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

Case Study: Can Video Quality be Quantitatively Measured?

Bill Reckwerdt, CTO
Video Clarity, Inc.

The Next Generation of Video Measurement Tools

Video algorithm designers, equipment manufacturers, and their customers need a quantitative, repeatable method to evaluate picture quality. Before compressed video, it was fairly obvious how to evaluate video quality: parameters such as color, brightness, and contrast were measured. With the advent of compressed digital video transmission, the process has changed considerably.

As an over-simplification, compression techniques knowingly discard a certain amount of the original content. Visible impairments such as "blockiness" or "softness" are inevitable. Traditional measurement techniques are no longer effective, since video content greatly affects the perceived quality.

To date, the solution has been subjective picture quality testing. Human subjects are gathered to view and judge video images based on guidelines outlined in ITU-R BT500. In spite of the controlled testing, subjective testing causes inconsistencies due to human judgment from viewer-to-viewer.

A new generation of video quality analysis systems needs to be defined to re-answer the basic question. Can we automatically assess video quality?

What is Video Quality

Since processing video may take many forms: compression, video enhancements, and color space conversions to name a few, a system must be put in place to normalize the video information so that a comparison can be done. It must be remembered that compression format is not important - Windows Media, DviX, MPEG-2, M-JP2 or MPEG-4. All that matters is whether the video left them with the desired impression.

For definition purposes, video is the technology of electronically capturing, recording, processing, storing, transmitting, and reconstructing a sequence of still images representing scenes in motion. Video sequences are the carriers of visual information from the outside world, and are used as input to the human visual perception stage of interaction. Technically, perception is the ability to measure and quantify one thing with respect to another. Based, on this definition of video, we can state that video quality is defined as adequately discerning one image from another.

To this end, we need a metric to quantifiably define the video sequence's discerning attribute strength. In this paper, we use 3 metrics. The goal is to assess the quality of a video without staring at it for hours.

Objective Measurement Setup

The standard MPEG relatively-uncompressed football sequence acts as the source for test. A 5-second portion of this sequence was brought into Adobe Premiere. Different impairments were entered every second. The impairments are listed below:

1. 5 frozen frames
2. 6 frames of black
3. 6 frames of unrelated video. Specifically, changing color bars.
4. 5 frames of the same scene, but compressed to MPEG-2 at 15Mbps.

The "original" and "processed" sequences were created as uncompressed AVI files. These were loaded into ClearView as 1920x1080, Y'CbCr, 8-bit video sequences, which was the native configuration of the AVI files. Even though ClearView can process the input, it was not asked to decode the image, perform color-space conversions, etc.

ClearView displays the 2 video sequences in multiple viewing modes, aligns the "original" sequence with respect to the "processed" sequence; then calculates, logs and graphs the following objective metrics:

- PSNR – Peak Signal to Noise
- Spatial – Std-Dev of the pixel values with respect to their neighbors per frame
- Temporal – RMS of the pixel values frame-to-frame.

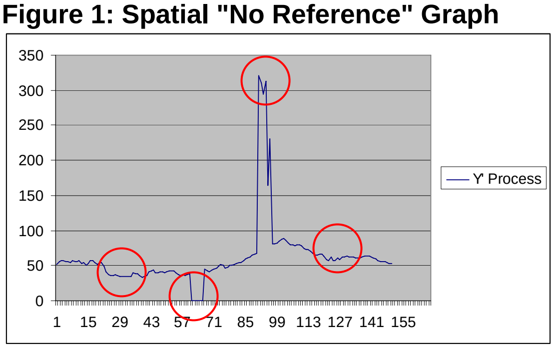
Interpreting the Objective Metrics

For these tests, PSNR is a reference objective measurement between the "original" and "processed" sequences. It is also possible to calculate PSNR with respect to black, which effectively calculates contrast and brightness, but this was not enabled for this experiment. Spatial and Temporal are no-reference objective metrics. They are later compared to form a quasi-reference objective measurement. For the purpose of data reduction, the data presented is Y' only. The video analysis is also done for Cb and Cr (or R, G, and B).

The Spatial Objective Metric is looking for sharp edges or transitions within a frame. Effectively, it is looking for areas with a great deal of disturbance within a frame. It is then plotted over time.

- Flat line around 1 second, which indicates frozen video.
- Flat line around 2 seconds, which indicates solid color (Spatial of 0).
- Huge spike around 3 seconds, which shows activity.
- No issues around 4 seconds

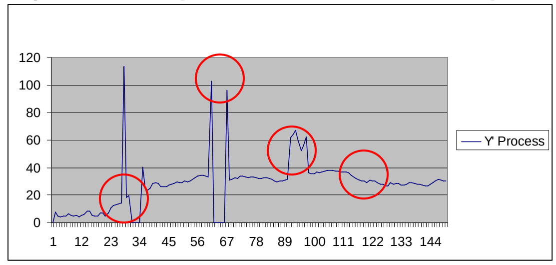
Figure 1: Spatial "No Reference" Graph



The Temporal Objective Measurement looks for transitions pixel-to-pixel over time. It finds scene changes and random noise introduced into the video sequence.

