View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository@USM

ENHANCED STATISTICAL MODELLING FOR VARIABLE BIT RATE VIDEO TRAFFIC GENERATED FROM SCALABLE VIDEO CODEC

SIMA AHMADPOUR

UNIVERSITI SAINS MALAYSIA

2016

ENHANCED STATISTICAL MODELLING FOR VARIABLE BIT RATE VIDEO TRAFFIC GENERATED FROM SCALABLE VIDEO CODEC

by

SIMA AHMADPOUR

Thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

February 2016

DEDICATION

I dedicate my thesis to my beloved Father and Mother

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my truthful respect and sincere appreciation to my supervisor, Associate Professor Dr. Wan Tat Chee for providing support, criticism and precious guidance throughout my research. He was always willing to help and support and kept encouraging me to finish the research.

My special thanks goes to my father and my mother for their endless and continuous encouragement and valuable support. Without their patience, understanding, and most of all love, the completion of this work would not have been possible. I would like to extend my gratitude to my sisters and my brother, Leila, Sara, and Reza for their understanding and support throughout these years.

Finally, I am grateful to all NAv6 members who contributed toward the success and gave moral support and friendly environment. I would like to acknowledge Universiti Sains Malaysia for supporting the research through the PhD fellowship and RUI grant number 1001/PNAV/817069 with the title “Modelling of SVC for Variable Bit Rate video transmission in wireless sensor network”.

Sima Ahmadpour

TABLE OF CONTENTS

Pages

CHAPTER 2 - LITERATURE REVIEW

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AVC	Advanced Video Codec
ACF	Autocorrelation Function
AODV	Ad hoc On-Demand Distance Vector Routing
ARIMA	Autoregressive Integrated Moving Average
AR	Autoregressive model
ATM	Asynchronous Transfer Mode
B frame	Bidirectional frame
BSN	Body Sensor Network
CBR	Constant Bit Rate
CGS	Coarse Grain Scalability
cdf	cumulative density functions
DSR	Dynamic Source Routing
DAR	Discrete Autoregressive
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DQP	Delta Quantization Parameter
EDAR	Enhanced Discrete Autoregressive
ECG	electrocardiograph
EL	Enhancement Layer
ffGN	fast fractional Gaussian Noise
fBm	fractional Brownian motion
FGS	Fine Granular Scalability
FSM	Finite State Machine
GAR	Gamma Autoregressive model

GBAR	Gamma-Beta Autoregressive model
GOP	Group of Picture
HDTV	high-definition television
HEVC	High Efficiency Video Coding
ITU	International Telecommunications Union
I frame	Intra frame
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
JCT-VC	Joint Collaborative Team on Video Coding
KLT	Karhunen-Loeve Transform
K-S	Kolmogorov-Smirnov
LRD	Long Range Dependence
MGS	Medium Grain Scalability
MLE	Maximum Likelihood Estimator
MPEG	Moving Pictures Expert Group
NE	Normal Equations
OLS	Ordinary Least Square
P frame	Predicted frame
pdf	probability density function
PSNR	Peak Signal to Noise Ratio
RE	Relative Efficiency
sd	standard deviation
SRD	Short Range Dependence
SNR	Signal to Noise Ratio
SSE	Sum of Squared Error

SVC	Scalable Video Codec
P2P	Pear to Pear
Q-Q	Quantile-Quantile
QoS	Quality of Service
TES	Transform-Expand-Sample
UHD	Ultra high definition
VBR	Variable Bit Rate
VCEG	Video Coding Experts Group
VLC	Variable Length Coding
VoD	Video on Demand
WSN	Wireless Sensor Network

PERMODALAN STATISTIK MAJU UNTUK TRAFIK VIDEO KADAR BERUBAH YANG DIHASILKAN OLEH KOD PENYAHKOD VIDEO BERSKALA

ABSTRAK

Mereka bentuk rangkaian yang berkesan dan berprestasi tinggi memerlukan pencirian dan pemodela punca trafik rangkaian yang tepat. Tesis ini menyediakan satu kajian tentang penghantaran, pemodelan dan analisis video variable bit rate (VBR) yang merupakan asas reka bentuk protokol dan penggunaan rangkaian yang cekap dalam penghantaran video. Dengan ini, satu model trafik video VBR yang dikodkan oleh scalable video codec (SVC) telah dicadangkan. EDAR (1) dapat menjana siri video dengan tepat di mana siri ini bersifat seakan-akan trafik video yang sebenar. Model ini telah disahkan dengan menggunakan pelbagai statistik untuk membandingkan jejak simulasi da asal. Pengesahan ini telah dilakukan melalui pengukuran grafik (Quantile-Quantile plot) dan statistik (Kolmogorov-Smirnov, Jumlah Ralat Berganda (SSE), dan Kecekapan Relatif (RE)) serta pengesahan secara bersilang. Tambahan pula, model EDAR (1) juga dibandingkan dengan tiga model yang berbeza dan sedia ada melalui teknik-teknik seperti yang dinyatakan di atas. Keempat-empat model dalam penyelidikan ini termasuk model EDAR (1) telah diimplementasikan untuk setiap wayang secara berasingan. Keputusan menunjukkan bahawa ralat SSE bagi model EDAR (1) adalah lebih rendah berbanding dengan tiga model yang lain sesebuah wayang. Kesemua wayang dalam penyelidikan ini juga turut

