Novel Entropy Coding and its application of the Compression of 3D Image and Video Signals

A Thesis submitted for the degree of Doctor of Philosophy

by

Amal Mehanna

School of Engineering and Design, Brunel University

June 2013

Abstract

The broadcast industry is moving future Digital Television towards Super high resolution TV (4k or 8k) and/or 3D TV. This ultimately will increase the demand on data rate and subsequently the demand for highly efficient codecs. One of the techniques that researchers found it one of the promising technologies in the industry in the next few years is 3D Integral Image and Video due to its simplicity and mimics the reality, independently on viewer aid, one of the challenges of the 3D Integral technology is to improve the compression algorithms to adequate the high resolution and exploit the advantages of the characteristics of this technology.

The research scope of this thesis includes designing a novel coding for the 3D Integral image and video compression. Firstly to address the compression of 3D Integral imaging the research proposes novel entropy coding which will be implemented first on 2D traditional images content in order to compare it with the other traditional common standards then will be applied on 3D Integra image and video. This approach seeks to achieve high performance represented by high image quality and low bit rate in association with low computational complexity.

Secondly, new algorithm will be proposed in an attempt to improve and develop the transform techniques performance, initially by using a new adaptive 3D-DCT algorithm then by proposing a new hybrid 3D DWT-DCT algorithm via exploiting the advantages of each technique and get rid of the artifact that each technique of them suffers from.

Finally, the proposed entropy coding will be further implemented to the 3D integral video in association with another proposed algorithm that based on calculating the motion vector on the average viewpoint for each frame. This approach seeks to minimize the complexity and reduce the speed without affecting the Human Visual System (HVS) performance. Number of block matching techniques will be used to investigate the best block matching technique that is adequate for the new proposed 3D integral video algorithm.

ACKNOWLEDGMENT

My deepest gratitude goes to Dr. Amar Aggoun, who has been such an inspirational supervisor, with his insightful criticism and patient encouragement, and giving me invaluable help with my thesis.

I wish to express my sincere thanks to a number of key people who, without their dedication and help, my thesis would not have been possible:

Prof. Sadka, Chair in Electronic and Computer Engineering.
Prof. John Cosmas, Professor of Multimedia Systems.
Dr. Emmanuel Tsekleves Lecturer and Course Director in Multimedia Technology and Design.

In addition to the afore mentioned, I would also like to convey my deepest appreciation to Dr. Alaister Duffy, Reader in Electromagnetics, School of Computer Science and Informatics, in DMU, for teaching me the principles of research and technology.

DEDICATION

To my parents, Prof. Safwat Mehanna and Dr. Mona El-Shebini, who I miss dreadfully, I dedicate my thesis. Their endless love, sacrifice and support, has enabled me to pursue my career and direction in life, with my beautiful family.

Sincere thanks to my wonderful husband, Ahmed Lelah, who has been such a support with his patience, understanding, and endless motivation. I would not have accomplished my thesis without his inspiration, energy and unconditional love, which I will forever be eternally grateful.

Many thanks to my beautiful children Janna and Farouk, for the sacrifices they have had to make to enable me to complete my thesis to such a high standard.

A final acknowledgement to all my colleagues, Ghaydaa, Rafiq, Obaid, Nawas, Sanusi, Moheb, Umar Abu-Bakr, Jawed, Carlos, Ibrahim, Abd-El-Kader, Peter, and my friends, Natasha, Peter Oliver and Karen for your understanding and support during my many moments of crisis.

Contents

List of Figures

Number *Page*

List of Tables

Number *Page*

LIST OF ABBREIATIONS AND ACRONYMS

2D DCT	Two.dimensional Discrete Cosine Transform
3D DCT	Three.dimensional Discrete Cosine Transform
3DTV	Three.dimensional Television
AEOB	Adaptive End of Block
ARPS	Adaptive Rood Pattern Search
bpp	Bits per pixel
CWT	Continuous Wavelet Transform
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DHT	Define Huffman Table
DPCM	Differential Pulse Code Modulation
dpi	Dots per inch
DS	Diamond Search
DWT	Discrete Wavelet Transform
EBCOT	Embedded Block Coding with Optimal Truncation
EZW	Embedded Zero Tree
FDCT	Forward Discrete Cosine Transform
FDWT	Forward Discrete Wavelet Transform
FNPA	Forward Negative Positive Algorithm
FS	Full Search
FSS	Four Step Search
HDTV	High.Definition Television
HT	Haar Transform
HVS	Human Visual System
IAEOB	Inverse Adaptive End of Block
IDCT	Inverse Discrete Cosine Transform
IDWT	Inverse Discrete Wavelet Transform
II	Integral Image
INPA	Inverse Negative Positive Algorithm
IP	Integral Photography
JPEG	Joint Photographic Experts Group
LCD	Liquid Crystal Display
MC	Motion Compensation
ME	Motion Estimation
MPEG	Moving Picture Experts Group
NTPA	Negative to Positive Algorithm
NTSS	New Three Step Search
OII	Omni.directional Integral Image
p.m.f	Probability Mass Function
PSNR	Peak Signal to Noise Ratio
QF	Quantization Factor
RD	Rate Distortion

RLE	Run Length
SH	Statistical Huffman
SPHIT	Set Partitioning in Hierarchical coding Techniques
UII	Unidirectional Integral Image
VP	Viewpoints

Chapter 1

Introduction

1.1 Overview

Recently, 3D technologies, such as stereoscopic technology have been at the forefront of the production and consumer market. At the same time and due to limitations of the current 3D technology, research has focused on the next generation of more sophisticated depth-illusion technologies, such as the promising Holoscopic technology (also referred to as Integral Imaging Technology) [1] developed within the project "3D VIVANT" [2][3]. Holoscopic technology mimics insects visual perception system, such as that of a fly, where a large number of micro-lenses are used to capture a scene, offering rich parallax information and enhanced 3D feeling without the need of wearing specific eyewear [4].

Due to the large amount of data required to represent the captured 3D integral image with adequate resolution, it is necessary to develop compression algorithms tailor to take advantage of the characteristics of the recorded integral image [5].

Data compression is still one of the hottest research areas in 2D and 3D image and video processing. Compression is based on exploiting the redundancy data, by converting the correlated data to de-correlated data [6]. This process occurs between each pixel and its neighbours in both space and time. However, due to the small angular disparity between adjacent micro-lenses, there is a high cross correlation between the sub-images of the captured integral image intensity distribution in the third dimension due to the recording micro-lens array. It is therefore necessary to evaluate both inter and intra sub-image correlation in order to maximise the efficiency of compression scheme for use with the integral image intensity distribution [5].

1.2 Research Challenges in image coding

Due to the need of data compression as shown in the above example, many studies have been done to improve the compression process in order to transmit it or store it. The original raw data (whether it is video or image, 2D or 3D) can be compressed in two ways, either lossless or lossy techniques.

Lossless compression leads to a perfect reconstructed data, but at the expense of a relatively low compression ratio [7], On the other hand the lossy technique will end up with high level of compression ratio, which is better for the channel bandwidth if the data is going to be transmitted through a channel or the memory capacity if it is going to be saved, but the reconstructed data will not be as perfect as the raw input data.

Many studies have been done to achieve efficient compression ratio (CR) and better quality of information. The main parameters that affect the compression process are the compression ratio (CR), data content quality and computational complexity. Reducing the computational complexity is very important especially if the application is going to be used in real-time and/or on mobile devices [8].

Video Coding standards are continuously being developed and improved to achieve better compression results but at the expense of the computational complexity and/or video quality. So for example in H.264 in order to get high compression rate and better quality, the whole process contains computational complexity that brought an additional the time of compression [9][10]. On the other hand, in image coding, one of the JPEG features is its simplicity, but still the system suffers from some drawbacks in terms of compression ratio especially at very low bit rate [11].

1.3 Compression Process

The compression process is mainly divided into levels or stages as shown in figure 1.1, first, a transform coding followed by quantization method then the coding technique is applied. Most of the Transform coding is lossless but because of the quantisation step, it becomes lossy. The entropy coding part is also a lossless part [12].

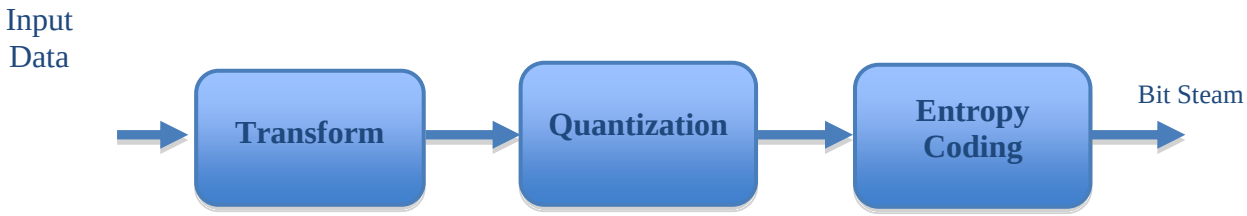


Figure 1.1 Encoder Block Diagram

The type of transform coding and the way it is used affects the image quality. It can be said that it has a direct effect on the frame quality of video codec and image quality in the image codec. While the Entropy coding part is affecting the file size and determines the compression ratio, so the way the quantized coefficient is encoded is the main factor that verifies the bit rate. This process is reversible.

Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are the two main transform techniques being used in most video coding standards such as JPEG, MPEG, and H.264. Also there are two main coding techniques that have been commonly used, Huffman Coding and Arithmetic Coding for JPEG and JPEG2000

respectively [13]. In H.264 another coding technique is used which is context-based adaptive variable length coding (CAVLC) and context-based adaptive binary arithmetic coding (CABAC) [14].

1.4 Objectives for research, Motivation and Contributions

1.4.1 Objectives:

The scope of the thesis is to improve the efficiency of the compression process by decreasing the data file size with keeping the same high image quality and also reduce the computational complexity. In usual, it is a trade off between quality and Computational complexity but the target here is to manage both and to optimum the three sides of the compression process with minimizing the compromises.

Briefly, in the thesis an investigation into the coding 3D integral images is being made. Several algorithms have been applied in order to achieve the balance between the three main parameters; Quality, Compression Ratio and Computational Complexity.

1.4.2 Motivation of the thesis:

The aim is to optimize the compression process through developing the transform technique and implementing new entropy coding that overcomes the drawbacks the other entropy coding deteriorates.

Due to this drawbacks of the other entropy coding techniques, denoted by the trade-off between the high coding ratio and the complexity of the algorithm. It is either obtain a high compression ratio but at expense of high computational complexity or lower data compression but with reduced computational time and cost. In view of this a proposed algorithm for entropy coding that tries to compromises the equation by optimizing the three factors at the same time.

The most common two transform techniques that have been proofed there efficiency in the last decade have been used in order to improve the compression process through developing them (DCT and DWT) via several ways that will be presented

later. For example, one of the well known facts that the DWT works in a very high levels in the medium and the high compression domain, but extremely bad with low compression domain in comparison with the DCT technique, this fact could be taken from another aspect that the DCT is working effectively in the low domain, despite the drawbacks of each system because each technique suffers from some artefacts, both techniques can be used together to exploit and get privilege of the advantages of each one and overcome the disadvantages. For example, the DCT suffers from blocking artefact due to its localization, at the same time the DWT endure ringing effect, so applying the DWT as a global technique first then applying the localize DCT technique can improve the performance of this hybrid technique rather than use each one separately.

1.4.3 Contributions of the thesis:

A novel entropy coding, is first be implemented on 2D images to proof the algorithm efficiency via comparing it to the 2D JPEG baseline standard first. Later, it will be implemented using 3D Integral content as an input considering the different types of the data weather image or video, full-parallax or unidirectional content. The rest of the thesis is applied only to 3D integral images and video.

The main purpose of the second part is to develop the performance of the 3D-DCT and this is achieved through several ways. First, by using a new proposed adaptive technique, however, the adaptive concept is well known and been used long time ago, but the way it is represented here in this thesis is different, and the results will show this differences. Beside that most of the time the input was 2D data, but in this thesis the 3D integral data is used which have a special characteristics. The 3D Integral image data have been used as an input to an adaptive algorithm proposed in the published paper [29], but the results of the proposed algorithm in this thesis will show more improvements in the compression process.

There is a well-known fact that DWT is working effectively in a medium and high compression but not good in low compression. If it is taken from a different aspect it could be said that the DCT is working better in the Low compression more than the DWT.

Both techniques got their drawbacks blocking artefacts in DCT and edging artefact (ringing effect) in DWT. But using first the global DWT then localize it by using DCT improves the compression performance.

The DCT have been used in [83] before but only for the low low but with testing and scaling the size of each band, it have been found that the high low suffers also from low compression ratio due to its high file size. So the proposed DCT proposed localization algorithm have been implemented to both low low band and high low band. It is also applied in association with the proposed entropy coding.

Also a number of experiments are implemented to proof that the order is right meaning that applying the global technique followed by localization it gives better results than the other way round, but also it have been approved that following this hybrid technique with one dimension DWT in the third domain one more time improve the performance more and that is due to the lack to de-correlation in this axis. It needs to de-correlate it more.

The new proposed entropy coding will be used as a block in the video block diagram a small change in the motion estimation will occurs built on the way the 3D integral video viewpoints extraction proposed. Instead of get the motion vector for all viewpoints; only one motion vector will be used from the middle viewpoints for all the other viewpoints.

The process will slightly decrease the image quality, the computational complexity will be reduced on account of the image quality but from the aspect of HVS the quality of image will be the same as the original, which will not affect the reconstructed video (Frames).

The following bulleted points summarised the novel contributions achieved in this thesis work:

- Anew entropy-coding algorithm has been developed and evaluated. A comparison with existing entropy coding techniques showed higher efficiency in term of compression ratio, image quality and computational

complexity.

- A proposed new coding technique is introduced and applied on 2D-Images, in order to evaluate its efficiency with other coding techniques, it is tested on different type of Integral Images.

- Two methods are introduced for quality optimization of 3D images as follows

    - First method is by applying 3D-Discrete Cosine Transform using two different adaptive block size schemes on the 3D Integral images.
    - Second method is by using hybrid transform techniques in different methodology, through applying the most two common Transform standards (3D-DCT and 2D-DWT) in a way to combine their advantages and overcome their drawbacks. Using both techniques in a cooperative way to achieve better results through different algorithms is represented.

- Finally, the proposed coding technique has been used with 3D video compression in order to reduce computational time. Additionally, there are several motion estimation techniques that have been implemented on the 3D-Integral Video to recognize the best block matching method that suits to the 3D- Integral Video by computing the computational complexity for each method and evaluate its PSNR.

1.5 Thesis Organization

Chapter 1 introduces the main subject of the thesis. In chapter 2, Literature review is represented starting with an overview of the ways and methods the 3D Integral image is captured and analyzing the different types of 3D Techniques. Later brief analysis of the compression standards and different entropy coding is shown in general and Huffman coding in particular.

In chapter 3, a new proposed coding algorithm is represented and its comparison with the Huffman coding has been applied on the 2D traditional Image using 2D-DCT and further implemented using 2D-DWT techniques.

Chapter 4, describes the application of the proposed entropy coding on 3D Integral images. An attempt to improve the 3D DCT transform coding is also represented by using two different proposed adaptive techniques. Comparing results with an adaptive technique is applied also on Integral Image. In addition a hybrid technique is presented using both DWT and DCT to improve the PSNR (the quality of the reconstructed Images).

In Chapter 5, a video coding technique is applied on 3D Integral video including extracting Viewpoints and the algorithm uses 3D-DCT transform with the proposed entropy coding technique. Different motion estimation techniques are applied to valuate the introduced technique.

Finally, in chapter 6, the overall conclusion and future work are presented.

Chapter 2

Literature Review

This chapter is divided into two sections, 3D-Display Technology and Image and Video compression techniques. The former section illustrates different methods of displaying and capturing 3D-Integral Images, while the latter represents, an overview of the different compression standards in general, illustrating some of the drawbacks and highlighting the "JPEG-Baseline Standard" that is to be used throughout the thesis to compare with the different proposed algorithms. Brief information about entropy coding will be mentioned, whilst highlighting and investigating the advantages and disadvantages of using only one

of these techniques which is the Huffman Coding, due to its popularity among the other techniques, finally the Measuring Quality method will be represented.

2.1 3D- Display Technology

Within the 3D industry, displays have a key role, with their advance technology and ways of visualising 3D information setting the rules of content, capturing and processing, resulting in a need for research in this domain as well. 3D perception of the human vision system is a complex process containing many different components. These include pre-learned interpretation of 2D images; unconscious actions that induce motion parallax, but most significantly the binocular parallax of the two eye positions.

3D Display technologies can be divided into two categories as shown in Figure 2.1, namely stereo display (requiring headgear) and auto-stereoscopic display (free viewing)[15][16].

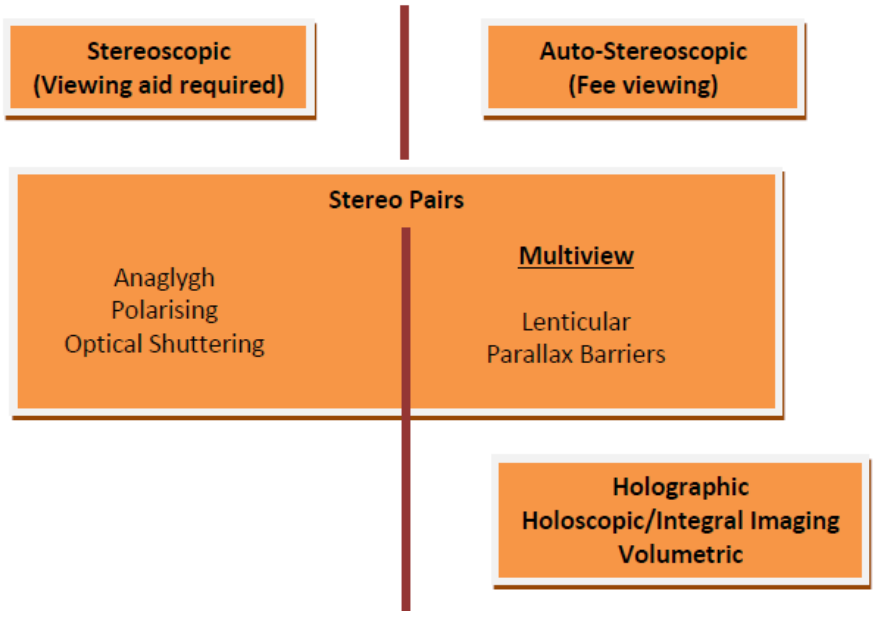


Figure 2.1 State of the art 3D Display technologies

2.1.1 Stereoscopic Displays

Indeed, today's 3D display technology is based on stereovision, where the viewer wears glasses to present the left and right eye image via spatial or temporal multiplexing [17].

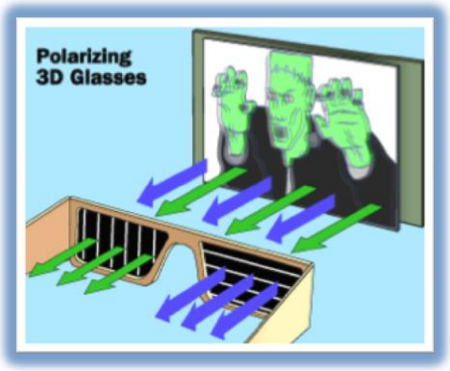


Figure 2.2 polarizing 3D glasses [2][3].

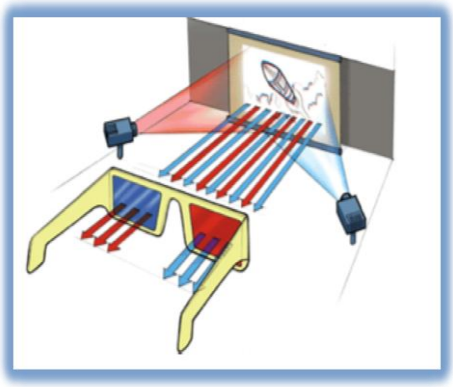


Figure 2.3 Anaglyphs [17].

2.1.2 Auto-stereoscopic displays

The alternative **Auto-stereoscopic displays** present 3D views without having to wear glasses or any other headgear; however *direction selective light emission* is required to provide the 3D perception. A various optical or lens raster technique is applied to the left and right images directly above the LCD surface.

2.2 Multiview autostereoscopic

The next step in the 3DTV development is the **Multiview autostereoscopic** systems. The combination of a lenticular and an electronic display provides an optically efficient way of creating an electronic 3D display which does not require special glasses [18]. The use of micro-optics with an LCD element becomes available with other several 3D display systems in commercial applications.

Auto-stereoscopic displays have been established by using a combination of optical elements and LCD:

- Parallax barriers, optical apertures aligned directly above LCD surface.
- Lenticular optics, cylindrical lenses aligned directly above the LCD surface.

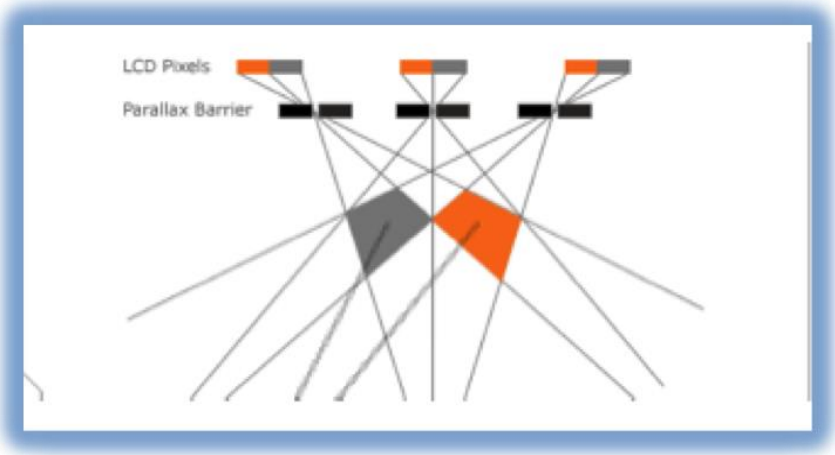


Figure 2.4 Parallax Barrier Technology [2][3].

For the multiview 3D displays, the cylindrical lenses aligned vertically with the 2D display similar to LCD. Capturing multiview images using different cameras allows the option of viewing 3D content by more than one viewer [19].

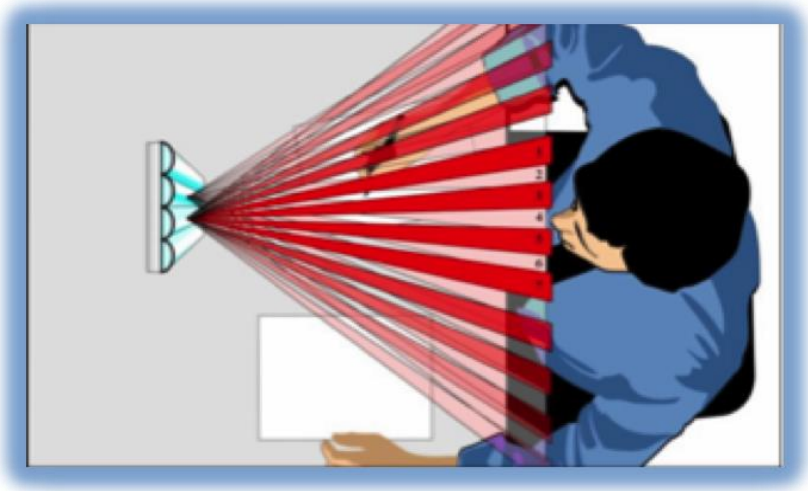


Figure 2.5 Lenticular Technology [19].

The multi-view approach has various drawbacks. One of them, referred to as a 'black mask', appears as dark lines for the viewer in each window. It happens when the viewer's eye crosses the window boundary. Another disadvantage is the 'image-flipping artefact', which occurs when the viewers move their eyes from one view window to the other [2], and there is also a fast decrease of the horizontal resolution with the number of views.

These problems can be solved by locating the lenticular array at an angle to the LCD pixel array, giving better features for each view by splitting the LCD display horizontally and vertically rather than horizontally only. This combination of the adjacent views eliminates the image flipping problems and reduces the black mask effect by spreading it [2]. "The slanted lenticular arrangement will require sub-pixel mapping to allow all pixels along a slanted line to be imaged in the same direction as well as increasing the resolution of the 3D images" [2].

One of the best multiview displays currently available is the Dimenco's Display, and despite the eye fatigue problem in traditional stereo displays, Dimenco's Display has been chosen by MPEG experts as a reference for defining auto-stereoscopic standards [20].

2.13 Holography & Holoscopic Imaging

The researchers have been motivated in seeking alternative methods for capturing true 3D content, two of the most recognised being ***holography*** and ***holoscopic*** Imaging [21]. The requirement to use a coherent light field to record holograms, limited the use of the Holographic technique, however, Holoscopic imaging (also referred to as Integral Imaging) is a simple technique. Each lens captures perspective views of the scene and the technique uses a number of lens arrays working in association with photographic film or digital sensors. Unlike Holography, no coherent light field is required, with 'holoscopic' colour images being achieved with full parallax [3][22][23].

Figure 2.6 Holographic Image [23].

2.4 Integral Image History

3D Holoscopic methodology imaging also referred to, as Integral Images [24] was first proposed by Lippman in 1908 [25]. During the last few decades the Integral images have been developed by various researches.

A glass or plastic sheet consisting of numerous small convex lenses is used by Lippman to construct a fly's eye lens sheet [26]. Before World War II the plastic materials were not available, leading Sokolov in 1911 to replace the plastic materials with a Pinhole sheet, which works like the lenses but instead of the diffraction process, it works on the refraction base [27].

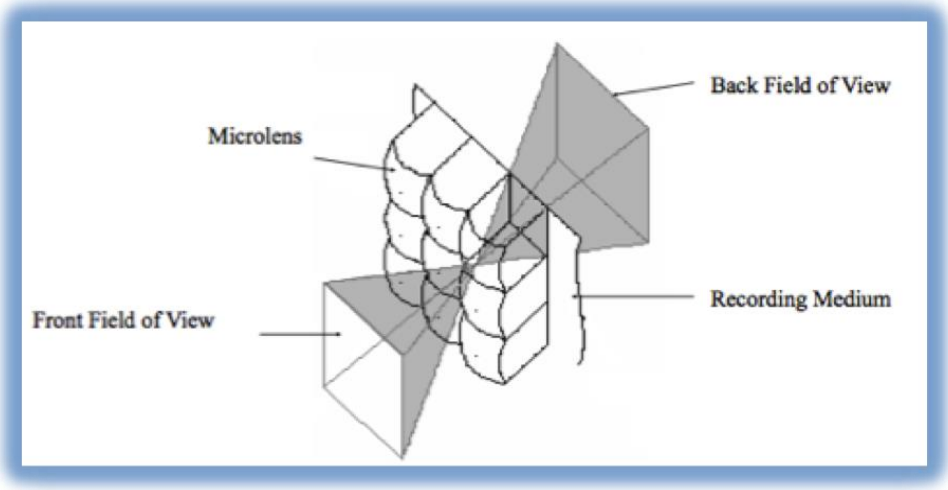


Figure 2.7 Fly's eye, the micro lens array [2].

One of the main drawbacks of Lippman's proposed scheme is the reconstructed images is Pseudoscopic, and in order to over-come this problem, Ives in 1931 [28] developed Lippman's idea by using a two-step integral capturing. A second integral image from the left side is captured, and viewed from the right side, results in the orthoscopic image [29].

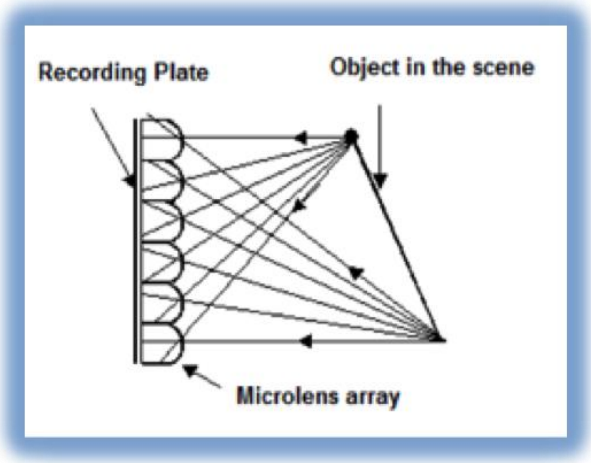


Figure 2.8 The recording of a 3D Holoscopic image [29].

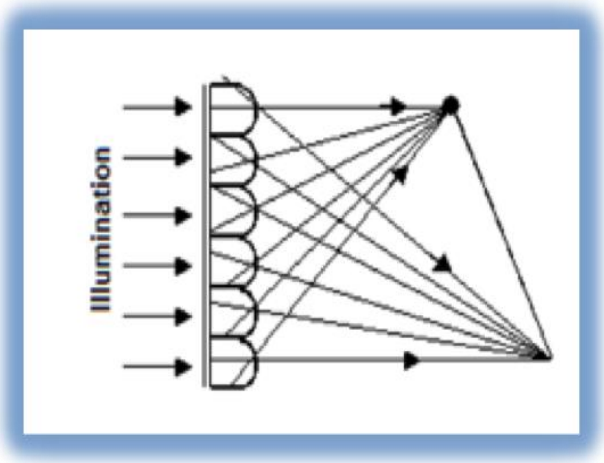


Figure 2.9 The replay of a 3D Holoscopic image [29].

In 1948 Ivanov Akimakiha experimented for the first time, Integral Photography, using a lens sheet, and increased the number of lenses more than lippmann, instead of 10,000, he used 2,000,000 lenslet with diameter 0.3 and focal length 0.5mm. Although it had been known that De Montebello in 1966 [30] had made a number of improvements to the IP, nothing had been documented, however, later in 1978 he cooperated with Ann Arbor and others to manufacture the IP and they named it MDH Products Corporation, achieving excellent picture quality.

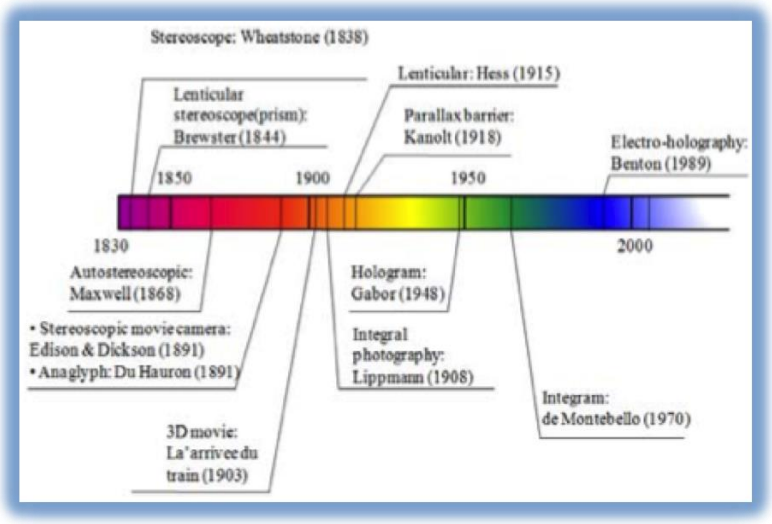


Figure 2.10 History of the 3D Display [24]

As it has been mentioned, the earlier principal of Integral Photography proposed by lippman was suffering from pseudoscopic, meaning the image was inverted in depth as illustrated in Figures 2.8, 2.9. In order to overcome this drawback a two-step method is applied as shown in Figure 2.11 [31][32][2].

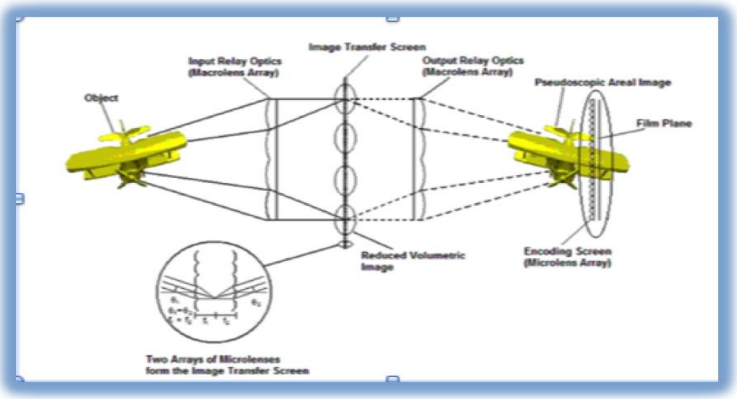


Figure 2.11 3D Holoscopic Imaging Camera [31][32][2].

2.5 3D Camera used in capturing the Full Parallax image

There are two types of cameras that have been constructed as part of the 3D VIVANT project [2]. Type 1 is represented in Figure 2.12 below; it consists of microlens array, relay lens and digital camera sensors. The drawback of the type 1 camera is that it only produces 3D virtual images or 3D real pseudocsopic images because it is unable to control the depth, resulting in the reconstructed image appearing in its original position.

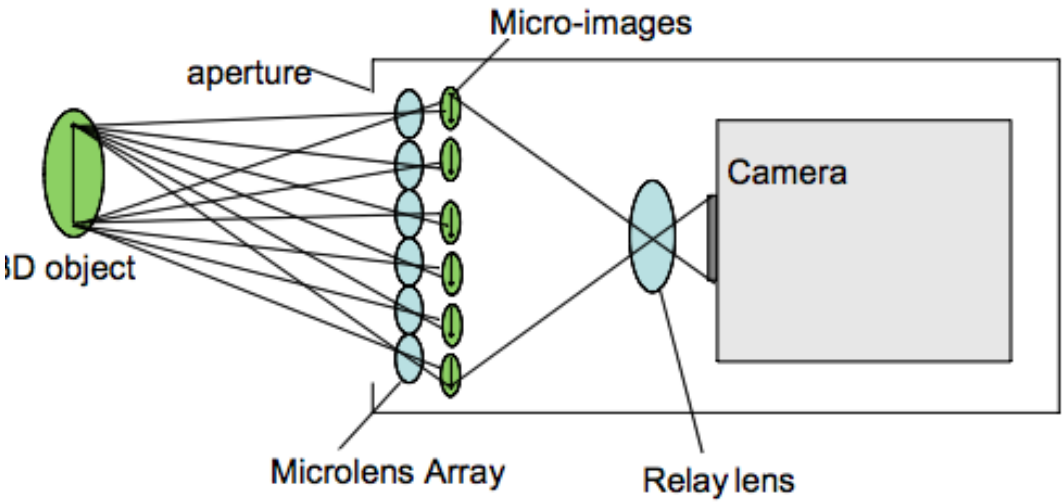
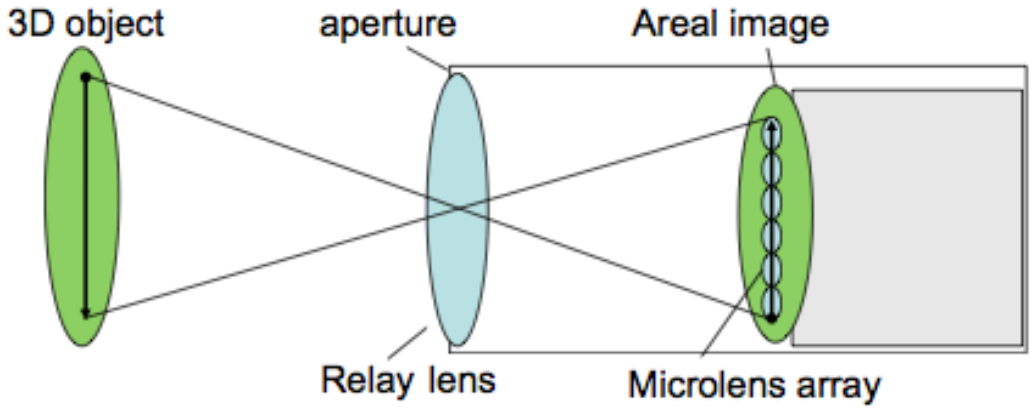


Figure 2.12 Type 1 Camera Outline [2][3].



*Fig*ure 2.13 Type 2 Camera Outline [2][3].

Type 2 Camera was used to generate 3D holoscopic content based on stop-motion. Type 2 and type 1 are similar except type 2 uses the relay lens instead of the field lens and the back focus of the microlens array instead of the sensor, see Figure 2.13, 2.14 and 2.15, so a number of elements have been reversed.

Figure 2.14 Camera-Types 2 [2], [3].

Figure 2.15 Camera-Types 2 [2], [3].

2.6 Image Compression

Exploiting the data redundancy in an effective way can optimize image compression. This process is achieved by reducing the Data redundancy, which is categorized into two types [33]: **Statistical** redundancy and **Psychovisual** Redundancy. The latter is associated with the characteristics of the human visual system (HVS); by eliminating the **Psychovisual** Redundancy (irrelevancy, Perceptual Redundancy [34]) the compression becomes more efficient. The former is divided into two types, the Interpixel redundancy and coding redundancy. The coding redundancy means the information that is represented in the way of codes like Huffman and Arithmetic Coding, while Interpixel redundancy is classified into Spatial and temporal redundancy.

Spatial redundancy implies the prediction of the pixel from its neighbouring and it is the basis of the differential transform. On the other side temporal redundancy represents the correlation between pixels from the sequence of the video frames.

Although spatial redundancy can be employed in image and video compression, Temporal can only be used in association with Video frames.

So we can classify Huffman as a Coding Statistical redundancy. Excluding Psychovisual Redundancy, all other redundancies depend on statistical correlation.

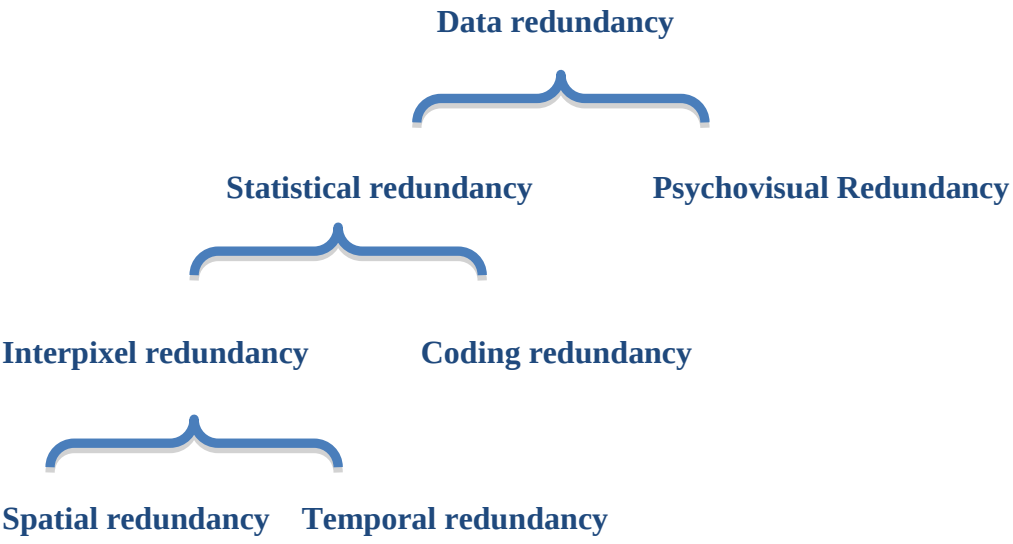


Figure 2.16 Data redundancy classification

2.7 Image Compression Classification

Image compression can fall into two classes: Lossless and Lossy [35]. Lossless compression keeps the image quality [36] the same by reversing the compressed data to the original one by exploiting the statistical redundancy. On the other hand Lossy compression affects the image quality by removing some details that is not noticable to the human's eye by exploiting the Psychovisual [36][37] redundancy that yields to minor error in the reconstructed image or Video data[38].

The JPEG standard includes DCT-based method (Lossy compression) and predictive method (Lossless compression). The DCT-based lossy scheme is followed by a lossless stage as a final step obtained by entropy coding[6][39] for redundancy reduction and it has been proposed for image, video and conferencing system like JPEG,MPEG,and H.261 [35][40] .

2.8 JPEG Standard

For the past few years, a joint ISO/CCITT committee known as JPEG (Joint Photographic Experts Group) has been working to establish the first international compression standard for continuous-tone still images, both gray scale and colour [11][40].

The **JPEG baseline compression system** [13] is by far the most widely used DCT-based system. The input image is partitioned into blocks of 8x8 pixels first and then each block is transformed using the forward DCT. Next, DCT coefficients are normalized using a preset quantization table and, finally, the normalized coefficients are entropy encoded. Figure 3.2 shows the block diagram of FDCT.

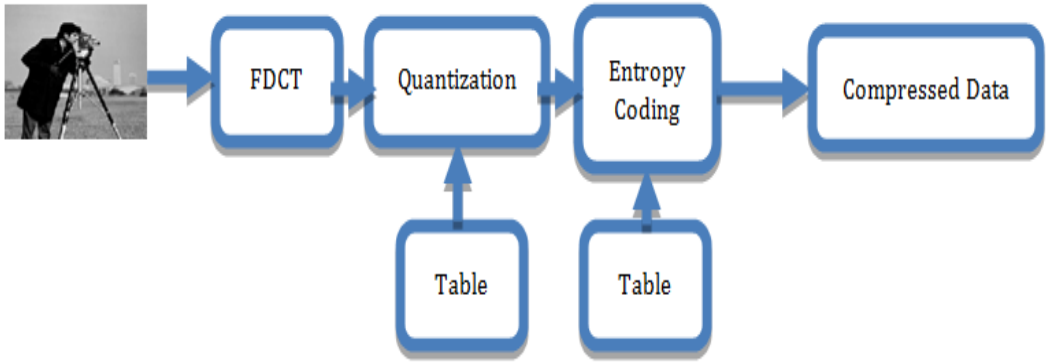


Figure 2.17 The JPEG Block Diagram.

2.8.1 Discrete Cosine Transform Overview

DCT has many advantages, most impressive of which is its capability to reduce the data into few coefficients, which refers to *Energy Compaction*, by squeezing the number of bits that are needed to be quantized [41]. Figure 2.18 shows the 2D- DCT coefficients after applying DCT to the entire Cameraman Image and Figure shows the DCT coefficients after applying 2D-DCT to (8x8) blocks.

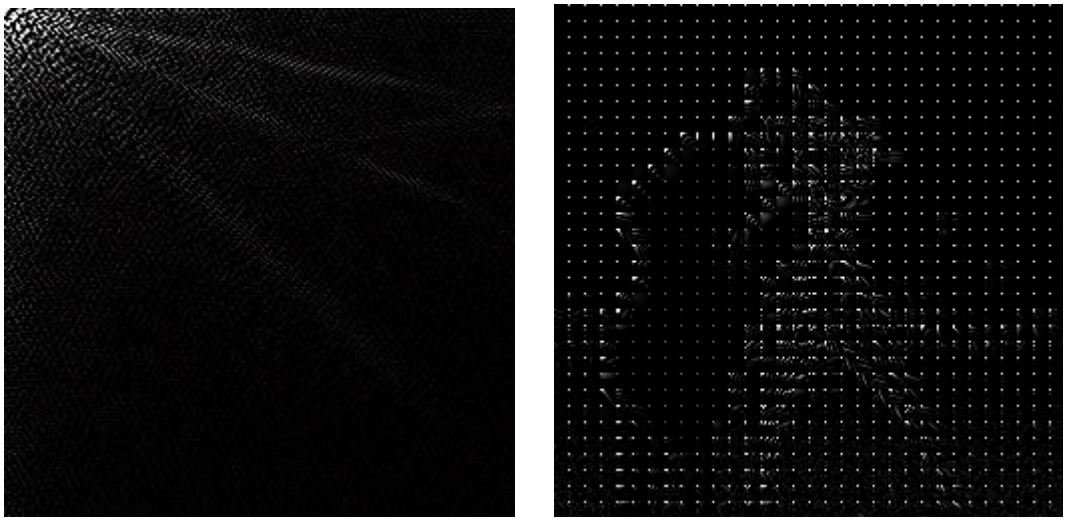


Figure 2.18 The 2D- DCT coefficients

Another advantage of the DCT is *Separability*, shown in figure2.19, which is exploited in this Chapter to reduce the overall computational complexity. So instead of applying the 3D-DCT in one step, it splits it into two steps first, 2D-DCT is applied to rows and columns for 2D-Array; it then applies 1D-DCT as a second step to accomplish the 3D dimension.

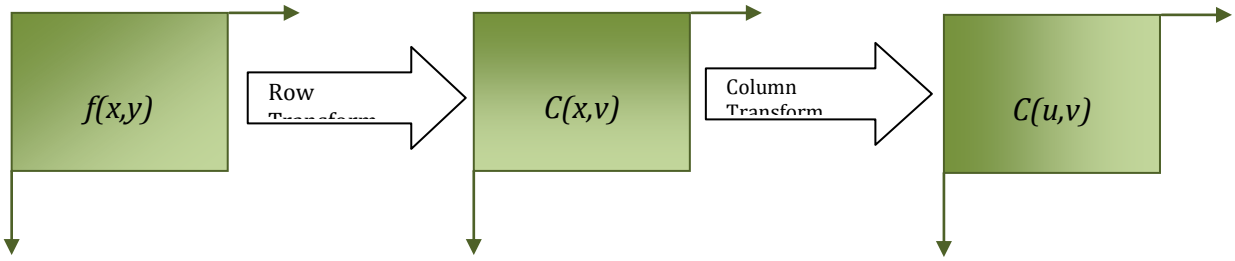


Figure 2.19 Traditional DCT Separability

DCT is also enhanced by *Symmetry* and *Orthogonality*, and both are very effective; the former gives the DCT the ability to be recalculated offline while the latter cooperated with the de-correlation property in reducing the complexity of calculation due to the fact that the inverse coefficients A are equal to its transpose $(1/A) = A_T$ and achieve a high level of flexibility [42].

The de-correlation property in the normal 2D Images, exploits the correlation between the adjacent pixels. The Upper Left-hand corner of the DCT coefficient array contains the Low frequency coefficients which according to the human eye is very imperative and contains the important data while the Lower right-hand corner of the DCT coefficient array lays the high frequency coefficients which are less important to the human eye and aimed to be zeros to achieve high compression ratio.

However, whilst the DCT has a number of advantages, it also has a serious drawback, which, related to the block scheme, has no correlation considered among neighbouring blocks causing the block artifacts. This particular aretefact is studied in Chapter 4 [42].

The following equations are the FDCT mathematical equations [11]:

$$F(u,v) = \frac{2}{N} C(u)C(v) \sum_{y=1}^{N} \sum_{x-1}^{N} f(x,y) \cos[\frac{(2x+1)u\pi}{2N}] \cos[\frac{(2y+1)v\pi}{2N}] \qquad (2.1)$$

where,

$C(u),C(v) = \frac{1}{\sqrt{2}}$ *for* $u,v = 0.$

$C(u),C(v) = 1$ 0therwise.

Individual blocks (NxN) are independently transformed into the frequency domain using the DCT [40]. Each block is transformed by the *forward DCT* (FDCT) into a set of 64 values referred to as *DCT coefficients*. The first value of each block is referred to as the *DC coefficient* and the other 63 as the *AC coefficients* [43]. Each of the 64 coefficients is then quantized using the Quantized Table. For the AC Quantized coefficients, a Zigzag scanning is applied resulting in a one-AC-array dimension followed by applying the Run Length Code. For the DC Quantized coefficients a different code called Differential Pulse Code Modulation is applied, which also results in one-Dc-Array. This process can be considered as a first step in entropy coding.

2.8.2 Differential Pulse Code Modulation

The DPCM is a simple lossless compression technique applied on the DC coefficients, and is ideal is to take the difference between the Dc from the current block and the previous one.

2.8.3 Run length Encode

While the DPCM is applied to the DC coefficients, the RLE is applied on the AC coefficients. It is a very common Compression code, and accounts for a number of long sequences, replacing them with shorter ones [44]. It is the number of repeated zero-valued AC coefficients in the zigzag sequence preceding the nonzero AC coefficient [11] [12].

Each DPCM-coded DC coefficient is represented by a pair of symbols (Size) and (Amplitude) [11]. Each RLE-coded AC coefficient is represented by two symbols. Symbol-1 represents two pieces of information, Run-length and Size. Symbol-2 represents the single piece of information Amplitude. Once the RLE and DPCM are applied the Huffman code is assigned.

2.9 JPEG baseline and JPEG2000

However JPEG2000 was developed after JPEG were tasked to provide a more efficient successor and to avoid the blocking artefacts of the DCT-based compression by using Discrete wavelet transform , however, JPEG baseline is still widely used for storage of images in digital cameras, PCs and on web pages.

Both JPEG and JPEG2000 [37] share features with MPEG-4 Visual and/or H.264 and at the same time as they are intended for compression of still images, the JPEG standards have made some impact on the coding of moving images.

2.10 Entropy Coding

Entropy coding depends on the frequency of occurrence, meaning that short code words are given to the most frequent data occurrence, on contrary; the long code words are given to the least frequently occurring data [45].

JPEG employs two entropy coding methods—arithmetic and Huffman coding [13], in general the challenge for entropy coding was to decrease the statistical redundancy in image or video compression [34], While The H.264/AVC provides two entropy coders: context-based adaptive variable length coder (CAVLC) and context-based adaptive binary arithmetic coder (CABAC) [14], CAVLC as a baseline profile and CABAC as the main profile.

The simple profile or baseline profile is supported by CAVLC and it is used for video-telephony, wireless communications and video-conferencing, while the Main

profile or High profile is supported by **CABAC for** video storage, television broadcasting and also video-conferencing [47].

H.264/AVC develops the system and achieves a better compression ratio than other video standards like MPEG-2, H.263, and MPEG-4 part 2 by saving approximately 50% of bits [10], by applying more complicated techniques [14], that results in the decoder becoming twice as complicated than the decoder in MPEG-4 Visual and also the encoder increased ten times more complicated than MPEG-4 Visual for simple profile and for the main profile the decoder is four times complicated than MPEG-2.

2.10.1 Huffman Coding

As previously mentioned, Huffman is a statistical code that affords the shortest average codeword length, leading it to achieve the best compression among the other statistical codes, in addition, the Huffman code is prefix-free, implying a simpler and easier decoder [45].

This results in a minimum bit rate with which the data source is encoded. However, in practice, the Huffman code suffers from problems.

1) If the pmf for a symbol is greater than 0.5 or close to one the compression efficiency will be poor [37] [48].

2) Transmission error, Spreading or propagation of an error is one of the serious problems that Huffman coding is prone to. An error at any stage of transmission leads the decoder to lose the sequence of the code or get different lengths and this can cause a de-synchronization or mismatching between the receiver and the transmission [37][48].

3) The real time applications is very difficult when using Huffman coding for instance a probability calculations for a very large video sequence requires a long time to obtain the table and that cause a delay.

4) Huffman Coding is working improbably when dealing with different probabilities of symbols that are not occasionally interconnected or complicated [48].

5) Both the encoder and the decoder must have exactly the same information regarding the probability distribution of the input data by sending the customized Huffman table in a way to be sure from the decoding process [35], and this increase the file size by adding extra data to the header and cause inefficient compression ratio [37][49].

Actually, the last disadvantage can be avoided using Adaptive Huffman coding (also called Dynamic Huffman) [36] where we don't need to know the symbol's frequency table in advance, however whilst it is more efficient, it has also got its drawbacks. The problem with adaptive coding is that it is too complex for real time Video applications, meaning, it is a challenge to develop an entropy code from an unknown, unpredictable source and may be also varying in probabilities to a real time Video, and the process will be sophisticated [48]. Additionally, more inefficient compression is added beginning of the data because of the shortage of the statistical information [49], and this problem increases if the segmentation is used to solve the transmission error. By segmenting the data the adaptive Huffman will work inefficiently, and will result in a lower compression ratio.

Another approach is to use the pre-defined table, in H.261 and MPEG, as it was used to optimized the compression and to get better coding efficiency [35], according to the experimental results in [35], customized Huffman tables gives a 2-8% advantage of data compression over the pre-defined ones, and this is because of the need to send the table to the decoder [37] that is the exactly the same one that is used in the encoder process and also the pre-defined table is not related directly to the image that is in use because the pre-defined table is constructed based on a symbol statistics of a large number of sample images, so the pre-defined table could contain statistical data that is not used in this particular image which makes the code gain more (weight) symbols that are not required and this causes a worthless enlarge of file size leading to insufficient compression. In MPEG-2, more than one table is used [35], for each intra/inter code there is a separate Huffman table.

