



Tools for Video Analysis
1566 La Pradera Dr
Campbell, CA 95008
www.videoclarity.com
408-379-6952

White Paper – Why Monitor?

*Bill Reckwerdt, CTO
Video Clarity, Inc.*

Measuring the audio and video quality has taken on great importance as we communicate and interact with video. The goal of quality assessment (QA) is to assess the quality in agreement with human observers. Considerable research continues to provide solutions to this problem.

New video services - IPTV, video on demand (VOD), peer-to-peer (P2P) video streaming – are continuously added to provide customers with more options. What happens if one of these services has poor video quality? Simply identifying that the data was received with errors is not sufficient. We need to define the affect on the video quality, which depends on categorizing the severity placement of the error.

Moreover, the customer knows that he is a video quality expert. If the quality degrades, he will immediately call customer support. A service provider who is not sensitive to this will lose revenue and eventually the customer to their competition. Thus, it is imperative to provide comprehensive video monitoring and analysis capabilities.

This paper will discuss a methodology for evaluating the video quality by comparing 2 points within the network using a Full Reference technique.

Terms Defined

CBR – Constant Bit-Rate Encoding/Compression.

DMOS – after comparing the A/V stream to another A/V stream, the video is judge using the Mean Opinion Scale. On this scale, 0 would be perfect.

Full Reference – Evaluating the video quality when you can compare the original to the processed. This method measures the quality difference as opposed to guessing at what the quality should be.

Grooming – Picking programs from multiple MPTS and forming a new MPTS

GS – Guaranteed Service – This refers to a network which allocates space for the streams data rate

Headroom – Encoders are allowed to allocate more bits than the average data rate if the scene is difficult to compress. Headroom refers to pre-allocated extra space just in case this happens

Lossless – the compression algorithm faithfully restores the original audio and video

Lossy – the compression algorithm does not attempt to faithfully restore the original audio and video

MPEG – Moving Pictures Experts Group – the informal name of ISO/IEC JTC1/SC29 WG11 responsible for standardizing MPEGX

MPTS – Multiple programs (streams) are combined together so that they can be sent as one combined stream when a certain data rate has been pre-allocated

PSNR – a metric that compares a set of reference values to a set of processed values usually incorporating the mean-square-error.

Quality of Experience (QoE) – Measuring the quality with respect to what the end user sees and hears.

STB – Set-top Box

VBR – Variable Bit-Rate Encoding/Compression.

VCEG – Visual (Video) Coding Experts Group – the informal name of the Visual Coding Working Party 3 of Study Group 16 of the ITU-T responsible for standardizing H.26X, JPEGX, and JBIG-X.

Video Quality – This will refer to both the image (picture) and the audio (speech).

Background

Most video is digitized and compressed at a data-rate too allow it to be transmitted over existing transmission paths: Satellite, microwave, fiber, and Internet. The only exception is when analog video is still transmitted, but that is going away.

The digital video formats defined by VCEQ and MPEG are the de-facto standard for entertainment video. They are popular because

- There are no restrictions on the implementation of the video encoder (compression device).
- The video decoder's (Set-top box, PC) capabilities are fully-defined based on levels and profiles.
- The standards include video, audio, transport, and timing functions.

These video formats include - MPEG-1, MPEG-2 (DVD), H.263 (video surveillance), MPEG-4/H.264 (combined next generation standard), JPEG (still pictures), JPEG-2000 (archival) – just to name a few. With the possible exception of JPEG-2000, all of them are lossy (information is lost during compression so the quality after encoding/decoding is not as good as the original). JPEG-2000 can be lossy or mathematically loss-less.

In practice all lossy encoders generate artifacts (areas of unfaithful visual/audible reproduction). If the encoder is designed well and the data rate is high enough, then these artifacts will be virtually invisible. The quality of the encoder and picking the appropriate settings can be checked offline (in non real-time) using a quantitative video quality analysis device like ClearView (www.videoclarity.com/products.html).

If a good encoder has been chosen and the settings are properly configured, then real-time errors can occur due to:

- Real-time Compression
- Ad Insertion
- Statistical Multiplexing
- Re-encoding
- Transmission Systems

Real-Time Compression

Real-time compression is needed for live transmissions (or retransmissions). The compression device (Encoder) runs making the best quality A/V streams possible. Two ways exist for encoding:

- 1) Constant Bit Rate (CBR)
- 2) Variable Bit Rate (VBR)

Video encoders which are based on inter and intra-frame compression (those mentioned above with the exception of JPEG) decrease bit rate by reducing redundant information within a frame and between one frame and the next. For slow moving scenes (highly redundant scenes), they do quite well. For high motion, then it is more difficult. Video, by its very nature, is dynamic. Imagine a scene where a young couple is walking through a park. This scene has very little motion. If all of a sudden, the scene changes to include a car come toward them at high speed, then this has considerable motion.

