

**3D Multiple Description Coding for Error
Resilience over Wireless Networks**

By

Abubakar Sadiq Umar

A thesis submitted for the degree of
Doctor of Philosophy
in
Electronic & Computer Engineering

School of Engineering and Design
Brunel University
September 2011

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public; therefore I authorise Brunel University to make available electronically to individual or institutions for the purpose of scholarly research.

Signature:.....

Abubakar Sadiq Umar

Date: 26th September 2011

I further authorise Brunel University to reproduce this thesis by photocopying or by any other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Signature:.....

Abubakar Sadiq Umar

26th September 2011

Copyright 2011 Umar Abubakar
All Rights Reserved

Abstract

Mobile communications has gained a growing interest from both customers and service providers alike in the last 1-2 decades. Visual information is used in many application domains such as remote health care, video –on demand, broadcasting, video surveillance etc. In order to enhance the visual effects of digital video content, the depth perception needs to be provided with the actual visual content. 3D video has earned a significant interest from the research community in recent years, due to the tremendous impact it leaves on viewers and its enhancement of the user’s quality of experience (QoE). In the near future, 3D video is likely to be used in most video applications, as it offers a greater sense of immersion and perceptual experience. When 3D video is compressed and transmitted over error prone channels, the associated packet loss leads to visual quality degradation. When a picture is lost or corrupted so severely that the concealment result is not acceptable, the receiver typically pauses video playback and waits for the next INTRA picture to resume decoding. Error propagation caused by employing predictive coding may degrade the video quality severely. There are several ways used to mitigate the effects of such transmission errors. One widely used technique in International Video Coding Standards is error resilience.

The motivation behind this research work is that, existing schemes for 2D colour video compression such as MPEG, JPEG and H.263 cannot be applied to 3D video content. 3D video signals contain depth as well as colour information and are bandwidth demanding, as they require the transmission of multiple high-bandwidth 3D video streams. On the other hand, the capacity of wireless channels is limited and wireless links are prone to various types of errors caused by noise, interference, fading, handoff, error burst and network congestion. Given the maximum bit rate budget to represent the 3D scene, optimal bit-rate allocation between texture and depth information rendering distortion/losses should be minimised. To mitigate the effect of these errors on the perceptual 3D video quality, error resilience video coding needs to be investigated further to offer better quality of experience (QoE) to end users.

This research work aims at enhancing the error resilience capability of compressed 3D video, when transmitted over mobile channels, using Multiple Description Coding (MDC) in order to improve better user's quality of experience (QoE).

Furthermore, this thesis examines the sensitivity of the human visual system (HVS) when employed to view 3D video scenes. The approach used in this study is to use subjective testing in order to rate people's perception of 3D video under error free and error prone conditions through the use of a carefully designed bespoke questionnaire.

Author's Contributions

- The thesis considered the design of 3D multiple description coding schemes for error resiliency. We designed multiple description coding by extending the optimised 3D multiple description coding with side information and pixel averaging (3D MDC SIPA) of [29]. We show that the new proposed scheme 3D multiple description coding with side information and motion interpolation (3D MDC SIMI) subsumes the 3D MDC SIPA, besides offering an additional benefit of error resiliency.
- We developed two side distortions (Description 1 and description 2) from the side encoders. These two side distortions restrict the amount of redundancy in the central encoder. Each motion vectors in the central distortion is calculated as a weighted sum of two motion vectors in the side distortion.
- The advantages of using MDC over Single Description Coding (SDC) with and without side information is investigated in this thesis.
- The author wrote the necessary program code and performed all the simulation runs obtaining the efficiency of our Scalable 3D MDC SIMI algorithm that performs better than the efficiency of the Scalable 3D MDC SIPA.
- The author has done the necessary derivations and carried out optimisation procedure of the new MDC Algorithm for varying channel conditions. The author has also performed simulations of video transmission over the network with packet loss rate (PLR).
- The author has prepared and written all the sections in this thesis from introductory chapter, chapter 2, chapter 3, chapter 4, 5 and 6 respectively. The author has completed the final PhD thesis and editing of the final document based on comments from my supervisors.
- **Journal papers:**
 - 1) A. Umar, R. Swash, A. Sadka, "Subjective Quality Assessment of 3D ", IEEE Trans. And Circuit for Video Communications. (Submitted for review on 25th August 2011).
 - 2) A. S. Umar, A. H. Sadka, "Scalable Multiple Description Coding for 3D Video", IEEE Trans. on Circuits (Submitted for review on 23rd September 2011).
- **Conference Papers:**

- 1) A. Umar, A. Sadka, "3D Video Communication Over Wireless Networks", RESCON 2009 Conference, Brunel University, Uxbridge, UK.
- 2) A. Umar, A. Sadka, "DIBR Algorithm for 3D video Services", RESCON2010 Conference, Brunel University, Uxbridge, UK.
- 3) A. Umar, A. Sadka, Scalable Multiple Description Coding for 3D video communication over wireless networks, IEEE Nigercon 2010, pp. 115-125, 17-19 June 2010.
- 4) A. Umar, A. Sadka, R. Swash, "Scalable 3D video with reduced resolution depth compression", Sixth International PhD & DLA Symposium, University of Pecs, Pollack Mihaly Faculty of Engineering, 25-26 October, 2010, Poland.
- 5) A. S. Umar, R. M. Swash, A. H. Sadka, "Subjective Quality Assessment of 3D Videos", IEEE Africon11 Signal and Image Processing, Livingstone, Zambia. 13-15 Sept. 2011.
- 6) A. S. Umar, A. H. Sadka, "3D Image and Video Quality Assessment", IEEE 4th International Congress on Image and Signal Processing (CISP 2011) and 4th International Conference on BioMedical Engineering and Informatics (BMEI 2011), Shanghai, China. 15th -17th October 2011.
- 7) A. S. Umar, R. M. Swash, A. H. Sadka, "Scalable Multiple Description Coding for 3D Videos and Images", IEEE 2nd International Conference on Electrical & Control Engineering (ICECE)", 16th -18th Sept. 2011. Ramada Yichang Hotel Yichang, China.

Acknowledgements

First and foremost I would like to thank Almighty Allah (SWA) for taking me through successfully over all these years, it's quite a journey and experience for the rest of my personal and professional life. I would also like to thank him for blessing me with the ability, strength, favour and fortitude to carry out this research work to completion.

This achievement wouldn't have been possible without help, support and inspiration of my supervisors. I would like to express my sincere appreciation and gratitude to my supervisors. I had the honour to work with two of the well-recognized researchers in the field of multimedia, image processing and 3D video communication. Prof. A. H. Sadka and Dr. Ammar Aggoun, are role models, talented and well respected researchers. Professor Sadka's comments and suggestions on my work have always made my work much easier. His knowledge, discipline, availability, patient, constructive criticism of my work and support are unique. My relationship with him from the beginning till end is exceptional; he is being very supportive especially when I get stacked in my research work at Brunel. He is always there for me for discussion and clarification meetings.

Dr. Ammar Aggoun is my second supervisor; he is well respected in the area of 3D video compression. His supervision and visionary advice helps a lot towards completing this research work. His comments and suggestions during visits to CMCR lab help a great deal in shaping me as a researcher.

My special thanks to Abba Iya, Engr. Nura Rafindadi, Engr. Ahmad Kawu, Dr. Aliyu Idi, Dr. Adamu Ibrahim, Dr. Mustapha Gidado and Abubakar Mansur.

Thanks go to Karen Thompson for bringing encouragement and joy to colour my life in CMCR Lab/ECE of Brunel University. I thank her for great help with some routine but important administrative work.

I also owe thanks to my colleagues in CMCR lab. I was very lucky to be surrounded by wonderful and bright people. Thank you Rafiq Swash for always being available to discuss my

research and rescue me from programming and computer related issues. Thank you, Ghaidaa Al-sultany, Nawaz, Sanusi, Mohib, Obaid and Amal. I appreciate the friendship and collaboration.

My brothers and sisters deserve special thanks for always supporting me, without which I would not have been able to face the challenges of this research work. I appreciate all the prayers.

Finally, I would like to thank my employer CSTD/NASRDA, for giving me this opportunity to further my studies upto this level. Particularly, the director CSTD and the Director General Dr. S. O Muhammed who kindly dealt with all the bureaucratic procedures and approvals needed to enable me secure study leave from the office.

My appreciation to Petroleum Technology Development Fund (PTDF) for funding this research work under the Overseas Scholarship Scheme (OSS) for PhD/Doctoral funding. Particularly Tijjani Galadima and his colleagues in the training department for working round the clock to make sure we receive payments as at when due.

Dedication

I dedicate this thesis to my beloved parents Sheikh Modibbo Umaru and Hajiya Aishatu Umaru for parental guidance, advice that sees me through as a growing child to manhood.

This work is also dedicated to my beloved wife, Asmau and my five children, Ahmad, Farouq, Mahmud, Aisha and Ibrahim. The sacrifices you all have made during the past three years in support of this work are greatly appreciated. You made it possible for me to complete this task and there is no way I would ever be able to have done it without you. Your understanding, love, and support brought me through the challenges and the long nights and provided the support necessary to finish this thesis. I love you all, always and forever. Thanks for being there for me, during the cause of my PhD studies here in UK.

Table of Contents

AUTHOR'S DECLARATION.....	ii
Abstract.....	iv
Author's Contributions	vi
Acknowledgements.....	viii
Dedication.....	x
Table of Contents.....	xi
List of Figures.....	xvii
List of Tables.....	xxi
List of Acronyms.....	xxii
List of Symbols.....	xxv
Chapter 1 : Introduction.....	1
1.1 The Context.....	1
1.2 The Problem Statement.....	1
1.2.1 Heterogeneity problems.....	5
1.2.2 Content Adaptation to overcome heterogeneity problems.....	6
1.3 The Solution.....	6
1.4 Thesis Objectives.....	7
1.5 Motivation for Multiple Description Coding.....	7
1.6 Thesis Outline.....	10
References.....	13
Chapter 2: Review of Wireless Technology.....	14
2.1 Introduction.....	14
2.2 Evolution from 1 st to 4 th generation networks.....	16
2.3 WiMAX.....	18
2.3.1 WiMAX Architecture.....	19
2.4 Wi-Fi.....	20
2.4.1 Wi-Fi Architecture.....	22
2.5 4G Network.....	22
2.5.1 4G Requirements.....	23

2.5.2	4G Network Architecture.....	24
2.6	Overview of UMTS Technology	25
2.6.1	UMTS Architecture.....	26
2.6.2	The User	28
2.6.3	Access Network	28
2.6.4	The Node B	30
2.6.5	The Radio Network Controller (RNC).....	31
2.6.6	UTRAN Protocols.....	32
2.6.7	UTRAN Transport Network	33
2.6.8	UMTS Core Network.....	33
2.7	UMTS Services.....	35
2.7.1	UMTS MBMS Services	35
2.7.2	UMTS Bearer Services	36
2.7.3	UMTS Teleservices.....	37
2.7.4	UMTS Supplementary Services (SS).....	38
2.8	UMTS QoS	38
2.8.1	Application/Services requirements to provide QoS for Videos	39
2.8.2	QoS at Various Levels	41
2.9	3D Video Transmission over UMTS	42
2.10	UMTS Simulator.....	44
2.10.1	OPNET Modeller	45
2.10.2	OPNET Simulation Technologies.....	51
2.10.3	OPNET UMTS Model	52
2.10.4	OPNET Scenarios	55
2.10.5	Applications and Profiles	56
2.10.6	Analysis of results.....	56
2.10.7	BER vs Eb/No over UMTS channel	57
2.10.8	Simulation of a Brunel Local Area Network	58
2.10.9	UMTS Simulation in Brunel University	62
2.11	UMTS Model Verification and Validation	65
2.11.1	UMTS Model Verification.....	65
2.11.2	UMTS Validation.....	66

2.12	Summary	69
	References	70
	Chapter 3: 3D Video Technology	73
3.1	Introduction of H.264/AVC-SVC	73
3.2	Review of H.264/AVC (Advanced Video Coding)	73
3.2.1	H.264/AVC Encoder/Decoder Architecture	75
3.3	Scalable Video Coding.....	76
3.3.1.	Layered Video Coding	79
3.3.2.	Temporal Scalability	80
3.3.3	Spatial Scalability and inter-layer prediction	80
3.3.4	SNR Scalability.....	81
3.3.5	Fine Grain Scalability	81
3.4	Scalable Extension of AVC/H.264	82
3.4.1	Spatial scalable results	83
3.5	Overview of 3D Technology.....	85
3.5.1	Introduction.....	85
3.5.2	Human Visual System (HVS).....	87
3.5.3	3D content generation	88
3.6	3D Display	91
3.6.1	Colour filtered anaglyph	91
3.6.2	Polarized glasses	91
3.6.3	Spectrum Filtered-Dolby 3D.....	93
3.6.4	Binocular with active glasses.....	94
3.6.5	Auto-stereoscopic Displays.....	94
3.6.6	Other methods for 3D video generation (Using Matlab)	96
3.7	Summary	100
	References.....	102
	Chapter 4: 3D MDC with Side Information and Motion Interpolation.....	105
4.1	Introduction.....	105
4.2	Literature Review of MDC Algorithms	107
4.2.1	MDC Quantisation	109
4.2.2	MDC Transform Coding.....	110

4.2.3	MDC Sub-sampling	110
4.2.4	MDC Even and Odd Frames	111
4.3	Motion Interpolation	113
4.3.1	FME and BME	115
4.3.2	Decision Mechanism	116
4.3.3	Motion Vector Selection	118
4.4	Proposed Scalable 3D MDC with Side Information & Motion Interpolation.....	118
4.4.1	Encoder Architecture	119
4.4.2	Decoder Architecture	121
4.5	Video Quality Performance Metrics/Parameters	122
4.5.1	Peak Signal to Noise Ratio.....	122
4.5.2	Structural Similarity Index Measure (SSIM)	123
4.5.3	Video Quality Metric (VQM)	123
4.5.4	Moving Picture Quality Metric (MPQM)	124
4.5.5	Packet Loss	124
4.5.6	Objective Performance Comparison	125
4.5.7	Packet Delay Variation or Jitter	125
4.5.8	Rate Distortion (R-D) Performance	126
4.6	Description of the experimental environment.....	126
4.6.1	Experiment.....	127
4.6.2	Simulation and discussions under error prone conditions.....	129
4.6.3	Objective Performance under Error Free Conditions for all the MDC algorithms	130
4.6.4	Objective Performance under Error Prone Conditions for all the MDC algorithms.....	132
4.6.5	Subjective Performance under Error Free Conditions	133
4.6.6	Subjective Performance under Error Prone Conditions	135
4.7	Conclusions.....	138
	References.....	140
	Chapter 5: Quality Assessment of 3D Videos	145
5.1	Introduction.....	145
5.2	Review of Subjective Quality Assessment Methods	145
5.3	Research Methodology	149
5.3.1	Equipment used and viewing conditions.....	151

5.3.2	Databases of Subjective results & test materials used in the study.....	151
5.3.3	Participants.....	152
5.3.4	Health and Safety.....	154
5.3.5	Protocol.....	155
5.3.6	Grading.....	155
5.4	Mean opinion scores and confidence interval.....	157
5.5	Results & Analysis.....	158
5.5.1	Tests of Normality.....	162
5.6	Analysis of Variance for MOS validation.....	166
5.7	Analysis & Discussion.....	170
5.7.1	Limitations of the survey.....	170
5.7.2	Other Findings of the survey.....	171
5.8	Relationship between Objective and Subjective Test Measure.....	171
5.8.1	Metric related to prediction accuracy.....	173
5.8.2	Metric related to prediction Monotonicity.....	173
5.8.3	Metrics related to prediction Consistency.....	173
5.8.4	Objective and Subjective comparison.....	173
5.9	Conclusion.....	177
	References.....	178
	Chapter 6: Conclusions & Future Work.....	182
6.1	Critical Summary.....	182
6.2	Conclusions.....	184
6.3	Future Work.....	184
6.3.1	3D MDC-SIMI with 4 descriptions.....	185
6.3.2	Improvement of the designed Algorithm.....	185
6.3.3	Understanding 3D User's Experience.....	185
6.3.4	3D Video Content Generation.....	186
6.3.5	Scalable extension of H.264/AVC.....	186
6.3.6	3D MDC –SIMI and channel interleaving.....	187
6.3.7	Motion Estimation Improvement.....	187
6.3.8	UMTS channel behaviour.....	187
6.3.9	3D Video Quality Metrics.....	187

References.....	189
Appendix A Sample Questionnaire.....	190
Appendix B: Methods of 3D Video Generation	194
Appendix C: H.264/AVC Encoder & Decoder Commands used in the thesis	197
Appendix D: Matlab Scripts for MDC.....	203
Appendix E: 3D MDC Comparison.....	205
Appendix F: Colour extraction from 3D image	208
Appendix G: 3D Video Quality Objective Assessment.....	209

List of Figures

Figure 1-1: Unicast vs Multicast.....	5
Figure 2-1: Basic Wireless Concepts [6].	14
Figure 2-2: WiMAX coverage with broadband internet [8]	19
Figure 2-3: Architecture of WiMAX [8].....	20
Figure 2-4: Architecture of Wi-Fi [11]	22
Figure 2-5: Evolution to 4G network [7].....	23
Figure 2-6: 4G Network Architecture [11]	24
Figure 2-7: UMTS Network Architecture [16].....	27
Figure 2-8: UMTS Domains	27
Figure 2-9: UMTS User Domain	28
Figure 2-10: UMTS Access Network and Interfaces.....	29
Figure 2-11: UTRAN Architecture [19]	30
Figure 2-12: UMTS RNC	31
Figure 2-13: UTRAN Protocol functions.....	32
Figure 2-14: UTRAN Transport Network [18].....	33
Figure 2-15: UMTS Core Network domain	34
Figure 2-16: UMTS Services [21]	35
Figure 2-17: MBMS in UMTS network [21].....	36
Figure 2-18: UMTS bearer service architecture [22].....	37
Figure 2-19: UMTS Traffic and Services classes [26].....	39
Figure 2-20: Architecture of QoS at Various Levels	42
Figure 2-21: UMTS video streaming scenarios [28].....	43
Figure 2-22: UMTS Video Transmission [28].....	43
Figure 2-23: Project Editor [22].....	46
Figure 2-24: Node Editor.....	47
Figure 2-25: Process Editor.....	48
Figure 2-26: UMTS network using application traffic (Advanced) [26]	52
Figure 2-27: UMTS network using raw packet generator (Simple) [26].....	53
Figure 2-28: UMTS station model [27]	54

Figure 2-29: Single sector Node B node model [27].	55
Figure 2-30: UMTS network architecture considering 5 Mobile Terminals (MTs).	56
Figure 2-31: UMTS channel performance over Eb/No ranges	58
Figure 2-32: Simulation of 3 wireless laptops.	59
Figure 2-33: 3 transmitting laptop in the simulation scenario.	59
Figure 2-34: Delay experience on the wireless network.	60
Figure 2-35: Average Queuing Delay experience on the wireless network.	60
Figure 2-36: Simulation of 5 wireless nodes.	62
Figure 2-37: Wireless queue size.	61
Figure 2-38: Simulation of 5 UMTS wireless nodes.	62
Figure 2-39: Simulation Sequence Editor.	62
Figure 2-40: Video Packet Queuing Delay.	63
Figure 2-41: Queue size graph in packets.	64
Figure: 2-42: Queue size Confidence Interval.	64
Figure 2-43: Simulation Log file generated	666
Figure 2-44: Analytical Model.	67
Figure 2-45: Simulation vs Theoretical results	69
Figure 3-1: Block diagram of H.264/AVC Encoder [5].	75
Figure 3-2: Scalable H.264/AVC MDC [15].	82
Figure 3-3: Subjective results for frame 15 of the orbi sequence for 2D video: (a) SDC, (b) MDC, and for stereoscopic 3D video: (c) SDC, (d) MDC when subjected to 10dB SNR UMTS channel.	85
Figure 3-4: 3D broadcast chain [28].	87
Figure 3-5: 3D visual depth perception (http://www.strabismus.org)	88
Figure 3-6: (a) Stereoscopic camera set up (b) Stereo images (Left & Right) [27]	90
Figure 3-7: (a) 3D Depth-range camera, (b) 2D luminance and depth [27].	90
Figure 3-8: Multiview video camera configuration [27].	90
Figure 3-9: Anaglyph glasses and anaglyph image [27].	91
Figure 3-10: Linear and circular polarizations (http://www.zalman.com)	92
Figure 3-11: Dolby 3D [29].	93
Figure 3-12: Parallax barrier displays [28].	95
Figure 3-13: Lenticular lens displays [28].	96
Figure 3-14: H.264/AVC Encoder	97

Figure 3-15: Combined left and right images	97
Figure 3-16: Generated depth map.....	98
Figure 3-17: Reconstructed 3D	98
Figure 3-18: Combined Left and Right images	99
Figure 3-19: Depth Map Generated	99
Figure 3-20: Generated 3D	100
Figure 4-1: Basic MDC Codec with two paths, wireless channels & three decoders.....	107
Figure 4-2: Contents of stream 1 and 2 at frame level	112
Figure 4-3: Proposed MI Architecture.....	115
Figure 4-4: Motion Vector Selection.....	117
Figure 4-5: Proposed 3D MDC Encoder.....	120
Figure 4-6: Proposed 3D MDC Decoder.....	121
Figure 4-7: Errors in both streams at the decoder.....	128
Figure 4-8: Comparison of 3D MDC-SIPA and 3D MDC-SIMI algorithms under channel errors.....	129
Figure 4-9: Comparison of 3D MDC-SIPA and 3D MDC-SIMI algorithms under errors free.....	130
Figure 4-10: Rate Distortion for Interview Luminance video sequence under error free conditions.....	131
Figure 4-11: Rate Distortion for Interview depth sequence under error free conditions.....	132
Figure 4-12: Rate Distortion for Interview 3D sequence under error free conditions.....	132
Figure 4-13: Mean PSNR vs Bitrate (Kbps) for Interview 3D video at 10% packet loss	133
Figure 4-14: Subjective quality-Interview at no packet loss of luminance for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 80.....	134
Figure 4-15: Subjective quality-Interview at no packet loss of depth for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 52.....	134
Figure 4-16: Subjective quality-Interview at no packet loss of Stereoscopic 3D video for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 16.....	135
Figure 4-17: Subjective quality-Interview at 10% packet loss of luminance for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 99.....	136
Figure 4-18: Subjective quality-Interview at 10% packet loss of depth for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 49.....	137
Figure 4-19: Subjective quality-Interview at 10% packet loss of Stereoscopic 3D video for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 13.....	137

Figure 5-1: Colour and Corresponding Depth Image (a) Breakdance, (b) Ballet, (c) Interview and (d) Orbi sequence.....	152
Figure 5-2: Group studies in terms of gender using bar chart.....	158
Figure 5-3: Group studies represented in pie chart.....	159
Figure 5-4: 3D Depth Perception, feel presence and overall 3D video quality.....	160
Figure 5-5 (a): MOS of 3D Video Quality in bar chart.....	162
Figure 5-6: Test for Normality of the MOS.....	163
Figure 5-7: MOS values and 95% Confidence Interval using Boxplot.....	164
Figure 5-8: MOS values for 3D video level of satisfaction; a) Males, b) Females.....	165
Figure 5-9: MOS values and 95% confidence interval obtained.....	166
Figure 5-10: F-distribution test for validation of raw MOS [31][32].....	168
Figure 5-11: Subjective vs Objective Performance Measures.....	175
Figure 5-12: DMOS vs PSNR (dB).....	175
Figure 6-1: 3D Videos in MDC with four descriptions.....	185

List of Tables

Table 2-1: Frequency bands for Wireless Communication.....	15
Table 2-2: QoS at Various Levels.....	41
Table 2-3: UMTS QoS Classes.....	41
Table 2-4: The Network and Links used in constructing the UMTS network.....	50
Table 3-1: Description of layers for the scalable MDC encoder.....	83
Table 3-2: Description of layers.....	84
Table 4-1: Comparison of various Objective Performances Metrics.....	125
Table 5-1: Distribution of participants in the survey.....	153
Table 5-2: Variable view of the MOS.....	156
Table 5-3: Data view of the MOS.....	156
Table 5-4: Frequency Distribution of the 3 Group studies.....	15959
Table 5-5: Total number of Participants.....	159
Table 5-6: Analysis of problems associated with 3D viewing.....	160
Table 5-7: Mean, Standard Deviation, Kurtosis and Skewness computed.....	161
Table 5-8: Overall User's needs, rqr, expectation and new technology.....	161
Table 5-9: Test for hypothesis.....	168
Table 5-10: Experimental result for validation.....	169
Table 5-11: Interpretation of results.....	169
Table 5-12: Summary of Analysis.....	174
Table 5-13: Continued Analysis.....	174
Table 5-14: ANOVA Test Results.....	174

List of Acronyms

ARQ: Automatic Repeat Request
AMPS: Advanced Mobile Phone System
ATM: Asynchronous Transmission Mode
ANOVA: Analysis of Variance
BME: Backward Motion Estimation
CN: Core Network
CDMA: Code Division Multiple Access
CI: Confidence Interval
CABAC: Context-Adaptive Binary Arithmetic Coding
CMCR: Centre for Media Communications Research
DIBR: Depth Image-Based Rendering
DPB: Decoded Picture Buffer
DL: Downlink
DF: Degree of Freedom
ECE: Electronic & Computer Engineering
EHF: Extra High Frequency
EVDO: Evolution Data Optimized
FEC: Forward Error Correction
FME: Forward Motion Estimation
FTV: Free-viewpoint Television
GBN: Go Back And
GSM: Global System for Mobile
GGSN: Gateway GPRS Support Node
GR: GPRS Register
GMM: GPRS Mobility Management
GDV: Global Disparity Vector
GOP: Group Of Pictures
HF: High Frequency
HSUPA: High Speed Uplink Packet Access

HSDPA: High Speed Downlink Packet Access
HLR: Home Location Register
ISDN: Integrated Services Digital Network
IPDV: Instantaneous Packet Delay Variation
JSVM: Joint Scalable Video Model
LOS: Line of Sight
LTE: Long Term Evolution
LAN: Local Area Network
MDC: Multiple Description Coding
ME: Mobile Equipment
MSC: Mobile Services Switching Centre
MAC: Medium Access Control
MDC: Multiple Description Coding
MDC OE: Multiple Description Coding Odd & Even
MDC OES: Multiple Description Coding Odd, Even & Side Information
MDC SIMI: Multiple Description Coding Side Information and Motion Interpolation
MDC-SIPA: Multiple Description Coding Side Information and Pixel Averaging
MPEG: Moving Pictures Experts Group
MSR: Microsoft Research
MV: Motion Vector
MVC: Multiview Video Coding
MOS: Mean Opinion Scores
NASRDA: National Space Research & Development Agency
OQA: Objective Quality Assessment
PLR: Packet Loss Rate
PLMN: Public Land Mobile Network
PDP: Packet Data Protocol
PSNR : Peak Signal-to-Noise Ratio
PDC: Personal Digital Cellular
QoE: Quality of Experience
QoS: Quality of Service
RTP: Real time Transport Protocol

RTD: Remote Transmission Diagnostic
RESCON: Research Student Conference
RANAP: Radio Access Network Application Part
RNSAP: Radio Network Subsystem Application Part
RAB: Radio Access Bearer
RAN: Radio Access Network
RNC: Radio Network Controller
RLC: Radio Link Control
SHF: Super High Frequency
SW: Stop and Wait
SR: Selective Repeat
SIM: Subscriber Identity Module
SGSN: Serving GPRS Support Node
SDC: Single Description Coding
SSIM: Structural Similarity
SAD: Sum of Absolute Differences
SSD: Sum of Square Differences
SQA: Subjective Quality Assessment
SPSS: Statistics Package for the Social Science
TDMA: Time Division Multiple Access
UMTS: Universal Mobile Telecommunications
USIM: UMTS Subscriber Identity Module
UE: User Equipment
UL: Uplink
VLR: Visitor Location Register
VCEG: Video Coding Experts Group
VQEG: Video Quality Expert Group
WiMAX: Worldwide Interoperability for Microwave Access
Wi-Fi: Wireless Fidelity
3D TV: Three Dimensional Televisions

List of Symbols

η	Throughput efficiency
ε	Bit error rate (BER)
P_b	Probability of bit error
$P_r, \gamma(\cdot)$	Power received (Watts).
N_b	Background Noise Power
N_i	Interference noise power.
E_b	Received Energy per bit
N_o	Noise power spectral density
L	Total number of bits per frame ($L = l + h$)
l	Number of information bits per frame
R or M	Number of receivers
N	Number of data frame per block
T	Time to transmit one block
ρ	Channel data rate, bits per unit time
q	Bit-error probability
D	Depth map
I_r	Right image
I_l	left image
k	Kurtosis
Skew	Skewness
k_d	differential depth
E	Expectation
df	Degree of freedom
CI	Confidence Interval
H_0 & H_1	Null and Alternate Hypothesis

Chapter 1: Introduction

1.1 The Context

The applications of wired and wireless video services have become in the past few years a challenging task, due to many factors, such as the availability of channel and high quality connections at affordable costs. The provisioning of such services have been in the past often constrained by the limited resources especially in terms of computational efficiency and network capacity; a key factor in the widespread use of these advanced services is the user's quality of experience (QoE), acceptance and awareness. Only recently, the big gap between the user's expectation and the technological evolution has been partially overcome making it possible to enhance the quality of service offering and to provide more sophisticated applications.

1.2 The Problem Statement

In this thesis, the main challenge of 3D video communication over wireless networks is to provide acceptable quality of service (QoS) to the mobile viewers/users. However, wireless networks have a limited bandwidth that may not be able to handle the 3D video frame sequence with the possibility that 3D video frames could be dropped, frozen or corrupted during transmission. 3D video transmission over wireless networks faces other challenges such as high error rates, bandwidth variations and limitation, and capability constraints on the handheld devices [1]. Among these, the unreliable and error-prone nature of the wireless channel is one of the major challenges in transmission of 3D video over wireless channels. The objective is the efficient delivery of 3D video services over wireless networks, capable of limiting the error introduced by the channel. In fact, there is often no return channel to inform the transmitter about the channel conditions and a compromise must be found, in order to allow an almost seamless reconstruction of video content at the decoder, even in the case of high error rates (worst channel conditions). Some of the problems identified are briefly described below.

➤ **Bandwidth variation and limitation**

Introduction

Real-time video transmission/streaming typically have quality of service (QoS) requirements characterized by bandwidth, delay and error packet loss. However, unreliability, bandwidth fluctuations and high bit error rates of wireless channels can cause severe degradation to video quality [1]. In addition, for video multicast, network heterogeneity and receiver heterogeneity make it difficult to achieve efficiency and flexibility. To address the bandwidth fluctuation, packet loss and heterogeneity problems, scalable (layered coding) is employed for delay and bandwidth sensitive videos [2]. Furthermore, to enhance the video quality, i.e error resilience in the presence of unavoidable packet loss and bit error, scalable multiple descriptions coding is employed.

To achieve acceptable quality of experience (QoE), a streaming application typically has minimum bandwidth requirement. However, the current Internet offers both the best effort service and bandwidth reservations across multiple Autonomous Domains[3]. Intra/Inter-domain bandwidth reservation standards (e.g., IntServ, NSIS) have been available and QoS mechanisms also implemented in routers. We can also get bandwidth reservations across Autonomous Systems (ASes) that belongs to a single provider but is up to interconnected NP/ISPs to offer bandwidth reservations across multiple ASes.

In wireless networks, the wireless channel suffers from both bandwidth fluctuation and bandwidth limitation: (1) The throughput of a wireless channel may be reduced due to multipath fading, shadowing, co-channel interference and noise disturbances; (2) When a mobile terminal moves between different networks, the available bandwidth may vary drastically from a few megabits/sec to a few kilobits/sec; (3) When a handoff takes place, a base station may not have enough unused radio resources to meet the demand of a newly admitted mobile host. Thus, the available bandwidth of wireless channel is time-varying and even unknown. If the transmission rate of streaming video is faster than the available bandwidth, the congestion will occur, resulting in bursty losses, excessive delay, and severe drop in video quality. On the contrary, congestion and packet retransmissions invoke the inefficient utilization of available bandwidth and the sub-optimal video quality. Thus, it is desirable for streaming video application to employ congestion control mechanisms to match video bit rate with available bandwidth.

Introduction

➤ **Packet loss and transmission error**

In the wired link of a mobile network or the Internet, most errors are caused by packet loss due to congestion. The effects of packet loss greatly depend on the types of packet loss: isolated single packet loss, burst packet loss and temporary outage such as loss of communication. Packet Reordering and Misrouting may also occur in the downlink during mobile handoffs, which will incur bursty packet loss and temporary outage [3].

Wireless channels are typically more error-prone at the bit level. The wireless link of a mobile network suffers from very high bit error rate (BER) due to multipath fading, shadowing, co-channel interference, noise disturbances and handoff. The types of transmission bit error also can be classified into single bit error and bursty bit error.

The effects of packet loss or bit error are significant for video streaming due to error propagation. Predictive video-encoding algorithms employ motion compensation to achieve high compression by reducing temporal redundancies between successive frames. When this motion information is lost at the decoder, a reconstruction error can occur. Such errors can propagate temporarily and spatially if the affected regions are subsequently used as a prediction for motion compensation. Furthermore, differential encoding is also employed to reduce statistical redundancies, for example in the encoding of motion and quantiser information. Loss of such information can cause additional spatial degradation throughout the affected video frames by producing incorrectly predicted motion vectors and quantiser levels. Because of motion compensation, these errors can also propagate temporally and spatially.

Because of error propagation of streaming video, isolated single packet loss or bit error is converted to bursty packet loss or bit error. Also, the video packet which arrives beyond a delay bound is useless and has to be considered lost in real time applications. Such loss or error can potentially make the visual presentation displeasing to human eyes or even make the presentation impossible. From a video communication perspective, it is important to reduce or eliminate the effects of burst loss/error. To enhance the video quality in presence of unavoidable packet loss or bit error, error control mechanisms should be used.

Introduction

➤ **Heterogeneity**

Before addressing the heterogeneity problems, we first compare unicast with multicast. The unicast delivers streaming videos through point-to-point transmission, where only one sender and one receiver are involved. In contrast, multicast delivers videos through point-to-multipoint transmission, where one sender and multiple receivers are involved. For applications such as video conferencing and Internet television, multicast delivery can achieve high bandwidth efficiency since the receivers can share connection/flow [4]. On the other hand, unicast delivery of such applications is inefficient in terms of bandwidth utilization. An example is given in figure 1-1(a) below, where for unicast, five copies of the video content flow across Link 1 and three copies flow across Link 2, for multicast in figure 1-1 (b), there is only one copy of the video content traversing any link in the network, resulting in substantial bandwidth savings. However, the efficiency of multicast is achieved at the cost of losing the service flexibility of unicast (i.e in unicast, each receiver can individually negotiate service parameters with the source). Such lack of flexibility in multicast can be problematic in a heterogeneous network environment. For example, the receivers in figure 1-1 (b) may attempt to request different video quality with different bandwidth. But only one copy of the video content is sent out from the source.

Introduction

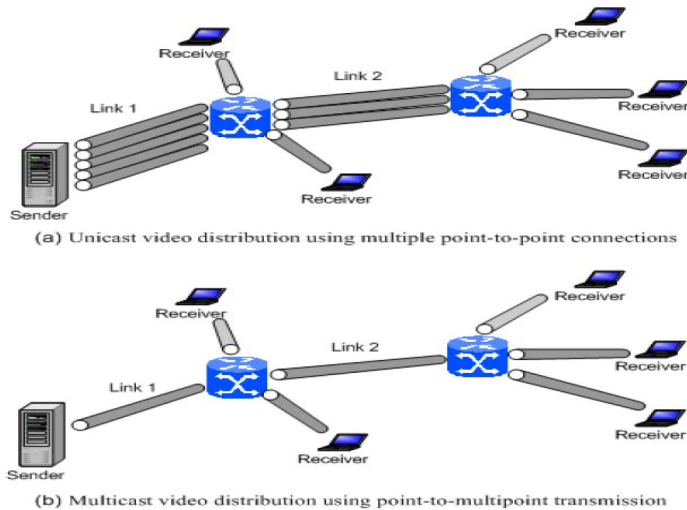


Figure 1-1: Unicast vs. Multicast.

1.2.1 Heterogeneity problems

A network connecting devices with different capabilities and/or protocols or a wireless network which provides a service through wireless LAN and is able to maintain the QoS when switching to a 3G network is called a heterogeneous network. Rapidly increasing volume of multimedia content has imposed restrictions on new applications and services such as varying transmission medium and access network characteristics, end terminal capabilities, and user's preferences to be deployed. These applications and services show very diverse characteristics which need accurate mapping to communication links or access terminals to enable content delivery with acceptable QoS [5]. Video is one of the most important applications of multimedia, as it targets the subjective experience of users through a human visual system (HVS). Some of the heterogeneous environments are:

- Growing amount of content produced by numerous content providers.
- Availability of increasing bandwidth for enhanced multimedia experience over wireless networks.
- Proliferation of a wide range of access devices such as PCs, PDA, laptops and fixed/mobile terminals.

Introduction

1.2.2 Content Adaptation to overcome heterogeneity problems

Content adaptation is employed to overcome the heterogeneity problems. These are: 1) Simulcast distribution model, as it produces several independently encoded copies of the same video content at varying features, such as different bit/frame rates, and spatial resolutions. It then provides the delivery of these copies to serve clients with different connection speeds, allowing them to decide and select which stream to use. Despite being the simplest method, results in over utilising the available transmission bandwidth, to solve this, the content provider has to generate excessive number of versions of video streams to be transmitted. 2) Scalable media model is a solution to the drawbacks mentioned above, as it provides a base layer for minimum requirements, and one or more enhancement layers to offer better qualities at the expense of bit/frame rates. It gives the content provider the opportunity to generate only one basic stream and its interdependent enhancements to cover a wide range of client bandwidth, terminal requirements, and preferences. 3) The third content adaptation strategy is the transcoding of video streams. It is a process to convert one video format into another with different features, and is used to tailor the video content to network characteristics, terminal capabilities, and user preferences. Several video transcoding schemes have been developed to provide a wide range of services, such as frame rate conversions, resolutions and error robustness [6]. This method provides flexible solutions, as transcoding is a middleware operation, and thus is performed on compressed video streams at the edges of different networks. Video transcoding provides more efficient adjustments of video content in response to varying network conditions. Transcoding can be perform on-line or off-line, this will help content providers to keep only one high quality copy of a video streams, thereby reducing the storage costs and the operational complexity.

1.3 The Solution

The focus of the thesis is error resiliency to combat packet loss and transmission errors. The problem can be solved from many different perspectives and several approached can be implemented in order to reduce the error rate at the decoder. Multiple Description Coding

Introduction

(MDC) provides solutions for error resilience techniques. The thesis presents innovative approaches to 3D Multiple Description Coding in chapter 4.

1.4 Thesis Objectives

These objectives are:

- To identify and retrieve all MDC algorithms available in the literature, both for 2D and 3D videos.
- To provide evaluation technique and most significantly develop analytical tools such as Matlab, JSVM software, Opnet simulator for the UMTS simulation and SPSS that facilitate the task.
- Familiarise with research skills, such as literature survey, comparative review, theoretical analysis and programming skills.
- Develop 3D video subjective assessment model.

1.5 Motivation for Multiple Description Coding

Presently, image and video compression are well developed and widely used in the field of signal processing. Modern state of the art coders provide better compression with better quality. This is particularly true for video, as a video sequence possesses correlation in both spatial and temporal domains. The most efficient video compression schemes use motion estimation as we can see in the later part of this thesis and compensation algorithms that exploit prediction to remove the correlation in temporal domain [7]. The correlation in spatial domain is usually removed by methods similar to those of image compression.

A new field in signal processing is representation of scenes in 3D. Recently, it has drawn significant attention of industry, research community and end users. 3D video scenes may be captured by stereoscopic or multi-view camera settings. The captured multi-view video can be compressed directly or converted to more representations and then compressed.

In any case, the efficiently compressed video data has to be transmitted over communication channels, such as wireless channels. This raises the issue of error protection, since most of these channels are error prone. Compressed videos/images and especially video sequences are

Introduction

vulnerable to transmission errors. If the error occurs in a video frame, it may propagate further into subsequent frames because of motion-compensated prediction. Multiview video compression methods utilize comprehensive temporal and inter-view prediction and therefore channel errors occurring in one view can propagate not only to the subsequent frames of the same view but also to the other views.

Common approach to error protection is to consider it as a pure channel problem, separate from the source compression problem. This approach is based on Shannon theory, which states that in principle the source and channel coding tasks can be carried out independently with no loss of efficiency. Following this approach, raw source video sequences are processed in a way which reduces the data rate as much as possible. Reliable transmission of the bit stream to the receiver is provided by a channel coder. The transport mechanism has to be as reliable as possible since a single error in the compressed bitstream might severely damage the reconstructed signal. According to the noisy channel coding theorem (establishes that for any given degree of noise contamination of a communication channel, it is possible to communicate discrete data i.e digital information nearly error-free up to a computable maximum rate through the channel), a near-optimum transmission may be achieved if the data rate does not exceed the channel capacity. However, this can be hardly achieved in practice.

Widely used error-protection methods operate at the data link layer of the OSI, reference model are Selective Repeat (SR), Stop and Wait (S &W) and go-back-N (GBN) [8]. Therefore, error-free transmission is achieved by retransmitting packets that have been lost or corrupted. A problem with such a mechanism is that it causes delays and thus requires larger memory buffers. The latency is at least a packet round trip delay time. A second problem arises when packet losses are caused by network congestions. Trying to retransmit lost packets generates extra data traffic and makes the network even more congested for. Furthermore, retransmissions are virtually impossible in digital broadcasting. During broadcasting, loss of even a single packet may cause the transmitter to receive multiple retransmission requests, an effect called a feedback retransmission such as the Go-Back-N (GBN), Automatic Repeat Request etc.

Introduction

This problem of congestion caused by retransmission is solve at the transport layer of OSI reference, as TCP backs off before retransmission starts. Congestion Avoidance (CA) is the mechanism designed in TCP to deal with lost packets. When TCP connection experiences packet loss, it then switches to CA, where the CA algorithm increments the size of the congestion window much more slowly than in low-start making the congestion window grows linearly than exponentially.

Another approach for reliable transmission over error-prone channels is so called forward error correction (FEC). The compressed bitstream data is distributed between packets and protected by block channel codes by adding extra bits. The data from the lost packets can be reconstructed from the received packets. The choice of the block code length is important. In terms of efficiency, long blocks are preferred since short blocks generate bitstreams with a relatively large number of additional symbols. The law of large numbers also dictates the choice of longer blocks since the number of errors is predicted more easily by long sequences. Just as with the previous approach based on retransmissions, a problem of delays and large memory buffers exists, this time caused by long blocks.

The decoders using the two above-mentioned approaches i.e., ARQ and FEC tolerate no errors. They assume that all the data transmitted is correctly received. Consequently, they spend a considerable amount of resources to guarantee this. Pure channel coding approaches might not be quite feasible in such cases. As an alternative, one can tolerate channel related losses. Assuming that not all data sent has reached the decoder, one can concentrate on ensuring efficient decoding of the correctly received data. In this case, one needs to change the source coding accordingly and, more broadly, to consider the error protection problem as a joint source-channel problem.

As the result multiple description coding (MDC) is an attractive coding approach as it provides a reliable and error resilience source reconstruction using only part of the data sent to the decoder and employs no priority network transmission mechanisms [9]. The MDC is especially advantageous in short-delay media streaming scenarios such as video conferencing and when broadcast over error-prone channels where it provides acceptable reconstruction quality and prevents feedback impulsion in case of packet loss [10][11].

Introduction

1.6 Thesis Outline

This is a manuscript-based thesis that follows the specifications required by the Brunel University for this format. Apart from chapter 1 which is the introductory, the thesis includes five more chapters. Please note some redundancy between chapters because of the manuscript-based format, though slightly modified to offer a logical progression in the thesis.. In this way, the reader can follow the development of MDC and its applications, from two-dimensional to three dimensional. In the following, we give a detailed summary for the following five chapters, which include the main work/contributions that we have made in this thesis:

In Chapter 2, Wireless Communication is reviewed, so also the Universal Mobile Telecommunication Services (UMTS) network architecture and UMTS protocols are discussed. OPNET simulation Modeller is reviewed. First of all, we describe its main features and functionalities, and the UMTS Model as well. Considerations about scenarios, network architecture, applications and profiles are also presented. Node and Global level results are processed. Simulation runs using Opnet, and test conditions for validation of the UMTS simulator using the theoretical results are also presented.

Chapter 3 reviews the fundamental techniques of H.264/AVC (Advanced Video Coding) coding architecture and introduces the basics of International Telecommunication Union (ITU)-T and Motion Picture Enhancement Group (MPEG) standards. Detailed review of scalable video coding technologies is also given in chapter 3 including spatial, temporal, quality/Signal-to-Noise ratio (SNR) and Fine Grain Scalability (FGS) scalabilities. Techniques of layered coding such as enhancement and video coding layers were presented. Detailed overview of 3D technologies, human visual system, 3D content generation methods (Depth-range camera, stereoscopic camera & multiview video configuration) and display technologies such as anaglyph, polarized glasses, spectrum filtered dolby 3D, binocular with active glasses and auto-stereoscopic displays (parallax barrier & Lenticular lens) are given in chapter 3.

Chapter 4 addresses the new 3D Multiple Description Coding (3D MDC with side information and motion interpolation) of stereoscopic video. The first sections of chapter 4 are devoted to introduction and description of previous related work in 2D and 3D multiple description coding

Introduction

(sections 4.1 and 4.2 respectively). Next, we present a method for motion interpolation solutions as applied in this thesis, the forward motion estimation, bidirectional motion estimation; spatial smoothing and bidirectional motion compensation are explained. Its comparison with Scalable 3D MDC SIPA (3D MDC with side information and pixel averaging) is also discussed in chapter 4. This chapter also present the new encoder/decoder architecture and estimates rate distortion performances, under error free and error prone conditions. Finally, chapter 4 presents and compares objective and subjective performances of the two algorithms under error free and error prone conditions.

Chapter 5 discusses the new 3D video subjective assessment technique. Introduction and review of current subjective video assessment is presented in section 5.1 and 5.2 respectively. The proposed methodology for 3D video assessment is discussed in detail. Experimental set up and requirement are described in this chapter. Three subjective video quality assessment experiments were conducted involving 45 human observers. The experiments revealed valuable insight into the human perception, needs, requirements and technology of 3D, so also the problems associated with 3D technology. Human observer mean opinion score (MOS/quality scores)were tabulated. MOS was analysed in detailed using SPSS software¹ as prescribed in Video Expert Quality Assessment Group (VEQG) assessment methods such as the Mean Opinion Scores, mean, Standard. Deviation, Skewness and Kurtosis were used. Other VQEG methods used are Analysis of Variance (ANOVA), Student t distribution method, Pearson linear correlation and Spearman rank order correlation.

We developed a validation model in chapter 5 based on a widely computed Confidence Interval (CI). CI is proposed to provide additional reliability of the computed MOS among the three different experiments conducted.

We finally discussed the results in chapter 5.

Chapter 6 provides the conclusions and discusses the issues for future work. We first present the summary of our research work and key challenges. Next, we present issues for related to the future work such as design improvement for the algorithm, understanding 3D user's experience,

Introduction

3D content generation, scalable extension of H.264/AVC, 3D MDC-SIMI channel interleaving, and motion estimation improvement.

Finally, the limitations of our algorithm are discussed.

References

- [1] M-T. Lu, C-K. Lin, J. Yao and H. Chen, “Multiple description coding with spatial-temporal hybrid interpolation for video streaming in peer-to-peer networks”, Journal of Zhejiang University of Science B, vol. 7, no. 5, pp. 894-899, April 2006.
- [2] Abdul H. Sadka, “Compressed Video Communications book”, John Willey & Sons Ltd, ISBN:0-470-84312-8, 2002.
- [3] J. G. Apostolopoulos, “Error-resilient video compression via multiple state streams” in Proc. of Int. Workshop on Very Low Bit rate Video Coding, VLBV99, Kyoto, Japan, October 1999.
- [4] S. Ekmekci and T. Sikora, “Multi-state video coding with side information“, In Proc. of Asilomer Conference on Signals, System and Computers, pp. 874-878, Oct. 28-Nov-1st Dec. 1 2005.
- [5] S. Dogan, S. Eminsoy, A. Sadka, A. Kondozi, “Video Content Adaptation using Transcoding for enabling UMA over UMTS.
- [6] A. Vetro, C. Christopoulos, H. Sun, “Video Transcoding Architectures and Techniques: An Overview, In IEEE Sig.
- [7] A. Norkin, A. Aksay, C. Bilen, G. Bozdagi Akar, A. Gotchev and J. Astola, “Schemes for multiple description coding for stereoscopic video”, in Proc. of LNCS, multi. Content, Representation and security, vol. 4105, pp. 730-737, Istanbul, Turkey, Sept. 2006.
- [8] A. Umar, “Reliable Delivery of Multimedia Services Via Satellite”, MSc Thesis, Univ. of Surrey, Guildford, UK, 2004.
- [9] G. Zhang, “Robust Scalable Video Compression using Multiple Description Coding”, PhD thesis, graduate program in Electrical Engineering, Notre Dame, Indiana, April 2007.
- [10] Y. Lee, J. Kim, Y. Altunbasak, R. Mersereau, “Layered coded vs. Multiple Description Coded over error-prone networks”, Elsevier Journal on Signal Processing: pp.337-356. November 2002. www.elsevier.com/locate/image . Issue 18.

Chapter 2: Review of Wireless Technology

2.1 Introduction

The basic concept of a wireless communication system is easy to understand. An electromagnetic signal or wave is generated, undergoes a process of modulation, amplifications and broadcast to one or more receivers that can be fixed or mobile. The data in that signal is received and demodulation process takes place in order to recover the original information that was initially sent through the wireless channel [1]. Figure 2 shows a basic system that normally consists of a transmitter, receiver and a radio frequency channel that utilizes different carrier frequencies for each baseband information signal that is transmitted. The basic issues that one must address in the design of wireless systems common to all of telecommunication networks are the effective use of the available frequency spectrum and power to provide high-quality communications. Some wireless systems often involve mobile services; this implies a constantly changing environment with rapidly changing interference conditions and dynamically variable multipath reflections [1]. This condition, plus the potential of conflicting demands for the use of radio frequencies in a free space medium, means difficult challenges for generating high quality signals.

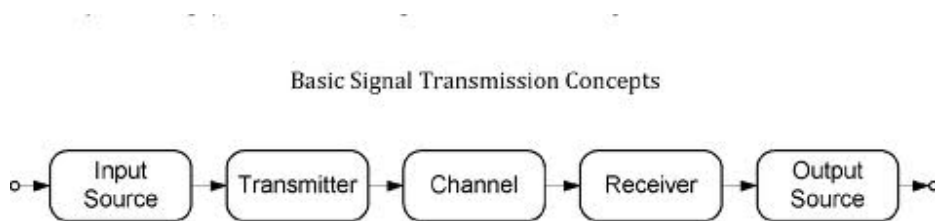


Figure 2-1: Basic Wireless Concepts [6].

Every wireless system must have the basics of transmitter, receiver and a channel to be able to transmit the signal from a stationary or mobile reference. Modulation is the process of varying one or more properties of a high-frequency periodic waveform, called the carrier signal, with a modulating signal which typically contains information to be transmitted. The three key parameters of a periodic waveform are its amplitude, its frequency and its phase. Demodulation

Review of Wireless Technology

which is also called detection is the inverse of modulation; it's the act of extracting the original information-bearing signal from a modulated carrier wave [2]. Its purpose is to restore the original baseband signal. The transmitter and receiver are very important parts of any radio communication link and at most base stations are separate components or they can be combined into a transceiver which transmit and receive signals simultaneously.

Transmitter is to generate a modulated signal with sufficient power, at the right frequency and to couple that signal into an antenna feed line. The modulation must be done in such a way that the demodulation process at the receiver can yield an error free copy of the original signal. Differences among transmitters result from variations in the required power level, carrier frequency and modulation type as well as some requirements such as portability and the ability to be controlled through a wireless device. Receiver performs an inverse function of transmitter. It must separate the desired signal from others present at the antenna, amplify and demodulated to recover the original baseband signal.

Radio waves are a form of electromagnetic radiation, as are infrared, visible light, ultraviolet light and gamma rays. The major difference is in the frequency of the waves. The part of the frequency spectrum that is useful for radio communication at present extends from 100 kHz to about 50GHz. Wireless communication mainly occupies the very high frequencies, ultra high frequencies and super high frequencies portions of the spectrum (see Table 1 below). Lower frequency systems need large or very large antennas and involve methods of signal propagation that are undesirable for wireless systems. Extremely high frequencies are still difficult to generate and amplify at reasonable cost.

Table 2-1: Frequency bands for Wireless Communication

1	Low Frequency	30-300KHz
2	High Frequency	3-30MHz
3	Very High Frequency (VHF)	30-300MHz
4	Ultra High Frequency (UHF)	300-3000 MHz
5	Super High Frequency (SHF)	3-30 GHz
6	Extremely High Frequency (EHF)	30-300 GHz

Review of Wireless Technology

The current uses of radio frequencies for wireless telecommunications include conventional fixed and mobile, direct broadcast services, conventional radio and TV broadcasting services, wireless cable TV, specialized mobile radio services, analogue and digital cellular radio services, Paging services, wireless Private Automated Brand Exchange and Local Area Network services etc. Sometimes these services have separate frequency allocations from each other and sometimes they compete with each other or with other public telecommunications, industrial or scientific applications.

The rapid growth of wireless services has the effect of creating rapidly growing demands on limited resources (Frequency bands). Furthermore, the desire to use frequencies in the Ultra High Frequency Band (UHF) has served to compound this problem. The physical characteristics of the radio spectra in terms of effective wavelengths have concentrated demand for most wireless services between the High Frequency (HF) bands and Ultra High Frequency (UHF) band. The next two bands in the Super High Frequency (SHF) and Extremely High Frequency (EHF) bands are presently being used for wireless services. These radio frequency bands are the most commonly used for wireless communications today.

The frequencies above 3000 MHz are most typically used for satellite communications and terrestrial microwave. These higher frequencies are subject to precipitation and attenuation losses, atmospheric heat scintillations, and propagation distortions. They are also unforgiving of any interruptions in direct line of sight (LOS) connections in the transmission path. Frequencies below the high frequency (HF) range are, on the other hand, very limited in bandwidth. They thus offer limited utility for many future services, especially broadband services, TV and high definition TV. Even with highly innovative frequency reuse concepts and digital compression, it seems unlikely that broadband applications for the VHF band or below would develop in the future simply because of the very limited spectra available. These conditions have combined to create a major problem in obtaining sufficient frequency allocations that are interference free or at acceptable interference levels.

2.2 Evolution from 1st to 4th generation networks

Tremendous changes are taking place in the area of mobile technologies, and the worldwide push towards 3rd generation services is currently at the forefront of these transformations [3]. Many questions surround the concept of 3G – not only in terms of what it means and what services it

Review of Wireless Technology

will offer, but also in terms of how to get there, which standard will be dominant, how long will it take to deploy, and whether it will be as lucrative as expected given the current rush of exorbitant spectrum fees. The successful development and deployment of GSM over the past two decades is most significant, if one is to accept the hypothesis that ‘experience counts’ in the mobile area. 3rd generation mobile technologies must, after all, in some way be the result of an evolution from pre-existing 2G systems, whether this is because they are developed from overlays on 2nd generation systems, or because operators deploying them must leverage pre-established 2G infrastructure or customer bases. The two are in many ways linked, and therefore examining one necessarily implies looking at the successes/shortcomings of the other.

The idea of cell-based mobile radio systems appeared at Bell Laboratories in the United States in the early 1970s [4]. However, mobile cellular systems were not introduced for commercial use until a decade later. Today, cellular systems still represent one of the fastest growing telecommunications systems. During development, numerous problems arose as each country developed its own system, producing equipment limited to operate only within the boundaries of respective countries, thus limiting the markets in which services could be sold.

First-generation cellular networks, the primary focus of the communications industry in the early 1980’s, were characterized by a few compatible systems that were designed to provide purely local cellular solutions. It became increasingly apparent that there would be an escalating demand for a technology that could facilitate flexible and reliable mobile communications. By the early 1990’s, the lack of capacity of these existing networks emerged as a core challenge to keeping up with market demand. The first mobile wireless phones utilized analogue transmission technologies, the dominant analogue standard being known as “AMPS”, (Advanced Mobile Phone System). Analogue standards operated on bands of spectrum with a lower frequency and greater wavelength than subsequent standards, providing a significant signal range per cell along with a high probability for interference [5].

Second-generation systems have proven to offer many advantages over analogue systems, including efficient use of radio spectrum, enhanced security, extended battery life, and video transmission capabilities. There are four main standards for 2G networks: Time Division Multiple Access (TDMA), Global System for Mobile Communications (GSM) and Code Division Multiple Access (CDMA); there is also Personal Digital Cellular (PDC), which is used

Review of Wireless Technology

exclusively in Japan [5][6]. In the meantime, a variety of 2.5G standards have been developed. “Going digital” has led to the emergence of several major 2G mobile wireless systems [5].

Third Generation, or 3G wireless networks are mobile telecommunication networks capable of transmitting and receiving voice and video information between mobile phones, landlines and the Internet at high speeds. The technology first launched with Verizon’s 3G network in early 2002. The 3G standard is being followed by the 4G standard featuring Long Term Evolution Technology (LTE) technology.

There are two standards of 3G technology in use: HSDPA/UMTS and Evolution Data Optimised (EVDO). HSUPA/UMTS allows for simultaneous voice and data transfer. This allows users to perform data tasks such as sending email or web browsing while also making voice calls. EVDO networks also allow users to transfer data and voice, but not simultaneously. Transmission speeds for 3G networks vary depending on location, device and carrier. Generally they are between 600 Kbps and 1.4 Mbps. As a point of reference, a device could transfer an average-sized MP3 file in approximately four seconds at 1.4 Mbps. These speeds are possible with mobile devices using EVDO. In addition to 3G capable cell mobile phones, many networks offer 3G computer expansions, allowing laptop users to access the Internet anywhere where there is network coverage. Devices that require wireless data and video transfer also use 3G technology. Classification of a network as 3G is based on the international Telecommunication Union’s (ITU) IMT-2000 standard. The standard is a framework set in place for developers to utilize the spectrum between 400 MHz and 3 GHz in a more concerted manner. Establishing a universal 3G standard reduces the differences between individual 3G networks. This reduces the challenges to global roaming by making it far easier to build a mobile phone that can work on any 3G network [7].

2.3 WiMAX

WiMAX (Worldwide Interoperability for Microwave Access) is a family of telecommunication protocols that provides fixed and mobile internet access [8]. The current Wimax revision provide upto 40 Mbps with the IEEE802.16m standard expected to offer up to 1 Gbps for fixed stations. The name “WiMAX” was created by the WiMAX Forum, which was formed in June 2001 to promote conformity and interoperability of the standard. The forum describes WiMAX as a

Review of Wireless Technology

standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and Digital Subscriber Line (DSL). WiMAX refers to interoperability implementation of the IEEE 802.16 wireless networks in similarity with Wi-Fi, which refers to interoperability implementations of the IEEE 802.11 Wireless LAN standard ratified by the Wi-Fi Alliance [9]. The WiMAX certification allows vendors to sell their equipment as WiMAX fixed or mobile certified, thus ensuring a level of interoperability with other certified products, as long as they fit the same profile. Figure 2-2 below shows the WiMAX coverage.

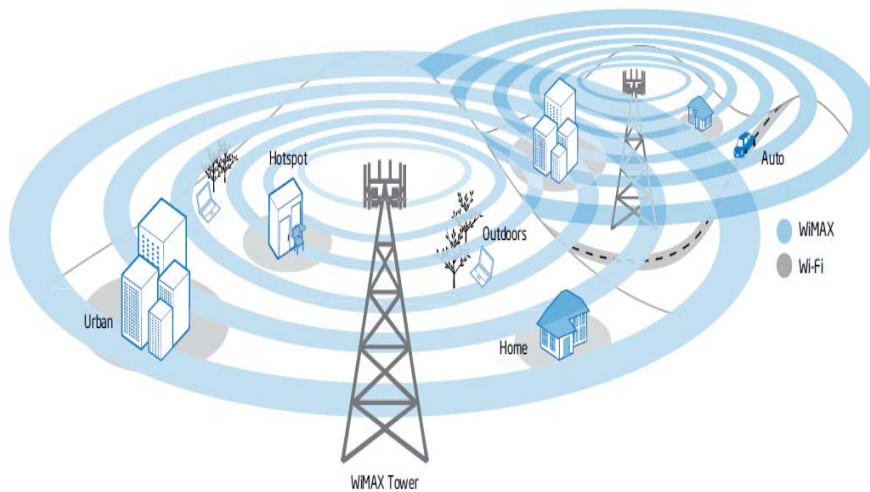


Figure 2-2: WiMAX coverage with broadband internet [8]

The IEEE 802.16 standards form the basis of WiMAX and is referred to Fixed WiMAX (802.16d) and Mobile WiMAX (802.16e) that is.

