Prediction of Quality of Experience for Video Streaming Using Raw QoS Parameters

Mohammed Alreshoodi

A thesis submitted for the degree of PhD

Department of Computer Science and Electronic Engineering

University of Essex

Date of Submission: May 2016

Dedication...

To my family

Abstract

Along with the rapid growth in consumer adoption of modern portable devices, video streaming is expected to dominate a large share of the global Internet traffic in the near future. Today user experience is becoming a reliable indicator for video service providers and telecommunication operators to convey overall end-to-end system functioning. Towards this, there is a profound need for an efficient Quality of Experience (QoE) monitoring and prediction. QoE is a subjective metric, which deals with user perception and can vary due to the user expectation and context. However, available QoE measurement techniques that adopt a full reference method are impractical in real-time transmission since they require the original video sequence to be available at the receivers end. QoE prediction, however, requires a firm understanding of those Quality of Service (QoS) factors that are the most influential on QoE.

The main aim of this thesis work is the development of novel and efficient models for video quality prediction in a non-intrusive way and to demonstrate their application

in QoE-enabled optimisation schemes for video delivery. In this thesis, the correlation between QoS and QoE is utilized to objectively estimate the QoE. For this, both objective and subjective methods were used to create datasets that represent the correlation between QoS parameters and measured QoE. Firstly, the impact of selected QoS parameters from both encoding and network levels on video QoE is investigated. The obtained QoS/QoE correlation is backed by thorough statistical analysis. Secondly, the development of two novel hybrid non-reference models for predicting video quality using fuzzy logic inference systems (FIS) as a learning-based technique. Finally, attention was move onto demonstrating two applications of the developed FIS prediction model to show how QoE is used to optimise video delivery.

Acknowledgements

I was once told that PhD is a long journey of transformation from a novice researcher to a professional researcher. The help and support of many made this journey possible for me. I would like to thank my wife for her encouragement, support and faith in me to undergo this huge task. I would also like to thank my parents for being a source of inspiration to me, and to my brothers and sisters for their love and support. Also, it is difficult to overstate my gratitude to my Ph.D. supervisor, Dr John Woods. With his enthusiasm, his inspiration, and his great efforts to make my PhD studies less painful for me.

I would also like to take the opportunity to thank the numerous reviewers who reviewed my papers. Without their constructive comments I would not have been able to raise the standard of my publications. Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this thesis work.

Contents

Abstract **ii**

Acknowledgements **iv**

List of Figures **x**

List of Tables **xiv**

Abbreviations **xvi**

1 Introduction **1**

1.1 Motivation 2

1.2 Thesis Objectives 7

1.3 Contributions of the thesis 8

1.4 Outline of Thesis 11

1.5 List of Publications 13

2 Background and Literature Review **16**
2.1 Introduction 16
2.2 QoS and QoE Layers 18
2.3 Video Streaming 20
2.4 Key Performance Indicators for Video Quality 22
2.4.1 Application level QoS parameters 23
2.4.2 Network level QoS parameters 26
2.5 QoE Measurement Methods 27
2.5.1 Subjective Methods 28
2.5.2 Objective Methods 29
2.6 Classification of Objective Quality Prediction Models 31
2.7 Exploring the Existing Video QoE Prediction Models 34
2.7.1 Models based on Statistical Approaches 35
2.7.2 Models based on Machine Learning Algorithms 36
2.8 Background to Fuzzy Logic Inference System 40
2.8.1 Fuzzifier 41
2.8.2 Fuzzy Inference Engine 43
2.8.3 Defuzzifier 46
2.9 Chapter Summary 46

3 Methodology **48**
3.1 Video Services 49
3.2 QoS Parameters 50

3.3 QoE Measurement 52
3.3.1 Objective Method 52
3.3.2 Subjective Method 53
3.4 Intelligent Machine Learning Algorithms 53
3.5 QoE Prediction 54
3.6 QoE-enabled Applications for Video Delivery 55
3.7 Chapter Summary 55

4 Studying The Impacts of QoS Parameters on Video QoE According to Content Types and Resolutions 57
4.1 Introduction 57
4.2 Related Work 59
4.3 Study 1: Impact of QoS parameters on 2D and 3D VBR Video Streaming 63
4.3.1 Experimental Set-Up 64
4.3.2 Video QoE Measurement 68
4.3.3 Results and Discussion 73
4.4 Study 2: Impact of QoS Parameters on H.264 and HEVC coded CBR Video Streaming 83
4.4.1 Experimental Set-up 84
4.4.2 QoE Measurement 87
4.4.3 Results and Discussion 88
4.5 Chapter Summary 97

5 Hybrid Non-reference Video Quality Prediction Model Using Fuzzy

Logic System **99**

5.1 Introduction 99

5.2 Experimental set-up and Dataset Generation 101

5.3 Methodology 102

5.3.1 Identifying the Inputs and Output 104

5.3.2 Design of Membership Functions 105

5.3.3 Fuzzy Rules Extraction 108

5.3.4 Defuzzification: Predicting the Output 110

5.3.5 Implementation of FIS-PRB model 110

5.4 Validation of FIS-PRB Model 112

5.4.1 Validation by Testing Dataset 112

5.4.2 Validation by External Dataset 114

5.5 Performance Comparision with RNN Method 116

5.6 3D Video QoE Prediction 119

5.7 Chapter Summary 119

6 Adaptive Hybrid Non-reference Model for Real-time Video Quality Prediction **121**

6.1 Introduction 121

6.2 Experimental set-up and Dataset Generation 123

6.3 Methodology 124

6.3.1 Identifying the Inputs and Output 126

6.3.2 Design of Membership Functions 127

6.3.3 Fuzzy Rules Extraction based on LFE Approach 127
6.3.4 Defuzzification: Predicting the Output 131
6.3.5 Implementation of FIS-A model 131
6.4 Validation of FIS-A Model 133
6.4.1 Validation on Testing Dataset 133
6.4.2 Validation on External Dataset 135
6.5 Performance Comparison with FIS-PRB Model 137
6.6 Model Evaluation on Real Test-bed 138
6.6.1 Experimental Set-Up 138
6.6.2 Experimental Results 140
6.7 Chapter Summary 142

7 QoE-enabled Applications for Optimising Video Delivery: Applicability Examples 144
7.1 Introduction 144
7.2 Application 1: QoE-enabled Optimisation Scheme for Real-Time Multilayer Video Delivery 145
7.2.1 Characteristics of scalability extension of H.264/AVC 146
7.2.2 QoE-based transport optimisation scheme 147
7.2.3 Experimental Set-Up 151
7.2.4 Results 153
7.3 Application 2: QoE-enabled Efficient Resource Utilisation Scheme for Mobile Video Delivery 155

7.3.1 QoE-enabled Resource Utilisation Scheme 157
7.3.2 Experimental Set-up 160
7.3.3 Results 163
7.4 Chapter Summary 167

8 Conclusion and Future Work 168

8.1 Conclusion 168
8.1.1 Critical Review of Existing QoE Prediction models 169
8.1.2 Study the Impact of QoS Parameter on Video QoE 169
8.1.3 Hybrid Non-reference Video QoE Prediction Models 170
8.1.4 Two QoE-enabled Applications for Video Delivery 171
8.2 Future Work 172

Appendices 202

A Gilbert-Elliot model 203

B Statistical Performance Metrics 205

B.1 Pearson Correlation Coefficient (R) 205
B.2 Root Mean Squared Error (RMSE) 206

C Correlation of QoE datasets for MUSIC and BMX videos 207

D 3D video frame size 209

E The Extracted Rules of FIS-A Model 211

List of Figures

1.1 The forecasted data for video consumption and adoption, 2013-2018 [1] 3

2.1 QoS/QoE layered architecture 19

2.2 The mapping curve between QoE and QoS 20

2.3 The range of QoS and QoE assessment [2] 21

2.4 Full-reference method 30

2.5 Reduced-reference method 30

2.6 No-reference method 31

2.7 Block diagram of the FIS controller 41

2.8 Membership function chart (triangular shape) 43

2.9 Fuzzy processing example showing (a) two fuzzy input values, (b) the four rules that they trigger, and (c) the resulting output [3]. 45

3.1 Components of the Objective QoE Evaluation Framework 49

4.1 One representative frame from each of the 3 source video sequences . 65

4.2 Conceptual illustration of the simulation scene 67
4.3 Subjective MOS scores with 95% confidence intervals 71
4.4 Correlation of subjective MOS and objective MOS 72
4.5 Comparison of the three datasets for Poker HD video sequence 73
4.6 MOS of SD, HD, and 3D video without loss, a) Music b) Poker c) BMX 74
4.7 Impact of PLR on 2D SD and HD video QoE, a) Music b) Poker c) BMX. 76
4.8 Impact of PLR on 3D SD and HD video QoE, a) Music b) Poker c) BMX. 77
4.9 Impact of MBL and PLR QoE for different 2D CT 80
4.10 Impact of MBL and PLR on QoE for different 3D CT 81
4.11 One representative frame from each of the 4 source video sequences . 85
4.12 Topology of the physical tes-tbed environment 86
4.13 Quality of HEVC and H.264 coded video without network impairment 89
4.14 Impact of PLR (2%) on H.264/AVC vs HEVC transmission 91
4.15 Impact of PDR (2%) on H.264/AVC vs HEVC transmission 92
4.16 Impact of PROR (2%) on H.264/AVC vs HEVC transmission 92
4.17 Impact of PLR+PDR+PROR at 2% each on 4kUHD HEVC video . . 94
4.18 Impact of 2% PLR on target bitrates of 4kUHD HEVC transmission . 94
4.19 The impact of PLR on 4kUHD HEVC transmission 96
4.20 The impact of PDR on 4kUHD HEVC transmission 96
4.21 The impact of PROR on 4kUHD HEVC transmission 97

5.1 Functional block diagram of the proposed FIS-based prediction model 103
5.2 Membership functions of the inputs 107
5.3 Membership functions of the output (QoE) 108
5.4 The FIS-PRB model in MATLAB toolbox 111
5.5 Predicted MOS vs.Objecively Measured MOS 113
5.6 Predicted MOS vs Subjectively Measured MOS 114
5.7 Predicted MOS vs. Measured MOS from External Dataset [Khan 2010] 115
5.8 Neural network architecture 117
5.9 Predicted MOS by RNN-based model vs Measured MOS 118

6.1 Functional block diagram for the proposed FIS-A prediction model . 125
6.2 Diagram of the java classes for the FIS-A model 132
6.3 Predicted MOS vs. Objectively Measured MOS 134
6.4 Predicted MOS vs Subjectively Measured MOS 135
6.5 Predicted MOS vs. Measured MOS from External Dataset [Khan 2010] 136
6.6 Topology of the physical testbed environment. 139
6.7 Predicted MOS vs.Objecively Measured MOS 141

7.1 Layer prioritization for SVC video streaming. 147
7.2 A conceptual diagram for the proposed scheme. 148
7.3 Functional block diagram for the proposed FIS-A prediction model . 149
7.4 QETOS scenario vs. best-effort scenario in case of different PLR . . . 153
7.5 QETOS scenario vs. best-effort scenario 154
7.6 The SNR thresholds for operational MCS regions 158

7.7 Functional block diagram of the FIS-A prediction model 159

7.8 The membership functions of MCS 159

7.9 The membership functions of RCE 160

7.10 Conceptual illustration of the simulation scene 161

7.11 Channel SNR vs. BER for the vehicular scenario 162

7.12 MOS vs. SNR, Interview video 163

7.13 MOS vs. SNR, GT fly video 164

7.14 MCS vs. SNR, Interview video 165

7.15 MCS vs. SNR, GT fly video 165

A.1 Two-State Markov Chain Model [4] 203

C.1 Comparison of the three datasets for Poker HD video sequence 208

C.2 Comparison of the three datasets for Poker HD video sequence 208

List of Tables

2.1 Comparison of objective quality prediction models [5] 33
2.2 Comparison of estimation learning techniques [6] 38
2.3 A summary of the selected evaluation approaches for each model . . . 39

4.1 Video sequences chosen and classes assigned 65
4.2 Simulated QoS Parameters 68
4.3 Subjective Mean Openion Scores 71
4.4 2D Input P-frame sizes by QP 79
4.5 Five-way ANOVA on QoE of 2D Video 82
4.6 Video sequences chosen and classes assigned 85
4.7 Encoding Parameters 85
4.8 QoS Measurement Parameters 87
4.9 Mapping of SSIM to MOS [7] 88

5.1 Video sequences chosen and classes assigned 102

5.2 The chosen QoS Input Parameters 105
5.3 Examples of the QoE index decision making rules 109
5.4 Models performance comparison 116
5.5 Performance comparison with the RNN-based model 118

6.1 Video sequences chosen and classes assigned 124
6.2 Models performance comparison 137
6.3 Performance comparison using external dataset 138
6.4 Experimental QoS Parameters 139
6.5 The Performance of FIS-A model in two scenario 142

7.1 The data bits capacity of each MCS for the tested WiMAX model . . 162
7.2 summary of the percentage of bandwidth increase and power saving . 166

D.1 3D Input P-frame sizes by QP 209
D.2 Five-way ANOVA on QoE of 3D Video 210

List of Abbreviations

QoE Quality of Experience

QoS Quality of Service

MOS Mean Opinion Score

VQM Video Quality Monitoring

PLR Packet Loss Ratio

PROR Packet Re-ordering Rate

PDR Packet Duplication Rate

CT Content Type

MBL Mean Burst Loss

BER Bit Error Rate

HD High Defination

SD Standard Defination

QCIF Quater Common Intermediate Format

CBR Constant Bit Rate

VBR Variable Bit Rate

NR No Reference

RF Reduce Reference

FR Full Reference

PSNR Peak Signal-to-Noise Ratio

SSIM Structural Similarity

AQoS Application Quality of Service

NQoS Network Quality of Service

SVM Support Vector Machine

DT Decision Tree

ANN Artificial Neural Network

ANFIS Adaptive Neural Fuzzy Inference System

DMOS Degraded Mean Opinion Score

GoP Group of Picture

PDF Probability Density Function

ANOVA Analysis of Variance

MPEG Moving Pictures Experts Group

RMSE Root Mean Squared Error

ST Spatio-Temporal

TI Temporal Information

SVC Scalable Video Coding

ITU International Telecommunication Union

IPTV Internet Protocol Television

KPI Key Performance Indicator

RNN Random Neural Network

OSI Open systems Interconnection Model

RTP Real Time Protocol

TCP Transmission Control Protocol

UDP User Datagram Protocol

Chapter 1

Introduction

The exponential growth of video services and applications has motivated the research conducted for this thesis. This chapter introduces the context and motivations of the proposed research in this thesis. It discusses the objectives as well as the primary contributions of this thesis. The chapter ends with a presentation of the organisation of the rest of this thesis.

This chapter is organised as follows: the motivation for the proposed research is described in Section 2. The research objectives are discussed in Section 3, while the major contributions of this thesis are summarized in Section 4. A brief overview and a description of the organisation of this thesis are given in Section 5. Section 6 presents the list of publications that have been presented to the research community.

1.1 Motivation

The motivation for the research conducted in this thesis stems from the search for answers to two basic questions.

Why video services?

Multimedia services, including 2D and 3D video streaming services, are a key component of next generation communication systems. According to a recent study [1], video is expected to dominate up to 80% of the global internet traffic by 2018. The same study projected that 66% of this global traffic will be carried over wireless and mobile devices at this time. Figure 1.1 shows the forecasted data for video consumption and adoption from 2013 to 2018. In addition to Standard-Definition (SD) TV, system-on-chips are in active development [8] for High-Definition (HD) TV display on smart phones. Indeed, there is also increasing demand for mobile 3D TV [9], with possible extensions to immersive video applications [10]. This growth has created a threat on network resources scarcity as well as opportunities for network operators and service providers for revenue generation.

Since video, in general, is considered a resource-hungry service, increasing video quality makes it more sensitive to any problems that occur along the video delivery chain. This sensitivity is translated into different levels of degradation in the user experience. When this is coupled with the highly sophisticated functionality offered by electronic devices nowadays, and the corresponding elevation of consumers' expectations to require broadcast level video quality that fully utilizes the capabilities of their devices, the development of robust, high quality video streaming technology becomes a priority. As a consequence, in the recent years, video traffic has diverted the attention of the research as well as industrial communities.

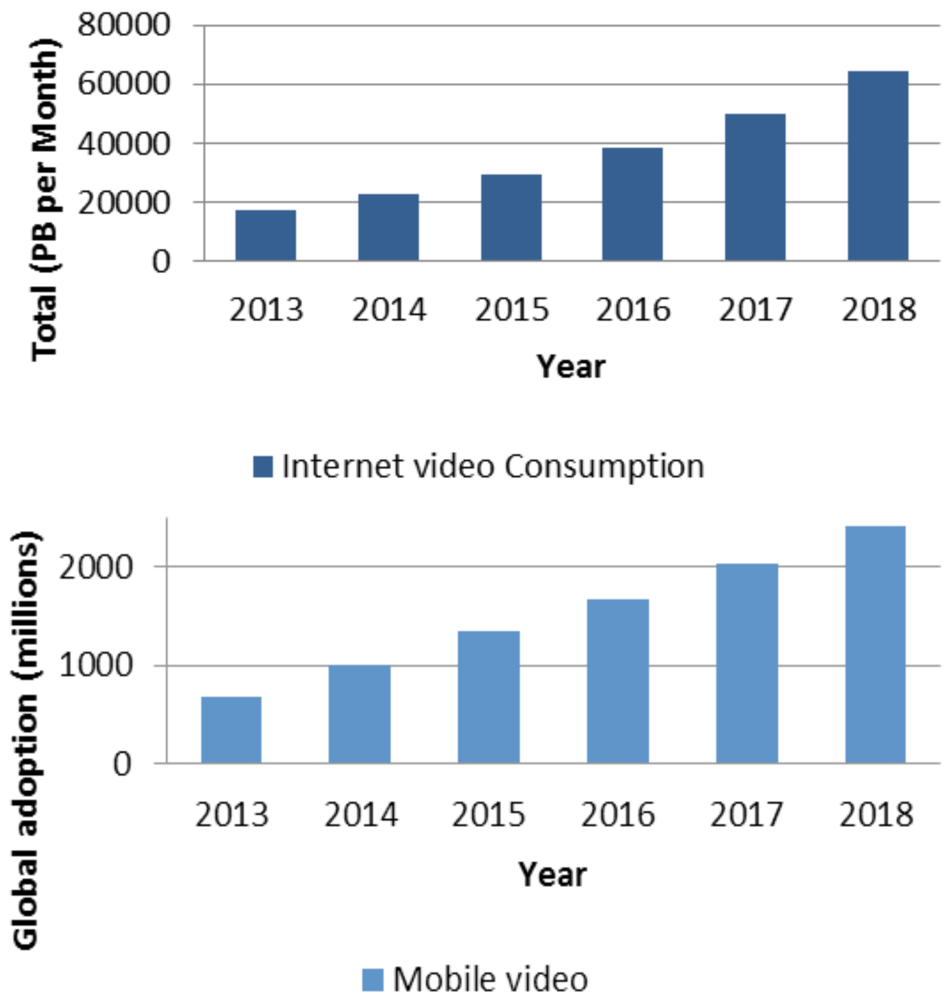


Figure 1.1: The forecasted data for video consumption and adoption, 2013-2018 [1]

Why video QoE prediction and monitoring?

Traditional approaches for measuring user satisfaction rely on Quality of Service (QoS) parameters collected from the network. These QoS parameters are monitored and controlled in order to provide a satisfactory level of service quality. Different QoS parameters such as bandwidth, delay, packet loss, etc., are essential metrics for determining the service quality from a technical point of view. However, QoS parameters do not necessarily reflect the users' satisfaction and feelings towards a particular service. There is now a realisation that the user's Quality of Experience (QoE) is an important metric that should be measured and studied during the design and management of content delivery systems and other engineering processes. There are many definitions of QoE [11–16], all of which are based on the underlying concept that QoE is related to the user's satisfaction with the offered service.

Adopting a more holistic understanding of quality as perceived by end-users (QoE) is becoming a vibrant area of research. When customers experience low quality video service, the service provider cannot afford to wait for customer complaints before taking remedial action. According to an Accenture survey [13], about 90% of users do not make complaints about low quality service, preferring to simply leave their current provider and go to another. To have a competitive advantage, service providers must be aware of this chain effect and develop the capacity to deliver video content through bandwidth limited and error-prone networks with a customer-agreed level of quality for specific services.

Video QoE prediction and monitoring is a multi-disciplinary approach which draws on fields such as engineering science, economics, and user psychology. The video QoE depends on different elements (i.e., content, application, network, etc.) that directly or indirectly affect the user's perception of the video services. Moreover, the diversity of these elements makes QoE estimation rather complex and unpredictable. In addition, there is a lack of accurate quantitative descriptions of QoE. This motivates the investigation of some prominent challenges related to video QoE prediction and monitoring as described below.

Subjective vs. objective QoE measurement: There are two methodologies for evaluating video QoE: subjective and objective measurements. The subjective method requires human participation and a controlled environment for the performance of the test. The disadvantages of this approach are that the methodology cannot be applied in real-time, is more expensive to implement, and requires more time than the objective approach. By contrast, the objective method of video QoE is based on objectively measured application/network parameters. It can be performed in an intrusive (full reference) or non-intrusive way (reference free). Its disadvantages are that it is computationally intensive as it may require complex mathematics and algorithms, and objective metrics can be hard to extract and may not correlate well with subjective QoE [17].

Impacts of QoS parameters on video QoE: In the research community, QoS is known to be the most influential factor on QoE [18, 19]. The interactions between various QoS parameters and their impacts on video QoE are still not well understood.

In addition, there is no standardized methodology that directly and quantitatively maps QoS to QoE. It is also difficult to select network related key performance indicators (KPIs) or QoS parameters that can impact on video QoE. However, the ability to identify the perceived degree of video impairment due to the impact of QoS parameters is a key ingredient in the prediction of video QoE. There is a need to quantitatively investigate the impacts of QoS parameters (arising from both application and physical layers) on video QoE.

Video QoE prediction models: Research on video QoE modelling is still limited. Hence, it is important to model the relationship between QoS and QoE so that QoE can be predicted in the absence of the reference video (no-reference). However, most of the current literature in this area discusses only partial solutions and overlooks some QoS parameters that influence video streaming quality. The existing proposals for video quality prediction tend to consider either encoder parameters, network impairment, or the features of the video content, but rarely consider all three in combination.

Learning and self-adaptive prediction models: Nowadays, researchers study human cognitive processes very carefully and try to develop models that exhibit similar behaviour to brain neurons. Since the human mind is known to be non-deterministic, it is challenging to develop a formal algorithm for human behaviour. This is the reason why researchers turn their attention to learning and self-adaptive models [20].

1.2 Thesis Objectives

The hypotheses tested in this thesis are that it is possible to measure and quantify QoE perceptions, and subsequently to derive a mapping between QoS parameters and the measured video QoE; and that it is possible to develop a QoS/QoE correlation framework for video quality prediction and monitoring. In testing these hypotheses, this thesis addresses the following general objectives:

- determine possible correlations between the raw QoS parameter values and the video QoE.
- undertake a fundamental investigation to quantify the impact of video encoder impairments (e.g. bitrate, spatial resolution, quantization parameters), compression ratio, video content type and access network impairments (e.g. packet/block loss, mean burst length) on perceived 2D and 3D video quality.
- develop novel and efficient hybrid non-reference video quality prediction models to predict perceived video quality from a combination of QoS parameters.
- demonstrate the benefit of the proposed video QoE prediction model by developing two practical applications in areas such as video quality optimization for real-time scalable video streaming and efficient resource utilisation scheme for mobile video delivery.

1.3 Contributions of the thesis

The key contributions of this thesis are:

- A detailed investigation of the relationships between video QoE, access network impairments, encoder impairments and video content type was performed. Both simulation and practical (test-bed) environments were utilized for video QoE evaluation. The investigation was divided into two main studies, which are presented in Chapter 4:

  - Study 1: Investigated the impact of QoS parameters on the video QoE by cross-layer simulations of the transmitted 2D and 3D H.264 video streams. Variable Bitrate (VBR) videos were used in this study. The video quality was measured using the objective and subjective metrics. This work is partially derived from the associated publications [21, 22].

  - Study 2: Investigated the impact of QoS parameters and compression efficiency on up-to 4kUHD of H.264 and H.265 coded videos, both transmitted and decoded simultaneously in real-time. Constant Bitrate (CBR) videos were used in this study. The video quality was measured by the objective metric which was mapped to MOS. This work is partially derived from the associated publications [23, 24].

- Novel hybrid non-reference models for predicting video quality using fuzzy logic inference systems (FIS) as a learning-based technique were developed. For end-to-end quality estimation, QoS parameters from both application and network layers were identified. The proposed models were developed using the measured objective and subjective dataset. The details of these contributions are:

- Model 1: This model is based on a predetermined FIS rule-based method (FIS-PRB). A semi-manual method was applied to develop the model, combining human knowledge and the behaviour of QoS parameters with testing on a simulator. The fuzzy toolbox in Matlab was used to implement the fuzzy controller. Details of this model can be found in Chapter 5. This work is partially derived from the associated publications [22, 25, 26].

- Model 2: This model is based on an automated rule-based method (FIS-A). The proposed prediction model can automatically adapt the fuzzy rules that are used for predicting the QoE. The Mendal-Wang method was used to apply the learning from example (LFE) approach. The model was implemented in the Java programming language. Details of this model can be found in Chapter 6. This work is partially derived from the associated publications [24, 27, 28].

The proposed models were validated through the correlation of predicted QoE and measured QoE data, using both testing and external datasets. Furthermore, the proposed models were validated against the RNN-based prediction model. The FIS-A model was evaluated on a real test-bed which was produced for the QoE prediction of real-time wireless H.265 video streaming.

- Two QoE-enabled applications of the developed FIS-A model are proposed:
  - Application 1 is a QoE-enabled transport optimisation scheme for real-time scalable video delivery. The proposed scheme optimises the video traffic by mapping video quality degradations (that are caused by the network) to the QoE without penetrating the video packets. The main objective is to maximize the QoE with respect to the capacity constraints. This work is partially derived from the associated publications [29, 30]
  - Application 2 is a QoE-enabled efficient resource utilisation scheme for mobile video delivery. This work shows that considering QoE as the basis for modulation scheme selection in AMC can be generally advantageous with respect to power and bandwidth efficiency, by contrast with techniques based solely on network parameters. This work is partially derived from th associated publications [31, 32].

These applications show how QoE is used to optimise video delivery and utilize existing network resources. In particular, QoE/QoS correlation is an indicator for network management and planning processes that will allow for the avoidance of resource over-provisioning. Details of these applications can be found in Chapter 7.

1.4 Outline of Thesis

The remainder of this thesis is organised as follows. Chapter 2 presents a detailed understanding of video QoE prediction. Different methodologies and techniques for video QoE measurement are discussed. The video coding techniques and codecs employed in this thesis are also presented. A number of existing QoE/QoS correlation models for the prediction of video quality are critically reviewed in order to provide a "broader picture" of this area of research.

In Chapter 3, the proposed video QoE evaluation framework is presented and discussed in general terms. This framework exploits the relationship between the QoS parameters and video quality to objectively predict the QoE.

Chapter 4 presents a deeper investigation which aims to quantify the impact of QoS parameters on video QoE. Video QoE is evaluated using both simulated and practical (test-bed) environments. Moreover, subjective and objective assessment methods are used to measure the QoE.

Two main studies are presented in this chapter. The first study investigates the impact of QoS parameters on the video QoE using cross-layer simulation of the transmitted 2D and 3D H.264 video streams. The second study investigates the impact of QoS parameters and compression efficiency on up-to 4kUHD of H.264 and H.265 coded videos which are both transmitted and decoded simultaneously in real-time.

Chapter 5 presents a hybrid non-reference model for predicting video quality using fuzzy logic inference systems (FIS) as a learning-based technique. The proposed model is based on a predetermined rule-based method (FIS-PRB). It is validated using both testing and external unseen datasets. Further validation of this model is performed against another artificial intelligence method known as the RNN-based prediction model.

Chapter 6 presents a hybrid non-reference video QoE prediction model based on an automated rule-based method (FIS-A). The proposed prediction model can automatically formulate the fuzzy rules that are used for predicting the QoE. An extended and further developed version of the Mendal-Wang method is used to apply the learning from example (LFE) approach in designing the proposed automated FIS-A model. The proposed model is validated by using both testing and external unseen datasets. Further evaluation of the proposed model is performed using a real practical experimental test-bed for H.265 coded video streaming.

Chapter 7 presents two QoE-enabled applications that demonstrate the benefits of the developed adaptive FIS-A model. The first application is a QoE-enabled transport optimisation scheme for real-time SVC video delivery. The second application is a QoE-enabled efficient resource utilisation scheme for mobile video delivery. These applications show how QoE is used to optimise video delivery and utilize existing network resources according to users' QoS requirements.

Finally, Chapter 8 discusses the significance of the results presented in this thesis and describes some future directions for research.

1.5 List of Publications

Journal papers:

1- Alreshoodi, M and Woods, J., Survey on QoE/QoS correlation models for multimedia services, Int. J. of Distrib. and Parallel Syst., vol. 4, no.3, pp. 5372, 2013. (published)

2- Alreshoodi, M.; Adeyemi-Ejeye, A.O.; Woods, J.; Walker, S.D., ”Fuzzy logic inference system-based hybrid quality prediction model for wireless 4kUHD H.265-coded video streaming,” in Networks, IET , vol.4, no.6, pp.296-303, 11 2015 doi: 10.1049/iet-net.2015.0018 . (published)

3- Alreshoodi, M.; Danish, E.; Woods, J.; Fernando, A.; De Alwis, C., ”Prediction of Perceptual Quality for Mobile Video Using Fuzzy Inference Systems”, IEEE Transactions on Consumer Electronics, Nov 2015 issue. (published)

4- Alreshoodi, M, Al-Jobouri, L., Woods, J., Fleury, M., Packet Loss Impact upon SD, HD, and 3D Video Streaming According to Content Type, Signal Processing: Image Communication, Elsevier. (under review)

Book Chapter:

5- Alreshoodi, M.; Adeyemi-Ejeye, A.; Walker, S., ”Ultra-High-Definition Video Transmission for Mission-Critical Communication Systems Applications: Challenges and Solutions”, Book chapter in Handbook of Research on Multimedia Services and

Applications in Mission Critical Communication Systems, IGI Global, 2016 (under review)

Conference papers:

6- Alreshoodi, M.; Woods, J., "An empirical study based on a fuzzy logic system to assess the QoS/QoE correlation for layered video streaming," in Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2013 IEEE International Conference on , 15-17 July 2013. (published)

7- Alreshoodi, M.; Woods, J., "QoE Prediction Model Based on Fuzzy Logic System for Different Video Contents,", 2013 European in Modelling Symposium (EMS), vol., no., pp.635-639, 20-22 Nov. 2013. (published)

8- Alreshoodi, M.; Woods, J., "QoE-enabled transport optimisation scheme for real-time SVC video delivery,", 2014 9th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), 23-25 July 2014. (published)

9- Alreshoodi, M.; Danish, E.; Woods, J.; Fernando, A., Resource Allocation Scheme based on Online QoE Estimation of Mobile H.264 Video Streaming, 2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies, Cambridge, UK, 9 Sep 2015. (published)

10- Danish, E., Alreshoodi, M., Fernando, A., Woods, J., "A hybrid prediction model for video quality by QoS/QoE mapping in wireless streaming,", 2015 IEEE International Conference on Communication Workshop (ICC), 8-12 June 2015. (published)

11- Alreshoodi, M.; Woods, J.;, "Optimising the delivery of Scalable H.264 Video stream by QoS/QoE correlation," 2015 IEEE International Conference on Consumer Electronics (ICCE), 9-12 Jan. 2015. (published)

12- Alreshoodi, M.; Danish, E.; Woods, J.; Fernando, A., "Prediction of perceptual quality for mobile 3D video using fuzzy inference systems," 2015 IEEE International Conference in Consumer Electronics (ICCE), 9-12 Jan. 2015, (published)

13- Alreshoodi, M.; Danish, E.; Woods, J.; Fernando, A., " QoE-enabled Efficient Resource Allocation for H.264 Video Streaming over WiMAX " 2016 IEEE International Conference in Consumer Electronics (ICCE), 8-11 Jan. 2016. (published)

Chapter 2

Background and Literature Review

2.1 Introduction

In the past, network quality has been determined by the objective measurement of a number of criteria. This quantification is called the Quality of Service (QoS) of the network. QoS measures the ability of the network to achieve more deterministic behaviour, so that data can be transported with minimal network impairments (e.g. packet loss, delay) and maximum bandwidth. One should note that QoS does not consider the user's perception of the service quality. Another measurement of quality, which takes the user's opinion into consideration, is the Quality of Experience (QoE). The QoE is a subjective metric that involves human dimensions; it incorporates user perceptions, expectations, and experiences of application and network performance into the measurement of network quality [12].

Adopting a more holistic understanding of quality as perceived by end-users (QoE) is becoming a vibrant area of research. When a customer experiences low quality service, the service provider cannot afford to wait for customer complaints. According to an Accenture survey [13], about 90% of users do not complain about low quality service, preferring to simply leave their current provider and seek better service elsewhere. Therefore, it is essential that the service provider has a means of continually measuring the QoE and improving it as necessary.

A variety of factors can affect the perceived quality of service including network reliability, the content preparation process, and the terminal performance. The QoS of multimedia streaming services over IP networks is determined by several interdependent parameters. Some of these parameters can be adjusted, such as bandwidth and image resolution, while others such as packet loss and delay can not. These missing parameters must be considered in order to increase the end user's satisfaction.

The literature describes a number of different video quality measurement methods that have different computational and operational requirements. This chapter presents background information on QoE/QoS correlation and video quality measurement, and highlights some key performance indicators relevant to video QoE. A review of the existing QoE/QoS correlation models will be provided in order to gain a "broader picture" of this area of research.

2.2 QoS and QoE Layers

The term QoS can mean different things to different bodies. The Reference Model for Open Distributed Processing refers to QoS as a set of quality requirements on the collective behaviour of one or more objects [33], while, according to the International Standards Organisation (ISO), QoS characteristics are intended for use in modelling the actual behaviour of systems. The definition of QoS is independent of the means by which it is represented or controlled [34]. Another view, expressed by Vogel et al, is that QoS is the set of those quantitative and qualitative characteristics of a distributed multimedia system which are necessary in order to achieve the required functionality of an application [35].

Different solutions for QoS have been proposed, each of which monitors a variety of layers in the OSI seven layers model. In video streaming services, the two layers that are generally used for QoS measurements are the application and network layers. At the application layer, QoS is concerned with parameters such as the frame rate, resolution, colour, video, and audio codec type. On the other hand, network layer QoS considers parameters such as delay, jitter, and packet loss. Definitions employed by different researchers suggest that a perceptual pseudo-layer can be imagined above both of these layers. This imaginary layer is concerned with the end-user's experience (QoE) [36]. Some researchers consider this pseudo-layer to be an extension of the application layer [37], whereas others view the QoE as an extension of traditional QoS because QoE provides information regarding the delivered services from the user's viewpoint [38]. Figure 2.1 shows the QoS/QoE layered architecture.

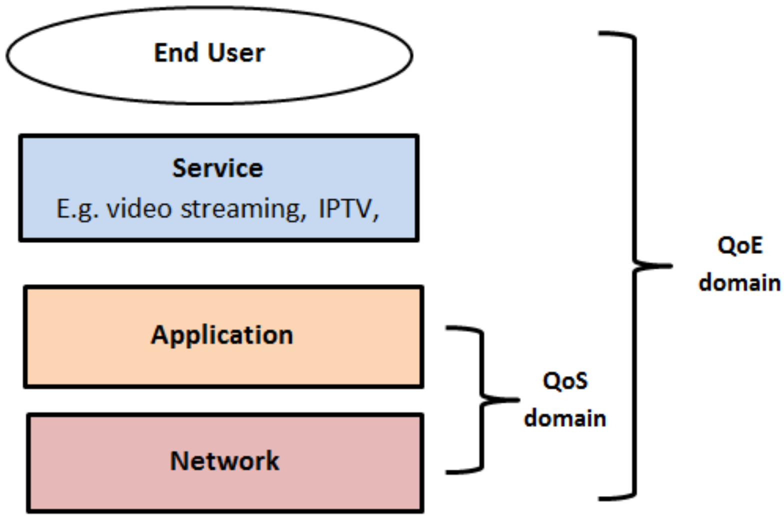


Figure 2.1: QoS/QoE layered architecture

The goal of QoS is to deliver the desired QoE. Delivering a desirable QoE depends on gaining an understanding of the factors that contribute to the overall user experience. Figure 2.2 gives a schematic representation of the relationship between QoS and QoE, which is divided into three zones. When the QoS disturbance lies inside zone 1, the QoE has a high value, i.e., the user's satisfaction is not affected. The QoE decreases when the QoS disturbance reaches zone 2. Finally, when the QoS disturbance enters zone 3, the QoE may fall, i.e., the user's satisfaction will be highly affected and they may stop using the service altogether. In general, when the QoS disturbance parameter increases, the QoE level and the user's perception of quality decrease [39].

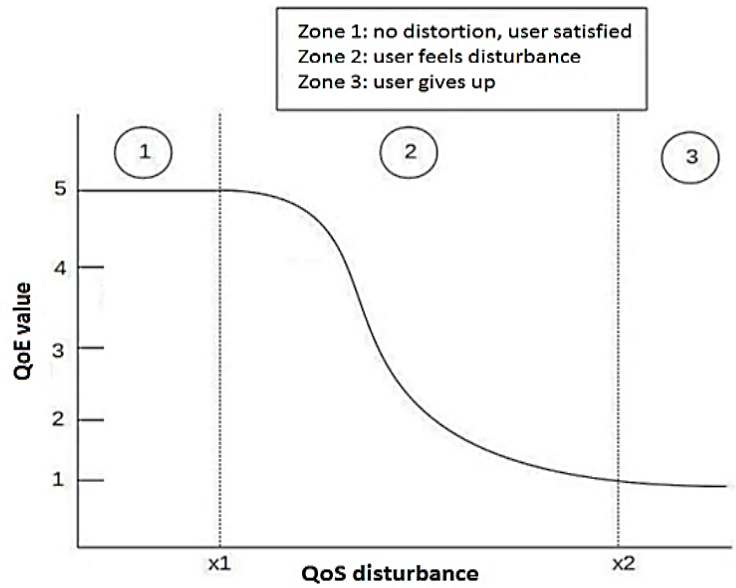


Figure 2.2: The mapping curve between QoE and QoS

2.3 Video Streaming

Video delivery by video streaming attempts to overcome the problems associated with file download, and also provides a significant amount of additional capabilities. The basic idea of video streaming is to split the video into parts, transmit these parts in succession, and enable the receiver to decode and playback the video as these parts are received, without having to wait for the entire video to be delivered. Upon the end-user's request, the video content is retrieved from the server and the channel encoder adapts the video stream to the network QoS requirements. After that, the encoded video stream is partitioned into packets and transmitted over the network.

At the end-user's terminal, the received digital data is transformed into continuous waveform in the source decoder, which can be viewed using different players at the application layer [40].

Unlike traditional services, such as web browsing, where the quality of network delivery is not critical, video streaming services need to deliver content with low distortion of the video quality from the user's point of view. Due to network congestion, the video stream can suffer from different types of perturbation, such packet loss and variation delay. Figure 2.3 gives an example of the QoS and QoE assessment range for video services. The QoS assessment is concerned with data management at the access point, as well as along the network. By contrast, the QoE measurement covers the whole path taken by packets from their source to their terminal diffusion in different modes (i.e., visualization, streaming, and downloading modes) [2, 41].

