



Tools for Video Analysis



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

White Paper: How to Do Objective Video Testing

*Bill Reckwerdt, CTO
Video Clarity, Inc.*



Over recent decades, the role of video images has grown steadily. Advances in technologies underlying the capture, transfer, storage, and display of images have created situations where communicating using images has become economically feasible. More importantly, video images are in many situations an extremely efficient way of communicating as witnessed by the proverb “a picture is worth a 1000 words.”

Notwithstanding these technological advances, the current state of the art requires many compromises. Examples of these compromises are temporal resolution versus noise, spatial resolution versus image size, and luminance/color range versus gamut. These choices affect the video quality of the reproduced images. To make optimal choices, it is necessary to have knowledge about how particular choices affect the impression of the viewer. This is the central question of all video quality research.

Current video quality research can be divided into 2 approaches: experimental evaluation and modeling.

Experimental Evaluation

A group of human subjects is invited to judge the quality of video sequences under defined conditions. Several recommendations are found in ITU-R BT.500.10 “Methodology for Subjective Assessment of the quality of Television Pictures” and ITU-T P.9210 “Subjective Video Quality Assessment methods for Multimedia Applications. The methods are summarized [here](#).

The main subjective quality methods are Degradation Category Rating (DCR), Pair Comparison (PC) and Absolute Category Rating (ACR). The human subjects are shown 2 sequences (original and processed) and are asked to assess the overall quality of the processed sequence with respect to the original (reference) sequence. The test is divided into multiple sessions and each session should not last more than 30 minutes. For every session, several dummy sequences are added, which are used to train the human subjects and are not included in the final score. The subjects score the processed video

sequence on a scale (usually 5 or 9) corresponding to their mental measure of the quality – this is termed Mean Observer Score (MOS).

Two serious drawbacks of this approach are:

1. It is extremely time consuming, and tiresome for the participants.
2. The obtained knowledge cannot be generalized because relationships between design choices and video quality are descriptive rather than based on understanding.

As a result, in a single series of experiments only a small fraction of the possible design decisions can be investigated. This makes the process even longer and more tedious.

Modeling

The second approach tries to address these drawbacks by means of developing models that describe the influences of several physical image characteristics on video quality, usually through a set of video attributes thought to determine video quality. When the influence of a set of design choices on physical video characteristics is known, then models can predict video quality. The models express video quality in terms of visible distortions, or artifacts introduced during the design process. Examples of typical distortions include flickering, blockiness, noisiness, or color shifts. Two types of models exist, where the fundamental difference between them is how the impairment is calculated.

In the first type, physiologically or psychophysically models of early visual processing are used to calculate impairment from a difference between the video sequences. Many well known metrics exist, which compare the “original” to the “processed” output:

- PSNR – Peak Signal to Noise Ratio
- JND – Just noticeable differences
- SSIM – Structural SIMilarity
- VQM – Video Quality Metric
- MPQM - Moving Picture Quality Metric
- NVFM - Normalize Video Fidelity Metric

The two most important drawbacks of this approach are



1. It is unclear what exactly the “original” version of a video is.
2. These algorithms are measuring visible differences not video quality.

The second type of model tries to estimate visible distortions directly from the “processed” video; instead of comparing it to the “original”. In this type of model, visible distortions of a video, such as unsharpness or noisiness are predicted by estimating physical attributes of the video. The advantage of this approach is that the “original” video sequence is not needed. The uncertain translation from visible distortions to video quality is an important drawback to this approach.

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>.