COMMITTEE T1

CONTRIBUTION

Document Number: T1A1.5/93-60

****

STANDARDS PROJECT: Digitally Transported Commercial Television Standard

****

TITLE: Objective Performance Parameters for NTSC Video at the DS3 Rate

****

ISSUE ADDRESSED: Objective and Subjective Video Performance Testing of DS3 Rate Transmission Channels

****

SOURCE: NTIA/ITS, Stephen Wolf, Arthur Webster

****

DATE: April 28, 1993

****

DISTRIBUTION TO: T1A1.5

****

KEYWORDS: Video Performance Testing, Objective Quality, Subjective Quality

****

NOTE: The methods and algorithms presented in this contribution are patent pending by the Department of Commerce. Licensing of the technology is currently available to all interested parties.

****

1. Overview

1.1 Background

The Institute for Telecommunication Sciences (ITS) has developed a number of objective performance parameters for digital video systems over the past 2 years. Contributions that document several of these parameters include T1Q1.5/92-112, T1A1.5/92-112, T1A1.5/92-135, T1A1.5/92-136, T1A1.5/92-138, T1A1.5/92-139, and T1A1.5/93-032. Most of these contributions have concentrated on objective performance parameters for video teleconferencing systems from 56 kbps to DS1 rates, but several commercial grade and contribution grade 45 Mbps systems were also addressed. Recently, ITS implemented the most promising objective video performance measurements in a real-time PC-based measurement system. Details of this can be found in T1A1.5/93-032.

During 1990, committee T1Y1.1 conducted extensive subjective tests of contribution quality 45 Mbps systems. D2 copies of the subjective viewing tapes and the subjective viewing results from these extensive subjective tests were made available to ITS. The T1Y1.1 video and subjective viewing test results have considerably expanded the amount of 45 Mbps data available to ITS.

1.2 Goal

The goal of ITS video quality research is to apply and extend the technology independent objective video quality metrics to a wide range of video systems. To this end, ITS has applied the video quality measurement methodology that is documented in previous contributions (T1A1.5/92-112, T1A1.5/92-135, and T1A1.5/92-136) to the T1Y1.1 45 Mbps data.

1.3 Approach

The first experiment that was conducted applied the ITS video quality metrics (which have been shown to work well on other data sets) to the T1Y1.1 subjectively rated video data. For convenience and ease of reference, the methods of measurement for these technology independent video quality parameters have been included in section 3. The subjective-objective correlation results from this first experiment are presented in section 2.1. The original ITS video quality metrics worked well in 33 of the 38 T1Y1.1 transmission channels that were tested. Inspection of the video scenes for the other 5 T1Y1.1 transmission channels revealed that they were introducing line-oriented noise. While the original parameters were detecting this noise, they were not penalizing the impairment as much as the T1Y1.1 viewers. Evidently, noise that has structure (e.g., noise lines) is more irritating to viewers than unstructured noise (i.e., noise that is randomly distributed throughout the picture area).

An investigation began as to the best method for imposing additional cost penalties (above those imposed by the original quality prediction model) when the noise has structure (such as lines). This investigation resulted in the development of a new perception-based parameter that can distinguish structured line-oriented noise from unstructured noise. Hence, this new parameter provides a valuable

addition to the technology independent parameters that have already been introduced. The method of measurement for this additional parameter is presented in section 3. The subjective-objective correlation results that include this additional parameter are given in section 2.2.

In addition to the above, section 3 of this contribution describes a modified method of measurement for a video delay parameter that can produce video delay measurements which are accurate to within one NTSC field (1/60 second). The method of measurement is similar to the one presented in contribution T1A1.5/92-139, except that NTSC field accuracy (1/60 second) is possible. The video delay parameter, in addition to being a useful performance parameter by itself, can be used to improve the accuracy of the technology independent video quality parameters. This is achieved by providing proper alignment between the source and destination video quality measurements.

1.4 Summary

In this contribution, we demonstrate the usefulness of the ITS technology independent performance parameters in the studio to studio broadcast community. An additional parameter has been developed which measures the perceptual decreases in video quality that result from line-oriented noise. Line-oriented noise, an impairment that was not significant in the ITS video data sets used to develop the original video quality metrics, appears in some contribution quality video systems.

2. Subjective and Objective Test Results

In producing the T1Y1.1 subjective test results, viewers were asked to rate the quality of both the source and destination video scenes on a scale of 0 to 50 points. For the T1Y1.1 subjective scale, 0 points represented “bad” quality and 50 points represented “excellent” quality. The presentation order of the T1Y1.1 source and destination video scenes was randomized so that the viewer did not know which scene was the source. We linearly mapped the mean opinion scores of the source and destination video scores to a 5 point scale such that the source video (highest score per scene) was mapped into a quality level of 4.9. A quality of 4.9 was chosen for the unimpaired source video quality since we have found that the average mean opinion score for unimpaired source video is about 4.9 (on a scale of 1 to 5).

One hundred thirty subjective data samples from the T1Y1.1 subjective viewing data were analyzed. These samples came from 38 different transmission channels that included the null or no impairment channel (channel 1), an anchor channel with noise (channel 2), one, two, three, four, five, and six passes through four different 45 Mbps codecs (channels 3-26), four different 45 Mbps codecs with a bit error rate of 10^{-5} in the digital transmission circuit (channels 27-30), four different 45 Mbps codecs with a bit error rate of 10^{-4} in the digital transmission circuit (channels 31-34), and four different 45 Mbps codecs with burst errors in the digital transmission circuit (channels 35-38).