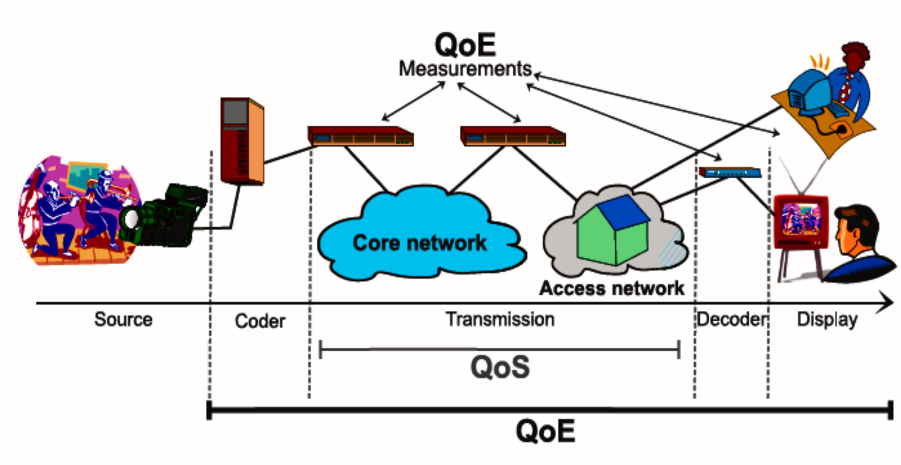


Figure 2.3: The range of QoS and QoE assessment [2]

Different streaming protocols for data transmission are used to control the data transfer between the video server and the clients. Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are the most popular protocols used to communicate between network devices. Most real-time video services employ UDP as the transport protocol [42]. Compared to TCP, UDP does not involve any retransmission mechanism, which makes it attractive to delay-sensitive applications. However, UDP is an unreliable protocol and the video streams can suffer packet loss which might cause distortion of the multimedia content [43]. In order to stream the video in real time, the Real-time Transport Protocol (RTP) [44] must run on top of UDP to make use of RTP services. RTP does not guarantee QoS or reliable delivery, but provides support for applications with time constraints by providing a standardized framework for common functionalities such as time stamps, sequence numbering, and payload specification. RTP enables detection of lost packets [42].

2.4 Key Performance Indicators for Video Quality

Different Key performance indicators (KPIs) are involved when video content is being transmitted from the server to a user. KPIs, which are also called QoS parameters, indicate the overall success or quality of a particular service, in this case, a video service. The measurement of KPI parameters at different transmission points can help to identify and locate the sources of problems. In general, video KPIs are classified into two groups namely: network level and application level QoS parameters [41].

2.4.1 Application level QoS parameters

Application level KPIs are the performance parameters that are directly associated with the application layer and the presentation of video content. Some of the important application level KPIs are described below.

2.4.1.1 Video Encoding Parameters

Video encoding is the process of taking the original video, as recorded by a camcorder, mobile phone or webcam, and preparing it for streaming in a digital format according to specific parameters. Several parameters are briefly discussed below [42]:

- Bit rate refers to the minimum rate at which video bits are transferred from a source to a destination. The quality of video content increases with higher video bit rates. Video streams are generally encoded at a Constant Bit rate (CBR) or a Variable Bit rate (VBR). CBR means that each frame uses the same amount of bits regardless of whether it needs them or not. CBR is most commonly used if the video contains a similar motion level across the entire duration. VBR means that you can vary the amount of bits used to represent a frame so that the overall average amount of bits-per-frame is achieved.

- Frame rate refers to the number of video frames presented per second. Higher frame rates lead to better video QoE as the resulting video appears to be smoother.

- Resolution refers to the number of pixels in both dimensions (width and height) of a video frame. A higher frame resolution yields a better video quality. The choice of the frame resolution is based on the available transmission capacity and the application.

- 2D/3D Video types, i.e., 2D/3D, refer to the visual dimensions of the video content. These content types have different service and network requirements.

- The quantization parameter (QP) has considerable influence on the number of coding bits required for the image block. As the QP increases, the number of bits required for output encoding becomes smaller. On the other hand, more output encoding bits are required if the QP is smaller.

2.4.1.2 Video Compression

Video compression reduces redundancies in temporal and spatial directions. In the temporal reduction, the amount of data needed to store a video frame is reduced by encoding only the pixels that change between consecutive frames in a sequence. For spatial reduction, the size of the video data is reduced by selectively discarding up to a fourth or more of unneeded parts of the original data in a frame. Video compression is likely to remain an essential component of video services, even with constant advances in storage and transmission capacities. However, highly compressed video streams are very susceptible to the effects of network impairment [2, 42].

The compression system normally includes an encoder and a decoder. The encoder converts the uncompressed raw media signals into bit streams, while the decoder converts the compressed form back into a representation of the original video data. The encoder/decoder pair is often referred to as a CODEC [45]. Examples of video codec standards are MPEG-4, H.264, and H.265 [46].

Two organizations dominate the video compression standardization industry. One is the ITU-T VCEG (Video Coding Experts Group) and the other is the ISO/IEC MPEG (Moving Picture Experts Group). A number of video compression standards have been proposed. The video streams evaluated in this thesis were encoded using the H.264/AVC (advanced video coding) [47] and H.265/HEVC (High Efficiency Video Coding) [48] video codecs, which are discussed briefly below.

- The H.264/AVC video coding standard developed by the ITU-T VCEG and the ISO/IEC MPEG has the main goal of enhancing compression performance and providing a "network-friendly" video representation addressing "conversational" (video telephony) and "non-conversational" (storage, broadcast or streaming) applications [49].

- H.265/HEVC is the successor codec to H.264, which, like H.264, was jointly developed by the ISO/IEC MPEG and ITU-T VCEG. The primary goal of the H.265/HEVC codec is to achieve 50 percent better compression efficiency than H.264 and to support resolutions of up to 8192×4320 [48].

2.4.2 Network level QoS parameters

Today's internet services rely on unreliable and best effort networks. The network bandwidth available between the source and destination is unknown in advance and can change over time. Therefore, there is no guarantee that the bandwidth, the packet loss, the burst loss, and the end-to-end delay will permit good video quality. Moreover, video streams are highly sensitive to network perturbations. The impact of network perturbations on video content streaming can result in distortions such as frame loss, freezing, and pixelization [50]. In order to provide better video QoE, transmission conditions at the network end should be reliable. The network condition is represented by different QoS parameters. Each video service has its own QoS requirements. Some of the network QoS parameters and their descriptions are listed below [41]:

- Packet loss rate (PLR) is the ratio of the total number of packets lost in transmission compared to the total number of packets sent. Higher packet loss rates result in lower multimedia QoE.

- Burst loss is the loss of a group of consecutive packets. A higher burst loss results in a lower video QoE.

- Average delay refers to an average time needed for a packet to reach from the source to the destination. Larger delay results in the increase of a start-up time in the multimedia playback.

- Packet duplication occurs when the same packet is received more than once. Duplicated packets appear as a result of configuration errors in the network or defective devices.

- Packet reordering occurs when the sequence number of the packet most recently received is smaller than the sequence number of a packet previously received. If the number of reordered packets increases, the video quality decreases.

2.5 QoE Measurement Methods

The current QoE measurement methods can be implicitly categorised into subjective and objective methods. Subjective methods consist of many participants viewing a sample video and rating their personal perception of its quality according to a predefined quality scale. 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 methods that predict subjective quality solely from the network/media parameters. Nevertheless, the objective methods are hard to correlate with human perception and some may require high computational power/time [17, 41].

There are a number of different methods for performing both subjective and objective measurements. The following subsections discuss the common QoE measurement methods.

2.5.1 Subjective Methods

Measuring and ensuring good QoE for video applications is very subjective in nature. Subjective video quality measurement methods require an appropriate test environment. Some of these methods are described in ITU-R BT.500-13 [51] and ITU-T Rec.P.910 [52]. These standards advise on the type of viewing conditions, the benchmark for observers, test material selection, assessment procedures, and methods for statistical analysis. ITU-R Rec.BT.500-13 describes subjective methods that are specialized for television applications, whereas ITU-T Rec.P.910 is intended for video applications.

The most commonly used subjective metric for quality measurement is the Mean Opinion Score (MOS). The standard for MOS is set in the ITU-T recommendations [52], where it is defined to be a numerical value in the range from 1 to 5. Here 1 is "poor" and 5 is "excellent". The participation of at least 15 viewers is considered to be statistically reasonable for these kinds of subjective tests [53]. Generally, the subjective approach provides the most accurate way to measure video quality. However, this approach has several limitations such as:

- Testing environment requires strict control.
- Real-time implementation is difficult.
- It is difficult to automate.

- It is costly and time consuming.

Therefore, there is a need to develop an objective method that produces results comparable to those provided by subjective testing.

2.5.2 Objective Methods

By definition, the objective approach is based on mathematical and/or comparative techniques that generate a quantitative measure of one-way video quality. This approach is useful for in-service quality monitoring and the design of networks/terminals, as well as in codec optimization and selection. The lack of a human presence increases the error margins of the corresponding measurements. However, the measurement process can be automated and, thus, objective measurements can be performed quickly enough for real time implementation. Therefore, service providers and network operators are interested in objective tools that can predict the subjective MOS given by users with a reasonable degree of accuracy. Objective methods can be divided into three main groups: full-reference, reduced-reference, and no-reference methods [17, 54].

Full-reference (FR) methods require reference to the source video. A distorted sample is compared with the original sample in terms of per-pixel processing and temporal/spatial alignment. These methods are impractical for online monitoring and prediction. Moreover, it may be impossible to access the source (original) video. The basic block diagram for the FR method is given in Figure 2.4. The Video Quality

Metric (VQM) [55], Structural Similarity Index Measurement (SSIM) [56] and Peak-Signal-to-Noise-Ration (PSNR) [57] are examples of full-reference methods. In this thesis, VQM and SSIM are used for video quality measurement in Chapter 4.

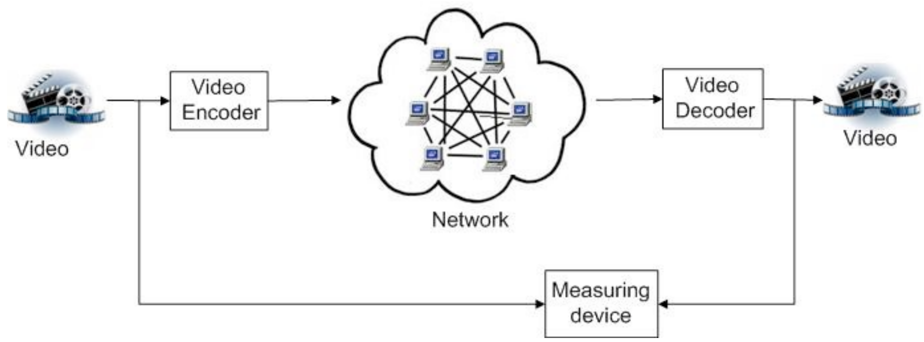


Figure 2.4: Full-reference method

Reduced-reference (RR) methods use only some features extracted from the original video clip. Consequently, if a reduced-reference method is to be used in an IPTV system, the user requirements should specify the side channel through which the featured data is transmitted. Figure 2.5 gives a basic block diagram for RR methods.

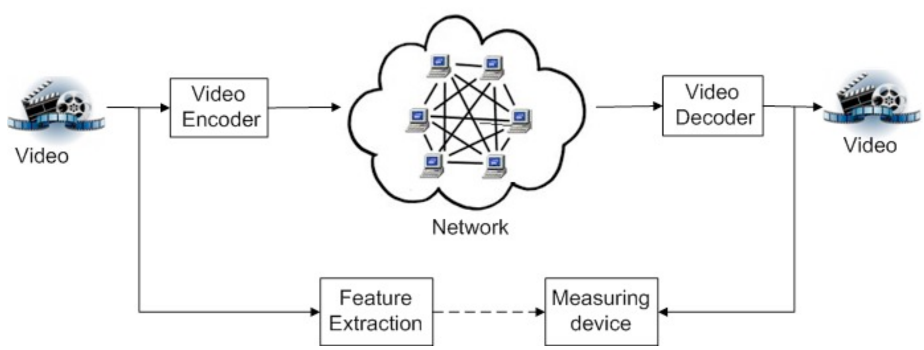


Figure 2.5: Reduced-reference method

No-reference (NR) methods use a degraded signal for the estimation of quality and do not rely on any information about the original reference sequence. This lack of a requirement for access to the source video clip makes no-reference methods particularly attractive. By contrast, it is impracticable to use the FR and RR methods for monitoring live traffic as they require access to the source video. Figure 2.6 shows a basic block diagram for NR methods.

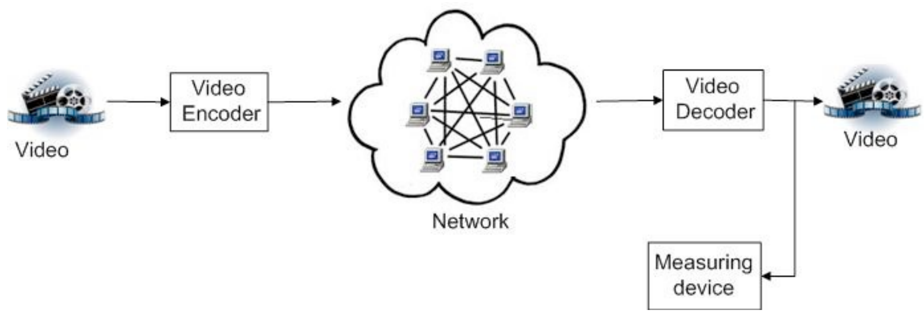


Figure 2.6: No-reference method

2.6 Classification of Objective Quality Prediction Models

It is important to investigate the relationship between end user-oriented QoE and network-oriented QoS parameters. This motivates the search for insight into the principal ways in which the quantitative parts of QoE are affected by the QoS network parameters. It is possible to measure and quantify the QoE, and subsequently derive a mapping correlating the QoS parameters with the measured QoE metrics. Hence, it is feasible to build an effective QoE-aware QoS model. A number of objective models have been devised for estimating QoE. The International Telecommunication Union (ITU) has developed a standardized classification [5] for these models, based on the focus of each model type as follows:

- Media layer models (MLM) predict the QoE by analysing the media signal via HVS. If media signals are not available, this type of model cannot be used. The FR and RR video quality methods fall into the category of media layer models.

- Parametric packet-layer models (PPLR) predict QoE from the packet-header information, without handling the media signal itself. They do not look at the payload information; consequently they have difficulty in evaluating the content dependence of QoE.

- Parametric planning models (PPM) take quality planning parameters for networks and terminals as their input. This type of model requires a priori information about the system under testing. These models typically use a mathematical formula, representing the quality estimation as a function of different parameters.

- Bit-stream layer models (BLM) are a new concept. They lie between parametric packet-layer models and media-layer models, extracting and analysing content characteristics from the coded bit-stream to perform their quality measurements. These models often need to decrypt the encrypted multimedia payload, and can, consequently, have high computational complexity.

- Hybrid models (HR) are combinations of some or all of these models. They are considered to be one of the most effective types of models for multimedia QoE estimation as they exploit as much information as possible to predict the QoE.

Table 2.1 allows the comparison of different objective quality prediction models with existing video standards.

Table 2.1: Comparison of objective quality prediction models [5]

	MLM	**PPLR**	**PPM**	**BLM**	**HM**
Input Information	Media signal	Packet header	Quality design parameters	Packet header and payload	Combination of any
Primary Application	Quality benchmarking	In-service nonintrusive monitoring	Network plan and application designing	In-service nonintrusive monitoring	In-service nonintrusive monitoring
e.g. Standard	ITU-T-J.vqhdtv	ITU-T-P.NAMS	ITU-T-G.1070	ITU-T-P.NBAMS	ITU-T-J.bitvqm

One of the main contributions of this thesis is to propose two hybrid non-reference video quality prediction models that are based on QoS parameters. The proposed models predict video quality directly from a combination of parameters including the access network, encoder related parameters and content types. These two QoE prediction models are discussed in Chapters 5 and 6.

2.7 Exploring the Existing Video QoE Prediction Models

This section presents a review of video QoE prediction models. Different approaches are discussed and analysed in order to obtain a "broader picture" of this area of research. Some of the reviewed models will be discussed in detail, while the features of others will be summarised in tabular form.

The ability to identify the perceived degree of video impairment due to QoS parameters is a key aspect of the QoE prediction of video traffic [18]. Moreover, as discussed in ITU-G. 1080 [58] and TR-126 [19], not all impairments of QoS parameters necessarily result in visible degradations. Therefore, measuring the impacts of a combination of QoS parameters, including parameters from the access network, encoder related parameters, and content types, on the quality of the video traffic is still a challenging task.

Video quality is predicted using either the impact of the NQoS or the AQoS. The papers, [59–63] consider solely NQoS parameters, such as random packet loss, burst loss and delay, for QoE estimation. Another group of research studies have focused on AQoS parameters such as quantization artifacts, frame rate and bit-rate [64–69]. Since the end-user is impacted by both the AQoS and the NQoS, it is important to understand their impacts in order to have a broader prediction of video quality.

2.7.1 Models based on Statistical Approaches

In [70–72], video quality prediction models based on a statistical analysis approach were proposed. The discriminate analysis method (DA) [73] is used in [70] to predict video quality. The authors argue that the inclusion of parameters related to the video content and coding can maximize the user-perceived quality and achieve efficient network utilization. However, their approach suffers from limited accuracy because it considers only two QoS parameters (bitrate and frame rate). Moreover, no specific implementation of the QoS parameters at the network level is considered. Work in [72] proposed a numerical formula for the evaluation of QoE using different QoS parameters such as packet loss, burst loss, jitter, delay, and GoP (Group of Picture) length. However, it also fails to consider user perception, and is lacking in experimental and validation results. Lastly, the authors did not discuss how new QoE parameters can be included in their model.

The work of Khan et al. [74, 75] proposes non-linear regression-based models to estimate video quality using the PSNR metric that is normalized to MOS. However, their work does not test spatial resolutions as a QoS factor. Moreover, [76] have presented a parametric model for estimating video quality for both SD and HD TV which does not consider video content. Joskowicz et al. [77] present a parametric mathematical model for video quality prediction that is based solely on objective quality metric and considers only low video resolutions (SD, CIF and QCIF).

2.7.2 Models based on Machine Learning Algorithms

Nowadays, people study human cognitive processes very carefully and try to elaborate models that exhibit similar behaviour to brain neurons. Since the human mind is known to be non-deterministic, it is challenging to develop a formal algorithm for human behaviour. This is the reason why researchers turn their attention to self-adaptive models and learning algorithms. In this regard, learning-based techniques including various types of machine learning have been the prime focus for the development of objective QoE prediction models [78]. One advantage of these algorithms is that they automatically learn from past observations to make accurate predictions in the future. They can adapt to changes in the QoS parameters because of their ability to learn. A number of intelligent algorithms have been proposed in the literature. However, there is still room for innovative mechanisms to efficiently correlate QoE from QoS in real time [41].

Most of the intelligent algorithms used for video quality prediction are based on artificial neural networks (ANN) [59, 79–82]. This is not ideal as neural networks are computationally complex, and require large training datasets and prolonged training time. Moreover, their reasoning processes are not transparent. The Decision Tree (DT) based learning approach was proposed in [83] for video quality prediction using bit stream information. However, DT only partially suit small datasets, small variations in the dataset require the regeneration of the tree, and the reasoning process is not completely transparent [84].

In addition, the work of [70] was further extended in [85], where the proposed models were built using two machine Learning methods: DT and Support Vector Machines (SVM) [86]. The authors found that both methods outperformed the Discriminate Analysis method which was used in [70]. The authors, however, measured the QoE in the form of "yes" or "no". They did not consider other QoE scales to predict users QoE. Further, they did not discuss how new context can be included in their method. Moreover, in [87], DT and SVM were also used to build an objective QoE model. The results were then compared with other machine learning methods including ANN, k-Nearest Neighbours (k-NN) and Random Forest (RF). RF was found to perform slightly better than the other examined methods. However, neither study considered different video resolutions or QoS network parameters.

Pokhrel *et al.* [60] presented a fuzzy logic model for QoE prediction. The estimated video quality showed a high correlation with the subjective QoE. Nevertheless, the proposed fuzzy model only considered QoS parameters from the network level, namely packet loss, burst loss and jitter. In [75, 88, 89], the Adaptive Neural Fuzzy Inference System (ANFIS) [90] was used to estimate the video QoE. The authors only considered a single video resolution, QCIF (176×144), and no other higher spatial resolutions were tested. Moreover, their video quality models used only simulated data for the video transmission and objective quality measurements.

The survey article [6] provides a comparison of different learning-based techniques in terms of their modelling capabilities, the contents of which are summarised in Table 2.2. It can be seen from this table that the fuzzy logic inference system (FIS) outperforms other estimation techniques in terms of modelling capabilities. The advantages of using FIS techniques are that they are simple, computationally less intensive and their reasoning processes are transparent. Moreover, FIS is good at making decisions with imprecise information. In general, the performance of learning algorithms depend highly on the size and coverage of the learning dataset and the number of input parameters [41].

Table 2.2: Comparison of estimation learning techniques [6]

Method	**Model Free**	**Can resist outliners**	**Explain output**	**Suits Small data sets**	**Adjustable for new data**	**Reasoning process is visible**	**Suit complex model**	**Include known facts**
LSR	N	N	P	N	N	Y	N	P
RRA	N	Y	P	P	N	Y	N	P
ANN	Y	N	N	N	P	N	Y	P
FIS	Y	P	Y	Y	P	Y	Y	Y
ANFIS	Y	P	Y	P	P	P	Y	Y
RBS	N	N/A	Y	N/A	N/A	Y	Y	Y
CR	Y	P	Y	P	Y	P	Y	N
RT	Y	Y	Y	P	Y	P	Y	P
DT	Y	Y	Y	P	Y	P	Y	P

Yes = Y, No= N, Partially = P. LSR= Least Square Regression, RRA= Robust Regression Analysis, RBS= Rule Based Systems CR= Case-Based Reasoning, RT= Regression Trees

The following Table 2.3 helps to show which aspects are evaluated by given model, as well as how many different aspects were evaluated in each model. The review of these models helps us to compare and make conclusions from the findings.

Table 2.3: A summary of the selected evaluation approaches for each model

Aspects / Model	AQoS Parameters	NQoS Parameters	Subjective Test	Objective Test	Learning Algorithms	Simulation (S) /Test-bed (T)	Environment Wired (W) Wireless(WL)	Reference Measurements
Fiedler et al [39]	x	x	x			S/T	W	RR
Siller et al [91]		x	x	x		T	W	FR
Wang et al [92]		x		x		S/T	W	NR
Agboma et al [70]	x		x			T	WL	NR
Menkovski et al [85]	x		x	x	x	S/T	WL	NR
Machado et al [82]		x	x	x	x	S	WL	FR
Du et al [93]		x	x	x	x	S/T	W	FR
Kim et al [72]		x	x			S/T	W	NR
Khan et al [74]	x	x		x	x	S	WL	NR
Han et al [94]	x	x		x		S	WL	NR
Laghari et al [95]	x	x	x		x	T	W	NR
Ramos et al [96]	x	x		x		S/T	W	NR
Koumaras et al [97]	x	x		x		S/T	W	NR
Frank et al [98]	x	x	x		x	S	W	NR
Mok et al [99]	x	x	x			T	W	NR
Elkotob et al [71]	x	x	x			T	WL	NR
Mushtaq et al [87]	x	x	x		x	S/T	W	NR
Hofeld et al [100]	x	x	x			T	W	NR
Staelens et al [83]	x	x	x		x	S	W	NR
Kang et al [79]	x	x		x	x	S	WL	NR
Cherif et al [81]		x	x		x	S	WL	NR
Pokhrel et al [60]		x	x		x	T	WL	NR

Overall, it is evident that, while a number of QoE/QoS correlation models appear in the literature, most of them provide only partial solutions to the QoE prediction problem. From the literature, we believe that machine learning-based techniques like the FIS can be valuable for QoE prediction since these techniques can discover relationships between several context and QoE parameters. FIS has been used in various telecommunications and engineering problems due to its ability to efficiently address the innate uncertainty, caused for example by the equipment's errors, or external factors. Finally, we conclude that QoE measurement and prediction over time largely remains an open area of research. In the following section, we briefly discuss the chosen learning algorithm that is used in this thesis to build a video QoE prediction model.

2.8 Background to Fuzzy Logic Inference System

Fuzzy logic is a well-known technique for user modelling that could imitate human reasoning using natural language in which words can imply ambiguous meanings [101]. It is considered as an extension to traditional set theory as statements could be partial truths, which means lying in between absolute truth and absolute falsity [102]. The idea of fuzzy logic was invented by Professor L. A. Zadeh of the University of California at Berkeley in 1965, and published a landmark paper entitled Fuzzy Sets. A number of fuzzy implementations have been used for making decisions based on imprecise/ambiguous information in various fields [103].

Typically, a FIS consists of 3 main components as depicted in Figure 2.7: fuzzifier, fuzzy inference engine, and defuzzifier. The fuzzification module receives crisp inputs and converts them into fuzzified inputs. These fuzzified inputs, called fuzzy sets, are then received by the inference engine where linguistic rules (in the form of if-then) are applied to them. The output of the inference engine is a collection of fuzzy conclusions. The defuzzication module converts these fuzzy conclusions back into crisp output [102].

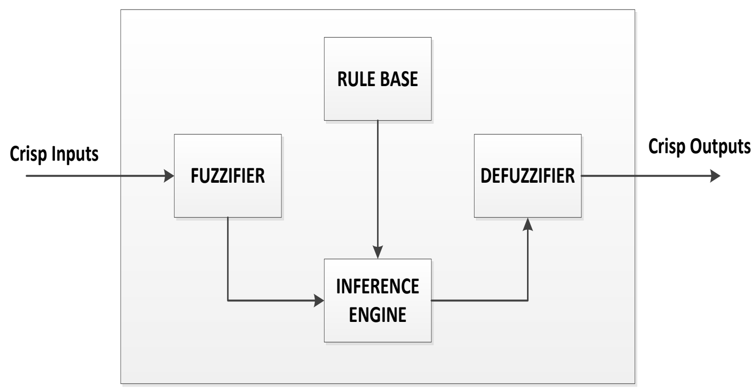


Figure 2.7: Block diagram of the FIS controller

2.8.1 Fuzzifier

Fuzzification is the first step to apply a fuzzy inference system. Generally, fuzzification involves two processes: derive the membership functions for input and output variables and represent them with linguistic labels. This process is equivalent to converting or mapping classical set to fuzzy set to varying degrees. The fuzzy set is converted into an equivalent form (shape) of the membership function. The curve values of the membership functions represent the degree to which a particular input parameter value belongs to the output fuzzy sets [104].

The membership functions can take different forms: triangles, trapezoids, bell curves, or any other shape so long as the shape accurately represents the distribution of information. For example, the triangular shapes of the fuzzy sets could be characterized by three values: the right boundary R, centre C, and left boundary L. For a crisp input value x, the membership value of x on a triangular fuzzy subset A was calculated using the formula [102]:

$$\mu_A(x) = \begin{cases} (x-a)/(b-a) & a \leq x \leq b \\ (c-x)/(c-b) & b \leq x \leq c \\ 0 & otherwise \end{cases} \tag{2.1}$$

Figure 2.8 shows a fuzzy set with five membership functions in a triangular shape. Although most fuzzy sets have an odd number of labels, a set can also have an even number of labels. For example, a fuzzy set may have four or six labels in any shape, depending on how the inputs are defined in relationship to the membership function [102].

The core of a fuzzy set is the set of elements whose degree of membership in that set is equal to 1, which is equivalent to a crisp set. The boundary of a fuzzy set indicates the range in which all elements whose degree of membership in that set is between 0 and 1 (0 and 1 are excluded) [104]. After the membership functions are defined for both input and output, the next step is to define the fuzzy control rule.

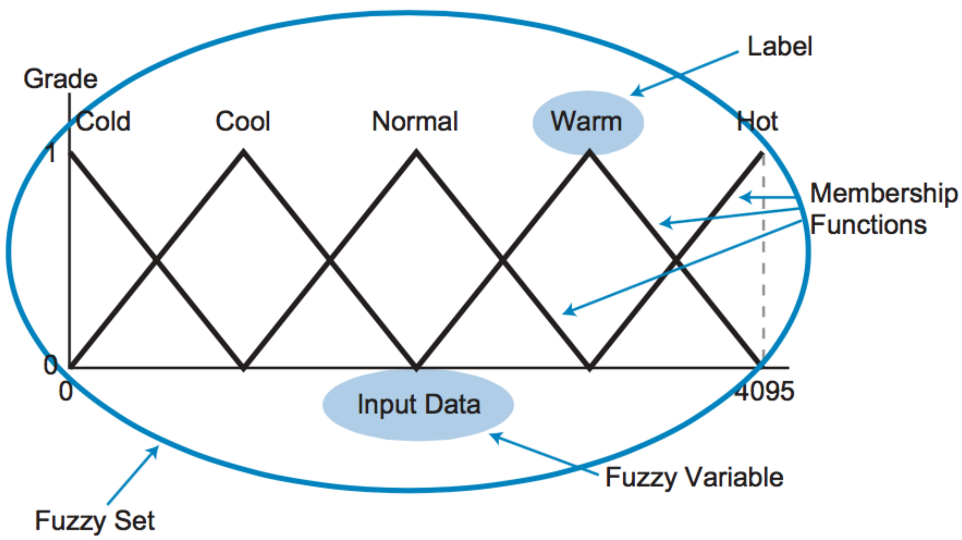


Figure 2.8: Membership function chart (triangular shape)

2.8.2 Fuzzy Inference Engine

In artificial intelligence, knowledge can be represented in different forms. One way to express human knowledge is by using natural language expressions of the form:

IF premise (antecedent), THEN conclusion (consequent)

Fuzzy logic uses a reasoning, or inferencing, process composed of IF...THEN rules, each providing a response or outcome. A fuzzy rule is a simple IF-THEN rule with a condition and a conclusion [102]. The fuzzy inference engine contains a collection of IF-THEN rules, obtained from experts or learnt using automated methods, such as Learning From Example (LFE) method. In this thesis, both techniques are used and will be discussed later.

The inputs taken from the fuzzifier (i.e., membership values) are applied to the antecedents of the fuzzy rules, and the obtained value is then applied to the consequent membership function (i.e., the output). The fuzzy rules and the combination of the results of the individual rules were evaluated using fuzzy set operations such as AND (intersection) and OR (union) [101]. Basically, a rule is activated, or triggered, if an input condition satisfies the IF part of the rule statement. This results in a control output based on the THEN part of the rule statement. Sometimes, more than one rule is triggered at a time in the FIS controller process. In this case, the controller evaluates all of the triggered rules in order to arrive at a single outcome value and then proceeds to the defuzzification process.

Figure 2.9(a) illustrates an example of two fuzzy inputs, X1 and X2, and one fuzzy output, Y1. The rules shown in Figure 2.9(b) represent four of nine possible rules that cover the two inputs. The four shown, however, cover the four possible triggering points for the two input readings, X1 and X2. Given the input values in Figure 2.9(a), the inputs will trigger rule 1 because X1 = ZR AND X2 = NL. This will generate two outputs for Y1 = NL, one at a grade of 0.6 (due to the input value of X1) and the other at a grade of 0.75 (due to the value of X2). In a fuzzy logic situation where a two-input rule with an AND relationship produces two outcome values, the controller will choose the outcome with the lowest grade, in this case 0.6NL, as shown in Figure 2.9(c) [3].

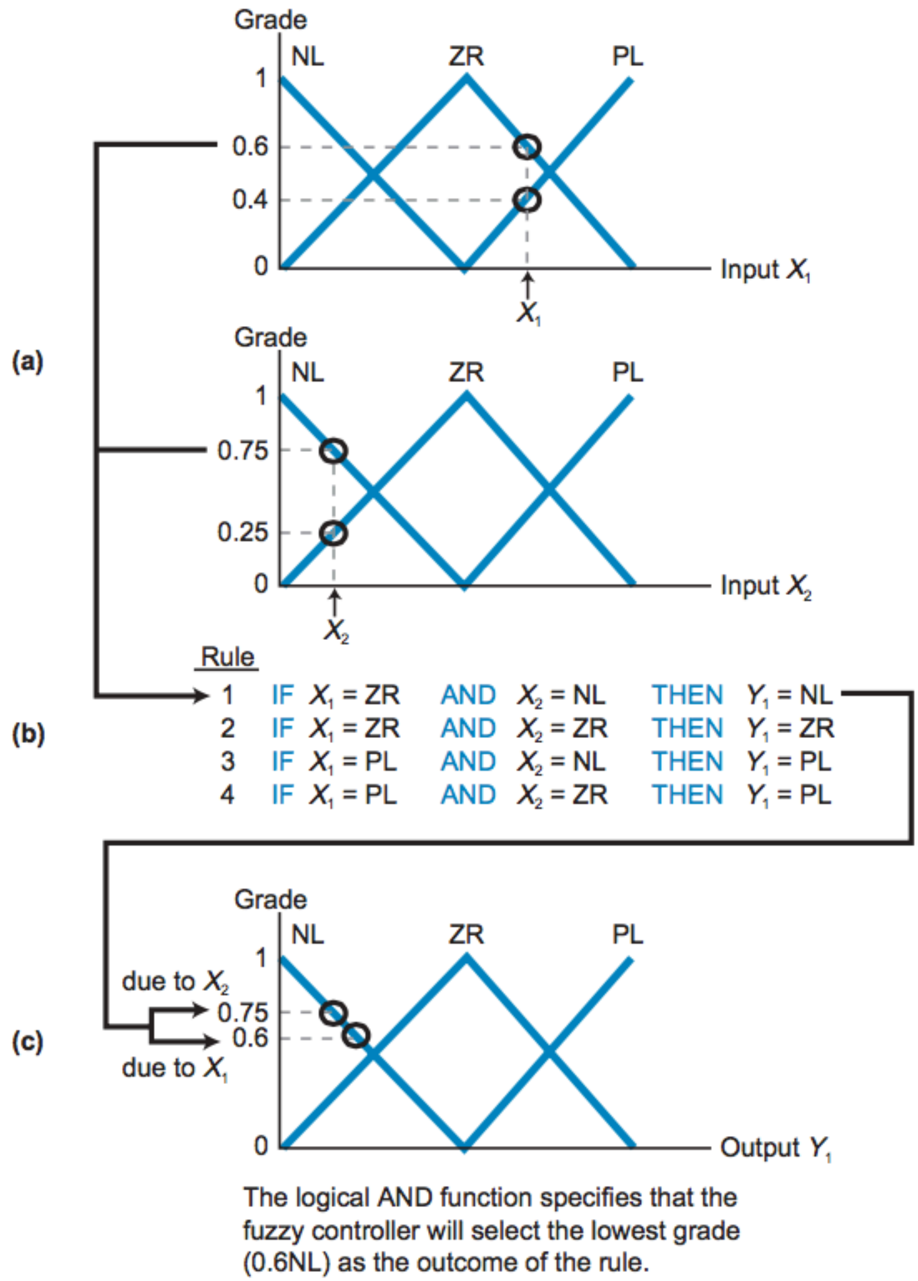


Figure 2.9: Fuzzy processing example showing (a) two fuzzy input values, (b) the four rules that they trigger, and (c) the resulting output [3].

2.8.3 Defuzzifier

Defuzzification is the reverse process of fuzzification. The fuzzifier converts a precise quantity to a fuzzy term. Likewise, the defuzzifier converts a fuzzy term back to a precise quantity. The defuzzification process examines all of the rule outcomes after they have been logically added and then computes a value that will be the final output of the fuzzy controller. The output of a fuzzy process can be the logical union/intersection of the membership functions. There are different defuzzification techniques such as the centroid method, the weighted average method, and the maximum method [101]. The centroid method, which is adopted in this work, relies mathematically on the centre of gravity (CoG), and is expressed by the following formula:

$$y(x) = f_s(x) = \frac{\sum_{i=1}^{M} y^1 \prod_{l=1}^{n} \mu F_l^i(x_l)}{\sum_{i=1}^{M} \prod_{l=1}^{n} \mu F_l^i(x_l)}. \tag{2.2}$$

Here, M is the rule number in the rule base, y^{-i} is the centroid of the ith output fuzzy set B^i, and $\Pi_{l=1}^{n} = \mu F_l^i(x_l)$ is the product of the membership values of each rule's inputs.

2.9 Chapter Summary

This chapter presented background information about QoS/QoE correlation and the KPIs for video streaming. Different methodologies and techniques for video QoE measurement were discussed. In addition, a number of existing QoE/QoS correlation models for the prediction of video quality were critically reviewed in order to obtain a "broader picture" of this area of research.

It is evident that, while a number of QoE/QoS correlation models appear in the literature, most of them provide only partial solutions to the QoE prediction problem. As such, some models are too specific for a particular kind of application, and each model has varying computational and operational requirements. From the literature, models that use intelligent machine learning techniques such as FIS outperform the other methods. Moreover, the existing proposals for video QoE prediction tend to consider either the video encoder's parameters, network impairments, or the features of the video content, but rarely all three in conjunction. There is a need for a novel hybrid video QoE prediction system that extends the existing work by enlisting a group of QoS parameters that has not been addressed so far in the context of video QoE prediction. The next chapter introduces and discusses the proposed video QoE evaluation framework in general terms.

Chapter 3

Methodology

This chapter introduces and discusses the proposed video QoE evaluation framework in general terms. The proposed framework exploits the relationship between QoS parameters and video quality to objectively predict the QoE. In order to understand how to exploit this relationship, a fundamental investigation of the relationship between video QoE and QoS parameters associated with the encoder, access network and content types, has been undertaken to quantifying their relative impacts. After the intelligent machine learning algorithms learn about the relationship between the QoS parameters and the QoE, they are used as an objective tool for QoE prediction. Figure 3.1 shows the components of the proposed video QoE evaluation framework. Each component will be described in the following subsections. Additional details about the framework are explored in the indicated chapters of this thesis.

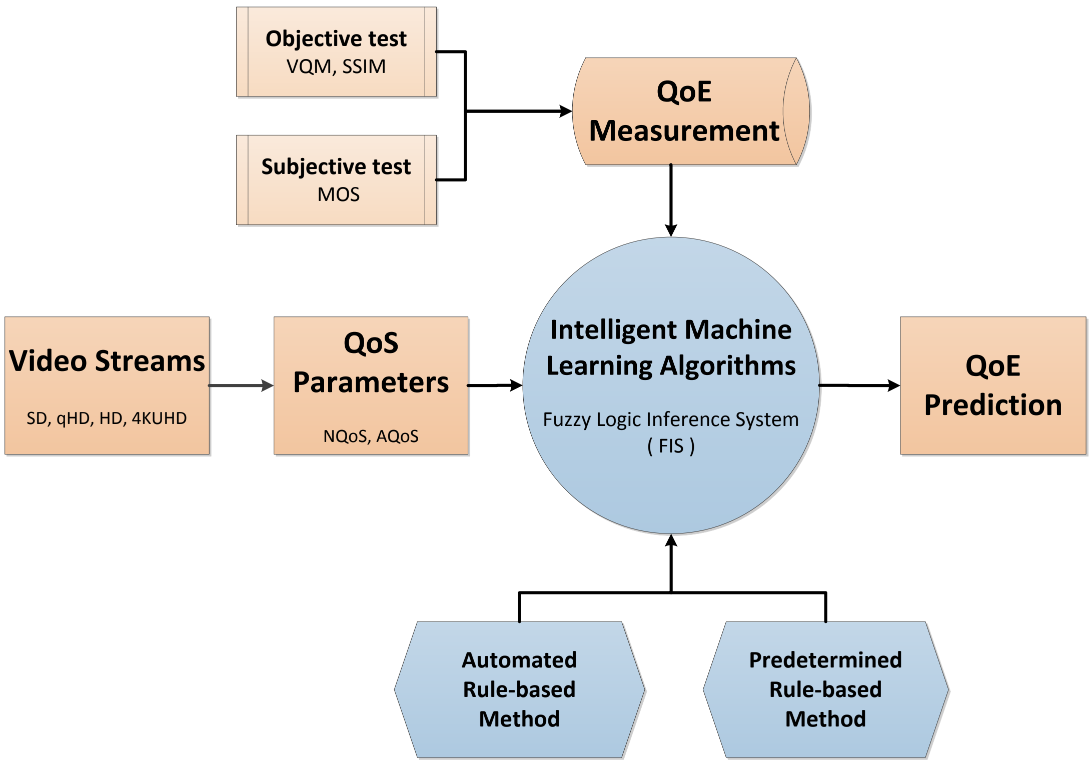


Figure 3.1: Components of the Objective QoE Evaluation Framework

3.1 Video Services

The dominance of video services, including 2D and 3D video streaming, on the internet continues to drive the evolution of internet access methods and core carrier networks. Consumers also own a growing number of video-capable devices, from mobile phones to UHD, and even 4k-capable, television sets. The popularity of the video applications introduces a higher demand for quality. Therefore, estimating the

QoE is a must for application/service providers in order to retain the highest possible service quality. The proposed work in this thesis focused on predicting the QoE of the video streaming service. In this thesis, a number of different video content types and resolutions are used for performance evaluation. In Chapter 4, the evaluated video streams were encoded by the H.264/AVC and H.265/HEVC video coding standards. Furthermore, the impact of QoS parameters on video quality in terms of different video content types and resolutions is investigated and analysed. For instance, "head & shoulder" video streams (such as in news broadcasting) declare acceptability thresholds that are substantially different from those used in the case of streaming an action movie. Further details of this investigation will be discussed in different chapters.

