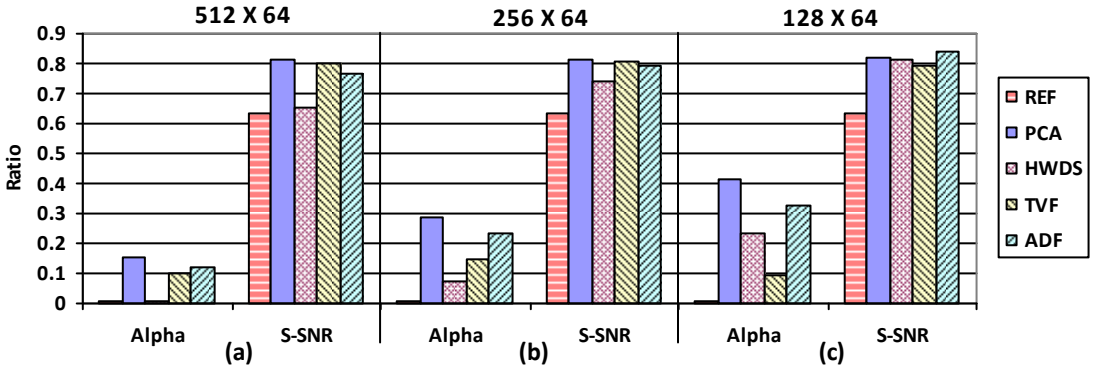


FIGURE 14: The bar graphs of the alpha and S-SNR values measured on carotid artery images of 512×64, 256×64, and 128×64 image sizes (respectively from a to c) before despeckling (REF) and after despeckling using the PCA, HWDS, TVF, and ADF methods. Vertical axis represents the magnitude.

The image visual quality of the results from the PCA, TVF, and ADF methods is far better than that from the HWDS method. For the S-SNR values, the despeckled images from the PCA method have better S-SNR values than most of those from the other methods, except for the S-SNR value from the ADF method applied on the image of 128×64 image size is better than those from the PCA method. For the α values, the despeckled images from the PCA method have higher α values than the other methods. The α parameter can be improved (lower α) by using more than one principle component per block but that will be on the account of the S-SNR. However, unlike the other denoising methods in this paper which provide better resolution accompanying reduction in the despeckling efficiency when they are applied to larger image sizes, the proposed PCA method improves the resolution while maintaining the superiority of speckle removal. Improving α while maintaining high S-SNR is observed for the PCA when moving from small to larger image sizes; from c to a in Fig. 14, or more clearly by comparing the despeckling performance of the methods for the 256×64 image size with the 512×64 image size. In general, the PCA method is a powerful speckle removing scheme regardless of the decimation factor or image size. Once the proper denoising block size is selected, the PCA-based approach provides high quality denoising results.

To further investigate the effective number of the principle components used in the PCA method on real ultrasound data, the carotid artery image of 256 × 64 image size is despeckled using the segment size of 16 × 8 and 2, 3, and 4 principle components, i.e., the parameters of (16/8-2), (16/8-3), and (16/8-4). The despeckled images are shown in Fig. 15. As expected, the higher number of principle components used, the more speckle noise becoming dominant. These results are in accordance with the simulation results previously shown.

The choice for the $q \times p$ denoising block size in this paper is based by default on the image size itself and the size of the speckle noise clusters. It follows $M \cdot \lfloor\sqrt{n}\rfloor_8 \times \lfloor\sqrt{m}\rfloor_8$ rule, where M is equal to 1 for small noise clusters and equal to 2 for large noise clusters; $\lfloor\sqrt{\cdot}\rfloor_8$ means the square root value is floored to the nearest multiple of eight; and n and m are the number of pixels in the axial and lateral dimensions of the image. For example, for the simulated phantom image of 256 × 128 image size in Fig. 1-b showing small noise clusters, therefore, the denoising block size is set to 16 × 8 (M = 1). For the simulated tissue mimicking image in Fig. 6-b after decimating its size to 128 × 64, the image shows large noise clusters; therefore, the block size is set to 16 × 8 (M = 2). For the carotid artery image in Fig. 12 after decimating its size to 128 × 64, the image shows small noise clusters; therefore, the block size is set to 8 × 8 (M = 1).

The complexity of these despeckling methods can be compared as follows. For the PCA method, there are three main stages. The first stage is to find N covariance matrices of vectors of $q.p$ length, where N is the number of image pixels and $q \times p$ is the segment size. Finding a covariance matrix of a vector requires the complexity of L^2, where L is the length of the vector. The complexity of the first stage is Nq^2p^2. From the rule to select the block size described above, $q^2p^2 \approx nm = N$. Therefore the complexity of the first stage is N^2. The second stage is to find the maximum eigenvector of the averaged covariance matrix. This requires the use of an eigen decomposition procedure and a sorting procedure. The eigen decomposition procedure is dominant and its complexity is M^3, where M is the number of elements in the matrix [25]. Therefore, the complexity of the second stage is $(q^2p^2)^3 \approx N^3$. Finally, the third stage is to projecting or filtering N vectors of $q.p$ length by a projection matrix of q^2p^2 size. This third stage requires a complexity of $NNq.p \approx N^{5/2}$. From these complexity approximations showing that the second stage is dominant, the complexity of the PCA method is N^3 and it is depending on the eigen decomposition procedure.

The complexity of the HWDS, TVF, and ADF methods can be approximated as follows. HWDS, wavelet decomposition and reconstruction require a complexity of $N \log N$ [26]. The TVF and ADF require a complexity of N for each iteration [27]. Note that the number of iteration is small compared to N. To enhance the performance of HWDS, TVF, and ADF there are two main preprocessing stages [6], which are point spread function estimation and outlier shrinkage stage. Point spread function estimation requires a complexity of $N \log N$ [28], while the outlier shrinkage stage requires a complexity of N. Note that the median filter size used in the outlier shrinkage is small compared to N. Therefore, the complexity of these despeckling methods including preprocessing is $N \log N$, which is lower than N^3 of the PCA method. However, the complexity of the PCA method could be reduced if we despeckled the image in the axial and lateral directions separately. Since the length of the vector for the covariance calculation is reduced to q or p for each direction. Therefore, the complexity could be reduced to $N^{3/2}$ and it is because of the domination of the eigen decomposition procedure. Nevertheless, this number is still larger than $N \log N$.

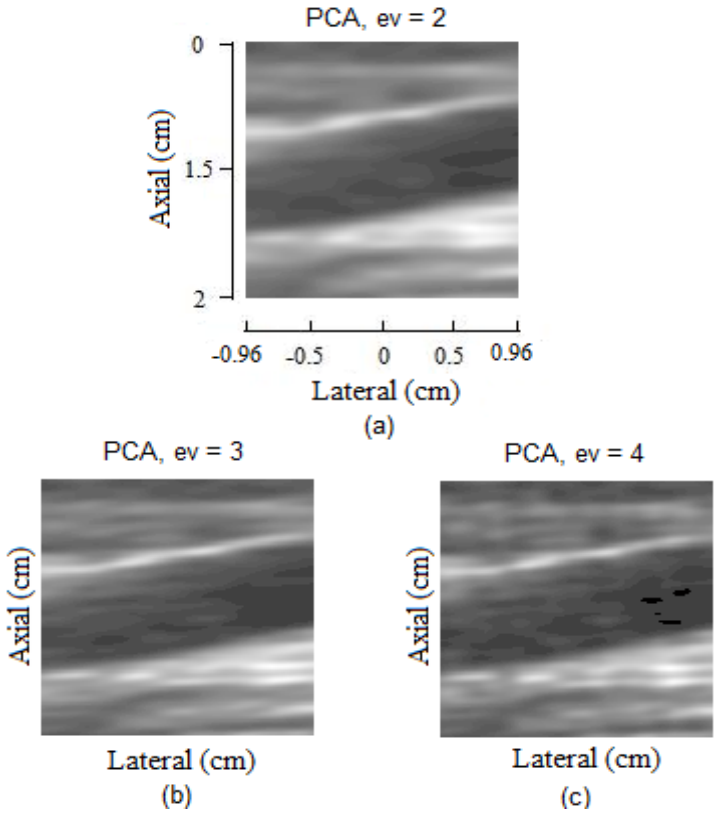


FIGURE 15: Comparison of the despeckled images using the PCA method with different numbers of principle components: ev = 2 (a), 3 (b), and 4 (c). The image size before despeckling is 256×64 and the segment size is 16×8. Note that the horizontal and vertical axes are the lateral and axial directions, respectively.

4. CONCLUSIONS

A PCA-based approach to filter out multiplicative noise from ultrasound images has been presented in this paper. It segments the envelope image into small overlapping blocks, finds a subspace representation of these blocks via principle component analysis, projects them back to the subspace, and averages the projections. The size of the block is dependent on the original size of the image and the size of the noise cluster found in the image. The proposed PCA-based approach outperforms many existing denoising approaches such as the homomorphic wavelet despeckling [4], the total variation filtering [10], and the anisotropic diffusion filtering [9] in both simulation and real ultrasound data. In terms of image resolution (α) the PCA-based approach lags the other denoising schemes. However, the S-SNR from PCA is the best for higher resolution (less decimated) images. It has been verified that, per segment size, very few principle components (in most cases the first or the first and the second principle components) are needed to represent the true signal. The proposed method depends solely on the size of the image and the number of the principle components used, while many parameters need to be adjusted for the other methods in order to obtain acceptable denoising results. Moreover, no preprocessing is needed for the proposed method, while the other methods need many preprocessing stages, such as decorrelating the speckle noise samples and then removing the outliers from the log transformed speckle noise [6]. It has been shown that the prior decimation factors performed on the noisy images have no effect on the despeckling quality of the results. In contrast to the homomorphic wavelet despeckling needs high decimation factors to improve the results. However, our approach needs to be optimized to improve the computational time and reduce the complexity of the algorithm.

5. FUTURE RESEARCH DIRECTIONS

Optical Coherence Tomography (OCT) is similar to the principle of ultrasound in the sense that OCT images are formed from the envelope of the measured interference signal. It has been shown that computation of the absolute magnitude of the signal for measurement of the envelope is a nonlinear process that destroys phase information. Processing the partially coherent OCT signals in the complex domain has provided the opportunity to correct phase aberrations responsible for speckle noise in OCT images. One of the future research directions for developing additional models for ultrasound image despeckling is to investigate the phase and spectra of the received ultrasound echo signals that are usually ignored when the magnitudes of the complex signals are being solely considered.

6. ACKNOWLEDGMENT

The authors would like to thank Prof. Jensen J., head of the center of fast ultrasound imaging at the technical university of Denmark, for his kind support in providing the medical ultrasound data used in this research.

7. REFERENCES

[1] X. Zong, A. F. Laine, and E. A. Geiser, “Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing.,” IEEE Trans. Med. Imaging, vol. 17, no. 4, pp. 532–40, Aug. 1998.

[2] A. Achim, A. Bezerianos, and P. Tsakalides, “Novel Bayesian multiscale method for speckle removal in medical ultrasound images.,” IEEE Trans. Med. Imaging, vol. 20, no. 8, pp. 772–83, Aug. 2001.

[3] L. Gagnon, A. Jouan, R. De, L. M. Canada, and A. Royalmount, “Speckle Filtering of SAR Images - A Comparative Study Between Complex-Wavelet-Based and Standard Filters,” 1997.

[4] S. Gupta, R. C. Chauhan, and S. C. Saxena, “Homomorphic wavelet thresholding technique for denoising medical ultrasound images.,” J. Med. Eng. Technol., vol. 29, no. 5, pp. 208–14.

[5] T. Loupas, W. N. McDicken, and P. L. Allan, “An adaptive weighted median filter for speckle suppression in medical ultrasonic images,” IEEE Trans. Circuits Syst., vol. 36, no. 1, pp. 129–135, Jan. 1989.

[6] O. V. Michailovich and A. Tannenbaum, “Despeckling of medical ultrasound images,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 53, no. 1, pp. 64–78, Jan. 2006.

[7] R. K. Mukkavilli, J. S. Sahambi, and P. K. Bora, “Modified homomorphic wavelet based despeckling of medical ultrasound images,” in 2008 Canadian Conference on Electrical and Computer Engineering, 2008, pp. 000887–000890.

[8] E. Nadernejad, “Despeckle Filtering in Medical Ultrasound Imaging,” Contemp. Eng. Sci., vol. 2, no. 1, pp. 17–36, 2009.

[9] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 7, pp. 629–639, Jul. 1990.

[10] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Phys. D Nonlinear Phenom., vol. 60, no. 1–4, pp. 259–268, Nov. 1992.

[11] Y. Wang, X. Fu, L. Chen, S. Ding, and J. Tian, “DTCWT based medical ultrasound images despeckling using LS parameter optimization,” 2013 IEEE Int. Conf. Image Process., pp. 805–809, Sep. 2013.

[12] M.-S. Lee, C.-L. Yen, and S.-K. Ueng, “Speckle reduction with edges preservation for ultrasound images: using function spaces approach,” IET Image Process., vol. 6, no. 7, p. 813, 2012.

[13] Y. Yu and S. T. Acton, “Speckle reducing anisotropic diffusion.,” IEEE Trans. Image Process., vol. 11, no. 11, pp. 1260–70, Jan. 2002.

[14] G. M. Henebry, “Advantages of principal components analysis for land cover segmentation from SAR image series,” in Third ERS Symposium on Space at the service of our Environment, 1997, pp. 14–21 .

[15] J.-S. Lee and K. Hoppel, “Principal components transformation of multifrequency polarimetric SAR imagery,” IEEE Trans. Geosci. Remote Sens., vol. 30, no. 4, pp. 686–696, Jul. 1992.

[16] J. Shlens, A Tutorial on Principal Component Analysis. 2009.

[17] J. F. Al-Asad, A. Moghadamjoo, and L. Ying, “Ultrasound image de-noising through Karhunen-Loeve (K-L) transformwith overlapping segments,” in 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009, pp. 318–321.

[18] J. A.K, Fundamental of Digital Image Processing. NJ: Prentice-Hall, 1989, pp. 267–330.

[19] X. Hao, S. Gao, and X. Gao, “A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing.,” IEEE Trans. Med. Imaging, vol. 18, no. 9, pp. 787–94, Sep. 1999.

[20] J. A. Jensen, D.- Lyngby, P. Medical, B. Engineering, and I. Technology, “Paper presented at the 10th Nordic-Baltic Conference on Biomedical Imaging : Field : A Program for Simulating Ultrasound Systems Field : A Program for Simulating Ultrasound Systems,” vol. 34, pp. 351–353, 1996.

[21] J. A. Jensen and N. B. Svendsen, “Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers.,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 39, no. 2, pp. 262–7, Jan. 1992.

[22] J. A. Jensen, O. Holm, L. J. Jerisen, H. Bendsen, S. I. Nikolov, B. G. Tomov, P. Munk, M. Hansen, K. Salomonsen, J. Hansen, K. Gormsen, H. M. Pedersen, and K. L. Gammelmark, “Ultrasound research scanner for real-time synthetic aperture data acquisition,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 52, no. 5, pp. 881–891, May 2005.

[23] G. Georgiou and F. S. Cohen, “Statistical characterization of diffuse scattering in ultrasound images.,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 45, no. 1, pp. 57–64, Jan. 1998.

[24] D. L. Donoho, “De-noising by soft-thresholding,” IEEE Trans. Inf. Theory, vol. 41, no. 3, pp. 613–627, May 1995.

[25] D. Tufts and C. Melissinos, “Simple, effective computation of principal eigenvectors and their eigenvalues and application to high-resolution estimation of frequencies,” IEEE Trans. Acoust., vol. 34, no. 5, pp. 1046–1053, Oct. 1986.

[26] Y. Yue, M. M. Croitoru, A. Bidani, J. B. Zwischenberger, and J. W. Clark, “Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images.,” IEEE Trans. Med. Imaging, vol. 25, no. 3, pp. 297–311, Mar. 2006.

[27] T. Brox and D. Cremers, “Iterated Nonlocal Means for Texture Restoration,” no. May, pp. 13–24, 2007.

[28] O. Michailovich and A. Tannenbaum, “Blind Deconvolution of Medical Ultrasound Images: A Parametric Inverse Filtering Approach,” IEEE Trans. Image Process., vol. 16, no. 12, pp. 3005–3019, Dec. 2007.

Script Identification In Trilingual Indian Documents

R. R. Aparna *skandhur@gmail.com*
Research Department of computer Science
S.D.N.B.Vaishnav college
AFF. Madras University
Chennai, 600043, India

R. Radha *radhasundar1993@gmail.com*
Research Department of computer Science
S.D.N.B.Vaishnav college
AFF. Madras University
Chennai, 600043, India

Abstract

This paper presents a research work in identification of script from trilingual Indian documents. This paper proposes a classification algorithm based on structural and contour features. The proposed system identifies the script of languages like English, Tamil and Hindi. 300 word images of the above mentioned three scripts were tested and 98.6% accuracy was obtained. Performance comparison with various existing methods is discussed.

Keywords: Script, Dilation, Boundary, Centroid, Zone, Blob, Bounding Box.

