

Volume 3, No 1; February 2015

Advances in Image And Video Processing

ISSN: 2054-7412

TABLE OF CONTENTS

EDITORIAL ADVISORY BOARD	I
DISCLAIMER	II
An Algorithm with Low Complexity for Image Compression and its Hardware Implementation using VHDL Shaktijeet Mahapatra and Mihir Narayan Mohanty	1
A 100 MHz 6th Order Continuous Time Band-Pass Sigma Delta Modulator with Active Inductor based Resonators Kevin Dobson, Shahrokh Ahmadi and Mona Zaghloul	10
Real Time Human Action Recognition using Kinematic State Model Geetanjali V. Kale and Varsha H. Patil	17
Predicting the Identity of a Person using Aggregated Features of Handwriting Revathi S V and Vijaya M S	23

EDITORIAL ADVISORY BOARD

Dr Zezhi Chen

Faculty of Science, Engineering and Computing; Kingston University London
United Kingdom

Professor Don Liu

College of Engineering and Science, Louisiana Tech University, Ruston,
United States

Dr Lei Cao

Department of Electrical Engineering, University of Mississippi,
United States

Professor Simon X. Yang

Advanced Robotics & Intelligent Systems (ARIS) Laboratory, University of Guelph,
Canada

Dr Luis Rodolfo Garcia

College of Science and Engineering, Texas A&M University, Corpus Christi
United States

Dr Kyriakos G Vamvoudakis

Dept of Electrical and Computer Engineering, University of California Santa Barbara
United States

Professor Nicoladie Tam

University of North Texas, Denton, Texas
United States

Professor Shahram Latifi

Dept. of Electrical & Computer Engineering University of Nevada, Las Vegas
United States

Professor Hong Zhou

Department of Applied Mathematics Naval Postgraduate School Monterey, CA
United States

Dr Yuriy Polyakov

Computer Science Department, New Jersey Institute of Technology, Newark
United States

Dr M. M. Faraz

Faculty of Science Engineering and Computing, Kingston University London
United Kingdom

DISCLAIMER

All the contributions are published in good faith and intentions to promote and encourage research activities around the globe. The contributions are property of their respective authors/owners and the journal is not responsible for any content that hurts someone's views or feelings etc.

An Algorithm with Low Complexity for Image Compression and its Hardware Implementation using VHDL

Shaktijeet Mahapatra¹ and Mihir Narayan Mohanty²

Department of ECE, Institute of Technical Education and Research (ITER), Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India

¹shaktijeetmahapatra@soauniversity.ac.in; ²mihirmohanty@soauniversity.ac.in

ABSTRACT

Image compression is highly essential for efficient transmission and storage of images in the field of communication engineering, bio-medical applications. Also, the compression technology is of special interest for the fast transmission and real-time processing on the internet. For reduced form and less capacity, the area of research growing day by day. The objective of image compression is to find a new representation in which pixels are less correlated, but with the original contents. In this paper, the existing as well as new algorithms are applied for compression for evaluation. The results have been compared for both techniques. On the basis of evaluating and analyzing the image compression techniques it presents the VHDL implementation of low complexity 2D-DWT approach applied to image compression. The decompression has to invert the transformations applied by the compression to the image data. When using the wavelet transform it is possible to exploit the unique properties of the wavelet coefficients to efficiently encode them.

Keywords. Image compression, real-time processing, VHDL implementation, 2D-DWT.

1 Introduction

Images constitute the significant part of data, particularly in remote sensing, biomedical and video conferencing applications. Image compression is the reduction or elimination of redundancy in data representation in order to achieve reduction in storage and communication cost. Image compression techniques reduce the number of bits required to represent an image by taking advantage of these redundancies. Image compression addresses the problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage/transmission requirements.

As the bandwidth requirement increase, the transmission or storage cost also increase simultaneously, so it is necessary to employ compression techniques, which reduce the data rate while maintaining the subjective quality of the decoded image or video signal. Even with the advances in bandwidth and storage capabilities, if images were not compressed many applications would be too costly. With the use of digital cameras, requirements for storage, manipulation, and transfer of digital images, has grown exponentially. These image files can be very large and can occupy a lot of memory.

It provides a potential cost savings associated with sending less data over switched telephone network where cost of call is really usually based upon its duration.

- It not only reduces storage requirements but also overall execution time.

- It also reduces the probability of transmission errors since fewer bits are transferred.
- It also provides a level of security against illicit monitoring.

2 Related Literature

A DCT-based method is specified for “lossy” compression, and a predictive method for “lossless” compression. JPEG features a simple lossy technique known as the Baseline method, a subset of the other DCT-based modes of operation. The Baseline method has been by far the most widely implemented JPEG method to date, and is sufficient in its own right for a large number of applications. [1]. In [2] for analyzing image compression methods that are based on wavelet decompositions. The theory relates the rate of decay in error between original image and the compressed images that is measured using a family of L^p norms, as the size of compressed image increases, to the smoothness of the image (in certain smoothness classes called Besov’s spaces). Within this theory, error incurred by the quantization wavelet transform coefficients is explained. An adaptive lossy LZW algorithm is proposed for palettised image compression in [3]. The algorithm employs an adaptive thresholding mechanism with human visual characteristics to lessen distortion. A probability model for natural images based on empirical observation of their statistics in wavelet domain has been developed and presented in [4]. Similarly some algorithms have been proposed by researchers for lossy and nearly lossy compression techniques [5-6]. An efficient VLSI design of one-dimensional direct discrete wavelet transform processor is presented in [7]. A real-time wavelet image compression algorithm using vector quantization and its VLSI architecture are proposed in [8]. The proposed zero-tree wavelet vector quantization (WVQ) algorithm focuses on the problem of how to reduce the encode wavelet images with high coding efficiency. Lossless compression is usually required in the medical image field is presented in [9]. The word length required for lossless compression makes too expensive the area cost of the architectures that appear in the literature. SPIHT algorithm is used to develop the codec, where authors claimed for their better efficiency [10]. The embedded block coding algorithm at the heart of the JPEG 2000 image compression standard is described in [11]. A lossless compression of images using coding schemes and patterns that include minterm, cube and coordinate data coding, Walsh, triangular and Reed–Muller weights based patterns, Reed–Muller spectra and reference row technique is proposed in [12]. A lossless wavelet-based image compression method with adaptive prediction is proposed in [13]. Also, wavelet transform is used with neural network by authors in past [14]. Signal processing methods have been used in FPGA implementation using VHDL by many authors [15-17]. But a considerable work on image processing methods has been implemented in FPGA [18-19]. An FPGA-based image and data processing core for future generation wireless capsule endoscopy (WCE) is presented in [20]. A resource efficient and high-performance architecture for a two-dimensional multi-level discrete wavelet transform processor is presented in [21]. Therefore authors were motivated to work in this field and hence taken the choice with wavelet transform.

The most commonly used metrics still remain simple, mathematically defined measures such as peak signal to noise ratio (PSNR) or mean squared error (MSE). When the quantization is varied on a single image in a straightforward manner, such as by varying the scale factor in JPEG compression, these metrics do correlate with image quality.

3 Method for Compression

The wavelet transform has been widely used in image and video compression since it allows localization in both the space and frequency domains [22]. The intensive computation of DWT due to its inherent multilevel data decomposition and reconstruction operations brings a bottleneck that drastically reduces its performance and implementations for real-time applications when facing large size digital images and/or high-definition videos. A hardware implementation of 1-D NHWT was presented in [23]. The architecture described in [20] involves only additions and subtractions along with normalization but only along one dimension.

A hardware implementation of 2-D Haar wavelet transform has been attempted in this work. The computing strategy finally chosen for coding the 2-D Haar wavelet transform is based on first evaluating the one-dimensional Haar wavelet transform to each row of pixels (a matrix of intermediate results is thus obtained), followed by the evaluation of the one-dimensional transform of the columns of the resulting intermediate matrix.

Two main aims have been pursued in the conception of this implementation: first, to exploit as much as possible the local properties of the Haar wavelet in order to eliminate possible redundant information during the coding process; and, second, to make use of the inherent arithmetical simplicity of the Haar wavelet transform in order to attain a low circuit complexity hardware implementation. The main aim is to implement a complete processor using one configurable general-purpose chip, yielding the lowest possible loss of quality. The simulation results have been presented thereafter. The VHDL implementation was done using Xilinx ISE 14.1.

3.1 Implementation of JPEG algorithm

The JPEG compression and decompression algorithm was coded without using any MATLAB library functions so that a hardware implementation using VHDL could be developed that conforms to the algorithm. The algorithm was implemented in following phases:

- 1) The uncompressed source data is separated into 8×8 blocks of pixels. 128 is subtracted from the value of each pixel so that the new effective range is from -128 to 127.
- 2) Each block is transformed into an 8×8 block of frequency coefficients using DCT.
- 3) These coefficients are quantized, by dividing the frequency coefficient matrix by the quantization matrix on an element-by-element basis.
- 4) An entropy encoder is applied to the quantized coefficients. The algorithm uses a zigzag ordered encoding, which collects the high frequency quantized values into long strings of zeros. DC values use delta encoding, which means that each DC value is compared to the previous value, in zigzag order. For all other values, i.e. high frequency values, Huffman coding (or any run-length coding is used).

The comparison between original and decompressed image has been based on parameters like mean square error (MSE), root mean square error (RMSE) and peak signal to noise ratio (PSNR) values.

3.2 Implementation of DWT (Haar Transform)

In this work, Haar wavelets compression is used. It provides for an efficient way to perform both lossless and lossy image compression. It relies on averaging and differencing values in an image matrix which may be sparse or nearly sparse. A sparse matrix has large number of entries as zeros,

which can be stored in an efficient manner, leading to smaller file sizes. The algorithm to compress the image is as follows:

1. An image is sub-divided into blocks of 8x8 pixels (with padding if necessary).
2. In each of the blocks, the pixels are grouped in twos. Then first 4 columns are replaced by their averages and last four by the half of their differences. The first four coefficients are known as approximation coefficients and last four are called detail coefficients.
3. In the next step, the first entries are again grouped in twos and first two are replaced by the averages and next two by the half of the differences, leaving last 4 entries unchanged.
4. In this step, the first 2 entries are grouped and replaced by their average and half of their difference.
5. The steps are repeated for the remaining rows of the block. After this, the process is repeated for the columns, grouping rows in the same manner as columns.

An easier implementation of 2D Haar Transform involves two 8x8 matrix multiplications, i.e., $H^T A H$, per 8x8 pixel block of the image. But from a hardware point of view, this procedure requires a total of 1024 multiplications and 896 additions and a subsequent division by 8, 4, and 2 operations for generating a resulting 8x8 matrix. For a lossless implementation, this also involves floating point implementation of this transform.