dibandingkan dan didapati model EDAR (1) mempunyai ralat SSE yang lebih rendah iaitu 11-30% kurang daripada model lain. Ini bermaksud data yang dijanakan oleh model EDAR (1) adalah lebih tepat dan mirip kepada trafik video yang sebenar. Daripada segi pengesahan bersilang, pengesahan bagi setiap model juga dilakukan secara berasingan. Ralat SSE bagi model EDAR (1) adalah berbeza, iaitu sebanyak 8-20% kurang daripada model lain. Dengan keputusan ini, model yang dicadangkan ini boleh dikatakan efektif dalam memperolehi jejak-jejak yang lebih tepat bagi analisis trafik video dan penilaian prestasi rangkaian.

ENHANCED STATISTICAL MODELLING FOR VARIABLE BIT RATE VIDEO TRAFFIC GENERATED FROM SCALABLE VIDEO CODEC

ABSTRACT

Designing an effective and high performance network requires an accurate characterization and modelling of the network traffic. This work involves the analysis and modelling of the Variable Bit Rate (VBR) of video traffic, usually described as the core of the protocol design and efficient network utilization for video transmissions. In this context, an Enhanced Discrete Autoregressive (EDAR (1)) model for the VBR video traffic model, which is encoded by a Scalable Video Codec (SVC), has been proposed. The EDAR (1) model was able to accurately generate video sequences, which are very close to the actual video traffic in terms of accuracy. The model is validated using statistical tests in order to compare simulated and original traces. The validation is done using graphical (Quantile-Quantile plot) and statistical measurements (Kolmogorov-Smirnov, Sum of Squared Error, and Relative Efficiency), as well as cross-validation. Furthermore, the EDAR (1) model was compared against three different other models using the aforementioned techniques. All four models under the study including the EDAR (1) model have been applied for each movie under the study separately. It is shown that the SSE of EDAR (1) model for one specific movie is less than the SSE of other three models. The same comparison is done for all movies under the study and the SSE of EDAR (1) model resulted in

about 11 - 30 % less error compared to the other models. This means that the data generated by the EDAR (1) model is more accurate and close to the actual video traffic than the other models. In terms of cross-validation, the validation has been done for each model in a specific movie separately. The SSE of the EDAR (1) model varied between 8 - 20 % less than the others. It has hereby been shown that the proposed model is effective in deriving a more accurate trace for both video traffic analysis and the network performance evaluation.

CHAPTER 1

INTRODUCTION

1.1 Research Background

Currently, video streaming applications are experiencing a huge growth. Internet users are used to upload and download videos via multiple Internet sites. The act of recording and sharing videos via cell phones burdened the big data traffic on networks. Furthermore, video calling over the Internet, such as Skype and Facetime, are quite popular. Big companies are also using video conferencing applications for face-to-face communication for meetings and other such endeavours. According to the recent report of Cisco (2015), it is reported that every form of videos, encompassing TV, Video on Demand (VoD), Internet, and P2P would make up 80-90% of the global consumer traffic by 2019. This is best illustrated in that for each second, a total of a million minutes of video are streamed. Taking into account all the aforementioned growth, consumers expect to receive quality service from their respective service providers.

Captured data such as digital video signal require large space for storage and increased bandwidth for transmission. To reduce the storage size, there are several compression techniques that are capable of compressing videos without negatively affecting the quality of the image. Normally, the video quality needs to be optimized at a given bit rate, which is provided by the network. Moreover, the network channel capacity keeps changing based on the network configuration and conditions. Therefore, a video compression technique that is capable of optimizing video quality within bit rates as opposed to a fixed rate is regarded as being necessary (Unanue et al., 2011b). The popular approach to manage efficiently is layered or scalable coding, which will be used in this work, and is commonly known as Scalable Video Coding (SVC). SVC represents an extension of H.264/AVC, which was expected to prop up bandwidth efficiency and loss resilient video streaming (Lin et al., 2008). The structure of SVC (as a VBR codec) was developed mostly for the optimization of the quality of videos possessing extended bit rates (Huang et al., 2009) (N.F Huang, 2009), which makes it attractive for use with low bandwidth networks.

