Dialogic
Making Innovation Thrive™

White Paper

Quality of Experience for Mobile Video Users

Quality of Experience for Mobile Video Users

White Paper

Executive Summary

Throughout the world, more and more people are choosing a mobile device as their platform of choice for watching video content. Some people use their mobile phones to catch the latest sports highlights, while others watch television shows on the train during their commutes. Some viewers use the WiFi in the local coffee shop to connect to Hulu on their netbooks, and others watch full-length movies on their iPods.

Regardless of the type of device, content viewed, or network used for access, each person has some basic expectations about the viewing experience. The level of expectation varies among viewers, depending on factors such as the amount of money spent for the service, type of content being requested, and type of device being used for viewing. It is the job of service providers to match the level of quality to their users' expectations.

This white paper introduces the topic of QoE (Quality of Experience) as it relates to mobile video. It discusses how video quality is typically measured, provides information on some of the tools and techniques in use today for measuring video quality, and suggests some future directions for measuring video quality. While the term QoE can be applied to a wide range of multimedia communications services, this discussion focuses on the world of video as it is displayed and viewed on a mobile device.

Quality of Experience for Mobile Video Users

White Paper

Table of Contents

Introduction

In the past, telecommunications carriers used several methods for measuring the Quality of Service (QoS) for the audio and video data provided to their consumers. Most of these methods relied on determining the quality of the underlying transport network. For example, carriers used service assurance systems and performance management applications to identify faults or areas of concern within a transport network. The common belief was that by identifying and isolating issues within the network, carriers could provide a high level of quality to users.

In recent years, though, this type of thinking has evolved. The concept of QoS continues to be important in the service provider environment, but a new concept called Quality of Experience (QoE) is rapidly gaining mindshare. The QoE concept differs from QoS in that it considers much more than the performance of the network — QoE is concerned with the overall experience the consumer has when accessing and using provided services.

This white paper examines QoE (Quality of Experience) as it relates to mobile video, discusses how video quality is typically measured, and describes Dialogic's perspective on QoE for mobile video.

Quantifying Quality

To most people, the concept of quality is a subjective term. When asked to describe how to measure quality, a person might say, "I know it when I see it." But in the telecommunications and multimedia world, quality has been measured objectively for decades. The commonly-used metric used to quantify the quality of a particular piece of multimedia content is the Mean Opinion Score (MOS).

The MOS is a numeric value between 1 and 5, with 5 representing the highest quality. It is used as a method for comparing different pieces (clips) of multimedia content. Traditionally, the MOS was used primarily for audio content, and it is still commonly used as a method for benchmarking speech quality in a communications channel. More recently, the same method and scoring system have been applied to video content.

When the concept of MOS was introduced, it was calculated in a manual, time-consuming way: groups of people were recruited and invited to listen to a set of different recordings. After listening to an audio clip, each person gave a numeric score to the quality of that clip. The average of the scores across the group became the MOS for that piece of content. Typically, a score greater than 4 was considered very good, while a score less than 3 was considered poor.

Today, the method of generating MOS values is automated. Vendors have created algorithms that can estimate a MOS based on the characteristics of the media stream and the network over which it is played. Instead of recruiting people to sit in a room and listen to or watch clips, interested parties can use measurement tools to generate MOSs automatically. The results are usually very reliable, and they correlate well with the results that would have been generated by humans.

MOS is, by definition, a **relative** measurement that is properly used as a **comparative** measure of quality. It is correct to say, "In this test, my new product has a MOS of 4.17, while the competition has a MOS of 4.08." However, it is technically incorrect to say, "My new product has a MOS of 4.1." Because the group of people generating the scores may use different grading criteria, it is important to consider the scores as a relative measure of quality.

Despite the fact that the MOS is a relative measurement, many companies that create products used to measure video quality refer to the MOS as an absolute value. Many vendors evaluate the quality of a specific piece of video content by using these products or by using their own algorithms that return a single numerical value for the quality measurement. These algorithms, however, are better used as a tool for comparing multiple video clips to determine a relative level of quality, rather than to produce an absolute result.