- A huge disturbance followed by a zero line around 1 second, which indicates a frozen video
- A huge disturbance followed by a zero line around 2 seconds followed by a huge disturbance, which indicates a frozen video and a probable solid color.
- Huge spike around 3 seconds, which shows multiple scene changes in a short period of time.
- No issue around 4 seconds.

Figure 2: Temporal "No Reference Graph"



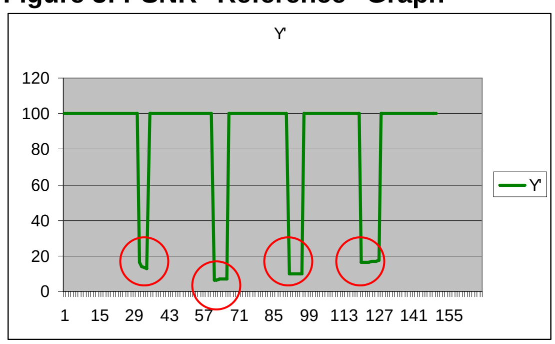
The next three graphs show a reference analysis of the impaired video with respect to the original (Y' only). The "original" video sequence is plotted as green, the "processed" video is red, and the delta is blue.

PSNR calculates and scores the "original" and "processed" video sequences with respect to each other.

PSNR shows:

- 4 disturbances once per second.

Figure 3: PSNR "Reference" Graph

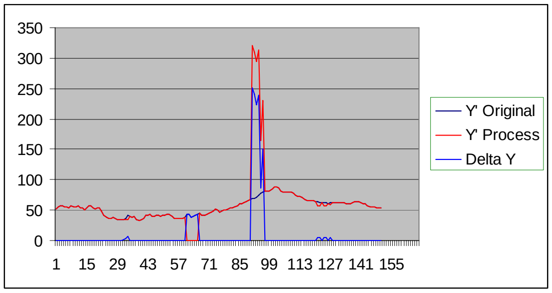


Spatial and Temporal calculate and score the "original" and "processed" video sequences separately, align, and then compare the results.

Spatial shows:

- 4 disturbances once per second. (blue lines)
- The disturbances around 1 and 4 seconds are fairly small, which shows that the original did not move much with respect to the impaired.
- The disturbances around 2 and 3 seconds are showing completely different video.

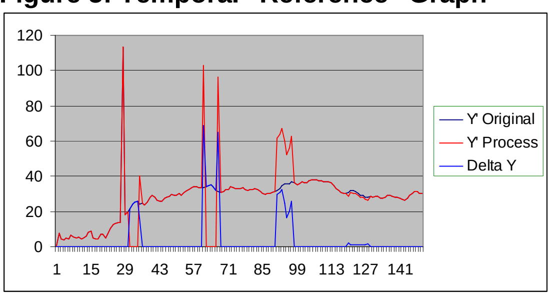
Figure 4: Spatial "Reference" Graph



Temporal shows:

- 4 disturbances once per second. (blue lines)
- The disturbances around 1, 2 and 3 seconds are large, which shows a big disturbance temporally.
- The disturbance around 4 seconds is small showing a fairly close correlation.

Figure 5: Temporal "Reference" Graph



From the non-reference metrics, we have learned

- Around 1 second, the video is frozen.
- Around 2 seconds, the video is a solid color.

From the reference metrics, we have learned:

- Around 3 seconds, temporal and spatial show a disturbance which looks like the wrong video is playing. This is referred to as hi-jacking the video signal and is a great concern in the IP community.
- Around 4 seconds, PSNR shows a problem, but the other metrics see no issue. This means that the video quality has been reduced, but not so much that it created a temporal or spatial disturbance. Most likely the average viewer will not care.

Summary

Objective picture quality measurements are an important part of equipment design and evaluation, bandwidth allocation, and QoS compliance programs. ClearView provides multiple objective scoring techniques while allowing the user to view the results in multiple viewing modes. Using ClearView's set of algorithms and setting up a controlled test, it is possible to create consistent, automated, quantitative objective video quality results.

ClearView Implementation

Video Clarity defined the ClearView to be an open-architecture solution for video test and measurement. As algorithms and interfaces change, ClearView grows to provide complete solutions for its customers.

- Capture video sequences in as many formats as possible.
- Matt/Crop all video sequences to user-selectable resolution.
- Translate all video sequences to uncompressed Y'CbCr 4:2:2 or RGB 4:4:4.
- Support 8 and 10-bit data paths with upgradeability to future 16-bit modes.
- Store the video sequences as frames (fields) so that they can be played at any rate.
- Display the video sequences in real time in multiple viewing modes.
- Playback controls include play, shuttle, jog, pause, pan, and zoom.
- Apply objective metrics to the video sequences.
- Export pieces of video sequences to further analyze off-line.

To simplify the work flow, any video sequence can be played; while capturing another video sequence, thus, combining the video server, capture device, and video analysis into one unit. By doing this ClearView controls the test environment, which allows for automated, repeatable, quantitative video quality measurements.

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.