Sudan University of Science & Technology

College of Graduate Studies

Assessment and Modelling of Quality of Experience for user of Mobile Devices

تقييم ونمذجة جودة التجربة لمستخدمي الأجهزة النقالة

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Computer Science

By:

Abubker Elmnsi Abdalla Mohamed

Supervisors:

Prof. Is-Haka Mkwawa

Dr.Niema Izeldeen

October 2017

Abstract

Quality of Experience (QoE) in multimedia networking and communication is a wider term that encompasses not only network impairments and video signal quality, but also covers other factors such as the end user devices, service infrastructure, network bandwidth, user expectation and the environment where the end user is communicating or consuming multimedia services. Smart phones and other portable devices such as tablets and game consoles have experienced exponential growth in numbers over the recent years. This tremendous growth has come with a lot challenges, amongst them are power consumption and quality of experience. The challenges in the quality of experience are due to the fact that there are many mobile device types from different manufactures having varying display technologies, size, memory and processors. These varying technologies make assessment, measurements and prediction of quality of experience a challenge for researchers in academia and industrial domain. In this study, mobile device screen size and mobile device user preferences are considered as mobile device context parameters for which their impact on the quality of experience over the video services is investigated. Through subjective tests, the impact of mobile device pixel density on the video quality is evaluated and the new quality of experience model is proposed that takes pixel density into consideration. The proposed quality model is validated through unseen video sequences, and the results have shown that the proposed model performs well regarding correlation coefficient. This study, through subjective tests, has also conducted an investigation on the impact of mobile device preferences on the quality of experience for video services. Based on this investigation, the experimental results has shown that quality of experience is highly correlated with user preferences on mobile devices. The new proposed quality of experience model and the findings on the impact of mobile device user preferences on the quality of experience have potential use for multimedia service providers such as YouTube, Facebook and Netflix in areas of quality control and optimization of network and multimedia services.

المستخلص

جودة التجربة في الشبكات متعددة الوسائط هو مصطلح واسع لا يشمل ضعف الشبكة وجودة إشارة الفديو فحسب، بل يغطي عوامل أخرى مثل أجهزة المستخدم، والبنية التحتية للخدمات، وعرض النطاق الترددي للشبكة، وتوقعات المستخدم والبيئة التي يتصل فيها المستخدم او يستهلك فيها خدمات الوسائط المتعددة.شهدت الهواتف الذكية وغيرها من الأجهزة المحمولة مثل الأجهزة اللوحية وأجهزة الألعاب نمواً هائلا في العدد على مدى السنوات الأخيرة. وقد جاء هذا النمو الهائل مع الكثير من التحديات، من بينها استهلاك الطاقة و جودة التجربة. تعزى التحديات في جودة التجربة إلى حقيقة أن هناك العديد من أنواع الأجهزة النقالة المصممة بواسطة مختلف المصنعين، ويؤدي ذلك إلى تتنوع في تقنيات العرض وحجم والذاكرة والمعالجة. تجعل هذه التقنيات المتنوعة التقييم والقياسات والتنبؤ بجودة التجربة تحديا للباحثين في المجال الأكاديمي والمجال الصناعي.في هذه الدراسة، يُعتبر حجم شاشة جهاز الهاتف النقال وتفضيل المستخدم لنوع الهاتف معاملات سياق للهاتف المحمول. ومن ثم يتم تقييم تأثير هذه المعاملات على جودة التجربة الخاصة بخدمات الفيديو التي أُجري عليها التحقُق. من خلال الاختبارات الذاتية، تم تقييم تأثير كثافة بكسل الهاتف المحمول على جودة الفيديو، كما تم إقتراح نموذج جديد لجودة التجربة يأخذ كثافة البكسل بعين الاعتبار. تم التحقق من صحة نموذج المقترح من خلال تسلسل فيديو غير مرئي. وقد أظهرت النتائج أن النموذج المقترح يؤدي أداءً جيداً فيما يتعلق بمعامل الارتباط.تم في هذه الدراسة أيضا، من خلال الاختبارات الذاتية، إجراء تحقيقا حول تأثير تفضيلات الجهاز المحمول على جودة التجربة لخدمات الفيديو. وبناء على هذا التحقيق، أظهرت النتائج التجريبية أن جودة التجربة ترتبط ارتباطا وثيقا بتفضيل المستخدم للهاتف المحمول.

يمكن أن يستخدم النموذج الجديد المقترح لجودة التجربة والنتائج المتعلقة بتأثير تفضيل المستخدم للهاتف المحمول على جودة التجربة في مجالات مراقبة الجودة وتحسين الشبكة والوسائط المتعددة لمقدمي خدمات الوسائط المتعددة مثل يوتيوب و فيسبوك ونيتفليكس.

Acknowledgements

There are a number of individuals without whom this thesis would not have been accomplished. First and foremost, I would like to thank my supervisory team; Dr Is-Haka Mkwawa and Dr Niemah Osman. I feel privileged to have had such a wonderful, knowledgeable and supportive supervisors, thank you for all your time, advices and guidance. Last, but not least I owe a massive gratitude's to my loving family who believed in me and gave me the strength and motivation during the course of this research. My friends Awadalaa, Ahmed, Ayman, Abdosh,Moaid. Ayoosh; you are my hidden true and you are my power. My sister Howaida we miss you, Mum and dad, you are my true heroes and I owe all my achievements to you all.

Declaration

"I, Abubakr Elmnsi, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Quality of Experience Assessment and Modelling over Mobile Devices.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree

at this University;

2. Where any part of this thesis has previously been submitted for a degree or anyother qualification at this University or any other institution, this has been clearly stated;
3. Where I have cited the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. parts of this work have been published as indicated on page iv:"

Name: ______________________________ Signature:_________________________

Date: __________________

Publications to Date

Abubkr Elmnsi, Niemah Osman and Is-Haka Mkwawa. The Impact of Mobile Device Preference on the Quality of Experience. International Journal of Computing and information Sciences, ISSN: 1708-0460. Volume 12, Number 1, P89-94, September 2016.

Abubkr Elmnsi, Niemah Osman and Is-Haka Mkwawa. Mobile Devices Pixel Density and Video Quality. The NAFOSTED Conference on Information and Computer Science. Hanoi, Vietnam November 24-25, 2017

Dedication

I dedicate this thesis to my beloved family without their support and prayers, I would never have completed my study.

List of Figures

List of Tables

CHAPTER I

1 Introduction

This chapter presents the motivations behind this research project, the research questions, and the aims and objectives of the project. In addition, it outlines the main contributions of this research. The chapter is organized as follows. Section 1.1 presents the motivations of this research. Section 1.2 discusses the e research questions. Aims and objectives of this research are presented in Section 1.3. The research contributions are outlined in in Section 1.4. The organization of the thesis is presented in Section 1.5.

1.1 Research Motivations

Advances in video coding have enabled significant reduction in video transmission, bandwidth and storage. For instance, for the same video quality, the HEVC video codec [2] can half the transmission bandwidth requirement compared to H.264 video codec [3]. Furthermore, advances in mobile network technologies from 2G to 4G have led to an increase in network capacity and hence, the reduction in video transmission bandwidth requirements. The increase in network bandwidth has prompted an increase in the proliferation of video applications and streaming services such as YouTube and Netflix.

Despite these advances in video coding and mobile networks, the success of video applications and services will rely on the perceived Quality of Experience

(QoE). In this context, QoE measurements and prediction is vital to multimedia and network service providers in order to avoid churn and increase business revenues.

Subjective and objective methods are used for QoE measurement and prediction. Subjective methods involve tests for which participants are asked to grade the video quality on a five point scale known as Mean Opinion Score (MOS). MOS values are ranged from “bad” (1) to “excellent” (5). Subjective tests are usually conducted in a laboratory also known as controlled environment. Although subjective tests can provide accurate results, they are slow, time-consuming and expensive. They also cannot be repeatable and cannot be used to monitor video quality in an on-line and real-time environment [4].

Objective methods can be conducted in two ways, intrusive or non-intrusive. The intrusive way needs access to the original video sequence to compare it with the impaired sequence. Peak- Signal-to-Noise-Ratio (PSNR), Structural Similarity (SSIM) and Video Quality Metric (VQM) are examples of intrusive ways of video quality measurements [5]. In Non-intrusive ways, measurements and prediction of video quality are done by using network layer and application parameters without the need to access the original video. The Non-intrusive way is the one that is mainly used in on-line and real-time video quality measurements and prediction.

Existing studies such as [6, 7, 8, 9, 10] on non-intrusive video quality assessment and predictions rely mainly on coding parameters while others such as [11, 12, 13, 14, 15, 16] are mainly based on network parameters. These studies are limited because mobile devices can have an impact on QoE in terms of their display size, resolution and device make and model. Therefore, there is a need to investigate and propose models that take device parameters into consideration and use them as variables in the QoE assessment and prediction.

This thesis specifically seeks to carry out an investigation of the impact of mobile device preferences on QoE and proposes a QoE model that takes into account mobile devices display parameters together with video content and H.264 video coding parameters. This research is important because it forms the basis of the mobile device displays to be quantified in order to develop video quality measurement metrics.

The measurement metrics are important for service and network service providers as they can be used for effective monitoring and provisioning for acceptable video quality for.

The proposed model based on mobile devices display has the following potential usage,

- Measurements, prediction of end-to-end video quality control and optimization over Dynamic Adaptive Streaming over HTTP (DASH).
- prediction of initial encoded video before transmission over the IP networks and storage

1.2 Research Questions

This thesis has the following research questions to address,

- What is the impact of mobile device preference on the video QoE?

  This question has triggered an investigation on the impact of mobile device preferences on the video QoE. The investigation to answer this question was carried out by the use of subjective tests on varying mobile device manufacturers over the H.264 codec.

- What is the impact of mobile device displays on the video QoE?

  This question has lead to investigate the impact of mobile device displays on the video QoE. The investigation to answer this question was carried out by the use of subjective tests on several mobile devices with various display sizes and resolutions. A novel model based on display parameters was also developed to answer this question.

1.3 Aims and Objectives

The main aims of this research are (1) to investigate the impact of the mobile device preferences on the quality of video (2) to investigate the impact of mobile device displays on the video quality (3) develop and evaluate an objective reference free model based on the mobile devices displays for video QoE measurements and prediction. In order to achieve these aims, the following out,

- Investigate video coding and content parameters that impact video quality on the delivery of H.264 video and further identify parameters that can be used for video quality measurements and predictions.
- Investigate mobile device preferences and their impact on the end to end video quality delivery of H.264.
- Investigate mobile device display parameters and their impact on the video quality.
- Develop and evaluate a novel objective model based on the display sizes over varying mobiles devices for QoE measurements and predictions.
- Implement the proposed model under DASH standards as a proof of concept to demonstrate its usefulness in the control and optimization of video quality over video streaming services.

1.4 Research Contributions

The contributions of this thesis are outlined as follows,

- Mobile device preferences have an impact on the overall end-to-end video quality transmission. Videos and network service providers should consider this on their video quality control and optimizations schemes. This contribution has been published in [17].
- The mobile device displays have an impact on the video quality. The proposed non-intrusive new model which takes display sizes into consideration can be used by video service and network providers such as YouTube and Netflix for video quality control and optimization. This contribution [18].
- The dataset generated as a result of this research can be used by other researchers in academia and industry as a public dataset.

1.5 Thesis Outline

Chapter 2 discusses the literature review related to this research, techniques, tools and software used in this study. It also reviews video quality assessment such as subjective and objective methods, video coding techniques and mobile devices and quality of experience.

Research methodology which includes approaches to methods, source of data, sample population, data collection, data analysis, validation of results and experimentation tools are discussed in Chapter 3.

Chapter 4 presents the impact of mobile device preference on the video quality which includes experimental setup, discussions and results.

Chapter 5 discusses the impact of mobile device displays on the quality of the video. It also proposes and evaluates the new model that takes into consideration display parameters for measurements and prediction of video QoE.

Chapter 6 demonstrates the implementation and application of the proposed model through the Libdash open source library which is proposed and discussed in Chapter 5.

Chapter 7 concludes the thesis and proposes future work. This chapter includes contributions to knowledge and limitations of the current research.

The outline of this thesis is depicted in Figure 1.

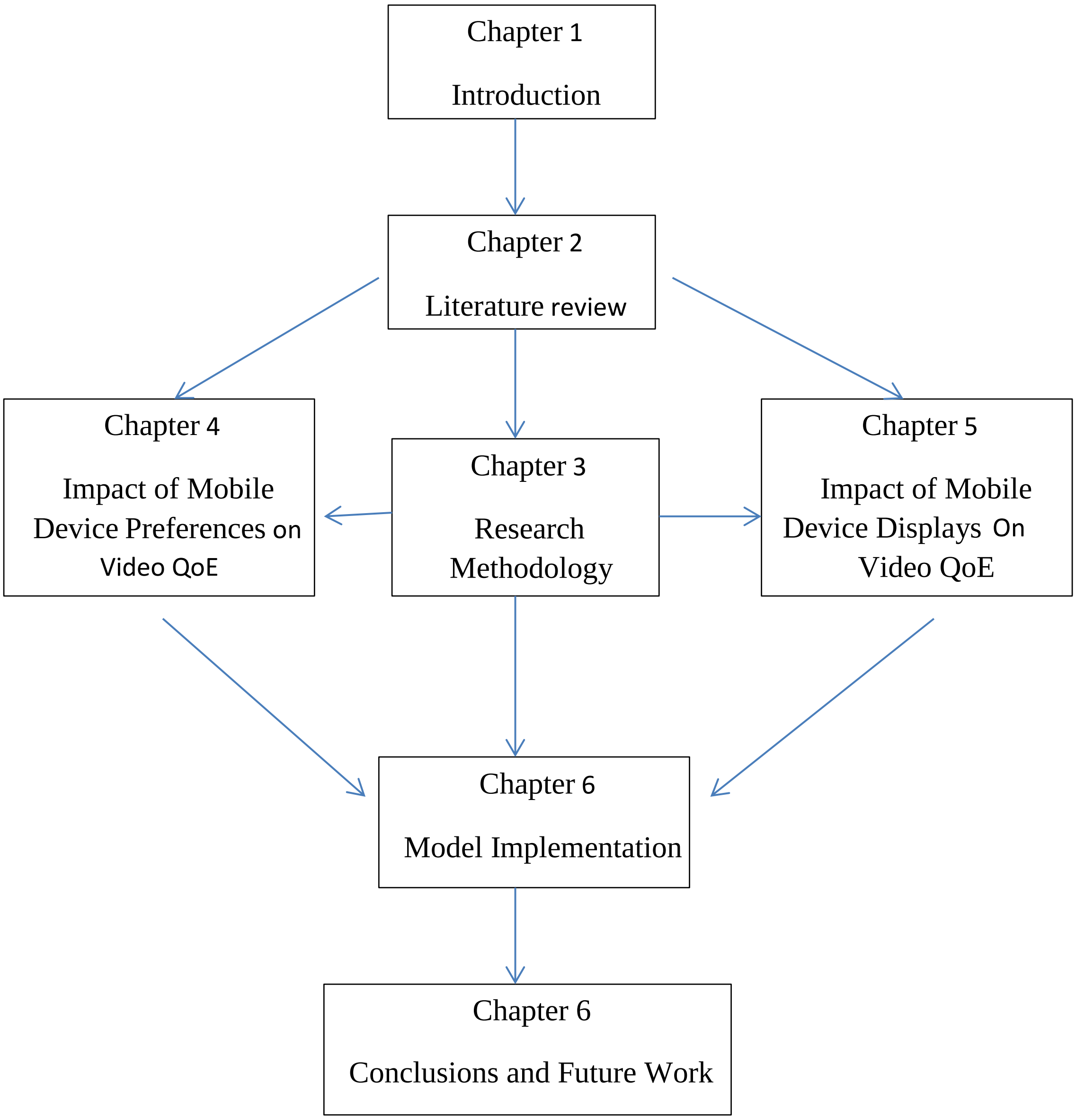


Figure 1.1: Thesis outline

CHAPTER II

2 Literature Review

2.1 Video Coding Techniques

Video compression techniques are deployed in raw video (uncompressed video) in order to reduce video size for storage and transmission while maintaining the perceived video quality for end users. For instance, an HD video of 720 x 1280 resolution with 30 frames per second of 1 hour long with full colour will need about 298GB of storage space and around 83MB/s transmission bandwidth. These requirements are not feasible to satisfy. Therefore, compression techniques are vital to reduce storage and transmission requirements.

Lossless and lossy techniques are categories of video coding techniques.

- Lossy:

  In this category, the high frequency components that are not apparent to human visual system are removed and therefore, the original data of the source video will not be fully retrievable.

- Lossless: In this category, the original data of the source video is fully retrievable.

The compression standards for transmission of video over IP networks are MPEG-4 part 2, H.264 also known as Advanced Video Coding (AVC) and MPEG-4 part 10. The latest one is HEVC.

In this research, the lossy compression has been used on H.264 codec.

2.1.1 Sampling and YUV Format

The first stage of video compression is sampling in temporal, spatial and colour domains. In temporal domain, sampling denotes the number of frames per second (fps). In the spatial domain, sampling refers to the number of pixels in each frame depending on the resolution. The colour space of Red, Blue and Green (RGB) and the number of bits required to represent each colour in each pixel is referred as sampling in the colour domain. Every pixel is composed of one luma (Y) component and two chroma components (U, V). Luma relates to the brightness of a pixel while U and V relate to the colour of the pixel.

The visual system of the human being is more sensitive to brightness (luma) than colour (chroma), therefore, the colour components are under sampled to reduce the transmission and storage of video sequences [3, 19]. Three parts ratio is used to represent chroma sub-sampling. For instance, YUV 4:2:0 denotes that for every 4 pixels of luma there is only one pixel of blue and one pixel for red. For YUV 4:2:0, there is half compression from its original YUV 4:4:4.

2.1.2 Video Compression Artefacts

In coding schemes that rely on motion compensation and on a block-based Discrete Cosine Transform (DCT) with the subsequent quantization of the coefficients, degradation in quality is typically caused by the quantization of the transform coefficients which controlled by the Quantization Parameter (QP). The QP controls the amount of spatial detail that is retained and the encoded bit rate [20, 21, 22, 23].

Increasing the QP value of an encoder leads to detail aggregation and drop in encoded video bit rate and quality. Even though, other factors such as motion prediction or decoding buffer size impact the visual quality of encoded video. These factors do not directly introduce visual degradation in quality. The following artefacts (though not exhaustive) are associated with video compression,

- Blurring:

  This is exhibited by the reduction of edge sharpness and a loss of spatial details. This is commonly due to coarse quantization, which suppresses high-frequency coefficients. Figure 2 illustrates an example of blurring.

Figure 2.1: Blurring example

- Blocking artefact:

  This refers to patterns of blocks in a compressed video. This is due to discontinuities at the boundaries of adjacent blocks in block-based coding schemes where individual blocks are independently quantized. Newer coding standards such as HEVC employ a deblocking filter followed by Sample Adaptive Offset (SAO) filter to reduce deblocking artefacts. Figure 3 shows an example of blocking.

Figure 2.2: Blocking example

- Jerkiness:

  This refers to object motion disorder caused by insufficient motion compensation or low temporal resolution. Typically, this happens as a result of poor performance of poor motion estimation.

- Colour bleeding:

  This is the smearing of colours between different strong areas of the chrominance. This happens because of the suppression of high-frequency coefficients chrominance components. Colour bleeding will typically extend over an entire Coding tree unit (CTU) or macroblock.

2.1.3 MPEG-4

ISO/IEC Moving Picture Experts Group defines MPEG4 video compression standard for encoding and decoding video storage and transmission over IP networks. Parts 2 and 10 out of 23 parts specify how the uncompressed video is compressed ready for transmission over the IP network. Part 10 defines Advanced Video Coding (AVC) most commonly known as H.264. In MPEG-4, a series of consecutive frames are divided in a Group of Pictures (GOP). Within a GOP, frames are of three different types, normally starting with I-frame followed by a number of P-frames and B-frames.

1. Intra coded frames (I-frame):

   I-frames are coded as a single picture. They do not have references to either P-frame or B-frame and are relatively bigger than other frame types.

2. Predictive coded frames (P-frame):

   P-frames are coded from previous I or P frames and are normally Smaller than the I-frames, but bigger than the B-frames.

3. Bi-directional coded frames (B-frame):

   B-frames are coded from preceding and succeeding I or P frames in the GOP structure. B-frames are normally the smallest.

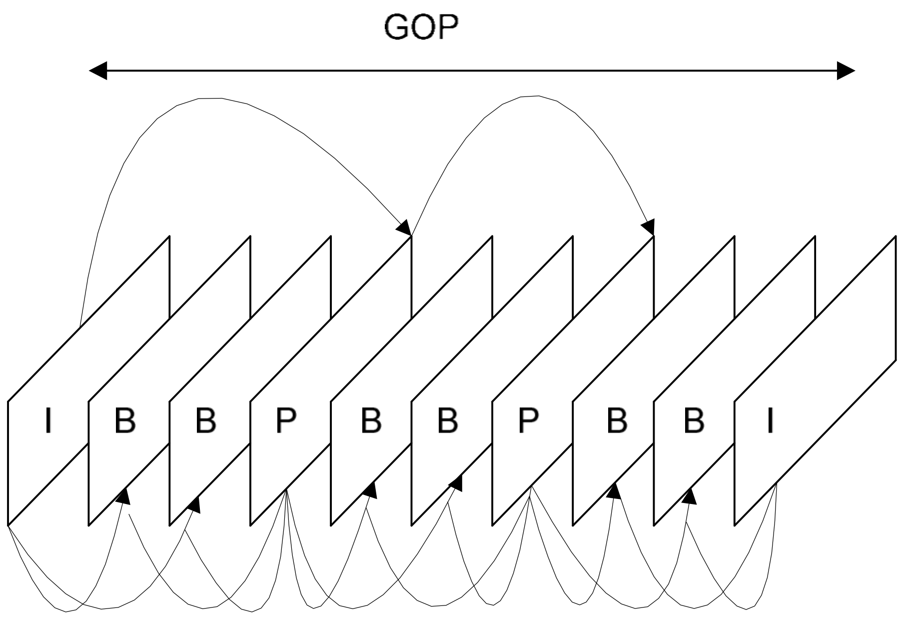


Figure 2.3: MPEG-4 GOP structure

A sample of MPEG-4 GOP structure is depicted in Figure 4. GOP(N,M) denoted N as the I to I frame distance and M is the I to P frame distance. Figure 4 can be presented as GOP(9,3), this means that one I-frame, 6 B frames and 2 P-frames are in the GOP. The second I-frame marks the start of the next GOP. Arrows in Figure 4 indicate that the B-frames and P-frames are decoded depending on the succeeding and preceding I-frames or
P-frames.

2.1.4 Advanced Video Coding H.264/AVC

There are seven profiles targeted to different classes of applications recommended for H.264. The baseline profile is used in this research because it is designed for mobile applications because of its low rate and complexity. The baseline profile only supports P-frames and I-frames.

H.264 has structured layers that consist of Video Coding Layer (VCL) and Network Abstraction Layer (NAL). VCL is structured to effectively represent the video content. NAL maps VCL data into different network transport layers such as RTP an RTSP and storage format such as ISO. Each frame in H.264 is partitioned into non-overlapping areas called macroblocks (MB). MBs are made up of 16×16 samples of the luma and 8×8 samples of each of the two chroma components. The MBs are arranged in slices representing subsets of MBs that can independently be decoded. The MBs can further be divided into smaller blocks of up to 4 x 4 pixels.

