Video Testing in a DTV World

By David K. Fibush

This paper describes the digital television system, its elements, and appropriate methods of video testing. A layered approach is defined including video, protocol, and transmission channel testing. With today's use of video compression there is still a need for traditional analog and full bandwidth digital testing methods. Compression systems require objective picture quality measurements of distortions caused by its nonlinear and time-variant digital processing. Test method terms such as in-service and realtime are defined for both video and protocol testing. The three internationally accepted system approaches to objective picture quality testing are discussed including the advantages and disadvantages of each.

In the days of composite analog television systems, a very limited set of video signal types was used throughout the production and distribution chain. A modern digital television system is much more complex, often requiring conversion of the video signal into a variety of signal types including nonlinear compressed forms. A simplified block diagram of a modern television system is shown in Fig. 1.

Program content production primarily uses baseband analog and full bandwidth digital signals for ease of content manipulation. However compressed video using MPEG-2, Motion JPEG, or DV is widely used for disk and tape storage, providing excellent quality results in a cost efficient manner.

Following program production, the television signal may be compressed for storage, efficient transmission, or intrafacility interconnection in digital form. Typically this will be MPEG compression resulting in an MPEG transport stream (MTS) which is then multiplexed with other MPEG transport streams for transmission or interconnection. Statistical multiplexing is often used for multiple programs providing the most effective use of the available digital bandwidth.

The broadband telecommunication system provides a variety of transmission methods. Traditionally these have been voice-channel oriented with special data mapping for digital television signals. Use of asynchronous transfer mode (ATM) hides such formatting from the user. In order to provide the quality of service expected for broadcast operation, forward error correction (FEC) is generally required. A number of different modulation methods are used for radio frequency (RF) transmission over cable, satellite, and terrestrial systems. FEC is always used to provide the required quality of service RF transmission.

At the receive end of a transmission system, the desired program is demultiplexed from the MTS and the program data is decompressed. One of the advantages of compression is the capability of providing different video picture quality levels based on bit rates. Distribution quality to the home may be adequate with bit rates of 2 to 5 Mbits/sec for standard definition television (SDTV) and 15 to 19 Mbits/sec for high-definition television (HDTV). Contribution quality is required if further production processing is planned for the received signal. In that case, bit rates of 18 to 50 Mbits/sec are required for SDTV and 100 to 300 Mbits/sec for HDTV.

Presented at the 34rd SMPTE Advanced Motion Imaging Conference (paper no. 33-S-3), in San Francisco, CA, February 3-5, 2000. David Fibush is with Tektronix Inc., Beaverton, OR.

Video testing in the digital television system is not just a matter of developing new techniques for the difficult nonlinear compression process. A significant portion of the system still uses analog and full-bandwidth digital signals requiring application of traditional analog and recently developed digital test methods.

Testing Layers

There are three major testing layers as shown in Fig 2. Baseband video consists of such signals as traditional analog NTSC or PAL video; component digital SDTV, as defined by ITU-R BT.601; and digital HDTV, as defined by ITU-R BT.709. Video quality is measured in two ways, signal quality and picture quality, as described next.

At the intrafacility level there are two aspects to consider: compression method and physical interconnect. MPEG-2 is one compression method within a program provider facility; however, it is commonly used in at least two forms: elementary streams or transport streams with PES streams and program streams additional possibilities. DV is another common compression method for both SDTV and HDTV program production. At this layer it is the formatting protocol of the compressed data and its mapping into the physical interconnection that are to be tested. Certainly the electrical operation of the interconnection must be maintained as well. However, this is usually accomplished with sufficient headroom to provide a completely error-free environment for all intrafacility interconnections. Common physical interconnections are:

- Serial digital interface (SDI) for both SDTV[1] and HDTV.[2]
- Serial data transfer interface (SDTI)[3] using the same physical interface as SDI.
- SMPTE 310M[4] a robust synchronous serial interface for ATSC transmitters.
- Asynchronous serial interface (ASI)[4] operating at the same bit rate as SDTV SDI but using a different channel code.