Since the ITS video quality laboratory does not have D2 format video equipment (the format of the T1Y1.1 subjective viewing tapes), a D2 to Betacam SP

dub was first performed on the T1Y1.1 subjective viewing tapes, and then the Y channel of the resulting Betacam SP dub was sampled at 4 times the color sub-carrier (4 x f_{sc}) using a high quality 8-bit frame-sampling device. The methods of measurement described in section 3 were then used to measure the parameters. Since the T1Y1.1 data did include transmission channels that contained up to 6 passes through 45 Mbps codecs, video gain shifts in the destination video were observed. These transmission channel gains were estimated according to

$$G = \frac{mean_{time}[std(Y_d(t_m))]}{mean_{time}[std(Y_s(t_n))]}, \quad (1)$$

where $std(Y_s(t_n))$ and $std(Y_d(t_m)$ are the standard deviations of the source and destination video images, respectively, and $mean_{time}$ denotes their average values over time. Although we used the video scene itself to estimate the transmission channel gain, one could also have used a standard video test signal (such as color bars). The gains estimated by equation (1) were used for G as specified in the methods of measurement for the video quality parameters in section 3.

2.1 Original Parameters Applied to T1Y1.1 Data

The ITS objective video quality metrics that were documented in previous contributions (T1A1.5/92-112, T1A1.5/92-135, and T1A1.5/92-136) were applied to the T1Y1.1 45 Mbps data. The 130 subjectively rated T1Y1.1 data samples were used to determine the optimal linear combination of the parameters discussed in these contributions. For convenience, the methods of measurement for two of these technology independent video quality parameters have been included in this contribution (see section 3). These parameters have been named to more accurately reflect what they measure. The Average Fraction Change in Edge Energy (AFCEE) parameter described in section 3 is identical to parameter m_1' in contribution T1A1.5/92-135, except that the AFCEE parameter does not include the normalization constant of 5.78. The Maximum Added Frame Noise Logarithmic Ratio (MAFNLR) parameter described in section 3 is identical to parameter m_3 in contribution T1A1.5/92-112, except that the MAFNLR parameter does not include the normalization constant of 4.2522. For the T1Y1.1 data samples, parameter m_2 in T1A1.5/92-112 did not significantly reduce the prediction error of the model for the T1Y1.1 data samples. Therefore, we have only included parameters AFCEE and MAFNLR in these results. For the T1Y1.1 data samples, the optimal linear combination of AFCEE and MAFNLR parameters is given by

$$\hat{s}_{T1Y1} = 4.72 - 14.1 \bullet AFCEE - 2.0 \bullet MAFNLR, \quad (2)$$

where $\hat{s}_{T1Y1}$ is the objective score of the T1Y1.1 data (on a scale from 1 to 5). Figure 1 presents a plot of this objective score (y-axis) versus the subjective score (x-axis). The coefficient of correlation for the entire set of 130 data samples was .79 and the root mean square (rms) error was .54 quality units. When the data samples were averaged across scenes to produce one point for each of the 38 transmission channels being tested (Figure 2), the coefficient of correlation increased to .88 and the rms error decreased to .43 quality units.