In [13] it illustrates that using the custom table reduces the code by only 1.38% and if an adaptive coding is used this percentage increased by only 1%bearing in mind the high level of computational.

The author in [44] used the same technique as JPEG, MPEG and H.263by having four ac code tables but applying only one to each block but his proposed algorithm differs in the way he categorized the DCT blocks depending on their ac coefficient. He then uses a different code table for each category, with his code dependent on the instantaneous coefficient content inside the block and the accumulated rate buffer content, on the other side [13] code each block using multiple ac code table, and that creates a problem for the decoder, because of the nonzero ac coefficient. So *Lakhani* proposed another algorithm instead of pairing a nonzero ac coefficient with the run-length of the preceding zero coefficients, the encoder pairs it with the run-length of subsequent zeros.

In [45] the author tried an approach that performs more effectively than [46] by showing a new one which is similar or often smaller decomposer hardware overhead to give better compression.

In [49] they tried to achieve a quicker Adaptive technique by dividing the process into two trees the front tree and the back tree. The front tree is an adaptive Huffman code, except the number of leaves (code-words) is a variable, while the back one is a symmetric or almost symmetric tree. Each symbol is encoded only in one tree and moves dynamically between two trees, according to the occurrence frequency. The symbols with higher frequency have more chance to be encoded in the front tree, and those with lower frequency often stay in the back tree. The faster adaptability is made possible by the relatively small size of the front tree, and due to this the data is more efficiently encoded.

In [39] they present an algorithm to generate Huffman coding tables for classes of images conceived using the Huffman optimizer. Applying the new generated tables instead of the standard ones, it is possible to obtain a further reduction in the final size

of the bit-stream without loss of quality. New tables for three classes of images (landscape, portrait and document) were generated taking into account statistical similarities, measured in the "event space".

In [50], they also propose an array of data structure to represent HCT and corresponding decoding algorithm. The proposed data structure overcomes the drawback in the SGHC approach proposed by Hashemian [36]. This adaptation still keeps the key advantages of the SGHC approach that provides memory efficient and fast Huffman decoding.

In [51], they used lossy compression approach for gray scale images. If we remove all the redundancies from the image, we can compress the image efficiently. Here psycho visual redundancy is removed by DWT, and coding redundancy is removed by Huffman coding. Our main objective is to optimize the compression ratio using the proposed algorithm.

2.11 Shifting Scheme

Before starting the whole process the samples should be level shifted to a signed representation by subtracting 2P – 1, where P is the precision parameter [43]. After a non-differential frame decoding process computes the IDCT and produces a block of reconstructed image samples, an inverse level shift shall restore the samples to the unsigned representation by adding 2P – 1 and clamping the results to the range 0 to 2P – 1. So, when P = 8, the level shift is by 128; when P = 12, the level shift is by 2048.

2.12 Measuring the Quality of the Compressed Image

In Image compression, there are many ways to measure the quality of the reconstructed Image [52]; one of them is the Peak to the signal noise ratio (PSNR). The PSNR method is not the best, and there are many other more effective methods, however, we use PSNR as most of the papers used for comparison with the proposed algorithms in the thesis are using it.

The MSE is defined as:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j)\right]^2 \qquad (2.2)$$

The PSNR is defined as:

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \qquad (2.3)$$

Where, *MAX* =255, *representing the pixels using 8 bits per sample.*

i,j: counter
I: original image
K: reconstructed image
m,n: dimensions of the image.

Chapter 3

Novel Entropy Coding Technique for 2D DCT and DWT based Image Compression

In this chapter a new entropy-coding algorithm is introduced in order to improve the quality of image and reduce the computational complexity. The new proposed algorithm will be used instead of the Huffman coding to achieve a high compression ratio and reduce the time. In order to proof the efficiency of the proposed algorithm, the new proposed entropy coding will be implemented on the 2D Images to compare it as like for like with the JPEG baseline Standard, in the first part of this chapter. In the second part of the chapter, the new entropy coding will be implemented with associated with 2D-DWT technique and the results will be compared to EZW, SPHIT and EBCOT. The algorithm will be tested on five common 2D Images. Later in the next chapter it will be implemented on the 3D-Integral Images.

3.1 Novel Entropy Coding Technique for 2D DCT based compression technique

Figure 3.1 represents the block diagram for the proposed algorithm. The input of any entropy coding is AC and DC Coefficients. The pre-processing block refers to the conventional 2D DCT transform as in JPEG encoding (as mention before in chapter 2) excluding the Huffman coding. The outputs are entered to two blocks, one with AC Coefficients as an input and the other block for the DC Coefficients.

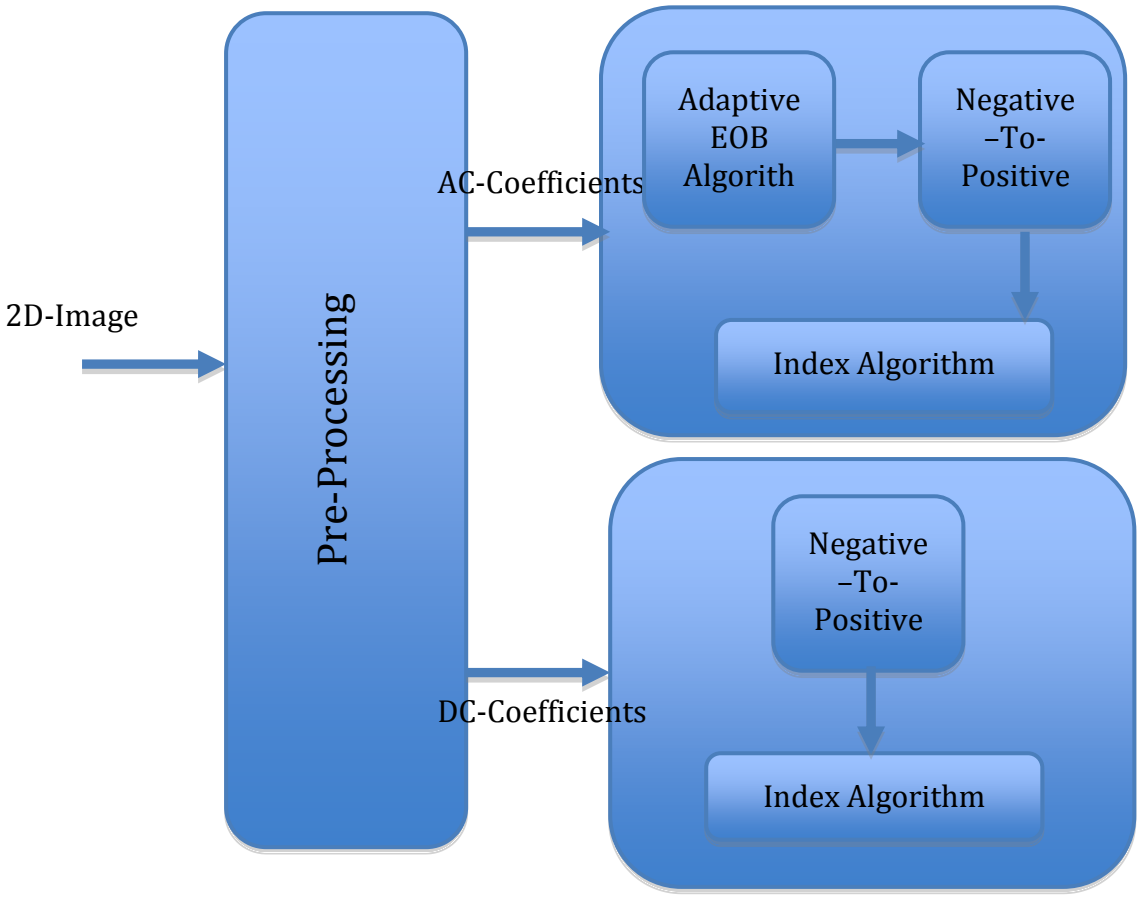


Figure 3.1 The Block Diagram for the Proposed N-To-P Algorithm

3.1.1 Concept

The most important objective of image compression is to minimize the file size before transmitting or storing. Amongst the factors that are deemed to affect the file size the following two have a higher effect on the file size:

1. The content of the file itself and how it is defined, be it small integers, large integers, double or float.

2. The sign of the variables inside the file.

By exploiting this fact, and to overcome this problem in this chapter a new algorithm is developed and applied on 2D images.

3.1.2 Proposed Negative-To-Positive Algorithm (N-To-P)

The proposed Entropy Coding Technique consists of three parts. Each part is considered, in order to increase effectively the compression ratio. The algorithm is applied to the Amplitude Part in Huffman Algorithm. Normally in JPEG base line after applying the quantization factor the result is DC and AC Values.

3.1.2.1 DC-Coefficients

For DC vector, the DPCM is applied to get smaller values so the second factor, which is the type of integer, is avoided.

The output of the DPCM includes negative values coming from subtracting each value from the values preceding it. Since this 1-D array is not necessary to be descending or ascending, it ends up with both signed and unsigned normalized coefficients. Hence this will result in an increase in the file size. To reduce the file size, all the negative coefficients are extracted and for each Negative Value, its position is saved in an index array.

This Negative array is multiplied by (-1) to get rid of the signs then return each Value to its position in the original array while keeping the index array as a header file. So in this level, two arrays for the DC Values are defined; (i) the original DC 1D-array without any negative values and (ii) the index array which contains the positions of those negative values.

3.1.2.2 AC-Coefficients

For the AC array, a proposed adaptive algorithm is represented and will be illustrated in the following section.

3.1.2.2.1 Adaptive EOB

In the case of the *AC-Coefficients part* in the Negative-to-positive algorithm is the **A**daptive **E**nd **o**f **B**lock. Normally in JPEG after quantizing the coefficient array, the RLE is used to count the zeros that are among the values and the EOB is used to define the last non-zero coefficient.

Normally JPEG is using "999" as an indicator of the End of Block. During the Huffman process the code "999" is treated differently than the other values. The Run/Category is not applied and even the amplitude is not equivalent to its binary value. Once, the program stops at "999" it encoded as "1010".

In the adaptive EOB algorithm, the "999" is not going to be used as an EOB indicator because it is going to be a large number among the AC values and repeating it will end up with increasing the size of the file. Another issue is that 999 could be a value among the AC values that will cause an error. So a new method will be used as a flag.

The algorithm is based on getting the maximum value of the AC 1–D array and increment it by 1 and this will be the EOB value for this array. It will be shown that by using the adaptive EOB algorithm with the Negative to Positive algorithm without Huffman the file size will drop dramatically comparing to the JPEG baseline algorithm. In addition, this process is reversible without any loss. The process of applying the Adaptive EOB to AC Array, followed by applying the Negative to

Positive algorithm to both AC and DC Arrays, will result into two Positive Arrays and two Index arrays; for the AC and DC coefficients, respectively.

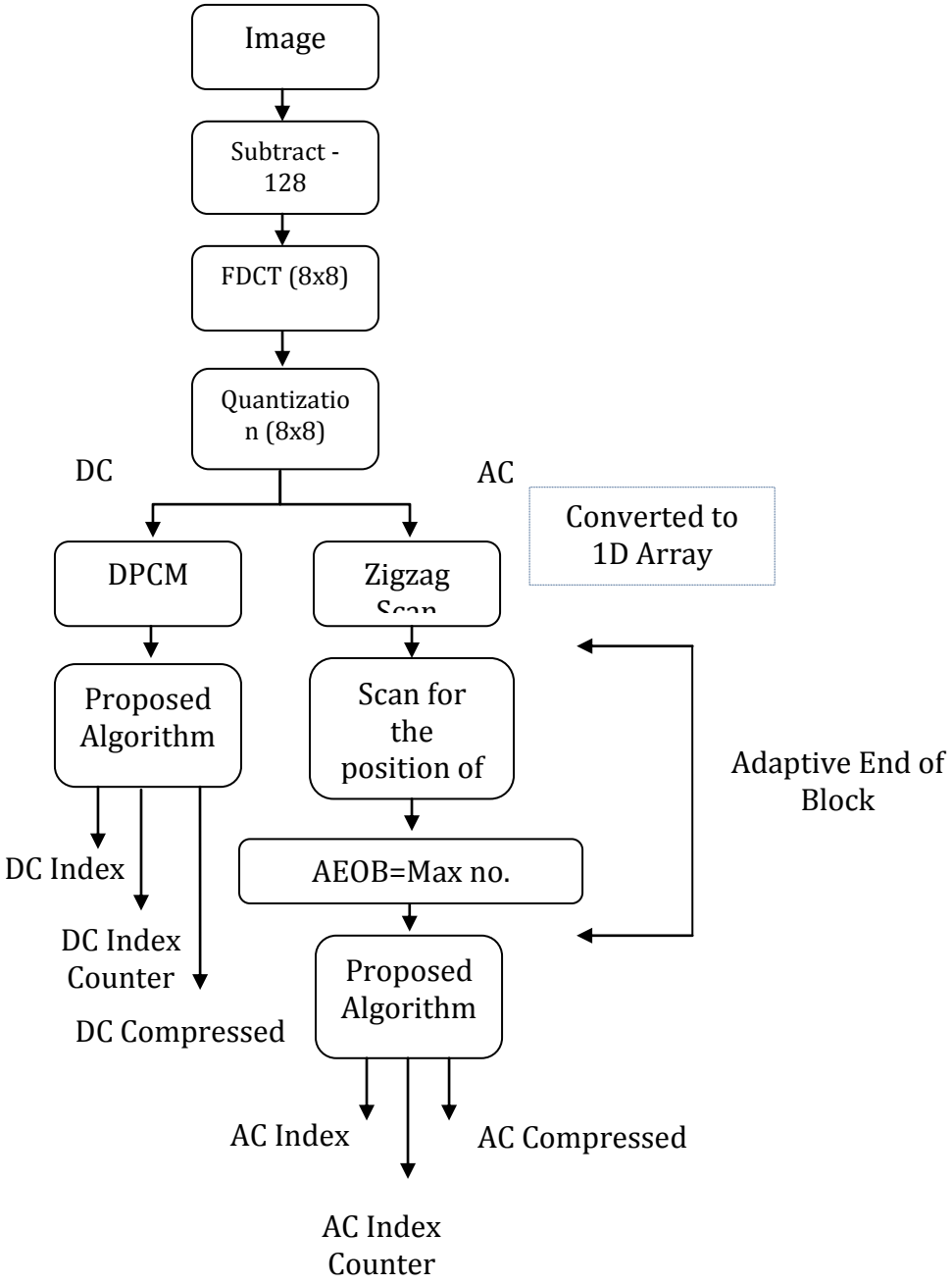


Figure 3.2 Block Diagram for the proposed N-To-P algorithm without using Index Algorithm

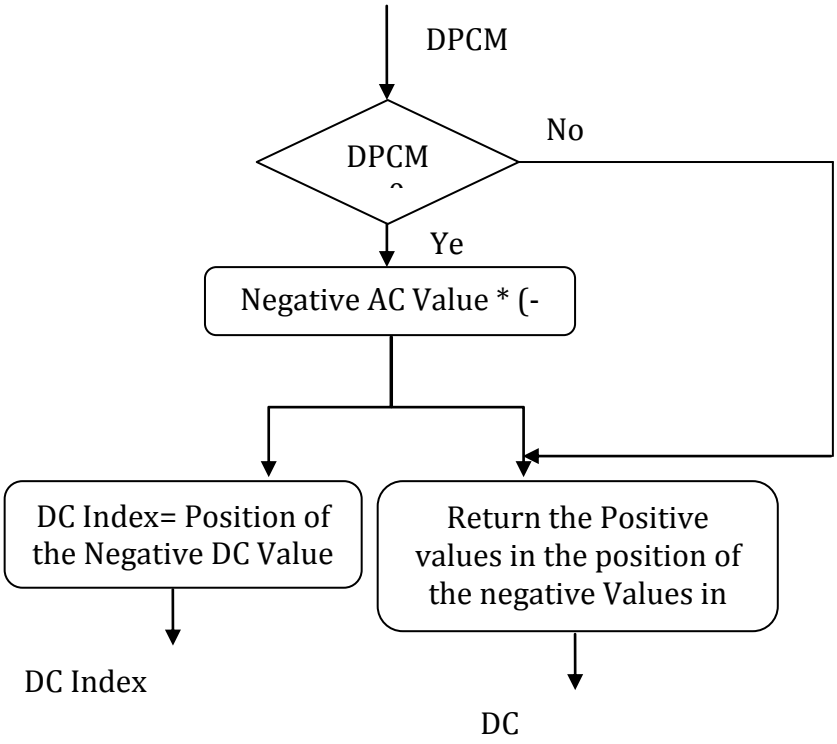


Figure 3.3 Proposed N-To-P Algorithm Flowchart DC Coefficients.

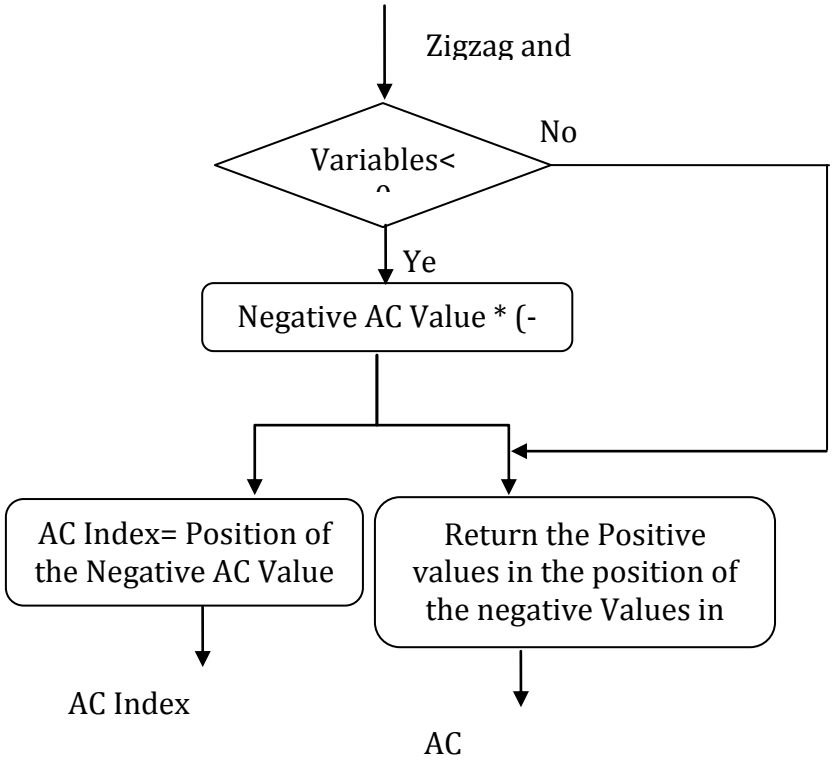


Figure 3.4 Proposed N-To-P Algorithm Flowchart AC Coefficients.

At this level the two positive arrays are quit low in size comparing to JPEG, but on the other hand the two indexed arrays are very big files as shown in table 3.1 (Lena Image), that cancels the privilege that have been gained by using the Negative to positive algorithm.

Q	4	8	16	24
AC Positive Array	**23.4kB**	**15.9kB**	**9.31kB**	**6.93kB**
DC Positive Array	**1.46kB**	**1kB**	**902Bytes**	**838Bytes**
AC&DC Index Arrays	36kB	24kB	15kB	11kB

Table 3.1 File sizes for Lena Image Without applying Index Algorithm.

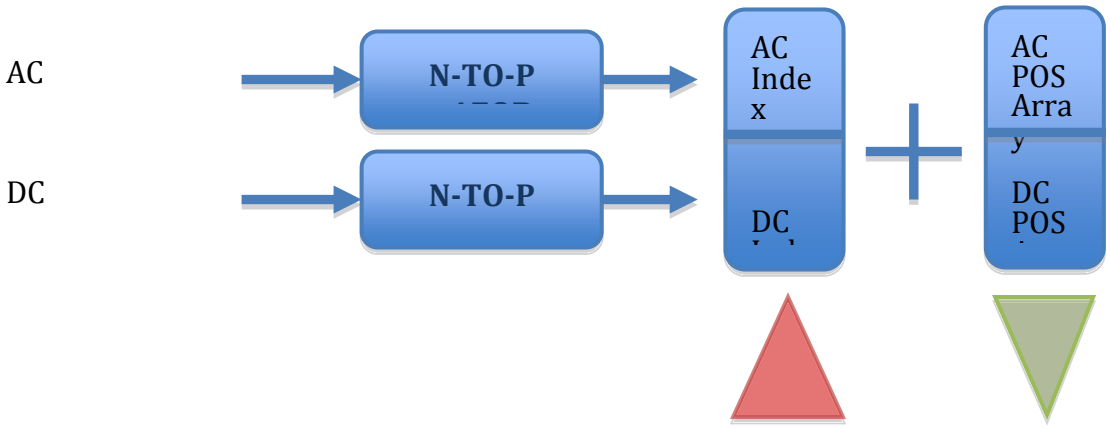


Figure 3.5 Forward Proposed Algorithm Block Diagram

In the following section new coding techniques to reduce the index array are investigated and compared to find the best solutions in terms the file size.

3.1.3 Inverse Processed Algorithm

Figure 4.6 shows the inverse process of the Negative-To-Positive algorithm for both the DC and the AC files. There is two inverse blocks. The inputs for each one of them are an Index Array and Compressed file for AC and DC respectively. The process is lossless. After applying the Inverse Positive-To-Negative algorithm for the AC Files, an inverse Adaptive-End-Of-Block is assigned.

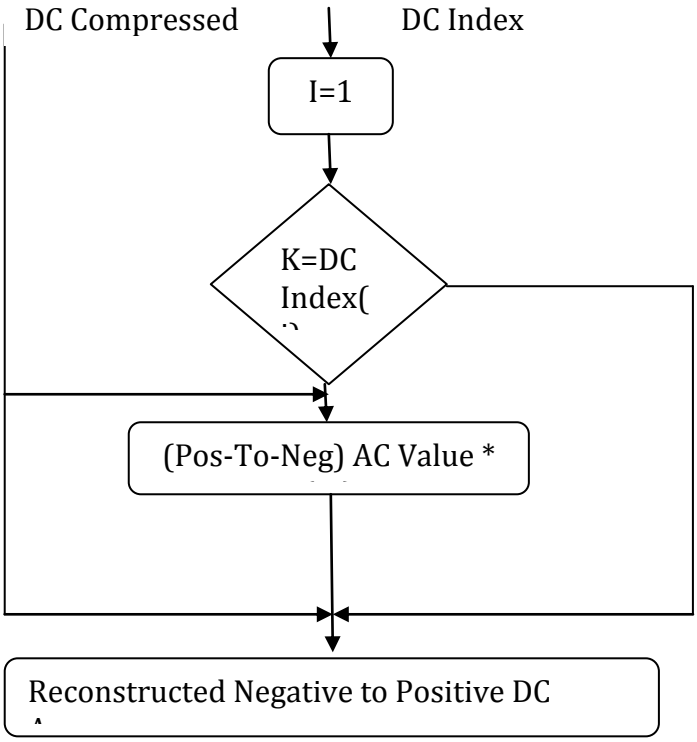


Figure 3.6 The Flowchart for the Inverse Positive to Negative Algorithm for the DC Array

3.1.3.1 Inverse Adaptive-End-Of-Block

As shown in figure 3.7, the process of the Inverse AEOB, as mentioned, is reversible, and it contains three steps:

1. Scanning the Inverse Negative-To-Positive array to get the maximum value.
2. Subtract 1 from the maximum value.
3. And finally, apply the inverse zigzag.

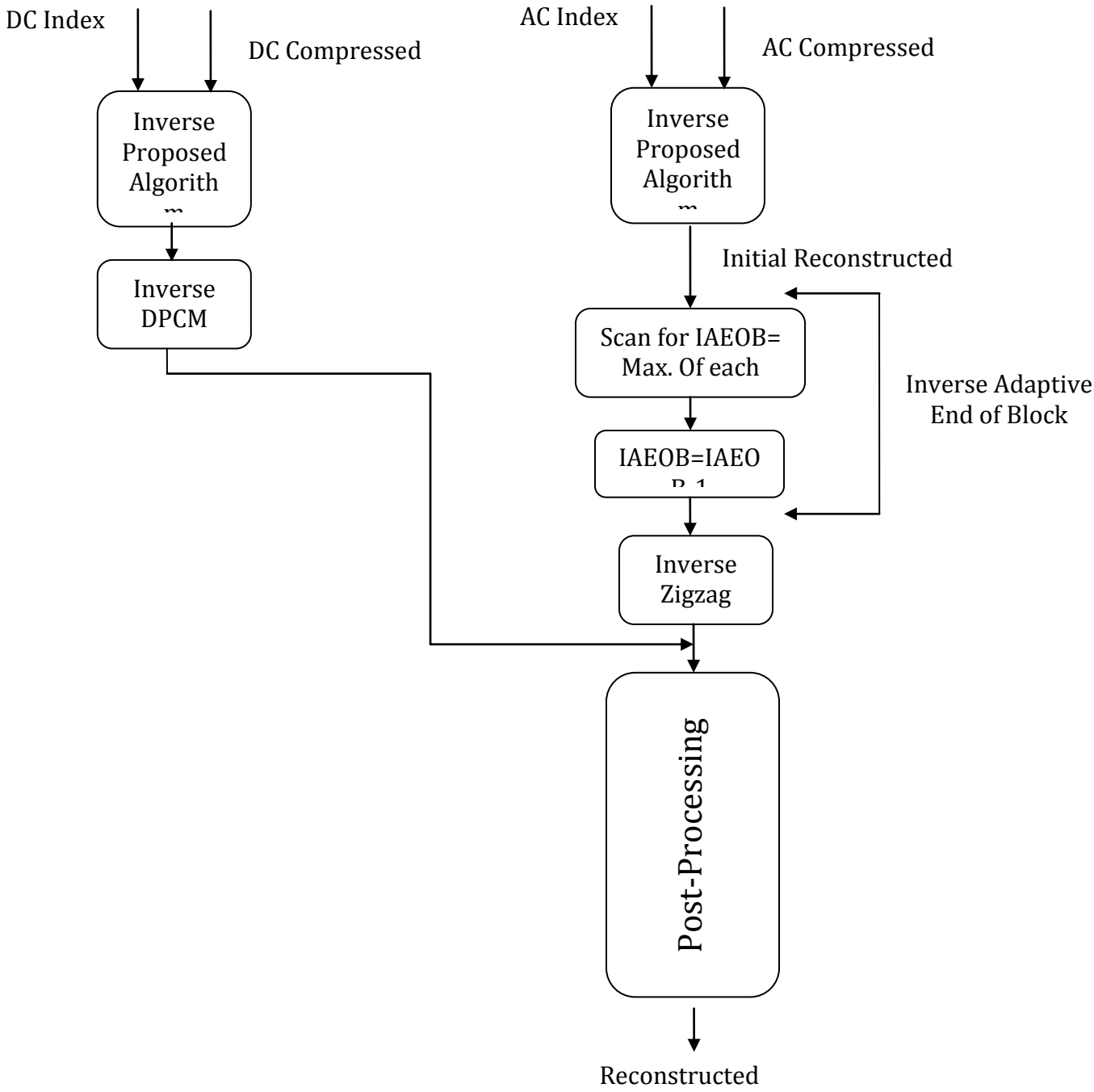


Figure 3.7 Inverse Proposed Algorithm Block Diagram

3.1.4 Proposed Index Algorithms

The index array has a high file size but it has some special characteristics:

1. The Values are ascending.
2. It starts with very small numbers and ends up with very huge number and that is enlarges the file size.

Several algorithms have been implemented in order to control the enlarging of the Indexed file size. Three of these algorithms are discussed in Appendix A. In this section two other algorithms are discussed in details.

3.1.4.1 Index Cumulative Algorithm

In this algorithm, the fact that the index is increasing in an ascending order is exploited. A flowchart of the algorithm is shown in figure 3.8.

Step 1: there are two main values are defined. (a) The Base Array (BA (i)) which is takes the values of the negative AC/DC array and (b) the Incremental Value (IV) which is used as a threshold. Test on several images showed that IV=20 gives the best results in term of bitrates vs. PSNR criteria.

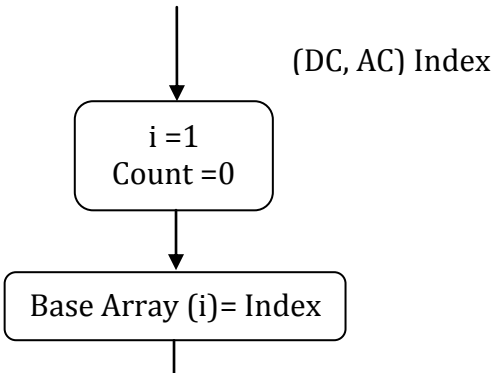
Step 2: an interval is allocated, the minimum limit is equal to BA (1) and the maximum limit is equal to the sum of BA (1) and the Incremented Values.

Step 3: an index scanning occurs, if the index variable is in the interval then the Subtracted Value is subtracted from the index value, but if it not between the interval limits then both the interval limits and the Subtracted Value increased by the Incremented Value and a counter is counting the number of times the Incremented Value increased and put it in an array. This Counter array is important in the decoder process to know when the incremented value should be used.

This algorithm reduce the both encoder and decoder time and it is a simple algorithm especially in the decoder process, and it minimize the huge size of the Index array.

At this point there are three arrays, one for the Negative Positive array and the other

two for the index (new Index and the counter array). The size of all of the three arrays is less than the input normalized Coefficient array, and if we take those arrays as an inputs to the to the Huffman coding using DHT, it will ends up with results either the input arrays will be less in size than the output arrays or the outputs of Huffman will be roughly the same size. That means that using Huffman doesn't affect the results with the Negative Positive algorithm.



Yes

Figure 3.8 The Cumulative Index Algorithm Flow Chart

3.1.4.2 Difference Index Algorithm

Another Index algorithm has been applied in an attempt to compress more the size of the Index array, and as mention before that the index array got two important features that is very useful in creating a new algorithm, the first one is that it is ascending and it is positive. In Difference Index algorithm, the idea of DPCM has been adopted in order to compress the file.

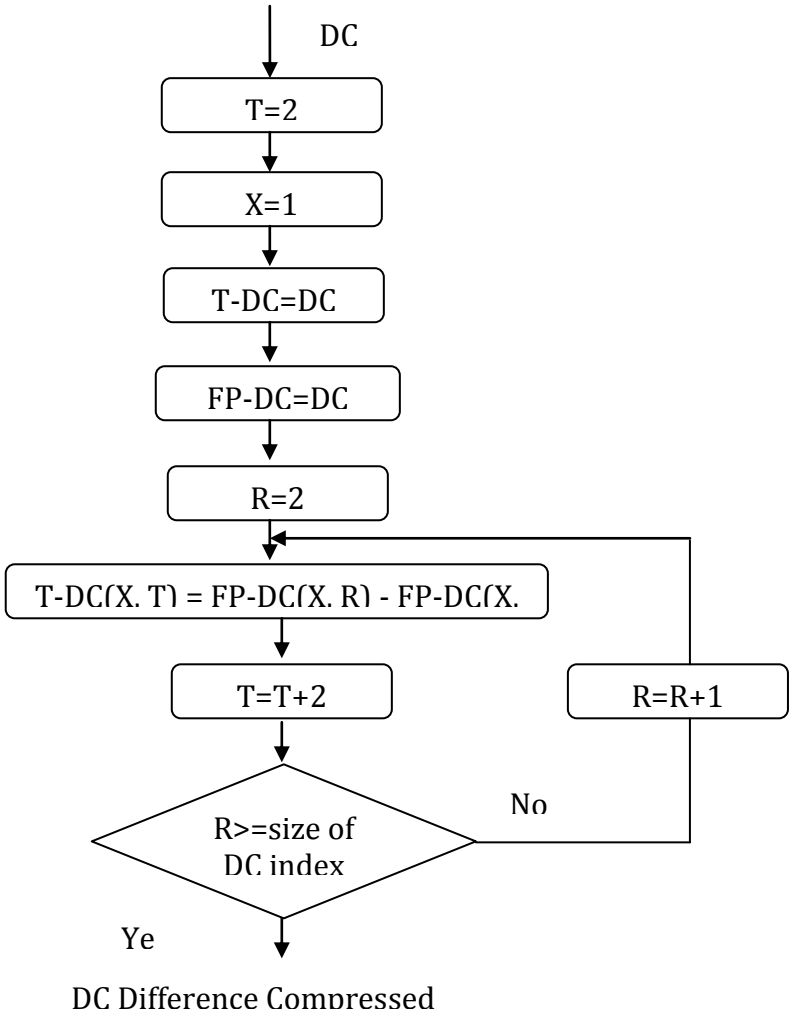


Figure 3.9 The DC Difference Index Array Flow Chart

The AC Index Array is treated in the same way as the DC Difference Index algorithm.

3.1.4.2.1 The Inverse Difference Index Algorithm

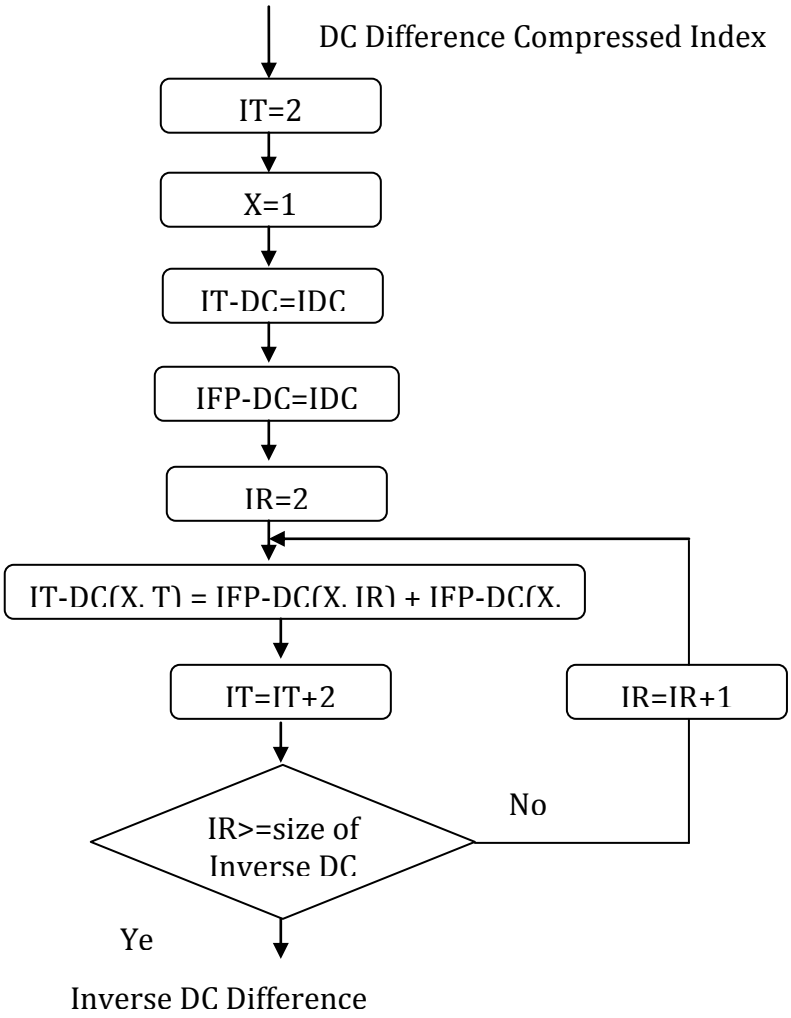


Figure 3.10 Inverse Difference index algorithm Flow Chart

The Difference Index algorithm shows its simplicity and also gives better results than the other proposed algorithms and achieves the goal of compressing the indexed arrays.

Figure 3.11 and figure 3.12 show the comparison between the Cumulative Index

algorithm and the Difference Index algorithm, for 2D and 3D images respectively. It can be observed that the Difference index algorithm is performed better than the pervious proposed Cumulative algorithm.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure 3.11 Comparison between Cumulative and Difference Index Algorithms for 2D Images

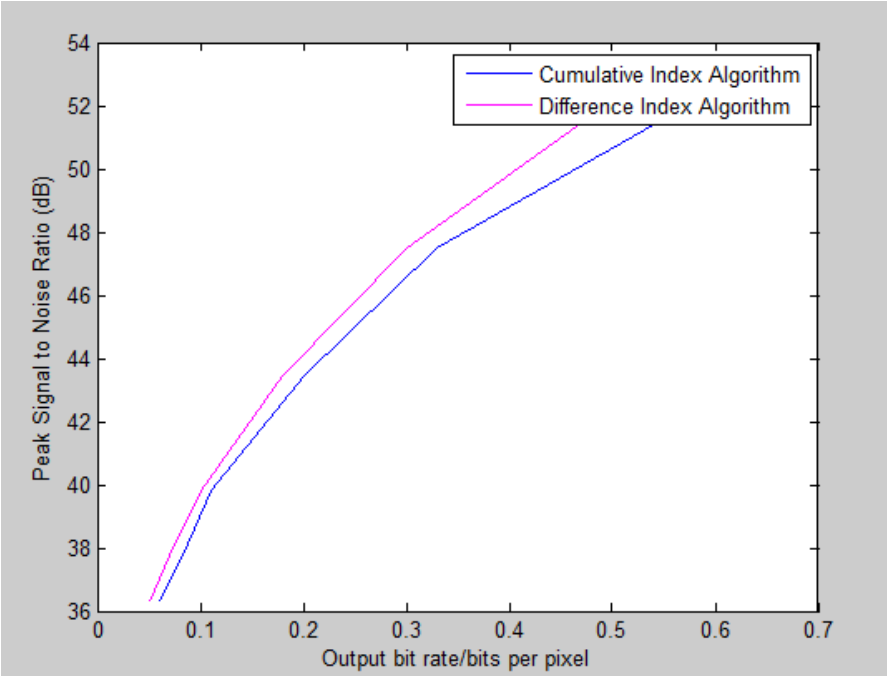


Figure 3.12 Comparison between Cumulative and Difference Index Algorithms for 3D Images

3.1.5 Results

From Table 3.2 and Table 3.3, it is shown that the size is decreased and there is a significant less in calculations resulting to significant time reduction with retaining the Image quality with same PSNR. Table 3.2 shows the average PSNR in decibel for all of the five the images (Lena, Cameraman, Barbara, Baboon and Peppers) in corresponding to the bit per pixel and also the time in seconds, for both the JPEG baseline system using Huffman Entropy Coding and the new proposed algorithm using the Negative-To-Positive Algorithm. So for each Quantisation factor the average results have been calculated for the five images.

	JPEG		Proposed Algorithm	
PSNR (db)	Bitrates (Bits/Pixel)	Time (sec)	Bitrates (Bits/Pixel)	Time (sec)
52.145	0.82134bpp	307.02	0.81bpp	29.034
46.8568	0.6266bpp	250.3	0.5282bpp	24.3
41.817	0.46213bpp	212.407	0.382bpp	19.106
36.89	0.33032bpp	132.312	0.278bpp	12.448
34.055	0.26694bpp	91.24	0.2bpp	7.936

Table 3.2 JPEG & N-To-P Average results

In Table 3.3 the results have been taken from another side that is for each image the average results had been calculated separately. So from table 3.3 it can be shown the effect of the new proposed algorithm on the different nature of images, for example the Baboon is a low redundant image with lots of details, which is normally difficult to compress. From this table, the percentage of the difference between the two technique; JPEG and N-To-P algorithm can be calculated for example the cameraman

is compressed better by 1.25% than the JPEG, Lena is better by 1.26%, Barbara better by1.249%, Peppers 1.17% and finally Baboon by 1.25%.

		JPEG		Proposed Algorithm	
Image	PSNR (db)	Bitrates (Bit/Pixel)	Time (sec)	Bitrates (Bit/Pixel)	Time (sec)
Cameraman (256x256)	40.995	0.4206	25.9	0.3345	2.7
Lena (256x256)	40.83518	0.4364	27.4	0.346	2.128
Barbara (256x256)	39.167	0.67954	51.9	0.543	4.118
Peppers (384x512)	43.068	0.21928	60.98	0.186	6.02
Mandrill (480x496)	41.122	0.7497	1008.01	0.6	67.45

Table 3.3 JPEG & N-TO-P Average results

3.1.5.1 File Sizes Factor

In this section, the results of using the whole proposed algorithm is shown compared to the JPEG standard using with the AEOB for the AC array followed by the Negative-To-Positive algorithm for both AC and DC, and finally the Difference Index algorithm is applied to all the following results. Figure 3.13 shows the difference between the performance of the JPEG and the N-TO-P algorithm via representing the quality of the image using PSNR in decibels versus the bitrates per pixel.

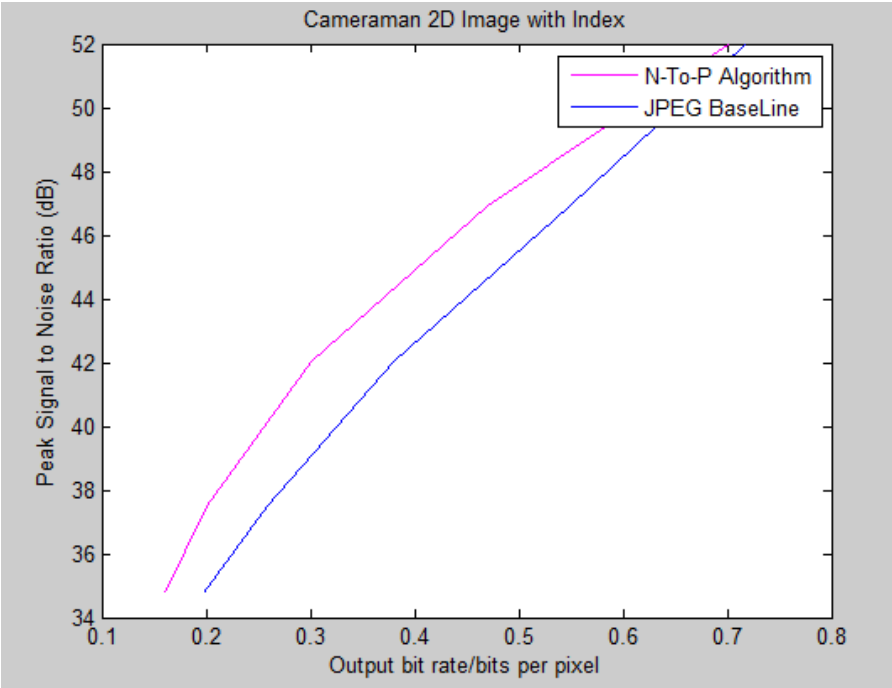


Figure 3.13 Performance of the proposed N-TO-P Algorithm, and Baseline JPEG for compression of 2D cameraman Image.

Figure 3.14 shows the difference between the performance of the JPEG and the N-To-P algorithm on Barbara image via representing the quality of the image using PSNR in decibels versus the bitrates per pixel.

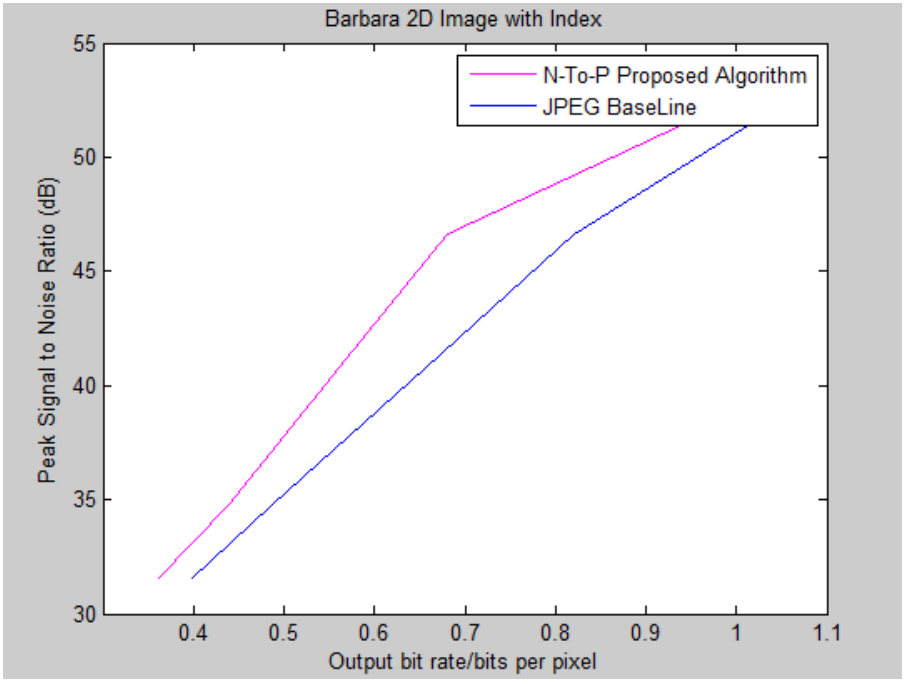


Figure 3.14 Performance of the proposed N-TO-P Algorithm, and Baseline JPEG for compression of 2D Barbara Image.

The same principle is applied on Lena and Baboon image, as will in figures 3.15 and 3.16. In 3.17 a comparison occurs on the baboon image between the N-to P Algorithm without using and Index algorithm and with using the Difference Index Algorithm, in order to show the effect of the index array size on the whole compression. Finally, figure 3.18 shows the average results for the five Images.

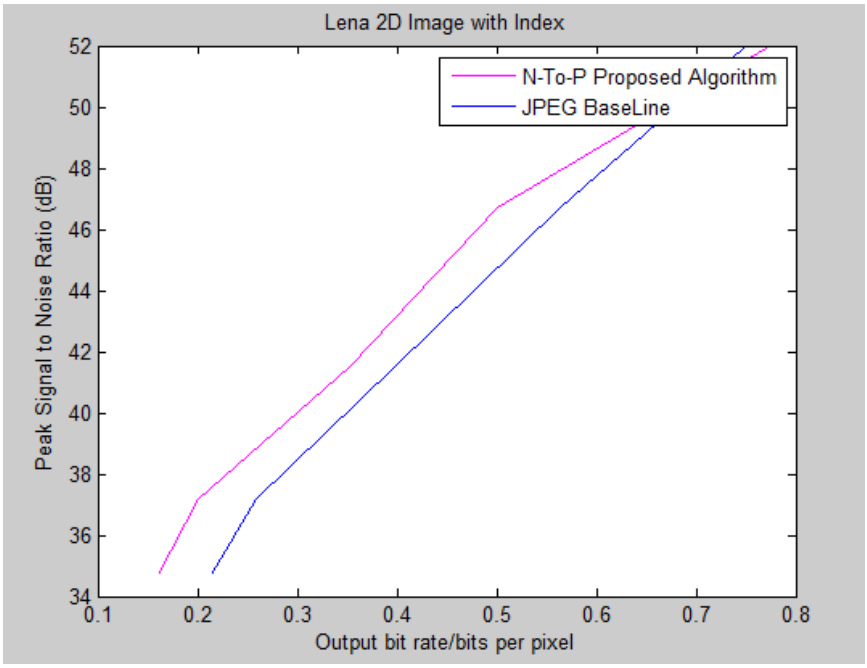


Figure 3.15 Performance of the proposed N-TO-P Algorithm, and Baseline JPEG for compression of 2D Lena Image.

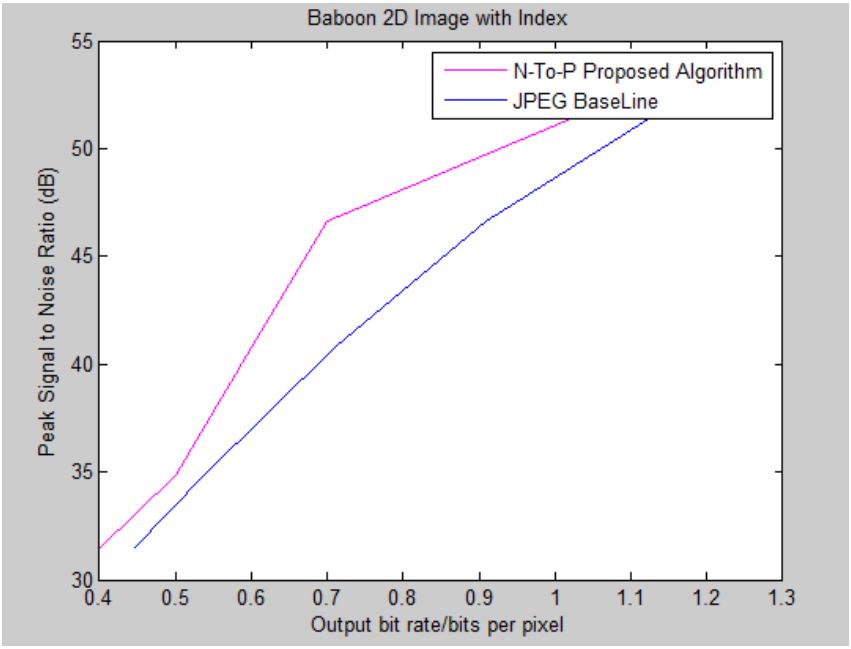


Figure 3.16 Performance of the proposed N-TO-P Algorithm, and Baseline JPEG for

compression of 2D Baboon Image.

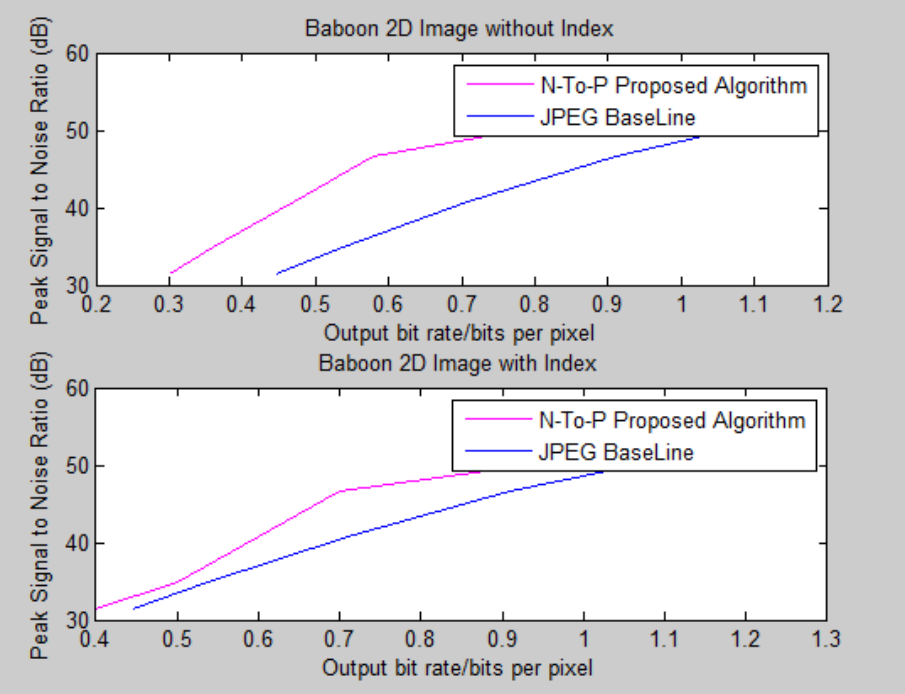


Figure 3.17 Performance of the proposed N-TO-P Algorithm, and Baseline JPEG for compression of 2D Baboon Image with and without Index arrays.

