VIDEO QUALITY IMPAIRMENTS 101 FOR MSO'S

Daniel Howard
VQLink, Inc.

Abstract

Key video quality impairments that impact video networks are described to help MSOs in monitoring and analysis of video quality over their networks. An in-depth discussion and taxonomy are given of video quality compression and network artifacts that are detectable by new, no-reference video quality technology that employs a hybrid of both bitstream and pixel processing and thus provides full video analysis of live and file-based MPEG2 and MPEG4 video content. The importance of measurement accuracy and minimal Type I and Type II errors for detection of these artifacts are developed and specific issues in transitioning from MPEG2 to MPEG4 are addressed with respect to these artifacts. Also discussed is how compression and network artifacts are perceived and detected differently in MPEG2 and MPEG4, and the specific video quality challenges for MSOs using transcoded video. Use cases for accurate measurement and classification of video impairments are given for network capacity planning, verification and maintenance of no material degradation (NMD) constraints, and stream bandwidth reductions for delivery of Internet video. Ultimately, MSOs can use newer video measurement and monitoring technology that provides accurate detection and classification of video quality impairments throughout their network to ensure that they affordably deliver the video quality required to remain competitive.

INTRODUCTION

The video quality and quantity wars are underway between cable operators and their telco and satellite competitors, much as the high speed data bandwidth wars began almost a decade ago, and continue to this day. Already the number of HD channels is a key market feature, and now video quality, especially of high definition content, has emerged as a market differentiator and competitive advantage. But just as MSOs discovered the key to effectively competing for high speed data subscribers was to offer the maximum amount of bandwidth they could affordably deliver, the key to winning the video quality wars will be for MSOs to be able to offer the maximum quality that they can affordably deliver. And just as in the bandwidth wars, where MSOs needed accurate tools to measure and adapt bandwidth delivered to subscribers, they will need an accurate tool to measure and monitor video quality in their networks, and this video quality measurement and monitoring tool can then also be used to adapt video quality to ensure network health as well as offer new features to subscribers over time.

Video quality measurement and monitoring technologies fall into several categories.

1) Network proxies for video quality monitoring and measurement: these technologies use network packet monitoring as a proxy for video quality measurement. The difficulties with this approach are as follows:

a) video quality impairments such as compression artifacts are not measured;

b) network artifacts in the video that do not have corresponding packet errors are missed, which happens when packet errors are re-encoded by an MSO or content provider, and packets are renumbered; and

c) packet errors that have little to no impact on actual video quality are reported, thereby 'crying wolf'.

2) Full reference video quality measurement: these technologies rely on access to a copy of the original video in order to compare pre and post processing versions, and use methods such as peak signal to noise ratio (PSNR) to evaluate the video quality. The difficulties with this method are as follows:

a) MSOs cannot measure the video quality of ingested video using full reference technologies since they do not have a copy of the original video from the content providers;

b) PSNR does not work when format changes in video occur because pixels are not in the same location; and

c) PSNR does not accurately reflect the human visual system (HVS) and thus may not correlate well with results from subjective testing.

3) Partial reference video quality measurement technologies: these are similar to full reference in that information derived from the original video is required, and thus they suffer from many of the same disadvantages.

4) No reference, full analysis video quality measurement technologies: these are ideally suited to deployment anywhere in the MSO's network since they do not require a copy of the original video. However, in the past they have suffered from inaccuracy, including the inability to detect even gross video quality impairments, and false alarms from structures in the video that are similar to codec induced blockiness or blurriness.

Fortunately, new no-reference technology based on based on the human visual system (HVS) and employing a hybrid bitstream-pixel processing approach is now available that can accurately detect a variety of video quality impairments anywhere in the network. MSOs can now fully and reliably characterize ingested video using machine algorithms instead of, or in addition to their so called 'golden eye' human video monitors, without requiring a copy of the original clean video. More importantly, they can place the equivalent of their golden eyes anywhere in their networks, and tune the results so that the maximum amount of video quality can be delivered under varying network conditions and competitive threats.