Sources of Video Quality Issues

Many quality issues experienced by consumers of mobile video are caused by issues within the transport network. Mobile telephony networks, especially pre-3G and older mobile telephony networks, are not well-suited for video transmission. The bandwidth in these networks is severely restricted, and the protocols used in the networks were not designed with video in mind.

There are many other sources of quality issues with mobile video, ranging from issues introduced when the video is created through issues introduced when displaying the video content.

Video quality issues are typically introduced at the following times:

- During video creation
- During transcoding
- During video transmission
- When displaying video on a device

Video Creation

Video content can be generated from a variety of sources. With professionally-produced content created by a television network or a movie studio, the cameras used to capture the video are expensive devices used by trained technicians. However, much of the video that is viewed today is user-generated content that can be created by anyone who has access to a recording device.

People who spend time on YouTube exploring the world of user-generated video content can attest that its quality varies widely. Most user-generated content is created on recording devices that do not produce a high-quality image. Low pixel values and dirty lenses can be blamed for many video issues that occur when content is recorded. Other quality issues can be traced to the environment in which the content was recorded or some form of impairment of the person doing the recording. For example, one frequent problem is that of people with shaky hands. Even with the powerful anti-shake algorithms embedded in many of today's recording devices, much of today's user-generated video has wildly shaky motions (and, subsequently, blurry images).

If a video is recorded on a poor quality device or recorded with shaky hands, the initial capture quality of the video will be poor. While there are some techniques for eliminating quality issues caused by poor capture, these quality issues will likely remain with the video over its lifetime.

Video Transcoding

When a recording device captures a video stream, the video is typically encoded (compressed) so that it can be transmitted efficiently through a network. This is the case for user-generated video, including video uploaded from a mobile recording device to a service such as YouTube. While today's compression algorithms are fairly effective at preserving video quality, there is some inherent loss of quality each time a video is compressed.

The best-performing compression algorithms (that is, those that retain the highest quality) typically require a lot of processing power. Since processing power is a scarce resource on mobile devices, the video that is typically captured and uploaded from a mobile device may be encoded with an older algorithm that does not do a very good job of preserving video quality.

Once a video is encoded, it must be decoded before it can be played. There are many algorithms to perform encoding and decoding, and each has a unique set of characteristics. As a rule, content encoded with a specific algorithm can only be decoded with that same algorithm. Many of today's playback devices contain multiple decoding algorithms, allowing them to play (decode) video content that was created with a variety of encoding algorithms.

Even though most playback devices support a wide variety of compression algorithms, it is still often necessary to perform a translation from one coding algorithm to another at some point between video capture and video display. This translation is called "transcoding," and it typically consists of decoding the video and then re-encoding it into the format required by the player. The decode/re-encode transaction causes a loss of some video fidelity, thereby reducing video quality.

In today's telecommunications networks, it is common for video content to be transcoded several times between the original entry to the network and when it ultimately arrives at a display device for viewing. Carriers expend a lot of effort attempting to retain the highest level of image quality during these multiple transcodes.

Video Transmission

Any time information is sent through a network, some of that information can be lost. In today's IP networks, it is common for IP packets to be lost, or at least delayed so long that they are effectively lost. Lost packets can cause any type of multimedia content to lose quality. In addition to packet loss, network characteristics such as jitter can also contribute to quality issues in multimedia content. Jitter affects quality by interfering with the standard algorithms used to buffer content played through a bandwidth-constrained network.

Network impairments such as packet loss and jitter have existed in voice networks for many years, but the effect of these issues in voice networks is frequently not serious. The human ear can recover from gaps in voice communications, using context to determine the missing portion of the conversation when the gap is small enough. Video communication is much more sensitive to packet loss than audio communication, because the human eye can often detect small glitches in a video stream caused by relatively minor packet loss, to an extent whereby enjoyment and/or understanding are more severely affected.

