UniversityABBES LAGHROUR Khenchela
Faculty of Sciences and Technology
Departement of Industrial Engineering
جامـــعة عبـــاس لــــغرور خـنـشـلـة
كليـــة العلـــوم والتكنـــولـــوجيـا
قسم الهنـدسة الصنـاعية

N° Serie :.....................

A Thesis

Presented to Fulfill the Partial Requirement for Master's Degree

Section: Telecommunications

Speciality: Telecommunications Systems

TITLE

Compressed video transmission using multiple description coding over AWGN Channel

Realized by: : - Kenzari Ouafa

- Merah Issam

Supported on 15 /09 /2020 In front of the jury composed of:

Mrs . FOUZIA MAAMERI	*President*	*University of Abbes Laghrour-Khenchela*
Mr. FAROUK BOUMEHREZ	*Superviser*	*University of Abbes Laghrour-Khenchela*
Mr. HAKIM SAHOUR	*Examiner*	*University of Abbes Laghrour-Khenchela*

Promotion 2019/2020

Acknowledgements

First, I thank Allah the Almighty for giving me courage, strength and patience to complete this modest work.

I wish to express my profound gratitude to Mr.BOUMEHREZ Farouk Rafik, my Respected supervisor, for his excellent guidance, caring, patience, I ask Allah to bless him.

My respect and gratitude to the jury members who gave me the honor of judging this work through their availability, observations and reports that have enabled me to enhance my scientific contribution.

Thanks to all the teachers of our faculty of Sciences and Technology

I thank all those who participate.

Dedicates

I would like to dedicate this work To the spring that never stops giving…

To my mother

To the big heart... My dear father.

To my dear sisters for their moral support AFAFE, RAHIFA, NADHJET and to

my niece ANAYIESS .

To my friends who encouraged me RANOU, SABAH

To the people who paved our way of science and knowledge All our distinguished teachers To every person who supported me in my studies.

I also dedicate this dissertation to my friend and partner ISSAM , and to all my classmates

Thank you

OUAFA

Dedicates

I dedicate my dissertation work to my family and friends. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears.

To every person who supported me in my studies.

To all our distinguished teachers

I also dedicate this dissertation to my friend and partner WAFAA, and to all my classmates

Thank you

ISSAM

ملخص:

تعد تقنية المعلومات من أهم مجالات التطوير مهم نظرًا لارتفاع الطلب من المستخدمين، الذين أصبحوا الآن جزءًا من يوميًا والتي أصبحت ضرورية لتقليل أو ضغط حجم الوسائط (صوت وفيديو وفيديو) وتسمى أيضًا (الوسائط المتعددة) .يعد ضغط الفيديو أحد التطبيقات المتطلبة حيث توجد كميات كبيرة من البيانات يجب التعامل معها. في هذه الدراسة، قمنا بتقييم معيار ترميز الفيديو الجديد [ترميز فيديو عالي الكفاءة]] /H.265) مع بعض المتغيرات، بما في ذلك محتوى الفيديو، نسبة الإشارة إلى الضوضاء، والمشكلات الرئيسية التي تظهر هي فقدان الحزمة، التقطع، عرض النطاق، تم اقتراح تقنية تشفير الوصف المتعدد (MDC) للاستخدام في أنظمة إرسال الصوت والفيديو كوسيلة لحل المشكلات، في مجموعة متنوعة من سيناريوهات التطبيق أجريت التجارب باستخدام منصة HM16.20.

الكلمات المفتاحية: ضغط الفيديو ، HEVC/H.265، MDC، EbNo، HM 16.20، AWGN .

Abstract:

Information technology is one of the most important areas of development because of the high demand of users, which are now part of our daily lives and have become essential to reduce the volume of media (audio, video and video), also called (Multimedia).

Video compression is a demanding application where large amounts of data must be processed. In this study, we evaluated the new video coding standard (HEVC [High Efficiency Video Coding] / H.265) with certain variables, in particular video content, signal to noise ratio (SNR), The main issues that appear are packet loss, delay, jitter, bandwidth, Multiple description coding (MDC) has been proposed for use in packet audio and video transmission systems as a means of combating the issues, in a variety of application scenarios the experiments were conducted using the HM16.20 platform.

Key words : video compression, H.265/ HEVC ,EbN0 ,MDC ,HM 16.20, AWGN.

Résume:

Les technologies de l'information constituent l'un des domaines de développement les plus importants en raison de la forte demande des utilisateurs, qui font désormais partie de notre quotidien et qui sont devenues indispensables pour réduire ou comprimer le volume des médias (audio, vidéo et vidéo), également appelé (MultiMedia).

La compression vidéo est une application exigeante où de grandes quantités de données doivent être traitées. Dans cette étude, nous avons évalué le nouveau standard de codage vidéo (HEVC [High EfficiencyVideoCoding] / H.265) avec certaines variables, notamment le contenu vidéo, le rapport signal sur bruit (SNR), Les principaux problèmes qui apparaissent sont la perte de paquets, retard, gigue, bande passante, codage de description multiple (MDC) a été proposé pour une utilisation dans les systèmes de transmission audio et vidéo par paquets comme un moyen de lutter contre les problèmes, dans une variété de scénarios d'application, les expériences ont été menées à l'aide de la plate-forme HM16.20.

Mots clés : Compression de la vidéo, H.265/ HEVC, EbN0 , MDC, HM 16.20, AWGN.

Tables list

Figures list

Abbreviation List

A

AVC: Advanced video coding

AP: Access point

ARQ: Automatic Repeat Request

AWGN: Additive White Gaussian Noise

AI: All intra

B

BMP: Bitmap

B-frame: Bidirectional predictive frame

BSS: Basic service set

BER: Bit error

C

CRT: Cathode ray tube

CSF: Contrast Sensitivity Function

D

DCT: Discrete cosine transform

DiffServ: Differentiated services

DSCP: Differentiated Services Code Point

dB: Decibel

E

ESS: Extend service set

ERC: Error Resilient Coding

F

FPS: frame per second

FEC: Forward Error Correction

G

GOP: Group of pictures

H

HDTV: High definition television

HVS: Human Visual System

HEVC: High Efficiency Video Coding

I

I-frame: Intra coded frame

IntServ: Integrative services

ISO: International Organization for Standardization

IBSS: Independent basic service set

IT: Information technology

IPv4: Internet Protocol version 4

ITU: International telecommunication union

J

JPEG: Joint photographic experts group

JVT : Joint video team

L

LD: Low delay

M

Mbps: Mega bit per second

MPEG: Moving Picture Experts Group

MAC: Medium access control

MOS: Mean opinion score

MDC: Multiple description coding

MDSQ: Multiple Description Scalar Quantizer

MSE: Mean squared error

N

NTSC: National Television Systems Committee

O

OSI: Open Systems Interconnection

P

PAL: Phase Alternate Line

P-frame: Predictive frame

PSNR: Peak Signal to Noise Ratio

PAR: Positive Acknowledgement with Retransmission

PSS: Poly-phase Spatial Sub-sampling

Q

QOS: Quality of service

QOE: Quality of experience

R

RGB: Red green bleu

RF: Radio frequency

RTP: Real time transport protocol

RSVP: Resource Reservation Protocol

RA: Random access

S

SECAM: Sequential Color with Memory

SSID: Service set identifier

SSIM: Structural similarity

SNR: Signal to noise ratio

SI: Spatial information

T

TC: Traffic class

TI: Temporal information

V

VSNR: Visual signal to noise ratio

VIF: Variance inflation factor

VQMT: Video quality measurement tool.

VIFP: Visual Information Fidelity, pixel domain version

W

WLAN: Wireless Local Area Network

WMAN: Wireless Metropolitan Area Networks

WWAN: Wireless Wide Area Network

WPAN: Wireless Personal Area Network

Y

Y: luminance

Summary

General Introduction

GENERAL INTRODUCTION

Wireless communication technology has developed significantly over the past few decades and has become one of the most important types of media transmission from one device to another. Without the use of wires or electronic conductors, wireless communication allows for information to be transmitted by using electromagnetic waves.

Wireless networks can be classified into two categories: networks with pre-existing fixed infrastructure, and networks without infrastructure. In the first category, an important logistical and material infrastructure is necessary for the deployment of the network.

One inherent problem with any communications system is that information may be altered or lost during transmission due to channel noise. The effect of such information loss can be devastating for the transport of compressed video because any damage to the compressed bit stream may lead to objectionable visual distortion at the decoder. In addition, real-time/interactivity requirements exclude the deployment of some well-known error-recovery techniques for certain applications.

HEVC (High Efficiency Video Coding), also standardized under the name H.265, this development takes place in a context of growing diversity of digital audiovisual services, as well as the growing popularity of HD (High Definition) video. and UHD (Ultra High Definition) "4K" or soon "8K". As well as the transmission capacities required by video services, are a big challenge for today's digital networks. Finally, the desire for greater video quality.

This transmission requires well-defined properties of service, such as jitter, packet loss, therefore, Quality of Service (QoS) has become a current and future development direction.

Our job is to study and analyze the impact of video and its effects by integrating different values of error ratio by transmission through AWGN channel, i.e. measuring the impact on videos by PSNR (Peak Signal to Noise Ratio).

Our brief is structured around four chapters:

Chapter one presents the general scheme of the various video compression techniques and some of its characteristics and the difference between these techniques. Video coding is the process of compressing and decompressing a digital video signal.

Chapter tow presents wireless network and its controlling tools QoS ans QoE.

Chapter three presents a description of, Multiple description coding which has been widely explored in the field of image and video coding.

Chapter four presents a detailed description of our simulation, and presents and the results obtained from experiment.

Our method has been proposed in order to provide sufficient high compression ratios with good performance, in terms of quality.

Finally, this dissertation ends with a general conclusion summarizing the fundamental ideas that this work has brought us.

Our goal in this paper was to review MDC and provide a taxonomy and analysis to aid practitioners and researchers for better understanding of MDC.

CHAPTER I:

Video compression

Introduction

Video is a rapidly developing field, especially with the progress made in video coding techniques. This progress has led to a high number of video applications, Here we are going to put more focus on Digital video communication such as High-Definition Television (HDTV), videoconferencing and real-time video transmission over multimedia. Due to the advent of multimedia computing, the demand for these video has increased, their storage and manipulation in their raw form is very expensive and it significantly increases the transmission time and makes storage costly. When an ordinary analog video sequence is digitized, it can consume up to 165 Mbps. With most surveillance applications infrequently having to share the network with other data intensive applications and data transfer of uncompressed video over digital networks requires very high bandwidth. To circumvent this problem, a series of techniques called video compression techniques have been derived to reduce the number of bits required to represent a digital Video data while maintaining an acceptable fidelity or Video quality. Their ability to perform this task is quantified by the compression ratio[1]. The higher the compression ratio is the smaller the bandwidth consumption is.

1-Image:

1-1-Introduction:

Since the dawn of time, humans used drawings to communicate and express their needs, and for ages, photos pictures and drawings were as equal as words at describing and communicating.

Images (including pictures, photos, animations, drawings, and even documents) can be obtained by using different observing and capturing systems from the real world in various forms and manners. They can act directly and/or indirectly, on human eyes and produce visual perception. The human visual system is a typical example of an observation system, from which humans perceive and recognize objects and scenes by observing and remembering the images.

The image contains a rich description of the objects it expresses, which is our most important source of information.

1-2-Image definition:

The real world is three-dimensional (3-D) in space. However, the images obtained from the real word are generally two dimensional (2D). Such a 2-D image, in most cases, is a 2-D light intensity function (2-D light pattern). An image can be represented by a 2D array f(x,y), where x and y are spatial coordinates and f is the amplitude at any pair of coordinates (x, y) representing certain properties of the scene projected on the image at that particular point. For example, in an image recording the brightness of scene, the amplitude f is proportional to the intensity of the image.

In general, the amplitude f is proportional to one or several attributes of the image. For a real image, the values of x and y as well as f are real values, which are limited to a certain range.

Mathematically, an image can be explained as a function f(x,y) with two variables. For the purpose of processing by computers, an analog Image f(x, y) should be digitalized to a digital image (r.c), in which r (row) and c (column) refer to the discrete position of any point in the digital image and the amplitude / refers to the discrete magnitude of the image at point (r.c). Since our main focus is to discusses digital Images, f(x,y) is used to represent a digital Image at any point (x,y) and f takes Integer values only. A digital image can be considered a matrix whose rows and columns refer to the position of any point in the image and the corresponding matrix value refers to the intensity at that point.

In the early days, an image was called a "picture”. When a picture is digitized, a sampling process is used to extract from the picture a discrete set of real numbers. The picture samples are usually quantized to a set of discrete gray level values, which are often considered to be equally spaced. The result of sampling and quantizing is a digital picture. It is assumed that a digital picture is a rectangular array of integer values. An element of a digital picture is called a picture element (often abbreviated as "pixel" or "pel"). Although nowadays the term “image” rather than picture is used, because computers store numerical images of a picture or scene, the element of an image is still called a pixel [2].

1-3-Pixel:

In a digital image, all the coordinates on 2-d function and the corresponding values are finite. Each value available in every location is considered as a pixel. In other words, a pixel is the smallest part of an image. So a digital image can be thought as 2-d array of pixels.

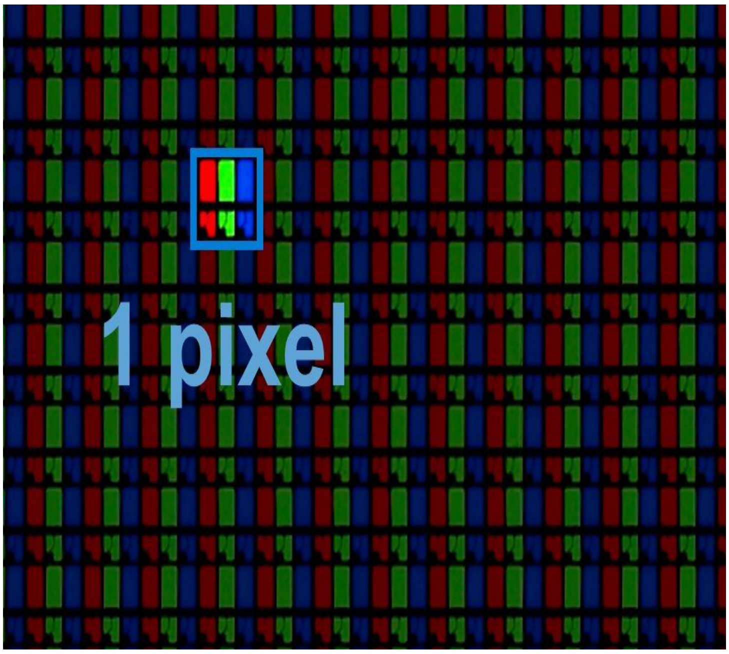


Figure I.1: pixels in a colorful image

1-4-Types of Images:

1-4-1-Binary Image: A binary image has only two possible values or intensities 0 and 1, there are no intermediate values. Binary images are used as masks for indicating the pixels of interest in many image processing tasks. Below is the example of binary image.

Figure I.2: Binary Image (only 0 and 255)

1-4-2-Grayscale Image: Grayscale image has range of values from 0 to 255, each pixel location can have any value between 0 and 255. If you watch old films around the 1950s, you are watching grayscale images, Here is the example below.

Figure I.3: Grayscale image (0–255 range)

1-4-3-Color Image: Both binary image and grayscale image are 2-dimensional arrays, where at every location you have one value to represent the pixel. Remember to represent a color image we need more than one value for each pixel. But how many values we need to represent a color? Typically you need 3 values for each pixel to represent any color. This came from the idea that any color can be formed by combining 3 basic colors Red, Blue and Green. Example: you get yellow by mixing red and green, violet can be formed by combining red and blue, etc. This is actually called RGB color space. There are many other ways to create color images which we will discuss in future discussions. Below is an example of a color image [3].

Figure I.4: Color Image (3-d array with 3 values at every pixel location)

1-5-Image compressing:

Image compression is the process of encoding or converting an image file in such a way that it consumes less space than the original file.

It is a type of compression technique that reduces the size of an image file without affecting or degrading its quality to a greater extent.

1-5-1-How does image compressing work:

There are two kinds of image compression methods:Lossy vs Lossless. Let's take a quick look at them both:

- **Lossless Compression:**

Lossless compression is a reversible form of compression in which there is no resulting loss of data to the code-stream. That means that the image recovered from the compressed data can be reproduced identically (bit per bit) just as it existed before compression was applied. There can also be visually lossless compression, which is a form of compression where the image cannot be exactly recovered. Images are mathematically distinguishable but not visually.

- **Lossy compression:**

Lossy Compression, oppositely to lossless, is an encoded method that compresses data discarding some of it. The advantage of lossy methods over lossless is that normally it is possible to obtain much smaller compressed files while still meeting the requirements of the application. The disadvantage is that it is not possible to recover the original data after a lossy compression.

Here it is a table showing a resume of the main differences between the two techniques:

Lossly compression	Lossless compression
1-The technique involves some loss of information.	1-involves no loss of information.
2-Data that has been compressed using this technique can't be recovered and reconstructed exactly.	2-If data has been (lossless) compressed, the original data can be recovered from the compressed data.
3-Used for application that can tolerate difference between the original and reconstructed data.	3-Used for application that can't tolerate any difference between original and reconstructed data.