1. INTRODUCTION

India being a multi-lingual country, most of the documents contains two (bilingual) or three language (trilingual) scripts. English and Hindi are the most prevalent language being used in India along with the state language. Important documents like, government documents and question papers of every state contains two or three languages. Scrip is referred as the graphics part of the writing system. The writing style and the characters are combined in the script class. A script can be used by only one language or many languages. Indian and South East Asian languages are based on Brahmic scripts. North Indian languages are based on Devanagri, Bengali, Manipuri, Gurumukhi, Gujarati and Oriya scripts. South Indian languages are based on Tamil, Telugu, Kannada and Malayalam scripts. Automatic identification of scripts is very essential before feeding into multi-script optical character recognition (OCR) system. Many methods exist in the literature for script identification. In section 2 existing methods in this field are discussed. Properties of English, Hindi and Tamil are discussed in section 3. In section 4 the proposed methodology is explained. Results and performance comparison with various existing methods are detailed in section 5. Section 6 deals with the conclusion and future work.

2. LITERATURE SURVEY

Various methods exist in the literature for multi-script identification. The script identification is performed in two ways as local approach and global approach. Identification of script at text line level or word level or character level is defined as local approach. In global approach the entire text block is considered [1]. The following methods identify multi-scripts at text line level. [2], [18] used shape features, statistical features and Water Reservoir based features and classified using a rule based classifier.[3] have used horizontal projection profile (HPP) and derived a rule based classifier by calculating the threshold using the successive maxima. Structural features, K-Nearest Neighbour (KNN) and support vector Machine (SVM) was used by [4] for classification. Profile based features were used by [5] and classified using KNN. Structural, topological, contour and Water Reservoir based features were used by [6] and classified using rule based classifier.

Character level script identification technique using Gabor and Gradient features and SVM for classification [7]. The following methods were performed at word level. Eight features were obtained by [8] based on top and bottom max row, top horizontal line, tick component, top and bottom holes, vertical lines and bottom component and classified using KNN. [9] performed Devanagri script identification using the presence of headlines as [2] and Kannada script identification using morphological operation. Their method was restricted to font style and size. [10] have used 400 and 64 dimensional features and SVM was used for classification. [11] has used Neural networks (NN) for classification and features based on histogram mean distance, pixel value and vertical zero crossing. [12] used Zone based Gabor features and SVM classifier was used.

3. PROPERTIES OF ENGLISH, HINDI AND TAMIL SCRIPTS

English language uses Latin or Roman script. English follows the Alphabetic system. Hindi language uses Devanagri script. The characters of the Hindi words are connected using a headline called “Shirorekha”. Printed Tamil and English words contain isolated characters. The structure of the three scripts is divided into three zones (Fig.1). The upper zone (UZ), middle zone (MZ) and lower Zone (LZ). The character set of English constitutes 52 characters comprising 26 uppercase and 26 lowercase characters. The uppercase English characters stay inside the middle zone. The lowercase English characters like b, d, f, h, k, l, t containing ascenders occupies UZ and those containing desecenders like g, j, p, q, y occupies the LZ. But the presence of pixels in the LZ and UZ by lower case English alphabets is of symmetrical pattern (Fig1.a). Tamil character contains 12 vowels and 18 consonants. 216 compound characters were formed by combining vowels and consonants. Almost all the character of Tamil occupies UZ and LZ. Tamil character that occupies the UZ and LZ has varied pixel density and patterns (Fig1.b). Hindi character contains 34 consonants and 10 vowels and form combined characters. Since Hindi characters are connected using headline, the probability of pixel density in the LZ and UZ is very less (Fig1.c).

FIGURE 1: a) English script b) Tamil script c) Hindi script.

4. PROPOSED METHODOLOGY

The input image is converted to greyscale image and binarized using Ostu method [17], (Fig.2). The proposed algorithm works by dividing the multi-script words into three zones and the pixel densities are calculated in the UZ above Headline (HL) and LZ below base line (BL) for discriminating multi-script (fig.1). The proposed algorithm is based on the contour features and classifies using a rule based classifier. First it checks whether the script belongs to Hindi by using the pixel density of UZ and LZ. Classification of English and Tamil script is performed using the contour features of the boundary pattern such as upper boundary (UB) and lower boundary (LB) pixel density calculation.

FIGURE 2: Binarized images: a) Tamil b) English c) Hindi.

4.1 Classification Phase - I

The given word image is classified in two phases. In phase - I the image is classified to check the presence of Hindi word. In phase - II, it is classified as English or Tamil.

Step 1: The heights (h_i) and widths (w_i) of all the blobs are calculated for the image (Fig2). Where i=1 to N, N is the number of blobs.

Step 2: Calculate the average width $avg(\mathrm{w}_i)$ and $avg(\mathrm{h}_i)$ average heights of all blobs.

Step 3: Divide every h_i by $avg(\mathrm{h}_i)$ and store in an array.

$$v_1 = \frac{h_i}{avg(h_i)} \tag{1}$$

Step 4: Choose the $\min(v_1)$ as it is closer to the average $avg(\mathrm{h}_i)$.

Step 5: Extract the $\min_x$, $\min_y$ and ht for the chosen h_i (blob).

Step 6: Calculate the height of the MZ using the following formula. (Fig.3)

$$mh = round\left(min_y + ht\right) \tag{2}$$

FIGURE 3: Middle Zone (MZ) Height.

Step 7: Count the pixels row wise for all rows of the entire image and store it in an array (s1).

Step 8: Count the pixels above HL and below BL using

$$\left(s2 = \sum_{i=1}^{\min_y} s1, \mathrm{s}3 = \sum_{i=mh}^{\max_y} s1 \right) \tag{3}$$

Step 9: Count the no of blobs

The word is classified as Hindi if pixels in BL and HL is zero and the total number of blobs is equal to one as the Hindi word is connected by a headline otherwise it is classified as English or Tamil in the next phase.

4.2 Classification Phase – II

Step 1: Extract the boundary features using Morphological operations
Step 2: Dilate the image.(Fig.4a)

Step 3: Calculate the centroid value c_x and c_y .

Step 4: Extract the boundary co-ordinates and store in the array, x_{ub} , y_{ub} and x_{lb} , y_{lb} . (Fig. 4 & 5)

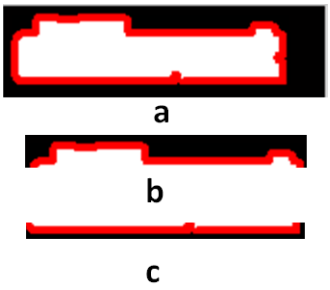


FIGURE 4: a) Obtained Boundary (English) b) UB c) LB.

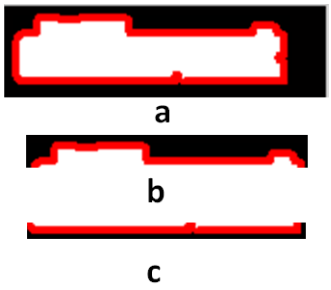


FIGURE 5: a) Obtained Boundary (Tamil) b) UB c) LB.

Step 5: Extract the upper boundary pixel co-ordinates (ub_y) and lower boundary (lb_y) pixel co-ordinates using the following formula,

$$ub_y = \sum_{i=1}^{\max(y_{ub})} y_{ub}, \mathrm{y}_{ub}(i) < c_y \quad (4)$$

$$lb_y = \sum_{i=1}^{\max(y_{lb})} y_{lb}, \mathrm{y}_{lb}(i) > c_y \quad (5)$$

Step 6: Calculate the total count of pixels in the upper boundary and lower boundary. If the total pixels is greater than the threshold then it is classified as Tamil else English. The basic idea is that, the upper boundary of Tamil words (Fig.6a) contains unsymmetrical patterns (peaks and valleys) by accumulating more pixels. But in English word (Fig.6b) the boundary shape pattern is more symmetrical without jerks. Hence the pixels between min_y and $\max_y$ will be zero or less for English. Always the characters of Tamil font will be little bigger in size compared to the English character of same font size.

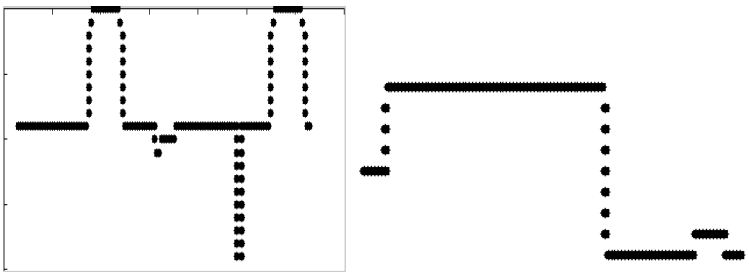


FIGURE 6: Scatter plot for UP of Tamil and b) Scatter plot for UP of English.

5 RESULTS AND DISCUSSION

Samples were collected from the trilingual documents. 300 word samples of Hindi, English and Tamil were tested using the proposed algorithm and 98.6% accuracy was obtained. In cases where the Tamil words were very small and having symmetrical pattern was misclassified as English. The error rate is very low as always the number of pixels will be greater for Tamil compared to English.

Script	Hindi	English	Tamil
Hindi	**100**	0	0
English	0	**98**	2
Tamil	0	2	**98**

TABLE 1: Accuracy of Script Identification.

There is lack of any comparative analysis of the results with most of the reported works in script recognition. Experimental results of every proposed method have not been compared with other benchmark works in the field [1]. As there was no standard evaluation measures available in the literature. The results were evaluated as using the following statistical measures (Table. 2 & 3). The proposed algorithm is evaluated using the statistical measures Recall, Precision, Accuracy and Fscore based on the following metrics. (H-Hindi, E-English and T-Tamil)

True Positive (TP): Correctly identified: (H, E, T identified as H, E, T).
False Positive (FP): Incorrectly identified: (E, T incorrectly identified as T, E).
True Negative (TN): Correctly rejected.
False Negative (FN): Incorrectly rejected.

5.1 Statistical Measure

Recall: This gives the positive cases obtained. It is same as the sensitivity to identify positive results.

$$\mathrm{Re}\,call = \frac{TP}{TP + FN}$$

Precision: This measure gives the percentage of positive predictions.

$$\mathrm{Precision} = \frac{TP}{FP + TP}$$

Accuracy: This measure gives the predictions that were correct.

$$\mathrm{Accuracy} = \frac{TN + TP}{TP + FN}$$

Fscore: This measure is used as a single measure of performance test by combining all the above results.

$$\mathrm{Fscore} = 2\left(\frac{\mathrm{Pr}\,ecision * \mathrm{Re}\,call}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call}\right)$$

	Predicted -ve	Predicted +ve
-ve case	TN: 0	FP: 4
+ve case	FN: 0	TP: 296

TABLE 2: Statistical Results.

Measure	Accuracy
Accuracy	100%
Recall	100%
Precision	98.6%
Fscore	99%

TABLE 3: Accuracy of Classification.

5.2 Performance Comparison with the Existing Methods

The proposed algorithm is compared with many of the existing methods in the literature for script identification of Hindi (H), English (E), Kannada (K), Devanagri (D), Oriya (O),Telugu (Te), Gurumukhi (G), and Tamil (T) using local approach.

Author & year	Script	Local Approach	Technique Used	Accuracy	Remarks
[14]-2002	E, H & K	Word-wise	NN	98%	Fixed font size and computationally complex.
[12]-2003	12 scripts	Text line	Local features and water reservoir based feature	98%	Many features were extracted from individual characters
[13]-2010	K, H & E	Text line	Top and bottom profile, KNN classifier	99.75%	Dependent on font size, time consuming
[16]-2002	T & E	Word-wise	Gabor features, SVM classifier	96%	Computationally complex
[5]-2005	D, E & U	Word-wise	Profile based features, Water reservoir method	97.51%	Character level features were extracted.
[9]-2009	E, D & B	Word-wise	64 & 400 dimensional features and SVM as classifier	98.51%	Error rate increases when number of character is less than three
[3]-2013	H, B, Te & K	Text line	Projection profile and rule based classifier	97.83%	Fixed font size was used
[6]-2013	G & E	Character level	Gabor, Gradient features and SVM classifier	99.45%	Complex and Time consuming
[10]-2013	D	Word-wise	Mean distance, pixel value and zero crossing, NN	90%	Computationally complex and execution time is very high
[7]-2011	H &E	Word-wise	Many features,	95%	

			KNN		
[8]-2011	K, E & H	Word-wise	Features based on Morphology and KNN	95.54%	Suitable only for fixed font size and font style
[15]-2002	D, T & O	Word-wise	Gabor filters	97.33	Computationally complex
Proposed method-2014	**H, E &T**	**Word-wise**	**Zonal and boundary features using rule based classifier**	**98.6%**	**Simple, fast, efficient and accurate**

TABLE 4: Performance Comparison.

6. CONCLUSION AND FUTUREWORK

The existing methods are computationally complex and have used more features based on structural and pixel density feature. But the proposed method is very simple, efficient, fast and accurate for discriminating the scripts in bilingual or trilingual documents. Majority of the Indian bilingual or trilingual documents contains 2 or 3 scripts. The documents contain Hindi, English and the state language. Hence the proposed algorithm can be used for script identification of trilingual documents containing Tamil, Hindi and English or Bilingual documents containing any two combinations of these three scripts. The future work will be concentrated towards script identification of Kannada, Telugu, and Malayalam in trilingual or bilingual documents with the combination of Hindi and English as the other two languages.

7. REFERENCES

[1] D.Ghosh, T.Dube and A.P.Shivaprasad. “Script Recognition: A review,” IEEE Transactions on pattern analysis and machine Intelligence, vol. XX, 2009.

[2] U.Pal and B.B.chaudhuri. “Identification of different script lines from multi-script documents,” Image and Vision Computing, vol. 20, pp. 945-954, 2002.

[3] O.Prakash, V.Shrivastava and A.Kumar. “An efficient approach for script identification,” International journal of computer Trends and Technology (IJCTT), vol. 4, 2013.

[4] Rajesh Gopakumar, N.V.SubbaReddy, Krishnamoorthi Makkithaya and U.Dinesh Acharya. “Script Identification from multilingual Indian documents using structural features,” Journal of computing, vol. 2, 2010.

[5] S.Chanda and U.Pal. “English , Devnagari and Urudu text Identification,” in Proc. International conference on Document Analysis and Recognition, pp. 538-545, 2005.

[6] R.Rani, R.Dhir and G. Lehal. “Script Identification of pre-segmented multifont characters and digits,” in Proc. ICDAR, 2013, pp.1150-1154.

[7] M.Swamy Das, D.Sndhya Rani, C.R.K.Reddy and A.Govardhan. “Script identification from multilingual Telugu, Hindi and English Text documents,” International Journal of Wisdom based computing, vol.1, 2011.

[8] B.V.Dhandra and M.Hangarge. “Morphological reconstruction for word level script identification,” International journal of computer science and security, vol.1, 2011.

[9] S.Chanda, S.Pal, K.Franke, U.Pal. “Two-stage approach for word wise script Identification,” in Proc. ICDAR, 2009.

[10] D.Yadav, S.Sanchez-cuadrado and J.Morato. “Optical acharacter recognition for Hindi language using a NN approach,” Journal of information processing systems, vol. 9, 2013.

[11] R.Rani, R.Dhir and G.Lehal. “Modified Gabor feature extraction method for word levelscript identification –Expermentation with Gurumukhi and English scripts,” International journal of signal processing, image processing and pattern recognition, vol. 6, pp. 25-38, 2013.

[12] U.pal and B.B.Chaudhri. “Indian script Character recognition: A survey,” Pattern recognition, vol. 37, pp.1887-1889, 2004.

[13] M.C.Padma and P.A Vijaya. “Script identification from trilingual documents using profile based feature,” International journal of computer science and applications, vol.7, pp.16-33, 2010.

[14] S.B. Patil and N.V.Subbareddy. “Neural network based system for script identification in Indian documents,” Sadhana, vol. 27, pp. 83-97, 2002.

[15] P.B.Pati, S.S.Raju, N.Pati and A.G.Ramakrishnan. “Gabor filters for Document analysis in Indian Bilingual Documents,” in Proc. IEEE ICISIP, 2004.

[16] D.Dhanya, A.G.Ramakrishnan and P.B.Pati, “Script identification in printed bilingual documents,” Sadhana, vol. 27, pp. 73-82, 2002.

[17] R.C.Gonzalez, R.E.Woods and S.L.Eddins. “Digital Image Processing using Matlab,” NewDelhi, Tata Mcgraw-Hill, 2011.

[18] U.Pal and B.B.Chaudhri. “Script line separation from Indian Multi script documents,” in Proc. ICDAR, pp. 406-409,1999.

