Investigating Wavelet Transform Technique for Transmitting Video Information over Mobile Devices

تطوير تقنية تحويل المويجة لنقل المعلومات الفديوية عبر جهاز المحمول

Prepared by
Talal Ali Hammouri

Supervised by
Prof. Muzhir Shaban Al-Ani

A Dissertation submitted in partial fulfillment of the requirement for the degree of Doctor of philosophy in Computer Science

Department of Computer Science
College of Computer Science and Informatics
Amman Arab University

April 2012

التفويض

انا طلال علي كريم الحموري افوض جامعة عمان العربية بتزويد نسخ من رسالتي للمكتبات او الشركات او الهيئات او الاشخاص عند طلبها.

الاسم : طلال علي كريم الحموري

التوقيع :

التاريخ :2012/6/16

DISSERTATION COMMITTEE APPROVAL

Date: 21-April-2012
Name: Talal Ali Kareem Hassan Al-Hammouri
Degree: Doctors of Philosophy in Computer Science
Title: "Investigating of Wavelet Transform Technique for Transmitting Video Information over Mobile Device"

This Dissertation, with the Title" Investigating of Wavelet Transform Technique for Transmitting Video Information over Mobile Device", was Presented, Examined and Approved on : 21-April-2012.

Examining Committee	**Title**	**Signature**
Prof. Nadeem Ali Abeid	Chair	
Prof. Muzhir Shaban Al-Ani	Supervisor & Member	
Prof. ALaa Hussein Al-hamami	Member	Alahamami
Mohamad Sleiman Qatawneh	Member	

Amman Arab University
April 2012.

Table of Contents

Table of Contents IV
List of Figures VIII
Table XI
List of Abbreviations XI
List of Tables XI
Acknowledgments XIV
Abstract XV
Arabic summary XVII
Chapter One Introduction 1
  1.1. Overview 1
  1.2 Practical needs for Image and Video Compression 2
  1.3 Video Compression 4
  1.4 Literature Reviews 4
  1.5 Statement of the Problem 14
  1.6 Objectives 15
  1.7 Scope of the Thesis 16
Chapter Two Image Compression 17
  2.1 Introduction 17
  2.2 Image Compression 18
  2.3 Compression Approaches 21
  2.3.1 Run Length Coding 22
  2.3.2. Huffman Coding 22
  2.3.3 Arithmetic Coding 23
  2.3.4. Lempel/Ziv Coding 24
  2.3.5 Area Coding 24
  2.4 Wavelet Image Compression 25
  2.4.1.Wavelet Families 29
  2.4.2 Choice of Wavelet Families 33

2.4.3 Two Dimensional Wavelet Masks.. 34
2.5 Fourier Transform.. 35
2.6.Discrete Wavelet Transform 38
2.7.Two Dimensional DWT.. 40
2.8 An Overview of Image Compression Standard....................... 43
2.9 Performance Measurement 49
2.9.1. Subjective Quality Measurement.. 49
2.9.2. Objective Quality Measurement.. 50
Chapter three video compression.. 52
3.1 Introduction 52
3.2 How Video Compression Works.. 58
3.3 Basic Methods of Reducing Data 59
3.4 Understanding Frames.. 60
3.5 Fundamentals of Video Compression..................................... 62
3.5.1 Removal of Spectral Redundancy 63
3.5.2 Removal of Spatial Redundancy 63
3.5.3 Removal of Temporal Redundancy 68
3.5.4 Removal of Statistical Redundancy..................................... 69
3.6 Digital Video/Image Coding Standard..................................... 69
3.6.1 JPEG Standard 74
3.6.2 JPEG2000.. 74
3.6.3 MPEG -1 75
3.6.4 MPEG-2 75
3.6.5 MPEG-4 76
3.6.6 H.261 76
3.6.7 H.263, H.263 Version 2 (H.263+), H.263 ++, and H.26L...... 76
3.6.8 MPEG -4 Part 10 Advanced Video Coding or H.264 /AVC .. 77
3.6.9 VC-1 77
3.6.10 Real Video.. 77
3.7 Video Quality Analysis.. 79

3.7.1 Subjective video assessment 80
3.7.2 Objective metrics of video quality 80
Chapter four system design... 82
4.1 Image Compression Approach 82
4.1.1 Images contents 83
4.1.2 Two Dimensional DWT.. 84
4.1.3 Two Dimensional Wavelet Masks....................................... 85
4.1.4 Performance Measurements 85
4.1.4.1 Objective Measure 86
4.1.4.2 Subjective Measure... 87
4.1.5 Image Compression System 89
4.2 Video Compression System 95
4.2.1 Video Content.. 96
4.2.2 Two Dimensional DWT.. 98
4.2.3 The Implemented System of Video Compression.............. 101
4.2.4 Frames Difference Approach.. 103
4.2.5 Video Compression Steps 105
4.2.6 Video Compression Approaches 112
Chapter Five Implementation and Results.................................... 120
5.1 Image Compression 120
5.1.1 Wavelet Families.. 121
5.1.2.Two Dimensional Wavelet Masks.................................... 121
5.1.3 Performance Measurements of Image Compression......... 121
5.1.4 Image compression results and analysis 121
5.1.4.1 Objective Results of Image Compression....................... 126
5.1.4.2 Subjective Results of Image Compression 138
5.2 Video Compression... 144
5.2.1 Frames difference approaches 145
Chapter 6 Conclusion and Future Works...................................... 177
6.1 Introduction 177

6.2 Conclusion 177
6.3 Future works 180
References.. 182

List of Figures

Figure	Page
(1.1): Image and video compression for visual transmission and storage	3
(2.1): Compression Framework	17
(2.2): Wavelet coding system: a) Encoder b) Decoder	24
(2.3): Haar wavelet	26
(2.4): Daubechies wavelets	27
(2.5): Biorthogonal wavelets	27
(2.6): Coiflets Wavelets	28
(2.7): Symlets Wavelets	28
(2.8): Morlet Wavelet	28
(2.9): Mexican Hat Wavelet	29
(2.10): Meyer Wavelet	29
(2.11): The two –level sub band decomposition used in the DWT	35
(2.12): Decomposition of image applying DWT	38
(2.13): Block diagrams of the JPEG2000 (a) encoder and (b) decoder	40
(3.1): Temporal structure of a video sequence	47
(3.2) Motion video sequence divided into information units	50
(3.3): Group of pictures "inter frame dependencies in a stream"	55
(4.1): Typical image compression system	76
(4.2): Decomposition of image applying DWT	78
(4.3): Implemented mask selection compression technique	84
(4.4) levels of Decomposition for Image compression.	85

(4.5): Different types of tested images	85
(4.6): Structural Diagram of Image Compression	87
(4.7): Basic decomposition steps of 2-Dimnsional DWT image	92
(4.8): The implemented frame selection key approach	95
(4.9): Color Image and its color histogram	98
(4.10): Structural Diagram of Video Compression	102
(4.11):(a):frames difference (b):the removed zero frames difference	104
(4.12): Structure Diagram of the first approach (Zero Difference approach)	105
(4.13): Frames difference (Mean Difference approach)	106
(4.14): Structure Diagram of the second approach (Mean Difference	107
(4.15): Structure diagram of the proposed frames difference percentage	108
(4.16): Structure Diagram of the third approach (Percentage Difference	109
(5.1): Original Plant image	112
(5.2): The compression image results using different wavelet families	113
(5.3): MSE for Plant Image	114
(5.4): PSNR for Plant Image	115
(5.5): Different types of images	115
(5.6): PSNR Measures	119-124
(5.7): MSE Measures	124-127
(5.8): Original, Compressed and Decompressed images	129-130
(5.9) Mean Opinion Scores for the test images	132
(5.10): Zero difference approach (10frames /second)	135-136
(5.11): Zero difference approach (15frames /second)	136-137
(5.12): Zero difference approach (20 frames /second	138
(5.13): Zero difference approach (25 frames /second)	139
(5.14): Zero difference approach (Scores)	141

(5.15): Mean difference approach (10frames/second)	142-143
(5.16): Mean difference approach (15frames/second)	143-144
(5.17): Mean difference approach (20frames/second)	145
(5.18): Mean difference approach (25frames/second)	146
(5.19): Mean difference approach (Scores)	147
(5.20): Percentage difference approach (10frames/second)	148-151
(5.21): Percentage difference approach (15 frames/second)	152-153
(5.22): Percentage difference approach (20 frames /second)	154-156
(5.23): Percentage difference approach (25 frames /second)	157-159
(5.24): Percentage difference approach Scores (10%)	161
(5.25): Percentage difference approach Scores (15%)	162
(5.26): Percentage difference approach Scores (20%)	163

List of Tables

Table	Page
(2.1): Some of the wavelet transforms masks	31
(3.1) Characteristics of a Variety of Computer Video Formats	49
(3.2): List of Some Organizations for Standardization	65
(3.3): Video/Image Coding Standards	66
(4.1): Mean Opinion Score (MOS)	81
(4.2): Image Quality Factors and Scores	82
(4.3): Video Quality Factors and Scores	100
(5.1): Experimental Results for test image Plant	114
(5.2): Image information's	116
(5.3): MSE & PSNR results for different types of images.	117-118
(5.4) Subjective Image quality results	131
(5.5) Video quality assessment (zero difference approaches)	140
(5.6) Video quality assessment (Mean difference approach)	147
(5.7) Video quality assessment (percentage difference approach (10%))	160
(5.8) Video quality assessment (percentage difference approach (15%))	162
(5.9) Video quality assessment (percentage difference approach (20%))	163

List of Abbreviations

2-DWT: Two-Dimensional Discrete Wavelet Transforms

3-D:Three-Dimensional

ATM: Asynchronous Transfer Mode

CCITT: International Telegraph and Telephone Consultative Committee

CIF: Common Intermediate Format

CPB: Cycles per Byte

CR: Compression Ratio

CW: Coiflets Wavelet

DBS: Direct Broadcasting System

DCT: Discrete Cosine Transform

DSP: Digital Signal Processors

DVD: Digital Versatile Disk

DW: Daubechies Wavelet

DWT: Discrete Wavelet Transform

FFT: Fast Fourier Transfer

FIR: Finite-Impulse Response

FMT: Faber-schauder Multi-scale Transform

FPS: Frames per Second

GOP: Group of Pictures

GSM: Global System for Mobile Communications
H.26X: Family of video coding standards published by the International Telecom Union

HDTV: High Definition Television

IEC: International Electro technical Commission

ISO: International Organization for Standardization

ITU: International Telecommunication Union

ITU-T: International Telecommunication Union

JPEG: Joint Photographic Experts Group

LAN: Local Area Network

LMSE: Laplacian Mean Square Error

LVF: Light weight Video Format

MD: Maximum Difference

MOS: Mean Opinion Score

MPEG: Motion Picture Experts Group

MSE: Mean Squared Error

NAE: Normalized Absolute Error

NTSC: National Television Standards Committee

PAL: Phase Alternating Line

PAL: Phase Alternating Line

PDE's: Partial Differential Equations

PSNR: Peak Signal-to-Noise Ratio

QMF: Quadrature Mirror Filters

QoS: Quality of Services

RGP: Red, Green, and Blue color space

ROI: Region of Interest

SC: Structural Content

SECAM: Sequential Color With Memory

SMPTE: Society of Motion Picture and Television Engineers

STDV: STandard-Definition Television

TTC: Tone Transfer Curve

TV: Television

VGA: Video Graphics Array

WMV: Windows Media Video

YUV: Color space- Y is luminance component, and U and V are chrominance components

Acknowledgments

Foremost, I would like to express my deepest gratitude to my advisor ***Prof. Muzhir Shaban Al-Ani*** for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study. This thesis would not have been possible without the help, support and patient of my Dean, ***Prof Alaa Al-Hamami*** not to mention his advice and unsurpassed knowledge but on both academic and a personal level, for which I am extremely grateful.

Besides my advisor, I would like to thank the rest of my thesis committee: *Prof. Nadeem Ali Abeid,* (Chairmen)*, and Prof. Mohamad Sleiman Qatawneh* for their encouragement, insightful comments, and hard questions.

I would like to thank my wife Ola for her personal support and great patience at all times.

I will be extremely grateful to my father, mother, brothers and sisters;

for all that they have done for me.

For any errors or inadequacies that may remain in this work, of course, the responsibility is entirely my own.

Abstract

The huge usage of digital multimedia via communications, wireless communications, Internet, Intranet and cellular mobile leads to incurable growth of data flow through these Media. The ability to transmit video and support related real-time multimedia applications are considered an important issue in mobile networks. Video streaming, video conferencing, online interactive gaming and mobile TV are applications expected to support the viability and survival of next generation mobile wireless networks.

Wavelet transform is an efficient method that can be used to perform an efficient compression technique. This dissertation deals with the developing of an efficient video compression based on frames difference approaches that concentrated on the calculation of frames near distances (difference between frames). The selection of the meaningful frame depends on many factors such as compression performance, frame details, frame size and near distance between frames. Three different approaches are applied for removing the lowest frame difference (Zero difference Approach, Mean Difference Approach and Percentage Difference Approach). Many factors are considered in system implementation to adapt image and video compression over mobile devices. In this work, many images and video contents are tested to insure the efficiency of this technique, in addition a good performance result has been obtained, in which the best MSE & PSNR gives by Coiflets wavelets transform (Coif4) 133 & 26.92 respectively. Also with 25 frames per second extracted, good results of the compressed video are obtained using subjective video quality assessment in which Mean

Opinion Score (MOS) gives on the average (4.5/5) for the case of removing all zero difference frames and the lowest 5% percentage of the frames difference of the extracted frames.

Arabic summary

الخلاصة

ادى الاستخدام الهائل للوسائط المتعددة الرقمية من خلال الاتصالات والاتصالات اللاسلكية والانترنت والانترانت والهاتف الخلوي، إلى وجود كم هائل من البيانات المتدفقه خلال هذه الوسائط. إن القدرة على نقل الفيديو والدعم ذات الصلة في نقل المعلومات الفديويه بشكل مباشر تعتبر مسألة هامة في شبكات المحمول. تدفق الفيديو ومؤتمرات الفيديو والألعاب التفاعلية على الانترنت والتلفزيون المحمول هي تطبيقات متوقع منها تحسين واستمرارية شبكات الجيل القادم من الشبكات اللاسلكية المحموله.في الآونة الأخيرة، أصبحت تقنيات ضغط الفيديو وتطبيقاتها في مجالات عديدة (التعليمية ، والزراعة ، والطبية...) اكثر المجالات الحيوية والاكثر اهمية. تعتبر تقنية تحويل المويجات هو وسيلة فعالة يمكن استخدامها لأداء تقنية ضغط فعالة. تقدم هذه الاطروحه طريقه مقترحه في ضغط الصور والفيديو والتي تعتمد على أساس الفروقات بين الصور المقطعة والتي ركزت على حساب قرب المسافات (الفرق بين مشاهد الصور). اختيار طريقة الضغط تعتمد على عوامل كثيرة مثل ضغط الأداء ، وتفاصيل الصورة ، وحجم الصورة ، وقرب المسافة بين مشاهد الصور. تم تطبيق ثلاثة أساليب لإزالة المشاهد المتشابهة. في هذا العمل ، **هناك العديد من العوامل اخذت بعين الاعتبار في تنفيذ وتهيئة ضغط الصور والفيديو عبر اجهزة المحمول.** تم اختبار العديد من الصور و الفيديوهات المختلفة للتأكد من كفاءة هذه التقنية ، حيث تم الحصول على نتائج جيده. البعض من تقنيات ضغط الصور تم اختبارها على انواع مختلفه من الصور حيث تم ضغطها واختبارها. حيث اظهرت اللنتائج ان عائلة Coiflets wavelets transform (Coif٤) ذات نتائج جيده بالمقارنة مع الأنواع الأخرى وأظهرت التحليلات إن قيمة MSE مساوية الى (133) في حين قيمة PSNR مساوية إلى (26.92) . كذلك تم ضغط الفيديو بعد تقطيعه الى 25 مقطع لكل ثانية وتم الحصول على نتائج جيده باستخدام مؤشر التقييم الشخصي حيث سجل مؤشر MOS في المتوسط علامه (4.5 / 5) وذلك في حال شطب المقاطع التي تعطي فروقات صفريه فيما بينها. وكذلك في حال شطب المقاطع ذات الفروقات الصغرى فيما بينها التي اقل من 5% من فروقات المقاطع التي تم احتسابها.

Chapter One
Introduction

1.1. Overview

The ability to transmit video and support related real-time multimedia applications is considered an important issue in mobile networks. Video streaming, video conferencing, online interactive gaming and mobile TV are applications expected to support the viability and survival of next generation mobile wireless networks. Therefore, it is significant to analyze the interaction of the particular media and applications [1].

The basic idea behind the success and the distribution of third generation mobile networks is the availability of attractive, useful, and low cost services for the final user. Recently, an increasing demand for mobile equipments and playback video over mobile devices were observed. Video over mobile devices has been an important media for communications and entertainment for a long time.

Video communication over a dynamic environment, such as a mobile and wireless network is much more difficult than over static channel, since the bandwidth, delay, and loss are not known in advance and are unbounded [2].

Mobile phones are really very important in our daily life. It is a long range, portable electronic device used for mobile communication. Mobile phones are affordable, easy to use and comfortable. They manage our task in industry, office and home.

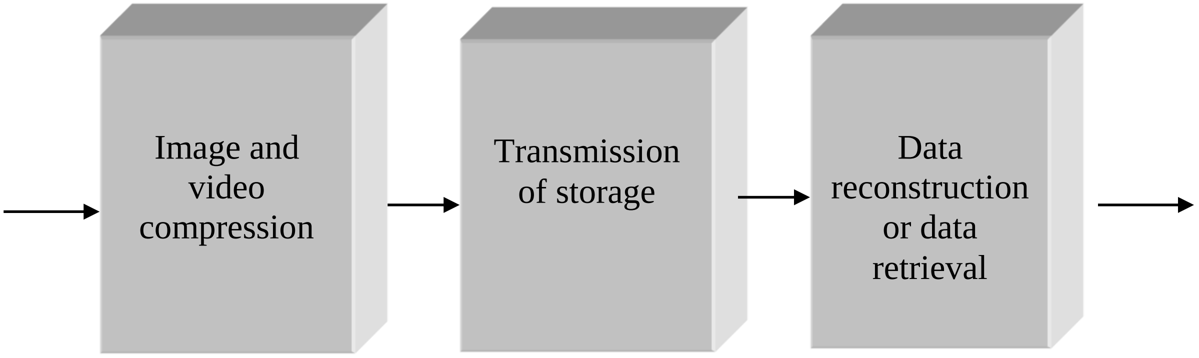
Mobile phones with video cameras and the ability to transmit and play videos are rapidly becoming popular and all over the word available. Since it is a modern –day wireless technology that has drastically change how people communicate, create, access, and share information [2].

1.2 Practical needs for Image and Video Compression

Needless to say, visual information is of vital importance for human beings to perceive, recognize, and understand the surrounding world. With the tremendous progress that has been made in advanced technologies, particularly in very large-scale integrated circuits, increasingly powerful computers and computations, it is becoming more possible than ever for video to be widely utilized in our daily life. Examples include video phony, video conferencing, high definition TV (HDTV), and digital video disk [3].

Video as a sequence of video frames, however, involves a huge amount of data. With increasingly demanding video services, such as three-dimensional (3-D) movies and games, and high video quality, such as HDTV, advanced image, and video data compression is necessary. Compressions become an enabling technology to bridge the gap between the required huge amount of video data and the limited hardware capability.

Image and video compression is not only necessity for rapid growth of digital visual communication, but is also feasible. Its feasibility rests with two types of redundancies, i.e. statistical redundancy and psychovisual redundancy. By eliminating these redundancies, one can achieve image and video compression. Figure 1.1 shows the functionality of image and video data compression in visual transmission and storage [3].



Input Output

Figure 1.1 Image and video compression for visual transmission and storage [3].

The required quality of the reconstructed image and video is application dependent. In medical diagnosis and some scientific measurements, we may need the reconstructed image and video to mirror the original image and video. In other words, only reversible, information-preserving schemes are allowed. This type of compression is referred to as **lossless** compression. In applications, such as motion picture and television (TV), a certain amount of information loss is allowed. This type of compression is called **lossy** compression. One can see that image and video data compression involves several fundamental concepts including information, data, visual quality of image and video and computational complexity [35].

1.3 Video Compression

Video compression refers to reducing the quantity of data used to represent digital images, and is a combination of spatial image compression and temporal motion compensation. In this work a video compression technique that reduces the size of bandwidth and size of data storage to facilitate the transmission operation over mobile devices will be illustrated.

Digital video compression is implemented because it makes the process of transmission fast and reliable. Compressed video is easy to store and download at low bandwidth rate. Digital compression techniques can compress video without affecting the quality of video. It converts it into parts that humans are unable to detect. For example, there are numberless colors but human can perceive 1024 shades. Video compression is the process of reducing data for such things that we cannot notice. Standard video compression cameras compress video data at a ratio of 5 to 1. There are different types of formats that can help you to compress video as much as you can. But too much compression is also not recommended. When you compress data too much, it throws away more data that can bring noticeable change in the video. Excessive compression can make video unrecognizable [3].

1.4 Literature Reviews

Many related pieces of research and papers are published; some of them are explained as the following:

P.G Santha and Arthanariee [4] proposed a Simple Wavelet Compression (SWC) processing of image coding to maximize compression and minimize energy cost in SWC. That gives an improved polyomines losses compression techniques and efficient wavelet compression, both increases the quality of image at receiving end of the wireless of communications by reducing the Peak to Signal Noise Ratio and Mean Square Error. The authors selected two parameters of the efficient wavelet image compression algorithm to vary, and presented the results of modifying the parameters on quality of image, computation and communications efficiency with respect to energy utilization.

Baluram et al [5] focused on selecting the most appropriate wavelet function for a given type of image compression and studied the behavior of different types of wavelet function with different type of images and suggested the most appropriate wavelet function that can perform optimum compression for a given type of images. The wavelet function that gives the maximum compression for a specific type of image will be the most appropriate wavelet for that type of image compression. Image quality is measured objectively, using peak signal-to-noise ratio or picture quality scale, and subjectively, using perceived image quality. The effects of different wavelet functions, image contents and compression ratios were assessed in the paper.

Sherin Kishk et al [6], proposed a technique that relies on applying the principle of component analysis, PCA, on the wavelet coefficients of the elemental images to improve the quality of the recovered 3D image while achieving high compression ratio.

The wavelet coefficients of the individual elemental images are stacked and rearranged before applying PCA compression. The PCA compression is applied to each sub-band individually to enhance the compression ratio. The quality of the reconstructed 3D images and received elemental images are calculated. Results show high compression ratio compared to PCA alone compression while maintaining the recovered 3D image quality. PSNR is used to measure the reconstructed 3D image quality.

Hojung Cha and Jongmin Lee [7] presented comprehensively the design and implementation details of a system architecture that makes video streaming possible on the state-of-the art cell phones and the mobile networks. The architecture has been actually deployed and presently services a large number of cell phones. The mobile phone is equipped with a browser and a lightweight video player with this phone, a subscriber connects to the server that contains video contents. A video clip is selected from the menu, and the clip is streamed or downloaded from the streaming server. The content authoring system converts video in existing format into our proprietary LVF (Light weight Video Format) file suitable for the video player running in a cell phone, the authoring system for the player, and the server system for pumping bit streams to the cell phone. Also in this research they have analyzed several video compression algorithms in terms of the processing requirements and the compression ratio, in which they developed a wavelet-based video codec, LVF, which reduces the computational complexity of video coding/decoding so as to run it on existing cell phones.

Ahmad & Azrol [8] explained the criteria for video transmission to mobile phone over wireless networks, in which they mentioned that there are many ways to develop multimedia application for mobile phones, but the issue is that target device and also the feeder application store in a server should support standard media format and acceptable to all mobile phone players, i.e. the best practice to deploy video application for mobile is that any server should deliver all media format to any type of mobile phone with standard video player.

M. D. Walker et al [9] have discussed a number of well-known and standardized techniques for audio/video-streaming. Also the paper has described the Fast nets multimedia streaming architecture for mobile IP networks. The system makes use of a hierarchy of independent video streams encoded according to ITU-T Recommendation H.264 and bit-rate adaptive transmission techniques to match the transmitted multimedia data to the instantaneous network capability.

Satyajayant Misra and others [10] presents a survey concerning multimedia streaming in wireless sensor networks; they categorize the requirements of multimedia traffic at each layer of the stack. Network protocol stack and further classify the mechanisms that have been proposed for multimedia streaming in wireless sensor networks at each layer of the stack. They also review the existing cross-layer approaches and propose a few possible cross-layer solutions to optimize the performance of a given wireless sensor network for multimedia streaming applications.

Muzhir et al [11] proposed a new communication system to integrate the use of GSM over the available satellite infrastructures in which the proposed communications system could be used to facility and get benefits of both systems (GSM and Satellites) to achieve competitive services over the world. the proposed system is concentrated on a global communication system that served all over the world and gave some specialization and privacy for each country, also this paper shows an algorithm on how to implement the GSM over satellite systems in an efficient, flexible, and cost-effective manner. The proposed integrated communication system will overcome a set of problems, such as, coverage area, handovers, mobility, bandwidths, and health concern.

Stephan Brumme [12] evaluated the opportunities and streaming multimedia contents on mobile phones from the technical, commercial and social point of view, a summary of the used MPEG technology and its modifications for mobile phone were explained, such are real-time requirements, power consumption and enhanced error correction are presented, also he examines the major problems of the user interface on small devices and how to overcome them.

Giovanni Gualdi, Andrea, and others [13] presents MoSES (Mobile Streaming for video Surveillance), an effective system for mobile video surveillance for both PC and PDA clients; it relies over H.264/AVC video coding and GPRS/EDGE-GPRS network. a new general-purpose methodology for streaming performance evaluation is also proposed and used to compare mobile streaming for video surveillance with existing solutions in terms of different parameters (latency, image quality, video quality, and frame losses), as well as in terms of performance in people segmentation and tracking.

Kosuke Numa et al [14] introduced a design and support system for collecting and connecting people's expressions using mobile phones. In their proposed workshop, facilitators shoot and collected videos from participants. Applying a talking format, each video has connections on a large screen. Participants can see a whole network of stories and add their own ones to it.

Kartin Verclas et al [15] discussed different activities, which includes the analysis of available and soon-to-be-available software applications that can be used to produce or consume media, including software and mobile-based web application facilitating journalistic activity, communication between mobile phones, secure communications, and audio/video platforms. Ideas and recommendations for strategies supporting the use of mobile telephony to promote freedom of information and engage in citizen media activities.

Edmund [16] presents in his thesis a new, low-complexity, wavelet video compression technique scheme. Noting from a complexity study that the generation of temporally decorrelated, residual frames represents a significant computational burden, in which the scheme uses the simplest technique, of difference frames of spatial clustering of significant coefficients by splitting the frame into tiles. The scheme is found to be suitable for implementation in mobile and embedded devices due to its moderate memory and computational requirements.

C. S. pattichis et al [17] provided a snapshot of the applications of wireless telemedicine systems followed by successful case studies in electronic patient record, emergency telemedicine, teleradiology, and home monitoring.

Mohammed et al [18] experimented with the choice of reference images the process of video compression by using the intra and predicted images extracted from sequences. They applied the Faber-schauder Multi-scale Transform (FMT) and compared each image with the other images by subtracting corresponding images transformed by FMT. The choice of the best reference image is based on the result of subtraction. The obtained results of this approach on video sequences revealed an improvement in data flow and average PSNR as compared to the original encoding and choosing reference images based on the mean square error.

S.Annoaduria et al [19] proposed an innovative technique for compressing color still images using wavelet compression scheme. The proposed scheme uses wavelet transformation, tree structured vector quantization and binary vector morphological prediction for compressing color image. The use of tree structured vector quantization reduced the search time for quantization and coding. This greatly enhanced the proposed algorithm in terms of compression time. The experiment results revealed that the proposed algorithm produced a high compression ratio with minimum loss.

Tmt.Nishat et al [20] discussed the important features of wavelet in compression of still images, including the extend to which the image quality is degraded by compression and decompression process.

The optimum method of wavelet transformation is explored. Performance measure of different wavelets is compared with and without shuffling scheme. And by using wavelets and compression they achieve an optimum balance between the performance metrics like PSNR and compression ratio and also reduce the Mean Square Error. The obtained results provide a good reference for application developers to choose a good wavelet compression system for their application.

Yogendra et al [21] examined and compared various wavelet families such as Haar, Daubechies, Symlets, Coiflts, Biorthognal and Reverse Biorthognal using variety of test images. The results are measured in terms of PSNR considering as subjective quality measures. Also analyzed the effects of wavelets belonging to each of these wavelet families on image quality at a compression ratio of 100:1 and 200: at decomposition level 5 on the test images. Conclusively they conclude that the "best wavelet" choice of wavelet in the image compression of images dependent on to the image contents and desired image quality.

G.K.Kharate et al [22] compared compression performance of Daubechies, Biorthognal, Coiflets and other wavelets along the results for different frequency images. Based on the result, it was proposed that proper selection of mother wavelet on the basis of nature of images, improve the quality as well as compression ratio remarkably. The prime objective is to select the proper mother wavelet during the transform phase to compress the color image. Also this paper includes the discussion on principals of image compression methodology, the basis of wavelet and orthogonal wavelet transforms. the results demonstrates that for line-based image percentage of zeros is more for db1 as compared to other wavelets and more energy is retained.

Sourav et al [23] proposed a system which is capable of compressing and streaming live videos over Bluetooth network. Three major aspects are to be taken into consideration namely video compression. Quality of services (QoS) controls and intermediate protocols. The proposed system concerns with transmitting live video streaming in between two mobile phones via built-in cell phone camera. This system can be useful for surveillance of a secured area or monitoring of any important work. This system can help to prevent tersest attack within a small range.

Eric et al [24] discussed the benefits of streaming with SP and SI frames and described SP and SI frame encoding and explained how perfect reconstruction may be achieved based on H.264. The performance analysis reveals that gains of up to 1.5 Db can be obtained for video rates between 100 kbps and 600 kbps. Experimental results obtained on a simulated bandwidth limited networks. Also the results show that streaming with SP and SI frames reduces the congestion created by the stream on the network by up to 40%.The results can help identify scenarios for which SI and SP frames provide an attractive alternative to streaming with I frames.

Rakesh et al [25] proposed a novel method for color video compression using key-frame based color transfer. In this scheme, compression is achieved by discarding the color information of all but few selected frames. These selected frames are either the key frames (frame selected by a key frame selection algorithm) or the Intra coded (I) frames. The partially colored video is compressed using a standard encoder thereby achieving higher compression. In the proposed decoder, a standard decoder first generates the partially colored video sequence from the compressed input. A color transfer algorithm is then used for generating the fully colored video sequence. The complexity of the proposed decoder is close to a standard decoder, allowing its use in wide variety of applications like video broadcasting, video streaming, hand-held devices etc.

A thesis written by Gyanendra [26] used discrete wavelet transform (DWT) in order to extract the vocal characteristics of the speakers in speech signal whereas Nearest Neighbor (KNN) algorithm is used for feature matching, which shows high improvement in the identification rate. The feature extraction is done by six levels wavelet decomposition and these features are extracted from wavelet coefficients by mean, standard deviation and ratios between them. The result shows improvement in identification rate as high as 100% using 10 speakers.

The above pieces of research and authors discussed different issues related to image and video compression conducted in different applications. This work attempts to find the most useful and efficient wavelet function among the existing members of the wavelet families for image and video compression. Important features such as PSNR, MSE and objectives measurement test at different decomposition levels (1- 4) were computed using the wavelet functions Daubechies, Symmlets, Coiflets, Biorthognal, and Reverse Biorthognal. Different types of image contents were implemented and tested, such as News, Cartoon, Series, Panorama, Sport and Music.

For video compression approach, Different Video files contents were selected and then extracted according to selected frames number. Based on frames difference approaches (difference between frames), a number of frames were removed. The remaining frames were compressed and re-extracted according to standard mobile screen sizes (128*128). Finally the compressed images and videos were tested to insure the efficiency of the proposed technique.

1.5 Statement of the Problem

Current transmission techniques used to transmit video signals via mobile devices lack the ability to deal with the required amount of data that makes the video quality acceptable and within real time.

Delivering real-time video over mobile phone is an important component of many mobile phone multimedia applications. With respect to real-time nature of video streaming, instable bandwidth, latency and packet loss, are all problems that effect video transmission over mobile phone, these problems challenge many researchers to works in this area.

In particular, this work is concerned with the adaptation of video over mobile devices. In addition, it provides an efficient compression technique that can be suitable for mobile devices. The quality of the received video sequence will be maintained to provide a required level of user satisfaction. Therefore, it is important to choose the compression parameters to maximize the end user quality.

This work proposes a method that adopts video over mobile device based on an efficient wavelet transform techniques of compression and frames difference approaches between the successive frames of the extracted video.

1.6 Objectives

Two major issues concerning the multimedia content currently available worldwide especially video over mobile are compression and quality in which this work attempts to address both these issues in details.

The display of video on mobile phone paves the road to several new applications. These services will be used to support applications such as video content including daily news, financial stories, sports highlights, short entertainments, clips, traffic report and weather reports, domestic uses include games, day care of security. In particular, our mainly objective concerns to achieve fast streaming video over mobile phone by using some compression techniques and at the same time to achieve acceptable video quality over the mobile phone.

An important application of this work can be implemented in health care medical video and images, since it is an effective way to share information that would make video production possible from anywhere, where phones in areas with good mobile bandwidth can even broadcast live video feeds to the internet and on other mobile phones. Videos on mobile phone can share information via Bluetooth and from mobile phone to mobile phone in which many countries practice these services easily and privately. In this work, different wavelet families with different levels of decomposition were used and tested for different types of images and videos to obtain a suitable image and video compression results.

1.7 Scope of the Thesis

This dissertation contains six chapters, including introduction and literature review (chapter-1). General information about image compression techniques concerning fundamentals of wavelet transforms and its implementation is given in (chapter 2). Chapter 3 provides general information about video and video compression techniques with some implementation. The proposed system design is explained in Chapter 4. In Chapter 5 results and analysis are presented. Finally Chapter 6 provides conclusions and future works.

Chapter Two
Image Compression

2.1 Introduction

Data compression can be defined as the process of encoding or representing information using fewer bits than a more conventional representation would use, through use of specific compression schemes and algorithms. Compression is possible because most real-world data are very statistically redundant. Compression is important because it facilitates optimum consumption of expensive resources, such as disk space or connection bandwidth. Some compression schemes are reversible so that the original data can be reconstructed (lossless data compression) while others accept some loss of data in order to achieve higher compression (lossy data compression). Due to applications, Compression mainly refers to reducing the quantity of data used to represent a file, image or video content without excessively reducing the quality of the original data [27].

The performance compression scheme and effect of compression quality of the decompressed data can be studied using certain measures of performance like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR).

Speed of compression is measured in terms of Cycles per Byte (CPB), the average number of machine cycles is taken to compress one byte. Also MSE and PSNR are used mainly to measure the distortion caused by lossy compression of images and videos.

2.2 Image Compression

Image compression is the application of data compression on digital images. The objective of image compression is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form [28]. Images contain large amounts of information that requires much storage space, large transmission bandwidths and long transmission times. Therefore it is advantageous to compress the image by storing only the essential information needed to reconstruct the image. An image can be thought of as a matrix of pixel (or intensity) values [29]. In order to compress the image, redundancies must be exploited, for example, areas where there is little or no change between pixel values. Therefore images having large areas of uniform color will have large redundancies, and conversely images that have frequent and large changes in color will be less redundant and harder to compress.

Wavelet analysis can be used to divide the information of an image into approximation and detail sub signals. The approximation sub signal shows the general trend of pixel values, and three detail sub signals show the vertical, horizontal and diagonal details or changes in the image. If these details are very small then they can be set to zero without significantly changing the image.

The value below whose details are considered small enough to be set to zero is known as the threshold. The greater the number of zeros the greater the compression that can be achieved [30]. The amount of information retained by an image after compression and decompression is known as the energy retained, and this is proportional to the sum of the squares of the pixel values. If the energy retained is 100% then the compression is known as **lossless**, in which the image can be reconstructed exactly. This occurs when the threshold value is set to zero, meaning that the detail has not been changed. If any values are changed then energy will be lost and this is known as **lossy** compression. Ideally, during compression the number of zeros and the energy retention will be as high as possible. However, as more zeros are obtained more energy is lost, so a balance between the two needs to be found. Ideally Figure 2.1 shows the Compression framework of image compression [31].

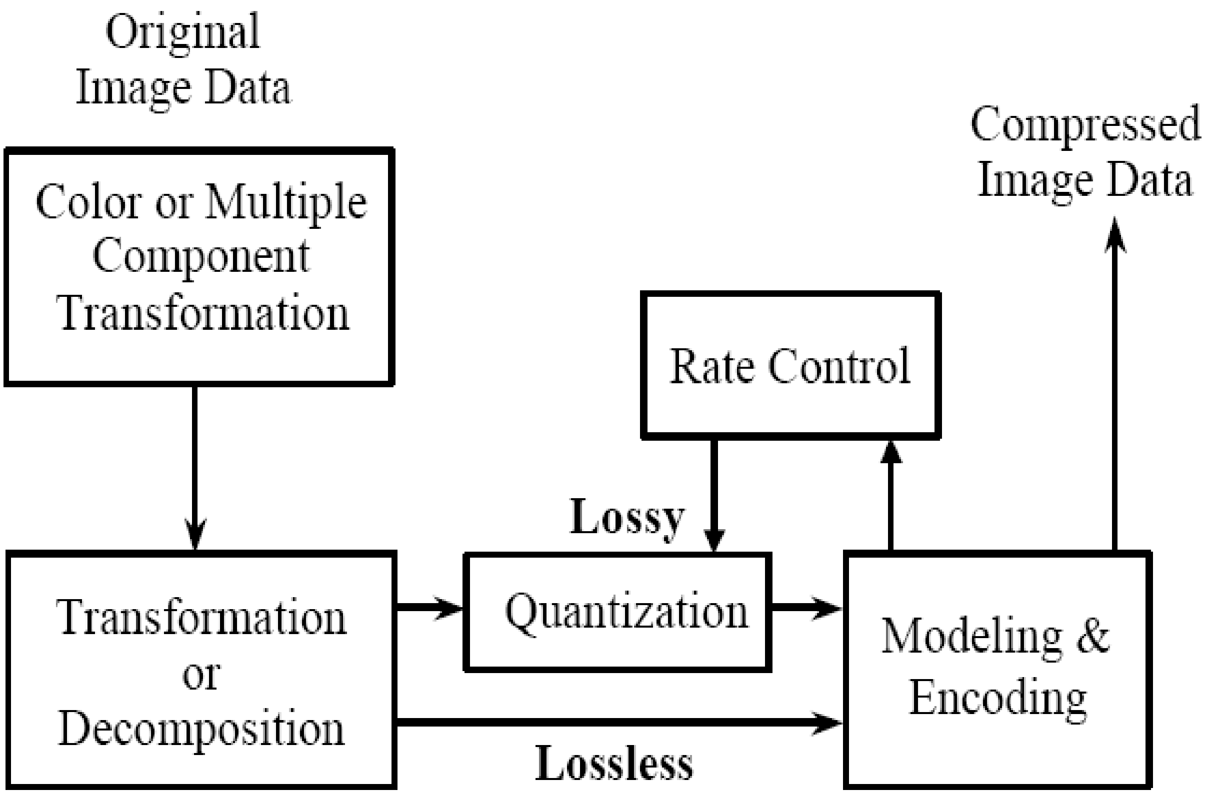


Figure (2.1): Compression Framework [31].

Different factors affecting achievable compression such as [31]:

- Sample parameters (spatial resolution, bit depth).
- Sensor characteristics (noise, spectral response).
- Scene content, including noise.
- Image size and viewing distance.
- Display characteristics (noise, light level, non-linearity's)
- Post Processing (Sharpening, Dynamic Range)

-Adjustment (DRA), Tone Transfer Curve (TTC))

- Pre-Processing (image formation, registration)
- Observer (IA, machine)
- Required task

2.3 Compression Approaches

Each compression approach is based on two main concepts: redundancy and irrelevancy. Redundancy is a reversible compression, and therefore lossless. It deals with information that is known from spatial or temporal similarities or statistical dependencies. With this approach compression about factor 2 is realizable. Irrelevancy concentrates on information that is not visible or not important for the receiver. This is an irreversible approach (lossy) but higher compression ratios about factor 40 are possible. So any compression scheme is based on removing redundancies and extracting relevant information.

Fundamentally there are 3 principles for removing redundancy [32].

- Exploiting temporal correlation
- Exploiting spatial correlation
- Exploiting statistical dependency of coded values

Reducing irrelevancy includes a data size reduction due to removing information which is unimportant to the receiver. The main fact for this approach is to exploit limitations of the human visual system. Therefore we use methods like thresholding, temporal sub sampling, spatial sub sampling, quantization of coefficients or vector quantization [32],

Entropy coding is the most important principle of redundancy reduction. It is totally lossless, so we cannot gain very high compression ratios (about factor 2). It is basic function is to eliminate data redundancy existing in source symbols and therefore supply an efficient and equivalent expression for the source, where "equivalent" means the expression is reversible and uniquely lossless decodable [33]. This approach exploits the properties of the signal source, especially the statistics of the signal levels described in an adequate source model. Huffman coding, Arithmetic coding, Lempel-Ziv Coding, Run Length Coding are kinds of Entropy coding approaches [29].

2.3.1 Run Length Coding

Probably the simplest coding scheme that takes advantage of the context is run length coding [34]. The simplest Run Length Coding saves bits by representing sequences of identical values by the value and the number of repetitions. For example, the message sequence acccbbaaabb could be transformed to (a,1), (c,3), (b,2), (a,3), (b,2) [34]. So, naturally Run Length Coding is most efficient with highly correlated signals, and is very sensitive to transmission errors. One advantage of this approach is, that we can gain possible improvement by entropy coding of the zero run length [32].

2.3.2. Huffman Coding

Huffman Code is a variable length code (similar to Morse code). The basic idea in Huffman coding is to assign shorter code words to those symbols occurring with higher probabilities and longer code words to those occurring with lower probabilities. A Huffman code is designed by merging together the two least probable characters, and repeating this process until there is only one character remaining.

A code tree is thus generated and the Huffman code is obtained from the labeling of the code tree. The branches are split as 0 or 1 and successively labeled backwards from 1 [35].

2.3.3 Arithmetic Coding

Arithmetic Coding (AC) was developed in the 1970's and 1980's, AC is a very complete concept that takes into account a Model, a statistics update strategy and the coder [36]. The main concept of Arithmetic coding is to take a stream of input symbols and replace them with a single floating point output number. The more complex the message, the more bits are needed in the output number. The output from an arithmetic coding process is a single number less than 1 and greater than or equal 0. This single number can be uniquely decoded to create the exact stream of input symbols.

Each symbol to encode is assigned a fixed probability value and the corresponding range along a "probability line", which is nominally 0 to 1. It doesn't matter which characters are assigned to which segment of the range, as long as it is done in the same manner by both the encoder and the decoder [32].In practice, it is usually necessary to represent the fractional numbers produced by an arithmetic encoder using fixed-point values within a limited dynamic range. Some implementation issues for the context-based arithmetic coder adopted for H.264 main profile [29].

2.3.4. Lempel/Ziv Coding

Lempel/Ziv Coding is a well known approach for entropy coding. It is an algorithm for lossless data compression. It is not a single algorithm, but a whole family of algorithms, stemming from the two algorithms proposed by Jacob Ziv and Abraham Lempel in their landmark papers in 1977 and 1978. Lempel Ziv algorithms are widely used in compression utilities such as gzip, GIF image compression and the V.42 modem standard [37]. Nowadays, there is a wide range of so called modified Lempel/Ziv coding. These algorithms all have a common way of working. It is based upon the source building up a dictionary of previously-seen strings, and transmitting only the “innovations" while creating new strings.

The LZ-Coding works as follows: The source stream is parsed until the shortest string is encountered that has not been encountered before. Since this is the shortest such string, all of its prefixes must have been sent before. The string can be coded by sending the index from the dictionary of the prefix string and the new bit. This string is then added to the dictionary. In terms of a programming language it is just a short loop [32]:

2.3.5 Area Coding

Area coding is an enhanced form of run length coding, reflecting the two dimensional character of images. This is a significant advance over the other lossless methods. For coding an image it does not make too much sense to interpret it as a sequential stream, as it is in fact an array of sequences, building up a two dimensional object. Therefore, as the two dimensions are independent and of the same importance, it is obvious that a coding scheme aware of this has some advantages [3].

The algorithms for area coding find rectangular regions with the same characteristics. These regions are coded in a descriptive form as an element with two points and a certain structure. The whole input image has to be described in this form to allow lossless decoding afterwards.

The possible performance of this coding method is limited mostly by the very high complexity of the task of finding largest areas with the same characteristics. Practical implementations use recursive algorithms for reducing the whole area to equal sized sub rectangles until a rectangle does fulfill the criteria defined as having the same characteristic for every pixel.

This type of coding can be highly effective but it bears the problem of a nonlinear method, which cannot be implemented in hardware. Therefore, the performance in terms of compression time is not competitive, although the compression ratio is [32].

2.4 Wavelet Image Compression

Wavelets are functions which allow data analysis of signals or images, according to scales or resolutions. The processing of signals by wavelet algorithms in fact works much the same way the human eye does; or the way a digital camera processes visual scales of resolutions, and intermediate details. But the same principle also captures cell phone signals, and even digitized color images used in medicine [38]

Wavelets are of real use in these areas, for example in approximating data with sharp discontinuities such as choppy signals, or pictures with lots of edges. Thus, applications of the wavelet idea include big parts of signal and image processing, data compression, fingerprint encoding, and many other fields of science and engineering.

Wavelet coding is based on the idea that the coefficient of a transform that de correlates the pixel of an image can be coded more efficiently than the original pixels themselves. If the transforms basis functions-in this case wavelets-back most of the important visual information into a small number of coefficients, the remaining coefficient can be quantized coarsely or truncated to zero with little image distortion. Figure (2.2) shows a typical wavelet coding system.

The implementation of wavelet compression method is very similar to that of sub band coding scheme: the signal is decomposed using filter banks. The output of the filter banks is down-sampled, quantized, and encoded. The decoder decodes the coded representation, up-samples and recomposes the signal [39].

The quantization can be adapted to exploit any positional correlation across the P decomposition levels. One or more of the lossless coding methods including run-length, Huffman, arithmetic, and bit-plane coding, can be incorporated into the final symbol coding metric, and bit-plane coding, can be incorporated into the final symbol coding step. Decoding is accomplished by inverting the encoding operations-with the exception of quantization, which cannot be reversed exactly [27].