4-In return for accepting this distortion in reconstructed data we obtain high compression rate.	4-No loss in information so compression rate is small.
5-Sound and image compression uses lossy compression.	5-Text compression uses lossless compression.

Table I.1: difference between lossy and lossless compression

1-5-2-RGB and YUV:

Medical researches proved that the human eye has different sensitivity to color and brightness, so in image compressing to have a much smaller size of an image we used conversion from color to brightness, RGB stands for red green and blue, this three colors are the three matrix combined that makes a color image, and for YUV, Y: luminance, UV: chrominance, Luminance is very similar to the grayscale version of the original image. The formulas for converting from RGB to YUV are given below.

$Y = (77/256)R + (150/256)G + (29/256)B$
$U = -(44/256)R - (87/256)G + (131/256)B + 128$
$V = (131/256)R - (110/256)G - (21/256)B + 128$

It is a known fact in image processing that most of the information that the human eye perceives is in luminance component Y and not in chrominance components UV, so generally we down sample UV components and still will be able to retain most of the information of the original image. So, RGB is the format in which data is actually produced by cameras, and YUV is the format used for algorithm processing, transmission and storage.

1-5-3-Transform coding:

Transform coding has been widely used to redundancy between data samples. In transform coding, a set of data samples is first linearly transformed into a set of transform coefficients. These coefficients are then quantized and entropy coded. A proper linear transform can de-correlate the input samples, and hence remove the redundancy. Another way to look at this is that a properly chosen transform can concentrate the energy of input samples into a small number of transform coefficients, so that the resulting coefficients are easier to encode than the original samples. The most commonly used transform for image compressing is the discrete cosine transform. The DCT is a unitary transform, that is, the transformation preserves the energy of the signal. Unitary transforms pack a large portion of the energy of the image into relatively few components of the transform coefficients. When the transform is applied to a block of pixels that are highly correlated, as in the case in a block of an image, the transform coefficients tend to be uncorrelated. Block processing yields good results when he bits allocated to encoding the frame is enough to guarantee a good reconstruction in the decoder [4].

1-5-4-Sampling and quantization:

In order to become suitable for digital processing, an image function f(x,y) must be digitized both spatially and in amplitude. Typically, a frame grabber or digitizer is used to sample and quantize the analogue video signal. Hence in order to create an image which is digital, we need to covert continuous data into digital form. There are two steps in which it is done:

- Sampling
- Quantization

The sampling rate determines the spatial resolution of the digitized image, while the quantization level determines the number of grey levels in the digitized image. A magnitude of the sampled image is expressed as a digital value in image processing. The transition between continuous values of the image function and its digital equivalent is called quantization.

The number of quantization levels should be high enough for human perception of fine shading details in the image. The occurrence of false contours is the main problem in image which has been quantized with insufficient brightness levels.

2-Video:

2-1-Definition of video:

A video is nothing more than a series of images in a sequence being flashed off a screen one at a time at such a speed that gives you the illusion of movement. So, by posting more than 20 frames per second, it is possible to deceive the human eye and does it look like a moving picture. We describe the fluidity of a video by the number of images per seconds (frame rate), expressed in FPS (frames per second) [5].

We have two types of video: analog video and digital video.

2-1-1-Analog video:

Analog video is created in a video camera by scanning an electron beam across a phosphor. The beam intensity is determined by the amount of light on each small area of the phosphor, which itself responds to the light being focused on it by a lens. That beam is then transmitted to a recording, switching, or display device. Analog switchers simply make a connection between devices by closing a relay dry contact. Recorders simply record the voltage changes of the electron beam onto tape, and display devices convert the voltage back into an electron beam and aim it at another phosphorus surface, which is the display monitor that is viewed.

The most common consumer display mechanism for video still uses analogue display devices such as CRT (Cathode ray tube). Until all terrestrial and satellite broadcasts become digital, analogue video formats will remain significant. The three principal Analogue Video Signal formats are: NTSC (National Television Systems Committee), PAL (Phase Alternate Line) and SECAM (Sequential Color with Memory). All the three are television video formats in which the information in each picture is captured by CRT is scanned from left to right to create a sequential intensity signal. The formats take advantage of the persistence of human vision by using interlaced scanning pattern in which the odd and even lines of each picture are read out in two separate scans of the odd and even fields respectively. This allows good reproduction of movement in the scene at the relatively low field rate of 50 fields /sec for PAL and SECAM and 60 fields /sec for NTSC.

2-1-2-Digital video:

In a digital video, the picture information is digitized both spatially and temporally and the resultant pixel intensities are quantized. The block diagram depicting the process of obtaining digital video from continuous natural scene is shown in figure I.5:

Figure I.5: digital video from natural scene

Digital video is composed of a series of orthogonal bitmap (BMP) images displayed in constant rapid succession with common frequencies of 15, 24, 30 and 60 frames per second (FPS), the more frames the digital video has, the more movement details are captured or displayed.

Digital video is captured and stored in a digital format as ones and zeros, rather than a series of still pictures captured in film. Digital, versus analog, signals are used. Information is processed and stored as a sequence of digital data for easy manipulation by computers, but the video is still presented to the viewer through a screen in analog form.

- **Spatial sampling:**

The sensitivity of Human Visual System (HVS) varies according to the spatial frequency of an image. In the digital representation of the image, the value of each pixel needs to be quantized using some finite precision. In practice, 8 bits are used per luminance sample.

- **Temporal sampling:**

A video consists of a sequence of images, displayed in rapid succession, to give an illusion of continuous motion. If the time gap between successive frames is too large, the viewer will observe jerky motion. The sensitivity of HVS drops off significantly at high frame rates. In practice, most video formats use temporal sampling rates of 24 frames per second and above [6].

- **Bitmap:**

Bitmap (BMP) is an image file format that can be used to create and store computer graphics. A bitmap file displays a small dots in a pattern that, when viewed from afar, creates an overall image. A bitmap image is a grid made of rows and columns where a specific cell is given a value that fills it in or leaves it blank, thus creating an image out of the data.

2-2-Video frames:

Three types of video frames are I-frame, P-frame and B-frame. ‘I’ stands for Intra coded frame, ‘P’ stands for Predictive frame and ‘B’ stands for Bidirectional predictive frame. ‘I’ frames are encoded without any motion compensation and are used as a reference for future predicted ‘P’ and ‘B’ type frames. ‘I’ frames however require a relatively large number of bits for encoding. ‘P’ frames are encoded using motion compensated prediction from a reference frame which can be either ‘I’ or ‘P’ frame. ‘P’ frames are more efficient in terms of number of bits required compared to ‘I’ frames, but still require more bits than ‘B’ frames. ‘B’ frames require the lowest number of bits compared to both ‘I’ and ‘P’ frames but incur computational complexity. Frames between two successive ‘I’ frames, including the leading ‘I’ frame, are collectively called as group of pictures (GOP). The GOP is illustrated in figure I.6. The illustrated figure has one ‘I’ frame, two ‘P’ frames and six ‘B’ frames. Typically, multiple ‘B’ frames are inserted between two consecutive ‘P’ or between ‘I’ and ‘P’ frames. The existence of GOPs facilitates the implementation of features such as random access, fast forward or fast and normal reverse playback.

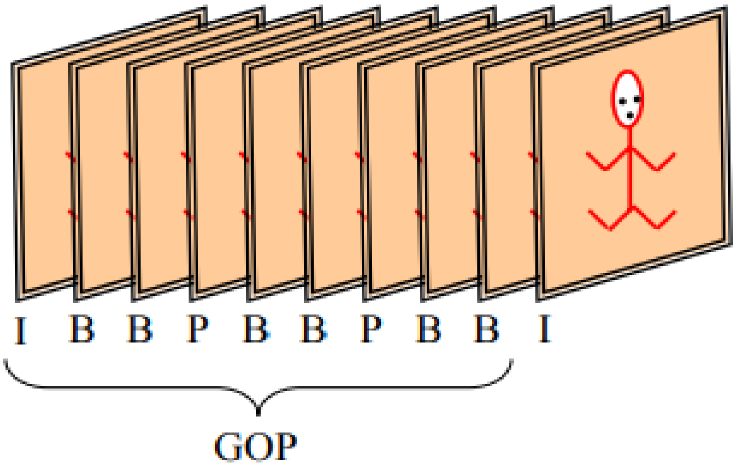


Figure I.6 : GOP (Group of pictures)

2-3-the reason behind compressing a video:

Nowadays it seems like we compress videos not just for having less space while storing it, in fact we have two more reasons for video compressing, The first, especially as it relates to streaming, is it makes it easier to transmit video over the Internet. This is because compression reduces the bandwidth required, while at the same time giving a quality experience. Without compression, raw video content would exclude many from being able to stream content over the Internet due to normal connection speeds not being adequate. The important aspect is bit rate, or the amount of data per second in the video.

The second reason for video encoding is compatibility. In fact sometimes content is already compressed to an adequate size but still needs to be encoded for compatibility, although this is often and more accurately described as transcoding. Being compatible can relate to certain services or programs, which require certain encoding specifications. It can also include increasing compatibility for playback with audiences [7].

3-Video compression:

Video compression plays an important role in many digital video applications such as digital libraries, video on demand, and high definition television. A video sequence with frame size of 176 X 144 pixels at 30 frames per second and 24 bits per pixel would require 18.25 Mbps, making it impractical to transmit the video sequence to transmit over standard telephone lines where data rates are typically restricted to 56,000 bits per second. This example illustrates the need for video compression. Effective video compression can be achieved by minimizing both spatial and temporal redundancy. A video sequence consists of a series of frames. In order to compress the video for efficient storage and transmission, the temporal redundancy among adjacent frames must be exploited. Temporal redundancy implies that adjacent frames are similar whereas spatial redundancy implies that neighboring pixels are similar. Video coding translates video sequences into an efficient bitstream. This translation involves the removal of redundant information from video sequence. Video sequence contains two kinds of redundancies spatial and temporal. Removal of spatial redundancy is generally termed as interframe coding and removal of temporal redundancy is termed as interframe coding. Video compression algorithms can be broadly classified as we said earlier into two types (i) Lossless video compression and (ii) Lossy video compression. Due to its importance in multimedia applications, most of the algorithms in video compression has centered on lossy video compression. Lossless video compression is important to applications in Luminance sample (Y) Chrominance sample (U, V) which the video quality cannot tolerate any degradation such as archiving of a video, compression of medical and satellite videos etc [8].

So the main basics of compression are:

- Reduce color nuances within the image.

- Reduce the color resolution with respect to the prevailing light intensity.

- Remove small, invisible parts, of the picture.

- Compare adjacent images and remove details that are unchanged between two images.

To make it simpler we resume video (image) compression in the figure I.7 below:

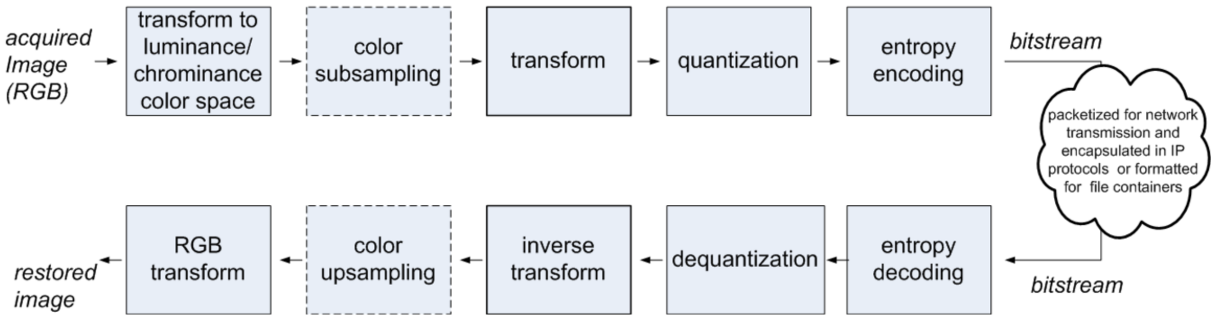


Figure I.7 : image compression steps

- **Inter frame:**

An Inter frame is a frame which is expressed in terms of one or more reference frames. Inter prediction takes advantage of the temporal redundancy between neighbor frames to achieve better compression. There are two kinds of Inter frames, depending on the reference frame they are allowed to use, P and B frames. P-frames use earlier pictures (mainly I-frames) and B-frames can predict or from earlier and 7 or later frames.

- **Intra frame:**

An Intra frame (I-frame) is a frame which uses only information contained in the current frame, so no temporal procession is performed outside the current frame. (I-frames) are the least compressible (compared to P-frames and B-frames) but do not require other video frames to decode themselves.

3-1-Video codecs:

Video codecs are video compression standards done through software or hardware applications. Each codec is comprised of an encoder, to compress the video, and a decoder, to recreate an approximate of the video for playback. The name codec actually comes from a merging of these two concepts into a single word: ENCOder and DECoder.

3-1-1-JPEG:

JPEG The JPEG standard, ISO/IEC 10918, is the single most widespread picture compression format of today. It offers the flexibility to either select high picture quality with fairly high compression ratio or to get a very high compression ratio at the expense of a reasonable lower picture quality. Systems, such as cameras and viewers, can be made inexpensive due to the low complexity of the technique. The artifacts show the “blockiness” as seen in Figure I.8. The blockiness appears when the compression ratio is pushed too high. In normal use, a JPEG compressed picture shows no visual difference to the original uncompressed picture. JPEG image compression contains a series of advanced techniques. The main one that does the real image compression is the Discrete Cosine Transform (DCT) followed by a quantization that removes the redundant information (the “invisible” parts).

Figure I.8 : Original image (left) and JPEG compressed picture (right).

3-1-1-A/ Motion JPEG:

A digital video sequence can be represented as a series of JPEG pictures. The advantages are the same as with single still JPEG pictures – flexibility both in terms of quality and compression ratio. The main disadvantage of Motion JPEG (a.k.a. MJPEG) is that since it uses only a series of still pictures it makes no use of video compression techniques. The result is a lower compression ratio for video sequences compared to “real” video compression techniques like MPEG. The benefit is its robustness with no dependency between the frames, which means that, for example, even if one frame is dropped during transfer, the rest of the video will be un-affected.

3-1-1-B/ JPEG 2000:

JPEG 2000 was created as the follow-up to the successful JPEG compression, with better compression ratios. The basis was to incorporate new advances in picture compression research into an international standard. Instead of the DCT transformation, JPEG 2000, ISO/IEC 15444, uses the Wavelet transformation.

The advantage of JPEG 2000 is that the blockiness of JPEG is removed, but replaced with a more overall fuzzy picture, as can be seen in Figure I.9.

Figure I.9: Original image (left) and JPEG 2000 compressed picture (right).

Whether this fuzziness of JPEG 2000 is preferred compared to the “blockiness” of JPEG is a matter of personal preference. Regardless, JPEG 2000 never took off for surveillance applications and is still not widely supported in web browsers either.

3-1-2-MPEG:

MPEG (Moving Picture Experts Group), It was developed specifically for motion images. There are five MPEG standards being used or in development (MPEG-1, MPEG-2, MPEG-4, MPEG-7, and MPEG-21). Each compression standard was designed with a specific application and bit rate in mind, although MPEG compression scales well with increased bit rates.

MPEG compression allows fast compression of video and audio, and decompression occurs in real time. MPEG is capable of decompressing data at a rate of 1.2 to 1.5M per second.

There are several MPEG standards:

- **MPEG-1**, developed in 1988, is a standard for the compression of video data and the associated audio channels (up to 2 channels for stereo listening). It allows the storage of videos at a rate of 1.5 Mbps.
- **MPEG-2**, a standard originally dedicated to digital television (*HDTV*) offering a high quality at a rate that may go up to 40 Mbps, and 5 surround audio channels. MPEG-2 moreover, allows identification and protection against ripping. This is the format used for DVD videos.
- **MPEG-4**, a standard intended to allow multi-media coding of data in the form of digital objects, in order to achieve greater interactivity, which makes it particularly suitable for the Web and for mobile peripheral devices.
- **MPEG-7**, a standard aimed at providing a standard representation of audio and visual data in order to allow the search for information in such data flows. This standard is thus also known as *Multimedia Content Description Interface*.
- **MPEG-21**, a standard still under development, whose goal is to provide a *framework* for all digital actors (producers, consumers) in order to standardize the management of these contents, as well as of the access rights, the copyrights.

In most video sequences, most of the scenes are fixed or change very little, this is what is called the temporal redundancy.

When only the lips of the actor move, it is almost only the pixels of the mouth that will be modified from one image to another; it is thus sufficient to just describe the change from one image to another. This is the main difference between MPEG (*Moving Pictures Experts Group*) and M-JPEG. However, this method will have much less impact on an action scene.

3-1-3-H.261:

was approved by ITU-T in early 1991. It was later revised in 1993 to include backward-compatible high resolution graphics transfer mode. It is a coding standard targeted to video conference and video telephony applications operating at bit rates between 64 Kbit/s and 2 M bit/s. This bit rate was chosen because of the availability of ISDN (Integrated Services Digital Network) transmission lines that could be allocated in multiples of 64 Kbit/s. The color space used by H.261 is YCbCr with 4:2:0 chrominance subsampling.

3-1-4-H.263:

H.263 standard is intended for video telecommunication. It was approved in early 1996. The key features of H.263 standard were variable block size compensation, overlapped block motion compensation. H.263 can achieve better video at 18-24 Kbps than H.261 at 64 Kbps and enable video phone over regular phone lines or wireless modem [9].