3.2 QoS Parameters

In the research community, QoS is known to be the most influential factor on QoE [18, 19]. The ability to identify the perceived degree of video impairment due to QoS parameters in video bit-streams is one of the key aims for video service providers. QoS is defined in terms of QoS parameters. These parameters are generally used to represent network layer quality, but they can also correspond to different OSI layers. In this thesis, the predominant QoS parameters and their ranges of significance are selected from different QoS layers. In this thesis, we proposed a more practical approach to investigate the effects of different QoS parameters, including packet loss, packet structure, compression efficiency, video codecs, content type, and spatial resolutions, on the QoE for video streaming. An understanding of the perceptual effects of these key parameters on video quality is important as it forms the basis for the development of video quality prediction models. We call the video content type and the parameters related to the encoding process (sender bitrate, quantization parameter (QP) and resolution) Application QoS parameters (AQoS). The access network parameters (packet loss, burst loss, packet duplication and reordering), grouped together, are referred to as Network QoS parameters (NQoS). There are other different QoS access network parameters that can be investigated on future work such as end-to-end delay, jitter, throughput, etc. It is never possible to examine every QoS parameters but the target on this thesis is to cover QoS parameters from different QoS layers to judge the likely impact of QoS on video QoE.

Today's multimedia applications are expected to run in physically heterogeneous environments composed of both wired and wireless components. Wireless links exhibit distinct characteristics, such as limited bandwidth, varying error-rates and potential hand-off operations. Consequently, QoS requirements in wireless networking are stringent and complicated, taking additionally into account the influencing mobile device characteristics and limitations. Because of these challenges of the wireless video streaming, both simulation and test-bed experiments were conducted in this thesis work using wireless environment. Furthermore, the wired environment was also used in this thesis work in order to present that the proposed model can be used on heterogeneous environments.

3.3 QoE Measurement

QoE is a broad concept that encodes users' levels of satisfaction with a product. Of the various aspects that constitute QoE, the user's perception of the content consumed is considered the most influential [19]. In order to develop a concrete objective QoE prediction model, a huge QoE dataset and test conditions were constructed for the proposed work in this thesis, using both objective and subjective methods for quality assessment. Each method of QoE measurement has its own shortcomings. To overcome these, various hybrid approaches that combine both subjective and objective methodologies are proposed. Later, the QoS/QoE mapping datasets are statistically analysed using 5-way ANOVA to confirm the impact of each chosen QoS parameter and to identify the most influential parameters.

3.3.1 Objective Method

QoE measurements using the objective method employ video quality metrics. In this thesis, two types of video quality metrics are used, which are VQM and SSIM. The VQM was independently evaluated by the video quality experts group (VQEG) [55]. The SSIM was developed by Wang et al [56]. Both quality metrics are normalized to the subjective MOS scores, using a method based on the one described in [77].

3.3.2 Subjective Method

The subjective method is the most accurate technique for measuring perceived video quality. However, the large number of test conditions required to formulate the proposed video QoE evaluation system make it extremely difficult to conduct subjective tests in which video sequences are assessed by viewers. Consequently, the main use of subjective assessment is to validate the measured objective scores, and to assure their credibility so that they can be used confidently for video QoE prediction. In this thesis, a standard subjective laboratory environment was utilized for the QoE evaluation of H.264/AVC video streams. A panel of 21 viewers of varying experience was selected as test subjects. After a training sequence was completed, the evaluations were conducted during individual test sessions held in a lab under controlled environmental conditions.

3.4 Intelligent Machine Learning Algorithms

Intelligent machine learning algorithms automatically learn from past observations in order to make more accurate predictions in the future. They have been the prime focus of researchers developing objective QoE prediction models [87, 105]. Some of the most popular machines learning algorithms are fuzzy logic inference system (FIS), decision tree, neural networks, etc. According to [6], the FIS outperforms other estimation techniques in terms of modelling capabilities and making decisions with imprecise information. In this thesis, the proposed video QoE evaluation framework used FIS to learn and map the correlation between QoS parameters and measured QoE. The learning process involves training the algorithm on the measured QoE dataset. Once the system has learnt, it can predict the QoE, based on any combination of input QoS parameters. The proposed work provides two novel and efficient reference-free models for the prediction of video quality in terms of different content types and resolutions. The first FIS model is based on a predetermined rule-based method (FIS-PRB). A semi-manual approach was applied to develop the FIS-PRB model, combining human knowledge and the behaviour of QoS parameters with testing on a simulator. The model is simulated in MATLAB. The second FIS model is based on an automated rule-based method (FIS-A). It is a self-learning system which enables the adaptive generation of fuzzy rules from the QoE dataset. The model was implemented in the Java programming language. The adaptive FIS-A model was evaluated on a real test-bed which was produced for the QoE prediction of real-time wireless H.265 video streaming. Further details of this model will be given in Chapter 5 and 6.

3.5 QoE Prediction

In this thesis, the output QoE scores are represented in terms of MOS scores. This MOS score is as close as possible to the mean score obtained from the objective and subjective tests. The predicted QoE output (MOS) is compared with the subjectively and objectively measured QoE through the calculation of the correlation coefficient and the Root Mean Squared Error (RMSE).

3.6 QoE-enabled Applications for Video Delivery

To demonstrate the benefits of the video QoE evaluation framework, two practical examples of QoE-enabled applications for optimising video delivery are given in Chapter 7. The first application is a QoE-enabled transport optimisation scheme for real-time SVC video delivery. The second application is a QoE-enabled resource utilisation scheme for mobile video delivery. These applications show how QoE is used to optimise video delivery and utilize existing network resources according to users' QoS requirements. QoS/QoE correlation can be used as an indicator for network management and planning processes in order to avoid resource over-provisioning.

3.7 Chapter Summary

In this chapter, we presented the proposed video QoE evaluation framework for predicting the QoE of different coded video streams. This framework learns about the QoS/QoE correlation in order to objectively predict the QoE using the FIS method. Furthermore, it incorporates both adaptive and non-adaptive approaches to FIS. The QoS parameters are associated with different layers of the OSI model. The framework utilizes laboratory based subjective and objective tests for correlating QoS parameters with the measured QoE. The proposed framework will later be applied to estimate the QoE of H.264/AVC video streams and 4kUHD H.265/AVC video streams. Later in this thesis, two examples of QoE-enabled applications for enhancing the quality of video delivery will be presented in order to demonstrate the benefits of the proposed video QoE evaluation framework.

Chapter 4

Studying The Impacts of QoS Parameters on Video QoE According to Content Types and Resolutions

4.1 Introduction

The trend towards video streaming with increased spatial resolutions and dimensions, SD, HD, 3D, and 4kUHD, even for portable devices has important implications for displayed video quality. There is an interplay between packetisation, packet loss visibility, choice of codec, and viewing conditions, which implies that prior studies at lower resolutions may not be as relevant. This chapter presents two sets of experimental studies, the one at a Variable BitRate (VBR) and the other at a Constant BitRate(CBR), which highlight different aspects of the interpretation. CBR means that each frame uses the same amount of bits regardless of whether it needs them or not. VBR means that you can vary the amount of bits used to represent a frame so that the overall average amount of bits-per-frame is achieved. One of the aims of the proposed studies on this chapter is investigating the impact of QoS parameters on the video streaming whether on VBR or CBR encoding modes.

The second experiment also compare and contrast encoding with either an H.264 or an High Efficiency Video Coding (HEVC) codec, with all results recorded as objective Mean Opinion Score (MOS). It is never possible to examine every configuration of wireless video streaming but in the two sets of experiments there is scope for extrapolation by the reader to a configuration of interest in order to judge the likely impact of changing spatial resolutions.

The remainder of this chapter is organized as follows. Section 2 presents the related work on video QoE evaluation. Section 3 discusses the experimental set-up and results of the first study. The second study is presented in section 4. Section 5 summarizes the Chapter.

4.2 Related Work

The majority of the existing video quality evaluation studies for both the H.264/AVC and HEVC codecs consider only low resolution video coded streams, and also rarely consider parameters associated with both the AQoS and the NQoS parameters. The research in [74] considered both wireless content dependency and network impairments, but the set of test videos had a resolution as low as QCIF (176×144 pixels/frame) with only a few packets per frame. The frame rate was also as low as 10 fps. The QoE was measured using the PSNR metric, which is not directly related to human perception [106]. In fact, the authors of [74] concede that video-frame resolution does have an impact on overall quality, even for low-bitrate, as presented in [106].

Much prior research has been conducted on lower resolution imagery (QCIF, CIF (352×288 pixels/frame), and VGA (640×288)) for which the relationship between the packet number and the frame size is very different from the one for HD. Nevertheless, in [107] the packet loss and bitrate were found to be more important than the frame rate for the subjective quality of lower-resolution videos. In addition, the authors of [108] presented an analysis of packet loss for compressed video. Their results showed that, in general, burst losses produce more distortion than an equal number of isolated losses. However, in [109], a contradictory result was obtained using a set of subjective experiments.

The impacts of packet structure and compression efficiency on different video resolutions have only rarely been investigated in the literature. In [110], a quality assessment comparison of H.264/AVC and MPEG-2 was presented. The results demonstrated that H.264 quality drops steeply for even low packet loss rates (0.02%), while MPEG-2 quality drops by much less. However, the two video codecs were evaluated using different bitrates, and only for HD resolution. Also, they only considered one type of video content.

The authors of [111] approached the impact of packet loss from the point of view of being able to predict which MPEG-2 Transport Stream (TS) packet losses within a video frame would be visible to subjective assessors. Amongst their interesting conclusions was the finding that the wider field of view made HD packet loss more visible than SD packet loss. The loss of an entire frame was also more visible in HD than in SD. They encoded the SD and HD video streams in a CBR mode. On the other hand, work in [112] argues that the degradation caused by a packet loss in HD is significantly lower than in SD. However, they did not mention the bitrate size or whether it was CBR or VBR video streams. Moreover, both studies conducted the comparison with only one type of video content.

Nightingale *et al* [113] discussed the impact of network impairment on HEVC encoded video streams at resolutions below HD. Their framework has since been superseded due to changes in the structure of the H.265 network abstraction layer (NAL) unit. The same framework was used in [62] and [114] to study the impact of network impairment on H.265 encoded video streams at resolutions below HD. Anegekuh *et al.* [65] carried out objective and subjective tests on video sequences to investigate the impact of video content type and encoding parameter settings on H.265 video quality. Their initial results showed that varying the video content type and encoding parameters has an impact on video quality. They did not consider NQoS parameters and the evaluated video was only at HD resolution.

The evaluation of 3D video QoE is more challenging because additional factors such as depth perception, comfort levels, and naturalness need to be considered. The effect of random packet losses on the overall 3D perception was studied in [115] using a subjective test. They found a negative trend in 3D perception when packet loss rates increase. Nasseralla *et al.* examined the quality degradation of 3D video transmitted over mobile networks through subjective tests in [116] and [117]. They investigated the effects of random packet losses on the overall 3D perception (i.e., distortions due to different packet loss rates). None of the above-mentioned studies considered the 3D video evaluation of VBR streams at HD resolutions.

It can be concluded from the related work presented in this section that video QoE is very sensitive to different values of the AQoS and NQoS parameters. These correspond to different levels of video impairment, and, consequently, to different QoE values. There are numerous features of QoS parameters, either content-dependent or content independent, that can influence their impact on the video quality. This makes it necessary to investigate the correlation between video QoE and QoS parameters associated with the access network, encoder, content type, and video resolutions beyond HD. It is also necessary to study the new video codec H.265/HEVC, since it provides a reduction in bandwidth and will be favoured as a codec for real-time video delivery.

The contributions of this chapter are twofold:

- Study 1: Investigated the impact of QoS parameters on the video QoE by cross-layer simulation of the transmitted 2D and 3D video streams. The coded videos were based on a Variable Bitrate (VBR). The video quality was measured using the objective (VQM) and subjective (MOS) metrics. The QoS/QoE mapping dataset was statistically analysed with 5-way ANOVA to confirm the impact of each chosen QoS parameter and to identify the most influential parameters.

- Study 2: Investigated the impact of QoS parameters and compression efficiency on up-to 4kUHD of the H.264 and H.265 coded videos, both transmitted and decoded simultaneously in real-time. The coded videos were based on a Constant Bitrate (CBR). The video quality was measured by the objective metric (SSIM).

The two studies are discussed in the following sections. The results of the experiments conducted for this chapter will be used in Chapters 5 and 6 for the construction of the objective non-reference models for perceptual video quality prediction.

4.3 Study 1: Impact of QoS parameters on 2D and 3D VBR Video Streaming

QoS parameters such as packet loss are recognized as a serious threat that can greatly prejudice the user's experience of a service. It turns out that there is a content-dependency aspect of video streaming, both in terms of content type and compression ratio, as mediated by the quantization parameter (QP). The QP determines the bitrate and, hence, the bandwidth required. The implication is that video service providers should carefully consider the content type and its QP when video is targeted at portable devices, along, of course, with considering techniques such as channel coding [118] to protect the video stream.

The contribution of this study is to examine the potential impact of QoS parameters across SD, HD and 3D video streams. Realistic video configurations were, thus, used and the extent of the influence of the content type was checked. Further, the need to carefully balance the QP/video quality against the effects of packet loss is justified in this study. The QoE measure of video quality employed herein was one that is known to approximate subjective assessments of quality, namely, the Video Quality Metric (VQM) [55]. Because the VQM was employed, the correlation between the packet-loss rate, the error burst length, and the content type was able to be directly mapped to MOS subjective ratings. Before doing so, however, subjective testing was used to check that the correlation between VQM adjusted to an MOS scale and HD video did, indeed, still exist at the higher spatial resolution.

4.3.1 Experimental Set-Up

4.3.1.1 Video Encoding

Three classes of SD/HD/3D video content formed on the basis of temporal activity were tested. The spatio-temporal classification metric from recommendation ITU-T P.910 [52] was used for this purpose. This technique extracts spatial and temporal features from a video sequence, and then assigns a spatial index (SI) and a temporal index (TI). The computed indices indicate the content type of each sequence. This technique is of low complexity and, thus, can classify videos in real-time.

Three video sequences were chosen for analysis, one from each class, as listed in Table 4.1. The video sequences were assessed in progressive SD, HD (1024×720 pixels/frame, i.e. 4:3 aspect ratio), and in 3D plus depth format in SD and HD. Figure 4.1 shows a sample frame and depth image for the 3D version of the sequences. Each sequence of length 200 frames (or 8 s in time) was captured in YUV 4:2:0 chroma format, at 25 fps, as this is typical for wireless video streaming [119]. Both the color image and the depth map were encoded with the H.264/Advanced Video Coding (AVC) [47] Joint Model (JM) reference codec [120] at the same QP. The Group of Picture (GOP) size was 16, where each group included one I-frame and all remaining frames were P-frames. This reduced the computation time arising from bi-predictive B-frames, and is the structure recommended for wireless video streaming [119].

Table 4.1: Video sequences chosen and classes assigned

Video Sequence	**TI**	**SI**	**Class**
Music	4.90	74.41	Low Motion
Poker	12.20	85.69	Moderate Motion
BMX	22.35	99.42	High Motion

Figure 4.1: One representative frame from each of the 3 source video sequences

Within the encoding process, network abstraction layer (NAL) units were RTP packetized. It is also assumed that the RTP packet was encapsulated in a UDP packet, and then in an IP packet on the network layer. Here the packet loss rate (PLR) denotes the loss of video slices (NAL units). The choice of UDP packet size matched the Ethernet maximum transmission unit (MTU) of a regular computer (1472 bytes payload). The codec output was provided at a Variable Bit Rate (VBR), according to the QP. VBR allows the quality setting, i.e. QP, of the video to be judged against the resulting bitrate.

4.3.1.2 Simulation Scenario and QoS Parameters

The simulation scenario is depicted in Figure 4.2. As previously mentioned, video sequences were assigned a class identifier according to their temporal activity. After coding with H.264/AVC outputting in the RTP mode, wireless transmission errors were introduced into the coded video packets by the simulator. Error concealment took the form of previous frame replacement. More sophisticated error concealment methods were avoided in order to allow comparisons with the results of others and to reduce the latency arising from some error concealment methods. The quality of the decoded video frames was then assessed using objective and subjective quality metrics to give an arithmetic mean over the sequence's frames.

The wireless channel was simulated both by the packet loss ratio (PLR) and by introducing mean burst losses (MBL) into the transmitted packet stream, in order to analyse a broader range of simulation conditions. The PLR was uniformly distributed along the packet loss trace, while the MBL was distributed as bursts with a mean burst length of MBL along the packet loss trace. Packet loss traces were generated, based on the Gilbert-Elliot model [121] (a two-state Markov chain model), with varying levels of the PLR and the MBL. Gilbert-Elliott type models do not emulate the physical channel, but do accurately model the application receiver's experience of packet losses resulting from fast fading [122]. The details of the Gilbert-Elliot model employed are outlined in Appendix A.

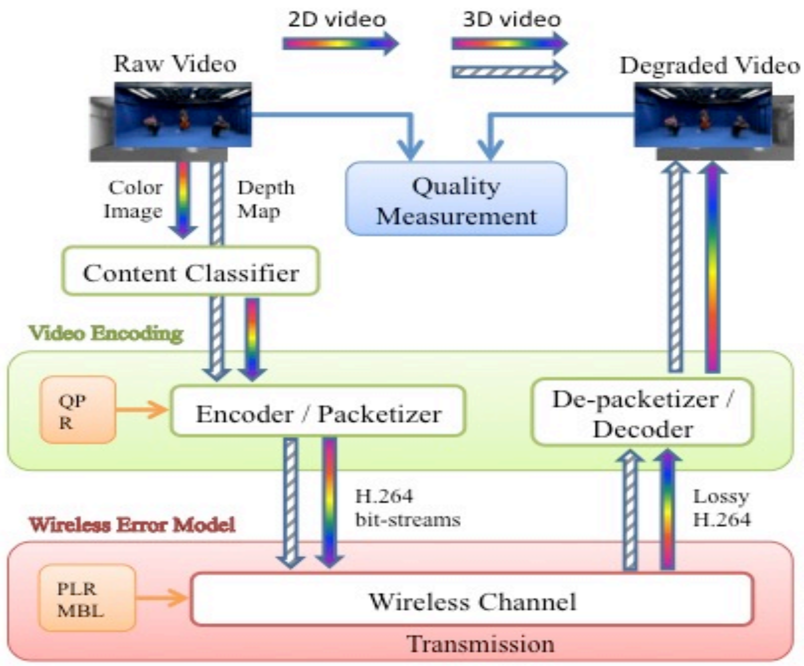


Figure 4.2: Conceptual illustration of the simulation scene

Table 4.2 presents the chosen and simulated AQoS parameters (CT, R, QP) and the selected NQoS parameters (PLR, MBL). The parameter values were selected carefully in order to generate a broad range of quality levels (QoE). For increased data confidence, the simulation of each tested condition was repeated ten times, so that the error trace began at a different displaced position of the coded bit-stream each time. This resembles real-life communications where errors can occur at any given point of the transmission time. It also ensured that 2000 video frames were considered for each simulation condition. Once the QoE of the 10 received videos was measured, the mean QoE and a 95% confidence interval were calculated.

Table 4.2: Simulated QoS Parameters

Parameters	**Values**
Content Type (CT)	Low, Moderate, High motion
Spatial resolution (R)	SD (720 × 540), HD (1280 × 720)
Quantization Parameter (QP)	16, 24, 32, 40, 48
Packet loss ratio (PLR)%	1, 2.5, 5, 7.5, 10
Mean burst length (MBL)	1, 2.5, 5, 7.5

4.3.2 Video QoE Measurement

QoE is a broad concept that encodes the user's level of satisfaction. Of the various dimensions that constitute QoE, the user's perception of the quality of the content consumed is considered the most influential [19]. Objective and subjective quality metrics were used as quality assessment methods in this study.

4.3.2.1 Objective Method

The objective measurements were conducted using a validated FR perceptual 3D video quality metric [123]. This metric adopts the NTIA General Model [55] for the assessment of 2D color images. Referred to in the literature as the video quality metric (VQM) [55], the NTIA General Model was independently evaluated by the video quality experts group (VQEG) and standardized by ANSI [124] and ITU [57]. For the 3D videos, the depth map was assessed using the depth quality model [123], which measures the quality of the depth signal based on the identification of dominant depth planes. The compound 3D quality was then determined by employing a joint mathematical model [123] which combined the measured VQM of the 2D color image with the corresponding depth map.

The VQM is measured on a continuous scale from 0 (complete loss) to 1 (original quality). The use of this 3D quality metric, which adopts VQM within its engine for the 2D component, makes the use of several methods applicable to VQM analysis available for rating 3D video quality. For example, 2D quality measurements of the same video sequences are made available within the collected 3D dataset. Moreover, this procedure makes it possible to map the 3D quality scale to the subjective Mean Opinion Score (MOS). MOS has five grades ranging from 1 to 5, which is the highest quality [52]. The assessed quality can be normalized to MOS [77] by means of the equations:

$$MOS = 5 - 4VQM \tag{4.1}$$

$$MOS = 5 - 4(1 - Q) \tag{4.2}$$

Equation 4.1 was employed for VQM because, on the VQM scale, 1 represents severe distortion and 0 represents original quality, whereas equation 4.2 was applied to Q, the 3D quality metric, as that scale's range is the inverse of the VQM scale. We assumed that, in practice, the VQM ranged from 0 to 1, rather than to 1.2. Making this assumption allowed us to convert Q to the VQM range by subtracting Q from 1.

4.3.2.2 Subjective Method

A subjective assessment test was conducted to assure the credibility of the objective assessment findings in this study. The measured subjective dataset will also be used for model validation in Chapters 5 and 6. The standard recommendation ITU-R BT.500-13 [51] was followed for this test. Because the total number of test conditions for the measured VQM dataset was 1080, a systematic approach was followed to select a subset of 64 video sequences for subjective testing. The 2D HD version of each sequence was chosen to perform a balanced selection of conditions that spanned the quality scale (0-to-1). This validation can be applied to 3D objective scores since the methods from ITU-R BT.500-13 are also applicable in 3D scenarios [125].

The selection approach used was based on the Kennard and Stone algorithm [126], which selects as the next sample the one that is most distant from those samples already selected. Thus, the selection covered the experimental region uniformly, yielding a flat distribution of the input data. This guaranteed that each value of each QoS parameter was achieved over the whole sample space. 20 samples were selected of each video content type, in addition to the reference video. A panel of 21 viewers were employed as test viewers for a single stimulus (SS) quality evaluation method, performed in a lab under controlled conditions. 2D videos were displayed on a 47" LED monitor, and the users marked their MOS responses on a continuous scale between 1 and 5 as illustrated in Table 4.3. The MOS scores for the 21 observers and their corresponding 95% confidence intervals are shown in Figure 4.3.

Table 4.3: Subjective Mean Openion Scores

Quality	Bad	Poor	Fair	Good	Excellent
MOS	1	2	3	4	5

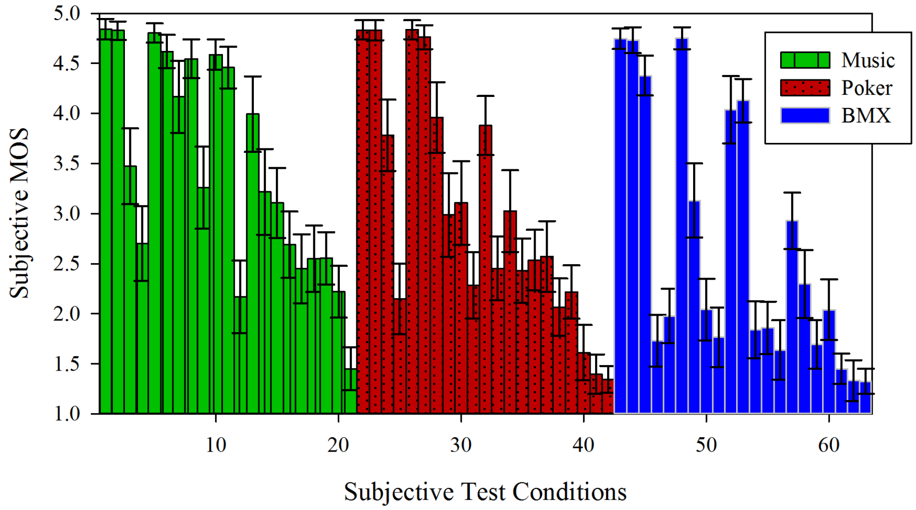


Figure 4.3: Subjective MOS scores with 95% confidence intervals

4.3.2.3 Correlation of Objective and Subjective Scores

As a validation measure for the objective scores obtained through simulations, the correlation between the subjective QoE and the objective QoE for the 2D videos is presented in Figure 4.4 and expressed in statistical terms using the Pearson Correlation Coefficient (PCC). A PCC of 0.92 indicates a high level of correlation and acknowledges the validity of the collected objective QoE dataset. A brief description about the PCC is given in Appendix B.

Furthermore, to provide a visualisation of the correlation between the three datasets (2D subjective, 2D objective, and 3D objective), Figure 4.5 portrays a comparison of the 21 test conditions for the "Poker" HD video with the scored MOS in each of the three datasets constructed. The 2D objective and 3D objective datasets were found to be highly correlated with a PCC of 0.99. Consequently, it can be concluded that the validation of the 2D objective dataset can be extended to validate the 3D objective dataset. Similar figures for the comparison of the "Music" and "BMX" videos are presented in Appendix C.

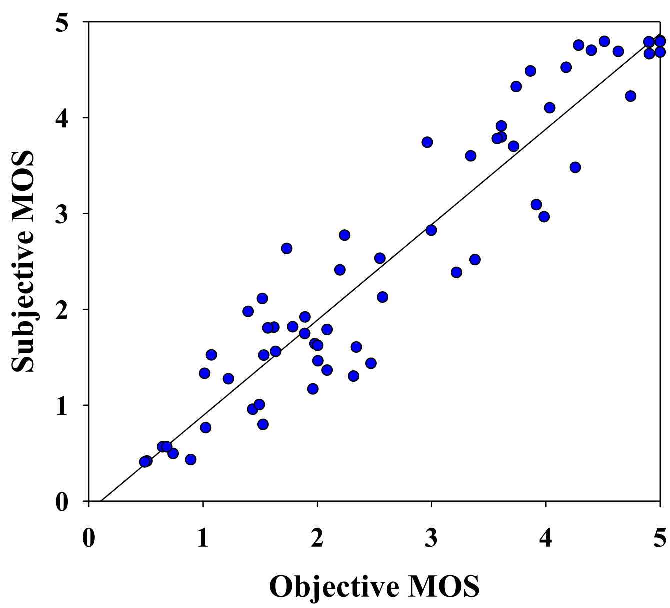


Figure 4.4: Correlation of subjective MOS and objective MOS

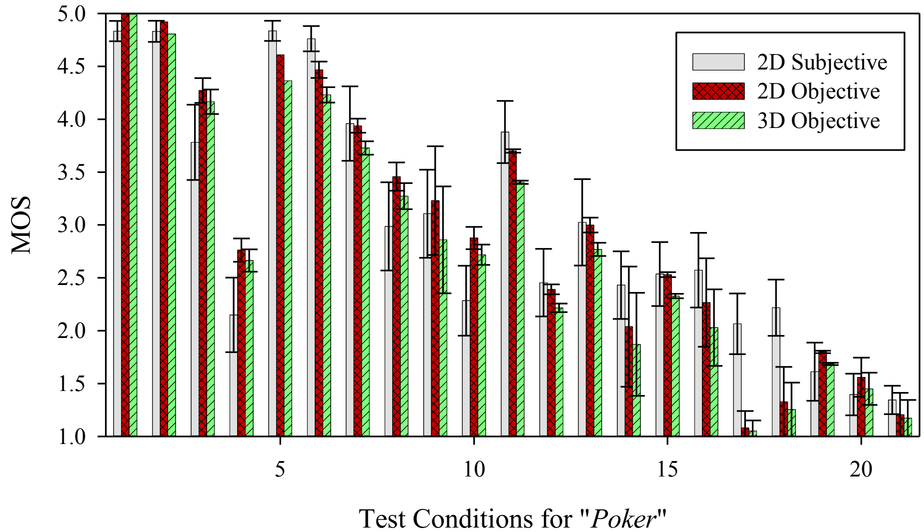


Figure 4.5: Comparison of the three datasets for Poker HD video sequence

4.3.3 Results and Discussion

For calibration with later results, coding loss with zero packet loss was assessed, as shown in Figure 4.6. As might be expected, the video quality was reduced with higher QP. The effect of including a depth image sequence was to degrade the quality (to some extent) in all cases. At the lowest QP (lowest compression ratio) it was difficult to distinguish between the three example videos. At the two higher QPs (greater compression), the temporal activity was an indication of the quality assessment. For example, the Music video at QP = 32 was of lower assessed quality than the high motion BMX video. An important finding was that the metrics (2D and 3D) reported very similar qualities regardless of whether the spatial resolution was SD or HD.

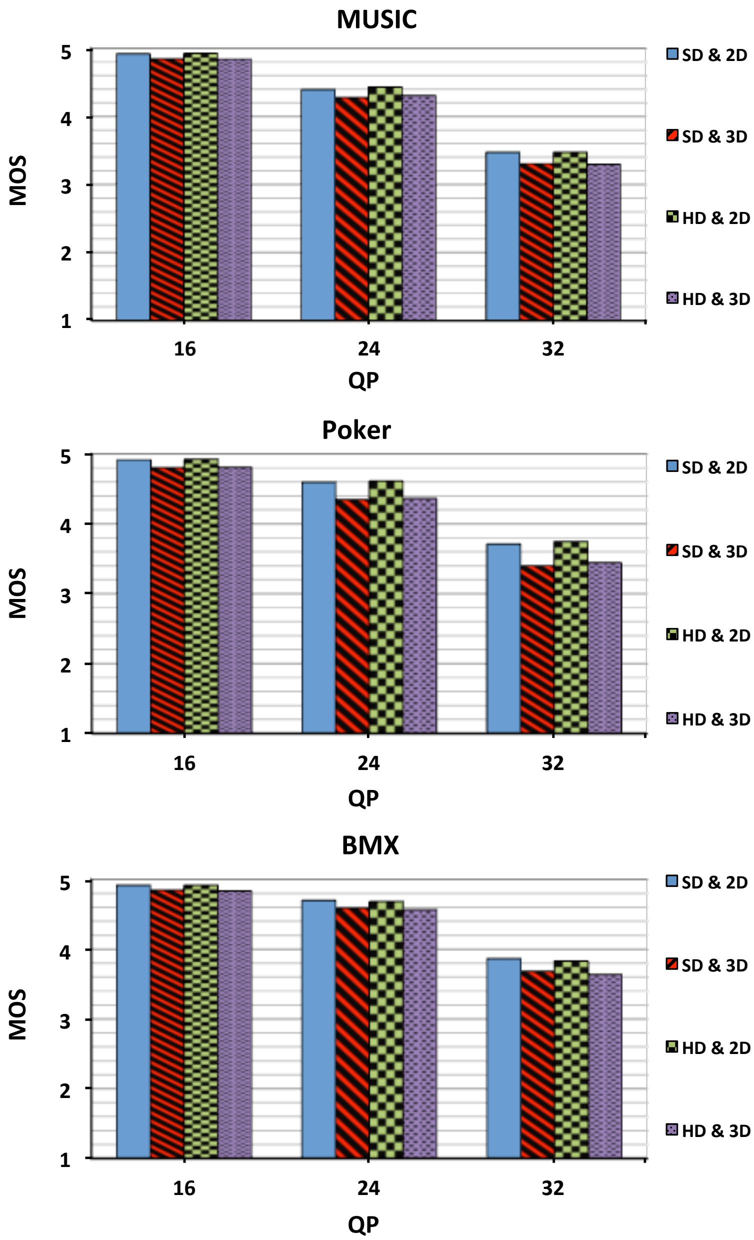


Figure 4.6: MOS of SD, HD, and 3D video without loss, a) Music b) Poker c) BMX

4.3.3.1 Impact of PLR on QoE According to Resolution and QP size

When Figure 4.7 is compared with Figure 4.6, it can be observed that the video resolution plays an important role in determining the impact of packet loss artifacts. The comparison of Figures 4.7 and 4.6 also shows that, on the whole, the distortion caused by a packet loss in SD is more than that in HD. We posit that the amount of information carried in an HD packet is lower than the amount carried in an SD packet, relative to the stream as a whole, and, thus, HD video is able to tolerate increased packet losses. As a result, a lost packet affects fewer macroblocks in HD, and, therefore, causes less spatio-temporal error propagation. In an SD video, an I packet, for example, may refer to the majority of, if not an entire image and so its loss will result in serious degradation and a poor MOS score. This observation has implications for related studies like [127] and [111] where the authors assumed that video on mobile/portable devices would be restricted to CIF or, even, QCIF resolution.

The corresponding results for 3D video, obtained by a comparison of Figures 4.8 and 4.7, present a similar pattern, although 3D assessments were consistently lower than those for the equivalent 2D video. Packet loss in 3D videos caused blockiness artifacts in different regions of each view. It was also observed that the impact of packet loss was to smooth out the effect of the QP settings in comparison to Figure 4.6. Further, as for 2D video, the assessed HD quality was somewhat higher than that for SD video (except for Poker at SD PLR=7.5).

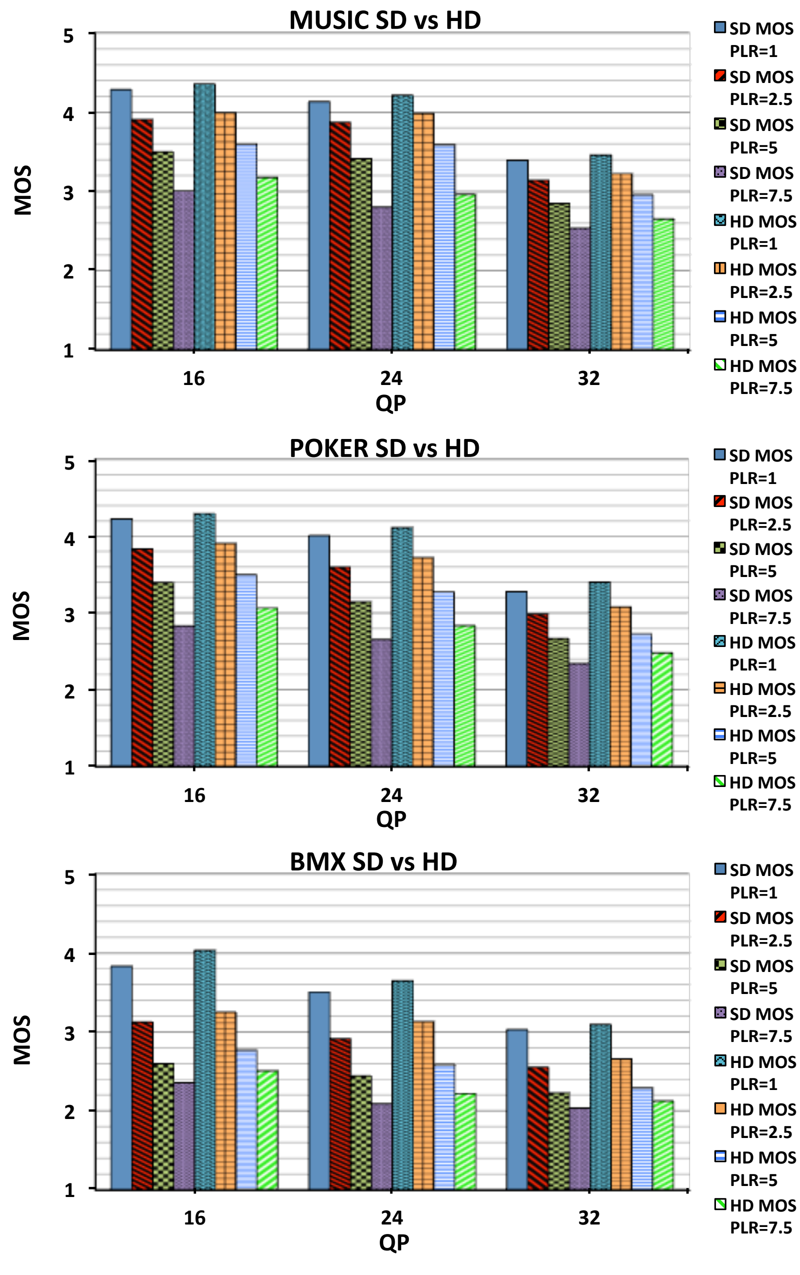


Figure 4.7: Impact of PLR on 2D SD and HD video QoE, a) Music b) Poker c) BMX.

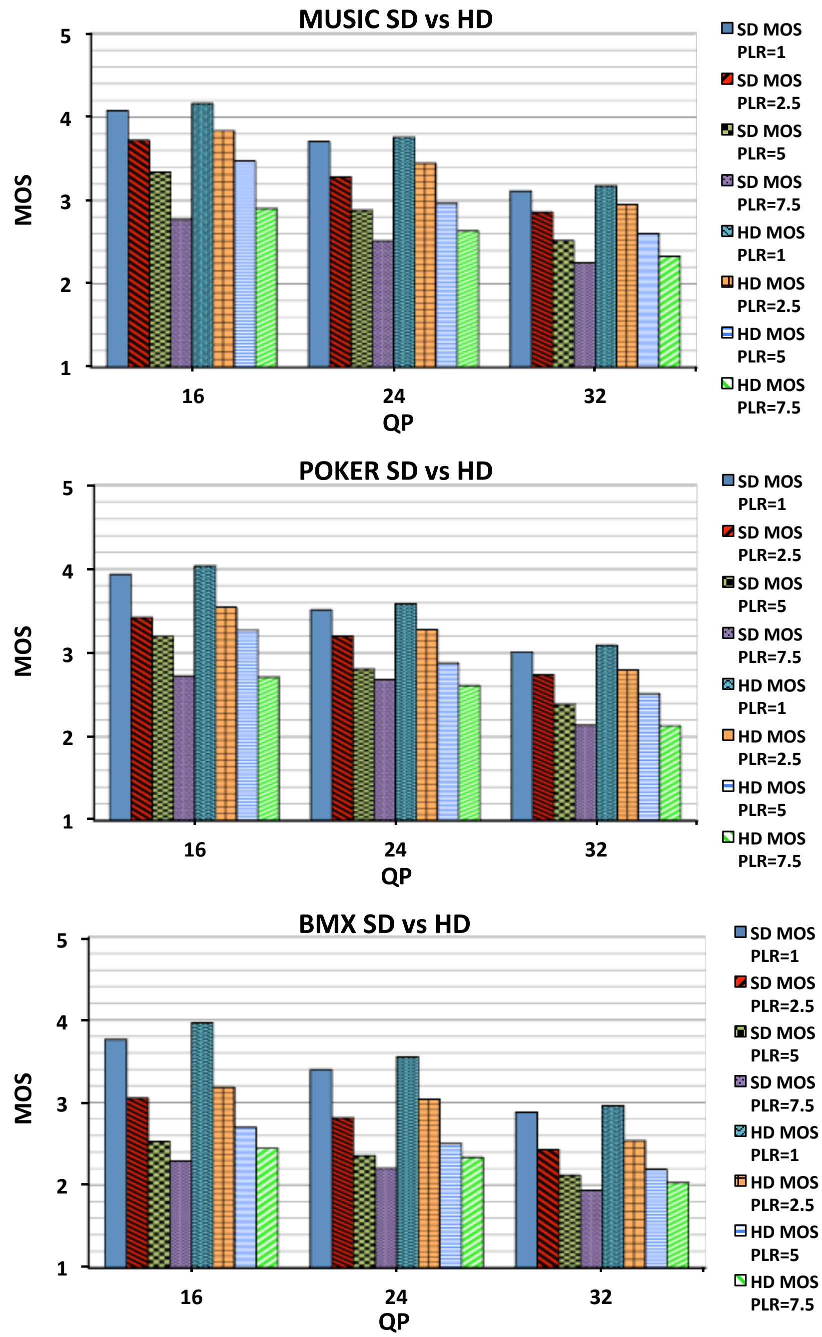


Figure 4.8: Impact of PLR on 3D SD and HD video QoE, a) Music b) Poker c) BMX.