Broadband digital networks and RF methods are used for interfacility transmission. Traditionally various proprietary and standard RF methods with powerful error correction have been used. More recently mapping the compressed data through an ATM layer or directly into PDH or SDH transport protocols with or without error correction is becoming more common. DVB[6] and ATSC[7] systems are most com-

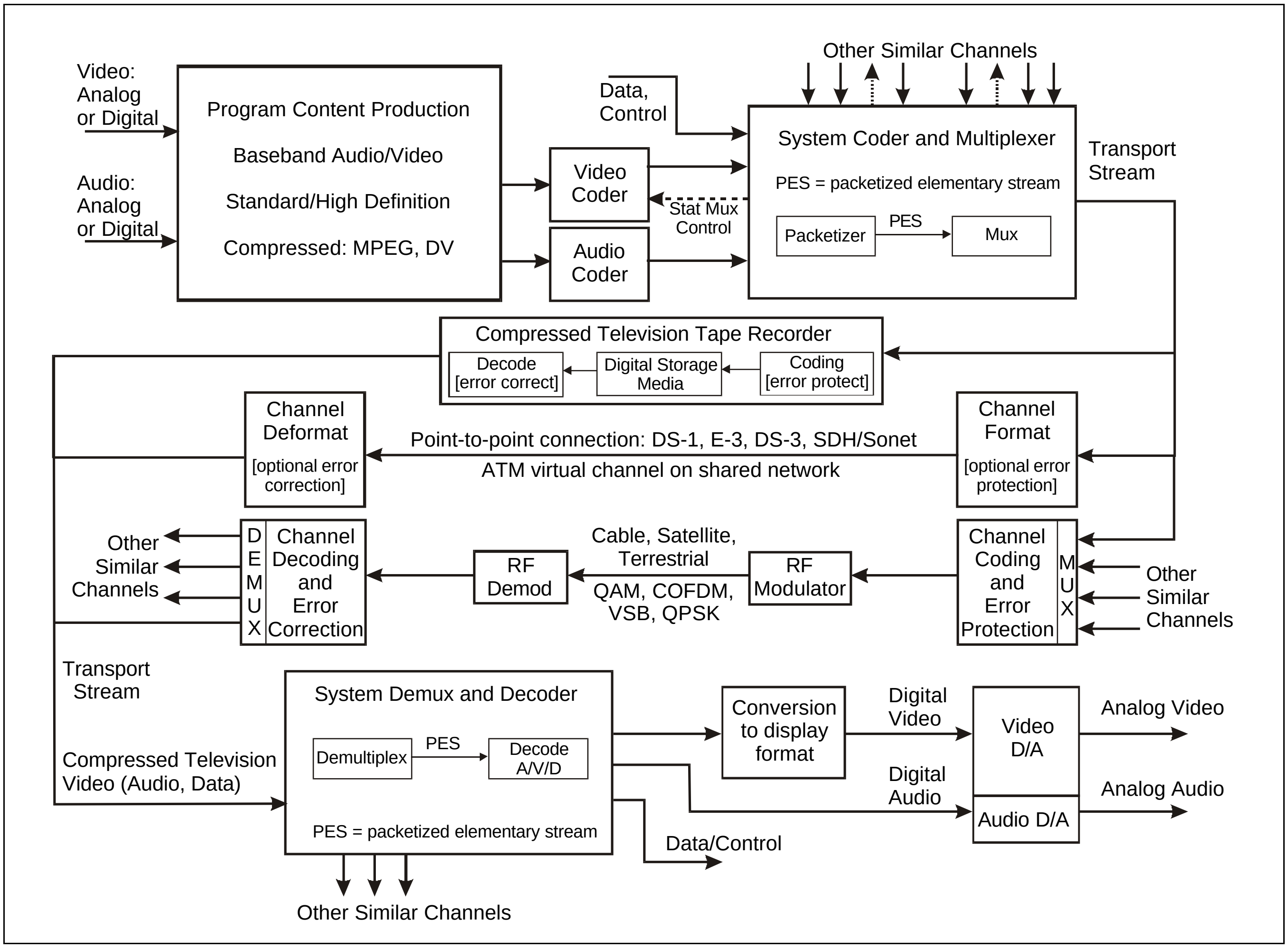

Figure 1. Digital television simplified block diagram.

monly used for consumer television terrestrial, satellite, and cable applications. They are both based on MPEG-2 data with different forms of information tables and RF modulation methods.

Video testing requirements for each application in the production and distribution of digital television programs is based on the testing layers appropriate for its part of the system. Program providers will be most interested in video quality and intrafacility interconnection operation, while transmission service providers will be most interested in error conditions of the channel. However, video quality parameters are important to all users as the ultimate test of providing the program to the final viewer. Although a broadband service provider might be most interested in error rates and ATM cell loss ratios, it is the effect they have on video quality that will be the topic of discussion with the ad agency that paid for the commercial. Similarly the program content provider will discuss guaranteed cell loss ratios when selecting a provider for interfacility transmission for a distributed studio design.

