

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

DEVELOPING A HIGH DEFINITION
VIDEO QUALITY OF EXPERIENCE
MODEL
BASED ON INFLUENTIAL
PARAMETERS

A THESIS PRESENTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
DOCTOR OF PHILOSOPHY
IN
COMPUTER SCIENCE
AT MASSEY UNIVERSITY, PALMERSTON NORTH,
NEW ZEALAND.

Syed Jawad Hussain

2015

Contents

Abstract	xi
Acknowledgements	xiii
1 Introduction	1
1.1 Brief Thesis Description	1
1.2 Research Context and Background	2
1.2.1 Evolution of QoE from QoS	4
1.3 Description of the Problem	6
1.4 Motivations for This Thesis	7
1.5 Thesis Aims and the Scope of Research	9
1.5.1 Research Aims	9
1.5.2 Scope	10
1.6 Summary of Contributions	11
1.7 Thesis Outline	12
2 Literature Review on Quality of Experience	15
2.1 Chapter Overview	15
2.2 Quality of Experience	15
2.2.1 Definitions of QoE	15
2.2.2 QoE vs QoS	16
2.2.3 QoE for IPTV	16
2.3 Artefacts and influencing factors	17
2.3.1 Classification of influencing factors	17
2.4 Content Domain Artefacts	19
2.4.1 Acquisition	19
2.4.2 Coding	20
2.5 Delivery Domain Artifacts	22
2.6 Customer Premises Equipment Domain Artefacts	23
2.7 Standards	25

2.8	Issues with Subjective Data Acquisition	26
2.9	Finding the most influential parameters	26
2.10	Research Questions	27
2.11	Summary and Conclusion	28
3	Framework and Experimental Design	31
3.1	Chapter Overview	31
3.2	Framework	31
3.3	Experiment	37
3.3.1	Limitations and Requirements	37
3.3.2	Experimental Design	38
3.3.3	Data Acquisition	46
3.3.4	Data Validation	48
3.3.5	Ranking of Parameters	49
3.4	Model for MOS Prediction	52
3.4.1	Predictive Models	52
3.4.2	Regression Model	52
3.4.3	Logistic Regression	52
3.4.4	Maximum Likelihood Estimations	53
3.5	Discriminant Analysis	53
3.5.1	Logistic vs. Discriminant	54
3.6	Validation	54
3.7	Ethics	54
3.8	Summary and Conclusion	55
4	Content Domain Influential Parameters	57
4.1	Introduction	57
4.2	Design of Experiment for content domain	58
4.2.1	Parameters	58
4.3	Experimental Setup	59
4.3.1	Test Bed	59
4.3.2	Selection of Test Sequences	60
4.3.3	Test Layout	60
4.3.4	User Task	61
4.3.5	Customization	61
4.4	Results and Analysis	62
4.4.1	Pairwise Correlation	62
4.4.2	Fleiss' Kappa	63
4.4.3	Signal to Noise Ratio Analysis	64

4.4.4	Quality Function Deployment	70
4.4.5	Comparison of Taguchi Method and HoQ Matrix	70
4.4.6	Feedback from observers	73
4.5	Summary and Conclusion	74
5	Network Domain Influential Parameters	75
5.1	Introduction	75
5.2	Design of Experiment for network domain	76
5.2.1	Parameters	76
5.3	Experimental Setup	78
5.3.1	Setup	78
5.3.2	Test Layout	78
5.3.3	User Task	79
5.3.4	Customization	79
5.4	Results and Analysis	79
5.4.1	Pairwise Correlation and Fleiss' Kappa	80
5.4.2	Signal to Noise Ratio Analysis	81
5.4.3	Quality Function Deployment	85
5.4.4	Comparison of Taguchi Method and HoQ Matrix	88
5.5	Summary and Conclusion	88
6	CPP Domain Influential Parameters	91
6.1	Introduction	91
6.2	Design of Experiment for CPP domain	92
6.2.1	Parameters	92
6.2.2	Test Bed	93
6.2.3	Selection of Test Sequences	94
6.2.4	Test Layout	95
6.2.5	User Task	95
6.2.6	Customization	95
6.3	Results and Analysis	96
6.3.1	Pairwise Correlation	96
6.3.2	Fleiss' Kappa	97
6.3.3	Signal to noise ratio analysis	97
6.3.4	Quality Function Deployment	101
6.3.5	Comparison of Taguchi method and HoQ matrix	103
6.4	Summary and Conclusion	104

