MISB MOTION IMAGERY STANDARDS BOARD	MISB ST 1108.2
STANDARD	
Motion Imagery Interpretability and Quality Metadata	**27 February 2014**

1 Scope

This document defines metadata keys necessary to express motion imagery interpretability and quality (IQ).

In addition it allows associated image chips or quality related features to be sent in an MPEG-2 transport stream to support reduced bandwidth, partial-reference techniques for IQ estimation at the receiver.

2 References

2.1 Normative References

The following references and the references contained therein are normative.

[1] MISB ST 0901.2, Video-NIIRS, Feb 2014
[2] MISB ST 0107.2, Bit and Byte Order for Metadata in Motion Imagery Files and Streams, Feb 2014
[3] MISB ST 0807.13, MISB KLV Metadata Dictionary, Feb 2014
[4] MISB RP 1203.3, Video Interpretability and Quality Measurement, Feb 2014

2.2 Informative References

[5] Objective perceptual assessment of video quality: Full reference television, online reference http://www.itu.int/ITU-T/studygroups/com09/docs/tutorial_opavc.pdf
[6] Subjective video quality assessment methods for recognition tasks, online reference, ITU-T P.912 Telecommunication Standardization Sector of ITU Series P: Terminal and Subjective and Objective Assessment Methods, (08/2008)

3 Definitions

Interpretability

Refers to the ability of the video to satisfy task-based criteria specified in the MISB ST 0901[1].

Quality

General perception of motion imagery quality measured on a scale of 0-100.

4 Acronyms

ID Identifier
IQ Interpretability and Quality
NIIRS National Imagery Interpretability Rating Scale
PNG Portable Network Graphics
PSNR Peak Signal to Noise Ratio

5 Revision History

Revision	Date	Summary of Changes
ST 1108.2	02/27/2014	• Promoted to ST; Updated to EARS format

6 Introduction

Video-NIIRS is a subjective task criteria scale similar to the well-known NIIRS used for still image quality measurement [1]. Each task in the Video-NIIRS includes a dynamic component that requires motion imagery of sufficient spatial and temporal resolution.

Reasons to include the IQ keys with the motion imagery:

- Monitoring proper operation of video collection and dissemination systems
- Discovery of motion imagery
- Automatic exploitation assistance

Quality is the perception of the overall appearance of the video. ITU distinguishes between methods to gauge video quality for tasks versus for general viewing in [5]. Video-NIIRS is a task based scale.

7 Interpretability and Quality (IQ) Metadata Local Set

Requirement	
ST 1108.2-01	All metadata shall be expressed in accordance with MISB ST 0107[2].
ST 1108.2-02	The Interpretability and Quality (IQ) metadata set shall use Local Data Set 16-byte Universal Key 06.0E.2B.34.02.03.01.01.0E.01.03.03.1C.00.00.00 (CRC 6633) for its implementation.

Table 1 lists the Interpretability and Quality Metadata Pack key and the elements of the local set.

Table 1: Interpretability and Quality Metadata Local Set

NAME		KEY			
Minimum Interpretability and Quality Pack		**06.0E.2B.34.02.03.01.01.0E.01.03.03.1C.00.00.00 (CRC 6633)**			
Constituent Keys					
Tag ID	**Key Value (hex)**	**Key Name**	**Units/ Range**	**Format**	**Length (Bytes)**
1	06.0E.2B.34.01.01.01.01.0E.01.01.03.3B.00.00.00 (CRC 20966)	Minimum Video Interpretability & Quality – Most Recent Frame Time	micro-seconds	UINT64	8
2	06.0E.2B.34.01.01.01.01.0E.01.01.03.39.00.00.00 (CRC 48270)	Minimum Video Interpretability & Quality – Video Interpretability	0 - 14	UINT8	1
3	06.0E.2B.34.01.01.01.01.0E.01.01.03.3A.00.00.00 (CRC 10066)	Minimum Video Interpretability & Quality – Video Quality	0 - 100	UINT8	1
4	06.0E.2B.34.01.01.01.01.0E.01.02.03.4A.00.00.00 (CRC 43975)	Minimum Video Interpretability & Quality – Interpretability Quality Method	0 - 255	UINT8	1
5	06.0E.2B.34.01.01.01.01.0E.01.02.03.4C.00.00.00 (CRC 35934)	PSNR Coefficient Identifier	0 - 255	UINT8	1
6	06.0E.2B.34.01.01.01.01.0E.01.02.03.4D.00.00.00 (CRC 64234)	Quality Coefficient Identifier	0 - 255	UINT8	1
7	06.0E.2B.34.01.01.01.01.0E.01.02.03.4B.00.00.00 (CRC 56691)	Minimum Video Interpretability & Quality – Rating Duration	0 - 65536 frames	UINT16	2
8	06.0E.2B.34.01.01.01.01.0E.01.01.03.3C.00.00.00 (CRC 203)	Minimum Video Interpretability & Quality – MIQ Pak Insertion Time	micro-seconds	UINT64	8