Performance Comparison of Hybrid Haar Wavelet Transform with Various Local Transforms in Image Compression using Different Error Metrics

H. B. Kekre *hbkekre@yahoo.com*
Sr. Professor / Computer Eng. Department / MPSTME
NMIMS University
Mumbai, 400056, India

Tanuja Sarode *tanuja_0123@yahoo.com*
Associate professor / Computer Eng. Department / TSEC
Mumbai University
Bandra, 400050, India

Prachi Natu *prachi.natu@yahoo.com*
Assistant Professor/Computer Engineering Department / MPSTME
NMIMS University
Mumbai, 400056, India

Abstract

A novel image compression using hybrid Haar wavelet transform has been proposed in this paper. Hybrid wavelet transform is generated using two different orthogonal transforms. Haar transform acts as a base transform and other sinusoidal transforms like DCT, DST, Hartley and Real-DFT are paired with Haar transform to generate hybrid Haar wavelet. Among these four pairs Haar-DCT hybrid wavelet gives lower error as compared to Haar-DST, Haar-Hartley and Haar-Real-DFT. Performance of Haar-DCT hybrid wavelet is further analyzed using multi resolution hybrid wavelet and Haar-DCT hybrid transform. Experimental results show that hybrid wavelet with component size 16-16 gives lower error at higher compression ratios than multi resolution analysis and hybrid transform performance. Performance is measured using RMSE which is traditional parameter to measure error. Lowest RMSE obtained is 9.77 at compression ratio 32 using Haar-DCT Hybrid Wavelet with component size 16-16. Various other error metrics like MAE, AFCPV and SSIM are used to measure error. Lowest MAE and AFCPV are observed at compression ratio 32 are in Haar (16x16) –DCT (16x16) hybrid wavelet having values 6.86 and 0.31 respectively. When blocked SSIM is applied on 16-16 Haar-DCT hybrid wavelet it gives value 0.993 at compression ratio 32 which is closer to one indicating that good quality of compressed image is obtained.

Keywords: Hybrid Transform, Haar Wavelet, SSIM, MAE, Image Compression.

1. INTRODUCTION

Digital images are inevitable part of today's multimedia world. Downloading and transmitting the images using internet consume considerable amount of time and bandwidth. Hence if these images are compressed and used, it will help to save time and bandwidth required for transmission. Image compression deals with reducing the number of bits required to store and transmit the image. It eliminates redundant information in image by retaining important information and intrinsic structure of the original image such that quality of compressed image is acceptable to human visual system. Wavelets have gained immense popularity in image compression during last two-three decades due to their high energy compaction property and multi-resolution analysis [1]. Basically wavelets are mathematical functions that represent time-frequency analysis of the data. Selecting a wavelet prototype function or mother wavelet is an

essential step in wavelet based analysis. Other wavelets are produced by translation and contraction of mother wavelet. These generated wavelet functions must be orthogonal which will comprise wavelet transform. Haar transform is a simple, orthonormal transform proposed by Alfred Haar in 1910 [2]. It serves as a prototype for wavelet transform. [3]. Many wavelet based image compression techniques have been used till now including Haar wavelet. Commonly used error metrics to judge the performance of compression methods are Mean Square Error (MSE) and Peak Signal to noise ratio (PSNR). This paper proposes hybrid Haar wavelet transform and compares its performance with its multi-resolution hybrid wavelet transform and hybrid transform using various error metrics like root mean square error (RMSE), mean absolute error (MAE), average fractional change in pixel value (AFCPV) and structural similarity index (SSIM).

2. REVIEW OF LITERATURE

A lot of research has been done on wavelet based image compression. Image compression using sparse Haar wavelet has been proposed by R. Mehala and Kuppusamy [***4***]. In this paper they have used 8x8 Haar matrix by inserting appropriate zeroes and ½ in Haar wavelet and it is applied on blocked gray scale image. Performance is measured using compression ratio and PSNR. But this method is applicable only on low intensity image. Haar wavelet and neural network based image compression is proposed by S. Shridhar et al. [5]. Using Haar wavelet image is decomposed into different frequency sub bands and then scalar quantization and Huffman coding are used for compression of different sub bands***.*** The coefficients in low frequency band are compressed by Differential Pulse Code Modulation (DPCM) and the coefficients in higher frequency bands are compressed using neural networks. Complexity of this method is high and it measures performance using traditional error parameter i.e. MSE. Wavelet based extension of JPEG 2000 standard is proposed by Singh and Sharma [6]. In this paper first level wavelet decomposes the image only in vertical direction and subsequent wavelet levels use full horizontal and vertical splitting for all image components. But performance of this method degrades when images are with low colour depth. Medical image compression using Haar wavelet, Daubechies wavelet and Coeflit wavelet has been proposed by Krishna Kumar et al. [7] Where performance has been measured using PSNR and SSIM. Different wavelets perform better for different kind of medical images and SSIM up to 0.7 is obtained for ECG images. Singular value decomposition combined with linear and quadratic interpolation has been proposed by J Hizadian, A Hosaini and M Jalili [8]. But this method is time consuming. A simple wavelet transform based image compression is proposed in [9] by H. B. Kekre, Tanuja Sarode and Prachi Natu. In this paper performance of full wavelet transform is compared with respective column wavelet and row wavelet transform. To save number of computations, column wavelet transform can be used with slight increase in error in reconstructed image. Hybrid wavelet based compression using DCT with RealDFT has been proposed in [10]. Hybridisation helps to incorporate properties of both transforms used to generate hybrid wavelet transform and hence error reduces drastically as compared to orthogonal transform and orthogonal wavelet transform.

In this paper a simpler method of image compression using Hybrid Haar wavelet has been proposed. Sinusoidal transforms like DCT, DST, Real DFT and Hartley are combined with Haar transform to generate hybrid Haar wavelet. This transformation matrix is generated in three different ways to study the effect of global, local and semi global properties on image compression. Performance of proposed method is measured using various fidelity criteria like RMSE, MAE, AFCPV and SSIM.

3. HAAR TRANSFORM

Haar transform is the simplest and basic transform used in image processing. It is faster to implement and helps to analyze local features of a signal. Haar transform uses Haar function as its basis function which varies in both scale and position. 8x8 Haar transform matrix is given below. It contains only real elements 1,-1 and 0.

$$H[m,n] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \quad (1)$$

Haar functions for 8x8 Haar matrix are shown in Fig. 1

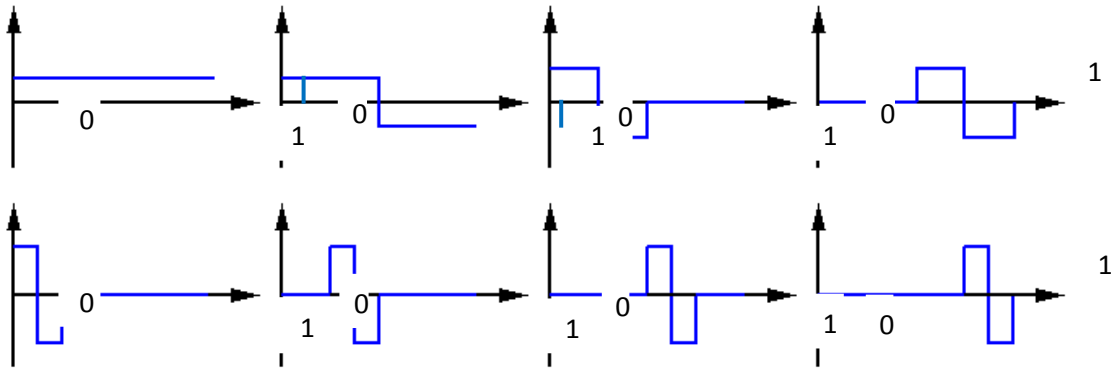


FIGURE 1: Haar functions of 8x8 Haar matrix.

4. FIDELITY CRITERIA

4.1 Root Mean Square Error

It is commonly used fidelity criteria to measure the distortion between original image and reconstructed image. Mathematically it is calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=p}\sum_{j=1}^{j=q}(x_{ij}-y_{ij})^2}{p*q}} \quad (2)$$

But it is not very well matched to perceived quality [11]. It is calculated with assumption that loss of perceptual quality is directly related to visibility of error signal.

4.2 Mean Absolute Error

It is average of absolute difference in pixel values. It gives better perceptibility than RMSE.

$$MAE = \frac{\sum_{i=1}^{i=p}\sum_{j=1}^{j=q}(|x_{ij}-y_{ij}|)}{p*q} \quad (3)$$

4.3 Average Fractional Change in Pixel Value

It represents the fractional change in pixel values and hence reflects perceptibility in better way.

$$AFCPV = \frac{\sum_{i=1}^{i=p}\sum_{j=1}^{j=q}(|x_{ij}-y_{ij}|)/x_{ij}}{p*q} \quad (4)$$

4.4 Structural Similarity Index (SSIM) [11]

Concept of SSIM is based on the assumption that human visual system is highly adapted for extracting structural information from still image and hence measure of structural similarity can provide a good approximation to perceived image quality. SSIM considers image degradation as perceived change in structural information. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene. SSIM is calculated as

$$SSIM\ (x, y) = (2\mu_x\mu_y+c_1)\ (2\sigma_{xy}+c_2)\ /\ (\mu_x^2+\mu_y^2+c_1)\ (\sigma_x^2+\sigma_y^2+c_2) \qquad (5)$$

5. PROPOSED METHOD

5.1 Hybrid Wavelet Transform [12]

In this paper hybrid Haar wavelet has been proposed using Kekre's algorithm to generate hybrid wavelet transform [12]. Haar transform is combined with sinusoidal transforms like DCT, DST, Hartley and Real-DFT. All these sinusoidal transforms are selected as local components of hybrid wavelet transform. Use of hybrid wavelet transform helps to incorporate traits of both component transforms. It is generated using following Kronecker product.

$$T_{AB} = \begin{pmatrix} A_p \otimes B_q(1) \\ I_p \otimes B_q\ (2) \\ I_p \otimes B_q\ (3) \\ . \\ . \\ . \\ . \\ I_p \otimes B_q\ (n) \end{pmatrix} \qquad (6)$$

Here 'A' is pxp 'base' transform and 'B' is qxq 'local' transform used to generate pqxpq hybrid wavelet transform denoted as T_{AB}. Size of component transforms p and q are selected such that size of T_{AB} is same as image size. B_q (1) indicates first row of matrix B whose Kronecker product with A is taken. It generates first p rows of transform matrix representing global features of an image. Kronecker product of Identity matrix 'I' of size pxp and each row of matrix B is used to translate the rows of matrix B which will contribute to local properties.

5.2 Multi-resolution Hybrid Wavelet [*13*]

Above generated hybrid wavelet matrix gives global and local features of an image hence give bi-resolution analysis. It is modified as below to include semi global features of an image.

$$T_{AB} = \begin{pmatrix} A_p \otimes B_q\,(0{:}i_1) & \text{Global} \\ I_{r0} \otimes (A_{p/r0} \otimes B_q\,(i_1{+}1{:}i_2)) & \text{Semi global 1} \\ I_{r1} \otimes (A_{p/r1} \otimes B_q\,(i_2{+}1{:}i_3)) & \text{Semi global 2} \\ \vdots & \vdots \\ I_{rn-1} \otimes (A_{p/rn-1} \otimes B_q\,(i_{n-2}{+}1{:}i_{n-1})) & \text{Semi global n} \\ I_p \otimes (B_q(i_n{:}q)) & \text{Local} \end{pmatrix} \quad (7)$$

In above Kronecker product A is pxp and B is qxq orthogonal transforms. Bq(i:j) represents i to j rows of transform B. Lower order matrix A of size p/r_n is generated where r_n is divisor of p except 1 and p itself. Lower order matrix A is used for scaling operation and Identity matrix is used for shifting operation in generation of hybrid wavelet transform.

5.3 Hybrid Transform

It is the limiting case of hybrid wavelet transform obtained by full Kronecker product of two component transforms.

It is given as $A \otimes B = a_{ij}\,[B]$ (8)

Where, a_{ij} is individual element of matrix A.

Generated hybrid wavelet transform is applied on each plane of color image. Coefficients in transformed plane are sorted in descending order of energy. Low energy coefficients are discarded. Image is reconstructed by applying inverse transform on transformed planes with retained high energy coefficients. Distortion between original and reconstructed image is measured using various fidelity criteria mentioned above. Performance of different hybrid Haar wavelets is compared using experimental database and required graphs are plotted below.

6. EXPERIMENTS AND RESULTS

Dataset used for experimental work is shown in Fig. 2. In contains colour images from different classes. All images are of 256x256x3 bytes.

FIGURE 2: Dataset of Color Images used for Experimental Work.

Fig. 3 shows plot of RMSE vs. compression ratio using hybrid Haar wavelet with different sinusoidal transforms. It shows that Haar-DCT hybrid wavelet gives less RMSE for all compression ratios. RMSE 9.91 is obtained at highest compression ratio 32. Performance of Haar-DCT is followed by Haar-RealDFT. DST combined with Haar gives higher error among all.

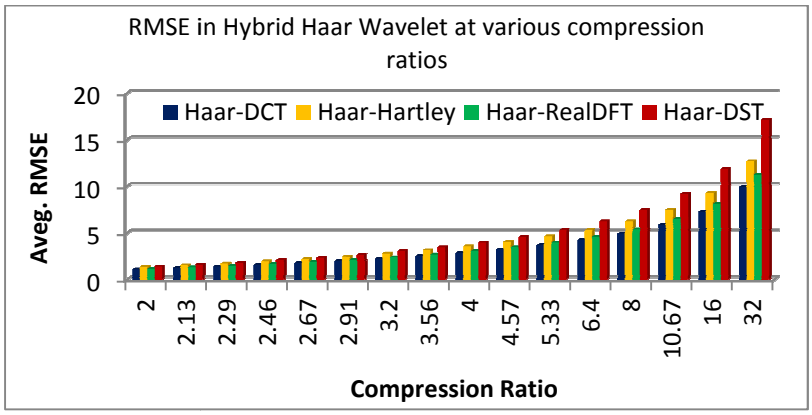


FIGURE 3: Average RMSE vs. Compression Ratio in Hybrid Wavelet using Haar Transform (8x8) with different Sinusoidal transforms (32x32).

As Haar-DCT gives less error, further it is analyzed to obtain the best component size combination of Haar and DCT that gives lowest error. Four combinations of component transforms are selected and RMSE in each of them is compared as shown in Fig. 4. Up to compression ratio 10.67, 16-16 and 32-8 size give nearly equal error. Slight increase in error is observed for size 32-8 at compression ratios 16 and 32. Thus 16-16 becomes better combination at higher compression ratios.

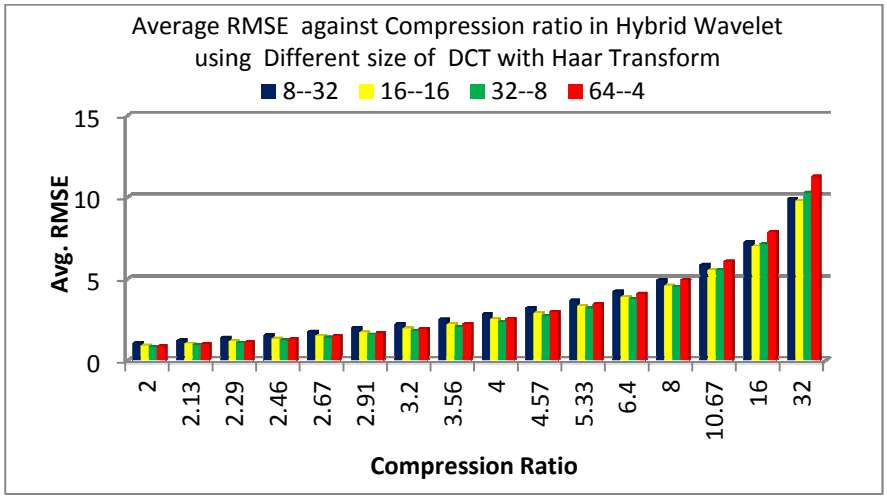


FIGURE 4: Comparison of RMSE in Haar-DCT hybrid wavelet using different component transform size.

Fig. 5 shows comparison of RMSE in multi-resolution hybrid wavelet. Here also Haar-DCT proves to be better giving less error. At compression ratio 32, RMSE 10.13 is obtained with component size 8-32. It is slightly higher than error in Haar-DCT hybrid wavelet shown in Fig. 3

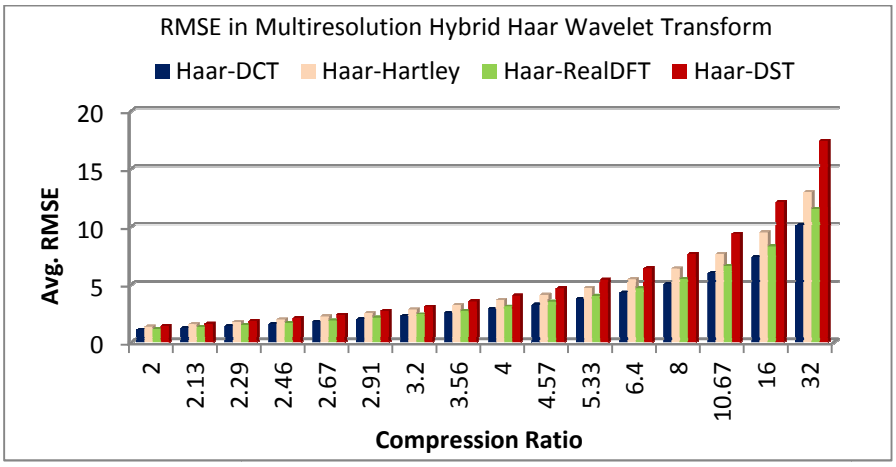


FIGURE 5: Avg. RMSE vs. Compression ratio in Multi-resolution using Haar (8x8) and sinusoidal component transforms of size 32x32.