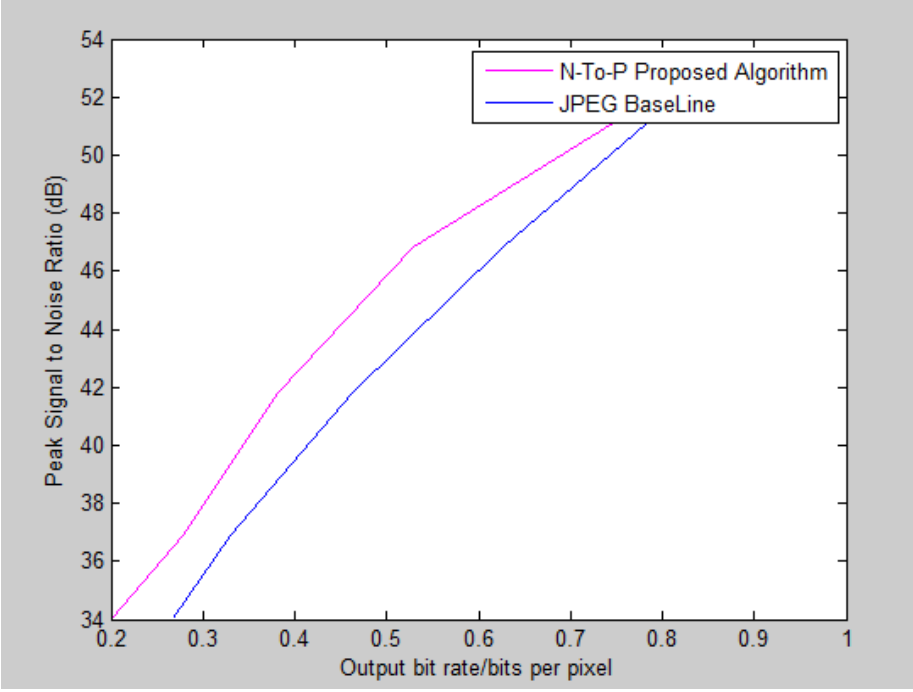


Figure 3.18 Average Performance of the proposed N-To-P Algorithm, and Baseline JPEG for compression of five different 2D Images.

3.1.5.2 Compression Ratio and Computational Complexity Factor

Table 3.4 represents the comparison between the compression ratio of the main file and the compression ratio of the coded JPEG in order to show how much data saved in the proposed algorithm and how it exploit the data redundancy efficiently.

Image	Compression Ratio	
	Proposed	Standard
Cameraman (256x256)	10.528:1	4.55164:1
Lena (256x256)	10.0634:1	4.56348:1
Woman (256x256)	5.91998:1	3.5796
Baboon (480x496)	5.75708:1	1.97342:1
Peppers (384x512)	21.52024:1	8.516:1

Table 3.4 Average Compression Ratios for the Proposed N-To-P and JPEG Baseline.

In figures 3.19, 3.20, the high compression ratio and the reduction in time is shown, respectively. It can be noticed from figure 3.20 two facts; *first* the huge difference in time between the proposed algorithm and the JPEG baseline, *second*, the time in

JPEG is increasing dramatically, however, on the proposed algorithm it is almost fixed according to the PSNR criteria.

A time investigation has been done by calculating the time for each part in the entropy coding process, the experiment was applied on JPEG standard using cameraman with quantization factor =2. The process is divided into six parts as shown in table 3.5 and for each part the time was calculated in order to know which part is taking most of the time.

Time slot	Category	Process	Time (sec)
Time1	DC	DPCM	0.0388
Time 2	DC	Huffman Encoder	0.2867
Time 3	AC	(Zigzag, RLE)+ Huffman Encoder	0.2147 239.512
Time 4	AC	Huffman Decoder	116.3792
Time 5	DC	Huffman Decoder + Inverse DPCM	2.5344
Time 6	AC + DC	Reconstructed Image	0.0760

Table 3.5 Time calculations for Cameraman image with Q= 2 (JPEG standard).

It has been shown that the Huffman encoder process (Time 3) is taking most of the time, and this is occurs due the large number of symbol that is needed to be scanned from the Huffman table, and this is one of the drawbacks that Huffman suffers from, and it is the main reason of the difficulty of using the Huffman in the Video applications. Also some applications needed to use a standard with minimum lose, in that case using Huffman will increase the quality of image by reducing the loss but this will be on the account of the Time [53].

From figure 3.20, it can be shown the significant gain of time reduction using the N-

TO-P algorithm, which gives the N-TO-P the advantages to be used in the video applications that need a fast compression technique, and also with the applications that need to keep the quality of the image as high as possible with the minimum loss, the N-TO-P will do that without coming over the complexity of the system.

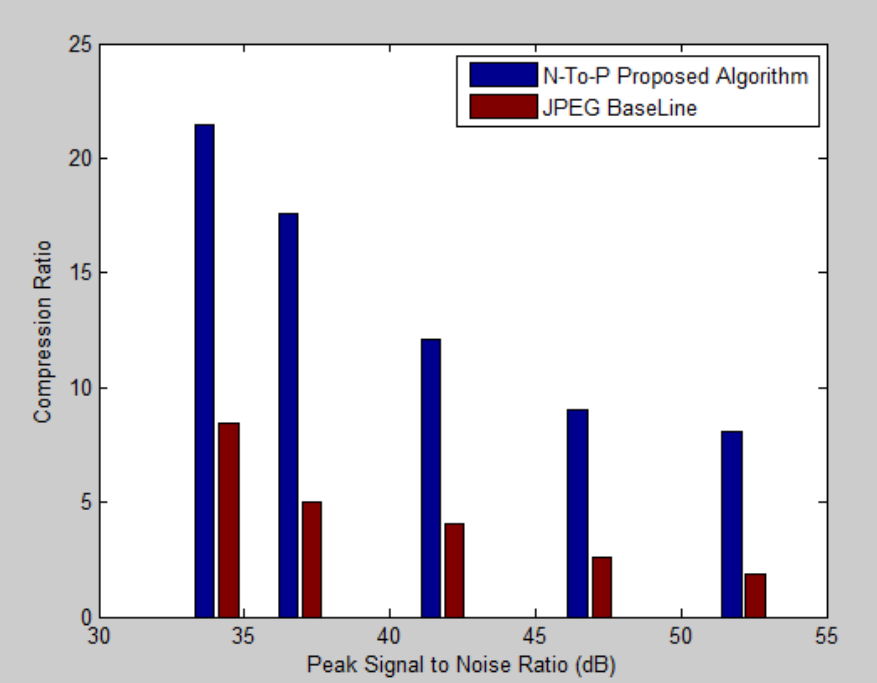


Figure 3.19 Compression ratios vs. PSNR for the N-To-P and JPEG baseline.

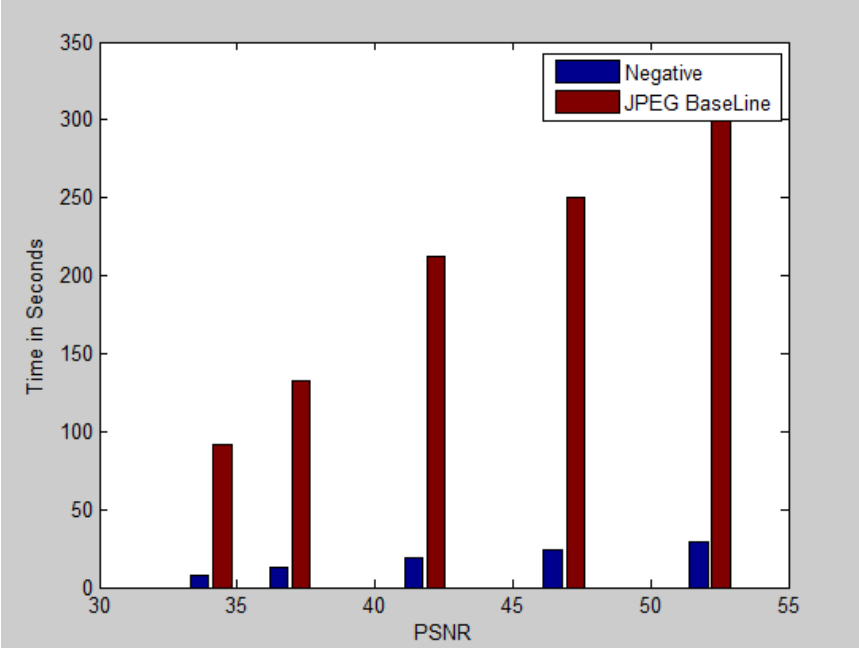


Figure 3.20 Time vs. PSNR for the N-To-P and JPEG baseline.

3.1.5.3 Effect of Quantization Methods

The variance of Quantization factor is shown in figure 3.21 and figure 3.22. In figure 3.21 a fixed Quantization factor of 24 is applied showing the results for both the original cameraman image and the reconstructed one after applying the proposed algorithm, Table 3.6 illustrates the differences.

Image	N-TO-P Algorithm Q Table			N-TO-P Algorithm Q=40 for Barbara Q=32 for Cameraman		
	PSNR (db)	File Size (kB)	Bitrates (Bits/Pixel)	PSNR (db)	File Size (kB)	Bitrates (Bits/Pixel)
Barbara	**27.4636**	**20**	**0.305**	**27.8905**	**16.7**	**0.25**
Cameraman	**31.6390**	**10.1**	**0.154**	**32.8092**	**10**	**0.152**

Table 3.6 Q table and Q Factor effect.

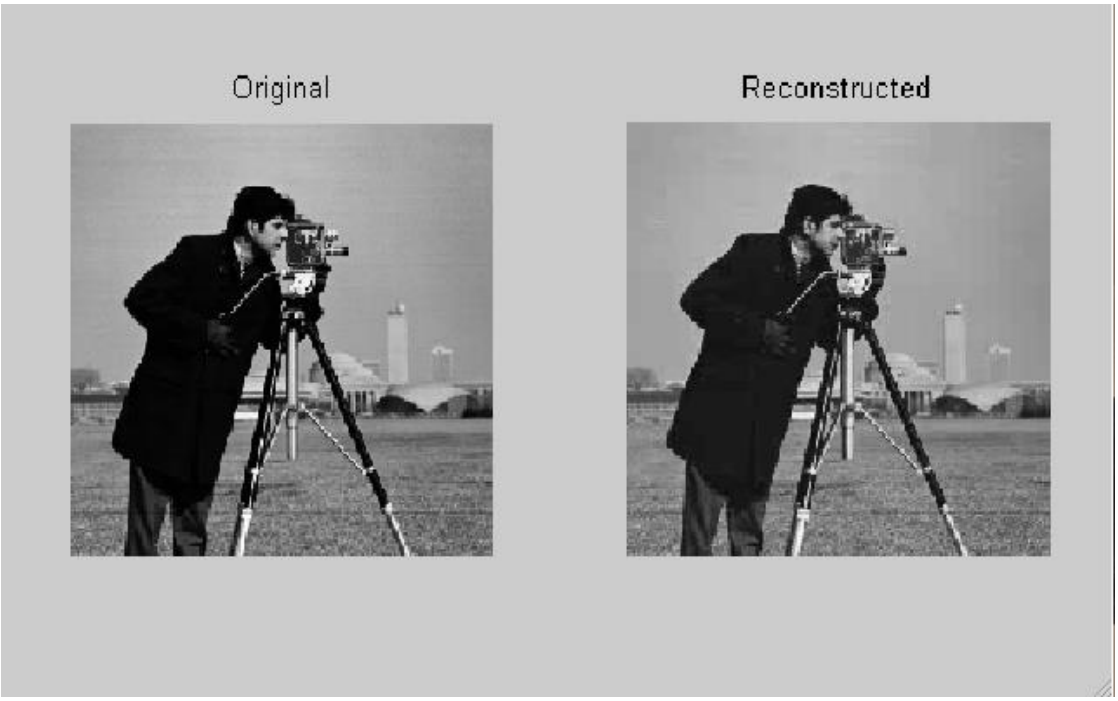


Figure 3.21 Cameraman Q=24 with bitrates=0.168bpp

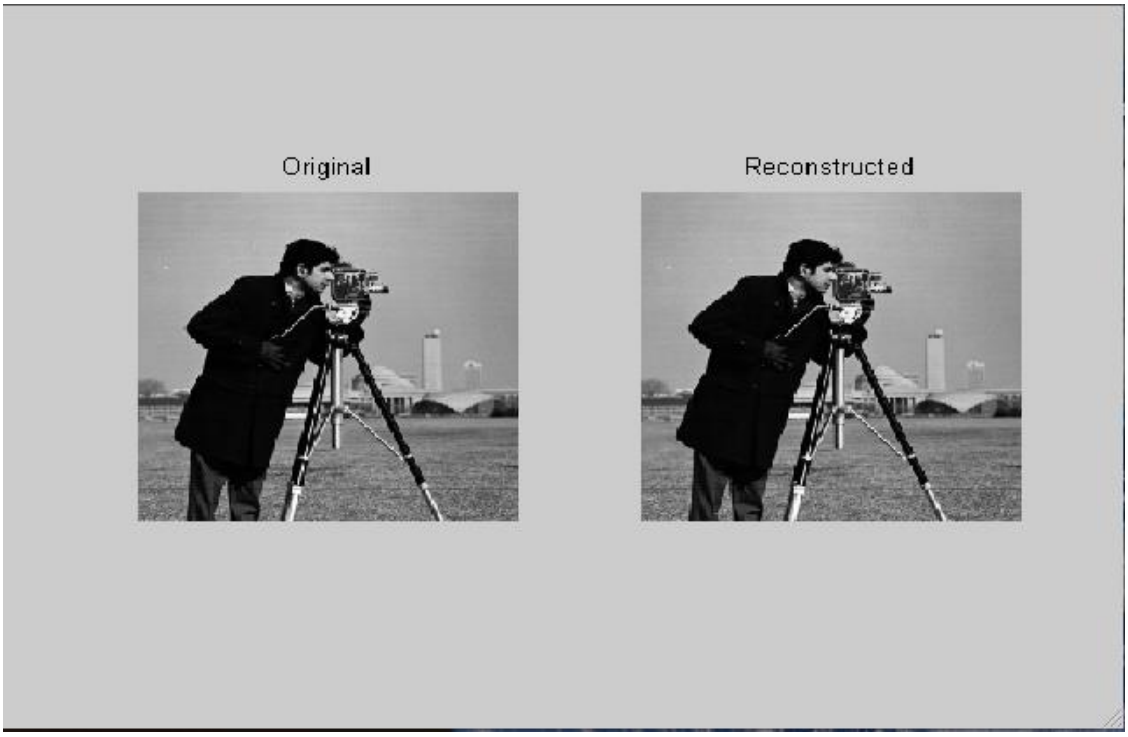


Figure 3.22 Cameraman Q=2 with bitrates=0.73bpp

While in Figure 3.23 and Figure 3.24 another different comparison occurs on Barbara image between a Quantization table (non-uniform Quantization) and a fixed

Quantization factor. By applying Table or Fixed factor on our algorithm that end up that both works efficiently with our proposed algorithm whatever Quantization method is used the results will be better than the original standard.

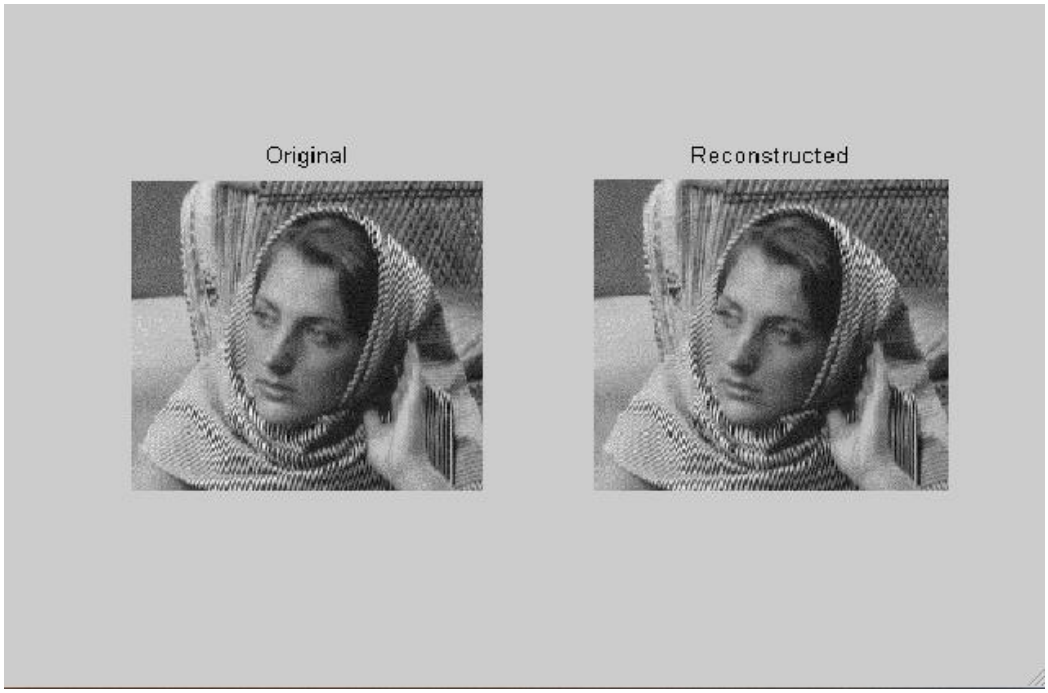


Figure 3.23 Woman Image (Barbara) Using QF=4with bitrates=0.68bpp

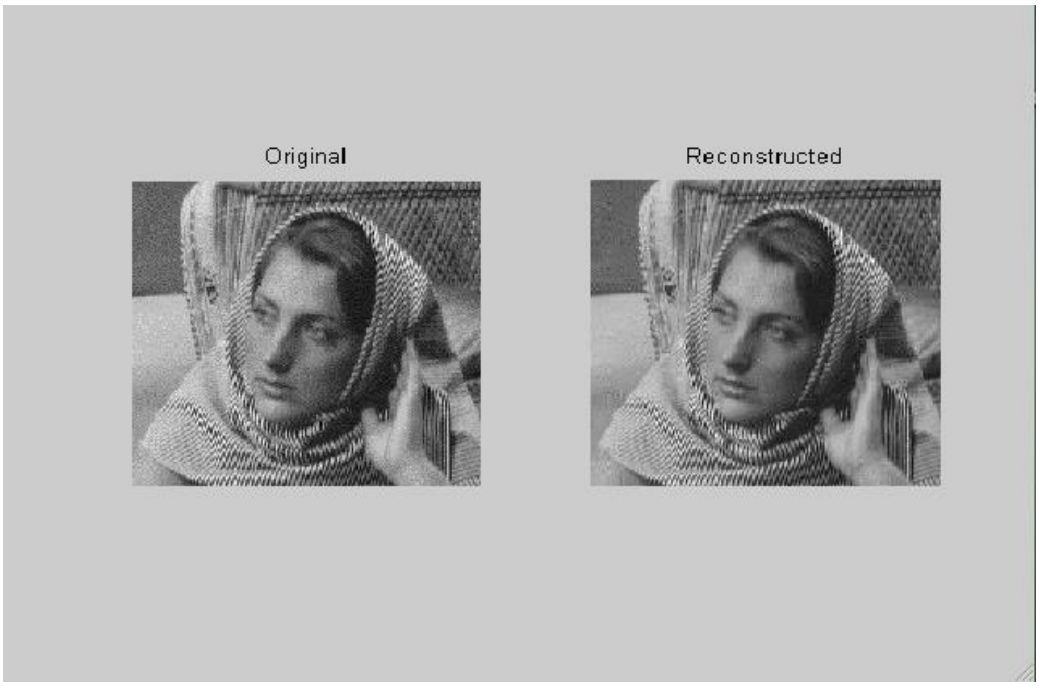


Figure 3.24 Barbara Image Using Q Table with bitrates=0.2334bpp

Figure 3.25 and Figure 3.26 shows both proposed algorithm and JPEG algorithm using same Quantization factor of 8.

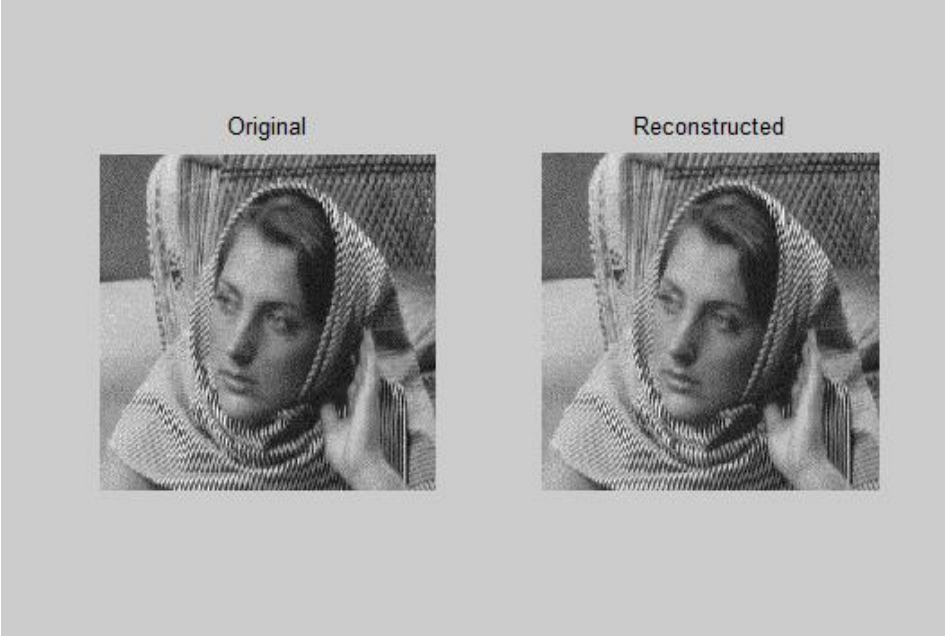


Figure 3.25 Barbara using N-TO-P algorithm Q=8 with bitrates=0.565bpp

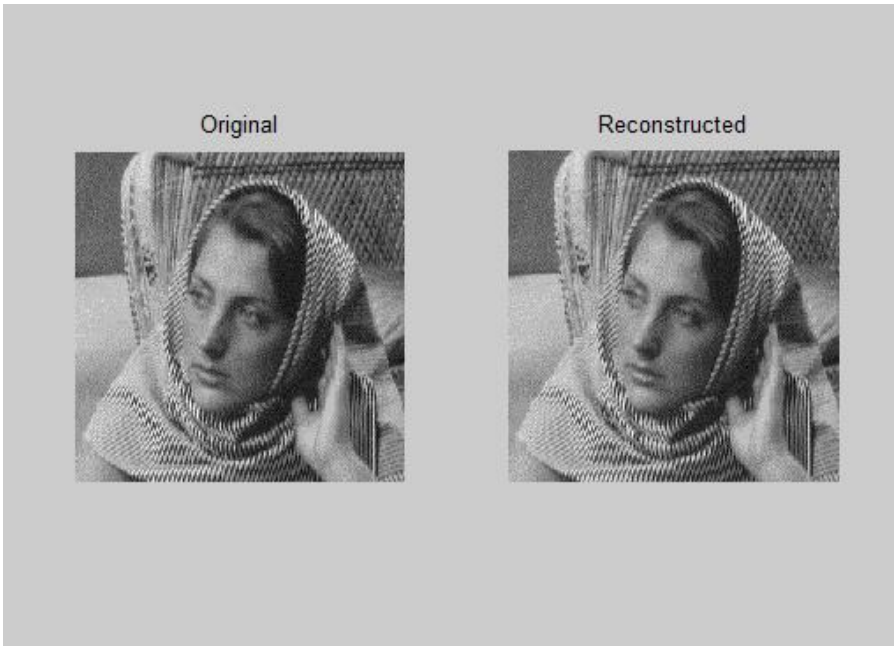


Figure 3.26 Barbara Image using JPEG with Q=8with bitrates=0.656bpp

Another example of the baboon Image is given in Figure 3.27 and Figure 3.28.

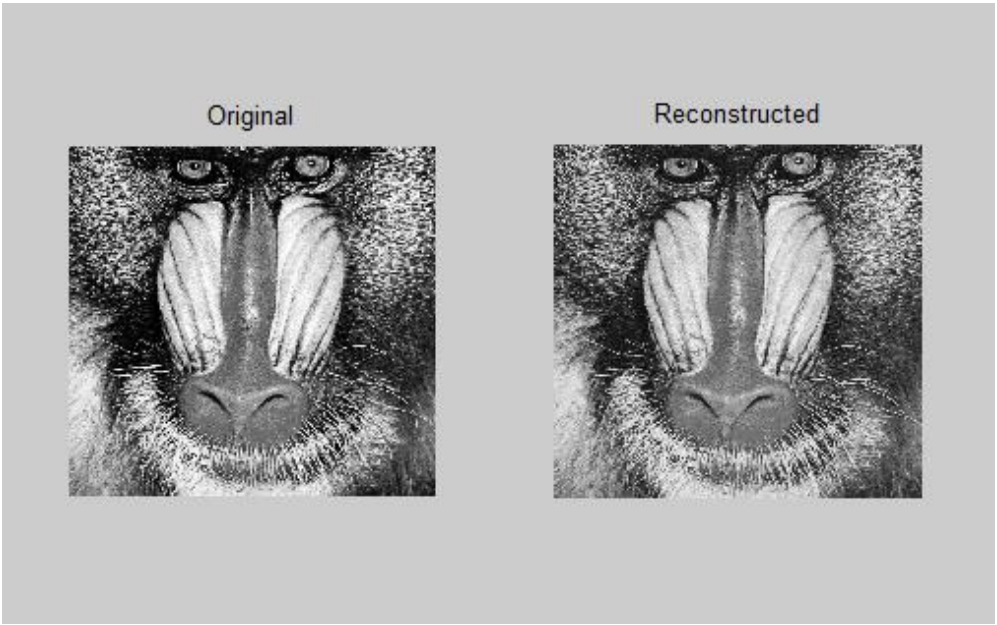


Figure 3.27 Baboon Image using N-TO-P Algorithm with Q=24 with bitrates=0.427bpp

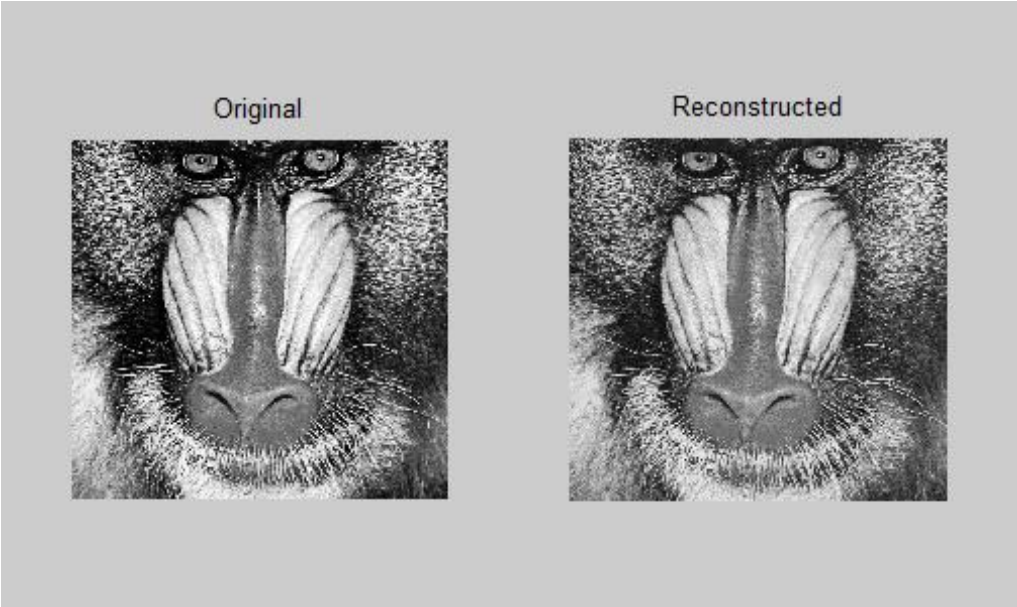


Figure 3.28 Baboon Image using JPEG with Q=24 with bitrates=0.537bpp

3.2 Novel Entropy Coding Technique for 2D DWT based compression technique

In this Section, a new proposed algorithm will be implemented by using the previous Negative to Positive algorithm, that already been tested with DCT in the first half of this chapter, but in this section, it will be tested in association with the 2D-DWT. The results will be compared to different techniques like; EZW, SPHIT and also EPCOT techniques. A detail review of DWT is described in the Appendix B.

3.2.1 2D-DWT Proposed Algorithm Flow Chart

A three levels of DWT is applied using Daubiches 1 (Haar), each sub-band is quantized by a fixed scale quantize factor, and then the encoding part is implemented. A pre-processing process is applied by shifting the values [54] [55][50][56].

A number of wavelets are applied but the best one was Haar (db1) and that is due to its simplicity, fastness, and memory efficiency and also it is not like the other wavelets that suffer from the edge effect [57].

The negative to positive algorithm is applied to all sub-bands, in all levels. This process ends up with two outputs to each sub-band, and in this case there are 10 sub-bands with 20 compressed files divided into two categories as illustrated in figure 3.29. The first outputs are the positive arrays and the second are the index arrays, a Difference index algorithm is applied to all index arrays in order to achieve an optimum compression ratio. The Advantages of the proposed algorithm, is that it is easy to implement, so no computational complexity beside an efficient compression ratio with high ratio distortion is achieved.

3.2.2 2D-DWT Block Diagram

The block diagram that is represented in figure 3.29 is a simple algorithm that is based on applying three levels of 2D-DWT, the outputs from each level is encoded using a forward N-to-P algorithm (FNPA), followed by the index N-To-P Algorithm (INPA) that has been explained earlier in chapter 3.

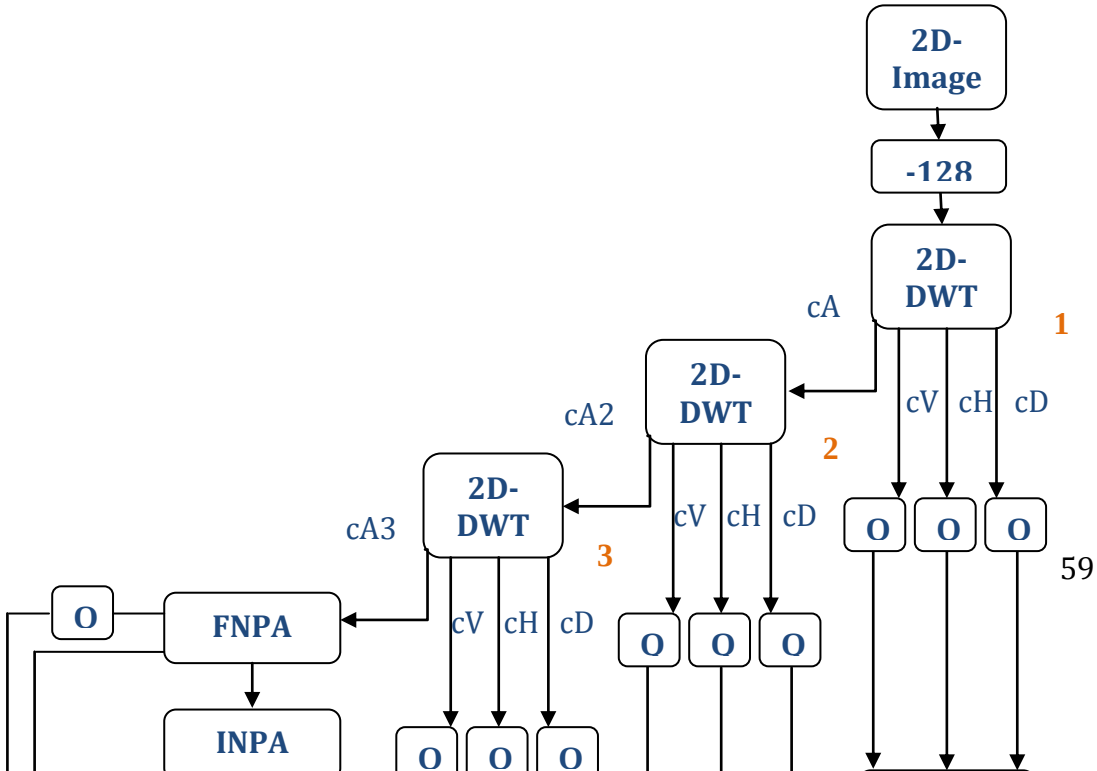


Figure 3.29 Forward Three Level 3D-DWT Algorithm.

3.2.3 Different Techniques

Comparison between the 2D-DWT proposed algorithm and three of different techniques, EZW, SPHIT and EBCOT will be represented in this section. For each technique the results will be shown from the aspect of time, quality and compression ratio.

3.2.3.1 EZW

The EZW is based on the concept of parents and children, in order to get high coding rate [38][58][59]. The first main step is to determine a Threshold which coefficients are compared to it; the output from this process will be divided to significant or insignificant coefficient. If the Threshold was greater than the determined data, the magnitude input data will be categorized as a significant coefficient and vice-verse. This significant coefficient will be categorized to POS if it is positive or NEG if is negative value, figure 3.30 illustrate it.

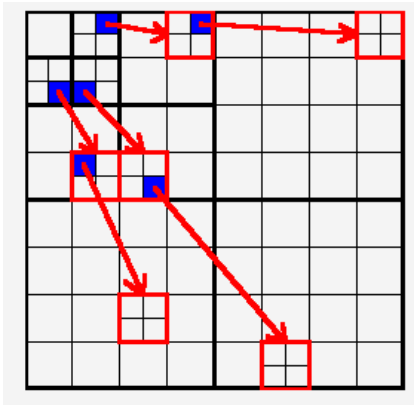


Figure 3.30 Relation between wavelet coefficients in different sub-bands as quad-trees

In case of there is an insignificant coefficient but with significant descendent it will be categorized as Isolated Zero, if it doesn't have significant descendent it will be categorized as zero-tree root. The Isolated Zero is coded separately, and the other zero-tree with all of its insignificant children will be represented to one symbol.
The process is keep repeated until all coefficients become lower than the threshold, even the process is repeated for the insignificant coefficient but with lower Threshold then all transmitted with the four categorized [58][60][61][62][63].

3.2.3.1.1 Results

Table 3.7 shows the results for both the EZW [62] and the proposed algorithm applied on the Cameraman Image from the criteria of PSNR verses Bitrates and Table 3.8 represent it from the PSNR verses Time criteria. The results that shown are after applying three level DWT for both algorithms.

The results can be more analysed from figure 3.32, as it shows the huge difference between both algorithms, and how much the proposed algorithm achieve a Low compression Bitrates and High Image Quality, also the time have been reduced dramatically as shown in figure 3.33. This two results leads to less computational complexity with high compression ratio.

EZW [62]	Proposed Algorithm

PSNR(db)	Bitrates	PSNR(db)	Bitrates
35.55	1.43	69.5081	1.02
30.75	0.77	52.6023	0.6
26.76	0.36	41.4153	0.36
23.18	0.15	32.3628	0.19
20.30	0.05	25.1961	0.09

Table 3.7 EZW and N-TO-P for Cameraman Image, PSNR vs. Bitrates.

EZW [62]		Proposed Algorithm	
PSNR(db)	TIME (sec)	PSNR (db)	TIME (sec)
35.55	182.62	69.5081	1.39
30.75	102.21	52.6023	1.37
26.76	64.59	41.4153	0.478
23.18	43.34	32.3628	0.279
20.30	35.47	25.1961	0.23

Table 3.8 EZW and N-TO-P for Cameraman Image, PSNR vs. Time.

Figure 3.31 shows an example of the original and reconstructed image after applying the 2D three level proposed algorithms with bitrates 0.36.

(a) (b)

Figure 3.31 (a) the Original Cameraman Image (b) Cameraman with 0.36bpp

Figure 3.32 the N-To-P 2D-DWT proposed algorithm shows that the rate distortion performs better than the EZW technique. A huge difference in results between both

techniques, moreover this big gap between both techniques increase more while dealing with low and medium compression ratio.

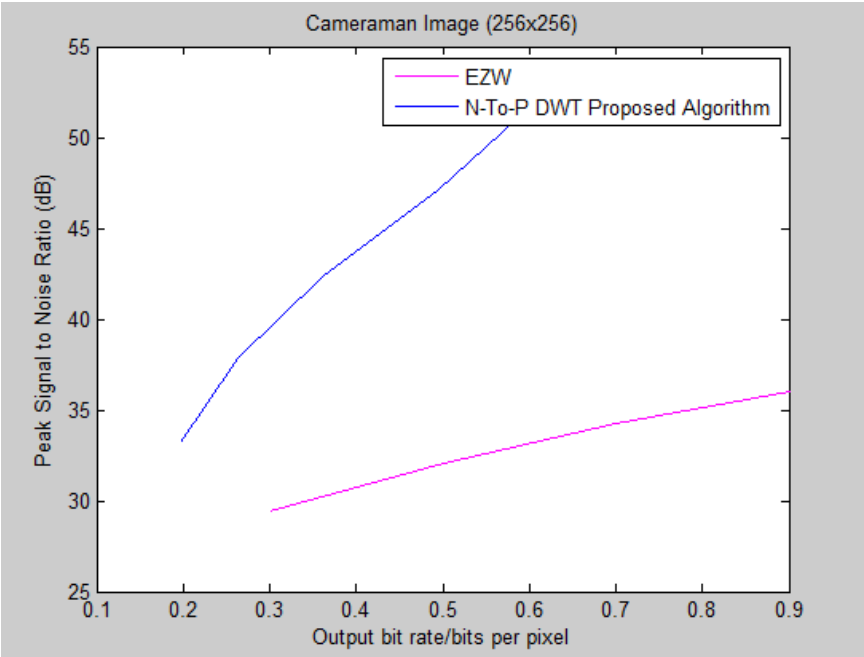


Figure 3.32 Performance of the 2D-DWT N-TO-P and EZW DWT compression for Cameraman Image, PSNR vs. Bitrates.

Another results are presented in figure 3.33 that shows the enormous Time difference between the proposed 2D-DWT algorithm and EZW. In the graph, it is can been noticed that the computational complexity represented by the time is increasing while the bit rate increase for the EZW technique, meaning that the increasing the number of the quantized coefficients, makes the EZW spending more time in coding this coefficients. On the other hand, the time of the N-To-P DWT proposed algorithm could be considered as negligible in complexity comparing to EZW, in addition, the complexity does not affected by the compression level, it could be considered as fixed in all levels, low or high compression, and this point had been illustrated in chapter 3.

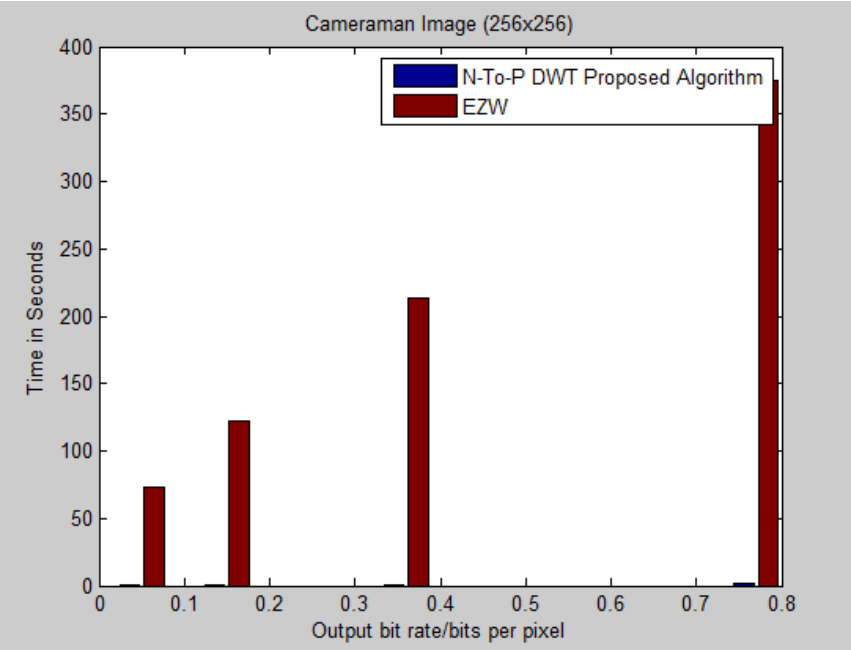


Figure 3.33 Cameraman Image Algorithm Complexity (Time is seconds)
N-TO-P 2D-DWT vs. EZW DWT

3.2.3.2 SPIHT

The SPHIT is an improved technique than EZW [64]. It gives better results than EZW [65]. SPHIT stands for Set Partitioning in Hierarchical coding Techniques [7].

After applying the DWT and getting the decomposed sub-bands, the SPHIT [65] is applied in order to exploit the redundancy and minimize the file size by coding it. SPHIT is an advanced technique over the EZW. It got the same basic idea as EZW,

but the only difference is the way it is treating the Zero-tree, which called spatial orientation tree in SPHIT [66].

In [66] the author applied different filters haar, db1, db6, bior 1.5 ...and cdf9/7. And he found that the Cohen-Daubechies-Feauveau 9/7-tap filters (CDF 9/7) give the best results. Table 3.9 compares the proposed algorithm with the results from [66] tested on Lena Image, and table 3.10 compares Lena time results compared to results from [13], and it shows the reduction in time in the proposed algorithm.

SPHIT [66]		Proposed Algorithm	
PSNR	**Bitrates**	**PSNR**	**Bitrates**
36.03	**0.9**	**69.4928**	**1kB**
34.22	**0.7**	**52.4633**	**0.6**
32.07	**0.5**	**46.8205**	**0.5**
29.43	**0.3**	**36.7346**	**0.29**

Table 3.9 Lena Image, SPHIT, PSNR and Bit rates.

Bitrates	DWT-SPHIT Time (sec) [13]	Proposed Algorithm Time (sec)
0.3	6.44	0.3
0.6	12.33	1.4
0.9	14.41	1.6

Table 3.10 Lena Image, SPHIT, Bit rates and Time.

3.2.3.2.1 Results

Figure 3.36 shows the difference between the proposed algorithm using 2D-DWT with the N-To-P algorithm and SPHIT using the PSNR in db versus the Bitrates (bit per pixels). Lena image has been shown in figure 3.34 and the reconstructed image with different PSNR is represented in figure 3.35.

Figure 3.34 Lena Image

Figure 3.35 (a) Lena DWT-SPIHT at 0.3bpp PSNR= 36.03dB

(b) Lena N-TO-P at 0.29bpp PSNR=36.73bB

Similar, to the results in figure 3.32, the N-To-P DWT proposed algorithm shows an efficient performance comparing to the SPHIT technique, figure 3.36, and this efficiency increases with the low compression ratios, but from the complexity point, the SPHIT is acting better than the EZW, but still the time the N-to-P algorithm is spending is less than the SPHIT, as shown in figure 3.37.

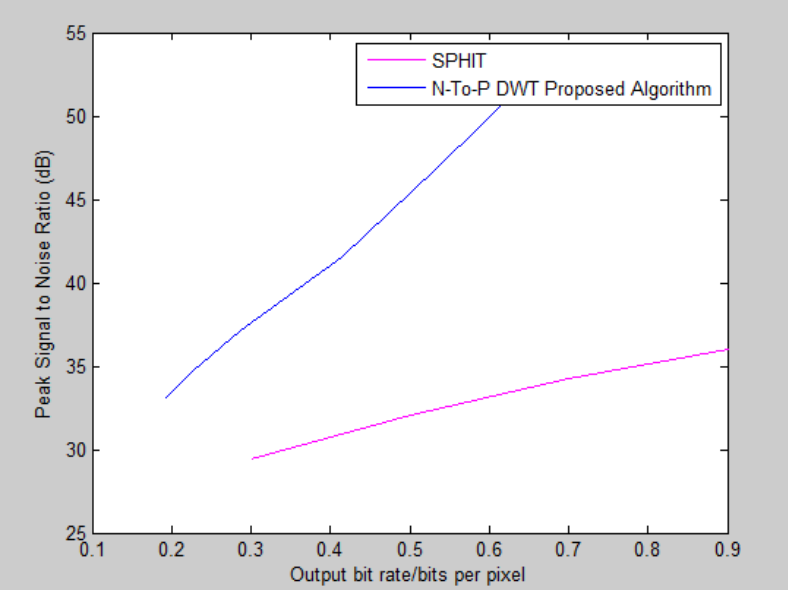


Figure 3.36 Rate Distorsion Performance of 2D-DWT vs. SPHIT DWT of Lena Image.

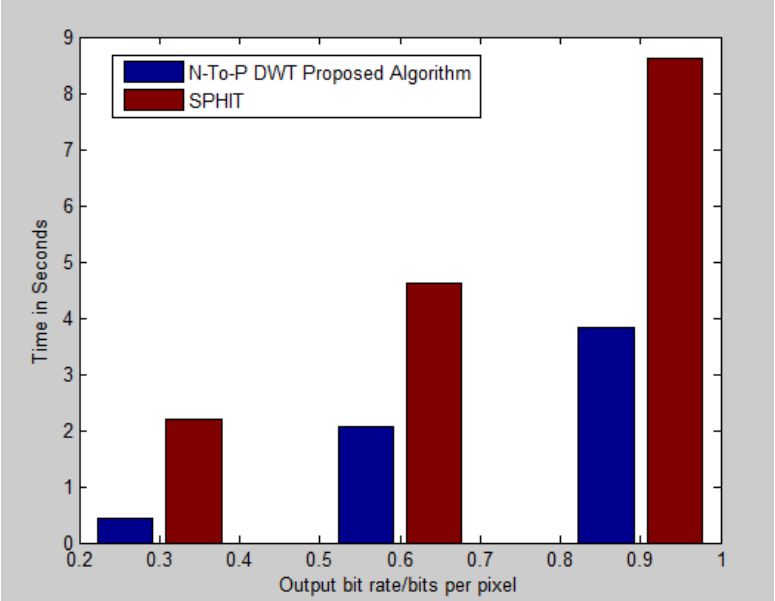


Figure 3.37 Performance of Lena Image Complexity (Time is seconds) (N-TO-P DWT vs. SPHIT DWT).

3.2.3.3 EBCOT

JPEG 2000 is one of the common standard that is been implemented to overcome the jpeg drawbacks [67]. It is like the JPEG standard divided into three levels, first applying the transform coding following by quantizing the coefficients and finally come the entropy coding. JPEG2000 uses the DWT as a transform coding, it

decompose the original image to different frequency band called sub-bands. Then the output coefficients quantized then the EBCOT coding is applied [68] [69].

EBCOT is stands for Embedded Block Coding with Optimal Truncation; it is basically divided into two parts, part one which is normally known by tier 1, the quantized coefficients from all sub-bands are divided into blocks, this blocks known as Code-Blocks. For every code-block there are three codes coming from the bit-plane technique is applied in order to gather a related data. Finally the arithmetic code is implemented to end up with the compress bit-stream. Part two (tier 2) its main target is to optimize the compression ratio by getting rid of some compressed data, and the rest of the compressed data gathered to end up with final bit-stream.

From table 3.11 and figure 3.38 for Lena Image and table 3.12 for Barbara Image, it is quit obvious that the rate distortion performance of the proposed N-To-P DWT algorithm performs better than EBCOT only in the low compression area, but getting lower PSNR meaning increasing the compression level causes weakness to the N-To-P DWT algorithm against the EBCOT. This bad performance is due to the way the

JPEG2000		Proposed Algorithm		
PSNR	Bitrates	Bitrates	Bitrates (Without Index)	PSNR
-	-	1kB	0.83	69.4928
52.06	0.6	0.6	0.5	52.4633
47.25	0.5	0.5	0.39	46.8205
43.22	0.4	0.4	0.289	41.4179
39.67	0.3	0.29	0.19	36.7346
36.10	0.2	0.23	0.15	34.1323
32.95	0.1	0.19	0.13	32.3565
28.47	0.01	0.12	0.08	28.47

Table 3.11 Lena Image, EBCOT

2D-DWT coefficients are quantized, as it been mentioned before in chapter 1, that the N-To-P proposed algorithm was first designed to simulate the DCT standard, so in this section the algorithm was implemented in a simple way using the same way of quantization that have been used with DCT, and for EBCOT there is a different way to quantize the coefficients in order to exploit the DWT characteristics which is not

used in this chapter nor in chapter 4 with the 3D-DWT. However, this could be one of the future works, to implement an algorithm that simulates and exploit the advantages of DWT. Although of this, this proposed algorithm that is represented in this chapter still gives good results comparing to the other techniques, and also better results in low compression level comparing to EBCOT. Another thing that is necessary to mention that in this thesis the entropy coding is the main block that was substituted by another proposed one, and in order to compare the performance of this new entropy coding it has been tested with a pre-processing blocks.

JPEG2000		Proposed Algorithm		
PSNR	Bitrates	Bitrates	Bitrates (Without Index)	PSNR
-	-	1.08	0.9	69.5989
50.39	0.75	0.79	0.62	52.5156
43.85	0.6	0.69	0.52	46.8456
40.06	0.5	0.57	0.4	41.7129
35.84	0.4	0.45	0.3	34.8214
31.26	0.3	0.38	0.24	31.4974
27.17	0.2	0.17	0.1	24.7098
23.71	0.1	0.14	0.08	23.6611

Table 3.12 Barbara Image, EBCOT

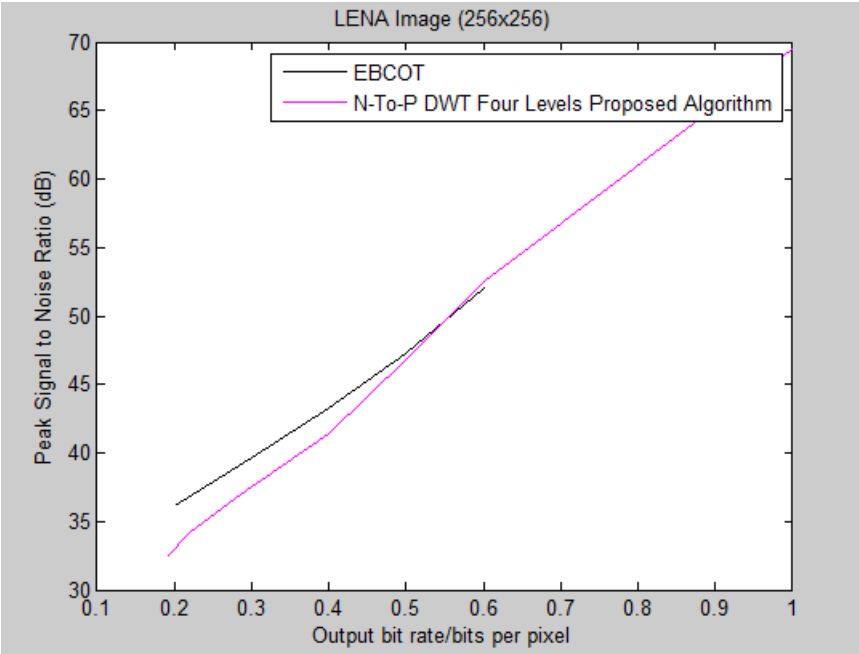


Figure 3.38 RD Performances of N-To-P and EBCOT algorithms for Lena Image.

Figures 3.39 and 3.40 are a zoom parts from the original figure 3.38 to show the performance of the proposed algorithm with the low compression ratio for Lena Image.

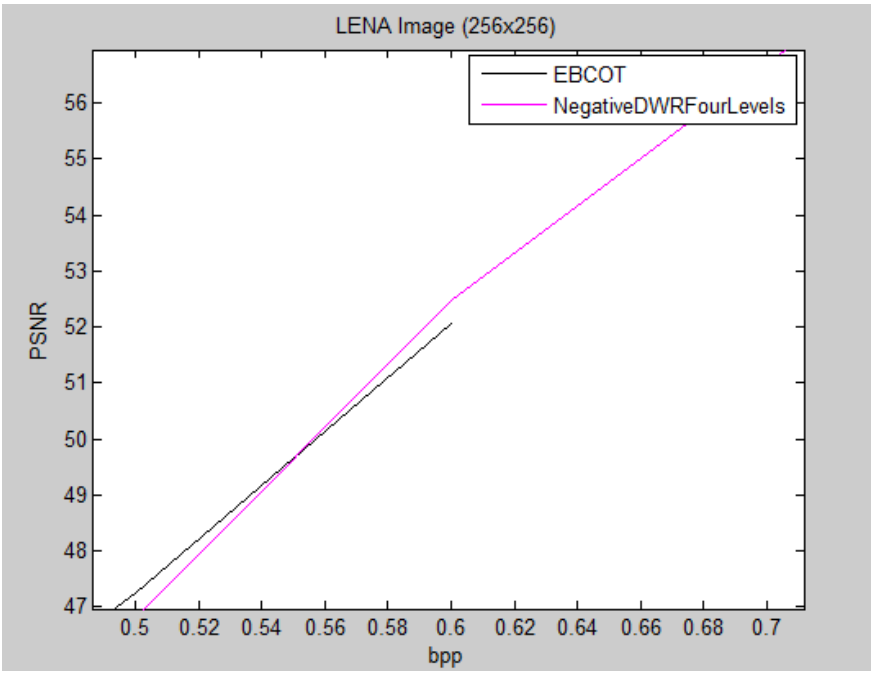


Figure 3.39 Performance of N-To-P and EBCOT for Lena Image from 0.5bpp to 0.7bpp.