Catastrophic network issues such as cable cuts or network outages also contribute to video quality issues. These types of issues, however, are typically easy to recognize and address.

Displaying Video on a Device

Even the best video content can be unsatisfying if it is displayed on a poor viewing device. Mobile video is typically displayed on small, fragile screens that can be damaged from being carried. In a screen with "dead pixels," the video quality is often judged to be inferior to the same content displayed on a higher-quality device. Also, many mobile devices have less-than-optimal video capabilities, because they were engineered to provide high-quality voice service, and video capabilities were a secondary consideration. High-quality video displays require a lot of power, so vendors may consider tradeoffs like offering a longer battery life by sacrificing display quality in the mobile device.

Common Video Quality Issues

Mobile video consumers commonly encounter the following types of video quality issues:

- Video blockiness
- Video blurriness
- Video freezing
- Video jerkiness
- Video blackout
- Audio sync issues

This list of issues is not exhaustive, but it does represent a sample of the types of issues that can affect mobile video quality. While many of the problems that affect today's viewers can be caused by bandwidth limitations, there are other causes as well, such as signal loss, poor capture quality, or poor device quality. It is important to realize that quality issues can be caused by a variety of impairments — some in the network and some in the device. Determining a root cause for a video quality issues requires a high level of skill and experience in dealing with these issues.

Video Blockiness

One of the most common quality issues in the world of mobile video is "blockiness," an impairment in which the image contains "artifacts" that resemble small blocks of a single color. To the viewer, this condition makes it appear like one large pixel is being displayed over a large area. Figure 1 shows an example of blockiness:

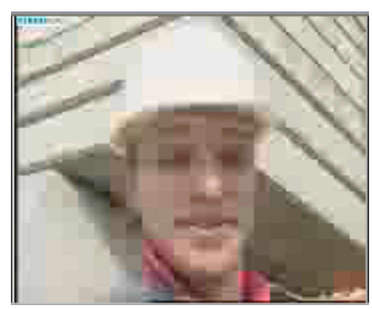

Figure 1. Example of Video Blockiness

Blockiness usually occurs when the video encoder software is not able to process the entire video data stream. As a result, the encoder cannot accurately represent the details of each image, and instead uses a single "mean value" to represent all pixels in a large area. When blockiness is severe, it can result in extreme user dissatisfaction; therefore, it is one of the more important video quality issues to identify.

Video Blurriness

Video blurriness typically occurs in two different areas: When capturing (recording) the video or when performing the encoding of the video stream. In user-generated video, a blurry image is most often caused by an issue during the initial recording of the video. Typically, the person doing the recording did not set the camera properly and as a result, the content recorded poorly.

In some cases, blurriness is introduced because of inadequacies in the compression algorithms used for encoding the video. This can be especially true in mobile video, in which blurriness is commonly seen when the video content contains an extreme amount of action and/or scene changes. At these times, the amount of bandwidth available in the network is not sufficient to handle the high demand required to encode the action properly.

Both blockiness and blurriness introduced by the video encoder have the same root cause — not having enough bits available to encode all details of an image. Advanced video coding algorithms that enable the encoder to use bits more efficiently can lessen the effect of video blurriness.

Figure 2 shows an example of blurriness:

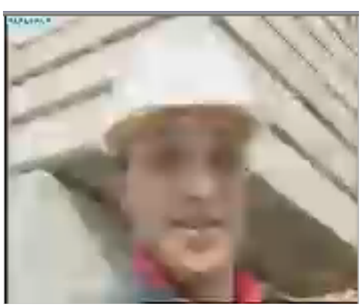

Figure 2. Example of Video Blurriness

Video Freezing

Sometimes a video can simply "freeze" while it is being watched. This is typically caused by insufficient bandwidth. If this happens while watching a video on a computer, the display often shows some indication that the video is "buffering" before the streaming will resume. In a mobile environment, similar conditions can also be caused by significant packet loss, which can cause some video frames to be delayed or lost in transit through the network. If the lost frames are particularly important frames (known as "intra-encoder" or "I-frames"), the video will freeze until the next I frame is received.

