

UIT - Secteur de la normalisation des télécommunications

ITU - Telecommunication Standardization Sector

UIT - Sector de Normalización de las Telecomunicaciones

Commission d'études ;Study Group;Comisión de Estudio} 12

Contribution tardive;Delayed Contribution ;Contribución tardía} **999**

Texte disponible seulement en ;Text available only in;Texto disponible solamente en} English

Question: 11/12, 10/12

SOURCE¹: Rapporteur

TITLE: Proposed Definitions for Video Classes and Video Terms from the VQEG
(Video Quality Experts Group)

Abstract

This contribution presents a set of definitions that have been drafted by members of the ITU VQEG (Video Quality Experts Group) ad hoc committee on video classes and definitions. These definitions are offered to the participating ITU Study Groups (ITU-T Study Groups 9 and 12 and ITU-R Study Group 11) for further review and comment. It is expected that the definitions of terms and video classes will be included in new Draft Recommendations in the area of video quality.

¹ Contacts: Arthur Webster, Rapporteur
Tel: +1 303 497 3567
Fax: +1 303 497 5323
E-mail: webster@its.bldrdoc.gov

David Fibush, Editor
Tel: +1 503-627-6289
Fax: +1 503-627-1707
E-mail: dave.fibush@bangate.tek.com

Introduction

The following definitions of terms and video classes represent an effort to standardize terminology in the area of video quality. Most terms are not new and some are included in previous Recommendations. The Video Classes are newly formulated, however, and attempt to divide the very large area of video (excluding HDTV) into seven categories. The division into these categories will help clarify and define important sections in ITU Recommendations in the area of video quality.

Video Classes Table and Definitions

Video classes are defined to provide an overview of applications whose range also implies levels of possible degradation. In broadcast applications, classes TV0 – TV3, it is the expected further processing of the decompressed video that is used to define the class. Video conferencing, internet video and related applications are covered by classes MM4 – MM5 where delivery parameters such as frame rate or raster size are used to define the class. Class MM6 is still pictures for a variety of applications with characteristics somewhat orthogonal from the rest of the list.

Table 1 Definitions of Classes

| | |
|------|---|
| TV 0 | Loss-less: ITU-R Recommendation 601 video used for contribution and primary distribution applications without compression. This is the reference source for subjective and objective measurements. |
| TV 1 | Contribution: Used for complete post production, many edits and processing layers, intra-plant transmission. Also used for remote site to plant transmission. |
| TV 2 | Primary distribution: Used for simple modifications, few edits, character/logo overlays, program insertion, and inter-facility transmission. A broadcast example would be network to affiliate transmission. An example in a cable system would be a regional downlink facility to a local headend. |
| TV 3 | Secondary distribution: Used for delivery to home/consumer (no changes). An example in a cable system would be from a local headend to a home. |
| MM 4 | All frames encoded. Low Artifacts relative to videoconferencing applications. Usually ≥ 30 fps. |
| MM 5 | Frames May be Dropped at Encoder. Perceptual artifacts possible, but quality level useful for designed tasks. |
| MM 6 | Series of Stills. Not Intended to provide full motion. (Examples: Surveillance, Graphics) |

Attributes of classes are presented to provide an idea of parameters typically encountered in today's applications. Values outside those listed are possible, particularly with respect to bit rate where improvements in encoding quality are continuously being made. This table includes all the columns in the original contribution to stimulate discussion and provide for future expanded description of the classes. Other changes have also been made to stimulate discussion.

Table 2 Attributes of Classes

| Class of Operation | Spatial Format | Coding Algorithm ¹ | Delivered Frame Rate ² | Typ.Latency ⁴ Delay Variable. | Channel Impairment | Nominal Bit Rate, Mbit/s |
|--------------------|----------------|-------------------------------|-----------------------------------|---|---------------------------------------|--------------------------|
| TV 0 | Rec. 601 | None | Max FR | (note 3) | None | 270 |
| TV 1 | Rec. 601 | MPEG-2 DV | Max FR | (note 3) | Infrequent Distortion ⁵ | 18 to 50 |
| TV 2 | Rec. 601 | MPEG-2 DV | Max FR | (note 3) | Infrequent Distortion ⁵ | 10 to 25 |
| TV 3 | Rec. 601 | MPEG-2 | Max FR Occasional frame repeat | (note 3) | Occasional ⁵ Distortion | 1.5 to 10 |
| MM 4 | Rec. 601 | H.262 | (15-) 30 fps | Delay = low var \approx 0 | application dependent ⁶ | 1.5 |
| MM 4 | CIF | H.263 | ~30 fps | Delay = low var \approx 0 | application dependent ⁶ | 0.768 |
| MM 5 | CIF | | 10-30 fps | Del \approx 400ms var \geq 100ms | application dependent ⁶ | |
| MM 5 | QCIF | | 1-15 fps | Del \approx var \geq 200ms | application dependent ⁶ | |
| MM 5 | Sub-QCIF | H.263 | >0 fps <30 fps | Del \approx var \geq | application dependent ⁶ | |
| MM 6 | 16CIF | H.263 | | > 1 sec | application ⁶ dependent | |

Note 1: Proprietary coding algorithms are also available. DV is seen as being used for TV1 and TV2 but not TV3. DV video is a DCT type compression system using only intra-frame compression designed to optimize storage applications such as tape recorders.