WiMAX can be used for a number of applications including broadband connections, cellular backhaul, hotspots etc. It is similar to Wi-Fi but it can also provide coverage at much greater distances. WiMAX is more effective on a large scale and it is more cost effective.

2.3.1 WiMAX Architecture

The WiMAX architecture defines how a WiMAX network can be connected to an IP based core network, which is typically chosen by network operators to serve Internet Service Providers (ISP). The WiMAX base station (BS) provides seamless integration capabilities with other types of network as with packet switched Mobile Networks. Figure 2-3 shows the WiMAX architecture

Review of Wireless Technology

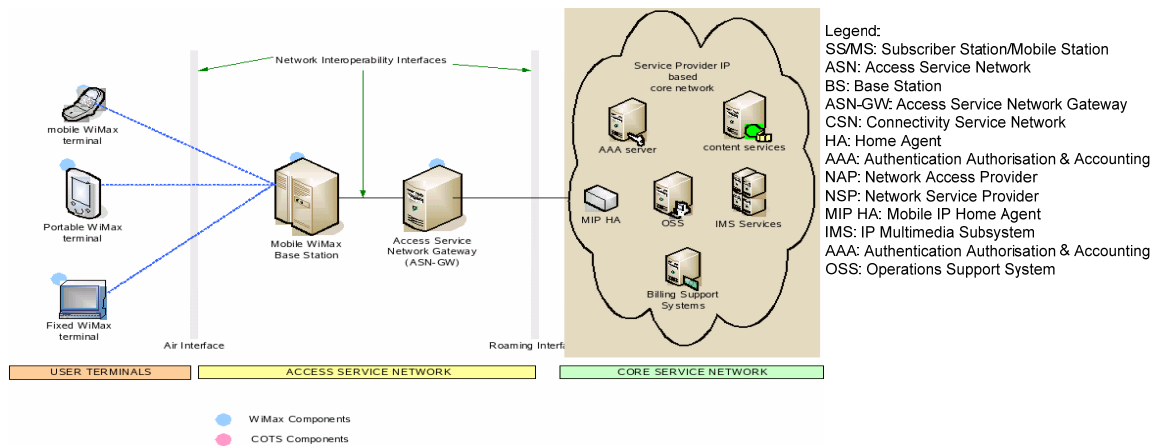


Figure 2-3: Architecture of WiMAX [8]

2.4 Wi-Fi

Wi-Fi (Wireless Fidelity) enabled devices such as a personal computer, video game console, smartphone, or digital audio player can connect to the Internet when it is within range of a wireless network connected to the Internet [11]. The coverage of one or more access points called hotspots when offering public access, generally comprises of an area the size of a few rooms but may be expanded to cover larger distances, depending on the number of access points with overlapping coverage.

Wi-Fi is a wireless local area network (WLAN) technology based on IEEE 802.11 standards, for device to device connectivity such as Wi-Fi Peer to Peer, Wi-Fi direct and a range of others [11]. Wi-Fi has been used interchangeably with IEEE 802.11. Wi-Fi is used by over 700 million people; there are over 750,000 hotspots with Wi-Fi internet connectivity around the world, and about 800 million new Wi-Fi devices every year.

A Wi-Fi enabled device can connect to the Internet within range of a wireless network as shown in Figure 5. In addition to private use in homes and offices, Wi-Fi hotspots provide access either free-of-charge or to subscribers of various commercial services. Organizations and businesses often provide free use hotspots to attract or help clients. In addition to the use of digital subscriber line modem, or Wi-Fi access point for connectivity to the Internet, the emergence of Mi-Fi and WiBro (portable Wi-Fi router) allows subscriber to easily create their own Wi-Fi hotspots and connect to Internet via Cellular networks such as iPhone.

Review of Wireless Technology

Wi-Fi technology allows the deployment of LAN without wires for client devices, typically reducing the costs of network deployment and expansion. Places where cables cannot be run, can host wireless LANs. Figure 2-4 showed the Wi-Fi architecture.

2.4.1 Wi-Fi Architecture

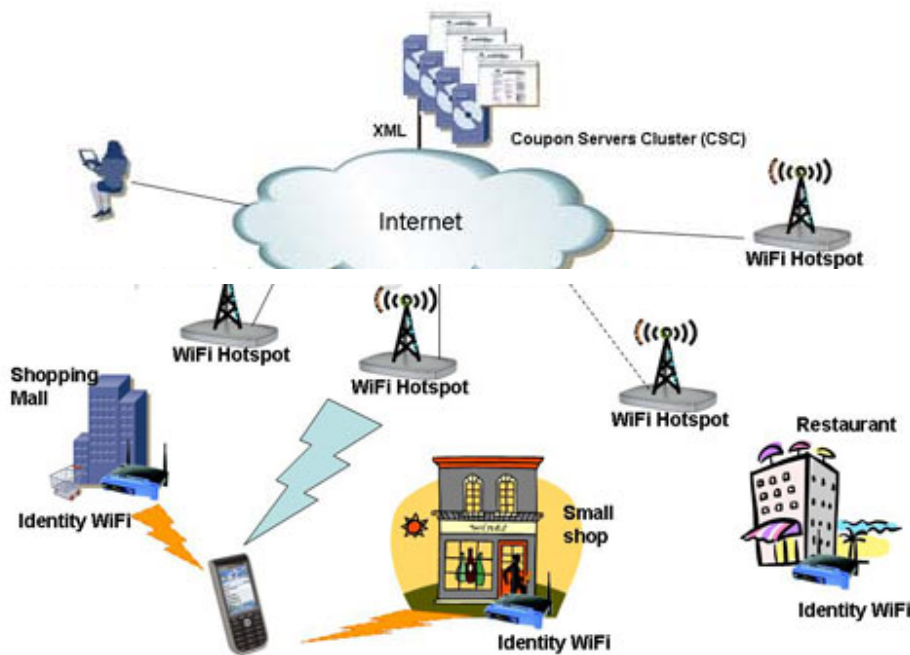


Figure 2-4: Architecture of Wi-Fi [11]

2.5 4G Network

Figure 2-5 show the evolution from 2G to 4G network in terms of stakeholder's involvements and business interactions. 4G is the fourth generation of cellular wireless standards otherwise known as Long Term Evolution (LTE), It is the latest step in moving forward from the cellular services such as GSM, UMTS, HSPA, LTE, and CDMA [11]. LTE is based on the standard developed by the 3rd Generation Partnership Project (3GPP) [12]. LTE may also be referred more formally as Evolved UMTS Terrestrial Radio Access (E-UTRA) and Evolved UMTS Terrestrial Radio Access Network (E-UTRAN). It is a successor to the 3G and 2G standards. In 2008, the ITU-R specified the IMT-Advanced (International Mobile Telecommunications Advanced) requirements for 4G standards, setting peak speed requirements for 4G services at 100 Mbps for high mobility communication such as trains and cars and 1 Gbps for low mobility communication such as pedestrians and stationary users.

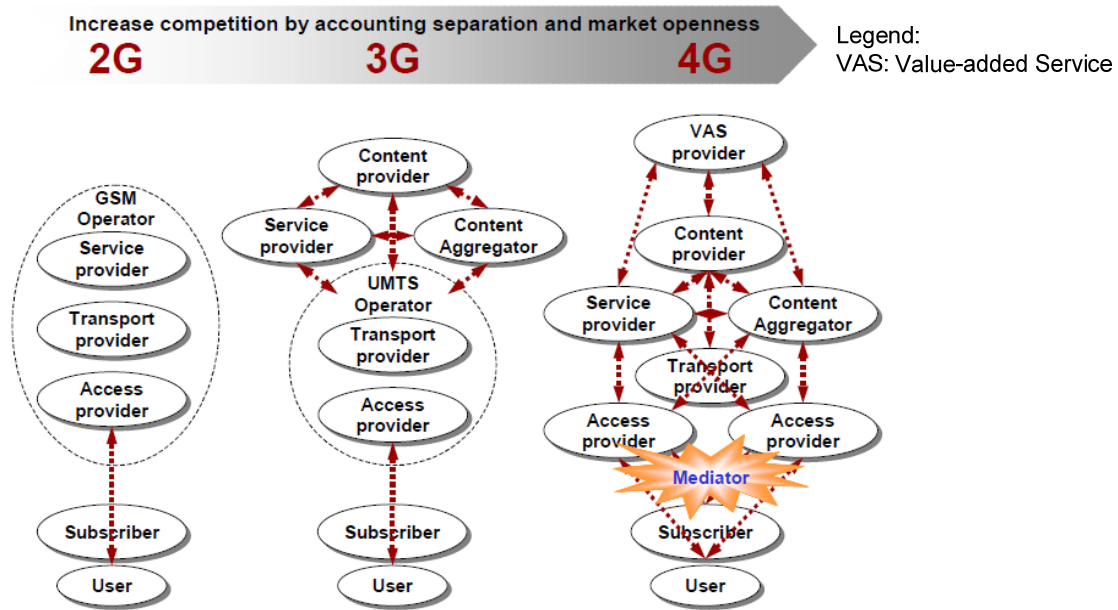


Figure 2-5: Evolution to 4G network [7]

A 4G system is expected to provide a comprehensive and secure all IP based mobile broadband solution to laptop computers, smart phones and other mobile devices. Services such as ultra-broadband Internet access, IP telephony, gaming services, and streamed multimedia may be provided to users.

Before 4G, technologies such as mobile WiMAX and 3G Long term evolution (LTE) have been on the market since 2006 and 2009 respectively, and are often branded as 4G. The current versions of these technologies did not fulfil the original ITU-R requirements of data rates approximately up to 1Gbps for 4G systems [14]. IMT-Advanced compliant versions of the two standards are under development and called LTE Advanced and WirelessMan-Advanced respectively. ITU announced that current versions of LTE, WiMAX and other 3G technologies that do not fulfill IMT-Advanced requirements, could be considered 4G provided they represent forerunners to IMT-Advanced with a substantial level of improvement in performance and capabilities with respect to the initial third generation systems that have already been deployed.

2.5.1 4G Requirements

4G or LTE (IMT-Advanced) systems must fulfill the following requirements:

- All-IP packet switched network

Review of Wireless Technology

- Peak data rates of up to approximately 100 Mbps for high mobility and up to approximately 1 Gbps for low mobility such as local wireless access.
- Dynamically share and utilize the network resources to support more simultaneous user's access per cell.
- Scalable channel bandwidth between 5 and 20 MHz, optionally up to 40 MHz.
- Peak link spectral efficiency of 15 bits/s/Hz in the downlink and 6.75 bits/s/Hz in the uplink.
- Ability to offer high quality of service for next generation multimedia support.
- High quality of services for 3D video transmission.

2.5.2 4G Network Architecture

Figure 2-6 provides an illustration of the future 4G mobile network architecture comprising hot-spot, cellular, ad-hoc and satellite radio components (Heterogeneous). The figure sketches a heterogeneous network infrastructure comprising different wireless access systems such as Media Gateway, SGSN, AAA, Accounting and Billing etc. Considering that a single system that optimally meets a wide range of user cases and satisfies diverse service requirements is likely to remain an engineering utopia, we understand that heterogeneous architectures that can exploit individual system capabilities to optimally serve the instant application and value-added service mix in a flexible manner are a plausible design approach. The goal of the future mobile communication systems will be to incorporate and integrate different wireless access technologies and mobile network architectures in a complementary manner so as to achieve a seamless wireless access infrastructure.

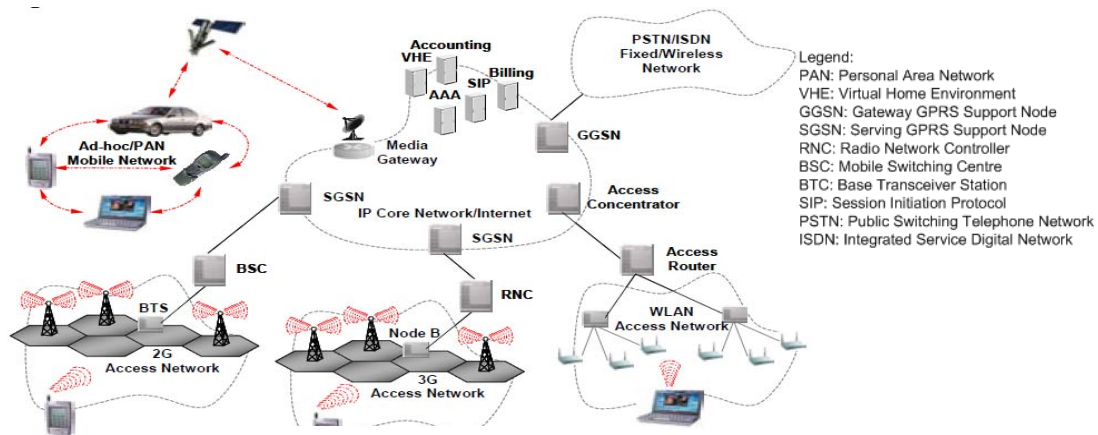


Figure 2-6: 4G Network Architecture [11]

2.6 Overview of UMTS Technology

Universal Mobile Telecommunications System (UMTS) is one of the third-generation (3G) mobile telecommunications technologies, which is also being developed into a 4G technology [15]. The first deployment of the UMTS is the release99 (R99) architecture. It is specified by 3GPP and is part of the global ITU IMT-2000 standard. The most common form of UMTS uses W-CDMA (IMT Direct Spread) as the underlying air interface but the system also covers TD-CDMA and TD-SCDMA (both IMT CDMA TDD) [15]. Being a complete network system, UMTS also covers the radio access network (UMTS Terrestrial Radio Access Network, or UTRAN) and the core network (Mobile Application Part), as well as authentication of users via SIM (Subscriber Identity Module) cards. Unlike EDGE (Enhanced data rates for GSM) and CDMA2000, UMTS requires new base stations and new frequency allocations. However, it is closely related to GSM/EDGE as it borrows and builds upon concepts from GSM. Furthermore, most UMTS handsets also support GSM, allowing seamless dual-mode operation. Therefore, UMTS is sometimes marketed as 3GSM, emphasizing the close relationship with GSM and differentiating it from competing technologies.

UMTS, using 3GPP, supports maximum theoretical data transfer rates of 45 Mbps with HSPA although at the moment users in deployed networks can expect a transfer rate of up to 384 Kbps (Uplink) and 7.2 Mbps for HSDA handsets in the downlink direction. This is still much greater than the 9.6 Kbps of a single GSM error-corrected circuit switched data channel or multiple 9.6 Kbps channels in High Speed Circuit Switch Data (HSCSD) and in competition to other network technologies such as CDMA2000, or WLAN- offers access to World Wide Web and other data services on mobile devices.

Prior to 3G are 2G mobile telephony systems such as GSM, CDMA, Digital AMPS and other 2G technologies deployed in different countries. In case of GSM, there is an evolution path from 2G to GPRS also known as 2.5G. GPRS supports a much better data rate up to a theoretical maximum of 140.8 Kbps, though typical rates are closer to 56 Kbps and is packet switched rather than connection oriented (circuit switched). It is deployed in many places where GSM is used. E-GPRS or EDGE is a further evolution of GPRS and is based on more modern coding schemes. With EDGE, the actual packet data rates can reach around 180 Kbps. EDGE systems are often referred to 2.75G Systems.

Review of Wireless Technology

Since 2006, UMTS networks in many countries have been or are in the process of being upgraded with High Speed Downlink Packet Access (HSDPA), sometimes known as 3.5G. Currently, HSDPA enables downlink transfer speeds of up to 21 Mbps. Work is also progressing on improving the uplink transfer speed with the High Speed Uplink Packet Access (HSUPA). For long term, the 3GPP Long Term Evolution project plans to move UMTS to 4G speeds of 100 Mbps downlink and 50 Mbps uplink, using the next generation air interface technology based on Orthogonal Frequency-Division Multiplexing (OFDMA) [16].

The first UK national consumer UMTS networks launched in 2002 with a heavy emphasis on mobile applications such as mobile TV and video. The high data speeds of UMTS are now most often utilized for Internet access. Experience in some other countries has shown that user demand for Video call is not high, and mobile audio/video content and high-speed access to the World Wide Web (WWW) has gained in popularity, either directly connected on a handset or connected to a computer via Wi-Fi, Bluetooth, Infrared or USB [17].

2.6.1 UMTS Architecture

Fundamental difference between GSM/GPRS and UMTS is the need for the later to support high bit rate services plus the notion of negotiated QoS and traffic characteristics [16]. Figure 2-7 shows the UMTS architecture. UMTS needs to support bursty traffic in an efficient way and to allow support of single and multimedia N-ISDN applications and single & multimedia IP applications. UMTS system is designed to be flexible, modular with network nodes defined that implement some specific functionality and open interfaces defined between such nodes. The aim is for an architecture which will minimize signaling traffic and optimize transmission infrastructure [17]. The architecture also needs to protect existing investments which operators have, and re-use as many elements of these as possible. The first release of UMTS builds directly upon an evolved GSM network including the addition of GPRS.

Review of Wireless Technology

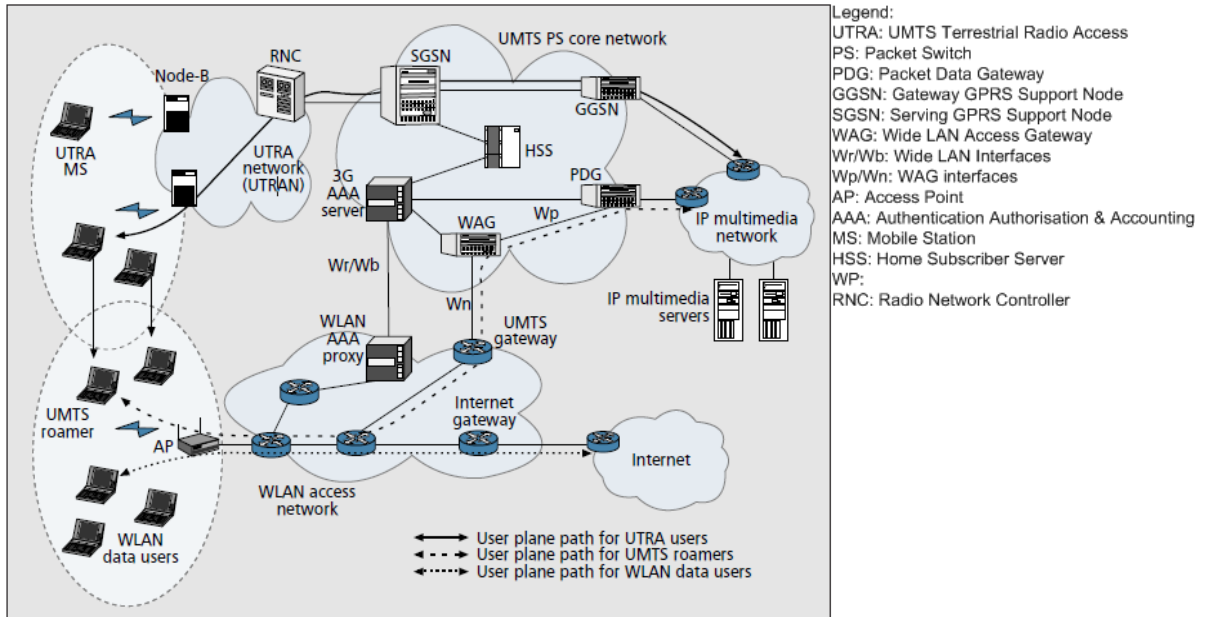


Figure 2-7: UMTS Network Architecture [16]

The UMTS architecture can be divided into the following domains as shown in figure 2-8:

- User
- Infrastructure which itself subdivided into Radio Access Network and Core Network.

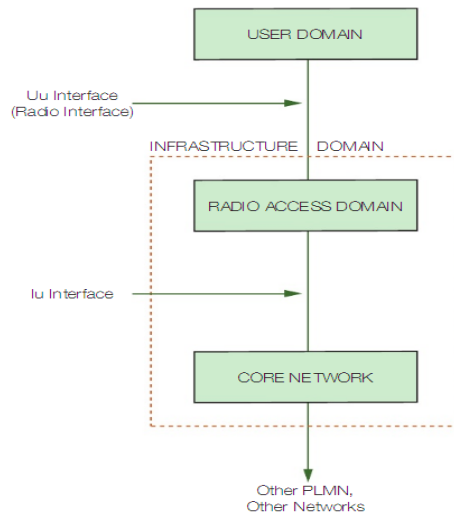


Figure 2-8: UMTS Domains

2.6.2 The User

The user describes the equipment needed by the user to access UMTS services. Within this we have Mobile Equipment (ME) and UMTS Subscriber Identity Module (USIM). The combined ME and USIM are called Mobile Stations (MS). The USIM contains the data and procedures allowing the ME to securely identify itself and is linked to the ME by the defined Cu interface. The ME performs radio transmission and contains applications. It may itself be further subdivided into the Mobile Terminal (MT) and Terminal Equipment (TE). The functionality of the MT is new in UMTS, in being able to interact with the access network over all UMTS radio interface (U_u). The mobile terminal must also be able to receive and transmit between both GSM-based and UMTS-based radio access. Figure 2-9 shows breakdown of what can be found in user domain.

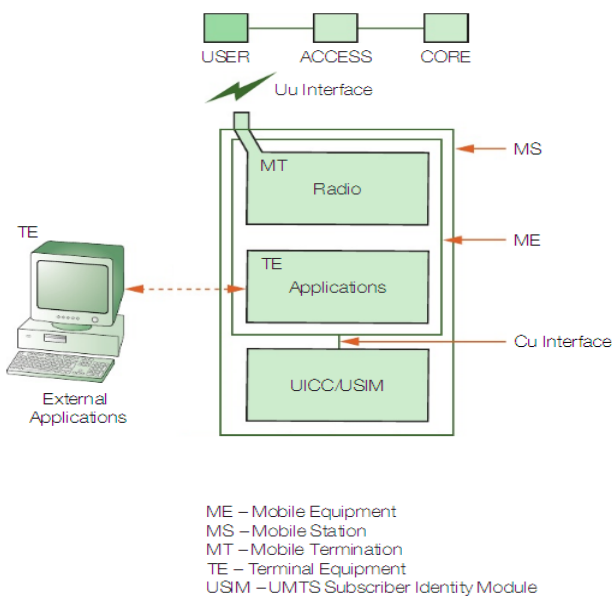


Figure 2-9: UMTS User Domain

2.6.3 Access Network

The Radio access domain is in direct contact with the User Equipment and the core network. This split is intended to separate access functionality from non-access functionality. The Access contains the physical entities to manage resources of the access network and provide the user with a way to access the core network. In UMTS, the Access domain refers to the radio access mechanism and is also known as the UTRAN (UMTS Terrestrial Access Network). All access

Review of Wireless Technology

networks will require use of the USIM. In the early roll-out of UMTS, it remains relevant to also include the GSM/EDGE radio network as an alternative within the access domain, since internetworking is required until UTRAN coverage is fully achieved. The UTRAN is connected via another standardized, open interface, the Iu to the Core Network Domain. Figure 2-10 shows what can be found in radio access network.

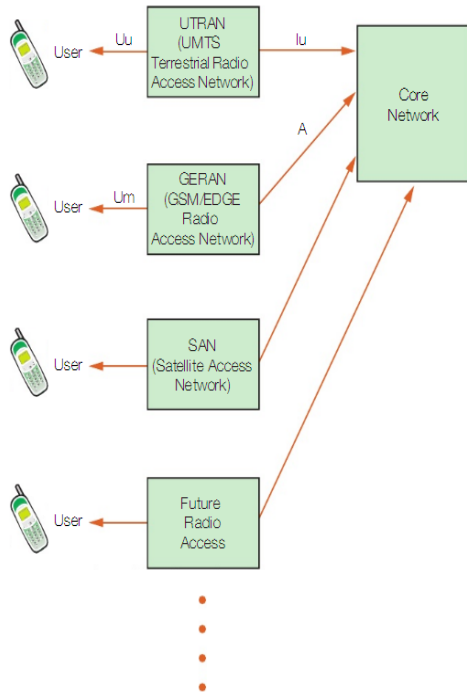


Figure 2-10: UMTS Access Network and Interfaces

The UTRAN is a separate entity to the core network, with a defined interface connecting them. UTRAN functions have been specified to provide support for all radio activities needed within the network infrastructure. These are divided into four parts, System Access, Mobility, Radio Channel Ciphering and Radio Resource Management and Control. This interface is designed to provide a logical separation of signaling and user data transport. All radio procedures are fully handled within the UTRAN including mobility. Figure 2-11 shows the UTRAN architecture connection to Core Network.

Review of Wireless Technology

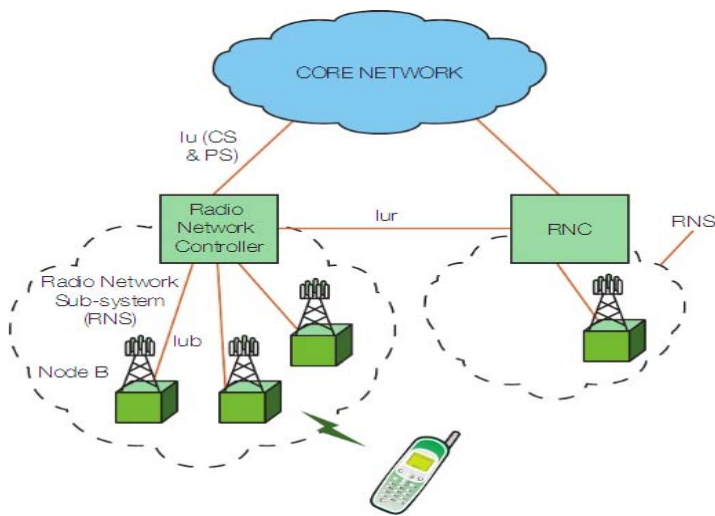


Figure 2-11: UTRAN Architecture [19]

Some of the most important requirements for UTRAN are:

- Logical Separation of Signalling and Data Transport
- CN and UTRAN functions are separate from Transport Functionality
- Macro diversity fully handled in UTRAN
- Mobility for Radio Resource Control (RRC) connection is fully controlled by UTRAN
- Interfaces based on logical model of the entities

The UTRAN architecture comprises of one or more Radio Network Controllers (RNCs) as shown in figure 12 above. Each group of RNC and its associated Node Bs are called Radio Network Sub-system (RNS). Standard interfaces connect each RNS to the Core Network and to the User Equipment. These interfaces are called IuCS, IuPS and Uu respectively.

2.6.4 The Node B

The term Node B refers to the base station equipment which communicates with the subscriber's handset via the radio link and also with the main network. It provides radio resources for a UMTS network, and uses UMTS channel allocation to communicate with the handset. It provides all the RF processing, enabling transmission and reception information to and from the mobile terminal. This information is encoded using the wide-code division multiple access (W-CDMA) scheme. A single UMTS channel can be used on adjacent Node B sites and in different sectors of the same Node B antenna system. A typical Node B may support a three sector

antenna and one or two UMTS carriers, although it is possible to configure up to six sectors and up to three UMTS carriers. Each sector can be used as a different cell.

2.6.5 The Radio Network Controller (RNC)

The RNC controls the operation of multiple Nodes Bs, managing resources such as allocating capacity for data calls and providing critical signaling such as connection set up, switching and traffic routing. Figure 2-12 show how RNC control multiple nodes. It enables autonomous Radio Resource Management by the UTRAN by allowing RNCs to directly communicate via the Iur interface.

Key features of the RNC are:

- Management of radio resources
- Channelization code allocation
- QoS monitoring
- Handover of users between cells on the same site
- Handover of users between cells on different sites (Soft handover).
- Handover of users between different UMTS carriers (Hard handover)
- Handover of users to GSM networks
- Power control management of user and Node B equipment.

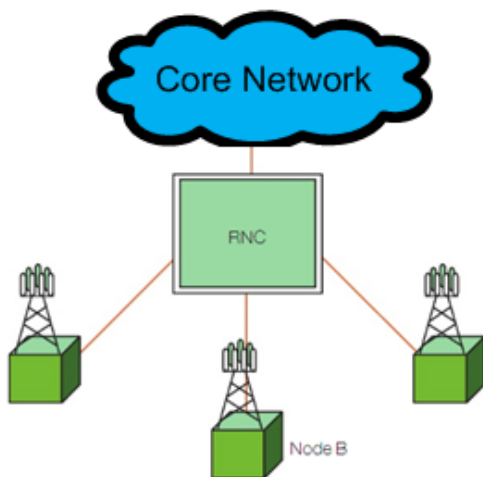


Figure 2-12: UMTS RNC and Node Bs

2.6.6 UTRAN Protocols

Figure 2-13 shows UTRAN protocols and their functionality. Radio Access Network Application Part (RANAP) includes those functions needed to manage location procedures which may need Core Network interaction such as hard handover and Service RNC (SRNC) relocation, Radio access bearer management, security, paging, Drift RNC (DRNC), identity management and transparent transfer.

Radio Network Subsystem Application Part (RNSAP) provides functions which are split into four modules. Basic inter RNC mobility is supported in order to provide soft handover between RNC and to transfer waiting data during SRNC relocations. Also, support is provided for both dedicated channel traffic and common channel traffic.

Node B application part (NBAP) functions are classed as either common or dedicated, depending on whether they are concerned with common or dedicated channels. The functions are concerned with the use or configuration of the radio channels including paging, access requests, radio link, handovers and fault management. Common channels continuously broadcast system identification and access control information, while dedicated channels is used to transfer user data.

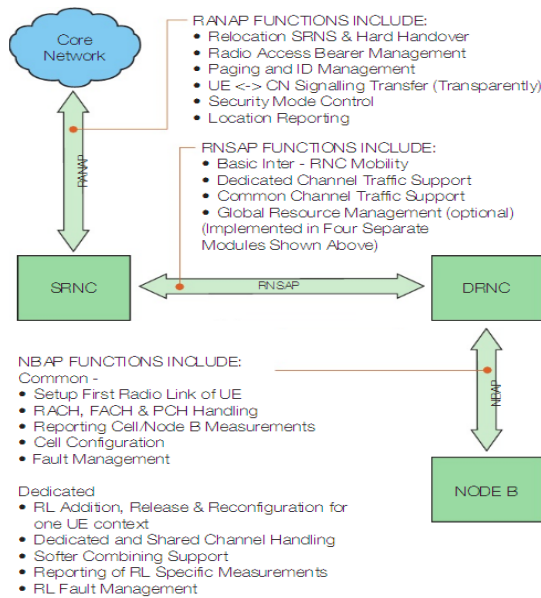


Figure 2-13: UTRAN Protocol functions

2.6.7 UTRAN Transport Network

The UTRAN provides the User Equipment (UE) with access to the Core Network (CN) for Circuit Switched and Packet Switched services as well as providing transport for all signaling interactions, including those confined within the UTRAN, those between the UTRAN and the Core Network and those being transferred through the UTRAN from UE to CN. The W-CDMA air interface has been designed to support services which vary widely in terms of acceptable quality of service (QoS). Hence services with varying data rates, delay tolerance, delay variance and acceptable error rates are all possible. Figure 2-14 shows the UTRAN transport network.

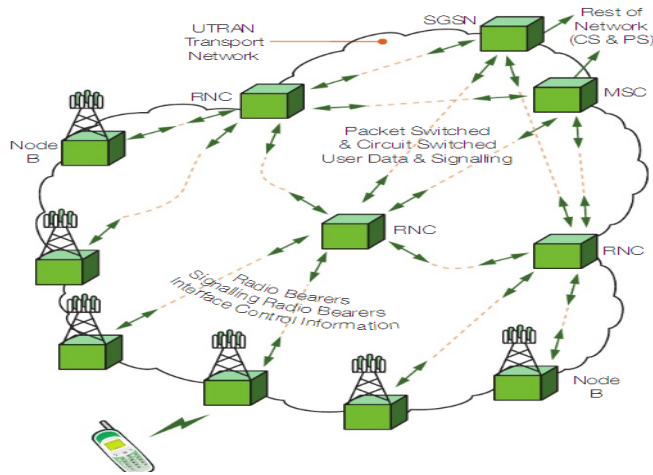


Figure 2-14: UTRAN Transport Network [18]

2.6.8 UMTS Core Network

The core network contains the physical entities providing the network features and services. For example, the management of user location, control of network services and switching & transmission mechanisms for signaling & user information. This is divided into circuit-switched and packet-switched networks [18][19].

- CS (circuit-switched) refers to all core network entities offering “CS connection” for which dedicated network resources are allocated at connection establishment and held until connection release. PSTN and ISDN are examples of other circuit-switched networks. In UMTS, the CS domain provides data services support of at least 64 Kbps.

Review of Wireless Technology

- PS (Packet-switched) refers to all core network entities offering “ PS connection” for which transport user information using autonomous concatenation of bits called packets, where each packet can be routed independently of the previous one. The Internet is the most well known example of a packet-switched network although other public data networks (PDNs) do exist. In UMTS, the PS provides support for data service capability of up to 2 Mbps.

The lu interface is therefore subdivided into luCS and luPS in order to support connection of each of this core network to the single access network (UTRAN). Figure 2-15 shows the core network of UTRAN.

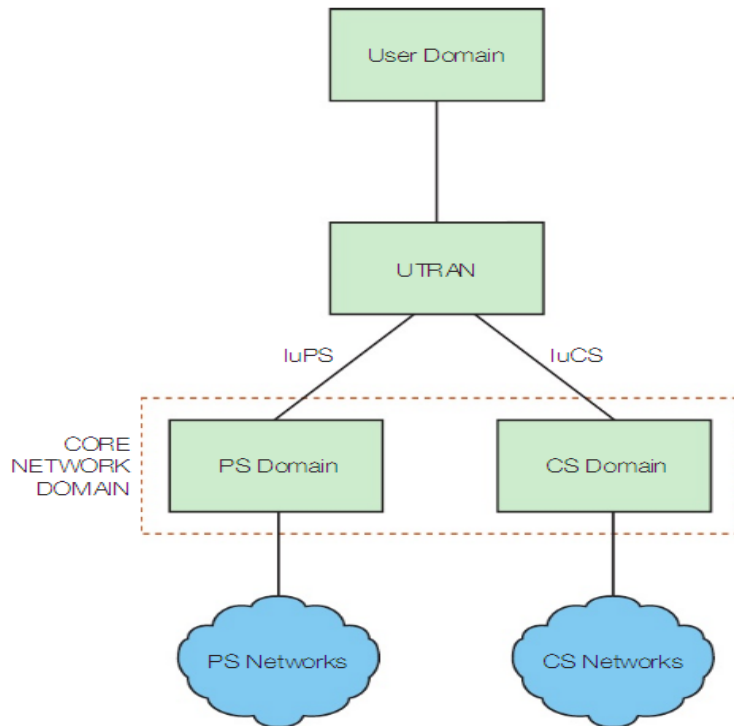


Figure 2-15: UMTS Core Network domain

Three categories of core network element are defined in UMTS release ‘99’. These are:

- GSM core network elements: Mobile Switching Centre (MSC), Visitor Location Register (VLR), Home Location Register (HLR), AuC interface and EIR (Equipment Identity Register).
- GSM enhancement (GSM) Phase 2+.
- GPRS to support packet switching.

- Customised Application for Mobile Networks Enhanced Logic (CAMEL) as a basis for the virtual home environment (VHE).

2.7 UMTS Services

The best known feature of UMTS is offering higher bit rates such as 384 kbps on circuit-switched connections and up to 2 Mbps on packet-switched connections [20]. Higher bit rates facilitate some new services such as video telephony, faster download of videos and data, Mobile TV, video on demand, Video conferencing, Telemedicine, location-based services etc. Figure 2-16 shows some of the current UMTS services and applications [21].

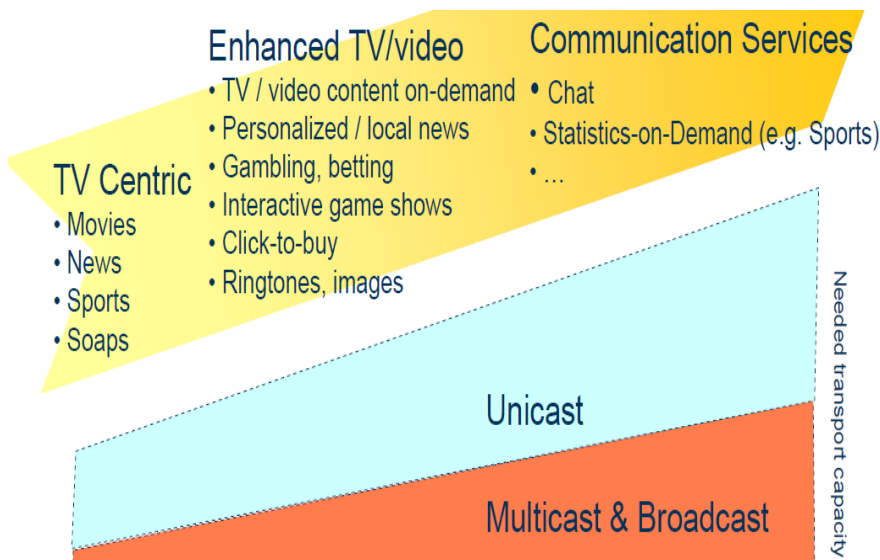


Figure 2-16: UMTS Services [21]

2.7.1 UMTS MBMS Services

Figure 2-17 shows some multimedia broadcast and multicast services (MBMS) for TV and on-demand services.

Review of Wireless Technology



Figure 2-17: MBMS in UMTS network [21]

2.7.2 UMTS Bearer Services

Bearer services provide the capability for information transfer between access points and involve only lower layer functions. These functions are sometimes referred to as lower layer capabilities in reference to OSI layers [22]. The user may choose any set of higher layer protocols for his communication and the public land mobile network (PLMN) does not ascertain compatibility at these layers between users. In the general case, a communication link between access points provides a general service for information transport. The communication link may span over different networks such as Internet, Intranet, LANs and ATM based networks, having network specific means for bearer control. Each network contributes to the end-to-end QoS perceived by the end user. Bearer services in UMTS are negotiable unlike GSM where they are not and can be used flexibly by applications. Figure 2-18 below shows the architecture of a UMTS radio access bearer services.

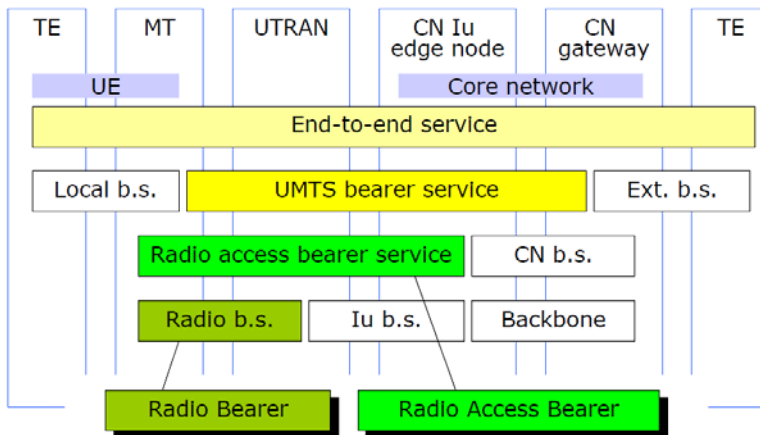


Figure 2-18: UMTS bearer service architecture [22]

Bearer services in UMTS are characterized by a set of end-to-end characteristics with requirements on QoS [23]. The characteristics and requirements shall cover major network scenarios, i.e the cases when the terminating network is PSTN, ISDN, IP networks, LANs, X.25 and a PLMN. QoS is the quality of the services that has been requested. The service characteristics include things like traffic type, supported bit rates and the quality of information. These parameters are negotiated when a connection is being established. If the network is unable to provide the required QoS, it will re-negotiate the QoS depending on what is available. After the connection has been established, if there is a need then these parameters can be negotiated again. The bearer services in UMTS are; information transfer (connection oriented/connectionless services), real time video, audio and speech (and non real time services), multimedia applications, non real time interactive and file transfers etc.

2.7.3 UMTS Teleservices

Teleservices provide the full capabilities for communications by means of terminal equipment, network functions and possibly functions provided by dedicated centres. Basically, it is a service that provides complete end-to-end capability for communication between mobile users according to standards. Teleservices contain both single media and multimedia services such as speech/telephony, Emergency calls, short message services and Internet access. Teleservices utilises the bearer services provided by lower layers. The Bearer Services and the Teleservices are not coupled to each other so as to aid independent development and changes to one may not necessarily mean changes to the other [24].

2.7.4 UMTS Supplementary Services (SS)

A supplementary service modifies or supplements a basic telecommunication service. Therefore, it cannot be offered to a user as a stand alone service. A stand alone service can be either Bearer Service or Teleservice but cannot be supplementary service. It shall be offered together or in association with a basic telecommunication service. The same supplementary service may be applicable to a number of basic telecommunication services. Also, one basic telecommunication service may use several SS simultaneously [24]. Some of the supplementary services are Call Forwarding, Call Deflection, Call Waiting, Call Hold, Call Restriction and Call barring and Number Identification.

2.8 UMTS QoS

UMTS bearer service attributes describes the service provided by the UMTS network to the user of the UMTS bearer service [25]. A set of QoS attributes or profiles specifies this service as shown in figure 2-20. At UMTS bearer service establishment or modification, different QoS profiles have to be taken into account. The User Equipment (UE) capabilities form a QoS which may limit the UMTS bearer service which can be provided. UE or the terminal equipment (TE) within the terminating network may request a QoS profile at UMTS bearer establishment or modification. The application using the UE may request the UE to provide a UMTS bearer service with a specific QoS profile. If the application requests no specific QoS, the UE may use a QoS profile configured within the UE. A QoS profile in the UMTS subscription describes the upper limits for the provided service if the service user requests specific values [26]. If the UE requests or modifies a UMTS bearer and one or more of the QoS attributes are not specified by the UE by setting the attribute to “subscribed”, the SGSN shall assume a request of values as specified in the QoS profile in the UMTS subscription. If the UE sets the traffic class to “subscribed”, the SGSN shall assume a request for Interactive class (see Figure 2-19). When the application in the UE requires streaming or conversational QoS, then the UE shall at least explicitly request the traffic class and should explicitly request the guaranteed bit rate and the maximum bit rate. For the rest of the QoS attributes, the network shall ensure that the negotiated QoS contains only values explicitly defined for the traffic class [26].

Review of Wireless Technology

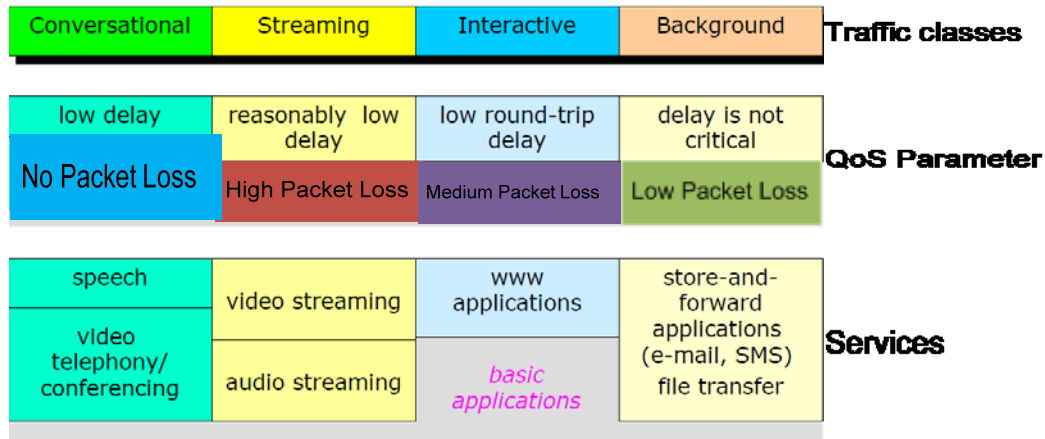


Figure 2-19: UMTS Traffic and Services classes [26]

2.8.1 Application/Services requirements to provide QoS for Videos

In packet switched network, while discussing QoS, there are two applications in respect of the network, one is the real-time and non real-time. Those applications deliver time-sensitive information are real-time, where the data blocks must be displayed consecutively at predetermined time intervals, thus require specific delay, jitter and error parameters. While non real-time applications don't include time-sensitive information which may be much more tolerable to delay and jitter but very sensitive to packet loss. Based on this, the applications in the proposed UMTS network are classified into four types based on the generated traffic as shown in figure 2-20. These are Conversational, Streaming, Interactive and Background.

Some well known QoS metrics for service/application requirements in UMTS networks are:

- Delay: It is the elapsed time for a video packet to traverse the network from the transmitter to the receiver.
- Jitter: It is defined as the variation in delay encountered by similar video packets following the same route through the UMTS network. The jitter requirement only affects real-time streaming applications because this QoS requirement arises from the continuous traffic characteristics of this class of applications. Jitter is included as a performance parameter since it is very important at the transport layer in packetized video data systems, due to the inherent variability in arrival times of individual video packets. Services intolerant of delay variation will usually try to reduce the delay variation by

Review of Wireless Technology

means of buffering. However, late data arrivals make data useless, resulting in receiver buffer underflow, and early arrival can lead to receiver overflow.

- **Packet Loss Rate:** Loss rate refers to the percentage of video data loss among all the delivered video data in a given transmission time interval, which can be evaluated in video frame level or packet level. Loss rate requirements apply to all classes of applications. Real-time applications might tolerate a limited amount of video data lost, depending on the error resiliency of the decoder, and the type of application. While, non real-time applications have much more strict requirement on video data loss.
- **Throughput:** It is defined as the rate at which video packets are transmitted in the UMTS network. It can be expressed as a maximum rate or an average rate.

Table 2-2: Error Tolerance among QoS classes

Service class	Conversational (delay << 1sec)	Interactive (delay << approx. 1 sec)	Streaming (delay < 10 secs.)	Background (delay > 10 secs.)
Error tolerant	Conversational, Voice and Video	Voice and Messaging	Streaming, Audio and Video	Fax
Error intolerant	Telnet, Interactive games	e-commerce, Web browsing	FTP, still images and Paging	e-mail arrival notification

Review of Wireless Technology

2.8.2 QoS at Various Levels

Table 2-3: QoS at Various Levels [26]

QoS at various levels	Definition/Characteristic/Parameter/Measurement
<i>User QoS Preferences</i>	Expressed by high-level terms e.g., Gold, silver, bronze (Olympic Services)
<i>Terminal chars & constraints</i>	Characterised by processing power, memory, jitter buffer size, codec, video playback frame rate, etc.
<i>Expected PQoS</i>	Defined as the perceived quality that a customer or end-user should expect from a contracted service described in cSLA taking into account all relevant resources and constraints.
<i>Delivered PQoS</i>	The quality that a customer or end-user actually perceives when consumes the service at his/her terminal measured by subjective/objective methods. It is expressed as Perfect, Excellent, Very Good, Good, Moderately Good, Fair, Somewhat Poor, Poor, Very Poor, Bad, or Useless. PQoS evaluated using: SNR, MOS, Structural Similarity Index (SSIM) techniques.
<i>Derived PQoS</i>	An approximation of PQoS delivered to a number of application streams deduced from <i>Measured NQoS</i> .
<i>ApQoS</i>	Defined as the application quality to express the application's needs and constraints in technical terms. ApQoS is characterised by encoding and transmission parameters such as frame rate, resolution, coding format, latency, jitter, loss ratio/BER, aspect ratio, etc.
<i>User Expected ApQoS</i>	Defined as the expected application quality the user should envisage which is obtained from the user's <i>Expected PQoS</i> via AQoS tool.
<i>AQoS (Adaptation QoS)</i>	AQoS is the means that provides the required metadata support for context-aware quality based adaptation/mapping expressing the relation between quality and resources, constraints, and adaptation operation.
<i>ApQoS Bound</i>	Specifies the boundaries that application can operate, based on the available application QoS profiles set by the application encoding and transmission requirement and the distortion that application can tolerate.
<i>Adopted ApQoS</i>	Is used to classify the applications in terms of their tolerable distortion and encoding and transmission quality requirements within <i>ApQoS Bound</i> .
<i>NQoS</i>	Generally defined as the quality targeted for (or experienced by) a network connectivity service and expressed by one or more performance parameters (i.e., one-way delay, round-trip delay, one-way packet loss, delay variation, offered load, throughput) that are quantified.
<i>Requested NQoS</i>	Defined as the networking quality needs of the application in terms of well-known metrics.
<i>Engineered NQoS</i>	Defined as the quality targeted for network connectivity services that an NP engineers its network to offer them to its customers.
<i>NQoS-class</i>	An <i>NQoS-class</i> is identified by a set of packet transfer performance parameters (attributes) associated with specific performance targets (values).
<i>Offered NQoS</i>	The quality actually set by the provider, assigned to <i>NQoS-classes</i> and deemed appropriate for creating competitive network connectivity service offerings to its customers or providers.
<i>Measured NQoS</i>	Defined as the network quality actually experienced by QoS-based connectivity services offered by one/more NPs and deduced by actual measurement during network operation.

This is shown in figure 2-20.

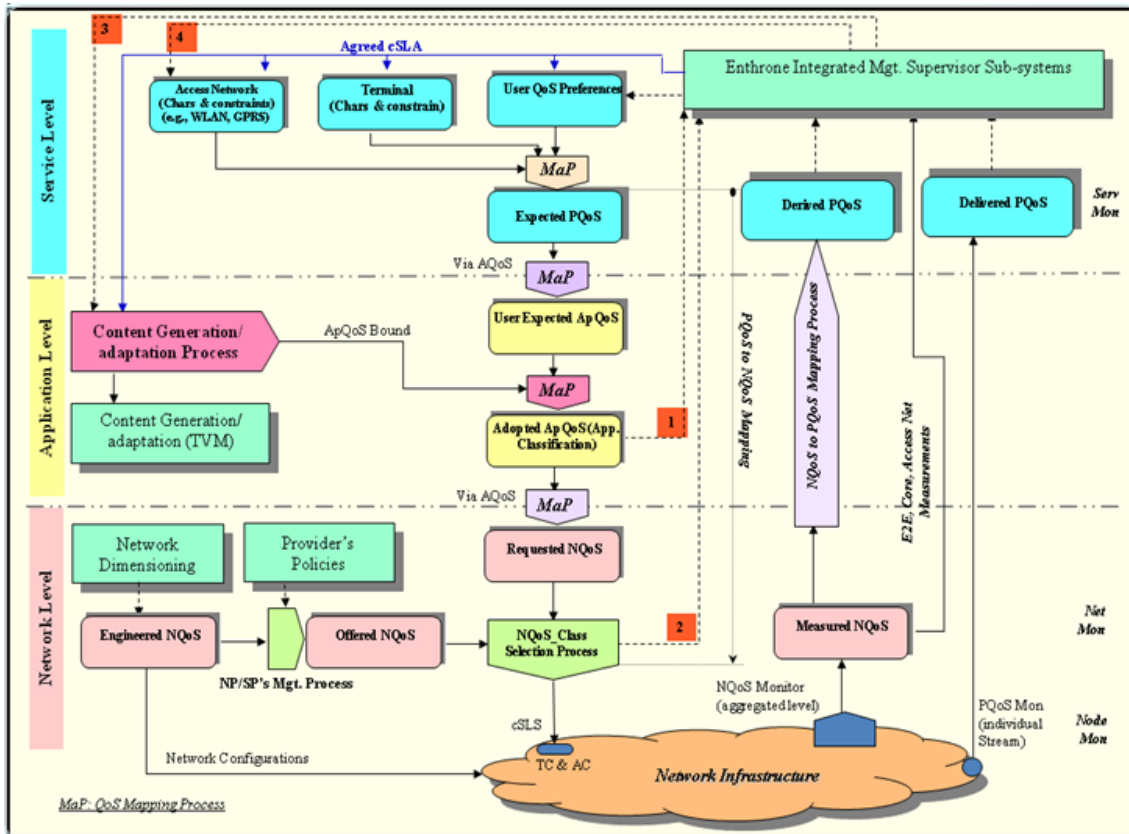


Figure 2-20: Architecture of QoS interaction at different layers [26]

2.9 3D Video Transmission over UMTS

Provision of video services over wireless networks is gaining more popularity. UMTS is capable of providing high mobility, whereas wireless local area networks (WLANs) are known for having relatively higher bandwidths. A WLAN can be considered an access network to 3G environment. In an interconnected environment, the end-to-end rate is imposed by the low capacity link. This will enable a user to access 3G cellular services via a WLAN. Figure 2-21 shows different video streaming scenarios over a UMTS network while figure 2-23 show the architecture of the UMTS for video streaming. Thus the WLANs can be considered as a complementary technology for the 3G cellular data networks [19].

Review of Wireless Technology



Figure 2-21: UMTS video streaming scenarios [28].

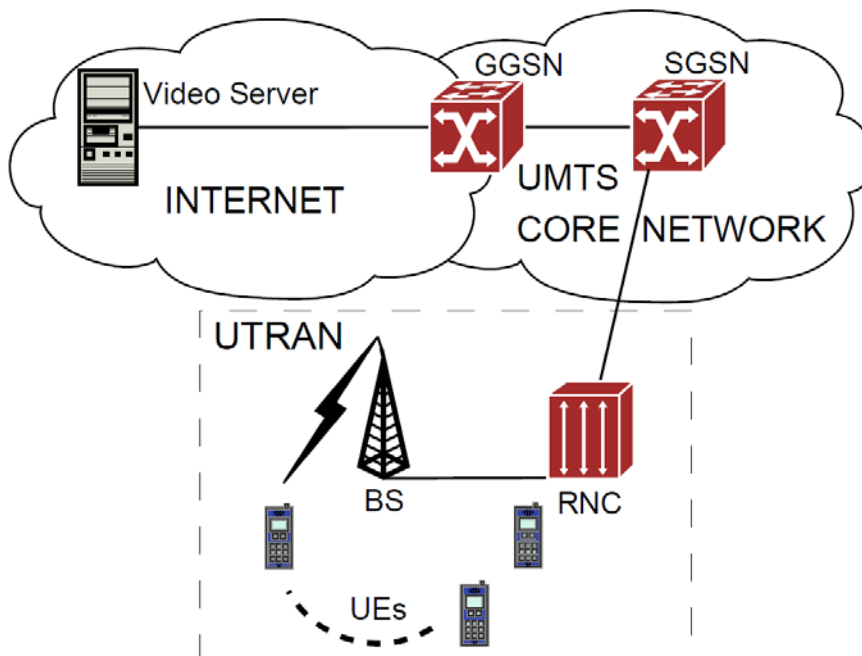


Figure 2-22: UMTS Video Transmission [28]

Review of Wireless Technology

The UMTS uses (WCDMA) as its radio interface technology. UMTS offers higher data rates with respect to older 2G and 2.5G networks and, with the Release 5 version, is evolving into an all-IP, wireless packet network [20]. The increased bandwidth provided by UMTS allows for the deployment of a wide range of services, like voice, data and multimedia streaming services.

Figure 2-22 shows a typical architecture for video streaming and a simplified UMTS architecture consisting of the core network and the UMTS Terrestrial Radio Access Network (UTRAN). The UTRAN consists of Radio Network Controllers (RNC) which control several base stations (BS, or Node B). A mobile user present in the UTRAN can stream video on her UE from a server in the Internet. The UTRAN is connected to the Internet through the Serving GPRS Support Node (SGSN) and the Gateway Serving GPRS Support Node (GSGSN) present in the core network as described earlier. In the UMTS network, all the links in the core network are usually over-provisioned. The fluctuations in radio channel quality that are inherent of wireless links will often make the UTRAN act as a bottleneck. Therefore, resource management will be required in UTRAN to provide good Quality of Service (QoS) to the users.

High Speed Downlink Packet Access (HSDPA) [20] can be regarded as a packet-based enhancement of UMTS. It supports data rates of several Mbps, making it suitable for data applications ranging from file transfer to multimedia streaming and video transmission. Such data rates are in principle high enough for supporting the video streaming of multimedia flows to several users at a time in a single cell. Nevertheless, due to its shared nature, the radio channel used to transfer data from the base station to mobile terminals remains a challenging environment for delay and loss sensitive applications like video. One of the main characteristics of HSDPA is the use of MAC-layer scheduling to perform resources management such as bandwidth allocation between terminals, taking into account the radio channel conditions of all users. Such bandwidth fluctuations cause packet loss and degradation in the quality of the received video. Additional factors like fairness between users, cell throughput or quality of Service (QoS) parameters are also considered in some scheduling mechanism proposed in the literature.

2.10 UMTS Simulator

This section gives an overview of the OPNET simulator [22][23][24] that was used during this research work, focusing mainly on aspects of the UMTS models of OPNET. All chosen options

Review of Wireless Technology

for simulations are described and the results are analysed. Comparison between the results is obtained from OPNET and the results from the theoretical model are used to validate/verify our simulation results.

2.10.1 OPNET Modeller

OPNET is a very powerful network simulation tool [24][25][26]. It has several modules for Application Performance Management, Network Operations, Cell Capacity Planning & Design, and network Research and Development (R&D) studies. These modules can fit different organisations: Enterprise IT, Defense, Service Providers, Network Equipment Manufacturers, Universities, and Research Institutes. OPNET Technologies, Inc. has free research software licenses for the two latter ones, and that is why OPNET Modeler, a commercial package for network modeling and simulation, has been used in my research. It allowed me to design and study communication networks, equipment, protocols, and applications such as UMTS, WiMAX, IPv6, MPLS, and host of others.

A network topology is a combination of nodes and links, and both of them can be user defined in order to study research problems. Simulations can then be performed and the results be analysed for any kind of network element in the simulation network. The OPNET Modeler 14.5 has key features such as scalable and efficient simulation engine, open model source code, different simulation technologies, object-oriented modeling, graphical user interface, integrated debugging and analysis, etc.

Modeler is based on a series of hierarchical editors that can be very efficient, simulating real networks, equipment and protocols. To create node and process models, build packet formats, create filters and parameters, one needs to use the following editors. The most important ones are described below.

Project Editor: The main editor for creating a network simulation is the Project Editor. This is used to create a network model using models from the standard library, collect statistics about the network, run the simulation and view the results. Using specialized editors accessible from the Project Editor via File> New, one can create node and process models, build packet formats and create filters and parameters. Figure 2-23 showed the project editor in OPNET environment.

Review of Wireless Technology

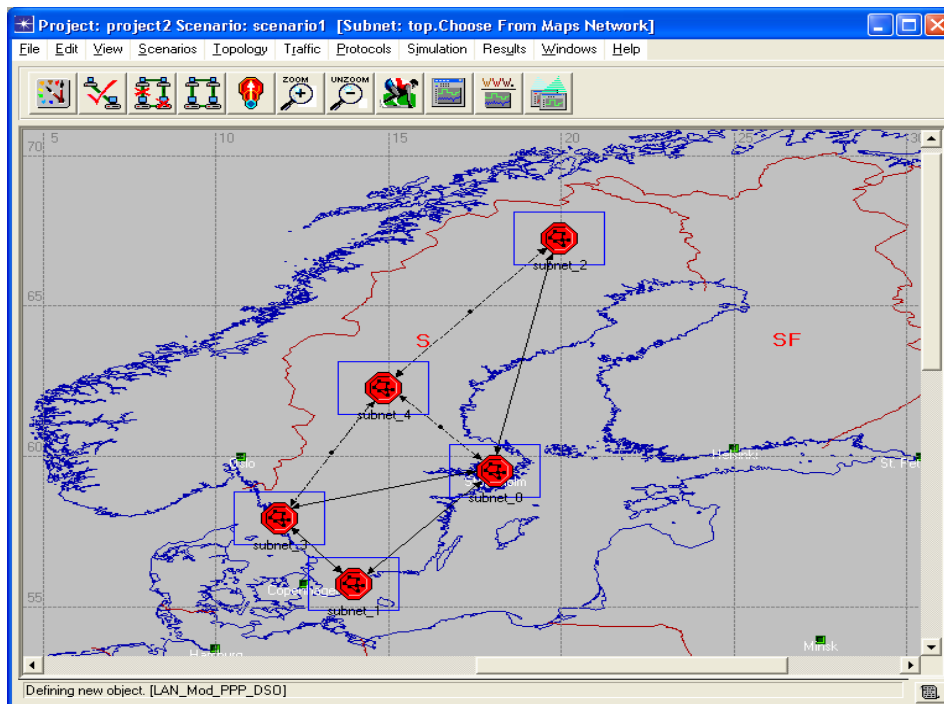


Figure 2-23: Project Editor [35]

Depending on the type of network being modeled, a network model may consist of subnetworks and nodes connected by point-to-point, bus, or radio links. Subnetworks, nodes, and links can be placed within subnetworks, which can then be treated as single objects in the network model. This is useful for separating the network diagram into manageable pieces and provides a quick way of duplicating groups of nodes and links.