For multi-resolution analysis, four different sized pairs of Haar-DCT are tried and their error is plotted against compression ratio in Fig. 6. Pair of size 8-32 gives lower RMSE than all other pairs at compression ratio 32. At lower compression ratios, difference in error is negligible.

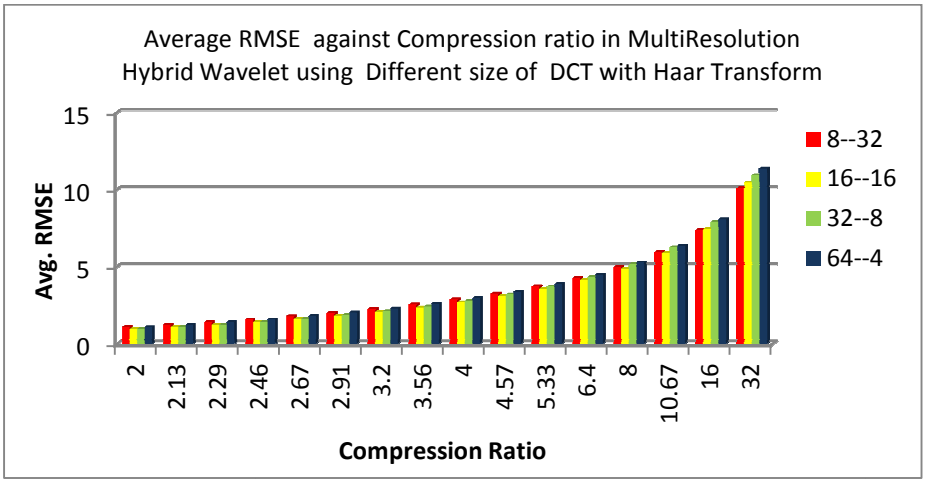


FIGURE 6: RMSE vs. Compression ratio for different component sizes of Haar-DCT in Multi-resolution analysis.

Fig. 7 shows RMSE obtained in Hybrid transform. In hybrid transform, which is obtained by full Kronecker product of two component matrices, all rows of transformation matrix represent global features of an image. Local and semi global features are absent in hybrid transform. In hybrid transform also Haar-DCT gives better performance in terms of RMSE.

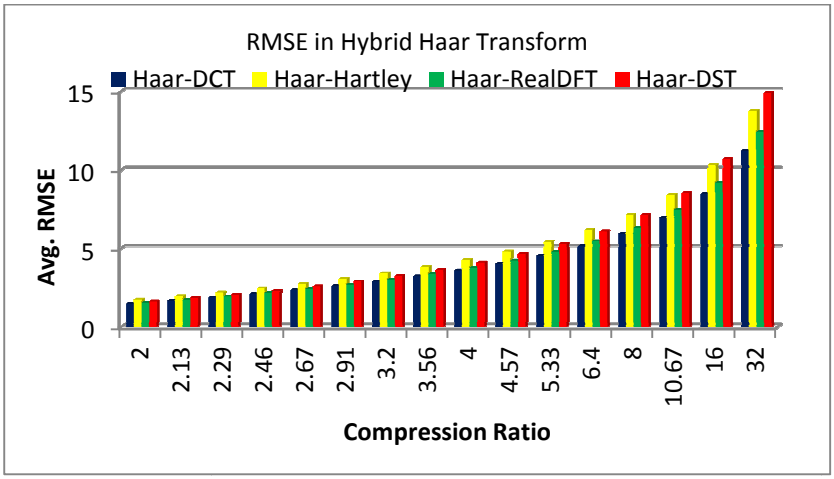


FIGURE 7: Avg. RMSE vs. Compression ratio in Hybrid transform using Haar transform (8x8) with different sinusoidal transforms (32x32).

Different size combinations of Haar-DCT hybrid transform are tried and error is plotted in Fig. 8. It shows that four different combinations give nearly equal error for selected compression ratio. Error increases with increase in compression ratio.

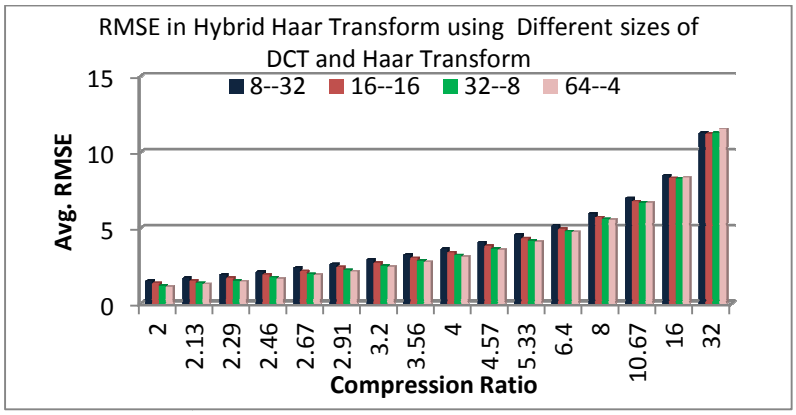


FIGURE 8: Average RMSE against Compression ratio in Hybrid Transform using different size of DCT with Haar Transform.

Fig. 9 shows overall comparison of RMSE in hybrid Haar wavelet, its multi resolution analysis and hybrid transform at various compression ratios. Up to compression ratio 8, hybrid wavelet with components Haar 32x32 and DCT 8x8 gives less error. Onwards, as compression ratio increases, Haar 16x16 and DCT 16x16 give less RMSE.

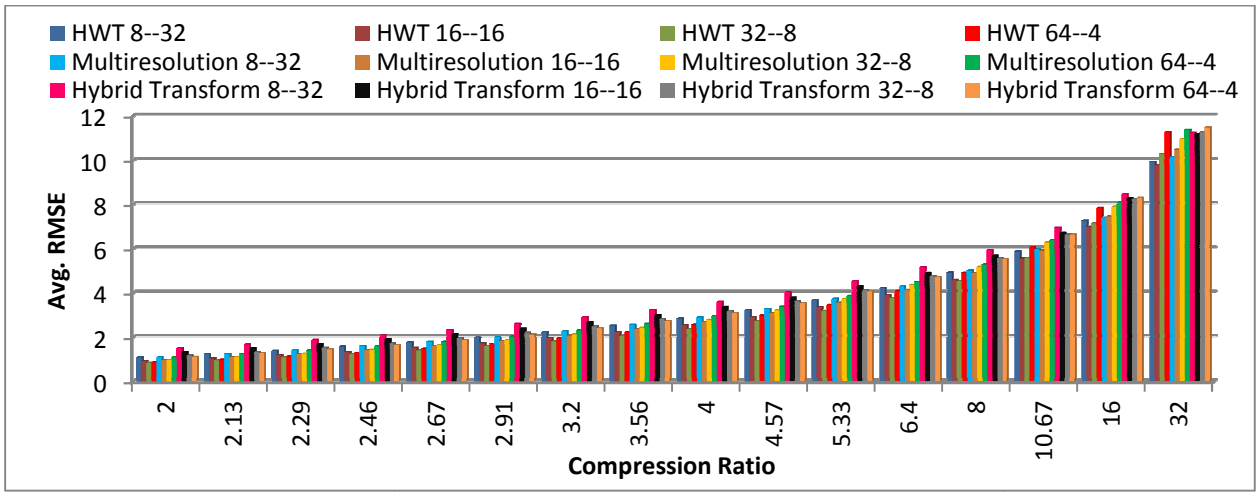


FIGURE 9: Average RMSE against compression ratio using different sizes of component transforms in Haar–DCT Hybrid Wavelet, Multi resolution analysis and Hybrid transform.

Fig. 10 shows performance comparison of Haar hybrid wavelet using MAE as error measurement criterion, Fig. 11 shows this comparison for multi resolution analysis and in Fig. 12 comparison of different hybrid transforms is done. Similar to RMSE, Haar-DCT gives better performance than other sinusoidal transforms in multi resolution as well as in Hybrid transform.

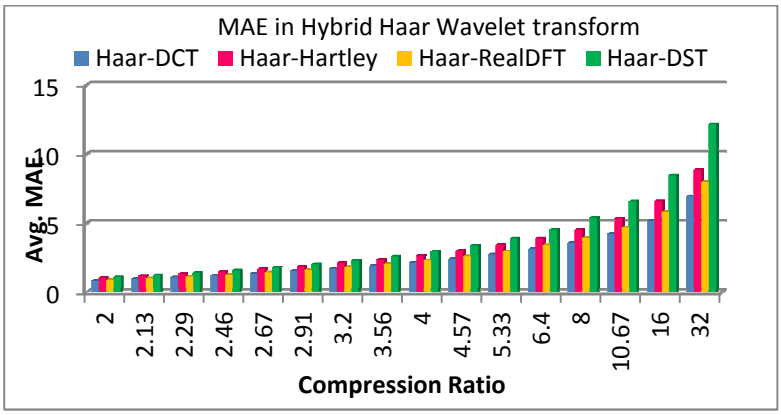


FIGURE 10: Avg. MAE against Compression Ratio in Hybrid Wavelet transform using different Sinusoidal transforms with Haar Transform.

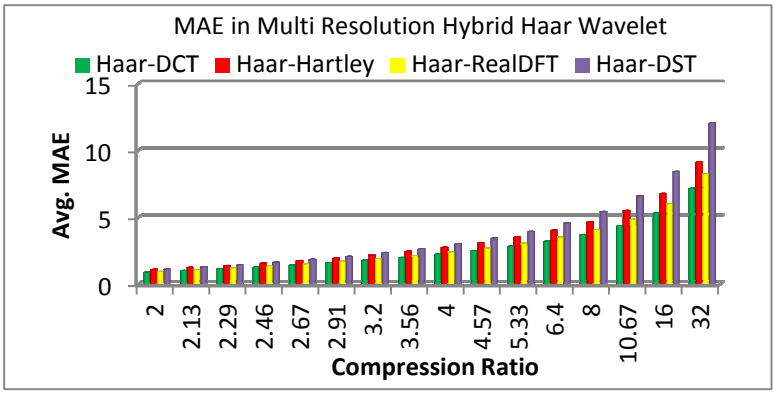


FIGURE 11: Avg. MAE against Compression ratio in Multi Resolution Hybrid Wavelet using different Sinusoidal transforms with Haar Transform.

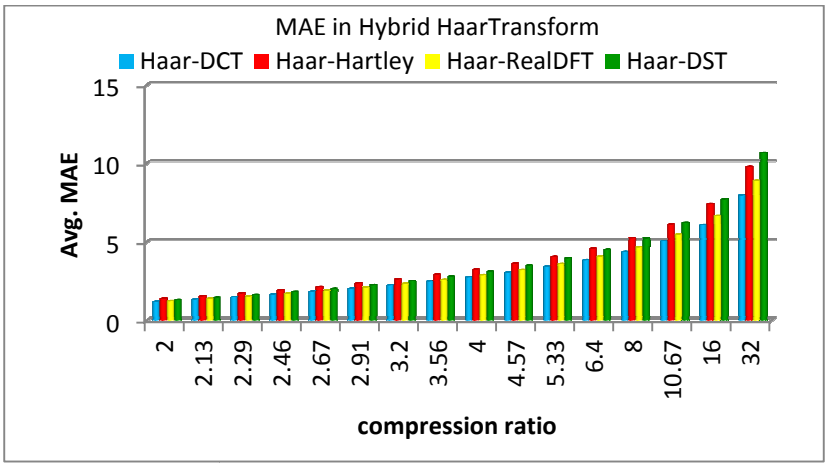


FIGURE 12: Avg. MAE vs. Compression ratio in Hybrid Transform using different Sinusoidal Transforms with Haar Transform.

Fig. 13, 14 and 15 show graph of MAE against compression ratio using various size combinations of Haar-DCT in Hybrid wavelet, its multi resolution and hybrid transform respectively. The best pair obtained in Hybrid wavelet is of size 16-16 and in Multi resolution analysis, it is of 8-32 for higher compression ratio 32.

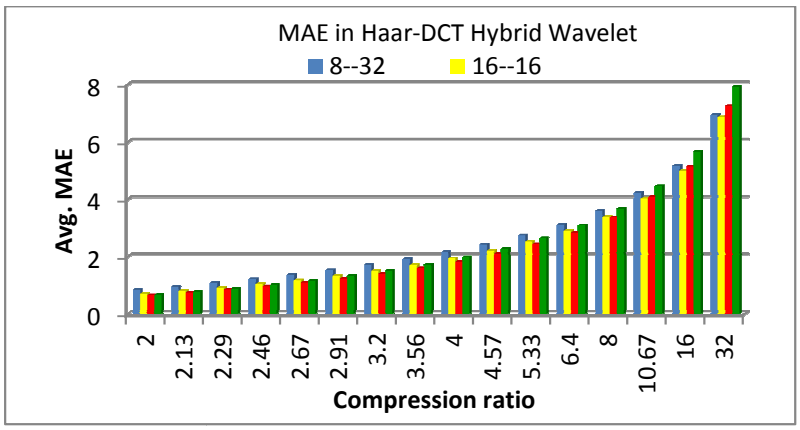


FIGURE 13: Avg. MAE against Compression ratio using Different size of Haar-DCT pairs in Hybrid Wavelet.

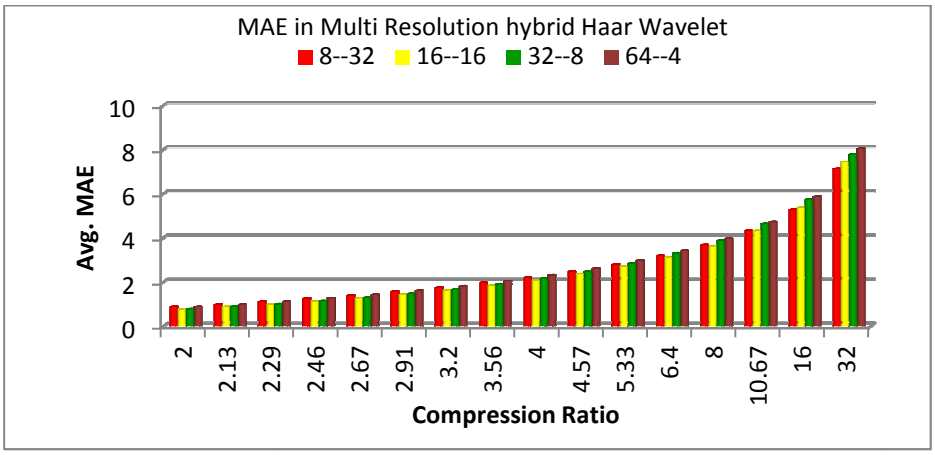


FIGURE 14: Avg. MAE in Multi Resolution hybrid Wavelet using different size variations in Haar-DCT.

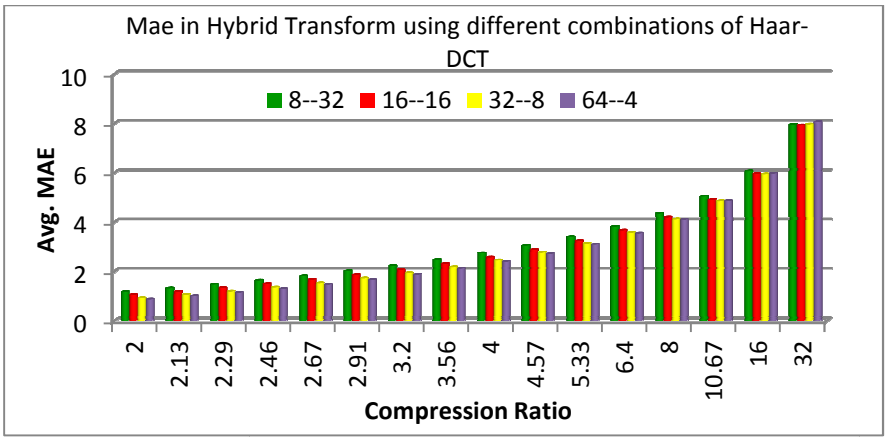


FIGURE 15: Avg. Mae vs. Compression ratio in Hybrid Transform using different combinations of Haar-DCT.

Fig. 16 shows overall comparison of MAE in Haar-DCT Hybrid Wavelet, Multi resolution analysis and hybrid Transform using various size combinations. It has been observed that Haar-DCT hybrid wavelet of size 16-16 gives lower MAE at higher compression ratio 32.