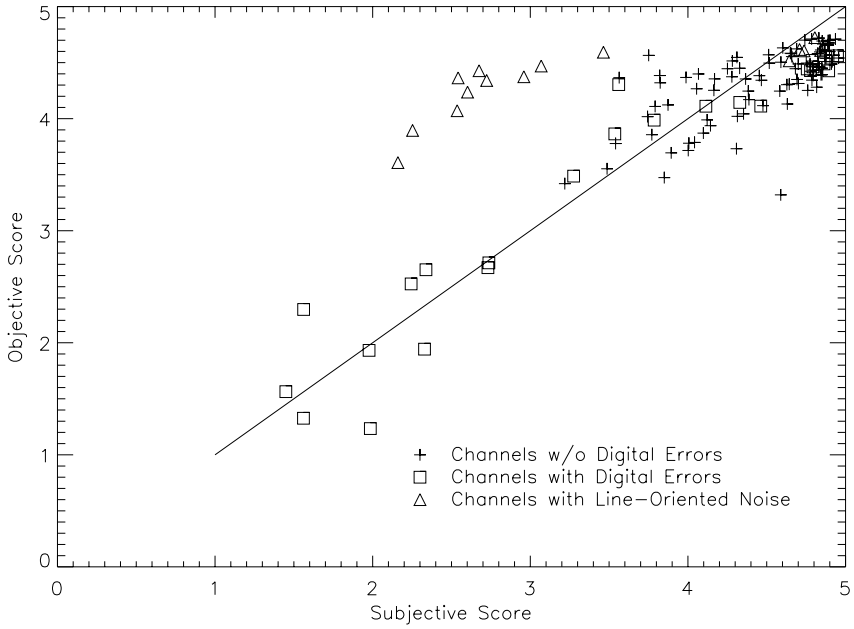

Figure 1 Model without MALNLR for 130 T1Y1.1 Test Scenes

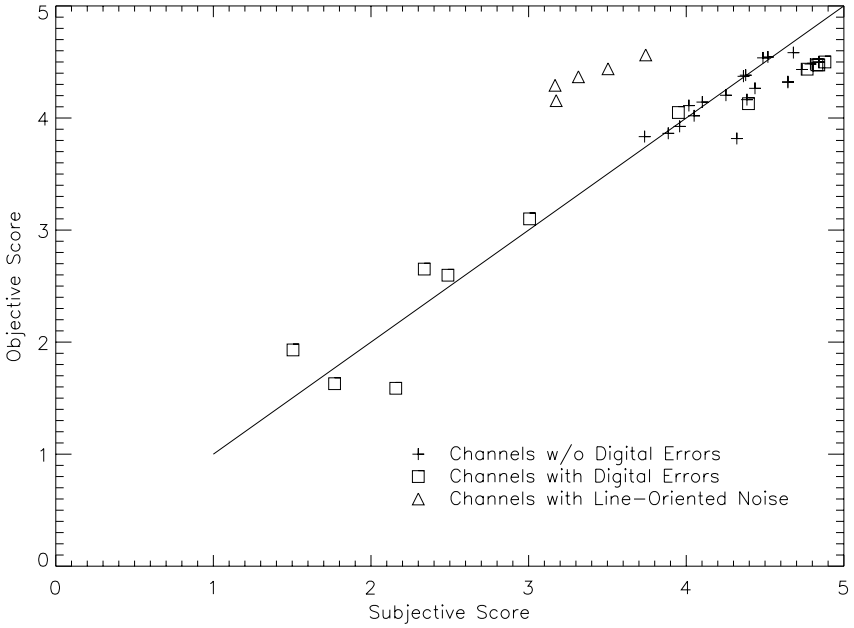

Figure 2 Model without MALNLR for 38 T1Y1.1 Transmission Channels

The plots in Figure 1 and Figure 2 separate the transmission channels that included error conditions (❑) from those that did not (+ and Δ). Upon inspection of the video scenes, it was noted that 5 of the T1Y1.1 transmission channels introduced line-oriented noise and that the original parameters, while detecting this noise, were not penalizing the impairment as much as the T1Y1.1 viewers. Data samples that belonged to these 5 transmission channels are represented with a Δ in the plots. When the transmission channels with line-oriented noise were removed from the 130 data samples (this reduced the number of data samples from 130 to 115 since 3 test scenes were used to test each of the 5 transmission channels), the coefficient of correlation for the set of 115 data samples and parameters was .92 and the root mean square (rms) error was .32 quality units. When the 115 data samples were averaged across scenes to produce one point for each of the 33 transmission channels, the coefficient of correlation increased to .97 and the rms error decreased to .22 quality units. Thus, the original ITS parameters performed very well on the T1Y1.1 data set, except for the 5 transmission channels that contained a perceptual line-oriented impairment. This is quite impressive considering that the original parameters have now been shown to perform well on 60 transmission channels that span a wide range of impairments (27 transmission channels in the original ITS test and 33 transmission channels in the T1Y1.1 test).

Analysis of the T1Y1.1 data has revealed that the line-oriented noise of the 5 transmission channels is dependent on the video scene. For a fixed transmission channel, some video scenes produced little added line noise whereas other video scenes produced a great deal of added line noise. This effect is very significant and can be responsible for more than 2 points of the 5 point CCIR-500 quality scale. One case was observed where the subjective quality varied from 2.16 to 4.64 for a fixed 45 Mbps transmission channel. The observation regarding line-oriented noise motivated the development of an additional technology independent performance parameter that could be used to complement the existing measures. This additional parameter, described as Maximum Added Line Noise Logarithmic Ratio (MALNLR) in section 3, can distinguish the perceptual video quality effects of line-oriented noise from random noise. MALNLR tracks perceptual decreases in video quality that result from increases in the amount of line-oriented noise. Inclusion of this additional parameter thus improves the accuracy of the objective model.

For completeness, the AFCEE and MAFNLR parameter coefficients were refitted using the original ITS data samples (recall that the original prediction model had AFCEE, MAFNLR, plus one additional parameter - see T1A1.5/92-112 and T1A1.5/92-135). This two-parameter fit is given by

$$\hat{s}_{ITS} = 4.89 - 7.1 \bullet AFCEE - 0.85 \bullet MAFNLR, \qquad (3)$$

where $\hat{s}_{ITS}$ is the objective score of the ITS data samples (on a scale from 1 to 5). The coefficient of correlation for the ITS data samples for equation (3) is .88 and the root mean square error is .64 quality units. If the ITS data samples are averaged across scenes to produce one point for each transmission channel, the coefficient of correlation increases to .97 and the rms error decreases to .29 quality units.

A interesting observation is that the weights on the parameters in equation (3)

are about one-half of the weights on the parameters in equation (2). One might conjecture from this observation that the T1Y1.1 viewers were about twice as critical as general viewers. This is understandable considering the context of the T1Y1.1 subjective tests (45 Mbps studio to studio transmission) and the T1Y1.1 viewers (broadcasters). Thus, the subjective rating of a particular video scene appears to depend upon the context of the subjective experiment and the criticality of the viewers. Test scenes that would have been rated high quality by general viewers (such as the subjective tests described in T1A1.5/92-112) may be rated as poor quality by critical viewers, such as broadcasters, in the context of testing contribution quality (i.e., studio to studio) video systems. Different subjective experiments may require a different set of weights for the individual objective parameters before they are combined into an overall objective quality score. To account for the fact that critical viewers are using a stricter quality criteria than general viewers, a technology independent quality assessment system with a viewer characterization switch may be desirable -- one switch position for general viewers and another for critical viewers. This might help to account for different classes of human viewers.

2.2 Original Parameters Plus MALNLR Applied to T1Y1.1 Data

The MALNLR parameter was added to the prediction model of equation (2) and the new optimal linear combination for the T1Y1.1 data samples is given by

$$\hat{s}_{T1Y1} = 4.97 - 9.34 \bullet AFCEE - 2.27 \bullet MAFNLR - 3.01 \bullet MALNLR \qquad (4)$$

where $\hat{s}_{T1Y1}$ is the objective score (on a scale from 1 to 5). Figure 3 presents a plot of this objective score (y-axis) versus the subjective score (x-axis). The coefficient of correlation for the entire set of 130 T1Y1.1 data samples and the model given by equation (4) is .92 and the root mean square (rms) error is .33 quality units. If the data samples shown in Figure 3 are averaged across scenes to produce a plot that contains 38 points, one for each of the transmission channels being tested (see Figure 4), the coefficient of correlation increases to .97 and the rms error decreases to .23 quality units. This demonstrates that the MALNLR parameter successfully measures the line-oriented noise of the 5 transmission channels, and thus complements the AFCEE and MAFNLR parameters. On the other hand, the addition of the MALNLR parameter to the prediction model of equation (3) for the ITS data samples does not change either the AFCEE weight, the MAFNLR weight, or the correlation results. This demonstrates that line-oriented noise was not a significant impairment in the original ITS data samples.