Node Editor: The Node Editor is used to create models of nodes. The node models are then used to create node instances within networks in the Project Editor. Internally, OPNET node models have modular structure. You define a node by connecting various modules with packet streams. The connections between modules allow packets and status information to be exchanged between modules. Each module placed in a node serves a specific purpose, such as generating packets, queuing packets, processing packets, or transmitting and receiving packets. Figure 2-24 showed the node editor in Opnet simulation environment.

Review of Wireless Technology

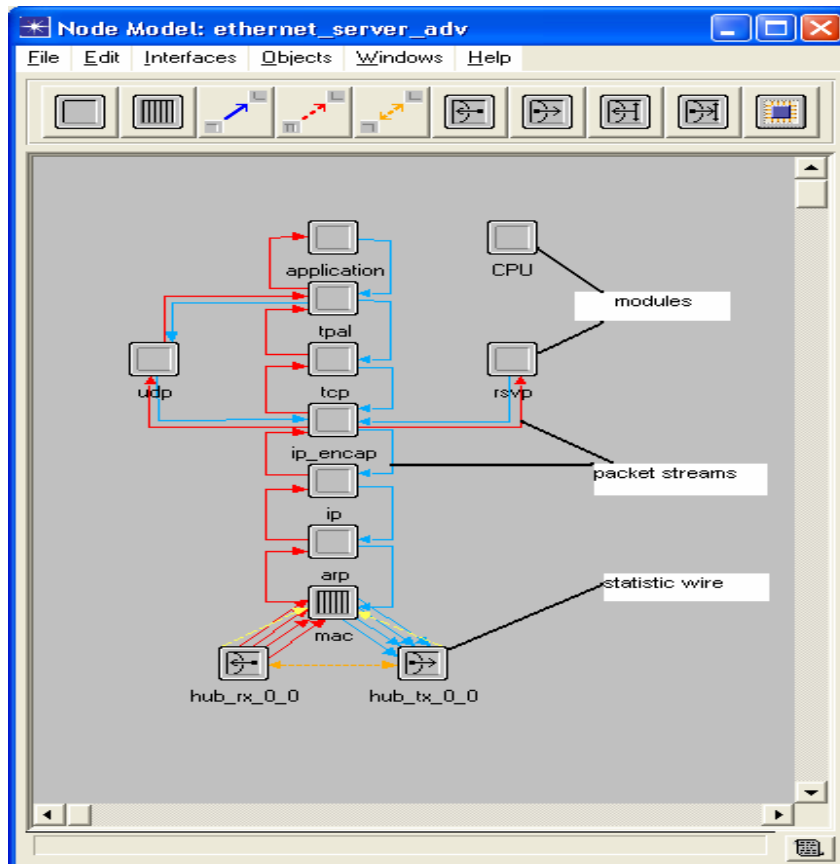


Figure 2-24: Node Editor [35]

Process Model Editor: To create process models which control the underlying functionality of the models created in the Node Editor, one can use the Process Model Editor. Process models are represented by finite state machines (FSMs) and are created with icons that represent states and lines that represent transitions between states. Operations performed in each state or for a transition are described in embedded C or C++ code blocks. It is used to describe the processes (Protocols, resources, applications, algorithms and queuing policies) that run inside the modules. Figure 23 maps the hierarchical structure among the main models. Figure 2-25 showed the process editor in Opnet simulation environment.

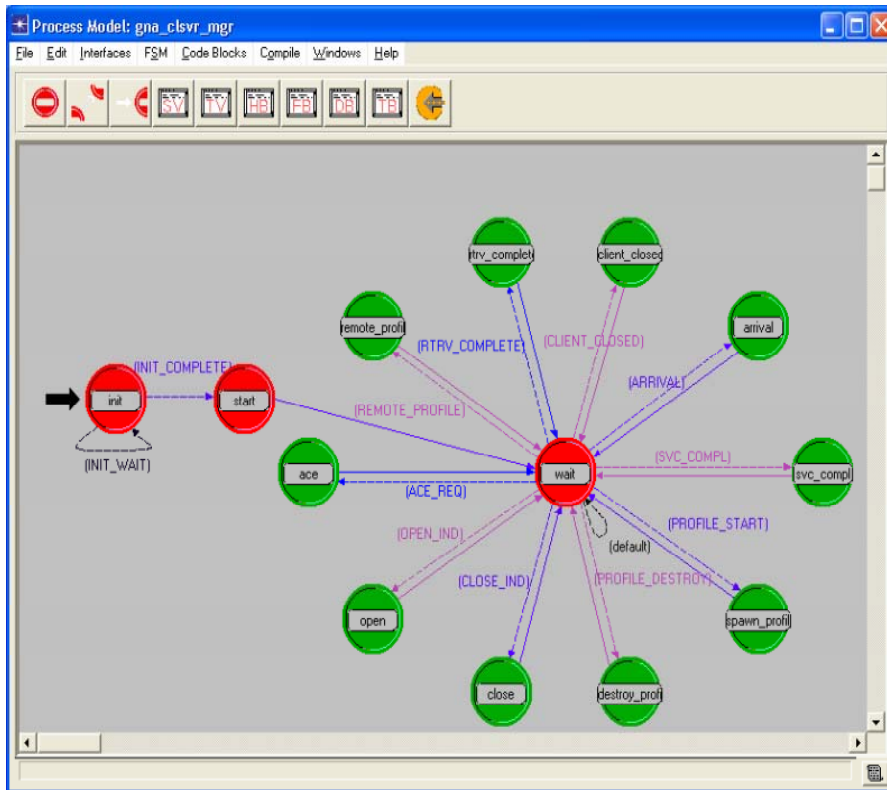


Figure 2-25: Process Editor [35]

There are many more editors in OPNET Modeler [24], such as Path Editor which creates new path objects in order to define a traffic route. Any protocol model that uses logical connections or virtual circuits such as MPLS, ATM, Frame Relay can use paths to route traffic. The Packet Format Editor defines the internal structure of a packet as a set of fields, a packet format contains one or more fields, represented in the editor as coloured rectangular boxed. The size of the box is proportional to the number of bits specified as the field's size. Link Model Editor creates new types of link objects, each new type of link can have different attribute interfaces and representation, and some important editors for simulating and analysing the results.

Probe Editor: The prob editor is used to specify the statistics to be collected, by using different probes. There are several different types of statistics that can be collected, including global statistics, link statistics, node statistics, attribute statistics, and several types of animation statistics.

While creating a new network model [25], one must first create a new project scenario. The project is a group of related scenarios that explore a different aspect of the network. The scenario

Review of Wireless Technology

is defined by the topology, the scale and size, and the background scenario for the network. One can use nodes (real-world network objects) and links (communication medium that connects nodes to one another). These take one to the next level: i.e., creating network topologies. It is possible to import the network or manually construct it. There are three different ways of doing it.

- Topology: although the model library that comes with OPNET provides models of various devices used in today's network, it is possible to import topologies directly from a number of vendor products.
- Rapid configuration: Allows selecting a network configuration, the type of nodes and the types of links.
- Place each individual node from the Object Palette into the workspace.

Once the general network topology is built, we need to add a server. This can be found in the object palette; otherwise we drag a universal server and configure it to our requirements. We then connect the server to the network.

Finally, we add configuration objects to specify the application traffic that will exist in the network. To do this, the Application Configuration and Profile Configuration objects should be dragged into the workspace. How do we configure them?

Communication networks enable applications to exchange data, and each of these applications generates its traffic, [21]. This means, different types of traffic cause and experience a different set of challenges, so we may want to accurately model the traffic patterns generated by a variety of applications. Each application can be enabled or disabled on the client nodes through the use of one or more profiles, and each can be specified as a supported application service type on the server nodes. FTP, E-mail, Remote Login, Video Conferencing, Database, HTTP, Print, Voice and Custom application are the components of the Application Model as shown in Table 2-4. Each of these applications has configurable attributes, such as start time, duration and repeatability.

Table 2-4: UMTS QoS Classes

QoS Class	Service	Paramter	Value Scenario
			Light Heavy

Review of Wireless Technology

Conversational	VoIP	Mean call duration/Hour	120	120
Streaming	Streamin g	Mean File Volume(kB)	17500	17500
		Calls/Hour	2	4
Interactive	HTTP	Mean Page Volume(kB)	34.4	34.4
		Calls/Hour	4	12
Backgroud	E-mail	Mean File Volume(kB)	100	100
		Calls/Hour	4	4

In order to configure a workstation to model the behaviour of a user or group of users, one needs to describe their profile. The profile is a set of applications used by that group; on how long and how often applications are used through the day. These profiles can represent different user groups. One can execute profiles at the same time and repeatedly and also configure applications within a profile. Applications can be executed at the same time, one after the other, in fact one has the choice to simulate what really happens in a communication network. We choose which statistics [24] to analyze. Statistics can be collected from individual nodes in the network called object statistics or from the entire network which is called global statistics. One can choose “Individual DES Statistics” to collect single node statistics or one can go to the Probe Editor. The Probe Editor allows one specify the statistics to be collected during simulations, as they can be global statistics, nodes statistics, link statistics, attribute statistics, etc.

2.10.2 OPNET Simulation Technologies

OPNET Modeler supports three kinds of simulation:

- Discrete Event Simulation (DES), the operation of a system is represented as a chronological sequence of events such as modeling a queue. Each event occurs at an instant in time and marks a change of state in the system. Common exercise in learning how to build discrete-event simulations is to model a queue, such as customers arriving at a bank to be served by a teller. In this example, the system entities are customer-Queue and Tellers. A number of mechanisms have been proposed for carrying out discrete-event simulation; among them are the event-based and process-based. DES provides highly detailed models that explicitly simulate packets and protocols messages. DES executes the protocol almost as a production environment, providing accurate results. Though, simulation runtimes are longer than other techniques.
- Flow Analysis uses analytical techniques to model steady state network behaviour. This technique can be very useful to study routing across the network in steady scenarios. Execution runtimes can be much faster.
- Hybrid Simulation combines DES and Flow Analysis techniques to provide accurate and detailed results for targeted flows. In order to achieve reasonable times, it is possible to fine-tune the balance between discrete and analytical models.

Although there are several simulation techniques, DES has been continuously enhanced to deliver faster and more efficient simulations that scale significantly with the amount of traffic in the model. DES is the recommended tool to study application performance, network capacity planning and resource utilisation analysis. To run a simulation, one can select “Run Discrete Event Simulation”, and then can configure the duration, update interval and simulation kernel. Run the simulation and a simulation progress dialog box will appear. Then, check Simulation Log for errors.

There are several ways to see the information collected for each statistics [24][25]. One can view simulation results in the Project Editor; the best way is to use the Analysis Tool. This feature can create scalar graphics and parametric studies, define template to which you apply statistical data, and create analysis configurations that can be saved and viewed later. The Filter Editor is used to define filters to mathematically process, reduce, or combine statistical data.

2.10.3 OPNET UMTS Model

OPNET's UMTS model suite is based on 3GPP Release 99 standards [28] supporting only packet switched (PS) traffic [27]. It allows modeling of UMTS networks to evaluate end-to-end QoS, throughput, drop rate, end-to-end delay variation (jitter) through the radio access network and core packet network. It can also evaluate the feasibility of offering a mix of service classes given QoS requirements. It is possible to configure a UMTS network model in two different ways:

- UMTS network using application traffic: figure 2-26 shows UMTS workstation nodes routing application traffic (Voice, ftp, etc.) to other UMTS workstation nodes or server nodes.
- UMTS network using raw packet generator: figure 2-27 shows UMTS station nodes sending data traffic (conversational, streaming, interactive and background classes) to other UMTS node stations through a single SGSN node. This network topology should be used when one wants to model raw data traffic within the UMTS network, and is not interested in the external IP network and does not wish to model CN to CN data transfer.

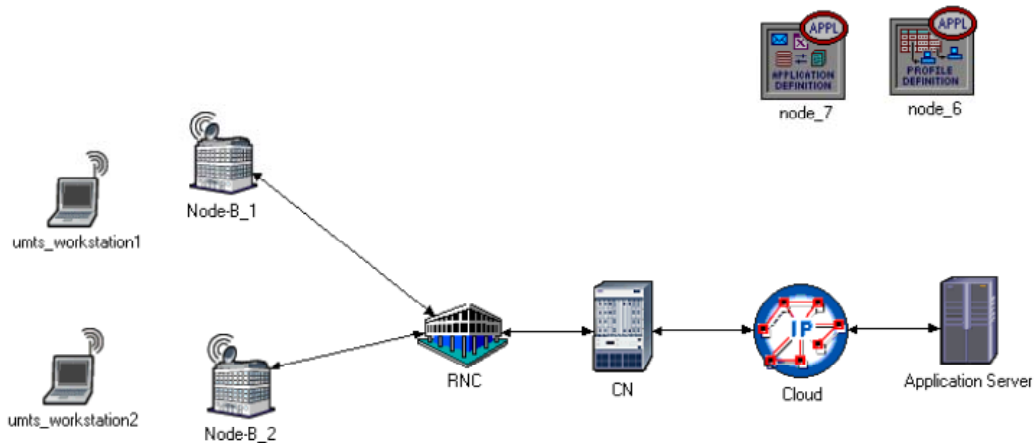


Figure 2-26: UMTS network using application traffic (Advanced) [26]

Review of Wireless Technology

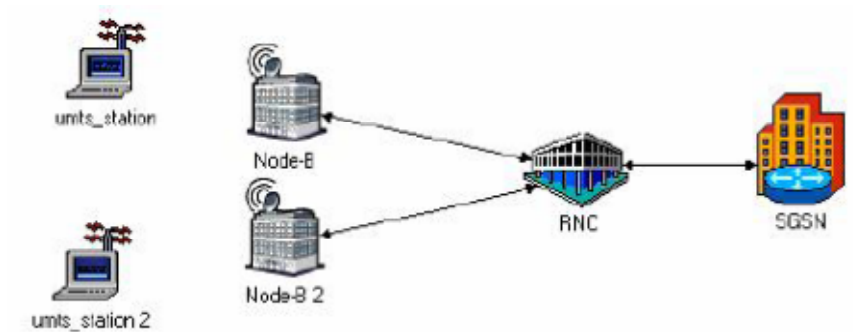


Figure 2-27: UMTS network using raw packet generator (Simple) [26]

When the UMTS station has traffic ready to send and queued on one of four QoS classes, it sends an Active Packet Data Protocol (PDP) Context Request to the SGSN, which includes the QoS request. The SGSN can grant or reject access. So, the SGSN sends a Radio Access Bearer (RAB) and the QoS request to the RNC. If there is sufficient uplink (UL) and downlink (DL) capacity, the RNC sends a Radio Bearer Setup (RAB) request to the UE. The UE sets up the channel as specified and sends a Radio Bearer Complete back to the RNC. At the end, the RNC sends a RAB assignment, which includes the granted QoS to the SGSN. The SGSN sends a Packet Data Protocol (PDP) Context Accept message to the UE, so it can send packets to the destination. There are three different node models for the User Equipment: simple mobile stations (umts_station), advanced workstations (umts_wkstn), and advanced servers (umts_server). In this thesis, the second node station is considered. Advanced workstations can be modeled as pedestrian outdoor, vehicular outdoor and indoor office. Looking at the UMTS work station model in figure 2-28, it includes the full TCP/IP protocol stack between the application layer and the GPRS Mobility Management (GMM) one. There is also a RLC/MAC layer, a radio receiver, a radio transmitter, and an antenna. Transport channels link the RLC/MAC layer, the transmitter and the receiver. The (GMM) layer contains mobility management functions, session management functions, and radio resource control functions. The RLC/MAC layer contains the RLC and MAC layers, including priority handling of data flows, the three types of RLC modes, and segmentation and reassembly of higher-layer packets.

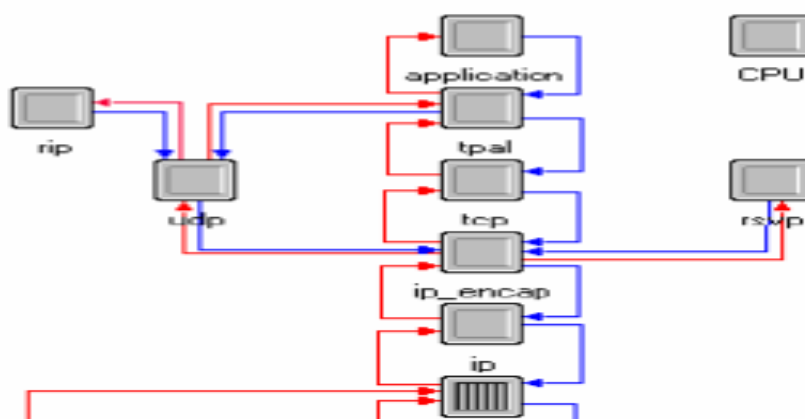


Figure 2-28: UMTS station model [27]

In order to manage the network's air interface, the UMTS model suite includes two Node-B models, the single-sector Node-B and the three sectors Node-B. Figure 2-29 represents the single-sector Node-B model. There is one Node-B processor module (node_b_mux), connected to an ATM stack, a transmitter module (pr_0) and receiver module (pt_0). Through the UL and DL, all packets travel over the transport channels.

The RNC is connected to one or more Node-Bs, and manages the resources of the air interface of all the UE's. Particularly, it coordinates the admission control process, manages handovers between its Node-Bs, buffer packets per QoS class, communicates with the SGSN, and manages the radio bearers. The RNC node model has nine ATM node stacks, each one of them is connected to the SGSN. More Node-Bs can be supported by adding more ATM stacks to the node structure.

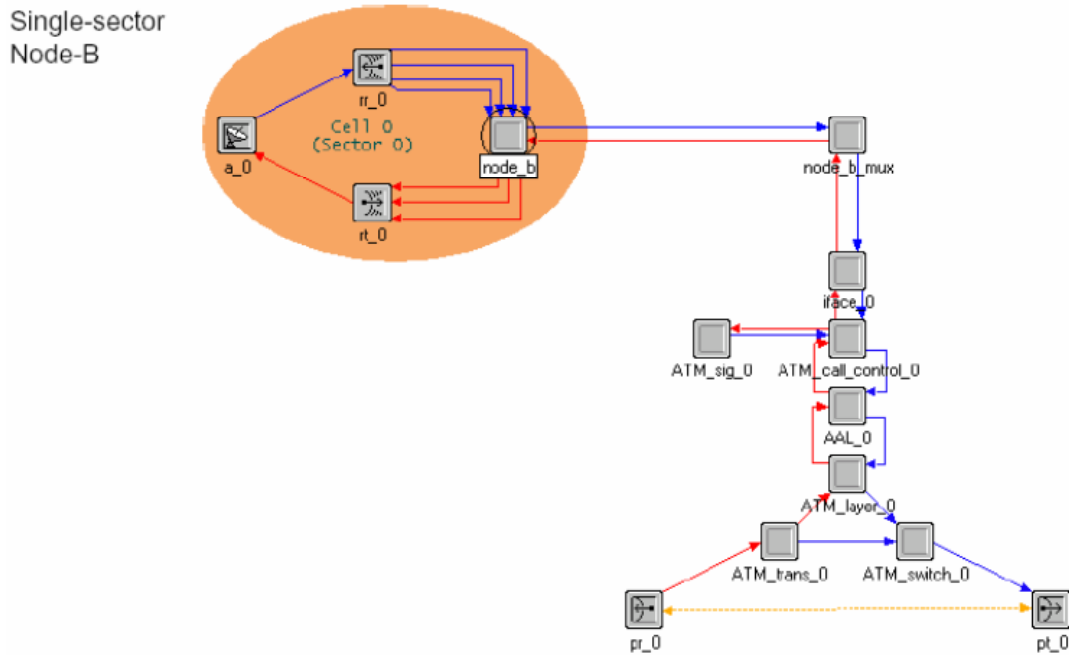


Figure 2-29: Single sector Node B node model [27].

It is possible to model the air interface between the UE and the UTRAN by modifying OPNET’s Modeler 14.5. Some of the most important pipelines are: the received power pipeline stage includes a path loss model and shadow fading model that depends on the environment; the propagation path loss models are based on Recommendation ITU-R M.1225 Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000, 1997; thermal noise and noise figure of the mobile and base station receiver can be modified due to the background noise pipeline stage; the interference noise pipeline stage will include same cell and other-cell interference calculations; the bit-error rate pipeline stage includes the Signal-to-Noise Ratio (SNR) versus Bit Error Rate (BER).

2.10.4 OPNET Scenarios

In this section, the network architecture and the applications and profile configuration will be described. Figure 2-30 presents the network architecture used in the simulation.

Review of Wireless Technology

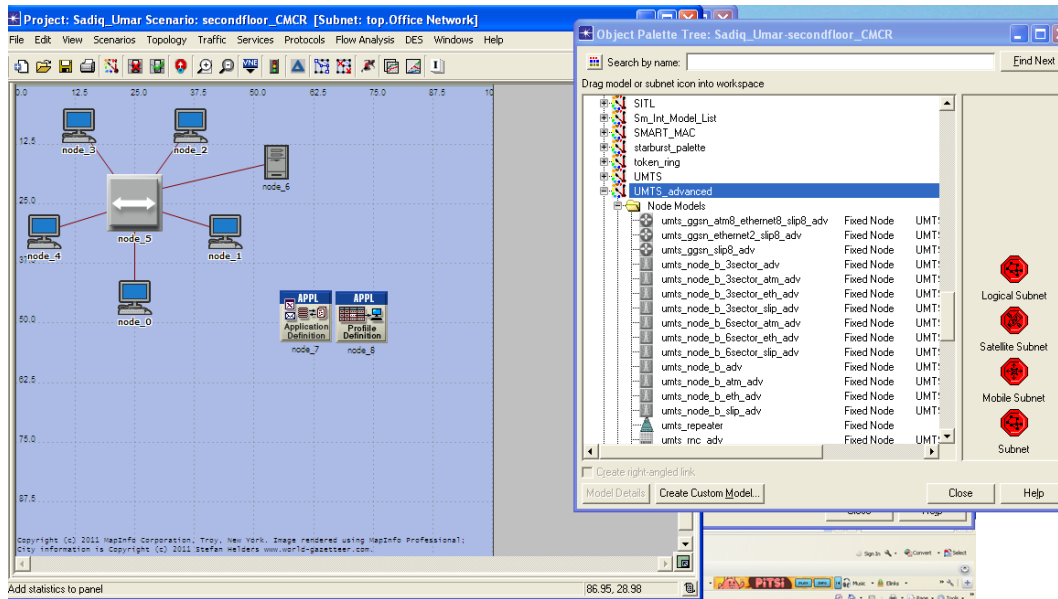


Figure 2-30: UMTS network architecture considering 5 Mobile Terminals (MTs).

2.10.5 Applications and Profiles

Seven different applications are considered: VoIP, Streaming, FTP, HTTP, E-mail, MMS and Video-conferencing [27][28][29]. OPNET does not differentiate between e-mail and MMS, so, the only difference between them is mainly the profile size. All applications were defined by [30][32][32]. The Profile configuration has one main parameter, the inter-repetition time, which is the time between sessions per user (one user, one application only).

One of the limitations of OPNET is the throughput limitation to each QoS class [33]. In order to study the network behaviour, several scenarios with different number of users were simulated considering Video/Data with the profiles. The number of users ranges from 10 to 50. The lower limit is a situation where everyone should be served, while the higher one presents a significant rejection probability. At the end, many scenarios were assessed. We conduct experiments for number of users per profile and per application for the video/data. One can see that some scenarios are quite demanding.

2.10.6 Analysis of results

This section presents the simulation results and analysis. As discussed earlier, OPNET offers a variety of network parameters that can be analysed. We selected some Node or Object Statistics such as throughput, packet delays and queuing delay. Global Statistics such as application's

Review of Wireless Technology

throughput and application's delay and point-multi-point statistics throughput analysis. For us to present convincing and validated results, each scenario was simulated upto 20 times. From figure 36 below, we can see that the total throughput for Node B which is covered by Node 5, which is the total throughput of the received traffic by the Node B in UL. These are mean values in one hour for each scenario for a six hour experiment. We also take the high standard deviation values into consideration while conducting these experiments. Therefore more simulations would have been run in order to reduce the standard deviation and increase the accuracy of our results. OPNET always provides service first to Conversational and Streaming traffic classes, so for scenarios with more interactive and background users not being served, because of VoIP and Streaming clients.

2.10.7 BER vs Eb/No over UMTS channel

The BER vs Energy/bit over Noise Spectral Density in UMTS channel is computed. In general, the bit error rate is a function of the type of modulation scheme used for the transmitted signal. We evaluate the BER both theoretically and experimentally.

The SNR (dB) is given by

$$SNR = 10 \log \left[\frac{P_r}{N_b + N_i} \right] \quad (2-1)$$

Where

P_r = received power (Watts), which is computed

N_b = background noise power (Watts)

N_i = Interference noise power (Watts)

The SNR value is added to the processing gain (dB) to obtain the effective SNR. This effective SNR is also written as E_b/N_o where E_b is the received energy per bit (Joules) and N_o is the power spectral density (in Watt/hertz). Bit error rate is derived from the effective SNR based on the QPSK (downlink)/BPSK (uplink) modulation curve assigned to the receiver in the UMTS error/fading channel. The probability of bit error P_b (BER) is given by [52].

$$P_b(E_b/N_0) = \frac{1}{2} [1 - E_b/N_0 / (1 + E_b/N_0)] \quad (2-2)$$

Figure 2-31 shows the BER curves vs Energy/bit over noise spectral density for the downlink in a UMTS fading channel.

Review of Wireless Technology

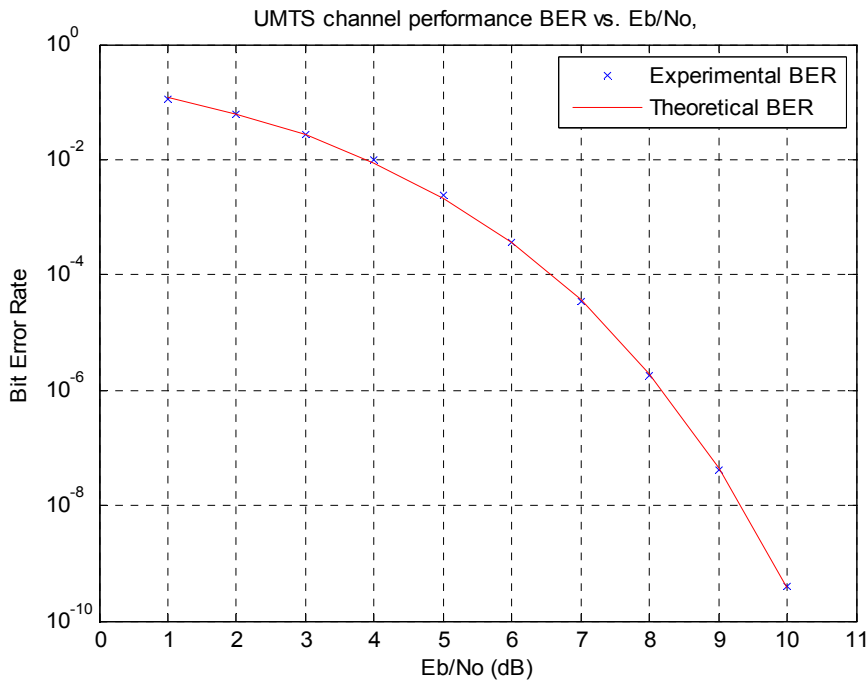


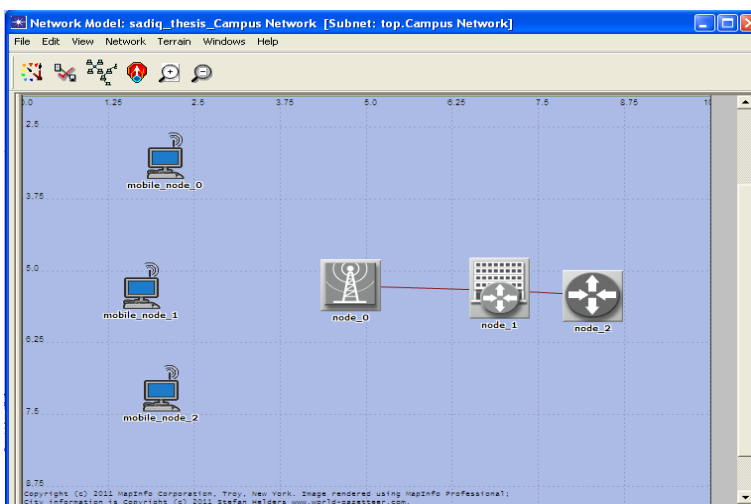
Figure 2-31: UMTS channel performance over Eb/No ranges

We conclude from figure 2-31 that, simulation results are demonstrated to confirm the theoretical/analytical results as the two curves passed over each other with slight differences. If E_b/N_0 increases from 1 to 10 dB, BER will be decreasing to the required value 10^{-10} .

2.10.8 Simulation of a Brunel Local Area Network

This experiment was to make a simulation model in OPNET of a simple Local area network with 3 computers. The experiment focused on building OPNET models for simulation of UMTS network within Brunel campus.

Scenario one, consists of 3 wireless nodes organized as clients, a wireless router and gateway to wired network. The wireless nodes also have the capability to communicate between themselves. The developed models have been used for performance evaluation in two scenarios depending on user traffic, number of clients, network size and radio channel allocations to the involved nodes. Figure 2-32 and 2-33 show the simulation environment. Figure 2-34 to 2-37 show the results achieved.



Review of Wireless Technology

Figure 2-32: Simulation of 3 wireless laptops.

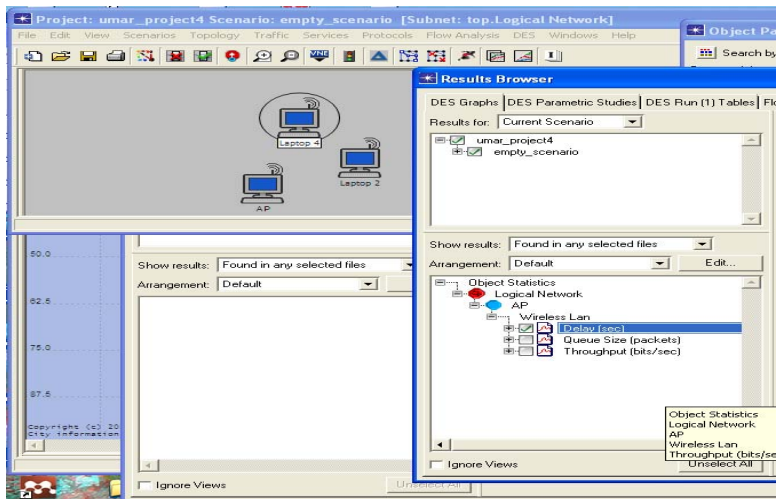


Figure 2-33: 3 transmitting laptop in the simulation scenario.

Review of Wireless Technology

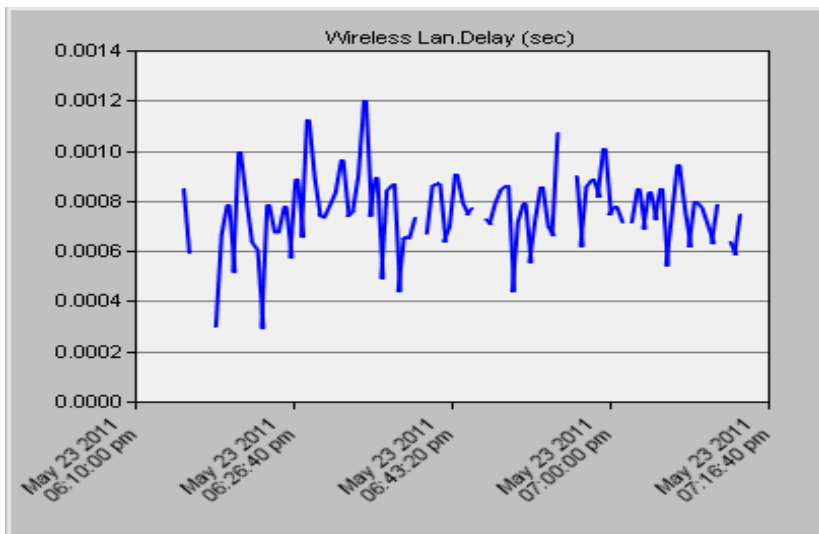


Figure 2-34: Delay experience on the wireless network.

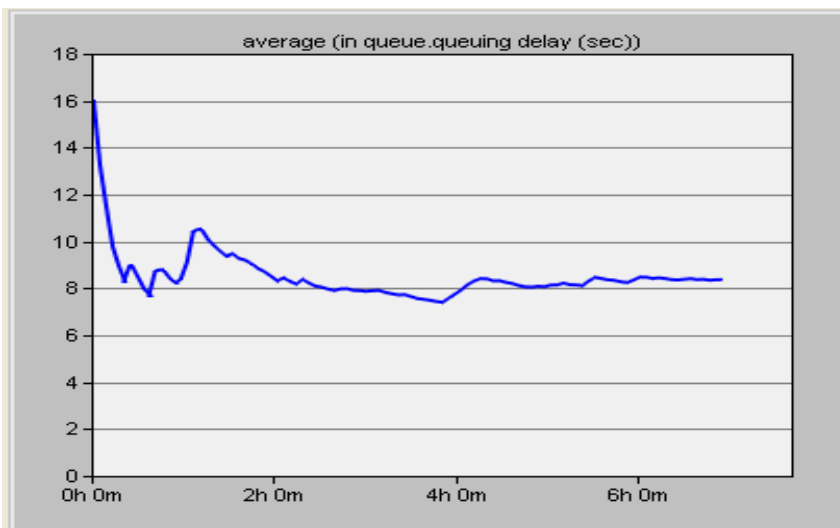


Figure 2-35: Average Queuing Delay experience on the wireless network

Review of Wireless Technology

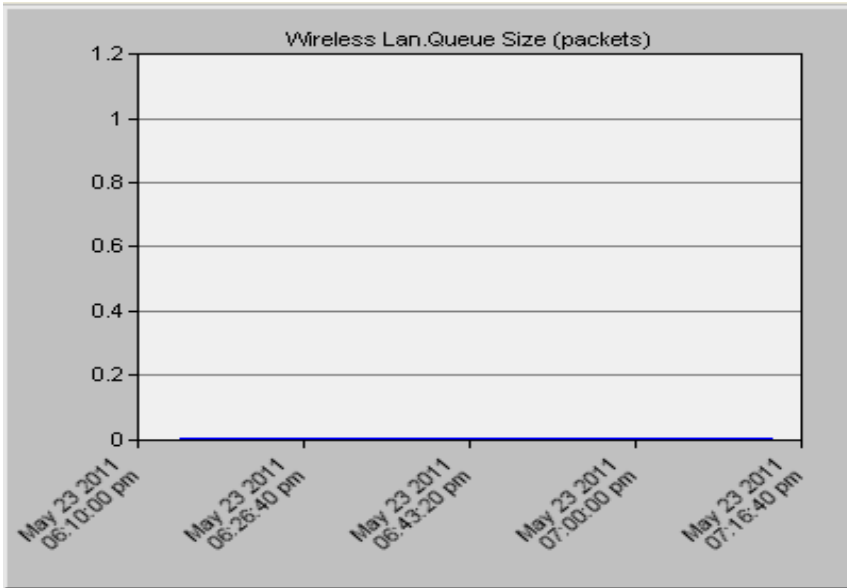


Figure 2-36: Wireless queue size.

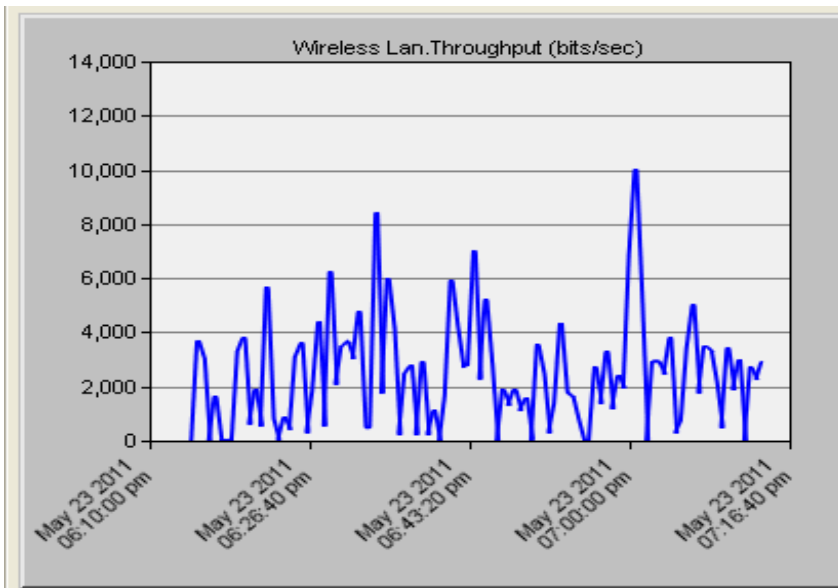


Figure 2-37: Throughput performance of video packets transmitted over a UMTS channel.

Review of Wireless Technology

2.10.9 UMTS Simulation in Brunel University

Figure 2-38 show the UMTS simulation environment, figure 2-39 shows the UMTS simulation sequence editor. Figure 2-40 to 2-42 show the results achieved.

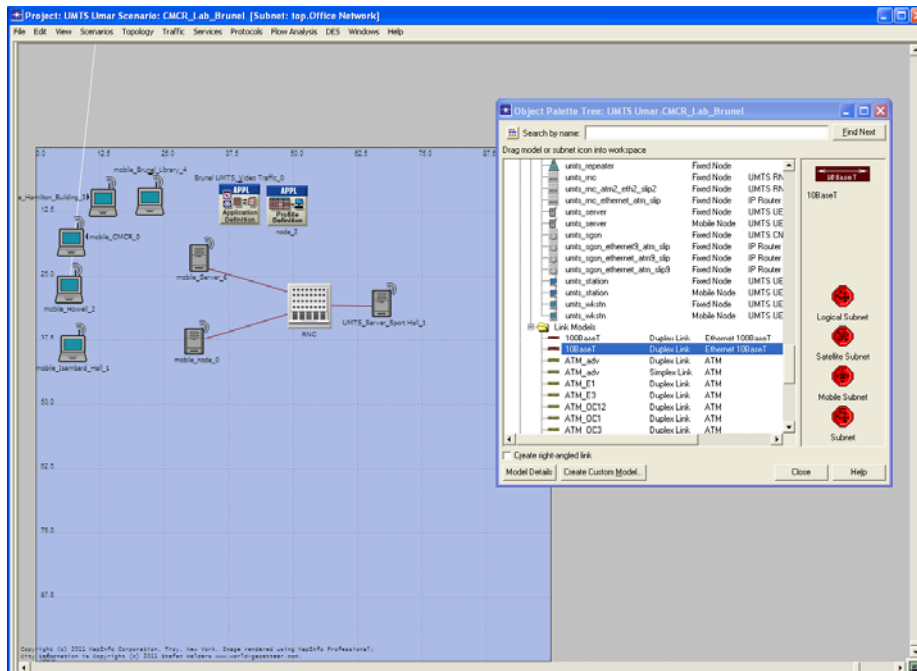


Figure 2-38: Simulation of 5 wireless nodes with Video Application Traffic.

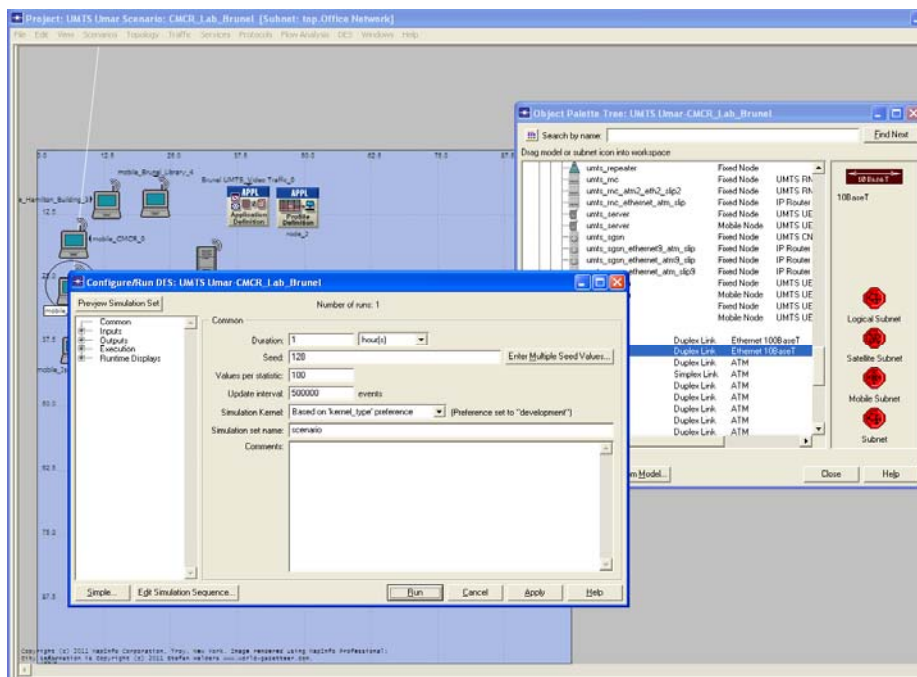


Figure 2-39: Simulation Sequence Editor.

Traffic Profiles:

- UMTS data rate = 1Mbps
- Simulation time = 6 hours
- File Size = 5000 bytes
- Latency = 0.06 sec
- Delay bound= 0.0051728 secs
- Tolarated jitter= Negligible

Review of Wireless Technology

Table 2-4: The nodes and links used in constructing the UMTS network.

1	Type of Network	UMTS Network
2	Node Mode	umts_wkstn
3	Link Model	10BaseT
4	Number of wireless nodes	5
5	Centre	X=25, Y=25
6	Radius	20, click ok

From the results, we can see average throughput increases with a higher number of users. These values are not high mainly due to OPNET limitations, because data rate values for video transmissions are limited to 12.2 Mbps and also because of DL limitations (Applications have a much heavier profile in DL). It is also important that these results are for a single sector and for 3G (UMTS).

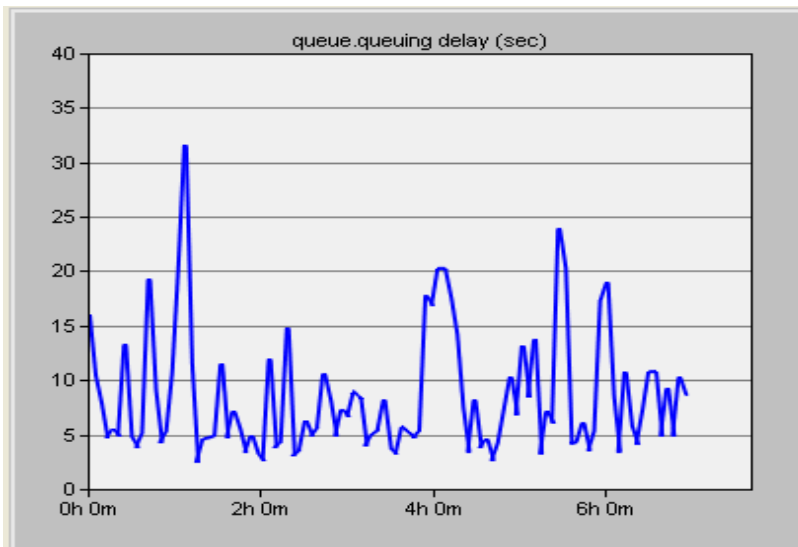


Figure 2-40: Video packet queuing delay

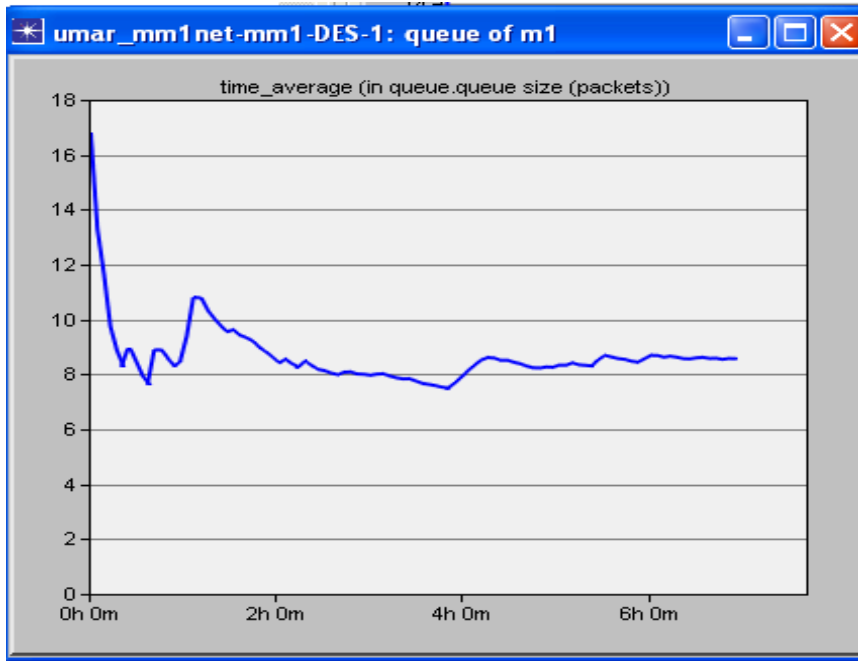


Figure 2-41: Queue size graph in packets

It is also important to know the confidence intervals (CI) of the simulation results. We will now plot the confidence interval.

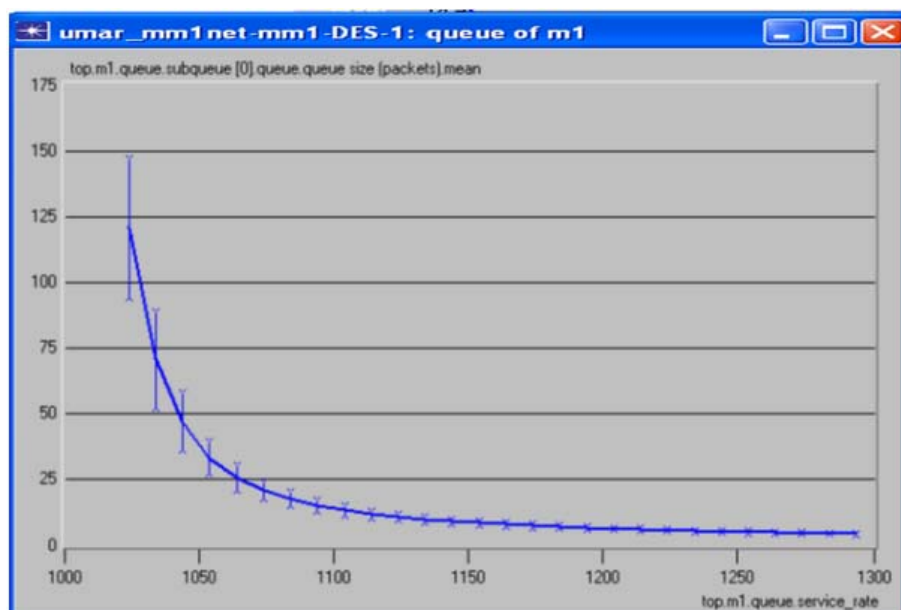


Figure 2-42: Queue size Confidence Interval

In order to get better confidence intervals you need to run more scenarios. Figure 2-42 shows the CI of scenarios. The service rate is between 1024 and 1300 with step size 10.

2.11 UMTS Model Verification and Validation

This section describes the methods used to ensure that our simulation model was both correctly implemented and representative of the real system. These two terms are called model verification and model validation [35][36][37]. Section 2.11.1 discusses the use of a single UMTS input signal to verify correct operation of the simulation. Section 2.11.2 presents theoretical analysis that was used to validate our UMTS simulation model. We compare theoretical and simulation results for a single client traffic input.

2.11.1 UMTS Model Verification

UMTS model verification is the process of determining if our simulation model functions correctly and all results obtained therein are correct. This includes such tasks as debugging the codes, testing for errors, and testing the functionality of different modules. The simulation was designed in OPNET Modeler 14.5. That simplified the task of model verification since each module was tested independently. Each node and process in the simulation was tested to verify that it functioned correctly. This was done by running simulation of shorter durations in the OPNET debugger mode. This mode provides an environment where the user has interactive control of the simulation in order to investigate its behaviour by setting breakpoints and traces to print out detailed simulation logs about events and objects. The user-defined authentication and process control was verified by creating traces for each of the messages. Each of the protocol transactions was verified by inspecting the proper message flow to ensure that it is correct.

It shows that our design mechanisms of the UMTS model functioned according to the 3GPP standards as designed [37].

Simulations conducted above were run to collect statistics at various points in the model to ensure that the model is functioning well. The standard OPNET processes were tested to ensure they function well as described in the software documentation. Figure 2-43 shows the simulation log file generated on the 21st May 2011.

Review of Wireless Technology

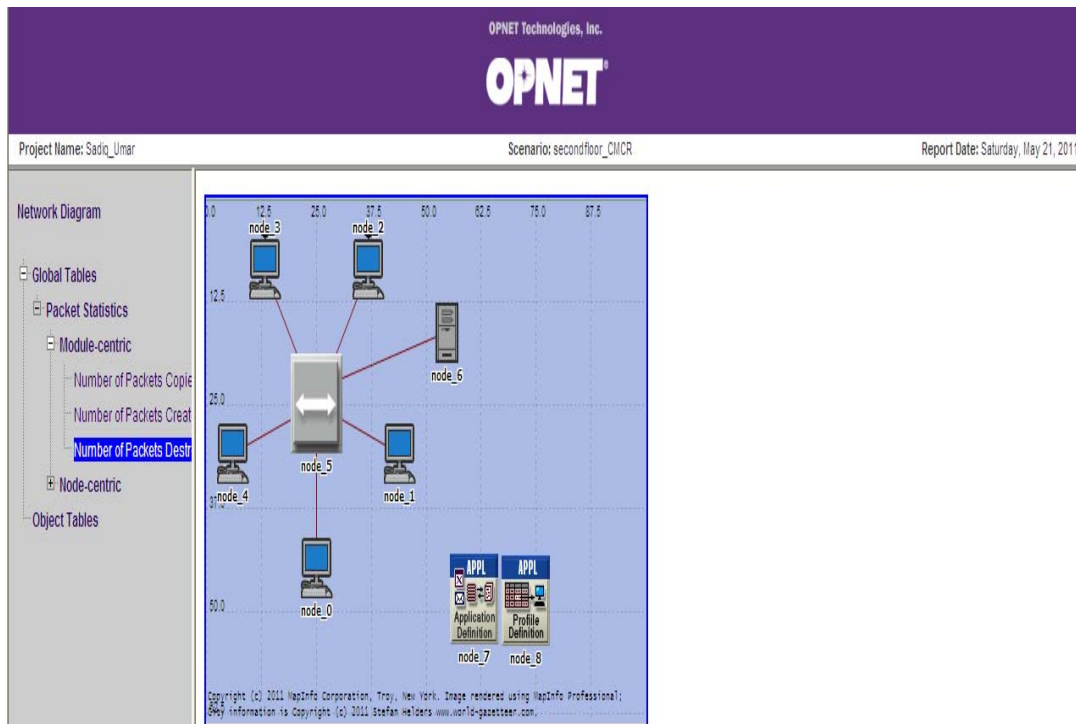


Figure 2-43: Simulation Log file generated

2.11.2 UMTS Validation

Model validation is the process of determining if a simulation model is representative of the real system. A simulation can be validated using expert intuition, real system measurements or theoretical results [37]. Comparing simulation outputs and measurements from a real system is the most reliable way of validating a simulation model. Real system measurements were not available through the course of my PhD in Brunel University, since resources were not available for prototype 3G wireless equipment in the Centre for Media Communication Research (CMCR) lab/Electronics & Computer Engineering department. Comparing simulation and theoretical results was the best option used to validate the simulation model. Theoretical analysis of the system was conducted using FTP application traffic.

To validate the video file transfer within the UMTS access network, a test was conducted with one wireless client data user running FTP application. The system consists of a wireless node operating in an infrastructure Basic Service Set (BSS), which is a set of all stations that can communicate with each other. The FTP file size was increased in order to vary the network

traffic load. The FTP download response time was measured for each increment of the FTP file size and compared with the analytical results.

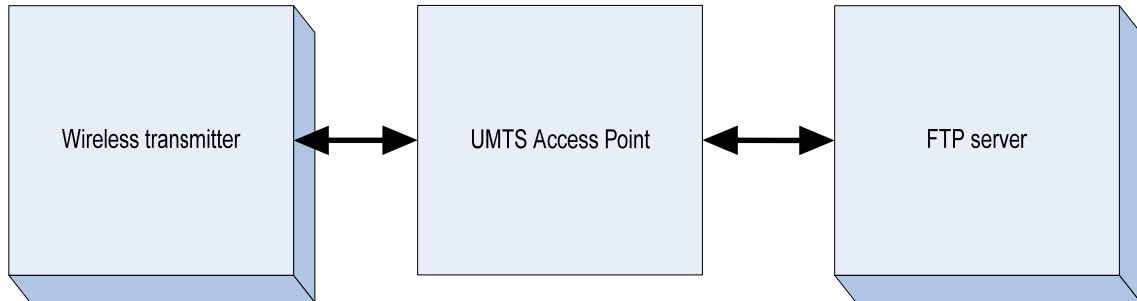


Figure 2-44: Analytical Model

The block diagram in figure 2-44 was used to construct the analytical model and determine all of the delays that an FTP packet will encounter in the system. The following assumptions were made in order to make the analytical model tractable.

Server Network: From 3GPP TR 25.853, the one-way delay in a 200 km link between SGSN and the GGSN would be 800 μ s; therefore, a total round-trip delay of 2ms was chosen between the UMTS access network and the server.

TCP Setup, Slow-start and Window size

The delays from TCP connection setup and the slow start algorithm were assumed to be constant. By assuming an infinite TCP window size, the delays from the TCP were minimized.

Node Processing Delay

The node processing delays were considered to be small as compared to the link transmission delays. Therefore, node processing delays were assumed to be zero.

Number of Packets

The number of packets was calculated by dividing the file size by the maximum segment size (MSS). The MSS for the access network was assumed to be 1500 bytes.

$$Num_Pkts = [Fil_Size/MSS] \quad (2-3)$$

UMTS timing and overheads

To simplify the timing overhead calculations, the term data exchange slot was defined as a complete data exchange including $Data + 2 * SIFS + ACK$ where SIFS means short inter frame space. UMTS Interval included ten data exchange slots. The *Data_Exchange_Slot_Ratio* was defined as the ratio of the number of data slots actually used to transmit divided by the total

Review of Wireless Technology

number of data exchange slots within an Interval Repetition [39]. The term *Slot_Data_OH_Ratio* was defined as the ratio of the data to overhead within a data exchange slot. The following equation defines this ratio.

$$Slot_Data_OH_Ratio = \frac{MSS}{Data+(2*SIFS)+ACK} \quad (2-4)$$

Where,

$$Data = \left[\frac{UMTS_MaxSDU+UMTS_MSDU_Header}{Data_Rate} \right] + plcp_overhead \quad (2-5)$$

Packet Delay

Using the terms defined above, the per packet delay was defined in equation below:

$$Pkt_delay = \frac{Pkt_size}{Data_Rate} * Slot_Data_OH_Ratio * Data_Exchange_Slot_Ratio \quad (2-6)$$

FTP Response time

The FTP response time was defined in equation below:

$$FTP_Rspns_Time = Conn_Setup + Num_Pkts * Pkt_delay \quad (2-7)$$

The theoretical results were calculated in MATLAB using equation (2-7) and the parameters defined earlier. Figure 2-45 shows a comparison between the theoretical and simulation results for various file sizes. The simulation results are very close to the theoretical results. The difference between the theoretical and the simulation results can be accounted for by the inability of the simple analytical models to capture the full effects of TCP. The similarity between the analytical and simulation results validates the correct operation of the data transfer for the UMTS access network. All aspects of the model passed the verification and validation process.

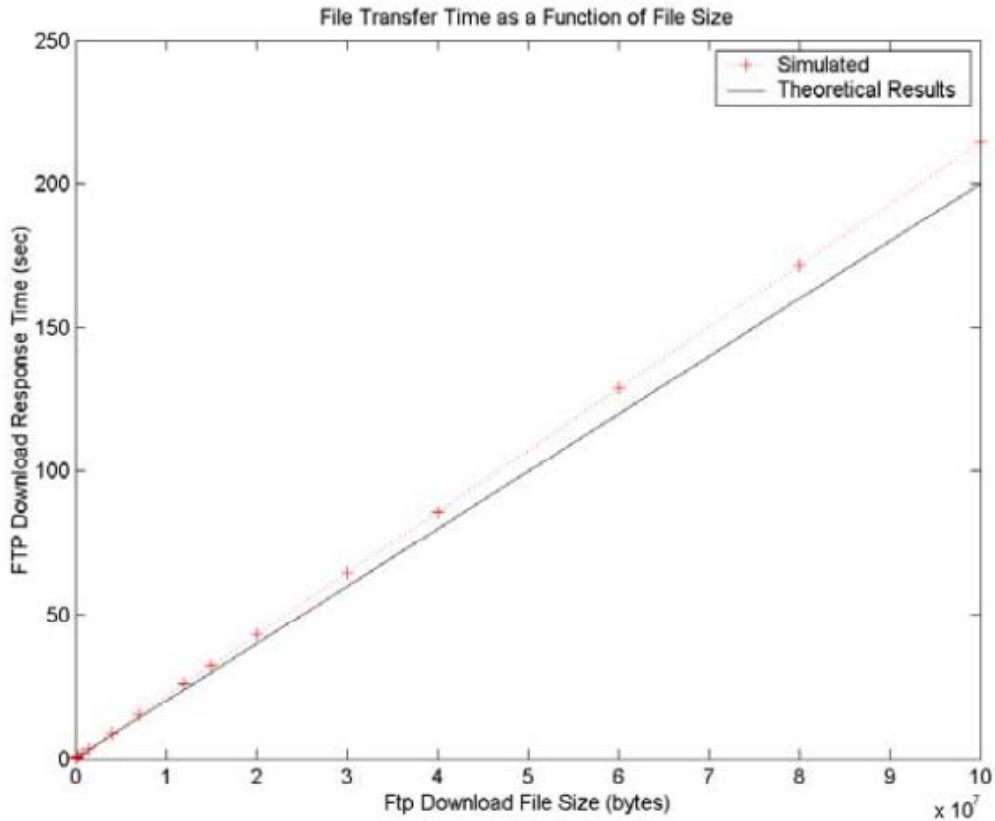


Figure 2-45: Simulation vs Theoretical results

2.12 Summary

In this chapter, we have presented some background information about wireless technology, evolution from 1st to 4th generation networks. We have also presented background information on WiMAX, WiFi and UMTS technology. The thesis examines the behaviour of UMTS networks and protocols and provides basic simulation procedure for UMTS by using OPNET Modeller simulation package. We use five wireless terminals located in various location within Brunel University as shown in the simulation environment.

UMTS model verification was also carried out to determine if our simulation model functions correctly and all the results obtained therein are correct. This includes such tasks as debugging the codes, testing for errors, and testing the functionality of different modules. Each node and process in the simulation were tested to verify that it functions correctly. UMTS model validation was done by comparing simulation and theoretical results.