To meet the expectation of customers in low bandwidth networks, the performance of the video service needs to be evaluated. This can be done via real networks and sources for a live experiment. Unfortunately, this can be quite an expensive undertaking. The trace-driven simulations are representative of real traffic load, however, they remain static and are representative of only a single point within the workload space (Al Tamimi et al., 2008). Traces can also be problematic, due to the fact that the simulation needs to take place within a designated number of packets/frames in the trace file, but altering these parameters and extending the traces can be quite complex (Tanwir & Perros, 2013). Furthermore, both the statistical and mathematical traffic models are regarded as being superior choice, due to the fact that they are more adept at describing the subtleties between multiple traffic characteristics (Tanwir and Perros, 2013).

Actually, video traffic modelling is intended to analyze the performance of a particular network (Rose, 1997). Generally, a traffic model represents the real behavior of a network for the development of telecommunication technology. A desired video traffic model should be accurate and mathematically tractable, with less computational complexity (Misra, 2008). Moreover, the match between the real video traffic and the corresponding results from the model dominates the efficiency of the video model.

Video traffic models are mostly applicable for mathematical analysis, simulations, and the generation of synthetic video traces that can be used for performance evaluation, testing, and assessment. Furthermore, it can be used to design synthetic loads that would be applicable towards network benchmarking, allocation of network resources, video streaming services, and delivery of certain Quality of Service (QoS) to guarantee end users (Salah et al., 2011, Tanwir and Perros, 2013).

The nature of traffic models is mainly stochastic; this basically means that data can be represented via the fluctuations in the model's parameters. Therefore, the video traffic model is able to characterize video traffics and represent a large range of video sources by varying only a few parameters. In the video traffic modelling filed, the researchers attempt to improve the accuracy of the models to a level that is reasonably realistic. This work describes the modelling of VBR video traffic originating from the streaming video traces, encoded by SVC compression technique for a low bit rate network environment, such as Wireless Sensor networks (WSN).

A WSN includes sensor nodes that can sense, measure, and gather the information, such as sounds, motions, or video (Misra, 2008). A WSN possessing multimedia capabilities are commonly made up of data sensor nodes, capable of sensing sound or motion and video sensor nodes capturing interesting events. The sensors enhance and complement the current surveillance system designed to address crime and terror attacks. Video sensors embedded in large-scale networks can enhance the capability of law enforcement in surveying areas, public events, private properties and borders (Akyildiz et al., 2007). WSN require a low resolution data to stream. Therefore, based on the characteristics of SVC, it seems to be a good choice of compression technique for WSN transmission environment.

Taking into account the aforementioned background, this work attempts to improve the accuracy of traffic modelling by proposing a Bayesian-based video traffic model. The proposed model is intended to increase the accuracy by adding one additional layer of information to the existing model. Therefore, the proposed model generates a video traffic with better coordination between the generated video traffic and a real video traffic compared to current models. Consequently, the performance of the video traffic can be accurately captured.

1.2 Research Motivations

In the event of a kidnapping, law enforcement personnel will be required to retrieve data, in the form of video footage, to identify the perpetrator. There are also cases where law enforcement personnel buffer images and streams in the event of an accident as a form of scene analysis post incident. Therefore, researchers are required to model traffic to determine whether or not WSN can support this kind of application, which runs over a low bit rate network link.

An efficient and reliable network will need to be aware of the traffic characteristics pertaining to the network. An accurate estimation of the performance of the network is vital towards its success. Performance modelling is salient for service providers, as it helps them improve their respective quality of service (QoS). It will also require a traffic model that is capable of defining the statistical characteristics of real traffic on the network. If this is not done, then the results are suspect; it can either be an over or under estimation of the performance of the network.

Therefore, it is necessary to understand the main characteristics of data traffic, which leads to enhanced network performance and the utilization of network resources. This prompts researchers to conduct statistical analysis and traffic modelling of encoded video traffic traces in different conditions.

1.3 Research Problems

To the best of the researcher's knowledge, there have not been enough research on traffic modelling to overcome the overall video traffic characteristics. Each of the existing models for video traffic modelling have their weaknesses, and a need for the accurate traffic model is highly in demand. Researchers struggle to overcome the obstacles originating from video traffic. The need for an accurate traffic model is acute, and up till now, there is no such model that fits the video traffic properly. However, it exhibits structural characteristic that is dynamic and complex form of multiple compression schemes. A model that can suitably demonstrate the multi-faceted statistical characteristics is scarce, to say the least (Tanwir & Perros, 2013).

From all of the aforementioned obstacles, one of the main problems is the lack of an accurate video traffic model. The research questions are formulated based on these assumptions:

1) What is the statistical characteristic of SVC video sequence to develop a representative statistical model?