In addition to the aforementioned general observations, it is apparent that packet loss has the effect of reducing the difference between the assessments for different QP sizes. This implies that there is less to be gained in terms of the video QoE from reducing the QP. Similarly, the bitrate rises with increasing QP, which, in turn, restricts the number of streams or simultaneous users that can share a wireless link.

To characterize the bitrate, the average P-frame size was found over the test sequences, as tabulated in Table 4.4 for 2D video. The bitrates given are, therefore, a lower bound on the required bandwidths, as I-frames result in bitrate spikes, owing to less efficient spatial coding. As mentioned before, at the lowest QP, it is apparent that the bitrate even of SD sequences would stress most wireless networks. For HD sequences, QP $=$ 16 introduces a steep change in the bitrate, implying that lower QPs should be avoided in many practical situations. These effects justify the need to carefully balance the QP/video quality against the effect of packet loss. The use of CBR video does not remove this requirement. For example, it is possible to set too high CBR and find a codec matching that rate even though the coding complexity does not merit it.

For the 3D video, the combined (C= color, D= depth) 3D bitrate tended to follow those of their 2D counterparts, but the requirement for separate coding of the two images for each frame resulted in an increase in the bitrate. The table of results for 3D video frame size is given in Appendix D. These observations are also made in [117, 128], which confirms that the requirements for 3D video are more stringent than those for 2D video, since the 3D video consumes a larger portion of the network's bandwidth. Generally, in order to maintain the right balance between good image quality and available bandwidth, it is necessary to select an appropriate QP for both 2D and 3D video encoding.

Table 4.4: 2D Input P-frame sizes by QP

2D	QP	Average size (bytes)	Average no. of packets	Approx. bitrate (Mbps)
BMX SD	16	297262	198	59.5
	24	82668	55	16.5
	32	26482	18	5.3
BMX HD	16	872130	581	174.4
	24	137963	92	27.6
	32	40923	27	8.2
POKER SD	16	277931	185	55.6
	24	43483	29	8.7
	32	13744	9	2.7
POKER HD	16	607188	504	121.4
	24	116278	78	23.3
	32	29952	20	6.0
MUSIC SD	16	166062	111	33.2
	24	20750	14	4.2
	32	7074	5	1.4
MUSIC HD	16	522905	349	104.6
	24	29866	20	6.0
	32	9295	6	1.9

4.3.3.2 Impact of PLR and MBL On QoE According to CT

Figures 4.9 and 4.10 show the correlations between the measured MOS and packet loss. It can be observed that the quality is highly dependent on the video content type. We found that, in the case of a lower motion video, the quality was acceptable up to a PLR of 4%. For high-motion videos, a PLR of greater than 2% reduced the quality significantly. It was also found that the overall impact of MBL is less obvious than that of PLR. This is because the total packet loss does, in fact, have a significant effect on the resulting distortion. In the case of a changing PLR with random loss, inter-frame dependencies play an influential role in propagating errors. However, for burst packet losses, the influence of inter-frame dependencies on temporal error propagation decreases with the growth in average burst length. Figure 4.10 shows no significant difference for 2D videos except that the PLR and MBL could cause more disparate distortions to 3D quality perception due to the complex nature of true 3D video perception. 3D video provides a sensation of depth by providing a disparate image of the same scene to each of the viewer's eyes.

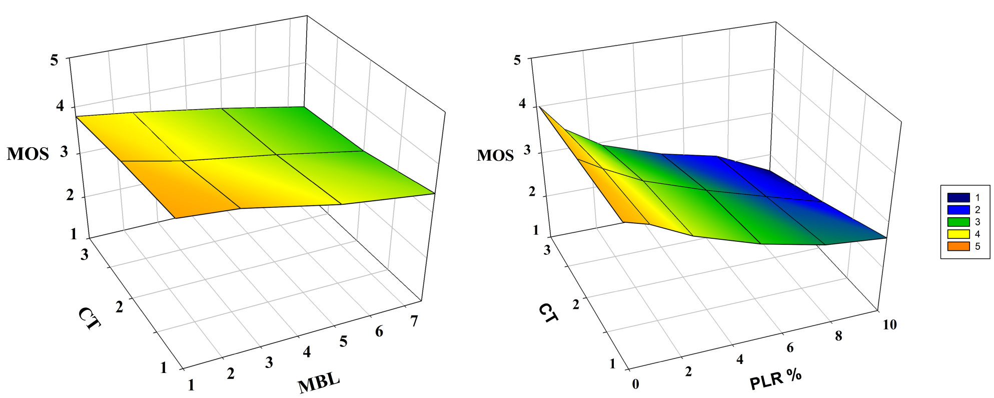


Figure 4.9: Impact of MBL and PLR QoE for different 2D CT
CT: 1=Low motion, 2=Moderate motion, 3=High motion
The key defines the colour mapping to MOS scores

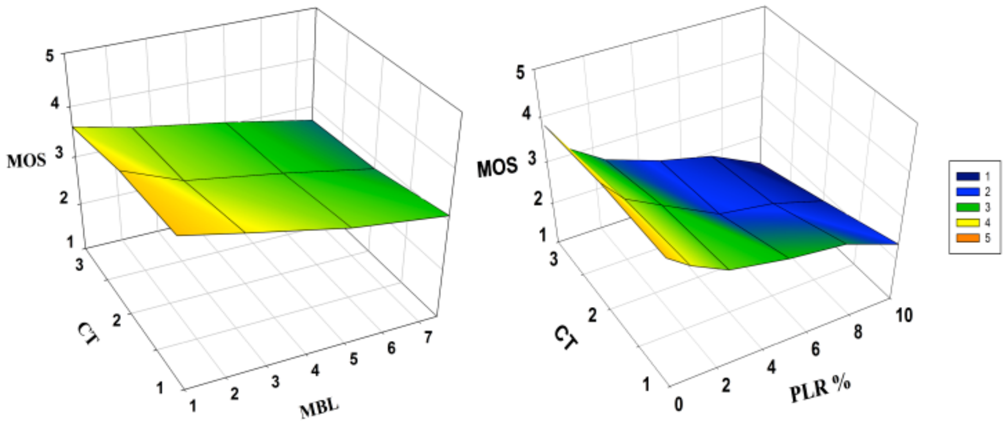


Figure 4.10: Impact of MBL and PLR on QoE for different 3D CT

CT: 1=Low motion, 2=Moderate motion, 3=High motion

The key defines the colour mapping to MOS scores

4.3.3.3 Analysis of Variance Test

The analysis of variance (ANOVA) test [129] is extremely useful for exploring the interactions between two or more independent variables. Harnessing ANOVA's analytical strength, a fundamental investigation was undertaken to quantify the impact of AQoS and NQoS parameters on the perceived video QoE. Tables 4.5 shows the results obtained from the ANOVA test of the 2D video datasets, where the degrees of freedom are shown in the first column, the second column is the F-statistic and the third column is the p-value. The p-value is determined from the cumulative distribution function (cdf) of F [130]. A small p-value (p 0.01) indicates that the video QoE is significantly affected by the corresponding parameter. Higher F-statistics correspond to higher proportions of the variance being caused by the independent variables [129].

The table of ANOVA's result for 3D video dataset is given in Appendix D.

It can be observed from the magnitudes of the p-values that all five parameters (p-value = 0) had significant effects on the video QoE. In particular, it can be seen that the PLR had the highest influence on the QoE (p-value=0) for both the 2D and 3D datasets, followed by QP and CT, while the MBL had the smallest influence on the QoE. Moreover, there were interactions between each pair of AQoS and NQoS parameters, each of which was significant. The two way interactions between PLR and CT, and PLR and QP had the highest influence on the QoE. In addition, the ANOVA results showed that the combined impact of MBL and CT was also significant.

Table 4.5: Five-way ANOVA on QoE of 2D Video

Source	Degree of freedom	F-statistics	p-value
CT	2	132.724	0.0
R	3	95.354	0.03132
QP	4	159.584	0.0
PLR	5	402.172	0.0
MBL	2	65.991	0.01068
PLR+CT	9	20.218	0.1180
PLR+R	9	5.182	0.3132
PLR+QP	15	26.955	0.11068
PLR+MBL	20	30.466	0.2301
MBL+CT	5	2.868	0.1541
MBL+R	5	7.940	0.48223
MBL+QP	11	13.533	0.28568

4.4 Study 2: Impact of QoS Parameters on H.264 and HEVC coded CBR Video Streaming

The majority of the related studies discussed earlier were conducted in an off-line video decoding scenario and, consequently, did not consider the effects of resource allocation while receiving and decoding a video stream in real-time. Moreover, exploring the effects of QoS parameters on video quality with respect to the packet structure and compression ratio have been rarely investigated in the literature.

To see that the packet structure is important in this regard, consider the following. Based on the available standards, the typical transmitted medium access (MAC) layer has a fixed payload size of 1500 bytes. If a video is encoded at a CBR of 20Mb/s, the number of packets generated when using the ETH/IP/UDP protocol stack will be approximately 1698, while a 10Mb/s video will generate 849 packets. This could mean that a 20Mb/s stream will be able to tolerate the same percentage of packet losses. The same scenario could also apply to video resolution, where different resolutions at the same bitrate would give rise to varied effects from the QoS parameters.

In addition, the bitrate reductions claimed for comparable video quality of H.264 and HEVC encoded streams invite scientific investigation: Are these claims valid for high resolutions video streams? In addition to examining the perceptual quality of coding-only impairments, there are other questions to consider: How does H.264 compare to HEVC when packet loss is present? Is there any potential disadvantage in the improved compression efficiency of HEVC in terms of robustness to dropped IP packets? To answer these questions, a practical wireless environment was implemented utilizing video encoders and a network impairment emulator. This study presents comparative results for the potential impact of packet loss and compression efficiency across SD, HD, and 4kUHD H.264/H.265 coded video streams, both transmitted and decoded simultaneously in real-time. In this study, the SSIM metric was used to objectively measure the video quality. SSIM is one of the best objective method of measuring video quality in real-time video streaming [131].

4.4.1 Experimental Set-up

4.4.1.1 Video Encoding

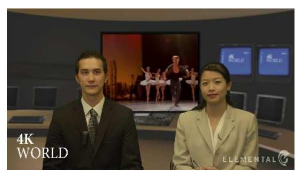
Four video sequences were chosen, one in each class, as listed in Table 4.6 and shown in Figure 4.11. These video sequences were classified, based on the SI and TI indexes for the luminance component of each piece of video content, as described in [52]. The video sequences were encoded using both of the H.264 and H.265/HEVC encoders. In the case of the H.264/AVC coded video, the implementation already available to FFMPEG was used, while the bespoke solution available in FFMPEG [132, 133] and LiBAV [134] was adopted for HEVC. The encoding parameters are given in Table 4.7. The choice of bitrate was based on the proposed average bitrate savings of 35.4% in comparison to H.264/AVC [46].

Table 4.6: Video sequences chosen and classes assigned

Video Sequence	**SI**	**TI**	**Class**
Coast	10.8370	16.9183	Low Motion
News	17.5219	21.2441	Low Motion
Foreman	16.3897	38.2870	Moderate Motion
Sintel	19.7101	72.2639	High Motion

Sintel

Coast



News

Foreman

Figure 4.11: One representative frame from each of the 4 source video sequences

Table 4.7: Encoding Parameters

Codec	**HEVC**	**H.264/AVC**
Profile	Main	High
Processing Unit	Coding Tree Block	Macroblocks
Processing Unit Size	64 x 64	16 x 16
Group of Picture (GOP) size	25	25
GOP structure	IPPP	IPPP
Bitrate	13.5Mb/s	20Mb/s

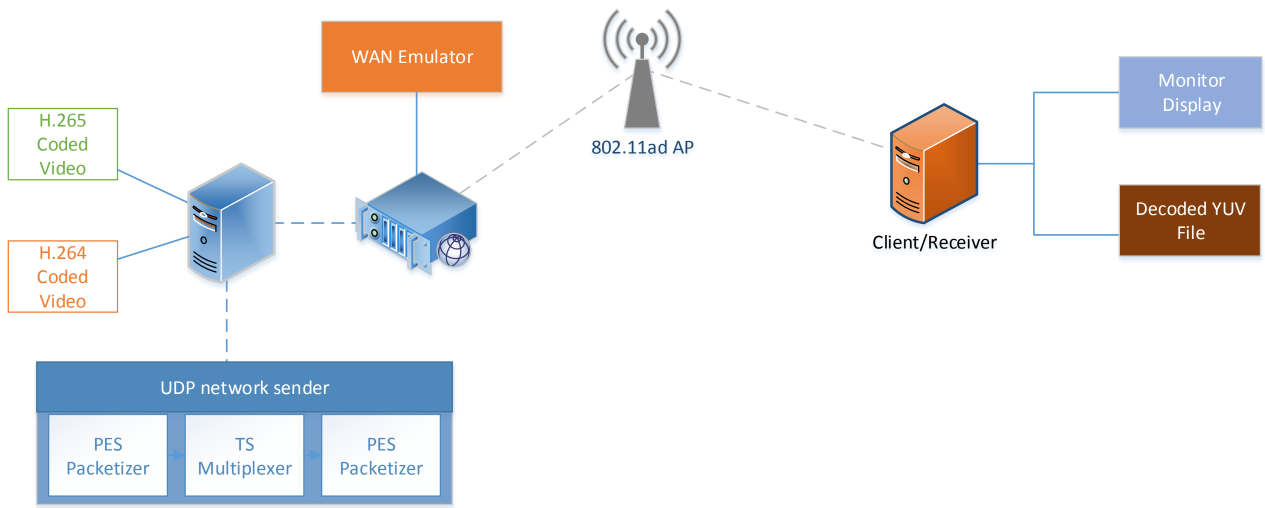


Figure 4.12: Topology of the physical tes-tbed environment

4.4.1.2 Test-bed Scene and QoS Parameters

To investigate the effect of QoS Parameters on the video QoE, the SCE [135] WAN emulator was adopted. The emulator was attached to the outbound connection of the sender using the bridging mode to enable packet transmission under controlled conditions. A schematic representation of the implementation can be seen in Figure 4.12. The NQoS parameters evaluated were the packet loss rate (PLR), duplication rate (PDR), and re-ordering rate (PROR), all of which were expressed as percentage values. Different values of these parameters were introduced to affect packets randomly, see Table 4.8.

In order to perform a live transmission of the coded video, the use of MPEG2-TS was adopted. The NAL units were fragmented into MPEG2-TS and were then re-

Table 4.8: QoS Measurement Parameters

QoS Parameters	Values
Packet Loss (PLR) (%)	1, 2, 5,.....,9
Packet Duplication (PDR) (%)	1, 2, 5,.....,9
Packet Reordering (PROR) (%)	1, 2, 5,.....,9
Combination PLR+PROR+PDR	2%
PLR (2%) + Bitrate	13.5Mb/s, 18Mb/s, 23Mb/s, 25Mb/s

encapsulated into user datagram packets (UDP). Since each MPEG2-TS packet has fixed size of 188 bytes, each UDP packet is only allowed to carry seven MPEG2-TS packets (1316 bytes), due to the MTU payload of 1472 bytes imposed by the operating system. Encoding, de-multiplexing and decoding were performed using open source applications [136]. With the exception of encoding, which was done offline, all other processes such as the modified solution for MPEG2-TS transmission (server side), the demultiplexing of received MPEG2-TS and decoding of the H.265 elementary stream (ES) (client side) operated in real-time. The decoded (distorted) YUV was then compared against the original (YUV) to obtain video quality results with respect to the SSIM quality metric.

4.4.2 QoE Measurement

The QoE of the received video sequences was measured using the structural similarity index metric (SSIM) [56]. SSIM provides a quality index measure of the similarity between two images. The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case of two identical sets of data. Owing to its excellent performance and extremely low compute cost, SSIM has become a dominant method of measuring video quality in real-time video streaming and broadcasting [131]. Zinner et al [7] proposed a mapping between SSIM and MOS in order to express the measured quality in corresponding subjective terms, as shown in Table 4.9.

Table 4.9: Mapping of SSIM to MOS [7]

Video sequence	SSIM
> 0.99	5 (excellent)
≤ 0.95 & < 0.99	4 (good)
≤ 0.88 & < 0.95	3 (fair)
≤ 0.5 & < 0.88	2 (poor)
< 0.5	1 (bad)

4.4.3 Results and Discussion

All results shown in this section are based on average video quality values with respect to the SSIM, obtained during the experiments. Each experiment was conducted ten times. The initial encoded video quality results are displayed in Figure 4.13. The results suggest a drop in video quality based on the motion complexity of each of the sequences, while an increase in video quality was observed as the resolution was decreased. This is a result of the reduced compression ratio provided by the encoder which made more pixel information available to the compressed image so that less coded video information was carried in each packet. In addition, the results also suggest increased performance with respect to video quality for HEVC over H.264.

This is negligible since the lowest mean SSIM value is 0.968 for Sintel (H.264/AVC).

The presented results in Figure 4.13 appear to differ from experiment ones results in two ways: 1) higher resolution video results in lower objective MOS ratings; and 2) lower motion videos have higher qualities. However, the results are not contradictory, because, in this set of tests, the CBR was fixed whatever the resolution. This naturally results in less compression for lower resolutions, as the QP varies to match the available bitrate. Moreover, the encoders have been able to take advantage of the additional bitrate to improve the quality of low-motion videos. In other words, both codecs avoid simply increasing the bitrate artificially, by, for example, including more intra-coded CTBs or MBs.

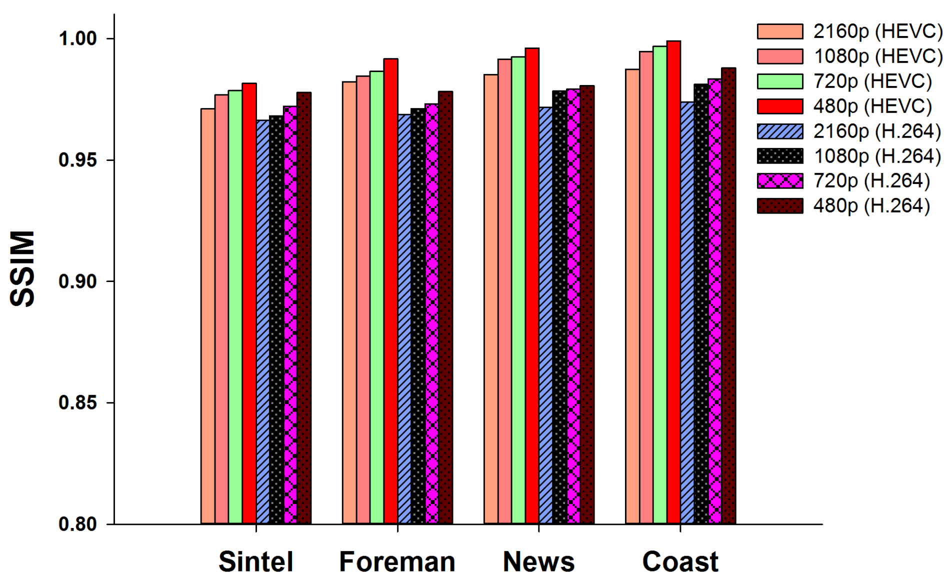


Figure 4.13: Quality of HEVC and H.264 coded video without network impairment

4.4.3.1 Impact of QoS Parameters According to Compression Ratio and Resolution

Figures 4.14-4.16 show the impact of the QoS parameters on both the HEVC and H.264/AVC coded content. The H.264/AVC codec now (compared to Figure 4.13) appears more resilient to packet loss than HEVC resulting in higher quality ratings for H.264/AVC encoding. This finding strongly suggests that HEVCs more efficient encoding makes it more sensitive to packet loss, once the relative coding gains have been allowed for by scaling the CBRs (with HEVC at 13.5 Mbps and H.264/AVC at 20 Mbps). For example, it can been seen that the most sensitive spatial resolution was 4kUHD (2160p), while the least sensitive was the standard definition (480p) resolution.

Packet loss now serves to exaggerate the difference in quality already starting to show in Figure 4.13 between lower (higher quality) and higher resolution (lower quality). Similarly in Figure 4.14, when packet losses occur higher motion sequences such as Sintel suffer in quality much more than lower motion sequences such as Coast. Again a possible explanation, when a fixed CBR is involved, is that lower spatial resolution and lower motion video sequences will tend to be less compressed. Consequently, with more coded content per packet for lower resolution and lower motion videos, error concealment is better able to reconstruct missing packets.

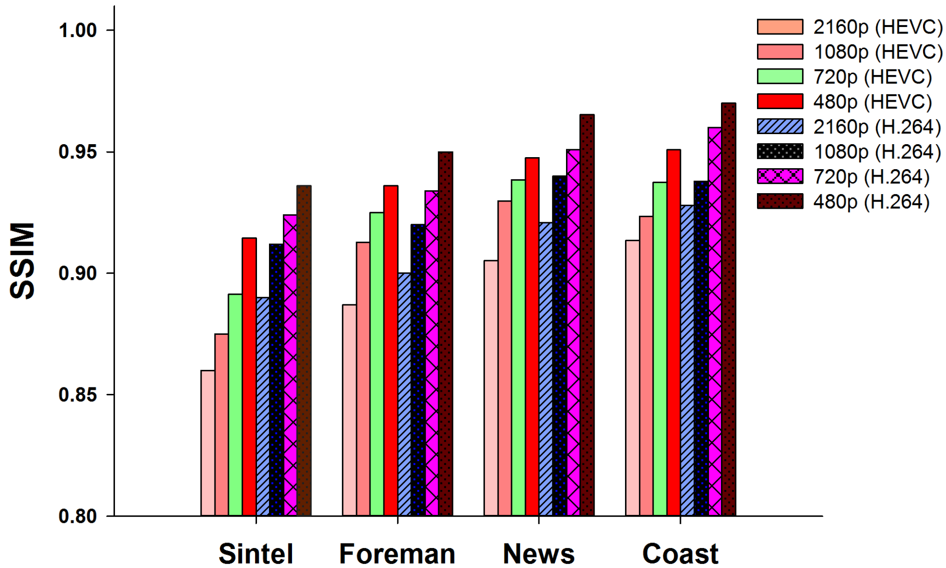


Figure 4.14: Impact of PLR (2%) on H.264/AVC vs HEVC transmission

Figures 4.15 and 4.16 show the correlation between SSIM and the combined effects of PROR and PDR. We observe that the impact of packet reordering and duplication was less than that of packet loss (the lowest values are 0.8862 and 0.8652 (Hevc-coded sintel) for reordering and duplication, respectively). We also noticed that the quality was dependent on the video content type, target bitrate and resolution, just as for packet loss. We further noticed that, due to the duplication of time stamps for packets, duplicate packets were dropped, thereby requiring additional processing overhead and, consequently, reducing the hardware resource allocation available to the decoder.

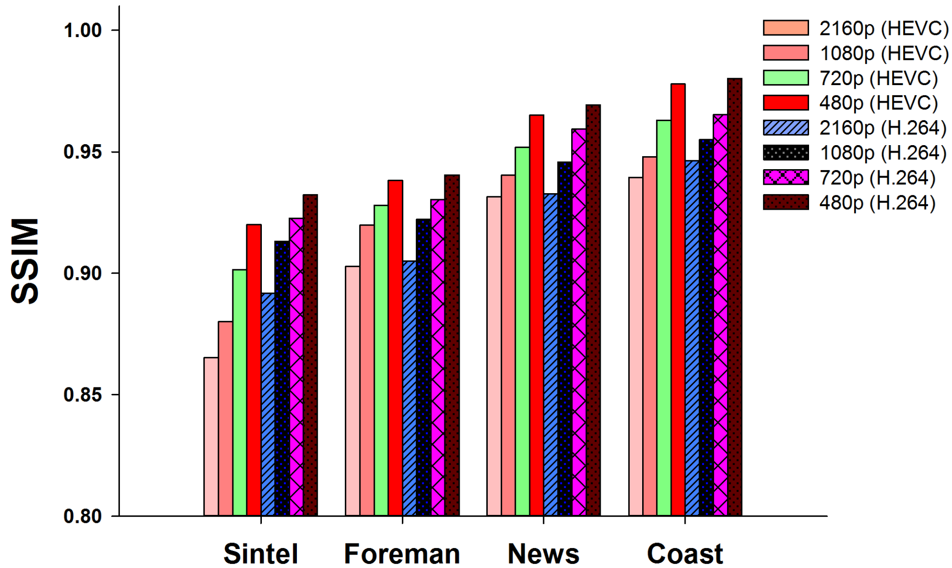


Figure 4.15: Impact of PDR (2%) on H.264/AVC vs HEVC transmission

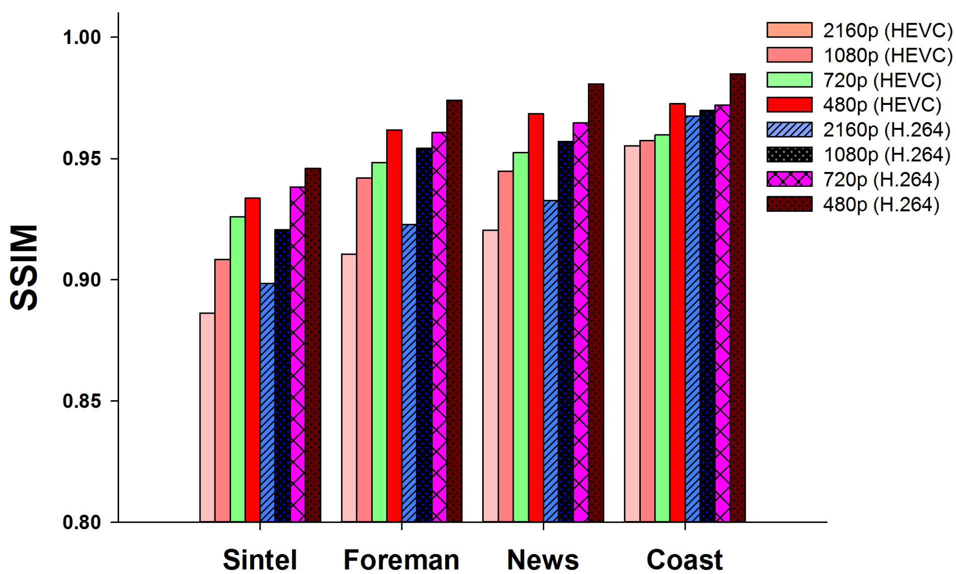


Figure 4.16: Impact of PROR (2%) on H.264/AVC vs HEVC transmission

4.4.3.2 Impact of QoS Parameters on 4k HEVC Video According to CT

The impact of all three QoS parameters, set at 2% each, on 4kUHD HEVC coded video can be seen in Figure 4.17. The results suggest that the combination of all three parameters had a more devastating impact on the test sequences than the results obtained for the scenarios presented in Figures 4.14-4.16.

In addition, Figure 4.18 illustrates the effect of 2% PLR on different choices of the target bitrate for 4kUHD HEVC coded video. The results suggest that increasing the target bitrate reduced the effect of packet loss, and thus increased the video quality. This can be attributed to the amount of coded information which was distributed amongst the packets. For example, a 13.5Mb/s video will have fewer packets, but this means that the amount of coded information in each packet is higher than for a 25Mb/s video, where this information could be split into two or more packets, resulting in a reduced sensitivity to packet loss. Moreover, as fast moving content is more sensitive to packet loss, a higher bitrate is required to achieve a comparable perceived quality for slow content.

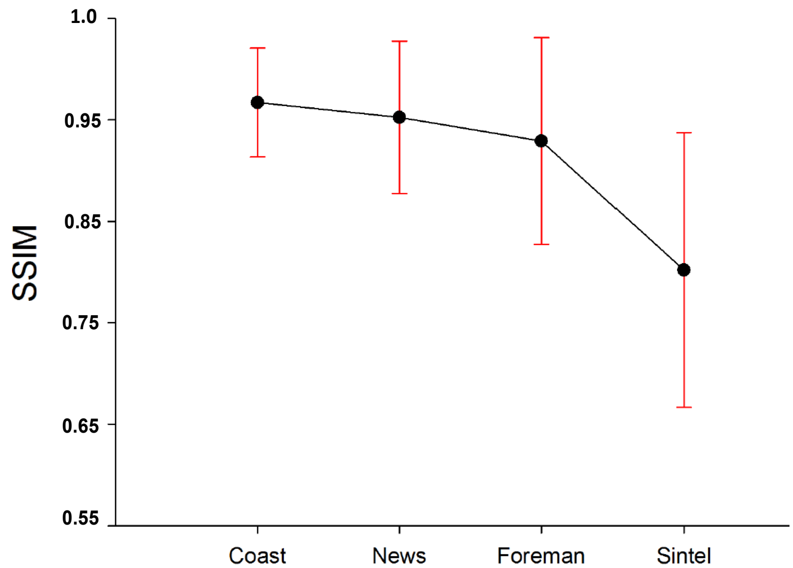


Figure 4.17: Impact of PLR+PDR+PROR at 2% each on 4kUHD HEVC video

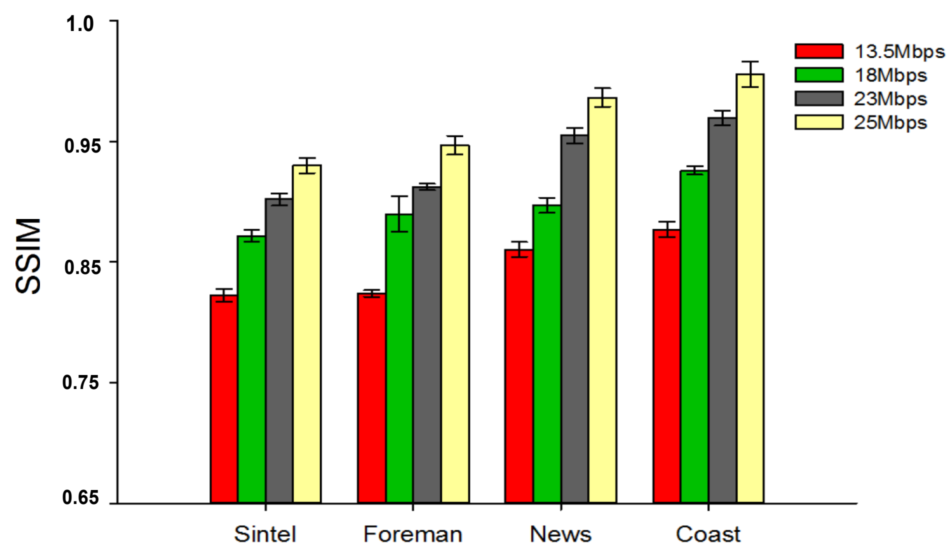


Figure 4.18: Impact of 2% PLR on target bitrates of 4kUHD HEVC transmission

Figures 4.19, 4.20 and 4.21, presents results on the impact of different level of PLR, PDR and PROR on 4kUHD HEVC coded video. In Figure. 4.19, the results presented show the impact of random packet loss on 4kUHD coded video. These results suggest that 4kUHD HEVC coded video is very sensitive to packet loss, especially for video content with higher TI values. It can be seen that from a PLR of 2%, the mean SSIM of such video contents (Sintel, Foreman, News) is impacted more than lowest TI value (Coast).

In Figure 4.20, the impact of random packet duplication with packet duplication ratios (PDR) is shown. Unlike the case of packet loss, the sequences seem tolerate up to 2% PDR, except in the case of Sintel, where the video quality falls in bad zone. All sequences gradually drop in quality after 2% PDR. While in Figure 4.21, the effects of random packet reordering with respect to PROR can be seen. These results suggest that video sequences with lower motion complexity (TI values), such as Coast and News can accommodate a maximum of 4% PROR. While those with higher motion are, the effects of PROR appears to be more obvious after 2%.

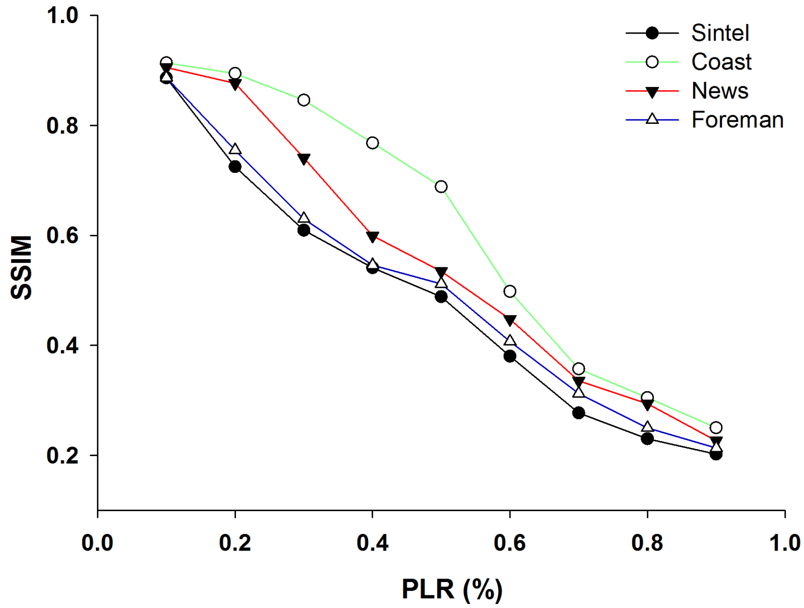


Figure 4.19: The impact of PLR on 4kUHD HEVC transmission

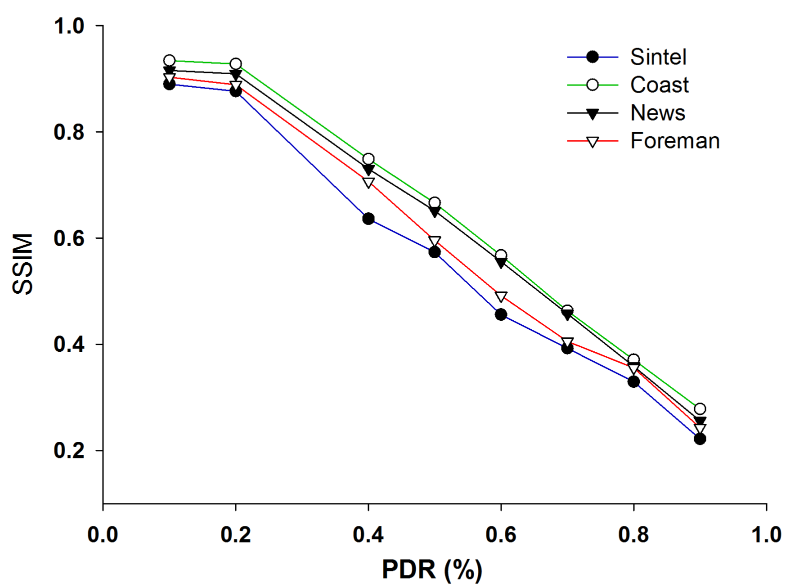


Figure 4.20: The impact of PDR on 4kUHD HEVC transmission

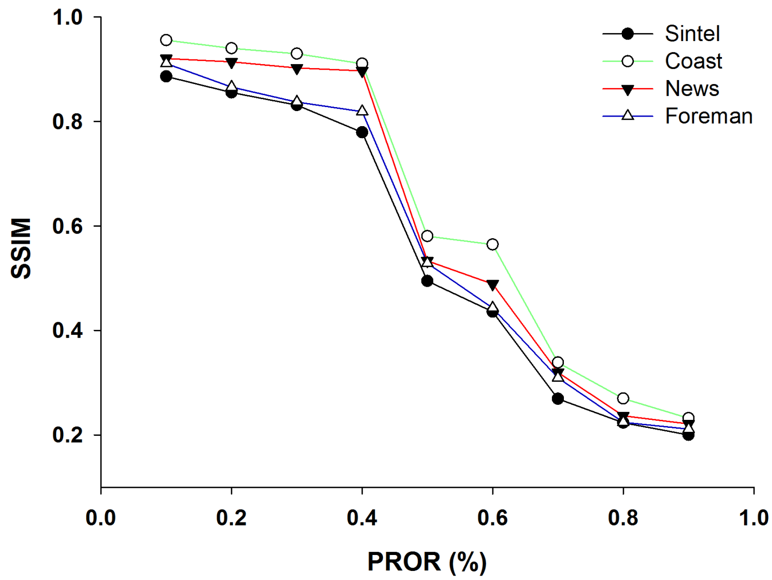


Figure 4.21: The impact of PROR on 4kUHD HEVC transmission

4.5 Chapter Summary

This chapter presents two sets of experiments, the one at a VBR and the other at a CBR, which highlight different aspects of the interpretation. The first study performed an investigation of the impacts of QoS parameters on the video QoE by cross-layer simulation of the transmitted SD, HD and 3D videos for the VBR scenario. The video quality was measured using objective and subjective metrics. The results showed that for a given packet loss rate is likely to have less impact on the QoE of a HD video than it has upon lower resolution videos when the bitrate is variable for both resolutions. That observation has implications for earlier related studies which assumed mobile/portable devices would be confined to CIF or even QCIF resolution.

The second study investigated the effects of QoS parameters on up-to 4kUHD resolution of H.264 and HEVC coded video streams which was both transmitted and decoded simultaneously in real-time. It was found that, at degradations of 2% in each of the QoS parameters, the video quality diminished, especially for the sequences with higher spatial resolution. However, this was because in the experiments the same CBR was used whether transmitting lower or higher resolution video, including 4kUHD video. Thus, the QP can be lower for lower resolution video than higher resolution video at the same CBR. Losing a packet from a lower resolution stream in those circumstances, has less of an impact, than losing a packet from a higher resolution stream. This highlights the need for careful interpretation of video quality results depending on the streaming modalities. The results presented also show that increasing the bitrate reduced the impact of packet loss. This artefact stemmed from the fragmentation of video data into more packets, and thus provided a trade-off between video bandwidth allocating and the impact of packet loss.

Overall, we found that video QoE is very sensitive to different values of the AQoS and NQoS parameters. Therefore, it is important to model the correlation between QoS and QoE so that the QoE can be predicted without the reference video. The results from this chapter enable the choice of the parameters required for developing the non-reference video quality prediction models presented in Chapters 5 and 6. We use the measured objective and subjective datasets from this chapter as a learning set to develop the proposed prediction models in Chapters 5 and 6.

Chapter 5

Hybrid Non-reference Video Quality Prediction Model Using Fuzzy Logic System

5.1 Introduction

In Chapter 4, it is observed that video QoE is very sensitive to different values of AQoS and NQoS parameters, with changes in these values corresponding to different levels of video impairments and differing QoE values (MOS scores). It is very difficult to accurately estimate the video QoE on the basis of observations or simple mathematical formulas as the relationship between QoS and QoE is fuzzy.