The implementation presented in this work, is a multiplier-less implementation of a lossy 2D Haar transform and involves only 112 additions, 112 subtractions and 224 shift-right-by-1-bit operations on 8-bit signed numbers. The addition and subtraction of the numbers have been implemented in the same module, thereby reducing the number of components having different design. The subtraction of two numbers and the addition of the negative numbers have been done using 2's complement of the binary numbers using 8-bit binary adder/ subtractor. The operations involving divide-by-2 has been achieved using shift-right-by-1-bit operation after each addition or subtraction. This solves two major purposes: first, this eliminates the need for floating point implementation; and, second, it drastically reduces computation load on the module. The data considered here, is a block of 8x8 pixels, which are of 8 bits each, corresponding to gray-scale intensity levels of 0-255 in case of unsigned numbers, or -128 – 127 in case of signed numbers.

The compression and decompression of the images using DWT was coded without using any MATLAB library functions so as to aid coding for hardware implementation using VHDL.

4 Simulation Results

From Table 1, it can be seen that JPEG algorithm takes considerable time for execution. This can be attributed to the fact that in phase 2, DCT is applied to the pixel blocks which is one of the most computation intensive phases; and in phase 4, wherein Huffman coding is applied which again is a computation intensive step. Same can be said for decompression times. Images with higher redundancy are compressed to a greater extent as compared to images with lower redundancy. Thus the compression ratio varies for different images. It can be seen that lossy JPEG compression produces error, but the error occurs mostly in high frequency elements.

As can be seen from Table 1, run times for compression and decompression using LZW algorithm are somewhat lesser than that of JPEG algorithm. This is because of the fact that, in LZW algorithm only

encoding or decoding of the sequence compresses or decompresses the sequences without applying any transforms, which saves some time.

Table 1: Observed Average Run-Time for JPEG Baseline, LZW and DWT (Haar) algorithms

Image	Average Run Time for Compression (in seconds)			Average Run Time for Decompression (in seconds)		
	JPEG Baseline Algorithm	LZW Baseline Algorithm	DWT Algorithm	JPEG Baseline Algorithm	LZW Baseline Algorithm	DWT Algorithm
Cameraman.tif	39.84	23.16	0.094	70.56	13.20	0.234
Eye_gray.bmp	41.96	15.03	0.091	42.83	10.76	0.228
Woman_blonde.tif	52.84	15.14	0.093	74.62	14.04	0.23
Woman_darkhair.tif	35.53	15.18	0.090	61.41	9.88	0.229
Pirate.tif	38.20	16.10	0.098	67.59	11.77	0.228
Mandrill_gray.tif	40.00	21.54	0.095	68.95	21.56	0.227

The images that have been used are 512x512 pixel size gray scale, standard test images. In Figure 1 and Figure 2, the left picture is original picture and the right one is the reconstructed picture after compression and decompression.

1. The Cameraman picture



Figure 1: Grayscale 512x512 Cameraman image [24] (a) Original image; Reconstructed image after lossy compression and decompression using DWT (b) with $\epsilon = 2$, (c) with $\epsilon = 5$, (d) with $\epsilon = 10$.

2. Other test pictures



Figure 2: (1) Grayscale 512x512 Eye gray image, (2) Grayscale 512x512 Woman blonde image, (3) Grayscale 512x512 Woman_darkhair image, (4) Grayscale 512x512 Pirate image, (5) Grayscale 512x512 Mandril_gray image [24]

It can be observed that the run time for compression and decompression for DWT algorithm are less than even a second. The observed run-time was equal both for lossless and lossy cases. This can be attributed to the fact that DWT using Haar transforms do not use any transforms like DCT or any encoding technique. This makes it a choice for hardware implementation in this work.

Table 2: Observed compression ratios and error in case of JPEG

Image	Compression Ratio	MSE	RMSE	PSNR (in dB)
Cameraman.tif	1.84	34.23	5.85	75.57
Eye_gray.bmp	2.18	27.50	5.24	77.76
Woman_blonde.tif	1.55	162.94	12.76	59.97
Woman_darkhair.tif	2.60	31.08	5.57	76.53
Pirate.tif	1.58	44.04	6.63	73.05
Mandrill_gray.tif	1.32	118.34	10.88	63.16

Table 3: Compression Ratio and Errors observed in case of DWT (lossy case, $\epsilon = 2$)

Image	Compression Ratio	MSE	RMSE	PSNR (in dB)
Cameraman.tif	6.20	4.99	2.23	94.82
Eye_gray.bmp	6.78	8.30	2.88	89.73
Woman_blonde.tif	3.22	8.79	2.96	89.17
Woman_darkhair.tif	8.91	6.84	2.61	91.66
Pirate.tif	3.16	8.05	2.83	90.04
Mandrill_gray.tif	2.09	7.37	2.71	90.93

Table 4: Compression Ratio and Errors observed in case of DWT (lossy case, $\epsilon = 5$)

Image	Compression Ratio	MSE	RMSE	PSNR (in dB)
Cameraman.tif	14.45	22.67	4.76	79.69
Eye_gray.bmp	22.12	28.83	5.37	77.28
Woman_blonde.tif	9.19	39.28	6.27	74.19
Woman_darkhair.tif	29.29	22.21	4.71	79.89
Pirate.tif	8.69	43.49	6.59	73.17
Mandrill_gray.tif	4.88	56.07	7.49	70.64

Table 5: Compression Ratio and Errors observed in case of DWT (lossy case, $\epsilon = 10$)

Image	Compression Ratio	MSE	RMSE	PSNR (in dB)
Cameraman.tif	29.35	64.43	8.02	69.25
Eye_gray.bmp	46.81	54.27	7.37	70.96
Woman_blonde.tif	23.06	97.80	9.88	65.07
Woman_darkhair.tif	49.56	45.28	6.73	72.77
Pirate.tif	23.66	113.96	10.67	63.54
Mandrill_gray.tif	15.29	190.65	13.80	58.39

In Table 2, it can be observed that the run time for compression and decompression are the less than even a second. The observed run-time was equal both for lossless and lossy cases. This can be attributed to the fact that DWT using Haar transforms do not use any transforms like DCT or any encoding technique. This makes it a choice for hardware implementation in this work.

In Tables 3 through 5, it can be clearly seen that compression ratio achieved using lossy compression techniques is greater than that achieved using lossless technique. The compression ratio can further be increased by making more number of entries to be zero. A non-negative threshold, ϵ can be fixed and any entry in the compressed matrix which is less than ϵ can be made zero. Hence resulting in a more sparse matrix and hence more saving in space. Compression ratios of about 20:1 are easily achievable. Higher compression ratios come at the cost of picture quality. It can be seen that the

mean square error (MSE) is highest when $\epsilon = 15$, thus picture quality degrades. The lossless case, by nature, will yield zero MSE.

4.1 VHDL Implementation

Figure 3 shows the schematic of how Haar transform has been applied in order to achieve the objective.

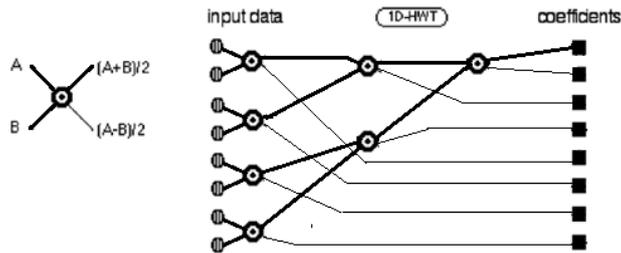


Figure 3: Computation scheme for discrete 1-D Haar wavelet transform

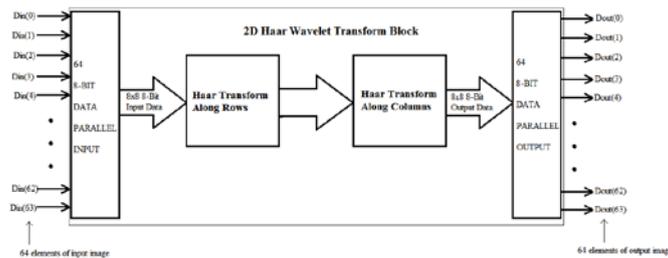


Figure 4: Schematic of the hardware implementation of 2D Haar Transform.

The data is brought into the module by 64 8-bit lines with a view to reduce the delay involved in fetching the data elements into the module.

Simulation Results

An 8 x 8 image has been used for the purpose of simulation.

The advanced HDL synthesis report as generated by Xilinx shows that 224 adder/subtractor blocks were totally needed.

```

=====
*                               Advanced HDL Synthesis                               *
=====

Advanced HDL Synthesis Report

Macro Statistics
# Adders/Subtractors           : 224
  8-bit addsub                 : 224
# Registers                    : 512
  Flip-Flops                   : 512
=====
    
```

Figure 5: Advanced HDL Synthesis report generated by Xilinx ISE

A 20 ns simulation shows that the results have been somewhat satisfactory. This simulation shows that there are many non-zero coefficients occurring in the matrix. This condition stems from the fact that no threshold co-efficient has been applied. No attempt has been made to make the detail coefficients to be zero.

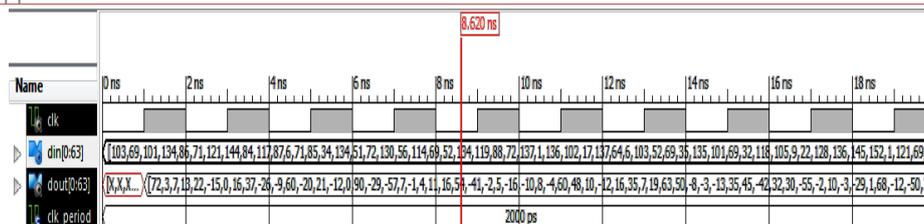


Figure 6: Test bench Simulation results

5 Conclusion

The DWT technique of compression has also been applied to the images. By the choice of Haar wavelet, it was found that there was minimal amount of error involved in case of lossy compression, which is responsible for higher PSNR values. A multiplier-less 2D Haar wavelet transform implementation was presented, which turns out to be very promising in reducing system complexity in terms of time and area, given the fact that DWT is computationally very costly. This can be taken care in future work by developing even simple algorithm.