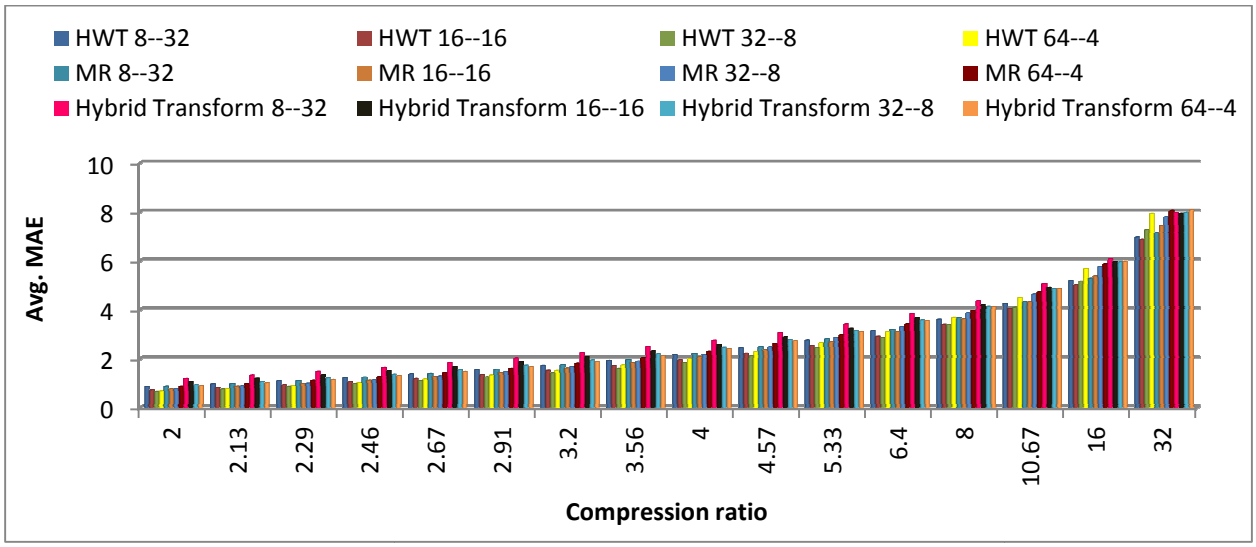


FIGURE 16: Average MAE against compression ratio using all possible sizes of Haar –DCT in Hybrid Wavelet, Multi resolution and hybrid transform.

Fig. 17, 18 and 19 show comparison of AFCPV in Haar hybrid wavelet, its multi resolution hybrid wavelet and hybrid transform respectively. AFCPV gives change in perceived value of a pixel. Lower the AFCPV value better is the image quality.

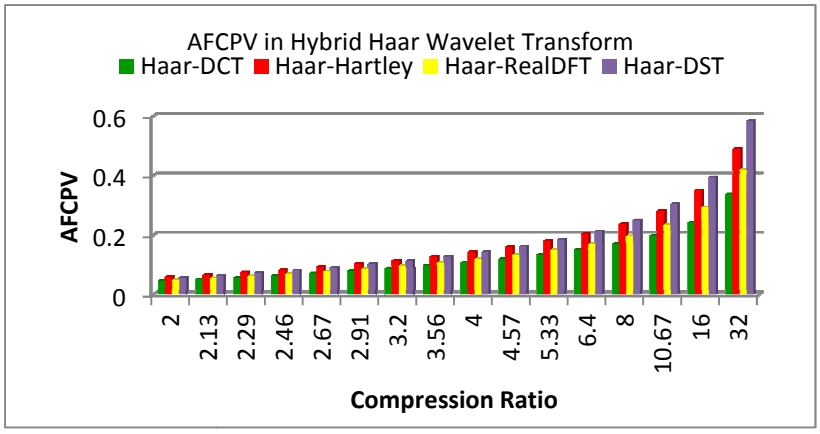


FIGURE 17: AFCPV vs. Compression ratio in Hybrid Wavelet using different Sinusoidal Transforms with Haar Transform.

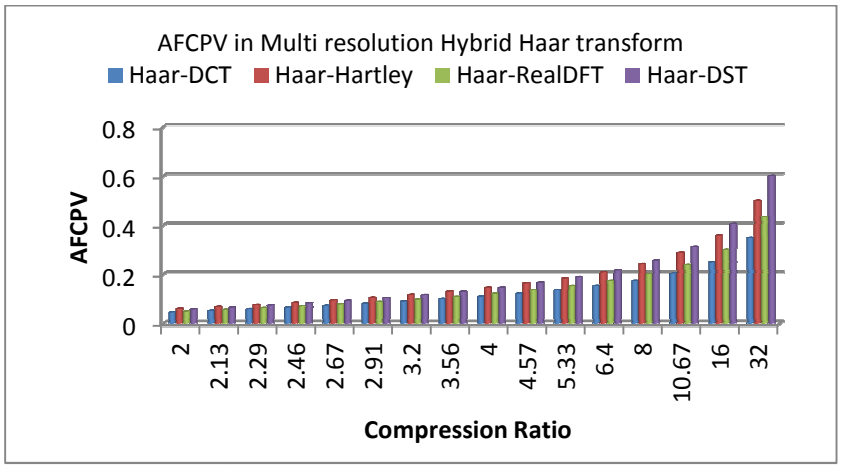


FIGURE 18: AFCPV vs. Compression Ratio in Multi resolution using different sinusoidal transforms with Haar transform.

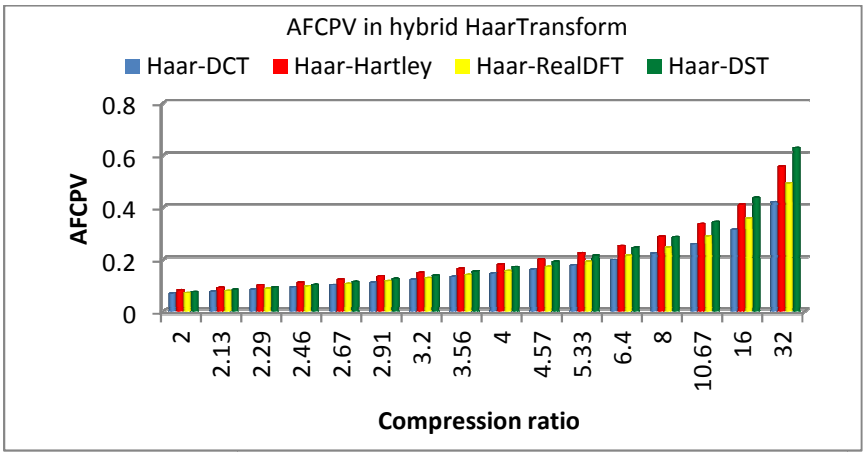


FIGURE 19: AFCPV against Compression ratio in hybrid Transform using different sinusoidal Transforms with Haar Transform.

Fig. 20, 21 and 22 show AFCPV in Haar-DCT using its Hybrid wavelet, Multi resolution analysis and hybrid transform with different component sizes. In hybrid wavelet, 32-8 pair of Haar-DCT gives low AFCPV. At compression ratio 32, Haar (32x32) and DCT (8x8) pair gives low AFCPV. In multiresolution 16-16 size gives better AFCPV up to compression ratio 16. Negligible difference is observed between AFCPV obtained using variations of component size. In Haar-DCT hybrid transform, 64-4 size gives lower value of AFCPV for all compression ratios.

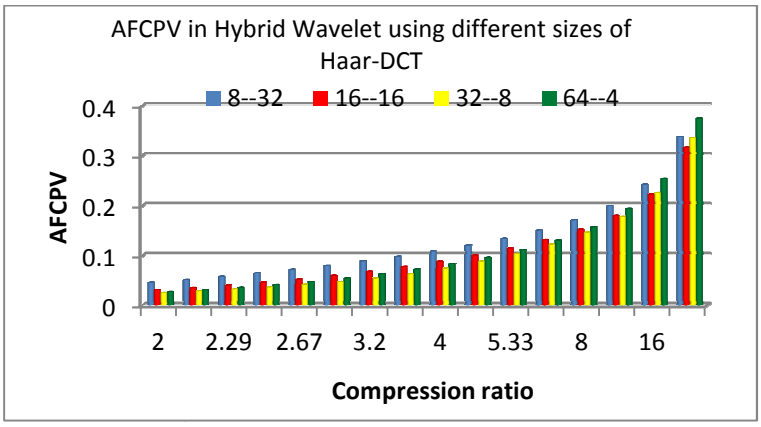


FIGURE 20: AFCPV vs. Compression Ratio in Hybrid Wavelet using different sizes of Haar-DCT.

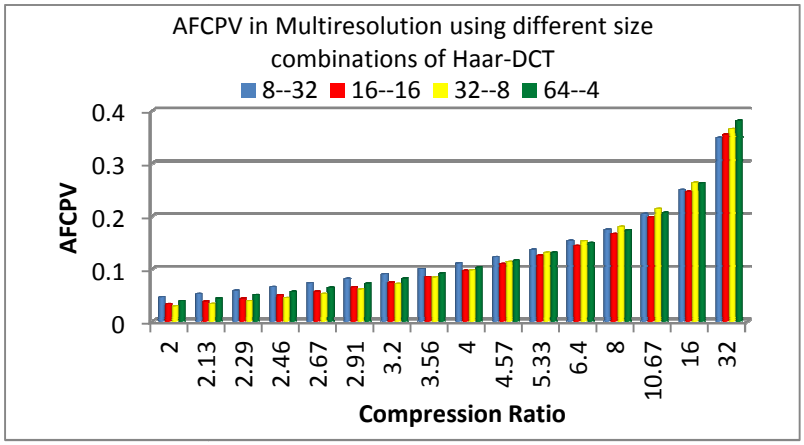


FIGURE 21: AFCPV vs. Compression Ratio in Multi resolution using different size combinations of Haar-DCT.

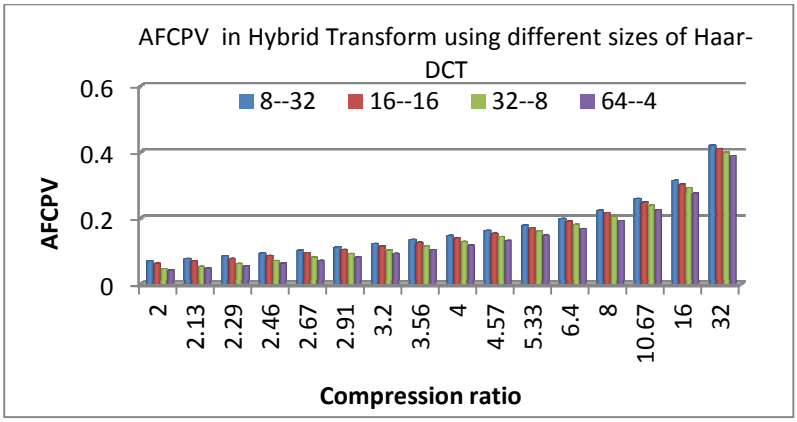


FIGURE 22: AFCPV vs. Compression Ratio in Hybrid Wavelet using different sizes of Haar-DCT.

Overall comparison of AFCPV in Haar-DCT hybrid wavelet, its multi resolution and hybrid transform is done in Fig 23. Hybrid wavelet again proves to be better for Haar 16x16 and DCT 16x16.

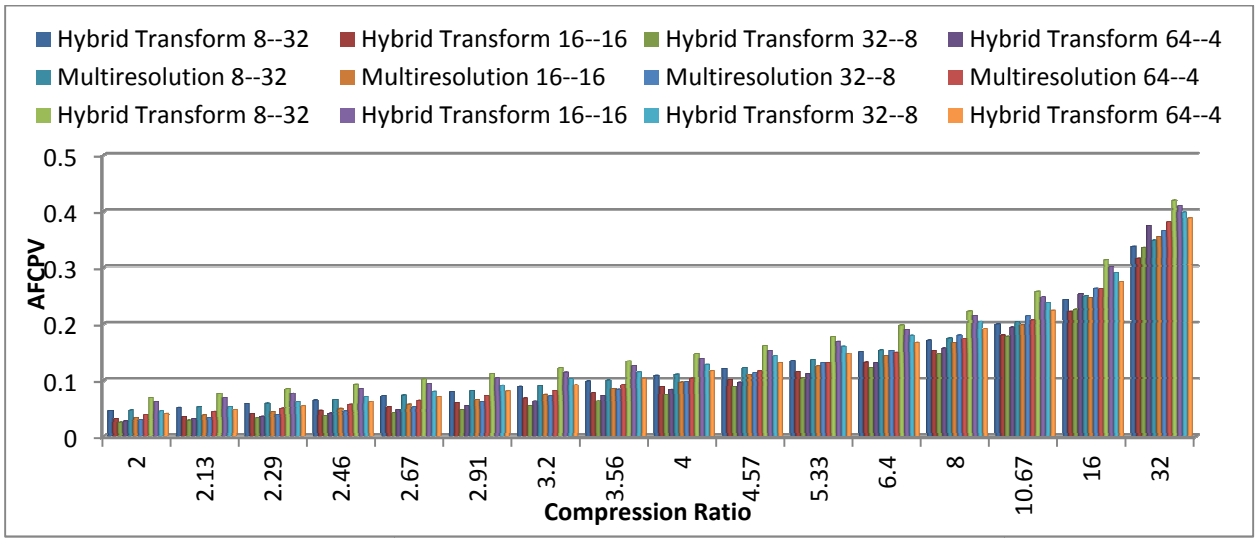


FIGURE 23: Comparison of AFCPV using various component transforms sizes in Haar-DCT Hybrid wavelet, Multi resolution hybrid wavelet and hybrid transform.

Fig. 24 shows Structural Similarity Index in Haar-DCT hybrid wavelet at various compression ratios for sample 'Lena' Image.

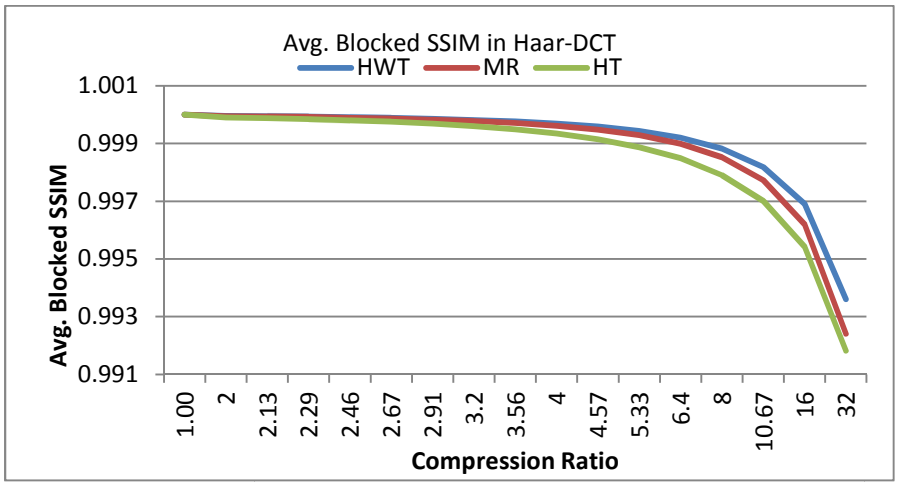


FIGURE 24: Average Blocked SSIM at different Compression Ratios in Haar-DCT hybrid Wavelet Transform, Its Multi resolution and Hybrid Transform with component size16-16 for 'Lena' image.

At compression ratio 1 i.e. when two images are exactly same, SSIM is one. As compression ratio increases, distortion in the reconstructed image increases and hence SSIM reduces. Image is divided into 16x16 blocks and SSIM of individual block is calculated to get more accuracy in perception of image to HVS. In hybrid wavelet SSIM is close to one at lower compression ratios and attains value 0.993 at compression ratio 32, indicating better perceived image quality. This objective measure of perceptibility is not obtained by traditional RMSE parameter. Multi resolution hybrid wavelet and hybrid transform show marginal difference in image quality than Hybrid wavelet in terms of SSIM.

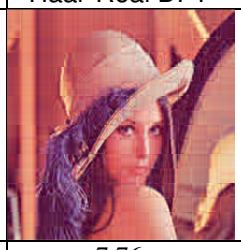
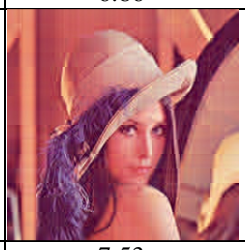
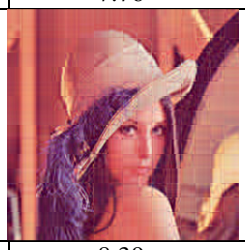
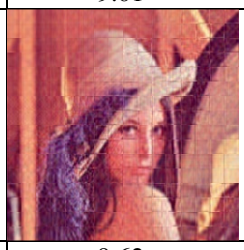
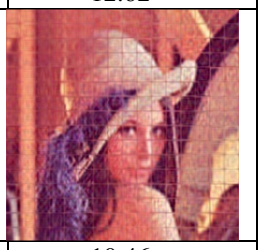
	Haar-DCT	Haar-Hartley	Haar-Real DFT	Haar-DST
Hybrid Wavelet				
MAE	6.86	8.55	7.76	13.75
Multi-resolution Hybrid Wavelet				
MAE	7.53	9.01	8.39	12.62
Hybrid Transform				
MAE	8.10	9.62	8.87	10.46

FIGURE 25: Reconstructed 'Lena' image at Compression ratio 32 using Hybrid Haar Wavelet, its Multi Resolution Analysis and Hybrid Transform with Haar 16x16 and different Local Component Transforms of Size 16x16.

Fig. 25 shows reconstructed Lena images at compression ratio 32 using different hybrid Haar wavelet transforms. Local component transform is varied keeping size of both components as 16x16. It shows that Haar-DCT pair gives lower MAE than other pairs. Also Hybrid wavelet transform gives less MAE than multi resolution analysis and hybrid transform. Haar-DST shows poor performance including grid effect in reconstructed image.

