NTNU
Norwegian University of
Science and Technology

Measuring perceptual quality in Internet television

Mathias Gjerstad Lervold

Master of Science in Communication Technology
Submission date: June 2009
Supervisor: Andrew Perkis, IET

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

Problem Description

Quality of Experience (QoE) aims at measuring the user perceived quality of an audiovisual presentation. When watching streaming videos over the Internet several factors may contribute in lowering the QoE for the end user. From the video is captured and produced at headend, transferred over the open Internet, and finally viewed on a screen connected to a computer at the user's end, monitoring the QoE is a major challenge for service providers. Standardization bodies, such as VQEG and ITU-T have developed standards for quality assessment of video. For IPTV there exist a variety of QoE assurance products, but at the present time there are no commercial end-to-end QoE solutions for Internet television. Studies have shown that QoE is one of the leading causes for customer churn, thus a service provider's ability to monitor and improve its level of QoE is key for the business. This thesis therefore aims to answer the following question:

How can we receive feedback to improve the QoE in an Internet television service?

The questions should be addressed following three phases; 1. An extraction of the critical QoE factors, 2. Revealing which parameters that can be controlled and improved by the service provider, and 3. Finding out how the service provider can receive the feedback.

The practical tests should be performed using Absolutt Fotball as a case considering the service architecture provided by Move Networks.

Assignment given: 13. January 2009
Supervisor: Andrew Perkis, IET

[1] Winkler, S. Digital Video Quality: Vision Models and Metrics, 1 ed. Wiley, Mar. 2005.

[2] Move Networks, http://www.movenetworks.com

Assignment given: 13. January 2009
Supervisor: Andrew Perkis, IET

Abstract

In this thesis we have evaluated the Quality of Experience (QoE) of Internet television through user tests of Absolutt Fotball. Absolutt Fotball is a Norwegian live football streaming service powered by adaptive streaming from Move Networks Inc.

In our tests we found that the users rated the overall quality better than other Internet video services, such as Youtube, TV2 Sumo and NRK Nett-TV, but worse than football on TV. The main problems were coding artifacts, such as blurring, edge ringing and color bleeding, as well as problems with the smoothness of playback. Response time and adaptation period were in general satisfactory; all users preferred adaptive streaming with quick starts and no interruptions over traditional streaming with constant quality and buffering in the start and sometimes during sequences.

The tests also revealed that factors other than video quality could have significance in the user's overall QoE. Most notably was the delay from other live services, such as SMS updates, radio and live updates on the Internet. We also found in our analyses tendencies of content and context dependencies to the QoE. E.g. the result of a user's favorite team, as well as his/her viewing environment, could have an impact on his/her perception of the quality.

In order to improve the QoE the service provider should evaluate the encoding stage in particular. By increasing the bit rate of the encoding, many of the problems related to coding artifacts and smoothness of playback could be reduced. The client should be optimized with regards to adaptation period, response time and live-delay, however there is a compromise to be made with the robustness and reliability of the media player.

The service provider can receive feedback on the QoE in three stages: 1. Full reference objective quality assessment at headend, such as VQM, 2. Bit rate statistics from the clients, and 3. An extended user profile and a QoE tool at user end. The proposed QoE tool in the form of a menu could include guides and tests related to user equipment and viewing environment, real-time feedback and support chat related to video quality problems, and service personalization in relation to quality/price and features.

We found that controlling the QoE in Internet television is very difficult. QoE monitoring is however possible for the service provider, but a true end-to-end solution would require a better integration of client and user than is today.

Contents

List of Figures

List of Tables

Abbreviations

FR	Full Reference quality assessment metrics
HD	High Definition television format
HVS	Human Visual System based quality assessment metrics
IPTV	Internet Protocol Television
MOS	Mean Opinion Score
NR	No Reference quality assessment metrics
PSNR	Peak signal-to-noise ratio
QoE	Quality of Experience
QoS	Quality of Service
RR	Reduced Reference quality assessment metrics
SD	Standard Definition television format
VoD	Video on Demand

Preface

This thesis holds the final outcome of my master's study in the field of multimedia signal processing at the Norwegian University of Science and Technology (NTNU). The report serves as a documentation of my work during the study, which was conducted in the spring of 2009. The assignment was suggested in cooperation with Erik Aaberg of Schibsted ASA and given by Professor Andrew Perkis at the Department of Electronics and Telecommunications, NTNU.

Working with QoE in Internet television has been a challenging and interesting task. QoE is becoming increasingly important for commercial video on the Internet, and at the same time little effort has been made in standardization of QoE assurance and monitoring for Internet television. In this sense the author was required to analyze existing work in related subjects in order to produce new and novel solutions tailored for Internet television, with its challenges and opportunities.

I would like to thank my supervisor Andrew Perkis for his feedback and support throughout the semester. Erik Aaberg of Schibsted ASA and Ketil Moland Olsen of Bergens Tidende have been supportive and very helpful in providing a lot of useful information on Absolutt Fotball. I would also like to thank the test users, many of whom close friends, who participated in the QoE tests and responded to our surveys, as well as Kine Walbye for final proofreading of my report. Finally, thanks to the participants and organizers of the NTNU Audio Visual Forum for inspiration as well as constructive feedback.

Trondheim,
June 2nd, 2009

Mathias Lervold

Chapter 1

Introduction

1.1 Background

Quality is in many cases a subjective measure, and a person's perception of quality is dependent on factors such as anticipation, earlier experience, preference, surroundings and state of mind. Quality is an elusive term describing certain properties of a service and a user's satisfaction with it. However in order to ensure a high degree of service quality it should also be objectively measured, thus we must find some universal parameters that can be related to the perceived quality.

Until recently Quality of Service (QoS) has been the Holy Grail by which IP services has been measured. QoS is usually a set of network parameters describing the bandwidth, transmission jitter, packet delay and loss, and other issues affecting the effectiveness and reliability of the network transmission. For the task of transmitting data from point A to point B this might be a good description of the quality, however when we evaluate the content in multimedia services the effect of transmission errors on the perceived quality at user end is not as straightforward.

Video over the Internet has become one of the biggest bandwidth consumers, and according to the trends transmission of video will dominate the IP traffic in the future [22]. Measuring the quality of the video from a user's point of view is a much more complex task than just measuring the network QoS. QoS is an important part of it, but at the same time we must take into account how the video is captured, processed and encoded before transmission, and also how it is decoded by the client and finally presented to the user. This leaves us with a calculation of many unknowns.

Quality of Experience (QoE) is a term describing the perceived quality from a user's point of view. It describes the user's overall satisfaction with the service, in this case video and audio, and takes into account all the unknowns mentioned in the previous paragraph. Many things can contribute in lowering the level of QoE; audiovisual content captured using bad equipment, errors in the encoding process, packet loss during transmission, or a bad viewing environment.

During the last years, video over the Internet has developed from grainy home made videos on Youtube and similar Internet video distributors, to prime-time shows, sporting events and major political events in HD on Internet television services. The mainstream media has seen business opportunities in taking their content online, where they can profit from pre-paid services or ads. This, however, brings a whole new focus on the quality of the services provided, since customers don't want to pay for a low quality service. The focus should also be shifted from QoS parameters, to the end user's experience of the quality.

1.2 Motivation

In most Internet video and music services today the customer chooses a suitable bit rate, which is a QoS parameter [35]. As briefly discussed, this does not represent the complete picture of how the user perceives the audiovisual quality, and in Internet television services where the quality is of high importance, a method of measuring the QoE is needed [31, 62, 68].

As mentioned in 1.1 measuring and ensuring the best possible QoE is not a straightforward task. There is a variety of video codecs, each with its advantages and disadvantages. Internet television is transferred over the open Internet, which means there are no guarantees of QoS. And finally there is a myriad of different clients as well as user equipment for presentation.

A variety of methods for video quality assessment has been proposed and standardized [21, 67, 70], and for IPTV services there exists headend as well as end-to-end QoE assurance products that can be implemented by the service provider [34, 63]. However at the present time there are no similar standards for Internet television.

Move Networks delivers an Internet television streaming service to some of the biggest TV broadcasters in the US [9]. They promise their customers "The quality difference", but their claim is difficult to measure, as there are no objective parameters that can be tested. In this thesis we chose to study their streaming service and QoE through Absolutt Fotball, delivering live football streaming over the Internet [20, 25, 26].

1.3 Purpose

This leads us to the present study. We believe that there is a need for a way of objectively measuring the QoE in Internet television services. We believe that if it is measurable, the QoE can also be improved, which can increase the profitability for the business.

We have therefore cooperated with Schibsted ASA and their football streaming service, Absolutt Fotball, to conduct user tests on commercial Internet television streaming to reveal whether there is a need for improved QoE, and if so, how this can be solved by QoE feedback to the service provider.

More concisely, we wish to answer the following question:

How can we receive feedback to improve the QoE in an Internet television service?

This will be divided into three parts; 1. An extraction of the critical QoE factors, 2. Revealing which parameters that can be controlled by the service provider, and 3. Finding out how the service provider can receive the feedback.

The study will primarily be focused on the Absolutt Fotball case and the service architecture provided by Move Networks.

1.4 Outline

The rest of this thesis is structured as follows. Chapter 2 gives an overview of the background theory we researched concerning Internet television, Quality of Experience and quality assessment. In chapter 3 we give an overview of the case that has been the focus of this study, as well as mini case studies on QoE assurance within IPTV. In chapter 4 we discuss the methodology and the experiments that have been conducted, and a presentation of the results of the study. We discuss the process, results and analysis in chapter 5, further work in chapter 6, and finally some general conclusions are given in chapter 7.

Chapter 2

Theory

2.1 Internet Television

The focus of this thesis will be on the QoE in Internet television. In this section we will give a definition of Internet Television, describe its basic architecture and technology, as well as look at its commercial opportunities.

2.1.1 Defining Internet television

Video over the Internet is one of the fastest growing IP services today, and will according to Cisco Systems [22] account for nearly 90 percent of all consumer IP traffic in 2012. It's common to describe such services as IPTV or Internet video. In this thesis however, we will use the two to define Internet television, which lies somewhere in between. We will make a clear distinction between what is described in [34] as IPTV from service providers and broadband video, which we will define as Internet television.

Table 2.1 gives a quick overview of the differences between IPTV and Internet television based on the differences between service provider and broadband video in [34]. The most important technical differences between Internet television and IPTV from service providers are the network QoS and user clients. Service providers usually have privately managed networks for their IPTV services where they can guarantee a certain level of QoS. The user may watch live television shows as well as order Video on Demand (VoD) from an Interactive Program Guide (IPG). For Internet television the user streams or downloads the video over the open Internet on a best effort basis, and thus have no guarantee of QoS. The client is often a flash-based streaming application or plug-in on a web site, or a custom desktop media player that downloads or streams the content to the user's computer. (More on the technical specifications in 2.1.2 and 2.1.3)

IPTV is more or less a revolution of the cable/satellite services with a walled garden approach, where users pay for the channels and videos they want to watch from the same service provider. Internet television however, is more open and the user is free to roam among a large number of free, pre-paid or ad-supported services from a multitude of content aggregators. In this fashion Internet television represents a more disruptive alternative to the old TV broadcasting model, and thus the business models will have to change. (More on this in 2.1.4)

IPTV	INTERNET TELEVISION
Controlled network	Public internet
Generally available on a TV	Generally available on a PC
Special decoder to TV-set (hardware)	Flash/plugin/media-player (software)
QoS guaranteed	Best effort basis
Broadcast channels	Streams
Video on Demand	Downloads
Interactive Program Guide	Web site
Prepaid, advertisement	Prepaid, advertisement, free
Walled garden	Open access

Table 2.1: *IPTV from service provider vs. Internet television, based on service provider vs. broadband video in [34].*

Figure 2.1 describes the value chain for video over IP [36] with the IPTV from service provider approach and the Internet television approach. What is typical in IPTV services is that the service provider controls most of the value chain, whereas a service provider of Internet television only controls the content aggregation and often times the content selection. This will play an important role in QoS and QoE as we will see in section 2.2.

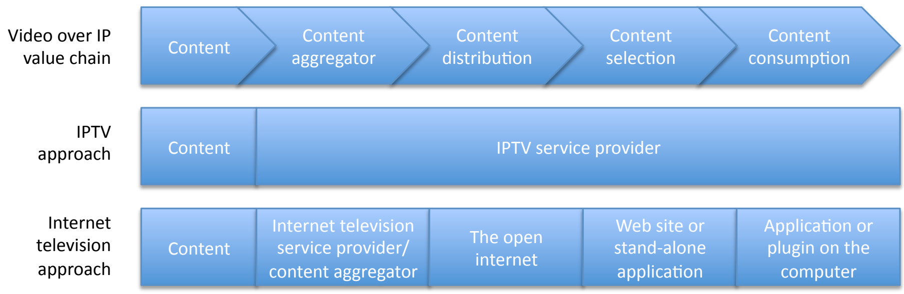


Figure 2.1: *Video over IP value chain with IPTV approach and Internet television approach.*

What separates Internet television from other Internet video services, such as Youtube, is the production and business model. Internet video is usually user generated or produced by non-professionals, while Internet television is provided by TV-networks. Internet video has so far been difficult to profit from and is usually free or ad-supported (in user interface), while Internet television in a much higher degree is dependent on income from subscriptions or ad revenue. Figure 2.2 gives an overview of the differences whereby Internet television and IPTV is separated by how the network is controlled and general openness, and Internet television and Internet video is separated by production and business model. Internet television therefore lies somewhere in between Internet video and IPTV on the quality scale. Internet television has also been known as Web TV.

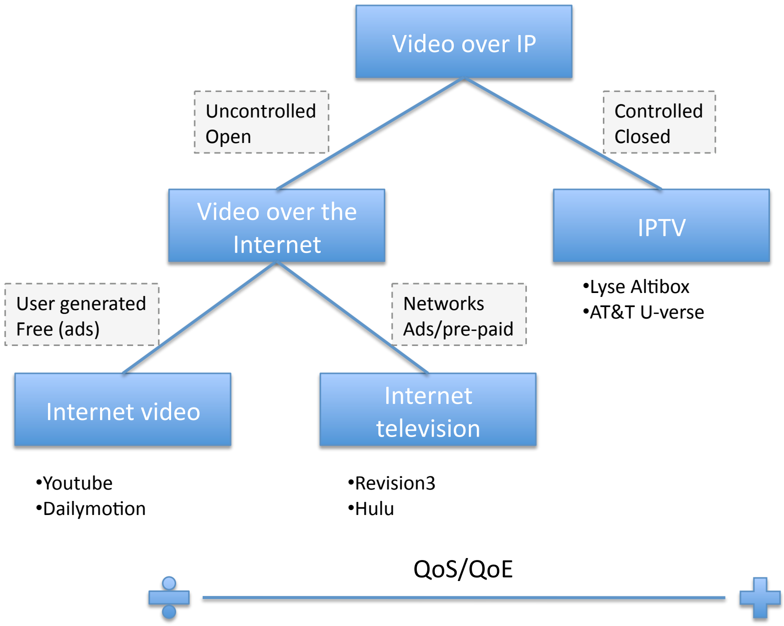


Figure 2.2: *An overview of how Internet television relates to IPTV and Internet video.*

2.1.2 Architecture

Internet television is usually based on a unicast architecture from a media server to several client nodes over TCP/IP. It's also possible to stream Internet television using multicast over UDP, however this is less widespread. The unicast architecture requires a great amount of bandwidth from the media servers, and peer-to-peer technologies are researched and developed to tackle this problem in the future [34, 36].

When looking at the Internet television architecture we can roughly divide the user clients into Flash, Silverlight or similar web-based video applications and plug-ins, and custom desktop video applications. Figure 2.3 shows how the network layers differ for web-based and desktop players, and table 2.2 highlights the main differences between the two.

Web-based video applications and plug-ins

Most web-based video players are based on the TCP/IP architecture where the video stream typically is sent from an HTTP-server in small packets. The packets are received at the user's computer and put into the right order in the video buffer, which can be played progressively. A simplified overview of the network stack can be seen in figure 2.3 [64].

The strength of TCP is its reliability. Packets are received in the right order, however there is no guarantee that the packets will be received on time. If packets are lost, the media-stream may stall until the packets are retransmitted. A buffer may fix the problem with lost and delayed packets to some extent, but this means that the video stream will never be "live". Because it is unicast, video over TCP/IP is technically more of a VoD service than a streaming service.

Adobe's Flash video player is used by a large number of Internet video/television services, among others Youtube and Revision3. Microsoft has developed a competitor to Flash called Silverlight, and there are other custom made players, such as Move Media Player, which we will take a closer look at in 3.1.1.

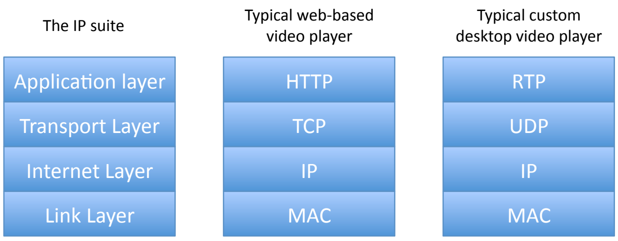


Figure 2.3: *Typical layer architecture for web-based and custom desktop video players.*

WEB-BASED	CUSTOM DESKTOP BASED
Unicasting	Unicasting and multicasting
Platform independent (mostly)	Platform dependent
Flash/Plugins almost ubiquitous	Must install custom player
Some delay on live programs	Short or no delay on live programs

Table 2.2: *Web-based video players vs. custom desktop based video players.*

Custom desktop video applications

Custom desktop video applications such as Windows Media Player, QuickTime or VLC may stream video using RTP over UDP, as seen in figure 2.3. UDP transmits datagrams without making an implicit handshake between sender and receiver, as in TCP, thus making it an unreliable service. There is no guarantee that the packets will arrive in the right order or arrive at all, however in time-sensitive applications such as video streaming, it is preferred to drop a packet rather than to wait for its retransmission. RTP supports both unicast and multicast [31], which can ease the bandwidth requirements from the server.

Internet QoS

Internet television is transported over the Internet on a best effort basis [35]. That means there is no guarantee of QoS from the media server to the user client.

Issues such as packet loss, packet delay, transport stream jitter and limited bandwidth must be taken into account when designing an Internet television service. The implications of this on the video stream will be further investigated in 2.2.

2.1.3 Codecs and compression

Because of the bandwidth limitations of the Internet, good compression algorithms are needed to minimize the data stream while maintaining the best possible video quality. The video is encoded and compressed before it is sent from the server, and then decoded at the user client.

Most encoders today are based on so-called hybrid video coding, i.e. compression based on a hybrid of motion compensation and transform coding. The aim of these algorithms is to reduce spatial, temporal, perceptual and statistical redundancies, thus reducing the bit rate while maintaining a satisfying level of quality. To reduce spatial redundancies in a macro-block one can use the information of surrounding blocks. Estimating block motion and using a mix of reference frames and predicted frames can reduce temporal information. Perceptual redundancies can be reduced by chroma sub sampling and quantization based on the human perception. Statistical redundancies can be reduced by Huffman or arithmetic coding [42]. Figure 2.4 shows a typical coding and decoding process for a hybrid coder [56].

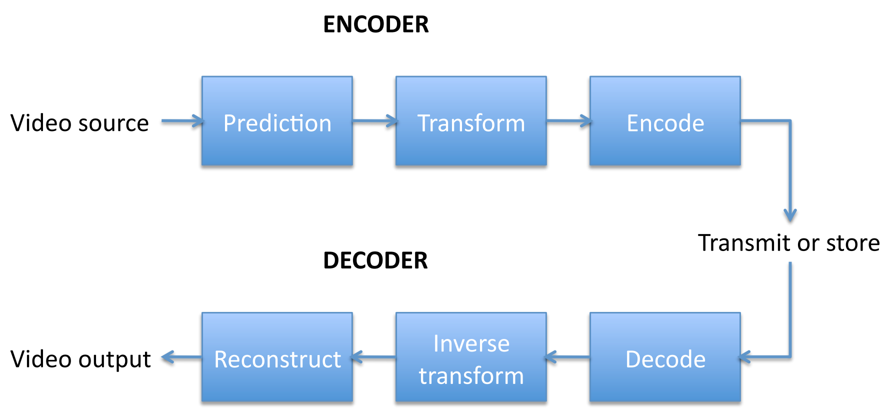


Figure 2.4: *Typical coding and decoding process for a hybrid coder [56].*

By codec we mean the combination of a coder and a decoder. In this thesis we will use the term in relation to the various chipsets and decoding standards used in video compression. The most used video codecs today for Internet video [37] are On2's VP6 and VP7, Sorenson 3 and Spark, H.264/AVC, Windows Media Video (WMV) 9 (VC-1), and DivX/Xvid.

On2 Technologies

On2 VP6 is one of 3 codecs supported by Adobe's Flash video player [6] and is also licensed to JavaFX. The codec delivers resolutions up to 720p HD and has surpassed WMV and many implementations of H.264 in PSNR (Peak signal-to-noise ratio) comparisons [50].

VP7 is licensed to among others Skype2.0 and Move Networks and supports resolutions up to 1080p Full-HD. It promises up to 50% image quality improvement over VP6 and better PSNR performances than competing codecs on data rates from dial-up (28.8 Kbps) to DVD and HD [51].

VP8 is the latest generation video codec from On2 and is supposed to replace VP7. It focuses on lower bandwidth usage while maintaining a video quality equivalent to leading H.264 implementations [52].

Sorenson Video Codecs

Sorenson video 3, also known as SVQ3, is exclusively licensed to Apple's QuickTime. It supports variable bit rate (VBR) compression, bi-directional prediction and packet loss correction [14]. Sorenson Spark is used by Adobe's Flash video player and has been incorporated into more video formats than any other video standard. It is resolution and hardware independent [15].

H.264/AVC

H.264, also known as Advanced Video Coding (AVC) or MPEG-4 Part 10, is an open video coding standard developed and standardized collectively by both the ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). It is said to match the best possible MPEG-2 quality at half the data rate and to deliver superior video quality to WMV-9. H.264 supports resolutions and bandwidths from videos on mobile phones at 40 kbps to 1080p Full-HD at 8 Mbps [1].

Because of its openness H.264 has been implemented in a large number of versions and different applications [10]. Some of the codec's features include: entropy coding, intra- and interframe prediction, transform coding, deblocking filter and network-friendly features [42].

WMV-9 and VC-1

WMV-9 is the video codec that is part of the Windows Media 9 series developed by Microsoft. Its aim is to enable effective media delivery through any network. It delivers video at bit rates and resolutions from 10 kbps for mobile applications to 4-8 Mbps HD video. It uses block based motion compensation and spatial transform scheme, similar to MPEG standards. It is said to deliver comparable video quality to H.264 and have some complexity advantages [58].

VC-1 is a proprietary video codec based on WMV-9, developed by Microsoft and standardized by Society of Motion Picture and Television Engineers (SMPTE). It is used in HD-video, such as Blu-Ray, as well as mobile applications, such as DVB-H [19]. VC-1 is said to offer a quality-complexity tradeoff advantage over H.264, especially in high-definition video services [43].

DivX and Xvid

The DivX codec is an implementation of MPEG-4 Part 2. Its latest release is 6.8.2. It has four certified profiles from 176 x 144, 15 fps Handheld at 200 kbps to 1080p Full-HD at 8 Mbps [2].

Xvid is an open source video codec similar to DivX and was created to promote the adoption of open standards. The codec is optimized towards high quality offline, multi-pass compression for storage and archival purposes and is claimed to deliver superior quality over any other implementation of MPEG-4 Part 2 [23].

2.1.4 Growth and business opportunities

Broadband Internet connections are increasing all over the world and Cisco Systems expect the data traffic over IP networks to double every two years through 2012 [22]. Internet video will be the biggest drive for this increase in traffic and is expected to constitute 50% of the total IP traffic in 2012. Better video compression and increased downstream access speed for consumers will allow streaming of HD video and in general provide a higher quality viewing experience.