The video QoE was evaluated in the previous chapter using full reference (FR) quality metrics in both of subjective and objective scenarios. However, in a real-time scenario, it is impractical to measure the QoE using FR methods since the reference video is absent. Consequently, it is important to model the correlation between QoS and QoE so that No-reference (NR) methods can be used to predict the QoE in the absence of the reference video, and, even, before the transmitted video is received by the end user. In order to provide a broader prediction of video quality, hybrid models that consolidate both NQoS and AQoS parameters have been investigated in the literature. In addition, many researchers have turned their attention to the use of machine learning algorithms for predicting video quality due to their ability to automatically learn from past observations. As discussed in Section 2.7, a number of video QoE prediction models have been proposed in the literature, but these have offered only partial solutions which overlook some key QoS parameters along the video delivery chain. Thus, there is still room for the development of innovative mechanisms to efficiently correlate the QoS with the QoE.

This chapter presents a hybrid non-reference model for predicting video quality which uses fuzzy logic inference systems (FIS) as its learning-based technique. The proposed model is based on a predetermined rule-based method (FIS-PRB). The learning dataset developed in Chapter 4 for the correlation of QoS parameters with the QoE measured using objective and subjective tests was employed to build the FIS-PRB model. For end-to-end quality estimation, QoS parameters from both the application and network layers were identified. The proposed model was validated through the correlation of the predicted and measured QoE found using the testing and external unseen datasets. Furthermore, the proposed model also evaluated by comparing its performance with another learning technique, which is the RNN.

The remainder of this Chapter is organized as follows. Section 2 presents the experimental set-up and dataset generation. In Section 3, the proposed methodology is discussed. Section 4 is devoted to the validation of the model. In Section 5, a performance comparison with the RNN method is discussed. The 3D video QoE prediction is outlined in section 6. Finally, Section 7 summarizes the Chapter.

5.2 Experimental set-up and Dataset Generation

The simulation set-up and the generated datasets were as described in Chapter 4 (Section 4.3.1). The objective dataset measured in Chapter 4 for the correlation of QoS parameters with the QoE were used as a learning set to build the proposed FIS-PRB model.

We extended the generated objective datasets to include three more video sequences in order to used them for testing the model. The new three video sequences were streamed via the same simulation and using the same QoS parameters. Consequently, the objective dataset includes 6 video sequences, two in each content class, as listed in Table 5.1. The Music, Poker and BMX video sequences were used for model learning, while the Fencing, Poznan and Pantomime sequences were chosen for validation.

Moreover, the subjective dataset was also used for model validation. A systematic approach was used to select a subset (from the generated objective dataset) designated for the subjective testing. The selection approach was based on the Kennard and Stone algorithm [126], which selects the sample that is most distant from those already selected as the next sample.

Table 5.1: Video sequences chosen and classes assigned

Video Sequence	**TI**	**SI**	**Class**
Music	4.90	74.41	Low Motion
Fencing	7.78	77.20	
Poker	12.20	85.69	Moderate Motion
Poznan	11.53	87.78	
BMX	22.35	99.42	High Motion
Pantomime	37.17	104.43	

5.3 Methodology

According to [6], the FIS outperforms other estimation techniques in terms of modelling capabilities and making decisions with imprecise information. In addition, FIS provides a way of constructing controller algorithms by means of linguistic labels and linguistically interpretable rules in a user-friendly way closer to human thinking and perception. Unlike neural networks or genetic algorithms, FIS does not need a period of online training or convergence, making it a proper tool for real-time control. Additionally, the calculations can be very simple, especially when triangular or trapezoidal membership functions are adopted [137].

The proposed prediction model is based on a predetermined rule-based method (FIS-PRB). A semi-manual method was applied to develop the FIS-PRB model, combining human knowledge and the behaviour of QoS parameters with testing on a simulator. Figure 5.1 illustrates the functional block diagram for the proposed FIS-PRB prediction model.

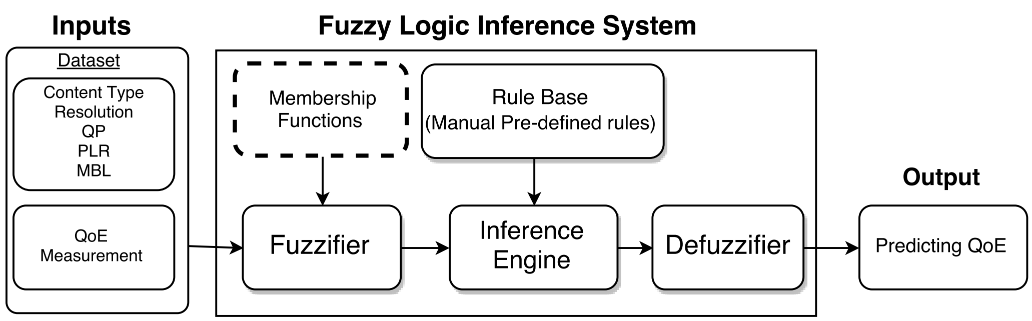


Figure 5.1: Functional block diagram of the proposed FIS-based prediction model

The membership functions and fuzzy rule extraction were conducted manually, and the fuzzy controller was implemented using the Matlab fuzzy toolbox. Once the MFs and a set of fuzzy rules were extracted from the dataset, the FIS could begin to perform the mapping between the QoS parameters and the QoE. Algorithm 1 gives an overview of the process of designing a fuzzy logic system. Background information about FIS was given in Chapter 2 (Section 2.8). The following sections describe the steps for the development of the proposed model, its membership functions and fuzzy rules extraction.

Algorithm 1: Fuzzy logic system

1. Define the linguistic expressions (Initialisation)
2. Design the membership function using triangle shape (Initialisation)
3. Convert crisp input value to fuzzy value using the MFs (Fuzzification)
4. Manually extract the fuzzy rule base (Fuzzy inference engine)
5. Evaluate the fuzzy rules in the rule base (Fuzzy inference engine)
6. Aggregate the results of each rule (Fuzzy inference engine)
7. Convert the fuzzy value to crisp output value (Defuzzification)

5.3.1 Identifying the Inputs and Output

The chosen input QoS parameters were content type (CT), resolution (Re), quantisation parameters (QP), packet loss ratio (PLR), and mean burst loss (MBL), while the output was the MOS scores (QoE). Table 5.2 outlines the chosen QoS parameters. Variations in these QoS parameters affect the quality of the delivered video and, consequently, the user satisfaction level. Once the inputs and the output were identified, both were categorized into linguistic expressions that represent the quantification of the QoE values (scores). This was achieved by the design of membership functions for the input and output variables.

Table 5.2: The chosen QoS Input Parameters

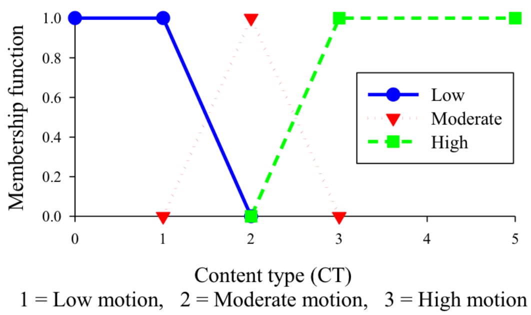
Parameters	Values
Content Type (CT)	Low, Moderate, High motion
Spatial resolution (R)	SD (720 $\times$ 540), qHD (960 $\times$ 576), HD (1280 $\times$ 720)
Quantization Parameter (QP)	16, 24, 32, 40, 48
Packet loss ratio (PLR)	1%, 2.5%, 5%, 7.5%, 10%
Mean burst length (MBL)	1, 2.5, 5, 7.5

5.3.2 Design of Membership Functions

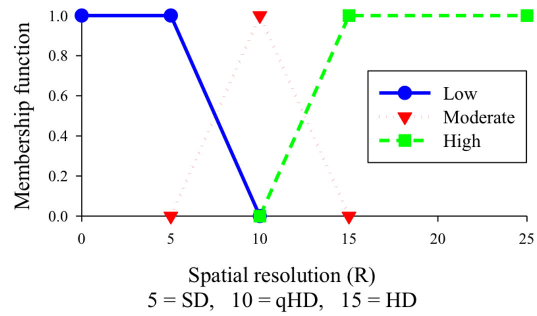
The data representing the correlation between QoS parameters and the measured QoE was translated into fuzzy membership functions. In this study, the membership functions were derived using probability distribution functions (PDF) [138]. Different PDFs were built for every QoS parameter. The probabilistic information was converted into a fuzzy set by dividing the PDF by its peak. Three fuzzy sets (low, moderate, high) were assigned to each of the chosen QoS input variables. These fuzzy sets were converted into equivalent forms (shapes) of the membership function using a curve fitting method [139]. The curve values of the membership functions represented the degree to which a particular QoS parameter value resulted in different MOS scores. In the proposed system, the fuzzy set was converted into an equivalent triangular shape, which was simple to implement and increased the speed of computation [137].

Figure 5.2 shows the membership functions for the QoS input parameters. We note that a membership value of 1 represents a high degree of membership in the corresponding class and a decreasing value represents deviation from the class. The relationships represented by the membership functions reflect the "effect" of each QoS parameter's value on the QoE. This "effect" is categorised into three fuzzy sets (low, moderate, high), which are interpreted as low effect, moderate effect, and high effect. For example, a PLR of 2% has a membership degree of 0.6 in the low effect on QoE set. In other words, a 2% PLR has a low effect on the QoE to a degree of 60% (or 0.6). Likewise, a 2% PLR has a moderate effect on the QoE to a degree of 30% (or 0.3), which is a lower degree. Finally, a 2% PLR has a high effect on the QoE to a degree of 0%, which means that a 2% PLR has no high effect on the QoE at all. The membership functions for the QoS input parameters were designed in direct relation to the values of these QoS parameters. That is, for these QoS parameters the fuzzy sets (low, moderate, and high) corresponded to numerical values of these parameters that are low, moderate, or high.

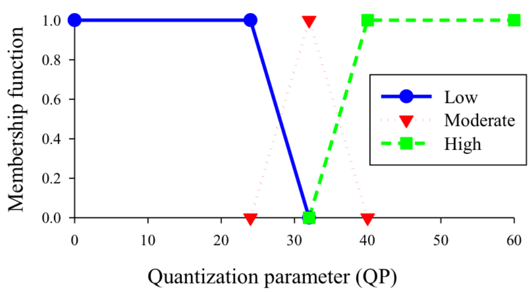
The membership function of the output (QoE) is presented in Figure 5.3. In this work, the measurement of the QoE was based on the MOS scores (1 to 5). Consequently, the QoE membership function was assigned five fuzzy sets, according to the standard MOS definition [52]. In Figure 5.3, each value of the MOS scale is related to the fuzzy sets (bad, poor, fair, good, excellent) to a different degree. For example, an MOS value of 3.2 is related to the fair set to a degree of 80% (0.8), and the good set to a degree of 20% (0.2). However, MOS 3.2 is not a bad, poor, or excellent value, since it is bad, poor, and excellent to the degree of 0%.



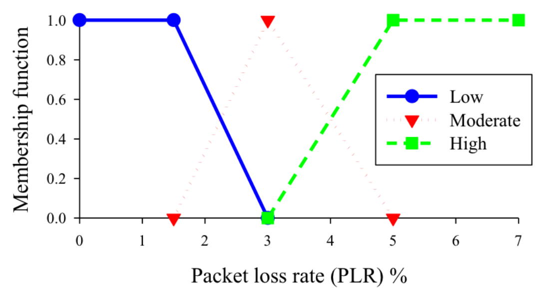
(a) Content type



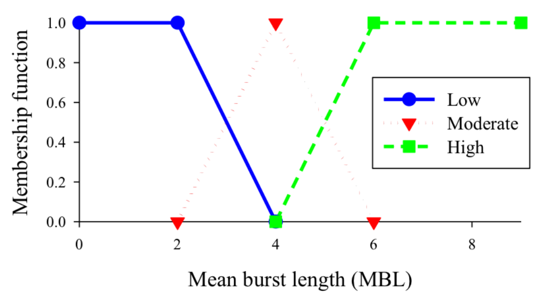
(b) Spatial resolution



(c) Quantisation parameter



(d) Packet loss rate



(e) Mean burst length

Figure 5.2: Membership functions of the inputs

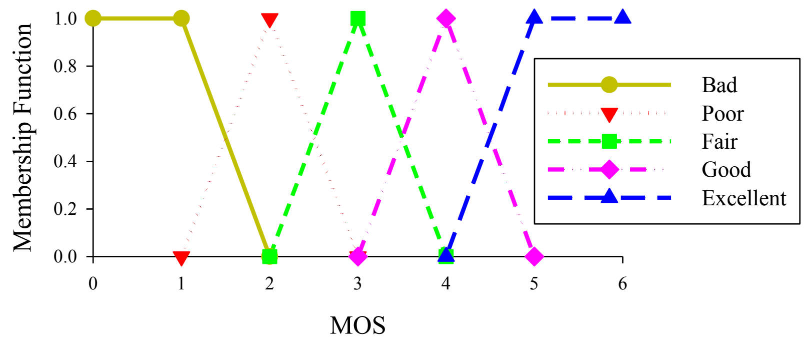


Figure 5.3: Membership functions of the output (QoE)

5.3.3 Fuzzy Rules Extraction

The proposed FIS-PRB employs the Mamdani inference type [140]. Based on the combinations of QoS parameters and their ratings, the impact of QoS variables on video quality (QoE) was estimated to give one of the MOS scores (QoE). That is, an estimated QoE score was associated with each combination of QoS parameter values. The fuzzy rules were generated by assigning weights to the video impairment scores. For each combination, the rule weight was calculated as the sum of the weights of the QoS parameter scores. In this work, the AND operator, which selects the minimum value of the fuzzy sets was used. After the results for each rule were evaluated, they were combined to produce a final result using the maximum algorithm [101]. This algorithm is the most commonly used accumulation method. It combines the results of the individual rules by selecting the fuzzy set that achieves the greater membership value in the IF part of the rule. An example of the processing of fuzzy rules extraction is given in Section 2.8.2.

The greatest number of rules that can be generated is X^n, where X is the number of fuzzy sets and n is the number of input variables. So, the maximum number of rules that can be extracted is 3^5. However, the actual extracted rules set usually contains fewer than X^n rules because, for many parameter values, the membership functions are zero for all fuzzy sets except one or two. A zero-valued membership function implies that the antecedent is not used in the rule [140]. The number of extracted fuzzy rules for the proposed model was 243. Table 5.3 shows a sample of the index fuzzy rules. The final fuzzy rule that was used in the FIS controller had the form:

IF (CT is Low motion) AND (MBL is Low) AND (PLR is Low) AND (QP is Low) AND (Resolution is Moderate) THEN (QoE is Good)

IF (CT is High motion) AND (MBL is High) AND (PLR is High) AND (QP is Moderate) AND (Resolution is Moderate) THEN (QoE is bad).

Table 5.3: Examples of the QoE index decision making rules

CT	**R**	**QP**	**PLR**	**MBL**	**QoE**
Moderate	Low	Low	Low	Low	Excellent
Moderate	High	Moderate	Low	Low	Good
Low	Moderate	High	Moderate	Moderate	Poor
High	Low	Moderate	Low	Moderate	Fair
High	Moderate	High	High	High	Bad
Low	Low	Moderate	Moderate	High	Poor
Low	Moderate	Low	Low	Moderate	Good

5.3.4 Defuzzification: Predicting the Output

The overall result after the inference step was a set of fuzzy values. This result was defuzzified to give a crisp output value (QoE), based on the MF of the output variable. The defuzzification process examined all of the rule outcomes after they were logically added and then computed a value that was the final output of the fuzzy controller. In this work, the defuzzification was conducted using the centroid method. The mathematical basis of this method relies on the centre of gravity (COG) defuzzification method that is expressed using the following formula [141]:

$$y(x) = f_s(x) = \frac{\sum_{i=1}^{M} y^1 \prod_{l=1}^{n} \mu F_l^i(x_l)}{\sum_{i=1}^{M} \prod_{l=1}^{n} \mu F_l^i(x_l)}. \tag{5.1}$$

Here, M is the rule number in the rule base, y^{-i} is the centroid of the ith output fuzzy set B^i, and $\Pi_{l=1}^{n} = \mu F_l^i(x_l)$ is the product of the membership values of each rule's inputs.

5.3.5 Implementation of FIS-PRB model

As discussed in Section 2.8, the FIS controller consists of three main components, or actions: fuzzification, fuzzy inference processing and defuzzification. These actions must be completed sequentially in order to determine the appropriate output value. There are different methods for implementing the FIS controller, one of which is the MATLAB FIS Toolbox [142] that provides computational and visual aids. In this study, the MATLAB FIS toolbox was used to develop a simulation scenario using the designed membership functions and the extracted fuzzy rules. Figure 5.4 shows the FIS-PRB model designed using the MATLAB FIS toolbox.

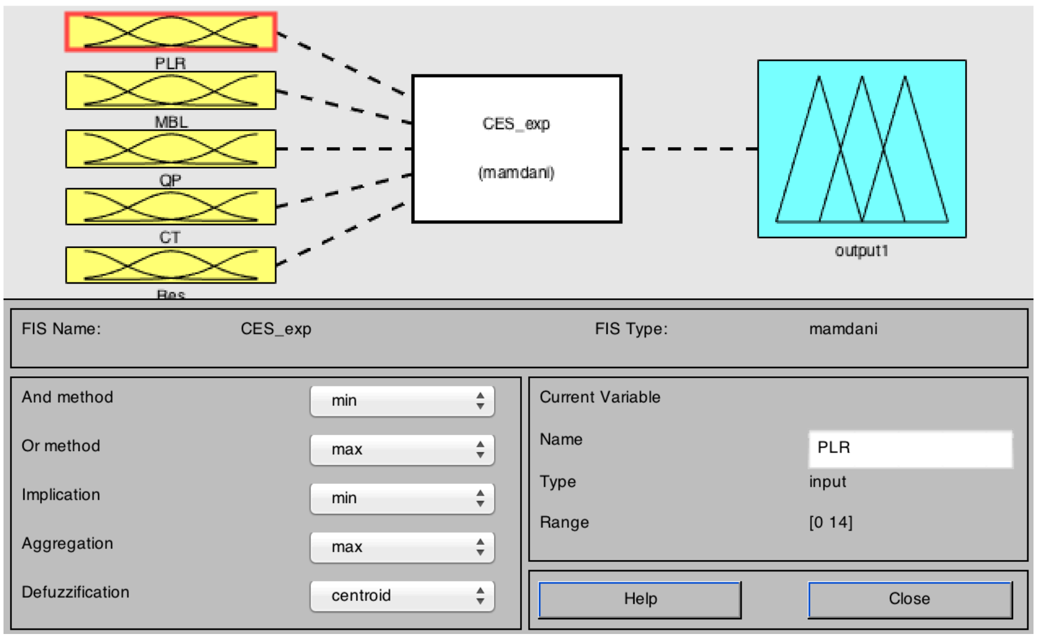


Figure 5.4: The FIS-PRB model in MATLAB toolbox

After this step, the QoE was predicted and compared with the measured QoE. The validation of the model is presented in the following section.

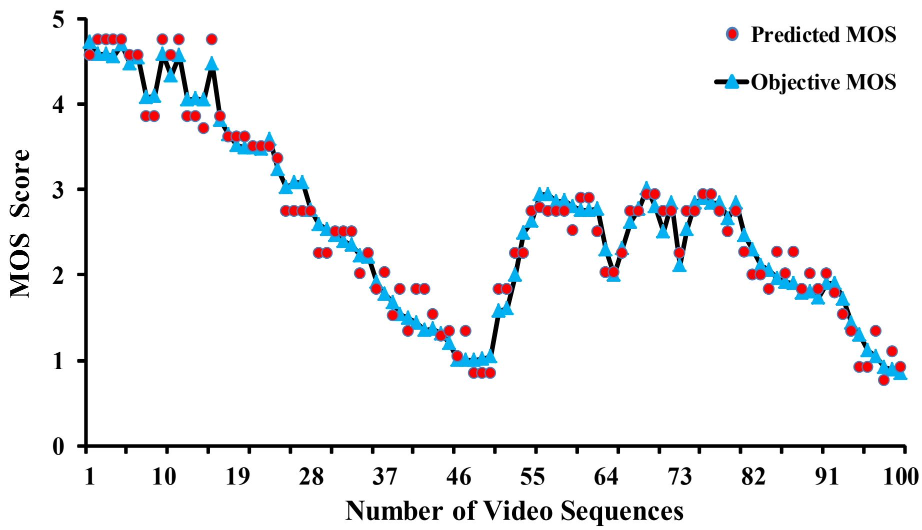
5.4 Validation of FIS-PRB Model

5.4.1 Validation by Testing Dataset

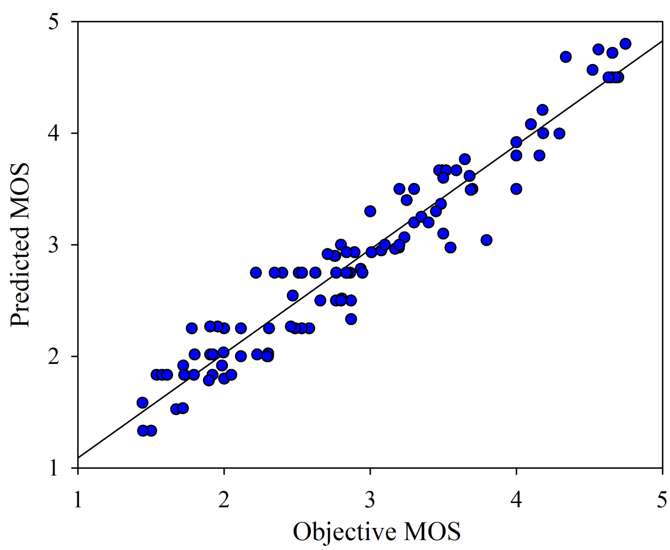
The proposed FIS-PRB model was validated against the objective and subjective datasets given in Chapter 4 (Section 4.3.1). The validation metrics used were R^2 correlation and the root mean squared error (RMSE). These metrics have been used in several related studies including [60,75,87,89]. The reader is referred to Appendix C for a brief description of the R^2 and $RMSE$ metrics.

Figure 5.5 (a-b) shows the correlation between the objectively measured MOS and the predicted MOS using line and scatter graphs. Each point in Figure 5.5 (a) represents the predicted MOS of a particular video clip, and the line represents the measured MOS. The obtained value of R^2 was 93.2% and of $RMSE$ was 0.192.

In addition, Figure 5.6 illustrates the correlation of the subjective MOS against the predicted MOS which reached around 90.6%, with an $RMSE$ value of 0.2661. Due to cost constraints, the subjective dataset was formed by taking a subset of the objective dataset. As a result, there were fewer test conditions generated from the subjective data than from the objective MOS. However, the results obtained using both datasets indicated that the predicted MOS was highly correlated with the measured MOS. Thus, the proposed FIS-PRB model had significant success in predicting the user's perception. These results show a consistent relationship between QoS and QoE for wireless video streaming.



(a)



(b)

Figure 5.5: Predicted MOS vs.Objecively Measured MOS

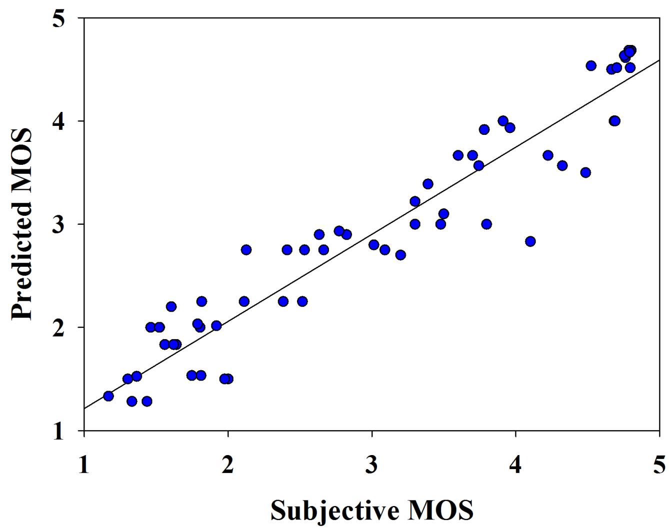


Figure 5.6: Predicted MOS vs Subjectively Measured MOS

5.4.2 Validation by External Dataset

The FIS-PRB model was further validated using an external MOS dataset that has been made publicly available at [143]. This external dataset is for H.264 encoded QCIF videos with network conditions of PLR, MBL for three types of video sequences (CT). In order to perform this validation, the number of input parameters for the FIS-PRB model was decreased to three: PLR, MBL and CT.

Figure 5.7 shows the correlation between the measured MOS and the predicted MOS. The FIS-PRB model achieved an R^2 score of 88.61%. The model showed a consistent relationship between the QoS parameters and the QoE, even with the unseen dataset. However, it is clear that its correlation rate was lower than the one shown in Figure

5.5. One of the reasons for this reduction in correlation could be the new variation in some parameters arising from the external dataset. For example, the considered MBL levels in this external dataset includes new high values in burst lengths compare with our dataset. Therefore, the rule base of the FIS-PRB model should be updated to cover the new MBL values.

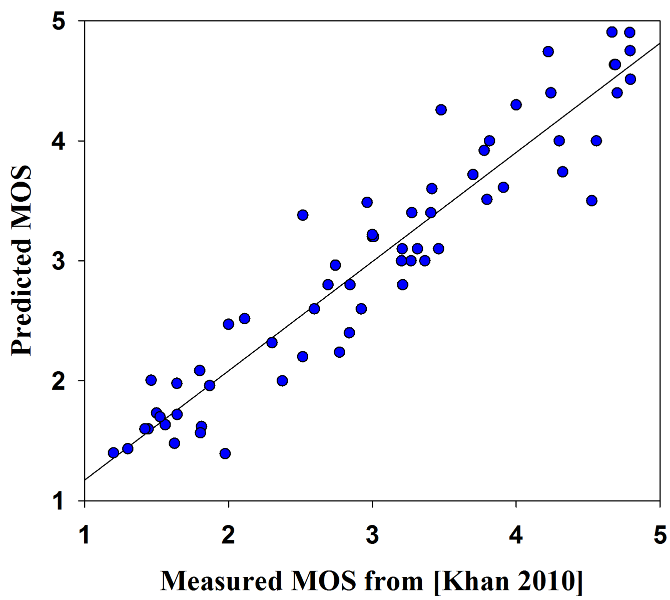


Figure 5.7: Predicted MOS vs. Measured MOS from External Dataset [Khan 2010]

In addition, this external dataset was first generated for building the video QoE prediction model that is proposed in [75]. The model of [75] is based on the regression analysis (RA) method. We compare the performance of the proposed FIS-PRB model with that of the RA-based model from [75] in Table 5.4. The FIS-PRB model outperformed the RA-based model in terms of the prediction accuracy rate. As shown the Table 5.4, the correlation rate of the two model is nearly similar, however, the

RA-based model was already built using this dataset, which is considered as unseen dataset in the case of the FIS-PRB model. Overall, the results indicate that the proposed FIS-PRB model performed well in terms of the correlation coefficient and achieved a reasonable fit, even with the unseen dataset.

Table 5.4: Models performance comparison

Model Type	**RMSE**	R^2
FIS-PRB model	0.329	88.61%
RA-based model	0.373	87.89%

5.5 Performance Comparision with RNN Method

In this section we compare the performance of the proposed FIS-PRB model with that of the random neural networks (RNN) technique [144]. We employ the QoE-RNN model developed in [145] for the comparison. The QoE-RNN model is an implementation of an RNN in the C programming language that is used in the NAPA-WINE project to estimate the QoE experienced by a peer. This model is publicly available at [145]. The measured objective datasets from Chapter 4 (Section 4.3.1) were used for the learning and testing of the QoE-RNN model, and then the results were compared with those for the FIS-PRB model.

The RNN is an artificial neural network (ANN) method that was invented by Erol Gelenbe [144]. Neural networks essentially consist of a system of adjustable parameters called "Weights" and "Biases" that are adjusted during a training phase to provide specific weights and biases for the network. A scalar input is multiplied by the effective weight and added to the bias to produce the target output. The neural network consists of an input layer, one or more hidden layers, and an output layer [144], as shown in Figure 2. We employed a neural network with an input layer of 5 input QoS parameters, two hidden layers with 4 neurons, and an output layer of one output.

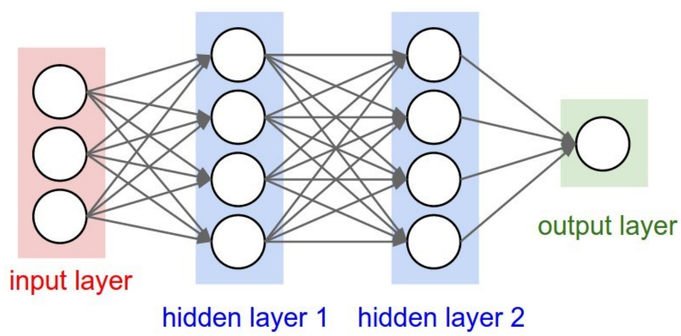


Figure 5.8: Neural network architecture

Figure 5.9 illustrates the correlation of the measured MOS against the predicted MOS which reached around 86.77%, with an $RMSE$ value of 0.399. Table 5.5 presents the comparison between the proposed FIS-PRB and the QoE-RNN models in terms of the R^2 and $RMSE$ metrics. The FIS-PRB model outperformed the RNN-based model in terms of prediction accuracy rate. This result confirms the findings of [6] that the FIS method outperforms other ANN technique due to its modelling capabilities and transparent reasoning process. By comparison with neural networks, the calculations for the FIS can be less complex, especially when triangular MFs are adopted [137].

However, the success of artificial intelligence learning-based techniques are dependant on their ability to fully learn the relationships between QoS and QoE.

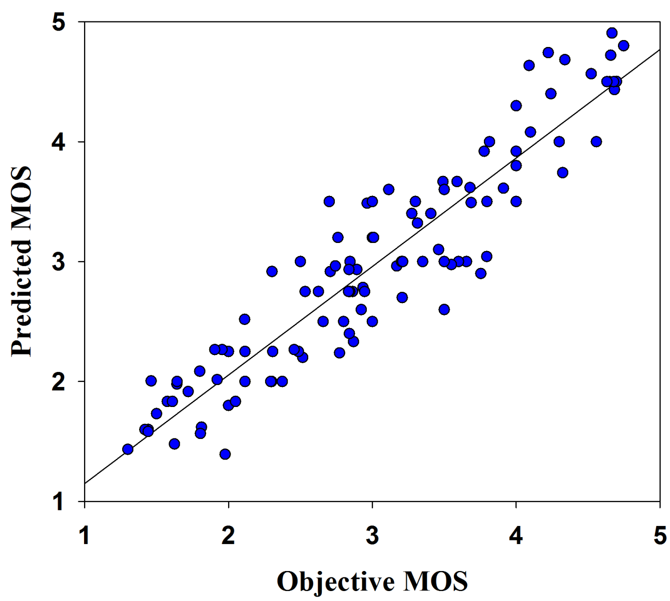


Figure 5.9: Predicted MOS by RNN-based model vs Measured MOS

Table 5.5: Performance comparison with the RNN-based model

Model	**RMSE**	R^2
FIS-PRB model	0.192	93.4%
QoE-RNN model	0.399	86.77%

5.6 3D Video QoE Prediction

As mentioned before, the objective and subjective learning dataset measured in Chapter 4 (Section 4.3.1) that correlates the QoS parameters with the measured QoE was used to build the proposed FIS-PRB model. This dataset includes QoE evaluations for both 2D and 3D H.264 video streams. This chapter only considers the FIS-based quality prediction model in the context of 2D video streaming. It should be mentioned that the FIS-PRB prediction model was also applied to 3D video streaming, and the results were published in [26]. The model shows similar performance as with the 2D video and there was no significant difference other than the observation that the same QoS parameters could cause disparate distortions in 3D quality perception. Note that the FIS-PRB in the context of 3D video model was only validated using the testing dataset because we could not find any public external 3D video QoE dataset that included both AQoS and NQoS parameters.

5.7 Chapter Summary

This chapter presented a hybrid non-reference FIS-based model for predicting video quality by mapping the impact of QoS parameters to the user perceived satisfaction level (QoE). The model developed was based on a predetermined rule-based method (FIS-PRB). The objective MOS learning dataset generated in Chapter 4 (Section 4.3.1) that correlates the QoS parameters with the measured QoE was used to build the proposed FIS-PRB model. For end-to-end quality estimation, QoS parameters from both of the encoding and access network layers were identified. The defined fuzzy membership functions were derived from the QoS/QoE correlation using probability distribution functions.

The proposed model was validated through the correlation of predicted measured QoE using both testing and external datasets. One conclusion drawn was that the validation results indicated that the proposed FIS-PRB model achieved an acceptable correlation rate, even with the unseen dataset. In addition, the FIS-PRB model was validated against an RNN-based prediction model using the same learning and testing datasets. The second important conclusion drawn from the results of this chapter was that the proposed FIS-PRB model showed a significant improvement in performance over the RNN model. This confirmed the known advantages of using fuzzy logic systems, brought about by their special features such as modelling capabilities, lower computational complexity, and transparent reasoning processes. The proposed work provides a proof of concept for a no-reference QoE prediction technique.

The developed FIS-PRB model is good at making decisions with imprecise information and easy to implement. However, it is static because of its fixed, pre-defined fuzzy rules. Thus, the fuzzy rules base needs to be manually updated when the input QoS parameters are changed. Therefore, a self-adaptive model needs to be developed in order to make progress towards the development of a more generic and practical prediction model. The next chapter presents an adaptive video QoE prediction model that can be used in real-time environments with changeable network conditions.

Chapter 6

Adaptive Hybrid Non-reference Model for Real-time Video Quality Prediction

6.1 Introduction

In general, two main criteria can be used to evaluate prediction techniques: prediction accuracy and self-adaptability [146]. The prediction accuracy refers to the ability of the model's estimated score to match that of the measured QoE. Self-adaptability refers to the ability of the model to automatically adapt to new datasets without much complexity and time [146]. The FIS-PRB prediction model proposed in Chapter 5 is good at making decisions based on imprecise information, but it cannot automatically formulate the fuzzy rules required for making such decisions. Indeed, each time the input QoS parameters are updated, the fuzzy rules base of the FIS-PRB must be updated manually.

Adaptive FIS prediction method has been used in several QoS and QoE related mechanisms, such as to perform a dynamic bandwidth and buffer allocation on multimedia traffic [147], to perform congestion control for wireless video streaming [148], and also to predict the QoE of web services [103]. The deep investigation of the literature presented in Section 2.6 found very few papers proposed automated no-reference models for video QoE prediction that take into account both AQoS and NQoS parameters. Also, the accuracy of these models is still questionable. In [75, 88, 89], the Adaptive Neural Fuzzy Inference System (ANFIS) [90] was used to estimate the video QoE. The authors in [75, 89] only considered a single video resolution, QCIF (176×144), and no other higher spatial resolutions were tested. Moreover, they also lack user perception and experimental and validation results.

This chapter presents an adaptive hybrid no-reference model, based on automated FIS, that predicts the QoE of real-time video streaming. The proposed model can automatically formulate the fuzzy rules that are required in order to predict the video QoE. This work provides an important step towards the development of a real time QoE prediction model that can be used at intermediate measurement points along the network path for video streaming. As in previous chapters, the video quality is predicted using a combination of parameters associated with the encoder and the access network for different types of content and resolutions. The QoE predictions are given in terms of the MOS obtained from the objective and subjective tests conducted for Chapter 4. The proposed model is validated by comparing the estimated QoE output with the measured QoE using both testing and external datasets. In addition, the model is evaluated on a real test-bed which was produced for the QoE prediction of real-time wireless H.265 video streaming.

The remainder of this chapter is organised as follows. The experimental set-up and procedure for dataset generation are presented in Section 2. Section 3 discusses the methodology for the proposed video quality prediction system. Sections 4 and 5 are devoted to the model validation and performance comparison. The proposed model is evaluated using a practical test-bed in Section 6. Finally, a summary of this chapter is given in Section 7.

6.2 Experimental set-up and Dataset Generation

The simulation set-up and the generated datasets are as described in Chapter 4, Section 4.2. The objective dataset developed in Chapter 4 for the correlation of the QoS parameters with the QoE are used as a learning set to build the proposed FIS models in this chapter.

Similar to the experiment in Chapter 5, the generated objective dataset is divided into learning and testing datasets. Six video sequences were selected for these datasets, two from each content class as listed in Table 6.1. The Music, Poker and BMX video sequences were used for model learning, while the Fencing, Poznan and Pantomime sequences were used for model validation. Moreover, the subjective dataset was also used for model validation.

Table 6.1: Video sequences chosen and classes assigned

Video Sequence	**TI**	**SI**	**Class**
Music	4.90	74.41	Low Motion
Fencing	7.78	77.20	
Poker	12.20	85.69	Moderate Motion
Poznan	11.53	87.78	
BMX	22.35	99.42	High Motion
Pantomime	37.17	104.43	

6.3 Methodology

The proposed prediction model is based on an automated FIS rule-base method (FIS-A). As mentioned before, the FIS consists of three main modules; fuzzifier, fuzzy inference engine and defuzzifier. In this chapter, the fuzzy inference engine module utilizes an unsupervised technique for extracting the rules from membership functions. Figure 6.1 illustrates the functional block diagram for the proposed FIS-A prediction model.

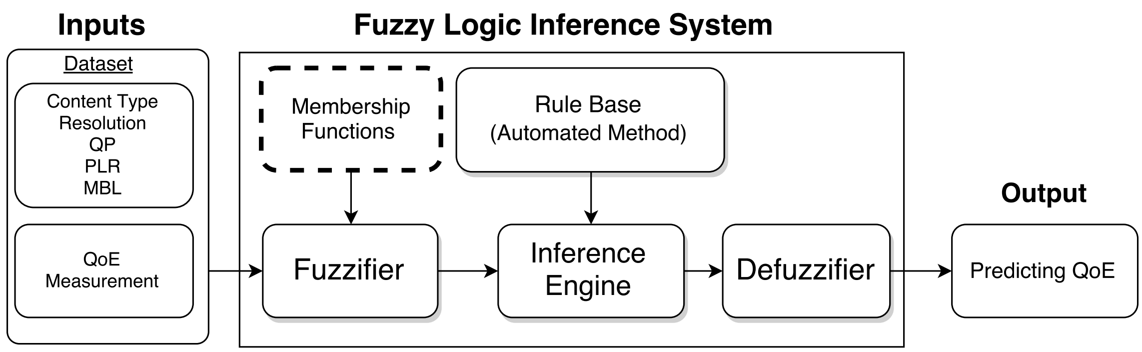


Figure 6.1: Functional block diagram for the proposed FIS-A prediction model

Unlike the model developed in Chapter 5, the FIS-A model automatically adapts the rules for the fuzzy inference system to accommodate changes in the dataset using the learning from example (LFE) approach [149]. An extended and further developed version of the Mendal-Wang method [141, 150] was used to implement the LFE approach. This method is a one-pass approach, centred on obtaining the fuzzy rules from the set of data under examination. The LFE learning procedure only constructs the rules, relying entirely on a complete specification of the membership functions by the analyst. The proposed FIS-A model implemented in the Java language environment. Algorithm 2 gives an overview of the process of designing a fuzzy logic system. In the following sections, we describe the steps for developing the proposed model, its membership functions, and fuzzy rule extraction. Background information about FIS is given in Chapter 2 (Section 2.7).