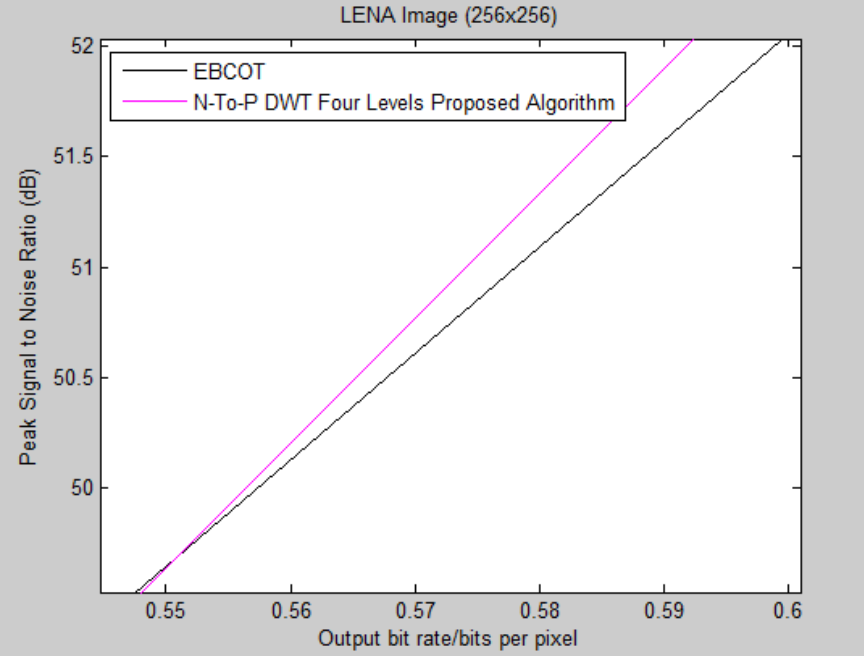


Figure 3.40 Performance of N-To-P and EBCOT for Lena Image from 0.55bpp to 0.6bpp.

Figure 3.41 shows the EBCOT and the N-To-P proposed algorithm with four DWT levels, but this time the N-To-P algorithm without the Index array file, in order to show how effective to compress the index algorithm. Followed by the Barbara results but including the Index array after compressing it using the Difference Index algorithm, in figure 3.42.

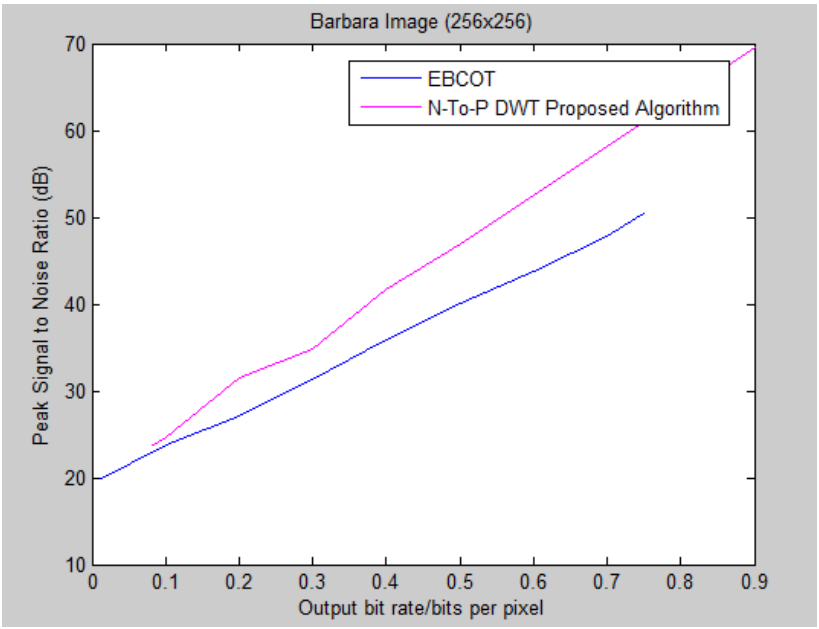


Figure 3.41 Barbara Image PSNR vs. bitrates N-To-P DWT Four Levels vs. EBCOT

DWT without the Index

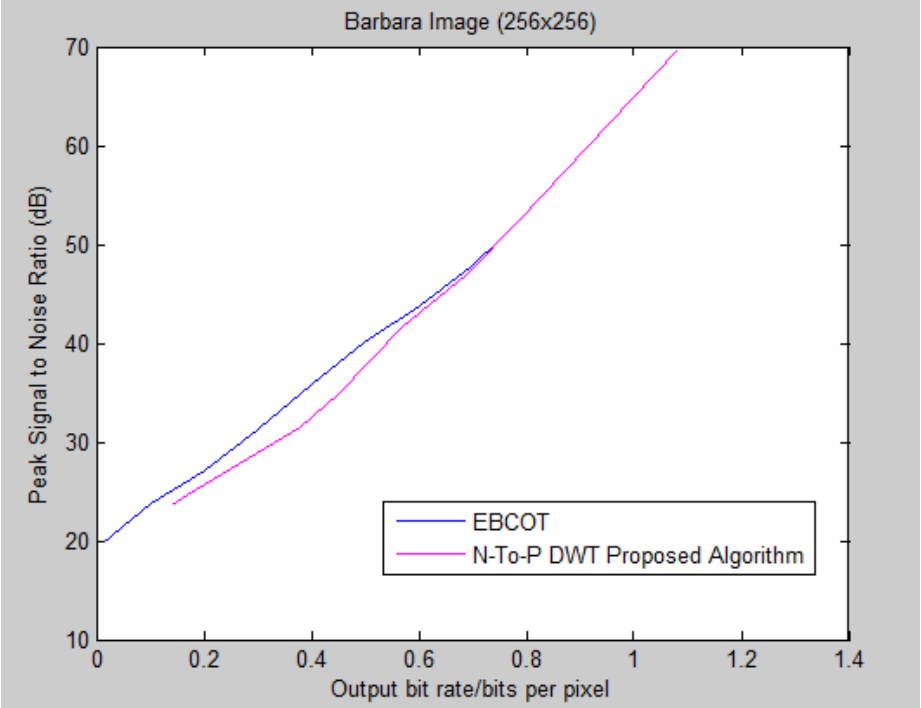


Figure 3.42 Barbara Image PSNR vs. bitrates
N-To-P Three Levels DWT vs. EBCOT DWT

Another factor that should be considered is the time; table 3.13 shows a comparison between the EBCOT and the proposed algorithm for both Lena and Barbara 2D-images. The time that is measuring in seconds for the EBCOT, is calculated for the encoding part only, however the proposed algorithm time is calculated for the whole algorithm, starting from the input raw image till the reconstructed output, includes encoding and decoding. The big difference between the time expending is clear.

Image	JPEG2000 (EBCOT)		Proposed Algorithm	
	PSNR	Time (sec) (Encoding)	PSNR	Time (sec) (Encoding + Decoding)
Lena	28.47	2.809	28.3361	0.3
Barbara	19.7	4.758	22.7185	0.2

Table 3.13 Time Comparison between JPEG2000 (EBCOT) & Proposed N-To-P Algorithm

PSNR	Time (sec)
69.5989	1.4
52.5156	1.39
46.8456	1.5
41.7129	1.08
34.8214	0.7
31.4974	0.5
24.7098	0.2
23.6611	0.2

Table 3.14 Barbara Image PSNR vs. Time in Seconds

PSNR	Time(sec)
69.4928	1.6
52.4633	1.4
46.8205	0.78
41.4175	0.46
36.7346	0.3
34.1323	0.29

Table 3.15 Lena Image PSNR vs. Time in Seconds

Tables 3.14 and 3.15 also show the time the proposed algorithm is taking for both Barbara and Lena image respectively. Figures 3.43 and 3.44 present a comparison between the performances of the 2D-DCT proposed earlier in this chapter and the 2D-DWT algorithm proposed in this section for Lena and Cameraman images. The 2D-DCT proposed algorithm performs better in the high compression levels comparing to the 2D-DWT proposed algorithm, and this was one target that was illustrated in chapter 1 is to reach to the level that the DCT works better in the high compression ratio, which is contrary to the fact that JPEG2000 works effectively more than the JPEG standards in the medium and high compression ratio.

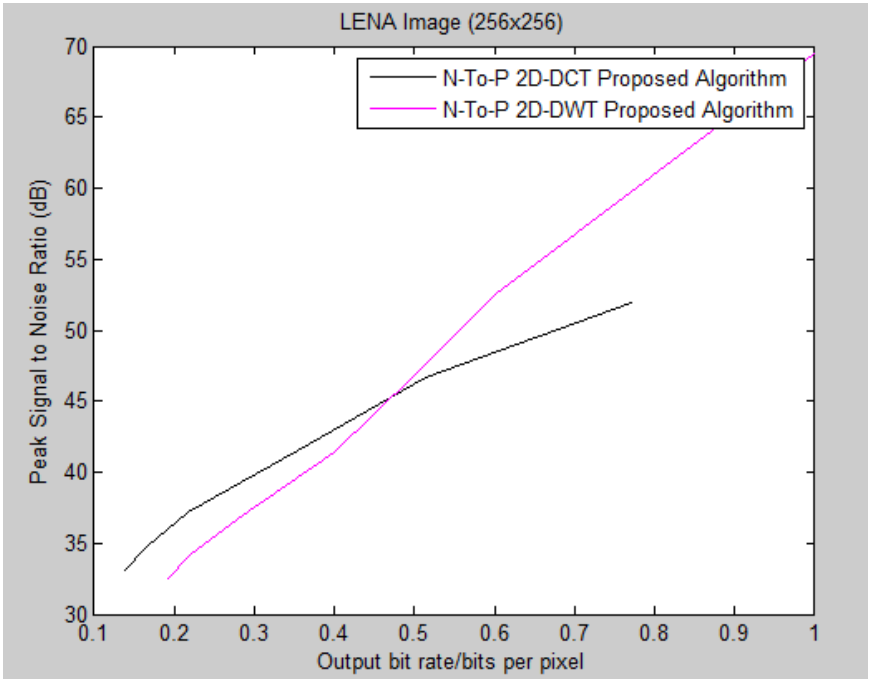


Figure 3.43 Performance of 2D-DCT and 2D-DWT N-TO-P algorithm for Lena Image.

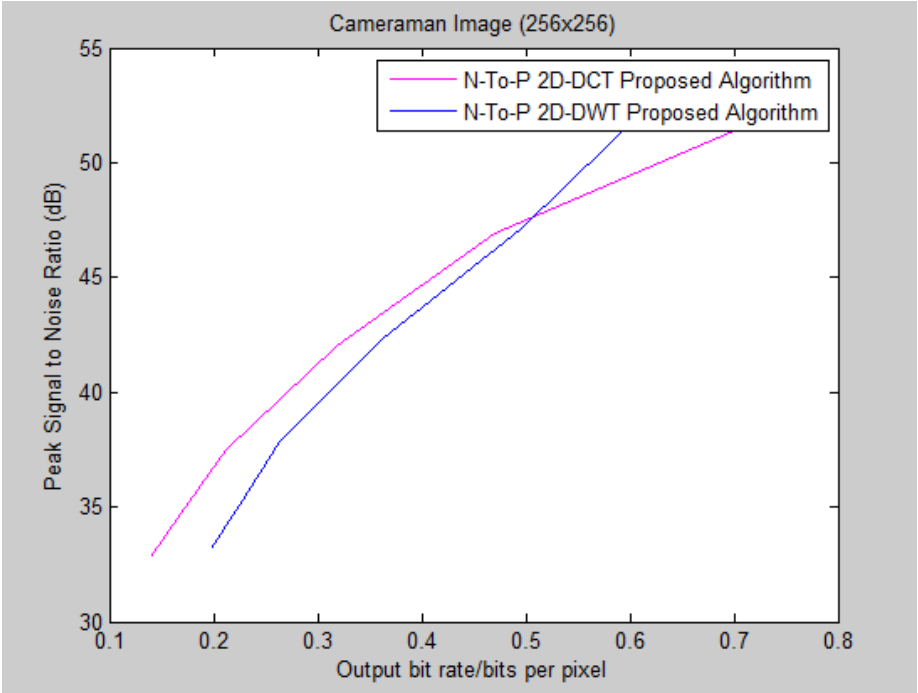


Figure 3.44 RD Performance of 2D-DCT and 2D-DWT N-TO-P algorithms for Cameraman Image.

Figure 3.45 illustrates the comparison between using 2D-DCT, 2D-DWT and EBCOT. It shows the DCT is acting better than DWT in high compression ratio, while the DWT performs very well in low compression ratio and DWT is better with

low compression ratio.

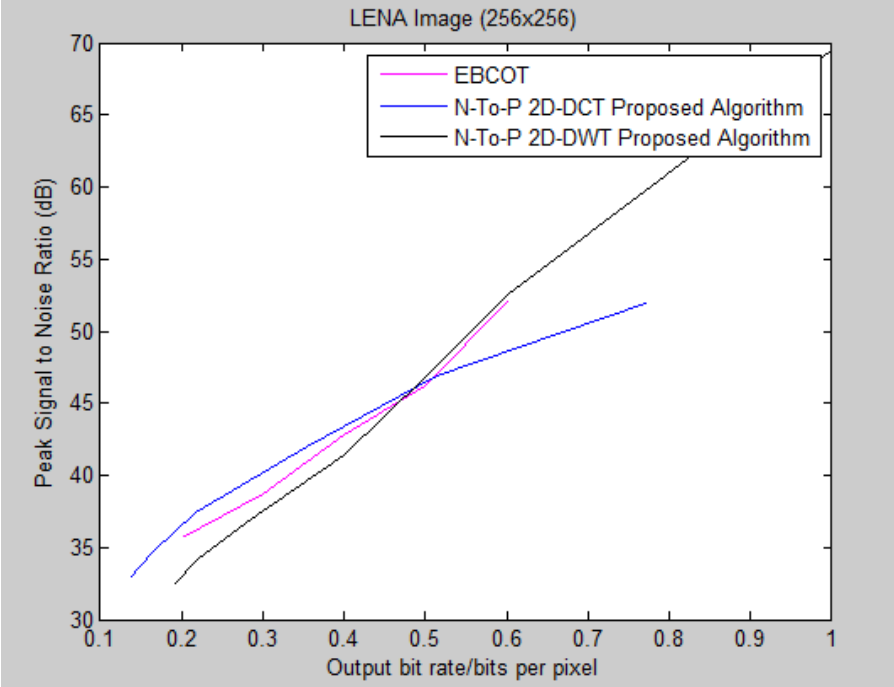


Figure 3.45 Comparison between 2D-DCT, 2D-DWT and EBCOT (JPEG2000) algorithms for Lena Image.

A comparison between both JPEG baseline, JPEG2000, N-To-P 2D-DCT and DWT are presented in figure 3.46,in high compression ratio, the worst scenario was the 2D-DWT N-To-P proposed algorithm and that is due its simplicity and the way the coefficients quantized, followed by the JPEG baseline standard and JPEG2000, then the 2D-DCT N-To-P algorithms performs better than the others. In the low compression ratio, with less data loss, the JPEG2000 and the 2D-DWT N-To-P proposed algorithms act better than the other two algorithms.

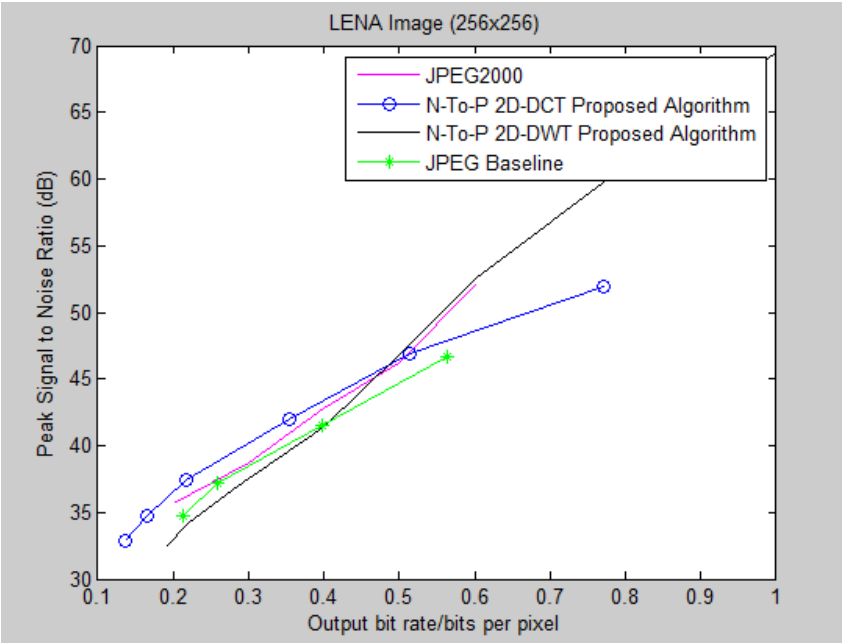


Figure 3.46 RD comparison between 2D-DCT, 2D-DWT, JPEG2000 and JPEG algorithms for Lena Image.

3.3 Conclusion

By applying the proposed algorithms a better Image compression is achieved. These algorithms have been applied to different sizes of images. Table 3.3 shows a comparison between the JPEG standard and the proposed one. The results confirm that the proposed algorithm is lossless and that it to get the same PSNR as performs better than the JPEG base line with respect to image quality and compression ratio. The results also show that the proposed algorithm has a much lower computational complexity. It also proves that there is a dramatically difference in Time reduction that is expended to implement both algorithm due to the less complexity and simplicity of our new one. In addition to that, it illustrates the difference of average ratio of the bit rate for each image, and shows that the proposed algorithm works more efficiently. This severely decrease in time is due to the time the JPEG uses in

RLE. Also the effect of the Quantization technique is represented and the results have been shown that the fixed Q Factor proof better results than the use of the table.

The 2D-DWT proposed algorithm shows its efficiency and reduction in time comparing to EZW and SPHIT in a massive way. The proposed encoding algorithm basically was design for the DCT JPEG standard, in this section a very simple way was implemented using the DWT without getting advantages from exploiting the advantages of the DWT, via using the proper quantization method that is working effectively with DWT and this can be shown from the comparison done with the EBCOT, So more future work should be done from the DWT vision, especially that there is a huge reduction in time between both algorithms. Another issue is that in the DWT proposed algorithm the coding technique is applied to all sub-bands without neglecting any one of them.

The way also the positive–Negative 2D-DWT technique is different than the 2D-DCT proposed algorithm that it is simpler, the coefficients were horizontally scanned, there was no need to use the adaptive End of Block for AC and also there was no need to apply the DPCM technique because we already didn't separate the coefficients. The algorithm is giving us an efficient performance, and the computational complexity is really reduced compared to all other techniques.

Also as mention before in the previous section in this chapter, that the different ways the index is treated could be optimized more and in the future work, it could give better result, as it was shown that using different technique to shrink the size of the index is affecting the size of it, especially that the index file extremely affect the Compression process.

Chapter 4

Novel Entropy Coding for 3D DCT and DWT based compression scheme on II and optimizing 3D-Integral compression process

4.1 Introduction

In this chapter the 3D-DCT Transform coding is applied, and it has been shown the difference between applying different encoding methods like JPEG with Define Huffman Table (DHT), Statistical Huffman, and Negative Difference Method described in chapter 3. A detail review of the DCT is described in Appendix. Also a 3D-DWT algorithm will be implemented and will be compared to the 3D-DWT lifting scheme that is represented in [76].

All of the above encoding techniques are applied after level shifting, extracting the Viewpoint Images from the original Integral Image, applying the Discrete Cosine Transform and finally quantize the Coefficient.

Applying the compression algorithms on extracted Viewpoints is less complex and faster than the one applied to the entire Image. Both schemes have been tested. A comparison between both techniques had been done and the results showed that using the viewpoints from the Integral Image rather than the entire Image is more effective.

In this chapter, the 3D Discrete Cosine Transform is applied on two different types of Integral Images, the first is the Unidirectional Integral Images (UII) and the Second is the Omni-Directional Integral Images (OII) [32]. Figure 4.1 show example of Unidirectional Image (Horseman Image), and Figure 4.2 shows example of Full parallax Image [2].

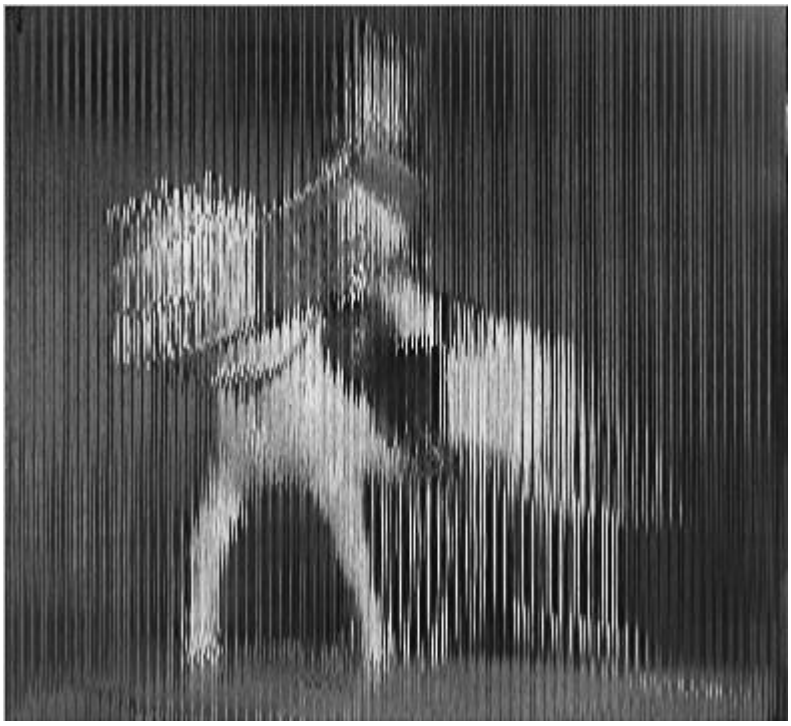
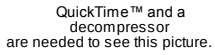


Figure 4.1 Horseman 3D-Unidirectional Integral Image.

Figure 4.2 Full Parallax Omni-Directional Integral Image.

The Unidirectional Images are captured using a lenticular sheet that consists of a number of micro lenses that arranged horizontally [29]. Each Micro lens record a 2D planar array that is totally different than the normal 2D Images. While the full

parallax images are capturing through Type 2 camera represented in chapter 2 with different lenses size as shown in figures 4.3.



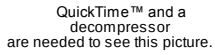


Figure 4.3 Full Parallax 250 micro lens and 90 micro lens Integral Images.

The Viewpoint is a parallel Orthogonal projection while the normal 2D is perspective projection [71]. Figure 4.4 and 4.5 Show the difference. The disparity between each microlens and its neighbour has a very small angle which leads to a very high cross-correlation between the adjacent micro-images.

The Viewpoint Parallel Projection, the 2D Traditional Image Perspective Projection

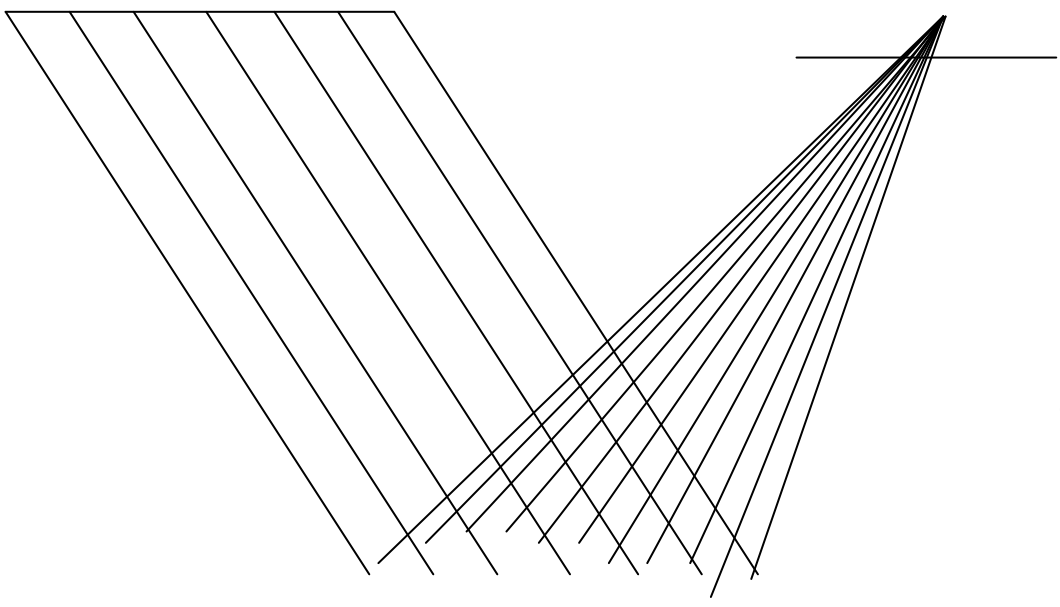


Figure 4.4 Difference between Parallel and Perspective Capturing.

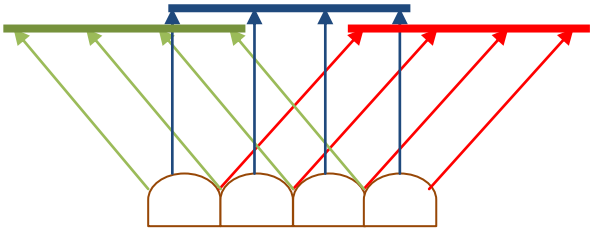
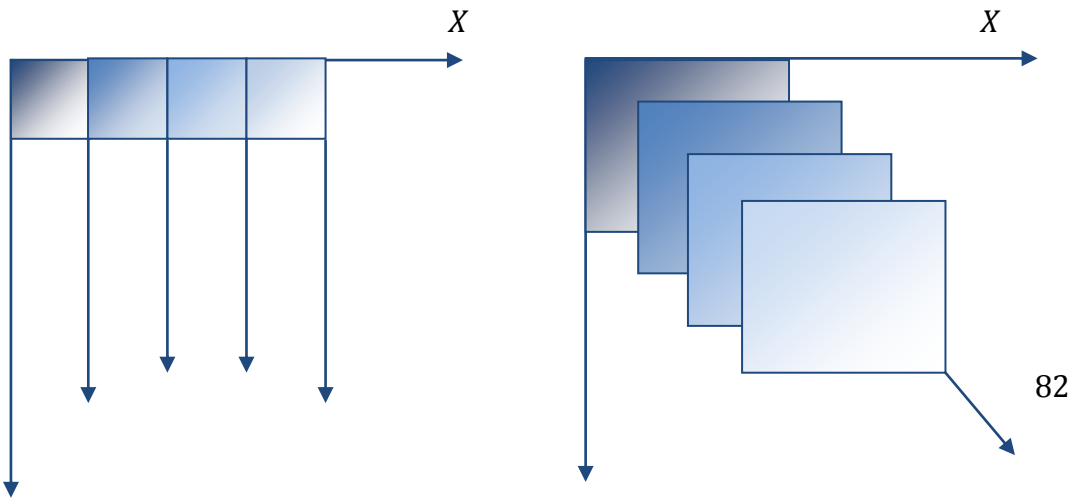


Figure 4.5 the Microlenses and the parallel Orothognal projection

In addition there is also a high correlation between the pixel in the same micro-images age and its neighbour. It is worth to mention that the Intra– micro-images are the correlation between the pixels in the 2D planar density and the intra- micro-images is the correlation between the Pixels in the third dimension that gives the Volume figure 4.6. The 3D-DCT Volume came from aligning the adjacent micro-images all together in order to achieve the third dimension [71].



Y *Y* *Z*

Figure 4.6 3D Volume

So, the separability property has been used in this chapter. The 3D-DCT is implemented in two levels, first to exploit the correlation between the neighbour pixels in the same (Viewpoints) intra- micro-images, a traditional 2D DCT (8x8) is used, Second in order to take advantage of the cross correlation between the micro-images a 1D-DCT is applied (Inter- micro-images) and that is what gives the volume for the image. The overall 3D-DCT is result of a considerable reduction of data [72][73].

In Omni-directional Integral Images: The process is similar to the Unidirectional Integral Images but the microlenses are laid in both directions vertically and horizontally [22] instead of one direction like the Unidirectional Images. Also in the Omnidirectional Integral Image the 3D-DCT will be used in order to utilize the intra and inter sub-image correlation [22] and gain a high compression ratio.

For Unidirectional Images, three Images were used; Horseman Image, Micro Image and Tank Image as shown in figure 4.7.

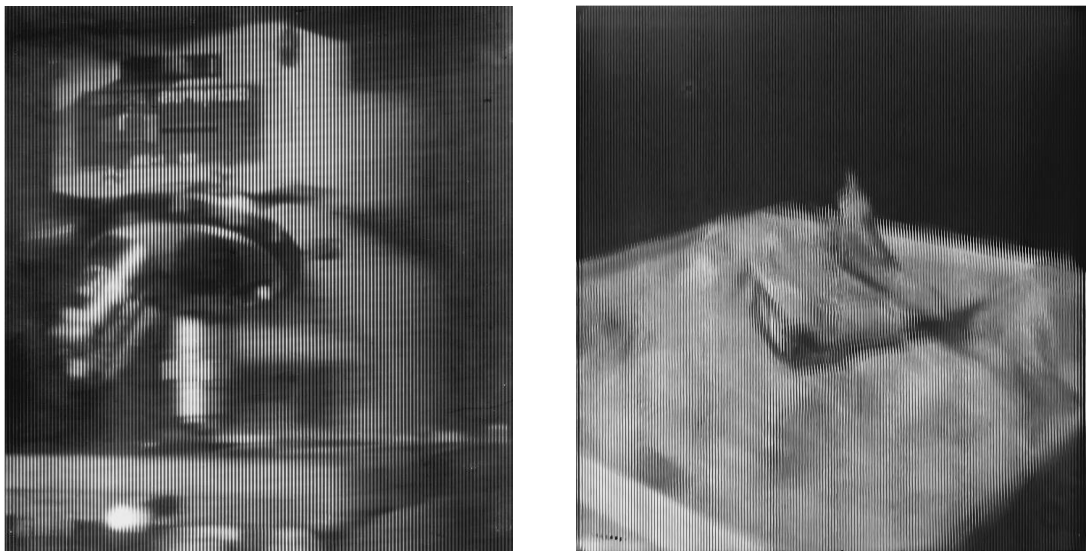


Figure 4.7 Micros and Tank integral Images

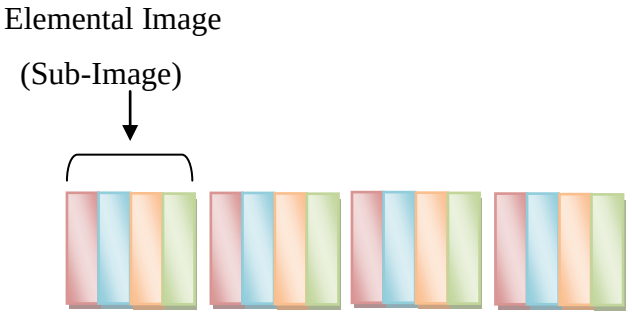
A comparison between three schemes is applied on the Integral Images using the 3D-Discrete Cosine Transform. First, a scheme similar to the JPEG, which is using Define Huffman Table (DHT), Second, a scheme that uses the Statistical Huffman Code, and finally, a scheme that uses the Negative Difference Algorithm that was illustrated in the previous Chapter but this time it will be applied on the 3D-Integarl Images instead of 2D-Images. All of these techniques will be applied after extracting Viewpoints from the original Input Image and also the data will be sampled Shifted.

4.1.1 Sampling Shifted

So First the Sampling Shift is applied on the original Integral Image, The pixel of a gray Image is ranged between 0 and 255, 0 indicates that the colour of the pixel is black and on the other hand the 255 indicates that the colour is White.
Because the DCT is working with sign pixels so the pixel is shifted by 128 and the range change to become between -127 and 128. This process occurs by subtracting the value P from each pixel in the image before applying the DCT and this process will be inversed after applying the Inverse DCT [74].

4.1.2 Viewpoint image extraction

As mentioned before, the microlense Sheet contains number of horizontal Microlenses in Unidirectional Integral Images and both horizontal and vertical Images in the Omni-directional Images. Each microlenses has a very small resolution capturing 2D planar density, which different than the traditional perspective capturing images, which is illustrated in figure 4.8.



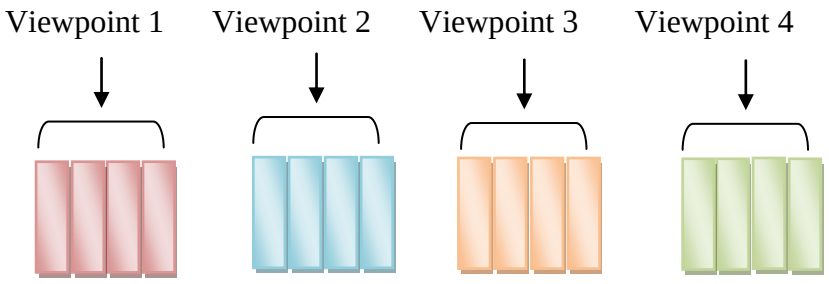


Figure 4.8 Extracting Viewpoints

The small images resulting from the each microlenses called Elemental Images or Sub-Images but another term is going to be used which is Viewpoint Images. The Viewpoint Images is integrated from the pixels in different microlenses but got the same position. The Viewpoint Image contains all the data needed from one certain view direction [70] [75].

The compression in this chapter and the following are going to be implemented on the Viewpoint Images. The entire Image compression has been tested instead of the viewpoints but the former was faster and simpler.

The number of the Viewpoints and its dimensions depends on three parameters, the number of the microlenses, the number of pixels in a microlens and the size of the image (X axis in Unidirectional and both (X, Y) in Omni-directional case).

$$VP = M / P.$$

Where , VP is the number of the Viewpoint Images, M is the number of the Microlenses in the original Image and P is the number of the pixels per Microlens.

Or.

$$VP = \text{size of X axis} / M$$

Number of microlese for Horseman = 160.

Number of microlese for Tank= 176.

Number of microlese for Micros= 144.

The Microlens Contains=1280 Pixel Horizontally in Horse Image.

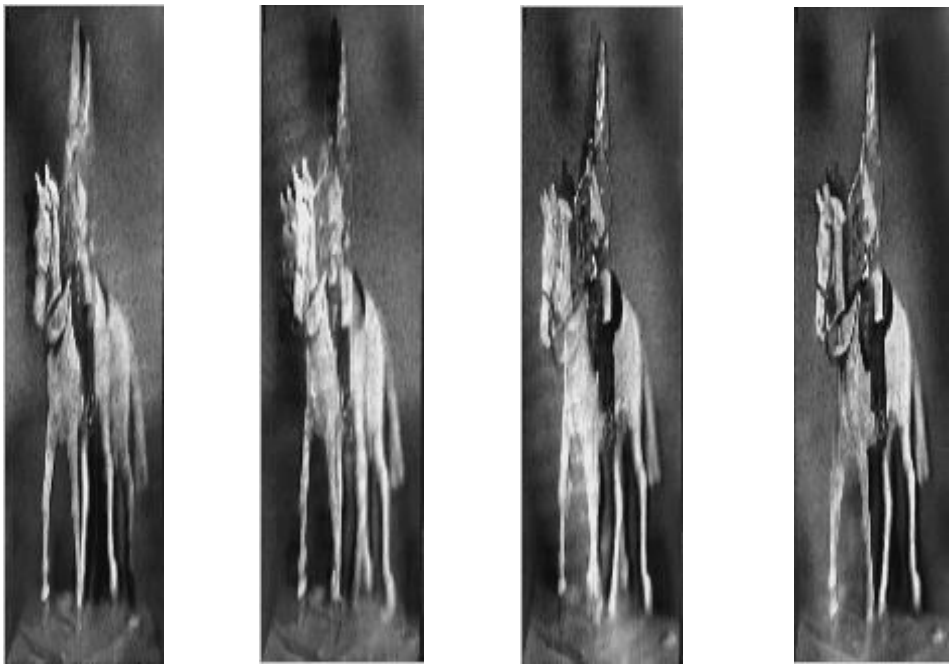
The Microlens Contains=1152 Pixel Horizontally in Micros Image.

The Microlens Contains=1408 Pixel Horizontally in Tank Image.

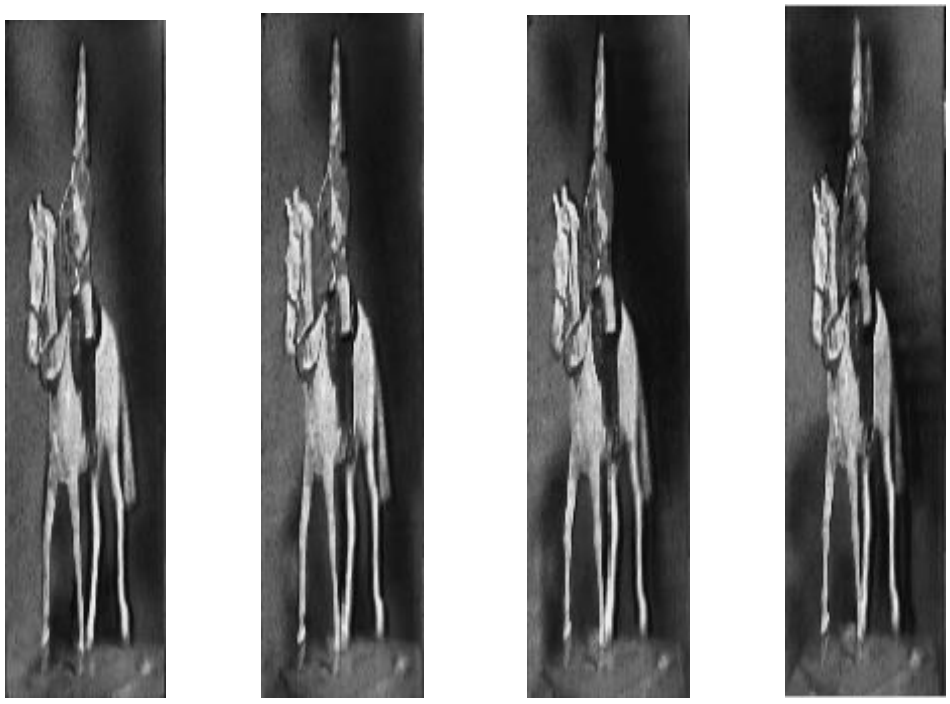
So the Number of Viewpoints is 8 Viewpoints in Horse Image, Micros Image and Tank Image.

The Omni-directional Image got 768 Viewpoints because its size is (3584 x 5376).

Figure 4.9 shows the Viewpoints for the Image horse



Viewpoint 1 **Viewpoint 2** **Viewpoint 3** **Viewpoint 4**



Viewpoint 5 **Viewpoint 6** **Viewpoint 7** **Viewpoint 8**

Figure 4.9 Horseman 8 Viewpoints

There are some advantages of extracting the viewpoints from the 3D-Integral input data:

a) The 2D content is more familiar than the 3D content, and extracting viewpoints end up with content similar to the traditional 2D Images.

b) Using Viewpoints improve the chances of reducing the memory requirements.

c) Viewpoints can be used separately, meaning that a part of the image can be for example decoded independently, rather than accessing the whole image, and this feature will be exploited later in this chapter in the proposed adaptive algorithm and also will be used later in the video coding chapter.

4.1.3 3D- DCT Theory

Each Viewpoint is divided into 8 by 8 blocks [59]. Figure 4.10 shows the block diagram of Forward 3D-DCT.

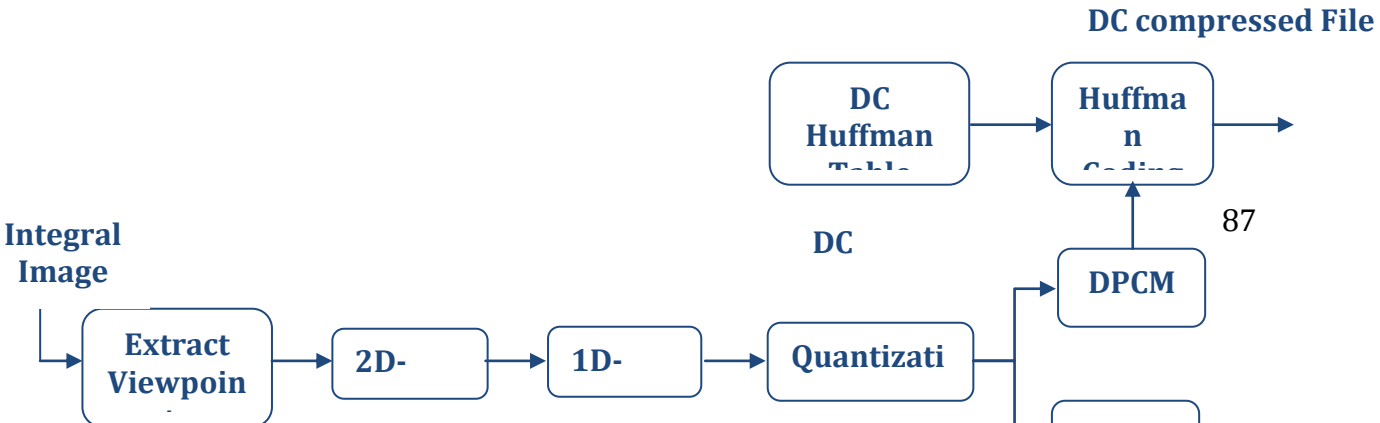


Figure 4.10 3D Forward DCT Diagram for Integral Images with DHT.

The following equations are the 3D-FDCT & 3D-IDCT mathematical equations:

Forward 3D-FDCT:

2D-FDCT

$$F(u,v) = \frac{2}{N} C(u)C(v) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x,y) \cos[\frac{(2x+1)u\pi}{2N}] \cos[\frac{(2y+1)v\pi}{2N}] \quad (3.1)$$

where,

$$C(u), C(v) = \frac{1}{\sqrt{2}} \; for \;\; u,v = 0.$$

$$C(u), C(v) = 1 \quad \text{Otherwise.}$$

Then, apply Forward 1D-DCT

$$X(k) = \sqrt{\frac{2}{N}} C(k) \sum_{i=0}^{N-1} x(i) [\frac{(2i+1)k\pi}{2N}] \qquad (3.2)$$

Inverse 3D-IDCT:

2D-IDCT

$$f(x,y)=\frac{2}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}C(x)C(y)F(u,v)\cos[\frac{(2u+1)x\pi}{2N}]\cos[\frac{(2v+1)y\pi}{2N}] \qquad (3.3)$$

1D-IDCT

$$x(i)=\sqrt{\frac{2}{N}}\sum_{k=0}^{N-1}C(k)X(k)\cos[\frac{(2i+1)k\pi}{2N}] \qquad (3.4)$$

4.2 Comparison between SH and DHT

4.2.1 3D JPEG system using Define Huffman Table (DHT)

After extracting the Viewpoints and subtracted 128, and extracting the Viewpoints, as mentioned in chapter 2 one of the characteristics of DCT is the Seperatable, meaning that the first 2D-DCT can be applied (8x8) then the 1D followed it to get the Volume from the third dimension and end up with 3D-DCT as shown in figure 4.10.

As mentioned before, since the 3D-DCT is applied to each (8x8) block, the coefficients resulting from this process is quantized by a fixed factor ranged from 2,4,8,16,24 and 36. The lower the quantization factor is the higher bitrates will be. Starting from this point all the following steps are similar to the JPEG system.

The DC coefficients are extracted from the 64 quantized coefficients in each block after reshaping all the quantized input data to a 1D-array. The Differential Pulse Code Modulation is applied and it results of two symbols (Size) and (Amplitude), on the other hand the rest of 63 AC coefficients are scanned first using Zigzag scanning.

The Zigzag scanning is implemented to all (8x8) blocks and that end up with a vector contains a number of zeros in each block coming from rounding the DCT coefficients after using the Quantization factor, so the output of the zigzag is taken as an input to level one AC encoding which is RLE.

RLE is stands for Run length Encode; each RLE-coded AC coefficient is represented by two symbols. Symbol-1 represents two pieces of information, Run-length and Size. Symbol-2 represents the single piece of information Amplitude as mentioned in the previous chapter. Now both the AC and DC Coefficients are ready for Huffman Define Table to be used. Once the data encoded both files are ready to be sent or stored.

4.2.2 Statistical Huffman

Applying the Statistical Huffman coding is similar to the above algorithm starting from the shifting, extracting Viewpoints till applying the 3D forward DCT and quantizes the coefficients. From this point, the quantized input is scanned horizontally rather than using Zigzag scanning that is due to the way the Unidirectional Integral Image is captured via the lenticular sheet which makes the horizontal coefficient had more correlation than the other directions.

Also both Ac and DC the coefficients are handled identically that means that both coefficients will be treated exactly the same way without needing to split them and no need also to apply different encoding techniques, So the DPCM will not be used for the DC Coefficients and the RLE will not be used for the AC coefficients. The input of the Horizontal Scanning will go directly to the Statistical Huffman coding process.

The statistical Huffman Coding is different than the Define Huffman in the way they encode the data. Instead of using a pre-define table, the Statistical Huffman is the traditional Huffman that depends only on the data of the Image rather than a table.

The advantage of the SH is that the probability of the Quantized Input for the entire Image is the only parameter that is used, However the pre-define table have been tested on many other image in order to get the table to be accurate and work probably with all images but still it is not depending on the characteristics of the image itself, and here in our case the Integral Image is different in characteristic than the normal traditional 2D Images that is why the SH is tested.

Another advantage is that there will be no need to send the table in the header. On the other hand the system is still suffering from some drawbacks during implementation. One of these draw backs is the time, However the system similar to the JPEG is taking a long time but the time for the statistical is a little bit more than the former, that is due to the time it need to scan the whole image to get the probability for the entire Image. It have been mentioned that the Table will not be attached to the header but still the dictionary of the Statistical Huffman that will be created after calculating the probability need to be included for the decoding purpose [39].

Finally, and there will be a drawback that will appear later when the encoding process is finished and the results came, that due to the way the image is scanned (unblock way), the EOB technique is not used, and this is not effecting the results massively because at the end the zeros will be just a number that will have its own code, but this drawback will be neglected while using a small Quantization factor which will results of a low Compression ratio and high Image quality because the number of zeros will not be a huge number (no sever loss) that makes the SH works better than the DHT, on the other side using a high Quantize value end up with high compression ratio with low Image quality, that indicates a huge number of zeros that makes the SH works worse than DHT not that huge difference but still not working probably. Figure 4.11 presents the performance for both the Statistical Huffman and the JPEG using Define Huffman Table.

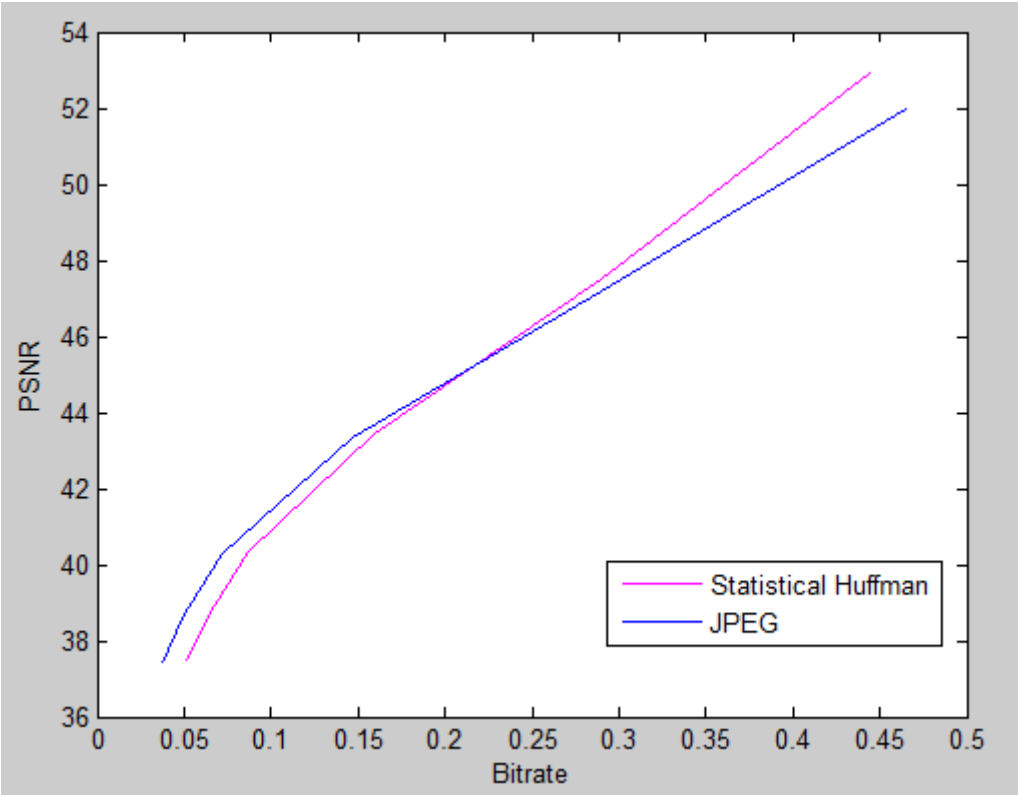


Figure 4.11 Comparison between Statistical Huffman and DHT 3D-DCT algorithms.

4.3 3D- Novel Entropy Coding Technique

4.3.1 Novel Entropy Coding Technique for 3D DCT based Image Compression

4.3.1.1 Unidirectional Images

In here, the system is combined between the System similar to the JPEG and the Negative Difference Algorithm that had been applied previously. The N-TO-P system is exactly the same as JPEG till it comes to the Huffman Code, instead of applying Huffman Code; the Negative Difference Algorithm is used.

Briefly the Negative difference is divided into three algorithms:

1) Negative to positive algorithm.
2) Difference Algorithm.
3) Adaptive End of Block Algorithm.

All of the above algorithms have been described and illustrated before in chapter 3, but the difference is that here it will be applied to 3D-DCT unidirectional Integral Images, which includes viewpoints extraction, so 3D-DCT will be applied in a seperablable way for each viewpoint.

Figure 4.12 shows a comparison in time between the tradition 3D-DCT that works with Huffman coding and the 3D-DCT that works in association with the N-To-P coding algorithm, The N-To-P algorithm performs between and achieved a reduction in computational complexity and thus elimination to the time that spend to run the algorithm.

4.3.1.1.1 Results

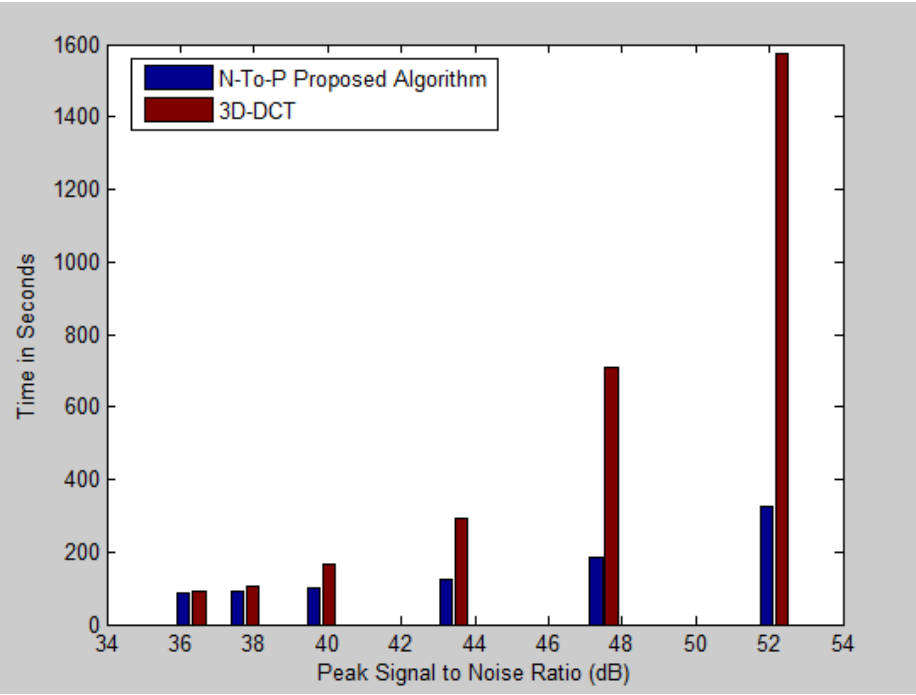


Figure 4.12 Average Time in seconds for both 3D-DCT and N-TO-P.