Products like Apple TV and Boxee are contributing to enhancing the user experience as they allow users to view Internet video on their television sets and browse or search through content in a more appealing fashion than in a regular Internet browser on the PC. Studies have shown that an increasing number of users wish to record/download TV-shows on their PCs if they can watch it on their television set contra on their PC monitors [36].

Internet television represents a new way of watching TV, where users can watch their favorite shows and videos any time, anywhere, and have access to millions of shows at the same time. The power is shifted from broadcasters and platform operators to producers and consumers. Also, whereas regular TV broadcasting is based on a fixed cost for the service provider, Internet television is based on marginal cost depending on the amount of viewers of each TV-show [34]. The old business model for TV operators will have to change.

According to Kagan research [41] "Internet video has long held potential to become a disruptive media force, allowing the entry of smaller and independent players worldwide and changing television viewing patterns and habits."

We will take a closer look at the emerging opportunities within ad revenues and niche markets, as well as concerns regarding copyrights and security.

Ad revenue

There has been a rapid growth in Internet advertising the last ten years. According to [33] it increased 28% compared to 4,8% for television, and is expected to rival TV broadcasting advertising by 2011.

One of the reasons for this remarkable growth is its superior targeting capabilities. The Internet is a two-way communication, where the advertisers can track how many people have clicked the links to their ads, and Internet television service providers can provide usage data to tailor the ads for their users. Studies have shown that viewers are 44% more engaged with commercials they watch online [47]. People in general are much more prone to watch/click an ad of something that matches their interests, and Google has had great success with this approach, based on people's searching habits.

Another reason why broadcasting advertising has become less attractive is the prevalence of the Digital Video Recorders (DVRs), such as TiVo. DVRs let the user skip through commercial breaks and surveys have shown that 62% of DVR users say this is one of the top benefits of the device [33].

The long tail

The long tail is a phrase popularized in 2004, by Chris Anderson of Wired magazine, which reflects the popularity of certain items when there is an abundance of choice. More specific it is a statistical distribution dominated by a few very popular items with mass appeal, and with a tail of a great number of less popular items with niche appeal [34]. The long tail is illustrated in figure 2.5.

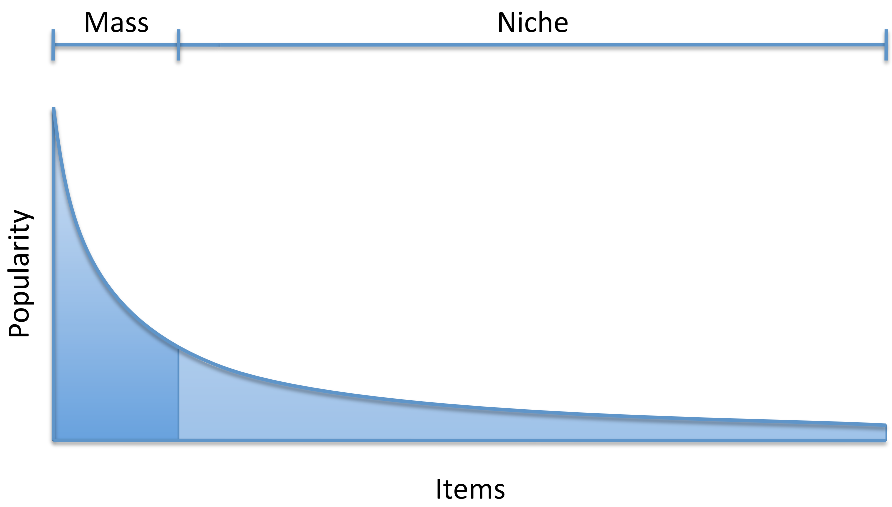


Figure 2.5: *The long tail with a few popular items with mass appeal and a large number of less popular items with niche appeal.*

There are a million niche markets, something which service providers may make great profits from. Niche markets may be old, local or foreign TV-shows and movies, independent films, or videos addressing special interests. New ways of searching and browsing for content enables the user to find new niche content, and similar to ads the service provider/content aggregator can tailor niche content specific to user based on his/her usage data and viewing history. Because of the marginal cost of Internet television and the fact that anyone anywhere in the world can have access to the service, it has made these niche markets much more profitable.

Security

One of the biggest challenges of the new digital era is security and rights management of media. What can be seen or heard on a computer, can be copied.

Digital Rights Management (DRM) is said to be a technological solution to a social and business problem [34]. DRM can be used to limit access to the content to specific users, on specific clients within given timeframes. Microsoft has incorporated this into their WMV format, however nothing on the Internet is 100% secure and their system has already been compromised.

There is a delicate balance between how much control the service provider wants over their content and how much freedom the users want with it. If the users are too restricted in how they may use the service they pay for, they will change their service provider. Also, restricting digital distribution disregards the considerable benefits of being able to maximize the audience [34].

Another way of somewhat controlling the distribution is by watermarking [34]. This may either be visible watermarking, such as a logo in one of the four corners of the screen, or invisible. Invisible watermarking or fingerprinting is robust to recompression and manipulation, and can be recognized and linked to its proper source.

2.2 Quality of Experience

In this section we will take a closer look at QoE. We will define the term as well as what we mean by video quality. We will look into different degradations to QoE and how it may have been caused. Finally we will look at the impact of QoE on the business model for Internet television, quickly describing the opportunities for service differentiation as well as overall growth and profitability.

2.2.1 Defining Quality of Experience

In section 2.1.2 we briefly discussed network QoS and how there is currently no guarantee for QoS on the open Internet. QoS is easily measurable with quantifiable parameters such as packet loss, jitter, delay, throughput and reliability [35]. Although these parameters may give an indication on the quality of the video streaming service, they do not paint a complete picture. QoS does not say anything about the quality of the captured video, the quality of the encoding process or the decoding and user equipment.

This is where QoE comes in. A definition of QoE is "The overall acceptability of an application or service, as perceived subjectively by the end user."[44] QoE aims to describe the quality from the video is captured until it is viewed by the user, which makes QoE a whole lot more complex compared to QoS, as illustrated in figure 2.6.

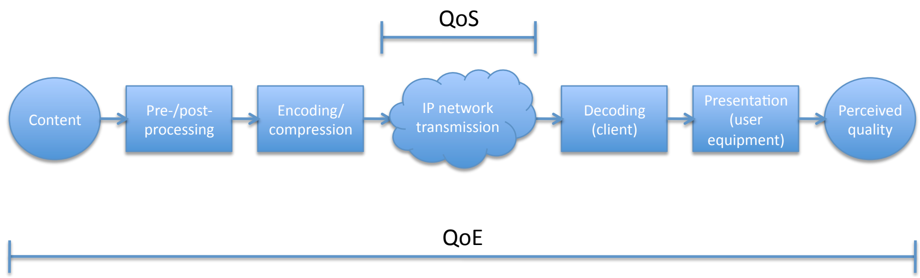


Figure 2.6: *QoS vs. QoE for video streaming, based on figure 3 in [31].*

QoE must take into account the quality of the captured video, degradations to the video in pre-/post-processing or in the encoding process, the network QoS, and the decoding and presentation of the video by client and user equipment.

2.2.2 Video quality

There are many factors contributing to a user's experience of the perceived video quality. Factors like quality expectations, how and where they watch it, the quality of their TV or PC monitor, former experiences, content preference etc. However, in this thesis we will focus on the perceived visual fidelity of the video. Visual distortions such as blurring, color problems and noise may degrade the video fidelity. Also, studies have shown that people tend to perceive colorful, well-lit, sharp pictures with high contrast to be of higher quality than dark and blurry pictures with low contrast [69]. Various scenarios of reduced video quality (QoE) are presented in 2.2.3.

Today there is no single standard to evaluate the video quality. There are several subjective and objective quality models with slightly different focus that are more or less suitable for different scenarios. Most quality models are based on a mean opinion score (MOS) that rate the video on a 1-5 scale where 1 equals bad quality and 5 equals excellent quality. Quality assessment methods are further discussed in 2.3.

The Video Quality Experts Group (VQEG) has made efforts to standardize video quality. They have conducted formal evaluations of video quality metrics for use on television as well as mobile applications. They are currently planning tests for HDTV [21].

2.2.3 QoE scenarios

As we briefly discussed in 2.2.1 there are many stages along the delivery path from capturing of the content to the viewer's perceived quality where the video quality may be degraded. We will categorize the QoE into four categories: audiovisual quality and sync, video quality, audio quality, and interactivity. Within these categories we will look at different QoE scenarios based on the distortions/degradations and what may have caused it.

Audiovisual quality and sync scenarios

Not satisfying overall quality. Overall quality may include such audiovisual parameters as resolution, details, dynamic, audio crispness, video colorfulness and contrasts, etc. Bad equipment and recordings in the capture-stage may degrade the overall quality. Bad pre-/post-processing as well as too much compression will also damage the video. Finally, if the user's equipment is of low quality, e.g. can't handle the resolution or the dynamic range, the experience of the overall quality will be lowered. Overall quality is usually measured on a MOS-scale with or without references (more on references in 2.3.3).

Audio and video out of sync (lip sync). Lip sync is one of the most important factors for the overall perception of QoE [62], and if the delay between audio and video is too large, it is perceived to be out of sync. Studies have shown that viewers perceive audio and video to be in sync with delays up to 80 ms [69]. There is also a higher tolerance for video ahead of audio, which may be due to the fact that light travels faster than sound and therefore seems more natural. As in the previous scenario the production stages are critical, as well as the presentation. Also, if packets are lost or there is much jitter in the transmission over the IP network, this may affect the sync-clock in the client, which can get audio and video out of sync. Lip sync could also be caused by jerkiness in the video, however this will be addressed later.

Too long adaptation time in adaptive streaming. Adaptive streaming is a way of adapting the video presentation to the current network throughput (as we will look further into in 3.1.1). Usually the stream will begin at a low quality and gradually adapt to the best possible quality by communication between client and server. The adaptation time is the time it takes to reach the best quality, and for short clips this may be too long to get a good user experience.

Video quality scenarios

Black or blue screen may occur for a short time ($<$0.5 sec) or for longer periods during video streaming as if the signal has been lost. It is usually caused by some defects during the capturing or the encoding/decoding process, or a packet loss during transmission.

Frame skips and freezes may seem like small glitches in the video where either some frames are dropped and the video suddenly "jumps" ahead, or the frame freezes for a moment. These types of errors are usually caused by problems with transmission, such as throughput, jitter or packet loss/delay. For hybrid coders with forward and backward prediction, the loss of a reference frame will freeze some parts of the picture.

Jerkiness is a degradation to the smoothness of the playback. That is, instead of running smoothly, the video stops a little bit at each frame as if the frame rate is low. This could contribute to making it seem as though the audio and video is out of sync (lip sync). Jerkiness is usually caused by some error in the pre-/post-processing or encoding stages, transmission jitter or poor user equipment (processor speed, memory, graphics card etc).

Block or slice losses may appear in block-based video coders, such as H.264. If a packet is lost during transmission, the frame may miss a block or slice.

Blockiness is often the result of too harsh compression or poor coding algorithm for block-based video encoders. It may seem as if there are many small blocks in the picture that doesn't match in the edges.

Blurring, ringing or loss of detail usually happen during the production stages; from capturing to encoding. A blurred image seems out of focus and lacks sharp edges and details. If the camera is out of focus during capture, there is little one can do to make the picture sharper. Also, fast moving objects may seem blurred because of the limited number of frames per second that can be taken by a CCD sensor in a camera. Harsh compression will also leave out more and more details in the picture, making it seem blurred and unsharp. The ringing artifact happens around high-frequency edges of a picture, i.e. areas where there is high contrast. Instead of sharp edges, ringing distorts the edges and the picture loses some of its detail. This usually happens during the encoding-process.

Color bleeding can be seen as an equivalent to blurring, only in the color channel (chrominance) instead of black-white (luminance). In a picture with color bleeding it may seem as the colors more or less flow into each other instead of having sharp borders. This is usually caused by errors in pre-/post-processing of the video or harsh compression.

Video artifacts or noise may have analog or digital sources. So-called mosquito noise can be described as an additive overlap of ringing and motion compen-

Figure 2.7: *Overview of where degradation to video stream may occur in the service delivery, based on [32, 62, 57, 67, 68, 69].*

sation artifacts from the encoding and appear as flickering in the luminance and chrominance channels over time. Analog noise during capturing or pre-/post-processing of the video may come from electric circuits and appear in the picture.

De-/interlacing effects can happen in the pre-/post-processing stage when the video is converted between formats. It looks as a distortion of the edges of moving objects because every second line from top to bottom is slightly skewed.

Audio quality scenarios

Silence for longer periods may be due to some error in the capturing stage or encoding, but there may also be a problem with the presentation if the user's equipment can't process the audio format.

Choppy audio is the equivalent to jerkiness in video and may be caused by bad encoding/decoding, not enough bandwidth or transmission jitter, but also with the presentation, as in the previous scenario.

Noise or loudness is usually caused by some error in the production stages. Bad microphones during capturing, gaining or compression during pre-/post-processing, or too harsh data compression during encoding.

Interactivity scenarios

Skip delay is the delay from the slider on the movie-progress bar (time) is moved to a new position, e.g. from 1:20 to 3:40, until the movie starts playing. The delay is mostly dependent on the available bandwidth to get the new data, but also the client buffer and user equipment's ability to process the new data fast.

Startup delay is the delay at the beginning of a video stream, from you press "play" until the video starts. The delay depends mostly on network throughput, client and user equipment, as in the previous scenario.

Figure 2.7 shows an overview of common QoE problems for Internet television related to where in the service delivery they may be imposed, based on [32, 62, 57, 67, 68, 69]. It is important to note that there are primary and secondary reasons for a QoE problem. One example might be that low bandwidth will result in low quality video, however it is the compression/encoding that is the primary reason and the low bandwidth is a secondary reason. Figure 2.7 focuses on the primary reasons.

The capture stage is where we get our original content, thus a bad recording caused by erroneous equipment or human errors yield a bad end result. During the pre-/post-processing it is possible to change certain properties or enhance the recording, as well as adding effects to or scale/transform the video or audio. If not done correctly, the video may be further degraded. As we can see the encoding stage is very critical for the end result, thus good compression algorithms are very important (see 2.1.3). Several coding artifacts may appear from harsh compression. However the production stages (headend), from capture to encoding, are usually under control by the service provider.

The transmission is difficult to control, but can be monitored and QoS parameters, such as network throughput/bandwidth, transmission jitter and packet loss/delay, can be extracted and related to QoE issues. The client is usually designed by the service provider, but its performance is often dependent on the user equipment, over which the service provider has no control. A badly designed client or low quality equipment may not be able to process all the content in a timely manner, thus degrading the perceived quality even further.

Because of all these factors it is very difficult for service providers to know exactly how good the perceived quality is for the end user, and how they can improve their QoE.

2.2.4 The impact of QoE on the business model

Studies on QoE for IPTV service providers have shown that video quality is the primary reason for customer churn as well as a major reason for customer calls [63]. Video quality is a competition differentiator, something of which service providers strive to control and maximize [60]. A high level of QoE translates into more satisfied customers, who will spend more time on the service, both per visit and the number of visits [48].

Customer expectations

Studies by Accenture [27] found that issues with QoS/QoE will be the biggest obstacles for IPTV adoption in the next years, and the same may be true for Internet television. Customers are used to 99,999% reliability from conventional broadcast TV, and at the present time Internet television cannot promise the same level of reliability. Also, viewers' quality expectations tend to be higher for a paid service than free content, thus if the service provider doesn't deliver up to user expectations, it may suffer audience retention [59].

If the service provider has good control over the level of QoE they can deliver, they will better know how they can meet customer expectations and how they can increase their customer satisfaction.

Service differentiation

Customers might not have the same expectations or the same willingness to pay for Internet TV, and in order to reach as many as possible, a service provider might consider service differentiation. A high level of QoE might be offered to their premium customers on a pay-per-view or subscription basis, and a lower level of QoE might be offered for customers with less willingness to pay, while still profiting from ads.

Ad revenues

Until now big advertisers that buy traditional TV broadcast time, have been wavering towards advertising on Internet television [59], and a major reason for this seems to be the quality of the content on the Internet. Well known brands don't want their name to be affiliated with low-quality, grainy videos. However, as the streaming technology gets better and the content reaches TV-level quality, this is a great source of ad income, as discussed in 2.1.4. The service provider who can promise the highest level of QoE, will have a major advantage over the competition.

According to Move Networks [46] "High-quality, uninterrupted viewing means that viewers will watch your content longer and more frequently. You will see increased audiences, increased ad impressions, and an overall increase in profitability."

2.3 Quality assessment

There are many methods for assessing the quality of a video, and in this section we will give an overview of the most popular subjective and objective quality assessment methods. We will also look at how reference information is used in quality assessment, and what the advantages and disadvantages are for the different methods.

2.3.1 Subjective quality assessment methods

There is a variety of subjective quality assessment methods differing in use of reference, stimulus and rating-method. The most used methods are *Double-stimulus impairment scale (DSIS)*, *Double-stimulus continuous quality-scale (DSCQS)* and *Single stimulus continuous quality evaluation (SSCQE)*.

Double-stimulus impairment scale (DSIS)

In DSIS the subjects are shown a reference sequence and a test sequence with the reference always first. The impairment of the test sequence is then rated discretely from "very annoying" to "imperceptible" (five grades). DSIS is well suited for evaluating clearly visible impairments such as artifacts caused by transmission errors, however it is limited to short sequences (10 seconds).

Double-stimulus continuous quality-scale (DSCQS)

DSCQS presents the reference- and test sequences twice in a randomly alternating fashion, so the subject doesn't know which is which. The sequences are rated on a continuous quality scale from "bad" to "excellent" (five points). This method is very sensitive to small differences in quality and also useful when it is not possible to provide test sequences that exhibit the full range of quality. DSCQS is also limited to short clips and criticized for putting the subjects in a situation too far from a normal viewing environment by repeating the content (familiarity with the content).

Single stimulus continuous quality evaluation (SSCQE)

SSCQE works for long sequences up to 30 minutes where the subjects continuously evaluate the quality on a scale similar to the one used in DSCQS. A drawback for this method is that the program content tends to have an influence of the scores, and also because of differences in subject reactions, the data can be more difficult to analyze.

Other methods

Double-stimulus comparison scale (DSCS) is similar to DSCQS, however the sequences are shown only once (in random fashion) and the differences between the two are rated on a discrete 7-grade scale.

Simultaneous double stimulus for continuous evaluation (SDSCE) is similar to SSCQE where the subject is presented with a reference sequence side-by-side with the test sequence, thus making it more sensitive to sparse impairments.

Absolute category rating (ACR) uses a short test sequence, without reference, that is evaluated on scale similar to the one used in DSCQS. Since it doesn't need reference it is quite efficient compared to DSIS and DSCQS.

Degradation category rating (DCR) is similar to DSIS and evaluates the degradation to a test sequence based on an explicit reference.

In *Pair comparison (PC)* test clips from the same sequence but with different quality are paired in all possible combinations. The subjects make a preference judgment for each pair.

Subjective Assessment Methodology for Video Quality (SAMVIQ) uses a multi-stimulus approach where many test sequences are evaluated each after the other on a scale similar to DSCQS. There is an explicit reference sequence and the subjects can reevaluate each test sequence as many times as they like. The method provides reliable discrimination for both low and high quality sequences, however it is very time consuming.

	REFERENCES	COMPARISON	MOMENT OF RATING	RATING SCALE	LENGTH OF STIMULI
DSIS	Explicit reference	Double	Retrospective	5-grade impairment scale	10 seconds
DSCQS	Hidden reference	Double	Retrospective	5-point continuous scale	10 seconds
SSCQE	No reference	Single	Continuous	5-point continuous scale	Long stimuli up to 30 minutes
DSCS	Hidden reference	Double	Retrospective	7-grade difference scale	10 seconds
SDSCE	Explicit reference	Double	Continuous	5-point continuous scale	Long stimuli up to 30 minutes
ACR	No reference	Single	Retrospective	5-point continuous scale (or higher if required)	10 seconds
DCR	Explicit reference	Double	Retrospective	5-grade impairment scale	10 seconds
PC	Hidden reference (no reference)	Double	Retrospective	2-grade preference scale	10 seconds
SAMVIQ	Explicit and hidden reference	Multi stimulus	Retrospective - rating can be adapted several times	5-point continuous scale	Maximum 15 seconds

Table 2.3: *Comparison of subjective quality assessment methods based on [11, 24, 40, 49, 54, 67, 70].*

Table 2.3 compares key distinctions between the different methods, based on [11, 24, 40, 49, 54, 67, 70]. For long sequences without reference SSCQE is the most popular. For short clips DSIS is the most popular when using an explicit reference for comparison purposes, while DSCQS is used with hidden reference for quality ratings.

2.3.2 Objective quality assessment methods

Subjective quality assessment is often time-consuming and difficult to implement real-time and over a large scale. The need for automatic quality measurement has spurred the development in objective quality assessment methods. Their aim is to best assess the video quality based on objective parameters, correlating to subjective MOS scores.

There are four main approaches; data metrics, packet- and bitstream-based metrics, human visual system (HVS) based picture metrics and engineering approach picture metrics. Data metrics measure the fidelity of the signal (data) without considering the content. Packet- and bitstream-based metrics extract parameters from the transport stream and the bitstream. HVS and the engineering approach measures the fidelity of the decoded video or picture. There also exists hybrid methods, which use a combination of the aforementioned approaches.

Data metrics

Data metrics only consider the fidelity of the data and not the content or its properties. This makes it a simple, easy and fast approach to quality assessment. The most widely used methods among data metrics are *Mean Squared Error (MSE)* and *Peak Signal to Noise Ratio (PSNR)*. MSE measures the average in which the test signal differs from the reference signal, and PSNR measures the ratio between the maximum possible power of the signal and the power of noise that affects the fidelity of its representation.

A disadvantage with data metrics is that the visual degradations to the picture with a data error varies greatly with its content. Two pictures can have the same PSNR but have a very different perceived quality, as illustrated by figure 2.8 from [70].

(a) (b)

Figure 2.8: *Two pictures with the same PSNR, but in (a) the distortions are in the high frequency parts of the picture (details) and in (b) the distortions are in the low frequency parts (sky). From [70].*

Packet- and bitstream-based metrics

Packet- and bitstream-based metrics are based on parameters that can be extracted from the transport- and bitstream with no or little decoding. And because of this they require much lower bandwidth and processing power compared to methods that require the entire reference data stream or decoded video. It is also possible to measure the quality of many videos in parallel.

V-Factor by Symmetricom uses the transport stream and the bitstream as input. It's a real-time end-to-end system that works on the packet-level and doesn't require a reference [70]. We will take a further look into the system in 3.2.1. Another method is the *Media Delivery Index (MDI)*, which is a cost effective method for indicating the video quality, although not very accurate [31].

Human visual system based picture metrics

HVS-based systems uses quality parameters modeled from human vision, such as color perception, contrast sensitivity and pattern masking. Such systems can be used on a wide variety of video distortions, however they are very complex and difficult to implement real-time [66].

Some notable implementations of HVS systems are: the *Visual Differences Predictor (VDP)* by Daly, the *Sarnoff Just Noticeable Differences (JND)* metric by Lubin, the *Moving picture quality metric (MPQM)* by van den Branden Lambrecht, the *Perceptual distortion metric (PDM)* by Winkler, the *Digital Video Quality (DVQ)* by Watson and the *Perceptual evaluation of video quality (PEVQ)* by Opticom [31, 70].

Engineering approach picture metrics

The engineering approach picture metrics are primarily based on the extraction and analysis of certain artifacts in the video, such as structural image elements (contours) or image distortions (e.g. block artifacts).