REFERENCES

- [1] G. K. Wallace, "The JPEG Still Picture Compression Standard", Vol. 38, no. 1, IEEE Transactions on Consumer Electronics, pp. xviii – xxxiv, Feb. 1992.
- [2] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445-3462, December 1993.
- [3] S.W. Chiang, and L.M. Po, "Adaptive Lossy LZW Algorithm for Palettised Image Compression", *Electronics Letters*, IEE, pp. 852-854, March 1997.
- [4] R. W. Buccigrossi, E.P. Simoncelli, "Image Compression via Joint Statistical Characterization in Wavelet Domain", Vol. 8, No. 12, pp. 1688-1701, December 1999.
- [5] M. Boliek, M. J. Gormish, E. L. Schwartz and A. Keith, "Next generation image compression and manipulation using CREW," *Proceedings of the IEEE International Conference on Image Processing*, October 26-29, Santa Barbara, CA, pp. III-567-III-357, 1997.
- [6] M. J. Gormish, E. L. Schwartz, A. Keith, M. Boliek and A. Zandi, "Lossless and nearly lossless compression for high quality images," *Proceedings of the SPIE/IS&T Conference on Very High Resolution and Imaging II*, vol. 3025, San Jose, CA, pp. 62-70, February 1997.
- [7] C. Souani, M. Abid, K. Torki, R. Tourki, "VLSI design of 1-D DWT architecture with parallel filters", *INTEGRATION, the VLSI Journal* 29, pp. 181-207, 2000.
- [8] S-K Paek, L-S Kim, "A Real-Time Wavelet Vector Quantization Algorithm and Its VLSI Architecture", pp. 475-489, Vol. 10, No. 3, *IEEE Transactions on Circuits and Systems for Video Technology*, April 2000.
- [9] Urriza, et al., "VLSI Implementation of Discrete Wavelet Transform for Lossless Compression of Medical Images", *Real-Time Imaging* 7, pp. 203-217, 2001.

- [10] A. Said and W. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243-50, June 1996.
- [11] D. Taubmann, E. Ordentlich, M. Weinberger, G. Seroussi, "Embedded Block Coding In JPEG2000", *Signal Processing: Image Communication* 17, pp. 49-72, 2002.
- [12] B. J. Falkowski, "Lossless Binary Image Compression Using Logic Functions and Spectra", *Computers and Electrical Engineering* 30, pp. 17-43, 2004.
- [13] Y-T Chen, D-C Tseng, "Wavelet Based Medical Image Compression with Adaptive Prediction", *Computerized Medical Imaging and Graphics* 31, pp. 1-8, 2007.
- [14] Rasmita lenka, Swagatika Padhi, Minakshee Behera, Naresh Patnaik, Mihir N. Mohanty, "Design of Neuro-wavelet based vector quantizer for image compression", *Special Issue of International Journal of Computer and Communication Technology*, Volume 1 Issue 2, 3, 4; pp- 214-221, August-2010.
- [15] Asit Kumar Subudhi, Biswajit Mishra, Mihir N. Mohanty, "VLSI Design and Implementation for Adaptive Filter using LMS Algorithm", *IJCTT*, Vol-2, Issue-6, 23rd -24th Feb, 2011.
- [16] Mihir Narayan Mohanty, Biswajit Mishra, Aurobinda Routray, "FPGA implementation of CLMS Algorithm", *ICEAS, IEEE Conference, Bhubaneswar*, 28th -30th December, 2011.
- [17] Mihir N. Mohanty, Hemanta Kumar Sahu, "FPGA Implementation of Variable Step-size LMS Algorithm", *IEEE Conf.- ICCSP' 13, Melmaruvathur, TN*, 03- 05 April 2013.
- [18] Panchami Padmasana, Mihir Narayan Mohanty, Hemanta Kumar Sahu, "VHDL Implementation of Spatial Filter for Image Enhancement", *IEEE Conf., Chennai, ICCSP – 3-5 April 2014*.
- [19] Panchami Padmasana, Mihir Narayan Mohanty, P. Kabisatpathy, "FPGA Implementation of Modified Median Filter for Impulse Noise Removal from Image", *IJECT Vol. 5.3 - Spl 1, July - September, 2014*.
- [20] P. Turzca, M. Duplaga, "Low Power FPGA-Based Image Processing Core for Wireless Capsule Endoscopy", *Sensors and Actuators A: Physical* 172, pp. 552-560, 2011.
- [21] S.M. Aziz, D.M. Pham, "Efficient Parallel Architecture for Multi-Level Forward Discrete Wavelet Transform Processors", *Computers and Electrical Engineering* 38, pp. 1325-1335, 2012.
- [22] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, July 1989.
- [23] F. Javier Diaz, A. M. Buron, J. M. Solana, "Haar wavelet based processor scheme for image coding with low circuit complexity", *Computers and Electrical Engineering* 33, pp. 109-126, 2007.
- [24] http://www.imageprocessingplace.com/root_files_V3/image_databases.htm

A 100 MHz 6th Order Continuous Time Band-Pass Sigma Delta Modulator with Active Inductor based Resonators

¹Kevin Dobson, ²Shahrokh Ahmadi and ³Mona Zaghoul

Department of Electrical and Computer Engineering, George Washington University,
Washington DC, USA;

¹kdobson@gwu.edu, ²ahmadi@gwu.edu, ³zaghloul@gwu.edu

ABSTRACT

This paper presents a 6th order, continuous time band-pass Sigma Delta modulator in IBM 0.18 um CMOS technology. We replace traditional RLC circuits, containing low quality factor spiral inductors with high quality factor, active inductor based resonators utilizing negative impedance circuits. Pad to pad simulation of the extracted layout in Cadence yields an enhanced SNDR of 78 dB with a noise bandwidth of 18 kHz and a power consumption of 12 mW. Our modulator occupies 7.5 mm² of chip area with pads.

Keywords: Sigma Delta; Active Inductor; Negative Impedance circuit.

1 Continuous Time Sigma Delta Modulator Loop Filter

In Figure-1 below we see the block diagram of a Continuous Time Sigma Delta (CT $\Sigma\Delta$) modulator. In order to design the loop filter $G(s)$ we start with a discrete time modulator transfer function $F(z)$ which provides the required noise shaping.

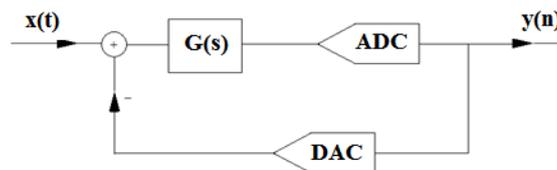


Figure-1: Continuous Time Sigma Delta Modulator

As outlined in [1], the impulse-invariant method is used to generate the equivalent CT loop filter $G(s)$ from $F(z)$.

$$F(z) = (1 - z^{-1})Z_T \left\{ L^{-1} \left[\frac{G(s)e^{-ds}}{s} \right] \right\} \quad (1)$$

Here d represents the delay introduced by the Analog to Digital Converter (ADC) and Digital to Analog Converter (DAC). For a sixth order band-pass Sigma Delta modulator, the equivalent continuous time loop filter transfer function is given by:

$$G(s) = \frac{(s - p)(s^2 + \frac{\omega_a}{Qa}s + \omega_a^2)(s^2 + \frac{\omega_b}{Qb}s + \omega_b^2)}{(s^2 + \frac{\omega_0}{Q_0}s + \omega_0^2)(s^2 + \frac{\omega_1}{Q_1}s + \omega_1^2)(s^2 + \frac{\omega_2}{Q_2}s + \omega_2^2)} \quad (2)$$

ω represents the normalized resonator frequencies with respect to the sampling frequency in radians per second and Q is the quality factor of the resonators. If we denote T as the sampling frequency, when d is equal to $1.4T$ and the sample rate is 4 times the frequency of the input signal, then the term p in the numerator is approximately equal to zero [2].

2 Active Inductor Resonator structure

The transfer-function of a parallel RLC resonator circuit $H(s)$ is given by:

$$H(s) = \frac{As}{s^2 + \frac{\omega_0}{Q} + \omega_0^2} \tag{3}$$

$G(s)$ cannot be realized as a cascade of resonators but can be realized by the structure shown in Figure 2 which appropriately places the poles and zeros. Here g , A_H and A_L represent amplifier gains, and H the resonators. The Σ block is an analog adder. H_0 , H_1 and H_2 are high Q resonators and provide noise shaping. H_3 is a low Q resonator which provides modulator stability.

A review of the available literature will show that most band-pass CT $\Sigma\Delta$ modulators have a loop filter containing parallel RLC circuits with spiral inductors. Such circuits occupy a large silicon area and the spiral inductors have a low Q . Active inductor based resonator circuits occupy a much smaller area leading to more economical designs. When Q enhancement techniques are used, high Q can be achieved. Active inductor based resonators are explained by the gyrator C theorem as outlined in Figure iii below.

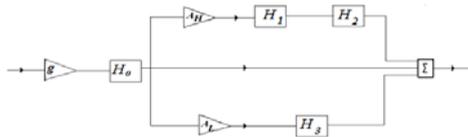


Figure 2: Sixth Order Loop Filter

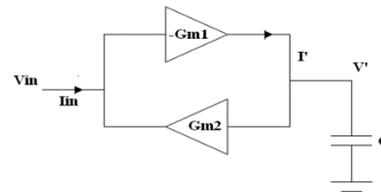


Figure 3: Gyrator Topology

It is easily shown that:

$$\frac{V_{in}}{I_{in}} = \frac{sC}{G_{m1}G_{m2}} \tag{4}$$

The s in the numerator of Eq. (4) indicates the inductive nature of the circuit. When MOSFETs are used to realize G_{m1} and G_{m2} additional parasitics are introduced forming a lossy parallel RLC resonator circuit as shown in Figure 4 below.

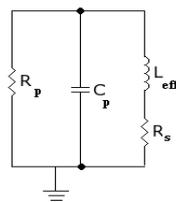


Figure 4: Lossy Parallel RLC Resonator

When two such gyrators are coupled with a Negative Impedance Circuit (NIC), as shown in Figure v, a high Q fully differential resonator, for use in a band-pass $\Sigma\Delta$ modulator, can be designed.

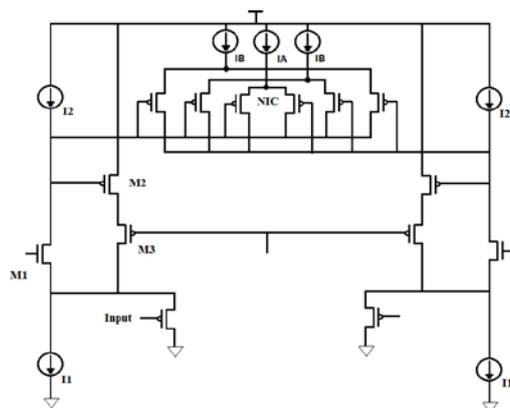


Figure 5: CMOS Active Inductor based Resonator with NIC

Upon performing a small signal analysis of the circuit in Figure v and separating the resistive, capacitive and inductive parts we get the following equations [3]. g_o is the drain source conductance and g_m the transistor transconductance. C_{nic} represents the additional capacitance introduced by the NIC and g_{nic} its transconductance.

$$R_p = \frac{1}{g_{o2}}; C_p = C_{gs1}; L_{eff} = \frac{C_{gs2}}{g_{m1}g_{m2}}; \text{ and } R_s = \frac{g_{o1}g_{o3}}{g_{m1}g_{m2}g_{m3}} \quad (5)$$

The enhanced self-resonant frequency and quality factor of the circuit is given respectively by:

$$\omega_{en} = \sqrt{\frac{C_p}{C_p + C_{nic}}} \sqrt{\frac{g_{m1}g_{m2}}{C_{gs1}C_{gs2}}} \quad (6)$$

$$Q_{en} = \frac{1}{g_{o2} - g_{nic}} \sqrt{\frac{C_p + C_{nic}}{L_{eff}}} \quad (7)$$

PMOS devices are used to couple the differential input to the circuit. They are biased to draw a small amount of current and do not disturb the gyrator function. Output gain is controlled by varying the size of these PMOS devices. Cascoding M3 and M2 reduces the output conductance which reduces R_s and increases the Q .