Video Jerkiness

Video content can sometimes appear jerky to the viewer. The video stream might not flow smoothly, and the action might appear to start and stop rapidly. While this is often a bandwidth issue, similar to that experienced with video freezing, jerkiness can have other causes as well.

For example, network jitter can cause a stream to appear jerky. In this case, the bandwidth may be sufficient to handle the video content, but some of the buffering algorithms within the network might not be operating properly, leading to overflow or underflow conditions. While the buffers attempt to regain stability, the video stream might be temporarily interrupted, leading to a jerky image.

Jerkiness can also be caused when the frame rate of a video stream is reduced. Professionally-produced video is typically recorded and played at 30 frames per second, which can be thought of as 30 still images that flash by so quickly that the eye interprets the result as motion. However, video sent to a mobile device is typically played at a lower frame rate — usually 10-15 frames per second. This lower frame rate is often sufficient for the eye to perceive the resulting video as smooth motion in the small display size of a mobile device. However, when the frame rate gets too slow, especially when the video contains fast motion or a high amount of detail, the eye perceives the resulting video as jerky. For a typical display on a mobile phone, this jerkiness can be experienced when frame rates fall below about 5 frames per second.

Frame rate reduction is sometimes a natural response by the compression algorithms to reductions in the bandwidth or other network conditions. Rather than keep a high frame rate (and buffer more often to wait for bits to be received), the algorithms might choose to keep the video streaming continually at a lower frame rate.

Video Blackout

A video blackout is also known as a loss of signal, because when it occurs, the video content disappears from the screen and the page appears black. Although a video blackout can occur due to a catastrophic network failure, such as a complete loss of the transmission signal or a "dropped call" failure, it is more often caused by severe bandwidth issues.

In mobile video, a video blackout can occur when the video signal is being handed off between adjacent coverage areas, especially when those coverage areas contain network equipment of different vintages. For example, one geographic area might have network equipment that can support certain video services, while an adjacent coverage area might not.

Audio Sync Issues

Although not purely a video issue, a loss of synchronization between the audio and video streams (most commonly noted by the viewer as an issue with lip sync) is one of the most annoying issues for viewers of mobile video.

Some audio sync issues can be attributed to bandwidth issues. The audio track and the video track are typically sent separately to the device, and if bandwidth issues slow down the video track, the synchronization can be lost.

Audio sync issues can also occur because of poor encoding at the time of recording. If the microphone or recording device is not working properly, the audio stream and the video stream might not be aligned properly at the time of content creation.

Measuring Mobile Video Quality

As video becomes more common as a method of communication and entertainment, vendors are creating products that can measure and report on video quality. This section describes some of the characteristics of these measurement techniques and explains the advantages and disadvantages of each technique.

Full-Reference versus Reference-Free Measurement Techniques

One way of categorizing video quality measurement techniques is as either full-reference or reference-free. A full-reference technique compares the measured video against the original, uncompressed video, and checks for differences. A reference-free technique analyzes the video under test and looks for impairments without making a comparison. Both of these approaches have positive attributes and drawbacks, and the appropriate choice depends on the desired application, cost constraints, and issues that can be caused by false readings.

Full-Reference Measurement Technique

An advantage of a full-reference technique is that it typically provides very powerful results for any changes in the quality of the video. Even subtle changes in the quality of the video image can be identified and treated as a potential issue area.

A disadvantage of a full-reference technique is that it requires a lot of processing power to perform the calculations necessary for the comparison. Full-reference techniques do not typically work in real-time. Analyzing a one-minute video stream can take several minutes, because each frame is compared pixel-by-pixel against the original version. Full-reference techniques also tend to be expensive, because they require powerful processors and very sophisticated algorithms to measure the quality changes.

Full-reference techniques are typically used in applications where the quality requirements are very high, price is not the primary consideration, and real-time streaming is not required.