References

- [1] Gans, J.S., King, S.P. & Wright, J., *Wireless Communications* *. *Communications*, 2.
- [2] H. Hwa, M. Guizani, “Next Generation wireless systems and networks”, John Wiley, ISBN:, 2006.
- [3] Zheng, R., 2003. “Scalable Multiple Description Coding and Distributed Video Streaming over 3G Mobile Networks”, PhD Thesis, Waterloo University, Ontario, Canada, 2003.
- [4] G. James. “Wireless Networks in Europe: A Three-Step Evolution”, MSc Thesis, Tufts University, the Fletcher School of Law & Diplomacy.
- [5] R. Ipsen, C. Long, A. Smith, “GPRS and 3G Wireless Applications”Wiley Computer Publishing, New York, pp.15-25. 2001
- [6] Bienaim, J., 2007. 3G/UMTS/HSPA Deployments Worldwide and Evolution – Perspectives for Central & Eastern Europe (CEE). *Forum American Bar Association*, (June), pp. 1-44.
- [7] Gazis, V. et al, “Evolving perspectives of 4th generation mobile communication systems Communication”, IEEE 13th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002).
- [8] http://www.google.co.uk/search?pq=ims+in+wimax&hl=en&cp=30&gs_id=78&xhr=t&q=wimax+coverage+with+broadband . Access in July 2009.
- [9] www.intel.com/go/wimax . Access in July 2009
- [10] C. S. Chong, M. Y. Low, A. I. Sivakumar, K L. Gay. A Bee Colony Optimisation Algorithm to Job Shop Scheduling, Proceedings of the Winter Simulation Conference (2006).
- [11] Paper, W., Wi-Fi (IEEE 802.11b) and Bluetooth. *Technology*, (February 2001).
- [12] Huang, V. & Zhuang, W., 2002. QoS-Oriented Access Control for 4G Mobile Multimedia CDMA Communications. *IEEE Communications Magazine*, pp. 118-125, March 2002.
- [13] www.itu.int/ITU-D/treg/Documentation/ITU-NGN09.pdf . Access in June 2010

Review of Wireless Technology

- [14] Belhouciet, M. Lazhar, M. Hakim, “4G Wireless Systems”. *Ieee Wireless Communications*, (January 2010), pp. 1-34.
- [15] Specification, Technical & Services, G., 2009. 3rd Generation Partnership Project (3GPP) TS 22.060. *Main*, 0, v 9, pp. 1-26.
- [16] A. Salkintzis, G. Dimiriadia, D. Skyrianoglou, “ Seamless Continuity of Real-time Video across UMTS and WLAN Networks: Challenges and Performance Evaluation” , 2nd IEEE IFIP Int. Workshop on Broadband Convergence Networks, 2007.
- [17] www.filesaveas.com/n95.html . Accessed on 6th Sept 2009
- [18] UMTS Forum, Report No. 5, September 1998 (http://www.ums-forum.org/component/option.com_docman/task.cat_view/gid, 152/Itemid,98/). Access in April 2010.
- [19] A. H. Sadka, “Compressed Video Communications ” book, John Wiley, 2002.
- [20] B. Walke, P Seidenberg, M. P. Althoff, “UMTS: the fundamentals” John Wiley 2003.
- [21] M. Loetscher, 2003. Simulative Performance Optimization for TCP over UMTS”, Thesis, Swiss Federal Institute of Technology, Zurich 2003.
- [22] I. Forkel, H. Klenner, “High Speed Downlink Packet Access (HSDPA): A means of increasing Downlink Capacity in WCDMA Cellular Networks”, Communication Networks, Aachen University, Germany.
- [23] <http://www.opnet.com> accessed in Jan. 2011.
- [24] http://www.opnet.com/solutions/network_rd/modeler.html, accessed in Feb 2011.
- [25] OPNET, introduction, OPNET Basic Modeler Tutorials, OPNET Technologies, Inc, Bethesda, MD, USA. 2006.
- [26] H. Asgari, P. Trimintzios, M. Irons, R. Egan, G. Pavlou, “Building Quality of Service Monitoring System for Traffic Engineering and Service Management” *Journal of Network and Systems Management*, vol. 11, issue 4, pp. 399-426, Dec. 2003.
- [27] OPNET, Basic Processes, OPNET Basic Modeler Tutorials, OPNET Technologies, Inc, Bethesda, MD, USA. 2006.
- [28] 3GPP, Release 99, December 1999 (<http://www.3gpp.org/>) . Access in Feb 2011

Review of Wireless Technology

- [29] Cisco System, Specialised Models User Guide-Software Release 11.5, San Jose, CA, USA, 2005 (<http://www.cisco.cm>). Access in Mar. 2011
- [30] J. W. Mark, W. Zhuang, “Wireless Communications and Networking”, Prentice Hall, New Jersey, 2002.
- [31] D. Sebastiao, “Algorithms for Quality of Service (QoS) in a Wi-Fi Network, MSc. Thesis, Instituto Superior Tecnico, Lisbon, Portugal, Dec. 2007.
- [32] J. Manuel, “Panning of UMTS Cellular Networks for Data Services Based on HSDPA”, MSc Thesis, Instituto Superior Tecnico, Lisbon, Portugal, June 2006.
- [33] P. Das, “Quality of Heterogeneous Services with Distributed Resource Management for a WCDMA Uplink, PhD Thesis, University of Newcastle, Callaghan, Australia, Dec. 2006.
- [34] E. Damoso, L. Correira, “Digital Mobile Radio Towards Future Generation, COST 231 Final Report, 1999, (<http://www.lx.it/cost231>). Access in April 2011
- [35] T. Svensson, A. Popescu, “Development of Laboratory Exercises on Opnet”, MSc Thesis, Blekinge Tekniska Hogskola (BTH) University, Sweden, Jun. 2003.

Chapter 3: 3D Video Technology

3.1 Introduction of H.264/AVC-SVC

The H.264/AVC (Advanced Video Coding) standard was developed by the Joint Video Team (JVT) of the Moving Picture Experts Group (MPEG) of the International Standardization for Organization (ISO) and Video Coding Experts Group (VCEG) of the International Telecommunication Union (ITU) [1][2]. H.264/AVC was published by MPEG as MPEG-4 part 10 Advanced Video Coding (AVC) and by ITU-T as ITU-T Recommendation H.264. Several versions of the standards have been released, some versions include new amendments. In particular, the version published in November 2007 refers to the standard that includes the Scalable Video Coding (SVC) Amendment and the version published in march 2009 refers to the standard that includes Multiview Video Coding (MVC) Amendment [3][4]. In this thesis, SVC refers to the scalable extension of H.264/AVC and MVC refers to the multiview extension of the H.264/AVC. They are specified as Amendments in the Annex G and annex H of the H.264/AVC specification respectively [4].

To deliver video over channels with limited bandwidth, coding efficiency is important. High efficiency corresponds to a high video quality with a high bandwidth, or a lower bandwidth with same video quality. As the coding efficiency of the video standards gets higher, the computational complexity required at the decoders becomes a concern for real-time decoding. In addition, video services are provided through lossy channels, so dealing with errors is necessary either at the encoder or the decoder. When these issues are jointly considered, interactivities of the coding layer and the transmission layer are necessary and thus it is required a good system level design for video coding schemes.

3.2 Review of H.264/AVC (Advanced Video Coding)

H.264/AVC includes the coding of the typical chroma subsampling (4:2:0), 8-bit, one-representation of video sequences in the earlier profiles. One extension of H.264/AVC is for high fidelity video, such as high bit-depth per sample, 4:4:2 and 4:4:4 chroma sampling formats. Two other extensions of video coding are for the scalable video applications and multiview video applications, where the video conveyed has, for one scene, different representations either with

3D Video Technology

different SNR (Signal-to-Noise Ratio) and spatial quality, or from different viewing perspectives. These two extensions are Scalable Video Coding (SVC) and Multiview Video Coding (MVC), both of which are part of the H.264/AVC standard [4].

H.264/AVC provides a support for the traditional Discrete Cosine Transform (DCT) plus Differential Pulse Code Modulation (DPCM) codec. As other video coding standards, such as MPEG-2 and MPEG-4 part 2 visual [2][5], a picture is coded as a series of macroblocks (MBs), each of which uses either intra prediction or inter prediction. When inter prediction is used, the previously decoded signal is employed to generate a predicted signal with the help of certain motion vectors. The difference between the intra/inter predicted signal and the original signal of an MB is DCT transformed, quantized and entropy coded.

The features of H.264/AVC are reviewed in [1][2][3][4][5]. A block diagram of H.264/AVC coding is shown in figure 42, where motion estimation are performed as encoding processes, and the decoder performs motion compensation based on the signaled motion vectors to get the predicted signal. Quantization, inverse quantization, transform and inverse transform of H.264/AVC are introduced in [5].

As most of video coding standards, H.264/AVC defines the syntax, semantics and decoding process for error-free bitstreams, any one of which is conforming to a certain profile or level. The encoder is not specified but needs to guarantee that the generated bitstreams are standard decoder compliant. For a video coding standard, the important design considerations are: coding efficiency, decoder complexity and the interactivities with the system.

In the context of video coding standards, a profile corresponds to a subset of algorithms, features or tools and constraints that apply to them; a level corresponds to the limitations of the decoder resource consumption, i.e. decoder memory and computation, which are related to the resolution of the pictures, bit rate and MB processing rate. A decoder conforming to a profile must support all the features defined in the profile. A decoder conforming to a level must be capable of decoding any bitstream that does not require resources beyond the limitations defined in the level. Profiles and levels are helpful for interoperability. For example, during video transmission, a pair of profile and level needs to be negotiated and agreed for a whole transmission session.

3.2.1 H.264/AVC Encoder/Decoder Architecture

In figure 3-1, the DCT and motion-compensated interframe prediction are combined. The coder subtracts the motion-compensated prediction from the input video source to form a prediction error. The prediction error is transformed with the DCT, the coefficients are quantised and these quantised values coded using a variable length coding (VLC). The coded luminance and chrominance prediction error is combined with side information required by the decoder, such as motion vectors (MV) and synchronizing information and formed into a bitstream for transmission.

In the decoder, the quantised DCT coefficients are reconstructed and inverse transformed to produce the prediction error. This is added to the motion-compensated prediction generated from previously decoded video to produce the decoded output

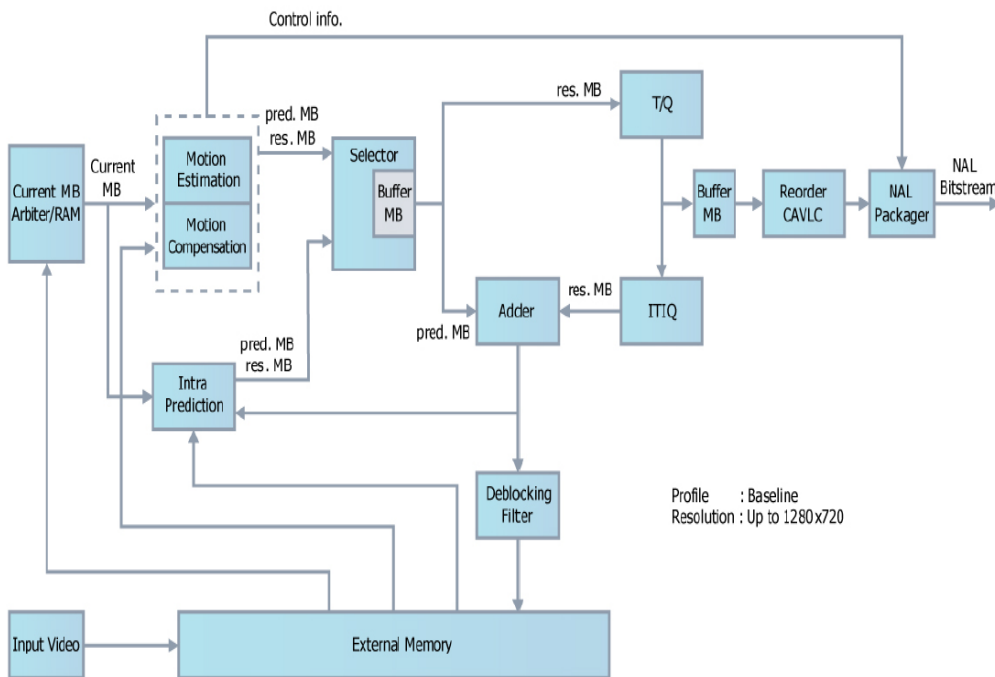


Figure 3-1: Block diagram of H.264/AVC Encoder [5].

The frame buffer stores previous frames to use for motion estimation. In JSVM, the frame buffer size is set before the encoding of a video. The frame buffer must be at least 24Kbits [1]. The ME block determines the best relationships, in terms of a motion vector between the current video frame and the ones stored in the frame buffer. Each macroblock can be split up into sub-blocks,

3D Video Technology

in intra prediction. In H.264, a macroblock may be split into 16x16, 16x8, 8x16, 8x8, 4x8, 8x4 and 4x4 respectively.

3.3 Scalable Video Coding

The problem of sharing video from a single source among a number of users in various settings using existing system and network resources raises new challenges to video coding. In many applications, such as video streaming, video conference, broadcasting, and surveillance, it becomes typical that same video source will be simultaneously sent to many clients in very different environments. A video coding solution being able to adapt to different requirements is desirable for quick and easy video transmission in these applications. [5][6]. For example, in a large video surveillance system such as airport, subway, and highway, there can be more than a thousand network cameras distributed in many important places. Meanwhile, there can also be several hundred monitoring points, such as central control room, on-site monitors, data analysis applications, mobile monitors in vehicles, or handheld devices, need to access any captured or stored video data from the system in real time. The cameras positioned far away can be connected to the system using IP networks, and other cameras and data storage centre may use large bandwidth networks. The connection of monitors to the system may vary from high speed cable, local area network, to wireless transmission. Each of these monitors may require the same video in different resolutions, frame rates, and decoding complexity simultaneously. A universal video solution is important in the above system to adaptively and efficiently meet the requirements.

Heterogeneity among video clients can be grouped into three aspects: network conditions, device settings, and quality requirements. Wireless network normally have smaller bandwidth and higher error rates than wired networks. Even different sub-networks of the same network may have different and varying quality of service. Some applications, such as video conferencing and broadcasting, are more sensitive and have less tolerance to transmission delay. Handheld and mobile devices, such as pocket PC, iPod, and cell phone, usually have less processing power, memory and storage space smaller size screens than computers or other high end video players, and therefore can only afford lower complexity decoding and low quality video. In such a system, low cost quality control coding scheme is highly desirable.

3D Video Technology

Streaming video over heterogeneous networks demands video services of varying spatial, temporal, and quality scalability (SNR). Transcoding a video into several formats to make it suitable for different clients requires special processing of video bit streams to meet end user's device requirements, which is complicated and time consuming. If a server keep several versions of the same video and switch during transmission, it also require additional storage and speedy data access for frequent bit stream switching. This is usually not practical for less powerful servers.

Scalable video coding provides an efficient way to generate adaptive video for transmission in heterogeneous environments [7]. Conventional scalable video coding techniques generate layers of video with different importance to the quality of video and fixed decoding order. In Layered Coding (LC) [10], effect of loss or errors in enhancement layers are limited. A minimum quality and normal decoding process is retained by the base layer. Four types of scalabilities: quality (SNR), spatial, temporal, and complexity are most commonly used and standardized, such as in MPEG, H.26x, etc. However, the base layer content requires stronger protection in order to benefit from scalability. Otherwise, losses in base layer still cause error propagation and render enhancement layers useless.

Scalable Video Coding has been an active research and standardisation area for the last 20 years. Prior international video coding standards such as H.262/MPEG-2 Video, H.263 and MPEG-4 Visual [8] already included several tools by which scalability can be supported. However, the scalable profiles of those standards have rarely been used. Reasons for that include the limitations of traditional video transmission system as well as the fact that the spatial and quality scalability features came along with significant loss in coding efficiency as well as a large increase in decoder complexity as compared to the corresponding non-scalable profiles. Scalable video coding attracts wide attention with the rapid growth of multimedia and media applications over wireless and Internet channels [9][10] [11].

The recently standardized SVC extension of H.264/MPEG4-AVC [12] [13] offers a significantly improved bit rate, distortion and complexity trade-offs. But the deployment of a new video coding technology is also dependent on the business environment in which it is planned to be used. SVC offers a number of features that include efficient methods of graceful degradation, bit rate adaptation and format adaptation.

3D Video Technology

In general, a video bit stream is called scalable when parts of the stream can be removed in a way that the resulting sub-stream forms another valid bit stream for some target decoder, and the sub-stream represents the source content with reconstruction quality that is less than that of the complete original bit stream but is high when considering the lower quantity of remaining data. Bit streams that do not provide this property are referred to as single-layer bit streams. The scalability in the context of video coding specifically corresponds to bitstream scalability, while a sub-stream can have a relatively lower representation with lower spatial resolution (Spatial Scalability), lower frame rate (Temporal Scalability) or lower SNR (SNR/quality scalability). In other words, a scalable bitstream consists of a compressed video content hierarchically organized in successive layers, corresponding to different levels of image quality, frame rate, and picture size. Besides these three typical dimensions, other dimensions that have been taken into consideration include complexity, colour bit-depth and chroma sampling format.

It is ideal to design a coding scheme with similar or better performance than H.264/AVC, but offering scalabilities with an equivalent or slightly higher complexity. Most widely exploited way is 3D wavelet coding. The principle is to apply the wavelet decomposition directly on the 3D spatio-temporal signal composed of a group of successive video frames. The 2D spatial decomposition is also applied to realize spatial scalability. MCTF (Motion Compensated Temporal Filtering) [14] is performed for each pixel of the reference frame of the group along its motion trajectory, to support temporal scalability. Wavelet based video coding uses embedded entropy coding based on embedded zero tree wavelet (EZW) and embedded block coding with optimized truncation (EBCOT) [14] algorithms, to get a SNR, spatial and temporal scalable stream. However, wavelet coding techniques, including those mentioned, have not been publicly standardized for video coding, although MPEG-4 part 2 has adopted wavelet coding for still texture.

In MPEG-2, SNR scalability is achieved by simply re-quantising the quantisation errors of the base layer. In MPEG-4 visual, SNR scalability is realized by Fine Granularity Scalability (FGS), where bit plane coding is used to generate several enhancement layers. However, each FGS enhancement layer can be truncated into any number of bits within each frame to provide partial enhancement proportional to the number of bits decoded for each frame. For spatial scalability, MPEG-2 up-samples the base layer reconstruction picture and makes it as one candidate for reference pictures.

3D Video Technology

Temporal scalability can be naturally supported by dropping the coded B pictures in MPEG-2 and MPEG-4 bitstreams. Besides, both of these standards provides the “base layer” and “enhancement layer” concepts to code two temporal layers, where the enhancement layer pictures can choose, for each prediction direction, a picture either from the base layer or the enhancement layer as a reference.

In recent years, Multiple Description Coding (MDC) techniques are developed to combat the errors in best effort networks by eliminating the error free assumption in the base layer of Layered Coding (LC). MDC supports equally important and independently decodable video bitstreams that are generated from a single video sequence. Errors in one description may not affect decoding of others. A video with higher fidelity can be obtained by combining any subset of the descriptions. Multiple Description (MD) is therefore considered providing parallel scalabilities to video sequences. Approaches for different types of scalabilities have been studied and reported, such as MD Scalar Quantisation, MD Transform Coding, MD Motion Compensation, and Multi State Video Coding (MSVC). These MDCs will be discussed in details in chapter 4. The first two approaches make use of spatial correlation of images and extend the MDC algorithms from image coding to video coding. MSVC mainly considers temporal redundancy between adjacent video frames. MDMC explore both redundancy in spatial and temporal domains and thus is able to control quality and distribute redundancy smoothly.

3.3.1. Layered Video Coding

Layered video coding generates several bit streams and arranges them in a number of layers. Generally there is a base layer and one or more enhancement layers. The idea is that the base layer should carry the most vital information and the enhancement layers carry the residual information that enhances the quality of the base layer video. At a video server, only one copy of the same video with scalabilities is necessary. Transmission bandwidth shall be guaranteed for the base layer and best-effort for enhancement layers.

According to the way that enhancement layers improve quality of base layer, there are three types of scalabilities: spatial, temporal and quality, each provides enhancement to the base layer in form of higher spatial resolution, temporal frame rate, and visual quality, respectively. MPEG-2 and H.263+ accept all three scalabilities, and the quality scalability is SNR scalability in both standards [4][5]. Theoretically, there can be as many layers as possible in each type of

scalability. But a large number of layers do not bring enough improvement in performance, considering dramatically increased complexity and overhead. In practice, only two or three layers are commonly used for performance-cost tradeoff.

In all the above-mentioned scalabilities, transition steps of video quality with different number of enhancement layers are not small, i.e video quality will jump noticeably in present or absence of an enhancement layer. A newly proposed scalable video coding method, is Fine Grain Scalability (FGS), enables small step of increase of video quality. It was standardized in MPEG-4 instead of SNR scalability [5].

3.3.2. Temporal Scalability

In each spatial layer, the temporal scalability is achieved by the motion compensated temporal filtering (MCTF) technique, which performs the wavelet decomposition/reconstruction along the motion trajectory [16]. The MCTF is mostly restricted to the short-length wavelet, which can be implemented by a lifting scheme with only one prediction step. In this case, the prediction and update can be realized using bidirectional. Inside the MCTF, an odd-indexed frame is predicted from the adjacent and even-indexed frames to produce the high pass frame. Accordingly, an even index frame is updated using the combination of adjacent high-pass frames to generate the low pass frame. To remove temporal redundancy, motion compensation is conducted before the prediction and update. By using n decomposition stages, up to n levels of temporal scalability can be provided.

3.3.3 Spatial Scalability and inter-layer prediction

For the spatial scalability, sequences of different spatial resolutions are coded in separated layers. To remove the redundancy among different spatial layers, the residues and motion vectors of an enhancement layer are predicted from the ones of the subordinate layer. In the prediction process, the residues and motion vectors of the subordinate layer are firstly interpolated if the subordinate layer has a lower resolution [17]. In addition, the partition of an inter MB can be derived from the relevant sub-blocks at the subordinate layer and the motion vectors can be obtained by refining and scaling the ones from the corresponding sub-blocks. On the other hand, for an intra MB, the inter-layer prediction is allowed only if the corresponding 8x8 block of the subordinate layer is within an intra coded MB.

3.3.4 SNR Scalability

For the SNR scalability, the residues after the inter-layer prediction are transformed with the 4x4 integer transform. Then the transform coefficients are successively quantized into multiple quality layers. The coefficients in a quality layer are coded by a hybrid approach of bit plane (Run, level) coding. For the scalability with fine granularity, the bit planes in a quality layer are coded using a cyclic block coding. The coding step is partitioned into the significant and refinement passes. The significant pass first encodes the insignificant coefficients that have zero values in the subordinate layers. Then, the refinement pass refines the remaining significant coefficients ranging from -1 to +1. During the significance pass, the transform blocks are coded in a cyclic and block-interleaved manner. On the other hand, the coding of refinement pass is conducted sub-band by sub-band [18]. To further reduce bit rate, a context adaptive binary arithmetic coder is applied (CABAC).

3.3.5 Fine Grain Scalability

A special scalability for progressive video transmission was introduced in [5] and adopted in MPEG-4. The so called Fine grained Scalability (FGS) provides smooth quality changes when video transmission suffers to channel conditions and thus very suitable for progressive video streaming.

FGS video coders in MPEG-4 have similar coding structures to quality scalable video coders. A base layer is also generated by coarsely quantizing the original video sequence. But the enhancement layers are coded using bitplane scan and shifting instead of conventional image coding. A bitplane is composed by a block of DCT coefficients of the enhancement layer pictures, ordered from significant bits to non-significant bits. In essence, the most significant bitplane is first coded and the least significant bitplane is last. Video server will send out the base layer and as much as possible bit-planes in the enhancement layer, according to availability of sources. Decoded video will guarantee a basic quality by the base layer, and be successively refined by received bitplanes. In practice, the FGS enhancement-layer encoder can be based on any fine-granular coding methods, such as wavelet methods. The enhancement layer signals can also be use in motion-compensated prediction in the base layer for better compression gains, if error resilience is not a concern.

Combinations of some scalabilities mentioned in this section will increase range of applications due to higher level of adaptivity. Hybrid scalability is more efficient and effective for video communication over different types of networks.

3.4 Scalable Extension of AVC/H.264

This section proposes the combination of two video coding aspects: multiple description coding and scalable video coding. Scalable MDC aims at combining the flexibility of scalable coding with the robustness of MDC. The proposed scalable MDC begins first with a standard H.264/SVC coder and the error resilience of the base layer of H.264/SVC is enhanced using temporal MDC [12]. The temporal MDC of the base layer is produced using the multi state video coding approach in [13], which separates the even and odd frames of a sequence into two MDC stream as will be discussed later in chapter 4.

Figure 3-2 shows the general block diagram of the H.264/SVC for stereoscopic 3D video. The depth data is combined with the texture data using a depth image base rendering technique [14] to produce left and right views and which is placed in the enhancement layer.

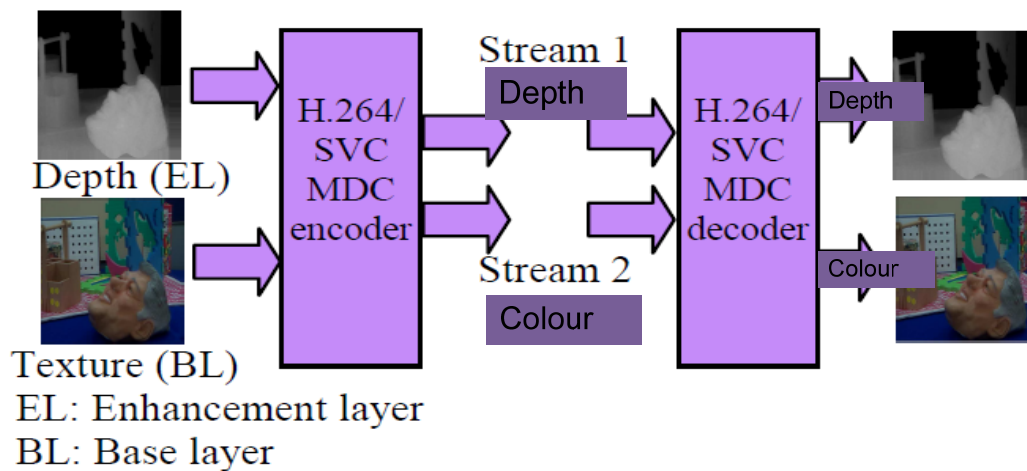


Figure 3-2: Scalable H.264/AVC MDC [15].

H.264/AVC can produce scalable layers that can be exploited for MDC. In [15], H.264/SVC used in a multicast scenario with interlayer prediction switched off. In the scalable coding, the even and odd frames are separated before the encoding process is preferred for both texture and depth. The even frames for the texture are coded in the base layer (layer 0) and the odd frames

3D Video Technology

for the texture are coded in the enhancement layer (layer 1). With the interlayer prediction switched off, it can be assumed that layer 1 is also the base layer for the scalable MDC. The even frames for the depth are coded in the enhancement layer (layer 2) and the odd frames for the depth are coded in the enhancement layer (layer 3). Table 3-1 shows the description of the layers that can be produced by the scalable MDC for a Quarter Common Intermediate Format (QCIF) pixels size image sequence for only one spatial resolution. Example with spatial scalability is given in sub-section 3.4.1.

Table 3-1: Description of layers for the scalable MDC encoder

Layer	Resolution (Pixel)	Description
0	176x144	Base layer, even-Texture
1	176x144	Base layer, odd-Texture
2	176x144	Enhancement layer, even-Depth
3	176x144	Enhancement layer, odd-Depth

A single standard compliant bit stream is produced from the above configuration. At the decoder, a bit stream extractor is used to extract the even and odd bit streams of the texture and the depth. Each bit stream can then be decoded by the standard H.264/SVC decoder. Finally, the decoded even and odd frames are merged together to produce a full resolution decoded sequence.

For both texture and depth, if both the even and odd streams are received, the decoder can reconstruct the coded sequence at full temporal resolution. If only one stream is received, the decoder can still decode the received stream at half the original temporal resolution. Since the even frames are predicted from previous even frames i.e independent from odd frames, there will be no mismatch if one of the streams is lost at the decoder.

3.4.1 Spatial scalable results

Table 4 below shows an example of the layers configuration which produces spatial scalable result of the proposed scalable MDC for Orbi sequence in figure 3-3. All layers are contained in one single bit stream. Layer 0 and Layer 1 are the MDC layers (base layer). The user can select to decode the required bit stream based on their terminal requirement. If a stereoscopic 3D terminal is present, the user can also decode the depth layer and used a Depth Image Based

3D Video Technology

Rendering (DIBR) to achieve stereoscopic 3D video. Table 3-2 showed different layers in QCIF and CIF format of the same video sequence.

Table 3-2: Description of layers

Layer	Resolution	Description	Grouping
0	176x144	Base layer, even-Texture	Texture
1	176x144	Base layer, odd-Texture	
2	176x144	Enhancement layer, even –Depth	Depth
3	176x144	Enhancement layer, odd-Depth	
4	352x288	Enhancement layer, even-Texture	Texture
5	352x288	Enhancement layer, odd-Texture	
6	352x288	Enhancement layer, even-Depth	Depth
7	352x288	Enhancement layer, odd-Depth	



(a)



(b)

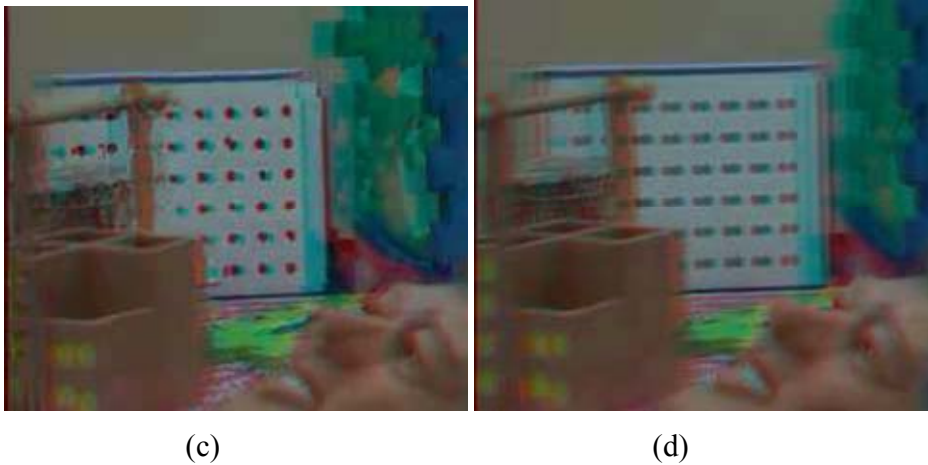


Figure 3-3: Subjective results for frame 15 of the orbi sequence for 2D video: (a) SDC, (b) MDC, and for stereoscopic 3D video: (c) Single Description Coding (SDC), (d) MDC when subjected to 10dB SNR UMTS channel.

3.5 Overview of 3D Technology

3.5.1 Introduction

Three-dimensional is believed by many to be the next logical development towards a more natural and life-like visual home entertainment experience [16]. Although the basic technical principles of stereoscopic TV were demonstrated in the 1920s by John Logie Baird, the step into the third dimension still remains to be taken. The many different reasons that have hampered a successful introduction of 3D TV so far could now be overcome by recent advances in the area of 3D display technology, image analysis and image-based-rendering (IBR) algorithms, as well as digital image compression and transmission. Building on these latest trends, a modern approach to three-dimensional television is described in the following, which aims at fulfilling important requirements for a future broadcast 3D TV system.

- Backwards- compatibility to today's digital 2D colour TV.
- Low additional storage and transmission overhead.
- Support for auto-stereoscopic on depth reproduction.
- Simplicity in producing sufficient, high-quality 3D content.

Several attempts have been made to introduce this technology into the market. Despite the immense desire towards 3D, the great expectations of viewers, content providers and distributors

3D Video Technology

were not fulfilled. The main disadvantages were the discomfort of the viewers such as headache and eye strain due to the poor quality content, the low-tech display systems and the high costs involved in the production and distribution of 3D content.

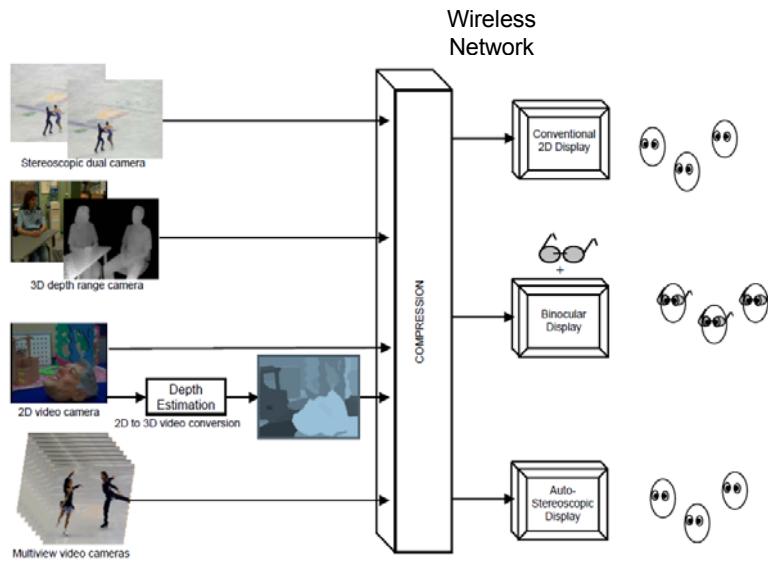
In spite of these very successful early demonstrations, it was not until the 1950s that Hollywood is at the forefront, leading the 3D technology revolution with a vested interest in seeing 3D succeed. Studios and content providers are aiming at an unprecedented 3D quality. Hollywood has pioneered many new 3D projects, including the gathering of various stakeholders, defining standards and funding research. 3D films have greater returns due to fees and charges, since consumers have proved to be willing to pay more for the enhanced 3D experience. Hollywood's revenue comes from home entertainment sales and rentals [17].

Recently, 3D TV has received tremendous attention among the research communities and technology developers. The showcase of advanced immersive 3D displays by major TV manufacturers in consumer electronics trade shows and the production of compelling 3D movies by Hollywood are evidence that the dream of watching 3D TV at home is not far from reality.

The sports industry also has a keen interest in the development of 3D technologies. 3D is especially applicable to sports broadcasts since it adds an immersive experience, allowing viewers to feel like they are in the actual stadium. A selling point for tickets to sporting events is the atmosphere, action and pace of the game, which cannot be recreated at home on normal television. 3D technology is much better equipped to provide a new perspective of sports in action and bring a stadium-like experience to the household [18]. 3D broadcasts are also attractive for sports leagues since they provide an alternate revenue stream, which can be priced at a premium.

Over the years, a consensus has been reached that a successful introduction of 3D TV broadcast services can only be a lasting success if the perceived quality and the viewing comfort are better than those of conventional 2D television. This is becoming increasingly feasible because of the recent advances in capturing, coding and display technologies [19]. The three key components of a future 3D broadcast chain are shown in figure 3-4.

3D Video Technology



Content generation Video coding/Compression Display technologies

Figure 3-4: 3D broadcast chain [28].

3.5.2 Human Visual System (HVS)

Human depth perception is based on a combination of many visual cues as well as internal mental templates and expectations. For most people, 3D depth perception is realized by two slightly different images projected on the left and right eye retinas (binocular parallax), each represents a slightly different viewpoint. The brain fuses the two images to give the depth perception. The viewer then sees one solid scene instead of two slightly different projections as shown in figure 3-5. The perceived image with depth contains all the information present in the two individual viewpoint images. It also conveys something else that is not present in either of them: an intrinsic feeling of depth, distance and solidity. The differences between the left and right eye viewpoint images arise because an object in a scene will not fall in the same spot in both images. This relative displacement is according to the object's distance from the viewer is called disparity.

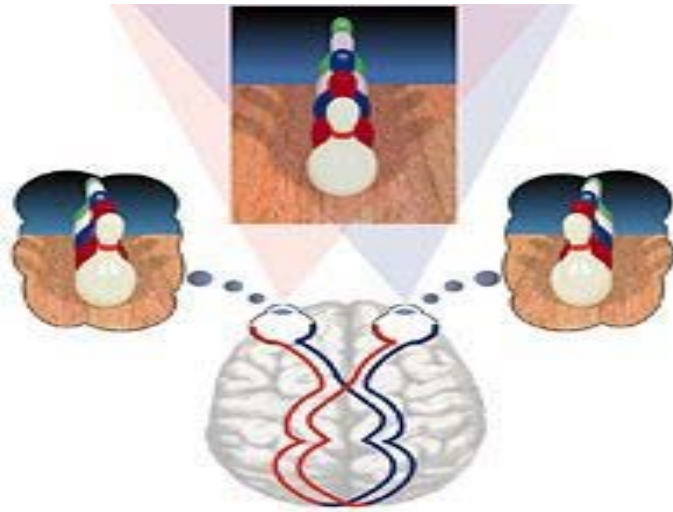


Figure 3-5: 3D visual depth perception (<http://www.strabismus.org>)

Although the binocular parallax is the most dominant cue for depth perception, there exist many other depth cues known as monocular depth cues, which do not require the observer to have two eyes to perceive depth. Over the years, the human brain has been trained to perceive depth using these cues. A number of cues are listed below:

- **Relative size:** If two objects in a scene are known to have the same size but located at different distances from the observer, then the projection of the near object onto the observer's retina will be larger than the projection of the far object.
- **Motion parallax:** The relative motion between the viewing camera and the observed scene provides an important cue to depth perception, near objects moves faster across the retina than distant objects do. This motion may be seen as a form of disparity over time.
- **Occlusion:** The principle of depth from occlusion has its roots in the phenomenon that an object which overlaps or partly obscures our view of another object is considered to be closer. Occlusion offers rich information in relative depth ordering of the objects.

3.5.3 3D content generation

Hollywood has spearheaded efforts to implements standards through the Society of Motion Picture and Television Engineers (SMPTE) in the early days of 3D. SMPTE has created then a task force to discuss possible standards for 3D content. The goal has been to make 3D system backward compatible with present 2D system, scalable and deliverable over cable, wireless,

3D Video Technology

satellite, disk and Internet. SMPTE delivered their report in April 2009. The report defined guidelines for the mastering standard, which will be used by 3D developers. Using these guidelines, SMPTE hopes to define the core standards by the summer of 2010 [20][21]. SMPTE hoped to work with other standardization bodies to develop the standards for complementary products, to ensure compatibility from end to end perspective [22][23]. It is expected that the interoperability standards are implemented across the industry within two years [23].

There are four types of 3D content generation, the stereoscopic dual-camera approach, which results in two separate views (left and right), the depth-range camera approach, which generates a 2D image plus a depth map, the 2D-to-3D video conversion approach, which converts existing 2D video into stereoscopic 3D by estimating a depth map from the 2D video sequence and then rendering the left and right sequence and last the multiview video camera approach [24][25][26]. These are shown in the figures 3-6, 3-7 and 3-8:

3D Video Technology



Figure 3-6: (a) Stereoscopic camera set up (b) Stereo images (Left & Right) [27]

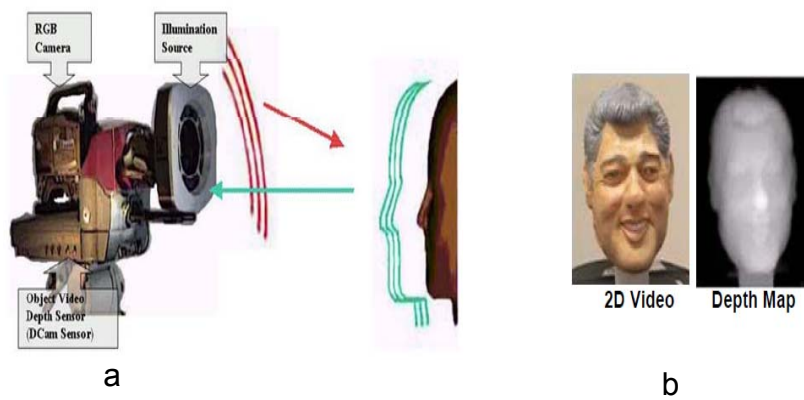


Figure 3-7: (a) 3D Depth-range camera, (b) 2D luminance and depth [27].

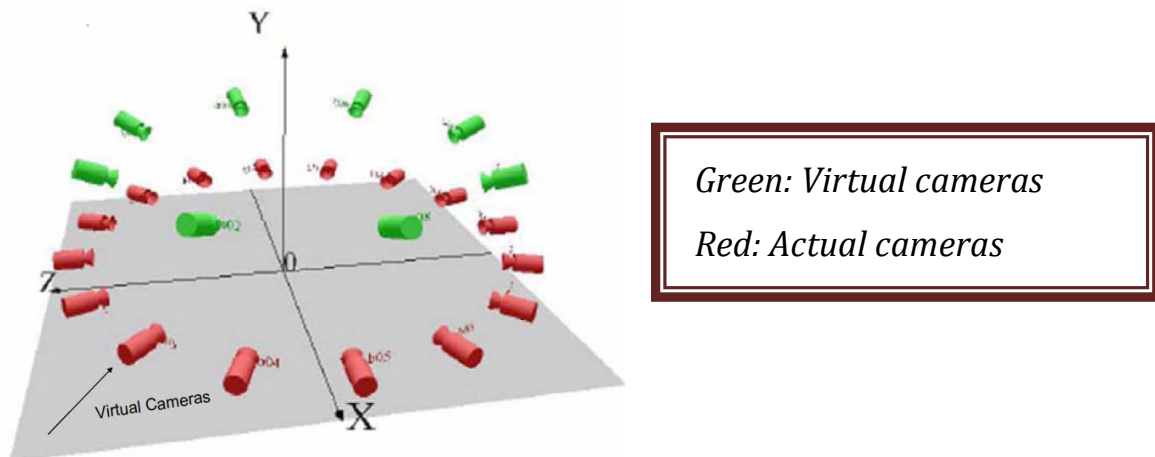


Figure 3-8: Multiview video camera configuration [27].

3.6 3D Display

Displaying 3D content is the last part in the 3D broadcast chain. 3D display falls outside the scope of this thesis; we feel that a short overview of this topic will help the reader obtain a better understanding and appreciation of the present status of the 3D technology. The 3D display is divided into two categories: Binocular with active or passive glasses, and auto-stereoscopic without glasses displays. The following sub-sections describes the different 3D displays reported in the literature.

3.6.1 Colour filtered anaglyph

Anaglyph is one of the first commercial methods for displaying 3D. In anaglyph displays, the left and right eye images are filtered with near-complementary colours such as red and green and red and blue [28][29][30]. The right and left eye images are superimposed over each other. Figure 3-9 shows this process. The viewers are required to wear colour filter glasses to filter the images and perceived depth.



Figure 3-9: Anaglyph glasses and anaglyph image [27].

This is a well known and inexpensive methods used for stereoscopic cinema and television, and is still popular for viewing stereoscopic images. Limitation of this method is that colour information is lost since it is used as a selection mechanism. Only limited colour rendition is possible through the mechanism of binocular colour mixture. The other disadvantages of this system are crosstalk. Crosstalk in 3D displays results in eyes seeing the wrong view (left eyes sees the right view image and vice versa). On the actual display crosstalk is seen as double contour otherwise called ghosting and is a potential cause of eyestrain and headaches [28][29].

3.6.2 Polarized glasses

Polarization-based displays separate left and right eye images by means of polarized light. Left and right output monitors or projectors are covered by orthogonally oriented filters, using either

3D Video Technology

linear or circular polarization. The polarized stereo images are projected and superimposed onto the same screen. The observer needs to wear polarized glasses to separate the different views again. When watching with glasses, since each lens passes only the light that is polarised in its polarizing direction and blocks the light polarized in the opposite direction, each eye sees its matching image and the observer perceives depth effect.

Linear polarized glasses use vertical polarization on one lens and horizontal polarization on the other (See figure 3-10). The 3D effects are perceived as long as the user's head is kept straight. By tilting the head, the 3D effect will break and some amount of ghosting or crosstalk may occur.

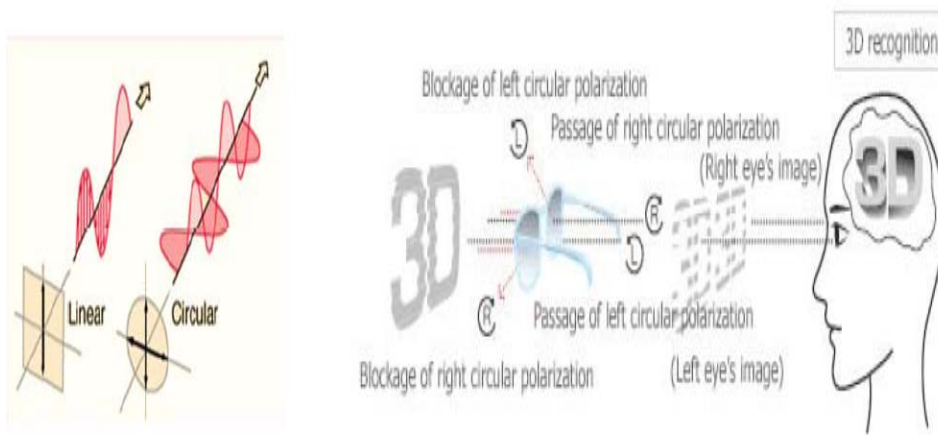


Figure 3-10: Linear and circular polarizations (<http://www.zalman.com>)

3D Video Technology

Circularly polarized lenses are polarized clockwise for one eye and counter clockwise for the other. This is shown in figure 51. This method of polarization will maintain the 3D effect even if the head is tilted. The polarized-based display system offers good quality stereoscopic image with full colour display at full resolution, and very little crosstalk in the stereo-pairs [28]. It is the system most commonly used in stereoscopic cinemas nowadays. Major drawback of this system is the loss of light output due to the use of polarizing filters which is more evident in the circular polarization.

3.6.3 Spectrum Filtered-Dolby 3D

Dolby 3D uses Infitec technology which stands for interference filter technology. This system encodes left and right images by projecting each with a differently filtered spectrum of light. In this case the light is filtered differently for each view, but both the left and right spectrums appear as white light or near-white light as shown in figure 3-11.

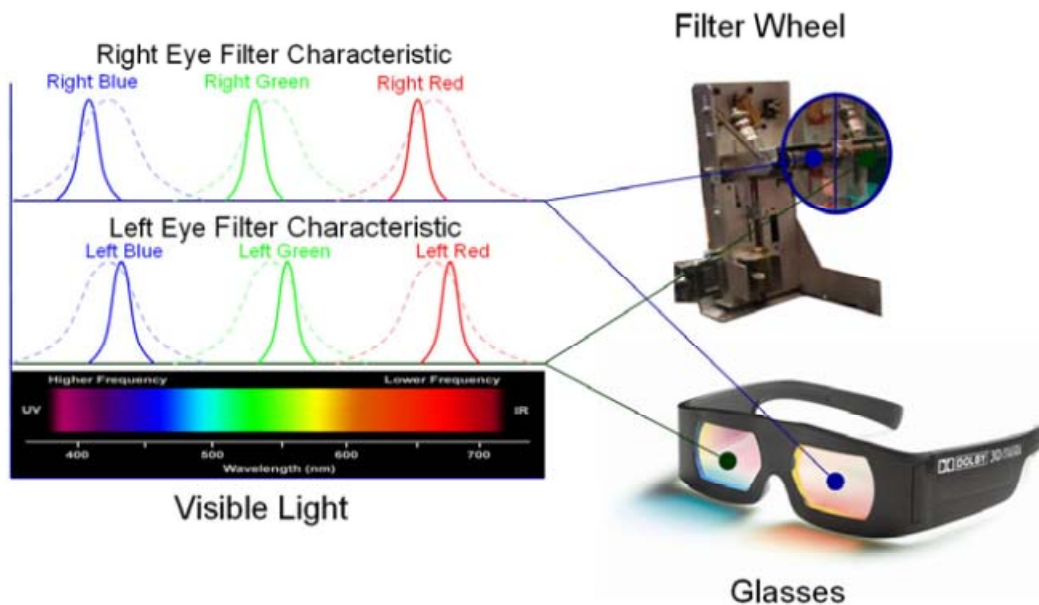


Figure 3-11: Dolby 3D [29].

This differentiates Infitec from the anaglyph method which uses red filters for one eye and blue filters for the other eye. In Dolby's implementation, the light path in the projector is modified with a filter wheel to achieve spectral division of the stereoscopic images. Prior to projection,

3D Video Technology

some colour balancing is applied to the image signal inside Dolby's digital cinema server. Complementary spectral division glasses are worn by audience members for decoding the images so that left eye images are seen only by left eye and right eye images are seen by only the right eye. To accomplish this, Dolby's glasses employ some 50 layers of thin-film coating to create the appropriate optical interference filters.

3.6.4 Binocular with active glasses

Shutter glasses are the most commonly used active 3D glasses. The lenses of shutter glasses are actually small LCD screens. When voltage is applied, the shutters close and the lens go dark. This behaviour is synchronized with the screen displaying the 3D content, usually through an infrared transmitter. When the viewer looks at the screen through shuttering eyewear each shutter is synchronized to occlude the unwanted image and transmit the wanted image. Thus, each eye sees only the appropriate perspective view. The left eye sees only the left view, and the right eye only the right view. On an LCD or LED television, this method of 3D displaying cuts the refresh rate in half and has been known to cause headaches for many people. To eliminate this problem, new display systems use a refresh rate double that of conventional displays (120Hz rather than 60Hz).

3.6.5 Auto-stereoscopic Displays

Auto-stereoscopic displays apply optical principles such as diffraction, refraction, reflection and occlusion to direct the light from the different perspective views to the appropriate eye [28], allowing multiple users to watch 3D content at the same time without wearing specialized 3D glasses mentioned above. This property makes them the best candidate for future consumer 3D TVs. One of the disadvantages of this display is that the resolution for each view drops as the number of views increases. The arrival of high resolution flat panel displays has made multiview applications more feasible [30]. The other important disadvantage of these displays is the fact that only under limited horizontal viewing angle the picture will be perceived correctly.

The two most important auto-stereoscopic techniques are based on parallax barriers and lenticular arrays. These techniques are being researched here in CMCR lab of Brunel University, under the VIVANT project [31].

3.6.5.1 Parallax barrier

Parallax barrier displays are based on the principle of occlusion, where part of the image is hidden from one eye but visible to the other eye. Figure 3-12 show this, at the right viewing distance and angle, each eye can only see the appropriate view, as the other view is occluded by the barrier effect of the vertical slits. Different implementations of this method are available, including parallax illumination displays where the opaque barriers are placed behind the image screen and moving slit displays which use time-sequential instead of stationary slits. The main advantage of these displays is their backward compatibility with a 2D display technology.

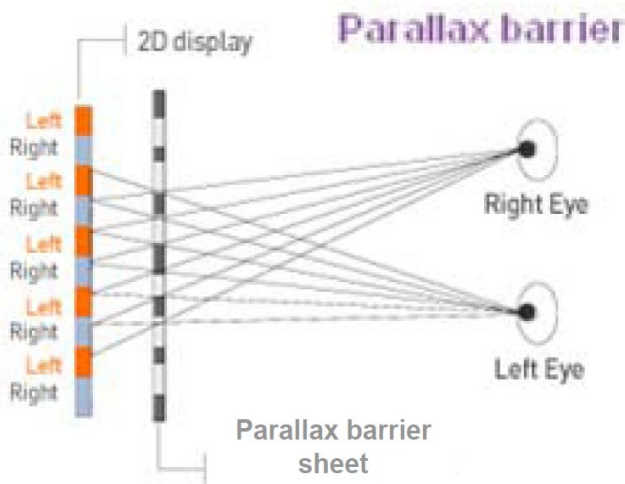


Figure 3-12: Parallax barrier displays [28].

3.6.5.2 Lenticular Lens

Lenticular technology is based on the principle of refraction. As it can be seen in figure 3-13, instead of using vertical gratings as with parallax barrier displays, an array or sheet of vertically oriented cylindrical lenses is placed in front of columns of pixels, representing parts of the left and right eye view. Through refraction, the light of each image point is emitted in a specific direction in the horizontal plane.

3D Video Technology

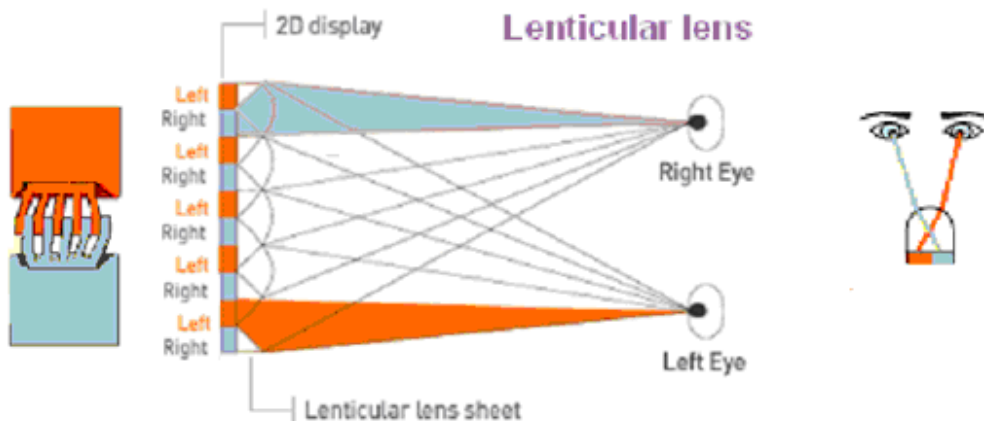


Figure 3-13: Lenticular lens displays [28].

In what is known as the sweet spot of a display, left and right images can be delivered to the corresponding eye to create a 3D effect. Older, less sophisticated systems required the viewer to sit at a specific distance and angle in order to properly view the image and avoid headaches and eyestrain. Current lenticular lens systems have corrected this by using a slanted lenticular sheet, allowing up to eight viewers to observe a 3D image with no ill effects.

3.6.6 Other methods for 3D video generation (Using Matlab)

H.264 Encoder can encode other video files to H.264/AVC format, with this very powerful tool; we can create an H.264 movie. This package is easy to use and helps in creating high quality Video files. The high quality and efficiency of H.264 bring video to life on Internet or mobile phones. The extension of this encoder is used for 3D video compression/3D encoding format. It can encode 3D videos with approximately 3 times fewer bits than MPEG encoders. This software offers fast encoding speed and high quality output. Figure 3-14 shows the encoder settings. Figures 3-15 to 3-20 shows some of the results obtained.

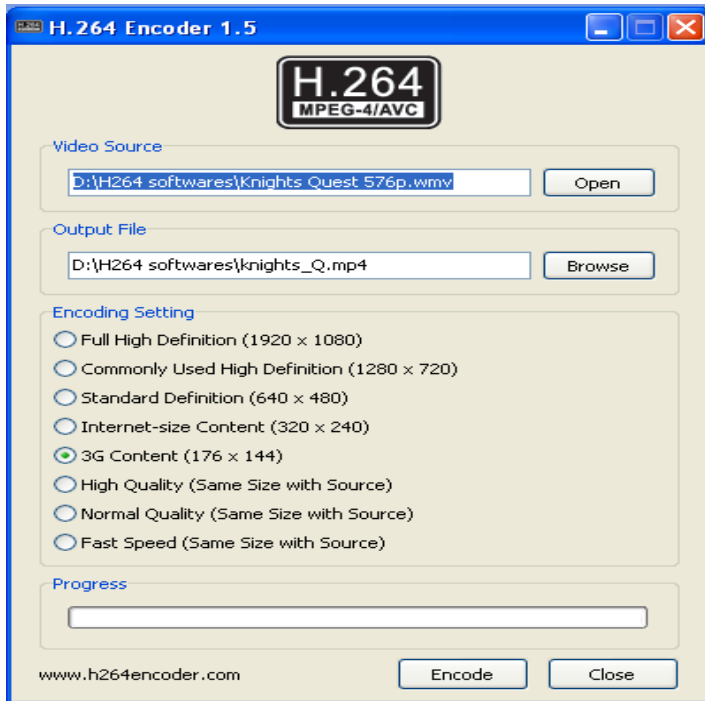


Figure 3-14: H.264/AVC Encoder

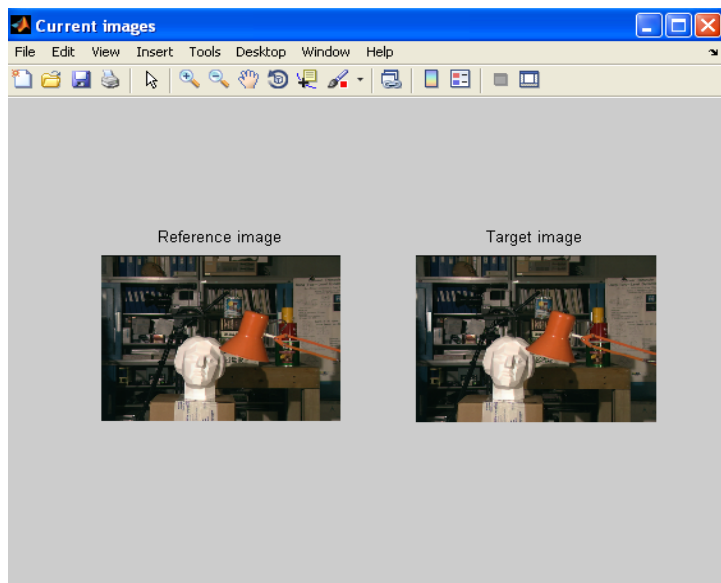


Figure 3-15: Combined left and right images

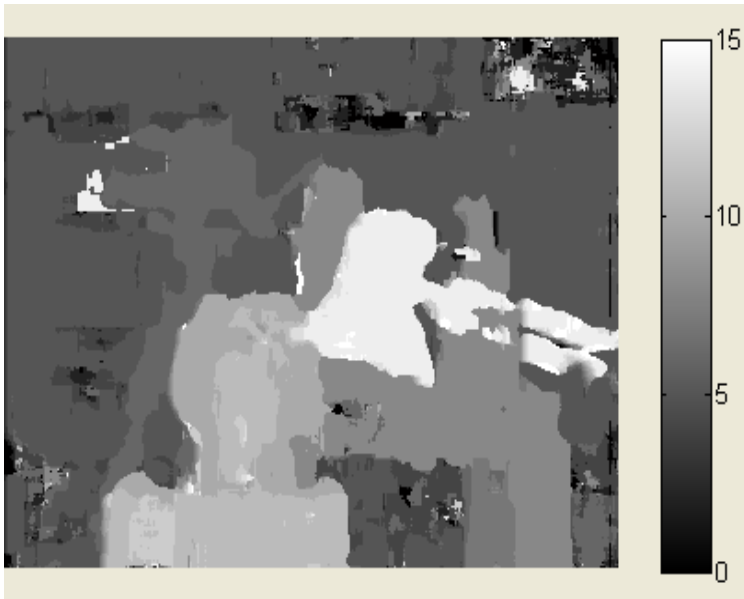


Figure 3-16: Generated depth map



Figure 3-17: Reconstructed 3D

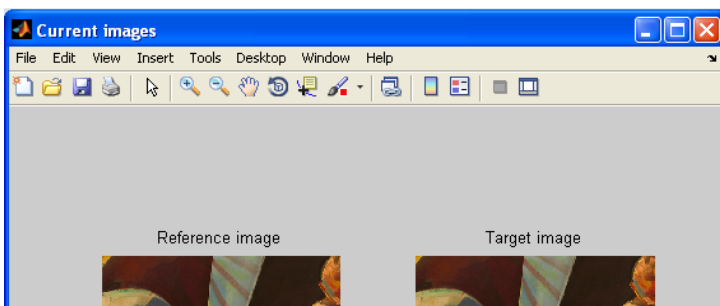


Figure 3-18: Combined Left and Right images

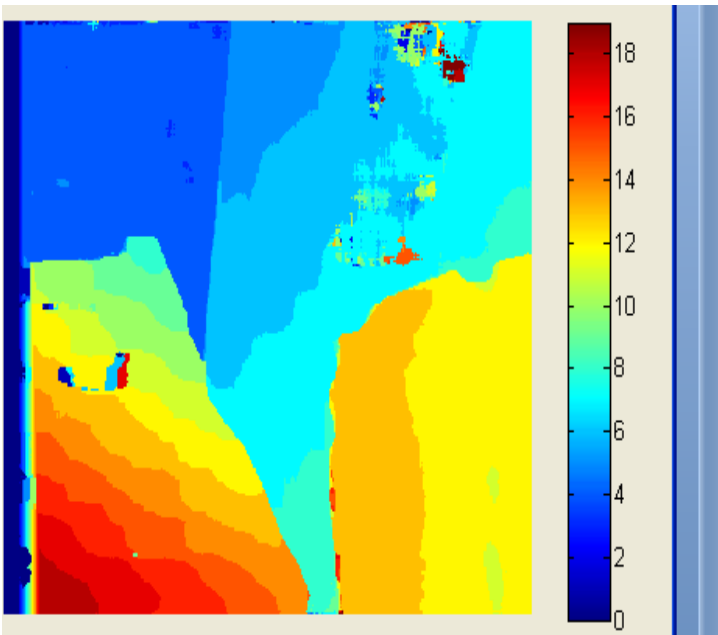


Figure 3-19: Depth Map Generated



Figure 3-20: Generated 3D

3.7 Summary

In this chapter, we have reviewed and analysed the Advanced Video Coding as described in the standard documents. We also presented latest development in H.264/AVC SVC. The H.264/AVC is capable of offering a fully scalable bit stream. Although the AVC based approach has been selected as the working draft in MPEG. We also reviewed the different scalabilities such as Temporal Scalability, Spatial Scalability, SNR Scalability and Fine Grain Scalability.