But using such a tool requires complete understandings of the video artifacts that can be detected by a no-reference, hybrid video quality technology, especially when grounded on human visual system modeling and testing, and so details and examples of these artifacts are given in this paper. Emphasis will be on the video image artifacts only; audio artifacts and synchronization issues are not covered.

VIDEO QUALITY IMPAIRMENTS

The two key types of visual digital video impairments that occur in otherwise properly functioning video networks are compression artifacts (CA) and network artifacts (NA). While less common, interlacing artifacts such as 'mice teeth' are also noticeable in modern digital video systems. Also far less common are impairments due to equipment malfunctions such as encoder errors and/or improper configurations, although they do occasionally occur. But to affordably offer video services to subscribers, some amount of CA and NA must be tolerated, and therefore a good understanding of the different types of CA and NA are needed.

It is important to note that even in top quality video, many video quality (VQ) impairments are visible on individual frames. But the human visual system (HVS) misses them if the video is properly encoded and the viewing distance is that of typical viewers, i.e., 2-3 times the diagonal length of the display. Humans can also miss artifacts if the impairment is away from the center of the screen, or if the temporal duration of the impairment is very short. On the other hand, if a video expert is viewing the digital video from a very close distance, many more artifacts can be seen, including weaker ones that would be missed at a greater distance or missed by untrained viewers, regardless of viewing distance. A properly designed video quality measuring and monitoring system permits a cable operator to vary the results so that the operator can automatically emulate either a video expert, or a typical subscriber, or something in between.

Compression Artifacts

Compression artifacts (CA) occur when the bit rate of the encoded video is insufficient to provide smooth video motion without unnatural jerks, blocks, blur, noise or jagged edges being visible in the video stream. In the worst cases (lowest bit rates), blocky CA can be seen even in static video, but this is very atypical of modern high quality video. The CA types most often visible to viewers (as evidenced by blogs on the subject) can be delineated into video artifacts that are blocky, blurry, and choppy. Unlike network artifacts that can appear somewhat randomly across the screen, most CA are typically seen in conjunction with motion in the video and are usually associated with the object in motion. Circular motion, explosions, scene changes, roaring fires, and fireworks often reveal compression artifacts when they are otherwise not detectable. Trails, which are vertical lines or colored blocks that trail from a moving object in the video, are actually detectable as either blocky or blurry artifacts and are thus included in these two categories. Jagged edges are also a form of blocky compression artifacts. Posterization, which is a loss of color depth, and color errors are also quite noticeable to viewers, but are less common in current video networks when properly configured. Mice teeth, which is an interlacing artifact, is also quite noticeable to viewers when present.

On the other hand, video experts looking up close can see blocky and blurry compression artifacts even when they are not visible to untrained viewers at normal viewing distance, which is to say that they can detect CA of far lower strengths. They can also detect more subtle CA that include the two manifestations of Gibbs effect, namely mosquito noise, which is random speckling around the edges of objects, and ringing, which is spatial ripple away from the edges of text or other sharp edges in the video. Another subtle artifact has been called occasional blur, or "background breathing", which is frequent sudden blur in, for example, background foliage in the video. Although more subtle and often missed by casual viewers, these artifacts are also described below since once artifacts can be detected by machine algorithms, it is a simple matter to convert these detections into appropriate metrics that reflect either expert viewers or more typical home viewers.

1) Blockiness: The most obvious compression artifact for both experts and typical viewers is blockiness, where the bit rate is too low for the level of action or spatial variation in the video. Strong blockiness is obvious to even untrained viewers, while slight blockiness is detectable only up close or by video experts. In Figure 1, there is actually slight blockiness in the first image on the top, but this would normally only be detectable by a video expert looking closely at the video. The middle image shows moderate blockiness that many viewers would miss if it were brief, while the bottom image shows blockiness that all viewers would notice.