The *Structural Similarity index (SSIM)* by Wang computes the mean, variance and covariance of small patches inside a frame and combines them into a distortion map. It measures luminance, contrast and structural similarity. The *Video Quality Metric (VQM)* by Pinson and Wolf measures both spatial and temporal features of small sequences. It takes into account such aspects as blurring, global noise, block distortion and color distortion [31, 70].

Hybrid metrics

Hybrid metrics may include data metrics and transport stream metrics, as well as HVS and engineering approach picture metrics. *Telechemy Video Quality Metrics (TVQM)* is an example of hybrid metrics. TVQM has 3 layers: *Perceptual Quality Metrics*, *Video Stream Metrics* and *Transport Metrics* [5].

The *Perceptual Quality Metrics* includes MOS for video and audio based on codec, frame/bit rates, packet loss and Group of Picture (GoP) structure. The MOS shows how the perceptual quality of video and audio is affected by network impairments. The *Video Stream Metrics* provides an insight to how the frame structure is impacted by data loss, by counting received, lost and impaired I/P/B frames and measuring the bandwidth. The *Transport Metrics* is focused on packet loss and jitter as well as gaps and bursts in the packet stream.

2.3.3 Reference information

Some of the methods in 2.3.1 and 2.3.2 require reference data and others do not. *Full Reference (FR)* methods assumes full access to the original sequence, while *Reduced Reference (RR)* methods only require a reference sequence from an earlier stage in the service delivery, as illustrated by figure 2.9. *No Reference methods (NR)* can evaluate the quality at any stage of the service delivery only based on the test-sequence.

Table 2.4 gives an overview of advantages and disadvantages of FR methods compared to RR- and NR methods [70]. FR methods are best used in lab testing and give a superior quality assessment performance [31], while RR and

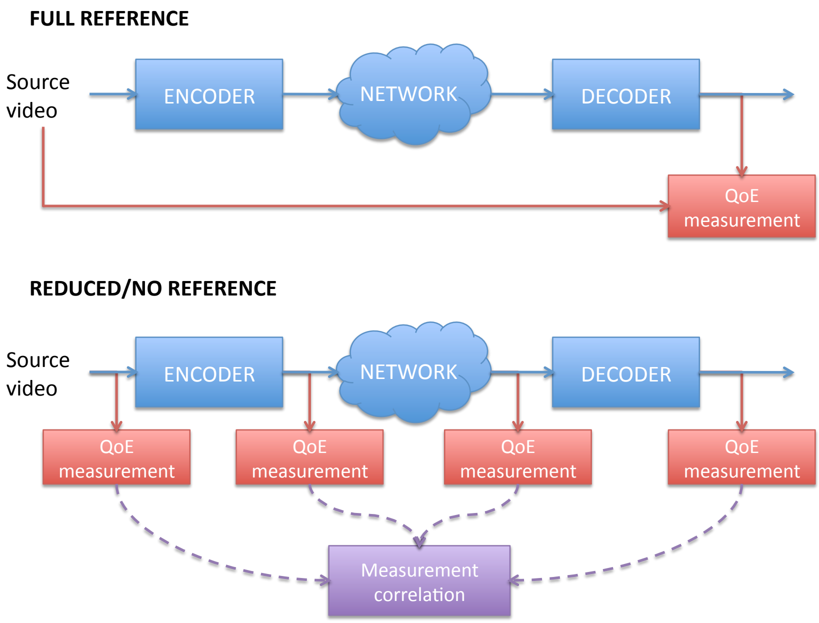


Figure 2.9: *Comparison of FR, RR and NR methods [68].*

NR methods are preferred for use in real-time systems (in-service video). FR methods are usually based on a frame-by-frame comparison, which requires precise spatial and temporal alignment. HVS metrics, data metrics and double-stimulus subjective quality assessment methods are typically FR. NR and RR methods are more flexible than FR, as NR need no reference at all and RR usually only need to extract reference parameters to compare between test-sequences. VQM is an example of an RR method and V-Factor is an example of a NR method.

	FULL REFERENCE	REDUCED/NO REFERENCE
+	+ Accurate + Very detailed analysis + Suitable for lab testing	+ Can use data from many stages of the delivery + Flexible (No Reference) + Suitable for in-service video
÷	- Requires full quality reference material - Difficult to implement real-time - Requires precise spatial and temporal alignment - Limited to short clips to minimize misalignment - Require brightness, color and contrast calibration	- Difficult to distinguish between distortions and content

Table 2.4: *Advantages (+) and disadvantages (÷) of FR, RR and NR [31, 70].*

Chapter 3

Case Studies

3.1 Absolutt Fotball

Absolutt Fotball is a football-streaming project conducted by Schibsted ASA, the largest media conglomerate in Norway. Absolutt Fotball consists of the streaming services *VG LIVE*, by Norway's largest newspaper Verdens Gang, and *100% Fotball*, which is owned by some of the largest regional newspapers as well as Aftenposten, the second largest newspaper in Norway. The structure of ownership in Absolutt Fotball is shown in figure 3.1.

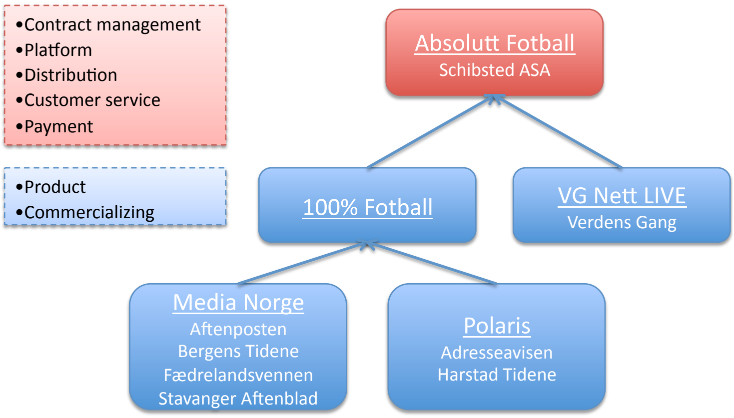


Figure 3.1: *Structure of ownership in Absolutt Fotball.*

VG LIVE and 100% Fotball is based on the same architecture and business model, which will be covered in section 3.1.2 and 3.1.3 respectively, however their products are slightly different. The streaming technology is provided by Move Networks, which is reviewed in 3.1.1. The most notable differences between VG LIVE and 100% Fotball are in the user interfaces, surrounding services and that 100% Fotball gives an option of local commentators at the football matches. In section 3.1.4 we will take a closer look at the VG LIVE product.

In previous years all the matches in the Norwegian top league *Tippeligaen* have been available on pay-TV. Before the start of the 2009 season the Norwegian electrical power company Lyse bought exclusive rights to all but one of the

Sunday matches, which is the primary day for Tippeligaen. Lyse sold the web-rights to Schibsted and TV2 (Norwegian), as well as broadcasting the matches on their own IPTV service *Altibox* [7]. This meant that 62,5% of the matches would be exclusively shown on the web or IPTV. There have been controversy regarding the choice of the Internet as the main channel for Tippeligaen, some of which will be covered in section 3.1.5. However, a week before the start of the football season the TV rights to the Sunday matches were sold to TV2 [4], taking away the exclusivity of the web-services.

3.1.1 Move Networks

Move Networks is a Utah-based company specializing in high quality Internet television services. They have some of the largest media companies in the world as customers, such as ABC, Fox, ESPN, the CW, Televisa, and others, and streams online 11 of the top 20 primetime shows in the US [9].

In August 2008 Move partnered with Level 3 Communications to provide live online HD coverage of the 2008 Democratic National Convention. Level 3 provided the backbone network solution, which had peak traffics of 60 Gbps during Barack Obama's acceptance speech. Users in 116 countries could experience HD live video stream and they had average bitrates of 1.2 Mbps [45].

The quality difference - QoE

Move's selling point compared to competing technologies such as Flash and Silverlight is what they call "The Quality Difference" [9]. To them the quality difference that will keep the audiences longer and more frequent to their services has three parts:

Fast start. They believe viewers want to watch a video straight away when they push "play", instead of having to wait for a buffer to fill up or a file to download.

TV-like quality. The viewer should experience the video quality as "smooth, crystal-clear video images. Like watching HDTV, but with more flexibility".

No buffering. The viewer shouldn't have to experience interruptions during his/her show because of buffering.

Adaptive streaming

To achieve their goals of a high level of QoE, the core of their system is based on adaptive streaming. The streaming quality adapts to the current network throughput as well as user processing power [46].

It starts with *Move Simulcode*, which after receiving the original content divides it into small segments called "streamlets". Then each streamlet is encoded using On2's VP7 codec into several bit rates, according to the content publisher's needs, and stored as a small file on a standard web-server [9]. The streamlet-files are downloaded to the client one by one over HTTP/TCP. The process is illustrated in figure 3.2.

When a user requests a video stream using the *Move Media Player* a low quality streamlet is sent to ensure the fastest possible startup time. Then *Move Adaptive Stream* will seamlessly adjust the image quality to the user's bandwidth and processing power, assuring the highest possible video quality. The

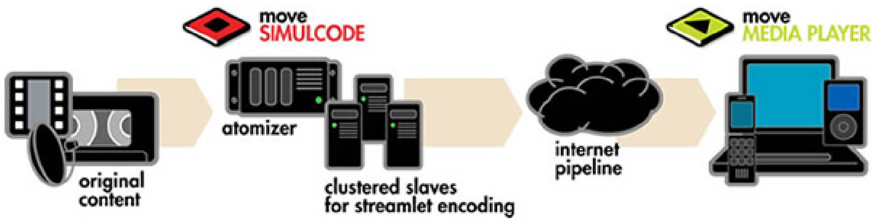


Figure 3.2: *An overview of the Move Simulcode process [9].*

intelligence in Move Adaptive Stream lies in the client, which is a small browser plug-in, and not in the server. If the bandwidth is compromised, the client automatically requests a lower bit rate from the server, avoiding buffering during a video stream, as illustrated in figure 3.3 [46].

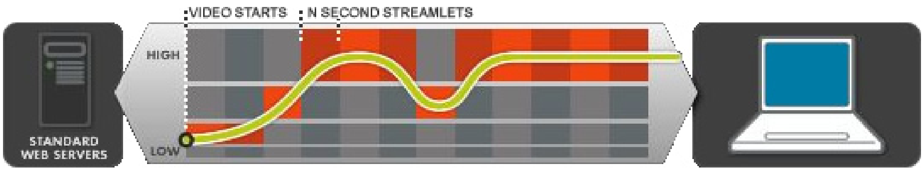


Figure 3.3: *An overview of the Move Adaptive Stream [46].*

Reporting

To analyze the viewer statistics Move has developed *Move Monetize*. The tool provides content publishers with reports on the number of unique viewers by time-periods and program, bit rates and GBs served, viewers by location, "Top" lists and custom reports [9].

The tool is a way of monitoring the QoS and to some extent the QoE and user satisfaction.

3.1.2 Service delivery

An overview of the service delivery from capture to user is shown in figure 3.4.

Usually on a Sunday there are several matches playing at the same time in different places around Norway. Each match is filmed using a multi-camera production. The match is produced in a production truck outside the arena and sent to Bergen through a dedicated fiber channel using lossless JPEG2000 encoding.

In Bergen the video stream is broken down into 2-second streamlets, and each streamlet is encoded into 8-10 different bit rates using Move Simulcode (On2 VP7). The bit rates vary from 32 kbps to 1360 kbps. The time it takes from capture at the stadium to encoded streamlets is approximately 4-6 seconds.

The encoded streamlets are then sent to Move Networks in USA where the stream is published to origin servers and later pushed to various cache servers in Europe. From Bergen to USA the files are sent over an open network using a proprietary transfer protocol that guarantees reliable high speed point-to-point transfer, and from USA back to Europe Level 3 is used as CDN with guaranteed QoS.

Users can access the stream through a web page using Move Media Player. They connect to the web-servers over the open Internet, and thus have no guarantee of QoS. Still, Move Adaptive Stream and a cache/buffer in both the server and client should preserve the QoE for the user. In some cases Level 3 is used all the way to the ISP's network, which gives a better QoS for the end user.

From the video is captured at the stadium until shown on the user's screen, there is a controlled delay of 60 seconds. Most of this delay is due to the cache in the web servers as well as in the client buffer. [26]

3.1.3 Business model

The Absolutt Fotball business model is primarily based on paid subscriptions (see figure 3.5), shared with the Norwegian football association (NFF/NTF). The customer generates a user account at VG LIVE or 100% Fotball where he/she can subscribe to the football matches. The customer may choose an annual, monthly or weekly (one match) subscription and the prices are respectively 999 NOK, 129 NOK and 99 NOK (introduction offer 499/89/59 NOK). The customer has access to all matches in Tippeligaen throughout the duration of his/her subscription. It's also possible to subscribe to all highlights for 39 NOK a month. Subscriptions are paid with credit card. After each match all the video highlights are available for free, as well as textual updates/highlights during the matches.

Schibsted or any of the part owners (VG LIVE, 100% Fotball) are not allowed to use ads in the video stream, but they are free to use ads in the user interface (UI) surrounding the video stream. Because the customer base is centered on Norwegian football and football interests in general, ad buyers are able to target their audience with high precision, allowing Schibsted to sell the ads in the streaming UI and related sites at a premium price.

On a much broader sense the streaming service is a calculated move by Schibsted and the Norwegian newspapers to gain access to moving pictures, which earlier were available only through TV and the TV stations' web sites. Herb Scannell, CEO and co-founder of Next New Networks, said in [16] about video on the Internet: "Video is a game changer. In the history of media, moving pictures trump everything. Moving pictures trump print, moving pictures trump audio, and so we are in the process of changing the game." The New York Times have also recognized this in the article "European Newspapers Find Creative Ways to Thrive in the Internet Age", where VG Nett is used as an example of how the newspapers can generate revenue through users in services such as VG LIVE [53].

Football is one of the leading entertainment industries in Norway and moving pictures will complete the football services for the newspapers, which earlier only had textual live updates as well as pictures (and some times radio) from the matches. Schibsted believes that this is an important part of their branding strategy and to increase customer loyalty and hits to their web sites. Absolutt Fotball is also an important part of Schibsted's further commitment to streaming services and competence in the field [25].

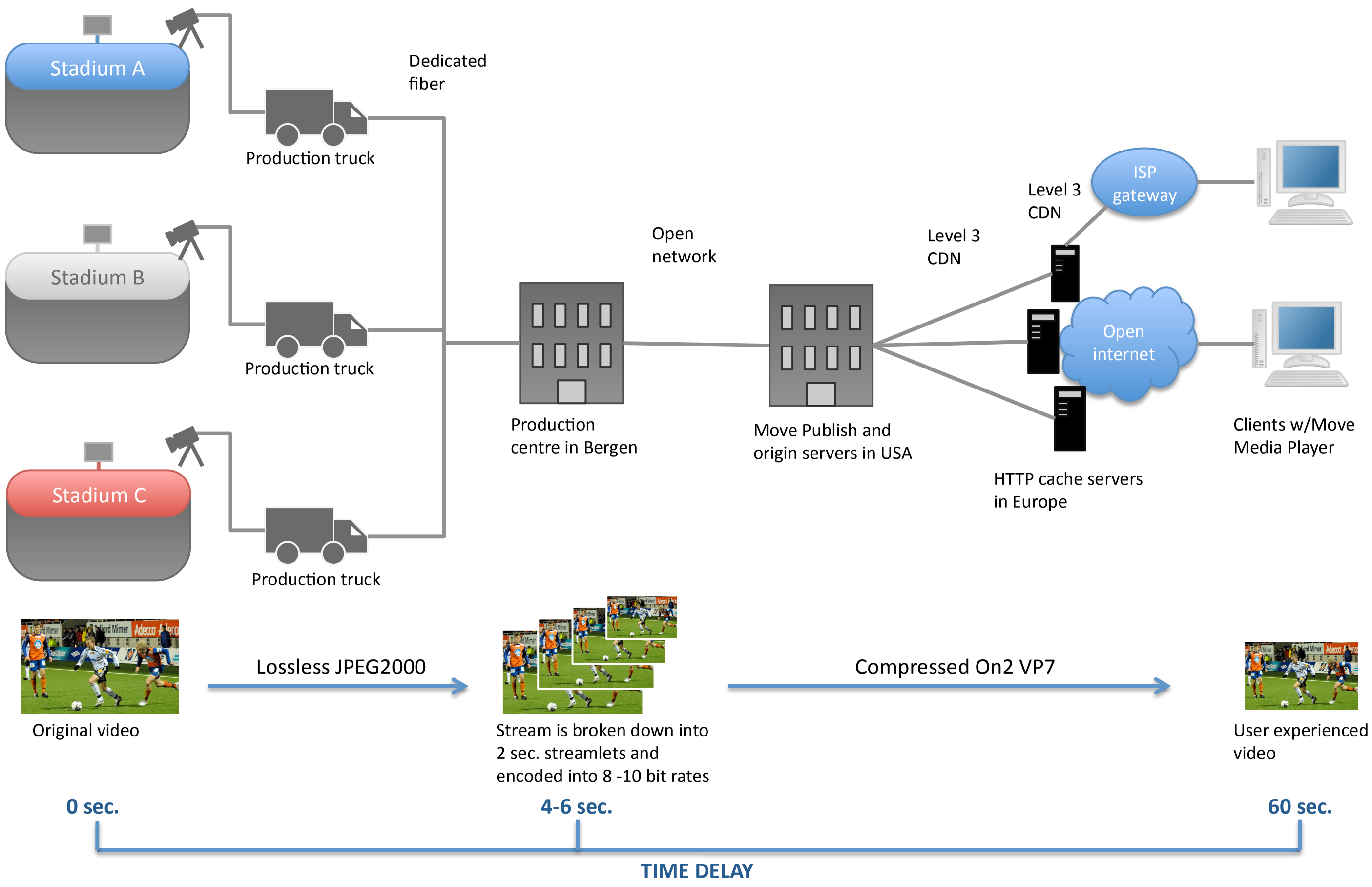


Figure 3.4: *An overview of the Absolutt Fotball service delivery.*

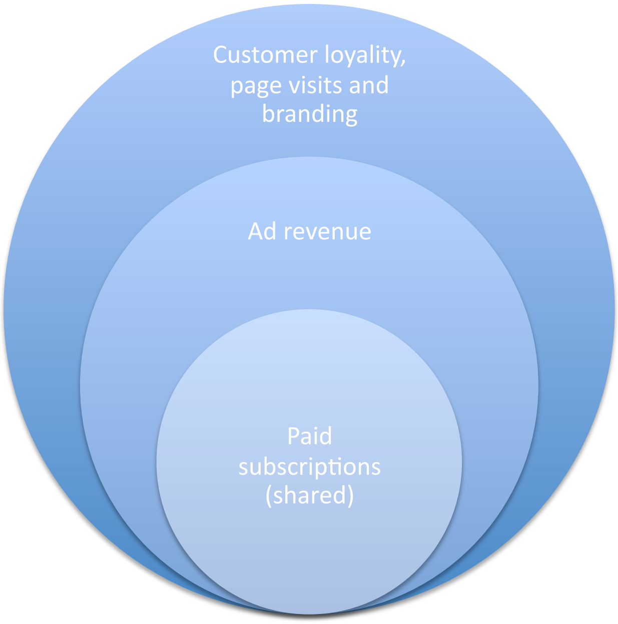


Figure 3.5: *An overview of the business model with paid subscriptions as the primary revenue source, and ads in UI and related sites as well as general branding, which will create revenue in a wider perspective.*

The centre of all this is the customer base, and as briefly discussed in 2.2.4 QoE is very important to prevent customer churn. As far as QoE in the streaming service VG LIVE promises the customer reduced probability of skipping (jerkiness and choppy audio) during playback thanks to Move's media player, as well as "the best picture quality your Internet connection allows". The ”pause”, ”forward/rewind” and "live" functions are also mentioned [20].

The video quality is at best SD (1360 kbps), but Schibsted is investigating the possibility of upgrading to HD in 2010. This will require a bandwidth per user of about 2,7 Mbps as well as increased costs in pre/post-production, distribution (to Bergen) and Move HD encoding [26].

3.1.4 Product and user interface

The VG LIVE user interface is shown in figure 3.6. The main aspects of the UI are: the *main video* (1), the *playback control panel* (2), the *secondary video* (3), *important events* (4) in the matches that are played at the same time, the *overview of other matches* (5), and the *match overview and highlights* (6) for the current match.

The *main video* (1) usually plays a live stream of the current match at 690 x 388 pixels. Using the *playback control panel* (2) it is possible to pause the video stream, skip forward or backward in time using the slider at the top, return to live mode, as well as choose to watch the movie full screen. The *secondary video* (3) usually plays highlights as selected by the user from either panel (4) or (5). The secondary video window is 388 x 162 pixels. The same functions from the playback control panel are available for the secondary video as for the main video. If a user chooses to replay a highlight in the secondary video at the same time as a live stream in the main video, he/she has to manually lower/mute the volume for the stream he/she doesn't want to hear.

The *important events* panel (4) gives a chronological overview of all highlights from the matches played at the same time. The user can access and view everything from goals to shots on/off target or fouls. The *overview of other matches* panel (5) gives the user an overview of all matches played that day, whereby the user may access other matches' highlights as well as live stream (red "play" icon). The *match overview and highlights* panel (6) shows all the important events during the current match. The player may access and view certain highlights by clicking the link labeled "SE KLIPP" with a red "play" button.

In selected matches the user may choose between two alternative match commentators in separate audio streams; one neutral and one locally biased.

Both streams in main and secondary video can be viewed in SD quality, when the user has good enough equipment and enough bandwidth. Audio is a stereo signal.

Customer service is offered through a FAQ page. If the user can't resolve his/her problem using one of the already answered questions, the problem may be submitted to a customer consultant [20].

3.1.5 Football on the Internet vs. pay-TV

The football media agreement

After the football media agreement was published in the summer of 2008, the choice of Internet TV and IPTV as the main channels for Tippeligaen rose big controversy in Norwegian media. In the previous three years TV2 had exclusive rights together with Telenor for Tippeligaen, and had gained a customer base of 85.000 to the pay-TV channel *TV2 Sport* [17]. The Norwegian football association (NFF/NTF) based their decision on the aim of reaching as many people as possible, however the money-side of the agreement of about 1,2-1,4 billion NOK also probably played an important role [7].

Niels Røine, CEO of Norsk Toppfotball (NTF), defends their choice of the web and IPTV with: "After analysis of the two scenarios we have concluded that it will be economically strategic and better for the football using the web and IPTV. We have seen Internet TV broadcasting from the US election and the Beijing Olympics that it's attractive for the users, which has increased our faith in Internet TV." (paraphrased from Norwegian [3])

The power of Internet TV is that it's available for almost anyone nowadays. You are no longer limited to location or cable service provider, and you can even watch it on the go using a 3G connection to the Internet. However until recently the viewing quality of Internet TV has been way below TV-quality, and for less tech-savvy customers who would like to watch football on their TV, connecting their PC to the TV can be quite a hassle compared to the traditional "all-in-one product" of cable-TV. The question remains if the change from pay-TV to Internet TV is appealing for the average Norwegian football viewer.

Figure 3.6: *The VG LIVE user interface.*

Last minute change

On Monday before the start of Tippeligaen (the following Saturday) Lyse and the Norwegian football association (NFF/NTF) re-evaluated the agreement and sold the rights to pay-TV coverage to TV2. This led to the holders of the web-rights losing their exclusivity and would have to compete with pay-TV along with each other on the web/IPTV. According to Jo Christian Oterhals of Verdens Gang (VG LIVE) this change in the agreement would give them less restrictions in their product pricing. Further on he says this will let them use the price of the subscription as a selling point against the competitors in the web arena, which in turn would make their product more appealing towards their targeted customer base [12].