2) How to generate more accurate video sequence with a better coordination with the real video traffic encoded by SVC?

3) How to validate the proposed model to make sure it is able to generate the video trace very close to the actual video traffic?

This research will move in the direction of addressing these problems.

1.4 Research Objectives

The overall goal of this research is to investigate the behavior of the SVC codec so as to develop a statistical model for low bit rate network environments, such as WSN. Accordingly, the objectives of the research are:

1) To derive the statistical properties of SVC that can be used to represent a video trace in order to find the proper statistical fit.

2) To develop a more accurate statistical traffic model in order to better represent actual SVC video traffic for presenting packet traces in network simulations.

3) To validate the model by comparing real video traffic against the generated video traffic, in terms of accuracy.

1.5 Research Scope

This research concerns the most important part of multimedia traffic, known as video traffic. In other words, the focus is on video traffic data extracted from (Video Trace Library), and is mainly limited to statistical modelling of video traffic profile generated by SVC video compression. SVC traces are limited to three types of scalability, such as coarse grain scalability (CGS), medium grain scalability (MGS), and spatial scalability.

Moreover, the proposed model was performed for H.265/HEVC, since it is regarded as the latest technique. The proposed model is shown to be compatible as well. In terms of traffic mode, this study intends to consider the VBR as a bit rate mode, since VBR video can provide better quality videos for the same average bandwidth compared to CBR.

As evaluation tools for the model, Quantile-Quantile (Q-Q) plot, Kolmogorov-Smirnov (K-S) test, SSE, RE and cross validation are compared. The performance of the proposed video traffic model is statistically evaluated by using MATLAB and R-software. The performance evaluation using the proposed model in WSN network simulation scenarios is not part of the scope of this thesis.

1.6 Contributions

This thesis contributes in video traffic modelling in the following way:

1) An empirical statistical analysis of SVC video sequence including fitting distribution and autocorrelation function (ACF) have been conducted.
2) An Enhanced Discrete Autoregressive model for SVC video traffic sequences known as EDAR (1) has been developed. EDAR (1) has been produced using Bayesian approach by adding one additional level of information into the existing DAR (1). Initially, to derive a DAR (1) model, one needs to characterize the sources which are strongly important due to the complexity and diversity of SVC video traffic. As a result, first, the best fit of marginal distribution of data traffic as a critical part of modelling is investigated. To the best of the researcher's knowledge, there is no specific research focusing on DAR (1) model in sources of scalable coders. In comparison with existing models, the proposed model is accurate, mathematically tractable and needs less computational complexity.
3) The validation of the model has been done in both graphical and statistical aspects. Meanwhile, Cross Validation is employed to verify the validity of the proposed model. This validation is commonly used in the area of "Artificial Intelligence".

1.7 Key Research Steps

This thesis was conducted using a combination of statistical analysis and the modelling to generate a video traffic model in order to have a better coordination with actual video traffic. The model developed is based on Bayesian approach which adds one extra layer of information to the existing model that makes it more accurate than before. **The first step** is to investigate the statistical characteristics of the actual video traffic. The results of this investigation can be used to model the video traffic. **The second step** is to develop a video traffic model based on Bayesian approach. **The final step** of this research is to validate the proposed model using validation techniques.

The following outline identifies the steps in order to develop the proposed model for addressing the problem statement:

1) Quantify the requirements for having an effective traffic model.

– Review the statistical analysis of existing models (by focusing on development of video traffic models).

– Review the conducted research in modelling video traffic for identifying the problem in order to finalize the problem, the type of codec in the considered area of research.

2) Find the statistical characteristics of video frame size sequences.

– The marginal probability density function (PDF) of frame sizes.

– The ACF of video traffic.

3) Develop a model based on defined parameters using in the existing SVC video traces.

4) Validate the model with the validation techniques.

– Compare the performance of the model in both actual data and data generated by the model for SVC streams using graphical test, statistical test, and cross-validation with the other three models.

1.8 Outline of the Thesis

After introducing the significance of this research by providing evidence and background information, stating existing problems, and clarifying its objectives, the rest of the thesis is divided into 6 chapters:

Chapter 2 comprehensively reviews the research work in the field of traffic modelling. The researcher will also discuss various video compression standards, video traffic models, overview of Bayesian theory, frame size distributions, as well as statistical evaluations. Overall, the chapter concludes by summarizing existing models and their corresponding algorithm and architectures, etc. and provides a perspective for introducing the proposed model in methodology.

Chapter 3 covers the methodology discussion on how the proposed model was designed. It defines the requirements and specifications for proposed model and how the enhancement was employed. Furthermore, the evaluation techniques used to evaluate the proposed model are explained here.

Chapter 4 elaborates the whole implementation of the proposed enhanced model; including frame size analysis, and the implementation details, as well as the format of source data.