3-1-5-H.264:

H.264 is a well-known video compression standard for high-definition digital video. Also known as MPEG-4 Part 10 or Advanced Video Coding (MPEG-4 AVC), H.264 is defined as a block-oriented, compensation-based video compression standard that defines multiple profiles (tools) and levels (max bitrates and resolutions). This format supports 4K and up to 8K Ultra High-Definition.

A codec based on the H.264 standard compresses a digital video file (or stream) so that it only requires half of the storage space (or network bandwidth) of MPEG-2. Through this compression, the codec is able to maintain the same video quality despite using only half of the storage space.

The purpose of creating H.264/AVC was to pioneer a new digital video standard capable of delivering good video quality at a substantially lower bitrate than previous standards without overly complicating the design so that implementation remains practical, and relatively inexpensive to implement. The next generation of compression is known as H.265 or HEVC, offering another quantum leap in efficiency.

The H.264 standard is also flexible enough to be applied to a wide variety of applications, networks, and systems, including those with low and high bit rates, low and high resolution video, broadcast, storage, IP packet networks, many types of networks, such as internet, satellite, cable, and also ITU-T multimedia telephony systems. H.264 is widely adopted within many verticals and by a wide range of devices—from professional decoders, all the way down to browsers and mobile devices [10].

3-1-6-HEVC:

HEVC stands for 'High Efficiency Video Coding' and is also known as H.265. HEVC is a recently developed video compression standard pioneered by the Joint Collaborative Team on Video Coding (JCT-VC).

This HEVC compression standard was developed to double the compression efficiency of the previous standard, H.264/AVC, and at typical consumer video distribution bitrates, HEVC has been largely successful in doing so and is the supported coded for the emerging Ultra HD Blu-Ray format. However, compression efficiency results will always vary according to the type of content and the encoder settings.

The power of most video compression standards is derived from a technique called motion compensated prediction. This technique encodes blocks of pixels by referencing another area in the same frame (called intra-prediction) or in another frame (called inter-prediction). HEVC/H.265 is an improvement over the previous standard largely because it can define a larger range of blocks sizes. For example, the previous standard defines blocks up to 16×16 pixels whereas HEVC/H.265 can define blocks up to 64×64 pixels. Similar to H.264, the new protocol defines a number of profiles (tools) and levels (max bitrates and resolutions) [11].

Conclusion

In this chapter we explained video signal and its characteristics, now anyone should know the basics of encoding and why it's done. They should also walk away knowing how content is compressed while not hugely impacting perceived quality.

We also defined some video compression codecs that the digital world uses today to make it easy and simple.

CHAPTER II:

Wireless Communications

Introduction

Over the past few decades, mobile and wireless communications as well as Internet are the most profound and important technology in information technologies that rapidly grow and continuously change human life. On the other hand the sheer volume of applications and devices on the market, paired with the rise of social media, has led to a flood of network traffic, putting network performance in jeopardy. Lately there has been an observation by the providers that even though they try to offer the best quality of service levels possible there are still some other factors that influence the perceived quality by the customer and as a consequence even the user satisfaction with the service was influenced. There started the need for defining what are the other factors that influences the perceived quality and another concept arouse causing one of the biggest changes in the telecommunication industry.

1-Wireless networks:

Wireless networks are computer networks that are not connected by cables of any kind. The use of a wireless network enables enterprises to avoid the costly process of introducing cables into buildings or as a connection between different equipment locations. The basics of wireless systems are radio waves, an implementation that takes place at the physical level of network structure.

Wireless networks use radio waves to connect devices such as laptops to the Internet, the business network and applications. When laptops are connected to Wi-Fi hot spots in public places, the connection is established to that business's wireless network.

There are four main types of wireless networks:

- **Wireless Local Area Network (WLAN)**: Links two or more devices using a wireless distribution method, providing a connection through access points to the wider Internet.
- **Wireless Metropolitan Area Networks (WMAN)**: Connects several wireless LANs.
- **Wireless Wide Area Network (WWAN)**: Covers large areas such as neighboring towns and cities.
- **Wireless Personal Area Network (WPAN)**: Interconnects devices in a short span, generally within a person's reach [12].

Type	Range	Applications	Standards
Personal area network (PAN)	Within a reach of a person	Cable replacement for peripherals	Bluetooth, ZigBee
Local area network (LAN)	Within a building or a campus	Wireless extension of wired network	IEEE 802.11 (Wifi)
Metropolitan area network (MAN)	Within a city	Wireless inter-network connectivity	IEEE 802.16 (WIMAX)
Wide area network (WAN)	Worldwide	Wireless access network	Cellular (UMTS, LTE, ect)

Table II.1: types of wireless network. [13]

Wireless networks consist of four major physical components, which are summarized in figure II.1:

1-Stations:

Networks are built to transfer data between stations. Stations are computing devices with wireless network interfaces. Typically, stations are battery-operated laptop or handheld computers. There is no reason why stations must be portable computing devices, though. In some environments, wireless networking is used to avoid pulling new cable, and desktops are connected by wireless LANs. Large open areas may also benefit from wireless networking, such as a manufacturing floor using a wireless LAN to connect components.

2-Access points:

Frames on an 802.11 network must be converted to another type of frame for delivery to the rest of the world. Devices called access points perform the wireless-to-wired bridging function. (Access points perform a number of other functions, but bridging is by far the most important.) Initially, access point functions were put into standalone devices, though several newer products are dividing the 802.11 protocol between "thin" access points and AP controllers.

3-Wireless medium:

To move frames from station to station, the standard uses a wireless medium. Several different physical layers are defined, the architecture allows multiple physical layers to be developed to support the 802.11 MAC. Initially, two radio frequency (RP) physical layers and one infrared physical layer were standardized, though the RP layers have proven far more popular. Several additional RF layers have been standardized as well.

4-Distribution system:

When several access points are connected to form a large coverage area, they must communicate with each other to track the movements of mobile stations. The distribution system is the logical component of 802.11 used to forward frames to their destination.

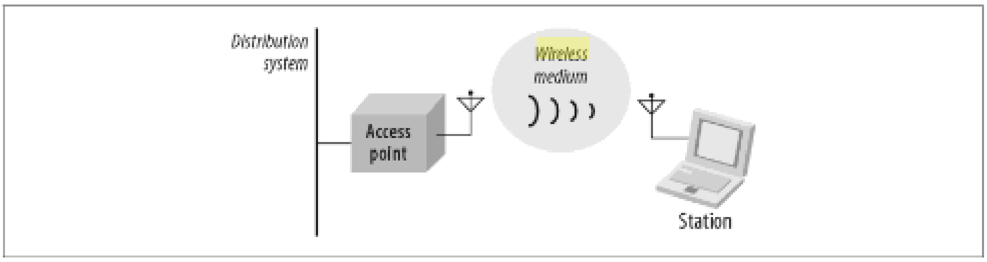


Figure II.1: components of 802.11 LANs

1-2-IEEE 802 wireless standards:

The IEEE 802 Standard comprises a family of networking standards that cover the physical layer specifications of technologies from Ethernet to wireless. IEEE 802 is subdivided into 22 parts that cover the physical and data-link aspects of networking. In the table below, we are just going to show you the IEEE 802 standards of the wireless network types.

802	Overview	Basics of physical and logical networking concepts
802.11	Wi-Fi	Wireless LAN Media Access Control and Physical Layer specification. 802.11a,b,g,etc. are amendments to the original 802.11 standard. Products that implement 802.11 standards must pass tests and are referred to as "Wi-Fi certified."
802.15	Wireless personal area network	Communications specification that was approved in early 2002 by the IEEE for wireless personal area networks (WPANs). Including 802.15.1 for the Bluetooth and 802.15.4 for ZigBee.
802.16	Wireless Metropolitan Area Networks	This family of standards covers Fixed and Mobile Broadband Wireless Access methods used to create Wireless Metropolitan Area Networks (WMANs.)

Table II.2: Some IEEE 802 standards. [14]

1-3-802.11 WLAN access mode:

The basic building block of a wireless network is the basic service set (BSS), which is simply a group of stations that communicate with each other. Communications take place within a somewhat fuzzy area, called the basic service area, defined by the propagation characteristics of the wireless medium.' When a station is in the basic service area, it can communicate with the other members of the BSS. BSSs come in two flavors, both of which are illustrated in Figure II.2.

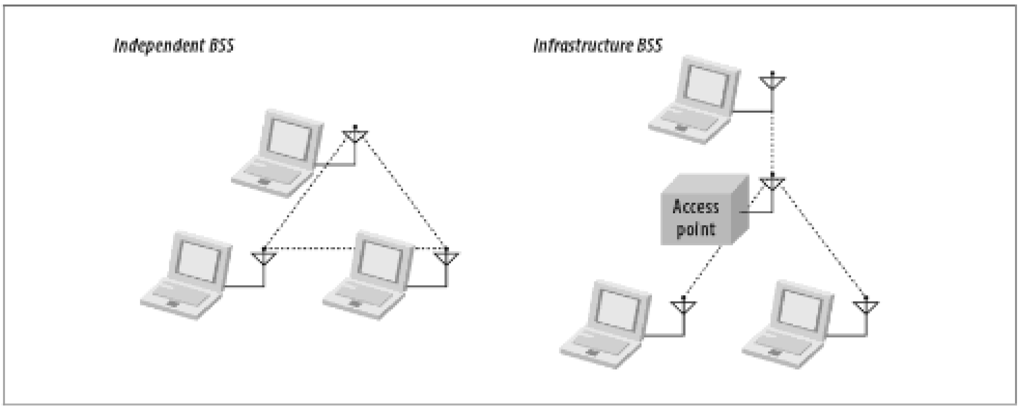


Figure II.2: Independent and infrastructure BSSs

1-Independent networks:

On the left is an independent BSS (IBSS). Stations in an IBSS communicate directly with each other and thus must be within direct communication range. The smallest possible wireless network is an IBSS with two stations. Typically, IBSSs are composed of a small number of stations set up for a specific purpose and for a short period of time. One common use is to create a short-lived network to support a single meeting in a conference room. As the meeting begins, the participants create an IBSS to share data. When the meeting ends, the IBSS is dissolved. Due to their short duration, small size, and focused purpose, IBSSs are sometimes referred to as ad hoc BSSs or ad hoc networks.

2-infrastructure networks:

On the right side of Figure is an infrastructure BSS, Infrastructure networks are distinguished by the use of an access point. Access points are used for all communications in infrastructure networks, including communication between mobile nodes in the same service area. If one mobile station in an infrastructure BSS needs to communicate with a second mobile station, the communication must take two hops. First, the originating mobile station transfers the frame to the access point. Second, the access point transfers the frame to the destination station. With all communications relayed through an access point, the basic service area corresponding to an infrastructure BSS is defined by the points in which transmissions from the access point can be received.

-Extended service set:

BSSs can create coverage in small offices and homes, but they cannot provide network coverage to larger areas. 802.11 allow wireless networks of arbitrarily large size to be created by linking BSSs into an extended service set (ESS). An ESS is created by chaining BSSs together with a backbone network. All the access points in an ESS are given the same service set identifier (SSID), which serves as a network "name" for the users [15].

2-Quality of Service:

Folded into most network monitoring tools is the ability to manage and monitor network traffic by a class of service methods. QoSis a set of technologies that work on a network to guarantee its ability to dependably run high-priority applications and traffic under limited network capacity. QoS technologies accomplish this by providing differentiated handling and capacity allocation to specific flows in network traffic. This enables the network administrator to assign the order in which packets are handled, and the amount of bandwidth afforded to that application or traffic flow. QoSis mainly referred to as the overall performance of the network but it is highly affected by parameters such as bandwidth (throughput), latency (delay), jitter (variance in latency), and error rate [16].

These QoS monitoring tools can empower system administrators to determine whether the QoS policies they have in place are effectively prioritizing traffic and providing a positive end-user experience.

2-1-QoS parameters:

- **Latency Reduction:** Network latency—any sort of delay in network system transactions—is an all-too-common occurrence for many IT technicians. If Real time Transport Protocol (RTP) packets, such as those for video conferencing, are left without QoS classification, they'll traverse the network unmarked and be treated as an ordinary piece of data. The absence of prioritization can lead to major consequences, especially in large networks prone to congestion, meaning videos and audio will be choppy, rendering them utterly ineffective for users on both ends. A system administrator's goal is to reduce latency as much as possible, which is why QoS in networking is needed.

- **Jitter Reduction:** Jitter refers to the irregular speed of packets as a result of deviations in signal pulses. Different factors can cause jitter, from electromagnetic interference to cross-talk with other signals. To the end user, these late and potentially out-of-sequence packets appear in the form of flickering monitors, blatant gaps in audio and video, and more. QoS drastically minimizes the occurrence of jitter.
- **Packet Loss Prevention:** I've yet to meet a system administrator who doesn't dread the occurrence of packet loss. Packet loss—or the failure of a packet to arrive at its destination—is the result of an onslaught of congestion within the networking device. When this happens, the router or switch simply disregards incoming packets until more space is available, resulting in broken images or unintelligible audio. QoS helps manage the flow of traffic by prioritizing packets and warding off potential traffic jams, so network devices don't have to drop packets entirely.
- **Bandwidth management:** QoS is used to classify and prioritize traffic throughout a network. QoS enables you to establish an end-to-end traffic priority policy to improve the control and throughput of important data. You can manage available bandwidth so that the most important traffic goes first.
- **Security Enhancements:** QoS is an integral part of secure network design. QoS mechanisms can stop traffic in its path, a powerful and useful capability if compromised traffic is about to enter the network.

2-2-QoS models:

Integrative services (IntServ) and **differentiated services** (DiffServ) models empower administrators to put QoS into play and manage their network traffic. IntServ homes in on a network's bandwidth and relies on the Resource Reservation Protocol (RSVP) to perform its duties. With IntServ, applications must request a resource reservation (such as bandwidth) for each data flow before sending data. Network devices then act as the "traffic cop" to determine what network resources are available and whether there are enough to accept the packet at hand. If accepted, the application data can flow as long as it remains within the initially requested traffic profile. To leverage IntServ, system admins must ensure they're using IntServ-capable routers. whileIntServ is an effective method of QoS policing, it consumes numerous network resources and thus is not recommended for companies looking to scale their operations.

DiffServ is probably the most common QoS model. With DiffServ, system administrators assign a Differentiated Services Code Point (DSCP) value, ranging from 0 to 63, to each network traffic type, classifying it at the Internet Protocol (IP) level based on a predetermined set of requirements and groupings known as Traffic Classes (TCs). This value can be placed in an IPv4 packet header as well as the TC field in IPv6. Applications with a high-priority DSCP value can travel the network without delay—they're the top "runners" the traffic cop is waving through. Unlike IntServ, DiffServ can easily be scaled, making it the go to choice for many system administrators.

3-Quality of Experience:

QoE is a user-oriented description of service quality. It refers to the "subjective" evaluation of the end users about the overall service quality and reflects the users' experiences during/ throughout the entire service utilizing. To dimension/evaluate the QoE of a service, the user's perceptions, such as “Good”, “Poor”, “Fair”, “Bad”, are often used. QoE is a “subjective” concept. Despite the objective reasons related to service delivering (transport network), end terminal installation, software configuration, etc., those subjective factors, such as the user’s expectations and experiences, can determine the user’s final (overall) perception.

QoE is sometimes a “long-term” concept. A user’s positive perception about a service often implies long-term good QoE performance covering the whole period of the entire service utilization. However, QoE is also a “short-term” concept. It reflects the user’s experiences, particularly negative perception over a short period. For instance, the bad experience about the speech quality even for a short period of a voice application is probably serious enough to let the user terminate the service utilization [17].

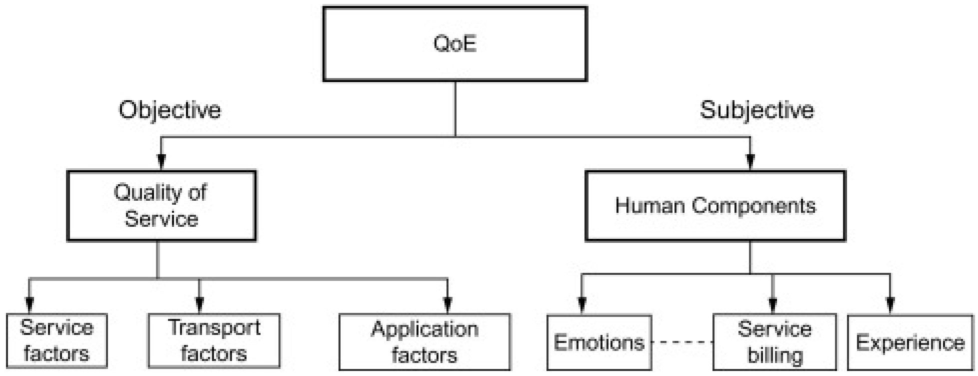


Figure II.3: Shows factors contributing to QoE

QoE often is referred as a correlation between Quality of Service QoS, which as we said determines the quality offered by an operator for a certain service, and the user perception towards a service .In figure II.3 shows that the service that the user get from the provider is just one part of the quality perception and the rest is defined by other important factors.