2.1.5 Slices:

A video frame can be divided into slices. Slices contain a sequence of macroblocks that are grouped together. This structure enables decoding and encoding of slices independently from other slices in a video frame. Slices are useful in a scenario whereby if a packet is lost during video streaming, only part of the frame will be lost and not the whole frame. The following types of slices are defined in H.264 standard:

1. Intra slices or I-slices: They contain MBs that are encoded using MBs in the same slice of the same frame. All the slices of the first frame are encoded as I-slices.

2. Predicted slices or P-slices: These are MBs that are encoded using MBs in a previously encoded and decoded frame. Some MBs in a P-slice may be encoded in intra mode.

3. Bi-directional or B-slices: These are predicted slices containing MBs encoded using MBs of the past and future I-slices or Pslices. The decoding order of B-slice is after the past and future I-slice or P-slice references.

2.2 Video Quality Assessment

There are several factors that impact video quality. These factors can be characterized as network, video coding and content parameters. The impact of these parameters on video quality as perceived by the user is referred to as the Quality of Experience (QoE). QoE is difficult to measure as it goes beyond the boundary of these parameters. According to the ITU-T P.10/G.100 [24], the definition of QoE is the overall acceptability of a service or an application as perceived subjectively by its users. This definition should include some other factors such as the end user device, service infrastructure, network bandwidth, user expectation and the environment where the end user is communicating [25, 26, 27, 28].

Subjective and objective methods are used to evaluate Video quality. Subjective quality is determined human perception of the quality of service as described by ITUT P.910 recommendations [29]. The Mean Opinion Score (MOS) is the QoE metric that has been widely used in telecommunication industry. Since subjective methods involve human beings, they are deemed to be the most reliable at measuring video quality. The disadvantages of subjective methods is that, can be expensive since volunteers have to be paid for participating in the tests. They can also be time consuming since participants have to be recruited. Therefore, to mitigate subjective method's drawbacks, there is a need for using objective method. Objective methods are used to predict video quality and hence, can lead to comparable results as those obtained by subjective methods. Results obtained through objective methods can be obtained by either intrusive or non-intrusive ways. The source of the video sequence is required (full or reduced reference) in the intrusive way to compare it with the impaired video sequence. Peak Signal-to-Noise Ratio (PSNR), Video Quality Monitor (VQM) [30] and Structure Similarity Index Measurement (SSIM) [31] are some of full reference quality metrics. Non-intrusive methods also known as reference free methods do not have access to the source video sequence and widely used because they are suitable for on-line quality measurements and prediction.

2.2.1 Subjective Video Quality Measurement Methods

Video Quality Experts Group (VQEG) and the International Telecommunication Union (ITU) have defined the subjective methods for video which involve a number of subjects, these subjects watch a video sequence in laboratory environment and the n score the video sequence after watching it. This score is defined as the Mean Opinion Score (MOS) and normally has a range from 1 to 5.

Subjective test methods have been clearly described in ITUT T.50011 and

ITUT Rec.P.910 (1999) [29] recommendations. Amongst other things, these recommendations propose viewing conditions, assessment procedures and testing materials. Subjective methods based on television applications are recommended in ITUT Rec. BT.50011 and multimedia applications are found in ITUT Rec. P.910. The main used subjective methods are:

2.2.1.1 Double Stimulus Impairment Scale (DSIS):

Pairs of the degraded video sequences along with video references are shown to the participants. The reference video sequences are supposed to be shown before degrade video sequences. The five point scale ranging from 1 to 5 are used for scoring as imperceptible, perceptible but not annoying, slightly annoying, annoying and very annoying.

2.2.1.2 Single Stimulus Methods:

The participants in this method are watching multiple separate video sequences. This method can have a single stimulus whereby video sequences are not repeated or a single stimulus with repetition whereby video sequences can be repeated several times. This method has three different scoring ways,

1. Adjectival: This is similar to DSIS, however, half scales are per-

mitted.

2. Numerical: This is comprised of 11-grade numerical scale.

3. Non-categorical: This is made up of a continuous scale with no numbers. For instance, 0 – 100.

2.2.1.3 Stimulus Comparison Method:

In this method is used when there are two matched displays in the test. The differences between the two scenes are scored by the either of the following two ways,

4. Adjectival: This is using a 7 grade scale marked from +3 to 3 which denotes "much better, better, slightly better, the same, slightly worse, worse, and much worse"
5. Non-categorical: There is continuous scale with no numbers as the one discussed before.

2.2.1.4 Single Stimulus Continuous Quality Evaluation (SSCQE):

The participants in this method watch a video sequence of about 2030 minutes without a reference of the original video sequence. The participants score by using a slider continuously at an instant of time from bad to excellent. This scale is an equivalent to numerical values between 0 and 100.

2.2.1.5 Double Stimulus Continuous Quality Scale (DSCQS):

In DSCQS method, the participants are required to watch several pairs of short videos sequences that can last for ten seconds. The video sequences are usually composed of sequence tests and reference video sequences. Both test and reference video sequences are randomly shown to participants twice. Participants are required to have no knowledge of the existence of reference video sequences. Participants are requested to score both test and the reference video sequences on a continuous quality scale ranging from bad to excellent equivalent to a numerical scale between 0 and 100. DSCQS method is mainly for the television signals and is explained in detail in the ITU-R Rec. T.500-13.

2.2.1.6 Absolute Category Rating (ACR) method:

The Absolute Category Rating (ACR) method is the widely used method telecommunication industry giving the Mean Opinion Score (MOS) as a measurement metric. While the Degradation Category Rating (DCR) method is deployed in some scenarios proving it uses the Degradation Mean Opinion Score (DMOS) as a measurement metric

The participants are required to watch video sequences without the original video sequences. After watching the video sequences, the participants are requested to give their opinion score. The quality evaluation by participants is based on the 5 points scale (Table 1). The Mean Opinion Score (MOS) will finally be computed as the average of the opinion scores from all participants' scores.

Table 2.1: ACR opinion scores

Category	Video quality
1	Bad
2	Poor

3	Fair
4	Good
5	Excellent

Figure 5 illustrates the ACR time pattern for the video sequences presentation. The voting time is required to be at most 10 seconds.

2.2.1.7 Degradation Category Rating (DCR) Method:

The DCR method is more suitable in the scenario whereby there is no much difference in the quality of the video sequences. The use of ACR method finds it hard to distinguish small differences in video sequence qualities such as 4 and 5.

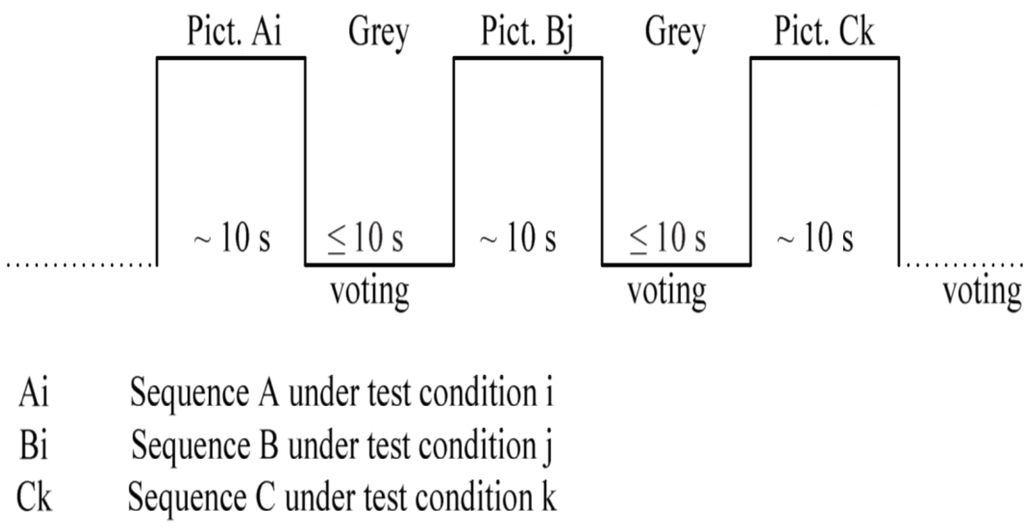


Figure 2.4: ACR method stimulus presentation

In this method, participants will compare the degraded video sequence against the original video sequence. It uses the Degradation Mean Opinion Score (DMOS) as a measurement metrics. The five scale points or the degradation levels are outlined in Table 2.

Table 2.2: DCR opinion scores

Category	Video quality
1	Very annoying
2	Annoying
3	Slightly annoying
4	Perceptible but not annoying
5	Imperceptible

Figure 6 illustrates the DCR time pattern for the video sequences presentation. The voting time is required to be at most 10 seconds.

Pair Comparison Method (PC):

In PC method, comparison of video test sequences are repeatedly conducted in pairwise manner. The video sequences in this method are put in all possible combinations and presented to participants in all possible orders. The participants give preferences between the pairs instead of grading the video sequences as in ACR or DCR. Figure 7 illustrates the PC time pattern for the video sequences presentation.

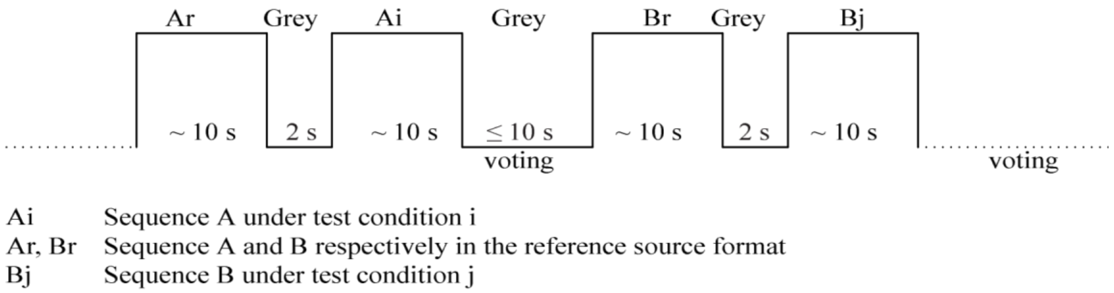


Figure 2.5: DCR method stimulus presentation

The voting time is required to be at most 10 seconds.

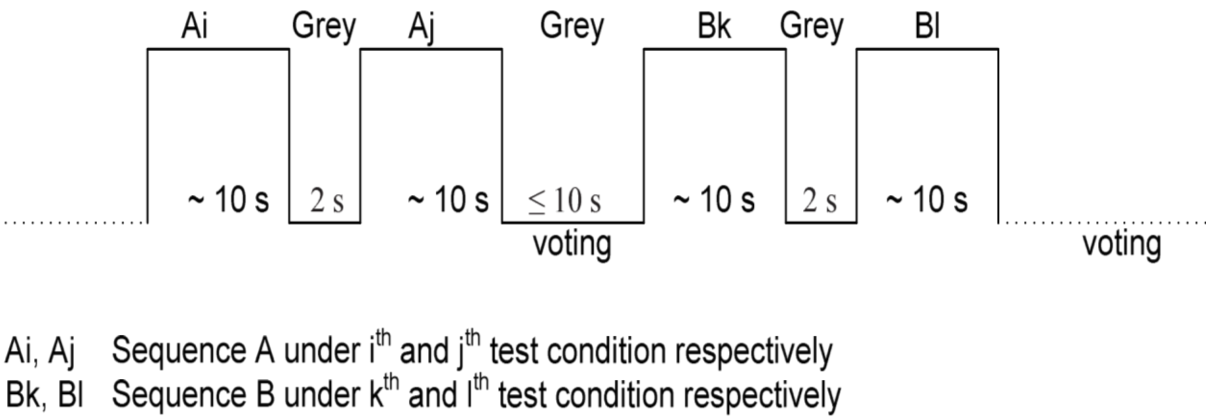


Figure 2.6: PC method stimulus presentation

2.2.2 Subjective Tests Design and Procedures

There are several aspects of subjective tests design and procedures that should be taken into account once data has been collected.

1. Scene characteristics:

   It is vital to have a choice of video test sequences that are a representation of the data that is to be collected. To avoid the boredom of participants in a test, video test sequences should be different. However, the video test sequences should be the same for all participants.

2. Replications:

   ITUT P.910 recommends that video test sequences should be replicated. Between two to four repetitions of the same video test sequences should be shown to participants in a test. This procedure is important because it will validate both subjects and the results they produce.

3. Presentation order:

   The order of presentation of all video test sequences are required to be randomized. The random order should be different to participants in the same test. After the test has been completed, the analysis of the results should follow the same presentation order. The reason behind this is that, participants might grade a fair sequence as good after watching a bad sequence.

4. Participants:

   According to ITU-T recommendations, the number of recommended participants are between 4 and 40. In general, a minimum of 15 nonexpert participants are recommended for better results. Participants should have no experience in video coding or multimedia communication.

5. Viewing conditions:

   The viewing conditions such as display and equipment for all participants in a test should be uniform.

6. Briefing:

   Participants in a test should be briefed about the aims and objectives of the test they are doing. The briefing should be in writing and clearly outlines what participants are supposed to do.

7. Training:

   The training session should be conducted prior to the test in order for the participants to familiarize themselves with the test.

8. Evaluation:

   Several evaluative scales can be considered depending on the subjective test method used. Grading scales or continuous scales can be used and should clearly be briefed to participants.

Subjective tests are made up of two phases, initial phase which has briefing and training procedures and test phase which should not last more than 30 minutes. The session should have breaks to avoid tiredness and fatigue.

2.2.3 Objective Video Quality Assessment

As it has been mentioned earlier that subjective tests are expensive and time consuming because a large number of participants and equipment should be involve for meaningful statistical results. Objective methods are easy to conduct and quick to setup and therefore, they are ideal for video quality evaluation. VQEG SG9 has been pioneering the research of objective methods in order to obtain objective results that are comparable with the subjective results for video quality evaluation. Objective methods can be classified as intrusive and non-intrusive video quality evaluation [71].

2.2.3.1 Intrusive Video Quality Assessment Methods

Intrusive methods are defined as full and reduce references.

1. Full Reference Method

The original video sequence is required in the full reference method in order to compare it with the degraded video sequence. The comparison is used as an indicator of the video quality. This method is not suitable for on-line and real-time scenario because it is impossible to get the reference video. PSNR and SSIM are the widely used in this method

- PSNR

  Peak Signal-to-Noise- Ration (PSNR) is widely used as an objective video quality metric. The objective value is found by comparing each video frame of the original sequence against the one in the degraded sequence. PSNR is defined as the Mean Squared Error (MSE) of two compared frames (1). The higher the PSNR the lower the MSE and hence, the higher the quality. PSNR values are given in (2).

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [(i, j - K(i, j)]^2 \tag{1}$$

$$PSNR = 20 \log \left(\frac{\max_i}{\sqrt{MSE}} \right) \tag{2}$$

  Where $\max_i$ denotes the maximum pixel value of the frame i.

- SSIM

  Structure Similarity Index Measurement (SSIM) provides the index Measure of the similarity between two frames. The measure between two frames x and y of size NxN is given by Equation (3).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\delta_{xy} + c_2)}{(\mu 2 + \mu 2y + c1)(\delta x2 + \delta 2y + c2)} \tag{3}$$

Where SSIM in Equation (3) is applied luma and the maximum value of 1 will denote that the video is of excellent quality. Structure Dissimilarity (DSSIM) is given in Equation (4).

$$DSSIM(x, y) = \frac{1}{1 - SSIM(x, y)} \tag{4}$$

2. Reduced Reference Method

In this method, only some features of the original video sequence are used. Video quality is predicted by using some features that are extracted from the original video sequence and some form the degraded video sequence.

2.2.3.2 Non-intrusive Video Quality Measurement

This method of quality measurement does not require accessing the original video sequence for video quality prediction, and is more suitable for real-time multimedia communication. Figure 8 depicts the block diagram of the non-intrusive method. This method can be categorized into signal, parameter and Hybrid based.

1. Signal based

Signal based measures the input signal of the video sequence such

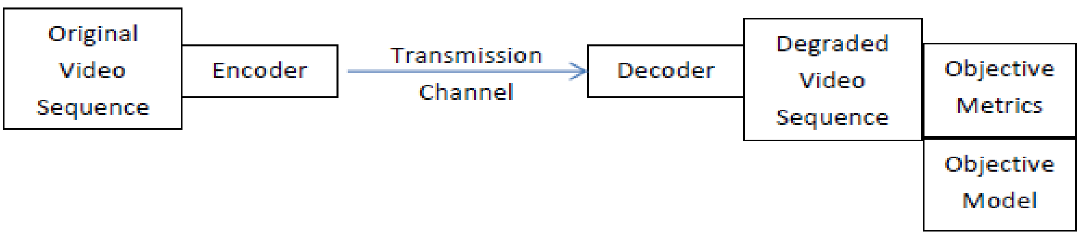


Figure 2.7: Non-intrusive measurement

Video content features.

2. Parameter based

This measurements gets parameters from the application layer such as codec and content type or from the network such as packet loss, delay and jitter or from device parameters such as display and luminance. Table 3 outlines some of the possible parameters that can be taken into consideration in quality measurements.

Table 2.3: DCR opinion scores

Category	Parameters
Network	type, bandwidth, delay, jitter, packet loss, RTT, loss burst size, protocols used, received signal strength, congestion levels
User and environment	location, temperature, heart rate, eye movement, social context, light, background noise, age, gender, education
Devices	screen size, design layout, resolution, general intuitiveness, buttons placement, input/output methods, appeal, usability
Application	Codec, content type

3. Hybrid based

This combines both signal and parameter based measurements.

This thesis proposes non-intrusive video quality prediction model using parameter based method.

2.2.4 Regression based methods

For parameter based measurements, a number of parameters will be the source of regression analysis and a model is fitted to get a goodness of the fit based on correlations coefficient and the Root Mean Squared Error (RMSE). Figure 9 illustrates the regression based method for video quality
Prediction.

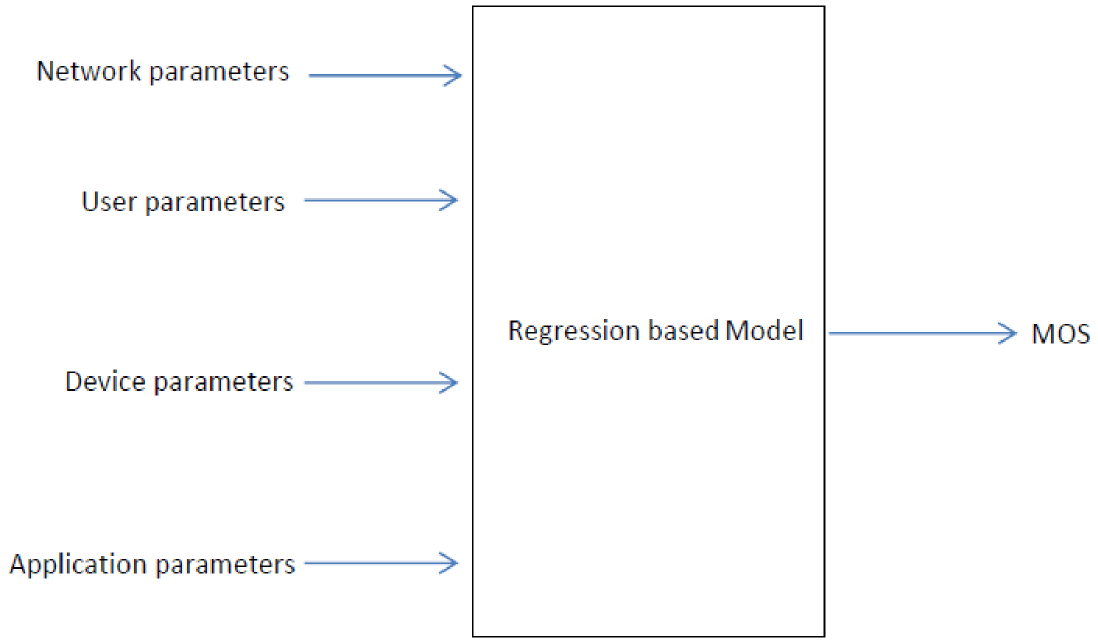


Figure 2.7: Regression based method for video prediction

2.3 Mobile devices and QoE

This research investigate the impact of mobile device preference and pixel density on the QoE of video services. Following is an overview of pervious work that consider these two factors.

2.3.1 Mobile Devices Preferences

The impact of video content preference have been investigated in several studies [32, 33, 34, 35], but none have investigated the impact of mobile device preference on the QoE over video service. With respect to mobile device specifications, there are some challenges for managing video traffic and displaying video to ensure an acceptable QoE for end users. This is due to the fact that the perceptual quality of video content depends on the properties of the display device and the viewing conditions [36].

Authors in [37, 38] aimed to assess the QoE in video streaming when the user is employing tablet devices that vary in terms of display size, resolution, hardware configuration and operating system. Authors conducted subjective video quality assessment on 216 video streams at two different bit rates (200 kbps and 400 kbps) with H.264/AVC. The videos were reproduced on Apple iPad 2 and the Samsung Galaxy Tab GTP1000.The values of the coefficients were very similar for both devices; and playout delay error was greater in the iPad 2 with respect to the Galaxy Tab. They also found that there is a strong correlation between the proposed quality index and the MOS for the iPad 2 and the Galaxy Tab.

The authors in [39] studied the user's preference for video content and how it can affect the video quality. The authors conducted subjective tests and proposed a video quality assessment method by taking the user preference for video content. From their experiment, they found that the values of QoE are highly correlated with the user's preference for video content type.

In [40], authors investigated the impact of image resolution, screen size, and screen resolution on user's perceived image quality. In the experiment, nine mobile phones and a quality monitor were used as test devices and allocated to evaluate the impact on perceived image quality (c.f., Table 4). 1360 data points were attained on the mobile phones. The authors also proposed an integrated assessment parameter to investigate the impact of mutual interaction between the device dependent image quality and image resolution. The assessment model was suggested to estimate the perceived image quality on different mobile devices. ANOVA was performed to check the significance of the influence of the screen size on the perceived image quality. The authors found that the improvement of screen resolution from 1080P to Quad HD did not have any impact. For the 1440P, 1080P, and 720P, the variation of perceived image quality for these images was similar to that for the 4K images.

To the best of our understanding, no prior study has considered the impact of device preferences on the Quality of Experience. In this thesis, we go beyond these and investigate the impact of device preferences on the Quality of Experience by using subjective tests.

2.3.2 Pixel Density

A systematic study regarding the impact of high resolution displays on the performance measures of subjective wellbeing such as the mood state and physical discomfort was conducted by Mayr et. al [41]. It was observed

Table 2.4: Parameters of display devices

Display device	Screen size (inch)	Resolution	Screen type
P1	4"	1136×640	IPS LCD
P2	4.3"	1280×720	IPS LCD
P3	4.9"	1920×1080	IPS LCD
P4	5.1"	1280×720	IPS LCD

P5	5.1"	1920×1080	AMOLED
P6	5.1"	2560×1440	AMOLED
P7	5.5"	1920×1080	TFT-LCD
P8	5.5"	2560×1440	SLCD
P9	5.7"	1920×1080	AMOLED
M1	30"	4096×2160	OLED

that reading speed and proofreading performance were not affected by the display resolution of 132 ppi or 264 ppi, and was further indicated that visual performance improves with pixel density up to around 130 ppi and 150 ppi, it then remains constant if pixel density is above 150 ppi. Although this study was based on visual reading speed and proofreading, similar investigation could be conducted for video streaming on different screens while varying with variable Pixel density.