VBR produces better video quality as it can change the compression rate dependent on the scene complexity. Of course, more bits require more bandwidth to stream. Most of the time, the streaming bandwidth is fixed over the network so VBR cannot be used.

Most people implement CBR for fixed bandwidth applications – Internet delivery, CableTV, Satellite TV, and IPTV. CBR is segmented into pieces, where the bit-rate over time is constant, but the instantaneous bit-rate is higher or lower depending on scene complexity. Buffers to smooth out variations in complexity are used to reduce the effects caused by complex scenes. This is known as allocating headroom and it must be pre-allocated sufficiently for the type of material.

For real-time compression, it is very important to allocate headroom. When the headroom is not sufficient errors will occur.

Non real time encoding is termed file-based encoding. The compressionist can take time to encode/re-encode the material based on their expertise. Headroom problems are eliminated by the compressionist's skill. The subsequent digital stream (file) is played out of a video server or written to a DVD.

Ad Insertion

Ad insertion is the process of inserting an advertising message into a stream. The ads can be inserted nationally, geographically, or demographically. Normally a digital tone (known as a cue tone) is generated which tells an ad

server to play the ad; instead of the normal programming. Another tone is generated to return to the normal programming.

Problems can occur during the switch if:

- 1) the resolution or aspect ratio between the programming and advertising is different,
- 2) the advertising starts or stops early or late,
- 3) the advertising causes the real-time encoder to need more headroom.

Statistical Multiplexing

A broadcaster purchases a fixed amount of bandwidth. To maximize the use, they pack as many channels as possible into this bandwidth. The normal technique for doing this is called Statistical Multiplexing. Statistical Multiplexing is a technique for combining a number of uncorrelated, bursty traffic sources together so that the sum of their peak rates does not exceed the link capacity.

A series of encoders are arranged so that their output can be combined by the multiplexer (combiner) into a single multi-program transport stream (MPTS). Each encoder is told its target bit-rate and the multiplexer monitors the sum of the traffic. When an encoder encounters a complex scene, it requests more bits. The multiplexer steals bits from the other encoders and allocates more to the requesting encoder. If many of the encoders encounter a challenging scene concurrently, then problems will occur. The multiplexer will either deny the encoders request or discard data (drop frames). Either way, the video quality is affected.

Statistical multiplexing is important when delivering video over a fixed pipe – as in satellite, microwave, and fiber transmission. The subscribed data rate is guaranteed and the user would like to use as much of the entire bandwidth for which they subscribed/paid.

Some multiplexers (pioneered by Divicom (now Harmonic – www.harmonicinc.com) use a look ahead statistical multiplexing technique. The encoding is done in 2 phases. The first phase calculates the bit-rate and passes this information to the multiplexer ahead of time so that it can change the bit-rate before the oversubscription happens.

Re-encoding

Another approach which is similar to Statistical Multiplexing is known as Re-Encoding. This is not a full decode and encode. If a full decode is done, then it is better to use a Statistical Multiplexer.

Re-encoding modifies an existing compressed digital stream in real-time without decoding. When a re-broadcaster is pulling programming from multiple sources, combining them, and sending them over their fiber, satellite, or microwave channel, they may choose to re-encode. A re-broadcaster would be a cableTV, satelliteTV, or IPTV operator.

Re-encoding parses the compressed syntax and removes some of the encoding details to fit multiple programs into a new MPTS. This is normally done in conjunction with a system multiplexer and when multiple MPTS are groomed (a new MPTS is formed by pulling programs out of multiple MPTS).

Once again, complex scenes can cause a situation where oversubscription will happen. In this case, the video quality will be affected.

Transmission systems

The video is transmitted over a guaranteed service (GS - microwave, satellite, or IP) or a controlled load service (IP). A controlled load service is a best-effort service. Due to the explosive growth of video on next generation IP networks, this method requires considerable data shaping.

Even in guaranteed service networks, bit errors do occur. The streams are sent over many routers and any one of them can delay the packets (causing jitter), reroute the packets (causing loss or reordering), or simply fail.

In a best-effort service, bit errors will occur.