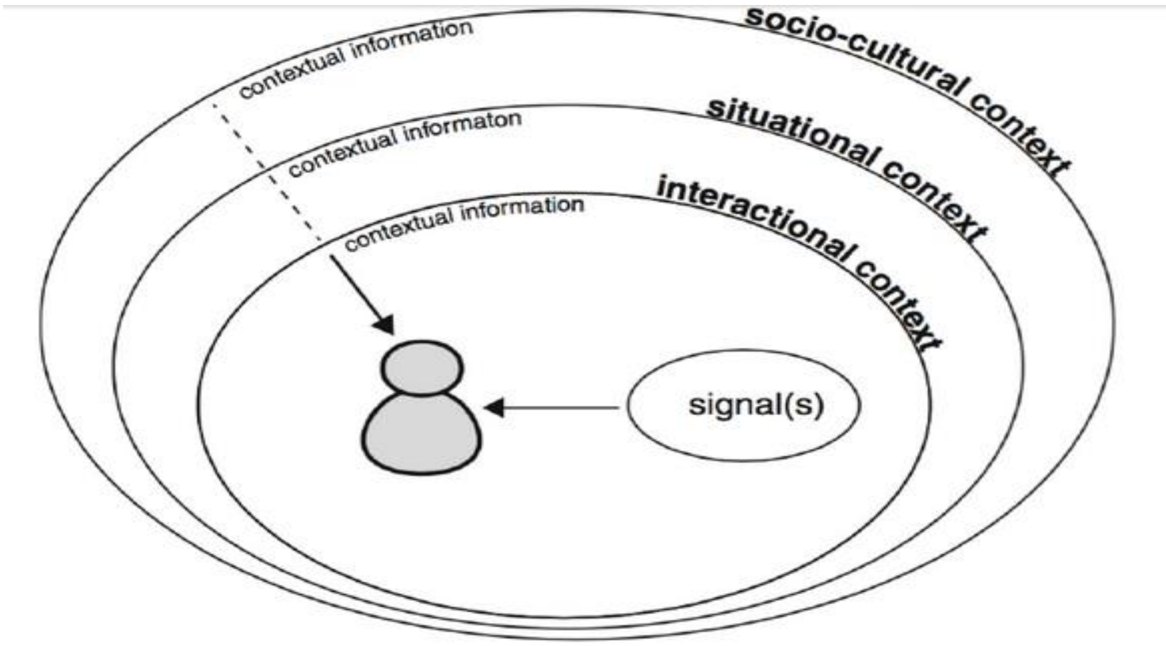


Figure II.4: Factors influencing user perception

As we can see from the figure II.4, there is an important factor named as signal, which represents what the user perceives and it is directly related to the actual content provided to the user but also with the quality that is delivered specifically by the network, which defines the QoS and it mainly includes bandwidth, dropped packet, errors, latency, jitter and out of order delivery. On the other hand we have other factors represented by the contextual parameters. Situational context refer to the environment in which the user is at the moment of the interaction with the service including physical and temporal context. Interactional context is directly related to the user that is interacting with another user through the service, mainly based on the intention of this interaction. While socio-cultural context is mostly related to the inter-personal relations existing between the user and another person that is presented on the user side at the moment of the experience.

These factors all together define the user experience with a service. There has been a lot of effort to also define Quality of Experience and one of the most valuable inputs comes by International Telecommunication Union (ITU) that defines Quality of Experience in its early stages of creation as “The overall acceptability of an application or service, as perceived subjectively by the end user”.

3-1-QoE Models:

Effective QoE management requires a combination of active and passive network traffic monitoring mixed with automation and real-time analysis.

First, IT teams should actively monitor traffic injected into the network that can test a large number of packets for key performance indicators. This should be combined with passive monitoring that analyzes every packet, looking at overall traffic characteristics.

This trove of network performance data then should be fed through platform that can analyze both passive and active data regardless of the source to detect for network behavior and troubleshoot performance variables on-the-fly. With this feedback loop, IT teams can respond to performance degradations that affect end user QoE.

So while QoS is gold standard for network performance management, it must be combined with QoE – and with good reason. The combination of QoS and QoE takes a much more holistic look at network performance and the end user, focusing on real world network outcomes. Bringing QoS and QoE together is truly the best of both worlds.

3-2-QoE Measurement approaches:

There are two main quality assessment methodologies, namely subjective and objective assessment. Measuring and ensuring good QoE of video applications is very subjective in nature. The most commonly used subjective method for quality measurement is the Mean Opinion Score (MOS). MOS is standardized in the ITU-T recommendations, and it is defined as a numeric value going from 1 to 5 (i.e. poor to excellent). The main drawbacks of this approach are: it is high in cost, time consuming, cannot be used in real time and lacks repeatability. These limitations have motivated the development of objective tools that predict subjective quality solely from physical characteristics. By definition, the objective approach is based on mathematical and/or comparative techniques that generate a quantitative measure of the one-way video quality. This approach is useful for in service quality monitoring or the design of networks/terminals, as well as in codec optimization and selection. An objective approach can be intrusive or non-intrusive. The intrusive methods are based on signals, while non-intrusive methods are based on network/application parameters. Generally, intrusive methods are accurate, but impracticable for monitoring live traffic because of the need for the original sequence, i.e. full reference quality measurement. Non-intrusive models do not require a copy of the original. Objective approaches usually ignore the content type and the way the content is perceived by Human Visual System (HVS). For example, some objective methods try to compare original and received signals pixel by pixel to detect signal distortions, such as Peak Signal to Noise Ratio (PSNR). A combination of the objective and subjective approaches can be performed to overcome the short comings of each individual technique. The PSNR-mapped-to-MOS technique is a commonly adopted method that is used to estimate video QoE which has been affected by network conditions. This technique was demonstrated by several researchers to be inaccurate in terms of the correlation to the perceived visual quality. However, several modifications have been proposed to enhance the estimation accuracy. Several communities have ratified the improved PSNR-mapped-to-MOS technique.

4-QoS and QoE layers:

Different solutions for QoS have been proposed at a variety of layers in the OSI seven layers Model. The two layers generally used for QoS are the application and network layers. The Application Layer includes services that are provided by the application in order to achieve the required QoS. Application layer QoS is concerned with parameters such as frame rate, resolution, colour, video and audio codec type, etc. On the other hand, network Layer services are provided by devices such as switches and routers. The network layer considers parameters such as delay, jitter, and packet loss, etc. Definitions from different authors suggest that a perceptual pseudo-layer can be imagined above both these two layers, which is concerned with the end-user’s experience (QoE) .Some authors consider this pseudo-layer as an extension to the Application layer, where as others view the QoE as an extension of traditional QoS because QoE provides information regarding the delivered services from the user’s viewpoint.

Conclusion

QoS is mainly oriented to analyze the performance of the network and the media involved. On the other side, Quality of Experience (QoE) examines the point of view of the end-user and how he perceives the value of the service. Though these two concepts are related, they have a substantial difference. In short, QoE is obtained from subjective tests, where human viewers evaluate the quality of tested videos under a laboratory environment, objective quality models are developed to predict QoE based on objective QoS parameters, but it is still an indirect way to estimate QoE.

CHAPTER III :

Multiple description coding

Introduction

With the explosion of the Internet, video transmission has become increasingly popular in the recent years and will continue to flourish in the future. However, network congestion and delay sensibility impose tremendous challenge on video communications. Due to network congestion, random bit errors and packet losses may cause substantial quality degradation of the compressed video sequence. In the case of real-time video application, delay sensibility has made the retransmission of corrupted data impossible. Therefore, this creates a need for coding approaches combining high compression efficiency and robustness. Multiple description coding (MDC) has emerged as an attractive framework for robust transmission over unreliable channels. It can effectively combat packet loss without any retransmission thus satisfying the demand of real time services and relieving the network congestion.

1. techniques used to counter the effect of data loss for video transmission:

Video transmission over noisy channels has been a challenging problem for more than two decades. First we have the very large bandwidth required for a video transmission. And second, compressed video is very sensitive to data loss which happens in best-effort networks such as the Internet. To counter the effect of data loss for video transmission over noisy networks, there are three categories of approaches: a reliable Automatic Repeat Request (ARQ), Forward Error Correction (FEC), and Error Resilient Coding (ERC). ARQ and FEC are channel level protections, while ERC can be used as a source level protection, such as Multiple Description Coding (MDC), in our chapter we are focusing on (MDC), put first we will understand this categories of protection [18].

1.1 Automatic repeat request (ARQ):

Automatic Repeat Request (ARQ) is a group of error – control protocols for transmission of data over noisy or unreliable communication network. These protocols reside in the Data Link Layer and in the Transport Layer of the OSI (Open Systems Interconnection) reference model. They are named so because they provide for automatic retransmission of frames that are corrupted or lost during transmission. ARQ is also called Positive Acknowledgement with Retransmission (PAR).

In ARQ the receiver sends an acknowledgement message back to the sender if it receives a frame correctly. If the sender does not receive the acknowledgement of a transmitted frame before a specified period of time, the sender understands that the frame has been corrupted or lost during transit. So, the sender retransmits the frame. This process is repeated until the correct frame is transmitted [19].

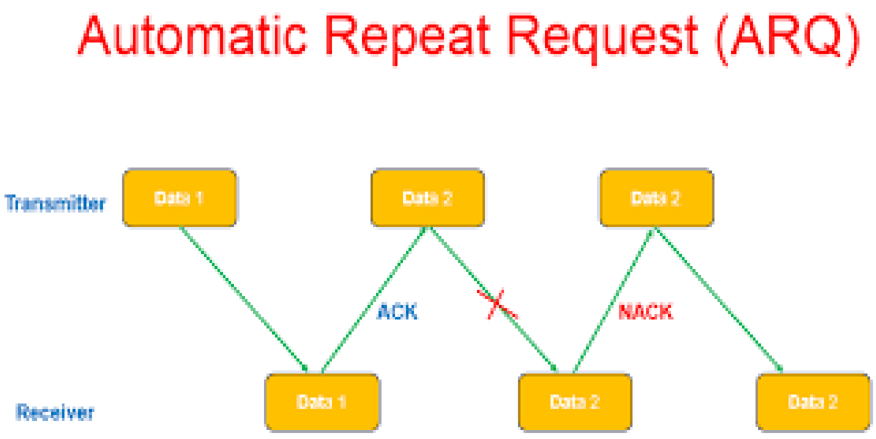


Figure III.1: ARQ back channels

1.2.Forward error correction (FEC):

Forward error correction (FEC) is an error correction technique to detect and correct a limited number of errors in transmitted data without the need for retransmission.

In this method, the sender sends a redundant error-correcting code along with the data frame. The receiver performs necessary checks based upon the additional redundant bits. If it finds that the data is free from errors, it executes error-correcting code that generates the actual frame. It then removes the redundant bits before passing the message to the upper layers [20].

1.3.Error resilient coding (ERC):

These coding options enable the decoder to conceal and regenerate the lost data from the received data, exploiting the correlation existed among blocks of the images [21].

ARQ and FEC are channel level protection, while ERC can be used as source level protection, such as:

- MDC (which is one of the most promising solutions for real-time video over lossy networks.

2. Multiple description code (MDC):

Multiple Description Coding (MDC) was originally used for speech communicating over circuit-switched network in the 1970's. Traditionally, to avoid communication interruption, an additional transmission link was on standby and would be activated in the case of the outage of the main link. This approach however was not cost efficient and therefore the idea of splitting the information over two channels: (speaking of MDC), was proposed.

In multiple description coding (MDC), independently-decodable and mutually-refundable streams of a video source are generated. The streams, also called descriptions, are then transmitted separately, possibly through different network paths. In MDC, as long as one or more descriptions arrive at the receiver, some video with certain quality can be displayed. If a packet is lost, the corresponding packets of the other descriptions, containing a different representation of the data in the lost packet, may be available and the video is decoded successfully but with a lower fidelity. This is depicted in figure III.3.

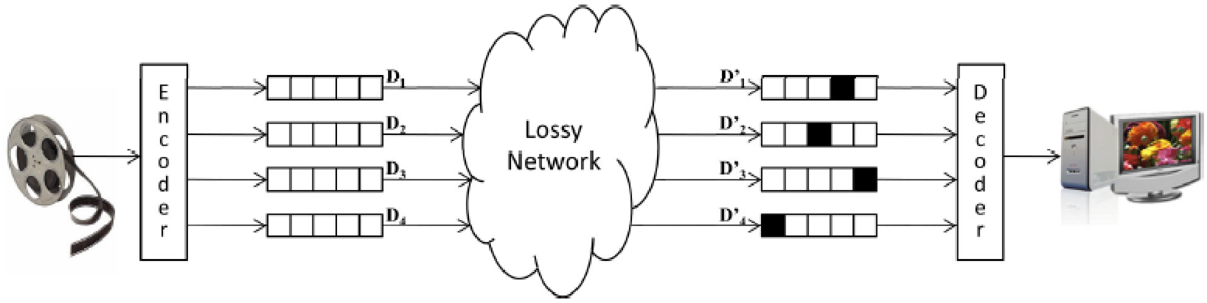


Figure III.2: The main rationale behind MDC video. Black boxes indicate lost information.

Figure III.3, shows the basic block diagram of MDC. This figure shows a two description case but a higher number of descriptions is possible. In the figure, a source is coded such that multiple complementary descriptions that are individually decodable are generated. After the descriptions are built, they can be transmitted separately, possibly through different network paths. At the receiver side, if only one description is available, it is decoded by the side decoder and the resulting quality (distortion) is called side quality (distortion). When both descriptions are available, they are decoded by the central decoder and the resulting quality (distortion) is called central distortion (quality). In central decoder the descriptions are merged and hence a video with higher quality is achieved. In other words, there exist two types of decoding at the receiver, when all descriptions are received the central decoding is used, and if one or more descriptions are not received the side decoder is used for the received description(s). Obviously, quality is enhanced by the number of received descriptions [22]. For robustness, it comes from the fact that it is unlikely that the same portion of the same picture is corrupted in all descriptions. The coding efficiency is reduced depending on the amount of redundancy left among descriptions [23].

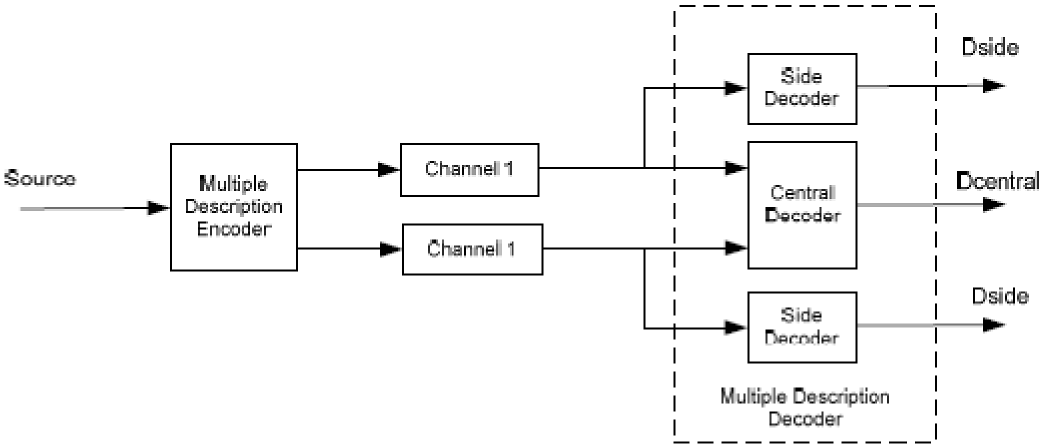


Figure III.3: block diagram of MDC

2.1. MDC types:

Our main focus is on video, even that these types were proposed and applied on images. It can also be applied on video. MDC types are based on their application domain, the domains are spatial, temporal, frequency, and compressed, in which the descriptions are generated by partitioning the pixels, the frames, the transformed data, and compressed data, respectively.

1-Spatial domain MDC:

In this category, the MDC process is carried out in the pixel domain. The simplest approach is to divide the image/frame into multiple sub-images and encode each one independently. Figure III.4, shows Poly-phase Spatial Sub-sampling (PSS) of a frame to generate four sub-images to achieve four-description coding. At the decoder side, if all descriptions are received, the sub-images are merged and the image in full resolution is reconstructed. Otherwise, any missed description must be recovered using interpolation or similar techniques.

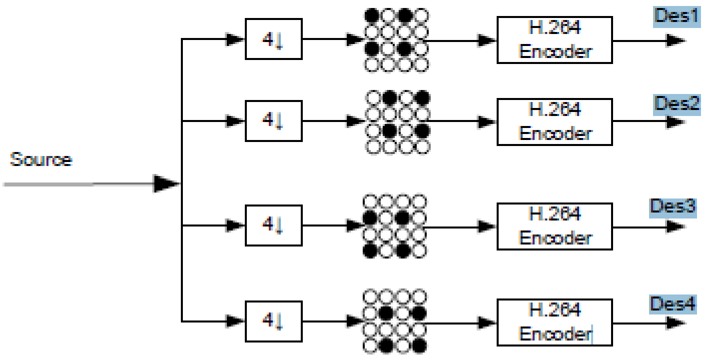


Figure III.4: spatial polyphase subsampling MDC

2-Temporel domain MDC:

In temporal domain MDC, the descriptions are generated by a process performed at the frame level. A simple case is frame splitting between descriptions: odd frames in one description and even frames in the other description, as shown in Figure III.5. At the side decoder, the lost frames are substituted by frame freezing or estimated by concealment methods. Motion estimation/compensation is performed intra description, meaning even (odd) frames are predicted from the even (odd) frames.