A standard compliance scalable MDC based on an odd and even frames is also presented for the H.264/SVC video coding targeting 3D video applications. It generates two descriptions for the base layer of SVC based on even and odd frame separation, which reduces its coding efficiency. The proposed scheme also achieves scalability through the layered coding of SVC. The scalable video sequence is composed of a so called base layer and one or more enhancement layer (s). The base layer is self-contained and fully decodable to a video of lower quality and/or lower resolution. Enhancement layers, on the other hand, cannot be decoded if the base layer is lost or damaged. They can only be used to improve the overall quality as described in section 3.4. We also conclude from section 3.4 that, scalability not only allow a stream to be sent over channels

3D Video Technology

with different bandwidth constraints or to devices having different capabilities, but also allow for different error protection schemes to be applied. The architecture of the encoder/decoder is shown in figure 43. Description of the different scalable layers was also presented. Subjective results of the scalability for Orbi sequence frame 15 was presented in figure 44.

Overview of 3D technology was introduced in section 3.5, effect of human visual system (HVS) on 3D perception was discussed, and efficient method for 3D content generation was presented. Then, we presented different display technologies.

Finally, extension of H.264/AVC to 3D video was presented. We used Matlab simulation to combine the left and right images (Stereoscopic images), we also generated the depth using Matlab and finally generated the 3D video using the same procedure. More results are presented in appendix B.

References

- [1] ITU-T Recommendation H.262/ISO/IEC International Standards 13818-2; 2000, “Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video”.
- [2] ISO/IEC International Standard 14496-2: 2001, “Information Technology – Coding of Audio-Visual Objects –Part2: Visual”.
- [3] ITU-T Recommendation H.264 (April 2003), “Advanced Video Coding for Generic Audio Visual Services”.
- [4] ISO/IEC JTC1/SC29/WG11, “Applications and requirements for scalable video coding”, Technical Report N5540, International Organisation for Standardisation, March 2003.
- [5] ISO/IEC JTC1/SC29/WG11. Applications and requirements for scalable video coding. Technical Report N6880, International Organisation for Standardisation. January 2005.
- [6] J. Ohm, “Advances in scalable video coding”, Proc. of IEEE International Conference on Image and Signal Processing, vol. 93, no. 1, pp. 42-56. Jan 2005.
- [7] A. Begen, Y. Altunbasak, O. Ergun, M. Ammar, “Multipath selection for Multiple Description Encoded Video streaming”, EURASIP signal processing: Image Comm., vol. 20, pp. 39-60, Jan. 2005.
- [8] http://www.wichiptech.com/h.264_encoder.html . Accessed on 4 June 2010
- [9] A. Norkin, “Multiple Description Coding of Visual Information”, PhD thesis, Tampere University of Technology, Finland. 4th Dec. 2007.
- [10] L. Yen-Chi, Y. Altunbasak, R.M. Merereau, “Layered coded vs Multiple Description coded video over error ” Signal Processing: Image Comm. Journal vol 18, pp. 337-356, 2005.
- [11] ISO/IEC JTC 1, “Coding of Audio-Visual Objects – Part 2: Visual”, ISO/IEC 14492-2 (MPEG-4 Visual), Version 1: April 1999, Version 2: Feb. 2000, Version 3: May 2004.

3D Video Technology

- [12] J. R. Ohm, “Registered Responses to the Call for Proposals on Scalable Video Coding”, ISO/IEC JTC1/SC29/WG11, M10569 (2004).
- [13] H. Schwarz, D. Marpe, T. Wiegand, “Overview of the scalable video coding extension of the H.264/AVC standard”, IEEE Trans. of Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding, vol. 17, no. 9, pp. 1103-1120, Sept. 2007.
- [14] D. Taubman, “High Performance Scalable Image Compression with EBCOT”, IEEE Transactions on Image Processing , pp 1158-1170, Sept. 2000.
- [15] H. A. Karim, C. T. E. R. Hewage, A. C. Yu, S. Worrall, S. Dogan, A. M. Kondoz, “Scalable Multiple Description 3D Video Coding Based on Even and Odd Frame“, I-Lab Technical Report, CCSR, University of Surrey, Guildford, UK.
- [16] O. Schreer, P. Kauff, T. Sikora, “3D Video Communication: Algorithms, Concepts and Real-time systems in human centred communication”, John Wiley & Sons, 1st edition, 2005.
- [17] <http://www.3dstereomedia.com/content/anaglyph-colorcode-chromadepth>, accessed on 24th November 2008.
- [18] <http://www.inition.co.uk/inition/product.php?> ,accessed on 27th January, 2010.
- [19] http://www.crc.gc.ca/en/html/crc/home/mediazone/eye_on_tech/2007/issue7/3dtv, accessed on 20th December 2009.
- [20] <http://3dcinecast.blogspot.com/2009/05/digital-hollywood-3d-needs-to-make.html>, accessed on 18th November 2008.
- [21] A. Woods, T. Docherty, R. Koch, “Image distortions in stereoscopic video systems”,in proceedings of Stereoscopic Displays and Applications IV, 1993.
- [22] L. M. J. Meesters, W. A. IJsselsteijn, P. J. H. Seuntjens, “A survey of perceptual evaluations and requirements of three-dimensional TV”, IEEE Trans. Circuits & systems for video technology, vol. 14, pp. 381-391. 2004.
- [23] C. Fehn, “Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV”, In Proc. of SPIE vol. 5291, pp. 93-104. 2004.
- [24] L Zhang, “Stereoscopic image generation based on depth image for 3D TV”, IEEE Trans. Broadcasting, vol. 51, no.2, pp 191-199, 2005.

3D Video Technology

- [25] ISO/IEC JTC1/SC29/WG11, “Depth Estimation Reference Software (DERS) 4.0”, M16605 report, July 2009.
- [26] ISO/IEC JTC1/SC29/WG11, “Reference Software for Depth Estimation and View Synthesis”, M15377 report, April 2008.
- [27] M. Talebpourazad, “3d-tv content generation and multi-view video coding”, PhD Thesis, Department of Electrical & Computer Engineering, University of British Columbia, Vancouver, Canada, June 2010.
- [28] S. Pastoor, M. W. opking, “3-D displays: A review of current technologies”, Displays, vol. 17, pp. 100-110, 1997.
- [29] Y. Y. Yeh, L. D. Silverstein, “Limits of fusion and depth judgment in stereoscopic color displays”, journal on Human Factors, vol. 32, issue 1, pp. 45-60. 1990.
- [30] I. Sexton, P. Surman, “Stereoscopic and autostereoscopic display systems”, IEEE Journal on Signal Processing Magazine, vol. 16, issue 3, pp. 85-99, 1999.
- [31] A. Aggoun, “3D VIVANT: Live Immerse Video-Audio Interactive Multimedia”: http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=94148. Accessed on 2 Nov. 2011.

Chapter 4: 3D MDC with Side Information and Motion Interpolation

4.1 Introduction

Multiple Descriptions Coding (MDC) has emerged as a promising technique to enhance the error resilience of a video transmission system. The MDC is a coding approach normally used over error-prone and unreliable channels. MDC has been introduced as a generalization of source coding subject to a fidelity criterion for communication systems that use diversity to overcome channel impairments. The source is encoded into one or more correlated coded representations called descriptions, which are transmitted over separate channels. The challenge in the design of an MDC scheme is thus how to achieve the best average rate distortion (R-D) performance when all channels operate (at the central decoder), subject to constraints on the average distortion when only one channel or a subset of channels is operational. In the presence of losses, e.g., when one description is lost, an acceptable quality can be achieved without making use of retransmission mechanisms such as the Selective Repeat, Go-back-N and Stop and Wait. Delay experienced by retransmissions may not be acceptable to certain traffic classes. This allows exploiting path diversity on communication networks. Transmission of the descriptions over multiple diverse paths can compensate the dynamic and unpredictable nature of the communication medium, different paths might have different error characteristics, introduce different delays and so on. It has been shown that MDC combined with multipath routing performs significantly better over diverse networks than the single path and single description coding (SDC), especially at low rates and under delay constraints [1]. Multiple description video coding schemes which make use of temporal prediction suffer from a number of limitations. Firstly, the redundancy which is introduced to achieve robustness is beneficial to improve the R-D performance of the side decoding but is in general wasted when both descriptions are received. Secondly, Good side R-D performance can, therefore be obtained, however, at the cost of degraded central R-D performance. In addition, the use of predictive coding in multiple description video coding schemes often leads to mismatch. At the decoder, the prediction signal may differ from the one used by the encoder. This means that it may not always be possible to reconstruct the video properly due to this predictive mismatch. A solution to eliminate mismatch

3D MDC with Side Information and Motion Interpolation

is to encode the given frames separately with respect to each possible predictor and transmit multiple descriptions; however, since the number of possible predictors grow exponentially with time, this solution could result in excessive bit rate and encoding complexity.

A certain amount of controlled redundancy is added to the compressed descriptions to make them useful even when received alone. The decoder can reconstruct the source from one description with low, yet acceptable quality. The more descriptions received, the higher is the reconstruction quality. Usually, the descriptions are balanced; that is, the descriptions are of equal rate and importance. In that case, the reconstruction quality depends on the number of received descriptions only and not on which particular descriptions that are received.

The simplest MDC architecture is shown in figure 4-1. The encoder creates two descriptions which are sent separately across the wireless channel as stream 1 and 2. The bit rates used to sent each description, in bits per second are R_1 and R_2 , and the total rate is $R = R_1 + R_2$. Three scenarios are possible: both descriptions are received by the MDC decoder or either one of the two descriptions is missing. The decoder has three decoders, each corresponding to three scenarios. This force the encoder to consider explicitly that the decoder may be in one of three states, even though the encoder cannot know which of the three states the decoder is in. The central decoder receives both descriptions and produces a high quality reconstruction with central distortion (X_0), while the two side decoders each receive only one of the two descriptions and produce lower but acceptable quality reconstruction with side distortions (X_1) and X_2). In many design balanced rate is useful i.e. $R_1 = R_2$.

When transmitted over a wireless network, each description is placed in a separate stream where video frames. The video frames in the receiver is considered either correctly received or lost. Increasing the number of descriptions increases the probability that at least one video frame (i.e one description) reaches the decoder. However, it also increases the coding redundancy and the decoder complexity.

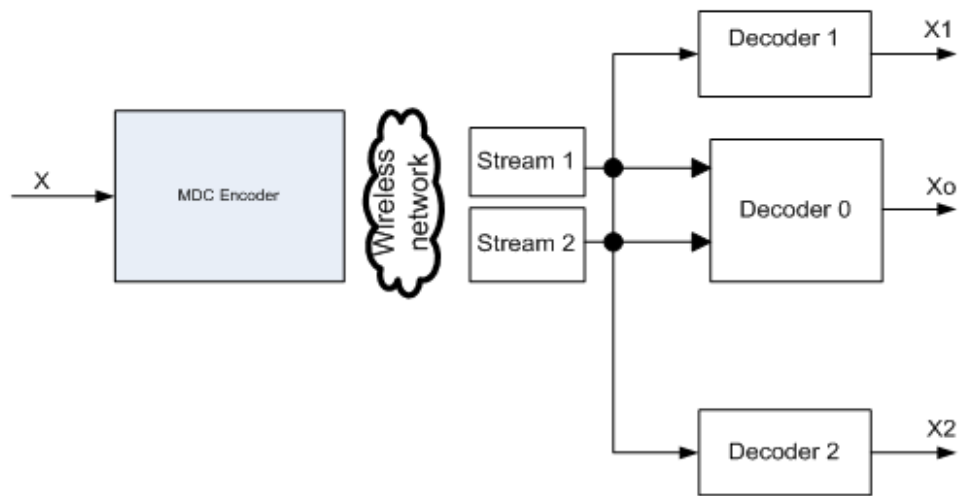


Figure 4-1: Basic MDC Codec with two paths, wireless channels & three decoders.

In summary, MDC is an attractive coding approach as it provides error resilience and scalability using only part of the data sent to the decoder and no need to employ no priority –enabled transmission mechanisms in the network. MDC is especially advantageous in short-delay media streaming scenarios such as video conferencing and when broadcast over error-prone channels where it provides acceptable reconstruction quality in case of packet loss.

4.2 Literature Review of MDC Algorithms

From the literature, one could get the impression that MDC arose as a curious analytical puzzle and then found application years later. More accurately, MDC started from explicit motivation to theoretical novelty and moved to engineering applications. Like many communication technologies, MDC was invented at Bell Laboratories in connection with voice communication over the telephone network [1]. However, some of this MDC work, though documented through citations in technical reports, was not archived. Little of it is known inside Bell Laboratories and almost none is known outside. The hallmark of the telephone system is reliability. But outages of transmission links are inevitable. They arise from device failures and also from routine maintenance and upgrades. Thus, achieving high reliability requires mechanisms to handle outages. In the 1970s, the primary mechanism for providing uninterrupted service despite link

3D MDC with Side Information and Motion Interpolation

outages was to divert calls to standby transmission links. The need for standby links increases cost and implies that not all the installed links are used in normal operation. To improve reliability without standby links, the information from a single call could be split and sent on two separate links or paths. In normal operation, both halves of the transmitted data would be combined for the usual voice quality; an outage of one link or the other would allow for communication at reduced quality. In [1], this idea of channel separation is attributed to W. S. Boyle. A citation of [2] and the later archived document [3] indicate that this idea may have been originated by Miller. Miller sketched a few simple methods for sending digital and analogue information over split, discrete-time analogue links. Miller's methods for digital information are all more or less equivalent to Gray coding [4]. Because of its decaying frequency spectrum, speech that is initially sampled at the Nyquist rate can be subsampled by two without too much aliasing. Thus, sending odd-numbered samples on one channel and even-numbered samples on the other channel.

Miller and Boyle worked primarily on the physical layer, designing optical equipments. In 1978 and 1979, the idea of channel splitting became popular with two groups in Bell Laboratories: Speech coders and information theorists. Gersho, spanning these two camps, was instrumental in this spread. He learned of the problem from Goodman, proposed an encoding technique and likely was the first to share the idea with Jayant, Ozarow, Witsenhausen, Wolf, and Wyner. Jayant, working independently with Miller, also proposed a separation of odd and even samples in a speech coding method for channel splitting. Jayant's simple, yet quite effective system was used for many years. A similar system had been motivated by random losses in packet-switched telephony in the early 1990s. The usual practice in telephony is for a speech signal to be band limited to 3.2 kHz and sampled at 8 kHz. In Jayant's system, the initial sampling is at 12 kHz so that sub-sampling by a factor of two results in only slight aliasing. The odd and even numbered samples are compressed separately with differential pulse code modulation (DPCM) and sent on two separate channels. Quantisation step sizes are adapted once for each block of several millisecond duration. Decoding from both channels requires a DPCM decoder for each channel and the interleaving of the samples, resulting in a signal with 12 kHz sampling and some amount of DPCM quantisation noise. To decode from a single channel, adaptive linear interpolation is used. It is almost the same thing to consider this to be speech sampled at 6 kHz with aliasing and quantisation noise.

3D MDC with Side Information and Motion Interpolation

Judging by both SNR and informal perceptual testing, this technique works very well for the range of 2 to 5 bits/sample (24 to 60 kbits/s): The quality of either half-rate reception in a system designed for total rate R is similar to that of the full-rate reception in a system optimized for total rate $R/2$. In particular, at 60 kbits/s even the half-rate receiver approaches “total quality”. The good performance is not surprising because, at the bit rates used for telephony, halving the sampling rate while keeping the quantisation unchanged is a reasonable way to halve the bit rate. Mathematical idealization of this system shows that the odd and even numbered samples have redundancy. If the redundancy was removed before odd/even separation for example with linear prediction, the performance would be unacceptable. Here, MDC is formed by splitting the quantized coefficients into two or more streams

Some MDC were later proposed, including [5]. In the second half of the 1990s, MDC became popular as an effective means to combat transmission errors in the best-effort Internet and wireless networks. Numerous MDC coders have since been proposed for coding multimedia (Speech, Image, and Video). For a comprehensive literature review of the development history of MDC, the rate-distortion (R-D) bound for MDC, and various MDC Algorithms developed primarily for images and later found its applications in videos. MDC is particularly promising for video because of the very stringent delay requirement in many video applications. As is well known, motion-compensated prediction can effectively exploit the temporal correlation between video frames.

MDC algorithms reported in the literature are categorized into a number of methods described in the following sub-sections.

4.2.1 MDC Quantisation

MDC quantisation splits the quantised coefficients into two or more streams. In the simplest form of implementation, the multiple descriptions are produced by using two quantisers whose decision regions are offset by one-half of a quantisation interval of each other.

In [6], it is shown that how quantisers can be used to produce two complimentary descriptions of the same scalar quantity. MDC scalar quantisation is flexible in that it allows a designer to choose the relative importance of the central distortion and each side distortion. The use of side distortion increases the redundancy of the codec. Reudink [6] invented several techniques with lower redundancy. Later, Vaishampayan [7] independently developed a theory for designing

3D MDC with Side Information and Motion Interpolation

MDC scalar quantisation. At high error rates, the central and side distortion can be traded off while keeping their product constant [7].

Extending the MDC scalar quantisation to vector quantisation is easy. However, the index assignment problem becomes more difficult because the code vectors/code books cannot be naturally ordered for videos; in addition, the encoding complexity increases with dimension especially when you are dealing with 3D videos. The technique that avoids these problems is the MDC lattice vector quantisation (MDCLVQ). The index assignment problem is simplified by lattice symmetries and the lattice structure also reduces encoding complexity [8]. The MDC quantisation algorithm can also be improved using embedded scalar quantiser which can achieve a fine grain scalable bit stream beside error resilience. The scalar quantiser will refine the source information successively using finer quantisers which can be created by further segmenting the steps of coarser quantizer.

Other MDC quantisation techniques are described in [9]. The proposed method in [9] is further enhanced in [10] by adding another quantiser in the central prediction loop.

4.2.2 MDC Transform Coding

In MDC transform coding, the multiple descriptions in the form of transform coefficients are produced from the output of a transform coded block. In [11], pair-wise correlating transform (PCT) is proposed to transform a set of coefficients into two sets of correlated coefficients with controlled redundancy and side information. Cascaded correlating transform as extension to the pair-wise correlating transform is proposed in [12].

In wavelet transform based MDC as in [13][14], the MDC streams are respectively produced from the arranged wavelet coefficients and the partitioned transform domain of the signals. Researchers in [15] investigated another wavelet based MDC, which produces multiple descriptions from the wavelet representation following a checker-board pattern. An MDC scheme and its application to multi-path transport have been investigated in [16]. Lapped orthogonal transform is used in the transform stage and the transformed coefficients are split into two descriptions using a checker-board pattern.

4.2.3 MDC Sub-sampling

In MDC sub-sampling, the original signal is decomposed into subsets, either in spatial, temporal or frequency domain [17] where each description corresponds to different subsets. This method

3D MDC with Side Information and Motion Interpolation

takes advantage of the correlation of the spatially or temporally adjacent video data samples. 3D Multiple description sub-sampling investigates the compression of depth information at reduced resolution for low bit rate applications such as the mobile networks. The frames are down-sampled prior to encoding and up-sampled after decoding. The application of reduced resolution for image compression has been investigated in [18] with the aim of improving subjective and objective perceived performance at low and very low bit rates. The application of reduced resolution for depth image sequences is also investigated in [19]. MDC sub-sampling for 2D video is also used for example in mobile applications with the sole aim of reducing bandwidth and providing portability. The down-sampling and up-sampling algorithms can be applied to the enhancement layer of SVC which is used to code the depth information.

Examples of algorithms include temporal frame interleaving and spatial pixel interleaving on image samples or motion vectors using quincunx sub-sampling [20]. In [20] for example, the MDC streams are generated by encoding the motion vector field into two descriptions using quincunx sub-sampling process.

4.2.4 MDC Even and Odd Frames

Many even and odd frames based MDC methods are investigated in the literature due to its simplicity in producing multiple streams. The even and odd MDC basically includes the even and odd video frames into description one and two respectively [21]. An odd frame is predicted from previously reconstructed odd frame only as shown in figure 4-2 . Prediction of the even frame is also similar to that of the odd frame. It is important to note that the two descriptions are independently coded so that when only a single stream is received at the decoder, it can be decoded with acceptable quality at lower temporal resolution. It also introduces no mismatch when only one of the descriptions is received because the decoder uses the same prediction signal as the encoder for each generated description. Compatibility with the existing video coding standard is another advantage for the even and odd frames MDC as the descriptions from this MDC can be decoded by the standard decoder, provided the descriptions are encoded using the standard encoder.

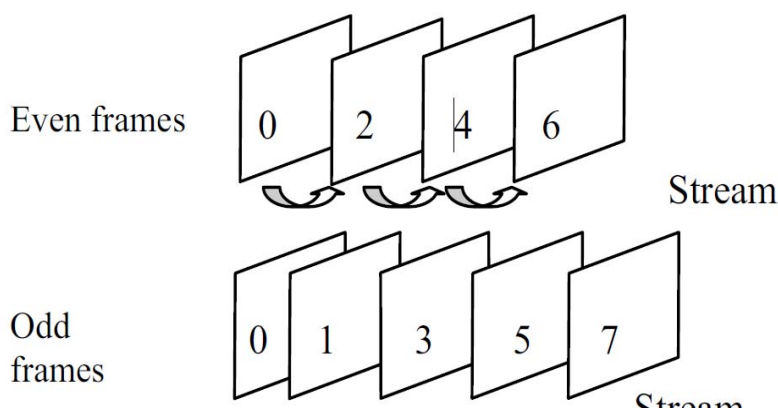


Figure 4-2: Contents of stream 1 and 2 at frame level.

Redundancy in even and odd MDC comes from the longer temporal prediction distance compared to standard video coder, which uses the nearest past frame for prediction. Hence, its coding efficiency is reduced. This method is similar to the video redundancy coding (VRC) proposed in [22].

For practical MDC scheme, it is necessary to control the redundancy [23] to match the network conditions. The MDC method in [23] is similar to the VRC method, but the mismatch between the predicted frames at the encoder and decoder is also coded and appended in both descriptions. The predictor and the mismatch signal quantizer can be varied to control the redundancies. In [24], it is stated that two streams of lower-resolution pictures are added to the multiple state video streams. In case one of the streams is lost, a spatial-temporal hybrid interpolation is used to recover the missing frames.

Performance of the multi-state video encoder proposed in [21] is improved by [25] using multi-hypothesis motion prediction at the encoder. Small additional block motion information is introduced, which helps fast error recovery at the decoder. Multi-state video encoder with side information is proposed in [26]. The side information, which is calculated offline at the encoder, will give good quality.

All the MDC methods discussed above can be applied to 2D video to provide error resilience and scalability. In case of 3D video, MDC schemes were proposed for 3D stereoscopic left and right views in [27]. Other MDC schemes which take advantage of encoding a source transmission over multiple channels have been used in [28]. This implores a novel MDC technique with side information for 3D video based on the even and odd video frames. The redundant side information consists of the difference between the interpolated frame and the locally reconstructed frame that can be quantized; hence the redundancies can be controlled by the

3D MDC with Side Information and Motion Interpolation

quantization parameter. This is then extended using Bi-directional frames (B-frame) coding technique to achieve reduced variable redundancies [29]. In [30] and [31], scalable MDC is proposed for the 2D plus depth format stereoscopic video. The method is then improved using motion interpolation and applied to 2D video in [32] and 3D video in [33].

In this thesis, the method in [32][33] is extended to improve the performance of stereoscopic video transmission in error free and error prone conditions using side information and motion interpolation.

4.3 Motion Interpolation

Previous work reported in [33] used a simple pixel averaging technique to interpolate the missing frames. Such a method gives an erroneous image with high errors/ghosting/blurring effect for 3D video sequences with higher motion and large computational complexity; it also produce large PSNR variation between frames when errors occur. Using Motion Interpolation (MI) approach reduces the complexity, buffering delay and the processing time in both the encoding and decoding of the redundant video data. This can be countered by interpolating at the encoder and sending the residual as described in [32]. However, such a method will cause high redundancy as the residual energy needed to represent a given frame will be high due to the poor prediction making such a method not viable for low bit rate applications over mobile channels. These observations motivate the need to use more powerful motion estimation techniques since the accuracy of the decoder frame interpolation module is a key factor to final compression performance. However, the traditional motion estimation and compensation techniques used at the encoder for video coding are not adequate to perform frame interpolation since they attempt to choose the best prediction for the current frame in the rate-distortion sense. For frame interpolation, we need to find an estimate of the current frame, and therefore a good criterion is to estimate the true motion, and based on that to perform motion compensation between temporally adjacent frames. To provide a solution, we propose to add a MI algorithm aside from the Side Information (SI) in the MDC-Even Odd and Side information [33] and call this new scheme Scalable 3D MDC-SIMI. By adding the new algorithm, the frame quality for an interpolated frame can significantly be improved and therefore it reduces the residual energy needed to represent a given frame when encoding, thus further reducing the bit rate. Figure 4-3 shows the proposed MI Algorithm to be added to the encoder.

3D MDC with Side Information and Motion Interpolation

This new approach employs a second-order predictor for motion compensation, which predicts a current frame from two previously coded frames. The mismatch between the predicted frames at the encoder and decoder is explicitly coded to avoid error propagation in the MDC channels. By using the second-order predictor and coding the mismatch signal, one can also suppress error propagation in packet lossy networks where packets in either description can be lost. The predictor and the mismatch signal quantizer can be varied to achieve a wide range of tradeoffs between coding efficiency and error resilience.

In the new algorithm, 3D MDC-SIMI, temporal correlation of video sequences is exploited through motion estimation at the encoder, so the estimated frames are known even before decoding. One of the most popular motion estimation algorithms is block-matching algorithm. In order to obtain high quality side information, series of measures is taken as shown in figure 4-3. These are low pass filter (LPF), forward motion estimation (FME), bidirectional motion estimation (BME), Spatial Smoothing and bidirectional motion compensation (BMC) [34][35]. The ME is used to simplify the motion search. The proposed scheme adopts two motion estimation algorithms: Forward motion estimation (FME) and bidirectional motion estimation (BME). In the motion compensation stage, the decision module is implemented to take full advantage of two motion estimations to improve side information. The proposed frame interpolation consists of several parts, i.e forward motion estimation (FME), forward motion compensation, bidirectional motion estimation (BME, filters and decision module.

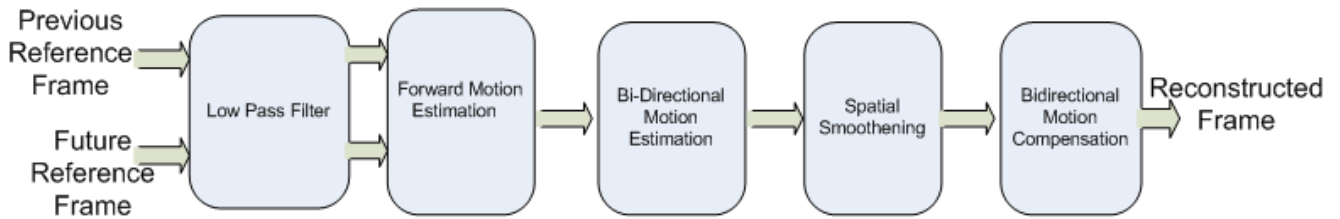


Figure 4-3: Proposed MI in MDC-SIMI Architecture.

One assumption is made for accurate motion estimation in block-matching methods. That is, all pixels in the entire block have the same motion vector (MVs). The smaller the block size the more possibly the assumption holds, but the stronger the local similarity, the less possibly the assumptions holds. In normal coding, inter-frame motion estimation is completed at the encoder, the estimated frame is known before encoding. The assumption has a significant impact on the interpolated results, however in our method, motion estimation is performed at the decoder and the estimated frame is unknown before encoding. Therefore, block size is of great importance in this new method.

4.3.1 FME and BME

In previous 3D coding techniques, key frames are reconstructed by the conventional intra frame decoder at the decoder side, while in our new method; key frames are estimated based on a new frame interpolation scheme. For a given block in the current frame, motion estimation is meant to search for the best matching block in the reference frame.

The new scheme includes two motion estimation algorithms: FME and BME, both estimation approaches is performed, by offsetting each other's weakness with each other's merit, greatly improving the quality of the side information and hence the overall quality of the decoded frame. In FME algorithm, the motion vectors are computed between the previous frame X_{n-1} and the following frame X_{n+1} . For a given current block, a reference block is obtained by full search method. It is found that FME cannot guarantee each pixel in the interpolated frame to be compensated. Motion Vector selection is shown in figure 4-4 and the process is explained in subsection of 4.3.3. Therefore, BME can be used to remove the problem associated with FME.

In order to cover all pixels in the interpolated frame, the BME algorithm is introduced. BME is composed of forward and backward estimation. However, the interpolated frame is not known to

3D MDC with Side Information and Motion Interpolation

overcome the shortcomings of the FME, the BME first needs to determine the search block location in the interpolated frame which will cover all blocks, one by one.

4.3.2 Decision Mechanism

Frame interpolation adopts two motion estimation algorithms: FME and BME. This results in two interpolated frames at the decoder. Uncovered areas in the interpolated frame from the FME will be filled up by BME. In order to make full use of both motion estimation algorithms, a decision mechanism is adopted to process the over lapped areas between two interpolated frames. In the overlapped areas, the bidirectional motion compensated value for current pixel is P_b with the corresponding sum of absolute difference(SAD) being SAD_b while the forward motion compensated value P_f with the corresponding SAD being SAD_f . The ultimate compensated value P is determined by

$$P = \frac{SAD_2}{SAD_1+SAD_2} P_b + \frac{SAD_b}{SAD_b+SAD_f} P_f \quad (4-1)$$

Where: SD_1 and SD_2 are sum of absolute differences and P_b the corresponding bidirectional motion compensated value while P_f is the motion compensated value respectively.

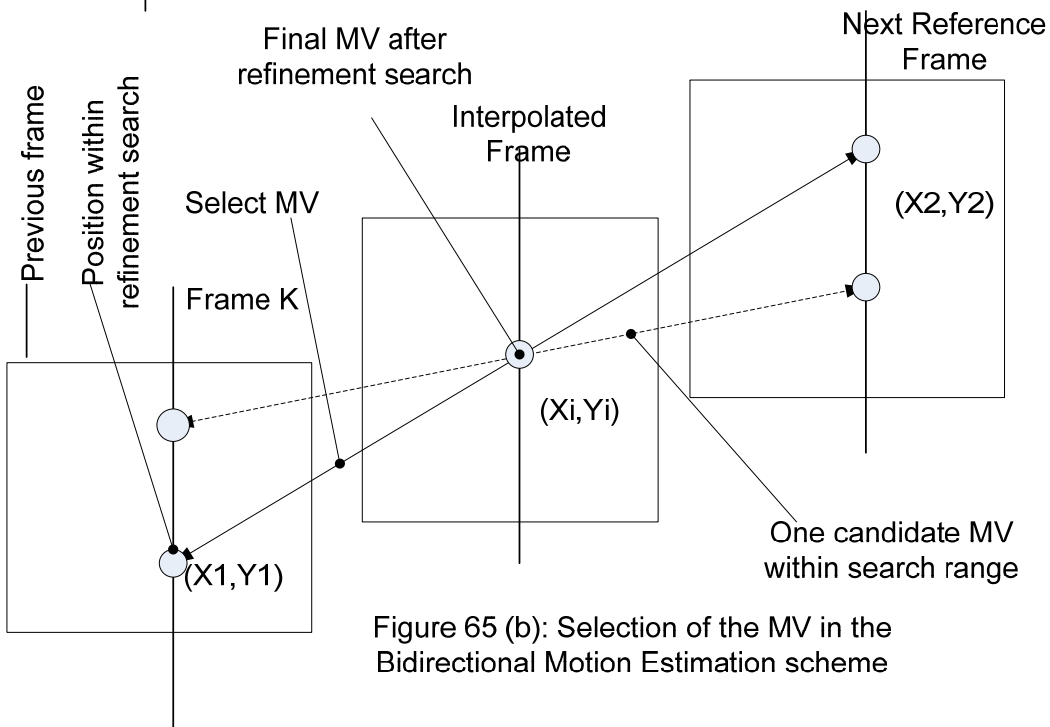
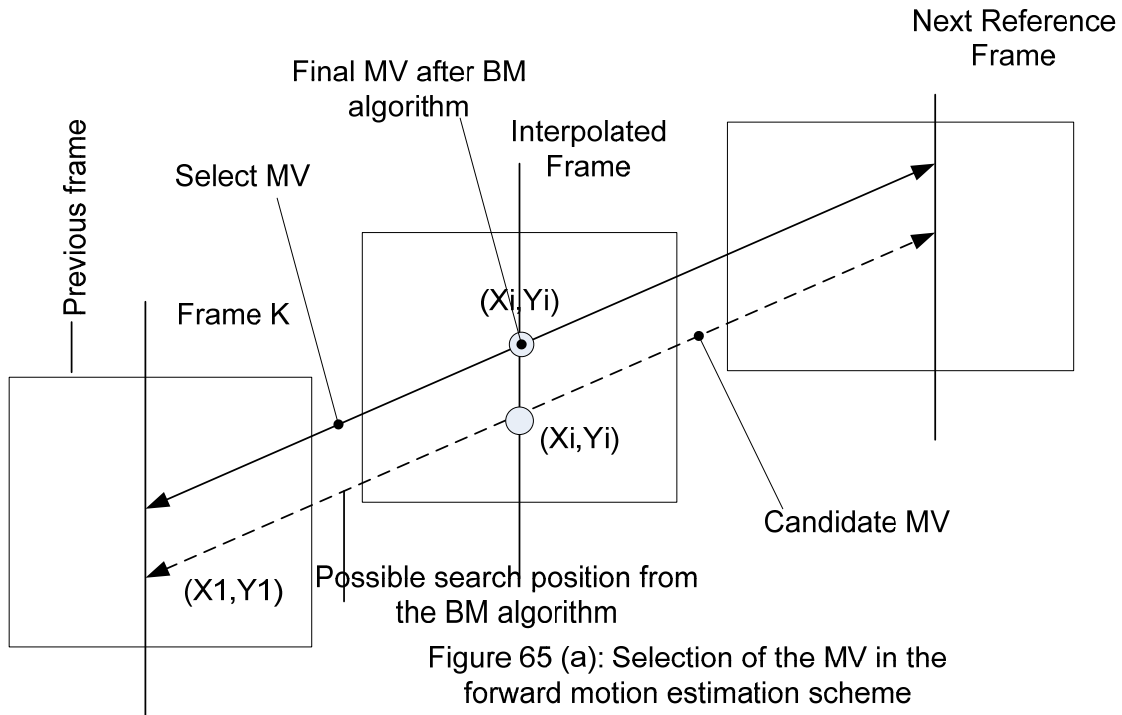


Figure 4-4: Motion Vector Selection.

4.3.3 Motion Vector Selection

Figure 4-4(a) shows the process in selecting a motion vector in the forward motion estimation scheme. A block matching algorithm is used to estimate the motion between the next and previous reference frames. Next motion vectors that intercept the interpolated frame closer to the centre are selected from motion vectors obtained from the previous block matching algorithm.

Figure 4-4 (b) shows the motion selection vector in the bidirectional motion estimation architecture. It selects a linear trajectory between next and previous reference frames that passes through the centre of the block located in the interpolated frame. It then performs a small refinement search around the initial block position and the motion vectors between three frames. In choosing new candidates, motion vectors must remain on the same linear trajectory thus motion vectors are symmetric between adjacent key frames generating:

$$(x_1, y_1) = (x_i, y_i) + MV(B_i) \quad (4-2)$$

$$(x_2, y_2) = (x_i, y_i) - MV(B_i) \quad (4-3)$$

Where (x_1, y_1) are the coordinates of the block in the previous reference frame, (x_2, y_2) are the coordinates of the block in the next frame and $MV(B_i)$ represents the motion vector obtained using the forward search which is determined by the distance between the forward and previous reference frames. In figure 4-4(b), the motion vector is divided by two, since the interpolated frame is the same distance between both reference frames.

After that, spatial smoothing is used to reduce the number of false motion vectors. It is based on weighted vector median filters as defined in [33]. For bidirectional motion compensation, the standard bidirectional motion compensation found in conventional video coding is used.

4.4 Proposed Scalable 3D MDC with Side Information & Motion Interpolation

Even and odd frames based MDC is popular due to its simplicity in producing multiple streams. It also introduces no mismatch when only one of the descriptions is received because the decoder uses the same prediction signal as the encoder for each generated description. Compatibility with the existing video coding standard is another advantage for the even and odd frames MDC as the descriptions from this MDC can be decoded by the standard decoder, provided the descriptions are encoded using the standard encoder.

3D MDC with Side Information and Motion Interpolation

Therefore, the proposed scalable MDC in this thesis is based on the Scalable even and odd frames based MDC for stereoscopic 3D that uses the SVC coder [31] and multiple description video coding using motion compensated frame interpolation (MCFI) [23] which enhances the base layer of SVC using temporal MDC. The new architecture is similar to the work in [33] but differs in the redundancy handling. This layer is produced by separating a sequence into even and odd thus creating two MDC streams.

In the previous work [23][31][33], if a stream became corrupted, it is replaced by interpolated frames from the other stream using a simple pixel averaging technique. However, this creates an erroneous image with severe ghosting especially with frames with high motion or severe differences between neighbouring frames. To solve this problem, it is possible to send extra information about missing frames located in the other stream. Doing so will improve the frame reconstruction at the decoder. However, such approach will add redundancy to the encoded bit stream. If this is not addressed adequately, the scalable MDC video delivery system will not be viable for low bit rate applications such as mobile video. The frame reconstruction can be improved by interpolating the adjacent frames in one stream and encoding the difference between the interpolated frame and corresponding frame located in the other stream. This information can be used to improve the quality of lost frames. This is known as 3D MDC Even and Odd frame with Side Information and Motion Interpolation.

Each description is separately transmitted and independently coded, so that loss of some of the description will not jeopardise the decoding of correctly received descriptions.

4.4.1 Encoder Architecture

The architecture for the proposed MDC SIMI for 3D video is shown in figure 4-5 for the encoder. The even and odd frames are encoded into stream 1 and 2 respectively. Each frame contains texture, motion and depth data. Side information for even and odd stream frames is also appended to their corresponding streams. At the encoder, the central encoder is used to produce even or odd frames. The frame buffer is used to store the reconstructed frames, $F'(n-1)$ and $F'(n-2)$. Even frames are predicted from previous reconstructed even frames and vice versa for odd frames.

Side encoder 1 and 2 are used to produce the side information for even and odd stream. Inside encoder 1, motion interpolation is performed between the current reconstructed even frame

3D MDC with Side Information and Motion Interpolation

$F_e'(n)$ and the previous reconstructed even frame, $F_e'(n-2)$. Only side encoder 1 is shown in figure 4-5, but side encoder 2 has similar structure.

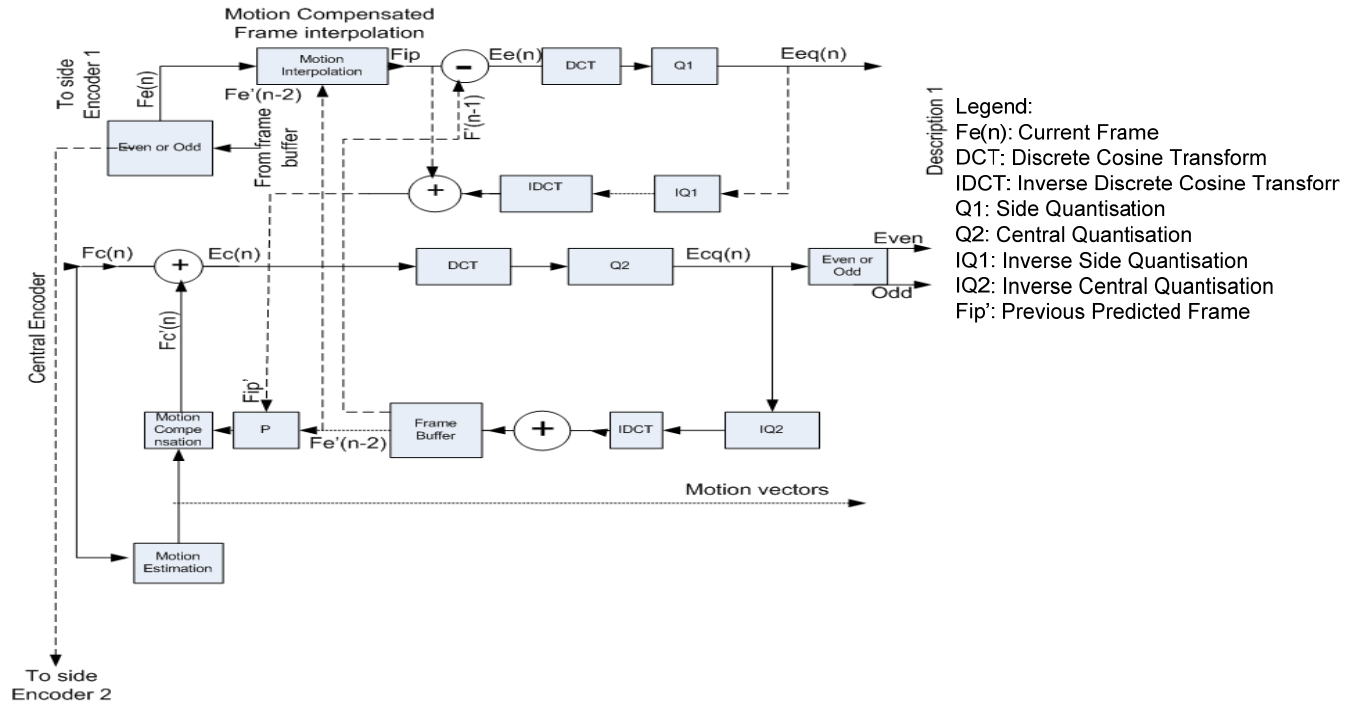


Figure 4-5: Proposed 3D MDC Encoder

For the central decoder,

- $F_c(n)$: Current frame
- $F_c'(n)$: Predicted current frame
- $E_c(n)$: Error for current frame
- $E_{cq}(n)$: Coded error

For Side Encoder 1,

- $F_e'(n)$: Current reconstructed even frame
- $F_e'(n-2)$: Previous reconstructed even frame
- $F_o'(n-1)$: Previous reconstructed odd frame
- $E_e(n)$: Side information error
- $E_{eq}(n)$: Coded side information error
- F_{ip} : Interpolated Frame

Side Encoder 2

- $F_o'(n)$: Current reconstructed odd frame

The interpolated frame is subtracted from the previous reconstructed frame, $F'(n-1)$ and the difference, $E_e(n)$, which is the side information, is coded using DCT and quantisation. There by controlling the redundancy by varying the quantization parameter (Q_1) of the side information.

4.4.2 Decoder Architecture

The architecture of the proposed MDC-SIMI decoder for 3D video is shown in figure 4-6. At the decoder, the central decoder is used to decode the central information. If only an even stream is received, side decoder 1 is used to recover the odd frame, $F_o'(n)$. The results of motion interpolation of the previous reconstructed even frame $F_c'(n-2)$, and previous reconstructed frame $F_c'(n)$, is added to the decoded side information, $E_e'(n)$, to get $F_o'(n)$ which is the recovered odd video frame.

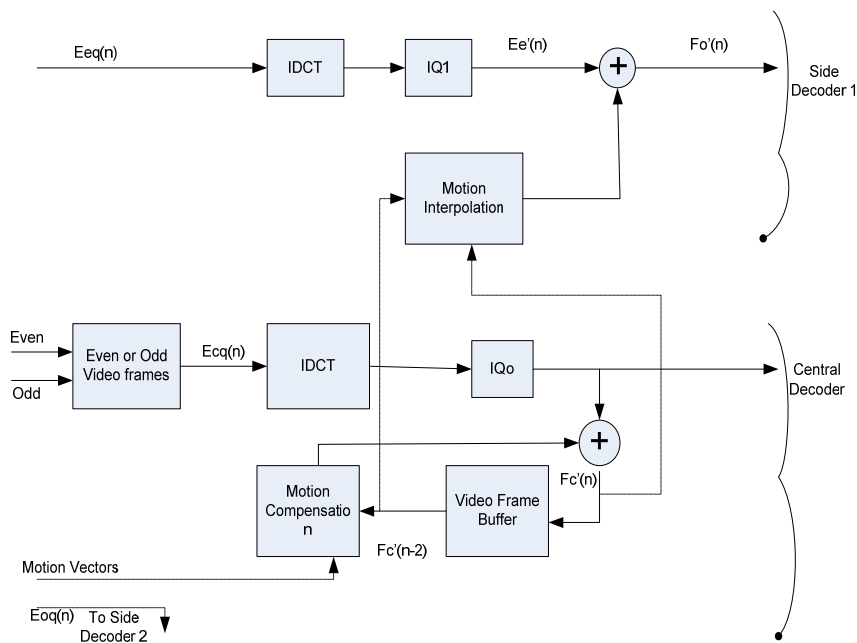


Figure 4-6: Proposed 3D MDC Decoder.

Central decoder:

- $E_{cq}(n)$: Received coded error
- $F_c'(n-2)$: Previous reconstructed even frame
- $F_c'(n)$: Decoded current frame

Side Decoder 1:

- $E_{eq}(n)$: Coded side information error
- $E_e'(n)$: Decoded Side Information
- $F_o'(n)$: Recovered odd frame

In the decoder, if both odd and even descriptions are received, the central decoder is used to reconstruct both even and odd 3D video frames. Each description is separately transmitted and

independently coded, so that loss of some of the description will not jeopardise the decoding of correctly received description

After decoding the even and odd frames with error signal, the frames are then merged with the decoded side information to produce a full resolution decoded sequence. Lost frames depending on the channel errors are interpolated at the decoder using MI. It is important to note that the two descriptions are independently coded so that when only a single stream is received at the decoder, it can be decoded with acceptable quality at lower temporal resolution. This has the advantage of not having to code the mismatch information.

4.5 Video Quality Performance Metrics/Parameters

The video quality metric proposed in recent years include PSNR, Video Quality Model (VQM), Moving Picture Quality Metric (MPQM), Noise Quality Measure (NQM), Structural Similarity Index Measure (SSIM), Subjective Quality of Experience (QoE) e.t.c. Peak-Signal-to-Noise-Ratio (PSNR) as video quality metric was used in this chapter due to its simplicity. The rate distortion (R-D) is measured using PSNR of the left and right views. This PSNR is computed by comparing the original left and right view data from the encoder with the left and right view data obtained from the decoder. A higher PSNR value provides a higher image and video quality. A lower PSNR value implies high numerical differences between images/videos. PSNR has been found to correlate well with human perception [34].

4.5.1 Peak Signal to Noise Ratio

The rate distortion is measured using PSNR of the left-and-right views. This PSNR is computed by comparing the original left-and-right view data with the left-and-right view data obtained from decoder. A higher PSNR value provides a higher image/video quality. A lower PSNR value implies high numerical differences between images. PSNR has been found not to correlate well with human perception. PSNR is derived by setting the mean squared error (MSE) in relation to the maximum possible value of the luminance for a typical 8-bit value that is $2^8-1 = 255$ as follows:

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [f(i,j) - F(i,j)]^2}{MXN} \quad (4-4)$$

$$PSNR = 20 \cdot \log_{10} \{ 255 / \sqrt{MSE} \} \quad (4-5)$$

Where,

3D MDC with Side Information and Motion Interpolation

$F(i, j)$ is the original signal at pixel (i, j) , $\hat{F}(i, j)$ is the reconstructed signal, and $M \times N$ is the picture size. The result is a single number in decibels, ranging from 30 to 50 for medium to high quality video. Despite several objective video quality models developed in the past two decades, PSNR continues to be the most popular evaluation of the quality difference among pictures, images and videos. (See appendix H for more results using MSU Video Quality Measurement tool.

4.5.2 Structural Similarity Index Measure (SSIM)

SSIM is used to measure the similarity between two images and considered to have high correlation with video quality [35] and with the quality perception of human visual system [36]. This approach is provided by [37]. This method differs with PSNR which is error based, by using the structural distortion measurement instead of the error. The idea behind this is that, the human visual system is highly specialized in extracting structural information from the viewing field and it is not specialized in extracting the errors. Thus, a measurement on structural distortion gives a good correlation to the subjective impression.

Many different quality assessment methods can be developed from this assumption but [4] proposes a simple but effective index algorithm. If you let $x = \{x_i \mid i = 1, 2, \dots, N\}$ be the original signal and $y = \{y_i \mid i = 1, 2, \dots, N\}$ be the distorted signal, the SSIM can be calculated as

$$SSIM = (2 \bar{x}\bar{y} + C_1)(2\sigma_{xy} + C_2) / [(\bar{x})^2 + (\bar{y})^2 + C_1](\sigma_x^2 + \sigma_y^2 + C_2) \quad (4-6)$$

In this equation, \bar{x} , \bar{y} , σ_x , σ_y , σ_{xy} are the estimates of the mean of x , mean of y , the variance of x , the variance of y and the covariance of x and y . C_1 and C_2 are constants. The value of SSIM is between -1 and 1 and gets the best value of 1 if $x_i = y_i$ for all values of i . The quality index is applied to every image/video using a sliding window with 11x11 circular symmetric Gaussian weighting function for which the quality index is calculated and the total index of the image/video is the average of all the quality indexes of the image/video.

4.5.3 Video Quality Metric (VQM)

VQM is developed to provide an objective measurement for perceived video quality. It measures the perceptual effects of video impairments including blurring, jerky/unnatural motion, global noise, block distortion and colour distortion, and combine them into a single metric. VQM has a

high correlation with subjective video quality assessment and has been adopted by American National Standard Institute (ANSI) as an objective video quality standard.

4.5.4 Moving Picture Quality Metric (MPQM)

MPQM is an objective quality metric for moving picture which incorporates two human vision characteristics. Two key human perception that have been available are contrast sensitivity and masking. The first accounts for the fact that a signal is detected by the eye only if its contrast is greater than some threshold. The eye sensitivity varies as a function of spatial frequency, orientation and temporal frequency. The second is related to the human vision response to the combination of several signals. A stimulus consists of two types of signals (foreground and background). The detection threshold of the foreground will be modified as a function of the contrast of the background.

It first decomposes an original video sequence and a distorted version of it into perceptual channels. A channel-based distortion measure is then computed, accounting for contrast sensitivity and masking. Finally, the data is pooled over all the channels to compute the quality rating which is then scaled from 1 to 5 (from bad to excellent).

4.5.5 Packet Loss

Packet Loss measures the reliability of a connection. In addition, packet loss is one of the main parameters affecting the quality of data services as the performance of a higher layer application is affected by packet loss. Different applications will vary in the extent to which packet loss affects their usability. Packet loss can occur for a number of reasons:

- Congestion of Network Elements (NEs) due to heavy loading, which lead to packet being discarded.
- Congestion preventive actions taken by NEs, discarding packets even if the buffer is not full.
- Network congestion due to bandwidth limitation of links or traffic overload.
- Delay in video packet transmission, with packet arriving too late at the receiver to be played back in real-time services.
- Transmission errors especially in wireless links caused by link obstructions, weather conditions, etc.

4.5.6 Objective Performance Comparison

Table 4-1: Comparison of various Objective Performances Metrics.

Quality Metric	Mathematical Complexity	Correlation with Subjective Methods	Accessibility
PSNR	Simple	Good	Easy
SSIM	Complex	Fairly Good	Available (MATLAB)
MPQM	Complex	Varying	Not Available
VQM	Very Complex	Good	Not Available
NQM	Complex	Unknown	Not Available
PLR	Simple	Fairly	Difficult (PLR Simulator)

Quite a few alternative performance models have been proposed but none of them have been commonly accepted. PSNR is still being widely used in literature. A comparison of the objective metrics is presented in Table 10?. From this table it can be seen that if only one evaluation method is used the VQM or the SSIM give the most reliable results [38][46]. However, the computational complexities of these two methods make them difficult to apply to real-time applications such as video conferencing. Moreover, the test results for SSIM are based on still images yet the performance of this metric on the video sequence remains unknown. This varying result on MPQM also makes it difficult to use in this thesis. Furthermore, previous published test results [39][40] showed that the performance of most objective video quality models are statistically equivalent to root mean squared error (RMSE) and PSNR. This is why PSNR is still the most commonly used performance metric in the literature. For this, the author use PSNR for objective quality measurements in this thesis and particularly in chapter 4.

4.5.7 Packet Delay Variation or Jitter

Jitter is defined as the difference between the longest and the shortest packet delay in some period of time or maximum delay difference between two consecutive packets in some period of time. The jitter, specified in the RFC 1889 (RTP), is defined to be the smoothed absolute value of the difference in relative transit times between two consecutive packets, the relative transit time is the difference between a packet's RTP timestamp and the receiver's clock at the time of

arrival, measured in the same units. The variation between the individual RTD in each sample may be called jitter. Instantaneous Packet Delay Variation (IPDV jitter) defined by IPPM is computed as the absolute value of the difference between the inter-leaving time of two subsequent packets at the sender and the inter-arrival time of those two packets at the receiver. The lower the packet waiting time in a queue, the less jitter will be. Jitter values will be longer if packets belonging to different sources are highly interleaved

4.5.8 Rate Distortion (R-D) Performance

Rate distortion is a method of improving video quality in video compression. The name refers to the optimization of the amount of distortion (loss of video quality) against the amount of data required to encode the video, the rate. While it is primarily used by video encoders, rate-distortion can be used to improve quality in any encoding situation such as image, video, audio or otherwise where decisions have to be made that affect both video size and quality simultaneously. The method of making encoding decisions is for the video encoder to choose the result which yields the highest quality output video or image. However, this has the disadvantage that the choice it makes might require more bits while giving little quality increase. Common example of this problem is in motion estimation, and in particular regarding the of quarter pixel precision motion estimation. Adding the extra precision to the motion of a block during motion estimation might increase quality, but in some case that extra quality isn't worth the extra bits necessary to encode the motion vector to a higher precision because of the large redundancy. Rate distortion measures both the deviation from the source and the bit cost. The deviation from the source is usually measured as the MSE, in order to maximize the PSNR video quality metric [45]. We used it in this chapter hence the need for the brief definition.

4.6 Description of the experimental environment

The UMTS simulation environment reported in chapter 2 is used for 3D MDC-SIMI performance measurements. The compressed stereoscopic 3D video is transmitted over a UMTS wireless channel. The effect of UMTS physical link layer on the encoder downlink was simulated by corrupting the video streams with appropriate error patterns. The error patterns are used for emulating the downlink channel conditions for a specific E_b/N_o value, channel coding scheme, spreading factor (SF), propagation environments, mobile speed, and power control

3D MDC with Side Information and Motion Interpolation

availability. The UMTS model parameters used are Convolution Codes with rate 1/3, Vehicular –A propagation environment at velocity of 50km/h with a spreading factors of 8 and power control availability. The Energy per spectral noise density described in chapter 2 section 2.10.7 which is E_b/N_o is used to quantify the errors. Mean PSNR, which is the average PSNR at each E_b/N_o over 20 simulations is used to measure the quality of the decoded sequence. The 3D video performance in this environment is similar to that in Pedestrian-B environment with power control.

In this thesis, it was observed that 20 simulation runs for each test are enough to represent the overall UMTS channel effects on the video quality. In other words, each simulation value is generated by calculating the average of 20 experimental values obtained by using different seeds. The obtained results are plotted below.

4.6.1 Experiment

Two descriptions are enough to correct the negative impact of packet loss since measurements on packets loss in the wireless networks show that most loss periods are short i.e., the average number of consecutively lost packets is small, and packet loss rates (PLR) are always lower than 50% at most [41]. Two video sequences, Interview and Orbi (QCIF) are used to test the performance of 3D MDC SIMI and 3D MDC SIPA [33]. Twenty experiments were carried out under different conditions i.e error free and under channels errors. Both or only one description from the two video sequences was received at the end of decoder. The subjective performance of the reconstructed video using the received one or two descriptions has been plotted.

If both even and odd streams are received, the decoder can reconstruct the coded sequence at full temporal resolution. If only one stream is received, the decoder can decode at full resolution by interpolating between the received frames using the following equations.

$$I_{ip}(i, j) = (1 - \alpha)I_{prev}(i, j) + \alpha I_{fut}(i, j) \quad (4-7)$$

$$\text{Where, } \alpha = I_{ip} - I_{prev} / (I_{ip} - I_{prev} + I_{fut} - I_{ip}) \quad (4-8)$$

Where:

$I_{ip}(i, j)$ is the frame to be interpolated at location (i, j) ,

$I_{prev}(i, j)$ is the previous frame located at (i, j) ,

$I_{fut}(i, j)$ is the future frame located at (i, j) .

3D MDC with Side Information and Motion Interpolation

To generate the interpolated frame for frame 1, frames 0 and 2 must first be reconstructed (See Figure 4-7). We used low pass filters (LPF) at the beginning for both reference frames to help estimate motion vectors that are spatially correlated.

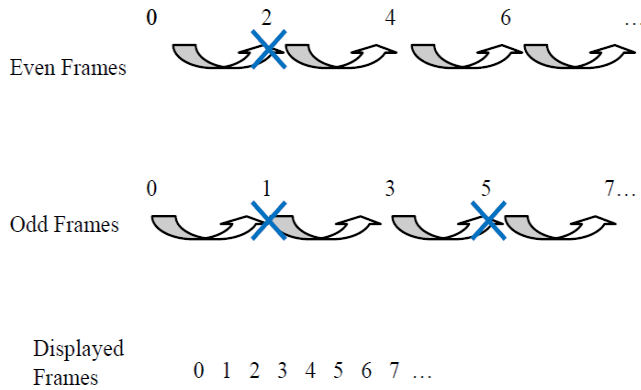


Figure 4-7: Errors on both streams at the decoder

However, in some situations the adjacent frame needed to construct a lost frame may also be lost. To solve this, the decoder waits until it receives the next erroneous frame and then constructs both lost frames using the equation 4-7. For example in figure 4-7, frame 1 is lost in the odd stream and frame 2 is lost in the even stream. The next available error free frame is frame 3. The decoder will generate 2 interpolated frames both for frame 1 and frame 2 by performing the MI and bidirectional motion compensation steps as described previously using different α values (this can be found according to the temporal distance between frames). In this problem, for frame 1, α will be 0.33 causing a weighting to I_{prev} and 0.67 for frame 2 causing a higher weighting to I_{fut} . To construct frame 5 in the odd stream α will be 0.5 as the distance between both I_{prev} , I_{fut} and the interpolated is equal.

The proposed MI has been compared with the pixel interpolation (PI) used in [33] for 3D and 2D video sequences. These sequences range from low to high motion. The even and odd frames with residual information has been separated into two independent bit streams with the SVC bit stream extractor used with two independent channels. Channel one transmits even frames with odd residuals while channel two transmits odd frames with even residuals. To simulate packet losses, VE NETWORK Packet Loss Simulator [42] was used. The packet losses rate ranges from [0 -20%] for each independent channel.

4.6.2 Simulation and discussions under error prone conditions

The error prone performance assessment was carried out on sequence “ Interview” with a QCIF resolution (174x144) for frames 13-50 with frame rate of 25 fps, Quantisation parameter (Q_p)= 20, fixed transmission rate of 300 Kbps. The compressed sequence was transmitted over a erroneous UMTS channel. The parameters for the channel are: OFDM modulation, FFT size: 64, carrier frequency of 2.4 GHz, sampling rate 20MHz and convolutional coding. The experiment was conducted 20 times with different BERs values and seed values and by averaging the result, we obtained the following plot. Figure 4-8 show the R-D performance of the two algorithms under error conditions. 3D MDC-SIMI outperforms the 3D MDC-SIPA under high error rates but under very low errors the 3D MDC-SIPA performs slightly better than our new algorithm. This is important to see how error propagation affects the quality performances of the two codec in a wireless network. Figure 4-8 plots the average PSNR as a function of channel bit error rates. Clearly, we can see that at low BER the performance of 3D MDC-SIPA is better than our new algorithm because it is approaching to SDC but under high error rates, 3D MDC-SIMI performs better than 3D MDC-SIPA. Also, the design of 3D MDC-SIMI requires approximately less computation times than 3D MDC-SIPA for error prone conditions.