The NIC is comprised of 3 cross-coupled differential pairs of PMOS devices with drains tied to the opposing gates. It provides a negative resistance that attempts to cancel the parallel resistance R_p , which further increases the Q . The use of PMOS devices allows for current reuse and results in less power consumption. Single pair cross-coupled NICs exhibit low linearity [4]. In order to obtain greater linearity a multi-tanh version of the NIC circuit is used. This requires the addition of two extra cross-coupled pairs of MOSFETs with a 4:1 size ratio [5][7]. When the signal is large and the symmetrical differential pair has saturated, the unbalanced differential pairs can still provide a differential current proportional to the input voltage. This scheme works effectively at high frequencies.

While there are no limits to the voltage that can be applied to spiral inductors, the maximum input voltage to active inductor based circuits cannot cause MOSFETs to cease operating in saturation mode. Active inductor based circuits are also noisier than circuits with real inductors by a factor of $2Q_0$, the intrinsic quality factor [5]. All current sources used are cascode in order to provide greater linearity.

3 Active Inductor Simulation

We were able to design and simulate active inductor based resonators in Cadence with resonant frequencies between 15 MHz and 1.5 GHz, and were consistently able to achieve a Q of 50 or greater. Simulation results for one active inductor designed are shown in Figure 6 and 7.

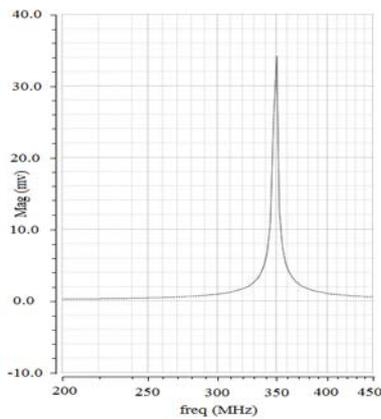


Figure 6: Active Inductor simulated AC response

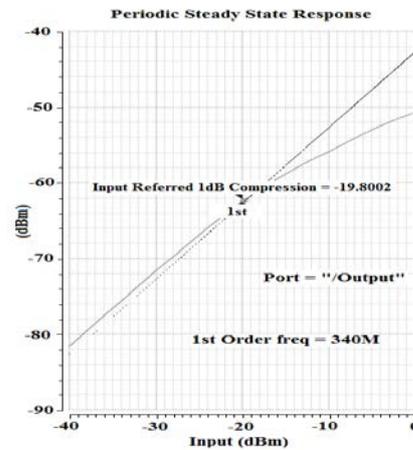


Figure 7: Active Inductor Linearity

4 Matlab Modulator Simulation

Initial values of the loop filter multiplying coefficients g , A_H and A_L of $G(s)$ were used to generate Pole-Zero plots and confirm modulator stability. The appropriate noise shaping characteristics were verified by the use of the Delta-Sigma toolbox [6]. Simulink simulations were then used to further refine the modulator design.

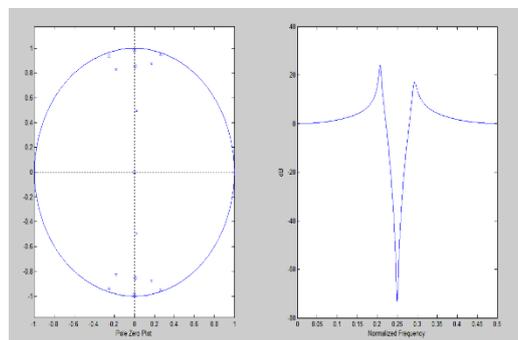


Figure 8: Modulator Pole Zero plot

The Simulink model was easily modified to reflect the non-idealities of an actual circuit such as limited gain due to nonlinearity and small delays introduced by each circuit component. A theoretical Signal to Noise-plus-Distortion Ratio (SNDR) of 95dB was obtained for the ideal case when there is a high gain in the path containing the most resonators. When the limitations of gain and circuit delay were taken into account a goal of 80 dB was deemed feasible.

5 Cadence Simulation

For the Cadence circuit design, resonators were designed to resonate at 25 MHz with a Q of 30. Linear operation was observed when the input was limited to less than 10mv p-p. A block diagram of the complete circuit simulated in Cadence is shown in Figure ix.

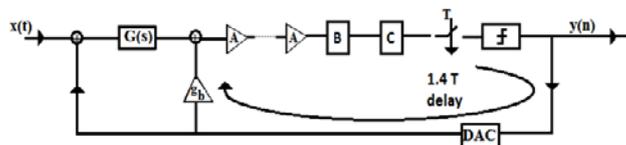


Figure 9: Active Inductor based Sixth Order Continuous Time Modulator

A series of differential amplifiers [8] A, provide effective amplification of the signal prior to quantization preventing clock feed-through [8] and contributing to the required delay of $1.4T$. A coupling capacitor with buffer for biasing the next stage of the circuit, B, eliminates dc offsets introduced after layout. This capacitor and input resistance of the buffer must be large enough to form an effective high pass filter thereby preventing attenuation. Pre-comparators C provide further amplification before quantization by a clocked comparator. Since there is a non-zero delay it is necessary to add a direct loop between the DAC and the ADC input [9]. The Adder used is described in [2]. A 2 input version of the adder is used to subtract the DAC output from the input signal. Conceptually the subtraction is the addition of the input and the negative DAC output. A clocked comparator [8] coupled with a SR flip flop generates the (NRZ) modulator output. The DAC is described in [10].

We use separate VDD and GND for these sub-circuits. We also surround our analog sub-circuits with two guard rings separated by BFMOAT in order to protect them from digital noise. [11]

Extraction of the layout yields extra parasitics which result in larger propagation delays and lower output signals, resonant frequency and Q . Several circuit modifications were made in order to negate the effects of the added parasitics. The first was to ensure a symmetrical layout so that the parasitics added on either side of the differential circuit were equivalent. The next was to increase the size of the input PMOS devices leading to better coupling of input signals. The third was to increase the current in the current mirror I2 which is necessary to overcome the added impedances. The fourth was to increase the size of M3. This decreases R_s which increases our deteriorated Q . The fifth was to lower our clock frequency to account for the increased propagation delays. When these changes were made to the layout and Input/ Output pads added, the layout was re-extracted and a pad to pad simulation was conducted.

This yielded a SNDR of 78dB with a noise bandwidth of 12 kHz. This was quite close to the 80dB observed in the Simulink simulations. The difference is believed to be due to circuit non-idealities, such as settling time of the adder output, and offset errors. This result is shown in Figure 10. It compares favorably with other non-active inductor based sixth order band-pass $\Sigma\Delta$ modulators such as [12], which yields a SNDR of 68dB. Our circuit consumes 12 mW and occupies 7.5 mm^2 , including pads.

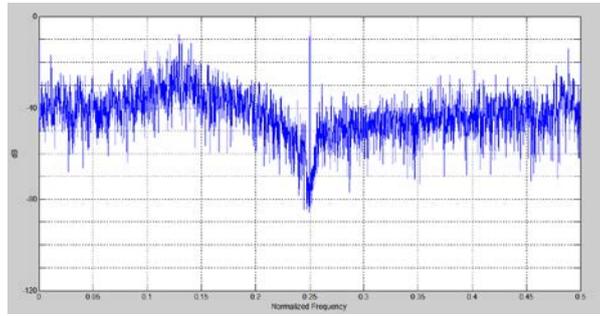


Figure 10: Extracted Layout Power Spectrum Density

6 Conclusion

The greatest design limitation encountered when designing CT $\Sigma\Delta$ modulators with active inductor based resonators is the limited linearity. This results in a design that is highly susceptible to offsets. Also many extra amplification stages are necessary to raise voltages to acceptable levels prior to quantization. The extra circuitry required to compensate for offsets and amplification negates the area gains made by the avoidance of spiral inductors. More research is needed to determine how to extend the linearity of active inductor based resonators. This would result in designs that are capable of operating at higher frequencies and occupying smaller areas.

A review of the available literature shows only one other active inductor based band-pass $\Sigma\Delta$ modulator [13]. The authors included schematic simulations but not simulation results for an extracted layout. We have succeeded in designing and simulating the first Sixth Order CT $\Sigma\Delta$ modulator, using active inductor based resonators in the loop filter and provided pad to pad simulations of the extracted layout. Our design has been submitted for fabrication with MOSIS.