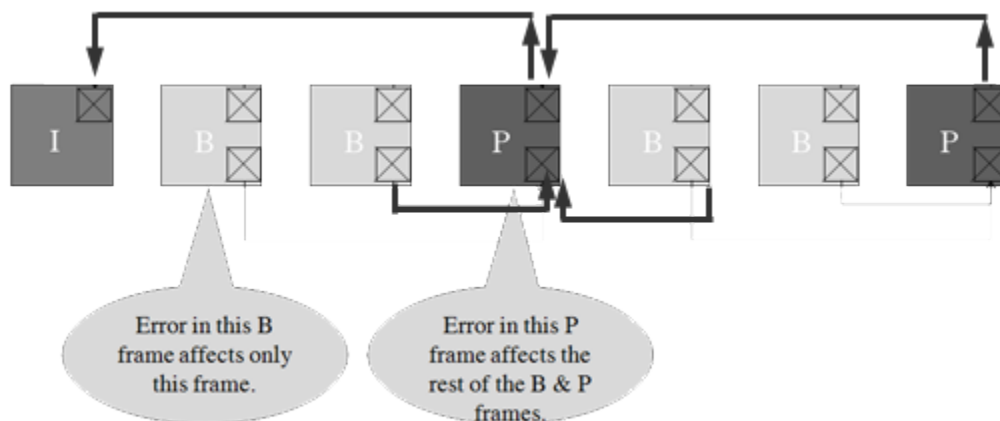
Monitoring for Real-time Errors

The simple truism is that errors will occur. What is the affect on the video quality?

This depends on the type of compression. In general for block based algorithms – MPEGx, H.26x, the frames are divided into 3 categories:

- Intra frames (I) – a fully specified picture
- Predicted frames (P) – holds the changes from the previous frame
- Bi Predictive frames (B) – holds the differences between the proceeding and following frames

If an I-frame is lost or corrupted, then this affects the video quality until the entire picture is redrawn. If a P-frame is lost, then the affected area has reduced video quality until it is redrawn by a subsequent P-frame or I-Frame. If a B frame is lost, then the affect is minimal.



How do you know which type of frame was lost? I frames are the largest followed by P then B so some algorithm attempt to intelligently look at the size of the packets. Others do a deep packet analysis and read the stream syntax. Deep packet analysis, of course, takes the most time and broadcasters encrypt their services rendering deep packet analysis impossible.

Set-top boxes (STBs) are computerized devices, which receive compressed digital signals, decrypt/decode them, and convert them to either an analog or digital format to be shown on your TV. The STB can be either an external box, built into the TV, a PC, a gaming console, etc. Regardless, it makes it possible to receive and display TV signals, connect to networks, play games, and surf the Internet. One of its primary functions is to detect errors, and fix or conceal them. It does this by:

- 1) Holding previous frame/partial picture
- 2) Asking for a retransmission (Microsoft's IPTV solution)

Some STBs do an exceptional job of hiding errors. This is why the monitoring must be done after the STB.

Why bother?

The answer is simple – competition. A poor quality service will affect sales to future customers and it will reduce existing customer satisfaction.

Monitoring should return 3 basic parameters:

- Knowledge that something bad has occurred
- The affect on the end customers perception
- Placement of the Error – which points caused the error

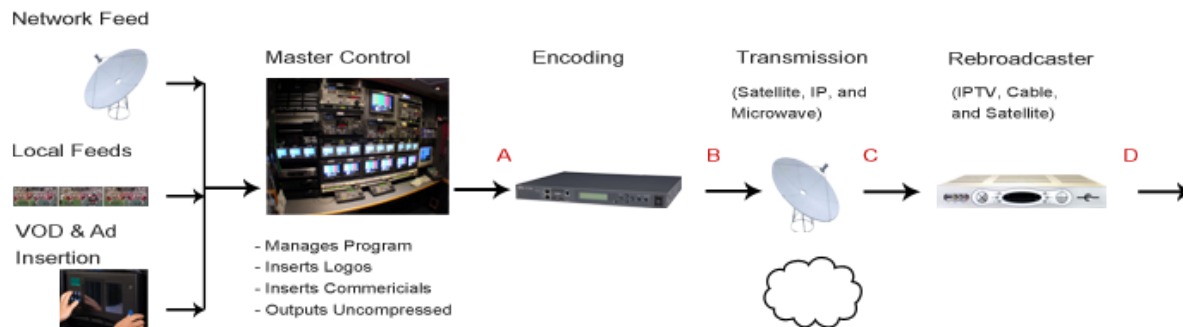
Armed with this knowledge the service provider can

- Fix the current Error
- Prevent future Errors

For these reasons, the best place to monitor is everywhere. Since this is impractical, the monitor should be placed after the

- Ad Insertion (Master Control)
- Real-time encoder
- Statistical multiplexer
- Re-encoder
- Transmitter

Figure 1: Master Control Workflow



Monitoring in the early phase can give a deeper understanding of the affect of errors. If the monitoring device saves the error states, then deeper analysis can occur to solve the error. In the end, a well devised monitoring systems, will cut costs, reduce customer churn, and provide a better long term solution.