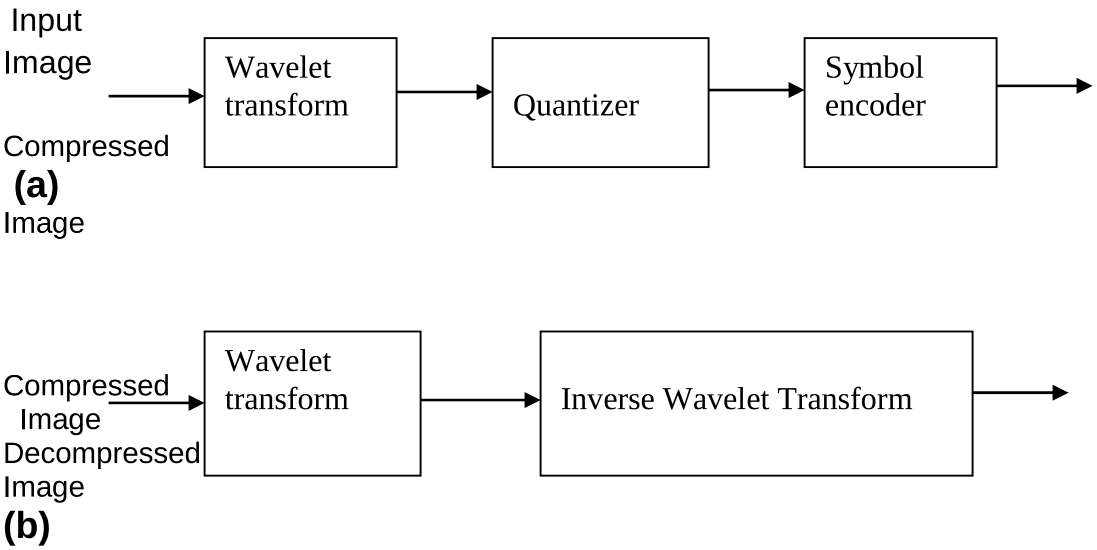


Figure (2.2): Wavelet coding system: a) encoder b) decoder

However wavelets transforms is applied to entire images, rather than sub images, so it produces no blocking artifacts. This is a major advantage of wavelet compression over other transform compression methods.

For some signal, many of the wavelet coefficients are close to or equal to zero. Thresholding can modify the coefficients to produce more zeros. In hard thresholding any coefficient below threshold λ is set to zero [35]. This should then produce many consecutive zero's which can be stored in much less space, and transmitted more quickly by using entropy coding compression.

The use of wavelets and thresholding serves to process the original signal, but, to this point, no actual compression of data has occurred, this explains that the wavelet analysis does not actually compress a signal, it is simply provides information about the signal which allows the data to be compressed by standard entropy coding techniques, such as Huffman coding. Huffman coding is good to use with a signal processed by wavelet analysis, because it relies on the fact that the data values are small and in particular zero, to compress data. It works by giving large numbers more bits and small numbers fewer bits. Long strings of zeros can be encoded very efficiently using this scheme. Therefore an actual percentage compression value can be stated in conjunction with an entropy coding technique. To compare different wavelets, the number of zeros is used. More zeros will allow a higher compression rate, if there are many consecutive zeros, this will give an excellent compression rate. Two rules are generally used for thresholding the wavelet coefficients (soft/hard thresholding). **Hard thresholding** sets zeros for all wavelet coefficients whose absolute value is less than the specified threshold limit. While a **soft threshold** is a preprocessing tool that reduces the back ground_in an image, in which intensity values below the threshold value are reduced (set to lower values or even zero) [40].

2.4.1.Wavelet Families

Several families of wavelets that have proven to be especially useful are included, Haar, Daubechies, Biothogonal, Coiflets, Symlets, Morlet, Mexican Hat, Meyer, Other Real Wavelets and Complex Wavelets [41].

The differences between different mother wavelet functions consist in how these scaling signals and the wavelets are defined. The choice of wavelet determines the final waveform shape; likewise, for Fourier transform, the decomposed waveforms are always sinusoid. To have a unique reconstructed signal from wavelet transform. We need to select the orthogonal wavelets to perform the transforms [40].

Haar

Any discussion of wavelets begins with Haar wavelet, the first and simplest. Haar wavelet is discontinuous, and resembles a step function. It represents the same wavelet as Daubechies db1.

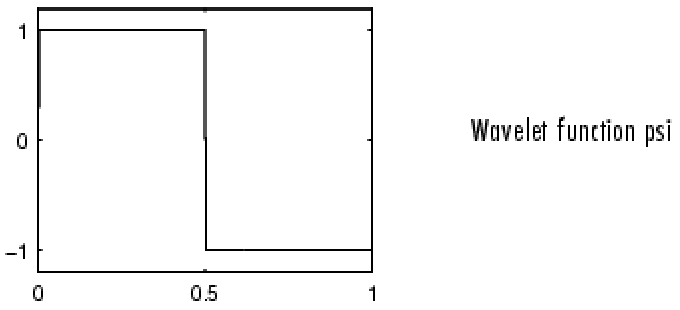


Figure 2.3 Haar wavelet

The Haar wavelet's mother wavelet function ψ(*t*) can be described as

$$\psi(t) = \begin{cases} 1 & 0 \le t < 1/2, \\ -1 & 1/2 \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{…… (2.1)}$$

And its scaling function φ(*t*) can be described as

$$\phi(t) = \begin{cases} 1 & 0 \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{…… (2.2)}$$

Daubechies

Ingrid Daubechies, one of the brightest stars in the world of wavelet research, invented what are called compactly supported orthonormal wavelets -- thus making discrete wavelet analysis practicable.

The names of the Daubechies family wavelets are written dbN, where N is the order, and db the "surname" of the wavelet. The db1 wavelet, as mentioned above, is the same as Haar wavelet. Here are the wavelet functions psis of the next nine members of the family:

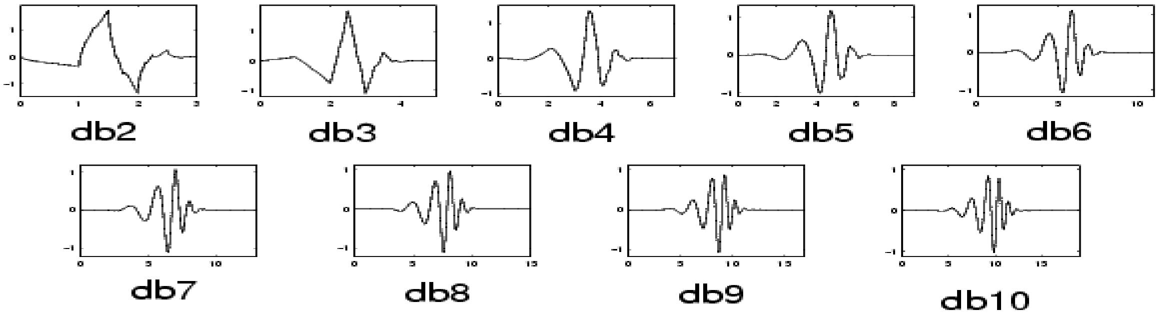


Figure (2.4): Daubechies wavelets

Biorthogonal

This family of wavelets exhibits the property of linear phase, which is needed for signal and image reconstruction. By using two wavelets, one for decomposition (on the left side) and the other for reconstruction (on the right side) instead of the same single one.

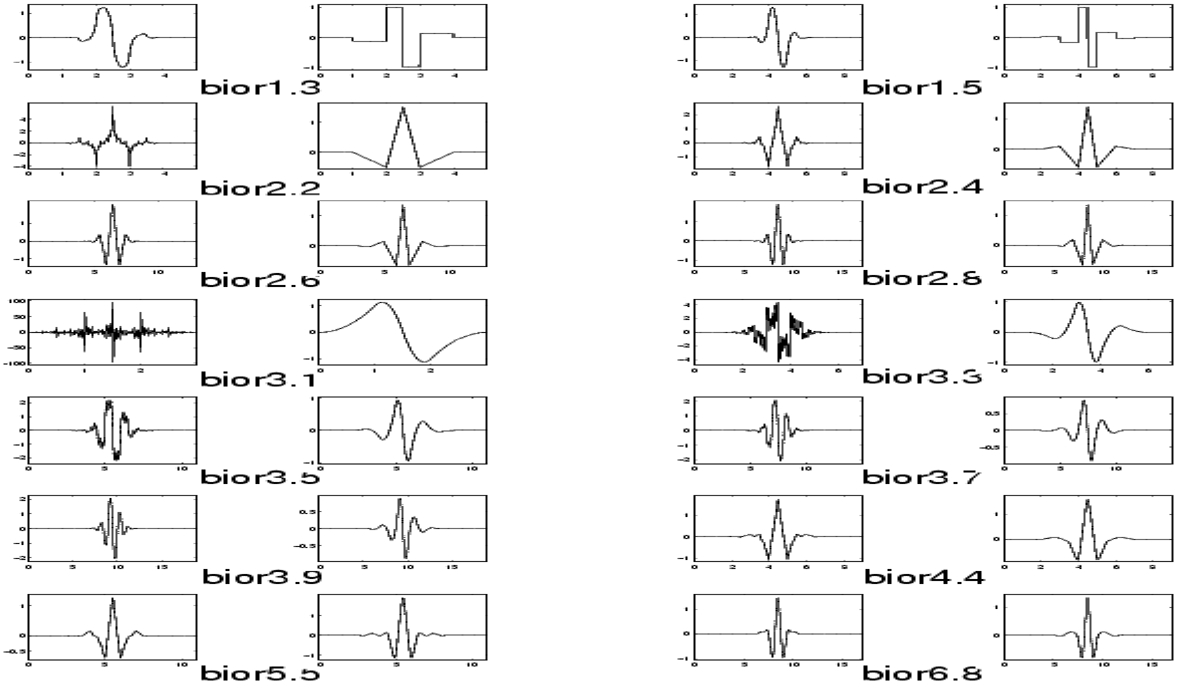
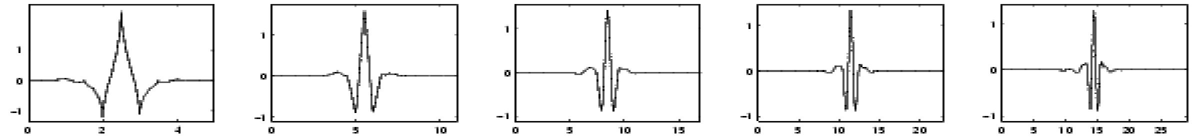


Figure (2.5): Biorthogonal wavelets

Coiflets

Built by I. Daubechies at the request of R. Coifman. The wavelet function has 2*N* moments equal to 0 and the scaling function has 2*N*-1 moments equal to 0. The two functions have a support of length 6*N*-1.

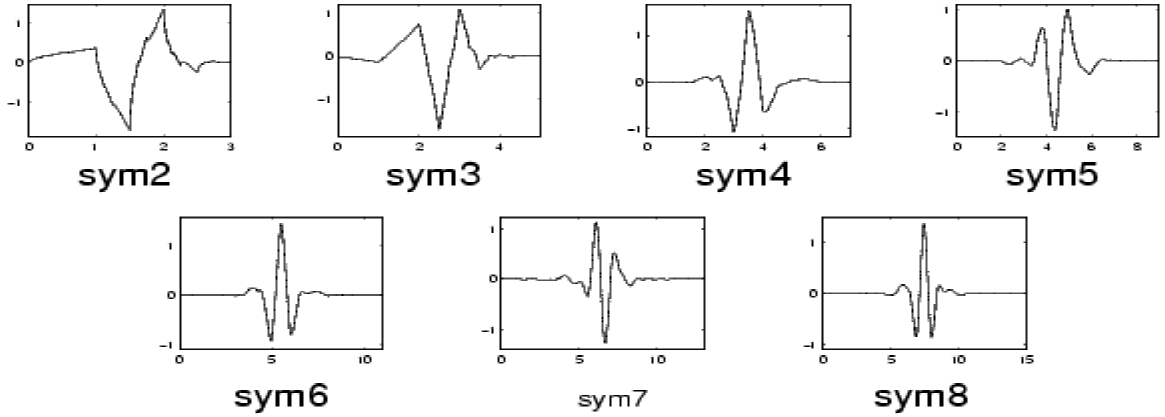


coif1 coif2 coif3 coif4 coif5

Figure (2.6): Coiflets Wavelets

Symlets

The symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the db family. The properties of the two wavelet families are similar. Here are the wavelet functions.



sym2 sym3 sym4 sym5

sym6 sym7 sym8

Figure (2.7): Symlets Wavelets

Morlet

This wavelet has no scaling function, but is explicit.

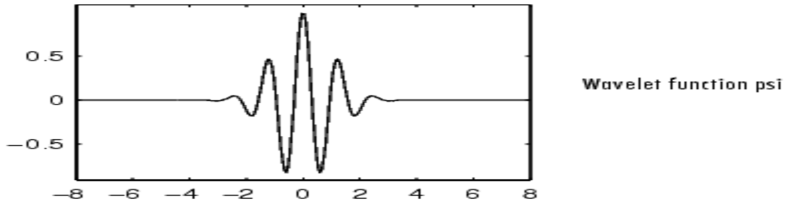


Figure (2.8): Morlet Wavelet

Mexican Hat

This wavelet has no scaling function and is derived from a function that is proportional to the second derivative function of the Gaussian probability density function.

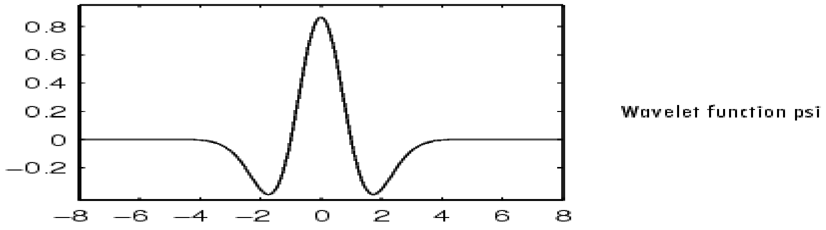


Figure (2.9): Mexican Hat Wavelet

Meyer

The Meyer wavelet and scaling function are defined in the frequency domain.

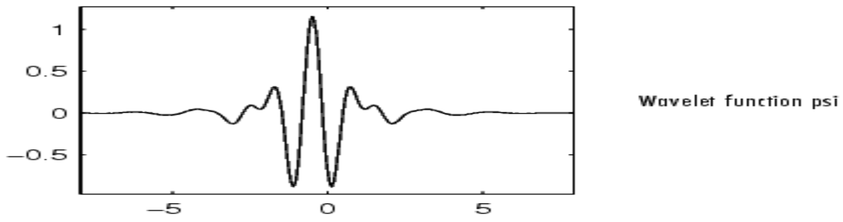


Figure (2.10): Meyer Wavelet

2.4.2 Choice of Wavelet Families

Important properties of wavelet functions in image compression applications are compact support (lead to efficient implementation), symmetry (useful in avoiding dephasing in image processing), orthogonally (allow fast algorithm), regularity, and degree of smoothness (related to filter order or filter length). In this work, six types of wavelet families are examined: Haar Wavelet (HW), Daubechies(db), Symlets(sym), Coiflets (coif), Biorthogonal (bior), Reverse Biorthogonal (rbior). Each wavelet family can be parameterized by integer N that determines filter order. Biorthogonal wavelets can use filters with similar or dissimilar orders for decomposition (Nd) and reconstruction (Nr). In this work, different filter orders are used inside each wavelet family. Daubechies and Coiflet wavelets are families of orthogonal wavelets that are compactly supported. Compactly supported wavelets correspond to Finite-Impulse Response (FIR) filters and, thus, lead to efficient implementation. Only ideal filters with infinite duration can provide alias-free frequency split and perfect interband decorrelation of coefficients. Since time localization of the filter is very important in visual signal processing, arbitrarily long filters cannot be used. A major disadvantage of Daubechies Wavelet (DW) and Coiflets Wavelet (CW) is their asymmetry, which can cause artifacts at borders of the wavelet sub bands. DW is asymmetrical while CW is almost symmetrical. Symmetry in wavelets can be obtained only if we are willing to give up either compact support or orthogonality of wavelet (except for haar wavelet), which is orthogonal, compactly supported, and symmetric). If we want both symmetry and compact support in wavelets, we should relax the orthogonality condition and allow nonorthogonal wavelet functions. An example is the family of biorthogonal wavelets that contains compactly supported and symmetric wavelets [53].

2.4.3 Two Dimensional Wavelet Masks

There are many different types of wavelet masks this work will be concentrated on the masks that are shown in table (2.1).

Table (2.1) some of the wavelet transforms masks [42]

	Decomposition low-pass filter	**Decomposition high-pass filter**	**Reconstruction low-pass filter**	**Reconstruction high-pass filter**
Haar				
h0	0.7071067812	-0.7071067812	0.7071067812	0.7071067812
h1	0.7071067812	0.7071067812	0.7071067812	-0.7071067812
Daubechies 2				

h0	-0.1294095226	-0.4829629131	0.4829629131	-0.1294095226
h1	0.2241438680	0.8365163037	0.8365163037	-0.2241438680
h2	0.8365163037	-0.2241438680	0.2241438680	0.8365163037
h3	0.4829629131	-0.1294095226	-0.1294095226	-0.4829629131
Symlets2				
h0	-0.1294095226	-0.4829629131	0.4829629131	-0.1294095226
h1	0.2241438680	0.8365163037	0.8365163037	-0.2241438680
h2	0.8365163037	-0.2241438680	0.2241438680	0.8365163037
h3	0.4829629131	-0.1294095226	-0.1294095226	-0.4829629131
Coiflets2				
h0	-0.0007205494	-0.0163873365	0.0163873365	-0.0007205494
h1	-0.0018232089	-0.0414649368	-0.0414649368	0.0018232089
h2	0.0056114348	0.0673725547	-0.0673725547	0.0056114348
h3	0.0236801719	0.3861100668	0.3861100668	-0.0236801719
h4	-0.0594344186	0.8127236354	0.8127236354	-0.0594344186
h5	-0.0764885991	0.4170051844	0.4170051844	0.0764885991
h6	0.4170051844	0.0764885991	-0.0764885991	0.4170051844
h7	0.8127236354	-0.0594344186	-0.0594344186	-0.8127236354
h8	0.3861100668	-0.0236801719	0.0236801719	0.3861100668
h9	-0.0673725547	0.0056114348	0.0056114348	0.0673725547
h10	-0.0414649368	0.0018232089	-0.0018232089	-0.0414649368
h11	0.0163873365	-0.0007205494	-0.0007205494	-0.0163873365
Biorthogonal2.2				
h0	0	0	0	0
h1	-0.1767766953	0.3535533906	0.3535533906	0.1767766953
h2	0.3535533906	-0.7071067812	0.7071067812	0.3535533906
h3	1.0606601718	0.3535533906	0.3535533906	-1.0606601718
h4	0.3535533906	0	0	0.3535533906
h5	-0.1767766953	0	0	0.1767766953
Reverse Biorthogonal2.2				
h0	0	0.1767766953	-0.1767766953	0
h1	0	0.3535533906	0.3535533906	0
h2	0.3535533906	-1.0606601718	1.0606601718	0.3535533906
h3	0.7071067812	0.3535533906	0.3535533906	-0.7071067812
h4	0.3535533906	0.1767766953	-0.1767766953	0.3535533906
h5	0	0	0	0

2.5 Fourier Transform

The Fourier transform is one of the most useful tools in image processing. It provides a realization of an image that is a composition of sinusoidal functions over an infinite band of frequencies.

This realization facilities many image processing techniques [43]. Filtering, enhancement, encoding, restoration, texture analysis, feature classification, and pattern recognition are but a few of the many areas of image processing that utilize the Fourier transform. There are many applications of Fourier transform image processing and filters, data processing and analysis, solving linear partial differential equations (PDE's), designing and using antennas, and pattern recognition. One of the most Fourier transforms types for the image compression is the two-dimensional Discrete Fourier Transform (2D DFT).The 2D DFT of an image array is a transform of pixels (sampled picture points) with a two-dimensional spatial location indexed by co-ordinate x and y. this implies that there are two dimensions of frequency, u and v, which are the horizontal and vertical spatial frequencies, respectively. The two-dimensional Fourier transform of an image function *F(x, y)* of size *M*N* is given by the equation (2.1) [44].

$$F(u,v) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\exp\left[-2\pi i(\frac{mu}{M}+\frac{nv}{N)}\right] \text{..........} 2.1$$

$$u = 0,1\text{..........}M-1$$

$$v = 0,1\text{..........}N-1$$

Where j=√-1.The inverse two Dimensional Discrete Fourier Transform (2D-IDFT) is given by:

$$f(m,n)=\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}F(u,v)\exp\left[2\pi i(\frac{mu}{M}+\frac{nv}{N)}\right]..........2.2$$

$$m=0,1..........M-1\text{, }n=0,1..........N-1$$

The variables u and v used in equation 2.1 are the frequency variables , m and n are the spatial or image variables. However equation (2.1) and equation (2.2) are very slow for large image sizes. They are usually implemented by using Fast Fourier Transfer (FFT).The (FFT) is a (DFT) developed by Tukey and Cooley in 1965, which is a wonderful development of the Fourier's computation that improves speed dramatically. The (FFT) can only be applied to square images whose size is an integer power of 2 (without special effort) by exploiting the separability property of the Fourier transform. Separability means that the Fourier transform is calculated in two stages: the rows are first transformed using 1D-FFT, and then this data is transformed in columns, again using a 1-D (FFT). This process can be achieved since the sinusoidal basis functions are orthogonal.

The full effect of the Fourier transform is shown by application to an image of much higher resolution.

2.6.Discrete Wavelet Transform

A fundamental shift in the compression approach came after the Discrete Wavelet Transform (DWT) became popular [28].

DWT is a transformation technique used to represent an image in a new time and frequency scale by decomposing the input image into low frequency, middle and high frequency bands. The value of low frequency band is averaging value of the filter whereas the high frequency coefficients are wavelet coefficients or detail values [45]. The advantage of DWT over existing transforms, such as Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT), is that the DWT performs a multi resolution analysis of a signal with localization in both time and frequency domain.

The discrete wavelet transform uses filter banks for the construction of the multi resolution time-frequency plane. The Discrete Wavelet Transform analyzes the signal at different frequency bands with different resolutions by decomposing the signal into an approximation and detail information. The decomposition of the signal into different frequency bands obtained by successive high pass g[n] and low pass h[n] filtering of the time domain signal. The combination of high pass g[n] and low pass filter h[n] comprise a pair of analyzing filters. The output of each filter contains half the frequency content, but an equal amount of samples as the input signal. The two outputs together contain the same frequency content as the input signal; however the amount of data is doubled. Therefore down sampling by a factor two, denoted by ↓2, is applied to the outputs of the filters in the analysis bank [46].

The advantages of wavelet transform, is that it divides the information of an image into decomposing images to approximate sub signals (LL) and detail sub signals(LH, HL, HH) parts as shown in figure (2.11). This enables to isolate and manipulate the data with specific properties. With this, it is possible to determine whether to preserve more specific details. For instance, keeping more vertical detail instead of keeping all the horizontal (LH), vertical details (HL), and diagonal (HH) of an image that has more vertical aspects. This would allow the Image to lose a certain amount of horizontal and diagonal details, but would not affect the image in human perceptions [6].

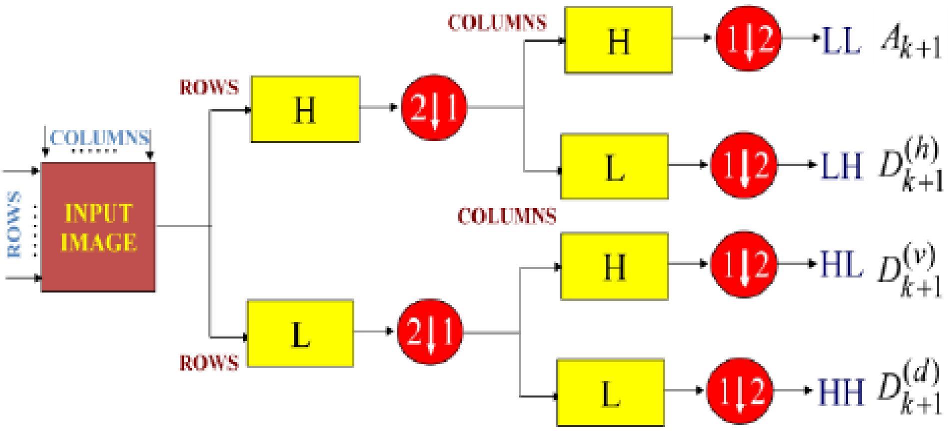


Figure (2.11): The two –level sub band decomposition used in the DWT

Once arrive at discrete wavelet coefficients, one need a way to reconstruct them back into the original signal (or a modified original signal if one played around with the coefficients). In order to do this, utilize the process that is known as the inverse discrete wavelet transform.

Much like the DWT can be explained by using filter bank theory, so can the reconstruction of the IDWT. The process is simply reversed. The DWT coefficients are first up sampled (the approximation and the detail coefficients are handled separately) by placing zeros in between every coefficient, effectively doubling the lengths of each. These are then convolved with the reconstruction scaling filter for approximation coefficients (the reconstruction scaling filter is simply the original scaling filter that has been flipped left to right) and the reconstruction wavelet filter for the detail coefficients. These results are then added together to arrive at the original signal [47].

Similar to how one made the signal periodic before the DWT calculations were carried out, one must make DWT coefficients periodic before convolving to obtain the original signal. This is done by simply taking the first N/2-1 coefficients from the DWT coefficients, and appending them to the end. Remembering that N is the length of the scaling filter [48].

2.7.Two Dimensional DWT

The Two-Dimensional DWT (2D-DWT) is a multi-level decomposition technique. It converts images from spatial domain to frequency domain. One-level of wavelet decomposition produces four filtered and sub-sampled images, referred to as sub bands. The sub band image decomposition using wavelet transform has a lot of advantages. Generally, it profits analysis for non-stationary image signal and has high compression rate. And its transform field is represented multi resolution like human's visual system so that can progressively transmit data in low transmission rate line. DWT processes data on a variable time-frequency plane that matches progressively the lower frequency components to coarser time resolutions and the high-frequency components to finer time resolutions, thus achieving a multi resolution analysis. The Discrete Wavelet Transform has become powerful tool in a wide range of applications including image/video processing, numerical analysis and telecommunication [46].

There are three basic architectures for the two-dimensional DWT: level by level, line-based, and block based architecture. In implementing the 2-D DWT, a recursive algorithm based on the line based is used. The image to be transformed is stored in a 2- D array. Once all the elements in a row are obtained, the convolution is performed in that particular row. The process of row-wise convolution will divide the given image into row parts with the number of rows in each part equal to half that of the image. This matrix is again subjected to a recursive line-based convolution, but this time column-wise.the result will DWT coefficients corresponding to the image, with the approximation coefficient occupying the top-left quarter of the matrix, horizontal coefficients occupying the bottom-left quarter of the matrix, vertical coefficients occupying the top-right quarter of the matrix and the diagonal coefficients occupying the bottom-right quarter of the matrix [28].

One can say 2D- DWT is obtained via the implementation of low pass and high pass filters on rows and columns of image respectively. A low pass filter and a high pass filter are chosen, such that they exactly half the frequency range between themselves.

This filter pair is called the Analysis Filter pair. First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the lpf is a half band filter, the output data contains frequencies only in the first half of the original frequency range. The high pass filter is applied for the same row of data, and similarly the high pass components are separated, and placed by the side of the low pass components. This procedure is done for all rows. As mentioned above, the LL band at the highest level can be classified as most important, and the other 'detail' bands can be classified as of lesser importance, with the degree of importance decreasing from the top of the pyramid to the bands at the bottom [35]. DWT is a multi spectral technique used for converting signal or image into four different bands such as low-low (LL), low-high (LH), high-low (HL) and high-high (HH) as demonstrated in Figure (2.12) [3].

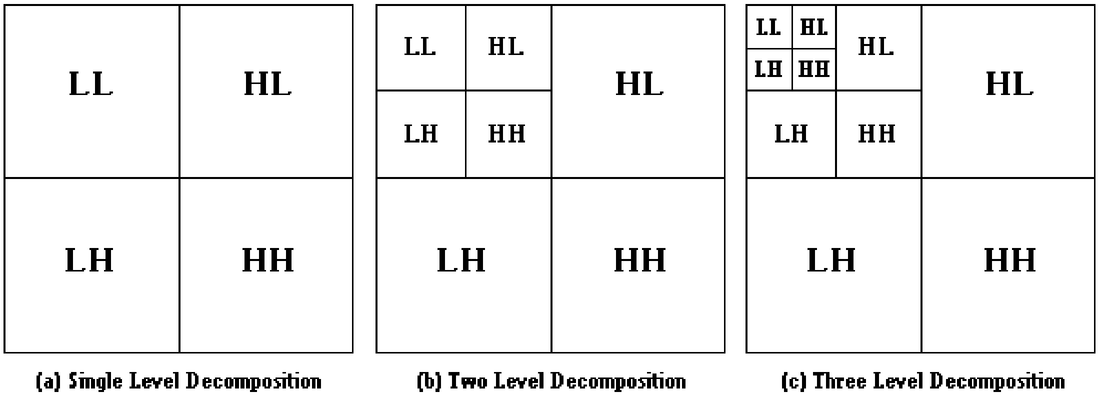


Figure (2.12) Decomposition of image applying DWT

2.8 An Overview of Image Compression Standard

Most of the standards of image compression are sectional by the International Standardization Organization (ISO) and the Consultative Committee of the International Telephone and Telegraph (CCTT). They address both binary and continuous-tone (monochrome and color) image compression, as well as both still-frame and video applications [35]. In this section an overview of the fundamental theory of two well-known image compression standards – JPEG and JPEG 2000 will be given. JPEG is an image compression standard that was developed by the "Joint Photographic Experts Group". JPEG was formally accepted as an international standard in 1992.

JPEG refers to a wide variety of possible image compression approaches that have been collected to a single standard that include:

- Lossy component
- Lossless component
- Entropy coding

The JPEG Process

JPEG is a commonly used method of compression for photographic images. The degree of compression can be adjusted, allowing a selectable tradeoff between storage size and image quality. JPEG typically achieves 10:1 compression with little perceptible loss in image quality. More comprehensive understanding of the process may be acquired as [49]:

1. The image is broken into 8x8 blocks of pixels.
2. Working from left to right, top to bottom, the DCT is applied to each block.
3. Each block is compressed through quantization.
4. The array of compressed blocks that constitute the image is stored in a drastically reduced amount of space.
5. When desired, the image is reconstructed through decompression, a process that uses the Inverse Discrete Cosine Transform (IDCT).

The JPEG 2000 Standard

JPEG 2000 is a standard for image compression produced by the ISO which defines "a set of lossless (bit-preserving) and lossy compression methods for continues-tone, bi-level, gray scale, or color digital still images. The block diagram of the JPEG2000 encoder is illustrated in Figure (2.13 (a)). The discrete transform is first applied on the source image data. The transform coefficients are then quantized and entropy coded, before forming the output code stream (bit stream). The decoder is the reverse of the encoder (Figure 2.13. b). The code stream is first entropy decoded, de quantized and inverse discrete transformed, and thus resulting in the reconstructed image data [50].

JPEG2000 makes use of the wavelet and sub-band technologies. Some of the markets targeted by the JPEG2000 standard are Internet, printing, digital photography, remote sensing, mobile, digital libraries and E-commerce [52].

JPEG 2000 provides a better rate-distortion tradeoff and improved subjective image quality, and can also provide additional functionalities lacking in the current JPEG standard.

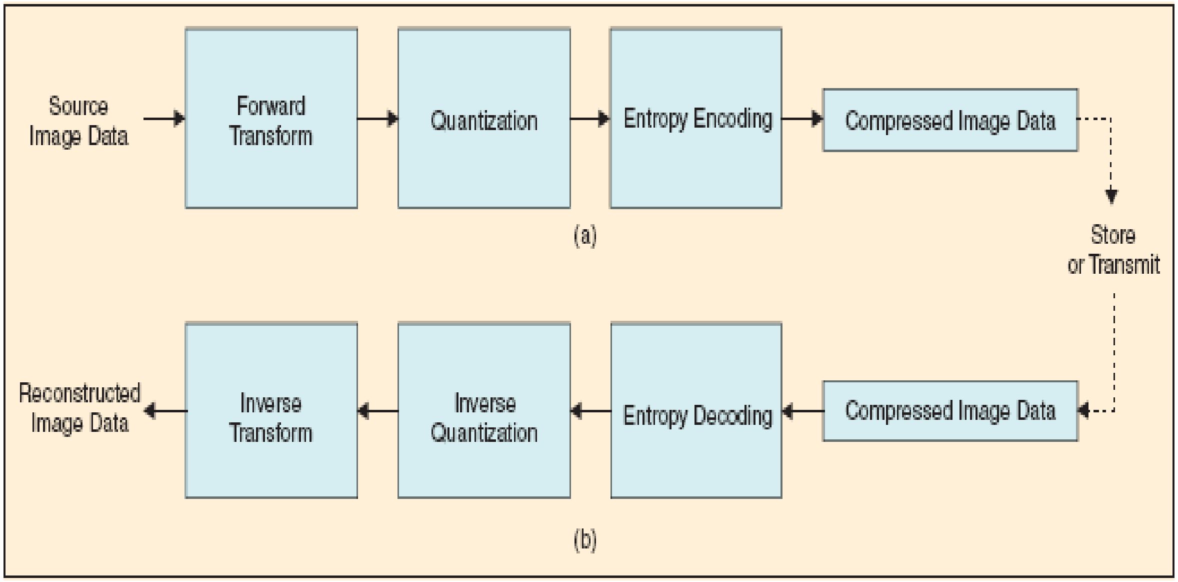


Figure (2.13): Block diagrams of the JPEG2000 (a) encoder and (b) decoder [53].

Features of JPEG2000

The JPEG2000 standard provides a set of features that are of importance to many high-end and emerging applications by taking advantage of new technologies [50]. It addresses areas where current standards fail to produce the best quality or performance and provides capabilities to markets that currently do not use compression.

The markets and applications better served by the JPEG2000 standard are Internet, color facsimile, printing, scanning (consumer and prepress), digital photography, remote sensing, mobile, medical imagery, digital libraries / archives and Ecommerce. Each application area imposes some requirements that the standard should fulfil. Some of the most important features that this standard should possess are the following [50, 51]:

Lossless and lossy compression: the standard provides lossy compression with a superior performance at low bit-rates. It also provides lossless compression with progressive decoding. Applications such as digital libraries/databases and medical imagery can benefit from this feature [50].

Protective image security: the open architecture of the JPEG2000 standard makes easy the use of protection techniques of digital images such as watermarking, labeling, stamping or encryption.

Region-of-interest coding: in this mode, regions of interest (ROI's) can be defined. These ROI's can be encoded and transmitted with better quality than the rest of the image.

Robustness to bit errors: The standards incorporate a set of error resilient tools to make the bit-stream more robust to transmission errors.

Progressive transmission by pixel accuracy and resolution: Progressive transmission that allows images to be reconstructed with increasing pixel accuracy or spatial resolution is essential for many applications.

This feature allows the reconstruction of images with different resolutions and pixel accuracy, as needed or desired, for different target devices. World Wide Web, image archival and printers are some application examples.

Random code stream access and processing: This feature allows user defined ROI's in the image to be randomly accessed and/or decompressed with less distortion than the rest of the image. Also, random code stream processing could allow operations such as rotation, translation, filtering, feature extraction and scaling.

Open architecture: It is desirable to allow open architecture to optimize the system for different image types and applications. With this feature, a decoder is only required to implement the core tool set and a parser that understands the code stream. If necessary, unknown tools could be requested by the decoder and sent from the source [50].

Content-based description: Image archival, indexing and searching is an important area in image processing. Standards like MPEG-7 ("Multimedia Content Description Interface") are addressing this problem currently [6]. Content based description of images might be available as part of the compression system (for example as metadata information).

Side channel spatial information (transparency): Side channel spatial information, such as alpha planes and transparency planes are useful for transmitting information for processing the image for display, printing or editing. An example of this is the transparency plane used in World Wide Web applications.

Continuous-tone and bi-level compression: It is desired to have a coding standard that is capable of compressing both continuous-tone and bi-level images. If feasible, this standard should strive to achieve this with similar system resources. The system should compress and decompress images with various dynamic ranges (e.g. 1 bit to 16 bit) for each color component. Examples of applications that can use this feature include compound documents with images and text, medical images with annotation overlays, and graphic and computer generated images with binary and near to binary regions, alpha and transparency planes, and facsimile [50].

While **GPEG2000** standard addresses the following problems **[**54**]:**

- Lossless and lossy compression: there is currently no standard that can provide superior lossless compression and lossy compression in a single bit streams.
- Low bit-rate compression
- Large images
- Single decomposition architecture
- Transmission in noisy environments
- Progressive transmission
- Region of interest coding
- Computer generated imagery
- Compound documents

In addition, JPEG2000 is able to handle up to 256 channels of information whereas the current JPEG standard is only able to handle three color channels [54].

2.9 Performance Measurement

In general, measurement of image quality usually can be classified into two categories, which are subjective and objective quality measurements. Subjective quality measurement, Mean Opinion Score (MOS), is truly definitive but too inconvenient, the most time taken and expensive. Therefore, objective measurements are developed such as MSE (Mean Square Error), MAE (Mean Average Error), PSNR (Peak Signal to Noise Ratio), Structural Content (SC), Maximum Difference (MD), Laplacian Mean Square Error(LMSE), and Normalized Absolute Error (NAE) [55]. In fact, MSE and PSNR are the most common measures of image quality in image compression systems, despite the fact that they are not adequate as perceptually meaningful measures, especially MSE variants do not correlation well with Subjective quality measure.

2.9.1. Subjective Quality Measurement

In image compression system, the truly definitive measure of image quality is perceptual quality [55]. The compressed image quality is specified by MOS, which is result of perception based on subjective evaluation. The meaning of the 5-level grading scales of MOS is 5-pleasant or excellent quality, 4-good, 3-acceptable, 2-poor quality and 1-unacceptable. MOS is defined as follow:

$$MOS = 1/S\sum_{i=1}^{S} ip(i).........2.3$$

Where ***i*** is image score ***p*(*i*)** is image score probability and S is number of observer.

2.9.2. Objective Quality Measurement

Two of the error metrics used to compare the various image compression techniques are the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error. The mathematical formulae for the two are:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[I(i,j) - K(i,j) \right]^2 \text{..........2.4}$$

The PSNR is defined as:

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20.\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \text{..........2.5}$$

Where I (i, j) is the original image, K (i, j) is the approximated version (which is actually the decompressed image) and m, n are number of pixels in row and column directions of image, respectively.

A lower value for MSE means less error, and as seen from the inverse relation between the MSE and PSNR, this translates to a high value of PSNR. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher. Here, the 'signal' is the original image, and the 'noise' is the error in reconstruction. So, if you find a compression scheme having a lower MSE (and a high PSNR), you can recognize that it is a better one [56].

Chapter three
video compression

3.1 Introduction

Video generally consists of a scene of physical objects whether the objects are real objects captured by a video camera or artificial objects drawn by an artist such as those in cartoons and 3D animations. The primary difference between using images and videos in representing a scene of physical objects is that video can portray the movement of objects [57].

The temporal structure of a video is an important aspect for video retrieval as it provides a logical hierarchy that allows the user to drill down to find the target object. Figure (3.1) classifies the hierarchical levels as either being syntactic or semantic. Syntactic levels should be extractable automatically with little domain knowledge. Semantic levels on the other hand require a knowledge base or annotation of the content for the levels to be appropriately constructed. An act or episode for example has little to do with the physical characteristics of the video and requires a semantic knowledge of film scripts. Syntactic levels may also require some domain knowledge but it is generally small. All videos consist of frames and are designed to be played back at a preset number of frames per second (fps). Extracting individual frames is not a challenge for video retrieval research as video decoders are designed to be able to present each frame from the video sequence. Broadcast quality videos generally have quite high frame rates ranging from 24-60 fps resulting in a large number of frames. For example a 30 minute at 30 fps will contain 54,000 frames.

Since the time difference between frames is small. Often the content between two consecutive frames is very similar. The small time period is used simply to provide the smoothest effect of motions. A larger time period could be used, such as .5s, but the smoothness of the motion would be lost [57]

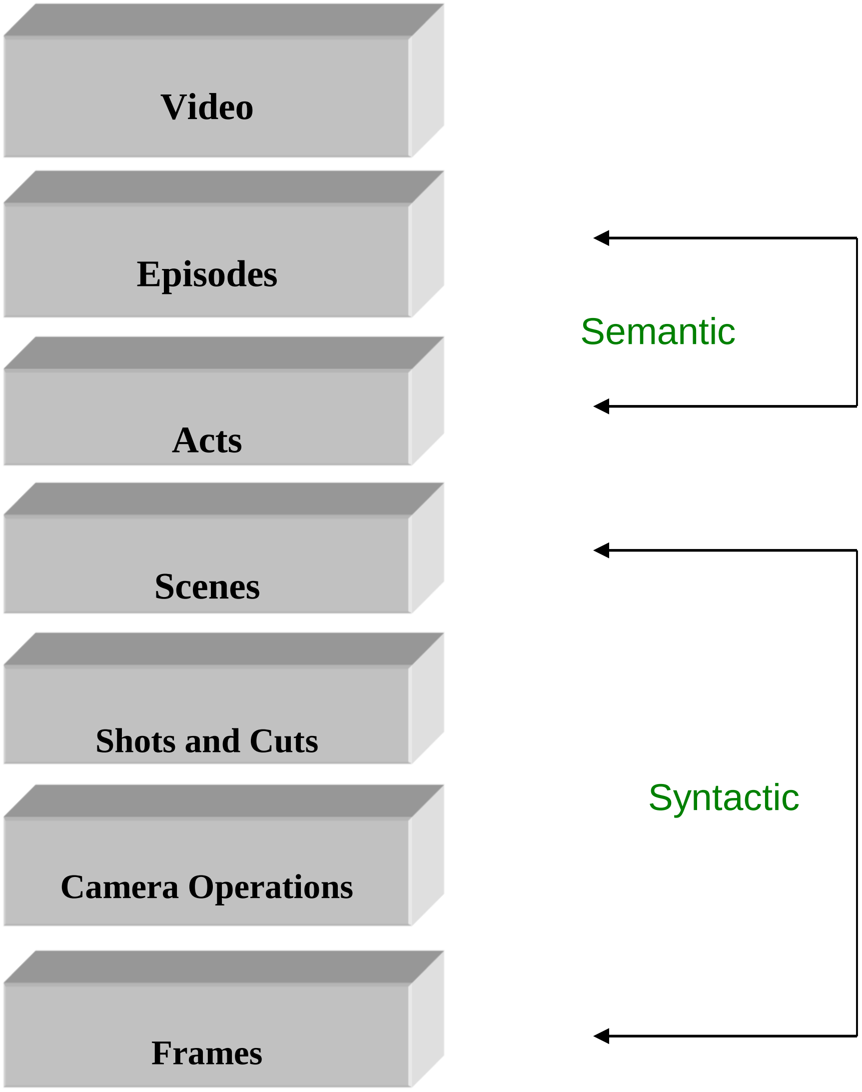


Figure (3.1): Temporal structure of a video sequence

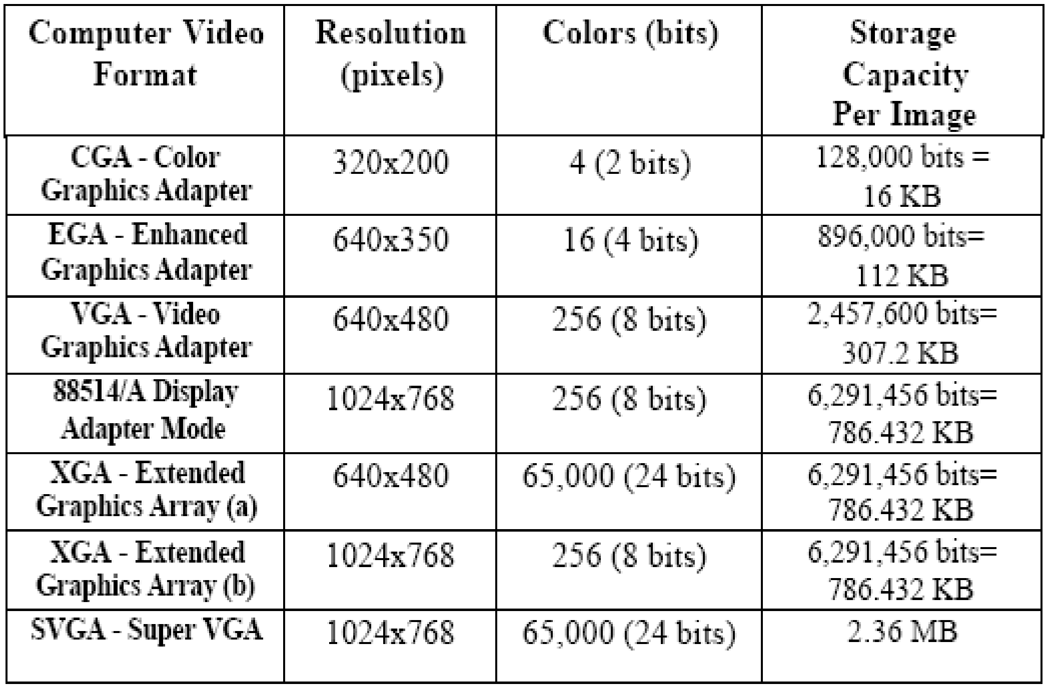
Digital video communication can be found today in many application scenarios, such as:

- Broadcast, subscription, and pay-per-view services over satellite, cable, and terrestrial transmission channels (e.g., using H.222.0 MPEG-2 systems);
- Wire-line and wireless real-time conversational services (e.g., using H.32x or Session Initiation Protocol (SIP);
- Internet or local area network (LAN) video streaming (using Real-Time Protocol/Internet Protocol (RTP/IP)
- Storage formats (e.g., digital versatile disk (DVD), digital camcorders, and personal video recorders) [58].

Video represents a collection of images or frames, where each image (frame) can be represented in one of the formats. These formats include RGB, YUV, and YCrCb. Continuous motion is produced at a frame rate of 15 frames per second (fps) or higher. Full motion video is typically referred as one at 30fps. While the traditional movies run at 24 fps. The NTSC television standard in U.S.A uses 29.97 Hz frequency, which is approximately 30fps, while PAL and SECAM television standards use 25 fps [59].

Resolution of an image or video system refers to its capability to reproduce fine detail. Higher resolution requires more complex imaging and video systems for representing these images (video frames) in real time. In computer systems, resolution is characterized by the number of pixels. Table (3.1) Summarizes popular computer video formats and related storage requirements.