Video Test Method Terms

Video systems are intended to display a picture that accurately represents the scene being scanned by the camera. With today's special effects capabilities, the displayed pictures may differ in an artistically defined manner from the original. Nonetheless, at some point in the processing, where the artistic changes are complete, the resulting pictures are to be accurately conveyed to the user with the desired quality. There are a number of terms that are used to describe video test methods:

Signal quality (indirect)	Picture quality (direct)
In-service	Out-of-service
Realtime	Non-realtime (deferred time)
Continuous	Sampled (or scanned for multiple programs)

It is important to note that each pair of terms is independent of the other pairs. As an example, in-service testing could use either direct or indirect methods, and it could be either real or non-realtime. Realtime tests may be either continuous or sampled.

Analog video and uncompressed digital video are evaluated using traditional signal quality tests. These are indirect with respect to the pictures passing through the system. That is, they measure channel response to a series of different high-quality test signals. Indirect tests are based on the premise of a time invariant linear system. Video distortions

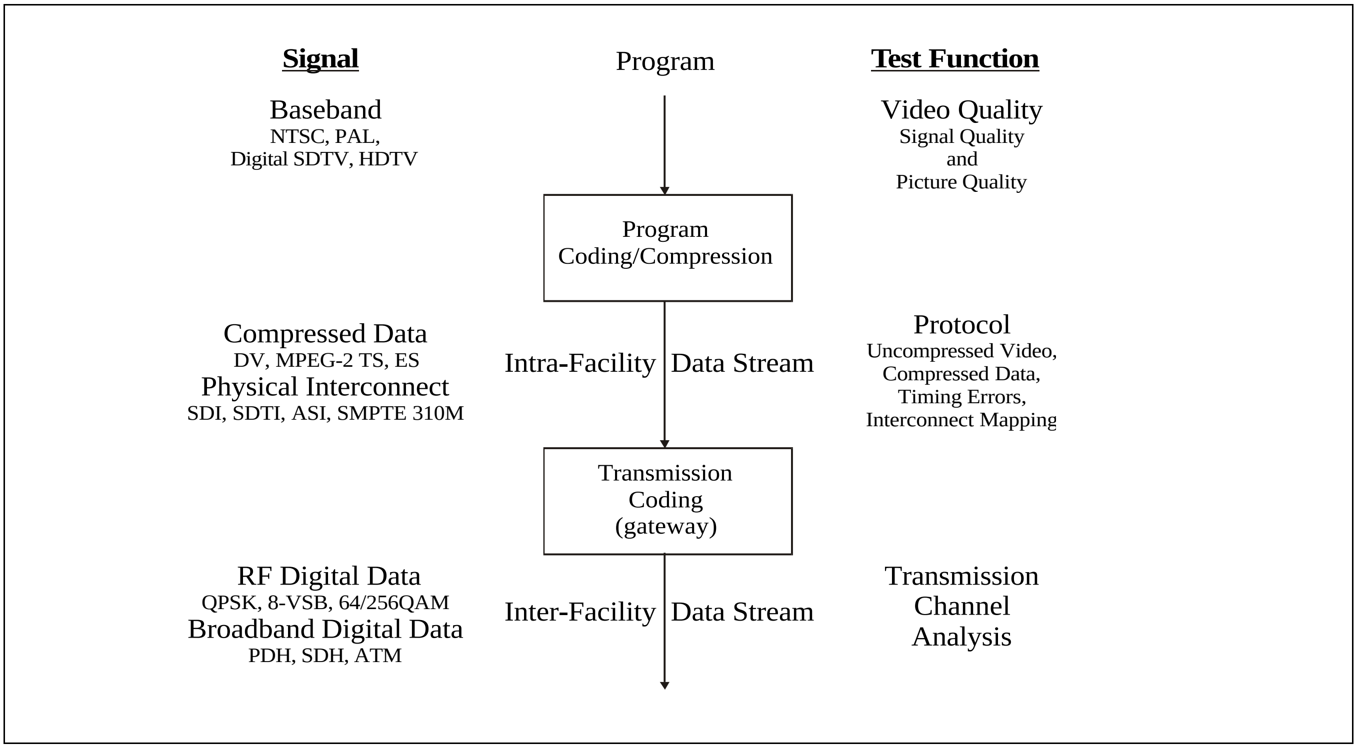

Figure 2. Digital television testing layers.

produced may be accurately determined by passing suitable test signals through the same system.