REFERENCES

- [1] C. Lelandais-Perrault, P. Benabes, J. De Gouy, R. Kielbasa, *A parallel structure of a continuous-time filter for bandpass sigma-delta A/D Converters*, 10th IEEE International Conference on Electronics, Sharjah (Emirates Arabes Unis), DEC. 2003.
- [2] S. Benabid, P. Benabes, *High linear integrated LC filter for a continuous-time bandpass sigma-delta ADC* Circuits and Systems, 2003 IEEE 46th Midwest Symposium on, pp. 291 - 294 Vol. 1, 30 Dec. 2003
- [3] Fei Yuan, *CMOS Active Inductors and Transformers, Principle, Implementation, and Applications*, Springer 2008
- [4] Byunghoo Jung, Harjani, R., *A wide tuning range VCO using capacitive source degeneration*, Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on, Volume 4, 23-26 May 2004, Page(s): IV - 145-8 Vol.4.
- [5] Yue Wu, Xiaohui Ding, Mohammed Ismail, and Håkan Olsson, *RF bandpass filter design based on CMOS active inductors*, IEEE Transactions on circuits and systems—II: Analog and Digital Signal Processing, vol 50, no.12, Dec. 2003.
- [6] Richard Schreier and Gabor C. Temes, *Understanding Delta-Sigma Data Converters*, IEEE Press

- [7] A. P. Ryan, O. McCarthy, *A novel pole-zero compensation scheme using unbalanced differential pairs*, *IEEE transactions on circuits and systems—I: Regular papers*, vol. 51, no. 2, Feb. 2004.
- [8] R. Jacob Baker, *CMOS, Circuit Design, Layout, and Simulation*, IEEE Press Series on Microelectronic Systems.
- [9] P. Benabes, M. Keramat and R. Kielbasa, *Synthesis and Analysis of Sigma-Delta Modulators Employing Continuous-Time Filters*, *Analog Integrated Circuits and Signal Processing*, no. 23, pp. 141-152, Jul. 1998.
- [10] Kevin Dobson, Shahrokh Ahmadi and Mona Zaghoul, *A 1.2 GHz Band-Pass Sigma Delta Analog to Digital Modulator with Active Inductor based Resonators*, *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2012, WCECS 2012, 24-26 October, 2012, San Francisco, USA*, pp 875-879
- [11] Haitao Dai, *Differential Sensing of Substrate Noise in Mixed-Signal 0.18- um BiCMOS Technology*, *Electron Device Letters, IEEE*, Aug. 2008, Volume 29: Issue 8, pp 898-901
- [12] Cho-Ying Lu, Silva-Rivas, J.F.; Kode, P.; Silva-Martinez, J.; Hoyos, S., *A Sixth-Order 200 MHz IF Bandpass Sigma-Delta Modulator With Over 68 dB SNDR in 10 MHz Bandwidth*, *Solid-State Circuits, IEEE Journal of*, Volume: 45, Issue: 6, 2010 , Page(s): 1122 – 1136
- [13] Quining Chen, Kamal El-Sankary and Ezz El-Masry, *A UHF continuous-time current-mode band-pass delta sigma modulator based on active inductor*, *Circuits and Systems*, 2008. *MWSCAS 2008. 51st Midwest Symposium on*

Real Time Human Action Recognition using Kinematic State Model

¹Geetanjali V. Kale and ²Varsha H. Patil

Department of Computer engineering, MCERC, Nashik, Maharashtra, India

¹gilkale@gmail.com, ²varsha.patil@gmail.com

ABSTRACT

Human action recognition has tremendous applications in interdisciplinary domain and it's challenges kept researchers busy worldwide. Variety of applications gave rise to different representation and recognition methods. Posture can either be represented by shape features or skeleton features. We have represented action using sequence of postures and skeleton features are used for posture representation. Proposed work recognizes Yogasana from real time video. Yogasana is a type of exercise, in which specific sequence of the postures needs to be performed. Regular practice of it shows tremendous benefits in physiological and psychological disorders. System uses skeleton data of twenty human joints for representation of posture. Asanas are represented by kinematic state model using skeleton data provided by Kinect sensor. System is tested on 120 real time video sequences captured on four different subjects performing three asanas. Our system gave 96% recognition rate.

Keywords: human motion recognition, kinematic human model, Yogasana, Kinect.

1 Introduction

Human motion recognition from video or images is interdisciplinary domain having social, industrial and educational benefits. It is key requirement for many applications like smart surveillance, human computer interaction (HCI), behavioral biometric, motion analysis [1-4]. General steps in any human motion recognition applications are (a) track the human from given video, (b) represent human and their actions efficiently, (c) recognize the action or activity performed i.e. classify action in its correct class.

Tracking human from given video stream is low level vision problem. With the advances in domain video sensors like ASUS Xtion Pro Live and Microsoft Kinect tracks the human in video. Microsoft Kinect provides depth as well as human skeleton stream of video data along with RGB [5]. Kinect is originally designed for gaming, but, soon it became popular in Computer Vision community due to quality of its skeleton and depth data.

Interdisciplinary applications gave rise to different representation and recognition strategies. Broadly, human posture representation methods are classified as humanoid body model approaches and image based approaches [6]. Humanoid body model uses skeleton data and body kinematics. Image based approach uses body silhouettes or contours. Body silhouettes or contours can be extracted from either depth data or RGB stream. Image based approaches are very sensitive to change in human action, camera view and human size. Whereas, kinematic approach is camera view invariant. Proposed work uses kinematic data for posture representation.

Section 2 discuss about previous work done in domain, section 3 explains proposed system along with methodology used, section 4 gives experimental details along with results and last section concludes paper with future scope.

2 Previous Work

Human motion recognition has a long history. Aggarwal and Ryoo, T. B. Moeslund et al., Poppe, and Pavan Turaga et al. have discussed insights of domain in detail [1]-[4]. This section discuss about work carried out on depth and skeleton data provided by Microsoft Kinect sensor after its launch in 2010. J. Han et.al. gave insights for researchers to exploit and improve computer vision algorithm using Kinect [13]. Authors also discussed object tracking and recognition, human activity analysis, hand gesture analysis, indoor 3D mapping using Kinect.

S. Nowozin and J. Shotton proposed concept of action points to recognize action from online low latency data [7]. Authors applied firing HMM and random forest direct classification on action points for low latency data. Results are analyzed on Weizmann repetitive actions, Kinect actions Dataset. The random forest recognition system based on the skeletal joint data has been successfully integrated into a game title that is currently being sold in retail stores. But proposed approach is suitable only for momentary and discrete movements. Assumptions do not fit for continuous data like walking or well defined steps in dance. M. Yang et al. has represented posture using angles made by body parts and achieved almost 100% results except for throw action [8]. X. Yang and Y. Tian used 3-D positions of body joints and combined action information using static posture, motion and offset [9]. Feature channel $f_c = [f_{cc}, f_{cp}, f_{ci}]$ where f_{cc} is posture feature, f_{cp} is motion feature and f_{ci} is offset feature, results in high dimensional data. They further used PCA to reduce redundancy and noise and also obtained EigenJoint representation for each frame. J. Sung et al. used a hierarchical maximum entropy Markov model (MEMM) for detection and recognition of twelve different activities performed by four people in different environments, such as a kitchen, a living room, an office, etc. [10]. Using RGBD data provided by Kinect sensor and achieved an average performance of 84.3% when the person was seen before in the training set (and 64.2% when the person was not seen before).

Z. Gao et al. have taken advantage of multi-modal information provided by Kinect and proposed multi-modality information collaborative representation and recognition (MMCRR) [11].

Major contour or silhouette based approaches fails to give anthropometric and view point invariance for action recognition. As persons shape changes with change in view point and size of person changes with change in anthropometry. 3-D kinematic approach can be used to overcome problem of view point invariance for recognition of human action. Proposed system uses kinematic approach and skeleton data obtained from the inexpensive Microsoft Kinect sensor. Our system recognizes performed *Yogasana* from given video. *Yogasana* is a type of exercise composed of sequence of postures. Regular practice of it shows tremendous benefits on health.

3 Proposed System

The present age of speed and competition has increased the stresses and strains resulting in an increasing prevalence of life style-related health problems such as hypertension, cardiac diseases, bronchial asthma, diabetes neurosis and depressive illness, low back pain [12][13]. With growing scientific evidence, *Yogasana* is emerging as an important health behavior-modifying practice to achieve good states of health. Regular practice demonstrated the beneficial effects on health behavior in many life style-related problems. Regular practice of *Yogasana* also reduces the stress,

which in result improves the performance of students in examination, Yoga also helps for improving memory, intelligence and creativity [14].

Yoga is the name of one of the six orthodox philosophical schools and the Yoga school was founded by Patanjali. The 29th Sutra of the 2nd book of Patanjali gave concept of "Ashtanga Yoga" ("Eight-Limbed Yoga") [15]. Asana is third limb of "Ashtanga Yoga" that concentrates on practice of specific breathing techniques and postures. Proposed work concentrates only on postures and does not consider breathing technique.

In proposed work we have considered standing postures without twist action. Three types of Asanas used are (i) Samasthiti Tadasana, (ii) Urdhva Hastasana Tadasana, (iii) Veerbhadrasana. Detailed posture sequence of Urdhva Hastasana Tadasana is shown in Figure 1. Figure 1 (a) shows skeleton representation of posture sequence to be followed in Urdhva Hastasana Tadasana and Fig 1 (b) shows state transition of one pose to another sequentially. Postures can be viewed as states and Asana is represented by sequence states. Posture sequence considered in Urdhva Hastasana Tadasana is Samasthiti, hands apart, hands up, whole body up (stand on toes), body down (stand on foot), Hands Apart (intermediate posture considered for hands down), Hands down. In Samasthiti Sama means upright, straight, unmoved and Sthiti means standing still or steadiness. Samasthiti Tadasana has three key postures Samasthiti, body up and body down. Veerbhadrasana has seven key postures i.e. Samasthiti, legs apart, hands apart, knee bend, knee straight, hands down, legs near. Yogasana has standing, seating and lying down postures. It is transition of posture sequence from one state to another.