An investigation based on subjective tests by Floris et., al [37] assessed the video quality during streaming over different tablet devices (Apple iPad 2 and Samsung Galaxy Tab GTP1000) while varying display sizes, resolutions, hardware configuration and operating systems. Based on their results, they observed that there is a high correlation between the proposed quality index and the MOS values for the two tablet devices in the experiment. Although display sizes and resolutions were part of the parameters in the assessment, pixel density could have been more appropriate to compare the two tablets in the experiment.

The impact of image resolution, screen size, and screen resolution on image quality was investigated by Zou et., al [40]. The image quality assessment model was proposed to estimate the perceived image quality on several mobile devices. Analysis of Variance (ANOVA) was applied to assess the significance of the impact of the screen size on the perceived image quality. It was concluded that increasing screen resolution from 1080P to Quad HD (QHD) did not affect perceived image quality. This study was based on the image quality which is different to video quality since the video has more complex parameters such as frame and bit rates. Therefore, a detailed study is important to investigate the impact of pixel density on the video quality.

ITU-T recommendation ITU-T P.1203.1 [42] has calculated the video quality degradation based on the quantization, upscaling and temporal degradation.

Upscaling video quality degradation was described as scaleFactor,

$$scaleFactor = Max(\frac{disRes}{codRes}, 1) \tag{5}$$

where, *disRes* is the device display resolution in pixels and *codRes* is the video encoding resolution in pixels. In this paper we argue that pixel density measured by the number of pixels per inch should be one of the parameters to be included in the *scaleFactor* because the same *disRes* for a 17 inch and a 5 inch screen would give different perceived video quality. It is only pixel density parameter that will fairly differentiate the perceived quality of the video depending on its *codRes*.

CHAPTER III

3 Research Methodology

This chapter starts by describing approaches to methods of data collection conducted in this study and comparing the advantages and disadvantages of using laboratory settings (controlled) and natural settings (uncontrolled or public environment). The sources of data and the reasons of sample population will be explained. The methods of data collection and analysis of the data will be outlined. The justification of the selected research methodology including reliability and validity issues will also be described.

3.1 Approached to Methods

Deciding the settings (laboratory or natural) in which to conduct subjective tests is crucial because it can determine the final results of the tests. Laboratory setting, also known as controlled setting, has an advantage because the sources of variability can be controlled [43]. In multimedia laboratory settings these variables can include lights, the size of device display, and the distance of viewing, for mobile devices, this can include luminance and the type of devices. In general, all tasks, time and physical space are controlled. By contrast, in natural setting, these variables are not under the control of the evaluator or experimenter.

The level of control depends on the nature of the research questions and the aims of the research. Real situations might require nature settings instead of laboratory settings.

In this study, based on the nature of video experiments, variables such as light, distance of viewing and time have to be controlled because a minor change in the lights or distance of viewing can affect the video quality.
Therefore, laboratory setting was selected to conduct the subjective tests.

3.2 Source of Data

Data was collected from participants of the subjective tests through questionnaires after watching each video sequence. Participants were asked to come to the laboratory of computer science at the National University in Khartoum, Sudan.

3.3 Sample Population

Undergraduate students of the National University in Khartoum, Sudan were recruited in this study as volunteers, no money was paid to them. This sample was selected because they are the ones who are more likely to watch videos streamed into their mobile devices form popular content providers such as YouTube and Facebook. So students feel comfortable to judge the quality of what they are watching in their daily lives.

Judgemental sampling [44] was used to select the sample of this research. In this sampling, the researcher selects of participants who were presumed to be representing the population in qualitative research. Participants were recruited through emails (Appendix B) sent in a mailing to the computer science students at the National University in Khartoum, Sudan.

3.4 Collection of Data

Following are the participants' tasks, the experimental setting and test procedures.

3.4.1 Tasks

Participants were given the following tasks that should be executed in order:

1. Sign the consent form.
2. Participate in a training session of no longer than 5 minutes
3. Watch video played back on a mobile which can last up to 60 seconds
4. Fill in the questionnaire by grading the video quality in the range of 1 (bad) to 5 (Excellent).
5. Answer any extra questions that might appear in the questionnaire
6. Have a break after watching a video before watching the next video

3.4.2 Experimental Setting

The subjective tests were conducted in the computer science laboratory at the National University in Khartoum, Sudan. This Laboratory has facilities such as personal computers, projectors and desks where participants can read and fill in consent forms and questionnaires. A timetable for participants was prepared and each participant knew his/her appointment. Particular days and times were set for each participant due to limited number of mobile phones availability.

Participants were supervised by the experimenter. The experimenter remained silent and unobtrusive during the course of subjective tests. Mobile devices were mounted on a stand in order to provide a more consistent viewing angle and viewing distance to all participants. The participants were allowed to adjust the viewing distance according to their preference, the display size and the video content quality. This is an ITU-T Recommendation P.913 [45]. In daily life the participants are not constrained when watching video content on their smartphone, however, they constrained when watching on TVs.

3.4.3 Test Procedures

The following procedures were followed in the subjective tests:

(a) On arrival, participants were briefed about the aims, objectives and the importance of the tests they were about to conduct (Appendix D).

(b) Participants were asked to read and sign the consent forms. Which amongst other things, participant's rights to withdraw are explicitly outlined.

(c) If a participant has signed the consent form, they will be given training to familiarize themselves with tools and tasks involved in the test.

(d) After the training session, participants were ready to undergo the real test by watching video sequences of not more than 60 seconds long.

(e) Once each video sequence has ended, participants fill in the MOS values score in the range of 1 (bad) to 5 (excellent). The voting was conducted by using paper ballots (Appendix C). Participants continue to watch video sequences in a particular device until all sequences are completed.

3.5 Analysis of Data

After the tests were completed, all the participants' forms are collected and the data are input into the Excel sheet for averaging to compute MOS for each video sequence for a particular mobile device.

To remove participant bias, the technique introduced by [46] was used with the following steps.

1. Estimate MOS for each video sequence,

$$MOS_j = \frac{1}{N_j}\sum_{i=1}^{N_j} p_{ij} \tag{6}$$

where, p_{ij} denotes observer rating of participant i for video sequence j and N_j is the total number of subjects that rated video sequence j. to estimate participant bias,

$$MOS_{\Delta_i} = \sum_{j=1}^{J_i}(p_{ij} - MOS_j) \tag{7}$$

where $MOS_{\delta i}$ estimates the overall shift between the i participant's scores and the true values (opinion bias) and J_i is the number of video sequences rated by participant i.

3. Compute the normalized ratings by removing participant bias from each rating,

$$\alpha_{ij} = p_i j - MOS_{\delta_i} \tag{8}$$

where α_{ij} is the normalized rating for paricipant i and video sequence j. MOS is then calculated normally. This normalization does not impact MOS,

$$MOS_j = \frac{1}{N_j} \sum_{i=1}^{N_j} \alpha_{ij} \quad — \quad —$$

For mobile device preferences subjective tests, the tests were done in two groups of 20 participants each from the National University in Khartoum, Sudan. Each participant in each group had to watch 7 short videos (arranged in random order according to ITU-T standard [47]) of not more than a minute each on each mobile device. The age of the participants varied from 18 to 23 years old, with average age of 22 years old. Participants had no knowledge of video signalling and processing. They were 55% men and 45% women; all had normal vision and clear understanding of the test.

For mobile device pixel density subjective tests, the tests involved 35 university students from the National University in Khartoum, Sudan. Each had to watch 10 short videos (rendered randomly as per ITU-T standard [47] recommendation). Each video length was 60 seconds. The age of the participants varied from 20 to 25 years old, with average age of 23 years old. All participants did not have any knowledge of video coding. All participants had normal eye vision, they were 22 men and 13 women.

3.6 Validation of Results

Analysis of Variance (ANOVA) is a statistical method that is used to test the differences of at least more than two means. The t-test (Independent Samples) has been used in this study to compare two groups of data, especially in the comparison of MOS values amongst group of participants ([48], [49]). In this research, values that are at most 0.05 would mean that there is statistical significant difference in the values under the experiments.

For the subjective tests regarding mobile device preferences, ANOVA was used to denote the statistical significant difference of MOS values with variation in mobile devices in this experiment.

For the subjective tests on pixel density, ANOVA was analyzed to demonstrate the statistical significant difference with variation in the pixel densities and MOS values.

Non-linear regression method was used to develop the proposed model based on pixel density. The regression based method of quality prediction deploys a number of parameters in the analysis and the proposed model is fitted depending on the correlation coefficient (R^2) which measures the goodness of the fit together with the root mean squared error (RMSE). If R^2 approaches 100% then the correlation coefficient is a very good fit. If RMSE approaches 0 then that means the error is very small on the proposed model.

3.7 Experimentation Tools

3.7.1 Mobile devices and OS

Android Operating System (OS) based smartphones were used in this study. Android is a mobile OS developed by Google. It is based on the Linux kernel and designed mainly for mobile devices such as smartphones and tablets. As of May 2017, Android OS has about 2 billion monthly active users, and it has the largest installed base of any OS [50].

For mobile device preference subjective tests, three mobile devices were used (Table 5).

Table 3.1: Properties of display devices

Device Model	size (in)	Resolution	Density
Sony Xperia Z	5.0	1080x1920	441
HTC One Max	5.9	1010x1920	373
Samsung A3	4.5	540x960	245

For pixel density subjective tests, eight Samsung mobile devices were used in the experiment (Table 6).

Table 3.2: Samsung Galaxy S series Smartphones

Mobile	Screen size	Screen Resolution	Pixel Density
Galaxy S	4.0	480x800	233.23
Galaxy S II	4.3	480x800	207.32
Galaxy S III	4.8	720x1280	305.96

Galaxy S 4	5.0	1080x1920	440.58
Galaxy S 5	5.1	1080x1920	431.94
Galaxy S 6	5.1	1440x2560	575.93
Galaxy S 7 edge	5.5	1440x2560	515.30
Galaxy S 8	5.8	1440x2960	506.42

3.7.2 Video Coding

FFmpeg is a free and open source software project that provides Application Programming Interfaces (APIs) and programs for handling multimedia data, such as coding and encoding video sequences. FFMPEG was used to encode YUV videos to MPEG-4 H.264/AVC with different bit rates and frame rates. All experiments to handle multimedia data under FFmpeg were conducted in Ubuntu 14.04.1 LTS trusty with Intel(R) Pentium(R) 4 CPU
2.80GHz, memory 8GB and 64 bits OS.

3.7.3 Video sequences

For mobile device preference subjective tests, two videos were used, FFMpeg was used to encode YUV videos to MPEG-4 H.264/AVC with the same frame rate (25 fps) and seven different bit rates 2000 Kbps, 2750 Kbps, 3000 Kbps, 3500 Kbps, 3750 Kbps, 4500 Kbps and 6000 Kbps. All videos were one minute long with 25 fps and 720p resolution.

For pixel density subjective tests, three videos were used, FFMpeg was used to encode YUV videos to MPEG-4 H.264/AVC at different bit rates and frame rates (Table 7).

Table 3.3: Video bit rate, resolution and frame rate

Sequence	Bit rate (kbps)	Resolution			Frame rate (fps)		
Vidyo	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
	2500	480p	720p	1080p	15	20	30
Johnny	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
——-	2500	480p	720p	1080p	15	20	30
BasketballDrive	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
——-	2500	480p	720p	1080p	15	20	30

3.7.4 DASH server and Client

Libdash [51] was used as a DASH server and client via the Apache 2 web server. The implementation of the proposed model based on pixel density enables content providers to optimally select video bit rate, frame rate and resolution depending on the mobile device pixel density and resolutions.

CHAPTER IV

The Impact of Mobile Device Preference on the Quality of Experience

This chapter present results obtained from the mobile device preference experiment. It also t-test and the distribution of mobile device preference as well as discussion.

Overview

Quality of Experience (QoE) is not only related to network impairments and video signal quality but it is a broad term that also covers other factors such as the end user devices, service infrastructure, network bandwidth, user expectation and the environment where the end user is communicating or consuming multimedia services. Recent studies have found that video content preference also has an impact on the QoE. To date, no research has reported the impact of device preference on the QoE. In this context, we investigate the impact of mobile devices preference on the QoE for video services. We evaluate the QoE of watching videos using mobile devices of different models, sizes, resolutions and densities. The experimental results based on subjective tests have shown that QoE is highly correlated with user preference on mobile devices.

4.1 Introduction

Mobile devices such as smart phones and tablets have witnessed exponential growth in number over the last decade. According to the Cisco Visual Networking Index (VNI) Global Mobile Data Traffic Forecast Update, more than half a billion mobile devices and Internet connections were added in 2015 [52]. In addition, video traffic accounted for 55% of the total mobile traffic in 2015 [52]. In this context, it is important to evaluate the Quality of Experience (QoE) with respect to different types of mobile devices in order for service providers and device vendors to increase revenue and avoid churn.

According to the ITU-T P.10/G.100 [24], the definition of QoE is the overall acceptability of a service or an application as perceived subjectively by its users. This definition should include some other factors such as the end user devices, service infrastructure, network bandwidth, user expectation and the environment where the end user is communicating [25]. The recommendations from the ITU-T for evaluating QoE with subjective tests are described in ITU-T Rec. P.800 [47]. The ITU-T recommends the PESQ model (ITU-T Recommendation, P.862) [55] and the E-model (Recommendation ITUT G.107) [56] to evaluate the quality of Voice over IP (VoIP). There are three classifications for QoE measurement; Machine Learning based techniques [57, 58, 59] liner and non-liner regression techniques [60, 61] and finally artificial intelligence techniques [57, 58, 59].

The impact of video content preference have been investigated in several studies [32, 33, 34, 32, 35], but none have investigated the impact of mobile device preference on the QoE over video service. With respect to mobile device specifications, there are some challenges for managing video traffic and displaying video to ensure an acceptable QoE for end users. This is due to the fact that the perceptual quality of video content depends on the properties of the display device and the viewing conditions [36].

In general, mobile devices impact QoE in terms of their display size, resolution and device make and model. In this study, the impact of mobile device size, resolution and device make and model is investigated to examine the impact of mobile devices preference on the QoE for video services.

Mobile devices with similar display size and resolution but from difference manufacturers were used in the experiment. From subjective tests conducted in Sudan at the National University in Khartoum, it was concluded that mobile device preference had an impact on QoE for video services.

This research is important for service providers such as YouTube, Facebook and FlexiNet because the results can be used as input to video streaming adaptation schemes in order to provide acceptable video quality under different network conditions.

The rest of this chapter is organized as follows, the experimental setup which includes participants, mobile devices and video sequences is discussed in Section 4.2. Section 4.3 presents the results obtained from the experiments which includes t-test and the distribution of the mobile devices preferences and discussion. The summary of this chapter is outlined in Section 4.4.

4.2 Experimental Setup

Objective video quality assessment and measurements tools such as PSNR and VQM are not computational intensive, however, they have drawbacks because they cannot capture the human visual perception such as the preferences of the devise used in conducting quality tests or even video content preferences. This is the reason why subjective tests are user because they can capture human visual perception of video quality.

In the subjective test, Absolute Category Rating (ACR) method was used whereby the participants were required to watch video sequences without the original video sequences. After watching the video sequences, the participants were requested to give their opinion score. The quality evaluation by participants was based on the 5 points scale. The Mean Opinion Score (MOS) were finally be computed as the average of the opinion scores from all participants' scores. The subjective video tests followed the ITU recommendations [29] and [62].

The first five minutes of the subjective test session was used to explain to participants the purpose and importance of the test. The grading scale and how to vote, the video sequences and timing of each video were also explained to participants. All participants were requested to provide written consent forms (see Appendix A). A training session was conducted to allow participants to familiarize themselves with all the procedures involved.

4.2.1 Participants

This study has been done in two groups of 20 participants each. Each participant in each group had to watch 7 short videos (arranged in a random order according to ITU-T standard [47]) of less than a minute each on each mobile device. The age of the participants varied from 18 to 23 years old, with average age of 22. All participants were students in computer sciences at the National University in Khartoum, Sudan. Participants had no knowledge of video signaling and processing. They were 55% men and 45% women; all had normal vision and clear understanding of the test.

4.2.2 Video Sequences

Two videos sequences, Big Buck Bunny and Elephant Dreams were selected (See Figure 7).

Figure 4.1: Encoded video sequences

Big Buck Bunny

Elephant Dreams

These videos are standard and are available with highest resolution. They were chosen because of their difference in content and complexity. Big Buck Bunny is medium movement video sequence, it is a short computer based animated comedy film where a large and adorable rabbit deals with three relatively tiny bullies including a flying and annoying squirrel, all of the bullies are adamant to crush the rabbit happiness. While the Elephant dreams video sequence is a fast movement video sequence. It is an animated science fiction film. The two main characters are on a journey in the folds of a giant machine, exploring the twisted and dark complex of wires and gears. FFMPEG version 3.3.3 under Linux was used to encode YUV videos to MPEG-4 H.264/AVC with the same frame rate (25 fps) and seven different bit rates 2000 Kbps, 2750 Kbps, 3000 Kbps, 3500

Kbps, 3750 Kbps, 4500 Kbps and 6000 Kbps. All videos were one minute long with 25 fps and 720p resolution. The selected range of these bit rates are recommended by YouTube for the 720p quality level for 24, 25 and 30 frame rates [34].

The participants in each group had to randomly watch video sequences, one group watched Big Buck Bunny and another group watched Elephant Dreams on three different mobile phones.

4.2.3 Mobile Devices

Participants were invited to watch video sequences and evaluate the quality of the video on three mobile devices. This experiment included three mobile devices, Sony Xperia Z, HTC One Max and Samsung Galaxy A3. Screen specifications for these devices are given in Table 8. Each mobile device was formatted and was installed with Android 5.0 (Lollipop) and Android Media Player to ensure that cache and memory were clear before playing the videos. The same media player was used in all tests and mobiles were set up on flight mode to avoid any interruptions during the experiment.

The brightness was adjusted on the same level on all mobile devices, the sleep and portrait mode were disabled. Participants were asked to fill in the Mean Opinion Score (MOS) based on a discrete level as stated in ITU-R quality ratings (from bad (1) to excellent (5)). The experiment used NonReference (NR) method because it is practical in multimedia communication.

Table 4.1: Properties of display devices

Device Model	size (in)	Resolution	Density
Sony Xperia Z	5.0	1080x1920	441

HTC One Max	5.9	1010x1920	373
Samsung A3	4.5	540x960	245

4.3 Results and Discussion

In this study the variables display density and resolutions were used to evaluate the quality of experience using MOS values. Table 9 and Figure 10 show the MOS scores for seven sequences of Big Buck Bunny video for the first group of participants.

Table 4.2: MOS values for Big Buck Bunny

Bit rates (Kbps)	Sony	HTC	A3
2000	2.56	2.91	3.91
2750	2.94	3.32	4.72
3000	3.08	3.21	4.79
3500	3.58	4.14	4.52
3750	4.01	4.58	4.64
4000	4.00	4.35	4.54
6000	4.10	4.22	4.39

The results show that MOS values for Galaxy A3 mobile phone are better than the rest of the devices at low bit rates between 2000 Kbps and 3740 Kbps and slightly higher at higher bit rates above 3750 Kbps. Although the Galaxy A3 mobile phone specifications are inferior compared to other mobile phones in the experiment, but participants gave higher MOS values. This is contrary to the expectations, given the screen densities of each mobile phone, it was expected that MOS values would be higher for Sony Xperia Z with 441 screen density, followed by HTC One Max and Samsung Galaxy A3 mobile phones.

To eliminate the influence of video content preference on the QoE, participants with no preference to any of the videos in the experiment were cho-

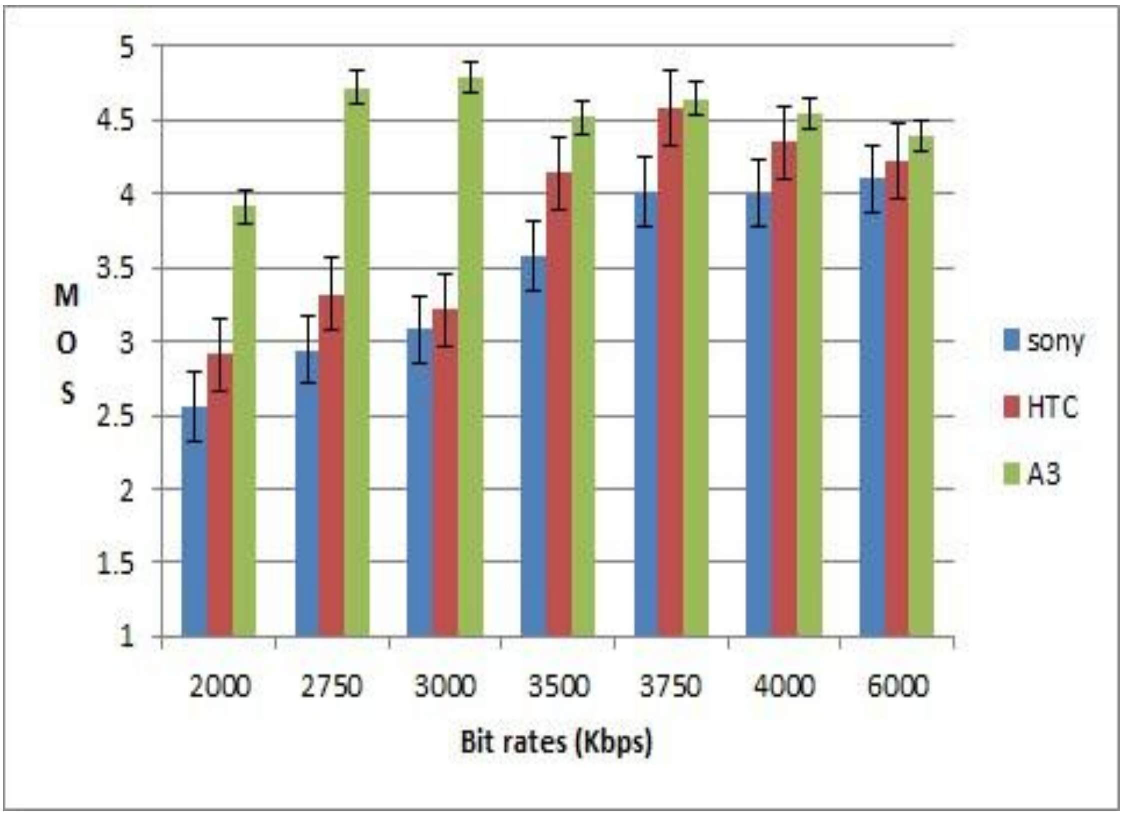


Figure 4.2: MOS values for different mobile phones-Group 1

sen. They were asked one question, "Amongst the three devices used in the experiment, which device you liked the most?" The response was 33% for Samsung Galaxy A3, 45% for HTC and 22% for Sony Xperia Z (c.f., Figure 11).