Digital compression systems are nonlinear, hence the resulting video quality will be a function of the picture content and other time varying attributes of the system (e.g., statistical multiplexing). Test signals are easily compressed, hence their changes through a system are not a meaningful measure of video quality. Therefore, direct testing of picture quality is required in addition to the signal quality measurements for the linear parts of the system. Direct evaluation of the system is based on quality changes between the video at the selected test point and the video at the source. The source video may be either a variety of defined (preferably standard) program-like picture sequences or general program material.

Note: the term “picture quality” is actually picture degradation through the system and does not imply the intrinsic quality of the picture. Beauty is in the eye of the beholder and not something we can measure.

In-service tests are made while the program is being displayed, directly by evaluating the program material, or indirectly for linear systems, by including vertical interval test signals with the program material. Out-of-service, appropriate test scenes are used for direct tests (picture quality) and full field test signals are used for indirect tests (signal quality).

Realtime testing provides results shortly after the video has been processed by the system where “shortly” is determined by the user requirements for the test. An example would be differential phase and gain measurements averaged over several fields using a Tektronix VM700T. For many applications, such as MPEG-2 transport stream protocol, verification, or high accuracy picture quality measurement, deferred time testing is appropriate to allow a more detailed analysis from data stored on disks.

In the application of realtime testing, it may be necessary to use a sampled approach where only certain time increments or samples of the data are analyzed due to the computing power required. This might also be used where several programs are scanned one after the other providing a sample of data for that program. Testing methods that are less complex can be accomplished in realtime based on a continuous stream of data.

Analog Video Testing

Measurement practices for analog television systems are well defined and very reliable.[8,9] They are based on the linear system premise that video distortions produced in a system may be accurately determined by passing suitable test signals through the same system. However, a series of different tests are required to verify that the system will produce acceptable pictures. One test will not do the job.

A minimum set of tests for an analog system might be: luminance amplitude, rise/fall times, bandwidth, group delay, signal-to-noise ratio, and waveform nonlinear distortions. For composite systems (PAL/NTSC): chroma amplitude, chroma phase, differential phase, and differential gain are also important. A matrix of tests and specifications can be associated with different qualities of video signal transmission, but no one number or mathematical combination of parameters will describe the resulting system operation. If adjustments are made, there is quite often interaction between processing elements resulting in changes of several measurement parameters. A television engineer must evaluate the combined results of the various tests as well as observe the resulting picture.

Full-bandwidth Digital Video Testing

Even with the rapid rise in the use of compression techniques, full-bandwidth digital television will continue to be a major element in television systems. Processing and interconnecting video in digital form eliminates most signal quality loss problems of an analog system. As the television industry transitions from analog to digital video, combinations of composite and component systems are being used; however, the preference is for all-component video operation.

For operational purposes, monitoring of equivalent analog video signal properties is based on a signal processed from the digital data. Where testing of the analog signal required only that various parameters be measured on a single waveform digital testing requires three types of analysis: digital signal coding, digital data formatting, and parameters relating to the digital waveform. Although all three measurement types can be performed with a single instrument, such as the Tektronix WFM 601M, there is significant processing between each pair of layers with different analysis methods for each layer as well.

Correct coding of the signal, with respect to the standard, can be determined by a waveform display of the resulting analog signal and a readout of corresponding digital values in defined areas such as black, peak white, zero-color, maximum color excursions, and non-use of certain excluded digital values. Digital formating of the nonvideo part of the signal can be quite complex particularly when digital audio is embedded in the digital video signal. Measurement includes displaying an analog rendition of the formatting digits, cursor selection of each sample for a readout of the digital value, as well as a logic-analyzer type display of a selected group of digital words.

Included in the digital formatting are error detection cyclic redundancy check (CRC) words that can be used in monitoring or operational equipment to determine if the digital signal at a receiver exactly matches that of the sending equipment, where the CRCs were inserted. Called EDH, for error detection and handling, this method provides an accurate way to determine the onset of errors when a stressing method is used to measure operational headroom in the virtually error-free environment of serial digital video.

Analysis of the serial digital waveform (which is really an analog signal) is, perhaps, the most complex of the three measurement types. An eye pattern display is used to measure basic parameters such as amplitude, rise time, and overshoot. The eye pattern can also measure jitter if the clock used to produce the eye pattern is carefully specified. In a typical serial digital television waveform monitor, eye pattern determination of jitter would require setting of cursors to make the measurement; however, numeric measurement values may be quickly obtained using a demodulator method. In many cases operational headroom for a serial digital video signal is simply related to the amount of coax in the transmission path and the amplitude of the launch signal. A measurement of these parameters can be derived at the receiver end based on an assumed type of coax thereby providing some important headroom information.