First verdict of the Internet TV services

After the first weekend of Tippeligaen NA24.no presented their verdict of the various Internet TV services [13]. Their conclusion was that VG LIVE was the only service to get the stamp of approval after a technically devastating Sunday for both the 100% Fotball service and TV2 Sumo.

VG LIVE were said to have "Very good picture quality with the new technology" (Move Networks) and "Surprisingly good commentators". The interactivity and second video screen for highlights were also mentioned as pros, while the delay of 1 min from "Live" was mentioned as a con.

TV2 Sumo also has an extra video screen for highlights and the best textual updates, however their video quality was said to be way below the Move technology. Their system crashed on Sunday leaving their users with a blank screen.

100% Fotball also had technical problems on Sunday and were sometimes unavailable for the users. However, their video quality and commentators were very good, since they use the same stream as VG LIVE.

3.2 IPTV QoE assurance mini case studies

Currently there aren't any commercial systems for QoE assessment of Internet television, but there are several for IPTV. Quality assessment for IPTV and Internet television share many common denominators and in this section we will take a closer look at some of the QoE assurance products for IPTV, with focus on the underlying quality assessment principles that may be used for Internet television and are relevant for this thesis.

3.2.1 Symmetricom QoE assurance

Symmetricom, a worldwide leader in precise time and frequency products and services, delivers an end-to-end QoE Assurance product they call *V-Factor*. V-Factor measures the video and audio quality over time for IPTV and rates the quality on a scale from 1: "Very annoying" to 5: "Imperceptible". It uses a set of hardware probes and analyzers, as well as software agents, which are deployed at various stages of the IPTV service delivery. It is all tied together with a management software, *Q-Advisor*, to give real-time reports back to the service provider about the user QoE. Figure 3.7 shows how V-Factor may be deployed in an IPTV service [62].

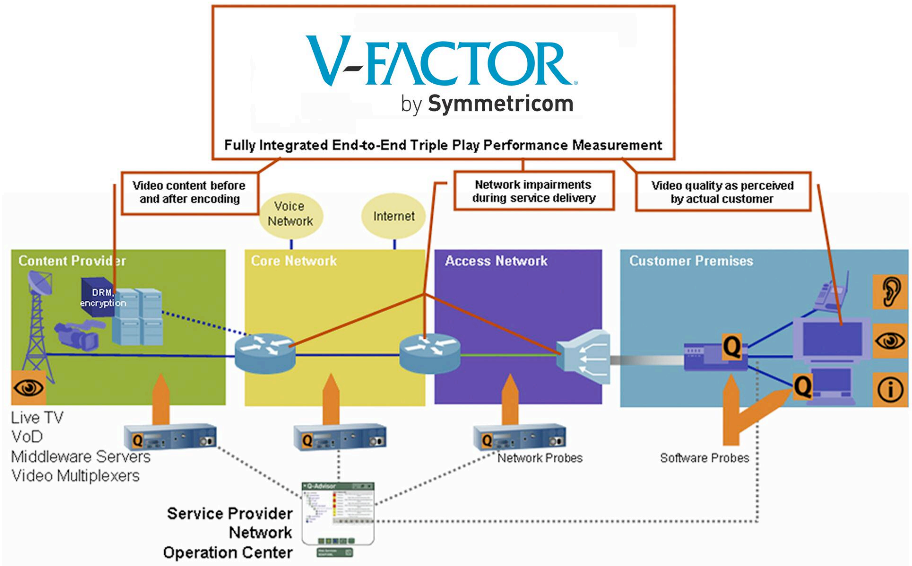


Figure 3.7: *An example of how V-Factor is deployed in an IPTV service [62].*

The V-Factor platform is based on a set of network-,transport stream- and content parameters, as well as A/V content quality metrics. Network parameters include latency, loss episode, loss episode length, jitter and packet counts. The transport stream parameters are based on the ETSI TR101290 standard. Content parameters include horizontal and vertical resolution, aspect ratio, codec profile and level, compression ratio, quantizer scale, and I/B/P frame and slice counts. A/V quality is rated on a MOS scale and takes into account video freeze, blockiness, blackout and motion activity [18].

Figure 3.8 is a block diagram of the various components in V-Factor. There we can see how the input stream is analyzed both in network related parameters and content related parameters, and how the quality is assessed using packet loss, Markov models, entropy and finally a HVS based MPQM model. The output also includes audio MOS as well as video specific measurements of the transport stream (ETSI TR 101290) [61].

Q-1000 is a headend in depth video analyzer that rates the source video based on NR-metrics in real-time and FR metrics offline. The NR metrics includes video MOS, blockiness, blurring, jerkiness, as well as scene temporal and spatial complexity. The FR metrics also includes analyses on noise, colorfulness, as well as PSNR and MSE. Most of the metrics (except for PSNR and MSE) are based on HVS [18]. Q-1000 together with the network probes and software agents are capable of giving the service provider a more or less complete picture of the perceived QoE at the user's end.

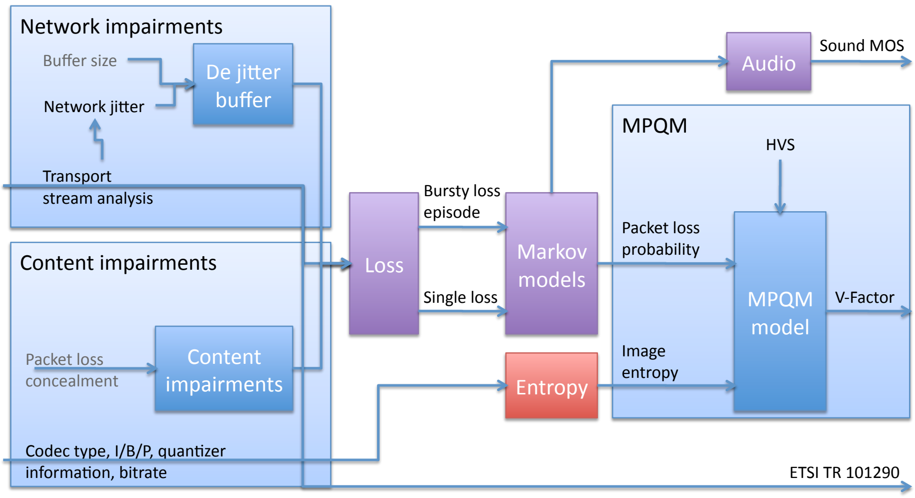


Figure 3.8: *V-Factor block diagram [61].*

3.2.2 Agilent N2X

Big scale lab testing of IPTV systems is difficult to achieve with traditional methods. It requires an extensive test bed of both video sources and receivers (set-top boxes), as well as configuration time and maintenance. It would require tens of thousands of receivers to represent a real world situation and test the outer limits of the system.

Agilent Technologies delivers a solution to this problem, that is the *Agilent N2X*. Agilent N2X is a QoE test solution for IPTV that can emulate hundreds of video sources and thousands of receivers per system to test the limits of its performance. It will give real-time reports on video quality and channel zap time on a per user basis [28, 29].

Agilent N2X is focused on the network performance and uses MDI to assess the video quality, which is a transport stream based metric (see section 2.3.2). The MDI metric reports on two parameters: network delay factor (transport jitter) and media loss rate (data loss). The advantage of MDI contra other video quality metrics is that it doesn't need to decode the video, which makes it cost effective. However, it doesn't always give the best characterization of QoE, thus Agilent N2X has an option of capturing video sequences for subjective visual examination.

3.2.3 Ixia IxLoad

Ixia IxLoad is a similar product to Agilent N2X that can emulate a highly scalable IPTV service for lab testing. IxLoad can be used to emulate clients, subscribers and servers to test devices or networks, or it can emulate clients or subscribers to test the performance of servers [39]. The system returns real-time statistics and post-analysis on application delivery, video delivery, voice delivery and multiplay delivery.

The video quality assessment in IxLoad is based on MDI and MOS-V from TVQM metrics (see section 2.3.2). MDI includes network delay factor (transport jitter) and media loss rate (data loss), while MOS-V rate the perceptual video quality on a scale from 1 to 5 where 1 is ”Useless” and 5 is ”Excellent”. MOS-V considers the effect of the video codec, frame rate, packet loss distribution and GoP structure on viewing quality [5].

3.2.4 Mariner xVu

The last mile in the IPTV service delivery has been difficult to control for service providers, and that’s why Mariner has developed the ”customer centric” approach to IPTV QoE assurance they call *Mariner xVu*. xVu is an end-to-end in-service proactive QoE assurance platform with four main features: *NetworkVu*, *SupportVu*, *ExecVu* and *HomeVu*.

NetworkVu monitors the service delivery from the headend all the way to the customer’s end and often times QoE problems are detected and resolved proactively using SupportVu before the customer even notices there is a problem. ExecVu tracks the customer activity on a macro level overlooking subscriber service uptime, critical customer satisfaction issues and customer churn.

The advantage Mariner xVu has over competing QoE assurance platforms is its presence in the home through HomeVu. HomeVu is an application that enables the customer to self-diagnose service performance issues, as well as report problems back to the service provider. The communication happens through on-screen messaging and in many cases the customers may resolve the problems themselves [8].

Chapter 4

Methodology and Results

In this chapter we will present the methodology and the results of the experiments that were conducted in order to test the hypothesis presented in chapter 1. The first phase was to identify the critical factors in the video QoE, which will be covered in 4.1. The next phase was to resolve which type of feedback that is useful for the service provider in order to better the QoE, which will be covered in 4.2. The last phase was to find out how the service provider of Internet TV can receive real-time feedback from the users, which will be covered in 4.3.

In order to perform the user experiments in 4.1 we have cooperated with Schibsted ASA as the service provider of Absolutt Football, focusing on the VG LIVE service, which were presented in 3.1.

4.1 Phase 1: Identifying the critical factors of the QoE

In the first phase we chose to conduct a user study to identify the critical factors in the VG LIVE video QoE. We chose to experiment through a user study because lab testing of the service simply wouldn't give a representative answer for the purpose of this study. Surveys are an effective and inexpensive way of gathering and processing feedback from a larger sample of the geographically spread population, and in our case this was important to get a good representation of the customer base to VG LIVE. A disadvantage with surveys is the response rate; since there are no-one interviewing the subjects or have any other form of control over their answers, a survey may get a variable response rate depending on the population.

A statistical survey is a type of quantitative research where the aim is to collect information about items in a population. Quantitative research gives us systematic data that can be analyzed or employed into mathematical or statistical models and theories. Surveys are an example of differential research methodology, where the individual differences are as important as the population mean. We chose to conduct a longitudinal survey, whereby the respondents respond to the survey several times over a longer period of time. Longitudinal research gives us information of the development over time for each user as well as a status quo.

The main purpose of the survey was to elicit the user's overall satisfaction with the video QoE, as well as find out how noticeable the various video degradations (see sec. 2.2) were to the end user. In order to address the correct aspects of the QoE we conducted a short pre-analysis of the video streaming service narrowing the scope of the survey, which can be found in appendix A.

Another concern was how to address the questions about the various aspects of the QoE to non-technical users. Terms like "jerkiness", "blockiness" or "ringing" don't say much to the average football supporter. To solve this, a number of pre-tests were conducted on people comparable to the users in the actual survey. After each test we received feedback on the questions in the questionnaire about the phrasing and how the user understood the questions.

4.1.1 Survey

The survey was composed of three parts. First a gathering of user data before the test period, secondly the quality analysis during the test period, and finally a post survey after the test period.

User data

The first part of the survey was to examine demographical and psychographic information about the users, as well as user equipment and network connection. The significance of the demographic and psychographic data is to describe the sample of the study resulting to external validity of the study, and to see if there is any relation between different demographical/psychographic parts of the population and their perception of QoE. User equipment and network connection is important as it gives a good picture of the "last mile" of the service delivery for each individual user and that certain problems with the QoE can be directly linked to this.

The demographical and psychographic questions were loosely based on [40] and included age, gender, occupation, technology attitude, as well as the user's football- and Internet video viewing habits. User equipment included screen size, resolution and type (CRT/LCD, computer/TV), computer age, operating system, Internet browser and Internet connection.

Quality analysis questionnaire

The quality analysis questionnaire was based on 2.2, the pre-analysis (appendix A) and the pre-tests. The questionnaire was to be answered after each watched match during the test-period from March 21st 2009 to May 4th 2009. Questions included overall quality of VG LIVE compared to other Internet video services (Youtube, NRK Nett-TV, TV2 Sumo etc.) and football on TV.

In order to address the various coding artifacts we asked about the user's satisfaction with details and general perception of players/referee/ball, textual information/messages (the score, substitutions etc.) and the field/crowd/surroundings. Blurring "removes" the details in the field/surroundings as well as make the players harder to recognize, ringing and color bleeding can make text elements harder to read and make it difficult to see the players at far range, and motion blurring affects especially fast moving objects (players/ball).

Questions also included satisfaction with the adaptation period for short and long sequences, satisfaction with the ”instant start”, and if the user noticed any problems with the video/audio flow. Finally, the user could describe any problems with the QoE that hadn't already been addressed.

Post survey

The purpose of the post survey was to clarify questions we might have had after reviewing the quality analysis data, as well as a summary of the user's impression throughout the test period. We were able to confirm or reject trends we saw in the quality analysis' results by simple questions in the post survey. It was also important to find out if the users after testing the service would have paid for it in the future, regarding Schibsted's pricing models, and finally what could have made the service better and more valuable for the users in the future.

4.1.2 Data collection

All data from the surveys were collected using an online questionnaire service, SurveyMonkey.com. The questions were in Norwegian and the questionnaires can be found in appendix B. The test group of 18 persons were recruited amongst friends and friends of friends while trying to get a representative demographic selection for VG LIVE. Both genders were represented and there were students from 18 and up to 55-year-old employees. They were identified by their e-mail addresses, which made it possible for us to relate test scores to demographical, psychographic and user equipment information in retrospect. The user data was collected from March 16th to March 19th 2009, the quality analysis data was collected from March 21st to May 6th 2009, and the post study data was collected between May 3rd and May 8th.

Several factors in the selection of the respondents may have caused biased results. First of all, we can't be sure that our selection is representative of the VG LIVE customer base demographically, in fact it probably isn't. However, based on feedback from Schibsted we were not far off, and their estimation that the majority of their customers would be male aged twenty to mid-thirties, fits well with our selection. A good portion of the respondents were students of technological studies at NTNU or had IT-related vocations, something that may have raised the general level of technological competence in the selected group compared to the VG LIVE customers. Also, since many of them already were well known friends, their answers may have been biased. Another selection might have given different results.

The data collection in itself was not time limited, meaning we weren't guaranteed that the respondents would respond to the questionnaire immediately after watching a football match. If they waited too long after the match to respond, it could have affected their memory as well as their state of mind. Also, there was no time limit of how short or long part of the match they would have to see before responding to the quality analysis, meaning some may have watched only the first or second half, or less, while others watched the entire match of 2x45 minutes. We had in other words no control over the content they were watching, other than it was a football match provided by Absolutt Football. The match they saw was reported in the quality analysis.

In relation to the various subjective quality assessment methods in 2.3.1, our method most closely resembles SSCQE, with long sequences and no reference, only we used a retrospective rating technique instead of continuous rating. This format fit our purpose for the quality analysis.

A concern when watching long sequences is that people tend to remember the beginning and the end the best, while the middle is less important. This is known as the recency effect [30]. Another problem when evaluating longer sequences is that people tend to evaluate the content instead of the video quality, and additionally the viewing context could be important. This will be further discussed in chapter 5.

4.1.3 Method of analysis

The data collected from the quality analysis questionnaire were discrete interval data measured on a 4- or 5-point rating scale. Rating scales are commonly used in surveys when rating an item from "none" or "poor" to "all the time" or "excellent", depending of the nature of the question. Each rating has a numerical score, positive or negative, which makes it easy to calculate the different statistical parameters.

First an analysis of the responses was conducted in order to sort out possible erroneous answers and unserious respondents. The errors may have come from respondents misunderstanding the questions, someone who have a radically different opinion of what "good quality" is and other conditions radically different compared to the general mean. These errors are called random errors and contribute to increasing the variance. Systematic error is called a "bias", and will change the mean. These errors are harder to find. A common source of systematic error when using a rating scale is the *central tendency*, whereby respondents tend to avoid the extreme ratings (both positive/high and negative/low).

We used descriptive analysis for the survey results, including average/mean, standard deviation and confidence intervals. We also analyzed the correlation between certain factors by comparing averages in relation to given prerequisites.

The purpose of our analysis was to see the totality of the survey, to get an overview of the perceived level of QoE amongst the test users. It was also interesting to see if the answers varied much from match to match for each user, and what technical or non-technical reasons might have correlated with their perception of QoE.

4.1.4 Survey results

There were collected 62 responses of the quality analysis from 18 respondents during the testing period. A summary of all survey data can be found in Appendix C. In this section we will present the main findings from the surveys.

Demographical and psychographic data

We had 18 respondents, 14 male and 4 female. They were aged from 18 to 55 years old, with a majority from 18-30, as seen in figure 4.1. An apparent weakness in our selection is the lack of respondents between 31 and 40, which is an important demographic for Absolutt Fotball.

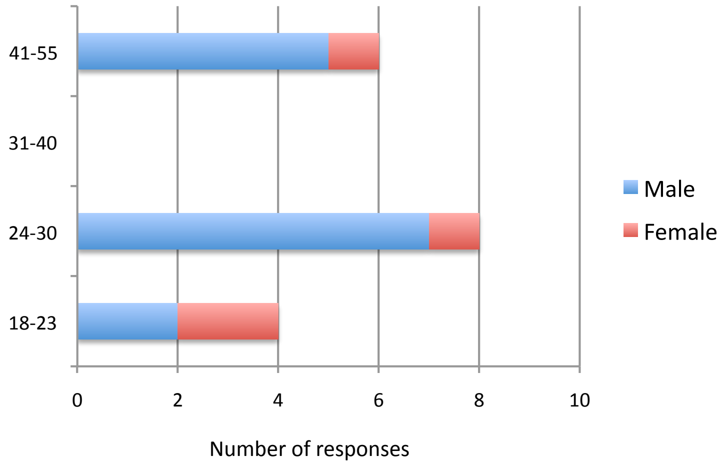


Figure 4.1: *Respondents in our survey categorized by age and gender.*

We had a 40/60 distribution of students and employees, a good portion of whom from technological studies (NTNU) and work, such as IT consultants and a professor at NTNU. This might have contributed to an above average interest in technology in the selection, as seen in figure 4.2.

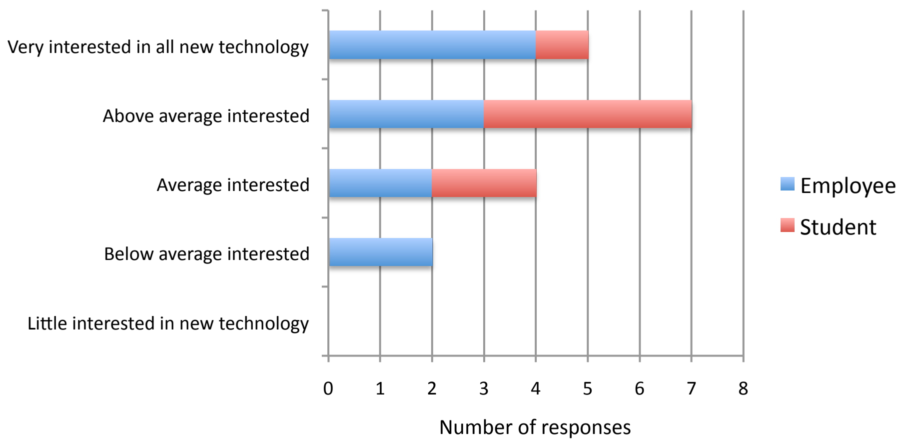


Figure 4.2: *Respondents in our survey categorized by technology interest and occupation.*

More than 80% watched football at least weekly, mainly on TV, and more than 50% watched videos on the Internet at least weekly.

User equipment data

4 of the 18 respondents would use their TV to watch the matches on VG LIVE, all of which were HDTVs above 32". The rest would use their computer screens ranging from small to medium sized (13" - 16"), medium to large sized (17" -

22") and large to extra large sized (23" - 32"), with resolutions from 1024x768 to 1920x1200.

Some of the respondents would use new computers (less than a year old) and others would use older computers (3-5 years old). Both Windows and Mac OSX was represented, as well as the supported browsers; Internet Explorer, Firefox and Safari. All respondents had a 2 Mbit or faster Internet connection, something that theoretically should give them the best quality at VG LIVE.

VG LIVE quality comparison

The overall impression of the video quality of VG LIVE was that it was slightly better than other Internet video services, such as Youtube, NRK Nett-TV and TV2 Sumo, and worse compared to football on the TV. However, as seen in figure 4.3 the answers were variable and almost the entire scale was used for both options. On a scale from -2 for ”VG LIVE was much worse” to 2 for ”VG LIVE was much better”, the average scores were 0,51 compared to Internet video (std.dev. 0,85) and -0,79 compared to TV (std.dev. 0,64).

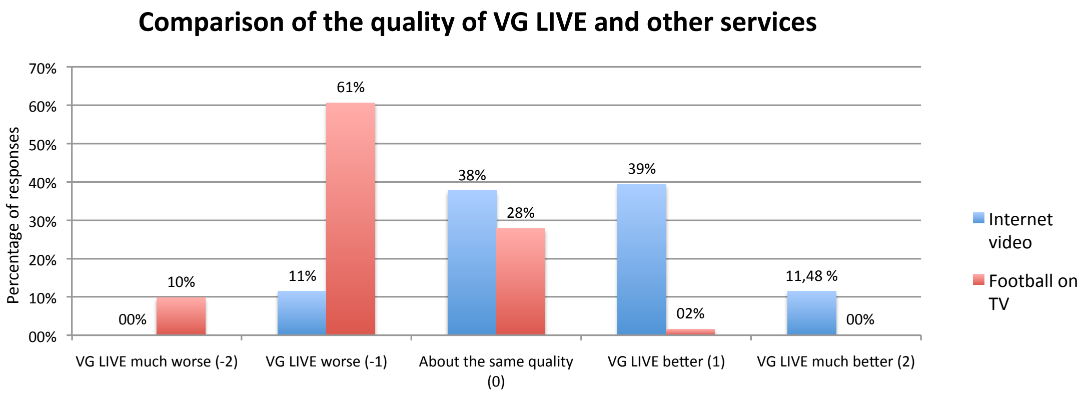


Figure 4.3: *Compared quality of VG LIVE to other services from the quality analysis.*

When we compared the results to the respondents' favorite teams we found a slight correlation whereby the average scores were higher if their favorite teams won than if they lost (see table 4.1). This is not a conclusive result since there are many other contributing factors, however this shows a tendency that the content is important for long-term quality analysis. We also found a correlation between the overall quality comparison and the respondents' Internet video habits. The results show that people who watch a lot of Internet video might have lower expectations to the quality, thus contributing to higher scores. These results require further testing to be conclusive.

Level of detail and sharpness

The respondents were in general satisfied with the details and sharpness of the video, with an average rating between "’OK” and ”Good” for all options as seen in figure 4.4. On a scale from 1 to 5, where 1 is ”Very bad” and 5 is ”Very good”, the respondents rated how well they saw the players/referee/ball at 3,48 (std.dev. 0,67), how well they saw the on-screen textual messages at 3,73

	VG LIVE vs. Internet video	VG LIVE vs. TV	Combined score
Average	0,51	-0,79	-0,14
Favorite team result			
Win	0,52	-0,61	-0,04
Draw	0,55	-0,82	-0,14
Loss	0,40	-1,00	-0,30
Internet video watching frequency			
Rarely	0,36	-0,85	-0,25
Weekly	0,58	-0,84	-0,16
Daily	0,71	-0,60	0,06

Table 4.1: *Compared quality of VG LIVE to other services based on the respondent's attitude towards the content and Internet viewing habits.*

(std.dev. 0,75), and the level of detail on the field/crowd etc. at 3,15 (std.dev. 0,65).