Table (3.1): Characteristics of a Variety of Computer Video Formats [59]



Computer Video Format	Resolution (pixels)	Colors (bits)	Storage Capacity Per Image
CGA - Color Graphics Adapter	320x200	4 (2 bits)	128,000 bits = 16 KB
EGA - Enhanced Graphics Adapter	640x350	16 (4 bits)	896,000 bits= 112 KB
VGA - Video Graphics Adapter	640x480	256 (8 bits)	2,457,600 bits= 307.2 KB
88514/A Display Adapter Mode	1024x768	256 (8 bits)	6,291,456 bits= 786.432 KB
XGA - Extended Graphics Array (a)	640x480	65,000 (24 bits)	6,291,456 bits= 786.432 KB
XGA - Extended Graphics Array (b)	1024x768	256 (8 bits)	6,291,456 bits= 786.432 KB
SVGA - Super VGA	1024x768	65,000 (24 bits)	2.36 MB

From the view of compression side and **w**hen the motion video is represented in digital form, it can be decomposed into a time dependent sequence of individual information units. For example, a motion video sequence can be divided into film, clips, frames, and pixels, as shown in Figure (3.2). A full motion video, or film, consists of a number of clips, which are characterized with a common thread. Each clip consists of a number of frames. Each frame can be divided into blocks. Typical sizes of the blocks, which are used in video processing systems (such are compression, retrieval and indexing, motion estimation, etc.) are 8 X 8 and 16 X 16 pixels. Pixels are the smallest pieces of information, which consists of 8, 16, or 24 bits [59].

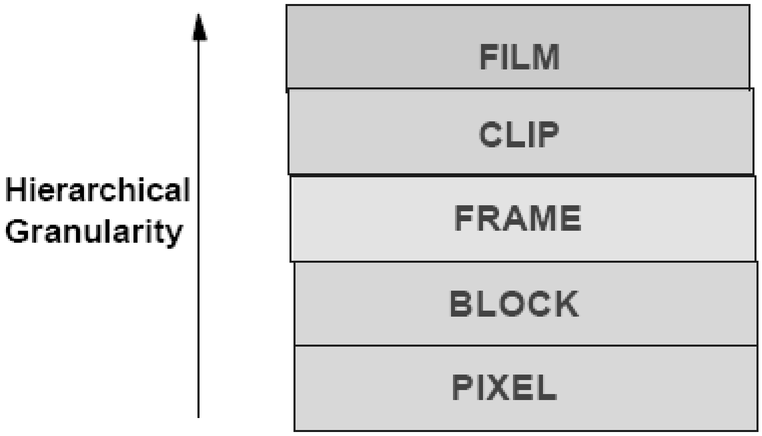


Figure (3.2): Motion video sequence divided into information units.

Video compression is the technique that reduces the size of data to convert it in digital video images. It is a collection of temporal motion compensation and image compression.

Digital video compression is implemented because it makes the process of transmission fast and high quality. Compression video is easy to store and download at low bandwidth rate. Digital compression techniques can compress video without affecting the quality of video display. It converts it into parts that humans are unable to detect. For example, there are numberless colors but human can perceive 1024 shades. Video compression is the process of reducing data for such things that we cannot notice. Standard video compression cameras compress video data at a ratio of 5 to 1. There are formats that help you to compress video as much as you can. But too much compression is also not recommended.

When you compress data too much, it throws away more data that can bring noticeable change in the video. Excessive compression can make your video unrecognizable.

At the time of compression video, you need to apply compression setting so that it alerts you about the recommended rate of video compression. In this way, you can maintain a right balance between the quality and size of the video.

Video as a sequence of video frames, however, involves a huge amount of data. Let us take a look at an illustrative example. Assume the Present Switch Telephone Network (PSTN) modem can operate at a maximum bit rate of 56,600 bits/s. Assume each video frame has a resolution of 288 X 352 (288 lines and 352 pixel/line), which is comparable with that of a normal TV picture and is referred to as Common Intermediate Format (CIF). Each of the three primary colors RGB (Red, Green, Blue) is represented for one pixel with 8 bits, as usual, and the frame rate in transmission is 30 frames /second to provide continuous motion video. The required bit rate, then, is 288 x 352 x 8 x 3 x 30=72,999,720 bits/second motion. Therefore, the ratio between the required bit rate and the largest possible bit rate is about 1289. This implies that we have to compress the video data by at least 1289 times in order to accomplish the transmission described in this example. Note that an audio signal has not been accounted for yet in this illustration [3].

3.2 How Video Compression Works

Video compression is about reducing and removing redundant video data so that a digital video file can be effectively sent and stored. The process involves applying an algorithm to the source video to create a compressed file that is ready for transmission or storage. To play the compressed file, an inverse algorithm is applied to produce a video that shows virtually the same content as the original source video. The time it takes to compress, send, decompress and display a file is called latency. The more advanced the compression algorithm, the higher the latency, given the same processing power. A pair of algorithms that works together is called a video codec (encoder/decoder). Video codec's that implement different standards are normally not compatible with each other; that is, video content that is compressed using one standard cannot be decompressed with a different standard. For instance, an MPEG-4 Part 2 decoder will not work with an H.264 encoder. This is simply because one algorithm cannot correctly decode the output from another algorithm but it is possible to implement many different algorithms in the same software or hardware, which would then enable multiple formats to be compressed [60].

Different video compression standards utilize different methods of reducing data, and hence, results differ in bit rate, quality and latency. Results from encoders that use the same compression standard may also vary because the designer of an encoder can choose to implement different sets of tools defined by a standard. As long as the output of an encoder conforms to a standard's format and decoder, it is possible to make different implementations. This is advantageous because different implementations have different goals and budget. Professional non-real-time software encoders for mastering optical media should have the option of being able to deliver better encoded video than a real-time hardware encoder for video conferencing that is integrated in a hand-held device. A given standard, therefore, cannot guarantee a given bit rate or quality. Furthermore, the performance of a standard cannot be properly compared with other standards, or even other implementations of the same standard, without first defining how it is implemented [3].

A decoder, unlike an encoder, must implement all the required parts of a standard in order to decode a compliant bit stream. This is because a standard specifies exactly how a decompression algorithm should restore every bit of a compressed video [60].

3.3 Basic Methods of Reducing Data

A variety of methods can be used to reduce video data, both within an image frame and between a series of frames. Within an image frame, data can be reduced simply by removing unnecessary information, which will have an impact on the image resolution. In a series of frames, video data can be reduced by such methods as difference coding, which is used by most video compression standards including H.264 [60]. In difference coding, a frame is compared with a reference frame (i.e. earlier I- or P-frame) and only pixels that have changed with respect to the reference frame are coded. In this way, the number of pixel values that are coded and sent is reduced. With Motion JPEG format, images are coded and sent as separate unique images (I-frames) with no dependencies on each other. With difference coding (used in most video compression standards including H.264), only the first image (I-frame) is coded in its entirety. In the images (P-frames), references are made to the first picture for the static elements, i.e. the house, and only the moving parts, i.e. the running man, is coded using motion vectors, thus reducing the amount of information that is sent and stored.

The amount of encoding can be further reduced if detection and encoding of differences is based on blocks of pixels (macro blocks) rather than individual pixels; therefore, bigger areas are compared and only blocks that are significantly different are coded. The overhead associated with indicating the location of areas to be changed is also reduced [60].

3.4 Understanding Frames

Depending on the H.264 profile, different types of frames such as I-frames, P-frames and B-frames, may be used by an encoder. These types of frames can be shown in figure (3.3). An I-frame, or intra frame, is a self-contained frame that can be independently decoded without any reference to other images. The first image in a video sequence is always an I-frame. I-frames are needed as starting points for new viewers or resynchronization points if the transmitted bit stream is damaged. I-frames can be used to implement fast-forward, rewind and other random access functions.

An encoder will automatically insert I-frames at regular intervals or on demand if new clients are expected to join in viewing a stream. The drawback of I-frames is that they consume much more bits, but on the other hand, they do not generate many artifacts [60].

A P-frame, which stands for predictive inter frame, makes references to parts of earlier I and/or P frame to code the frame. P-frames usually require fewer bits than I-frames, but a drawback is that they are very sensitive to transmission errors because of the complex dependency on earlier P and I reference frames.

A B-frame, or bi predictive inter frame, is a frame that makes references to both an earlier reference Frame and a future frame [60].

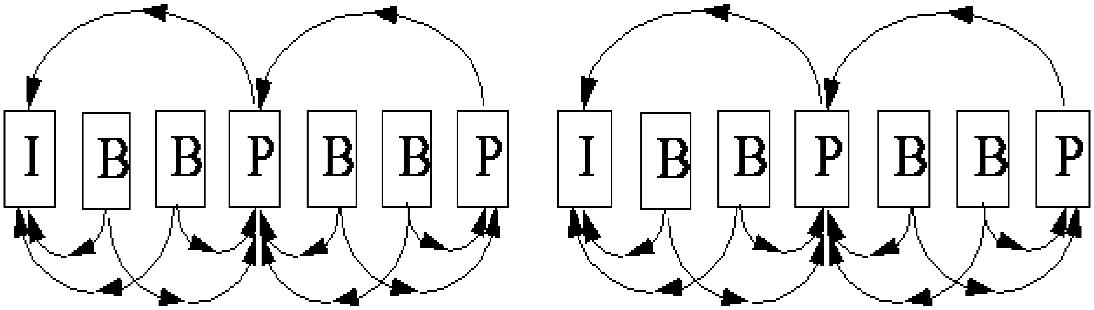


Figure (3.3): Group of pictures "inter frame dependencies in a stream"

The goal of the temporal model is to reduce redundancy between transmitted frames by forming a predicted frame and subtracting this from the current frame. The output of this process is a residual (difference) frame and the more accurate the prediction process, the less energy is contained in the residual frame. The residual frame is encoded and sent to the decoder which re-creates the predicted frame, adds the decoded residual and reconstructs the current frame. The predicted frame is created from one or more past or future frames (‘reference frames’).

The accuracy of the prediction can usually be improved by compensating for motion between the reference frame(s) and the current frame.

The simplest method of temporal prediction is to use the previous frame as the predictor for the current frame [61].

Changes between video frames may be caused by object motion (rigid object motion, for example a moving car, and deformable object motion, for example a moving arm), camera motion (panning, tilt, zoom, rotation), uncovered regions (for example, a portion of the scene background uncovered by a moving object) and lighting changes. With the exception of uncovered regions and lighting changes, these differences correspond to pixel movements between frames. It is possible to estimate the trajectory of each pixel between successive video frames, producing a field of pixel trajectories known as the optical flow (optic flow) [61].

3.5 Fundamentals of Video Compression

While investigating the properties of video signals, one should consider the properties of the human visual system. The aim of a video compression is to decrease the information found in the video signal by removing the information which cannot be perceived by visual system and eliminating the redundant information that is repeating itself. There are several solutions to reduce those redundancies and irrelevant information.

3.5.1 Removal of Spectral Redundancy

In general, the input for video compression schemes, i.e. the output of a professional video camera, is a full color video signal in RGB (Red – Green – Blue) color space. RGB space is spectrally redundant. Since human eye is more sensitive to luminance than to chrominance, color space is transformed into the YUV space as the first step of compression, in order to eliminate spectral redundancies. In the original digitized video signal, there are red (R), green (G) and blue (B) components for each pixel. Each of these components is usually 9 represented with 8 bits. In YUV color space, there are one luminance (Y) and two chrominance (U and V) components for each pixel. Since U and V spaces are redundant, these signals can be down sampled. In 4:2:2 YUV system, down sampling is applied only in horizontal direction, giving a 3:2 compression ratio. Whereas in 4:1:1 YUV conversion, U and V signals are down sampled in both horizontal and vertical directions and a 2:1 compression ratio is achieved [62].

3.5.2 Removal of Spatial Redundancy

Besides spectral redundancy, video signals have also spatial information that cannot be perceived by humans. Human visual system is more sensitive to lower frequencies in the spatial domain. Hence, one can compress higher frequencies in the signal better, without distorting the visual quality. This can be achieved by first transforming the signal into another domain. By this transformation, a better energy compaction is possible. After the transformation the coefficients in the transformed domain can be more efficiently coded using different quantization. Two main techniques are used to remove spatial redundancy:

DCT and DWT. Discrete Cosine Transform (DCT) is the most widely used transform in image and video coding algorithms. Nowadays, Digital Wavelet Transform (DWT) is becoming popular as it outperforms DCT. In either transform the aim is to localize signal energy at different frequencies. In DWT localization in both spectral and frequency domain is possible [62].

Discrete Cosine Transform (DCT):

The DCT is closely related to Discrete Fourier Transform (DFT) and it is of some importance to realize that the DCT coefficients can be given a frequency interpretation close to the DFT. Thus low DCT coefficients relate to low spatial frequencies within image blocks and high DCT coefficients to higher frequencies [53]

For DCT-based methods, each frame is separated into 8x8 blocks and processed separately. Each block is transformed by DCT. The coefficients with higher indexes are higher frequencies in both horizontal and vertical dimensions 10 and the lowest indexed coefficient in both dimensions is the average value of the block and called as DC coefficient. The DC coefficient is the most important coefficient in the block and should be coded without any loss. The other coefficients are called AC coefficients and they are scanned in the order of their frequencies. Since Human Visual System (HVS) is insensitive to details, higher frequency coefficients can be compressed more using a higher quantization. The quantization level in DCT based codec's are adjusted based on the target bit rate. While the target bit rate gets smaller, higher frequencies are more suppressed. Since each block is coded and quantized independently, the block boundaries can be detected for low bit rates. This is one of the main disadvantages of DCT [62].

Discrete Wavelet Transform (DWT):

Wavelet compression technology is a very new technology. Wavelet transforms remained a mathematicians' delight until mid 1980s. During the 80's, 11 Ingrid Daubechies made a breakthrough in designing elegant wavelets that could be implemented in practice. During the same period, Stephane Mallat Designed and implemented wavelet transforms of real signals on Digital Signal Processors (DSP). Their work caused an upsurge in applications of wavelets in all areas of signal processing, compression and transmission. Wavelet transforms rely on short basis functions that integrate to zero during their existence. The basis function is sometimes called 'mother wavelet' [62].

To produce the wavelet transform of a signal, mother wavelet is contracted, stretched and shifted. Summation of all these signals should achieve a perfect reconstruction of the original signal. Wavelet Transform searches for similarities between the mother wavelet and the original signal. Conventional transform techniques such as Fourier and Discrete Cosine transforms search for similarities between a fixed basis function (namely sine or cosine) and the original signal. This is the fundamental difference between wavelet and conventional transform techniques. This enables us to use basis functions which are finite in time. Wavelet transforms have a huge number of basis functions that can be used in an application as long as they satisfy certain mathematical criteria. Fourier transform however, can only use sine and cosine functions. This feature presents an opportunity for signal processing engineers. Basis functions for the transform can now be selected for the application in hand. In practice, wavelet transforms are taken by filtering and down sampling the signal. Original signal reconstructed signal Shifted wavelets Stretched wavelet [62].

The group of low and high pass filters is called 'Quadrature Mirror Filters' (QMF). Down sampling the signal introduces aliasing. However specially designed QMF counteract this during reconstruction and 'repair' the signal. Therefore perfect reconstruction is always possible with wavelet transforms. An important point to note here is that the transformed signal is still in time domain.

Wavelet transforms may also be used on images. This requires taking the wavelet transform in two directions but all the basic operations remain the same. Sub sampling operates in two directions reducing the image pixel count to a quarter in the transformed image. The original image can be rebuilt exactly from the transformed image. Examination of the transformed image reveals that each transformed quarter contains different frequency information, which means that wavelet transform enables localization in both time and frequency domains [35].

Further wavelet transforms of the low frequency portion can be taken by increasing the depth (level) of the transformation. The original image can still be reconstructed exactly but the inverse transform will have to be taken 3 times [35].

Wavelet transform opens up new possibilities in image compression. One of the reasons is that wavelet transform separates image data into directional information. The human eye is not sensitive to high frequency information. Even without threshold, the high frequency quarter of the transformed image could be discarded. Images reconstructed without the high frequency information will very much look like the original due to the weakness in our vision system but can be constructed using less data. When compared to DCT based compression schemes such as JPEG, main advantages of wavelet compression is that wavelet transform coding can effectively eliminate the blocking effect, and fundamentally resolve the problem of mosquito noise. There is no blocking effect since the basis in adjacent blocks overlap one another. Moreover, mosquito noise is significantly reduced since its higher frequency bases are narrower. Another approach may be to apply DCT to the whole image in order to eliminate these properties. However applying DCT to the whole image will drastically increase the computation time when compared with DWT. In addition, with this approach higher frequency bases will increase, which will increase the mosquito noise [62].

In general the DWT consists of applying a wavelet coefficient matrix *hierarchically*, first to the full data vector of length *N*, then to the “smooth” vector of length *N*/2, then to the “smooth-smooth” vector of length *N*/4, and so on until only a trivial number of “smooth-*. . .*-smooth” components (usually 2) remain. The procedure is sometimes called a pyramidal algorithm, for obvious reasons. The output of the DWT consists of these remaining components and all the “detail” components that were accumulated along the way [50].

Another advantage of wavelet transforms over DCT is that the basis functions can be changed to suit the range of images [62].

3.5.3 Removal of Temporal Redundancy

The video signals are sampled with a rate of 25 or 30 frames/second. This makes the consecutive frames of a video sequence to be almost similar. Especially, the background information will be the same, if the camera movements and illumination changes are insignificant. This causes a temporal redundancy between frames. In order to remove this redundancy, the first approach is to subtract the previous frame from the current frame to reduce the information to be transmitted by assuming no camera motion. A better algorithm is to divide the image into blocks and try to estimate the current frame from previous frame with the motion information of each block. By sending this motion information, camera movement and the motion of objects can be taken into account while reducing the temporal redundancies. This operation is called motion compensation. After estimating the motion, current frame is generated using the motion information, and the previous frame. This method decreases the bit rate drastically, since instead of the whole image only the error between the blocks is sent to the decoder [62, 3].

There are several approaches for estimating motion in video. The most widely used system is to use fixed size rectangular blocks with a logarithmic or exhaustive search within a search window. Another approach is to use hierarchical motion estimation.

Several resolutions of the image are obtained by filtering and down sampling operations. Motion estimation is applied at the lowest resolution and refined either by fixed block size or variable block size at each upper level.

Another approach is to utilize three-dimensional transforms instead of motion estimation and 2-D DCT. Frequency domain transforms are applied in time and spatial domains. The main disadvantage of this scheme is its complexity, since 3-D transform requires lots of computation. Moreover, 17 temporal redundancies cannot be eliminated as successful as in motion estimation-based approaches [62, 35].

3.5.4 Removal of Statistical Redundancy

Although the above redundancies are exploited, there is still redundancy in the compressed signal. This is due to the statistical properties of the video signals. In order to reduce these redundancies, some statistical lossless coding schemes are used. Run-length coding, Huffman coding and Arithmetic coding are examples of lossless coding techniques used to eliminate these statistical redundancies.

Arithmetic coding is computationally more complex than the above but gives better results. It should be noted that in arithmetic coding, predefined tables, which limit compression performance, are not used, which necessary in Huffman is coding [62, 3].

3.6 Digital Video/Image Coding Standard

The fast growth of digital transmission services has generated a great deal of interest in the digital transmission of video signals. Some digitized video source signals require very high bit rates, ranging from more than 100 Mbits/s for broadcast-quality video to more than 1 Gbits/s for HDTV signals. Owing to this, video compression algorithms, which reduce the bit rates to an affordable level on practical communication channels, are required. Digital video coding techniques have been investigated over several decades. There are two factors that make video compression possible: the statistical structure of the data in the video source and the psychophysical redundancy of human vision. Video compression algorithms can remove the spatial and temporal correlation that is normally present in the video source. In addition, human observers are subject to perceptual limitations in amplitude, spatial resolution, and temporal acuity. By proper design of the coding system, it is possible to discard information without affecting perceived image quality, or at least, with only minimal degradation [3].

Video compression standards extend to transform-based, still image compression techniques to include methods for reducing temporal or frame-to-frame redundancies. Although there are a variety of video coding standards in use today most rely on similar video compression techniques. Depending on the intended application, the standards can be grouped into broad categories: (1) video teleconferencing standards and ((2) multimedia standard [35].

Several traditional techniques have been developed for image and video data compression. Recently, with advances in data compression and VLSI techniques, the data compression techniques have been extensively applied to video signal compression. Video compression techniques have been under development for over 20 years and have recently emerged as the core enabling technology for a new generation of DTV (both SDTV and HDTV) and multimedia applications. Digital video systems currently being implemented (or under active consideration) include terrestrial broadcasting of digital HDTV in the United States, satellite DBS (Direct Broadcasting System), computer multimedia, and video via packet networks. In response to the needs of these emerging markets for digital video, several national and worldwide standards activities have been started over the last few years. These organizations include ISO (International Standards Organization), ITU (International Telecommunication Union, formally known as CCITT (International Telegraph and Telephone Consultative Committee), JPEG (Joint Photographic Experts Group), and MPEG (Moving Picture Experts Group) as shown in Table (3.2). The related standards include JPEG standards, MPEG-1, 2, 4 standards, and H.261 and H.263 video teleconferencing coding standards as shown in Table (3.3). It should be noted that the JPEG standards are usually used for still image coding, but they can also be used in video coding. Although the coding efficiency would be lowered, they have shown to be useful in some applications [3].

Table (3.2): List of Some Organizations for Standardization [3]

Organization	**Full Name of Organization**
ITU	International Telecommunication Union
JPEG	Joint Photographic Experts Group
MPEG	Moving Picture Experts Group
ISO	International Standards Organization
IEC	International Electro technical Commission

Table (3.3): Video/Image Coding Standards [3]

Name	Year of Completion	Major Features
JPEG	1992	For still image coding, DCT based
JPEG2000	2000	For still image coding, DWT based
H.261	1990	For videoconferencing, 64 kbits/s–1.92 Mbits/s
MPEG-1	1991	For CD-ROM, 1.5 Mbits/s
MPEG-2	1994	For HDTV, 2-15 Mbits/s; for ATSC HDTV ,19.2 Mbits/s most extensively used
H.263	1995	For very low bit rate coding, below 64 kbits/s
MPEG-4 Part 2	1999	For multimedia, content-based coding, its simple profile and advanced simple profile is applied to mobile video and streaming
H.264/AVC (MPEG-4 Part 10)	2005	For many applications with significant improved coding performance over MPEG-2 and MPEG-4 Part 2
VC-1	2005	For many applications, coding performance close to H.264
Real Video	1997	For many applications, coding performance similar to MPEG-4 Part 2
MPEG-7	2000	Content description and indexing
MPEG-21	2002	Multimedia framework

e.g., studio editing systems. Though JPEG standards are not video coding standards, and then include them to give a full picture of all international image and video coding standards.

3.6.1 JPEG Standard

Since the mid-1980s, the ITU and ISO have been working together to develop a joint inter national standard for the compression of still images. Officially, JPEG [jpeg] is the ISO/IEC international standard, ‘‘Digital compression and coding of continuous-tone still images’’ or the ITU- T recommendation T.81. JPEG became an international standard in 1992. JPEG is a DCT- based coding algorithm. It continues to work on future enhancements, which may adopt wavelet-based algorithms [3].

3.6.2 JPEG2000

JPEG 2000 [jpeg2000] is a type of image coding system under development by JPEG for still image coding. JPEG2000 is considered using the wavelet transform as its core technique. This is because the wavelet transform can provide not only excellent coding efficiency but also wonderful spatial and quality scalable functionality. This standard is intended to meet a need for image compression with great flexibility and efficient interchange ability. It is also intended to offer unprecedented access into the image while still in compressed domain. Thus, an image can be accessed, manipulated, edited, transmitted, and stored in a compressed form [3].

3.6.3 MPEG -1

In 1988, ISO established the MPEG to develop standards for the coded representation of moving pictures and associated audio information for digital storage applications. MPEG completed the first phase of its work in 1991. It is known as MPEG-1 [mpeg1] or ISO standard 11172, ‘‘Coding of moving picture and associated audio.’’ The target application for this specification is digital storage media at bit rates up to about 1.5 Mbits/s [3]. MPEG-1 was primarily targeted for multimedia CD-ROM applications, requiring additional functionality supported by both encoder and decoder [64]

3.6.4 MPEG-2

MPEG started its second phase of work, MPEG-2, in 1990. MPEG-2 is an extension of MPEG-1 that allows for greater input-format flexibility, higher data rate for SDTV or HDTV applications, and better error resilience.

Basically MPEG-2 can be seen as a superset of the MPEG-1 coding standard and was designed to be backward compatible to MPEG-1 - every MPEG-2 compatible decoder can decode a valid MPEG-1 bit stream. Many video coding algorithms were integrated into a single syntax to meet the diverse applications requirements. New coding features were added by MPEG-2 to achieve sufficient functionality and quality, thus prediction modes were developed to support efficient coding of interlaced video, in addition to scalable video coding [63]

3.6.5 MPEG-4

MPEG-4 Part 2 Visual standard has been approved in 1999. The MPEG-4 Part 2 Visual supports object-based coding technology and it aims to provide enabling technology for a variety of functionalities and multimedia applications:

Universal accessibility and robustness in error-prone environments

High interactive functionality

Coding of natural and synthetic data or both

Compression efficiency [64].

3.6.6 H.261

This model was adopted in 1990 and the final revision was approved in 1993 by the ITU- T. It is designed for video teleconferencing and utilizes a DCT-based motion compensation scheme. The target bit rates are from 64 to 1920 kbits /s.

3.6.7 H.263, H.263 Version 2 (H.263+), H.263 ++, and H.26L

The H.263 [h263] vide o coding standard is specifically designed for very low bit rate applications such as video conferencing. Its technical content was completed in late 1995 and the standard was approved in early 1996. It is based on the H.261 standard with several added features: unrestricted motion vectors , syntax- based arithmetic coding, advanced prediction, and PB-frame s. The H.263 version 2 video coding standard, also known as H.263 , was approved in January 1998 by the ITU- T. H.263 includes a number of new optional features based on the H.263. These new optional features are added to provide improved coding efficiency, a flexible video format, scalability, and backward compatible supplemental enhancement information. H.263

is the extension of H.263 and was completed in the year 2000. H.26L is a long-term project, which is looking for more efficient video coding algorithms. Finally, the activity of H.26L ended because the joint video team (JVT) of MPEG and ITU-T VC EG developed a new video coding standard H.264, which has greatly improved the coding efficiency over MPEG-2 and H.263 [3].

3.6.8 MPEG -4 Part 10 Advanced Video Coding or H.264 /AVC

Recently, the JVT of MPEG and ITU- T VC EG has developed new video coding standard H264. Because many new tools have been used, H.264 /AVC has achieved higher coding efficiency, which is almost twice better than MPEG-2 [3].

3.6.9 VC-1

VC-1 is a video codec developed by Microsoft and late has been standardized by SMPTE (Society of Motion Picture and Television Engineers). It is implemented by Microsoft as Windows Media Video (WMV) 9. Its coding performance is close to the H.264/AVC.

3.6.10 Real Video

Real Video is a video codec developed by Real Net Works. It was first released in 1997 and its version 10 is released in 2006. Real Video is supported on many platforms, including Window s, Mac, Linux, Solaris, and several mobile phones. Its coding performance is close to MPEG -4 Part 2.

It should be noted that MPEG-7 and MPEG-21 are not a coding standard; MPEG-7 is a multimedia content description standard, which can be used to fast indexing and searching for multimedia content; and the MPEG-21 is a multimedia framework, which aims at defining an open framework for multimedia applications.

The VC-1 is a SMPTE standard and Real Video is not an international standard, but it is extensively supported by many platforms.

It is also interesting to note that in terms of video compression methods, there is a growing convergence towards motion compensated (MC), inter-frame DCT algorithms represented by the video coding standards. However, wavelet-based coding techniques have found recent success in the compression of still image coding in both the JPEG2000 and MPEG-4 standards. This is because it posses unique features in terms of high coding efficiency and excellent spatial and quality scalability [3, 35]. The wavelet transform has not successfully been applied to video coding due to some several difficulties. First, it is not clear how the temporal redundancy can be removed in this domain. Motion compensation is an effective technique for DCT-based video coding scheme; however, it is not so effective for wavelet-based video coding. This is because the wavelet transform uses large block size or full frame, but motion compensation is usually performed on a limited block size. This mismatch would reduce the inter-frame coding efficiency. Many engineers and researchers are working on these problems. Among these standards, MPEG-2 has had a great impact on the consumer electronics industry because the DVD (digital versatile disk) and DTV have adopted it as core technology. But recently developed new coding standard is attracted many applications, including HD-DVD, mobile TV, and others [3].

3.7 Video Quality Analysis

As video becomes a fundamental part of advanced networking applications, being able to measure its quality has significant importance at all stages: starting from the development of new video codec's and ending to the monitoring the quality of the transmission system. Indeed, in order to appreciate and compare the performance of all involved components and devices, methods to assess and quantify the quality are proposed.

In measuring video quality, it is useful to address separately the varies component of video quality such as Picture quality, Audio Quality, Audiovisual (Multimedia) Quality, Transmission Quality[65].

In general, two classes of methods are available to measure video quality: **subjective tests**, where human subjects are asked to assess or rank the viewed material, and **objective models**, which are computational models that measure the quality by comparing the original and distorted video sequences. Subjective tests may produce the most accurate ratings, but, they require costly and complex setup and viewing conditions and thus, they are inflexible to use. Objective quality metrics, on the other hand, are based purely on mathematical methods, from quite simplistic yet inaccurate models, like PSNR, to sophisticated ones that exploit models of human visual perception and produce far more reliable results. While objective models of quality appear to be very promising, both methods are considered useful in the process of measuring the quality of video applications (and, in general, multimedia applications). The next sections provide a review of work on subjective and objective Video Quality Assessment, presenting their respective advantages and weak points [66].

3.7.1 Subjective video assessment

Subjective quality tests aim to capture the user's perception and understanding of quality. As we pointed out earlier, the user's perception of quality in not uni-dimensional and it depends on many factors. Besides the quality of the viewed material per se, user perception is also content specific, i.e., whether the video material is interesting and intriguing or not. It has also been recognized that what determines quality also depends on the purpose of the interaction and the level of user's engagement. The extent to which QoS is perceived as degraded depends upon the real-world task that the user is performing. Furthermore, depending on the application, the quality of the background sound is also highly important. For example, it is shown that subjective quality ratings of the same video sequence are usually higher when accompanied by good quality sound, as this may lower the viewers' ability to detect impairments [56].

3.7.2 Objective metrics of video quality

We stressed earlier that although subjective procedures for measuring video quality still constitute the most reliable method for gaining insight knowledge about the performance of digital video transmission systems, the complicated and costly setup of subjective tests makes this method particularly unattractive for automating the assessment procedure. The involvement of human subjects in this process makes this approach unusable when the quality monitoring systems have to be embedded into practical processing systems. For this reason, quality metrics are able to produce objectively obtained ratings present an attractive alternative. Objective quality metrics have been the subject of research for several years.

The first models were designed to work on analogue video transmission. However, the recent advent of digital manipulation and transmission of video means that video material is effected in a completely different way, leading to different types of impairments. The simplest form of measuring the quality is by calculating the distortion at the pixel level. The peak signal-to-noise (PSNR) measures the mean squared error (MSE) between the reference and test sequences have been extensively used by the image processing community. Due to its simplicity it is still being used [66].

Chapter four
system design

The huge uses of digital images and videos via communication media include the internet leads to volumetric growth of data flow through these media. This enforced researchers to develop efficient data compression techniques. Discrete Wavelet Transform (DWT) is an efficient method that can be used to perform the compression technique.

In this chapter a detailed description of the implemented system is introduced, which focuses on image compression approaches in the first part, including DWT, Image Compression, Wavelet Masks, Wavelet Families, while the second part describes the methods developed for video compression approach including frames extraction, frame difference, frame selection key and color analysis.

4.1 Image Compression Approach

The implemented approach of image compression developed via efficient image compression method is based on the selection of the wavelet mask. Many wavelet masks are applied in this work to different types of images. The selection of the mask type depends on many factors such as compression performance, image details and image size. A typical image compression system is shown in figure (4.1). This system consists of three closely connected components named

Transformer, Quantizer and Encoder. Image compression is achieved by applying a linear transform in order to decor relate image data, then quantizing the resulting transform coefficients and then applying entropy coding on the quantized values.

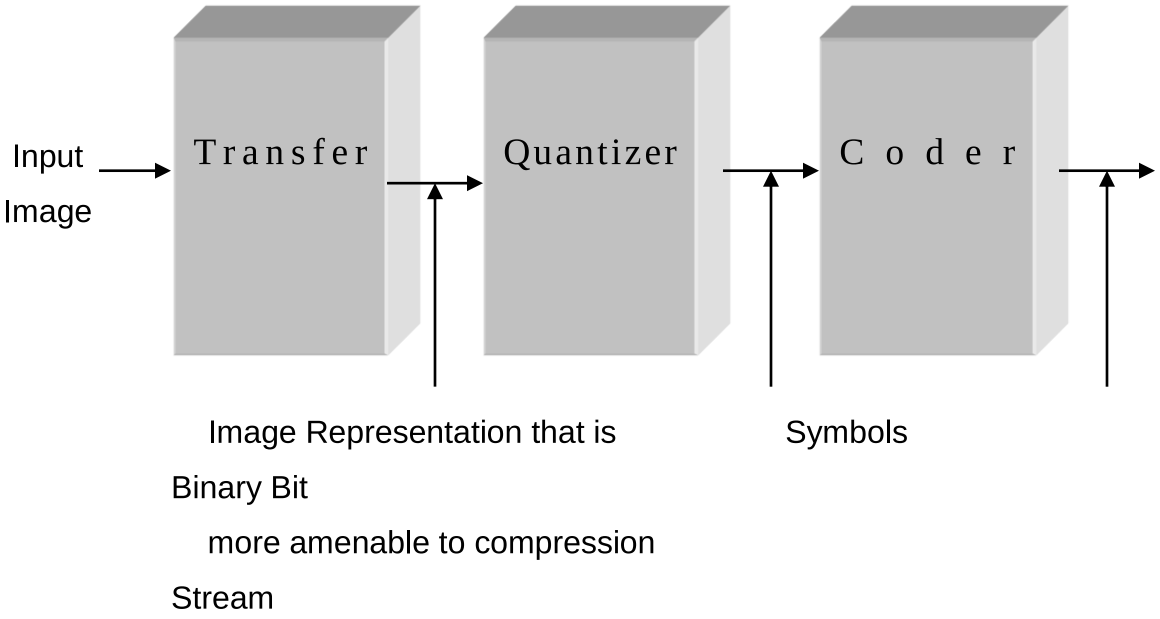


Figure (4.1) typical image compression system

4.1.1 Images contents

The image content being viewed influences the perception of quality irrespective of technical parameters of the system. Normally, a series of pictures, which are average in terms of how difficult they are for system being evaluated, has been selected. To obtain a balance of critical and moderately critical material, different types of test images with different frequency content were used. Choice of wavelet function is crucial for coding performance in image compression. However, this choice should be adjusted to image content. The compression performance for images with high spectral activity is fairly insensitive to choice of compression method (for example, test image original). On the other hand, coding performance for images with moderate spectral activity (for example, test image nature) are more sensitive to choice of compression method.

4.1.2 Two Dimensional DWT

Two Dimensional Discrete Wavelet Transform (2D-DWT) is obtained via the implementation of filter pair i.e. low pass and high pass filters on rows and columns of the input image respectively. A Low Pass Filter (LPF) and a High Pass Filter (HPF) are chosen, such that they exactly half the frequency range between themselves. This filter pair is called the analysis filter pair. First, the low pass filter is applied on each row of matrix data, thereby getting the low frequency components of the rows. But since the LPF is a half band filter, the output data contains frequencies only in the first half of the original frequency range. The HPF is applied on the same rows of data, and similarly the high pass components are separated, and placed by the side of the low pass components. This procedure is done for all rows. As mentioned above, the Low-Low (LL) band at the highest level can be classified as the most important band, and the other 'detail' bands can be classified as of less important bands, with the degree of importance decreasing from the top of the pyramid to the bands at the bottom. DWT is a multispectral technique used for converting signal or image into four different bands such as Low-Low (LL), Low-High (LH), High-Low (HL) and High-High (HH) as demonstrated in figure (4.2).

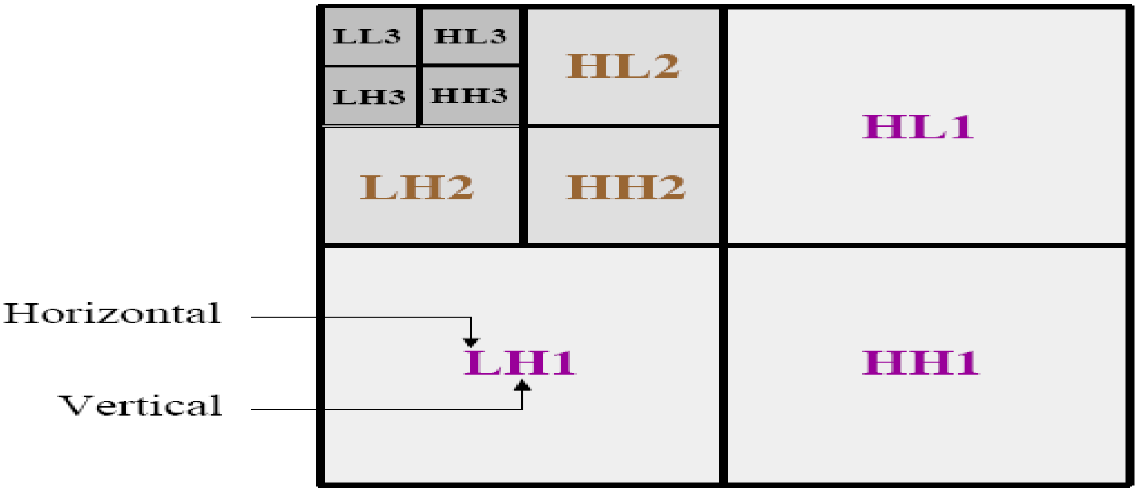


Figure (4.2) Decomposition of image applying DWT

4.1.3 Two Dimensional Wavelet Masks

There are many different types of wavelet masks, this work concentrated on the masks that shown in chapter 2 table (2.1) using wavelet families (Haar, Daubechies, Symlets, Coiflets, Biorthogonal and Reverse Biorthogonal).

4.1.4 Performance Measurements

In general, two classes of methods are available to measure image quality, **Subjective tests**, in which human subjects are asked to assess or rank the viewed material, and **Objective tests**, which are computational models that measure the quality by comparing the original and the compressed images.

Subjective tests may produce the most accurate ratings but, they require costly and complex setup and viewing conditions and thus, they are inflexible to use. Objective quality metrics, on the other hand, are based purely on mathematical methods, which will give important insights into the tradeoffs between file size and image quality.

4.1.4.1 Objective Measure

Two of the error metrics used to compare the various image compression techniques are the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error. The mathematical formulas for the two measures are given in equation (4.3) and (4.4) respectively:

$$MSE = \frac{1}{NM}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x,y) - f*(x,y)\right]^2 ....4.3$$

Where ***f(x, y)*** is the original image data and ***f*(x, y)*** is the compressed image value. While M, N are the matrix dimensions in x, y respectively.

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20.\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)..........4.4$$

Here, ***MAX_I*** is the maximum possible pixel value of the image.

The MSE & PSNR are two error metrics used to compare image compression quality between original images to the decompressed image. Since the MSE is a measure of error, so a lower value means better quality. While PSNR is a ratio in logarithm scale between (2^k-1) ^2 (the square of the highest gray level in the image and the MSE. A higher value of the PSNR suggests less distortion and hence a better decompressed image.

4.1.4.2 Subjective Measure

Subjective evaluation by viewers is still a method commonly used in measuring image quality where the observers are simply asked to assess the overall quality. When taking subjective test, viewer's focus on the difference between reconstructed image and the original image. A **Questionnaire** check list is implemented then the results are collected for different viewers. Each viewer compared the compressed image, reconstructed image and the original image and gave the score. The proposed **Questionnaire** for evaluating image compressions is given in Table (4.2):

In multimedia (image, video and voice) especially when codec's are used to compress the image or video, the **mean opinion score (MOS)** provides a numerical indication of the perceived quality of received image after compression and/or transmission. The MOS is expressed as a single number in the range 1 to 5, where 1 is lowest perceived image quality, and 5 is the highest perceived image or video quality measurements.

The MOS is generated by averaging the results of a set of standard, subjective tests where a number of observers rate the quality of test images or videos. An observer is required to give each sentence a rating using the following rating scheme given in Table (4.1).

Table (4.1): Mean Opinion Score (MOS)

MOS	**Quality**	**Impairment**
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

The MOS is the arithmetic mean of all the individual scores, and can range from 1 (worst) to 5 (best) as given in the following equation.

Suppose we have sample space

$$Xn = \{x_1, \ldots, x_n\} \qquad \ldots\ldots\ldots.4.5$$

Where x1, x2,…..,xn represented the individual scores obtained from the rate of test images or videos, whereas n represent the total number of the selected observers.

Then the arithmetic mean *A* is defined via the equation

$$A = \frac{1}{n}\sum_{i=0}^{n} xi\ldots\ldots\ldots.4.6$$

An image quality factors were collected and tested for different viewers as shown in Table (4.2).

Table (4.2) Image Quality Factors and Scores

	score				
Image Quality Factors	**1 Bad**	**2 Poor**	**3 Fair**	**4 Good**	**5 Excellent**
Degree of the overall clearance of image.					
Degree of recognizing overall of image.					
Degree of clearance details of image.					
Degree of clearance the colors in the image.					
Degree of non distortion in the image.					
Degree of the non noise in the image.					
Degree of the non noise effect in the overall image.					
Degree of the non distortion in the image.					

4.1.5 Image Compression System

The implemented system of image compression is concentrated on the adequate selection of wavelet mask that implemented for image compression technique. This system is implemented via many stages as illustrated in Figure (4.3) and these stages are listed below:

- Filtering stage that generates the de-nosing image which deals with suppressing of high frequency noise from the original image.
- 2D DWT stage that generates the LL-band. This concentrated on applying LPF of rows and columns of the image obtained from step one. Then 2D DWT is obtained via the

- implementation of low pass and high pass filters on rows and columns of image respectively. A low pass filter and a high pass filter are chosen such that they exactly half the frequency range between themselves. Thus the 2-D DW decomposes the image into sub-images, three details and one approximation. The DWT separates an image into a lower resolution approximation image (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail components. For example, HL means that we used a high-pass filter along the rows, and a low-pass filter along the columns.
- Performance measure that generates the measures of MSE and PSNR. Both measures are applied on the original and compressed image.
- Threshold calculation stage that generates the selected threshold, to be prepared for decision making.
- Compare measures that generate the results measured to find the difference between the frames.
- Decision making that denotes the selected mask in appropriate one to be applied.

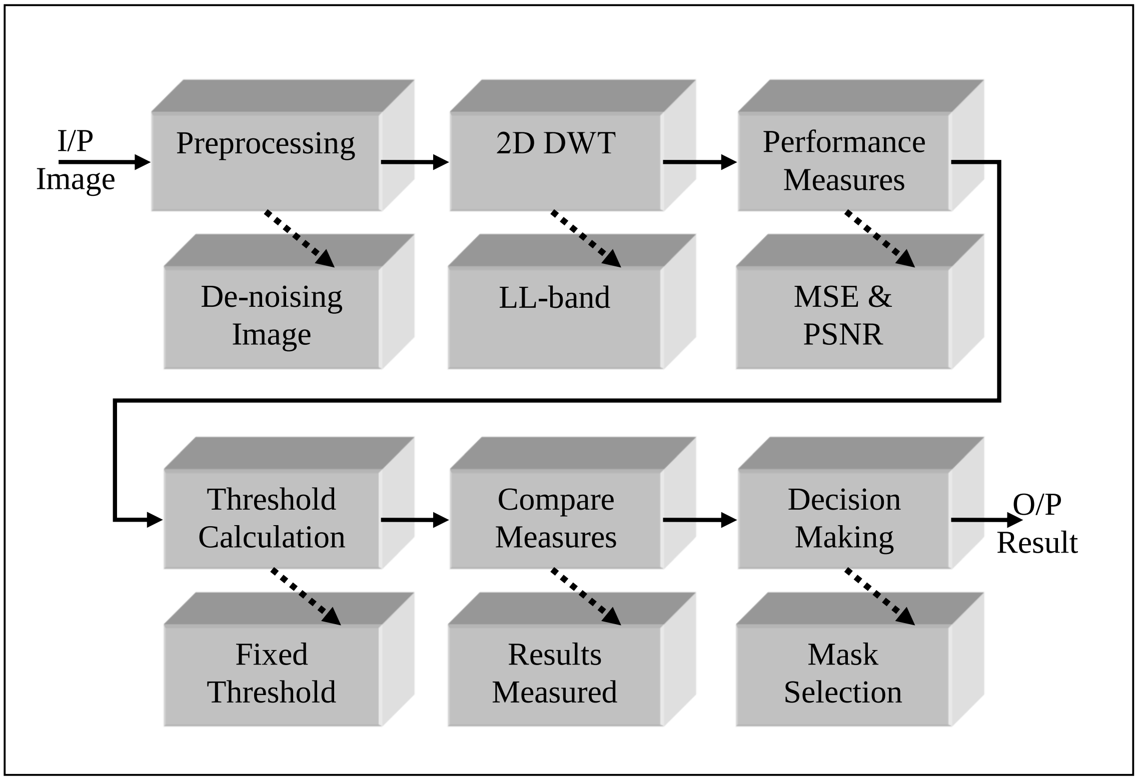


Figure (4.3): Implemented mask selection compression technique

Also different levels of decomposition were selected for Image & Video Compression as shown in figure (4.4).

Different types of images are collected and prepared to be used as test images as shown in Figure (4.5).