Chapter 5 presents a detailed analysis of the simulation results and validation of the enhanced model alongside existing ones.

Chapter 6 concludes the research findings and suggests possible future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides the relevant information for the current research, encompassing video compression standards, important works on video traffic modelling, well-known distribution for frame size analysis, statistical evaluation techniques, and video frame size source. Section 2.2 presents the fundamentals of video coding including the compression strategies and review of its standards. Since a Scalable Video Coding (SVC) was regarded as a main compression technique in this thesis, a general description containing different types of scalability have been investigated in Section 2. 4. Meanwhile, a review of existing models in the literature, including their respective advantages and disadvantages will be presented in Section 2.5. Section 2.6 details a Bayesian approach, which was used to propose an enhanced model. Sections 2.7 and 2.8 will discuss the statistical characteristics and evaluation techniques that will be utilized throughout this thesis. This chapter will end with comprehensive summary pertaining the work.

2.2 Fundamentals of Video Coding

A video includes a combination of several frames, where each frame is displayed for a small amount of time to represent the illusion of a moving image. Usually, a video communication system encompasses compression, transmission, and reconstruction. The transmission of video is governed by the steps as follows:

1) The raw video compresses (encodes) into a data stream.

2) The sender recaptures the compressed data from storage devices and send them through the network.

3) The receiver receives the data and decompresses (decodes) and reconstructs them into a video.

Obviously, the raw video requires high bandwidth and large storage spaces. Therefore, it needs to be compressed (encoded) prior to being sent through media using codecs (Golston, 2004). Compression ratio is directly proportional to computational power. Having high redundancy is a usual characteristic for digital images. An efficient compression technique will also be able to reduce redundant information while storing the ones that are to be transmitted. The main aim of compression standards is to enhance the efficiency of coding. This is mostly related to the ability to compress a video to the low bit rate video quality. Thus, compression techniques are regarded as vital towards video transmission.

The main reason for image compression is the high correlation between a pixel and its neighbor pixels. In other words, adjacent pixels have similar values. This is called spatial redundancy, due to the correlation in single frame (Wei et al., 2008).

Moreover, temporal correlation is another issue that needs to be taken into account. A video contains several images, with short time distance between them. Hence two neighboring images are very similar to each other and have a high correlation among images or frames in a period time distance. This kind of correlation in the time direction is called the temporal redundancy interframe correlation. The video compression is achieved only if the interframe correlation can be reduced in an efficient manner (Wang et al., 2001). There three frame types are as follows:

1) I frame: as Intra-coded picture, also known as key frame. It includes the main part of image information. I frames are coded themselves without using information from any other frames.

2) P frame: as a Predictive frame utilizes previous I-type or P-type frame as a reference to code the differences.

3) B frame: as a Bi-directional uses both previous and forward frames such as I frame or P frame as well as the next I frame or P frame as references to code the pixel information.

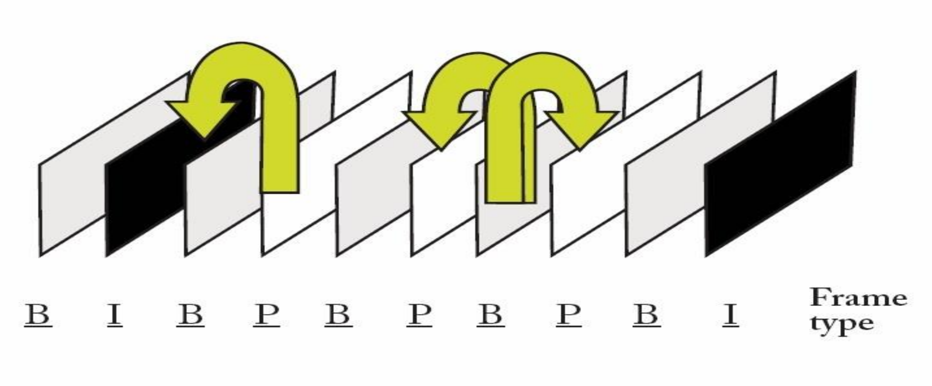

Figure 2.1: I , B, and P frames

As it can be seen from Figure 2.1 for producing B frames, encoder can reference both forward and backward frames to get the highest amount of data compression.

2.2.1 Data Compression Strategies

There are two compression algorithms based on the requirements of reconstruction, which are known as the lossless compression and Lossy compression (Sayood, 2006).

From their respective names the lossless compression techniques basically imply that there will be no loss of information. This means that the recovery of the compressed data will also generate the original data. This kind of technique is developed for applications that are very sensitive towards loss of information between the original and reconstructed data, such as text compression, radiological images, and financial data.