Related to video artifacts this could mean that people notice a good amount of blurring and motion blurring, because of the lower score on level of details on the field/crowd etc., as well as some edge ringing and color bleeding. There were also explicit responses from the comment field of the survey about motion blurring and lack of details (grainy video).

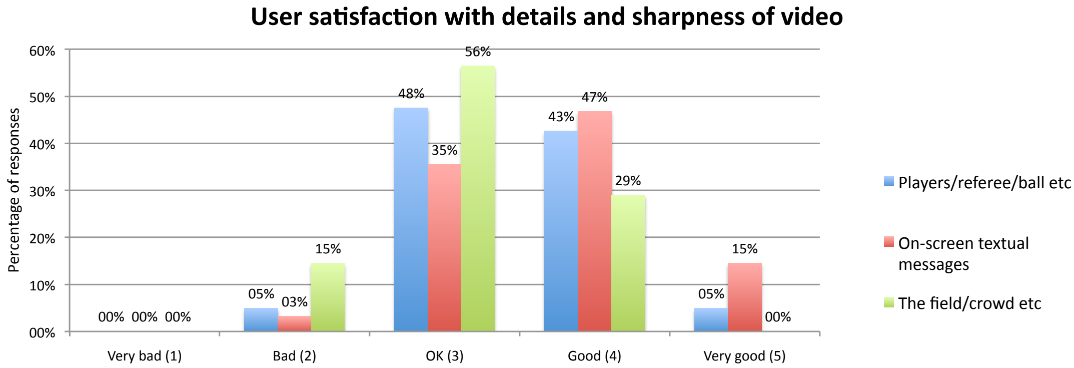


Figure 4.4: *The respondents rated their satisfaction with details and sharpness of video.*

Adaptation period

The results that the respondents didn't feel much inconvenience with the adaptation time for long sequences (the main match), however as expected, the ratings for short sequences (highlights) were lower, meaning the adaptation period was more inconvenient for short sequences (see figure 4.5). On a scale from

1 to 5 where 1 equals ”Unacceptable” and 5 equals ”No inconvenience”, the long sequences received an average score of 4,11 (std.dev. 0,83) and short sequences 3,40 (std.dev. 1,04).

The short sequences were usually between 40 and 60 seconds long, and the time of interest (e.g. when a goal was scored) could be 10 seconds into the clip. If the adaptation period was longer than the time of interest, it would be inconvenient for the user.

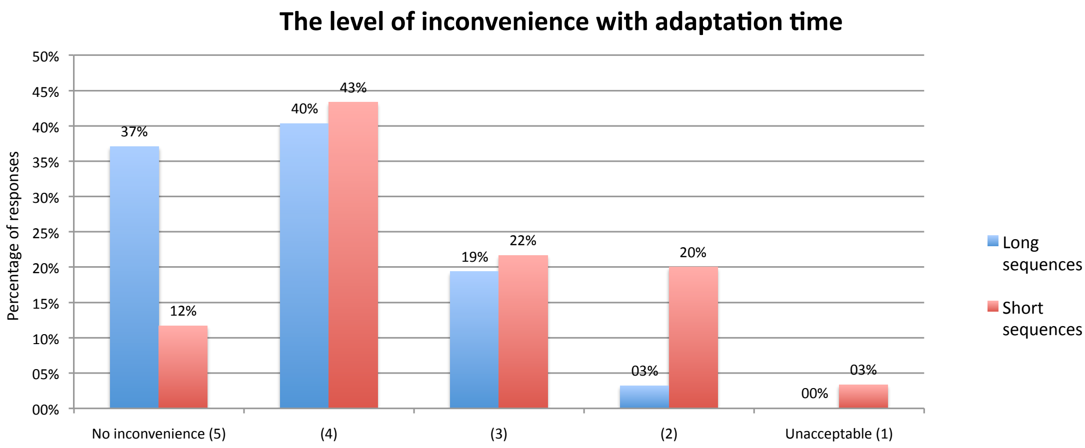


Figure 4.5: *The respondents rated their inconvenience with the adaptation period.*

Response time

The users reported a general satisfying response time between ”OK” and ”Good” on the rating scale. On a scale from 1 to 5 where 1 equals ”Very bad” and 5 equals ”Very good” the average scores were 3,62 for starting a new clip (std.dev. 0,71), 3,69 for the ”Live”-function (std.dev. 0,70) and 3,54 for skipping/forwarding/rewinding (std.dev. 0,75). Higher scores (towards 5) equals better QoE. They were all close in rating, which is reasonable since they are all based on the same mechanism; requesting and receiving the first low quality streamlet.

Video and audio flow

The results showed some problems with the flow of audio and video, however in general it was pretty good. On a rating scale from 1 to 4 where 1 was jerkiness ”all the time” and 4 was ”no jerkiness”, audio rated in average at 3,52 (std.dev. 0,59) and video at 3,29 (std.dev. 0,69), meaning there are some problems with choppy audio and slightly more problems with video jerkiness.

There were reported few problems with audio silence or noise, but there seemed to have been some problems with video freezes. On a scale from 1 to 5 where 1 equals occurrences "often and for longer periods", and 5 equals "no occurrences", video freezes had an average rating of 4,40 (std.dev. 0,84), audio silence 4,77 (std.dev. 0,61), black screen 4,84 (std.dev. 0,52) and noise/repeated sound at 4,79 (std.dev. 0,52). Higher scores (towards 5) equals better QoE. See

Appendix C for further details of the rating scales and results of the quality analysis.

Overall impression

In the post-survey the users were asked to rate their overall impression of the video quality on a scale from 1 to 10 where 1 equals grainy low-quality videos on Youtube and 10 equals TV-quality. This gave an average rating of 6,72. However, as can be seen in figure 4.6 the answers varied from 4 to 8 on the rating scale, meaning there were some variance in the experiences.

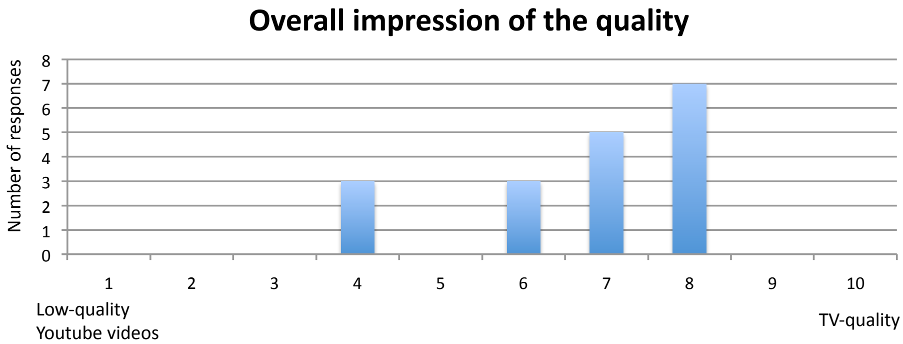


Figure 4.6: *The respondents rated their overall impression of the video quality, with an average score of 6,72.*

VG LIVE functions

All respondents preferred adaptive streaming to traditional streaming with constant quality, some startup delay and occasionally interruptions mid-stream. They were also asked to evaluate the various functions of VG LIVE on a scale of "Very bad"(-2), "Bad"(-1), "OK"(0), "Good"(1) and "Very good"(2), the result of which is shown in table 4.2.

What is interesting to see from the table is that both "quick start" and "no interruptions" were amongst the best rated, which is 2 of 3 criteria for Move Networks' "Quality difference". However the image quality, their third criteria, did not receive as good scores. Worst scores were given to the VG LIVE web page and the delay from real time services. This was commented on in the quality analysis questionnaires as well.

The VG LIVE web page was described as "messy with too many banners and animations" and that it was difficult to find the important information. The delay from other live services, such as SMS updates, radio and TV were reported as annoying. Although these things are outside of the video quality, they seemed to be important factors of the user's overall impression of the QoE.

Some of the respondents may have mistaken the meaning of "Live function (delay)" as the response time when pushing the "Live"-button in the menu, instead of the delay from other live services. This means it could have received lower scores if it had been 100% clear. Also, during the first visits to the VG LIVE web page for the users, we had provided detailed step-by-step descriptions

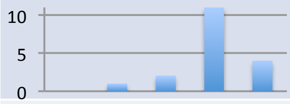
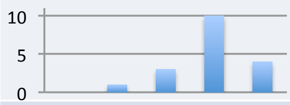
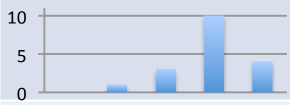
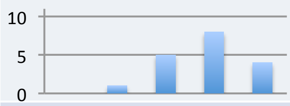
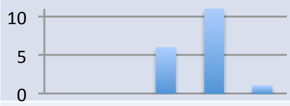
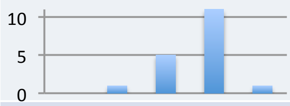
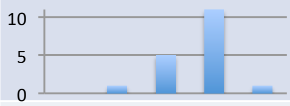
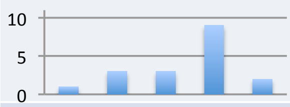
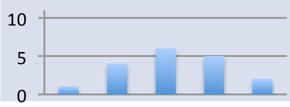
Function	Description	Rating/variability	Average score	Verdict
Quick start	The time from pushing e.g. "play" until the video stream starts		1,00	Good
No interruptions	There should be no interruptions or stops during the video stream		0,94	Good
Forwarding/ rewinding	The possibility of forwarding/rewinding (skipping) in the video stream		0,94	Good
Short highlight sequences	40-60 second highlight sequences from goals, shots on target, etc.		0,83	Good ÷
Adaptation period	The time it takes to reach the best possible quality when starting a video stream		0,72	Good ÷
Image quality in general	General impression of the fidelity of the image quality in the video		0,67	Good/ OK
Player menus	The ease of use and functionality of the menus surrounding the player		0,67	Good/ OK
VG LIVE web page	The user interface surrounding the video stream		0,44	OK/ Good
Live function (delay)	The delay from true real-time streaming.		0,17	OK +

Table 4.2: *The various functions on VG LIVE rated on a scale of "Very bad"(-2), "Bad"(-1), "OK"(0), "Good"(1) and "Very good"(2).*

in how to register profiles and use the service. This may have skewed the ratings of the web page towards more positive.

When we compared the results between people who usually watch football on an HDTV in HD and people who watch football using their PCs, we found that the people who usually watched HDTV rated the overall video quality and picture quality lower, as seen in table 4.3. Also when we compared the average scores of people who used small to medium PC screens (13" - 16"), medium to large PC screens (17" - 22") and TVs (above 32"), we found that the ones who used TV gave the overall video- and picture quality higher scores. The differences are not great and these results are not conclusive and should be further tested, but they show a tendency that the viewing context might be important to the quality evaluation. Sitting in the sofa in front of a big TV screen and enjoying the football match might be a better experience than sitting in front of a smaller PC-screen.

	Overall video quality (1 to 10)	Evaluate picture quality (-2 to 2)
Usually watches football on..		
HDTV (in HD)	5,67	0,17
PC (Internet streaming)	6,50	0,67
Screen used when watching football in the tests		
PC (S-M)	6,67	0,67
PC (M-L)	6,71	0,71
TV	7,50	0,75

Table 4.3: *Overall video- and image quality based on the user's football viewing habits and presentation screen.*

Impressions throughout the test period

11 of the 18 users were positively surprised by the video quality, based on their expectations beforehand, the rest were neutral. 7 of them reported that their impression of the video quality became more positive during the test period, 10 were neutral and only one reported a more negative impression.

8 users said there were ”some variations” in quality from match to match, 3 said there were ”much variations”, while 7 said their video quality was the same in each match.

Price vs. value

The ordinary VG LIVE subscriptions cost 999 NOK annually, 129 NOK monthly or 99 NOK per round (w/loss insurance), however there was an introductory offer the first couple of months at 499/89/59. Based on their experiences with the service 10 out of 18 would pay the ordinary prices for subscriptions (annually or monthly), as shown in figure 4.7. 4 said it was too expensive, 2 of which would consider the introductory prices. 2 said VG LIVE didn't have good enough quality and 2 had other options (stadium, pub, cable TV).

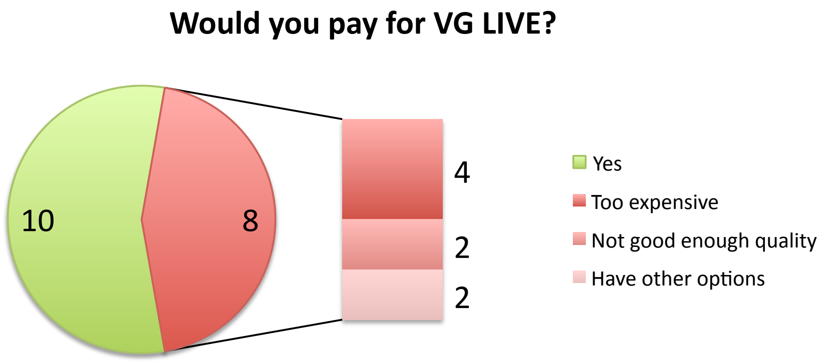


Figure 4.7: *10 of 18 would pay the ordinary price for VG LIVE.*

In order to increase the value of the service, i.e. making it better for the user, all users reported some interest in HD quality video streams. On a scale from 0 to 3 where 0 equals "Not interested" and 3 equals "Very interested", HD quality received an average score of 2,06 ("Interested"), the option to choose quality and price received an average of 1,22, and the option to pay less for low quality and a support-chat in the user interface both received average scores of 0,78, as shown in figure 4.8.

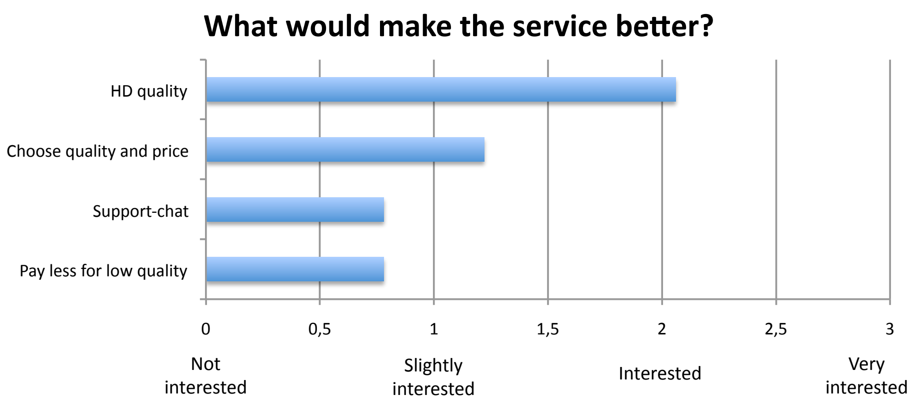


Figure 4.8: *In order to increase the value of the service the respondents in the survey showed interest in HD quality video streams and a slight interest in the option to choose quality and price, paying less for low quality and support-chat in the UI.*

4.1.5 Informal objective analysis

In order to better understand the subjective results we conducted a series of objective tests on the system.

Buffer size and live delay

When the Move Media Player retrieves the streamlets, they are first put in a buffer, and after being played sent to the player cache. These cache files are possible to find on the computer and has the extension .qss.

We were able to measure the buffer size of the implemented Move Media Player by closing down the internet connection and then clock the time it took for the stream to stall. We were also able to count the number of 2-second streamlets that entered the cache after the connection was broken, meaning they were the buffered streamlets at the time of disconnection. We found the buffer size to be 30 seconds. Before the disconnection we made sure we had a stable buffer.

We also tested the total delay from an event happened at the stadium until it was streamed on VG LIVE. This was done using mobile phones where one person reported from the stadium and a second would clock the delay on VG LIVE. It was found to be 60 seconds, as expected.

Adaptation period

We were able to clock the adaptation time by counting the number of streamlets whose quality was lower than the best quality (1360 kbps). We tested both starting a new stream and forwarding/rewinding within the same stream, and found the adaptation time to be between 8-16 seconds, that is 4 to 8 lower quality streamlets.

The tests were conducted on an Apple iMac with 2.4 GHz Intel Core 2 Duo processor, 3 GB DDR2 SDRAM, ATI Radeon HD 2600 Pro graphics card and a 7 Mbit Internet connection, a system well within the maximum requirements of VG LIVE.

4.1.6 Summary

In phase 1 we found that the general perception of the video quality of VG LIVE is that it is better than other Internet video services, but notably worse than TV quality. Coding artifacts such as blurring, motion blurring, edge ringing and color bleeding may be the main reasons for the lower perception compared to TV, especially the blurring, since it "removes" a lot of details in the picture. There were some problems with the audio and video flow, especially lack of smoothness in the video playback and occurrences of frozen frames (likely related issues). The adaptation time was barely inconvenient for long sequences, but for short sequences it was a little more annoying. The general response time was good, as was the video stream without interruptions.

Factors outside the video quality, such as user interface and delay from other live-services, were reported to be important for the total QoE. The surveys showed that there are variable opinions on what good quality is and how it compares to similar services. The analysis also showed that the content of the video stream, as well as the context in which it's being watched will affect the perception of quality. In other words, people's preferences, expectations and earlier experiences as well as viewing environment may play an important role in their overall evaluation of QoE.

4.2 Phase 2: How to improve the QoE

In this section we will analyze the importance and implications of the results from 4.1 related to the service delivery. Based on the QoE wheel (figure 2.7) in 2.2.3 we will extract the critical factors in the service delivery and find solutions to improve the QoE.

4.2.1 Encoding vs. bandwidth

From 4.1.4 we found the biggest problems with the video quality to be artifacts such as (motion) blurring, edge ringing and color bleeding, as well as the smoothness of the stream, or lack thereof caused by video jerkiness and frozen pictures. Referring to figure 2.7 we can assume that most of these problems stem from harsh compression in the encoding stage.

Blurring and motion blurring could also stem from the capturing stage or pre-/post-processing stage, however these problems lie within the production, which is the same for the TV-sent matches, thus we will assume that these problems are minor and out of our control at the moment.

Jerkiness could be caused by a low bandwidth or transmission jitter, however the built-in buffer of the Move Media player should counteract these problems, and if the bandwidth is too low (less than 160 kbps) the video quality would be so bad anyway that the jerkiness wouldn't be the biggest problem. Jerkiness could also be caused by low processing power on the user end, something that we will cover in 4.2.3.

A packet loss might cause a frozen frame, but since the transmission is based on TCP/HTTP, the missing packets would be re-transmitted, and if received before the buffer became empty, the user will not notice this.

The question however is where the bottleneck lies? If the bottleneck lies in the access network, meaning the user experience bad video quality because he/she doesn't receive the highest possible bit rate, there is no point in encoding the video in a higher bit rate. In this case the efficiency of the compression algorithm used in the encoding must be evaluated related to the aforementioned artifacts and video degradations. The codec could be tweaked to be more efficient for its purpose, or new codecs could be implemented and replace the older one, such as On2 VP8 or H.264.

In the other case, if the bottleneck lies in the encoded bit rate, the QoE could be improved by offering a higher bit rate with less compression, and therefore with less artifacts, jerkiness and frozen frames. In this case the added costs caused by higher traffic and higher quality encoding must be evaluated related to the increased user satisfaction.

From our tests it is most likely the latter case, since all our subjects had Internet connections fast enough for the highest video stream bit rate and they reported little variations throughout the period.

4.2.2 Optimizing the client

In our case adaptation time and response time is largely based on the settings in the client. Adaptation time is decided by how fast the client advances its bit rate requests from low to high, based on the current buffer status. Response time is based on how quick the first streamlet is transferred from the server and displayed by the client. A lower bit rate of the first streamlet would reduce the response time, but at the same time could increase the adaptation time.

By tweaking the client's buffer by how it chooses to advance the bit rates in the start, we could reduce the adaptation time. However this could effect the reliability of the player, since it would be less robust against transmission errors. A smaller buffer would also reduce the delay of the service related to other live services.

By our estimations an optimization of the client would be a compromise between various QoE factors, as shown in figure 4.9. The figure shows how there is a balance between the different QoE factors when modifying the starting bit rate and buffer size (or thresholds for increasing the bit rate). Further research would be required in order to find the optimal combination for Absolutt Fotball, however from our results and limited knowledge of the mechanisms it seems the service could increase its overall level of QoE by slightly reducing the buffer size from 30 seconds, and also do some minor tweaks in how the client chooses to request the bit rates (adaptation).

For smaller sequences, such as 40-60 second highlights, adaptive streaming doesn't seem to be as effective, since often times the time of interest (e.g. when a goal is scored) is earlier in the sequence than the time it takes to adapt to the highest bit rate. A standard flash-based video player should be considered for these sequences, since this would eliminate the adaptation problem and such small sequences wouldn't take as much time to download/buffer.

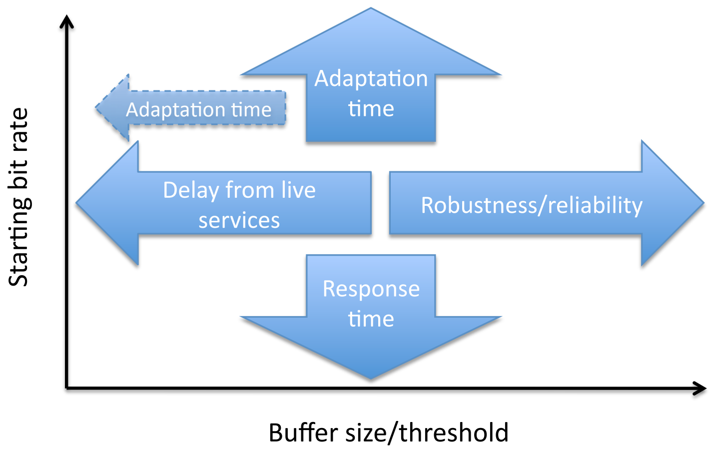


Figure 4.9: *Diagram of how various factors of the QoE can be increased by tweaking the client.*

4.2.3 User equipment recommendations, guides or tests

On the VG LIVE web pages there are information to the user on which type of computer hardware and software that is recommended for the service. The requirements are limited to processor, memory, graphics card, operating system and web browser. These are minimum requirements and a customer using a computer with these specifications is not guaranteed a good video experience.

This could be improved by adding an interactive guide where the user can perform a real time test on his/her system and see where the bottleneck might be. The test could include CPU/GPU processing speeds and Internet bandwidth, since the Move Media Player plug-in already tests this. It could also gather information about the user's operating system and Internet browser in order to ensure compatibility with the service. Additionally the guide could include recommendations on screen resolution and viewing environment in general in order to make the video experience as good as possible.

By real time compatibility tests of a user's equipment we would find out if the jerkiness, choppy audio or other QoE issues were caused by the processing power, and if so how it could be improved for the user.

4.2.4 Maximizing the user experience

As briefly discussed in 4.1.4 a user's total QoE is based on more factors than purely video stream quality. The user interface was reported as one of these factors and in order to maximize the user experience this should be dealt with.

From our tests we found that the VG LIVE interface was "messy" and that it was difficult to find the important information. A redesign could bring more focus to the content (the video stream) and the information that the user wants. This however is out of our thesis scope and should be further researched.

4.3 Phase 3: How to receive feedback on the QoE

In this section we will describe some practical methods of monitoring the QoE in Internet television, based on the system used in our case.

4.3.1 Headend

By headend we mean the stages before the compressed video content is transmitted over the Internet, in our case the headend is in Bergen.

At headend we can assess the video quality using one of the objective quality assessment methods mentioned in 2.3.2. Since the reference video is available we can use a full reference (FR) method (see section 2.3.3).