Picture Quality Test Methods

Television programs are produced for the enjoyment or education of human viewers so it is their opinion of the quality that is important. Formal subjective tests as defined by ITU-R BT.500[10] have been used for many years. With the advent of digital video compression, the number of different test methods in BT.500 have increased every year.

Advantages of subjective testing are: a test may be designed to accurately represent a specific application; valid results are produced for both conventional and compressed television systems; a scalar mean opinion score (MOS) is obtained; and a wide range of still and motion picture applications are accommodated.

Weaknesses of subjective testing are: a wide variety of possible methods and test element parameters must be considered; meticulous setup and control are required; many observers must be selected and screened, and complexity makes it very time consuming.

The result is subjective tests are only applicable for development purposes; they do not lend themselves to operational monitoring, production line testing, trouble shooting, or repeatable measurements required for equipment specifications.

The need for an objective testing method of picture quality is clear, subjective testing is too complex and provides too much variability in results. However, since it is the observer's opinion of picture quality that counts, any objective measurement system must have good correlation with subjective results for the same video system and test sequences.

With a new paradigm such as compressed digital video, subjective test methods were used before objective test methods became available. Certainly subjective test methods are the starting point for evaluating objective test methods. However, it is an objective test method that provides accurate measurements, with proven correlation to subjective assessments, which will be the benchmark for development of test materials and calibrating less capable objective methods.

There is general agreement in the industry that there are three methodologies for objective picture quality measurement that provide three levels of measurement accuracy.[11,12] They are identified as:

• Complete source video (also called Picture comparison and mow officially known as Full Reference).

• Reduced source video information (also called Feature extraction and now officially known as Reduced Reference).

• No source video (also called Single-ended and now officially known as No Reference).

The first two methods are double-ended, that is the actual source video, an exact copy of the source video or some reduced information extracted from the source video must be available to the instrument making the picture quality calculations. (Note: Source video is sometimes called reference video. In this paper the term "source" is used since "reference" might imply a defined video sequence whereas quality will often be determined based on program video.)

Picture comparison makes a measurement of picture quality (degradation) using the full source video and the video processed by the system under test as shown in Fig. 3. It uses a matrix-based mathematical computation to process each picture or sequence of pictures. The resulting data represents a filtered version of the pictures containing an amount of data similar to the original pictures. Typically, the pixel-by-pixel difference between filtered versions of the source and degraded pictures is used to determine an objective quality score. This is the most accurate method because it has complete information about the changes in the pictures and generally uses a very sophisticated computation algorithm based on a model of the human vision system including temporal and color response.

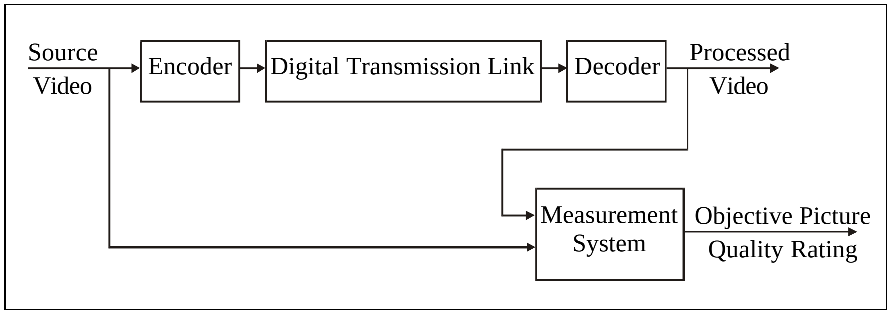

Figure 3. Picture comparison test method.

Knowledge of compression or other processing applied to the video is not required. Although a primary application of this measurement method is for codec evaluation using standard test sequences it can be used in-service for monitoring of statistical multiplexer operation or at a remote location if a copy of the source material is available. Realtime continuous operation is not precluded although significant compute power is required.

It is the sophisticated processing algorithm that gives the picture comparison method an accuracy superior to the other two methods. Because the complete source video is available another less sophisticated calculation can be made. Peak signal-to-noise ratio (PSNR)[13] has traditionally been used for evaluating differences between a source and processed picture. Although it is known to produce inconsistent results on sequences with different complexity of spatial and temporal pictures, it is a useful benchmark. It has been shown to have good accuracy for pictures with low degradation and can be used to locate small differences between pictures that would not necessarily be seen by a human observer.