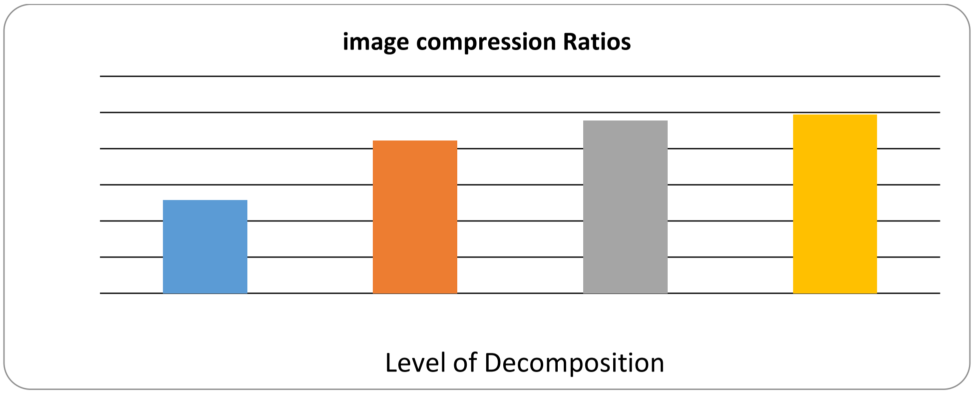
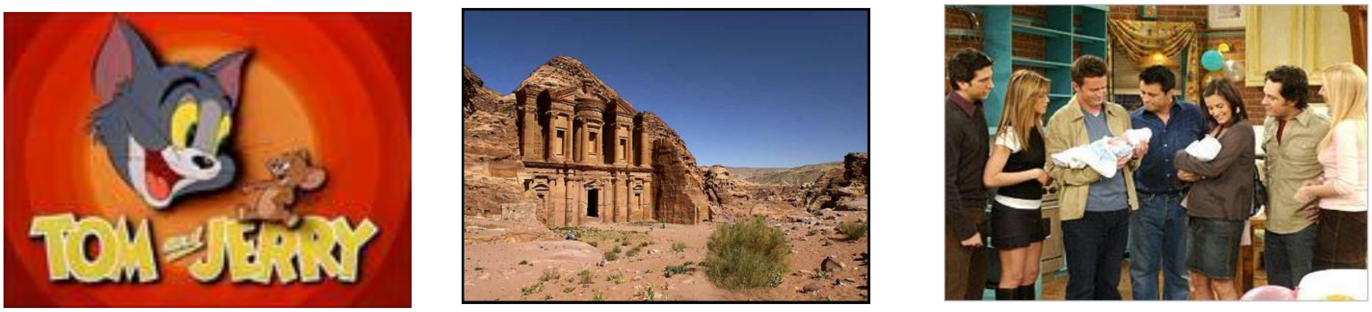


Figure (4.4) levels of Decomposition for Image compression.

(a) NEWS (b) SPORT (c) MUSIC



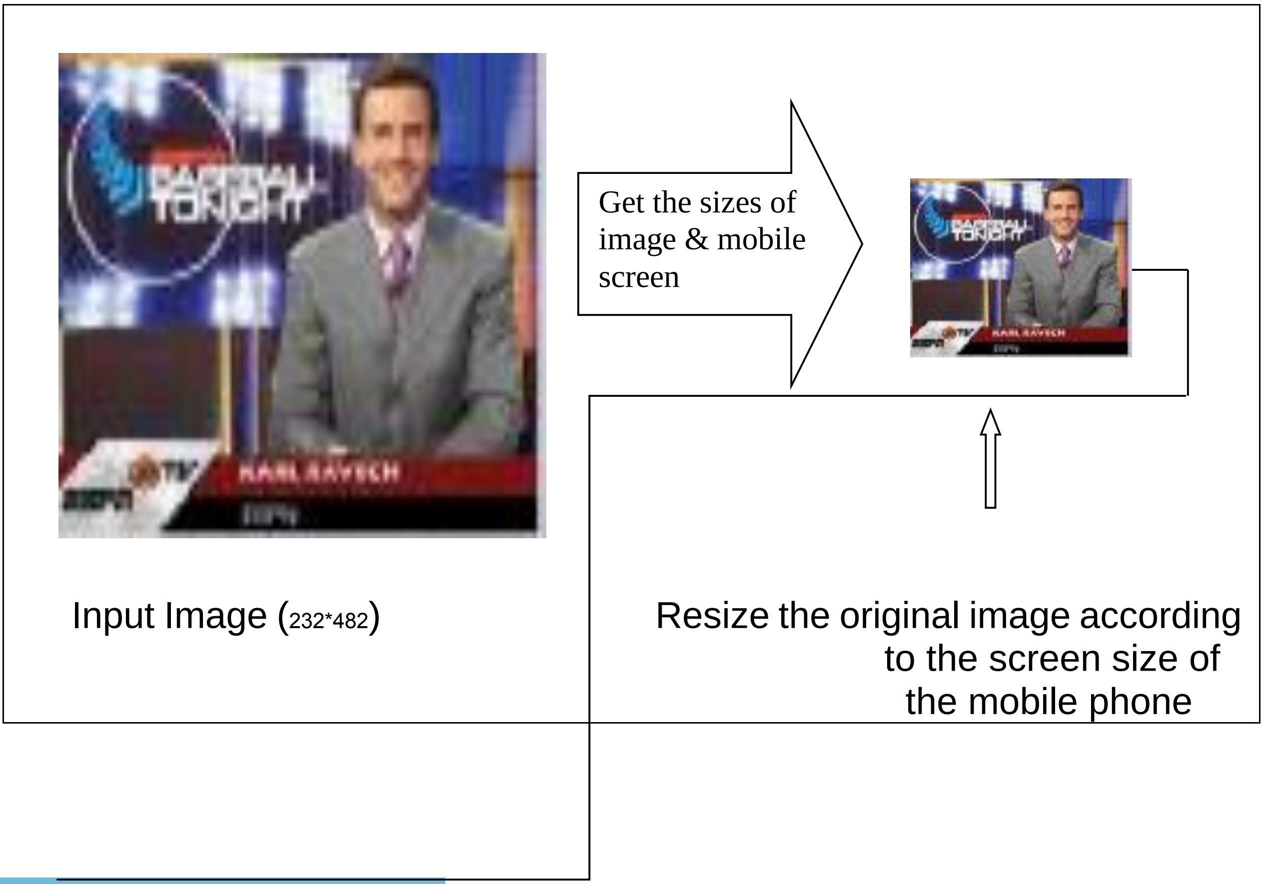
(d) NEWS (e) PANORAMA (f) SERIES

Figure (4.5): Different types of tested images

Algorithm (4.1) summarizes the main steps of image compression to be adopted in mobile devices. In which the original image is read using Mat lab code. Then transfer the size of the original image into adequate size using image resize function and store it in a variable called I where I represents the two dimensions of the rows and columns of the original image. After that the screen size of the mobile phone must be measured. Then the integer ratio called D must be calculated by dividing the original size of original image to the screen size of the mobile phone. Apply the 2D-DWT compression up to level D that obtained from the previous step using different types of wavelet families(Haar, Daubechies, Symlets, Coiflets, Biorthogonal and Reverse Biorthogonal). Finally apply the performance assessment measures that generate the measures of MSE& PSNR in addition, a subjective evaluation measures by viewers is used to measure the image quality where the observers are simply asked to assess the overall quality of the compressed image with the comparison of the original image. The structural diagram of image compression can be shown in Figure (4.6).

Algorithm (4.1): Image Compression Steps to be adopted on mobile devices

```
Step 1: Read i/p image
Step 2: Get i/p original image size (I size)
Step 3: Get the mobile image size (M size)
Step 4: Calculate the division factor Int ((D) =
            I size/ M size)
Step 5: Apply 2D-DWT up to level D.
Step 6:  Apply Performance measures that generate the measures of MSE
          & PSNR,  in addition to Viewer Assessment /Evaluation.
Step 7: Decision making for the obtained results.
Step 8:  Adopt the compressed image to mobile devices.
```



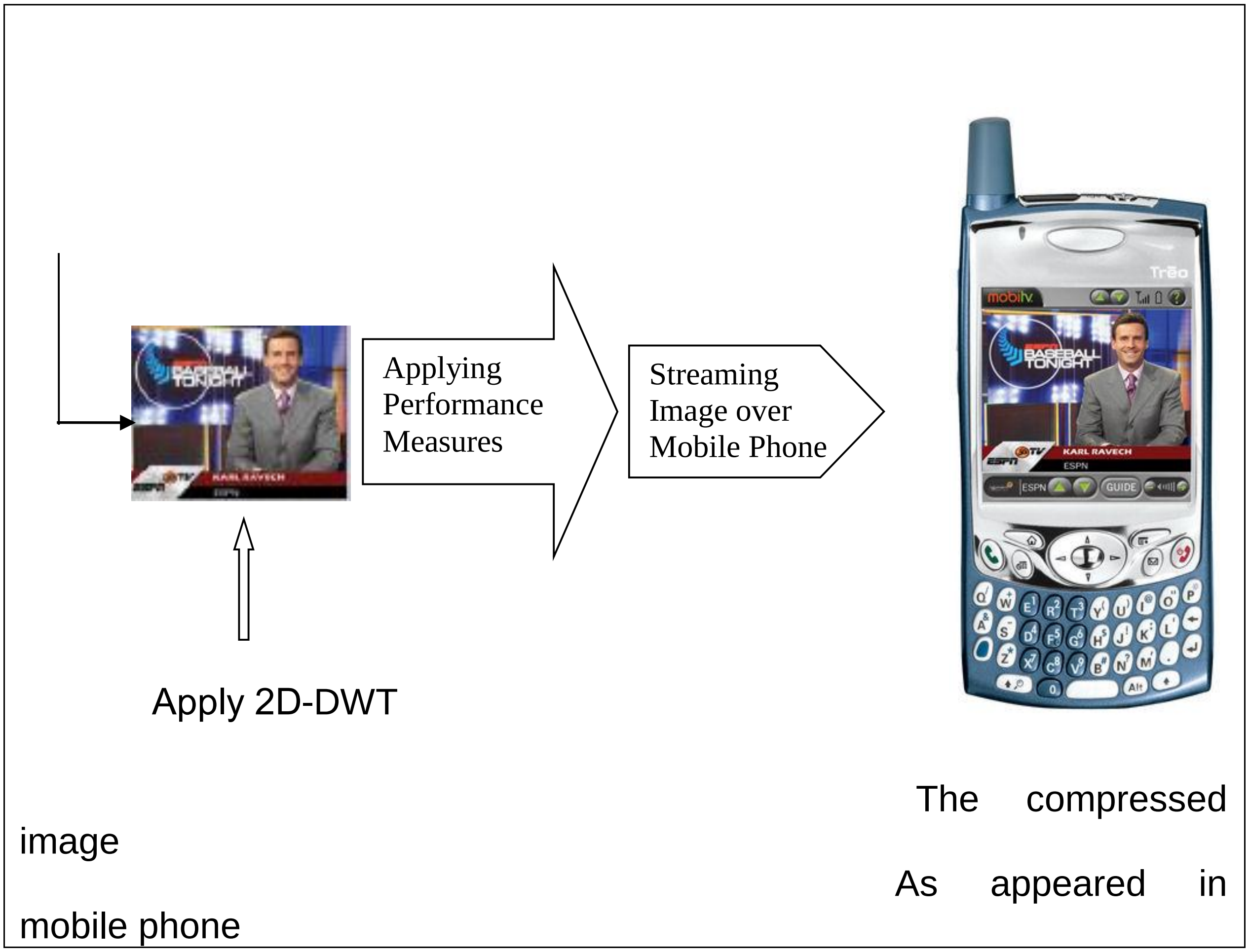


Figure (4.6): Structural Diagram of Image Compression to be adopted in mobile devices.

4.2 Video Compression System

The implemented approach of video compression deals with the developing of an efficient video compression approach based on frame selection key that concentrated on the calculation of frame near distance (or difference between frames). The selection of the meaningful frame depends on many factors such as compression performance, frame details, frame size and near distance between frames. For a given video sequence, a set of low-level spatial features is extracted as a low-level description of the video sequence. These features include color, texture, edges, shape, etc.

The focus here is on the most expressive and widespread spatial feature; color. In this work, many videos are used and tested to insure the efficiency of the implemented technique.

4.2.1 Video Content

Content video scenario is specified by the environment of usage, streamed content and the screen size of the mobile terminal. Therefore, the mobile scenario is strictly different in comparison with the classical TV broadcasting services or broadband IP-TV services. The mostly provided mobile contents are news; cartoons, sport, music, series and panorama (see figure 4.5).

The character of motion is determined by the amount and direction of the motion between two scene changes. For mobile video streaming we define the five most frequent contents with different impact on the user perception.

1) Content class (News): The first content class includes sequences with a small moving region of interest (face) on a static background. The movement in the Region of Interest (ROI) is mainly determined by eyes, mouth and face movements. The ROI covers up to approximately 15% of the screen surface.

(a) Snapshot of typical class (News).

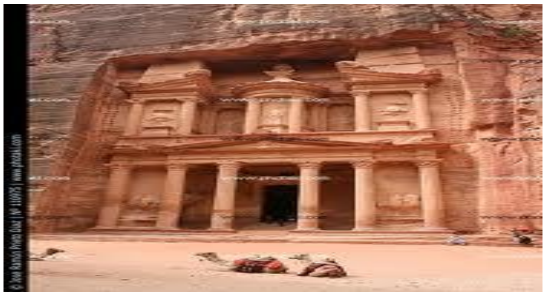
Content class (Sport): This content class contains wide angle camera sequences with uniform camera movement (panning). The camera is tracking small rapid moving objects (ball, players) on the uniformly colored (typically green) background.

(b): Snapshot of typical class (Sport).

3) Content class (Cartoon): In this content class, object motion is dominant, the background is usually static. The global motion is almost not present due to its artificial origin of the movies (no camera). The movement object has no natural character.

(c): Snapshot of typical class (Cartoon).

4) Content class (Panorama): Global motion sequences taken with a wide angle panning camera. The camera movement is uniform and in a single direction.



(d): Snapshot of typical class (Panorama).

5) Content class (Rest: music, series,...): the content class contains a lot of global and local motion or fast scene changes. Scenes shorter than three seconds are also associated to this content class. The content class covers scenes which do not fit any of the previous four classes.

(e) & (f): Snapshot of typical class (Music & Series).

4.2.2 Two Dimensional DWT

The 2-D DWT decomposes the image into sub-images via the implementation of low pass and high pass filters on rows and columns of image respectively,

three details and one approximation, the 2-Dimensional wavelet transform can be done in a separable fashion, meaning that we can use a 1-Dimensional DWT and apply it horizontally, then vertically.

To accomplish the 1-D DWT, we split a 1-D input signal into two streams, and filter each. We have a low-pass filter (h) and a high-pass filter (g), corresponding to the scaling and wavelet functions, respectively. Filtering is a basic operation in digital signal processing where the input signal (i.e. a row or column from an image) is convolved with a set of coefficients.

DWT Algorithm

For images, there exist an algorithm similar to the one-dimensional case for two-dimensional wavelets and scaling functions obtained from one- dimensional ones by tonsorial product.

This kind of two-dimensional DWT leads to a decomposition of approximation coefficients at level *j* in four components: the approximation at level j + 1, and the details in three orientations (horizontal, vertical, and diagonal).

Lo_D is the decomposition low-pass filter.

Hi_D is the decomposition high-pass filter.

Lo_D and Hi_D must be the same length.

[cA,cH,cV,cD] = dwt2(X,Lo_D,Hi_D) computes the two-dimensional wavelet decomposition as above, based on wavelet decomposition filters that you specify.

Figure (4.7) describes the basic decomposition steps for images. Whereas algorithm (4.2) summarizes the main steps of 2D-DWT.

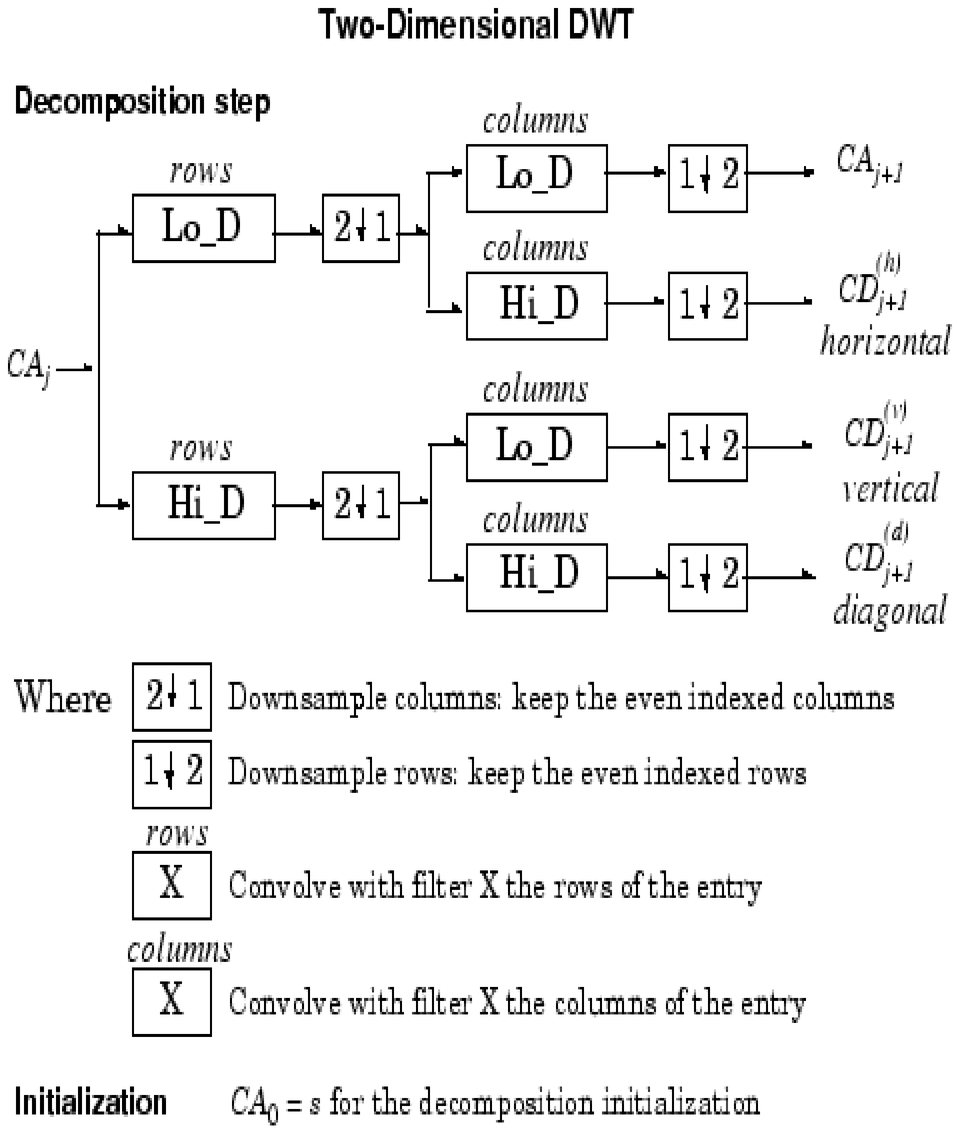


Figure (4.7): Basic decomposition steps of 2-Dimnsional DWT image.

Algorithm (4.2): 2D DWT

Step 1: Read i/p image
Step 2: Apply LPF on rows
Step 3: Apply LPF on columns → LL
Step 4: Apply HPF on rows
Step 5: Apply HPF on columns of step 1 →LH
Step 6: Apply LPF on columns → HL
Step 7: Apply HPF on columns → HH

4.2.3 The Implemented System of Video Compression

The implemented system of video compression concentrate on the adequate frame selection key in which used for the compression technique. This system is implemented via many stages as illustrated in figure (4.8) and these stages are listed below:

- Preprocessing stage that applying LPF to generate the de-nosing video that illuminates high frequency noise and redundancies.
- Frame extraction stage that generates frames for a certain time, in which different frames per second were selected in the tested videos (10 frames per second, 15 frames per second, 20 frames per second and 25 frames per second).Varies types of video including news, sports, cartoons, etc are used to evaluate the proposed approaches.
- Frame selection stage that generates effective frames, depending on the near distance between consecutive frames. By

- calculating frame difference between each consecutive frames. In which frame selection is a popular approach for key-frame extraction. For that frame difference
- measures using absolute difference of the color histogram were applied to find the near distance between each consecutive frame.
- Frame reordering that prepared frames to be compressed. Different approaches were implementing via removing certain frames before prepared to compression. Three different approaches are suggested for removing the lowest frame difference in which we remove the very similar frames.
- Applying 2D DWT stage that generates the LL-band, via applying LPF for rows and columns. Different wavelet families are applied with different level of decomposition.
- Video construction that construct the final compressed video, to perform video streams.

The main core of this system is the frame selection key, in which pass the similar frames and select the different frames depends on a certain specified threshold. In addition the selected frames are compressed via applying two dimensional discrete wavelet transform as illustrated in image compression system.

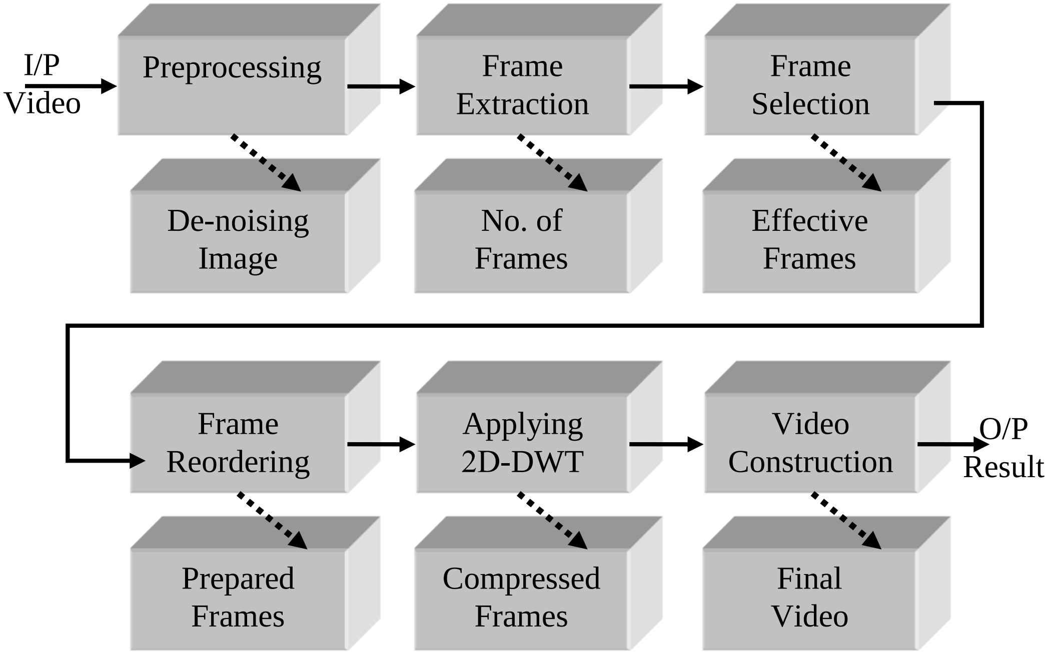


Figure (4.8) the implemented frame selection key approach

4.2.4 Frames Difference Approach

A color histogram is a representation of the distribution of colors in an image, Color histograms are frequently used to compare images because they are simple to compute, and tend to be robust regarding to small changes in camera view point. Computationally, the color histogram is formed by counting the number of pixels belonging to each color. Usually a color quantization phase is performed on the original image in order to reduce the number of colors to be considered in computing the histogram and thus the size of the histogram itself. There are many methods to compare color histograms; one simple method is the absolute difference between two color histograms.

The color histogram difference Diff rgb(Ii,Ij) between two consecutive frames Ii,Ij can be calculated as below:

$$\text{Diff } \mathbf{rgb} \text{ (Ii,Ij)} = \sum_{k=1}^{n}(| H^{r}i\ (k)-H^{r}j\ (k) | + | H^{g}i\ (k)-H^{g}j\ (k) | + | H^{b}i\ (k)-H^{b}j(k) | \quad \ldots \quad (4.7)$$

Where, $H^{r}i(k)$ and $H^{r}j(k)$ stand for the histogram of Ii and Ij respectively.

A short transition occurs when Diff rgb (Ii, Ij) is greater than a given threshold, and selecting an appropriate threshold is the key to the method.

The implemented system for calculating frame difference can be summarized as follows:

1- Read the consequences of frames.

2- Calculate the color histogram for every frame.

3- Take the absolute difference between the color histogram of each two consequences of frames as given in equation (4.7).

4- Summing the total absolute differencing where, the absolute difference between each two consecutive frames is greater than some thresholds. Then different threshold values (3, 4, and 5) are selected to be applied.

4.2.5 Video Compression Steps

Different stages where implemented via video compression system, in this section we will briefly illustrate the main function of each stage.

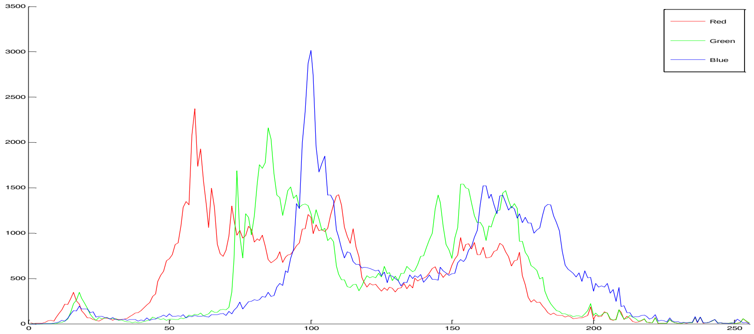
***Images resizing**

The software image resizing is a tool to increase or decrease the dimensions of frames or images. This tool is used for adjusting the dimensions of the images before starting the compression procedure. Images resize scans recursively the directory provided and find all the images located in the specific directory in the hard disk. This step helps to decrease the time and the number of operations needed in image compression. This step aims to decrease the size of frames according to the original size of video frames and the detailed of the video frames, in order to accommodate with the size of mobile screen, therefore different resizing where taken in our experiments. The screen size of mobile devices is considered in this work. The most popular size of mobile sizes is (128*128) for different types of mobile phones.

*Color histogram

Color histograms are frequently used to compare images because they are simple to compute, and tend to be robust regarding small changes in camera view point. An image histogram H ( ) refers to the probability mass function of image intensities. Computationally, the color histogram is formed by counting the number of pixels belonging to each color. Usually a color quantization phase is performed on the original image in order to reduce the number of colors to consider in computing the histogram and thus the size of the histogram itself. There are a number of ways to compare color histograms. Figure (4.9) shows a sample image and its color histogram

(a): Sample color image news.jpg



(b): color histogram

Figure (4.9): **(a)** Color Image and **(b)** its color histogram

***Frames Difference Measure**

There are a number of methods to compare histograms (absolute difference, Euclidean distances, histogram intersection). One simple method is the absolute difference between two color histograms as mentioned in equation (4.7).

In this way a lower distance value represents a greater similarity between images. This means we can remove the very similar frames according to the value of frame difference obtained. While less similarity means we have to keep the frames.

***Threshold Calculation**

It is clear that color histogram is organized into a number of pixels that fall within each color range where the histogram allows images with similar color distribution to be retrieved. And as it was explained earlier the absolute difference measure different selections threshold were tested like 3, 4, 5 and 6 to get the suitable threshold in calculating the absolute difference. It was concluded that 4 is a suitable threshold for obtaining good difference between each two frames color histogram.

***Remove Frames**

After calculating frames difference between each consecutive frame, then the frame difference method is applied to select the removable frame

- Zero difference
- Mean difference
- Percentage difference

***Re-extract frames**

After removing the lowest frames difference, re-extract the remaining frames which means keeping the highest frames difference to be ready for compression.

***Frames compression**

Different wavelets families and levels of decomposition are applied for

the extracted frames obtained from the previous step as explained in image compression.

***Reconstruct Video**

After applying the frames compression process, then the compressed frames were constructed to obtain the required video compression

***Measurement Video Quality**

Different criterion were applied to measure the equity of the obtained compressed video, in which, subjective evaluation by viewers is used to measure video quality in which the observers are simply asked to assess the overall quality of the obtained compressed videos . Such criterions are shown in table (4.3).

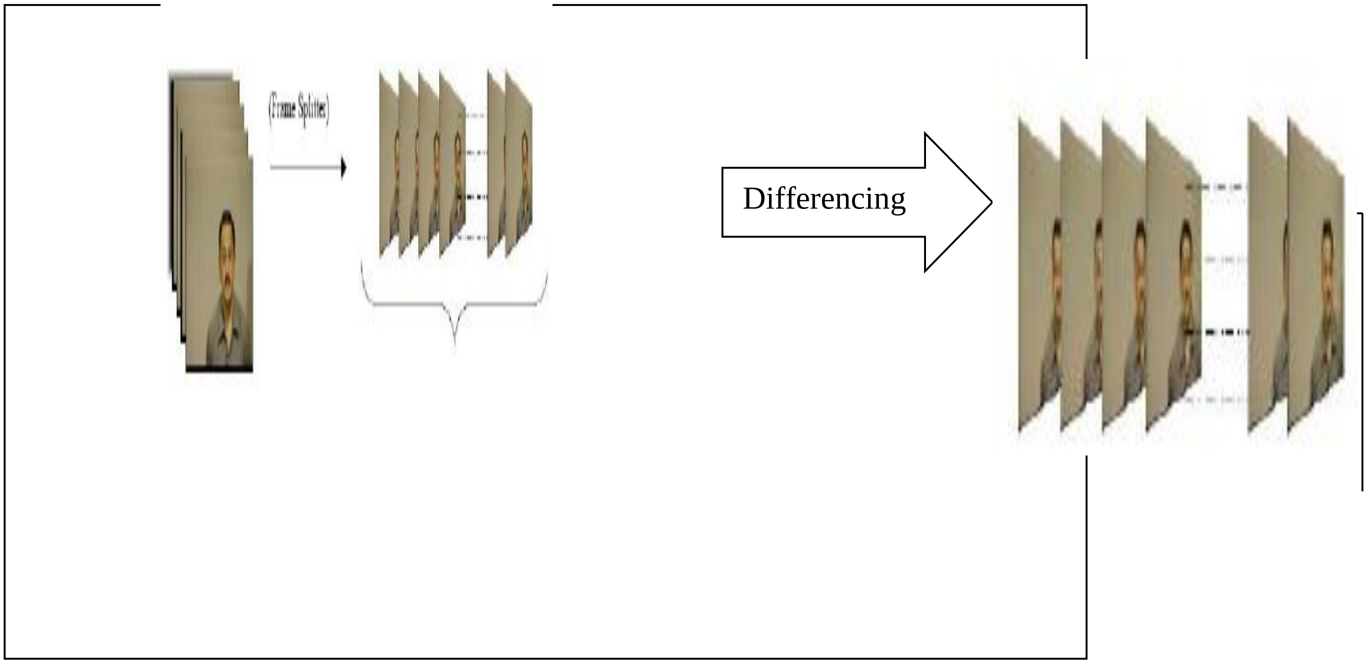
Table (4.3) Video Quality Factors and Scores

Video Quality Factors			**scores**		
	Bad	**Poor**	**Fair**	**Good**	**Excellent**
How would you rate video colors?					
How would you rate contrast?					
How would you rate video borders?					
How would you rate the movement continuity?					
Did you notice any flicker in the sequence?					
Did you notice any smearing in the sequence?					

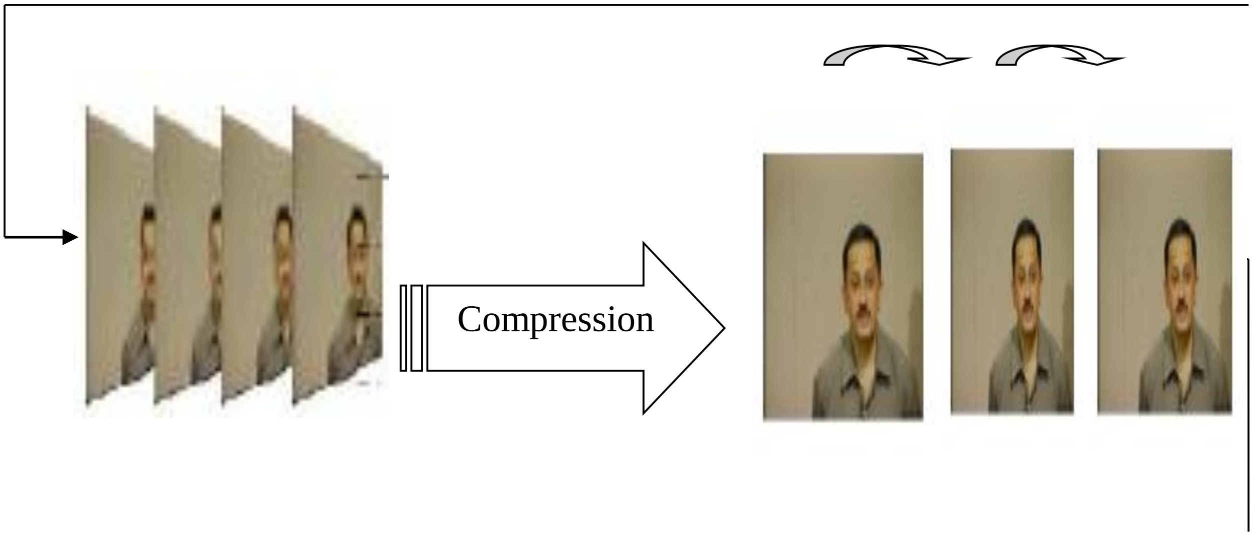
Algorithm (4.3) Summarizes the Video Compression Steps. Different types of video contents can be tested. For mobile video, the most frequent content are news; cartoons, sport, music, series and panorama. In which the original video were read using Mat lab code. Then extract the frames according to frames number starting from 10, 15, 20, 25 and 30 frames per second. Then calculate the frame difference between each consequence of frames, different values of thresholds were calculated and tested to obtain the frames differencing appropriately. Then remove the very similar frames according to video compression approaches given in section 4.2.6 (Zero difference, Mean difference, Percentage difference). For the remaining frames, apply image compression steps as given in the algorithm (4.2) to obtain the compressed frames. Then rebuild the compressed frames as a new compressed video. Finally apply the subjective assessment test for the obtained compressed video. In which the observers are simply asked to assess the overall quality of the compressed video with the comparison of the original video. Figure (4.10) shows the structural diagram of video compression.

Algorithm (4.3): Video Compression Steps to be adopted in mobile devices

```
Step 1: Get the i/p Video
Step 2: Extract frames from the video.
Step 3: Calculate the near difference.
Step 4: Remove the similar frames.
Step 5: Apply algorithm (4.2) for the remaining frames.
Step 6: Re-Extract the new frames as a compressed video.
Step 7: Apply subjective test for the compressed videos
Step 8: Decision making for the obtained results.
Step 9: Adopt the obtained compressed video to mobile devices.
```



(Input Video) (Frame splitter) (Calculate the Frames difference)



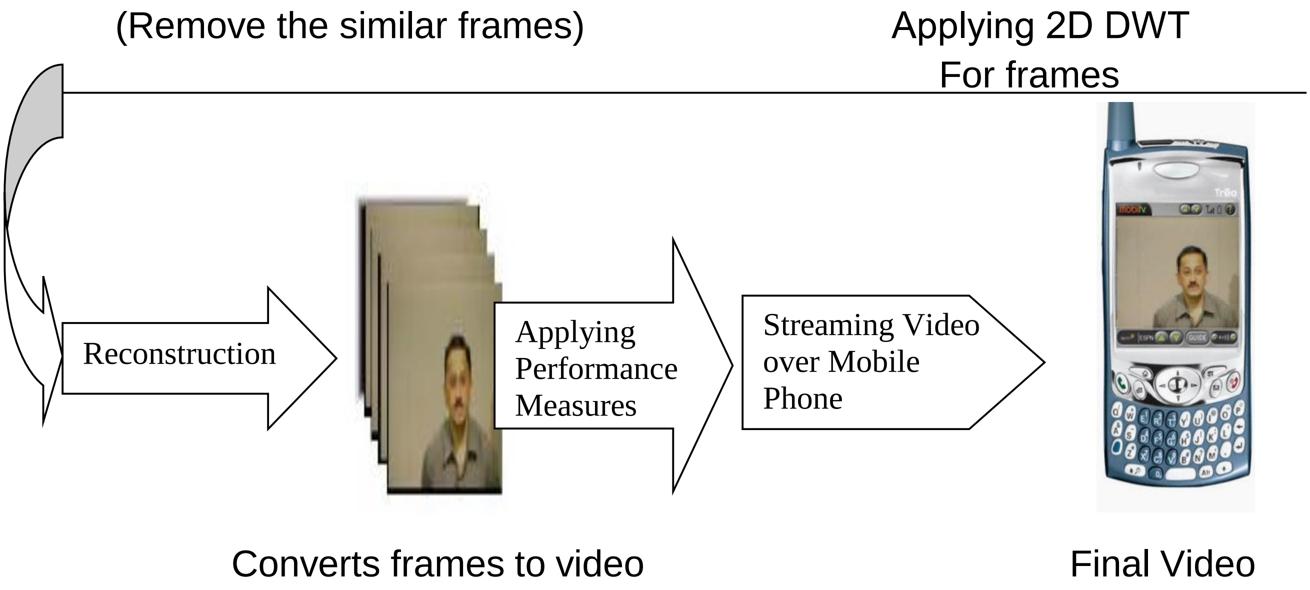


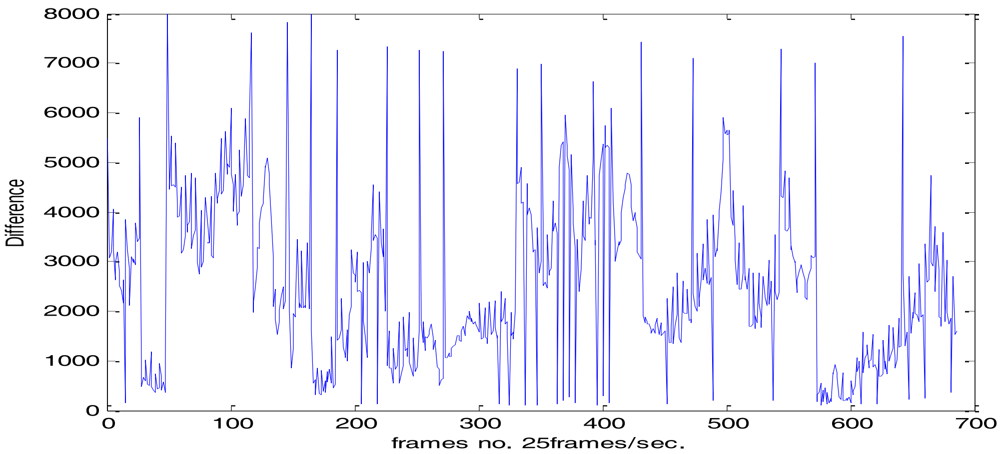
Figure (4.10): Structural Diagram of Video Compression to be adopted in mobile devices.

4.2.6 Video Compression Approaches

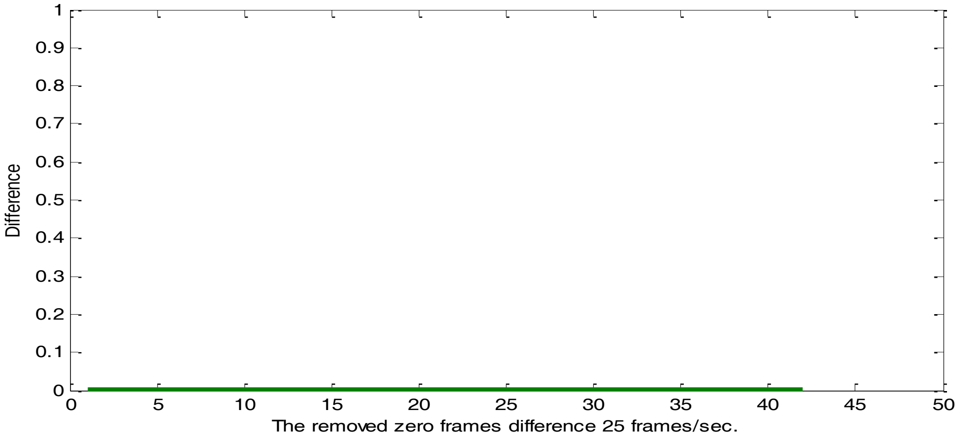
In this work three different approaches are applied for removing the lowest frame difference, the near similar frames are mainly removed, in which frames difference between each consecutive frame of the extracted frames are applied. Those approaches are Zero Difference, Mean Difference and Percentage Difference Approach.

a) Zero Difference Approach

In this approach frames are removed when the distance between any two consecutive frames is zero. It is clear in extracting video frames that many of the consecutive pair of frames has zero difference, especially when the numbers of frame extracted above twenty frames per second. In those cases this approach will remove that frames in which the frames difference is equal to zero. That will minimize the number of frames to be compressed and re-extracted. Figure (4.11) (a) shows the frames difference between each consecutive frames in which, the x-axis represents frames number that obtained after the extraction operation while the y-axis represents the values of differences between each two consecutive frames some of the frames difference have large values and some are less and others equal to zero. In this approach, values of frames difference equal to zero were removed. Figure (4.11) (b) shows that many of the frames number in which frames difference are equal to zero are removed as shown in the green line. The number of frames to be removed after the extraction operation (25 frames per second) is about 40 frames.

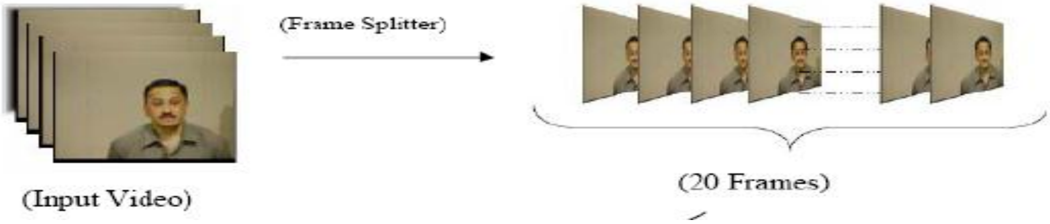


(a): Frames difference between each consecutive frame (Zero method).



(b) : The removed Zero difference frames

Figure (4.11): (a) frames difference & (b) the removed zero frames difference



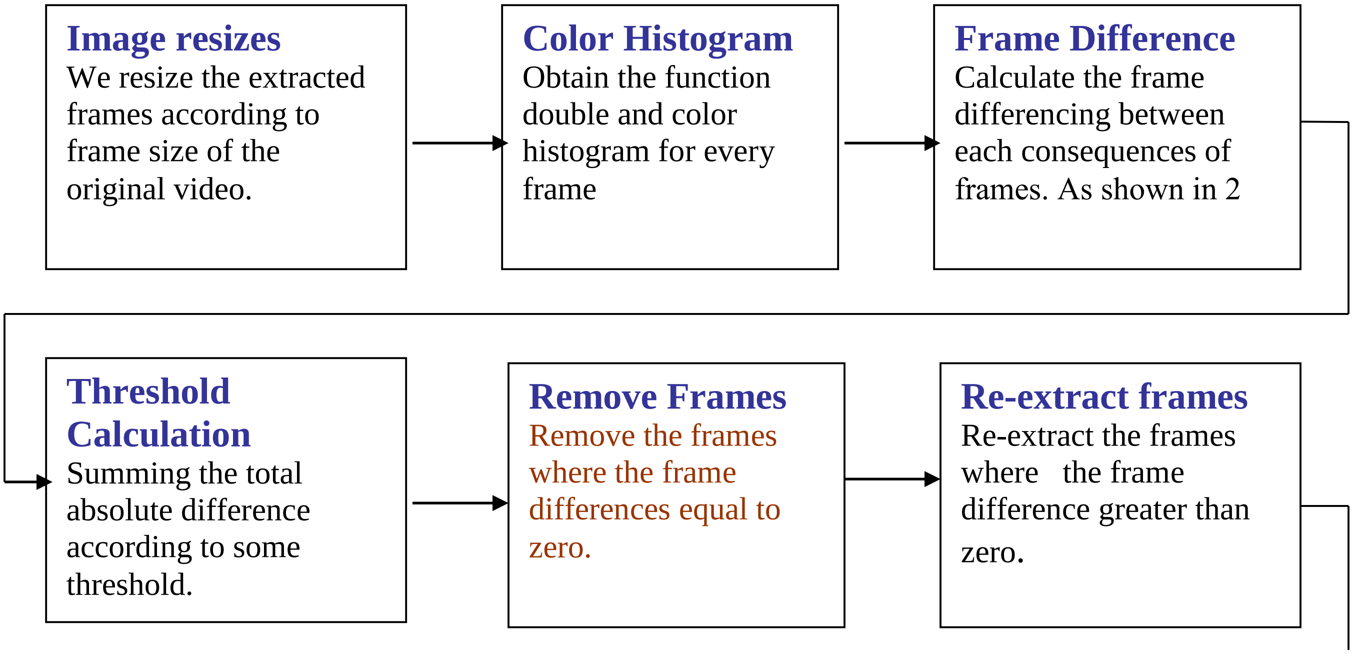


Image resizes
We resize the extracted frames according to frame size of the original video.

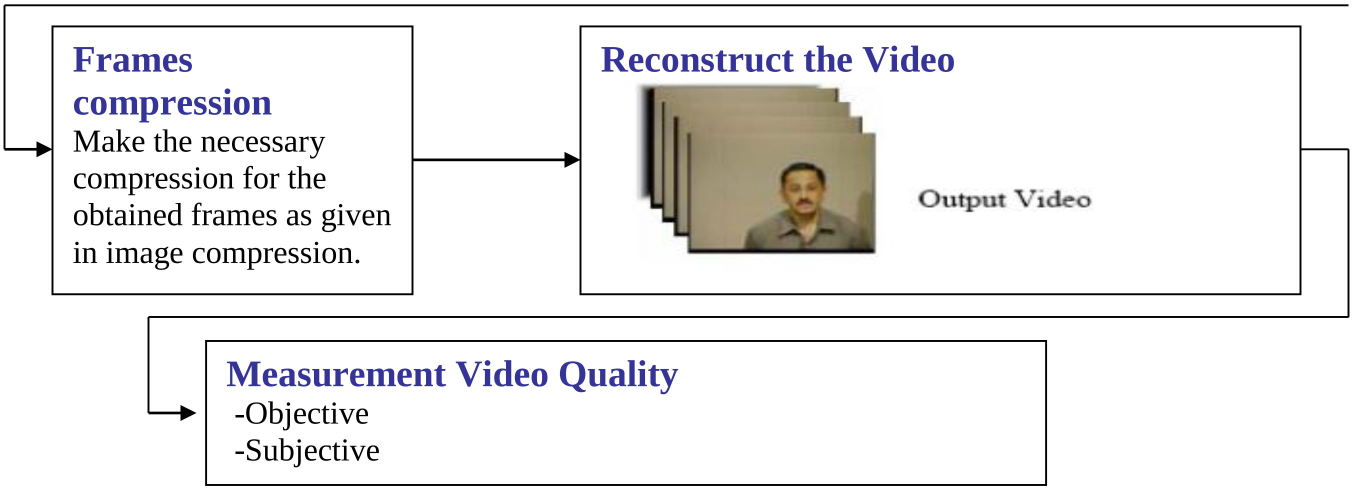
Color Histogram
Obtain the function double and color histogram for every frame

Frame Difference
Calculate the frame differencing between each consequences of frames. As shown in 2

Threshold Calculation
Summing the total absolute difference according to some threshold.

Remove Frames
Remove the frames where the frame differences equal to zero.

Re-extract frames
Re-extract the frames where   the frame difference greater than zero.



Make the necessary compression for the obtained frames as given in image compression.