In addition to reduction in the time, figure 4.13 shows that the average rate distortion performance for the N-To-P proposed algorithm acts better, but the results can be improved better by developing two factors, first the way the image is scanned, instead of using the horizontal way, another methods can be implemented by using a 3D scanning method rather than applying the 2D scan for each viewpoint separately, the second approach is to develop and improve the way the Index Array is encoded.

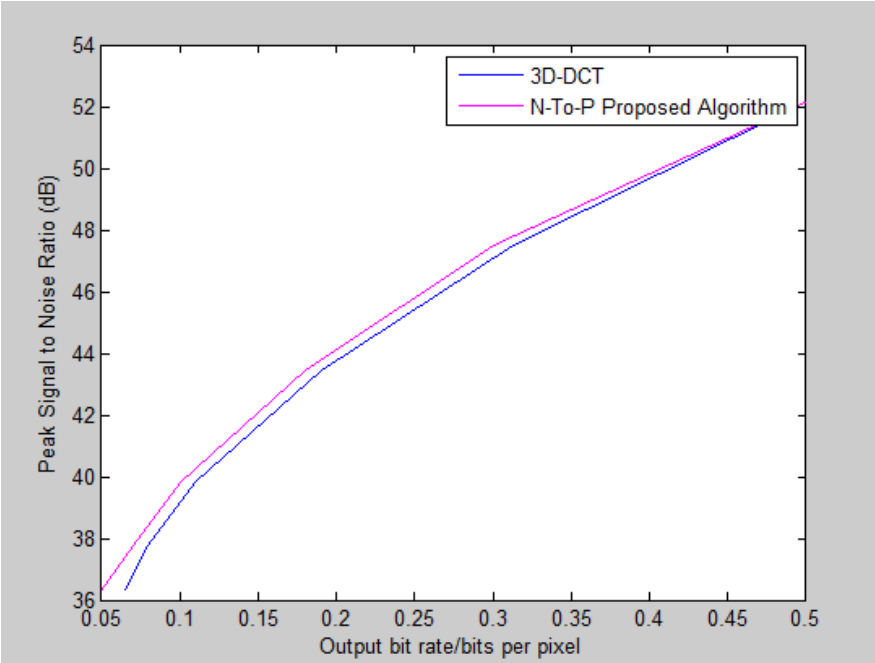


Figure 4.13 Average results for 3D-DCT and N-TO-P.

Figure 4.14 and 4.15 show the performance also of the peak signal to noise ratio versus the bit rates for the Tank image and Micros image, respectively, and the proposed N-To-P algorithm shows its efficiency.

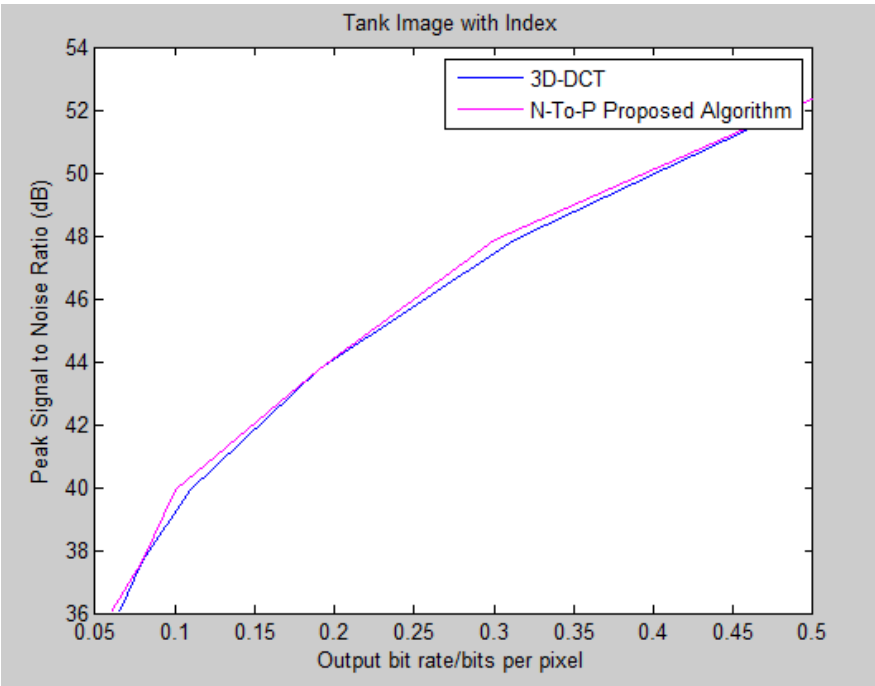


Figure 4.14 Tank results for 3D-DCT and N-TO-P .

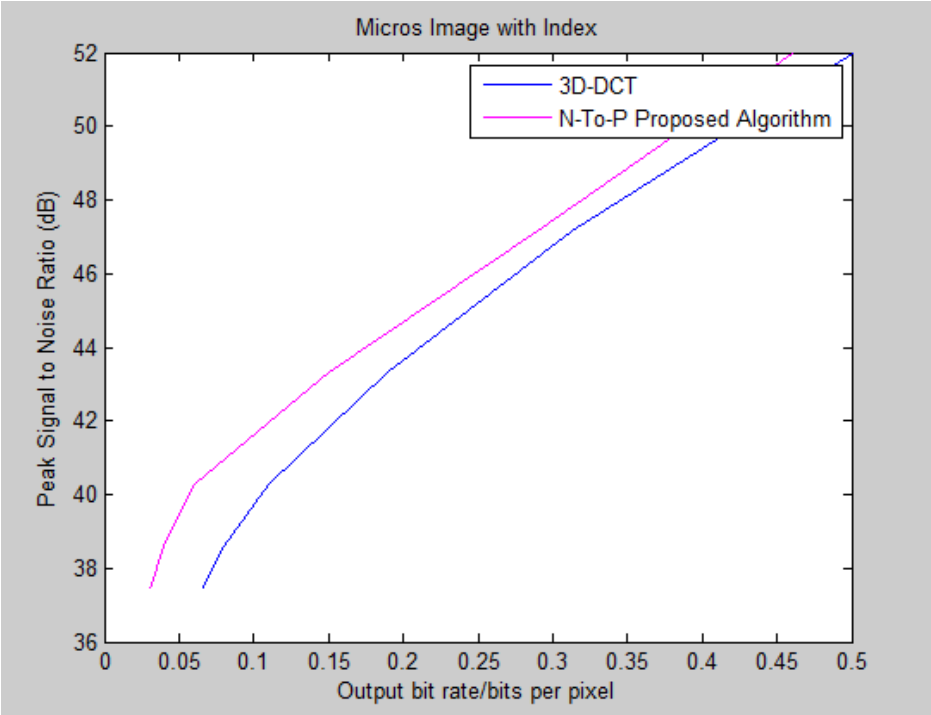


Figure 4.15 Micros results for 3D-DCT and N-TO-P.

Figure 4.16 shows the input and the output unidirectional images, Horse, Tank and Micros, after applying the N-To-P 3D-DCT algorithm for different bit rates. From this image, it can be noticed that the compression ratios are high and at the same time there is no perceptual loss, meaning that the reconstructed image quality is as high as the compression ratio.

Another comparison occurs in figure 4.16, to distinguish the difference between the original image in comparison with both the 3D-DCT with DHT coding technique and the proposed 3D-DCT N-To-P proposed algorithm. While figure 4.17 makes a comparison between different bit rates for the same Tank image.

Figure 4.16 a) Micros Original Image b) 3D-DCT with Bitrates=0.049bpp

Figure 4.17 a) Horseman Original Image b) 3D-DCT with Bitrates=0.113bpp

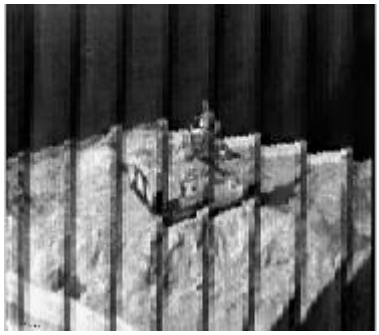


Figure 4.18 a) Tank Original Image b) The 3D-DCT with Bitrates=0.088bpp

Horseman Original

3D-DCT

N-TO-P

Figure 4.19 Comparison between different encoding techniques

With Q=32

Original Tank Image

Bitrates=0.5bpp

Bitrates=0.3bpp

Bitrates=0.19bpp

Bitrates=0.11bpp

Bitrates=0.08bpp

Bitrates=0.06bpp

Figure 4.20 Tank Image with different bitrates.

4.3.1.2 Full Parallax Integral Image

For the Full Parallax, the same method is applied apart from the way the viewpoint is extracted. As in Full Parallax the viewpoint is extracted vertically and horizontally [2]. A new proposed N-To-P algorithm had been implemented to three full parallax images. In general, the results of the full parallax image are not very accurate that is due to the way the images are captured. All the full parallax images are suffers from Barrel effects, which makes the viewpoint extraction process is not accurate, and that is will appear from the results shown in figure 4.22.

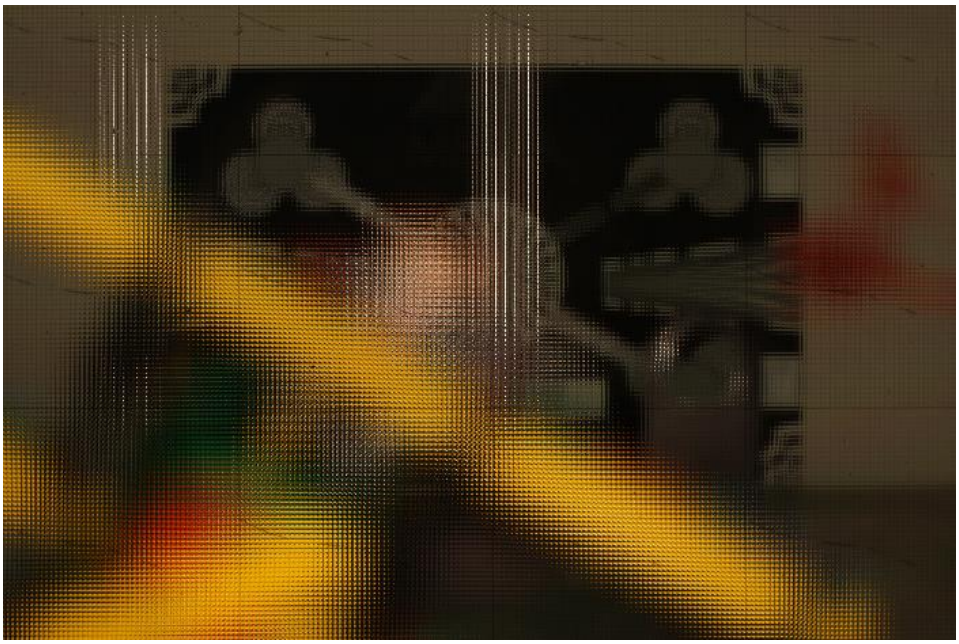


Figure 4.21 Full Parallax Image

However, the image suffers from the barrel effect but still the results in Figure 4.22 shows that using the proposed algorithm achieve better results than applying JPEG on the Omni-directional Integral Image.

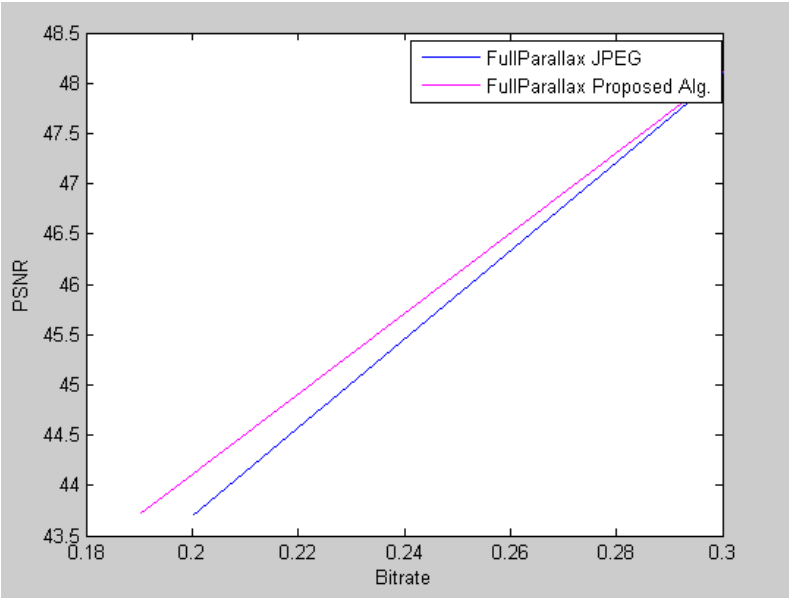


Figure 4.22 Performance of the 3D-DCT and N-TO-P for the full parallax Image

Figure 4.23 represents the rate distortion performance for the 250 micro lens full parallax images, although the image suffers from Barrel effect due to the way its captured but still the N-To-P algorithm proofs its efficiency over the 3D-DCT algorithm.

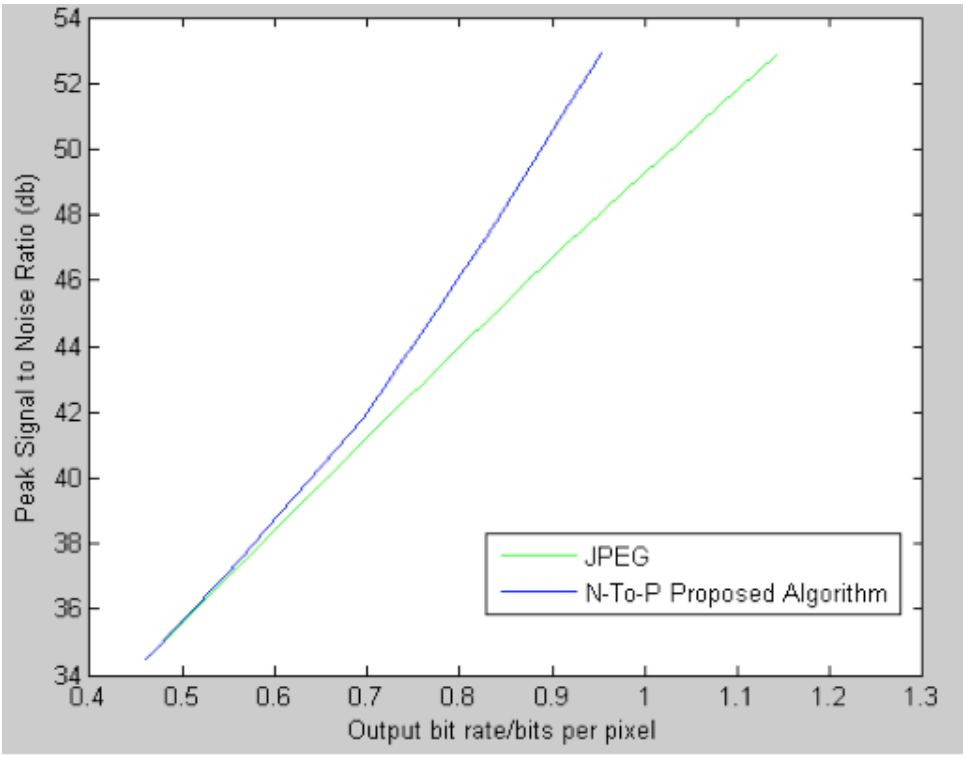


Figure 4.23 Performance of the 3D-DCT and N-TO-P for the 250 micros Lens full parallax Image.

4.3.2 Novel Entropy Coding Technique for 2D-3D DWT based Image Compression

4.3.2.1 Proposed algorithm for 3D Integral Images

The 3D DWT applied on three Integral Images, Horseman, Tank and Micros. A Similar algorithm for to 2D-Images is applied, but a pre-process step need to be applied, which is extracting the viewpoints, then 3D-DWT is applied to each Viewpoint, this step followed by the previous proposed encoding technique, figure 4.24.

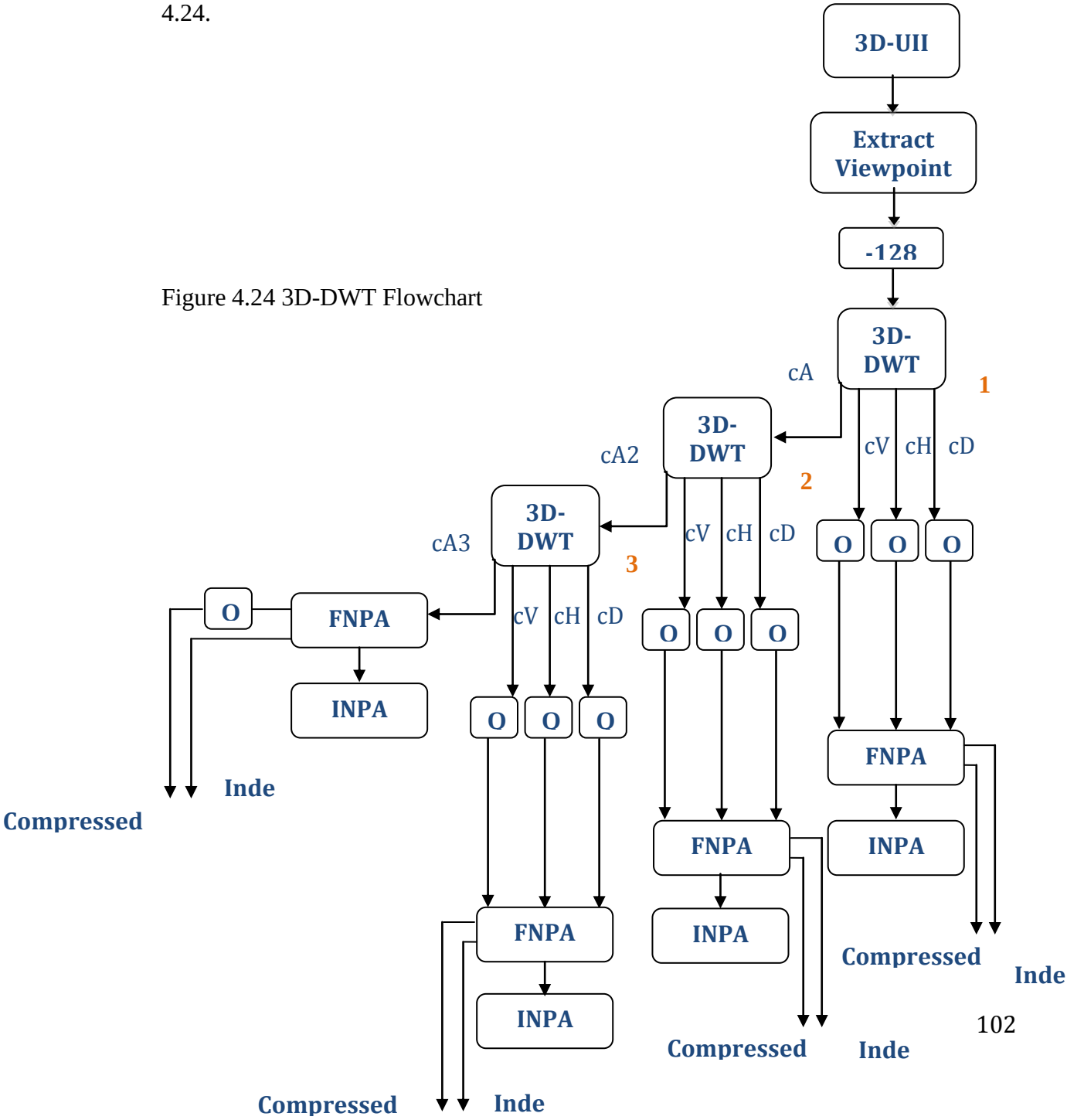


Figure 4.24 3D-DWT Flowchart

4.3.2.2 Comparison between the N-TO-P Algorithm and the Lifting Scheme

In [76], the author applied 3D-DWT on Integral Images using Lifting Scheme. Figure 4.25 represents the difference between the algorithms that proposed in [76] that based on the lifting scheme and the new N-To-P proposed algorithm that is proposed in this chapter. The N-To-P algorithm is using the same Images that have been already used in [76], so that makes the comparison more compatible. The new N-To-P algorithm proofs how efficient it is comparing to the lifting scheme algorithm from the quality aspect and also compression ratio.

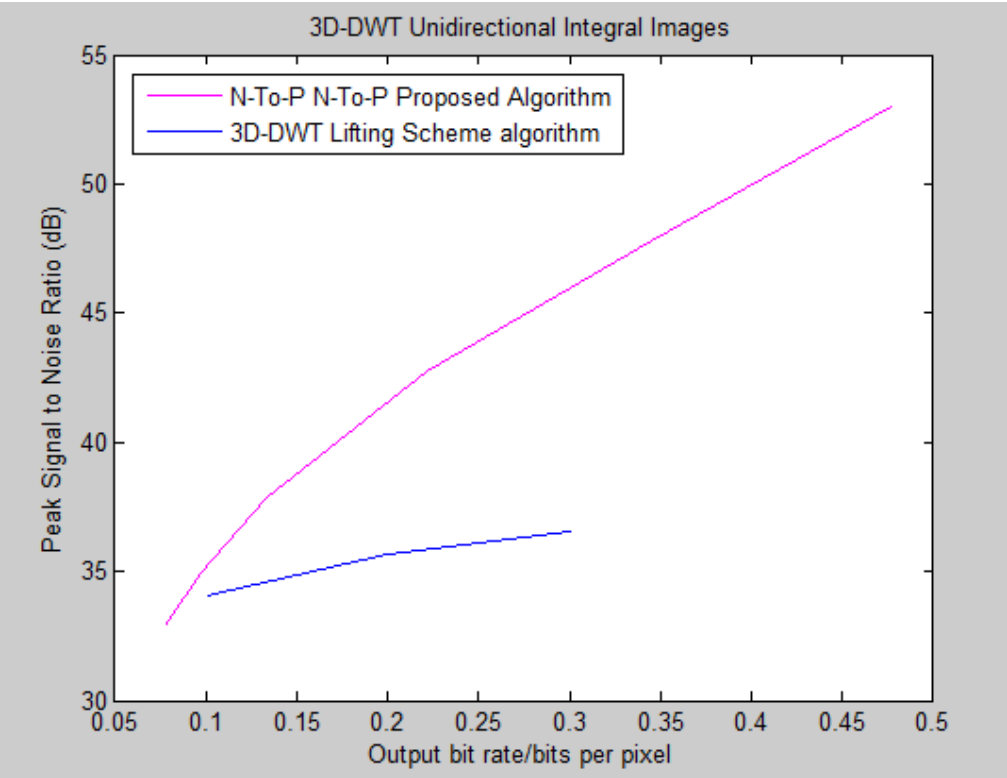


Figure 4.25 comparison between the 3D-DWT proposed algorithm and the 3D-DWT lifting scheme proposed in [76].

4.4 Hybrid DWT-DCT Techniques on Integral Images

In this section, different hybrid techniques built on DWT and DCT is going to be implemented on unidirectional Images. As it has been mentioned before, a lot of Transform coding in Compression standards had been improved in the last few decades, starting from the improves occurred to JPEG in 1992 The Joint Photo-graphic Expert Group, with adopting the Discrete Cosine Transform (DCT), coming to the DWT using Embedded Zero Tree Wavelet coefficients EZW that was established by Shapiro, then the Set Partitioning in Hierarchical coding Techniques (SPIHT) and finally the EBCOT. The most important target in this section is to achieve better compression through minimizing the file size and optimizing the Quality of the reconstructed Image.

JPEG standards is very common standards, actually it is the most common standards, that is due to its simplicity in hard ware implementation and also its energy compaction, but it's still suffering from some disadvantages. Each one of these systems has its advantages that make it better than the others and also the drawbacks that affect the final results. For example, the EZW however, it was developed to get different resolutions and to have high compression ratio, but to reach this number of resolutions levels, more scales needed to be applied, and that increases the computational complexity. Then when, Set Partitioning in Hierarchical coding Techniques (SPIHT) appeared, it approved its ability to get better results than EZW and in less time but its compression level was not that promising level comparing to JPEG2000.

One of the main advantages of hybrid technique between DWT and DCT is that it will solve the blocking problem due to the lack of inter-block correlation due to the localization that coming from using the block scheme. Using blocks however, it solve ringing artefacts that caused by the non-localized scheme, which will make the system suffers from a major problem if it is not been solved, and make the reconstructed image be blurry, but it caused also blocking artefact. The blocking artefacts not only

affect the reconstructed Image, and allow the human eye to notice this sever a discontinuity in the Image, but also will affect the compression process from the point of correlation, the correlation inside each block will be exploit, but the correlation among each block and its neighbour will be missed.

So instead of applying DCT directly and face the problem of no inter blocking correlation, the DWT will be used at the beginning, in order to find the correlation first on the global Image then localize it by using DCT Transform and exploit the energy compaction again to both the inter correlation inside each block after exploiting the correlation in a global manner then also exploit the correlation between the Different Viewpoints with applying third DCT, more information can be found in Appendix C.

Briefly, the DCT gathers all the high frequency (smooth data) details in the low left part and often this is the part that got the lost data from it after quantization, while DWT concept is to gathers all of the similar frequency in a titling way. So if a hybrid scheme done an extreme uses of the correlation will be achieved.

Lots of work had been done using different hybrid transform standards, using two or more transform discrete systems have been used in the last few years. Some researchers combine three systems [77], by using DFT, DCT and HT in order to get higher compression. Others used only two systems as in [78], he implements DWT and DCT to achieve better compression in video coding, and also [79]adopted scalable DWT-DCT on Videos. There are also numbers of researchers that make a comparison between the different transform Coding like [80], [81]. In [82] and [83] a study has been done, on the unidirectional images.

Two-hybrid technique is going to be used:

1) 2D-DWT 3DCT
2) 2D-DWT 3DCT 1D-DWT

In both techniques, comparison will be done between traditional 3D-DCT Huffman algorithm and the proposed algorithm, before, the coding system all coefficients coming from the hybrid schemes will be quantized using fixed quantization factors.

4.4.1 2D-DWT_3D-DCT Hybrid Algorithm

The hybrid proposed algorithm starting by pre-processing process that is containing subtracting 128 and extracting the Viewpoints. All the latter process will be applied to all of the Viewpoints. Then a 2D-FDWT is implemented to each viewpoint, giving four sub-bands, and since the number for our three unidirectional images is 8, so we got 32 sub-bands. The features of these 32 sub-bands are that it is four groups; each group contains 8 sub-bands with the same frequency band but for different viewpoints.

The four main groups are categorized as Low-Low bands (cA), Low-High (cH), High-Low (cV), High-High (cD). The experiment shows that the two bands LL and HL carry the most important details and this can be proofed figure 4.26.

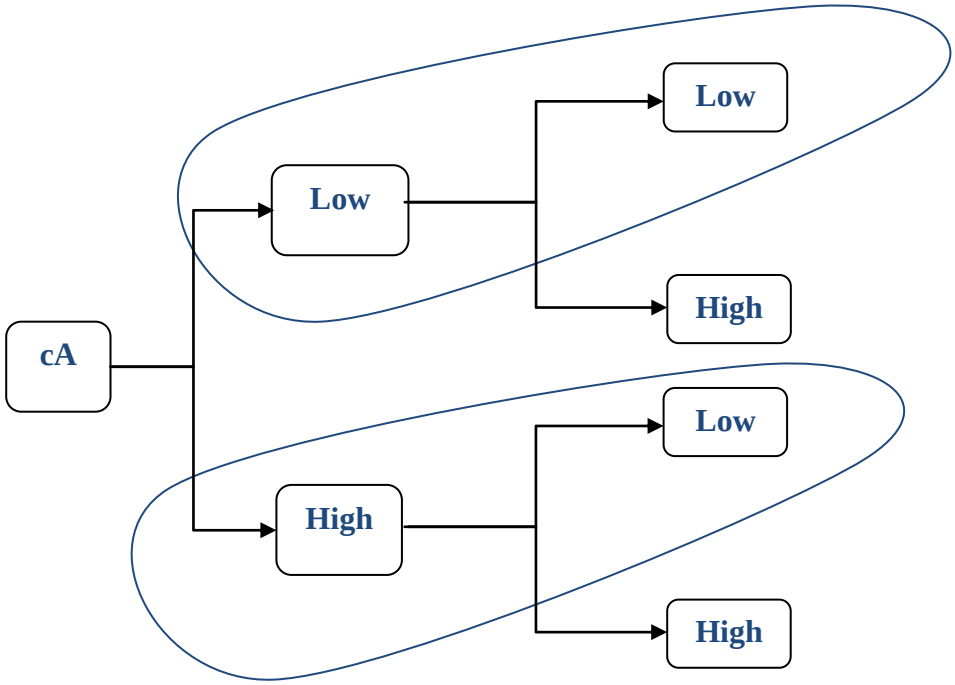


Figure 4.26 The different 2D-DWT sub-bands

So, in order to increase the efficiency of the compression, the 3D-DCT is applied to both Low-Low bands and the group of High-Low bands. So the DCT exploit the correlation inside each block after taking advantage of the global 2D-DWT and

exploit the image for the whole Image, as illustrated in the block diagram in figure 4.27.

4.4.1.1 Block Diagram

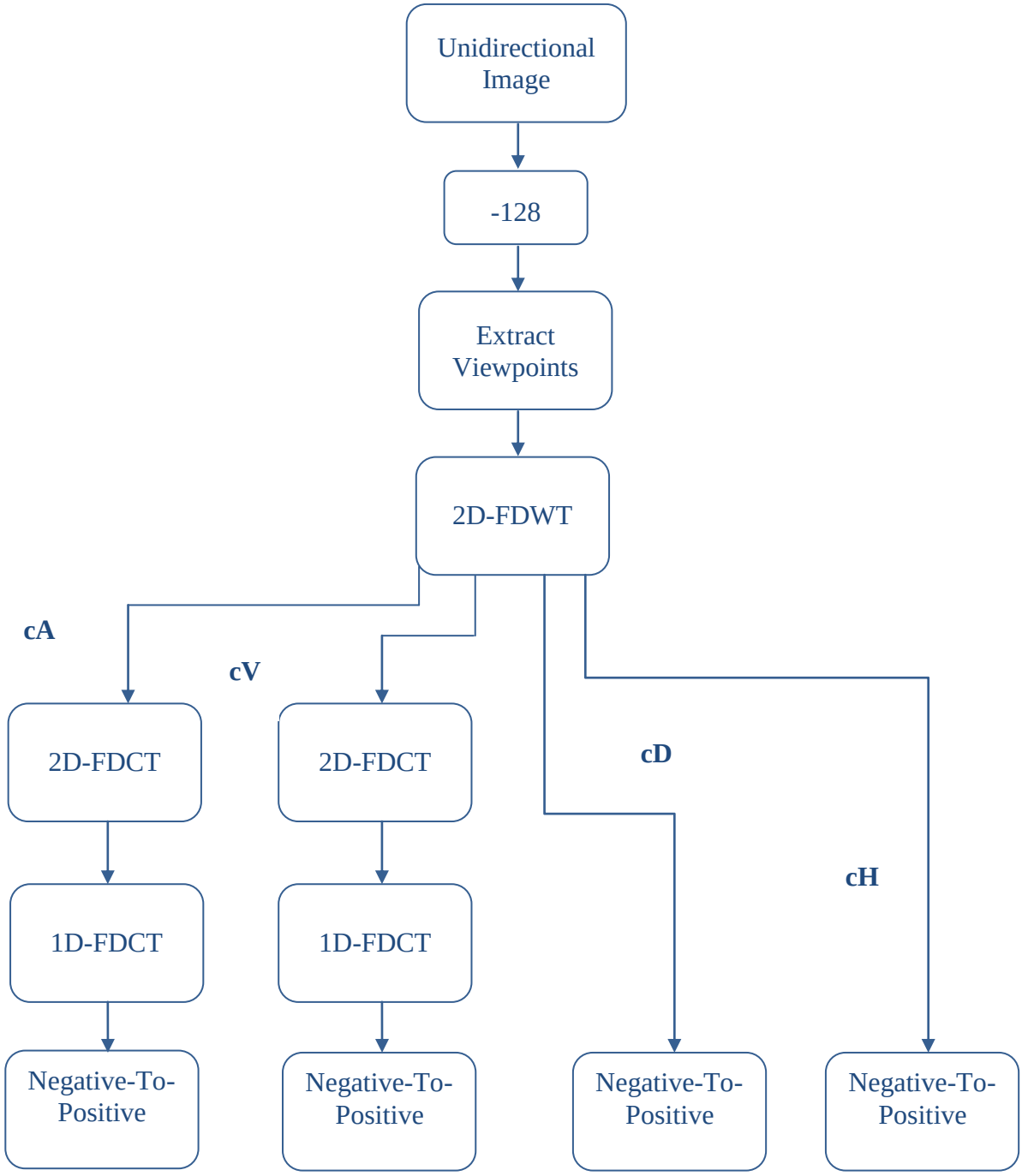


Figure 4.27 2D-DWT 3D-DCT Hybrid Proposed Algorithm

4.4.1.2 Results

Figure 4.28 represents the difference between applying Hybrid 2D-DWT-3D-DCT to the Micro unidirectional Image and applying JPEG with 3D-DCT to the same Image. The average results for the three unidirectional images are represented in figure 4.29, where the hybrid proposed algorithm performs better than the 3D-DCT algorithm with both using the same N-To-P entropy coding algorithm, but the proposed algorithm acts better in the low compression ratio.

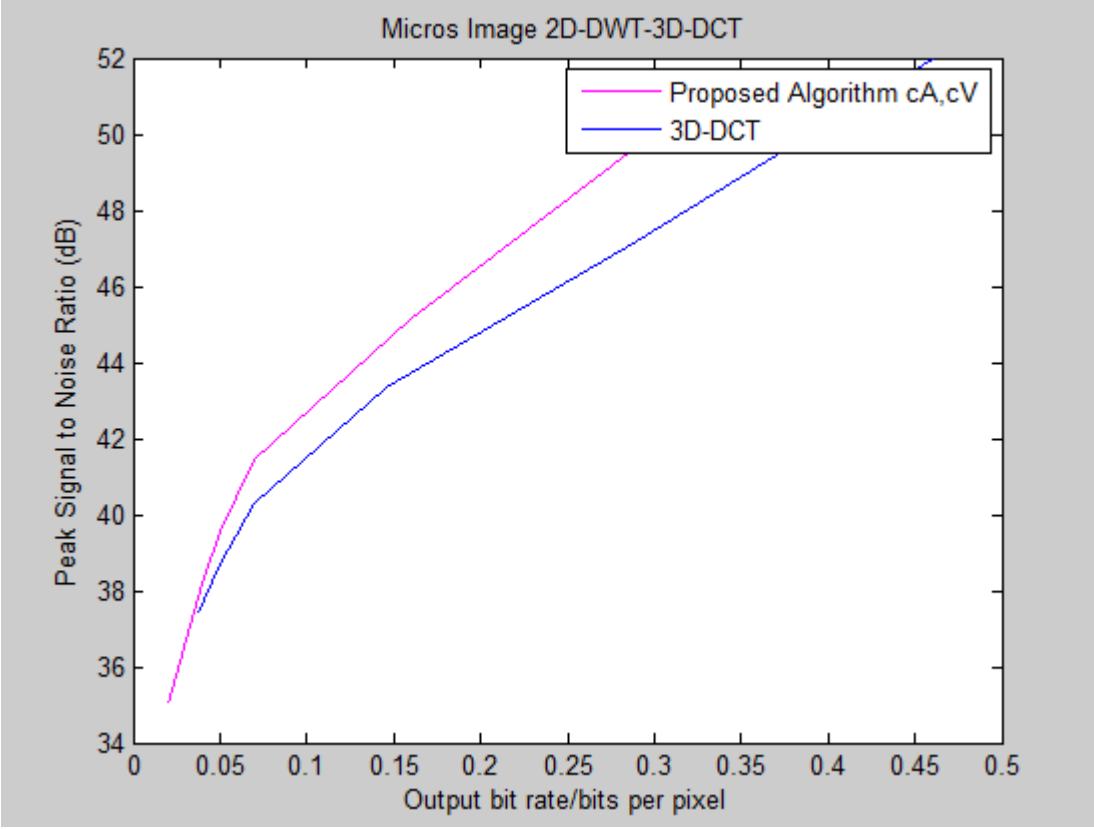


Figure 4.28 Comparison between the proposed algorithm and 3D-DCT DHT algorithm.

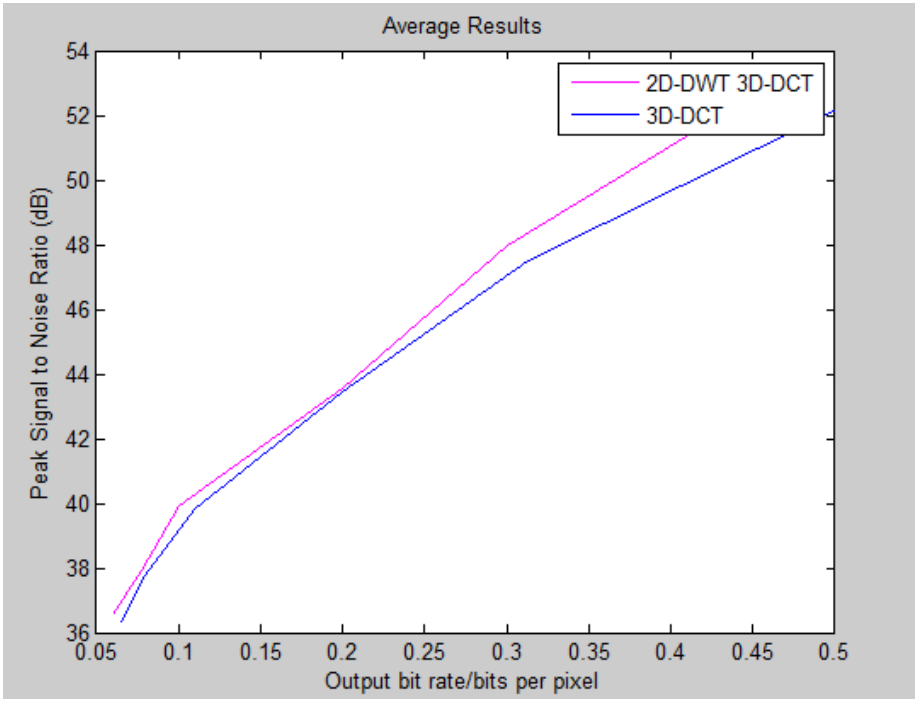


Figure 4.29 Comparison between the averages proposed algorithm vs. the average 3D-DCT.

In [82], the author applied the same steps but first she used a Quantization table scale, then she applied the 3D-DCT only to the Low-Low band only, then she used for this Coefficients Huffman table, while the three rest bands, she used arithmetic coding. Figure 4.30 shows a comparison between the algorithm proposed in [82], and the algorithm proposed here in this chapter.

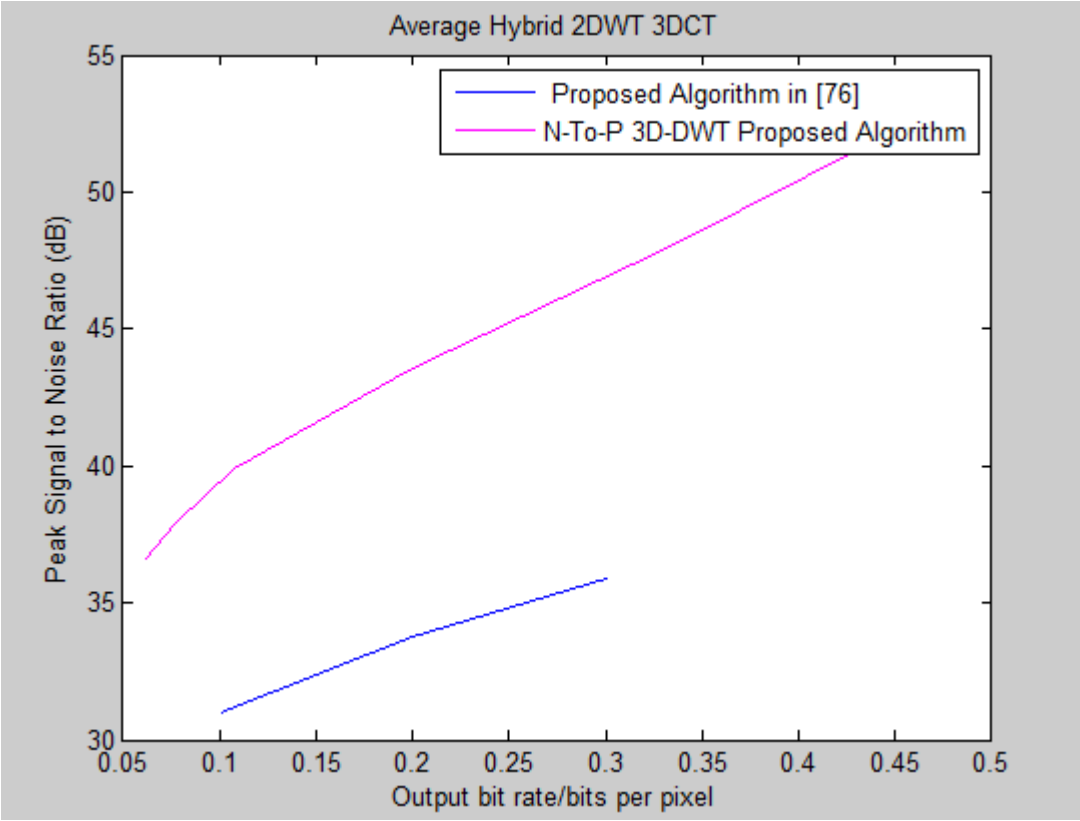


Figure 4.30 Comparison between the proposed algorithm and algorithm applied in [82]

4.4.2 2D-DWT_3D-DCT_1D-DWT Hybrid Algorithm

In this proposed algorithm the same main idea is applied with adding the 1D-DWT after finishing implementing the 3D-DCT, but in this time the coding way is different, as there is no dividing to the coefficients to AC and DC, all of the coefficients is treated in the same way by applying the Negative-Positive algorithm to the Transform coefficients in case of cA (Low Low band) and cV (High Low band) or the quantization coefficients to both cH (Low High band) and cD (High High band). The block diagram of the proposed algorithm is illustrated in figure 4.31.

4.4.2.1 Block Diagram

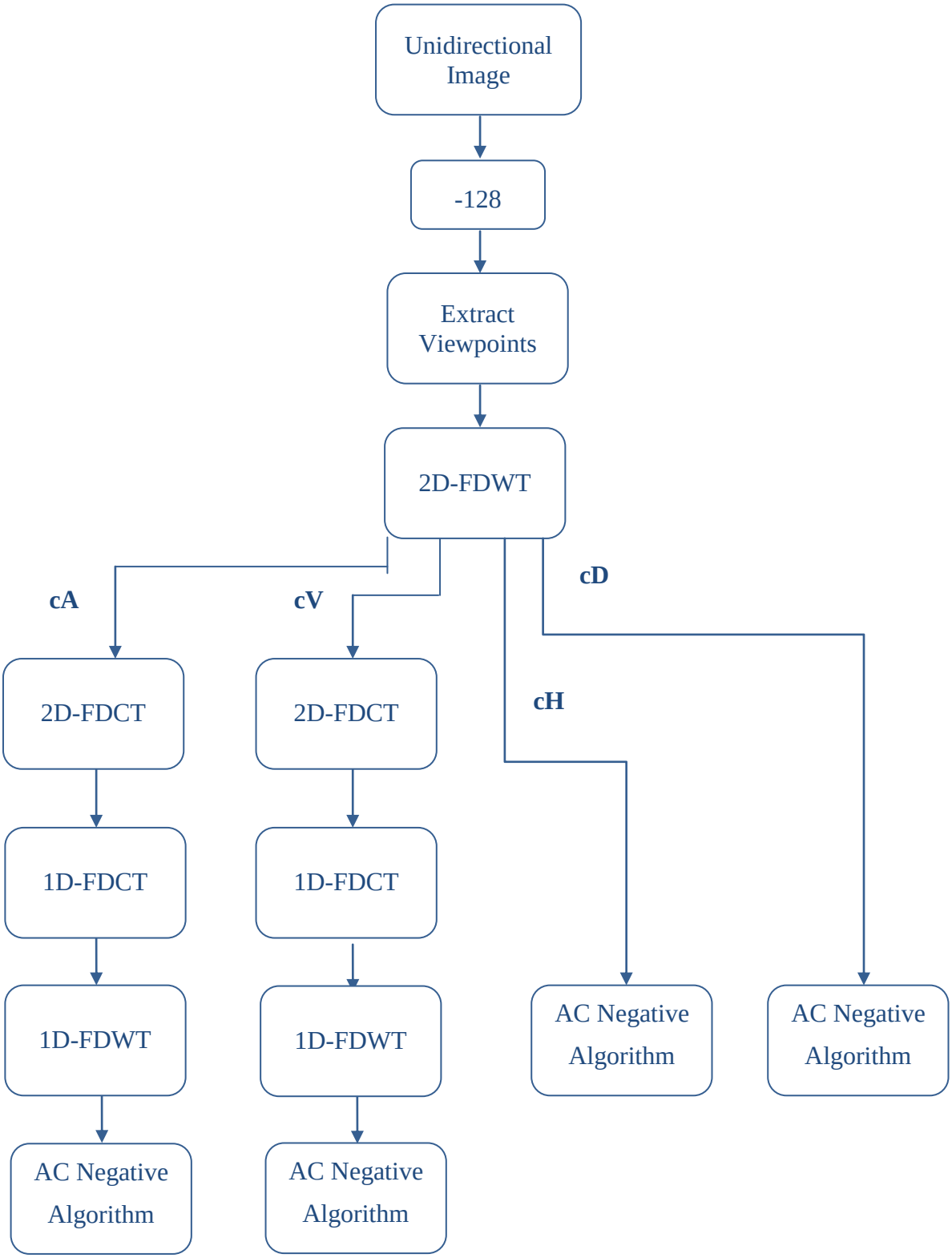


Figure 4.31 2D-DWT 3D-DCT Hybrid Proposed Algorithm

4.4.2.2 Results

From the Tank performance in figure 4.32, the proposed Hybrid algorithm, 2D-DWT-3D-DCT-1D-DWT, proofs its efficiency comparing to the 3D-DCT. Both algorithms used the same N-To-P entropy coding proposed in chapter 3, and also both follow the same way of quantization.

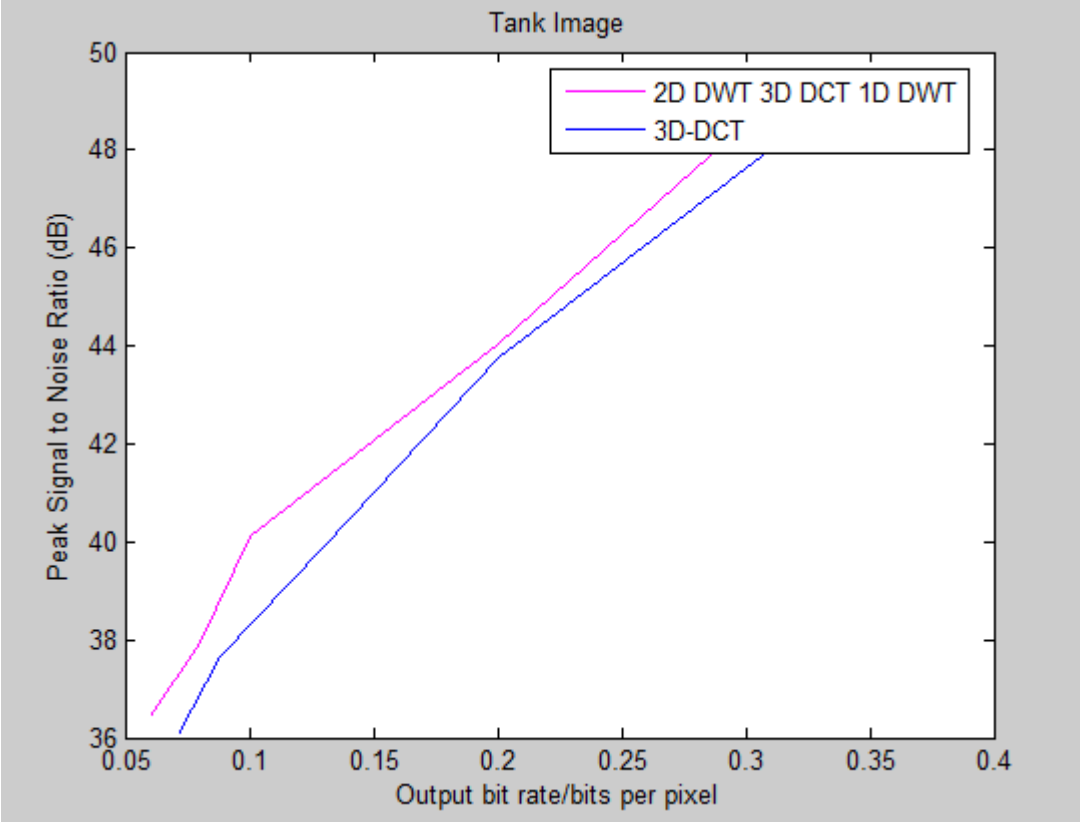


Figure 4.32 Comparison between the proposed algorithm and 3D-DCT.

Figures 4.33 and 4.34, both for the Horse image, the first graph shows the performance of the 2D-DWT-3D-DCT-1D-DWT against the 3D-DCT algorithm, and the second graph represents the performance of the 2D-DWT-3D-DCT against the

3D-DCT algorithm. From both figures, it can be observed that the former algorithm acts better than the latter one.

Also figure 4.35 and 4.36 proofs that there is no loss from the perceptual aspect for both Horse and Micros image, respectively.