7	End to End Study of Influential Parameters	107
7.1	Introduction	107
7.2	Design of Experiment for end to end influence	108
7.2.1	Parameters	108
7.2.2	Test Bed	109
7.2.3	Selection of Test Sequences	110
7.2.4	Test Layout	111
7.2.5	User Task	111
7.2.6	Customization	112
7.3	Results and Analysis	112
7.3.1	Pairwise Correlation	112
7.3.2	Fleiss' Kappa	113
7.3.3	Signal to Noise Ratio Analysis	113
7.3.4	Quality Function Deployment	116
7.3.5	Comparison of Taguchi Method and HoQ Matrix	118
7.4	Summary and Conclusion	119
8	Modeling & Discussion	121
8.1	Chapter Overview	121
8.2	Model Development	121
8.2.1	Ordinal Logistic Regression	121
8.2.2	Model Validation	128
8.2.3	Model Deployment	128
8.3	Summary	129
9	Conclusions & Future Work	135
9.1	Chapter Overview and Summary	135
9.2	Summary of Research Outcome	137
9.3	Conclusions	137
9.3.1	Novel Contributions	142
9.4	Future Work	143
A	Code Fragments and additional tables	147
A.1	Code Fragments	147
A.2	Tables	148
	Bibliography	173

List of Tables

3.1	HoQ Questionnaire	47
4.1	Control Factors and Levels.	58
4.2	Noise Factors and Levels.	59
4.3	Array with control factors.	60
4.4	Content Effect Table.	64
4.5	MOS scores per condition per trial.	64
4.6	Response Table.	66
4.7	Predicted SNR values for optimal configuration	69
4.8	HoQ with raters and experts feedback.	71
4.9	Calculated Weights for HOQ.	72
4.10	P-value table using permutation sampling.	73
5.1	Control Factors and Levels.	76
5.2	Noise Factors and Levels.	78
5.3	L9 Array with control factors.	79
5.4	MOS scores per condition per trial.	81
5.5	Response Chart	82
5.6	Effect Table	83
5.7	Predicted SNR values for optimal configuration	85
5.8	HoQ with raters and experts feedback.	87
5.9	Calculated Weights for HOQ.	88
5.10	P-value table using permutation sampling.	89
6.1	Control Factors and Levels.	92
6.2	Noise Factors and Levels.	93
6.3	Array with control factors.	94
6.4	Content Effect Table.	98
6.5	MOS scores per condition per trial.	98
6.6	Response Table.	100
6.7	Predicted SNR values for optimal configuration	101

6.8	HoQ with raters and experts feedback.	102
6.9	Calculated Weights for HOQ.	103
6.10	P-value table using permutation sampling.	105
7.1	Control Factors and Levels.	108
7.2	Noise Factors and Levels.	109
7.3	Array with control factors.	110
7.4	Content Effect Table.	113
7.5	MOS scores per condition per trial.	114
7.6	Response Table.	116
7.7	HoQ with raters and experts feedback.	117
7.8	Calculated Weights for HOQ.	118
7.9	P-value table using permutation sampling.	120
8.1	Response Information	123
8.2	Response Information	124
8.3	Ordinal Logistic Regression Table	125
8.4	Goodness of Fit Tests	126
8.5	Measures of Association	126
8.6	Ordinal Logistic Regression Table for Model OLR2	127
8.7	Goodness of Fit Tests for Model OLR2	127
8.8	Measures of Association for Model OLR2	127
A.1	Calculation for Network Domain Fleiss' Kappa	148
A.2	Calculation for CPP Domain Fleiss' Kappa	150
A.3	Calculation for End to End Fleiss' Kappa	152
A.4	Factorial design for 5 parameters	155
A.5	End to end experiment additional data for validation from 16 raters with mean	163
A.6	MOS per video per condition	169

List of Figures

2.1	Utility cycle of QoE	16
2.2	A framework highlighting factors within domains which affect IPTV service delivery	29
3.1	Experiment flow diagram	32
3.2	Methodology	34
3.3	Comparison between Taguchi method and HoQ method	35
3.4	Experiment flow diagram	36
3.5	House of Quality Matrix	51
4.1	Combination of noise factors	60
4.2	Correlation among raters	63
4.3	Interaction Plot	65
4.4	Pie Chart of influential parameters of content domain	67
4.5	Combination of noise factors	68
4.6	Optimal parameter configuration for content domain	69
4.7	Pie Chart of influential factors using HoQ matrix	72
5.1	Correlation among raters	80
5.2	Interaction Plot	83
5.3	Pie Chart of influential factors	84
5.4	Optimal parameter configuration for network domain	85
5.5	Pie Chart of influential factors using HoQ matrix	86
6.1	Correlation among raters	97
6.2	Interaction Plot	99
6.3	Pie Chart of influential factors	100
6.4	Optimal parameter configuration for CPP domain	103
6.5	Pie Chart of influential factors using HoQ matrix	104
7.1	Test bed setup	109
7.2	Correlation among raters	112