Algorithm 2: Fuzzy logic system
1. Define the linguistic expressions (Initialisation)
2. Design the membership function using triangle shape (Initialisation)
3. Convert crisp input value to fuzzy value using the MFs (Fuzzification)
4. Automatically extract the fuzzy rule base (Fuzzy inference engine)
5. Evaluate the fuzzy rules in the rule base (Fuzzy inference engine)
6. Aggregate the results of each rule (Fuzzy inference engine)
7. Convert the fuzzy value to crisp output value (Defuzzification)

6.3.1 Identifying the Inputs and Output

The chosen input QoS parameters were the content type (CT), resolution (Re), quantisation parameters (QP), packet loss ratio (PLR), and mean burst loss (MBL). The output was the MOS scores (QoE). Table 4.2 outlines the chosen QoS parameters. Unlike the model in Chapter 5, the proposed method in this chapter allows the incorporation of additional parameters without much complexity and time. The number of input parameters can be can easily increased due to the FIS-A model ability to automatically extract the fuzzy rules from the membership functions. The collected dataset consisted of multiple input and output data pairs of the form:

$$(x^{(t)}; y^{(t)})(t = 1, 2, ..., N), \tag{6.1}$$

where N is the number of data instances, $x^{(t)} \in R^n$, and $y^{(t)}\ R^k$. Once the inputs and the output were identified, both were converted into linguistic expressions to represent the quantification of the QoE values (scores). This was achieved by the design of membership functions for the input and output variables.

6.3.2 Design of Membership Functions

The correlation between the QoS parameters and the measured QoE was transferred into fuzzy membership functions. Membership functions are curves that explain how each linguistic term is linked to a membership value (or degree of membership) between 0 and 1. In this study, the membership functions were derived using probability distribution functions (PDF) [138]. Different PDFs were built for every QoS parameter. The probabilistic information was converted into a fuzzy set by dividing the PDF by its peak. The design of the membership functions was the same as in Chapter 5, Section 5.3.2.

6.3.3 Fuzzy Rules Extraction based on LFE Approach

The inputs and output were divided into fuzzy areas, based on the antecedent and consequent fuzzy sets associated with the rules. The extracted rules took a number of different forms. In the proposed system, fuzzy *IF-THEN* rules of the following form were used to represent the relationship between the input pattern $x = (x_1, ..., x_n)^T$ and the output $y = (y_1, ..., y_n)^T$:

$$IF\ x_1\ is\ A_1^{(i)}\ and\ ....\ x_n\ is\ A_n^{(i)}\ THEN\ y_k\ is\ B_k^{(i)}, \tag{6.2}$$

where $i = (1, 2, ...M)$, M is the number of rules, and i is the index for the rules. There were V fuzzy sets A_s^q, for $q \in \{1, 2, \ldots, V\}$, defined for each input X_s, and

W fuzzy sets B^h, for $h \in \{1, 2, \ldots, W\}$, defined for each output y_c. Moreover, the AND operator, which selects the minimum value of the fuzzy sets was used in this model. The process of extracting the fuzzy rules from the data was performed using the following two steps, as described in [141, 150]:

Step 1:

For a fixed input-output pair $(x^{(t)}; y^{(t)})$ from the dataset (6.1) of N pairs, the membership values $\mu_{A_s^q}(x_s^{(t)})$ are computed for each membership function ($q \in \{1, \ldots, V\}$), and each input variable $s \in \{1, \ldots, n\}$ to give $q^* \in \{1, \ldots, V\}$ such that:

$$\mu_{A_s^{q^*}}\ (x_s^{(t)}) \ \geq \ \mu_{A_s^q}\ (x_s^{(t)}). \tag{6.3}$$

The idea is to select the fuzzy set that achieves the maximum membership value at the data point as the one in the *IF part* of the rule. We call the following rule

$$IF\ x_1^t\ is\ A_1^{q^*}\ and\ ....\ x_n^t\ is\ A_n^{q^*}\ THEN\ y\ is\ centred\ at\ y^{(t)}, \tag{6.4}$$

the rule generated by $(x^{(t)}; y^{(t)})$. Note that each of the fuzzy sets X_s associated to the input variables is characterised by V fuzzy sets A_s^q, where $q \in \{1, \ldots, V\}$, and so the maximum number of possible rules that can be generated is V^n. However, depending on the dataset, only those rules from the V^n possibilities whose dominant regions contain at least one data point will be generated.

In step 1, one rule is generated for each input and output data pair. For each input, the fuzzy set that achieves the maximum membership value at the data point is selected for the *IF part* of the rule, as shown in equations (6.3) and (6.4). This rule will be modified to create its final form in step 2. The weight of the rule is computed using the formula:

$$W^{(t)} \;=\; \prod_{s=1}^{n} \mu_{A_s^q}(x_s^{(t)}). \tag{6.5}$$

The weight $W^{(t)}$ of rule t is a measure of the strength of the points $x^{(t)}$ belonging to the fuzzy region covered by rule t.

Step 2:

Step 1 is repeated for all $t \in \{1, \ldots, N\}$ in order to generate N rules using equation (6.5). Since the number of data points is usually large, many of the rules generated in Step 1 will share the same IF part, but have differing THEN parts, i.e., they will share the same antecedent membership functions, but dissimilar consequent values. In this step, the N rules are divided into groups on the basis of their IF parts. Suppose there are M such groups, and that group i ($i \in \{1, \ldots, M\}$) contains N_i rules of the form:

$$IF\ x_1\ is\ A_1^{(q^i)}\ and\ ....\ x_n\ is\ A_n^{(q^i)}\ THEN\ y\ is\ centred\ at\ y^{(t_u^i)}, \tag{6.6}$$

where $u \in \{1, \ldots, N_i\}$ and t_u^i is the index for the data points in group i. The weighted average of the rules in this conflict group is then calculated using the formula [150]:

$$average^{(i)} = \frac{\sum_{u=1}^{N_i} y^{(t_u^i)} w^{(t_u^i)}}{\sum_{u=1}^{N_i} w^{(t_u^i)}}, \tag{6.7}$$

where the weights $w^{(t_u^i)}$ are computed using equation (6.7) from Step (1). These N_i

rules are then combined into a single rule which has the form:

$$IF\ x_1\ is\ A_1^{(i)}\ and\ ....\ x_n\ is\ A_n^{(i)}\ THEN\ y\ is\ B^{(i)}. \tag{6.8}$$

The output fuzzy set B^i is selected on the basis of finding the set B^{h^*} among the W output fuzzy sets B^1, ..., B^w that satisfies:

$$\mu_{B^{h*}}\ (average^{(i)})\ \geq\ \mu_{B^h}\ (average^{(i)}), \tag{6.9}$$

where $h \in \{1, 2, ..., W\}$, and $B^{(i)}$ is chosen to be B^{h*}.

Overall, the task performed in Step 1 can be summarized as generating one rule from each input-output data pair. The idea is to select the fuzzy set that achieves the maximum membership value at the data point as the one in the *IF part* of the rule, see (6.3) and (6.4). The fuzzy set for the *THEN part* is centered at the data point $y^{(t)}$, and the weight of the rule $W^{(t)}$, is measured by the agreement between the data point $x^{(t)}$ with the *IF part* of the rule, as computed in (6.5).

Since the number N of data points is usually large, many rules generated in Step(1) will share the same *IF part.* The task performed in Step 2 is to merge the rules with the same *IF part* into a single rule. Since the *IF part* has already been determined, this task simply involves determining the fuzzy set $B^{(i)}$ for the *THEN part.* The idea is to place the center of $B^{(i)}$ at the weighted average of the output values for the data points that generate this group of rules, as in equation (6.7), where the weight is the rule weight computed in Step 1. The calculations invoke Equations (6.7) and (6.9), and are repeated for each output value [141,150]. The number of extracted rules was 243 rules. Samples of the extracted fuzzy rules are listed in Appendix E. The final fuzzy rule that was used in the FIS controller had the form:

IF (CT is High motion) AND (QP is Moderate) AND (Resolution is Moderate) AND (PLR is High) AND (MBL is High) THEN (QoE is Bad)

IF (CT is Low motion) AND (QP is Moderate) AND (Resolution is Low) AND (PLR is Low) AND (MBL is Moderate) THEN (QoE is Fair)

6.3.4 Defuzzification: Predicting the Output

The overall result after the inference engine step was a set of fuzzy values that need to be defuzzified to give a crisp output value (QoE), based on the MF of the output variable. In this work, the defuzzification was conducted using the centroid method [141], which was defined in Chapter 5 (section 5.3.4).

6.3.5 Implementation of FIS-A model

Unlike the FIS-PRB model (which was implemented using the Matlab toolbox), the designed membership functions, fuzzy rule extraction and fuzzy controller for the FIS-A model were all implemented using a Java programming language environment.

In terms of software, the cross-platform versatility of the Java programming language allowed it to be embedded into real-time video streaming applications, as presented in Chapter 7. The main components of FIS controller were programmed as Java classes. The main java class (Fuzzy_inference_engine) includes the equations of Mendal-Wang method. Figure 6.2 illustrates the block diagram of the java classes for the FIS-A model. The predicted QoE is then compared with the measured QoE using the same validation metrics that were used for the FIS-PRB model. The validation of the model is discussed in the next section.

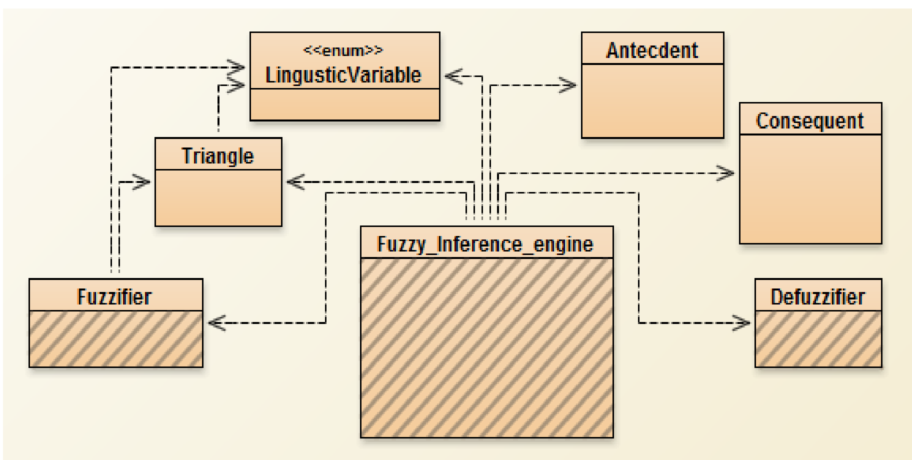


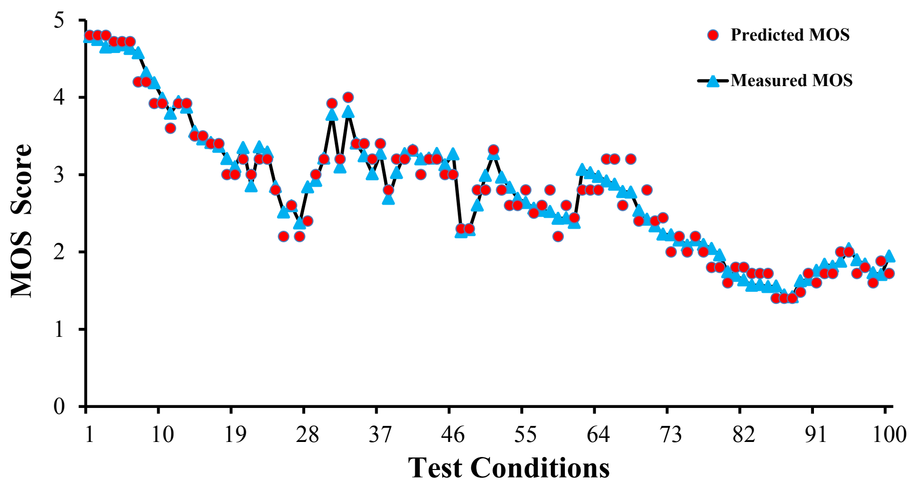
Figure 6.2: Diagram of the java classes for the FIS-A model

6.4 Validation of FIS-A Model

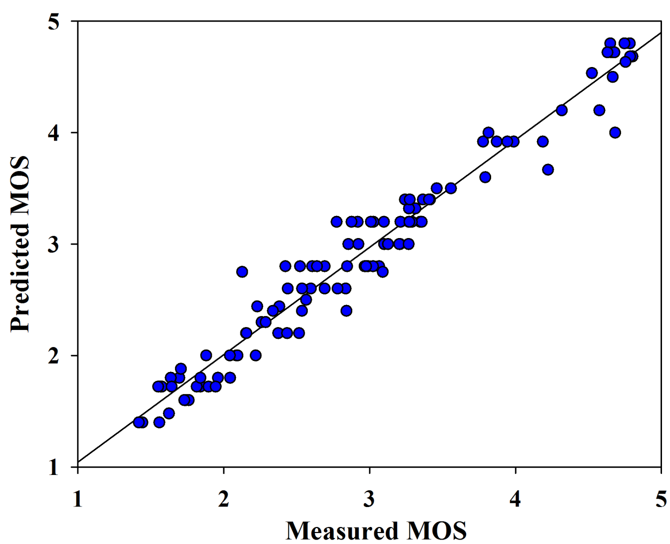
6.4.1 Validation on Testing Dataset

The proposed FIS-A model was validated against the objective and subjective datasets given in Chapter 4 (Section 4.2), which were also used for the FIS-PRB model. The validation metrics used were R^2 and RMSE. Figure 6.3 (a-b) shows the correlation between the objectively measured MOS and the predicted MOS using line and scatter graphs. Each point in Figure 6.3 (a) represents the predicted MOS of a particular video clip and the line represents the measured MOS. The corresponding R^2 value was 95.2% and the value of $RMSE$ was 0.1098.

In addition, Figure 6.4 shows the correlation between the subjective MOS and the predicted MOS which reached around 93.1%, with an $RMSE$ value of 0.2181. The results shown in these figures indicate that the predicted MOS was highly correlated with the measured MOS. Thus, the proposed FIS-A displayed a significant amount of success in predicting the user's perception. These results show a consistent relationship between the QoS and QoE for real-time video streaming.



(a)



(b)

Figure 6.3: Predicted MOS vs. Objectively Measured MOS

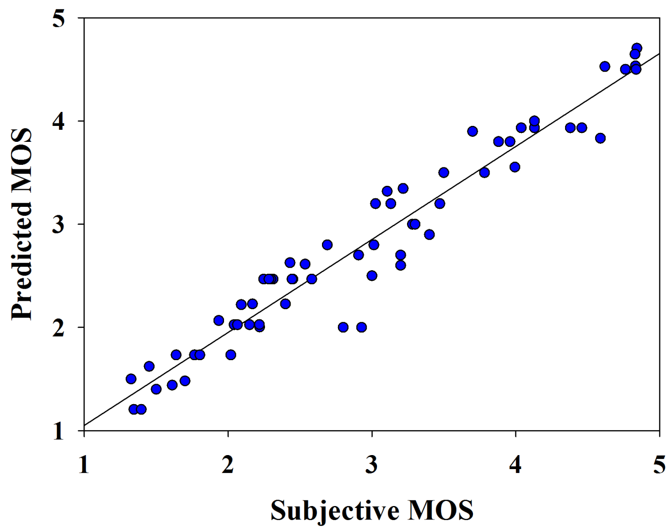


Figure 6.4: Predicted MOS vs Subjectively Measured MOS

6.4.2 Validation on External Dataset

The FIS-A model was further validated using an external MOS dataset that has been made publicly available at [143]. This external dataset is for H.264 encoded QCIF videos with network conditions of PLR, MBL for three types of video sequences (CT). In order to work with this dataset, the number of input parameters for the FIS-A model was reduced to three: PLR, MBL and CT. Figure 6.5 shows the correlation between the measured MOS and the predicted MOS for this dataset. The FIS-A model scored an R^2 value of 91.97%, and demonstrated a consistent relationship between the QoS parameters and the QoE, even with the unseen dataset.

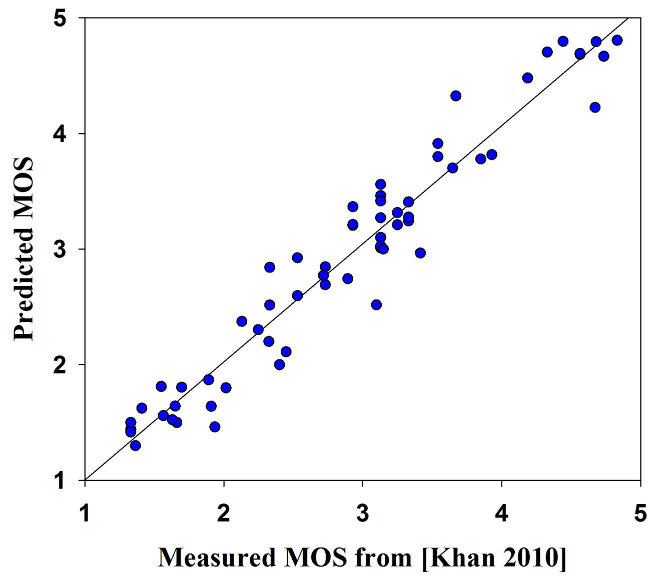


Figure 6.5: Predicted MOS vs. Measured MOS from External Dataset [Khan 2010]

This external dataset was initially generated for the video QoE prediction model that was proposed in [151]. The model in [151] is based on the ANFIS method. Thus, our analysis in this section compares the performance of the proposed FIS-A model with that of the ANFIS-based model from [151]. Table 6.2 shows the performance comparison for these models. Unlike the ANFIS-based model which was already learnt from this external dataset, the proposed FIS-A model performed well in terms of the correlation coefficient, and achieved a reasonable fit with the external dataset.

Table 6.2: Models performance comparison

Model Type	**RMSE**	R^2
FIS-A model	0.217	91.97%
ANFIS-Based	0.315	90.12%

6.5 Performance Comparison with FIS-PRB Model

The validation results for the proposed FIS-PRB and FIS-A models showed that the measured MOS scores were highly correlated with the predicted MOS scores. This indicates that the proposed FIS-based models were successful in reflecting the user's perception. It can be observed that the correlation rates are very similar when the validation was based on the testing dataset. However, the two models did not exhibit similar performances when they were validated by the external unseen dataset that included new variations in the values of the QoS parameters. The result displayed in Table 6.3 indicates that the FIS-A model outperformed the FIS-PRB model. This is because of its ability to automatically adapt the fuzzy rules necessary for predicting the output using the LFE method. By contrast, as the FIS-PRB model is based on predefined fixed fuzzy rules, the fuzzy rules base needed to be manually updated when the values of input QoS parameters were changed.

Table 6.3: Performance comparison using external dataset

Model	RMSE	R^2
FIS-PRB model	0.329	88.61%
FIS-A model	0.217	91.97%

6.6 Model Evaluation on Real Test-bed

The proposed FIS-A model was further tested with close to real network conditions, using a real-time transmission of H.265/HEVC video over a wireless network in order to measure its performance for an on-line video transmission scenario. The test-bed created in Chapter 4 (Section 4.4) was used in this chapter to evaluate the FIS-A model. In the following subsections, the evaluation of the test bed is described; then the analysis of the experimental results is presented.

6.6.1 Experimental Set-Up

The experimental set-up is the same as in Chapter 4, Section (4.4.1). Figure 6.6 shows the topology of the physical test-bed environment. To introduce network impairments, the SCE [135] WAN emulator was attached, virtually, to the outbound connection of the sender using the bridging mode. The chosen QoS parameters are listed in Table 6.4. The collection of input parameters for the proposed FIS-A model was decreased to the three inputs: PLR, R, and CT.

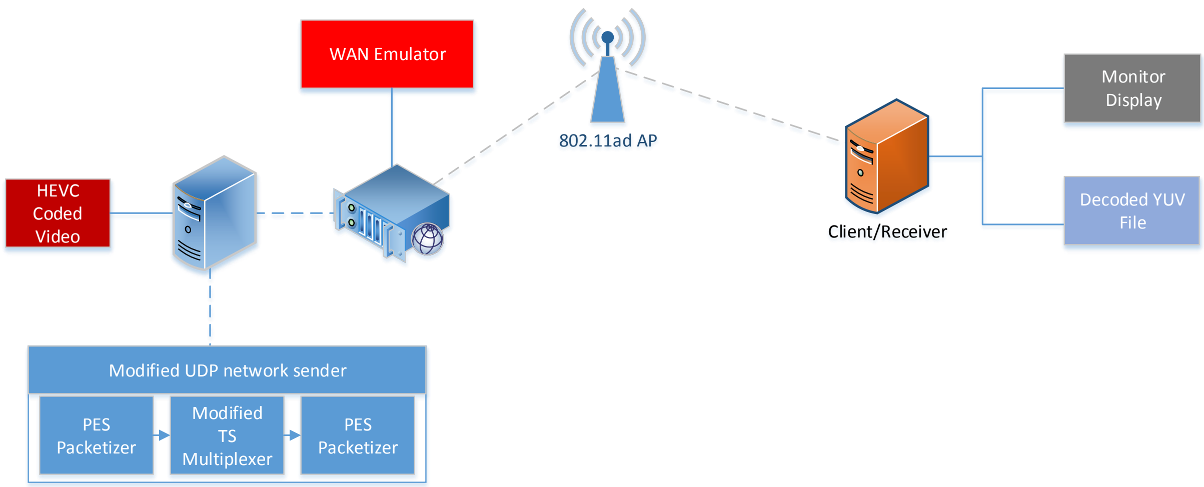


Figure 6.6: Topology of the physical testbed environment.

Table 6.4: Experimental QoS Parameters

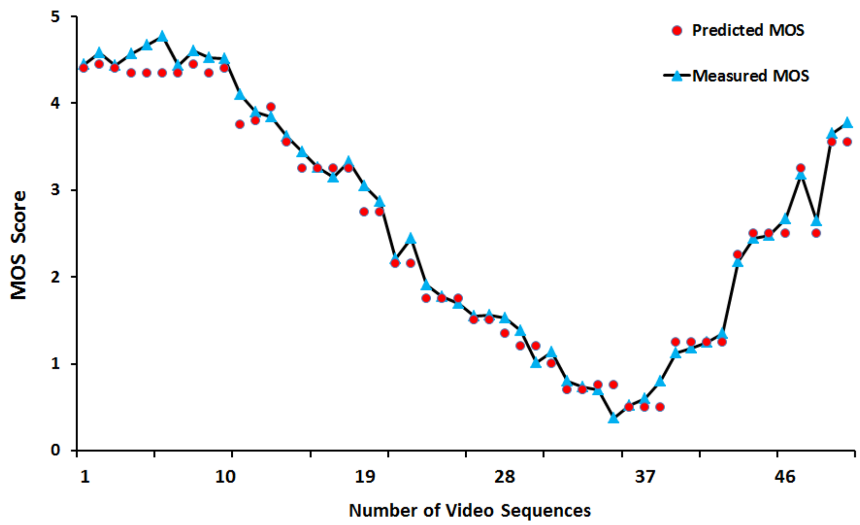
Parameters	**Values**
Content Type (CT)	Low, Moderate, High motion
Spatial resolution (R)	SD (480p), HD (720p), 4kUHD (2160p)
Packet loss ratio (PLR)	1, 2, 5,.....,9

With the exception of encoding, which was done off-line, all other processes such as the modified solution for MPEG2-TS transmission (server side), the demultiplexing of received MPEG2-TS, and the decoding of the H.265/HEVC elementary stream (ES) (client side) operated in real-time. The decoded (distorted) YUV was then compared against its original (YUV) to obtain video quality results with respect to the structural similarity index metric (SSIM) [56]. More details of the test-bed's implementation can be found in Section (4.4.1).

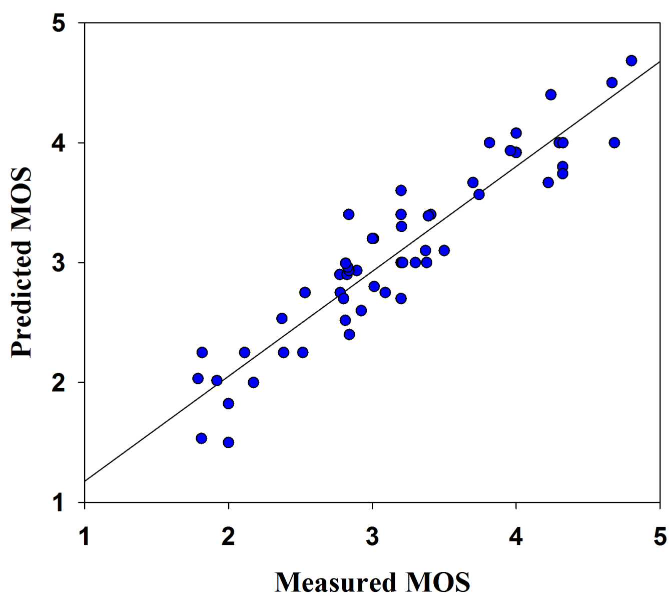
6.6.2 Experimental Results

The main objective of this approach was to identify the relationships between the QoS parameters that affect QoE and the overall perceived QoE. We achieved high accuracy and made dynamic adaptation possible using this model. Figure 6.7 illustrates the correlation of the measured MOS with the predicted MOS. The correlation factor (R^2) and the RMSE were used to validate the proposed FIS-A model. Since the R^2 score was 90.88% with an RMSE value of 0.3297, the proposed system achieved an acceptable degree of success in predicting video quality in the real-time scenario.

Table 6.5 shows a further comparison between the test-bed's result (on-line scenario) with the result of the model validation by the external dataset (off-line scenario), which is presented in Section 6.4.2. In both scenarios, the results demonstrate a consistent relationship between QoS and QoE for real-time video streaming. However, the FIS-A model is expected to be more accurate in the off-line scenario. This is because, in case of real-time streaming, the computational resources required for processes such as decoding complexity increase with the spatial resolution. In addition, other processes such as inbound network queue packet processing and video display at the required refresh rate compete for resource allocation. This is quite different from the off-line scenario, where the QoS parameters are normally fixed.



(a)



(b)

Figure 6.7: Predicted MOS vs.Objecively Measured MOS

Table 6.5: The Performance of FIS-A model in two scenario

Model	**RMSE**	R^2
Off-line scenario	0.217	91.97%
real-time scenario	0.329	90.88%

6.7 Chapter Summary

This chapter presented an adaptive hybrid non-reference FIS-based model for predicting video quality by measuring the impact of QoS parameters on the user perceived satisfaction level (QoE). The developed model was based on an automated rule-base method (FIS-A). The objective MOS learning dataset presented in Chapter 4 for the correlation of QoS parameters with the measured QoE, was used to build the proposed FIS-A model. QoS parameters from both of the encoding and access network layers were identified for end-to-end quality estimation. The fuzzy membership functions were derived from the QoS/QoE correlation using probability distribution functions.

The proposed model was validated through the correlation of the predicted and measured QoE using both testing and external datasets. The results showed a high prediction accuracy. The model achieved a reasonable fit, even when the external dataset. In all cases, the FIS-A model outperformed the FIS-PRB model, especially when the models validated by the external unseen dataset. Although the FIS-PRB model is good at making decisions on the basis of imprecise information and is easy to implement, its fixed pre-defined fuzzy rules mean that it is static. By contrast, even though the FIS-A model is more complicated, it can adapt to changing network conditions due to its ability to automatically formulate the fuzzy rules necessary for predicting the output.

The FIS-A model was also evaluated on a real test-bed for the QoE prediction of real-time H.265 video streaming. The results showed a strong correlation that could effectively define the encoding and network impacts on the video QoE. The FIS-A showed a consistent mapping between the QoS parameters and the QoE. This work achieved our research objective of developing an adaptive non-reference QoE prediction model. This is significant because the higher accuracy in determining the expected level of QoE enables us to make efficient decisions regarding the provisioning of network resources whilst keeping the customer satisfied. Thus, our model will be of significant help in the design of applications for the QoE enabled optimisation of video delivery discussed in Chapter 7.

Chapter 7

QoE-enabled Applications for Optimising Video Delivery: Applicability Examples

7.1 Introduction

This chapter demonstrates two applications of the adaptive video quality prediction model developed in Chapter 6. These are (1) a QoE-enabled transport optimisation scheme for real-time SVC video delivery and (2) a QoE-enabled resource utilisation scheme for mobile video delivery. The proposed applications demonstrate the use of the non-reference QoE prediction model for optimising mobile video delivery. These applications show how QoE is used to optimise video delivery and utilize existing network resources. The following sections discuss the two applications.

7.2 Application 1: QoE-enabled Optimisation Scheme for Real-Time Multi-layer Video Delivery

As discussed in Chapter 2, the QoS is the major determining factor for the QoE. Thus, finding the correlation between QoS and QoE is a significant first step towards a system that can manage video services in an efficient way. Most studies in the literature that investigate the optimization of video delivery mainly focus their attention on throughput maximization [152, 153]. However, real-time applications such as video streaming are highly sensitive to dynamic changes in the application requirements (e.g., data rate, packet loss, etc). As a result, throughput-based optimization provides a suboptimal solution with respect to the user-perceived quality [154, 155].

Other researchers have focused their attention on QoE rather than on traditional QoS parameters [74, 75, 156–158]. However, these algorithms are difficult to implement in real-world scenarios (i.e., uncontrolled environments) or only provide partial solutions such as in [159]. The work in [159] demonstrates the QoE-aware traffic management for mobile video delivery within the MEDIEVAL architecture (The EU project MultimEDia transport for mobIlE Video AppLications). The MEDIEVAL architecture applies a crosslayer framework to efficiently handle video traffic in the mobile network. The framework proposed a feedback mechanism to link all the QoS and QoE aspects. However, the framework does not address specific device participation in the framework. For instance, the proposal does not define the functions of interconnecting elements in the network in realizing QoS management. Moreover, the authors did not consider multi-layer video stream in high resolution.

We present a QoE-enabled Transport Optimization Scheme (QETOS) for real-time multi-layer video delivery. The objective of this study is to give an applicable example that allows the online scenario to be considered. The proposed scheme optimises the video traffic by mapping QoS parameters (from both application and access network level) to the QoE without penetrating the video packets. It takes the advantage of the partitions of the scalable video coding (SVC) [160] that organize the video into layers of different importance, thereby facilitating the rate adaptation of the video streams. The proposed work is an application of the developed FIS-A model in Chapter 6.

7.2.1 Characteristics of scalability extension of H.264/AVC

The scalable extension of H.264/AVC [160] is related to H.264/AVC [49], and so it is also divided into two parts: the Video Coding Layer (VCL) and the Network Abstraction Layer (NAL) [161]. There are three main scalability aspects, i.e., temporal, spatial and quality scalability, for VCL. A typical SVC stream includes one base layer, and one or more enhancement layers as shown in Figure 7.1. The SVC base layer has the lowest bitrate and requires the least resources. Adding more enhancement layers to the base layer provides better video quality at the expense of a higher overall bitrate. The SVC layers are mapped to a finite number of priorities which are marked in the video packet headers for transmission [160]. The proposed system in this study takes the advantage of the H.264/SVC partitions. Layered encoding is used for adapting the video streams to the network dynamics.

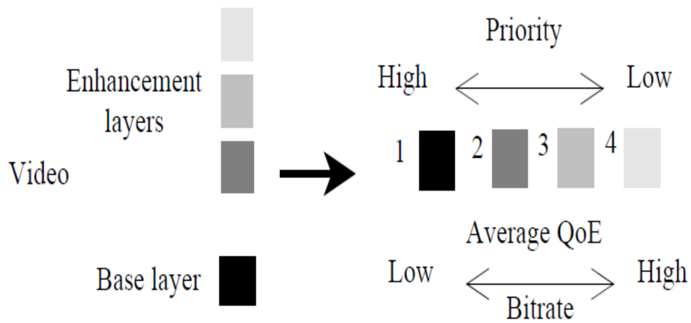


Figure 7.1: Layer prioritization for SVC video streaming.

7.2.2 QoE-based transport optimisation scheme

The proposed QoE-based transport optimisation scheme for real-time SVC delivery (QETOS) is a practical example of optimising the video traffic by mapping the QoS parameters to the QoE without penetrating the video packets. In this study, video streams are encoded in a layered manner in such a way that every additional layer increases the perceived quality of the stream.

7.2.2.1 System Structure

Normally, the implementation of advanced functionalities such as prediction and policing on each router in a network introduces a prohibitive workload for the router. This consequently affects the performance of the router in terms of delay and packet loss. Furthermore, the operating system of real-world routers has a limited capacity and implements only forwarding and switching logic. An approach for the introduction of additional functionalities into a network is the use of Software Defined Networking (SDN) logic. This approach allows all functions to be implemented by a remote server, effectively separating the forwarding and switching logic of devices in the network. Figure 7.2 shows a conceptual diagram for the proposed scheme.

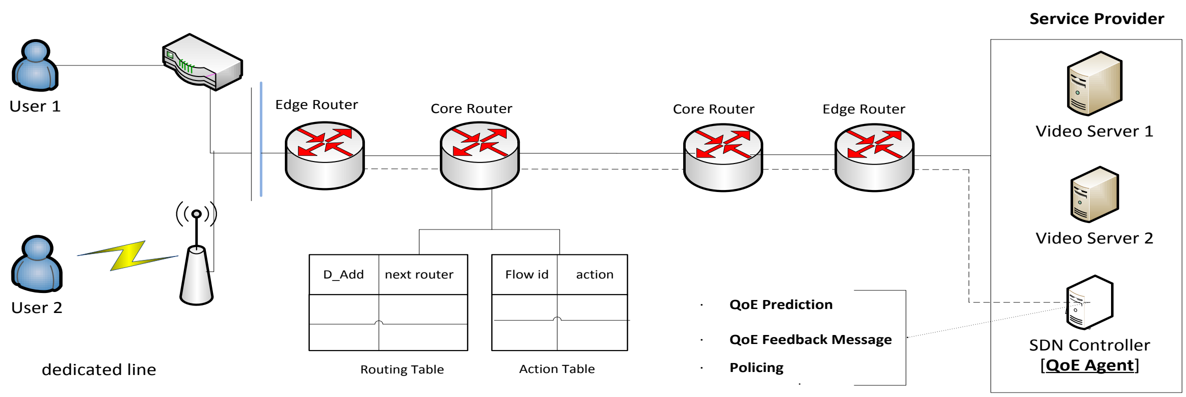


Figure 7.2: A conceptual diagram for the proposed scheme.

We propose that the main functionalities of the QETOS will be enabled by a QoE agent that is placed in the SDN server. The QoE agent will perform a per-flow analysis and make scheduling decisions. We assume that the routers use a Deep Packet Inspection (DPI) engine at its entry point to inspect the network traffic for the identification of video flows and extract relevant per-flow QoS parameters. These parameters are the packet loss rate (PLR) and mean burst length (MBL). The QoE agent requests the QoS parameters from the router by sending a RTP Control Protocol (RTCP) report. The primary function of RTCP is to provide feedback on the

QoS in media distribution by periodically sending statistics information to participants in a streaming multimedia session [162]. Specifically, the proposed QoE agent will perform three main functions:

- **QoE Prediction**: the QoE estimation performed by the FIS-A model developed in Chapter 6. The measured QoS parameters were constantly fed to the QoE agent for use as inputs to the FIS-A model in order to estimate the video QoE on a per flow basis. The number of input parameters for the FIS-A model was reduced to three: PLR, MBL and CT. Figure 7.3 shows the functional block diagram for the proposed QoE estimation model.

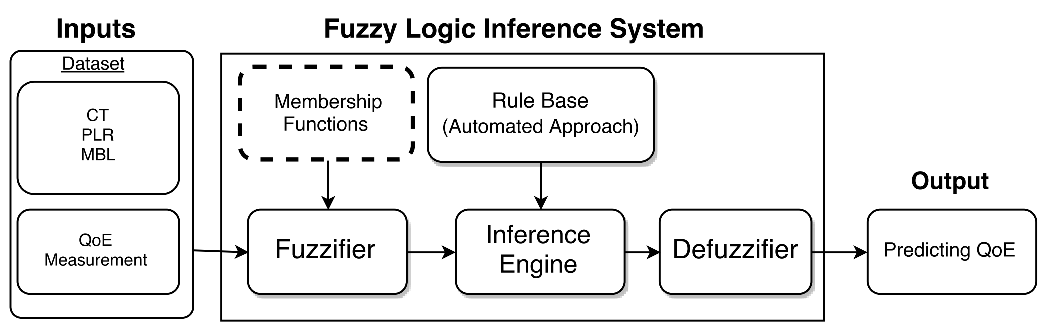


Figure 7.3: Functional block diagram for the proposed FIS-A prediction model

- **Video Layer Dropping**: In the case of high network congestion, if the predicted QoE is below the pre-determined threshold, the low priority video layer will be selectively dropped in order to ensure the correct transmission of the video layer with higher priority, thereby guaranteeing a minimum video quality. The policy will be applied the next time the flow arrives at the edge router.

- **QoE Feedback Messages**: at every time step, the QoE agent automatically sends a QoE feedback message about the traffic status to the network administrator to solve the congestion issue. If reducing the video layers had no effect on the MOS, the network administrator will change the traffic route to the least congested route.

7.2.2.2 Optimisation Mechanism

The optimisation was based on changing the video layer number according to the current predicted QoE. The work was evaluated by the following procedure:

- At the beginning, the bandwidth was set at 5 Mb/s which was just enough to allow the video to stream and to accommodate other background traffic.

- During video streaming, the QoE agent runs periodically, i.e. every 5 seconds. In each period, the QoE agent retrieves information about the video stream (content type). It also retrieves the QoS parameters (PLR and MBL) from the router by sending RTCP report. Both types of parameters are required by the QoE agent to predict the current QoE of the video streams.

- After a few minutes of the video streaming session had elapsed, the network was deliberately congested causing significant degradation in the video quality.

- Once the QoE is dropped to or below the minimum threshold MOS value (3.5) for a pre-determined duration, the QoE agent sends a request to the router to drop the least important video layers (i.e., the enhancement layers) from the stream. According to Jammeh et al [156], the minimum threshold for acceptable quality corresponds to the MOS value of 3.5. Algorithm 1 shows simple pseudo-code for the policing action that is sent from the QoE agent to the router. The policy will be applied the next time the flow arrives at the edge router.

- The newly determined layer number was subsequently used in order to maximise the delivered video QoE above the threshold.

- The QoE agent sends a feedback message to the network administrator to deal with the congestion issue. If the route is highly congested, the administrator changes the current route to the least congested route.

Algorithm 1: Simple pseudo-code for the switching logic

Input= the predicted current video QoE value

If (QoE >3.5) **Then** Action =”No action”

If ($3 \leqslant$ QoE $\leqslant 3.5$) **Then** Action =”Drop enhancement layer 2”

If (QoE <3) **Then** Action =”Drop enhancement layers 2 and 1”

Output= sendToRouter (Action)

7.2.3 Experimental Set-Up

The YUV video sequences used for this particular evaluation were ”Animation” (low motion) and ”Sport” (high motion) (238 frames) in HD format (1280×720 pixels).

The Joint Scalable Video Model (JSVM) [163] was used to encode/decode the SVC video streams. The video sequences possessed three layers of spatial, temporal and quality availabilities. They were encoded using a GOP size of 16 frames and the IPPP structure. After encoding the raw YUV video with different encoding parameters, BitStreamExtractor (provided in JSVM) and F-N Stamp (provided in SVEF [164]) were used to generate an original NALU Trace file, which is used as an input file for the simulator.

To demonstrate the strength of our proposed framework, we integrated the SVEF (Scalable Video-streaming Evaluation Framework) [164] with network elements and the fuzzy logic engine (the FIS-A model) from which was written in Java. The key network components simulated using Java classes included a video server, 3 clients (end-users), three routers, links between the routers, and an SDN enabled controller. Each router was capable of communicating with the SDN controller in order to share important network conditions. The QoE agent (which includes the FIS-A model) was placed in the SDN controller that was connected to the routers.

The ability of the proposed scheme to maximize the delivered QoE for each user was tested by varying the available bandwidth. Specifically, we varied the available bandwidth (βe-a) between the router and the video server. The βe-a was assumed to be limited with a high probability of a bottleneck. The βe-a bandwidth was set to different capacities from 1 Mbps to 10 Mbps. Two scenarios were considered. The first was video streaming in a Best Effort (BE) scenario (without control). The second scenario included the application of the QETOS to manage the video traffic.