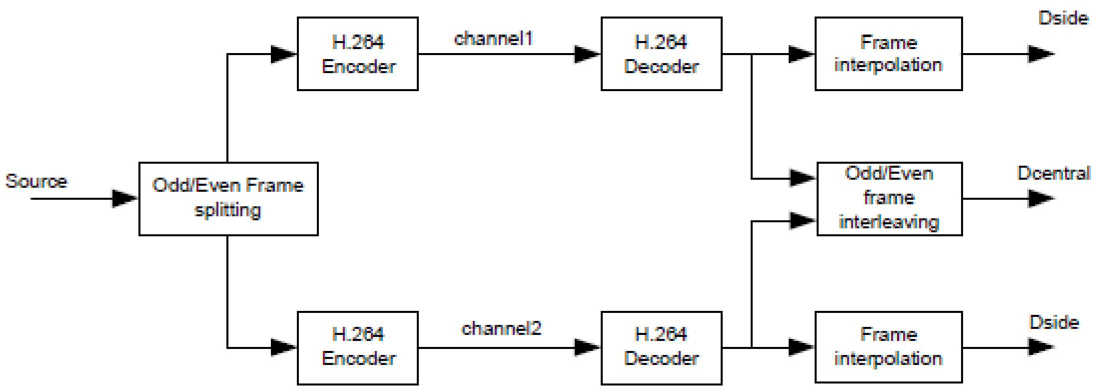


Figure III.5: Block diagram of temporal domain MDC

3-Frequency domain MDC:

Several approaches exist, like MDSQ (Multiple Description Scalar Quantizer). The concept of MDSQ is to use different quantization methods such that they refine each other at the central decoder. The simplest way of MDSQ is shifting the quantization intervals of each side encoder by half, as shown in figureIII.6:

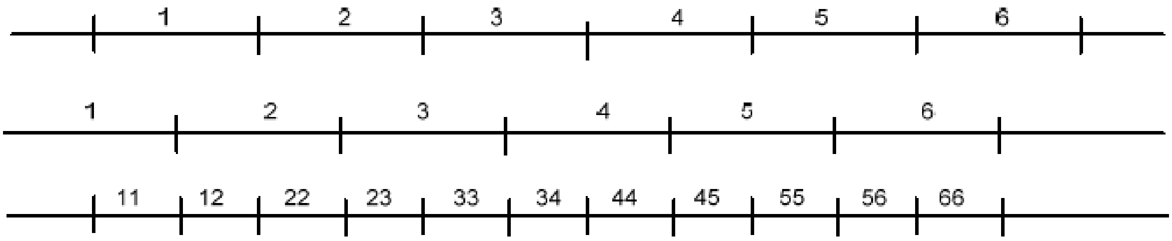


Figure III.6: The simplest MDSQ, upper two lines: side encoder (decoder) quantization (dequantization), lowest line: central decoder dequantization.

4-compressed domain MDC:

As the name implies, this MDC type is applied after the video has been encoded already. This can also be called packet domain MDC. The main idea here is to partition a layer into *K* segments which are then expanded to *N*>*K* segments using FEC codes. Then, these 2 segments are partitioned again into multiple descriptions which are sent independently. It is known that with any *K* out of these *N* segments, the *K* source segments are recoverable. As is shown in Figure III.7, the video or image is encoded into a scalable bit stream where layer (i) is partitioned into (i) segments. The descriptions are composed by collecting one and only one segment (source or FEC code) from each layer.

Generally, the number of segments for each layer determines the trade-off between loss resiliency and redundancy. The lower the number of segments in a layer, the higher the redundancy of that layer and the more likely for it to be reconstructed at the receiver.

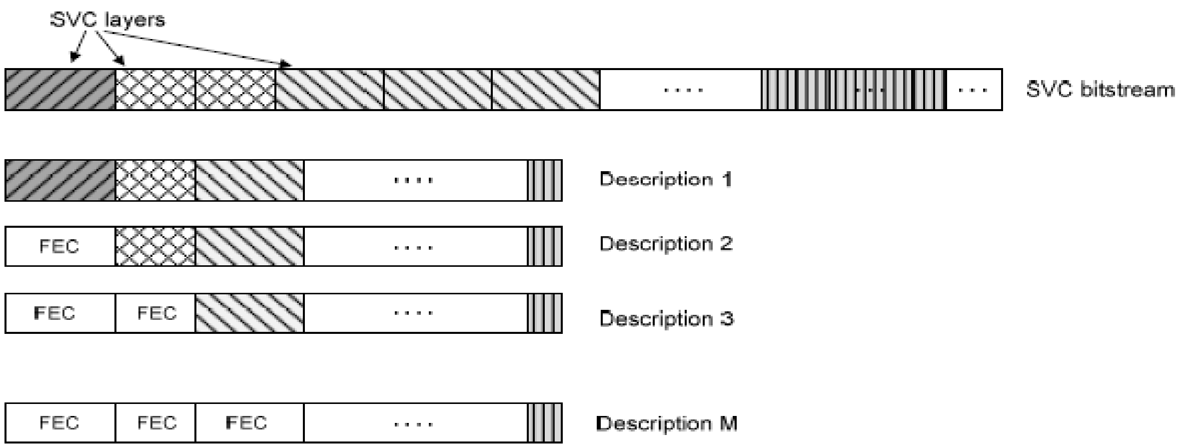


Figure III.7: MDC-FEC: with equal number of descriptions and SVC layers.

Conclusion

MDC is one of the promising solutions for real-time video streaming over lossy networks. The MDC process generating and dividing information into descriptions can be performed in several areas. The domains are spatial, temporal, frequency and compressed, in which the descriptions are generated by partitioning the pixels, frames, transformed data and compressed data respectively. Under each area, we present a number of categories of MDC.

CHAPTER IV :

Simulation and results

1 Introduction

Video Compression is the term used to define a method for reducing the data used to encode digital video content. This reduction in data translates to benefits such as smaller storage requirements and lower transmission bandwidth requirements, for a clip of video content. In this thesis we tried to compare multiples videos with and without MDC method by changing the values of their SNR but digital videos are subject to a wide variety of distortions during acquisition, compression storage, transmission and reproduction. Any of these may result in degradation of the video quality. To ensure an acceptable level of quality control, video metrics are applied to judge the quality. In reality subjective tests are carried out which require human experts to view the video and assess it for quality applying the Mean Opinion Score(MOS) weighting. This is time consuming and expensive. Objective tests can be carried out automatically, these objective quality metric systems predict visual quality by using information known about the human visual system to compare signals such as (PSNR, SSIM, MSE, VSNR,VIF……..).

2. FADING CHANNELS

The role of any communication system is to ensure that the receiver understands all the messages transmitted by the transmitter, regardless of the compression, the size or data type, but also the disturbances induced on the transmission channel and its parasitic effect.

Fading channel is a channel that fades due to obstacle in the transmission path. Multiple signals are received by the receiver because of the multipath component. The originally received signal is the vector sum of the individual signals and they get faded owing to various factors such as reflection, diffraction and scattering.

Fading is classified into large scale fading and small scale fading based on the distance. When the distance between the transmitter and receiver is long, then it is known as large scale fading channel. Similarly, when the distance between the transmitter and the receiver is short, it is called small scale fading that has three types namely:

- Additive White Gaussian Noise channel
- Rayleigh channel
- Rician channel.

2.1 Rayleigh Channel Model

Rayleigh fading channel model is used as there are many scatters between transmitter and receiver. It is used in large areas where there is no line of sight and many buildings and other obstacles reflect, diffract and attenuate the signal between the transmitter and the receiver.as its shown in figure .1

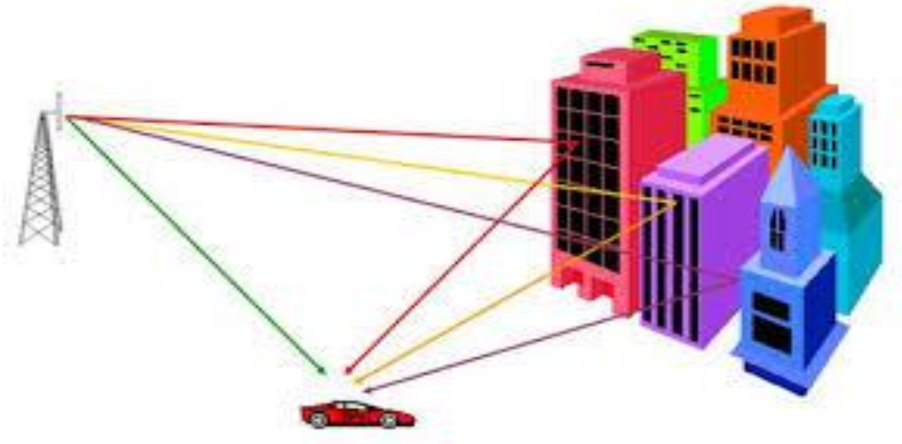


Figure IV.1: Rayleigh fading channel model with many scatters between transmitter and receiver

2.2 Rician fading channel model

Rician fading is a channel model to propagate the signal with partial cancellation of radio frequency signal so that the signal reaches the receiver with different multipath interference. The line of sight signal is stronger than the other signals that occur in Rician fading.

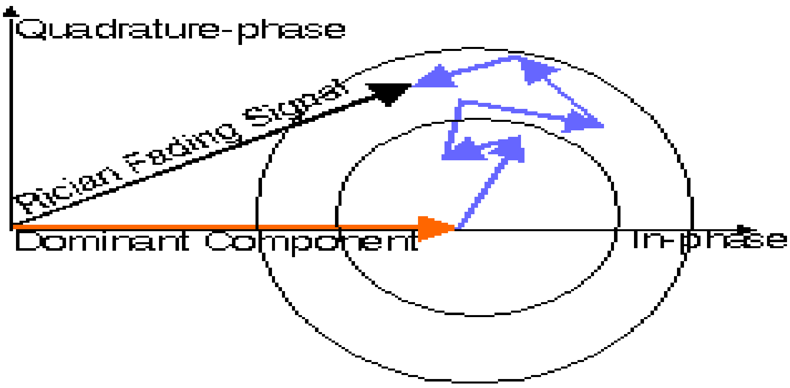


Figure IV.2: Diagram of Rician fading signal

Rician fading channel model that has non fading signal or stationary signal known as line of sight. When there is a stationary signal dominant component present in the line of sight, propagation path with envelope distribution in small scale fading is called Rician.

2. 3 Additive White Gaussian Noise Channel

Additive White Gaussian Noise (AWGN) is a channel model that has a linear addition of wideband or white noise with a constant spectral density and a Gaussian distribution of amplitude and it does not depend on multipath fading, frequency selectivity, interference and non-linearity or dispersion. AWGN is a basic channel model to every communication channel that has statistically random radio noise characterized by a wide frequency range with regard to a signal in a communication channel. This noise has additive, white and noise samples with a Gaussian distribution. It is a simplest wireless communication system employed at additive white Gaussian noise environment, and the signals are generated and some amount of background noise is added in the channel. Mathematical expression for the received signal in additive white Gaussian noise channel is:

$$r(t) = s(t) + n(t) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.. (1)$$

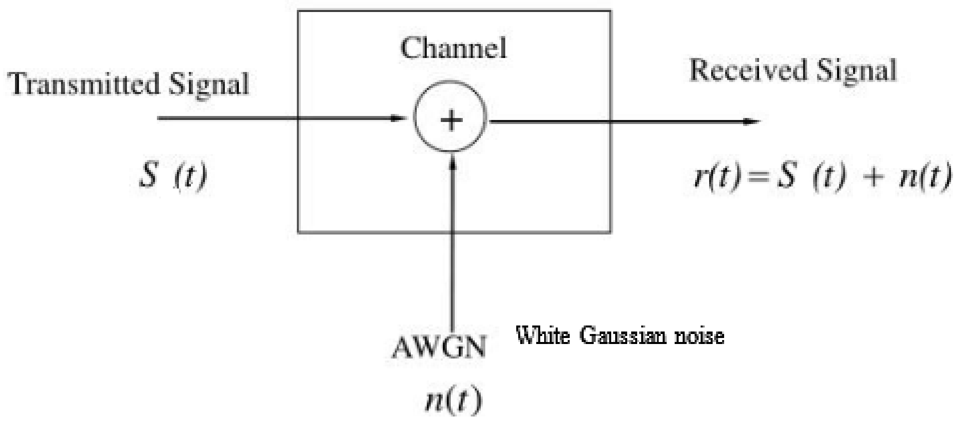


Figure IV.3: Block diagram of AWGN channel

Where:

s(t): Transmitted signal

r(t): Received signal

n(t): Background noise

In this work we used AWGN channel because of its simplicity and because studies showed that it has lower BER rate compared to the other channels which helps us have a better performance and less noise.

In digital transmission, the number of bit errors (BER) is the number of received bits of a data stream over a communication channel that has been altered due to noise, interference, distortion or bit synchronization errors.The BER is the divided by the total number of transferred bits during a particular time interval. BER is a unit less performance measure, often expressed as a percentage.

BER= (Bits in Error) / (Total bits received),.......……………………………(2)

Signal-to-noise ratio is defined as the ratio of the power of a signal (meaningful input) to the power of background noise (meaningless or unwanted input) SNR is defined as the ratio of signal power to the noise power, often expressed in decibels.

$$SNR = \frac{P_{signal}}{P_{noise}}, \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.. \dots\dots\dots\dots\dots\dots (3)$$

3 General operation of a transmission

A digital communication system consists of four basic blocks. The functional blocks at the transmitter are responsible for processing the input message (original video), encoding, modulating, and transmitting over the communication channel. The functional blocks at the receiver perform the reverse process to retrieve the original message. The purpose of a digital communication system is to transmit the message efficiently over the communication channel by incorporating various data compressions (e.g., DCT, JPEG, MPEG.X) encoding, in order to reproduce the message in the receiver (reconstructed video) with the least errors.

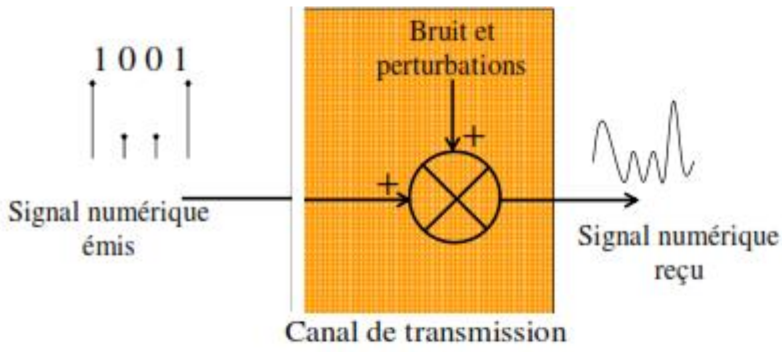


Figure IV.4: General model of an additive noise transfer channel

Figure 4 illustrates the general model of an AWGN channel. The channel is characterized by a transfer function or an impulse response. External disturbances and noise couple to the channel are added to the transmitted signal. This model is the most used for sizing telecommunications equipment because it perfectly reproduces their behavior with respect to noise. This noise has a detrimental effect on analog and digital communications, as long as the signal does not have a much greater amplitude than this noise.

While the signal-to-noise ratio requirements for analog transmissions are very high, they are much lower for digital communications. The noise levels required to induce a bit error must be very large and of the same order as the signal amplitude. In general, it is possible to receive a digital signal with acceptable quality with a slightly negative signal to noise ratio.

Thus, the SNR is not the best metric for qualifying the quality of a digital signal. It is preferred to employ a standardized signal to noise ratio called the signal to noise ratio per bit, denoted Eb / No. This is the ratio between the energy conveyed by a bit Eb and the power spectral density of noise No, this quantity is directly related to the binary error rate, and to set a constraint in terms of rate of bit error amounts to setting a constraint on the Eb/No ratio. Intuitively, we feel that the degradation of a digital signal, in this case a bit interpretation error, will depend on the ratio between the energy transported by a bit and that of the noise.

Here in figure5 , we have summarized the basic steps for a typical video communication system .

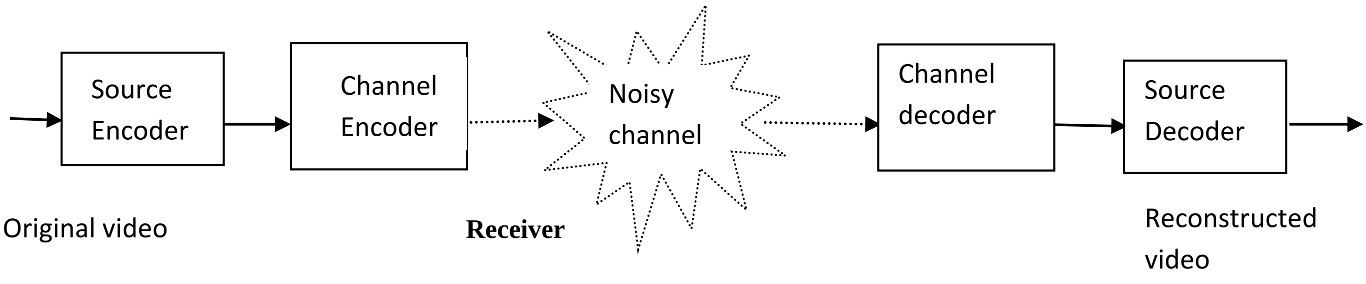


Figure IV.5: A typical video communication system

- **The *source encoder*** is responsible for compressing the original video to represent it with less redundancy. The compressed data is passed to the *channel encoder*.
- ***The channel encoder*** introduces some redundancy in the binary information sequence that can be used by the channel decoder at the receiver to overcome the effects of noise and interference encountered by the signal while in transit through the communication channel. Hence, the

redundancy added in the original video helps in increasing the reliability of the data received and also improves the fidelity of the received signal. Thus, the channel encoder aids the receiver in decoding the desired information sequence.