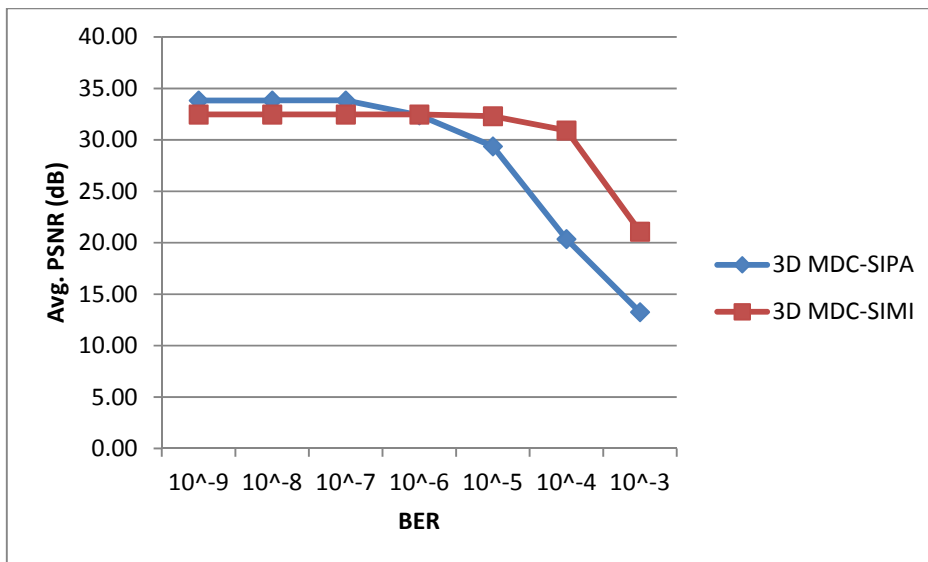


Figure 4-8: Comparison of 3D MDC-SIPA and 3D MDC-SIMI algorithms under channel errors.

3D MDC with Side Information and Motion Interpolation

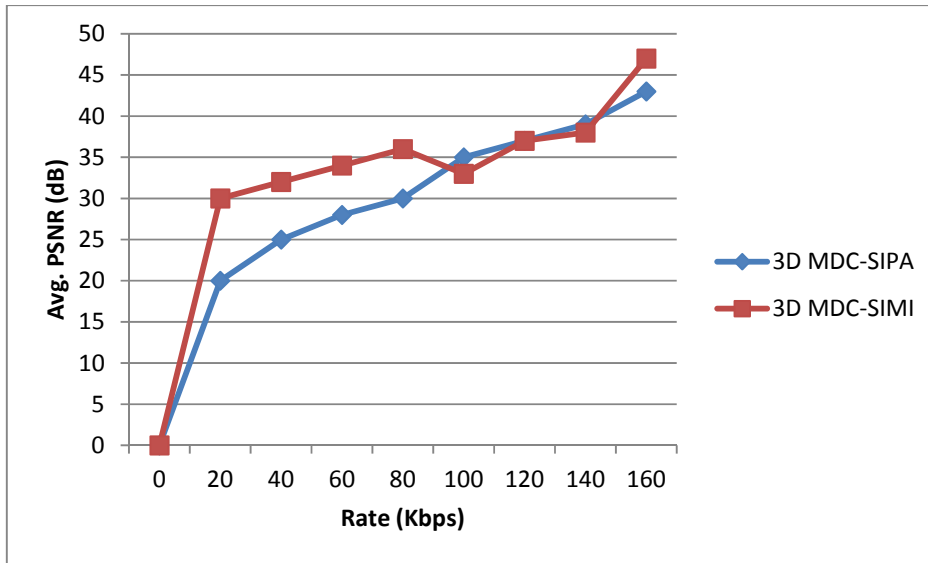


Figure 4-9: Comparison of 3D MDC-SIPA and 3D MDC-SIMI algorithms in terms of transmission rates

The above experiment in Figure 4-9 was also carried out 20 times with different seed values by varying the channel transmission rates (variable transmission in Kbps). Figure 4-9 compared only the MDC-SIPA and MDC-SIMI comparison under error free condition. It can be seen that most of the time, MDC SIMI outperforms MDC-SIPA, particularly in the low bit rates ranges upto 100Kbps. The PSNR vs frame rate for a number of frames in the interview sequence. It can be deduced that that the MI can give up to 2dB gain or more over PI. MDC-SIMI also has a better coding efficiency compared to MDC-SIPA in the low bit rate range.

4.6.3 Objective Performance under Error Free Conditions for all the MDC algorithms

The rate distortion performance of SDC, MDC-EO, MDC EO SIPA and MDC EO SIMI are compared in an error free environment. The distortion is measured using the Average PSNR of the luminance and depth. In some cases, the distortion is measured using the Average PSNR of the left and right views. The 3D sequence is produced from the original 2D texture and its associated original depth. Also a compressed left and right view sequence is obtained from the 2D and depth data reconstructed at the decoder. The luminance and depth PSNR is obtained by comparing the original luminance view data with the 2D plus depth data output from the decoder.

3D MDC with Side Information and Motion Interpolation

The Interview sequence with resolution 176 x 144, 353 x 288 and 704 x 576 are used in the error free environment. Fixed quantization is used in order to obtain the variable bit rates. The rate distortion for the depth information of the video sequence simulation is shown in the figures below. The bit rate is varied using the quantisation parameter. Some quantisation parameter is used for colour and depth information. In all the objective performance figures, the horizontal axis shows the bit rate in Kbps and the vertical axis shows the average PSNR corresponding to the bit rate.

The averaged PSNR of the luminance and depth for the SDC, MDC –EO, MDC-SIPA and MDC SIMI are plotted over a range of bit rates as shown in figures 4-10, 4-11 and 4-12. It can easily be shown that MDC SIPA and MDC SIMI is less efficient than SDC under error free conditions. Both are about 1-2 dB lower than SDC for the same bit rate. This is due to large redundancy/residue incurs by MDC SIPA and MDC-SIMI.

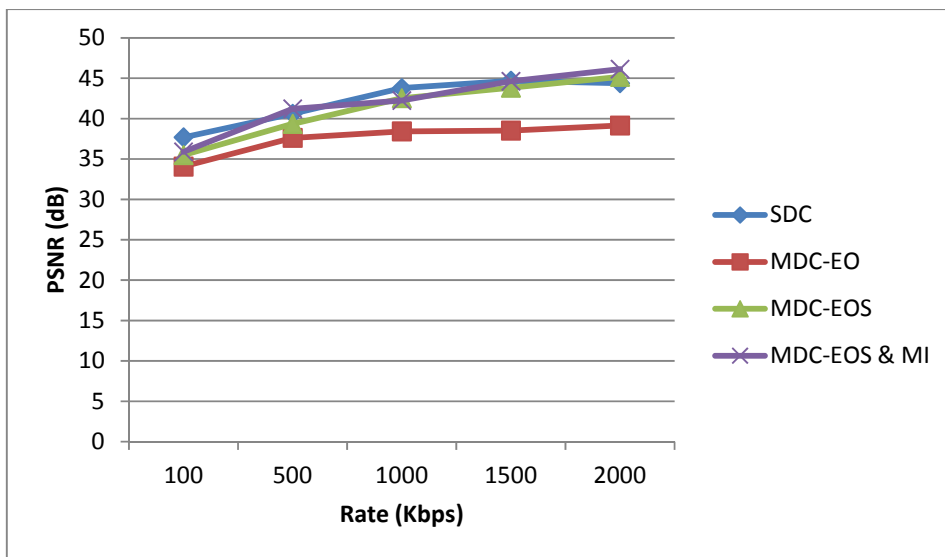


Figure 4-10: Rate Distortion for Interview Luminance video sequence under error free conditions

3D MDC with Side Information and Motion Interpolation

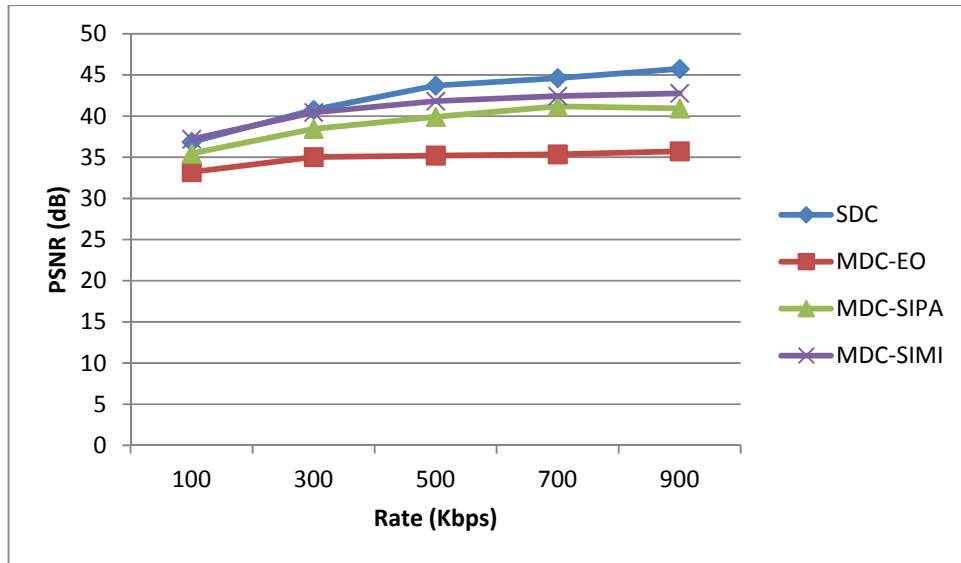


Figure 4-11: Rate Distortion for Interview depth sequence under error free conditions

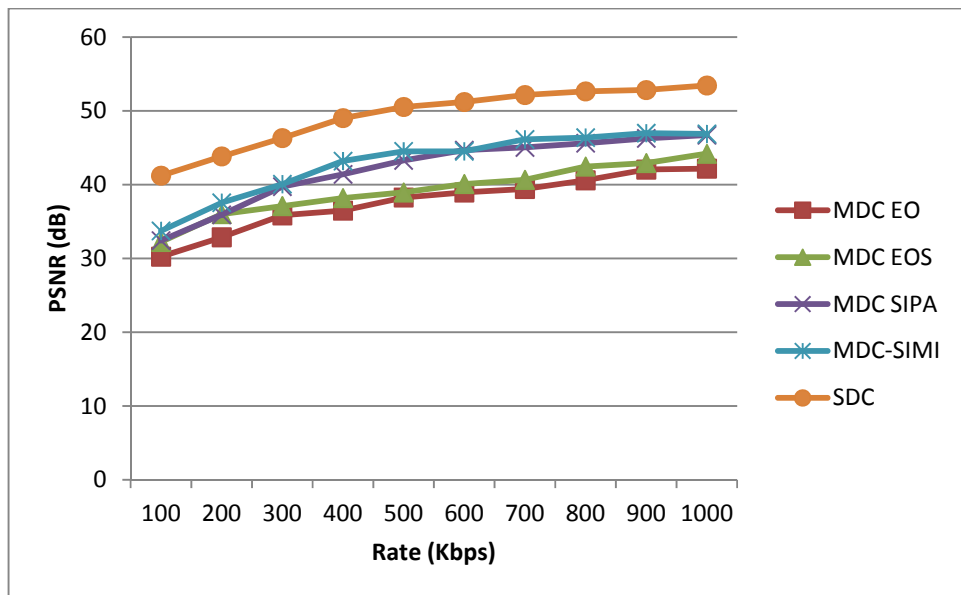


Figure 4-12: Rate Distortion for interview 3D sequence under error free condition.

4.6.4 Objective Performance under Error Prone Conditions for all the MDC algorithms

Here, the performance of SDC, MDC-EO, MDC-SIPA and MDC-SIMI are evaluated under error conditions. The compressed 3D video is transmitted over a simulated UMTS channel. Three packet loss error patterns were used i.e., 5%, 10% and 20%. In the simulation runs, the loss of a packet is equivalent to the loss of one video frame. Frame copy error concealment is used in

3D MDC with Side Information and Motion Interpolation

SDC and MDC-EO, while MDC-SIPA and MDC-SIMI uses frame and motion interpolation techniques when error occurs. The Interview sequence is used in the simulation.

Figure 4-13 shows the rate distortion performance of the Interview sequence for packet loss of 10%. In this figure, the decoded luminance, depth and 3D video Averaged PSNR for SDC, MDC-EO, MDC SIPA and MDC-SIMI are plotted. It can easily be seen that MDC –SIMI has better performance compared to MDC-SIPA at 10% packet loss. Our method results in better compression performance by up to 2dB PSNR or more in some cases in a network with packet losses.

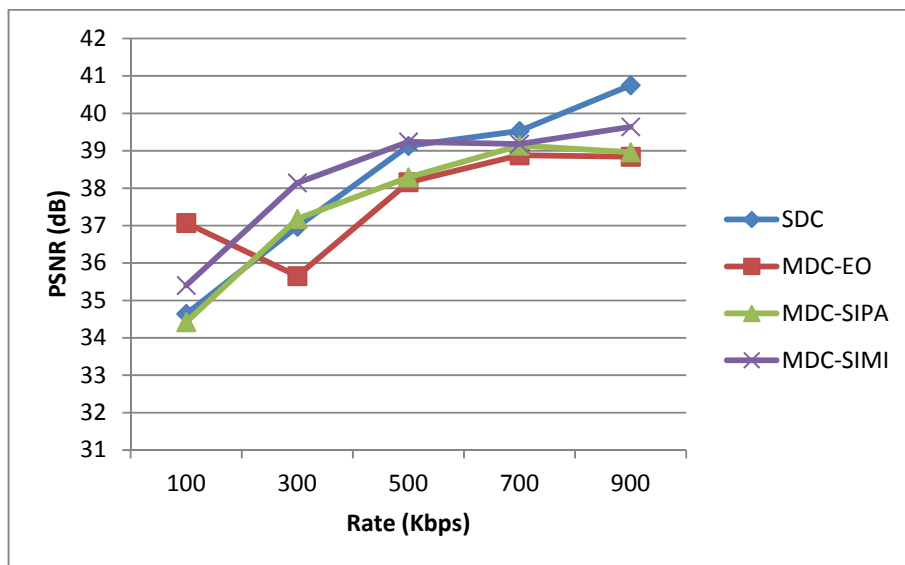


Figure 4-13: Mean PSNR vs Bitrate (Kbps) for Interview 3D video at 10% packet loss under error conditions.

4.6.5 Subjective Performance under Error Free Conditions

Figures 4-14, 4-15, and 4-16 show selected decoded frames from Interview sequence under error free conditions. Each video frame has the same number of packets. We use the same test conditions for all the video frames. The luminance subjective quality of frame 80 for the Interview sequence under error free conditions is shown in figure 4-14. The depth of frame 52 PSNR for MDC-SIPA and MDC-SIMI are plotted in figure 4-15. The 3D Stereoscopic video quality can be obtained from the combination of the luminance and depth as reported in [7][37]. Figure 4-16 shows stereoscopic 3D video quality with MDC-SIMI and MDC-SIPA under error free condition for frame 16. The subjective quality for 4-16(a) and 4-16(b) are almost the same

3D MDC with Side Information and Motion Interpolation

for the two algorithms. Figure 4-16 can be viewed using the red and blue stereoscopic glasses which are readily available and discussed in chapter 3.

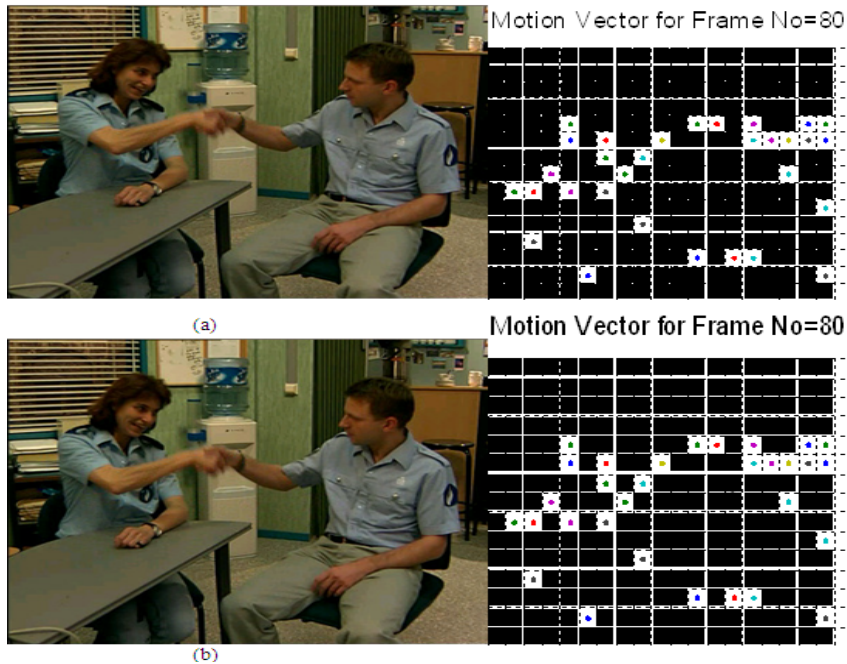


Figure 4-14: Subjective quality-Interview at no packet loss of luminance for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 80

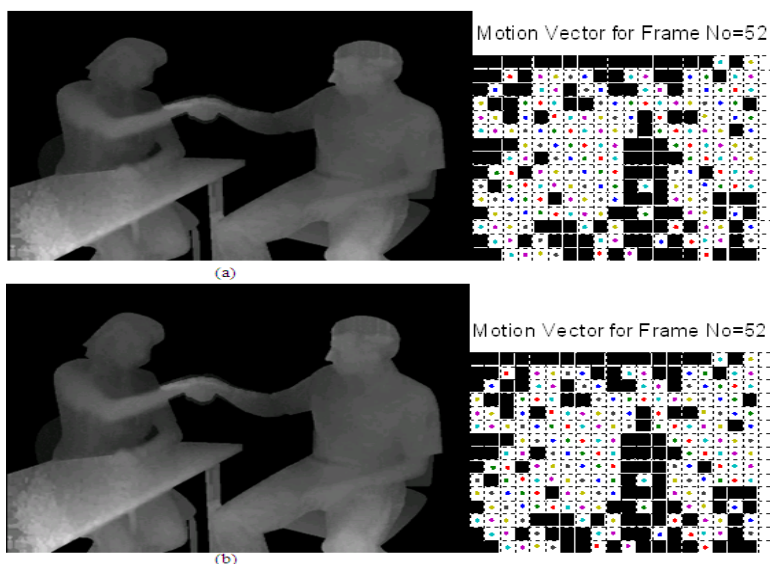


Figure 4-15: Subjective quality-Interview at no packet loss of depth for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 52.

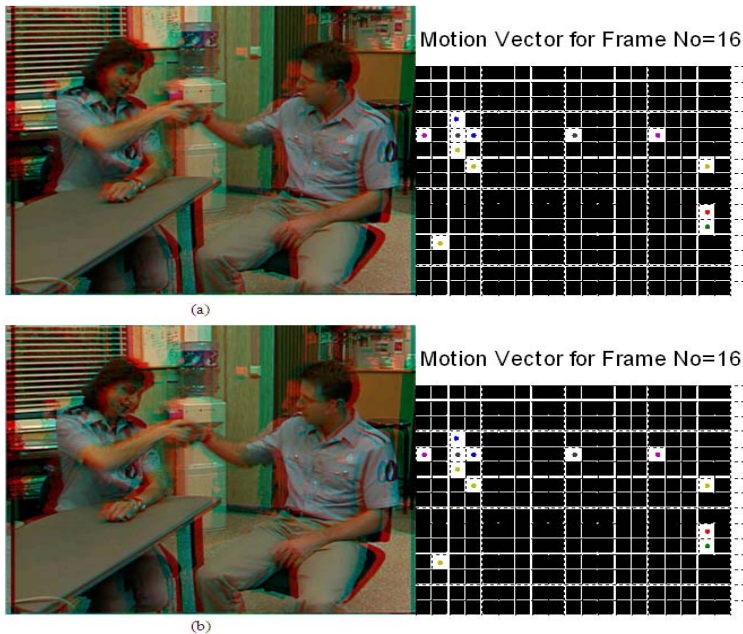


Figure 4-16: Subjective quality-Interview at no packet loss of Stereoscopic 3D video for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 16.

4.6.6 Subjective Performance under Error Prone Conditions

Here, both luminance and depth video packets are subjected to the same channel conditions and same packet loss-rate, which is 10%. The corrupted bitstream file is decoded. The decoded luminance and depth videos are used to generate the 3D videos. In order to obtain the PSNR, the reconstructed 3D video under packet losses are compared with the luminance and depth video sequences as carried out in [37]. If the depth is lost, by decoding the luminance (YUV) video packets coded at base layer the users still can experience 3-D. Hence, it is important that the luminance video packets are given higher priority than depth video packets in the network. The video quality are shown in figures, 4-17, 4-18 and 4-19 for luminance, depth and reconstructed 3D videos for frame 99, 49 and 13 respectively. When the packets of both luminance and depth videos are exposed to the same error conditions, video quality of the reconstructed 3D video degrades rapidly with increasing packet loss rate.

3D MDC with Side Information and Motion Interpolation

Error prone performance of MDC-SIMI is better than MDC-SIPA as it is clearly visible in figures 4-17, 4-18 and 4-19. Performance gain is achieved by MDC-SIMI over MDC-SIPA for luminance, depth and reconstructed 3D videos. 3D MDC-SIMI is a promising technique to provide better error resilience for 3D video transmission through the use of side information and motion interpolation. Visually, picture shown in figure 4-19(a) has been more affected by packet losses compared to 4-19(b). Figure 4-19 (a & b) can be viewed using the red and blue stereoscopic glasses which are readily available and discussed in chapter 3.

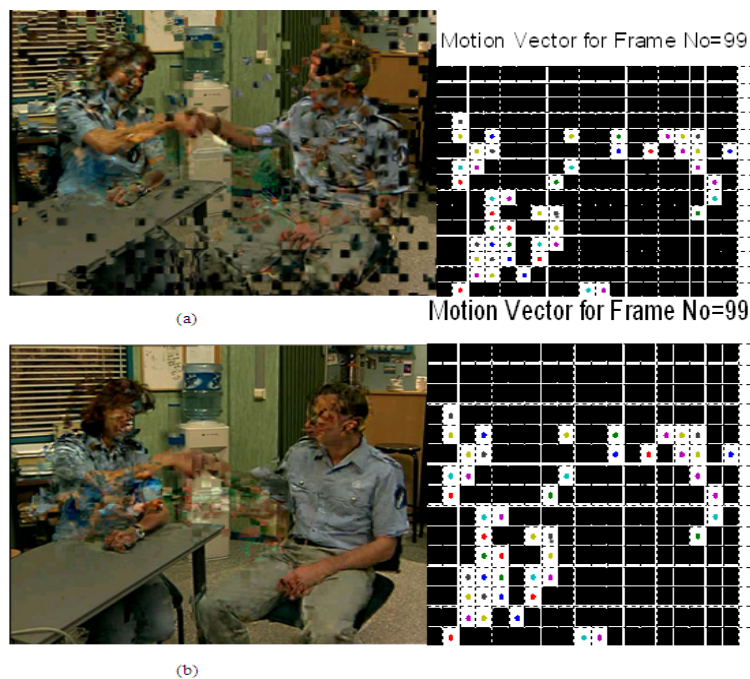


Figure 4-17: Subjective quality-Interview at 10% packet loss of luminance for (a) 3D MDC-SIPA) (b) 3D MDC-SIMI for frame 99.

3D MDC with Side Information and Motion Interpolation

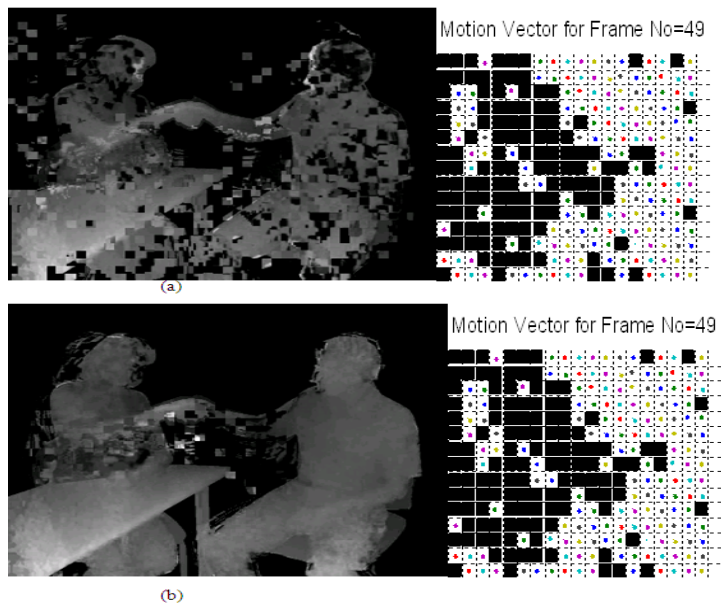


Figure 4-18: Subjective quality-Interview at 10% packet loss of depth for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 49.

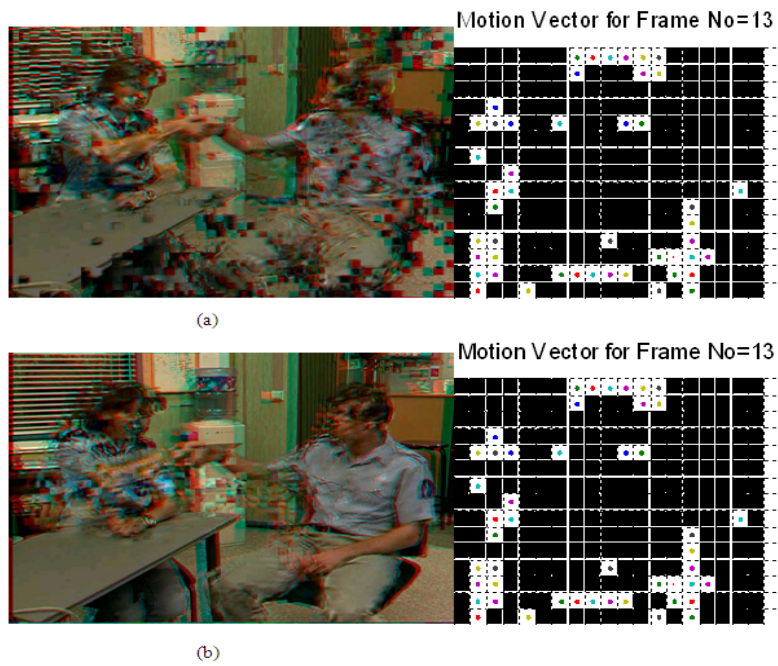


Figure 4-19: Subjective quality-Interview at 10% packet loss of Stereoscopic 3D video for (a) 3D MDC-SIPA (b) 3D MDC-SIMI for frame 13.

4.7 Conclusions

In this chapter, we presented a scalable MDC scheme for stereoscopic 3D video based on even and odd frames with side information and motion interpolation. The first two sections of chapter 4 covered introduction and description of previous related work in 2D and 3D multiple description coding. Next, we presented a method for motion interpolation solutions as used in this thesis. The forward motion estimation, bidirectional motion estimation, spatial smoothing and bidirectional motion compensation were explained.

We proposed 3D MDC-SIMI having the following processes. We start with generating even and odd frames, these are encoded into two streams i.e., 1 and 2. The contents of the two streams include the texture and depth information. We then add side information (overhead) to improve the quality of the decoded frames thereby reducing coding efficiency. Two side decoders are used to perform the interpolation process between the current frame and the previous frame to produce the interpolated frame. The difference between the interpolated frame and the previous reconstructed frame is now DCT quantized to produce coded side information. The combination of the coded side information with either odd or even frames from the central encoder will now produce description 1 or 2 respectively.

To further enhance the error resiliency and concealment of MDC-SIMI, the shape, motion, and texture information in the bit-stream of the videos are organized into different layers such as enhancement and coding layers. This new MDC can take the advantage of classification and priority assignment in the network.

3D MDC-SIMI has been built on top of the highly successful H.264/AVC SVC codec. The proposed algorithm can improve the performance of 3D MDC and gives reliable error resiliency. The R-D behaviour of our algorithm outperforms that of 3D MDC described in [33]. The application of 3D MDC-SIMI results in an improvement in the rate distortion performance particularly in the low bit rate range. Our method results in better compression performance by up to 2dB PSNR or more in some cases in a network with packet losses. The gain is more significant when the packet size is large and the encoding frame rate (Kbps) is lower.

Objective and subjective evaluations confirm that the performance of MDC-SIMI compared to MDC-SIPA, MDC EO and MDC-EOS for stereoscopic 3D video in error free and error prone conditions is performs better than the MDC algorithms retrieved from the literature. Most of the

3D MDC with Side Information and Motion Interpolation

time, 3D MDC-SIMI algorithm performs better in error prone conditions than 3D MDC-SIPA algorithm, most especially at high channel error rates.

However, the proposed method is only applicable to two descriptions. When there are up to four descriptions as against the two descriptions we used, there is a possibility of losing both descriptions as our 3D MDC-SIMI cannot support this.

References

- [1] A. Gersho, “The channel splitting problem and modulo-PCM coding”, Bell Labs, Technical report, Oct. 1979.
- [2] V. K. Goyal, “Multiple Description Coding: Compression Meets the Network”, IEEE Signal Processing Magazine, vol. 18, issue 5, pp. 21-24 Sept. 2011.
- [3] S. E. Miller, “New transmission configuration”, Bell Labs, Technical report 55637, May 1978.
- [4] M. Schwartz, “Information Transmission, Modulation and Noise” 4th edition, McGraw-Hill, Newyork, 1990.
- [5] R. Balan, I. Deubechies and V. Vaishampayan, “The analysis and design of windowed Fourier frame based MDC source coding schemes”, IEEE Transactions. Information Theory, vol. 46, no. 7, pp. 2491-2536, Nov. 2000.
- [6] D. O. Reudink, “The channel splitting problem with interpolative coders”, Bell Labs, Technical Report TM80-134-1, Oct. 1980.
- [7] V. A. Vaishampayan, “Design of MDC scalar quantizers”, IEEE Trans. Inform. Theory, vol. 39, pp. 821-834, May 1993.
- [8] J. H. Conway and N. J. A. Sloane, “Fast quantizing and decoding algorithms for lattice and codes”, IEEE Transactions Information Theory, vol. 28, no. 2, pp. 227-232, Mar. 1982.
- [9] A. R. Reibman, H. Jafarkhani, Y. Wang, M. T. Orchard and R. Puri., “Multiple description video coding using motion-compensated temporal prediction”, IEEE Transactions Circuits And System For Video Technology, vol. 12, no. 3, pp. 193-294, Mar. 2002.
- [10] Y. C. Lee, Y. Altunbasak, and R. M. Mersereau, “An enhanced two stage multiple description video coder with drift reduction“, IEEE Trans. Circ. And Sys. for Video Technology, vol. 14, no. 1, pp. 122-127, Jan. 2004.

3D MDC with Side Information and Motion Interpolation

- [11] Y. Wang, M. T. Orchard, V. Vaishampayan, R. Reibman, "Multiple description coding using pairwise correlating transform", IEEE Trans. Image Processing, vol. 10, no. 3, pp. 351-366. March 2001.
- [12] R. Knoerig, T. Sikora, "MDC image coding using cascaded correlating transform", In Proc. of Multimedia on Mobile Devices, IS&T/SPIE 19th Annual Symp. Electronic Image, CDROM, California, USA. Jan. 2007.
- [13] S. Somasundaram, K. P. Subbalakshmi, "3D Multiple Description Video Coding for packet switch networks", In Proc. of Int. Conf. On Multimedia and Expo, vol. 1, pp. 589-592. Jul. 2003.
- [14] I. V. Bajic, J. Woods, "Domain-based Multiple Description Coding of Images and Video", IEEE Trans. Image Processing, vol. 12, no. 10, pp. 1211-1225. Oct. 2003.
- [15] N. V. Boulgouris, K. E. Zachariadis, A. N. Leontaris, M. G. Strintzis, "Drift-free Multiple Description Video Coding", In Proc. of IEEE Workshop on Multimedia Signal Processing, pp. 4034-4037, France. Oct. 2001.
- [16] N. Gogate, D. M. Chung, S. S. Panwar, Y. Wang, "Supporting Image and Video Applications in a Multihop Radio Environment, using path diversity and multiple description coding", IEEE Trans. Circuits And Systems for Video Technology, vol. 12, no. 9, pp. 777-792, Sept. 2002.
- [17] N. Franchi, M. Fumagalli, G. Gatti, R. Lancini, "A novel error-resilience scheme for a 3D Multiple Description Video Coder", in Proc. Of Picture Coding Symposium (PCS04), pp. 373-376, CA, USA. Dec. 2004.
- [18] H. A. Karim, S. Worrall, and A. M. Kondoz, "Reduced resolution depth compression for Scalable 3D Video Coding", 5th International Conference on Visual Information Engineering (VIE 2008), July 29 -Aug. 1 2008, Xi'an, China.
- [19] H. A. Karim, C. Hewage, S. Worrall, and A. M. Kondoz, "Scalable multiple description video coding for stereoscopic 3D", IEEE Trans. Consumer Electronics, vol. 54, no. 2, pp. 745-752, May 2008.
- [20] C. S. Kim, S. U. Lee, "Multiple Description Coding of Motion Fields for Robust Video Transmission", IEEE Trans. Circuits And System for Video Technology, vol. 11, no. 9, pp. 999-1010, Sept. 2001.

3D MDC with Side Information and Motion Interpolation

- [21] J. G. Apostolopoulos, "Error-resilient video compression via multiple state streams" In Proc. of Int. Workshop on Very Low Bit rate Video Coding (VLBV99), Kyoto, Japan, October 1999.
- [22] S. Wenger, "Video redundancy coding in H.263+", In Proc. of Workshop on Audio-Visual Services for Packet Networks, Aberdeen, Scotland, September 2007.
- [23] Y. Wang, S. Lin, "Error-resilient video coding using multiple description motion compensation", IEEE Transactions Circuits And System For Video Technology, vol. 12, no. 6, pp.438-452. June 2002.
- [24] M-T. Lu, C-K. Lin, J. Yao and H. Chen, "Multiple description coding with spatial-temporal hybrid interpolation for video streaming in peer-to-peer networks", Journal of Zhejiang University of Science B, vol. 7, no. 5, pp. 894-899, April 2006.
- [25] G. Zhang, R. L. Stevenson, "Efficient error recovery for multiple Descriptions Video Coding", In Proc. of IEEE Int. Conf. on Image Processing, pp. 829-832. Oct. 2004.
- [26] S. Ekmekci and T. Sikora, "Multi-state video coding with side information", in Proc. Of Asilomer Conference on Signals, System and Computers, pp. 874-878, Oct. 28-Nov. 1 2005.
- [27] A. Norkin, A. Aksay, C. Bilen, G. Bozdagi Akar, A. Gotchev and J. Astola, "Schemes for multiple description coding for stereoscopic video", In Proc. LNCS, Multiple Content, Representation and security, vol. 4105, pp. 730-737, Istanbul, Turkey, Sept. 2006.
- [28] H. A. Karim, A. S. Sadka, S. Worrall, and A. M. Kondo, "3D Video error resilience using B-frame multiple description coding" 2nd Workshop on Immersive Communication and Broadcast Systems, (ICOB 2005), Berlin, Germany, October 2005.
- [29] H. A. Karim, A. S. Sadka, S. Worrall, and A. M. Kondo, "3D Video error resilience using B-frame multiple description coding" In Proc. of Source Coding and Reliable Delivery of Multimedia Contents, Toulouse, pp. 21-24, May 2006.

3D MDC with Side Information and Motion Interpolation

- [30] H. A. Karim, C. Hewage, S. Worrall, A. Kondo, “Scalable multiple description video coding for stereoscopic 3D”, IEEE Trans. Consumer Electronics, vol. 54, no. 2, pp. 745-752, May 2008.
- [31] H. A. Karim, C. T. E. R. Hewage, A. C. Yu, S. Worrall, S. Dogan, A. Kondo, “Scalable multiple description 3D video coding based on even and odd frame”, Picture Coding Symposium 2007, Lisbon, Portugal, Nov. 2007.
- [32] S. Adedoyin, W. A. C. Fernando, H. A. Karim, C. T. E. R. Hewage, A. M. Kondo, “Scalable Multiple Description Coding with Side Information Using Motion Interpolation”, IEEE Trans. On Consumer Electronics, vol.54, no. 4, pp. 2045 -2052, November 2008.
- [33] H. Abdul Karim, S. Worrall, Abdul H. Sadka, “ Multiple Description Video Coding for Stereoscopic 3D ”IEEE Transactions on Consumer Electronics, vol. 55, no. 4, pp. 2048-2056, November 2009.
- [34] S. Adedoyin, W. A. C. Fernando, “Scalable MDC for 3D Stereoscopic Video Using Motion Vector Encoding”, In Proc. of ICME 2010, Suntec city, Singapore, pp. 1718-1723, 19-23rd July 2010.
- [35] M. Pinson, S. Wolf, “A New Standardised Method for Objectively Measuring Video Quality”, IEEE Trans. On Broadcasting, vol. 50, no.3, pp. 312-322, Sept. 2004.
- [36] S. Muhammed. <http://www.irisa.fr/armor/lesmembres/Mohamed/Thesis.pdf> , Accessed Nov. 2011.
- [37] Z. Wang, L. Lu, A. Bovic, “Video Quality Assessment using structural distortion measurement”, Signal Processing: Image Communication, special issue on “Objective video quality metrics”, vol. 19, no.2, pp. 121-132, Feb. 2004.
- [38] D. Niranjan, et al, “Image Quality Assessment based on a degradation model”, IEEE Trans. On Image Processing, vol. 9, no. 4, Apr. 2000.
- [39] J. Martens, L. Meesters, “Image dissimilarity”, Signal Processing, vol. 70, pp. 155-176, nov. 1998.
- [40] VQEG.. (March 2000) Final Report From the Video Quality Expert Group on the validation of Objective Models of Video Quality Assessment. Available on: <http://www.vqeg.org/> . Accessed Nov. 2011.

3D MDC with Side Information and Motion Interpolation

- [41] C. Fan, L. Fan, Y. Wang, “A New Side Information Algorithm based on Block Interpolation for Distributed Video Coding”, In Proc. of Information Communication and Signal Processing, Macau city, China, 8-10 December 2009.
- [42] http://www.akmalabs.com/downloads_netsim.php . Accessed Apr. 2010.
- [43] J. Ascenso, C. Brites, F. Pereira, “Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding” 5th EURASIP, Slovak Republic, July 2005.
- [44] H. Margaret, S. Wolf, G. Cermak, “HDTV Subjective Quality of H.264 vs. MPEG-2, with and without Packet Loss: www.its.bldrdoc.gov/pub/n3/video/ieee_10.pdf , Accessed March 2011.
- [45] <http://www.hhi.fraunhofer.de/en/departments/image-processing/image-communication/video-coding/rate-distortion-optimization/> . Accessed on 2 dec. 2011.
- [46] A. Hore, D. Ziou, “Image quality metrics: PSNR vs SSIM”, In Proc. of Int. conference on Pattern Recognition”, pp. 2366-2369, 2010.

Chapter 5: Quality Assessment of 3D Videos

5.1 Introduction

Subjective Quality of Experience (QoE) is defined as the perceived quality of a service received by users. Subjective tests are conducted under controlled conditions to rate video quality by users/subjects. The rating expresses the subjective QoE described by the Mean Opinion Scores (MOS). Colour plus depth map based stereoscopic video has attracted significant attention in the last 10 years. Quality assessment of coded 3D video sequence can currently be performed reliably using expensive and inconvenient subjective tests [1]. The main goal of many subjective video quality assessments is to automatically estimate average user or viewer opinion on the quality of video processed by the system. However, measurement of subjective video quality can be challenging because it may require a trained expert to judge it. Many subjective video quality metrics are described in ITU-T recommendation BT.500 [2]. The main idea is to obtain the Mean Opinion Score (MOS) for video sequences which are showed to the group of viewers. The viewers opinions are recorded and averaged to evaluate the quality of each video sequence. Optimisation of 3D video systems in a timely manner is very important; it is therefore necessary that reliable subjective measures are calculated based on statistical analysis. This thesis investigates subjective assessments for four standard 3D video sequences. Subjective tests are performed to verify the 3D video quality and depth perception of a range of differently coded video sequences, with packet loss rates ranging from 0% to 20%. The proposed measurement of 3D perception and 3D quality of experience (QoE) is shown to correlate well with human perception of quality on a publicly available dataset of 3D videos and human subjective scores.

5.2 Review of Subjective Quality Assessment Methods

Quality assessment of 3D video and stereo images is achievable either through subjective assessment or using objective metrics. The best way to assess video quality would surely be to run subjective tests according to standardized procedures, which are defined in order to obtain correct universal and reliable video quality evaluation by participants [2]. The human, as the receiver of video sequences, is the basis of subjective video quality assessment. The perfect and

Quality Assessment of 3D Videos

desired quality measures by human. However, the use of subjective test is time consuming due to subjective test preparation, requirements and cost. Furthermore, the analysis of collected data is not a straight forward process, and there is a growing demand to use objective quality assessment (OQA) approach. With the OQA approach, computer algorithms are used to estimate the quality of the product instead of human subjects [2]. The existing 2-D objective quality measurements of individual colour and depth map sequences may not represent the true image quality as perceived by the human viewers. Although some researchers use PSNR for assessing 3-D videos, the limitations of PSNR for 2-D video have been demonstrated in [1][2]. These limitations are likely to be similar for the assessment of image quality in 3-D video; in addition, PSNR does not give information about the depth perception. Therefore, we use the subjective quality assessment (SQA) to confirm the results obtained by objective quality assessment (OQA) of 3D video sequences or images.

Efforts has been made by both the academic and industrial communities to develop subjective test models to evaluate both 2D and 3D video sequences. Moreover, when image and video delivery takes place in an error prone environment, subjective quality assessment can be used as side information for the image and video server to take the necessary actions to improve the quality of reconstructed data. With the widespread application of 3D technologies to different fields such as sports/life events, films, TV series and documentaries, medical applications, gaming etc, 3D images and videos need to be processed, transmitted and distributed to various users. Therefore, it is important to define both subjective procedures and objective metrics to use for assessment of the quality of the processed stereoscopic videos. Indeed, compared to 2D videos, perception of stereoscopic video content involves several factors. The multidimensional attributes such as blocking and noise associated with 2-D video cannot be used in measuring the perceptually important qualities of 3-D video such as notion of presence, sharpness, naturalness and depth perception which is related to the sensation of immersion in the 3D visual scene [1]. We consider two issues when conducting a subjective test: 1) The impact of technology on the observer's viewing experience 2) The factors to be taken into account to measure 3D video quality and how they impact on visual perception. Subjective experiments must be conducted to understand these issues and other related problems.

Taking into account these considerations, we limit our study to stereoscopic 3D videos. The subjective assessment is addressed within this context taking care of the heritage of 2D image

Quality Assessment of 3D Videos

and video quality assessment. We also use depth information to design our subjective procedures for 3D quality assessment.

In this research reported in the thesis, subjective testing using human viewers is used to measure the mean opinion scores (MOS), 3D video satisfaction, user experience, quality of experience (QoE), preference, comfort, depth presences etc. based on the perceived quality of the reconstructed stereoscopic video and using the designed questionnaire for data collection. The perceptual attributes such as overall quality and depth perception are used to measure the response from human observers. The relationship between the perceived overall image quality and the perceived depth is also discussed in this thesis. Finally, the perceived attributes of reconstructed stereoscopic video are predicted using different objective quality metric (OQM). The results help to answer important questions concerning the correlation between objective and subjective measurements of 3D video quality, and also concerning the best methods for measuring the quality of 3D video subjectively in error free and error prone conditions.

In [3], a wide variety of subjective tests have been conducted to identify how depth information retrieval, crosstalk, depth representation and 3D compression impact on the 3D video quality, naturalness, viewing experience, presence and visual strain. In this research/thesis however, the final viewer Difference Mean Opinion Scores (DMOS) depends on the video having the highest quality. Therefore, the perceived quality of a stereopair whose images have been distorted strictly depends on the applied distortions, which is related to the level of the human visual system (HVS) masking effects. Tests were also carried out in order to identify the impact of eye dominance. In [4][5][6], no effect of eye dominance was noticed for the image quality evaluation. In [7], it was observed that eye dominance improves the performance of visual search tasks by aiding visual perception in binocular vision, and the eye dominance effect in 3D perception and coding was also analyzed.

In [8], a depth perception threshold model was designed. The impact of depth information on the perceived 3D image quality is one of the main issues that have to be investigated. Recent studies show that depth is not related to the perceived 3D effect. However, other studies point out the importance of depth for quality perception. In [9], a blurring filter whose blur intensity depends on the depth of the area where it is applied is used to enhance the viewing experience. This was validated in [6] which show that blurring 3D images reduces discrepancy between responses of

Quality Assessment of 3D Videos

accommodation and convergence, so that blur increases viewer's experience. Also, methods which aim at enhancing the local depth information on objects have been proposed where the algorithm directly impacts on the quality by taking into account depth information.

Research in the field of 3D video quality assessment relies on the availability of subjective scores, otherwise called Mean Opinion Scores (MOS), collected by means of experiments where groups of people are asked to rate the quality of 3D video sequences. In order to gather reliable and statistically significant data, subjective tests have to be carefully designed and performed, and require a relevant number of participants. These tests are quite time consuming, however, subjective data is fundamental to test and compare the performance of the objective algorithms, i.e metrics, which try to predict human perception of 3D video quality by analyzing the 3D video streams.

Examples of comparative studies of objective video quality performance are those carried out by VQEG [1][2], based on the results of two extensive campaigns of subjective tests which involved many laboratories. Unfortunately, the subjective results and the test material used to perform these studies have not been made public, thus VQEG subjective results cannot be used by independent researchers for testing of more recent and future performance. Also, many studies are available in the literature, reporting results of subjective experiments, such as [3][4] which investigate quality degradation in 3D video streaming applications, but none of these two provides access to the collected data for usage.

At the best of author's knowledge, the only publicly available databases of subjective results and related test material, in the field of visual assessment, are the LIVE database [12] for standard definition images, and the Politechnique Federale de Lausanne (EPFL) database for high resolution images [13]. The EPFL database includes CIF video sequences coded with H.264/AVC reference software and impaired by simulating packet losses over an error-prone channel. The subjective data has been gathered at the premises of two academic institutions: Politecnico di Milano (PoliMI) – Italy, and Ecole EPFL. In addition to subjective data, the database includes the original video contents and configuration files used to encode them, the original and corrupted H.264/AVC bitstreams, as well as the network simulator used to generate the test materials.

5.3 Research Methodology

We applied survey methodology for the elicitation of participants/user requirements. Surveys are commonly used as a method to perform subjective analysis of 3D video sequences based on the previously designed questionnaire. It helps us identify needs, requirements, expectations, overall 3D video quality, viewing comfort, perception, preference and satisfaction of 3D users. The designed survey also helps to identify weakness of this new technology [14]. The designed survey contained 12 sections and background information of each participant. Prior to the use of the questionnaire, an expert evaluation and review was undertaken by experts in the field of 3D video communication to improve its content. It should be noted that this research was aimed at elicitation of participant's view on the subjective assessment of 3D video. It is neither market research nor cross-cultural research since the samples were not representative of the whole population.

Our hypothesis is that natural 3D videos/images have certain statistical properties that are interpreted as natural by the human observer. Deviations from this naturalness may lead to discomfort in the perception thereby reducing the QoE. We will attempt to capture these deviations from naturalness using simple but very powerful tools/statistical measures such as the mean, variance, skew and other indicators. Changes in camera distance will change the statistical distributions of 3D images and our hypothesis is that these changes are related to the perceived 3D quality and QoE. Apart from statistics computed from the depth maps, we also compute spatial statistics from the left-right views in order to ensure that masking effects due to content [15] which generally influence perception. For each 3D image/video (Left-Right pair, I_l , I_r + *depth map D*), we compute the following statistical features from the depth:

1. Mean depth $\mu = E[D]$,
2. Median depth $med = median [D]$,
3. Depth standard deviation $\sigma = \sqrt{E[(D - \mu)^2]}$
4. Kurtosis of depth $k = E[(D - \mu)^4]/(E[(D - \mu)^2])^2$
5. Skewness of depth $skew = E[(D - \mu)^3]/(E[(D - \mu)^2])^{3/2}$
6. Mean differential depth $\mu_d = E[\sigma D]$
7. Differential depth standard deviation $\delta_d = \sqrt{E[(\sigma D - \mu_d)^2]}$

Quality Assessment of 3D Videos

8. Kurtosis of differential depth $k_d = E[(\sigma D - \mu_d)^4]/(E[(\sigma D - \mu_d)^2])^2$
9. Skewness of differential depth $skew_d = E[(\sigma D - \mu_d)^3]/(E[(\sigma D - \mu_d)^2])^{3/2}$

Where the differential depth (σD) was computed using a Laplacian operator on the depth map. Differential depth statistics are computed in order to capture changes in depth information [16]. To capture the nature of spatial content of the scene, we compute spatial activities for I_l and I_r from the left and right images/views. Specifically, we compute the gradient of the image and estimate the variance of non overlapping 8x8 blocks across the image. We compute the following from the map of spatial activities:

1. Mean depth $\mu = E[D]$
2. Kurtosis of depth $k = E[(D - \mu)^4]/(E[(D - \mu)^2])^2$
3. Skewness of depth $skew = E[(D - \mu)^3]/(E[(D - \mu)^2])^{3/2}$

Such computation is undertaken for all the images/videos. However, 3D perception is characterized by eye dominance effect hence we choose to retain all the individual statistics. Apart from these spatial features, 3D videos are characterized by motion information. Motion information is very important for human perception, therefore human visual system (HVS) devotes a significant amount of processing to extract motion estimation [17]. Once the motion estimation is computed, they are pooled across each frame by computing the coefficient of variation within the video frame. We note that the coefficient of variation has been used for pooling quality scores within a frame with success [18]. Finally, in order to pool these frame level scores across the videos, the median, standard deviation, kurtosis and skewness of these motion-compensated differences are computed across frames and are stacked together with the computed spatial statistics. Such computation is performed twice since flows are computed on both the left and right videos separately.

The subjective human trials based on the ITU-R recommendation 500, assessed the quality of compressed 3D videos. The assessors viewed the video clips and were asked to complete the questionnaire immediately afterwards and rate the videos based on answering the questionnaire. After the completion of the sessions, the mean opinion scores (MOS) for each test was calculated and all data was entered into SPSS for analysis [31].

5. 3.1 Equipment used and viewing conditions

In the experiment, we displayed the videos and their depth on a 32" display monitor. The screen had a resolution of 1152x 900 pixels, frequency of 60 Hz x 60 Hz and a measured size of 360 x 270mm. The optical path length is 320mm and the horizontal FOV is 0.8 rad. The monitor was at a viewing distance of 2m. Each subject had to judge all the image scenes. The 3D videos were viewed using colour filter anaglyph. This displayed a full resolution to each eye. Participants were asked to sit in a particular place to see the 3D effects. All the depth maps shown were stored as disparity from the original images. Experiment was conducted in a room with minimum light level and based on VQEG requirements as contained in [2].

5. 3.2 Databases of Subjective results & test materials used in the study

The datasets that we used for this experiment were those that have recently been made public by researchers at EPFL [19][20], Henrich-Hertz- Institute [21] with resolution 720x576 pixels which is the resolution for standard definition TV. These sequences were encoded using 3D MDC-SIMI with QP =5. The original frame rates of these sequences is 25 fps (frames per second) and 250 frames i.e. 10 seconds long sequence. We also used datasets from Visual Media Research Group at Microsoft Research [22] as shown in figure 5-1. They are 10-second long sequences (150 frames). The EPFL 3D video database consists of stereoscopic images. We use C_6^4 images for training, C_6^2 images for validation and C_6^3 images for testing. Since we wanted to show that the proposed method is robust across contents, all possible combinations of the dataset were used to form the above mentioned training, validation and test sets.

Quality Assessment of 3D Videos

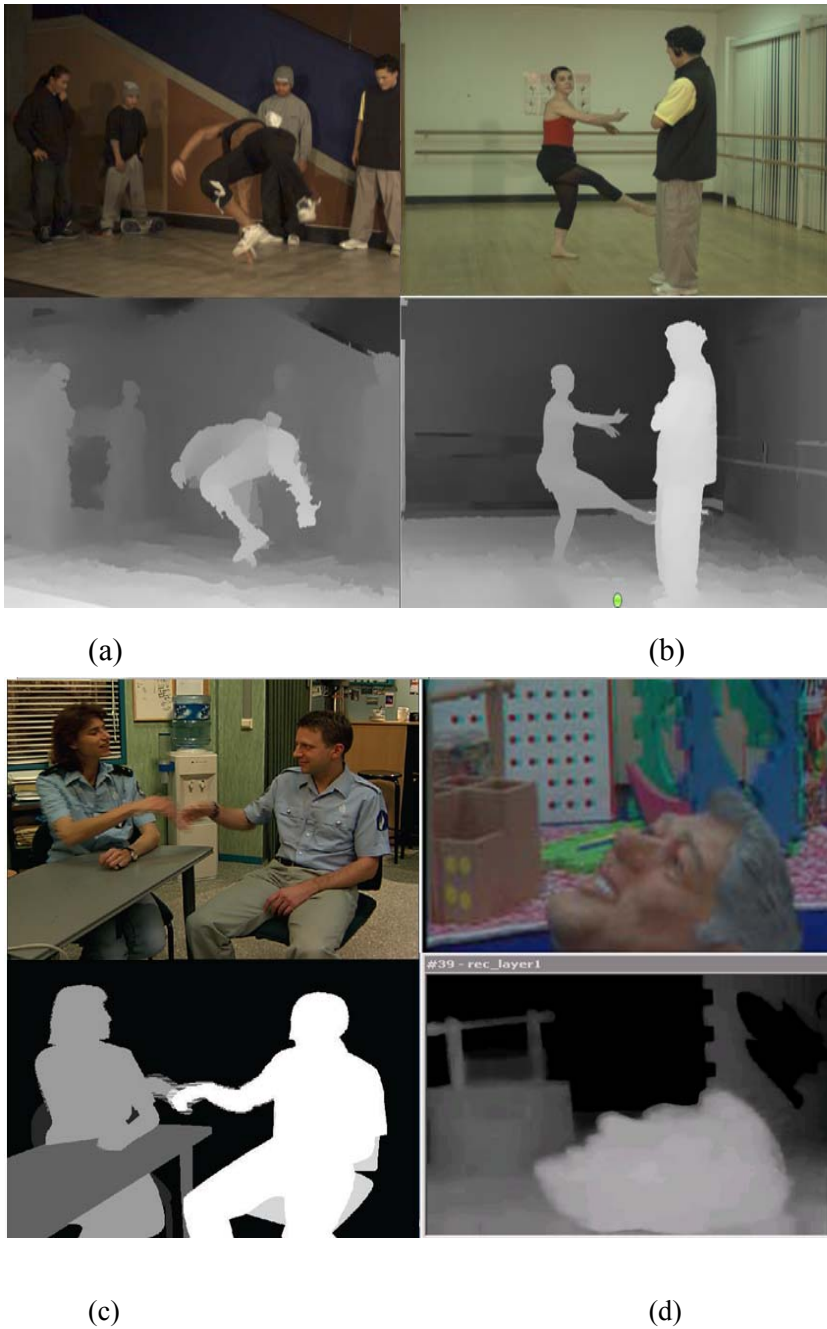


Figure 5-1: Colour and Corresponding Depth Image (a) Breakdance, (b) Ballet, (c) Interview and (d) Orbi sequence.

5.3.3 Participants

In order to gauge human perception of quality and perceived depth, a total of 25 candidates (Group 1) with 18 males and 7 females were recruited from the Charville Lane estates, of

Quality Assessment of 3D Videos

Middlesex, UK. Their ages varied from 15 to 50 with a mean age of 23 years. All the participants had a visual acuity of ≥ 1 , good stereo vision < 60 seconds of arc and good colour vision. The stimulus set contains SVC-coded video sequences and the original uncompressed version of each scene was used as the reference in the evaluation test. Stereo image pairs were obtained by coding of colour and depth image sequences.

Group 2 comprised of 9 non expert observers (6 males and 3 females) who volunteered to participate in the CMCR lab experiment at Brunel University; the observers in this category were mostly research students including few undergraduate students with a relevant technical background. Their ages range from 20-40. Four observers had prior experience with 3D video content using different viewing aids such as shutter glasses and red/blue anaglyph glasses.

Group 3 experiments were conducted in Nigeria, NASRDA office, Abuja with 11 non expert observers (7 male 4 females) who agreed to participate in the experiment. The observers in this category were mostly NASRDA staff, some with technical background. Their ages range from 26-45. Seven of them had prior experience with 3-D format. All participants had a visual acuity of ≥ 1 , good stereo vision < 60 s of arc and good colour vision [23]. The attributes namely perceived image quality and perceived depth were assessed by all the observers. Table 5-1 showed the distribution of the three groups

Table 5-1: Distribution of participants in the survey

S/N	Group	Age Range	Male	Female	Sub Total
1	Charville Estate,	15-50	18	7	25

Quality Assessment of 3D Videos

	Middlesex				
2	CMCR Lab, Brunel	20-40	6	3	9
3	NASRDA office, Abuja, Nigeria	26-45	7	4	11
4	Total Participants				45

The negative effect of viewing 3-D for a long time was also evaluated; symptoms such as general discomfort, fatigue, headache, eye-strain, difficulty in focusing and blurred vision were all included in the questionnaire and are assessed by the observers. All experiments were conducted based on ITU-R requirement of subjective video test, in that the participants were not directly involved in 3D video quality assessment or analysis in their normal work, and were not experienced video/image assessors. Their opinion of 3D video quality may differ from that of people familiar with this technology. Participants were not aware of the purpose of the experiment. They all received a nominal payment of £5 each for participating and spending their time.

5. 3.4 Health and Safety

In addition to human factors related to 3D perceptions, it is important to identify all the cues related to preventing human vision performance degradation of such display technologies. Indeed, some recent studies [24][25] revealed some possible problems caused by 3D displays like decline of visual functions after experiments, requiring vengeance adaptation on 3D content. Also, asymmetrical image distortions can cause vision degradation such as myopia increase [26]. Some ophthalmologists remain concerned that viewing stereoscopic images may cause strabismus, an abnormality in binocular alignment in young children. However, there is no evidence that the fruition of stereoscopic images causes strabismus except for what is reported in [27]. An extensive survey of the potential health problems related to 3D technologies is given in [24].

Quality Assessment of 3D Videos

5.3.5 Protocol

Prior to the start of the experiments, participants received briefing and performed trial tests. This included four 3D videos viewing and rating in, but these practice trials were not included in the final experimental data analysis, as they were just for familiarity purposes.

The participants started the main experiment individually after completing the trial sessions. Participants were requested to be as accurate as possible in their judgment of each 3D video but not to spend too much time on each video, although no time limit was imposed. Afterwards, they were asked to complete the questionnaire based on their judgment. Answers could not be changed once scores were recorded, even before submission.

After the whole experiments and the completion of the questionnaire, participants were debriefed and given a chance to ask questions. The practical viewing of the 3D videos and the completion of the questionnaire took 10 minutes on average and the experiments lasted 3 hours including small breaks and the trial session.

5.3.6 Grading

For each experiment, the 3D videos were rated by participants on a sliding scale of:

- Excellent, Good, Fair, Don't Know, bad
- High impact, moderate impact, less impact, Don't know, No impact.
- Very important, Important, Less important, Don't know, Not important.
- Strongly Agree, Agree, Disagree, Don't know, Strongly disagree.

Participants were asked to assess the overall 3D video quality and overall 3D video Perception Table 6 shows observed results (See heading values column). All the scales were based on ITU_R five point quality scales [28]. Subjective opinion scores obtained from the above studies were averaged across subjects to produce MOS which are representative of the perceived quality of 3-D experience. A total of 24 videos with associated MOS scores are available as part of the two datasets. The approach to no-reference QoE assessment involves extracting relevant features from these visual stimuli and regressing these features onto the MOS. In order to calibrate the process, we divided these datasets into various train test combinations, trained our model and

Quality Assessment of 3D Videos

then tested how well the learned features perform in assessing the QoE. Tables 5-2 and 5-3 shows the variable view and data view respectively.