2.3 Future Work

This contribution has presented a new parameter, MALNLR, which is another technology independent video performance parameter that has been found to be useful for measuring the perceptual video quality effects of line-oriented noise. Recent research has revealed that other technology independent performance parameters may produce increases in objective video quality measurement accuracy. New parameters will either replace existing ones or become additions to the set of technology independent video performance parameters. ITS is continuing the refinement of the video quality prediction models.

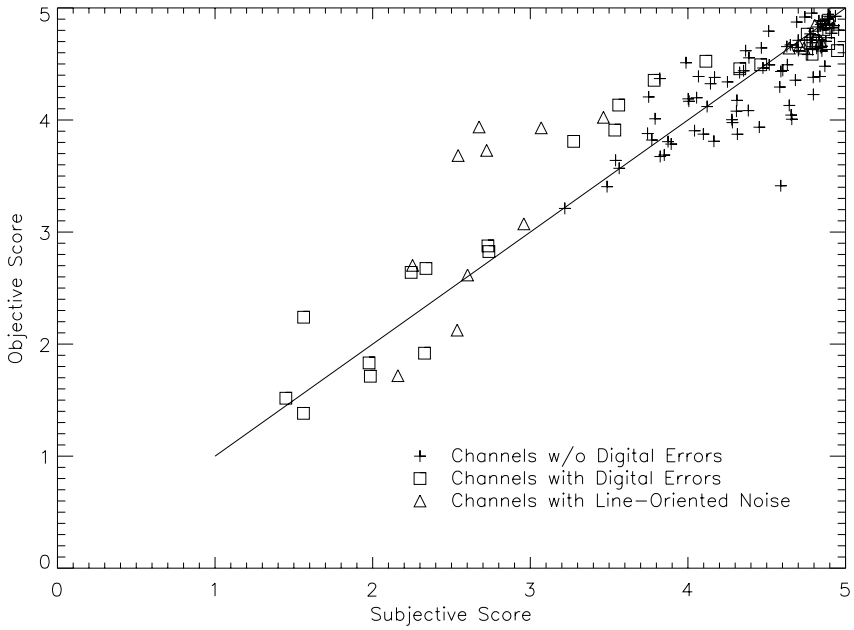

Figure 3 Model with MALNLR for 130 T1Y1.1 Test Scenes

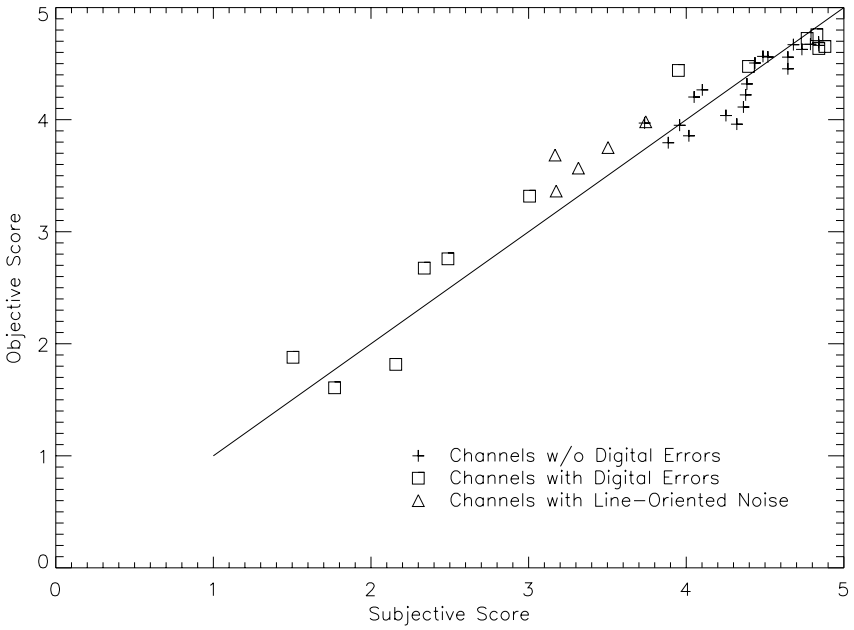

Figure 4 Model with MALNLR for 38 T1Y1.1 Transmission Channels

3. Methods of Measurement

3.1 Video Delay

The motion energy in a video scene can be used to accurately measure the video delay of a transmission channel. We have devised a technique that uses the motion energy in a video scene to non-intrusively measure the video delay to an accuracy of one NTSC field (1/60 second). The video delay measurement technique can be applied to any transmission channel, even when the source and destination locations are separated by thousands of miles. The video delay parameter presented here is closely related to the video delay parameter that was presented in contribution T1A1.5/92-139.

The following is a description of the method of measurement for the video delay.

1. Non-intrusively sample the luminance portion of the source video signal and the destination video signal. The source video signal is the video signal that is input to the transmission channel and the destination video signal is the video signal that is output from the transmission channel. Either one of the following two methods for digitally sampling these signals may be used. (1) If the luminance signal is directly available, it may be sampled using D1 format (CCIR Recommendation 601-2). (2) If the NTSC signal is the only signal that is available, it may be sampled using D2 format, and a digital filter may be used to separate the luminance signal.