In properly configured video networks, a slight amount of blockiness is often acceptable because it either happens infrequently, or is only detectable when video experts look up close at a screen.

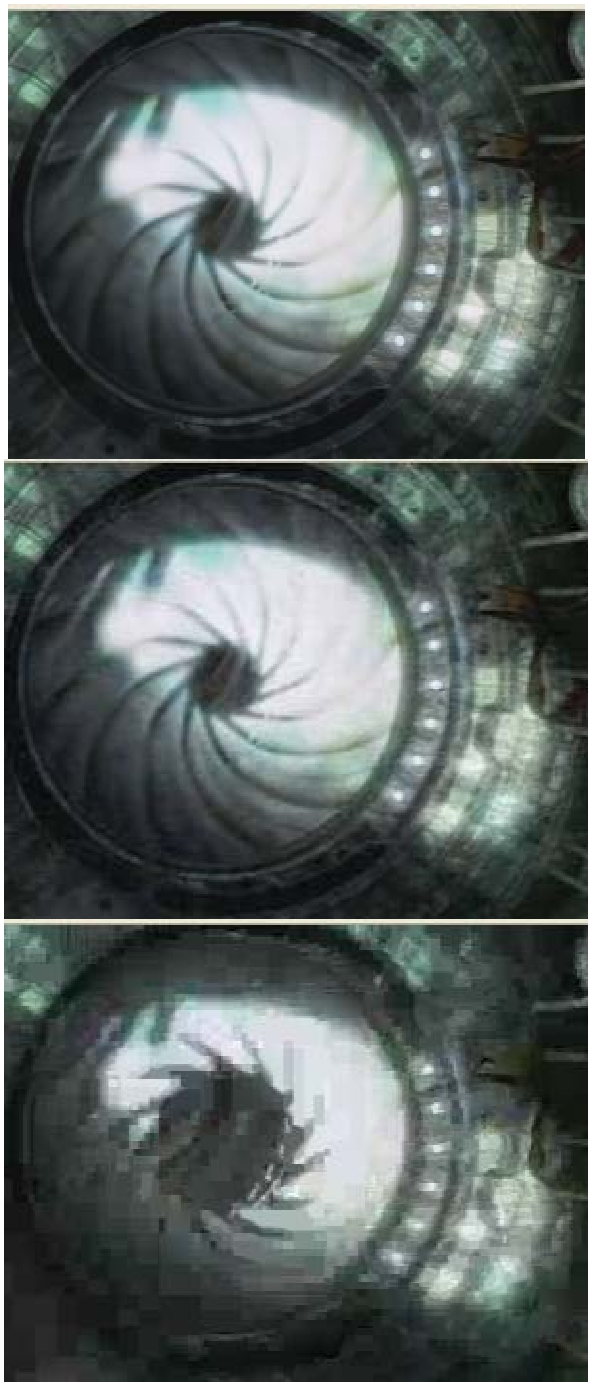

Figure 1. Slight Blockiness (top), moderate blockiness (middle) and strong blockiness (bottom)

2) Blurriness: Codec induced blurriness similarly occurs when the bit rate is too low, however the effect is often more subtle to viewers and thus a greater amount can be tolerated by untrained or less picky viewers. Especially in fast moving scenes, even moderate blurriness can be missed entirely by many viewers. Video experts however, can not only detect even slight blurriness, but can readily tell the difference between naturally blurry features in the video (objects that are intentionally out of focus, e.g.) and codec induced blurriness. Therefore, video quality systems should also be capable of detecting even slight codec-induced blurriness if they are to mimic experts.

Figure 2 shows natural blurriness in the image as seen near the portal opening at the top left in the original image on the top, while the image on the bottom shows codec-induced blurriness as seen around the four lights on the rightmost image in Figure 2.