Unfortunately, the convenience of performing measurement level picture quality evaluation anywhere in the system, albeit by indirect means, is not available for compressed systems as it is for uncompressed linear systems. There are two approaches to this situation.

First, since the video is being carried throughout the system as digital data, if the bits don't change and the timing of the transmission is held within given limits, the video and its quality will not change either. Intrafacility transmission links are error free if operating properly.

For broadcast applications, interfacility transmission links will use forward error correction (FEC) to provide the same error free operation. Even video conferencing transmission links are expected to have quite low error rates and use sophisticated error concealment.[14] In this case, measurement of picture quality at the original encoding location provides the necessary quality information. For satellite transmission, the link can be included in the measurement by providing the picture comparison measurement instrument with downlink video as the processed video at the uplink location.

The second approach is the use of known video sequences. Application of standardized reference video test sequences[15] allows out-of-service picture comparison evaluation of system performance at a location remote from the input video source. Standardized video sequences or any predetermined video sequence of the user's choice (perhaps

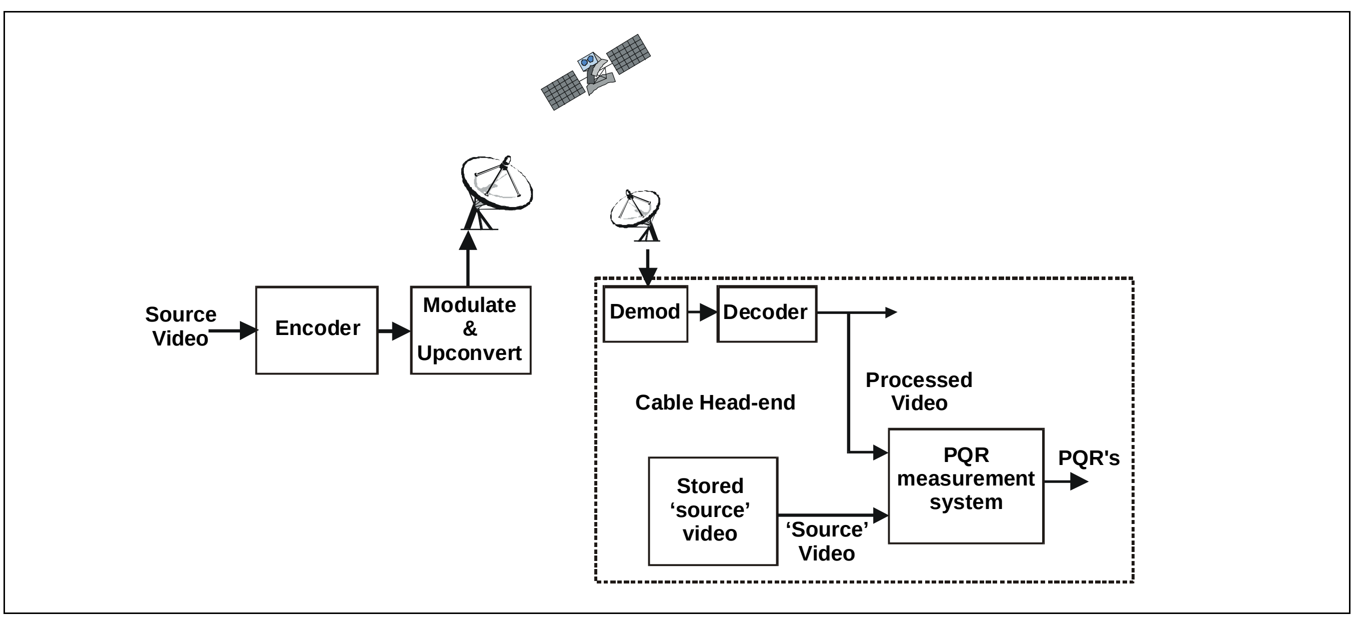

Figure 4. Example of remote PQR measurement.

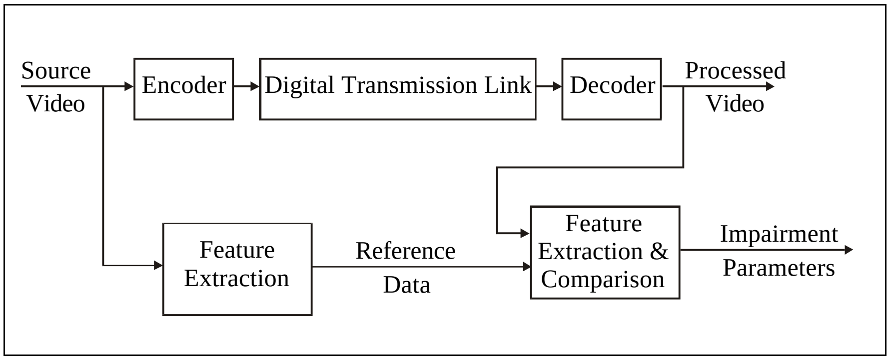

Figure 5. Feature extraction test method.