Figure (4.12) Structure Diagram of the first approach (Zero Difference Approach).

b) Mean Difference Approach

In this method, the mean value of the frames difference is calculated. Mathematically the mean (Average) is obtained by dividing the sum of the observed values of frames difference by the number of observations. Then remove the frames where the frames difference between any consecutive frames is lower than the mean value of the overall frames difference. An input video were read and extracted with the rate of 15 frames per second. Then calculations of frames difference are obtained for each consecutive frame. The dotted line in Figure (4.13) shows the mean value of the frames difference. All frames difference less than the mean value that locates under the green line is removed. The structure diagram of the Mean difference approach is shown in Figure (4.14).

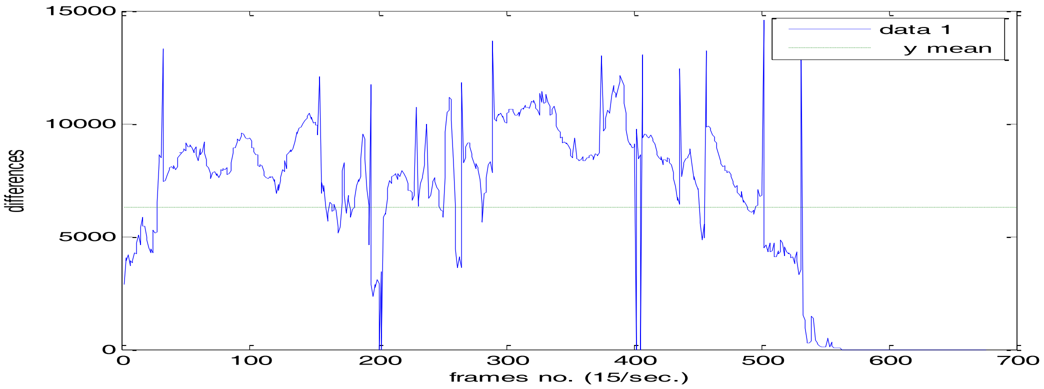
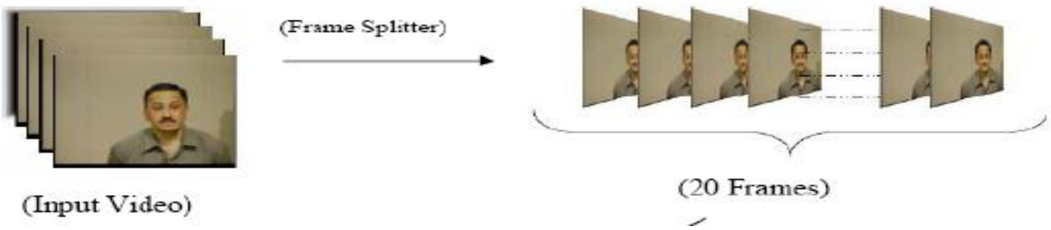
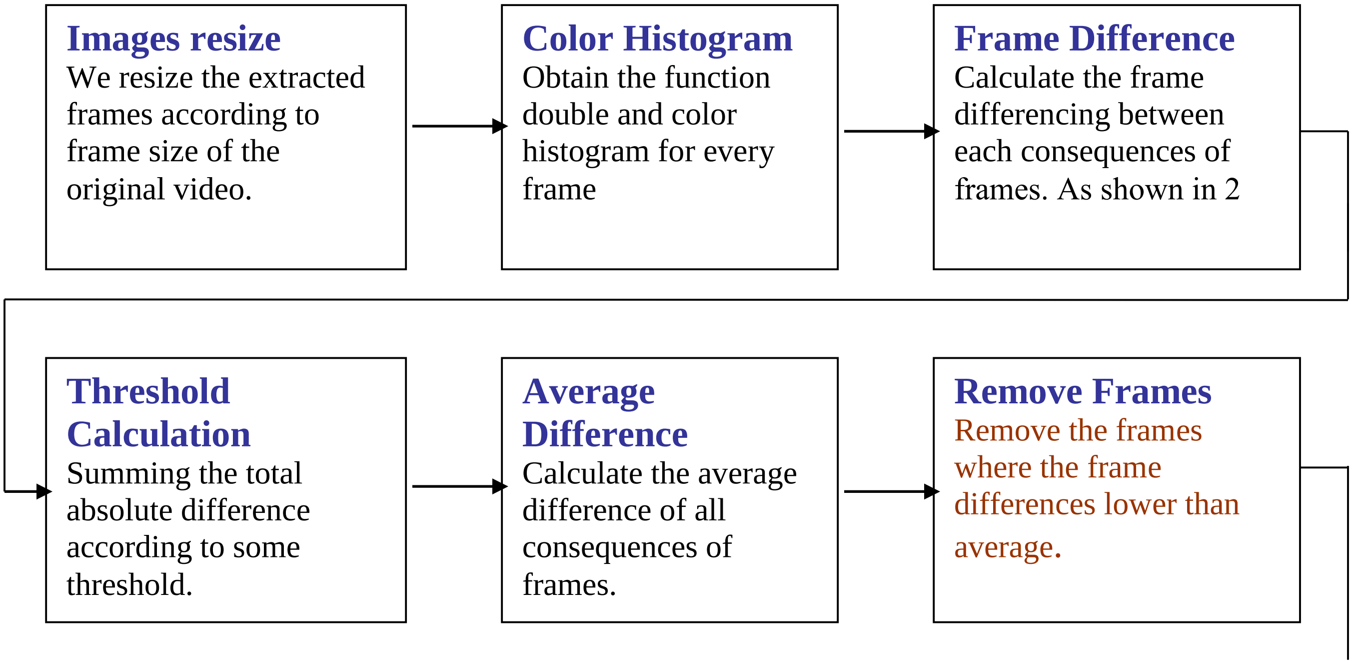


Figure (4.13) Frames difference between each consecutive frame (Mean Difference approach).





Images resize
We resize the extracted frames according to frame size of the original video.

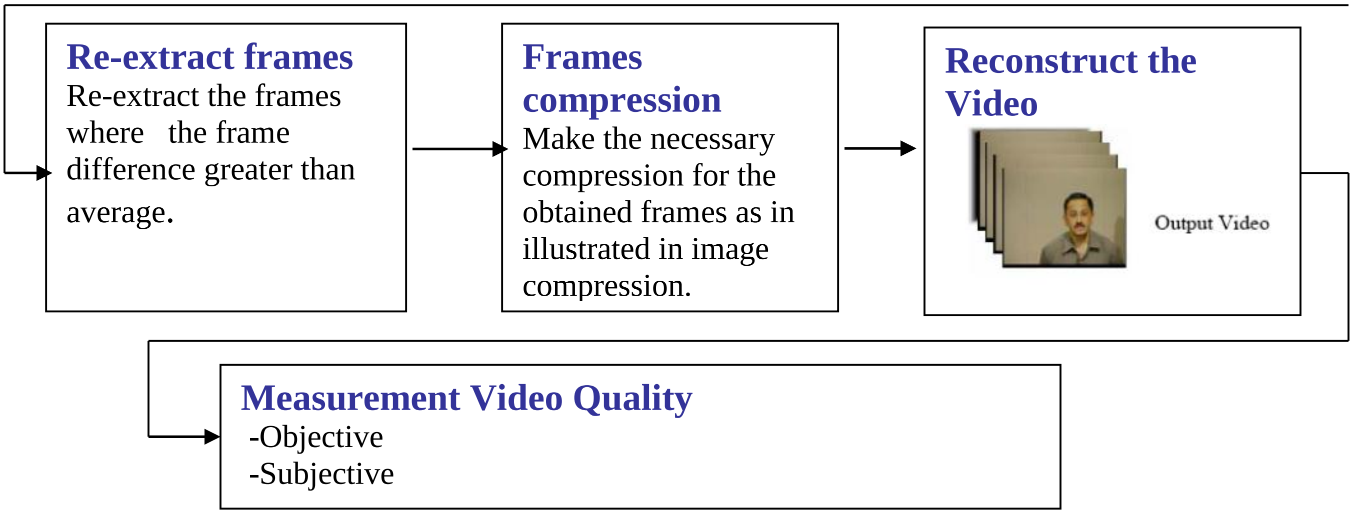
Color Histogram
Obtain the function double and color histogram for every frame

Frame Difference
Calculate the frame differencing between each consequences of frames. As shown in 2

Threshold Calculation
Summing the total absolute difference according to some threshold.

Average Difference
Calculate the average difference of all consequences of frames.

Remove Frames
Remove the frames where the frame differences lower than average.



Re-extract frames
Re-extract the frames where the frame difference greater than average.

Frames compression
Make the necessary compression for the obtained frames as in illustrated in image compression.

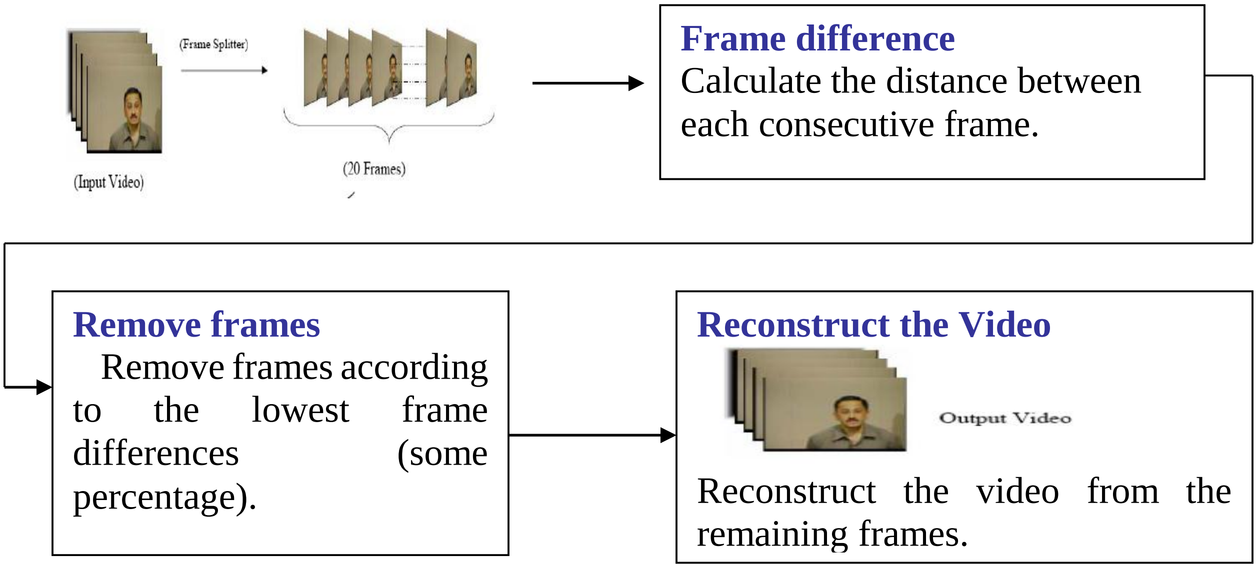
Reconstruct the Video

Measurement Video Quality
-Objective
-Subjective

Figure (4.14) Structure Diagram of the second approach (Mean Difference)

c) Percentage Difference Approach

In this approach different types of videos are examined according to frame details, frame size and the obtained frames differences. The implemented approach concerns mainly in removing the lowest frames difference as shown in figure (4.15).



Frame difference
Calculate the distance between each consecutive frame.

Remove frames
Remove frames according to the lowest frame differences (some percentage).

Reconstruct the Video
Reconstruct the video from the remaining frames.

Figure (4.15) structure diagram of the proposed frames difference percentage.

Let the distance between any consecutive frames be denoted by d(Ii,Ij) for all i=1,2,…..,n where n is the number of the extracted frames of the original video file.

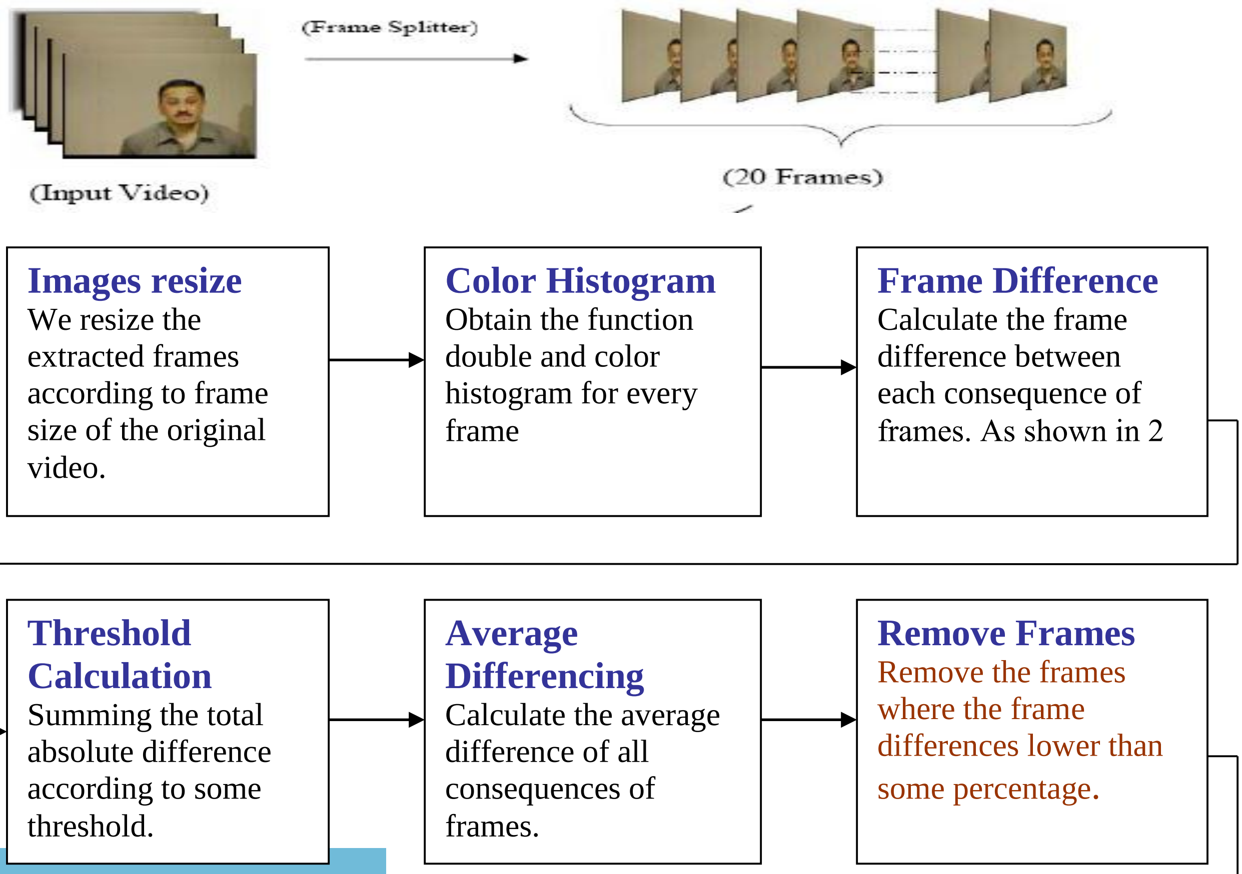
Define D as the set of all frame differences obtained from the extracted video file

The bigger value of the obtained frame difference di is the best represented by the original video as shown in equation (4.8):

D=max (di), i=1, 2….. n. …….. (4.8)

From equation (4.8), remove the lowest frames difference according to some specific percentage. The percentage depends upon many factors such as compression performance, frame details, frame size and distance between frames.

The structure diagram of the Percentage Difference approach can be shown in Figure (4.16).



Images resize
We resize the extracted frames according to frame size of the original video.

Color Histogram
Obtain the function double and color histogram for every frame

Frame Difference
Calculate the frame difference between each consequence of frames. As shown in 2

Threshold Calculation
Summing the total absolute difference according to some threshold.

Average Differencing
Calculate the average difference of all consequences of frames.

Remove Frames
Remove the frames where the frame differences lower than some percentage.

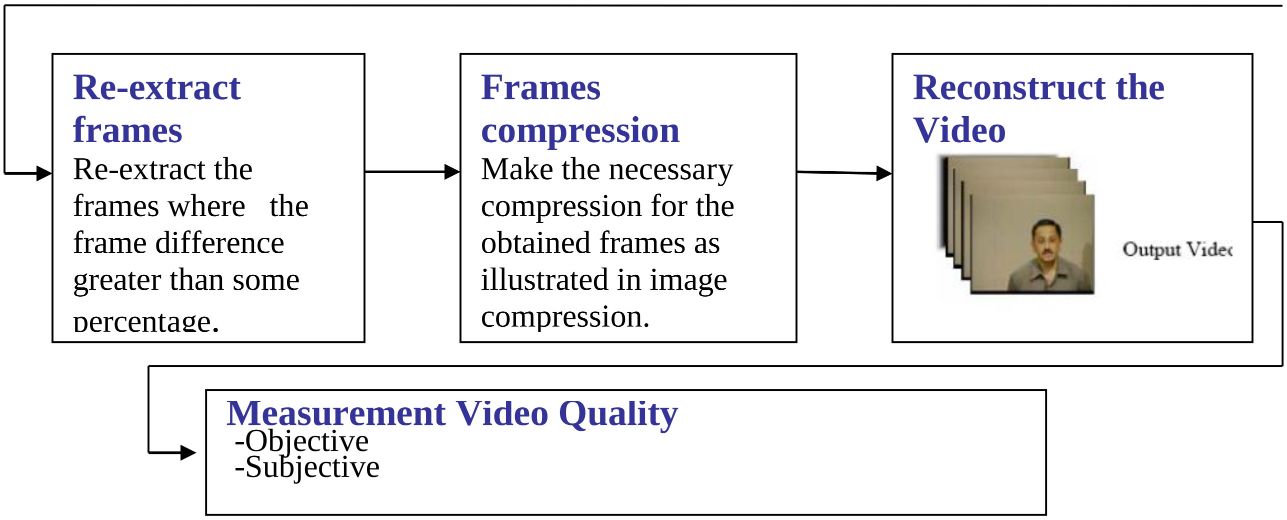


Figure (4.16) Structure Diagram of the third approach (Percentage Difference).

Chapter Five
Implementation and Results

This chapter will focus on the obtained results and the performance of the implemented system. The first part of this chapter discusses the results of image compression using discrete wavelet transform in which different wavelet families are selected for different types of images. While the second part is concentrated on the results and their performance obtained from different types of videos.

The implemented approaches were tested using mat lab programming codes (Mat lab 2010a) on windows platform (Windows 2007). All experiments were performed on Intel i5 Toshiba 2.68 MHZ.

5.1 Image Compression

Image Compression is possible because most real-world data are very statistically redundant. Compression is important because it facilitates optimum consumption of expensive resources, such as disk space or connection bandwidth. Some compression schemes are reversible so that the original data can be reconstructed (lossless data compression) while other scheme accept some loss of data in order to achieve higher compression (lossy data compression).

The performance compression scheme and effect of compression quality of the decompressed data can be studied using certain measures of performance like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) and can be used to measure the distortion caused by lossy compression of images and videos.

5.1.1 Wavelet Families

Several types of wavelet families are implemented and tested for different types of images. The most common ones are Haar, Daubechies, Symlets, Coiflets, Biorthogonal and Reverse Biorthogonal.

However wavelets transforms is applied to entire images, rather than sub images, so it produces no blocking artifacts. This is a major advantage of wavelet compression over other transform compression methods.

5.1.2.Two Dimensional Wavelet Masks

There are many different types of wavelet masks, this work concentrated on the masks that shown in table (4.1) given in chapter 4.

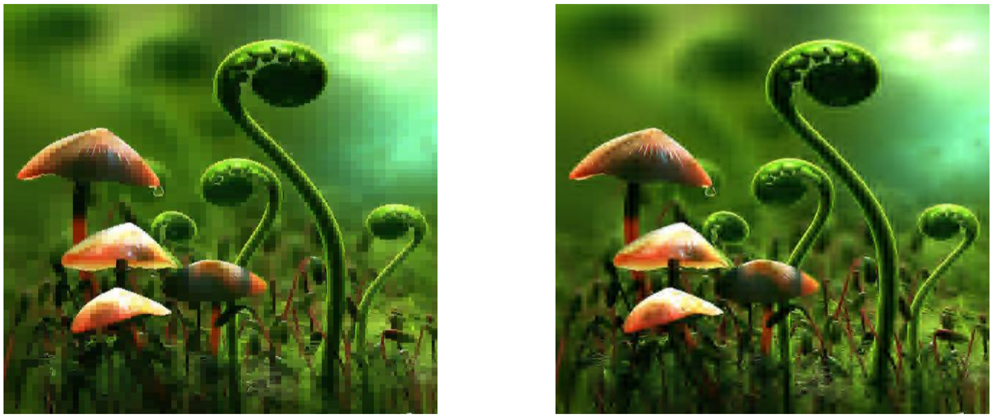
5.1.3 Performance Measurements of Image Compression

Two classes of methods are conducted to measure image quality, **Subjective tests**, in which human subjects are asked to assess or rank the viewed material, and **Objective tests**, which are computational models that measure the quality by comparing the original and the compression image.

5.1.4 Image compression results and analysis

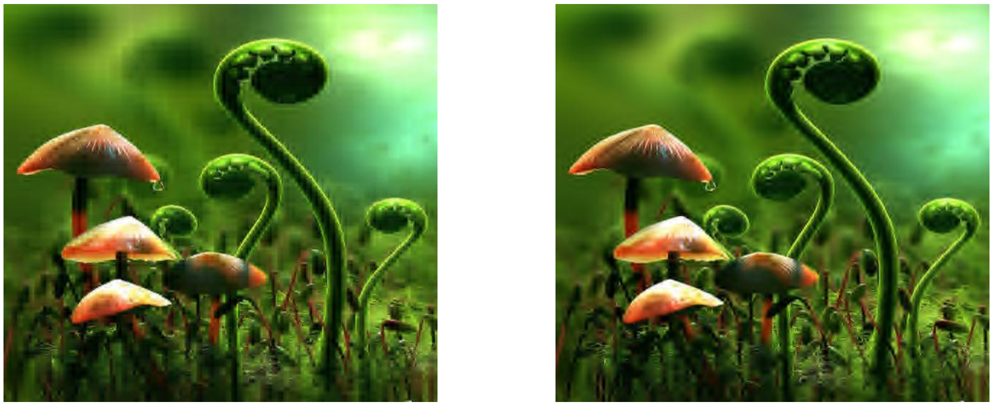
The natural plant image of specification is used here as test image as shown in figure (5.1) (Dimensions 600*794 Size 93.9KB). Then different types of wavelets are applied to this image, these wavelets are Haar wavelet (HW), Duabaechies(db), Symlets (sym2), Coiflets(coif2), Biorthogonal (2.2), and Reverse biorthognal (rbio2.2). Results are observed in terms of Compression Ratio (CR), Mean Square Error (MSE) and Peak Signal to Ratio (PSNR). Table (5.1) summarizes the calculated MSE & PSNR for different wavelet families. The best MSE & PSNR gives by Coiflet wavelet transform 133 & 26.92 as shown in figure (5.2) (d). While Ribo2.2 gives maximum MSE (245) & less PSNR (24.26) as shown in figure (5.2) (f). Among the compressed images shown in figure (5.2), one can notice easily that coiflets2 is the highest quality image and it is very close to the original plant. Also figure (5.3) & figure (5.4) shows results of MSE and PSNR for plant test image respectively.

Figure (5.1) Original Plant image

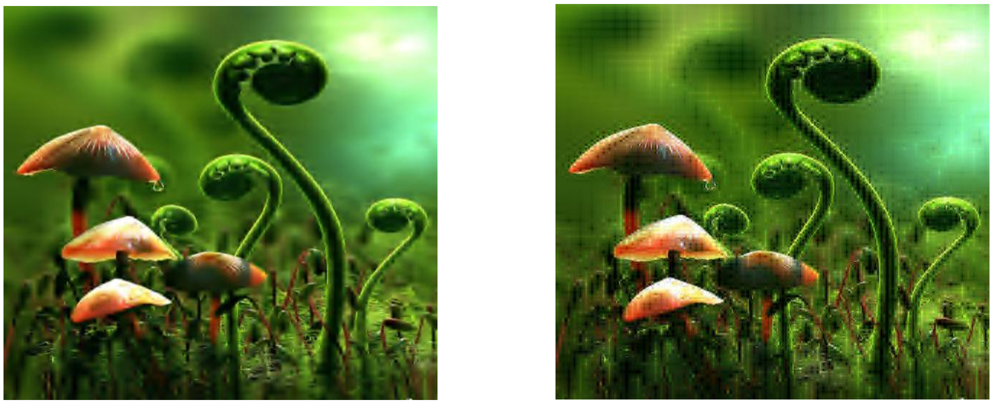


(a) Haar (98.78%)

(b) Daubechies 2 (98.68%)



(c) Symlets (98.68%) (d) Coiflets2 (98.31%)



(e) Biorthognal 2.2 (98.58%) (f) Reverse Biorthogonal 2.2 (98.58%)

Figure (5.2) the compression image results using different wavelet families.

Table (5.1) Experimental Results for test image Plant.

Wavelet Transform	Compression Ratio	Mean Square Error	Peak Signal to Ratio
	(%)	(MSE)	**(**PSNR**)**
Haar	98.78	213.0	24.88
Daubechies	98.68	164.98	25.99
Symlets	98.68	164.98	25.99
Coiflet	98.32	133.07	26.92
Bioorthognal	98.58	147.74	26.469
Reverse biorthognal	98.58	245.0	24.265

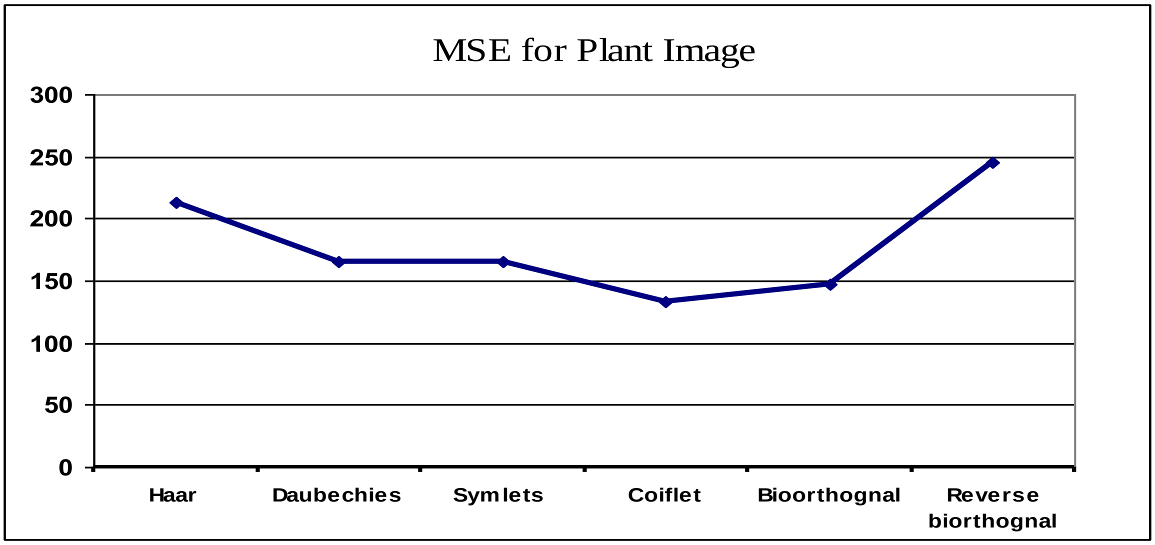


Figure (5.3): MSE for Plant Image.

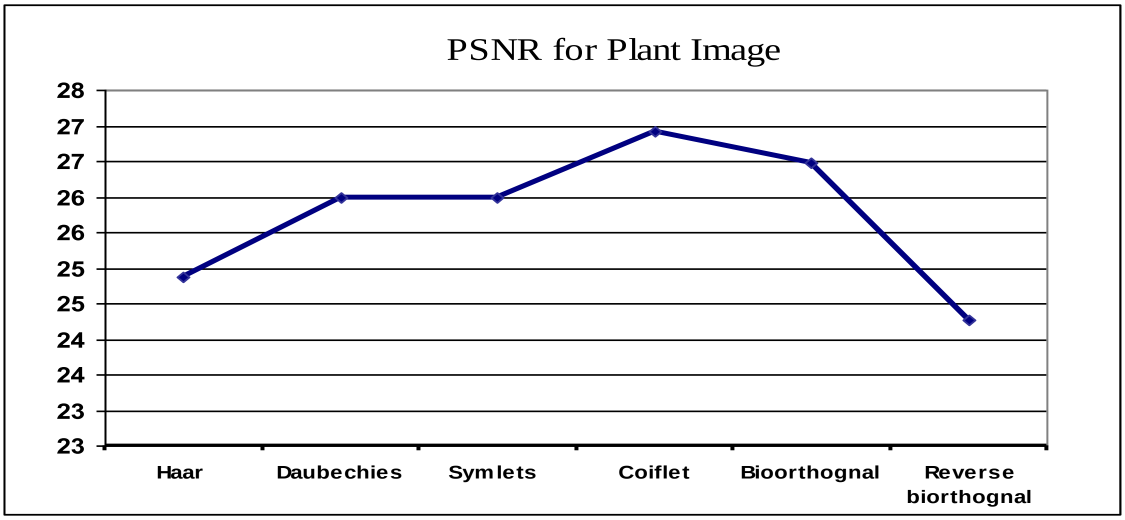


Figure (5.4): PSNR for Plant Image.

Also, in this work, different types of images are used with different details such as News, Sport, Music, Cartoon, Series, and Panorama. Figure (5.5) shows these types of images, while table (5.2) describes detail descriptions (file size, format, Width, Height, Bit depth, and Color type) for each image.

(a) NEWS (b) SPORT (c) MUSIC

(d) CARTOON (e) SERIES IMAGES (f) PANORAMA

Figure (5.5) Different types of images

Table (5.2) Images information's

	News	Sport	Music	Cartoon	Series	Panorama
File size(KB)	45.2	81.3	78.1	66.1	67.7	45.2
Format	jpg	jpg	jpg	jpg	jpg	jpg
Width	500	548	531	379	400	500
Height	201	389	411	285	300	201
Bit Depth	24	24	24	24	24	24
Color Type	true color	true color	true color	true color	true color	true color

5.1.4.1 Objective Results of Image Compression

To compare the objective results measures, PSNR and MSE are calculated for all tested images using four wavelet transform levels of compression in which figures 5.6 shows PSNR measures for different types of images. While figures 5.7 shows MSE measures for different types of images. Also Table (5.3) gives summary of PSNR & MSE measures for different wavelet families, levels of decomposition and types of images.

The obtained results of PSNR indicate that there is a small difference indicated for Haar, Db2, Ribo2.2 and Ribo2.4 for all types of images and all types of wavelet levels. Also the obtained results show that PSNR decreases with the increasing of compression ratio, also MSE increases with the increasing of compression ratio. On the average the maximum value of PSNR gives by News image while the lowest value gives by Panorama's image. The obtained results of MSE indicate that it is normal that MSE increase with the increasing of compression ratio or wavelet levels. MSE measures indicate that Ribo2.2 gives maximum MSE for all wavelet levels and for each image type. The next MSE after Ribo2.2 is Haar then Ribo2.4 then the rest of wavelet masks for all wavelet levels and for each image type as shown in Table (5.3) and figure (5.6) and figure (5.7).Also the obtained results shows that the News results gives the minimum MSE. While the maximum MSE gives by Series image. Also as the level of decomposition increases (1, 2, 3 up to 4) the PSNR decreases and vice versa with respect to MSE. The best wavelet family results give by coif4 for the most image types (News, Sport, Music, Cartoon, and Series).

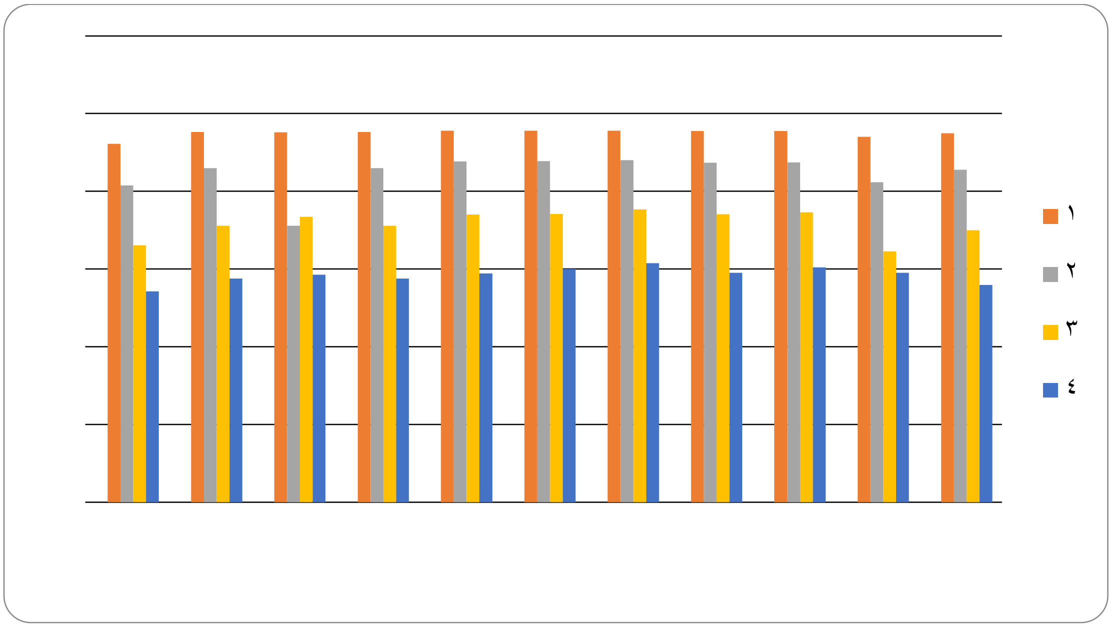
Table (5.3) MSE & PSNR results for different types of images.

wavelet family		NEWS				SPORT				MUSIC			
		1	2	3	4	1	2	3	4	1	2	3	4
Haar	MSE	1.6	5.5	32.5	127.2	7.9	72.3	191.2	361.7	2.0	16.4	77.8	250.2
	PSNR	46.1	40.8	33.0	27.1	39.2	29.6	25.4	22.6	45.1	36.0	29.3	24.2
db2	MSE	1.1	3.3	18.2	86.7	7.0	64.7	166.5	311.9	1.3	7.7	40.7	151.0
	PSNR	47.6	43.0	35.6	28.8	39.7	30.1	26.0	23.2	47.0	39.3	32.1	26.4
db4	MSE	1.1	2.9	13.9	77.3	6.7	60.3	152.8	283.6	1.3	6.0	30.7	115.1
	PSNR	47.6	43.6	36.7	29.3	39.9	30.3	26.3	23.6	47.1	40.4	33.3	27.6
sym2	MSE	1.1	3.3	18.2	86.7	7.0	64.7	166.5	311.9	1.3	7.7	40.7	151.0
	PSNR	47.6	43.0	35.6	28.8	39.7	30.1	26.0	23.2	47.0	39.3	32.1	26.4
sym4	MSE	1.1	2.7	13.1	74.7	6.6	59.1	149.7	276.7	1.2	5.8	29.5	110.0
	PSNR	47.8	43.8	37.0	29.4	40.0	30.4	26.4	23.7	47.3	40.5	33.5	27.7
coif2	MSE	1.1	2.7	12.8	65.0	6.6	57.8	144.8	263.4	1.2	5.5	26.8	95.7
	PSNR	47.8	43.9	37.1	30.0	39.9	30.5	26.6	24.0	47.3	40.7	33.9	28.4
coif4	MSE	1.1	2.6	11.3	55.1	6.7	54.8	133.3	232.6	1.2	4.7	21.3	73.2
	PSNR	47.8	44.0	37.6	30.8	39.9	30.8	26.9	24.5	47.4	41.5	34.9	29.5
bior2.2	MSE	1.1	2.8	12.8	73.6	7.0	61.3	154.0	285.8	1.3	6.4	31.2	116.1
	PSNR	47.8	43.7	37.1	29.5	39.7	30.3	26.3	23.6	47.2	40.1	33.2	27.5
bior2.4	MSE	1.1	2.8	12.2	62.4	6.9	59.3	148.1	268.7	1.2	6.1	27.9	95.9
	PSNR	47.8	43.7	37.3	30.2	39.8	30.4	26.5	23.9	47.2	40.3	33.7	28.3
rbio2.2	MSE	1.3	5.0	39.0	176.9	9.8	81.3	214.1	411.7	1.5	12.8	31.2	312.0
	PSNR	47.0	41.1	32.3	25.7	38.3	29.1	24.9	22.0	46.3	37.1	29.0	23.2
rbio2.4	MSE	1.2	3.4	20.8	105.2	8.4	68.3	174.8	323.3	1.3	7.4	40.5	155.4
	PSNR	47.5	42.8	35.0	27.9	39.9	29.8	25.7	23.1	47.0	39.5	32.1	26.2

Table (5.3) MSE & PSNR results for different types of images (Continued).

wavelet family		CARTOON				SERIES				PANORAMA			
		1	2	3	4	1	2	3	4	1	2	3	4
haar	MSE	2.3	35.9	184.9	500.2	8.1	89.5	309.2	676.5	28.6	33.4	62.2	152.7
	PSNR	44.6	32.6	25.5	21.2	39.1	28.6	23.3	19.9	33.6	32.9	30.2	26.3
db2	MSE	1.8	27.0	145.0	378.2	6.6	74.9	258.0	574.6	27.7	32.1	56.3	121.2
	PSNR	45.6	33.9	26.6	22.4	40.0	29.4	24.0	20.6	33.7	33.1	30.7	27.3
db4	MSE	1.7	23.8	125.4	327.3	6.2	69.1	241.2	517.2	27.7	32.3	52.8	102.5
	PSNR	45.7	34.4	27.2	23.0	40.2	29.8	24.3	21.0	33.7	33.1	30.9	28.1
sym2	MSE	1.8	33.9	145.0	378.2	6.6	74.9	258.0	574.6	27.7	32.1	56.3	121.2
	PSNR	45.6	33.9	26.6	22.4	40.0	29.4	24.0	20.6	33.7	33.1	30.7	27.3
sym4	MSE	1.7	23.4	119.3	306.2	6.1	65.8	228.3	491.5	27.7	83.6	94.6	98.7
	PSNR	45.8	34.5	27.4	23.3	40.3	30.0	24.6	21.2	33.7	33.1	31.0	28.2
coif2	MSE	1.7	83.1	106.8	284.0	6.0	63.0	216.8	462.1	27.7	31.7	49.8	92.0
	PSNR	45.9	34.9	27.8	23.6	40.4	30.2	24.8	21.5	33.7	33.2	31.9	28.5
coif4	MSE	1.6	17.6	83.2	210.8	5.7	56.8	187.8	392.8	27.6	31.0	45.2	75.2
	PSNR	46.1	35.7	29.0	24.9	40.6	30.6	25.4	22.2	33.8	33.2	31.6	29.4
bior2.2	MSE	1.7	22.9	121.6	327.4	6.1	67.0	234.2	511.4	27.7	31.8	52.3	104.3
	PSNR	45.8	34.6	27.3	23.0	40.3	29.9	24.5	21.1	33.7	33.1	31.0	28.0
bior2.4	MSE	1.7	20.9	107.9	285.5	6.0	63.4	215.3	457.3	27.6	31.5	48.7	89.3
	PSNR	45.8	35.0	27.8	23.6	40.4	30.1	24.8	21.6	33.8	33.2	31.3	28.7
rbio2.2	MSE	2.2	41.2	209.8	533.5	9.4	105.2	369.0	811.0	27.8	34.4	72.2	169.0
	PSNR	44.8	32.0	24.9	20.9	38.4	27.9	22.5	19.1	33.7	32.8	29.6	25.9
rbio2.4	MSE	1.9	29.5	150.8	369.7	7.8	81.4	274.7	577.1	27.7	33.0	57.2	110.0
	PSNR	45.4	33.5	26.4	22.5	39.2	29.1	23.8	20.6	33.7	33.0	30.6	27.8

The results of the content class (**News**) shown in tables (5.3) and figures (5.6) (a) and (5.7) (a) indicates that as the levels of decomposition increases, values of MSE increases and values of PSNR decreases for all types of wavelet families. The obtained results of MSE measures indicate that coif 4 gives minimum value of MSE (55.1) with 4 level of decomposition, while Rbio gives maximum value of MSE (176.9) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (30.2), while minimum value of PSNR was given by Rbio2.2 (25.7).

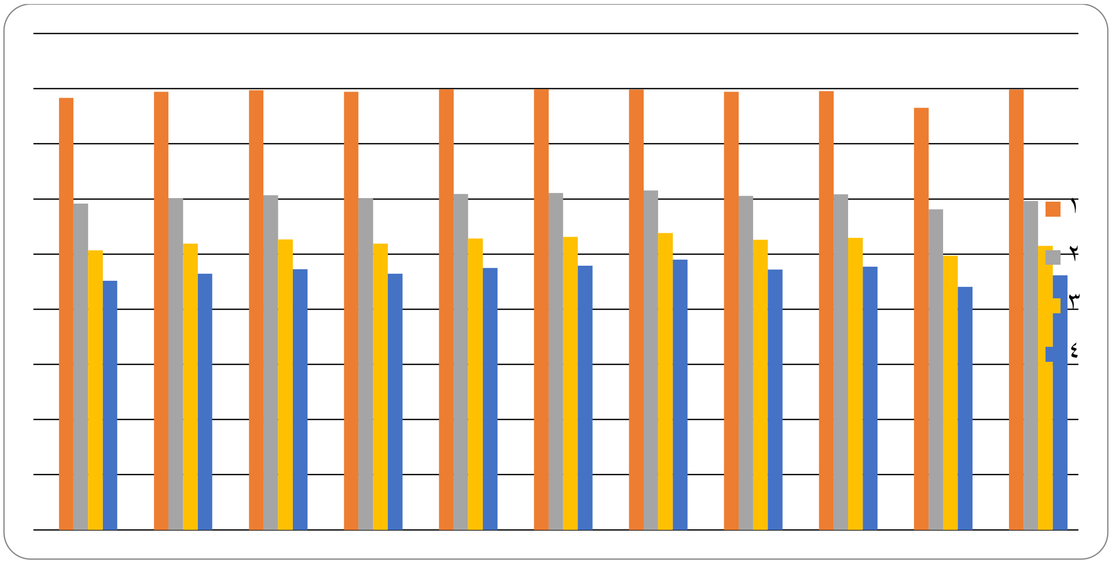


(a) News image.

Figures (5.6) PSNR measures

The results of the content class (**Sport**) shown in tables (5.3) and figures (5.6) (b) and (5.7) (b) indicates that as the levels of decomposition increases, values of MSE increases and values of PSNR decreases for all types of wavelet families.

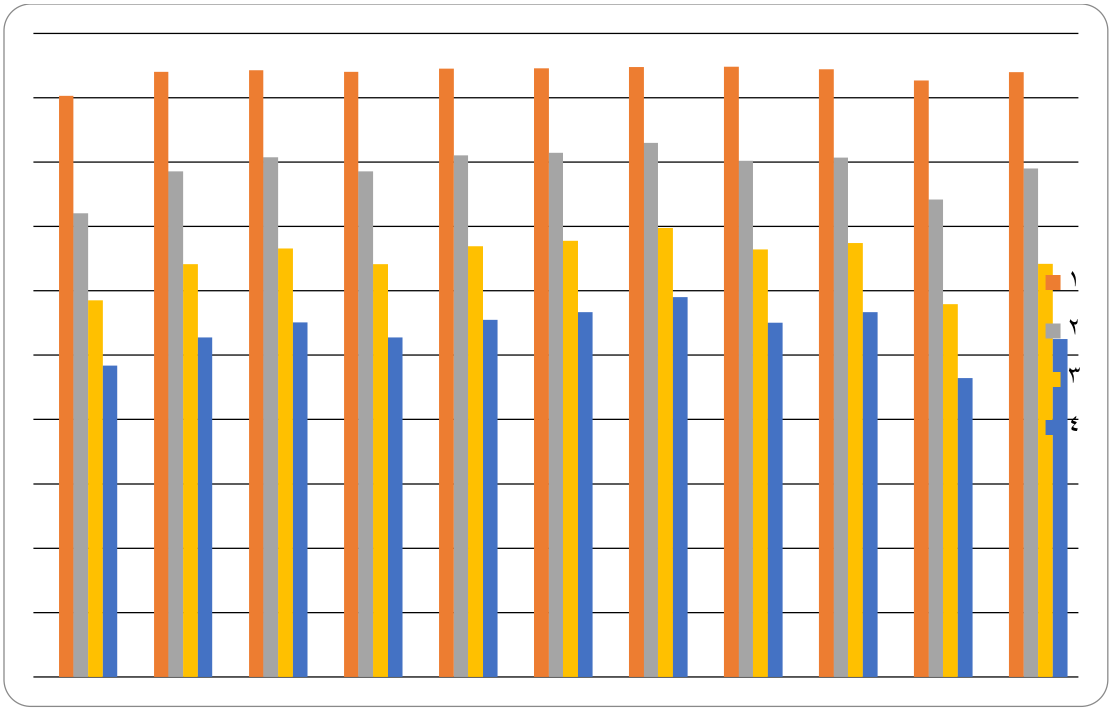
The obtained results of MSE measures indicate that coif 4 gives minimum value of MSE (55.1) with 4 level of decomposition, while Rbio gives maximum value of MSE (176.9) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (30.2), while minimum value of PSNR was given by Rbio2.2 (25.7).



(b) Sport image.

Figures (5.6) PSNR measures (Continued)

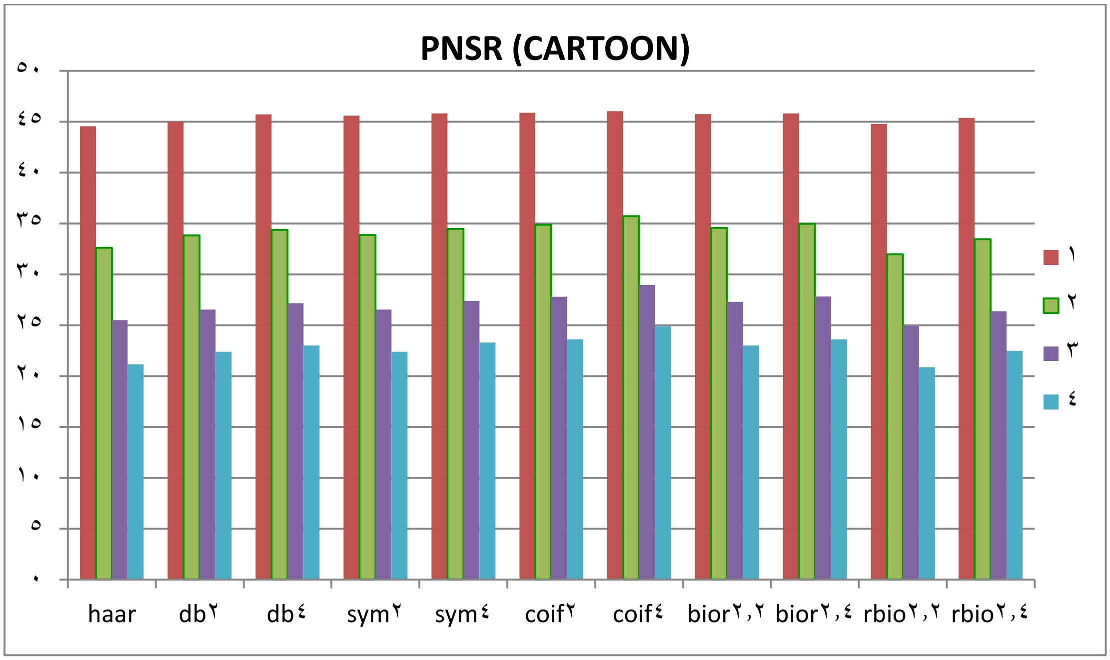
The results of the content class (**MUSIC**) shown in tables (5.3) and figures (5.6) (c) and (5.7) (c) indicates that values of MSE measures shows that coif 4 gives minimum value of MSE (73.2) with 4 level of decomposition, while Rbio2.2 gives maximum value of MSE (312) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (29.5), while minimum value of PSNR was given by Rbio2.2 (23.2).