   Let $Y_s(t_n)$ represent the active portion of the source luminance field of video, sampled at time t_n. Let $Y_d(t_m)$ represent the active portion of the destination luminance field of video, sampled at time t_m. These luminance fields occur at the rate of approximately 60 times per second for NTSC video. For D1 sampling, the active video portion of the luminance field will comprise 720 horizontal pixels by 243 lines. For D2 sampling, the active video portion of the luminance field will comprise approximately 764 horizontal pixels by 243 lines.

   Note: The displayed video portion of the luminance field (that part of the video that the viewer actually sees) may be used instead of the active video portion. This displayed video portion is typically only 93 to 94 percent of the active video portion.

2. Compute the absolute value of the field difference images ($|Y_s(t_n) - Y_s(t_{n+2})|$ and $|Y_d(t_m) - Y_d(t_{m+2})|$) as shown in Figure 5 and arrange the field difference images in the order shown.

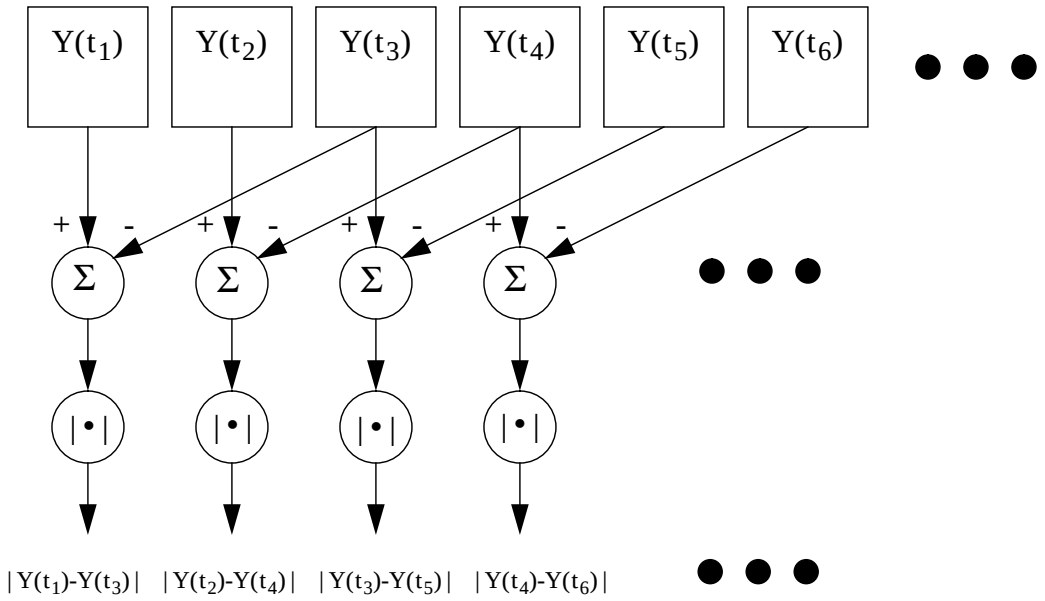

Figure 5 Computation of Field Differences

3. Compute the mean of the absolute value difference image. This will be represented by

$$m_s(t_n) = mean\left|Y_s(t_n) - Y_s(t_{n+2})\right| \tag{5}$$

for the source video and

$$m_d(t_m) = mean\left|Y_d(t_m) - Y_d(t_{m+2})\right| \tag{6}$$

for the destination video. The subtraction and absolute value functions in equations (5) and (6) are performed pixel by pixel, and the mean is computed using the resulting difference image pixels.

4. To measure the video delay at time n = N for the source video, form the time history vector for the source video defined by

$$\vec{m}_s(t_N) = [m_s(t_{N-K}), ..., m_s(t_{N-1}), m_s(t_N)], \tag{7}$$

which contains $m_s(t_N)$ and the previous K time samples. A set of time history vectors for the destination video is also formed, defined by

$$\vec{m}_d(t_m) = [m_d(t_{m-K}), ..., m_d(t_{m-1}), m_d(t_m)], \tag{8}$$

for every m such that $(max_{delay}+t_N) \geq t_m \geq t_N$,

where max_{delay} is the maximum expected video delay.

5. Compute the correlation function (c_m) given by

$$c_m = std[\vec{m}_s(t_N) - \vec{m}_d(t_m)], \tag{9}$$

where the standard deviation (std) is computed using the K+1 components of the

difference vector.

Note: The computation of the correlation function (c_m) should only include the source and destination time samples with non-zero motion energy.

6. The m = M that minimizes c_m is the alignment estimate and the video delay (v_{delay}) is computed according to the following equation:

$$v_{delay} = t_M - t_N. \tag{10}$$

Dynamic changes in the video delay of the transmission channel that occur slower than every K fields can be tracked by the parameter. When the video delay varies with time, an overall average video delay and histogram of the measured video delays should be computed.

The video delay method of measurement provides a robust method for non-intrusively measuring the video delay of a 45 Mbps transmission channel to an accuracy of 1/60 of a second. Figure 6 presents an example time history plot of $m_s(t_n)$ and $m_d(t_m)$ given by equations (5) and (6). Here, the source video ($m_s(t_n)$) is shown with a solid line and the destination video ($m_d(t_m)$) is shown with a dashed line. The x-axis for the plot gives the time in fields. The amplitude of the time histories quantify the amount of motion in the video scene at each time sample. As can be seen by examining the plot, the transmission channel has introduced added motion energy that may have resulted from added noise or a non-unity transmission channel gain.

The correlation function given by equation (9) is plotted in Figure 7. The x-axis for the plot gives the shift in fields for correct alignment of the source and destination waveforms shown in Figure 6. As can be seen from Figure 7, a sharp null in the correlation function for a shift of 0 fields indicates that the source and destination waveforms are properly aligned.