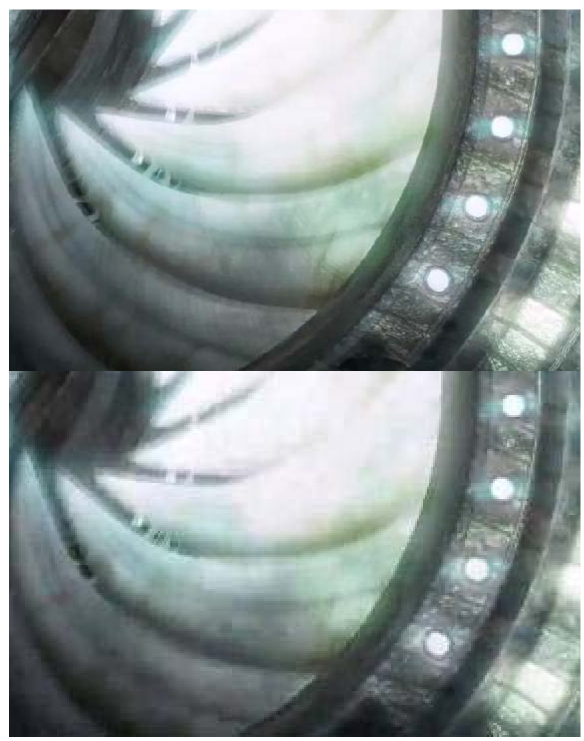

Figure 2. Natural vs. MPEG4 codec-induced blurriness in original (top) and low bit rate (bottom) video.

While previously it was difficult for machine algorithms to detect such subtle differences between codec-induced blurriness and natural blurriness, new video quality measuring technology is quite capable of discerning codec-induced blurriness from natural blurriness and alerting operators to the presence of blurry compression artifacts without undue false alarms from natural blurriness in the video.

Note that in the bottom image in Figure 2 there is both blockiness and blurriness produced by the lower bit rate. It is often the case that several compression artifacts are simultaneously present in a video when the bit rate is too low, and a properly designed video quality measuring system should indicate specifically which artifacts are present. This is so that an operator can, for example, control the bit rate carefully to minimize blockiness that can be seen by most viewers but permit a higher level of blurriness since far fewer viewers will notice the latter.

3) Choppiness: A lack of smoothness in motion in a video (sometimes called jerky video) produces choppiness in the video that can also be detected by newer video quality measuring technologies. While difficult to depict in still images here, it can be imagined as slow motion video effects where they should not be. New video quality measuring technology can also detect this type of choppy compression artifact.

4) More subtle compression artifacts: Unless the bit rate is grossly under the required level for quality viewing, more subtle compression artifacts are often missed by most viewers, but are nonetheless detectable by video experts. Examples include mosquito noise and ringing (which are both manifestations of Gibbs effect), and sudden blurriness or ‘background breathing’ which is often noticeable in MPEG4 encoded video at lower bit rates. Mosquito noise and ringing are shown in Figure 3. On the top, mosquito noise is seen as random speckle around the text, while on the bottom there is a periodicity to the smearing of the edge of the image, especially at the top of the letter “O” for example.

Figure 3. Mosquito noise (top) and ringing artifact (bottom) around text.

Another compression artifact which can be subtle is the so-called ‘background breathing’ a somewhat frequent blurring of the background, which can be seen in MPEG4 video in particular when there is foliage in the video background. Figure 4 shows a closeup of the fixed background in two consecutive frames of MPEG4 video. The entire region appears to ‘jump’ slightly in the video due to a sudden increase in blurriness, which can be seen in the highlighted region in particular, but in actuality the entire background is perceptibly different.

Figure 4. Background breathing example in MPEG4 video.

5) Interlacing artifacts: Another video artifact in interlaced formats (480i, 1080i) that is not so subtle and also detectable by modern video quality monitoring technology is the so-called 'mice teeth', shown in Figure 5 below, and when present, it is also highly visible to even untrained viewers from a distance.

Figure 5. Interlacing artifacts (mice teeth).