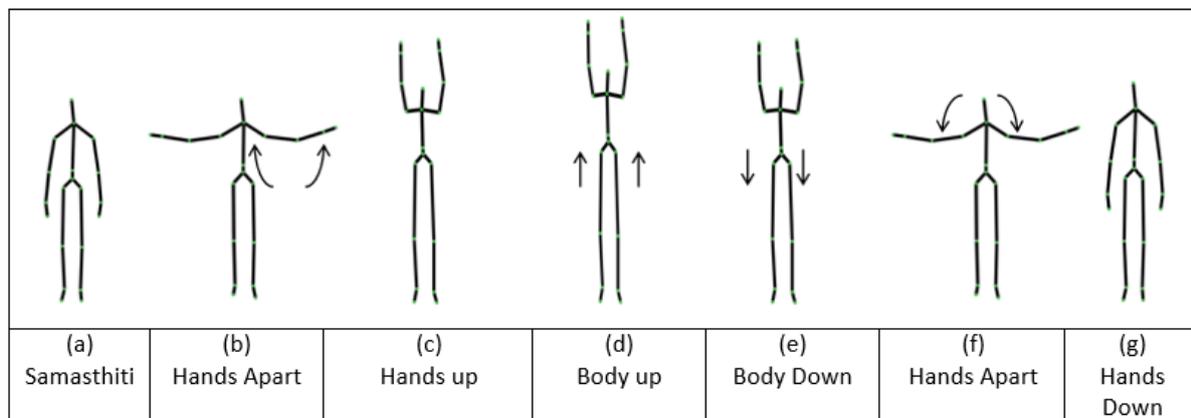


Fig 1(a): States using skeleton representation

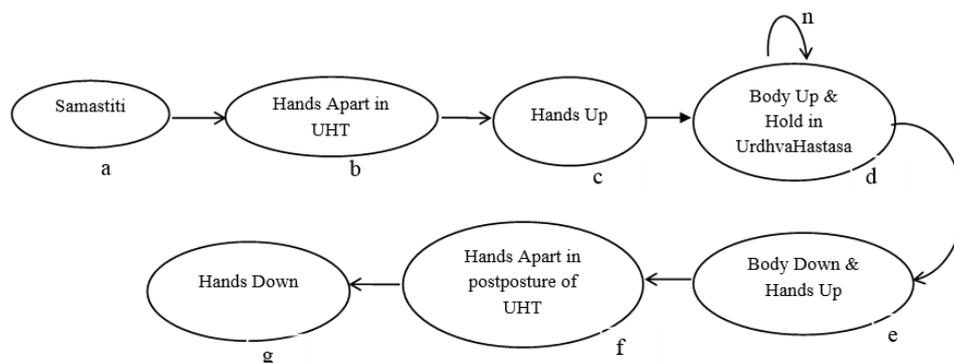


Figure 1(b): Sequence of State transition
Figure 1: Urdhva Hastasana Tadasana

$Yogasana = \{P_1, P_2, P_3, \dots, P_n\}$, where P_i is posture in given Asana $1 \leq i \leq n$, n is number of keyframes in particular *Asana*. Each posture is given by articulated human model represented using 20 joints, $P_i = \{J_{i,1}, J_{i,2}, J_{i,3}, \dots, J_{i,20}\}$. Each joint can be described in space using three parameter triplet $\langle X_{ij}, Y_{ij}, Z_{ij} \rangle$, where $1 \leq j \leq 20$ and i is posture $1 \leq i \leq n$. \therefore Posture $P_i = \{\{J_{i,x}^1, J_{i,y}^1, J_{i,z}^1\}, \dots, \{J_{i,x}^{20}, J_{i,y}^{20}, J_{i,z}^{20}\}\}$. Motion is represented by change in joint positions with respect to *Samasthiti*. *Samasthiti* is taken as reference posture for initialization $P_1 = \{\{J_{1,x}^1, J_{1,y}^1, J_{1,z}^1\}, \dots, \{J_{1,x}^{20}, J_{1,y}^{20}, J_{1,z}^{20}\}\}$. Difference in two postures can be given by, $diff(\partial P_i, \partial P_{i+1})$. This will result in vector $\{\partial J_1, \partial J_2, \partial J_3, \dots, \partial J_n\}$. Next posture / state in sequence is identified from difference of most informative joints of reference posture and recent posture.

$$\{\delta_{i,x}^k, \delta_{i,y}^k, \delta_{i,z}^k\} = diff(\{J_{1,x}^1, J_{1,y}^1, J_{1,z}^1\}, \{J_{i,x}^{20}, J_{i,y}^{20}, J_{i,z}^{20}\})$$

Change will depend on type of action performed and sequence of postures.

4 Experimentation and Results

System is tested on three types of Asana - *Samasthiti Tadasana*, *Urdhva Hastasana Tadasana*, *Veeerbhadrasana* (Warrior Pose - II). X and Y co-ordinates of twenty joint positions of human body obtained from Kinect are considered as feature for recognition. Joint positions considered are center, left and right point of hip and shoulder, left and right point of hand, wrist, elbow, knee, ankle, and foot, spine and head. All three asanas are tested on 4 female subjects and each asan is repeated 10 times i.e. $3 \times 4 \times 10 = 120$ video sequences. Table 1 shows results obtained for three asanas repeated 40 times.

Table 1: Recognition results for three Asana

	Samasthiti Tadasana	Urdhva Hastasana Tadasana	Veeerbhadrasana
Samasthiti Tadasana	39 (97.5%)	0	0
Urdhva Hastasana Tadasana	0	39 (97.5%)	0
Veeerbhadrasana	0	0	38 (95%)

Our system shows average 96% recognition rate. It is observed that failure of recognition is due to non-identification of some intermediate posture in *asana*.

Fundamental requirement of any action recognition system or algorithm is minimize computational complexity as well as feature vector dimensionality. Our algorithm satisfies both, as we consider only most informative joints for action recognition.

Computational Complexity of each Asana is $3 \times \sum_{i=1}^m N_i$ where, m - number of postures in *Yogasana* and N_i is number of joints changed in posture i (most informative joints)

5 Conclusion and Future Work

Many researchers are working worldwide on human motion recognition, but as of now, no off the shelf technology is available for it. This is due to different applications shows success for different representation and recognition methods. Proposed kinematic state model for identification of *Yogasana* from video sequence has shown good recognition rate. Rare failure is due to non-

recognition of intermediate states of posture. Model can be applied for different exercises like aerobics and dances like bharatnatyam, where action is sequence of specific postures.

In future, we will try to make recognition rate 100% and system can be applied for other applications.

ACKNOWLEDGEMENT

Authors are thankful to Dr. Ruta Kulkarni of 'Yogapoint' for her guidance and students of Pune Institute of Computer Technology for performing Yogasana to create dataset. We are also thankful to anonymous reviewers for their time for review and expert comments on work.

REFERENCES

- [1] J.K. Aggarwal and M.S. Ryoo, "Human activity analysis: a review", ACM Computing Surveys, vol. 43, No. 3, Article 16, 2011, pp. 16:1–16:43
- [2] T. B. Moeslund et al., "A survey of advances in vision-based human motion capture and analysis", Computer Vision and Image Understanding 104, 2006, pp. 90–126.
- [3] R. Poppe, "Vision-based human motion analysis: An overview", Computer Vision and Image Understanding 108, 2007, pp 4–18.
- [4] P. Turaga et.al, "Machine Recognition of Human Activities: A Survey", IEEE Trans. on Circuits and Systems for video technology, vol. 18, no. 11, Nov. 2008, pp. 1473-1488.
- [5] J. Han et al., "Enhanced Computer Vision with Microsoft Kinect Sensor: A Review " , IEEE trans. on Cybernetics, Vol. 43 , no. 5, 2013, pp 1318-1343.
- [6] G. V. Kale and V. H. Patil, "Vision Based Human Motion Representation and Recognition Methods a Survey", Proceedings of Third Post Graduate Conference on "Computer Engineering" cPGCON 2014, Elsevier, Vol X, pp XX - XX.
- [7] S. Nowozin and J. Shotton, "Action points: A representation for low-latency online human action recognition," Microsoft Research Cambridge, Tech. Rep. MSR-TR-2012-68, 2012.
- [8] M. Yang et al., "Human Action Recognition Based on Kinect", Journal of Computational Information Systems 10: 12, 2014, pp 5347–5354.
- [9] X Yang and Y Tian, "Effective 3D action recognition using EigenJoints", Journal of Visual Communication and Image Recognition, 25 2014, pp 2–11.
- [10] J. Sung et al., "Human Activity Detection from RGBD Images", AAAI Workshop on Pattern, Activity and Intent Recognition, July 2011.
- [11] Zan Gao et.al, "Human Action Recognition Via Multi-modality Information", J Electr Eng Technol Vol. 9, No. 2, 2013, pp 739-748.

- [12] Sudheer Deshpande, Nagendra H R, Raghuram Nagarathna, "A randomized control trial of the effect of yoga on Gunas (personality) and Health in normal healthy volunteers" International Journal of Yoga Vol. 1:1, Jan-Jun-2008, pages:2-10

- [13] K. Williamsa et.al., "Effect of Iyengar yoga therapy for chronic low back pain", International Association for the Study of Pain, 115 (2005) 107–117.

- [14] A. Kauts and N. Sharma, "Effect of yoga on academic performance in relation to stress," International Journal of Yoga Vol. 2:1 Jan-Jun-2009 Pages:39-43.

- [15] B. K. S. Iyengar, "The Illustrated Light on Yoga", HarperCollins Publisher India, Tenth impression, published in 2005.

Predicting the Identity of a Person using Aggregated Features of Handwriting

¹Revathi S V and ²Vijaya M S

PSGR Krishnammal College for women, Coimbatore
sv.revathi15@gmail.com; msvijaya@grgsact.com

ABSTRACT

The identification of an individual based on handwriting is a useful biometric modality. The biometric modalities are broadly classified into two types, namely psychological and behavioral characteristics. The physiological characteristics are fingerprint, face, iris, retina, hand geometry and the behavioral characteristics are voice, signature, gait, handwriting. Handwriting recognition plays vital role in forensic analysis, signature verification and network security. The automatic writer identification will be a valuable and relevant tool in forensic analysis and biometric authentication. Hence it is proposed to design and develop a model for automatic recognition of a person based on handwriting using pattern recognition technique.

Keywords: Classification, Feature Extraction, Prediction, Training, Writer Identification.

1 Introduction

The significance and scope of writer identification has become more prominent in these days. Writer identification is used in most of the areas like digital rights management, forensic expert decision making system and biometric authentication. By combining the writer identification with the authentication system it is used to access the confidential site or data where large amounts of documents, notes, forms and meeting minutes are constantly being processed and managed.

In forensic science writer identification is used to authenticate documents such as records, diaries, wills, signatures and also in criminal justice. The copyrights of electronic media are protected in the digital rights administration system. Two broad categories of biometric modalities are physiological biometrics and behavioral biometrics. The physiological biometrics performs person identification based on measuring a physical property of the human body. The behavioral biometrics use individual traits of a person's behavior for identification. Therefore writer identification is the category of behavioral biometrics.

Writer identification can be generally classified into two types as online and offline. In online, the writing behavior is directly captured from the writer but in offline the handwritten text is used for identification in the form of scanned images. Off-line writer identification is extensively considered as more challenging than on-line because it contains more information about the writing style of a person, such as pressure, speed, angle which is not available in the off-line mode.

Writer identification approaches can be categorized into two type's namely text-dependent and text-independent methods. Dependent on the text content, text-dependent methods only matches the same characters and requires the writer to write the same text consequently.

The text-independent methods are able to identify writers independent of the text content and it does not require comparison of same characters. If it requires the same writing content, then this method is not apt for many practical situations. Even though it got a wider applicability, text-independent methods do not obtain the same high accuracy as text-dependent methods do. The basic property of handwriting is that there exists writer invariant which makes writer identification possible. The writer's invariants reflecting the writing style or writing individuality of handwriting can be defined as the set of similar patterns.

2 Literature Survey

Based on the study of various literatures available on writer identification, a brief report is presented in this section about the research directions in writer identification.

In the research work [8], the authors developed the writer identification model using character level and word level. Scanned images of handwritten words were segmented into words and it is further classified into characters. 26 features were used for both word and character. RBF, Polynomial and Kernel were used to train the dataset using SVMLight.

In [9] the author developed a writer identification model using the handwritten documents. The noise in the scanned images is removed by locating the text lines and empty spaces and the height of the text are normalized. The features are extracted using the generalized gaussian density wavelet transform. Weighted Euclidean distance (WED) were used for classification.

Authors in [11] implemented writer identification model in which normalization and Empirical mode decomposition technique was used. Intrinsic Mode function was used to extract the features. IMFs of each function differ from one another. It contains significant information of each handwritten document. The feature vector was classified using K-nearest neighbor.