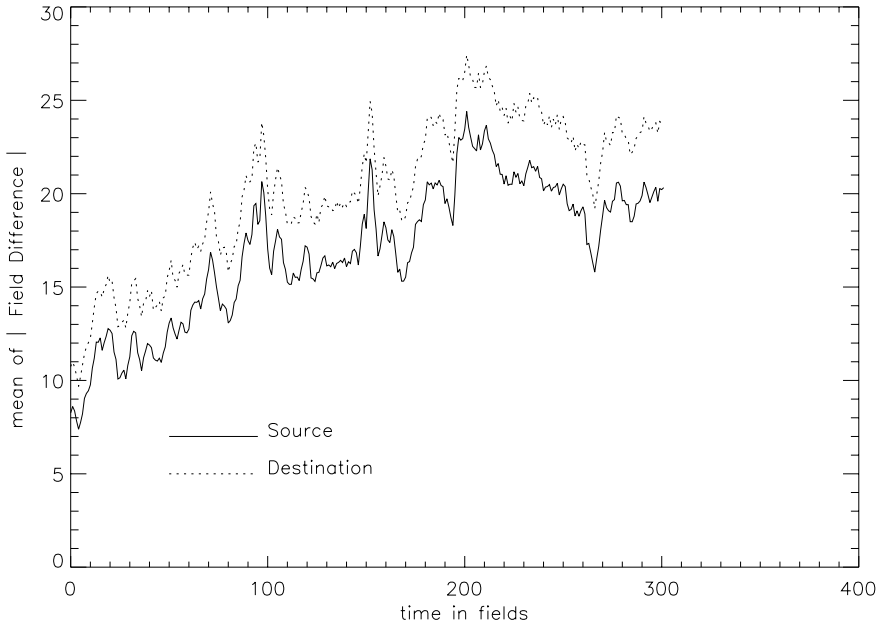

Figure 6 Aligned Time Histories of $m_s(t_n)$ and $m_d(t_m)$

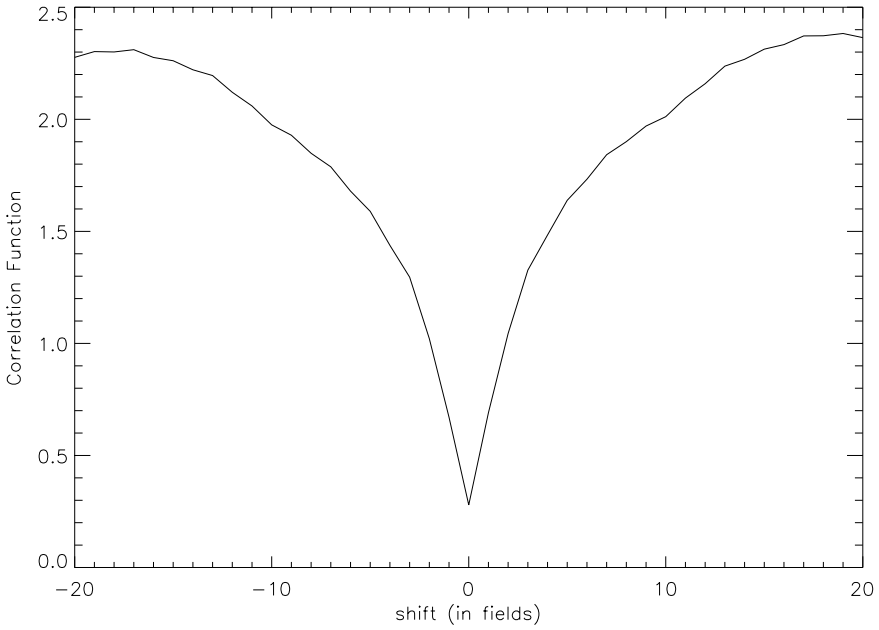

Figure 7 Correlation Function for Waveforms in Figure 6.

3.2 Average Fraction Change in Edge Energy (AFCEE)

The Average Fraction Change in Edge Energy (AFCEE) parameter measures the average fraction change in edge energy of the destination video with respect to the source video. Since an entire field of video is used, this parameter responds to global spatial blurring and added noise energy. Except for dropping a normalization constant of 5.78, the AFCEE parameter presented here is identical to parameter m_1' in contribution T1A1.5/92-135. The following is a description of the method of measurement for the AFCEE parameter.

1. Perform the method of measurement for video delay in section 3.1.

2. Next, $Y_s(t_n)$ and $Y_d(t_m)$ are filtered with the two 3 x 3 masks that are given in Figure 8. The filtering operation that detects horizontal edges is given by the convolution of the horizontal mask (H) with the luminance images, represented by $H*Y_s(t_n)$ and $H*Y_d(t_m)$. Likewise, the filtering operation that detects vertical edges is given by the convolution of the vertical mask (V) with the luminance images, represented by $V*Y_s(t_n)$ and $V*Y_d(t_m)$. The pseudo-sobel (PSobel) filtered luminance images combine the outputs of the two filtering operations as

$$PSobel(Y_s(t_n)) = |H*Y_s(t_n)| + |V*Y_s(t_n)|, \quad (11)$$

$$PSobel(Y_d(t_m)) = |H*Y_d(t_m)| + |V*Y_d(t_m)|, \quad (12)$$

where the absolute value and addition operations are performed on a pixel by pixel basis.

-1	-2	-1
0	0	0
1	2	1

H

-1	0	1
-2	0	2
-1	0	1

V

Figure 8 Horizontal and Vertical Edge Filters

Note: The pseudo-sobel filtering operations may be applied to the entire frame of video rather than just the individual fields of video. [Frame Sobel filtering was used to obtain the results in section 2]

3. Select $t_m = t_n + v_{delay}$, where v_{delay} has been found from step 1, and compute the standard deviation of the PSobel filtered images ($\sigma_{fs}(t_n)$ and $\sigma_{fd}(t_m)$) as

$$\sigma_{fs}(t_n) = std[PSobel(Y_s(t_n))], \quad (13)$$

and

$$\sigma_{fd}(t_m) = \frac{std[PSobel(Y_d(t_m))]}{G}. \quad (14)$$

Note: Normally the video gain, G, will be 1.0 and thus gain compensation may be ignored. However, if the transmission channel includes multiple passes through a video codec, then any non-unity gain of the video codec is greatly amplified and thus should be compensated for by including the G term in equations (14).