The choice of metric would be depending on the accuracy, complexity, cost and the need for real-time measurements. PSNR, a data metric, is simple to implement real-time, however it lacks in accuracy. HVS-based metrics, such as MPQM, are very accurate, but likewise complex in implementation.

A compromise between the two could be an engineering approach picture metric, such as an implementation of VQM as proposed by ITU-T or VQEG [38, 65]. VQM measures blurring, noise, color distortions and problems with motion flow, something that fits well with the QoE problems we found in our results. The metric compares the reference with encoded video over short sequences and calculates the distortion factor.

Using VQM at the headend would give the service provider a good approximate of the video quality for the various bit rates.

4.3.2 Network transmission

The open Internet has no guarantee of QoS, and variable bandwidth, transmission jitter and packet loss makes it very difficult to control.

Through the Move Media Player clients the service provider is able to monitor the user's bit rate, start, stop and pause of segments, the number of segments watched and CPU/GPU usage on the user's equipment. This gives us some clue towards the user's available bandwidth and network reliability:

1. A very variable bit rate would suggest an unreliable network connection with variable available bandwidth and probably also some transmission jitter and packet delays.

2. A generally low bit rate would suggest a bottleneck in the access network (given there are enough processing resources).

3. Re-requesting of packets would suggest a high packet loss rate.

By knowing the bit rate a user receives we would also know the maximum video quality he/she would have, by the objective quality assessment at headend. Also, a very variable bit rate would mean that the video quality periodically would get better and worse, something that for many is more annoying than a constantly low quality.

4.3.3 User end

The user end is usually the most difficult to measure. We propose two practical ways of gaining an insight into the user's equipment and preferences/expectations.

User profile

Since all users of Absolutt Fotball have to register a profile in order to use the service, this profile could be extended to include user equipment and psychographic information.

As of today users register their favorite teams in Absolutt Fotball, something that gives them a "losing insurance" when buying subscription to single matches. This is a step towards understanding and satisfying the user preferences towards content. It could also include information about the user's screen and viewing context, processing power and other preferences. The information could be parameterized in order to get a quality measure, or impairment measure. This would require further research.

Of course this would have to be voluntary and it would not be guaranteed that all users would submit the information.

QoE menu

To get a further insight into the user's real-time QoE we propose a QoE menu that could be implemented into the service UI (i.e. on the web page). This menu could consist of 3 main parts: *(1) QoE guide and test*, *(2) Support and feedback* and *(3) Service personalization/differentiation.* A mock-up of the QoE menu can be seen in figure 4.10.

(1) could include the interactive user guide and real time test as described in 4.2.3. This would allow the user some influence on his/her own QoE, and a guidance in how to maximize it.

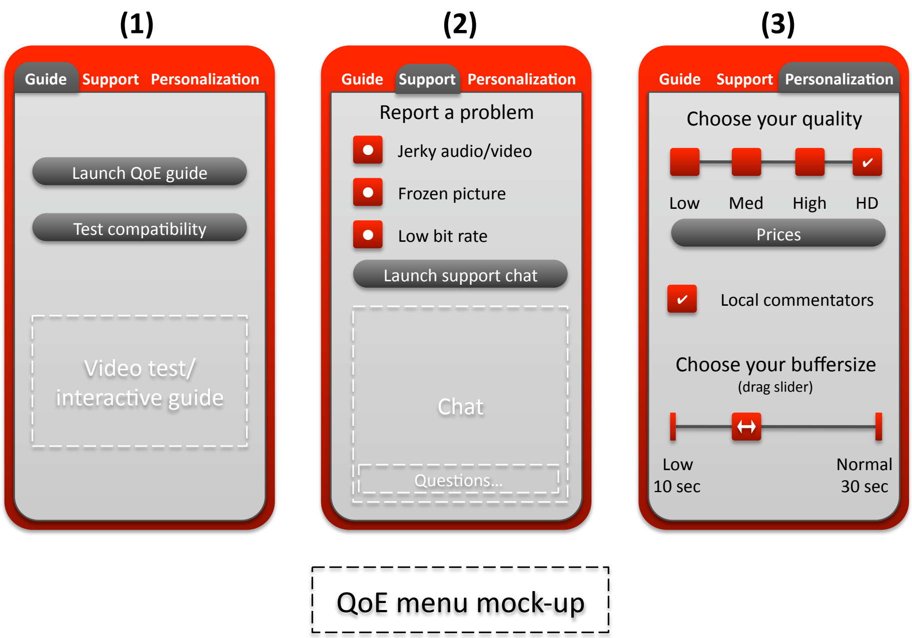


Figure 4.10: *A mock-up of the proposed QoE menu.*

(2) could include specific feedback on known QoE problems, which could be submitted real-time to service provider. It could also include a real-time support-chat where the user could get help to localize and resolve compatibility and QoE problems.

(3) could allow the users to somewhat control their own QoE in relation to price, according to their own preferences and needs. It could include the option to get HD quality content (for a premium price), reduce or improve quality/price, add features such as local commentators, or even choose their buffer size in relation to reliability (for the more advanced users).

The individual features, descriptions and labels (text) would of course have to be modified to fit the general user's interest and technical knowledge, however figure 4.10 shows how the QoE menu could be implemented theoretically.

Our tests have shown that users have a wide variety of preferences; for some a reduction in buffer size would be very important, others would gladly pay a little bit more for a higher quality. A QoE menu could go a long way in satisfying individual needs.

4.3.4 Parameterization

From headend to user end we can extract quantifiable parameters that rate the quality or degradations at each stage. These parameters can be combined into one QoE parameter unique for each individual user at a given time. As a feedback device it would provide an effective way of monitoring the general level of QoE.

Raake et. al. proposed a model for parameter-based prediction of IPTV quality called the T-V-model [55]. Their model predicted the video quality on a scale from 0 to 100 (100 for best quality) based on an input quality and degradations from picture resolution, coding impairments, transmission errors, and processing and presentation impairments. A similar model could fit our case, based on the extracted parameters. This would require further work in order to assess the quantitative impairment factors related to subjective perceptions.

At present there is no clear link between a user's profile on VG LIVE or 100% Fotball and the Move Media Player client. The data from the clients are anonymous. Thus there is a need for further integration between client and user in order to get true end-to-end QoE monitoring for each individual customer.

Chapter 5

Discussion

As briefly discussed in 4.1.2 our user selection for the test may not cover the entire demography of Absolutt Fotball's customer base. Our selection was also above average interested in technology, as well as technological competence. This may have given some bias to the results, whereby we might have had different averages with a different selection. The recruitment of friends and friends of friends may also have resulted in biased answers. However, in order to make a more representative selection there would have to be done market analysis of the customer base as well as a much larger user selection. Based on Schibsted's estimations of their core customer group (male 20 to 35), we covered it adequately.

Since our selection is limited, and perhaps not entirely representative, the results should not be read as universal facts by its numbers. The results have in all generality showed the tendencies within the group, and many of them have confirmed our expectations previous to the tests. We believe that these tendencies would also be found within the Absolutt Fotball customer base, and that the QoE problems highlighted in 4.2 are the most important to improve in order to increase customer satisfaction with the service.

An important question that arose from our results is the importance of the content and context in which it is being watched. Content is likely more important for long sequences, as the user's focus and attention may drift over time. As we saw in 4.1.4 our results may have been influenced by wins or losses of favorite teams. As far as context goes, we didn't see any clear correlation between screen sizes and average quality scores. However there are more factors concerned with the viewing context, and a complete study would require much more resources and time what was available to us.

One respondent had seen a football match on his 17" laptop using an unstable 3G connection in a car, and reported decent overall quality scores, whereas other respondents would sit on their computer in their homes with a stable connection and report lower scores. This could be evidence that a person's expectations of quality is affected by viewing context, and that watching live streaming football in a car could be considered a luxury in itself, while sitting at home raises the expectations of comfort and picture quality. Also, for some the match was a background/side entertainment while e.g. surfing the Internet, while for others the match was the main entertainment of the evening. These aspects were not adequately covered in the questionnaires.

There may have been some lacks in our questionnaire in order to cover all QoE aspects. For instance our questionnaire didn't cover the intra-match quality variance (quality variations within the same match/sequence), only inter-match quality variance (quality variations from match to match). The quality variations within the same segment can be very annoying to the viewer, and a correlation between quality variance and overall quality score could have shown some interesting results. Also, the questions with regards to coding artifacts could have been more specified, in order to distinguish blurring from edge ringing and color bleeding. Other problems included audio and video out of sync, delay from other live services and user interface, however these problems were limited to certain users or video sequences.

Our analysis in phase 2 (4.2) and phase 3 (4.3) is based on our limited technical knowledge of the Move Media Player. The technical information we have gathered has mostly come from our contacts in Schibsted, who have been working with implementing the player and streaming service. Our assumptions on how the client settings could have been tweaked may be wrong, however based on our information from Schibsted and general knowledge they are valid assumptions in concept.

Because of our limited time and resources, and that we didn't have access to Move's source code, we haven't been able to make a proof of concept. A good cooperation with Move (or any other Internet TV streaming provider) would have been preferable in order to get a good integration of user and client to implement the proposed QoE tools.

Finally, in our thesis we have shown how the QoE can be improved and how it can be monitored, but one important question we haven't asked is: "Is it worth it?" That is; is the improved quality and customer satisfaction worth the extra cost? This has been out of our scope for this thesis, but is without a doubt an important question in relation to our work. Our results showed that more than half of our selection would be willing to pay for the service at its present state. About half of those who didn't think it was worth it might have reconsidered with lower prices or higher quality, but should the service provider try to satisfy all, or will the general population be satisfied with the quality as is? This should be further researched in relation to costs of improvement, which we haven't covered in this thesis.

Chapter 6

Further Work

As discussed in the previous chapter further work is needed in order to assess the significance of content and context in the QoE. An important question here is if this can be parameterized and also how the service provider can handle content and context problems. Absolutt Fotball has already done some to improve the QoE with regards to content by offering a ”loss insurance” for the user's favorite team on their single round-subscriptions, although the question still remains if it is effective.

We found some tendencies that the user's favorite team, internet video habits, football viewing habits and presentation screen could have an influence on their quality ratings. Further work could address these questions specifically and expand to include the user's viewing environment and purpose of viewing (entertainment, wasting time, background noise while surfing the Internet or working) to get a greater picture of how the content and context affected their impression of the video quality and QoE. Another question is also how important the video quality would be in these situations for the overall QoE.

There need to be done more work in the parameterization and weighing of the different QoE factors. Thorough subjective measures could provide an insight onto how much the overall QoE is affected/degraded by watching the content on the ”wrong” screen or with varying degrees of different coding artifacts (related to bit rates). This could be used in a quantifiable end-to-end QoE measurement model as described in 4.3.

The surrounding factors, such as user interface and delay from other live services, showed to have some importance in our results. There needs to be done more research on how these surrounding factors affects the QoE in video streaming, and how it can be improved by the service provider.

The QoE menu as proposed by this thesis is only a rough sketch at the present time and further work is needed in order to optimize it for various Internet TV services. The QoE menu's content should cover various user needs, from basic to advanced, and at the same time be simple and user friendly to promote its use.

Chapter 7

Conclusions

In this thesis we have conducted user tests on live streaming Internet television. We have found that although providing a satisfying quality compared to other Internet video services there is a need for improving the QoE in order to gain a greater share of the customers who usually watch football on TV. Our results show that quality and price are important factors for potential customers.

We found ways to improve the QoE based on our case in Absolutt Fotball. The most critical factors seemed to be coding artifacts, video flow, response time and adaptation time for short sequences. Coding artifacts and video flow can in most cases be improved by increasing the bit rate of the encoding or changing to more effective codecs. The user's processing power and bandwidth could be a problem if the bit rate is too high, and in the cases of this being the bottleneck, increasing the bit rate wouldn't help. Response time can be improved by reducing the starting bit rate, but this is also much up to the user's available bandwidth. Finally, using a regular constant quality video stream instead of adaptive streaming could eliminate the problem with adaptation time for short sequences. All respondents in our survey preferred adaptive streaming for long sequences.

Factors other than video quality, such as content, context, user interface and delay from other live services, seemed to affect the user's overall impression of the QoE. Content, context and user interface are factors that need to be further researched in order to assess its impact on the QoE and how it could be improved. The delay from other live services could be improved by reducing the client buffer, however this would come at the cost of reduced robustness against network errors.

We found that the QoE is difficult to control because of the lack of QoS in the network transmission and the huge variance in user equipment. The QoE could be objectively measured and monitored by using a combination of objective quality assessment at headend, an extensive user profile and real-time user feedback through a QoE menu at user-end, and a better integration between client and user in order to assess the network QoS. The integration between client and user is key in order to monitor the end-to-end QoE. At the present time Move Networks provide their clients with user statistics on bandwidth and segments watched etc. This is not linked up to individual users and is therefore not enough in assessing the user's QoE.

We have proposed a QoE tool, in the form of a menu in the user interface, which would give the users more influence on their own QoE. The menu could include real-time compatibility tests and guides in how to improve the QoE in their viewing environment. It could include real-time feedback on certain QoE issues as well as a support chat. Finally, it could include a way of personalizing the service by being able to choose quality (related to price), service features (e.g. commentators) and more advanced service settings, such as buffer size in relation to delay from other live services.

Our results have shown that users have different preferences and needs. A QoE menu would support this and able them to optimize their own QoE. Further on, together with the QoE monitoring as proposed by this thesis, it would allow the service provider to differentiate its services by selling premium quality at a premium price and vice versa, thus reaching out to new customers, who earlier were dissatisfied with the quality or prices, and satisfying individual needs.

Appendix A

Pre-analysis

The Move Media Player is based on adaptive streaming, thus the user will experience a period in the start of each stream where the quality is quite low at first and gradually improving towards the best quality possible on the user's internet connection. The time of this *adaptation-period* could be critical for the user's overall perception of the video quality, especially since VG LIVE offers both long sequences (whole matches) as well as short sequences of 1 minute (highlights).

Certain coding artifacts became apparent, especially at low bit rate (e.g. during the adaptation period). *Blurring*, as illustrated in figure A.1, will make a lot of details disappear on the players and ball, as well as on the grass field and in the surroundings. *Ringing* around the edges, as shown in figure A.2, together with *color bleeding*, as shown in figure A.3, are also contributing to the loss of detail, as the edges become less clear. This makes it more difficult to recognize the players by face or player number, and text-elements become harder to read. Players and the football often move very fast on the field, whereby *motion blurring* becomes apparent. This makes it harder for the viewer to follow the ball or players and is illustrated in figure A.4 where the ball seems oval instead of round.

Figure A.1: *An example of blurring in the VG LIVE stream. The grass field loses its details and the players become more difficult to recognize.*

Figure A.2: *An example of ringing in the VG LIVE stream. The edges become less distinct (high frequency parts of the picture).*

There was also some issues with the natural flow of the video and audio, namely *jerkiness* and *choppy audio*. Sometimes the video or audio would even stop for a little while leaving a *freezed image*, *black screen* or *loss of audio*.

One of the advantages with the Move Media Player is the instant start. ”Instant start” is a subjective term and some users might find it not quite as ”instant”. This could become apparent when starting a new clip, forwarding or rewinding in a clip, or when pushing the ”Live”-button to return to live mode.

Stress testing

We stress tested the adaptive streaming using a bandwidth limiter software (Netlimiter) in Windows XP on an Apple MacBook. There we could see how the video quality gradually would be getting worse from highest bit rate to lowest bit rate.

Descending from highest bit rate to approximately 160 kbps the video would flow as normal (little or no jerkiness) while the image quality gradually was getting worse. Especially blurring were apparent, but also some more edge ringing and color bleeding.

Below 160 kbps the video would become more jerky (audio still good) until below 80 kbps it would start to freeze for short periods. On the lowest bit rate, 32 kbps, the video barely continued. There were constant stops in audio and video, and the image was so blurry it was impossible to distinguish the players from each other. The audio was still pretty good on low bit rates.

Apple MacBook specifications: 2.0 GHz Intel Core Duo processor, 2GB DDR2 SDRAM, NVIDIA GeForce 9400M (256MB VRAM).

Figure A.3: *An example of color bleeding in the VG LIVE stream. The color from the grass and football player is mixed together around the edges, making it harder to separate the two.*

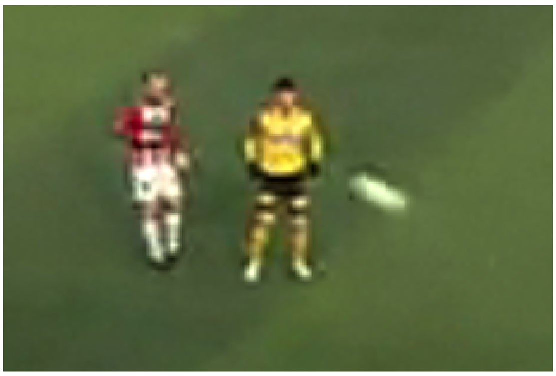


Figure A.4: *An example of motion blurring in the VG LIVE stream. The ball is more oval than round.*

Appendix B

Questionnaires

Brukerregistrering

Exit this survey

1. Brukerdetaljer

***1. E-post**

***2. Brukernavn på VGlive**

***3. Alder**

***4. Kjønn**

- Mann
- Kvinne

***5. Yrke/studie**

- Skoleelev
- Student
- I fast arbeid
- Arbeidsledig
- Pensjonert

Hvilket yrke/studie?

***6. Ditt forhold til (ny) teknologi**

- Veldig interessert i all ny teknologi
- Over gjennomsnittlig interessert
- Gjennomsnittlig interessert
- Under gjennomsnittlig interessert
- Lite interessert i ny teknologi

***7. Hvor ofte ser du fotball?**

- Flere ganger i uka
- Hver uke

- En eller flere ganger i måneden
- Sjelden

***8. Hvor/hvordan ser du som oftest på fotball? (velg en eller flere)**

- På stadion
- På pub/bar
- Hjemme i stua på TV (ikke HD)
- Hjemme i stua på TV i HD
- På PC, streaming over internett

Annet

***9. Hvor ofte ser du video på internett?**

- Hver dag
- Flere ganger i uka
- En gang i uka
- En eller flere ganger i måneden
- Sjelden

***10. Hvilke internett-video tjenester bruker du?**

- Youtube
- NRK Nett-TV
- TV2 (Sumo)
- Nettaviser (VG, Dagbladet, Aftenposten etc.)

Andre tjenester

Done

Utstyr spesifisering

Exit this survey

1. Noen spørsmål ang. det utstyret du kommer til å bruke under testperioden

*** 1. E-post**

*** 2. Hvilken skjermtype?**

- Dataskjerm, LCD (flatskjerm/laptop)
- Dataskjerm, CRT
- LCD TV (flatskjerm)
- CRT TV
- Plasma TV (flatskjerm)
- Projektor

Annet/kommentar

*** 3. Hvor stor er skjermen?**

- 12" eller mindre
- 13" - 16"
- 17" - 22"
- 23" - 32"
- Større enn 32"

*** 4. Hvilken oppløsningsklasse har skjermen? (velg nærmeste)**

- 800x600
- 1024x768
- 1280x800
- 1280x960
- 1440x900
- 1440x1050
- 1600x1200
- 1680x1050
- 1920x1200
- Standard TV-format
- HDTV 720p
- Full-HDTV 1080p

*** 5. Hvor gammel er datamaskinen?**

- Mindre enn 1 år
- 1-3 år
- 3-5 år
- Mer enn 5 år

Vennligst spesifiser prossessor (CPU) og minne (MB/GB RAM) - hvis du vet

***6. Hvilket operativsystem bruker du?**

- Windows
- Mac OSX
- Linux

***7. Hvilken internettbrowser bruker du?**

- Internet Explorer
- Firefox
- Opera
- Safari
- Google Chrome

Annet/kommentar

***8. Hvilken internettoppkobling har du?**

- EDGE/3G/Super-3G
- ADSL/Kabel under 2 Mbit
- ADSL/Kabel 2-6 Mbit
- ADSL/Kabel/Fiber over 6 Mbit

Annet/kommentar

Done

Kvalitetsanalyse

Exit this survey

1. Default Section

***1. E-post**

2. Hvilken kamp så du?

Hjemmelag

Bortelag

3. Hvor god vil du si den generelle kvaliteten på videostreamingen på VGlive var hvis du sammenligner med:

	VGlive var mye dårligere	VGlive var litt dårligere	Omtrent like bra	VGlive var litt bedre	VGlive var mye bedre
Andre videotjenester på internett (Youtube, NRK Nett-TV, TV2 Sumo, nettaviser, etc.)					
Fotballkamper på TV (generelt)					

Evt. kommentar

4. Hvor fornøyd var du med:

	Veldig dårlig	Dårlig	Tilfredsstillende	Bra	Veldig bra
Hvor godt du så spillere/dommer/ball					
Hvor godt du så tekstlige beskjeder ved scoring/bytte etc.					
Detaljenivået på gress/publikum etc.					

Evt. kommentar

5. Den nye streamingteknologien som er i bruk gjør at bildekvaliteten i streamingen først er litt dårlig for så å bli bedre etter litt tid. Hvor sjenerende synes du dette var på en skala fra 1 til 5 hvor 1 er "ikke sjenerende i det hele tatt"

og 5 er "helt uakseptabelt"?

	1 - Ikke sjenerende i det hele tatt	2	3	4	5 - Helt uakseptabelt
Lange klipp (hovedkampen)					
Korte klipp (1 minutts repriser)					

Evt. kommentar

6. Tiden det tar fra du f.eks trykker "Se kampen her" eller "Live" til videoen spiller kalles responstid. Hvor fornøyd var du med responstiden på:

	Veldig dårlig	Dårlig	Tilfredsstillende	Bra	Veldig bra
Start av nytt klipp - fra du trykte "Se kampen her" til videoen spilte					
Live-funksjonen - fra du trykte "Live" til videoen spilte live					
"Spoling" - fra du flyttet tidsmarkøren til videoen spilte fra den nye posisjonen					

Evt. kommentar

7. La du merke til hakking eller ujevnheter i lyd eller video i løpet av streamingen?

	Nei	Noen få ganger	Mange ganger	Hele tiden
Lyd				
Video				

Evt. kommentar

8. Hendte det seg at lyd eller bilde forsvant/hang seg opp under streamingen?

	Nei	Noen få ganger	Noen få ganger, av og til over lengre	Mange ganger	Mange ganger, av og til over lengre tid

			tid		
Lyd forsvant (stille)					
Bilde forsvant (svart/blå skjerm)					
Lyd hang seg opp (støy/gjentatt lyd)					
Bilde hang seg opp (det samme bildet over lengre tid)					

Evt. kommentar

9. Evt. andre feil eller problemer med videostreamingen som du la merke til

10. Andre relevante kommentarer. (eks. om du satt på et annet nett enn du spesifiserte i før-undersøkelsen, brukte annet utstyr/skjerm, kjørte tunge prossesser i bakgrunnen eller lastet ned data samtidig, så flere kamper samtidig, etc.)

Done

Etterundersøkelse

Exit this survey

1.