In [19] the authors proposed writer identification and verified using tamil handwritten words. In scanned images the following operations were performed such as segmentation, noise removal, binarization, edge detection and thinning. Global Features and Local Features are used to extract the features. Three supervised learning algorithms such as Naïve bayes classifier, Decision Tree induction and k-Nearest Neighbor were used for learning the classification model.

In this research work [20] the authors proposed a writer identification model using document images. The normalization technique was applied to normalize the skewed word and features are extracted using the multi-channel Gabor filtering and gray scale co-occurrence matrix. Here two classifiers k-nearest neighbor (KNN) and weighted Euclidean distance (WED) were used for classification.

Author in [21] proposed a writer identification model using document images. The noise in the scanned images was removed by locating the empty spaces and text lines. Gabor, GGD and Contourlet GGD are used to extract the features from the preprocessed image and the performance of the model was evaluated using the features.

In the research work [22] the authors developed a writer identification model using the handwritten document. The noise in the images is removed on both the foreground and background of the image.

The edge based directional feature and edge hinge distribution was used to extract the features from the preprocessed image. K-nearest neighbors (KNN) were used for classification.

From the background study, it was observed that the writer identification problem can be modeled as pattern classification task and can be solved using supervised learning techniques. As machine learning technique can automatically learn the model by taking intelligent hints from the training data

and predicts the output more accurately, it has been influenced in this work to pool the various features of handwritings for making the training dataset. Various phases of the proposed implementation are described in Section 3.

3 Proposed Model for Writer Identification

This work aims to develop a discriminative model for identifying a person using the handwritten documents. The language considered here is English and the handwritings have been collected from different writers and preprocessing is done using the normalization technique. The features are extracted using Gabor, GLCM, GGD, Contourlet GGD and directional features. Standard classification algorithms have been employed to build the model and to predict the writer's identity.

3.1 Data Acquisition

The data acquisition is an important task in writer identification. Text dependent data has been acquired from ten different writers. The paragraphs are scanned using scanner of resolution 300 dpi and a total of 300 JPEG text images are collected from 10 writers at a rate of 30 paragraphs per writer are obtained.

3.2 Preprocessing

Preprocessing is an important task in any mining activity. Here normalization technique is performed prior to preprocessing in order to correct the skewed words in the handwriting image. The space between vertical and horizontal lines has been normalized to produce a well-defined pattern for texture analysis. Then the scanned images are preprocessed to remove the noise and converted into grayscale images to carry further preprocessing tasks. The preprocessing tasks carried out here are edge detection, image dilation and box bounding.

Edge detection: Edges in the binary image are detected using sobel method. The Sobel method finds edges using the Sobel approximation to the derivative. It returns edges at those points where the gradient of I is maximum. Edge ignores all edges that are not stronger than threshold. If threshold do not specified, or if threshold is empty ([]), edge chooses the value automatically.

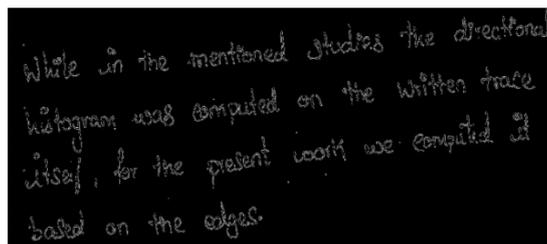


Figure 1: Edge detection

Dilation: The dilation is a fundamental morphological operation. It helps to add the pixels to the boundaries of the images. Based on the size and shape of the structuring element the pixels are added. Once the image is converted into grayscale image, the dilation operation is performed.

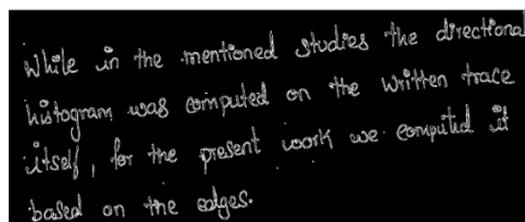


Figure 2: Dilated image

Box bounding: After the image is converted into dilated image, the pixels in the image are converted into white pixels. To box bound a word, the white pixels are taken for each character with a pixel space of 30 pixels. Once it exceeds the pixel space of 30 then a word is box bounded and it starts from the next word to box bound until it exceeds the space of 30 pixels.

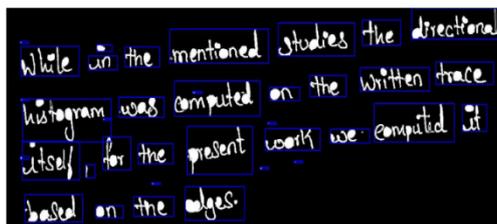


Figure 3: Box bounded image

3.3 Feature extraction

Feature extraction plays a vital role in improving the classification effectiveness and computational efficiency. It is used to extract the writer invariance in the form of a pattern. The various features used in preparing the training dataset are described below.

Gabor Filter

The Gabor filter is a bandpass filter used in the image processing for feature extraction. The Gabor filter requires an input image with $N \times N$ pixel image along with the frequency (f) and an angle θ . Here the θ and f specifies the location for the Gabor filter. As the size of the image is $N \times N$, the frequency used here is 4, 8, 16 and 32 cycles/degree. The parameters used in this filter are bandwidth, phase shift and lambda.

For each central frequency f , filtering is performed at 0, 45, 90 and 135 degrees. This gives a total of 16 output images for each frequency, from which the writer's features are extracted. These features are the mean and the standard deviation of each output image and the features are extracted from the images.

GLCM

GLCM is a matrix where number of rows and columns is equal to number grey levels G in an image. It is defined over an image to be the distribution of co-occurring values in the given offset. It is a way of extracting second order statistical features. It is used to measure the spatial relationships between pixels. This method is based on the belief that texture information is contained in such relationships. The GLCMs are constructed by mapping the grey level co-occurrence probabilities based on spatial relation of pixels in different angular direction. Greycomatrix function creates the GLCM by calculating how often pixels with grey-level value i occurs horizontally adjacent to a pixels with the value j . Each element (i, j) in GLCM specifies the number of times that the pixels with values i occurred horizontally adjacent to a pixels with value j .

There are 22 texture features associated with GLCM. They are autocorrelation, contrast, correlation, Cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares, sum of average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1, information measure of correlation 2, inverse difference homogeneity, inverse difference normalized, Inverse difference moment normalized.

Generalized Gaussian density

The basic idea of the wavelet-based GGD method is to establish corresponding wavelet-based GGD model for a handwriting image. Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. The parameters α, β are regarded as the features of the handwritten image. For each handwritten image, the GGD model cut the image using different frequency components are cut into regions. Each region is called as a wavelet subband. For each wavelet subband the probability is maximized for the estimated parameters α, β which are optimal for improving accuracy of writer identification. The probability is estimated as,

$$P(\{\alpha, \beta\}/X)$$

Each wavelet subband has coefficients as $X=\{x_1, x_2, \dots\}$. After the error probability is estimated the likelihood function is defined as,

$$L(X|\{\alpha, \beta\}) = \log \prod_{i=1}^N P(x_i|\{\alpha, \beta\}).$$

According to the langrange optimization, the likelihood equation is obtained as follows,

$$\frac{\partial L(x|\{\alpha, \beta\})}{\partial \alpha} = -\frac{N}{\alpha} + \sum_1^N \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha},$$

After the likelihood function is estimated then α, β values are obtained using the following equations,

$$\begin{aligned} \frac{\partial L(x|\{\alpha, \beta\})}{\partial \beta} &= -\frac{N}{\beta} + \frac{N\varphi(1/\beta)}{\beta^2} \\ &\quad - \sum_{i=1}^N \left(\frac{x_i}{\alpha}\right) \log\left(\frac{x_i}{\alpha}\right) \end{aligned}$$

where $\varphi(z) = \tau(z)/\tau(z)$

$$\hat{\alpha} = \left(\frac{\beta}{N} \sum_{i=1}^N |x_i|^\beta\right)^{1/\beta} \quad \text{--1}$$

The estimation of the β is be calculated by using the following equation,

$$\begin{aligned} 1 + \frac{\varphi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^N |x_i|^{\hat{\beta}} \log|x_i|}{\sum_{i=1}^N |x_i|^{\hat{\beta}}} \\ + \frac{\log((\hat{\beta}/N) \sum_{i=1}^N |x_i|^{\hat{\beta}})}{\hat{\beta}} = 0. \end{aligned}$$

Once the value of the β is calculated then it is substituted in the equation 1 to find the value of α . By this process α, β values are obtained for each pixel and the features have been obtained for all the input image.

Contourlet Generalized Gaussian density

In GGD, wavelet transform is used to decompose the image into subbands in different frequency and orientation. But wavelet is able to capture only limited directional operation which is an important issue in image analysis. To overcome this problem, multiscale and directional representations can be used to efficiently capture the image's geometrical structures such as edges or contours.

Contourlet transform based on an efficient two-dimensional multiscale and directional filter bank can deal effectively with images having smooth contours. The contourlets are implemented using the double filter bank named pyramidal directional filter bank (PDFB). Laplacian pyramid is used in PDFB to decompose the images into multiscale using 9-7 filter bank. The directional filter bank is used to analyze the multiscale into a four directional subbands. Multiscale and directional decomposition stages in the contourlet transform are independent of each other, because while using PFBF filter it gets a cascade structure. The parameters α , β of the GGD model is taken and used to represent the contourlet subband. The GGD is given as below,

$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\tau(1/\beta)} \exp^{-1(|x|/\alpha)^\beta}$$

Where $\tau(\cdot)$ is the Gamma function, i.e.

$$\tau(\cdot) = \int_0^\infty \exp^{-t} t^{Z-1} dt, Z > 0.$$

Various methods are available to estimate the parameters α and β . Here the maximum likelihood estimation is used to estimate the parameters α , β by converting the subband image into multi-dimensional vector and it is defined as given below,

$$L(x; \alpha, \beta) = \log \prod_{i=1}^L p(x_i; \alpha, \beta)$$

By using the following equation the features are extracted for the subband images.

$$\frac{\partial L(x; \alpha, \beta)}{\partial \alpha} = -\frac{L}{\alpha} + \sum_1^L \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha}$$

The values of the GGD are taken as input and the filters mentioned above are applied to extract the features from a given input image. In this manner the features are obtained for all the images.

Directional features

The normalized feature vector for classification is obtained using the directional features here. First the input image is characterized into four types, such as vertical line, horizontal line, left diagonal and right diagonal. The values are calculated from the four directions, as fraction of the distance traversed across the image. If the transition is computed from left to right, a transition found close to the left is assigned a high value compared to a transition computed further to the right. A maximum value (MAX) is defined as the largest number of transitions that is recorded in each direction. If there are less than MAX transitions recorded, then the remaining MAX transitions are assigned values of 0.