Network Artifacts

Network artifacts in video may occur when packet errors occur such as loss, excessive delay or jitter, or even that packets are sent out of order. The video may or may not be affected depending on the type of MPEG frame involved, error masking techniques in use, and so on. Thus, not all network errors lead to actual video artifacts. On the other hand, if network artifacts in the video are re-encoded such that packets are renumbered, there may be network errors in the video that do not have corresponding packet error indications, which means the network artifacts can only be detected via pixel analysis of the video.

The worst network errors result in stuck frames, lost frames, and blank frames, and these are all detectable by most modern video quality measurement systems. Amazingly, entire frames can be lost and still the viewer not notice it if the video is relatively static. This is also true of the most typical network errors, where portions of the frame are affected, especially those in motion. Error concealment can essentially repeat the affected portion of the frame from a previous frame and thus the network error looks like a slight jerk in one portion of the video. With error concealment turned off however, the network errors appear as either streaks, blocks around the edge of a moving object, or a checkerboard type pattern. Figure 6 shows several network artifacts (NA) visible in MPEG2 and MPEG4 video.

Figure 6. Checkerboard and streaky (top) MPEG2 network artifacts, and H.264 network artifacts (bottom).

Note that these figures depict the errors as they actually appear, i.e., when error concealment in the decoder is disabled. In the checkerboard artifact, the entire screen is affected, however even this can be concealed or missed if it occurs in a single frame only. Network streaks, as seen in the middle image of Figure 6 are far easier to mask when they are localized, and if they occur at the edge of the screen they may not even be perceived by typical viewers.

In a sense, MPEG2 network artifacts are more straightforward to detect, even in the absence of packet errors in the video stream, because of the regularity of macroblock features seen in the video. Network artifacts in H.264 (MPEG4 part 10 or AVC) video, on the other hand, are much more challenging to detect via pixel analysis because the artifact pattern is much more varied and non-regular, as seen in the bottom image in Figure 6.

Nevertheless, network artifacts in H.264 can be detected by advanced video analysis algorithms, and alternately when packet errors occur, the video can be analyzed to determine what impact, if any, the error had on the H.264 video.

Other Video Artifacts

The other class of video artifacts that occur occasionally are due to errors or failures in the encoders themselves, which can take the form of super macroblocks in the video, often seen at the edge. However, many of these patterns are specific to the encoder and thus are not shown here.

MEASURING IMPAIRMENTS WITH NO-REFERENCE TECHNOLOGIES

As mentioned in the introduction, the key benefit of using no-reference, hybrid bitstream-pixel video quality measurement technology grounded in subjective testing is that the video may be analyzed anywhere in the network, from initial video ingestion point to the home, and it provides an accurate match to what humans would have perceived. And unlike approaches that rely solely on packet errors, and therefore require such devices throughout the network for locating the source of video problems, no-reference hybrid techniques can detect and classify compression and network artifacts even after subsequent packet processing, which means monitoring devices can potentially be more sparsely deployed in a network and still accurately detect the vast majority of artifacts.

However, since no-reference techniques do not have the original, assumedly clean video for comparison, there will a non zero error rate in both the probability of false alarm (Type I errors) and the probability of missed detections (Type II errors). It will be up to the individual MSO to decide where best to draw the line between minimizing Type I errors vs. minimizing Type II errors, depending on whether maximum visibility of artifacts is desired, or a mimic of untrained viewers who are likely to miss many subtle errors. The required missed detection and false alarm probabilities are also likely to be a function of the value of the video asset, level of artifacts already present in the video, the time of day, and tradeoffs between certain video services and other services such as high speed data for business customers. MSOs should therefore tune the quality monitoring system to their specific headend needs or the needs of their subscribers, which may vary considerably.