7. CONCLUSION

In this paper Hybrid Haar wavelet with bi resolution analysis, multi resolution analysis and with global features is implemented and compared. Sinusoidal transforms DCT, DST, Hartley and Real DFT are used as local component transforms and combined with Haar transform. Haar-DCT hybrid wavelet gives less error than Haar-DST, Haar-Hartley and Haar-Real DFT wavelet. In multi resolution analysis and hybrid transform also Haar-DCT gives less error. In Haar- DCT hybrid wavelet, size variation of component transforms is done to observe the changes in different fidelity criteria. AFCPV is used to measure fractional change in pixel values. It gives better perceptibility of image than traditional RMSE metric. Lowest AFCPV is obtained for Haar-DCT pair with size 16-16. SSIM is objective fidelity criteria used to approximate the perceived image quality. SSIM equal to 1 indicates that original and reconstructed images are exactly similar. Using hybrid Haar wavelet SSIM 0.993 is obtained at compression ratio 32. In multi resolution analysis and hybrid transform, values of SSIM are 0.992 and 0.991 for same compression ratio. Thus there is marginal difference in image quality when it is compressed using Hybrid wavelet, multi resolution and hybrid transform.

8. REFERENCES

[1] M. Ashok, T. Bhaskarareddy, "Image compression Techniques using Modified High Quality MultiWavelets", International Journal of Advanced Computer Science and Applications(IJACSA), Vol. 2, No. 7, pp. 153-158, 2011.

[2] R.S. Stanković and B.J. Falkowski. "The Haar wavelet transform: its status and achievements". Computers and Electrical Engineering, Vol. 29, No.1, pp.25-44, Jan. 2003.

[3] R. Wang. "Haar transform". http://fourier.eng.hmc.edu/e161/lectures/Haar/index.html, Dec 04, 2008.

[4] R. Mehala, K. Kuppusamy, "A New Image Compression Algorithm using Haar Wavelet Transformation ", In Proc. of International Conference on Computing and information Technology (IC2IT), International Journal of Computer Applications, pp. 1-4, 2013.

[5] S. Sridhar, V.Venugopal, S. Ramesh, S.Srinivas and Sk. Mansoob, "Wavelets and Neural Networks based Hybrid Image Compression Scheme", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), vol. 2, Issue 2, pp. 195-200, Apr.2013.

[6] Singara Singh, R.K. Sharma & M.K. Sharma, "Use of Wavelet Transform Extension for Graphics Image Compression using JPEG2000 Standard", International Journal of Image Processing (IJIP), Volume 3, Issue 1, pp. 55 to 60.

[7] Krishna Kumar, Basant Kumar & Rachna Shah, "Analysis of Efficient Wavelet Based Volumetric Image Compression", International Journal of Image Processing (IJIP), Volume 6, Issue 2, pp.113-122, 2012.

[8] J. Izadian , A. Hosaini & M. Jalili, "A Hybrid SVD Method Using Interpolation Algorithms for Image Compression", International Journal of Image Processing (IJIP), Volume 6, Issue 5, pp.273-282, 2012.

[9] H.B. Kekre, Tanuja Sarode, Prachi Natu, " Image Compression Using Column, Row and Full Wavelet Transforms Of Walsh, Cosine, Haar, Kekre, Slant and Sine and Their Comparison with Corresponding Orthogonal Transforms". International Journal of Engineering Research and Development, Vol. 6, Issue 4, pp. 102-113, Mar.2013.

[10] H.B. Kekre, Tanuja Sarode, Prachi Natu, "Image Compression using Real Fourier Transform, It's Wavelet Transform and Hybrid Wavelet with DCT". International Journal of Advanced Computer Science and Applications (IJACSA). Vol. 4, Issue 5, pp. 41-47, May 2013.

[11] Alan Brooks, Thrasyvoulos Pappas, "Using structural similarity quality Metrics to Evaluate Image Compression Techniques", In Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), vol. 1, (Honolulu, Hawaii), pp. I-873 - I-876, Apr. 2007.

[12] H.B. Kekre, Tanuja Sarode, Sudeep Thepade, "Inception of Hybrid Wavelet Transform using Two Orthogonal Transforms and its Use for Image Compression". International Journal of Computer Science and Information Security (IJCSIS). Vol. 9 Issue (6), pp. 80-87, Jun 2011.

[13] H.B. Kekre, Tanuja Sarode, Rekha Vig, "Multi-resolution Analysis of Multispectral palm prints using Hybrid Wavelets for Identification". International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 4 Issue 3, pp.192-198, Mar 2013.

Use of Discrete Sine Transform for A Novel Image Denoising Technique

Malini. S *malinivipin@gmail.com*
Marian Engineering College, Thiruvananthapuram
(Research center: L.B.S), 695 582, India

Moni. R. S *moni2006rs@gmail.com*
Professor, Marian Engineering College, Thiruvananthapuram
695 582, India

Abstract

In this paper, we propose a new multiresolution image denoising technique using Discrete Sine Transform. Wavelet techniques have been in use for multiresolution image processing. Discrete Cosine Transform is also extensively used for image compression. Similar to the Discrete Wavelet and Discrete Cosine Transform it is now found that Discrete Sine Transform also possess some good qualities for image processing; specifically for image denoising. Algorithm for image denoising using Discrete Sine Transform is proposed with simulation works for experimental verification. The method is computationally efficient and simple in theory and application.

Keywords: Denoising, Multiresolution, Image Transform, Discrete Sine Transform, Sub Bands.

1. INTRODUCTION

Different types of noise corrupt images during its capture, storage, transmission, etc. When noise creeps in image, its visual quality is reduced. Moreover, depending on the amount of noise, important features of the image become disguised and are not readily available for identification and use[1]. The only way to solve these problems is to denoise (remove noise) the image.

Several denoising techniques have been proposed during the last two/three decades. However, so called perfect denoising algorithm is yet to be developed; especially in the case of color images. Removal of noise is not yet complete in all respects [2].

One of the comparatively recent successful denoising techniques makes use of multiresolution capability of image transforms. It was originally proposed by Stephen Mallat (1988/89) with the use of Wavelet Transform (WT) for image processing [3]. However, WT is not the only transform which exhibit multiresolution capability [4, 5]. Discrete Sine Transform (DST) is now seen to possess multiresolution property. In this paper, Multiresolution DST is proposed for denoising of images. Rest of the paper is organized as Noise in Images in section 2, Some Denoising Techniques in Images in section 3, Multiresolution Techniques in Image Processing in section 4, DST and Multiresolution Analysis in section 5, Implementation of the DST Algorithm in section 6, Experimental results in section 7, and Conclusions in section 8.

2. NOISE IN IMAGES

Generally noise is generated along with the signal in most of the image production mechanisms. Photographic image is obtained using digital camera. Almost all cameras now available use Charge Coupled Devises (CCD) or Charge Injection Devices (CID) or Complementary Metal Oxide Semiconductor (CMOS) in its construction, making use of integrated circuit (IC) technology. Due to the imperfections in its fabrication, dark currents, unequal currents, clock

noise, etc. produce noise (Gaussian, Impulse, etc.) in the images photographed by these cameras [6, 7]. In addition, varying levels of illumination on the objects and back ground light also introduce noise in images. Further, signal conditioning (amplification, quantization, etc.) applied to the image signal add electronic device noise due to imperfect operation of active and passive devices. Medical images are another category of images which are widely used by clinicians for diagnostic purposes. Techniques such as Ultrasound (US), Magnetic Resonance(MR), Computed Tomography(CT), X-rays, Positron Emission Tomography(PET), etc. are used in the generation of medical images[8, 9]. All of these images contain different types of noise (speckle, Poisson, etc.) due to the imperfect operation of respective instruments. Satellite images are remote sensed images obtained by high resolution/ low resolution cameras and provide amazing information about earth's surface, roads, rivers, buildings, etc. These images are basically noisy. Variations in natural environments, limitations of cameras, etc are some of the reasons for noise in such images [10].

3. SOME DENOISING TECHNIQUES IN MAGES

Image denoising techniques can be broadly classified as i) Spatial Domain and ii) Frequency Domain. In spatial domain, pixel by pixel processing is done for removing noise [11]. Some kind of moving masks are used to locate each pixel in the image and neighborhood operation is done to filter the noise. Using suitable weights [12] in elements of mask, convolution or correlation is done with the neighborhood of pixel and elements of mask to get new pixel values for replacing the noisy pixel. This operation is known as filtering. The method works well in some environments. But it processes all pixels whether noisy or not and hence introduces some distortions. Noise detection before denoising improved the situation to some extent.[13, 14].

In the frequency domain method, the image is transformed in to its frequency domain, using suitable image transform and then processing is done. One of the early and widely used transform is 2dimensional Discrete Fourier Transform (2DDFT) implemented [1] using fft2 in MATLAB. Using this transform, the image pixel coefficients are represented by equivalent sine and cosine coefficients having different frequencies and amplitudes. This transform gives complex coefficients. Also DFT does not give information about the time at which a particular frequency component has occurred. Using short windows of time for groups of pixels in the image, and computing DFT (named as STFT), some amount of non- stationary analysis could be carried out. STFT has the limitation of constant resolution. A better method is to get variable resolution (multiresolution) for non-stationary images. Wavelet transform is the appropriate solution in this respect.[15] It has multiresolution analysis capability. Wavelet transform in its various forms (decimated, undecimated, orthogonal, biorthogonal, etc) and new generations (curvelet, contourlet, etc) have been in use in multiresolution approach for several image processing, including image denoising[16, 17] .

4. MULTIRESOLUTION TECHNIQUES IN IMAGE PROCESSING

Multiresolution refers to characteristics of the image analysis such that a feature in an image can appear at different resolutions and scales.[1, 3]. In multiresolution analysis, some features that go undetected at one resolution may be easy to spot at another resolution. Wavelet transform is a pioneer in this category of transforms. In its discrete implementation, known widely as Discrete Wavelet Transform (DWT), digital FIR filters are used to separate (decompose) the images in to its frequency bands.

In one level of decomposition, using filters and down sampling by two, four subimages are obtained. They are called as LL1, LH1, HL1 and HH1 bands. The band LL1 contains low frequency information and is called approximation sub band. The other three sub bands (LH1, HL1 and HH1) contain higher frequency information of the image and are called as detailed sub bands. Denoising of image relies on the strategy that noise is of higher frequency as well as noise is of smaller amplitudes compared to signal amplitudes. Hence the approximation coefficients (in LL1 band) more or less represent signal and some parts (low amplitude coefficients) of detailed sub bands represent the unwanted noise. To remove noise from image, we threshold[18] the

detailed sub bands, keeping higher amplitudes (signals)in these bands and removing the low amplitude (noise). Wavelet synthesis of LL1 band with thresholded detailed sub bands gives the denoised image. Results can be improved by decomposing LL1 band to the second level of wavelet decomposition giving rise to more detailed sub bands and processing these bands further on the same lines as first level decomposition.

5. DST and MULTIRESOLUTION ANALYSIS

The Discrete Sine Transform (DST) is related to DFT (and FFT). DST is used to represent signal in terms of a sum of sinusoids with its different frequencies and amplitudes [19]. However DST is not the imaginary part of FFT. To obtain DST of a digital signal, elements of the signal are reconfigured as an odd (anti-symmetric) extension of the input and then applying FFT. Total numbers of input as well as total number of DST coefficients are the same. All the DST coefficients are real, which is an advantage of DST processing over FFT (FFT gives complex coefficients). A companion transform of DST is Discrete Cosine Transform (DCT). DCT is widely used for image compression (JPEG).

The one dimensional DST of a vector of N elements can be obtained by reconfiguring the input and including the odd symmetric extension of its elements resulting in 2N elements. Depending on the symmetry used for extension of elements, there are 8 versions of DST (DST-1 to DST-8) equations. Each of them has somewhat different properties (in addition to some common properties)

The N point DST-1 is defined as in [19]

$$X(k)=\sum_{n=0}^{N-1} x(n)\sin\left[\frac{\pi}{N+1}(n+1)(k+1)\right] \qquad 0\le k\le N-1 \qquad (1)$$

where x(n) is the input signal.

The corresponding Inverse Discrete Sine Transform (IDST) is given by

$$x(n)=\sum_{k=0}^{N-1} X(k)\sin\left[\frac{\pi}{N+1}(n+1)(k+1)\right] \qquad 0\le n\le N-1 \qquad (2)$$

However, the DST is defined in MATLAB as

$$Y(k)=\sum_{n=1}^{N} x(n)\sin[\pi kn/(N+1)] \qquad 1\le k\le N \qquad (3)$$

and the corresponding IDST is

$$x(n)=\frac{2}{(N+1)}\sum_{k=1}^{N} Y(k)\sin[\pi kn/(N+1)] \qquad 1\le n\le N \qquad (4)$$

The 2D DST is an extension of 1D DST for two dimensional signals.

The 1D DST can be evaluated using fast algorithm of FFT. The DST available in MATLAB (dst) is 1D DST.

The DST has some more desirable properties as a transform for image processing. It has high energy compaction (most of the energy is confined to small number of coefficients) and sparse representation (large numbers of its coefficients are zeros.). Further, DST coefficients are real, symmetric and orthogonal. Symmetric and orthogonal property indicates that for forward and reverse transformation, computation is same except for normalization. As DST is a separable transform, the 2D DST can be implemented using twice the 1D DST. The 1D DST is first applied

column wise and its result is used as the input for a second 1D DST now row wise. It is a fast transform and hence computation time is less.

Multiresolution technique can be implemented using DST to separate the frequency components of an image in to one low frequency band and three high frequency bands (as in the case of wavelet transform) in its first level decomposition. Whereas filters are used in DWT multiresolution analysis, DST and IDST operations are done selectively so as to separate low frequency part of image and high frequency parts at each level. BY repeating the operation on the low frequency part of first level, further separation of the image into coarse and fine frequency can be done.

6. IMPLEMENTATION OF THE DST ALGORITHM

Image analysis is done as shown in Figure 1. It is based on the method proposed in [20] for DCT multiresolution application in image fusion. The 1D DST available in MATLAB is made use of. Frequency domain information of noisy image is obtained by applying DST column wise on the noisy image. For an MxN image, the image is divided into N columns of Mx1 image. DST operation is done column by column and the result is stored as two (M/2)xN matrices. The first half consists of frequency components from 0 to π/2 and the second half represents frequency components from π/2 to π. IDST is now applied column wise to the first half to get L1 band which represents the low frequency components of the image. Again IDST is applied column wise to the second half to get H1 band which represents the high frequency components of the image. Now DST is applied row wise to the L1 band. To the first half of this result, IDST is applied row wise to get the LL1 band. Application of IDST row wise to the second half gives LH1 band. Similar row wise IDST operations on the first half and second half after DST operation of H1 band gives HL1 and HH1 bands. This is the first level of decomposition. The process can be repeated by taking the LL1 band as the input to give four sub bands namely LL2, LH2, HL2 and HH2. The LL1 band consists of frequency components of 0 to π/2 row wise and 0 to π/2 column wise. The LH1 band consists of 0 to π/2 row wise and π/2 to π column wise. The HL1 band consists of π/2 to π row wise and 0 to π/2 column wise. The HH1 band consists of π/2 to π row wise and π/2 to π column wise.

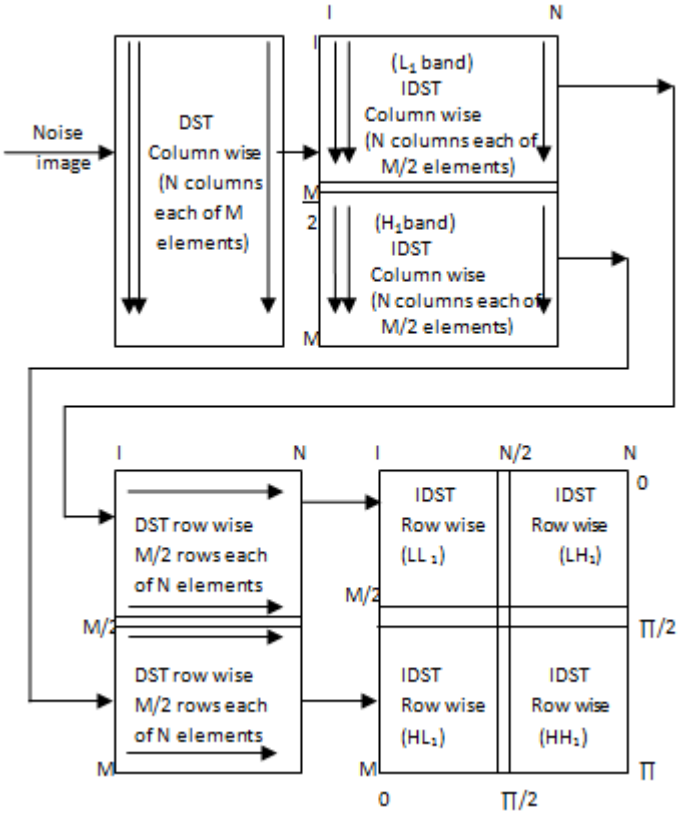


FIGURE 1: 1D DST Multiresolution Analysis.