- **The *source decoder*** reconstructs the original the video The reconstructed video at the receiver is probably an approximation of the original video because of errors involved in channel decoding and the distortion introduced by the source encoder and decoder.
- ***The communication channel*** is the physical medium used to transfer signals carrying the encoded video from the transmitter to the receiver. A range of noise and interferences can affect the video during transmission depending on the type of the channel medium, e.g., thermal noise, atmospheric noise, man-made noise. The communication channel can be air, wire, or optical cable.

4 Multiple Description Coding

Multiple Description Coding (MDC) is an error resilient coding approach to address the packet loss issue on communication channels. In MDC, the video is encoded into multiple streams (also known as descriptions) where these streams are individually decodable and mutually refinable. On the receiver side, if one of the streams is received correctly, then it can be decoded to offer a degraded but acceptable quality. Otherwise, if more than one stream is received then they are combined to achieve higher video quality. This way, the video is available to the receivers.

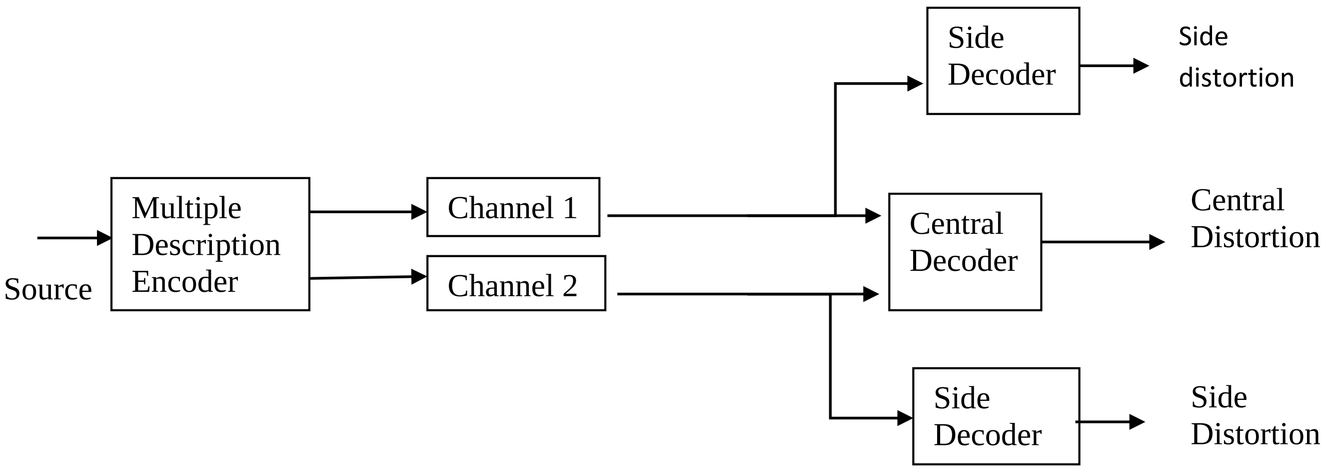


Figure IV.6: Multiple description coding technique with two description

As shown in Figure 6, there are two types of MDC decoders, namely Side decoder and Central decoder. Central decoder is used when all descriptions of the MDC coded video are received correctly. Otherwise, side decoder is used at the receiving end. If there is no packet loss, the prediction reference frame is reconstructed by the central decoder, and this reference frame is matched with the reference used for encoding. On the other hand, in the case of side decoder reconstructing this reference frame, there is a mismatch between the references at the decoder and encoder. If a reference frame is erroneous, then the next frames that directly or indirectly are referred to this frame are also erroneous. In other words, due to the mismatch in side and central decoder outputs, side decoding of one reference frame produces an error that is propagated to the future frames, and this error remains in the sequence until an I-frame is received.

5 Video Quality Metrics

Many accepted methods are used to evaluate and examine video quality. There are two classes of assessment methods: subjective and objective.

5.1 Subjective test methods

The traditional video quality metrics, such as signal-to-noise ratio (SNR), peak-signal-to-noise ratio (PSNR), and mean squared error (MSE), though computationally simple, are known to disregard the viewing conditions and the characteristics of human visual perception. Subjective video quality assessment methods are able to reliably measure the video quality that is perceived by the Human Visual System(HVS) and are crucial for evaluating the performance of objective visual quality assessment metrics. The subjective video quality methods are based on groups of trained/untrained users viewing the video content, and providing ratings for quality. Also, to meet the ITU-T recommendations for subjective quality evaluation, the tests have to follow strict evaluation conditions, including conditions on viewing distance, room illumination, test duration, and evaluators' selection. Subjective test methods require human viewers. Subjective assessment can be a costly and time-consuming process, but one, however, that yields accurate results for any given evaluation.

5.2 Objective test methods

Objective test methods do not use human subjects but rather measure and analyze the video signal. Traditional analog methods are able to accurately measure and assess analog impairments of the video signal. However, with the introduction and development of digital technologies, visually noticeable artifacts are manifested differently than analog artifacts. This change has led to the need for new objective test methods.

The new objective measurement methods analyze the video signal, some utilizing knowledge of the human visual system. These methods implement an algorithm that measures video quality based (usually) on the comparison of a source and a processed sequence. The algorithms, referred to as models, may incorporate characteristics of the human visual system in art attempt to systematically measure the perceptible degradation occurring in the video imager. Performance of the objective models was evaluated with respect to three aspects of their ability to estimate. Subjective. Assessment of video quality:

- prediction accuracy - the ability to predict the subjective quality ratings with low error.
- prediction monotonicity - the degree to which the model's predictions agree with the relative magnitudes of subjective quality ratings.
- prediction consistency - the degree to which the model maintains prediction accuracy over the range of video test sequences, Le., that its response is robust with respect to a variety of video impairments.

5.3 Peak-Signal-to-Noise-Ratio (PSNR)

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad \text{………………………..……….(4)}$$

PSNR is derived by setting the mean squared error (MSE) in relation to the maximum possible value of the luminance (for a typical 8-bit value this is 255). Here, MAX_I is the maximum possible pixel value of the image.

MSE represents the cumulative squared error between the compressed and the original image, whereas PSNR represents a measure of the peak error. The lower the value of MSE, the lower the error.

the latter of which can be almost mapped to MOS.

PSNR(dB)	MOS
> 37	Excellent
31-37	Good
25-31	Fair
20-25	Poor
< 20	Bad

Table IV.1: PSNR to MOS

In this thesis work, when processing videos, we mainly used PSNR, because it is the most widely used objective assessment of quality; it performs a pixel-by-pixel comparison between the reference and the distorted content,

6 MSU Video Quality Measurement tools

MSU is served to compare the perceived video with the original, so that the impact of the use of the network could be analyzed, allows you to use different metrics for comparison (common - PSNR, Delta, VQM, SSIM and our own metrics for measure blurring and blocking). the implementation of the calculation model was carried out in the programming language C.
In our work we used VQMT Video Quality Measurement Tool. This software provides fast implementations of the following objective metrics:

- PSNR: Peak Signal-to-Noise Ratio,
- SSIM : Structural Similarity,
- MS-SSIM : Multi-Scale Structural Similarity,
- VIFP : Visual Information Fidelity, pixel domain version,
- PSNR-HVS: Peak Signal-to-Noise Ratio taking into account Contrast Sensitivity Function (CSF),
- PSNR-HVS-M: Peak Signal-to-Noise Ratio taking into account Contrast Sensitivity Function (CSF) and between-coefficient contrast masking of DCT basis functions.

In this software, the above metrics are implemented in OpenCV (C++) based on the original MATLAB implementations provided by their developers. The source code of this software can be compiled on any platform and only requires the OpenCV library (core and imgprocmodules). This software allows performing video quality assessment without using MATLAB and shows better performance than MATLAB in terms of run time.

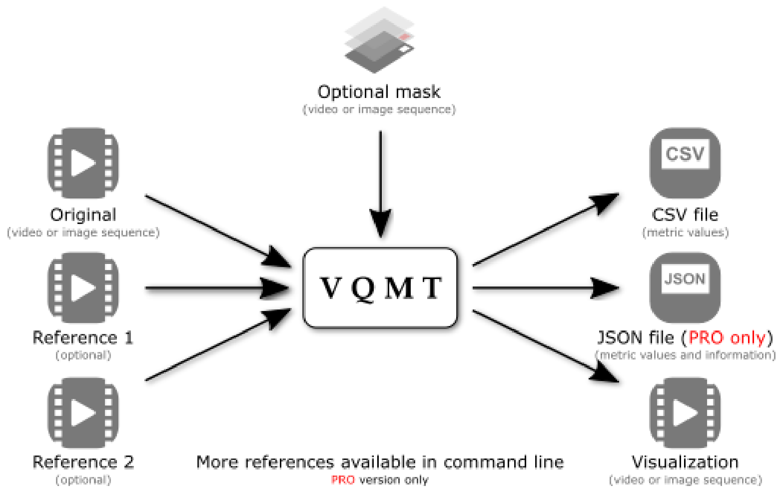


Figure IV.7: Video Quality Measurement tools VQMT

7 YUV player

The application comes in an extremely lightweight package that takes up little space and uses a barely noticeable amount of system resources. What's more, no installation is required so you can benefit from its features wherever you are. Files stored under the YUV format are used as a better means to store color information, in other words, as an alternative to RGB. Both image and video files can be encoded this way and the application can load both. Once your file is loaded, each frame can be accessed, either by manually scrolling through each one or using a dedicated function. The integrated context menu is the only way through which images can be exported. Possibilities let you save the luminance, YUV444, YUV422, YUV420, as well as RGB formats.

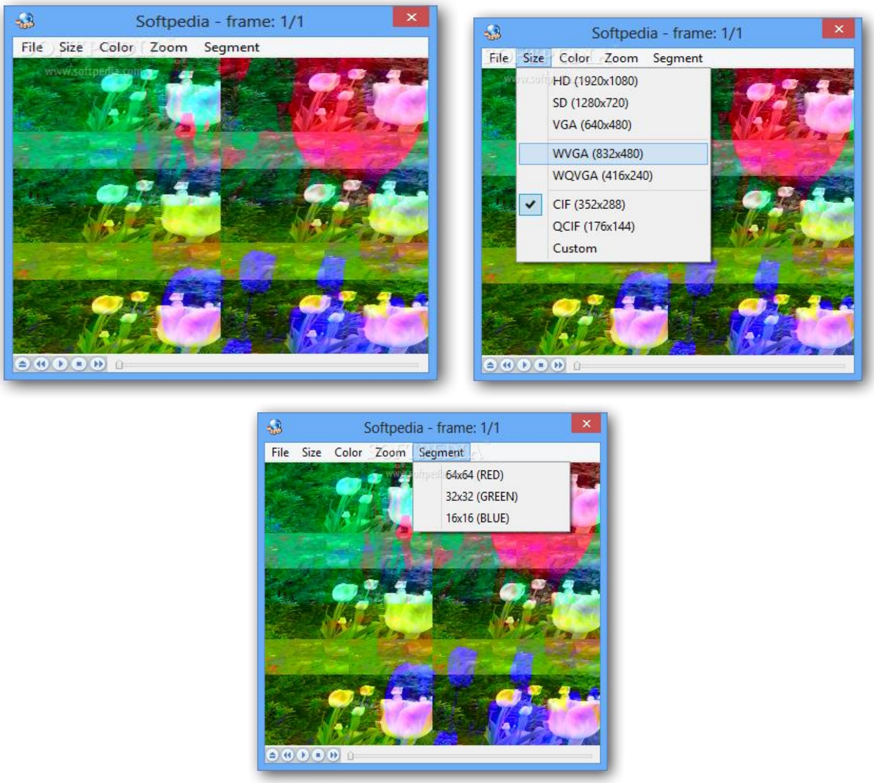


Figure IV.8: Multiple image setup options for YUV player

In addition, view size can be set to several standard formats, such as VGA or 1080p, and even input your own size specifications. Color is also one of the aspects up for selection. Depending on the section you want to be analyzed, options range from multiple YUV formats

Taking everything into consideration, we can say that YUV Player is a lightweight, yet practical video analysis tool. It's not for everyone, with solid knowledge required in order to properly use its features. It's good with system resources and you can work on your projects on the go because no setup is required.

8 YUV Video Sequences

Is a WEBSIT were they provide video sequences of commonly used video test sequences in the 4:2:0 YUV format. we used four videos among others Foreman,Coatguard,News,Silent.

Videos	Pixels	Frame rate FR(fps)	Frames
Foreman	352 x288	30	300
Coastguard	352 x288	30	300
News	352 x288	30	300
Silent	352 x288	30	300

Table IV.2: Four video sequences original with pixels, frame rate (fps), frames

Silent

Coastguard

News

Foreman

Figure IV.9: Samples of test videos

This graph presents temporal information and spatial information for the original videos:

TI: presents movement activities.
SI: presents the complexity of video sequences.

We notice that the SI and TI indices vary from relatively small to relatively large for the selected content.

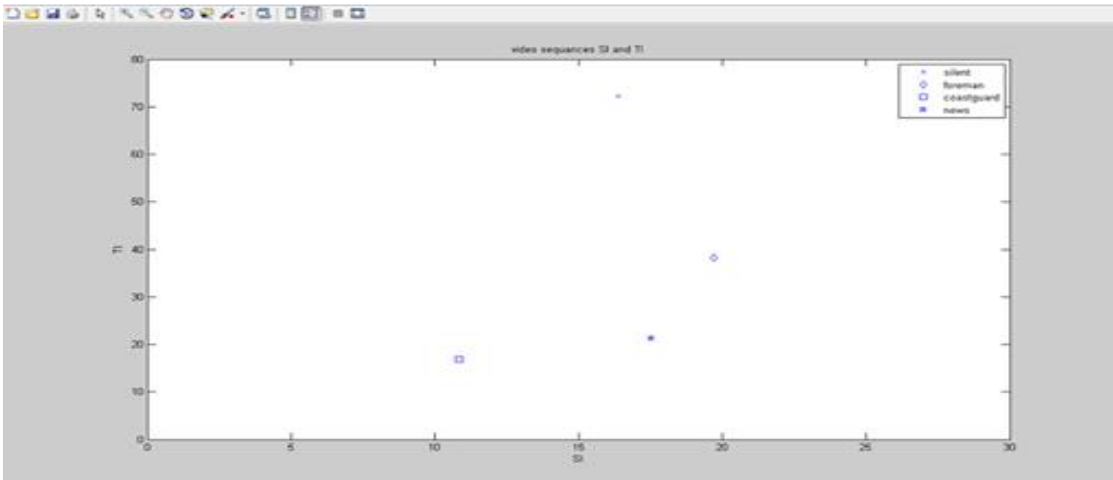


Figure IV. 10: TI and SI plan for original videos

Samples News

Frame 1

Frame 11

Frame 114

Frame 206

Frame 246

Frame 300

Samples of silent

Frame 1

Frame 46

Frame 67

Frame 107

Frame 203

Frame 300

Samples of foreman

Frame 1

Frame 71

Frame 110

Frame 177

Frame 200

Frame 300

Samples of coastguard

Frame 1

Frame 46

Frame 76

Frame 129

Frame 212

Frame300

9 The encoder and decoder implementation under HEVC

The purpose of HEVC-platforms is to allow the user to explore various encoding / decoding options of HEVC and to compare video encoders.

9.1 x265 Encoder

x265 is a commercially funded open source implementation of theH.265/High Efficiency Video Coding (HVEC) compression standard The x265 project is led by MulticoreWare, a leading provider of high performance video software libraries.

HEVC was developed with the goal of providing twice the compression efficiency of the previous standard, H.264 / AVC. Although compression efficiency results vary depending on the type of content and the encoder settings, at typical consumer video distribution bit rates HEVC is typically able to compress video twice as efficiently as AVC. End-users can take advantage of improved compression efficiency in one of two ways (or some combination of both);

- At an identical level of visual quality, HEVC enables video to be compressed to a file that is about half the size (or half the bit rate) of AVC, or
- When compressed to the same file size or bit rate as AVC, HEVC delivers significantly better visual quality.

The encoder x265 offer us:

- The highest possible quality at any given bit rate.
- The lowest possible bit rate for a given quality level.
- Highest visual quality, whether you are doing high quality offline encoding orhigh speed real-time encoding.