4. The AFCEE parameter is computed as

$$AFCEE = \frac{\left| rms_{time}(\sigma_{fs}(t_n)) - rms_{time}(\sigma_{fd}(t_m)) \right|}{rms_{time}(\sigma_{fs}(t_n))}, \quad (15)$$

where rms_{time} denotes the root mean square (rms) of the time history (nominally, 5 to 10 seconds of time history data are used).

3.3 Maximum Added Frame Noise Logarithmic Ratio (MAFNLR)

The Maximum Added Frame Noise Logarithmic Ratio (MAFNLR) parameter measures how much frame noise and/or motion distortion has been introduced into the video. The logarithmic ratio is used to produce an approximately linear relationship between the parameter and subjective quality. Since an entire field of video is used, this parameter responds to globally added noise and/or motion energy. The MAFNLR parameter presented here is identical to parameter m_3 in contribution T1A1.5/92-112, except that the MAFNLR parameter does not include the normalization constant of 4.2522. The following is a description of the method of measurement for the MAFNLR parameter.

1. Perform the method of measurement for video delay in section 3.1.

2. For the odd source fields, compute the standard deviation of the source field difference images ($Y_s(t_n)$ - $Y_s(t_{n+2})$) at time t_n, and the destination field difference images ($Y_d(t_m)$ - $Y_s(t_{m+2})$) at time $t_m = t_n + v_{delay}$. These will be represented by $\sigma_s(t_n)$ and $\sigma_d(t_m)$ and computed according to

$$\sigma_s(t_n) = std[Y_s(t_n) - Y_s(t_{n+2})] \quad (16)$$

$$\sigma_d(t_m) = \frac{std[Y_d(t_m) - Y_d(t_{m+2})]}{G} \quad (17)$$

$$n = 1, 3, 5, \ldots$$

where G the is the video gain of the transmission channel. The subtraction in equations (16) and (17) is performed pixel by pixel, and the standard deviation is computed over the resulting difference image pixels.

Note: Normally the video gain G will be 1.0 and thus gain compensation may be ignored. However, if the transmission channel includes multiple passes through a video codec, then any non-unity gain of the video codec is greatly amplified and thus should be compensated for by including the G term in equations (17). The entire frame of video may be used instead of just the odd fields in equations (16) and (17). [Frame differencing was used to obtain the results in section 2]

3. The MAFNLR parameter is then computed as

$$MAFNLR = max_{time}\{\log 10\left[\frac{\sigma_d(t_n)}{\sigma_s(t_m)}\right]\}. \quad (18)$$

where max_{time} is the maximum observed value over time (nominally, 5 to 10 seconds of time history data are used).

3.4 Maximum Added Line Noise Logarithmic Ratio (MALNLR)

The Maximum Added Line Noise Logarithmic Ratio (MALNLR) parameter measures how much line-oriented noise has been added to the video. The logarithmic ratio is used to produce an approximately linear relationship between the parameter and subjective quality. Since the added line noise parameter uses a single line of video, it responds to localized line-oriented disturbances in the image. The MALNLR parameter imposes additional quality penalties when line-oriented noise is present in the destination video. To increase the sensitivity of the measurement, the video line with the smallest amount of motion is used and the maximum line noise that has been added to this line is measured by the parameter. Apparently, T1Y1.1 viewers partially based their quality decisions on how much line-oriented noise was added to the stationary (non-moving) portion of the video scene. A possible reason for this may be that motion in the video scene masks the added line noise impairment. The following is a description of the method of measurement for the MALNLR parameter.

1. Perform the method of measurement for video delay in section 3.1.

2. From the absolute value of the source field difference images ($|Y_s(t_n) - Y_s(t_{n+2})|$) shown in Figure 5, compute the mean and standard deviation for each line of video in the odd fields. For line *l* of the odd source fields (n=1, 3, 5, etc), let $m_s(l, t_n)$ and $\sigma_s(l, t_n)$ represent this mean and standard deviation, respectively.

$$m_s(l, t_n) = mean|Y_s(l, t_n) - Y_s(l, t_{n+2})| \quad (19)$$

$$\sigma_s(l, t_n) = std|Y_s(l, t_n) - Y_s(l, t_{n+2})| \quad (20)$$

$$n = 1, 3, 5, ...$$

3. At time t_n for the source video, find the video line that has the minimum motion energy. This line will be called l_n and is determined according to the rule

$$m^2{}_s(l_n, t_n) + \sigma^2{}_s(l_n, t_n) \le m^2{}_s(l, t_n) + \sigma^2{}_s(l, t_n) \text{ for every } l \ne l_n. \quad (21)$$

Selecting the minimum motion energy line increases the sensitivity of the measurement.

Note: Although just one line of video was selected in this step, the number of lines that could be used for the MALNLR parameter is still a topic for research. One could apply the measurement to a horizontal region that was composed of several consecutive video lines.

4. Find the corresponding line l_n in the destination video at time $t_m = t_n + v_{delay}$

and compute $m_d(l_n, t_m)$ and $\sigma_d(l_n, t_m)$ according to

$$m_d(l_n, t_m) = \frac{mean\left|Y_d(l_n, t_m) - Y_d(l_n, t_{m+2})\right|}{G}, \quad (22)$$

and

$$\sigma_d(l_n, t_m) = \frac{std\left|Y_d(l_n, t_m) - Y_d(l_n, t_{m+2})\right|}{G}, \quad (23)$$

where G is the video gain of the transmission channel.