*** 1. E-post**

[ ]

*** 2. Ut i fra dine forventninger før du begynte å bruke VG LIVE, ble du positivt eller negativt overrasket av videokvaliteten?**

- Positivt
- Nøytralt
- Negativt

Evt. kommentar

[ ]

*** 3. På en skala fra 1 til 10, hvor 1 er lav-kvalitets kornete videoer på Youtube eller lignende og 10 er kvaliteten på TV-sendte kamper, hvor god vil du si videokvaliteten på VG LIVE er?**

	1 (kornete Youtube-kvalitet)	2	3	4	5	6	7	8	9	10 (TV-kvalitet)
Videokvalitet										

Evt. kommentar

[ ]

*** 4. Vurder kvaliteten på følgende funksjoner ved tjenesten**

	Veldig dårlig	Dårlig	Nøytral/har ingen mening	Bra	Veldig bra
Generell bildekvalitet					
Rask start på streamingen					
Streamingen stopper ikke opp under kampen					
Reprise-klippene fra mål/sjanser fra andre kamper					
Adapsjonstiden fra dårlig kvalitet i starten til best kvalitet					

	Veldig dårlig	Dårlig	Nøytral/har ingen mening	Bra	Veldig bra
Menyene rundt spilleren	◯	◯	◯	◯	◯
VG LIVE nettsiden generelt	◯	◯	◯	◯	◯
Muligheten for å "spole" i kampene	◯	◯	◯	◯	◯
Live-funksjonen, dvs. hvor lang tid etter TV, radio etc. VG LIVE sine sendinger lå	◯	◯	◯	◯	◯

Annet/kommentar

[ ]

***5. Den nye streamingteknologien som er brukt gjør at det tar en periode før bildekvaliteten er på topp, men samtidig starter kampen kjapt og den stopper ikke midt i for å "bufre". Dette kalles "adaptiv streaming".**

Det andre alternativet er en fast videokvalitet, men der man må vente litt på videoen i starten og der den av og til stopper midt i for å "bufre". (eks. på Youtube eller NRK Nett-TV)

Synes du VG LIVE sin løsning med adaptiv streaming er ønskelig?

◯ Ja (foretrekker streaming med kjapp start og ingen bufring)

◯ Nei (foretrekker streaming med fast kvalitet)

Evt. kommentar

[ ]

***6. I løpet av testperioden, ble inntrykket ditt av kvaliteten på videostreamen bedre eller dårligere?**

	Mye dårligere	Dårligere	Like bra	Bedre	Mye bedre
Inntrykket ble	◯	◯	◯	◯	◯

Evt. kommentar

[ ]

***7. I løpet av testperioden, var kvaliteten på videostreamen stort sett den samme fra kamp til kamp, eller varierte den? (gitt at du satt med samme datamaskin og samme internettilkobling)**

	Var stort sett den samme	Varierte litt	Varierte en del	Varierte hver kamp	La ikke merke til det eller så

	fra kamp til kamp				ikke nok kamper på samme utstyr
Kvaliteten	◯	◯	◯	◯	◯

Evt. kommentar

*** 8. Prisen for årskort på VG LIVE er 999,-, månedsbillett 129,- og rundebillett (m/tapsforsikring) 99,-**

Med det du vet nå etter å ha testet tjenesten, ville du ha betalt for et abonnement i fremtiden?

- ◯ Årskort (999,-)
- ◯ Månedsbillett (129,-)
- ◯ Rundebillett (99,-)
- ◯ Nei

Hvis nei, hvorfor ikke? (Hadde introduksjonstilbudet på 499/89/59,- endret svaret ditt?)

*** 9. Hva ville gjort tjenesten bedre for deg som kunde? (Hva kunne bedret opplevelsen din av tjenesten?)**

	Ikke interessant i det hele tatt	Lite interessant	Interessant	Veldig interessant
HD-kvalitet	◯	◯	◯	◯
Mulighet til å betale mindre ved dårligere kvalitet	◯	◯	◯	◯
Mulighet til å velge kvalitet og tilhørende pris	◯	◯	◯	◯
Support-chat (MSN e.l.) for spørsmål ang. kvalitet og bruken av tjenesten	◯	◯	◯	◯

Annet/kommentar

Done

Appendix C

Survey results

Demographical/psychographical survey		
Answer Options	Response Frequency	Response Count
Respondents	100,0%	18
Age		
18-23	22,2%	4
24-30	44,4%	8
31-40	0,0%	0
41-55	33,3%	6
Gender		
Male	77,8%	14
Female	22,2%	4
Occupation		
Pupil	0,0%	0
Student	38,9%	7
Working	61,1%	11
Out of work	0,0%	0
Retired	0,0%	0
Attitude towards (new) technology		
Very interested in all new technology	27,8%	5
Above average interested	38,9%	7
Average interested	22,2%	4
Below average interested	11,1%	2
Little interested in new technology	0,0%	0
How often do you watch football?		
Several times a week	44,4%	8
Weekly	38,9%	7
Once or more a month	16,7%	3
Rarely	0,0%	0
Where/how do you usually watch football? (multiple answers)		
At the stadium	33,3%	6
In a pub/bar	22,2%	4
In the living room (non-HD)	72,2%	13
In the living room in HD	33,3%	6
On a computer, streaming over the Internet	33,3%	6
How often do you watch video on the Internet?		
Daily	22,2%	4
Several times a week	27,8%	5
Weekly	5,6%	1
Once or more a month	16,7%	3
Rarely	27,8%	5
Which Internet video services do you use? (multiple answers)		
Youtube	83,3%	15
NRK Nett-TV	55,6%	10
TV2 (Sumo)	11,1%	2
News sites (VG, Dagbladet, Aftenposten etc.)	66,7%	12

User equipment specification

Which type of screen will you be using?		
Computer screen, LCD	72,2%	13
Computer screen, CRT	5,6%	1
LCD TV	22,2%	4
CRT TV	0,0%	0
Plasma TV	0,0%	0
Projector	0,0%	0
What is the size of the screen you will be using?		
12" or less	0,0%	0
13" - 16"	33,3%	6
17" - 22"	38,9%	7
23" - 32"	5,6%	1
Bigger than 32"	22,2%	4
What is the resolution of the screen you will be using?		
1024x768	5,6%	1
1280x800	22,2%	4
1280x960	11,1%	2
1440x900	16,7%	3
1440x1050	0,0%	0
1600x1200	11,1%	2
1680x1050	0,0%	0
1920x1200	11,1%	2
Standard TV-format	0,0%	0
HDTV 720p	5,6%	1
Full-HDTV 1080p	16,7%	3
How old is the computer you will be using?		
Less then 1 year	33,3%	6
1-3 years	55,6%	10
3-5 years	11,1%	2
More than 5 years	0,0%	0
What type of OS will you be using?		
Windows	66,7%	12
Mac OSX	33,3%	6
Linux	0,0%	0
Which type of Internet browser will you be using?		
Internet Explorer	27,8%	5
Firefox	50,0%	9
Opera	0,0%	0
Safari	22,2%	4
Google Chrome	0,0%	0
How fast is your Internet connection?		
EDGE/3G/Super-3G	0,0%	0
ADSL/Kabel under 2 Mbit	0,0%	0
ADSL/Kabel 2-6 Mbit	77,8%	14
ADSL/Kabel/Fiber over 6 Mbit	22,2%	4

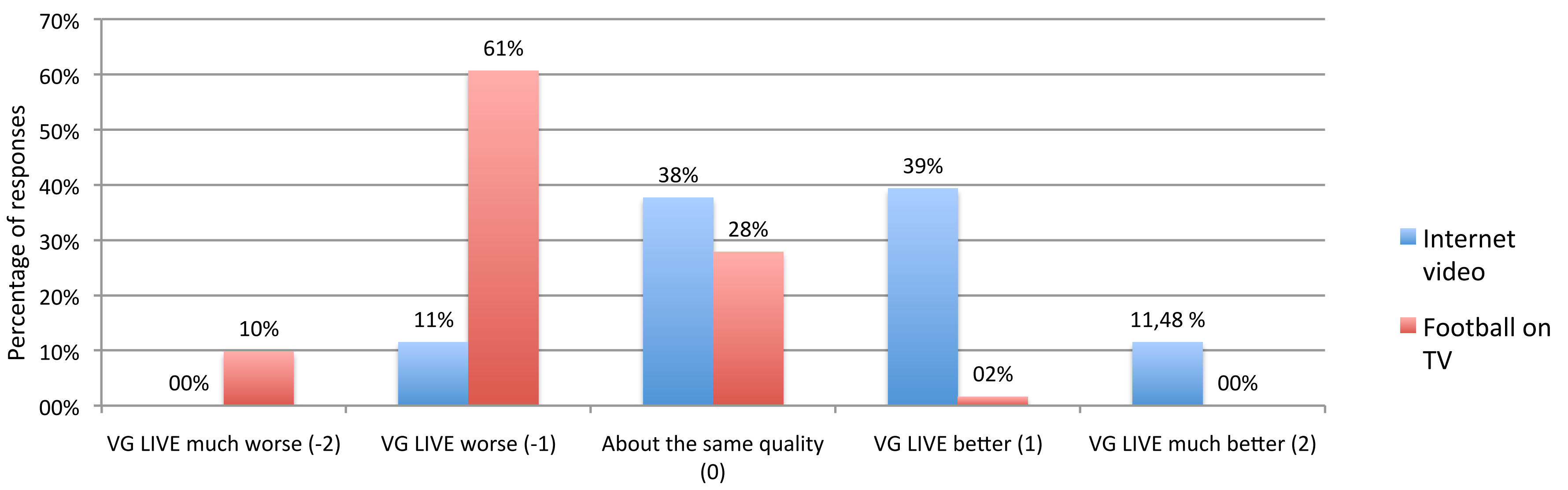
Quality analysis

How good would you say the video quality of VG LIVE is compared to:

Answer Options	VG LIVE much worse (-2)	VG LIVE slightly worse (-1)	About the same (0)	VG LIVE better (1)	VG LIVE much better (2)	Rating Average	Response Count
Other Internet video services	0	7	23	24	7	0,51	61
Football on TV	6	37	17	1	0	-0,79	61

	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Internet video	0,51	0,85	0,21
Football on TV	-0,79	0,64	0,16

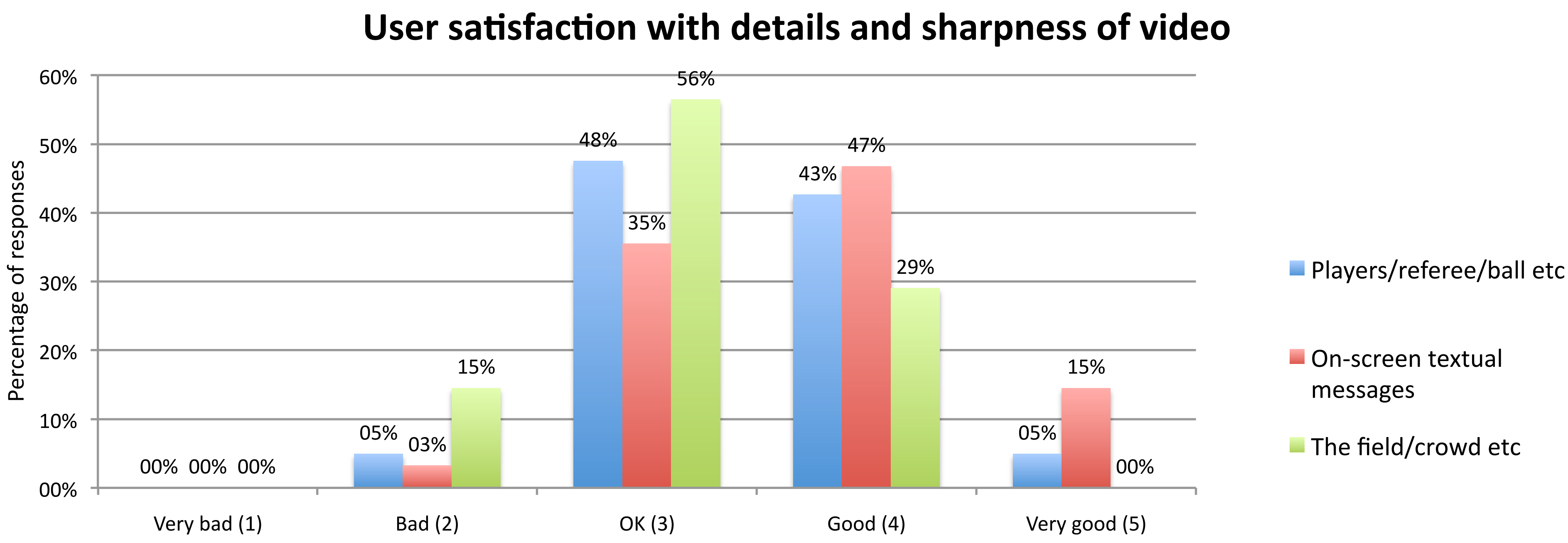
Comparison of the quality of VG LIVE and other services



How satisfied were you with:

Answer Options	Very bad (1)	Bad (2)	OK (3)	Good (4)	Very good (5)	Rating Average	Response Count
How well you saw players/referee/ball	0	3	29	26	3	3,48	61
How well you saw on-screen textual messages	0	2	22	29	9	3,73	62
Level of detail on the field/crowd etc.	0	9	35	18	0	3,15	62

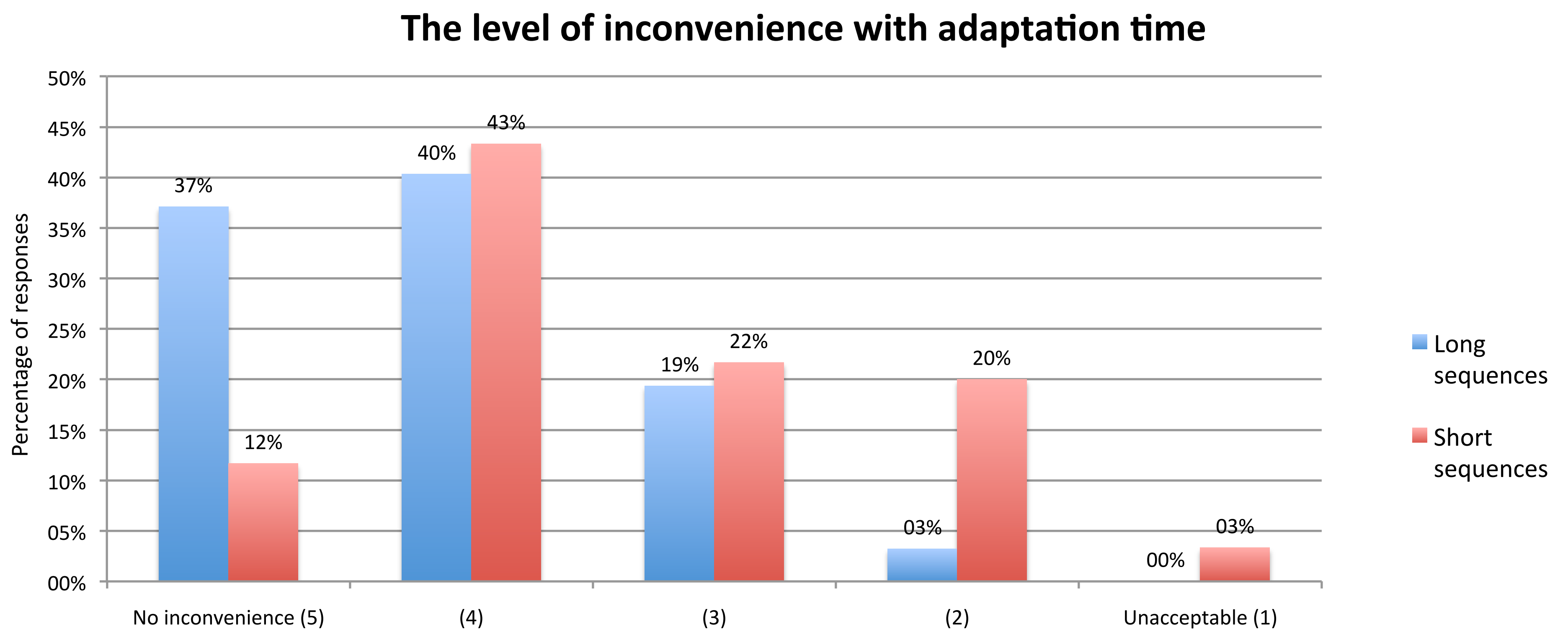
	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Players/referee/ball etc	3,48	0,67	0,17
On-screen textual messages	3,73	0,75	0,19
The field/crowd etc	3,15	0,65	0,16



The adaptation time is the period from the beginning of a sequence until it has reached maximum quality. How much of an inconvenience was the adaptation time?

Answer Options	No inconvenience (5)	(4)	(3)	(2)	Unacceptable (1)	Rating Average	Response Count
Long sequences (main match)	23	25	12	2	0	4,11	62
Short sequences (highlights)	7	26	13	12	2	3,40	60

	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Long sequences	4,11	0,83	0,21
Short sequences	3,40	1,04	0,26

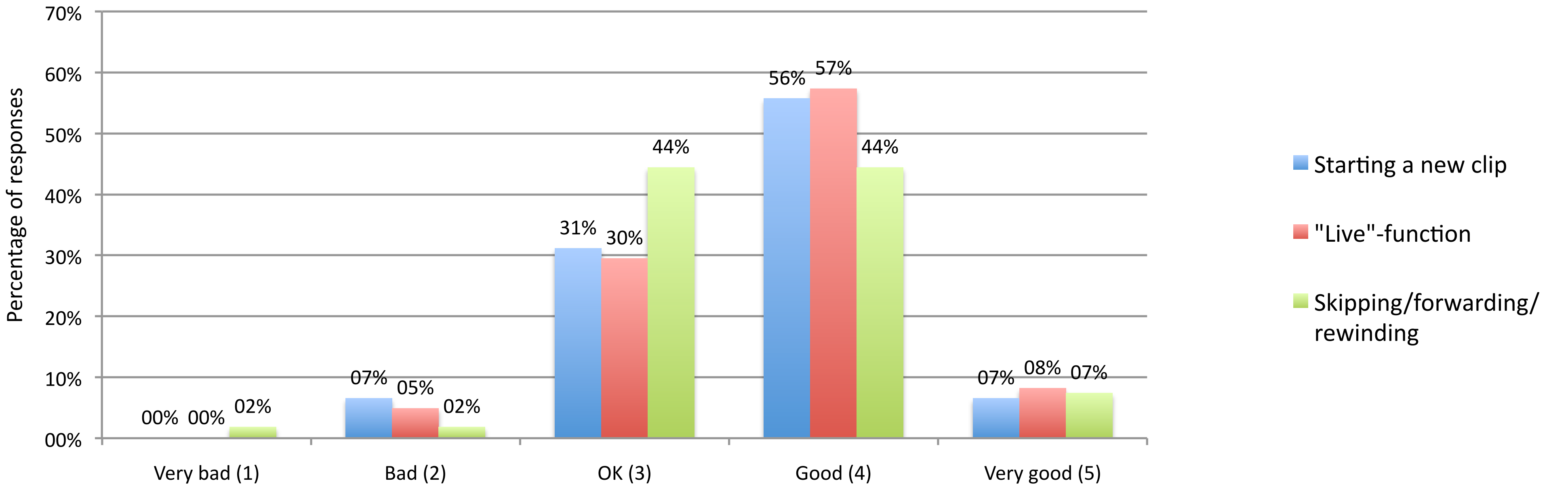


The time it takes from you push e.g. "play" until the video stream starts is called the response time. How satisfied are you with the response time?

Answer Options	Very bad (1)	Bad (2)	OK (3)	Good (4)	Very good (5)	Rating Average	Response Count
Starting a new clip	0	4	19	34	4	3,62	61
"Live"-function	0	3	18	35	5	3,69	61
Skipping/forwarding/rewinding	1	1	24	24	4	3,54	54

	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Starting a new clip	3,62	0,71	0,18
"Live"-function	3,69	0,70	0,17
Skipping/forwarding/rewinding	3,54	0,75	0,20

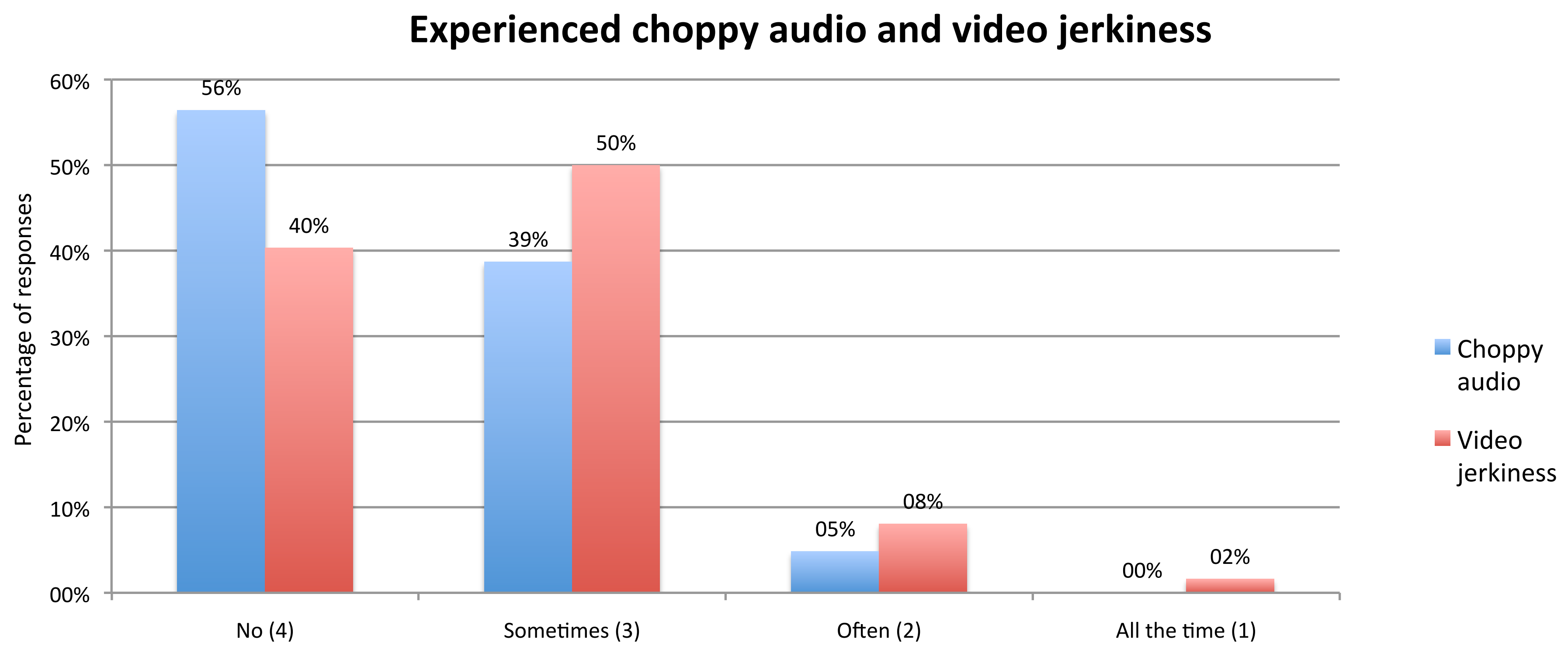
User satisfaction with response time



Did you notice any skipping or lack of smoothness in sound or video during the stream?

