

Submission to G-2.1.6

**Report of Task Force
to define
A Unit of Measure and Means of Calibration
for
Video Quality Analysis**

Author: Leon Stanger
10 Jul 98

Task Force Members:

Alan Godber
Roger Green
Paul Jones
John Libert
Bruce Lilly
Al Morton
Mihir Ravel
Leon Stanger
Andrew Watson

Although participation has been light in recent months, some concepts have been put forward and considered by the task force. This report will outline the basic concepts and summarize the findings and recommendations of the group.

Goal:

To define scale or unit of measure and a calibration method for video quality impairments. The unit of measure must be defined in terms of human perception and calibrated by statistical analysis of viewer test results.

Concepts:

Three concepts have been put forward for consideration. These concepts are summarized as follows:

1. A continuum of picture impairment.

This method is based on having at least five steps to define the picture degradation after passing through a system under test. The definitions could range from excellent or no impairment to objectionable.

While this method would carry far more information concerning the degree of impairment, there is a great deal of concern about having repeatable results with different groups or labs.

2. A threshold measurement based on A/B comparisons.

This method will determine a statistically significant point where viewers can see a difference between, or will prefer picture A compared to B. The most common terminology for the threshold point is the just noticeable difference (JND).

The greatest difficulty of this method is to define a scale that spans multiple and fractional JNDs. One method is to use the same comparison process such that the first JND point B is compared to the second JND point C.

3. A threshold measurement using distance as the metric.

This method is based on the concept that as the viewer moves closer, differences become easier to see and a statistical threshold can be determined where a preference could emerge between A and B. In practice, the “distance” could be a virtual function by changing the image size relative to the viewer.

The concern raised is that results could be influenced by the viewer’s eye sight. It may also favor or penalize some artifacts. For example, edge artifacts are best seen up close where luminance variations on large areas can be seen from a distance.

Of the three methods discussed, a casual survey indicated a preference for 2, then 3, then 1 for focusing efforts for future work. The responses to the survey were, however, very light so this point should receive additional discussion at the next general meeting.

Conclusions of the Task Force:

1. That a JND based scale should be established as a calibration reference for video quality measurements.
2. That the scale have a specific definition for multiple and fractional JND units.
3. That processes used for viewer testing focus on repeatable results. This may imply closer viewing distances and trained observers as opposed to what may be seen in an average home with greater distances and untrained observers.

Recommendations:

1. That G-2.1.6 assign additional task forces to further define the specific requirements for a JND based calibration reference system.

This work should include the detailed viewer test specifications and computational methods to derive a multiple step JND scale.

2. That a task force be assigned to find suitable test material.

Test material must include sequences with monotonically increasing degradation from typical DCT based compression algorithms.

3. That a task force be assigned to identify labs willing to conduct viewer tests.