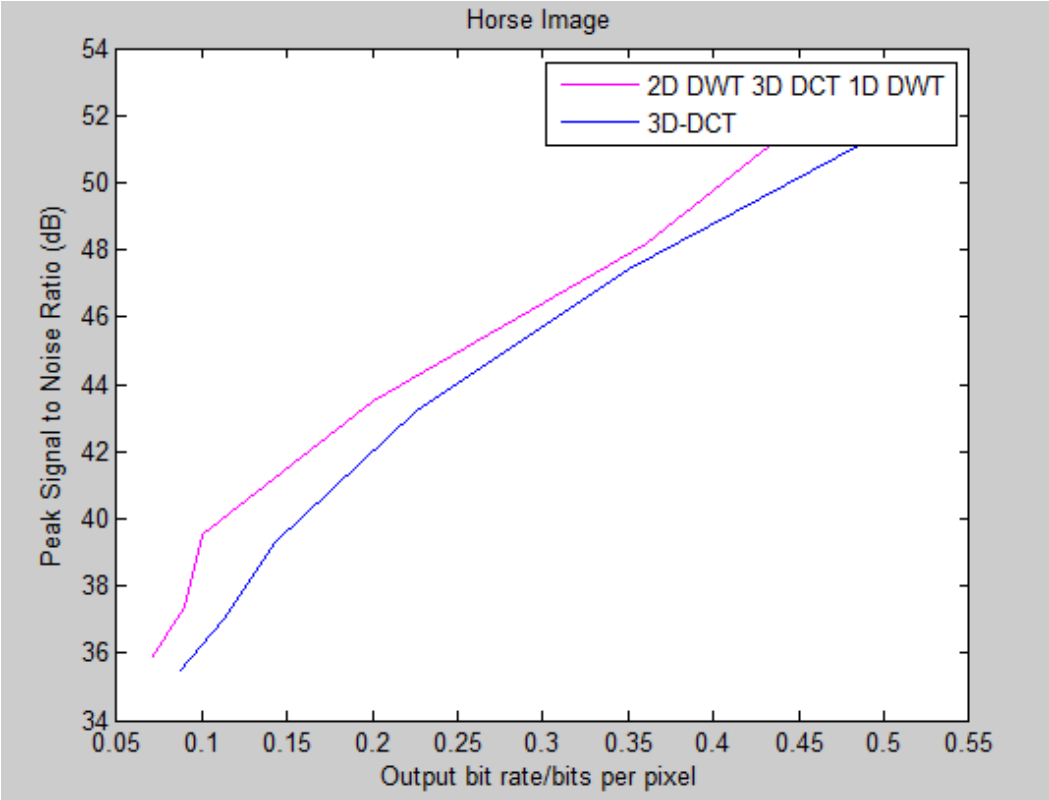


Figure 4.33 Comparison between the 2D DWT 3D-DCT 1D-DWT proposed algorithm and 3D-DCT.

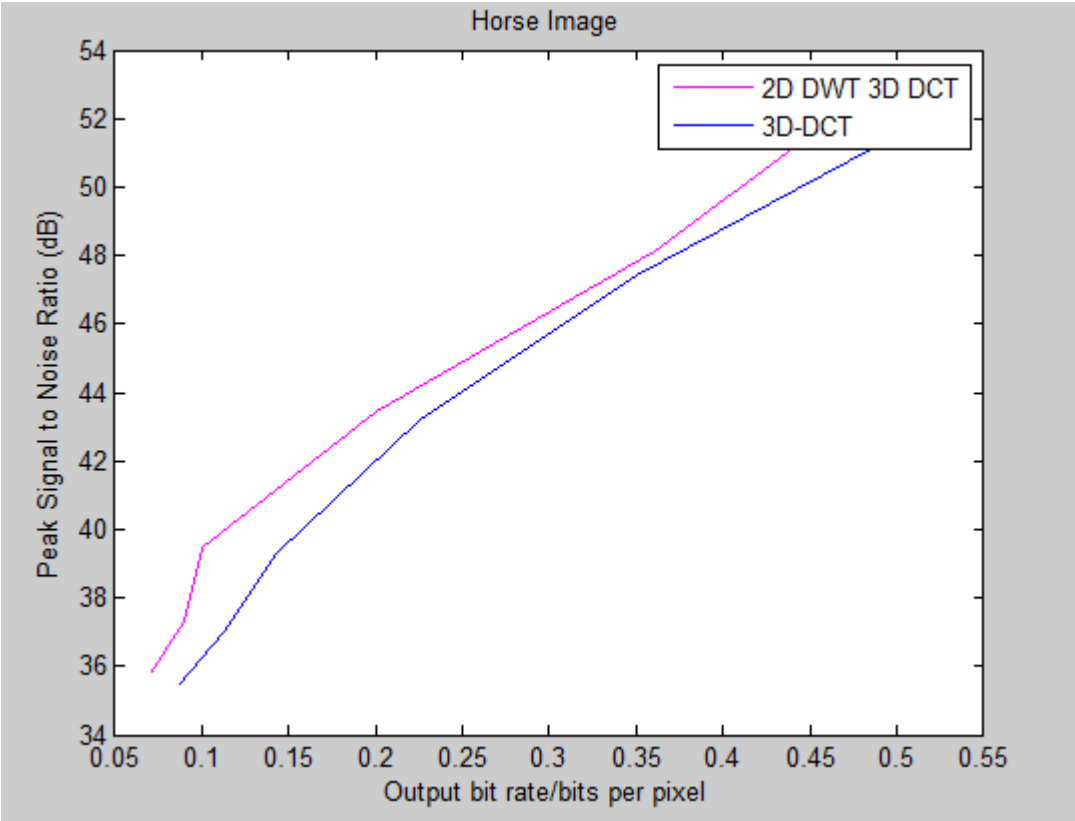


Figure 4.34 Comparison between the 2D-DWT 3D-DCT proposed algorithm and 3D-DCT.

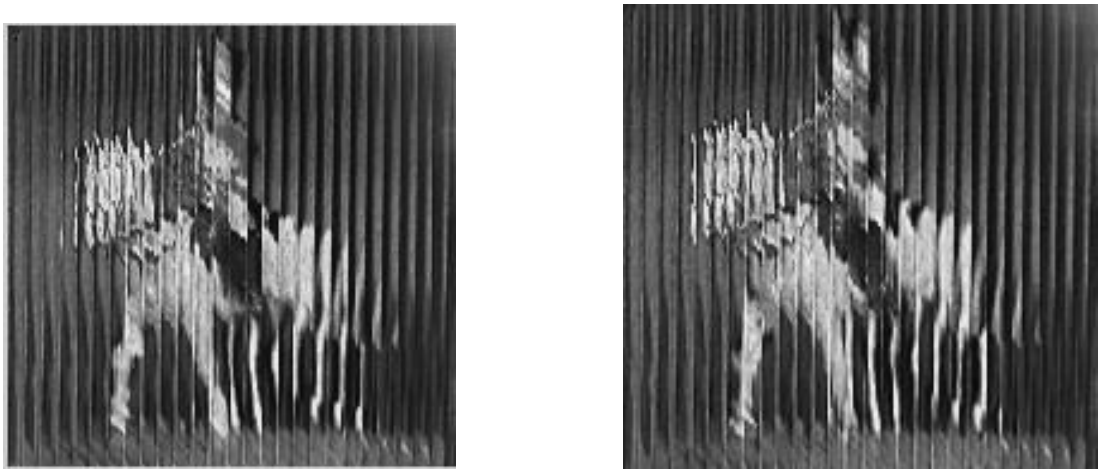


Figure 4.35 (a) Original Horseman Image *(b)* reconstructed Image with Bitrates=0.09bpp 2D-DWT 3D-DCT 1D-DWT

Figure 4.36 (a) Original Micros Image *(b)* reconstructed Image with Bitrates=0.19bpp

2D-DWT 3D-DCT 1D-DWT

4.5 Adaptive 3D-DCT Based compression scheme for Integral Images

The aim of this section is to reduce the artefacts blocking by using a new adaptive algorithm. Most of the compression standards aim to reduce the bitrates in order to store files or transmit it. Standards like JPEG and MPEG used DCT because of its ability to compact energy, There are two methods to implement the DCT, either by applying a global technique via considering the entire image as a one block, or by dividing the whole image into fixed number of blocked that are non-overlapped blocks. The former technique called Global technique and the latter called block technique. Each one of those techniques got their own drawbacks. For example the Global technique suffers from ringing effect as it can be observed from figure 4.37.b and 4.38 b that makes it rarely used. On the other hand the blocking technique, which is the one that is commonly used in association with the DCT, bears from blocking

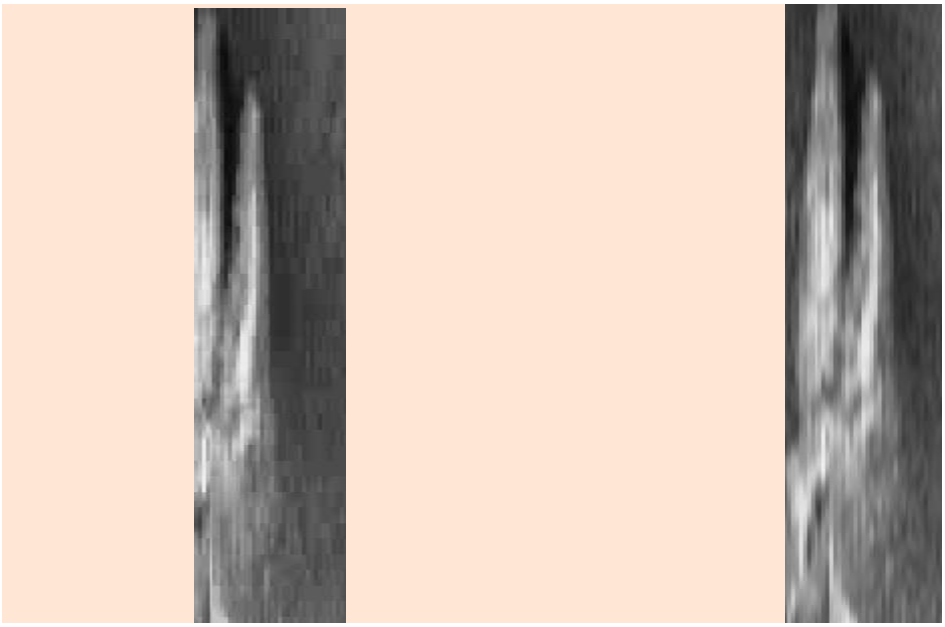
artefacts that can be noticed from figure 4.37.a as a maximizing part of the figure 4.38 .a, another drawback related also to the Blocking technique which is the lack of inter-block correlation which means each block is isolated from the others so each block treated separately and that is leads to the microcosm technique that is applied in DCT, as shown in figure 4.39. A lot of studies have been done in a way to avoid or alleviate the blocking artefacts and attempt to find a solution for the miss correlation between the blocks [84].

On the other side each one of those methods got some privilege, for example the Global method can attain the same compression ratio but with higher image quality but on the account of time and complexity, the same thing occurs with block method with big block size.

In order to exploit the advantage of each technique and avoid the drawbacks to achieve a high compression ratio with high quality, many studies have been done, and they have been classified mainly into two categories. First, improving the Global technique, second, developing a blocking technique, the latter have been shown better results.

4.5.1 Optimizing Blocking technique

The trade between both ringing and blocking artefacts in a blocking technique is still tracing us. So in order to achieve a good compression ratio an efficient blocking technique need to be used.



(a) (b)

Figure 4.37 Horseman Unidirectional Integral Image Viewpoint1, a) Blocking Artefact b) Ringing artefact

(a) *(b)*

Figure 4.38 shows the ringing artefact DCT is applied to the entire Image (a)The original Image (b) the ringing effect

Figure 4.39 The Blocking artefact

As it has mentioned before, The pixels in the top left corner contained the smooth data the maximum horizontal and vertical details in the original image excluding the first pixel value as in gives the average of each block and the low right corners contained the highest frequency values with more diagonal details [85]. To end up with a reasonable compression ratio, the regions that contained more details, a small size of block should be applied otherwise a sever distortion will occur if this details thrown away by applying large block, on the opposite side the regions with less details, a large block is used in order to achieve the desired compression level.

In JPEG a dyadic 8x8 block is used, it has been approved that this size is proper to be used with both level of variations, but because it is periodic the image is not smooth enough due to the blockalization effect. On the other hand, the H.264 video coding standard, the system overcome the ringing effect by applying a small block size 4x4 and avoid the blocking artefact by using a smoothing filter afterwards, making it very sufficient with I-frame.

Window media video 9 (WMV-9) varies the blocking size exploiting the 8x8 block by applying either two horizontal blocks (8x4) or two vertical Blocks (4x8) or four blocks (4x4).

It is clearly obvious that using one block size is less in computational process but still will be also less in the quality of the reconstructed image. It is also quit clear that using adaptive algorithm will be one main way to achieve the trade that is targeted, simply by applying small blocks to the more detailing area (non –stationary regions) and larger blocks to the background or less details area (homogenous regions) [86].

Figure 4.40 represents the rate-distortion performance with applying different block size in 2D while the third dimension is fixed to 8; it shows 2x2, 4x4, 8x8 and 16x16. In this graph a statistical Huffman coding is applied. It appears that the 2x2 block size is worthless to use, that is why we didn't implement our adaptive algorithm with 2x2 sizes, and the later adaptive algorithms didn't reach to level 2x2 block size, when a condition is true the splitting is occurred until it reaches to 4x4 block size and the reason behind it is that choosing a bigger block size that means reducing the coefficients number which leads to increasing the Image efficiency for the same compression rate.

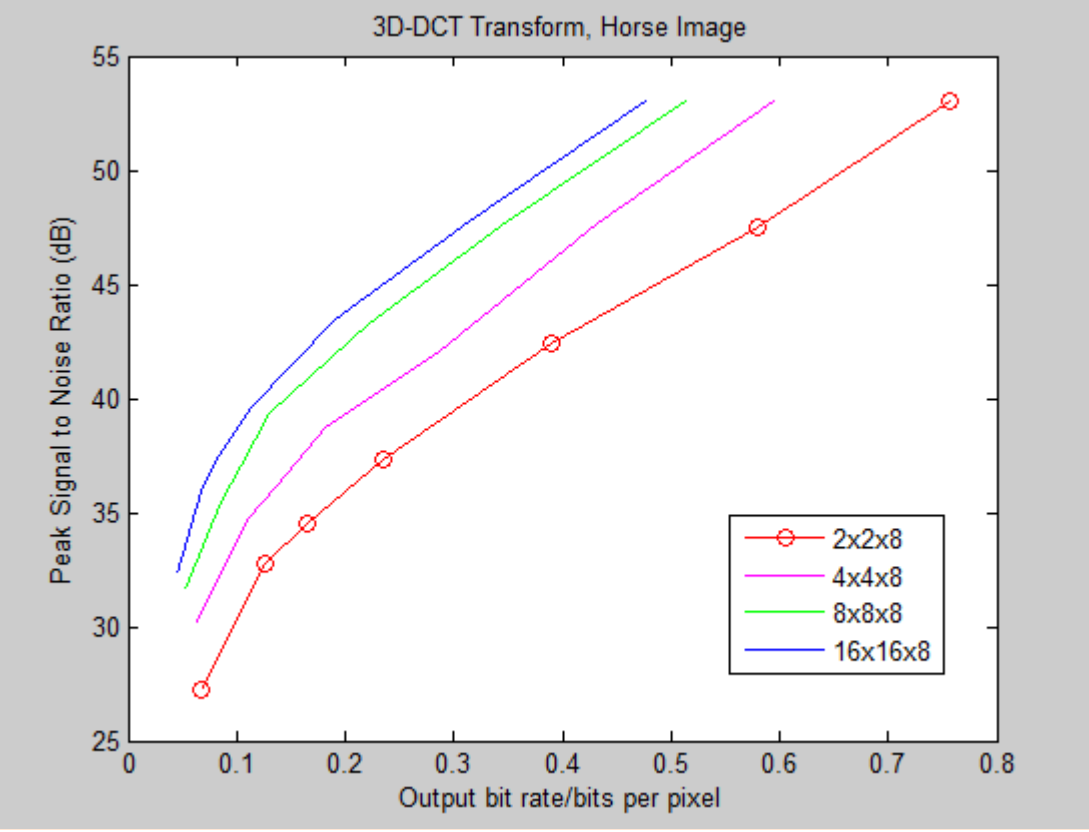


Figure 4.40 different fixed block size (2x2), (4x4), (8x8) and (16x16).

4.5.2 Adaptive Schemes

The adaptive scheme can be categorized mainly in three ways: Using Quad tree, Mean segmentation and Pre-Post filtering Method.

[A1]

4.5.2.1 Quad tree scheme

One of the familiar techniques that are used to represent the Blocks is the quad tree scheme. It adopts the same idea of a tree with parents and children. It is based on the quadratic division, starting from divide the image into four equal blocks, and these four blocks become the children of the original parent Image. The splitting process depends on a threshold or a certain condition that is applied. If the condition is true for one of these blocks then this child block split again into four other children if not then the block will remain with the same size and so on till reach to the end [87] the main advantages of this scheme is its simplicity, the tree with nodes makes it easy to implement, figure 4.41 shows different level of quad tree.

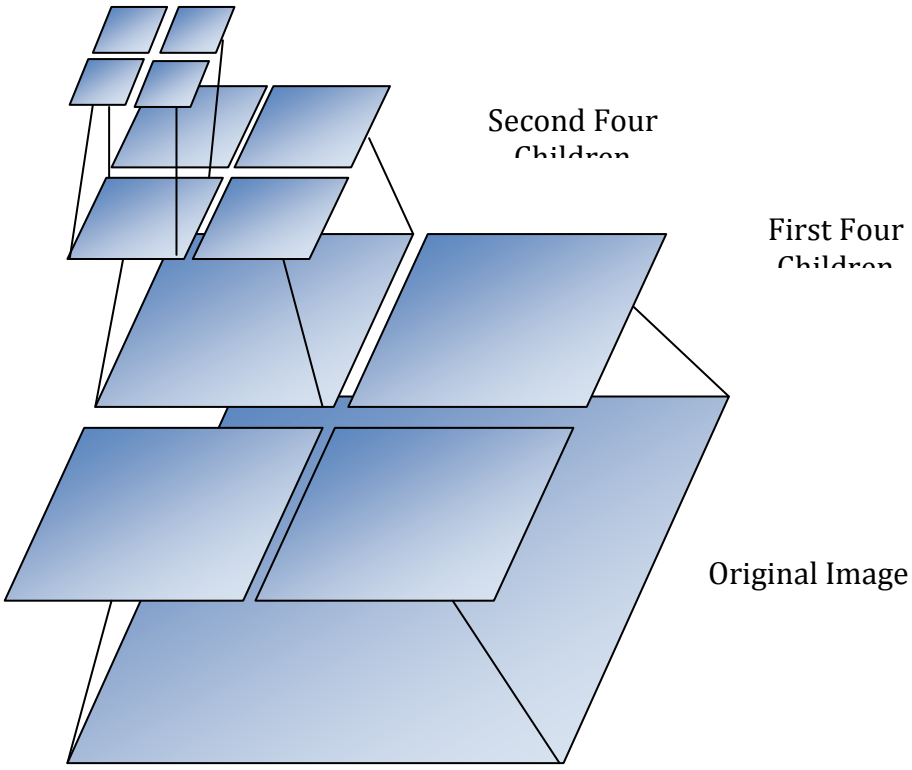


Figure 4.41 Quad Tree.

4.5.2.2 Mean segmentation

In the mean segmentation algorithm, figure 4.42, the main point is the Image is split to a fixed number of blocks as a start as a platform, the size of the block depends on the algorithm, it could be 8x8 or 16x16 or even 32x32.

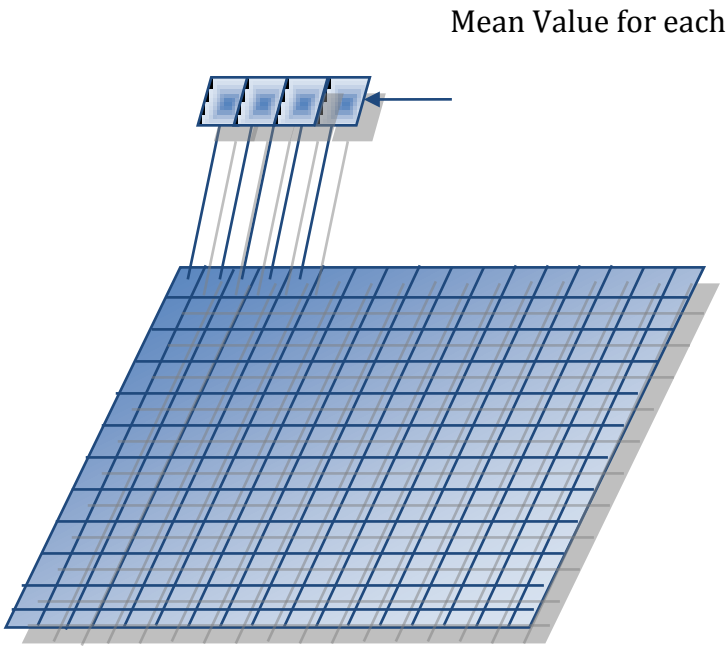


Figure 4.42 Mean Segmentation.

Instead of having 8x8 block; the mean Segmentation algorithm will create a mean planar array containing only one pixel for each block, meaning that each 8x8 block will be represented by only one local mean value [32], from this mean planar platform it can been known how much is the disparity degree in this area (variation Level) [88].

Hybrid methods for the compression of 2D images using both quad tree and mean method have been reported in literature [88]. First, the quad tree is used then the local mean is calculated for each 32x32 block and quantized then subtracted from the original parent block, after that a condition is applied, and the authors found that the threshold should be 80 after a lot of experiments, if the condition is positive then quad tree is applied on the determine block of this local mean, this process is keep running until the block size become 4x4.

4.5.2.3 Pre-Post filtering

Normally, the pre-post Filtering method demolish the original data, the author in [84] applied a lapped transform pre-post filtering to make sure that the data will be inverted correctly [87]. They proofed that the time domain pre and post processing for both Forward DCT and Inverse DCT respectively, can generate a large class of Lapped transform with a uniformed number of channel and overlapping samples. The two process Pre-filtering and the post-filtering are used to make all input DCT blocks homogeneous and alleviate the blocking artefact in the output. Their algorithm show a little bit improve in the coding performance and also less complexity.

In [86], a Pre-/post- filtering scheme to mitigate the blocking artefact and also increase the efficiency was presented. The process is divided into two levels similar to approach reported in [84]. First the DCT energy compactness is improved by using Pre-filter in order to reduce the correlation between the blocks, then the blocking artefact is eliminated by using Post-filter to smooth the reconstructed data. The technique is separable like DCT, and gives better results that JPEG2000 and H.264 I frame coding. However it has one drawback, which is that, the high complexity comparing to the JPEG2000 and lower quality comparing to H.264.

.

4.5.3 Proposed Adaptive Algorithms

In order to improve the compression ratio and mitigate the blocking artefact with also limit the ringing effect, two algorithms are proposed and applied on the 3D Unidirectional Image. First the Adaptive Binary Mapping algorithm and the second is the Adaptive Mean Algorithm.

4.5.3.1 Binary Mapping Proposed Algorithm

It is pre-processing algorithm, its notion based on using filters to smooth the binary image represented from the original image. The algorithm illustrated in figure 4.43 is implemented on the Viewpoints extracted from the Unidirectional Integral Image.

The viewpoints are converted to zeros and ones as a binary map with keeping the original viewpoints, for each viewpoint both the viewpoint and its map is going to be used. First, number of filters is applied on the map array in order to make it smoother, and then the technical adaptive algorithm will start depending on those maps as input and applying the results on the actual viewpoints.

Both Viewpoints and the Binary maps are going to be split to 16x16 blocks, and if condition is applied using a flag, investigating whether the binary map 16x16 block is all zeros or all ones. If flag is 1 which indicates true that means that this area is smooth region and doesn't have lots of details which leads to apply the 3D-DCT on the whole 16x16 block in the determined Viewpoint, otherwise if the flag was 0 this is an indicator that there is a kind of variation and details in this region, meaning that this area needed to be split again into four 8x8 blocks, the same process is applied to all blocks, the blocks will only be reached to 4x4 levels not more, from other studies it has been shown that going to 2x2 block level doesn't worth it.

This Algorithm is a hybrid algorithm, using the quad tree method with the pre-post pre-processing filtering, and the binary maps size is nearly negligible and it is also simple. The Negative to positive encoding algorithm is applied with the adaptive End of Block and also the Difference Algorithm for the index array.

4.5.3.1.1 Binary Mapping Algorithm

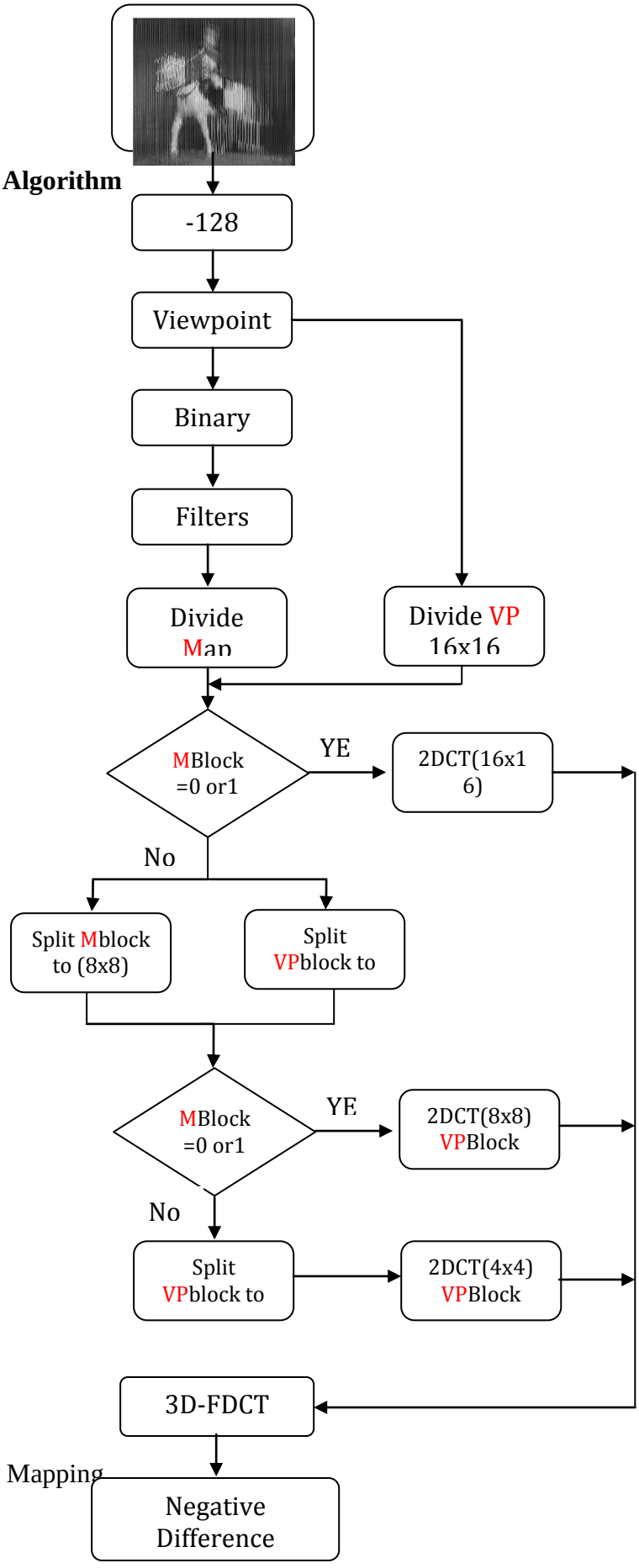


Figure 4.43 Adaptive Binary Mapping proposed algorithm.

4.5.3.1.2 Results

The adaptive Binary Mapping algorithm didn't give the results that were expected that leads to look for another method that can give better performance with high RD. Figure 4.44 and figure 4.45 show how bad the performance of the Binary Mapping Adaptive technique comparing to the 3D-DCT that is proposed former in this chapter. Figure 4.44 is representing the algorithm implemented on the Micros Image and figure 4.45 for the Tank unidirectional Image. This bad performance is due to the number of filters that have been used in the pre-processing process of the Adaptive Binary Map algorithm. Although these filters are changing the input data to make it smoother and make it easier for the adaptive algorithm to process but it is irreversible process, so it is lossy which make it gives worse than the 3D-DCT algorithm that proposed previously in this chapter.

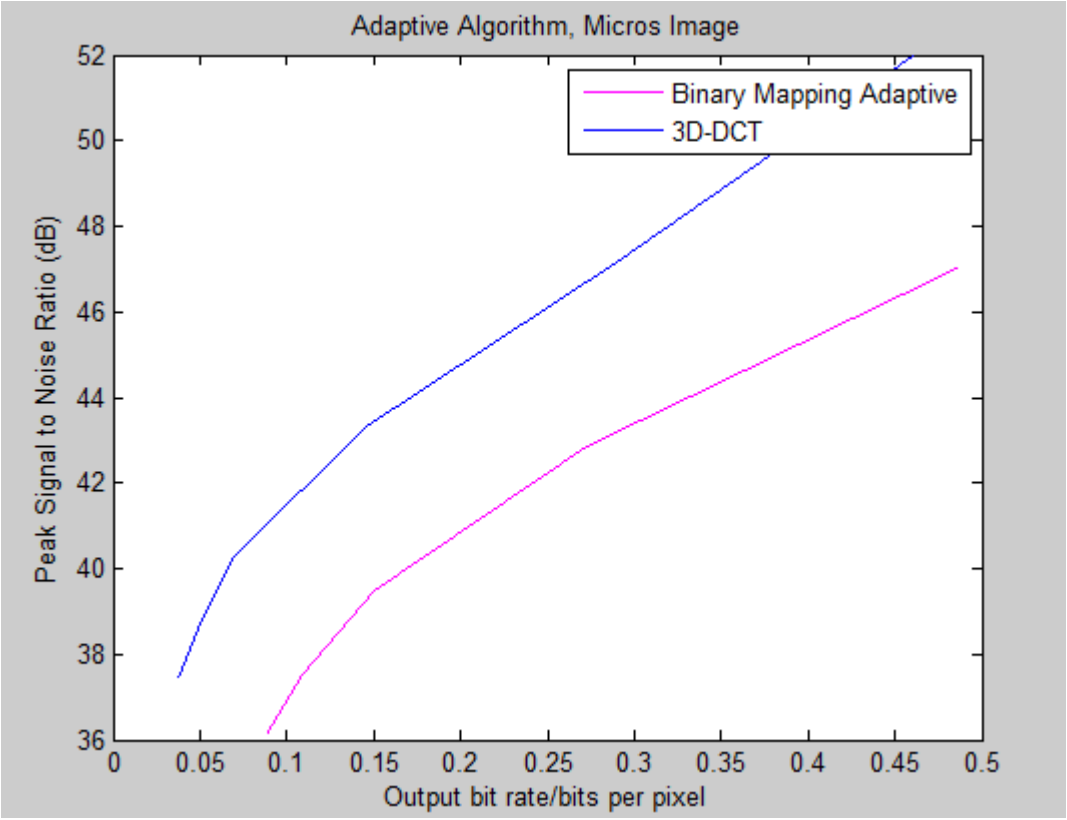


Figure 4.44 RD Performances for Micros Image

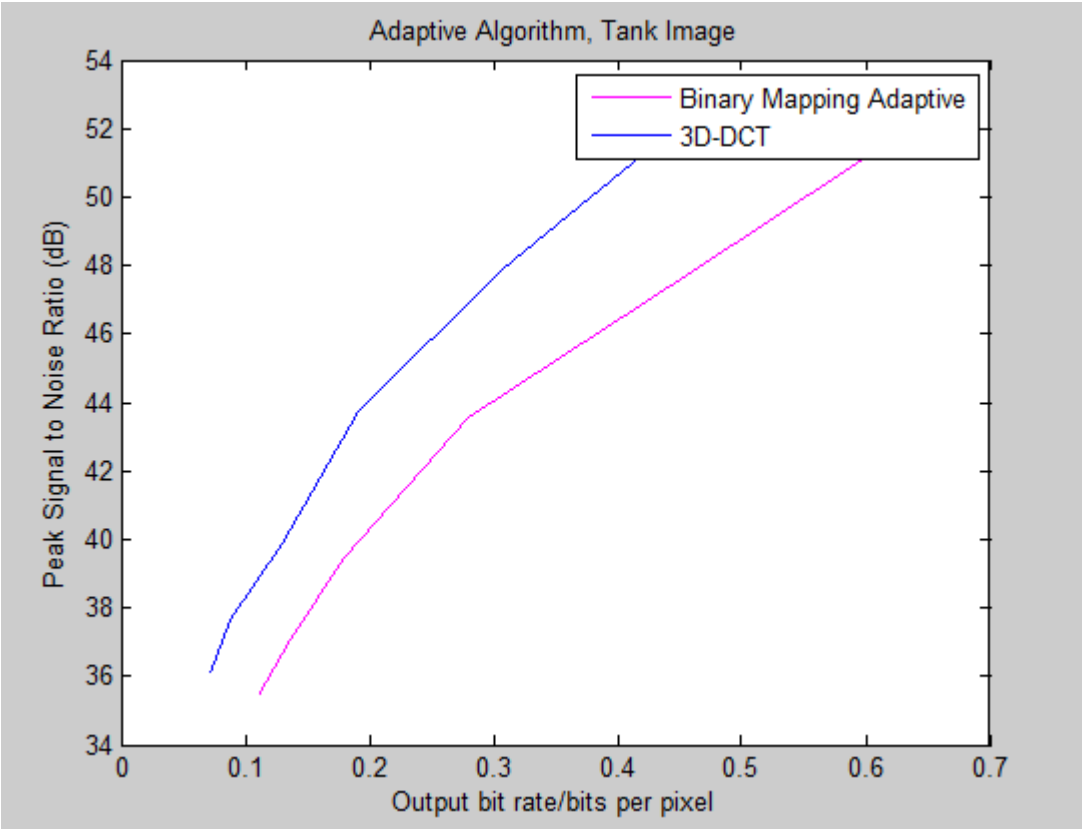


Figure 4.45 RD Performances for Tank Image

Figures 4.46 and 4.47 show the original and reconstructed images for both the Micros and the Tank unidirectional images using the proposed binary mapping adaptive algorithm.

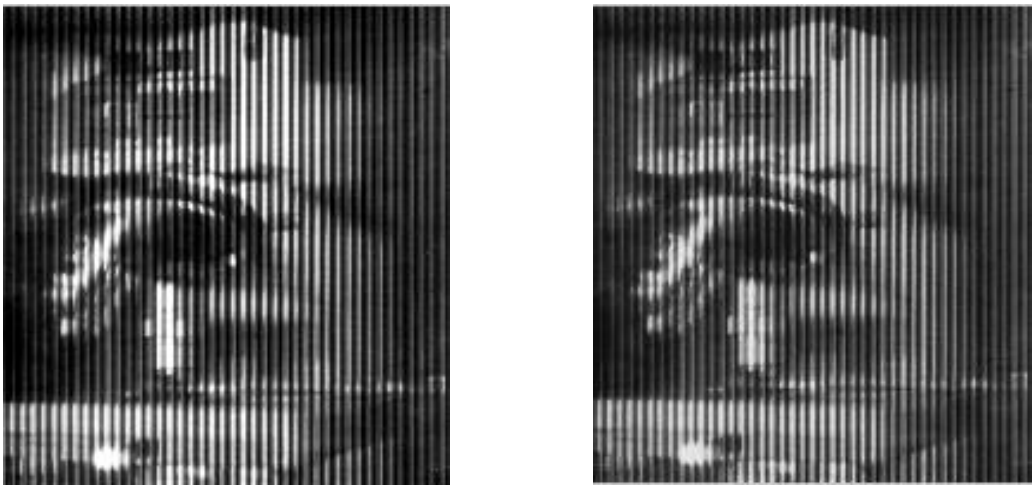


Figure 4.46 a) The original Micros Image b) The Adaptive Binary Mapping Reconstructed Image with Q=32 using Negative Difference Coding Algorithm

Figure 4.47 a) The original Micros Image b) The Adaptive Binary Mapping Reconstructed Image with Q=32 using Negative Difference Coding Algorithm

4.5.3.2 Mean Algorithm

In the mean algorithm, the viewpoints split into 16x16 and a mean value is calculated for the first viewpoint, then for each 16x16 block of all Viewpoints, the mean value is calculated. Figure 4.48 shows the Mean Adaptive algorithm block diagram.

4.5.3.2.1 Mean Adaptive Proposed Algorithm

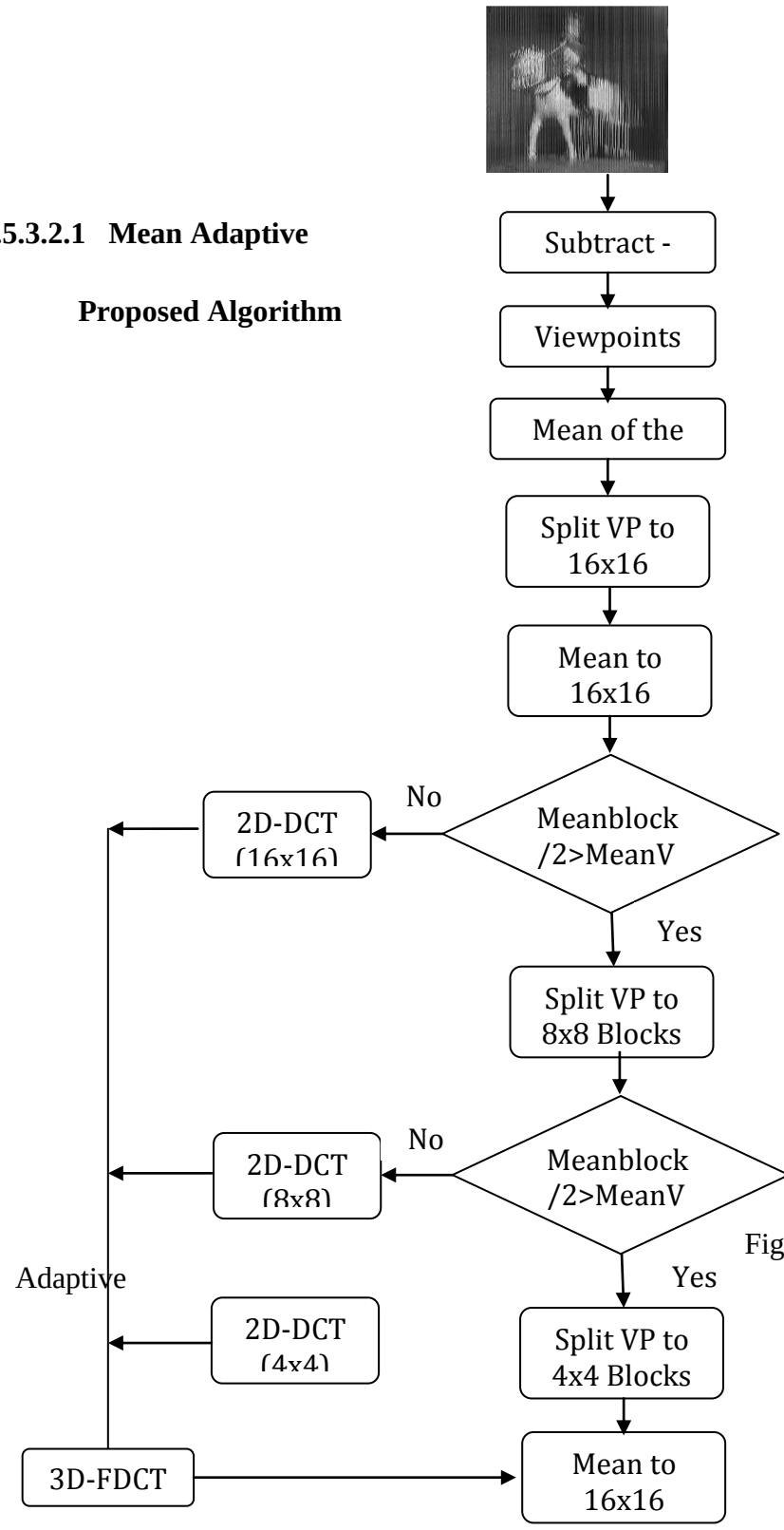


Figure 4.48 Mean Adaptive proposed Algorithm

Following this, an IF condition is applied in order to take a decision if another split is going to occur or not. The condition is a comparison between the local mean value of the block and the mean value for the entire Viewpoint, so this non-local mean value is act as the threshold that is the splitting scheme depends on it. Later another 1D-DCT with fixed size of eight is applied on the third dimension to perform a 3D-DCT as a final step.

4.5.3.2.2 Results

Figure 4.49 shows the difference between applying the binary Mapping algorithm, Mean algorithm and also the JPEG standard, applied on Horse Image. The big difference between both the Mean Adaptive and 3D-DCT algorithm and the Binary adaptive is really big and this is due to the reasons that is explained in the previous section. Another comparison is between the performance of the Mean adaptive and the 3D-DCT algorithm, figure 4.49 and also figure 4.50 that present the average results show that the Mean Adaptive is working more efficiently than the 3D-DCT.

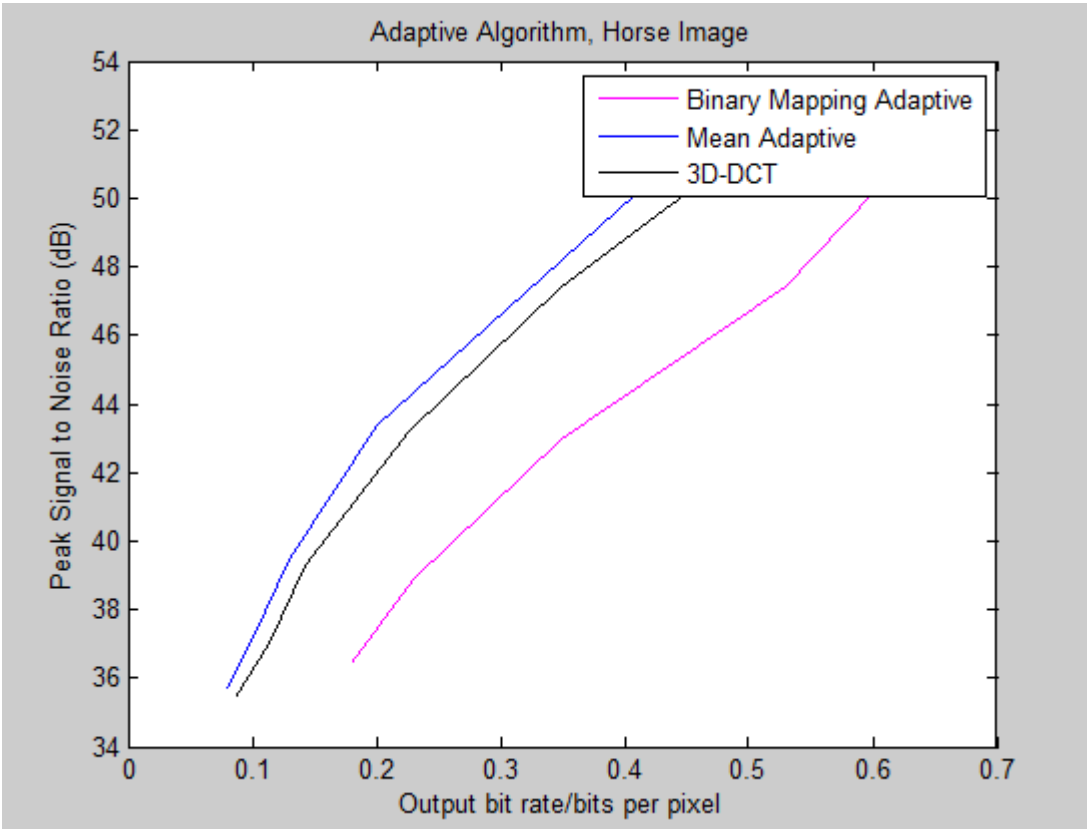


Figure 4.49 The Horse Image PSNR vs. bit rates.

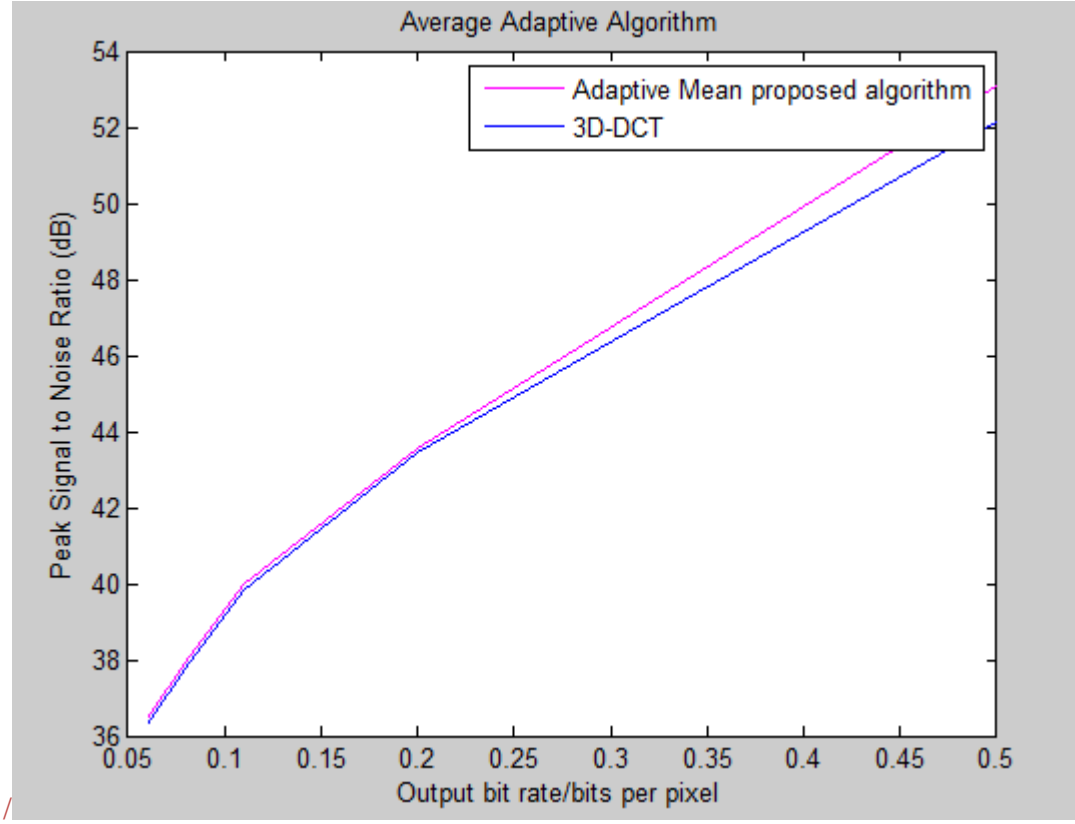


Figure 4.50 Average result PSNR vs. Bitrates between 3D-DCT adaptive Mean algorithm and 3D DCT (8x8x8) block.

In figure 4.51, different 3D-DCT-block size is implemented on Micros image, starting from 2x2 and ending with 16x16 for 2D-DCT with using fixed size 8 for the third Dimension. This graph represented to show the difference in performance between the different 2D size block and the Mean adaptive 2D size block with fixed size of the third dimension for both categories. Indeed, using bigger size block size is better because it leads to decrease in the number of coefficient, but still using a big block size makes the chances of facing a blocking artefact more possible to occur.

As has been observed from figure 4.50 that the adaptive mean algorithm perform better, especially at low compression ratio, comparing to 3D-DCT with 8x8 block size but from figure 4.51 it is clear that over the range evaluated the 16x16 block size afford best results over the other ranges even the adaptive.

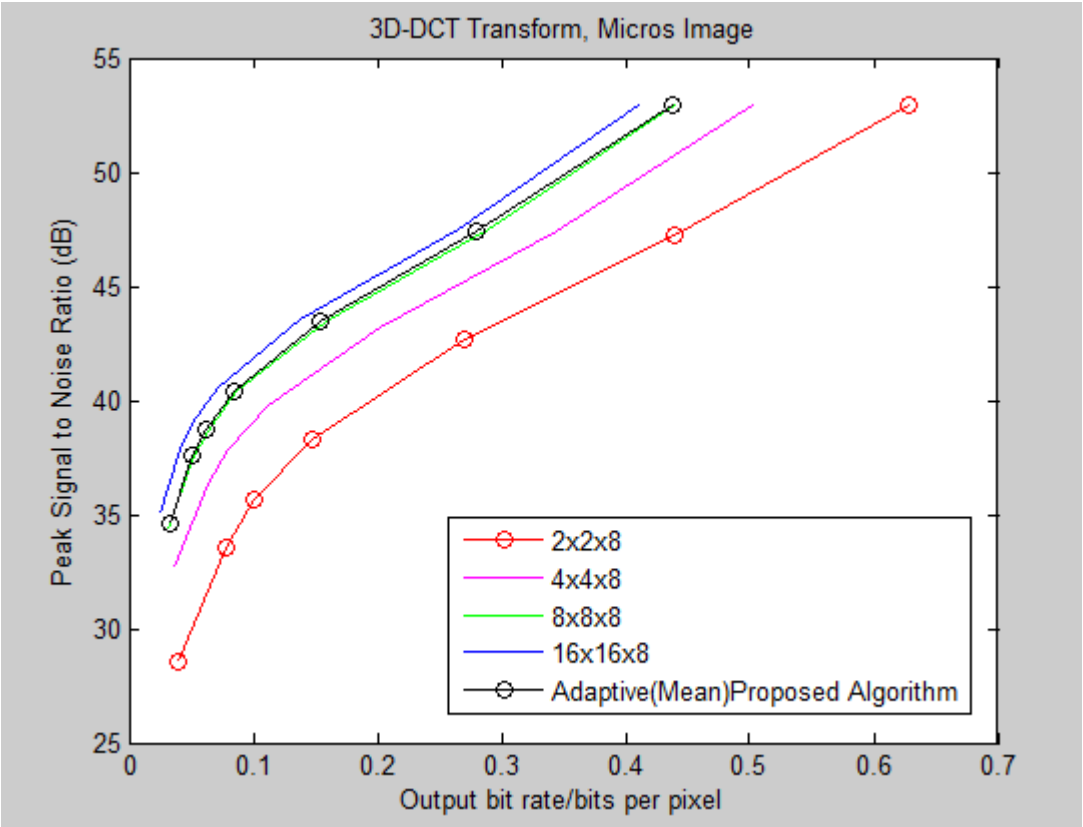


Figure 4.51 Micros Image using Statistical Huffman & 3D-DCT adaptive Mean algorithm.

Figures 4.52 and 8.53 show the original and reconstructed images for both the Micros and the Tank unidirectional images using the proposed mean adaptive algorithm.

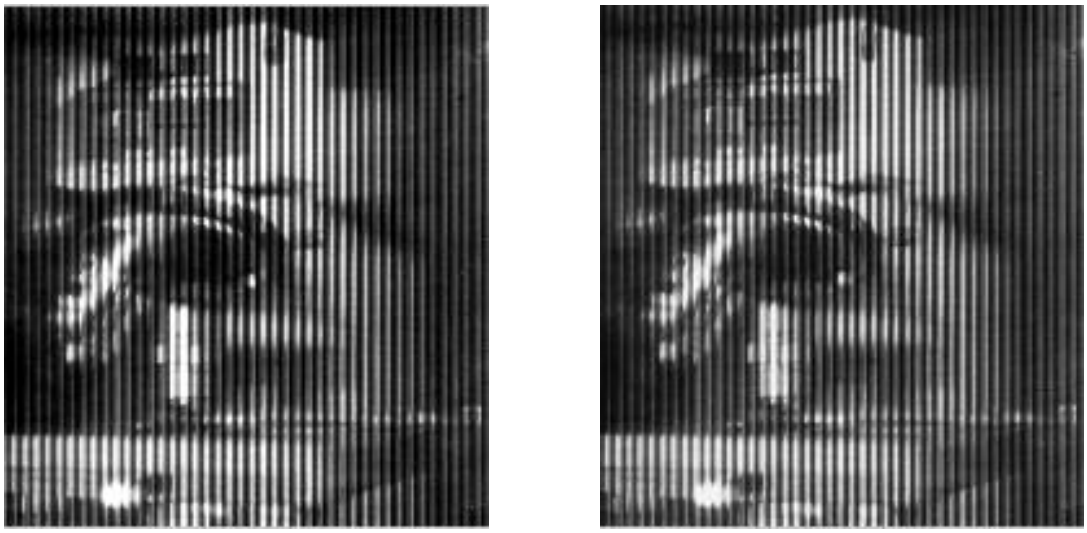


Figure 4.52 a) the Original Micro Image b) the Adaptive Mean reconstructed Image Q=16

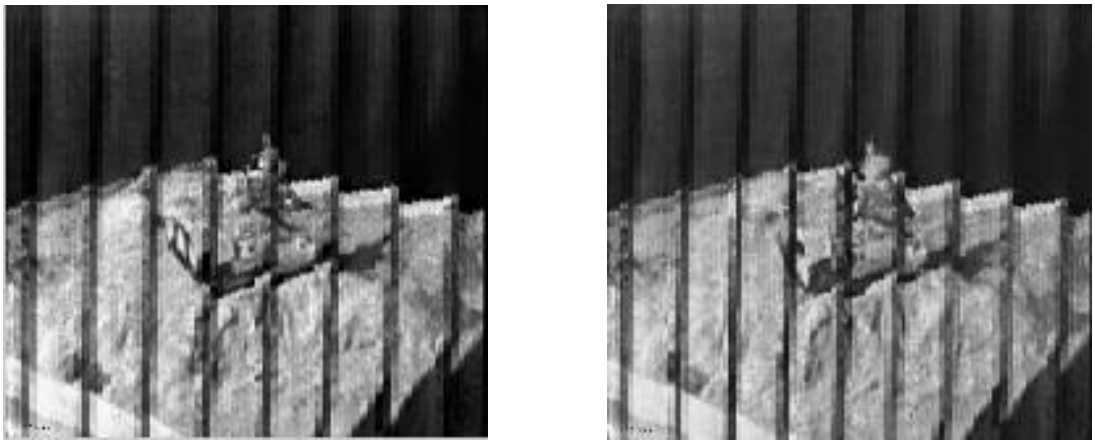


Figure 4.53 a) the Original Tank Image b) the Adaptive Mean reconstructed Image Q=32

4.5.3.2.3 Comparison between the algorithm in represented [32] and the proposed algorithm

In [32] the author adopted the idea of dividing the Unidirectional Integral to 8x8 blocks, this process is applied to the entire source image without extracting the viewpoints. Then there proposed algorithm is based on the Mean Adaptive scheme, for each block, she calculate the mean value and put all together to form a planar mean image that is create low resolution image, then Horse man image is segmented and finally a split-merge algorithm is applied. It is important to mention that Ramona's paper can be considered that it is the only paper that applied adaptive algorithm on a 3D Integral Content, in addition it is using the same images that have been used in this chapter. So the comparison between her proposed algorithm and the new algorithm that is represented here is really vital. Table 4.1 represents the Bitrates and the PSNR different values for both proposed algorithm and the algorithm proposed in [32].