If 2D DST is readily available, implementation of the algorithm can be modified as illustrated in Figure 2. The 2D DST is applied to the MxN size image. Four partitions are now made on the result. The first partition consists of 1 to M/2 rows and 1 to N/2 columns. IDST is applied to this partition giving rise to LL1 band. The second partition consists of 1 to M/2 rows and (N/2) +1 to N columns. IDST applied to this partition gives the LH1 band. The third partition consists of (M/2) +1 to M rows and 1 to N/2 columns. IDST applied to this partition gives HL1 band. The fourth partition is of (M/2) +1 to M rows and (N/2)+1 to N columns. IDST applied to this partition gives HH1 band. A second level of decomposition can now be obtained by repeating the steps with LL1 band as the input, giving rise to LL2, LH2, HL2and HH2 bands.

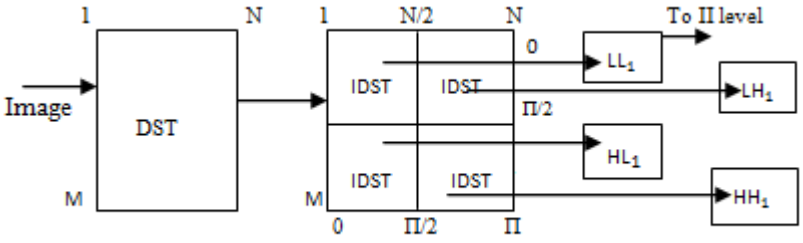


FIGURE 2: 2D DST Multiresolution Analysis.

The coarsest sub band is LL1. It consists of low frequency DST coefficients of the image and is the approximation band. All very important features of the image are represented by this band. Magnitudes of the coefficients in this band are larger than that in any other bands. Noise in this band can be assumed to be very small as the noise is generally of high frequency .The other three sub bands (LH1, HL1 and HH1) consist of DST coefficients representing high frequency

information such as edges, curves, etc. Noise is also present in these bands. However coefficients representing noise are of smaller magnitudes compared to that of signal in these bands. The denoising strategy is that the low amplitude noisy coefficients are removed from these bands by suitable thresholding [1]. According to the requirements, hard thresholding or soft thresholding can be employed. Adaptive thresholding [18] is employed in this paper. The thresholded detailed sub bands and the approximation sub band are synthesized to get the denoised image. As the approximation sub band has most of the features of the original image, except for some high frequency details, resizing of approximation band to size of original image is to some extent the denoised image.

DST is as old as DCT. However whereas DCT has been extensively used for image processing(eg. image compression; JPEG) DST is not seen used, at least as much as DCT, for signal processing, although both have many similar properties [19]. In this paper, utility of DST is brought to the lime-light in the form of a multiresolution signal processing tool; specifically for image denoising. Implementation of the algorithm is much simpler than many of the recently proposed algorithms. In spite of this, attractive feature is that the performance, especially in visual quality, is as good as many of the existing methods. Computational complexity is less and hence, time taken for implementation is comparatively low.

7. EXPERIMENTAL RESULTS

Multi resolution DST is applied to the “cameraman” image and the first level of decomposition is done. Resulting LL1 band is again decomposed to get the second level of analysis. Results of the two levels of decomposition along with the original image are shown in Figure 3. Gaussian noise is added to the “Barbara” image with zero mean and 10 % standard deviation. The noisy image is decomposed to get the LL1, LH1, HL1 and HH1 bands. Denoising is done as explained in section 7. Results are shown in Figure 4. A 10 % impulse noise is added to “circles” image and denoising is done. Results are shown in Figure 5. Again impulse noise is added to “rice” image and denoising is done. Results are shown in Figure 6.

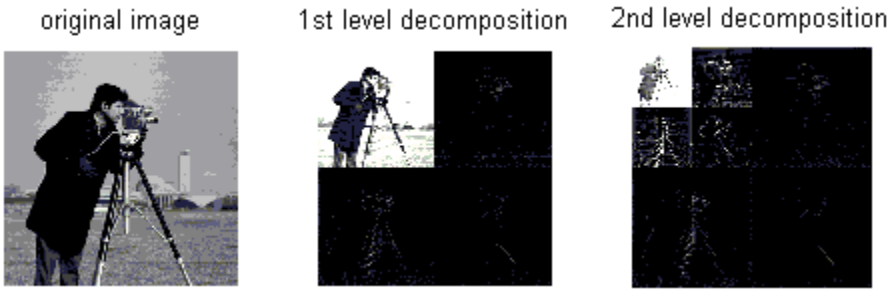


FIGURE 3: Two Level DST Decomposition.

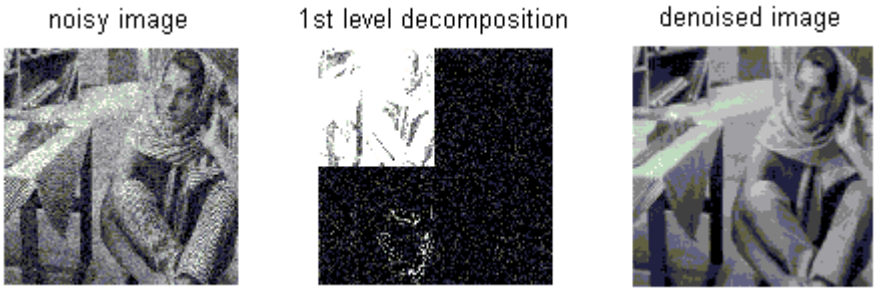


FIGURE 4: Denoising of Gaussian Noise.

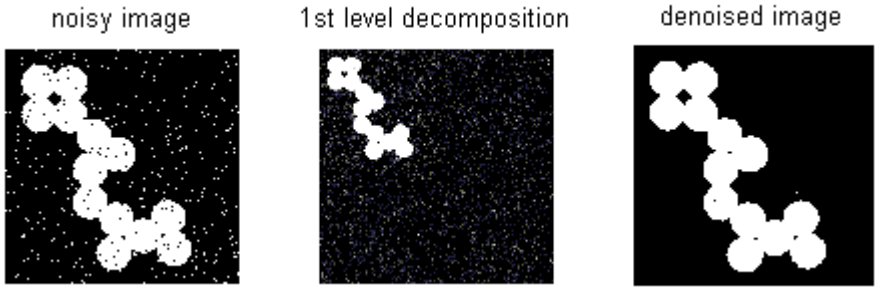


FIGURE 5: Denoising of Impulse Noise.

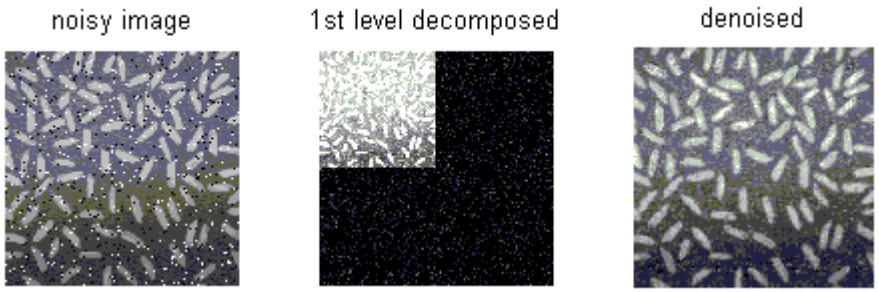


FIGURE 6: Denoising of Impulse Noise.

For color image (RGB model) denoising, as there is correlation between color components, color model has to be changed (to Ycbcr, HSV, etc.) and then processing has to be done. A color image ("autumn') is used for demonstration of denoising. Speckle noise is added in this image. The resulting RGB image is converted to Ycbcr image. After denoising process on Y components, with the multiresolution DST, Ycbcr image is converted back to RGB image. Experimental results of denoising are shown in Figure 7. Results show that noise is removed, but there is small amount of color artifacts. Further work is to be done to rectify this defect.

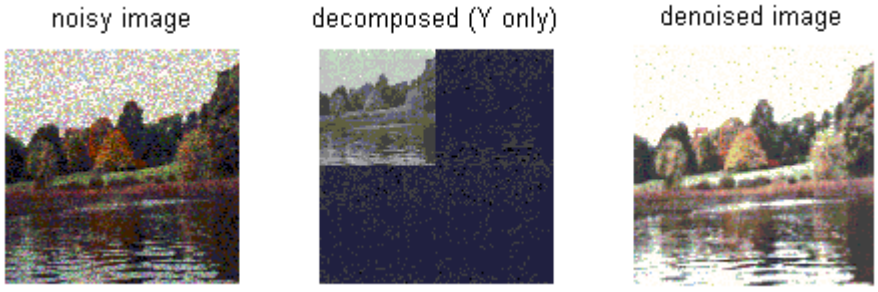


FIGURE 7: Denoising of Speckle Noise from Color Image.

Performance comparison of the proposed denoising method with that of a few of the existing methods are done (1) by observing the visual quality and (2) by measuring the Peak Signal to Noise Ratio (PSNR) of denoised image. The metric PSNR is defined as

$$PSNR = 10 \log_{10} [\ 255 * 255 / MSE] \quad dB \tag{5}$$

$$\text{where MSE (Mean Square Error)} = \text{sum (sum ([} Y(i,j) - Y_1(i,j)]^2)) / (M*N*D) \tag{6}$$

Y(i , j) is the original (non-noisy) image , Y_1 (i , j) is the denoised image and M, N,D are the size of the image. For Gaussian denoising, noise density is varied from 10 to 40 in steps of 10 and for each value, PSNR of Gaussian denoised image is measured. The respective PSNR for three existing methods wavelet thresholding [18], Multiresolution Bilateral Filter [21] and LPG-PCA are also obtained. Table 1 shows these PSNR values. It can be seen that the proposed method has PSNR values better than or very close to the other existing methods referred. However the second and third methods are computationally more involved (compared to the proposed method) and take much more execution time.

Methods	**PSNR**			
	σ = 10	**σ = 20**	**σ = 30**	**σ = 40**
Wavelet Threshold	30.8	27.8	25.4	23.8
Multi resolution Bilateral filter	34.1	31.5	28.8	27.2
LPG-PCA	33.2	30.5	27.2	25.3
Proposed Method	33.4	30.4	27.2	25.4

TABLE 1: PSNR of Gaussian noise removal (Barbara Image).

For Impulse denoising experiment, impulse noise is added to "circles" image with a noise level from 10% to 40 % in steps of 10 in the proposed method and PSNR are measured in each case. The procedure is repeated for "rice" image also. The respective PSNR using three existing methods, Median filter, Adaptive median filter and Modified Decision Based [14] method are also obtained. Table 2 shows all the PSNR values. It can again be seen that the performance of the proposed method is better than the first two existing methods and very close to that of the third method. However the third method is computationally much involved and requires more computer time.

Image	**"Circles"**				**"Rice"**			
Methods	**PSNR for noise levels**				**PSNR for noise levels**			
	10%	20%	30%	40%	10%	20%	30%	40%
Median Filter	26.2	25.1	24.0	22.1	27.8	26.3	25.1	24.0
Adaptive Median Filter	30.1	29.6	27.4	26.1	32.3	31.1	29.6	28.2
Modified Decision Based	32.5	31.8	30.4	28.7	33.2	32.4	30.1	29.8
Proposed Method	31.2	29.4	28,5	27.4	32.6	31.4	29.7	28.6

TABLE 2: PSNR of Impulse noise removal (Circles and Rice images) .

In terms of visual quality, the proposed method has superior visual quality compared to the existing methods referred.

8. CONCLUSIONS

Multiresolution image denoising using DST is proposed in this paper as a new method of image denoising. It has close similarity with DWT denoising. Algorithm for generating sub bands of low frequency and high frequencies using DST and IDST is given. As MATLAB does not provide function for 2D DST, implementation of algorithm using 1D DST is explained. Also if 2D DST is readily available, algorithm of implementation of denoising using 2D DST and 2D IDST is also given. Strategy of denoising is also discussed. It has close similarity with that using DWT. Approximation sub band which is substantially the original signal and thresholded detailed sub bands are synthesized to get denoised image. If some reduction in the high frequency component of the image can be tolerated, the approximation band can represent the denoised image. Experimental results using MATLAB simulation are given. It is seen that the method can remove different types of noises such as Gaussian, Impulse, Poisson, Speckle, etc. The visual qualities of the denoised images are good while its computational complexity is low. Future work has to be for getting good quality denoised color images.

This work might be extended to add sophistications to further reduce noise when the noise level is very high. Also a universal denoising algorithm can be pursued to denoise images when they are corrupted simultaneously by different types of noises. Further, improvement in the qualities of color images by modifying the proposed work could be a worth future work.

9. REFERENCES

[1] R.C.Gonzalez, R.E.Woods, "*Digital Image Processing"* Pearson, 3rd Edition, India, 2009.

[2] P.Chatterjee and P.Milanfar, "Is Denoising Dead", IEEE Transactions on Image Processing, vol.19, No.4, pp. 895-910, Apr. 2010.

[3] S.Mallat, "A theory of multi resolution signal decomposition, the wavelet representation", IEEE Transactions Pattern Anal Machine intelligence, Vol.PAM1-11, pp. 674-693., Jul 1989.

[4] Rajeesh, R.S.Moni, S.Palanikumar and T.Gopalakrishnan "Noise reduction in Magnetic Resonance Images using Wave atom shrinkage" International Journal of Image Processing (IJIP) Vol.4, Issue 2, pp 131-141, Jun 2010.

[5] M.Do, M.Vitterli "The Contourlet transform: an efficient directional multiresolution image representation" IEEE Transactions on Image Processing, Vol. 14. Issue 12, pp2091-2106, 2005.

[6] T.R.Vijayakumar, P.T.Vanathi, P.K.Kanagasabhapathy "Fast and efficient algorithm for removal of Gaussian noise in digital images" IAENG International Journal of computer Science, 37:1, Feb.2010.

[7] A.Buades, B.Coll and J.Morel "A review of image denoising algorithm with a new one" Multiscale modeling simulation ,Vol. 4, Issue. 2, pp490-530, 2005.

[8] A.Pizuria, W.Philips, I.Lemahieu and Acheroy "A versatile wavelet domain noise filtering technique for medical imaging" IEEE Transactions on medical imaging, Vol.22, Issue 3, pp 323-331, 2003.

[9] Z.Yang and M.D.Fox, "Speckle reduction and structure enhancement by multichannel median boosted anisotropic diffusion", EURASIP J. Appl. Signal processing, Vol.2004, issue. 1, pp2492-2502, Jan.2004.

[10] Lizy Abraham, M.Sasikumar "Unsupervised building extraction from high resolution satellite images irrespective of roof top structures" International Journal of Image Processing (IJIP), vol.6, Issue.4, pp 219-232, Aug. 2012.

[11] T.S.Huang and G.Y.Tang "A fast two-dimensional Median Filtering algorithm" IEEE Transactions on Acoustics, speech and signal processing, ASSP-27, pp13-18, 1979.

[12] Sung-Jea and Y. H. Lee "Center weighted median filters and their applications to image enhancement" IEEE Transactions on circuits and systems , Vol.38, No.9, pp984-993, Sep.1991.

[13] S. Zhang and M.A. Karim "A new impulse detector for switching median filters",IEEE Signal Processing Letters, Vol.9, No.11, pp 360-363, Nov.2002.

[14] S.Esakirajan, T.Veerakumar,A.N.Subramanyan and C.H.PremChand "Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter", IEEE Signal Processing Letters , Vol. 18, No.5, 11287-290, May 2011.

[15] Denver, I.K.Fodor, and C.Kamath "Denoising through wavelet shrinkage, an empirical study", Electronic Imaging, Vol.12, pp151-160, 2003.

[16] A.L..Pogam, H.Hanzouli, M.Halt, C.C.L.Rost and D.Visvikis "Denoising of PET images by combining wavelets and curvelets for improved preservation of resolution and quantization" Medical Imaging Analysis, Vol.17, Issue 3, pp877-891, Dec.2013.

[17] J.Starck, E.J.Cande and D.L.Donho "The curvelet transform for image denoising", IEEE Transactions on Image Processing, Vol.11, Issue 6, pp 670-684, June 2002.

[18] S.G.Chang, B.Yu, V.M.Vitterli, "Adaptive wavelet thresholding for image denoising and compression", IEEE Trans. Image Processing, Vol.9, No.9, pp. 1532-1546., Sep.2000.

[19] A.K.Jain, Fundamentals of Digital Image Processing ,PHI, India, 1995.

[20] V.P.S.Naidu, "Discrete cosine Transform based Image Fusion", Defence Science Journal, Vol.60, No.1, pp.48-54., Jan.2010.

[21] M. Zhang and B.K.Gunturk "Multiresolution Bilateral filter for image denoising", IEEE Transactions on Image Processing, Vol.17, No.12,pp 2324-2333, Dec.2008.

[22] L.Dhang, W. Dong D.Zhang and G.Shi "Two stage image denoising by principal component analysis with local pixel grouping" *Pattern Recognition,* Vol.43, pp1531-1549, 2010.