In lossy compression techniques, there will be some unrecoverable loss of information. This is designed for applications who do not mind the lack of precise reconstruction, such as video and sound, where most users will not notice the loss of certain information. In fact, the difference between reconstruction and the original video is not burdensome, as long as it does not corrupt the services. As a result of this, lossy compression is generally used to compress videos. The lossy compression techniques form the main concern of this work. Subsequent sections will review and discuss common video compression standards.

2.3 Video Compression Standards

In this section, various video compression techniques are reviewed, starting from H.261 series, and are briefly presented in Table 2.1.

Table 2.1: A brief history of video compression standards based on the year, publisher, and the common application (ITU-T, 2015; Tanwir & Perros, 2014,).

Standard	Year	Publisher	Popular implementation
H.261	1990	ITU-T	Video conferencing, Videotelephony over ISDN
MPEG-1	1992	ISO	Video on digital storage media -CD
MPEG-2/H.262	1994	ISO, ITU-T	DVD video, Digital video broadcasting, SVCD
H.263	1995	ITU-T	Video conferencing, Videotelephony, Video on mobile phones (3GP)
MPEG-4	1998	ISO	Video on Internet, Object-based coding, Synthetic content, DivX
H.264/MPEG-4AVC	2003	ISO, ITU-T	Blu-ray, HD DVD, Digital video broadcasting, HDTV, iPod video, Apple TV
H.264/MPEG-4SVC	2007	ISO, ITU-T	Polycom video conferencing
H.265/HEVC	2013	ISO, ITU-T	Broadcast (cable TV on optical networks / copper, Satellite, Terrestrial, etc.), Digital cinema, TV broadcasting, Internet streaming

H.261 was introduced by ITU-T for the first time in 1990 (ITU-T, 1990). It is the earliest standard of H.26x family of video coding standards, and is widely regarded as the initial practical video codec. Its operational bit rate range is 64 – 2048 kb/s. H.261 is regarded as the earliest standard that began developing the basic building macroblocks. The blocks includes motion-compensated prediction, block Discrete Cosign Transform (DCT), and two-dimensional run-level VLC coding.

Simultaneously, another group, known as Moving Pictures Expert Group (MPEG) was provided different compression standards. This group was established by the ISO in 1988 for the purpose of developing standards for compressing moving pictures (video) and audio for digital storage media. The prior system was finalized, and called MPEG-1 (e.g., CD-ROM), with a bit rate range between 1.2 Mb/s to 1.5 Mb/s in 1991(ISO/IEC 11172, 1993). It was the first compression standard for both audio and video (Watkinson, 2012). From a quality perspective, **MPEG-1** delivers better quality than H.261in 261in the context of high bit rates. The technical features of MPEG-1 consists of bi-directional predicted frames (B-frames) and half-pixel motion prediction. International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), known as ISO/IET, is a joint technical committee that was formed in 1987.

MPEG-2 was designed to be a cooperative system that worked with both ISO/IEC and ITU-T , which was completed in 1994 (ITU-T and ISO/IEC JTC1, 1994). Its target is to support high-definition television (HDTV) and deliver field-based coding and scalability tools. From a technical perspective, it also helps handle interlaced-scan pictures and hierarchical bit-usage scalability in an efficient manner.

MPEG-3 was intended to standardize the scalable and multi-resolution compression. It was originally meant to handle coding for high-definition video. However, this was also the objective of MPEG-2, and this aspect of MPEG-3 was incorporated in MPEG-2. Currently, MPEG-4 is under development, which led to the omission of ‘3’ from the name (Richardson, 2003).

H.263 represents the initial standard that was designed for the purpose of handling low bit rate videos (ITU-T, 1998). The resulting encoded video is similar in terms of quality to the H.261, however, this is achieved at a much lower bit rate. Its corresponding technical features include variable block-size motion compensation, overlapped-block motion compensation, picture extrapolation motion vectors, three-dimensional VLC coding, and median motion vector prediction.

Unlike MPEG-1/2, H.261/263 are designed specifically to handle video telephony, which means that it only includes video coding and lacks audio coding and systems multiplex. Furthermore, these standards are meant to govern conversational applications (i.e., low bit rate and low delay), and mostly lacks supporting stored data.

MPEG-4 was designed to address the needs of a new generation of highly interactive multimedia applications and produce tools for object-based coding of natural and synthetic audio and videos (JTC1, 1999). Basically, the features of MPEG-4 are made up of object-based coding, synthetic content, and interactivity.

H.264 is one of the recent video standards that is more efficient at coding compared to MPEG-4. This represents a joint effort between ITU and MPEG, and can also be regarded as a subset of the MPEG-4 standard. The emerging H.264 recommendation (also known as MPEG-4 Part 10, 'Advanced Video Coding' and formerly known as H.26L) is known to be a joint effort between MPEG and the Video Coding Experts Group (VCEG), which is a study group of the International Telecommunications Union (ITU).