(c) Music image.

Figures (5.6) PSNR measures (Continued)

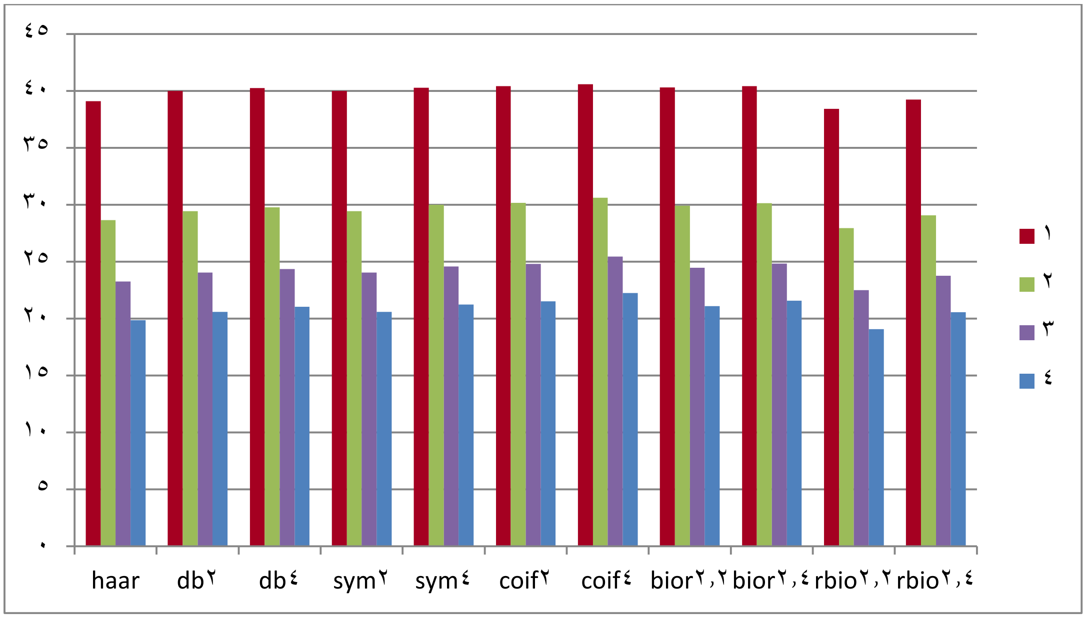
The results of the content class (**CARTOON**) shown in tables (5.3) and figures (5.6) (d) and (5.7) (d) indicates that values of MSE measures shows that coif 4 gives minimum value of MSE (210.8) with 4 level of decomposition, while Rbio2.2 gives maximum value of MSE (533.5) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (24.9), while minimum value of PSNR was given by Rbio2.2 (20.9).



(d) Cartoon image.

Figures (5.6) PSNR measures (Continued)

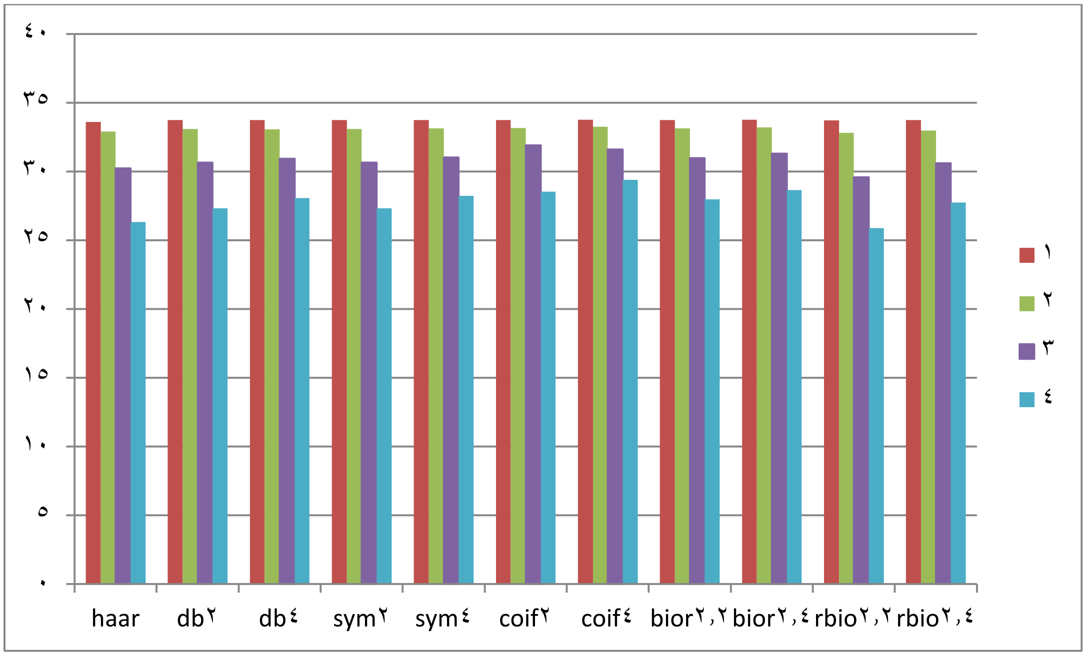
The results of the content class (**SERIES**) shown in tables (5.3) and figures (5.6) (e) and (5.7) (e) indicates that values of MSE measures shows that coif 4 gives minimum value of MSE (392.8) with 4 level of decomposition, while Rbio2.2 gives maximum value of MSE (811.0) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (22.2), while minimum value of PSNR was given by Rbio2.2 (19.1).



(e) Series image.

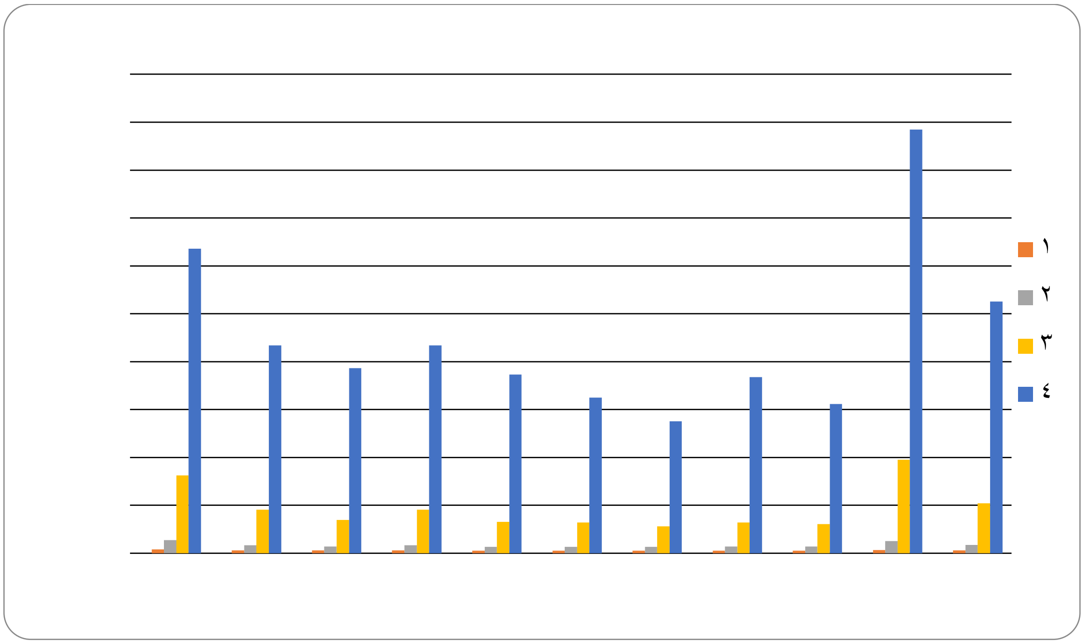
Figures (5.6) PSNR measures (Continued)

The results of the content class (**PAMORAMA**) shown in tables (5.3) and figures (5.6) (f) and (5.7) (f) indicates that values of MSE measures shows that coif 4 gives minimum value of MSE (75.2) with 4 level of decomposition, while Rbio2.2 gives maximum value of MSE (169.0) with 4 level of decomposition. With respect to PSNR measures, Coif 4 gives maximum values of PSNR (29.4), while minimum value of PSNR was given by Rbio2.2 (25.9).



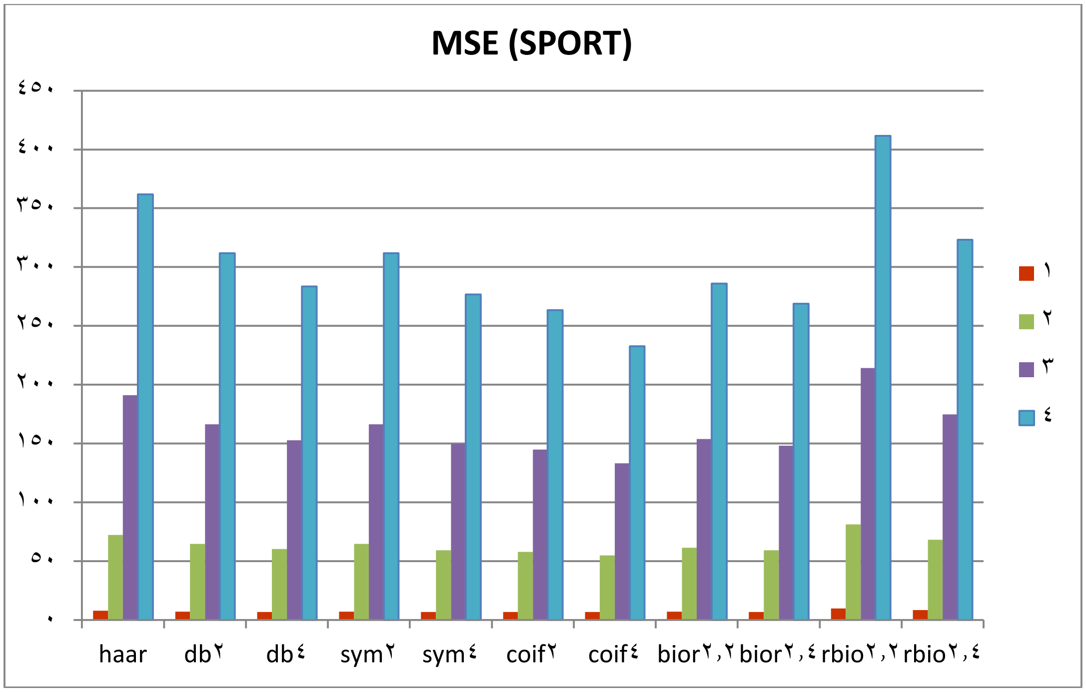
(f) Panorama image

Figures (5.6) PSNR measures (Continued)

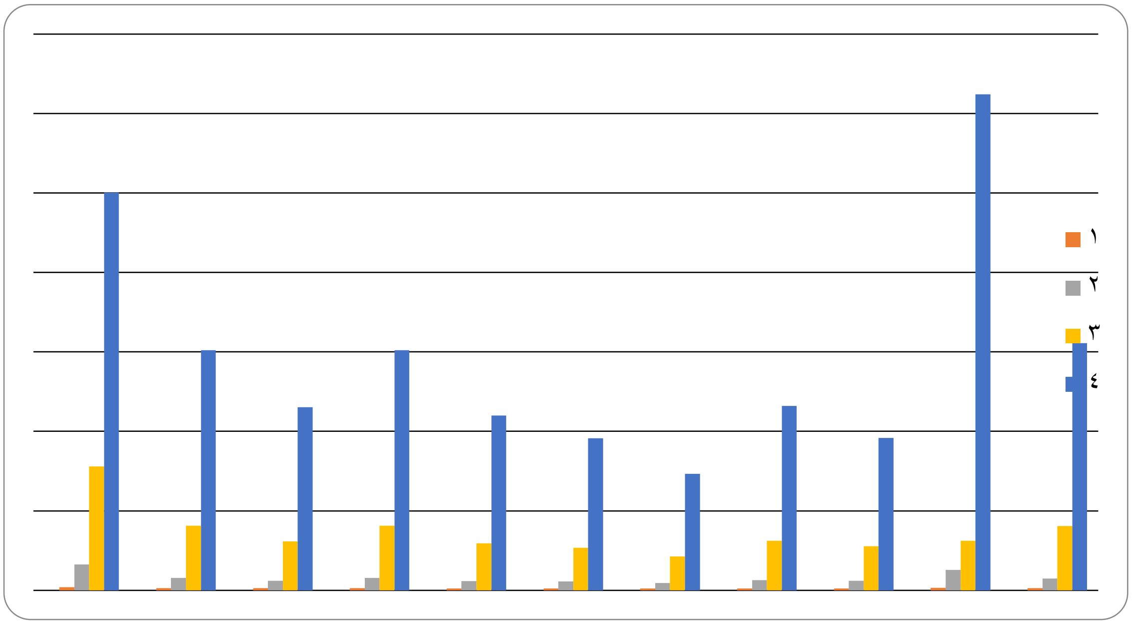


(a) News image.

Figures (5.7) MSE Measures

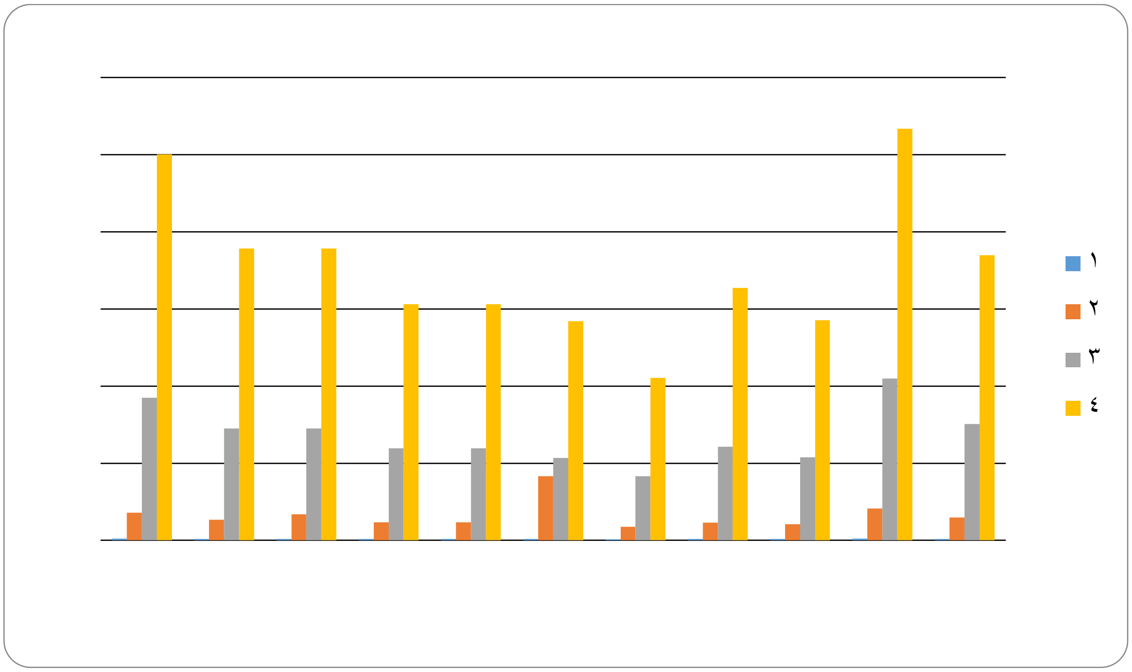


(b) Sport image.

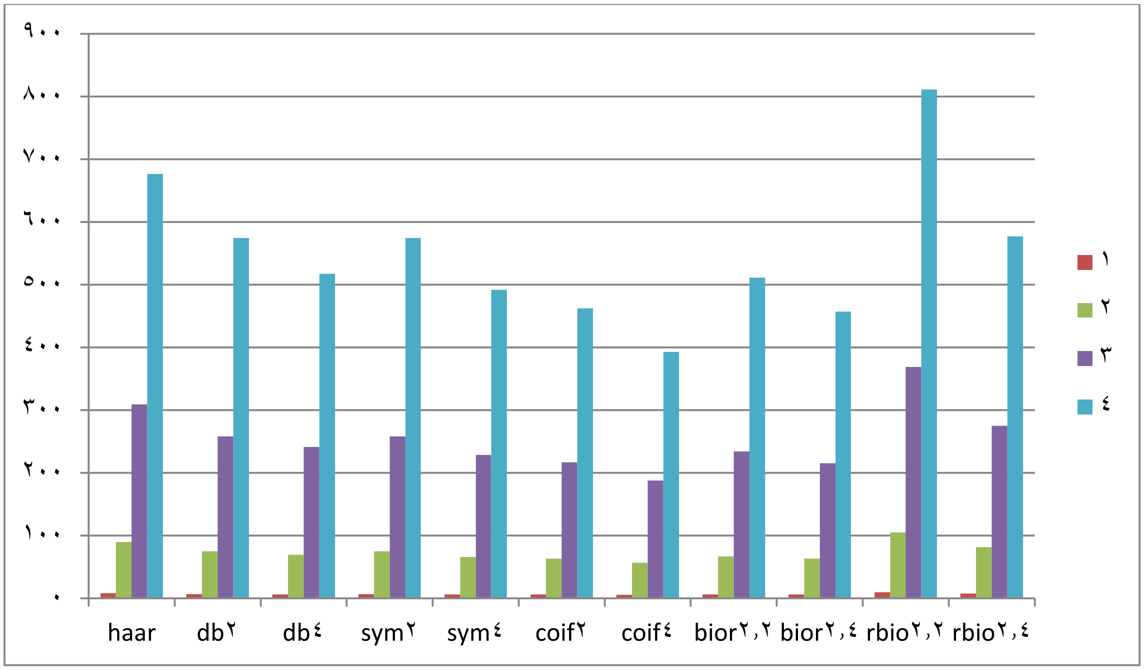


(c) Music image.

Figures (5.7) MSE Measures (Continued)

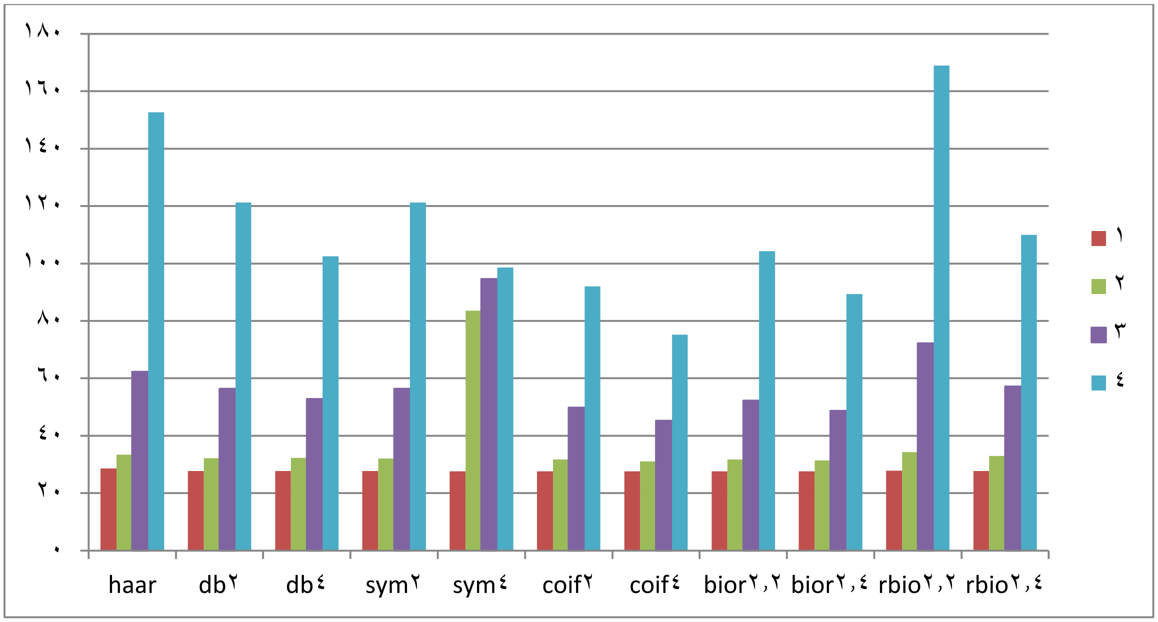


(d) Cartoon image.



(d) Cartoon image.

Figures (5.7) MSE Measures (Continued)



(e) Series image.

Figures (5.7) MSE Measures (Continued)

5.1.4.2 Subjective Results of Image Compression

When talking about subjective test, viewer's focus on the difference between the reconstructed image and the original image. Figure (5.8) shows some types of images original, compressed and reconstructed. In addition fourth level of wavelet compression are applied to these images.

A **Questionnaire** check list is implemented on different types of images, and then the results are collected for different viewers. Each viewer compares the compressed image, reconstructed image and the original image and then a point that score at a check list. The proposed **Questionnaire** for evaluating image compressions is given in chapter 4. In this work the evaluation tests were collected from different sets of the selected six images classes.

The chosen group of test persons ranged from different ages (between 25 to 45), gender, education and some of them have experience with image processing. People evaluated image quality after each sequence using five grade scales (1-bad, 2-Poor, 3-Fair, 4-Good, 5-Excellent) in a prepared form. To compare the subjective quality, we can say that the obtained results denoted that there is no big difference mentioned between the original images and their corresponding compressed and reconstructed images as shown in table (5.4).

For the image (Series) human observers give an average mean opinion score (MOS) of 4.6 (out of 5). This score corresponds to an 'excellent or very good quality". For the images (Cartoon & Music) human give an average MOS of 4.2 and 4.1 respectively these scores corresponds to 'good quality". While the MOS score for the images (News & Sport) is 3.9 this score corresponds to "almost good quality". Finally human observers give Image (Panorama) score 3.4 which correspond to "fair quality". The obtained results can be summarized in figure (5.9). Therefore, image type, wavelet families, and levels of decomposition play an important rule in image compression.

Original Image Compressed image De-compressed image

A-NEWS IMAGES (Haar level 3)

B-CARTOON IMAGES (sym3 level 2)

C-SPORT IMAGES (coif2 level 4)

Figure (5.8) original, compressed and decompressed images.

Original Image Compressed image De-compressed image

D-MUSIC IMAGES (db3 level 3)

E-SERIES IMAGES (bior2.4 level 2)

F-PANORAMA (rbio3.3 Level 3)

Figure (5.8) Original, Compressed and Decompressed images (continued).

Table (5.4) Subjective Image quality results

Image Quality Factors	Types of images					
	NEWS	CARTOON	SPORT	MUSIC	SERIES	PANORAMA
	Scores					
1- Degree of the overall clearance of image	4.2	4.4	4.0	4.2	4.6	3.3
2- Degree of recognizing overall of image	3.8	4.0	3.8	4.0	4.6	3.4
3-Degree of clearance details of image	3.7	4.2	3.8	3.8	4.6	3.4
4-Degree of clearance the colors in the image	4.2	4.4	4.0	4.2	4.5	3.5
5-Degree of non distortion in the image	3.8	4.4	4.0	4.4	4.6	3.4
6-Degree of the non noise in the image	3.7	4.0	3.8	4.2	4.6	3.3
7-Degree of the non noise effect in the overall image	3.9	4.2	4.0	4.0	4.4	3.4
8-Degree of the non distortion in the image	3.7	4.2	3.8	4.2	4.4	3.3

Averages	3.9	4.2	3.9	4.1	4.6	3.4

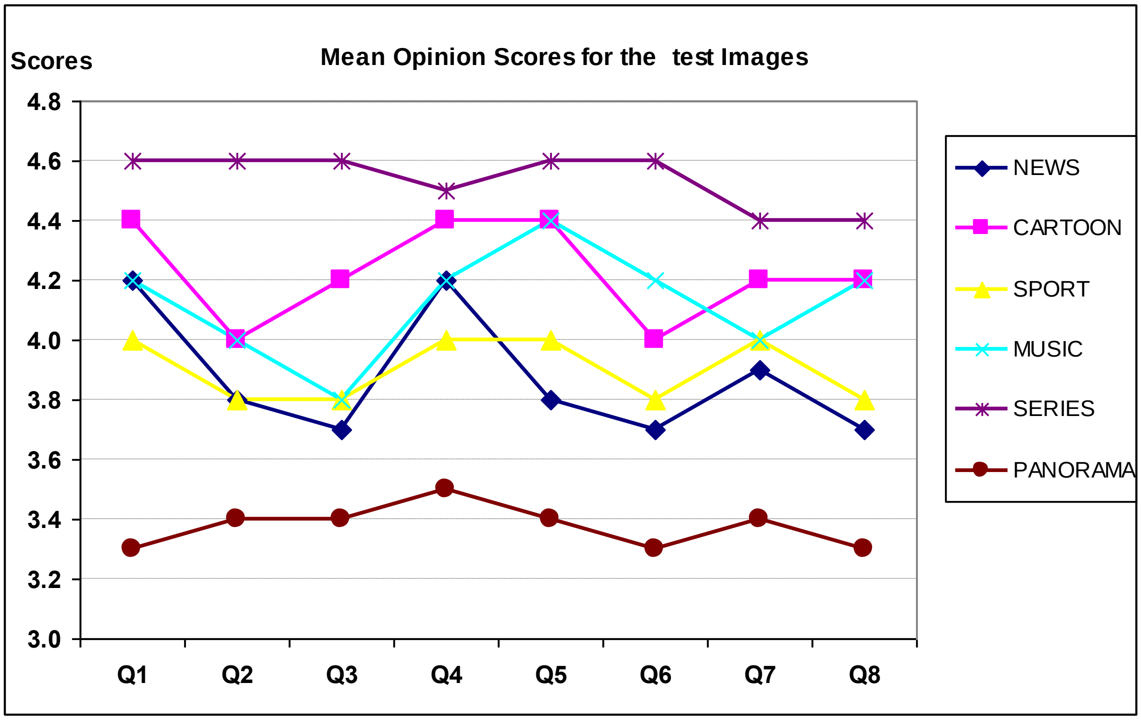


Figure (5.9) Mean Opinion Scores for the test images.

This work explains important features of wavelet transforms masks wer applied on image compression. The implemented algorithm based on an effective image compression technique that utilizes wavelet mask selection approach to reach high performance results. A comparative study of different types of DWT (with different masks) are implemented to select the perfect mask for a certain image depending on image size, compression ratio, wavelet level and reconstructed image quality. The objective measures (MSE and PSNR) are obtained that lead to optimal wavelet mask selection.

5.2 Video Compression

In this section, different approaches of video compression were implemented and tested for different video contents using frames difference approaches. Also subjective video quality assessment method is performed to evaluate the compressed videos.

These approaches were applied for removing the lowest frame difference; in which the very similar frames were removed. Three different methods are suggested to remove the similar frames in which frames difference approaches between each consecutive frame of the extracted frames are applied. Different types of videos were tested according to:

-Video types.

- Wavelet families' i.e. haar, db2, sym2, coif2, bior2.2, and rbio2.2.
- Levels of decomposition i.e. 1, 2, 3 and 4.

-Frames extraction i.e. frames per second 10frames/sec, 15frames/sec, 20frames/sec and 25frames/sec.

Subjective quality assessments were applied as a good method to measure video quality. A **Questionnaire** check list is implemented then the results are collected from different viewers. Each viewer compares the original video with the compressed one. Video quality experienced by users depends on spatial (edges, colors …) and more considerably on temporal (movement speed, direction …) features of video sequence, most of the well-known methods are based on spatial features.

Results of video quality assessment are given according to frames difference approaches.

5.2.1 Frames difference approaches

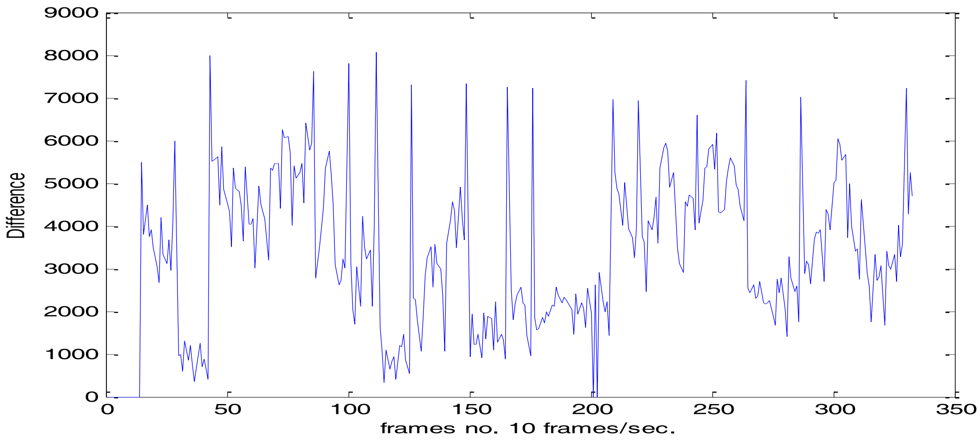
In this section three different approaches are implemented and tested as shown below:

A) **Zero Difference Approach**

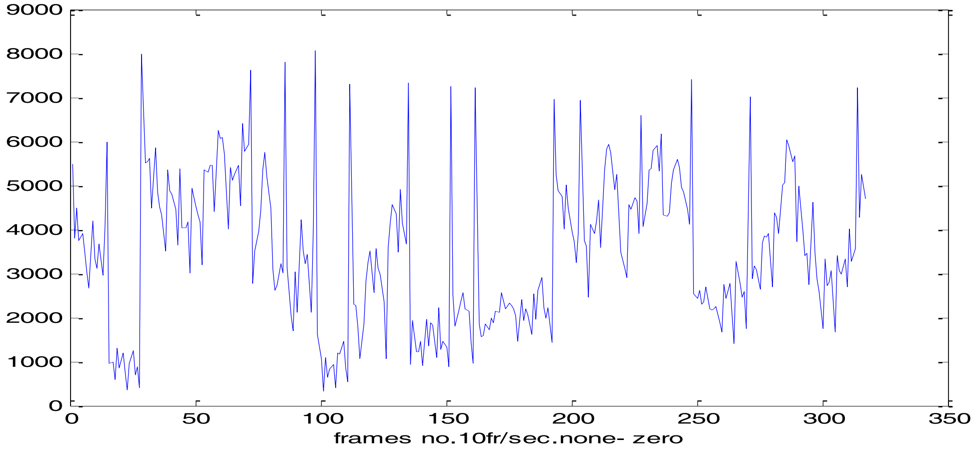
In this approach frames are removed were the distance between any two consecutive frames is zero. The obtained results are shown in figures (5.10), (5.11), (5.12) and (5.13).

Figures 5.10(a), 5.11(a), 5.12(a) and 5.13(a) represent the frames difference including zero's difference. While part b of the mentioned figures 5.10(b), 5.11(b), 5.12(b) and 5.13(b) represents frames difference excluding zero's difference frames. Also the removed frames in which the frames difference equal zero are shown in figures 5.10(c), 5.11(c), 5.12(c) and 5.13(c). As increasing extracting frames per second 10, 15, 20 and 25 frames per second, the frames difference equal zero not necessarily increased as can be shown from figures 5.12(c). Since the frames differences in general decreased as the frames per second increased but not exactly equal zero. After removing the frames in which the frames difference equal zero, then implementing video compression using different wavelet families and different level of decomposition after that measuring the obtained compressed video using subjective quality assessment. Using different wavelet families and level of decomposition to compress the obtained frames and a reconstructed video were made to obtain the compressed video. A subjective quality assessment tests were applied for the obtained compressed video.

As shown in figure (5.10) (a), the horizontal line represents frames number and the vertical line represents the value of frames differences between each consecutive frame after extracting video with 10 frames per second. One can notice that the range of values between each consecutive frame is zero and others above zero. For more details Figure (5.10) (b) represents the frames difference after excluding zero differencing also some of value differencing were very close to zero. While Figure (5.10) (c) indicates the frames to be removed (15 frames).

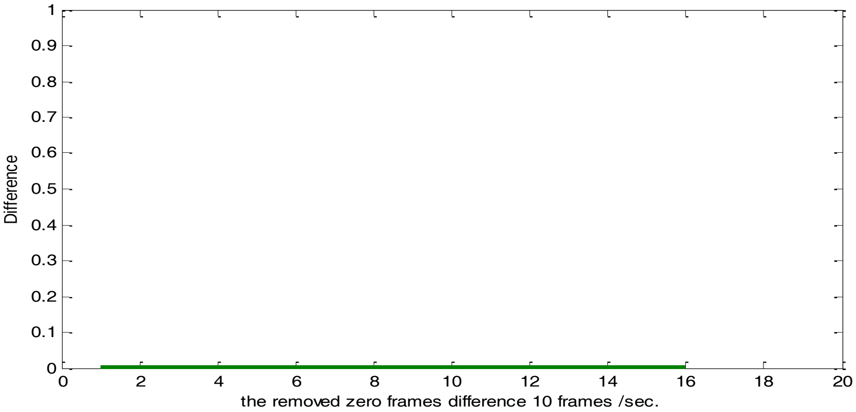


(a): Frames difference



(b): Frames difference after removing the zero difference frames.

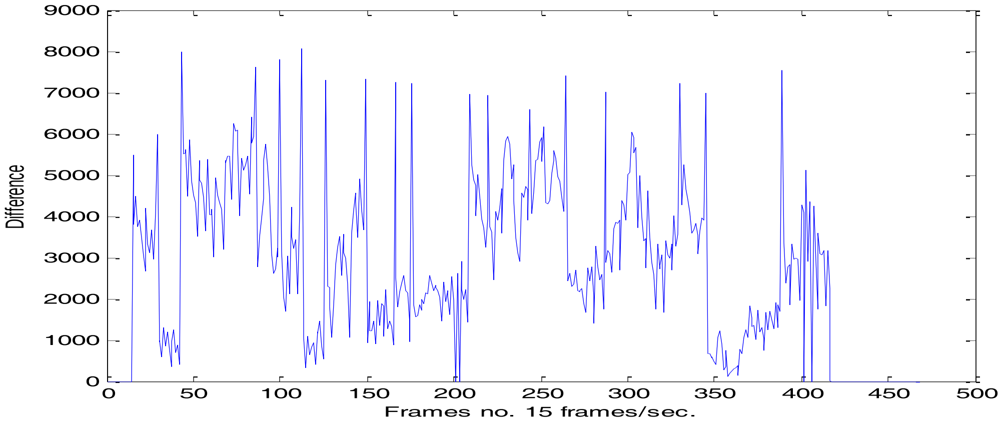
Figures (5.10) Zero difference approach (10frames /second)



(c): The removed zero difference frames.

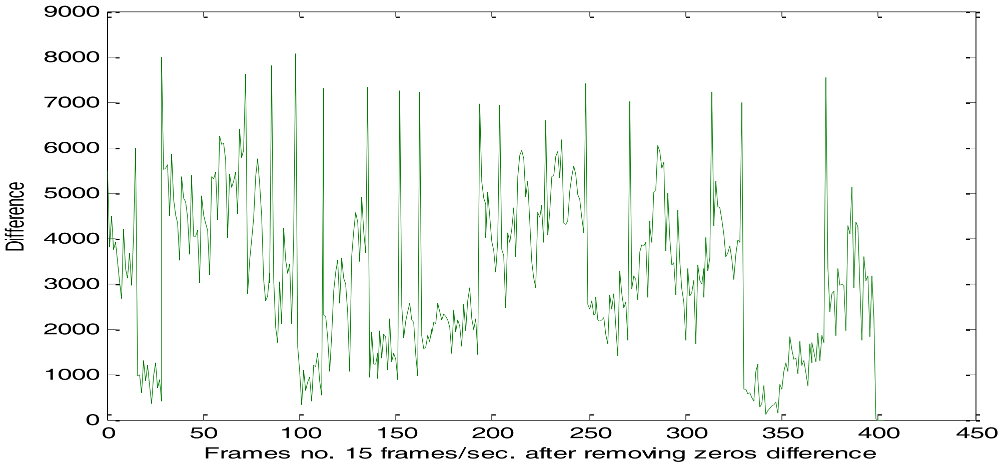
Figures (5.10) Zero difference approach (10frames /second) (cont.)

As shown in Figure (11) (a) with 15 frames per one can notice that the values frames differencing was decreased. Figure (5.11) (b) represents the frames difference after excluding zero differencing, also some of value differencing were very close to zero. While Figure (5.10) (c) indicates the frames to be removed (about 60 frames).

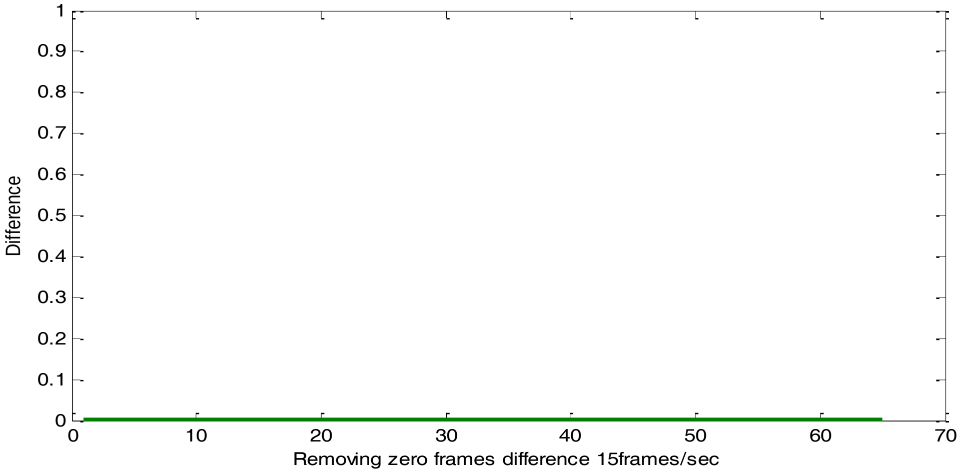


(a): Frames difference

Figures (5.11) Zero difference approach (15frames /second)



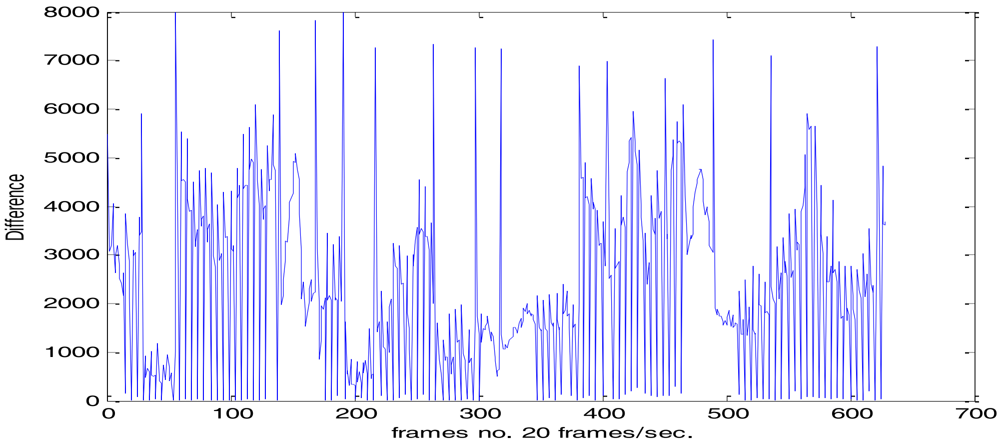
(b): Frames difference after removing the zero difference frames.



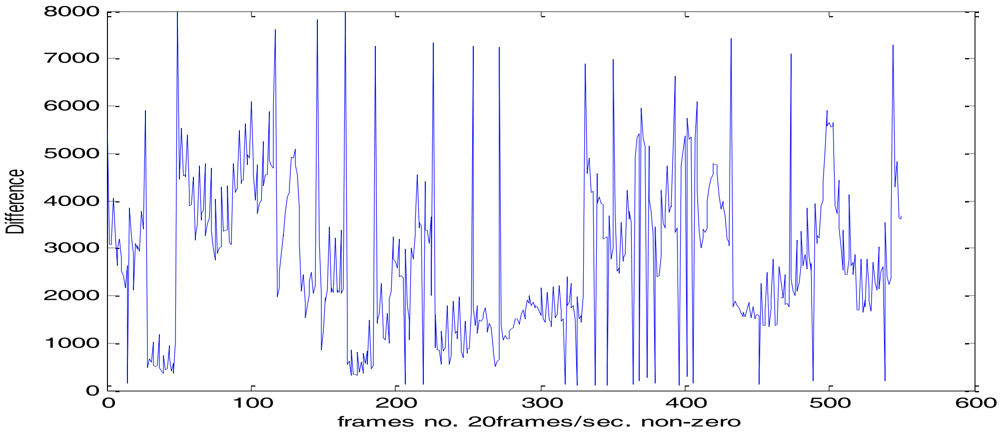
(c): The removed zero difference frames.

Figures (5.11) Zero difference approach (15frames /second) (cont.)

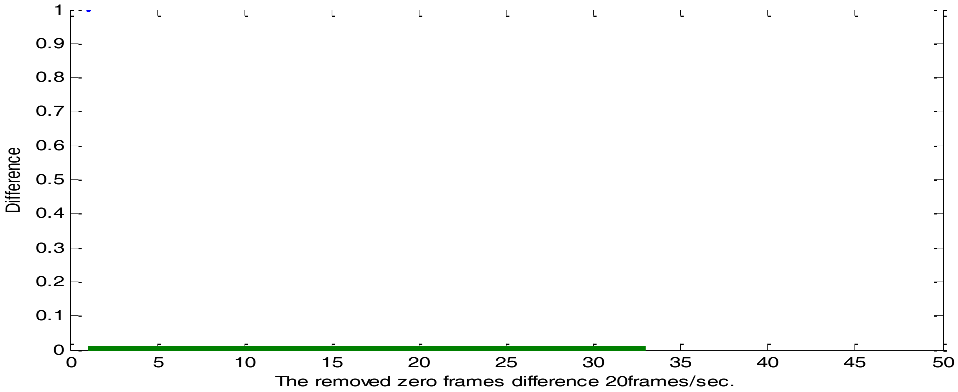
With 20 and 25 frames per second values of frames differencing were decreased comparing with 10 and 15 per second extraction, this can be shown in figures (5.12) (a), (5.13) (a) but not exactly equal zero, Figures (5.11) (c), (5.12) (c) shows the frames to be removed as given by the green line.



(a): Frames difference

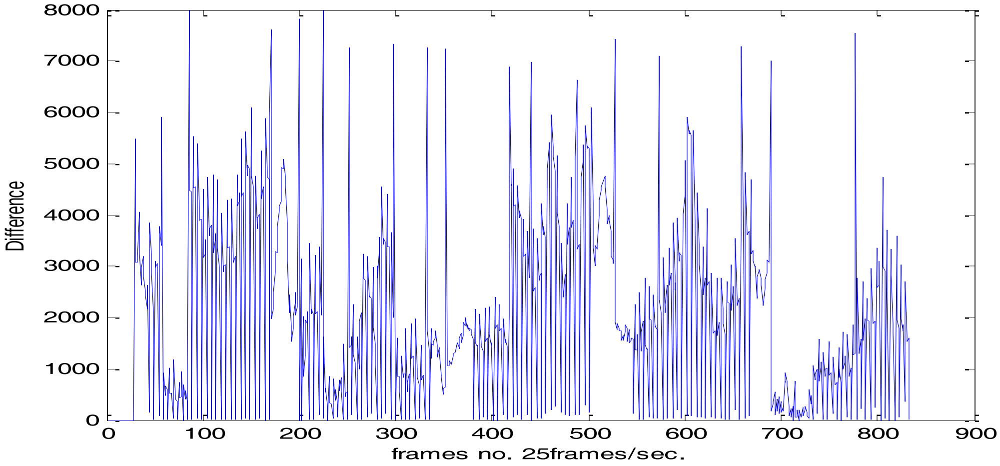


(b): Frames difference after removing the zero difference frames.

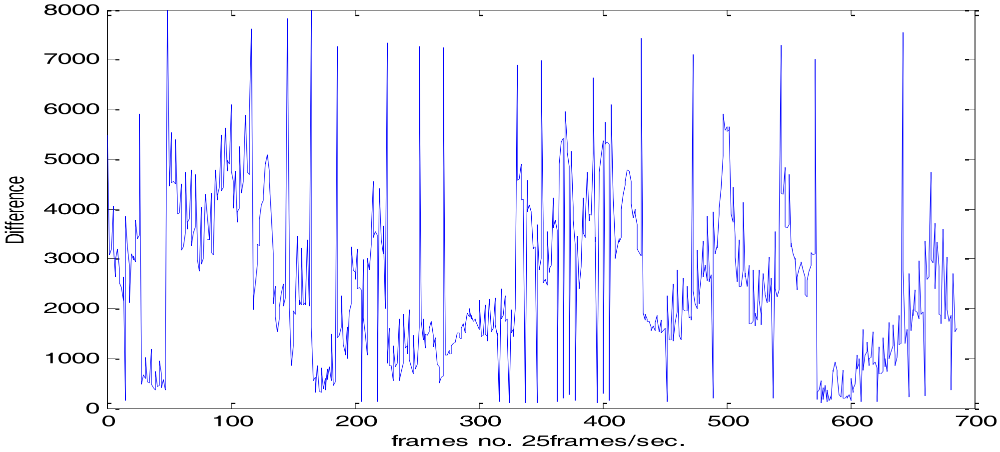


(c): The removed zero difference frames.