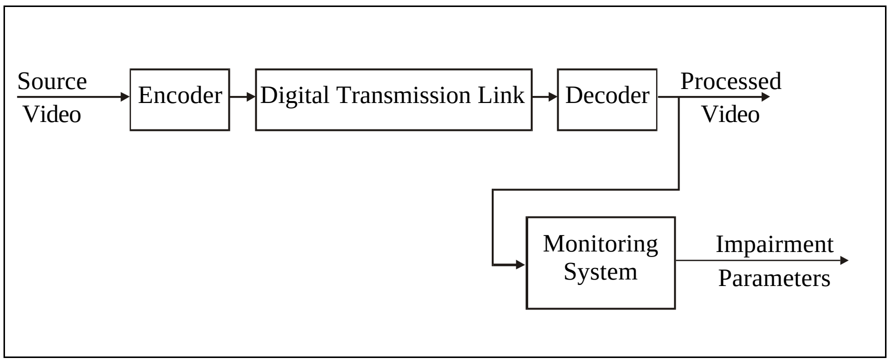

Figure 6. Single-ended test method.

a show-opening moving graphic) can also be used as the "source" video for the measurement. An example is measurement of picture quality rating (PQR) for a cable headend (Fig. 4). Another very important benefit of using standardized reference video sequences is that they enable the manufacturer and purchaser of equipment for compressed digital television to understand and measure performance specifications using common material.

Feature extraction (Fig. 5), uses a mathematical computation to derive characteristics of a single picture (spatial features) or a sequence of pictures (temporal features). This usually results in an amount of reference data per picture that is considerably less than used to transmit the compressed picture. The calculated characteristics of the source and degraded pictures are then compared to determine changes in the pictures related to the type of features extracted. This provides a measure of picture impairments. In this low reference data mode, incorrect video, lost frames, and other types of errors not necessarily due to the compression process, can be detected.[11] To facilitate application of the feature extraction method, a special MPEG-2 transport stream PID (packet identification) has been defined by the DVB to carry the reference data.

By using a relatively high level of reference data, perhaps as much as the compressed picture, a useful objective picture-quality calculation can be made. Knowledge of compression or other processing applied to the video may be used to determine what features are to be extracted in order to increase the measurement accuracy of the calculation. Because bits equal cost, applications for feature extraction are primarily for monitoring inconsistencies of pictures sent and received through a transmission system as described above.

Single-ended testing is the most appropriate method for picture monitoring (Fig. 6). Based on knowledge of the compression system being used, the processed video canalyzed for artifacts and other defects. The most common artifact to be detected is blockiness, a result of the discrete cosine transform (DCT) compression system used by MPEG-2 and DV. As the compression system works harder, due to either more complex program material or less bits available for data, more blockiness will be generated. The higher the level of blockiness the greater the picture degradation for a given source video sequence.

Much like PSNR, this method does not provide strong correlation with subjective picture-quality assessment over a variety of video sequences. Nonetheless it has some significant operational advantages. Since no information from the source video is required, the single-ended monitor may be placed anywhere in the system. Detection of picture defects is not limited to compression artifacts; a measure of gaussian noise would be an example. This could be useful at a compression encoder input to increase coding efficiency and resulting picture quality.

Compute power required for the test is modest allowing an economic approach to multichannel operation. Artifact level detection can be used as a trigger to warn of possible quality problems. A log of detected artifacts versus time can be a valuable system trouble-shooting tool.

Protocol Test Methods

Intrafacility transmission links require several different sublayers of testing and the extended application of digital analysis techniques. Although there are a variety of interconnection methods and types of data to be carried, each is defined in specifications developed by standards organizations. While protocol testing is the primary type of analysis to be performed there are other sublayers to be considered:

1. Protocol analysis of formatting and compliance to standards of uncompressed video and included ancillary data. Compressed programs in such formats as MPEG-2 elementary streams (ES), transport streams (TS), and program streams (PS); DV-based data streams; and physical layer data mapping of SDTI and ASI.
2. Physical layer electrical characteristics, waveform type specifications for SDI, ASI, and SMPTE 310M.
3. Timing errors due to interfacility transmission links: jitter and wander in received data timing information and itter and wander in baseband video resulting from the timing information.
4. Response to custom data streams.