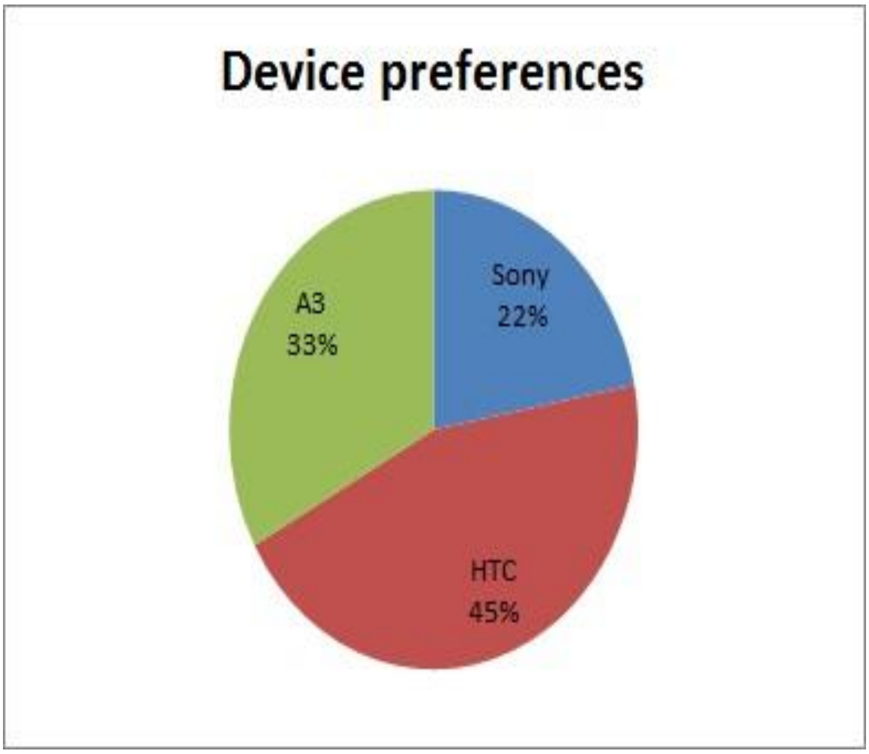


Figure 4.3: Device preferences distribution-Group 1

To support the significance of these results, the P-value from Analysis of Variance (ANOVA) was 0.00056≤0.05 (Table 10) which denotes that the MOS values show the statistically significant difference with variation between mobile phones at each bit rate of Big Buck Bunny video sequence in the experiment. This means that a null hypothesis is rejected because 0.00056≤0.05, 0.05 is a significance threshold level (α). A null hypothesis is a hypothesis that shows there is no statistical significance of the MOS values between the mobile devices in the experiment.

Table 4.3: P-value for devices with all bit rates

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	3.88	2	1.94	14.89	0.00056	3.89
Within devices	3.77	6	0.63	4.83		

Total	9.21	8				

In Table 10 SS in the first row is the sum of the squares between mobile devices sample MOS means. While in the second row SS is the sum of squares MOS means within the mobile devices. MS is the weighted average of the mobile devices sample variances. df is the degree of freedom and F-value > F-critical.

Table 4.4: P-value for devices at 2000 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	19.75	2	9.87	55.69	3.92E-14	3.16
Within devices	10.11	57	0.18			
Total	29.85	59				

Results are further broken down into individual bit rates in order to support the significance of these results, the P-value in Table 11 denotes that the MOS values show the significant difference with variation between mobile phones at 2000 Kbps of Big Buck Bunny video sequence.

P-value in Table 12 shows that the MOS values have significant difference with variation between mobile phones at 2750 Kbps of Big Buck Bunny video sequence. The P-value for this bit rate is smaller (3.92E-14) compared to the significance level of 0.05 and therefore, the null hypothesis is rejected.

Table 4.5: P-value for devices at 2750 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	35.15	2	17.58	164.45	2.13E-24	3.16
Within devices	6.09	57	0.11			
Total	41.24	59				

P-value in Table 13 demonstrates that the MOS values have significant difference with variation between mobile phones at 3000 Kbps of Big Buck Bunny video sequence.

Table 4.6: P-value for devices and 3000 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	36.25	2	18.12	257.76	2.79E-29	3.16
Within devices	4.01	57	0.07			
Total	40.26	59				

P-value in Table 14 demonstrates that the MOS values have significant difference with variation between mobile phones at 3500 Kbps of Big Buck Bunny video sequence.

Table 4.7: P-value for devices at 3500 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	8.94	2	4.47	58.57	1.50E-14	3.16
Within devices	4.35	57	0.076			
Total	13.30	59				

P-value in Table 15 demonstrates that the MOS values have significant difference with variation between mobile phones at 3750 Kbps of Big Buck Bunny video sequence.

Table 4.8: P-value for devices at 3750 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	4.84	2	2.44	23.21	4.23-08	3.16
Within devices	5.94	57	0.10			
Total	10.77	59				

P-value in Table 16 demonstrates that the MOS values have significant difference with variation between mobile phones at 4500 Kbps of Big Buck Bunny video sequence.

Table 4.9: P-value for devices at 4500 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	3.00	2	1.50	15.01	5.79E-06	3.15
Within devices	5.70	57	0.10			
Total	8.70	59				

P-value in Table 17 demonstrates that the MOS values have significant difference with variation between mobile phones at 6000 Kbps of Big Buck Bunny video sequence.

Table 4.10: P-value for devices at 6000 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between Groups	0.85	2	0.42	6.89	0.044	3.16
Within Groups	8.37	57	0.15			
Total	9.22	59				

Table 18 and Figure 12 illustrate the MOS scores for seven sequences of Elephant Dream video for the second group of participants.

Table 4.11: MOS values for Elephant Dream

Bit rates	sony	HTC	A3
2000	3.01	3.56	3.21
2750	3.50	4.31	3.82
3000	4.09	4.31	4.11
3500	4.15	4.52	4.25
3750	4.19	4.25	4.19
4000	4.10	4.19	4.15
6000	4.02	4.12	4.105

The results show that MOS values for HTC One Max mobile phone are better than Sony Xperia Z and Samsung Galaxy A3 devices at bit rates between 2000 Kbps and 3740 Kbps and slightly higher at higher bit rates more than 3750 Kbps. It was expected that Sony Xperia Z would have better MOS values because it has higher screen density than the rest, but its MOS values were the lowest. Although Samsung Galaxy A3 mobile phone specifications

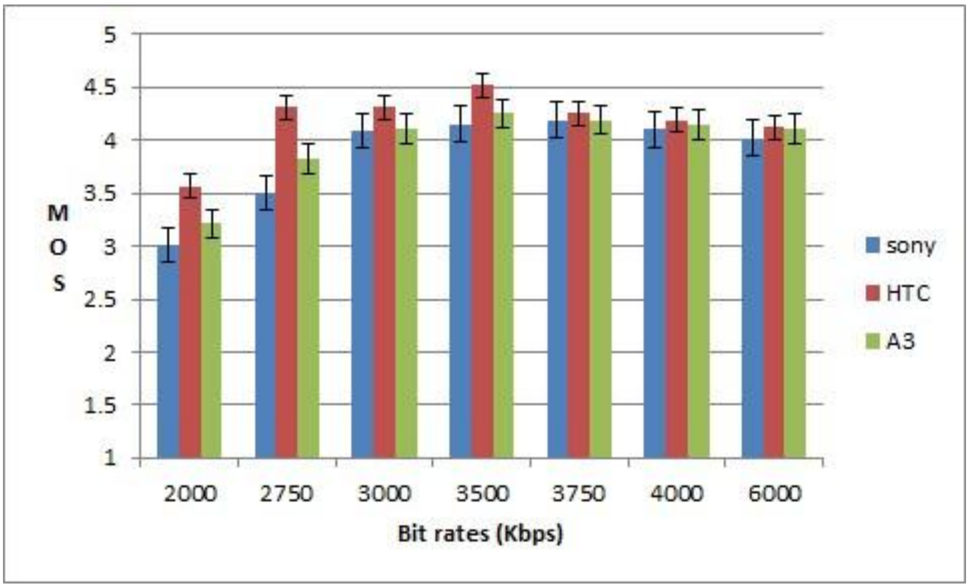


Figure 4.4: MOS values for different mobile phones-Group 2 are inferior compared to other mobiles phones in the experiment, its MOS values came second. It was observed that these results are correlated to the participants' preferences to three mobile phones in the experiment as in the first group.

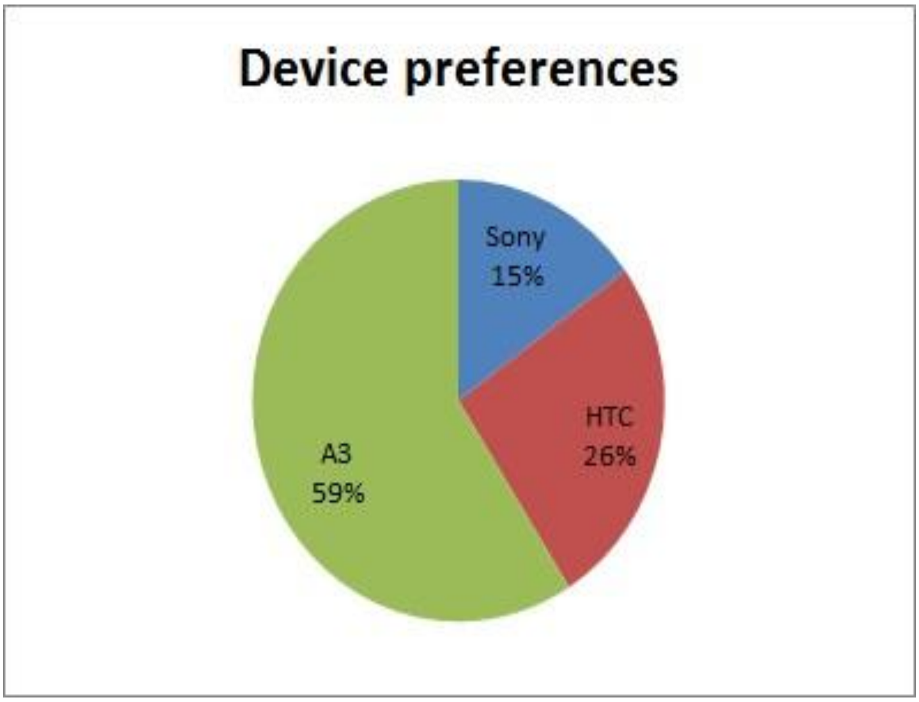


Figure 4.5: Device preferences distribution-Group 2

Similarly to Group 1 participants, participants in this second group were also asked one question, "Amongst the three devices used in the experiment, which device you liked the most?" The response was 59% for Samsung Galaxy A3, 26% for HTC One Max and 15% for Sony Xperia Z (c.f., Figure 13).

To illustrate the statistical significance of these results, P-value from Analysis of Variance (ANOVA) was $0.032 \leq 0.05$ (Table 19) which denotes that the MOS values for this second group show the statistically significant difference with variation in the mobile phones for Elephant dream video sequence.

Table 4.11: P-value for devices with all bit rates

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	2.50	6	0.42	11.21	0.00026	3.00
Within devices	0.345	2	0.17	4.65	0.032	3.89
Total	3.29	20				

In order to support the significance of these results, the P-value in Table 20 denotes that the MOS values show the significant difference with variation between mobile phones at 2000 Kbps of Elephant Dream video sequence.

Table 4.12: P-value for devices at 2000 Kbps

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	0.85	2	0.42	12.89	0.03	3.15
Within devices	8.37	57	0.15			
Total	9.22	59				

The videos quality based on the PSNR values were also provided for both video sequences for each bit rates selected in the test (Figure 14). The Big Buck Bunny as slow movement video demonstrated higher quality than Elephant Dreams video sequence as fast movement video sequence. This is because fast movement videos have more complex features such as motions compared to medium and slow

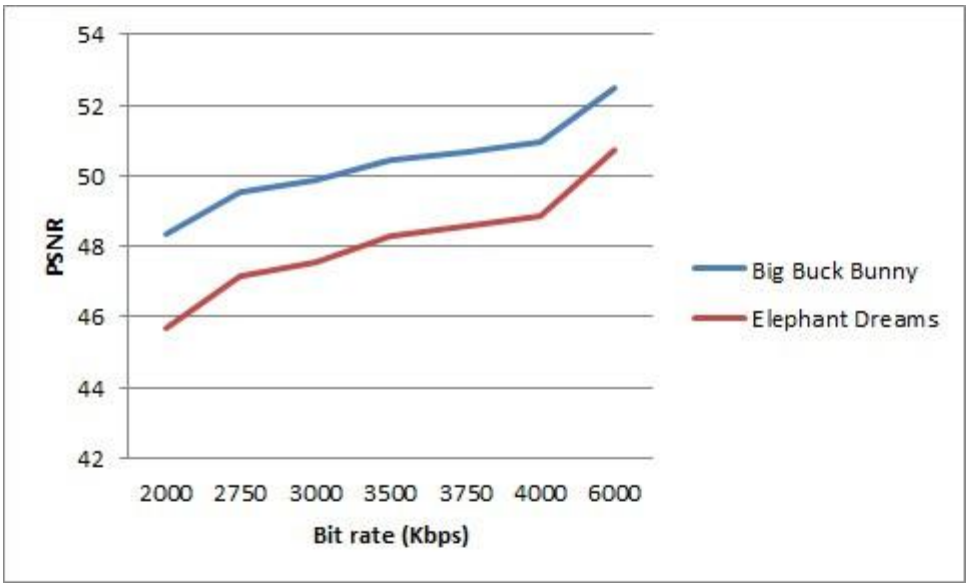


Figure 4.6: PSNR values at each bit rate

movement video sequences. Therefore, at the same bit rates, slow or medium video sequence will have better quality than fast video sequences [7].

4.4 Summary

This chapter has investigated the impact of device preferences on the Quality of Experience for video services. Results based on subjective tests have shown that there is a correlation between device preferences and QoE. The subjective test conducted had followed ITU-T recommendations [62, 29]. Apart from MOS values, PSNR values were also provided for both video sequences. Medium movement video sequenced (Big Buck Bunny) demonstrated to have higher quality at the same bit rates compared to fast movement video sequences (Elephant Dreams). Based on these results, there is a need to consider human factors such as device and content preferences when objective models are proposed by video service providers such as YouTube and Netflix. This investigation is important in the multimedia services networking and communications because it can be used in the video quality control and optimization in order not to over provision the video quality for limited bandwidth and storage usage.

CHAPTER V

Mobile Device Pixel Density and Video QoE

This chapter explain the experiment setup and demonstrates results. It also proposes and evaluate a scale factor density based method.

5 Mobile Device Pixel Density and Video QoE

Overview

Device screen size, design layout and resolution are some of the device context parameters which impact video quality on mobile devices. In this thesis, through subjective tests, we evaluate the impact of mobile device pixel density on the video quality. We then, through the subjective test, propose a video quality model that takes pixel density into consideration. The proposed model was evaluated, and the preliminary results have shown that the proposed model performs well regarding correlation coefficient. The model has potential use for multimedia service providers such as YouTube and Netflix in areas of quality control and optimization of network and multimedia services.

5.1 Introduction

Smart phones and other portable devices such as tablets and game consoles have experienced exponential growth in number in recent years. According to the Cisco Mobile Visual Networking Index (VNI) Forecast (20162021) released in February 2017, there will be 5.5 billion mobile device users, 12 billion mobile-ready devices and connections and 587 Exabytes annual rate of Mobile IP traffic [63]. The forecast also predicts that just over 75% of the video traffic of the world will be mobile data traffic [63]. These figures put enormous pressure to multimedia service providers such as Netflix and YouTube to provide acceptable video quality to its customers.

According to OpenSignal report which tracks sensors and networks, in 2015, there were 24,093 Android based mobile devices. These devices come with different shapes and sizes, with diverse performances, display sizes and display resolution [64]. This level of fragmentation put another pressure to multimedia service providers on how to accurately provide acceptable video quality to their customers.

To this end, it is vital to conduct thorough evaluation of the Quality of Experience (QoE) in order to satisfy customers with this level of mobile device fragmentation. Customers satisfaction will make them happy with the service they consume and hence increase revenue for service and network providers.

QoE represents the customer's experience measure of consumed services such as voice and video calling, IPTV and gaming services [53]. Three groups can be used to classify parameters that impact QoE [54],

- quality of multimedia content where they are produced
- network parameters, also known as Quality of Service (QoS)
- human perception, which includes among other things emotions, the degree of annoyance and past experience

QoE can be quantified by using either subjective assessments or objective tools. Most of these techniques are based on available on freely tools and methodologies [65]. QoE is normally measured by the Mean Opinion Score (MOS) metric. The MOS metric has a 1 to 5 points scale representing five terms of multimedia quality (bad, poor, fair, good and excellent) [29].

Although subjective assessments can give accurate MOS values, they require stringent environment, and they are expensive and time consuming if they are to be repeated frequently. They are also impossible to use in real time multimedia communication. Therefore, objective modelling is preferred to predict multimedia quality based on subjective tests.

Pixels per inch (ppi) is used to measure pixels density (PD). Pixel density has an impact on the visual quality, the higher the pixel density the higher the quality of the video [41]. Pixel density is computed by using the display diagonal size and the display resolution (width × height) in pixels. The display resolution in pixels (d_p) is given by,

$$dp = \sqrt{w_p^2 + h_p^2} \qquad (10)$$

where, w_p and h_p are width and high resolutions, respectively. The pixel density (pd) is then calculated as,

$$pd = \frac{d_p}{d_i} \qquad (11)$$

Where, d_i is the display diagonal size in inches. The diagonal size in inches is the one normally advertised as the screen/display size.

Display technology has experienced tremendous increase in Pixel density over the last three decades. In 1980s, the common pixel densities were between 60 and 90 ppi, while in 1990s, it was up to 120 ppi. By now we have pixel density of more than 260 ppi [41]. For a large screen of lets say 46 inch HDTV with a 1920x1080 resolution, the pixel density is only 52 ppi. This can be perceived to be lower than common smart phone screens. However, such big screens are not viewed at the same distance as mobile devices. When watching big screens from the comfort of your an arm chair, individual pixels are impossible to distinguish, making pixel density more important for small displays than big displays.

Table 21 depicts some of the mobile phone Pixel Density.

Table 5.1: Pixel density of some mobile phones

Mobile	Screen size (d_i)	Screen Resolution ($w_d \times h_d$)	Pixel Density (ppi)
Sony Xperia Z5 Premium	5.5	3840×2160	801.06
Samsung Galaxy S6	5.1	2560×1440	575.92
Samsung Galaxy S8	5.8	2960×1440	567.53
LG G6	5.7	2880×1440	564.90
Motorola Droid Turbo	5.2	2560×1440	564.85
Samsung Galaxy Note 4	5.5	2960×1440	515.30
BlackBerry KeyONE	4.5	1620×1080	432.67
Apple iPhone 7 Plus	5.5	1920×1080	400.53

This study is vital for multimedia service providers such as YouTube and NetFlix because the proposed results can be applied in their video streaming quality adaptation algorithms in order to provide acceptable video quality under different level of pixel density fragmentations. Knowing the pixel density of mobile devices is vital for service providers not to over provision video quality and therefore, saving limited bandwidth and storage.

The rest of this chapter is organized as follows, the experimental setup is presented in Section 5.2. Section 5.3 will outline the results and discussion. The proposed model based on scale factor pixel density is derived in Section 5.4 and evaluated in Section 5.5. The summary is presented in Section 5.6.

5.2 Experimental Setup

Absolute Category Rating (ACR) method was used in this study whereby the participants were required to watch video sequences without the original video sequences. After watching the video sequences, the participants were requested to give their opinion score. The quality evaluation by participants was based on the 5 points scale. The Mean Opinion Score (MOS) were

finally be computed as the average of the opinion scores from all participants scores. The subjective video tests followed the ITU recommendations [29] and [62].

5.2.1 Participants

This study involved 35 university students from the National University in Khartoum, Sudan. Each had to watch 10 short videos (rendered randomly as per ITU-T standard recommendation [47]). Each video length was 60 seconds. The age of the participants varied from 20 to 25 years old, with average age of 23 years old. All participants did not have any knowledge of video coding. All participants had normal eye vision, they were 22 men and 13 women.

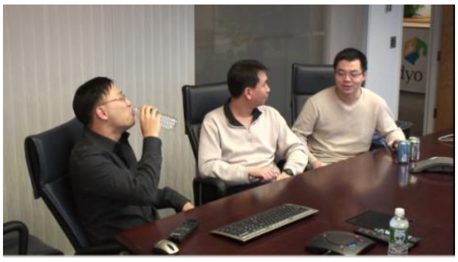
The first five minutes of the subjective test session was used to explain to participants the purpose and importance of the test. The grading scale and how to vote, the video sequences and timing of each video were also explained to participants. All participants were requested to provide written consent forms (see Appendix A). A training session was conducted to allow participants to familiarize themselves with all the procedures involved.

5.2.2 Video Sequences

H264/AVC encoded video sequences were used in this experiment. The encoding and decoding was done by FFMpeg software. The videos were categorized into slow (Johnny sequence), medium (Vidyo sequence) and fast movements (BasketballDrive sequence). Figure 22 illustrates the thumbnails of video sequences used in the study. Video sequences bit rates and frame rates are listed in Table 7 Each video sequence was 1 minute long.

Figure 22: Encoded video sequences

Johnny



Vidyo

BasketballDrive

Johnny sequence is a slow movement video because the news reader is seated, the upper part of the body moves slowly while he is talking. The Vidyo sequence is a medium movement because it shows men chatting while seated and the upper parts of their bodies are gently moving. The BasketballDrive video sequence is categorized as fast movement videos due to rapid movements of the basketball players.

Table 5.2: video sequences

Sequence	Bit rate (Kbps)	Resolution			Frame Rate(fps)		
Vidyo	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
	2500	480p	720p	1080p	15	20	30
Johnny	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
——-	2500	480p	720p	1080p	15	20	30
Basketball	500	480p	720p	1080p	15	20	30
	1000	480p	720p	1080p	15	20	30
——-	1500	480p	720p	1080p	15	20	30
——-	2000	480p	720p	1080p	15	20	30
——-	2500	480p	720p	1080p	15	20	30

5.2.3 Mobile Devices

Samsung Galaxy S series smartphones were used in this study, they are all based on Android operating system. They are all based on the same display technology (Super AMOLED) with varying display sizes and resolutions. Samsung Galaxy S series smartphones used in this experiment are listed in Table 23.

All smartphones were formatted to make sure that cache and memory were freed before playing the video sequences. In all the smartphones, flight mode was set up in order to avoid any interference that can be caused by cellular and wireless signals during the experiment.

Table 5.3: Samsung Galaxy S series Smartphones

Mobile	Screen size (d_i)	Screen Resolution ($w_d \times h_d$)	Pixel Density (ppi)
Galaxy S	4.0	480x800	233.23
Galaxy S II	4.3	480x800	207.32
Galaxy S III	4.8	720x1280	305.96
Galaxy S 4	5.0	1080x1920	440.58
Galaxy S 5	5.1	1080x1920	431.94
Galaxy S 6	5.1	1440x2560	575.93
Galaxy S 7 edge	5.5	1440x2560	515.30
Galaxy S 8	5.8	1440x2960	506.42

In all the smartphones, the brightness was set on the same level (full), the sleep and portrait mode were disabled. All participants were requested to choose the Mean Opinion Score as stipulated in ITU-R quality ratings from 1 to 5. The study deployed Non-Reference (NR) method which is ideal for end-to-end multimedia communication.