7.2.4 Results

The performance of the QETOS scheme was compared to the best-effort scenario (without control) in terms of the delivered quality MOS. Figure 7.4 clearly demonstrates the improvement in the MOS using the QETOS, especially for higher PLRs (greater than 0.1 (10%)). Nevertheless, it was difficult for the QETOS to greatly improve the delivered MOS for PER beyond 10%. This is because the video stream was so greatly impacted by the network congestion, and reducing the video layers had no effect on the MOS.

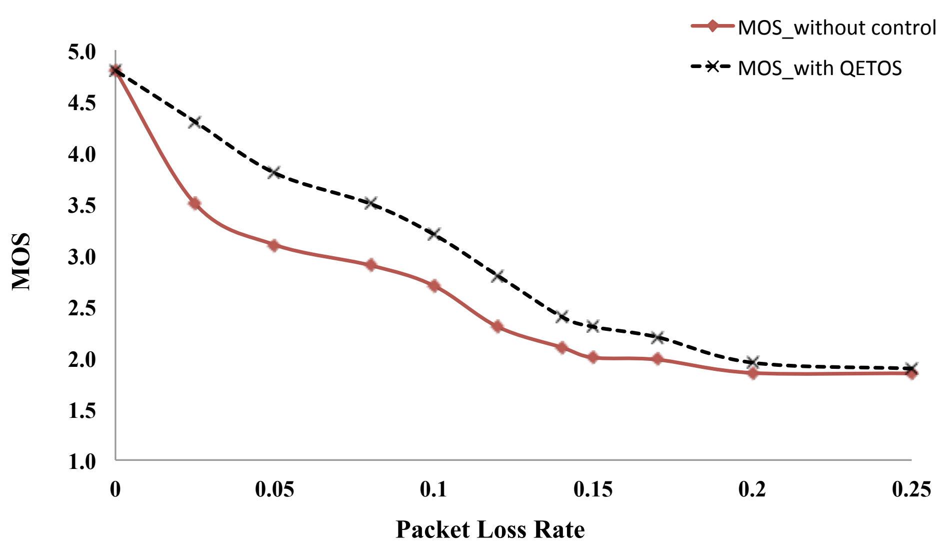


Figure 7.4: QETOS scenario vs. best-effort scenario in case of different PLR

In Figure 7.5, the QETOS quickly detected a change in the available bandwidth and adapted the numbers of sent video layers accordingly, resulting in an optimized QoE for the available bandwidth. The results showed that the proposed scheme provided an MOS gain of over 1 compared to the best-effort scenario. In the best-effort scenario (without control), the QoE suffered a sudden, drastic drop as the bandwidth was constricted. By contrast, with the QETOS scenario, the delivered quality was gracefully adapted to the available network bandwidth.

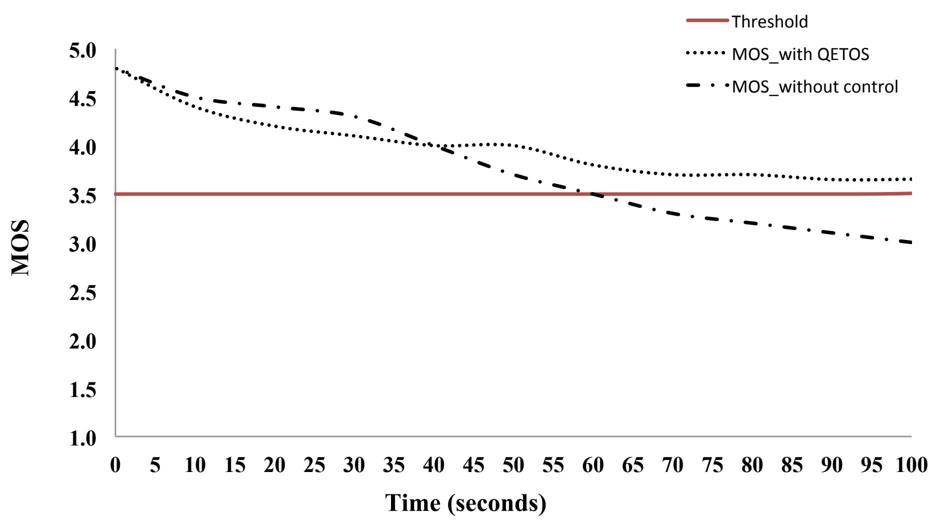


Figure 7.5: QETOS scenario vs. best-effort scenario

Overall, the results showed that the scheme was responsive to the available network conditions and delivered the optimum quality for a given available network bandwidth. There was a clear improvement in video quality when the less important video layers (enhancement layers) were dropped to compensate for network congestion. The proposed QETOS system enhanced video streaming performance and minimized the side effects of congestion on user perceived quality.

7.3 Application 2: QoE-enabled Efficient Resource Utilisation Scheme for Mobile Video Delivery

Wireless resources are still at a premium, firstly, because of lower average cellular bandwidths under variable link conditions, and, secondly, due to the increasing traffic demands, especially in the form of streaming video [1]. This trend will have a highly significant impact on both the user's perception and network resources. Increasing the number of base stations, instituting data caps, and limiting the access to some mobile services are not the only solutions for this challenge. Consequently, the issue of resource utilisation has become a more pressing issue than ever for networks in general, and wireless access networks in particular.

Accordingly, research on network resource utilisation has introduced several techniques for more efficient power and bandwidth consumption. The majority of these techniques are based on QoS and network parameters, where improving the spectral efficiency, jitter, service latency, etc., is the way to provide an acceptable level of service to the consumers. For instance, the content-aware approach in [165] used content quality as the basis for rate control to improve the energy efficiency for mobile IPTV. In [166], a bitQoS-aware resource utilisation was presented, which aimed to increase user throughput and reduce the probability of packet drop. Their method was based on adaptively matching the QoS requirements of the user application bits to the characteristics of the Orthogonal Frequency Division Multiplexing (OFDM) subcarriers in a mixed-traffic environment. The authors of [167] presented an adaptive bandwidth allocation scheme based on the queue length and the probability of packet loss. However, little attention has been paid to the resource utilisation approach that is based on the user's QoE. The QoE has become the prime performance criterion for media delivery technologies.

Most existing video QoE-aware approaches for wireless resource management, such as [154, 168–170], have mainly focused on offline video quality metrics (e.g. PSNR and SSIM), which are full-reference methods. These visual quality assessment metrics require the output of the decoder and the original video reference for their measurements. Metrics of this type are more suitable for performance analysis rather than practical quality assessment which can be applied for a real-world application [171, 172]. The offline QoE estimation method does not provide a practical solution because it assumes that the QoE is known to the access network before transmission. By contrast, the online QoE estimation method (no-reference) fills this gap by providing relatively less accurate, but sufficiently reliable, measurements for real-time video streaming.

In this work, we present a novel resource utilisation scheme based on the online QoE estimation of mobile video streaming using the FIS-A model that is presented in Chapter 6. The main idea is identifying the receiver's transition between the Adaptive Modulation and Coding (AMC) regions based on the video QoE, rather than network parameters (QoS). A mobile video transmission is simulated, through which the correlation between the receiver's signal-to-noise ratio (SNR) and perceived video quality is identified.

It should be noted that this work was completed in collaboration with the multimedia group from the Surrey University, UK. They made their WiMAX model available for this research. This work has been published in [31] and [32].

7.3.1 QoE-enabled Resource Utilisation Scheme

The proposed scheme facilitates the proposal of bandwidth and power efficient resource utilisation applications on WiMAX access networks, based on online QoE estimation. Offline QoE estimation relies on a full-reference quality metric. This method lacks a practical solution because it assumes that the QoE is known to the access network before transmission. Moreover, in real-time video streaming, it is impractical to assess the QoE at the user's end. The online (no-reference) estimation model fills this gap by providing relatively less accurate, but sufficiently reliable, measurements for real-time video streaming.

7.3.1.1 SNR thresholds for band AMC transitions

Adaptive Modulation and Coding (AMC) permits the transition of a mobile receiver between MCS regions. Hence, the receiver's SNR threshold levels provide a lower bound for the operational downlink MCS region, as illustrated in Figure 7.6. If the SNR steps out of the designated operating region, the receiver requests a change to a new operational MCS region. The IEEE standard [173] suggests that the SNR thresholds bounding each operational region shall be identified at a bit-error-rate (BER) of 10^{-6}. By contrast, our proposed scheme identifies the SNR thresholds based on the user's perceptual QoE. Accordingly, both methods for assigning the operational MCS regions were compared using our tested WiMAX model environment. The first method was based on the threshold of BER=10^{-6}, and the second was based on the threshold of VQM = 0.8, which corresponds to 4 on the scale. We chose this QoE threshold since it is regarded by the research community to be a sufficiently high and acceptable perception level for users, as emphasized by [7].

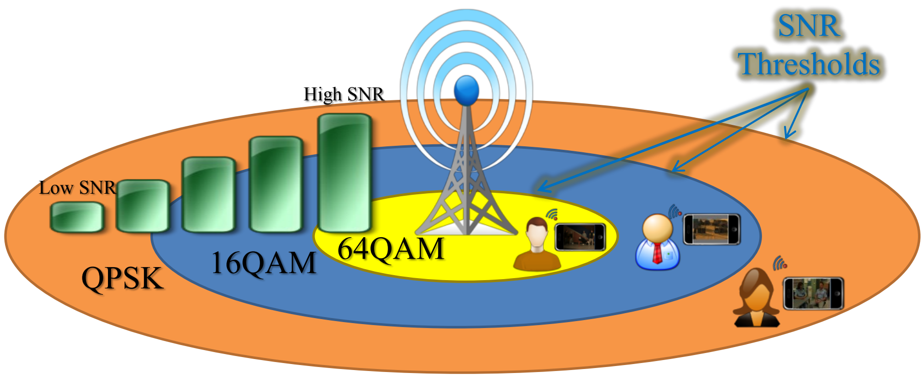


Figure 7.6: The SNR thresholds for operational MCS regions

7.3.1.2 QoE Prediction Methodology

In this study, the online QoE estimation was conducted by the FIS-A model developed in Chapter 6. This prediction model can automatically formulate the fuzzy rules that are used for predicting the output (QoE). The FIS-A model was updated to include two additional parameters, which are the MCS, and radio channel environments (RCE) (pedestrian and vehicular). Thus, the input parameters were the video content type (CT), packet error rate (PER), MCS, and RCE. The output was

MOS scores (VQM mapped to MOS). Figure 7.7 shows the functional block diagram for the QoE prediction model. Once the inputs and outputs were identified, we categorized the gathered input/output data into linguistic labels (low, moderate, high) to represent the quantification of the values. This was achieved by the design of membership functions for the input and output variables. The membership functions of the CT and PER are already designed and presented in Chapter 6, while for the MCS and RCE are shown in Figure 7.8 and 7.9 .

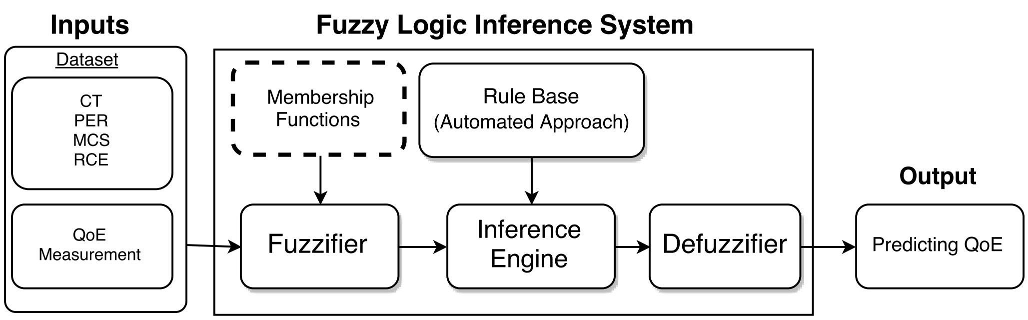


Figure 7.7: Functional block diagram of the FIS-A prediction model

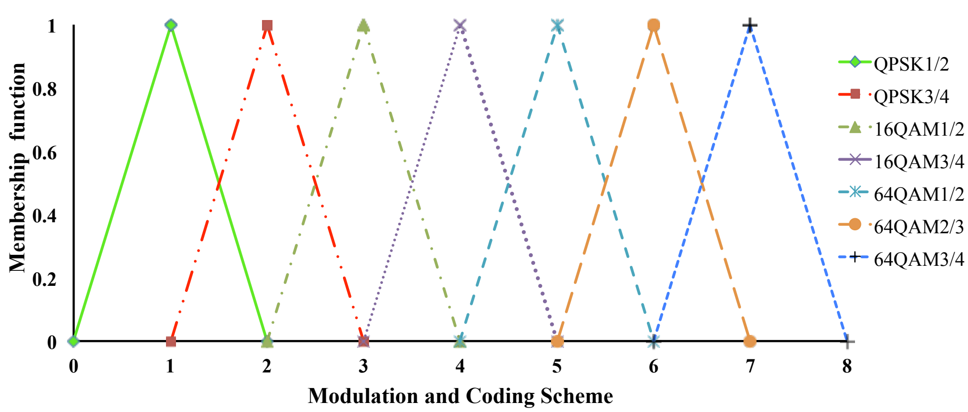


Figure 7.8: The membership functions of MCS

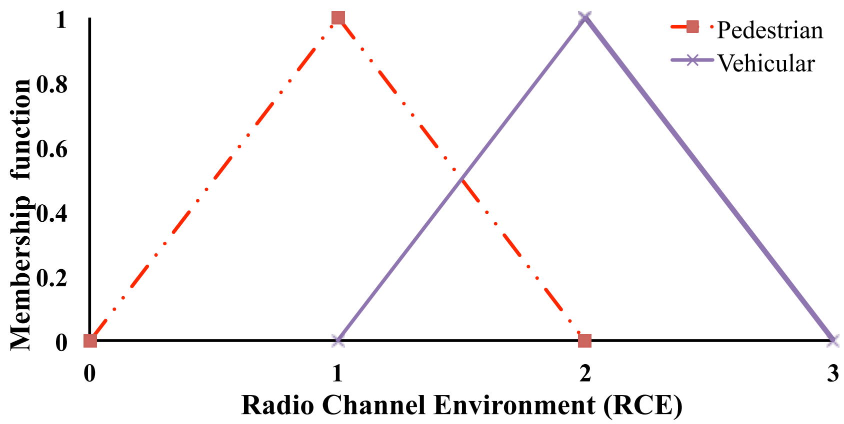


Figure 7.9: The membership functions of RCE

7.3.2 Experimental Set-up

Two test video sequences were used (*Interview* and *GT Fly*) at a resolution of 720×576 Standard Definition (SD) with 25 frames/second (fps), giving a total of 250 frames. Based on the H.264 standard [47], the H.264/AVC Reference JM Software encoder [120] was used for the source video coding. The two video sequences were classified in terms of two types of video content: high and low motion. Based on a recommended bitrate for SD video [42], the bitrate allocated to the Interview video is 1Mbps, whereas the GT Fly video is allocated 2Mbps.

Wireless transmission of the encoded H.264 bit-stream was simulated over a WiMAX (IEEE 802.16e) channel. The WiMAX simulation model presented in [174] was used for the simulations. This model explores two radio channel environments (pedestrian and vehicular) in several modulation and coding schemes (MCS) at different levels of SNR. The block diagram in Figure 7.10 depicts the simulations conducted to map the chosen independent variables (QoS) to the targeted dependent variable (QoE).

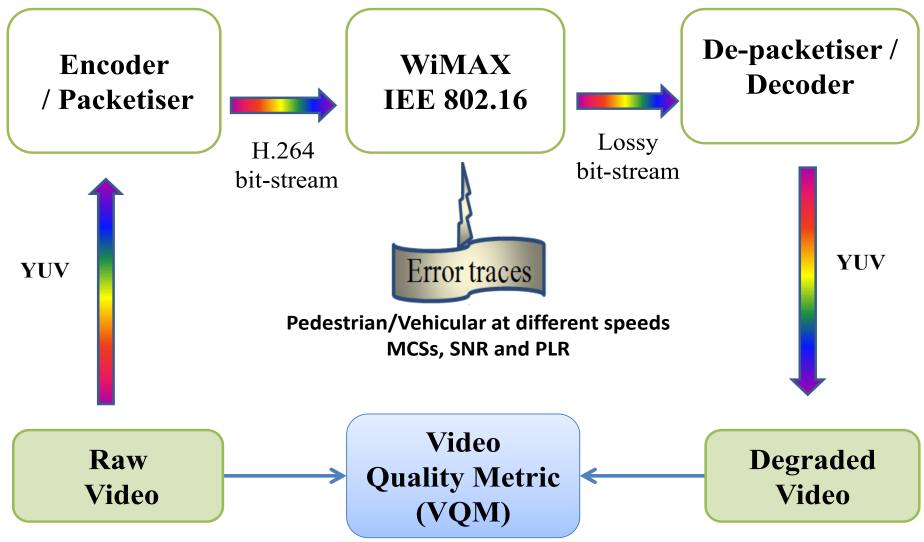


Figure 7.10: Conceptual illustration of the simulation scene

The video bit-stream was channel-coded, modulated, and exposed to different error traces. A number of error bits are introduced in the 15 seconds error trace depending on the MCS used. Each error trace was repeated 10 times at randomly chosen starting positions of the video bit-stream. This is because in real-life scenarios bit errors could occur at any time during transmission. The SNR-to-BER relationship in this model is shown in Figure 7.11 for the vehicular scenario. Later, based on the target video packet size, the bit-error traces are used to produce packet-error rate traces. Thus, in the end, the simulation model [174] measured the PER for each BER. The PER parameter is required for the prediction model.

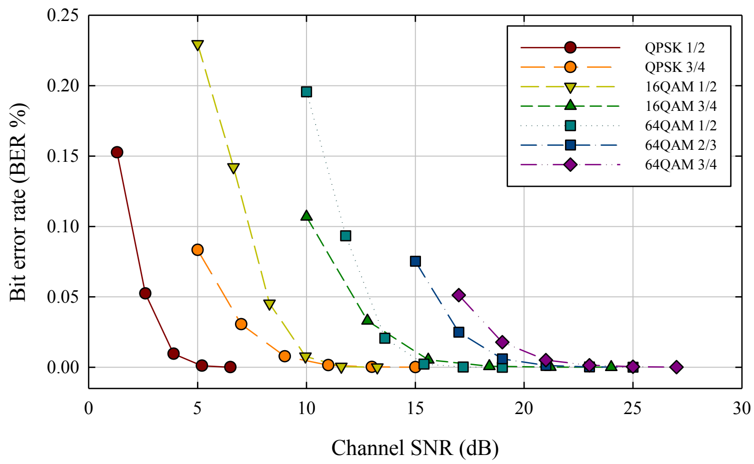


Figure 7.11: Channel SNR vs. BER for the vehicular scenario

The used WiMAX model offers a downlink capacity of 390 data slots (30 subcarriers x 13 time symbols) in a Time Division Duplex (TDD) frame every 5ms. The maximum channel transmission rate ranges from 3.744 Mbps for QPSK 1/2 (48 bits per data slot), to 16.848 Mbps for 64QAM 3/4 (216 bits per data slot). The data bits capacity of each MCS for the tested WiMAX model are given in Table 7.1.

Table 7.1: The data bits capacity of each MCS for the tested WiMAX model

Modulation	QPSK		16QAM		64QAM		
Code Rate	1/2	3/4	1/2	3/4	1/2	2/3	3/4
Data TX capacity (bits/data slot)	48	72	96	144	144	192	216

7.3.3 Results

To visualise the simulation results, Figure 7.12 and 7.13 shows the quality achieved over different levels of SNR for each MCS. It can be noted from these two Figures that the "Interview" video scored a higher maximum performance (QoE) than the "GT Fly" video since it was allocated double the bitrate. This, however, was at the expense of extra power transmission requirements, which are signified by the increase in the SNR. The data collected in Figure 7.12 and 7.13 is used to infer the SNR thresholds suggested at QoE=0.8 (MOS=4), in order to compare it with those thresholds based on BER.

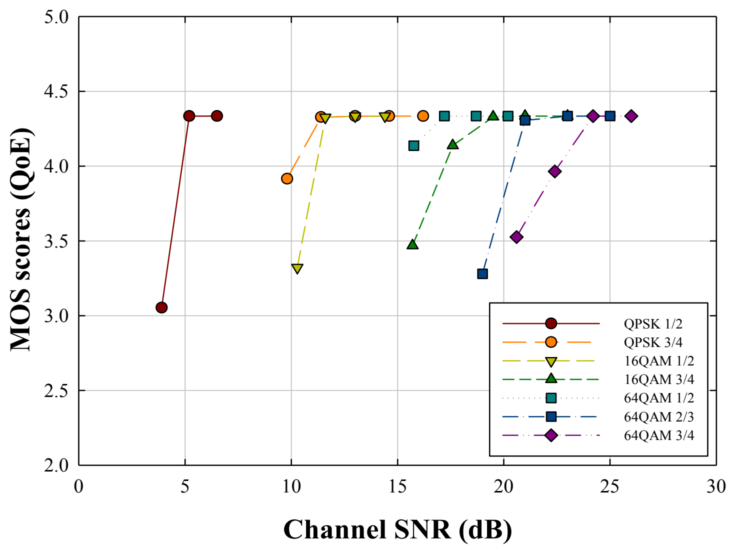


Figure 7.12: MOS vs. SNR, Interview video

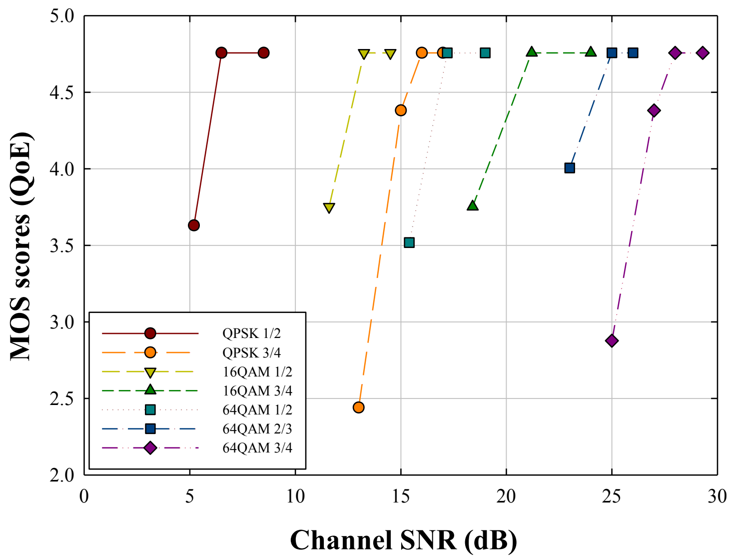


Figure 7.13: MOS vs. SNR, GT fly video

Figures 7.14 and 7.15 illustrate the relative performances of the BER-based method and the QoE-based method for assigning the operational MCS regions. It is clear that significant efficiencies in bandwidth and power requirements can be achieved by considering QoE as the basis for MCS selection (through AMC), rather than using the typical BER-based method. The proposed QoE-based selection for AMC bands not only is power-efficient, but is also bandwidth-efficient since it keeps a MS allocated to a higher-order band for longer than it used to be (with the BER-based method). This conclusion becomes evident when comparing the receiver's SNR for BER$= 10^{-6}$ and for QoE $= 0.8$. Thus, if QoE is considered, the lower value of the SNR implies that the transition to a higher-order MCS can occur earlier. This early transition to a higher AMC band means that more data bits per slot can be carried at the same bandwidth.

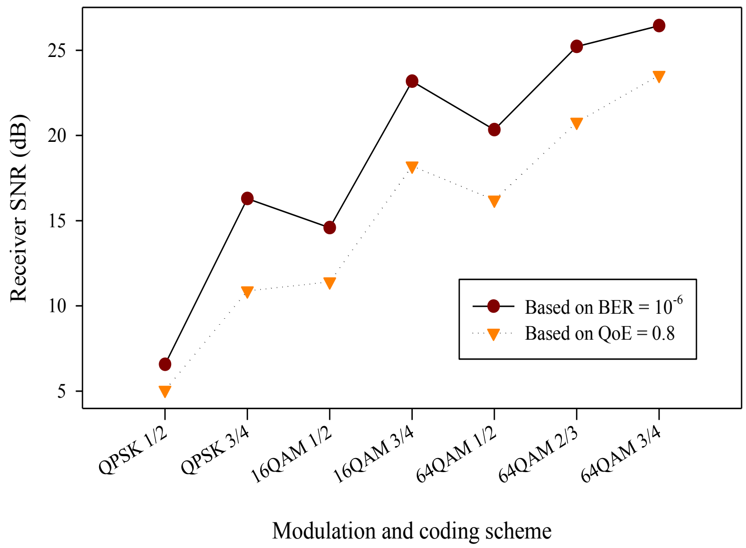


Figure 7.14: MCS vs. SNR, Interview video

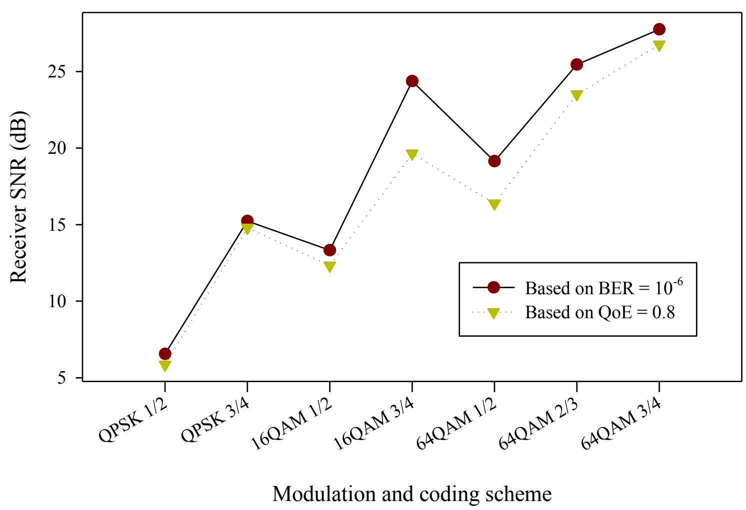


Figure 7.15: MCS vs. SNR, GT fly video

Table 7.2 summarizes the percentage of bandwidth increase and power savings that can be made as a result of such early transition. For example, 50% additional bandwidth is made available to the consumer at a lower SNR for a transition code from 16-QAM 1/2 to 16-QAM 3/4. With regard to power efficiency, as shown in Figures 7.14 and 7.15, the lower SNR value shows that the consumer can operate on the same MCS at a lower power level, which is controlled by the transmitting base station. For instance, in case of the GT Fly clip, the operating SNR threshold for 64QAM 1/2 based on QoE requires 2.75 dB less power than the BER-based method. In contrast, for Interview clip, the QoE-based method requires 4.12 dB less power than the BER based method. The percentage of power that could be saved for each MCS is shown in Table 7.2.

Table 7.2: summary of the percentage of bandwidth increase and power saving

Modulation	Code Rate	Percentage of additional data bits/slot to carry	Percentage of Power Saved	
			Interview	GT Fly
QPSK	1/2	———	22.87%	10.53%
	3/4	50%	33.13%	2.82%
16QAM	1/2	33.33%	21.66%	7.43%
	3/4	50%	21.40%	19.37%
64QAM	1/2	0%	20.26%	14.37%
	2/3	33.33%	17.57%	7.58%
	3/4	12.5%	10.97%	3.57%

7.4 Chapter Summary

This chapter presented two applications of the FIS-A model that is developed in Chapter 6. The first application was a QoE-enabled transport optimisation scheme for real-time SVC video delivery. The main objective of this application is to give an applicable example that allows the real-world scenario to be considered. It took advantage of the SVC Video partitions that organize video into layers of different importance, thereby facilitating the rate adaptation of video streams. The results showed that the proposed scheme enhanced video streaming performance and minimized the side effects of congestion on user perceived quality.

The second application was a QoE-enabled resource utilisation scheme for mobile video delivery. The receiver's transition between AMC regions was identified, based on the video QoE rather than QoS parameters. The video QoE was estimated in an online scenario using the FIS-A model. The results confirmed that considering the video QoE as the basis for modulation scheme selection in AMC can be generally advantageous with respect to power and bandwidth efficiency.

These applications showed how QoE is used to optimise video delivery and utilize existing network resources, according to the user's QoS requirements. QoE/QoS correlation is an indicator for network management and planning processes which allows the avoidance of resource over-provisioning. Furthermore, the results of the proposed applications provide further evidence that FIS-A performed well in terms of QoS/QoE correlation.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

This thesis explored the interesting, but complex, concepts of QoS/QoE correlation for video QoE prediction with respect to different video resolutions and content types. The research was broken down into three main sub-goals or challenges: (1) Investigate the impacts of the AQoS and NQoS related parameters that affect video QoE over wireless networks for different types of video content and resolutions, (2) Develop hybrid non-reference video quality prediction models using FIS, (3) Develop QoE-enabled optimisation applications for video delivery as applicability examples for the developed video quality prediction model. The contributions of this thesis are summarized in the following subsections.

8.1.1 Critical Review of Existing QoE Prediction models

A number of existing QoE/QoS correlation models for the prediction of video quality were critically reviewed in order to obtain a "broader picture" of this area of research. The survey article [175] was published for the research community and has been cited in several papers.

8.1.2 Study the Impact of QoS Parameter on Video QoE

We have provided a deeper insight into the impacts and a quantification of the effects of different QoS parameters associated with different layers of the OSI model. These parameters either directly or indirectly affected the video QoE. This work was divided into two studies that were presented in Chapter 4. The first study investigated the impact of QoS parameters on the video QoE by means of a cross-layer simulation of the transmission of SD, HD and 3D videos, encoded at a VBR mode. It was observed from the results that, for both 2D and 3D videos, the resultant for lower motion videos was much higher than for high motion complex videos. Moreover, due to the complex nature of true 3D video perception, the PLR and MBL caused more noticeable distortions to 3D video than to 2D video. In the case of VBR video streaming, packet loss was likely to have less of an impact upon the quality of HD video than upon lower resolution video. Furthermore, the bitrate rose with increasing QP which, in turn, restricted the number of streams or simultaneous users that could share a wireless link.

The second study investigated the effects of QoS parameters, packet structure, and compression efficiency on SD, HD, and 4kUHD H.264/H.265 coded videos that were both transmitted and decoded simultaneously in real-time. These videos were coded at a CBR mode. It was observed from the results that SD videos were more tolerant to packet loss than videos of higher resolution. This can be attributed to the fact that, since the bitrate is constant for both SD and HD, the bits-per-pixel value required to deliver equivalent quality drops as resolutions increase. Moreover, the target bitrate provided the key to determining the effects of QoS parameters. There was a trade-off between the compression ratio and the sensitivity of the QoE to the QoS parameters. The results suggested that, in general, if there was packet loss, the H.264 video codec outperformed the H.265 video codec. However, if there was no packet loss, then the H.264 codec performed better.

8.1.3 Hybrid Non-reference Video QoE Prediction Models

This thesis has contributed to the development of new hybrid non-reference models for predicting video quality using fuzzy logic inference systems (FIS) as a learning-based technique. Our proposed prediction models were based on a combination of QoS parameters related to the encoder, access network and content types. The learning dataset developed in Chapter 4 for the correlation of QoS parameters with the measured QoE from objective and subjective tests was used to build the two proposed prediction models, thereby avoiding time consuming subjective tests.

The first proposed model was based on a predetermined rule-based method (FIS-PRB). While the second proposed model was based on an adaptive FIS rule-based method (FIS-A). Both models were validated through the correlation of predicted QoE and measured QoE on both testing and external unseen datasets. The validation results showed a high correlation between the measured QoE and the predicted QoE scores. The FIS-A model was also evaluated on a real test-bed for the QoE prediction of real-time wireless video streaming. The experimental results showed a reasonable fit, even in the real-time scenario. Overall, the FIS-A model outperformed the FIS-PRB model, especially on the external dataset. The proposed models discussed in Chapters 5 and 6 provide a proof of concept hybrid non-reference video QoE prediction technique. Both FIS models showed a consistent mapping between the QoS parameters and the QoE. A second important conclusion that can be drawn from this work is that the selection of appropriate QoS parameters has a significant impact on the prediction of video quality. The work has been contributed to the research community in the following publications [22, 24–28].

8.1.4 Two QoE-enabled Applications for Video Delivery

The usefulness of the developed non-reference FIS-A model was demonstrated in Chapter 7 by its application to two important areas. First, a QoE-enabled transport optimisation scheme for real-time SVC video delivery was developed. The proposed application provides a crucial design choice that will optimise the video traffic by mapping video quality degradations (that are caused by the network) to the QoE without penetrating the video packets. The results showed that the proposed scheme enhanced video streaming performance and minimized the side effects of congestion on the user perceived quality.

The second application was a QoE-enabled efficient resource utilisation scheme for mobile video delivery. This work relied on the fact that the use of QoE as the basis for modulation scheme selection in AMC can be generally advantageous with respect to power and bandwidth efficiency, when compared to techniques that are solely based on the BER parameter. The results reflected a significant improvement in performance when the receiver's transition between AMC regions was identified using the video QoE. This efficiency reached up to 33% less power and 50% more bandwidth. These applications showed how the QoE/QoS correlation be an indicator for network management and planning processes that helps network manages to avoid resource over-provisioning. This work has been presented to the research community in the following publications [29–32].

8.2 Future Work

This thesis has made several advances in the prediction of QoE, based on QoS parameters, and has provided a framework for monitoring and estimating the video QoE. At the same time, it has unlocked the following future directions for research:

- It would be interesting to work towards the development of a more generic video

QoE prediction model. This would include the consideration and, potentially, the incorporation of additional QoS parameters into the developed models. Special care will be taken to measure the computational overhead as we aim to provide real-time estimation of video QoE over high speed data links.

- The models developed in the thesis for predicting video quality non-intrusively can be extended to consider voice. This will require the consideration of new parameters that impact on audiovisual quality. For example, parameters such as the voice bit rate (BR) may be useful for the prediction of voice quality over wireless access networks.

- Hybrid approaches have attracted considerable attention in the Computational Intelligence community. One of the most popular approaches is the hybridization between fuzzy logic and GAs leading to genetic fuzzy systems (GFSs). As we mentioned, the more accurately defined the membership functions the higher the prediction ability of the system. Genetic algorithms have demonstrated to be a robust and very powerful tool to perform tasks such as the generation and tuning of membership functions.

- In the Chapter 7 (appliacation 1), we only considered one policing action which is dropping video layers. However, this function can be improved to incorporate additional policing actions such as routing and prioritisation using Differentiated services (DiffServ) mechanism [176]. We will also investigate the use of the feedback message from FIS systems to implement corrective actions on network and application levels in order to return the QoE to satisfactory levels.

- In this thesis, three types of video contents were considered. They broadly covered video sequences from slow moving (head and shoulder) to fast moving (sports type). However, cartoon clips and movies were not considered. The spatio-temporal features of cartoon and movie clips may have an impact on end-to-end quality.

- It would be interesting to use the video quality prediction model in the area of mobile Ad hoc network (MANET), considering a new parameters like mobility and scalability.

- The emergence of cloud computing offers a new approach to media content delivery. Using the service concept in cloud computing, a specific media streaming service can be delivered to a user at a price charged per use, similar to what is obtainable in a common utility service. This approach makes it possible for the provider to distinguish between classes of users based on the price and service level agreed between the parties. This user-specific classification can be achieved by, for example, implementing a network aware traffic management that guarantees a certain QoE to a priority user. for future work, it would be interesting to apply the QoE-enabled transport optimisation scheme (that is presented in Chapter 7) on a cloud computing environment. Moreover, modelling the dynamic QoE parameters and including them as part of the service, strict SLA can easily be delivered to each end-user.

Bibliography

[1] I. Cisco, "Cisco visual networking index: Forecast and methodology, 2013–2018," *CISCO White paper*, 2014.

[2] G. Tsihrintzis and L. Jain, *Multimedia Services in Intelligent Environments: Advanced Tools and Methodologies*, vol. 120. Springer Science & Business Media, 2008.

[3] L. Bryan and E. Bryan, "Programmable controllers; fuzzy logic and plc," *Industrial Text and Videocompany*, 1988.

[4] G. Haßlinger and O. Hohlfeld, "The gilbert-elliott model for packet loss in real time services on the internet," in *Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 2008 14th GI/ITG Conference-*, pp. 1–15, VDE, 2008.

[5] D. Hands, O. V. Barriac, and F. Telecom, "Standardization activities in the itu for a qoe assessment of iptv," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 78–84, 2008.

[6] A. R. Gray and S. G. MacDonell, "A comparison of techniques for developing predictive models of software metrics," *Information and software technology*, vol. 39, no. 6, pp. 425–437, 1997.

[7] T. Zinner, O. Hohlfeld, O. Abboud, and T. Hoßfeld, "Impact of frame rate and resolution on objective qoe metrics," in *Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on*, pp. 29–34, IEEE, 2010.

[8] M. Mehendale, S. Das, M. Sharma, M. Mody, R. Reddy, J. Meehan, H. Tamama, B. Carlson, and M. Polley, "A true multistandard, programmable, low-power, full hd video-codec engine for smartphone soc," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pp. 226–228, IEEE, 2012.

[9] C. T. Hewage, S. T. Worrall, S. Dogan, S. Villette, and A. M. Kondoz, "Quality evaluation of color plus depth map-based stereoscopic video," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 304–318, 2009.

[10] C. T. Hewage and M. G. Martini, "Reduced-reference quality assessment for 3d video compression and transmission," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1185–1193, 2011.

[11] A. Takahashi, D. Hands, and V. Barriac, "Standardization activities in the itu for a qoe assessment of iptv," *Communications Magazine, IEEE*, vol. 46, pp. 78–84, February 2008.

[12] K. Brunnström, S. A. Beker, K. De Moor, A. Dooms, S. Egger, M.-N. Garcia, T. Hossfeld, S. Jumisko-Pyykkö, C. Keimel, C. Larabi, *et al.*, "Qualinet white paper on definitions of quality of experience," 2013.

[13] I. Nokia, "Quality of experience (qoe) of mobile services: can it be measured and improved?," *Nokia White paper*, 2004.

[14] S. Patrick, J. Singer, B. Corrie, S. Noe, K. El Khatib, B. Emond, T. Zimmerman, and S. Marsh, "A qoe sensitive architecture for advanced collaborative environments," in *Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004. First International Conference on*, pp. 319–322, IEEE, 2004.

[15] D. Soldani, M. Li, and R. Cuny, *QoS and QoE management in UMTS cellular systems*. John Wiley & Sons, 2007.

[16] M. Siller, *An agent-based platform to map quality of service to experience in active and conventional networks*. PhD thesis, University of Essex, 2006.

[17] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE transactions on Image Processing*, vol. 19, no. 6, pp. 1427–1441, 2010.

[18] U. Reiter, "Perceived quality in consumer electronics-from quality of service to quality of experience," in *2009. ISCE'09. IEEE 13th International Symposium on Consumer Electronic*, pp. 958–961, IEEE, 2009.

[19] T. Rahrer, R. Fiandra, S. Wright, *et al.*, "Triple-play services quality of experience (qoe) requirements," in *DSL-Forum. Technical Report TR-126*, 2006.

[20] D. Kanellopoulos, *"Emerging Research on Networked Multimedia Communication Systems"*. IGI-globale, 2015.

[21] M. Alreshoodi, L. AlJobouri, J. Woods, and M. Fleury, "Packet loss impact upon sd, hd, and 3d video streaming according to content type," *IEEE Transactions on Broadcasting*, 2015.

[22] M. Alreshoodi, E. Danish, J. Woods, A. Fernando, and C. De Alwis, ""prediction of perceptual quality for mobile video using fuzzy inference systems"," *IEEE Transactions on Consumer Electronics*, Nov 2015.