9.2 HM encoder

Reference software is being made available to provide a reference implementation of the HEVC standard being developed by the Joint Collaborative Team on Video Coding (JCT-VC) regrouping experts from ITU-T SG 16 and ISO/IEC SC29 WG11. One of the main goals of the reference software is to provide a basis upon which to conduct experiments in order to determine which coding tools provide desired coding performance. the HM encoder allows the user to control to:

- Video resolution (width and height)
- Target bit rate
- Quantization profile for quantization parameters
- Intra-I period
- Number of images to encode.
- Maximum size of the encoding unit
- Maximum depth of the encoding unit
- Slice size determined by CTU, bytes or pixel.
- Various options for motion estimation, such as full search,

An illustration of the user interface for the HM encoder is given in figure.11

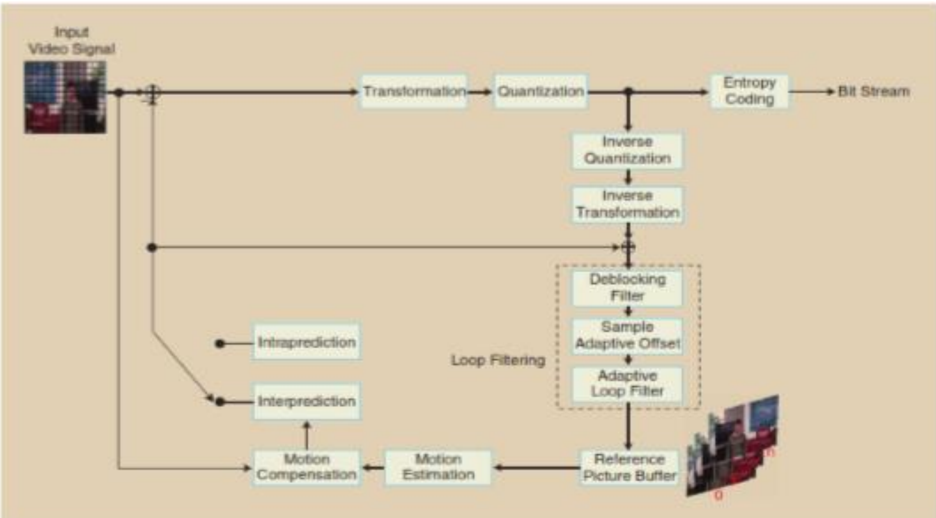


Figure IV.11: Diagram of HM encoder

Additionally, packet loss has been implemented in HM and x265 to help study the effects of packet loss on decoded video quality.

- **Group of Pictures (GOP)**

In all video coding standards, GOPs are used to define coding relationships in a sequence of video frames. Generally speaking, all GOP, modes start with an I Intra/Key frame that isindependently encoded without referencing any other frames. Random access depends on the use of the first key frame within the GOP. Generally, we expect that the use of larger GOP sizes will lead to more compression efficiency. Effective DRASTIC control can be accomplished using a variable GOP structure that supports different inter-mode prediction modes. For this work, the GOP size will be adaptively changed using the HM16.20 Reference Software. HEVC like other video coding standards primarily utilizes three frame types: I (Intra), P(Predictive), and B (Bi-Predictive) types as shown in figure.11. More specifically, we have:

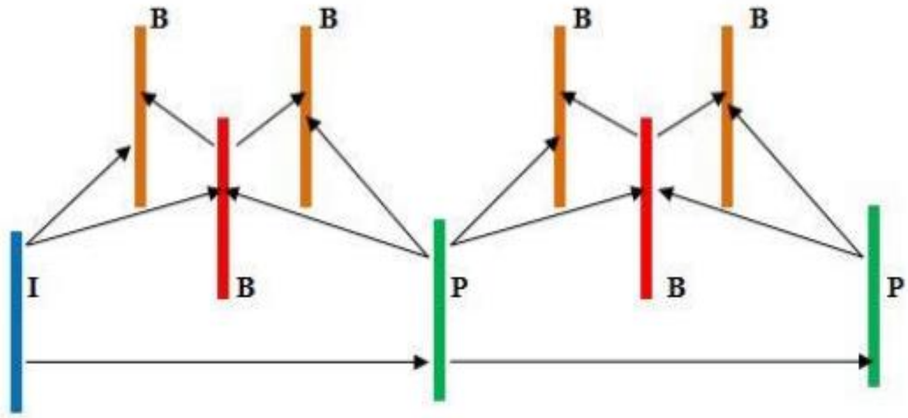


Figure IV.12: An Example of a Closed GOP Structure in HEVC

I-Pictures: Intra (I) pictures, also known as reference or key frames contain all the necessary video information for decoding, without referencing any other frames. Typically, I pictures require more bits to encode than other frame types. Each GOP has anI picture that may or may not be the first frame in the coding video sequence (CVS). I pictures are encoded using angular spatial prediction with modes ranging from 0 to 34 and take lesser time to encode compared to P or B pictures**.**

P-Pictures: Predicted (P) pictures are coded using inter-prediction with one motion compensation signal per PU (i.e. uni-direction) based on the availability of a closest preceding I or P. P pictures only use reference picture list 0 (L0) and P pictures take mode rate amount of time to encode compared to I-frames.

B-Pictures: Bi-directional (B) pictures are coded using inter-prediction with two motion compensated signals per PU (i.e. bi-direction) based on the availability of I and P pictures that come before and after them. B pictures use reference picture list 0 (L0) and list 1 (L1). B pictures take longer time to encode compared to both I and P-pictures.

9.3 Encoder Configuration Modes

The HM 16.20 reference software encoder has three primary configurations. These configurations, include: all intra (AI) mode, random access (RA) mode, and a lowdelay (LD) mode.

All Intra Mode (AI): In this mode, all the frames are encoded only using I slices. It can be an IDR or CRA picture according to the configuration file. And there is no temporal prediction done in this mode as all the pictures are intra-encoded using spatial angular prediction. The quantization parameter for each I picture can be modified within the sequence. Figure .13 shows an all intra mode representation.

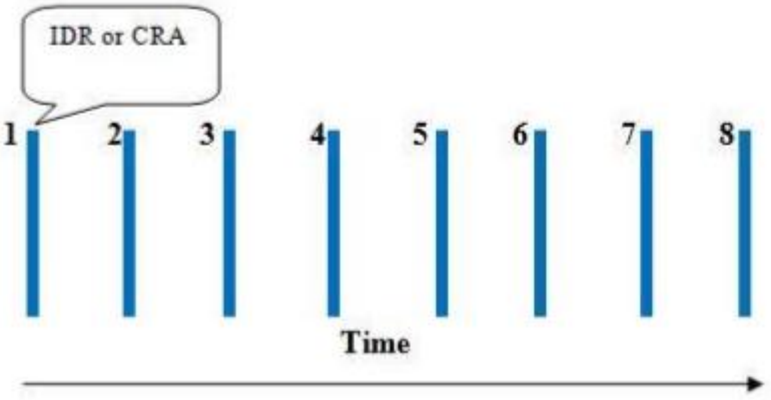


Figure IV .13: All Intra Mode Representation

Random Access (RA): This mode uses a pyramidal or a dyadic relationship among the B slices. Hierarchical B pictures are used for the coding with a random access picture used every 1s. This mode can be used in digital video broadcasting applications. Figure .14 shows a random access configuration. Each picture is depicted with a number showing the encoding order. The first I picture of a video sequence is always encoded as an IDR picture and the other intra pictures are encoded as non-IDR intra pictures ("Open GOP"). Since this configuration follows a pyramid like structure, there are several temporal layers within pictures and each of the intermediate pictures is encoded as a B picture. The Intra period, GOP size and type can always be changed and with this Generalized P and B (GPB) structure that can also be used to define the lowest temporal layer that refers to I or GPB picture for inter prediction.

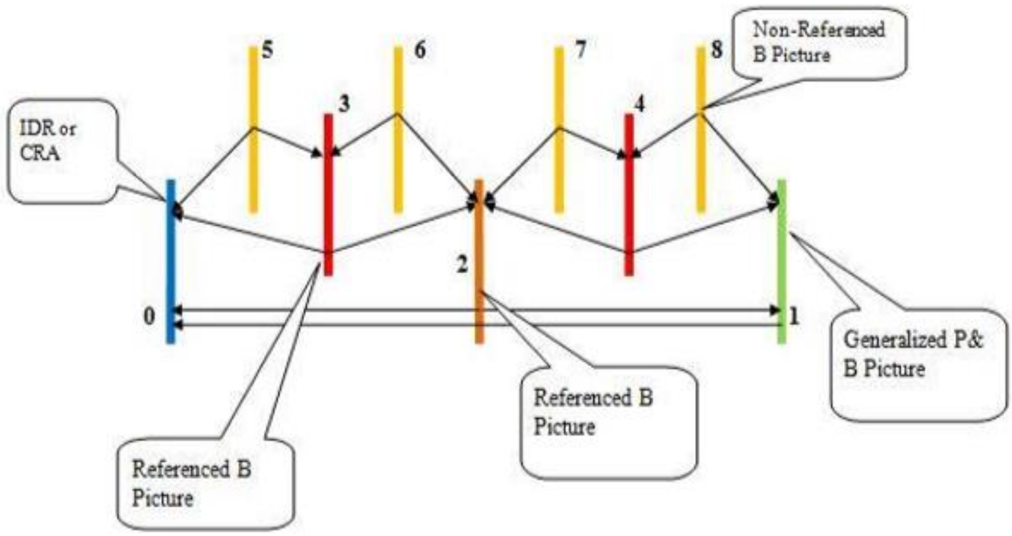


Figure IV.14:. Random Access Representation

The other two layers viz., second and third temporal layers have referenced B pictures, and the last temporal layer from the IDR picture contains only the non-referenced B pictures. QP fo reach inter coded picture is derived by adding a QP offset to QP of Intra coded picture depending on temporal layer.

Low Delay (LD): This configuration slightly differs from the random access mode as there is no picture reordering and only the first frame is encoded using I slices. It is suitable for livestreaming and video conferencing applications. Similar to random access, in low-delay coding conditions, the first picture in the CVS is always encoded as an IDR picture. Other pictures in the sequence are coded as encoded as Generalized P and B-pictures (GPB). Since there is no picture reordering as in random access, the GPB uses only the reference pictures, where the picture order or the display order is smaller than the current picture being encoded. In low delay, there are two reference pictures List0 and List1 in the decoded buffers to be used while in the reconstruction of the original picture. The contents of List0 and List1 are identical, and are updated with a sliding window management process.

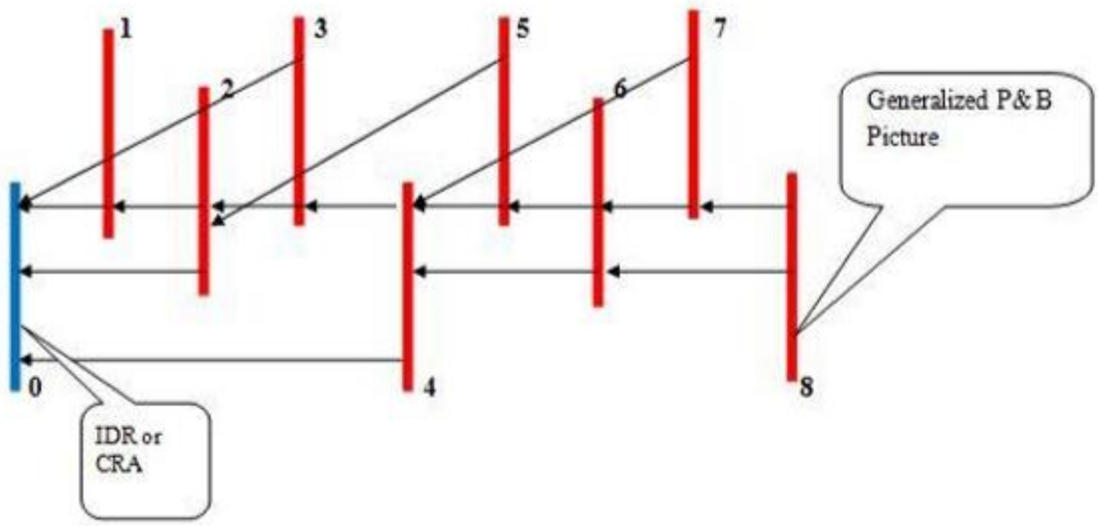


Figure IV.15: Low Delay Representation

The reference pictures shown in the figure .15 demonstrate that there is no picture reordering or in other words no cyclic (dyadic) relationship among reference B pictures. Each picture is associated with a number representing the encoding order and QP for each picture can be modified corresponding to the QP offset derived from the IDR or CRA picture in the beginning of the GOP. In [HM reference software] there is a provision to use low delay configuration with P-picture encoding using the only the reference picture List0 for inter prediction.

10 SIMULATION RESULTS

10.1 Configuration parameters

The defined video content type is used with different a QP=32 to model the impact of the HEVC encoding process for a given test case. Video sequences were encoded using the HEVC / H.265 encoder, the HM 16.20 platform.
Figure IV.16 shows the adaptation diagram of the system that has been designed:

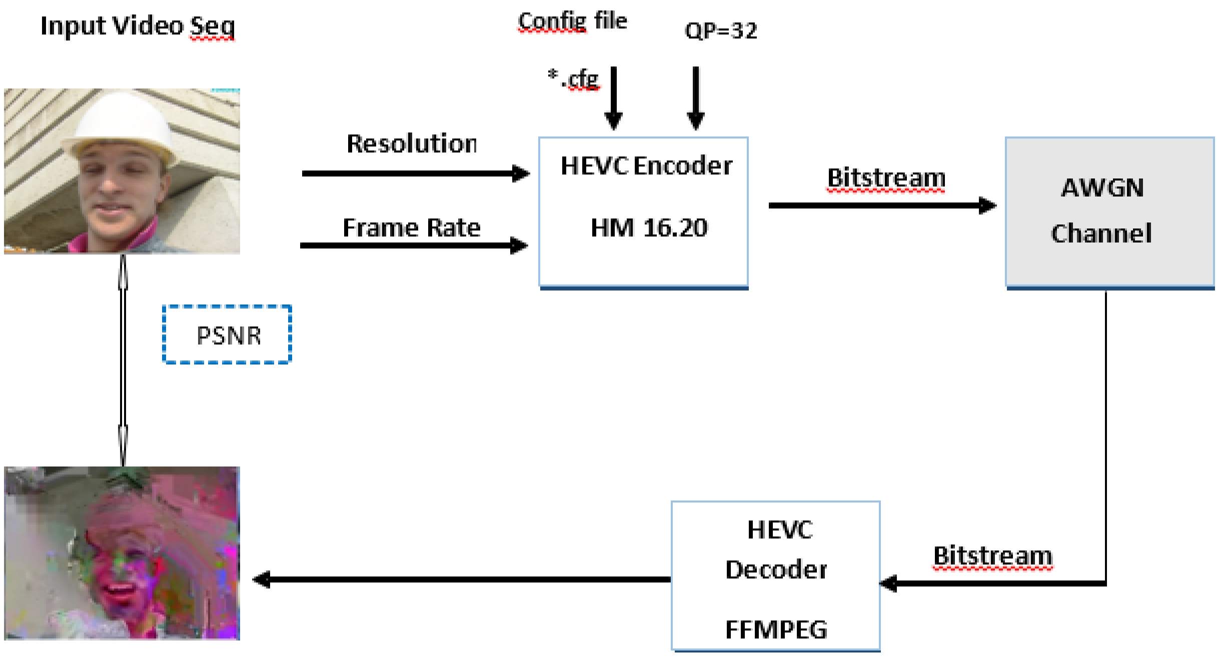


Figure IV .16: Adaptation scheme for the objective / subjective evaluation of video sequences

In this experiment, each frame of the test video sequences is encoded in the low delay mode with B (LB) slices, as it is suitable for live broadcast and video conferencing applications. The configuration file: encoder_lowdelay_main.cfg is used, as shown in Table IV.3.

Item	Value
MaxCUWidth	64
Height	64
MaxPartitionDepth	4
QuadtreeTULog2MaxSize	5
QuadtreeTULog2MinSize	2
QuadtreeTUMaxDepthInter	3
QuadtreeTUMaxDepthIntra	3
IntraPeriod	1
GOPSize	4
RDOQ	1
LoopFilterDisable	1(enabled)
InternalBitDepth	8
SAO	1
QP	32

Table IV.3: the HEVC file configuration

FFmpeg is the leading multimedia framework, able to decode, encode, transcode, mux, demux, stream, filter and play pretty much anything that humans and machines have created. It supports the most obscure ancient formats up to the cutting edge. No matter if they were designed by some standards committee, the community or a corporation. It is also highly portable: FFmpeg compiles, runs, and passes our testing infrastructure FATE across Linux, Mac OS X, Microsoft Windows, the BSDs, Solaris, etc. under a wide variety of build environments, machine architectures, and configurations.

In our work the FFmpeg tool is used on the decoder side, because in some cases the HM platform couldn't decode frames influenced by the alteration of the network, which makes the quality assessment impossible.

This work is based on the comparison between four videos (FOREMAN, COATGUARD, NEWS, SILENT) with and without MDC by changing their EbN0.

10.2 MDC based HEVC encoder parameter

two descriptions of the same video are created; the latter two takes the even and the odd fields respectively, as shown in Figure IV.6, for example, news_even.yuv and news_odd.yuv. Then, under the same conditions as mentioned in Table 4, the HEVC / H.265 encoder was applied to each description. On the decoder side, the central description is reconstructed from the bit streams with the HEVC / H.265 decoder and the side decoder with interpolation (even and odd).