Reference-Free Measurement Technique

Reference-free techniques attempt to analyze quality issues by looking only at the video stream being tested. Instead of making comparisons, reference-free techniques try to identify or predict impairments such as blockiness or jerkiness by analyzing the characteristics of the underlying network transport and/or by analyzing the video itself.

One shortcoming of a reference-free technique is that it can generate false positives. In other words, it might identify an issue when the “issue” might actually be something that is inherent within the video stream. For example, sometimes video content can intentionally contain blocky images. In this situation, a reference-free technique might interpret that as a quality impairment and report blockiness, while a full-reference technique would realize that this is part of the desired content and would not report a problem.

A reference-free technique has the following advantages:

- It can be used in real-time.
- The processing requirements are not as high as for a full-reference technique.

Reference-free techniques are typically used in applications where cost is a consideration and some false positives can be allowed.

Pre-Service versus In-Service Testing

Today's video quality measurement techniques include products specializing in pre-service (lab-based) testing and products specializing in the real-time monitoring of "live" video services (in-service testing).

Pre-Service Testing

As a general rule, the full-reference techniques described above are used for pre-service testing. For example, consider a situation in which a carrier wants to purchase new network equipment to support video services. As part of the purchase decision, the carrier may want to evaluate the quality of the video that the network equipment will transmit. To perform this evaluation, the carrier might:

1. Install some of the new equipment into a laboratory environment.
2. Have the equipment process a video stream, simulating performance in a live network.
3. Direct the video output to the equipment performing the measurement for analysis.
4. Compare the resulting video content to the original video content to see if the network equipment adversely affected the video quality.

Another application for the pre-service measurement technique is comparing the video processing capabilities of equipment from several vendors. For example, a carrier may want to conduct a controlled experiment where the same video content is processed by network equipment from several vendors. To perform this experiment, the carrier might:

1 Install some of the new equipment into a laboratory environment.

2 Have each piece of equipment process the same video stream.

3 Direct each video output to the measurement technique for analysis.

4 Evaluate the statistics from the measurement technique to see which piece of network equipment retained the highest video quality.

In-Service Testing

While the above examples are common scenarios for pre-service and full-reference measurement techniques, other vendors provide different techniques that are focused on in-service measurement and monitoring. These are often reference-free techniques, because in-service monitoring requires real-time capabilities. The objective of reference-free techniques is often to identify impairments as they happen, or at least to gather and retain measurement data. In this way, the data is available at a later time for analysis against reported quality issues.

Many in-service testing techniques include interfaces to higher-level analysis systems. For example, a video quality monitoring technique might be focused on analyzing and evaluating the performance of the sub-network that handles video traffic. If an issue is identified within that network, the measurement system might feed that information to a network management system that can correlate that sub-network issue into an overall diagnosis to isolate a potentially larger issue within the carrier network.

Characteristics of Video Quality Measurements

One of the most important trends in video quality measurement over the last few years has been the increasing importance of QoE as compared to QoS. Traditionally, carriers were concerned with the quality of the service offered, which was usually determined by monitoring the transport network and ensuring that it was working properly. Carriers spent millions of dollars to buy network operations systems that monitored the network. If the network was operating properly, these systems would typically report that all was well.

More recently, carriers and vendors have been expanding the scope of measurement and monitoring to consider the quality of the overall user experience. By the broadest definition, QoE can refer to any portion of the user experience at the time of using a mobile video service. Was the call completion time acceptable? Was the user interface used to make the call acceptable? Was the price paid for the video content acceptable? This is a broad interpretation of QoE, and for the purpose of this paper, the QoE measurement refers only to the quality of the video image itself as it is perceived by the user.

Two important differences between QoS and QoE are where the quality measurements are taken and the role of user perception.

Where the Quality Measurements Are Taken

In a QoS technique, measurements are usually taken at the network. In a QoE technique, on the other hand, measurements are best taken at the device used to view the video content, as that is the area most closely related to the user experience.