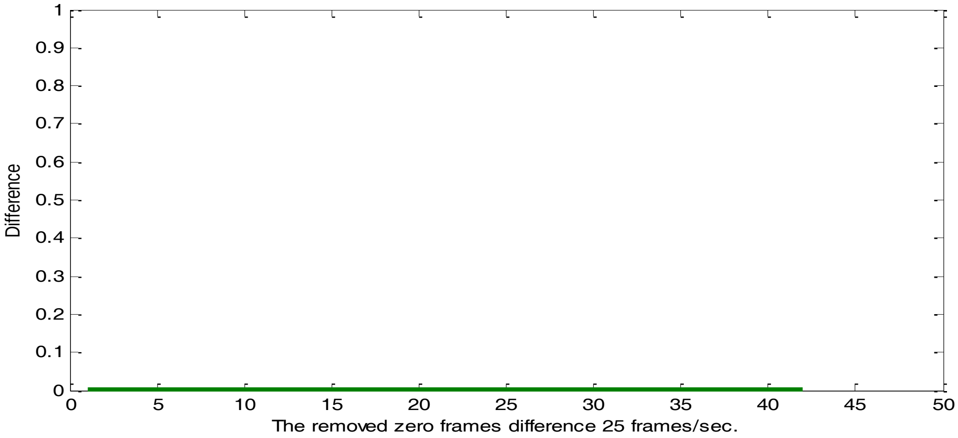
Figures (5.12) Zero difference approach (20 frames /second)



(a): Frames difference



(b): Frames difference after removing the zero difference frames.



(c): The removed zero difference frames.

Figures (5.13) Zero difference approach (25 frames /second)

Subjective Video quality assessment

To evaluate the dimension of the overall video quality, a special questionnaire was used for about 30-40 persons. Examples of question were asked after the presentation of the original and the compressed video, results of video quality assessment are given in table (5.5) and figure (5.14). From the obtained results we can notice that video quality increased with the increasing of frames per second. Evaluations of 25 frames give maximum video quality scores compared with (10, 15, and 20) frames per second.

Table (5.5) Video quality assessment (zero difference approaches).

Video Quality Factors	**10 frames**	**15 frames**	**20 frames**	**25 frames**
How would you rate video colors?	3.9	4.0	4.2	4.4
How would you rate video contrast?	4.0	4.2	4.3	4.5
How would you rate video borders?	4.0	4.1	4.4	4.6
How would you rate the movement continuity?	3.8	4.0	4.2	4.4
Did you notice any flicker in the sequence?	annoying	not annoying	not annoying	not annoying
Did you notice any smearing in the sequence?	annoying	little annoying	not annoying	not annoying

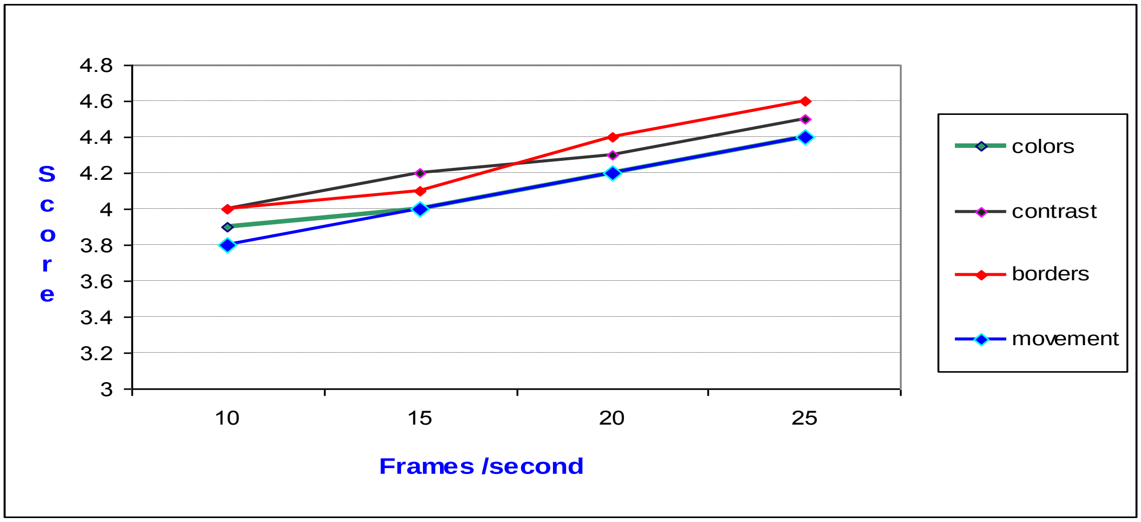


Figure (5.14) zero difference approach (scores)

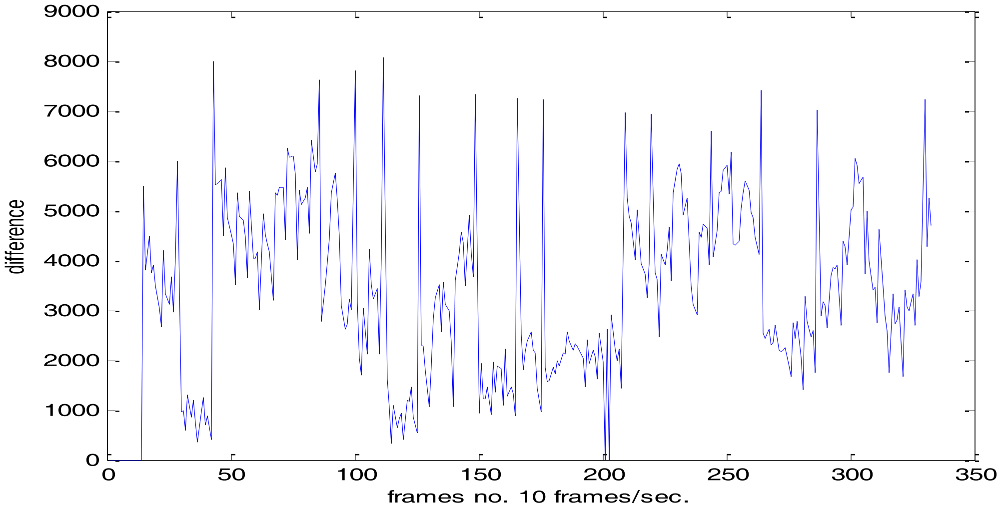
B) Mean Difference Approach

In this method, the mean value of the frames difference is calculated. Mathematically the mean (Average) is obtained by dividing the sum of the observed values of frames difference by the number of observations.

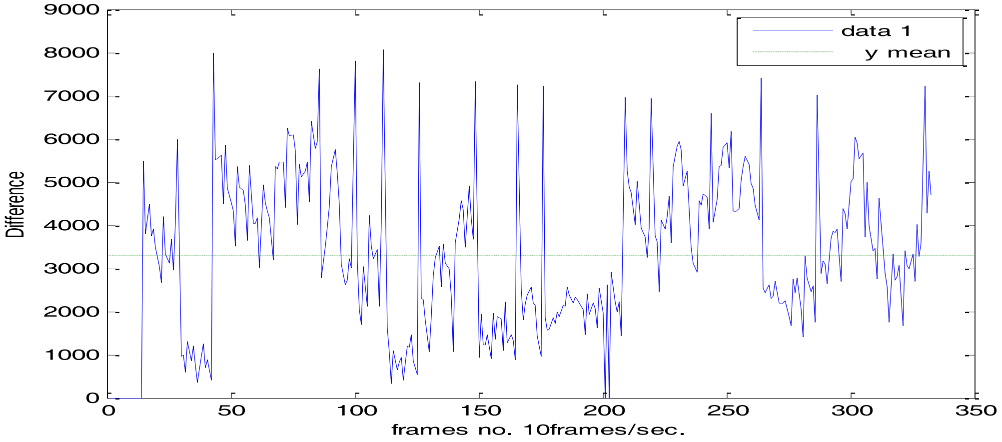
Then remove the frames where the frames difference between any consecutive frames is lower than the mean value of the frames difference.

Figures 5.15(a), 5.16(a), 5.17(a) and 5.18(a) represent the frames difference including mean's difference. Part b of the figures 5.15(b), 5.16(b), 5.17(b) and 5.18(b) represents frames difference including frames below the mean of the overall difference frames. In which below the line represents the mean of the frames differences. The frames difference above the green line will be maintained and compressed, while the frames difference below the n line will be removed. Part c of the figures 5.15(c), 5.16(c), 5.17(c) and 5.18(c) represents the frames to be removed. As increasing extracting frames per second 10, 15, 20 and 25 frames per second, the frames difference decreased and the removed frames are larger.

With 10 frames per second extraction as shown in figure (5.15) (a), we calculate the average of overall values of frames differencing as shown in Figure (15.5) (b) were the green line represent average value of frame differencing. In this approach, frames were removed if the frames difference below the average as shown in figure (15.5) (c) and the frames were the frame difference above the green line will be maintained and compressed.

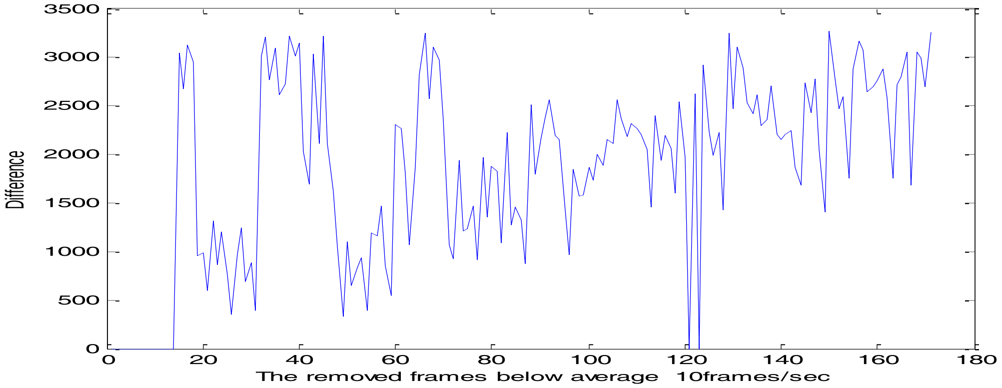


(a) Frames difference.



(b): The frames difference below the Mean (under the line)

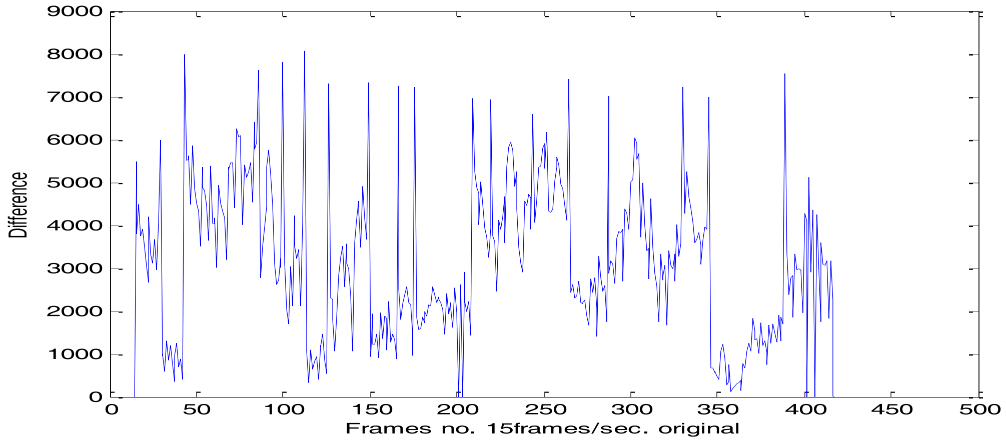
Figure (5.15) Mean difference approach (10frames/second.).



(c): The removed frames below the Mean difference.

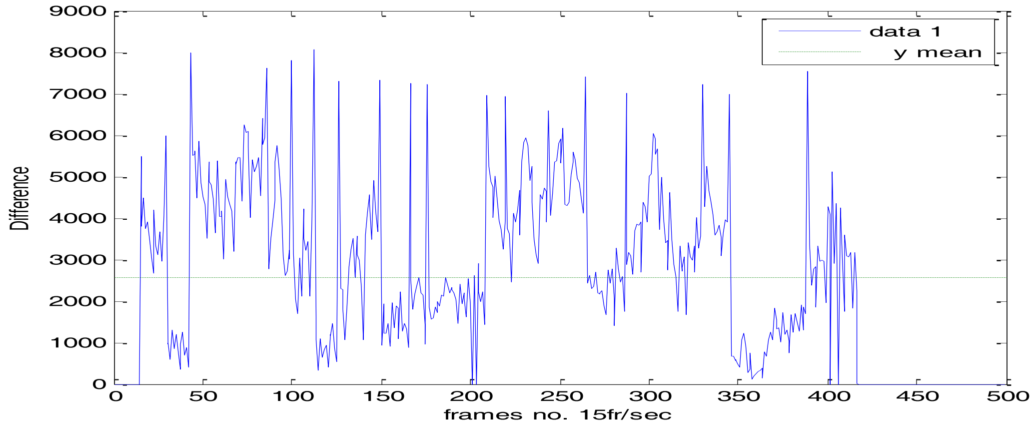
Figure (5.15) Mean difference approach (10frames/second.).

Also with 15 frames per second extraction shown in figures (5.16) (a) and (5.16) (b), the green line represents the average of overall frames differencing, some of the frames were lies above the average and others below, the removed frames below the green line were shown in figure (5.16) (c).

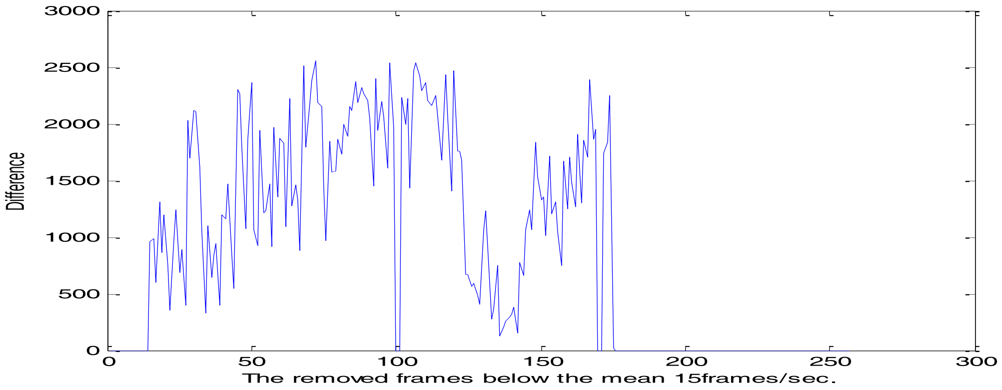


(a) Frames difference.

Figure (5.16) Mean difference approach (15frames/second.).



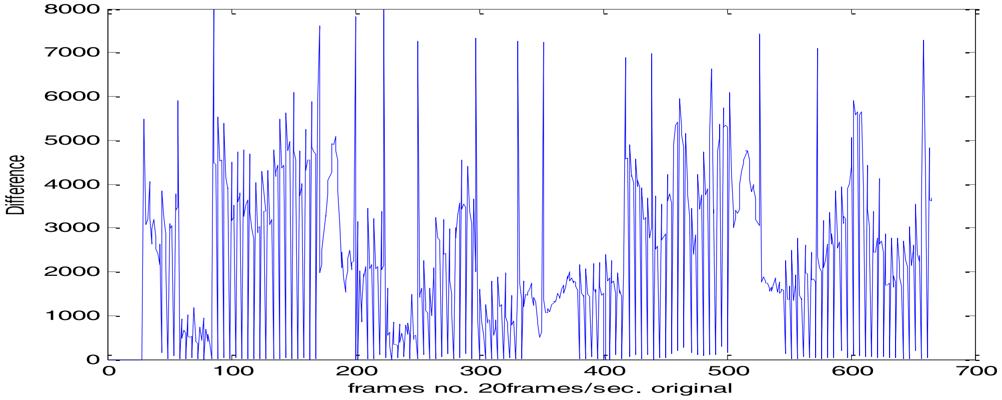
(b): The frames difference below the Mean (under the line)



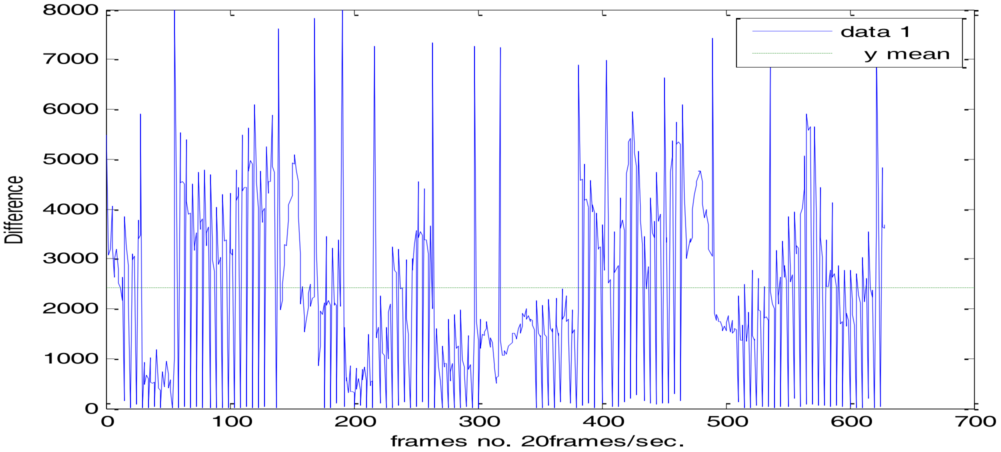
(c): The removed frames below the Mean difference.

Figure (5.16) Mean difference approach (15frames/second.)(cont.)

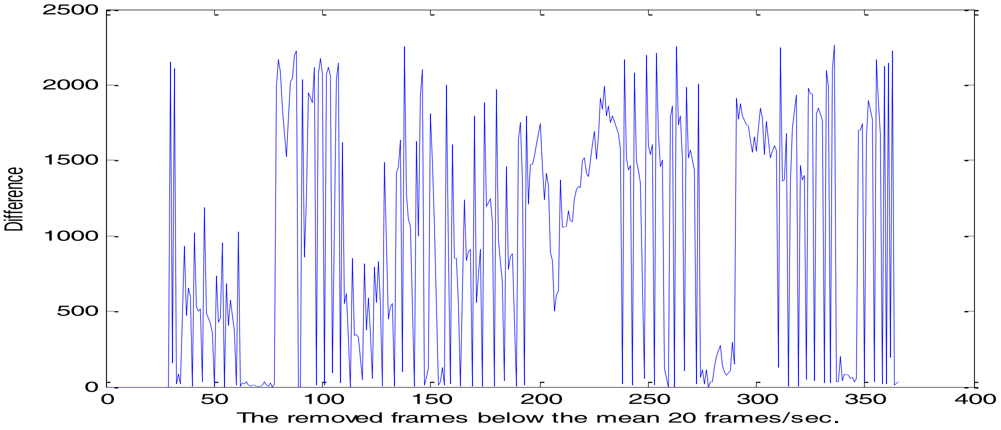
With 20 and 25 of extraction frames per second shown in figures (5.17) (a), (5.18) (a) the green line represents the average of overall frames differencing, some of the frames were lies above the average and others below, the removed frames below the green line were shown in figures (5.17) (c), (5.18) (c) and in this approach many of the frames difference were removed.



(a) Frames difference.

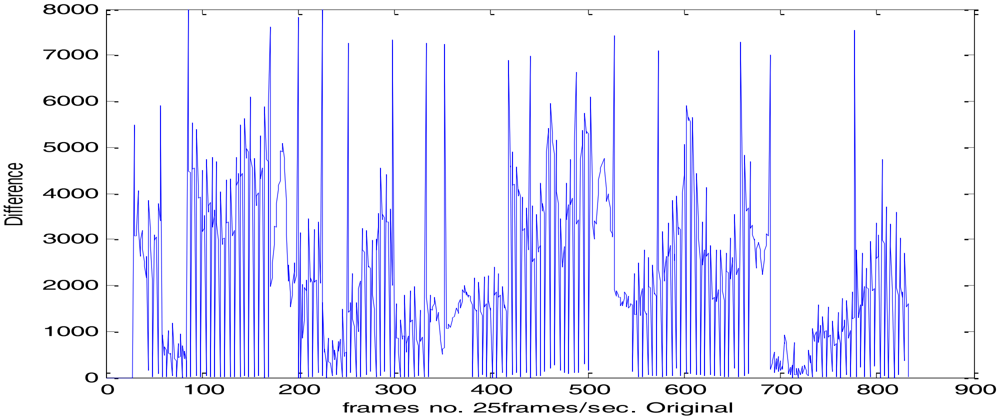


(b): The frames difference below the Mean (under the line)

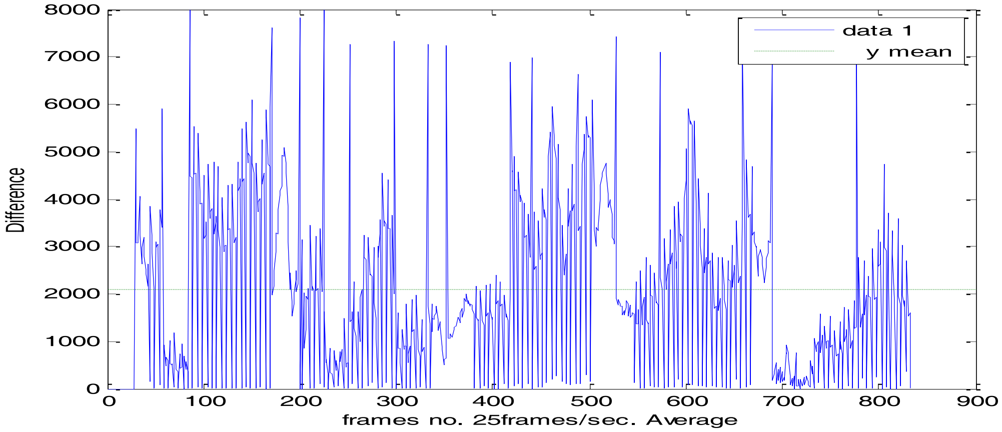


(c): The removed frames below the Mean difference.

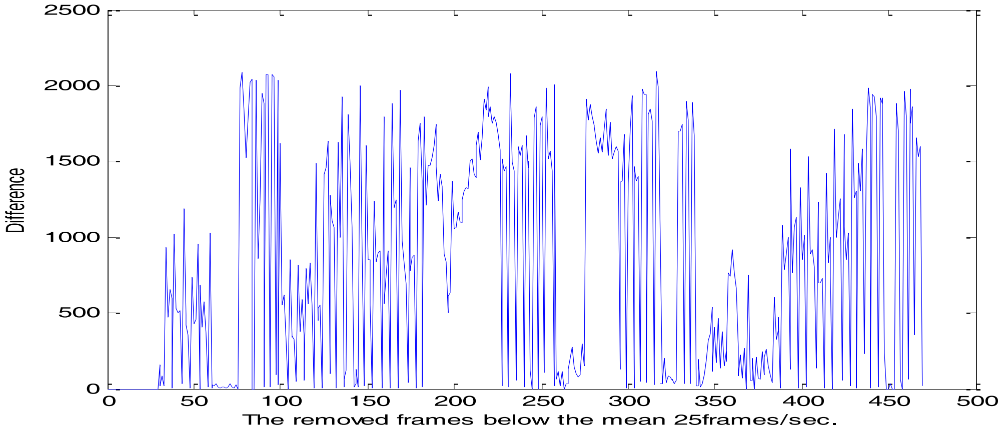
Figure (5.17) Mean difference approach (20frames/second.).



(a) Frames difference.



(b): The frames difference below the Mean (under the line)



(c): The removed frames below the Mean difference.

Figure (5.18): Mean difference approach (25frames/second.).

Subjective Video quality assessment (Mean difference approach)

Results of video compression using mean difference approach is given by table (5.6) and figure (5.19). It is clear that numbers of removed frames are big; therefore, the assessment scores of compressed videos are small compared with the results of zero difference approach results.

Table (5.6) Video quality assessment (Mean difference approach)

Video Quality Factors	**10 frames**	**15 frames**	**20 frames**	**25 frames**
How would you rate video colors?	3.7	3.8	4.0	4.1
How would you rate video contrast?	3.8	4.0	4.0	4.2
How would you rate video borders?	3.8	3.9	4.1	4.3
How would you rate the movement continuity?	3.6	3.8	3.9	4.1
Did you notice any flicker in the sequence?	annoying	Little annoying	Little annoying	Little annoying
Did you notice any smearing in the sequence?	annoying	Little annoying	Little annoying	Little annoying

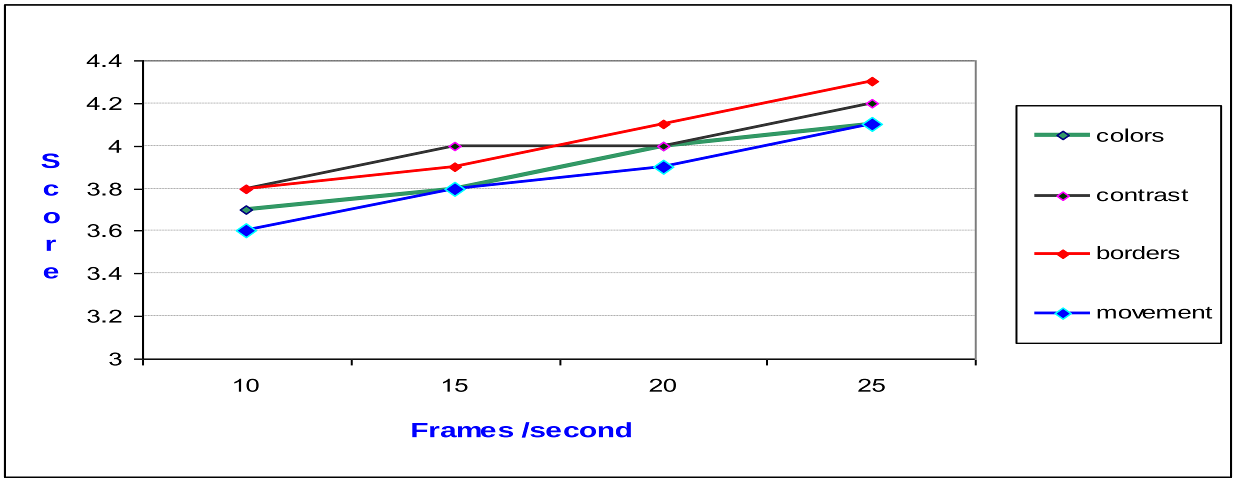
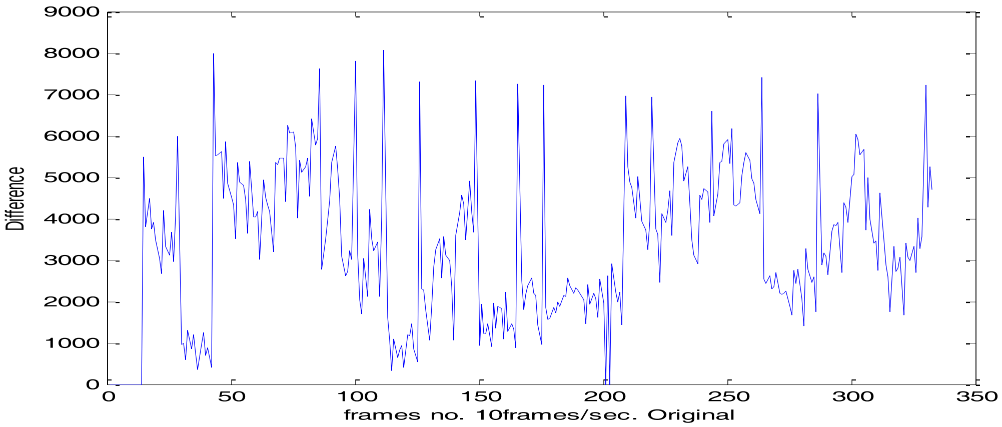


Figure (5.19) Mean difference approach (Scores)

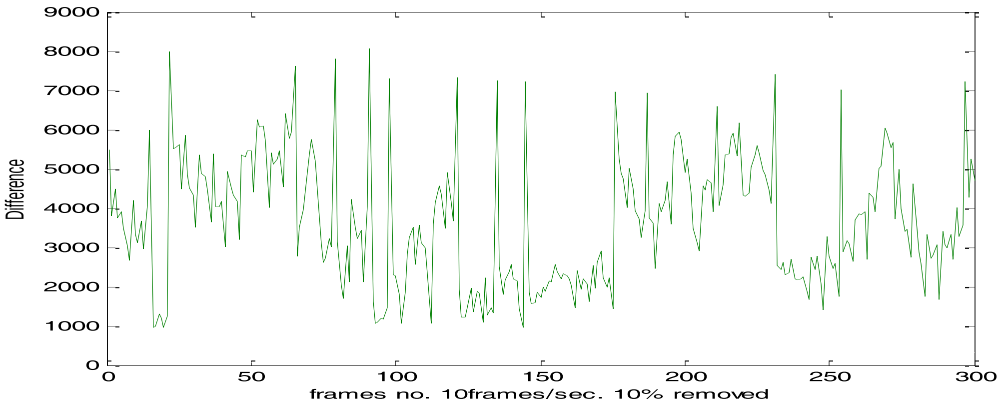
C) Percentage Difference Approach

The implemented approach is concerned with the removed of the lowest frames difference according to some specific percentages. Figures 5.20 (a), 5.21 (a), 5.22 (a) and 5.23 (a) represent the frames difference for different frames per second 10, 15, 20 and 25. Part b of the figures 5.20 (b), 5.21 (b), 5.22 (b) and 5.23 (b) represents frames difference after removing 10% of the lowest frames difference. While figures 5.20 (c), 5.20 (e), 5.20 (g), 5.21 (c), 5.21 (d), 5.21 (e) and 5.22 (c), 5.22 (e), 5.22 (f), 5.23 (c), 5.23 (e), 5.23 (g) represents the frames to be removed. Also different percentages like 15%, 20% were removed as can be shown in figures 5.22d, 5.22e, 5.22f ,5.23e, 5.23f.

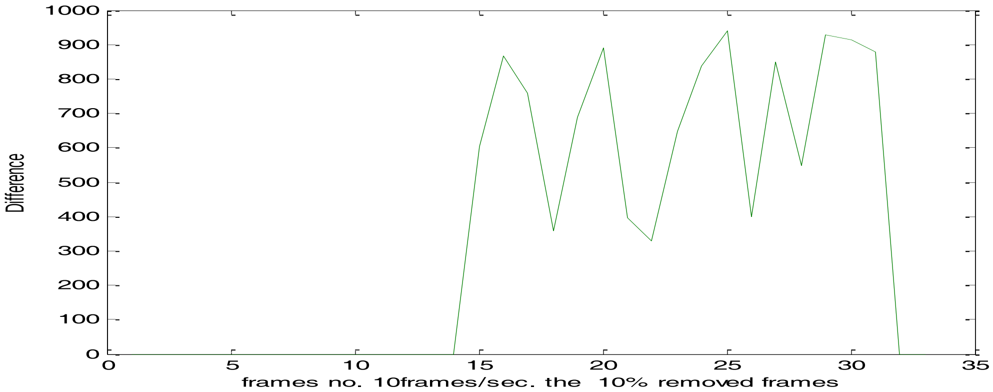
With 10 frames per second extraction as shown in figure (5.20) (a) and with 10 percent of removed frames one can see the remaining 90 percent of the frames difference. While figure (5.20) shows the lowest 10 percent of the lowest frames difference to be removed.



(a) Frames difference (10 frames/second).



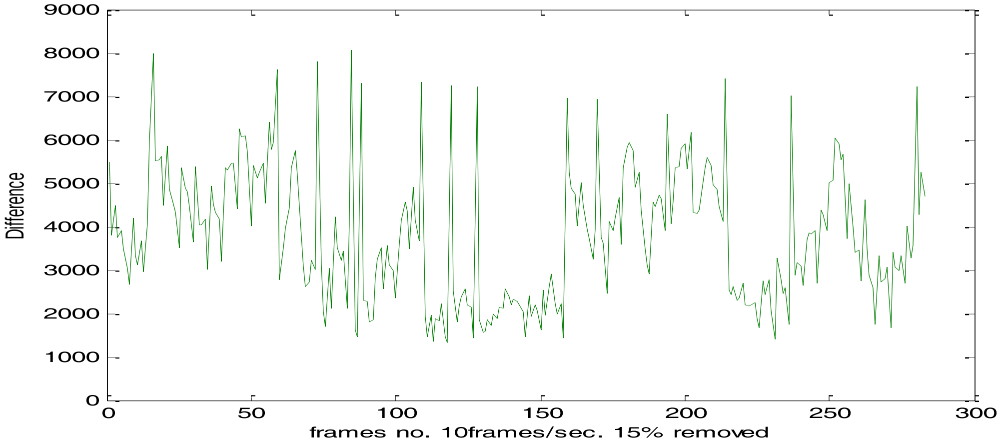
(b): Frames difference after removing 10% of the lowest frames difference.



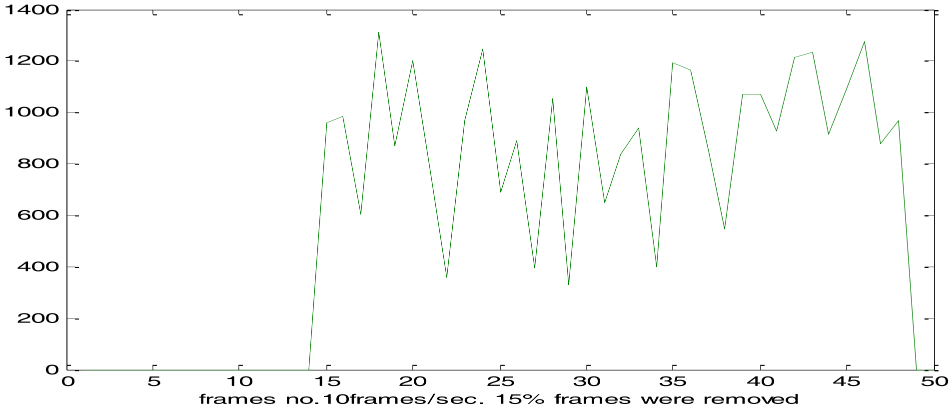
(c): The 10% removed frames.

Figure (5.20) Percentage difference approach (10frames/second)

Also with 10 frames per second extraction and with 15 percent of the removed frames of the lowest frames as shown in figure (5.20) (d). the removed frames were seen in figure (5.20) (e) with the green line. It is clear that the number of the removed frames are higher than the previous one (10percent).



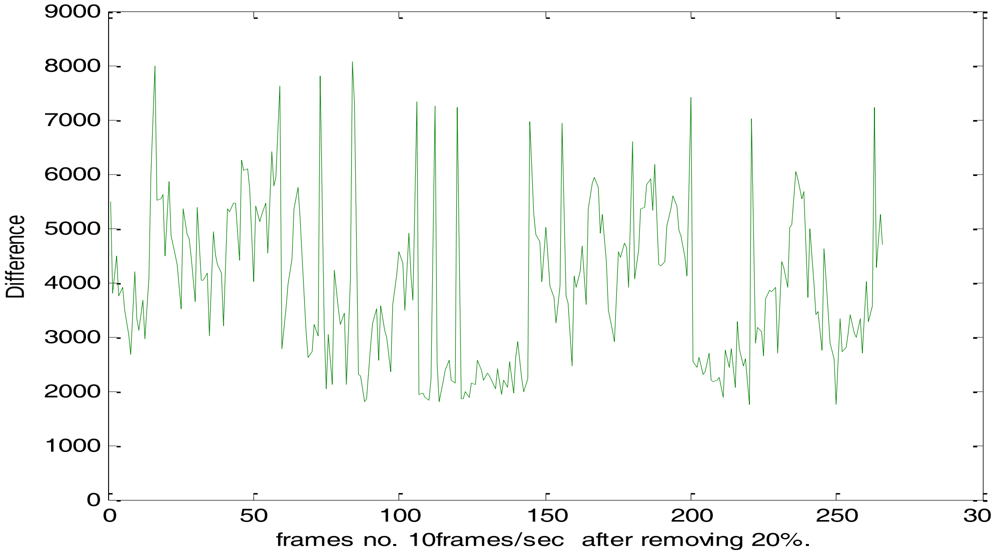
(d): Frames difference after removing 15% of the lowest frames difference.



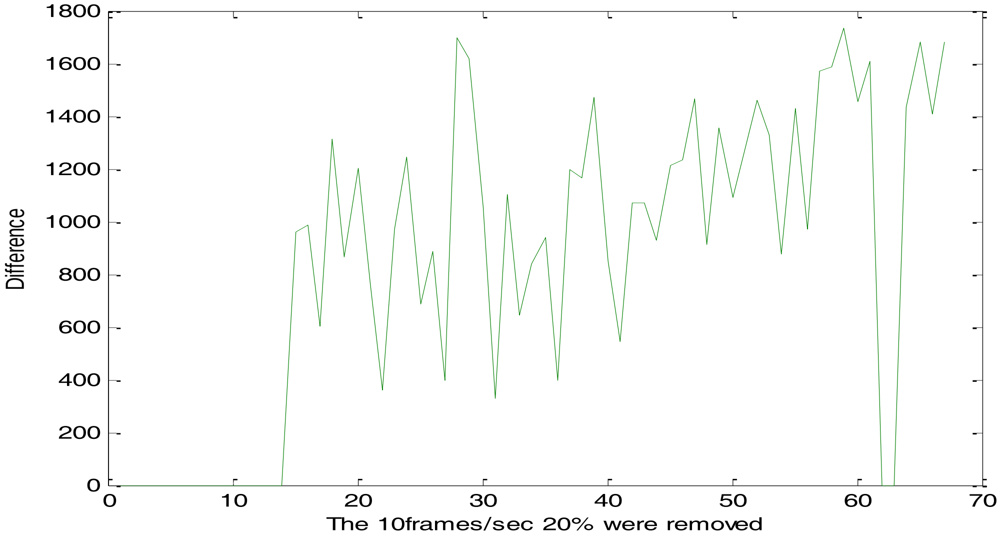
(e): The 15% removed frames.

Figure (5.20) Percentage difference approach (10frames /second) (continued).

With 10 frames per second extraction and with 20 percent of the removed frames of the lowest frames as shown in figure (5.20) (f). The removed frames were seen in figure (5.20) (g) as indicated with the green line. It is clear that the number of the removed frames are higher compared with (10 percent, 15 percent).



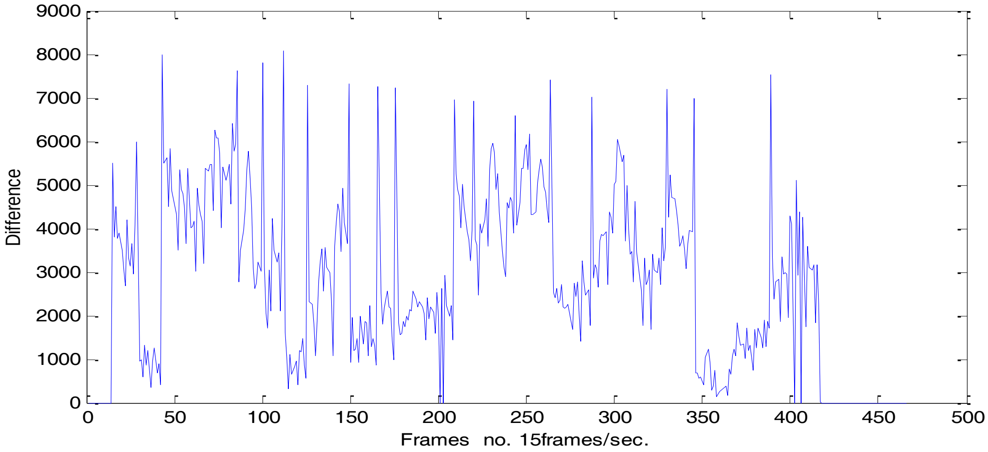
(f): frames difference after removing 20% of the lowest frames difference.



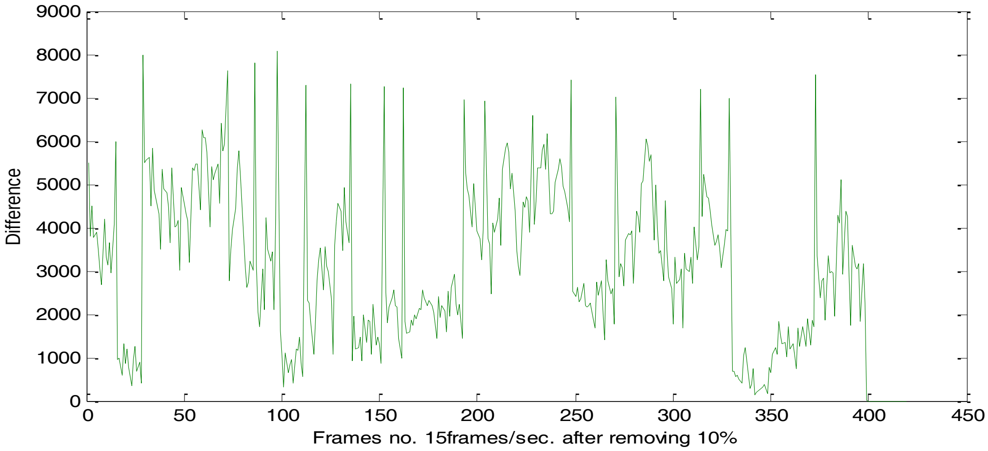
(g): The 20% removed frames.

Figure (5.20) Percentage difference approach (10 frames /second.) (Continued).

In this approach, 15 frames per second extraction (Figure 5.21) (a), 10 percent of removed frames of the lowest frames as shown in figure (5.21) (b). The removed frames were seen in figure (5.21) (c) as indicated with the green line. It is clear that the removed frames have zero differences in all cases. Also with 15, 20 percent of the removed lowest frames can be shown in figures (5.21) (d), (e) respectively.

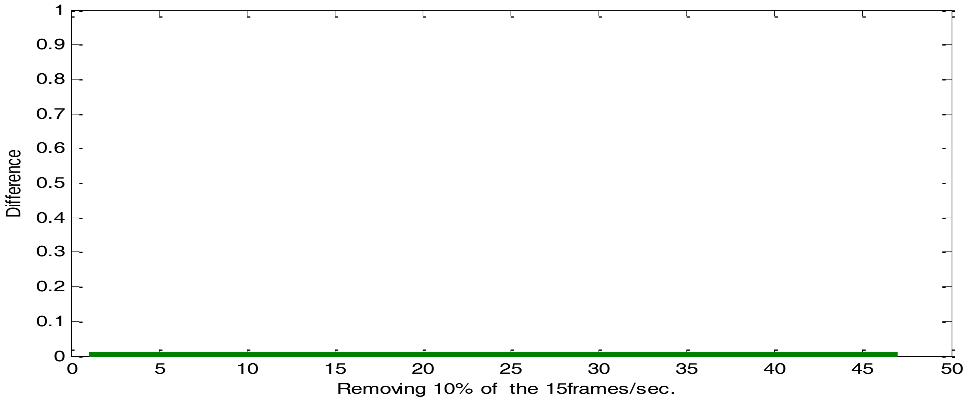


(a) Frames difference (15 frames/second).



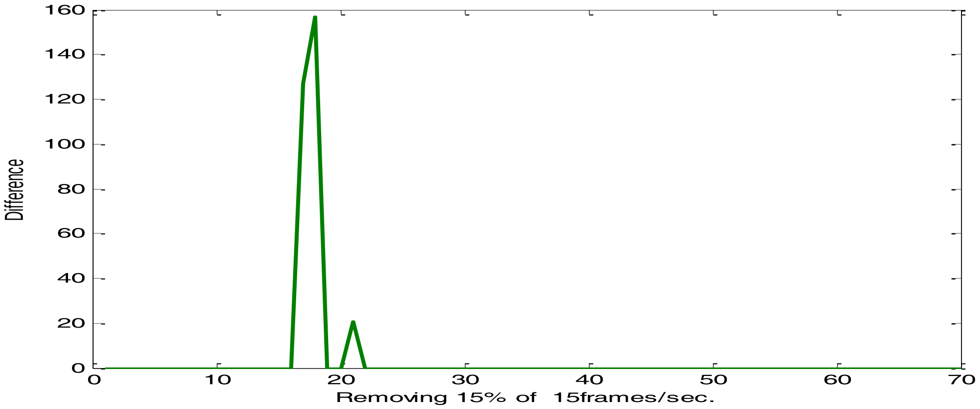
(b): Frames difference after removing 10% of the lowest frames difference.

Figure (5.21) Percentage difference approach (15 frames/second)

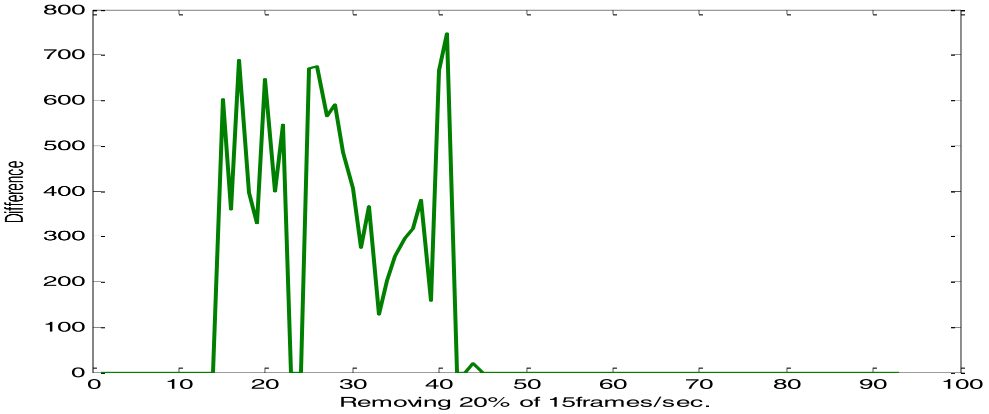


(c): The 10% removed frames.

Figure (5.21) Percentage difference approach (15 frames/second) (Cont.)



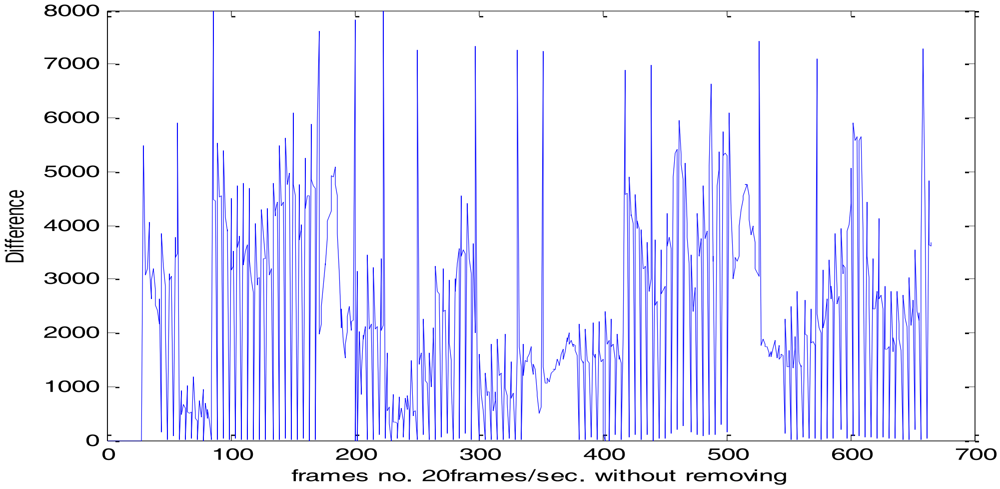
(d): The 15% removed frames.