Answer Options	No (4)	Sometimes (3)	Often (2)	All the time (1)	Rating Average	Response Count
Sound	35	24	3	0	3,52	62
Video	25	31	5	1	3,29	62

	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Choppy audio	3,52	0,59	0,15
Video jerkiness	3,29	0,69	0,17

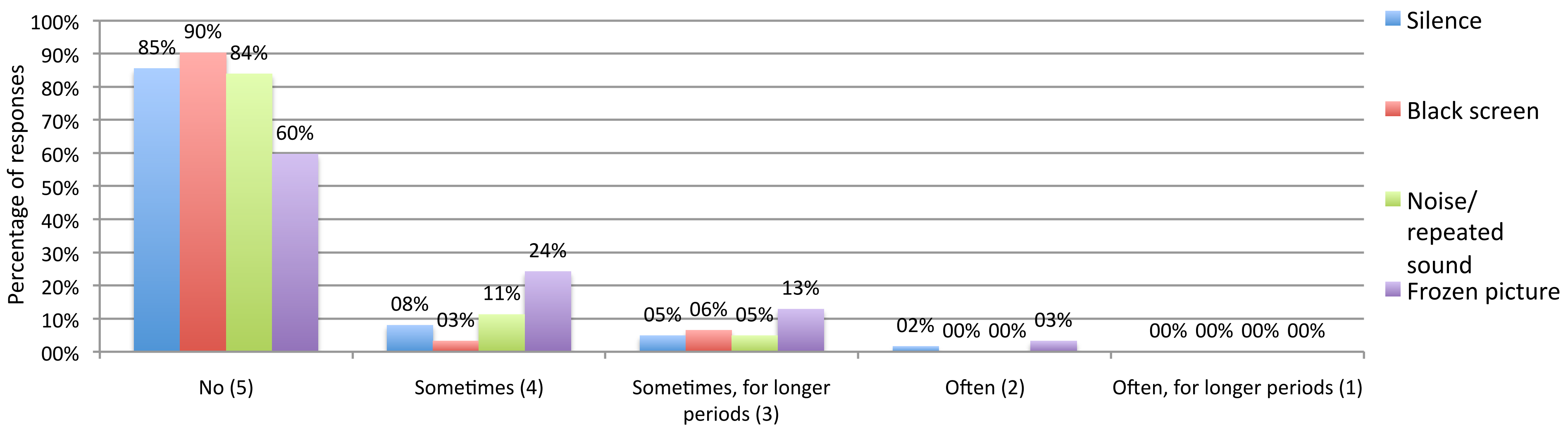


Did sound or picture disappear or freeze during the stream?

Answer Options	No (5)	Sometimes (4)	Sometimes, for longer periods (3)	Often (2)	Often, for longer periods (1)	Rating Average	Response Count
Sound disappeared (silence)	53	5	3	1	0	4,77	62
Picture disappeared (black screen)	56	2	4	0	0	4,84	62
Sound froze (noise/repeated sound)	52	7	3	0	0	4,79	62
Picture froze	37	15	8	2	0	4,40	62

	MEAN	STANDARD DEVIATION	95% CONFIDENCE
Silence	4,77	0,61	0,15
Black screen	4,84	0,52	0,13
Noise/repeated sound	4,79	0,52	0,13
Frozen picture	4,40	0,84	0,21

Experienced sound and video blackouts and freezes

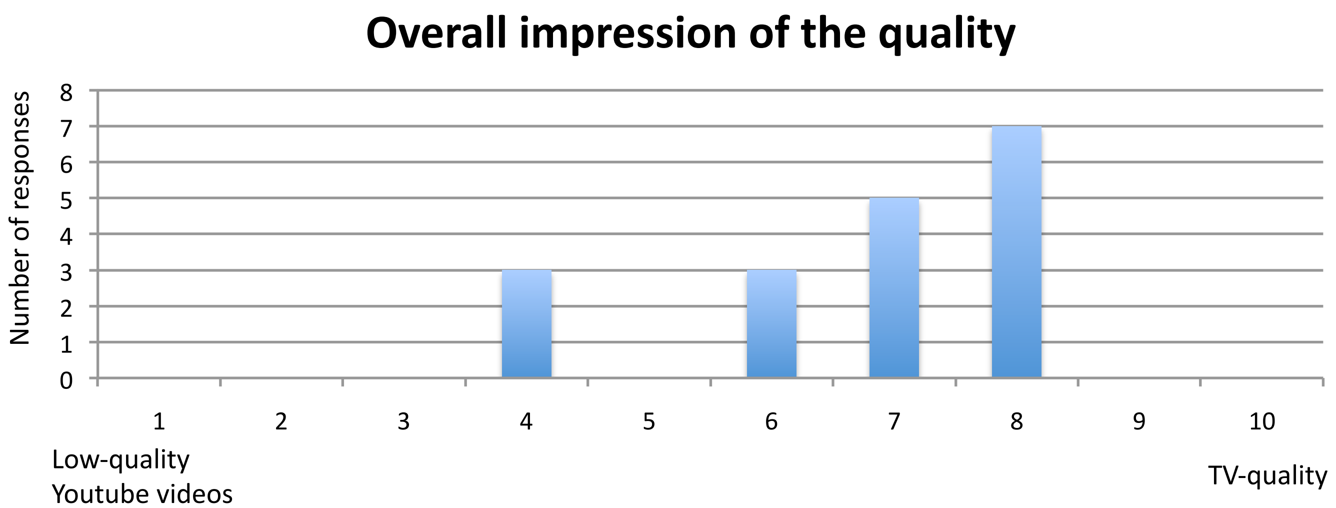


Post-survey

Answer Options	Response Frequency	Response Count
Respondents	100,0%	18
Based on your expectations before testing the service, were you posively of negatively surprised?		
Positive	61,1%	11
Neutral	38,9%	7
Negative	0,0%	0

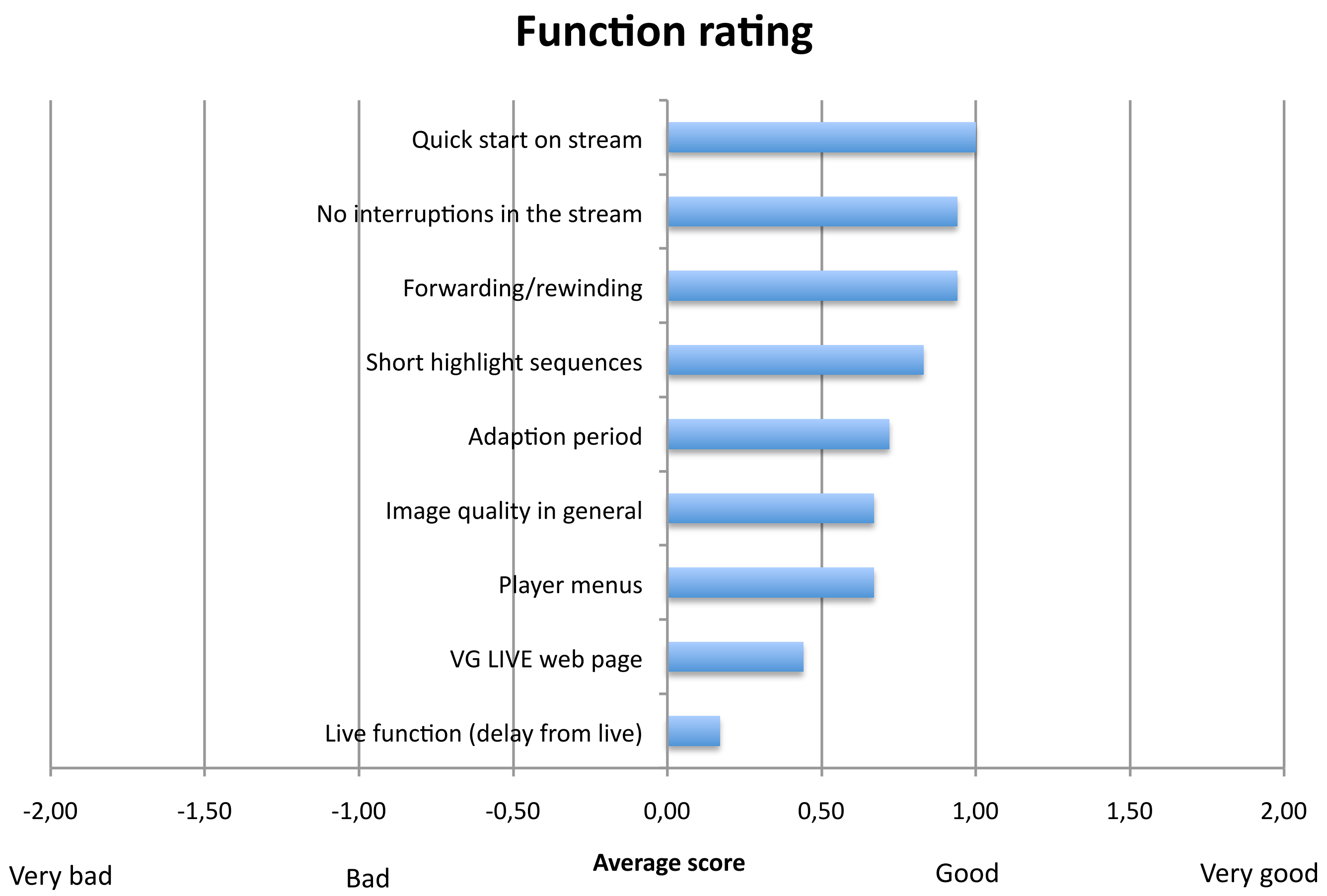
On a scale from 1 to 10, where 1 equals low-quality Youtube videos and 10 the quality of TV matches, how good would you say the overall quality of VG LIVE is?

Answer Options	1	2	3	4	5	6	7	8	9	10	Rating Average	Response Count
Video quality	0	0	0	3	0	3	5	7	0	0	6,72	18



Evaluate the following functions in VG LIVE

Answer Options	Very bad (-2)	Bad (-1)	Neutral/ No opinion (0)	Good (1)	Very good (2)	Rating Average	Response Count
Image quality in general	0	1	5	11	1	0,67	18
Quick start on stream	0	1	2	11	4	1,00	18
No interruptions in the stream	0	1	3	10	4	0,94	18
Short highlight sequences	0	1	5	8	4	0,83	18
Adaption period	0	0	6	11	1	0,72	18
Player menus	0	1	5	11	1	0,67	18
VG LIVE web page	1	3	3	9	2	0,44	18
Forwarding/rewinding	0	1	3	10	4	0,94	18
Live function (delay from live)	1	4	6	5	2	0,17	18



Do you prefer adaptive streaming with quick start and no interruptions over streaming with a constant quality, some startup delay and interruptions?

Answer Options	Response Frequency	Response Count
Yes (prefer adaptive streaming)	100,0%	18
No, (prefer constant quality)	0,0%	0

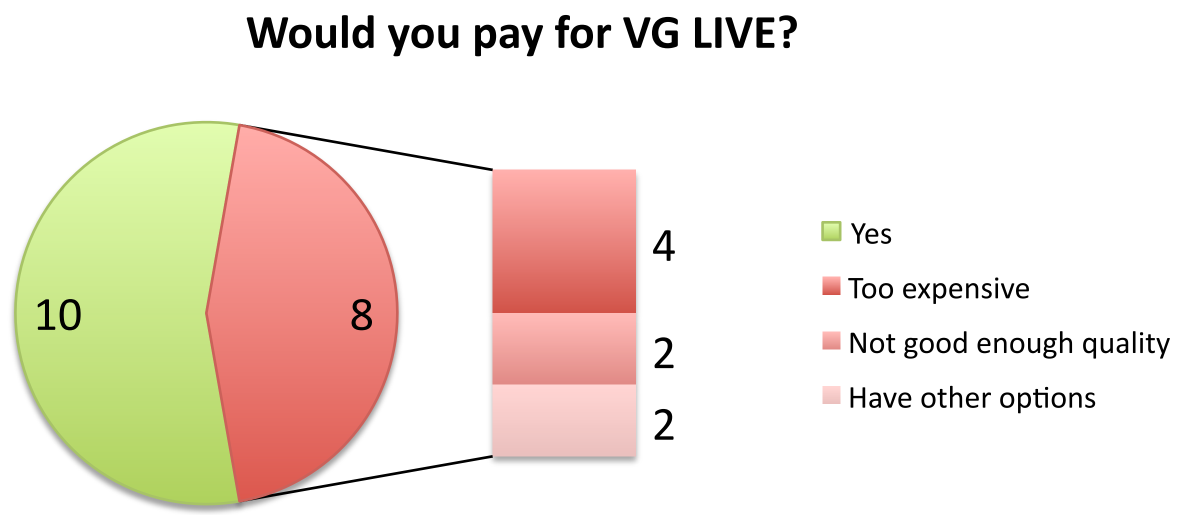
During the test period, did your impression of the video quality become better or worse?

Answer Options	Much worse (-2)	Worse (-1)	Neutral (0)	Better (1)	Much better (2)	Rating Average	Response Count
Impression was	0	1	10	6	1	0,39	18

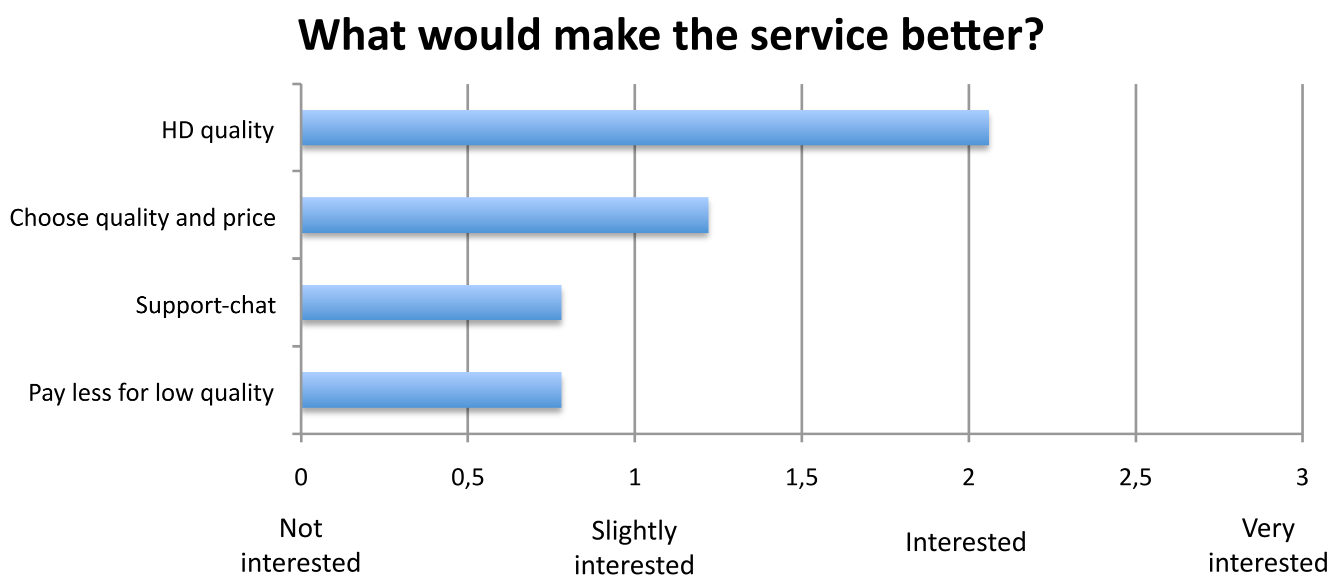
During the test period, did the video quality vary much from match to match or was it the same?

Answer Options	The same (4)	Some variations (3)	Much variations (2)	Varied for each match (1)	Did not notice (0)	Rating Average	Response Count
The video quality	7	8	3	0	0	3,22	18

Given the prices of subscriptions at 999,- annually, 129,- monthly or 99,- per round (w/losing insurance), would you pay for VG LIVE in the future? (multiple answers possible)		
Answer Options	**Response Frequency**	**Response Count**
Annually (999,-)	22,2%	4
Monthly (129,-)	44,4%	8
Per round (99,-)	0,0%	0
No	44,4%	8



What would make the service better?						
Answer Options	**Not interested (0)**	**Slightly interested (1)**	**Interested (2)**	**Very interested (3)**	**Average Rating**	**Response Count**
HD quality	0	4	9	5	2,06	18
Pay less for low quality	6	10	2	0	0,78	18
Choose quality and price	3	8	7	0	1,22	18
Support-chat	7	8	3	0	0,78	18



References

[1] Apple - QuickTime - h.264 - frequently asked questions. http://www.apple.com/quicktime/technologies/h264/faq.html.

[2] DivX certified profiles \textbar DivX labs - everywhere communication occurs community happens. http://labs.divx.com/node/72.

[3] Fotballen dropper betal-tv - kampanje.com. http://www.kampanje.com/medier/article377755.ece.

[4] Fotballen tilbake til TV-seerne - NA24.no. http://arkiv.na24.no/Nyhet/349999/Fotballen+tilbake+til+TV-seerne.html.

[5] IPTVTroubleshooter.com - TVQM. http://www.iptvtroubleshooter.com/tvqm.html.

[6] List of codecs supported by adobe flash player. http://kb.adobe.com/selfservice/viewContent.do?externalId=kb402866&sliceId=1.

[7] Lover mer fotball til alle - NA24.no. http://arkiv.na24.no/Nyhet/314828/Lover+mer+fotball+til+alle.html.

[8] Mariner xVu. http://www.marinerxvu.com/.

[9] Move networks » home. http://www.movenetworks.com/.

[10] Overview of H.264/AVC and MPEG-4 codec products. http://radio.irt.de/docs/codecs.html.

[11] RECOMMENDATION ITU-R BT.500-11: methodology for the subjective assessment of the quality of television pictures. Tech. rep., INTERNATIONAL TELECOMMUNICATION UNION.

[12] Schibsted kutter fotballkostnader - NA24.no. http://pub.tv2.no/nettavisen/na24/propaganda/article2569366.ece.

[13] Sleivspark for nettfotballen - NA24.no. http://pub.tv2.no/nettavisen/na24/propaganda/article2576816.ece.

[14] Sorenson media - related products. http://www.sorensonmedia.com/products/?pageID=1&ppc=

[15] Sorenson media - sorenson store - sorenson video 3 pro codec. http://www.sorensonmedia.com/pages/?pageID=73.

[16] Streaming media west 2008: New trends in online video measurement.

[17] TV 2 skuffet over NFF - kampanje.com. http://www.kampanje.com/medier/article377890.ece.

[18] V-FACTOR - QoE assurance & quality management. http://www.symmetricom.com/products/qoe-assurance/v-factor/.

[19] VC-1 technical overview. http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.as

[20] VG Nett LIVE. http://vglive.no/.

[21] The video quality experts group web site. http://www.its.bldrdoc.gov/vqeg/.

[22] Visual networking index - cisco systems. http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html.

[23] Xvid.org: Project info. http://www.xvid.org/Project-Info.46.0.html.

[24] ITU-T recommendation p.910: Subjective video quality assessment methods for multimedia applications. Tech. rep., INTERNATIONAL TELECOMMUNICATION UNION, 1999.

[25] Erik Aaberg from Schibsted ASA about the Absolutt Fotball business model, 2009.

[26] Ketil Moland Olsen, producing manager of Absolutt Fotball, about the absolutt fotball technical architecture, 2009.

[27] Accenture. IPTV Monitor—Issue 4, 2007.

[28] Agilent Technologies. Agilent N2X technical data sheet: IPTV & IPTVv6 quality of experience (QoE) test solution.

[29] Agilent Technologies. White paper: Validating IPTV service quality under multiplay network conditions.

[30] Aldridge, R., Davidoff, J., Ghanbari, M., Hands, D., and Pearson, D. Recency effect in the subjective assessment of digitally-coded television pictures.

[31] Boavida, F., and Cerqueira, E. Benchmarking the quality of experience of video streaming and multimedia search services: the CONTENT network of excellence. *Telecommunication Review and Telecommunication News Journal, Technical Journals and Books Publisher SIGMA NOT Ltd.* (Dec. 2008).

[32] Boev, A., Hollosi, D., and Gotchev, A. Classification of stereoscopic artefacts. Tech. rep., MOBILE3DTV, July 2008.

[33] Bronstein, G., Garon, H., Ricci, J., Talarico, D., Taub, G., and Scherer, W. Identifying growth opportunities using information analytics: A case study in five industries. In *Systems and Information Engineering Design Symposium, 2007. SIEDS 2007. IEEE* (2007), pp. 1–6.

[34] Cooper, W., and Lovelace, G. IPTV guide. In *IPTV Conference 2007 - Deployment and Service Delivery, IET* (2007), p. 1–66.

[35] El-Gendy, M., Bose, A., and Shin, K. Evolution of the internet QoS and support for soft real-time applications. *Proceedings of the IEEE 91*, 7 (2003), 1086–1104.

[36] Hulsen, P., Vlemmix, E., and Baker, K. Peer-to-peer internet television. In *Automated Production of Cross Media Content for Multi-Channel Distribution, 2005. AXMEDIS 2005. First International Conference on* (2005), p. 4 pp.

[37] Ireland, G., and Ward, L. *Transcoding Internet and Mobile Video: Solutions for the Long Tail.* IDC, Sept. 2007.

[38] ITU-T. J.247 : Objective perceptual multimedia video quality measurement in the presence of a full reference, Aug. 2008.

[39] Ixia. IxLoad overview: Customer presentation application and Multiplay-Delivery testing.

[40] Jumisko-Pyykkö, S., and Strohmeier, D. Report on research methodologies for the experiments. Tech. rep., MOBILE3DTV, 2008.

[41] Kagan Research. Media trends (12th ed.), Mar. 2006.

[42] Kalva, H. The h.264 video coding standard. *Multimedia, IEEE 13*, 4 (2006), 86–90.

[43] Kalva, H., and Lee, J. The VC-1 video coding standard. *Multimedia, IEEE 14*, 4 (2007), 88–91.

[44] Kishigami, J. The role of QoE on IPTV services. In *Proceedings of the Ninth IEEE International Symposium on Multimedia* (2007), IEEE Computer Society, pp. 11–13.

[45] Level 3. Joint level 3/Move networks DNC case study - technical, 2008.

[46] Move Networks. Move adaptive stream.

[47] Move Networks. It's not web video, it's television: The power of internet TV, Sept. 2008.

[48] Move Networks. Move internet television services: Move networks provides a new standard for broadcasters, Sept. 2008.

[49] Nguyen, T. H. Assessment of video, May 2005.

[50] On2 Technologies. WHITE PAPER: on2 VP6 for flash 8 video, Sept. 2005.

[51] On2 Technologies. WHITE PAPER: TrueMotion VP7 video codec, 2005.

[52] On2 Technologies. On2 video: Advanced video compression solutions, from web to wireless and handset to home cinema, 2009.

[53] Pfanner, E. European newspapers find creative ways to thrive in the internet age. *The New York Times* (Mar. 2009).

[54] Pinson, M. H., and Wolf, S. Comparing subjective video quality testing methodologies, 2003.

[55] Raake, A., Garcia, M., Moller, S., Berger, J., Kling, F., List, P., Johann, J., and Heidemann, C. T-V-model: parameter-based prediction of IPTV quality. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on* (2008), pp. 1149–1152.

[56] Richardson, I. Vcodex white paper: An overview of h.264 advanced video coding, 2007.

[57] Shenick Network Systems. Testing MPEG based IP video QoE/QoS.

[58] Srinivasan, S., Hsu, P. J., Holcomb, T., Mukerjee, K., Regunathan, S. L., Lin, B., Liang, J., Lee, M., and Ribas-Corbera, J. Windows media video 9: overview and applications. *Signal Processing: Image Communication 19*, 9 (Oct. 2004), 851–875.

[59] Streaming Media Magazine. How to profit from internet TV: technologies and strategies for monetizing your video content.

[60] Symmetricom. IPTV video quality considerations.

[61] Symmetricom. V-Factor introduction webcast.

[62] Symmetricom. Delivering quality of experience to IPTV and TVoIP customers, 2007.

[63] Symmetricom. 2007 IPTV video quality study, 2008.

[64] Tanenbaum, A. S. *Computer Networks*, 4 ed. Prentice Hall PTR, Aug. 2002.

[65] VQEG. FRTV phase II.

[66] Wang, Z., Sheikh, H. R., and Bovik, A. C. *Objective Video Quality Assessment.* Sept. 2003.

[67] Winkler, S. *Digital Video Quality: Vision Models and Metrics*, 1 ed. Wiley, Mar. 2005.

[68] Winkler, S. Quality of experience (QoE): an important measure of success for IP-based video services, 2007.

[69] Winkler, S. Symmetricom white paper: Video quality and beyond, Oct. 2007.

[70] Winkler, S., and Mohandas, P. The evolution of video quality measurement: From PSNR to hybrid metrics. *Broadcasting, IEEE Transactions on 54*, 3 (2008), 660–668.