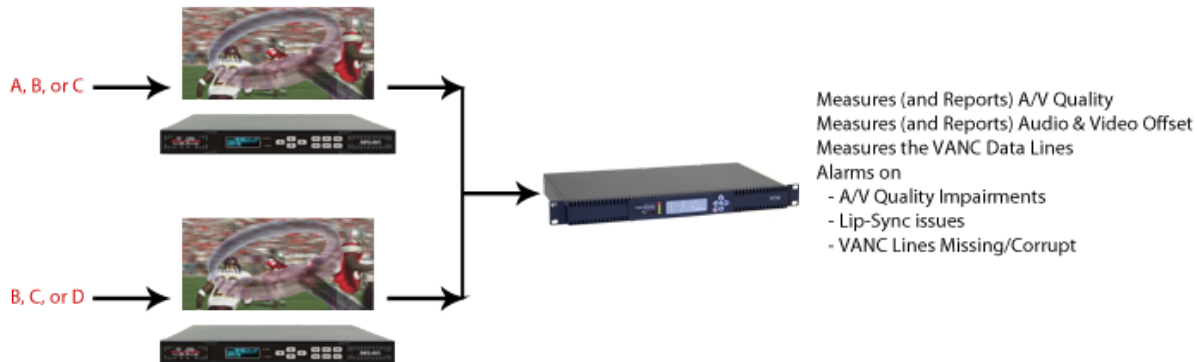
Video Clarity RTM Solution

In real life, we know that A/V errors occur. RTM compares transmission feeds after the STB; alerts when error causes visual, audio, and ancillary data glitches, reports if lip-sync occurs, and saves the results around errors. RTM provides information to choose a recovery plan.

To summarize RTM

- Reads & Aligns 2 live feeds
- Monitors the A/V Quality
- Monitors the VANC
- Calculates A/V Sync (Lip-sync)
- Saves streams around the error for off-line analysis

Figure 2: Hardware Video Quality Monitoring

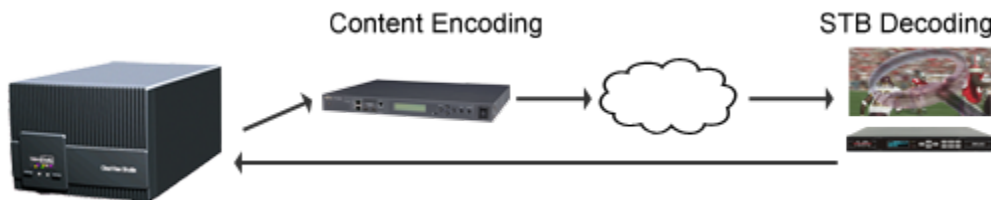


The alerts can be used in a variety of ways depending on the complexity of the broadcaster:

- switch from the A feed to the B (alternative) feed automatically
- trigger a retransmission
- the error can be saved for later analysis to prevent future issues

Video Clarity ClearView Solution

ClearView Video Analysis generates test signals, captures live inputs, and inputs compressed or uncompressed files. It calculates the DMOS, JND, and/or PSNR scores. It uses the Sarnoff/PQR algorithm ported to JND (using the VQEG database) and the MS-SSIM algorithm ported to DMOS (using the University of Texas' LIVE database). It also lets you play test patterns for subjective testing, view the “reference” and “processed” signals side-by-side or for your own evaluation.



The Sarnoff and MS-SSIM algorithms are further discussed on our website at www.videoclarity.com/WhitePapers.html.

The Author

Bill Reckwerdt has been involved in digital video since the early 90's from digital compression, video on demand, to streaming servers. He received his MS specializing in Behavioral Modeling and Design Automation from the University of Illinois Urbana-Champaign.

He is currently the VP of Marketing and the CTO for Video Clarity, which makes quantitative, repeatable video quality testing tools. For more information about Video Clarity, please visit their website at <http://www.videoclarity.com>.