Role of User Perception

Although a QoS technique might try to estimate the quality of the video that the viewer receives, its estimate is based on measurements gained from the network, and these measurements may not always be relevant to the perceived video quality. The QoS approach has some notable drawbacks, especially for video applications, because video is more sensitive than other multimedia applications to the effects of minor, but strategically-placed, network impairments.

The most common method of encoding video is to start by encoding a complete representation of the initial image of a video stream. This complete initial image is known as an I-frame. After that, the next frames of the video stream may just contain encodings of the differences that have occurred since the start. If the video content does not have a lot of movement, the amount of information contained in these subsequent frames is relatively low. At some interval (perhaps once a second), another refresh frame (I-frame) is encoded and transmitted.

Some frames are more important than others in maintaining a smooth and clear picture. If network packet loss causes the loss of a frame that does not contain much information, the viewer probably would not notice the effect. However, if packet loss affects one of the I-frames, the effect to the viewer could be very serious.

In the video world, not all packets are created equal. If a network-centric measurement system reports some packet loss, there is no way to determine whether the viewer was affected by this issue or not. A QoS system would report the packet loss as a quality issue, whether that network issue affected the image presented to the viewer or not. A true QoE system would be able to determine whether the packet loss was inconsequential to the user, caused a slight impairment of quality, or created a major issue.

Efforts at Standardization

As any technology matures, industries that support that technology define standards. This helps to ensure that the industry uses a common nomenclature and common system architectures for defining features and applications.

Video QoE is now recognized as a field where the technology has matured to the point where standards efforts are underway. Within the International Telecommunication Union (ITU), industry experts have defined several standards related to video quality, including:

- ITU-T Rec. G.1070 (04/2007) Opinion model for videophone applications.
- ITU-T Rec. J.247 (08/2008) Objective perceptual multimedia video quality measurement in the presence of a full reference.
- ITU-T Rec. J.246 (08/2008) Perceptual audiovisual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference.

Complementing the effort in the ITU, an industry group known as the Video Quality Experts Group (VQEG) has been working since 1997 to define international standards for measuring and improving video quality.

Less formally, other industry and academic groups are holding conferences and presenting research related to perceptual video quality. For example, the QoMEX (Quality of Multimedia Experience) conference was held for the first time during July, 2009, and organizers are planning to make it an annual event.

These activities are just an indication of the types of work currently focused on defining common standards in the mobile video space. The growth of these efforts signals an increasingly important role for QoE.

Dialogic's Perspective on QoE for Mobile Video

The increasing focus on QoE for video, and especially for mobile video, is an important step forward for the industry. Dialogic focuses on providing high-quality, high-performance video features for its customers, so the increased industry focus on perceptual video quality aligns well with Dialogic's strategies and corporate goals.

Dialogic's video research and development lab, Dialogic® Media Labs, is currently creating techniques for measuring video QoE. Dialogic is taking a different approach from many other vendors in this space, who claim QoE expertise while still using network performance measurements to estimate the viewing experience. Dialogic is exploring QoE techniques that measure the quality of the video image in the pixel domain. Rather than performing network measurements and estimating the effect on the video stream, Dialogic accesses the video stream directly and performs measurements on the pixels to get an unambiguous view of the quality of the video being displayed. Dialogic believes that pixel-level measurement is the only approach that can truly be called QoE, because it is the only approach that examines the video stream itself to determine the viewer's experience.

Dialogic believes that Video is the New Voice™, and it continues to provide its customers with the latest innovations in video technology to foster their continued success in the expanding role of mobile video. To learn more about Dialogic's plans for mobile video and the current developments related to QoE, please contact your sales representative or indicate your request on the Dialogic website at www.dialogic.com.

Acronyms

IP	Internet Protocol
ITU	International Telecommunication Union
MOS	Mean Opinion Score
QoE	Quality of Experience
QoMEX	Quality of Multimedia Experience
QoS	Quality of Service
VQEG	Video Quality Experts Group

Making Innovation Thrive™

www.dialogic.com

Dialogic Corporation
9800 Cavendish Blvd., 5th floor
Montreal, Quebec
CANADA H4M 2V9

12/09 11681-01