The results of video quality measurements can be mapped to a variety of metrics for MSOs, such as number of artifacted seconds, or Mean Opinion Score (MOS). Many more granular and higher level metrics are also possible and have been implemented in new technology from VQLink, and these metrics can be used to characterize video streams absolutely, relative to each other, and also over time for trend analysis. Ideally, the measurement system should not only detect artifacts, but also classify them. But for any of these metrics to have meaning, the video quality measuring system must have adequate Type I and II error performance for detection and classification of video artifacts, and this should be grounded on, and verified with human subjective testing to ensure accuracy.

CHALLENGES WITH MEASURING VIDEO QUALITY OF TRANSCODED CONTENT

One of the challenges for no-reference video quality measurement is when the video has been transcoded. In the simplest example, low bit rate video could be reencoded at a higher bit rate prior to ingestion by the MSO, thereby preventing the MSO from knowing the original bit rate. An accurate video quality measurement system would not be fooled by the higher bit rate, but rather would detect the video artifacts correctly regardless of the bit rate of the video. It is much more challenging when not only the bit rate, but also the type of encoder used is changed by the transcoding. What can then happen is that the compression artifacts typical of one type of encoder at a low bit rate are present in a video stream which uses a different type of encoder and bit rate. While this would previously mean higher Type I/II error rates, new video quality measurement algorithms are much more intelligent and can perform adequately even for this type of content, although the error rates will likely vary depending on the specifics of the original encoder, the final encoder, and any bit rate changes that occur.

USING ACCURATE VIDEO QUALITY MEASUREMENTS TO IMPROVE NETWORK EFFICIENCY

Once an accurate video quality (VQ) measurement and monitoring system is in place, there are several use cases for this technology of interest to MSOs, content providers and content aggregators. First, the MSO can use VQ measurements to characterize the level of artifacts in ingested video. This gives the MSO the ability to groom the channel as needed and compare the results to the ingested stream to ensure that no material degradation has occurred, both in the headend and at the edge of the network. Since the quality of ingested video varies considerably, characterizing the video at this point with maximum accuracy is particularly important.

Second, with accurate VQ measurements, content providers and MSOs that offer video streams over the Internet can trim the bit rates so that quality needs specific to Internet delivery are met, which may be different from those on their cable networks, for example. For content providers, this represents potential savings of bandwidth costs for delivery over the Internet. As before, if the video delivered is already artifacted to some extent, there is generally room for additional bandwidth savings as long as the video quality can be accurately measured and maintained.

Third, MSOs seeking to add more HD channels, or increase the number of QAMs allocated to high speed data service (in the daytime for business customers, for example) may need to alter their QAM lineups using accurate video quality measurement of channels and their trends over time. While the 3DTV frame compatible system currently proposed for cable is designed to avoid additional bandwidth requirements for 3D content, there are proposals for enhancements to frame compatible 3DTV that would add additional bandwidth requirements. And if the market demands full 3DTV over time, the bandwidth needs of 3D channels will certainly increase.

CONCLUSIONS

New, no-reference video quality measurement technology that employs a hybrid of bitstream and pixel processing to perform full analysis of video is available that offers MSOs the ability to accurately detect, classify, and monitor video artifacts anywhere in their networks, from initial ingestion to edge delivery. However, to use such technology maximally, the MSO must understand the types of video artifacts common in cable networks, their impact on video quality as a function of artifact type and strength, and use this information to configure video quality systems so that the data reported is both accurate and actionable. Too many false alarms mean that such systems can become ignored by technicians, while inability to detect artifacts under all conditions, including transcoded video, leads to a false sense of security and lack of trust in the monitoring system.

With the emerging video quality wars between competing video service providers, the provider who can *affordably* deliver superior video quality to subscribers, and monitor and enhance that quality over time will have the edge. Further, as more home viewers get ever larger TVs, and blogs, forums and other social networks as well as marketing tactics focus the viewer's attention on video quality, the differences between the professional video expert and the home viewer is likely to decrease over time. Hence, understanding video artifacts fully and having an ability to accurately measure and monitor them will be critical to both the current and the future success of MSOs in their delivery of video services.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation under Grant No. 0848558. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.