5.3 Results and Discussions

The bit rate, frame rate, mobile display size, mobile display resolution and video codec resolution where used as variables in this study. Smartphones from the manufacture (Samsung) were used in the study in order to keep the same display technology for reliable results. For Galaxy S smartphone with 480X800 display resolution and and 4" display size, the relationship between frame rate and quality of video in terms of MOS values is evaluated. This is done by setting the bit rate and resolution constant at each frame rate it can be seen that the higher the frame rate the better the quality (c.f., Figure 5.1). The same trend is seen for other bit rates such as for 1500 Kbps (c.f., Figure 5.2).

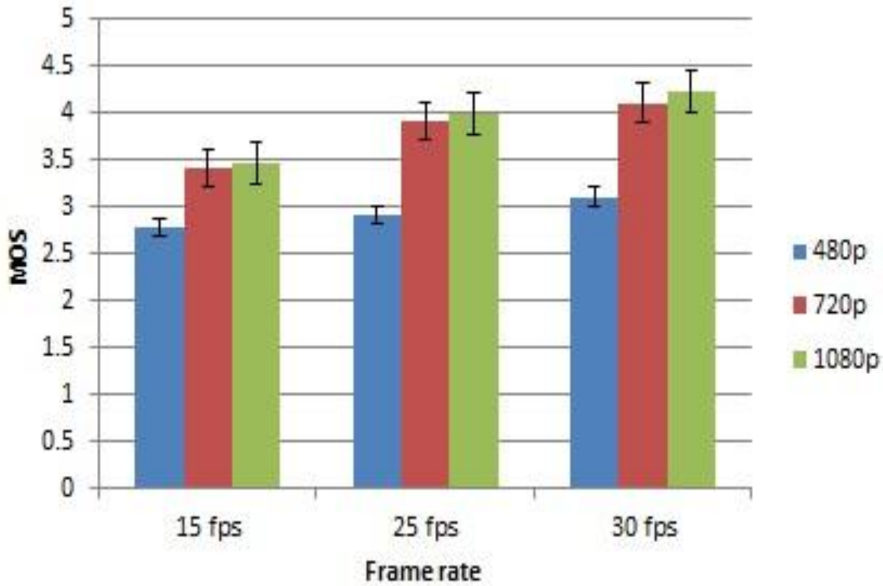


Figure 5.1: MOS values for Galaxy S at 500 Kbps for Johnny sequence

To demonstrate the impact of pixel density on the video quality, the Galaxy S III smart phone with 720X1280 display resolution and and 4.8" display size is used. The relationship between frame rate and MOS values is evaluated, it can be observed that the same trend occurs as with the Galaxy S.

Figure 5.2 and Figure 5.3 illustrate this trend for Johnny slow movement se-

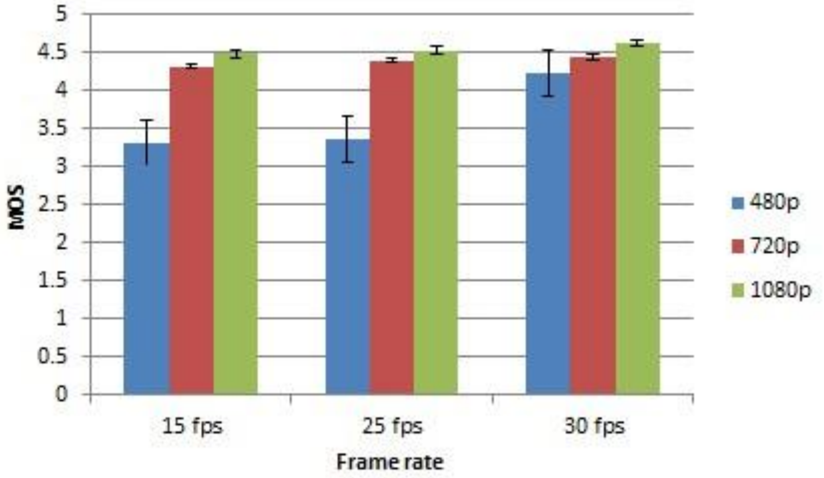


Figure 5.2: MOS values for Galaxy S at 1500 Kbps for Johnny sequence

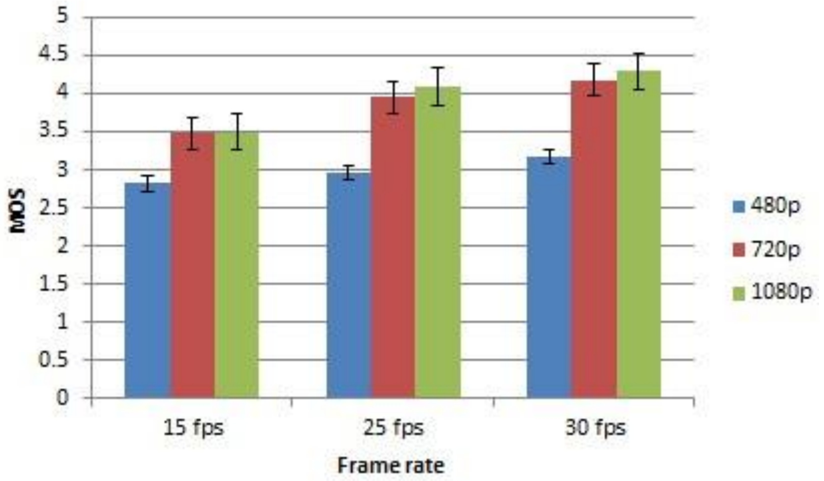


Figure 5.3: MOS values for Galaxy S III at 500 Kbps for Johnny sequence

quence. The same trend was observed for Vidyo medium movement sequence with not much difference in the video quality compared to Johnny video sequence.

Smartphones that are listed in Table 6 are evaluated based on their pixel densities against the MOS values. This chapter has introduced scaling factor based on pixel density. This is done by first mapping the video enconding resolution into the device display resolution, which can be called video encoding pixel density (codPd),

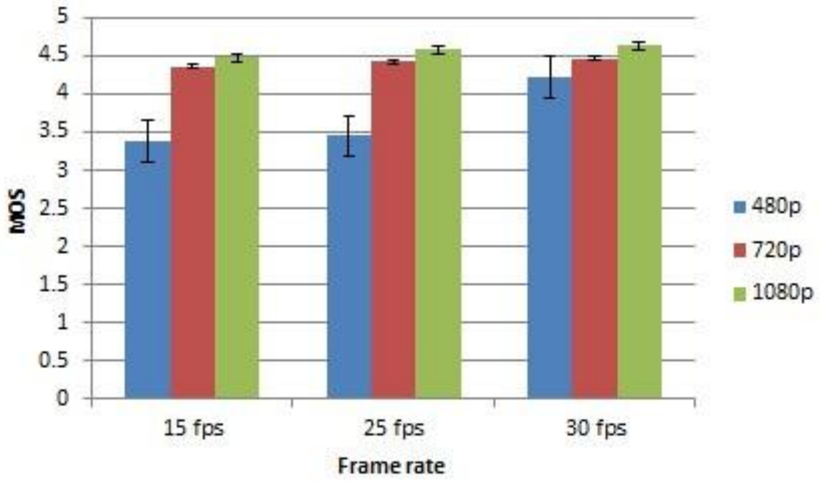


Figure 5.4: MOS values for Galaxy S III at 1500 Kbps for Johnny sequence

$$codpd = \frac{cd_p}{d_i} \qquad (12)$$

where, d_i is the display diagonal size in inches and cd_p is the encoding video display resolution in pixels an as given as

$$cd_p = \sqrt{cw_p^2 + ch_p^2} \qquad (13)$$

where, cw_p and ch_p are width and hight resolutions of the video sequence, respectively. The scale factor based on pixel density sfp is therefore,

$$sfp = \frac{pd}{cpd} \qquad (14)$$

Where, *pd* is pixel density as formulated in (11).

The scale factor proposed in [42] does not consider the display size and there for will not perform well in a scenario where mobile devices have the same resolution but variable display sizes. For instance, if we take Galaxy S 5 with display size of 4" and resolution of 480x800 and S II with display size of 4.3" and resolution of 480x800, for a video sequence of 480p, the scale factor derived in [42] will be the same (0.19) for both mobile devices and therefore, will give the same video quality. However, the scale factor *sfp* proposed in this thesis will be different for Galaxy S (1.17) and Galaxy II (1.11) and therefore, will give different video quality as expected.

Samsung Galaxy S with pixel density 233.24 is compared to S III with pixel density 305.96 for Johnny video sequence with 480p resolution. It can be observed that the Galaxy S III MOS values are better than the ones in Galaxy S at each frame rate. This is because the pixel density of Galaxy S III is higher than Galaxy S. Figure 19 depicts the comparison of the video quality between Galaxy S and S III. The results emphasise on the already existing literature on the impact of frame rate on the video quality, the higher the frame rate the better the quality [66, 67].

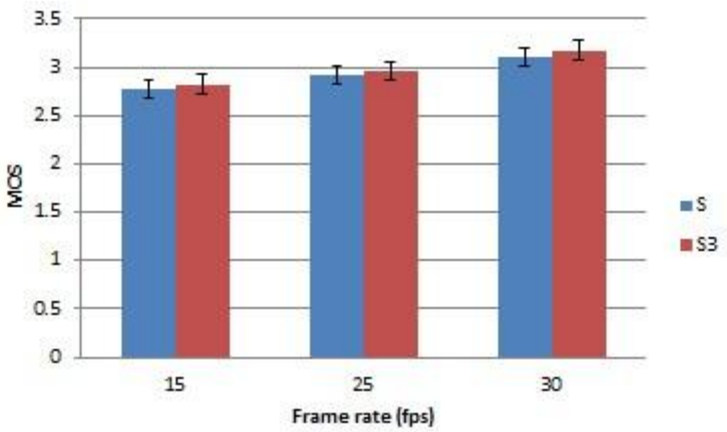


Figure 5.5: Video quality: Galaxy S and S III at 500 Kbps for Johnny sequence

The same trend can be seen for other bit rates (c.f., Figure 20-23) which are

1000 Kbps, 1500 Kbps, 2000 Kbps and 2500 Kbps at each frame rate of 15, 25 and 30.

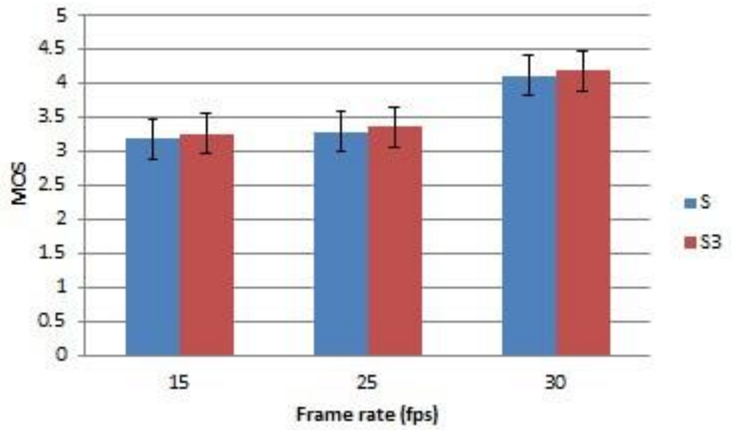


Figure 5.6: Video quality: Galaxy S and S III at 1000 Kbps for Johnny sequence

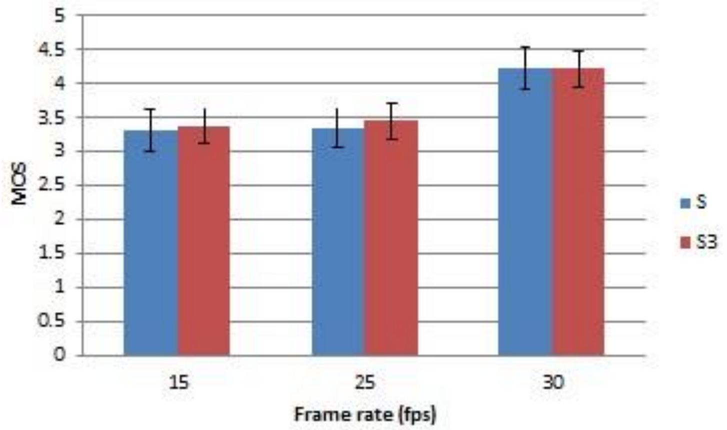


Figure 21: Video quality: Galaxy S and S III at 1500 Kbps for Johnny sequence

The results shown in from Figure 19 to Figure 23 have emphasised on the already existing results in the literature that show the impact of bit rates on the video quality. As the bit rate increases, the video quality increases [5, 68, 69].

The statistical significance of pixel density on the video quality is illustrated by the use of P-value from Analysis of Variance (ANOVA). The P-value was 1.7E-06 $\leq$ 0.05 (Table 24) which shows that the MOS values for all smartphones in the experiment have demonstrated the statistical significant difference with variation in the pixel densities for all bit rates and frame rates

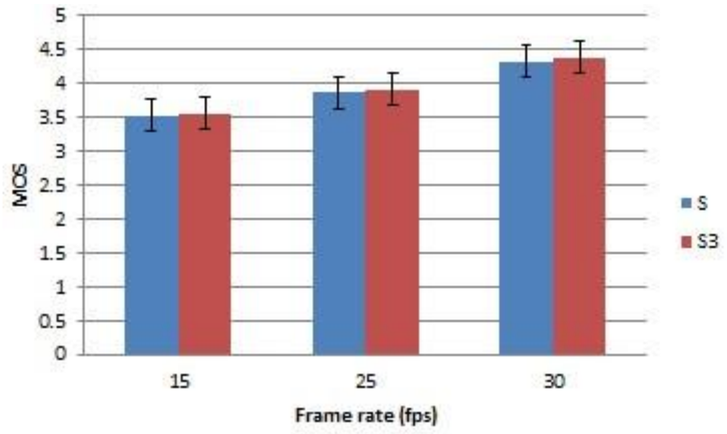


Figure 5.7: Video quality: Galaxy S and S III at 2000 Kbps for Johnny sequence

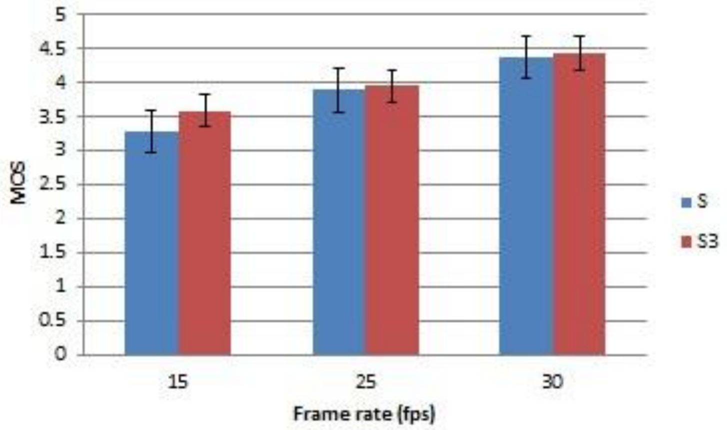


Figure 5.7: Video quality: Galaxy S and S III at 2500 Kbps for Johnny sequence

used in the experiment for the Johnny video sequence as the slow movement video.

The P-value result demonstrates that a null hypothesis is rejected because 1.7E-06 ≤ 0.05, 0.05 is a significance threshold level (α). A null hypothesis is a hypothesis that shows there is no statistical significance of the MOS values between the mobile devices in the experiment.

The P-value for mobile devices and medium movement video is illustrated in Table 25. The P-value is $0.002 \leq 0.05$. This results rejects the null hypothesis and accept the alternative hypothesis which shows that there is a

Table 5.4: P-value for devices and slow movement video

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	49.69	7	7.69	6.90	1.7E-06	2.20
Within devices	120.92	112	0.25			
Total	170.61	119				

statistical significance of the MOS values between all mobile devices in the experiment.

Table 5.5: P-value for devices and medium movement video

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	21.74	7	3.11	3.53	0.002	2.10
Within devices	98.42	112	0.88			
Total	120.16	119				

The P-value for mobile devices and medium movement video is illustrated in Table 26. The P-value is 9.4E-05 ≤ 0.05. This results rejects the null hypothesis and accept the alternative hypothesis which shows that there is a statistical significance of the MOS values between all mobile devices in the experiment.

Table 5.6: P-value for devices and fast movement video

Source of Variation	SS	df	MS	F-value	P-value	F-critical
Between devices	25.72	7	3.67	4.80	9.4E-05	2.09
Within devices	85.76	112	0.77			
Total	111.48	119				

5.4 QoE Modelling based on Regression Technique

This section models the QoE based on scale factor pixel density *sfpd*, bit rate and frame rate. A non-linear regression technique was used to derive QoE model. The dataset was divided into 70% and 30% for training and testing, respectively. Based on the subjective results, the MOS values can be modelled as,

$$MOS = f(BR, FR, SFPD) \tag{15}$$

where *BR*, *FR* and *SFPD* denote bit rate, frame rate and scale factor pixel density, respectively. The proposed model is then derived as,

$$MOS = \alpha + \gamma * FR + \beta * ln(BR) + \delta e^{SFPD} \tag{16}$$

where α, γ, β and δ are coefficients of the proposed model (c.f., Table 27).

Table 5.7: Coefficients of the proposed model

	α	γ	β	δ
Slow movement	-2.18	0.05	0.63	0.37
Medium movement	0.95	0.06	0.24	-0.07
Fast movement	0.64	0.07	0.28	0.01

5.5 Model Evaluation

The propose QoE model based on scale factor pixel density was evaluated by the use of 30% of the dataset for validation. The Root Mean Squared Error (RMSE) and correlation coefficient R^2 were used to find the accuracy of the proposed model in (16). The correlation coefficients and RMSE values are tabulated in Table 28.

Table 5.8: R^2 and RMSE of the proposed model

	Slow movement	Medium movement	Fast movement
*R*2	83%	93%	93%
RMSE	0.24	0.02	0.14

Figure 24, Figure 25 and Figure 26 illustrate the correlations of subjective MOS values against the proposed scale factor pixel density model. From these results, it can be concluded that the proposed scale factor pixel density model can be deployed by multimedia and network service providers as a simple yet accurate objective assessment.

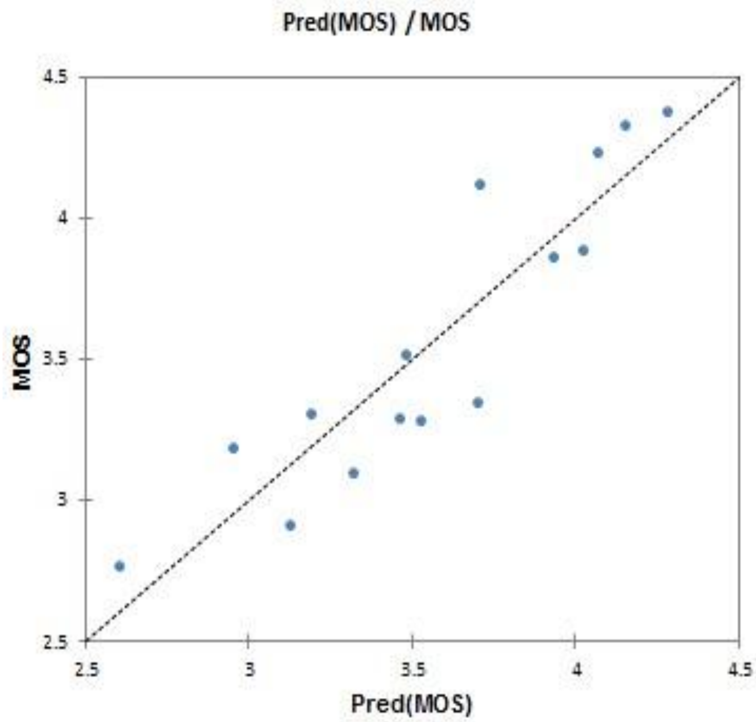


Figure 5.8: Validation of the proposed model for slow movement video

5.6 Summary

This chapter has evaluated the impact of mobile device pixel density on the Quality of Experience for video services. Through subjective tests, results have demonstrated that there is a correlation between mobile device pixel density and video quality. This chapter has also proposed QoE model based on scale factor pixel density. It has been shown that the proposed model performed well in terms of the correlation coefficient. This model can be applicable to video service providers such as YouTube, Netflix and Facebook for control and optimization of video quality. By getting user smartphones manufacturer and model parameters, service providers can easily provision video streams with an acceptable quality based on the smartphones display futures.

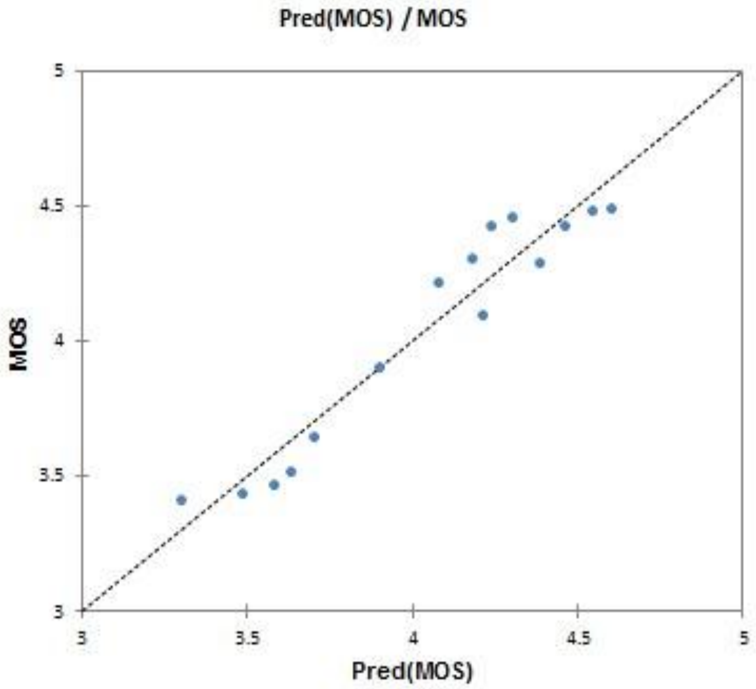


Figure 5.9: Validation of the proposed model for medium movement video

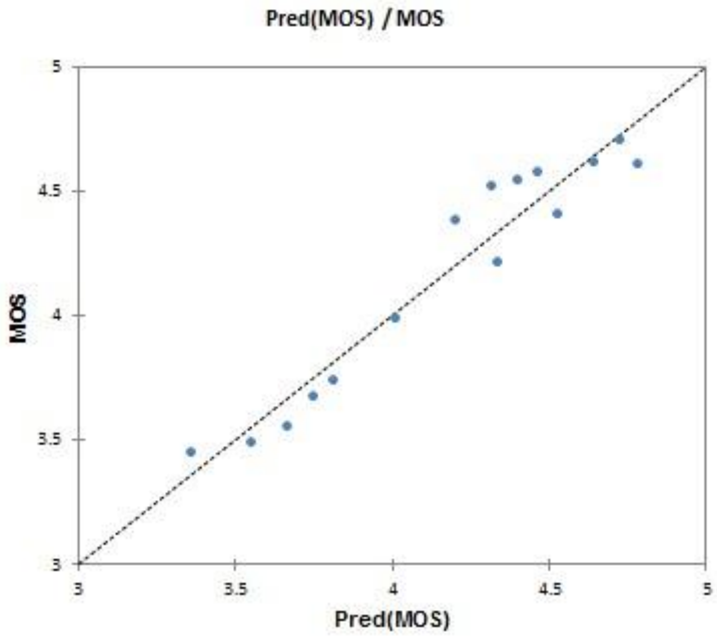


Figure 5.10: Validation of the proposed model for fast movement video

CHAPTER VI

Pixel Density Based Model Implementation

Overview

Dynamic Adaptive Streaming over HTTP popularly known as DASH, is a standard developed by ISO/IEC MPEG. DASH implements the convenient and smooth transmission of video streams over heterogeneous end user devices over the IP network with limited bandwidth. This chapter demonstrates the implementation and the application of the new proposed model over DASH by using an open source libdash library.