Table 5-2: Variable view of the MOS.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	2	0	Group id	None	None	8	Right	Scale
2	Group	Numeric	8	0	group of participants	{1, Charville Estate}...	None	8	Right	Nominal
3	gender	Numeric	2	0	Gender	{1, MALES}...	None	8	Right	Nominal
4	age	Numeric	2	0	Participant age	None	None	8	Right	Scale
5	videoequal	Numeric	2	0	3D video quality	{1, Poor}...	None	8	Right	Ordinal
6	Comfort	Numeric	2	0	Viewing comfort	{1, Not important at all}...	None	8	Right	Ordinal
7	depth	Numeric	2	0	perceived depth	{1, No impact at all}...	None	8	Right	Ordinal
8	presencs	Numeric	2	0	feel presence	{1, Strongly disagree}...	None	8	Right	Nominal
9	problems	Numeric	2	0	problems associated with 3D viewing	{1, Strongly disagree}...	None	8	Right	Nominal
10	perceptions	Numeric	2	0	human perceptions of 3D	{1, No}...	None	8	Right	Nominal
11	impact	Numeric	2	0	impact of eye dominance	{1, No impact at all}...	None	8	Right	Ordinal
12	users	Numeric	2	0	overall users view(needs, rgrs, expec, new tech)	{1, Not important at all}...	None	8	Right	Ordinal
13	qoe	Numeric	2	0	quality of experience	{1, Not important at all}...	None	8	Right	Ordinal
14	preference	Numeric	2	0	3D or 2D	{1, 2D}...	None	8	Right	Nominal
15	viewing	Numeric	8	0	type of viewing glasses prefer	{1, Analgypic}...	None	8	Right	Nominal
16	satisfaction	Numeric	8	0	level of satisfaction	None	None	10	Right	Scale
17	agegp3	Numeric	5	0	Age divided into 3 groups	{1, 15-29}...	None	10	Right	Ordinal
18	Influence	Numeric	8	0	Influence of encoding parameters on viewers experience	None	None	8	Right	Scale
19										
20										

Table 5-3: Data view of the MOS

	id	Group	gender	age	videoequal	Comfort	depth	presencs	problems	perceptions	impact	users	qoe	preference	viewing	satisfaction
2	2	1	1	16	4	1	1	1	5	1	1	2	3	3	1	40
3	3	1	1	33	5	2	5	2	4	1	4	2	2	2	2	60
4	4	1	2	18	3	3	5	5	4	1	4	3	3	0	1	30
5	5	1	1	46	1	3	5	4	3	1	4	5	2	4	2	70
6	6	1	1	45	2	4	4	3	3	2	4	3	1	3	2	70
7	7	1	2	50	4	1	3	3	4	2	1	4	1	3	2	80
8	8	1	1	35	4	1	1	4	4	2	1	2	1	4	2	55
9	9	1	1	36	4	5	2	4	5	2	3	5	3	2	2	60
10	10	1	1	38	3	5	4	4	5	2	3	2	1	3	3	80
11	11	1	1	39	3	5	4	4	1	1	5	4	3	4	3	80
12	12	1	1	41	3	5	4	5	1	2	1	4	2	2	1	85
13	13	1	1	42	3	5	1	5	1	2	5	3	1	2	1	70
14	14	1	1	43	2	5	4	5	1	1	5	2	4	3	2	65
15	15	1	2	44	2	4	4	5	2	2	5	4	3	3	2	80
16	16	1	2	45	2	4	1	5	2	1	4	4	1	2	2	55
17	17	1	1	21	5	4	5	5	2	2	1	5	1	3	2	60
18	18	1	2	22	3	3	5	5	2	1	2	5	2	3	2	70
19	19	1	1	23	4	3	4	5	4	2	3	4	1	2	2	80
20	20	1	1	17	4	2	4	5	4	1	1	3	2	2	3	85
21	21	1	1	25	4	2	4	2	4	1	1	4	3	3	3	55
22	22	1	2	19	4	5	3	1	4	1	3	3	2	4	3	40
23	23	1	1	27	5	5	3	2	4	2	3	1	3	3	3	45
24	24	1	1	28	4	5	3	2	4	2	4	3	3	2	3	50
25	25	1	2	29	4	4	3	2	4	2	3	4	3	2	2	80
26	26	2	1	25	5	4	5	5	5	5	4	3	4	3	3	75
27	27	2	2	23	4	5	5	5	4	4	4	4	4	4	4	75
28	28	2	1	23	5	5	4	4	5	4	4	3	3	3	3	50
29	29	2	2	37	4	4	4	4	2	2	3	4	4	1	2	75
30	30	2	1	32	4	3	4	4	4	4	5	3	4	4	4	75
31	31	2	2	35	5	4	5	4	4	3	5	4	4	3	3	75
32	32	2	1	27	4	4	5	5	5	5	4	4	4	4	3	75
33	33	2	1	25	4	4	4	4	5	4	4	4	4	3	4	70
34	34	2	1	37	4	4	4	4	3	3	4	4	4	4	4	80
35	36	3	1	22	5	5	5	5	4	4	5	5	4	3	3	80
36	37	3	2	32	5	5	5	4	4	4	5	5	4	4	2	80
37	37	3	1	41	5	5	5	5	4	4	4	5	4	4	2	80
38	38	3	1	39	4	4	5	4	5	4	4	4	4	4	3	80
39	39	3	1	38	5	5	5	3	3	3	5	5	4	4	4	80
40	40	3	1	42	5	5	4	4	4	4	4	5	4	4	3	80
41	41	3	1	39	5	5	5	5	3	3	5	5	5	4	2	80
42	42	3	1	38	4	4	5	4	4	5	5	4	4	4	3	80
43	43	3	2	28	4	3	3	3	3	4	4	5	2	2	3	60
44	44	3	2	30	3	5	3	4	3	3	3	3	3	2	1	65
45	45	3	2	26	4	3	3	3	2	3	2	2	3	1	2	50

5.4 Mean opinion scores and confidence interval

After screening, the results of the test campaign can be summarized by computing the mean MOS for each test condition as:

$$MOS_j = \frac{\sum_{i=1}^N m_{ij}}{N} \quad (5-1)$$

Where N is the number of subjects participants and m_{ij} is the score by subject i for the test condition j . The relationship between the estimated mean values based on a sample of the population (i.e., the subjects that took part in our experiment and completed the survey questionnaire) and the true mean values of the entire population is given by the confidence interval of estimated mean [29]. The 100 x (1- α)% confidence interval(CI) for mean opinion scores were computed using the student t-distribution below:

$$CI_j = \left(\frac{t(1-\alpha)}{2}, N \right) x \sigma_j / \sqrt{N} \quad (5-2)$$

Where

α is the significance level,

$\left(\frac{t(1-\alpha)}{2}, N \right)$ is the t -value corresponding to a two-tailed t-Student distribution with N-1 degree of freedom and a desired significance level α (equal to 1 -deg of confidence). N corresponds to the number of subjects, and σ_j is the standard deviation of a single test condition across subjects and/or participants. The interpretation of a confidence interval is that if the same test is repeated for a large number of times, using a random sample of the population, each time and a confidence interval is constructed every time, then 100 x(1- α)% of these intervals will contain the true value. We computed our confidence intervals for an α equal to 0.05, which corresponds to a degree of significance of 95%. The result from the study clearly shows that the experiments and/or questionnaire have been properly designed, since the subjective rates uniformly span the entire range of quality levels. Also, the confidence intervals are reasonably small, hence proving that the effort required from each participant was appropriate and participants were consistent in their grading [30].

5.5 Results & Analysis

A detailed statistical analysis of the subjective results has been performed. All data was statistically processed using SPSS tool to obtain the MOS by averaging the data of all respondents. In addition the Standard Deviation and the 95% Confidence Interval (CI) were computed. Figure 5-2 shows the distribution of the 3 group under studies. Figure 5-3 shows the population distribution of the 3 groups under studies. Table 5-4 shows the frequency distribution of participants. Table 5-5 and 5-6 shows total valid number of participants and analysis of problems associated with 3-D viewing respectively. Figure 5-4 show the bar chart distribution of the perceived depth, notion of presence and overall 3D video quality. The scores above 50% show that the quality of the compressed videos is good.

Table 5-7 shows the analysis of the MOS in terms of mean, standard deviation, variance, standard error, kurtosis and skewness for a valid number of participants as computed by the software. Figure 5-5 shows the MOS of the 3D video quality.

Through the series of experiments conducted, we analysed all the respondents' feedback and observed how each of the respondent answered the questionnaire questions. The results plotted shows that most of the respondents rated high the subjective qualities of the 3D videos. This was clearly seen in figure 5-4.

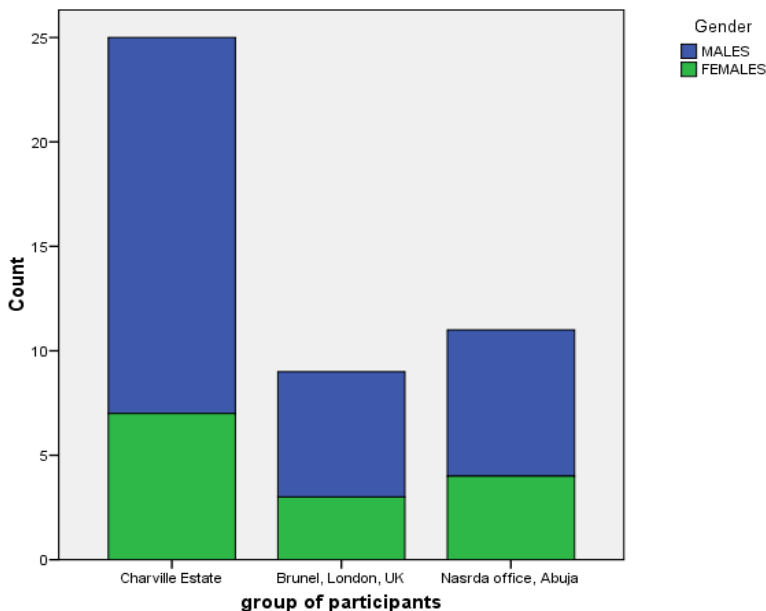


Figure 5-2: Group studies in terms of gender using bar chart

Quality Assessment of 3D Videos

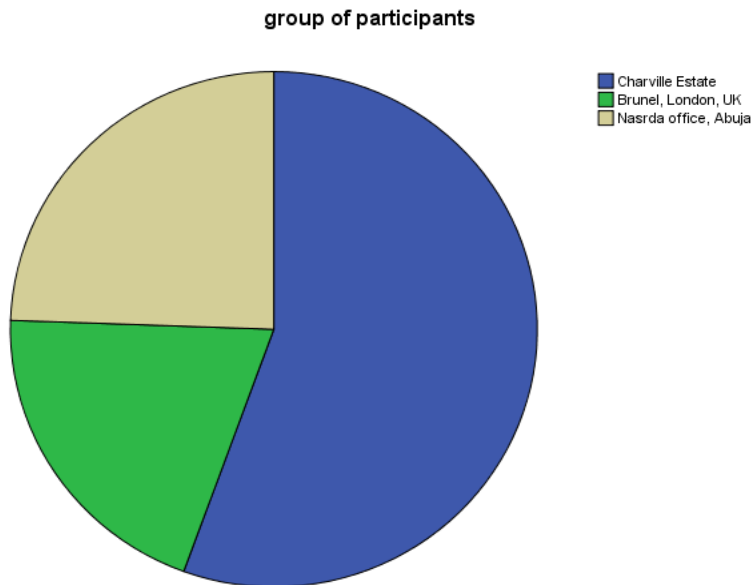


Figure 5-3: Group studies represented in pie chart.

Table 5-4: Frequency Distribution of the 3 Group studies

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Charville Estate	25	55.6	55.6	55.6
	Brunel, London, UK	9	20.0	20.0	75.6
	Nasrda office, Abuja	11	24.4	24.4	100.0
	Total	45	100.0	100.0	

Table 5-5: Total number of Participants

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
group of participants * problems associated with 3D viewing	45	100.0%	0	.0%	45	100.0%

Quality Assessment of 3D Videos

Table 5-6: Analysis of problems associated with 3D viewing

			problems associated with 3D viewing					Total
			Strongly disagree	Disagree	Dont know	Agree	Strongly agree	TOTAL
group of participants	Charville Estate	Count	5	4	2	11	3	25
		Expected	2.8	3.3	3.9	10.6	4.4	25.0
	Brunel, London, UK	Count	0	1	1	3	4	9
		Expected	1.0	1.2	1.4	3.8	1.6	9.0
	Nasrda office, Abuja	Count	0	1	4	5	1	11
		Expected	1.2	1.5	1.7	4.6	2.0	11.0
Total	Count	5	6	7	19	8	45	
	Expected	5.0	6.0	7.0	19.0	8.0	45.0	
	Count							

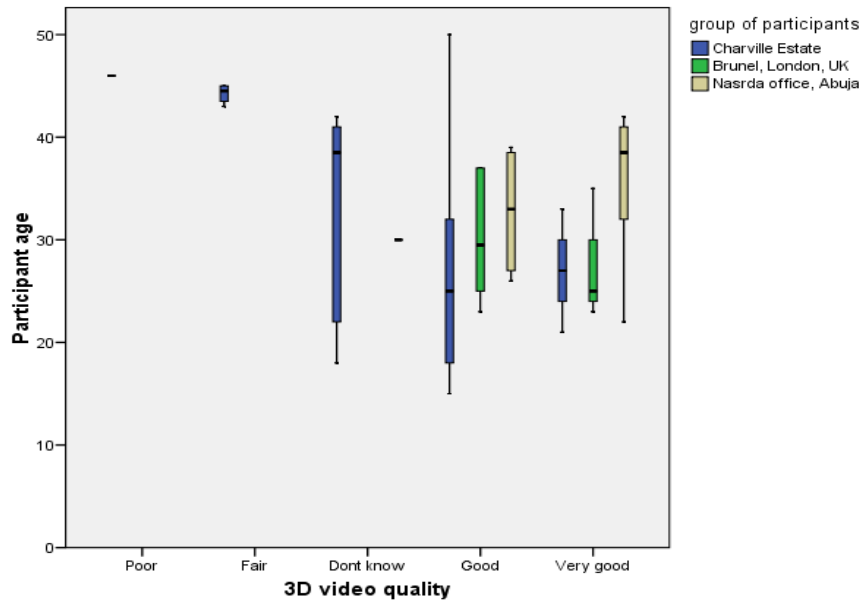


Figure 5-4: 3D Depth Perception, feel presence and overall 3D video quality.

Quality Assessment of 3D Videos

Table 5-7: Mean, Standard Deviation, Kurtosis and Skewness computed

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
level of satisfaction	45	55	30	85	68.00	2.093	14.037	197.045	-.904	.354	-.124	.695
Gender	45	1	1	2	1.31	.070	.468	.219	.844	.354	-1.349	.695
Valid N (listwise)	45											

Table 5-8: Overall User's needs, rqrs, expectation and new technology

overall users view(needs, rqrs, expec, new tech)		Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
level of satisfaction	Not important at all	.260	2	.			
	Less important	.146	6	.200(*)	.988	6	.985
	Dont know	.195	10	.200(*)	.935	10	.502
	Important	.298	16	.000	.719	16	.000
	Very important	.335	11	.001	.733	11	.001

df:degree of freedom

Sig: Significance level

Quality Assessment of 3D Videos

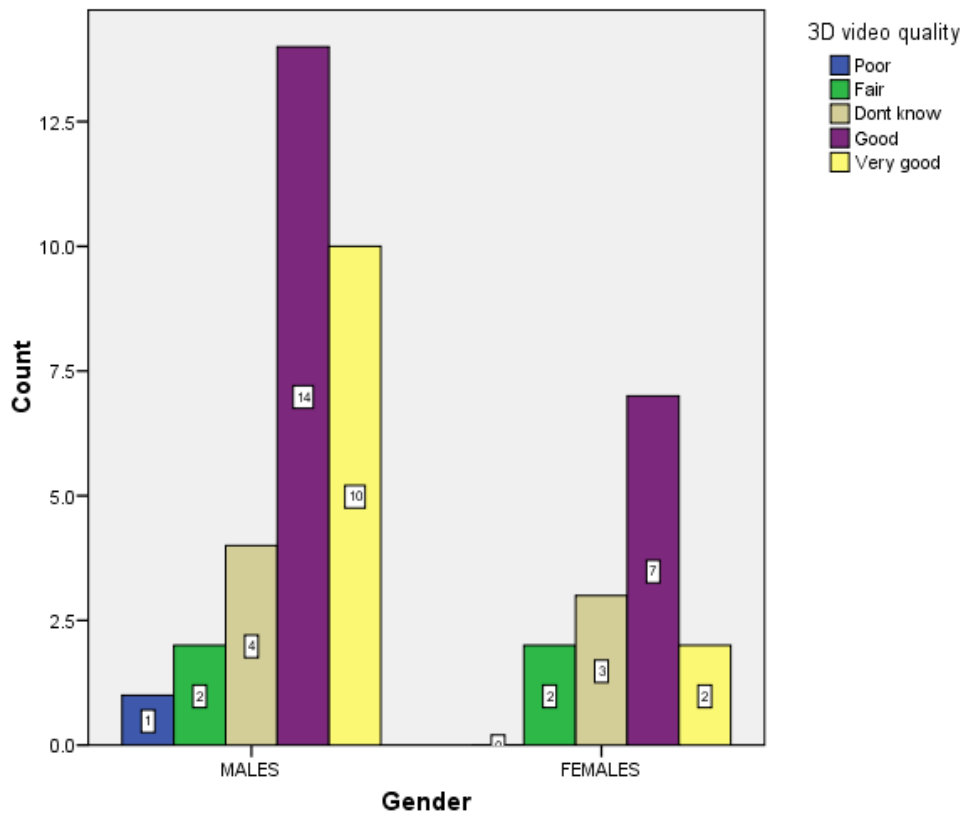


Figure 5-5 (a): MOS of 3D Video Quality in bar chart.

5.5.1 Tests of Normality

Figure 5-6 shows the test for normality of the total MOS. Figure 5-7 shows the MOS represented in box plot, the scores returned for important and very important are higher across the three experiments conducted as compared to “don’t know”, “less important” and “not important at all”.

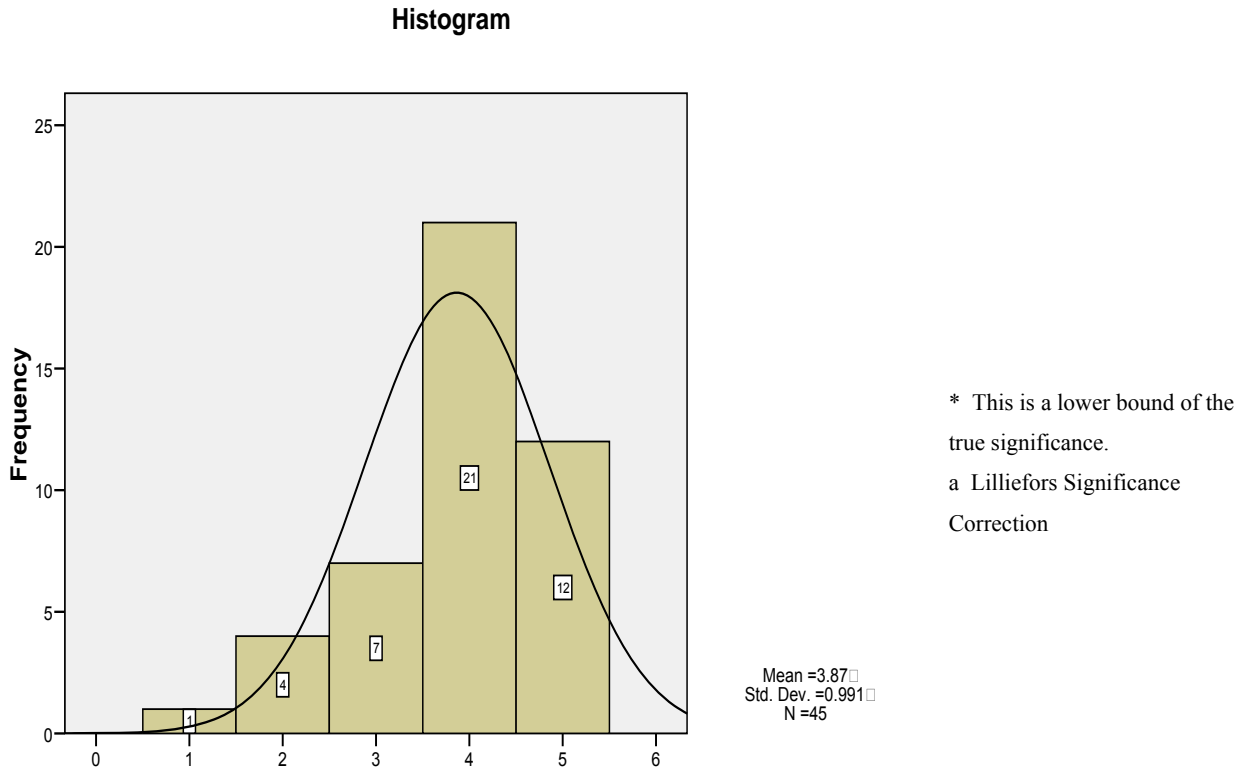


Figure 5-6: Test for Normality of the MOS

In statistical analysis, a box plot is a graph that can be a valuable source of easy-to-interpret information about a sample of study. It gives information about our sample range, median, normality and skew of our distribution. In figure 5-7 below, the box plot shows a box encased by two outer lines known as whiskers. The box represents the middle 50% of our sample, the remaining 50% of the sample is contained within the areas between the box and the whiskers with some exceptions called outliers. The bottom 25% of our MOS is represented by the space between the lower whiskers and the box, the middle 50% is within the box, and the remaining 25% contained between the box and the upper whisker. The location of the box within the whiskers gives us insight on the normality of the MOS of our distribution. The MOS may be positively or negatively skewed. For both male and female participants, in our survey the results show that the box is shifted significantly to the high end, it is negatively skewed.

Quality Assessment of 3D Videos

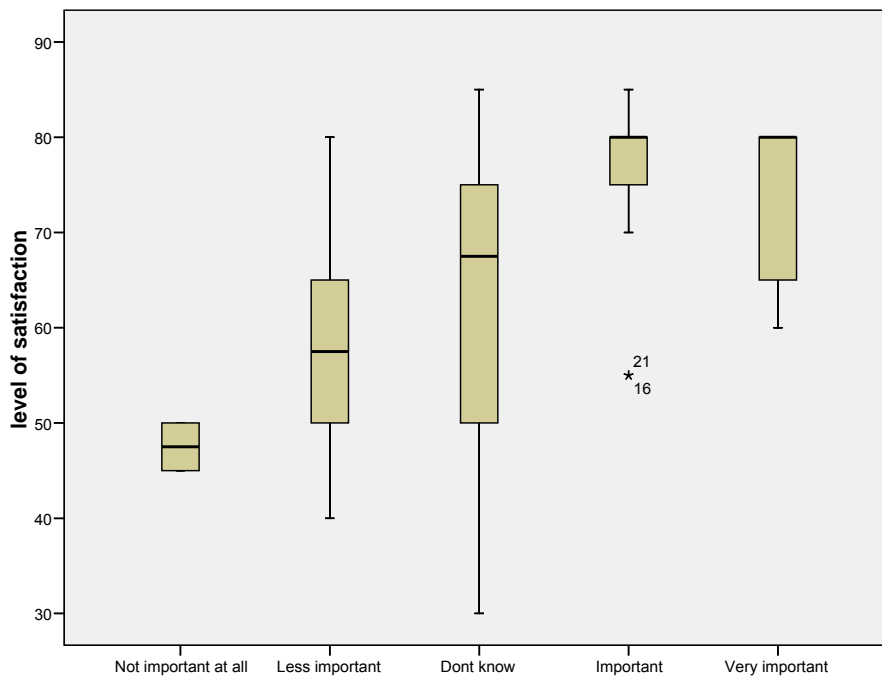


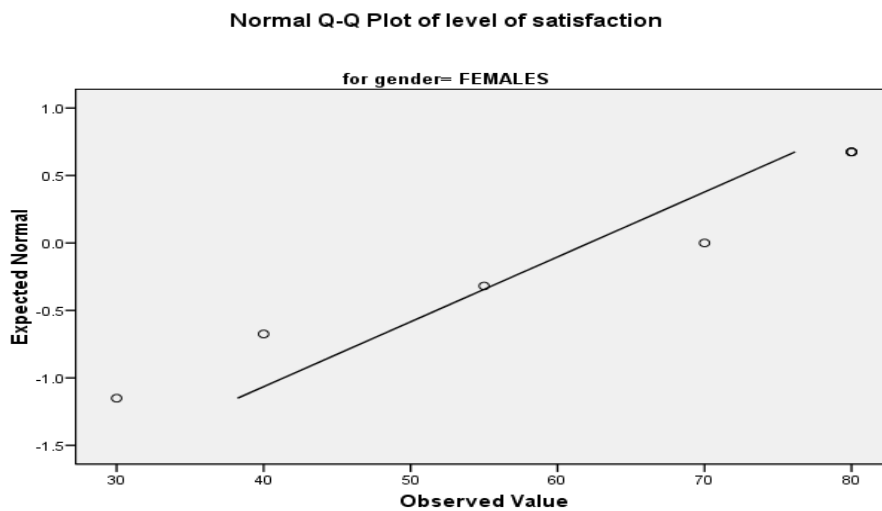
Figure 5-7: MOS values and 95% Confidence Interval using Boxplot.

The normal Q-Q plot graphically, compares the distribution of the expected MOS and observed MOS to the normal distribution. The straight line represents what our data would look like if it were perfectly normally distributed. Our actual observed data is represented by the dots circle plotted along this line. The closer the dot circles are to the line, the more normally distributed our data looks. Here in figure 5-8 and figure 5-9, most of our points fall almost perfectly along the line. This is a good indicator that our MOS from both male and female participants is normally distributed.

Quality Assessment of 3D Videos



(a)



(b)

Figure 5-8: MOS values for 3D video level of satisfaction; a) Males, b) Females.

Quality Assessment of 3D Videos

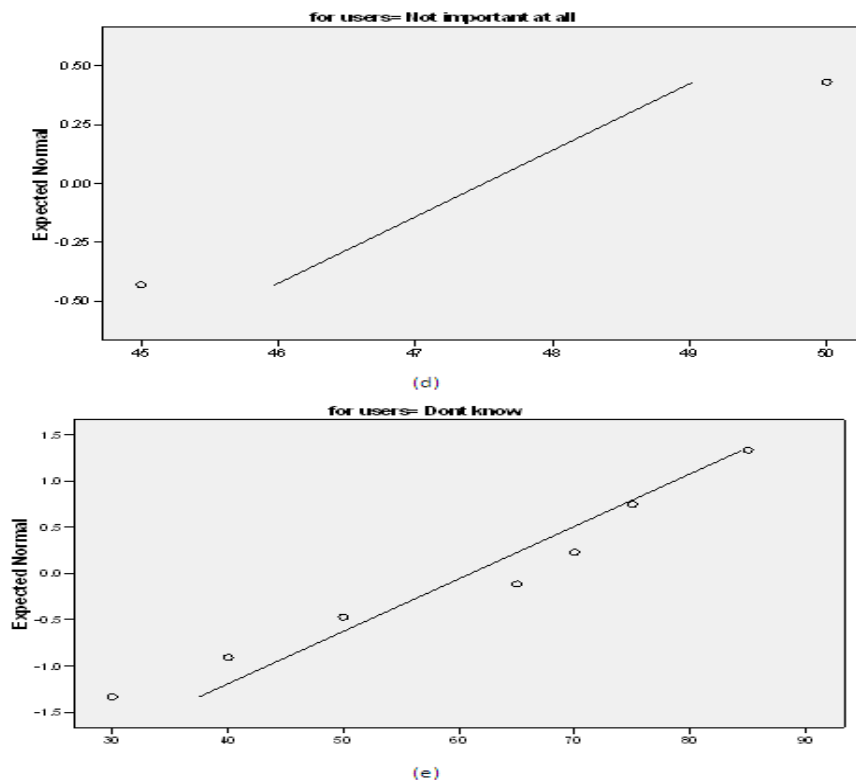


Figure 5-9: MOS values and 95% confidence interval obtained.

5.6 Analysis of Variance for MOS validation

The experiment for validation was conducted based on ITU-R compliant. The strategy for the subjective evaluation was also based on ITU-R recommendation. In data analysis, one-way analysis of variance is a technique used to compare means of two or more samples using F distribution as shown in figure 5-10. The ANOVA tests the null hypothesis between the 3 study groups. To do this, estimates were made of the population variance. Some assumptions were also made. The ANOVA produces an F statistic, the ratio of the variance calculated among the means to the variance within the samples [31]. Typically, however, the ANOVA was used to test for correlation among the 3 study groups, for 2 groups we simply used student t-distribution test.

The results of a one-way ANOVA can be considered reliable and acceptable as long as the following assumptions are met:

Quality Assessment of 3D Videos

- Response variable from participants/population must be normally distributed or approximately normally distributed.
- Samples are independent
- Responses for a given group are independent and identically distributed normal random variables.

In this research work, we used one-way Analysis of variance (one way-ANOVA) to test the validity of MOS between the 3 study groups to find out whether or not there is a significant difference in the MOS between the 3 groups. We observe each variable from different participants. Each participant was only used once for each test/experiment, no participant participated twice either in group 1 and 2, or group 1 and 3 or group 2 and 3. The survey was conducted once for each member of a particular group. MOS was calculated by adding up responses to several different survey questions as shown in Table 5-2 and their responses in Table 5-3.

We identified the hypothesis test and verify some of the following assumptions:

1. The populations which we sampled from are approximately normal
2. The MOS distribution for each group are approximately normal
3. The population which we sampled from have equal variance
4. The variance of the MOS for each group is approximately equal

The assumptions are tested when we run the one-way ANOVA test. The claim for this test is:

There is no significant difference between the MOS of the 3 groups (Charville, CMCR & Abuja, Nigeria).

H_0 : No significant difference between the MOS across the 3 groups as shown in figure 5-10.

H_1 : There is a significant difference between the MOS across the 3 groups as shown in figure 5-10.

We now choose a confidence level (typically 95% or 99%). We determine the level of significance α , where $\alpha = 1 - CI$.

Quality Assessment of 3D Videos

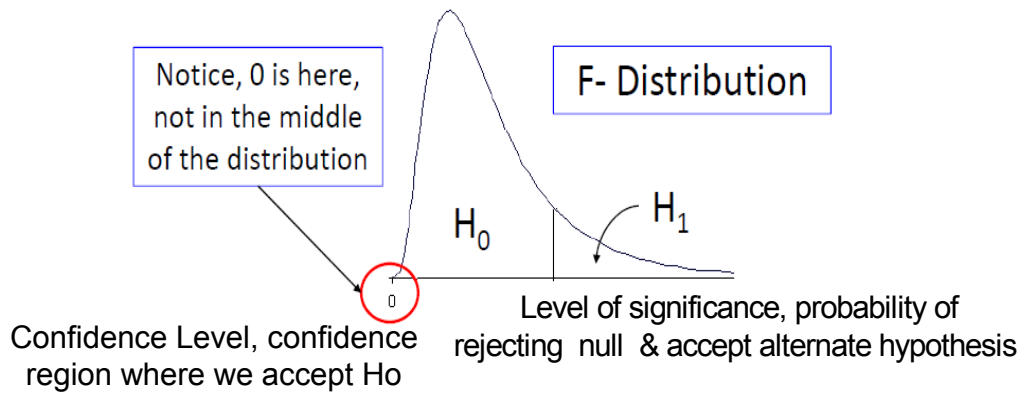


Figure 5-10: F-distribution test for validation of raw MOS [31][32].

Therefore $\alpha = 5\%$ or 0.05 with 95% CI

We now relate it back to the original problem statement, interpret the results and state the final conclusion. The MOS difference is significant at the 0.05 level. Do we accept H_0 ? or reject H_0 in favour of H_1 ? At the end of the experiment, we will be able to answer this question. Table 5-9 showed the test for hypothesis, Table 5-10 showed the experimental results for validation and Table 5-11 showed how the results are interpreted.

Table 5-9: Test for hypothesis

Test for hypothesis	P –value vs α
Accept H_0	If P –value $> \alpha$
Reject H_0 in favour of H_1	If P –value $< \alpha$

Quality Assessment of 3D Videos

Table 5-10: Experimental result for validation

		Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.	Monte Carlo Sig.		
		Sig.	99% Confidence Interval	Sig.	99% Confidence Interval		Sig.	
		Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound
Ordinal by Ordinal	Kendall's tau-b	.135	.115	1.170	.242	.302(c)	.290	.314
	Spearman Correlation	.161	.137	1.071	.290(d)	.282(c)	.270	.294
Interval by Interval	Pearson's R	.191	.118	1.274	.209(d)	.225(c)	.215	.236
N of Valid Cases		45						

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

c Based on 10000 sampled tables with starting seed 2000000.

d Based on normal approximation.

Table 5-11: Interpretation of results

Interpretation of results	P –value vs α
Accept H_0	P –value(0.270 > $\alpha(0.05)$)
Reject H_1 in favour of H_0	P –value(0.215 > $\alpha(0.05)$)

The MOS between the 3 groups of studies was investigated. Since $P - value(0.270 > \alpha(0.05))$ for Spearman correlation test and $P - value(0.215) > \alpha(0.05)$ for Pearson's Rank correlation, this suggests that there is no significant difference between the MOS at 5% level, we therefore accept the null hypothesis. H_0 states that there is no significant difference between the MOS across the 3 study groups and reject H_1 which indicates that there is a significant difference between the MOS. This means that people who are the consumers of 3D videos, irrespective of their background should be able to judge whether the quality of the 3D videos is good or bad and will determine how rapidly 3D video market grows.

5.7 Analysis & Discussion

The most reliable way of assessing the quality of a coded image or video is subjective evaluation, because human beings are the final receivers in most applications. Since it is the observer's opinion of video quality that counts, the subjective tests were conducted in this thesis.

In this chapter, three groups of forty five participants took part in the experiments and all completed the questionnaire at the end of the experiments. The three groups are from Charville Estate in Hayes, CMCR Lab in Brunel University and NASRDA office in Abuja.

The MOS obtained from the 3 groups shows good correlation among individual scores when spearman correlation coefficient and Pearson Rank correlation coefficient were analyzed (See Table 5-11). The Monte Carlo with 95% CI shows that $p > \alpha$, $0.282 > 0.05$ and also $0.225 > 0.05$ (level of significance or 5%). This shows there is no significant difference in the MOS recorded from the three groups, and we accepted the null hypothesis and rejected the alternate hypothesis since both p values $> \alpha$. It is clear that the proposed approach performs well in terms of correlation with subjective human perception of 3D videos.

5.7.1 Limitations of the survey

There are a number of disadvantages of conducting a questionnaire survey. Self-administered questionnaires are frequently criticized because of low response rates [33].

When conducting the survey, the researcher cannot guarantee that all the respondents understand the questions clearly. In addition, it is hard to identify which participants might misunderstood the questions when analysing the results. During the analysis, it was found that a few of the responses were not directly relevant to what was being asked, which might be the result of misunderstandings; however, most of the time the questions worked well.

It is hard to identify if the respondents treat the questionnaire seriously [34]; however, from the answers given to the questions, it was concluded that most of the respondents sincerely tried to share their observations.

Two groups were frequently compared during the analysis of the questions: (1) those with prior experience of 3D videos and (2) those with no experience. However, the numbers of the respondents were not equally distributed for the two groups. There were fewer respondents who

have prior experience of 3D videos or films. This should be borne in mind when interpreting the results statistically.

5.7.2 Other Findings of the survey

The results of the survey showed that the majority of the respondents who took part in the survey suggests that there is a difference in terms of the approach to the design adopted by the researcher as compared to everyday survey tools such as the survey monkey which does not provide detailed statistical analysis [35]. All respondents appreciate the method adopted by the researcher.

The results also suggest that both the participants who had prior experience and those without prior experience responded similarly.

5.8 Relationship between Objective and Subjective Test Measure

The thesis composed of two evaluation of test video material but objective and subjective as demonstrated in chapter 4 and 5. One aspect of test in chapter 4 is the objective test and the results are plotted while, chapter 5 dealt with subjective evaluation, which is by human observers, subjective result are presented in chapter 4 under the subjective test while the experiment and data analysis was performed in chapter 5. The objective models are meant to predict the subjective judgements. Each test spans a wide range of quality, so that the evaluation criteria are able to determine statistical differences in objective and subjective test performance. The results of the tests are given in terms of MOS which is a quantitative measure of the subjective quality of a video sequence as judged by human observers. The test had a wider coverage of typical video content such as spatial details, motion complexity and colour, and typical video processing conditions to assess the ability of models to perform reliably over a UMTS network discussed in chapter 2.

We conducted three independent tests on the subjective evaluation. Two were conducted in UK, while the third was conducted in Nigeria. In parallel, we produced the objective test measurements of the video quality of the same video sequences tested with human observers, using PSNR which is simple compared to SSIM which is very complex (See chapter 4 for PSNR preferences to other objective performance measurement in section 5.1). The results of the three tests are similar but not identical. We computed the correlation coefficients in this chapter. By

Quality Assessment of 3D Videos

this criterion, all the test measurements on the videos both objective and subjective perform equally well.

The Pearson correlation coefficients range from 0.294 to 0.270 and Spearman correlation coefficients range from 0.236 to 0.215.

PSNR was calculated (See plots in chapter 4 and appendix G), the results were analysed. For all the data sets, the PSNR fit significantly worse than the subjective test results of the same video sequences.

Also processed video sequences using PSNR contained no information relative to normalization i.e., no correction for gain and level offset, spatial shifts, or temporal shifts. In other words, unlike the subjective test, the video sequence files did not contain any alignment patterns to facilitate the normalization operation. Therefore, we used the subjective test measurements in chapter 5 for the normality test. The output of the objective assessment test should correlated with the participant's viewer DMOS in a predictable manner. The objective test and DMOS need not be linear as subjective testing can have nonlinear quality rating. To remove any nonlinearity due to the subjective viewer rating process and to facilitate comparison of the two models in a common analysis space, the relationship between each prediction and the subjective ratings was estimated using a nonlinear regression between the objective and subjective test (i.e regression between the PSNR and the DMOS).

The nonlinear regression was fitted to the DMOS and PSNR data set and restricted to be monotonic over the range of PSNR. The following function was used [2].

$$DMOS_p = b_1 / (1 + \exp(-2b * (PSNR - b_3))) \quad (5-3)$$

Fitted to the data [DMOS, PSNR]

The nonlinear regression function was used to transform the set of PSNR values to a set of predicted MOS values, $DMOS_p$, which were then compared with the actual DMOS values from the subjective tests. Once the nonlinear transformation was applied, the objective test prediction performance was then evaluated by computing various metrics on the actual sets of subjectively measured DMOS and the predicted $DMOS_p$. The test mandates some objective metrics of the correspondence between the PSNR and the subjective data (DMOS). In addition, it requires the

Quality Assessment of 3D Videos

checks of the quality of the subjective data. The test does not include the statistical test of the difference between different PSNR's fit to DMOS.

5.8.1 Metric related to prediction accuracy

The Pearson linear correlation coefficient.

5.8.2 Metric related to prediction Monotonicity

Spearman rank order correlation coefficient.

PSNR performance was assessed by correlating subjective scores and the corresponding PSNR predicted scores after the subjective data were averaged over subjects.

The Spearman correlation and Pearson correlation and all other statistics were calculated across all the group studies. In particular, these correlations were not calculated separately for individual group.

5.8.3 Metrics related to prediction Consistency

Outlier Ratio of "Outlier points" to total points N.

$$\text{Outlier Ratio} = \text{total number of outliers}/N \quad (5-4)$$

Where an outlier is a point for which:

$$\text{ABS}[Q_{\text{error}}[i]] > 2 * \text{DMOSStandardError}[i] \quad (5-5)$$

Twice the DMOS Standard Error was used as the threshold for defining an outlier point

5.8.4 Objective and Subjective comparison

Below is the model program for computing the results

```
* NonLinear Regression.
MODEL PROGRAM PSNR vs DMOS .
COMPUTE PRED_ = videoqual.
NLR Performance comparison between the two models
  /OUTFILE= 'C:\DOCUME~1\eepegasu\LOCALS~1\Temp\spss2580\SPSSFNLR.TMP'
  /PRED PRED_
  /CRITERIA SCONVERGENCE 1E-8 PCON 1E-8 .
```

Quality Assessment of 3D Videos

Table 5-12: Summary of analysis

Parameters	Group 1	Group 2	Group 3	Total
PSNR	0.804	0.811	0.127	0.3125
DMOS	0.2	0.4	0.6	0.8
Outlier Ratio	30/45=0.6 66	33/45=0.7 33	34/45=0.75 55.	38/45=0.8 44.
RMS error	0.139	0.075	0.11	0.117
Pearson Correlation	0.759	0.937	0.857	0.835
Spearman Correlation	0.759	0.934	0.875	0.814
MSE null hypothesis	0.01790	0.01790	0.01790	0.01790

Table 5-13: Continued Analysis

Iteration Number(a)	Residual Sum of Squares	Parameter					
		DMOS	Outlier Ratio	RMS error	Pearson Correlation	Spearman Correlation	PSNR
20	353.000	0.2-1	0.74	0.880	45.000	0.759	0.767

Derivatives are calculated numerically.

a Major iteration number is displayed to the left of the decimal, and minor iteration number is to the right of the decimal.

b Run stopped after 1 model evaluations and 1 derivative evaluations because the relative reduction between successive parameter estimates is at most PCON = 1.00E-008.

ANOVA(a,b)

Table 5-14: ANOVA Test Results

Source	Sum of Squares	df	Mean Squares
Regression	-266.000	6	-44.333
Residual	353.000	39	9.051
Uncorrected Total	87.000	45	
Corrected Total	9.644	44	

Dependent variable: Gender

a The solution results in negative sum of squares. Check model specification.

b R squared = 1 - (Residual Sum of Squares) / (Corrected Sum of Squares) =

These results show that the correlations of the PSNR measures are lower than the lowest subjective quality measure for all the iterations. Figures 5-11 and 5-12 show the plot for the DMOS vs PSNR using equation 5-3 above with SPSS software tool [31][2].

Quality Assessment of 3D Videos

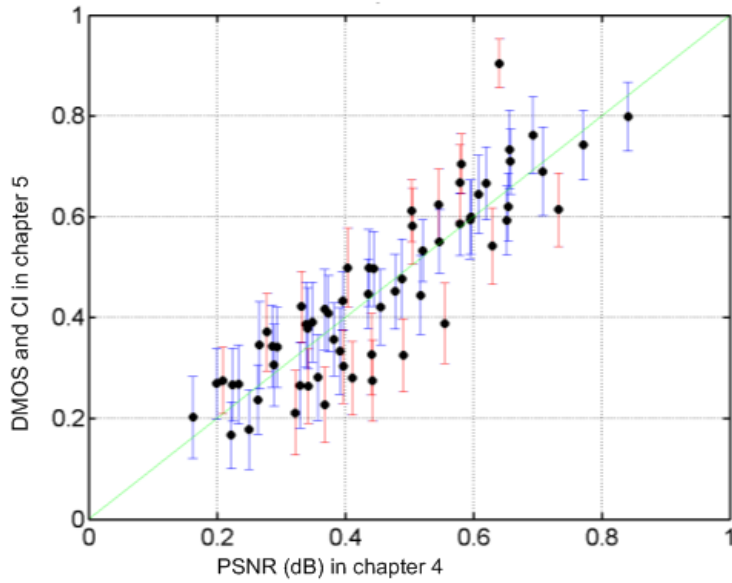


Figure 5-11: Subjective vs Objective Performance Measure

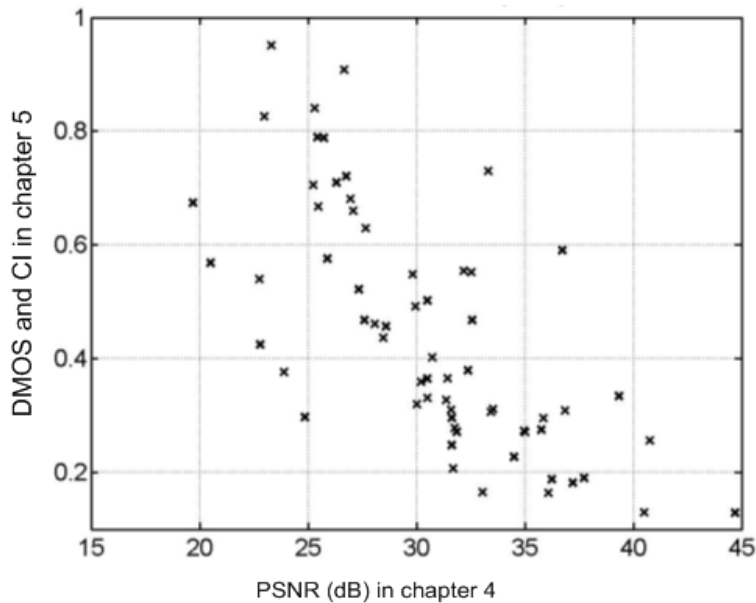


Figure 5-12: DMOS vs PSNR (dB)

We conclude that the fit metrics appear to show differences among the models. Which of the differences are statistically significant? A test for differences between correlation coefficients was carried out. The sensitivity analysis of this test statistics depends on the size of the sample of

Quality Assessment of 3D Videos

observations or participants, N- which is true of many statistics. For two correlations, both based on 45 subjects, the test for the difference is

$$\text{Sigma} = \text{SQRT} (1/45 + 1/45) = 0.21.$$

Experiment repeated 20 times in chapter 4 therefore, sigma is $\text{SQRT} (1/20 + 1/20) = 0.316$.

Differences of two sigmas are taken as significant. Thus correlation in Table 5-12 must differ by a very large amounts to be considered significant. Another approach to testing significance of differences uses the idea of an optimal model and the F-tests used in analysis of variance [2]. An optimal model would predict each of the DMOS values for the 45 subjects. The residual differences of individual subjects ratings from the 45 scores cannot be predicted by any objective model, an objective model usually makes prediction for each subjective testing using PSNR. This residual is the baseline against which any objective performance is tested. The optimal model is also a “null” hypothesis. The null hypothesis achieves optimal fit to the subjective data by not doing any prediction at all. The mean rating for the particular stimulus is what the null hypothesis predicts. When an objective performance is tested against the individual subjective responses, a residual variance is obtained. When the “null hypothesis”: Response = Stimulus is computed, the residual variance is calculated around the mean DMOS for each stimulus. Here, the stimulus is just an identifier with one degree of freedom (df) for each subjective testing. The residual for the null hypothesis is the baseline minimal residual. It is given in Table 5-12, line 7. The ratio of these two residual variances is an F-test. In the F-distribution statistics, values of F smaller than about 1.07 indicates that a performance metric is not statistically different from the null hypothesis or optimal model.

The reason the F-test is able to discriminate between PSNR and subjective performance better than when we compared correlation coefficients is that the F-test directly makes use of the number of stimulus as well as the number of subjects, while the correlation sensitivity test depends only on the number of subjects.

5.9 Conclusion

In this chapter, we conducted a subjective quality assessment to evaluate human perception and quality of experience (QoE) for 3D videos. We performed the assessment on publicly available datasets for 3D quality. In this chapter, we provided information sources frequently used when designing a questionnaire, and the expectations from the users. Based on the requirements and expectations of 3-D users, we developed a questionnaire. The work addressed 3-D quality assessment performance using the subjective test. Subjective results are analyzed to demonstrate the most effective metrics for determining the QoE of compressed 3-D colour and depth. The results show that the output of the 2D-colour plus depth correlate strongly with overall viewer perception of image quality and depth perception and quality of experience (QoE). The row mean opinion scores (MOS) from the experiment outperform objective tests using PSNR and root mean square error (RMSE) found in the literature [1][2][23]. Assessment of the depth alone provides no useful information on viewer ratings of image/video quality. The statistical features seem to perform well in terms of correlation with human perception across contents. The histogram, mean, median and standard deviation of the scores correlate well with the MOS.

The ANOVA model described in this chapter was validated using the correlation coefficient. It was observed that there is a strong correlation among the studies.

Also, the degree of compression during the video encoding expressed as the Quantisation Parameter (QP) has great impact on the MOS. Both MDC-SIMI and H.264/AVC show a similar trend, the assessed quality level decreases when the encoding QP was increased.

We also use curve fitting, correlations, Sensitivity analysis and F-distribution test to compare the objective performances in chapter 4 with the subjective performances in chapter 5 which meets the assumption of normality and the F-distribution which assumes that the residuals comes from a normal Gaussian distribution and conclude that subjective test might gives better quality measures as compared to PSNR. Sensitivity analysis test conducted also confirms the superiority of subjective test over the objective test (PSNR).

References

- [1] Hewage, C. T, et al, “Quality Evaluation of Colour Plus Depth Map-Based Stereoscopic Video”, IEEE Journal of Selected Topics in Signal Processing, vol. 3, issue 2, pp. 304-318. April 2009.
- [2] Video Quality Expert Group (VQEG), “ Final Report From the Video Quality Expert Group on the Validation of Objective Models of Video Quality Assessment, Phase II”, 2003. Available: <http://www.vqeg.org>.
- [3] S. Khalil, Yonis, “Objective Video Quality Assessment using the 3D Dual-Tree Complex Wavelet Transform”, MSc Thesis, Department of Systems and Computer Engineering, Carleton University, 2009.
- [4] F. Okuyama, “Evaluation of stereoscopic display with visual function and interview”, In proceedings of the SPIE 3639, pp. 28-35, 1999.
- [5] P. Seuntiens, “Visual experience of 3D TV”, PhD. Thesis, Eindhoven University, Eindhoven, The Netherland, 2006.
- [6] H. Kalva, L. Christodoulou, and B. Furht, “Evaluation of 3D TV service using asymmetric view coding based on MPEG-2”, in Proceedings of 3DTV-Conference, Kos Island, pp. 1-4, Greece, May 2007.
- [7] A. P. Mapp, H. Ono, and R. Berbeito, “What does dominant eye dominate? A brief and somewhat contentious review”, Perception and Psychophysics, vol. 65, no. 2, pp. 310-317, 2003.
- [8] E. Shneur and S. Hoshstein, “Effects of eye dominance in visual perception”, In Proceedings of the Vision International Congress, vol. 1282 of International Congress Series, pp. 719-723, London, UK, April 2005.
- [9] N. Holliman, B Froner, and S. Liversedge, “An application driven comparison of depth perception or desktop 3D displays”, in Stereoscopic Displays and Virtual Reality Systems XIV, vol. 6490 of proceedings of SPIE, pp. 1-22, San Jose, California, USA, January 2007.

Quality Assessment of 3D Videos

- [10] A. Berthold, “The influence of blur on the perceived quality and sensation of depth of 2D and stereo images”, Technical. Report, ATR Human Information Processing Research Laboratories”, 1997.
- [11] P. Seuntings, L. Meesters, and W. IJsselsteijn, “Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation”, *ACM Trans. Appl. Perception (TAP)*, vol. 3, pp. 95-109, 2006.
- [12] H. A. Karim, A. H. Sadka, and A. M. Kondo, “Reduced Resolution Depth Compression for 3D video over wireless networks”, In *Proc. of IET 2nd Int. Workshop Signal Process. Wireless Comm*, London, UK, pp. 96-99, June 2004.
- [13] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, “Live image quality assessment database release 2”, Available online at: <http://live.ece.utexas.edu/research/quality/subjective.htm>. Accessed on 4th Feb. 2011.
- [14] F. D. Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, and T. Ebrahimi, “Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel”, in *First International Workshop on Quality of Multimedia Experience*, San Diego, CA, USA, July 2009.
- [15] Y. Yang, R. Blake, “Spatial frequency tuning of human stereopsis”, *Vision Research*, vol. 31, no. 7-8, pp. 1176-1189, 1991.
- [16] Y. Liu, L. K. Cormack, A. C. Bovik, “Natural scene statistics at stereo fixation”, In *Proc. of the 2010 Symposium on Eye-Tracking Research & Applications*, ACM, pp. 161-164, 2010.
- [17] K. Seshadrinathan, A. C. Bovik, “Motion-based perceptual quality assessment of Video”, *Pro. SPIE-Human Vision and Electronic Imaging*, 2009.
- [18] S. E. Palmer, “*Vision Science: Photons to phenomenology*”, MIT press Cambridge, MA. 1999.
- [19] A. Mittal, K. Moorthy, J. Ghosh and A. Bovic, “Algorithmic Assessment of 3D Quality of Experience for Images and Videos”, *Proceedings of the 14th IEEE*

Quality Assessment of 3D Videos

Digital Signal Processing & Signal Processing Education Workshop, Enchantment Resort, Sedona, Arizona, USA. Jan 4-7, 2011.

- [20] S. Jumisko, M. Weitzel, D. Strohmeier, “Designing for User Experience: What to Expect from Mobile 3D TV and Video?”, Proceeding of the 1st international conference on Designing interactive user experiences for TV and video, New York, NY, USA, 2008.
- [21] I. Aljoscha, “3D Innovation Centre, Berlin: http://www.hhi.fraunhofer.de/en/press/press-and_media/the-3d-innovation-centre-berlin-platform-and-launch-pad-for-the-market/ . Accessed on 1 Sept. 2011
- [22] Visual Media Research Group of Microsoft: <http://research.microsoft.com/en-us/groups/interactivevisualmedia/> . Accessed on 1st Sept.2011.
- [23] M. Zwicker, S. Yea, A. Vetro, X. Forlines, W. Matusik, H. Pfister, “Display pre-filtering for multiview video compression”, In Proc. of 15th International Conference on Multimedia, pp.1046-1053, Augsburg, Germany, Sept. 2007.
- [24] K. Ukai, P. A. Howarth, “Visual fatigue caused by viewing stereoscopic motion images: Background theories, and observation”, Displays, vol. 29, no.2, pp. 106-116, 2008.
- [25] M. Emoto, T. Niida, F. Okano, “ Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television”, Journal of Display Technology, vol. 1, no. 22, pp. 328-340, 2005.
- [26] K. Ukai, “Human factors for stereoscopic images”, Proc. Of the IEEE International Conference on Multimedia and Expo (ICME’06), pp. 1697-1700, Toronto, Canada. July 2006.
- [27] S. Tsukuda, Y. Murai, “A case of manifest esotropia after viewing anaglyph stereoscopic movie”, Japanese Orthopedic Journal, vol. 18, pp. 69-72, 1988.
- [28] ITU-R BT.500-11, Methodology for the Subjective Assessment of the Quality of Television Pictures.
- [29] F. D. Simone, M. Tagliasacchi, M. Naccir, S. Tubaro, T. Ebrahim, “A H.264/AVC Video Database for the Evaluation of Quality Metrics”, 35th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), Dallas, Texas, U.S.A, March 14–19, 2010.

Quality Assessment of 3D Videos

- [30] A. Benoit, P. Callet, P. Campisi, R. Cousseau, “Quality Assessment of Stereoscopic images”, EURASIP Journal on images and Video Processing, vol. 2008, Article ID 659024, 14th Oct. 2008.
- [31] C. Gregory, “One-Way between groups ANOVA in SPSS”, Spring ASK Week Training, Brunel Library, 2011.
- [32] F. Andy, “Discovering Statistics Using SPSS, 3rd Edition, SAGE Publications Ltd, London, 2009.
- [33] R. Sapsford, V. Jupp, “ Data Collection and Analysis”, London: Sage Publications, 2006.
- [34] N. Stanton, “ Product Design with People in Mind”, Human Factors in Consumer Products, pp. 1-17. London, 1998.
- [35] A. Vermeeren, P. Bouwmeester, J. Aasman, “ A Tool for Detailed Video Analysis of User Data” Behaviour & Information Technology Conference, pp. 403-423, vol. 21, 2002.

Chapter 6: Conclusions & Future Work

In this chapter, the whole thesis is summarized, and conclusions are drawn. Suggestions for future work are also presented. Error resilience and compression efficiency to transmission errors are the major factors affecting video quality in many communication systems, especially when streaming delivery is used. The thesis gave in depth overview of MDC as a tool for error resiliency and scalability of 3D video communication. Main motivation for using multiple description coding arises from the vulnerability of the compressed bitstream to channel errors when transmitting video data over error prone channels. The proposed 3D MDC-SIMI scheme greatly improves the error robustness of the encoder/decoder.

6.1 Critical Summary

We presented background information about wireless technologies, i.e., the evolution from 1st to 4th generation networks. The thesis examined the behaviour of UMTS networks and protocols and provided some basic simulations for UMTS using OPNET Modeler simulation tool. We achieved in establishing QoS at different layers such as services, applications and Network e.t.c. We developed a UMTS model through OPNET with specified/services requirements to provide QoS for video streams in wireless environment. We also achieved in developing a WLAN wireless network model in Brunel University to provide QoS for video transmission. We performed verification of the UMTS model to ensure that our simulation model was both correctly implemented and representative of the real system.

Model validation was carried out in order to determine if our simulation model is representative of the real system. A simulation can be validated using expert intuition, real system measurements or theoretical results. We compared our simulation and theoretical results. The simulation results performed better than the theoretical results as shown in figure 2-45. The similarity between the analytical and simulation results validates the correct operation of the data transfer for the UMTS access network. Limitations: no real system measurements available for prototype 3G wireless equipment through the course of my PhD in Brunel University.

In chapter 3, we reviewed and analysed the Advanced Video Coding as described in the standard documents. We also reviewed the different scalabilities such as Temporal Scalability, Spatial

Conclusions and Future Work

Scalability and SNR Scalability. We use a standard compliance scalable MDC based on odd and even frames was also presented for the H.264/SVC video coding targeting 3D video applications to obtain the scalable results and layering techniques. It generates two descriptions for the base layer of SVC based on even and odd frame separation.

We reviewed the 3D technology as retrieved from the literature, effect of human visual system on 3D perception was discussed and efficient methods for 3D content generation and display was presented.

Matlab scripts were used in simulation to combine the left and right images (Stereoscopic images). We also generated the depth and finally the 3D video using Matlab (See more results in appendix B).

In chapter 4, we designed 3D multiple description coding schemes for both error free and error prone environments. The designed algorithm (3D MDC-SIMI) is achieved by extending the work in [33] in order to combat channel-induced impairments across the transmission channel. We included some quality comparisons with other proposed 3D MDC and concluded that our scheme performs better by around 1 to 2 dB in average PSNR as discussed in chapter 4 and in some cases even more, for different sequences, different bit rates and under different test conditions. See all the objective and subjective results in chapter 4 and 5. Achievements from the new algorithms listed below:

- 3D MDC-SIMI is a good candidate for wireless channels and performs better under error prone conditions.
- 3D MDC-SIMI requires approximately less computation times than 3D MDC-SIPA for error prone conditions.
- We compared the performance of the two MDC schemes for error free and error prone channels. Our MDC-SIMI algorithm performs better than the MDC-SIPA especially under high bit error rates. For low BERs, both have comparable performance.

Limitation include high redundancy required in the side decoders

In chapter 5, we performed a subjective quality assessment to evaluate human perception, requirements, expectation, and quality of experience (QoE) for 3D videos from users. We carried out the assessments on publicly available datasets for 3D quality through a well authored questionnaire. The results showed that the output of the 2D-colour plus depth correlate strongly

Conclusions and Future Work

with overall viewer perception of image quality and depth perception and QoE. The statistical results showed correlation with human perception across contents. The histogram, mean, median and standard deviation of the scores correlate well with the mean opinion scores (MOS). We used ANOVA model to validate the MOS generated from the three study groups from UK and Nigeria. We finally, presented the relationship between results in chapter 4 and chapter 5 (Objective and subjective performances) using curve fitting technique and ANOVA. We found that objective and subjective test correlate well, but the subjective test performs better (See subjective results in chapter 4). Main limitation is that, the SSIM cannot be used for performance evaluation due to its complexity, this could have given better subjective results.

6.2 Conclusions

- The proposed 3D MDC SIMI scheme was validated through the simulation runs presented in chapter 4 and compared its performance with 3D MDC-SIPA for both error free and error prone conditions. However, verification of the 3D MDC SIMI have not been undertaken because of time limitation and technical complexity involved in the process.
- The significant performance improvement of the 3D MDC –SIMI is achieved at the cost of large redundancies created by the two side information streams.
- The 3D MDC SIMI encoder and decoder are more complex and this translates to cost.
- The 3D MDC-SIMI is more applicable in a client-server scenario for 3D video streaming service, but not for end-to-end real time 3D video services.