The obtained measurements are shown in Figure 17 respectively, for test video sequences with a HEVC encoder with MDC and without MDC method (see figure IV.6) through a AWGN channel. These measurements are given in terms of PSNR according to the EbN0, respectively.

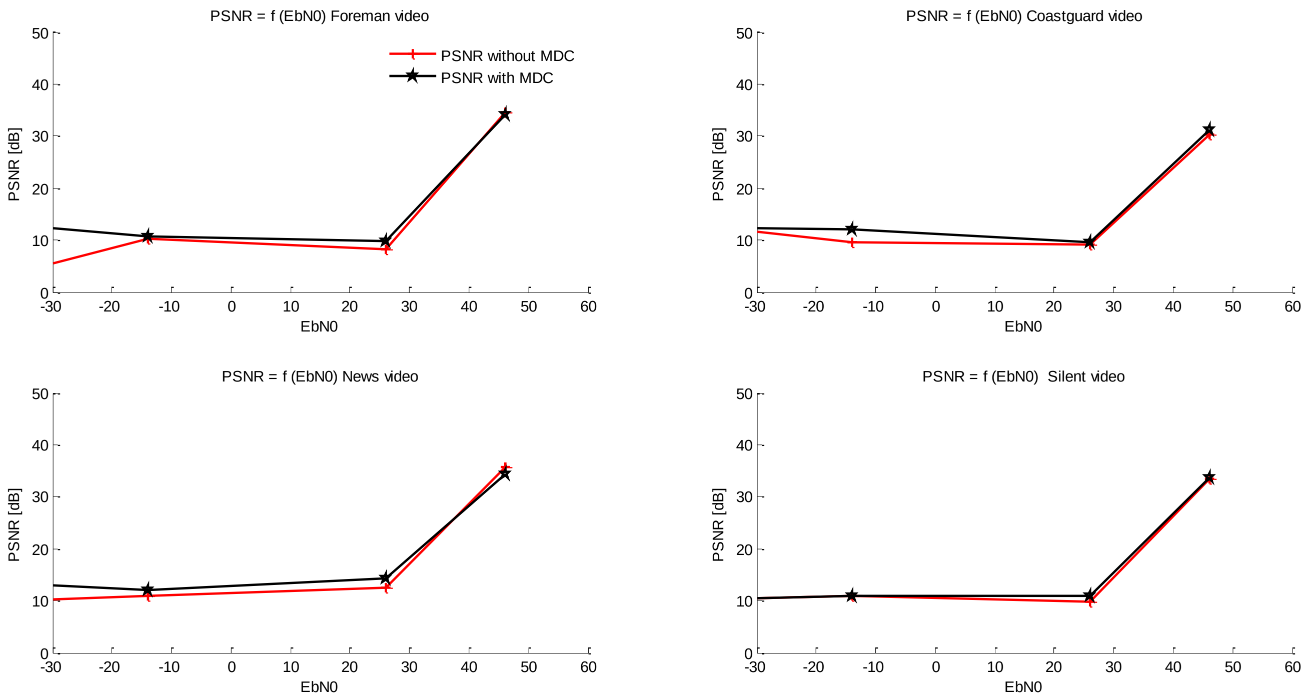


Figure IV.17: HEVC video coding video with MDC and without MDC.

However, any missed description must be recovered using interpolation. However, in this basic approach there is no way to add redundancy and hence no provision for improving the side quality. The following solutions were therefore proposed to address this problem.

If only one of the two descriptions is received, an Interpolation algorithm has to be applied on which description (left or right) is received, we just replicated the frames.

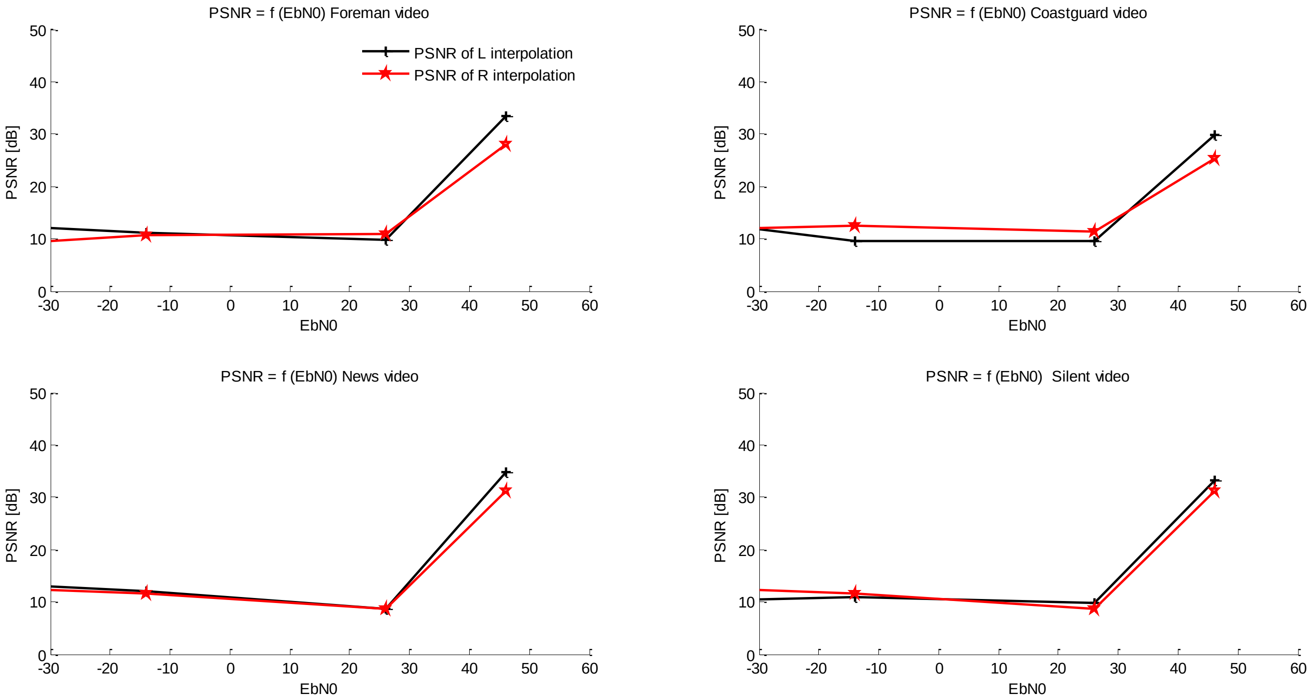
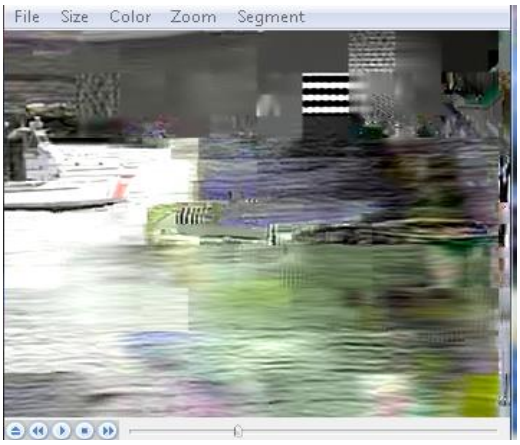


Figure IV .18: HEVC video coding video with MDC interpolation

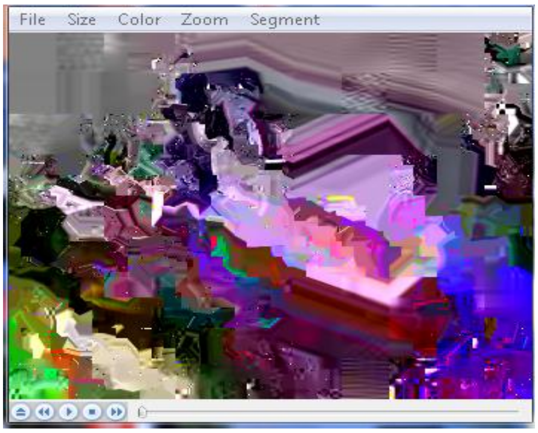
10.3 Discussion

However, Figure 17 and 18 shows that higher values of EbNo results in higher values of PSNR, indicating better use of video perception. As the EbN0 decreases, the PSNR decreases sharply, but it remains acceptable. Therefore, the video quality will be degraded if videos with a very small percentages of bit error, users might consider not watching them again. Videos are more susceptible to data loss.



Coastguard

Foreman



News Silent

Figure IV.19: shows the impact of low SNR(-13) on our tested videos

Starting with foreman video, we have foreman video without mdc, and with mdc (reconstructed), video sideL (using MDC) after interpolation, video sideR (using MDC) after interpolation, these graphs show that we have a difference of PSNR values in the part where EbN0(SNR) is less than (-13). moving on to the news video graph it's obvious that it is different from the others (4 videos), and as we can see the video where we didn't use MDC has low values of PSNR along the line of EbN0 values, and for the 2 videos with interpolation we can see that they are identical and have a bit higher PSNR values than the video without MDC. The video reconstructed using MDC seems to be the best one by having a perfect PSNR values, which we notice that it is higher than any of the last videos, giving us a with good quality that looks more like the video at the transmission.
the coastguard graph seems to have a 4 identical video signals (see figures….), with a slight change in the videos with interpolation, the PSNR dropped a bit but the quality is still good.
the silent graph shows 4 identical signal videos with the same values of PSNR. so the videos seem to be the same while the SNR values are increasing.

We notice that the PSNR of the videos with MDC is way better than the other videos without MDC which means That the videos with the MDC have good quality unlike the videos without it, so that proves that when we use MDC on videos, it doesn't matter if the SNR is decreased the quality of the video is better than without MDC. the robustness provided by the MDC is more important.

The Eb/No ratio depends not only on the power of the signal and the noise, but also on the properties of the signal. The Eb/No ratio therefore takes into account the effect of other influencing parameters on the robustness of a digital communication. This ratio is therefore a more interesting parameter for comparing different communication systems.

In addition, video content has different characteristics due to SI and TI [27], (see figure IV.10), the Silent video sequence has the highest TI, is more sensitive to noise with higher quality degradation than the video, which has the lowest TI and SI compared to silent and news video sequence.

It can be noted that when the EbN0 decreases the degraded perceptual quality, independent of TI of the video sequences.

As well, MDC is a helpful tool that we use on wireless networks, so we can provide a better QOS to the users, and it is the best solution right now to get a better quality and to compact the packet loss, frame, data…. etc.

Conclusion

To understand the impact of video content, encoding parameters such as SNR and packet loss on perceived video quality, we used videos compressed by HEVC standard and MDC method. The results show that for videos with MDC it doesn't matter if we have high SNR value because as it showed above the values of PSNR Estimated is always increasing which gives us a good quality, as for the videos without MDC we see that it's not the same in low SNR but MDC can give a good result. Nevertheless, it can be concluded that MDC is one of the promising solutions for live video delivery over lossy networks in disturbed channel because of its resilient against variable loss rates, and this can be used to get the best performance; so as not to waste the available bandwidth.

General Conclusion

General conclusion

The whole purpose of this study is to satisfy the clients, the users of the network, to give them different kinds of services with a better quality, and to do that, we are obligated to go step by step starting by the transmitter moving to the channel we use and ending up with the receptor. By studying each step improvements, we made sure that the end users are getting a much better quality of their services.

Videos are the better and the easiest way possible to exchange information, putting that as a ground, and knowing that videos are much more harder to control during a transmission, regarding their data size, we used HEVC (HM 16.20) a codec to downsize the data, and before sending the compressed data we knew we will have channel problems such as error ratio, packet loss which will down grade our service quality, so we used the multiple description coding (MDC), then compressed the data, and we send it in a AWGN channel, we add Gaussian noise to the channel, and as a receptor we decompressed the data and had as a result a video with a good quality looking much more like the original video.

The goal of this study is obvious, we used the multiple description coding (MDC) as a tool to maintain a good quality of service and to satisfy the users of the network. Which we did with success.

Perspectives and future work:

HEVC is a really promising new standard. it's already starting to spread and for a least some applications it seems the best standard at the moment. For future works, it would be interesting to test higher resolutions (like 3840x2160) as they are needed in Digital Cinema with more 3D, HD and UHD.

Improve the robustness of HEVC video transmission in real time in vehicular ad-hoc networks. Indeed, these channels are the most hostile to video transmission characterized by a variable topology, a variable number of nodes, variable and fast mobility, etc. To this end, two points must be addressed, namely: real-time video streaming and reduction of the video streaming delay.

References

REFERENCES

[1]: https://scialert.net/fulltextmobile/?doi=jas.2010.1834.1840, consulted on may 28, 2020.

[2] : YJ Zhang, image engineering vol 1 and image processing, 3rd edition, January 1, 2000, consulted on may 28, 2020.

[3]: https://towardsdatascience.com/introduction-to-images-c9c7abe6bfd2, consulted on may 29, 2020.

[4]: http://dl.icdst.org/pdfs/files/da090a75f2b3c3179de82d428b33ef4d.pdf, consulted on june 02, 2020.

[5]:https://www.commentcamarche.net/contents/1493-introduction-a-la-video-numerique, onsulted on june 03, 2020.

[6]: http://dl.icdst.org/pdfs/files/da090a75f2b3c3179de82d428b33ef4d.pdf, consulted on june 03, 2020.

[7] :https://blog.video.ibm.com/streaming-video-tips/what-is-video-encoding-codecs-compressiontechniques/#:~:text=The%20first%2C%20especially%20as%20it,time%20giving%20a%20quality%20experience.&text=The%20process%20of%20video%20encoding,codecs%2C%20or%20video%20compression%20standards. consulted on june 04, 2020.

[8] : -http://dl.icdst.org/pdfs/files/da090a75f2b3c3179de82d428b33ef4d.pdf, consulted on june 04, 2020.

[9] :http://dl.icdst.org/pdfs/files/da090a75f2b3c3179de82d428b33ef4d.pdf, consulted on june 05, 2020.

[10] :https://www.haivision.com/resources/streaming-video-definitions/h-264/ consulted on june 10, 2020.

[11] :https://www.haivision.com/resources/streaming-video-definitions/hevc-h-265/ consulted on 10 june 2020.

[12]: https://www.techopedia.com/definition/26186/wireless-network, consulted on june 28, 2020.

[13] :https://hpbn.co/introduction-to-wireless-networks/ consulted on june 28, 2020.

[14]: https://searchmobilecomputing.techtarget.com/definition/IEEE-802-Wireless-Standards-Fast-Reference, consulted on july 02, 2020.

[15]:,https://books.google.dz/books?hl=fr&lr=&id=9rHnRzzMHLIC&oi=fnd&pg=PR3&dq=wireless+networks&ots=3zASI6cIx&sig=GRwwXLxng0pEeab6RsCBmIIgkDU&redir_esc=y#v=onepage&q&f=false, consulted on july 05, 2020.

[16] :https://www.paloaltonetworks.com/cyberpedia/what-is-quality-of-service-qos, consulted on july 06, 2020.

[17] :https://link.springer.com/chapter/10.1007%2F978-3-642-23541-2_16, consulted on july 06, 2020.

[18] : Mohammad Kazemi, Khosrow Haj Sadeghi, ShervinShirmohammadi. A Review of Multiple Description Coding Techniques for Error-Resilient Video Delivery, web article published on june 2013, consulted on june 12, 2020.

[19]: https://www.tutorialspoint.com/automatic-repeat-request-arq, consulted on june 15, 2020.

[20]: https://www.tutorialspoint.com/forward-error-correction-fec. consulted on june 15, 2020.

[21] : Mohammad Kazemi, Khosrow Haj Sadeghi, ShervinShirmohammadi. A Review of Multiple Description Coding Techniques for Error-Resilient Video Delivery, web article published on june 2013, consulted on june 18, 2020.

[22]: Mohammad Kazemi, Khosrow Haj Sadeghi, ShervinShirmohammadi. A Review of Multiple Description Coding Techniques for Error-Resilient Video Delivery, web article published on june 2013, consulted on june 18, 2020.

[23]: multiple description coding, a new technologie for video streaming over the internet, ANDREA VITALI. IEEE SIGNAL PROCESSING MAGAZINE

[24]: BER PERFORMANCE OF AWGN, RAYLEIGH AND RICIAN CHANNEL consulted 20.2020

[25]: https://en.wikipedia.org/wiki/Bit_error_rate, consulted on august 13, 2020.

[26]:VikramArkalgudChandrasetty, Syed Mahfuzul Aziz, in Resource Efficient LDPC Decoders, 2018

[27]: Video Quality Experts Group: The Quest for Valid Objective Methods Philip Corriveau,a,e Arthur Webster,b Ann Marie Rohalyc and John Libertd a Communications.

[28]:Resehttp://www.compression.ru/video/quality_measure/video_measurement_tool.htmlarch Centre, Ottawa, Canada

[29]:https://www.softpedia.com/get/Multimedia/Video/Video-Players/YUV-Player.shtml, consulted on august 19, 2020.

[30]: Discussion 6.2.1: What is redundancy, temporal redundancy, and spatial redundancy? | Robert Foxx, consulted on august 23, 2020.

[31]: http://x265.org/hevc-h265, consulted on august 24, 2020.

[32]: Esakki, Gangadharan. "Dynamic Switching of GOP Configurations in High Efficiency Video Coding (HEVC) using Relational Databases for Multi-objective Optimization." (2014). http://digitalrepository.unm.edu/ece_etds/80