6.1 Libdash

Libdash is an open source library providing an object-oriented interface to the DASH standard. It has become and official reference implementation of the ISO/IEC MPEG-DASH standard [1]. The libdash source code is licensed under the GNU Lesser General Public License 2.1+.

Figure 27 depicts the general architecture of MPEG-DASH. The orange parts of the figure are standardized as the Media Presentation Description (MPD) and segment formats. The orange parts which include the transmission of the MPD, streaming control, media player and segment parser are not standardized, which allows different solutions to be implemented. Libdash resides at the client side encapsulating the MPD parsing and HTTP module. The HTTP module is responsible for handling the HTTP download via the streaming logic.

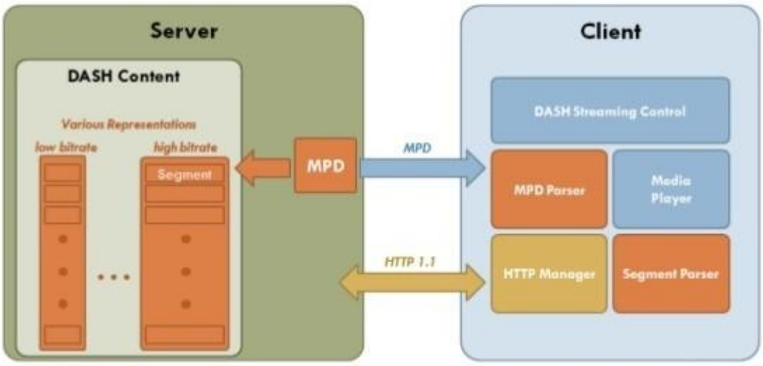


Figure 6.1: MPEG-DASH architecture [1]

DASH server enables the transmission of segments in several video bit rates and resolutions. At the beginning of the HTTP request, a client receives the MPD by using libdash [1] through an object-oriented interface to that MPD. The MPD has the information description of several video segments of varying qualities and durations. The client is able to download individual video segments. At any time in this context, variable bandwidth conditions can be mitigated by switching quality levels of video segments. This approach provides smooth streaming experience by avoiding re-buffering and hence improves QoE.

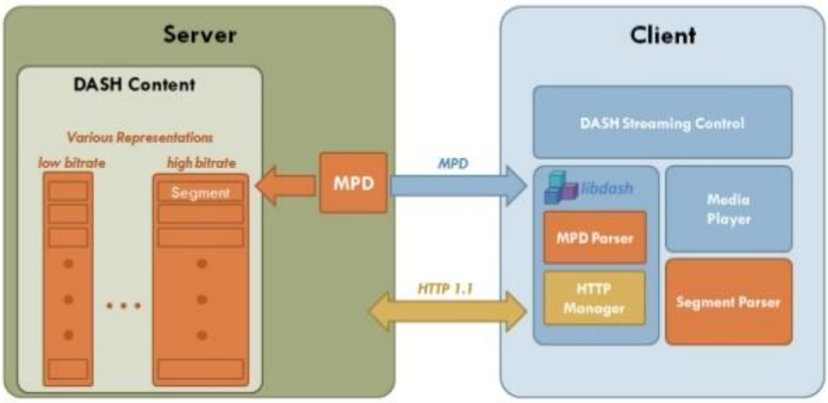


Figure 6.2: Libdash within MPEG-DASH architecture [1]

The new proposed model based on pixel density is implemented in the client side and based on the pixel density, frame rate and bit rate, the client can request varying quality levels of video segments.

6.2 Demonstration Setup

The DASH client with implemented pixel density based model requests video sequences via the WiFi router from the DASH server running under Apache 2 on Ubuntu 14.04.1 64 bit with 8GB RAM. This demonstration setup is depicted in Figure 29.

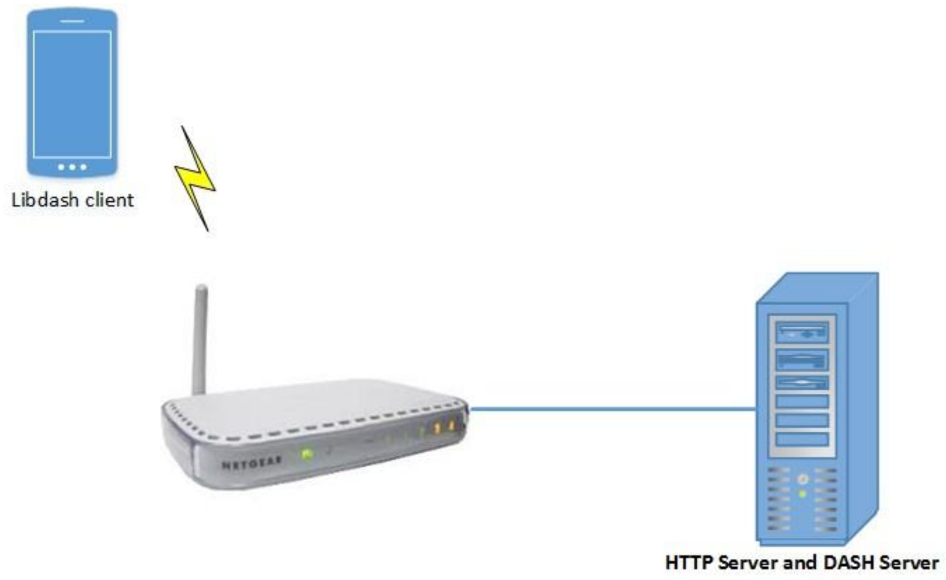


Figure 6.3: Demonstration setup

6.2.1 Segments creation

MP4Box was used to create DASH content to conform with the MPEGDASH specification. Before applying MP4Box, FFMpeg is used to create multi bit rate streams for DASH.

Producing MP4 files to work with HTML5 video on devices which supports H.264 video codec and AAC-LC audio, keyframes aligned video files at different bit rates and resolutions should be generated. This is the main requirement for adaptive bit rate in DASH.

For instance, a 1920x1080 30fps input video file input.mov, the following command in Figure 30 can be executed. The parameters are explained as,

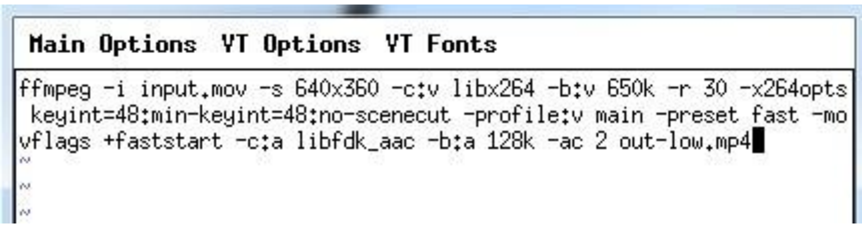


Figure 6.4: Multi bit rate generation

- -i input.mov: Denotes the input video file
- -s 640x360: Resizes ther input file to 640x360
- -c:v libx264: Denotes that x264 is used as the video encoding
- -b:v 650k: Denotes the video bitrate of 650 kbps
- -r 30: Denotes the constant framerate at 30 fps
- -x264opts keyint=48:min-keyint=48:no-scenecut: Denotes that one keyframe at every 48 frames.
- -profile:v main: Denotes the use of H.264 main profile
- -preset fast: Denotes the usage of fast preset for x264 transcoding
- -movflags +faststart: Denotes that the file should be web ready

- -c:a libfdk_aac: Denotes that usage of libfdk_aac for audio encoding
- -b:a 128k: Denotes the target audio bitrate at 128 kbps
- -ac 2: Denotes a stereo output
- out-low.mp4: Denotes the output file as an MP4 file named

Once the output MP4 file is produced then, MP4Box is used to generate DASH segments,

MP4Box -dash 4000 -frag 4000 -rap -segment-name out-low.mp4

The parameters are explained as,

- -dash 4000: Denotes the 4000 ms segment duration
- -frag 4000: Denotes the duration of sub segments. This duration is always less than the segment duration
- -rap: This parameter forces segments to begin with random access points
- -segment-name: Denotes the segment name for generated segments

  • Three video sequences with different spatio and temporal characteristics where used for the implementation of the proposed model . Four People video sequence is a slow movement video, Park Joy is the medium movement video sequence and Stockholm is the fast movement video sequence.

  • These videos are presented in Table 38.

Figure 6.5: Encoded unseen video sequences

Four People

Park Joy

Stockholm

6.3 Video Player via Libdash

Figure 31 illustrates the libdash client playing the Park Joy video sequence. The model was implemented in the client to measure the MOS values for the video quality as the video sequences were played.

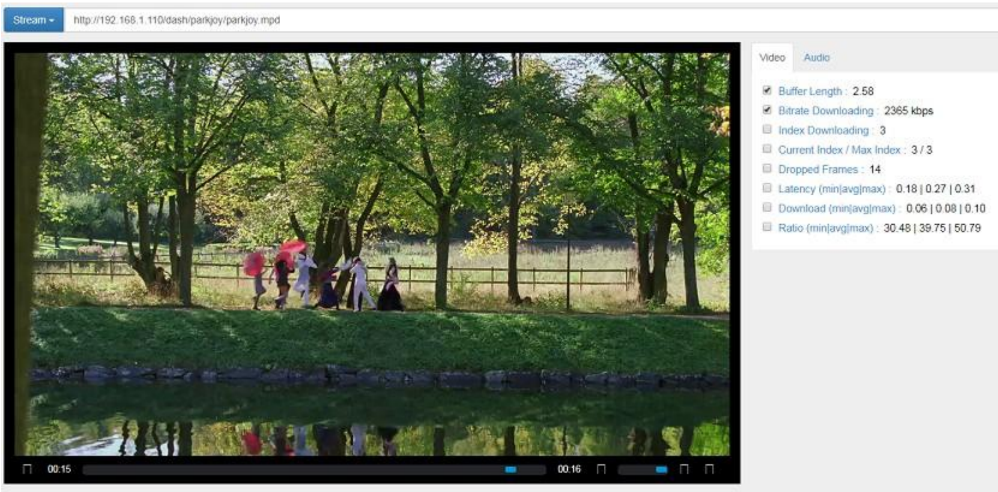


Figure 6.6: Libdash client screen-shot of park Joy video sequence

The inputs of the model are the bit rates which are obtained by using the libdash API, frame rates which are also read from the API and the pixel density which is obtained from the mobile itself. The proposed model will be used to trigger the switch of various quality levels of video segments from the DASH server whenever the measured MOS values drop to the predetermined threshold, for this demo this threshold was set at MOS value of 3.5 which is generally referred as an acceptable quality for video streaming.

Figures 32 and 33 illustrate the libdash client playing the Stockholm and Four People video sequences. The statistics on the right hand side of the implemented player show,

- Buffer Length: Is the length of forwarding buffer in seconds • Bitrate Downloading: The bitrate of the presentation being downloaded.
- Index Downloading: The index of the presentation being downloaded.
- Current Index/Max Index: The index of the current presentation being rendered.

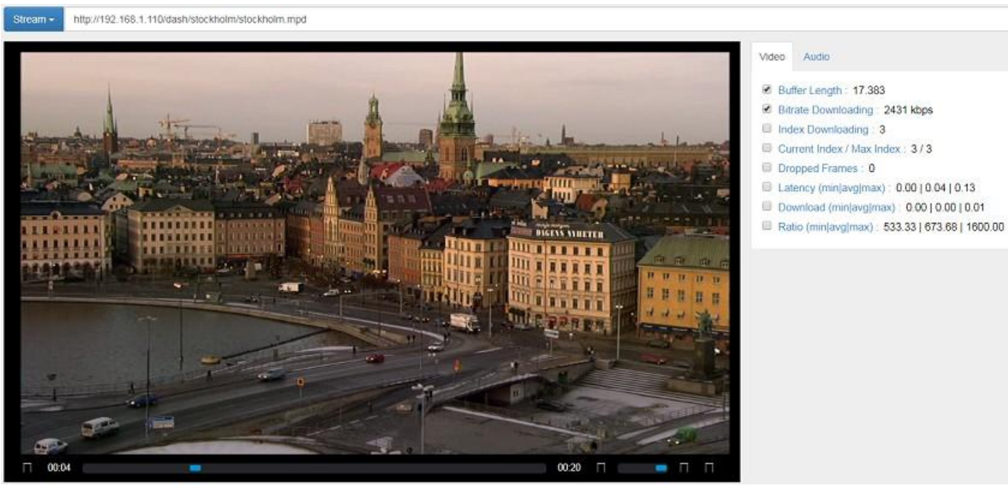


Figure 6.7: Libdash client screen-shot of Stockholm video sequence

- Dropped Frames: The number of frames dropped during rendering. This can be caused by overflowing buffer.

- Latency (min|avg|max): The minimum, average and maximum latency over the last four requested segments. Latency is the time in seconds from the request of segment to receipt of first byte.

- Download (min|avg|max): The minimum, average and maximum download time for the last four requested segments. Download time is the time in seconds from the first byte being received to the last byte.

- Ratio (min|avg|max): The minimum, average and maximum ratio of the segment playback time to total download time over the last 4 segments.

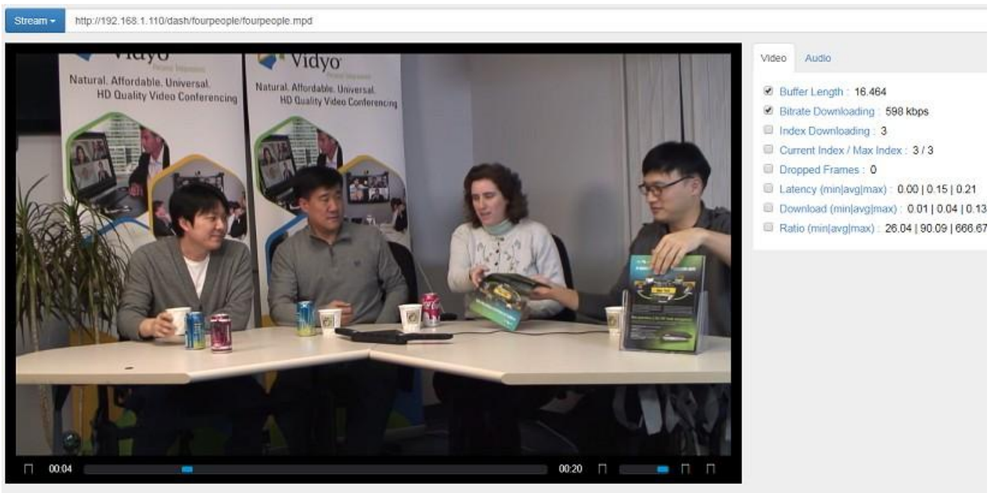


Figure 6.8: Libdash client screen-shot of Four People video sequence

6.4 Media Presentation Description

The DASH media player uses the Media Presentation Description (MPD) [104] is an Extensible Markup Language (XML) formatted document which contains all the information regarding media segments, and other necessary instructions on how select media segments and other metadata that can be used by the DASH media player. The MPD file has five main parts,

- Periods: Describes a part of the video content containing a starting time and its duration. Several Periods can be used to describe scenes and chapters.
- Adaptation Sets: Contains a media stream or set of media streams with different qualities in order to reduce bandwidth.
- Representations: Permits an Adaptation Set to have the same content encoded at different qualities. The most used varying qualities are video resolutions and bit rates. Apart from resolution and bit rates, representation can also be encoded at different codecs. This will allow media player to select the preferred codec for media playback.
- Media Segments: These are the actual media files that are played by the DASH player. These segments are played back to back as if they are in the same file.
- Index Segment: These are of two types, Representation Index Segment for the whole Representation and a Single Index Segment for each Media Segment. In the case of a Representation Index Segment, this is normally a separate file and in the case of a Single Index Segment, this takes a byte range in the same file. Index Segment contains information regarding Media Segment durations in bytes and time by using ISO Base Media File Format (ISOBMFF) Segmentation IndexboX (SIDX) [105].

The MPD file format for Four People video sequence is depicted in Figure 34. In this MPD, there is one Period which consists of one Adaptation Set with three Representations. The duration of the media is 20 seconds. The Adaptation Set has Maximum Frame Rate of 30 and Resolution of 1280x720.

The first Representation has bandwidth of 598369 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 598369 bits/sec bit rate.

The second Representation has bandwidth of 1187292 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 1187292 bits/sec bit rate.

The third Representation has bandwidth of 2365295 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 2365295 bits/sec bit rate.

The MPD file format for Stockholm video sequence is depicted in Figure 35. In this MPD, there is one Period which consists of one Adaptation Set with three Representations. The duration of the media is 20 seconds. The Adaptation Set has Maximum Frame Rate of 30 and Resolution of 1280x720.

The first Representation has bandwidth of 623340 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 623340 bits/sec bit rate.

The second Representation has bandwidth of 1234006 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 1234006 bits/sec bit rate.

The third Representation has bandwidth of 2430909 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 2430909 bits/sec bit rate.

The MPD file format for Park Joy video sequence is depicted in Figure 36.

In this MPD, there is one Period which consists of one Adaptation Set with three Representations. The duration of the media is 16 seconds. The Adaptation Set has Maximum Frame Rate of 30 and Resolution of 1280x720.

The first Representation has bandwidth of 600907 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 600907 bits/sec bit rate.

The second Representation has bandwidth of 1195963 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 1195963 bits/sec bit rate.

The third Representation has bandwidth of 2365207 bits/sec, resolution of 1280x720 and Frame Rate of 30. This Representation has seven segments of 4000 milliseconds each in length at 2365207 bits/sec bit rate.

6.5 Quality Assessment

The video quality assessment is performed to evaluate the performance of the proposed pixel density model implementation. Extensive experiments have been conducted by using several available bandwidth parameters and on three different video sequences (Park Joy, FourPeople and Stockholm) for three quality levels (600 Kbps, 1200 Kbps and 2400 Kbps bit rates). The bandwidths of 2400 Kbps, 1200 Kbps and 600 Kbps are depicted in Figures 37 and 38 for for Park Joy video sequence, respectively. The performance measures used to assess the video quality are average number of stalls and average number of stall time.

6.5.1 DASH without the Proposed Model

The maximum available was fixed at 100 Mbps to assess the performance of DASH without the proposed model. The available bandwidth was varied by introducing the Transmission Control Protocol (TCP) traffic by using iperf traffic generator tool [106].

At 512 Kbps of available bandwidth, Figure 39 depicts the average number of stalls experienced by each video sequence. If the DASH client buffer is starved due to bandwidth fluctuation, the re-buffering will commence, this behaviour causes video to stall. Video stalling plays an important part in the evaluation of video streaming quality. The less the stalling the better the quality.

At 2400 Kbps quality level, more number of stalls occurred for fast movement video sequences (Stockholm = 12) than medium movement (Park Joy = 7) and slow movement (Four People = 5) (c.f., Figure 39). Similar trend can be seen at 1200 Kbps and 600 Kbps quality levels. This is due to the fact that fast movement video sequences occupy more bandwidth than medium and slow movements videos due to high bit rates.

More number of stalls was also observed for fast movement video sequences at 1200 Kbps quality level (Stockholm = 7) than medium movement (Park Joy = 4) and slow movement (Four People = 3) (c.f., Figure 40). Similar trend can be seen at 600 Kbps quality level. This is due to the fact that fast movement video sequences occupy more bandwidth than medium and slow movements videos due to high bit rates.

At 600 Kbps quality level, average number of stalls was observed for fast movement video sequences at 600 Kbps quality level (Stockholm = 3) than medium movement (Park Joy = 1) and slow movement (Four People = 0)
(c.f., Figure 40).

The average stalling time also plays a major role in the evaluation of video quality over DASH, the longer the stalling time, the more dissatisfied users will be. The average stall time in seconds for DASH is taken into consideration in the video quality assessment over DASH.

The average stalling time in seconds is illustrated in Figure 42 for all video sequences at 2400 Kbps quality level. Because of a high bandwidth requirements, fast movement video sequence (Stockholm) had longer average stalling time than medium and slow video sequences (Park Joy and Four People).

For the 1200 Kbps quality level, Figure 43 illustrates the average stalling time in seconds for all video sequence, similar trend is demonstrated in this quality level as in 2400 Kbps.

For the 600 Kbps quality level, no stalls were experienced for slow movement video sequence (Four People) (c.f., Figure 44).

6.5.2 DASH with the Proposed Model

The proposed mobile pixel density model was implemented in the Libdash framework for measurement and control of the video quality. Based on the the proposed model which takes into account the codec resolution, frame rate and device pixel density, Libdash will use it to provide video quality without over provisioning of bandwidth. This is necessary for fair consumption and the increase of multimedia consumption by more users. At 24000 Kbps quality level, the average number of stalling and average stalling time are compared to the default Libdash adaptation. The results show that the proposed model performs better (c.f., Figure 45 and Figure).

At 1200 Kbps quality level, the same trend was observed. The fast moving video sequence (Stockholm) had more number of stalls and at the same time the duration of stalling was longer than the medium movement video sequence (Park Joy) (c.f., Figure 47 and Figure 48). For the 600 Kbps video quality level, the proposed model implementation of the video quality measurement and control did not experience any stalling.