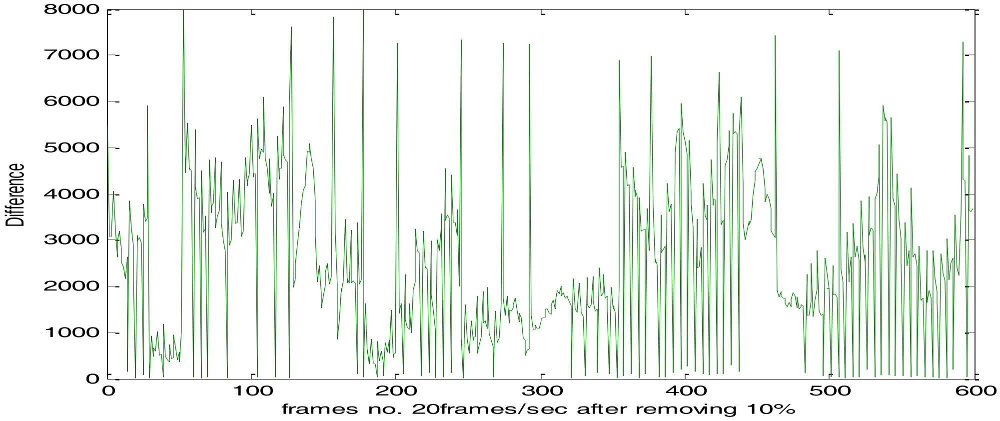
(e): The 20% removed frames.

Figure (5.21) Percentage difference approach (15 frames /second.) (Continued).

In this approach, 20 frames per second extraction (Figure 5.22) (a), 10 percent of removed frames of the lowest frames as shown in figure (5.22) (b). The removed frames were seen in figure (5.22) (c) as indicated with the green line. It is clear that the removed frames are almost zero. Also with 15, 20 percent of the removed lowest frames can be shown in figures (5.22) (d), (e) respectively.

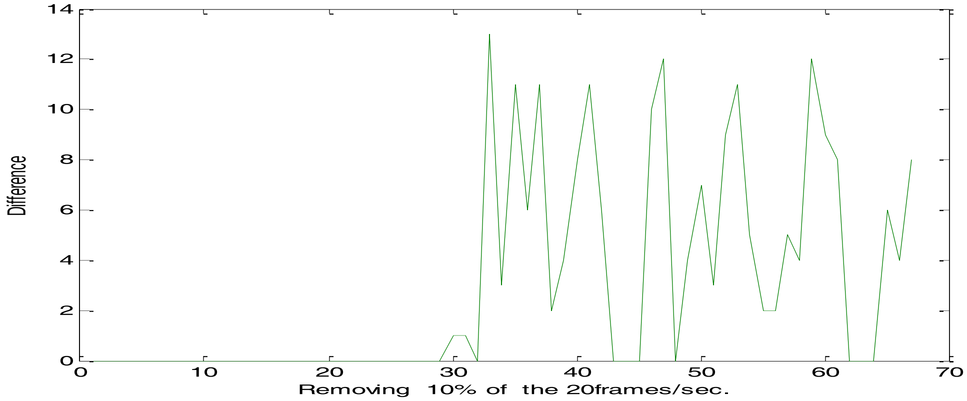


(a) Frames difference (20 frames /second).

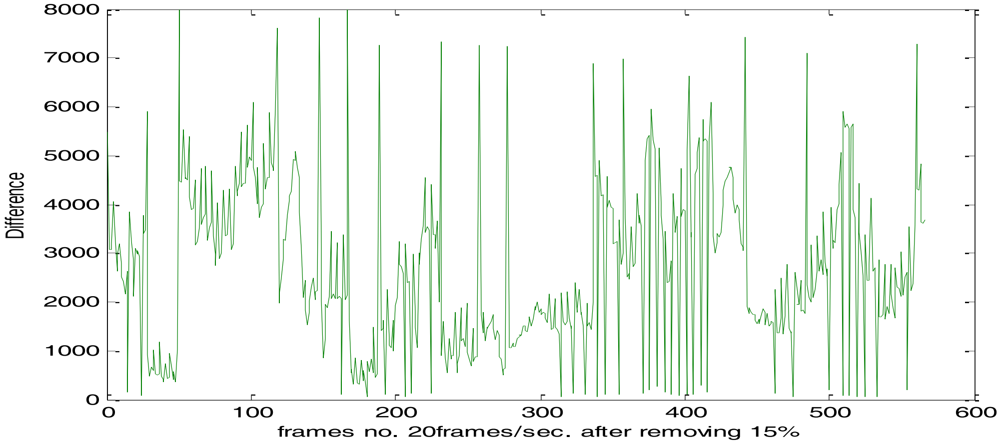


(b): Frames difference after removing 10% of the lowest frames difference.

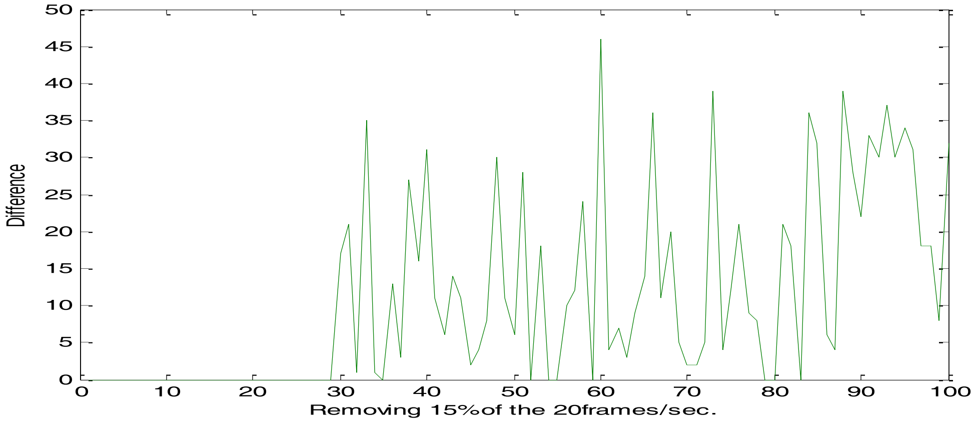
Figure (5.22) Percentage difference approach (20 frames /second.)



(c): The 10% removed frames.

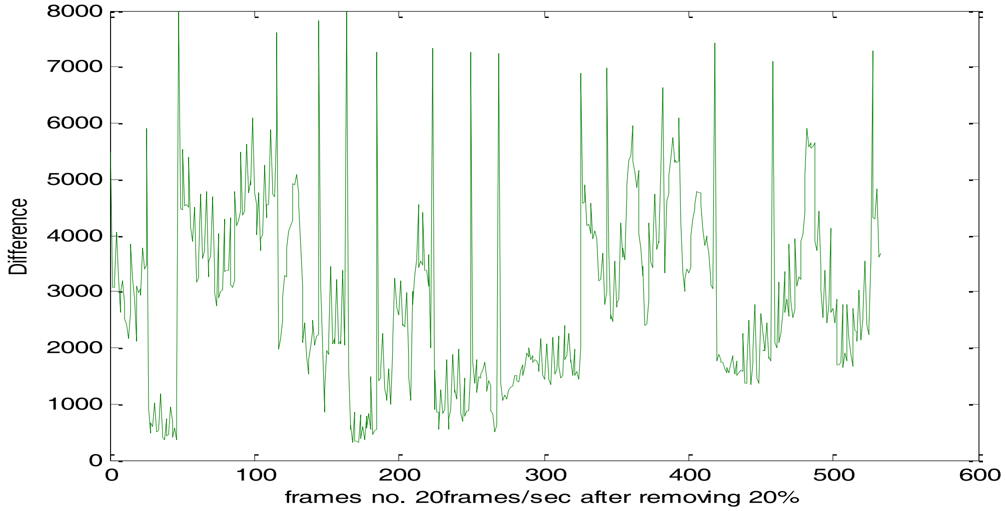


(d): Frames difference after removing 15% of the lowest frames difference.

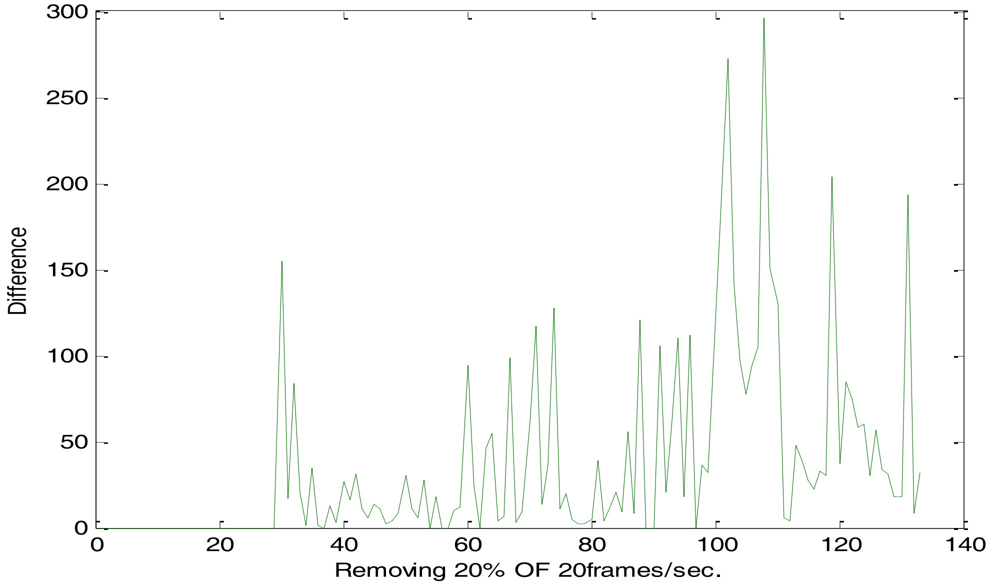


(e): The 15% removed frames.

Figure (5.22) Percentage difference approach (20 frames /second.)(cont.)



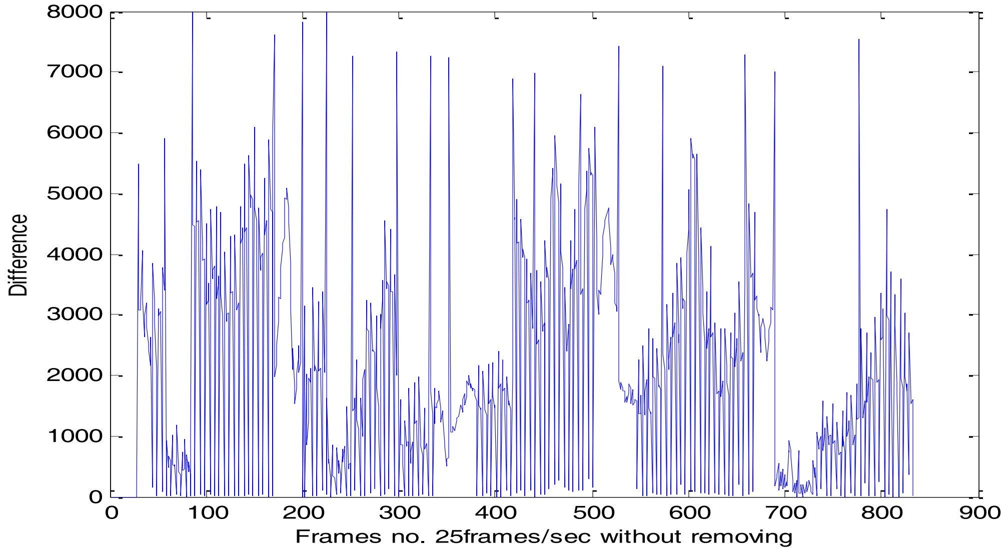
(f): Frames difference after removing 20% of the lowest frames difference.



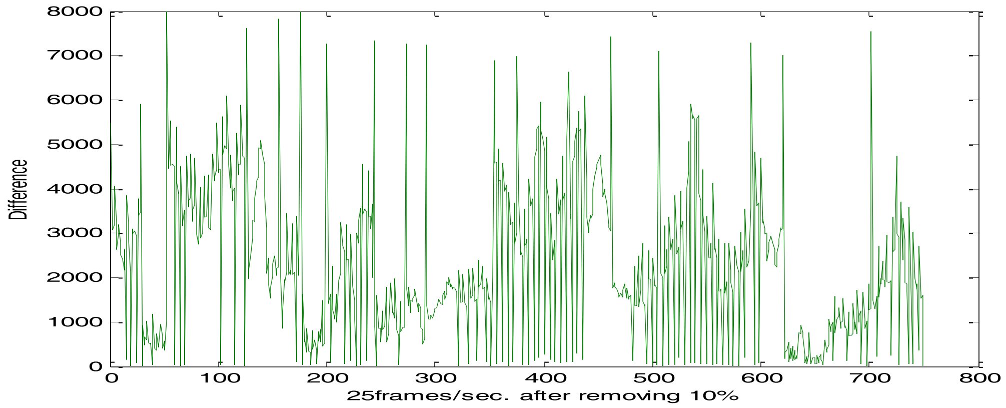
(g): The 20% removed frames.

Figure (5.22) Percentage difference approach (20 frames /second) (cont.)

In this approach, 25 frames per second extraction (Figure 5.23) (a), 10 percent of removed frames of the lowest frames as shown in figure (5.23) (b). The removed frames were seen in figure (5.22) (c) as indicated with the green line. It is clear that the removed frames are almost zero. Also with 15, 20 percent of the removed lowest frames can be shown in figures (5.23) (d), (e) respectively.

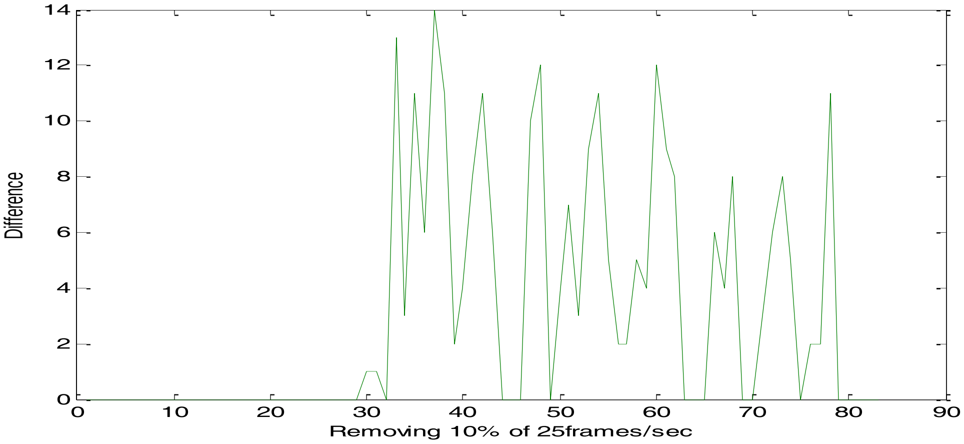


(a) Frames difference (25 frames /second).

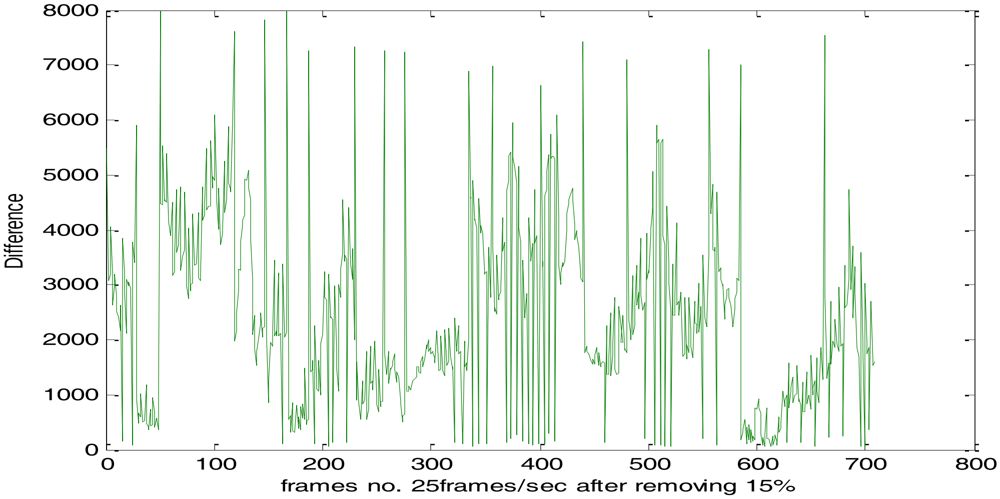


(b): Frames difference after removing 10% of the lowest frames difference.

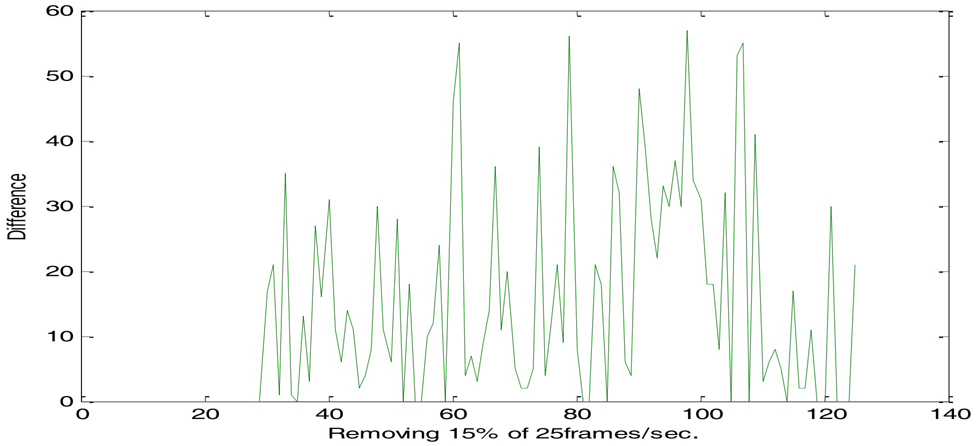
Figure (5.23) Percentage difference approach (25 frames /second)



(c): The 10% removed frames.

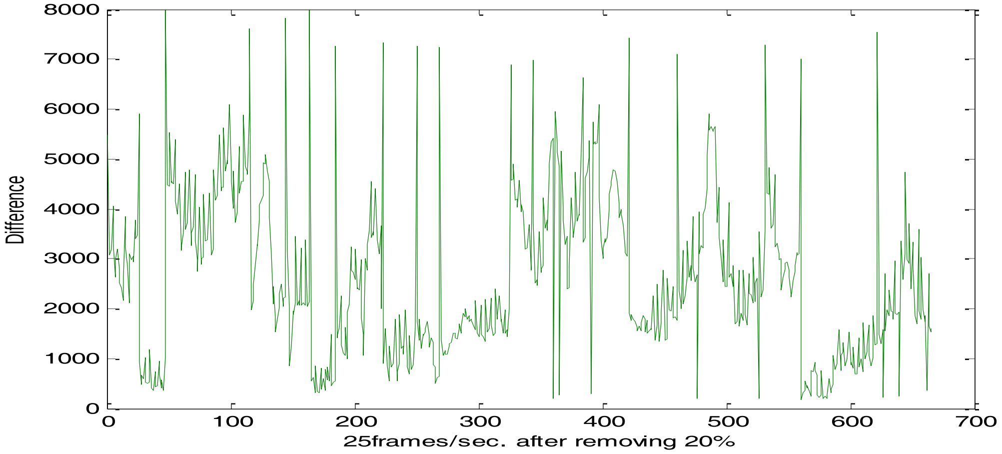


(d): Frames difference after removing 15% of the lowest frames difference.

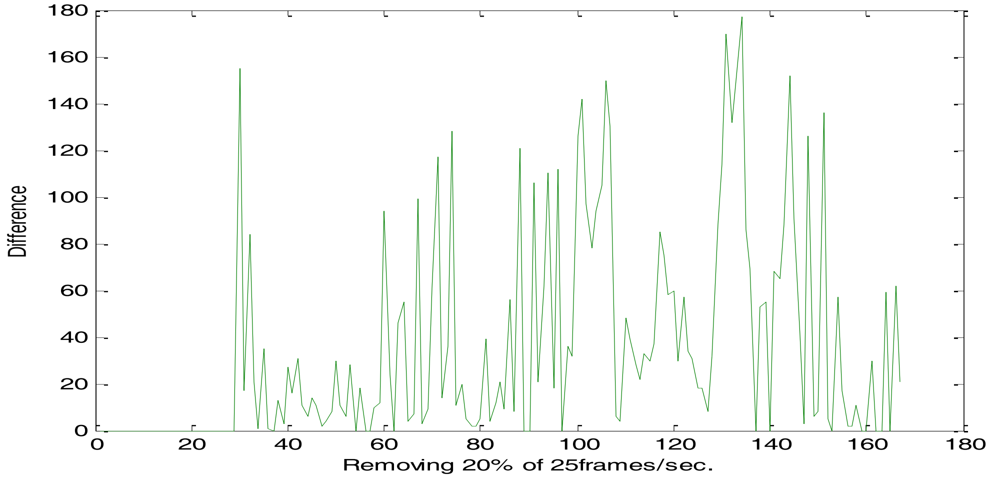


(e): The 15% removed frames.

Figure (5.23) Percentage difference approach (25 frames /second) continued.



(f): Frames difference after removing 15% of the lowest frames difference.



(g): The 20% removed frames.

Figure (5.23) Percentage difference approach (25 frames /second) continued.

Subjective Video quality assessment (Percentage difference approach)

The obtained results of tables (5.7), (5.8) and (5.9) and Figures (5.24), (5.25) and (5.26) shows that the subjective video quality assessment measures increases with the increasing of frames per second (10, 15, 20, and 25), also as percentage of removed frames increased (10%, 15% and 20%) video quality assessment measures decreased as shown in table (5.9) and Figure (5.26).Therefore frames per second and the percentage of removing the lowest frames difference play an important role in video compression.

As shown in table (5.7), figure (5.24), (10, 15) frames per second and with 10 percent of the removed lowest frames difference, the score is given below 4 out of 5. While with (20, 25) per second and with 10 percent removed lowest frames difference, the score gives above 4 out of 5.

Table (5.7) Video quality assessment (percentage difference approach (10%))

Video Quality Factors	10 frames	15 frames	20 frames	25 frames
How would you rate video colors?	3.9	4.0	4.1	4.2
How would you rate video contrast?	4.0	4.0	4.0	4.2
How would you rate video borders?	3.9	3.9	4.1	4.3
How would you rate the movement continuity?	3.9	4.0	4.3	4.5
Did you notice any flicker in the sequence?	annoying	Little annoying	Not annoying	Not annoying
Did you notice any smearing in the sequence?	annoying	Little annoying	Not annoying	Not annoying

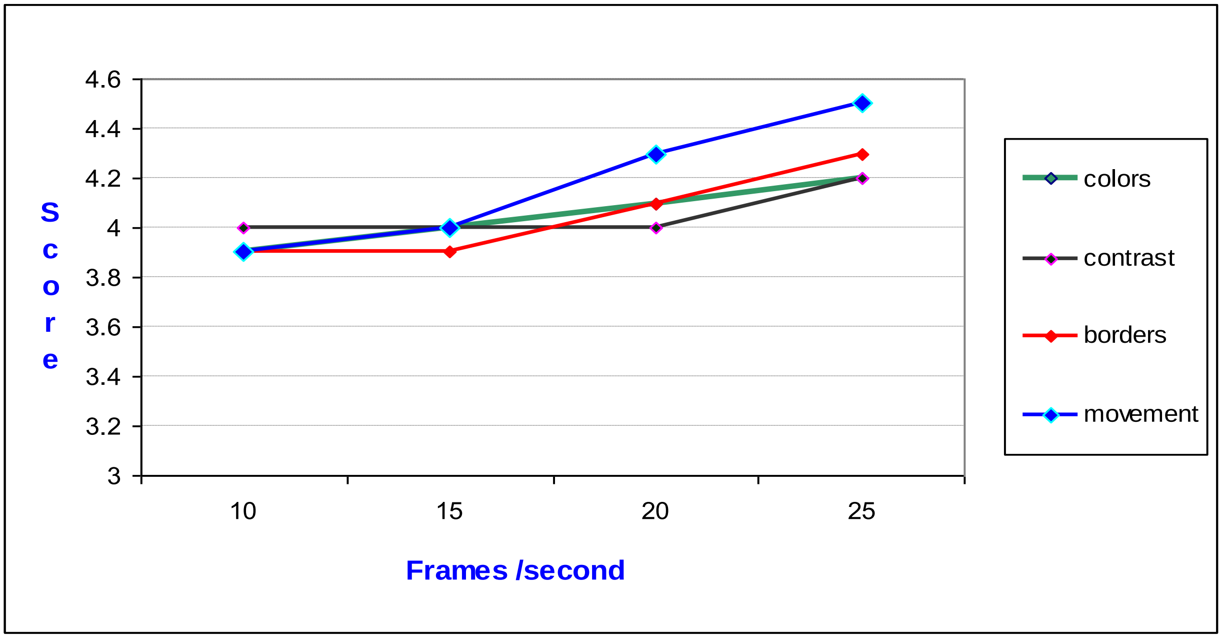


Figure (5.24) Percentage difference approach (10%)

As shown in table (5.8), figure (5.25), (10, 15) frames per second and with 15 percent of the removed lowest frames difference, the average score of video quality assessment is given below 4 out of 5. While with (20, 25) per second and with 15 percent removed lowest frames difference, the score gives above 4 out of 5.

Table (5.8) Video quality assessment (percentage difference approach (15%))

Video Quality Factors	**10 frames**	**15 frames**	**20 frames**	**25 frames**
How would you rate video colors?	3.8	3.9	4.0	4.1
How would you rate video contrast?	3.8	3.9	4.0	4.1
How would you rate video borders?	3.8	3.8	4.1	4.2
How would you rate the movement continuity?	3.6	3.8	3.9	4.3
Did you notice any flicker in the sequence?	annoying	Little annoying	Little annoying	Little annoying
Did you notice any smearing in the sequence?	annoying	Little annoying	Little annoying	Little annoying

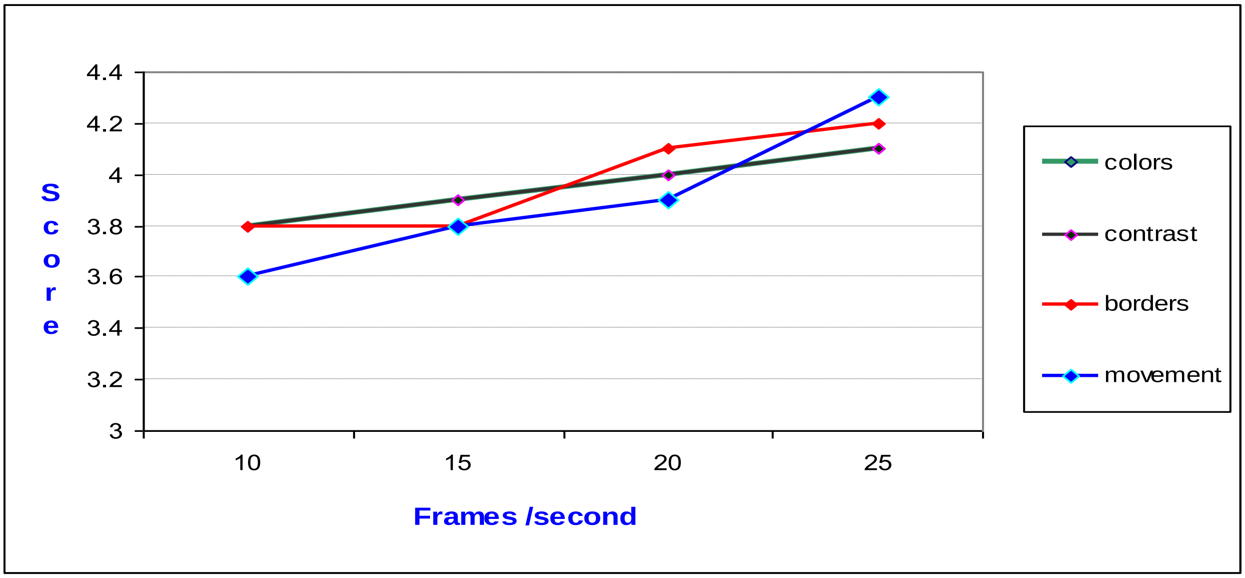


Figure (5.25) Percentage difference approach (15%).

As shown in table (5.9), figure (5.26), (10, 15, 20) frames per second and with 20 percent of the removed lowest frames difference, the average score of video quality assessment is given below 4 out of 5. While with 25 per second and with 20 percent removed lowest frames difference, the score gives above 4 out of 5.

Table (5.9) Video quality assessment (percentage difference approach (20%))

Video Quality Factors	**10 frames**	**15 frames**	**20 frames**	**25 frames**
How would you rate video colors?	3.7	3.8	3.9	4.0
How would you rate video contrast?	3.7	3.8	3.9	4.1
How would you rate video borders?	3.7	3.7	4.0	4.2
How would you rate the movement continuity?	3.5	3.7	3.8	4.0
Did you notice any flicker in the sequence?	annoying	Little annoying	Little annoying	Not annoying
Did you notice any smearing in the sequence?	annoying	Little annoying	Little annoying	Not annoying

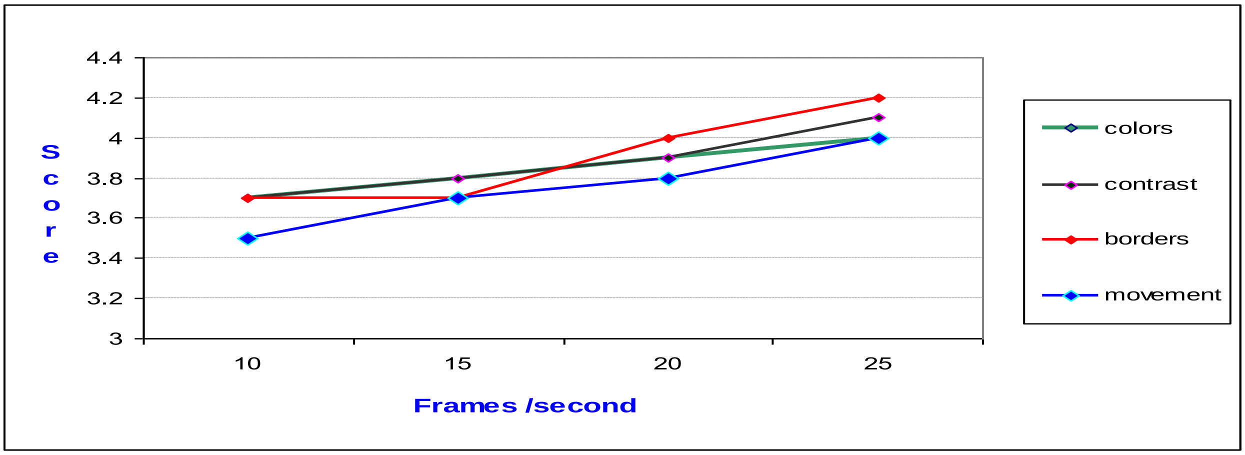


Figure (5.26) percentage difference approach (20%)

From the above results one can notice that the percentage difference approach gives good result comparing with mean difference approach. Mean while it is possible to determine the percentage of the removed frame from the extraction operation. In addition, video types in case of video contents play an important rule in determining the size of compression to be executed.

Chapter 6

Conclusion and Future Works

6.1 Introduction

This work deals with investigating an efficient video compression based on frames difference approaches that concentrated on the calculation of frame near distance (difference between frames) using different wavelet families, where different types of video contents are compressed and tested to insure the efficiency of the proposed technique, in addition a good performance results have been obtained.

6.2 Conclusion

In the first part of the implemented system, different types of color image contents are tested with different types of wavelet families and levels of decomposition. Two classes of methods are applied to measure the image quality of the obtained compressed images, **Subjective tests**, in which human subjects are asked to assess or rank the viewed material, and **Objective tests**, which are computational models that measure the quality by comparing the original and the compressed images.

Compression ratio and selected effective frames are considered as a main feature to adapt information over mobile devices. Image and video contents plays an important role in the quality of the compressed image and video to adapt information via mobile devices. Also sizes of mobile screens were considered in image and video compression.

This research gives the choice of wavelet family types for image compression.

The effects of Haar, Daubechies, Symlets, Coiflets, Biorthogonal and Biorthogonal wavelet family on six different test images have been examined. The Compression Ratios and Visual Image Quality are also presented. Peak Signal to Noise Ratio (PSNR) is used as the objective quality measure. The results are analyzed for a wide range of wavelets. An example of a natural plant image was compressed and tested for different wavelet families , the obtained results shows that the best MSE & PSNR gives by Coiflets wavelets transform (Coif2) 133 & 26.92 respectively
While Ribo2.2 gives maximum MSE (245) & less PSNR (24.26).

Also different types of images are used with different details such as News, Sport, Music, Series, Cartoon and Panorama. And up to fourth level of wavelet compression are applied to these images. To compare the subjective quality, it can be concluded that the obtained results denoted that there are no mentioned difference between the original images and their corresponding compressed and reconstructed images.

To compare the objective results measures, PSNR and MSE are calculated, the obtained results of PSNR indicate that there is a small difference indicated for Haar, Db2, Ribo2.2 and Ribo2.4 for all types of images and all types of wavelet levels. The obtained results of MSE indicate that it is normal that MSE increase with the increasing of compression ratio or wavelet levels. MSE measures indicate that Ribo2.2 gives maximum MSE for all wavelet levels and for each image type.

The next MSE after Ribo2.2 is Haar then Ribo2.4. The best wavelet family results give by coif4 for the most image types (News, Sport, Music, Cartoon, and Series).

The obtained results explained the important features of wavelet transforms masks that applied on image compression. The proposed algorithm based on an effective image compression technique that utilizes wavelet mask selection approach to reach high performance results. A comparative study of different types of DWT (with different masks) are implemented then select the perfect mask for a certain image depending on image size, compression ratio, wavelet level and reconstructed image quality. The objective measures (MSE and PSNR) and the subjective measure are obtained that lead to optimal wavelet mask selection.

Hence it can be concluded that choice of wavelet depends not only on the size of the image but also on the content of the image. Hence, that the choice of wavelet in the process of image compression depends on size of the image and content of the image for desired image quality.

In the second part of the implemented system of video compression, an efficient video compression method based on frames difference approaches are developed that concentrated on the calculation of frame near distance. Many factors are applied in the selection of meaningful frames, in which eliminate the similar frames. The implemented system passes through many steps; preprocessing, frame extraction, frame selection, frame reordering, 2D-DWT, then video construction.

Different types of video contents were compressed and tested. The obtained results shows that with 25 frames per second extracted, good results of the compressed video are obtained using subjective video quality assessment in which MOS gives on the average (4.5/5) for the case of removing all zero difference frames and the lowest 5% percentage of the frames difference of the extracted frames.

6.3 Future works

There are many possible extensions to this research, these include the following:

- Finding the best thresholding strategy in determining frames difference approaches and generalized according to types of video and images contents.
- Finding the best wavelet families for a given image or video.
- Investigating other wavelet families.

One should be aware that there are many simple algorithms or mathematical equations which claim to be able to compute video quality scores like PSNR and MSE. But the difficulty is to create an algorithm that computes scores that are highly correlated with subjective quality scores (which means creating a program that output scores which are very close to the quality scores given by human observers during subjective video quality assessment tests in normalized conditions).Therefore one should avoid using these simple methods (like PSNR) because, even if they give you quantitative scores, these scores have no relation (or very few) with subjective judgments of video quality.

However, data compression methods in general might be far away from the ultimate limits. Interesting issues like obtaining accurate models of images, optimal representations of such models, and rapidly computing such optimal representations are the grand challenges facing the data compression community.

Image coding based on models of human perception, scalability, robustness, error resilience, and complexity are a few of the many challenges in image coding to be fully resolved and may affect image data compression performance in the recent years.

References

[1] Vasos Vassiliou, et al, "Requirements for the transmission of streaming video in mobile wireless networks University of Cyprus. . May 2006.

[2] Thomas McNeal et al " Any where, anytime: Using mobile phone for learning" Journal of the research for educational technology. Vol. 3, No. 2 (2006, winter).

[3] Yun Q.Shi and Huifang Sun A book entitled " Image and Video Compression for Multimedia Engineering, Fundamental Algorithm, and Standards" second edition.

[4] P. Santha et al, “Efficient Wavelet based Image Compression Technique for Wireless Communication”, International Journal of Innovative Technology & Creative Engineering. Vol. 1 No. 3 March 2011.

[5] Baluram Nagaria et al, ‘Comparative Analysis of Fast Wavelet Transform for Image Compression for optimal Image Quality and Higher Compression Ratio. International Journal of Engineering Science and Technology. Vol. 3 No. 5 May 2011.

[6] Sherin Kishk et al, ‘Integral Images Compression using Discrete wavelets and PCA’, “International Journal of Signal Processing and Pattern Recognition”, Vol. 4 No.2 June 2011.

[7] Hojung Cha, et al "Video Streaming System for Mobile Phone: Practice and Experience, Wireless Network vol. 11, No. 3, May 2005.

[8] Ahmed Ali, et al. "Video transmission to mobile phones over wireless Network". Telekom Research & Development Sdn, December 2004

[9] M D Walker et al. "Mobile video-streaming" BT Technology Journal. Vol. 21 No. July 2003..

[10] Satyajayant Misra, et al." A survey of multimedia streaming in wireless sensor networks". Communications surveys & tutorials IEEE, vol. 10 issue 4. Fourth quarter of 2008.

[11] Muzhir Sh. Al-Ani et al " Constructing a New communication system by integrating the GSM to the Satellites Infrastructure". WSEAS transactions on communications. Issue 2 Vol. 8 February 2009.

[12] Stephan Brumme, "Multimedia streaming on mobile phones", advanced data communications, University of Technology, Sydeny,2004

[13] Giovanni Gualdi, et al "Video Streaming for Mobile Video Surveillance" IEEE Transactions on multimedia, vol. 10, no 6 October 2008.

[14] Kosuke Numa , et al " Using Mobile Phone Function as expression tool" Design of Keitai Trial Workshop. www.qwik.jp/PerCAS/Program.html. accessed at 20/8/2009

[15] Patricia Michael “A mobile Voice: The Use of Mobile Phones in Citizen Media". Mobileactive.org. produced by USAID, November 2008.

[16] Emund Stephen A thesis entitled “High Ratio Wavelet Video Compression through Real-Time Rate-Distortion Estimation “ University of Natal, Durban , July 2003.

[17] C.S Patchichis " Wireless Telemedicine Systems : An Overview" IEEE vol. 44 No. 2 , April 2002.

[18] Mohammed Benabdalla " Choice of reference images for video compression" applied mathematical sciences Vol.1 , 2007, no. 44 pages 2187-2201

[19] S. Annoaduria et al "Wavelet based enhanced color image compression relying on sub-band vector quantization", ICGST-Gvip Journal ISSN 1687-398X Volume (9), ISSUE (I) February 2009.

[20] Tmt. Nishat et al " Performance measure of different wavelets for a shufflled image compression scheme" IGCSNS International Journal om Computer Science and Network Security, Volume 9 Number 3 h 2009 India

[21] Yogendra et al " Choice of wavelet from wavelet families to facilitate imag Video compression 2005].Smarat Ashok Technological Institute, Vidis Vidisha(M.P) 464001 India.

[22] G.k.Kharate et al "Selection of mother wavelet for image compression on basis of nature of image" Journal of Multimedia Vol. 2 No 6 November 2007.

[23] Sourav et al "Real time video streaming over Bluetooth network between two mobile nodes" IGCSI International Journal of Computer Science Issues Vol. 7 Issues 3, No. 11 MAY 2010.

[24] Eric etal " Vide streaming with SP and SI frames " Information system laboratory" Stanford University , Stanford, CA 94305

[25] Rakesh et al, "A color video compression technique using key frames and a low complexity color transfer" Visual Information Processing and Communication II. Edited by Said, Amir; Guleryuz, Onur G.; Stevenson, Robert L. Proceedings of the SPIE, Volume 7882, pp. 78820P-78820P-10 (2011).

[26] Gyanendra Kr. Verma" a thesis entitled" Speaker Identification System using Wavelet Transform" Indian Institute of Information Technology, 2009.

[27] Mamta Sharms " Compression Using Huffman Coding", IJCSNS International Journal of Computer Science and Network Security, Volume . 10, No. 5, May 2010.

[28] Arun Kumar PS A master thesis Entitled "Implementation of Image Compression Algorithm using Verilog with area, Power and Timing Constraints", National Institute of Technology, Rourkela , 2009.

[29] Vinay U.Kale et al "Performance Evaluation of Various Wavelets for Image Compression of Natural and Artificial Images." IJCSC International Journal of Computer Science and Communication Vol. 1, No. January-June 2010 pp. 179-184.

[30] Karen Lees A thesis Entitled " Image Compression using Wavelets" accessed at : http://web.cecs.pdx.edu/~mperkows/CAPSTONES/HAAR/u9kvl.pdf

[31] Bernie Brower "Image Compression basics' IMA system engineering services.
http://www.math.cuhk.edu.hk/~lmlui/CompressionTutorial.pdf

[32] PRESTO – Preservation Technologies for
European Broadcast Archives
IST-1999-20013 Accessed at:
http://presto.joanneum.ac.at/Public/D5_4.pdf

[33] Qiang Wang et al "Content –Based 2D-VLC Entropy Coder in AVS Video Coding Standard", Journal of Computer Science & Technology ,May 2006, Vol. 21, No. 3, pp 315-322.

[34] Guy E. Blelloch "Introduction to Data Compression", Carnegie Mellon University. September 25, 2010. Accessed at :
http://www.cs.cmu.edu/~guyb/realworld/compression.pdf

[35] Rafael C. Gonzalez, Richard E. Woods a Book entitled "Digital Image Processing " 2nd Edition.

[36] Xavier Alemeda i Pineda " Entropy Coding for image Compression Based on Generlized Lifting and SPECK", TECHNICAL University of Catalonia, September 2009.

[37] Christina Zeeh, "Seminar "Famous Algorithms"
January 16, 2003.

[38] _Myug –Sing Song "Wavelet Image Compression”, 1991 Mathematics Subject Classification, Primary 42c40. http://www.siue.edu/~msong/Research/article.pdf

[39] G. Boopathi et al “Image Compression: An approach using t Transform and Modified FCM”, International Journal of Computer Applications “, Volume 28- no. 2 August 2011.

[40] SVI Support Wiki Pages Scientific Volume Imaging B.V http://www.svi.nl/SoftThreshold

[41] Wavelet Toolbox for use with MATLAB http://web.mit.edu/1.130/WebDocs/wavelet_ug.pdf

[42] Wavelet Browser by Bywavelets
http://wavelets.pybytes.com/

[43] Gerhard X. Ritter and Joseph N. Welson “Handbook of Computer Computer Vision Algorithms in Image Processing “, CRC Press 2000.

[44] Andrew B Watson “Image Compression Using the Discrete Cosine Transform “, Mathematics Journal, 4(1), 1994, P. 81-88.

[45] V. Santhi , “DC Coefficients based watermarking technique for color images using singular value decomposition'. International Journal of Computer and Electrical Engineering Vol. 3 No.1 , February 2011.

[46] Sugreev Kaur1 et al ,"High Speed and Area Efficient 2D Dwt Processor Based Image Compression Signal & Image Processing: An International Journal (SIPIJ) Vol.1, No.2, December 2010.

[47] **A. Wims et al ," Blind Source Separation using Wavelet”, 2010 IEEE International Conference Intelligence and Computer Research.**

[48] **Mark Eastaway ," The Inverse Discrete Wavelet accessed at: Transform http:www//cnx.org/content/m18987/latest/**

[49] Ken Cabeen and Peter Gent " Image Compression and the Discrete Cosine Transform", College of the Redwoods, accessed at: http://www.lokminglui.com/dct.pdf

[50] Charilaos Christopoulos1 " The JPEG2000 still image Coding System: An overview ",IEEE Transactions on Consumer Electronics, Vol. 46, No. 4, pp. 1103-1127, November 2000.

[51] David S. Taubman et al,"JPEG2000: Standard for Interactive Imaging", proceedings of the IEEE, Vol. 90, No. 8, August 2002.

[52] V.Sanchez et al ' The JPEG2000 image compression standard', University Alberta' Accessed at 29/7/2011: http://webdocs.cs.ualberta.ca/~anup/Courses/604/NOTES/slide_jpeg2000.p

[53] Athanassios Skodraset et tal "The JPEG 2000 Still Image Compression Standard' IEEE Xplore. March 31, 2009.

[54] Li & Drew a book entitled" 'Image Compression Standards' Chapter 9 Copy Right Prentice Hall 2003.

[55] Ratchakit Sakuldee et al “Objective Performance of Compressed Image Quality Assessments, “World Academy of Science, Engineering and Technology 35 2007.

[56] S.Medouakh et al "Study of the Standard JPEG2000 in Image Compression',
International Journal of Computer Applications (0975 – 8887) ,Volume 18, No. 1, March 2011.

[57] Multimedia Systems and Applications Series Consulting Editor Barko Furht A book entitled “The VC-1 and H.264 Video Compression Standards for Broadband Video Services.

[58] Garry C. Sullivan et al "Video compression –from concepts to the h.264/av
Standard",Proceeding of IEEE vol. 93 no.1, January 2005.

[59] Borko Farht et al "Video presentation and compression chapter 9" Florida Atlantic University accessed at: http://www.cse.fau.edu/~borko/Chapter9_mc.pdf

[60] AXIS communications ,”H.264 Video Compression Standard”, 2008, accessed at : http://www.hellascams.gr/grc/products/videoser/pdf

[61] Lain E.G Richsrdson , a book entitled "H.264 and MPEG-4 Video Compression" copy right 2003. Chapter 7,Accessed at : http://mirror05.x264.nl/Dark/videobook.pdf

[62] Anil Aksay "Motion Wavelet Video Compression" Middle East Technical University, master thesis 2001.

[63] Cambridge University press , Sample pages of “The art of scientific computing”, Wavelet Transforms Copy right © 1988-1992 pages 591 accessed at : http://cacs.usc.edu/education/cs653/NRC-c13.10-Wavelets.pdf

[64] T.Sicora “MPEG-1 and MPEG-2 Digital Video Coding Standards”, digital consumer electronic handbook Mc Grow hill book company.

[65] Telchemy Actively managing multimedia,” Understanding IP Video Quality metrics”, march 2011.

[66] **[**Vinay U.Kale & Nikkoo N. Khalsa**,”**Performance evaluation of various wavelets for image compression of natural and artificial images” IJCSC Vol. No.1 Januaray - june 2010 pp 179-184.