Note: Normally the video gain G will be 1.0 and thus gain compensation may be ignored. However, if the transmission channel includes multiple passes through a video codec, then any non-unity gain of the video codec is greatly amplified and thus should be compensated for by including the G term in equations (22) and (23). Also note that if vertical shifts of the field image due to the transmission channel are present, this condition should be corrected for in the selection of the corresponding line l_n in the destination video.

5. Compute the mean ratio for line l_n at source time t_n and destination time $t_m = t_n + v_{delay}$. This mean ratio is called $MR(t_n)$ and is given by

$$MR(t_n) = \frac{m_d(l_n, t_m)}{m_s(l_n, t_n)}, \quad (24)$$

where $m_s(l_n, t_n)$ and $m_d(l_n, t_m)$ are from equations (19) and (22), respectively. Also compute the standard deviation ratio for line l_n at source time t_n and destination time $t_m = t_n + v_{delay}$. This standard deviation ratio is called $SR(t_n)$ and is given by

$$SR(t_n) = \frac{\sigma_d(l_n, t_m)}{\sigma_s(l_n, t_n)}, \quad (25)$$

where $\sigma_s(l_n, t_n)$ and $\sigma_d(l_n, t_m)$ are from equations (20) and (23), respectively. These two ratios have proven to be extremely valuable for 45 Mbps video quality assessment. Both ratios are sensitive to added line noise and both ratios will be greater than 1.0 when added line noise is present in the destination video. However, the mean ratio $MR(t_n)$ is sensitive to time varying intensity shifts of the video line. Such shifts seem to be much more irritating than random, unstructured noise, which is measured by $SR(t_n)$. If $MR(t_n) > SR(t_n)$, then the added noise has line-oriented structure. It should be noted that this same technique could be modified to detect added noise with other structural shapes. For instance, if the transmission channel includes a video codec that uses blocks of size 16 pixels x 16 lines, then this block region could be used instead of the video line region. In this case, a maximum added block noise parameter could be computed.

6. Find the time $t_n = t_{max}$ where $MR(t_n)$ reaches its maximum value, i.e.,

$$MR(t_{max}) \geq MR(t_n) \text{ for every } t_n \neq t_{max}, \quad (26)$$

where 5 to 10 seconds of time history data are nominally used.

7. The MALNLR parameter is then computed as

$$MALNLR = \frac{1.0}{[1.0 + e^{(MR(t_{max}) - 10.0)}]} \log 10 \left[\frac{MR(t_{max})}{SR(t_{max})} \right]. \quad (27)$$

If MALNLR < 0.0, then set MALNLR=0.0

A few words of explanation are in order. We have found that if $MR(t_{max})>10$, this is indicative of a 45 Mbps system that has severe perceptual errors. For the T1Y1.1 video data, this included errors in the video picture that resulted from burst errors and bit error rates of 1.0×10^{-4} in the digital transmission channel. In this case, no further additional penalties are imposed by the MALNLR parameter (i.e., the parameter approaches 0). When $MR(t_{max})<10$, the MALNLR parameter imposes additional penalties for impairments which result from time varying intensity fluctuations (rather than just random noise fluctuations). This is because the ratio $[MR(t_{max})/SR(t_{max})]$ is greater than 1.0 for noise that has line-oriented structure.

4. Conclusions

The ITS video quality metrics have been applied to the T1Y1.1 DS3 video data. These video quality parameters have been used to measure the quality of the T1Y1.1 45 Mbps video to an accuracy of .2 to .4 root mean square (rms) quality units (on a scale from 1 to 5, where 5 is the best quality). It is important to note that the quality of a given 45 Mbps transmission channel depends upon the video scene that is being transmitted. Some test scenes can produce little or no impairment while other test scenes can produce a large amount of impairment. This effect is very significant and can result in subjective quality ratings that vary by more than two points over the 5 point CCIR-500 scale. One case was observed where the subjective mean opinion score varied from 2.16 to 4.64 for a fixed 45 Mbps transmission channel. Thus, these parameters should be applied directly to the video scene being transmitted by the contribution quality 45 Mbps channel.

An improved video delay parameter has been presented that can be used to measure the video delay of the transmission channel with an accuracy of 1/60 second. This is much more accurate than required from a perceptual standpoint.

Critical viewers that took part in the T1Y1.1 45 Mbps subjective tests seemed to base their quality judgements partly on the amount of noise that was added to the still background portion of the video scene. In addition, when this noise had structure (such as horizontal lines), the video scene was rated much lower in quality than if the same amount of noise was randomly distributed throughout the picture area. The MALNLR parameter presented in this contribution can be used to measure the perceptual quality degradation due to line-oriented noise.

The subjective quality judgements made by critical viewers on contribution quality 45 Mbps video systems contrasts with results obtained when low bit rate systems were rated by general viewers. Here, general viewers seemed to base quality judgements on distortions that were added to the motion portion of the video scene (such as jerky motion). Yet very similar measurements can be used to quantify both low bit rate and high bit rate distortions. The excellent results that have been obtained suggest that the ITS objective video performance parameters measure fundamental perceptual quantities and thus are applicable to a wide range of digital video systems. To account for the fact that critical viewers are using stricter quality criteria than general viewers, a technology independent quality assessment system with a viewer characterization switch might be desirable -- one switch position to emulate general viewers, and another for critical viewers.

5. Acknowledgments

ITS would like to express appreciation to PBS for providing D2 copies of the T1Y1.1 subjective viewing tapes and to Bellcore for providing an electronic copy of the T1Y1.1 subjective viewing results.