Note 2: Normally 30 fps for 525 systems and 25 fps for 625 systems

Note 3: Channel impairments will be expressed in terms of the packet/cell loss ratio, the errored packet/cell ratio, or an appropriate bit error specification.

Note 4: Broadcast systems all have constant, but not necessarily low, latency and constant delay variation. For broadcast applications latency will be low, say 0.045 or 0.09 sec, for multi-source news to high, say >1 sec, for TV3. MM4 systems have a low and constant latency with low or zero delay variation.

Note 5: "Infrequent Distortion" in the TV classes indicates that transmission impairment is considered to mean the system is broken and needs repair. "Occasional Distortion" indicates that if transmission impairment is present, it will cause video distortion not objectionable to the program viewer. Although this is somewhat subjective, this clearly excludes cases of continuous (or near-continuous) video distortion.

Note 6: Multimedia classes generally have some distortion in the error-free compression/decompression process and additional distortion if there are transmission errors. Video carried by higher bit rates usually has less distortion, however this is very much a factor of the complexity of the scene and action. Acceptable levels of distortion vary widely for different applications.

Terms and Definitions

Source sequence: Input motion or still sequence to a system under test. Generally a high quality sequence with a desired set of attributes.

Processed sequence: Output sequence from a system under test. Usually a degraded version of the source sequence. In some unusual instances the output sequence can actually have better quality than the input sequence.

Video quality: Difference in characteristics between the source and processed sequence. Consists of two parts, signal quality and picture quality.

Signal quality: Difference in measured parameters of a test signal between the input and output of a linear (non-compressed) part of a system under test. This is an indirect measurement of picture quality as the changes in the test signal are used to infer the difference in characteristics of source sequences passing through the same system. Typical parameters and test signals are defined in documents such as EIA/TIA 250C.

Picture quality: Difference in objectively or subjectively measured characteristics between the source and processed sequence. Typical characteristics are defined in documents such as ANSI T1.801.02 and an ITU contribution to be specified.

Intrinsic picture quality: Picture quality measured without an apparent reference. Called single stimulus in ITU-R Recommendation 500 subjective measurements, the implied reference is the observer's expectations for the application and comparison with other sequences in the measurement series.

Objective measurement: A measurement made by an instrument. For signal quality the measurement may be either manual by observation of a waveform on a calibrated instrument, or automatic by the instrument supplying numeric results. For picture quality the instrument provides a result that is correlated to the average expected from a series of human observers under the same measurement conditions.

In-service testing: Video quality measurements made on a system under test while the same processing channel is being used to provide program material for the intended application. This can sometimes be accomplished for signal quality measurements in systems that don't use compression by adding test signals to the vertical interval of the program material. The picture quality (and hence the in-service measurement result) is likely to be a function of program scene content (e.g. spatial information and temporal information content) as well as transmission channel operating characteristics that might change by taking the system under test out of service. An example might be bit rate allocations between different channels in a statistically multiplexed channel.

Out-of-service testing: Video quality measurements made on a system under test by supplying a defined (perhaps standard) source sequence or test signal(s) fully utilizing the channel to be evaluated.

Video

- (1) The visually displayed images of television or video teleconferencing/video telephony.
- (2) A signal that contains timing/synchronization information as well as luminance (intensity) and chrominance (color) information that when displayed on an appropriate device gives a visual representation of the original image sequence.
- (3) Of or pertaining to visually displayed images of television or video teleconferencing/video telephony.

Video frame

One complete scanned image or picture from a set comprising video imagery. A video frame is usually composed of two interlaced fields.

Video imagery

A sequence of video frames.

Video teleconferencing/video telephony service (VTC/VT)

The transmission of video signals capable of portraying motion and the accompanying audio signal(s) between two or more locations using bi-directional transmission facilities. Both analogue and digital transmission may be used. A typical example of this service is interactive video teleconferencing between groups of personnel located at two or more locations.

Motion video

Temporally varying visual imagery intended to communicate or convey movement or change.

Frame repetition

A condition during the display of motion video wherein identical video frames are repeated.

Spatial perceptual information (SI)

A measure that generally indicates the amount of spatial detail of a picture. It is usually higher for more spatially complex scenes. It is not meant to be a measure of entropy nor associated with the information defined in communication theory.

Temporal perceptual information (TI)

A measure that generally indicates the amount of temporal changes of a video sequence. It is usually higher for high motion sequences. It is not meant to be a measure of entropy nor associated with the information defined in communication theory.

Transparency (fidelity)

A concept describing the performance of a codec or a system in relation to an ideal transmission system without any degradation. Two types of transparency can be defined. The first type describes how well the processed signal conforms to the input signal, or ideal signal, using a mathematical criterion. If there is no difference the system is fully transparent. The second type describes how well the processed signal conforms to the input signal, or ideal signal, for a human observer. If no difference can be perceived under any experimental condition the system is perceptually transparent. The term transparent without explicit reference to a criterion will be used for systems that are perceptually transparent.