Proposed algorithm in [32]		**Proposed algorithm**	
Bitrates	**PSNR**	**Bitrates**	**PSNR**
0.6	39.33	0.55	53.0964
0.4	38.22	0.19	43.57
0.2	34.77	0.06	36.46

Table 4.1 Comparison between the algorithm in represented [32] and the proposed algorithm.

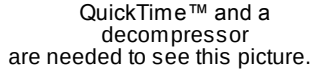


Figure 4.54 comparison between adaptive proposed algorithm in [32] and proposed algorithm

4.6 Conclusion

Chapter 4 came through different aspects in order to improve the 3D-Integral process. Different methods and algorithms have been using to optimize the compression technique. The chapter is divided into four sections; each section includes a proposal for one algorithm or more.

The chapter started by investigating the best way to implement the Huffman coding. A comparison between Statistical Huffman (SH) coding and Define Huffman Table (DHT) coding have been applied on the 3D-Unidirectional Integral Images.

The results show that the DHT saves more time than the SH. The SH is performing better than DFH from the compression aspects but without statistical Huffman dictionary, which cannot be occur. However, that in the SH coding there is no need to

send a define table in the header, but still the dictionary must be send to decode the data sent.

The work represented in the second section is the main core work of this chapter; it is the effect of the N-To-P entropy-coding algorithm on the 3D-Integral images with both of its types, the five unidirectional images and the three full parallax images. First, it implemented with the 3D-DCT techniques, and from the results that shown, the proposed entropy coding proofs it's efficiency regarding compression ratio, image quality and time. So, by applying this new proposed algorithm, the compromises between the three factors have been achieved. The full parallax images can be improved better if a correction of the Barrel effect has been solved.

The second part of this section is the performance of the N-To-P entropy-coding algorithm in association with the 3D-DWT technique, it is characterized by its simplicity, but it showed a very good results comparing to the proposing lifting scheme in [76].

From this stage, and due to the good performance of the new proposed N-To-P entropy-coding algorithm, it had been used to all of the following proposed algorithms as the entropy coding block, even it will be used as the entropy coding in the next chapter (Video Coding).

After choosing N-To-P algorithm as an entropy-coding standard, it was time to improve the performance of the transform technique. Each block in the compression process block diagram, starting from Transform technique, coming through quantization and end up with entropy coding tries to get rid of the data that will not affect the image to be understandable but costs extra number of bits that lead to increase the file size.

The N-To-P is a lossless entropy-coding block, but the transform block and the quantization blocks are sensitive parts to deal with. The quantize block wiped out a lot of bits but without affecting the perceptual. In general, throwing away bits can cause severe distortion to the image and the quality could be degraded dramatically.

So, there were two new algorithms to improve the performance of the Transform technique. First algorithm was explained and implemented in section 3, by using a

Hybrid DWT-DCT algorithm by exploiting their advantages and attempt to eliminate there drawbacks. One of the privileges of the DCT is its good ability of energy compaction, but also it suffers from blockalization artefacts plus that the correlation between each block is not exploited properly. The DWT on the other side achieves better results in the medium and high compression ratio but the DCT obtained better results in the low level. Another disadvantage of the DWT is the ringing effect that is due to the wavelets basis. So in order to exploit and takes advantages of both techniques, a hybrid technique was used by attained the DWT technique first as a global technique to de-correlate each viewpoint first, then apply DCT technique to de-correlate the localized bits inside each block. This method have been implemented before on 3D-Integral images but the novelty in this section is that the hybrid technique applied to both LL DWT band as long as the HL DWT band, in addition to that the other blocks have been treated differently as explained in subsection 4.4. . Also there was one more algorithm have been represented that based on the same 2D-DWT-3D-DCT algorithm, but this time it was applied with more correlation to the third domain by applying one more DWT, in order to re-de-correlate the coefficients in this axis and the results showed an improvement in performance more than the former algorithm.

In the final section, two algorithms are presented to improve and develop the adaptive technique for the 3D Images. The first algorithm didn't show the results that were expected, so the second algorithm was implemented. The Mean proposed algorithm shows better results than the algorithm that was implemented in [32].

Chapter 5

Video Coding

5.1 Introduction

In this chapter a new 3D-Integral Video algorithm is represented in order to keep high video quality, in addition to this, optimize the performance of the Video speed. This is achieved by working in association with the previous proposed N-To-P coding technique. Another target is to investigate the best matching technique that adequate the 3D-Integral Video characteristics from different aspects, the quality and the complexity. Five different block-matching schemes are tested on 3D Integral video.

Video applications are one of the vital technologies nowadays. The demand of video application with 2D or 3D technologies increased rapidly. It has been used in a lot of application that is used daily. Video compression becomes an imperative enquiry. So in order to optimize the compression technique, the nature of video needed to be understated and illustrated to be able to compress it effectively [89].

Normally, the 2D-Video consists of a number of frames, each frame represented a 2D traditional image, to shrink the size of it, and an exploiting to the redundancy should be done. There are three different redundancies in video, temporal, statistical and spatial redundancy. The statistical and spatial both can be removed by applying encoding technique and transform technique (TC) respectively to each frame separately. To de-correlate the data between the frames and exploit the temporal redundancy, the Motion estimation and motion compensation techniques are applied.

The changes between the frames depend on the movement of the objects. Sometimes the change of the position or location of the object remain the same, other times it displaced slightly, and sometimes there is a big change in the position of the objects. In order to set the changing in the object position, a motion vector is determined by splitting each frame into a number of macro-blocks (i.e. (16x16)) and applying Motion estimation technique. Two frames types are used to de-correlate between the frames [90].

Intra- Frame and Inter- Frame are used in order to determine the Motion vector, Intra-Frame refers to the reference frame or I Frame, and Inter- Frame refer to current frame or P- Frame and other times it could be B-frame [37],[91],[92]As it has mentioned before that the video consists of a sequence of frames. These frames are divided to a Group of Pictures (GOP). Each GOP contains references frames and other frames depending on the reference frames.

I P B P I P P I

Not all of the frames will be coded in the same way, in order to increase the compression efficiency. Some frames will be treated in the same way as the normal still image by applying DCT then quantize the coefficients and finally encode them, these frames are the called I-Frames or reference frames as mentioned above and it comes usually every 3 or 4 frames an I-Frame should be treated in that way. The rest of the frames P-Frames (Predicted frame) and B-Frames (Bi-Prediction Frame) are using the Motion Estimation and Motion Compensation in order to reconstruct a predicted Image from the reference Image. In this chapter, a 3D-Integral video is used; a similar algorithm is applied as illustrated above but with slightly different changes. For simplicity, only I Frames and P-Frames are used. Normally, the DCT is used for transform technique and Huffman coding is used in entropy code

5.2 Proposed Algorithm

Indeed, the nature of the integral images is different that the normal images in the way of capturing which leads to a little bit of changes in compressing method. The integral Video consists of number of frames, but in this case each frame does not represent 2D-Images, because the way the video is captured, each frame captured by a parallel projection method not perspective method. So, in 3D-omnidirectional integral video, each frame consists of number of viewpoints depending of the number of lenses and the number of pixel per each micro-lens [2].

Figure 5.1 3D Video [2], [3]

A pre-processing step will be added to extract the viewpoints from the each frame. The way the viewpoints is treated will be different depend on the order of the frame. The sequence of I P I P I P I P will be followed, that means that the viewpoints of the first frame will be treated as a references frame. So each I Viewpoint will get into the same steps as if it is a 3D- Integral Image, staring from applying 3D-DCT then quantize each Viewpoint coefficient by a quantization factor, and finally enter to the encoder, which has previously proposed in chapter 4.

For the next Frame, which will be treated as P-Frame, the viewpoints are also extracted from it and then it enters to Motion Estimation in order to get the motion vectors. To get the Motion vectors, the motion estimation takes the middle viewpoint from the P-Frame and from frame store. The motion estimation needs to calculate the motion vector by comparing the middle viewpoint from the P-Frame (current Frame) and from the Inverse de-quantized coefficients of the middle viewpoint from the reference frame that has been stored.

The motion vector output will be applied to all viewpoints from stored reference I frame in order to get the predicted frame with predicted Viewpoints. Each Predicted Viewpoint will be subtracted from the original P-Viewpoints that have been extracted. The results will end up of subtracted predicted Viewpoints that will be transformed to spatial domain using 3D-DCT and quantized and finally encoded using the previous proposed algorithm. The motion Vectors will also be encoded in the same way, and the whole process in the decoder will be revered in order to reconstruct the video.

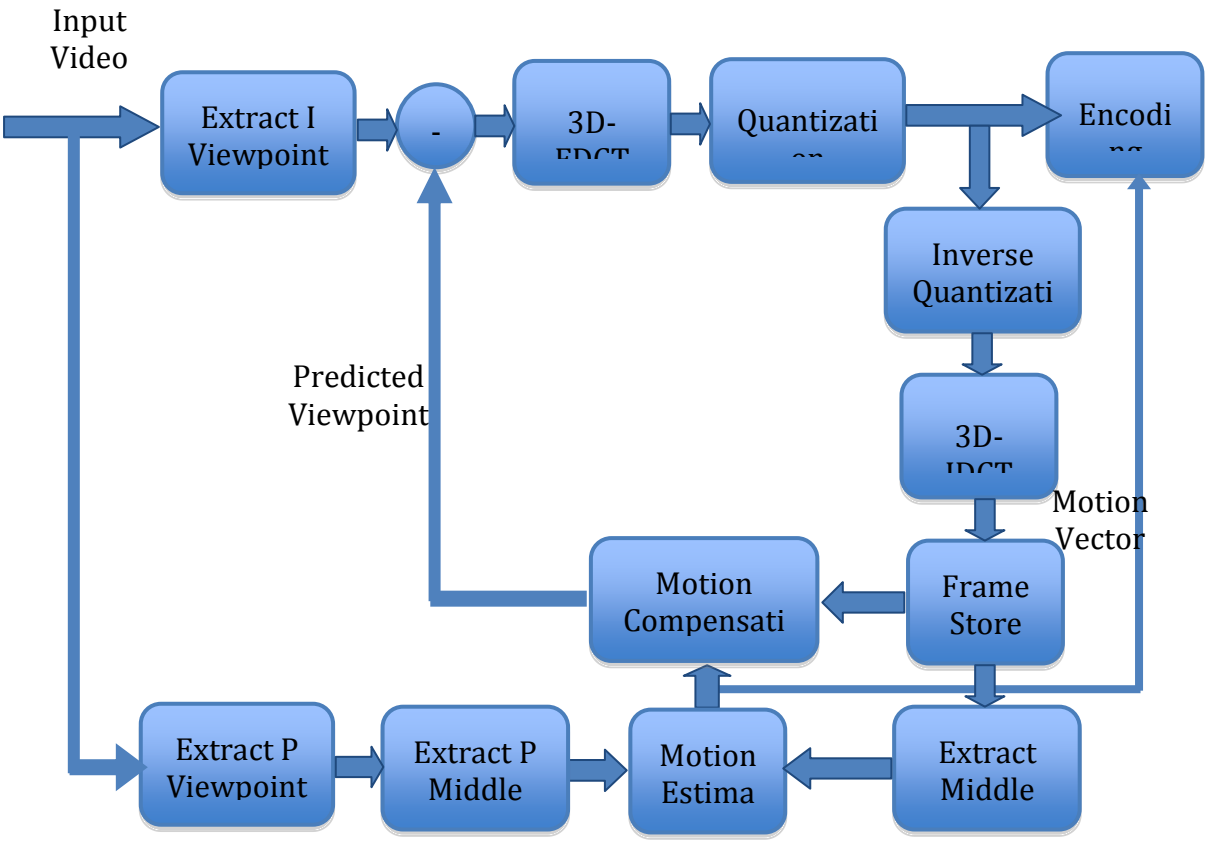


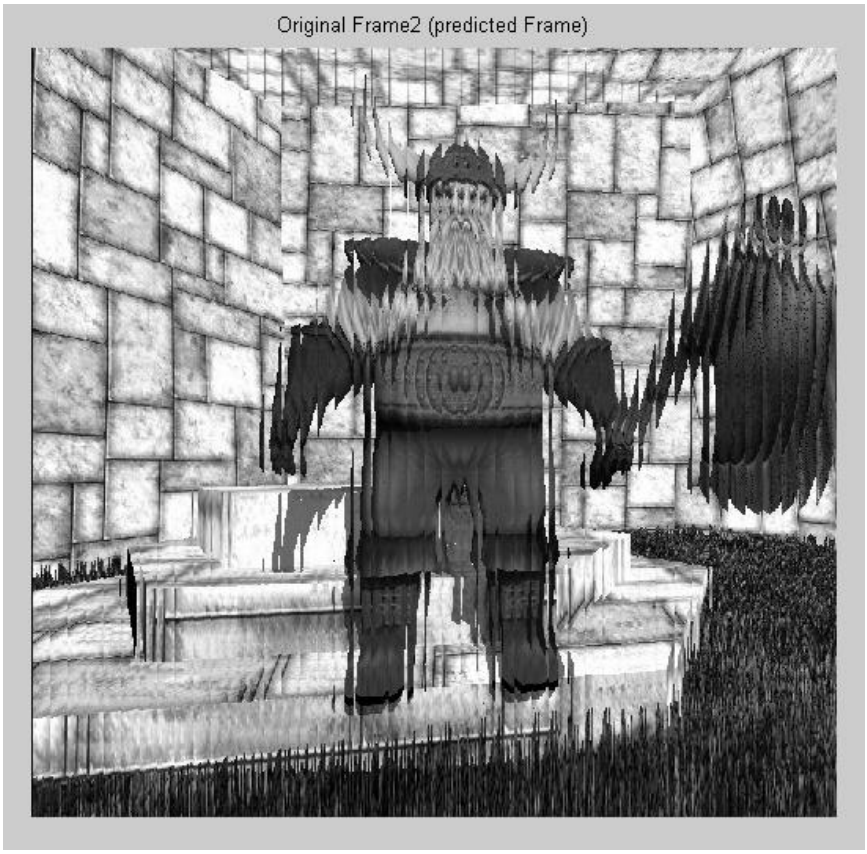
Figure 5.2 3D Integral Video Coding.

This proposed video algorithm aims to increase the speed by diminishing the computational complexity. As mentioned in 3D-Integral Video each frame got a number of viewpoints, so instead of calculating the motion vector for each viewpoint, the motion vector for the middle viewpoint only will be computed and used for the others.

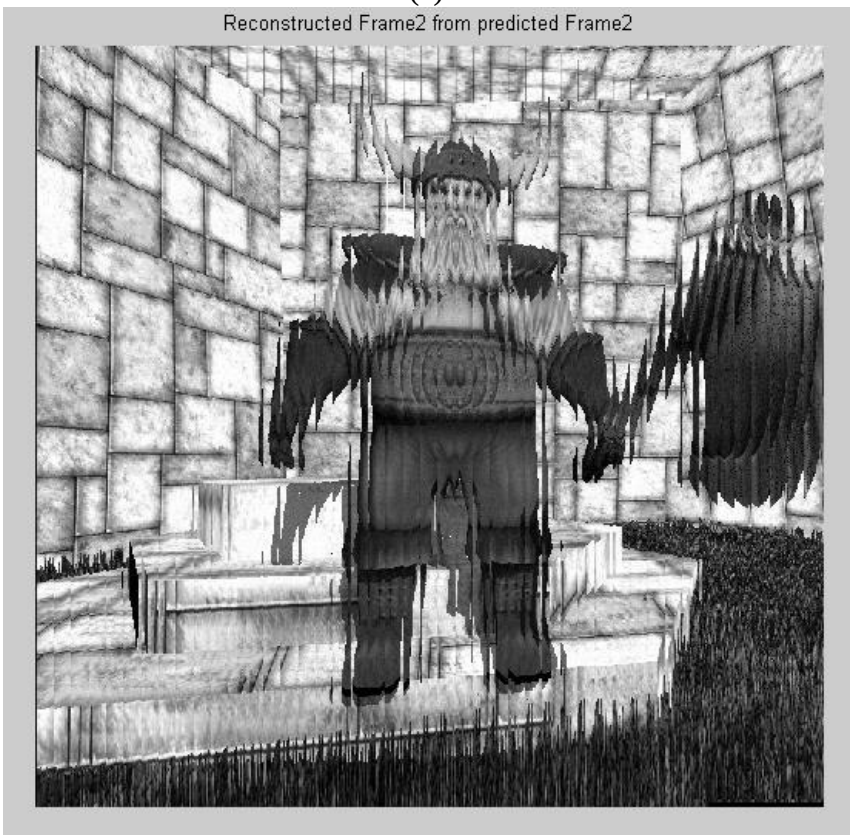
The following images will show the original and the reconstructed frames, the human visual system will not be able to notice the difference in quality between both, perceptually, at the same time there is reduction in complexity and that saves more time for the overall video.

First, figure 5.3 shows the original and reconstructed Frame 2. Frame 2 is a predicted frame and it can be observed that the notion algorithm did not affect the reconstructed frame perceptually. Also another predicted frame can be shown in figure 5.4 for frame 10, and the reconstructed frame, from the HVS aspect, is not affected.

Finally, the original and the reconstructed reference frame number 1, is represented in figure 5.5. From all the figures that are extracted from the video and shown the differences between the inputs and the outputs, it can be noticed that the proposed algorithm literally did affect the Peak Signal To noise ratio, but did not affect the human vision which is more important.

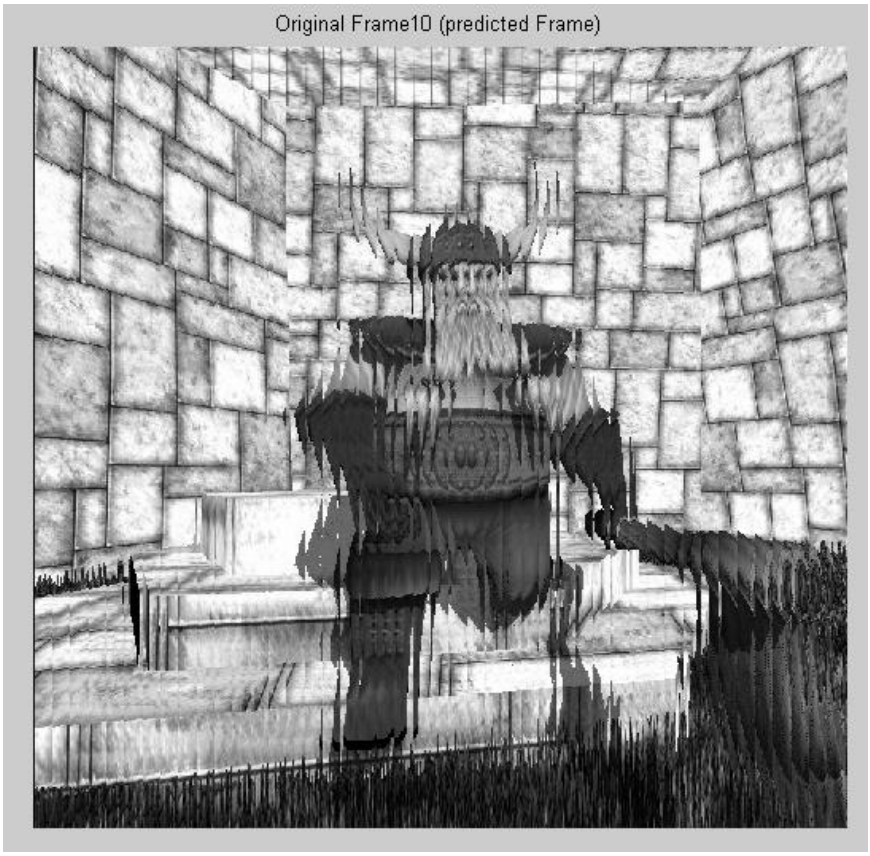


(a)



(b)

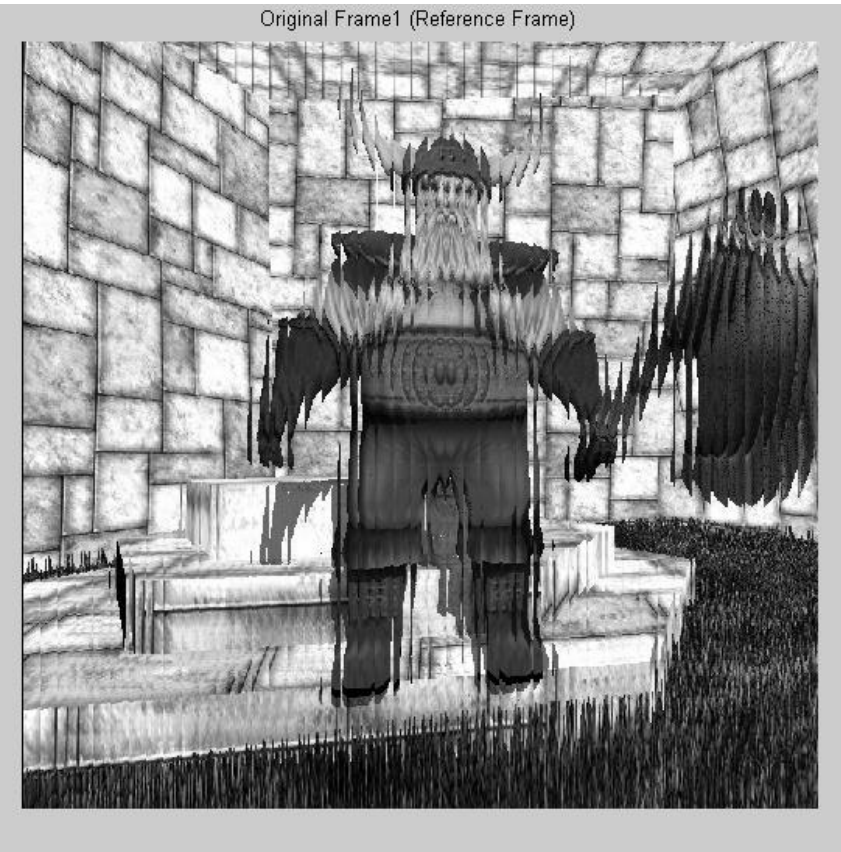
Figure 5.3 (a) the original frame 2 (b) the reconstructed frame 2.



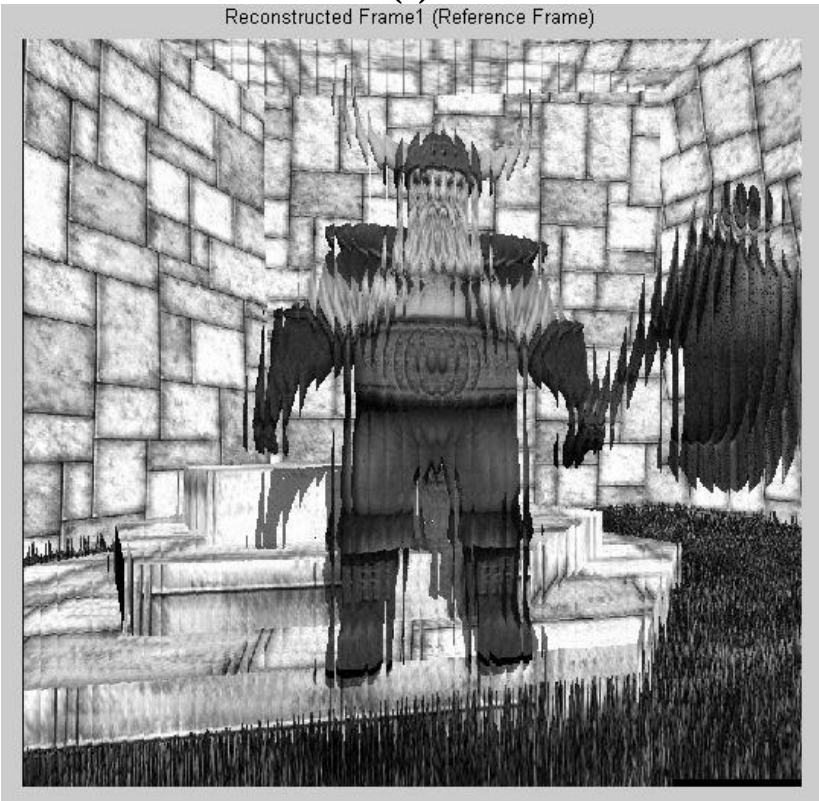
(a)

(b)

Figure 5.4 (a) the original frame 10 (b) the reconstructed frame 10.



(a)



(b)

Figure 5.5 (a) the original frame 1 (b) the reconstructed frame 1.

5.3 Block Matching Algorithms

There are many studies that have been done to improve the Motion estimation process. The two main factors that affect the performance of the process is the quality of the reconstructed result and the computational complexity, there is trade-off between both, so in order to get the best match a comparison between the current block and all the other possible locations in the search window should be done, and this results of a high number of computations which leads to complexity in computations, which is the case of Full Research Algorithm. Other algorithms minimize the number of computations but on account of the quality of the data for example like the Three Step Search algorithm [93], that decrease the number of positions searched in order to reduce the complexity that ends up to low the expenses cost.

Five Different Motion Estimation techniques are applied in order to obtain the best algorithm that matches the Integral input Data.

5.3.1Full Search Motion Estimation

Is also known as Executive Search, However the Full search gives the best results but one of drawbacks is its complexity in computations. Each Viewpoint is divided into (16x16) macro block. The search window is set to +7 and -7. So the search area is determined by this specific window size, as illustrated in figure 5.6.

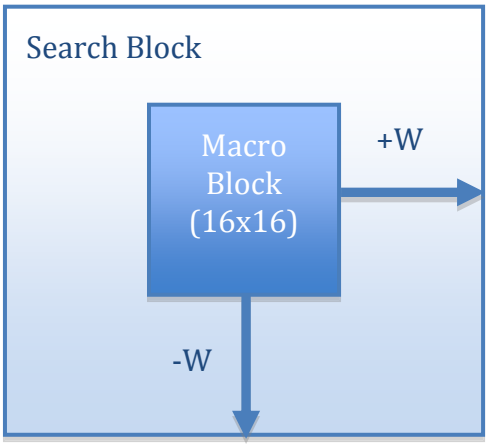


Figure 5.6 Search Window

In order to get the best block match, all possible positions in the search window is calculated, the main drawback is the bigger the search window, the more block possibilities are calculated, which ends up with more computational is done.

The parameter that set the best block matching is the cost function. The Mean Squared Error (MSE) [94], and the PSNR sets the performance of the motion compensation image that is results from the motion vectors (the output of the Motion Estimation) and the coordinate Macro block from the I-frame [95].

The following three schemes based on the central biased search.

5.3.2New Three Step Search

Actually, NTSS is based on the TSS; Koga et al introduced the later in 1981[96]. It was very common and popular used; it is based on choosing eight blocks as an initial step, and then divides the step size by two, and the centre shifted to the least deformation.

NTSS It is one of the Central biased search methods; so the first step is different than TSS, it inserts eight more checking points [89]. NTSS is used in MPEG-1 and H.261 [94].

5.3.3Four Step Search

It is the second Central biased search methods; it is similar to the NTSS but instead of using 9x9 windows, the size of the FSS window is 5x5, and the number of checking points is nine. The Cost function is calculated, and if it is found to be the minimum, then the process is moved to the final step by reducing the search window to 3x3 and the motion vector is determined, otherwise, the searching window is minimized to 5x5. The checking points is increased to five more if the cost function founded in one of the search window corners, and increase by three if the cost function founded in the middle of one of the horizontal or vertical axis, but if it is founded in the centre than no need to add more checking points and move to the final step which is reducing the window search to 3x3, figure 5.7 represent it.

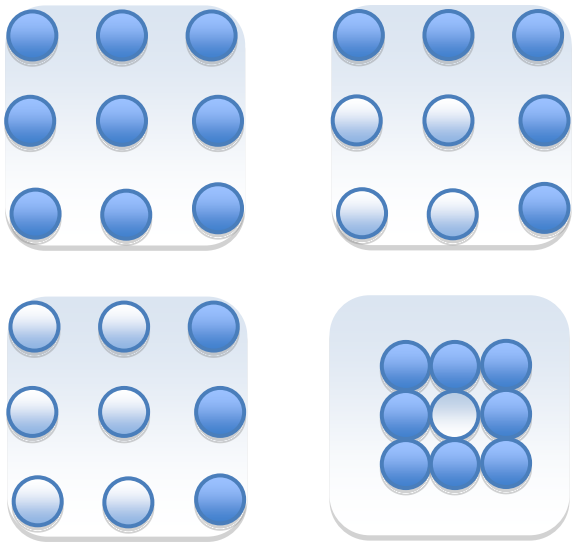


Figure 5.7 Four Step Search

5.3.4Diamond Search

It is our last central biased search scheme, figure 5.8. It follows the same method of the four steps search but instead of square it is diamond. Two patterns are used, the first one with nine checking points and the second one with five checking points. Both are of diamond shape but with different sizes. The larger one keep repeated until the minimum cost function founded in the centre, then it is replaced by the smaller diamond shape with five checking points [89].

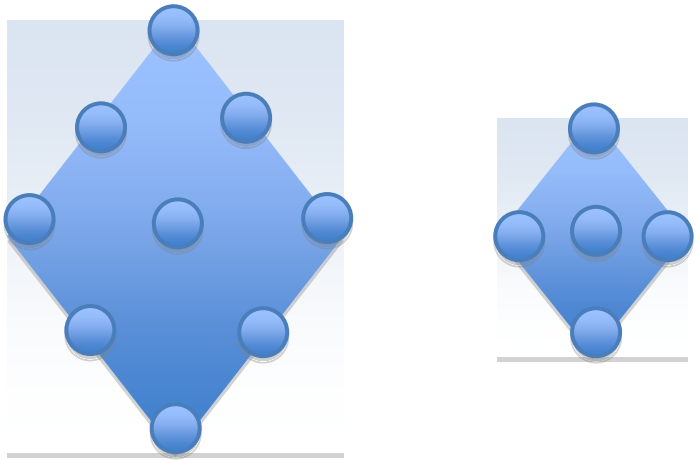


Figure 5.8 Diamond Search, the large search window with 9 checking points and the small search window with 5 checking points

5.3.5Adaptive Rood Pattern Search (ARPS)

In ARPS the main idea based on the concept of rood shape that have main four highest points, So the ARPS got these four highest points in addition to the other search points that predicted by the motion vector. The author in [97] his vision is that the current block will move in the same direction as the surrounding blocks directed to move to, and he depends on this probability, that is why he is taking the motion vector as a reference to calculate the cost function [94].

The motion vector is possible to be on the same line with the four highest points. So, it can get four overlapping points or five overlapping points.

5.4 Results

The following results are represented to investigate two issues the compression ratio and computational complexity, the former to measure the quality of the frames for different quantization levels and the latter to estimate the time the whole process is taking for each scheme for each frame.

Figure 5.9 shows a comparison for the performance of the five different block-matching schemes separately with quantization factor equals to 32, figure 5.10 shows the same results but all in the same graph to make the comparison more clear and accurate. Figures 5.11,5.12 and 5.13 presents the same comparison as figure 5.10 but with different quantization factors, 16,8 and 2 respectively. The results for the different quantization levels proof that the best performance is for the algorithms that uses Full search and that uses the adaptive Rood Pattern search. However the technique for both schemes are totally different but the way they scan for the best matching points touches the target, either by applying a fully search method or by applying an adaptive one to reach to the maximum optimization.

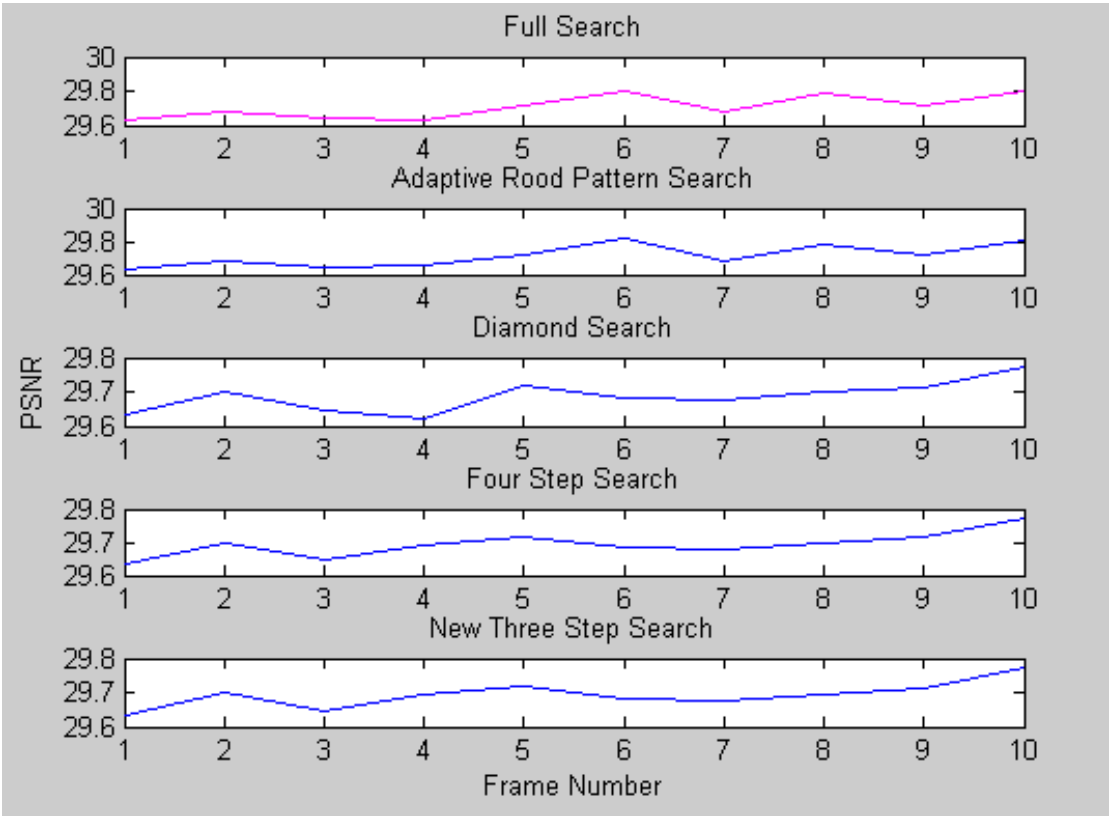


Figure 5.9 Different Motion Estimation techniques with Q factor=32.

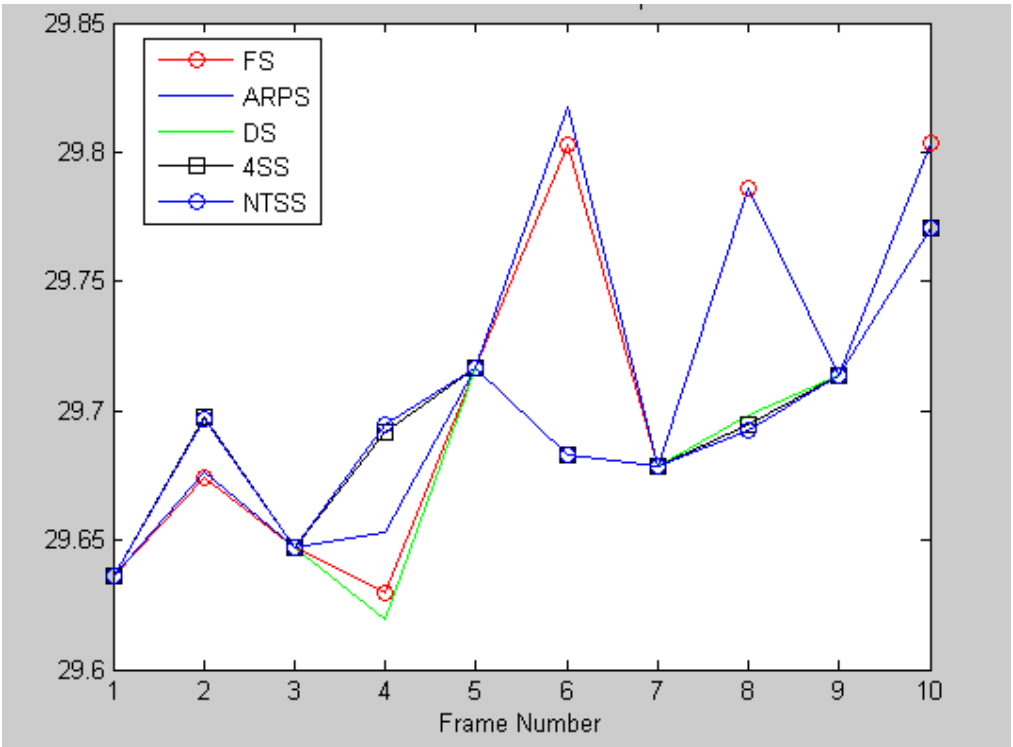


Figure 5.10 Different Motion Estimation techniques with Q factor=32.

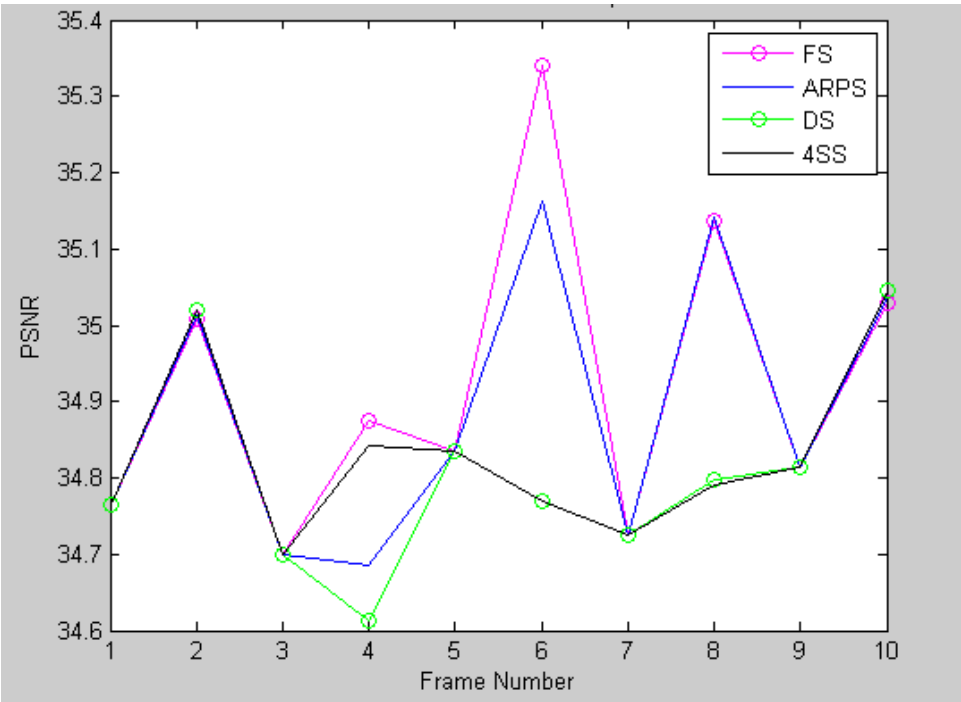


Figure 5.11 Different Motion Estimation techniques with Q factor=16.

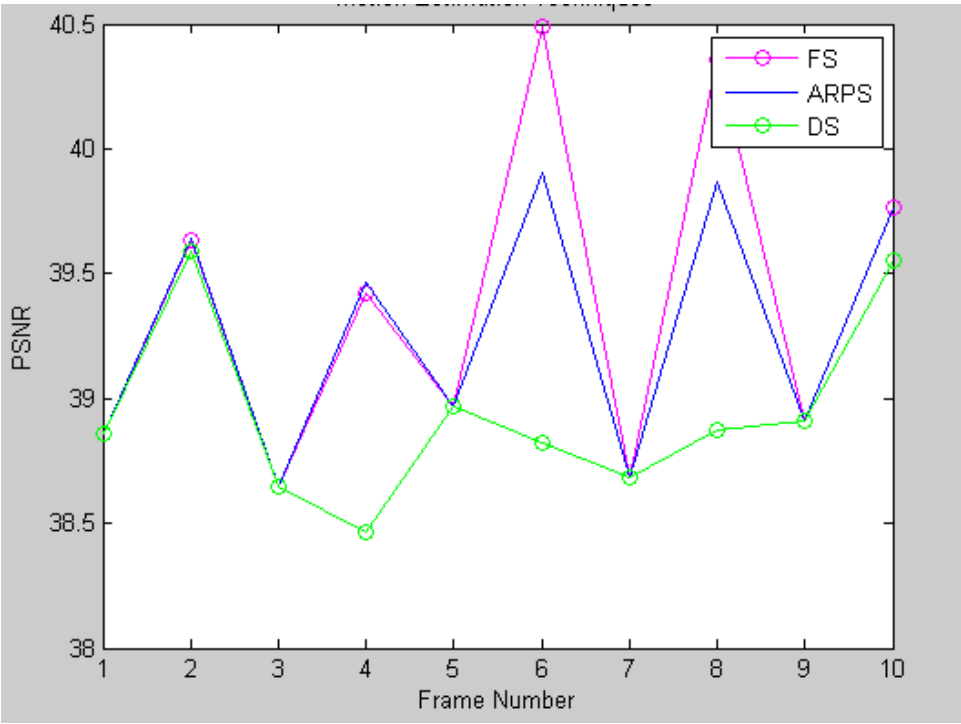


Figure 5.12 Different Motion Estimation techniques with Q factor=8.

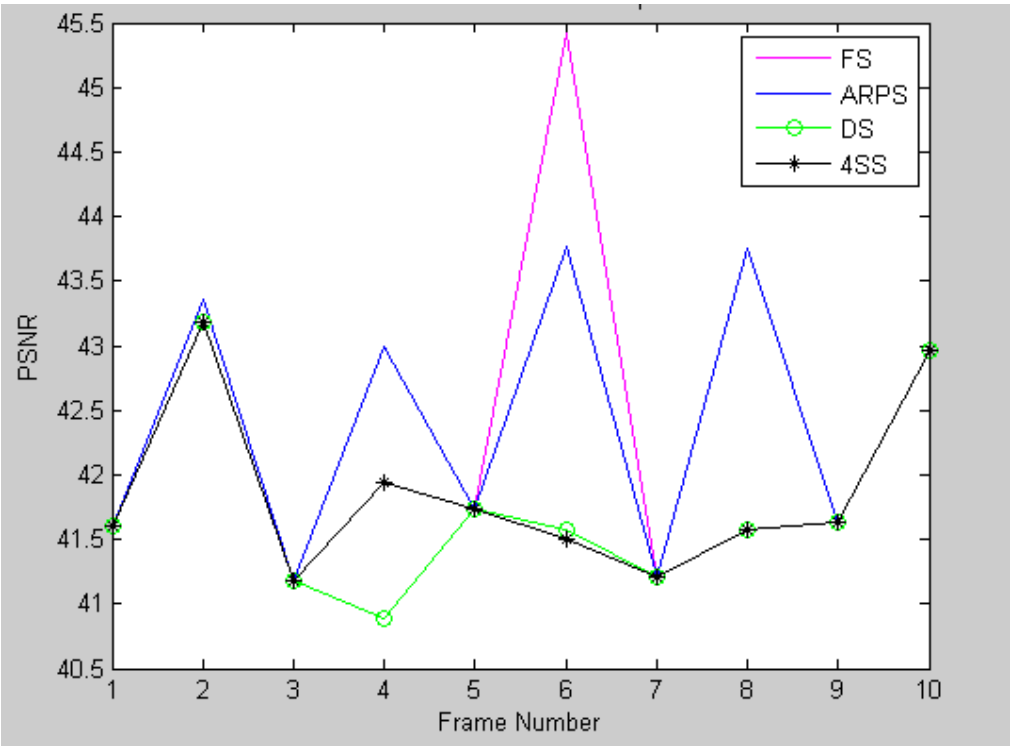


Figure 5.13 Different Motion Estimation techniques with Q factor=2.

5.5 Computational Complexity

Despite the good performance of the full search scheme over the other scheme, but from the complexity aspect it appears from figure 5.14 that it acts as the worst scenario that can occurs. Figure 5.15 represents the huge difference in complexity between the Full Search and the other four schemes, in figure 5.16 a maximize part of the other four schemes is presented and it shows that Diamond search is the next one that followed the Full search in its complexity followed by 4 Step Search and New Three Step Search and finally comes the Adaptive Rood Pattern Search with the least computational complexity. That makes the ARPS is the best matching block scheme that proof its efficiency in quality and complexity.

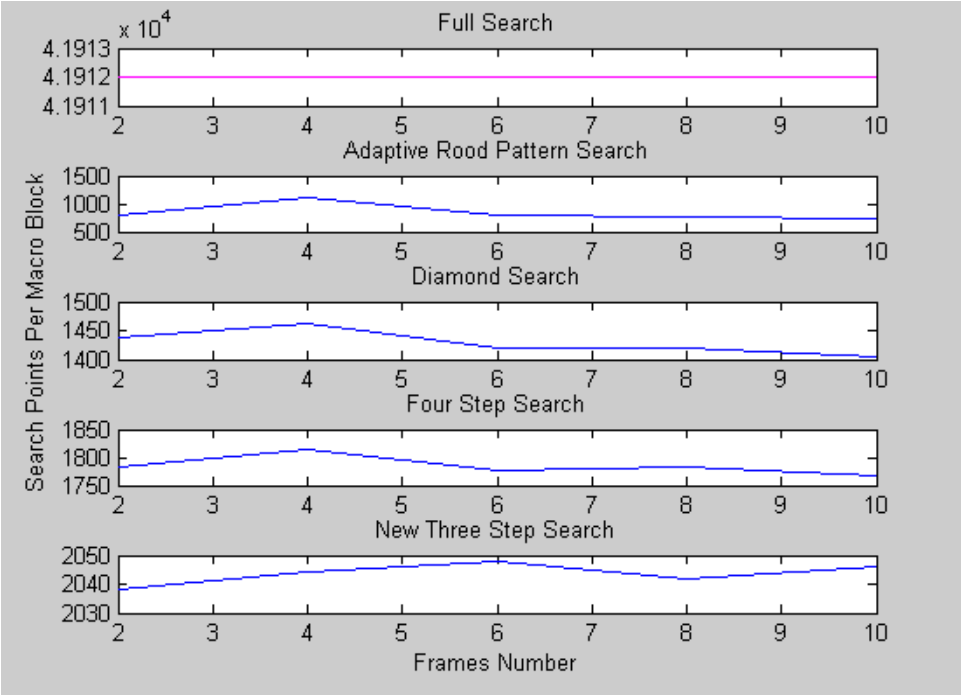


Figure 5.14 Different Motion Estimation techniques Computational compexity.

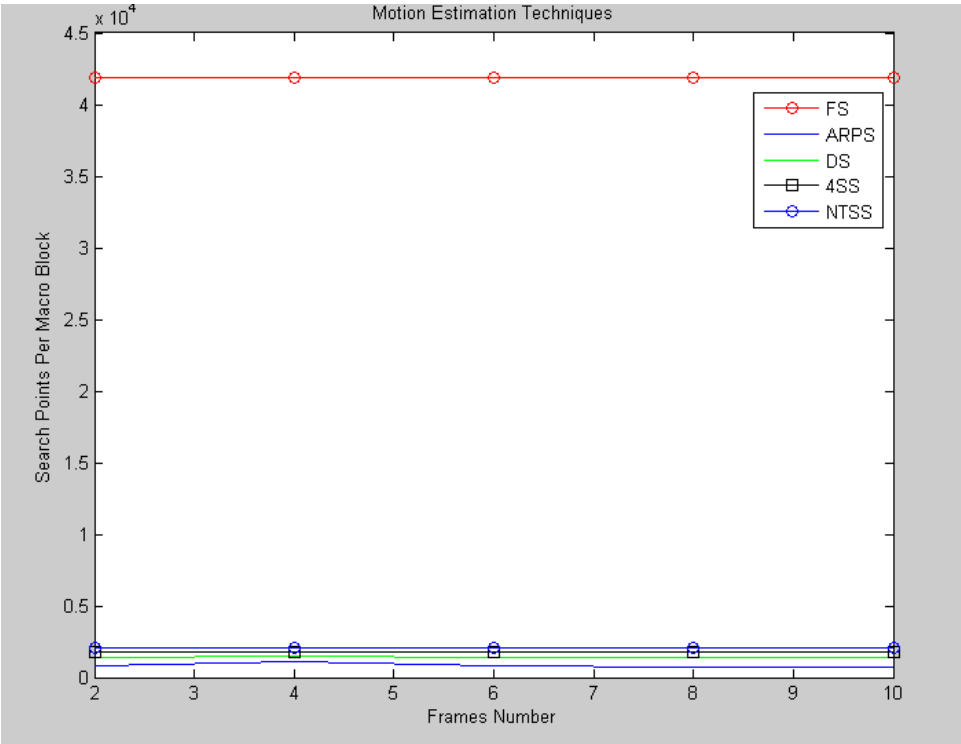


Figure 5.15 Different Motion Estimation techniques Computational complexity

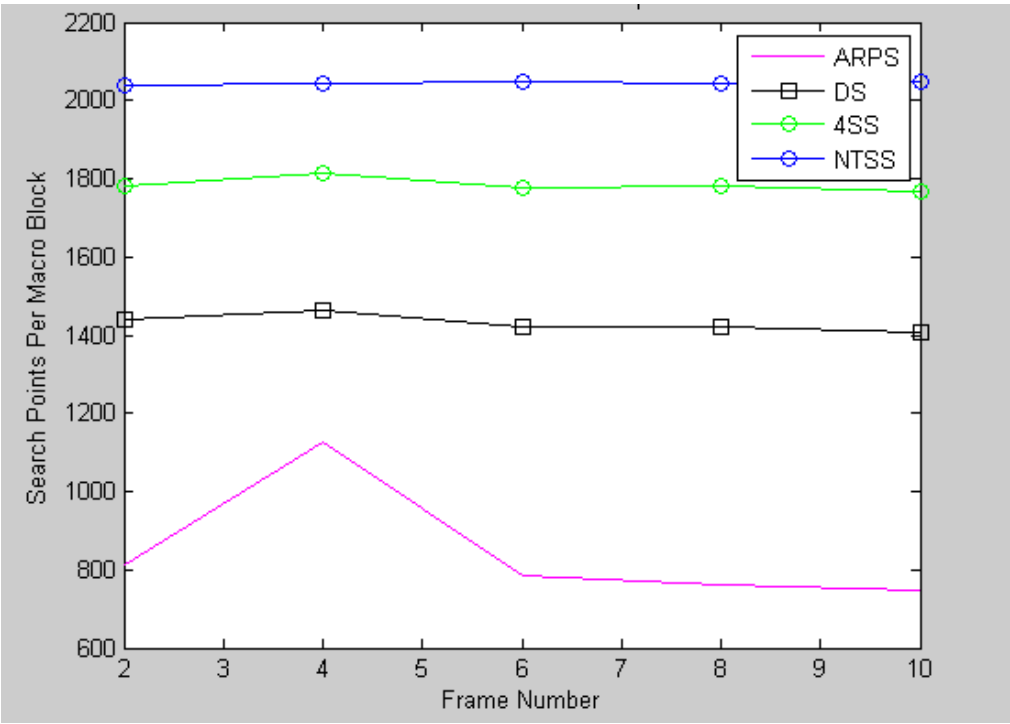


Figure 5.16 Different Motion Estimation techniques Computational complexity

5.6 Conclusion

This chapter have been working on improving the compression performance of the 3D-Integarl Video and also investigating the best block matching technique that suits the Integral Video features.