6.6 Summary

This chapter has illustrated and implemented the proposed model based on the pixel density of the smart phones display. The implementation has shown that, the model can be used in DASH standard by video content service providers such as YouTube, Facebook and Netflix to measure and control the video quality (quality adaptation). The implementation has shown that the proposed model performs better than the default Libdash quality adaptation in terms of average stalling numbers and average stalling time.

```xml
<?xml version="1.0"?>
<!-- MPD file Generated with GPAC version 0.5.2-DEV-revVersion: 0.5.2-426-gc5ad4e4+dfsg5-3build1  at 2017-09-28T14:47:09.775Z-->
<MPD xmlns="urn:mpeg:dash:schema:mpd:2011" minBufferTime="PT1.500S" type="static" mediaPresentationDuration="PT0H0M20.033S" maxSegmentDuration="PT0H0M4.000S" profiles="urn:mpeg:dash:profile:full:2011">
 <ProgramInformation moreInformationURL="http://gpac.sourceforge.net">
  <Title>fourpeople.mpd generated by GPAC</Title>
 </ProgramInformation>

 <Period duration="PT0H0M20.033S">
  <AdaptationSet segmentAlignment="true" bitstreamSwitching="true" maxWidth="1280" maxHeight="720" maxFrameRate="30" par="16:9" lang="und">
   <SegmentList>
    <Initialization sourceURL="fourpeople_init.mp4"/>
   </SegmentList>
   <Representation id="1" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" frameRate="30" sar="1:1" startWithSAP="1" bandwidth="598369">
    <SegmentList timescale="30000" duration="100000">
     <SegmentURL media="segment_600_1.m4s"/>
     <SegmentURL media="segment_600_2.m4s"/>
     <SegmentURL media="segment_600_3.m4s"/>
     <SegmentURL media="segment_600_4.m4s"/>
     <SegmentURL media="segment_600_5.m4s"/>
     <SegmentURL media="segment_600_6.m4s"/>
     <SegmentURL media="segment_600_7.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="2" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" frameRate="30" sar="1:1" startWithSAP="1" bandwidth="1187292">
    <SegmentList timescale="30000" duration="100000">
     <SegmentURL media="segment_1200_1.m4s"/>
     <SegmentURL media="segment_1200_2.m4s"/>
     <SegmentURL media="segment_1200_3.m4s"/>
     <SegmentURL media="segment_1200_4.m4s"/>
     <SegmentURL media="segment_1200_5.m4s"/>
     <SegmentURL media="segment_1200_6.m4s"/>
     <SegmentURL media="segment_1200_7.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="3" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" frameRate="30" sar="1:1" startWithSAP="1" bandwidth="2365295">
    <SegmentList timescale="30000" duration="100000">
     <SegmentURL media="segment_2400_1.m4s"/>
     <SegmentURL media="segment_2400_2.m4s"/>
     <SegmentURL media="segment_2400_3.m4s"/>
     <SegmentURL media="segment_2400_4.m4s"/>
     <SegmentURL media="segment_2400_5.m4s"/>
     <SegmentURL media="segment_2400_6.m4s"/>
     <SegmentURL media="segment_2400_7.m4s"/>
    </SegmentList>
   </Representation>
  </AdaptationSet>
 </Period>
</MPD>
```


Figure 6.9: MPD file of Four People video sequence

```xml
<?xml version="1.0"?>
<!-- MPD file Generated with GPAC version 0.5.2-DEV-revVersion: 0.5.2-426-gc5ad4e4+dfsg5-1build1
 at 2017-09-28T15:12:11.976Z-->
<MPD xmlns="urn:mpeg:dash:schema:mpd:2011" minBufferTime="PT1.500S" type="static" mediaPresentati
onDuration="PT0H0M20.133S" maxSegmentDuration="PT0H0M3.200S" profiles="urn:mpeg:dash:profile:full
:2011">
 <ProgramInformation moreInformationURL="http://gpac.sourceforge.net">
  <Title>stockhom.mpd generated by GPAC</Title>
 </ProgramInformation>

 <Period duration="PT0H0M20.133S">
  <AdaptationSet segmentAlignment="true" bitstreamSwitching="true" maxWidth="1280" maxHeight="720
" maxFrameRate="30" par="16:9" lang="und">
   <SegmentList>
    <Initialization sourceURL="stockhom_init.mp4"/>
   </SegmentList>
   <Representation id="1" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="623340">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_600_1.m4s"/>
     <SegmentURL media="segment_600_2.m4s"/>
     <SegmentURL media="segment_600_3.m4s"/>
     <SegmentURL media="segment_600_4.m4s"/>
     <SegmentURL media="segment_600_5.m4s"/>
     <SegmentURL media="segment_600_6.m4s"/>
     <SegmentURL media="segment_600_7.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="2" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="1234006">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_1200_1.m4s"/>
     <SegmentURL media="segment_1200_2.m4s"/>
     <SegmentURL media="segment_1200_3.m4s"/>
     <SegmentURL media="segment_1200_4.m4s"/>
     <SegmentURL media="segment_1200_5.m4s"/>
     <SegmentURL media="segment_1200_6.m4s"/>
     <SegmentURL media="segment_1200_7.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="3" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="2430909">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_2400_1.m4s"/>
     <SegmentURL media="segment_2400_2.m4s"/>
     <SegmentURL media="segment_2400_3.m4s"/>
     <SegmentURL media="segment_2400_4.m4s"/>
     <SegmentURL media="segment_2400_5.m4s"/>
     <SegmentURL media="segment_2400_6.m4s"/>
     <SegmentURL media="segment_2400_7.m4s"/>
    </SegmentList>
   </Representation>
  </AdaptationSet>
 </Period>
</MPD>
```


Figure 6.10: MPD file of Stockholm video sequence

```xml
<?xml version="1.0"?>
<!-- MPD file Generated with GPAC version 0.5.2-DEV-revVersion: 0.5.2-426-gc5ad4e4+dfsg5-1build1
 at 2017-09-20T13:48:01.347Z-->
<MPD xmlns="urn:mpeg:dash:schema:mpd:2011" minBufferTime="PT1.500S" type="static" mediaPresentati
onDuration="PT0H0M16.667S" maxSegmentDuration="PT0H0M3.200S" profiles="urn:mpeg:dash:profile:full
:2011">
 <ProgramInformation moreInformationURL="http://gpac.sourceforge.net">
  <Title>parkjoy.mpd generated by GPAC</Title>
 </ProgramInformation>

 <Period duration="PT0H0M16.667S">
  <AdaptationSet segmentAlignment="true" bitstreamSwitching="true" maxWidth="1280" maxHeight="720
" maxFrameRate="30" par="16:9" lang="und">
   <SegmentList>
    <Initialization sourceURL="parkjoy_init.mp4"/>
   </SegmentList>
   <Representation id="1" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="600907">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_600_1.m4s"/>
     <SegmentURL media="segment_600_2.m4s"/>
     <SegmentURL media="segment_600_3.m4s"/>
     <SegmentURL media="segment_600_4.m4s"/>
     <SegmentURL media="segment_600_5.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="2" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="1195963">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_1200_1.m4s"/>
     <SegmentURL media="segment_1200_2.m4s"/>
     <SegmentURL media="segment_1200_3.m4s"/>
     <SegmentURL media="segment_1200_4.m4s"/>
     <SegmentURL media="segment_1200_5.m4s"/>
    </SegmentList>
   </Representation>
   <Representation id="3" mimeType="video/mp4" codecs="avc3.4d401f" width="1280" height="720" fra
meRate="30" sar="1:1" startWithSAP="1" bandwidth="2365207">
    <SegmentList timescale="30000" duration="96000">
     <SegmentURL media="segment_2400_1.m4s"/>
     <SegmentURL media="segment_2400_2.m4s"/>
     <SegmentURL media="segment_2400_3.m4s"/>
     <SegmentURL media="segment_2400_4.m4s"/>
     <SegmentURL media="segment_2400_5.m4s"/>
    </SegmentList>
   </Representation>
  </AdaptationSet>
 </Period>
</MPD>
```