The transition value is calculated for a particular direction. To calculate the directional transition, the transition value is divided by a predetermined number. Here the predetermined number is 10. Thus eight features are obtained for one transition. Then this process is repeated for the remaining transitions and the features are obtained, using the following formula,

$$\text{nrFeatures} * \text{nrTransitions} * \text{nrVectors} * \text{resampledMatrixHeight (Width)}$$

The above features are extracted by developing MATLAB code and feature vectors are generated for all the input images. Particle swarm optimization (PSO) method is used to select the contributive feature and to reduce the dimension of the feature vector.

Feature Selection

The feature selection method used here is particle swarm optimization (PSO) method. The PSO method is used to reduce the number of features obtained during the feature extraction. The feature selection is done in order to speed up the processing rate and predictive accuracy. The features are extracted for each and every pixel for a given input image. So to reduce the dimension of feature vector, feature selection method is used. If there are n number of features, then the threshold value of the feature selection is $n < 10$. The features are reduced based on the weight assigned to each feature. After the weight is assigned to each feature, the features are selected in the descending order based on the weight of the features.

4 Experiments and Results

The writer identification model proposes the application of computational intelligence technique to develop discriminative model for a person identification using their handwriting. The language considered here is English and the handwritings have been collected from 10 persons as described in section 3(A). Scanned images of handwritten documents are used to prepare the dataset. The scanned images are normalized as explained in section 3(B). Features like Gabor, Gray level co-occurrence matrix, generalized gaussian density, Contourlet generalized gaussian density and Directional features are taken into account as described in the section 3(C).

Since the total number of handwritten document images considered is 300, the training dataset with 250 instances is developed. Three independent experiments have been carried out with three different datasets and by training the standard classification algorithms naive bayes, decision tree, KNN, support vector machine and neural networks. The dataset is prepared in .CSV format in order to classify the instances using R. The built classifiers have been evaluated using hold out method and the results are analyzed.

For the first experiment, the features associated with Gabor filter, Gray level co-occurrence matrix, generalized gaussian density and Contourlet generalized gaussian density are aggregated and the training dataset is developed. The dataset is then learned using above mentioned classification techniques and their performance is evaluated using measures such as accuracy, precision, recall and F measure. The results of the experiment are shown Table I and illustrated in Figure 1.4.

Table 1: Results of classifiers based on GGD and Contourlet GGD features

Alg	Accuracy	Precision	Recall	Fmeasure
NB	63%	0.723	0.515	0.638
DT	85%	0.912	0.749	0.852
KNN	82%	0.901	0.721	0.823
NN	69%	0.689	0.678	0.699
SVM	75%	0.735	0.755	0.744

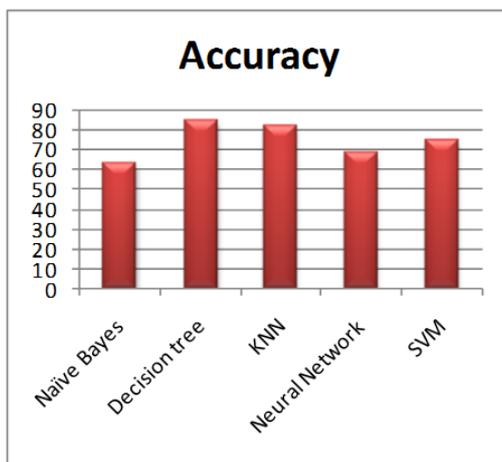


Figure 4: Performance of classifiers based on GGD and Contourlet GGD features

To carry out the second experiment, along with texture features i.e. Gabor and GLCM, directional features are combined to develop training dataset. The dataset is then trained and the classifiers are evaluated. The results of the classifiers in terms of accuracy, precision, recall and F measure are shown in Table II and illustrated in Figure 1.5.

Table 2: Results of classifiers based on Directional features

Alg	Accuracy	Precision	Recall	Fmeasure
NB	70%	0.826	0.635	0.719
DT	61%	0.768	0.524	0.618
KNN	77%	0.723	0.756	0.788
NN	70%	0.711	0.704	0.707
SVM	60%	0.725	0.512	0.663

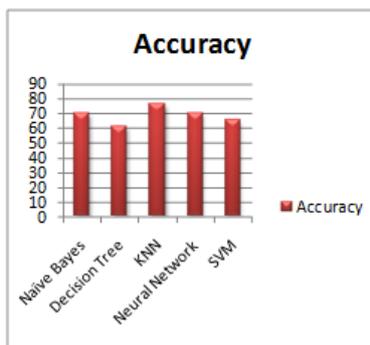
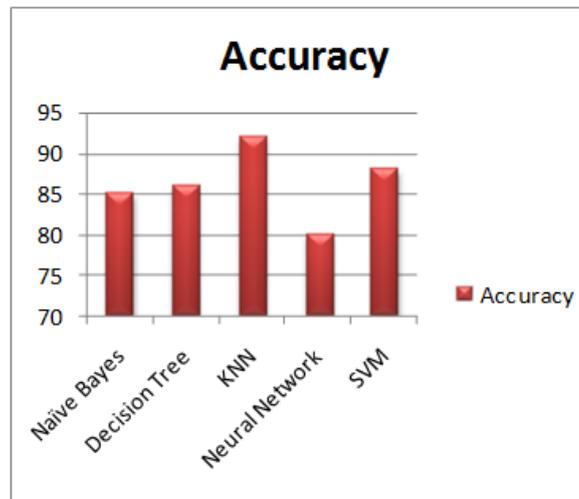


Figure 5: Performance of classifiers based on Directional features

In the third case, all five categories of features are pooled to form a training dataset. The results of learning the classifiers are shown in Table III and illustrated in below.

Table 3: Results of classifiers based on aggregated features

Alg	Accuracy	Precision	Recall	Fmeasure
NB	85%	0.895	0.876	0.852
DT	86%	0.923	0.756	0.861
KNN	92%	0.922	0.927	0.929
NN	80%	0.974	0.715	0.823
SVM	88%	0.865	0.871	0.867

**Figure 6: Performance of classifiers based on aggregated features**

From the above experiments it was observed that the performance of the classifiers is high when training dataset contains aggregated features. The classification models built using GGD and Contourlet GGD features produced an accuracy of about 82%. The classification models built using directional features produced an accuracy of about 77%. When all the features Gabor, GLCM, GGD, Contourlet GGD and directional features are combined together, the models showed an accuracy of about 92%. Hence it is concluded that the pooled features can better produce a trained model for accurate writer identification.

5 Conclusion

This paper demonstrates the modeling of writer identification as classification task and describes the implementation of supervised learning approach for identifying the writer based on their handwriting. Various features such as Gabor, GLCM, GGD, Contourlet GGD and directional features are pooled to develop the training datasets and are used to fabricate the models. The outcome of the experiments proves that, the writer identification model is effectual when the collective features are used in learning. Further work can be extended for text independent handwriting.

REFERENCES

- [1] Eibe Frank, Ian H. Witten. (2005), Data Mining – Practical Machine Learning Tools and Techniques. Elsevier Gupta GK “Introduction to Data Mining with Case Studies”.
- [2] Mitchell T. “Machine learning”, Mc Graw-Hill International edition.
- [3] Sreeraj.M and Sumam Mary Idicula. (2011), “A Survey on Writer Identification Schemes”, International journal of computer applications, Vol. 26, No. 2.

- [4] Marius Bulacu, Lambert Schomaker, Louis Vuurpijl. (2003), “Writer Identification Using Edge-Based Directional Features”, Proceedings of the Seventh International Conference on Document Analysis and Recognition.
- [5] Saranya K, Vijaya M.S (2013) “An interactive tool for writer identification based on offline text dependent approach”, International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No. 1.
- [6] Plamondon, R., Lorette, G. (1989) Automatic Signature Verification and Writer identification —The State of the Art, || Pattern Recognition, vol. 22, no. 2, pp. 107-131.
- [7] H. E. S. Said¹, G. S. Peake¹, T. N. Tan² and K. D. Baker¹ “Writer Identification from Non-uniformly Skewed Handwriting Images”.
- [8] Saranya K, Vijaya M.S (2013) “Text Dependent Writer Identification using Support Vector Machine”, International Journal of Computer Applications (0975 8887) Volume 65 No.2
- [9] Zhenyu, H., Xinge, Y., Tang, Y.Y. (2008) “Writer Identification using global wavelet-based features” neurocomputing 71, 1832–1841.
- [10] Plamondon, R., Lorette, G. (1989) Automatic signature verification and Writer identification—The State of the Art, Pattern Recognition, vol. 22, no. 2, pp. 107-131
- [11] Karukara K and Dr. B.P. Mallikarjunasamy “Writer Identification based on offline handwritten Document images in Kannada language using Emprical mode decomposition method” Volume 30– No.6, September 2011.
- [12] Zhu, Y., Tan, T., Wang, Y. (2001) Font Recognition Based on Global Texture Analysis, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 10, pp.1192- 1200.
- [13] Zhang, B., Srihari, S. (2003) Analysis of Handwritten Individuality Using Word Features, Proc. Seventh Int’l Conf. Document Analysis and Recognition (ICDAR), pp.1142-1146.
- [14] Bensefia, A., Paquet, T., Heutte, L. (2005) —A Writer Identification and Verification system, Pattern recognition system, vol. 26, no. 10, pp. 2080-2092
- [15] Bensefia, A., Nosary, A., Paquet, T., Heutte, L. (2002) Writer Identification by Writer’s Invariants, Proc. Eighth Int’l Workshop Frontiers in Handwriting Recognition, pp. 274-279.
- [16] Marti, U.V., Messerli, R., Bunke, H. (2001) —Writer Identification Using Text Line Features, Proc. Sixth Int’l Conf. Document Analysis and Recognition (ICDAR), pp. 101- 105.
- [17] Hertel, C., Bunke, H. (2003) —A Set of Novel Features for Writer Identification, Proc Fourth Int’l Conf. Audio and Video-Based Biometric Person Authentication, pp. 679-687.

- [18] Schlapbach, A., Kilchherr, V., Bunke, H. (2005) Improving Writer Identification by Means of Feature Selection and Extraction, Proc. Eighth Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 131-135.

- [19] Thendral, Vijaya.M.S., "Supervised Learning Approach for Tamil Writer Identity Prediction using Global and Local Features".

- [20] H. E. S. Said¹, G. S. Peake¹, T. N.Tan² and K. D. Baker¹ "Writer Identification from Non-uniformly Skewed Handwriting Images".

- [21] Z.Y.He, " A Contourlet based method for writer identification"

- [22] Marius Bulacu, Lambert Schomaker, Louis Vuurpijl. (2003), "Writer Identification Using Edge-Based Directional Features", Proceedings of the Seventh International Conference on Document Analysis and Recognition