DCT and Simulink Based Realtime Robust Image Watermarking

Durgansh Sharma *durgansh@gmail.com*
Ph.D. Student, UPES, Dehradun India
Asst. Professor, Jaipuria Institute of Management, Noida, India

Manish Prateek *mprateek@ddn.upes.ac.in*
Professor, Centre For Information Technology,
College of Engineering Studies, UPES, Dehradun, India

Tanushyam Chattopadhyay *t.chattopadhyay@tcs.com*
Senior Scientist, R&D, Innovation Lab,
Tata Consultancy Services, Kolkata, India

Abstract

Ownership of digital content has become a serious matter, due to the exponential raise in the global repository of digital multimedia content, like images are to be considered in this paper. The validated proof as an imperceptible and robust watermark is needed to be embedded in the digital images. This paper proposes a simulation of DCT with Fuzzy Logic based HVS model for Realtime Robust Image Watermarking technique using Simulink.

Keywords: Realtime, Digital Image Watermarking, Fuzzy logic, HVS, Simulink.

1. INTRODUCTION

Digital image watermarking is a process used for embedding a set of information in an image for enhancing its authenticity. It has been used through various optimized techniques but mainly on gray scale images already available in the repository. In this paper, we propose a simulink based model for realtime robust image watermarking. The robustness of the process of image watermarking includes DCT and Fuzzy Inference System, which implements HVS (Human Visual System) for embedding the watermark in an image, captured using webcam in realtime. It transforms the captured image from spatial domain to frequency domain using DCT and applies fuzzy logic to implement the HVS logic for embedding the watermark in the host image.

Within a span of few years, we have observed the exponential raise in the usage of digital media for distribution of images. This model could be extended for the realtime image watermarking in camera enabled mobile devices to improve the authorization of the image distributed using digital media which seems to have no boundaries with the open ended environment called internet.

2. REVIEW

The model in this paper proposes the technique to incorporate the authorization watermark in the host image captured in real time. Motwani et al. [1] used MAMDANI type Fuzzy Inference System (FIS), its input parameters are derived from Human Visual System (HVS) using the sensitivity towards brightness, edge and contrast of the gray scale image which has been further improved to use blue frame of a colour image captured using webcam. Charu et al. [2] further used and extended the research work using the three layered Fuzzy-BPN having a layer configuration of (3-3-1) for learning mechanism through 50 iterations. Charu et al. [2] divided an image of size 256x256 into 1024 blocks of size 8x8 and compute its sensitivity, on the basis of the variance computed using Fuzzy-BP, the blocks were filtered and the random sequence of numbers are embedded as watermark. This procedure generated a good quality imperceptible

watermarked image, but, in this paper we try to insert the system identification numbers using the suggested techniques of Zhao, Jian, and Eckhard Koch [5][6]. Saraju et al.[3] has suggested a simulink based realtime perceptual watermarking architecture for video broadcasting through FPGA. Saraju et al. [3] has suggested very useful methods of combining cryptography and watermarking for better results in hiding useful data in a video frame as a visible transparent watermark, and Diffie et al. [7] has suggested various new ways of cryptography which can give better results in data hiding. We try to use his suggestive method in an image captured in a single shot through webcam using simulink.

3. MODELLING AND ANALYSIS

We consider a 640x480 pixel image captured using webcam for this presented work. The characteristics are modeled using Eckhard et. al. [5, 6] presented that the multimedia data must contain a label or code, which could identify it uniquely as property of the copyright holder. The embedded watermark extracted from the signed image using proposed algorithm was compared for the similarity correlation using SIM(X, X^*) proposed by Cox et al. [4]. this parameter is determined for recovered watermark. Computed values show a good significance level of optimization in the process of embedding and extraction of watermark.

Sharma et al. [9] proposed Fuzzy inference system (FIS) is used to embed the watermark in the host image in the DCT domain. The FIS is based on a set of 27 inference rules using SIGMOID way of interpreting the logical inputs, based on the facts of HVS based sensitivity towards noise in the image with respect to brightness, texture or contrast, edges.

Rule No.	Luminance Sensitivity	Contrast Sensitivity	Edge Sensitivity	Weighting Factor
1	DARK	LOW	SMALL	LEAST
2	DARK	MEDIUM	SMALL	LEAST
3	DARK	HIGH	SMALL	LEAST
4	MEDIUM	LOW	SMALL	LEAST
5	MEDIUM	MEDIUM	SMALL	LEAST
6	MEDIUM	HIGH	SMALL	LEAST
7	BRIGHT	LOW	SMALL	LEAST
8	BRIGHT	MEDIUM	SMALL	LEAST
9	BRIGHT	HIGH	SMALL	LEAST
10	DARK	LOW	MEDIUM	LESS
11	DARK	MEDIUM	MEDIUM	HIGH
12	DARK	HIGH	MEDIUM	HIGHER
13	MEDIUM	LOW	MEDIUM	LESS
14	MEDIUM	MEDIUM	MEDIUM	AVERAGE
15	MEDIUM	HIGH	MEDIUM	AVERAGE
16	BRIGHT	LOW	MEDIUM	LESS
17	BRIGHT	MEDIUM	MEDIUM	AVERAGE
18	BRIGHT	HIGH	MEDIUM	HIGHER
19	DARK	LOW	LARGE	LESS
20	DARK	MEDIUM	LARGE	HIGHER
21	DARK	HIGH	LARGE	HIGHEST

22	MEDIUM	LOW	LARGE	LESS
23	MEDIUM	MEDIUM	LARGE	AVERAGE
24	MEDIUM	HIGH	LARGE	HIGHER
25	BRIGHT	LOW	LARGE	LESS
26	BRIGHT	MEDIUM	LARGE	HIGHER
27	BRIGHT	HIGH	LARGE	HIGHEST

TABLE 1: HVS based 27 Rules for Fuzzy Inference System.

In this paper, we propose a technique to embed imperceptible watermark in an image in realtime. The model constitutes of webcam needed to acquire an image in realtime, Matlab version 8.0 with Simulink running on a computer. The acquired image constitutes of RGB colour frames; we propose to extract the Blue frame for embedding the watermark in it and merge it with other two Red and Green colour frames to reconstitute the image. This process makes the image watermarking robust and optimized.

The host image captured in realtime in spatial domain having the size of 640x480 pixels is divided into the blocks of 8x8 pixels each. Discrete Cosine Transformation (DCT) is used for the transformation of these blocks in the frequency domain. All the three HVS characteristics mentioned formerly are computed over these blocks as follows:

The Luminance Sensitivity: It is derived from the DC coefficients from the DCT blocks of the host image according to following formula:

$$L_i = \frac{X_{DC,i}}{X_{DCM}} \qquad (1)$$

Where, $X_{DC,i}$ denotes the DC coefficient of the i^{th} block and X_{DCM} is the mean value of the DC coefficients of all the blocks put together.

The Contrast Sensitivity: The contrast sensitivity is derived from the texture content of a region of 8x8 blocks in an image. The value of variance computed of an image block is provided to the direct metric for the quantification of the texture as a parameter. A routine proposed by Gonzalez et. al. [9] is used through MATLAB. The execution of this routine is given by (2).

```
t = statxture(f)
```
(2)

where, f is the input image or the sub-image (block) and t is the 7 – element row vector, one of which is the variance of the block in question.

The Edge Sensitivity: The edge could be detected in an image using the threshold operation; edge sensitivity can be quantified as a natural effect to the calculation of the block threshold T. The Matlab image processing toolbox implements bluethresh(f) routine which computes the block threshold using histogram – based on Otsu's method of computing graythresh(f) [8]. The implementation of this routine is given by (3)

```
T = bluethresh(f)
```
(3)

Where, f is the host sub-image (block) in question and T is the computed threshold value. These three parameters are fed into the FIS shown in Figure 3.

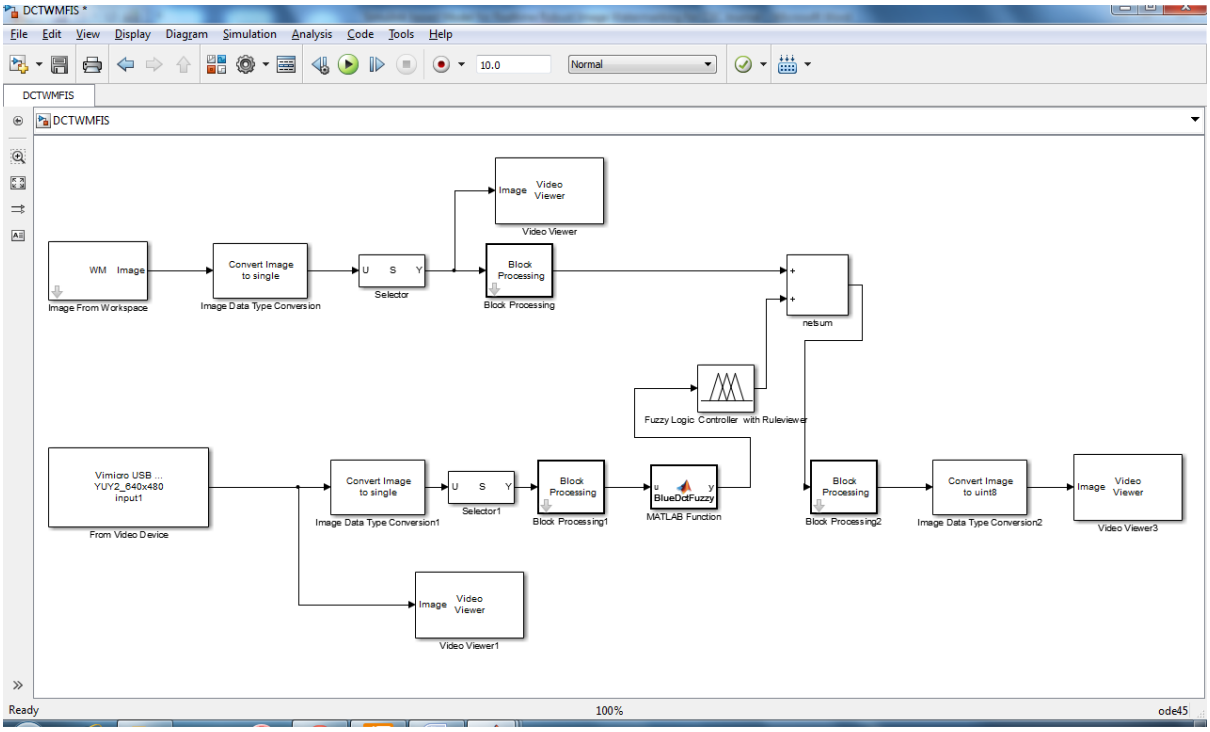


FIGURE 1: Simulink Model for DWM.

Once the watermark is embedded then Quality assessment of the signed image is done by computing Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR).

Extracting Watermark from Signed Image and Computing SIM(X, X^*) Parameter: Firstly, the DCT of both host and signed images are computed block wise. Thereafter, the computed coefficients are subtracted from each other and the watermark is recovered. Let the original and recovered watermarks be denoted as X and X^* respectively. A comparison check is performed between X and X^* using the similarity correlation parameter given by eq. (1).

$$SIM(X, X^*) = \frac{\sum_{i=1}^{n}(X, X^*)}{\sum_{i=1}^{n}\sqrt{(X, X^*)}} \qquad (1)$$

4. RESULTS

The profile summary shows the time consumed in the entire process of realtime image watermarking using the proposed method.

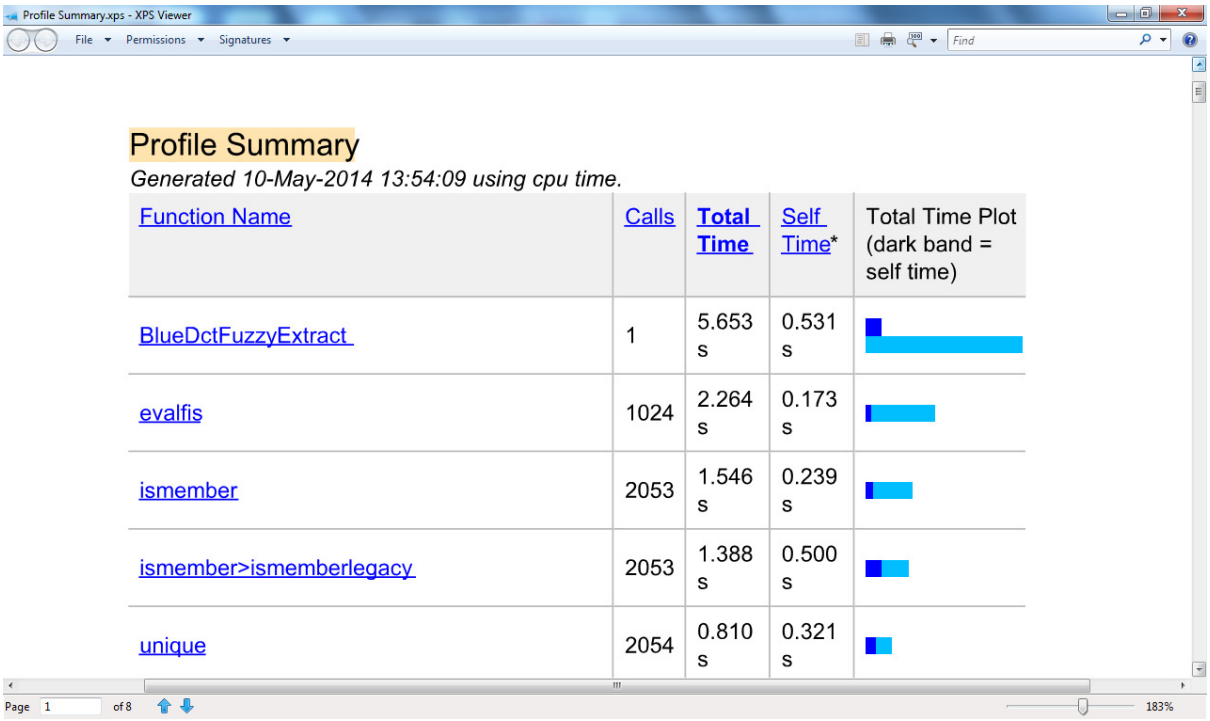


FIGURE 2: Profile Summary of Time Consumed.

Following are the result of watermark embedding process adopted in this paper using SIMULINK

FIGURE 3: Original Image Captured in Realtime.

FIGURE 4: Watermarked Image in Blue Channel.

MSE: 4.2053dB; PSNR: 41.957dB

5. CONCLUSION

Computed value of SIM(X, X*) parameter for the image depicted in Figure 4 (Singed Image) is 18.5987 which indicates a good watermark recovery process. The time consumed in image watermarking is computed as approx. 12 seconds, this model could be extended for the realtime digital image watermarking in camera enabled mobile devices for improving the authenticity of images captured and shared using Smartphone.

6. REFERENCES

[1] M.C. Motwani, and C.H. Frederick Jr. "Fuzzy Perceptual Watermarking For Ownership Verification." In IPCV, pp. 321-325. 2009.

[2] C. Agarwal, and A. Mishra. "A Novel Image Watermarking Technique using Fuzzy-BP Network." In Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on, pp. 102-105. IEEE, 2010.

[3] S.P. Mohanty, and K. Elias "Real-time perceptual watermarking architectures for video broadcasting." Journal of Systems and Software 84.5 (2011): 724-738.

[4] I.J. Cox, K. Joe, F. T. Leighton, and S. Talal "Secure spread spectrum watermarking for multimedia." Image Processing, IEEE Transactions on 6, no. 12 (1997): 1673-1687.

[5] J. Zhao, and K. Eckhard "Embedding Robust Labels into Images for Copyright Protection." In KnowRight, pp. 242-251. 1995.

[6] J. Zhao, and K. Eckhard. "Towards robust and hidden image copyright labeling." In IEEE Workshop on Nonlinear Signal and Image Processing, pp. 452-455. Neos Marmaras, Greece, 1995.

[7] W. Diffie and M. Hellman, “New directions in cryptography”, IEEE Transactions on Information Theory, vol. IT-22, pp. 644-654, 1976.

[8] R.C. Gonzalez, R.E. Woods and S.L. Eddins, Digital Image Processing Using MATLAB, Pearson Education (2005), pp 406 and 467.

[9] D. Sharma, M. Prateek, T. Chattopadhyay, “Optimized Robust Image Watermarking”, Proceedings of 4th International Conference on Emerging Trends in Engineering & Technology, October 25th -27th, 2013, IETET (2013), pp 99-106.

INSTRUCTIONS TO CONTRIBUTORS

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

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

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

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 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
- Photographic emulsions
- Autonomous vehicles
- Chemical and spectral sensitization
- Coating technologies
- Cognitive aspects of image understanding
- Communication of visual data
- Display and printing
- Generation and display
- Image analysis and interpretation
- Image generation, manipulation, permanence
- Image processing: coding analysis and recognition
- Imaging systems and image scanning
- Latent image
- Network architecture for real-time video transport
- Non-impact printing technologies
- Photoconductors
- Photopolymers

- Prepress and printing technologies
- Remote image sensing
- Storage and transmission
- Protocols for packet video
- Retrieval and multimedia
- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 9 - **Issue: 1**

i. Submission Deadline : November 30, 2014 **ii. Author Notification:** December 31, 2014

iii. Issue Publication: January 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 © 2014
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