Figure 6.11: MPD file of Park Joy video sequence

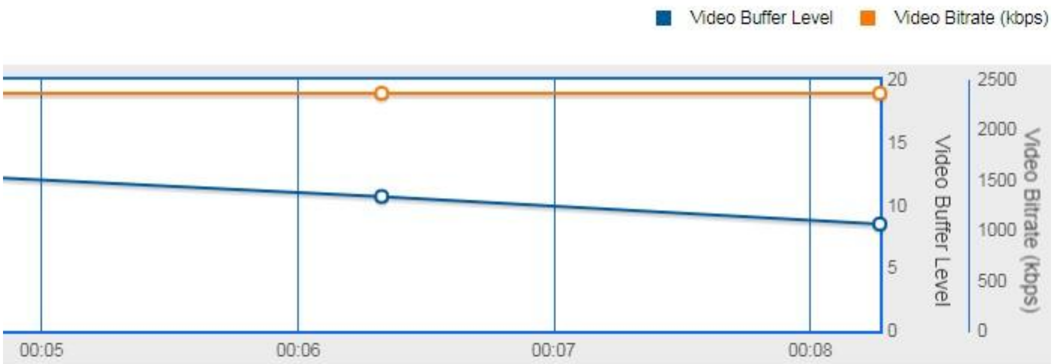


Figure 6.12: Video bandwidth display at 2400 Kbps

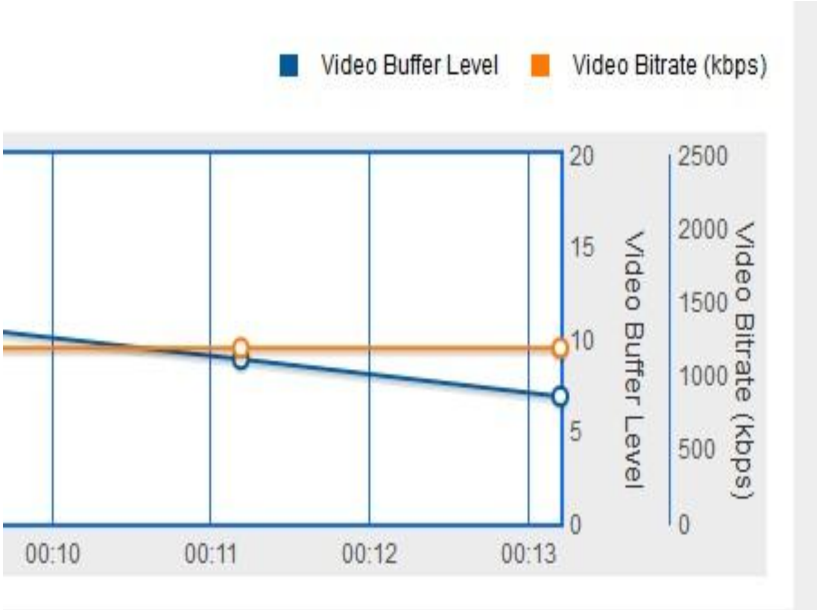


Figure 6.13: Video bandwidth display at 1200 Kbps

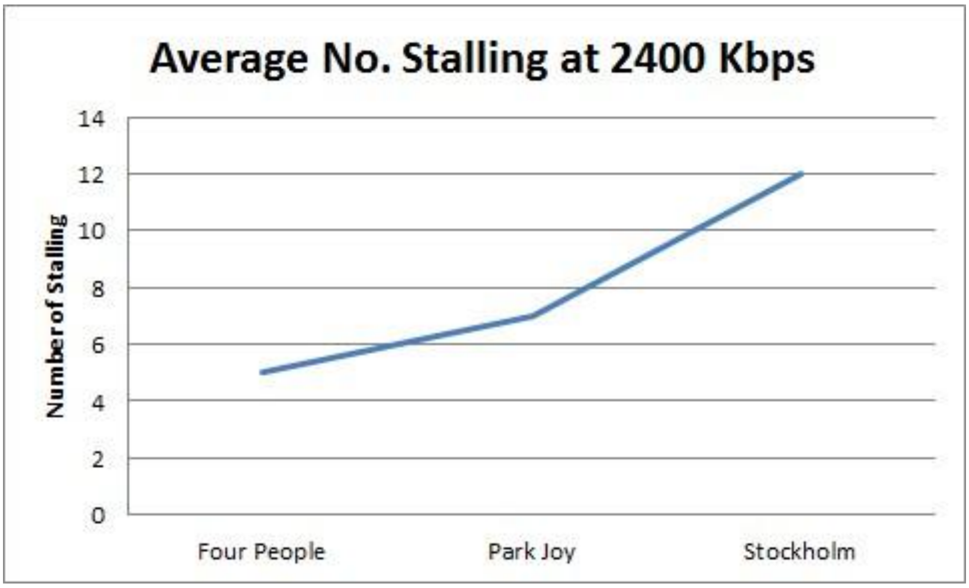


Figure 6.14: Number of stalls at 2400 Kbps bit rates

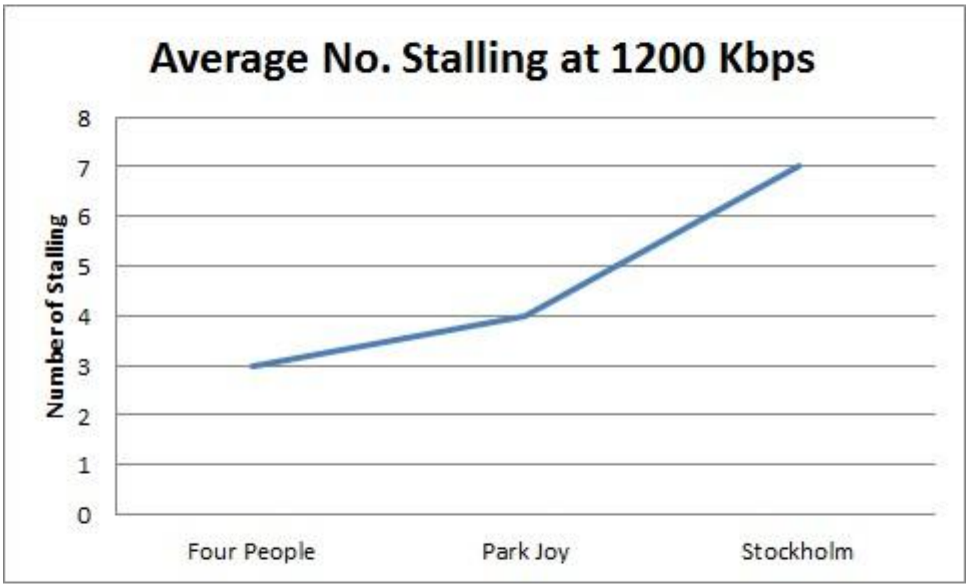


Figure 6.15: Number of stalls at 1200 Kbps bit rates

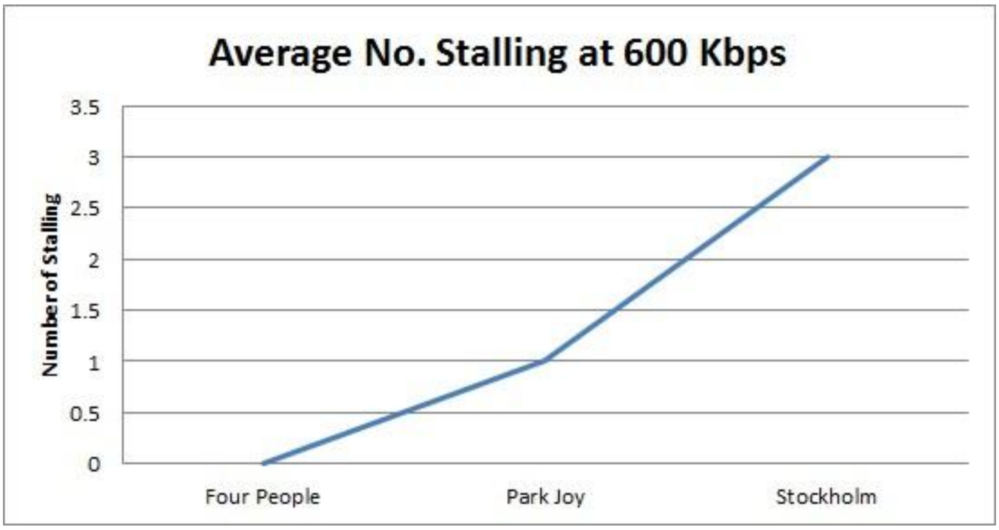


Figure 41: Number of stalls at 600 Kbps bit rates

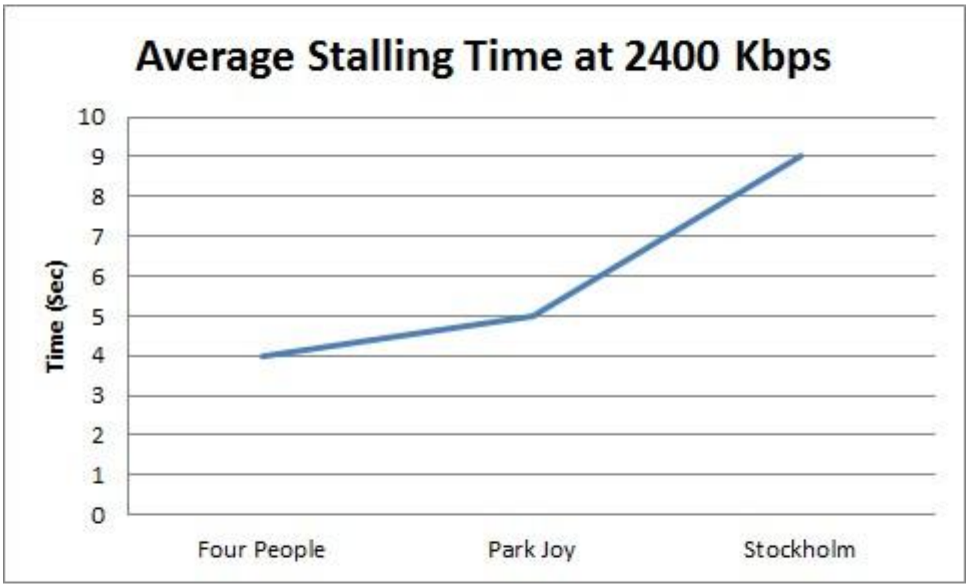


Figure 6.16: Average stalling time at 2400 Kbps bit rate

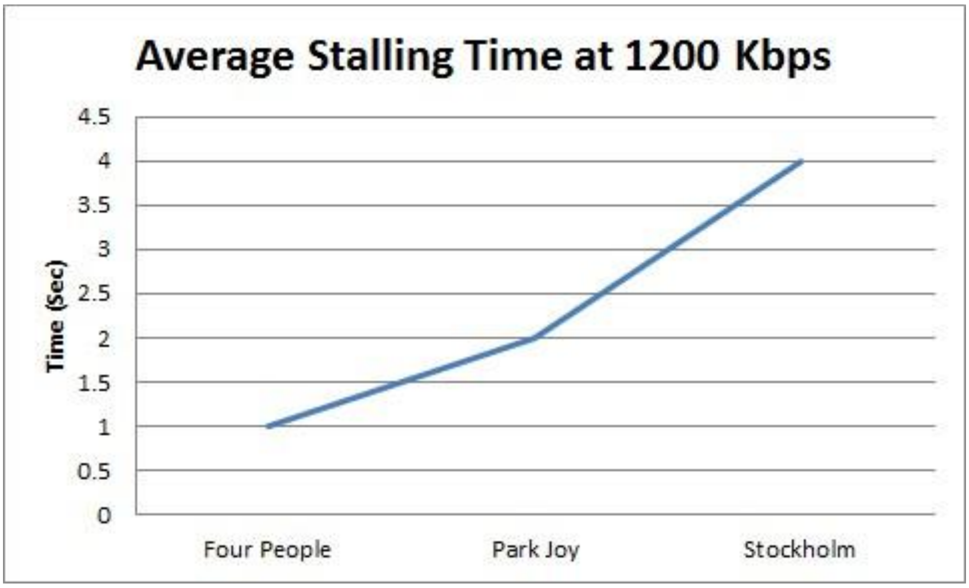


Figure 6.17: Average stalling time at 1200 Kbps bit rate

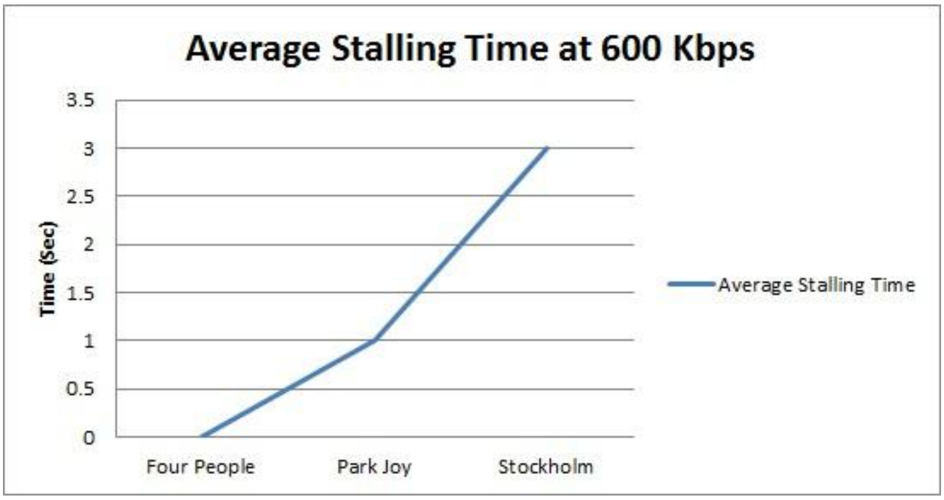


Figure 6.18: Average stalling time at 600 Kbps bit rate

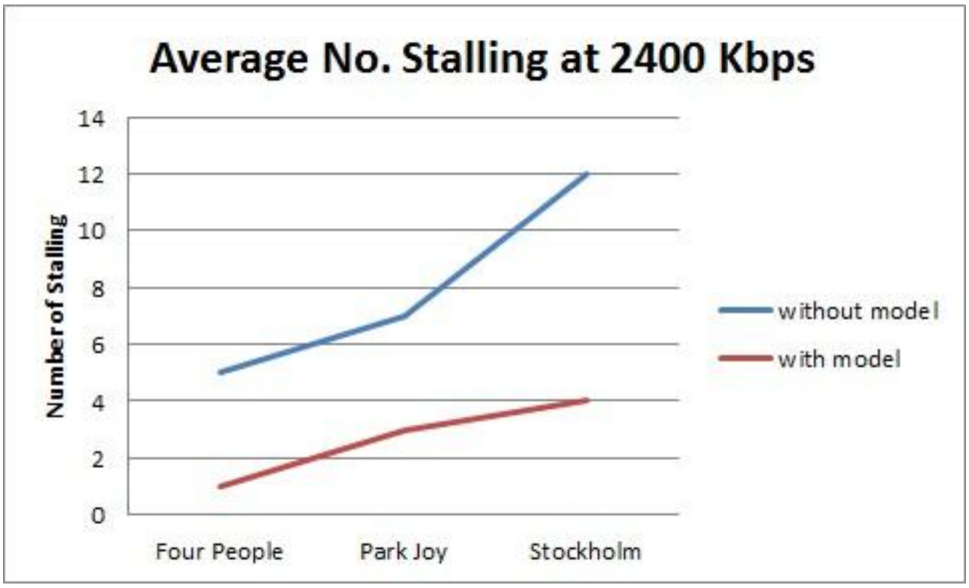


Figure 6.19: Average number of stalls: Proposed model at 2400 Kbps

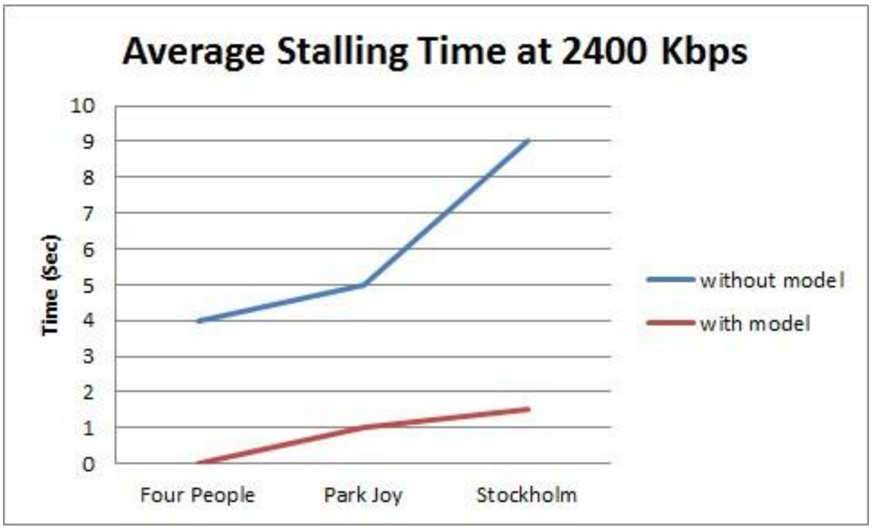


Figure 6.20: Average stalling time: Proposed model at 2400 Kbps

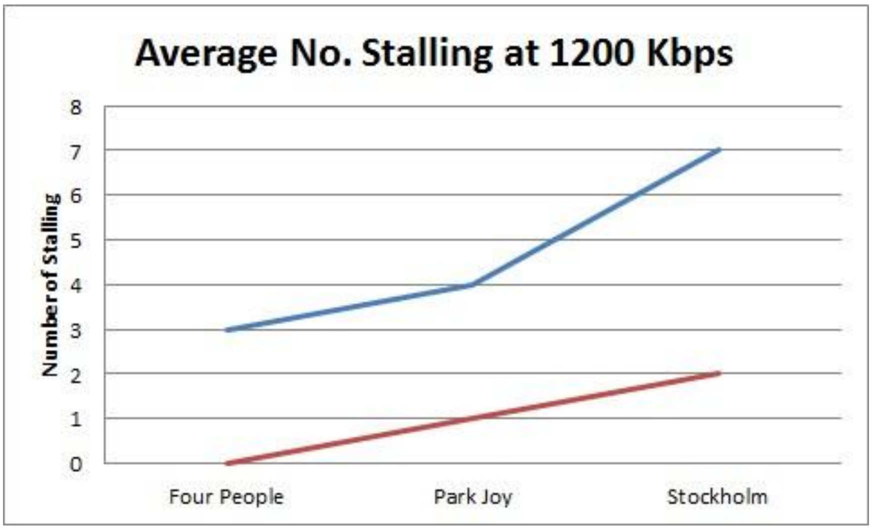


Figure 6.21: Average number of stalls: Proposed model at 1200 Kbps

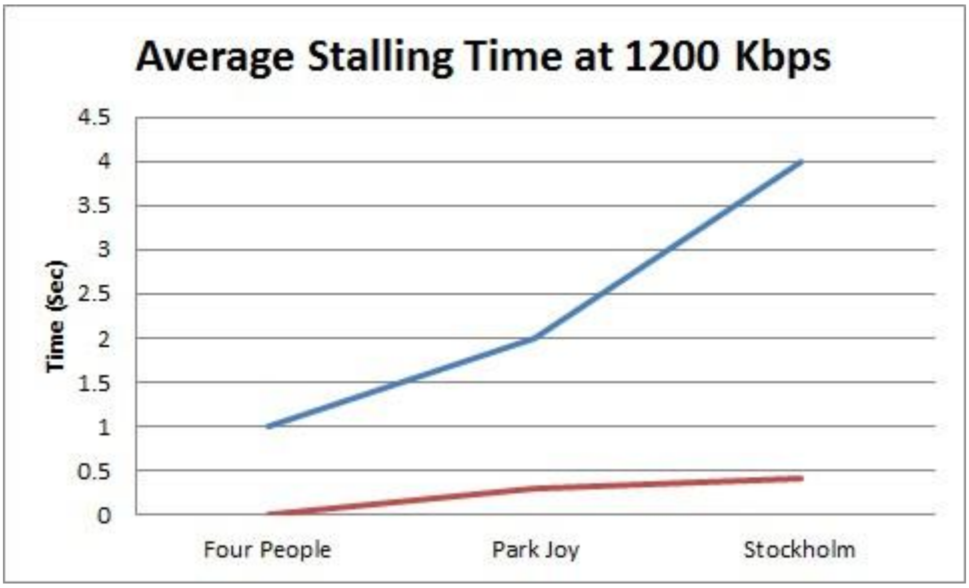


Figure 6.22: Average stalling time: Proposed model at 1200 Kbps

CHAPTER VII

7 Conclusions and Future Work

The advances in computational power and memory storage of mobile devices, video coding efficiency and advances experienced in wireless communication networks has increased the widespread usage of mobile devices and mobile video applications for several video streaming service providers. However, the success of the current and future mobile video applications will rely on the perceivable Quality of Experience of end users. Several end to end factors impact QoE. These factors can be categorized into network, devices and users. For devices, parameters such as display size and resolution can have an effect on the quality of the video. It is therefore, important for multimedia service and network providers to measure, monitor or predict the QoE especially of end users mobile devices.

Subjective tests or objective methods are used to measure video quality. The Mean Opinion Score is recommended by the ITU [29] to be used in subjective tests for measuring video quality. The drawback of subjective tests is that it expensive and it is time consuming. Subjective test cannot be deployed to measure video quality in real time scenarios such as in live video steaming. Therefore, objective methods are ideal for live video streaming.

7.1 Contributions to knowledge

This thesis has the following main contributions:

1. Provision of the detailed understanding of the relationship between video quality and users' mobile device preferences:

   This work has contributed to a detailed understanding through subjective tests of human preference of mobile devices and the quality of video when viewed on preferred mobile devices. This understanding can provide a foundation for the development of efficient objective models that can be deployed by multimedia service providers such as YouTube, Netflix and Facebook for video quality monitoring, control and optimization.

   This work has been contributed to the research community through dissemination as a publication published in the International Journal of Computing and Information Sciences [17].

2. New non-intrusive model to predict video quality non-intrusively based on mobile device pixel density:

   This work has contributed to the development of a new non-intrusive model for video quality measurement and prediction. The model was based on the combination of coding parameters (bit rate and frame rate) and device parameters (pixel density). The model was derived by a non-linear regression method from the subjective tests on Samsung mobile phones with variable display sizes and resolutions.

   This study has been contributed to the research community through a publication to be published by the Conference on Information and Computer Science [18].

3. Application of the proposed model in the Dynamic Adaptive Streaming over HTTP (DASH):

The new model was applied in DASH to demonstrate its usefulness. DASH is a popular standard used by major video streaming providers such as YouTube to measure, control and optimize video quality. Libdash [51] was used as a DASH server via the Apache 2 web server. The implementation enables content providers to optimally select video send bit rate, frame rate and resolution depending on the mobile device pixel density and resolution.

4. Availability of the subjective tests dataset to the research community:

   The subjective tests results of the mobile device preferences and the pixel densities will be available to the research community for further research by either evaluating or validating their studies against the ones obtained in this research. These dataset will include raw videos, encoded vidoes and MOS values.

7.2 Limitations of the Current Research

All this thesis has successfully met all the requirements and objective set out in the Introduction chapter, there are some limitations that can be addressed in the future work.

1. considerations of mobile device options:

   For mobile device preferences subjective tests, three mobile devices were used. More mobile devices with varying manufactures can be used in future work in order to capture the widely used mobile devices. This also applies to the subjective tests for pixel density experiment.

2. Improved validation of the work:

Although the new Scale Factor model has been validated by the data that was not used to derive the model, external dataset and large scale dataset is still needed to further validate the model.

3. Content types consideration:

   In this thesis, five different video sequences were considered. These video sequences covered slow moving (head and shoulder like news broadcaster) to fast moving (sports type like basketball). Although cartoons video sequence were considered, movie video sequences were considered. The spatio and temporal features of movie sequences may have an impact on the content type and therefore, on the new proposed quality model.

7.3 Suggested Future work

There are some aspects of this research that can further be investigated in the future.

1. Wide range of video send bit rates, frame rates and resolutions:

   The send bit rates used in this research ranged from 1000 Kbps to 6000 Kbps. The target was to consider HD video transmission ranging from 1000 Kbps to 6000 kbps for HD quality of 480p to 1080p. Future work could consider low bit rates from 300 Kbps for 240p, 360p to 480p. Higher resolutions of 1440p to 4K could be considered in the future with corresponding bit rates from 6000 Kbps to 51000 Kbps. The frame rate of 60 fps could also be considered in the future corresponding to 4K video quality. This will make the proposed model more generic for a wide range of values of frame rates, bit rates and resolutions.
2. Objective methods of measuring and predicting video quality based on mobile device preferences:

This study has conducted subjective tests to find the impact of mobile device preferences on the quality of experience for video services. As mentioned in the literature review chapter regarding the importance of objective methods in real time scenarios and cost, there is a need to develop non-intrusive video quality model that take into account the device preference parameters. In order to improve this model, ideas from the business and economics disciplines can investigated on how modelling consumer preferences are conducted [70].

7.4 Conclusions

Exponential growth of mobile devices especially smartphones have come with challenges which include varying display sizes, resolutions, hardware configuration and operating systems. Inspired by these challenges, this research was initiated to investigate the impact of mobile devices preferences, pixel density and resolutions on the quality of video services.

Based on subjective tests, results have shown that there is a correlation between device preferences and QoE for video services. Medium movement video sequences (Big Buck Bunny) demonstrated to have higher quality at the same bit rates compared to fast movement video sequences (Elephant Dreams). Based on these results, there is a need to consider human factors such as device and content preferences when objective models are proposed by video service providers such as YouTube and Netflix. This investigation has shown the importance of its results in the multimedia services networking and communications because it can be used in the video quality control and optimization over the limited bandwidth and storage environment.

Based on the subjective tests, a new non-linear regression model was developed which takes pixel density into considerations together with video coding parameters such as send bit rates, frame rates and resolutions. It has been shown that the proposed model performed well in terms of the correlation coefficient.

The proposed model can be used by video service providers such as YouTube, Netflix and Facebook for control and optimization of video quality. By obtaining user smartphones manufacturer and model parameters, service providers can easily provision video streams with an acceptable quality based on the smartphones display futures.

As a proof of concept, the proposed model was deployed and integrated in the DASH test-bed. The implementation was successful which shown the usefulness of the proposed model in the control and optimization of video quality over video streaming services.

This thesis has successfully met all the objectives and answered all questions that have been posed in the introduction chapter.

References

[1] C. Mueller, S. Lederer, J. Poecher, and C. Timmerer, “Demo paper: Libdash-an open source software library for the mpeg-dash standard,” in *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on*, pp. 1–2, IEEE, 2013.

[2] H. Choi, J. Nam, D. Sim, and I. V. Bajic, “Scalable video coding based´ on high efficiency video coding (hevc),” in *Communications, Computers and Signal Processing (PacRim), 2011 IEEE Pacific Rim Conference on*, pp. 346–351, IEEE, 2011.

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

[4] V. Verdot, A. Gonguet, N. Bouché, and U. Lucena, “Intention-aware multimedia modeling for optimized quality of experience,” in *Globecom Workshops (GC Wkshps), 2012 IEEE*, pp. 1298–1303, IEEE, 2012.

[5] S. Winkler and P. Mohandas, “The evolution of video quality measurement: From psnr to hybrid metrics,” *IEEE Transactions on Broadcasting*, vol. 54, no. 3, pp. 660–668, 2008.

[6] B. Lee and M. Kim, “No-reference psnr estimation for hevc encoded video,” *IEEE Transactions on Broadcasting*, vol. 59, no. 1, pp. 20–27, 2013.

[7] L. Anegekuh, L. Sun, E. Jammeh, I. H. Mkwawa, and E. Ifeachor, “Content-based video quality prediction for hevc encoded videos streamed over packet networks,” *IEEE Transactions on Multimedia*, vol. 17, pp. 1323–1334, Aug 2015.

[8] A. Khan, L. Sun, and E. Ifeachor, “Qoe prediction model and its application in video quality adaptation over umts networks,” *IEEE Transactions on Multimedia*, vol. 14, pp. 431–442, April 2012.

[9] J. Bruneau-Queyreix, M. Lacaud, D. Negru, J. M. Batalla, and E. Borcoci, "Qoe enhancement through cost-effective adaptation decision process for multiple-server streaming over http," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, July 2017.

[10] A. Asan, W. Robitza, I. h. Mkwawa, L. Sun, E. Ifeachor, and A. Raake, "Impact of video resolution changes on qoe for adaptive video streaming," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 499–504, July 2017.

[11] I. H. Mkwawa, A. A. Barakabitze, and L. Sun, "Video quality management over the software defined networking," in *2016 IEEE International Symposium on Multimedia (ISM)*, pp. 559–564, Dec 2016.

[12] M. Naccari, M. Tagliasacchi, and S. Tubaro, "No-reference video quality monitoring for h. 264/avc coded video," *IEEE Transactions on Multimedia*, vol. 11, no. 5, pp. 932–946, 2009.

[13] A. Khan, L. Sun, and E. Ifeachor, "Content clustering based video quality prediction model for mpeg4 video streaming over wireless networks," in *Communications, 2009. ICC'09. IEEE International Conference on*, pp. 1–5, IEEE, 2009.

[14] C. Bezerra, A. D. Carvalho, D. Borges, N. Barbosa, J. Pontes, and E. Tavares, "Qoe and energy consumption evaluation of adaptive video streaming on mobile device," in *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*, pp. 1–6, Jan 2017.

[15] N. Eswara, K. Manasa, A. Kommineni, S. Chakraborty, H. P. Sethuram, K. Kuchi, A. Kumar, and S. S. Channappayya, "A continuous qoe evaluation framework for video streaming over http," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2017.

[16] M. Usman, N. Yang, M. A. Jan, X. He, M. Xu, and K. M. Lam, "A joint framework for qos and qoe for video transmission over wireless multimedia sensor networks," *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[17] A. Elmnsi, , N. Osman, I.-H. Mkwawa, and and, “The impact of mobile device preference on the quality of experience,” *International Journal of Computing and Information Sciences,* vol. 12, pp. 89–94, sep 2016.

[18] A. Elmnsi, , N. Osman, I.-H. Mkwawa, and and, “Mobile devices pixel density and video quality,” *Conference on Information and Computer Science*, vol. 12, pp. 1–6, sep 2017.

[19] J. Ni, X. Tan, Y. Shao, X. Wu, and J. Zhu, “Qoe-driven rate adaptation algorithm for fair dynamic adaptive video streaming in named data networking,” in *2017 36th Chinese Control Conference (CCC),* pp. 7599–7604, July 2017.

[20] Y.-F. Ou, Z. Ma, and Y. Wang, “A novel quality metric for compressed video considering both frame rate and quantization artifacts,” *City,* vol. 80, p. 100, 2009.

[21] L. Zhang, S. Wang, F. Yang, and R. N. Chang, “Qoecenter: A visual platform for qoe evaluation of streaming video services,” in *2017 IEEE International Conference on Web Services (ICWS)*, pp. 212–219, June 2017.

[22] B. Dey and M. K. Kundu, “Enhanced macroblock features for dynamic background modeling in h.264/avc video encoded at low-bitrate,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2016.

[23] C. Sloan, N. Harte, D. Kelly, A. C. Kokaram, and A. Hines, “Bitrate classification of twice-encoded audio using objective quality features,” in *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX),* pp. 1–6, June 2016.

[24] I. T. Rec", “Definitions used in recommendations on general characteristics of international telephone connections and circuits,” in *G.100,* 2006.

[25] D. Soldani, “Bridging qoe and qos for mobile broadband networks,” in *ETSI workshop on QoS, QoE and User Experience focusing on*

speech, multimedia conference tools, 2010.

[26] M. Robinson, M. Milosavljevic, P. Kourtessis, S. Fisher, G. P. Stafford, J. Treiber, M. J. Burrell, and J. M. Senior, "Qoe based holistic traffic engineering in sdn enabled heterogeneous transport networks," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, July 2017.

[27] T. Mangla, E. Halepovic, M. Ammar, and E. Zegura, "Mimic: Using passive network measurements to estimate http-based adaptive video qoe metrics," in *2017 Network Traffic Measurement and Analysis Conference (TMA)*, pp. 1–6, June 2017.

[28] A. Schwind, M. Seufert, O. Alay, P. Casas, P. Tran-Gia, and F. Wamser, "Concept and implementation of video qoe measurements in a mobile broadband testbed," in *2017 Network Traffic Measurement and Analysis Conference (TMA)*, pp. 1–6, June 2017.

[29] P. ITU-T Recommendation, "Subjective video quality assessment methods for multimedia applications," 1999.

[30] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE transactions on broadcasting*, vol. 57, no. 2, pp. 165–182, 2011.

[31] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal processing: Image communication*, vol. 19, no. 2, pp. 121–132, 2004.

[32] Y. Chtouki, H. Harroud, M. Khalidi, and S. Bennani, "The impact of youtube videos on the student's learning," in *Information Technology Based Higher Education and Training (ITHET), 2012 International Conference on*, pp. 1–4, IEEE, 2012.

[33] P. N. Satgunam, R. L. Woods, P. M. Bronstad, and E. Peli, "Factors affecting enhanced video quality preferences," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5146–5157, 2013.

[34] A. Molnar, "Factors affecting user preference for mobile video quality," 2016.

[35] A. Khan, L. Sun, and E. Ifeachor, "Impact of video content on video quality for video over wireless networks," in *2009 Fifth International Conference on Autonomic and Autonomous Systems*, pp. 277–282, IEEE, 2009.

[36] V. Menkovski, G. Exarchakos, A. Liotta, and A. C. Sánchez, "Measuring quality of experience on a commercial mobile tv platform," in *Advances in Multimedia (MMEDIA), 2010 Second International Conferences on*, pp. 33–38, IEEE, 2010.

[37] A. Floris, L. Atzori, G. Ginesu, and D. D. Giusto, "Qoe assessment of multimedia video consumption on tablet devices," in *2012 IEEE Globecom Workshops*, pp. 1329–1334, IEEE, 2012.

[38] L. Yu, T. Tillo, and J. Xiao, "Qoe-driven dynamic adaptive video streaming strategy with future information," *IEEE Transactions on Broadcasting*, vol. 63, pp. 523–534, Sept 2017.

[39] D. Z. Rodríguez, R. L. Rosa, E. A. Costa, J. Abrahão, and G. Bressan, "Video quality assessment in video streaming services considering user preference for video content," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 3, pp. 436–444, 2014.

[40] W. Zou, J. Song, and F. Yang, "Perceived image quality on mobile phones with different screen resolution," *Mobile Information Systems*, vol. 2016, 2016.

[41] S. Mayr, M. Köpper, and A. Buchner, "Effects of high pixel density on reading comprehension, proofreading performance, mood state, and physical discomfort," *Displays*, vol. 48, pp. 41–49, 2017.

[42] I.-T. Rec", "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport – video quality estimation module," in *P.1203-1*, 2016.

[43] J. Tague-Sutcliffe, "The pragmatics of information retrieval experimentation, revisited," *Information Processing & Management*, vol. 28, no. 4, pp. 467–490, 1992.

[44] L. A. Ponemon and J. P. Wendell, "Judgmental versus random sampling in auditing: An experimental investigation," *Auditing*, vol. 14, no. 2, p. 17, 1995.

[45] I. T. Rec", "Methods for the subjective assessment of video quality, audio quality and audiovisual quality of internet video and distribution quality television in any environment," in *P.913*, 2016.

[46] L. Janowski and M. Pinson, "Subject bias: Introducing a theoretical user model," in *Quality of Multimedia Experience (QoMEX), 2014 Sixth International Workshop on*, pp. 251–256, IEEE, 2014.

[47] I. T. Rec, "Methods of subjective determination of transmission quality," 1996.

[48] M. J. Crawley, "Statistical computing: An introduction to data analysis using," 2002.

[49] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2015.

[50] B. Simon, "Smartphone os market share, 2017 q1," 2017.

[51] C. Mueller, S. Lederer, J. Poecher, and C. Timmerer, "Demo paper: Libdash - an open source software library for the mpeg-dash standard," in *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pp. 1–2, July 2013.

[52] C. V. N. Index, “Global mobile data traffic forecast update 2014–2019 white paper, feb 2015,” *See: http://www. cisco. com/c/en/us/solutions/collateral/service-provider/visual-networkingindex-vni/white_paper_c11-520862. html*, 2015.

[53] K. Mitra, A. Zaslavsky, and C. Åhlund, “Context-aware qoe modelling, measurement, and prediction in mobile computing systems,” *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 920–936, 2015.

[54] M. Alreshoodi and J. Woods, “Survey on qoe\qos correlation models for multimedia services,” *CoRR*, vol. abs/1306.0221, 2013.

[55] I.-T. Rec", “Perceptual evaluation of speech quality (pesq): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” in *P.862*, 2001.

[56] I.-T. Rec, “The e-model, a computational model for use in transmission planning,” in *G.107*, 2009.

[57] K. Mitra, C. Åhlund, and A. Zaslavsky, “A decision-theoretic approach for quality-of-experience measurement and prediction,” in *2011 IEEE International Conference on Multimedia and Expo*, pp. 1–4, IEEE, 2011.

[58] K. Mitra, A. Zaslavsky, and C. Åhlund, “A probabilistic context-aware approach for quality of experience measurement in pervasive systems,” in *Proceedings of the 2011 ACM symposium on applied computing*, pp. 419–424, ACM, 2011.

[59] V. Menkovski, G. Exarchakos, and A. Liotta, “Online qoe prediction,” in

Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on, pp. 118–123, IEEE, 2010.

[60] L. Sun and E. C. Ifeachor, “Voice quality prediction models and their application in voip networks,” *IEEE transactions on multimedia*, vol. 8, no. 4, pp. 809–820, 2006.

[61] L. Janowski and Z. Papir, "Modeling subjective tests of quality of experience with a generalized linear model," in *Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on*, pp. 35–40, IEEE, 2009.

[62] I. ITU-T Recommendation, "500-11,"methodology for the subjective assessment of the quality of television pictures," recommendation itu-r bt. 500-11," *ITU Telecom. Standardization Sector of ITU*, vol. 7, 2002.

[63] C. V. N. Index, "Cisco visual networking index: Global mobile data traffic forecast update, 2016–2021 white paper, feb 2017," *See: http://www.cisco.com/c/en/us/solutions/collateral/serviceprovider/visual-networking-index-vni/mobile-white-paper-c11520862.html*, 2017.

[64] O. Index, "Android fragmentation report, 2015," *See: https://opensignal.com/reports/2015/08/android-fragmentation/*, 2015.

[65] F. Kuipers, R. Kooij, D. D. Vleeschauwer, and K. Brunnström, "Techniques for measuring quality of experience," *Proceeding WWIC'10 Proceedings of the 8th international conference on Wired/Wireless Internet Communications*, vol. 38, no. 5, pp. 216 –227, 2010.

[66] Y.-F. Ou, T. Liu, Z. Zhao, Z. Ma, and Y. Wang, "Modeling the impact of frame rate on perceptual quality of video," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pp. 689–692, IEEE, 2008.

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

[68] H. Riiser, P. Vigmostad, C. Griwodz, and P. Halvorsen, "Bitrate and video quality planning for mobile streaming scenarios using a gpsbased bandwidth lookup service," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pp. 1–6, IEEE, 2011.

[69] T. C. Thang, Q.-D. Ho, J. W. Kang, and A. T. Pham, "Adaptive streaming of audiovisual content using mpeg dash," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 1, 2012.

[70] M. Wan, D. Wang, M. Goldman, M. Taddy, J. Rao, J. Liu, D. Lymberopoulos, and J. McAuley, "Modeling consumer preferences and price sensitivities from large-scale grocery shopping transaction logs," in *Proceedings of the 26th International Conference on World Wide Web*, pp. 1103–1112, International World Wide Web Conferences Steering Committee, 2017.

[71] Markus Fiedler, Sebastian Möller, and Peter Reichl. "Quality of Experience: From User Perception to Instrumental Metrics," in Report from Dagstuhl Seminar 12181, 2017.