In order to improve the compression process the proposed algorithm in chapter 3 is implemented in association with a new proposed algorithm that adequate the characteristics of the Integral Video frames to decrease the number of calculations for each frame, as a consequence, it speed up the whole process and eliminate the complexity overall the video frames. This proposed algorithm based on selecting the middle Viewpoint then estimating the motion vector for it then popularize this particular motion vector for all of the viewpoints in the same frame, this process is a repeated for all frames.

From the results that represented previously, it is shown that the human vision will not affected by the new proposed algorithm, but on the other side this new proposed algorithm will reduce the time it is taking for compressing each frame thus reducing the whole video compression process. So literally there is not only keening the vision quality of the video but also achieving time reduction at the same time.

In the second part of the chapter, different block matching techniques is tested on the 3D-Integral Video, in order to find the best scheme that works effectively and ensembles the features of the 3D-Integral data. Five block matching techniques have been tested on the Axe Video shown in figure 5.1.

The huge difference in computational complexity between the Full search Estimation algorithms and the other algorithms has been illustrated, however the PSNR results is better but it will not substitute the big difference in cost and time. The second best algorithm that tried to make a balance between the two most important factors -the quality and the complexity- is the Adaptive Rood Pattern Search. The Adaptive Rood Pattern Search shows reasonable results according to the 3D-Integral content.

Chapter 6

Conclusion and Future Work

During the last decades the computer and Internet technology improved in an amazing way. Nowadays, the use of the different types of digital devices involved vividly in the human life, different kind of data is transmitted daily and needed, starting from the traditional still images, coming through 3D Images, Video conferences, 3D Videos, etc.

The demand of achieving a high compression rate is increasing day after day due to the various high technology applications that been in use. But a trade off between achieving efficient compression ratio and getting high quality is been a challenge in the last few years.

The raising of the challenges ceiling motivate the researchers to seek for applying new algorithms to follow the market requirements.

6.1 Scope of the thesis

The scope of this thesis was mainly focused on the following aspects:

- Designing entropy coding for 2D Images and compare it with JPEG standards. To achieve the trade between the quality of the Image and the high compression ratio and at the same time reduce the computational complexity.

- Implement the entropy coding represented in chapter 4, to the 3D Integral Images using 3D DCT technique.

- Design a new 2D and 3D DWT algorithm in associate with the new entropy coding.

- Improve the compression ratio by developing a hybrid DCT/DWT technique. To exploit the advantages from both the DCT and DWT systems and avoid there drawbacks in order to improve the compression process.

- Implementing new adaptive 3D-DCT to overcome the artefacts blocking. Eliminate the artefacts effects that the 3D-DCT Images facing by developing a new techniques.

- New 3D Video coding algorithm for 3D Integral Video. Improving the performance of the 3D-Video coding technique via using new algorithm and investigated the most proper block matching algorithm that proof better results when working in cooperation with this proposed algorithm.

6.2 Results and analysis:

6.2.1 Proposed Entropy Coding:

- The proposed entropy-coding algorithm proofs a high compression ratio with the same PSNR comparing to the JPEG standards. In addition, a dramatically decrease in time is achieved.

  The reduction in time refers to the simplicity of the algorithm and less in computational process comparing to the complicated process that the other standards is facing.

  In JPEG baseline, the time is increasing dramatically with the increase in the bit rates. On contrary, in N-TO-P, not only the time is reduced but also, there is no variation in the time for different bitrates level. The time scale is almost fixed.

  This stability on the time is due to the simplicity and symmetric of the encoding process, it is not depending on the quantization method or the quantization scale like the JPEG standard. The increase in time in JPEG is refer to the number of zeros that is results from applying the quantization process to the coefficients and the way it is treated by using the RLE technique and also Huffman table. So the variations of the level of time are varied according to the bitrates. The amount of data, which required to be scanned, for the high quality, is larger than the amount of data that is with lower PSNR. This is due to the number of zeros that has been neglected according to the RLE. In N-TO-P, the amount of data is unchangeable weather it was for high or low quality.

  In N-TO-P, the process depends on the sign of the values, and the difference between the indexes values, where both are simple calculations.

  Comparing the complexity of these two calculations in N-TO-P , with the scanning process of the RLE, in order to account the number of zeros then giving each value a specific symbol, and finally scan this values with Huffman table and implement it, shows from where the reduction in both time and complexity come from.

  The N-TO-P is a kind of symmetric operation rather than variant operation. Where JPEG depends on the content of the coefficient working in association

of number of tables that need to be scanned with each value which is the cause of the time increase especially with high bitrates where there is no big loss in data.

This in the case of the JPEG baseline, on the other standards more computational process is required in order to a get the desired compression ratio.

The compression rate is increased because it takes advantage of the fact that the sign values take more space than the other unsigned values and also the fact that smaller integer is much effective in the reduction of the size of the file. Exploiting these two facts with two simple calculations leads to this dramatic reduction in time and bitrates. In addition, the entropy algorithm that is used is lossless meaning that the quality of the image will remain the same.

- It also shows a higher improvement when used with the 3D Integral Images as shown in chapter 5, in terms higher compression ration and decrease in the computational complexity. However the results are not at the same level as when the algorithm is applied to 2D Images. This is mainly due to the way the 3D DCT coefficients were scanned. In the experimentation carried out in chapter 5, 2D planar zigzag is used for each plane and hence the correlation between the different planes was not exploited. As 3D zigzags scan is needed to be applied to the 3D volume will increase the efficiency of the N-TO-P applied when applied on the 3D Images.

- For the 2D-DWT, the new algorithm shows a very good results comparing to EZW and SHPIT as shown in chapter 6 from aspects of PSNR, bitrates and computational complexity.

- When it is compared to EBCOT, the proposed 2D-DWT algorithm shows massive reduction in the complexity due to the high complexity level of the EBCOT is facing. Furthermore the proposed algorithm achieved better image quality results in term of PSNR values at low bit rates. Although at high bit rate, the EBCOT algorithm tends to achieve image quality.

- The algorithm that has been applied to the 2D-DWT is quit simple; a number of correlation parameters have not been exploited unlike the EBCOT or even the proposed 2D DCT case. This can be shown from the comparison between 2D-DCT and EBCOT as DCT shows better results in the low bitrates than EBCOT, which is in contradiction to the traditional JPEG baseline. So in that case the advantage of power exemption is increased and the algorithm overcome the disadvantage of the lower bitrates comparing to JPEG2000, with time and complexity reduction.

6.2.2 Proposed Adaptive Algorithm

- Two algorithms were proposed in chapter 7, in order to reduce the artifacts errors that 3D-DCT Images is facing. The first one didn't achieve the results that were expected, that lead us to seek for a better algorithm. The results were compared to the results that were proposed in [29] and it proofs its effectiveness.

6.2.3 Video Coding

- Finally, the proposed 3D Integral Video Coding shows good results with both the Adaptive Rood Pattern Search and the full search from the Quality aspect. But the factor that makes the ARPS better is the huge difference in complexity between the full search and the rest of the implemented blocking search algorithms.

6.3 Future Work

6.3.1 Hybrid Index Algorithm

As shown from the results in chapter 4 and 5, the new proposed entropy-coding algorithm represents a good compression ratio with a significant time reduction. But

comparing both of the results of the 2D and the 3D, it can be shown that due to the large amount of data in the 3D Integral image, the efficiency of the algorithm decreased by a very small percentage but still better than the traditional entropy coding. The main factor that affects the Integral image compression ratio is the size of the Index arrays

In order to improve the compression ratio not for the 3D Integral images only but for both 2D and 3D a new Index algorithm can be implemented. This new algorithm can be implemented by combining the Cumulative index algorithm that presented in chapter 4 and the Difference Index algorithm represented in the same chapter. This can improve the overall compression ratio without increasing the computational complexity that is due to the fact that the computational complexity of both algorithms are already very small which will not affect the time effectively if they are become together in a hybrid index algorithm.

6.3.2 The 3D-DWT algorithm

Improve the 3D-DWT algorithm that have been proposed in chapter 6, for example by applying more DWT levels.

6.3.3 H.264 and HEVC

Implement the H.264 and also HEVC using the proposed entropy coding that represented in chapter 4. Using this new algorithm can raise the compression ratio more and also will decrease the high computational complexity that H.264 is suffering from.

6.3.4 Predicted Viewpoint Algorithm

It is an algorithm that can increase the efficiency of the Integral Images by achieving higher compression ratio in association with a high PSNR. Treating the extracted

Viewpoints as frames in a video could do this. Using Intra Viewpoint and Inter Viewpoint following the same technique as the video frames.

Not all of the Viewpoints will be coded in the same way, in order to increase the compression efficiency. Some Viewpoints will be treated in the same way as the normal still image by applying DCT then quantize the coefficients and finally encode them, these Viewpoints will be called I-Viewpoint or reference Viewpoints.

The rest of the frames P-Viewpoints (Predicted Viewpoints) are using the Motion Estimation and Motion Compensation in order to reconstruct a predicted Image from the reference Image.

REFERENCES

[1] Marta, M., Kunt, M., and Grgic, M. (2010) “Signals and Communication Technology: High-Quality Visual Experience: Creation, Processing and Interactivity of High-Resolution and High-Dimensional Video Signals”, 10th edn, springer, verlag berlin heidelberg.

[2] 3D VIVANT (2013). Available at: http://www.3dvivant.eu/.

[3] Aggoun, A., Tsekleves, E., Zarpalas, D., Daras, P., Dimou, A., Soares, L. and Nunes, P. (2013) "Immersive 3D Holoscopic Video System[A2]" *Multi Media IEEE* vol. 20, no.1, pp. 28-37.

[4] Adedoyin, S., Fernando, W. and Aggoun, A. (2007) "A joint motion & disparity motion estimation technique for 3D integral video compression using evolutionary strategy", *Consumer Electronics, IEEE Transactions* on, vol. 53, no. 2, pp. 732-739.

[5] Adedoyin, S., Fernando, W., Aggoun, A. and Kondoz, K. (2007) "Motion and disparity estimation with self adapted evolutionary strategy in 3D video coding", *Consumer Electronics, IEEE Transactions on,* vol. 53, no. 4, pp. 1768-1775.

[6] Aggoun, A., Mehanna, A., Tsekleves, E., Cosmas, J. and Loo, J. (2010) "Live immerse video-audio interactive multimedia", *Broadband Multimedia Systems and Broadcasting (BMSB), 2010 IEEE International Symposium on IEEE*, pp. 1-5

[7] Shrestha, S. (2010) "Hybrid DWT-DCT algorithm for image and video compression applications*" phd diss. University of Saskatchewan*[A3]

[8] Santoso, A. J., Nugroho, L. E., Suparta, G. B., Hidayat, R. (2011) "Compression Ratio and Peak Signal to Noise Ratio in Grayscale Image Compression using Wavelet", *InternatIonal Journal of Computer Science and technology, IJCSt,* vol. 2, no. 2, pp. 7-11.

[9] Bystrom, M., Richardson, I,. Zhao,Y,. (2008) "Efficient mode selection for H.264 complexity reduction in a Bayesian framework", *Signal Processing: Image Communication*, vol. 23, pp. 71-86.

[10] Ostermann, J., Bormans, J., List, P., Marpe, D., Narroschke, M., Pereira, F., Stockhammer, T. and Wedi, T. (2004) "Video coding with H. 264/AVC: tools, performance, and complexity", *Circuits and Systems magazine, IEEE*, vol. 4, no. 1, pp. 7-28.

[11] Wallace, G.K. (1992) "The JPEG still picture compression standard", *Consumer Electronics, IEEE Transactions on*, vol. 38, no. 1, pp. 18-34.

[12] Wallace, G.K. (April 1991) "The JPEG Still Picture Compression Standard", *Communication of the ACM* vol. 34, no. 4.

[13] Lakhani, G. (2004) "Optimal Huffman coding of DCT blocks", *Circuits and Systems for Video Technology, IEEE Transactions* on, vol. 14, no. 4, pp. 522-527.

[14] Kim, S., Heo, J. and Ho, Y. (2010) "Efficient entropy coding scheme for H. 264/AVC lossless video coding", *Signal Processing: Image Communication*, vol. 25, no. 9, pp. 687-696.

[15] Okoshi, T. (1980) "Three-dimensional displays", *Proceedings of the IEEE*, vol. 68, no. 5, pp. 548-564.

[16] Dudnikov,Y.A. (1970) "Autostereoscopy and integral photography", *Opt.Tech*, vol. 37, no. 7.

[17] http://science.howstuffworks.com/3-d-glasses2.htm.

[18] Lawton, G. (2011) "3D displays without glasses: coming to a screen near you", *Computer*, vol. 44, no. 1, pp. 17-19.

[19] http://www.sundoginteractive.com/sunblog/posts/3d-digital-signage-without-the-glasses.

[20] Reichelt, S., Häussler, R., Fütterer, G. and Leister, N. (2010) "Depth cues in human visual perception and their realization in 3D displays", *SPIE Defense, Security, and Sensing International Society for Optics and Photonics*, , pp. 76900B.

[21] Hong, J., Kim, Y., Choi, H., Hahn, J., Park, J., Kim, H., Min, S., Chen, N. and Lee, B. (2011) "Three-dimensional display technologies of recent interest: principles, status, and issues [Invited]", *Applied Optics*, vol. 50, no. 34, pp. H87-H115.

[22] Aggoun, A. (2006) "A 3D DCT compression algorithm for omnidirectional integral images", *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on IEEE*, pp. II-II.

[23] http://www.3pointd.com/20070522/3d-holographic-tv-closer-than-you-think

[24] Kim, Y. and Hong, K. (2009) "Recent researches based on integral imaging display method", *3D RESEARCH*, vol. 1, no. 1, pp. 17-27.

[25] Lippmann, G. (1908) "Epreuves reversibles donnant la sensation du relief", *J.Phys.Theor.Appl.*, vol. 7, no. 1, pp. 821-825.

[26] Eljdid, M., Aggoun, A. and Youssef, O. (2007) "Computer Generated Content for 3D TV", *3DTV Conference, 2007 IEEE*, pp. 1-4

[27] Sokolov, A. (1911) "Autostereoscopy and Integral Photography by Professor Lippmann's Method" *MGU, Moscow State Univ. Press*.

[28] Ives, H.E. (1931) "Optical properties of a Lippman lenticulated sheet", *Journal of the Optical Society of America* (1917-1983), vol. 21, pp. 171.

[29] Wu, C., Wang, Q., Yang, Y., Fu, G. and Wang, H. (2007) "3D reconstruction

from integral images", *3rd International Symposium on Advanced Optical Manufacturing and Testing Technologies: Advanced Optical Manufacturing Technologies International Society for Optics and Photonics*, pp.67220E-67220E.

[30] Montebello, R.L.D. (1970) PRODUCTION OF LENTICULAR SHEETS FOR INTEGRAL PHOTOGRAPHY, *US Patent 3,538,198.*

[31] Milnthorpe, G., McCormick, M., Aggoun, A., Davies, N. and Forman, M. (2002) "Computer generated content for 3D TV displays", *International Broadcasting Convention, Amsterdam, Netherlands.*

[32] Zaharia, R., Aggoun, A. and McCormick, M. (2002) "Adaptive 3D- DCT compression algorithm for continuous parallax 3D integral imaging", *Signal Processing: Image Communication*, vol. 17, no. 3, pp. 231-242.

[33] Jayaraman, S. Esakkirajan, T. Veerakumar, (2011) "Digital Image Processing", *Tata McGraw-Hill Education*.

[34] Kalva, H. (2006) "The H. 264 video coding standard", *MultiMedia, IEEE*, vol. 13, no. 4, pp. 86-90.

[35] Jeon, B., Park, J. and Jeong, J. (1998) "Huffman coding of DCT coefficients using dynamic codeword assignment and adaptive codebook selection *", Signal Processing: Image Communication*, vol. 12, no. 3, pp. 253-262.

[36] http://www.mvnet.fi/index.php?osio=Tutkielmat&luokka=Yliopisto&sivu=Image_compression

[37] Richardson, I. E. (2004) "H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia", *John Wiley & Sons.*

[38] http://www.britannica.com/EBchecked/topic/152168/data- compression.

[39] Battiato, S., Bosco, C., Bruna, A., Di Blasi, G. and Gallo, G. (2005) "Statistical

modeling of huffman tables coding " *in Image Analysis and Processing–ICIAP 2005 Springer,* pp. 711-718.

[40] http://www.cs.sfu.ca/CourseCentral/365/li/interactive- jpeg/Ijpeg.html.

[41] Swaminathan, A. and Agarwal, G. "A comparative study of image compression methods".

[42] Gan, L., Tu, C., Liang, J., Tran, T.D. and Ma, K. (2007) "Undersampled boundary pre-/postfilters for low bit-rate DCT- based block coders", *Image Processing, IEEE Transactions on*, vol. 16, no. 2, pp. 428-441.

[43] Chen, W. and Pratt, W. (1984) "Scene adaptive coder", *Communications, IEEE Transactions on,* vol. 32, no. 3, pp. 225-232.

[44] Shahbahrami, A., Bahrampour, R., Rostami, M.S. and Mobarhan, M.A. (2011) "Evaluation of Huffman and Arithmetic Algorithms for Multimedia Compression Standards*", arXiv preprint arXiv:1109.0216.*

[45] Kavousianos, X., Kalligeros, E. and Nikolos, D. (2007) "Optimal selective Huffman coding for test-data compression", *Computers, IEEE Transactions on,* vol. 56, no. 8, pp. 1146-1152.

[46] Jas, A., Ghosh-Dastidar, J., Ng, M. and Touba, N.A. (2003) "An efficient test vector compression scheme using selective Huffman coding", *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions* on, vol. 22, no. 6, pp. 797-806.

[47] Marpe, D., Schwarz, H. and Wiegand, T. (2003) "Context-based adaptive binary arithmetic coding in the H. 264/AVC video compression standard", *Circuits and Systems for Video*

[48] Wiegand, T. and Schwarz, H. (2011) "Source Coding: Part I of Fundamentals of Source and Video Coding", *Foundations and Trends in Signal Processing,*

vol. 4, no. 1–2, pp. 1-222.

[49] Lu, W. and Gough, M. (1993) "A fast-adaptive Huffman coding algorithm", *Communications, IEEE Transactions on*, vol. 41, no. 4, pp. 535-538.

[50] Setia, V. and Kumar, V. (2012) "Coding of DWT Coefficients using Run-length coding and Huffman Coding for the purpose of Color Image Compression", *International Journal of Computer and Communication Engineering,* vol. 6.

[51] Sharma, M. (2010) "Compression using Huffman coding", *IJCSNS International Journal of Computer Science and Network Security,* vol. 10, no. 5, pp. 133-141.

[52] Zhilin, L., Xiuxia, Y. and Lam, K.W. (2002) "Effects of. JPEG compression on the accuracy of photogrammetric point determination", *Photogrammetric Engineering and Remote Sensing,* vol. 68, no. 8, pp. 847-853.

[53] Nebot, J.P., Morbee, M. and Delp, E.J. (2012) "Generalized PCM Coding Of Images". *Image Processing, IEEE Transaction,* vol. 21, no. 8, pp. 3801-3806.

[54] Rabbani, M. and Joshi, R. (2002) "An overview of the JPEG 2000 still image compression standard", *Signal Processing: Image Communication,* vol. 17, no. 1, pp. 3-48.

[55] Santa-Cruz, D., Ebrahimi, T., AskelÖf, J., Larsson, M. and Christopoulos, C. (2000) "JPEG 2000 still image coding versus other standards", *Proceedings of SPIE, the International Society for Optical Engineering Society of Photo-Optical Instrumentation Engineers,* vol.4115, pp. 446-454.

[56] Bénéteau, C. and Van Fleet, P.J. (2011) "Discrete wavelet transformations and undergraduate education", *Notices of the AMS,* vol. 58, no. 05.

[57] Dubey, V.G. and Singh, J. (2012) "3D Medical Image Compression Using

Huffman Encoding Technique". *International Journal Of Scientific and Research Publication*, vol. 2, no. 9.

[58] Muzaffar, T. and Choi, T. (2000) "Simplified EZW image coder with residual data transmission", *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on IEEE,* , pp. 111.

[59] Rehna, V. and Kumar, M. (2012) "Wavelet Based Image Coding Schemes: A Recent Survey*", International Journal on Soft Computing (IJSC)* Vol.3, No.3, pp.101-118.

[60] Ahuja, A. and Singh, A. "Image Coding using EZW and QM coder".

[61] http://www.polyvalens.com/blog/?page_id=13.

[62] shingata . V. S and sontakke. T. R (2010) "Still Image Compression Using Embedded Zerotree Wavelet Encoding", *International Journal of Computer Applications*, vol. 1, no. 7, pp. 62-67.

[63] Monro, D. and Dickson, G. (1997) "Zerotree coding of DCT coefficients", *Image Processing 1997, Proceedings International Conference on IEEE*, pp. 625.

[64] Pateriya, K. and Singh, D. (2012) "Comparison of Image Compression Through Neural Network SPIHT, STW, EZW Wavelets", *International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN:* 2249-9555, vol. 2, no. 1.

[65] Zayed, H.H., Kishk, S.E. and Ahmed, H.M. (2012) "3D Wavelets with SPIHT Coding for Integral Imaging Compression", *IJCSNS*, vol. 12, no. 1, pp. 126.

[66] J. Malý, P. Rajmic, "DWT-SPIHT IMAGE CODEC IMPLEMENTA- TION" *Department of Telecommunications, Brno University of Technology, Brno,*

Czech Republic".

[67] Zhang, H. and Fritts, J. (2004) "EBCOT coprocessing architecture for JPEG2000*", Electronic Imaging 2004 International Society for Optics and Photonics,* , pp. 1333.

[68] Sanguankotchakorn, T. and Fangtham, J. (2003) "A new approach to reduce encoding time in EBCOT algorithm for JPEG2000*", TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region IEEE,* , pp. 1338.

[69] Lian, C., Chen, K., Chen, H. and Chen, L. (2001) "Analysis and architecture design of lifting based DWT and EBCOT for JPEG 2000", *VLSI Technology, Systems, and Applications. Proceedings of Technical Papers. 2001 International Symposium on IEEE,* pp. 180.

[70] Wu, C., McCormick, M., Aggoun, A. and Kung, S. (2008) "Depth mapping of integral images through viewpoint image extraction with a hybrid disparity analysis algorithm", *Journal of display technology*, vol. 4, no. 1, pp. 101-108.

[71] Forman, M. and Aggoun, A. (1997) "Quantisation strategies for 3D- DCT-based compression of full parallax 3D images", *DE MONFORT UNIVERSITY.*

[72] Forman, M.C. and Aggoun, A. (1997) "Compression of full parallax integral 3D-TV image data", *Proceedings of SPIE*, vol. 3012 , pp. 222.

[73] Khayam, S.A. (2003) "The discrete cosine transform (DCT): theory and application", *Michigan State University*.

[74] *Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines - ISOIEC - 1992*.

[75] Aggoun, A. and Tabit, M. (2007) "Data compression of integral images for 3D

TV", *3DTV Conference*, 2007 IEEE, pp. 1.

[76] Mazri, M. and Aggoun, A. (2003) "Compression of 3D integral images using wavelet decomposition", *Visual Communications and Image Processing 2003 International Society for Optics and Photonics*, pp. 1181.

[77] Paul, P.J. and Girija, P. (2006) "A novel VLSI architecture for image compression", *Multimedia, 2006. ISM'06. Eighth IEEE International Symposium on IEEE*, , pp. 794-795.

[78] Ezhilarasan, M., Thambidurai, P. (2006) "A Hybrid Transformation Technique for Advanced Video Coding", *LNCS*, vol. 4308, pp. 503-508.

[79] Zhijun, F., Yuanhua, Z. and Daowen, Z. (2003) "A scalable video coding algorithm based DCT-DWT", *Signal Processing and Information Technology, 2003. ISSPIT 2003. Proceedings of the 3rd IEEE International Symposium on IEEE*, , pp. 247-249.

[80] Ibrahimpatnam, R. (2011) "A High Performance Novel Image Compression Technique using Hybrid Transform for Multimedia Applications", *IJCSNS*, vol. 11, no. 4, pp. 119.

[81] Nivedita, P.S., Jindal, S. and Ferozepur, P.I. (2012) "A Comparative Study of DCT and DWT-SPIHT", *IJCEM International Journal of Computational Engineering and Management*, vol. 15, no. 2, pp. 26-32.

[82] Aggoun, A., Mazri,M. (2008) "Wavelet-based compression algorithm for still omnidirectional 3d integral images", *Signal, Image and Video Processing* vol.2, no.2, pp. 141-153.

[83] Aggoun, A. (2011) "Compression of 3D integral images using 3D wavelet transform", *Journal of Display Technology*, vol. 7, no. 11, pp. 586-592.

[84] Tran, T.D., Liang, J. and Tu, C. (2003) "Lapped transform via time- domain pre-and post-filtering", *Signal Processing, IEEE Transactions on*, vol. 51, no. 6, pp. 1557-1571.

[85] http://stackoverflow.com/questions/1955663/dct-compression-block-size-choosing-coefficients.

[86] Dai, W., Liu, L. and Tran, T.D. (2005) "Adaptive block-based image coding with pre-/post-filtering", *Data Compression Conference, 2005. Proceedings. DCC 2005 IEEE*, pp. 73.

[87] Colbert, M. (2005) "Adaptive Block-based Image Coding with Pre-/post-filtering", *university of central Florida.*

[88] Vaisey, J. and Gersho, A. (1992) "Image compression with variable block size segmentation", *Signal Processing, IEEE Transactions on*, vol. 40, no. 8, pp. 2040-2060.

[89] Pandian, S.I.A., Bala, G.J. and George, B.A. (2011) "A study on block matching algorithms for motion estimation", *International Journal*, vol. 3.

[90] Sikora, T. (1997) "MPEG digital video-coding standards", *Signal Processing Magazine, IEEE*, vol. 14, no. 5, pp. 82-100.

[91] He, Z. (2007) "Video compression and data flow for video surveillance", *Texas Instruments, SPRY104.*

[92] Servais, M. and De Jager, G. (1997) "Video compression using the three dimensional discrete cosine transform (3D-DCT)", *Communications and Signal Processing, 1997. COMSIG'97, Proceedings of the 1997 South African Symposium on IEEE*, , pp. 27.

[93] Chen, Y., Hung, Y. and Fuh, C. (2001) "Fast block matching algorithm based on the winner-update strategy", *Image Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 1212- 1222.

[94] Barjatya, A. (2004) "Block matching algorithms for motion estimation", *IEEE Transactions Evolution Computation*, vol. 8, no. 3, pp. 225-239.

[95] Jayaswal, D.J., Zaveri, M.A. and Chaudhari, R. (2010) "Multi step motion estimation algorithm", *Proceedings of the International Conference and Workshop on Emerging Trends in Technology ACM*, , pp. 471.

[96] KOGA, T. and K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro (1981) "Motion-compensated interframe coding for video conferencing", *Proc.NTC81, New Orleans, LA, Nov./Dec.1981*.

[97] Nie, Y. and Ma, K. (2002) "Adaptive rood pattern search for fast block-matching motion estimation", *Image Processing, IEEE Transactions on*, vol. 11, no. 12, pp. 1442-1449.

[98] Xiong, Z., Guleryuz, O.G. and Orchard, M.T. (1996) "A DCT-based embedded image coder", *Signal Processing Letters, IEEE*, vol. 3, no. 11, pp. 289-290.

[99] Strang, G. and Nguyen, T. (1996) "Wavelets and filter banks", *Wellesley Cambridge Press.*

[100] Shensa, M.J. (1992) "The discrete wavelet transform: wedding the a trous and Mallat algorithms", *Signal Processing, IEEE Transactions on*, vol. 40, no. 10, pp. 2464-2482.

[101] *MathWorks T (2008). "Matlab - The Language of Technical Computing." URL http://www.mathworks.com.*

[102] Chun-Lin, L. (2010) "A Tutorial of the Wavelet Transform", *NTUEE, Taiwan.*

[103] Malvar, H.S. (1998) "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts", *Signal Processing, IEEE Transactions on*, vol. 46, no. 4, pp. 1043-1053.

[104] Katkovnik, V., Foi, A., Egiazarian, K. and Astola, J. (2010) "From local kernel

to nonlocal multiple-model image denoising", *International journal of computer vision*, vol. 86, no. 1, pp. 1-32.

Appendix A

Algorithm A consists of 3 step. Step 1: An ascending even index array is created by multiplying the original index array by 2. Step 2: The even index array values are multiplied by the maximum value in the original index array. Step 3: the indexed array is divided by the values in the corresponding AC or DC coefficient arrays.

This Algorithm results with a fraction array which will have enlarge the size more than the indexed file itself although all the values are less than the original values, so a rational fraction function is used. This process results in two arrays, one for the numerator and the other one for the denominator with small integers.

At this point there are three arrays, one for the Negative Positive array and the other two for the index. The size of all of the three arrays is less than the input normalized Coefficient array, and if we take those arrays as an inputs to the to the Huffman coding using DHT, it will ends up with results either the input arrays will be less in size than the output arrays or the outputs of Huffman will be roughly the same size. That means that using Huffman doesn't affect the results with the Negative Positive algorithm.

For the AC Values, the same algorithm is applied but with a small change, instead of creating even array at the beginning, a normal ascending array will be created and this is due to the nature of the DC array after applying DPCM.

The drawback of this algorithm is that it got a percentage of loss that is make it a lossy algorithm. This loss came from the step that the rational fraction function is applied, as this step is not always reversible.

2.3.2 Second Algorithm, instead of creating new ascending array either normal or even, a similar concept as DPCM but in this case the input is ascending array with positive values.1) flip the 1-D array to start with the bigger value 2) then subtracting the second value from the preceding one and store the differences with keeping the first value.

2.3.3 Third Algorithm, Prime Algorithm factorized the index array in to two arrays then added the prime numbers by one and save it in another array.

Appendix B

B.1 Discrete Wavelet Transform Overview

As mention before, the Most Transform coding level is reversible, non-lossy step Like DFT, DCT, the aim of the transform is to have a de-correlated coefficients [56]. Another type of transform coding that is also popular and well known is the Discrete Wavelet Transform.

It converts the *spatial* input signal in time-frequency domain [98]. One of the main good advantages of the DWT is it is multi-resolution technique meaning that different

frequencies can be analyzed in different resolution. DWT uses number of wavelet to represent an Image [7].

The wavelet according to [99] is a small wave, "wavelets are localized waves. Instead oscillating forever, they drop to zero".

The wavelet localization property of the DWT provides advantage over DCT, which is non-localize Transform, as it going to be more explained in Chapter (7).

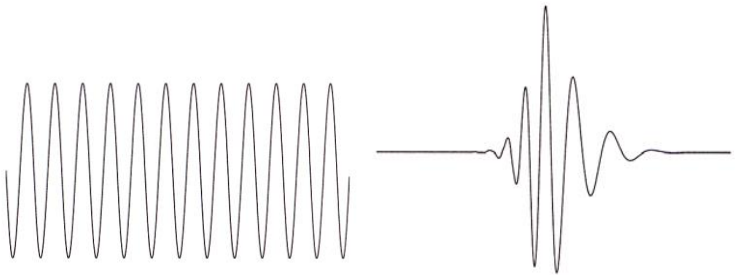


Figure B.1 Localization

The Continuous Wavelet Transform

The basis function sometimes referred to Mother Wavelet, is the function that is used to calculate and create all the wavelet functions. This occurs by Shifting and scaling [100][101].

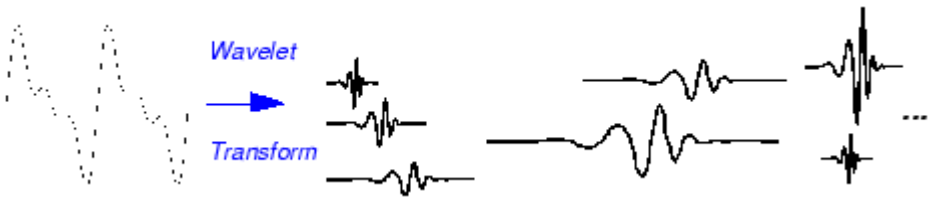


Figure B.2 represents the signal and CWT [101]

The basis function (mother wavelet) depends on the shifting parameter and the Scale parameter. For the basis function to works effectively it needs to determine the location of the wavelet in the input signal as well as the Scaling factor. This is relating to the Time and the frequency respectively.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure B.3 the different scale wavelet 1,2,4

$$S=1/f$$

Where, S is the scaling Factor

F is the frequency

Large scale with Low frequency stretching the Image and give the option to realize more details and make small data more clear while small scale with high frequency shrank the image or compress it in order to give a chance to have a overall view for the image.

Figure B.4 represents the Low scale and large scale [101]

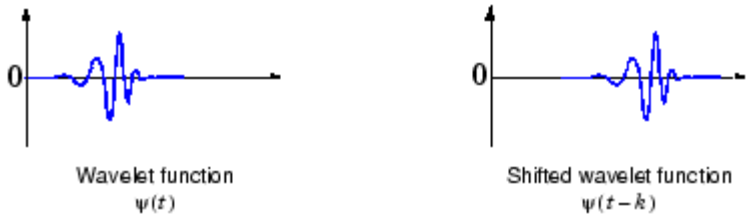


Figure B.5 Shifting wavelet

Briefly, High frequency indicates low scale; compressed wavelet that gives quickly changing details, while Low frequency indicates large scale with gradually changes [101].

B.2 The Discrete Wavelet Transform:

Figure (B.6) shows the difference between the CWT and the DWT.

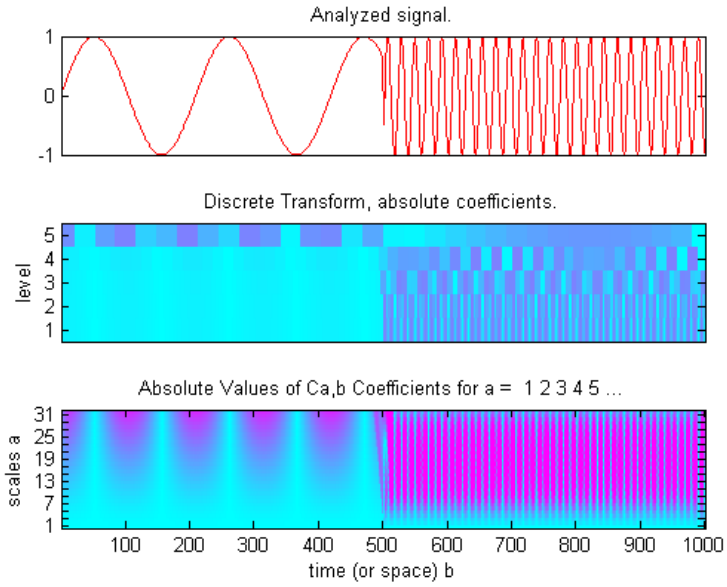


Figure B.6 the CWT and the DWT.

In the Continuous Wavelet transforms, the basis function is used to analysis the signal using simple scaling and shifting. On the other side the Discrete Wavelet Transform is using a digital filters with various Cut-offs and frequencies to analysis the digital signals [100].

The main feature of the DWT is the ability to represent the Image in different resolution, this property referred to multi-resolution, it happened through using different filters, while the scaling option occurs by using up-sampling and down-sampling technique.

The low- frequency contains the important data. It is what gives the signal its identity, while the high-frequency (the low scale) gives the flavour instants [101] it is more about the details that can be removed.

1D-DWT

The DWT filters consist of two filters, Low-pass filter and High-pass filter.

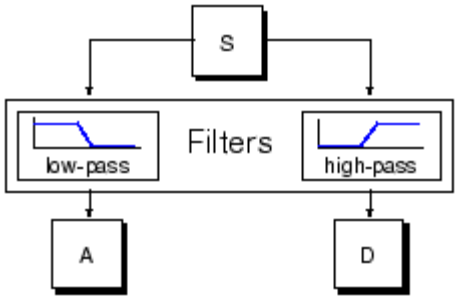


Figure B.7 Low and High-pass filters.

During the filtering process the, input data become double samples than the input signal, So an up-sampling in the composition process and down-sampling in the decomposition process is needed in order to return back to the actual size.

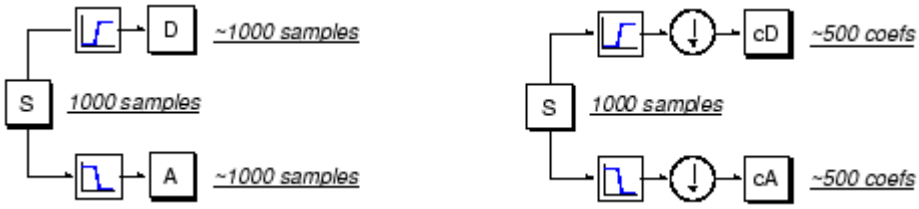


Figure B.8 up-sampling and down-sampling.

Where, cD referred to High coefficient, cA Low Coefficient, as illustrated in Figure (B.7). For 1D-DWT, the output from low pass filter is an approximate coefficient and the output from the high pass filter is a detail coefficient [7].

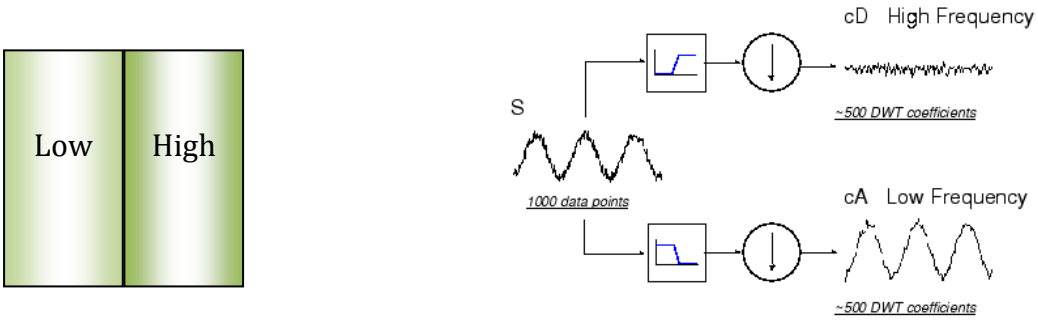


Figure B.9 High, Low Coefficients

2D-DWT

The output from 1D-DWT is a High coefficient and Low coefficient, while using 2D-DWT results with four frequencies coefficients, Low-Low, High-Low, Low-High, and High-High.

Low-High coefficient comes as an output from low pass filter applied to row and a high pass filter applied to column. Hence, LH signal contains horizontal coefficients (cH) and HL contains vertical coefficients (cV), finally the HH represent the diagonal coefficients (cD) [7], where, L, H indicates Low Pass and High Pass filters respectively. As in the 1D DWT, the outputs from each branch are down sampling by two.

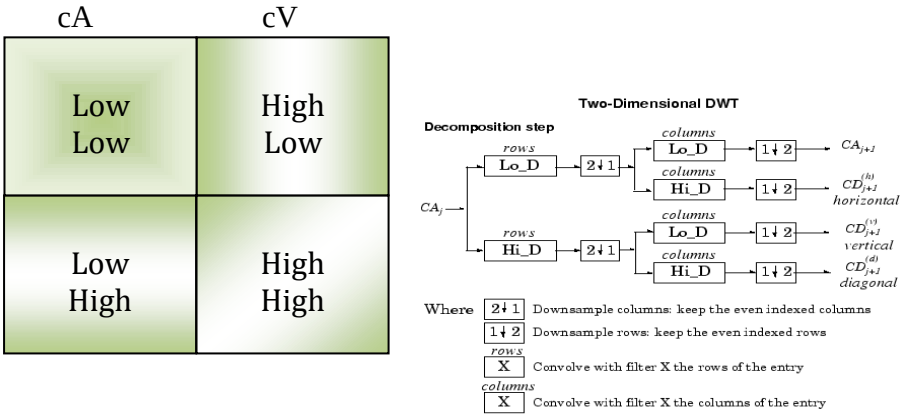


Figure B.10 2D-DWT coefficients.

Decomposing the Low-Low coefficient into more levels will lead to the multi-resolution property, as illustrated in Figure (B.9). The important data is in the Large Scale with Low frequency, so high compression ratio can be achieved by neglecting one of the High Pass coefficients; the more high Pass coefficient abandon the better compression ratio achieved, meaning passing over the HH, HL, LH gives better compression.

The processed signal is either stored or transmitted. For most compression applications, processing involves quantization and entropy coding to yield a compressed image. During this process, all the wavelet coefficients that are below a chosen threshold are discarded. These discarded coefficients are replaced with zeros during reconstruction at the other end. To reconstruct the signal, the entropy coding is decoded, then quantized and then finally Inverse Wavelet Transformed.

Classification of wavelets

There are two types of wavelets, either Orthogonal or Bi- Orthogonal [100].

Orthogonal wavelet filter

Both filters got the same length but they are not symmetric, and their coefficients are real number. The relation between the low pass filter and the high pass filter is

$$H_0(z) = z^{-N} G_0(-z^{-1})$$

Both low and high pass filters are perpendicular to each other meaning that they are swapped flip to each other. If the result of Shifting by two between the two filters is zero that indicate perfect reconstructed data.

Also to get a perfect reconstructed data, the synthesis filters got to be the same as the analysis apart from the time. One of the orthogonal features is that they got a customary construction (regular structure) that makes it easy to implement table B.1 shown Daubechies 9/7 analysis and synthesis for both low and high pass filters.

Table 3. Daubechies 9/7 Analysis and Synthesis Filter Coefficients.

Analysis Filter Coefficients		
i	Low-Pass Filter $h_L(i)$	High-Pass Filter $h_H(i)$
0	0.6029490182363579	1.115087052456994
±1	0.2668641184428723	–0.5912717631142470
±2	–0.07822326652898785	–0.05754352622849957
±3	–0.01686411844287495	0.09127176311424948
±4	0.02674875741080976	
Synthesis Filter Coefficients		
i	Low-Pass Filter $g_L(i)$	High-Pass Filter $g_H(i)$
0	1.115087052456994	0.6029490182363579
±1	0.5912717631142470	–0.2668641184428723
±2	–0.05754352622849957	–0.07822326652898785
±3	–0.09127176311424948	0.01686411844287495
±4		0.02674875741080976

Table B.1 Daubechies 9/7 analysis and synthesis

Bi-orthogonal wavelet filter

The Bi-Orthogonal wavelet filters is contrary to the orthogonal in all of its features, both low and high pass filters are not the same length like the orthogonal, also the high pass filter could be symmetric or not but the low is always symmetric, final difference is that the coefficients of bi- Orthogonal filters could be integers or real numbers.

Wavelet Families

As it mention before in the CWT, the mother wavelet (basis function) is the function that control the process depending on shifting and scaling, so selecting a proper Mother wavelet results on achieving effective Wavelet Transform [100].

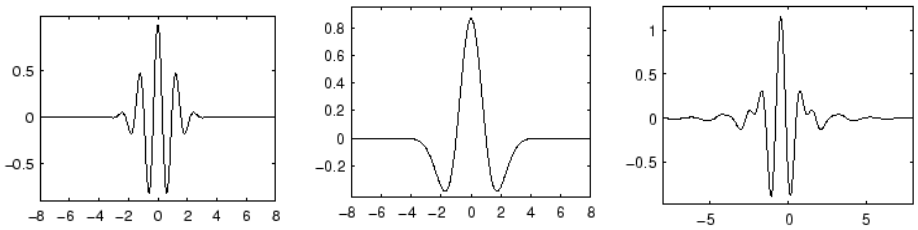
Year	*Wavelet name*
1910	Haar families.
1981	Morlet wavelet concept.
1984	Morlet and Grossman "wavelet".
1985	Meyer "orthogonal wavelet".
1987	International conference in France.
1988	Mallat and Meyer multiresolution.
1988	Daubechies compact support orthogonal wavelet.
1989	Mallat fast wavelets transform.

Table B.2 History of Wavelets Transform [102].

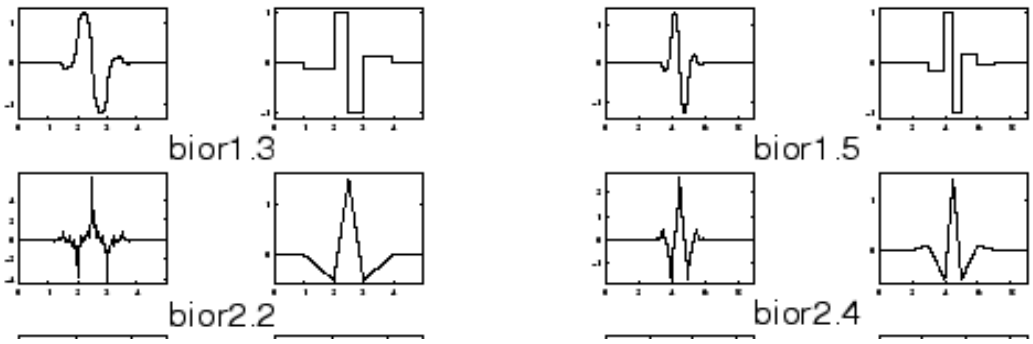
In 1910, Alfred Haar represented ***Haar wavelet***; it was the first time the wavelet exists, and till now it is the simplest wavelet [100]. Then in 1981 Morlet restart searching and representing the wavelets [102], since that time a number of wavelets have been proposed as shown in Table B.2.

One of the most common wavelets is ***the Daubechies wavelet*** Known also as *Maxflat wavelets* because they got the maximum flatness at frequencies 0 and π. This family is well known and used extendedly among the other wavelets.

The ***Symlets*** and ***Coiflets*** are efficiently supported orthogonal wavelets like the Haar and the Daubechies. ***Meyer wavelets*** and all the previous wavelets are capable of perfect reconstruction. The ***Meyer, Morlet*** and ***Mexican Hat wavelets*** are symmetric in shape [100].



Morlet Mexican Hat Meyer



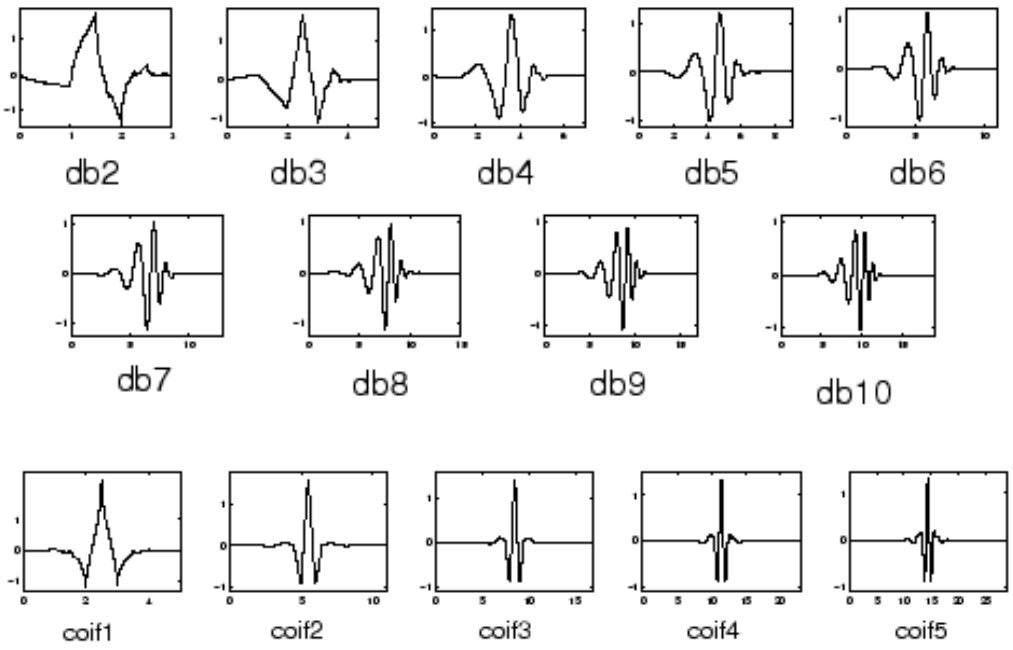


Figure B.11 Wavelet Families.

Appendix C

There are two types of transform techniques that could be applied in order to solve these two problems, first the Global Transform (non-Local) like DWT and Lapping Transform and second, the Local Technique based on Blocking but in an adaptive way.

Using the Global Technique like DWT which is applied in JPEG2000, works efficiently in reducing the Blocking artefacts, but it still facing a ringing artefacts due to the long basis especially in the areas that is have lots of edges and also another draw back is because it is a global transform it is not exploiting the local correlations in the areas with low frequency in the background of the image which leads to improper compression ratio [86].

Another Global Transform that can be applied is the Lapping Transform. It is very efficient in reducing and eliminating the Blocking artefacts but still facing ringing artefacts and even more than the DCT due to its long basis [103].

So, briefly there is a trade between the ringing artefacts and the blocking artefacts, either to have a ghost affect that will appear on the edge of the image or as a pre-echo in audio or to have a visual discontinuity effect [87].

The long basis of the global transform (non-Local) makes it suffers from the ringing artefacts and on the other side the local transforms systems suffers from Visual discontinuity due to the Blocking technique especially on the edges in the horizontal and vertical boundaries.

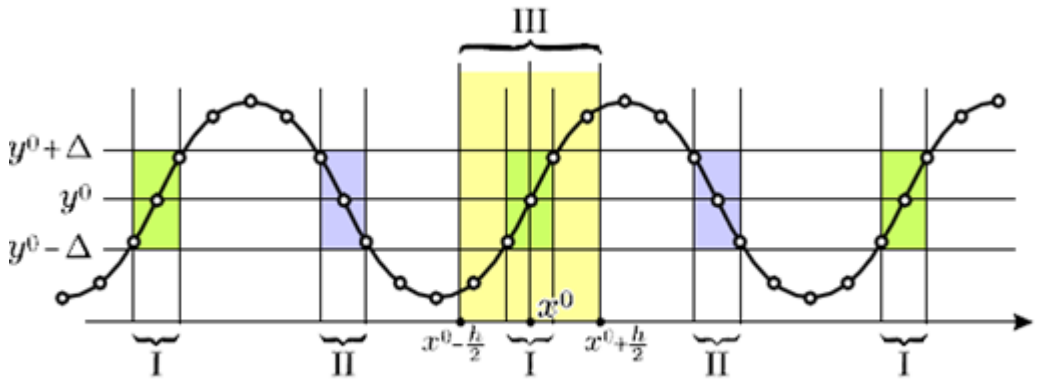


Figure C.1 Local & non-Local model [104]

Papers

- Amar Aggoun, Amal Mehanna, Emmanuel Tsekleves, John Cosmas, Jonathan Loo (2010) Live Immerse Video-Audio Interactive Multimedia. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2010 (BMSB 2010). 24-26 March 2010, Shanghai, China. IEEE.

- A. Mehanna, A. Aggoun, O. Abdulfatah, M. R. Swash, E. Tsekleves (2013) Adaptive 3D-DCT based compression algorithms for Integral Images. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2013 (BMSB 2013).