The Scalable Video Coding (SVC), is regarded as an extension of the H.264/MPEG-4 AVC video compression standard. Its design mostly pertained to playing the role of supporting the bandwidth efficiency and loss resilient video streaming. Its multilayer predictive encoding helps user devices adapt their respective video reception via the extraction and decoding of several selected code layers based on their devices' display capability and network throughput (Feldmann, 1997). This will be explained in detail in Secession 2.4.

High Efficiency Video Coding the latest standard in video compression, commonly known as H.265. HEVC represents the next direction for MPEG video coding. HEVC is a successor to H.264/MPEG-4 AVC (Advanced Video Coding), and is currently under joint development by ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG) as ISO/IEC 23008-2 MPEG-H Part 2 and ITU-T H.265. MPEG and VCEG have established a Joint Collaborative Team on Video Coding (JCT-VC) in their quest to develop HEVC standard (ITU-T, 2014). HEVC aimed to enhance video quality and increase data compression ratio by doubling that of H.264/MPEG-4 AVC. HEVS is capable of supporting 8K UHD (Ultra high definition), and resolutions of up to 8192×4320 pixels (digital video format) (Han et al., 2012). HEVC are present in an extended applications such as mobile TV, home cinema and Ultra High Definition TV (UHDTV) (De Simone et al., 2011). To the best of author's knowledge, there have not been done a traffic modeling for video traces encoded by H.265 so far.

Generally, all of the aforementioned video compression standards are frame-based and block motion-compensated DCT coding. Furthermore, the standards help determine the syntax of the bitstream and decoding semantics, and ease the implementation of encoder and decoder helps make it flexible. New encoding and decoding strategies are devised to precipitate a standard-compatible manner.

2.4 Scalable Video Coding

A desired video streaming system involves these characteristics:

1) The available network is stable.

2) At a given bandwidth, the encoder compresses the video.

3) The decoder decodes all received data.

However, the bandwidth is unstable in real network, and the encoder and decoder should be able to adapt to the quality of the video based on the given bit rate instead of a specific bit rate. Addressing the time restriction behavior that is related to video streaming, the decoder should eschew utilizing packets that are sent just prior to the deadline of their playback. To overcome the aforementioned obstacles, scalable coding is introduced for video streaming. SVC is an effective solution against the problems of modern video transmission systems (Schwarz et al., 2007). The SVC codec receives the bits of network data stream, which are then translated into pictures and videos, and vice versa. Therefore, video bit streams are broken up into subsets. These subsets contain the layers of quality and resolution (enhancement layers), which will be added to the video. SVC codec drops these subsets or packets to prevent the picture from breaking up. This is done by reducing the frame rate, resolution, and the usage of bandwidth of the picture. For example, a HD video conferencing console is capable of receiving base and enhancement layers, while a cellphone is only capable of getting either one, but not both simultaneously. SVC is also backwards compatible, which enables it to communicate with an H.264 codec (SearchUnifiedCommunications, 2012).

Scalability refers to the ability of recovering partial compressed bitstreams. It concerns the removed part of the bitstream that will allow it to cater to the needs of the users and alter the terminal capabilities or network conditions ((Pellan & Concolato, 2009). For instance, take into account MPEG-1 video codec at 1.5 Mbps, which can be downloaded for play back in real time and is connected to the server with a high speed link (for example, ADSL modem). Obviously, having a modem connection of 56Kbps will not allow us to receive enough bits for real time for playback. Scalable video streams allow users with high bandwidth connections to download whole bitstream for a full quality video, whereas users with a 56 Kbps connection are only able to download parts of bitstream, resulting in lower quality videos. This example represents what is called bandwidth scalability.

Scalable coding techniques include coarse granularity (spatial, temporal, quality scalability) and fine granularity (fine granular scalability (FGS)) (Wang et al., 2001). In terms of both categories, the lower priority layer is coded with the residual of the original and reconstructed image of higher priority layers, such as base or lower enhancement layers. The main difference between coarse and fine granularity is that the former improves the quality when a complete enhancement layer has been received, while the latter enhances the quality of the video in real time while receiving codewords, which duly help enhance the layer's bitstream (Dai, 2009). The following subsections discuss the individual scalability and describe their features based on the standardized specifications of the H.264/SVC video codec.

2.4.1 Temporal Scalability

Temporal scalability represents the same video in various frame rates. The encoder proceeds to code the base layer at a lower frame rate while utilizing temporal up-sampled pictures from the lower layer in the form of prediction within the higher layer. The easiest way of realizing temporal up-sampling and down-sampling is via frame copying and skipping.