7.3	Interaction Plot	115
7.4	Pie Chart of influential factors	118
7.5	Pie Chart of influential factors using HoQ matrix	119
8.1	MOS, categorical data from Group1 and prediction by Model OLR1 . . .	130
8.2	Cumulative probability plot for each video	131
8.3	MOS, categorical data from Group1 and prediction by Model OLR2 . . .	132
8.4	MOS, categorical data from Group1 and prediction by Model OLR1 and Model OLR2	133

Abstract

Multimedia applications are quickly becoming a basic necessity for both business and personal situations. They are being used on daily basis, for a big part of the population, it's becoming the main source of entertainment and communication. Internet based Television and IPTV are now being preferred over standard Television due to the added functionality that it provides. Users not only expect good quality they also require it to be available continuously without fail. The research reported in this thesis is focused on the "Development of a high definition video Quality of Experience (QoE) model based on influential parameters". We have proposed a model for predicting QoE for high definition videos. We sought and justified appropriate parameters for inclusion into our prediction model. Our analysis has helped us in identifying parameter quantization as the most influential parameter from content domain. Whereas the parameters such as packet loss, packet reorder and jitter were found to be equally influential on QoE from the network domain. We also were able to show that parameter buffer size was the most influential parameter from the customer premises processing domain. We also showed that by integrating these parameters from the three domains of content, network and customer premises processing has enabled us to predict QoE with greater accuracy. The model enables the service provider to predict QoE and act before user perception goes beyond a predetermined threshold level. By including all the most influential parameters from all three domains we have ensured that we were accounting for a complete end to end effect of a user's perception. We were also able to identify the optimal configuration for good quality. In addition, we determined configurations that reduced the output quality to unacceptable levels that should be avoided. We were also able to find out a mathematical relationships between these parameters.

We have discussed the current state-of-the-art knowledge for QoE in the three identified domains and then proceeded to identify the shortcomings of models and methods found in the current literature. We made use of Taguchi robust designs for reducing the run time of experiments. This also helped us in analyzing many different configurations but only conducting experiments on a subset of these configurations. For three experiments the Taguchi DoE technique helped us reduce the number of test combinations from 81 to only 9 combinations. For the fourth experiment we were able to reduce the

number of combinations from 256 to only 18 combinations. We were able to reduce known boredom and memory effects by careful design of our experiments. Moreover, we were able to analyze the results using an appropriate signal to noise ratio (SNR) analysis method. This method helped us identify the most influential parameters for QoE in the three identified domains and it also assisted in the identification of optimal configurations. The method utilize quality loss model for identifying optimum configurations. By this method we found the SNR for each configuration and those individual parameter configurations were selected which ensured the highest SNR values. This helped identify the optimum configurations from each domain. In addition, we used the House of Quality (HoQ) method to validate the results of our SNR analysis. We were able to show that for individual domain the HoQ method was able to verify the results of our SNR analysis. HoQ only failed when we tried to use it to verify the end to end video quality degradation. We identified that the reason for HoQ method's failure in generating acceptable results was the use of weights to identify the most influential parameters in situation were all parameters under investigation where already found to be influential. These heavy weight parameters canceled the effect of each other and the parameters became statistically insignificant.

We proposed a QoE prediction model by using the most influential parameters identified from each domain. Our results show that these influential parameters were the major reasons for degradation in quality. Hence, using such parameters can help in better prediction of video quality degradation.

We made use of ordinal logistic regression for developing the prediction model. We proposed a complete model which was able to predict quality upto an 88% accuracy. Later, we developed a reduced version of the model for low computing solutions which was able to predict quality with 84.5% accuracy.

Acknowledgements

I would like to express my deepest and sincere appreciation and gratitude to my supervisor Prof. Richard J. Harris for his constant and continuous support, encouragement, understanding, patience, advice, criticisms and opinions. His unconditional support was the main reason I was able to finish this study.

I would also like to thank Dr. Amal Punchihewa for his polite and comforting mentoring and support. He was able to discuss issues with such a calm manner that it atleast momentarily pulled me out of the research blues.

I also would like to thank Higher Education Commission Pakistan for funding my studies and support me during this PhD study. I am once again thankful to Prof. Richard J. Harris for his financial support for conferences and living expenses at the end of my studies in New Zealand.

I would also like to thank my friends at Massey University especially Dr. Stephen Lean for sharing ideas and long discussions.

Lastly, I would like to thank my family for their unconditional support and motivation. It would not be possible for me to finish this PhD study without my mother. Her continuous support, encouragement, prayers and financial help pushed me towards achieving this goal. My wife who was always there when I needed a shoulder for support. My three kids who made me forget my work issues the moment I entered my home. A Special thanks to my brother and sister who prayed for my success since forever. A very special thank you to my father Syed Sajjad Hussain (Late) who planted the idea for my higher studies. I dedicate my successes to him.