9	06.0E.2B.34.01.01.01.01. 0E.01.02.03.4E.00.00.00 (CRC 24886)	Chip Location, Size, & Bit Depth	Pixels	4x UINT16	8
10	06.0E.2B.34.01.01.01.01. 0E.01.02.03.4F.00.00.00 (CRC 6018)	Chip Y-values Uncompressed	0 - 2n-1 n = 8/16	UINT8 or UINT16	Min: 16x16x 1 Max: 128x128x2
11	06.0E.2B.34.01.01.01.01. 0E.01.02.03.50.00.00.00 (CRC 55499)	Chip Y-values PNG	Length of PNG	Wrapper	Variable
12	06.0E.2B.34.01.01.01.01. 0E.01.02.03.51.00.00.00 (CRC 44671)	Chip Edge Intensity	0 - 1000	UINT16	2 Bytes
13	06.0E.2B.34.01.01.01.01. 0E.01.02.03.52.00.00.00 (CRC 13731)	Chip Frequency Ratio	0 - 0.5	UINT16	2 Bytes
14	06.0E.2B.34.01.01.01.01. 0E.01.02.03.53.00.00.00 (CRC 17175)	Chip PSNR	0 - 100 dB	UINT8	1 Byte

7.1 Time Stamp of Last Frame in Sequence

The time stamp of the last frame in the rating sequence is the frame with the largest time stamp value.

If a chip, or feature from a single frame is sent, then the time stamp is that frame and the duration is set to 1.

7.2 Estimated Interpretability

The estimated interpretability is for an interval of video, nominally at least 10-seconds, played at the sample frame rate speed.

The estimated interpretability can be either manual or automatic as indicated by the Interpretability-Quality (IQ) Method ID.

MISB RP 1203[4] describes computational details of computing video interpretability.

7.3 Estimated Quality

The estimated quality is expressed using a score in the range of 0-100, where the following adjectives describe the rating: Excellent 100-80; Good 79-60; Fair 59-40; Poor 39-20; Bad 19-0.

The estimated quality can be either manual or automatic as indicated by the Interpretability-Quality (IQ) Method ID.

7.4 IQ Method ID

Manual ratings are indicated by a zero. Non-zero entries indicate the version number of MISB RP 1203, "Video Interpretability and Quality Measurement" used to produce the inserted rating.

7.5 PSNR Coefficient Identifier

MISB RP 1203 estimates the PSNR without a reference by applying a coefficients to a set of image features. The coefficients are updated as additional test materials are analyzed. This key designates the coefficient set number. It may be in the future, that each system, or scene type may have its own coefficient set, e.g. one set for snow, another for urban, etc.

For consistency with MISB RP 1203 [4] it is recommended that a PSNR coefficient file have the filename of PSNRCOFxxx.csv, where xxx is the PSNR coefficient identifier.

7.6 Quality Coefficient Identifier

MISB RP 1203 estimates the Quality without a reference by applying a coefficients to a set of image features. The coefficients are updated as additional test materials are analyzed. This key designates the coefficient set number. It may be in the future, that each system, or scene type may have its own coefficient set, e.g. one set for snow, another for urban, etc.

For consistency with MISB RP 1203 [4] it is recommended that a Quality Coefficient file have the filename of QUALCOFSxxx.csv, where xxx is the PSNR coefficient identifier

7.7 Rating Duration

The rating duration is expressed in frames. The maximum value is 65535. When chips or features are inserted from a single frame, the duration is set to one.

7.8 IQ Local data set Creation Time

The IQ Local data set Creation time is the current time when the IQ key value(s), chips, or features are inserted into the stream.

7.9 Chip Location, Size and Bit-Depth

The chip location is indicated by the column and row index of the top-left corner of the chip.

The first location value is the column index. The second location value is the row index.

The third value is the length of the edge of the chip. Chips are square.

The fourth value is the bit-depth of the imagery. Bit-depths up to 16 are supported.

Requirement	
ST 1108.2-03	Chip sizes shall be 32, 64 or 128.
ST 1108.2-04	For a given image frame only one chip shall be allowed.

7.10 Chip Luminosity Values Uncompressed

The chip values can be 8-bit or 16-bit depending on the bit depth. Imagery greater than 8-bits is represented in two-bytes with the first byte being the most significant.

Chip features such as frequency ratio and / or edge energy can be sent instead of the actual pixel values. The decision is based on available link bandwidth versus onboard computing power as shown in Table 2:

Table 2: Chip frequency based on bandwidth and available compute power

		Bandwidth	
		HIGH	**LOW**
Available Onboard Computing Power	**HIGH**	Scan for best chip location, send chip AND send features too.	Send just features or use scan results for most compressible PNG chip location
	LOW	Send larger uncompressed chips	Send just the chip features (frequency ratio, edge intensity, and PSNR)

7.11 Chip Luminosity Values PNG

Chip image luminosity values can be losslessly compressed using PNG format and inserted into the metadata stream.

7.12 Chip Edge Intensity Feature

Edge intensity (EI) is defined as,

$$\text{EI} = \frac{1}{RC}\sum_{i,j}\sqrt{g(i,j)_x^2 + g(i,j)_y^2},$$

where, g_x and g_y are the Sobel horizontal and vertical filtered versions of the input sub-image (chip) rows (R) and columns (C).

7.13 Chip Frequency Ratio Feature

The Chip Frequency Ratio is computed by taking the ratio of high-pass to low-pass energy. Matlab code listing is in the Appendix of MISB RP 1203, "Video Interpretability and Quality Measurement and Prediction".

Requirement	
ST 1108.2-05	The Chip Frequency Ratio shall have a cut-off frequency of 0.15.

7.14 Chip PSNR Feature

The Chip PSNR is calculated by taking 10*log10 of the ratio of the square of the greatest possible pixel intensity value to the mean-square-error between the compressed and uncompressed chip. If the PSNR is computed over the entire frame, then the chip size is the frame size.