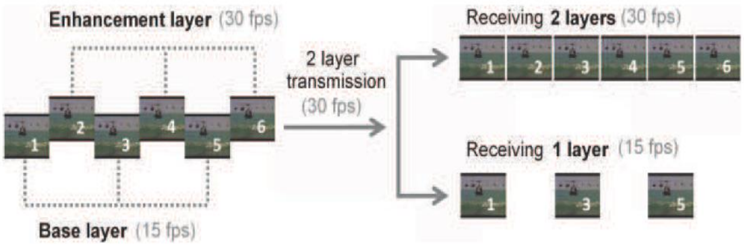

Figure 2.2: An example of temporal scalability approach in H.264/SVC. Source: (Unanue et al., 2011b)

Figure 2-2 illustrates an example of scalable coding using temporal scalability.

2.4.2 Spatial Scalability

Spatial scalability involves the characteristics of similar videos within multiple spatial resolutions. The base layer is generated directly via the image with the lowest resolution. The raw video is spatially down-sampled, DCT-transformed, and quantized. The base layer image is subsequently reconstructed, up-sampled, and used as a prediction for the enhancement layer. The residual between the prediction and the original image is then DCT-transformed, quantized, and coded into the enhancement layer coded bitstream over the Internet (Unanue et al., 2011b).

Figure 2.3: An example of spatial Scalability approach in H.264/SVC. Source: (Unanue et al., 2011b)

Figure 2.3 displays an example of spatial scalability for the purpose of visual identification with regards to spatial scalability encoding.

2.4.3 SNR/Quality Scalability

Quality scalability defines the mechanism used to realize multiple qualities via successive refinement in the quantization of DCT coefficients (Figure 2.4). The encoder subsequently codes the base layer using coarse quantizer while the enhancement layer was coded using a finer quantizer. Due to the fact that multiple quantization accuracies lead to different PSNRs between the original video and the one reconstructed from different layers, this quality scalability is termed SNR scalability (Wang et al., 2002). The H.264/SVC is capable of supporting three distinct SNR scalability modes (Sayood, 2006).

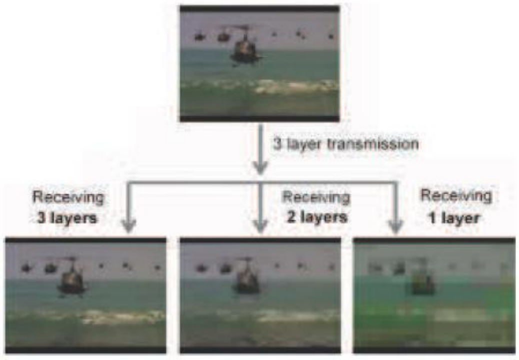

Figure 2.4: An example of quality scalability approach in H.264/SVC. Source: (Unanue et al., 2011b)

2.4.3.1 Coarse Grain Scalability

In coarse grain scalability (CGS), each layer has their respective independent prediction procedures (all references possess similar quality levels), analogous to SNR's scalability of MPEG-2. As a matter of fact, CGS strategy can be assumed to be a special predisposition in the case of spatial scalability, especially when the consecutive layers share similar resolutions (ITU-T, 2014).

2.4.3.2 Medium Grain Scalability

The medium grain scalability (MGS) approach increases efficiency by using a more flexible prediction module. However, this approach has been known to produce a drifting effect (i.e. introducing synchronism offsets between encoder and decoder) upon receipt of base layers. This issue is addressed via the MGS specification proposing the utilization of periodic key pictures, which assist in the instantaneous resynchronization of the prediction module (Unanue et al., 2011b).

2.4.3.3 Fine Grain Scalability

Fine Grain Scalability (FGS) is a version of the SNR scalability that intends to provide a continuous adaptation to the outputs' bit rate vis-à-vis the real network bandwidth. FGS utilizes an advance bit-plane technique, where multiple layers deal with transmitting distinct subsets of bits that are linked to data information. This scheme enables data truncation upon any arbitrary point, which will support the progressive refinement of the transform coefficients. In this particular type of scalability, the motion prediction techniques are casted by the base layers (Unanue et al., 2011a).

SVC is regarded as the target compression technique in this work, due to the fact that it is capable of supporting low resolution video stream. The author is of the opinion that it can be used within a WSN environment due to its similar characteristics. This work only took into account the SVC types that are present in (Video Trace Library), also known as temporal scalability, spatial scalability, CGS, and MGS.

2.5 Video Traffic Modelling

The main evaluation of video network transport with video traces are evaluations with actual video or evaluations with video traffic models. Evaluations with actual video start with the uncompressed source video, carry out the encoding of the source video, simulate the transmission of the actual encoded video bit stream through the transport network, and evaluate the quality of the received video through comparison with the source video.

Such evaluations have the advantage that they allow for the detailed analysis of the received video bit stream. However, these evaluations are very computationally