There are a number of terms are used to define the test methods. They are similar to those used for video testing but are restated here as applicable for protocol testing:

Realtime	Non-realtime (deferred time)
Continuous	Sampled (or scanned for multiple programs).
In-service	Out-of-service

Realtime testing is primarily a monitoring function where key characteristics of one or more data streams are analyzed to determine if they meet required specifications. Deferred-time testing is implemented by storing a data stream and generally performing an extensive in-depth protocol analysis. Realtime testing may include timing analysis, whereas the storage of the data for deferred-time testing usually does not provide any timing information.

Continuous testing is the preferred method for realtime monitoring. This means that every set of data, such as TS packets, are analyzed to a level of detail allowed by the compute power available. Sampled testing would only analyze a selected subset of data. This could be due to a lack of compute power to analyze every packet or due to scanning of multiple channels where increments of time for analysis are allocated to each channel.

In-service tests have no specific limitations as are incurred with video quality testing. The data may be analyzed while programs are being transmitted at any location in the system and in a single-ended manner. The only limits relate to compute power for realtime and continuous monitoring as discussed above. Out-of-service testing is also an important aspect for maintaining DTV systems but the emphasis is on custom streams, such as TS, to determine system response to known valid or specifically invalid data sets.

Conclusion

This paper has given an overview of video test methods including protocol analysis of the compressed data stream. Traditional test methods continue to be important for the analog and full bandwidth portions of today's hybrid systems of compressed and uncompressed video. The system approach to a number of objective picture quality and data protocol test methods has been described. While individual test methods are important, it is their integration into a "Quality of Service" monitoring system that will provide the greatest overall benefit to maintenance of digital television systems.

References

1. Rec. ITU-R BT.656-4, "Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601 (Part A)."
2. Rec. ITU-R BT.1120-2, "Digital Interfaces for 1125/60and 1250/50 HDTV Studio Signals."
3. Rec. ITU-R BT.1381, "SDI-based Transport Interface for Compressed Television Signals in Networked Television Production Based on Rec. ITU-R BT.656 and ITU-R BT.1302."
4. SMPTE 310M, "Synchronous Serial Interface for MPEG-2 Digital Transport Streams, *SMPTE J., 107:*274, Apr. 1998.
5. DVB A010, "Interfaces for CATV/SMATV Headends and Similar Professional Equipment."
6. DVB standards are published by ETSI, go to http://www.etsi.org and access "Publications and Products" then "Publications Download Area" then search for DVB.
7. http://www.atsc.org.
8. *Television Measurements for NTSC Systems,* Tektronix Inc., Beaverton, OR, 1994, Lit No. 25W-7049-3.
9. *Television Measurements for PAL Systems,* Tektronix Inc., Beaverton, OR, 1994, Lit No. 25W-7075-1.
10. Rec. ITU-R BT.500, Methodology for the Subjective Assessment of the Quality of Television Pictures.
11. J. Baina and G. Goudezeune, "Equipment and Strategies for Service Quality Monitor-ing for Digital Television Networks," *SMPTE J., 108:*632, Sept. 1999.
12. ITU-T Study Group 9 Draft Rec., *J.OVQ,* April 1999.
13. ANSI Standard T1.801.03-1996, "Digital Transport of One-way Video Signals, Parameters for Objective Performance Assessment."
14. Draft Standard on Quality of Service for Business Multimedia Conferencing, T1A1.5/99-201, www.t1.org/index/0629.htm.
15. Rec. ITU-R BT.802-1, "Test Pictures and Sequences for Subjective Assessments of Digital Codecs Conveying Signals Produced According to Rec. ITU-R BT.601." (Note: The sequences are available from SMPTE.)

The Author

David Fibush has been involved in television engineering for more than 30 years. He was a major contributor to standardization of the SMPTE Type C format and development of television recorders for various tape formats.

At Tektronix, Fibush was involved in the development of test signal generators and video frame synchronizers and the planning of compressed television testing products. Presently, he is a consultant to Tektronix in the area of video quality testing for compressed television systems.

Fibush is a SMPTE Fellow and Chairman of the Committee on Video Compression Technology. In 1995, he received the SMPTE Progress Medal, which recognizes outstanding technical contributions to the progress of engineering phases of the motion picture and/or television industries.