[23] A. Adeyemi-Ejeye, M. Alreshoodi, S. Walker, J. Woods, and M. Ghanbari, "Evaluation of encoded sd, hd, and 4kuhd video streaming: A packet layer perspective," *Circuits and Systems for Video Technology, IEEE Transactions on*, 2015.

[24] M. Alreshoodi, A. Adeyemi-Ejeye, J. Woods, and S. Walker, ""fuzzy logic inference system-based hybrid quality prediction model for wireless 4kuhd h.265-coded video streaming"," *Networks, IET*, vol. 4, no. 6, pp. 296–303, 2015.

[25] E. Danish, A. Fernando, M. Alreshoodi, and J. Woods, “A hybrid prediction model for video quality by qos/qoe mapping in wireless streaming,” in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 1723–1728, June 2015.

[26] M. Alreshoodi, E. Danish, J. Woods, A. Fernando, and C. De Alwis, “”prediction of perceptual quality for mobile 3d video using fuzzy inference systems”,” in *2015 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 130–131, Jan 2015.

[27] M. Alreshoodi and J. Woods, “”an empirical study based on a fuzzy logic system to assess the qos/qoe correlation for layered video streaming”,” in *2013 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pp. 180–184, July 2013.

[28] M. Alreshoodi and J. Woods, “”qoe prediction model based on fuzzy logic system for different video contents”,” in *2013 European Modelling Symposium (EMS)*, pp. 635–639, Nov 2013.

[29] M. Alreshoodi, J. Woods, and I. Musa, “Qoe-enabled transport optimisation scheme for real-time svc video delivery,” in *Communication Systems, Networks Digital Signal Processing (CSNDSP), 2014 9th International Symposium on*, pp. 865–868, July 2014.

[30] M. Alreshoodi, J. Woods, and I. Musa, “”optimising the delivery of scalable h.264 video stream by qos/qoe correlation”,” in *2015 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 253–254, Jan 2015.

[31] M. Alreshoodi, E. Danish, J. Woods, and A. Fernando, “”resource utilisation scheme based on online qoe estimation of mobile h.264 video streaming”,” in *2015 Ninth International Conference on Next Generation Mobile Apps, Services and Technologies (NGMAST)*, Sep 2015.

[32] M. Alreshoodi, E. Danish, J. Woods, and A. Fernando, “”qoe-enabled efficient resource utilisation for h.264 video streaming over wimax”,” in *2016 IEEE International Conference on Consumer Electronics (ICCE)*, Jan 2016.

[33] K. Raymond, “Reference model of open distributed processing (rm-odp): Introduction,” in *Open Distributed Processing*, pp. 3–14, Springer, 1995.

[34] ISO, “Quality of service framework, iso/iec jtc1/sc21/wg1 n9680,” 1995.

[35] A. Vogel, B. Kerhervé, G. v. Bochmann, and J. Gecsei, “Distributed multimedia applications and quality of service: a survey,” in *Proceedings of the 1994 conference of the Centre for Advanced Studies on Collaborative research*, p. 71, IBM Press, 1994.

[36] E. Gallo, M. Siller, and J. Woods, “An ontology for the quality of experience framework,” in *2007. ISIC. IEEE International Conference on Systems, Man and Cybernetics*, pp. 1540–1544, IEEE, 2007.

[37] B. Bauer and A. S. Patrick, “A human factors extension to the seven-layer osi reference model,” *Retrieved January*, vol. 6, p. 2004, 2004.

[38] D. Lopez, F. Gonzalez, L. Bellido, and A. Alonso, “Adaptive multimedia streaming over ip based on customer oriented metrics,” in *2006 International Symposium on Computer Networks*, pp. 185–191, IEEE, 2006.

[39] M. Fiedler, T. Hossfeld, and P. Tran-Gia, “A generic quantitative relationship between quality of experience and quality of service,” *Network, IEEE*, vol. 24, no. 2, pp. 36–41, 2010.

[40] F. Kozamernik, “Media streaming over the internet–an overview of delivery technologies,” *EBU Technical Review*, vol. 10, 2002.

[41] J. Pokhrel, B. Wehbi, N. Kushik, N. Yevtushenko, and A. Cavalli, “Multimedia quality of experience,” *Emerging Research on Networked Multimedia Communication Systems*, p. 250, 2015.

[42] J. G. Apostolopoulos, W.-t. Tan, and S. J. Wee, “Video streaming: Concepts, algorithms, and systems,” *HP Laboratories, report HPL-2002-260*, 2002.

[43] H. Zheng and J. Boyce, “An improved udp protocol for video transmission over internet-to-wireless networks,” *IEEE Transactions on Multimedia*, vol. 3, no. 3, pp. 356–365, 2001.

[44] R. RFC3550-IETF, “A transport protocol for real-time applications,” *Internet*

Engineering Task Force, 2003.

[45] I. E. Richardson, *H. 264 and MPEG-4 video compression: video coding for next-generation multimedia.* John Wiley & Sons, 2004.

[46] J. Ohm, G. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards 2014;including high efficiency video coding (hevc)," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1669–1684, Dec.

[47] ITU-T, "Advanced video coding for generic audiovisual services, recommendation itu-t h.264," 2013.

[48] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, pp. 1649–1668, Dec 2012.

[49] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the h. 264/avc video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.

[50] M. van der Schaar and P. Chou, "Multimedia networking and communication: Principles and challenges," 2011.

[51] I. Rec, "Bt. 500-13, methodology for the subjective assessment of the quality of television pictures," *International Telecommunication Union*, 2012.

[52] P. ITU-T RECOMMENDATION, "Subjective video quality assessment methods for multimedia applications," 2008.

[53] I. R. Assembly, *Methodology for the subjective assessment of the quality of television pictures.* International Telecommunication Union, 2003.

[54] S. Chikkerur, V. Sundaram, M. Reisslein, and L. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE Transactions on Broadcasting*, vol. 57, pp. 165–182, June 2011.

[55] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp. 312–322, 2004.

[56] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "The ssim index for image quality assessment," *MATLAB implementation available online from: http://www. cns. nyu. edu/˜ lcv/ssim*, 2003.

[57] I. Recommendation, "objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," *International Telecommunication Union*, 2004.

[58] I.-T. I. ITU, "Quality of experience requirements for iptv services," 2008.

[59] I. Paudel, J. Pokhrel, B. Wehbi, A. Cavalli, and B. Jouaber, "Estimation of video qoe from mac parameters in wireless network: A random neural net-

work approach," in *2014 14th International Symposium on Communications and Information Technologies (ISCIT)*, pp. 51–55, IEEE, 2014.

[60] J. Pokhrel, B. Wehbi, A. Morais, A. Cavalli, and E. Allilaire, "Estimation of qoe of video traffic using a fuzzy expert system," in *Consumer Communications and Networking Conference (CCNC), 2013 IEEE*, pp. 224–229, IEEE, 2013.

[61] F. You, W. Zhang, and J. Xiao, "Packet loss pattern and parametric video quality model for iptv," in *Eighth IEEE/ACIS International Conference on Computer and Information Science (ICIS) 2009*, pp. 824–828, June 2009.

[62] J. Nightingale, Q. Wang, C. Grecos, and S. Goma, "The impact of network impairment on quality of experience (qoe) in h. 265/hevc video streaming," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 242–250, 2014.

[63] J. Nightingale, Q. Wang, C. Grecos, and S. Goma, "Modeling qoe for streamed h. 265/hevc content under adverse network conditions," 2013.

[64] H. Malekmohamadi, W. Fernando, E. Danish, and A. Kondoz, "Subjective quality estimation based on neural networks for stereoscopic videos," in *2014 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 107–108, IEEE, 2014.

[65] L. Anegekuh, L. Sun, and E. Ifeachor, "Encoding and video content based hevc video quality prediction," *Multimedia Tools and Applications*, pp. 1–24, 2013.

[66] X. Lin, H. Ma, L. Luo, and Y. Chen, "No-reference video quality assessment in the compressed domain," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 505–512, 2012.

[67] Y.-F. Ou, Z. Ma, T. Liu, and Y. Wang, "Perceptual quality assessment of video considering both frame rate and quantization artifacts," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 3, pp. 286–298, 2011.

[68] S.-O. Lee, K.-S. Jung, and D.-G. Sim, "Real-time objective quality assessment based on coding parameters extracted from h. 264/avc bitstream," *Consumer Electronics, IEEE Transactions on*, vol. 56, no. 2, pp. 1071–1078, 2010.

[69] A. Eden, "No-reference estimation of the coding psnr for h. 264-coded sequences," *Consumer Electronics, IEEE Transactions on*, vol. 53, no. 2, pp. 667–674, 2007.

[70] F. Agboma and A. Liotta, "Qoe-aware qos management," in *Proceedings of the 6th international conference on advances in mobile computing and multimedia*, pp. 111–116, ACM, 2008.

[71] M. Elkotob, D. Grandlund, K. Andersson, and C. Ahlund, "Multimedia qoe optimized management using prediction and statistical learning," in *2010 IEEE 35th Conference on Local Computer Networks (LCN)*, pp. 324–327, IEEE, 2010.

[72] H. J. Kim and S. G. Choi, "A study on a qos/qoe correlation model for qoe evaluation on iptv service," in *2010 The 12th International Conference on Advanced*

Communication Technology (ICACT), vol. 2, pp. 1377–1382, IEEE, 2010.

[73] W. R. Klecka, *Discriminant analysis.* No. 19, Sage, 1980.

[74] A. Khan, L. Sun, and E. Ifeachor, "Qoe prediction model and its application in video quality adaptation over umts networks," *IEEE Transactions on Multimedia*, vol. 14, no. 2, pp. 431–442, 2012.

[75] A. Khan, L. Sun, E. Ifeachor, J. O. Fajardo, and F. Liberal, "Video quality prediction model for h. 264 video over umts networks and their application in mobile video streaming," in *Communications (ICC), 2010 IEEE International Conference on*, pp. 1–5, IEEE, 2010.

[76] A. Raake, M.-N. Garcia, S. Möller, J. Berger, F. Kling, P. List, J. Johann, and C. Heidemann, "Tv-model: Parameter-based prediction of iptv quality," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 1149–1152, IEEE, 2008.

[77] J. Joskowicz, R. Sotelo, and J. C. Lopez Ardao, "Towards a general parametric model for perceptual video quality estimation," *Broadcasting, IEEE Transactions on*, vol. 59, no. 4, pp. 569–579, 2013.

[78] S. Aroussi and A. Mellouk, "Survey on machine learning-based qoe-qos correlation models," in *Computing, Management and Telecommunications (ComManTel), 2014 International Conference on*, pp. 200–204, April 2014.

[79] Y. Kang, H. Chen, and L. Xie, "An artificial-neural-network-based qoe estimation model for video streaming over wireless networks," in *Communications in China (ICCC), 2013 IEEE/CIC International Conference on*, pp. 264–269, IEEE, 2013.

[80] E. Aguiar, A. Riker, M. Mu, S. Zeadally, E. Cerqueira, and A. Abelem, "Real-time qoe prediction for multimedia applications in wireless mesh networks," in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, pp. 592–596, IEEE, 2012.

[81] W. Cherif, A. Ksentini, D. Négru, and M. Sidibé, "A_psqa: Efficient real-time video streaming qoe tool in a future media internet context," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pp. 1–6, IEEE, 2011.

[82] V. A. Machado, N. Carlos, R. S. O. Silva, A. M. Melo, M. Silva, C. R. Francês, J. C. Costa, N. L. Vijaykumar, and C. M. Hirata, "A new proposal to provide estimation of qos and qoe over wimax networks," 2011.

[83] N. Staelens, V. Glenn, K. Crombecq, N. Vercammen, J. De Cock, B. Vermeulen, V. Rik, T. Dhaene, and P. Demeester, "No-reference bitstream-based visual quality impairment detection for high definition h. 264/avc encoded video sequences," *Broadcasting, IEEE Transactions on*, vol. 58, no. 2, pp. 187–199, 2012.

[84] J. R. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1, no. 1,

pp. 81–106, 1986.

[85] V. Menkovski, A. Oredope, A. Liotta, and C. Sánchez, "Predicting quality of experience in multimedia streaming," in *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*, pp. 52–59, ACM, 2009.

[86] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[87] M. S. Mushtaq, B. Augustin, and A. Mellouk, "Empirical study based on machine learning approach to assess the qos/qoe correlation," in *Networks and Optical Communications (NOC), 2012 17th European Conference on*, pp. 1–7, IEEE, 2012.

[88] C. N. Pitas, D. E. Charilas, A. D. Panagopoulos, and P. Constantinou, "Adaptive neuro-fuzzy inference models for speech and video quality prediction in real-world mobile communication networks," *Wireless Communications, IEEE*, vol. 20, no. 3, pp. 80–88, 2013.

[89] A. Khan, L. Sun, and E. Ifeachor, "An anfis-based hybrid video quality prediction model for video streaming over wireless networks," in *The Second International Conference on Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST'08*, pp. 357–362, IEEE, 2008.

[90] J.-S. R. Jang, "Anfis: adaptive-network-based fuzzy inference system," *Sys-*

tems, Man and Cybernetics, IEEE Transactions on, vol. 23, no. 3, pp. 665–685, 1993.

[91] M. Siller and J. Woods, "Improving quality of experience for multimedia services by qos arbitration on a qoe framework," in *in Proc. of the 13th Packed Video Workshop 2003*, Citeseer, 2003.

[92] T. Wang, A. Pervez, and H. Zou, "Vqm-based qos/qoe mapping for streaming video," in *Broadband Network and Multimedia Technology (IC-BNMT), 2010 3rd IEEE International Conference on*, pp. 807–812, IEEE, 2010.

[93] H. Du, C. Guo, Y. Liu, and Y. Liu, "Research on relationship between qoe and qos based on bp neural network," in *Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on*, pp. 312–315, IEEE, 2009.

[94] B. Han, X. Zhang, Y. Qi, Y. Gao, and D. Yang, "Qoe model based optimization for streaming media service considering equipment and environment factors," *Wireless Personal Communications*, vol. 66, no. 3, pp. 595–612, 2012.

[95] K. Laghari, I. Khan, and N. Crespi, "Quantitative and qualitative assessment of qoe for multimedia services in wireless environment," in *Proceedings of the 4th Workshop on Mobile Video*, MoVid '12, pp. 7–12, ACM, 2012.

[96] R. L. P. Ramos, J. Salmern and F. Vidal, ""estimating perceived video quality from objective parameters in video over ip services"," 2012.

[97] H. Koumaras, C. Lin, C. Shieh, and A. Kourtis, “A framework for end-to-end video quality prediction of mpeg video,” *Journal of Visual Communication and Image Representation*, vol. 21, no. 2, pp. 139–154, 2010.

[98] P. Frank and J. Incera, “A neural network based test bed for evaluating the quality of video streams in ip networks,” in *Electronics, Robotics and Automotive Mechanics Conference, 2006*, vol. 1, pp. 178–183, IEEE, 2006.

[99] R. Mok, E. Chan, and R. Chang, “Measuring the quality of experience of http video streaming,” in *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*, pp. 485–492, IEEE, 2011.

[100] T. Hoßfeld, M. Seufert, M. Hirth, T. Zinner, P. Tran-Gia, and R. Schatz, “Quantification of youtube qoe via crowdsourcing,” in *Multimedia (ISM), 2011 IEEE International Symposium on*, pp. 494–499, IEEE, 2011.

[101] J. Bih, “Paradigm shift-an introduction to fuzzy logic,” *IEEE Potentials*, vol. 25, no. 1, pp. 6–10, 2006.

[102] J. M. Mendel, “Fuzzy logic systems for engineering: a tutorial,” *Proceedings of the IEEE*, vol. 83, no. 3, pp. 345–377, 1995.

[103] J. Pokhrel, F. Lalanne, A. Cavalli, and W. Mallouli, “Qoe estimation for web service selection using a fuzzy-rough hybrid expert system,” in *Advanced Information Networking and Applications (AINA), 2014 IEEE 28th International Conference on*, pp. 629–634, IEEE, 2014.

[104] Y. Bai and D. Wang, "Fundamentals of fuzzy logic controlfuzzy sets, fuzzy rules and defuzzifications," in *Advanced Fuzzy Logic Technologies in Industrial Applications*, pp. 17–36, Springer, 2006.

[105] R. Schapire, "Machine learning algorithms for classification," *Princeton University*, vol. 10, 2006.

[106] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, and M. Etoh, "Cross-dimensional perceptual quality assessment for low bit-rate videos," *Multimedia, IEEE Transactions on*, vol. 10, no. 7, pp. 1316–1324, 2008.

[107] G. Cermak, "Subjective video quality as a function of bit rate frame rate, packet loss, and codec," in *Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on*, pp. 41–46, IEEE, 2009.

[108] Y. Liang, J. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: does burst-length matter?," in *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, vol. 5, pp. V–684–7 vol.5, April 2003.

[109] F. Boulos, B. Parrein, P. Le Callet, and D. Hands, "Perceptual effects of packet loss on h. 264/avc encoded videos," in *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics VPQM-09*, 2009.

[110] M. H. Pinson, S. Wolf, and G. Cermak, "Hdtv subjective quality of h. 264 vs. mpeg-2, with and without packet loss," *Broadcasting, IEEE Transactions on*,

vol. 56, no. 1, pp. 86–91, 2010.

[111] J. Shin and P. C. Cosman, "Classification of mpeg-2 transport stream packet loss visibility," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 910–913, IEEE, 2010.

[112] S. Argyropoulos, A. Raake, M.-N. Garcia, and P. List, "No-reference bit stream model for video quality assessment of h. 264/avc video based on packet loss visibility," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 1169–1172, IEEE, 2011.

[113] J. Nightingale, Q. Wang, and C. Grecos, "Hevstream: a framework for streaming and evaluation of high efficiency video coding (hevc) content in loss-prone networks," *IEEE Transactions on Consumer Electronics*, vol. 58, pp. 404–412, May 2012.

[114] J. Nightingale, Q. Wang, C. Grecos, and S. Goma, "Subjective evaluation of the effects of packet loss on hevc encoded video streams," in *IEEE Third International Conference on Consumer Electronics, Berlin (ICCE-Berlin), 2013*, pp. 358–359, IEEE, 2013.

[115] C. T. Hewage, M. G. Martini, M. Brandas, and D. De Silva, "A study on the perceived quality of 3d video subject to packet losses," in *Communications Workshops (ICC), 2013 IEEE International Conference on*, pp. 662–666, IEEE, 2013.

[116] M. M. Nasralla, C. T. Hewage, and M. G. Martini, "Subjective and objective evaluation and packet loss modeling for 3d video transmission over lte networks," in *Telecommunications and Multimedia (TEMU), 2014 International Conference on*, pp. 254–259, IEEE, 2014.

[117] M. M. Nasralla, O. Ognenoski, and M. G. Martini, "Bandwidth scalability and efficient 2d and 3d video transmission over lte networks," in *Communications Workshops (ICC), 2013 IEEE International Conference on*, pp. 617–621, IEEE, 2013.

[118] L. Al-Jobouri, M. Fleury, and M. Ghanbari, "Protecting h. 264/avc data-partitioned video streams over broadband wimax," *Advances in Multimedia*, vol. 2012, p. 10, 2012.

[119] R. A. Farrugia, *Multimedia Networking and Coding*. IGI Global, 2012.

[120] J. V. T. J. of ISO/IEC MPEG & ITU-T VCEG, "H.264/avc jm reference software.," 2011.

[121] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell system technical journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[122] M. Zorzi, R. Racd, and L. B. Milstein, "A markov model for block errors on fading channels," in *Personal, Indoor and Mobile Radio Communications, 1996. PIMRC'96., Seventh IEEE International Symposium on*, vol. 3, pp. 1074–1078, IEEE, 1996.

[123] S. Yasakethu, S. Worrall, D. De Silva, W. Fernando, and A. Kondoz, "A compound depth and image quality metric for measuring the effects of packet loss on 3d video," in *Digital Signal Processing (DSP), 2011 17th International Conference on*, pp. 1–7, IEEE, 2011.

[124] ANSI/IEEE, ""digital transport of one-way video signals-parameters for objective performance assessment", ansi standard atis-0100801.03.2003(r2013)," 2003.

[125] I.-R. B. Recommendation, "Subjective assessment of stereoscopic television pictures," 2000.

[126] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, pp. 137–148, 1969.

[127] M. Shahid, A. Rossholm, B. Lövström, and H.-J. Zepernick, "No-reference image and video quality assessment: a classification and review of recent approaches," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1–32, 2014.

[128] C. Hewage, M. Martini, M. Brandas, and D. De Silva, "A study on the perceived quality of 3d video subject to packet losses," in *Communications Workshops (ICC), 2013 IEEE International Conference on*, pp. 662–666, June 2013.

[129] G. W. Snecdecor and W. G. Cochran, *Statistical Methods.* Wiley, 1991.

[130] W. Krzanowski, *Principles of multivariate analysis.* Oxford University Press, 2000.

[131] Y. Wang, “Survey of objective video quality measurements,” 2006.

[132] A.Adeyemi-Ejeye, “Hevc mpeg ts definition in mpeg ts header,” 03 2014.

[133] A.Adeyemi-Ejeye, “Hevc mpeg ts: Allowing hevc video using ts stream type 0x24,” 03 2014.

[134] A.Adeyemi-Ejeye, “[patch] defined hevc for mpegts,” 03 2014.

[135] SoftPerfectResearch, “Softperfect connection emulator (sce),” 10 2014.

[136] F. Bellard, M. Niedermayer, *et al.*, “Ffmpeg,” *ht tp://ffmpeg. org*, 2012.

[137] W. Pedrycz, “Why triangular membership functions?,” *Fuzzy sets and Systems*, vol. 64, no. 1, pp. 21–30, 1994.

[138] B. S. de Lima and N. F. Ebecken, “A comparison of models for uncertainty analysis by the finite element method,” *Finite Elements in Analysis and Design*, vol. 34, no. 2, pp. 211–232, 2000.

[139] M. Anoop, K. B. Rao, and S. Gopalakrishnan, “Conversion of probabilistic information into fuzzy sets for engineering decision analysis,” *Computers & structures*, vol. 84, no. 3, pp. 141–155, 2006.

[140] E. H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller,” *International journal of man-machine studies*, vol. 7, no. 1, pp. 1–13, 1975.

[141] L.-X. Wang, “The wm method completed: a flexible fuzzy system approach to data mining,” *Fuzzy Systems, IEEE Transactions on*, vol. 11, no. 6, pp. 768–782, 2003.

[142] Mathworks, “Fuzzy inference systems: Tutorial (fuzzy logic toolbox).”

[143] A. Khan, “subjective mos,” 2010.

[144] E. Gelenbe, “Random neural networks with negative and positive signals and product form solution,” *Neural computation*, vol. 1, no. 4, pp. 502–510, 1989.

[145] L. Muscariello, “Qoe-rnn - quality of experience estimation using random neural networks,” 2011.

[146] N. Kushik, J. Pokhrel, N. Yevtushenko, A. Cavalli, and W. Mallouli, “Qoe prediction for multimedia services: Comparing fuzzy and logic network approaches,” *International Journal of Organizational and Collective Intelligence (IJOCI)*, vol. 4, no. 3, pp. 44–64, 2014.

[147] F. Fauzi and D. Yang, “Fuzzy logic for bandwidth allocator applies on ip multimedia traffic,” in *Information and Communication Technology (ICoICT), 2013 International Conference of*, pp. 416–421, IEEE, 2013.

[148] M. Fleury, E. Jammeh, R. Razavi, and M. Ghanbari, "Resource-aware fuzzy logic control of video streaming over ip and wireless networks," in *Pervasive Computing*, pp. 47–75, Springer, 2010.

[149] K. M. Passino, S. Yurkovich, and M. Reinfrank, *Fuzzy control*, vol. 42. Citeseer, 1998.

[150] L.-X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 22, no. 6, pp. 1414–1427, 1992.

[151] A. Khan, L. Sun, E. Ifeachor, J. O. Fajardo, and F. Liberal, "An anfis-based hybrid quality prediction model for h. 264 video over umts networks," in *Communications Quality and Reliability (CQR), 2010 IEEE International Workshop Technical Committee on*, pp. 1–6, IEEE, 2010.

[152] X. Liu, E. Chong, and N. Shroff, "Transmission scheduling for efficient wireless utilization," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 776–785, IEEE, 2001.

[153] V. Tsibonis, L. Georgiadis, and L. Tassiulas, "Exploiting wireless channel state information for throughput maximization," *Information Theory, IEEE Transactions on*, vol. 50, no. 11, pp. 2566–2582, 2004.

[154] S. Thakolsri, S. Khan, E. Steinbach, and W. Kellerer, "Qoe-driven cross-layer

optimization for high speed downlink packet access," *Journal of Communications*, vol. 4, no. 9, pp. 669–680, 2009.

[155] S. Luo, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over lte," *Communications Magazine, IEEE*, vol. 48, no. 2, pp. 102–109, 2010.

[156] E. Jammeh, I. Mkwawa, A. Khan, M. Goudarzi, L. Sun, and E. Ifeachor, "Quality of experience (qoe) driven adaptation scheme for voice/video over ip," *Telecommunication Systems*, vol. 49, no. 1, pp. 99–111, 2012.

[157] B. Fu, G. Kunzmann, M. Wetterwald, D. Corujo, and R. Costa, "Qoe-aware traffic management for mobile video delivery," in *Communications Workshops (ICC), 2013 IEEE International Conference on*, pp. 652–656, IEEE, 2013.

[158] M. Shehada, S. Thakolsri, Z. Despotovic, and W. Kellerer, "Qoe-based cross-layer optimization for video delivery in long term evolution mobile networks," in *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*, pp. 1–5, IEEE, 2011.

[159] N. Amram, B. Fu, G. Kunzmann, T. Melia, D. Munaretto, S. Randriamasy, B. Sayadi, J. Widmer, and M. Zorzi, "Qoe-based transport optimization for video delivery over next generation cellular networks," in *Computers and Communications (ISCC), 2011 IEEE Symposium on*, pp. 19–24, IEEE, 2011.

[160] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding

extension of the h. 264/avc standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1103–1120, 2007.

[161] H. ITU-T RECOMMENDATION, "Advanced video coding for generic audiovisual services," *ISO/IEC*, vol. 14496, 2003.

[162] T. Koistinen, "Protocol overview: Rtp and rtcp,"

[163] H. Schwarz, M. Wien, and J. Vieron, "Jsvm software manual," *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q*, vol. 6, 2006.

[164] A. Detti, G. Bianchi, C. Pisa, S. Proto, P. Loreti, W. Kellerer, S. Thakolsri, and J. Widmer, "Svef: an open-source experimental evaluation framework for h. 264 scalable video streaming," in *Computers and Communications, 2009. ISCC 2009. IEEE Symposium on*, pp. 36–41, IEEE, 2009.

[165] S. Lee, J. Koo, and K. Chung, "Content-aware rate control scheme to improve the energy efficiency for mobile iptv," in *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*, pp. 445–446, IEEE, 2010.

[166] C. Huang and C. Leung, "Bitqos-aware resource allocation for multi-user mixed-traffic ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 5, pp. 2067–2082, 2012.

[167] K. Dong-Hoi and J. Kyungkoo, "Dynamic bandwidth allocation scheme for

video streaming in wireless cellular networks," *IEICE transactions on communications*, vol. 89, no. 2, pp. 350–356, 2006.

[168] K. Piamrat, K. Singh, A. Ksentini, C. Viho, and J. Bonnin, "Qoe-aware scheduling for video-streaming in high speed downlink packet access," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, pp. 1–6, IEEE, 2010.

[169] D. Bethanabhotla, G. Caire, and J. Neely, "Joint transmission scheduling and congestion control for adaptive streaming in wireless device-to-device networks," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, pp. 1179–1183, IEEE, 2012.

[170] V. Joseph and G. De Veciana, "Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs," in *INFOCOM, 2012 Proceedings IEEE*, pp. 567–575, IEEE, 2012.

[171] R. Kuschnig, I. Kofler, and H. Hellwagner, "Evaluation of http-based request-response streams for internet video streaming," in *Proceedings of the second annual ACM conference on Multimedia systems*, pp. 245–256, ACM, 2011.

[172] A. Vishwanath, P. Dutta, M. Chetlu, P. Gupta, S. Kalyanaraman, and A. Ghosh, "Perspectives on quality of experience for video streaming over wimax," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13, no. 4, pp. 15–25, 2010.

[173] I. . W. Group *et al.*, "Ieee standard for local and metropolitan area networks. part 16: Air interface for fixed broadband wireless access systems," *IEEE Std*, vol. 802, pp. 16–2004, 2004.

[174] C. Liew, S. Worrall, and M. Goldshtein, "Wimax modelling," 2006.

[175] M. Alreshoodi and J. Woods, "Survey on qoe\qos correlation models for multimedia services," *International Journal of Distributed and Parallel Systems (IJDPS)*, vol. 17, no. 3, 2013.

[176] E. Davies, M. Carlson, W. Weiss, D. Black, S. Blake, and Z. Wang, "An architecture for differentiated services," 1998.

[177] W. Karner, O. Nemethova, P. Svoboda, and M. Rupp, "Link error analysis and modeling for video streaming cross-layer design in mobile communication networks," *ETRI journal*, vol. 29, no. 5, pp. 569–595, 2007.

Appendices

Appendix A

Gilbert-Elliot model

Packet loss traces are generated based on the Gilbert-Elliot model [121] (a two-state Markov chain model) with varying the PLR and the MBL. Gilbert-Elliott type models do not emulate the physical channel but do accurately model the application receivers experience of packet loss resulting from fast fading [122]. The used Gilbert-Elliot model is outlined in in Appendix D.

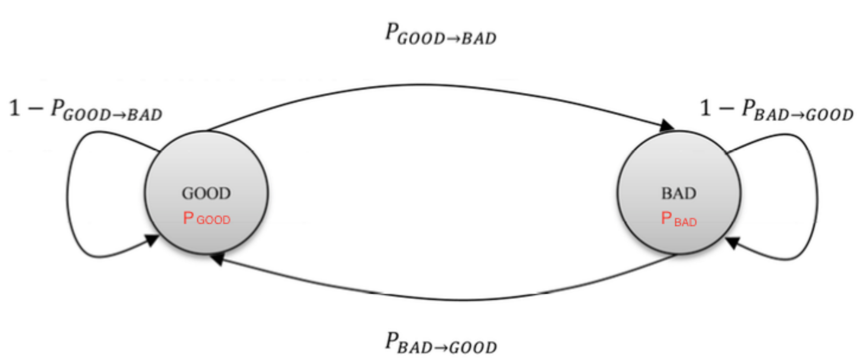


Figure A.1: Two-State Markov Chain Model [4]

According to this model, no packet losses occur in state GOOD and all packets are lost in state BAD. The probability of switching from state GOOD to BAD is denoted as p, while the probability of switching from state BAD to GOOD is denoted by q. The PLR is calculated by [121]:

$$PLR = p/(p + q) \tag{A.1}$$

The MBL value of 1 depicts random packet losses, whereas other MBL values represent increasingly bursty conditions, as shown in Table 4.2. The MBL is selected based on the mean error burst length measured in [177] for typical roaming scenarios from real-world wireless communication measurements. The MBL of the packet loss trace is calculated as follow [121]:

$$MBL = 1/q \tag{A.2}$$

In addition, each combination of QP, PLR, and MBL are simulated over ten times to ensure over 3000 video frames are considered per each simulation configuration. The average video quality of the received video frames are considered to present simulation results.

Appendix B

Statistical Performance Metrics

B.1 Pearson Correlation Coefficient (R)

PCC reveals whether or not a linear relationship exists between two variables. Furthermore, it indicates the direction and the strength of this relationship. In other words, it measures the similarity between two data measures. This is obtained by dividing the covariance of the two variables by the product of their standard deviations. Hence, for two given data series x_i and y_i having n observations each, PCC is calculated by:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{B.1}$$

x^- and y^- are the mean of x_i and y_i respectively, and $-1 \leq r \leq 1$. Accordingly, if $r = 1$, then the two data series are perfectly linear (similar). If $r = 0$, then they are not linear (not similar) at all. If $r = -1$, then they are perfectly inverse to each other. Any values in between indicate the degree of linear relationship. The squared PCC, denoted R^2, known as the coefficient of determination, indicates how much of the variance between the two variables is given by the linear fit.

B.2 Root Mean Squared Error (RMSE)

RMSE measures the individual differences (residuals) between two variables (data series). For example, values predicted and values measured in a model under study. RMSE aggregates these individual differences into a single measure. Hence, for two given data series x_i and y_i having n observations each, RMSE is calculated by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{B.2}$$

The unit of RMSE is that of the data series. As a second performance measure in the QoE prediction model in this thesis, RMSE was used to measure the residuals between the predicted and the measured QoE. Hence, the unit of RMSE measured is a QoE unit on the MOS scale.

Appendix C

Correlation of QoE datasets for MUSIC and BMX videos

In order to visualise the correlation between the three datasets (2D subjective, 2D objective, and 3D objective), Figures C.1 C.2 portrait a comparison of 21 test conditions for Music and BMX HD videos with the scored MOS in each of the three datasets constructed.

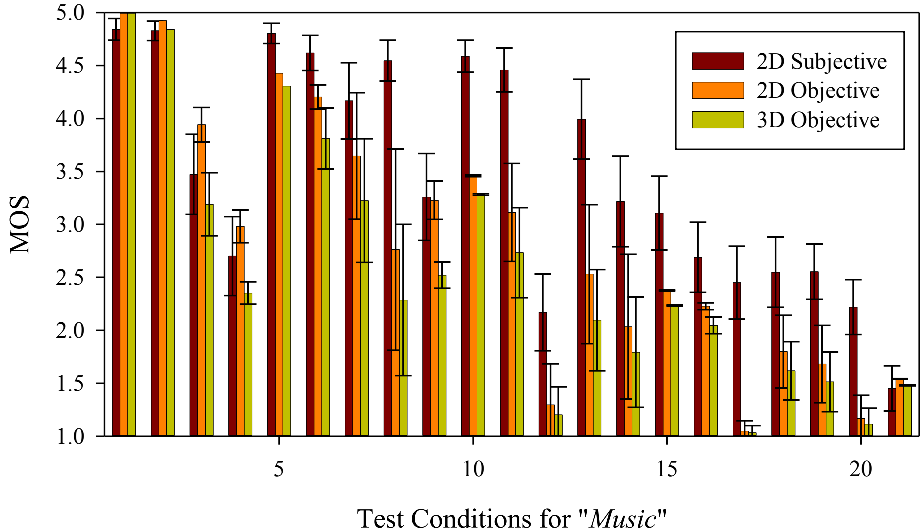


Figure C.1: Comparison of the three datasets for Poker HD video sequence

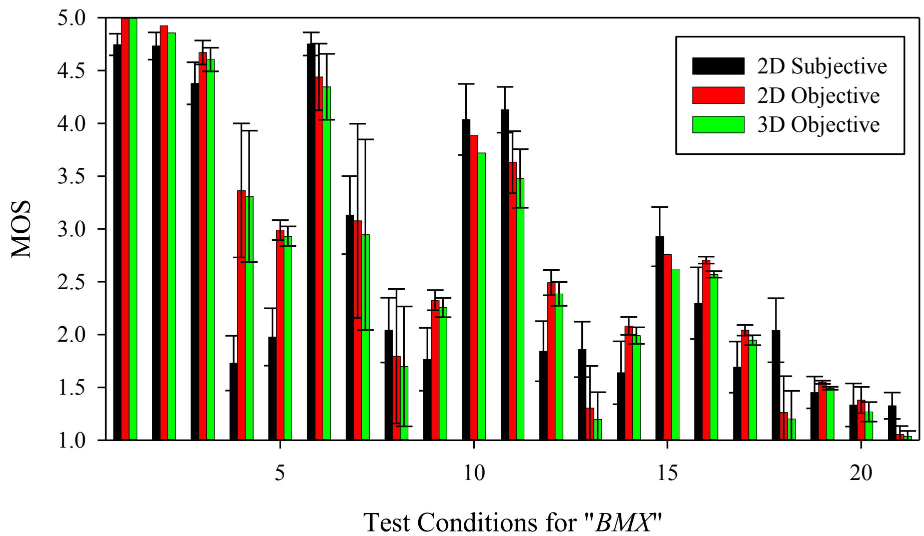


Figure C.2: Comparison of the three datasets for Poker HD video sequence

Appendix D

3D video frame size

Table D.1: 3D Input P-frame sizes by QP

3D	QP	Average size (bytes)		Average no. of packets		Approx. bitrate (Mbps)
		C	D	C	D	
BMX SD	16	299950	32997	200	22	66.7
	24	78639	11174	52	7	18.0
	32	24719	5507	16	4	6.0
BMX HD	16	851549	62397	568	42	182.8
	24	116278	48969	78	33	33.0
	32	29952	19694	20	13	9.9
POKER SD	16	255257	28137	170	19	56.7
	24	42519	23897	28	16	13.3
	32	12879	9133	9	6	4.4
POKER HD	16	522905	140687	349	94	132.7
	24	125554	18101	84	13	28.7
	32	38054	8555	25	6	9.3
MUSIC SD	16	17116	66811	114	45	36.0
	24	20119	11987	13	8	6.4
	32	6429	39	4	1	1.3
MUSIC HD	16	535290	47054	374	31	116.5
	24	29170	21991	19	15	10.2
	32	8642	7346	6	5	3.2

Table D.2: Five-way ANOVA on QoE of 3D Video

Source	Degree of freedom	F-statistics	p-value
CT	2	141.7391	0
R	2	124.416	0.02975
QP	4	165.981	0
PLR	5	639.172	0
MBL	3	73.354	0.01003
PLR+CT	10	10.218	0.11909
PLR+R	8	5.182	0.3342
PLR+QP	18	71.955	0.1264
PLR+MBL	15	30.466	0.2921
MBL+CT	5	2.868	0.1942
MBL+R	5	7.940	0.52301
MBL+QP	12	13.533	0.29968

Appendix E

The Extracted Rules of FIS-A Model

Due to the high number of the extracted fuzzy rules, which are about 245 rules, we listed some of them in the following:

if the video motion is low and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is low Then QoE is poor

if the video motion is low and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is moderate Then QoE is poor

if the video motion is low and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is high Then QoE is poor

if the video motion is low and the QP is moderate and the Resolution is low and the PLR is low and the MBL is high Then QoE is poor

if the video motion is low and the QP is moderate and the Resolution is low and the PLR is low and the MBL is moderate Then QoE is poor

if the video motion is low and the QP is moderate and the Resolution is low and the PLR is low and the MBL is low

Then QoE is fair

if the video motion is moderate and the QP is low and the Resolution is low and the PLR is moderate and the MBL is low Then QoE is poor

if the video motion is moderate and the QP is low and the Resolution is low and the PLR is moderate and the MBL is moderate Then QoE is bad

if the video motion is moderate and the QP is low and the Resolution is low and the PLR is moderate and the MBL is high Then QoE is bad

if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is high Then QoE is bad

if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is moderate Then QoE is bad

if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is moderate and the MBL is low Then QoE is poor

if the video motion is low and the QP is low and the Resolution is moderate and the PLR is low and the MBL is low Then QoE is good

if the video motion is low and the QP is low and the Resolution is moderate and the PLR is low and the MBL is moderate Then QoE is good

if the video motion is low and the QP is low and the Resolution is moderate and the PLR is low and the MBL is high Then QoE is good

if the video motion is low and the QP is low and the Resolution is low and the PLR is moderate and the MBL is high Then QoE is bad

if the video motion is moderate and the QP is moderate and the Resolution is high and the PLR is moderate and the MBL is moderate Then QoE is bad

if the video motion is moderate and the QP is moderate and the Resolution is high and the PLR is moderate and the MBL is high Then QoE is bad if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is low and the MBL is high Then QoE is poor

if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is low and the MBL is moderate Then QoE is poor

if the video motion is high and the QP is moderate and the Resolution is moderate and the PLR is low and the MBL is low Then QoE is poor

if the video motion is high and the QP is low and the Resolution is high and the PLR is low and the MBL is low Then QoE is fair