6.3 Future Work

3D video communication over wireless networks is an under explored area. This research specifically focused on error resilience MDC for 3D video transmission. However, there are remaining aspects to be explored, which could extend the research work in the future. The researcher had a number of ideas from the supervisor and other CMCR lab members. The researcher compiled all the suggestions/ideas in the future, and intends to pursue funding and professional help for further work in this area of 3D videos.

Some of these important ideas identified which offered opportunities for further investigation within ongoing future research are as follows:

Conclusions and Future Work

6.3.1 3D MDC-SIMI with 4 descriptions

Future research shall extend the study of 3D MDC-SIMI beyond the two descriptions (Odd & Even). Figure 6-1 shows scenario where four descriptions are sent to two clients through a wireless channel which is an issue for future research. This research focused only on 2 descriptions (Odd and Even).

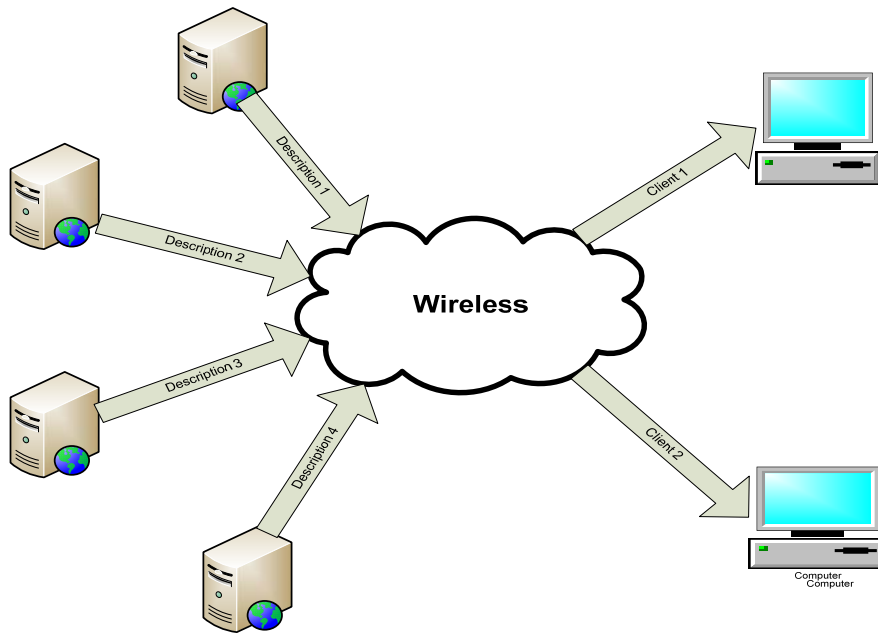


Figure 6-1: 3D Videos in MDC with four descriptions.

6.3.2 Improvement of the designed Algorithm

During the course of this research, a number of suggestions for improvement to the 3D MDC – SIMI algorithm were made. The knowledge and skills of the researcher in terms of algorithms development depends on the availability of resources and software which ultimately prolonged the work. The designed algorithm should be further developed and improved.

6.3.3 Understanding 3D User's Experience

Participant's views have been investigated in order to obtain an in-depth understanding of their requirements and expectation. However, what is the rationale for 3D user's decision when purchasing a product? Understanding this may help designers to better address user's expectations.

Conclusions and Future Work

Environmental challenges for 3D experience such as temperature, humidity, pressure and electromagnetic fields were not investigated in this thesis; however, future research in this area needs to bring these factors into consideration. There might be other factors affecting user's experience such as their lifestyle, cultural diversity, and emotional factors on users, capabilities and preferences which are also subject for future research. The survey conducted as part of this PhD research work also did not seek to understand the following:

- What 3D users think about the emerging trends in 3D technology?
- What is the role of users in this emerging trend?
- What sort of help do 3D users frequently seek from professionals about the 3D technology?
- What should designers consider when designing 3D?

Future work shall focus on the improvement of the design questionnaire to include the above stated points, and the development of an in-depth understanding of users without 3D experience and those with 3D experience on how they perceived 3D experiences.

6.3.4 3D Video Content Generation

The availability of a variety of 3D contents is one of the major requirements for the successful introduction of 3D TV to the consumer market. Although there are several studies regarding the guidelines for generation and recording of suitable 3D content for viewers [3][5], in practice 3D content generation is very challenging, expensive and also time consuming. More studies in the future are required about the impact of different camera setup parameters such as baseline distance, focal length etc. on the quality of perceived 3D content. The appropriate camera parameters should be defined for scenes with different depth ranges. The type and resolution of the display device (theatre, HD, active or passive glasses, etc.) should also be considered when choosing these parameters.

6.3.5 Scalable extension of H.264/AVC

Recently, the scalable extension of H.264, or MPEG-4 AVC, is under intensive research in Joint Video Team (JVT) for the next generation of video codec and applications [8]. Based on H.264, the so call Motion-Compensated Temporal Filtering (MCTF) provides an efficient and standard compatible scalable coding in both temporal and quality domain [9]. It could be a very

Conclusions and Future Work

interesting research topic to explore the possibility of using the Multiple Description Coding principle in the MCTF settings.

6.3.6 3D MDC –SIMI and channel interleaving

The two descriptions in 3D MDC-SIMI also provides possible channels for interleaving as suggested by Dr. Nichos in one of my CMCRC presentation [10]. Our work presented in chapter 4 does not include interleaving. The future research could explore the way to interleave information of block coefficients.

6.3.7 Motion Estimation Improvement

Motion estimation improvement may come from variable block size, as in H.264. More accurate block matching is expected in other to give a better motion estimation. Although this increase computation cost, in the future a properly selected threshold can be used to determine whether the current block search or the proper block size?.

6.3.8 UMTS channel behaviour

Study of channel behaviour shall give more insight to the design of the video codec. Commonly used two-state Markov models can only describe ideal channels; other models that are more accurate and practical shall be helpful in analysis of the transport system and therefore suggest possible improvements of our coding algorithms.

6.3.9 3D Video Quality Metrics

Another key factor for the successful penetration of 3D technology to the users/consumer market is to ensure that the new experience is superior to the one presently offered to the users. Assessing 3D quality is a huge challenge. It seems that the perceived user experience is psychological in nature and viewing 3D content introduces a new dimension of different environmental and display conditions. Therefore, new techniques are needed to assess this kind of experience. Currently the subjective test standards are not fully dedicated to the evaluation of 3D content quality. More study and research in the future are required for defining the quality of experience in 3D and for developing a 3D quality measure metric.

Future studies will consider other resolutions of the same video test sequences, as well as other subjective evaluation methodologies.

Conclusions and Future Work

References

- [1] L. Zhang, W. J. Tam, “Stereoscopic image generation based on depth images for 3D TV”, In IEEE Trans. Broadcasting, vol. 51, no.2, pp. 191-199, 2005.
- [2] P. Harman, J. Flack, S. Fox, M. Dowley, “Rapid 2D to 3D Conversion”, In Proceedings of SPIE, vol. 4660, pp.78-86, 2002.
- [3] S. H. Lai, C. W. Fu, S. Chang, “A generalized depth estimation algorithm with a single image”, PAMI, vol 14(4), pp. 405-411, 1992.
- [4] P. Merkle, K. Muller, A. Smolic, T. Wiegand, “Efficient Compression of Multiview video exploiting inter-view dependencies based on H.264/MPEG4-AVC”, In proc. ICIP, Canada. pp. 9-12, Jul. 2006.
- [5] A. Vetro, P. Pandit, H. Kimata, A. Smolic, “Joint Multiview Video Model (JMVM) 8.0”, ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-Technical Report AA207, Apr. 2008.
- [6] T. Weigand, G. J. Sullivan, G. Bjontegaard and A. Luthra, “Overview of the H.264/AVC Video Coding Standard”. IEEE Transaction on Circuits and System for Video Technology, vol. 13, no 7, pp. 645-656, July 2003.
- [7] M. Flierl, B. Girod, “Generalized B pictures and the draft H.264/AVC video compression standard”, IEEE Trans. Circuits Syst. Video Tech., Vol. 13, pp. 587-597, July 2003.
- [8] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. Scalable video coding-working draft 1. Document JVT-N020, Jan. 2005.
- [9] R. Schafer, H. Schwarz, D. Marpe, T. Schierl, T. Wiegand, “MCTF and scalability extension of H.264/AVC and its application to video transmission, storage, and surveillance”, In international Symposium on Visual Communication and Image Processing (VCIP), July 2005.
- [10] U. Abubakar, “Scalable 3D MDC coding based on Even and Odd Frames”, CMCR presentation, department of Electronic & Computer Engineering, Brunel University, March 2009.

Appendix A
Sample Questionnaire

What is your current status?

- College student
- Undergraduate
- Masters student
- PhD student
- Researcher
- Others (Specify)

Gender:

Age:

When was your first experience with 3D videos or 3D images?

Years/months/weeks

Section 1: Perceptual attributes associated with 3D perception:

	Extremely Important	Important	Less Important	Not Important
Video Quality	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Presence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comfort	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Perceived depth	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stereo impairments	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix

Viewing experience

Section 2: Problems associated with watching 3D videos:

	Extremely Important	Important	Less Important	Not Important
General discomfort	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Visual Fatigue/				
Visual discomfort	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Headache	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eye-strain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Simulator sickness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Difficulty focusing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Blurred vision	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 3: Human perception for 3D videos:

	Extremely Important	Important	Less Important	Not Important
Accommodation issues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Masking effect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interpupillary	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Age	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 4: How Important are the following effects to you?

	Extremely Important	Important	Less Important	Not Important
Blurring	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jerky/un-natural motion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix

Global noise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Block distortion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Colour distortion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 5: Impact of eye dominance effect in 3D perception:

High impact	Moderate impact	Less impact	No impact at all
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 6: Impact of depth information on the perceived 3D effect (enhance viewing experience):

High impact	Moderate impact	Less impact	No impact at all
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 7: How good is the quality of our 3D videos?

Extremely Good	Good	Fair	Poor
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 8: Please rate your level of satisfaction with our 3D videos:

Extremely Satisfied	Satisfied	Dissatisfied	Not satisfied
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 9: Overall users view of 3D videos:

	Extremely Important	Important	Less Important	Not Important
User's need	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
User's requirement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix

User's expectation

Enthusiasm for new technology

Section 10: Why did you choose 3D over 2D videos?:

QoE

Design

First user experience

Others (Specify please)

Section 11: Overall users QoE and QoS:

Excellent

Good

Fair

Poor

Bad

Section 12: What type of viewing glasses you prefer to use most?

Anaglyphic lenses

Polarized glasses

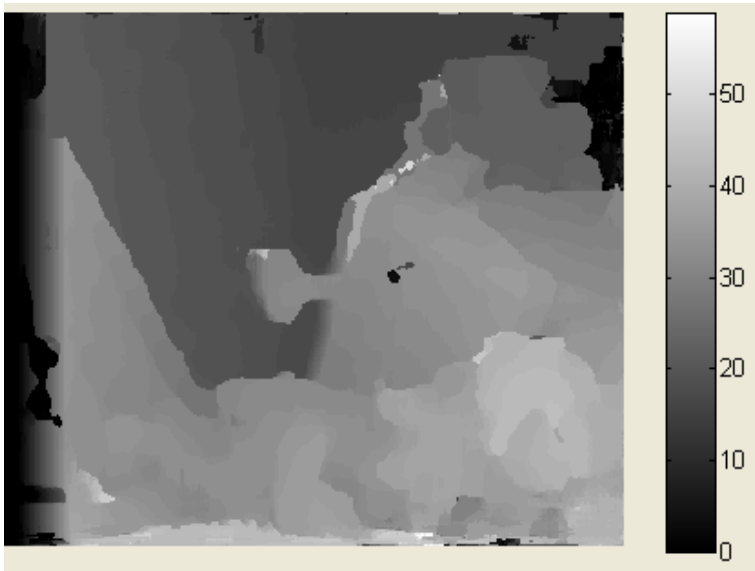
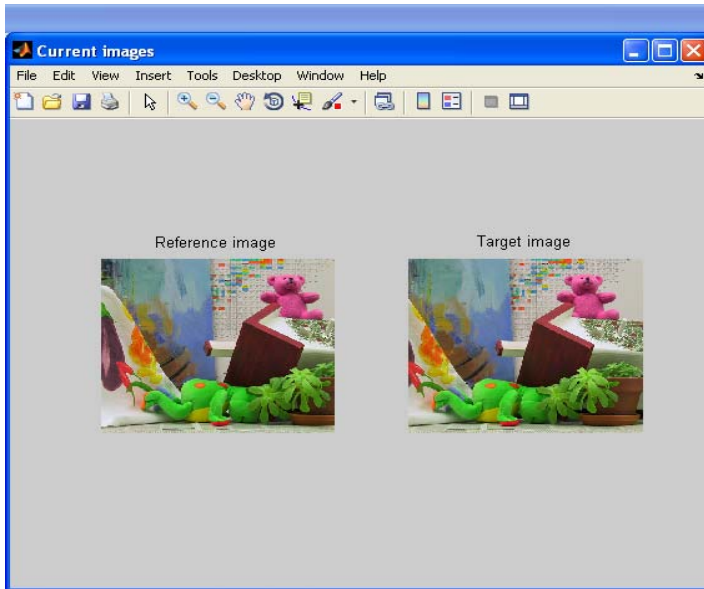
Shutter glasses

Others (if any)

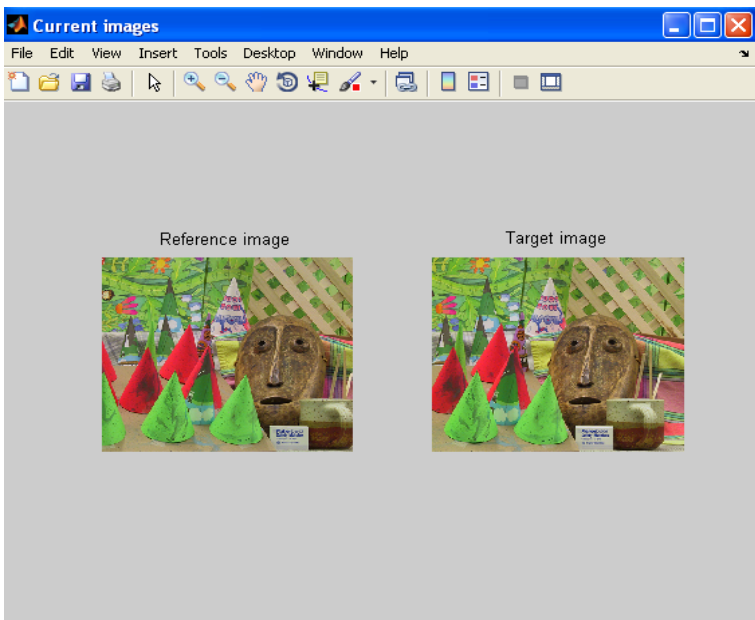
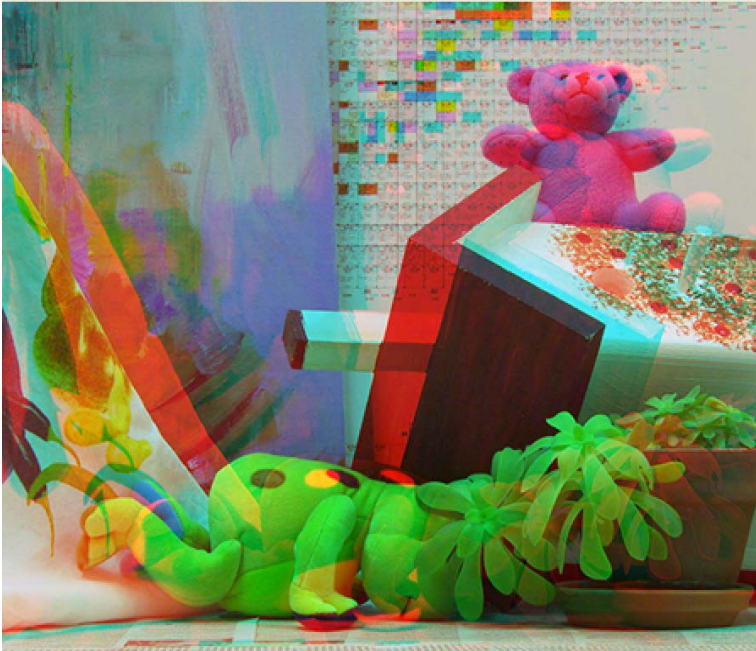
That's more or less it. Thanks for your time and participation. Hope you've enjoyed it.

Appendix

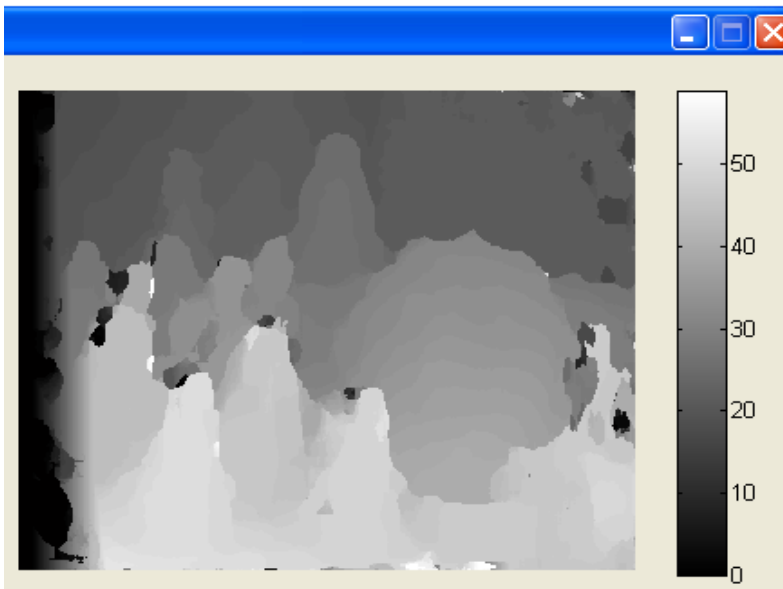
Appendix B: Methods of 3D Video Generation



Appendix



Appendix



Appendix

Appendix C: H.264/AVC Encoder & Decoder Commands used in the thesis

Running H.264/AVC Encoder commands

```

C:\WINDOWS\system32\cmd.exe
D:\>cd \jsum\bin
D:\jsum\bin>H264AVCEncoderLibTestStaticd.exe
JSUM 9.19.8 Encoder

supported options:
-pf      Parameter File Name
-bf      BitStreamFile
-frms    Number of total frames
-gop     GOPSize - GOP size (2,4,8,16,32,64, default: 1)
-iper    Intra period (default: -1) : must be a power of 2 of GOP size (or -1)
-numl    Number Of Layers
-cabac   CABAC for all layers as entropy coding mode
-vlc     ULC for all layers as entropy coding mode
-ecmf    <Layer> <entropy_coding_mod_flag>
-org     <Layer> <original file>
-rec     <Layer> <reconstructed file>
-ec      <Layer> <entropy coding mode>
-rqp     <Layer> <ResidualQP>
-mqp     <Layer> <Stage> <MotionQP>
-lqp     <Layer> <ResidualAndMotionQP>
-meqppl  <Layer> <MotionQPLooppass>
-ilpred  <Layer> <InterLayerPredictionMode>
-blid    <Layer> <BaseLayerId>
-bcip    Constrained intra prediction for base layer (needed for single-loop) i
n scripts
-anasip  <Layer> <SIP Analysis Mode>[0: persists all inter-predictions, 1: forb
ids all inter-prediction.] <File for storing bits information>
-encsip  <Layer> <File with stored SIP information>
-icsei   <IntegrityCheckSEIEnableFlag>[0: IntegrityCheckSEI is not applied, 1:
IntegrityCheckSEI is applied.]
-tlnest  <TlevelNestingFlag>[0: temporal level nesting constraint is not applie
d, 1: the nesting constraint is applied.]
-tlidx   <Tl0DepRepIdxSeiEnable>[0: tl0_dep_rep_idx is not present, 1: tl0_dep_r
ep_idx is present.]
-ciu     <Constrained intra upsampling>[0: no, 1: yes]
-kpm     <mode> [0:only for FGS(default), 1:FGS&MGS, 2:always]
-mgsctrl <mode> [0:normal encoding(default), 1:EL ME, 2:EL ME+MC]
-edqpc  <layer> <value> sets explicit QP cascading mode for given laye
r [0: no, 1: yes]
-dqpc   <layer> <level> <value> sets delta QP for given layer and temporal lev
el (in explicit mode)
-aeqpc  <value> sets explicit QP cascading mode for all layers
[0: no, 1: yes]
-adqpc  <level> <value> sets delta QP for all layers and given tempora
l level (in explicit mode)
-xdqp   <DQP0> <DDQP1> <DDQPN> sets delta QP for all layers (in explicit mode)
>
-mbaff  <Layer> <Mb Adaptive Frame Field Coding>
-paff  <Layer> <Picture Adaptive Frame Field Coding>
-ml     <mode> [0:disabled(default), 1:multi-layer lambda selection, 2:mode1x0.8]
]
-h      Print Option List
D:\jsum\bin>
  
```

Project: H264AVCEncoderLibStatic_vc8 Output File

Table 1: Project 1: H264AVCEncoderLibStatic_vc8 Output File

Filename	Status	Errors	Warnings
..\..\src\lib\H264AVCEncoderLib\H264AVCEncoderLibStatic_vc8.vcproj	Converted	0	1

Conversion Issues - ..\..\src\lib\H264AVCEncoderLib\H264AVCEncoderLibStatic_vc8.vcproj:

Project file successfully executed up as 'C:\Documents and

Appendix

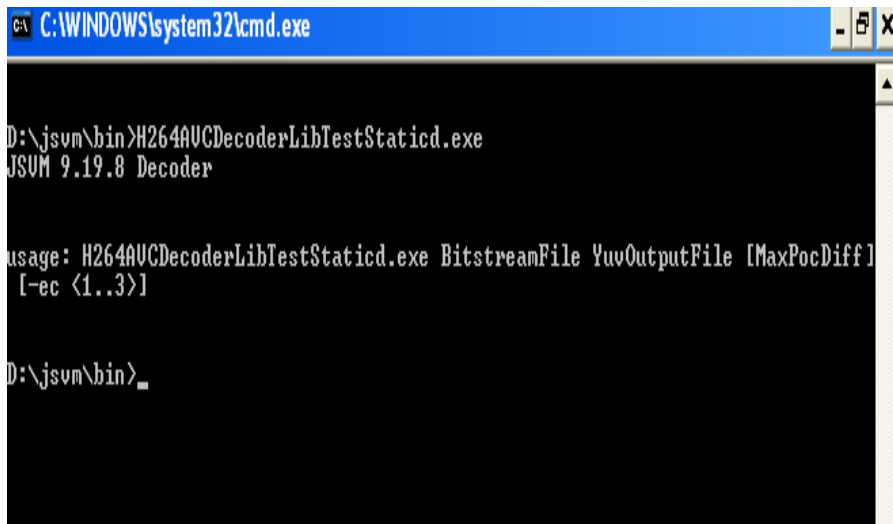
Settings\eepegasu\Desktop\JSVM\H264Extension\src\lib\H264AVCEncoderLib\H264AVCEncoderLibStatic_vc8.vcproj.8.00.old'.

Project compiled successfully.

This application has been updated to include settings related to the User Account Control (UAC) feature of Windows Vista. By default, when run on Windows Vista with UAC enabled, this application is marked to run with the same privileges as the process that launched it. This marking also disables the application from running with virtualization. You can change UAC related settings from the Property Pages of the project.

1 file	Converted: 1 0 1
	Not converted: 0

Running H.264/AVC Decoder commands



```
C:\WINDOWS\system32\cmd.exe
D:\jsvm\bin>H264AVCDecoderLibTestStaticd.exe
JSUM 9.19.8 Decoder

usage: H264AVCDecoderLibTestStaticd.exe BitstreamFile YuvOutputFile [MaxPocDiff]
[-ec <1..3>]

D:\jsvm\bin>_
```

Project 2: H264AVCDecoderLibTestStatic_vc8

Table 2: H264AVCDecoderLibTestStatic_vc8

Appendix

Filename	Status	Error s	Warning s
+ ..\..\src\test\H264AVCDecoderLibTest\H264AVCDecoderLibTestStatic_vc8.vcproj	Converted	0	1
<p>Conversion Issues - ..\..\src\test\H264AVCDecoderLibTest\H264AVCDecoderLibTestStatic_vc8.vcproj:</p> <p>Web deployment to the local IIS server is no longer supported. The Web Deployment build tool has been removed from your project settings.</p> <hr/> <p>Project file successfully executed up as 'C:\Documents and Settings\eepgasu\Desktop\JSVM\H264Extension\src\test\H264AVCDecoderLibTest\H264AVCDecoderLibTestStatic_vc8.vcproj.8.00.old'.</p> <hr/> <p>Project compiled successfully.</p> <hr/> <p>This application has been updated to include settings related to the User Account Control (UAC) feature of Windows Vista. By default, when run on Windows Vista with UAC enabled, this application is marked to run with the same privileges as the process that launched it. This marking also disables the application from running with virtualization. You can change UAC related settings from the Property Pages of the project.</p> <hr/>			
1 file	Converted:	0	1
		1	
	Not converted:	0	

Appendix

Running H.264/AVC DownConvert commands

```
C:\WINDOWS\system32\cmd.exe
D:\jsvm\bin>DownConvertStatic.exe
ERROR: wrong number of arguments
Usage: DownConvertStatic.exe <win> <hin> <in> <wout> <hout> <out> [<method>] [<t>]
 [<skip>] [<frms>]]] [<crop> <args>] [<phase> <args>] [<resample_mode> <arg>]]
win      : input width <luma samples>
hin      : input height <luma samples>
in       : input file
wout     : output width <luma samples>
hout     : output height <luma samples>
out      : output file

----- OPTIONAL -----
method   : rescaling methods <default: 0>
          0: normative upsampling
          1: non-normative downsampling <JUT-R096>
          2: dyadic upsampling <AUC 6-tap (1/2 pel) on odd samples>
          3: dyadic downsampling <MPEG-4 downsampling filter>
          4: crop only
          5: upsampling <Three-lobed Lanczos-windowed sinc>
          6: upsampling <JUT-004: AUC 6-tap 1/2 pel + bilinear 1/4 pel>
t        : number of temporal downsampling stages <default: 0>
skip     : number of frames to skip at start <default: 0>
frms     : number of frames wanted in output file <default: max>

----- OVERLOADED -----
-crop <type> <params>
type     : 0: Sequence level, 1: Picture level
params  : IF Sequence level: <x_orig> <y_orig> <crop_width> <crop_height>
          cropping window origin <x,y> and dimensions <width and height>
          IF Picture level: <crop file>
          input file containing cropping window parameters.
          each line has four integer numbers separated by a comma
          as following: "x_orig, y_orig, crop_width, crop_height"
          for each picture to be resampled;
-phase <in_uv_ph_x> <in_uv_ph_y> <out_uv_ph_x> <out_uv_ph_y>
in_uv_ph_x : input chroma phase shift in horizontal direction <default:-1>
in_uv_ph_y : input chroma phase shift in vertical direction <default: 0>
out_uv_ph_x: output chroma phase shift in horizontal direction <default:-1>
out_uv_ph_y: output chroma phase shift in vertical direction <default: 0>
-resample_mode <resample_mode>
resample_mode : resampling modes, present when method=0 <default: 0>
              0: low-res-frm = progressive, high-res-frm = progressive
              1: low-res-frm = interlaced, high-res-frm = interlaced
              2: low-res-frm = progressive <top-coincided>
              3: high-res-frm = interlaced
              4: low-res-frm = progressive <bot-coincided>
              5: high-res-frm = interlaced
              6: low-res-frm = interlaced <top-first>
              7: high-res-frm = progressive <double frm rate>
              8: low-res-frm = interlaced <bot-first>
              9: high-res-frm = progressive <double frm rate>

D:\jsvm\bin>
```

Running H.264/AVC Bit Stream Extractor Commands

```
C:\WINDOWS\system32\cmd.exe
D:\jsvm\bin>BitStreamExtractorStatic.exe
JSUM 9.19.8 BitStream Extractor

No arguments specified
Usage: BitStreamExtractorStatic.exe [-pt trace] InputStream [OutputStream [-e]
 [-s]] [-l] [-t] [-f] [-b] [-et]]
options:
-pt trace  -> generate a packet trace file "trace" from given stream
-sl SL     -> extract the layer with layer id = SL and the dependent low
er layers
-l L       -> extract all layers with dependency_id <= L
-t T       -> extract all layers with temporal level <= T
-f F       -> extract all layers with quality_level <= F
-b B       -> extract a layer (possibly truncated) with the target bitrate = B
-e AxBEC:D -> extract a layer (possibly truncated) with
              - A frame width [luma samples]
              - B frame height [luma samples]
              - C frame rate [Hz]
              - D bit rate [kbit/s]
-et        -> extract packets as specified by given <modified> packet trace
ace file
-r RATE    -> extract rate of highest layer wo trunc. (use QL when present)
            - may be specified as percentage <0% no ql, 100% all qls>
-dtql     -> do not truncate quality layers for "-r" <only remove>
-sip       -> extract using SIP algorithm
-ql        -> information about quality layers are used during extraction
-n         -> ordered/toplayer quality layer extraction
            - simulates truncation using normal ql even if MLQL assignment
            was used
-keepf     -> use with "-l" and "-f" options: extract all included layers
of the layer L specified with "-l" and all quality levels below quality level
F specified with "-f" of the layer L
Options "-l", "-t" and "-f" can be used in combination with each other.
Other options can only be used separately.

D:\jsvm\bin>
```

Appendix

Running H.264/AVC Fixed Quantisation Parameter Commands

```
C:\WINDOWS\system32\cmd.exe

D:\jsvm\bin>FixedQPEncoderStaticd.exe
JSUM SUC FixedQPEncoder

usage:
FixedQPEncoder <ConfigFile>

D:\jsvm\bin>_
```

Running H.264/AVC Motion Compensation Temporal Filtering Commands

```
C:\WINDOWS\system32\cmd.exe

D:\jsvm\bin>MCTFPreProcessorStaticd.exe
JSUM MCTF Pre-Processor

Usage: MCTFPreProcessor -w Width -h Height -f frms -i Input -o Output
[-gop GOPSize] [-qp QP]

-w Width - frame width in luma samples (multiple of 16)
-h Height - frame height in luma samples (multiple of 16)
-f frms - number of frames to be processed (>1)
-i Input - input sequence
-o Output - output sequence
-gop GOPSize - GOP size for MCTF (2,4,8,16,32,64, default: 16)
-qp QP - QP for motion estimation and mode decision
(>0, default: 26)

D:\jsvm\bin>_
```

Running H.264/AVC Peak Signal-to-Noise Ratio Commands

```
C:\WINDOWS\system32\cmd.exe

D:\jsvm\bin>PSNRStaticd.exe
Usage: PSNRStaticd.exe <w> <h> <org> <rec> [<t> [<skip> [<strm> <fps> [strg]]]]
[-r]

w : original width <luma samples>
h : original height <luma samples>
org : original file
rec : reconstructed file
t : number of temporal downsampling stages <default: 0>
skip : number of frames to skip at start <default: 0>
strm : coded stream
fps : frames per second
strg : prefix string for summary output
-r : return luma psnr <default: return -1 when failed and 0 otherwise>

D:\jsvm\bin>
```

Appendix

Running H.264/AVC QualityLevelAssigner commands

```
D:\jsvm\bin>QualityLevelAssignerStaticd.exe
JSVM Quality Assigner

Info: This tool relies on the scalable SEI message
      This tool requires prefix NAL units in front of each H.264 NAL unit
      This tool assumes a fixed GOP size (non-AGS) throughout a sequence

Usage: QualityLevelAssigner -in Input -org L Original [-org L Original]
                             [-out Output [-sei] [-wp DatFile] [-dep] [-ind] [-ml
q1]
or   QualityLevelAssigner -in Input -out Output -rp DatFile [-sei]

-in Input      - input bit-stream
-out Output    - output bit-stream with determined quality layer id's
-org L Original - original image sequence for layer L
-wp DatFile    - data file for storing rate and distortion values
-rp DatFile    - data file with previously computed rate and
                distortion values
-sei           - provide quality layer info using SEI messages
-dep           - determine only dependent distortions
                (speed-up by factor of 2, slight coding eff. losses)
-ind           - determine only independent distortions
                (speed-up by factor of 2, slight coding eff. losses)

-mlq1         - determine Multi Layer quality layer id's

D:\jsvm\bin>
```

Running H.264/AVC YUV Comparison Commands

```
D:\jsvm\bin>YUVCompareStaticd.exe

Usage: YUVCompare (width) (height) (test.yuv) (ref.yuv) [(max_count)=0] [(ignore_
frames)=0]

D:\jsvm\bin>
```

Appendix

Appendix D: Matlab Scripts for MDC

```
% CMCR 3D Multiple description coding techniques
%Multidescription. m
clc;
or_path=' interview.yuv' ;

%? Original? file
dest_pathL = 'interview_sideL.yuv';
dest_pathR = ' interview_sideR.yuv' ;

height=288;
width=352;
N=15;
frame_size=height*width*1.5;
f_or = fopen(or_path, 'rb');
original = fread(f_or);
fclose(f_or);
flag = 0;
sideL = [];
sideR = [];
for i=0:N-1

if(flag == 0)
sideL= [sideL ; original(i*frame_size+1: (i+1)*frame_size)];
flag = 1;
else
sideR= [sideR ; original(i*frame_size+1: (i+1)*frame_size)];
flag=0;
end
end

f_dst=fopen(dest_pathL, ' wb' );
fwrite(f_dst, sideL, ' uint8' );
fclose(f_dst);
f_dst=fopen(dest_pathR, ' wb' );
fwrite(f_dst, sideR, ' uint8' );
fclose(f_dst);

% CMCR video reconstruction Algorithms
%Reconstruction. m
clc
f_orL=fopen( ' interview_sideL. yuv' , ' rb' ) ;
f_orR=fopen( ' interview_sideR. yuv' , ' rb' ) ;
```

Appendix

```
originalL=fread(f_orL);
originalR=fread(f_orR);
fclose(f_orL);
fclose(f_orR);

height=288;
width=352;
N=30;
frame_size=height*width*1.5;
f_rec=fopen( 'interview_recon.yuv' , 'a' );
for i=0:N/2-1
fwrite(f_rec, originalL(i*frame_size+1:(i+1)*frame_size), 'uint8');
fwrite(f_rec, originalR(i*frame_size+1:(i+1)*frame_size), 'uint8');
end
fclose(f_rec);

% CMCR 3D MDC Frame Interpolation Techniques for side encoder Right (Even)
% InterpolationR.m
clc
or_path='interview_sideR_50.yuv'; % Original? file
dest_pathL = 'interview_sideR_50_30frames.yuv';
f_or = fopen(or_path, 'rb');
original = fread(f_or);
fclose(f_or);

height=288;
width=352;
N=30;
frame_size=height*width*1.5;
new = [];
i=0;
new = [new ; original(i*frame_size+1: (i+1)*frame_size)];
for i=0:N/2-2
new = [new ; original(i*frame_size+1: (i+1)*frame_size)];
new = [new ; (original(i*frame_size+1: (i+1)*frame_size)+original((i+1)*frame_size+1:(i+2)*frame_size) )/2];
end
new = [new ; original(i*frame_size+1: (i+1)*frame_size)];

f_dst=fopen(dest_pathL, 'wb');
fwrite(f_dst, new, 'uint8');
fclose(f_dst);
```

Appendix

Appendix E: 3D MDC Comparison

```
% Function file for 3D MDC protocols and the
% new scheme developed in CMCR lab 15/11/2010
function [MDCSIPA,MDCSIMI,SDC]=comparison_MDC(N,ps,pd,K)
pu=ps;
% find Codec efficiency for 3D MDC using Pixel Averaging technique (3D MDC SIPA) and
% 3D single description coding
for i = 1:length(K)
    r=K(i);
    MDCSIPA(i) = pu*power(pd,r)/(pu*power(pd,r)+N*(1-pu*power(pd,r)));
    k=K(i);
    R = 1+(k-2)*(1-pd);
    pin = (ps*pd^R)/(ps*pd^R+N*(ps-ps*(1-pd)^k-ps*pd^k));
    pir = 1-pin;
    SDC(i)= (pin*ps*pd^k + pir*ps*pd^R)/(pin*(ps*pd^k+N*(1-ps*pd^k))+pir*(ps*pd^R+N*(1-
ps*pd^R)));
end
% find Codec Efficiency for 3D MDC using Motion Interpolation (3D MDC-SIMI)
pu=ps;
for i=1:length(K)
    r=K(i);
    MDCSIMI(i)=pu*power(pd,r)/(1+(N-1)*pu*power(pd,r));
end
% estimate probabilities Qi
function [upj,Q]=estimate_Qis(p,N,m,R)
upQ=2;
for j=1:upQ
    f(j)=(1-(1-p)^(j*m+1))^R;
    f_(j)=((1-(1-p)^(j*m+1))^R - p^R)/(1-p^R);
end
Q(1)=power(f(1),N-1)*f_(1);
for i=2:upQ
    sumQ=0;
    for k=1:i-1
        sumQ=sumQ+Q(k);
    end
    Q(i)=f(i)*f_(i)-sumQ;
end
baseupQ=upQ;
while sum(Q) < 0.999999
    upQ=upQ+1;
    j=upQ;
    f(j)=(1-(1-p)^(j*m+1))^R;
    f_(j)=((1-(1-p)^(j*m+1))^R - p^R)/(1-p^R);
    sumQ=sumQ+Q(j-1);
    Q(j)=f(j)*f_(j)-sumQ;
end
upj=upQ;
```

Appendix

```
% Comparison of 3D MDC script file
clear all;clc;
k=[1 2 5 10 10 20 50 100 200 500 1000];
ps = 1;
pd = 0.999;
% for N = 50
% CODEC (Encoder & Decoder) efficiency of all schemes at N = 50
N= 50;
grid;
fig(1)=figure;
[MDCSIPA MDCSIMI SDC]=comparison_MDC(N,ps,pd,k);
semilogx(k,MDCSIPA,'kd-');hold on
semilogx(k,MDCSIMI,'bo-');hold on;
semilogx(k,SDC,'rs-');hold on;
xlabel('Number of receivers K');
ylabel('Encoder/Decoder efficiency');
title('CODEC Efficiency vs. number of receivers ( N= 50)');
legend('3D MDC-SIPA Algorithm','3D SDC Algorithm','3D MDC-SIMI Algorithm',1);
grid on;
fprintf(1,'Figure 1 \n');
%pause
% for N = 250
% CODEC Efficiency of all 3D MDC for N = 250
N = 250;
grid;
[MDCSIPA MDCSIMI SDC]=comparison_MDC(N,ps,pd,k);
fig(2)=figure;
semilogx(k,MDCSIPA,'kd-');hold on
semilogx(k,MDCSIMI,'bo-');hold on;
semilogx(k,SDC,'rs-');hold on;
xlabel('Number of receivers K');
ylabel('Encoder/Decoder efficiency');
title('CODEC Efficiency vs. number of receivers ( N= 250)');
legend('3D MDC-SIPA Algorithm','3D SDC Algorithm','3D MDC-SIMI Algorithm',1);
grid on;
fprintf(1,'Figure 1 \n');
%pause;
grid;
% for N = 500
%CODEC efficiency of all 3D MDC for N = 500
N= 500;
grid;
[MDCSIPA MDCSIMI SDC]=comparison_MDC(N,ps,pd,k);
fig(3)=figure;
semilogx(k,MDCSIPA,'kd-');hold on
semilogx(k,MDCSIMI,'bo-');hold on;
semilogx(k,SDC,'rs-');hold on;
xlabel('Number of receivers K');
ylabel('Encoder/Decoder efficiency');
```

Appendix

```
title('CODEC Efficiency vs. number of receivers ( N= 500)');
legend('3D MDC-SIPA Algorithm','3D SDC Algorithm','3D MDC-SIMI Algorithm',1);
grid on;
fprintf(1,'Figure 1 \n');
%pause
% for N = 1000
% CODEC efficiency of all MDC for N = 1000
N = 1000;
grid;
[MDCSIPA MDCSIMI SDC]=comparison_MDC(N,ps,pd,k);
fig(4)=figure;
semilogx(k,MDCSIPA,'kd-');hold on;
semilogx(k,MDCSIMI,'bo-');hold on;
semilogx(k,SDC,'rs-');hold on;
xlabel('Number of receivers K');
ylabel('Encoder/Decoder Efficiency');
title('CODEC Efficiency vs. number of receivers ( N= 1000)');
legend('3D MDC-SIPA Algorithm','3D SDC Algorithm','3D MDC-SIMI Algorithm',1);
grid on;
fprintf(1,'Figure 1 \n');
```


Appendix

Appendix F: Colour extraction from 3D image

```
%Colour extraction from 3D images
%greyscale image.
%Image_rgb=imread('No corruption 1.bmp');
%Image_rgb=imread('texture corrupted 3.bmp');
%Image_rgb=imread('motion corrupted.bmp');
%Image_rgb=imread('color-cam2-f009.bmp');
Image_rgb=imread('breakdancers output.bmp');
Image_r = Image_rgb(:,:,1);
Image_g = Image_rgb(:,:,2);
Image_b = Image_rgb(:,:,3);

subplot(2,4,1);
imshow(Image_rgb);
subplot(2,4,2);
imshow(Image_r, []);
subplot(2,4,3);
imshow(Image_g, []);
subplot(2,4,4);
imshow(Image_b, [])

set(gcf, 'Position', get(0,'Screensize')); % Maximize figure.

% Save it to a file
imwrite(Image_rgb, 'texture corrupted 1.bmp', 'BitDepth', 16);
% Recall it.
recalledImage = imread('texture corrupted 1.bmp');
% See what it's class is.
class(recalledImage)
imshow(Image_g);
```

Appendix

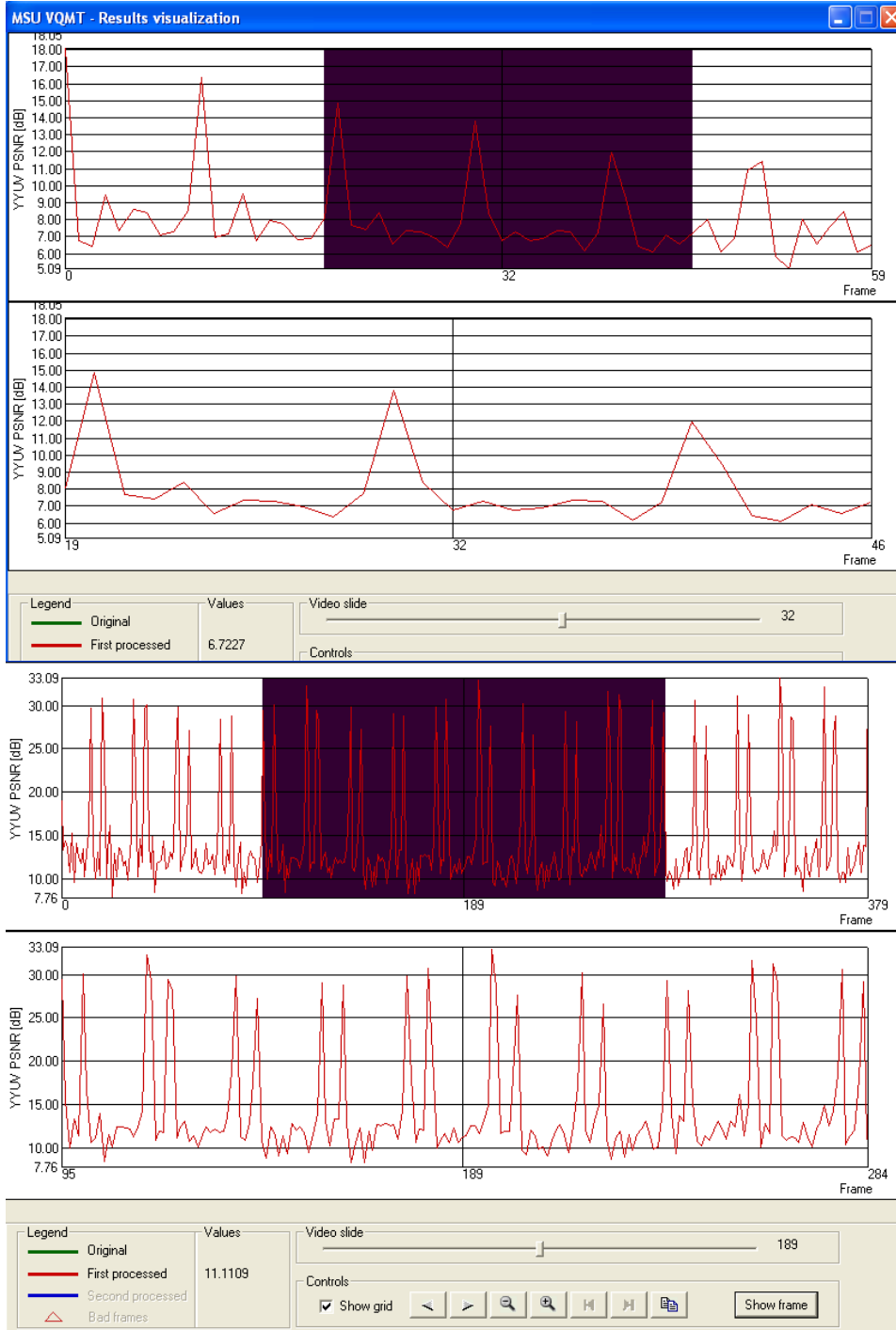
Appendix G: 3D Video Quality Objective Assessment

Frame	SNR_Y	SNR_U	SNR_V
1	12.3097	12.2585	12.3183
2	12.0319	12.0747	11.9572
3	13.6934	13.7072	13.6785
4	13.8813	13.8465	13.8611
5	11.9026	11.807	11.9327
6	12.2273	12.1811	12.1925
7	13.4165	13.4207	13.445
8	12.1127	12.0353	12.1512
9	11.714	11.7645	11.69
10	13.6943	13.5465	13.9366
11	12.3699	12.345	12.4035
12	10.9401	10.8459	11.08
13	11.8058	11.8106	11.8015
14	13.8791	13.7689	13.9279
15	13.0285	13.04	13.0178
16	10.6822	10.6805	10.7028
17	12.9519	12.922	13.0011
18	13.6711	13.6849	13.5849
19	10.8002	10.7069	10.8437
20	11.542	11.5327	11.5031
21	13.0724	13.1045	13.0011
22	12.6075	12.6062	12.5771
23	10.8487	10.799	10.8671
24	12.1186	12.1762	12.1543
25	13.8033	13.8439	13.7189
26	11.2494	11.2576	11.2536
27	11.5067	11.4996	11.493
28	13.3631	13.3393	13.41
29	12.776	12.7903	12.7921
30	10.3621	10.3746	10.3424
31	11.7624	11.7156	11.7914
32	13.5779	13.5266	13.5646
33	12.1683	12.151	12.1166
34	10.7883	10.7634	10.7504
35	12.8222	12.9121	12.7899
36	13.9652	13.9853	14.0113
37	10.2471	10.211	10.2115
38	11.2832	11.2875	11.3236
39	12.994	12.9956	12.9871
40	12.307	12.3388	12.2653
41	10.7996	10.8046	10.823
42	12.251	12.4069	12.1542
43	13.5926	13.5294	13.5633
44	11.367	11.3297	11.3918
45	11.5876	11.6308	11.4791
46	13.6568	13.669	13.6871
47	13.0277	13.1061	13.0017
48	10.5015	10.4468	10.5141
49	11.8818	11.8403	11.8794
50	13.1234	13.1428	13.1304
51	12.3579	12.4494	12.2813
52	10.9462	10.9399	10.9642
53	13.6826	13.5445	13.8524
54	13.5928	13.589	13.6483
55	10.6025	10.6373	10.5871
56	11.8047	11.8151	11.8179
57	12.7859	12.8765	12.779
58	13.0112	13.0336	13.007
59	10.772	10.78	10.7823
60	12.5828	12.5664	12.5998
61	14.0749	14.0808	14.0927
62	11.3639	11.4003	11.3088
63	11.6099	11.6334	11.6607
64	14.1791	14.2233	14.2177
65	13.4537	13.5299	13.4187
66	10.6361	10.6749	10.6037
67	12.2908	12.3021	12.3326
Average	12.2957	12.2931	12.2989

Appendix



Appendix



Appendix

Frame	SNR_Y	SNR_U	SNR_V
1	18.1407	18.2286	18.1369
2	17.4647	17.4996	17.4023
3	19.2249	19.3875	19.0971
4	20.0985	20.1659	20.1109
5	17.9426	18.0279	17.9675
6	16.7603	16.5657	16.873
7	21.6571	21.6717	21.5898
8	18.8985	18.8157	18.967
9	17.8393	17.9221	17.833
10	18.3724	18.3398	18.5399
11	20.4675	20.4446	20.5195
12	18.0857	18.1409	17.9511
13	17.15	17.3289	17.0608
14	19.7734	19.8254	19.7134
15	19.6269	19.563	19.7266
16	17.998	17.9475	18.0838
17	17.5739	17.6766	17.5239
18	21.3509	21.28	21.5075
19	18.0234	18.114	18.0385
20	17.3839	17.3579	17.4135
21	19.0412	19.1681	19.1388
22	20.0187	20.1501	19.9339
23	17.8896	17.8762	17.8365
24	16.7552	16.7223	16.6709
25	21.0123	20.9168	21.0659
26	18.9716	18.9739	18.8983
27	17.6935	17.7234	17.6871
28	18.4258	18.3308	18.6069
29	19.9445	19.9608	19.9495
30	17.9708	18.0948	17.9248
31	17.0626	17.0749	16.9706
32	19.4616	19.5034	19.4476
33	19.3767	19.4397	19.4191
34	17.9844	18.025	17.9739
35	17.6327	17.5751	17.6265
36	20.5566	20.6804	20.5685

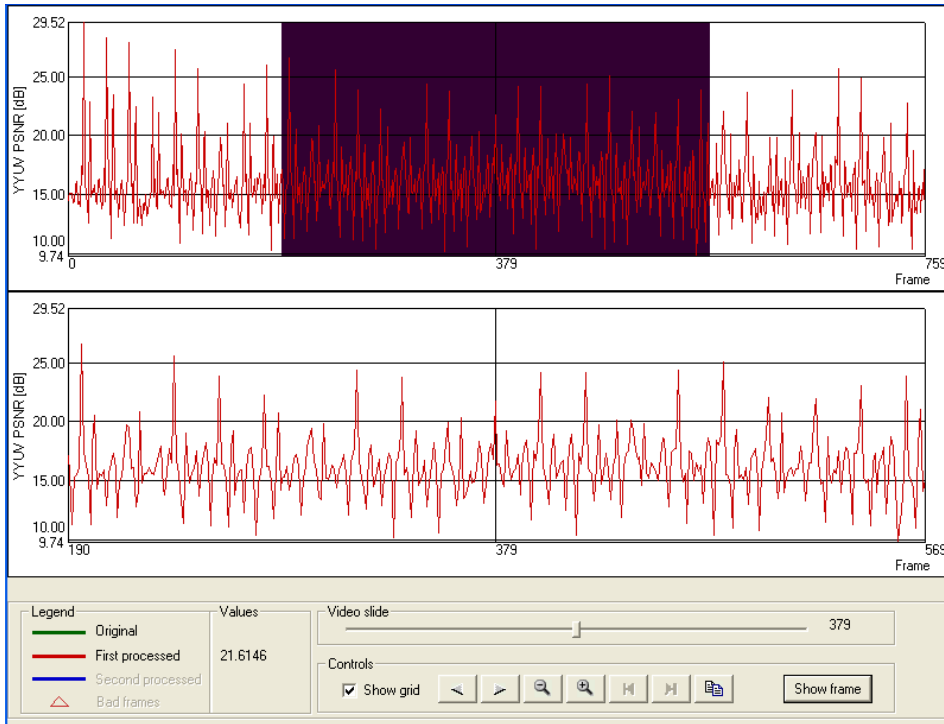
Appendix

37	18.1143	18.1342	17.9936
38	17.3633	17.3336	17.3985
39	18.953	18.9247	19.0806
40	19.5771	19.5954	19.6757
41	17.9803	18.0948	17.9005
42	16.7849	16.6901	16.8149
43	20.5788	20.485	20.7565
44	19.1187	19.0363	19.2371
45	17.8736	17.7787	17.9531
46	18.3794	18.2303	18.4669
47	19.7581	19.7947	19.7752
48	18.0678	18.1018	18.0797
49	17.1346	17.1441	17.0872
50	19.2752	19.3309	19.3294
51	19.5527	19.6392	19.5173
52	18.1721	18.1781	18.1137
53	17.508	17.4969	17.4215
54	20.4209	20.4594	20.44
55	18.1754	18.2885	18.032
56	17.7135	17.6911	17.6889
57	18.7935	18.7451	18.8763
58	19.6513	19.6081	19.704
59	18.116	18.2696	17.9602
60	16.876	16.9046	16.8786
61	20.5877	20.6249	20.6546
62	18.9676	18.97	18.9503
63	18.1442	18.2416	18.0334
64	18.4207	18.444	18.4251
65	19.6424	19.7118	19.6402
66	18.1603	18.0487	18.2654
67	17.3088	17.3408	17.2567
68	19.3374	19.3276	19.4251
69	19.4582	19.4748	19.5192
70	18.1132	18.1609	18.1177
71	17.7694	17.829	17.7802
72	20.5233	20.5671	20.6311
73	18.0037	17.9022	18.1689
74	17.7936	17.9226	17.7386

Appendix

75	19.0033	18.9126	19.0112
76	19.636	19.6652	19.6014
77	17.9451	18.0017	17.9481
78	17.0869	17.12	17.0898
79	20.3465	20.446	20.3385
80	18.9392	18.9473	18.9118
81	18.1758	18.2245	18.152
82	18.418	18.543	18.3464
83	19.2269	19.2156	19.2796
84	18.0496	17.9993	18.0803
85	17.2549	17.276	17.308
86	19.0286	19.0175	19.0511
87	19.3727	19.2525	19.4839
88	18.1509	18.229	18.0636
89	17.6594	17.6878	17.6562
90	20.1229	20.1897	20.0924
91	18.1139	18.1142	18.167
92	17.7306	17.6793	17.7929
93	18.9908	19.0756	19.037
94	19.634	19.6119	19.7129
95	18.1033	18.1137	18.1648
96	16.9751	16.9162	17.0309
97	20.3246	20.3375	20.2901
98	18.9832	18.9005	19.1437
99	17.9763	18.008	18.0083
100	18.3179	18.2567	18.4083

Appendix



Appendix

Frame	SNR_Y	SNR_U	SNR_V
1	14.5164	14.6245	14.593
2	14.5605	14.6446	14.6041
3	16.858	17.0037	16.9995
4	15.3263	15.6121	15.1638
5	14.919	15.0007	14.8784
6	14.9583	14.9661	15.0414
7	15.9471	16.0596	16.0102
8	15.5975	15.9103	15.5781
9	14.5706	14.6624	14.5949
10	16.6332	16.7067	16.7304
11	14.8551	14.9411	14.8947
12	13.3851	13.4618	13.3714
13	14.0582	14.2208	13.9193
14	15.408	15.3141	15.3683
15	15.2884	15.5189	15.287
16	14.9452	15.0213	14.8707
17	15.2658	15.434	15.2215
18	15.5538	15.6492	15.5975
19	14.4697	14.6582	14.5909
20	14.5988	14.7133	14.7451
21	16.9181	17.0903	17.0729
22	14.762	14.8652	14.7635
23	15.1085	15.1585	15.0445
24	14.74	14.8338	14.6477
25	15.2076	15.2278	15.1929
26	15.2616	15.5398	15.352
27	14.684	14.7545	14.5835
28	15.0398	15.1161	14.9797
29	14.7935	14.8775	14.887
30	14.8523	15.0217	14.9831
31	15.0247	15.1147	15.1031
32	15.4569	15.4674	15.5274
33	15.1685	15.1886	15.2373
34	15.6555	15.6854	15.6157

Appendix

35	15.5626	15.557	15.5218
36	15.139	15.106	15.0489
37	14.8774	14.8122	14.8853
38	14.9977	15.1655	15.0344
39	17.0776	17.1673	17.0177
40	14.6885	14.7301	14.6494
41	15.6688	15.7532	15.7388
42	15.0165	15.1489	15.0879
43	14.9545	14.9202	14.9836
44	16.0268	16.0274	15.7648
45	14.7968	14.7153	14.8069
46	16.7618	16.7269	16.7584
47	14.8758	14.8688	14.8258
48	15.2098	15.2338	15.1879
49	14.6338	14.7126	14.6133
50	15.6385	15.6152	15.6421
51	15.4914	15.5369	15.6778
52	15.7835	15.8591	15.8856
53	15.3687	15.3272	15.4516
54	15.1875	15.1549	15.1674
55	14.8285	14.8996	14.99
56	14.7355	14.7817	14.7851
57	16.9871	17.0635	17.0079
58	14.8583	14.8451	14.8554
59	15.5685	15.6882	15.3334
60	14.9525	15.0267	14.8658
61	15.3592	15.3249	15.3062
62	15.8145	15.7695	15.95
63	14.4333	14.3952	14.435
64	17.0485	17.1878	17.0652
65	14.6237	14.5695	14.6547
66	14.8563	14.8232	14.6776
67	15.3073	15.3514	15.294
68	16.6653	16.747	16.5373
69	15.9144	15.8989	15.8498
70	16.045	16.1155	16.1356
71	15.722	15.7686	15.6495
72	15.6705	15.6817	15.6682

Appendix

73	15.1612	15.3483	15.1969
74	15.4496	15.4964	15.4928
75	17.2754	17.2831	17.2349
76	15.1952	15.3499	15.1864
77	16.0773	16.3071	16.0702
78	15.4076	15.5203	15.3429
79	16.7316	16.7655	16.6527
80	16.2195	16.4997	16.1958
81	14.9687	14.9845	14.9925
82	17.8164	17.7744	17.8423
83	14.9962	15.0183	14.9926
84	15.5893	15.812	15.703
85	15.5883	15.6385	15.612
86	15.9086	15.9157	15.8753
87	15.2406	15.2388	15.1167
88	15.9749	15.9965	15.9133
89	15.7461	15.7657	15.7549
90	15.1833	15.1395	15.2162
91	15.0497	14.9419	15.1617
92	14.7523	14.7821	14.7287
93	16.888	16.9476	16.8531
94	14.7514	14.7358	14.8715
95	15.7935	15.9952	15.8061
96	14.9167	15.0135	14.89
97	15.0183	15.016	15.0089
98	16.1951	16.1043	16.0943
99	14.3907	14.3803	14.4064
100	14.0547	13.9607	14.1226
101	14.338	14.3882	14.3097
102	15.4257	15.5884	15.4545
103	14.9814	15.1133	14.9437
104	16.5544	16.5896	16.5441
105	14.6821	14.7042	14.6411
106	15.623	15.5909	15.6563
107	15.2274	15.2174	15.2158
108	15.2372	15.2076	15.3023
109	13.3559	13.3611	13.3084
110	13.7994	13.9407	13.7262

Appendix

111	15.7164	15.7702	15.8095
112	14.622	14.7705	14.6271
113	15.3816	15.5738	15.4175
114	14.7586	14.8276	14.7619
115	15.3344	15.4027	15.2901
116	15.9582	16.1481	15.9295
117	14.3702	14.4285	14.3926
118	17.3524	17.3952	17.3587
119	14.6516	14.7217	14.5983
120	15.429	15.5314	15.4282
121	14.8188	14.8709	14.7977
122	16.3917	16.3605	16.4311
123	15.2379	15.3432	15.1226
124	15.6122	15.6703	15.5879
125	15.2361	15.2497	15.2518
126	15.1411	15.1634	15.1265
127	14.5579	14.6616	14.4659
128	14.597	14.6118	14.6247
129	16.6982	16.814	16.6778
130	13.8794	13.8312	13.891
131	14.8013	14.8849	14.7553
132	14.5826	14.6141	14.5429
133	14.0881	14.1415	14.0945
134	14.7171	14.7727	14.6924
135	13.8876	13.98	13.8001
Average	15.2957	15.3564	15.2946

Appendix

